

吴福朝 编著

计算机视觉中的数学方法

内容简介

本书由射影几何、矩阵与张量、模型估计三个部分组成，它们是三维计算机视觉所涉及到的基本数学理论与方法。I. 射影几何学是三维计算机视觉的数学理论基础，是从事计算机视觉研究所必备的数学知识。本书着重介绍射影几何学和它在视觉中的应用，主要内容包括：平面与空间射影几何，摄像机几何，两视点几何，自标定技术和三维重构理论。II. 矩阵与张量是描述和解决计算机视觉问题的必要数学工具，视觉领域研究人员都应该掌握这门数学。本书着重介绍与视觉有关的矩阵、张量理论与它的应用，主要内容包括：矩阵分解，矩阵分析，张量代数，运动与结构，多视点张量。III. 模型估计是三维计算机视觉的基本问题，通常涉及到变换或某种数学量的估计。本书着重介绍与视觉估计有关的数学理论与方法，主要内容包括：迭代优化理论，参数估计理论，视觉估计的代数方法、几何方法、鲁棒方法和贝叶斯方法。上述三部分涉及的数学内容是相对独立的，但三维计算机视觉将它们组成一个有机的整体。通过阅读本书，读者能掌握三维计算机视觉中的基本数学内容与方法，增强数学素养、提高分析和解决视觉问题的数学能力。

目 录

第一篇 射影几何

第 1 章	平面射影几何	
1.1	射影平面	3
1.1.1	射影平面	3
1.1.2	两点、两线的叉积	5
1.1.3	共线点、共点线的交比	5
1.2	二次曲线	8
1.2.1	矩阵表示	9
1.2.2	切线	9
1.2.3	配极对应	10
1.2.4	对偶二次曲线	13
1.2.5	圆环点及其对偶	14
1.3	二维射影变换	16
1.3.1	二维射影变换	16
1.3.2	直线与二次曲线的射影变换	20
1.4	变换群与不变量	21
1.4.1	等距变换群	21
1.4.2	相似变换群	23
1.4.3	仿射变换群	24
1.4.4	射影变换群	27
第 2 章	空间射影几何	
2.1	射影空间	31
2.1.1	空间点	31
2.1.2	空间平面	31
2.1.3	空间直线	34
2.1.4	共线平面束的交比	37
2.2	三维射影变换	38
2.2.1	三维射影变换	38
2.2.2	平面与直线的变换规则	39
2.3	二次曲面与变换规则	40
2.3.1	基本性质	40
2.3.2	二次曲面的对偶	42
2.3.3	绝对二次曲线与绝对二次曲面	45
2.4	空间射影变换群的子群	49
2.4.1	仿射变换群	49

	2.4.2 相似变换群	51
	2.4.3 等距变换群	52
	2.4.4 二次曲面分类	53
	2.5 射影坐标系与射影坐标变换	56
第3章	摄像机几何	
	3.1 摄像机模型	59
	3.1.1 摄像机模型	59
	3.1.2 摄像机矩阵元素的几何意义	63
	3.1.3 摄像机矩阵估计	66
	3.1.4 欧氏空间与射影空间	67
	3.2 投影与反投影	69
	3.2.1 空间点	69
	3.2.2 空间直线	70
	3.2.3 空间平面	72
	3.2.4 二次曲线	73
	3.2.5 二次曲面	74
	3.3 恢复平面景物的几何结构	77
	3.3.1 仿射结构	77
	3.3.2 相似结构	78
	3.3.3 绝对欧氏结构	79
第4章	两视点几何	
	4.1 基本矩阵	81
	4.1.1 极几何	81
	4.1.2 基本矩阵	83
	4.1.3 几何解释	87
	4.2 单应矩阵	88
	4.2.1 单应矩阵	88
	4.2.2 与基本矩阵的关系	91
	4.2.3 不动点与不动线	93
	4.3 基本矩阵估计	94
	4.3.1 8-点算法	94
	4.3.2 最小点对应算法	96
	4.4 恢复摄像机矩阵	97
	4.4.1 射影相关	97
	4.4.1 摄像机矩阵的恢复	100
第5章	自标定理论	
	5.1 正交性与摄像机内参数	101

5.1.1	影消点与影消线	101
5.1.2	摄像机内参数	102
5.2	圆环点与摄像机内参数	105
5.2.1	内参数约束方程	105
5.2.2	确定圆环点的图像	105
5.2.3	圆环点与正交方向	107
5.3	平行性与摄像机内参数	108
5.3.1	平行四边形的射影	108
5.3.2	平行六面体的射影	111
5.3.2	摄像机内参数	113
5.4	Kruppa 方程与摄像机内参数	117
5.4.1	Kruppa 方程	117
5.4.2	由 Kruppa 方程求焦距	119
5.5	绝对二次曲线与摄像机内参数	121
5.5.1	基本约束方程	121
5.5.2	变化内参数	121
5.5.3	恒定内参数	122
5.5.4	尺度因子计算公式	123
5.6	绝对二次曲面与摄像机内参数	125
5.6.1	基本约束方程	125
5.6.1	约束的等价性	126
第 6 章	三维重构理论	
6.1	三角原理	130
6.2	基本矩阵与射影重构	132
6.3	无穷远平面与仿射重构	133
6.4	绝对二次曲线与度量重构	135
6.5	绝对二次曲面与度量重构	138
6.6	实现分层重构的实例	140
6.6.1	仿射点对应	141
6.6.2	准仿射重构	142
6.6.3	仿射重构	143
6.6.4	度量重构	146
6.7	多摄像机系统标定	147
6.7.1	一维标定物	147
6.7.2	确定仿射摄像机矩阵	148
6.7.3	确定欧氏摄像机矩阵	149

第二篇 矩阵与张量

第 7 章	正交对角化	
7.1	内积空间与正交矩阵	154
7.1.1	内积空间	154
7.1.2	正交矩阵	155
7.2	酉空间与酉矩阵	158
7.2.1	酉空间	158
7.2.1	酉矩阵	159
7.3	正规矩阵	162
7.3.1	Schur 引理	162
7.3.2	正规矩阵	163
7.3.3	正交谱分解	166
7.4	厄米特矩阵	169
7.4.1	特征值的极性	169
7.4.2	半正定厄米特矩阵	172
7.5	反对称矩阵	176
第 8 章	矩阵分解	
8.1	正交—三角分解	178
8.1.1	Givens 方法	179
8.1.2	Householder 方法	180
8.1.2	内参数与外参数的分解	182
8.2	乔里斯基 (Cholesky) 分解	183
8.3	奇异值分解	184
8.3.1	正交对角分解	184
8.3.2	奇异值分解	185
8.3.3	奇异值的极性	187
8.3.4	极分解	188
8.4	最小二乘问题	189
8.4.1	满秩最小二乘问题	191
8.4.2	亏秩最小二乘问题	193
8.4.3	数值秩的定义和确定方法	193
8.4.4	齐次最小二乘问题	194
8.4.5	约束齐次最小二乘问题	196
第 9 章	矩阵分析	
9.1	向量与矩阵范数	198
9.1.1	向量范数	198
9.1.2	矩阵范数	202

9.1.3	矩阵条件数	207
9.2	矩阵级数与矩阵函数	210
9.2.1	矩阵序列	210
9.2.2	矩阵级数	212
9.2.3	矩阵函数	213
9.3	矩阵导数	217
9.3.1	函数矩阵的导数	217
9.3.2	向量值映射关于向量的导数	219
9.3.3	函数关于矩阵的导数	221
9.3.4	函数矩阵关于矩阵的导数	224
9.4	矩阵直积	228
9.4.1	概念与基本性质	228
9.4.2	特征值与特征向量	233
第 10 章	张量代数	
10.1	张量概述	236
10.2	张量积	240
10.2.1	线性映射	240
10.2.2	多重线性映射	243
10.3	张量	259
10.3.1	张量及其代数运算	259
10.3.2	对称与反对称张量	253
10.4	外代数 (Grassmann 代数)	255
10.4.1	外积运算	256
10.4.2	外代数	260
10.4.3	Plucker-Grassmann 坐标	264
第 11 章	运动与结构	
11.1	欧氏运动与结构	267
11.1.1	本质矩阵	267
11.1.2	欧氏运动	269
11.1.3	欧氏结构	271
11.2	仿射运动与结构	273
11.2.1	仿射摄像机	273
11.2.2	仿射运动与结构	275
11.3	射影运动与结构	277
11.3.1	基本原理	278
11.3.2	选择射影深度	279
11.3.3	迭代分解算法	280

第 12 章	多视点张量	
12.1	双线性关系	283
12.1.1	基本矩阵的张量形式	283
12.1.2	极点的张量形式	285
12.2	三线性关系	286
12.2.1	三点对应	286
12.2.2	点、线对应	289
12.2.3	三线性约束的独立数	291
12.2.4	恢复摄像机矩阵	293
12.3	四线性关系	295
12.3.1	四线性关系	295
12.3.2	四线性约束的独立数	298

第三篇 模型估计

第 13 章	迭代优化	
13.1	最优性条件	301
13.1.1	最优性条件	301
13.1.2	迭代格式	305
13.2	一维搜索	305
13.2.1	精确一维搜索	306
13.2.2	非精确一维搜索	308
13.3	无约束优化	310
13.3.1	最速下降法	310
13.3.2	NEWTON 法	310
13.3.3	变度量法	311
13.3.4	共轭方向法	314
14.3.5	LM 法	315
13.4	约束优化	316
13.4.1	惩罚法	316
13.4.2	乘子法	319
第 14 章	参数估计	
14.1	最大似然估计	325
14.1.1	基本概念	325
14.1.2	相合性与渐近正态性	328
14.1.3	混合模型	330
14.2	Bayes 估计	332

14.2.1	贝叶斯估计模型	332
14.2.2	无信息先验密度	334
14.2.3	共轭先验密度	337
14.2.4	贝叶斯估计	339
14.3	期望最大化 (EM) 算法	343
14.3.1	EM 算法	343
14.3.2	收敛性与估计精度	346
14.3.3	EM 法推广	349
14.4	混合模型的 EM 算法	351
14.4.1	一般混合模型	351
14.4.2	混合高斯模型	354
第 15 章	代数方法	
15.1	模型估计概述	356
15.1.1	模型	356
15.1.2	模型参数化	358
15.2	直接线性算法	361
15.2.1	线性计算框架	361
15.2.2	视觉估计问题	361
15.3	因子化线性算法	365
15.3.1	因子化计算框架	365
15.3.2	视觉估计问题	366
15.4	归一化线性算法	374
第 16 章	几何方法	
16.1	几何方法	376
16.1.1	直线与二次曲线	376
16.1.2	几何距离最小化	379
16.2	视觉估计问题	382
16.2.1	单应矩阵	382
16.2.2	基本矩阵	384
16.2.3	三焦张量	387
16.2.4	FOE	388
16.3	最大似然方法	392
16.3.1	高斯分布	392
16.3.2	最大似然方法	393
16.3.3	残差与估计误差	395
16.3.4	参数的协方差	400
第 17 章	鲁棒方法	

17.1	RANSAC 方法	408
17.1.1	直线估计	408
17.1.2	RANSAC 方法	409
17.1.3	基本矩阵的RANSAC估计	414
17.1.4	卡方分布	415
17.2	M-估计方法	416
17.3	最小中值方法	419
17.4	鲁棒最大后验估计	420
17.4.1	鲁棒最大后验估计	418
17.4.2	似然项与先验项的形式	420
17.4.3	最大化边缘后验	423
17.4.3	MAPSAC 算法	426
第 18 章 模型选择		
18.1	似然比检验	429
18.1.1	基本运动模型	429
18.1.2	似然比检验	431
18.2	AIC 与模型选择	434
18.2.1	AIC 标准	434
18.2.2	用AIC选择模型	436
18.3	BIC 与模型选择	440
18.3.1	Bayes 证据	440
18.3.1	BIC 标准	414
18.3.2	用BIC选择模型	443
18.4	GRIC 与模型选择	444
18.4.1	GRIC 标准	444
18.4.2	用GRIC选择模型	446
参考书目与文献		450
索引		453

第一篇 射影几何

本篇提要

本篇分为两个部分。第一部分由第 1、2 章所构成，主要介绍有关射影几何理论，这些理论不是射影几何所涉及的全部内容，而是从事三维计算机视觉研究所必须掌握的那些几何知识。第二部分由第 3—6 章所构成，主要论述射影几何理论在三维计算机视觉中的应用，读者将会看到射影几何的重要作用。

各章具体内容如下：

第 1 章，介绍平面射影几何。主要内容包括点、线，二次曲线的齐次表示，二次曲线与对偶二次曲线的性质；二维射影变换与基本几何元素的射影变换规则；二维射影变换群及其子群的不变量与不变性质。

第 2 章，介绍三维射影几何。主要内容包括三维点、线、面和二次曲面的齐次表示，二次曲面与对偶二次曲面的性质；三维射影变换与基本几何元素的射影变换规则；三维射影变换群及其子群的不变量与不变性质。

第 3 章，介绍摄像机几何。首先，对摄像机进行数学建模；然后，应用前两章射影几何知识，给出空间基本几何元素的投影性质，以及图像平面基本几何元素的反投影性质。这些投影与反投影性质，是从图像恢复物体三维几何结构的基础，尤其是绝对二次曲线与绝对二次曲面的投影性质。

第 4 章，介绍两视点几何。应用射影几何知识引进两幅图像的点、线关联关系，即所谓的极几何。基本矩阵是极几何的代数描述，同时它也给出了射影意义下的摄像机矩阵。

第 5 章，介绍摄像机自标定方法。主要着重介绍正交性、圆环点、平行性、KRUPPA 方程、绝对二次曲线和绝对二次曲面自标定与摄像机内参数的关系。所有自标定方法均来源于射影几何理论，它们都归结为绝对二次曲线或绝对二次曲面的投影性质。

第 6 章，介绍三维重构理论。三维重构是三维计算机视觉中的核心问题。本章应用射影几何知识来建立三维重构理论。不需要图像以外的知识，仅从图像就能获得射影重构；仿射重构，确定无穷远平面的射影坐标，确定无穷远单应三者是相互等价的；度量重构，确定绝对二次曲线，确定绝对二次曲面，确定摄像机内参数四者是相互等价的。

1. 平面射影几何

1.1 射影平面

1.1.1 射影平面

齐次坐标

在本章中，除特别说明外，我们均假定在平面上建立了欧氏坐标系。平面上的点可用一个二维有序数组 $\tilde{\mathbf{p}} = (x, y)^T$ 来表示，即该点的欧氏坐标。平面上的直线方程可以表示为

$$ax + by + c = 0 \quad (1.1.1)$$

在方程(1.1.1)两边同乘以任一非零常数 t ，得到下述方程：

$$axt + byt + ct = 0 \quad (1.1.2)$$

方程(1.1.1)与方程(1.1.2)有相同的几何意义，它们表示同一条直线。令

$$\mathbf{p} = (xt, yt, t)^T, \quad \mathbf{l} = (a, b, c)^T$$

则方程(1.1.2)可写成

$$\mathbf{l}^T \mathbf{p} = 0 \quad (1.1.3)$$

其中 \mathbf{p} 是变量，表示直线上的点； \mathbf{l} 是一个固定的向量，代表该直线。

一般地， $\mathbf{p} \equiv (xt, yt, t)^T$ 称为点的齐次坐标， $\mathbf{l} \equiv (a, b, c)^T$ 称为直线的齐次坐标。这里的“齐次”也可以这样来理解：在这种表示下，直线方程(1.1.3)关于点或直线变量都是齐次的，而方程(1.1.1)则是非齐次的。

齐次坐标可以相差任意的非零常数因子，即 $\forall s \neq 0$ ， \mathbf{p} 和 $\mathbf{q} = s\mathbf{p}$ 表示同一个点，因为它们的非齐次坐标相等：

$$\tilde{\mathbf{p}} = (x/t, y/t)^T = (sx/st, sy/st) = \tilde{\mathbf{q}}$$

对于直线的齐次坐标也一样，因为方程 $(s\mathbf{l})^T \mathbf{p} = 0$ 与方程(1.1.3)确定同一条直线。

射影平面

齐次坐标为 $\mathbf{p}_\infty = (x, y, 0)^T$ 的点称为无穷远点，其中 x, y 至少有一个不为零。注意：无穷远点没

有欧氏坐标, 这是因为 $x/0 = \infty$, $y/0 = \infty$, 同时也可以看出为什么将它称为无穷远点。平面上所有无穷远点所构成的集合称为无穷远直线。由于所有无穷远点 $p_\infty = (x, y, 0)^T$ 都满足方程:

$$0 \cdot x + 0 \cdot y + 1 \cdot 0 = 0$$

所以, 无穷远直线的齐次坐标为 $l_\infty = (0, 0, 1)^T$ 。

由欧氏平面与无穷远直线的并集所形成的扩展平面称为射影平面, 有时也称为二维射影空间。

1.1.2 两点、两线的叉积

三维向量的叉积

令 $\mathbf{x}_1 = (x_1, y_1, t_1)^T$, $\mathbf{x}_2 = (x_2, y_2, t_2)^T$ 是两个三维向量, 它们的叉积定义为

$$\mathbf{x}_1 \times \mathbf{x}_2 = \det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & y_1 & t_1 \\ x_2 & y_2 & t_2 \end{pmatrix} = \left(\det \begin{pmatrix} y_1 & t_1 \\ y_2 & t_2 \end{pmatrix}, -\det \begin{pmatrix} x_1 & t_1 \\ x_2 & t_2 \end{pmatrix}, \det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix} \right)^T \quad (1.1.4)$$

叉积和反对称矩阵相关联。由向量 $\mathbf{x} = (x, y, t)^T$, 按下述方式定义反对称矩阵 $[\mathbf{x}]_\times$:

$$[\mathbf{x}]_\times = \begin{pmatrix} 0 & -t & y \\ t & 0 & -x \\ -y & x & 0 \end{pmatrix}$$

并称为由向量 \mathbf{x} 所确定的反对称矩阵。矩阵 $[\mathbf{x}]_\times$ 具有下述性质:

- (1) 对任意非零向量 \mathbf{x} , 有 $\text{rank}([\mathbf{x}]_\times) = 2$;
- (2) 对任意两个 3 维向量 $\mathbf{x}_1, \mathbf{x}_2$, 有 $\mathbf{x}_1 \times \mathbf{x}_2 = [\mathbf{x}_1]_\times \mathbf{x}_2$;
- (3) \mathbf{x} 是 $[\mathbf{x}]_\times$ 的右零空间, 同时也是它的左零空间, 即 $[\mathbf{x}]_\times \mathbf{x} = 0, \mathbf{x}^T [\mathbf{x}]_\times = 0$;
- (4) 对任意 3 维向量 \mathbf{y} , 有 $\mathbf{y}^T [\mathbf{x}]_\times \mathbf{y} = 0$ 。

性质(1)是明显的, (3)与(4)可由性质(2)导出, 因此下面仅验证性质(2):

令 $\mathbf{x}_1 = (x_1, y_1, t_1)^T$, $\mathbf{x}_2 = (x_2, y_2, t_2)^T$, 则

$$[\mathbf{x}_1]_\times \mathbf{x}_2 = \begin{pmatrix} 0 & -t_1 & y_1 \\ t_1 & 0 & -x_1 \\ -y_1 & x_1 & 0 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \\ t_2 \end{pmatrix} = \begin{pmatrix} -t_1 y_2 + y_1 t_2 \\ t_1 x_2 - x_1 t_2 \\ -y_1 x_2 + x_1 y_2 \end{pmatrix} = \mathbf{x}_1 \times \mathbf{x}_2$$

性质(2)表明, 两个向量的叉积可以用其中一个向量的反对称矩阵左乘另一个向量来表达。

两点、两线的叉积

如果 p_1, p_2 是射影平面上两点, 则 $l = p_1 \times p_2$ 表示通过这两点的直线。这是因为, 对直线 l 上的任一点, 它的齐次坐标可以表示为 $p = s_1 p_1 + s_2 p_2$, 根据反对称矩阵的性质(3)和(4), 可以得到:

$$l^T p = \sum_{j=1}^2 s_j ([p_1]_{\times} p_2)^T p_j = 0$$

若三点 p_1, p_2, p_3 共线, 则必有 $p_2^T [p_1]_{\times} p_3 = -([p_1]_{\times} p_2)^T p_3 = 0$ 。反之, 若 $p_2^T [p_1]_{\times} p_3 = 0$, 则三点 p_1, p_2, p_3 必共线。因此, 我们有下述命题:

命题 1.1.1 (1) 两点 p_1, p_2 连线的坐标是 $l = [p_1]_{\times} p_2$; (2) 三点 p_1, p_2, p_3 共线的充要条件是 $p_2^T [p_1]_{\times} p_3 = 0$ 。

对偶原理 在射影平面内, 点和线是一对互为对偶元素。在包含“点”和“线”元素的命题中, 如果将两个元素的角色互换, 则对应的命题也成立, 并称它们是一对互为对偶命题。

例如: 命题 1.1.1 有如下对偶命题:

命题 1.1.2 (1) 两线 l_1, l_2 的交点的坐标是 $p = [l_1]_{\times} l_2$; (2) 三线 l_1, l_2, l_3 共点的充要条件是 $l_2^T [l_1]_{\times} l_3 = 0$ 。

1.1.3 共线点、共点线的交比

共线点的参数化与交比

平面上的点有两个自由度, 用三维非零向量(齐次坐标)来表示。而直线上的点仅有一个自由度, 因此, 直线上点的齐次坐标仅需要二维向量来表示。如何用二维向量来表示直线上点的齐次坐标, 就是所谓共线点的参数化问题。

给定直线 l 上两个不同点的齐次坐标 p_1, p_2 ^{*)}, 则直线 l 上任何一个点 p 的坐标均可以表示为

$$p = u p_1 + v p_2$$

这样, 利用直线 l 上两个点 p_1, p_2 , 直线 l 上所有点都可以用二维向量来表示:

$$\hat{p} = (u, v)^T$$

并称这个二维向量为直线 l 上点的参数化。

显然, p_1, p_2 的(二维)齐次坐标分别是 $\hat{p}_1 = (1, 0)^T, \hat{p}_2 = (0, 1)^T$ 。因此, 这种参数化过程实际上

^{*)} “给定齐次坐标”的意义是指已经给出了齐次坐标分量的具体数值, 在本书其它地方都遵循这个约定。

是建立直线坐标系的过程。直线上点的参数化不是唯一的，不同的参数化对应不同的坐标系。

共线点的交比 假定 p_1, p_2, p_3, p_4 是 4 个共线点，它们在某种参数化下的齐次坐标为 $\hat{p}_j = (u_j, v_j)^T$ 。定义：

$$(p_1, p_2; p_3, p_4) = \frac{\det(\hat{p}_1, \hat{p}_3)}{\det(\hat{p}_2, \hat{p}_3)} : \frac{\det(\hat{p}_1, \hat{p}_4)}{\det(\hat{p}_2, \hat{p}_4)} \quad (1.1.10)$$

为该 4 点的交比。

下面证明：共线点 p_1, p_2, p_3, p_4 交比不依赖于点参数化的选择，或者说不依赖于直线坐标系的选择。

若点 p_1, p_2, p_3, p_4 在两个给定点 q_1, q_2 的参数化下的齐次坐标分别为 $\hat{p}_j = (u_j, v_j)^T$, $j=1,2,3,4$ ，而在另外两个给定点 q'_1, q'_2 的参数化下的齐次坐标分别为 $\hat{p}'_j = (u'_j, v'_j)^T$, $j=1,2,3,4$ 。令点 q'_1, q'_2 在 q_1, q_2 的参数化下齐次坐标分别为 $\hat{q}'_1 = (\alpha_1, \beta_1)^T$, $\hat{q}'_2 = (\alpha_2, \beta_2)^T$ ，即

$$q'_1 = \alpha_1 q_1 + \beta_1 q_2, q'_2 = \alpha_2 q_1 + \beta_2 q_2$$

记 $H = \begin{pmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{pmatrix}$ ，则 $\det(H) \neq 0$ 。否则， $\{(\alpha_1, \beta_1)^T, (\alpha_2, \beta_2)^T\}$ 必线性相关，从而 q'_1, q'_2 表示直线上的同一个点。于是，

$$p_j = u'_j q'_1 + v'_j q'_2 = (u'_j \alpha_1 + v'_j \alpha_2) q_1 + (u'_j \beta_1 + v'_j \beta_2) q_2$$

所以， $\begin{cases} u_j = u'_j \alpha_1 + v'_j \alpha_2 \\ v_j = u'_j \beta_1 + v'_j \beta_2 \end{cases}$ ，即 $\hat{p}_j = H^T \hat{p}'_j$ 。因此，我们有

$$\begin{aligned} & \frac{\det(\hat{p}_1, \hat{p}_3)}{\det(\hat{p}_2, \hat{p}_3)} : \frac{\det(\hat{p}_1, \hat{p}_4)}{\det(\hat{p}_2, \hat{p}_4)} \\ &= \frac{\det(H^T (\hat{p}'_1, \hat{p}'_3))}{\det(H^T (\hat{p}'_2, \hat{p}'_3))} : \frac{\det(H^T (\hat{p}'_1, \hat{p}'_4))}{\det(H^T (\hat{p}'_2, \hat{p}'_4))} = \frac{\det(\hat{p}'_1, \hat{p}'_3)}{\det(\hat{p}'_2, \hat{p}'_3)} : \frac{\det(\hat{p}'_1, \hat{p}'_4)}{\det(\hat{p}'_2, \hat{p}'_4)} \end{aligned}$$

故 4 个共线点 p_1, p_2, p_3, p_4 的交比不依赖于参数化的选择。

根据交比不依赖于参数化选择的性质，可以简化交比的计算。

例如：在 4 个共线点 p_1, p_2, p_3, p_4 中，可以将 p_3, p_4 在平面的齐次坐标分别表示为 $p_3 = p_1 + \lambda_1 p_2$ ， $p_4 = p_1 + \lambda_2 p_2$ ，则有

$$\hat{p}_1 = (1, 0)^T, \hat{p}_2 = (0, 1)^T, \hat{p}_3 = (1, \lambda_1)^T, \hat{p}_4 = (1, \lambda_2)^T$$

于是，

$$(p_1, p_2; p_3, p_4) = \frac{\det(\hat{p}_1, \hat{p}_3)}{\det(\hat{p}_2, \hat{p}_3)} : \frac{\det(\hat{p}_1, \hat{p}_4)}{\det(\hat{p}_2, \hat{p}_4)} = \frac{\lambda_1}{\lambda_2} \quad (1.1.11)$$

这是常用的交比计算公式。

若 $(p_1, p_2; p_3, p_4) = -1$, 则称 p_1, p_2 与 p_3, p_4 成调和共轭。例如, 通过圆心的直线交圆上的两个点与圆心以及该直线上的无穷远点成调和共轭。

共点直线的参数化与交比

平面上的直线有两个自由度, 用三维非零向量(齐次坐标)来表示。但是, 对于共点直线束中的直线仅有一个自由度, 因此共点直线束中的直线可以用二维向量来表示。

给定共点直线束 l_λ 中两条不同直线的齐次坐标 l_1, l_2 , 则直线束 l_λ 中任何一条直线 l 的坐标均可以表示为

$$l = al_1 + bl_2$$

这样, 利用直线束 l_λ 中的两条直线, 直线束 l_λ 中的所有直线 l 都可以用二维向量来表示:

$$\hat{l} = (a, b)^T$$

称这个二维向量为直线束 l_λ 中直线 l 的参数化。

共点直线的交比 与定义共线点的交比一样, 也可以定义 4 条共点直线的交比。假定 l_1, l_2, l_3, l_4 是直线束 l_λ 中的 4 个条直线, 它们在某种参数化下的齐次坐标为 $\hat{l}_j = (a_j, b_j)^T$ 。定义:

$$(l_1, l_2; l_3, l_4) = \frac{\det(\hat{l}_1, \hat{l}_3)}{\det(\hat{l}_2, \hat{l}_3)} : \frac{\det(\hat{l}_1, \hat{l}_4)}{\det(\hat{l}_2, \hat{l}_4)} \quad (1.1.10)$$

为该 4 直线的交比。同样, 4 条共点直线的交比也不依赖于直线束参数化的选择。

在实际应用中, 通常应用下述方法来计算 4 个共点直线的交比: 在 4 个共点直线 $l_1, l_2; l_3, l_4$ 中, 将 l_3, l_4 齐次坐标分别表示为 $l_3 = l_1 + \lambda_1 l_2$, $l_4 = l_1 + \lambda_2 l_2$, 则它们的交比为

$$(l_1, l_2; l_3, l_4) = \frac{\lambda_1}{\lambda_2} \quad (1.1.11)$$

命题 1.1.3 如果 4 条共(有穷)点直线 $l_1, l_2; l_3, l_4$ 的斜率分别为 k_1, k_2, k_3, k_4 , 则它们的交比为:

$$(l_1, l_2; l_3, l_4) = \frac{k_1 - k_3}{k_2 - k_3} : \frac{k_1 - k_4}{k_2 - k_4} \quad (1.1.12)$$

证明: 假定直线 $l_1, l_2; l_3, l_4$ 共(有穷)点 $\tilde{p} = (x, y)^*$, 则直线 $l_1, l_2; l_3, l_4$ 齐次坐标分别为

*) 以后用 \tilde{p} 表示有穷点的非齐次坐标

$$l_1 = (-k_1, 1, k_1 x_0 - y_0)^T, l_2 = (-k_2, 1, k_2 x_0 - y_0)^T, l_3 = (-k_3, 1, k_3 x_0 - y_0)^T, l_4 = (-k_4, 1, k_4 x_0 - y_0)^T$$

以直线 l_1, l_2 为基准, 对过点 p 的直线束进行参数化, 我们有:

$$l_3 = \frac{k_3 - k_2}{k_1 - k_2} l_1 + \frac{k_3 - k_1}{k_1 - k_2} l_2, l_4 = \frac{k_4 - k_2}{k_1 - k_2} l_1 + \frac{k_4 - k_1}{k_1 - k_2} l_2$$

所以,

$$\hat{l}_1 = (1, 0)^T, \hat{l}_2 = (0, 1)^T, \hat{l}_3 = \left(\frac{k_3 - k_2}{k_1 - k_2}, \frac{k_3 - k_1}{k_1 - k_2} \right)^T, \hat{l}_4 = \left(\frac{k_4 - k_2}{k_1 - k_2}, \frac{k_4 - k_1}{k_1 - k_2} \right)^T$$

因此,

$$(l_1, l_2; l_3, l_4) = \frac{\det(\hat{l}_1, \hat{l}_3)}{\det(\hat{l}_2, \hat{l}_3)} : \frac{\det(\hat{l}_1, \hat{l}_4)}{\det(\hat{l}_2, \hat{l}_4)} = \frac{\frac{k_3 - k_1}{k_1 - k_2}}{\frac{k_3 - k_2}{k_1 - k_2}} : \frac{\frac{k_4 - k_1}{k_1 - k_2}}{\frac{k_4 - k_2}{k_1 - k_2}} = \frac{k_1 - k_3}{k_2 - k_3} : \frac{k_1 - k_4}{k_2 - k_4}$$

证毕。

关于 4 条共点直线的交比与 4 共线点的交比之间的关系, 我们有下述命题:

命题 1.1.4 如果 4 条共点直线 l_1, l_2, l_3, l_4 被任一直线 l 截于 4 点 p_1, p_2, p_3, p_4 , 则

$$(l_1, l_2; l_3, l_4) = (p_1, p_2; p_3, p_4)$$

在 1.3 节, 我们将会看到交比是射影变换的不变量。

1.2 二次曲线

1.2.1 矩阵表示

我们知道二次曲线的方程可表示为

$$ax^2 + by^2 + 2cxy + 2dx + 2ey + f = 0$$

写成矩阵形式, 有

$$(x, y, 1) \begin{pmatrix} a & c & d \\ c & b & e \\ d & e & f \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0$$

令 $C = \begin{pmatrix} a & c & d \\ c & b & e \\ d & e & f \end{pmatrix}$, 它是一个对称矩阵。如果点使用齐次坐标, 则二次曲线的方程可表示为

$$p^T C p = 0 \quad (1.2.1)$$

我们称对称矩阵 C 是二次曲线的矩阵表示。

矩阵 C 虽然有 6 个不同的元素, 但由于方程(1.2.1)的齐次性, 所以仅有 5 个独立的元素, 即起确定作用的是 5 个比值, 例如: $a/f, b/f, c/f, d/f, e/f$ 。因此, 二次曲线 C 有 5 个自由度。在一般情况下^{*)}, 射影平面上的 5 个点唯一确定一条二次曲线, 并且可以通过求解下述线性方程组得出:

$$p_j^T C p_j = 0, j = 1, 2, \dots, 5 \quad (1.2.2)$$

二次曲线根据它的秩(即 C 的秩)是否是满秩分为非退化与退化的两种情况。非退化二次曲线是正常二次曲线, 退化二次曲线或者是由两条直线所构成(rank $C=2$), 或者由二条重合直线所构成(rank $C=1$)。

如果二次曲线 C 退化为两条直线 l 与 m , 则它的矩阵表示为

$$C = lm^T + ml^T \quad (1.2.3)$$

这是明显的: 如果 $p \in l$, 则 $l^T p = p^T l = 0$, 因此, $p^T C p = p^T (lm^T + ml^T) p = 0$ 。同理, 当 $p \in m$ 时, 也有 $p^T C p = 0$ 。反之, 如果点 p 使得 $p^T C p = 0$, 则

$$2(l^T p)(m^T p) = p^T lm^T p + p^T ml^T p = p^T C p = 0$$

因此, $l^T p = 0$ 或者 $m^T p = 0$, 即 $p \in l$ 或者 $p \in m$ 。所以, 式(1.2.3)是退化二次曲线的矩阵表示。

1.2.2 切线

切点与切线

假定 C 是一条非退化二次曲线, 下面讨论切点与切线的代数表示。

令 p 是 C 上的任一点, 则 $l=Cp$ 确定平面上的一条直线。下面证明直线 l 是 C 在点 p 处的切线。首先, 点 p 必在直线 l 上, 这是因为 $l^T p = (Cp)^T p = p^T C p = 0$ 。如果能证明: 除点 p 外, 直线 l 与二次曲线 C 不再有另外的交点, 那么就证明了 l 是点 p 处的切线。反证: 若直线 l 还交 C 于另外一点 q , 则必有 $q^T C q = 0, p^T C q = l^T q = 0$ 。将此式与 $p^T C p = 0$ 结合起来, 可导致等式: $(sp + tq)^T C (sp + tq) = 0$ 对任何标量 s, t 都成立, 这表明直线 l 在二次曲线 C 上, 与 C 非退化矛盾。

假定 l 是非退化二次曲线 C 的任一条切线, 令切点为 p , 由上面的讨论必有 $l=Cp$ 。因此, 切点

^{*)} 平面上给定 5 个点, 如果任意 3 点不共线, 则这 5 个点唯一确定一条二次曲线。

是 $p = C^{-1}l$, 并且还可以得到 $l^T C^{-1}l = 0$ 。

如果直线 l 使得 $l^T C^{-1}l = 0$, 则坐标为 $p = C^{-1}l$ 的点一定在二次曲线 C 上, 这是因为有 $p^T Cp = l^T C^{-T} C C^{-1}l = l^T C^{-1}l = 0$ 。由于 $l = Cp$, 所以直线 l 是二次曲线 C 的切线。

总结上述讨论, 我们有下述命题:

命题 1.2.1 (1) 非退化二次曲线 C 在点 p 处切线为 $l = Cp$; (2) 若直线 l 是非退化二次曲线 C 的切线, 则切点为 $p = C^{-1}l$; (3) 直线 l 为非退化二次曲线 C 的切线当且仅当 $l^T C^{-1}l = 0$ 。

由于退化二次曲线是由两条直线 l 与 m 所构成, 它的矩阵表示为 $C = lm^T + ml^T$ 。如果直线 l 与 m 相交(包括交点在无穷远处的情况), 即二次曲线 C 的秩为 2, 则除交点外每一点处的切线是该点所在的直线。如果直线 l 与 m 重合, 则每一点的切线是该直线本身。

若点 p 在退化二次曲线 C 上, 例如在直线 l 上, 则必有 $Cp = (lm^T + ml^T)p = (m^T p)l$, 这说明退化二次曲线 C 在点 p 处的切线坐标仍为 Cp , 即命题 1.2.1(1)对退化二次曲线仍成立。

过二次曲线外一点的两条切线

由初等几何, 我们知道对于非退化二次曲线 C 外部任一点 p , 必存在两条切线 l 与 m 过点 p , 它们构成平面上的一条退化二次曲线 T , 如图 1.2.1 所示。

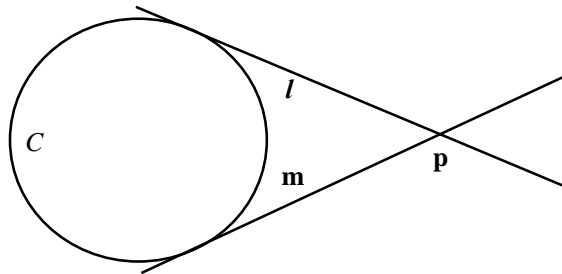


图 1.2.1 过二次曲线外部一点的两条切线

现在我们希望由 C 和 p 给出退化二次曲线 T 的矩阵表示, 对此有下述命题:

命题 1.2.2 对于非退化二次曲线 C 外部任一点 p , 过点 p 的两条切线 l 与 m 所构成的退化二次曲线的矩阵表示是 $T = [p]_x C^{-1} [p]_x$ 。

证明 设 q 是 T 上的任一点, 例如它是切线 l 上的点, 则 $l = [p]_x q$ 。于是, 我们有

$$q^T T q = q^T [p]_x C^{-1} [p]_x q = l^T C^{-1} l = 0$$

即 T 上的任一点 q 满足方程 $q^T T q = 0$ 。反之，若点 q 满足方程 $q^T T q = 0$ ，即

$$q^T [p]_{\times} C^{-1} [p]_{\times} q = 0$$

由命题 1.2.1 (3)，坐标为 $[p]_{\times} q$ 的直线是过点 p 的切线，因此 q 在切线 l 上，或者在切线 m 上。故命题成立。证毕。

1.2.3 配极对应（变换）

配极对应

给定一条二次曲线 C ，则对平面上的任一点 p ， $l=Cp$ 确定了一条直线。直线 l 称为点 p 关于 C 的极线，而点 p 称为直线 l 关于 C 的极点。如果点 p 在 C 上，则它关于 C 的极线是通过它的切线 l ，而切线 l 关于 C 的极点是切点 p 。

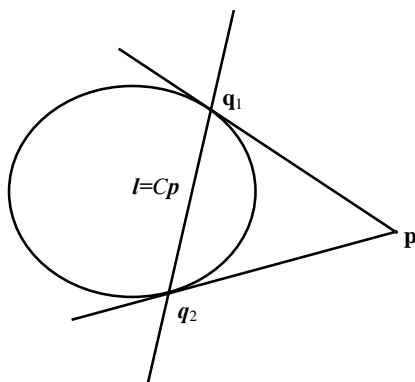


图 1.2.2 二次曲线的配极对应。过点 p 的两条切线的切点的连线 $q_1 q_2$ 是点 p 的极线 $l=Cp$ 。

由二次曲线所确定的这种点与直线之间的对应关系称为二次曲线的配极对应。可以证明：非退化二次曲线的配极对应是点与直线之间的一一对应。配极对应，也可以给出它的几何描述：

命题 1.2.3 点 p 关于非退化二次曲线 C 的极线 $l=Cp$ 交 C 于两个点^{*)}，且 C 在这两个交点的切线交于点 p ，如图 1.2.2 所示。

证明 直线 l 与二次曲线 C 总交于二个点，交点的坐标是方程组 $\begin{cases} q^T C q = 0 \\ l^T q = 0 \end{cases}$ 的解。令 q 是一个交点，则有 $(Cp)^T q = 0$ 。由矩阵 C 的对称性， $p^T C q = 0$ ，这说明 p 必在切线 Cq 上，同理 p 也在另一

^{*)} 可能是两个虚点或重点。如果是虚点，则必为一对共轭虚点，而一对共轭虚点确定一条（实）直线；如果是重点，则极线是切线。

个交点的切线上。证毕。

共轭点

如果两个点 p, q 使得 $p^T C q = 0$ ，则称点 p, q 关于 C 互为共轭。不难看出点 p 关于 C 的所有共轭点所构成的集合是点 p 关于 C 的极线。

命题 1.2.4 若点 p, q 关于二次曲线 C 是一对共轭点，直线 $l = p \times q$ 交 C 于两点 r_1, r_2 ，如图 1.2.3 所示，则 $(r_1, r_2; p, q) = -1$ 。

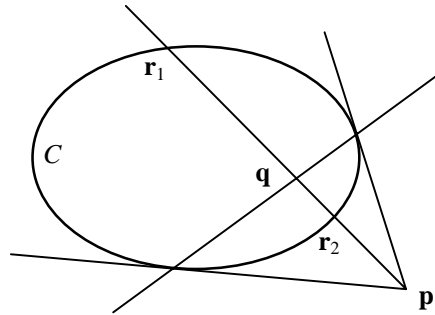


图 1.2.3: 共轭点 p, q

证明 以 p, q 为基点，对直线 $l = p \times q$ 上的点进行参数化，则 $\hat{p} = (1, 0)^T, \hat{q} = (0, 1)^T$ 。下面计算 r_1, r_2 的参数化坐标。由于直线 l 上的任一点均可表示为 $r = up + vq$ ，且点 p, q 是二次曲线 C 的一对共轭点，所以有

$$r^T C r = u^2 p^T C p + v^2 q^T C q$$

因此，直线 l 与 C 的两个交点的坐标满足方程： $r^T C r = 0$ ，于是，

$$u^2 p^T C p + v^2 q^T C q = 0$$

不妨假定 $u = 1$ ，则 r_1, r_2 的参数化坐标必为

$$\hat{r}_{1,2} = \left(1, \pm \sqrt{-\frac{p^T C p}{q^T C q}} \right)^T$$

由交比计算公式(1.1.11)，得到 $(r_1, r_2; p, q) = -1$ 。

自极三角形 如果一个三角形的三个顶点都是其对边关于二次曲线 C 的极点，则称它为 C 的自极三角形。例如：二次曲线上的四点构成的完全四边形的对边三角形是该二次曲线的自极三角形，如图 1.2.4 所示。

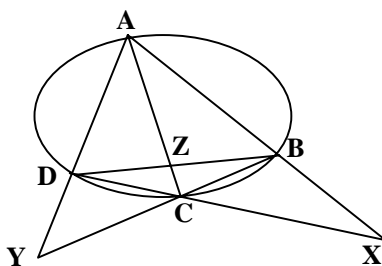


图 1.2.4 完全四点形 $ABCD$ 的对边三角形 XYZ 是该二次曲线的自极三角形

1.2.4 对偶二次曲线

前面所讨论的二次曲线被看作是由平面上的点所构成的集合，以点作为二次曲线的基本元素。平面上的点与直线构成一对互为对偶元素，如果将二次曲线方程 $p^T C p = 0$ 中的点元素换成线元素，矩阵 C 也换成对偶形式 C^* ，则我们得到线元素的二次方程：

$$l^T C^* l = 0 \tag{1.2.5}$$

其中 C^* 是对称矩阵，它是矩阵 C 的对偶。

方程(1.2.5)也表示射影平面内的一条二次曲线，这条二次曲线 C^* 是由直线生成的。由直线生成的二次曲线通常称为对偶二次曲线。在几何上，对偶二次曲线是直线族的包络，即 C^* 的几何元素是二次曲线的切线，如图 1.2.5 所示。

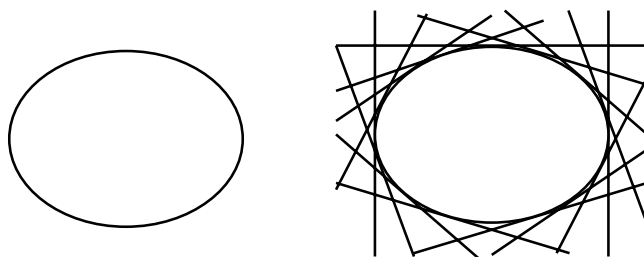


图 1.2.5 二次曲线(左)；对偶二次曲线(右)

下面讨论二次曲线与其对偶二次曲线之间的关系。

非退化情况 射影平面上任一条(点)二次曲线 C , 都可以作为其切线的包络, 即同时可用(线)二次曲线来表示, 记为 C^* , 并称 C 与 C^* 互为对偶。下面考虑 C^* 与 C 之间的代数关系。对于 C 上的任一点 p , 该点的切线为 $l=Cp$ 。由于 C 是满秩的, 所以有 $C^{-1}l=p$, 又因切点必在切线上, 即 $l^T p=0$, 于是有 $l^T C^{-1}l=0$ 。因此 $C^*=C^{-1}$ 。另外, 还可以证明 $(C^*)^*=C$ 。

综合上面所讨论的结果, 我们有下述命题:

命题 1.2.5 非退化二次曲线 C 与其对偶 C^* 之间的关系是 $C^*=C^{-1}$, 并且 $(C^*)^*=C$ 。

退化情况 退化二次曲线 C 由两条直线所组成。令这两条直线为 l, m , 则 C 可表示为

$$C = lm^T + ml^T$$

根据对偶原理, 对偶二次曲线 C^* 包含两个点 x, y , 并且 C^* 可表示为

$$C^* = xy^T + yx^T$$

注意: 对于退化情况, $(C^*)^* \neq C$ 。

1.2.5 圆环点及其对偶

圆环点

无穷远直线 l_∞ 上的两个点:

$$I = \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}, \quad J = \begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix} \quad (1.2.6)$$

称为圆环点或绝对点。其中 $i = \sqrt{-1}$, 所以它们是一对共轭虚点。圆环点的方程可以表示为

$$\begin{cases} x^2 + y^2 = 0 \\ t = 0 \end{cases}$$

因此, 圆环点可以看作是平面上的一条(退化)二次曲线。注意: 在平面上, 圆环点必须用两个方程来表示。如果限制在无穷远直线上, 即, 论域是无穷远直线而不是整个平面, 则圆环点可由单个方程: $x^2 + y^2 = 0$ 来表达, 它的矩阵表示是一个 2 阶单位矩阵。

圆环点的命名来源于平面上任何圆与无穷远直线均交于 I, J 。事实上, 圆的方程可表示为

$$x^2 + y^2 + 2dxt + 2eyt + ft = 0$$

而无穷远直线的方程是 $t = 0$ 。将这两个方程联立求解, 可以得到 $x^2 + y^2 = 0$ 。因此, 交点的齐次坐

标必为 I, J 。现在, 也可以解释为什么给定三个点能唯一确定一个圆, 因为圆总是通过两个圆环点。所以, 3 点确定一个圆与 5 个点才能确定一条二次曲线并不矛盾。

圆环点的对偶

圆环点可以看作平面上的一条(退化)二次曲线, 它的方程是 $x^2 + y^2 = 0$ 。它的对偶二次曲线 C_{∞}^* 称为圆环点的对偶, 可以表示为

$$C_{\infty}^* = IJ^T + JI^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (1.2.7)$$

对偶二次曲线 C_{∞}^* 是由以圆环点 I, J 为中心的两个平行(虚)直线束所构成的, 两个平行直线束的斜率分别为 i 和 $-i$ 。这两个直线束中的直线称为迷向直线 (或极小直线)。通过平面内任一有穷点, 存在两条迷向直线, 分别属于这两个直线束。

命题 1.2.6 令 $l_1 = (a_1, b_1, c_1)^T, l_2 = (a_2, b_2, c_2)^T$ 是两条非迷向直线, 它们之间的交角记为 θ , 则

$$\cos \theta = \frac{l_1^T C_{\infty}^* l_2}{\sqrt{l_1^T C_{\infty}^* l_1} \cdot \sqrt{l_2^T C_{\infty}^* l_2}} \quad (1.2.8)$$

证明 这个命题的证明是容易的。因为由欧氏几何, 我们有

$$\cos \theta = \frac{a_1 a_2 + b_1 b_2}{\sqrt{a_1^2 + b_1^2} \cdot \sqrt{a_2^2 + b_2^2}} \quad (1.2.9)$$

而

$$l_1^T C_{\infty}^* l_2 = a_1 a_2 + b_1 b_2, \quad \sqrt{l_1^T C_{\infty}^* l_1} = \sqrt{a_1^2 + b_1^2}, \quad \sqrt{l_2^T C_{\infty}^* l_2} = \sqrt{a_2^2 + b_2^2}$$

所以, 式(1.2.8)成立。

命题 1.2.6 给出了对偶二次曲线 C_{∞}^* 的度量性质, 其重要性在于它在一般射影空间中仍成立。也就是说, 如果知道对偶二次曲线 C_{∞}^* 在一般射影空间中的表示, 仍可以利用式(1.2.8)来计算两条直线之间的夹角, 而公式(1.2.9)仅适用于欧氏坐标系。

如果两条直线中有一条是迷向直线(或两条都是迷向直线), 则它们之间的夹角是不定的。例如: l_1 是一条迷向直线, 则必有 $l_1^T C_{\infty}^* l_1 = 0$ (这是因为 l_1 通过一个圆环点), 于是得到 $\cos \theta = */0$ 。因此, 包含迷向直线的两条直线的夹角是不能确定的, 即迷向直线不象通常直线那样具有方向。

拉格尔(Laguerre)定理

命题 1.2.7(拉格尔定理) 设两条非迷向直线的夹角为 θ , 并且这两条直线与过它们交点以 $i, -i$ 为斜率的两条迷向直线所成的交比为 μ , 则必有

$$\theta = \frac{1}{2i} \ln \mu \quad (1.2.11)$$

证明 设两条非迷向直线 l_1, l_2 的斜率为 λ_1, λ_2 , 以 $i, -i$ 为斜率的两条迷向直线为 m_1, m_2 , 则有

$$\mu = (l_1, l_2; m_1, m_2) = \frac{\lambda_1 - i}{\lambda_2 - i} : \frac{\lambda_1 + i}{\lambda_2 + i} = \frac{1 + i \frac{\lambda_2 - \lambda_1}{1 + \lambda_1 \lambda_2}}{1 - i \frac{\lambda_2 - \lambda_1}{1 + \lambda_1 \lambda_2}}$$

由于 $\operatorname{tg} \theta = \frac{\lambda_2 - \lambda_1}{1 + \lambda_1 \lambda_2}$, 所以, $\mu = \frac{1 + i \operatorname{tg} \theta}{1 - i \operatorname{tg} \theta}$ 。又因 $\operatorname{tg} \theta = \frac{1}{i} \left(\frac{e^{2i\theta} - 1}{e^{2i\theta} + 1} \right)$, 即 $e^{2i\theta} = \frac{1 + i \operatorname{tg} \theta}{1 - i \operatorname{tg} \theta}$, 因此, $e^{2i\theta} = \mu$ 。

故 $\theta = \frac{1}{2i} \ln \mu$ 。

拉格尔定理与命题 1.2.6 一样 (事实上, 它们是相互等价的), 也十分重要, 因为这个定理用交比射影概念表达了角度量概念, 形成了角度的射影解释, 从而将欧氏几何与射影几何联系起来。

推论 1.2.1 两条非迷向直线相互垂直的充要条件是这两条直线与过交点以 $i, -i$ 为斜率的两条迷向直线成调和共轭, 或等价地说, 两条非迷向直线相互垂直的充要条件是这两条直线上的无穷远点与两个圆环点成调和共轭。

1.3 二维射影变换

1.3.1 二维射影变换

射影变换

射影变换是射影平面上的可逆齐次线性变换, 这个变换可由 3×3 的矩阵来描述:

$$\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad (1.3.1)$$

或更简略地记为 $\mathbf{x}' = H\mathbf{x}$ 。

射影变换有时又称为单应, 而矩阵 H 称为射影变换矩阵或称为单应矩阵。由于变换是齐次的(点使用了齐次坐标), 所以同一个射影变换矩阵 H 可以相差一个非零常数因子, 因此射影变换仅有 8 个自由度, 即射影变换矩阵可由它的元素所构成的 8 个比值所确定。

例 1.3.1 投影中心不在物体平面上的中心投影是一个射影变换(如图 1.3.1 所示)。中心投影将物体平面上的点投影到图像平面上得到像点, 像点是物体平面点和投影中心的连线与像平面的交点, 物体平面点到像点之间的变换是一个射影变换。物体平面上的无穷远点的像点是该无穷远点与投影中心的连线(平行于物体平面)与像平面的交点, 一般地该交点是像平面上的有限点(即该点在图像平面中的齐次坐标的第三个分量不为零)。物体平面上的无穷远线的像是通过投影中心且平行于物体平面的平面与像平面的交线, 一般地它是像平面上的一条有限直线。

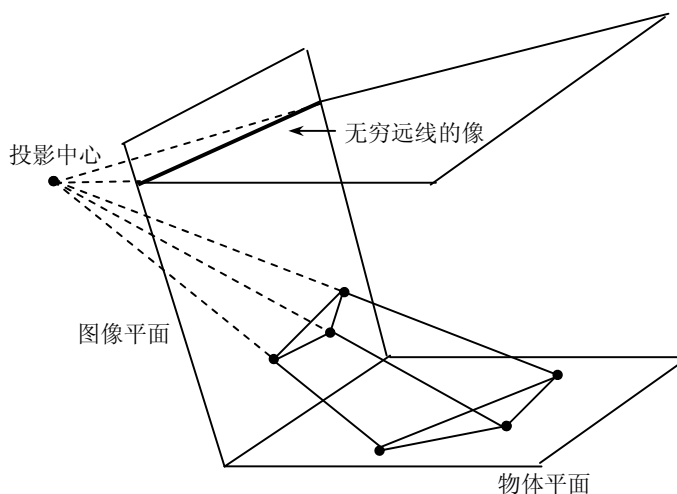


图 1.3.1: 投影中心不在物体平面上的中心投影

任何射影变换都将点变换到点, 并且保持点的共线性质, 因而将直线变为直线。射影变换将点变换到点的性质是由其定义所确定的。下面说明射影变换保持点的共线性: 令 $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ 是三个共线点, 即 $\det(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = 0$, 经射影变换 H 后的三个点分别为 $\mathbf{x}'_1, \mathbf{x}'_2, \mathbf{x}'_3$, 于是,

$$\det(\mathbf{x}'_1, \mathbf{x}'_2, \mathbf{x}'_3) = \det(H\mathbf{x}_1, H\mathbf{x}_2, H\mathbf{x}_3) = \det(H) \det(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = 0$$

因此, 三点 $\mathbf{x}'_1, \mathbf{x}'_2, \mathbf{x}'_3$ 共线。任何射影变换的逆变换(对应于单应矩阵的逆)都是射影变换、任意两个的合成(对应于两个单应矩阵的积)也都是射影变换(如图 1.3.2 所示), 因此射影变换的全体构成射影平面上的一个变换群。

例 1.3.2 中心投影变换的合成是射影变换。如图 1.3.2 所示，图中第一个中心投影变换是 H ，第二个中心投影变换是 G ，由这两个投影得到一个从第一个像平面到第二个相平面的变换是 F 。由于 H, G 都是射影变换，它们的逆变换是像点沿投影线反投到物体平面上的点，对应的变换矩阵分别是 H 与 G 的逆矩阵，因此逆变换也是射影变换。变换 F 是 H 的逆变换与变换 G 的合成，它可以用 3×3 的可逆矩阵 GH^{-1} 来描述，所以也是一个射影变换。但它不再是中心投影变换而是一般的射影变换。

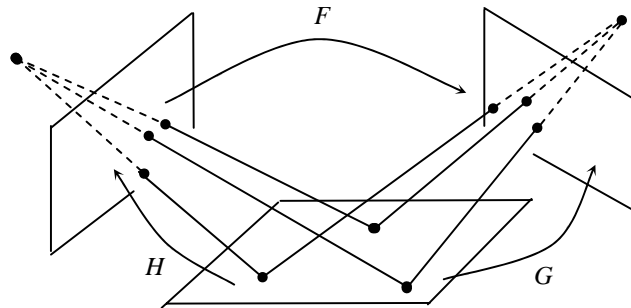


图 1.3.2 中心投影变换合成是一般的投影变换。

如果被变换点 \mathbf{x} 是欧氏坐标系下的齐次坐标，则无穷远点 $\mathbf{x}_\infty = (x_1, x_2, 0)^T$ 的射影变换是

$$\mathbf{x}'_\infty = H\mathbf{x}_\infty = (h_{11}x_1 + h_{12}x_2, h_{21}x_1 + h_{22}x_2, h_{31}x_1 + h_{32}x_2)^T$$

一般地， $h_{31}x_1 + h_{32}x_2 \neq 0$ 。这样，无穷远点变换后的坐标不再有第 3 个分量为零的形式。事实上，射影变换(1.3.1)等价于坐标基的变换，变换后的坐标基称为射影坐标基。上述观察等价于无穷远点在一般坐标基下，第 3 个分量不为零。由于射影几何主要是讨论射影变换群下的不变几何性质与不变量的理论，这就是在研究射影性质时不使用无穷远点术语而把无穷远点和非无穷远点同等对待的理由。但是无穷远点在计算机视觉中具有特别重要的作用，所以在本书中仍使用无穷远点的术语。

满足式 (1.3.1) 的一对点 $\{\mathbf{x}, \mathbf{x}'\}$ 称为射影变换的一个点对应，记为 $\mathbf{x} \leftrightarrow \mathbf{x}'$ 。在一般情况下，4 个点对应唯一确定一个二维射影变换。下述命题是更确切的陈述：

命题 1.3.1 从 4 个点对应唯一确定二维射影变换的充要条件是 4 个点对应中任意三点不共线，并且可以由下述公式计算这个射影变换：

$$H = s_4(\mathbf{x}'_1, \mathbf{x}'_2, \mathbf{x}'_3) \text{diag} \left(\frac{p'_1}{p_1}, \frac{p'_2}{p_2}, \frac{p'_3}{p_3} \right) (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)^{-1} \quad (1.3.2)$$

其中: $(p_1, p_2, p_3)^T = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)^{-1} \mathbf{x}_4$, $(p'_1, p'_2, p'_3)^T = (\mathbf{x}'_1, \mathbf{x}'_2, \mathbf{x}'_3)^{-1} \mathbf{x}'_4$ 。

证明 由于 $\mathbf{x}_j \leftrightarrow \mathbf{x}'_j, (1 \leq j \leq 4)$ 是 4 个点对应, 所以存在常数 $s_j (1 \leq j \leq 4)$ 使得

$$s_j \mathbf{x}'_j = H \mathbf{x}_j, (1 \leq j \leq 4) \quad (1.3.3)$$

充分性: 由式(1.3.3), 可得到:

$$H(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = (s_1 \mathbf{x}'_1, s_2 \mathbf{x}'_2, s_3 \mathbf{x}'_3) = (\mathbf{x}'_1, \mathbf{x}'_2, \mathbf{x}'_3) \text{diag}(s_1, s_2, s_3)$$

于是,

$$H = (\mathbf{x}'_1, \mathbf{x}'_2, \mathbf{x}'_3) \text{diag}(s_1, s_2, s_3) (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)^{-1} \quad (1.3.4)$$

由 $\mathbf{x}'_4 = s_4 H \mathbf{x}_4$, 得到:

$$(\mathbf{x}'_1, \mathbf{x}'_2, \mathbf{x}'_3) \text{diag}(s_1, s_2, s_3) (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)^{-1} \mathbf{x}_4 = s_4 \mathbf{x}'_4$$

所以,

$$\text{diag}(s_1, s_2, s_3) (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)^{-1} \mathbf{x}_4 = s_4 (\mathbf{x}'_1, \mathbf{x}'_2, \mathbf{x}'_3)^{-1} \mathbf{x}'_4$$

令

$$\mathbf{p} = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)^{-1} \mathbf{x}_4 = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}, \quad \mathbf{p}' = (\mathbf{x}'_1, \mathbf{x}'_2, \mathbf{x}'_3)^{-1} \mathbf{x}'_4 = \begin{pmatrix} p'_1 \\ p'_2 \\ p'_3 \end{pmatrix}$$

则对任意的 p_j 必有 $p_j \neq 0$ 。否则, 例如: $p_1 = 0$, 则有

$$\mathbf{x}_4 = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \begin{pmatrix} 0 \\ p_2 \\ p_3 \end{pmatrix} = p_2 \mathbf{x}_2 + p_3 \mathbf{x}_3$$

因此 $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ 共线, 矛盾。于是, $s_j = \frac{s_4 p'_j}{p_j}, (1 \leq j \leq 3)$, 将它代入式(1.3.4)可得式 (1.3.2)。

必要性: 反证: 若存在三个共线点, 不妨假定 $\mathbf{x}_j (1 \leq j \leq 3)$ 是三个共线点, 则必有

$$\mathbf{x}_3 = a \mathbf{x}_1 + b \mathbf{x}_2$$

于是,

$$\lambda H \mathbf{x}_3 = a H \mathbf{x}_1 + b H \mathbf{x}_2$$

因此, (1.3.3)中至多有三组方程是独立的, 故不可能在相差一个常数因子的意义下确定单应矩阵 H 。证毕。

1.3.2 直线与二次曲线的射影变换

直线的变换规则

射影变换(1.3.1)是由点的变换规则来定义的，以后说射影变换 H 均是指满足点变换规则(1.3.1)的射影变换。下面讨论直线在射影变换 H 下变换规则。

令 l 是平面上的一条直线， l' 是经过射影变换 H 后的直线。由于 $\forall x \in l, x' = Hx \in l'$ ，所以，

$$l'^T x' = l'^T Hx = 0 \quad (1.3.5)$$

因此， $l'^T = l'^T H$ ，即 $l' = H^{-T} l$ 。射影变换 H^{-T} 称为变换 H 的对偶。于是，直线的变换规则由点变换的对偶给出，即我们有下述命题：

命题 1.3.2 射影变换 H 对直线的变换规则，由 H 的对偶所确定：

$$l' = H^{-T} l \quad (1.3.6)$$

满足式(1.3.6)的一对直线 $\{l, l'\}$ 称为射影变换 H 的一个线对应，并记作 $l \leftrightarrow l'$ 。与点对应一样，一般地，4 个线对应唯一确定一个射影变换并且有类似于式(1.3.2)的计算公式。

二次曲线的变换规则

令 C 是平面上的一条二次曲线， C' 是经过射影变换后的曲线。由于 $\forall x \in C, x' = Hx \in C'$ ，所以，

$$x'^T Cx' = x'^T H^{-T} C H^{-1} x' = 0$$

这样，必有 $C' = H^{-T} C H^{-1}$ 。因此，二次曲线 C 经过射影变换后仍是一条二次曲线。设 D 是一个可逆矩阵，则称矩阵变换 $Y = D X D^T$ 为合同变换。因此，射影变换 H 对二次曲线的变换规则是对偶合同。

命题 1.3.3 射影变换 H 对二次曲线的变换规则，由 H 的对偶合同所确定：

$$C' = H^{-T} C H^{-1} \quad (1.3.7)$$

应用对偶原理，有下述命题：

命题 1.3.4 射影变换 H 关于对偶二次曲线 C^* 的变换规则，由 H 合同所确定：

$$C^{*'} = H C^* H^T \quad (1.3.8)$$

二次曲线的射影分类

二次曲线的射影分类是指二次曲线在射影变换下的等价类。二次曲线由对称矩阵 C 来表示，根

据第 7 章关于对称矩阵的特征分解理论, 不难证明对任意对称矩阵 C , 必存在可逆矩阵 H 使得 $H^{-T}CH^{-1} = \text{diag}(s_1, s_2, s_3)$, 其中 $s_j = \pm 1$, 或 0 。根据二次曲线的变换规则, 任何二次曲线 C 都可以通过射影变换变为具有上述对角矩阵形式的二次曲线。因此, 可以得到二次曲线的如下射影分类:

表 1.1.1: (实)二次曲线的射影分类

对角元素	方程	类型
(1,1,1)	$x^2 + y^2 + t^2 = 0$	无实点 (假二次曲线)
(1,1,-1)	$x^2 + y^2 - t^2 = 0$	圆
(1,1,0)	$x^2 + y^2 = 0$	实点 (0,0,1)
(1,-1,0)	$x^2 - y^2 = 0$	两条直线
(1,0,0)	$x^2 = 0$	二重直线

1.4 变换群与不变量

平面上的所有射影变换构成一个变换群, 通常称这个群为射影变换群。几何学的主要内容是研究在各种变换群作用下的几何不变量 (包括不变几何性质), 在射影变换群中包含两类重要的子群: 欧氏变换群与仿射变换群。本节重点介绍这两类子群以及相应的不变量。

1.4.1 等距变换群

等距变换群

等距变换是指保持距离不变的变换, 其定义如下:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} \sigma \cos \theta & -\sin \theta & x_0 \\ \sigma \sin \theta & \cos \theta & y_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (1.4.1)$$

其中, $\sigma = \pm 1$ 。使用非齐次坐标, 上式可以写成下面的形式:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \sigma \cos \theta & -\sin \theta \\ \sigma \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \quad (\tilde{\mathbf{x}}' = U\tilde{\mathbf{x}} + \tilde{\mathbf{x}}_0)$$

由此可以看出，等距变换是先作正交变换，然后再作平移变换所得到的变换。正交变换与平移变换都保持距离不变，因此等距变换也保持距离不变。等距变换的逆变换仍是一个等距变换，两个等距变换的合成变换也是一个等距变换，所以等距变换的全体构成一个变换群，称为等距变换群。

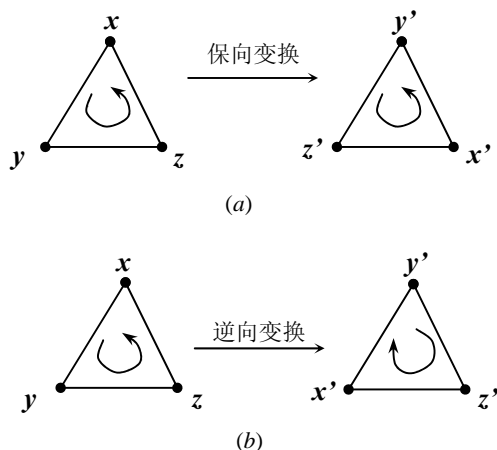


图 1.4.1 保向变换与逆向变换：(a) 保持点的顺序不变，是一个保向变换；
(b) 变换后对应三点的顺序与原来反向，是一个逆向变换。

正交变换 U 根据它的行列式是否等于 1 而分为旋转变换与反射变换。当 $\det(U)=1$ 时，是旋转变换； $\det(U)=-1$ 时，是反射变换。它们的几何意义是旋转变换不但保持两点的距离不变，而且还保持方向(保向)不变，而反射变换是一个逆向变换(如图 1.4.1 所示)。

欧氏变换群

在等距变换(1.4.1)中，如果矩阵 U 是一个旋转矩阵，则这个等距变换称为欧氏变换。不难验证欧氏变换的全体也构成一个变换群，通常称它为欧氏变换群(简称欧氏群)。欧氏变换群是距变换群的子群。欧氏群可以更简洁地表示为

$$\mathbf{x}' = H_e \mathbf{x} = \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix} \mathbf{x} \quad (1.4.2)$$

平面欧氏变换有 3 个自由度(因为在平面上，旋转有 1 个自由度，平移有 2 个自由度)。因此，两个点对应可确定欧氏变换。值得注意的是，矩阵 U 为反射的等距变换不能构成等距变换群的子群，因为两个这样变换的合成是一个欧氏变换。

欧氏不变量

等距变换群的不变量主要有：两点的距离、两线的夹角、图形的面积等。由于欧氏群是等距变

换群的子群，因此等距变换群的不变量也是欧氏不变量。下面给出一个在计算机视觉中经常使用的欧氏不变性质：

命题 1.4.1 欧氏变换保持圆环点不变，因此也保持无穷远直线不变。

反射等距变换将两个圆环点互换，即反射等距变换只能保持两个圆环点的整体不变。当然，它也是保持无穷远直线不变的。

1.4.2 相似变换群

相似变换

相似变换是等距变换与均匀伸缩变换的合成变换，所谓均匀伸缩变换是指下述变换：

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

其中 s 是均匀伸缩因子。

相似变换，顾名思义，它是保持图形相似的变换。在初等几何中，相似分为旋转相似(保向)和对称相似(逆向)。旋转相似是欧氏变换与均匀伸缩变换的合成，而对称相似是反射等距变换与均匀伸缩变换的合成。

在计算机视觉中最关心的是旋转相似，它可用下面的矩阵形式来表示：

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} s \cos \theta & -s \sin \theta & x_0 \\ s \sin \theta & s \cos \theta & y_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad \left(\mathbf{x}' = H_s \mathbf{x} = \begin{pmatrix} sR & t \\ 0 & 1 \end{pmatrix} \mathbf{x} \right) \quad (1.4.3)$$

旋转相似变换有 4 个自由度，因为它比欧氏变换多一个均匀伸缩因子。与欧氏变换一样，两个点对应也可以确定相似变换。相似变换的全体也构成一个变换群，通常称为相似变换群。旋转相似变换是相似变换群的子群，而欧氏群又是旋转相似变换群的子群。非旋转相似变换不能构成相似变换群的子群。

相似不变量

相似变换群的不变量有：两直线的夹角，长度的比值，面积的比值。这些性质是很容易验证的。下面的命题是非常重要的，因为它在计算机视觉中扮演着非常重要的角色。

命题 1.4.2 (1) 射影变换保持圆环点不动的充要条件是它为相似变换; (2) 射影变换保持对偶二次曲线 C_{∞}^* 不动的充要条件是它为相似变换。

证明 因为

$$\begin{pmatrix} s \cos \theta & -s \sin \theta & x_0 \\ s \sin \theta & s \cos \theta & y_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix} = s e^{i\theta} \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix} = \mathbf{I} \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}$$

所以, 相似变换保持圆环点不变。类似地, 可以证明对另一个圆环点也保持不变。

反之, 如果

$$H \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix} = \begin{pmatrix} a & b & x_0 \\ c & d & y_0 \\ e & f & 1 \end{pmatrix} \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}$$

则必有

$$a + ib = \lambda, c + id = i\lambda, e - if = 0$$

最后的等式表明 $e=f=0$ 。令 $\lambda = \lambda_1 + i\lambda_2$, 则由前两等式得到:

$$a = \lambda_1, b = \lambda_2, c = -\lambda_2, d = \lambda_1$$

令

$$s = \sqrt{\lambda_1^2 + \lambda_2^2}, \cos \theta = \frac{\lambda_1}{\sqrt{\lambda_1^2 + \lambda_2^2}}, \sin \theta = -\frac{\lambda_2}{\sqrt{\lambda_1^2 + \lambda_2^2}}$$

则有

$$H = \begin{pmatrix} a & b & x_0 \\ c & d & y_0 \\ e & f & 1 \end{pmatrix} = \begin{pmatrix} s \cos \theta & -s \sin \theta & x_0 \\ s \sin \theta & s \cos \theta & y_0 \\ 0 & 0 & 1 \end{pmatrix}$$

因此, H 是一个相似变换。

(2) 可由对偶原理从(1)直接得到, 因为相似变换 H 的对偶 H^{-T} 仍是一个相似变换。

1.4.3 仿射变换群

仿射变换

仿射变换定义为

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a & b & x_0 \\ c & d & y_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad \left(\mathbf{x}' = H_a \mathbf{x} = \begin{pmatrix} A & t \\ 0 & 1 \end{pmatrix} \mathbf{x} \right) \quad (1.4.4)$$

其中 A 是一个 2 阶可逆矩阵。仿射变换有 6 个自由度，3 个不共线的点对应唯一确定仿射变换。仿射变换的全体也构成一个变换群，称为仿射变换群。相似变换群是它的子群。

仿射变换的分解

除平移变换外，只须对矩阵 A 进行分解。对矩阵 A 作奇异值分解(SVD 分解，见第 8 章)，我们得到 $A = UDVT^T$ ，其中 U, V 是正交矩阵， D 是对角元为正数的对角矩阵

$$D = \begin{pmatrix} s_x & \\ & s_y \end{pmatrix}$$

由此可以看出，仿射变换是一个等距变换 V^T 、一个非均匀伸缩变换 D 以及另一个等距变换 U 的合成，因此它与相似变换的差别在于非均匀伸缩。

仿射变换(1.4.4)是否保向，根据矩阵 A 的行列式 $\det(A)$ 是否大于零来确定。为了看出这一点，我们将 A 写成 $A = (UV^T)(VDV^T)$ ，由于 (VDV^T) 总是一个保向的变换(不论 V 是否为旋转矩阵)， (UV^T) 是否保向是由它的行列式是否为 +1(即是否为旋转矩阵)来确定，而行列式 $\det(A)$ 的符号与 (UV^T) 的符号是一致的。

所有保向仿射变换构成仿射变换群的子群，而旋转(保向)相似变换群又是它的子群。

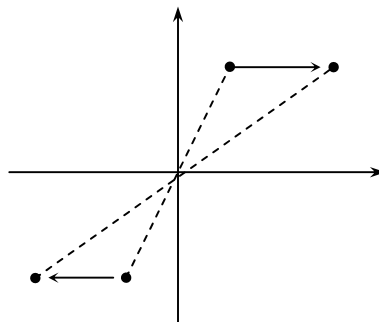


图 1.4.2 推移变换 P : 位于 x 轴上的点保持不动，不在 x 轴上的点沿 x 轴平行移动(在轴两侧移动，方向相反)

仿射变换的另一种分解。对 A 作 QR 分解^{*)}得到 $A=UK$ ，其中 U 是一个正交阵， K 是一个对角元素均大于零的上三角阵：

$$K = \begin{pmatrix} s_x & e \\ & s_y \end{pmatrix}$$

再对 K 再作如下分解 $K = \underbrace{\begin{pmatrix} s_x & \\ & s_y \end{pmatrix}}_D \underbrace{\begin{pmatrix} 1 & e/s_x \\ & 1 \end{pmatrix}}_P = DP$ ，于是，仿射变换可以表示为

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = UDP \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

变换 $P = \begin{pmatrix} 1 & e/s_x \\ & 1 \end{pmatrix}$ 通常称为推移变换(如图 1.4.2 所示)。因此，仿射变换(除一个平移变换外)是推移变换、非均匀伸缩变换与正交变换的合成。

仿射不变量

- **仿射变换保持平行性不变** 不难验证：仿射变换将无穷远点变换到无穷远点。所以，仿射变换保持平行性不变，也就是说平面上任何两条平行线(或线段)经过仿射变换后的两条直线(或线段)仍然是平行的。因此，平行性是仿射变换的不变性质。
- **仿射变换保持面积的比值不变** 不难验证：若仿射变换将图形 G 变换到图形 $G' = H_a G$ ，则 $\sigma(G') = \sigma(H_a G) = |\det(A)| \cdot \sigma(G)$ (σ 代表面积)，这是因为，仿射变换的 Jacobi 行列式为 $\det(A)$ 。所以，仿射变换保持面积的比值不变。
- **仿射变换保持平行线段长度的比值不变** 这一性质可由前两个性质导出，请读者自己给出证明。

此外，在计算机视觉中经常用到的一个结论是下述命题：

命题 1.4.3 射影变换 H 保持无穷远直线不动的充要条件是 H 为仿射变换。

注意：这里所说保持无穷远直线不动的意思是将无穷远点变换到无穷远点(可能不是同一点)，即将直线看作一个整体是保持不变的。

证明 令 $H = \begin{pmatrix} A & t \\ 0^T & 1 \end{pmatrix}$ 是仿射变换，则根据射影变换对线的变换规则(命题 1.3.6)，可知：

^{*)}见 8.1 节

$$l' = H^{-T} l_\infty = \begin{pmatrix} A^{-T} & 0 \\ -t^T A^{-T} & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = l_\infty$$

反之, 若射影变换 $H = \begin{pmatrix} A & a \\ b^T & 1 \end{pmatrix}$ 使得

$$\begin{pmatrix} A & a \\ b^T & 1 \end{pmatrix}^{-T} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

则必有 $b=0$ 。因此, H 为仿射变换。证毕。

二次曲线的仿射分类

对于(非退化)二次曲线在欧氏变换下的分类: 椭圆、抛物线与双曲线, 对仿射变换仍然是有效的。由于椭圆与无穷远直线没有实交点、抛物线与无穷远直线相切, 即有两个接触点、双曲线与无穷远直线有两个实交点, 而仿射变换保持无穷远直线不动且保持交点性质不变(实变实、虚变虚), 所以前面的性质是仿射不变的。因此, 二次曲线的仿射分类仍然是: 椭圆、抛物线与双曲线三类。

1.4.4 射影变换群

射影变换与其它变换一样也可以写成分块矩阵的形式:

$$x' = Hx = \begin{pmatrix} A & t \\ v^T & k \end{pmatrix} x \quad (1.4.5)$$

$k=0$ 当仅当这个射影变换将无穷远直线变换为通过坐标原点的直线, 因此在一般情况下, $k \neq 0$ 。当 $k \neq 0$ 时, H 可分解为下述形式:

$$H = \begin{pmatrix} sR & t/k \\ 0 & 1 \end{pmatrix} \begin{pmatrix} K & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} I & 0 \\ v^T & k \end{pmatrix} \quad (1.4.6)$$

其中 K 是行列式等于 1 且对角元素均大于零的上三角矩阵, R 是正交矩阵。

显然, $H_p = \begin{pmatrix} I & 0 \\ v^T & k \end{pmatrix}$ 是改变无穷远直线的射影变换, $H_a = \begin{pmatrix} K & 0 \\ 0 & 1 \end{pmatrix}$ 是保持面积比不变的仿射变

换，而 $H_s = \begin{pmatrix} sR & t/k \\ 0 & 1 \end{pmatrix}$ 是相似变换。式(1.4.6)是不难证明的：因为

$$\begin{aligned} H \begin{pmatrix} I & 0 \\ \mathbf{v}^T & k \end{pmatrix}^{-1} &= \begin{pmatrix} A & t \\ \mathbf{v}^T & k \end{pmatrix} \begin{pmatrix} I & 0 \\ -\mathbf{v}^T/k & 1/k \end{pmatrix} \\ &= \begin{pmatrix} A - t\mathbf{v}^T/k & t/k \\ 0^T & 1 \end{pmatrix} = \begin{pmatrix} sRK & t/k \\ 0^T & 1 \end{pmatrix} = \begin{pmatrix} sR & t/k \\ 0 & 1 \end{pmatrix} \begin{pmatrix} K & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

倒数第二个等式，是利用 QR 分解并将分解中的上三角矩阵行列式归一化所得的结果。所以，我们有

$$H = \begin{pmatrix} sR & t/k \\ 0 & 1 \end{pmatrix} \begin{pmatrix} K & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} I & 0 \\ \mathbf{v}^T & k \end{pmatrix}$$

与仿射变换不同的是射影变换不再有保向与逆向之分，这是因为一般的射影变换将无穷远直线 l_∞ 变到一条有限直线 l'_∞ ，在源平面上的两个有序图形，如果被变换到直线 l'_∞ 的两侧，则必存在一个图形与原来的图形反序，而另一个图形与原来的同序。如图 1.4.3 所示。

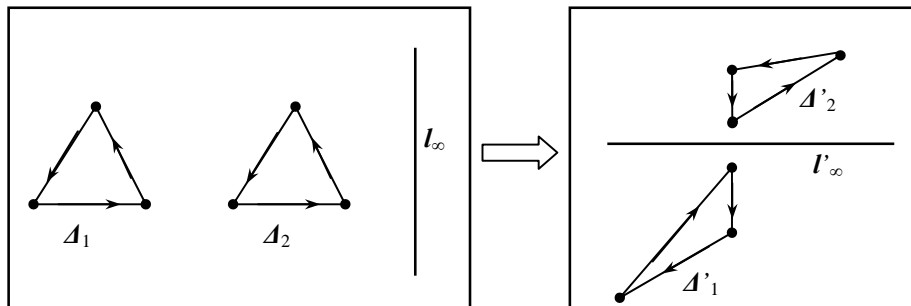


图 1.4.3：（非仿射的）射影变换不是保向变换也不是逆向变换。

射影不变量

基本射影不变量是四共线点的交比。

关于共线点的交比，在 1.1.3 节已给出定义。如果直线 l 上 4 个点 $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4$ 的齐次坐标为 $\mathbf{x}_j = (x_{1j}, x_{2j})^T$ ，则它们的交比是

$$(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}_3, \mathbf{x}_4) = \frac{\det(\mathbf{x}_1, \mathbf{x}_3)}{\det(\mathbf{x}_2, \mathbf{x}_3)} : \frac{\det(\mathbf{x}_1, \mathbf{x}_4)}{\det(\mathbf{x}_2, \mathbf{x}_4)} \quad (1.4.7)$$

4 点交比在一维射影变换下是不变的(如图 1.4.4 所示)，换句话说交比的定义不依赖于直线 l 的坐标系的选择。

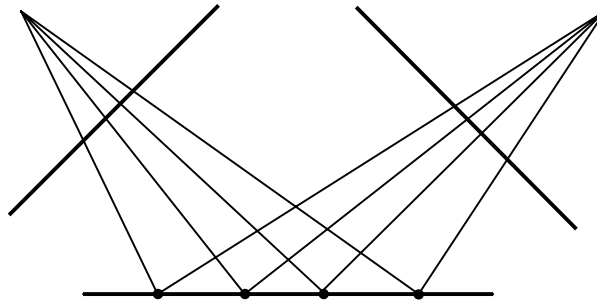


图 1.4.4: 与图中两个线束相交的所有直线上的 4 个交点的交比均相等，因为它们之间满足一维射影变换。

所谓一维射影变换，在代数上与二维射影变换类似，是指线 l 上的可逆齐次线性变换，这个变换由 2×2 的矩阵 H 来描述：

$$\begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

显然 $\det(\mathbf{x}'_i, \mathbf{x}'_j) = \det(H\mathbf{x}_i, H\mathbf{x}_j) = \det(H)\det(\mathbf{x}_i, \mathbf{x}_j)$ ，而 $\det(H)$ 在交比的比值中自动消除。因此一维射影变换保持交比不变。如果 4 个点都是有限点，则可以将它们第二个坐标归一化，写成 $\mathbf{x}_j = (x_{1j}, 1)^T$ ，则 $\det(\mathbf{x}_i, \mathbf{x}_j) = x_{1i} - x_{1j}$ ，它表示两点之间的有向距离。这样，就可以通过有向距离来计算交比。

在平面上，任何二维射影变换 H 都可以诱导出直线的一维射影变换，由此，立即得到平面射影变换保持交比不变的结论。下面提供了一种诱导一维射影变换方法。在直线 l 上取两个不同的点，并给定齐次坐标 $\mathbf{x}_1, \mathbf{x}_2$ 。令 $\mathbf{x}'_j = H\mathbf{x}_j$ ($j=1,2$)，如 1.1.3 节的线参数化那样，利用下式：

$$\mathbf{x} = u\mathbf{x}_1 + v\mathbf{x}_2 = (x_1, x_2) \begin{pmatrix} u \\ v \end{pmatrix}, \quad \mathbf{x}' = u'\mathbf{x}'_1 + v'\mathbf{x}'_2 = (x'_1, x'_2) \begin{pmatrix} u' \\ v' \end{pmatrix}$$

定义点 \mathbf{x} 和对应点 \mathbf{x}' 的齐次坐标分别为 $\begin{pmatrix} u \\ v \end{pmatrix}, \begin{pmatrix} u' \\ v' \end{pmatrix}$ ，于是，

$$(\mathbf{x}'_1, \mathbf{x}'_2) \begin{pmatrix} u' \\ v' \end{pmatrix} = \mathbf{x}' = H\mathbf{x} = H(\mathbf{x}_1, \mathbf{x}_2) \begin{pmatrix} u \\ v \end{pmatrix}$$

因此,

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = (\mathbf{x}'_1, \mathbf{x}'_2)^+ H(\mathbf{x}_1, \mathbf{x}_2) \begin{pmatrix} u \\ v \end{pmatrix} = \hat{H} \begin{pmatrix} u \\ v \end{pmatrix}$$

$\hat{H} = (\mathbf{x}'_1, \mathbf{x}'_2)^+ H(\mathbf{x}_1, \mathbf{x}_2)$ 是一个 2 阶可逆矩阵(这里矩阵的上标 “+” 表示矩阵的广义逆), 且它对线 l 的作用与 H 对线 l 的作用是相同的。因此, 它是由 H 在线 l 上诱导的一维射影变换。

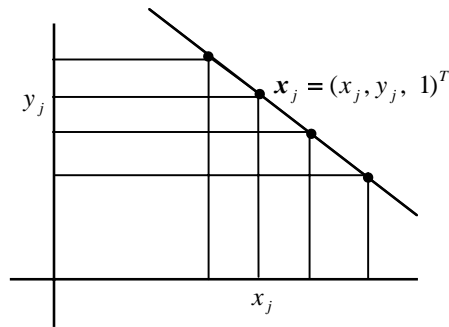


图 1.4.5: 平面上 4 个共线点的有限点的交比, 可以通过它的非齐次坐标分量来计算, 因为直线点到坐标轴的投影是一维射影变换。

如果 4 个共线点是平面上的有限点, 我们不需要通过线参数化, 再利用式(1.4.7)计算交比。因为此时可以将点的第三个坐标归一化, 写成 $\mathbf{x}_j = (x_j, y_j, 1)^T$, 于是 $x_j(y_j)$ 是沿 $y(x)$ 方向在轴 $x(y)$ 上的投影(如图 1.4.5 所示), 而投影变换(直线到坐标轴的投影)是一维射影变换, 因此 4 点的交比与它们在各坐标轴上的投影点的交比相同。也就是说, 我们可以通过非齐次坐标分量来计算平面上共线点的交比。

2. 空间射影几何

2.1 射影空间

2.1.1 空间点

假定在空间建立了欧氏坐标系, 空间点的欧氏坐标记为 $\tilde{\mathbf{X}} = (x, y, z)^T$, 令

$$\frac{x_1}{x_4} = x, \frac{x_2}{x_4} = y, \frac{x_3}{x_4} = z, x_4 \neq 0$$

定义空间点的齐次坐标为 $\mathbf{X} = (x_1, x_2, x_3, x_4)^T$ 。

当 $s \neq 0$ 时, $s\mathbf{X}$ 与 \mathbf{X} 表示同一空间点的齐次坐标, 即空间点的齐次坐标可以相差一个非零常数因子。令 $x_4 \rightarrow 0$, 除 $x_1 = 0, x_2 = 0, x_3 = 0$ 外, 下述三式至少有一式成立:

$$x = \frac{x_1}{x_4} \rightarrow \infty, y = \frac{x_2}{x_4} \rightarrow \infty, z = \frac{x_3}{x_4} \rightarrow \infty$$

因此, 定义齐次坐标第 4 个分量 $x_4 = 0$ 的点为无穷远点。

这样, 只要 x_1, x_2, x_3, x_4 不同时为零, $\mathbf{X} = (x_1, x_2, x_3, x_4)^T$ 就代表扩展空间(包括所有无穷远点的三维空间)中的一个点, 反之扩展空间中的每一点都可以用不同时为零的 4 个数构成的齐次坐标 $\mathbf{X} = (x_1, x_2, x_3, x_4)^T$ 来表示; $x_4 \neq 0$ 时代表有穷点(非无穷远点), $x_4 = 0$ 时代表无穷远点。称这样扩展的三维空间为**三维射影空间**。注意: $(0,0,0,0)^T$ 不能作为三维射影空间中任何点的齐次坐标。

2.1.2 空间平面

在三维射影空间中, 平面方程可以写成

$$\pi_1 x + \pi_2 y + \pi_3 z + \pi_4 w = 0 \quad (2.1.1)$$

其中 $\mathbf{X} = (x, y, z, w)^T$ 表示空间点的齐次坐标。称 4 维向量 $\boldsymbol{\pi} = (\pi_1, \pi_2, \pi_3, \pi_4)^T$ 为该平面的齐次坐标。显然, 方程(2.1.1) 两边同乘以一个非零常数仍表示该平面, 所以, 平面的齐次坐标 $\boldsymbol{\pi}$ 仅依赖于三对独立的比值 $\{\pi_1 : \pi_2 : \pi_3 : \pi_4\}$, 也就是说, 平面在三维空间中有 3 自由度。方程(2.1.1)可以写成更简

洁的形式

$$\pi^T \mathbf{X} = 0 \quad (2.1.2)$$

如果 $\pi = (0,0,0,1)^T$, 则方程(2.1.2)的解集为

$$\{\mathbf{X} = (\tilde{\mathbf{X}}^T, 0) \mid \tilde{\mathbf{X}} \in R^3 - \{0\}\}$$

它是所有无穷远点所构成的集合。因此, 称平面 $\pi = (0,0,0,1)^T$ 为无穷远平面, 并记为 π_∞ 。

如果 $\pi \neq \pi_\infty$, 则该平面上的有限点 $\mathbf{X} = (\tilde{\mathbf{X}}^T, 1)^T$ 满足方程:

$$\mathbf{n}^T \tilde{\mathbf{X}} + d = 0$$

其中 $\mathbf{n} = (\pi_1, \pi_2, \pi_3)^T$, $d = \pi_4$, $\frac{|d|}{\|\mathbf{n}\|}$ 是坐标原点到该平面的距离。不难看出它是欧氏几何中的平面法式方程。该平面上的无穷远直线由下述方程给出:

$$\mathbf{n}^T \tilde{\mathbf{X}} = \pi^T \begin{pmatrix} \tilde{\mathbf{X}} \\ 0 \end{pmatrix} = 0$$

即, 平面 π 法向量 \mathbf{n} 是该平面上无穷远直线的表示。因此, 平面上的无穷远直线代表了该平面的法向。

下述结论是明显的几何事实:

1. 两平面平行的充要条件是它们的交线为无穷远直线, 或者说它们有相同的方向;
2. 直线与直线(面)平行的充要条件是它们相交于无穷远点。

三点确定一个平面

假定 $\mathbf{X}_j, j=1,2,3$ 是平面 π 上的三个点, 则必有

$$\begin{pmatrix} \mathbf{X}_1^T \\ \mathbf{X}_2^T \\ \mathbf{X}_3^T \end{pmatrix} \pi = 0 \quad (2.1.3)$$

如果三点 $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$ 不共线(通常称它们处于一般位置), 则方程(2.1.3)中系数矩阵的秩必为 3, 此时平面 π 是系数矩阵的一维(右)零空间的元素(相差一个齐次因子), 因此一般位置上的三个点唯一确定一个平面。

如果三点 $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$ 共一条直线 L , 则系数矩阵的秩为 2, 因此系数矩阵有二维(右)零空间, 此时不能唯一确定平面 π 。实际上, 通过直线 L 的所有平面都满足方程(2.1.3), 即方程(2.1.3)确定了以直线 L 为轴的平面束。

假定三点 $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$ 处于一般位置, 令 $\mathbf{X} = (x, y, z, w)^T$ 是这三个点所确定的平面 π 上的任一点, 则 \mathbf{X} 是 $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$ 的线性组合, 即 $\det(\mathbf{X}, \mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3) = 0$, 而

$$\det(\mathbf{X}, \mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3) = xd_{234} - yd_{134} + zd_{124} - wd_{123}$$

其中 d_{jkl} 是由矩阵 $(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3)$ 的第 j, k, l 行所构成的行列式。所以, 三点 $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$ 所确定的平面 π 的坐标为

$$\pi = (d_{234}, -d_{134}, d_{124}, -d_{123})^T \quad (2.1.4)$$

实际上, 它是方程(2.1.3)的非零解向量。

如果三点 $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$ 是有穷点, 则它的齐次坐标可以写成下面的形式:

$$\mathbf{X}_1 = \begin{pmatrix} \tilde{\mathbf{X}}_1 \\ 1 \end{pmatrix}, \mathbf{X}_2 = \begin{pmatrix} \tilde{\mathbf{X}}_2 \\ 1 \end{pmatrix}, \mathbf{X}_3 = \begin{pmatrix} \tilde{\mathbf{X}}_3 \\ 1 \end{pmatrix}$$

根据式(2.1.4), 它们所确定的平面 π 坐标为

$$\pi = \begin{pmatrix} (\tilde{\mathbf{X}}_1 - \tilde{\mathbf{X}}_3) \times (\tilde{\mathbf{X}}_2 - \tilde{\mathbf{X}}_3) \\ -\tilde{\mathbf{X}}_3^T (\tilde{\mathbf{X}}_1 \times \tilde{\mathbf{X}}_2) \end{pmatrix} \quad (2.1.5)$$

这与欧氏几何中的结果是一致的, 例如平面 π 的法向量为 $\mathbf{n} = (\tilde{\mathbf{X}}_1 - \tilde{\mathbf{X}}_3) \times (\tilde{\mathbf{X}}_2 - \tilde{\mathbf{X}}_3)$ 。

三个平面确定一点

在空间中点与平面对偶的, 而直线是自对偶的。对换式(2.1.3)中的点与面元素, 我们有

$$\begin{pmatrix} \pi_1^T \\ \pi_2^T \\ \pi_3^T \end{pmatrix} \mathbf{X} = 0 \quad (2.1.6)$$

如果三个面 π_1, π_2, π_3 不共线(此时称三个面处于一般位置), 则其系数矩阵的秩为 3, 此时点 \mathbf{X} 是方程(2.1.6)系数矩阵的一维(右)零空间中的元素, 因而能唯一确定(相差一个齐次因子), 即一般位置上的三个平面能唯一确定一个点(有可能是无穷远点)。

如果三个面 π_1, π_2, π_3 共一条直线 L , 则方程(2.1.6)系数矩阵的秩为 2, 此时系数矩阵有二维(右)零空间, 因而不能唯一确定点 \mathbf{X} , 实际上在直线 L 上的所有点都满足方程(2.1.6)。如果三个面 π_1, π_2, π_3 处于一般位置, 就可以由系数矩阵的 3 阶子行列式来计算点 \mathbf{X} , 其计算公式类似于公式(2.1.4)。

平面点的参数化

空间平面 π 上的点仅有两个自由度, 如果将空间平面 π 上的点 \mathbf{X} 作为一个射影平面上的点, 则

点 X 可以用三维向量 \mathbf{x} 来表示, 三维向量 \mathbf{x} 称为平面 π 上的点 X 的参数化表示。

给定平面 π 上不共线三个点的齐次坐标 X_1, X_2, X_3 , 则平面 π 上的任一点 X 可以表示成

$$X = \alpha X_1 + \beta X_2 + \gamma X_3 = (X_1, X_2, X_3) \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} \quad (2.1.7)$$

这样, 就得到了平面 π 上的点 X 的一种参数化表示 $\mathbf{x} = (\alpha, \beta, \gamma)^T$, 有时也称它为平面点的二维齐次坐标。显然, 平面点的参数化不是唯一的。

2.1.3 空间直线

在三维空间中, 直线不如点、平面那样可以非常简单地用一个四维向量(齐次坐标)来表示, 因为三维空间中的直线有 4 个自由度。下面将介绍直线的若干种表示方法。

直线的点表示

以点为基本几何元素来表示直线, 即将直线作为两个点的连线。

假定 X_1, X_2 是空间中两个不重合点, 令 W 是以这两个点的齐次坐标作为行所构成 2×4 矩阵 $W = \begin{pmatrix} X_1^T \\ X_2^T \end{pmatrix}$ 。于是, 有下述结论:

1. 点束 $L = \{X = W^T \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \mid (\alpha, \beta) \in R^2\}$ 是连结两个空间点 X_1, X_2 的一条直线 (通常简述为矩

阵 W 生成点束 L);

1. 矩阵 W 的二维右零空间是以直线 L 为轴的平面束。因为这个右零空间中的平面都通过空间点 X_1, X_2 , 所以连结这两点的直线 L 必在这些平面上。

由直线 L 上的另外两点 X'_1, X'_2 所定义的 W' 和 W 有相同的右零空间, 即它们生成同一个点束。因此, 空间直线 L 可以由它上面的两个点所构成的矩阵 W 来表示。在这种表示下, 连结两点的直线 L 也说成是直线 W 。

直线的面表示

类似地, 也可以用平面作为基本几何元素来表示直线, 即将直线定义为两个平面 π_1, π_2 的交, 它是直线的对偶表示。由不重合的面 π_1, π_2 定义一个 2×4 矩阵 W^* :

$$W^* = \begin{pmatrix} \pi_1^T \\ \pi_2^T \end{pmatrix}$$

我们有下述结论:

1. $\alpha\pi_1 + \beta\pi_2 = W^{*T} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$, $\alpha, \beta \in R$, 是以一条直线 L 为轴的平面束;
2. W^* 的二维右零空间是一条直线 L 构成的空间点束。

空间中的点 X 和直线 W 按下述方式定义了一个 3×4 矩阵:

$$M = \begin{pmatrix} W \\ X^T \end{pmatrix}$$

如果点 X 不在直线 W 上, 则 M 的右零空间是一维的, 这个零空间确定一个平面, 即没有结合性质的点与直线确定一个平面, 或者不共线的三点确定一个平面; 如果点 X 在直线 W 上, 则 M 的右零空间是二维的。

空间中的直线 W^* 与平面 π 也定义了一个 3×4 矩阵:

$$M^* = \begin{pmatrix} W^* \\ \pi^T \end{pmatrix}$$

如果直线 W^* 不在平面 π 上, 则 M^* 的右零空间是一维的, 并且这个零空间确定一个点 X , 即直线 W^* 与平面 π 的交点; 直线 W^* 在平面 π 上, 则 M^* 的右零空间是二维的。

直线的 Plucker 矩阵表示

连结两点 A, B 的直线可由下述 Plucker 矩阵 L 表示:

$$L = AB^T - BA^T \quad (2.1.8)$$

它是一个 4×4 反对称矩阵。

在平面上两点的连线可表示为 $l = x \times y$, 因此, $L = AB^T - BA^T$ 是二维平面中直线向量积的表示在三维空间中的推广。

连结 A, B 的直线是矩阵 L 的二维(右)零空间所确定的平面束的轴线。事实上, 若 $A, B \in \pi$, 即 $A^T \pi = B^T \pi = 0$, 所以 $L\pi = (B^T \pi)A - (A^T \pi)B = 0$ 。反之, 若 $L\pi = 0$, 则 $(B^T \pi)A - (A^T \pi)B = 0$, 由于 A, B 是两个不同点的齐次坐标, 所以必有 $A^T \pi = B^T \pi = 0$, 即 $A, B \in \pi$ 。因此, 连结 A, B 的直线是矩阵 L 的二维(右)零空间所确定的平面束的轴线。

直线的 Plucker 矩阵 L 表示与选取该直线上点 A, B 无关。这是因为对于直线上任何一个异于 $A,$

B 的点 C , 有 $C = A + sB$, 从而得到

$$AC^T - CA^T = A(A^T + sB^T) - (A + sB)A^T$$

从这里也可以看出空间直线有 4 个自由度。因为反对称矩阵 L 有 6 个非零元素, 但仅有 5 个比率是有意义的, 另外 L 满足约束 $\det(L)=0$, 所以直线的自由度是 4。

直线的对偶 Plucker 表示 L^* 由两个平面 P, Q 的交所确定:

$$L^* = PQ^T - QP^T \quad (2.1.9)$$

由于矩阵 L^* 与 L 在空间中表示同一条直线, 所以必有

$$L^*L = (PQ^T - QP^T)(AB^T - BA^T) = 0$$

由此, 可推知矩阵 L^* 与 L 有下述关系:

$$l_{12} : l_{13} : l_{14} : l_{23} : l_{24} : l_{34} = l_{34}^* : l_{42}^* : l_{23}^* : l_{14}^* : l_{13}^* : l_{12}^* \quad (2.1.10)$$

关系规则非常简单: 对偶和原来元素的下标总包含所有的数字 {1,2,3,4}, 如果原来元素的下标是 ij , 那么对偶元素的下标是 {1,2,3,4} 中不包含 ij 的数, 例如 $12 \mapsto 34$ 。

直线在 Plucker 矩阵的表示下, 有下述结论:

1. 如果点 X 不在直线 L 上, 则它们所确定的平面是 $\pi = L^*X$; 而 $L^*X=0$ 的充要条件是 X 在直线 L 上;
2. 直线 L 和平面 π 交点是 $X=L\pi$; 而 $L\pi=0$ 的充要条件是直线 L 在 π 上;
3. 两(或更多)条线 L_1, L_2 的性质, 可以由矩阵 $M = (L_1, L_2, \dots)$ 的零空间推出。例如: 三线 L_1, L_2, L_3 共面的充要条件是 $M^T = (L_1, L_2, L_3)^T$ 有一维零空间。

直线的 Plucker 坐标^{*)}

Plucker 直线坐标是 4×4 反对称 Plucker 矩阵 L 的六个非零元素, 即

$$L = (l_{12}, l_{13}, l_{14}, l_{23}, l_{24}, l_{34}) \quad (2.1.11)$$

它是齐次 6 维向量, 因而是 5 维射影空间中的元素。

由于 $\det L=0$, 并且不难计算:

$$\det L = (l_{12}l_{34} + l_{13}l_{42} + l_{14}l_{23})^2$$

所以, Plucker 直线坐标满足下述方程:

^{*)} 参考 10.4.3 节, 在那里使用外积给出点、线、面坐标的统一表达, 即 Plucker-Grassmann 坐标。

$$l_{12}l_{34} + l_{13}l_{42} + l_{14}l_{23} = 0 \quad (2.1.12)$$

反之, 如果向量 L 满足(2.1.12), 则它对应于三维空间中的一条直线。

假定两直线 L, \hat{L} 分别是点 A, B 和 \hat{A}, \hat{B} 的连线, 定义:

$$(L | \hat{L}) = \hat{l}_{12}l_{34} + l_{12}\hat{l}_{34} + l_{13}\hat{l}_{42} + \hat{l}_{13}l_{42} + l_{14}\hat{l}_{23} + \hat{l}_{14}l_{23}$$

不难验证:

$$(L | \hat{L}) = \text{Det}(A, B, \hat{A}, \hat{B}) \quad (2.1.13)$$

由于两直线 L, \hat{L} 相交的充要条件是 4 点 A, B, \hat{A}, \hat{B} 共面, 而 4 点共面等价于 $\text{det}(A, B, \hat{A}, \hat{B})=0$ 。于是, 我们有下述结论:

1. 两直线 L, \hat{L} 相交(即共面)的充要条件是 $(L | \hat{L}) = 0$;
2. 假定两线 L, \hat{L} 分别是平面 P, Q 和 \hat{P}, \hat{Q} 的交线, 则 $(L | \hat{L}) = \text{Det}(P, Q, \hat{P}, \hat{Q})$;
3. 如果 L 是两平面 P, Q 的交线, \hat{L} 是两点 A, B 的连线, 则有

$$(L | \hat{L}) = (P^T A)(Q^T B) - (Q^T A)(P^T B) \quad (2.1.14)$$

2.1.4 共线平面束的交比

由命题 1.1.3 知, 平面内四条共点直线所构成的线束(line pencil) $\{l_j : j=1, \dots, 4\}$ 被任一条直线所截, 得到四共线点 $\{x_j : j=1, \dots, 4\}$, 四该共线点的交比等于该线束的交比。设线束的交点为 o , $\{\hat{x}_j : j=1, \dots, 4\}$ 分别为这四条直线上不同于 o 的点, 则四共点线的交比可由 Möbius 公式计算:

$$(l_1, l_2; l_3, l_4) = \frac{\text{det}(o, x_1, x_3) \cdot \text{det}(o, x_1, x_4)}{\text{det}(o, x_2, x_3) \cdot \text{det}(o, x_2, x_4)} \quad (2.1.15)$$

令 $\{\Pi_j : j=1, \dots, 4\}$ 是四个共线的平面所构成的面束, Π_5 是第 5 个平面, 它与平面束 $\{\Pi_j : j=1, \dots, 4\}$ 相截得到四条共点直线 $\{l_j : j=1, \dots, 4\}$, 见图 2.1.1 所示。定义平面束的交比为

$$(\Pi_1, \Pi_2; \Pi_3, \Pi_4) = (l_1, l_2; l_3, l_4) \quad (2.1.16)$$

由于平面束被不同平面相截得到线束的交比均相等, 所以上述定义确实是有意义的。设 P, Q 是平面束轴线上的不同两点, 令 X_j 为第 j 个平面不在轴线上的点, $j=1, \dots, 4$, 则共线平面束的交比可由 Möbius 公式计算:

$$(\Pi_1, \Pi_2; \Pi_3, \Pi_4) = \frac{\text{det}(P, Q, X_1, X_3) \cdot \text{det}(P, Q, X_1, X_4)}{\text{det}(P, Q, X_2, X_3) \cdot \text{det}(P, Q, X_1, X_4)} \quad (2.1.17)$$

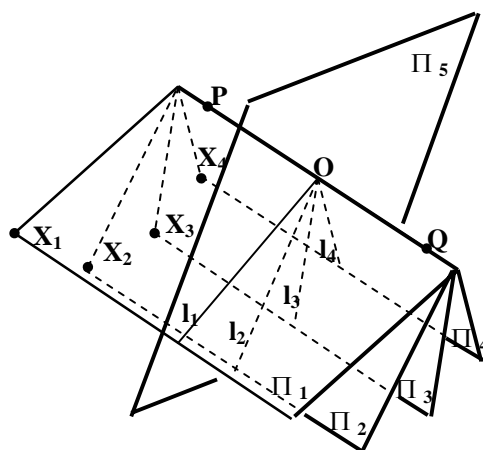


图 2.1.1 共线四面的交比

2.2 三维射影变换

2.2.1 三维射影变换

三维射影变换是射影空间上的可逆齐次线性变换，这个变换可由 4×4 的矩阵 H 来描述：

$$X' = HX \tag{2.2.1}$$

矩阵 H 称为射影变换矩阵或称为单应矩阵。由于变换是齐次的，所以射影变换矩阵 H 可以相差一个非零常数因子，因此三维射影变换有 15 个自由度，即射影变换矩阵可由它的元素所构成的 15 个比值唯一确定。

三维射影变换将空间上的点(线、面)变换到点(线、面)，并且保持点的共线(面)性、线的共面性等性质。任何三维射影变换的逆变换都是三维射影变换、任意两个三维射影变换的合成(对应于两个单应矩阵的积)也是三维射影变换。因此，三维射影变换的全体构成三维射影空间上的一个变换群，称它为三维射影变换群。

5 点确定三维射影变换

满足式(2.2.1)的点 $X \leftrightarrow X'$ 称为射影变换 H 的一个点对应。给定点对应的齐次坐标, 由于(2.2.1)是一个齐次等式, 即, 式(2.2.1)表示在相差一个常数意义下的相等, 所以有

$$sX' = HX$$

其中 s 为未知的非零齐次因子。消去上式中齐次因子 s , 可得到关于 H 的 3 个线性齐次方程。因此, 在一般情况下, 5 个点唯一确定一个三维射影变换。下述命题是更确切的陈述。

命题 2.2.1 如果 5 个点对应中任意 4 点不共面, 则它们唯一确定一个三维射影变换。

由于两个三维射影变换的合成仍是一个射影变换, 因此为了证明此命题, 只须证明它的下述特殊形式:

设 X'_1, X'_2, \dots, X'_5 为三维射影空间中任意给定的 5 个点, 其中任何 4 个点不共面, 则存在唯一的射影变换 H 将下述 5 个点:

$$X_1 = (1,0,0,0)^T, X_2 = (0,1,0,0)^T, X_3 = (0,0,1,0)^T, X_4 = (0,0,0,1)^T, X_5 = (1,1,1,1)^T$$

依次变为 X'_1, X'_2, \dots, X'_5 。

证明 因为 $s_j X'_j = HX_j$, $s_j \neq 0, j=1,2,\dots,5$, 所以有

$$(s_1 X'_1, s_2 X'_2, s_3 X'_3, s_4 X'_4) = H(X_1, X_2, X_3, X_4)$$

因此,

$$H = (X'_1, X'_2, X'_3, X'_4) \text{diag}(s_1, s_2, s_3, s_4) (X_1, X_2, X_3, X_4)^{-1} = (X'_1, X'_2, X'_3, X'_4) \text{diag}(s_1, s_2, s_3, s_4)$$

于是, 有

$$s_5 X'_5 = (X'_1, X'_2, X'_3, X'_4) \text{diag}(s_1, s_2, s_3, s_4) X_5 = (X'_1, X'_2, X'_3, X'_4) S$$

其中 $S = (s_1, s_2, s_3, s_4)^T$ 。由于 X'_1, X'_2, X'_3, X'_4 不共面, 所以 (X'_1, X'_2, X'_3, X'_4) 是可逆矩阵, 因此,

$$S = s_5 (X'_1, X'_2, X'_3, X'_4)^{-1} X'_5$$

于是,

$$H = s_5 (X'_1, X'_2, X'_3, X'_4) \text{diag}((X'_1, X'_2, X'_3, X'_4)^{-1} X'_5)$$

即 H 在相差一个常数因子的意义下有唯一解。证毕。

2.2.2 平面与直线的变换规则

变换 H 的对偶是 H^{-T} , 由于在空间中点与平面是一对互为对偶元素, 所以有下述命题:

命题 2.2.2 射影变换 H 对平面的变换规则是

$$\pi' = H^{-T} \pi \quad (2.2.2)$$

如果直线用 Plucker 矩阵表示, 则有如下变换规则:

命题 2.2.3 射影变换 H 对直线 L 的变换规则是

$$L' = HLH^T \quad (2.2.3)$$

其对偶形式是

$$L^{*'} = H^{-T} L^* H^{-1} \quad (2.2.4)$$

证明是容易的, 因为 $L' = HA(HB)^T - HB(HA)^T = H(AB^T - BA^T)H^T = HLH^T$ 。

2.3 二次曲面与变换规则

2.3.1 基本性质

二次曲面由下述方程所定义:

$$X^T Q X = 0 \quad (2.3.1)$$

其中 Q 是 4×4 的对称矩阵。如果 Q 是降秩的, 则称它为退化二次曲面, 否则称为非退化二次曲面。为了陈述方便, 通常用“二次曲面 Q ”来代替陈述“由对称矩阵 Q 所确定的二次曲面”。图 2.3.1 给出了欧氏空间中的几种常见的二次曲面。

下面是二次曲面的一些常用性质:

1. 二次曲面有 9 个自由度, 即由它的 10 个不同元素的比值所确定, 因此空间中 9 个点可确定一个二次曲面; 如果二次曲面是退化的, 则可用较少的点来确定;
2. 直线与二次曲面交于两个点(可能是重点或虚点);
3. 平面 π 与二次曲面 Q 的交是一条二次曲线;
4. 在一般情况下, 两个二次曲面的交是一条空间 4 次曲线。如果两个二次曲面都是锥面, 则它们的交线由两条二次曲线所构成。
5. 对于非退化的二次曲面 Q 上的每一点 X 都存在切平面 π , 切平面的坐标由 $\pi = QX$ 给出; 如果平面 π 是切平面, 则切点 X 的坐标由 $X = Q^{-1}\pi$ 给出。锥面 Q 在顶点处不存在切平面, 其他任何一点 X 都存在切平面 π , 切平面的坐标也由 $\pi = QX$ 。与非退化二次曲面不同的是锥面同一条母线上的点有相同的切平面, 也就是说给定锥面的切平面不能唯一确定它的切点。

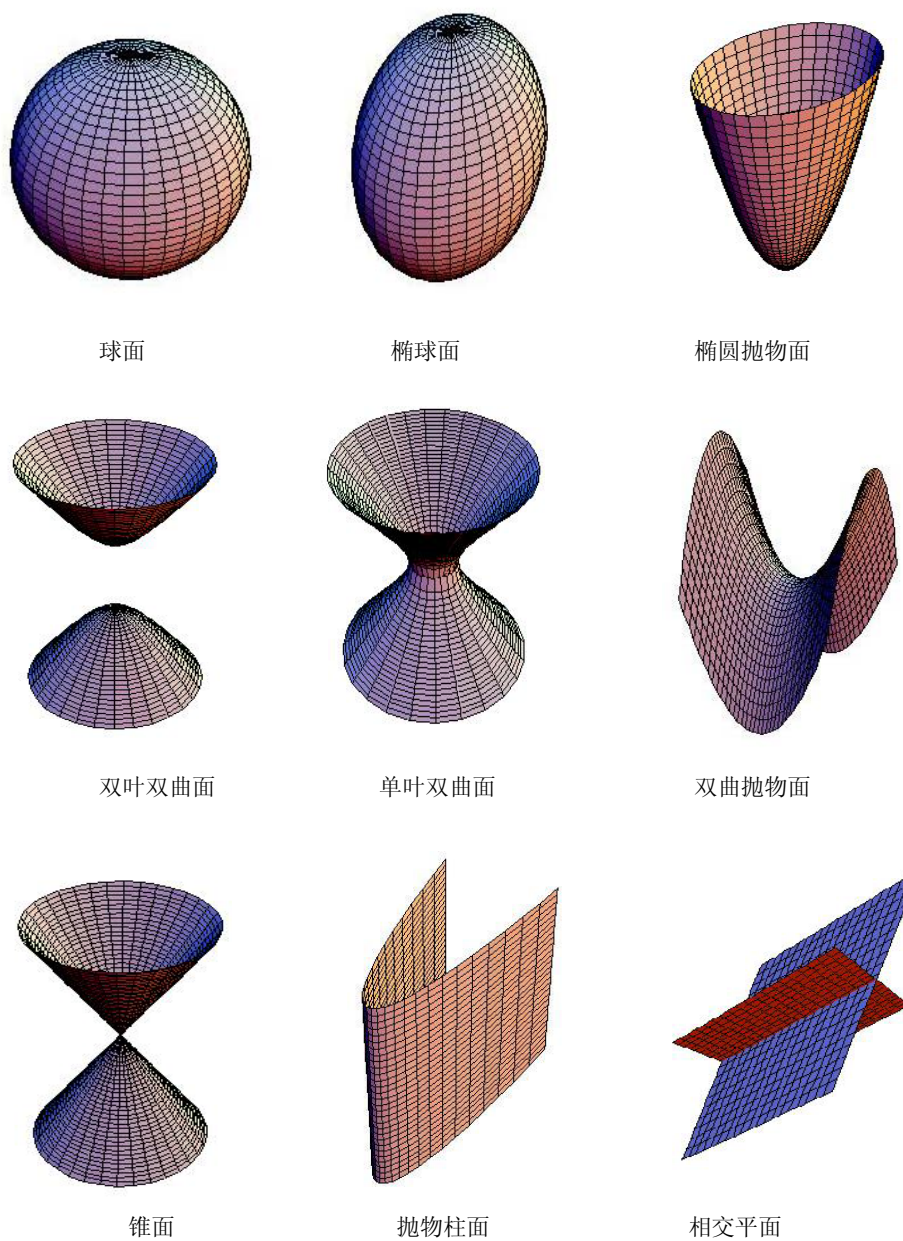


图 2.3.1: 欧氏空间中的几种常见曲面

6. 给定一个二次曲面 Q , 则 $\pi = QX$ 确定了空间点与平面的的一个对应关系, 通常称为由二

次曲面 Q 的配极对应。如果二次曲面 Q 是非退化的, 则它的配极对应是点与平面之间的一一对应。在几何上, 如果点 X 在二次曲面 Q 上, 则它的极平面是点 X 的切平面; 如果点 X 不在(非退化)二次曲面 Q 上, 则点 X 的极平面是以 X 为顶点的锥面与 Q 的切点所在的平面(如图 2.3.2 所示);

7. 在射影变换 $X' = HX$ 下, 二次曲面变换规则是

$$Q' = H^{-T} Q H^{-1} \tag{2.3.2}$$

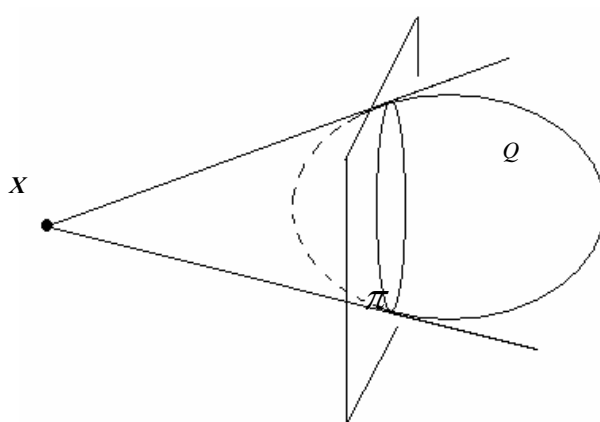


图 2.3.2: 点 X 关于二次曲面的极平面是过该点的锥与 Q 的切点所在的平面 π 。

2.3.2 二次曲面的对偶

空间曲面的对偶是指以该曲面的切平面为基本元素在对偶空间(面空间)中所构成的曲面, 通常称对偶曲面。下面着重考虑二次曲面的对偶。在一般情况下, 二次曲面的对偶仍为一个二次曲面。令 Q 是一个二次曲面, 它的对偶曲面记为 Q^* , 按照对偶曲面的定义 Q^* 的基本元素是 Q 的切平面, 也就是说它是 Q 的所有切平面所构成的平面集合, 而 Q 是 Q^* 中的所有平面所形成的包络。在计算机视觉中, 二次曲面的对偶, 尤其是锥面与空间二次曲线的对偶具有特别的重要性。

非退化二次曲面的对偶

考虑非退化二次曲面的对偶。令 Q 是一个非退化的二次曲面, 即 $\det(Q) \neq 0$, 它在(点)空间的方程为 $X^T Q X = 0$ 。根据上面的定义, 它的对偶是它的所有切平面构成的集合, 下面证明这个集合在对偶空间中也构成一个非退化的二次曲面。

任取 Q 的一个切平面 π , 切点为 X , 则必有 $\pi = QX$, 因此, $Q^{-1}\pi = X$ 。又因 X 在平面 π 上, 所以必有 $\pi^T X = 0$ 。于是, 得到 $\pi^T Q^{-1}\pi = \pi^T X = 0$ 。因此, 对 Q 的任一切平面 π , 等式 $\pi^T Q^{-1}\pi = 0$

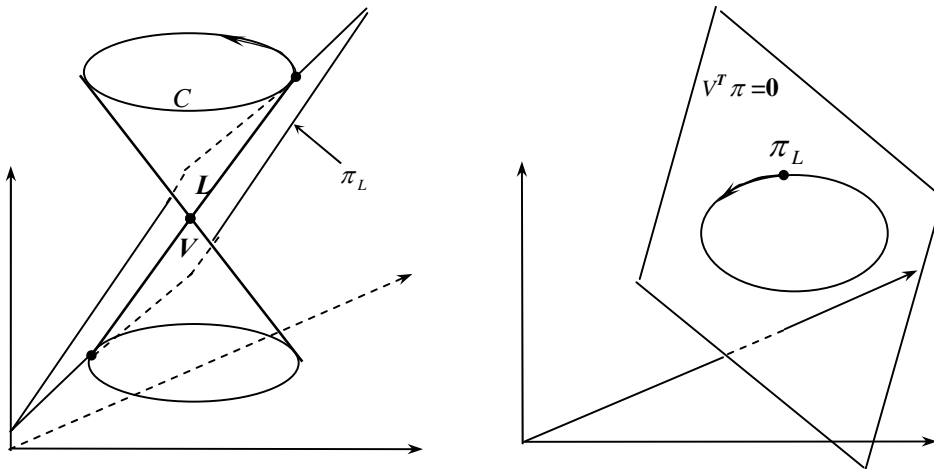
成立。反之，假定平面 π 满足方程 $\pi^T Q^{-1} \pi = 0$ ，下面证明平面 π 必为 Q 的切平面。令 $X = Q^{-1} \pi$ ，则必有 $\pi = QX$ 。为了证明 π 为 Q 的切平面，现在只须证明点 X 在二次曲面上。由于 $X^T Q X = (Q^{-1} \pi)^T \pi = \pi^T Q^{-T} \pi = \pi^T Q^{-1} \pi = 0$ ，其中倒数第二个等式利用了 Q 的对称性，因此，点 X 在二次曲面 Q 上。从上面的论证，有下述命题：

命题 2.3.1: 非退化二次曲面的对偶 Q^* 仍是二次曲面，并且 $Q^* = Q^{-1}$ 。

注：非退化二次曲面与它的对偶二次曲面互为对偶，即有 $(Q^*)^* = Q$ 。

锥面的对偶

令 Q 是一个锥面，即 $rank(Q) = 3$ ，它是一个退化二次曲面。下面考虑它的对偶 Q^* 。由于 $rank(Q) = 3$ ，所以 Q 有一维零空间，并且零空间的元素是锥面 Q 的顶点 V 的齐次坐标，即锥面 Q 的顶点 V 是方程 $QV = 0$ 的非零解。现在考虑 Q 的切平面集合在对偶空间中所构成的曲面形式。首先注意到：锥面 Q 在顶点 V 处不存在切平面。从代数上也可以看出这一点：由于顶点 V 使得 $QV = 0$ ，而四维零向量不能作为任何平面的齐次坐标，因此锥面 Q 在顶点 V 处不存在切平面。



(a) 点空间中的锥面；

(b) 锥面对偶在面空间(对偶空间)是一条二次曲线。

图 2.3.3: 锥面及其对偶。

参考图 2.3.3: 点 V 的对偶在对偶空间中表示一个“平面” V ，即在对偶空间中满足方程 $V^T \pi = 0$ 的所有“点” π 的集合。锥面母线 L 上的点，除顶点外，都有相同的切平面 π_L ，即母线 L 上所有点的对偶是同一“点” π_L ，换句话说，母线 L 在对偶空间中被压缩成一个“点” π_L 。由于在点空间

中平面 π_L 过顶点 V , 所以在对偶空间中“点” π_L 必在“平面” V 上。当母线 L 绕基线 C 运动时, π_L 在对偶空间中的轨迹将是“平面” V 上的一条“点”曲线。所以, 锥面的对偶是一条平面曲线。

下面证明这条曲线是一条二次曲线。令 X 是锥面 Q 上任一异于顶点 V 的点, 则它的切平面为 $\pi = QX$ 。显然, 一个点 Y 在母线 VX 上, 当且仅当 $QY = \pi$ 。令 $\tilde{X} = Q^+ \pi$, 则它必在母线 VX 上, 这是因为

$$Q\tilde{X} = QQ^+ \pi = QQ^+ QX = QX = \pi$$

因此, 母线的参数方程为 $\tilde{X} = V + sX$ 。于是, 我们有

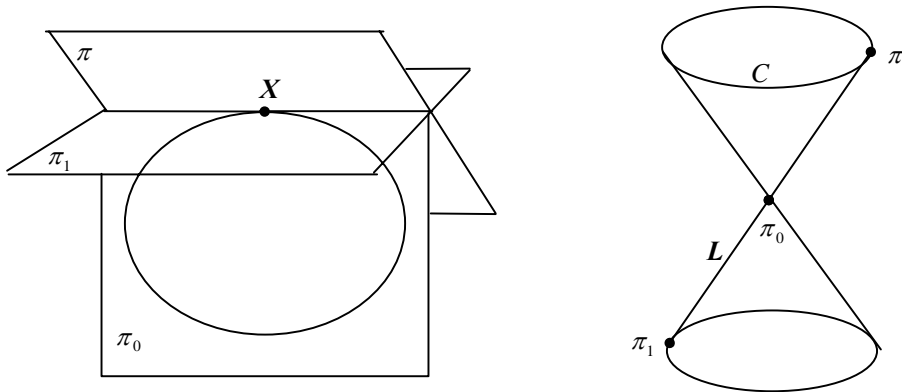
$$\pi^T Q^+ \pi = X^T Q^T (V + sX) = X^T QX = 0$$

所以, 锥面的对偶由下述方程表示: $\begin{cases} \pi^T Q^+ \pi = 0 \\ V^T \pi = 0 \end{cases}$ 。由于 $rank(Q^+) = 3$, 所以锥面的对偶(在对偶空间中)

中)是一个锥面与平面的交线, 因此它是一条平面二次曲线。

命题 2.3.2: 锥面 Q 的对偶在对偶空间中是一条二次曲线, 这条二次曲线的支撑面是锥面顶点的对偶, 锥面 Q 的母线在对偶空间中被压缩为二次曲线上的一个点。锥面 Q 的对偶可以用下述方程来描述:

$$\begin{cases} \pi^T Q^+ \pi = 0 \\ V^T \pi = 0 \end{cases} \tag{2.3.3}$$



(a) 点空间中的二次曲线

(b) 二次曲线的对偶在面空间(对偶空间)是锥面

图 2.3.4: 空间二次曲线及其对偶

空间二次曲线的对偶

参考图 2.3.4, 令平面 π_0 是空间二次曲线的支撑平面, 它的对偶在对偶空间中表示一个“点” π_0 。二次曲线上任一点 X 的切平面是以该点的切线为轴的一个平面束(但不包括支撑平面 π_0)。

令 π_1 是这个面束中的一个成员, 则这个平面束的参数方程为: $\pi = \pi_0 + s\pi_1$ 。在对偶空间中, 它表示经过“点” π_0 的一条直线 $L(s) = \pi_0 + s\pi_1$ 。当点 X 沿二次曲线运动时, L 在对偶空间中的轨迹将是形成一个锥面。

命题 2.3.3: 空间二次曲线的对偶曲面是一个锥面, 二次曲线的支撑平面的对偶是这个锥面的顶点, 二次曲线上的一个点在对偶空间中被扩展为锥面的一条母线, 二次曲线的切线与锥面的母线构成一一对应关系。

对偶二次曲面的变换规则

在(点)变换 $X' = HX$ 下, 应用平面的变换规则 $\pi' = H^{-T}\pi$, 立即得到对偶二次曲面 Q^* 的变换规则:

$$Q^{*'} = HQ^*H^T \quad (2.3.4)$$

注意: 由于锥面的对偶曲面 Q^* 是一条空间二次曲线, 它不能由一个矩阵来表示, 即它的变换规则不能统一在上述公式中, 但可以由对偶锥面和平面的变换规则来联合表达:

$$\begin{cases} \pi^T Q^+ \pi = 0 \\ V^T \pi = 0 \end{cases} \rightarrow \begin{cases} \pi'^T H Q^+ H^T \pi' = 0 \\ V^T H^T \pi' = 0 \end{cases}$$

2.3.3 绝对二次曲线与绝对二次曲面

绝对二次曲线

绝对二次曲线 Ω_∞ 是 π_∞ 上的一条(点)二次曲线。在欧氏坐标系下 $\pi_\infty = (0, 0, 0, 1)^T$, Ω_∞ 是下述方程的解集:

$$\begin{cases} x_1^2 + x_2^2 + x_3^2 = 0 \\ x_4^2 = 0 \end{cases} \quad (2.3.5)$$

它是 π_∞ 上的一条虚二次曲线。尽管 Ω_∞ 没有实点, 但它具有二次曲线的共同性质, 如: 切线、配极对应等等。

下面不加证明地引进一些常用性质:

1. 无穷远直线交绝对二次曲线于两点, 这两个点是通过该无穷远线的平面上的两个圆环点;
2. 绝对二次曲线是空间中所有平面的圆环点所构成的集合, 因而任意一个圆与绝对二次曲线相交于两个圆环点;
3. 任意一个球与无穷远平面的交是绝对二次曲线;
4. 如果绝对二次曲线在无穷远平面上的矩阵表示为 Ω_∞ (在欧氏坐标系下, Ω_∞ 是一个 3 阶单位矩阵), 则它的任一点 x_∞ 的切线为 $l_\infty = \Omega_\infty x_\infty$, 反之若 l_∞ 是 Ω_∞ 的任一条切线, 则切点为 $x_\infty = \Omega_\infty^{-1} l_\infty$;
5. 如果绝对二次曲线在无穷远平面上的矩阵表示为 Ω_∞ , 则 $l_\infty = \Omega_\infty x_\infty$ 确定了绝对二次曲线的配极对应;
6. 空间两条正交直线的方向(即两条正交直线与无穷远平面的交点)是绝对二次曲线的一对共轭点(如图 2.3.5 所示), 三正交方向(即三条两两正交直线与无穷远平面的三个交点)构成绝对二次曲线的一个自极三角形。

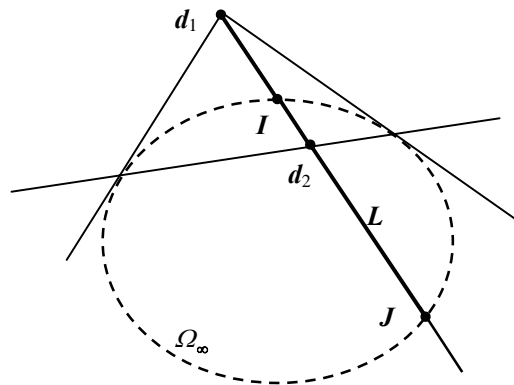


图 2.3.5: d_1, d_2 是空间中两条相互正交直线与无穷远平面的交点, 则它们的连线是一条无穷远直线 L 。直线 L 交绝对二次曲线于两个点 I, J , 它们是以 L 为无穷远直线的空间平面 π 上的两个圆环点, 在平面 π 上通过无穷远点的 d_1, d_2 两条直线必相互正交, 因此 d_1, d_2 与圆环点 I, J 调和共轭, 所以 d_2 必在 d_1 的极线上。故 d_1, d_2 是绝对二次曲线的一对共轭点。

在一般三维射影空间中, 通过绝对二次曲线 Ω_∞ 可以度量两条直线的夹角, 确切地说, 我们有下述命题。

命题 2.3.4 三维射影空间中, 令 d_1 和 d_2 是两条直线与二次曲线 Ω_∞ 所在平面 π_∞ 的交点, 它表

示这两条直线在射影空间中的方向, Ω_∞ 是绝对二次曲线在平面 π_∞ 上的矩阵表示。则两条直线交角可以通过下述公式来计算:

$$\cos \theta = \frac{(d_1^T \Omega_\infty d_2)}{\sqrt{(d_1^T \Omega_\infty d_1)(d_2^T \Omega_\infty d_2)}} \quad (2.3.6)$$

证明: 先验证欧氏空间的情形: 由于在欧氏空间中 Ω_∞ 在无穷远平面上的矩阵表示为 $\Omega_\infty = I$, 两条直线与无穷远平面的交点 $D_1 = (d_1^T, 0)^T, D_2 = (d_2^T, 0)^T$ 在无穷远平面上的表示必为 d_1 和 d_2 , 它们是两条直线的欧氏方向。由欧氏几何, 立即得到两直线的交角公式:

$$\cos \theta = \frac{(d_1^T d_2)}{\sqrt{(d_1^T d_1)(d_2^T d_2)}} = \frac{(d_1^T \Omega_\infty d_2)}{\sqrt{(d_1^T \Omega_\infty d_1)(d_2^T \Omega_\infty d_2)}}$$

因此, 公式(2.3.6)成立。下面证明一般情况:

由于通过一个射影变换 H 可以将欧氏空间变换到一般射影空间, 并且射影变换将平面映射为平面。记无穷平面 π_∞ 被映射到平面 π'_∞ , 则射影变换 H 必诱导出从平面 π_∞ 到平面 π'_∞ 上的一个二维射影变换 $\tilde{H}_{3 \times 3}$ 。于是, 根据二次曲线的变换规则, 绝对二次曲线 Ω_∞ 在平面 π'_∞ 上的矩阵表示必为 $\Omega'_\infty = \tilde{H}_{3 \times 3}^{-T} I \tilde{H}_{3 \times 3}^{-1}$, 且两直线的方向 d_1, d_2 被变换到射影方向: $d'_1 = \tilde{H}_{3 \times 3} d_1, d'_2 = \tilde{H}_{3 \times 3} d_2$ 。因此, 我们有

$$\cos \theta = \frac{d_1^T \tilde{H}_{3 \times 3}^{-T} H^{-1} d'_2}{\sqrt{d_1^T \tilde{H}_{3 \times 3}^{-T} H^{-1} d'_1} \sqrt{d_2^T \tilde{H}_{3 \times 3}^{-T} H^{-1} d'_2}} = \frac{(d_1^T \Omega'_\infty d'_2)}{\sqrt{(d_1^T \Omega'_\infty d'_1)(d_2^T \Omega'_\infty d'_2)}}$$

即公式(2.3.6)成立。证毕。

利用这个命题, 可以直接得到绝对二次曲线的基本性质(6)。

绝对二次曲面

绝对二次曲线 Ω_∞ 的对偶是三维空间中的退化对偶二次曲面, 称它为绝对二次曲面并记为 Q_∞^* 。在几何上, Q_∞^* 是所有与 Ω_∞ 相切的平面所构成的集合。在代数上, Q_∞^* 可由秩 3 的 4×4 的齐次矩阵来表示, 在欧氏坐标系下它的表示为

$$Q_\infty^* = \begin{pmatrix} I & 0 \\ 0^T & 0 \end{pmatrix} \quad (2.3.7)$$

平面 π 在 Q_∞^* 上的充要条件是 $\pi^T Q_\infty^* \pi = 0$ 。令平面的坐标为 $\pi = (v^T, k)^T$, 给定 Q_∞^* 形式(2.3.7)后, $\pi^T Q_\infty^* \pi = 0$ 等价于 $v^T v = 0$ 。而 v 表示平面 $\pi = (v^T, k)^T$ 与无穷远平面的交线, 该直线与绝对二次曲线相切的充要条件是 $v^T I v = 0$ 。因此, Q_∞^* 正好由这些与绝对二次曲线相切的平面所组成。

通过式(2.3.7), 可直接验证无穷远平面是 Q_∞^* 的右零空间, 即 $Q_\infty^* \pi_\infty = 0$ 。

绝对二次曲面在射影坐标系下有 8 个自由度, 因为它是退化的对偶二次曲面。这 8 个自由度也是在一般射影坐标系下确定度量性质所要确定的自由度。

命题 2.3.5 在三维射影空间中, 若绝对二次曲面的矩阵表示为 Q_∞^* , 则两平面 π_1 和 π_2 之间的夹角由下式给出:

$$\cos \theta = \frac{\pi_1^T Q_\infty^* \pi_2}{\sqrt{(\pi_1^T Q_\infty^* \pi_1)(\pi_2^T Q_\infty^* \pi_2)}} \quad (2.3.8)$$

特别地, 在欧氏空间中, 若两平面的坐标 $\pi_1 = (\mathbf{n}_1^T, d_1)^T$, $\pi_2 = (\mathbf{n}_2^T, d_2)^T$, 则两平面的夹角计算公式简化为:

$$\cos \theta = \frac{\mathbf{n}_1^T \mathbf{n}_2}{\sqrt{(\mathbf{n}_1^T \mathbf{n}_1)(\mathbf{n}_2^T \mathbf{n}_2)}}$$

证明是容易的, 欧氏空间的夹角公式直接由欧氏几何得到, 对于一般射影空间可由射影变换关于对偶二次曲面的变换规则和平面的变换规则得到。

关于绝对二次曲线和绝对二次曲面的一些更深刻的度量性质将在第 2.4 节给出。

代数表示的几何解释

由于无穷远平面的特殊性, 所以绝对二次曲线和绝对二次曲面也有很多特殊性质。在三维计算机视觉中, 尤其在摄像机自标定与三维重构理论中, 绝对二次曲线和绝对二次曲面处于十分重要的地位。为了理解它们在欧氏坐标下的表示, 下面从另一个角度来考虑。

首先, 考虑绝对二次曲线。令 Q_r 是中心在原点半径 r 为的球面, 则它的矩阵表示为

$$Q_r = \text{diag}(1, 1, 1, -r^2)$$

即球面 Q_r 上的点 $X = (x, y, z, w)^T$ 满足方程:

$$\left(\frac{x}{r}\right)^2 + \left(\frac{y}{r}\right)^2 + \left(\frac{z}{r}\right)^2 = w^2$$

当 r 逐渐增大时, 球面 Q_r 上的点就逐渐接近与无穷远平面。记 Q_∞ 是球面 Q_r 在 $r \rightarrow \infty$ 时的极限。当 $r \rightarrow \infty$ 时, 对任意 x, y, z , 有 $\frac{x}{r} \rightarrow 0, \frac{y}{r} \rightarrow 0, \frac{z}{r} \rightarrow 0$ 。由上式, 必有 $w \rightarrow 0$, 因此, Q_∞ 上的点

$X = (x, y, z, w)^T$ 必是无穷远点, 且满足方程:

$$\begin{cases} x^2 + y^2 + z^2 = 0 \\ w = 0 \end{cases}$$

这正是绝对二次曲线在欧氏坐标系下的方程, 所以可以将绝对二次曲线作为球面 Q_r 在 $r \rightarrow \infty$ 时的极限。

再考虑对偶二次曲面 Q_∞^* 。由于绝对二次曲面 Q_∞^* 是绝对二次曲线的对偶, 所以从上面的讨论可以将它看作作为球面 Q_r 对偶 Q_r^* 的极限。由于 Q_r 的对偶可以表示成

$$Q_r^* = \text{diag}(1, 1, 1, -1/r^2)$$

当 $r \rightarrow \infty$ 时, 必有

$$Q_r^* = \text{diag}(1, 1, 1, -1/r^2) \rightarrow \text{diag}(1, 1, 1, 0) = Q_\infty^*$$

这样, 就得到了绝对二次曲面在欧氏坐标系下表示式(2.3.7)。从这里也可以看出, 我们为什么将绝对二次曲面记成 Q_∞^* 而不是按习惯那样记成 Ω_∞^* 。

2.4 三维射影变换群的子群

2.4.1 仿射变换群

三维仿射变换是

$$X' = H_a X = \begin{pmatrix} A & t \\ 0 & 1 \end{pmatrix} X \quad (2.4.1)$$

其中 A 是一个 3 阶可逆矩阵。三维仿射变换有 12 个自由度。

所有三维仿射变换的全体构成一个变换群, 称为仿射群。也就是说, 仿射变换的逆变换也是仿射变换, 两个仿射变换的合成也是仿射变换。

仿射不变量

对于仿射不变量, 我们有下述结论:

1. 保持无穷远平面不变, 即将无穷远点变换到无穷远点;
2. 保持直线与直线、直线与平面以及平面与平面之间的平行性;
3. 保持物体的体积比、平行图形(或在同一平面上的图形)的面积比、平行线段(或在同一直线

上的线段)的长度比不变。

性质(1)与(2)是明显的。下面证明: 仿射变换保持体积比的不变性质。假定 V_1, V_2 是两个空间物体, 其体积分别为 $v(V_1), v(V_2)$, 经过仿射变换(2.4.1)后的物体记为 V'_1, V'_2 , 其体积分别为 $v(V'_1), v(V'_2)$ 。将仿射变换(2.4.1)写成非齐次形式:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = A \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}$$

则仿射变换的 Jacobi 行列式为:

$$J = \det \left(\frac{\partial(x', y', z')}{\partial(x, y, z)} \right) = \det(A)$$

所以,

$$v(V'_j) = \iiint_{V'_j} dx' dy' dz' = \iiint_{V_j} \left| \frac{\partial(x', y', z')}{\partial(x, y, z)} \right| dx dy dz = |\det(A)| \iiint_{V_j} dx dy dz = |\det(A)| v(V_j)$$

因此, 我们有

$$\frac{v(V_1)}{v(V_2)} = \frac{v(V'_1)}{v(V'_2)}$$

平行图形面积比的不变性质可由体积比不变性质导出, 而平行线段长度比不变性质可由面积比不变性质导出。请读者自己证明。

命题 2.4.1 射影变换 H 保持无穷远平面不变的充要条件是 H 为仿射变换。

证明 仿射变换保持无穷远平面不变, 所以仅需证明必要性。为此, 将射影变换 H 写成分块形式:

$$H = \begin{pmatrix} A & \mathbf{b} \\ \mathbf{c}^T & d \end{pmatrix}$$

假定 H 保持无穷远平面 $\pi_\infty = (0^T, 1)^T$ 不变。由平面的射影变换规则, 我们有

$$\pi_\infty = \begin{pmatrix} A & \mathbf{b} \\ \mathbf{c}^T & d \end{pmatrix}^{-T} \pi_\infty$$

即,

$$\begin{pmatrix} A^T & \mathbf{c} \\ \mathbf{b}^T & d \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

因此, $c = 0, d = 1$, 所以 $H = \begin{pmatrix} A & b \\ 0^T & 1 \end{pmatrix}$ 是一个仿射变换。证毕。

命题 2.4.1 表明, 保持无穷远平面不变是仿射变换的基本特征。

2.4.2 相似变换群

相似变换由下述变换所定义

$$X' = \begin{pmatrix} sU & t \\ 0^T & 1 \end{pmatrix} X \quad (2.4.2)$$

其中, U 是三维正交矩阵, s 是相似比例因子。所有三维相似变换的全体构成一个群, 通常称为相似变换群, 它是三维仿射群的子群。如果限制 U 是一个三维旋转矩阵, 则上述变换称为旋转相似变换。旋转相似变换的全体构成相似变换群的子群。

相似不变量

相似变换除了仿射不变量作为它的不变量之外, 最本质的不变性质是绝对二次曲线与绝对二次曲面。

命题 2.4.2 射影变换 H 保持绝对二次曲线不变的充要条件是 H 为相似变换^{*)}。

证明 相似变换必为仿射变换, 所以 H 可以写成下述形式:

$$H = \begin{pmatrix} A & t \\ 0 & 1 \end{pmatrix}$$

仿射变换将无穷远点变为无穷远点, 因此 H 限制在无穷远平面上的二维变换是 A 。而绝对二次曲线 Ω_∞ 是无穷远平面上的二次曲线, 在无穷远平面上它的矩阵表示是一个 3 阶单位矩阵 I , 因此 H 将 Ω_∞ 变为 Ω_∞ 的充要条件是 $A^{-T}IA^{-1} = sI$, 而 $A^{-T}IA^{-1} = sI$ 等价于 $sA^T A = I$, $sA^T A = I$ 等价于 A 是一个与正交矩阵相差常数倍的矩阵。所以, 射影变换 H 保持绝对二次曲线不变的充要条件是 H 为相似变换。证毕。

下面的命题是命题 2.4.2 的对偶命题:

命题 2.4.3 射影变换 H 保持绝对二次曲面 Q_∞^* 不变的充要条件是 H 为相似变换。

^{*)} 这里的保持绝对二次曲线不变, 是指整体不变性, 而不是说二次曲线上的每一点都保持不变。

可以给出命题 2.4.3 的直接证明: 因为 Q_{∞}^* 是对偶二次曲面, 所以它的变换规则为式(2.3.4)。于

是, Q_{∞}^* 在变换 H 下不变的充要条件是 $Q_{\infty}^* = HQ_{\infty}^*H^T$ 。令 $H = \begin{pmatrix} A & \mathbf{t} \\ \mathbf{v}^T & k \end{pmatrix}$, 下述齐次等式:

$$\begin{pmatrix} I & 0 \\ 0^T & 0 \end{pmatrix} = \begin{pmatrix} A & \mathbf{t} \\ \mathbf{v}^T & k \end{pmatrix} \begin{pmatrix} I & 0 \\ 0^T & 0 \end{pmatrix} \begin{pmatrix} A^T & \mathbf{v} \\ \mathbf{t}^T & k \end{pmatrix} = \begin{pmatrix} AA^T & A\mathbf{v} \\ \mathbf{v}^T A^T & \mathbf{v}^T \mathbf{v} \end{pmatrix}$$

成立的充要条件是 $\mathbf{v}=0$ 且 A 与正交矩阵相差一个非零常数因子, 从而 H 是一个相似变换。

2.4.3 等距变换群

等距变换由下式所定义:

$$\mathbf{X}' = \begin{pmatrix} U & \mathbf{t} \\ 0^T & 1 \end{pmatrix} \mathbf{X} \quad (2.4.3)$$

其中 U 是三维正交矩阵。所有三维等距变换的全体构成一个群, 通常称为等距变换群, 它是三维射影群的子群。如果限制 U 是一个三维旋转矩阵, 则上述变换称为欧氏变换。欧氏变换的全体构成等距变换群的子群。等距变换群的重要不变量是保持物体形状和体积不变。等距变换是特殊的相似变换, 它具有相似变换的一切特性, 如保持绝对二次曲线和绝对二次曲面不变。

下面给出关于欧氏变换的不动点性质。一个射影变换的不动点是指在这个变换下保持不动的空间点。在代数上, 一个空间点 \mathbf{X} 是射影变换 H 的不动点的充要条件是 \mathbf{X} 为 H 的特征向量, 即 $H\mathbf{X} = \lambda\mathbf{X}$ (注意: 这是一个齐次等式, 齐次因子是与特征向量 \mathbf{X} 对应的特征值)。

命题 2.4.4 (1) 设 E 是一个欧氏变换, 则正交于旋转轴的平面上的两个圆环点是 E 的两个不动点, 它们是 E 的两个互为共轭复特征向量; 旋转轴与无穷远平面的交点是 E 的另一个不动点, 它是 E 的特征值 1 的特征向量。如果 E 为一般欧氏运动, E 仅有上述三个不动点。(2) 如果平移向量在与旋转轴正交的平面上(通常称为平面运动), 则 E 还存在另外的不动点, 它是 E 的特征值 1 的另一个特征向量。

在几何上, 空间旋转变换 R 有两个不变子空间, 一个是旋转轴构成 R 的一维不变子空间, 其上的每一点在旋转变换 R 下是不动的; 另一个是与旋转轴正交且通过坐标原点的平面, 它构成 R 的一个二维不变子空间(R 在这个不变子空间上不是点点不动的), 由于 R 限制在这个平面上是一个二维旋转变换, 所以这个平面上的两个圆环点是保持不动的。空间的平移变换是一个保持平行性的变换, 并且在无穷平面上保持点点不动。由于欧氏变换 E 是先进行旋转 R 再作平移所构成的变换, 所以旋

转轴上的无穷远点是它的一个不动点, 与旋转轴正交平面上的两个圆环点也是它的不动点。如果平移向量在与旋转轴正交的平面上, 除了这三个不动点外还有另一个的不动点, 这个不动点是由平移向量所确定的。

下面将给出这个命题的代数证明:

证明 令欧氏变换为

$$E = \begin{pmatrix} R & t \\ \theta^T & 1 \end{pmatrix}$$

考虑欧氏变换 E 的特征向量。由于 $\det(sI_{4 \times 4} - E) = (s-1)\det(sI_{3 \times 3} - R)$, 所以 E 的特征值必为 $\{e^{i\theta}, e^{-i\theta}, 1, 1\}$, 其中 θ 是 R 的旋转角。显然, E 有如下三个线性无关的特征向量:

$$E_1 = (i_R, 0)^T, E_2 = (\bar{i}_R, 0)^T, E_3 = (a_R^T, 0)^T$$

其中: i_R, \bar{i}_R 是 R 的共轭复特征值的特征向量, 所以 E_1, E_2 在与旋转轴正交的平面 π_E 上, 并且 π_E 是 R 的不变子空间, 因此 E_1, E_2 是平面 π_E 上的两个无穷远点。由于 E 在 π_E 上的限制 E_{π_E} 是一个二维欧氏变换, 平面 π_E 上的两个圆环点在变换 E_{π_E} 下保持不变, 于是 E_1, E_2 必是平面 π_E 上的两个圆环点。 a_R 是 R 的特征值 1 的特征向量, 它是 R 的旋转轴方向, 所以 E_3 是旋转轴与无穷远平面的交点。

如果 E 对应于特征值 1 还有另外的特征向量 E_4 , 则 E_4 必有形式 $E_4 = (x^T, 1)^T$ 。于是从

$$\begin{pmatrix} R & t \\ \theta^T & 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ 1 \end{pmatrix}$$

可知 $(I - R)x = t$, 所以 $a_R^T t = a_R^T (I - R)x = a_R^T x - (R^{-1} a_R)^T x = a_R^T x - a_R^T x = 0$, 这表示平移向量 t 在与旋转轴正交的平面 π_E 上(注意: 这里使用了三维欧氏坐标)。反之, 若平移向量 t 在与旋转轴正交的平面 π_E 上, 则 E 必有形如 E_4 的特征向量。证毕。

2.3.5 二次曲面的分类

度量分类

在三维欧氏空间中, 二次曲面的一般方程为

$$\varphi(x) = (x^T, 1) Q \begin{pmatrix} x \\ 1 \end{pmatrix} = x^T B x + 2c^T x + l$$

其中:

$$Q = \begin{pmatrix} a & d & e & g \\ d & b & f & h \\ e & f & c & k \\ g & h & k & l \end{pmatrix}, B = \begin{pmatrix} a & d & e \\ d & b & f \\ e & f & c \end{pmatrix}, \mathbf{c} = (g, h, k)^T, \mathbf{x} = (x, y, z)^T$$

由于 B 是一个对称矩阵, 所以存在旋转变换 $\mathbf{y} = \mathbf{R}\mathbf{x}$, 使 $\mathbf{x}^T B \mathbf{x}$ 化为标准形

$$\mathbf{y}^T \text{diag}(d_1, d_2, d_3) \mathbf{y}$$

于是, 二次曲面 $\varphi(\mathbf{x})$ 化为

$$\varphi(\mathbf{y}) = \mathbf{y}^T \text{diag}(d_1, d_2, d_3) \mathbf{y} + 2\mathbf{c}'^T \mathbf{y} + l$$

欧氏变换不改变矩阵的秩, 通过 Q, B 的秩、行列式以及 B 的特征值的符号, 可得到二次曲面的度量分类, 见表 2.3.1。注意, 度量分类是二次曲面的一种形状分类, 它不是欧氏变换的等价性质, 即同一种类型的二次曲面不能通过欧氏变换相互转化。

表 2.3.1: 二次曲面的度量分类

rank B	rank Q	曲面	有附加条件的细分类
3	4	有唯一中心的曲面	椭球面, d_1, d_2, d_3 都有和 $\det A / \det B$ 异号
			虚椭球面, d_1, d_2, d_3 都有和 $\det A / \det B$ 同号
			单叶双曲面, d_1, d_2, d_3 中仅有一个和 $\det A / \det B$ 同号
			双叶双曲面, d_1, d_2, d_3 中正好有两个和 $\det A / \det B$ 同号
3	3	锥面	退化锥面, d_1, d_2, d_3 同号
			非退化锥面, d_1, d_2, d_3 异号
2	4	抛物面	椭圆抛物面, d_1, d_2 同号
			双曲抛物面, d_1, d_2 异号
2	3	中心为一条直线的柱面	双曲柱面, d_1, d_2 异号
			椭圆柱面(实或虚), d_1, d_2 同号
2	2	相交二平面	实的, d_1, d_2 异号
			虚的, d_1, d_2 同号
1	4	不可能出现	
1	3	抛物柱面	
1	2	平行两平面(实或虚)	
1	1	二重合平面	

仿射分类

在仿射空间(斜坐标系)内, 二次曲面一般方程为

$$\varphi(\mathbf{x}) = \mathbf{x}^T B \mathbf{x} + 2\mathbf{b}^T \mathbf{x} + c = 0, \text{ 其中 } B = (a_{ij})_{3 \times 3}, \mathbf{x} = (x_1, x_2, x_3)^T$$

其中 B 是 3 阶对称矩阵。

存在满秩变换(仿射变换) $\mathbf{y} = P^{-1}\mathbf{x}$, 使得二次曲面方程变为

$$\varphi(\mathbf{y}) = \mathbf{y}^T \text{diag}(d_1, d_2, d_3)\mathbf{y} + 2\mathbf{b}'^T \mathbf{y} + c' = 0$$

其中 $d_j = 0, \pm 1$, 但不全为零。再对它进行平移与伸缩变换(它们都是仿射变换), 上述方程可变为下述 17 个标准形之一, 见表 2.3.2。

表 2.3.2: 二次曲面的仿射分类

序号	方程	曲面
1	$x_1^2 + x_2^2 + x_3^2 + 1 = 0$	虚椭球面(虚曲面)
2	$x_1^2 + x_2^2 + x_3^2 - 1 = 0$	椭球面
3	$x_1^2 + x_2^2 - x_3^2 + 1 = 0$	双叶双曲面
4	$x_1^2 + x_2^2 - x_3^2 - 1 = 0$	单叶双曲面
5	$x_1^2 + x_2^2 + x_3^2 = 0$	仅有一个实点的虚锥面
6	$x_1^2 + x_2^2 - x_3^2 = 0$	二次锥面
7	$x_1^2 + x_2^2 - x_3 = 0$	椭圆抛物面
8	$x_1^2 - x_2^2 + x_3 = 0$	双曲抛物面
9	$x_1^2 + x_2^2 + 1 = 0$	虚柱面
10	$x_1^2 + x_2^2 - 1 = 0$	椭圆柱面
11	$x_1^2 - x_2^2 + 1 = 0$	双曲柱面
12	$x_1^2 + x_2 = 0$	抛物柱面
13	$x_1^2 + x_2^2 = 0$	交一实直线的二个虚面
14	$x_1^2 - x_2^2 = 0$	相交的二个平面
15	$x_1^2 + 1 = 0$	一对平行虚面
16	$x_1^2 - 1 = 0$	一对平行实面
17	$x_1^2 = 0$	二个重合的平面

射影分类

由于二次曲面的矩阵 Q 是对称的, 所以它可以分解为 $Q=UDU^T$, 这里 U 是一个实正交矩阵而 D 是一个实对角矩阵。通过对 U 的四列各自进行适当的伸缩, 可以将 Q 分解成 $Q=H^{-T}DH^{-1}$, 这里 D 是对角元素取 0, 1, 或 -1 的对角矩阵, 并且使 D 的零对角元素出现在对角线的最后, $+1$ 出现在最前面, 而 -1 次之 (如表 2.3.3 所给的那种形式)。显然, H 是一个射影变换, 因此二次曲面 Q 通过射影变换 H 必(射影)等价于二次曲面 D 。因此, 对角矩阵 D 的每一种形式代表了二次曲面的一种射影等价类, 见表 2.3.3。

表 2.3.3: 二次曲面的射影分类(注: 符号差是指 D 的对角元素中 1 的个数与 -1 个数的差)

Q 的秩	符号差	方程	曲面
4	4	$x_1^2 + x_2^2 + x_3^2 + x_4^2 = 0$	虚椭球面(虚曲面)
4	2	$x_1^2 + x_2^2 + x_3^2 - x_4^2 = 0$	椭球面或双叶双曲面或椭圆抛物面
4	0	$x_1^2 + x_2^2 - x_3^2 - x_4^2 = 0$	单叶双曲面或双叶抛物面
3	3	$x_1^2 + x_2^2 + x_3^2 = 0$	虚锥面
3	1	$x_1^2 + x_2^2 - x_3^2 = 0$	实锥面或柱面
2	2	$x_1^2 + x_2^2 = 0$	一对虚平面(虚曲面)
2	0	$x_1^2 - x_2^2 = 0$	一对实平面
1	1	$x_1^2 = 0$	二个重合的平面

2.5 射影坐标系与射影坐标变换

在前面各节中, 我们所使用的坐标系都是欧氏坐标系。本节将讨论一般射影坐标系以及射影坐标系之间的变换(射影坐标变换)。

射影坐标系

这里不打算用纯几何的方法来建立射影坐标系, 而是从给定的欧氏齐次坐标系 σ 来建立一般射影坐标系。

设 A, B, C, D 是空间中的 4 个不共面的点, 它们在欧氏坐标系 σ 下的齐次坐标分别为 $X_1^e, X_2^e, X_3^e, X_4^e$, 则对于空间任一点 P 的欧氏齐次坐标均可以表示成

$$X^e = u_1 X_1^e + u_2 X_2^e + u_3 X_3^e + u_4 X_4^e \quad (2.5.1)$$

其中 u_j 不全为零。这样, 任一点 \mathbf{P} 有一个分量不全为零的 4 维有序数组 $\{u_j\}$ 与之对应。但是, $\{u_j\}$ 还不能作为 \mathbf{P} 点的新齐次坐标, 因为 u_j 的比值 $u_1:u_2:u_3:u_4$ 不能唯一确定, 例如对欧氏坐标系下的齐次坐标 \mathbf{X}_j^e 选择不同的齐次因子 s_j , 则 u_j 将变成 u_j/s_j , 且

$$u_1:u_2:u_3:u_4 \neq (u_1/s_1):(u_2/s_2):(u_3/s_3):(u_4/s_4)$$

为了确定比值, 必须再加约束条件。设空间中的第五个点 \mathbf{E} , 它在欧氏坐标系 σ 下的齐次坐标记为 \mathbf{X}_5^e , 并且它与原来 4 点中的任何 3 个点都不共面, 因此存在 4 个都不全为零的数 v_1, v_2, v_3, v_4 , 使得

$$\mathbf{X}_5^e = v_1\mathbf{X}_1^e + v_2\mathbf{X}_2^e + v_3\mathbf{X}_3^e + v_4\mathbf{X}_4^e$$

令 $\hat{\mathbf{X}}_j^e = v_j\mathbf{X}_j^e$, 则式(2.5.1)可以写成

$$\mathbf{X}^e = x_1\hat{\mathbf{X}}_1^e + x_2\hat{\mathbf{X}}_2^e + x_3\hat{\mathbf{X}}_3^e + x_4\hat{\mathbf{X}}_4^e, (x_j = u_j/v_j) \tag{2.5.2}$$

这样对于 $\{x_j\}$ 中的 4 个元素就有确定的比值, 即不依赖于 $\mathbf{X}_j^e (j=1,2,3,4)$ 的齐次因子 s_j 的选择。因为对任意的 $\{s_j\}$, 总有

$$\begin{aligned} x_1:x_2:x_3:x_4 &= (u_1/v_1):(u_2/v_2):(u_3/v_3):(u_4/v_4) \\ &= (u_1s_1/v_1s_1):(u_2s_2/v_2s_2):(u_3s_3/v_3s_3):(u_4s_4/v_4s_4) \end{aligned}$$

于是, 对于每一个空间点 \mathbf{P} 都有一个新的齐次坐标 $(x_1, x_2, x_3, x_4)^T$ 。特别地, $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{E}$ 的新齐次坐标分别为

$$\mathbf{X}_1^p = (1,0,0,0)^T, \mathbf{X}_2^p = (0,1,0,0)^T, \mathbf{X}_3^p = (0,0,1,0)^T, \mathbf{X}_4^p = (0,0,0,1)^T, \mathbf{X}_5^p = (1,1,1,1)^T$$

这样建立起来的坐标系称为射影坐标系, 并称 $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ 所构成的 4 面形为射影坐标系的 4 面形; \mathbf{E} 称为单位点; $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{E}$ 称为射影坐标系的基点(如图 2.5.1 所示)。不难看出以欧氏坐标(或仿射坐标)为基础的齐次坐标系是一种特殊的射影坐标系, 其坐标 4 面形的顶点是三个坐标轴的无穷远点和坐标原点, 而单位点是非齐次坐标为 $(1,1,1)^T$ 的空间点。

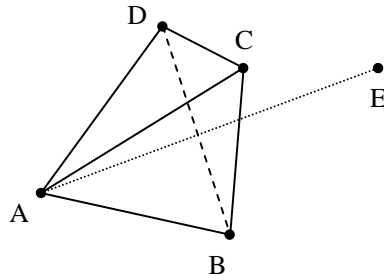


图 2.5.1: 射影坐标系的 4 面形 $\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$ 与单位点 \mathbf{E}

值得说明的是以下几点:

1. 在以一般射影坐标系为基础的三维射影空间中, 无穷远点、无穷远直线与无穷远平面在以欧氏坐标系(或仿射坐标系)为基础的代数形式都消失了, 即不再有表示它们的那种(欧氏坐标系中的)特殊代数形式, 所有的几何元素的地位都是同等的。
2. 可以在一般射影坐标系下讨论平面的齐次坐标, 二次曲面, 射影变换等等。
3. 在以一般射影坐标系为基础的三维射影空间中的射影变换不再有层次之分, 如相似变换、仿射变换、欧氏变换等等。因此, 上节中关于射影变换的分层以及特殊变换的特征, 都只是在以欧氏坐标系为基础的射影空间中才成立, 正如我们不能在仿射坐标系下讨论欧氏变换一样。

射影坐标变换

同一个空间点在两个不同射影坐标系中的射影坐标之间的变换, 是很容易获得到。令 σ_X, σ_Y 是两个一般射影坐标系, σ_e 是以欧氏坐标为基础的特殊射影坐标系, 点 \mathbf{P} 在这三个坐标系下的坐标分别为 $\mathbf{X}, \mathbf{Y}, \mathbf{X}^e$, 根据式(2.5.2), $\sigma_X \rightarrow \sigma_e, \sigma_Y \rightarrow \sigma_e$ 的射影坐标变换分别为

$$\mathbf{X}^e = x_1 \hat{\mathbf{X}}_1^e + x_2 \hat{\mathbf{X}}_2^e + x_3 \hat{\mathbf{X}}_3^e + x_4 \hat{\mathbf{X}}_4^e = (\hat{\mathbf{X}}_1^e, \hat{\mathbf{X}}_2^e, \hat{\mathbf{X}}_3^e, \hat{\mathbf{X}}_4^e) \mathbf{X}$$

$$\mathbf{X}^e = y_1 \hat{\mathbf{Y}}_1^e + y_2 \hat{\mathbf{Y}}_2^e + y_3 \hat{\mathbf{Y}}_3^e + y_4 \hat{\mathbf{Y}}_4^e = (\hat{\mathbf{Y}}_1^e, \hat{\mathbf{Y}}_2^e, \hat{\mathbf{Y}}_3^e, \hat{\mathbf{Y}}_4^e) \mathbf{Y}$$

因此 $\mathbf{Y} = (\hat{\mathbf{Y}}_1^e, \hat{\mathbf{Y}}_2^e, \hat{\mathbf{Y}}_3^e, \hat{\mathbf{Y}}_4^e)^{-1} (\hat{\mathbf{X}}_1^e, \hat{\mathbf{X}}_2^e, \hat{\mathbf{X}}_3^e, \hat{\mathbf{X}}_4^e) \mathbf{X}$, 这样就证明了 $\sigma_X \rightarrow \sigma_Y$ 的射影坐标变换是一个可逆的齐次线性变换。给定空间 5 个点, 其中任意 4 个点不共面, 如果已知它们在 σ_X, σ_Y 的坐标为 $\mathbf{X}_j, \mathbf{Y}_j$, 则由命题 2.2.1 就可以唯一确定这个可逆的齐次线性变换。

可以看出射影坐标变换也具有射影变换的形式。射影坐标变换具有射影变换的形式, 不是偶然的, 因为可以给射影变换以两种解释。第一种解释: 射影变换是同一个坐标系中空间点之间的变换, 坐标系没有发生变化, 图形发生变化, 不但位置发生了变化, 连整个图形的形状也发生了变化。当然, 变换前、后的图形在同一个坐标系下的代数形式也发生了变化。另一解释: 射影变换是不同坐标系之间的变换, 图形不发生变化, 而是坐标系发生变化, 这种变化使得同一个图形具有不同的代数形式。为区别前一种解释, 后一种解释的射影变换通常称为射影坐标变换。这两种对射影变换的解释没有本质上的差异, 只是观察的角度不同而已。前一种是立足于坐标系观察变换(运动), 后一种是立足于变换(运动)观察坐标系。

3. 摄像机几何

三维计算机视觉的主要任务是利用三维物体的二维图像所包含的信息，获取三维物体的空间位置与形状等几何信息，并在此基础上识别三维物体。图像上每一点的亮度与物体某个表面点的反射光的强度有关，而图像点在图像平面上的位置仅与摄像机与空间物体的相对方位和摄像机的内部结构有关，摄像机的内部结构是由摄像机的内部参数所决定的。为了描述摄像机的几何成像关系，需要对摄像机进行数学建模。本章所介绍的摄像机模型是计算机视觉中广泛使用的针孔模型，通常也称为线性模型。这种模型在数学上是三维空间到二维平面的中心投影，由一个 3×4 矩阵来描述，可以说这种模型是一个(退化的)射影变换，因此通常又称它为射影摄像机。在本章，利用前两章的射影几何知识，给出摄像机关于空间点、直线、平面、二次曲线和二次曲面的投影性质，以及图像平面点、直线与二次曲线的反投影性质。这些投影与反投影性质，是从图像恢复物体三维几何结构的基础，尤其是绝对二次曲线与绝对二次曲面的投影性质。从本章可以看出，摄像机关于空间平面的投影是平面到平面的一个二维中心投影变换，因此可以使用第一章所介绍的二维射影变换的知识从平面景物图像恢复它的几何结构。对于空间物体，由于摄像机将三维物体表面投影到二维平面上，是一个(退化的)射影变换，因此不可能从三维物体的单幅图像恢复其三维结构。能否从多幅图像恢复物体的三维结构？这是三维计算机视觉中三维重构问题，将在第 6 章讨论。

3.1 摄像机模型

3.1.1 摄像机模型

基本模型

摄像机的基本成像模型，通常称为基本针孔模型，由三维空间到平面的中心投影变换所给出。令空间点 O_c 是投影中心，它到平面 π 的距离为 f 。空间点 X_c 在平面 π 上的投影(或像) m 是以点 O_c 为端点并经过点 X_c 的射线与平面 π 的交点，如图 3.1.1(a)所示。平面 π 称为摄像机的像平面，点 O_c 称为摄像机中心(或光心)， f 称为摄像机的焦距，以点 O_c 为端点且垂直于像平面的射线称为光轴或主轴，主轴与像平面的交点 p 称为摄像机的主点。

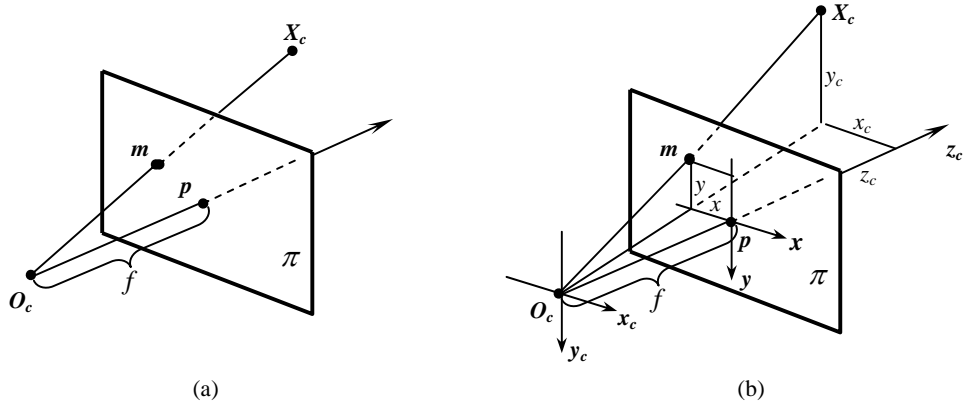


图 3.1.1:基本针孔模型

为了从代数上描述这种投影关系，需要建立摄像机(欧氏)坐标系和图像平面(欧氏)坐标系。在图像平面上，以主点 p 为像平面坐标系的坐标原点 o ，以水平线与铅直线分别为 x 轴和 y 轴，建立图像坐标系 $o-xy$ 。在空间中，以摄像机中心 O_c 为摄像机坐标系的坐标原点，以主轴为 z_c 轴，以平行于 x 轴且通过摄像机中心 O_c 的直线为 x_c 轴，以平行于 y 轴且通过摄像机中心 O_c 的直线为轴 y_c ，建立摄像机坐标系 $O_c - x_c y_c z_c$ ，如图 3.1.1(b)所示。空间点 X_c 在摄像机坐标系中的欧氏坐标记为 $\tilde{X}_c = (x_c, y_c, z_c)^T$ ，它的像点 m 在图像坐标系中的坐标记为 $\tilde{m} = (x, y)^T$ 。根据三角形相似原理，可推知空间点 X_c 与它的像点 m 满足下述关系：

$$\begin{cases} x = \frac{fx_c}{z_c} \\ y = \frac{fy_c}{z_c} \end{cases} \quad (3.1.1)$$

上式可表述为下述矩阵形式：

$$z_c \mathbf{m} = \begin{pmatrix} fx_c \\ fy_c \\ z_c \end{pmatrix} = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \mathbf{X}_c \quad (3.1.2)$$

其中 $\mathbf{X}_c = (x_c, y_c, z_c, 1)^T$ ， $\mathbf{m} = (x, y, 1)^T$ 分别为空间点和图像点的齐次坐标。它是从空间到像平面的一个齐次线性变换。如果记

$$P = \text{diag}(f, f, 1)(I, \mathbf{0}) \quad (3.1.3)$$

则这个齐次线性变换可表示为更简洁的形式：

$$\mathbf{m} = P\mathbf{X}_c \quad (3.1.4)$$

其中, 矩阵 P 是一个 3×4 矩阵, 通常称它为摄像机矩阵。这就是基本成像模型的代数表示。注意: (3.1.4) 是一个齐次等式, 表示在相差一个非零常数因子的意义下相等。

主点偏离图像中心

在实际应用中, 由于事先不知道主点的确切位置, 通常都是以图像中心或者图像的左上角作为图像坐标系的原点来建立图像坐标系的(在本书中除特别说明外都是以图像中心作为图像坐标系的原点), 在此坐标系下, 由于主点可能不是图像坐标系的原点, 因此摄像机矩阵不再可能具有(3.1.3)的形式。

若主点在上述坐标系下的坐标为 $\mathbf{p} = (x_0, y_0, 1)^T$, 则摄像机的投影关系变为:

$$z_c \mathbf{m} = \begin{pmatrix} f & 0 & x_0 & 0 \\ 0 & f & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \mathbf{X}_c = P\mathbf{X}_c \quad (3.1.5)$$

摄像机矩阵的形式为 $P = K(I, \theta)$, 其中

$$K = \begin{pmatrix} f & 0 & x_0 \\ 0 & f & y_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (3.1.6)$$

并称它为摄像机内参数矩阵。

CCD 摄像机

用于计算机处理的图像通常都是用 CCD 摄像机所获取的数字图像, 或者是由模拟信号摄像机获取的图像再经过特别设备进行数字离散化的数字图像。一般地, CCD 摄像机内参数矩阵不具有(3.1.6)的形式。为了得到 CCD 摄像机的模型, 必须刻划 CCD 摄像机的数字离散化过程。

假定 CCD 摄像机数字离散化后的像素是一个矩形, 矩形的长与宽分别为 d_x, d_y 。设图像点 $(x, y, 1)^T$ 在离散化后的坐标为 $(u, v, 1)^T$, 则必有

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} 1/d_x & 0 & 0 \\ 0 & 1/d_y & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

如果离散化后的图像坐标仍用 \mathbf{m} 表示, 则摄像机的投影关系可以写成:

$$\mathbf{m} = K(I, 0)\mathbf{X}_c = P\mathbf{X}_c \quad (3.1.8)$$

其中

$$K = \begin{pmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (3.1.9)$$

矩阵 K 称为 CCD 摄像机的内参数矩阵, $f_x = f/d_x, f_y = f/d_y$ 称为 CCD 摄像机在 u 轴和 v 轴方向上的尺度因子, $(u_0, v_0)^T = (x_0/d_x, y_0/d_y)^T$ 称为 CCD 摄像机的主点。

由于制造工艺的限制, 一般情况下, CCD 摄像机数字离散化后的像素不是一个矩形而是一个平行四边形, 四边形的一边平行于 u 轴, 而另一边与 u 轴形成一个 θ 角。令平行四边形两边长分别为 d_x, d_y , 图像点 $(x, y, 1)^T$ 离散化后的坐标为 $(u, v, 1)^T$, 则必有

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} 1/d_x & -ctg\theta/d_x & 0 \\ 0 & \sin\theta/d_y & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

结合(3.1.5)式, 可得到

$$z_c \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} f/d_x & -fctg\theta/d_x & (x_0 - y_0ctg\theta)/d_x & 0 \\ 0 & fsin\theta/d_y & y_0sin\theta/d_y & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} X_c$$

如果离散化后的图像坐标仍用 m 表示, 则上式可写成:

$$m = K(I, 0)X_c = PX_c \quad (3.1.9)$$

其中:

$$K = \begin{pmatrix} f_x & s & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (3.1.10)$$

是摄像机内参数矩阵, $f_x = f/d_x, f_y = fsin\theta/d_y$ 称为一般 CCD 摄像机在 u 轴和 v 轴方向上的尺度因子, $(u_0, v_0)^T = ((x_0 - y_0ctg\theta)/d_x, y_0sin\theta/d_y)^T$ 称为一般 CCD 摄像机的主点, 而 $s = -fctg\theta/d_x$ 称为一般 CCD 摄像机的畸变因子或倾斜因子。

摄像机矩阵的一般形式

上面所介绍的摄像机矩阵是在摄像机坐标系下的结果。由于摄像机的中心和主轴等事先都是未知的, 这个坐标系不能给出空间点的具体坐标值, 另外摄像机可安放在环境中的任何位置, 所以需要有一个基准坐标系来描述空间点和摄像机的位置。这个基准坐标系通常称为世界坐标系。

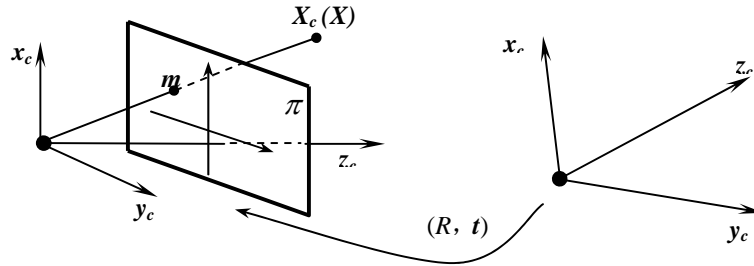


图 3.1.2：世界坐标系与摄像机坐标系之间的欧氏变换

世界坐标系与摄像机坐标系之间的关系可以用一个旋转矩阵和一个平移向量来描述，如图 3.1.2 所示。令空间点在世界坐标系与摄像机坐标系的坐标分别为 $\mathbf{X} = (x, y, z, 1)^T$ ， $\mathbf{X}_c = (x_c, y_c, z_c, 1)^T$ ，则它们之间的关系为

$$\mathbf{X}_c = \begin{pmatrix} R & -R\tilde{C} \\ \mathbf{0}^T & 1 \end{pmatrix} \mathbf{X} \quad (3.1.11)$$

其中 \tilde{C} 表示摄像机中心在世界坐标系中的非齐次坐标，即摄像机中心的齐次坐标为 $\mathbf{C} = (\tilde{C}^T, 1)^T$ 。以后，点 \mathbf{X} 的非齐次坐标总用 $\tilde{\mathbf{X}}$ 来表示。将(3.1.11)代入式(3.1.9)，则有

$$\mathbf{m} = K(I, 0) \begin{pmatrix} R & -R\tilde{C} \\ \mathbf{0}^T & 1 \end{pmatrix} \mathbf{X} = KR(I, -\tilde{C})\mathbf{X} \quad (3.1.12)$$

这样，就得到了摄像机矩阵的一般形式：

$$P = KR(I, -\tilde{C}) \quad (3.1.13)$$

矩阵 $R(I, -\tilde{C})$ 称为摄像机的外参数矩阵。有时，也用 $\tilde{\mathbf{X}}_c = R\tilde{\mathbf{X}} + \mathbf{t}$ 来描述世界坐标系与摄像机坐标系之间的关系，此时摄像机矩阵为

$$P = K(R, \mathbf{t}) \quad (3.1.14)$$

其中， $\mathbf{t} = -R\tilde{C}$ 。

摄像机矩阵是一个秩 3 的 3×4 矩阵，因为它的前三列所构成的子矩阵是一个可逆矩阵。另外，由于摄像机矩阵的齐次性，所以它仅有 11 个独立元素。

3.1.2 摄像机矩阵元素的几何意义

摄像机中心

考虑摄像机中心在世界坐标系中的坐标。由下式：

$$PC = KR(I, -\tilde{C}) \begin{pmatrix} \tilde{C} \\ 1 \end{pmatrix} = KR(\tilde{C}, -\tilde{C}) = 0$$

可知摄像机中心 $C = (\tilde{C}^T, 1)^T$ 是方程 $PC=0$ 的一个解。另一方面， P 仅有一维右零空间，因为它的秩等于 3。于是，摄像机中心的齐次坐标构成 P 的右零空间。在已知摄像机矩阵 P 的情况下，可以通过求解方程 $PX=0$ 得到摄像机中心在世界坐标系中的坐标。事实上，如果令 $P = (H, p_4)$ ，其中 H 为 P 的前三列所构成的 3×3 矩阵， p_4 是 P 的第四个列向量，则从方程 $PX=0$ 可得到摄像机中心在世界坐标系中的齐次坐标：

$$C = \begin{pmatrix} -H^{-1}p_4 \\ 1 \end{pmatrix} \quad (3.1.15)$$

坐标原点与坐标轴方向

记摄像机矩阵为 $P = (p_1, p_2, p_3, p_4)$ ，其中 p_j 为 P 的第 j 列向量。

世界坐标系的原点坐标为 $X = (0, 0, 0, 1)^T$ ，所以它的图像点坐标为

$$s_0 m_0 = PX = (p_1, p_2, p_3, p_4) \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = p_4$$

即摄像机矩阵的第 4 列向量是世界坐标原点图像的齐次坐标。

考虑世界坐标系三个坐标轴方向的图像，即三个坐标轴与无穷远平面交点的图像。显然，三个坐标轴与无穷远平面交点分别为

$$X = (1, 0, 0, 0)^T, Y = (0, 1, 0, 0)^T, Z = (0, 0, 1, 0)^T$$

所以，它们的图像分别为

$$s_1 m_1 = PX = p_1, \quad s_2 m_2 = PY = p_2, \quad s_3 m_3 = PZ = p_3$$

因此，摄像机矩阵的前三个列向量分别是世界坐标系三个坐标轴方向的图像点的齐次坐标。

主平面与轴平面

记摄像机矩阵为 $P = \begin{pmatrix} p^{1T} \\ p^{2T} \\ p^{3T} \end{pmatrix}$ ，其中 p^{1T}, p^{2T}, p^{3T} 分别为 P 的三个行向量。

主平面 摄像机的坐标平面 $O_c - x_c y_c$ ，即与像平面平行的坐标平面，通常称为主平面。主平面在

世界坐标系中可用摄像机矩阵的第 3 行向量 \mathbf{p}^{3T} 来表示。因为主平面与像平面平行, 所以它们的交线是一条无穷远直线, 即主平面的像是像平面上的无穷远直线。令 \mathbf{X} 为主平面上的任一点, 则它在摄像机下的图像必为

$$\begin{pmatrix} u \\ v \\ 0 \end{pmatrix} = P\mathbf{X} = \begin{pmatrix} \mathbf{p}^{1T} \mathbf{X} \\ \mathbf{p}^{2T} \mathbf{X} \\ \mathbf{p}^{3T} \mathbf{X} \end{pmatrix}$$

所以 $\mathbf{p}^{3T} \mathbf{X} = 0$, 于是主平面在世界坐标系中的坐标为摄像机矩阵的第 3 行 \mathbf{p}^{3T} 。

轴平面 考虑由方程 $\mathbf{p}^{1T} \mathbf{X} = 0$ 所确定的平面, 即在世界坐标系中坐标为 \mathbf{p}^{1T} 的平面。它是图像平面的 v 轴与摄像机中心所确定的平面, 通常称它为轴平面。令 \mathbf{X} 为这个轴平面上的任一点, 则它在摄像机下的图像必为

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = P\mathbf{X} = \begin{pmatrix} \mathbf{p}^{1T} \mathbf{X} \\ \mathbf{p}^{2T} \mathbf{X} \\ \mathbf{p}^{3T} \mathbf{X} \end{pmatrix} = \begin{pmatrix} 0 \\ \mathbf{p}^{2T} \mathbf{X} \\ \mathbf{p}^{3T} \mathbf{X} \end{pmatrix}$$

因此, 这个轴平面的图像点集合是 $\{(0, v, 1)^T\}$, 即像平面上的 v 轴。

同理, 由 $\mathbf{p}^{2T} \mathbf{X} = 0$ 所确定的另一个轴平面是图像平面的 u 轴与摄像机中心所构成的平面。轴平面与主平面的不同之处在于它依赖于图像坐标系的选择, 即对于不同的图像坐标系, 对应的轴平面是不同的。

主轴与主点

主轴 主轴与主平面是正交的, 因此主轴必为主平面的法线。一般地, 一个平面 $\pi = (\pi_1, \pi_2, \pi_3, \pi_4)^T$ 的法线有两个方向(正向与负向), 它们是 $\hat{\pi} = \pm(\pi_1, \pi_2, \pi_3)^T$, 在无穷远平面上它们表示同一个点, 即法线与无穷远平面的交点。由于主平面是 $\mathbf{p}^{3T} = (p_{31}, p_{32}, p_{33}, p_{34})^T$, 所以主轴两个方向为 $\hat{\mathbf{p}}^3 = \pm(p_{31}, p_{32}, p_{33})^T$ 。通常所讲的主轴方向是它的正方向, 即指向摄像机前方的方向, 由于摄像机矩阵可以相差一个常数因子, 所以 $\hat{\mathbf{p}}^3 = \pm(p_{31}, p_{32}, p_{33})^T$ 中的正号并不代表主轴的正向。如果摄像机矩阵 $P = (H, \mathbf{p}_4)$ 与标准摄像机矩阵 $K(R, \mathbf{t})$ 相差一个正常数, 必有 $\det(H) > 0$, 否则 $\det(H) < 0$ 。因此, 主轴的正向是

$$\mathbf{v} = \det(H) \mathbf{h}^3 \quad (3.1.16)$$

其中 \mathbf{h}^{3T} 是矩阵 H 的第三行向量。

主点 主点是主轴与像平面的交点, 由于主轴过摄像机中心, 因此主点必为主轴方向的图像点, 所以主点坐标为

$$\mathbf{p} = P(\mathbf{h}^{3T}, 0)^T = H\mathbf{h}^3 \quad (3.1.17)$$

3.1.3 摄像机矩阵估计

在经典立体视觉中，需要在欧氏坐标系下估计摄像机矩阵，才能够完成立体视觉系统的标定。常用的方法是根据一些空间点在欧氏坐标系下的坐标与其图像坐标之间的对应关系，建立摄像机矩阵的约束方程，从而确定摄像机矩阵。在实践中，为了得到一些空间点的欧氏坐标，需要制作一个标定参考物，在标定参考物上经过精确测定的特征点作为估计摄像机矩阵时所需要的空间点，如图 3.1.3 所示。

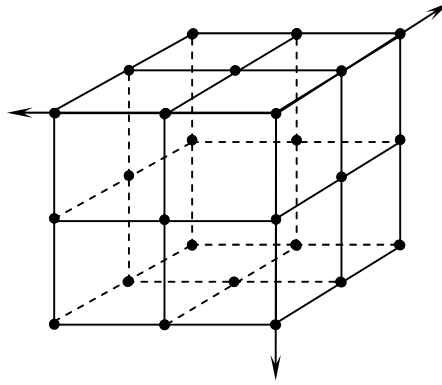


图 3.1.3：用于求解摄像机矩阵的立方体：正交的两条棱作为世界坐标系的两个坐标轴，各面上直线的交点的三维欧氏坐标已精确测定，根据这些特征点与其图像点的对应可以求解摄像机投影矩阵。

记摄像机矩阵为

$$P = \begin{pmatrix} \mathbf{p}^{1T} \\ \mathbf{p}^{2T} \\ \mathbf{p}^{3T} \end{pmatrix}$$

其中 \mathbf{p}^{jT} 为矩阵 P 的第 j 行向量。令 $\mathbf{X}_j = (x_j, y_j, z_j, 1)^T$ 是特征点在世界坐标系下的坐标，对应的图像点坐标为 $\mathbf{m}_j = (u_j, v_j, 1)^T$ ，于是根据摄像机的投影关系，得到

$$s_j \mathbf{m}_j = P \mathbf{X}_j = \begin{pmatrix} \mathbf{p}^{1T} \mathbf{X}_j \\ \mathbf{p}^{2T} \mathbf{X}_j \\ \mathbf{p}^{3T} \mathbf{X}_j \end{pmatrix} \quad (3.1.18)$$

因此，消去上式中的常数因子后，可得到下述方程：

$$\begin{cases} \mathbf{p}^{2T} \mathbf{X}_j - v_j \mathbf{p}^{3T} \mathbf{X}_j = 0 \\ \mathbf{p}^{1T} \mathbf{X}_j - u_j \mathbf{p}^{3T} \mathbf{X}_j = 0 \\ v_j \mathbf{p}^{1T} \mathbf{X}_j - u_j \mathbf{p}^{2T} \mathbf{X}_j = 0 \end{cases} \quad (3.1.19)$$

在这个方程组中, 第三个方程可由前两个方程线性表示, 因此只有两个方程是线性独立的。因此, 给定 $N \geq 6$ 个以上的特征点与其图像点的对应, 可线性求解摄像机矩阵 P 。

最小二乘解

当图像数据存在测量误差时, 方程组(3.1.19)一般不存在非零解。此时, 通常以它的最小二乘解作为摄像机矩阵的估计。对每一个点对应, 记

$$A_j = \begin{pmatrix} 0 & \mathbf{X}_j^T & -v_j \mathbf{X}_j^T \\ \mathbf{X}_j^T & 0 & -u_j \mathbf{X}_j^T \\ v_j \mathbf{X}_j^T & -u_j \mathbf{X}_j^T & 0 \end{pmatrix} \quad (3.1.20)$$

它是一个 2×12 的矩阵 A_j 。给定 n 个点对应, 得到 n 个形如这样的矩阵, 再将这 n 个矩阵组合起来得到一个 $2n \times 12$ 的矩阵 $A = (A_1^T, \dots, A_n^T)^T$, 对 A 作奇异值分解 $A = UDV^T$, 则 V 的最后一个列向量 $\mathbf{p} = \mathbf{v}_{12}$ 是方程 $A\mathbf{p} = \mathbf{0}$ 的最小二乘解^{*)}, 再将 \mathbf{p} 写成矩阵的形式就得到摄像机矩阵 P 。

3.1.4 欧氏空间与射影空间

如果世界坐标系是一般射影坐标系, 摄像机矩阵又具有什么样的形式呢? 这是不难回答的。因为摄像机坐标系和射影坐标系之间的变换仍然可以用一个 4×4 的齐次矩阵 H 表示, 即 $\mathbf{X}_c = H\mathbf{X}$, 并且由三维射影空间到图像平面的映射仍然可以用一个秩为 3 的 3×4 矩阵 P 表示。事实上, 对于最一般的情形, 摄像机模型可以被看作是从三维射影空间到二维射影平面的映射, 并且这个映射能够用下述矩阵合成的方式来表达:

$$P = K \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} H \quad (3.1.21)$$

如果 $H = \begin{pmatrix} A & \mathbf{b} \\ \mathbf{0}^T & 1 \end{pmatrix}$, 其中 $\text{rank}(A)=3$, 它表示三维空间的一个仿射变换, 式(3.1.21)变为

^{*)} 见 8.4 节

$$P=K \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} A & \mathbf{b} \\ 0^T & 1 \end{pmatrix} \quad (3.1.22)$$

它是世界坐标系为仿射坐标系的摄像机矩阵, 称它为仿射空间中的摄像机矩阵。

(2) 如果 $H = \begin{pmatrix} sR & \mathbf{t} \\ 0^T & 1 \end{pmatrix}$, 其中 R 是旋转矩阵, s 为非零常数, 它表示三维空间的一个相似变换,

此时式(3.1.21)变为

$$P=K \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} sR & \mathbf{t} \\ 0^T & 1 \end{pmatrix} \quad (3.1.23)$$

它是世界坐标系为欧氏坐标系(但度量单位为 s)的摄像机矩阵, 称它为相似空间中的摄像机矩阵。

(3) 如果 $H = \begin{pmatrix} R & \mathbf{t} \\ 0^T & 1 \end{pmatrix}$, 其中 R 是旋转矩阵, 它表示三维空间的一个欧氏变换, 此时式(3.1.21)

变为

$$P=K \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} R & \mathbf{t} \\ 0^T & 1 \end{pmatrix} \quad (3.1.24)$$

它是世界坐标系为欧氏坐标系(其度量是绝对度量)的摄像机矩阵, 称它为欧氏空间中的摄像机矩阵, 即在以前各节所介绍的摄像机矩阵。

(4) 如果 $H=(I, \mathbf{0})$, 此时式(3.1.21)变成

$$P=K \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (3.1.25)$$

它是以摄像机坐标系为世界坐标系的摄像机矩阵。

3.2 投影与反投影

3.2.1 空间点

正向投影 空间点 X 通过摄像机 P 被作用到图像平面的图像点 $m = PX$ ，这种投影关系称为摄像机的正向投影，简称为投影。在正向投影中，无穷远点的投影是非常重要的，这是因为从无穷远点的投影可以恢复景物的仿射结构。由于无穷远点的齐次坐标为

$$X_{\infty} = (d^T, 0)^T$$

其中 d 是三维向量，它表示通过点 X_{∞} 的直线方向，所以在摄像机矩阵 $P = (H, p_4)$ 的作用下，无穷远点在像平面上的投影为

$$PX_{\infty} = (H, p_4)X_{\infty} = Hd \quad (3.2.1)$$

因此，无穷远点的投影仅与摄像机矩阵的前三列有关，而与第4列无关。

反向投影 反向投影是针对图像平面的基本几何元素而言的，图像平面点 m 的反投影是指在摄像机 P 的作用下具有像点 m 的所有空间点的集合，即

$$l_b = \{X \mid m = PX\}$$

在几何上，不难看出图像点 m 的反投影是从摄像机中心出发并通过图像点 m 的一条射线。这时我们将图像平面视为三维空间中的一张平面，即将图像点看作三维空间中的点，如图 3.2.1 所示。

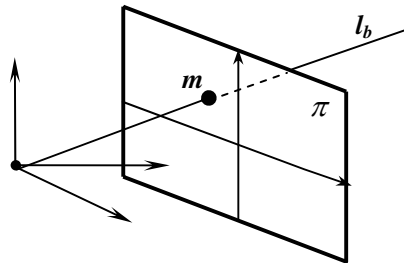


图 3.2.1：图像点 m 的反投影是射线 l_b

下面考虑图像点 m 的反投影 l_b 在世界坐标系中的方程。这是非常重要的，因为三维计算机视觉中将从多幅图像的对应点的反投影射线恢复空间点的三维坐标。由于两点确定一条直线，所以，如果能确定摄像机中心 C 和射线 l_b 上另一点的空间坐标，就可以得到射线 l_b 在空间中的方程。由式 (3.1.15)，我们知道摄像机中心 C 在世界坐标系的坐标为

$$C = \begin{pmatrix} -H^{-1}p_4 \\ 1 \end{pmatrix}$$

再考虑由图像点 m 和 P 的广义逆 $P^+ = P^T(PP^T)^{-1}$ 所定义的另一空间点 P^+m , 该空间点必在射线 l_b 上, 因为 $P(P^+m) = PP^T(PP^T)^{-1}m = m$ 。于是, 根据 l_b 上的两个点 C 和 P^+m 就得到 l_b 的下述参数方程:

$$X(u) = u(P^+m) + C \quad (3.2.2)$$

另一种方法是通过射线 l_b 上的无穷远点来确定它的参数方程。令 $X_\infty = (d^T, 0)^T$ 是这条射线上的无穷远点, 则 $PX_\infty = m$ 。因此, 根据式(3.2.1), $d = H^{-1}m$ 。于是, 射线 l_b 的参数方程为:

$$X(u) = u \begin{pmatrix} H^{-1}m \\ 0 \end{pmatrix} + \begin{pmatrix} -H^{-1}p_4 \\ 1 \end{pmatrix} = \begin{pmatrix} H^{-1}(um - p_4) \\ 1 \end{pmatrix} \quad (3.2.3)$$

3.2.2 空间直线

正向投影 由于空间直线 L 可从它上面的两个点唯一确定, 因此 L 的参数方程可以通过它上面的两个点的坐标 X_1, X_2 来表达:

$$X(u) = X_1 + uX_2$$

它在摄像机矩阵 P 的作用下, 必有

$$m(u) = P(X_1 + uX_2) = PX_1 + uPX_2 = m_1 + um_2 \quad (3.2.4)$$

其中 m_1, m_2 分别是 X_1, X_2 的图像点。所以, 直线 L 的图像是连结这两个像点 m_1, m_2 的直线。因此, 空间直线在摄像机的作用下的是像平面上的直线。

利用空间直线的 Plucker 矩阵和 Plucker 坐标也可以表达空间直线的投影, 确切地说, 我们有下述结论:

命题 3.2.1 如果空间直线的 Plucker 矩阵为 L , 即 $L = X_1X_2^T - X_2X_1^T$, 其中 X_1, X_2 是 L 上的两个点, 则它的像直线 l 的坐标满足:

$$[l]_x = PLP^T \quad (3.2.5)$$

证明 令 m_1, m_2 分别是 X_1, X_2 的图像点, 即 $PX_1 = m_1, PX_2 = m_2$ 。由于 $l = m_1 \times m_2$, 所以,

$$[l]_x = [m_1 \times m_2]_x = m_1m_2^T - m_2m_1^T = PLP^T$$

证毕。

对于图像直线 l 的坐标, 也可以象空间点的投影那样由直线投影矩阵给出。由摄像机矩阵 P , 我们定义一个秩 3 的 3×6 矩阵:

$$P_L = \begin{pmatrix} P^2 \wedge P^3 \\ P^3 \wedge P^1 \\ P^1 \wedge P^2 \end{pmatrix} \quad (3.2.6)$$

其中 p^{jT} 为摄像机矩阵 P 的第 j 行向量， $P^i \wedge P^j$ 是两平面 $\{P^i, P^j\}$ 交线的 Plucker 坐标。由上一节的讨论我们知道 P^1, P^2 是摄像机的两个轴平面， P^3 是摄像机的主平面。

命题 3.2.2 如果空间直线使用 Plucker 坐标 L 表示，则它在摄像机 P 下的像直线 l 是：

$$l = P_L L = \begin{pmatrix} (P^2 \wedge P^3 / L) \\ (P^3 \wedge P^1 / L) \\ (P^1 \wedge P^2 / L) \end{pmatrix} \quad (3.2.7)$$

证明 记 X_1, X_2 是 L 上的两个点，其像点 $PX_1 = m_1, PX_2 = m_2$ ，则

$$l = m_1 \times m_2 = (PX_1) \times (PX_2) = \begin{pmatrix} (p^{2T} X_1)(p^{3T} X_2) - (p^{2T} X_2)(p^{3T} X_1) \\ (p^{3T} X_1)(p^{1T} X_2) - (p^{3T} X_2)(p^{1T} X_1) \\ (p^{1T} X_1)(p^{2T} X_2) - (p^{1T} X_2)(p^{2T} X_1) \end{pmatrix}$$

而，

$$(P^i \wedge P^j / L) = (P^{iT} X_1)(P^{jT} X_2) - (P^{iT} X_2)(P^{jT} X_1)$$

所以，

$$l = m_1 \times m_2 = \begin{pmatrix} (P^2 \wedge P^3 / L) \\ (P^3 \wedge P^1 / L) \\ (P^1 \wedge P^2 / L) \end{pmatrix} = P_L L$$

证毕。

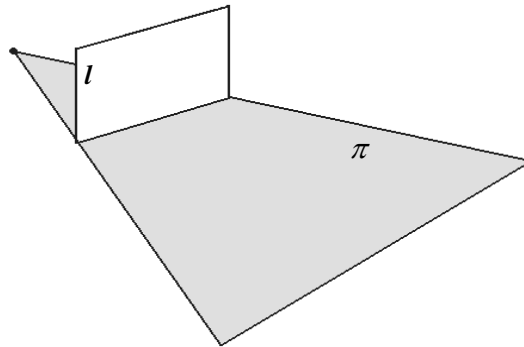


图 3.2.2: 图像直线 l 的反投影平面是 $\pi = P^T l$ 。

反向投影 在几何上，像平面上一条直线的反投影是空间中通过摄像机中心的一张平面，如图 3.2.2 所示。下面的命题是这一几何事实的代数描述。

命题 3.2.3 在摄像机 P 下的图像直线 l 的反投影是空间平面 $\pi = P^T l$ 。

证明 令 X 是摄像机 P 将它投影到直线 l 上的任一空间点，则必有 $X^T (P^T l) = l^T (PX) = 0$ 。因此，像直线 l 的反投影是空间平面 $\pi = P^T l$ 。证毕。

3.2.3 空间平面

考虑摄像机关于空间平面 π 的投影。空间点 X 在摄像机 P 的作用下，其图像点是 $m = PX$ ，由于图像点 m 在图像平面的坐标仅依赖于图像坐标系的选择而与世界坐标系的选择无关。因此，可以自由地选择世界坐标系。以空间平面 π 为世界坐标系的 O - xy 平面，如图 3.2.3 所示，则平面 π 上点 X 的坐标为 $X = (x, y, 0, 1)^T$ ，因此，

$$m = PX = (p_1, p_2, p_3, p_4) \begin{pmatrix} x \\ y \\ 0 \\ 1 \end{pmatrix} = (p_1, p_2, p_4) \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

记 $H = (p_1, p_2, p_4)$ ， $X_\pi = (x, y, 1)^T$ ，则上式可简写成

$$m = HX_\pi \tag{3.2.8}$$

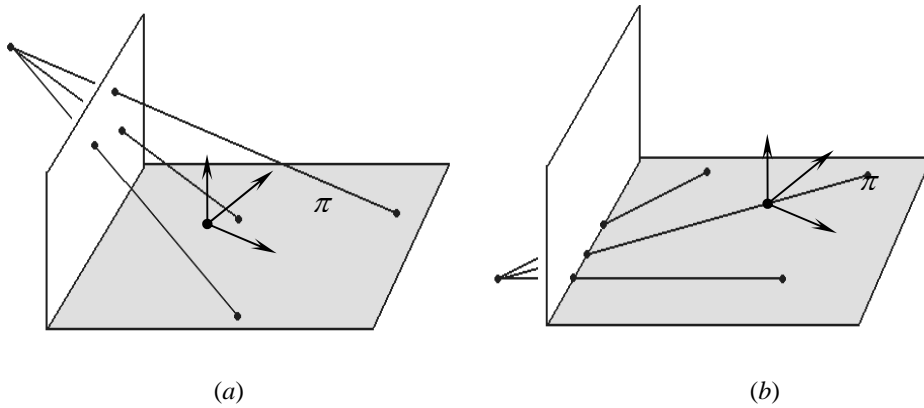


图 3.2.3: (a)空间平面到图像平面的单应变换 H ; (b)退化情况：摄像机中心在空间平面上。此时空间平面的像是该平面与像平面的交线， $\text{rank}(H)=2$ 。

显然， $\text{rank}(H) > 1$ ，这是因为摄像机矩阵是秩 3 的。 $\text{rank}(H)=2$ 的充要条件是平面 π 通过摄像机中心。事实上， $C = (x, y, z, 1)^T$ 为摄像机光心当且仅当 $PC = 0$ 。若 $\text{rank}(H)=2$ ，则 H 有一维右零空间，

于是存在 $C_\pi = (x, y, 1)^T \in \pi$ 使得 $P(x, y, 0, 1)^T = HC_\pi = 0$, 因此平面 π 一定通过摄像机光心。反之, 若平面 π 通过摄像机光心, 则 H 有一维右零空间, 于是 $\text{rank}(H)=2$ 。

当通过摄像机光心时, 平面 π 的投影是一条直线, 这条直线是平面 π 与像平面的交线; 当平面 π 不通过摄像机光心时, H 是一个秩 3 的矩阵, 因此它是从平面 π 到像平面二维射影变换, 而且是中心投影变换, 如图 3.2.3 (a) 所示。以后在讨论平面投影时, 除特别说明外均假定该平面不通过摄像机光心。

通常称矩阵 H 为平面 π 到像平面的单应矩阵, 或简称单应。由于单应矩阵是齐次的, 因此它有 8 个自由度。

无穷远单应

假定无穷远点 $X = (x, y, z, 0)^T$ 的图像为 m , 则有

$$m = PX = (p_1, p_2, p_3)X_\infty$$

其中 $X_\infty = (x, y, z)^T$ 。所以, 无穷远平面到像平面的单应, 简称无穷远单应, 可表示为

$$H_\infty = (p_1, p_2, p_3) = KR \quad (3.2.9)$$

式(3.2.9)同时也说明无穷远单应是摄像机矩阵前三列所构成的子矩阵。由此, 摄像机矩阵可表示为 $P = (H_\infty, p_4)$ 。下述命题是明显的:

命题 3.2.6 令 $\varpi^* = KK^T$, 则 $H_\infty H_\infty^T = \varpi^*$ 。

3.2.4 二次曲线

正向投影

考虑空间中二次曲线在摄像机下的投影。二次曲线是平面曲线, 将二次曲线的支撑平面记为 π , 如上节那样建立世界坐标系, 则摄像机对平面 π 的作用可由一个单应矩阵 H 来描述, 即 $m = HX_\pi$ 。二次曲线可以用一个 3×3 对称矩阵 C 来表示, 即它的方程可写成

$$X_\pi^T C X_\pi = 0, \quad (X_\pi = (x, y, 1)^T \in \pi)$$

由于对任意 $X_\pi \in C$, 其图像点 $m = HX_\pi$, 所以

$$m^T H^{-T} C H^{-1} m = X_\pi^T C X_\pi = 0 \quad (3.2.13)$$

由于 $C_m = H^{-T} C H^{-1}$ 仍为一个 3×3 对称矩阵, 所以它表示图像平面上的一条二次曲线。因此, 二次曲线的图像仍是二次曲线。

命题 3.2.6 二次曲线 C 的图像仍是一条二次曲线 C_m 。如果二次曲线 C 的支撑平面到像平面的单应矩阵为 H , 则有

$$C_m = H^{-T} C H^{-1} \quad (3.2.14)$$

绝对二次曲线的图像 (IAC) 绝对二次曲线 Ω_∞ 的支撑平面是无穷远平 π_∞ , 可由下述方程来描述:

$$(x, y, z) I (x, y, z)^T = 0 \quad (3.2.15)$$

它是 π_∞ 上一条虚二次曲线。下面考虑绝对二次曲线在摄像机下的像曲线。

令摄像机矩阵为 $P = (p_1, p_2, p_3, p_4)$ 。从式(3.2.9), 无穷远单应矩阵为 $H_\infty = (p_1, p_2, p_3)$ 。于是, 由命题 3.2.6, Ω_∞ 的图像是二次曲线: $\varpi = H_\infty^{-T} I H_\infty^{-T}$ 。因 $H_\infty = (p_1, p_2, p_3) = \alpha K R$, 所以,

$$\varpi = K^{-T} K^{-1} \quad (3.2.16)$$

这里略去了常数因子 α^2 , 因为二次曲线的矩阵表示与常数因子无关, 即 C 与 sC 表示同一条二次曲线)。这样, 就得到了下述非常重要的命题:

命题 3.2.7 绝对二次曲线在摄像机下的像曲线为 $\varpi = K^{-T} K^{-1}$ 。

绝对二次曲线的图像 (IAC) 与世界坐标系的选择无关(或者说与摄像机位置无关), 仅与摄像机内参数有关。命题 3.2.7 是摄像机自标定的理论基础。

反向投影

令 C_m 是图像平面上的二次曲线, 考虑它在摄像机 P 下的反投影。在几何上, C_m 的反投影是顶点在摄像机中心并通过二次曲线 C_m 的一个锥面 Q , 它是一个退化的二次曲面。这一结论的代数表述是下面的命题:

命题 3.2.8 设摄像机矩阵为 P , 则二次曲线 C_m 的反投影 Q 为

$$Q = P^T C_m P \quad (3.2.17)$$

证明 图像点 m 在二次曲线 C_m 上, 当且仅当 $m^T C_m m = 0$ 。由于空间点 X 的投影是 $m = P X$, 所以 X 被投影到二次曲线 C_m 上当且仅当 $X^T P^T C_m P X = 0$ 。由于对称矩阵 $P^T C_m P$ 是秩 3 的, 所以二次曲线 C_m 的反投影是锥面 $Q = P^T C_m P$ 。证毕。

2.2.4 二次曲面

二次曲面 Q 可用一个 4×4 的对称矩阵来表示, 它的方程为 $X^T Q X = 0$ 。当 $\text{rank}(Q)=4$ 时, Q 表示一个非退化的二次曲面; 当 $\text{rank}(Q)=3$ 时, Q 是一个锥面; 当 $\text{rank}(Q)=2$ 时, Q 表示两张不重合的

平面；当 $rank(Q)=1$ 时， Q 表示两张重合的平面。在考虑摄像机对二次曲面的作用时，一般假定 Q 是非退化的或是一个锥面。

令 O 是空间中的一个点， Q_c 是以 O 为顶点且与 Q 相切的所有射线组成的集合，它是一个以 O 为顶点的锥面，通常称 Q_c 为二次曲面 Q 的视锥面。锥面 Q_c 与二次曲面 Q 相切于一条二次曲线 Γ ，它是 Q_c 中所有母线与 Q 的切点的集合，如图 3.2.4 所示。称 Γ 是 Q 的一条轮廓线。显然， Q 的轮廓线与视锥面的顶点有关。

二次曲面的投影

在几何上，二次曲面 Q 的轮廓 Γ 在图像平面上的投影 C 是顶点在摄像机光心的视锥面与像平面的交线。二次曲面 Q 的图像是交线 C 所包含的区域，如图 3.2.4 所示。通常称 C 是二次曲面图像的轮廓线。因 Γ 是一条二次曲线，所以 C 也是一条二次曲线。如果二次曲面上没有纹理，它的轮廓线 Γ 的图像 C 是唯一可以利用的图像信息。因此，以后称 C 是二次曲面 Q 的图像。在代数上，我们有下述命题。

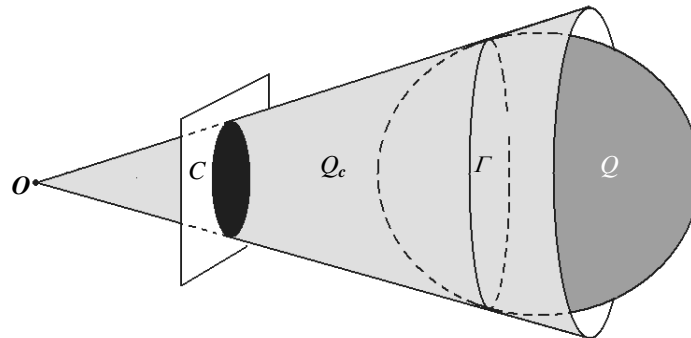


图 3.2.4：二次曲面的投影

命题 3.2.9 (1) 令摄像机矩阵为 P ，二次曲面 Q 的对偶为 Q^* ，它的图像 C 的对偶为 C^* ，则有

$$C^* = PQ^*P^T \tag{3.2.18}$$

(2) 二次曲面的轮廓线 Γ 所在的平面为 $\pi_\Gamma = QO$ ，其中 O 是摄像机的光心坐标。

证明 结论(2)可从二次曲面对极关系推出，因为摄像机的中心 C 关于二次曲面的极面是 π_C 。下面仅证明(1)。令 l 是 C 的任一条切线，则必有

$$l^T C^* l = 0 \tag{3.2.19}$$

由命题 3.2.3， l 的反投影面为 $\pi = P^T l$ ，并且它与二次曲面 Q 相切。因此，有

$$l^T P Q^* P^T l = 0$$

结合(3.2.19)式, 就得到 $C^* = P Q^* P^T$ 。证毕。

关于二次曲面的反投影锥面 Q_c , 由式(3.2.17) 直接给出。

绝对二次曲面的图像 (DIAC) 绝对二次曲线 Ω_∞ 的对偶是三维空间中一个退化的对偶二次曲面, 称为绝对二次曲面并记为 Q_∞^* 。在几何上, Q_∞^* 由 Ω_∞ 的切平面所组成。在代数上 Q_∞^* 由一个秩为 3 的 4×4 的齐次矩阵来表示:

$$Q_\infty^* = \begin{pmatrix} I & 0 \\ 0^T & 0 \end{pmatrix} \quad (3.2.20)$$

绝对二次曲面是绝对二次曲线的对偶, 所以它在摄像机 P 下的图像是绝对二次曲线图像的对偶。因此, 绝对二次曲面 Q_∞^* 的图像是

$$\omega^* = \omega^{-1} = K K^T \quad (3.2.21)$$

这样, 就得到了下述命题:

命题 3.2.10 绝对二次曲面在摄像机下的图像为 $\omega^* = K K^T$ 。

与绝对二次曲线图像(IAC)一样, 绝对二次曲线的对偶图像(DIAC), 即绝对二次曲面的图像, 也是图像平面上的一条(对偶)二次曲线, 与摄像机位置无关而仅与内参数有关。因此, 它也构成了摄像机自标定的理论基础。

3.3 恢复平面景物的几何结构

本节主要讨论如何从平面景物的图像, 恢复景物的仿射结构、相似(或度量)结构以及绝对欧氏结构。从上一节我们知道, 摄像机关于平面景物的成像过程是景物平面到图像平面之间的二维射影变换, 这个可用一个 3 阶可逆矩阵 H 来表示。如果从图像能获得在景物平面的某个仿射坐标系下的二维射影变换 H , 则这个变换的逆 H^{-1} 作用到图像点 m 就可以得到它所对应的景物点在景物平面的仿射坐标系下的坐标 $X = H^{-1} m$, 从而可以得到平面景物的仿射结构。类似地, 如果从图像能获得在景物平面的欧氏坐标系下的二维射影变换 H , 如果这个欧氏坐标系的度量单位与绝对(或实际)度量相差一个常数倍, 则可以通过景物的图像点和变换 H 可以得到景物的相似结构; 如果这个欧氏坐标系的度量单位与绝对度量相等, 则通过景物的图像点和变换 H 可以得到景物的绝对欧氏结构。因此, 恢复平面景物的仿射结构, 等价于某个仿射坐标系下确定景物平面到图像平面的射影变换; 恢复平

面景物的度量结构，等价于在某个欧氏坐标系下确定景物平面到图像平面的射影变换；恢复平面景物的绝对欧氏结构，等价于在某个绝对欧氏坐标系下确定景物平面到图像平面的射影变换。

3.3.1 仿射结构

并非对任何平面场景的图像都能恢复它的仿射结构，那么需要什么信息才能确定仿射结构呢？对此，我们有下列命题。

命题 3.3.1 从图像能恢复平面景物仿射结构的充要条件是能确定物体平面上无穷远直线的图像直线。

证明 必要性： 令 $\{X_j^{(A)}\}$ 是平面景物的图像 $\{m_j | j=1,2,\dots,n\}$ 的仿射结构，即 $\{X_j^{(A)}\}$ 是平面景物点在某个仿射坐标系下的坐标，由点对应 $\{X_j^{(A)} \leftrightarrow m_j\}$ 可以得到景物平面到像平面的射影变换 $H^{(A)}$ 。齐次坐标为 $X_1^{(A)} = (1, 0, 0)^T$ ， $X_2^{(A)} = (0, 1, 0)^T$ 的点是景物平面上的两个无穷远点，并且它们的图像点坐标分别为 $v_1 = H^{(A)} X_1^{(A)} = h_1^A$ ， $v_2 = H^{(A)} X_2^{(A)} = h_2^A$ ，因此景物物体平面上无穷远直线的图像直线为 $l = h_1^{(A)} \times h_2^{(A)}$ 。

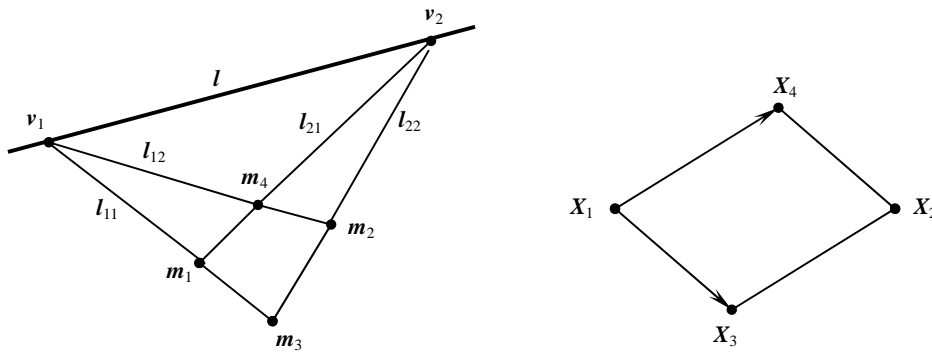


图 3.3.1：从无穷远直线的图像直线恢复景物的仿射几何。

充分性： 已知景物物体平面上无穷远直线的图像直线 l 。在 l 上任取两个点 v_1, v_2 ，另外在直线 l 外再在图像平面内选取两个点 m_1, m_2 。这样，在图像平面上可得到 4 条直线 $l_{11} = v_1 \times m_1, l_{12} = v_1 \times m_2$ ， $l_{21} = v_2 \times m_1, l_{22} = v_2 \times m_2$ ，显然，与它们对应的景物平面直线构成两组平行线，因此可得到景物平面上的一个平行四边形 $\{X_1, X_2, X_3, X_4\}$ 和它的图像 $\{m_1, m_2, m_3, m_4\}$ ，如图 3.3.1 所示。在景物平面上建立仿射坐标系，使得这个平行四边形的四个顶点的分别坐标为

$$X_1 = (0, 0, 1)^T, X_3 = (1, 0, 1)^T, X_4 = (0, 1, 1), X_2 = (1, 1, 1)$$

四个点对应 $X_j \leftrightarrow m_j, j=1,2,3,4$ ，确定了物体平面到像平面的唯一射影变换 $H^{(A)}$ ，即存在唯一的射影变换 $H^{(A)}$ 使得 $m_j = HX_j, j=1,2,3,4$ 。该射影变换可使用 1.1.3 节所给的方法计算出来。令 $\{\hat{m}_j | j=1,2,\dots,n\}$ 为所有景物点的图像所构成的集合，由射影变换 $H^{(A)}$ 可计算出这些图像点所对应的景物点在上述仿射坐标系下的仿射坐标 $X_j^{(A)} = (H^{(A)})^{-1} m_j, j=1,2,\dots,n$ 。这样，我们就得到了景物的仿射几何结构。

3.3.2 相似结构

命题 3.3.2 从图像能恢复平面景物的相似结构的充要条件是在像平面内能确定景物平面两个圆环点的图像。

证明 必要性： 令 $\{X_j^{(s)}\}$ 是平面景物的图像 $\{m_j | j=1,2,\dots,n\}$ 的相似结构，即 $\{X_j^{(s)}\}$ 是平面景物点在某个欧氏坐标系下的坐标，由点对应 $\{X_j^{(A)} \leftrightarrow m_j\}$ 可以得到景物平面到像平面的射影变换 $H^{(s)}$ 。由于两个圆环点在任何欧氏坐标系下的坐标总是 $X_1^{(s)} = (1, i, 0)^T, X_2^{(s)} = (1, -i, 0)^T$ ，所以，两个圆环点的图像分别为 $m_1 = H^{(s)} X_1^{(s)} = h_1^{(s)} + ih_2^{(s)}, m_2 = H^{(s)} X_2^{(s)} = h_1^{(s)} - ih_2^{(s)}$ 。

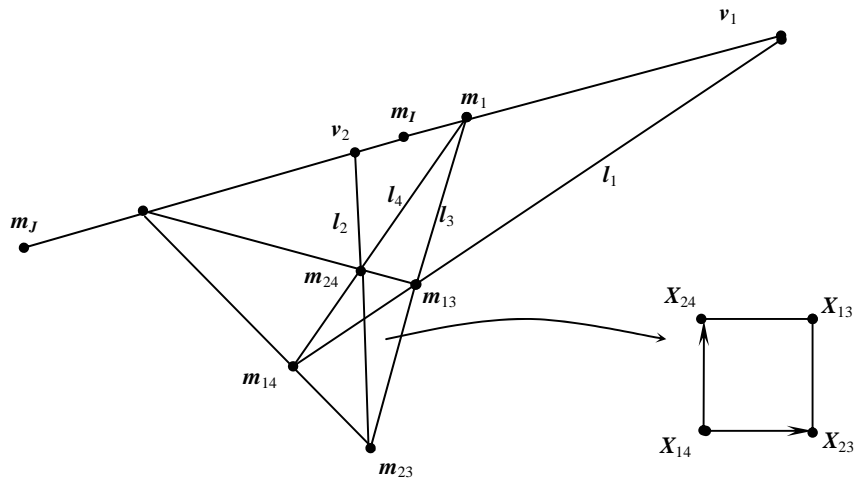


图 3.3.2：从两个圆环点的图像恢复景物的相似结构

充分性： 已知物体平面圆环点图像 m_1, m_2 ，则无穷远直线的图像直线必为 $l = m_1 \times m_2$ 。在 l 上取两个点 v_1, v_2 使得它们与 m_1, m_2 成调和共轭，则通过 v_1, v_2 的任意两条直线所对应的景物平面上的两条直线是相互正交的，如图 3.3.2 所示。另外，再在直线 l 上取一个点 m_1 使得 $[\ln(v_1 m_1; m_2 m_1)]/2i = \pi/4$ ，则由拉盖尔 (Laguerre) 定理，分别通过 v_1, m_1 的任意两条直线所对应

的景物平面上的两条直线之间的夹角等于 $\pi/4$ 。令 l_1, l_2 是分别通过 v_1, v_2 的两条直线, l_3, l_4 是通过 m_1 的两条直线, 则图像平面上的四个点:

$$m_{13} = l_1 \times l_3, \quad m_{14} = l_1 \times l_4, \quad m_{23} = l_2 \times l_3, \quad m_{24} = l_2 \times l_4$$

所对应的景物平面上的 4 个点 $X_{13}, X_{14}, X_{23}, X_{24}$ 必为一个正方形的 4 个顶点。在景物平面上, 以正方形的两条邻边为坐标轴、以边长为度量单位建立欧氏坐标系(如图 3.3.2 所示), 则正方形的四个顶点的坐标分别为

$$X_{14} = (0, 0, 1)^T, \quad X_{23} = (1, 0, 1)^T, \quad X_{24} = (0, 1, 1)^T, \quad X_{13} = (1, 1, 1)^T$$

上述 4 个点对应唯一确定景物平面到像平面的一个射影变换 H 。令 $\{m_j \mid j=1,2,\dots,n\}$ 为平面景物所有点的图像所构成的集合。由单应矩阵 H 可计算出景物体点在上述欧氏坐标系下的欧氏坐标 $X_j^{(s)} = H^{-1}m_j, j=1,2,\dots,n$ 。这样, 我们就得到了景物的相似结构。

3.3.3 绝对欧氏结构

命题 3.3.3 从图像能恢复平面景物的欧氏结构的充要条件是能确定物体平面两个圆环点的图像, 并且已知某两个图像点所对应的两个物体平面点之间的距离。

证明 必要性: 令 $\{X_j^{(e)}\}$ 是平面景物图像 $\{m_j \mid j=1,2,\dots,n\}$ 的一个欧氏结构, 即 $\{X_j^{(e)}\}$ 是景物点在绝对欧氏坐标系下的坐标。由点对应 $\{X_j^{(e)} \leftrightarrow m_j\}$ 可以得到景物平面到像平面的射影变换 $H^{(e)}$ 。于是, 物体平面上两个圆环点的图像为

$$m_1 = H^{(e)}(1, i, 0)^T = h_1^{(e)} + ih_2^{(e)}, \quad m_2 = H^{(e)}(1, -i, 0)^T = h_1^{(e)} - ih_2^{(e)}$$

并且对任意两个图像点 m_1, m_2 所对应的景物平面上两个点 $X_1^{(e)}, X_2^{(e)}$ 之间的实际距离为

$$d(X_1^{(e)}, X_2^{(e)}) = d(H^{(e)-1}m_1, H^{(e)-1}m_2)$$

充分性: 已知 m_1, m_2 所对应的物体平面上两个点 X_1, X_2 之间的实际欧氏距离为 d , 以及两个圆环点图像 m_1, m_2 。在图像上, 直线 $l_1 = m_1 \times m_2$ 必相交于直线 $l = m_1 \times m_2$ 上的一个点 v_1 。令 o 在直线 l 上, 并且使得 v_1, o 与 m_1, m_2 成调和共轭, 于是 o 所对应的物体平面点 O 必是 X_1, X_2 的中点。在 l 上选取另一个点 v_2 使得 v_1, v_2 与 m_1, m_2 成调和共轭, 则直线 $l_1, l_2 = o \times v_2$ 所对应的物体平面上的两条直线相互正交, 如图 3.3.3 所示。与命题 3.3.2 的证明类似, 再在直线 l 取一个点 m 使得 $[\ln(v_1 m, m_1 m_2)]/2i = \pi/4$, 令 l_3, l_4 是分别通过 $(m, m_1), (m, m_2)$ 的两条直线, 则图像平面上的 4 个点:

$$m_1 = l_1 \times l_4, \quad m_2 = l_1 \times l_3, \quad m_3 = l_2 \times l_3, \quad m_4 = l_2 \times l_4$$

所对应的物体平面上的 4 个点 X_1, X_2, X_3, X_4 必构成一个对角线长度为 d 的正方形顶点。在物体平

面上建立欧氏坐标系如图 3.3.3 所示，则正方形 X_1, X_2, X_3, X_4 的坐标分别为

$$X_1 = \begin{pmatrix} d/2 \\ 0 \\ 1 \end{pmatrix}, \quad X_2 = \begin{pmatrix} -d/2 \\ 0 \\ 1 \end{pmatrix}, \quad X_3 = \begin{pmatrix} 0 \\ d/2 \\ 1 \end{pmatrix}, \quad X_4 = \begin{pmatrix} 0 \\ -d/2 \\ 1 \end{pmatrix}$$

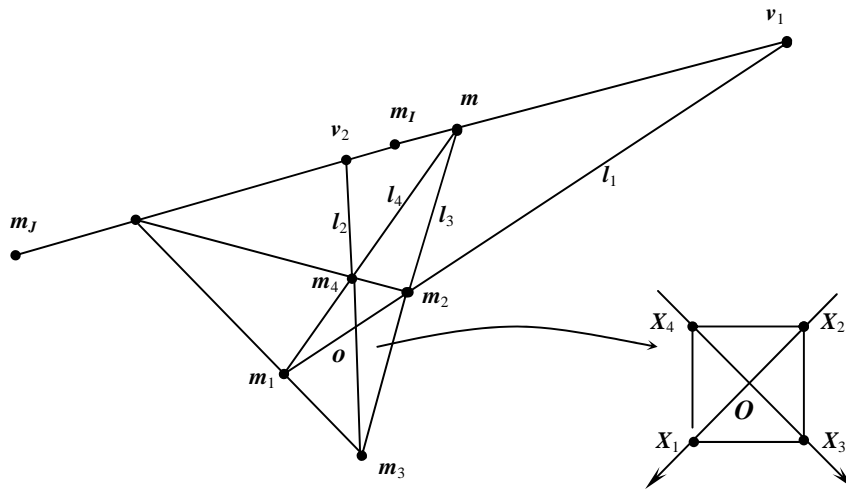


图 3.3.3: 恢复绝对欧氏结构

从点对应 $X_j \leftrightarrow m_j, j=1,2,3,4$ 可计算出景物平面到像平面的一个射影变换 H 。
 $\{m_j \mid j=1,2,\dots,n\}$ 为景物所有点的图像所构成的集合，由射影变换 H 可计算出这些图像点所对应的
 物体平面点的欧氏坐标 $X_j^{(e)} = H^{-1}m_j, j=1,2,\dots,n$ 。由于此时物体平面坐标系是一个欧氏坐标系，
 并且与真实欧氏坐标系有相同的度量单位，所以 $\{X_j^{(e)} \mid j=1,2,\dots,n\}$ 是景物的绝对欧氏结构。

4. 两视点几何

在本章, 主要介绍两个视点图像间的几何关系, 即所谓的两幅图像间的极几何。极几何是两幅图像的点、线关联关系(对应点在对应极线上), 这种关联关系可用所谓的基本矩阵进行代数描述。基本矩阵有 7 个自由度, 从 8 对图像点对应可以线性唯一确定。在射影(等价)意义下, 基本矩阵确定了两幅图像所对应的摄像机矩阵。

4.1 基本矩阵

4.1.1 极几何

先介绍极几何中所涉及的有关几何元素。

极平面(Epipolar Plane)

通过两个摄像机光心的平面称为极平面。两个摄像机光心的连线称为基线, 任意两张极平面均相交于基线, 所以极平面的全体构成共基线的平面束。两个摄像机的极几何(Epipolar Geometry)由极平面导入。

极线(Epipolar Line)

极平面与摄像机像平面的交线称为极线。同一张极平面 π 与两个摄像机像平面的交线 l, l' 称为一对极线对应, 如图 4.1.1 所示。显然, 极平面 π 上任一点 X 在第一个摄像机像平面上的投影 m 必位于极线 l 上, 即 $m \in l$ 。同样, 点 X 在第二个摄像机像平面上的投影 m' 必位于极线 l' 上, 即 $m' \in l'$ 。

极点(Epipolar Point)

基线与像平面的交点称为极点。两个摄像机像平面上的极点分别记为 e, e' 。显然, e 是第二个摄像机光心在第一个摄像机像平面上的投影, e' 是第一个摄像机光心在第二个摄像机像平面上的投影。由于所有极平面都通过基线, 所以第一个像平面上的所有极线都相交于极点 e , 从而第一个像

平面上的所有极线构成共极点 e 的平面线束, 同样第二个像平面上的所有极线构成共极点 e' 的平面线束。

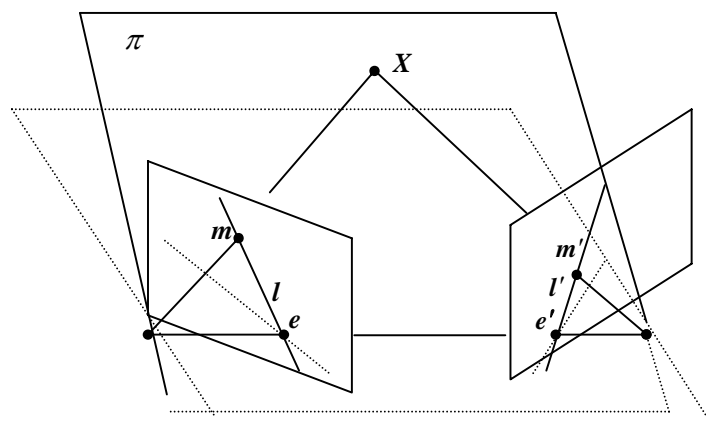


图 4.1.1: 两个摄像机的极几何

令 m, m' 是空间点 X 分别在两个摄像机下的投影, 通常称它们为点对应 (Point Correspondence), 并记作 $m \leftrightarrow m'$ 。若 $m \leftrightarrow m'$, 则 m 的反投影线 $l_m^{(b)}$ 与 m' 的反投影线 $l_{m'}^{(b)}$ 必相交于一个空间点 X 。因此, 反投影线 $l_m^{(b)}, l_{m'}^{(b)}$ 确定一张通过两摄像机光心的平面 π , 它是一张极平面, 由它所确定的一对对应极线分别记为 $l_{m'}, l_m$, 通常称 $l_{m'}, l_m$ 为极线对应 (Epipolar Line Correspondence), 并记作 $l_{m'} \leftrightarrow l_m$ 。显然 l_m 是反投影线 $l_{m'}^{(b)}$ 在第二个摄像机下的投影, 且 m' 在极线 l_m 上, 同样 $l_{m'}$ 是反投影线 $l_m^{(b)}$ 在第一个摄像机下的投影, 且 m 在极线 $l_{m'}$ 上。如图 4.1.2 所示。极线 l_m 称为像点 m 对应的极线, $l_{m'}$ 称为像点 m' 对应的极线。

极几何约束

总结上述讨论, 得到下述所谓的极几何约束。

命题 4.1.1 令 $m \leftrightarrow m'$ 是一个点对应, 则 m' 必位于 m 对应的极线 l_m 上, m 必位于 m' 对应的极线 $l_{m'}$ 上, 即 $m' \in l_m, m \in l_{m'}$ 。

给定两个摄像机下的图像 (I, I') , 极几何约束表明:

1. $\forall m \in I$, 在 I' 上存在一条极线 l_m 与之对应, 并且与它对应的像点 $m' \in l_m$;
2. $\forall m' \in I'$, 在 I 上存在一条极线 $l_{m'}$ 与之对应, 并且与它对应的像点 $m \in l_{m'}$ 。

在上述讨论中, 没有利用场景的任何几何结构信息, 因此极几何约束与场景的几何结构无关,

它是两幅图像间的固有射影性质。

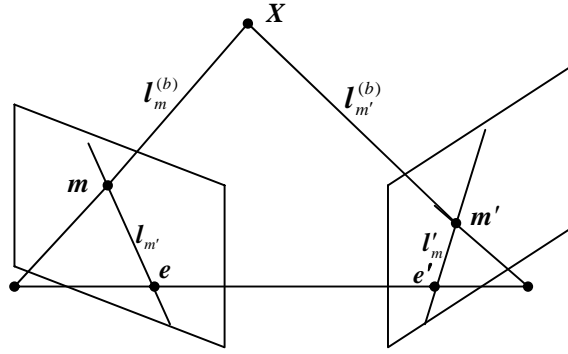


图 4.1.2: 极线对应 $l'_m \leftrightarrow l_m$

4.1.2 基本矩阵

极几何所描述的图像点与其极线的对应关系，可以用所谓的基本矩阵（Fundamental Matrix）来刻画，换句话说基本矩阵是极几何的代数表示。

基本矩阵

假定两个摄像机矩阵分别为 P, P' ，记两个摄像机的像平面分别为 I, I' ，则 $\forall m \in I$ 的反投影线 $l_m^{(b)}$ 的参数方程为

$$X(s) = P^+m + sC, \quad s \in (-\infty, \infty)$$

其中， P^+ 是 P 的广义逆，即 $PP^+ = I$ ， C 是第一个摄像机的光心，即 $PC = 0$ 。于是可推知：

$$l'_m = e' \times m' = (P'C) \times (P'X(s)) = (P'C) \times (P'P^+m + sP'C) = (P'C) \times (P'P^+m) = [e']_x P'P^+m$$

即，

$$l'_m = [e']_x P'P^+m$$

记

$$F = [e']_x P'P^+ \tag{4.1.1}$$

并称它为两个摄像机间或两幅图像 (I, I') 间的基本矩阵。基本矩阵 F 描述了图像点 m 与其对应极线 l'_m 间的对应关系：

$$l'_m = Fm \tag{4.1.2}$$

由于图像点 m 在第二幅图像上的对应点 m' 在极线 l'_m 上，所以必有

$$m'^T Fm = 0 \tag{4.1.3}$$

并且

$$Fe = 0, F^T e' = 0 \quad (4.1.4)$$

因此, 基本矩阵 F 刻划了两幅图像间的极几何。

基本矩阵 F 将二维空间(第一个图像平面)点(齐次)线性映射到共点线束(第二幅图像平面上的极线束), 而共点线束构成一维空间, 因此 F 的秩必为 2。基本矩阵是一个齐次变换矩阵, 它有 8 个参数需要确定, 而 $\text{rank}(F) = 2$ 导致 F 的一个约束 $\det(F) = 0$, 因此基本矩阵有 7 个自由度。对每一个图像点对应, 式(4.1.3)为基本矩阵提供一个约束。所以, 在一般情况下, 利用 7 对以上的图像点对应就可以确定基本矩阵。

基本矩阵的性质

下述命题概括了基本矩阵的一些性质。

命题 4.1.2 令 F 是两幅图像 (I, I') 间的基本矩阵, e, e' 是极点, $m \leftrightarrow m'$ 是任意一个点对应, 则

1. $l'_m = Fm$ 是 m 对应的极线, $l_m = F^T m'$ 是 m' 对应的极线; 或等价地表示为 $m'^T Fm = 0$;
2. 两极点分别满足下述方程: $Fe = 0, F^T e' = 0$;
3. 基本矩阵有 7 个自由度且 $\text{rank}(F) = 2$ 。

例 4.1.1 假定第一个摄像机的内参数矩阵为 K , 第二个摄像机的内参数矩阵为 K' , 第二个摄像机相对于第一个摄像机的方位为 (R, t) , 即第二个摄像机与第一个摄像机坐标系之间的关系为 $X' = RX + t$, 其中 R 是旋转矩阵, t 是平移向量, 则第一与第二个摄像机的投影矩阵可表示为

$$P = K(I, 0), \quad P' = K'(R, t) \quad (4.1.5)$$

由于 $P^+ = \begin{pmatrix} K^{-1} \\ 0^T \end{pmatrix}$, 且 $e' = K't$, 所以

$$F = [K't]_x P' P^+ = [K't]_x (K'R, K't) \begin{pmatrix} K^{-1} \\ 0^T \end{pmatrix} = [K't]_x K'RK^{-1}$$

因 $[K't]_x$ 是一个秩为 2 的矩阵, 而 $K'RK^{-1}$ 是一个秩为 3 的矩阵, 所以从这里再次得到基本矩阵 F 的秩为 2 的结论。

对 $[K't]_x K'RK^{-1}$ 通过代数运算, 还可以得到下述一系列等式:

$$F = [K't]_x K'RK^{-1} = K'^{-T} [t]_x RK^{-1} = K'^{-T} RK^T [KR^T t]_x \quad (4.1.6)$$

事实上, 式(4.1.6)利用摄像机内参数和相对方位给出了基本矩阵的三种表示。由于 $e = KR^T t$,

$e' = K't$ ，所以，由式(4.1.6)又可以得到 $Fe = 0, F^T e' = 0$ 。

显然， $F^T = [KR^T t]_{\times} KR^T K'^{-1}$ 。矩阵 F^T 描述了图像点 m' 与它对应的极线 l_m 间的对应关系 $l_m = F^T m'$ 。

极线对应

下面利用基本矩阵来描述极线之间的对应关系。

如前所述，一张极平面 π 与两个摄像机像平面的交线 l, l' 构成一个极线对应，这样极平面在几何上确定了两幅图像极线之间的对应关系。如何使用基本矩阵来刻画这个对应关系呢？为此，令 q 是第一幅图像上不通过极点的任一条直线， $l \leftrightarrow l'$ 是任一极线对应，如图 4.1.3 所示。因 $q \times l = [q]_{\times} l$ 是两直线 q, l 的交点 m ，因而在 l 上。于是， $l' = F(q \times l) = F[q]_{\times} l$ 。特别地，取 $q = e$ ，它是以极点 e 作为线坐标的直线，从而不通过极点 e （因为 $e^T e \neq 0$ ）。因此，有下述命题：

命题 4.1.3 令 F 是两摄像机间的基本矩阵， e, e' 分别是两个摄像机的极点。对任一极线对应 $l \leftrightarrow l'$ ，必有

$$l' = F[e]_{\times} l, l = F^T [e']_{\times} l' \tag{4.1.7}$$

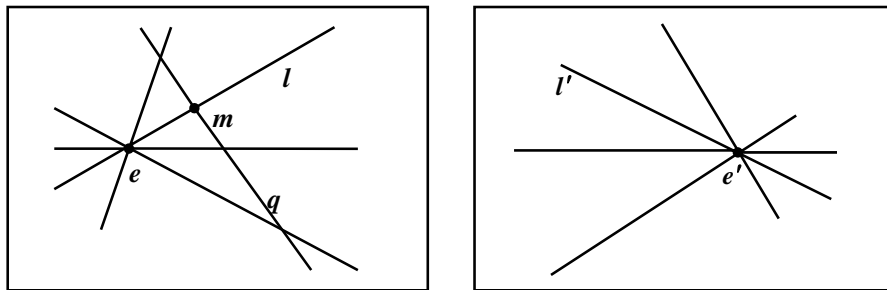


图 4.1.3: 极线对应关系 $l' = F[e]_{\times} l$

纯平移运动下的基本矩阵

两个摄像机在不同位置下所拍摄的图像可以视为同一个摄像机运动前、后所拍摄的图像。如果两个摄像机的内参数不相同，此时可以看作同一个摄像机作刚体运动时内参数同时发生变化。在本书中，这种等价性简称为这两个摄像机是刚体运动下的两个摄像机。如果两个摄像机的摄像机内参数相同，我们称这两个摄像机是纯刚体运动下的摄像机，这里的“纯”是指摄像机在作刚体运动时，其内参数保持不变。下面讨论在纯平移运动下的基本矩阵。

由(4.1.6)式, 可知纯平移运动 (I, \mathbf{t}) 下两个摄像机间的基本矩阵为

$$F = [\mathbf{e}']_{\times} K K^{-1} = [\mathbf{e}']_{\times} \quad (4.1.8)$$

这就是说, 纯平移运动的基本矩阵 F 是由极点所确定的反对称矩阵 $[\mathbf{e}']_{\times}$, 并且两幅图像上的极点相同, 即 $\mathbf{e}' = \mathbf{e}$ 。

例 4.1.2 当摄像机沿 x 轴作纯平移时, 两幅图像的极点 $\mathbf{e}' = \mathbf{e} = (1\ 0\ 0)^T$, 因此基本矩阵必为

$$F = [\mathbf{e}']_{\times} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

此时, 对任意一点对应 $(u, v, 1)^T \leftrightarrow (u', v', 1)^T$ 均有 $v = v'$ 。这说明两幅图像上的极线都是相互平行的, 且对应极线为同一条水平线。这是立体视觉中的一种典型的摄像机配置。

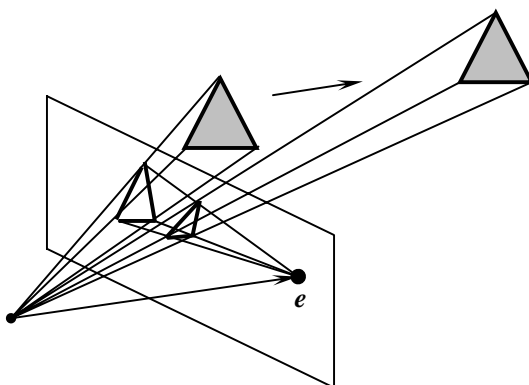


图 4.1.4: 纯平移运动摄像机的极几何

关于纯平移运动的基本矩阵, 可以作如下几何解释: 当摄像机作平移运动且内参数保持不变时, 某个静止物体的两幅图像可视为物体作相对平移运动而摄像机在固定位置下所得到的两幅图像, 如图 4.1.4 所示。平移运动物体上的任何一对对应点的连线必与平移方向平行, 因此所有对应点的连线在空间中相交于同一(无穷远)点, 该点表示了平移方向。因此两幅图像对应点的连线也必交于一点, 该点是平移方向在图像平面上的投影, 由极几何可以知道它是两幅图像的极点。于是, 摄像机作纯平移运动时, 两幅图像的极点相同并且是平移方向在图像平面上的投影, 对应的极线为同一条直线。在运动分析中, 平移运动下的极点通常又称为汇聚点 (Focus of Expansion, FOE)。

4.1.3 几何解释

记 F_s, F_a 分别为基本矩阵 F 的对称部分和反对称部分, 即

$$F_s = \frac{F + F^T}{2}, F_a = \frac{F - F^T}{2}, (F = F_s + F_a)$$

考虑图像对的不动点集合。所谓图像对的不动点是指有相同图像坐标的图像对应点 $m = m'$ 。所有不动点构成的集合可以表示为:

$$\Phi_F = \{m \mid m^T F m = 0\}$$

由于 $m^T F m = m^T F_s m + m^T F_a m$, 而 F_a 是反对称矩阵, 对任意的 m 均有 $m^T F_a m = 0$, 所以不动点集合由 F 的对称部分所确定, 即

$$\Phi_F = \{m \mid m^T F_s m = 0\}$$

在一般的情况下 $\text{rank}F_s = 3$, 这就是说在一般的情况下图像对的不动点集合是一条非退化的二次曲线 F_s , 它是三维空间中一条三次扭线在两个摄像机下的图像^{*)}。由于

$$\begin{aligned} e^T F_s e &= e^T F_s e + e^T F_a e = e^T F e = 0 \\ e'^T F_s e' &= e'^T F_s e' + e'^T F_a e' = e'^T F e' = 0 \end{aligned}$$

所以, 图像的两个极点 e, e' 在二次曲线 F_s 上。

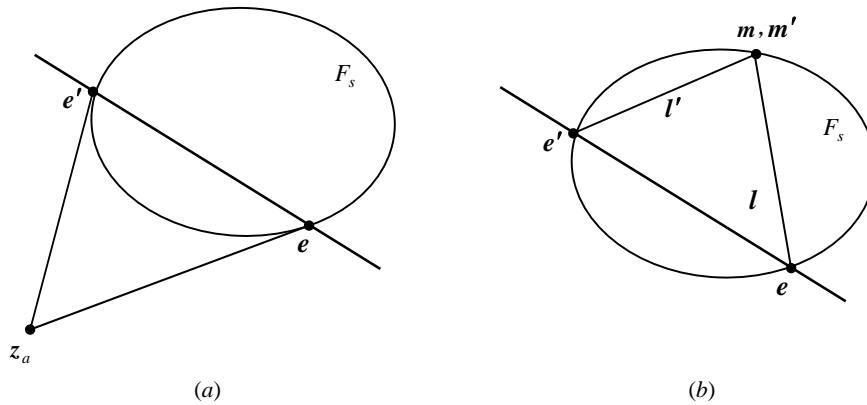


图 4.1.5: 基本矩阵的几何解释

再考虑 F 的反对称部分 F_a 。由于它是反对称矩阵, 所以存在三维向量 z_a 使得 $F_a = [z_a]_{\times}$ 。因

^{*)} 三次扭线是空间中的一类曲线, 其参数方程为 $\mathbf{X} = A(1, t, t^2, t^3)^T$, 其中 A 是可逆 4 阶矩阵, t 是参数。三次扭线在摄像机退化配置的研究中具有非常重要的作用。如果摄像机运动是平面运动 (即平移分量在与旋转轴正交的平面上), $\text{rank}F_s = 2$ 。

$F_a e = [z_a]_x e = F e - F_s e = F_s e$ ，所以将 z_a 作为一个图像点的齐次坐标，极点 e 关于二次曲线 F_s 的极线 $F_s e$ 是直线 $z_a \times e$ 。由于极点 e 在二次曲线 F_s 上，所以 $F_s e = z_a \times e$ 是点 e 处的切线。同理， $z_a \times e'$ 是点 e' 处的切线。所以点 z_a 关于二次曲线的极线是直线 $l_a = e \times e'$ 。如图 4.1.5(a)所示。

最后，考虑两幅图像的极线对应关系。令 m 是二次曲线上 F_s 的一个点，由于它是两幅图像的不动点，即 $m \leftrightarrow m' = m$ 是一个点对应，因此 $l = e \times m \leftrightarrow l' = e' \times m$ 是一个极线对应。如图 4.1.5(b)所示。于是，极线对应关系由二个极点和二次曲线 F_s 所确定，它是一维射影对应。

4.2 单应矩阵

上节所讨论的极几何约束与场景结构无关，也就是说极几何约束对于任何三维场景结构的两幅图像都成立，因此它不能给出两幅图像点间的一一对应关系，只能给出点对应的必要条件，即图像点在另一幅图像上的对应点位于对应的极线上，不可能给出对应点的确切位置。如果已知场景结构的某些信息，无疑两幅图像对应点的约束会得到加强。本节主要讨论空间平面在两个摄像机下的投影几何。空间平面在两个摄像机下图像点具有一一对应关系，这个一一对应关系是齐次线性的，可由一个 3 阶矩阵即所谓的单应矩阵来描述。在本节，还给出了单应矩阵与基本矩阵之间的关系。

4.2.1 单应矩阵

设 π 是不通过两摄像机任一光心的空间平面，它在两个摄像机下的图像分别记为 I, I' ，如图 4.2.1 所示。令 X 是平面 π 上的任一点，它在两个摄像机下的像分别记为 m, m' 。

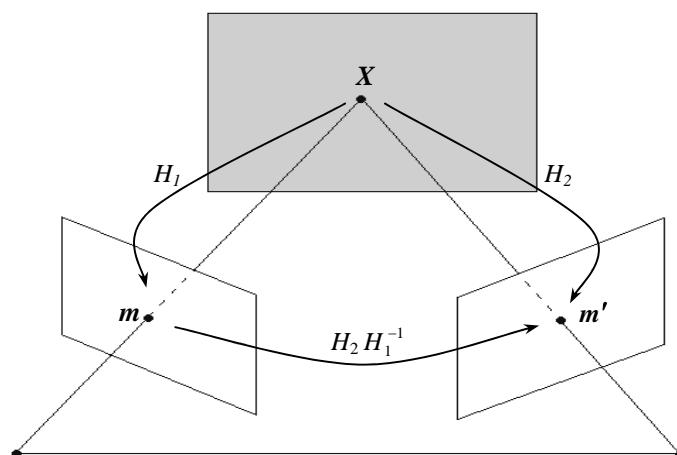


图 4.2.1: 两幅图像的单应变换

从 3.2.3 节可知, 在空间平面 π 到两个像平面之间存在两个单应矩阵 H_1, H_2 使得 $m = H_1 X$, $m' = H_2 X$ 。由于平面 π 不通过两摄像机的任一光心, 所以 H_1, H_2 实现平面 π 到对应的像平面之间的一一变换。因此 m, m' 之间存在一个二维射影变换 $H = H_2 H_1^{-1}$ 使得

$$m' = Hm \tag{4.2.1}$$

即平面在两个摄像机下的图像之间的关系也是一个二维射影变换。矩阵 H 实现从第一个像平面与第二个像平面的一一变换, 从而矩阵 H 为可逆矩阵。

由于两个摄像机的极点 e, e' 是摄像机基线与平面 π 的交点在两个摄像机下的投影, 因此它们是一个点对应, 于是总有 $e' = He$ 。即变换 H 总是将第一幅图像上的极点变换到第二幅图像上的极点。

满足式(4.2.1)的矩阵称为平面 π 诱导的两个摄像机间(或两幅图像间)的单应矩阵, 或简称平面 π 的单应矩阵, 相应的变换称为单应。单应矩阵 H 是一个齐次矩阵, 它有 8 个自由度。根据(4.2.1)式, 利用单应矩阵 H , 从一幅图像上的点可以得到另一幅图像上的对应点, 这是基本矩阵所不具备的性质。

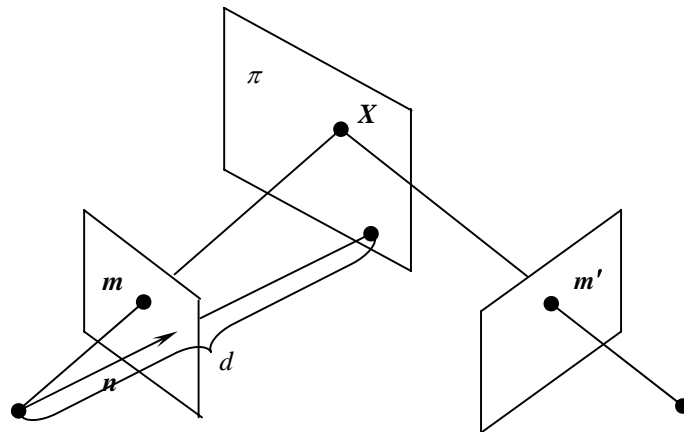


图 4.2.2: 平面 π 在第一个摄像机坐标系下的单位法向量为 n , 坐标原点到平面的距离为 d 。

对于单应矩阵的代数表示, 有下述命题:

命题 4.2.1 假定第一个摄像机的内参数矩阵为 K , 第二个摄像机的内参数矩阵为 K' , 第二个摄像机相对于第一个摄像机的方位为 (R, t) 。若 n 为平面 π 在第一个摄像机坐标系下的单位法向量, d 为坐标原点到平面 π 的距离, 如图 4.2.2 所示, 则平面 π 的单应矩阵可表示为

$$H = K'(R + tn_d^T)K^{-1} \tag{4.2.2}$$

其中: $n_d = \frac{n}{d}$ 。

证明 取第一个摄像机坐标系为世界坐标系, 根据假定, 第一和第二个摄像机矩阵分别为

$$P = K(I, \mathbf{0}), \quad P' = K'(R, t)$$

对于平面 π 上的任一点 $\mathbf{X} = (\tilde{\mathbf{X}}^T, 1)^T$, 考虑它在两个摄像机下的投影。显然, 在第一个摄像机下的投影必为

$$\mathbf{m} = P \begin{pmatrix} \tilde{\mathbf{X}} \\ 1 \end{pmatrix} = K\tilde{\mathbf{X}}$$

因此, 平面 π 到第一个像平面的单应矩阵为 $H_1 = K$ 。在第二个摄像机下的投影为

$$\mathbf{m}' = P' \begin{pmatrix} \tilde{\mathbf{X}} \\ 1 \end{pmatrix} = K'(R, t) \begin{pmatrix} \tilde{\mathbf{X}} \\ 1 \end{pmatrix} = K'(R\tilde{\mathbf{X}} + t)$$

由于平面 π 的方程为 $\mathbf{n}^T \tilde{\mathbf{X}} = d$, 即 $\mathbf{n}_d^T \tilde{\mathbf{X}} = 1$, 因此 t 可以写成 $t = \mathbf{t}\mathbf{n}_d^T \tilde{\mathbf{X}}$ 。于是, 在第二个摄像机下的投影为

$$\mathbf{m}' = K'(R\tilde{\mathbf{X}} + t) = K'(R + \mathbf{t}\mathbf{n}_d^T)\tilde{\mathbf{X}}$$

即平面 π 到第二个像平面的单应矩阵 $H_2 = K'(R + \mathbf{t}\mathbf{n}_d^T)$, 所以两幅图像间的单应矩阵可表示为

$$H = H_2 H_1^{-1} = K'(R + \mathbf{t}\mathbf{n}_d^T) K^{-1}$$

证毕。

无穷远单应

由无穷远平面所诱导的两幅图像的单应矩阵, 简称为无穷远单应矩阵。由于无穷远平面与坐标原点的距离为无穷大, 所以在式(4.2.2)中取 $d \rightarrow \infty$ 的极限值, 就可以得到无穷远单应矩阵。显然,

$$H = K'(R + \mathbf{t}\mathbf{n}_d^T) K^{-1} \rightarrow K'RK^{-1} \quad (d \rightarrow \infty)$$

因此, 无穷远单应矩阵可表示为

$$H_{\infty} = K'RK^{-1} \quad (4.2.3)$$

从上式可以看出: 无穷远单应矩阵与摄像机的平移运动无关, 仅与旋转运动参数有关。

结合式(4.2.2)与(4.2.3), 任何平面 π 的单应矩阵都可以写成下面的形式

^{*)}在第3章曾用符号 H_{∞} 来标记无穷远平面到像平面的单应矩阵, 这里也使用符号 H_{∞} 来标记无穷远平面所诱导的两幅图像间的单应矩阵。这可从上下文来理解符号 H_{∞} 意义, 而这不致于引起混淆。

$$H = H_\infty + e'a^T, \quad a \in R^3 \quad (4.2.4)$$

其中:

- $H_\infty = K'RK^{-1}$ —— 无穷远单应矩阵: 与平面 π 无关, 仅与两个摄像机内参数和相对姿态 (旋转) 有关;
- $e' = K't$ —— 第二幅图像的极点: 与平面 π 无关, 仅与第二个摄像机内参数和相对平移有关;
- $a = K^{-T}n_d$ —— 平面 π 在第一幅图像上的隐消线 (即平面上无穷远直线在第一幅图像上的像直线): 与平面 π 相对第一个摄像机的位置和第一个摄像机内参数有关, 而与第二个摄像机无关。

式(4.2.4)表明, 任何一个平面的单应矩阵, 均可以由无穷远单应、极点的一个三维向量来表示。从下面的命题可以看出, 任何一个平面的单应都可以充当式(4.2.4)中无穷远单应的角色。

命题 4.2.2 令 H^* 为两个摄像机下任一给定平面 π^* 的单应矩阵, 则 3 阶矩阵 H 为某个平面 π 的单应矩阵的充要条件是存在三维向量 $a \in R^3$ 使得

$$H = H^* + e'a^T \quad (4.2.5)$$

证明 必要性: 由式(4.2.4), 存在三维向量 $a_1, a_2 \in R^3$ 使得

$$H^* = s_1(H_\infty + e'a_1^T), \quad H = s_2(H_\infty + e'a_2^T)$$

这里, 显式地给出了齐次因子 s_1, s_2 。于是,

$$\frac{1}{s_2}H - \frac{1}{s_1}H^* = e'(a_2^T - a_1^T)$$

$$H = \frac{s_2}{s_1}(H^* + s_1 e'(a_2^T - a_1^T)) = H^* + e'a^T, \quad a = s_1(a_2 - a_1) \in R^3$$

充分性: 由

$$H = s(H^* + e'a^T) = sK'(R + tn_d^{*T})K^{-1} + \lambda K't(K^T a)^T K^{-1} = sK'(R + t(n_d^* + \lambda(K^T a)^T))K^{-1}$$

知 H 是以方程 $(n_d^* + \lambda(K^T a)^T)\tilde{X} = 1$ 所确定的平面的单应矩阵。证毕。

注: 式(4.2.5)中的三维向量 a 不再是平面 π 在第一幅图像上的隐消线, 而是平面 π^* 与平面 π 的交线在第一幅图像上的像直线。

4.2.2 与基本矩阵的关系

令 π 是不通过两摄像机光心的任一张空间平面, 其图像间的单应矩阵为 H , $m \leftrightarrow m'$ 是两幅图

像的任一点对应, 则有 $m' = Hm$, 因此 m 对应的极线为

$$l'_m = Fm = e' \times m' = e' \times Hm = [e']_x Hm$$

所以基本矩阵可以表示为

$$F = [e']_x H \tag{4.2.6}$$

从下式:

$$F = [K't]_x K'RK^{-1} = [e']_x H_\infty = [e']_x (H_\infty + e'a^T) = [e']_x H$$

同样可以得到式(4.2.6)。

由于 $[e']_x H = H^{-T} [H^{-1}e']_x = H^{-T} [e]_x$, 所以基本矩阵又可以表示为

$$F = H^{-T} [e]_x \tag{4.2.7}$$

式(4.2.6)与(4.2.7)表明基本矩阵由平面单应矩阵与极点唯一确定。事实上, 有下述更深刻的结论:

命题 4.2.3 3 阶矩阵 H 使得 $[e']_x H (H^{-T} [e]_x)$ 为基本矩阵的充要条件是 H 为某个平面的单应矩阵。

证明 从上面的讨论, 充分性已经得到了证明。下面证明必要性: 令基本矩阵 $F = [e']_x H$ 。不难证明齐次矩阵方程 $[e']_x X = 0$ 的通解为

$$X = e'a^T, \quad a \in R^3$$

显然, $X = H_\infty$ 是方程 $F = [e']_x X$ 的一个特解, 因此 $F = [e']_x X$ 的通解为

$$X = H_\infty + e'a^T, \quad a \in R^3$$

所以存在三维向量 a^* 使得

$$H = H_\infty + e'a^{*T}$$

由命题 4.2.2, H 必为平面的单应矩阵。证毕。

表 4.2.1 给出了在各种条件下, 基本矩阵的计算公式。

表 4.2.1: 基本矩阵计算公式

已知条件	基本矩阵
两个摄像机矩阵 P, P'	$F = [P'C]_x P'P^+, PC = 0$
两个摄像机的内参数 K, K' , 两个摄像机的相对运动 (R, t)	$F = [K't]_x K'RK^{-1} = K'^{-T} [t]_x RK^{-1} = K'^{-T} RK^T [KR^T t]_x$
单应矩阵 H , 极点 e, e'	$F = [e']_x H = H^{-T} [e]_x$

4.2.3 不动点与不动线

下面考虑一些特殊单应的不动点与不动直线性质。

纯平移运动

对于纯平移运动下的两个摄像机, 平面的单应矩阵有下述更简单的表示式:

$$H = I + ea^T$$

其中 $e = Kt$ 是两个像平面的共同极点^{*)}, $a = K^{-T}n_d$ 是平面 π 在第一幅图像上的隐消线。

不难看出: $He = (I + ea^T)e = (1 + a^T e)e$, 因此极点是平面单应的一个不动点。令 m 是隐消线 $a = K^{-T}n_d$ 上的任一点, 则 $a^T m = 0$, 所以 $Hm = (I + ea^T)m = m$, 因此隐消线上的任一点都是平面单应的不动点, 而隐消线构成一条不动直线(保持点点不动)。于是, 我们有下述命题:

命题 4.2.4 对于纯平移运动下的两个摄像机, 极点是平面单应的一个不动点, 它是平面单应矩阵 H 的单重特征值的特征向量; 隐消线是平面单应的不动直线(保持点点不动的), 它是平面单应矩阵 H 的二重特征值对应的二维特征子空间。

在纯平移运动下, 平面单应除平面的隐消线作为它的一条(点点不动的)不动直线外, 还有另外的不动直线吗? 为此, 考虑过点 m 的极线 $l = e \times m$ 。 $\forall x \in l$, 即 $(e \times m)^T x = 0$ 。由于 $(e \times m)^T Hx = m^T [e]_x (I + ea^T)x = m^T [e]_x x = (e \times m)^T x = 0$, 所以 $Hx \in l$ 。因此, 每一条极线都是平面单应的一条不动直线。也就是说, 平面单应的所有(非点点不动的)不动直线构成极线束。由于二维射影变换对直线的变换规则由它的对偶所确定, 因此有下述命题:

命题 4.2.5 对于纯平移运动下的两个摄像机, 平面隐消线是平面单应的(保持点点不动的)不动直线, 它是对偶单应矩阵 H^{-T} 单重特征值的特征向量; 极线束是单应的(非点点不动的)不动直线束, 它是对偶单应矩阵 H^{-T} 二重特征值的二维特征子空间。

相对单应矩阵

考虑在一般运动 (R, t) 下的两个摄像机。假定平面 π_1, π_2 在这两个摄像机下的单应矩阵分别为 H_1, H_2 , 则矩阵 $H_{12} = H_2^{-1}H_1$ 确定了第一个像平面到自身的一个二维射影变换(如图 4.2.3 所示), 这个变换称为相对(平面 π_2 的)单应, 矩阵 $H_{12} = H_2^{-1}H_1$ 称为相对(平面 π_2 的)单应矩阵。

由命题 4.2.3, 可知: $H_1 = H_2 + e'a^T, a \in R^3$, 于是, 相对单应矩阵可表示为

$$H_{12} = H_2^{-1}H_1 = H_2^{-1}(H_2 + e'a^T) = I + ea^T, a \in R^3$$

$e = H_2^{-1}e'$ 是第一幅图像上的极点。从上式可以看出相对单应矩阵等价于纯平移运动下的单应矩阵, 因此相对单应有与类似命题 4.2.4 和 4.2.5 的不动点与不动线性质。事实上, 只需要将这两个命题中

^{*)} 实际上, 在纯平移运动下, 由于两个摄像机的内参数相同, 所以它们的像平面是同一张空间平面。

的“平面单应”替代为“相对单应”和“平面隐消线”替代为“平面 π_1, π_2 的交线在第一幅图像上的像直线”，对应的命题成立。

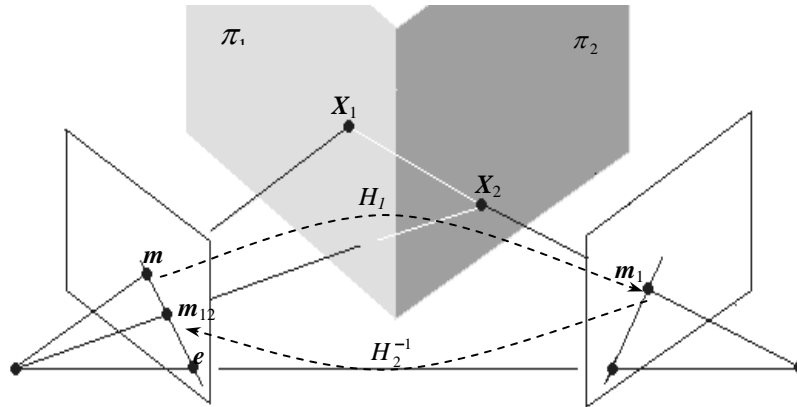


图 4.2.3: 相对单应: 对第一个图像平面上的点 m , 单应 H_1 将它变到第二幅图像的点 m_1 , 它是平面 π_1 上的点 X_1 在第二个摄像机下的投影。 H_2^{-1} 再将 m_1 变换到第一幅图像上的点 m_{12} , 它是 m_1 的反投影线与平面 π_2 的交点 X_2 在第一个摄像机下的投影。因此, 相对单应 $H_{12} = H_2^{-1}H_1$ 是第一幅图像上的一个射影变换。可以看出: m_{12} 是极线 $e \times m$ 上的点, 因为 X_1, X_2 在同一个极平面上。相对单应总是将第一幅图像点 m 变换到它所确定的极线 $e \times m$ 上, 因此极线是相对单应的不动直线。除此之外, 相对单应还有一条保持点点不动的不动直线, 它是两平面交线的图像直线。

4.3 基本矩阵估计

4.3.1 8-点算法(8-Point Algorithm)

给定一对点对应 $m = (u, v, 1)^T, m' = (u', v', 1)^T$, 基本矩阵 F 满足下述方程

$$m'^T F m = 0 \tag{4.3.1}$$

给定充分多(至少 ≥ 8)的点对应, 由上式可线性估计基本矩阵 F 。

令 $F = (f_{ij})$, 则基本矩阵的约束方程(4.2.1)可以写成下述形式:

$$u'uf_{11} + u'vf_{12} + u'f_{13} + v'uf_{21} + v'vf_{22} + v'f_{23} + uf_{31} + vf_{32} + f_{33} = 0$$

记

$$f = (f_{11}, f_{12}, f_{13}, f_{21}, f_{22}, f_{23}, f_{31}, f_{32}, f_{33})^T$$

它是由 F 的 3 个行向量所构成的 9 维列向量, 则上述方程可写成向量内积的形式:

$$(u'u, u'v, u', v'u, v'v, v', u, v, 1)f = 0$$

这样, 给定 N 个对应点可以得到线性方程组:

$$Af = \begin{pmatrix} u'_1 u_1 & u'_1 v_1 & u'_1 & v'_1 u_1 & v'_1 v_1 & v'_1 & u_1 & v_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ u'_N u_N & u'_N v_N & u'_N & v'_N u_N & v'_N v_N & v'_N & u_N & v_N & 1 \end{pmatrix} f = 0 \quad (4.3.2)$$

其中 A 是一个 $N \times 9$ 矩阵。

由于基本矩阵是非零的, 所以 f 是一个非零向量, 即线性方程组有非零解。因此在一般情况下, 当点对应精确时, $\text{rank}(A)=8$, f 是矩阵 A 的零空间, 或者说是矩阵 A 的零特征值对应的特征向量。因此, 当点对应精确时, 直接通过求解线性方程组可以确定基本矩阵 F 。然而, 在实际情况中, 不可能得到精确的图像点对应。在点对应不精确时, 矩阵 A 通常是一个秩 9 的矩阵(这里点对应数 >8 。因为当点对应数 $=8$ 时, A 是一个 8×9 矩阵。不论点对应是否有误差, A 的秩总等于 8), 这样方程组只有零解。

在实际情况中, 不能直接通过求解线性方程组来确定基本矩阵, 而是求在约束条件 $\|f\|=1$ 下方程组(4.2.2)的最小二乘解, 即求解下述问题:

$$\begin{cases} \min \|Af\| \\ \text{subject to } \|f\|=1 \end{cases} \quad (4.3.3)$$

问题(4.3.3)的解是 A 的最小奇异值的右奇异向量。令 A 的奇异值分解为 $A=UDV^T$, 则(4.2.3)的解是 V 的最后一个列向量, 即 $f=v_9$, 于是由 $f=v_9$ 可构造矩阵 F 。

秩 2 约束 基本矩阵秩为 2 的约束是它的一个重要性质。如果用一个秩不等于 2 的矩阵作为基本矩阵, 并用它估计图像极线, 则所估计的极线不交于同一个图像点。这就是说如果用秩不为 2 的矩阵作为基本矩阵的估计, 我们不能得到两幅图像的极几何, 这是因为在极几何中所有的极线均交于极点。然而, 通过(4.3.3)所确定的矩阵 F 一般是满秩的, 因此还必须用一个秩 2 的矩阵 \bar{F} 去逼近矩阵 F 作为基本矩阵的估计, 也就是说将下述最小化问题:

$$\begin{cases} \min \|F - \bar{F}\| \\ \text{subject to } \text{rank}(\bar{F}) = 2 \end{cases} \quad (4.3.4)$$

的解作为基本矩阵的最终估计。问题(4.3.4)的解可通过对 F 进行奇异值分解得到。令 F 的最小奇异值分解为 $F = U \text{diag}(s_1, s_2, s_3) V^T$, ($s_1 \geq s_2 \geq s_3$), 则(4.3.4)的解为 $\bar{F} = U \text{diag}(s_1, s_2, 0) V^T$ 。

总结上述讨论, 有下述估计基本矩阵的 8-点算法:

算法 4.3.1: 估计基本矩阵的 8-点算法

目标: 给定 $N(>=8)$ 个点对应 $m_i \leftrightarrow m'_i$, 计算基本矩阵 F

8-点算法:

- (1) 由 N 个对应点集构造矩阵 A ;
 - (2) 对 A 进行奇异值分解 $A = UDV^T$, 并由 V 的最后一个列向量 v_9 可构造矩阵 F ;
 - (3) 对 F 进行奇异值分解 $F = Udiag(s_1, s_2, s_3)V^T, (s_1 \geq s_2 \geq s_3)$, 得基本矩阵的估计 $\bar{F} = Udiag(s_1, s_2, 0)V^T$ 。
-

4.3.2 最小点对应算法

7-点算法

7-点算法是指仅用 7 个点用来估计基本矩阵的方法。在一般情形下, 7 个点对应能够估计基本矩阵所需要的最少点对应数, 它的重要性在于:

在图像中仅有 7 个点对应时, 只能用这种方法; 在基本矩阵的鲁棒估计中, RANSAC 方法的随机抽样通常需要抽取最小的点对应集。这将在 17.1 节详细介绍。

正如 8-点算法那样, 使用 7 个点对应可得到下述方程:

$$A\mathbf{f} = \begin{pmatrix} u'_1u_1 & u'_1v_1 & u'_1 & v'_1u_1 & v'_1v_1 & v'_1 & u_1 & v_1 & 1 \\ & & & \vdots & & & & & \\ u'_7u_7 & u'_7v_7 & u'_7 & v'_7u_7 & v'_7v_7 & v'_7 & u_7 & v_7 & 1 \end{pmatrix} \mathbf{f} = 0$$

一般地, 系数矩阵 A 的秩为 7, 所以上述方程的解集是 9 维空间中的通过坐标原点的一张 2 维平面。令 $\mathbf{f}_1, \mathbf{f}_2$ 是方程的两个单位正交解, 则它的单位解集合是一个单参数簇 $\mathbf{f} = s\mathbf{f}_1 + (1-s)\mathbf{f}_2$, 于是基本矩阵可以表示为

$$F = sF_1 + (1-s)F_2$$

其中, F_1, F_2 是由 $\mathbf{f}_1, \mathbf{f}_2$ 确定的 3 阶矩阵。由于基本矩阵的秩为 2, 因此获得参数 s 的一个约束方程:

$$\det(sF_1 + (1-s)F_2) = 0$$

这是一个关于参数 s 的 3 次方程, 因此有一个解或有 3 个解(如果有复解则删去, 因为基本矩阵是一个实矩阵)。于是, 我们有下列结论:

命题 4.2.1 使用 7 个点对应, 基本矩阵有一个解或者有 3 个可能的解。

若图像中仅有 7 个对应点, 当基本矩阵有 3 个解时, 不能断定哪一个是真解。如果图像中有多

于7个点对应, 当基本矩阵有3个解时, 可选取匹配点对应数最多的矩阵作为基本矩阵的真解。

6-点算法

如果场景能提供某些可以利用的信息的话, 则在估计基本矩阵时可以减少点对应的个数。例如, 当场景中包含有平面信息时, 仅利用6个点对应, 其中4个是来自同一平面上的点对应, 就可以计算基本矩阵。

首先, 利用平面上的4个点对应计算该平面的单应矩阵 H 。由于基本矩阵可表示为 $F = [e']_x H$, 而极点 e' 仅有两个自由度, 所以根据另外的两个对应点 $m_j \leftrightarrow m'_j, j=1,2$, 通过下述方程:

$$m_j'^T [e']_x H m_j = 0, j=1,2 \quad (4.3.5)$$

可以线性地确定极点 e' 。

值得指出的是, 利用一张平面上的点对应, 无论有多少个都不可能确定基本矩阵。这是因为对于该平面上的任一个点对应 $m \leftrightarrow m'$, 总是有 $m' = Hm$, 于是式(4.3.5)化为恒等式: $m'^T [e']_x H m = m'^T [e']_x m' \equiv 0$, 因而不能构成对极点 e' 的约束。

关于基本矩阵更精细的估计, 请参考第三篇: 模型估计。

4.4 恢复摄像机矩阵

4.4.1 射影相关

设 $\langle P, P' \rangle, \langle \hat{P}, \hat{P}' \rangle$ 为两个摄像机对, 若存在一个可逆的4阶矩阵 M 使得下式成立:

$$\hat{P} = PM, \hat{P}' = P'M \quad (4.4.1)$$

则称 $\langle P, P' \rangle, \langle \hat{P}, \hat{P}' \rangle$ 是两个射影相关的摄像机对, 并记为 $\langle P, P' \rangle \propto \langle \hat{P}, \hat{P}' \rangle$ 。

可逆的4阶矩阵 M 确定一个三维空间的射影变换, 所谓射影相关实际上就是摄像机对之间的射影变换。射影相关是摄像机对的一种等价关系, 即它满足下述性质:

自反性: $\langle P, P' \rangle \propto \langle P, P' \rangle$

对称性: $\langle P, P' \rangle \propto \langle \hat{P}, \hat{P}' \rangle \Rightarrow \langle \hat{P}, \hat{P}' \rangle \propto \langle P, P' \rangle$

传递性: $\langle P_1, P'_1 \rangle \propto \langle P_2, P'_2 \rangle, \langle P_2, P'_2 \rangle \propto \langle P_3, P'_3 \rangle \Rightarrow \langle P_1, P'_1 \rangle \propto \langle P_3, P'_3 \rangle$

因此, 根据射影相关可以将摄像机对分为不同的射影相关类, 同一类的摄像机对都是射影相关的, 不同类的摄像机对是不射影相关的。由于射影相关是一种等价关系, 所以有时又称射影相关为射影等价。

令 $\langle P, P' \rangle$ 是一个摄像机对, 由于摄像机矩阵 P 是一个秩 3 的 3×4 矩阵, 所以由矩阵 P 可以构造一个 4 阶可逆矩阵 $M_p = \begin{pmatrix} P \\ \mathbf{q}^T \end{pmatrix}^{-1}$, 其中 \mathbf{q} 是与 P 的行向量组线性无关的一个向量, 使得 $PM_p = (I, \mathbf{0})$ 。令 $\langle \tilde{P}, \tilde{P}' \rangle = \langle PM_p, P'M_p \rangle$, 则 $\langle P, P' \rangle$ 必与下述形式的摄像机对射影相关:

$$\tilde{P} = (I, \mathbf{0}), \quad \tilde{P}' = (A, \mathbf{b}) \quad (4.4.2)$$

由此, 可以看出 $\langle \tilde{P}, \tilde{P}' \rangle$ 属于 $\langle P, P' \rangle$ 所在的射影相关类。摄像机对 $\langle \tilde{P}, \tilde{P}' \rangle$ 称为该射影相关类的典型摄像机对。

注意: 同一个射影相关类的典型摄像机对不是唯一的。但是, 有下述命题:

引理 4.4.1 $\langle \tilde{P}_1 = (I, \mathbf{0}), \tilde{P}'_1 = (A_1, \mathbf{b}_1) \rangle$ 与 $\langle \tilde{P}_2 = (I, \mathbf{0}), \tilde{P}'_2 = (A_2, \mathbf{b}_2) \rangle$ 为同一射影相关类的两个典型摄像机对的充要条件是存在三维向量 $\mathbf{a} \in R^3$ 使得

$$A_2 = A_1 + \mathbf{b}_1 \mathbf{a}^T, \quad \mathbf{b}_2 = \mathbf{b}_1 \quad (4.4.3)$$

证明 必要性: 由 $\tilde{P}_2 = \tilde{P}_1 M$ 知 $M = \begin{pmatrix} I & \mathbf{0} \\ \mathbf{a}^T & k \end{pmatrix}$, 其中 $k \neq 0$, \mathbf{a} 是一个三维向量。再由 $\tilde{P}'_2 = \tilde{P}'_1 M$ 可以得到式(4.4.3)。充分性: 令 $M = \begin{pmatrix} I & \mathbf{0} \\ \mathbf{a}^T & k \end{pmatrix}$, 由于 $k \neq 0$, 所以 M 为可逆矩阵。显然 $\tilde{P}_2 = \tilde{P}_1 M$, 由式(4.4.3)可得 $\tilde{P}'_2 = \tilde{P}'_1 M$, 所以 $\langle \tilde{P}_1, \tilde{P}'_1 \rangle \propto \langle \tilde{P}_2, \tilde{P}'_2 \rangle$, 因此它们属于同一射影相关类。证毕。

基本矩阵与射影相关类

命题 4.4.2 若 $\langle P, P' \rangle \propto \langle \hat{P}, \hat{P}' \rangle$, 则它们有相同的基本矩阵。

证明 因 $\langle P, P' \rangle, \langle \hat{P}, \hat{P}' \rangle$ 射影相关, 所以存在一个可逆的 4 阶矩阵 M 使 $\hat{P} = PM, \hat{P}' = P'M$ 。令 $\mathbf{m} \leftrightarrow \mathbf{m}'$ 是对应于摄像机对 $\langle P, P' \rangle$ 的一个点对应, 即存在空间点 \mathbf{X} 使得

$$\mathbf{m} = P\mathbf{X}, \quad \mathbf{m}' = P'\mathbf{X}$$

于是,

$$\mathbf{m} = P\mathbf{X} = (PM)\mathbf{X} = \hat{P}\mathbf{X}, \quad \mathbf{m}' = P'\mathbf{X} = (P'M)\mathbf{X} = \hat{P}'\mathbf{X}$$

即 $\mathbf{m} \leftrightarrow \mathbf{m}'$ 也是摄像机对 $\langle \hat{P}, \hat{P}' \rangle$ 的一个点对应。同理可知摄像机对 $\langle \hat{P}, \hat{P}' \rangle$ 的点对应也是摄像机对 $\langle P, P' \rangle$ 的点对应。所以, $\langle P, P' \rangle$ 与 $\langle \hat{P}, \hat{P}' \rangle$ 确定相同的点对应集合, 从而有相同的基本矩阵。证毕。

注 命题 4.4.2 说明每一个射影相关类的摄像机对有相同的基本矩阵。基本矩阵不依赖于世界坐标系的选择。一个可逆的 4 阶矩阵 M 确定一个三维投影变换, 反之一个三维射影变换也可以通过

一个可逆的4阶矩阵 M 来表示。在命题 4.4.2 的证明中, 变换 $\hat{X} = M^{-1}X$ 表示从 $\langle P, P' \rangle$ 世界坐标系到 $\langle \hat{P}, \hat{P}' \rangle$ 世界坐标系之间的变换。因此, 命题 4.4.2 说明, 基本矩阵不依赖于世界坐标系的选择。

例 4.4.1 两个摄像机对 $\langle P, P' \rangle$, $\langle \hat{P}, \hat{P}' \rangle$ 定义如下:

$$P = (I, \mathbf{0}), P' = (K'RK^{-1}, K't); \quad \hat{P} = K(I, \mathbf{0}), \hat{P}' = K'(R, t)$$

令 $M = \begin{pmatrix} K & 0 \\ 0 & 1 \end{pmatrix}$, 则 $\langle \hat{P}, \hat{P}' \rangle = \langle PM, P'M \rangle$, 因此 $\langle P, P' \rangle \propto \langle \hat{P}, \hat{P}' \rangle$, 从而上述两个摄像机对有相同的基本矩阵。矩阵 M 的逆:

$$M^{-1} = \begin{pmatrix} K^{-1} & 0 \\ 0 & 1 \end{pmatrix}$$

是 $\langle P, P' \rangle$ 的世界坐标系到 $\langle \hat{P}, \hat{P}' \rangle$ 的世界坐标系的变换。因 $\langle \hat{P}, \hat{P}' \rangle$ 的世界坐标系是一个欧氏坐标系(第一个摄像机坐标系), 而变换 M 是一个仿射变换, 因此 $\langle P, P' \rangle$ 的世界坐标系是一个仿射坐标系。

命题 4.4.3 令 $\langle \tilde{P} = (I, \mathbf{0}), \tilde{P}' = (A, \mathbf{b}) \rangle$ 是一个典型的摄像机对, 则 A 是一个平面的单应矩阵, $\mathbf{b} = \mathbf{e}'$, 从而这个摄像机的基本矩阵为

$$F = [\mathbf{b}]_{\times} A \quad (4.4.4)$$

证明 记两个摄像机的内参数矩阵分别为 K, K' , 它们之间的相对运动为 (R, t) , 则

$$\langle P = (I, \mathbf{0}), P' = (K'RK^{-1}, K't) \rangle$$

是一个与 $\langle \tilde{P}, \tilde{P}' \rangle$ 射影相关的摄像机对, 由引理 4.4.1,

$$A = K'RK^{-1} + K't\mathbf{a}^T = H_{\infty} + \mathbf{e}'\tilde{\mathbf{a}}^T$$

是一个平面的单应矩阵, 且 $\mathbf{b} = K't = \mathbf{e}'$, 由命题 4.2.3, $F = [\mathbf{b}]_{\times} A$ 。证毕。

命题 4.4.4 若摄像机对 $\langle P, P' \rangle$, $\langle \hat{P}, \hat{P}' \rangle$ 有相同的基本矩阵 F , 则 $\langle P, P' \rangle \propto \langle \hat{P}, \hat{P}' \rangle$ 。

证明 令

$$\langle \tilde{P}_1 = (I, \mathbf{0}), \tilde{P}'_1 = (A_1, \mathbf{b}_1) \rangle \propto \langle P, P' \rangle$$

$$\langle \tilde{P}_2 = (I, \mathbf{0}), \tilde{P}'_2 = (A_2, \mathbf{b}_2) \rangle \propto \langle \hat{P}, \hat{P}' \rangle$$

我们只须证明 $\langle \tilde{P}_1, \tilde{P}'_1 \rangle \propto \langle \tilde{P}_2, \tilde{P}'_2 \rangle$ 。由命题 4.4.3, $F = [\mathbf{b}_1]_{\times} A_1 = [\mathbf{b}_2]_{\times} A_2$, 所以

$$F^T \mathbf{b}_1 = F^T \mathbf{b}_2 = 0$$

由于 $\text{rank}(F)=2$, 所以存在非零常数 k 使得 $\mathbf{b}_2 = k\mathbf{b}_1$ 。于是 $[\mathbf{b}_2]_{\times} A_2 = [\mathbf{b}_2]_{\times} (kA_1)$, 由此得

$$A_2 = kA_1 + \mathbf{b}_1 \mathbf{a}^T = k \left(A_1 + \mathbf{b}_1 \left(\frac{\mathbf{a}}{k} \right)^T \right), \mathbf{a} \in R^3$$

由引理 4.4.1, $\langle \tilde{P}_1, \tilde{P}'_1 \rangle \propto \langle \tilde{P}_2, \tilde{P}'_2 \rangle$ 。证毕。

将命题 4.4.2 与命题 4.4.4 结合起来, 有下述结论:

命题 4.4.5 摄像机对 $\langle P, P' \rangle, \langle \hat{P}, \hat{P}' \rangle$ 有相同的基本矩阵 F , 当且仅当 $\langle P, P' \rangle \propto \langle \hat{P}, \hat{P}' \rangle$ 。

4.4.2 摄像机矩阵的恢复

基本矩阵由两幅图像间的点对应所确定, 两组摄像机对射影相关的充要条件是它们有相同的基本矩阵, 所以, 由两幅图像点对应, 可以在射影的意义下确定两幅图像的摄像机对, 即利用基本矩阵可以恢复射影意义下的摄像机对。在射影意义下恢复摄像机对, 通常也称为两幅图像的射影重构。

给定基本矩阵的一个分解 $F = [\mathbf{e}']_{\times} H$, 构造典型摄像机对 $\tilde{P} = (I, 0), \tilde{P}' = (H, \mathbf{e}')$ 。由命题 4.4.3, 这个摄像机对的基本矩阵必为 F , 再由引理 4.4.1, 与这个摄像机对射影相关的所有典型摄像机对均可表示为

$$\tilde{P} = (I, 0), \tilde{P}'(s, \mathbf{a}) = (H + \mathbf{e}' \mathbf{a}^T, s\mathbf{e}'), \quad s \neq 0, \mathbf{a} \in R^3$$

于是, 从上节的讨论, 我们有下述命题:

命题 4.4.6 给定基本矩阵 F , 则它对应的所有摄像机对为

$$\tilde{P} = (I, 0)M, \tilde{P}'(s, \mathbf{a}) = (H, \mathbf{e}')M \quad (4.4.5)$$

其中 M 为 4 阶可逆矩阵。

5. 自标定理论

摄像机内参数刻画了摄像机的内部结构。为了从图像恢复物体的度量结构，确定摄像机内参数是不可缺少的步骤，因为从图像恢复物体的度量结构与确定摄像机内参数是等价的(见第 6 章)。确定摄像机内参数通常又称为摄像机标定。传统标定方法是使用经过精密加工的标定块，通过标定块上三维坐标已知的点与其图像点间的对应关系计算出摄像机矩阵，然后将摄像机矩阵进行分解得到内参数。所谓自标定，是指仅从场景的多幅图像来确定摄像机内参数，这种方法不需精密加工的标定块。本章主要介绍摄像机自标定技术，在本质上所有自标定技术均来源于射影几何理论，说得更具体些，所有自标定技术理论基础是绝对二次曲线和绝对二次曲面的投影性质。本章除了介绍基于绝对二次曲线与绝对二次曲面的自标定方法外，还介绍了基于正交性、圆环点与平行性的自标定方法，虽然这些方法可以归结为基于绝对二次曲线或绝对二次曲面的方法，但它大多是线性的。

5.1 正交性与摄像机内参数

5.1.1 隐消点与隐消线

隐消点 直线上无穷远点的图像称为该直线的隐消点。由于平行直线与无穷远平面相交于同一个无穷远点，因此平行直线有一个相同的隐消点，即隐消点只与直线的方向有关而与直线的位置无关。令 $D = (d^T, 0)^T$ 是直线 L 的隐消点，则向量 d 是该直线的方向，有时也用 D 来表示直线的方向。因此，隐消点通常也说成是直线方向的隐消点。设摄像机矩阵为 $P = (H_\infty, p_4)$ ，其中 H_∞ 是无穷远单应矩阵，于是直线 L 的隐消点为 $v = H_\infty d$ 。

记两条直线 L_1, L_2 的方向分别为 d_1, d_2 ，由欧氏几何可知它们之间的夹角可通过下述公式来计算：

$$\cos \theta = \frac{d_1^T d_2}{\sqrt{d_1^T d_1} \cdot \sqrt{d_2^T d_2}} \quad (5.1.1)$$

若直线 L_1, L_2 的隐消点分别为 v_1, v_2 , 则根据式(3.2.9), 得到

$$d_1 = H_{\infty}^{-1}v_1 = R^{-1}K^{-1}v_1, d_2 = H_{\infty}^{-1}v_2 = R^{-1}K^{-1}v_2$$

于是,

$$\cos \theta = \frac{v_1^T (H_{\infty}^{-T} H_{\infty}^{-1}) v_2}{\sqrt{v_1^T (H_{\infty}^{-T} H_{\infty}^{-1}) v_1} \cdot \sqrt{v_2^T (H_{\infty}^{-T} H_{\infty}^{-1}) v_2}} = \frac{v_1^T \varpi v_2}{\sqrt{v_1^T \varpi v_1} \cdot \sqrt{v_2^T \varpi v_2}} \quad (5.1.2)$$

其中 $\varpi = K^{-T}K^{-1}$ 是绝对二次曲线的图像 (IAC)。

如果已知两条直线的夹角和它们的隐消点, 则式(5.1.2)就构成摄像机内参数的约束, 从而可被用于标定摄像机内参数。

隐消线 平面 π 上的无穷远直线 L_{∞} 在像平面上的投影称为该平面的隐消线。平行平面相交于无穷远平面上的同一条直线, 因而平行平面有相同的隐消线。隐消线只与平面的法向量(或称为平面的方向)有关而与平面的位置无关。在几何上, 平面 π 的隐消线 l 是与平面 π 平行且通过摄像机光心的平面与像平面的交线。下面给出隐消线的代数表达。

令 L_1, L_2 是平面 π 上的相交于有限点的两条直线, 其无穷远点分别记为

$$D_1 = (d_1^T, 0)^T, D_2 = (d_2^T, 0)^T$$

则平面 π 上的无穷远直线必通过 D_1, D_2 , 并且平面 π 的方向是 $n = d_1 \times d_2$ 。令 l 是平面 π 的隐消线, 则直线 L_1, L_2 的隐消点 v_1, v_2 是 l 上两个不同的点。于是,

$$l = v_1 \times v_2 = H_{\infty} d_1 \times H_{\infty} d_2 = H_{\infty}^{-T} (d_1 \times d_2) = H_{\infty}^{-T} n = K^{-T} R n \quad (5.1.3)$$

其中: $n = d_1 \times d_2$ 是平面 π 的法方向。

令平面 π_1, π_2 的方向分别为 n_1, n_2 , 则它们之间的夹角可表示为

$$\cos \theta = \frac{n_1^T n_2}{\sqrt{n_1^T n_1} \cdot \sqrt{n_2^T n_2}} \quad (5.1.4)$$

如果 l_1, l_2 分别为平面 π_1, π_2 的隐消线, 则利用式(5.1.3)和(5.1.4), 就可得到下述平面夹角表达式:

$$\cos \theta = \frac{l_1^T \varpi^* l_2}{\sqrt{l_1^T \varpi^* l_1} \cdot \sqrt{l_2^T \varpi^* l_2}} \quad (5.1.5)$$

其中 $\varpi^* = K K^T$ 是绝对二次曲面的图像, 它是绝对二次曲线图像 ϖ 的对偶(DIAC)。

同样, 在已知平面夹角和隐消线的情况下, 式(5.1.5)也可以被用于确定摄像机内参数。

5.1.2 正交性与摄像机内参数

两条正交直线的隐消点, 称为一对正交隐消点(注意: 这里的正交是指隐消点对应的空间直线相互正交, 并非指表示两个隐消点的向量相互正交)。类似地, 也有正交隐消线的概念。在一般情况下式(5.1.2)与式(5.1.5)是摄像机的 IAC, DIAC 的二次约束, 对于正交隐消点或正交隐消线, 可以建立摄像机的 IAC 或 DIAC 的线性约束方程, 从而能线性标定摄像机内参数。

命题 5.1.1 (1)假定 $\mathbf{v}_1, \mathbf{v}_2$ 是相互正交的隐消点, 则 $\mathbf{v}_2^T \boldsymbol{\omega} \mathbf{v}_1 = 0$; (2)假定 l_1, l_2 为相互正交的隐消线, 则 $l_2^T \boldsymbol{\omega}^* l_1 = 0$; (3)假定隐消点 \mathbf{v} 与隐消线 l 是相互正交的, 则 $l = \boldsymbol{\omega} \mathbf{v}$, 或等价地, $\mathbf{v} = \boldsymbol{\omega}^* l$ 。

注意: 上述等式都是齐次等式, 即对于给定点(或直线)的齐次坐标, 上述等式在相差一个非零常数因子意义下成立。

结论(1)与结论(2)直接由公式(5.1.2)和(5.1.5)得到。下面给出结论(3)的证明: 由于隐消点 \mathbf{v} 与隐消线 l 是相互正交的, 所以对任意的 $\mathbf{v}_1, \mathbf{v}_2 \in l$, $\{\mathbf{v}, \mathbf{v}_1\}$, $\{\mathbf{v}, \mathbf{v}_2\}$ 是两对正交隐消点。于是, 从式(5.1.2), 可知:

$$\mathbf{v}^T \boldsymbol{\omega} \mathbf{v}_1 = 0, \quad \mathbf{v}^T \boldsymbol{\omega} \mathbf{v}_2 = 0$$

从而, 可推知

$$\mathbf{v} = \boldsymbol{\omega} \mathbf{v}_1 \times \boldsymbol{\omega} \mathbf{v}_2 = \boldsymbol{\omega}^* (\mathbf{v}_1 \times \mathbf{v}_2) = \boldsymbol{\omega}^* l$$

证毕。

上述结论表明, 如果从图像中能获得 5 组以上的正交隐消点或隐消线, 就能得到 5 个以上的关于 IAC 或 DIAC 的线性约束方程, 由于 IAC 或 DIAC 只含有 5 个独立的参数, 所以将这些方程联立起来就可以解出 IAC 或 DIAC, 从而再通过 IAC 或 DIAC 的 Cholesky 分解^{*)}就可以确定摄像机的内参数矩阵。例如通过 Cholesky 分解将 $\boldsymbol{\omega}$ 分解成 $\boldsymbol{\omega} = \tilde{K}^{-T} \tilde{K}^{-1}$ (其中 \tilde{K} 是对角元均大于零的上三角矩阵)。此时, \tilde{K} 与摄像机内参数 K 相差一个常数因子。由于 K 的最后一个元素是 1, 所以摄像机内参数必为 $K = \tilde{k}_{33}^{-1} \tilde{K}$ 。

三正交隐消点 如果 $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ 是两两相互正交的隐消点, 则称它们为三正交隐消点。对于三正交隐消点, 在形式上可提供下述三组关于内参数的约束方程:

$$x \mathbf{v}_1 \times \mathbf{v}_2 = \boldsymbol{\omega} \mathbf{v}_3, \quad y \mathbf{v}_1 \times \mathbf{v}_3 = \boldsymbol{\omega} \mathbf{v}_2, \quad z \mathbf{v}_2 \times \mathbf{v}_3 = \boldsymbol{\omega} \mathbf{v}_1$$

这里, 显式地写出了常数因子 x, y, z , 它们都是未知的非零常数。因此, 在上述每一组中消去常数因子后可以得到二个关于内参数的独立约束方程, 但下述命题表明: 在它们联立的 6 个约束方程中仅有三个是独立的。

^{*)}见 8.2 节

命题 5.1.2 令 $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ 是三正交方向的隐消点, 则摄像机的 IAC 和 DIAC 的解分别为

$$\boldsymbol{\omega} = (\mathbf{v}_1 \times \mathbf{v}_2)(\mathbf{v}_1 \times \mathbf{v}_2)^T + s_1(\mathbf{v}_1 \times \mathbf{v}_3)(\mathbf{v}_1 \times \mathbf{v}_3)^T + s_2(\mathbf{v}_2 \times \mathbf{v}_3)(\mathbf{v}_2 \times \mathbf{v}_3)^T \quad (5.1.6)$$

$$\boldsymbol{\omega}^* = \mathbf{v}_1 \mathbf{v}_1^T + s_1^* \mathbf{v}_2 \mathbf{v}_2^T + s_2^* \mathbf{v}_3 \mathbf{v}_3^T \quad (5.1.7)$$

其中 s_j 是未知常数。

证明 由于 $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ 是三正交隐消点, 所以存在未知常数 x, y, z 使得

$$x\mathbf{v}_1 \times \mathbf{v}_2 = \boldsymbol{\omega} \mathbf{v}_3, \quad y\mathbf{v}_1 \times \mathbf{v}_3 = \boldsymbol{\omega} \mathbf{v}_2, \quad z\mathbf{v}_2 \times \mathbf{v}_3 = \boldsymbol{\omega} \mathbf{v}_1$$

因此,

$$\boldsymbol{\omega}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = (z\mathbf{v}_2 \times \mathbf{v}_3, y\mathbf{v}_1 \times \mathbf{v}_3, x\mathbf{v}_1 \times \mathbf{v}_2)$$

从 $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ 不共线, 可得到

$$\boldsymbol{\omega} = (z\mathbf{v}_2 \times \mathbf{v}_3 \quad y\mathbf{v}_1 \times \mathbf{v}_3 \quad x\mathbf{v}_1 \times \mathbf{v}_2)(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)^{-1}$$

所以,

$$\begin{aligned} \boldsymbol{\omega} &= \frac{1}{\det(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)} (z\mathbf{v}_2 \times \mathbf{v}_3, y\mathbf{v}_1 \times \mathbf{v}_3, x\mathbf{v}_1 \times \mathbf{v}_2) \begin{pmatrix} (\mathbf{v}_2 \times \mathbf{v}_3)^T \\ (\mathbf{v}_3 \times \mathbf{v}_1)^T \\ (\mathbf{v}_1 \times \mathbf{v}_2)^T \end{pmatrix} \\ &= \frac{z(\mathbf{v}_2 \times \mathbf{v}_3)(\mathbf{v}_2 \times \mathbf{v}_3)^T + y(\mathbf{v}_1 \times \mathbf{v}_3)(\mathbf{v}_1 \times \mathbf{v}_3)^T + x(\mathbf{v}_1 \times \mathbf{v}_2)(\mathbf{v}_1 \times \mathbf{v}_2)^T}{\det(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)} \end{aligned}$$

于是, 在相差一个常数因子的意义下, 摄像机的 IAC 可以表示为

$$\boldsymbol{\omega} = (\mathbf{v}_1 \times \mathbf{v}_2)(\mathbf{v}_1 \times \mathbf{v}_2)^T + s_1(\mathbf{v}_1 \times \mathbf{v}_3)(\mathbf{v}_1 \times \mathbf{v}_3)^T + s_2(\mathbf{v}_2 \times \mathbf{v}_3)(\mathbf{v}_2 \times \mathbf{v}_3)^T$$

同理可证摄像机的 DIAC 为 $\boldsymbol{\omega}^* = \mathbf{v}_1 \mathbf{v}_1^T + s_1^* \mathbf{v}_2 \mathbf{v}_2^T + s_2^* \mathbf{v}_3 \mathbf{v}_3^T$ 。证毕。

命题 5.1.2 不但给出了 IAC 和 DIAC 的计算公式, 同时也说明从单幅图像上三正交隐消点只能提供关于内参数的三个独立约束, 尽管在形式上提供了 6 个方程。因此, 需要两组三正交隐消点, 才能完全确定摄像机内参数。当摄像机的畸变因子为零, 且像素的纵横比为 1 时, 可根据这些约束完全确定式(5.1.6)和(5.1.7)中的未知常数, 即此时从单幅图像上的三正交隐消点可以标定摄像机。值得指出的是, 如果像素的纵横比等于 1, 即摄像机的两个尺度因子相等, 则摄像机的主点是以 $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ 为顶点的三角形三条高线的交点。

例 5.1.1 利用两个长方体的图像确定摄像机内参数。

利用长方体平行棱相交于同一个无穷远点的性质, 可以从图像上得到三正交隐消点 $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ 。根据命题 5.1.2, 可得到 $\boldsymbol{\omega}$ 的双参数簇:

$$\boldsymbol{\omega} = (\mathbf{v}_1 \times \mathbf{v}_2)(\mathbf{v}_1 \times \mathbf{v}_2)^T + s_1(\mathbf{v}_1 \times \mathbf{v}_3)(\mathbf{v}_1 \times \mathbf{v}_3)^T + s_2(\mathbf{v}_2 \times \mathbf{v}_3)(\mathbf{v}_2 \times \mathbf{v}_3)^T$$

如果另一个长方体不与当前的长方体平行，则根据它的图像可确定另一组三正交隐消点，并得到 ω 的另一组线性约束方程，从而能唯一确定常数 s_1, s_2 。这样，就确定了 ω 。

例 5.1.2 从摄像机的正交纯平移运动确定内参数。当摄像机作纯平移运动（在平移运动中，内参数保持不变），对某个静止物体的两幅图像可视为物体作相对平移运动而摄像机在固定位置下所得到的两幅图像，如图 4.1.4 所示。平移运动物体上的任何一对对应点的连线必与平移方向平行，因此所有对应点的连线在空间中相交于同一（无穷远）点，该点表示平移方向。因此两幅图像对应点的连线也必交于一点，该点是平移方向在图像平面上的投影，由极几何可以知道它是两幅图像的极点。于是，当摄像机作纯平移运动时，两幅图像的极点相同并且是平移方向在图像平面上的投影。

令 $\{t_1, t_2\}$ 是摄像机一组正交纯平移运动，即平移运动参数满足 $t_1^T t_2 = 0$ 。令 e_1, e_2 分别为作运动 $\{t_1, t_2\}$ 后所拍摄的图像关于初始图像的极点，它们是一对正交隐消点，于是可得到内参数的约束方程 $e_1^T \omega e_2 = 0$ 。这样，从 5 组正交纯平移运动就可以求解摄像机内参数。

5.2 圆环点与摄像机内参数

5.2.1 内参数的约束方程

平面 π 上的两个圆环点 I, J 的坐标可以分别为 $I = (1, i, 0)^T, J = (1, -i, 0)^T$ 。由于平面 π 的两个圆环点 I, J 都在绝对二次曲线(AC)上，所以它们的像 m_I, m_J 必在 IAC 上。因此，可得到下述 IAC 的约束方程：

$$m_I^T \omega m_I = 0, m_J^T \omega m_J = 0 \quad (5.2.1)$$

由于 m_I, m_J 是一对共轭点，在上述两个方程中，实际上只能提供下述两个关于 IAC 的实线性约束方程：

$$\operatorname{Re}(m_I^T \omega m_I) = 0, \operatorname{Im}(m_I^T \omega m_I) = 0 \quad (5.2.2)$$

其中：Re, Im 分别表示复数的实部和虚部。

如果能从图像上获得两两不平行的三个平面上的圆环点图像，则可以得到 IAC 上的 6 个点。由于平面上 5 个点确定一条二次曲线，因此这 6 个点足以拟合出 IAC，从而可以确定摄像机内参数。这里要求每两个平面不平行，是因为平行平面有相同的圆环点。

5.2.2 确定圆环点的图像

基于圆环点的标定方法的关键问题是如何在图像平面上确定圆环点的图像。由于圆环点是虚点

从而无法直接从图像获取, 因此需要借助图像中所包含的相关场景结构信息才能得到圆环点图像。

正方形 令正方形的 4 个顶点坐标分别为:

$$\mathbf{X}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \mathbf{X}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \mathbf{X}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \mathbf{X}_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

应用这些顶点的图像坐标, 可以求出正方形所在平面到像平面的单应矩阵 H 。于是, 该平面圆环点的像坐标为:

$$\mathbf{m}_{1,j} = H \begin{pmatrix} 1 \\ \pm i \\ 0 \end{pmatrix} = \mathbf{h}_1 \pm i \mathbf{h}_2$$

两个圆 两个相交的圆有两个实交点与两个虚交点, 显然这两个虚交点即为圆环点, 因为平面上任何一个圆都相交于圆环点。于是两条图像曲线(二次曲线)的两个虚交点即为圆环点在图像平面上的投影。已知一个圆及其中心的图像, 也可以确定圆环点的图像。

圆与直线 两条通过圆心直线分别交无穷远直线于 P 、 P' 两点, 则交比 $(AB; OP) = -1$, $(A'B'; OP') = -1$ 。其中 $\{A, B\}$, $\{A', B'\}$ 分别是两条直线与圆的两个交点, O 是圆心。由于投影变换保持交比不变, 所以可求出 P 和 P' 的图像点坐标, 从而可求得圆支撑平面的隐消线方程。圆的图像是一条二次曲线, 并且可以用曲线拟合的方法, 求出它在像平面上的方程。由隐消线方程与这个二次曲线方程联立求解, 就得到圆环点的图像坐标。当然, 也可以从两条像直线的交点获得圆心的图像, 从而由配极对应得到隐消线。

在包含圆与直线的场景中, 确定圆环点图像的条件是通过圆的图像曲线和直线的图像直线能够确定平面的消影线, 或能够确定圆心的图像。

平面相似图形 设 A, B 是同一张平面 π 上的两个相似图形, 记 $X = \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_4, \dots\}$, $Y = \{\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_4, \dots\}$ 是一组相似的点对应, 于是存在二维相似变换 S 使得 $\mathbf{Y}_j = S\mathbf{X}_j$ 。令 X, Y 的图像点分别为 $\mathfrak{M}_x = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_4, \dots\}$, $\mathfrak{M}_y = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_4, \dots\}$, 则存在二维射影变换 H 使得 $\mathbf{x}_j = H\mathbf{X}_j$, $\mathbf{y}_j = H\mathbf{Y}_j$ 。于是, 我们有 $\mathbf{y}_j = HSH^{-1}\mathbf{x}_j$, 即 $\mathfrak{M}_x \leftrightarrow \mathfrak{M}_y$ 是一组射影对应。记 $H_p = HSH^{-1}$, 当 $\mathfrak{M}_x \leftrightarrow \mathfrak{M}_y$ 中包含 4 个以上的点对应时, 可以唯一确定射影变换 H_p 。在一般情况下, H_p 的两个互为共轭复特征向量是平面 π 上的两个圆环点的图像。事实上, 由于平面 π 上任一点 X 的图像必为 $\mathbf{x} = HX$, 所以平面 π 的一个圆环点 J 的像是 $\mathbf{m}_j = HJ$ 。因相似变换 S 保持圆环点不变, 即 $SJ = J$, 所以有

$$H_p m_J = H_p HJ = HSJ = HJ = m_J$$

又因, 圆环点 J 是虚点, 并且 H_p 是一个实矩阵, 所以 m_J 是 H_p 的复特征向量。因此 H_p 的两个互为共轭复特征向量是平面 π 上的两个圆环点的图像。值得指出的是, 两个相似图形之间必须有平面旋转运动并且旋转角不是 $\pi/2$ 的整数倍, 否则这个结论不成立。

5.2.3 圆环点与其正交方向

假定平面 π 上的两个圆环点 $\{I, J\}$ 的图像分别为 m_I, \bar{m}_I , 且平面 π 方向 D 的隐消点为 v 。由于通过 I, J 的无穷远直线的图像是平面 π 的隐消线, 所以根据命题 5.1.1 (3), 我们有

$$(m_I \times \bar{m}_I) = x\omega v$$

平面 π 的方向 D 关于绝对二次曲线的极线是通过 I, J 的无穷远直线, 而 I, J 在绝对二次曲线上, 因此通过 $\{D, I\}$ 与 $\{D, J\}$ 的两条直线是绝对二次曲线的两条切线, 所以,

$$(m_I \times v) = y\omega m_I$$

$$(\bar{m}_I \times v) = \bar{y}\omega \bar{m}_I$$

其中: \bar{X} 表示 X 的共轭。这样, 就得到 IAC 的线性约束方程:

$$\begin{cases} (m_I \times \bar{m}_I) = x\omega v \\ (m_I \times v) = y\omega m_I \\ (\bar{m}_I \times v) = \bar{y}\omega \bar{m}_I \end{cases} \quad (5.2.3)$$

上述方程的对偶是

$$\begin{cases} \omega^* (m_I \times \bar{m}_I) = xv \\ \omega^* (m_I \times v) = ym_I \\ \omega^* (\bar{m}_I \times v) = \bar{y}\bar{m}_I \end{cases} \quad (5.2.4)$$

命题 5.2.2 方程(5.2.3)和(5.2.4)的解分别为

$$\omega = 2\lambda(a \times b)(a \times b)^T + (a \times v)(a \times v)^T + (b \times v)(b \times v)^T \quad (5.2.5)$$

$$\omega^* = 2\lambda(aa^T + bb^T) + vv^T \quad (5.2.6)$$

其中 λ 为一个非零实数, a, b 分别为 m_I 的实部和虚部。

证明 仅证明(5.2.6), 因为用类似的方法可得到(5.2.5)。由

$$\begin{cases} \omega^* (m_I \times \bar{m}_I) = xv \\ \omega^* (m_I \times v) = ym_I \\ \omega^* (\bar{m}_I \times v) = \bar{y}\bar{m}_I \end{cases}$$

可推知:

$$\begin{aligned}
\varpi^* &= (xv, ym_1, \overline{ym_1})(m_1 \times \overline{m_1}, m_1 \times v, \overline{m_1} \times v)^{-1} \\
&= (xv, ym_1, \overline{ym_1}) \cdot \frac{1}{\det(v, m_1, \overline{m_1})} \begin{pmatrix} v^T \\ -\overline{m_1}^T \\ m_1^T \end{pmatrix} \\
&= \frac{1}{\det(v, m_1, \overline{m_1})} (xvv^T - ym_1\overline{m_1}^T + \overline{ym_1}m_1^T) \quad (y = y_1 + iy_2) \\
&= xvv^T - (y_1 + iy_2)m_1\overline{m_1}^T + (y_1 - iy_2)\overline{m_1}m_1^T \quad (\text{齐次等式, 即相差一个常数意义下相等}) \\
&= xvv^T - y_1(m_1\overline{m_1}^T - \overline{m_1}m_1^T) - iy_2(m_1\overline{m_1}^T + \overline{m_1}m_1^T) \\
&= xvv^T - iy_2(m_1\overline{m_1}^T + \overline{m_1}m_1^T) \quad (\text{因为 } \varpi^* \text{ 对称, 可知 } y_1 = 0) \\
&= vv^T + \lambda(m_1\overline{m_1}^T + \overline{m_1}m_1^T) \quad (\text{齐次等式. } \lambda \text{ 为一个非零实数, 因为 } x \text{ 是个纯虚数}) \\
&= vv^T + 2\lambda(aa^T + bb^T)
\end{aligned}$$

命题 5.2.2 表明: 通过一个平面上的圆环点与其正交方向, 可以提供关于摄像机内参数的四个独立约束。如果摄像机使用 4 参数模型, 例如畸变因子为零, 则根据命题 5.2.2, 可以确定摄像机内参数矩阵 K 。因为, 此时 IAC 中的元素 $\varpi_{12} = 0$, 所以由这个约束可以确定(5.2.6)式中的未知常数 λ 。

例 8.2.1 从圆柱的图像标定内参数。假定摄像机畸变因子为零, 则从一个圆柱的图像就足以确定内参数。从圆柱上、下面的两个圆的图像, 可确定两个圆环点的图像, 而与两个圆环点正交方向的隐消点是两条母线的图像的交点, 因此可从圆柱图像的轮廓线中获取。

5.3 平行性与摄像机内参数

5.3.1 平行四边形的不变量与射影

平行四边形的不变量

下面讨论平行四边形的不变量以及射影性质。尽管引进的不变量是简单的, 但是从它们可以得到一些有趣性质和重要结果。假定 $\{\tilde{X}_i : i = 1, 2, 3, 4\}$ 为空间中平行四边形的四顶点的欧氏坐标, 如图 5.1.1 所示, 并且约定 $\overline{X_1X_2} = \overline{X_3X_4}$ 。令

$$t = \frac{\|\tilde{X}_3 - \tilde{X}_1\|}{\|\tilde{X}_2 - \tilde{X}_1\|} \quad (5.3.1)$$

$$\cos \theta = \frac{(\tilde{\mathbf{X}}_3 - \tilde{\mathbf{X}}_1)^T (\tilde{\mathbf{X}}_2 - \tilde{\mathbf{X}}_1)}{\|\tilde{\mathbf{X}}_3 - \tilde{\mathbf{X}}_1\| \cdot \|\tilde{\mathbf{X}}_2 - \tilde{\mathbf{X}}_1\|} \quad (5.3.2)$$

由于相似变换保持线段长度和夹角不变,因此参数 $\{t, \theta\}$ 均为平行四边形的相似不变量。进一步,参数 $\{t, \theta\}$ 确定了平行四边形的形状,即如果两个平行四边形有相同的参数 $\{t, \theta\}$,则它们一定是相似的。我们再给出平行四边形的一个仿射不变量,它在确定摄像机内参数中具有极其重要的作用。

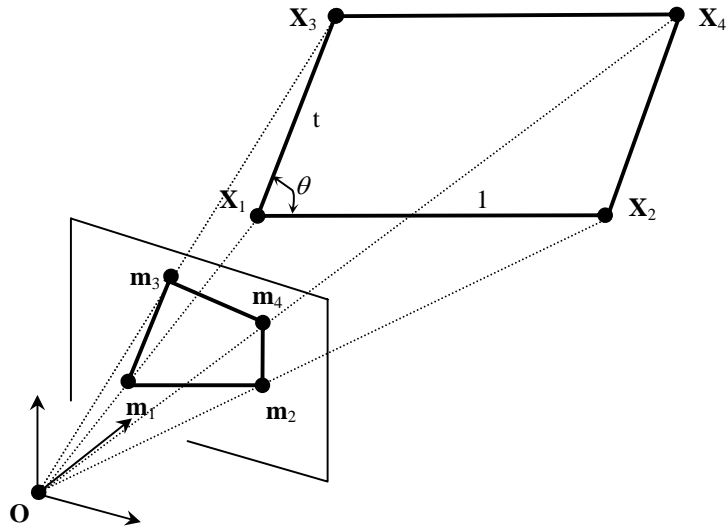


图 5.1.1: 平行四边形和它的图像

从 $\overline{\mathbf{X}_1 \mathbf{X}_2} = \overline{\mathbf{X}_3 \mathbf{X}_4}$ 知,

$$\tilde{\mathbf{X}}_4 - \tilde{\mathbf{X}}_3 = (\tilde{\mathbf{X}}_2 - \tilde{\mathbf{X}}_1)$$

因此,

$$\tilde{\mathbf{X}}_4 = \tilde{\mathbf{X}}_2 - \tilde{\mathbf{X}}_1 + \tilde{\mathbf{X}}_3 = (\tilde{\mathbf{X}}_1, \tilde{\mathbf{X}}_2, \tilde{\mathbf{X}}_3)(-1, 1, 1)^T$$

记 $X = (\tilde{\mathbf{X}}_1, \tilde{\mathbf{X}}_2, \tilde{\mathbf{X}}_3)$, 则有

$$X^{-1} \tilde{\mathbf{X}}_4 = (-1, 1, 1)^T \quad (5.3.4)$$

因仿射变换保持平行性且保持平行线段的长度比,所以式(5.3.4)是平行四边形的仿射不变量。事实上,也可以给出如下的直接的证明:令 $\tilde{\mathbf{X}}' = A\tilde{\mathbf{X}} + t$ 是任一三维仿射变换,则

$$\begin{aligned} \tilde{\mathbf{X}}'_4 &= A\tilde{\mathbf{X}}_4 + t = (-A\tilde{\mathbf{X}}_1 + A\tilde{\mathbf{X}}_2 + A\tilde{\mathbf{X}}_3) + t \\ &= (-\tilde{\mathbf{X}}'_1 + t + \tilde{\mathbf{X}}'_2 - t + \tilde{\mathbf{X}}'_3 - t) + t \end{aligned}$$

$$= -\tilde{\mathbf{X}}_1' + \tilde{\mathbf{X}}_2' + \tilde{\mathbf{X}}_3'$$

因此, $X^{-1}\tilde{\mathbf{X}}_4' = (-1, 1, 1)^T$ 。所以, 式(5.3.4)是平行四边形的仿射不变量。

式(5.3.4)描述了平行四边形的基本属性, 这是因为有下述命题:

命题 5.3.1 四边形 $\{\tilde{\mathbf{X}}_i : i = 1, 2, 3, 4\}$ 为平行四边形的充要条件为(5.3.4)成立。

平行四边形的射影

利用仿射不变量(5.3.4), 极易导入平行四边形的射影、及其与相似不变量 $\{t, \theta\}$ 之间的关系。

命题 5.3.2 假定 $\{\tilde{\mathbf{X}}_i : i = 1, 2, 3, 4\}$ 为平行四边形的四个顶点, 其图像点为 $\{\mathbf{m}_i : i = 1, 2, 3, 4\}$, 则

(1) 令 $(q_1, q_2, q_3)^T = (-\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3)^{-1}\mathbf{m}_4$, 在摄像机坐标系下, 有

$$\tilde{\mathbf{X}}_j = \alpha_4 q_j K^{-1} \mathbf{m}_j, j = 1, 2, 3, 4 \quad *) \quad (5.3.5)$$

(2) 令 $L = (q_2 \mathbf{m}_2 - q_1 \mathbf{m}_1, q_3 \mathbf{m}_3 - q_1 \mathbf{m}_1)$, 则相似不变量与摄像机内参数之间有下列关系:

$$(\|\tilde{\mathbf{X}}_2 - \tilde{\mathbf{X}}_1\|^2 / \alpha_4^2) \eta = L^T \varpi L \quad (5.3.6)$$

其中: α_4 是空间点 $\tilde{\mathbf{X}}_4$ 的深度, $\eta = \begin{pmatrix} 1 & t \cos \theta \\ t \cos \theta & t^2 \end{pmatrix}$ 称为平行四边形的相似参数矩阵, $\varpi = K^{-T} K^{-1}$ 是绝对二次曲线的图像。

证明: 在摄像机坐标系下, 有

$$\alpha_j \mathbf{m}_j = K \tilde{\mathbf{X}}_j, j = 1, 2, 3, 4$$

即,

$$\tilde{\mathbf{X}}_j = \alpha_j K^{-1} \mathbf{m}_j, j = 1, 2, 3, 4 \quad (5.3.7)$$

所以,

$$X = (\tilde{\mathbf{X}}_1, \tilde{\mathbf{X}}_2, \tilde{\mathbf{X}}_3) = K^{-1} (\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3) \text{diag}(\alpha_1, \alpha_2, \alpha_3)$$

于是,

$$X^{-1} = \text{diag}(1/\alpha_1, 1/\alpha_2, 1/\alpha_3) (\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3)^{-1} K$$

因此,

$$\begin{aligned} X^{-1} \tilde{\mathbf{X}}_4 &= \text{diag}(1/\alpha_1, 1/\alpha_2, 1/\alpha_3) (\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3)^{-1} K (\alpha_4 K^{-1} \mathbf{m}_4) \\ &= \text{diag}(\alpha_4 / \alpha_1, \alpha_4 / \alpha_2, \alpha_4 / \alpha_3) (\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3)^{-1} \mathbf{m}_4 \\ &= \text{diag}(\alpha_4 / \alpha_1, \alpha_4 / \alpha_2, \alpha_4 / \alpha_3) (-q_1, q_2, q_3)^T \\ &= (-\alpha_4 q_1 / \alpha_1, \alpha_4 q_2 / \alpha_2, \alpha_4 q_3 / \alpha_3)^T \end{aligned}$$

*) 约定 $q_4 = 1$

根据式(5.3.4), 我们有

$$\alpha_j = \alpha_4 q_j, j=1,2,3 \quad (5.3.8)$$

代入式(5.3.7), 得到式(5.3.5)。从式(5.3.5), 有

$$\tilde{\mathbf{X}}_2 - \tilde{\mathbf{X}}_1 = \alpha_4 K^{-1} (q_2 \mathbf{m}_2 - q_1 \mathbf{m}_1) \quad (5.3.9)$$

$$\tilde{\mathbf{X}}_3 - \tilde{\mathbf{X}}_1 = \alpha_4 K^{-1} (q_3 \mathbf{m}_3 - q_1 \mathbf{m}_1) \quad (5.3.10)$$

由此可推知

$$\|\tilde{\mathbf{X}}_2 - \tilde{\mathbf{X}}_1\|^2 \eta = \begin{pmatrix} (\tilde{\mathbf{X}}_2 - \tilde{\mathbf{X}}_1)^T (\tilde{\mathbf{X}}_2 - \tilde{\mathbf{X}}_1) & (\tilde{\mathbf{X}}_2 - \tilde{\mathbf{X}}_1)^T (\tilde{\mathbf{X}}_3 - \tilde{\mathbf{X}}_1) \\ (\tilde{\mathbf{X}}_2 - \tilde{\mathbf{X}}_1)^T (\tilde{\mathbf{X}}_3 - \tilde{\mathbf{X}}_1) & (\tilde{\mathbf{X}}_3 - \tilde{\mathbf{X}}_1)^T (\tilde{\mathbf{X}}_3 - \tilde{\mathbf{X}}_1) \end{pmatrix} = \alpha_4^2 L^T \varpi L$$

因此, 式(5.3.6)成立。

注 由于 $\tilde{\mathbf{X}}_4 - \tilde{\mathbf{X}}_3 = \tilde{\mathbf{X}}_2 - \tilde{\mathbf{X}}_1$ 和 $\tilde{\mathbf{X}}_4 - \tilde{\mathbf{X}}_2 = \tilde{\mathbf{X}}_3 - \tilde{\mathbf{X}}_1$, 所以从 (5.3.5), 可得到

$$q_2 \mathbf{m}_2 - q_1 \mathbf{m}_1 = q_4 \mathbf{m}_4 - q_3 \mathbf{m}_3 \triangleq \mathbf{v}_1, \quad q_3 \mathbf{m}_3 - q_1 \mathbf{m}_1 = q_4 \mathbf{m}_4 - q_2 \mathbf{m}_2 \triangleq \mathbf{v}_2.$$

图像点 $\mathbf{v}_1(\mathbf{v}_2)$ 是平行边 $\mathbf{X}_1 \mathbf{X}_2 // \mathbf{X}_3 \mathbf{X}_4$ ($\mathbf{X}_1 \mathbf{X}_3 // \mathbf{X}_2 \mathbf{X}_4$) 的隐消点, 这是因为

$$\mathbf{v}_1^T (\mathbf{m}_1 \times \mathbf{m}_2) = \mathbf{v}_1^T (\mathbf{m}_3 \times \mathbf{m}_4) = 0, \quad \mathbf{v}_2^T (\mathbf{m}_1 \times \mathbf{m}_3) = \mathbf{v}_2^T (\mathbf{m}_2 \times \mathbf{m}_4) = 0$$

在摄像机内参数已知的情况下, 有下述推论:

推论 5.3.1: 从平行四边形 $\{\tilde{\mathbf{X}}_i : i=1,2,3,4\}$ 的图像点 $\{\mathbf{m}_i : i=1,2,3,4\}$ 能恢复平行四边形的形状, 即可以确定相似不变量 $\{t, \theta\}$ 。

5.3.2 平行六面体的不变量及其射影

考虑平行六面体的射影, 如图 5.3.2 所示。令

$$t_1 = \frac{\|\tilde{\mathbf{X}}_3 - \tilde{\mathbf{X}}_1\|}{\|\tilde{\mathbf{X}}_2 - \tilde{\mathbf{X}}_1\|}, \quad t_2 = \frac{\|\tilde{\mathbf{X}}_5 - \tilde{\mathbf{X}}_1\|}{\|\tilde{\mathbf{X}}_2 - \tilde{\mathbf{X}}_1\|}, \quad (5.3.11)$$

$$\cos \theta = \frac{(\tilde{\mathbf{X}}_3 - \tilde{\mathbf{X}}_1)^T (\tilde{\mathbf{X}}_2 - \tilde{\mathbf{X}}_1)}{\|\tilde{\mathbf{X}}_3 - \tilde{\mathbf{X}}_1\| \cdot \|\tilde{\mathbf{X}}_2 - \tilde{\mathbf{X}}_1\|}, \quad \cos \phi = \frac{(\tilde{\mathbf{X}}_5 - \tilde{\mathbf{X}}_1)^T (\tilde{\mathbf{X}}_2 - \tilde{\mathbf{X}}_1)}{\|\tilde{\mathbf{X}}_5 - \tilde{\mathbf{X}}_1\| \cdot \|\tilde{\mathbf{X}}_2 - \tilde{\mathbf{X}}_1\|}, \quad \cos \varphi = \frac{(\tilde{\mathbf{X}}_5 - \tilde{\mathbf{X}}_1)^T (\tilde{\mathbf{X}}_3 - \tilde{\mathbf{X}}_1)}{\|\tilde{\mathbf{X}}_5 - \tilde{\mathbf{X}}_1\| \cdot \|\tilde{\mathbf{X}}_3 - \tilde{\mathbf{X}}_1\|} \quad (5.3.12)$$

则参数 $\{t_1, t_2, \theta, \phi, \varphi\}$ 是平行六面体的相似不变量, 它确定了平行六面体的形状。

在通常情况下至少有三条平行棱上的 6 个顶点是可见的, 在图像平面上可通过这 6 个可见点的图像恢复出不可见点的图像。在这里, 我们说平面六面体的图像是指三条平行棱上的 6 个可见顶点

$\{\tilde{X}_i : i=1,2,\dots,6\}$ 的图像, 并且约定 $\overline{X_1 X_2} = \overline{X_3 X_4} = \overline{X_5 X_6}$ 。

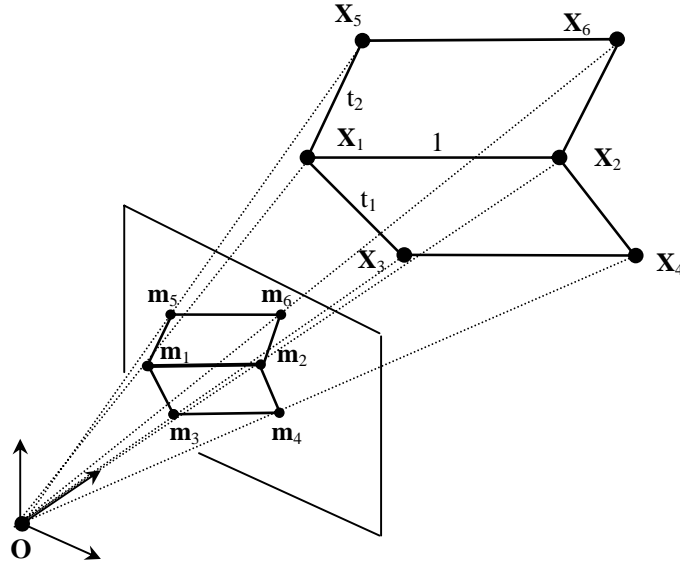


图 5.2.1 平行六面体和它的图像

设平行六面体 $\{\tilde{X}_i : i=1,2,\dots,6\}$ 的图像为 $\{m_i : i=1,2,\dots,6\}$, 记

$$(q_1, q_2, q_3)^T = (-m_1, m_2, m_3)^{-1} m_4, (q_5, q_6, \tilde{q}_3)^T = (-m_5, m_6, m_3)^{-1} m_4$$

根据命题 5.3.2, 有

$$\alpha_4 q_3 K^{-1} m_3 = X_3 = \alpha_4 \tilde{q}_3 K^{-1} m_3$$

所以, 必有 $q_3 = \tilde{q}_3$ 。于是,

$$(q_1, q_2, q_3, q_5, q_6)^T = (A^T A)^{-1} A^T \begin{pmatrix} m_4 \\ m_4 \end{pmatrix} \quad (5.3.13)$$

其中

$$A = \begin{pmatrix} -m_1 & m_2 & m_3 & 0 & 0 \\ 0 & 0 & m_3 & -m_5 & m_6 \end{pmatrix} \quad (5.3.14)$$

因此, 在摄像机坐标系下, 平行六面体的顶点 X_i 可表示为

$$\tilde{X}_i = \alpha_4 q_i K^{-1} m_i, i=1,2,\dots,6 \quad (5.3.15)$$

令

$$X = (\tilde{X}_2 - \tilde{X}_1, \tilde{X}_3 - \tilde{X}_1, \tilde{X}_5 - \tilde{X}_1) \quad (5.3.16)$$

$$M = (q_2 m_2 - q_1 m_1, q_3 m_3 - q_1 m_1, q_5 m_5 - q_1 m_1) \quad (5.3.17)$$

根据式(5.3.15), 有

$$X = \alpha_4 K^{-1} M \quad (5.3.18)$$

由于 X 是可逆矩阵, 所以 M 也是可逆矩阵, 因此,

$$M^{-T} X^T X M^{-1} = \alpha_4^2 K^{-T} K^{-1} = \alpha_4^2 \varpi \quad (5.3.19)$$

并且

$$X^T X = \begin{pmatrix} (\tilde{X}_2 - \tilde{X}_1)^T (\tilde{X}_2 - \tilde{X}_1) & (\tilde{X}_2 - \tilde{X}_1)^T (\tilde{X}_3 - \tilde{X}_1) & (\tilde{X}_2 - \tilde{X}_1)^T (\tilde{X}_5 - \tilde{X}_1) \\ (\tilde{X}_3 - \tilde{X}_1)^T (\tilde{X}_2 - \tilde{X}_1) & (\tilde{X}_3 - \tilde{X}_1)^T (\tilde{X}_3 - \tilde{X}_1) & (\tilde{X}_3 - \tilde{X}_1)^T (\tilde{X}_5 - \tilde{X}_1) \\ (\tilde{X}_5 - \tilde{X}_1)^T (\tilde{X}_2 - \tilde{X}_1) & (\tilde{X}_5 - \tilde{X}_1)^T (\tilde{X}_3 - \tilde{X}_1) & ((\tilde{X}_5 - \tilde{X}_1)^T (\tilde{X}_5 - \tilde{X}_1)) \end{pmatrix} \triangleq \|\tilde{X}_2 - \tilde{X}_1\|^2 \mu \quad (5.3.20)$$

其中

$$\mu = \begin{pmatrix} 1 & t_1 \cos \theta & t_2 \cos \phi \\ t_1 \cos \theta & t_1^2 & t_1 t_2 \cos \varphi \\ t_2 \cos \phi & t_1 t_2 \cos \varphi & t_2^2 \end{pmatrix} \quad (5.3.21)$$

于是, 有

$$M^{-T} \mu M^{-1} = \frac{\alpha_4^2}{\|\tilde{X}_2 - \tilde{X}_1\|^2} \varpi \quad (5.3.22)$$

总结上述讨论, 有下述命题:

命题 5.3.4 设平行六面体 $\{\tilde{X}_i : i=1, 2, \dots, 6\}$ 的图像为 $\{m_i : i=1, 2, \dots, 6\}$, 则

- (1) 在摄像机坐标系下, 顶点 \tilde{X}_i 由式(5.3.15)表示;
- (2) 相似不变量与摄像机内参数矩阵之间具有约束关系(5.3.22)。

注: 矩阵 M 与相似不变量无关, 可直接从图像点得到。通常, 称矩阵 μ 为平行六面体的相似参数矩阵, 式(5.3.22)说明平行六面体的参数矩阵与摄像机内参矩阵是互为对偶的, 对偶关系由矩阵 M 来刻画。

推论 5.3.2 从平行六面体图像确定摄像机内参数矩阵与确定相似不变量 $\{t_1, t_2, \theta, \phi, \varphi\}$ 是等价的。

推论 5.3.3 令 ${}^a X_i = q_i m_i, i=1, 2, \dots, 6$, 则 ${}^a X_i = q_i m_i, i=1, 2, \dots, 6$ 是平行六面体的 $\{\tilde{X}_i : i=1, 2, \dots, 6\}$

的仿射重构^{*)}。

5.3.3 摄像机内参数

下面仅讨论如何从平行四边形图像求解摄像机内参数问题, 因为平行六面体可归结为多个平行四边形的情况。

m 个共面平行四边形的 n 幅图像

给定 m 个共面平行四边形 $\{\tilde{\mathbf{X}}_{ki}^{(j)} : i = 1, 2, 3, 4\}$ 的 n 幅图像 $\{\mathbf{m}_{ki}^{(j)} : i = 1, 2, 3, 4; j = 1, 2, \dots, n\}$, $k = 1, 2, \dots, m$ 。
令

$$(q_{k1}^{(j)}, q_{k2}^{(j)}, q_{k3}^{(j)})^T = (-\mathbf{m}_{k1}^{(j)}, \mathbf{m}_{k2}^{(j)}, \mathbf{m}_{k3}^{(j)})^{-1} \mathbf{m}_{k4}^{(j)}, \quad L_{kj} = (q_{k2}^{(j)} \mathbf{m}_{k2}^{(j)} - q_{k1}^{(j)} \mathbf{m}_{k1}^{(j)}, q_{k3}^{(j)} \mathbf{m}_{k3}^{(j)} - q_{k1}^{(j)} \mathbf{m}_{k1}^{(j)})$$

根据命题 5.3.2, 有

$$\|\tilde{\mathbf{X}}_{k2}^{(j)} - \tilde{\mathbf{X}}_{k1}^{(j)}\|^2 \eta_k = \alpha_{k4}^{(j)} L_{kj}^T \varpi_j L_{kj}, \quad j = 1, 2, \dots, n; k = 1, 2, \dots, m, \quad (5.3.23)$$

其中

$$\tilde{\mathbf{X}}_{ki}^{(j)} = \alpha_{k4}^{(j)} q_{ki}^{(j)} K_j^{-1} \mathbf{m}_{ki}^{(j)} : i = 1, 2, 3, 4 \quad (5.3.24)$$

是第 k 个平行四边形顶点在第 j 个摄像机坐标系下的坐标。从式

$$\|\tilde{\mathbf{X}}_{k2}^{(j)} - \tilde{\mathbf{X}}_{k1}^{(j)}\| = \|\tilde{\mathbf{X}}_{k2} - \tilde{\mathbf{X}}_{k1}\|, \quad j = 1, 2, \dots, n,$$

得到

$$\alpha_{k4}^{(j)} L_{kj}^T \varpi_j L_{kj} = \alpha_{k4}^{(1)} L_{k1}^T \varpi_1 L_{k1}, \quad j = 2, 3, \dots, n; k = 1, 2, \dots, m.$$

消去上述方程中的标量因子, 得到 $2m(n-1)$ 个内参数的二次约束方程

$$\frac{(L_{kj}^T \varpi_j L_{kj})_{11}}{(L_{kj}^T \varpi_j L_{kj})_{22}} = \frac{(L_{k1}^T \varpi_1 L_{k1})_{11}}{(L_{k1}^T \varpi_1 L_{k1})_{22}}, \quad \frac{(L_{kj}^T \varpi_j L_{kj})_{12}}{(L_{kj}^T \varpi_j L_{kj})_{22}} = \frac{(L_{k1}^T \varpi_1 L_{k1})_{12}}{(L_{k1}^T \varpi_1 L_{k1})_{22}}. \quad (5.3.25)$$

在上述 $2m(n-1)$ 个二次约束方程至多有 $2n$ 个是独立的, 因为度量平面 (即圆环点图像已被确定的平面) 的 n 幅图像仅能提供 ϖ_j 的 $2n$ 个独立约束, 所以在给定共面四边形的 n 幅图像情况下, 独立约束数不可能超过 $2n$ 。

对于共面平行四边形, 下述命题是有用的。

命题 5.3.5. 如果两个平行四边形有相同的相似不变量, 则从它们的 n 幅图像能导致 $2n$ 个线性

^{*)} 仿射重构的概念见 6.3 节。

约束方程。

证明 令 $\{\mathbf{m}_{ki}^{(j)} : i=1,2,3,4; j=1,2,\dots,n\}$ 两个平行四边形 $\{\tilde{\mathbf{X}}_{ki}, k=1,2\}$ 的 n 幅图像, 根据命题 5.3.2, 第 k 个四边形顶点 \mathbf{X}_{ki} 在第 j 个摄像机坐标系下的坐标为

$$\tilde{\mathbf{X}}_{ki}^{(j)} = \alpha_{k4}^{(j)} q_{ki}^{(j)} K_j^{-1} \mathbf{m}_{ki}^{(j)}, i=1,2,3,4; k=1,2. \quad (5.3.26)$$

先考虑标量因子 $\alpha_{24}^{(j)}$, $\alpha_{14}^{(j)}$ 之间的关系。

不难看出点集 $\{\hat{\mathbf{X}}_{ki}^{(j)} = \alpha_{k4}^{(j)} q_{ki}^{(j)} \mathbf{m}_{ki}^{(j)} : i=1,2,3,4; k=1,2\}$ 是点集 $\{\tilde{\mathbf{X}}_{ki} : i=1,2,3,4; k=1,2\}$ 的仿射重构, 因此它们必是共面点。所以, 有

$$\det(\alpha_{24}^{(j)} q_{21}^{(j)} \mathbf{m}_{21}^{(j)} - \alpha_{14}^{(j)} q_{11}^{(j)} \mathbf{m}_{11}^{(j)}, \alpha_{14}^{(j)} q_{12}^{(j)} \mathbf{m}_{12}^{(j)} - \alpha_{14}^{(j)} q_{11}^{(j)} \mathbf{m}_{11}^{(j)}, \alpha_{14}^{(j)} q_{13}^{(j)} \mathbf{m}_{13}^{(j)} - \alpha_{14}^{(j)} q_{11}^{(j)} \mathbf{m}_{11}^{(j)}) = 0.$$

解这个方程, 得到

$$\alpha_{24}^{(j)} = \frac{\det[q_{11}^{(j)} \mathbf{m}_{11}^{(j)}, q_{12}^{(j)} \mathbf{m}_{12}^{(j)}, q_{13}^{(j)} \mathbf{m}_{13}^{(j)}]}{\det[q_{21}^{(j)} \mathbf{m}_{21}^{(j)}, q_{12}^{(j)} \mathbf{m}_{12}^{(j)} - q_{11}^{(j)} \mathbf{m}_{11}^{(j)}, q_{13}^{(j)} \mathbf{m}_{13}^{(j)} - q_{11}^{(j)} \mathbf{m}_{11}^{(j)}]} \alpha_{14}^{(j)} \triangleq p_j \alpha_{14}^{(j)}. \quad (5.3.27)$$

令 $\tilde{q}_{2i}^{(j)} = p_j q_{2i}^{(j)}$, $\tilde{q}_{1i}^{(j)} = q_{1i}^{(j)}$, $i=1,2,3,4$, 则等式(5.3.26)能被改写成

$$\tilde{\mathbf{X}}_{ki}^{(j)} = \alpha_{14}^{(j)} \tilde{q}_{ki}^{(j)} K_j^{-1} \mathbf{m}_{ki}^{(j)}, i=1,2,3,4; k=1,2$$

并且点集

$$\{\hat{\mathbf{X}}_{ki}^{(j)} = \tilde{q}_{ki}^{(j)} \mathbf{m}_{ki}^{(j)} : i=1,2,3,4; k=1,2\} \quad (5.3.28)$$

仍是 $\{\tilde{\mathbf{X}}_{ki} : i=1,2,3,4; k=1,2\}$ 的仿射重构。根据命题 5.3.2, 有

$$\|\tilde{\mathbf{X}}_{k2}^{(j)} - \tilde{\mathbf{X}}_{k1}^{(j)}\|^2 \eta_k = \alpha_{14}^{(j)} \tilde{L}_{kj}^T \varpi_j \tilde{L}_{kj}, j=1,2,\dots,n; k=1,2.$$

其中 $\tilde{L}_{kj} = [q_{k2}^{(j)} \mathbf{m}_{k2}^{(j)} - q_{k1}^{(j)} \mathbf{m}_{k1}^{(j)}, q_{k3}^{(j)} \mathbf{m}_{k3}^{(j)} - q_{k1}^{(j)} \mathbf{m}_{k1}^{(j)}]$ 。于是, 由 $\eta_2 = \eta_1$ 和 $\|\tilde{\mathbf{X}}_{k2}^{(j)} - \tilde{\mathbf{X}}_{k1}^{(j)}\| = \|\tilde{\mathbf{X}}_{k2} - \tilde{\mathbf{X}}_{k1}\|$, 得到

$$\tilde{L}_{2j}^T \varpi_j \tilde{L}_{2j} = s \tilde{L}_{1j}^T \varpi_j \tilde{L}_{1j}, j=1,2,\dots,n, \quad (5.3.29)$$

其中 $s = \|\tilde{\mathbf{X}}_{22} - \tilde{\mathbf{X}}_{21}\|^2 / \|\tilde{\mathbf{X}}_{12} - \tilde{\mathbf{X}}_{11}\|^2$ 。如果能确定标量因子 s , 则方程(5.3.29) 能提供 $2n$ 个线性约束。下面讨论如何确定这个标量因子。

因两个平行四边形是相似的, 所以有

$$s = \frac{\|\tilde{\mathbf{X}}_{22} - \tilde{\mathbf{X}}_{21}\|^2}{\|\tilde{\mathbf{X}}_{12} - \tilde{\mathbf{X}}_{11}\|^2} = \frac{\sigma(\{\tilde{\mathbf{X}}_{2i} : i=1,2,3,4\})}{\sigma(\{\tilde{\mathbf{X}}_{1i} : i=1,2,3,4\})},$$

其中 σ 表示平行四边形的面积。由于仿射变换保持两个共面图形的面积比, 所以从仿射重构(5.3.28)

得到标量因子：

$$s = \frac{\sigma(\{\tilde{\mathbf{X}}_{2i}^{(j)} : i = 1, 2, 3, 4\})}{\sigma(\{\tilde{\mathbf{X}}_{1i}^{(j)} : i = 1, 2, 3, 4\})} = \frac{\|(\tilde{q}_{23}^{(j)} \mathbf{m}_{23}^{(j)} - \tilde{q}_{21}^{(j)} \mathbf{m}_{21}^{(j)}) \times (\tilde{q}_{22}^{(j)} \mathbf{m}_{22}^{(j)} - \tilde{q}_{21}^{(j)} \mathbf{m}_{21}^{(j)})\|}{\|(\tilde{q}_{13}^{(j)} \mathbf{m}_{13}^{(j)} - \tilde{q}_{11}^{(j)} \mathbf{m}_{11}^{(j)}) \times (\tilde{q}_{12}^{(j)} \mathbf{m}_{12}^{(j)} - \tilde{q}_{11}^{(j)} \mathbf{m}_{11}^{(j)})\|} \quad (5.3.30)$$

注 在上述命题中，两个相似平行四边形之间应存在一个 2D 旋转，否则该命题不真。此外，这个命题能推广到任意两个共面的相似四点形，见 5.2 节关于从相似图形确定圆环点图像的论述。

命题 5.3.6. 如果两个共面平行四边形有相同的边长（不一定相似）则从它们的 n 幅图像能导致 $2n$ 个线性约束方程。

证明 令 $\{\mathbf{m}_{ki}^{(j)} : i = 1, 2, 3, 4; j = 1, 2 \dots n\}$ 两个平行四边形 $\{\tilde{\mathbf{X}}_{ki}, k = 1, 2\}$ 的 n 幅图像。不失一般性，假定 $\|\tilde{\mathbf{X}}_{22} - \tilde{\mathbf{X}}_{21}\| = \|\tilde{\mathbf{X}}_{12} - \tilde{\mathbf{X}}_{11}\|$, $\|\tilde{\mathbf{X}}_{23} - \tilde{\mathbf{X}}_{21}\| = \|\tilde{\mathbf{X}}_{13} - \tilde{\mathbf{X}}_{11}\|$ 。根据命题 5.3.5 的证明，有

$$\|\tilde{\mathbf{X}}_{k2} - \tilde{\mathbf{X}}_{k1}\|^2 \eta_k = \alpha_{14}^{(j)} \tilde{L}_{kj}^T \boldsymbol{\omega}_j \tilde{L}_{kj}, j = 1, 2 \dots n; k = 1, 2.$$

于是，可得到 $2n$ 个线性约束方程

$$(\tilde{L}_{2j}^T \boldsymbol{\omega}_j \tilde{L}_{2j})_{11} = (\tilde{L}_{1j}^T \boldsymbol{\omega}_j \tilde{L}_{1j})_{11}, (\tilde{L}_{2j}^T \boldsymbol{\omega}_j \tilde{L}_{2j})_{22} = (\tilde{L}_{1j}^T \boldsymbol{\omega}_j \tilde{L}_{1j})_{22}, j = 1, 2 \dots n. \quad (5.3.31)$$

m 个非共面平行四边形的 n 幅图像

下面主要讨论如何利用摄像机或平行四边形的先验知识，从 m 个非共面平行四边形的 n 幅图像线性确定摄像机内参数、平行四边形的相似不变量。

假定在空间中有 m 个平行四边形 $\{\tilde{\mathbf{X}}_{ki}\}, k=1, 2 \dots m$ ，在它们当中至少存两个不共面，它们的 n 幅图像记为 $\{\mathbf{m}_{ki}^{(j)}\}$ 。令

$$(q_{k1}^{(j)}, q_{k2}^{(j)}, q_{k3}^{(j)})^T = (-\mathbf{m}_{k1}^{(j)}, \mathbf{m}_{k2}^{(j)}, \mathbf{m}_{k3}^{(j)})^{-1} \mathbf{m}_{k4}^{(j)}, L_{kj} = (q_{k2}^{(j)} \mathbf{m}_{k2}^{(j)} - q_{k1}^{(j)} \mathbf{m}_{k1}^{(j)}, q_{k3}^{(j)} \mathbf{m}_{k3}^{(j)} - q_{k1}^{(j)} \mathbf{m}_{k1}^{(j)}).$$

由命题 5.3.2，第 k 个四边形顶点 \mathbf{X}_{ki} 在第 j 个摄像机坐标系下的坐标为

$$\tilde{\mathbf{X}}_{ki}^{(j)} = \alpha_{k4}^{(j)} q_{ki}^{(j)} K_j^{-1} \mathbf{m}_{ki}^{(j)} : i = 1, 2, 3, 4. \quad (5.3.32)$$

根据命题 5.3.2 的注，图像点 $\mathbf{v}_{k1}^{(j)} = q_{k2}^{(j)} \mathbf{m}_{k2}^{(j)} - q_{k1}^{(j)} \mathbf{m}_{k1}^{(j)}$, $\mathbf{v}_{k2}^{(j)} = q_{k3}^{(j)} \mathbf{m}_{k3}^{(j)} - q_{k1}^{(j)} \mathbf{m}_{k1}^{(j)}$ 是第 k 个四边形两平对平行边在第 j 个像平面上的消隐点，因此从点对应 $\{\mathbf{v}_{k1}^{(1)} \leftrightarrow \mathbf{v}_{k1}^{(j)}, \mathbf{v}_{k2}^{(1)} \leftrightarrow \mathbf{v}_{k2}^{(j)}, k = 1, 2 \dots m\}$ ，能线性确定第一幅图像与第 j 幅图像间的无穷远单应 H_{1j} 。于是，得到 $5n-5$ 个关于 $\boldsymbol{\omega}_j$ 的约束方程（参考 5.5 节）：

$$\boldsymbol{\omega}_j = s_j H_{1j}^{-T} \boldsymbol{\omega}_1 H_{1j}^{-1}, j = 2, 3 \dots n, \quad (5.3.33)$$

其中 s_j 是未知标量。另一方面根据命题 5.3.2，得到关于 $(\eta_k, \boldsymbol{\omega}_j)$ 的约束：

$$\eta_k = t_{kj} L_{kj}^T \varpi_j L_{kj}, j = 1, 2 \dots n; k = 1, 2 \dots m, \tag{5.3.34}$$

其中 $t_{kj} = (\alpha_{k4}^{(j)} / \| \mathbf{X}_{k2} - \mathbf{X}_{k1} \|^2)$ 也是未知标量。

注意式 (5.3.33) 和(5.3.34) 中的所有约束都是非线性的。然而，利用摄像机的先验知识，从 (5.3.33) 能得到关于 ϖ_1 线性约束(见表 5.3.1); 利用平行四边形的先验知识，从 (5.3.34) 也能得到关于 ϖ_1 线性约束(见表 5.3.2)。

从上面讨论，可以看出使用摄像机或平行四边形先验知识，可线性确定 ϖ_1 。一旦 ϖ_1 被确定， ϖ_j 可由等式 5.3.33 确定，于是 η_k 能从式(5.3.34)确定。

表 5.3.1 来自摄像机先验的线性约束

摄像机先验	ϖ_1 的线性约束
零畸变因子	$(H_{1j}^{-T} \varpi_1 H_{1j}^{-1})_{12} = 0$
主点在原点	$(H_{1j}^{-T} \varpi_1 H_{1j}^{-1})_{13} = (H_{1j}^{-T} \varpi_1 H_{1j}^{-1})_{23} = 0$
已知纵横比 $\tau = f_v / f_u$	$\tau^2 (H_{1j}^{-T} \varpi_1 H_{1j}^{-1})_{22} - (H_{1j}^{-T} \varpi_1 H_{1j}^{-1})_{11} = 0$

表 5.3.2 来自平行四边形先验的线性约束

平行四边形先验	ϖ_1 的线性约束
$\theta_k = \pi / 2$	$(L_{k1}^T \varpi_1 L_{k1})_{12} = 0$
$t_{k1} = 1$	$(L_{k1}^T \varpi_1 L_{k1})_{11} = (L_{k1}^T \varpi_1 L_{k1})_{22}$
共面相似	线性约束方程(5.3.29)
共面等边长	线性约束方程(5.3.31)

5.4 Kruppa 方程与摄像机内参数

5.4.1 Kruppa 方程

令空间二次曲线 A 在两个视点下的像曲线分别为 A_l, A_r ，并假定二次曲线 A 支撑平面的单应矩阵为 H 。参考图 5.4.1， π_a, π_b 是与二次曲线 A 相切的两个外极面，它们与两个像平面的交线分别为 l_{la}, l_{lb} 和 l_{ra}, l_{rb} ，显然， $l_{la} \leftrightarrow l_{ra}, l_{lb} \leftrightarrow l_{rb}$ 是两对对应极线。

由于 l_{la}, l_{lb} 均为二次曲线 A_l (它是 A 在第一个摄像机下的图像二次曲线)的切线，并且相交于极点 e_l ，所以根据命题 1.2.2，两条直线 l_{la}, l_{lb} 所确定的退化二次曲线为 $A_{lab} = [e_l]_x A_l^{-1} [e_l]_x^T$ 。同样 l_{ra}, l_{rb}

所确定的退化二次曲线为 $A_{rab} = [e_r]_x A_r^{-1} [e_r]_x^T$ 。

根据二次曲线的射影变换规则，有

$$A_{rab} = H^{-T} A_{lab} H^{-1}$$

所以，

$$[e_r]_x A_r^{-1} [e_r]_x^T = A_{rab} = H^{-T} A_{lab} H^{-1} = H^{-T} [e_l]_x A_l^{-1} [e_l]_x^T H^{-1}$$

因两幅图像间的基本矩阵为 $F = H^{-T} [e_r]_x$ ，所以，

$$F A_l^{-1} F^T = s [e_r]_x A_r^{-1} [e_r]_x^T \tag{5.4.1}$$

这里，显式地写出了齐次因子 s 。

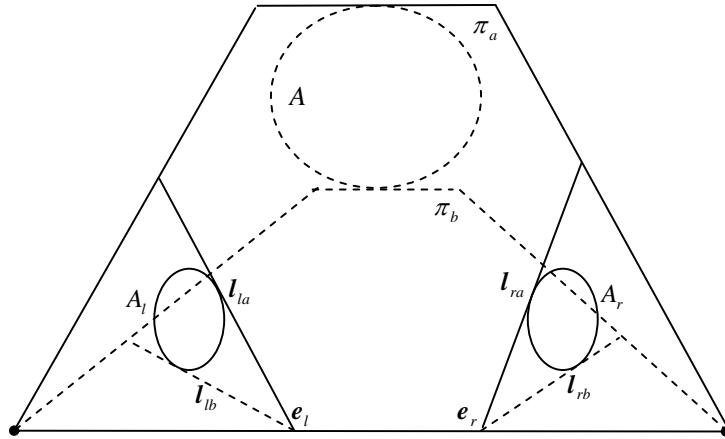


图 5.4.1：二次曲线的极切线。

Kruppa 方程

由于绝对二次曲线在两个视点下的像曲线分别为 $\varpi = K^{-T} K^{-1}$, $\varpi' = K'^{-T} K'^{-1}$ ，所以将它们代入式 (5.4.1) 得到 DIAC 的约束方程：

$$F \varpi^* F^T = s [e_r]_x \varpi'^* [e_r]_x^T \tag{5.4.2}$$

通常称这个方程为 Kruppa 方程。不难看出，Kruppa 方程的标量因子 s 必是非零正数。

根据基本矩阵 F 的秩为 2 的性质， F 的 SVD 分解必有下述形式：

$$F = U D V^T = U \text{diag}(\sigma_1, \sigma_2, 0) V^T$$

其中 u_3 和 v_3 满足 $F^T u_3 = 0$ 和 $F v_3 = 0$ ，即极点 $e_l = v_3$ 和 $e_r = u_3$ 。将它们代入式 (5.4.2) 得到：

$$[\mathbf{u}_3]_x \varpi'^* [\mathbf{u}_3]_x = \mathbf{U} \mathbf{D} \mathbf{V}^T \varpi^* (\mathbf{U} \mathbf{D} \mathbf{V}^T)^T \quad (5.4.3)$$

对上式左乘以 \mathbf{U}^T 并且右乘以 \mathbf{U} , 根据 \mathbf{U} 的正交性, 方程(5.4.3)的两边分别化为:

$$\mathbf{U}^T [\mathbf{u}_3]_x \varpi'^* [\mathbf{u}_3]_x \mathbf{U} = (\mathbf{u}_2, -\mathbf{u}_1, 0)^T \varpi'^* (\mathbf{u}_2, -\mathbf{u}_1, 0) = \begin{pmatrix} \mathbf{u}_2^T \varpi'^* \mathbf{u}_2 & -\mathbf{u}_1^T \varpi'^* \mathbf{u}_2 & 0 \\ -\mathbf{u}_1^T \varpi'^* \mathbf{u}_2 & \mathbf{u}_1^T \varpi'^* \mathbf{u}_1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\mathbf{D} \mathbf{V}^T \varpi^* \mathbf{V} \mathbf{D} = \text{diag}(\sigma_1, \sigma_2, 0) \mathbf{V}^T \varpi^* \mathbf{V} \text{diag}(\sigma_1, \sigma_2, 0) = \begin{pmatrix} \sigma_1^2 \mathbf{v}_1^T \varpi^* \mathbf{v}_1 & \sigma_1 \sigma_2 \mathbf{v}_1^T \varpi^* \mathbf{v}_2 & 0 \\ \sigma_1 \sigma_2 \mathbf{v}_1^T \varpi^* \mathbf{v}_2 & \sigma_2^2 \mathbf{v}_2^T \varpi^* \mathbf{v}_2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

因此, 有

$$\begin{pmatrix} \mathbf{u}_2^T \varpi'^* \mathbf{u}_2 \\ -\mathbf{u}_1^T \varpi'^* \mathbf{u}_2 \\ \mathbf{u}_1^T \varpi'^* \mathbf{u}_1 \end{pmatrix} = s \begin{pmatrix} \sigma_1^2 \mathbf{v}_1^T \varpi^* \mathbf{v}_1 \\ \sigma_1 \sigma_2 \mathbf{v}_1^T \varpi^* \mathbf{v}_2 \\ \sigma_2^2 \mathbf{v}_2^T \varpi^* \mathbf{v}_2 \end{pmatrix}$$

消去上式标量因子 s 后, 可知 Kruppa 方程(5.4.1)等价于下述两个二次方程:

$$\frac{\mathbf{u}_2^T \varpi'^* \mathbf{u}_2}{\mathbf{u}_1^T \varpi'^* \mathbf{u}_1} = \frac{\sigma_1^2 \mathbf{v}_1^T \varpi^* \mathbf{v}_1}{\sigma_2^2 \mathbf{v}_2^T \varpi^* \mathbf{v}_2}, \quad \frac{\mathbf{u}_2^T \varpi'^* \mathbf{u}_2}{\mathbf{u}_1^T \varpi'^* \mathbf{u}_1} = -\frac{\sigma_1 \mathbf{v}_1^T \varpi^* \mathbf{v}_2}{\sigma_2 \mathbf{v}_2^T \varpi^* \mathbf{v}_2} \quad (5.4.4)$$

Kruppa 方程是对 DIAC 的约束, 它与场景的几何结构无关, 仅依赖于两个视点的极几何。每个 Kruppa 方程仅能提供关于 DIAC 的 6 个未知参数的两个独立约束。因此, 给定三幅视图并已知每对视图之间的基本矩阵时, 一般可得到 6 个二次约束。如果摄像机内参数在运动过程中保持不变, 则这 6 个二次约束能够确定 ϖ^* , 从而能获得摄像机内参数。如果没有摄像机内参数的知识, 则不论多少幅图像都无法确定摄像机内参数, 见第 5.5 节的讨论。

值得指出的是: 在纯平移情形下, 式(5.4.1)变为 $[\mathbf{e}_r]_x \varpi^* [\mathbf{e}_r]_x = [\mathbf{e}_r]_x \varpi^* [\mathbf{e}_r]_x$, 它是一个恒等式。因此纯平移运动的 Kruppa 方程不能对内参数构成约束。

5.4.2 由 Kruppa 方程求焦距

两幅图像的 Kruppa 方程至多能提供内参数的两个约束, 下述命题在两个摄像机的畸变因子为零, 并且像素的纵横比等于 1 的情况下, 给出了由主点表示焦距的显式表达式。它的重要性在于当主点已知时, 可以利用它直接计算焦距。

命题 5.4.1 假定两个摄像机的畸变因子为零, 并且像素的纵横比等于 1, 即左、右摄像机内参数

矩阵有下述形式：

$$K = \begin{pmatrix} f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{pmatrix}, \quad K' = \begin{pmatrix} f' & 0 & u'_0 \\ 0 & f' & v'_0 \\ 0 & 0 & 1 \end{pmatrix}$$

则可由 Kruppa 方程计算出焦距：

$$f^2 = -\frac{\mathbf{p}'^T [e_r]_{\times} \bar{I} F \mathbf{p} \mathbf{p}'^T F^T \mathbf{p}'}{\mathbf{p}'^T [e_r]_{\times} \bar{I} F \bar{I} F^T \mathbf{p}'} \quad (5.4.5)$$

$$f'^2 = -\frac{\mathbf{p}^T [e_l]_{\times} \bar{I} F^T \mathbf{p}' \mathbf{p}'^T F \mathbf{p}}{\mathbf{p}^T [e_l]_{\times} \bar{I} F^T \bar{I} F \mathbf{p}} \quad (5.4.6)$$

其中， $\bar{I} = \text{diag}(1, 1, 0)$ ， $\mathbf{p} = (u_0, v_0, 1)^T$ ， $\mathbf{p}' = (u'_0, v'_0, 1)^T$ 。

证明 由于两个摄像机的畸变因子为零，并且像素的纵横比等于 1，所以两个视点的 DIAC 可分别表示成

$$\omega^* = f^2 \bar{I} + \mathbf{p} \mathbf{p}^T, \quad \omega'^* = f'^2 \bar{I} + \mathbf{p}' \mathbf{p}'^T$$

将它们代入 Kruppa 方程，我们有

$$f^2 F \bar{I} F^T + F \mathbf{p} \mathbf{p}^T F^T = s f'^2 [e_r]_{\times} \bar{I} [e_r]_{\times}^T + s [e_r]_{\times} \mathbf{p} \mathbf{p}'^T [e']_{\times}^T$$

两边同时右乘以主点 $\mathbf{p}' = (u'_0, v'_0, 1)^T$ ，可得

$$f^2 F \bar{I} F^T \mathbf{p}' + F \mathbf{p} \mathbf{p}^T F^T \mathbf{p}' = s f'^2 [e_r]_{\times} \bar{I} [e_r]_{\times}^T \mathbf{p}'$$

两边再同时左乘以 $\mathbf{p}'^T [e_r]_{\times} \bar{I}$ ，可知

$$f^2 \mathbf{p}'^T [e_r]_{\times} \bar{I} F \bar{I} F^T \mathbf{p}' + \mathbf{p}'^T [e_r]_{\times} \bar{I} F \mathbf{p} \mathbf{p}'^T F^T \mathbf{p}' = s f'^2 \mathbf{p}'^T ([e_r]_{\times} \bar{I} [e_r]_{\times} \bar{I} [e_r]_{\times}^T) \mathbf{p}' = 0$$

因此，我们得到

$$f^2 = -\frac{\mathbf{p}'^T [e_r]_{\times} \bar{I} F \mathbf{p} \mathbf{p}'^T F^T \mathbf{p}'}{\mathbf{p}'^T [e_r]_{\times} \bar{I} F \bar{I} F^T \mathbf{p}'}$$

类似地，可推知：

$$f'^2 = -\frac{\mathbf{p}^T [e_l]_{\times} \bar{I} F^T \mathbf{p}' \mathbf{p}'^T F \mathbf{p}}{\mathbf{p}^T [e_l]_{\times} \bar{I} F^T \bar{I} F \mathbf{p}}$$

证毕。

5.5 绝对二次曲线与摄像机内参数

5.5.1 基本约束方程

无穷远单应矩阵 H_∞ 是无穷远平面 π_∞ 诱导出的两幅图像间的二维射影变换，由二次曲线的射影变换规则，两个视点下的 IAC $\{\varpi, \varpi'\}$ 必满足下述关系：

$$\varpi' = s H_\infty^{-T} \varpi H_\infty^{-1} \quad (5.5.1)$$

这里显式地写出了齐次因子 s 。这样，就得到了 IAC 的约束方程。

对(5.5.1)两边求逆，有

$$\varpi'^* = u H_\infty \varpi^* H_\infty^T \quad (5.5.2)$$

其中齐次因子 $u=1/s$ 。这是关于 DIAC 的约束方程。

不难看出，约束方程(5.5.1)与(5.5.2)是相互等价的。这两个约束方程是摄像机自标定中重要的约束关系，同时也是由仿射重构提升到度量重构的基础。一旦确定了 ϖ 或 ϖ^* ，利用(5.5.1)或(5.5.2)，可以确定 ϖ' 或 ϖ'^* ，再对它们进行 Cholesky 分解就可以得到 K 和 K' 。

对于两个视点，不论是 IAC 约束还是 DIAC 约束，在摄像机内参数不变的情况都至多能提供内参数的 4 个独立约束。因此，至少需要三个视点才可能确定摄像机的 5 个内参数，但求解方法是线性的。在变化内参数情况下，对于每对视点，不论是 IAC 约束还是 DIAC 约束，都至多能提供内参数的 5 个独立约束，每增加一个视点会增加关于摄像机的 5 个新参数，因此如果没有内参数或其它的先验知识，不论使用多少幅图像，都不可能确定所有摄像机的内参数。

5.5.2 变化内参数

在这种情况下，由于每个形如(5.5.1, 5.5.2) 的方程至多能提供内参数的 5 个独立约束。正如上面所述，如果没有关于内参数的一些先验知识，无法确定摄像机的所有内参数。因此，本节主要讨论在已知某些有关内参数知识的情况下，如何标定变参数的摄像机。

基本原理 将第 j 个摄像机的 IAC 或 DIAC 元素的约束关系，通过式(5.5.1, 5.5.2)传递给第 1 个摄像机的 IAC 或 DIAC，从而得到第 1 个摄像机的 IAC 或 DIAC 的约束。当视点足够多时，可得到第 1 个摄像机 IAC 或 DIAC 的足够多约束，于是可求解第 1 个摄像机的 IAC 或 DIAC，再通过式(5.5.1)或(5.5.2)得到其它摄像机的 IAC 或 DIAC。

假定有 n 个视点，无穷远平面关于第 1 个视点与第 j 个视点间的单应矩阵记为 $H_{j\infty}$ ，第 j 个视点的 IAC 和 DIAC 矩阵分别记为 ϖ_j 和 ϖ_j^* 。先考虑 IAC 约束，由(5.5.1)，有

$$\varpi_j = s_j H_{j\infty}^{-T} \varpi H_{j\infty}^{-1}, \quad j=2,3,\dots,n \quad (5.5.3)$$

其中: ϖ 是第 1 个摄像机的 IAC。

记 $(g_{kl}^{(j)}(\mathbf{c})) = H_{j\infty}^{-T} \varpi H_{j\infty}^{-1}$, 其中 $\mathbf{c} = (\varpi_{11}, \varpi_{12}, \varpi_{13}, \varpi_{22}, \varpi_{23}, \varpi_{33})^T$ 。于是, 有

$$s'_j \varpi_j = (g_{kl}^{(j)}(\mathbf{c})), \quad j=2,3,\dots,n \quad (5.5.4)$$

其中 $s'_j = s_j^{-1}$ 。显然, $g_{kl}^{(j)}(\mathbf{c})$ 均为 \mathbf{c} 的线性函数。

畸变因子为零 当摄像机的畸变因子为零时, 可知第 j 个摄像机的 IAC 矩阵 ϖ_j 元素 $(\varpi_j)_{12} = 0$, 由(5.5.4)式, 有

$$g_{12}^{(j)}(\mathbf{c}) = 0, \quad j=2,3,\dots,n \quad (5.5.5)$$

这样, 就有关于 \mathbf{c} 的齐次线性约束方程组。当 $n \geq 6$ 时可以确定 \mathbf{c} , 从而确定了 ϖ , 将它代入(5.5.1)式, 就得到了 ϖ_j 。

主点在原点 当摄像机的主点在原点时, 可知 $(\varpi_j)_{13} = (\varpi_j)_{23} = 0$, 于是根据(5.5.4)式得到 \mathbf{c} 的齐次线性约束方程组:

$$g_{13}^{(j)}(\mathbf{c}) = g_{23}^{(j)}(\mathbf{c}) = 0, \quad j=2,3,\dots,n \quad (5.5.6)$$

已知摄像机的畸变因子为零并且已知像素的纵横比 假定第 j 个摄像机像素的纵横比 $r_j = \frac{f_v}{f_u}$,

可以知 $(\varpi_j)_{22} = r_j^2 (\varpi_j)_{11}$, 于是根据(5.5.4)式得到 \mathbf{c} 的齐次线性约束方程组:

$$g_{22}^{(j)}(\mathbf{c}) - r_j^2 g_{11}^{(j)}(\mathbf{c}) = 0, \quad g_{12}^{(j)}(\mathbf{c}) = 0, \quad j=2,3,\dots,n \quad (5.5.7)$$

已知摄像机的畸变因子为零并且每个摄像机像素纵横比相同(但未知) 根据上面的类似讨论, 可得到 \mathbf{c} 的约束方程组:

$$g_{12}^{(j)}(\mathbf{c}) = 0, \quad j=2,3,\dots,n; \quad g_{11}^{(j)}(\mathbf{c})g_{22}^{(1)}(\mathbf{c}) - g_{22}^{(j)}(\mathbf{c})g_{11}^{(1)}(\mathbf{c}) = 0, \quad j=3,4,\dots,n \quad (5.5.8)$$

注意, 后一组约束方程是非线性的。

对于 DIAC, 可以作类似的分析。

5.5.3 恒定内参数

如果内参数是恒定的, 即 $\varpi_j^* = \varpi^*$, $j=1,\dots,n$, 那么 ϖ^* 的约束方程(5.5.2)变成为

$$\varpi^* = \eta_j H_{j\infty} \varpi^* H_{j\infty}^T, \quad j=1,\dots,n \quad (5.5.9)$$

与变参数的情况不同, 这里的齐次因子 η_j 可以不通过求解方程(5.5.9)而事先确定。事实上, 对(5.5.9)两边取行列式, 就可以得到

$$\eta_j = \left(\frac{1}{\sqrt[3]{\det(H_{j\infty})}} \right)^2 \quad (5.5.10)$$

令 $\bar{H}_{j\infty} = \frac{H_{j\infty}}{\sqrt[3]{\det(H_{j\infty})}}$, 则 $\det(\bar{H}_{j\infty}) = 1$ 且 (5.5.9)变为 $\varpi^* = \bar{H}_{j\infty} \varpi^* \bar{H}_{j\infty}^T$ 。这就是说, 事先将 H_{∞} 的行列式归一化 $\det(H_{\infty}) = 1$, 方程(5.5.9)总可写成下面的形式:

$$\varpi^* = H_{j\infty} \varpi^* H_{j\infty}^T \quad (5.5.11)$$

式(5.5.11)构成关于对称矩阵 ϖ^* 的 6 个元素的 6 个线性约束方程, 将这 6 个方程写成齐次线性形式:

$$A\mathbf{c}^* = 0 \quad (5.5.12)$$

其中 A 是由 $H_{j\infty}$ 的元素组成的一个 6×6 矩阵, 而 \mathbf{c}^* 是二次曲线 ϖ^* 构成的 6 维向量。 \mathbf{c}^* 不能由单个这样的方程组所唯一确定, 因为此时 A 的秩最多为 4。如果将 $n \geq 2$ 对视点所对应的线性约束方程(5.5.12)联立起来, 使得 A 是一个 $6n \times 6$ 的矩阵, 并且视点之间的旋转具有不同的旋转轴, 则在一般情况下 \mathbf{c}^* 能被唯一确定。

5.5.4 两幅图像求解尺度因子

下面给出两个视点在畸变因子为零且主点已知时的求解尺度因子的计算公式。

命题 5.5.1 已知两个视点的畸变因子为零, 则有下列求解尺度因子 f_u, f_v, f'_u, f'_v 的计算公式:

$$f_u'^2 = \frac{\mathbf{p}'^T (\mathbf{h}_1 \times \mathbf{h}_2) (\tilde{\mathbf{h}}_1^T \mathbf{p} - \tilde{\mathbf{h}}_3^T \mathbf{p} \mathbf{e}_1^T \mathbf{p}')}{\tilde{\mathbf{h}}_3^T \mathbf{p} \mathbf{e}_1^T (\mathbf{h}_1 \times \mathbf{h}_2)}, \quad f_v'^2 = \frac{\mathbf{p}'^T (\mathbf{h}_1 \times \mathbf{h}_2) (\tilde{\mathbf{h}}_2^T \mathbf{p} - \tilde{\mathbf{h}}_3^T \mathbf{p} \mathbf{e}_2^T \mathbf{p}')}{\tilde{\mathbf{h}}_3^T \mathbf{p} \mathbf{e}_2^T (\mathbf{h}_1 \times \mathbf{h}_2)},$$

$$f_u^2 = \frac{\mathbf{p}^T (\mathbf{g}_1 \times \mathbf{g}_2) (\tilde{\mathbf{g}}_1^T \mathbf{p} - \tilde{\mathbf{g}}_3^T \mathbf{p}' \mathbf{e}_1^T \mathbf{p})}{\tilde{\mathbf{g}}_3^T \mathbf{p}' \mathbf{e}_1^T (\mathbf{g}_1 \times \mathbf{g}_2)}, \quad f_v^2 = \frac{\mathbf{p}^T (\mathbf{g}_1 \times \mathbf{g}_2) (\tilde{\mathbf{g}}_2^T \mathbf{p} - \tilde{\mathbf{g}}_3^T \mathbf{p}' \mathbf{e}_2^T \mathbf{p})}{\tilde{\mathbf{g}}_3^T \mathbf{p}' \mathbf{e}_2^T (\mathbf{g}_1 \times \mathbf{g}_2)}$$

其中: $\mathbf{p} = (u_0, v_0, 1)^T$, $\mathbf{p}' = (u'_0, v'_0, 1)^T$ 分别为第 1 个和第 2 个摄像机的主点; $\mathbf{h}_j, \tilde{\mathbf{h}}_j$ 分别为 H_{∞} 的第 j 列和第 j 行向量; $\mathbf{g}_j, \tilde{\mathbf{g}}_j^T$ 分别为 H_{∞}^{-1} 的第 j 列和第 j 行向量; $\mathbf{e}_1 = (1, 0, 0)^T$, $\mathbf{e}_2 = (0, 1, 0)^T$, $\mathbf{e}_3 = (0, 0, 1)^T$ 。

证明 根据假定, 两个摄像机的 DIAC 矩阵可表示为

$$\varpi = \text{diag}(f_u^2, f_v^2, 0) + \mathbf{p}\mathbf{p}^T, \quad \varpi' = \text{diag}(f_u'^2, f_v'^2, 0) + \mathbf{p}'\mathbf{p}'^T$$

将上式代入(5.5.2), 有

$$\text{diag}(f_u'^2, f_v'^2, 0) + \mathbf{p}'\mathbf{p}'^T = \eta(f_u^2 \mathbf{h}_1 \mathbf{h}_1^T + f_v^2 \mathbf{h}_2 \mathbf{h}_2^T + H_\infty \mathbf{p} \mathbf{p}^T H_\infty^T)$$

用向量 $(\mathbf{h}_1 \times \mathbf{h}_2)$ 右乘上式两边, 得到

$$\text{diag}(f_u'^2, f_v'^2, 0)(\mathbf{h}_1 \times \mathbf{h}_2) + \mathbf{p}'\mathbf{p}'^T (\mathbf{h}_1 \times \mathbf{h}_2) = \eta H_\infty \mathbf{p} \mathbf{p}^T H_\infty^T (\mathbf{h}_1 \times \mathbf{h}_2)$$

而 $\mathbf{p}^T H_\infty^T (\mathbf{h}_1 \times \mathbf{h}_2) = \mathbf{p}^T \begin{pmatrix} 0 \\ 0 \\ \mathbf{h}_3^T (\mathbf{h}_1 \times \mathbf{h}_2) \end{pmatrix} = \det(H_\infty)$, 所以,

$$\text{diag}(f_u'^2, f_v'^2, 0)(\mathbf{h}_1 \times \mathbf{h}_2) + \mathbf{p}'\mathbf{p}'^T (\mathbf{h}_1 \times \mathbf{h}_2) = \eta \det(H_\infty) H_\infty \mathbf{p} \quad (5.5.13)$$

显然,

$$\mathbf{e}_1^T \text{diag}(f_u'^2, f_v'^2, 0)(\mathbf{h}_1 \times \mathbf{h}_2) = \mathbf{e}_1^T (\mathbf{h}_1 \times \mathbf{h}_2) f_u'^2$$

$$\mathbf{e}_2^T \text{diag}(f_u'^2, f_v'^2, 0)(\mathbf{h}_1 \times \mathbf{h}_2) = \mathbf{e}_2^T (\mathbf{h}_1 \times \mathbf{h}_2) f_v'^2$$

$$\mathbf{e}_3^T \text{diag}(f_u'^2, f_v'^2, 0)(\mathbf{h}_1 \times \mathbf{h}_2) = 0$$

用 \mathbf{e}_3^T 左乘式(5.5.13), 有

$$\mathbf{e}_3^T \mathbf{p}'\mathbf{p}'^T (\mathbf{h}_1 \times \mathbf{h}_2) = \eta \det(H_\infty) \mathbf{e}_3^T H_\infty \mathbf{p}$$

$$\eta = \frac{\mathbf{e}_3^T \mathbf{p}'\mathbf{p}'^T (\mathbf{h}_1 \times \mathbf{h}_2)}{\det(H_\infty) \mathbf{e}_3^T H_\infty \mathbf{p}} = \frac{\mathbf{p}'^T (\mathbf{h}_1 \times \mathbf{h}_2)}{\det(H_\infty) \tilde{\mathbf{h}}_3^T \mathbf{p}}$$

用 \mathbf{e}_1^T 左乘(5.5.13)式, 并由上式, 可得到

$$\mathbf{e}_1^T (\mathbf{h}_1 \times \mathbf{h}_2) f_u'^2 = -\mathbf{e}_1^T \mathbf{p}'\mathbf{p}'^T (\mathbf{h}_1 \times \mathbf{h}_2) + \eta \det(H_\infty) \mathbf{e}_1^T H_\infty \mathbf{p}$$

$$f_u'^2 = \frac{\eta \det(H_\infty) \mathbf{e}_1^T H_\infty \mathbf{p} - \mathbf{e}_1^T \mathbf{p}'\mathbf{p}'^T (\mathbf{h}_1 \times \mathbf{h}_2)}{\mathbf{e}_1^T (\mathbf{h}_1 \times \mathbf{h}_2)} = \frac{\mathbf{p}'^T (\mathbf{h}_1 \times \mathbf{h}_2) (\tilde{\mathbf{h}}_1^T \mathbf{p} - \tilde{\mathbf{h}}_3^T \mathbf{p} \mathbf{e}_1^T \mathbf{p}')}{\tilde{\mathbf{h}}_3^T \mathbf{p} \mathbf{e}_1^T (\mathbf{h}_1 \times \mathbf{h}_2)}$$

同理, 用 \mathbf{e}_3^T 左乘(5.5.13)式, 可以得到

$$f_v'^2 = \frac{\mathbf{p}'^T (\mathbf{h}_1 \times \mathbf{h}_2) (\tilde{\mathbf{h}}_2^T \mathbf{p} - \tilde{\mathbf{h}}_3^T \mathbf{p} \mathbf{e}_2^T \mathbf{p}')}{\tilde{\mathbf{h}}_3^T \mathbf{p} \mathbf{e}_2^T (\mathbf{h}_1 \times \mathbf{h}_2)}$$

用类似的方法, 可求出 f_u, f_v :

$$f_u^2 = \frac{\mathbf{p}^T (\mathbf{g}_1 \times \mathbf{g}_2) (\tilde{\mathbf{g}}_1^T \mathbf{p} - \tilde{\mathbf{g}}_3^T \mathbf{p} \mathbf{e}_1^T \mathbf{p}')}{\tilde{\mathbf{g}}_3^T \mathbf{p}' \mathbf{e}_1^T (\mathbf{g}_1 \times \mathbf{g}_2)}$$

$$f_u^2 = \frac{\mathbf{p}^T (\mathbf{g}_1 \times \mathbf{g}_2) (\tilde{\mathbf{g}}_2^T \mathbf{p} - \tilde{\mathbf{g}}_3^T \mathbf{p}' \mathbf{e}_2^T \mathbf{p})}{\tilde{\mathbf{g}}_3^T \mathbf{p}' \mathbf{e}_2^T (\mathbf{g}_1 \times \mathbf{g}_2)}$$

证毕。

该命题的重要性在于，两个摄像机的尺度因子都可以表示为主点的显函数。当主点已知时，可由它们计算出尺度因子。

5.6 绝对二次曲面与摄像机内参数

5.6.1 绝对二次曲面约束

基本约束方程

由命题 3.2.8 和 3.2.9，绝对二次曲面 Q_∞^* 的投影是 DIAC，即

$$s \varpi^* = P Q_\infty^* P^T \quad (5.6.1)$$

式(5.6.1)就是基于绝对二次曲面的自标定的基本约束方程^{*)}，在这里齐次因子 s 被显式给出。

基于绝对二次曲面 Q_∞^* 的自标定的基本思想与基于绝对二次曲线是类似的，即通过(5.6.1)式中的摄像机矩阵 P 把 ϖ^* 上的约束转换到 Q_∞^* 上的约束。一旦得到 Q_∞^* 在当前射影坐标系下的矩阵表示后，反过来再通过式(5.6.1)就可得到 ϖ^* 中的未知参数，从而能标定出摄像机的所有内参数。因此，基于 Q_∞^* 的自标定的关键是确定绝对二次曲面在当前射影坐标系下的矩阵表示。由于绝对二次曲面 Q_∞^* 有 8 个自由度，因此需要关于它的 8 个约束方程才能非线性求解，这里的非线性是由于使用了 Q_∞^* 的退化性质即它的行列式等于零的约束。如果要想线性求解 Q_∞^* ，则至少需要 8 个以上的线性约束。因此，基于绝对二次曲面的自标定通常被用于图像序列。

变化内参数

在变化内参数的情况下，通常都是直接由内参数元素的约束来估计 Q_∞^* ，再根据基本约束方程(5.6.1)得到摄像机内参数矩阵。

在摄像机都有零畸变因子的情况下，可得到 ϖ_j^* 元素之间如下约束

$$(\varpi_j^*)_{12} (\varpi_j^*)_{33} = (\varpi_j^*)_{23} (\varpi_j^*)_{13}$$

^{*)} 基本约束方程中的投影矩阵是射影意义下的摄像机矩阵，对于多幅图像可使用点对应进行计算。如给定两幅图像，可利用基本矩阵进行计算，见 4.4 节；给定 3 幅图像，可利用三焦张量进行计算，见 12.2 节；给定 3 幅以上的图像，可应用矩阵分解方法进行计算，见 11.3 节。

这导致关于 Q_{∞}^* 的一个二次方程：

$$(P_j Q_{\infty}^* P_j^T)_{12} (P_j Q_{\infty}^* P_j^T)_{33} = (P_j Q_{\infty}^* P_j^T)_{23} (P_j Q_{\infty}^* P_j^T)_{13} \quad (5.6.2)$$

从 n 幅图像中，可以得到 n 个二次方程。然而，还有一个另外的约束方程 $\det Q_{\infty}^* = 0$ 能被使用，因为绝对二次曲面是退化的。由于 Q_{∞}^* 是对称的，从而有 10 个齐次线性参数，即位于对角线及其上方的 10 个元素。因此，在仅知摄像机有零畸变因子的情况下，至少需要 8 幅图像才有可能非线性地计算出 Q_{∞}^* 。

在已知摄像机的主点的情况下，可以选择图像坐标系的原点与主点相重合，即 $u=0, v=0$ 。由此可得到： $(\omega_j^*)_{13} = (\omega_j^*)_{23} = 0$ ，于是由基本约束(5.6.1)可得到两个线性约束方程：

$$(P_j Q_{\infty}^* P_j^T)_{13} = (P_j Q_{\infty}^* P_j^T)_{23} = 0 \quad (5.6.3)$$

将 Q_{∞}^* 的 10 个元素用一个 10 维向量 \mathbf{q} 表示，按通常方法将 Q_{∞}^* 的线性方程(5.6.3)组合在一起，形成一个形如 $A\mathbf{q}=0$ 的线性方程。对于每幅图像，(5.6.3)的两个方程提供矩阵 A 的两行。从 5 幅图像总共得到 10 个方程，由此得到一个线性解。这个线性解可通过 SVD 分解得到（见 8.4 节）。从 4 幅图像所产生的 8 个方程，它的解是单参数簇。于是，由条件 $\det Q_{\infty}^* = 0$ 所给出的 4 次方程可知 Q_{∞}^* 至多有四个解。

如果除了已知主点外，还知道像素的长宽比并且畸变参数为零，也就是说仅在焦距未知的情况下，可知每个图像有 4 个关于 Q_{∞}^* 的线性约束。在有两幅图像时，可导致 8 个线性约束，再利用约束 $\det Q_{\infty}^* = 0$ ，则 Q_{∞}^* 至多有四个解。如果有三幅图像，则能唯一线性求解 Q_{∞}^* 。

恒定内参数

如果所有摄像机的内参数是相同的，即对于所有的 i 和 j 有 $\omega_i^* = \omega_j^*$ 。此时关于绝对二次曲面的约束方程化为

$$P_i Q_{\infty}^* P_i^T = s_{ij} P_j Q_{\infty}^* P_j^T \quad (5.6.4)$$

消去齐次因子后，可得到关于 Q_{∞}^* 的五个二次方程。于是给定三幅图像，总共能产生 10 个方程。因此，在恒定内参数的情况下，给定三幅图像可非线性求解 Q_{∞}^* 。

5.6.2 自标定约束的等价性

基于 Kruppa 方程的约束、基于绝对二次曲线和基于绝对二次曲面的约束是摄像机自标定的三种重要方法，现在讨论它们之间的关系。

基于 Q_∞^* 与 Ω_∞ 约束的等价性

假定各个视点的摄像机矩阵为

$$P_0 = (I, \mathbf{0}), \quad P_j = (H_j, \mathbf{a}_j), \quad j=0,1,2,\dots,n \quad (5.6.5)$$

其中, 视点 0 作为参考视点, 它的内参数记为 K , 相应的 DIAC 记为 ϖ^* 。第 j 个视点的内参数记为 K_j , 相应的 DIAC 记为 ϖ_j^* , 则有基于绝对二次曲面的约束方程:

$$s_j \varpi_j^* = P_j Q_\infty^* P_j^T, \quad j=0,1,2,\dots,n \quad (5.6.6)$$

这里 $\varpi_0^* = \varpi^*$ 。

如果 $\{P_0, P_j\}$ 的射影坐标系到欧氏坐标系的变换是 H^{-1} , 则必有 $Q_\infty^* = H \tilde{I} H^T$ 。取欧氏坐标系与参

考视点的摄像机坐标系重合, 则必有 $P_0 = (I, \mathbf{0})H = (K, \mathbf{0})$, 所以 $H = \begin{pmatrix} K & \mathbf{0} \\ \mathbf{v}^T & 1 \end{pmatrix}$ 。于是,

$$Q_\infty^* = \begin{pmatrix} KK^T & K\mathbf{v} \\ (K\mathbf{v})^T & 1 \end{pmatrix}$$

并且无穷远平面 π_∞ 在当前射影坐标系下的坐标为

$$\pi_\infty = H^{-T} \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} = \begin{pmatrix} K^{-T} & -K^{-T}\mathbf{v} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} = \begin{pmatrix} -K^{-T}\mathbf{v} \\ 1 \end{pmatrix}$$

令 $\mathbf{p} = -K^{-T}\mathbf{v}$, 则 $\pi_\infty = (\mathbf{p}^T, 1)^T$, $\mathbf{v} = -K^T\mathbf{p}$ 。于是,

$$Q_\infty^* = \begin{pmatrix} KK^T & K\mathbf{v} \\ (K\mathbf{v})^T & 1 \end{pmatrix} = \begin{pmatrix} KK^T & -KK^T\mathbf{p} \\ -(KK^T\mathbf{p})^T & 1 \end{pmatrix} = \begin{pmatrix} \varpi^* & -\varpi^*\mathbf{p} \\ -(\varpi^*\mathbf{p})^T & 1 \end{pmatrix}$$

将上式代入(5.6.6)式, 得到

$$s_j \varpi_j^* = (H_j, \mathbf{a}_j) \begin{pmatrix} \varpi^* & -\varpi^*\mathbf{p} \\ -(\varpi^*\mathbf{p})^T & 1 \end{pmatrix} (H_j, \mathbf{a}_j)^T = (H_j - \mathbf{a}_j \mathbf{p}^T) \varpi^* (H_j - \mathbf{a}_j \mathbf{p}^T)^T$$

即,

$$s_j \varpi_j^* = H_{j\infty} \varpi^* (H_{j\infty})^T$$

这样, 就得到了基于绝对二次曲线的自标定约束方程。反向推导也成立。因此, 基于绝对二次曲线与绝对二次曲面的约束是等价的。

基于 Q_{∞}^* 约束与 Kruppa 方程约束之间的关系

由于从绝对二次曲线约束可推出 Kruppa 方程约束, 而绝对二次曲面约束与绝对二次曲线约束等价, 所以由绝对二次曲面约束可导致 Kruppa 方程约束。但是, 从下面的讨论将看出从 Kruppa 方程约束不能导致绝对二次曲面约束, 它们之间的差别仅在二次曲面的退化性约束。

假定有 N 幅图像。由这 N 幅图像可导致 $N(N-1)/2$ 个 Kruppa 方程所构成的方程组:

$$F_{ij}\varpi_i^*F_{ij}^T = s_{ij}[\mathbf{e}_{ij}]_{\times}\varpi_j^*[\mathbf{e}_{ij}]_{\times}^T, \quad 1 \leq i < j \leq N \quad (5.6.6)$$

其中: F_{ij} 是第 i 幅图像与第 j 幅图像的基本矩阵; \mathbf{e}_{ij} 是第 i 幅图像关于第 j 幅图像的极点, 即第 i 个摄像机的光心在第 j 幅图像上的投影; s_{ij} 是未知的非零常数因子; $\varpi_j^* = K_j K_j^T$ 是第 j 个摄像机的 DIAC。

令 P_i 是第 i 幅图像的射影意义下的摄像机矩阵, 即

$$P_i = (I, \mathbf{0}), \quad P_i = (H_{1i}, \mathbf{e}_{1i})$$

其中 H_{1i} 是第一幅与第 i 幅图像关于某个空间平面 π 的单应矩阵, \mathbf{e}_{1i} 是第一幅图像关于第 i 幅图像的极点。可以通过 N 幅图像的射影重构得到它们。于是, 第一幅图像与第 i 幅图像之间的基本矩阵为

$$F_{1i} = [\mathbf{e}_{1i}]_{\times} H_{1i}$$

显然, 第 i 幅图像的光心在第 j 幅图像的极点可以表示为

$$\mathbf{e}_{ij} = \mathbf{e}_{1j} - H_{1j} H_{1i}^{-1} \mathbf{e}_{1i}, \quad i \neq j, i \neq 1$$

第 i, j 幅图像关于平面 π 的单应矩阵可以表示为

$$H_{ij} = H_{1j} H_{1i}^{-1}, \quad i \neq j, i \neq 1$$

所以, 第 i, j 幅图像之间的基本矩阵为

$$F_{ij} = [\mathbf{e}_{ij}]_{\times} H_{1j} H_{1i}^{-1}$$

将上述特定 F_{ij} , \mathbf{e}_{ij} 代入(5.6.6)得到 Kruppa 方程的特定形式, 并称它为一致化 Kruppa 方程组。

不仅能证明一致化 Kruppa 方程与一般 Kruppa 方程是等价的, 而且还能进一步得到下述命题:

命题 5.6.1 一致化 Kruppa 方程组与下述方程组等价:

$$s_{1i} \varpi_i^* = (H_i, \mathbf{e}_i) \begin{pmatrix} \varpi_1^* & \mathbf{a} - s\mathbf{g}_2 \\ (\mathbf{a} - s\mathbf{g}_2)^T & 2s \end{pmatrix} (H_i, \mathbf{e}_i)^T, \quad i = 2, 3, \dots, N \quad (5.6.7)$$

其中: $\mathbf{g}_2 = H_{12}^{-1} \mathbf{e}_{12}$; \mathbf{a} 是未知的三维向量, s 是未知标量。

这个命题的证明是相当复杂的, 这里从略。

记 $Q^* = \begin{pmatrix} \varpi_1^* & \mathbf{a} - s\mathbf{g}_2 \\ (\mathbf{a} - s\mathbf{g}_2)^T & 2s \end{pmatrix}$, 则它是三维空间中的二次曲面, 并且(5.6.7)可化为

$$s_i \varpi_i^* = (\mathbf{H}_i, \mathbf{e}_i) Q^* (\mathbf{H}_i, \mathbf{e}_i)^T, \quad i=1,2..N \quad (5.6.8)$$

式(5.6.8)的几何意义为: ϖ_i^* 是二次曲面 Q^* 在第 i 幅图像上的投影。而基于绝对二次曲面 Q_∞^* 的约束是

$$s_i \varpi_i^* = (\mathbf{H}_i, \mathbf{e}_i) Q_\infty^* (\mathbf{H}_i, \mathbf{e}_i)^T, \quad i=1,2..N, \quad (5.6.9)$$

且绝对二次曲面 Q_∞^* 满足退化性约束 $\det Q_\infty^* = 0$ 。

从上面的讨论可以看出: Kruppa 方程约束与绝对二次曲面 Q_∞^* 约束之间的差别仅在于二次曲面的退化性约束。事实上, 如果在(5.6.8)中增加约束 $\det(Q^*) = 0$, 则必有

$$\mathbf{a}_2 = \varpi_1^* \mathbf{a} + \frac{\mathbf{a}^T \varpi_1^* \mathbf{a}}{2} \mathbf{H}_2^{-1} \mathbf{e}_2, \quad s = \frac{\mathbf{a}^T \varpi_1^* \mathbf{a}}{2}$$

从而有

$$Q^* = \begin{pmatrix} \varpi_1^* & \mathbf{a} - s\mathbf{g}_2 \\ (\mathbf{a} - s\mathbf{g}_2)^T & 2s \end{pmatrix} = \begin{pmatrix} \varpi_1^* & \varpi_1^* \mathbf{a} \\ (\varpi_1^* \mathbf{a})^T & \mathbf{a}^T \varpi_1^* \mathbf{a} \end{pmatrix} = Q_\infty^*$$

代入式(5.6.8)就得到绝对二次曲面 Q_∞^* 约束(5.6.9)。

Kruppa 方程约束的独立数

给定 N 幅图像, Kruppa 方程关于摄像机内参数能构成多个独立的约束? 对此, 有下述结论:

命题 5.6.2 假定摄像机的运动为一般运动, 且 5 个内参数都是变化的。对于 N 个视点的图像, 我们有:

- (1) Kruppa 方程关于摄像机内参数的独立约束数为 $5N-9$;
- (2) 绝对二次曲面(或绝对二次曲线)关于摄像机内参数的独立约束数为 $5N-8$ 。

该命题的证明沉长而复杂, 我们略去证明。

6. 三维重构理论

如何从图像点恢复与它对应的空间点在世界坐标系中的坐标，称为三维重构问题，它是三维计算机视觉中的核心问题。摄像机投影变换是从三维空间到二维平面的(退化)射影变换，无法从空间景物的单幅图像得到它的几何结构。事实上，图像点的反投影射线上的任何空间点的投影都与该图像点重合，因此不能确定与图像点对应的空间点的确切位置。所以，重构问题的确切描述是：从空间景物的两幅图像或多幅图像恢复景物的几何结构。在已知摄像机矩阵的情况下(通常也称为图像是已标定的)，重构的基本方法是所谓的三角原理。本章主要讨论未标定图像的三维重构问题。从未标定图像重构三维物体，主要方法是分层重构方法：首先从图像获得三维物体的射影结构(即射影重构)，然后利用先验知识(摄像机的运动或景物几何结构知识)由射影重构获得三维物体的仿射结构(即仿射重构)，最后进一步利用先验知识由仿射重构获得三维物体的度量结构(即度量重构)。本章主要目的是应用射影几何来建立三维重构的理论。对于射影重构，不需要图像以外的知识，仅从图像就能获得。对于仿射重构，如果摄像机在运动过程中保持内参数不变的话，理论上由射影重构能获得仿射重构，不仅如此，在这种情况下还能得到度量重构；如果对摄像机内参数的知识一无所知的话，必须利用摄像机的某种运动信息或景物的某种几何结构信息才能获得仿射重构。那么，究竟需要何种知识才能得到仿射重构呢？本章研究表明，仿射重构、确定无穷远平面在射影重构坐标系的坐标以及确定无穷远单应矩阵三者是相互等价的，因此仿射重构所需要的知识是足以确定无穷远平面或无穷远单应矩阵的知识。对于度量重构，它与确定绝对二次曲线、确定绝对二次曲面、确定摄像机内参数是相互等价的，因此度量重构所需要的知识是足以能确定摄像机内参数的知识，或确定绝对二次曲线的知识，或确定绝对二次曲面的知识。

6.1 三角原理

给定两幅图像在同一世界坐标系下的摄像机矩阵 P 和 P' ， $m \leftrightarrow m'$ 是两幅图像的一个点对应，即它们满足极几何约束 $m'^T F m = 0$ ，现在要根据 P 和 P' 计算点对应 $m \leftrightarrow m'$ 的空间点。正如第 3 章所指出， m 的反投影射线 $l_m^{(b)}$ 与 m' 的反投影射线 $l_{m'}^{(b)}$ 确定了通过两摄像机光心的一张平面 π_p ，即一张极

平面，所以两条反投影线 $l_m^{(b)}$ ， $l_{m'}^{(b)}$ 必相交于一个空间点。也就是说，对应点的反投影射线与两个摄像机的基线构成一个三角形，这个三角形的顶点是两个摄像机的光心和两条反投影线的交点，这个交点正是我们要确定的空间点，如图 6.1.1 所示。

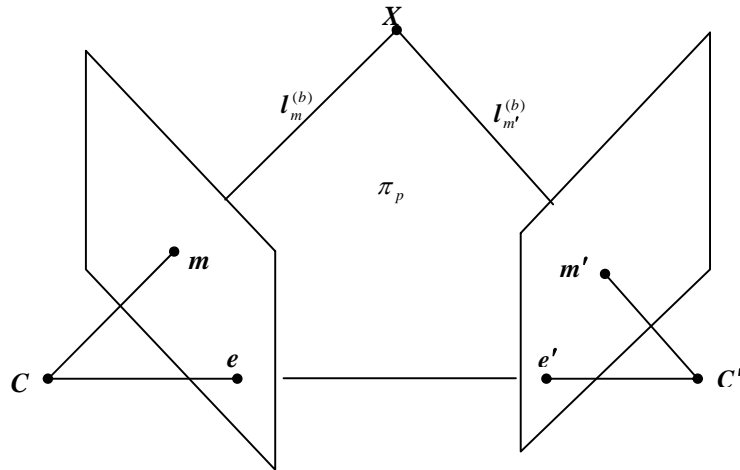


图 6.1.1：三角原理：空间点是图象点 x 和 x' 反向投影射线的交点。

唯一例外的情况是，三维空间中在两个摄像机基线上的点不可能由它们的图像点对应(两幅图像上的极点)而得到恢复，因为在这种情况下，两条反投影的射线与基线重合，所以空间点不能唯一确定。

空间点的重构可用下述方法来实现：令

$$P = \begin{pmatrix} p^{1T} \\ p^{2T} \\ p^{3T} \end{pmatrix}, P' = \begin{pmatrix} p'^{1T} \\ p'^{2T} \\ p'^{3T} \end{pmatrix} \quad (6.1.1)$$

分别是两幅图像的摄像机矩阵。若 $m = (u, v, 1)^T \leftrightarrow m' = (u', v', 1)^T$ 是一个点对应，则它们的反投影射线在世界坐标系中的方程分别为

$$\begin{cases} p^{2T} X - v p^{3T} X = 0 \\ p^{1T} X - u p^{3T} X = 0 \end{cases} \quad (6.1.2a)$$

$$\begin{cases} p'^{2T} X - v' p'^{3T} X = 0 \\ p'^{1T} X - u' p'^{3T} X = 0 \end{cases} \quad (6.1.2b)$$

这里 \mathbf{X} 是空间点的齐次坐标。根据三角原理，点对应 $m \leftrightarrow m'$ 的空间点 \mathbf{X} 是下述方程的解：

$$\begin{pmatrix} \mathbf{p}^{2T} & -v\mathbf{p}^{3T} \\ \mathbf{p}^{1T} & -u\mathbf{p}^{3T} \\ \mathbf{p}'^{2T} & -v'\mathbf{p}'^{3T} \\ \mathbf{p}'^{1T} & -v'\mathbf{p}'^{3T} \end{pmatrix} \mathbf{X} = 0 \quad (6.1.3)$$

因此，通过求解这个线性方程可确定空间点 \mathbf{X} 。

三角原理的重构结果，是否为空间景物的欧氏结构取决于摄像机所处的世界坐标系。也就是说，如果世界坐标系是欧氏的，则重构结果是空间景物的欧氏结构；否则如果世界坐标系是仿射的(或射影的)，则重构结果是空间景物的仿射结构(或射影结构)。

6.2 基本矩阵与射影重构

由上节讨论可知，从图像得求解摄像机在同一世界坐标系下的投影矩阵，构成三维重构的关键问题。在摄像机内参数和运动参数都未知的情况下，极几何约束是两幅图像间的唯一约束，从第 3 章可知这个约束可由基本矩阵 F 来代数刻画，并且可以由下述方法构造相应的摄像机矩阵。令 $F = [e']_{\times} H$ 是基本矩阵 F 的一个分解，对应两幅图像的典型摄像机矩阵对是

$$\tilde{P} = (I, \mathbf{0}), \tilde{P}' = (H, e') \quad (6.2.1)$$

这样，根据三角原理由图像点对应就能获得景物的几何结构。但这个结构一般是景物在某个射影空间中的结构，即景物的射影结构。因此，在理论上，从两幅图像的点对应就能得到空间景物的射影结构，而不需要任何有关摄像机内参数和相对运动以及空间景物结构等信息。

由于人们总是习惯在三维欧氏空间中描述物体，所以希望在获得景物射影结构的基础上给出景物的欧氏结构。由命题 4.4.6，与摄像机矩阵对(6.2.1)射影相关的所有摄像机矩阵对均可表示为

$$\tilde{P} = (I, \mathbf{0})M^{-1}, \tilde{P}' = (H, e')M^{-1} \quad (6.2.2)$$

其中， M 为 4 阶可逆矩阵，而 $\mathbf{X}' = M\mathbf{X}$ 表示三维空间点的射影变换。因欧氏坐标系是一种特殊的射影坐标系，因此对应于欧氏坐标系的摄像机矩阵必包含在(6.2.2)之中，即存在射影变换 M 使得式(6.2.2)是某个欧氏坐标系下的摄像机矩阵对。如何求解这样的射影变换 M ？这一问题可以分层解决，即先求解使得式(6.2.2)为某个仿射坐标系下摄像机矩阵对的射影变换 M ，得到景物的仿射结构，然后在仿射结构的基础上，求解一个仿射变换 M 使得式(6.2.2)为某个欧氏坐标系下的摄像机矩阵对，

从而最终获得景物的欧氏结构。然而，在摄像机内参数和运动参数都未知的情况下，是无法获得这样的射影(或仿射)变换 M ，那么在什么条件下能够求解出这样的射影变换 M ？这就是以后各节所要讨论的主题。

6.3 无穷远平面与仿射重构

定义 6.3.1 令 P_i, P'_i 观察三维欧氏空间点集 $\{X_j^{(i)}\}$ 的实际摄像机对，并记为 $[P_i, P'_i, \{X_j^{(i)}\}]$ ，它们观察空间点 $\{X_j^{(i)}\}$ 所得到的图像点对应记为 $\{m_j \leftrightarrow m'_j\}$ 。设 $[P, P', \{X_j\}]$ 是由图像点对应 $\{m_j \leftrightarrow m'_j\}$ 所得到的一个三维重构，如果存在仿射变换 A 使得

$$(P_i, P'_i) = (PA^{-1}, P'A^{-1}), X_j^{(i)} = AX_j$$

则称 $[P, P', \{X_j\}]$ 为仿射重构，并记为 $[P_A, P'_A, \{X_j^{(A)}\}]$ 。

仿射重构 $[P_A, P'_A, \{X_j^{(A)}\}]$ 中的点 $\{X_j^{(A)}\}$ 是三维仿射空间中的点，它由摄像机对 $[P_A, P'_A]$ 和图像点对应 $\{m_j \leftrightarrow m'_j\}$ 唯一确定，因此有时也简称摄像机对 $[P_A, P'_A]$ 为仿射重构，并且称重构所在的仿射坐标系为仿射重构坐标系。

定理 6.3.1(仿射重构定理) 令 $[P, P', \{X_j\}]$ 是景物的一个射影重构，则三维射影变换 H 使得 $[PH^{-1}, P'H^{-1}, \{HX_j\}]$ 为仿射重构的充要条件是 H 具有下述形式：

$$H = \begin{pmatrix} A & b \\ \mathbf{p}^T & w \end{pmatrix} \quad (6.3.1)$$

其中 $\pi_{\infty}^{(p)} = (\mathbf{p}^T, w)^T$ 为无穷远平面在射影重构坐标系下的坐标。

证明 必要性： 因为仿射重构空间与欧氏空间相差一个仿射变换，即空间点的仿射重构坐标与它的欧氏坐标之间满足一个仿射变换，并且仿射变换是保持无穷远平面不动的，所以，无无穷远平面在仿射重构空间中的坐标必为 $\pi_{\infty}^{(i)} = (0, 0, 0, 1)^T$ 。于是，根据平面的射影变换规则，立即得到：

$$\pi_{\infty}^{(p)} = (\mathbf{p}^T, w)^T = H^T \pi_{\infty}^{(i)} = \mathbf{h}^{4T} \quad (\mathbf{h}^4 \text{ 表示 } H \text{ 的第 4 行向量})$$

因此， H 必有(6.3.1)的形式。

充分性： 由于射影变换 H 使得

$$H^T \pi_{\infty}^{(i)} = (\mathbf{p}^T, w)^T = \pi_{\infty}^{(p)}$$

所以， $H^{-T} \pi_{\infty}^{(p)} = \pi_{\infty}^{(i)} = (0^T, 1)^T$ ，即 H 将无穷远平面在射影重构空间中的坐标 $\pi_{\infty}^{(p)} = (\mathbf{p}^T, w)^T$ 变换为 $\pi_{\infty}^{(i)} = (0, 0, 0, 1)^T$ 。于是，无穷远平面在重构 $[PH^{-1}, P'H^{-1}, \{HX_j\}]$ 坐标系下和在欧氏坐标系下有相

同的坐标 $\pi_{\infty}^{(i)} = (0, 0, 0, 1)^T$ 。这表明重构 $[P H^{-1}, P' H^{-1}, \{HX_j\}]$ 坐标系与欧氏坐标系之间满足一个仿射变换, 因此它是一个仿射重构。证毕。

定理 6.3.1 表明, 从射影重构确定仿射重构与在射影重构空间中确定无穷远平面的射影坐标是等价的。

例 6.3.1 令 $P = (I, \mathbf{0})$, $P' = (A, \mathbf{e}')$ 是一个射影重构, 无穷远平面在这个重构空间中的坐标为 $\pi_{\infty}^{(p)} = (\mathbf{p}^T, 1)^T$, 取

$$H = \begin{pmatrix} I & \mathbf{0} \\ \mathbf{p}^T & 1 \end{pmatrix}$$

则 $PH^{-1} = (I, \mathbf{0})$, $P'H^{-1} = (A - \mathbf{e}'\mathbf{p}^T, \mathbf{e}')$ 是一个仿射重构。由于 $H_{\infty} = A - \mathbf{e}'\mathbf{p}^T$ 是无穷远单应矩阵, 因此这个例子说明一旦知道无穷远单应矩阵, 就可以得到一个仿射重构。事实上仿射重构与确定无穷远单应矩阵也是等价的。

定理 6.3.2 确定仿射重构与确定无穷远单应矩阵 H_{∞} 是等价的。

证明 令摄像机对 $P=(A, \mathbf{b})$, $P'=(A', \mathbf{b}')$ 是一个仿射重构, 下面证明无穷远单应矩阵 $H_{\infty} = A'A^{-1}$ 。因为无穷远平面上的点 $X=(\mathbf{x}^T, 0)^T$, 在第一个摄像机下的图像点是 $\mathbf{m}=\mathbf{A}\mathbf{x}$, 在第二个摄像机下的图像点是 $\mathbf{m}' = A'\mathbf{x}$, 因此对于无穷远平面上的点 X , 在两幅图像上的点对应 $\mathbf{m} \leftrightarrow \mathbf{m}'$, 有 $\mathbf{m}' = A'A^{-1}\mathbf{m}$ 。所以, $H_{\infty} = A'A^{-1}$ 。

反之, 若无穷远单应矩阵 $H_{\infty} = A'A^{-1}$, 证明摄像机对 $P=(A, \mathbf{b})$, $P'=(A', \mathbf{b}')$ 是一个仿射重构。对这个摄像机对作仿射变换:

$$H = \begin{pmatrix} A & \mathbf{b} \\ \mathbf{0}^T & 1 \end{pmatrix}$$

得到

$$PH^{-1}=(A, \mathbf{b})H^{-1}=(I, \mathbf{0}), \quad P'H^{-1}=(A', \mathbf{b}')H^{-1}=(A'A^{-1}, A'A^{-1}\mathbf{b} + \mathbf{b}')=(H_{\infty}, \mathbf{e}')$$

由例 6.3.1, 摄像机对 $[PH^{-1}, P'H^{-1}]$ 是一个仿射重构。因此, 摄像机对 $[P, P']$ 也是一个仿射重构。证毕。

推论 6.3.1 摄像机对 $P = (I, \mathbf{0})$, $P' = (A, \mathbf{e}')$ 为仿射重构的充要条件是 $H_{\infty} = A$ 。

定理 6.3.1 和 6.3.2 表明, 如果能确定无穷远平面在射影重构空间中的射影坐标, 或无穷远单应矩阵, 则从射影重构就可以确定仿射重构, 因此确定无穷远平面的射影坐标或它的单应矩阵就构成了由射影重构确定仿射重构的关键问题。如果没有摄像机的知识(包括内参数和运动)或场景结构的知识作为约束条件, 则无法从射影重构, 或者说从图像(因为射影重构仅需要图像点对应), 所提供

的信息来确定无穷远平面的射影坐标或它的单应矩阵。

例 6.3.2 利用平行性由射影重构确定仿射重构。

平行直线 平行性是仿射变换的不变性质，因此可以通过景物中的平行直线来确定景物的仿射重构。空间两条平行直线的交点是无穷远平面上的点，该交点的图像是这组平行线的隐消点，而且是两条像直线的交点。如果在场景中能够确定不在同一平面上的不同方向上的三组平行直线，则通过这三组平行直线可得到无穷远平面上的三个不同点。又因三点决定一平面，故这个信息足够决定无穷远平面。因此，由射影重构我们能确定仿射重构。

平行平面 空间两个平行平面的交线是无穷远平面上的直线，该直线的图像是平行平面的消隐线。隐消线可以通过下述方法来计算：记两个平行平面的单应矩阵分别为 H_1 和 H_2 ，平行平面在两幅图像上的隐消线分别记为 l 和 l' ，则必有 $H_1^{-T}l = l' = H_2^{-T}l$ ，因此 $H_2^T H_1^{-T}l = l$ ，即隐消线 l 是 $H_2^T H_1^{-T}$ 的一维特征子空间，同样 l' 是 $H_1^T H_2^{-T}$ 的一维特征子空间。并且这样的特征子空间是唯一的^{*)}。于是分别求解 $H_2^T H_1^{-T}$ ， $H_1^T H_2^{-T}$ 的单重特征值的特征向量可确定隐消线 l 和 l' 。当已知两组不平行的平行平面时，通过这两组平行平面所确定两个(消隐)线对应，就可以在射影重构的基础上确定无穷远单应矩阵，因此由推论 6.3.1 就可以确定仿射重构。

6.4 绝对二次曲线与度量重构

定义 6.4.1 令 P_i, P'_i 观察三维欧氏空间点集 $\{X_j^{(0)}\}$ 的实际摄像机对，并记为 $[P_i, P'_i, \{X_j^{(0)}\}]$ ，它们观察空间点 $\{X_j^{(0)}\}$ 所得到的图像点对应记为 $\{m_j \leftrightarrow m'_j\}$ 。设 $[P, P', \{X_j\}]$ 是由图像点对应 $\{m_j \leftrightarrow m'_j\}$ 所得到的一个三维重构，如果存在相似变换 S 使得

$$(P_i, P'_i) = (PS^{-1}, P'S^{-1}), X_j^{(0)} = SX_j$$

则称 $[P, P', \{X_j\}]$ 为欧氏重构，并记为 $[P_e, P'_e, \{X_j^{(e)}\}]$ 。

欧氏重构 $[P_e, P'_e, \{X_j^{(e)}\}]$ 中的点 $\{X_j^{(e)}\}$ 是三维欧氏空间中的点，它由摄像机对 $[P_e, P'_e]$ 和图像点

^{*)}事实上，隐消线 l 是 $H_2^T H_1^{-T}$ 的一维特征子空间，而第一幅图像上通过极点的线束，即极线束，是 $H_2^T H_1^{-T}$ 的二维特征子空间。我们只须证明极线束是 $H_2^T H_1^{-T}$ 的二维特征子空间。假定第一幅图像上异于隐消线 l 的直线 \hat{l} 满足 $H_2^T H_1^{-T} \hat{l} = \hat{l}$ ，则必有 $H_1^{-T} \hat{l} = H_2^{-T} \hat{l} = \hat{l}'$ ，于是 \hat{l} 的反投影平面与两个平行平面的交线 L_1, L_2 在第二幅图像上有相同的像直线 \hat{l}' 。显然 $L_1 // L_2$ 且不重合，且它们所确定的空间平面 $\hat{\pi}$ 必过第二个摄像机的光心，因此平面 $\hat{\pi}$ 是极平面，故 \hat{l}, \hat{l}' 是一对对极线。反之，如果 \hat{l} 是第一幅图像上的极线，则显然有 $H_2^T H_1^{-T} \hat{l} = \hat{l}$ 。因此第一幅图像上的极点的线束构成 $H_2^T H_1^{-T}$ 的一个二维特征子空间。于是，隐消线 l 是 $H_2^T H_1^{-T}$ 的一维特征子空间，而极线束是 $H_2^T H_1^{-T}$ 的二维特征子空间。注意：对空间中两张相交的平面，只要将隐消线替换为两平面交线的图像，对应的结论也成立。

对应 $\{m_j \leftrightarrow m'_j\}$ 唯一确定，因此有时也简称摄像机对 $[P_e, P'_e]$ 为欧氏重构，并且称重构所在的欧氏坐标系为欧氏重构坐标系。欧氏重构也称为相似重构或度量重构。

射影重构是最一般意义下的重构，它仅保持空间结构的射影性质，如交比；仿射重构是一个中间层次，它保持空间结构的仿射性质，如平行、体积比等性质；度量重构是一种最精细的重构，是在没有绝对空间结构信息情况下是最好的一种重构，它保持空间几何结构的相似性质。

定理 6.4.1 (度量重构定理 I) 设 $[P_A = (I, \mathbf{0}), P'_A = (H_\infty, \mathbf{e}')]^T$ 是一个仿射重构，则仿射变换 H 使得 $[P_A H^{-1}, P'_A H^{-1}]^T$ 为度量重构的充要条件是

$$H = H_s \begin{pmatrix} K^{-1} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix}$$

其中， K 是第一个摄像机的内参数矩阵， H_s 为相似变换矩阵。

证明 令 $\varpi = K^{-T} K^{-1}$ 是第一个摄像机的 IAC，注意 IAC 仅与摄像机的内参数有关，而与摄像机的方位以及所处的坐标系无关，即它与任何重构的变换都是无关的。在仿射重构坐标系下，对任意位于绝对二次曲线上的点 $(\mathbf{x}^T, 0)^T \in \Omega_\infty$ ，它在摄像机 $P_A H^{-1}$ 下的像点记为 \mathbf{m} ，则必有

$$(\mathbf{x}^T, 0) H^{-T} P_A^T \varpi P_A H^{-1} \begin{pmatrix} \mathbf{x} \\ 0 \end{pmatrix} = \mathbf{m}^T \varpi \mathbf{m} = 0 \quad (6.4.1)$$

记 $H = \begin{pmatrix} A & \mathbf{b} \\ \mathbf{0}^T & 1 \end{pmatrix}$ ，则

$$H^{-T} P_A^T \varpi P_A H^{-1} = \begin{pmatrix} A^{-T} \varpi A^{-1} & -A^{-T} \varpi A^{-1} \mathbf{b} \\ -\mathbf{b}^T A^{-T} \varpi A^{-1} & -\mathbf{b}^T A^{-T} \varpi A^{-1} \mathbf{b} \end{pmatrix}$$

将它代入式 (6.4.1)，就得到绝对二次曲线在重构 $[P_A H^{-1}, P_A H^{-1}]$ 空间中的方程

$$\mathbf{x}^T A^{-T} \varpi A^{-1} \mathbf{x} = 0$$

重构 $[P_A H^{-1}, P_A H^{-1}]$ 为度量重构的充要条件是该重构坐标系与欧氏坐标系相差一个相似变换，而一个变换为相似变换的充要条件是绝对二次曲线在这个变换下保持不动，因此重构 $[P_A H^{-1}, P_A H^{-1}]$ 为度量重构的充要条件是

$$A^{-T} \varpi A^{-1} = sI \quad (6.4.2)$$

其中 s 为任意的正实数。由上式可推知 $A = \sqrt{s} R K^{-1}$ ，其中 R 为任意旋转矩阵。故 $[P_A H^{-1}, P_A H^{-1}]$ 为度量重构的充要条件为

$$H = \begin{pmatrix} A & \mathbf{b} \\ \mathbf{0}^T & 1 \end{pmatrix} = \begin{pmatrix} sR & \mathbf{b} \\ \mathbf{0}^T & 1 \end{pmatrix} \begin{pmatrix} K^{-1} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix} = H_s \begin{pmatrix} K^{-1} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix}$$

证毕。

例 6.4.1 我们知道 $H_\infty = \alpha K' R K^{-1}$, $\mathbf{e}' = \beta K' \mathbf{t}$, 其中 (R, \mathbf{t}) 是摄像机的运动。取

$$H = \begin{pmatrix} K^{-1} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix}$$

由定理 6.4.1, 就得到一个度量重构的摄像机对:

$$P_A H^{-1} = K(I, \mathbf{0}), P'_A H^{-1} = K'(R, \hat{\boldsymbol{\alpha}}) \quad (6.4.2)$$

这是以第一个摄像机坐标系为世界坐标系的度量重构。

定理 6.4.2 由仿射重构实现度量重构与确定其中的一个摄像机内参数矩阵是等价的, 即与确定其中一个摄像机的 IAC 是等价的。

证明 不妨假定仿射重构为 $[P_A = (I, \mathbf{0}), P'_A = (H_\infty, \mathbf{e}')]$, 因为任一仿射重构均可以通过仿射变换变换为这种形式。若已知第一个摄像机内参数矩阵 K , 令

$$H = \begin{pmatrix} K^{-1} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix}$$

则 $[P_A H^{-1}, P'_A H^{-1}]$ 是一个度量重构。若已知第二个摄像机内参数矩阵 K' , 则必有

$$\varpi = \alpha H_\infty^T \varpi' H_\infty$$

所以对 $H_\infty^T \varpi' H_\infty$ 进行 Cholesky 分解, 可计算内参数矩阵 K , 这样仍能得到度量重构 $[P_A H^{-1}, P'_A H^{-1}]$ 。

反之, 若 $[P_A H^{-1}, P'_A H^{-1}]$ 是一个度量重构, 则由定理 6.4.1, 有

$$H = \begin{pmatrix} A & \mathbf{b} \\ \mathbf{0}^T & 1 \end{pmatrix} = H_s \begin{pmatrix} K^{-1} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix} = \begin{pmatrix} sR & \mathbf{b} \\ \mathbf{0}^T & 1 \end{pmatrix} \begin{pmatrix} K^{-1} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix}$$

所以, $A = sR K^{-1}$, 于是对 A 进行 RQ 分解可计算内参数矩阵 K 。证毕。

定理 6.4.1 和 6.4.2 表明, 如果能确定某个视点的摄像机内参数矩阵, 或某个视点摄像机的 IAC, 就可以直接通过仿射重构确定度量重构。根据命题 6.4.1, 从仿射重构能得到无穷远单应矩阵。如果每个视点的摄像机有相同的内参数, 就可以应用第 5.5.3 节的方法线性地求解摄像机内参数。因此,

在摄像机内参数恒定的情况下，由仿射重构而不需要其它任何知识就可以用线性方法实现景物的度量重构。如果摄像机内参数是变化的，则还需要来自仿射重构之外的知识来求解摄像机的内参数，而获得度量重构。

6.5 绝对二次曲面与度量重构

假定 $[P_i, \{X_j\}]$ 是一个射影重构，为了将它提升到度量重构，需要寻求一个三维射影变换 H 使得重构 $[P_i H^{-1}, \{HX_j\}]$ 的空间坐标系是一个欧氏坐标系。那么，这样的 H 具有什么性质呢？对此，有下述定理：

定理 6.5.1(度量重构定理 II) 假定 $[P_i, \{X_j\}]$ 是一个射影重构，则三维射影变换 H 使得 $[P_i H^{-1}, \{HX_j\}]$ 为度量重构的充要条件是绝对二次曲面 Q_∞^* 在射影重构坐标系的表示是 $Q_\infty^* = H^{-1} \tilde{I} H^{-T}$ 。其中 $\tilde{I} = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}$ 。

证明 必要性：若射影变换 H 使得 $[P_i H^{-1}, \{HX_j\}]$ 是一个度量重构，则绝对二次曲面在度量重构空间中的表示必为

$$\tilde{I} = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}$$

这是因为度量重构坐标系是一个欧氏坐标系。由于 H 是射影重构坐标系到度量重构坐标系的变换，根据二次曲面的射影变换规则，有 $H Q_\infty^* H^T = \tilde{I}$ 。因此， $Q_\infty^* = H^{-1} \tilde{I} H^{-T}$ 。

充分性：若射影变换 H 使得 $Q_\infty^* = H^{-1} \tilde{I} H^{-T}$ ，则绝对二次曲面在重构 $[P_i H^{-1}, \{HX_j\}]$ 坐标系下的表示必为 $H Q_\infty^* H^T = \tilde{I}$ ，因此重构 $[P_i H^{-1}, \{HX_j\}]$ 坐标系与欧氏坐标系相差一个相似变换，所以 $[P_i H^{-1}, \{HX_j\}]$ 是一个度量重构。证毕。

命题 6.5.1 假定射影重构为

$$P = (I, \theta), \quad P_j = (A_j, a_j), \quad j=1,2,\dots,n \quad (6.5.1)$$

若第一个摄像机内参数矩阵为 K ，无穷远平面在射影重构空间中的坐标为 $\pi_\infty = (p^T, 1)^T$ ，则 $[P_i H^{-1}, \{HX_j\}]$ 为度量重构的充要条件是三维射影变换 H 为下述形式：

$$H = H_s \begin{pmatrix} K^{-1} & \mathbf{0} \\ \mathbf{p}^T & 1 \end{pmatrix} \quad (6.5.2)$$

其中， H_s 为任一相似变换。

证明 记第一个摄像机的 DIAC 为 ϖ^* ，则有

$$\varpi^* = P Q_\infty^* P^T \quad (6.5.3)$$

所以， $Q_{33}^* = \varpi^* = K K^T$ ，其中 Q_{33} 是 Q_∞^* 左上角的 3 阶子矩阵。因此， Q_∞^* 必有下述形式：

$$Q_\infty^* = \begin{pmatrix} \varpi^* & \mathbf{q} \\ \mathbf{q}^T & w \end{pmatrix}$$

由于无穷远平面 $\pi_\infty = (\mathbf{p}^T, 1)^T$ 是 Q_∞^* 的左零空间，即 $(\mathbf{p}^T, 1) Q_\infty^* = \mathbf{0}^T$ ，因此，有 $\mathbf{q}^T = -\mathbf{p}^T K K^T$ ， $w = \mathbf{p}^T \varpi^* \mathbf{p}$ ，于是，

$$Q_\infty^* = \begin{pmatrix} \varpi^* & -\varpi^* \mathbf{p} \\ -\mathbf{p}^T \varpi^* & \mathbf{p}^T \varpi^* \mathbf{p} \end{pmatrix} = \begin{pmatrix} K & 0 \\ -\mathbf{p}^T K & 1 \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} K^T & -K^T \mathbf{p} \\ 0 & 1 \end{pmatrix}$$

这样 $Q_\infty^* = H^{-1} \tilde{H}^{-T}$ 有一个解：

$$\hat{H} = \begin{pmatrix} K^{-1} & \mathbf{0} \\ \mathbf{p}^T & 1 \end{pmatrix} \quad (6.5.5)$$

由定理 6.5.1，可以得到一个度量重构：

$$P^e = (I, \mathbf{0}) \hat{H}^{-1} = K(I, \mathbf{0}), \quad P_j^e = (A_j, \mathbf{a}_j) \hat{H}^{-1} = ((A_j - \mathbf{a}_j \mathbf{p}^T) K, \mathbf{a}_j), \quad j=1,2,\dots,n$$

它是以第一个摄像机坐标系为世界坐标系的度量重构。

由于 $[P_i H^{-1}, \{H X_j\}]$ 为度量重构当且仅当它与上面的重构相差一个相似变换 H_s ，因此 $[P_i H^{-1}, \{H X_j\}]$ 为度量重构的充要条件是

$$H = H_s \begin{pmatrix} K^{-1} & \mathbf{0} \\ \mathbf{p}^T & 1 \end{pmatrix}$$

证毕。

命题 6.5.1 将度量重构问题又带回到求解摄像机内参数和确定无穷远平面问题，而我们希望不通过求解摄像机内参数来确定度量重构，即希望通过直接求解 H 将射影重构提升到度量重构。由定理 6.5.1，当 Q_∞^* 已知时，如果有算法能将 Q_∞^* 分解为 $Q_\infty^* = H^{-1} \tilde{H}^{-T}$ ，就能得到射影变换 H 将射影重构坐标系提升到一个欧氏坐标系，从而得到一个度量重构。

在任意射影坐标系下，绝对二次曲面 Q_{∞}^* 是一个秩为 3 的 4×4 对称矩阵，因此它的奇异值分解必有如下形式： $Q_{\infty}^* = U \text{diag}(\sigma_1, \sigma_2, \sigma_3, 0) U^T$ 。因此，将 Q_{∞}^* 分解为 $Q_{\infty}^* = H^{-1} \tilde{I} H^{-T}$ 是不困难的，用 SVD 分解技术就能实现这种分解。

命题 6.5.2 假定在射影重构坐标系下绝对二次曲面 $Q_{\infty}^* = U \text{diag}(\sigma_1, \sigma_2, \sigma_3, 0) U^T$ ，则方程

$$Q_{\infty}^* = H^{-1} \tilde{I} H^{-T} \quad (6.5.6)$$

的所有解为

$$H = H_s U^T \text{diag} \left(\frac{1}{\sqrt{\sigma_1}}, \frac{1}{\sqrt{\sigma_2}}, \frac{1}{\sqrt{\sigma_3}}, 1 \right) \quad (6.5.7)$$

其中 H_s 是任意三维相似矩阵(即它是相似变换)。

证明 由于

$$\begin{aligned} Q_{\infty}^* &= U \text{diag}(\sigma_1, \sigma_2, \sigma_3, 0) U^T \\ &= \left(U \text{diag}(\sqrt{\sigma_1}, \sqrt{\sigma_2}, \sqrt{\sigma_3}, 1) \right)^T \left(U \text{diag}(\sqrt{\sigma_1}, \sqrt{\sigma_2}, \sqrt{\sigma_3}, 1) \right)^T \end{aligned}$$

并且 $A^{-1} \tilde{I} A^{-T} = \tilde{I}$ 的充要条件是 $A = H_s$ 为相似变换，所以方程：

$$\left(U \text{diag}(\sqrt{\sigma_1}, \sqrt{\sigma_2}, \sqrt{\sigma_3}, 1) \right)^T \left(U \text{diag}(\sqrt{\sigma_1}, \sqrt{\sigma_2}, \sqrt{\sigma_3}, 1) \right)^T = H^{-1} \tilde{I} H^{-T}$$

的所有解为

$$\begin{aligned} H &= \left(\left(U \text{diag}(\sqrt{\sigma_1}, \sqrt{\sigma_2}, \sqrt{\sigma_3}, 1) \right) H_s^{-1} \right)^{-1} \\ &= H_s U^T \text{diag} \left(\frac{1}{\sqrt{\sigma_1}}, \frac{1}{\sqrt{\sigma_2}}, \frac{1}{\sqrt{\sigma_3}}, 1 \right) \end{aligned}$$

证毕。

定理 6.5.1 和命题 6.5.2 表明，从射影重构直接提升到度量重构等价于确定射影重构坐标系下的绝对二次曲面 Q_{∞}^* 。如何在射影重构坐标系下确定绝对二次曲面 Q_{∞}^* 是从射影重构直接得到度量重构的关键问题。如从射影重构到仿射重构类似，没有射影重构之外的知识是不能得到度量重构的。也就是说，从射影重构直接实现度量重构的方法并不能克服从仿射重构实现度量重构所遇到的困难。

6.6 实现分层重构的实例

本节将给出如何从图像的仿射点对应计算度量重构的实例，在某种程度上仿射点对应是度量重

构所需要的最起码的结构信息，如基于相似物体的度量重构、基于镜面反射的度量重构、基于运动物体的度量重构以及基于消失点(线)的度量重构都可以归结为这种情况。从仿射点对应计算度量重构的方法是：首先从点对应计算射影重构；再根据仿射点对应的射影重构建立一个射影变换，由这个射影变换的实特征向量确定无穷远平面在射影重构空间中的坐标，从而得到仿射重构；从仿射重构计算出摄像机的内参数矩阵，最终得到度量重构。

6.6.1 仿射点对应

如果两幅图像(I, I')间的两组点对应

$$\mathfrak{M}_x = \{m_1, m_2, \dots, m_k\} \leftrightarrow \mathfrak{M}'_x = \{m'_1, m'_2, \dots, m'_k\},$$

$$\mathfrak{M}_y = \{n_1, n_2, \dots, n_s\} \leftrightarrow \mathfrak{M}'_y = \{n'_1, n'_2, \dots, n'_s\}$$

满足下述性质：点对应 $\mathfrak{M}_x \leftrightarrow \mathfrak{M}'_x$ 的空间点集 $X = \{X_1, X_2, \dots, X_s\}$ 与 $\mathfrak{M}_y \leftrightarrow \mathfrak{M}'_y$ 的空间点集 $Y = \{Y_1, Y_2, \dots, Y_s\}$ 之间相差一个仿射变换，即存在仿射变换

$$A = \begin{pmatrix} B & b \\ \mathbf{0}^T & 1 \end{pmatrix}$$

使得 $Y_j = AX_j, j = 1, 2, \dots, 5$ ，简记为

$$Y = AX \tag{6.6.1}$$

则称两组点对应 $\mathfrak{M}_x \leftrightarrow \mathfrak{M}'_x, \mathfrak{M}_y \leftrightarrow \mathfrak{M}'_y$ 为仿射点对应。例如两个六面体顶点的两幅图像构成的两组点对应是仿射点对应，因为任何两个长方体之间都存在一个仿射变换将其中的一个顶点集变换到另一个顶点集。

在下面所陈述的一系列问题中均存在仿射点对应，因此它们都可以应用本节所给出的方法来解。

相似物体 已知两个相似物体一组对应特征点的图像，计算景物的度量重构。因为物体 X 与 Y 相似，则必存在一个相似变换 $S = \begin{pmatrix} sU & \mathbf{u} \\ \mathbf{0}^T & 1 \end{pmatrix}$ 使得

$$Y = SX = \begin{pmatrix} sU & \mathbf{u} \\ \mathbf{0}^T & 1 \end{pmatrix} X \tag{6.6.2}$$

由于相似变换是仿射变换，因此两个相似物体 Y 与 X 的任一組对应特征点的图像均构成仿射点对应。

镜面反射 已知一个物体及其镜面反射在图像中的投影，计算景物的度量重构。令 Y 是物体 X 的镜面反射，则必存在一个反射变换 $S = \begin{pmatrix} U & \mathbf{u} \\ \mathbf{0}^T & 1 \end{pmatrix}$ 使得

$$Y = SX = \begin{pmatrix} U & u \\ \mathbf{0}^T & 1 \end{pmatrix} X \quad (6.6.3)$$

由于反射变换是仿射变换，因此物体 X 与它的镜面反射 Y 的任一组对应特征点的图像都构成仿射点对应。

运动物体 已知运动物体在两个摄像机下不同时刻的图像(每个摄像机内参数不变且位置是固定的)，计算景物的度量重构。令运动物体在时刻 1 关于世界坐标系的表示为 X ，并且关于两个摄像机的图像分别为 $\mathfrak{M}_x, \mathfrak{M}'_x$ ；在时刻 2 关于世界坐标系的表示为 Y ，并且关于两个摄像机的图像分别为 $\mathfrak{M}_y, \mathfrak{M}'_y$ 。由于物体在两个时刻之间的运动可以用欧氏运动来表示，即存在欧氏变换 $E = \begin{pmatrix} R & t \\ \mathbf{0}^T & 1 \end{pmatrix}$ 使得

$$Y = EX = \begin{pmatrix} R & t \\ \mathbf{0}^T & 1 \end{pmatrix} X \quad (6.6.4)$$

因此，点对应 $\mathfrak{M}_x \leftrightarrow \mathfrak{M}'_x, \mathfrak{M}_y \leftrightarrow \mathfrak{M}'_y$ 必为仿射点对应。

双眼装置 假设双眼装置作一般刚体运动 E ，从双眼摄像机下的图像计算景物的度量重构。这个问题与上面运动景物的重构问题是等价的。因为双眼装置作一般刚体运动 E 所得到的静止物体的图像与双眼装置在固定的位置上景物作相对的刚体运动所得到的图像是相同的。

三隐消点 已知空间三个方向的隐消点在图像中的投影，计算度量重构问题。记三个无穷远点 $X_1 = (\mathbf{V}_1^T, 0)^T, X_2 = (\mathbf{V}_2^T, 0)^T, X_3 = (\mathbf{V}_3^T, 0)^T$ 在两幅图像中的投影分别为 m_1, m_2, m_3 和 m'_1, m'_2, m'_3 。另外，取两幅图像的 4 个点对应 $\{m_4, m_5, m_6, m_7\} \leftrightarrow \{m'_4, m'_5, m'_6, m'_7\}$ ，它们对应的空间点记为 $\{X_4, X_5, X_6, X_7\}$ 。令

$$\begin{aligned} \mathfrak{M}_x &= \{m_1, m_2, m_3, m_4, m_5\} \leftrightarrow \mathfrak{M}'_x = \{m'_1, m'_2, m'_3, m'_4, m'_5\} \\ \mathfrak{M}_y &= \{m_1, m_2, m_3, m_6, m_7\} \leftrightarrow \mathfrak{M}'_y = \{m'_1, m'_2, m'_3, m'_6, m'_7\} \end{aligned}$$

它们对应的两组空间点 $X = \{X_1, X_2, X_3, X_4, X_5\}$ 与 $Y = \{X_1, X_2, X_3, X_6, X_7\}$ 所确定的 3D 射影变换是一个仿射变换，因为这个射影变换将无穷远平面变换为无穷远平面。因此 $\mathfrak{M}_x \leftrightarrow \mathfrak{M}'_x, \mathfrak{M}_y \leftrightarrow \mathfrak{M}'_y$ 是仿射点对应。

6.6.2 准仿射重构

准仿射重构是一种特殊的射影重构，它是由 Hartley 首先引进并深入研究的。令无穷远平面在一个射影重构空间中的坐标为 π_∞^p ，如果 $\{X_j^p\}$ 的射影重构 $\{X_j^p\}$ 的凸包 $\text{hull}(\{X_j^p\})$ 都位于平面 π_∞^p 的同一侧，则称这个射影重构为准仿射重构(Quasi-Affine Reconstruction)。Hartley 指出，任何一个射影重

构都可以提升到一个准仿射重构。下面简要介绍准仿射重构算法。

给定 n 幅图像的点对应 $\{\mathbf{m}_j^i\}$ ，令

$$P_1 = (I, \mathbf{0}), P_i = (H_i, \mathbf{e}_i), i = 2, 3, \dots, n; \{\mathbf{X}_j^p \mid j = 1, 2, \dots, m\} \quad (6.6.5)$$

是这 n 幅图像的一个射影重构。其中 P_i 是射影重构空间中第 i 幅图像的摄像机矩阵， \mathbf{X}_j^p 是点对应 $\{\mathbf{m}_j^i\}$ 的射影重构。

射影重构为准仿射重构的充要条件是(6.6.5)中的摄像机矩阵和空间点满足下述条件：对所有的 i, j ，式 $P_i \mathbf{X}_j^p = s_j^i \mathbf{m}_j^i$ 中的射影深度 s_j^i 有相同的符号。由此，可以通过改变射影重构的摄像机矩阵的符号以及空间点的符号得到准仿射重构。

6.6.3 仿射重构

下面介绍如何从仿射点对应计算一个仿射重构。令

$$X^{qa} = \{\mathbf{X}_1^{qa}, \mathbf{X}_2^{qa}, \dots, \mathbf{X}_k^{qa}\}, Y^{qa} = \{\mathbf{Y}_1^{qa}, \mathbf{Y}_2^{qa}, \dots, \mathbf{Y}_k^{qa}\} \quad (6.6.6)$$

是仿射点对应 $\mathfrak{M}_x \leftrightarrow \mathfrak{M}'_x, \mathfrak{M}_y \leftrightarrow \mathfrak{M}'_y$ 的一个准仿射重构。如果已知射影重构空间中的无穷远平面 $\pi_{\infty}^{qa} = (\mathbf{a}^T, 1)^T$ ，则可以得到一个仿射重构：

$$P_i^a = P_i^{qa} \begin{pmatrix} I & \mathbf{0} \\ \mathbf{a}^T & 1 \end{pmatrix}^{-1} = (H_i^{qa} - \mathbf{e}_i^{qa} \mathbf{a}^T, \mathbf{e}_i^{qa}), i = 1, 2, \dots, n \quad (6.6.7)$$

$$X^a = \begin{pmatrix} I & \mathbf{0} \\ \mathbf{a}^T & 1 \end{pmatrix} X^{qa} = \{\mathbf{X}_1^a, \mathbf{X}_2^a, \dots, \mathbf{X}_k^a\}, Y^a = \begin{pmatrix} I & \mathbf{0} \\ \mathbf{a}^T & 1 \end{pmatrix} Y^{qa} = \{\mathbf{Y}_1^a, \mathbf{Y}_2^a, \dots, \mathbf{Y}_k^a\} \quad (6.6.8)$$

因此，为了计算仿射重构，只需要在准仿射重构空间中确定无穷远平面 $\pi_{\infty}^{qa} = (\mathbf{a}^T, 1)^T$ 。

考虑 X, Y 的准仿射重构 $X^{qa} = \{\mathbf{X}_1^{qa}, \mathbf{X}_2^{qa}, \dots, \mathbf{X}_k^{qa}\}, Y^{qa} = \{\mathbf{Y}_1^{qa}, \mathbf{Y}_2^{qa}, \dots, \mathbf{Y}_k^{qa}\}$ 。记

$$\hat{X}^{qa} = \{\hat{\mathbf{X}}_1^{qa}, \hat{\mathbf{X}}_2^{qa}, \dots, \hat{\mathbf{X}}_k^{qa}\}, \hat{Y}^{qa} = \{\hat{\mathbf{Y}}_1^{qa}, \hat{\mathbf{Y}}_2^{qa}, \dots, \hat{\mathbf{Y}}_k^{qa}\}$$

其中 $\hat{\mathbf{X}}_j^{qa}, \hat{\mathbf{Y}}_j^{qa}$ 表示 $\mathbf{X}_j^{qa}, \mathbf{Y}_j^{qa}$ 最后一个坐标归一化的向量。由对应关系 $\hat{X}^{qa} \leftrightarrow \hat{Y}^{qa}$ ，当 $k \geq 5$ 时可得到准仿射重构空间中的一个(点)射影变换 H_p ，即射影变换 H_p 使得 $\hat{X}^{qa} = H_p \hat{Y}^{qa}$ ，注意：这是一个齐次变换。如果下述严格等式成立：

$$\hat{Y}_1^{qa} = H_p \hat{X}_1^{qa} \quad (6.6.9)$$

则变换矩阵 H_p 是唯一的，这个变换矩阵记为 \hat{H}_p 。

命题 6.6.1 在射影重构空间中，无穷远平面 $\pi_{\infty}^{qa} = (\mathbf{a}^T, 1)^T$ 是矩阵 \hat{H}_p 的正特征值的特征向量。

证明 由于 $\hat{X}^{qa}, \hat{Y}^{qa}$ 与它们的真实几何结构 X, Y 之间相差一个(点)射影变换, 即存在矩阵 H_{ep} 使得

$$X = H_{ep} X^{qa}, \quad Y = H_{ep} Y^{qa} \quad (6.6.10)$$

由式(6.6.1)和(6.6.10), 可推知

$$Y^{qa} = H_{ep}^{-1} Y = H_{ep}^{-1} A X = H_{ep}^{-1} A H_{ep} X^{qa}$$

结合式(6.6.9), 我们有

$$\hat{H}_p = \mu H_{ep}^{-1} A H_{ep} \quad (6.6.11)$$

所以

$$\hat{H}_p^{-T} H_{ep}^T = \mu H_{ep}^T A^{-T} \quad (6.6.12)$$

从(6.6.12)可以看出 \mathbf{v} 为 A^{-T} 的特征向量当且仅当 $\mathbf{v}_p = H_{ep}^T \mathbf{v}$ 为 \hat{H}_p^{-T} 的特征向量。由 3D 空间中点面的对偶性, A^{-T} 是 3D 欧氏空间中的一个面变换, 因此 A^{-T} 的特征向量 \mathbf{v} 是它的一个不动面。而 H_{ep}^T 表示从欧氏空间到准仿射重构空间的一个面变换, 所以 H_{ep}^T 将欧氏空间中的平面 \mathbf{v} 变换到准仿射重构空间中的平面 $\mathbf{v}_p = H_{ep}^T \mathbf{v}$ 。由于 A 是仿射变换, 所以无穷远平面是 A^{-T} 的一个不动平面, 即 $\pi_\infty = (\mathbf{0}^T, 1)^T$ 是 A^{-T} 的一个特征向量。于是, $\pi_\infty^{qa} = H_{ep}^T \pi_\infty$ 是 \hat{H}_p^{-T} 的一个实特征向量, 它是准仿射重构空间中的无穷远平面。由于 H_p^T 和 H_p^{-T} 有相同的特征向量, 所以 $\pi_\infty^{qa} = H_{ep}^T \pi_\infty$ 也是 \hat{H}_p^T 的一个实特征向量。

最后, 证明 $\pi_\infty^{qa} = (\mathbf{a}^T, 1)^T$ 一定是 \hat{H}_p^T 的正特征值的特征向量。反证: 假定 $\pi_\infty^{qa} = (\mathbf{a}^T, 1)^T$ 是 \hat{H}_p^T 的负特征值的特征向量, 即 $\hat{H}_p^T \pi_\infty^{qa} = \alpha \pi_\infty^{qa}$, $\alpha < 0$ 。于是, 有

$$(\mathbf{a}^T, 1) \hat{H}_p = \alpha (\mathbf{a}^T, 1) \quad (6.6.13)$$

令 $L(s) = s \hat{X}_1^{qa} + (1-s) \hat{Y}_1^{qa}$, $0 \leq s \leq 1$ 是连结两点 $\hat{X}_1^{qa}, \hat{Y}_1^{qa}$ 的直线段。从式(6.6.13), (6.6.9), 有

$$\begin{aligned} (\pi_\infty^{qa})^T (s \hat{X}_1^{qa} + (1-s) \hat{Y}_1^{qa}) &= s (\mathbf{a}^T, 1) \hat{X}_1^{qa} + (1-s) (\mathbf{a}^T, 1) \hat{Y}_1^{qa} \\ &= s \frac{1}{\alpha} (\mathbf{a}^T, 1) \hat{H}_p \hat{X}_1^{qa} + (1-s) (\mathbf{a}^T, 1) \hat{Y}_1^{qa} \\ &= s \frac{1}{\alpha} (\mathbf{a}^T, 1) \hat{Y}_1^{qa} + (1-s) (\mathbf{a}^T, 1) \hat{Y}_1^{qa} \\ &= (s \frac{1}{\alpha} + (1-s)) (\mathbf{a}^T, 1) \hat{Y}_1^{qa} \end{aligned}$$

取 $s = \frac{-\alpha}{1-\alpha}$, 则 $0 < s < 1$ 且 $(\pi_\infty^{qa})^T (s \hat{X}_1^{qa} + (1-s) \hat{Y}_1^{qa}) = 0$, 即 $\hat{X}_1^{qa}, \hat{Y}_1^{qa}$ 位于无穷远平面 $\pi_\infty^{qa} = (\mathbf{a}^T, 1)^T$ 的

两侧, 与准仿射重构矛盾。所以 $\pi_\infty^{qa} = (\mathbf{a}^T, 1)^T$ 一定是 \hat{H}_p^T 的正特征值的特征向量。证毕。

当 A 为一般的仿射变换时, 矩阵 \hat{H}_p^T 可能有 4 个正实值的特征向量 $\mathbf{v}_{pk} = (\mathbf{a}^T, 1)^T, 1 \leq k \leq 4$ 。为了得到仿射重构, 必须从 $\{\mathbf{v}_{pk}\}$ 确定出 π_∞^{qa} 。在讨论一般情况之前, 先讨论 A 为相似变换的特殊情况, 在这种情况下, 除了平面(运动)变换外, 均可以唯一确定在准仿射重构空间中的无穷远平面 π_∞^{qa} 。

命题 6.6.2 假定 A 是一个相似变换 S 。

(1) 如果 S 不是一个平面(运动)变换, 则在准仿射重构空间中可以唯一确定无穷远平面 π_∞^{qa} ;

(2) 当 S 为平面(运动)变换时, 则在准仿射重构空间中无穷远平面 π_∞^{qa} 有两个解。

证明 相似变换 S 可以表示为

$$S = \begin{pmatrix} sU & \mathbf{u} \\ \mathbf{0}^T & 1 \end{pmatrix}$$

其中: s 是相似尺度因子, U 是一个正交矩阵。不难计算

$$S^{-T} = \begin{pmatrix} s^{-1}U & \mathbf{0} \\ -s^{-1}\mathbf{u}^T U & 1 \end{pmatrix}$$

它有形如 $\{1, \varepsilon s^{-1}, s^{-1}e^{-i\theta}, s^{-1}e^{i\theta}\}$ 的特征值, 其中: $\varepsilon = \pm 1$, 当 U 为旋转矩阵时, 取正号; 当 U 为反射矩阵时, 取负号。容易验证无穷远平面 $\pi_\infty = (\mathbf{0}^T, 1)^T$ 是 S^{-T} 的对应于特征值 1 的特征向量。考虑变换

$$\hat{H}_p^{-T} = H_{ep}^T S^{-T} H_{ep}^{-T} \quad (6.6.14)$$

其中: \hat{H}_p, H_{ep} 的意义如前面所述。

(1) 当 $s \neq 1$ 时, \hat{H}_p^{-T} 必有形如 $\{\alpha, \varepsilon\beta, \beta e^{-i\theta}, \beta e^{i\theta}\}$ 的特征值且 $|\alpha| \neq |\beta|$ (一般地, $|\alpha| \neq 1, |\beta| \neq s$, 这是因为式(6.6.14)是齐次等式, 即在相差一个常数倍的意义下相等)。不难看出 \hat{H}_p^{-T} 关于特征值 α 的特征向量 $\mathbf{v}_p = (\mathbf{a}^T, 1)^T$ 是在拟仿射重构空间中无穷远平面 π_∞^{qa} 。由于 \hat{H}_p^{-T} 的特征值是 H_{ep}^{-T} 的特征值共轭的倒数且对应的特征向量相同, 所以 \hat{H}_p^{-T} 有形如 $\{\alpha^{-1}, \varepsilon\beta^{-1}, \beta^{-1}e^{-i\theta}, \beta^{-1}e^{i\theta}\}$ 的特征值且对应于特征值 α^{-1} 的特征向量 $\mathbf{v}_p = (\mathbf{a}^T, 1)^T$ 是无穷远平面在准仿射重构空间中的坐标。注意到 α^{-1} 是其模与其它三个特征值的模都不相等的唯一特征值。因此在这种情况下, 我们可以唯一确定在拟仿射重构空间中的无穷远平面 π_∞^{qa} 。

(2) 当 $s = 1$ 时, S 是一个等距变换。当 U 为旋转矩阵时, S 是一个欧氏变换; 当 U 为反射矩阵时, S 是一个反欧氏变换, 通常称为反射变换(如镜面反射)。下面我们分 S 为欧氏变换和反射变换两种情况来讨论。

当 S 为欧氏变换时, 它必有形如 $\{1, 1, e^{-i\theta}, e^{i\theta}\}$ 的特征值, 因此 S^{-T} 的特征值也是 $\{1, 1, e^{-i\theta}, e^{i\theta}\}$ 。若 S 为非平面(运动)变换, 即 S 的平移向量不与旋转轴正交的平面平行, 则由命题 2.4.4, 可推知无穷远平面 $\pi_\infty = (\mathbf{0}^T, 1)^T$ 是 S^{-T} 对应于二重实特征值 1 的唯一的特征向量, 此时 \hat{H}_p^{-T} 有唯一的实特征向量 $\mathbf{v}_p = (\mathbf{a}^T, 1)^T$, 它是准仿射重构空间中的无穷远平面 π_∞^{qa} ; 若 S 为平面(运动)变换, 即 S 的平移向量

与旋转轴正交的平面平行, 则 S^{-T} 对应于二重实特征值 1 有两个线性无关的特征向量, 从而 \hat{H}_p^T 必有一个正的二重特征值, 且对应这个二重特征值有两个线性无关的特征向量。所以, 此时在线性无关仿射重构空间中, 无穷远平面 π_∞^{qa} 有两个解。

当 S 为反射变换时, 它必有形如 $\{1, -1, e^{-i\theta}, e^{i\theta}\}$ 的特征值, 因此 S^{-T} 的特征值也是 $\{1, -1, e^{-i\theta}, e^{i\theta}\}$ 。因此 \hat{H}_p^T 仅有一个正特征值, 根据命题 6.6.1, 可以唯一确定准仿射重构空间中的无穷远平面 π_∞^{qa} 。证毕。

当 A 为一般仿射变换时, 根据摄像机内参数为恒定和变化的两种情况, 给出选择 π_∞^{qa} 的方法。

恒定内参数 如果所有摄像机的内参数相等即 $K^i = K$, 则第 1 与第 j 个视点间的无穷远平面的单应矩阵 H_∞^{1j} ($j=2,3,\dots,n$) 的 3 个特征值的模相等(通常称为模约束)。由摄像机投影矩阵

$$P_i^a = P_i^{qa} \begin{pmatrix} I & \mathbf{0} \\ \mathbf{a}^T & 1 \end{pmatrix}^{-1} = (H_i^{qa} - \mathbf{e}_i^{qa} \mathbf{a}^T, \mathbf{e}_i^{qa}), i=1,2,\dots,n$$

可知第 1 与第 j 个视点间的无穷远平面的单应矩阵所有可能的解为

$$H_k^{1j} = (H_i^{qa} - \mathbf{e}_i^{qa} \mathbf{a}_k^T) (H_{qa}^1 - \mathbf{e}_{qa}^1 \mathbf{a}_k^T)^{-1}, 1 \leq k \leq 4 \quad (6.6.15)$$

在通常的情况下, 仅存在一个 K 使得对每一个 j 矩阵 H_k^{1j} 满足模约束。这样, 可以通过验证矩阵 H_k^{1j} ($j=2,3,\dots,n$) 的所有特征值的模是否相等来确定 π_∞^{qa} 。

变化内参数 两个摄像机内参数不相等时, 模相等约束不再成立。此时, 计算所有点对应的准仿射重构:

$$\mathbf{X}_j^{qa} = H_\delta^{qa} \mathbf{X}_j^p, \quad j=1,2,\dots,m \quad (6.6.16)$$

如果 $\pi_\infty^{qa} = \mathbf{v}_{pk}$, 则 $f(\mathbf{X}_j^{qa}) = \mathbf{v}_{pk}^T \mathbf{X}_j^{qa}$, $j=1,2,\dots,m$ 必有相同的符号。于是, 可以通过验证 $f(\mathbf{X}_j^{qa})$, $j=1,2,\dots,m$ 是否有相同的符号来确定 π_∞^{qa} 。

注 上面用来确定 π_∞^{qa} 的条件是必要条件, 有时不能唯一的确定 π_∞^{qa} 。

6.6.4 度量重构

假定有 $n+1$ 个视点的图像。从 $n+1$ 个视点的图像, 可以得到一个仿射重构:

$$P_i^a = (H_i^{qa} - \mathbf{e}_i^{qa} \mathbf{a}^T, \mathbf{e}_i^{qa}), \{\mathbf{X}_j^a\}, i=1,2,\dots,n, j=1,2,\dots,m \quad (6.6.17)$$

如果已知第 1 个视点摄像机的内参数矩阵 K , 则可以得到一个度量重构:

$$P_i^e = P_i^e \begin{pmatrix} K & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix}, \mathbf{X}_j^e = \begin{pmatrix} K^{-1} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix} \mathbf{X}_j^a, i = 1, 2, \dots, n, j = 1, 2, \dots, m; \quad (6.6.18)$$

因此为了计算度量重构，只需要计算第 1 个视点摄像机的内参数矩阵 K 。

记第 j 个视点摄像机的绝对二次曲线图像(IAC)为 $\omega^i = (K^i)^{-T} (K^i)^{-1}$ ，由仿射重构可以得到 IAC 的约束方程：

$$\eta_i \omega^i = (H_\infty^{li})^{-T} \omega (H_\infty^{li})^{-1}, i = 2, \dots, n \quad (6.6.19)$$

其中 $H_\infty^{li} = (H_i^{qa} - \mathbf{e}_i^{qa} \mathbf{a}^T) (H_1^{qa} - \mathbf{e}_1^{qa} \mathbf{a}^T)^{-1}$ 是第 1 与第 i 个视点间的无穷远平面的单应矩阵， ω 是第 1 个视点摄像机的 IAC。这样用 5.5 节的方法可以求解摄像机内参数。

6.7 多摄像机系统标定

本节主要利用分层重构的思想给出一种多摄像机系统的标定方法：基于一维标定物的多摄像机系统的标定方法。多摄像机系统的标定通常出现在大场景视觉监控和大场景视觉测量等应用中。基于一维标定物标定的优点主要表现在：一维标定物的构造非常容易，在实际应用中可以利用一根杆上的三个以上的标志点构造一维标定物；所有摄像机能同时观察到整个标定物，这是为了提高标定精度在多摄像机系统标定时应满足条件，基于三维或二维标定物的标定方法因自遮挡而无法满足此条件。

6.7.1 一维标定物

如果线段 AC 包含一些点 B, D, E, F, \dots ，并且任意两点之间的距离已知，则称线段 AC 为一维标定物。研究表明为了标定摄像机，一维标定物至少应包含 3 个彼此距离已知的共线点，包含 4 个或更多个共线点并不能提供任何新的关于摄像机内参数的独立约束，考虑到数据冗余性可抑制噪声的影响，在实际应用中更多的点能提高标定精度。这里只考虑由三个点构成一维标定物这一最小配置，但结论可以扩展到一维标定物包含 4 个或更多个共线点的情况。假定一维标定物由三个点 A, B 和 C 构成，并且 $\|A - C\| = d_1, \|B - C\| = d_2$ 。为陈述方便，一维标定物也称为线段 (ABC) ，由线段 (ABC) 定义的直线记为 L_{ABC} 。

如图 6.7.1 所示，线段 (ABC) 在第 j 个刚体运动下的位置记为 $(A_j B_j C_j)$ ，其在第 i 个摄像机下的图像记为：

$$\{a_{ij}, b_{ij}, c_{ij} \mid j = 1, 2, \dots, n, i = 0, 1, 2, \dots, m\}。$$

我们的目的是根据 $(m+1)$ 个摄像机之间的图像点对应：

$$\{a_{0j} \leftrightarrow a_{1j} \leftrightarrow \dots \leftrightarrow a_{mj}, b_{0j} \leftrightarrow b_{1j} \leftrightarrow \dots \leftrightarrow b_{mj}, c_{0j} \leftrightarrow c_{1j} \leftrightarrow \dots \leftrightarrow c_{mj} \mid j = 1, 2, \dots, n\}$$

计算每个摄像机在第 0 个摄像机坐标系下的摄像机矩阵：

$$P_0^{(e)} = K_0(I \mid 0), P_1^{(e)} = K_1(R_1 \mid t_1), \dots, P_m^{(e)} = K_m(R_m \mid t_m),$$

从而可得到所有摄像机的内参数和这些摄像机之间的相对位姿关系。

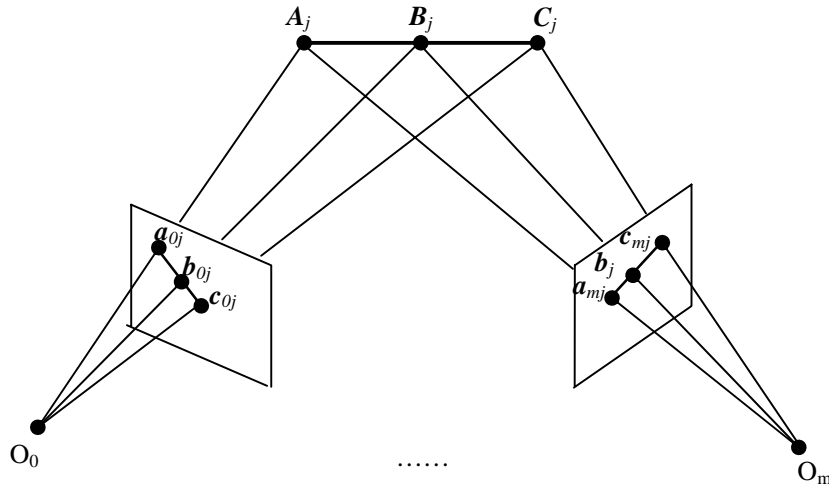


图 6.7.1 用于多摄像机标定的一维标定物

标定原理类似于分层重建：首先计算一维标定物所在直线的无穷远点在每个摄像机下的投影点，可得到仿射意义下摄像机矩阵；然后根据一维标定物的距离信息，计算欧氏摄像机矩阵。

6.7.2 确定仿射摄像机矩阵

根据 6.3 节的讨论，仿射意义下的摄像机矩阵具有如下形式：

$$P_0^{(a)} = (I \mid 0), P_1^{(a)} = (H_{\infty 1} \mid e_1), \dots, P_m^{(a)} = (H_{\infty m} - e_m a^T \mid e_m)$$

其中 $H_{\infty i}$ 为第 0 个摄像机到第 i 个摄像机的无穷远单应， e_i 是第 i 个摄像机关于第 0 个摄像机的极点。

先计算无穷远单应 $H_{\infty i}$ 。为此，首先确定在第 i 个摄像机图像中直线 $L_{A_j B_j C_j}$ 的隐消点 v_{ij} 。由于共

线点 $\{A_j, B_j, C_j\}$ 的单比为 $Simple(A_j, B_j; C_j) = d_1/d_2$ ^{*}，所以共线点 $\{A_j, B_j, C_j, P_{j\infty}\}$ 的交比也为 d_1/d_2 ，即

$$Cross(A_j, B_j; C_j, P_{j\infty}) = d_1/d_2$$

根据射影变换保持交比不变可以得到如下关于 v_{ij} 的线性约束：

$$Cross(a_{ij}, b_{ij}; c_{ij}, v_{ij}) = d_1/d_2. \quad (6.7.1)$$

由此可以计算隐消点 v_{ij} 。

隐消点是无穷远点的图像，所以从第0个和第*i*个摄像机之间的对应隐消点得到 $H_{\infty i}$ 的约束方程：

$$H_{\infty i} v_{0j} = s_{ij} v_{ij}, i=1, 2, \dots, m; j=1, 2, \dots, n \quad (6.7.2)$$

因此，我们可以线性确定 $H_{\infty i}$ 。

下面确定 e_i 。令 $A_j^{(a)} = \begin{pmatrix} \tilde{A}_j^{(a)} \\ 1 \end{pmatrix}, B_j^{(a)} = \begin{pmatrix} \tilde{B}_j^{(a)} \\ 1 \end{pmatrix}, C_j^{(a)} = \begin{pmatrix} \tilde{C}_j^{(a)} \\ 1 \end{pmatrix}$ 为空间点 A_j, B_j, C_j 的仿射重构，则有

$$\begin{aligned} H_{\infty i} \tilde{A}_j^{(a)} + e_i &= P_i^{(a)} A_j^{(a)} = s_{ij}^a a_{ij}, \quad i=0, 1, \dots, m; j=1, 2, \dots, n \\ H_{\infty i} \tilde{B}_j^{(a)} + e_i &= P_i^{(a)} B_j^{(a)} = s_{ij}^b b_{ij}, \quad i=0, 1, \dots, m; j=1, 2, \dots, n \\ H_{\infty i} \tilde{C}_j^{(a)} + e_i &= P_i^{(a)} C_j^{(a)} = s_{ij}^c c_{ij}, \quad i=0, 1, \dots, m; j=1, 2, \dots, n. \end{aligned}$$

其中： $H_{\infty 0} = I, e_0 = 0$ 。消去上述方程中标量因子，可以得到如下关于 $\{\tilde{A}_j^{(a)}, \tilde{B}_j^{(a)}, \tilde{C}_j^{(a)}, e_i\}$ 线性方程组：

$$\begin{aligned} [a_{ij}]_{\times} H_{\infty i} \tilde{A}_j^{(a)} + [a_{ij}]_{\times} e_i &= 0, \quad i=0, 1, \dots, m; j=1, 2, \dots, n \\ [b_{ij}]_{\times} H_{\infty i} \tilde{B}_j^{(a)} + [b_{ij}]_{\times} e_i &= 0, \quad i=0, 1, \dots, m; j=1, 2, \dots, n \\ [c_{ij}]_{\times} H_{\infty i} \tilde{C}_j^{(a)} + [c_{ij}]_{\times} e_i &= 0, \quad i=0, 1, \dots, m; j=1, 2, \dots, n. \end{aligned}$$

通过求解上述线性方程组，我们就得到了仿射意义下的摄像机矩阵

$$P_0^{(a)} = (I | 0), P_1^{(a)} = (H_{\infty 1} | e_1), \dots, P_m^{(a)} = (H_{\infty m} - e_m a^T | e_m)$$

和空间点 $\{A_j, B_j, C_j\}$ 的仿射重构 $\{A_j^{(a)}, B_j^{(a)}, C_j^{(a)}\}$ 。

6.7.3 确定欧氏摄像机矩阵

在得到仿射意义下的摄像机矩阵后，欧氏意义下的摄像机矩阵必有下述形式：

$$P_i^{(e)} = P_i^{(a)} \text{diag}(K_0, 1), i=0, 1, 2, \dots, m,$$

^{*}单比定义为 $Simple(A, B; C) = \frac{|\overline{AC}|}{|\overline{BC}|}$

并且空间点的欧氏重构必为

$$\mathbf{A}_j^{(e)} = \text{diag}(K_0^{-1}, 1)\mathbf{A}_j^{(a)}, \mathbf{B}_j^{(e)} = \text{diag}(K_0^{-1}, 1)\mathbf{B}_j^{(a)}, \mathbf{C}_j^{(e)} = \text{diag}(K_0^{-1}, 1)\mathbf{C}_j^{(a)}, j = 1, 2, \dots, n. \quad (6.7.3)$$

因此，我们只需求解第 0 个摄像机的内参数矩阵 K_0 。为此，令

$$\mathbf{A}_j^{(e)} = \begin{pmatrix} \tilde{\mathbf{A}}_j^{(e)} \\ 1 \end{pmatrix}, \mathbf{B}_j^{(e)} = \begin{pmatrix} \tilde{\mathbf{B}}_j^{(e)} \\ 1 \end{pmatrix}, \mathbf{C}_j^{(e)} = \begin{pmatrix} \tilde{\mathbf{C}}_j^{(e)} \\ 1 \end{pmatrix}$$

则根据式(6.7.3)，得到如下等式：

$$\tilde{\mathbf{A}}_j^{(e)} = K_0^{-1}\tilde{\mathbf{A}}_j^{(a)}, \tilde{\mathbf{B}}_j^{(e)} = K_0^{-1}\tilde{\mathbf{B}}_j^{(a)}, \tilde{\mathbf{C}}_j^{(e)} = K_0^{-1}\tilde{\mathbf{C}}_j^{(a)}, j = 1, 2, \dots, n.$$

由于 $\|\mathbf{A}_j^{(e)} - \mathbf{C}_j^{(e)}\| = d_1, \|\mathbf{B}_j^{(e)} - \mathbf{C}_j^{(e)}\| = d_2$ ，由从可以得到关于 K_0 的线性方程：

$$\begin{cases} (\tilde{\mathbf{C}}_j^{(a)} - \tilde{\mathbf{A}}_j^{(a)})^T \varpi_0 (\tilde{\mathbf{C}}_j^{(a)} - \tilde{\mathbf{A}}_j^{(a)}) = d_1^2 \\ (\tilde{\mathbf{C}}_j^{(a)} - \tilde{\mathbf{B}}_j^{(a)})^T \varpi_0 (\tilde{\mathbf{C}}_j^{(a)} - \tilde{\mathbf{B}}_j^{(a)}) = d_2^2 \end{cases}, j = 1, 2, \dots, n, \quad (6.7.4)$$

其中 $\varpi_0 = K_0^{-T}K_0^{-1}$ 。根据方程组(6.7.4)可以得到 ϖ_0 的线性解，通过对 ϖ_0^{-1} 进行 Cholesky 分解可以得到摄像机内参数矩阵 K_0 。这样，我们就得到了欧氏意义下的摄像机矩阵：

$$P_i^{(e)} = P_i^{(a)} \text{diag}(K_0, 1), i = 0, 1, 2, \dots, m, \quad (6.7.5)$$

和空间点 $\{\mathbf{A}_j, \mathbf{B}_j, \mathbf{C}_j\}$ 的欧氏重重构：

$$\mathbf{A}_j^{(e)} = \text{diag}(K_0^{-1}, 1)\mathbf{A}_j^{(a)}, \mathbf{B}_j^{(e)} = \text{diag}(K_0^{-1}, 1)\mathbf{B}_j^{(a)}, \mathbf{C}_j^{(e)} = \text{diag}(K_0^{-1}, 1)\mathbf{C}_j^{(a)}, j = 1, 2, \dots, n. \quad (6.7.6)$$

根据摄像机矩阵 $P_i^{(e)} = P_i^{(a)} \text{diag}(K_0, 1) = (H_{i0} K_0 | \mathbf{e}_i)$ ，利用 QR 分解可以得到第 i 个摄像机内参数矩阵 K_i ，以及第 i 个摄像机相对第 0 个摄像机的位姿参数 R_i 和 \mathbf{t}_i 。

注：对于每个 j ，方程组(6.7.4)中的两个方程不是相互独立的。由于仿射变换保持共线点的单比不变， $(\tilde{\mathbf{C}}_j^{(a)} - \tilde{\mathbf{B}}_j^{(a)}) = (d_2/d_1)(\tilde{\mathbf{C}}_j^{(a)} - \tilde{\mathbf{A}}_j^{(a)})$ ，因此方程组(6.7.4)的两个方程 $(\tilde{\mathbf{C}}_j^{(a)} - \tilde{\mathbf{A}}_j^{(a)})^T \varpi_0 (\tilde{\mathbf{C}}_j^{(a)} - \tilde{\mathbf{A}}_j^{(a)}) = d_1^2$ 和 $(\tilde{\mathbf{C}}_j^{(a)} - \tilde{\mathbf{B}}_j^{(a)})^T \varpi_0 (\tilde{\mathbf{C}}_j^{(a)} - \tilde{\mathbf{B}}_j^{(a)}) = d_2^2$ 是相互等价的。因此，我们需要一维标定物至少运动 6 次才能标定摄像机系统。

上述线性标定算法的结果是最小化代数距离的解，利用最大似然可以对线性算法的结果进行优化。在计算机视觉研究中，对线性解后处理的常用方法是捆绑调整(bundle adjustment)技术。在这里，捆绑调整是一个以摄像机矩阵和空间点为优化变量以最小化重投影误差为优化目标的非线性优化过程。当图像点测量噪声服从各向同性零均值的高斯分布且独立同分布时，捆绑调整可得到最大似然意义下的欧氏重构。捆绑调整是求解如下非线性最小化问题：

$$\min_{\{P_i, A_j, B_j, C_j\}} \sum_{i=0}^m \sum_{j=1}^n (\|a_{ij} - P_i A_j\|^2 + \|b_{ij} - P_i B_j\|^2 + \|c_{ij} - P_i C_j\|^2) \quad (6.7.7)$$

这是一个具有 $9n+11m+5$ 个变量的非线性问题。由于 $\{A_j, B_j, C_j\}$ 共线，即它们不是相互独立的，因此我们可以进行化简。利用球面坐标 ϕ_j 和 θ_j 表示线段的方向，可以将 B_j, C_j 表示为：

$$B_j = A_j + (d1 - d2) \begin{pmatrix} \sin \phi_j \cos \theta_j \\ \sin \phi_j \sin \theta_j \\ \cos \phi_j \end{pmatrix}, \quad C_j = A_j + d1 \begin{pmatrix} \sin \phi_j \cos \theta_j \\ \sin \phi_j \sin \theta_j \\ \cos \phi_j \end{pmatrix}$$

因此，最小化问题(6.7.7)可以表示为 $5n+11m+5$ 个变量的最小化问题：

$$\min_{\{P_i, A_j, \phi_j, \theta_j\}} \sum_{i=0}^m \sum_{j=1}^n (\|a_{ij} - P_i A_j\|^2 + \|b_{ij} - P_i B_j\|^2 + \|c_{ij} - P_i C_j\|^2) \quad (6.7.8)$$

以前面所得到的欧氏重构结果作为初始值，利用 Levenberg-Marquardt 算法求解非线性最小化问题(6.7.8)，最终得到摄像机投影矩阵和空间点的优化解。

第二篇 矩阵与张量

本篇提要

本篇分为两个部分。第一部分由第 7-10 章构成，介绍矩阵和张量的有关知识，它们是描述、解决计算机视觉问题的强有力数学工具。第二部分由第 11, 12 两章构成，主要介绍矩阵和张量在计算机视觉中的几个典型应用。在第一篇已经看到了矩阵理论对处理计算机视觉问题的作用，在本篇将会进一步看到矩阵和张量的重要作用。

本篇各章主要内容如下：

第 7 章，介绍正交对角化理论。主要内容包括欧氏空间与正交矩阵，正规矩阵正交对角化理论，正规矩阵的正交谱分解及其函数演算。

第 8 章，介绍矩阵分解理论，主要内容包括矩阵的正交三角分解，对称矩阵与反对称矩阵的特征分解，正定矩阵的乔里斯基分解，以及矩阵的 SVD 分解与极分解。

第 9 章，介绍矩阵分析理论，主要内容包括向量和矩阵的范数，矩阵级数与矩阵函数，以及矩阵的微分和矩阵直积。

第 10 章，介绍张量代数。主要内容包括张量积，张量与外代数的一些基本概念与性质。

第 11 章，介绍矩阵奇异值分解在计算机视觉中的应用。主要内容包括欧氏运动与结构，仿射运动与结构，以及射影运动与结构。

第 12 章，介绍多视点张量。用张量来描述多视点几何的多重线性关系，主要内容包括二重线性关系，三重线性关系和四重线性关系。

7. 正交对角化

7.1 内积空间与正交矩阵

7.1.1 内积空间

定义 7.1.1 设 V 是实数域 R 上的线性空间， V 上的内积是满足下述性质的实值函数 $\langle \cdot, \cdot \rangle: V \times V \rightarrow R$:

- (a): 对称性质: $\langle x, y \rangle = \langle y, x \rangle$;
- (b): 线性性质: $\langle sx + uy, z \rangle = s \langle x, z \rangle + u \langle y, z \rangle$;
- (c): 非负性质: $\langle x, y \rangle \geq 0$ 并且 $\langle x, x \rangle = 0 \Leftrightarrow x = 0$.

若在线性空间 V 上定义了一个内积，则称这个空间为**内积空间**。由(a)和(b)，内积关于第二变元也是线性的。在内积空间中有向量的长度、夹角和正交等几何概念：

长度(模或范数): $\|x\|_2 = \sqrt{\langle x, x \rangle}$

距离: $d(x, y) = \|x - y\|_2$

Cauchy 不等式: $\langle x, y \rangle^2 \leq \langle x, x \rangle \cdot \langle y, y \rangle$

夹角: $\cos^2 \theta_{xy} = \frac{\langle x, y \rangle^2}{\langle x, x \rangle \cdot \langle y, y \rangle}$

正交: $x \perp y \Leftrightarrow \langle x, y \rangle = 0$

标准正交基: $\{u_1, \dots, u_n\}$, $\langle u_i, u_j \rangle = \delta_{ij} = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases}$

Fourier 展开式: $\forall x \in V, x = \langle x, u_1 \rangle u_1 + \dots + \langle x, u_n \rangle u_n$

Bessel's 等式: $\forall x \in V, \|x\|_2^2 = \langle x, u_1 \rangle^2 + \dots + \langle x, u_n \rangle^2$

Parseval's 等式: $\forall x, y \in V, \langle x, y \rangle = \langle x, u_1 \rangle \langle y, u_1 \rangle + \dots + \langle x, u_n \rangle \langle y, u_n \rangle$

Schmidt 正交化: $B = \{v_1, v_2, \dots\}$ 是线性无关的向量组，定义：

$$u_1 = v_1, u_k = v_k - \sum_{j=1}^{k-1} \frac{\langle v_k, u_j \rangle}{\langle u_j, u_j \rangle} u_j, k = 2, 3, \dots$$

则 $O = \{u_1, u_2, \dots\}$ 是正交向量组, 并且有

$$\text{span}\{u_1, u_2, \dots, u_k\} = \text{span}\{v_1, v_2, \dots, v_k\}, \quad \forall k$$

欧氏空间 R^n 在 R^n 中的(标准)内积定义为

$$\langle x, y \rangle = y^T x = \sum_{i=1}^n x_i y_i$$

以后, 在没有特别说明的情况下, 所讨论的欧氏空间是指上式定义的内积空间 R^n 。

7.1.2 正交矩阵

定义 7.1.2 如果实方阵 Q 满足 $Q^T Q = Q Q^T = I$ (或等价地 $Q^{-1} = Q^T$), 则称 Q 为正交矩阵。

正交矩阵具有下述性质:

1. Q 为正交矩阵 \Leftrightarrow 任意两个列向量是单位正交的且任意两个行向量也是单位正交的。
2. Q 为正交矩阵 \Leftrightarrow 保持向量的内积不变, 即 $\forall x, y \in R^n, \langle Qx, Qy \rangle = \langle x, y \rangle$ 。
3. Q 为正交矩阵 \Leftrightarrow 保持向量间的距离不变。
4. 任多个正交矩阵的积是正交的。
5. 正交矩阵的行列式等于 ± 1 。

如果正交矩阵的行列式等于 1, 则称为正常正交矩阵(或旋转矩阵), 否则称为非正常正交矩阵(或反射矩阵)。

正交矩阵的几何意义

2 阶正交矩阵 令 $Q = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 是正交阵, 则

$$a^2 + b^2 = 1 = c^2 + d^2, \quad a^2 + c^2 = 1 = b^2 + d^2, \quad ac + bd = 0 = ab + cd$$

从 $a^2 + b^2 = 1$ 知, 存在角 θ , 使得 $a = \cos \theta, b = \pm \sin \theta$ 。如果 $b = \sin \theta$, 令 $\theta' = -\theta$, 则有 $a = \cos \theta', b = -\sin \theta'$ 。总之, 可选取一个角 θ 使得

$$a = \cos \theta, b = -\sin \theta \tag{7.1.1}$$

再利用 $a^2 + c^2 = 1$, 得 $c = \pm \sin \theta$ 。最后, 由 $ac + bd = 0$, 得到 $d = \pm \cos \theta$ ($\sin \theta \neq 0$)。这说明, 当 $\sin \theta \neq 0$ 时,

$$c = \sin \theta, \quad d = \cos \theta \quad \text{或} \quad c = -\sin \theta, \quad d = -\cos \theta$$

所以，当 $\sin \theta \neq 0$ 时， Q 仅有下述两种类型：

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \quad \begin{pmatrix} \cos \theta & -\sin \theta \\ -\sin \theta & -\cos \theta \end{pmatrix} \quad (7.1.2)$$

至于 $\sin \theta = 0$ 时， Q 必为下述四种情况之一：

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

前两种情况，可归为(7.1.2)中的第一个矩阵($\theta = 0, \pi$)；后两种情况可归为(7.1.2)中的第二个矩阵($\theta = 0, \pi$)。综上所述，2阶正交阵 Q 可表示为(7.1.2)中的两种形状。

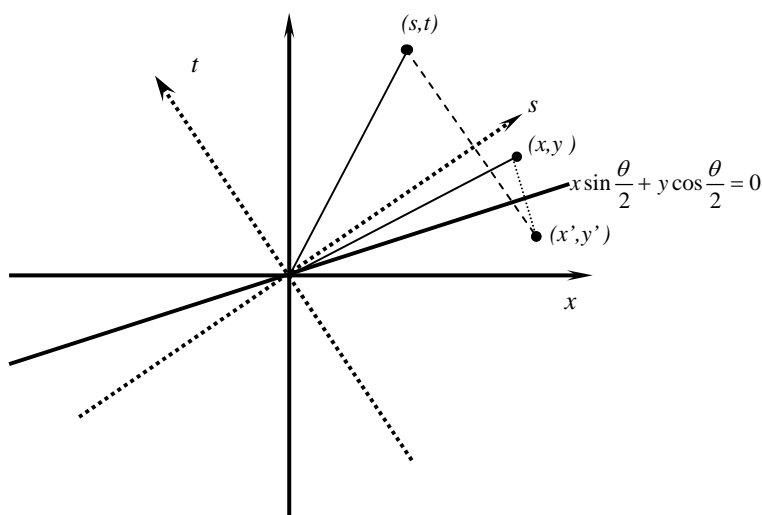


图 7.1.1：2 阶反射矩阵的几何意义

变换 $\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$ 的几何意义：是平面上绕原点旋转 θ 角把点 \mathbf{x} 变到 \mathbf{x}' 的旋转变换。

此时， $\det Q = 1$ ， Q 是正常的正交矩阵。

变换 $\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ -\sin \theta & -\cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$ 的几何意义：这个变换可以看作下述两个变换的复合：

$$\begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \quad \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} s \\ t \end{pmatrix}$$

前者表示旋转, 后者表示关于 s 轴的反射变换, 如图 7.1.1 所示。因此 Q 确定的变换是平面上以直线 $x \sin \frac{\theta}{2} + y \cos \frac{\theta}{2} = 0$ 为对称轴的反射变换。此时 Q 的行列式等于 -1, 它是非正常正交阵。

总之, 2 阶正交矩阵在几何上或是绕坐标原点的旋转, 或是通过原点直线的反射, 根据它是正常或非正常来确定。

3 阶正交矩阵 先分析 3 阶正交矩阵 Q 的特征值与特征向量的性质。

如果 λ 是 Q 的一个实特征值, \mathbf{x} 是相应的特征向量, 即 $Q\mathbf{x} = \lambda\mathbf{x}$, 则 $\mathbf{x}^T \mathbf{x} = \mathbf{x}^T Q^T Q \mathbf{x} = \lambda^2 \mathbf{x}^T \mathbf{x}$ 。于是, 正交矩阵的实特征值必等于 ± 1 。

因 Q 的特征多项式是一个实系数多项式, 若有复特征值, 则其共轭也是特征值。假定 $\lambda_1, \lambda_2, \lambda_3$ 是 Q 的 3 个特征值, 其中 λ_2, λ_3 为互为共轭复特征值。由 $\lambda_1 \lambda_2 \lambda_3 = \det(A) = \pm 1$, 知 $|\lambda_2| = |\lambda_3| = 1$, 因此复特征值的模必等于 1。

令 $\lambda_1 = \pm 1, \lambda_2 = \cos \theta + i \sin \theta, \lambda_3 = \cos \theta - i \sin \theta$ 是 Q 的特征值, \mathbf{x}_1, f, \bar{f} 为相应的单位特征向量。

记 $\mathbf{x}_2 = \frac{f + \bar{f}}{\sqrt{2}}, \mathbf{x}_3 = \frac{f - \bar{f}}{i\sqrt{2}}$, 则有

$$Q\mathbf{x}_1 = \pm \mathbf{x}_1$$

$$\begin{aligned} Q\mathbf{x}_2 &= \frac{1}{\sqrt{2}} Q(f + \bar{f}) = \frac{1}{\sqrt{2}} (\lambda_2 f + \lambda_3 \bar{f}) = \frac{\cos \theta (f + \bar{f})}{\sqrt{2}} + \frac{\sin \theta (f - \bar{f})}{i\sqrt{2}} \\ &= \cos \theta \mathbf{x}_2 + \sin \theta \mathbf{x}_3 \end{aligned}$$

$$Q\mathbf{x}_3 = -\sin \theta \mathbf{x}_2 + \cos \theta \mathbf{x}_3$$

由此, 可推知

$$QX = XR(\theta)$$

其中:

$$X = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3), \quad R(\theta) = \begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}$$

不难验证: X 是一个正交矩阵, 并且可以适当选取 \mathbf{x}_1 的符号使得 X 是一个旋转矩阵。因此, 正交 Q 旋转相似于下述矩阵之一:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}, \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}$$

第二个矩阵可分解为反射矩阵与旋转矩阵的积：

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}$$

综上所述，3 阶正交矩阵所确定的变换或是以一直线为旋转轴的旋转(旋转轴：实特征值的特征向量 \mathbf{x}_1 ，旋转角：复特征值的幅角 θ)，或以一直线为轴的旋转之后再作一个以垂直于此直线的平面的镜面反射(旋转轴：实特征值的特征向量 \mathbf{x}_1 ，旋转角：复特征值的幅角 θ ，反射面：复特征向量实部与虚部张成的平面 $\text{span}\{\mathbf{x}_2, \mathbf{x}_3\}$)。

7.2 酉空间与酉矩阵

7.2.1 酉空间

定义 7.2.1 设 V 是复数域 C 上的线性空间， V 上的内积是满足下述性质的复值函数 $\langle \cdot, \cdot \rangle: V \times V \rightarrow C$ ：

- (a): 共轭对称性或厄米特性: $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$;
- (b): 对第一变元的线性: $\langle s\mathbf{x} + u\mathbf{y}, \mathbf{z} \rangle = s\langle \mathbf{x}, \mathbf{z} \rangle + u\langle \mathbf{y}, \mathbf{z} \rangle$;
- (c): 非负性: $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ ，并且 $\langle \mathbf{x}, \mathbf{x} \rangle = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$ 。

若在复线性空间 V 上定义了一个内积，则称这个线性空间为酉空间(或复内积空间)。由(a)和(b)，内积对于第二变元具有共轭线性：

$$\langle \mathbf{z}, s\mathbf{x} + u\mathbf{y} \rangle = \overline{\langle s\mathbf{x} + u\mathbf{y}, \mathbf{z} \rangle} = \overline{s\langle \mathbf{x}, \mathbf{z} \rangle + u\langle \mathbf{y}, \mathbf{z} \rangle} = \bar{s}\langle \mathbf{z}, \mathbf{x} \rangle + \bar{u}\langle \mathbf{z}, \mathbf{y} \rangle$$

通常，称酉空间的内积是一个半线性的。

欧氏空间中向量的长度、夹角和正交等几何概念与相关性质都可以平行地推广到酉空间：

长度(模或范数): $\|\mathbf{x}\|_2 = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$

距离: $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_2$

Cauchy 不等式: $\langle x, y \rangle \leq \sqrt{\langle x, x \rangle \langle y, y \rangle}$

夹角: $\cos^2 \theta_{xy} = \frac{\langle x, y \rangle^2}{\langle x, x \rangle \langle y, y \rangle}$

正交: $x \perp y \Leftrightarrow \langle x, y \rangle = 0$

标准正交基: $\{u_1, \dots, u_n\}, \langle u_i, u_j \rangle = \delta_{ij} = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases}$

Fourier 展开式: $\forall x \in V, x = \langle x, u_1 \rangle u_1 + \dots + \langle x, u_n \rangle u_n$

Bessel's 等式: $\forall x \in V, \|x\|_2^2 = |\langle x, u_1 \rangle|^2 + \dots + |\langle x, u_n \rangle|^2$

Parseval's 等式: $\forall x, y \in V, \langle x, y \rangle = \langle \langle x, u_1 \rangle u_1 + \dots + \langle x, u_n \rangle u_n, \langle y, u_1 \rangle u_1 + \dots + \langle y, u_n \rangle u_n \rangle$

Schmidt 正交化: $B = \{v_1, v_2, \dots\}$ 是线性无关的向量组, 定义:

$$u_1 = v_1, u_k = v_k - \sum_{j=1}^{k-1} \frac{\langle v_k, u_j \rangle}{\langle u_j, u_j \rangle} u_j, k = 2, 3, \dots$$

则 $O = \{u_1, u_2, \dots\}$ 是正交向量组, 并且有

$$\text{span}\{u_1, u_2, \dots, u_k\} = \text{span}\{v_1, v_2, \dots, v_k\}, \forall k$$

酉空间 C^n 在 C^n 中的标准内积定义为

$$\langle x, y \rangle = \bar{y}^T x = \sum_{i=1}^n \bar{y}_i x_i$$

以后, 在没有特别说明的情况下, 所讨论的酉空间是上式定义的复内积空间 C^n 。

7.2.2 酉矩阵

定义 7.2.2 设 U 为复矩阵, 若满足

$$U^H U = U U^H = I \quad (U^H = \bar{U}^T)$$

则称 U 为酉矩阵(或复正交矩阵)。

欧氏空间 R^n 是酉空间 C^n 的实子空间, 正交矩阵是酉矩阵的特别情形。下述陈述是等价的:

1. U 为酉矩阵;
2. U 的任两列是单位正交的且任两行是单位正交的;
3. 对 C^n 的任意两个向量有 $\langle Ux, Uy \rangle = \langle x, y \rangle$;
4. 对 C^n 的任意向量有 $\|Ux\|_2 = \|x\|_2$ 。

酉矩阵具有下述性质:

1. 酉矩阵的逆是酉矩阵;
2. 酉矩阵的有限积是酉矩阵;
3. 酉矩阵的任一特征值均为单位复数;
4. 不同特征值的特征向量是正交的;
5. 酉矩阵行列式的模等于 1。

定理 7.2.1 设 U 为酉矩阵, 则存在酉矩阵 V 使得 $V^{-1}UV (= V^H UV)$ 为对角矩阵。

证明: 令 $d_1 (|d_1|=1)$ 为 U 的一个特征值, \mathbf{x}_1 为相应的单位特征向量。构造以 \mathbf{x}_1 为第一列的正交阵 $V_1 = (\mathbf{x}_1, \dots)$ (先将 \mathbf{x}_1 扩张为 n 个线性无关的向量组, 再将这个向量组进行 Schmidt 单位正交化可得到 V_1), 于是有 $V_1^{-1}U\mathbf{x}_1 = d_1\mathbf{e}_1$, 因此,

$$V_1^{-1}UV_1 = \begin{pmatrix} d_1 & \mathbf{s} \\ \mathbf{0} & B \end{pmatrix}$$

由 V_1 为酉矩阵知 $\mathbf{s}=\mathbf{0}$, 所以,

$$V_1^{-1}UV_1 = \begin{pmatrix} d_1 & \mathbf{0} \\ \mathbf{0} & B \end{pmatrix}$$

其中 B 必为 $n-1$ 阶酉矩阵。归纳地假定: 存在 $n-1$ 阶正交阵 V_2 , 使得

$$V_2^{-1}BV_2 = \text{diag}(d_2, d_3, \dots, d_n)$$

于是 $V = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & V_2 \end{pmatrix} V_1$ 是酉矩阵, 并且有

$$V^H UV = V^{-1}UV = \text{diag}(d_1, d_2, \dots, d_n)$$

证毕。

推论 7.2.1 (正交矩阵的复标准形) 正交矩阵 Q 必酉相似于一个对角矩阵, 即存在酉矩阵 V 使得

$$V^H QV = V^{-1}QV = \text{diag}(d_1, d_2, \dots, d_n)$$

正交矩阵的实标准形

当正交矩阵有复特征值时, 推论 7.2.1 中的矩阵 V 不可能是一个正交矩阵。现在的问题是能否选择一个适当的正交矩阵 V 使得 $V^{-1}QV$ 有一种简单的形状呢?

正交矩阵的特征多项式是一个实系数多项式, 其复根必共轭成对出现, 并且有

$$Q\mathbf{v} = \lambda\mathbf{v} \Rightarrow Q\bar{\mathbf{v}} = \bar{\lambda}\bar{\mathbf{v}}$$

矩阵 R 称为正交矩阵 Q 的**实标准形**。综上所述, 我们有

定理 7.2.3 对任一正交矩阵 Q 必正交相似于一个形状如 R 的正交矩阵。

7.3 正规矩阵

7.3.1 Schur 引理

定义 7.3.1 设 A, B 是复方阵(或实方阵), 若存在酉(或正交)矩阵 U , 使得

$$U^H A U = U^{-1} A U = B \text{ (或 } U^T A U = U^{-1} A U = B \text{)} \quad (7.3.1)$$

则称 A 酉相似(或正交相似)于 B 。如果 A 酉相似(或正交相似)于对角矩阵, 则称 A 可酉(或正交)角化。

定理 7.3.1(Schur) 任何复方阵均酉相似于上三角矩阵; 如果实方阵的特征值均为实数, 则它正交相似于实上三角矩阵。

证明 下面仅证明定理的前半部分, 用类似方法可证明后半部分。使用数学归纳法: 对于一阶矩阵, 定理显然成立。设定理对 $n-1$ 阶方阵成立, 下面证明对 n 阶矩阵也成立。

令 λ_1 是 n 阶矩阵 A 的一个特征值, \mathbf{a}_1 为相应的单位特征向量, 构造以 \mathbf{a}_1 为第一列的 n 阶酉矩阵 $U_1 = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$, 则

$$A U_1 = (A \mathbf{a}_1, A \mathbf{a}_2, \dots, A \mathbf{a}_n) = (\lambda_1 \mathbf{a}_1, A \mathbf{a}_2, \dots, A \mathbf{a}_n)$$

因为 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ 是 C^n 的正交基, 所以 $A \mathbf{a}_i = \sum_{j=1}^n x_{ij} \mathbf{a}_j$ ($i=1, 2, \dots, n$), 于是,

$$A U_1 = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) \begin{pmatrix} \lambda_1 & x_{21} & \cdots & x_{n1} \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix}$$

其中 A_1 是 $n-1$ 阶矩阵, 由归纳假定存在 $n-1$ 阶酉矩阵 W 和 $n-1$ 阶上三角矩阵 R , 使得

$$W^H A_1 W = R$$

令 $U = U_1 \begin{pmatrix} 1 & \\ & W \end{pmatrix}$, 则 U 是 n 阶酉矩阵, 且

$$U^H A U = \begin{pmatrix} \lambda_1 & x_{21} & \cdots & x_{n1} \\ 0 & & & \\ \vdots & & R & \\ 0 & & & \end{pmatrix}$$

证毕。

定理 7.3.2(Schur 不等式) 设 $A = (a_{ij})_{n \times n}$, 它的特征值为 $\lambda_1, \dots, \lambda_n$, 则

$$\sum_{j=1}^n |\lambda_j|^2 \leq \sum_{i,j=1}^n |a_{ij}|^2 \quad (7.3.2)$$

其中等号成立的充要条件为 A 酉相似于对角矩阵。

证明 由 Schur 引理知, 存在酉矩阵 U 使得

$$U^H A U = R, \text{ 且 } U^H A^H U = R^H$$

其中 $R = (r_{ij})$ 是上三角矩阵。于是,

$$U^H A A^H U = R R^H$$

因此,

$$\sum_{i,j} |a_{ij}|^2 = \text{tr}(A A^H) = \text{tr}(R R^H) = \sum_{i,j} |r_{ij}|^2 \geq \sum_j |r_{jj}|^2 = \sum_j |\lambda_j|^2$$

并且, 等号成立 $\Leftrightarrow r_{ij} = 0 (i \neq j) \Leftrightarrow R$ 是对角矩阵。

7.3.2 正规矩阵

定义 7.3.2 设 A 是复矩阵。若 $A^H = A$, 则称 A 是厄米特矩阵; 若 $A^H = -A$, 则称 A 是反厄米特矩阵; 若 $A^H A = A A^H$, 则称 A 是正规矩阵。

显然, 厄米特、反厄米特、酉矩阵都是正规矩阵, 因而对称矩阵、反对称、正交矩阵也是正规矩阵; 正规的酉相似矩阵也是正规的。

酉对角化

引理 7.3.1 正规的上三角矩阵必为对角矩阵。

证明 令 $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & a_{nn} \end{pmatrix}$ 。由正规性, 知

$$A^H A - AA^H = \begin{pmatrix} \bar{a}_{11}a_{11} - \sum_{j=1}^n \bar{a}_{1j}a_{1j} & & & & * \\ & \sum_{j=1}^2 \bar{a}_{j2}a_{j2} - \sum_{j=2}^n \bar{a}_{2j}a_{2j} & & & \\ & & \ddots & & \\ * & & & \ddots & \\ & & & & \sum_{j=1}^n \bar{a}_{jn}a_{jn} - \bar{a}_{nn}a_{nn} \end{pmatrix} = 0$$

由此可推知：

$$a_{1j} = 0, j = 2, 3, \dots, n; \quad a_{2j} = 0, j = 3, 4, \dots, n; \dots; \quad a_{n-1n} = 0$$

因此， A 为对角矩阵。

定理 7.3.3 矩阵 A 与对角矩阵酉相似的充分必要条件为 A 是正规矩阵。

证明 必要性：假定存在酉矩阵 V 使得

$$V^H AV = V^{-1}AV = \text{diag}(d_1, d_2, \dots, d_n)$$

即 $A = V \text{diag}(d_1, d_2, \dots, d_n) V^H$ ，于是，

$$\begin{aligned} A^H A &= V \text{diag}(\bar{d}_1, \bar{d}_2, \dots, \bar{d}_n) \text{diag}(d_1, d_2, \dots, d_n) V^H \\ &= V \text{diag}(d_1, d_2, \dots, d_n) \text{diag}(\bar{d}_1, \bar{d}_2, \dots, \bar{d}_n) V^H \\ &= AA^H \end{aligned}$$

故 A 是正规矩阵。

充分性：设 A 是正规矩阵。由 Schur 引理，任一矩阵均与上三角矩阵酉相似，因此存在酉矩阵 V 与上三角矩阵 R 使得

$$V^{-1}AV = R$$

显然 R 是正规矩阵，再由引理 7.3.1， R 必为对角矩阵，因此正规矩阵必与对角矩阵相似。证毕。

定理 7.3.3 说明，正规矩阵可表示为

$$A = V \text{diag}(d_1, d_2, \dots, d_n) V^H = d_1 v_1 v_1^H + d_2 v_2 v_2^H + \dots + d_n v_n v_n^H$$

因此，正规矩阵有 n 个两两相互正交的特征向量。

推论 7.3.1 设 A 是正规矩阵，

- (1) A 为厄米特矩阵的充要条件是 A 的特征值均为实数；
- (2) A 为反厄米特矩阵的充要条件是 A 的非零特征值均为纯虚数；
- (3) A 为酉矩阵的充要条件是 A 的特征值均为单位复数。

证明 因 A 为正规矩阵, 所以存在酉矩阵 V 使得

$$A = V \text{diag}(d_1, d_2, \dots, d_n) V^H$$

下面仅证明(1), 用类似的方法可证明(2)和(3)。

必要性: 因为

$$A = V \text{diag}(d_1, d_2, \dots, d_n) V^H, \quad A^H = V \text{diag}(\bar{d}_1, \bar{d}_2, \dots, \bar{d}_n) V^H$$

所以,

$$0 = A - A^H = V \text{diag}(d_1 - \bar{d}_1, d_2 - \bar{d}_2, \dots, d_n - \bar{d}_n) V^H$$

因此,

$$d_1 - \bar{d}_1 = d_2 - \bar{d}_2 = \dots, d_n - \bar{d}_n = 0$$

故 A 的特征值均为实数。

充分性: 由于 A 的特征值均为实数, 所以有

$$A^H = V \text{diag}(\bar{d}_1, \bar{d}_2, \dots, \bar{d}_n) V^H = V \text{diag}(d_1, d_2, \dots, d_n) V^H = A$$

即 A 为厄米特矩阵。证毕。

定理 7.3.4 实正规矩阵 A 与对角矩阵正交相似的充要条件是 A 为实对称矩阵。

证明 充分性: 对于实对称矩阵 A , 因特征值均为实数, 所以根据 Schur 引理, 存在正交矩阵 Q 使得 $Q^T A Q = D$ 是实上三角矩阵, 并且仍是实对称的, 所以 D 是实对角矩阵。

必要性: 令 $A = Q \text{diag}(d_1, d_2, \dots, d_n) Q^T$, 其中 Q 是正交阵, 于是 A 的特征向量是实向量。又因 A 是实矩阵, 所以 $\text{diag}(d_1, d_2, \dots, d_n) = Q^T A Q$ 实对角矩阵。因此

$$A^T = \left(Q \text{diag}(d_1, d_2, \dots, d_n) Q^T \right)^T = Q \text{diag}(d_1, d_2, \dots, d_n) Q^T = A$$

故 A 是实对称矩阵。证毕。

同时酉对角化

设 A, B 是 n 阶正规矩阵。对于 A 存在酉矩阵 Q 使得

$$Q^H A Q = \text{diag}(\lambda_1, \dots, \lambda_n)$$

同样对于 B , 存在酉矩阵 R 使

$$R^H B R = \text{diag}(\mu_1, \dots, \mu_n)$$

一般来说, Q 与 R 是不同的酉矩阵。在什么时候有 $Q=R$? 即何时存在一个共同的酉变换使两个正规矩阵同时化为对角形? 这就是所谓同时对角化问题。显然, A, B 同时对角化的充要条件是它们有相同的特征向量。下面定理给出同时对角化的本质特征。

定理 7.3.5 设 A, B 是 n 阶正规矩阵, 它们同时对角化的充要条件是 $AB=BA$, 即 A 与 B 可交换。

证明 必要性: 若 $Q=R$, 则

$$R^T A R = \text{diag}(\lambda_1, \dots, \lambda_n), \quad R^T B R = \text{diag}(\mu_1, \dots, \mu_n)$$

所以,

$$\begin{aligned} AB &= R \text{diag}(\lambda_1, \dots, \lambda_n) R^H R \text{diag}(\mu_1, \dots, \mu_n) R^H \\ &= R \text{diag}(\lambda_1, \dots, \lambda_n) \text{diag}(\mu_1, \dots, \mu_n) R^H \\ &= R \text{diag}(\mu_1, \dots, \mu_n) \text{diag}(\lambda_1, \dots, \lambda_n) R^H \\ &= R \text{diag}(\mu_1, \dots, \mu_n) R^H R \text{diag}(\lambda_1, \dots, \lambda_n) R^H \\ &= BA \end{aligned}$$

充分性: 设 $AB=BA$ 。若 λ_1, \mathbf{x}_1 使得 $A\mathbf{x}_1 = \lambda_1 \mathbf{x}_1$, 则对任意正整数有

$$AB^k \mathbf{x}_1 = B^k A \mathbf{x}_1 = \lambda_1 B^k \mathbf{x}_1$$

因此, 当 \mathbf{x}_1 为 A 的相应于 λ_1 的特征向量时, $B\mathbf{x}_1, B^2\mathbf{x}_1, \dots, B^k\mathbf{x}_1, \dots$ 均为 A 的相应于 λ_1 的特征向量, 即

$$\{\mathbf{x}_1, B\mathbf{x}_1, B^2\mathbf{x}_1, \dots, B^k\mathbf{x}_1, \dots\} \subset V_{\lambda_1}(A)$$

其中 $V_{\lambda_1}(A)$ 表示 λ_1 的特征子空间。令 B 的特征值为 μ_1, \dots, μ_n , 则必有

$$(B - \mu_1 I)(B - \mu_2 I) \dots (B - \mu_n I) \mathbf{x}_1 = 0$$

不妨假定 $\mathbf{q}_1 = (B - \mu_2 I) \dots (B - \mu_n I) \mathbf{x}_1 \neq 0$ (否则考虑 $(B - \mu_2 I) \dots (B - \mu_n I) \mathbf{x}_1 = 0, \dots$), 则 $\mathbf{q}_1 \in V_{\lambda_1}(A)$ 且

$(B - \mu_1 I) \mathbf{q}_1 = 0$, 即 \mathbf{q}_1 是 A, B 的一个公共的特征向量。于是 $\mathbf{r}_1 = \mathbf{q}_1 / (\mathbf{q}_1^H \mathbf{q}_1)^{\frac{1}{2}}$ 是 A, B 的公共单位特征向量。

再考虑与 \mathbf{r}_1 正交的子空间 $S^\perp(\mathbf{r}_1)$, 一定存在 $\mathbf{x}_2 \in S^\perp(\mathbf{r}_1)$ 是 A 的特征向量, 用上面类似的方法, 可找到 A, B 的第 2 个公共的单位特征向量 $\mathbf{r}_2 \in S^\perp(\mathbf{r}_1)$, 如此继续下去, 一直找到 n 个 A, B 的两两相互正交的公共特征向量为止。这样就证明了所需要的结论。证毕。

7.3.4 正交谱分解

正交投影

定义 7.3.3 令 S 是 C^n 的子空间, S^\perp 是 S 在 C^n 正交补, 即 $S \perp S^\perp$, 且 $C^n = S + S^\perp$ 。 $\forall \mathbf{x} \in C^n$, 存在唯一的 $\mathbf{x}_s \in S, \mathbf{x}_{s^\perp} \in S^\perp$ 使得 $\mathbf{x} = \mathbf{x}_s + \mathbf{x}_{s^\perp}$ 。如果矩阵 P_s 使得

$$P_s \mathbf{x} = P_s (\mathbf{x}_s + \mathbf{x}_{s^\perp}) = \mathbf{x}_s \quad (7.3.3)$$

则称 P_S 是沿方向 S^\perp 到 S 上的正交投影，如图 7.3.1 所示。

令 $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ 和 $\{\mathbf{x}_{k+1}, \mathbf{x}_{k+2}, \dots, \mathbf{x}_n\}$ 分别为 S 和 S^\perp 的一组标准正交基，当然它们的全体构成 C^n 的标准正交基。记 $X_S = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k)$ ， $X_{S^\perp} = (\mathbf{x}_{k+1}, \mathbf{x}_{k+2}, \dots, \mathbf{x}_n)$ ，则正交投影 P_S 使得

$$P_S \mathbf{x}_j = \mathbf{x}_j, j = 1, 2, \dots, k; \quad P_S \mathbf{x}_j = 0, j = k + 1, k + 2, \dots, n$$

所以， $P_S(X_S, X_{S^\perp}) = (X_S, \mathbf{0})$ ，因此 S 上的正交投影有下面的矩阵表示：

$$P_S = (X_S, \mathbf{0})(X_S, X_{S^\perp})^H = X_S X_S^H = \mathbf{x}_1 \mathbf{x}_1^H + \mathbf{x}_2 \mathbf{x}_2^H + \dots + \mathbf{x}_k \mathbf{x}_k^H$$

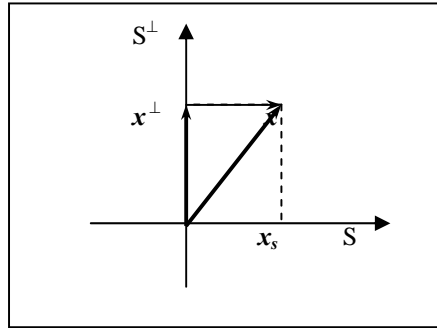


图 7.3.1: 正交投影

若 P 一个正交投影，则存在 C^n 的子空间 S 使得 $P = X_S X_S^H$ ，因此 $P^H = P$ ，即 P 是厄米特矩阵。由

$$P^2 \mathbf{x} = P(P(\mathbf{x}_s + \mathbf{x}_{s^\perp})) = P\mathbf{x}_s = P\mathbf{x}$$

知 P 是一个幂等(idempotent)矩阵^{*)}。反之，若 P 是厄米特的幂等矩阵，令

$$R(P) = \{P\mathbf{x} \mid \mathbf{x} \in C^n\}, \quad N(P) = \{\mathbf{x} \mid P\mathbf{x} = 0, \mathbf{x} \in C^n\}$$

分别为矩阵 P 的值空间与零空间。 $\forall \mathbf{x} \in N(P), \mathbf{y} \in C^n$ ，有

$$\langle P\mathbf{y}, \mathbf{x} \rangle = \mathbf{x}^H P\mathbf{y} = \mathbf{x}^H P^H \mathbf{y} = \mathbf{0}^H \mathbf{y} = 0$$

即 $R(P) \perp N(P)$ 。

令 $S = \{\mathbf{x} \mid \mathbf{x} = P\mathbf{x}, \mathbf{x} \in C^n\}$ ，则 $S \subset R(P)$ 。另一方面 $\forall \mathbf{y} \in C^n, \mathbf{x} = P\mathbf{y} = P^2 \mathbf{y} = P\mathbf{x}$ ，即有 $S \supset R(P)$ ，于是 $S = R(P)$ 。因此， $\forall \mathbf{y} \in C^n$

$$\mathbf{y} = P\mathbf{y} + (\mathbf{y} - P\mathbf{y}) \in R(P) + N(P)$$

^{*)} 若 $P^2 = P$ ，则称 P 是幂等矩阵

于是, 可以看出 P 是 $R(P)$ 上的正交投影。于是, 我们得到正交投影的特征:

定理 7.3.5 矩阵 P 是正交投影的充要条件为它是厄米特的幂等矩阵。

关于正交投影的运算, 有下述定理:

定理 7.3.6(正交投影的运算)

- (1) 设 P_1, P_2 是两个正交投影, 则 $P_1 + P_2$ 是正交投影当且仅当 $R(P_1) \perp R(P_2)$ 。当 $P_1 + P_2$ 是正交投影时, 则它必 $R(P_1) + R(P_2)$ 上的正交投影;
- (2) 设 P_1, P_2 是两个正交投影, 则 $P_1 - P_2$ 是正交投影当且仅当 $P_1 P_2 = P_2 P_1 = P_2$ 。当 $P_1 - P_2$ 是正交投影时, 则它必为 $N(P_1) + R(P_2)$ 上的正交投影;
- (3) 设 P_1, P_2 是两个正交投影且 $P_1 P_2 = P_2 P_1$, 则 $P_1 P_2$ 是 $R(P_1) \cap R(P_2)$ 上的正交投影;
- (4) 若 P_1, \dots, P_k 是正交投影, 则 $P_1 + \dots + P_k$ 是正交投影当且仅当 $R(P_i) \perp R(P_j), \forall i \neq j$;
- (5) 设 $S_1, S_2, S_3, \dots, S_k$ 是 C^n 的两两正交的子空间, 且 $C^n = S_1 + S_2 + S_3 + \dots + S_k$, 则我们有

$$I = P_{s_1} + P_{s_2} + \dots + P_{s_k}$$

正交谱分解

定理 7.3.7 (正交谱分解) 设 A 是正规矩阵, 假定 $\lambda_1, \lambda_2, \dots, \lambda_k$ 是 A 的所有相异的特征, 相应的特征子空间为 $S_1, S_2, S_3, \dots, S_k$, 则有

$$A = \lambda_1 P_{s_1} + \lambda_2 P_{s_2} + \dots + \lambda_k P_{s_k} \quad (7.3.4)$$

并且 P_{s_j} 是 S_j 上的正交投影; $P_{s_1} + P_{s_2} + \dots + P_{s_k} = I$ 。

证明 根据定理 7.3.3, 有

$$A = V \text{diag}(d_1, d_2, \dots, d_n) V^H = d_1 v_1 v_1^H + d_2 v_2 v_2^H + \dots + d_n v_n v_n^H$$

将上式右端按相同特征值进行合并, 则它可改写为:

$$A = \lambda_1 P_1 + \lambda_2 P_2 + \dots + \lambda_k P_k$$

其中: $P_j = \sum_{k=1}^{r_j} v_j v_j^H = P_{s_j}$ 是 A 的特征子空间 S_j 上的正交投影, 并且

$$P_{s_1} + P_{s_2} + \dots + P_{s_k} = v_1 v_1^H + v_2 v_2^H + \dots + v_n v_n^H = I$$

证毕。

函数演算

设 A 是正规矩阵, 由定理 7.3.7,

$$A = \lambda_1 P_1 + \lambda_2 P_2 + \dots + \lambda_k P_k$$

由于 $(P_j)^m = P_j (m \geq 1)$, $P_i P_j = 0 (i \neq j)$, 所以对任意多项式 $p(t)$, 有

$$p(A) = p(\lambda_1)P_1 + p(\lambda_2)P_2 + \dots + p(\lambda_k)P_k \quad (7.3.5)$$

这个结果可以推广到解析函数 $f(t)$ ^{*)}:

$$f(A) = f(\lambda_1)P_1 + f(\lambda_2)P_2 + \dots + f(\lambda_k)P_k \quad (7.3.6)$$

上式通常称为正规矩阵的函数演算。

逆运算: 若 A 是可逆的正规矩阵, 则特征值都不为零, 因此对它的谱分解作函数 $1/t$ 演算, 可以得到 A 的逆矩阵:

$$A^{-1} = (1/\lambda_1)P_1 + (1/\lambda_2)P_2 + \dots + (1/\lambda_k)P_k \quad (7.3.7)$$

共轭转置: 正规矩阵的共轭转置可表示为:

$$A^H = \bar{\lambda}_1 P_1 + \bar{\lambda}_2 P_2 + \dots + \bar{\lambda}_k P_k \quad (7.3.8)$$

厄米特矩阵的指数函数: 令 $A = \lambda_1 P_1 + \lambda_2 P_2 + \dots + \lambda_k P_k$ 是厄米特矩阵的谱分解, 则

$$e^{iA} = e^{i\lambda_1} P_1 + e^{i\lambda_2} P_2 + \dots + e^{i\lambda_k} P_k \quad (7.3.9)$$

定理 7.3.8 U 为酉矩阵的充要条件是存在厄米特矩阵 A 使得 $U = e^{iA}$ 。

证明 由于

$$\begin{aligned} e^{iA}(e^{iA})^H &= (e^{i\lambda_1} P_1 + e^{i\lambda_2} P_2 + \dots + e^{i\lambda_k} P_k)(e^{i\lambda_1} P_1 + e^{i\lambda_2} P_2 + \dots + e^{i\lambda_k} P_k)^H \\ &= P_1 + P_2 + \dots + P_k = I \end{aligned}$$

所以, 厄米特矩阵的指数 e^{iA} 是酉矩阵。反之, 如果 U 是一个酉矩阵, 则必有

$$U = e^{i\lambda_1} P_1 + e^{i\lambda_2} P_2 + \dots + e^{i\lambda_k} P_k$$

令 $A = \lambda_1 P_1 + \lambda_2 P_2 + \dots + \lambda_k P_k$, 则 A 必为厄米特矩阵, 并且 $U = e^{iA}$ 。

7.4 厄米特矩阵

7.4.1 特征值的极性

定义 7.4.1 设 A 为 n 阶厄米特矩阵, 定义

^{*)}若 $f(t)$ 在区间 (a, b) 内有幂级数表示, 则 $f(t)$ 称在区间 (a, b) 内解析。

$$r(\mathbf{x}) = \frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{x}}, \mathbf{x} \in C^n - \{0\} \quad (7.4.1)$$

为矩阵 A 的 Rayleigh 商。

Rayleigh 商有下述性质:

- (1) $r(\mathbf{x})$ 是 $C^n - \{0\}$ 上的连续函数;
- (2) 对任意 $\lambda \neq 0$, $r(\lambda \mathbf{x}) = r(\mathbf{x})$;
- (3) $\forall \mathbf{x} \in \text{span}\{\mathbf{x}_0\}$, $r(\mathbf{x}) = r(\mathbf{x}_0)$;
- (4) 若 d 为 A 的特征值, 对应的特征子空间记为 V_d , 则 $\forall \mathbf{x} \in V_d - \{0\}$, $r(\mathbf{x}) = d$;
- (5) $r(\mathbf{x})$ 在 $C^n - \{0\}$ 上存在最大值和最小值, 并且能够在单位球面 $S = \{\mathbf{x} \in C^n : \|\mathbf{x}\|_2 = 1\}$ 上达到最大值和最小值。

性质(1)至(4)是明显的。下面考虑性质(5): 事实上, 由于 S 是闭集且 $r(\mathbf{x})$ 是 S 上的连续函数, 所以在 S 上存在 $\mathbf{x}_1, \mathbf{x}_2$ 使得

$$\min_{\mathbf{x} \in S} r(\mathbf{x}) = r(\mathbf{x}_1), \max_{\mathbf{x} \in S} r(\mathbf{x}) = r(\mathbf{x}_2)$$

$\forall \mathbf{y} \in C^n - \{0\}$, 令 $\mathbf{x} = \mathbf{y} / \|\mathbf{y}\|_2 \in S$, 根据性质(2), 我们有 $r(\mathbf{x}) = r(\mathbf{y} / \|\mathbf{y}\|_2) = r(\mathbf{y})$ 。所以

$$\min_{\mathbf{x} \in S} r(\mathbf{x}) = r(\mathbf{x}_1) \leq r(\mathbf{y}) \leq r(\mathbf{x}_2) = \max_{\mathbf{x} \in S} r(\mathbf{x})$$

这样, 就证明了性质(5)。

根据性质(5), 我们有

$$\max_{\mathbf{x} \neq 0} \frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{x}} = \max_{\|\mathbf{x}\|=1} \mathbf{x}^H \mathbf{A} \mathbf{x}, \quad \min_{\mathbf{x} \neq 0} \frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{x}} = \min_{\|\mathbf{x}\|=1} \mathbf{x}^H \mathbf{A} \mathbf{x} \quad (7.4.2)$$

下面讨论 Rayleigh 商最大值和最小值与 A 的特征值之间的关系。

定理 7.4.1 设厄米特矩阵 A 的特征值为 $d_1 \leq d_2 \leq \dots \leq d_n$, 则

$$\max_{\mathbf{x} \neq 0} \frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{x}} = d_n, \quad \min_{\mathbf{x} \neq 0} \frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{x}} = d_1 \quad (7.4.3)$$

证明 由定理 7.3.3, 存在酉矩阵 $Q = (\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n)$ 使得

$$A = V \text{diag}(d_1, d_2, \dots, d_n) V^H = d_1 \mathbf{q}_1 \mathbf{q}_1^H + d_2 \mathbf{q}_2 \mathbf{q}_2^H + \dots + d_n \mathbf{q}_n \mathbf{q}_n^H$$

其中 \mathbf{q}_j 是特征值 d_j 的单位特征向量, 并且 $\{\mathbf{q}_j\}$ 构成 C^n 的标准正交基, 所以 $\forall \mathbf{x}, \mathbf{x} = \sum_{j=1}^n a_j \mathbf{q}_j$ 。于是,

$$\frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{x}} = \frac{\sum_{j=1}^n d_j |a_j|^2}{\sum_{j=1}^n |a_j|^2} \leq \frac{d_n \sum_{j=1}^n |a_j|^2}{\sum_{j=1}^n |a_j|^2} = d_n$$

特别取 $\mathbf{x} = \mathbf{q}_1$, 有 $\frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{x}} = d_n$, 因此

$$\max_{\mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{x}} = d_n$$

用同样的方法可证明:

$$\min_{\mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{x}} = d_1$$

证毕。

下面是较为一般的定理:

定理 7.4.2 设 $\mathbf{x} \in \text{span}\{\mathbf{q}_r, \mathbf{q}_{r+1}, \dots, \mathbf{q}_s\}$, 其中 $1 \leq r \leq s \leq n$, 则

$$\max_{\mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{x}} = d_r, \quad \min_{\mathbf{x} \neq \mathbf{0}} \frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{x}} = d_s \quad (7.4.4)$$

更为一般的定理, 通常称为瑞米特矩阵特征值的极小极大原理:

定理 7.4.3(Courant-Fischer) 设瑞米特矩阵 A 的特征值为 $d_1 \leq d_2 \leq \dots \leq d_n$, V_k 为 C^n 的任一 k 维子空间, 则

$$d_k = \min_{V_k} \max \left\{ \frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{x}} : \mathbf{x} \in V_k - \{0\} \right\} \quad (7.4.5)$$

$$d_{n-k+1} = \max_{V_k} \min \left\{ \frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{x}} : \mathbf{x} \in V_k - \{0\} \right\} \quad (7.4.6)$$

证明 下面仅证明式(7.4.5), 式(7.4.6)的证明类似。令

$$W_k = \text{span}\{\mathbf{q}_k, \mathbf{q}_{k+1}, \dots, \mathbf{q}_n\}$$

其中 \mathbf{q}_j 是特征值 d_j 的单位特征向量。显然, $\dim(W_k) = n - k + 1$ 。因 $V_k + W_k \subset C^n$, 所以

$$n \geq \dim(W_k + V_k) = \dim(V_k) + \dim(W_k) - \dim(V_k \cap W_k) = n + 1 - \dim(V_k \cap W_k)$$

于是,

$$\dim(V_k \cap W_k) \geq 1$$

因此，存在非零向量 $\mathbf{x}_0 \in V_k \cap W_k$ ，从而 $\mathbf{x}_0 = \sum_{j=k}^n c_j \mathbf{q}_j$ ，并且

$$\frac{\mathbf{x}_0^H A \mathbf{x}_0}{\mathbf{x}_0^H \mathbf{x}_0} = \frac{\sum_{j=k}^n d_j |c_j|^2}{\sum_{j=k}^n |c_j|^2} \geq d_k$$

所以，对 C^n 的任一 k 维子空间 V_k ，有

$$\max \left\{ \frac{\mathbf{x}^H A \mathbf{x}}{\mathbf{x}^H \mathbf{x}} : \mathbf{x} \in V_k - \{0\} \right\} \geq d_k \tag{7.4.7}$$

另一方面，取 $V_k = \text{span}\{\mathbf{q}_k, \mathbf{q}_{k+1}, \dots, \mathbf{q}_n\}$ ，则根据定理 7.4.2，有

$$\max \left\{ \frac{\mathbf{x}^H A \mathbf{x}}{\mathbf{x}^H \mathbf{x}} : \mathbf{x} \in V_k - \{0\} \right\} = d_k \tag{7.4.8}$$

所以，根据式(7.4.7)和(7.4.8)，得到

$$d_k = \min_{V_k} \max \left\{ \frac{\mathbf{x}^H A \mathbf{x}}{\mathbf{x}^H \mathbf{x}} : \mathbf{x} \in V_k - \{0\} \right\}$$

证毕。

7.4.2 半正定厄米特矩阵

定义 7.4.2 厄米特矩阵 A 称为半正定的当且仅当 $\langle A\mathbf{x}, \mathbf{x} \rangle = \mathbf{x}^H A \mathbf{x} \geq 0, \forall \mathbf{x} \in C^n$ 。半正定的厄米特矩阵记为 $A \geq 0$ 。如果厄米特矩阵 A 使得 $\langle A\mathbf{x}, \mathbf{x} \rangle = \mathbf{x}^H A \mathbf{x} > 0, \forall \mathbf{x} \neq 0$ ，则称它为正定厄米特矩阵，并记为 $A > 0$ 。如果 A 是对称矩阵，可类似定义半正定对称矩阵和正定对称矩阵。

下述性质是明显的：

- (1) $A \geq 0 \Leftrightarrow A$ 的特征值是非负的；
- (2) $A > 0 \Leftrightarrow A$ 的特征值均为正数 $\Leftrightarrow A \geq 0, \det(A) \neq 0$ ；
- (3) $A \geq 0, B \geq 0 \Rightarrow \forall a, b \in R^+, aA + bB \geq 0$ 。

令 A 的谱分解为

$$A = V \text{diag}(d_1, d_2, \dots, d_n) V^H = d_1 \mathbf{v}_1 \mathbf{v}_1^H + d_2 \mathbf{v}_2 \mathbf{v}_2^H + \dots + d_n \mathbf{v}_n \mathbf{v}_n^H$$

对任意 $\alpha \geq 0$ ，定义

$$A^\alpha = d_1^\alpha \mathbf{v}_1 \mathbf{v}_1^H + d_2^\alpha \mathbf{v}_2 \mathbf{v}_2^H + \dots + d_n^\alpha \mathbf{v}_n \mathbf{v}_n^H \tag{7.4.9}$$

特别称 $A^{\frac{1}{2}} = \sqrt{d_1} \mathbf{v}_1 \mathbf{v}_1^H + \sqrt{d_2} \mathbf{v}_2 \mathbf{v}_2^H + \dots + \sqrt{d_n} \mathbf{v}_n \mathbf{v}_n^H$ 为 A 的算术平方根。

如果 $A > 0$ ，则 A^α 对任何实数 α 都有意义，且 $A^\alpha > 0$ 。

对于半正定厄米特矩阵 A 的幂，显然有下述性质：

$$(1) A^\alpha \geq 0;$$

$$(2) A^\alpha A^\beta = A^{\alpha+\beta};$$

$$(3) (A^\alpha)^\beta = A^{\alpha\beta};$$

$$(4) \text{ 如果 } A, B \text{ 是可交换的半正定厄米特矩阵, 则 } A^\alpha B^\alpha = (AB)^\alpha = (BA)^\alpha$$

相对特征值

定义 7.4.3 设 A 为厄米特矩阵， B 为正定的厄米特矩阵，若 (d, \mathbf{x}) 是下述方程的解：

$$A\mathbf{x} = dB\mathbf{x} \quad (7.4.10)$$

则 d 称为 A 相对 B 的特征值， \mathbf{x} 称为 A 相对 B 的特征向量。

当 B 为单位矩阵时，相对特征值与相对特征向量就是通常的特征值与特征向量。因此，相对特征值与相对特征向量通常又称为广义特征值与广义特征向量。

显然， A 相对 B 的特征值 d 是下述方程的解：

$$\det(A - dB) = 0 \quad (7.4.11)$$

通常，方程(7.4.11) 称为相对特征值问题(7.4.10)的特征方程。

定理 7.4.4 设 A 为厄米特矩阵， $B > 0$ ，则存在非奇异矩阵 P 使得

$$P^H A P = \text{diag}(d_1, d_2, \dots, d_n), \quad P^H B P = I \quad (7.4.10)$$

其中 d_j 是 A 相对 B 的特征值， P 的第 j 个列向量 \mathbf{p}_j 是对应的相对特向量。

当 $A \geq 0$ 时， A 相对于 B 的特征值 d_j 均为非负实数。

证明 令 $L = B^{-\frac{1}{2}} > 0$ 。由于矩阵 $L^{-1} A L^{-1}$ 仍是厄米特矩阵，所以存在酉矩阵 Q 使得

$$Q^H L^{-1} A L^{-1} Q = \text{diag}(d_1, d_2, \dots, d_n)$$

取 $P = L^{-1} Q$ ，则

$$P^H A P = \text{diag}(d_1, d_2, \dots, d_n), \quad P^H B P = Q^H Q = I$$

从上式，我们有

$$P^H A P = P^H B P \text{diag}(d_1, d_2, \dots, d_n)$$

所以，

$$A P = B P \text{diag}(d_1, d_2, \dots, d_n)$$

因此, d_1, d_2, \dots, d_n 为相对特征值, P 的列向量为对应的相对特向量。

另外, 由于 $\det(xB - A) = \det(B)\det(xI - B^{-1}A)$, 所以 $B^{-1}A$ 的特征值就是 A 相对于 B 的特征值。注意到矩阵 $B^{-1/2}AB^{-1/2}$ 与 $B^{-1}A$ 相似, 因而有相同的特征值。当 $A \geq 0$ 时, 必有 $B^{-1/2}AB^{-1/2} \geq 0$, 因此 A 相对于 B 的特征值均为非负实数。证毕。

相对特征值的极性

定义 7.4.3 设 A 为 n 阶厄米特矩阵, B 为 n 阶正定厄米特矩阵, 定义

$$R(\mathbf{x}) = \frac{\mathbf{x}^H A \mathbf{x}}{\mathbf{x}^H B \mathbf{x}}, \mathbf{x} \in C^n - \{0\} \quad (7.4.11)$$

为矩阵 A 的广义 Rayleigh 商。

广义 Rayleigh 商与 Rayleigh 商有类似的性质:

- (1) $R(\mathbf{x})$ 是 $C^n - \{0\}$ 上的连续函数;
- (2) 对任意 $\lambda \neq 0$, $R(\lambda \mathbf{x}) = R(\mathbf{x})$;
- (3) $\forall \mathbf{x} \in \text{span}\{\mathbf{x}_0\}, R(\mathbf{x}) = R(\mathbf{x}_0)$;
- (4) 若 d 为 A 的相对特征值, 对应的相对特征子空间记为 V_d , 则 $\forall \mathbf{x} \in V_d - \{0\}, R(\mathbf{x}) = d$;
- (5) $R(\mathbf{x})$ 在 $C^n - \{0\}$ 上存在最大值和最小值, 并且能够在单位球面 $S = \{\mathbf{x} \in C^n : \|\mathbf{x}\|_2 = 1\}$ 上达到它的最大值和最小值。

关于广义 Rayleigh 商最大值和最小值与 A 的相对特征值之间的关系, 可由下列定理来描述。

定理 7.4.5 设厄米特矩阵 A 的相对特征值为 $d_1 \leq d_2 \leq \dots \leq d_n$, 则

$$\max_{\mathbf{x} \neq 0} \frac{\mathbf{x}^H A \mathbf{x}}{\mathbf{x}^H B \mathbf{x}} = d_n, \quad \min_{\mathbf{x} \neq 0} \frac{\mathbf{x}^H A \mathbf{x}}{\mathbf{x}^H B \mathbf{x}} = d_1 \quad (7.4.12)$$

定理 7.4.6 设 $d_r \leq d_{r+1} \leq \dots \leq d_s$ 为厄米特矩阵 A 的相对特征值, 相应的相对特征向量为 $\{\mathbf{p}_r, \mathbf{p}_{r+1}, \dots, \mathbf{p}_s\}$, 其中 $1 \leq r \leq s \leq n$ 。若 $\mathbf{x} \in \text{span}\{\mathbf{p}_r, \mathbf{p}_{r+1}, \dots, \mathbf{p}_s\}$, 则

$$\max_{\mathbf{x} \neq 0} \frac{\mathbf{x}^H A \mathbf{x}}{\mathbf{x}^H B \mathbf{x}} = d_r, \quad \min_{\mathbf{x} \neq 0} \frac{\mathbf{x}^H A \mathbf{x}}{\mathbf{x}^H B \mathbf{x}} = d_s \quad (7.4.13)$$

定理 7.4.7 设厄米特矩阵 A 的相对特征值为 $d_1 \leq d_2 \leq \dots \leq d_n$, V_k 为 C^n 的任一 k 维子空间, 则

$$d_k = \min_{V_k} \max \left\{ \frac{\mathbf{x}^H A \mathbf{x}}{\mathbf{x}^H B \mathbf{x}} : \mathbf{x} \in V_k - \{0\} \right\} \quad (7.4.14)$$

$$d_{n-k+1} = \max_{V_k} \min \left\{ \frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{B} \mathbf{x}} : \mathbf{x} \in V_k - \{0\} \right\} \quad (7.4.15)$$

可用证明厄米特矩阵极性定理的类似方法给出上述定理的证明，也可以通过下述方法将它们作为厄米特矩阵极性定理的推论。首先，不难证明：

$$(1) \quad \forall \mathbf{x} \in \mathbb{C}^n, \quad \frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{B} \mathbf{x}} = \frac{\mathbf{y}^H \mathbf{B}^{-1/2} \mathbf{A} \mathbf{B}^{-1/2} \mathbf{y}}{\mathbf{y}^H \mathbf{y}}, \quad \text{其中: } \mathbf{y} = \mathbf{B}^{1/2} \mathbf{x};$$

(2) \mathbf{p} 为 \mathbf{A} 的对应于相对特征值 d 的相对特征向量的充要条件是 $\mathbf{v} = \mathbf{B}^{1/2} \mathbf{p}$ 为 $\mathbf{B}^{-1/2} \mathbf{A} \mathbf{B}^{-1/2}$ 的对应于特征值 d 的特征向量；

(3) 对于 \mathbf{A} 的任意 r 个相对特征向量 $\{\mathbf{p}_{k_1}, \mathbf{p}_{k_2}, \dots, \mathbf{p}_{k_r}\}$ ，则

$$\text{span}\{\mathbf{p}_{k_1}, \mathbf{p}_{k_2}, \dots, \mathbf{p}_{k_r}\} = \text{span}\{\mathbf{v}_{k_1}, \dots, \mathbf{v}_{k_r}\}$$

其中 $\mathbf{v}_{k_j} = \mathbf{B}^{1/2} \mathbf{p}_{k_j}$ ，并且 $\{\mathbf{v}_{k_1}, \dots, \mathbf{v}_{k_r}\}$ 构成 $\text{span}\{\mathbf{p}_{k_1}, \mathbf{p}_{k_2}, \dots, \mathbf{p}_{k_r}\}$ 的正交基。

应用上述三个性质，由定理 7.4.1，7.4.2 和 7.4.3，可以推知：

$$\begin{aligned} \max_{\mathbf{x} \neq 0} \frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{B} \mathbf{x}} &= \max_{\mathbf{x} \neq 0} \frac{\mathbf{y}^H \mathbf{B}^{-1/2} \mathbf{A} \mathbf{B}^{-1/2} \mathbf{y}}{\mathbf{y}^H \mathbf{y}} = d_1, \\ \min_{\mathbf{x} \neq 0} \frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{B} \mathbf{x}} &= \min_{\mathbf{x} \neq 0} \frac{\mathbf{y}^H \mathbf{B}^{-1/2} \mathbf{A} \mathbf{B}^{-1/2} \mathbf{y}}{\mathbf{y}^H \mathbf{y}} = d_n, \\ \max_{\mathbf{x} \in \text{span}\{\mathbf{p}_r, \mathbf{p}_{r+1}, \dots, \mathbf{p}_s\}} \frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{B} \mathbf{x}} &= \max_{\mathbf{x} \in \text{span}\{\mathbf{v}_r, \mathbf{v}_{r+1}, \dots, \mathbf{v}_s\}} \frac{\mathbf{x}^H \mathbf{B}^{-1/2} \mathbf{A} \mathbf{B}^{-1/2} \mathbf{x}}{\mathbf{x}^H \mathbf{x}} = d_r, \\ \min_{\mathbf{x} \in \text{span}\{\mathbf{p}_r, \mathbf{p}_{r+1}, \dots, \mathbf{p}_s\}} \frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{B} \mathbf{x}} &= \min_{\mathbf{x} \in \text{span}\{\mathbf{v}_r, \mathbf{v}_{r+1}, \dots, \mathbf{v}_s\}} \frac{\mathbf{x}^H \mathbf{B}^{-1/2} \mathbf{A} \mathbf{B}^{-1/2} \mathbf{x}}{\mathbf{x}^H \mathbf{x}} = d_s, \\ \min_{V_k} \max \left\{ \frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{B} \mathbf{x}} : \mathbf{x} \in V_k - \{0\} \right\} &= \min_{V_k} \max \left\{ \frac{\mathbf{x}^H \mathbf{B}^{-1/2} \mathbf{A} \mathbf{B}^{-1/2} \mathbf{x}}{\mathbf{x}^H \mathbf{x}} : \mathbf{x} \in V_k - \{0\} \right\} = d_k, \\ \max_{V_k} \min \left\{ \frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{B} \mathbf{x}} : \mathbf{x} \in V_k - \{0\} \right\} &= \max_{V_k} \min \left\{ \frac{\mathbf{x}^H \mathbf{B}^{-1/2} \mathbf{A} \mathbf{B}^{-1/2} \mathbf{x}}{\mathbf{x}^H \mathbf{x}} : \mathbf{x} \in V_k - \{0\} \right\} = d_{n-k+1}. \end{aligned}$$

作为练习，请读者完成证明细节。

7.5 反对称矩阵

在计算机视觉中特别感兴趣的是 3 阶反对称矩阵。

3 阶反对称矩阵与三维向量 $\mathbf{a} = (a_1, a_2, a_3)^T$ 之间有下述关系：

$$[\mathbf{a}]_{\times} = \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix} \quad (7.5.1)$$

即对任何一个向量 \mathbf{a} ，可以按上述公式确定一个反对称矩阵；反之，任何一个反对称矩阵，都可以用一个向量 \mathbf{a} 写成式(7.5.1)的形式。由于矩阵 $[\mathbf{a}]_{\times}$ 是奇异的，并且 \mathbf{a} 是它的(右、左)零向量，所以 3 阶反对称矩阵由它的零向量确定到相差非零因子的程度。

三维向量 \mathbf{a}, \mathbf{b} 的叉乘被定义为

$$\mathbf{a} \times \mathbf{b} = (a_2 b_3 - a_3 b_2, a_3 b_1 - a_1 b_3, a_1 b_2 - a_2 b_1)^T$$

不难验证叉乘与反对称矩阵的关系是：

$$\mathbf{a} \times \mathbf{b} = [\mathbf{a}]_{\times} \mathbf{b} = (\mathbf{a}^T [\mathbf{b}]_{\times})^T \quad (7.5.2)$$

矩阵相对于叉乘的分配律与矩阵的余因子矩阵有关，确切地说，有下述等式：

$$(\mathbf{A}\mathbf{x}) \times (\mathbf{A}\mathbf{y}) = \mathbf{A}^* (\mathbf{x} \times \mathbf{y}) \quad (7.5.3)$$

其中 \mathbf{A} 是任意的 3×3 矩阵， $\mathbf{A}^* = ((-1)^{i+j} \det M_{ij})$ 是 \mathbf{A} 的余因子矩阵，即 \mathbf{A}^* 的 (i, j) 元素是 \mathbf{A} 的 (i, j) 元素的代数余子式，通常称它为 \mathbf{A} 的对偶。

当 \mathbf{A} 可逆时， $\mathbf{A}^* = \det(\mathbf{A})\mathbf{A}^{-T}$ ；当 \mathbf{A} 不可逆时，此公式不成立。因此， \mathbf{A} 为可逆矩阵时，式(7.5.3)变为

$$(\mathbf{A}\mathbf{x}) \times (\mathbf{A}\mathbf{y}) = \det(\mathbf{A})\mathbf{A}^{-T} (\mathbf{x} \times \mathbf{y}) \quad (7.5.4)$$

在式(7.5.3)中，令 \mathbf{y} 分别为 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ ，我们得到

$$[\mathbf{A}\mathbf{x}]_{\times} \mathbf{A} = \mathbf{A}^* [\mathbf{x}]_{\times} \quad (7.5.5)$$

如果 \mathbf{A} 是可逆矩阵，在(7.5.5)中令 $\mathbf{t} = \mathbf{A}\mathbf{x}$ ，我们得到下述重要公式：

$$[\mathbf{t}]_{\times} \mathbf{A} = \det(\mathbf{A})\mathbf{A}^{-T} [\mathbf{A}^{-1}\mathbf{t}]_{\times} \quad (7.5.6)$$

下面讨论反对称矩阵的对角化问题。由于反对称矩阵是正规矩阵，所以它一定能酉对角化，而定理 7.3.4 表明，反对称矩阵不能正交对角化。那么反对称矩阵的正交相似标准形是什么？由推论

7.3.1, 反对称矩阵的非零特征值均为纯虚数, 反对称矩阵有偶数个非零特征值, 这样奇数阶的反对称矩阵必为奇异矩阵。

定理 7.5.1 若 A 是反对称矩阵, 必存在正交矩阵 Q 使得

$$Q^T A Q = \text{diag}(\lambda_1 Z, \lambda_2 Z, \dots, \lambda_r Z, 0, \dots, 0) \quad (7.5.7)$$

其中 $i\lambda_j, -i\lambda_j, j = 1, 2, \dots, r$ 是 A 的非零特征值, $Z = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ 。

证明 由推论 7.3.1, 反对称矩阵的非零特征值均为纯虚数, 因此反对称矩阵有偶数个非零特征值。不失一般性, 假定 A 的非零特征值为 $i\lambda_j, -i\lambda_j, j = 1, 2, \dots, r$, 对应的单位特征向量为

$$f_j, \bar{f}_j, j = 1, 2, \dots, r$$

并假定 $\{x_{2r+1}, x_{2r+2}, \dots, x_n\}$ 是 A 的零空间的一组标准正交基(如果存在的话), 则

$$V^H A V = \text{diag}(i\lambda_1, -i\lambda_1, i\lambda_2, -i\lambda_2, \dots, i\lambda_r, -i\lambda_r, 0, \dots, 0) \quad (7.5.8)$$

其中: $V = (f_1, \bar{f}_1, f_2, \bar{f}_2, \dots, f_r, \bar{f}_r, x_{2r+1}, \dots, x_n)$ 是酉矩阵。令

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1/i \\ 1 & -1/i \end{pmatrix}$$

则它是一个 2 阶酉矩阵, 并且 $W = \text{diag}(\underbrace{U, U, \dots, U}_r, I_{n-2r})$ 是 n 阶酉矩阵。于是

$$\begin{aligned} VW &= (f_1, \bar{f}_1, f_2, \bar{f}_2, \dots, f_r, \bar{f}_r, x_{2r+1}, \dots, x_n) \text{diag}(\underbrace{U, U, \dots, U}_r, I_{n-2r}) \\ &= \left(\frac{f_1 + \bar{f}_1}{\sqrt{2}}, \frac{f_1 - \bar{f}_1}{i\sqrt{2}}, \dots, \frac{f_r + \bar{f}_r}{\sqrt{2}}, \frac{f_r - \bar{f}_r}{i\sqrt{2}}, x_{2r+1}, \dots, x_n \right) \end{aligned}$$

是正交矩阵, 并且

$$W^H \text{diag}(i\lambda_1, -i\lambda_1, i\lambda_2, -i\lambda_2, \dots, i\lambda_r, -i\lambda_r, 0, \dots, 0) W = \text{diag}(\lambda_1 Z, \lambda_2 Z, \dots, \lambda_r Z, 0, \dots, 0)$$

其中 $Z = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ 。令 $Q = VW$, 则由式(7.5.8), 得到

$$\begin{aligned} Q^T A Q &= W^H V^H A V W = W^H \text{diag}(i\lambda_1, -i\lambda_1, i\lambda_2, -i\lambda_2, \dots, i\lambda_r, -i\lambda_r, 0, \dots, 0) W \\ &= \text{diag}(\lambda_1 Z, \lambda_2 Z, \dots, \lambda_r Z, 0, \dots, 0) \end{aligned}$$

证毕。

8. 矩阵分解

8.1 正交三角分解^{*)}

定义 8.1.1 如果非奇异实矩阵 A 能够表示为正交矩阵 Q 与上三角矩阵 R 的积，即

$$A=QR \quad (8.1.1)$$

则称式(8.1.1)是 A 的 QR 分解。

与 QR 分解类似，还有 QL, LQ, RQ 分解，其中 L 表示下三角矩阵。矩阵的 QR 分解，RQ 分解，QL 分解，LQ 分解统称为矩阵的正交三角分解。下面仅讨论实矩阵的 QR 分解，其它类型分解可类似的获得。

定理 8.1.1 对任意非奇异实矩阵 A 总可以分解为正交矩阵 Q 与上三角矩阵 R 的积，如果要求上三角阵 R 的对角元素均为正数，则分解是唯一的。

证明 令 $A=(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ ，由于 A 是非奇异矩阵，因此它的列向量线性无关，所以对列向量组进行 Schmidt 正交化可得到 n 个两两相互正交的单位向量 $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n$ 使得

$$\begin{cases} \mathbf{q}_1 = c_{11}\mathbf{a}_1 \\ \mathbf{q}_2 = c_{12}\mathbf{a}_1 + c_{22}\mathbf{a}_2 \\ \dots\dots\dots \\ \mathbf{q}_n = c_{1n}\mathbf{a}_1 + c_{2n}\mathbf{a}_2 + \dots + c_{nn}\mathbf{a}_n \end{cases}$$

其中 $c_{jj} \neq 0$ 。将上式写成矩阵形式，有

$$\underbrace{(\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n)}_Q = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) \begin{pmatrix} c_{11} & \dots & c_{1n} \\ & \ddots & \vdots \\ & & c_{nn} \end{pmatrix}$$

即 $Q=AC$ 。由于 $C^{-1}=R$ 也是上三角矩阵，所以有 $A=QR$ 。

唯一性：设 A 有两种分解 $QR=A=\tilde{Q}\tilde{R}$ ，则必有 $\tilde{R}\tilde{R}^{-1}=Q^{-1}\tilde{Q}$ ，于是 $\tilde{R}\tilde{R}^{-1}$ 是正交的上三角矩阵。

^{*)} 在本章中，除特别说明外矩阵均指实矩阵。

下面仅给出 3 阶矩阵 RQ 分解的 Givens 旋转方法, 因为读者毫无困难地将这种方法推广到 n 阶矩阵的 QR、RQ、LQ、QL 分解。令

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$$

(1) 从等式

$$AG_x = (\mathbf{a}_1, c\mathbf{a}_2 + s\mathbf{a}_3, -s\mathbf{a}_2 + c\mathbf{a}_3)$$

我们可以选择一个角度 θ 使得 AG_x 对角线以下的任何一个指定的元素变为零。例如, 为了使 AG_x 的 (3, 2) 元素变为零, 即

$$ca_{32} + sa_{33} = 0$$

可选取

$$c = -\frac{a_{33}}{\sqrt{a_{32}^2 + a_{33}^2}}, s = \frac{a_{32}}{\sqrt{a_{32}^2 + a_{33}^2}}$$

显然, 有 $c^2 + s^2 = 1$ 。因上式定义的 c, s 确实使 G_x 是一个旋转矩阵。即, 找到了一个 Givens 旋转矩阵 G_x 使 AG_x 的 (3, 2) 元素变为零, 且 AG_x 的第一列与 A 的第一列相等。

(2) 类似地, 可以找到一个 Givens 旋转矩阵 G_y 使 $AG_x G_y$ 的 (3, 1) 元素变为零, 并且 $AG_x G_y$ 的第二列与 AG_x 的第二列相同;

(3) 同样可以找到一个 Givens 旋转矩阵 G_z 使 $AG_x G_y G_z$ 的 (2, 1) 元素变为零, $AG_x G_y G_z$ 的第三列与 $AG_x G_y$ 的第三列相同并且它们的第三行也相同。于是, 我们有

$$AG_z G_y G_x = R$$

它是上三角矩阵。于是, A 是上三角矩阵 R 与旋转矩阵 $Q = G_x^T G_y^T G_z^T$ 的乘积。这样, 我们就完成了矩阵 A 的 RQ 分解。

8.1.2 Householder 方法

Givens 旋转方法, 对于 n 阶矩阵需要作 $n(n-1)/2$ 个 Givens 旋转矩阵的积, 计算量较大。对于高阶矩阵而言, 用下述 Householder (反射) 矩阵变换进行 QR 分解更有效。

定义 8.1.2 $\forall \mathbf{v} \in R^n \setminus \{0\}$, n 阶矩阵

$$H_v = I - 2 \frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T \mathbf{v}}$$

称为 Householder 矩阵。

不难验证 $H_v^T H_v = I$ 且 $\det(H_v) = -1$, 因此 H_v 是正交矩阵, 更确切地说它是反射矩阵。

引理 8.1.1 对任意给定的非零列向量 $\mathbf{x} \in \mathbb{R}^n$ 及单位向量 $\mathbf{z} \in \mathbb{R}^n$, 存在 Householder 矩阵 H 使得 $H\mathbf{x} = \|\mathbf{x}\| \mathbf{z}$ 。

证明 当 $\mathbf{x} = \|\mathbf{x}\| \mathbf{z}$ 时, 取单位向量 \mathbf{v} 满足 $\mathbf{v}^T \mathbf{x} = 0$, 则必有

$$H_v \mathbf{x} = (I - 2\mathbf{v}\mathbf{v}^T)\mathbf{x} = \mathbf{x} - 2\mathbf{v}(\mathbf{v}^T \mathbf{x}) = \mathbf{x} = \|\mathbf{x}\| \mathbf{z}$$

当 $\mathbf{x} \neq \|\mathbf{x}\| \mathbf{z}$ 时, 取 $\mathbf{v} = \frac{\mathbf{x} - \|\mathbf{x}\| \mathbf{z}}{\|\mathbf{x} - \|\mathbf{x}\| \mathbf{z}\|}$, 则

$$H_v \mathbf{x} = (I - 2\mathbf{v}\mathbf{v}^T)\mathbf{x} = \mathbf{x} - 2 \frac{(\mathbf{x} - \|\mathbf{x}\| \mathbf{z})(\mathbf{x} - \|\mathbf{x}\| \mathbf{z})^T}{\|\mathbf{x} - \|\mathbf{x}\| \mathbf{z}\|^2} \mathbf{x}$$

因 $\|\mathbf{x} - \|\mathbf{x}\| \mathbf{z}\|^2 = 2 \langle \mathbf{x} - \|\mathbf{x}\| \mathbf{z}, \mathbf{x} \rangle$, 所以

$$H_v \mathbf{x} = \mathbf{x} - (\mathbf{x} - \|\mathbf{x}\| \mathbf{z}) = \|\mathbf{x}\| \mathbf{z}$$

证毕。

根据引理 8.1.1, 对 n 阶非奇异方阵 $A = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ 作一系列的 Householder 变换, 可化将它化为上三角矩阵。具体步骤如下:

(1) 由引理 8.1.1 可以找到 n 阶 Householder 矩阵 H_1 使得

$$H_1 \mathbf{a}_1 = \|\mathbf{a}_1\| \mathbf{e}_1 \quad (\mathbf{e}_1 = (1, 0, \dots, 0)^T \in \mathbb{R}^n)$$

令 $a_{11}^{(1)} = \|\mathbf{a}_1\|$, 则有

$$H_1 A = \begin{pmatrix} a_{11}^{(1)} & \mathbf{a}^{(1)T} \\ \mathbf{0} & A_1 \end{pmatrix}$$

(2) 再对 $n-1$ 阶矩阵 A_1 施行步骤 1), 有

$$H_2 A_1 = \begin{pmatrix} a_{22}^{(2)} & \mathbf{a}^{(2)T} \\ \mathbf{0} & A_2 \end{pmatrix}$$

.....

($n-1$) 将上述过程进行到 $n-1$ 步, 有

$$H_{n-1} A^{(n-2)} = \begin{pmatrix} a_{n-1,n-1}^{(n-1)} & \mathbf{a}^{(n-1)} \\ \mathbf{0} & a_{n,n}^{(n-1)} \end{pmatrix}$$

$$n) \text{ 令 } Q = \begin{pmatrix} I_{n-2} & \mathbf{0} \\ \mathbf{0} & H_{n-1} \end{pmatrix} \cdots \begin{pmatrix} I_2 & \mathbf{0} \\ \mathbf{0} & H_3 \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & H_2 \end{pmatrix} H_1, \text{ 则}$$

$$QA = \begin{pmatrix} a_{11}^{(1)} & & * & & \\ & a_{22}^{(2)} & & & \\ & & \ddots & & \\ & & & & a_{nn}^{(n-1)} \end{pmatrix} R$$

是上三角矩阵。这样，就将矩阵 A 分解为正交矩阵与上三角矩阵的乘积。

上述分解方法通常称为 Householder 方法，它只需要作 $(n-1)$ 个 Householder 矩阵的积，其计算量大约是 Givens 旋转方法的一半。

8.1.3 内参数与外参数的分解

第 3 章已给出估计出（欧氏）摄像机矩阵 P 的线性方法。在有些应用场合，如运动分析，需要通过矩阵 P 求解摄像机的内外参数，再确定视觉系统关于世界坐标系的位置关系。RQ 分解是实现摄像机内参数、外参数求解的最为简便的方法。

令摄像机内参数矩阵为 K ，外参数矩阵为 (R, \mathbf{t}) ，由于此时的摄像机矩阵是欧氏的，所以可写成下述形式：

$$P = \alpha K(R, \mathbf{t})$$

将 P 表示成 $P = (H, \mathbf{p}_4)$ ，则有

$$H = \alpha KR, \quad \mathbf{p}_4 = \alpha \mathbf{t}$$

对 H 作 RQ 分解： $H = \hat{K}\hat{R}$ ，其中 \hat{K} 是对角元均为正数的上三角矩阵， \hat{R} 为正交矩阵，并且种分解是唯一的。三角矩阵 \hat{K} 与摄像机参数矩阵相差一个正常数倍，由于内参数矩阵最后一个元素为 1，所以内参数矩阵必为： $K = \hat{k}_{33}^{-1}\hat{K}$ ，其中 \hat{k}_{33} 是矩阵 \hat{K} 的第(3, 3)元素。如果正交矩阵 \hat{R} 是一个旋转矩阵，则它就是摄像机关于世界坐标系的姿态。如果 \hat{R} 不是旋转矩阵，则表明所估计的摄像机矩阵与实际的(欧氏)摄像机矩阵反号，即 P 中的齐次因子的符号 $\text{sgn } \alpha = -1$ ，于是 $R = -\hat{R}$ 就是所要求解旋转矩阵。对于平移参数 \mathbf{t} ，可通过下式来确定：

$$\begin{cases} \mathbf{t} = \hat{k}_{33}^{-1} \mathbf{p}_4, & \text{若 } \hat{R} \text{ 是旋转矩阵;} \\ \mathbf{t} = -\hat{k}_{33}^{-1} \mathbf{p}_4, & \text{若 } \hat{R} \text{ 不是旋转矩阵。} \end{cases}$$

这样，就得到了摄像机内、外参数。

8.2 乔里斯基 (Cholesky) 分解

根据对称矩阵的特征值分解, 我们有

$$A = Q \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) Q^T \quad (8.2.1)$$

所以, $\forall j$,

$$\mathbf{q}_j^T A \mathbf{q}_j = \mathbf{e}_j^T \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) \mathbf{e}_j = \lambda_j \quad (8.2.2)$$

因此, 正定对称矩阵的特征值 λ_j 均为正数。

定理 8.2.1(Cholesky 分解) 正定对称矩阵 A 可以唯一地分解为对角元素均大于零的上三角矩阵 K 与其转置 K^T 的乘积, 即 $A = KK^T$, 其中 K 为对角元素均大于零的上三角矩阵。

证明 根据式(8.2.1),

$$A = \underbrace{[Q \text{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_n})]}_V \underbrace{[Q \text{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_n})]^T}_{V^T} = VV^T$$

对 V 作 RQ 分解, 有 $V = \tilde{K}Q$, 其中 \tilde{K} 是上三角矩阵, Q 是正交矩阵, 于是,

$$A = VV^T = \tilde{K}Q Q^T \tilde{K}^T = \tilde{K}\tilde{K}^T$$

令 $D = \text{diag}(\text{sign}k_{11}, \text{sign}k_{22}, \dots, \text{sign}k_{nn})$, 则 $K = \tilde{K}D$ 是对角元素均为正数的上三角矩阵。因此,

$$A = \tilde{K}\tilde{K}^T = (KD^{-1})(KD^{-1})^T = KK^T$$

下面证明唯一性: 若存在 K_1, K_2 使得 $K_1 K_1^T = A = K_2 K_2^T$, 则必有

$$K_2^{-1} K_1 = K_2^T K_1^{-T} = (K_2^{-1} K_1)^{-T}$$

因上式两边分别为上三角和下三角矩阵, 因此它们是同一个对角矩阵, 从而

$$K_2^{-1} K_1 = (K_2^{-1} K_1)^{-T} = (K_2^{-1} K_1)^{-1}$$

因此, 必有 $K_2^{-1} K_1 = I$, 故 $K_1 = K_2$ 。证毕。

递推算法

令 $K = (k_{ij})$, 则由 $A = KK^T$ 得到下述等式:

$$\begin{cases} a_{ji} = k_{jj}k_{ij} + k_{jj+1}k_{ij+1} + \cdots + k_{jn}k_{in} = k_{jj}k_{ij} + \sum_{r=j+1}^n k_{jr}k_{ir} & (j > i) \\ a_{jj} = k_{jj}^2 + k_{jj+1}^2 + \cdots + k_{jn}^2 = k_{jj}^2 + \sum_{r=j+1}^n k_{jr}^2 & (j = i) \end{cases}$$

从上式, 可得到 k_{ij} 的下述递推关系:

$$k_{ij} = \begin{cases} \sqrt{a_{jj} - \sum_{r=j+1}^n k_{jr}^2}, & (j = i) \\ \frac{1}{k_{jj}}(a_{ji} - \sum_{r=j+1}^n k_{jr}k_{ir}), & (j > i) \\ 0, & (j < i) \end{cases}$$

特别地, 对于 3 阶矩阵 $A = (a_{ij})$, 有

$$\begin{aligned} k_{33} &= \sqrt{a_{33}}, \quad k_{23} = \frac{1}{k_{33}}a_{32}, \quad k_{13} = \frac{1}{k_{33}}a_{31}, \\ k_{11} &= \sqrt{a_{11} - \sum_{r=2}^3 k_{1r}^2}, \quad k_{22} = \sqrt{a_{22} - k_{23}^2}, \quad k_{12} = \frac{1}{k_{22}}(a_{21} - k_{23}k_{13}). \end{aligned}$$

Cholesky 分解在计算机视觉中的一个典型应用, 是利用它可以非常方便地从绝对二次曲线 ω 或绝对二次曲面的图像 ω^* 分解出摄像机内参数。

8.3 奇异值分解

矩阵的奇异值分解在最优化问题, 特征值问题, 最小二乘问题及矩阵的广义逆问题中都具有非常重要的应用。在计算机视觉中, 常常使用它来求解运动分析和线性最优化问题。

8.3.1 正交对角分解

在 7.3 节, 已经看到对称矩阵的正交对角分解, 即对任意的对称矩阵 A , 总存在正交矩阵 Q 使

$$A = Q \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) Q^T$$

对于一般的非对称矩阵, 不再有上述分解, 但是存在两个正交矩阵 P 和 Q 使得上式成立。确切

地说, 有下述结论。

定理 8.3.1 (正交对角分解) 设 A 是一个非奇异方阵, 则存在正交矩阵 P 和 Q 使得下式成立:

$$A = P \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n) Q^T \quad (8.3.1)$$

其中, $\sigma_i > 0 (i = 1, 2, \dots, n)$ 。

事实上, 因为矩阵 A 是非奇异的, 所以 $A^T A$ 为正定实对称矩阵, 因此存在正交矩阵 Q 使

$$Q^T (A^T A) Q = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

其中, $\lambda_i > 0 (i = 1, 2, \dots, n)$ 为 $A^T A$ 的特征值。令

$$\sigma_i = \sqrt{\lambda_i}, \quad D = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n)$$

则有

$$Q^T (A^T A) Q = D^2$$

或写成

$$\underbrace{(A Q D^{-1})^T}_{P} A Q = D$$

因 $P^T P = (A Q D^{-1})^T (A Q D^{-1}) = D^{-T} Q^T A^T A Q D^{-1} = I$, 即 P 是正交矩阵, 所以 $P^T A Q = D$, 故有

$$A = P \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n) Q^T$$

8.3.2 奇异值分解

奇异值分解是将定理 8.3.1 推广到任意实矩阵, 即不要求矩阵 A 是可逆的方阵, 也不要求它是方阵。

设 $A \in R_r^{m \times n}$ *, $A^T A$ 的特征值为

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > \lambda_{r+1} = \dots = \lambda_n = 0$$

则称 $\sigma_i = \sqrt{\lambda_i} (i = 1, 2, \dots, n)$ 为矩阵 A 的奇异值。

当 A 为零矩阵时, 它的所有奇异值均为零。一般地, 矩阵 A 的奇异值的个数等于 A 的列数, A 的非零奇异值的个数等于 A 的秩。

定理 8.3.2 (奇异值分解, SVD) 设 $A \in R_r^{m \times n}$, 则存在 m 阶正交矩阵 U 和 n 阶正交矩阵 V 使**)

*) $R_r^{m \times n}$ 表示秩为 r 的所 $m \times n$ 阶实矩阵的集合。

**) 对复矩阵 $A \in C_r^{m \times n}$, 定理也成立, 即存在 m 阶酉矩阵 U 和 n 阶酉矩阵 V 使 $A = U D V^H$ 。证明也类似。

$$A = U \underbrace{\begin{pmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}}_D V^T \quad (8.3.2)$$

其中 $\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$, 即 $\sigma_1, \sigma_2, \dots, \sigma_r$ 是 A 的非零奇异值。式(8.3.2)通常记为

$$A = UDV^T \quad (8.3.3)$$

证明 记对称矩阵 $A^T A$ 的特征值为

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > \lambda_{r+1} = \dots = \lambda_n = 0$$

由对称矩阵的正交分解, 我们有

$$V^T (A^T A) V = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) = \begin{pmatrix} \Sigma^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \quad (8.3.4)$$

将 V 分块为

$$V = (V_1 | V_2), \quad V_1 \in R^{n \times r}, V_2 \in R^{n \times (n-r)}$$

并将(8.3.4)改写为

$$(A^T A) V = V \begin{pmatrix} \Sigma^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

则有

$$A^T A V_1 = V_1 \Sigma^2, \quad A^T A V_2 = \mathbf{0} \quad (8.3.5)$$

由(8.3.5)中第一式得到

$$V_1^T A^T A V_1 = \Sigma^2, \quad \text{或写成 } (A V_1 \Sigma^{-1})^T (A V_1 \Sigma^{-1}) = I_r$$

由(8.3.5)中第二式得到

$$(A V_2)^T (A V_2) = \mathbf{0}, \quad \text{或写成 } A V_2 = \mathbf{0}$$

令 $U_1 = A V_1 \Sigma^{-1}$, 则 $U_1^T U_1 = I_r$, 即 U_1 的 r 个列向量是两两正交的单位向量, 记为

$$U_1 = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r)$$

将 $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r\}$ 扩充为 m 维空间的一组正交基 $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r, \mathbf{u}_{r+1}, \dots, \mathbf{u}_m\}$ 。令

$$U_2 = (\mathbf{u}_{r+1}, \dots, \mathbf{u}_m)$$

则 $U = (U_1 | U_2)$ 是个正交矩阵, 且

$$U_1^T U_1 = I_r, \quad U_2^T U_1 = \mathbf{0}$$

这样, 我们有

$$U^T AV = U^T (AV_1 | AV_2) = \begin{pmatrix} U_1^T \\ U_2^T \end{pmatrix} (U_1 \Sigma | \mathbf{0}) = \begin{pmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

故有 $A = U \begin{pmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} V^T$ 。证毕。

记 $U = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m)$, $V = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$, 由 A 的 SVD 分解, 不难验证:

(1) A 的零空间:

$$N(A) = \text{span}\{\mathbf{v}_{r+1}, \mathbf{v}_{r+2}, \dots, \mathbf{v}_n\} \quad (8.3.6)$$

(2) A 的值空间:

$$R(A) = \text{span}\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r\} \quad (8.3.7)$$

(3) A 可分解为:

$$A = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \dots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T \quad (8.3.8)$$

通常也称式(8.3.8)为 A 的奇异值分解, 并且称 \mathbf{u}_i 为奇异值 σ_i 的左奇异向量, \mathbf{v}_i 为奇异值 σ_i 的右奇异向量。

8.3.3 奇异值的极性

矩阵 A 的奇异值 $\sigma(A)$ 和半正定矩阵 $A^T A$ 的特征值 $\lambda(A^T A)$ 有下述关系 $\sigma(A) = \sqrt{\lambda(A^T A)}$ 。因此, 可利用半正定矩阵特征值的极性研究奇异值的极性。

定理 8.3.3 设 $A \in R_r^{m \times n}$ 的奇异值为

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > \sigma_{r+1} = \dots = \sigma_n = 0$$

则第 k 个奇异值与第 $n-k+1$ 个奇异值具有下述性质:

$$\sigma_k = \max_{V_k} \left(\min_{\substack{\mathbf{x} \in V_k \\ \mathbf{x} \neq \mathbf{0}}} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \right), \quad \sigma_{n-k+1} = \min_{V_k} \left(\max_{\substack{\mathbf{x} \in V_k \\ \mathbf{x} \neq \mathbf{0}}} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \right),$$

其中 V_k 为 R^n 的任一 k 维子空间。

证明 设 $A^T A$ 的特征值为

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > \lambda_{r+1} = \dots = \lambda_n = 0$$

于是有 $\sigma_j = \sqrt{\lambda_j}$ ($j=1, 2, \dots, n$), 对 $A^T A$ 应用定理 7.4.3, 有

$$\sigma_k = \sqrt{\lambda_k} = \sqrt{\max_{V_k} \left(\min_{\substack{\mathbf{x} \in V_k \\ \mathbf{x} \neq \mathbf{0}}} \frac{\mathbf{x}^T A^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \right)} = \max_{V_k} \left(\min_{\substack{\mathbf{x} \in V_k \\ \mathbf{x} \neq \mathbf{0}}} \sqrt{\frac{\mathbf{x}^T A^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}}} \right) = \max_{V_k} \left(\min_{\substack{\mathbf{x} \in V_k \\ \mathbf{x} \neq \mathbf{0}}} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \right)$$

同理可证

$$\sigma_{n-k+1} = \min_{V_k} \left(\max_{\substack{x \in V_k \\ x \neq 0}} \frac{\|Ax\|_2}{\|x\|_2} \right)$$

定理 8.3.4 设 $A \in R_r^{m \times n}$ 的奇异值为

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > \sigma_{r+1} = \dots = \sigma_n = 0$$

$(A+Q) \in R_r^{m \times n}$ 的奇异值为 $\tau_1 \geq \tau_2 \geq \dots \geq \tau_r > \tau_{r+1} = \dots = \tau_n = 0$ ，则必有

$$|\sigma_j - \tau_j| \leq \|Q\|_2 \quad (j=1,2,\dots,n)$$

定理表明，矩阵 A 在摄动 Q 下，奇异值的变化量不超过 $\|Q\|_2$ 。因此，矩阵奇异值的计算具有良好的数值性质。

证明 设 $A^T A$ 的特征值为 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > \lambda_{r+1} = \dots = \lambda_n = 0$ ，相应的单位特征向量为 x_1, x_2, \dots, x_n ，记

$$V_j = \text{Span}\{x_1, x_2, \dots, x_j\}, j=1,2,\dots,n$$

则有

$$\sigma_j = \sqrt{\lambda_j} = \sqrt{\left(\max_{\substack{x \in V_j \\ x \neq 0}} \frac{x^T A^T A x}{x^T x} \right)} = \max_{\substack{x \in V_j \\ x \neq 0}} \frac{\|Ax\|_2}{\|x\|_2}$$

对于矩阵 $(A+Q)$ 有

$$\tau_j \leq \max_{\substack{x \in V_j \\ x \neq 0}} \frac{\|(A+Q)x\|_2}{\|x\|_2} \leq \max_{\substack{x \in V_j \\ x \neq 0}} \frac{\|Ax\|_2}{\|x\|_2} + \max_{\substack{x \in V_j \\ x \neq 0}} \frac{\|Qx\|_2}{\|x\|_2} \leq \sigma_j + \|Q\|_2$$

同理，考虑 $(A+Q)^T(A+Q)$ 的特征值与单位特征向量，可以得到 $\sigma_j \leq \tau_j + \|Q\|_2$ ，因此有

$$|\sigma_j - \tau_j| \leq \|Q\|_2 \quad (j=1,2,\dots,n)$$

8.3.4 极分解

对于非零复数 z 可表示为极（坐标）形式： $z = \rho e^{i\theta}$ 。下面，指出对于非零复方矩阵也可以表示为极形式，通常称为矩阵的极分解：

$$A = P e^{iB}$$

其中 P 为半正定的厄米特矩阵， B 为厄米特矩阵。

定理 8.3.5(极分解) 令 A 为非零复(实)方阵，则存在唯一的半正定厄米特(实)矩阵 P 和酉(正交)矩阵 U 使得

$$A = PU = Pe^B$$

证明 对 A 作奇异值分解：

$$A = Q \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V^H$$

令

$$P = Q \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} Q^H, U = QV^H$$

则 P 是半正定厄米特矩阵， U 是酉矩阵且 $A = PU$ 。证毕。

8.4 最小二乘问题

考虑线性系统：

$$Ax = \mathbf{b} \quad (A \in R^{m \times n} (m > n), \mathbf{b} \in R^m, \mathbf{x} \in R^n)$$

的最小二乘解。即求 $\mathbf{x} \in R^n$ 使得

$$\|Ax - \mathbf{b}\|_2 = \min \{ \|Av - \mathbf{b}\|_2 : v \in R^n \} \quad (8.4.1)$$

记

$$\mathbf{X}_{LS} = \{ \mathbf{x} \in R^n : \mathbf{x} \text{ 是 (8.4.1) 的解} \} \quad (8.4.2)$$

则称 \mathbf{X}_{LS} 是最小二乘问题(8.4.1)的解集； \mathbf{X}_{LS} 中范数最小者称为最小范数解，并记作 \mathbf{x}_{LS} ，即

$$\|\mathbf{x}_{LS}\|_2 = \min \{ \|\mathbf{x}\|_2 : \mathbf{x} \in \mathbf{X}_{LS} \}$$

命题 8.4.1 $\mathbf{x} \in \mathbf{X}_{LS} \Leftrightarrow A^T(A\mathbf{x} - \mathbf{b}) = 0$

证明 $\forall \mathbf{x}, \mathbf{y} \in R^n$, 有

$$\|\mathbf{b} - A(\mathbf{x} + \mathbf{y})\|_2^2 = \|\mathbf{b} - A\mathbf{x}\|_2^2 - 2\mathbf{y}^T A^T(\mathbf{b} - A\mathbf{x}) + \|\mathbf{A}\mathbf{y}\|_2^2$$

因此，

$$\mathbf{x} \in \mathbf{X}_{LS} \Leftrightarrow \forall \mathbf{y} \in R^n, \|\mathbf{A}\mathbf{y}\|_2^2 - 2\mathbf{y}^T A^T(\mathbf{b} - A\mathbf{x}) \geq 0 \Leftrightarrow A^T(\mathbf{b} - A\mathbf{x}) = 0。$$

方程 $A^T(A\mathbf{x} - \mathbf{b}) = 0$ 称为 $A\mathbf{x} - \mathbf{b} = 0$ 的正规方程。

根据命题 8.4.1，有下述推论：

推论 8.4.1: (1) \mathbf{X}_{LS} 是凸集；(2) \mathbf{x}_{LS} 是唯一的；(3) $\mathbf{X}_{LS} = \{\mathbf{x}_{LS}\}$ 的充分必要条件是 $\text{rank}(A) = n$ 。

Moore-Pseudo 广义逆 为了给出最小二乘解的一般表示, 需要矩阵的广义逆的概念。

定义 8.4.1 $A \in R^{m \times n}$, 若 $X \in R^{n \times m}$ 满足

$$AXA = A, \quad XAX = X, \quad (AX)^T = AX, \quad (XA)^T = XA$$

则称 X 是 A 的 Moore-Pseudo 逆, 简称广义逆, 并记为 A^+ 。

矩阵 A 的广义逆是唯一的, 并且可以利用 A 的 SVD 分解进行计算。令 A 的 SVD 分解为

$$A = U \begin{pmatrix} \Sigma_r & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} V^T$$

不难验证:

$$A^+ = V \begin{pmatrix} \Sigma_r^{-1} & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} U^T \quad (8.4.3)$$

命题 8.4.2 最小二乘问题(8.4.1)的一般解为

$$\mathbf{x} = A^+ \mathbf{b} + (I - A^+ A) \mathbf{z}, \quad \mathbf{z} \in R^n \quad (8.4.4)$$

最小范数解是

$$\mathbf{x}_{LS} = A^+ \mathbf{b} \quad (8.4.5)$$

证明 由命题 8.4.1, 最小二乘问题(8.4.1)的解可由它的正规化方程:

$$A^T A \mathbf{x} = A^T \mathbf{b} \quad (8.4.6)$$

给出。利用广义逆, 可验证 $\mathbf{x} = A^+ \mathbf{b}$ 是(8.4.6)的一个解。另外, 由 SVD 分解可证明 $N(A^T A) = N(A)$, 且

$$N(A) = \{(I - A^+ A) \mathbf{z} : \mathbf{z} \in R^n\}$$

所以, 根据线性方程组理论, 方程(8.4.6)的所有解为

$$\mathbf{x} = A^+ \mathbf{b} + (I - A^+ A) \mathbf{z}, \quad \mathbf{z} \in R^n$$

此外, 注意到

$$[(I - A^+ A) \mathbf{z}]^T A^+ \mathbf{b} = \mathbf{z}^T (I - A^+ A) A^+ \mathbf{b} = 0$$

所以,

$$\|\mathbf{x}\|_2^2 = \|A^+ \mathbf{b}\|_2^2 + \|(I - A^+ A) \mathbf{z}\|_2^2 \geq \|A^+ \mathbf{b}\|_2^2$$

即 $\mathbf{x}_{LS} = A^+ \mathbf{b}$ 。

在求解最小二乘问题(8.4.1)中, 我们通常关心的是极小范数解 \mathbf{x}_{LS} 。

最小二乘解的扰动分析

假定 $A, \delta A \in R^{m \times n}$ 和 $\mathbf{b}, \delta \mathbf{b} \in R^m$ ，设 $\mathbf{x}, \mathbf{x} + \delta \mathbf{x} \in R^n$ 分别为最小二乘问题

$$\|A\mathbf{x} - \mathbf{b}\|_2 = \min \{ \|A\mathbf{v} - \mathbf{b}\|_2 : \mathbf{v} \in R^n \}$$

和

$$\|(A + \delta A)(\mathbf{x} + \delta \mathbf{x}) - (\mathbf{b} + \delta \mathbf{b})\|_2 = \min \{ \|(A + \delta A)\mathbf{v} - (\mathbf{b} + \delta \mathbf{b})\|_2 : \mathbf{v} \in R^n \}$$

的最小范数解，即

$$\begin{aligned} \mathbf{x} &= A^+ \mathbf{b} \\ \mathbf{x} + \delta \mathbf{x} &= (A + \delta A)^+ (\mathbf{b} + \delta \mathbf{b}) \end{aligned}$$

现在考虑 δA 和 $\delta \mathbf{b}$ 的大小对 $\delta \mathbf{x}$ 的影响。

由于广义逆的不连续性，即 $\delta \rightarrow 0$ 时，不一定有 $(A + \delta A)^+ \rightarrow A^+$ ，这使得最小二乘解的扰动分析变得复杂化。但是，广义逆的连续性与保秩扰动之间具有内在联系。实际上， $\lim_{\|\delta A\|_2 \rightarrow 0} (A + \delta A)^+ = A^+$ 的充要条件是当 δ 充分小时， $\text{rank}(A + \delta A) = \text{rank}(A)$ 。因此，通常都在扰动 δA 不改变 A 的秩的前提下，考虑最小二乘解的扰动分析。记

$$\begin{aligned} \bar{A} &= A + \delta A, \quad \bar{\mathbf{x}} = \mathbf{x} + \delta \mathbf{x}, \quad \bar{\mathbf{b}} = \mathbf{b} + \delta \mathbf{b}, \\ \varepsilon_A &= \|\delta A\|_2 / \|A\|_2, \quad k = k_2(A) = \|A\|_2 \|A^+\|_2, \quad \eta = k \varepsilon_A. \end{aligned}$$

命题 8.4.3 如果 $\text{rank}(A) = \text{rank}(\bar{A})$ ，且 $\eta < 1$ ，则有

$$\|\delta \mathbf{x}\|_2 \leq \frac{k}{1 - \eta} \left(\varepsilon_A \|\mathbf{x}\|_2 + \frac{\|\delta \mathbf{b}\|_2 + \varepsilon_A k \|\mathbf{b} - A\mathbf{x}\|_2}{\|A\|_2} \right) + \varepsilon_A k \|\mathbf{x}\|_2 \tag{8.4.7}$$

这个结论的证明比较复杂，证明略。

命题 8.4.3 表示 $k_2(A)$ 的大小，在一定程度上反映了最小二乘问题解对扰动的敏感程度，称 $k_2(A)$ 为最小二乘问题(8.4.1)的条件数。当 $k_2(A)$ 很大时，就说问题(8.4.1)是病态的；否则就说它是良态的。

8.4.1 满秩最小二乘问题

如果(8.4.1)中的矩阵 A 是列满秩的，即 $\text{rank}(A) = n$ ，则称它为满秩最小二乘问题。下面考虑满秩最小二乘问题的数值算法。

正规化方法

将求解问题(8.4.1)转化为求解正规化方程组

$$A^T A \mathbf{x} = A^T \mathbf{b} \quad (8.4.8)$$

由于 $\text{rank}(A) = n$ ，所以 $A^T A$ 是对称正定矩阵，因而(8.4.8)的唯一解 \mathbf{x}_{LS} 可用 Cholesky 分解法求得。正规化的基本步骤为：

- ① 计算 $C = A^T A$, $\mathbf{d} = A^T \mathbf{b}$;
- ② 对 C 进行 Cholesky 分解 $C = GG^T$;
- ③ 求解三角方程 $G\mathbf{y} = \mathbf{d}$ 以及 $G^T \mathbf{x} = \mathbf{y}$ 。

QR 分解方法

设 A 有 QR 分解：

$$A = Q \begin{pmatrix} R \\ 0 \end{pmatrix} = Q_1 R$$

其中 $Q \in R^{m \times n}$ 是正交矩阵， Q_1 是 Q 的前 n 列组成的矩阵，即 $Q = \begin{pmatrix} Q_1 & Q_2 \\ & \end{pmatrix}$ ， $R \in R^{m \times n}$ 是对角线上元素均为正数的上三角矩阵。

由于正交矩阵保持范数不变，所以问题(8.4.1)等价于

$$\|Q^T(A\mathbf{x} - \mathbf{b})\|_2 = \min \{ \|Q^T(A\mathbf{v} - \mathbf{b})\|_2 : \mathbf{v} \in R^n \}$$

记

$$\mathbf{d} = Q^T \mathbf{b} = \begin{pmatrix} Q_1^T \\ Q_2^T \end{pmatrix} \mathbf{b} = \begin{pmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \end{pmatrix}$$

则有

$$\|Q^T(A\mathbf{x} - \mathbf{b})\|_2^2 = \left\| \begin{pmatrix} R \\ \mathbf{0} \end{pmatrix} \mathbf{x} - \begin{pmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \end{pmatrix} \right\|_2^2 = \|R\mathbf{x} - \mathbf{d}_1\|_2^2 + \|\mathbf{d}_2\|_2^2$$

因此， \mathbf{x} 是(8.4.1)的解当且仅当 \mathbf{x} 是方程 $R\mathbf{x} = \mathbf{d}_1$ 的解。这样，(8.4.1)的解可由上三角方程组 $R\mathbf{x} = \mathbf{d}_1$ 求得。

QR 分解方法的基本步骤如下：

- ① 求 A 的 QR 分解；
- ② 计算 $\mathbf{d}_1 = Q_1^T \mathbf{b}$ ；
- ③ 解方程组 $R\mathbf{x} = \mathbf{d}_1$ 。

值得指出的是，QR 分解方法比正规化方法有较好的数值稳定性，并且计算结果比正规化方法要精确。当然，QR 方法比正规化方法会付出更大的计算代价。

SVD 分解方法

最常用的方法是 SVD 分解方法。由于 $\text{rank}(A) = n$, 所以 A 必有下述形式的 SVD 分解

$A = U \begin{pmatrix} \Sigma_n \\ 0 \end{pmatrix} V^T$, 于是, $A^+ = V(\Sigma_n^{-1}, 0)U^T$ 。所以, 问题(8.4.1)的解为

$$\mathbf{x} = A^+ \mathbf{b} = V(\Sigma_n^{-1}, 0)U^T \mathbf{b} = \frac{\mathbf{u}_1^T \mathbf{b}}{\sigma_1} \mathbf{v}_1 + \frac{\mathbf{u}_2^T \mathbf{b}}{\sigma_2} \mathbf{v}_2 + \cdots + \frac{\mathbf{u}_n^T \mathbf{b}}{\sigma_n} \mathbf{v}_n = \sum_{j=1}^n \frac{\mathbf{u}_j^T \mathbf{b}}{\sigma_j} \mathbf{v}_j \quad (8.4.9)$$

上式给出了 SVD 分解方法关于最小二乘解的计算公式。

8.4.2 亏秩最小二乘问题

如果在最小二乘问题(8.4.1)中, 矩阵 A 是亏秩的, 即 $\text{rank}(A) < n$ 。此时(8.4.1)有无穷多解, 在上面介绍的处理满秩问题(8.4.1)的正规化方法, QR 分解方法都将失败, 不能给出最小范数解 \mathbf{x}_{LS} 。但是, SVD 分解方法仍然有效。具体地说, 若 $\text{rank}(A) = r$, 则 A 有 SVD 分解:

$$A = U \begin{pmatrix} \Sigma_r & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} V^T$$

因此,

$$\mathbf{x} = A^+ \mathbf{b} = V \begin{pmatrix} \Sigma_r^{-1} & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} U^T \mathbf{b} = \sum_{j=1}^r \frac{\mathbf{u}_j^T \mathbf{b}}{\sigma_j} \mathbf{v}_j \quad (8.4.10)$$

在这里, 可以看到 SVD 分解在数值计算中的作用, 不论是满秩的还是亏秩的最小二乘问题, SVD 分解方法总能给出它们的求解计算公式, 并且有统一的形式。下面将会进一步看到 SVD 分解的重要作用。

8.4.3 数值秩的定义和确定方法

从上面的讨论, 可以看出亏秩最小二乘问题(8.4.1)求解与矩阵的秩密切相关。然而“秩”这一在数学上精确定义的概念, 在数据有误差时或者在计算机上进行浮点运算时, 就变得模糊不清了。这样就自然地引进了所谓的数值秩的概念。

定义 8.4.1 设 $A \in R^{m \times n}$, 如果对某一正数 ε , 有

$$r = \min\{\text{rank}(B) : B \in R^{m \times n}, \|B - A\|_2 < \varepsilon\} \quad (8.4.11)$$

则称 r 为矩阵 A 的 ε 数值秩。

下面考虑数值秩的确定问题。

设 $A \in R^{m \times n}$ ($m \geq n$) 的奇异值 $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n$, 对应的左, 右奇异向量分别为 $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$; $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$, 则

$$\inf_{\text{rank}(\mathbf{B}) \leq k} \|\mathbf{A} - \mathbf{B}\|_2 = \sigma_{k+1}, \quad k = 1, 2, \dots, n$$

且下确界在 $\mathbf{B} = \sum_{j=1}^k \sigma_j \mathbf{u}_j \mathbf{v}_j^T$ 达到。因此, 矩阵 A 的 ε 数值秩是 r 的充分必要条件是

$$\sigma_r \geq \varepsilon \geq \sigma_{r+1} \quad (8.4.12)$$

于是, 可以使用 SVD 分解来确定矩阵的 ε 数值秩。

由于 SVD 分解的良好数值性态, 现在人们普遍认为奇异值分解方法是确定数值秩的最可靠方法。

值得指出的是: 要使(8.4.12)成立, σ_{r+1} , σ_r 之间必须具有一定的距离, 如果矩阵 A 的奇异值分离不明显, 而且又是亏秩时, 应用 SVD 分解确定矩阵 A 的 ε 数值秩就有一定的困难, 此时需要更复杂的方法来处理。

8.4.4 齐次最小二乘问题

考虑齐次线性方程组:

$$\mathbf{Ax} = \mathbf{0} \quad (\mathbf{A} \in R^{m \times n} (m > n), \mathbf{x} \in R^n, m > n)$$

对应的最小二乘问题是

$$\|\mathbf{Ax}\|_2 = \min \{ \|\mathbf{Av}\|_2 : \mathbf{v} \in R^n \} \quad (8.4.13)$$

显然, $\mathbf{x} = \mathbf{0}$ 总是上述最小二乘问题的最小范数解。在实际中, 人们所关心的并非是齐次最小二乘问题的零解, 而是它的非零解。因此, 总是考虑相应的约束最小二乘问题:

$$\|\mathbf{Ax}\|_2 = \min \{ \|\mathbf{Av}\|_2 : \mathbf{v} \in R^n, \|\mathbf{v}\|_2 = 1 \} \quad (8.4.14)$$

或者等价地写成:

$$\begin{cases} \min \|\mathbf{Ax}\|_2^2 \\ \text{subject to } \|\mathbf{x}\|_2 = 1 \end{cases} \quad (8.4.15)$$

命题 8.4.4 若 $\text{rank}(A) = r$, 则(8.4.15)的所有解为

$$\mathbf{x} = \sum_{j=1}^{n-r} s_j \mathbf{v}_{r+j}, \quad \left(\sum_{j=1}^{n-r} s_j^2 = 1 \right)$$

其中, $\mathbf{v}_{r+1}, \mathbf{v}_{r+2}, \dots, \mathbf{v}_n$ 是 $A^T A$ 的零特征值的 $n-r$ 个线性无关的单位特征向量。

证明 问题(8.4.15)等价于

$$\begin{cases} \min \mathbf{x}^T A^T A \mathbf{x} \\ \text{subject to } \|\mathbf{x}\|_2 = 1 \end{cases} \quad (8.4.16)$$

因 $\text{rank}(A^T A) = \text{rank}(A) = r$, 所以 $A^T A$ 有特征值分解

$$A^T A = V \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_r, 0, \dots, 0) V^T$$

其中 V 是正交矩阵, 所以 $\mathbf{v}_{r+1}, \mathbf{v}_{r+2}, \dots, \mathbf{v}_n$ 是 $A^T A \mathbf{x} = 0$ 的 $n-r$ 个相互正交的单位解。因此(8.4.16)的所有解为

$$\mathbf{x} = \sum_{j=1}^{n-r} s_j \mathbf{v}_{r+j}, \quad \left(\sum_{j=1}^{n-r} s_j^2 = 1 \right)$$

利用矩阵 A 的 SVD 分解, 也可以求解(8.4.15)。

命题 8.4.5 若 $\text{rank}(A) = r$, 则(8.4.15)的所有解为

$$\mathbf{x} = \sum_{j=1}^{n-r} s_j \mathbf{v}_{r+j}, \quad \left(\sum_{j=1}^{n-r} s_j^2 = 1 \right)$$

其中, $\mathbf{v}_{r+1}, \mathbf{v}_{r+2}, \dots, \mathbf{v}_n$ 是 A 的 $n-r$ 个零奇异值的右奇异向量。

推论 8.4.6 若 $\text{rank}(A) = n-1$ 时, (8.4.15)有唯一解, 这个唯一解是 $A^T A$ 的零特征值的单位特征向量, 或者说是 A 的零奇异值的右奇异向量。当数据有误差时, $A^T A$ 的最小特征值的单位特征向量(或者说是 A 的最小奇异值的右奇异向量)是(8.4.15)的一个解。

推论 8.4.6 的前半部分可由命题 8.4.4 和命题 8.4.5 直接推出。当数据有误差时, $A^T A$ 的特征值分解必为下述形式

$$A^T A = V \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) V^T$$

其中 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$, 且 $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ 构成 R^n 的一组单位正交基。 $\forall \mathbf{x} \in R^n, \|\mathbf{x}\| = 1$ 则必有

$$\mathbf{x} = \sum_{j=1}^n s_j \mathbf{v}_j, \quad \left(\sum_{j=1}^n s_j^2 = 1 \right)$$

且

$$\begin{aligned} \|A^T A \mathbf{x}\|_2^2 &= \left\| V \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) V^T \cdot \sum_{j=1}^n s_j \mathbf{v}_j \right\|_2^2 = \left\| V \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) \cdot \sum_{j=1}^n s_j \mathbf{e}_j \right\|_2^2 \\ &= \sum_{j=1}^n \|\lambda_j s_j \mathbf{v}_j\|_2^2 = \sum_{j=1}^n \lambda_j^2 s_j^2 \geq \lambda_n^2 = \|A^T A \mathbf{v}_n\|_2^2 \end{aligned}$$

故, \mathbf{v}_n 是(8.4.15)的一个解。

8.4.5 约束齐次最小二乘问题

在本节, 将会看到 SVD 分解在求解一般约束齐次最小二乘问题中的重要作用。

(I) 考虑约束最小二乘问题

$$\begin{cases} \min \|A\mathbf{x}\|_2^2 \\ \text{subject to } \|\mathbf{x}\|_2 = 1, C\mathbf{x} = \mathbf{0} \end{cases} \quad (8.4.17)$$

不妨假定 $\text{rank}(C) = r < n$, 对 C 作 SVD 分解

$$C = U \begin{pmatrix} \Sigma_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} V^T$$

令 $V_{n-r} = (\mathbf{v}_{r+1}, \mathbf{v}_{r+2}, \dots, \mathbf{v}_n)$ 为 V 的最后 $n-r$ 个列向量所构成的矩阵, 由(8.3.6)式知 $C\mathbf{x} = \mathbf{0}$ 的所有解可表示为:

$$\mathbf{x} = V_{n-r}\mathbf{y}, \quad \mathbf{y} \in \mathbf{R}^{n-r}$$

并且 $\|\mathbf{x}\|_2 = 1 \Leftrightarrow \|\mathbf{y}\|_2 = 1$, 这是因为 V_{n-r} 是列正交的。于是(8.4.17)化为下述最小化问题

$$\begin{cases} \min \|AV_{n-r}\mathbf{y}\|_2^2 \\ \text{subject to } \|\mathbf{y}\|_2 = 1 \end{cases}$$

这是上节已解决的问题。

(II) 考虑约束最小二乘问题

$$\begin{cases} \min \|A\mathbf{x}\|_2^2 \\ \text{subject to } \|\mathbf{x}\|_2 = 1, \mathbf{x} = C\mathbf{y} \end{cases} \quad (8.4.18)$$

对 C 作 SVD 分解 $C = U \begin{pmatrix} \Sigma_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} V^T$, 则由(8.2.7)式 C 的值空间可表示为

$$R(C) = \{U_r\mathbf{x}' \mid \mathbf{x}' \in \mathbf{R}^r\}$$

其中 U_r 为 U 的前 r 列所构成的矩阵。于是 $\mathbf{x} = C\mathbf{y}$ 可表示为 $\mathbf{x} = U_r\mathbf{x}'$ 且有 $\|\mathbf{x}\|_2 = 1 \Leftrightarrow \|\mathbf{x}'\|_2 = 1$ 。于是(8.4.18)化为

$$\begin{cases} \min \|AU_r\mathbf{x}'\|_2^2 \\ \text{subject to } \|\mathbf{x}'\|_2 = 1 \end{cases}$$

这是已解决的问题。

(III) 考虑约束最小二乘问题

$$\begin{cases} \min \|A\mathbf{x}\|_2^2 \\ \text{subject to } \|C\mathbf{x}\|_2 = 1 \end{cases} \quad (8.4.19)$$

对 C 作 SVD 分解 $C = U \begin{pmatrix} \Sigma_r & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} V^T$, 并令 $\mathbf{x}' = V^T \mathbf{x}$, 则

$$C\mathbf{x} = U \begin{pmatrix} \Sigma_r & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} \mathbf{x}'$$

且 $\|C\mathbf{x}\|_2 = 1 \Leftrightarrow \left\| \begin{pmatrix} \Sigma_r & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} \mathbf{x}' \right\|_2 = 1 \Leftrightarrow \|\Sigma_r \mathbf{x}'\|_2 = 1$, 其中 $\mathbf{x}' = \begin{pmatrix} \mathbf{x}'_r \\ \mathbf{x}'_{n-r} \end{pmatrix}$ 。所以,

$$A\mathbf{x} = AV\mathbf{x}' = A'\mathbf{x}' = \begin{pmatrix} A'_r & A'_{n-r} \end{pmatrix} \begin{pmatrix} \mathbf{x}'_r \\ \mathbf{x}'_{n-r} \end{pmatrix} = A'_r \mathbf{x}'_r + A'_{n-r} \mathbf{x}'_{n-r}$$

取 $\mathbf{y} = \Sigma_r \mathbf{x}'_r$, 即 $\mathbf{x}'_r = \Sigma_r^{-1} \mathbf{y}$, 于是(8.4.19)化为最小化问题

$$\begin{cases} \min \|A'_r \Sigma_r^{-1} \mathbf{y} + A'_{n-r} \mathbf{x}'_{n-r}\|_2^2 \\ \text{subject to } \|\mathbf{y}\|_2 = 1 \end{cases} \quad (8.4.20)$$

由于 $\min_{\mathbf{x}'_{n-r}} \|A'_r \Sigma_r^{-1} \mathbf{y} + A'_{n-r} \mathbf{x}'_{n-r}\|_2$ 的最小化问题的解为

$$\mathbf{x}'_{n-r} = -A'^{+\dagger}_{n-r} A'_r \mathbf{x}'_r$$

于是(8.4.20)归结为求最小化问题:

$$\begin{cases} \min \|(A'_r - A'_{n-r} A'^{+\dagger}_{n-r} A'_r) \Sigma_r^{-1} \mathbf{y}\|_2^2 \\ \text{subject to } \|\mathbf{y}\|_2 = 1 \end{cases}$$

这样就归结为所熟知的最小化问题。

9. 矩阵分析

9.1 向量与矩阵范数

9.1.1 向量范数

定义 9.1.1 设 V 为数域 F 上的线性空间，若实值函数 $\|\cdot\|: V \rightarrow R^+ \cup \{0\}$ 满足下述三个条件：

- (a) 非负性： $\forall x \in V, \|x\| \geq 0$ ，且 $\|x\| = 0 \Leftrightarrow x = \theta$ ；
- (b) 齐次性： $\forall x \in V, k \in P, \|kx\| = |k| \cdot \|x\|$ ；
- (c) 三角不等式： $\forall x, y \in V, \|x + y\| \leq \|x\| + \|y\|$ 。

则称 $\|\cdot\|$ 为 V 上的范数， $\|x\|$ 称为向量 x 的范数。

p -范数 $\forall x = (x_1, x_2, \dots, x_n)^T \in C^n(R^n)$ ，定义

$$\|x\|_p = \left(\sum_{j=1}^n |x_j|^p \right)^{1/p} \quad (1 \leq p < +\infty) \quad (9.1.1)$$

则称 $\|\cdot\|_p$ 为 $C^n(R^n)$ 上的范数，这种范数通常称为 p -范数。

证明 根据范数的定义，只须验证 $\|\cdot\|_p$ 满足条件(a), (b), (c)。条件(a)显然满足。对于条件(b)，有

$$\|kx\|_p = \left(\sum_{j=1}^n |kx_j|^p \right)^{1/p} = |k| \left(\sum_{j=1}^n |x_j|^p \right)^{1/p} = |k| \cdot \|x\|_p$$

下面证明条件(c)满足：因为

$$\|x + y\|_p^p = \sum_{j=1}^n |x_j + y_j|^p \leq \sum_{j=1}^n |x_j + y_j|^{p-1} |x_j| + \sum_{j=1}^n |x_j + y_j|^{p-1} |y_j|$$

所以，当 $p > 0$ 时，对上式后两项分别使用 Holder 不等式^{*)}，我们有

^{*)} $\sum_{j=1}^n |a_j b_j| \leq \left(\sum_{j=1}^n |a_j|^p \right)^{1/p} \cdot \left(\sum_{j=1}^n |b_j|^q \right)^{1/q}$ ， $(p, q > 1, \frac{1}{p} + \frac{1}{q} = 1)$

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|_p^p &\leq \left(\sum_{j=1}^n |x_j|^p\right)^{1/p} \cdot \left(\sum_{j=1}^n |x_j + y_j|^{(p-1)q}\right)^{1/q} + \left(\sum_{j=1}^n |y_j|^p\right)^{1/p} \cdot \left(\sum_{j=1}^n |x_j + y_j|^{(p-1)q}\right)^{1/q} \\ &= \left[\left(\sum_{j=1}^n |x_j|^p\right)^{1/p} + \left(\sum_{j=1}^n |y_j|^p\right)^{1/p} \right] \cdot \left(\sum_{j=1}^n |x_j + y_j|^p\right)^{1/q} \\ &= (\|\mathbf{x}\|_p + \|\mathbf{y}\|_p) \|\mathbf{x} + \mathbf{y}\|_p^{p/q} \end{aligned}$$

因此,

$$\|\mathbf{x} + \mathbf{y}\|_p = \|\mathbf{x} + \mathbf{y}\|_p^{p-1/q} \leq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p$$

当 $p=1$ 时, 显然有

$$\|\mathbf{x} + \mathbf{y}\|_1 = \sum_{j=1}^n |x_j + y_j| \leq \sum_{j=1}^n |x_j| + \sum_{j=1}^n |y_j| = \|\mathbf{x}\|_1 + \|\mathbf{y}\|_1$$

所以, 式(9.1.1)是 $C^n(\mathbb{R}^n)$ 上的范数。证毕。

在 $C^n(\mathbb{R}^n)$ 上定义

$$\|\mathbf{x}\| = \max_j |x_j| \quad (9.1.2)$$

不难证明, 它也是 $C^n(\mathbb{R}^n)$ 上的范数。事实上, 式(9.1.2)是式(9.1.1)在 $p \rightarrow +\infty$ 时的极限, 即

$$\lim_{p \rightarrow +\infty} \|\mathbf{x}\|_p = \max_j |x_j|$$

自然地, 引进记号 $\|\mathbf{x}\|_\infty = \max_j |x_j|$ 。这样, 式(9.1.1)就扩张到 $p = +\infty$ 的情况, 即 $C^n(\mathbb{R}^n)$ 上的 p -范数:

$$\|\mathbf{x}\|_p = \left(\sum_{j=1}^n |x_j|^p\right)^{1/p}, \quad 1 \leq p \leq +\infty \quad (9.1.3)$$

当 $p=2$ 时, $\|\mathbf{x}\|_2 = \left(\sum_{j=1}^n x_j \bar{x}_j\right)^{1/2} = \langle \mathbf{x}, \mathbf{x} \rangle^{1/2}$, 这是通常的内积范数或欧氏范数。

根据 $C^n(\mathbb{R}^n)$ 的 p -范数, 我们可诱导 n 维线性空间的 p -范数。

给定 n 维线性空间 \mathbf{V} 的一组基向量 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$, 设 $\mathbf{x} \in \mathbf{V}$ 在这组基下的坐标向量为 $\tilde{\mathbf{x}} = (\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n)^T$, 即 $\mathbf{x} = \tilde{x}_1 \varepsilon_1 + \tilde{x}_2 \varepsilon_2 + \dots + \tilde{x}_n \varepsilon_n$, 定义

$$\|\mathbf{x}\|_p = \|\tilde{\mathbf{x}}\|_p, \quad 1 \leq p \leq +\infty$$

不难证明它是 \mathbf{V} 上的范数, 称它为 \mathbf{V} 上的 p -范数。

事实上, 对于 $C^n(\mathbb{R}^n)$ 上的任何一种范数, 按照上面的方法都可以诱导出 \mathbf{V} 上的一种范数; 反

过来也是一样，对于 V 上的任何一种范数，也可以诱导出 $C^n(R^n)$ 上的一种范数。

椭球范数(加权范数) 设 A 是 n 阶正定厄米特矩阵，定义 C^n 上的非负函数

$$\|x\|_A = (x^H Ax)^{1/2} \quad (9.1.5)$$

则 $\|\cdot\|_A$ 是 C^n 上的范数。

事实上，由于 A 是正定，所以 $(x^H Ax)^{1/2} = 0 \Leftrightarrow x = \mathbf{0}$ ，即非负性成立。 $\forall k \in C$ ，

$$\|kx\|_A = ((kx)^H A(kx))^{1/2} = |k| (x^H Ax)^{1/2} = |k| \|x\|_A$$

因此，齐次性成立。下面证明三角不等式也成立：由于 A 是正定厄米特矩阵，所以存在非奇异性矩阵 P 使得 $P^H AP = I$ ，即存在非奇异性矩阵 B 使 $A = B^H B$ ，因此 $\forall x \in C^n$ ，

$$\|x\|_A = (x^H Ax)^{1/2} = ((Bx)^H (Bx))^{1/2} = \|Bx\|_2$$

所以，

$$\|x+y\|_A = \|B(x+y)\|_2 \leq \|Bx\|_2 + \|By\|_2 = \|x\|_A + \|y\|_A$$

证毕。

在 R^n 中，范数 $\|\cdot\|_A$ 的单位球 $b = \{x \in R^n : \|x\|_A = 1\}$ 是一个椭球，因此通常称 $\|\cdot\|_A$ 为椭球范数。

向量范数的等价性

对于同一个线性空间，可以在上面定义无穷多种范数，下面讨论这些范数之间的关系。

定义 9.1.2 设 $\|\cdot\|_\alpha, \|\cdot\|_\beta$ 为 n 维线性空间 V 上的两个范数，若存在正数 c_1, c_2 使 $\forall x \in V$ 有

$$c_1 \|x\|_\beta \leq \|x\|_\alpha \leq c_2 \|x\|_\beta \quad (9.1.7)$$

则称 $\|\cdot\|_\alpha, \|\cdot\|_\beta$ 是两个相互等价的范数。

例如：在 $C^n(R^n)$ 中，不难验证

$$\|x\|_\infty \leq \|x\|_1 \leq n \|x\|_\infty, \quad \|x\|_\infty \leq \|x\|_2 \leq \sqrt{n} \|x\|_\infty$$

更一般地，

$$\|x\|_\infty \leq \|x\|_p \leq n^{1/p} \|x\|_\infty \quad (1 \leq p < +\infty)$$

由此可推知

$$n^{-1/p_1} \|x\|_{p_1} \leq \|x\|_{p_2} \leq n^{1/p_2} \|x\|_{p_1}$$

因此， $C^n(R^n)$ 上的所有 p -范数都是相互等价的。这不是偶然的的结果，因为有下列范数的等价性定理。

定理 9.1.1 n 维线性空间 V 中的任意两种范数都是等价的，即对任意范数 $\|\cdot\|_\alpha, \|\cdot\|_\beta$ ，必存在

正常数 c_1, c_2 使式(9.1.7)成立。

证明 先证: 对 V 上的任意范数 $\|\cdot\|$, 均存在正常数 a, b 使 $a\|\mathbf{x}\|_2 \leq \|\mathbf{x}\| \leq b\|\mathbf{x}\|_2$ 。

设 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ 是 V 的一组基向量, 则 $\forall \mathbf{x} \in V$, $\mathbf{x} = x_1\varepsilon_1 + x_2\varepsilon_2 + \dots + x_n\varepsilon_n$ 。所以

$$\|\mathbf{x}\| = \|x_1\varepsilon_1 + x_2\varepsilon_2 + \dots + x_n\varepsilon_n\|$$

记 $f(x_1, x_2, \dots, x_n) = \|\mathbf{x}\|$, 则 f 是 \mathbf{x} 坐标向量的连续函数, 这是因为

$$\begin{aligned} |f(x_1, x_2, \dots, x_n) - f(x'_1, x'_2, \dots, x'_n)| &= \|\mathbf{x}\| - \|\mathbf{x}'\| \leq \|\mathbf{x} - \mathbf{x}'\| \\ &= \|(x_1 - x'_1)\varepsilon_1 + (x_2 - x'_2)\varepsilon_2 + \dots + (x_n - x'_n)\varepsilon_n\| \\ &\leq \|\varepsilon_1\| \cdot |x_1 - x'_1| + \|\varepsilon_2\| \cdot |x_2 - x'_2| + \dots + \|\varepsilon_n\| \cdot |x_n - x'_n| \end{aligned}$$

由于 $\|\varepsilon_j\|$ 均为常数, 所以上面的不等式表明 f 是连续函数。令

$$S_{n-1} = \{(x_1, x_2, \dots, x_n)^T \mid |x_1|^2 + |x_2|^2 + \dots + |x_n|^2 = 1\}$$

则 S_{n-1} 是有界闭集。根据连续函数的性质, f 在 S_{n-1} 上达到最大值 b 及最小值 a 。由于在 S_{n-1} 上 $(x_1, x_2, \dots, x_n)^T \neq 0$, 所以 $b > a > 0$ 。令

$$\mathbf{y} = \frac{x_1}{\|\mathbf{x}\|_2} \varepsilon_1 + \frac{x_2}{\|\mathbf{x}\|_2} \varepsilon_2 + \dots + \frac{x_n}{\|\mathbf{x}\|_2} \varepsilon_n$$

则必有

$$\left(\frac{x_1}{\|\mathbf{x}\|_2}, \frac{x_2}{\|\mathbf{x}\|_2}, \dots, \frac{x_n}{\|\mathbf{x}\|_2} \right)^T \in S_{n-1}$$

所以,

$$a \leq f\left(\frac{x_1}{\|\mathbf{x}\|_2}, \frac{x_2}{\|\mathbf{x}\|_2}, \dots, \frac{x_n}{\|\mathbf{x}\|_2}\right) = \|\mathbf{y}\| \leq b$$

而

$$\|\mathbf{y}\| = \frac{\|x_1\varepsilon_1 + x_2\varepsilon_2 + \dots + x_n\varepsilon_n\|}{\|\mathbf{x}\|_2} = \frac{\|\mathbf{x}\|}{\|\mathbf{x}\|_2}$$

于是,

$$a\|\mathbf{x}\|_2 \leq \|\mathbf{x}\| \leq b\|\mathbf{x}\|_2$$

再证: 式(9.1.7)成立。由上面的证明, 存在正常数 $a, b; a', b'$ 使

$$a\|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_\alpha \leq b\|\mathbf{x}\|_2, \quad a'\|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_\beta \leq b'\|\mathbf{x}\|_2$$

令 $c_2 = \frac{b}{a'}, c_1 = \frac{a}{b'}$, 则必有

$$c_1 \|x\|_\beta \leq \|x\|_\alpha \leq c_2 \|x\|_\beta$$

证毕。

向量序列的极限

定义 9.1.3 设 $\{x^k\}$ 是线性空间 V 中的向量序列，若存在向量 $x^* \in V$ 使

$$\lim_{k \rightarrow +\infty} \|x^k - x^*\| = 0$$

则称 $\{x^k\}$ 是范数 $\|\cdot\|$ 下的收敛序列， x^* 为向量序列 $\{x^k\}$ 的极限，记为 $x^* = \lim_{k \rightarrow +\infty} x^k$ 。

由范数的等价性定理，向量序列的收敛性质不依赖于范数的选择，即如果向量序列在某种范数 $\|\cdot\|$ 下收敛，则对任何范数都是收敛的并且有相同的极限。这是不难证明的，假定 $\{x^k\}$ 在范数 $\|\cdot\|_\alpha$ 下收敛，即

$$\lim_{k \rightarrow +\infty} \|x^k - x^*\|_\alpha = 0$$

对任何范数 $\|\cdot\|_\beta$ ，由范数等价性定理存在 c_1, c_2 使

$$c_1 \|x^k - x^*\|_\alpha \leq \|x^k - x^*\|_\beta \leq c_2 \|x^k - x^*\|_\alpha$$

所以，

$$\lim_{k \rightarrow +\infty} \|x^k - x^*\|_\beta = 0$$

证毕。

尽管不同的向量范数可能具有不同的大小，然而在范数下考虑向量序列收敛问题时，却表现出明显的一致性。

例 9.1.1 C^n 中的向量序列 $\{x^k\}$ 收敛到向量 x^* 的充要条件为每一个分量 x_j^k 收敛到 x_j^* ，即

$$\forall j, \lim_{k \rightarrow +\infty} x_j^k = x_j^*。$$

证明 $\forall j, \lim_{k \rightarrow +\infty} x_j^k = x_j^* \Leftrightarrow \lim_{k \rightarrow +\infty} \max_j |x_j^k - x_j^*| = 0$ ，即 $\lim_{k \rightarrow +\infty} \|x^k - x^*\|_\infty = 0 \Leftrightarrow \forall$ 范数 $\|\cdot\|$ ， $\lim_{k \rightarrow +\infty} \|x^k - x^*\| = 0$ 。证毕。

9.1.2 矩阵范数

定义 9.1.4 设 $\|\cdot\|: C^{m \times n} \rightarrow R$ 是实值函数，若它满足下述三个条件：

- (a) 非负性： $\forall A \in C^{m \times n}$ ， $\|A\| \geq 0$ ，且 $\|A\| = 0 \Leftrightarrow A = \mathbf{0}$ ；
- (b) 齐次性： $\forall k \in C, A \in C^{m \times n}$ ， $\|kA\| = |k| \cdot \|A\|$ ；
- (c) 三角不等式： $\forall A, B \in C^{m \times n}$ ， $\|A+B\| \leq \|A\| + \|B\|$ 。

则称 $\|\cdot\|$ 为广义矩阵范数。

若 $\|\cdot\|$ 还满足第四个性质：

(d)相容性: $\forall A \in \mathbf{C}^{m \times n}, B \in \mathbf{C}^{n \times l}, \|AB\| \leq \|A\| \cdot \|B\|$ 。

则称 $\|\cdot\|$ 为矩阵范数。

注意: 在相容性的定义中 B 为 $\mathbf{C}^{n \times l}$ 中的矩阵, AB 为 $\mathbf{C}^{m \times l}$ 中的矩阵,实数 $\|B\|, \|AB\|$ 的定义规则与实数 $\|A\|$ 的定义规则是相同的。

矩阵范数的连续性 与向量的情况一样,对于矩阵序列也有极限概念。设矩阵序列 $\{A^{(k)}\}$,其中 $A \in \mathbf{C}^{m \times n}$,若 $A^{(k)}$ 的每一个元素 $a_{ij}^{(k)}$ 均有极限 a_{ij} ,则称 $\{A^{(k)}\}$ 有极限 $A = (a_{ij})$,或者说 $\{A^{(k)}\}$ 收敛到矩阵 A ,记为

$$\lim_{k \rightarrow +\infty} A^{(k)} = A \text{ (或 } A^{(k)} \rightarrow A \text{)}$$

不收敛的矩阵序列称为发散的。

由三角不等式,可推知 $\forall A, B \in \mathbf{C}^{m \times n}$,有 $\| \|A\| - \|B\| \| \leq \|A - B\|$ 。这是因为

$$\|A\| = \|A - B + B\| \leq \|A - B\| + \|B\|$$

即

$$\|A\| - \|B\| \leq \|A - B\|$$

同样,考虑矩阵 B ,也有 $\|B\| - \|A\| \leq \|A - B\|$ 。于是,有下述范数连续性定理:

定理 9.1.2(矩阵范数的连续性)若 $\lim_{k \rightarrow +\infty} \|A^{(k)} - A\| = 0$,则 $\lim_{k \rightarrow +\infty} \|A^{(k)}\| = \|A\|$ 。

矩阵范数与向量范数的相容性 对于 $A \in \mathbf{C}^{m \times n}$,它定义了 \mathbf{C}^n 到 \mathbf{C}^m 上的一个线性变换,因此在数值估计中,不但要考虑矩阵范数,还需要考虑矩阵范数与定义域、值域空间的向量范数之间的联系,这种联系由矩阵范数与向量范数相容的概念来实现。为此引入下面定义。

定义 9.1.5 对于 $\mathbf{C}^{m \times n}$ 上的矩阵范数 $\|\cdot\|_M$ 以及 \mathbf{C}^m 与 \mathbf{C}^n 上的向量范数 $\|\cdot\|_V$,如果 $\forall A \in \mathbf{C}^{m \times n}, x \in \mathbf{C}^n$,有

$$\|Ax\|_V \leq \|A\|_M \cdot \|x\|_V$$

则称矩阵范数 $\|\cdot\|_M$ 与向量范数 $\|\cdot\|_V$ 是相容的。

Frobenius 范数

设 $A \in \mathbf{C}^{m \times n}$,Frobenius 范数定义为

$$\|A\|_F = \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \right)^{1/2} = (\text{tr} A^H A)^{1/2} \quad (9.1.8)$$

下面证明 **Frobenius** 范数是与向量范数 $\|\cdot\|_2$ 是相容的矩阵范数。

显然, $\|\cdot\|_F$ 有非负性和齐次性。验证 $\|\cdot\|_F$ 满足三角不等式: 记矩阵 A, B 的第 j 列分别为 a_j, b_j ,

则

$$\begin{aligned} \|A+B\|_F^2 &= \sum_{j=1}^n \|\mathbf{a}_j + \mathbf{b}_j\|_2^2 \leq \sum_{j=1}^n (\|\mathbf{a}_j\|_2 + \|\mathbf{b}_j\|_2)^2 = \sum_{j=1}^n \|\mathbf{a}_j\|_2^2 + 2\sum_{j=1}^n \|\mathbf{a}_j\|_2 \cdot \|\mathbf{b}_j\|_2 + \sum_{j=1}^n \|\mathbf{b}_j\|_2^2 \\ &\leq \sum_{j=1}^n \|\mathbf{a}_j\|_2^2 + 2\left(\sum_{j=1}^n \|\mathbf{a}_j\|_2^2\right)^{1/2} \left(\sum_{j=1}^n \|\mathbf{b}_j\|_2^2\right)^{1/2} + \sum_{j=1}^n \|\mathbf{b}_j\|_2^2 \\ &= \|A\|_F^2 + 2\|A\|_F \|B\|_F + \|B\|_F^2 = (\|A\|_F + \|B\|_F)^2 \end{aligned}$$

因此, $\|A+B\|_F \leq \|A\|_F + \|B\|_F$ 。再验证矩阵乘法的相容性: $\forall B \in C^{n \times l}$, 有

$$\|AB\|_F^2 = \sum_{i=1}^m \sum_{j=1}^l \left| \sum_{k=1}^n a_{ik} b_{kj} \right|^2 \leq \sum_{i=1}^m \sum_{k=1}^n \left(\sum_{j=1}^l |a_{ik} b_{kj}| \right)^2 \leq \left(\sum_{i=1}^m \sum_{k=1}^n |a_{ik}|^2 \right) \left(\sum_{j=1}^l \sum_{k=1}^n |b_{kj}|^2 \right) = \|A\|_F^2 \cdot \|B\|_F^2$$

所以

$$\|AB\|_F \leq \|A\|_F \cdot \|B\|_F$$

最后验证与向量范数 $\|\cdot\|_2$ 的相容性: 在上式中, 令 $B = \mathbf{x} \in C^{n \times 1}$, 则有

$$\|A\mathbf{x}\|_2 = \|AB\|_F \leq \|A\|_F \cdot \|B\|_F = \|A\|_F \cdot \|\mathbf{x}\|_2$$

证毕。

Frobenius 范数具有下面的重要特性。

定理 9.1.3 设 $A \in C^{m \times n}$, 若 $Q \in C^{m \times n}$, $R \in C^{n \times n}$ 是酉矩阵, 则必有

$$\|QA\|_F = \|A\|_F = \|AR\|_F$$

此定理表明, 用酉矩阵右乘或左乘一个矩阵后, 不改变矩阵的 **Frobenius** 范数。

证明 因为

$$\|QA\|_F^2 = \|Q(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)\|_F^2 = \|(Q\mathbf{a}_1, Q\mathbf{a}_2, \dots, Q\mathbf{a}_n)\|_F^2 = \sum_{j=1}^n \|Q\mathbf{a}_j\|_2^2 = \sum_{j=1}^n \|\mathbf{a}_j\|_2^2 = \|A\|_F^2$$

所以, $\|QA\|_F = \|A\|_F$ 并且有 $\|A\|_F = \|A^H\|_F = \|R^H A^H\|_F = \|(AR)^H\|_F = \|AR\|_F$ 。证毕。

推论 9.1.1 Frobenius 范数是酉(正交)相似不变范数, 即 $\forall A \in C^{n \times n}$, $Q \in U^{n \times n}$, 有

$$\|A\|_F = \|Q^H A Q\|_F$$

算子范数(从属范数)

定理 9.1.4 已知 C^m 和 C^n 上的同类向量范数 $\|\cdot\|$, 定义

$$\|A\|_{op} = \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\| \quad (9.1.9)$$

则范数 $\|\cdot\|_{op}$ 是与向量范数 $\|\cdot\|$ 相容的矩阵范数。

称 $\|\cdot\|_{op}$ 为算子范数(或从属范数, 或 $\|\cdot\|$ 诱导的矩阵范数)。

证明 因 $\|Ax\|$ 为有界闭集 $S_{n-1} = \{x \in C^n \mid \|x\|=1\}$ 上的连续函数, 所以在 S_{n-1} 上达到最大值, 即存在 $x_0 \in S_{n-1}$ 使

$$\|Ax_0\| = \max_{\|x\|=1} \|Ax\| = \|A\|_{op}$$

(a) 非负性: 当 $A \neq 0$ 时, 存在 $x_0 \in S_{n-1}$ 使 $\|A\|_{op} = \|Ax_0\| = \max_{\|x\|=1} \|Ax\|$; 当 $A = 0$ 时,

$$\|A\|_{op} = \max_{\|x\|=1} \|0x\| = 0$$

(b) 齐次性: $\forall k \in C, \|kA\|_{op} = \max_{\|x\|=1} \|kAx\| = |k| \max_{\|x\|=1} \|Ax\| = |k| \|A\|_{op}$ 。

(c) 三角不等式: $\forall B \in C^{m \times n}$, 对矩阵 $A+B$, 必存在 $x_1 \in S_{n-1}$ 使

$$\|A+B\|_{op} = \|(A+B)x_1\| \leq \|Ax_1\| + \|Bx_1\|$$

所以,

$$\|A+B\|_{op} \leq \|A\|_{op} + \|B\|_{op}。$$

(d) 与向量范数 $\|\cdot\|$ 的相容性: $\forall y \in C^n$, 若 $y = 0$, 则必有 $\|Ay\| \leq \|A\|_{op} \cdot \|y\|$ 成立。若 $y \neq 0$, 令 $x_0 = y / \|y\|$, 则 $\|x_0\|=1$ 。所以,

$$\|Ax_0\| \leq \max_{\|x\|=1} \|Ax\| = \|A\|_{op} \|Ax_0\| = \|Ay\| / \|y\|$$

因此,

$$\|Ay\| \leq \|A\|_{op} \cdot \|y\|$$

(e) 矩阵与乘法的相容性: $\forall B \in C^{m \times l}$, 对于矩阵 AB 必存在 $x^* \in C^l, \|x^*\|=1$ 使得

$$\|AB\|_{op} = \|ABx^*\|$$

并且

$$\|ABx^*\| \leq \|A\|_{op} \cdot \|Bx^*\| \leq \|A\|_{op} \max_{\|x\|=1} \|Bx\| = \|A\|_{op} \cdot \|B\|_{op}$$

注 对于 $C^{n \times n}$ 上的任何一种算子范数, 总有 $\|I\|_{op} = \max_{\|x\|=1} \|Ix\| = 1$ 。但是, 对于一般的矩阵范数(与向量范数相容), 仅能保证 $\|I\| \geq 1$ 。因为 $\|x\| = \|Ix\| \leq \|I\| \cdot \|x\|$, 所以 $\|I\| \geq 1$ 。

定理 9.1.4 表明任何一种向量范数均可以诱导出一种与该向量范数相容的矩阵范数——算子范数, 但并非每一种算子范数都可以用矩阵的元素来解析表示。

下面给出由向量范数 $\|\cdot\|_1, \|\cdot\|_2, \|\cdot\|_\infty$ 所诱导的三种常用矩阵范数的具体表达式, 证明留给读者。

定理 9.1.5 设 $A \in C^{m \times n}, x \in C^n$, 则从属于向量的三种范数 $\|\cdot\|_1, \|\cdot\|_2, \|\cdot\|_\infty$ 矩阵范数依次

为

$$\|A\|_1 = \max_j \sum_{i=1}^n |a_{ij}| \quad (9.1.10)$$

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^H A)} \quad (9.1.11)$$

$$\|A\|_{\infty} = \max_i \sum_{j=1}^n |a_{ij}| \quad (9.1.12)$$

其中 $\lambda_{\max}(A^H A)$ 表示矩阵 $A^H A$ 的最大特征值。

范数 $\|A\|_1$, $\|A\|_2$, $\|A\|_{\infty}$ 分别称为列和范数, 谱范数, 行和范数。

谱范数的性质 谱范数具有非常好的性质, 因此在矩阵分析和系统理论中通常使用它作为分析工具。下面讨论谱范数的性质。

定理 9.1.6 设 $A \in C^{m \times n}$, 则

$$(1) \|A\|_2 = \max_{\|x\|_2=\|y\|_2=1} |y^H A x|, \quad x, y \in C^n$$

$$(2) \|A^H\|_2 = \|A\|_2$$

$$(3) \|A^H A\|_2 = \|A\|_2^2$$

证明 (1) $\forall \|x\|_2 = \|y\|_2 = 1$ 有 $|y^H A x| \leq \|y\|_2 \cdot \|A x\|_2 \leq \|A\|_2$ 。另一方面, 设 $\|x_0\|_2 = 1$ 使

$\|A x_0\|_2 = \|A\|_2 \neq 0$, 并令 $y_0 = \frac{A x_0}{\|A x_0\|_2}$, 则 $\|y_0\|_2 = 1$, 且 $|y_0^H A x| = \frac{\|A x_0\|_2^2}{\|A x_0\|_2} = \|A\|_2$ 。所以,

$$\|A\|_2 = \max_{\|x\|_2=\|y\|_2=1} |y^H A x|$$

$$(2) \|A\|_2 = \max_{\|x\|_2=\|y\|_2=1} |y^H A x| = \max_{\|x\|_2=\|y\|_2=1} |x^H A^H y| = \|A^H\|_2。$$

(3) 由 $\|A^H A\|_2 \leq \|A^H\|_2 \cdot \|A\|_2$, $\|A^H\|_2 = \|A\|_2$, 知 $\|A^H A\|_2 \leq \|A\|_2^2$ 。另一方面,

$$\|A^H A\|_2 \geq \max_{\|x\|_2=1} |x^H A^H A x| = \max_{\|x\|_2=1} \|A x\|_2^2 = \|A\|_2^2$$

所以,

$$\|A^H A\|_2 = \|A\|_2^2$$

证毕。

定理 9.1.7 设 $A \in C^{m \times n}$, 且 $Q \in U^{m \times m}$, $R \in U^{n \times n}$ 是酉矩阵, 则 $\|QAR\|_2 = \|A\|_2$ 。

证明 令 $v = R^H x, u = Q^H y$, 则 $\|x\|_2 = 1 \Leftrightarrow \|v\|_2 = 1$, $\|y\|_2 = 1 \Leftrightarrow \|u\|_2 = 1$ 。所以,

$$\|A\|_2 = \max_{\|x\|_2=\|y\|_2=1} |y^H A x| = \max_{\|x\|_2=\|y\|_2=1} |u^H QARv| = \|QAR\|_2$$

定义 9.1.6 设 $A \in C^{n \times n}$ 的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_n$ ，称

$$\rho(A) = \max_j |\lambda_j| \quad (9.1.13)$$

为矩阵 A 的谱半径。

谱半径是数值代数中经常使用的概念，它的几何意义是以原点为圆心，包含 A 的所有特征值的最小圆。

定理 9.1.8 对任意 $A \in C^{n \times n}$ ，有

$$\rho(A) \leq \|A\| \quad (9.1.14)$$

其中 $\|\cdot\|$ 为 A 的任何一种范数。

这个定理说明矩阵 A 的所有特征值均在以原点为圆心，以 A 的范数为半径的圆内，也就是说此定理给出矩阵特征值上界的一个估计。

证明 令 λ 为 A 的任何一个特征值， \mathbf{x} 为相应的特征向量，则有 $A\mathbf{x} = \lambda\mathbf{x}$ 。因此，

$$|\lambda| \cdot \|\mathbf{x}\| = \|\lambda\mathbf{x}\| \leq \|A\| \cdot \|\mathbf{x}\|$$

即 $|\lambda| \leq \|A\|$ ，所以 $\rho(A) \leq \|A\|$ 。证毕。

定理 9.1.9 设 $A \in C^{n \times n}$ 是正规矩阵 ($A^H A = A A^H$)，则

$$\rho(A) = \|A\|_2 \quad (9.1.15)$$

证明 因 A 是正规矩阵，所以存在酉矩阵 U 使 $U^H A U = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ 。于是，根据定理 9.1.7，有

$$\|A\|_2 = \|U^T A U\|_2 = \|\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)\|_2 = \sqrt{\max_j \bar{\lambda}_j \lambda_j} = \sqrt{\max_j |\lambda_j|^2} = \rho(A)$$

证毕。

定理 9.1.10 设 $A \in C^{n \times n}$ 是可逆矩阵，则 A 的谱范数为

$$\|A\|_2 = \sqrt{\rho(A^H A)} = \sqrt{\rho(A A^H)} \quad (9.1.16)$$

证明留给读者。

9.1.3 矩阵的条件数

定义 9.1.7 如果线性方程组 $A\mathbf{x} = \mathbf{b}$ 的系数矩阵 A 或常数项 \mathbf{b} 的微小变化引起方程组解的巨大变化，则称方程组为病态方程组，其系数矩阵 A 就解方程组(或求逆)而言是病态矩阵；反之，称方程组为良态方程组， A 称为良态矩阵。

值得指出的是“病态矩阵”的概念针对具体问题而言的，例如：对于解方程(或求逆)来说是病态的，但是对于求特征值而言未必是病态的，不能笼统地说某个矩阵是“病态”的。在本节所说

的“病态”都是针对解方程组而言的。我们希望能给出衡量矩阵是否病态的标准。为此，先给出一个引理。

引理 9.1.1 若 $\|A\| < 1$ ，则 $I \pm A$ 为非奇异矩阵，且

$$\|I \pm A\|^{-1} \leq (1 - \|A\|)^{-1} \quad (9.1.17)$$

其中 $\|\cdot\|$ 为矩阵的任何一种范数。

证明 由定理 9.1.8， $\rho(A) \leq \|A\| < 1$ ，因此 $\lambda = \pm 1$ 不可能为矩阵 A 的特征值，所以，

$$(I - A)(I - A)^{-1} = I$$

于是

$$(I - A)^{-1} = ((I - A) + A)(I - A)^{-1} = I + A(I - A)^{-1}$$

所以

$$\|(I - A)^{-1}\| \leq 1 + \|A\| \cdot \|(I - A)^{-1}\|$$

即

$$\|(I - A)^{-1}\| \leq (1 - \|A\|)^{-1}$$

同理可证明 $I + A$ 的情形。

设有方程组 $Ax = b$ ， $b \neq 0$ 。下面分析系数矩阵与常数项有一个微小变化时，方程组的解是如何变化的。这个问题通常称为线性方程组的摄动分析。

假定 A 是精确的， b 有误差 δb ，解为 $x + \delta x$ ，则

$$A(x + \delta x) = b + \delta b, \delta x = A^{-1}\delta b, \|\delta x\| = \|A^{-1}\| \cdot \|\delta b\|$$

另一方面，由 $Ax = b$ ，有 $\|b\| \leq \|A\| \cdot \|x\|$ ，所以，

$$\frac{1}{\|x\|} \leq \frac{\|A\|}{\|b\|}$$

因此，可得到下述命题：

命题 9.1.1 如果 A 为非奇异矩阵， $Ax = b \neq 0$ ，且 $A(x + \delta x) = b + \delta b$ ，则

$$\frac{\|\delta x\|}{\|x\|} \leq (\|A^{-1}\| \cdot \|A\|) \frac{\|\delta b\|}{\|b\|} \quad (9.1.18)$$

上式给出了解的相对误差的上界，常数项 b 的相对误差在解中可能放大 $\|A^{-1}\| \cdot \|A\|$ 倍。

下面假定 b 是精确的， A 有微小误差(摄动) δA ，解为 $x + \delta x$ ，则

$$(A + \delta A)(x + \delta x) = b, (A + \delta A)\delta x = -(\delta A)x \quad (9.1.19)$$

若 δA 不受限制的话， $A + \delta A$ 可能是非奇异的，因此不能直接求逆。但是，有

$$(A + \delta A) = A(I + A^{-1}\delta A)$$

所以, 由命题 9.1.1 时, 当 $\|A^{-1}\delta A\| < 1$ 时, $(I + A^{-1}\delta A)^{-1}$ 存在。于是, 从(9.1.19)式, 有

$$\delta \mathbf{x} = (I + A^{-1}\delta A)^{-1} A^{-1}(\delta A)\mathbf{x}$$

因此,

$$\|\delta \mathbf{x}\| \leq \frac{\|A^{-1}\| \cdot \|\delta A\| \cdot \|\mathbf{x}\|}{1 - \|A^{-1}\delta A\|}$$

如果矩阵 A 的微小变化 δA 满足

$$\|A^{-1}\| \cdot \|\delta A\| < 1 \quad (9.1.20)$$

则有

$$\frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \frac{(\|A^{-1}\| \cdot \|A\|) \cdot \frac{\|\delta A\|}{\|A\|}}{1 - (\|A^{-1}\| \cdot \|A\|) \cdot \frac{\|\delta A\|}{\|A\|}} \quad (9.1.21)$$

这个不等式说明了解的相对变化与系数矩阵的相对变化之间的关系。不难看出, 当条件(9.1.20)满足时, 式(9.1.21)表明, $(\|A^{-1}\| \cdot \|A\|)$ 越大, 解的相对变化也越大, $(\|A^{-1}\| \cdot \|A\|)$ 越小, 解的相对变化也越小。综合上述讨论, $(\|A^{-1}\| \cdot \|A\|)$ 在某种程度上刻划了方程组的解对于原始数据变化的敏感度, 也就是说刻划了方程组的“病态”程度。因此, 自然地引进下面的概念。

定义 9.1.8 设 A 为非奇异矩阵, 则称

$$Cond(A) = \|A^{-1}\| \cdot \|A\|$$

为矩阵 A 的条件数。

矩阵条件数, 刻划了方程组的相对误差可能的放大率。当 $Cond(A) \gg 1$ 时, 则方程组是“病态”的(即 A 是“病态”矩阵, 或者说 A 是坏条件的); 当 $Cond(A)$ 相对地小, 则方程组是“良态”的(或者说 A 是好条件的)。但是, 条件数究竟多大是矩阵才算是病态的, 一般来说没有具体的标准, 也就是说“病态”概念只是一种相对概念。

在数值分析中, 通常使用的条件数有:

$$(1) Cond(A)_{\infty} = \|A^{-1}\|_{\infty} \cdot \|A\|_{\infty};$$

$$(2) Cond(A)_2 = \|A^{-1}\|_2 \cdot \|A\|_2 = \sqrt{\frac{\lambda_{\max}(A^H A)}{\lambda_{\min}(A^H A)}} \quad (\text{谱条件数});$$

特别地, A 为实对称矩阵时, 谱条件数为 $Cond(A)_2 = |\lambda_1| / |\lambda_n|$, 其中, λ_1, λ_n 分别为模最大和最小的特征值。

$$(3) \text{Cond}(A)_F = \|A^{-1}\|_F \cdot \|A\|_F.$$

9.2 矩阵级数与矩阵函数

9.2.1 矩阵序列

矩阵序列的极限

为了叙述方便, 假定所考虑的矩阵是 n 阶方阵。本节中的所有概念与结果, 只要不涉及矩阵的乘法与逆矩阵时, 对任意矩阵都成立。如果仅涉及矩阵乘法时, 只要作相应的修改, 也有相应的结论。

在第 9.1.2 节中, 已经涉及了矩阵序列极限的概念, 在本节将集中讨论。

定义 9.2.1 设有矩阵序列 $\{A^{(k)}\}$, $A^{(k)} = (a_{ij}^{(k)}) \in C^{m \times n}$, 当 $k \rightarrow +\infty$ 时, $a_{ij}^{(k)} \rightarrow a_{ij}$, 则称 $\{A^{(k)}\}$ 收敛, 并将矩阵 $A = (a_{ij})$ 称为 $\{A^{(k)}\}$ 的极限, 或者说 $\{A^{(k)}\}$ 收敛到 A , 记为

$$\lim_{k \rightarrow +\infty} A^{(k)} = A \text{ (或 } A^{(k)} \rightarrow A)$$

不收敛的矩阵序列称为发散的。

当然, 也可按照范数定义矩阵的收敛性。即若

$$\lim_{k \rightarrow +\infty} \|A^{(k)} - A\| = 0$$

则称 $\{A^{(k)}\}$ 在范数 $\|\cdot\|$ 意义下收敛于 A 。然而, 这种定义与上面的定义是等价的。

定理 9.2.1 定义 9.2.1 与范数 $\|\cdot\|$ 意义下收敛性是等价的。

证明 取矩阵的 F 范数:

$$\|A\|_F = \left(\sum_{i,j=1}^n |a_{ij}|^2 \right)^{1/2}$$

若 $\|A^{(k)} - A\|_F \rightarrow 0$, 则有 $\sum_{i,j=1}^n |a_{ij}^{(k)} - a_{ij}|^2 \rightarrow 0$ 。因此, $\forall i, j, a_{ij}^{(k)} \rightarrow a_{ij}$ 。反之, 若 $\forall i, j, a_{ij}^{(k)} \rightarrow a_{ij}$,

则必有 $\sum_{i,j=1}^n |a_{ij}^{(k)} - a_{ij}|^2 \rightarrow 0$ 。因此 $\|A^{(k)} - A\|_F \rightarrow 0$ 。

设 $\|\cdot\|$ 为矩阵的任一种范数, 由矩阵范数的等价性, 存在正数 c_1, c_2 使

$$c_1 \|A^{(k)} - A\|_F \leq \|A^{(k)} - A\| \leq c_2 \|A^{(k)} - A\|_F$$

因此, 在 F 范数意义下的收敛性等价于在任何矩阵范数 $\|\cdot\|$ 意义下收敛性。所以定义 9.2.1 与任何范数意义下的收敛性是等价的。证毕。

对于矩阵序列的极限运算与普通数列的极限运算完全相同, 即

(1) 若 $\lim_{k \rightarrow +\infty} A^{(k)} = A$, $\lim_{k \rightarrow +\infty} B^{(k)} = B$, 则 $\forall a, b \in C$, 有

$$\lim_{k \rightarrow +\infty} (aA^{(k)} + bB^{(k)}) = aA + bB$$

(2) 若 $\lim_{k \rightarrow +\infty} A^{(k)} = A$, $\lim_{k \rightarrow +\infty} B^{(k)} = B$, 则

$$\lim_{k \rightarrow +\infty} A^{(k)} B^{(k)} = AB$$

(3) 设 $\{A^{(k)}\}$ 为每一项都是可逆的矩阵序列, 若 $\lim_{k \rightarrow +\infty} A^{(k)} = A$, 则 $\{(A^{(k)})^{-1}\}$ 也收敛, 并且

$$\lim_{k \rightarrow +\infty} (A^{(k)})^{-1} = A^{-1}$$

证明 (1),(2) 是明显的。下面证明(3): 记

$$\text{adj}(A^{(k)}) = \begin{pmatrix} A_{11}^{(k)} & A_{21}^{(k)} & \cdots & A_{n1}^{(k)} \\ \vdots & \vdots & \vdots & \vdots \\ A_{1n}^{(k)} & A_{2n}^{(k)} & \cdots & A_{nn}^{(k)} \end{pmatrix}$$

为 $A^{(k)}$ 的伴随矩阵。其中 $A_{ij}^{(k)}$ 是 $A^{(k)}$ 中元素 $a_{ji}^{(k)}$ 的代数余子式, 从而 $A_{ij}^{(k)}$ 是 $A^{(k)}$ 元素的 $n-1$ 次多项式。因此, 由 $\lim_{k \rightarrow +\infty} A^{(k)} = A$ 知

$$\lim_{k \rightarrow +\infty} A_{ij}^{(k)} = A_{ij}, \quad \lim_{k \rightarrow +\infty} \text{adj}(A^{(k)}) = \text{adj}(A)$$

而 $(A^{(k)})^{-1} = \frac{\text{adj}(A^{(k)})}{\det(A^{(k)})}$, 所以,

$$\lim_{k \rightarrow +\infty} (A^{(k)})^{-1} = \lim_{k \rightarrow +\infty} \frac{\text{adj}(A^{(k)})}{\det(A^{(k)})} = \frac{\text{adj}(A)}{\det(A)} = A^{-1}$$

证毕。

下面考虑矩阵 A 的序列

$$A^1, A^2, \dots, A^n \dots$$

的收敛性。即矩阵序列 $\{A^k\}$ 的收敛性(注意: 上标不加括号, 表示矩阵的幂)。

定理 9.2.2 $\lim_{k \rightarrow +\infty} A^k = 0$ 的充要条件是 $\rho(A) < 1$ (或者说 A 的所有特征值的模均小于 1)。

证明 令 A 的约当标准形为

$$J = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_s \end{pmatrix}, \quad J_r = \begin{pmatrix} \lambda_r & 1 & & \\ & \lambda_r & 1 & \\ & & \ddots & \ddots \\ & & & \ddots & 1 \\ & & & & \lambda_r \end{pmatrix}_{n_r \times n_r}$$

因此，存在可逆矩阵 B 使 $A = BJB^{-1}$ 。显然 $A^k = BJ^k B^{-1}$ ，所以，

$$\lim_{k \rightarrow +\infty} A^k = 0 \Leftrightarrow \lim_{k \rightarrow +\infty} J^k = 0$$

而

$$J_r^k = \begin{pmatrix} \lambda_r^k & c_k^1 \lambda_r^{k-1} & \cdots & \cdots & c_k^{n_r-1} \lambda_r^{k-n_r+1} \\ & \lambda_r^k & \ddots & & \vdots \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & c_k^1 \lambda_r^{k-1} \\ & & & & \lambda_r^k \end{pmatrix}_{n_r \times n_r}$$

所以， $\lim_{k \rightarrow +\infty} J^k = 0 \Leftrightarrow \forall r, \lim_{k \rightarrow +\infty} J_r^k = 0_{n_r \times n_r} \Leftrightarrow \forall r, \lim_{k \rightarrow +\infty} \lambda_r^k = 0 \Leftrightarrow \forall r, |\lambda_r| < 1 \Leftrightarrow \rho(A) < 1$ 。

推论 9.2.1 若 $\|A\| < 1$ ，则 $\lim_{k \rightarrow +\infty} A^k = 0$ 。

9.2.2 矩阵级数

矩阵级数与数学分析中的数项级数的定义类似，并且有类似的性质。

定义 9.2.2 设有矩阵序列 $\{A^{(k)}\}$ ，称无穷和 $\sum_{k=0}^{\infty} A^{(k)}$ 为矩阵级数， $A^{(k)}$ 称为矩阵级数的一般项。前 $k+1$ 项的和： $S^{(k)} = \sum_{j=0}^k A^{(j)}$ 称为矩阵级数的部分和，如果 $\{S^{(k)}\}$ 收敛，且有极限 S ，即 $S = \lim_{k \rightarrow \infty} S^{(k)}$ ，则称矩阵级数 $\sum_{k=0}^{\infty} A^{(k)}$ 收敛， S 称为级数的和，记作 $S = \sum_{k=0}^{\infty} A^{(k)}$ 。不收敛的矩阵级数称为发散的矩阵级数。

显然，矩阵级数收敛的充要条件为 $\forall i, j$ ，数项级数 $\sum_{k=0}^{\infty} a_{ij}^{(k)}$ 收敛，并且收敛时，它的和为 $S = (\sum_{k=0}^{\infty} a_{ij}^{(k)})_{n \times n}$ 。下面的性质也是明显的。

- (1) 若 $\sum_{k=0}^{\infty} A^{(k)}$ 收敛，则 $\lim_{k \rightarrow +\infty} A^k = 0$ ；
- (2) 若 $\sum_{k=0}^{\infty} A^{(k)} = A$ ， $\sum_{k=1}^{\infty} B^{(k)} = B$ ，则 $\forall a, b \in C$ ，有 $\sum_{k=1}^{\infty} (aA^{(k)} + bB^{(k)}) = aA + bB$ 。

定义 9.2.3 设矩阵级数 $\sum_{k=0}^{\infty} A^{(k)}$ ， $\forall i, j$ ， $\sum_{k=0}^{\infty} a_{ij}^{(k)}$ 为绝对收敛的数项级数，则称 $\sum_{k=0}^{\infty} A^{(k)}$ 绝对收敛。

定理 9.2.3 矩阵级数 $\sum_{k=0}^{\infty} A^{(k)}$ 绝对收敛的充要条件为数项级数 $\sum_{k=0}^{\infty} \|A^{(k)}\|$ 收敛，其中 $\|\cdot\|$ 为任何矩阵范数。

证明 若 $\sum_{k=0}^{\infty} A^{(k)}$ 绝对收敛，则存在正数 M 使 $\forall i, j$ 有 $\sum_{k=0}^{+\infty} |a_{ij}^{(k)}| \leq M$ 。从而有

$$\sum_{k=0}^{+\infty} \|A^{(k)}\| = \sum_{k=0}^{+\infty} \left(\sum_{i=1}^n \sum_{j=1}^n |a_{ij}^{(k)}| \right) \leq n^2 M$$

因此， $\sum_{k=0}^{\infty} \|A^{(k)}\|$ 收敛。反之，必有 $\sum_{k=0}^{+\infty} |a_{ij}^{(k)}| \leq \sum_{k=0}^{+\infty} \|A^{(k)}\|, \forall i, j$ 。即 $\forall i, j$ ， $\sum_{k=0}^{+\infty} |a_{ij}^{(k)}|$ 收敛，因此 $\sum_{k=0}^{\infty} A^{(k)}$ 绝对收敛。证毕。

对于绝对收敛的矩阵级数的乘积，有

定理 9.2.4 若 $\sum_{k=0}^{\infty} A^{(k)}$ ， $\sum_{k=0}^{\infty} B^{(k)}$ 分别绝对收敛于 A, B ，则它们的乘积 $\sum_{k=0}^{\infty} C^{(k)}$ 绝对收敛于 AB 。

其中， $C^{(k)} = \sum_{j=1}^k A^{(j)} B^{(k-j+1)}$ 。

证明 因为 $\sum_{k=1}^{\infty} \|C^{(k)}\| = \sum_{k=1}^{\infty} \left\| \sum_{j=1}^k A^{(j)} B^{(k-j+1)} \right\| \leq \sum_{k=1}^{\infty} \sum_{j=1}^k \|A^{(j)}\| \cdot \|B^{(k-j+1)}\| = \sum_{k=1}^{\infty} \|A^{(j)}\| \cdot \sum_{k=1}^{\infty} \|B^{(k)}\|$ ，所以

$\sum_{k=1}^{\infty} C^{(k)}$ 绝对收敛。为了计算 $\sum_{k=1}^{\infty} C^{(k)}$ 的和，我们可以取它的特殊排法：

$$A^{(1)}B^{(1)} + (A^{(1)}B^{(2)} + A^{(2)}B^{(1)}) + \cdots + \left(\sum_{i=1}^k A^{(i)} \sum_{j=1}^k B^{(j)} - \sum_{i=1}^{k-1} A^{(i)} \sum_{j=1}^{k-1} B^{(j)} \right) + \cdots$$

记 $\sum_{k=0}^{\infty} A^{(k)}$ ， $\sum_{k=0}^{\infty} B^{(k)}$ 的部分和分别为 $S_A^{(n)}$ ， $S_B^{(n)}$ ，则按上述排法规则 $\sum_{k=1}^{\infty} C^{(k)}$ 的部分和为

$S_C^{(n)} = S_A^{(n)} \cdot S_B^{(n)}$ ，因此， $\lim_{n \rightarrow \infty} S_C^{(n)} = \lim_{n \rightarrow \infty} S_A^{(n)} \cdot \lim_{n \rightarrow \infty} S_B^{(n)} = AB$ ，故 $\sum_{k=1}^{\infty} C^{(k)} = AB$ 。证毕。

9.2.3 矩阵函数

矩阵的幂级数

下面建立矩阵的幂级数理论。先从类似于数的几何级数的简单矩阵级数开始：

定理 9.2.5 矩阵 A 的级数：

$$\sum_{k=0}^{\infty} A^k = I + A + A^2 + \cdots + A^k + \cdots \tag{9.2.1}$$

收敛的充要条件为 $\rho(A) < 1$, 在收敛时, 有 $\sum_{k=0}^{\infty} A^k = (I - A)^{-1}$ 。

证明 级数(9.2.1)收敛 $\Rightarrow \forall i, j, \sum_{k=0}^{\infty} a_{ij}^{(k)}$ 收敛 $\Rightarrow \forall i, j, a_{ij}^{(k)} \rightarrow 0 \Rightarrow \rho(A) < 1$ 。反之, 若 $\rho(A) < 1$,

则 $I - A$ 是非奇异矩阵, 从而 $(I - A)^{-1}$ 存在。又因

$$S^{(k)}(I - A) = (I + A + \cdots + A^k)(I - A) = I - A^{k+1}$$

所以,

$$S^{(k)} = I + A + \cdots + A^k = (I - A)^{-1} - A^{k+1}(I - A)^{-1}$$

因此,

$$\lim_{k \rightarrow \infty} S^{(k)} = (I - A)^{-1} - \lim_{k \rightarrow \infty} A^{k+1}(I - A)^{-1} = (I - A)^{-1}$$

定理 9.2.6 若矩阵 A 对某种矩阵范数有 $\|A\| < 1$, 则

$$\left\| (I - A)^{-1} - \sum_{j=1}^k A^j \right\| \leq \frac{1}{1 - \|A\|} \|A\|^{k+1} \quad (9.2.2)$$

证明 由定理 9.2.5, $(I - A)^{-1} - \sum_{j=1}^k A^j = \sum_{j=k+1}^{\infty} A^j$ 。又因

$$\left\| \sum_{j=k+1}^{k+l} A^j \right\| \leq \sum_{j=k+1}^{k+l} \|A\|^j = \frac{\|A\|^{k+1} (1 - \|A\|^l)}{1 - \|A\|}$$

所以,

$$\left\| \sum_{j=k+1}^{\infty} A^j \right\| = \lim_{l \rightarrow \infty} \left\| \sum_{j=k+1}^{k+l} A^j \right\| \leq \frac{1}{1 - \|A\|} \|A\|^{k+1}$$

证毕。

上面给出了一种特别的矩阵幂级数, 并讨论了它的收敛性。下面给出矩阵幂级数的一般定义:

定义 9.2.4 设 $A = (a_{ij}) \in C^{n \times n}$, 称矩阵级数:

$$\sum_{k=0}^{\infty} c_k A^k = c_0 I + c_1 A + \cdots + c_k A^k + \cdots \quad (9.2.3)$$

为矩阵幂级数。

定理 9.2.7 设级数 $\sum_{k=0}^{\infty} c_k Z^k$ 的收敛半径为 r , A 为 n 阶方阵, 若 $\rho(A) < r$, 则矩阵幂级数 $\sum_{k=0}^{\infty} c_k A^k$

绝对收敛; 若 $\rho(A) > r$, 则级数 $\sum_{k=0}^{\infty} c_k A^k$ 发散。

证明 令 J 是 A 的 Jordan 标准形, 则存在可逆矩阵 P 使得

$$A = PJP^{-1} = P \text{diag}(J_1(\lambda_1), J_2(\lambda_2), \cdots, J_s(\lambda_s)) P^{-1},$$

$$J_r = \begin{pmatrix} \lambda_r & 1 & & \\ & \lambda_r & 1 & \\ & & \ddots & \ddots \\ & & & \ddots & 1 \\ & & & & \lambda_r \end{pmatrix}_{n_r \times n_r}, \quad \sum_{r=1}^s n_r = n$$

于是,

$$A^k = P \text{diag}(J_1^k, J_2^k, \dots, J_s^k) P^{-1}, \quad J_r^k = \begin{pmatrix} \lambda_r^k & c_k^1 \lambda_r^{k-1} & \dots & c_k^{n_r-1} \lambda_r^{k-n_r+1} \\ & \lambda_r^k & \ddots & \vdots \\ & & \ddots & c_k^1 \lambda_r^{k-1} \\ & & & \lambda_r^k \end{pmatrix}_{n_r \times n_r}$$

所以,

$$\sum_{k=0}^{\infty} c_k A^k = \sum_{k=0}^{\infty} c_k (P J^k P^{-1}) = P \left(\sum_{k=0}^{\infty} c_k J_r^k \right) P^{-1} = P \text{diag} \left(\sum_{k=0}^{\infty} c_k J_1^k, \sum_{k=0}^{\infty} c_k J_2^k, \dots, \sum_{k=0}^{\infty} c_k J_r^k \right) P^{-1}$$

其中

$$\sum_{k=0}^{\infty} c_k J_r^k = \begin{pmatrix} \sum_{k=0}^{\infty} c_k \lambda_r^k & \sum_{k=1}^{\infty} c_k C_k^1 \lambda_r^{k-1} & \dots & \sum_{k=n_r-1}^{\infty} c_k C_k^{n_r-1} \lambda_r^{k-n_r+1} \\ & \ddots & \ddots & \vdots \\ & & \ddots & \sum_{k=1}^{\infty} c_k C_k^1 \lambda_r^{k-1} \\ & & & \sum_{k=0}^{\infty} c_k \lambda_r^k \end{pmatrix}_{n_r \times n_r}$$

$$\begin{cases} C_k^l = \frac{k(k-1)\cdots(k-l+1)}{l!}, & k \geq l \\ C_k^l = 0, & k < l \end{cases}$$

当 $\rho(A) < r$ 时, 幂级数 $\sum_{k=0}^{\infty} c_k \lambda_r^k$, $\sum_{k=0}^{\infty} c_k C_k^1 \lambda_r^{k-1}$, ..., $\sum_{k=0}^{\infty} c_k C_k^{n_r-1} \lambda_r^{k-n_r+1}$ 都是绝对收敛的, 因此矩阵的幂级数 $\sum_{k=0}^{\infty} c_k A^k$ 绝对收敛。

当 $\rho(A) > r$ 时, 幂级数 $\sum_{k=0}^{\infty} c_k \lambda_r^k$ 发散, 所以矩阵的幂级数 $\sum_{k=0}^{\infty} c_k A^k$ 发散。

矩阵函数的幂级数表示

设函数 $f(x)$ 在 $|x| < r$ 内可以表示为幂级数:

$$f(x) = \sum_{k=0}^{\infty} c_k x^k \quad (|x| < r) \quad (9.2.4)$$

若矩阵 A 满足 $\rho(A) < r$, 由定理 9.2.7, 矩阵幂级数 $\sum_{k=0}^{\infty} c_k A^k$ 收敛。因此, 可定义矩阵函数的幂级数表示:

$$f(A) = \sum_{k=0}^{\infty} c_k A^k \quad (9.2.5)$$

矩阵函数的 **Jordan** 表示

令 A 的 Jordan 标准形为 $J = \text{diag}(J_1(x_1), \dots, J_s(x_s))$, 变换矩阵为 P , 即 $A = PJP^{-1}$, 则

$$\sum_{k=0}^{\infty} c_k A^k = P \text{diag} \left(\sum_{k=0}^{\infty} c_k J_1^k(x_1), \dots, \sum_{k=0}^{\infty} c_k J_s^k(x_s) \right) P^{-1} \quad (9.2.6)$$

其中

$$\sum_{k=0}^{\infty} c_k J_r^k(x_r) = \begin{pmatrix} \sum c_k x_r^k & \sum c_k C_k^1 x_r^{k-1} & \cdots & \sum c_k C_k^{n_r-1} x_r^{k-n_r+1} \\ & \sum c_k x_r^k & \ddots & \vdots \\ & & \ddots & \sum c_k C_k^1 x_r^{k-1} \\ & & & \sum c_k x_r^k \end{pmatrix} \quad (9.2.7)$$

不难计算:

$$f(x_r) = \sum_{k=0}^{\infty} c_k x_r^k, \quad f'(x_r) = \sum_{k=0}^{\infty} c_k C_k^1 x_r^{k-1}, \quad \frac{1}{2!} f''(x_r) = \sum_{k=0}^{\infty} c_k C_k^2 x_r^{k-2}, \quad \dots, \\ \frac{1}{(n_r-1)!} f^{(n_r-1)}(x_r) = \sum_{k=0}^{\infty} c_k C_k^{n_r-1} x_r^{k-n_r+1}, \quad \dots$$

所以,

$$\sum_{k=0}^{\infty} c_k J_r^k(x_r) = \begin{pmatrix} f(x_r) & f'(x_r) & \cdots & \frac{1}{(n_r-1)!} f^{(n_r-1)}(x_r) \\ & f(x_r) & \ddots & \vdots \\ & & \ddots & f'(x_r) \\ & & & f(x_r) \end{pmatrix} = f(J_r)$$

将上式代入(9.2.6), 有

$$f(A) = \sum_{k=0}^{\infty} c_k A^k = P \text{diag} (f(J_1), f(J_2), \dots, f(J_s)) P^{-1} \quad (9.2.8)$$

式(9.2.8)称为矩阵函数的 **Jordan** 表示。

9.3 矩阵导数

9.3.1 函数矩阵的导数

定义 9.3.1 若矩阵 A 的元素 a_{ij} 均为实变量 t 的函数, 即 $A(t) = (a_{ij}(t))_{m \times n}$, 则称 $A(t)$ 为函数矩阵。

对于函数矩阵, 可以与普通函数那样, 引入极限、连续、导数等概念。

如果对所有元素 $a_{ij}(t)$ 在 t_0 点存在极限 a_{ij} , 即 $\lim_{t \rightarrow t_0} a_{ij}(t) = a_{ij}$, 则称矩阵 $A(t)$ 在 t_0 点有极限, 且以矩阵 $A = (a_{ij})$ 为极限值, 即 $\lim_{t \rightarrow t_0} A(t) = (\lim_{t \rightarrow t_0} a_{ij}(t))_{m \times n} = (a_{ij})$; 与普通函数类似, 若 $A(t)$, $B(t)$ 在 t_0 点有极限 A, B , 则

$$\lim_{t \rightarrow t_0} (aA(t) + bB(t)) = aA + bB, \quad a, b \text{ 为常数}$$

$$\lim_{t \rightarrow t_0} (A(t)B(t)) = AB$$

类似地, 可引进 $A(t)$ 的连续性概念。

定义 9.3.2 设 $A(t) = (a_{ij}(t))_{m \times n}$, 若 $\forall i, j$, $a_{ij}(t)$ 在 $t = t_0$ 处(或在闭区间 $[a, b]$ 上)可导, 则称 $A(t)$ 在 t_0 处(或在 $[a, b]$ 上)可导, 且记为

$$A'(t_0) = \left. \frac{dA(t)}{dt} \right|_{t_0} = (a'_{ij}(t_0))_{m \times n}$$

因函数矩阵导数本身还是一个函数矩阵, 因此我们可以定义高阶导数:

$$\frac{d^k A(t)}{dt^k} = \frac{d}{dt} \left(\frac{d^{k-1} A(t)}{dt^{k-1}} \right), \quad (k = 1, 2, \dots, n)$$

对于函数矩阵导数, 有下述性质:

(1) $A(t)$ 为常数矩阵 $\Leftrightarrow A'(t) = 0$;

(2) 若 $A(t)$, $B(t)$ 可导, 则 $\frac{d}{dt}(A(t) \pm B(t)) = \frac{dA(t)}{dt} \pm \frac{dB(t)}{dt}$;

(3) 若 $\varphi(t)$ 是可导的实值函数, $A(t)$ 可导, 则 $\frac{d}{dt}(\varphi(t)A(t)) = \frac{d\varphi(t)}{dt} \cdot A(t) + \frac{dA(t)}{dt} \varphi(t)$;

(4) 若 $A(t)$, $B(t)$ 可导, 且可乘, 则 $\frac{d}{dt}(A(t)B(t)) = \frac{dA(t)}{dt} \cdot B(t) + A(t) \cdot \frac{dB(t)}{dt}$;

(5) 若 $A(t)$, $A^{-1}(t)$ 均可导, 则 $\frac{dA^{-1}(t)}{dt} = -A^{-1}(t) \cdot \frac{dA(t)}{dt} \cdot A^{-1}(t)$;

(6) 设 $A(t)$ 关于变量 t 可导, 而 $t = \varphi(x)$ 是变量 x 的可导函数, 则

$$\frac{dA(t)}{dt} = \frac{dA(t)}{dt} \cdot \frac{d\varphi(x)}{x} = \frac{d\varphi(x)}{dx} \cdot \frac{dA(t)}{dt}$$

这里只证明(5), 其它各条的证明是容易的。

证明 因 $A(t) \cdot A^{-1}(t) = I$, 所以 $\frac{d[A(t) \cdot A^{-1}(t)]}{dt} = 0$ 。另一方面,

$$\frac{d[A(t) \cdot A^{-1}(t)]}{dt} = \frac{dA(t)}{dt} \cdot A^{-1}(t) + A(t) \frac{dA^{-1}(t)}{dt}$$

所以,

$$\frac{dA^{-1}(t)}{dt} = -A^{-1}(t) \cdot \frac{dA(t)}{dt} \cdot A^{-1}(t)$$

例 4.5.1 设 $A(t) = \begin{pmatrix} 1 & t^2 \\ t & 0 \end{pmatrix}$, 求 $\frac{d^2 A}{dt^2}$, $\frac{dA^{-1}}{dt}$ 。

显然,

$$\frac{dA}{dt} = \begin{pmatrix} 0 & 2t \\ 1 & 0 \end{pmatrix}, \quad \frac{d^2 A}{dt^2} = \frac{d}{dt} \left(\frac{dA}{dt} \right) = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$$

为了计算 $\frac{dA^{-1}}{dt}$, 先计算:

$$A^{-1} = \begin{pmatrix} 0 & \frac{1}{t} \\ \frac{1}{t^2} & -\frac{1}{t^3} \end{pmatrix}$$

所以,

$$\frac{dA^{-1}}{dt} = - \begin{pmatrix} 0 & \frac{1}{t} \\ \frac{1}{t^2} & -\frac{1}{t^3} \end{pmatrix} \begin{pmatrix} 0 & 2t \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{t} \\ \frac{1}{t^2} & -\frac{1}{t^3} \end{pmatrix} = \frac{1}{t^4} \begin{pmatrix} 0 & -t^2 \\ -2t & 1 \end{pmatrix}$$

例 9.3.2 证明 $\frac{d}{dt}(\text{tr}A(t)) = \text{tr} \frac{dA(t)}{dt}$ 。

证明: $\frac{d(\text{tr}A(t))}{dt} = \text{tr} \left[\left(\frac{da_{ij}(t)}{dt} \right) \right]$

$$= \frac{da_{11}(t)}{dt} + \cdots + \frac{da_{mm}(t)}{dt} = \frac{d}{dt} (a_{11}(t) + \cdots + a_{mm}(t)) = \text{tr} \frac{dA(t)}{dt}$$

9.3.2 向量映射关于向量的导数

函数关于向量的导数

定义 9.3.3 设 $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$, $g = g(x_1, x_2, \dots, x_n)$, 定义:

$$\frac{dg}{d\mathbf{x}} = \left(\frac{\partial g}{\partial x_1} \quad \dots \quad \frac{\partial g}{\partial x_n} \right)^T; \quad \frac{dg}{d\mathbf{x}^T} = \left(\frac{\partial g}{\partial x_1} \quad \dots \quad \frac{\partial g}{\partial x_n} \right) \quad (9.3.1)$$

通常称上述两个式子分别为函数 g 关于列向量 \mathbf{x} 与行向量 \mathbf{x}^T 的梯度。

例 9.3.3 数量函数 $f(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$ 关于向量 \mathbf{x} 的导数为:

$$\frac{df}{d\mathbf{x}} = (A + A^T) \mathbf{x} \quad (9.3.2)$$

证明 因为 $f(\mathbf{x}) = \mathbf{x}^T A \mathbf{x} = \sum_{i,j=1}^n a_{ij} x_i x_j$, 所以,

$$\begin{aligned} \frac{df}{d\mathbf{x}} &= \left(\frac{\partial}{\partial x_1} \sum_{i,j=1}^n a_{ij} x_i x_j, \frac{\partial}{\partial x_2} \sum_{i,j=1}^n a_{ij} x_i x_j, \dots, \frac{\partial}{\partial x_n} \sum_{i,j=1}^n a_{ij} x_i x_j \right)^T \\ &= \begin{pmatrix} \sum_{j=1}^n a_{1j} x_j \\ \vdots \\ \sum_{j=1}^n a_{nj} x_j \end{pmatrix} + \begin{pmatrix} \sum_{j=1}^n a_{j1} x_j \\ \vdots \\ \sum_{j=1}^n a_{jn} x_j \end{pmatrix} = A \mathbf{x} + A^T \mathbf{x} = (A + A^T) \mathbf{x} \end{aligned}$$

例 9.3.4 令 $\mathbf{x} = (x_1(t), x_2(t), \dots, x_n(t))^T$, $f(\mathbf{x}) = f(x_1, x_2, \dots, x_n)$, 则

$$\frac{df}{dt} = \frac{df}{d\mathbf{x}^T} \frac{d\mathbf{x}}{dt} = \left(\frac{df}{d\mathbf{x}} \right)^T \frac{d\mathbf{x}}{dt} \quad (9.3.3)$$

证明
$$\frac{df}{dt} = \frac{\partial f}{\partial x_1} \frac{dx_1}{dt} + \dots + \frac{\partial f}{\partial x_n} \frac{dx_n}{dt} = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right) \begin{pmatrix} \frac{dx_1}{dt} \\ \vdots \\ \frac{dx_n}{dt} \end{pmatrix} = \frac{df}{d\mathbf{x}^T} \frac{d\mathbf{x}}{dt} = \left(\frac{df}{d\mathbf{x}} \right)^T \frac{d\mathbf{x}}{dt}.$$

对于 n 维向量的实值函数有下述求导法则:

(1) 线性法则:
$$\frac{d(af(\mathbf{x}) + bg(\mathbf{x}))}{d\mathbf{x}} = a \frac{df(\mathbf{x})}{d\mathbf{x}} + b \frac{dg(\mathbf{x})}{d\mathbf{x}};$$

$$(2) \text{ 乘积法则: } \frac{d(f(\mathbf{x})g(\mathbf{x}))}{d\mathbf{x}} = g(\mathbf{x})\frac{df(\mathbf{x})}{d\mathbf{x}} + f(\mathbf{x})\frac{dg(\mathbf{x})}{d\mathbf{x}};$$

$$(3) \text{ 商法则: } \frac{d(f(\mathbf{x})/g(\mathbf{x}))}{d\mathbf{x}} = \frac{1}{g^2(\mathbf{x})}\left(g(\mathbf{x})\frac{df(\mathbf{x})}{d\mathbf{x}} - f(\mathbf{x})\frac{dg(\mathbf{x})}{d\mathbf{x}}\right), (g(\mathbf{x}) \neq 0);$$

$$(4) \text{ 链式法则: } \frac{d(f(g(\mathbf{x})))}{d\mathbf{x}} = \frac{dg^T(\mathbf{x})}{d\mathbf{x}} \cdot \frac{df(g)}{dg}.$$

其中: $\frac{dg^T(\mathbf{x})}{d\mathbf{x}} = \left(\frac{\partial g_1(\mathbf{x})}{\partial x}, \dots, \frac{\partial g_n(\mathbf{x})}{\partial x}\right)$ 是 n 阶方阵。

向量映射关于向量的导数

定义 9.3.4 设 $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$, $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_m(\mathbf{x}))$ 为 m 维行向量值函数, 则定义 $\mathbf{f}(\mathbf{x})$ 对于 \mathbf{x} 的导数为:

$$\frac{d\mathbf{f}(\mathbf{x})}{d\mathbf{x}} = \begin{pmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \frac{\partial f_2(\mathbf{x})}{\partial x_1} & \dots & \frac{\partial f_m(\mathbf{x})}{\partial x_1} \\ \frac{\partial f_1(\mathbf{x})}{\partial x_2} & \frac{\partial f_2(\mathbf{x})}{\partial x_2} & \dots & \frac{\partial f_m(\mathbf{x})}{\partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_1(\mathbf{x})}{\partial x_n} & \frac{\partial f_2(\mathbf{x})}{\partial x_n} & \dots & \frac{\partial f_m(\mathbf{x})}{\partial x_n} \end{pmatrix} \quad (9.3.4)$$

并称它为 m 维行向量值函数 $\mathbf{f}(\mathbf{x})$ 在点 \mathbf{x} 处的梯度矩阵。

当 $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_m(\mathbf{x}))^T$ 为 m 维列向量值函数时, 定义 $\mathbf{f}(\mathbf{x})$ 对于 \mathbf{x}^T 的导数为

$$\frac{d\mathbf{f}(\mathbf{x})}{d\mathbf{x}^T} = \begin{pmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \frac{\partial f_1(\mathbf{x})}{\partial x_2} & \dots & \frac{\partial f_1(\mathbf{x})}{\partial x_n} \\ \frac{\partial f_2(\mathbf{x})}{\partial x_1} & \frac{\partial f_2(\mathbf{x})}{\partial x_2} & \dots & \frac{\partial f_2(\mathbf{x})}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m(\mathbf{x})}{\partial x_1} & \frac{\partial f_m(\mathbf{x})}{\partial x_2} & \dots & \frac{\partial f_m(\mathbf{x})}{\partial x_n} \end{pmatrix} \quad (9.3.5)$$

它是 m 维列向量值函数 $\mathbf{f}(\mathbf{x})$ 在点 \mathbf{x}^T 处的梯度矩阵, 通常也称它为 $\mathbf{f}(\mathbf{x})$ 的 Jacobi 矩阵。

显然, 有

$$\frac{d\mathbf{x}^T}{d\mathbf{x}} = \frac{d\mathbf{x}}{d\mathbf{x}^T} = I_n \quad (9.3.6)$$

这个结果虽然简单, 但它却是非常有用的。

下面给出一些常用的求导公式。

(1) 向量 c 与向量 x 无关, 则

$$\frac{dx^T c}{dx} = \frac{dc^T x}{dx} = c \quad (9.3.7)$$

$$\frac{dc^T f(x)}{dx} = \frac{df^T(x)c}{dx} = \frac{df^T(x)}{dx} c \quad (9.3.8)$$

(2) A, y 均与向量 x 无关, 则

$$\frac{dx^T Ay}{dx} = \frac{dx^T}{dx} Ay = Ay \quad (9.3.9)$$

$$\frac{dy^T Ax}{dx} = \frac{dx^T A^T y}{dx} = A^T y \quad (9.3.10)$$

(3) A 与向量 x 无关, 则

$$\frac{dx^T A}{dx} = A \quad (9.3.11)$$

$$\frac{dx^T Ax}{dx} = (A + A^T)x \quad (9.3.12)$$

$$\frac{dx^T Ax}{dx} = 2Ax, A \text{ 为对称矩阵} \quad (9.3.13)$$

(4) A 与向量 x 无关, $f(x), g(x)$ 与 x 有关, 则

$$\frac{d(f(x))^T Af(x)}{dx} = \frac{d(f(x))^T}{dx} (A + A^T)f(x) \quad (9.3.14)$$

$$\frac{d(f(x))^T Ag(x)}{dx} = \frac{d(f(x))^T}{dx} Ag(x) + \frac{d(g(x))^T}{dx} A^T f(x) \quad (9.3.15)$$

9.3.3 函数关于矩阵的导数

定义 9.3.3 设 $A \in R^{m \times n}$, $f(A)$ 为矩阵 A 的数量函数, 即 f 是 $m \times n$ 元函数, 则定义 $f(A)$ 关于矩阵 A 的导数为:

$$\frac{df}{dA} = \left(\frac{\partial f}{\partial a_{ij}} \right)_{m \times n} = \begin{pmatrix} \frac{\partial f}{\partial a_{11}} & \cdots & \frac{\partial f}{\partial a_{1n}} \\ \vdots & & \vdots \\ \frac{\partial f}{\partial a_{m1}} & \cdots & \frac{\partial f}{\partial a_{mn}} \end{pmatrix} \quad (9.3.16)$$

例 9.3.5 求 $f(A) = x^T Ax$ 关于矩阵 A 的导数, 其中 $A \in R^{n \times n}$ 是实对称矩阵。

解: $\frac{d}{dA}(x^T Ax) = \left(\frac{\partial}{\partial a_{ij}} \sum_{i,j=1}^n a_{ij} x_i x_j \right)_{n \times n} = (x_i x_j)_{n \times n} = xx^T$ 。

对于矩阵变元的实值函数有下述求导法则：

$$(1) \text{ 线性法则: } \frac{d(af(A) + bg(A))}{dA} = a \frac{df(A)}{dA} + b \frac{dg(A)}{dA};$$

$$(2) \text{ 乘积法则: } \frac{d(f(A)g(A))}{dA} = g(A) \frac{df(A)}{dA} + f(A) \frac{dg(A)}{dA};$$

$$(3) \text{ 商法则: } \frac{d(f(A)/g(A))}{dA} = \frac{1}{g^2(A)} \left(g(A) \frac{df(A)}{dA} - f(A) \frac{dg(A)}{dA} \right), (g(A) \neq 0);$$

$$(4) \text{ 链式法则: } \frac{d(g(f(A)))}{dA} = \frac{dg(y)}{dy} \cdot \frac{df(A)}{dA}, \text{ 其中 } g \text{ 是单变量的实值函数, } y = f(A).$$

行列式的导数

设 A 为 n 阶方阵，下面考虑 A 的行列式的导数 $\frac{d}{dA} \det(A)$ 。令 A_{ij} 表示元素 a_{ij} 的代数余子式，则有

$$\det(A) = \sum_{i=1}^n a_{ij} A_{ij}, \quad j = 1, 2, \dots, n$$

由于 A_{ij} 不含有元素 a_{ij} ，所以，

$$\frac{\partial}{\partial a_{ij}} \det(A) = A_{ij}$$

于是，

$$\frac{d}{dA} \det(A) = (\text{adj}A)^T \quad (9.3.17)$$

其中 $\text{adj}A = (A_{ij})$ 是 A 的伴随矩阵。

(1) 如果 A 可逆，则 $A^{-1} = \frac{1}{\det(A)} \text{adj}(A)$ ，这样就得到

$$\frac{d}{dA} \det(A) = \det(A) A^{-T} = \det(A^T) A^{-T} \quad (9.3.18)$$

同时，有

$$\frac{d}{dA} \det(A^{-1}) = \frac{d}{dA} (1/\det(A)) = \frac{1}{(\det(A))^2} \left(-\frac{d}{dA} \det(A) \right) = -\frac{1}{\det(A)} A^{-T}$$

因此，

$$\frac{d}{dA} \det(A^{-1}) = -\frac{1}{\det(A)} A^{-T} \quad (9.3.19)$$

(2) 如果 A 为正定对称矩阵，则

$$\frac{\partial}{\partial a_{ij}} \det(A) = \begin{cases} A_{ij}, & i = j \\ 2A_{ij}, & i \neq j \end{cases}$$

所以,

$$\frac{d}{dA} \det(A) = \det(A)(2A^{-1} - \text{diag}(A^{-1})) \quad (9.3.20)$$

(3) 对非奇异矩阵 A , 根据链式法则, 有

$$\frac{d}{dA} \log(\det(A)) = \frac{1}{\det(A)} \cdot \frac{d}{dA} \det(A) \quad (9.3.21)$$

如果 A 又为对称矩阵, 则有

$$\frac{d}{dA} \log(\det(A)) = 2A^{-1} - \text{diag}(A^{-1}) \quad (9.3.22)$$

下面, 不加证明的引进矩阵乘积行列式的导数:

$$\frac{d}{dA} \det(AA^T) = 2 \det(AA^T)(AA^T)^{-1} A, \text{rank}(A_{m \times n}) = m \quad (9.3.23)$$

$$\frac{d}{dA} \det(A^T A) = 2 \det(A^T A)A(A^T A)^{-1}, \text{rank}(A_{m \times n}) = n \quad (9.3.24)$$

$$\frac{d}{dA} \det(A^2) = 2 \det^2(A)A^{-T}, \text{rank}(A_{n \times n}) = n \quad (9.3.25)$$

迹的导数

设 $A \in R^{n \times m}$, $B \in R^{m \times n}$, 下面考虑 $\frac{d}{dA} \text{tr}(AB)$ 。

由于 $\text{tr}(AB) = \sum_{i=1}^n \sum_{j=1}^m a_{ij} b_{ji}$, 所以,

$$\frac{\partial}{\partial a_{ij}} \text{tr}(AB) = b_{ji}$$

于是,

$$\frac{d}{dA} \text{tr}(AB) = B^T \quad (9.3.26)$$

特别地,

$$\frac{d}{dA} \text{tr}(BA) = \frac{d}{dA} \text{tr}(AB) = B^T \quad (9.3.27)$$

$$\frac{d}{dx} \text{tr}(\mathbf{x}\mathbf{y}^T) = \frac{d}{dx} \text{tr}(\mathbf{y}\mathbf{x}^T) = \mathbf{y}, \quad \mathbf{x}, \mathbf{y} \in R^n \quad (9.3.28)$$

$$\frac{d}{dA} (\mathbf{x}^T A \mathbf{x}) = \frac{d \text{tr}(\mathbf{x}^T A \mathbf{x})}{dA} = \frac{d \text{tr}(A \mathbf{x} \mathbf{x}^T)}{dA} = \mathbf{x} \mathbf{x}^T, \quad \mathbf{x} \in R^n, A \in R^{n \times n} \quad (9.3.29)$$

$$\frac{d}{dA} \text{tr}(A) = I_n, \quad A \in R^{n \times n} \quad (9.3.30)$$

$$\frac{d}{dA} \text{tr}(AB) = B + B^T - \text{diag}(B), \quad A \text{ 为对称矩阵} \quad (9.3.31)$$

下面不加证明地引进一些常用的迹导数计算公式。

$$\frac{d}{dA} \text{tr}(AA^T) = \frac{d}{dA} \text{tr}(A^T A) = 2A \quad (9.3.32)$$

$$\frac{d}{dA} \text{tr}(ABA^T) = A(B + B^T) \quad (9.3.33)$$

$$\frac{d}{dA} \text{tr}(ABA^T) = 2AB, \quad B \text{ 为对称矩阵} \quad (9.3.34)$$

$$\frac{d}{dA} \text{tr}(A^T BA) = (B + B^T)A \quad (9.3.35)$$

$$\frac{d}{dA} \text{tr}(A^T BA) = 2BA, \quad B \text{ 为对称矩阵} \quad (9.3.36)$$

$$\frac{d}{dA} \text{tr}(A^2) = \frac{d}{dA} \text{tr}(AA) = 2A^T, \quad A \in R^{n \times n} \quad (9.3.37)$$

$$\frac{d}{dA} \text{tr}(BA^{-1}) = -(A^{-1}BA^{-1})^T, \quad A, B \in R^{n \times n} \quad (9.3.38)$$

$$\frac{d}{dA} \text{tr}(BA^{-1}C) = -(A^{-1}CBA^{-1})^T, \quad A, B, C \in R^{n \times n} \quad (9.3.39)$$

9.3.4 函数矩阵关于矩阵的导数

定义 9.3.4 设 $X = (x_{ij})_{m \times n}$ ，函数矩阵

$$F(X) = \begin{pmatrix} f_{11}(X) & \cdots & f_{1s}(X) \\ \vdots & & \vdots \\ f_{r1}(X) & \cdots & f_{rs}(X) \end{pmatrix}$$

其中 $f_{ij}(X)$ 是 $m \times n$ 元实值函数，定义

$$\frac{dF}{dX} = \begin{pmatrix} \frac{\partial F}{\partial x_{11}} & \frac{\partial F}{\partial x_{12}} & \cdots & \frac{\partial F}{\partial x_{1n}} \\ \frac{\partial F}{\partial x_{21}} & \frac{\partial F}{\partial x_{22}} & \cdots & \frac{\partial F}{\partial x_{2n}} \\ \vdots & \vdots & & \vdots \\ \frac{\partial F}{\partial x_{n1}} & \frac{\partial F}{\partial x_{n2}} & \cdots & \frac{\partial F}{\partial x_{nn}} \end{pmatrix} \quad (9.3.40)$$

其中

$$\frac{\partial F}{\partial x_{ij}} = \begin{pmatrix} \frac{\partial f_{11}}{\partial x_{ij}} & \frac{\partial f_{12}}{\partial x_{ij}} & \cdots & \frac{\partial f_{1s}}{\partial x_{ij}} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_{r1}}{\partial x_{ij}} & \frac{\partial f_{r2}}{\partial x_{ij}} & \cdots & \frac{\partial f_{rs}}{\partial x_{ij}} \end{pmatrix} \quad (9.3.41)$$

例 9.3.6 设 $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$, $F(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_n(\mathbf{x}))$, 求 $\frac{dF}{d\mathbf{x}}$

$$\text{解 } \frac{dF}{d\mathbf{x}} = \left(\frac{\partial F}{\partial x_1}, \frac{\partial F}{\partial x_2}, \dots, \frac{\partial F}{\partial x_n} \right)^T = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_2}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_1} \\ \frac{\partial f_1}{\partial x_2} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_2} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_1}{\partial x_n} & \frac{\partial f_2}{\partial x_n} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}.$$

利用矩阵的直积^{*)}, 可以给出函数矩阵乘积关于矩阵导数表达式。

定理 9.3.1 (函数矩阵乘积的导数) 设 $X = (x_{pq})_{m \times n}$, $F(X) = (f_{ij}(X))_{s \times r}$, $G(X) = (g_{ij}(X))_{r \times t}$,

则

$$\frac{d(FG)}{dX} = \frac{dF}{dX} (I_n \otimes G) + (I_m \otimes F) \frac{dG}{dX} \quad (9.3.42)$$

证明 令 $A(X) = F(X)G(X)$, 则

$$a_{ij}(X) = \sum_{k=1}^r f_{ik}(X)g_{kj}(X), \quad i = 1, 2, \dots, s; j = 1, 2, \dots, t$$

根据导数定义, 有

^{*)} 见 9.4 节。

$$\begin{aligned} \frac{dA}{dX} &= \left(\frac{\partial A}{\partial x_{pq}} \right)_{m \times n}, \\ \frac{\partial A}{\partial x_{pq}} &= \left(\frac{\partial a_{ij}}{\partial x_{pq}} \right)_{s \times t} = \left(\sum_{k=1}^r \frac{\partial (f_{ik} g_{kj})}{\partial x_{pq}} \right)_{s \times t} = \left(\sum_{k=1}^r \frac{\partial f_{ik}}{\partial x_{pq}} \cdot g_{kj} \right)_{s \times t} + \left(\sum_{k=1}^r f_{ik} \cdot \frac{\partial g_{kj}}{\partial x_{pq}} \right)_{s \times t} \\ &= \frac{\partial F}{\partial x_{pq}} \cdot G + F \cdot \frac{\partial G}{\partial x_{pq}} \end{aligned}$$

所以,

$$\begin{aligned} \frac{dA}{dX} &= \left(\frac{\partial F}{\partial x_{pq}} \cdot G \right)_{m \times n} + \left(F \cdot \frac{\partial G}{\partial x_{pq}} \right)_{m \times n} \\ &= \begin{pmatrix} \frac{\partial F}{\partial x_{11}} \cdot G & \cdots & \frac{\partial F}{\partial x_{1n}} \cdot G \\ \vdots & & \vdots \\ \frac{\partial F}{\partial x_{m1}} \cdot G & \cdots & \frac{\partial F}{\partial x_{mn}} \cdot G \end{pmatrix} + \begin{pmatrix} F \cdot \frac{\partial G}{\partial x_{11}} & \cdots & F \cdot \frac{\partial G}{\partial x_{1n}} \\ \vdots & & \vdots \\ F \cdot \frac{\partial G}{\partial x_{m1}} & \cdots & F \cdot \frac{\partial G}{\partial x_{mn}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial F}{\partial x_{11}} & \cdots & \frac{\partial F}{\partial x_{1n}} \\ \vdots & & \vdots \\ \frac{\partial F}{\partial x_{m1}} & \cdots & \frac{\partial F}{\partial x_{mn}} \end{pmatrix} \begin{pmatrix} G & & \\ & \ddots & \\ & & G \end{pmatrix} + \begin{pmatrix} F & & \\ & \ddots & \\ & & F \end{pmatrix} \begin{pmatrix} \frac{\partial G}{\partial x_{11}} & \cdots & \frac{\partial G}{\partial x_{1n}} \\ \vdots & & \vdots \\ \frac{\partial G}{\partial x_{m1}} & \cdots & \frac{\partial G}{\partial x_{mn}} \end{pmatrix} \\ &= \frac{dF}{dX} (I_n \otimes G) + (I_m \otimes F) \frac{dG}{dX} \end{aligned}$$

因此, 式(9.3.42)成立。证毕。

例 9.3.7 令 $f(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$, 其中 $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$, $A = (a_{ij})_{n \times n}$, 求 $\frac{df}{d\mathbf{x}}$ 。

解: 由定理 9.3.1, 我们有 $\frac{df}{d\mathbf{x}} = \frac{d\mathbf{x}^T}{d\mathbf{x}} (I_1 \otimes A \mathbf{x}) + (I_n \otimes \mathbf{x}^T) \frac{dA\mathbf{x}}{d\mathbf{x}}$ 。因为,

$$\frac{d\mathbf{x}^T}{d\mathbf{x}} = \begin{pmatrix} \frac{\partial \mathbf{x}^T}{\partial x_1} \\ \vdots \\ \frac{\partial \mathbf{x}^T}{\partial x_n} \end{pmatrix} = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} = I_n, \quad (I_1 \otimes A \mathbf{x}) = A \mathbf{x}$$

所以,

$$\frac{dx^T}{dx}(I_1 \otimes Ax) = Ax$$

又因，

$$(I_n \otimes x^T) = \begin{pmatrix} x^T & & \\ & \ddots & \\ & & x^T \end{pmatrix}, \quad \frac{dAx}{dx} = \begin{pmatrix} \frac{\partial Ax}{\partial x_1} \\ \vdots \\ \frac{\partial Ax}{\partial x_n} \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

所以，

$$(I_n \otimes x^T) \frac{dAx}{dx} = \begin{pmatrix} x^T & & \\ & \ddots & \\ & & x^T \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} x^T a_1 \\ \vdots \\ x^T a_n \end{pmatrix} = A^T x$$

于是，

$$\frac{d(x^T Ax)}{dx} = \frac{dx^T}{dx}(I_1 \otimes Ax) + (I_n \otimes x^T) \frac{dAx}{dx} = (A + A^T)x \quad (9.3.43)$$

例 9.3.8 令 $f(A) = A^T A$ ，其中 $A = (a_{ij})_{n \times n}$ ，求 $\frac{df}{dA}$ 。

解 由定理 9.3.1，我们有 $\frac{df}{dA} = \frac{dA^T}{dA}(I_n \otimes A) + (I_n \otimes A^T) \frac{dA}{dA}$ 。不难计算：

$$\frac{dA}{dA} = \begin{pmatrix} \frac{\partial A}{\partial a_{ij}} \end{pmatrix} = (E_{ij})_{n \times n}, \quad \frac{dA^T}{dA} = \begin{pmatrix} \frac{\partial A^T}{\partial a_{ij}} \end{pmatrix} = (E_{ji})_{n \times n},$$

所以，

$$\frac{d(A^T A)}{dA} = (E_{ji})_{n \times n} (I_n \otimes A) + (I_n \otimes A^T) (E_{ij})_{n \times n} = (E_{ji} A + A^T E_{ij})_{n \times n} \quad (9.3.44)$$

例 9.3.9 求 $\frac{dA^{-1}}{dA}$ 。

解 由于 $A^{-1}A = I$ ，所以， $\frac{dA^{-1}}{dA}(I_n \otimes A) + (I_n \otimes A^{-1}) \frac{dA}{dA} = 0$ 。于是，

$$\frac{dA^{-1}}{dA} = -(I_n \otimes A^{-1}) \frac{dA}{dA} (I_n \otimes A)^{-1} = -(I_n \otimes A^{-1}) (E_{ij})_{n \times n} (I_n \otimes A^{-1})$$

即

$$\frac{dA^{-1}}{dA} = -(A^{-1} E_{ij} A^{-1})_{n \times n} \quad (9.3.45)$$

例 9.3.10 $f(A) = \mathbf{x}^T A^{-1} \mathbf{x}$ ，求 $\frac{df}{dA}$ 。

$$\begin{aligned} \text{解 } \frac{df}{dA} &= \frac{d\mathbf{x}^T}{dA} (I \otimes A^{-1} \mathbf{x}) + (I \otimes \mathbf{x}^T) \frac{dA^{-1} \mathbf{x}}{dA} = (I \otimes \mathbf{x}^T) \left[\frac{dA^{-1}}{dA} (I \otimes \mathbf{x}) + (I \otimes A^{-1}) \frac{d\mathbf{x}}{dA} \right] \\ &= (I \otimes \mathbf{x}^T) \frac{dA^{-1}}{dA} (I \otimes \mathbf{x}), \end{aligned}$$

因此，

$$\frac{d(\mathbf{x}^T A^{-1} \mathbf{x})}{dA} = (I \otimes \mathbf{x}^T) \frac{dA^{-1}}{dA} (I \otimes \mathbf{x}) \quad (9.3.46)$$

9.4 矩阵直积

9.4.1 基本性质

定义 9.4.1 设 $A = (a_{ij}) \in C^{m \times n}$, $B = (b_{ij}) \in C^{p \times q}$ ，称分块矩阵：

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{pmatrix} \in C^{mp \times nq}$$

为矩阵 A 与 B 的直积，或 Kronecker 积，或张量积。 $A \otimes B$ 是一个 $m \times n$ 块的分块矩阵，简记为 $A \otimes B = (a_{ij}B)_{m \times n}$ 。

例 9.4.1 设 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $B = \begin{pmatrix} x \\ y \end{pmatrix}$ ，则

$$A \otimes B = \begin{pmatrix} aB & bB \\ cB & dB \end{pmatrix} = \begin{pmatrix} ax & bx \\ ay & by \\ cx & dx \\ cy & dy \end{pmatrix}_{4 \times 2}$$

而，

$$B \otimes A = \begin{pmatrix} xA \\ yA \end{pmatrix} = \begin{pmatrix} xa & xb \\ xc & xd \\ ya & yb \\ yc & yd \end{pmatrix} = \begin{pmatrix} ac & bx \\ cx & dx \\ ay & by \\ cy & dy \end{pmatrix}_{4 \times 2}$$

因此，矩阵的直积运算一般不满足交换律。但是，对于单位矩阵，有

$$I_n \otimes I_m = I_m \otimes I_n = I_{mn}$$

不难验证直积满足下述运算律：

- (1) 数乘律 $k(A \otimes B) = kA \otimes B = A \otimes kB, k \in \mathbb{C}$;
- (2) 分配律 $(A+B) \otimes C = A \otimes C + B \otimes C$;
- (3) 结合律 $(A \otimes B) \otimes C = A \otimes (B \otimes C)$

定理 9.4.1 设 $A = (a_{ij})_{m \times n}, B = (b_{ij})_{s \times r}, C = (c_{ij})_{n \times p}, D = (d_{ij})_{r \times t}$ ，则

$$(A \otimes B)(C \otimes D) = AC \otimes BD \quad (9.4.1)$$

证明：通过直接运算，有

$$(A \otimes B)(C \otimes D) = (a_{ij}B)(c_{ij}D) = \left(\sum_{k=1}^n a_{ik}c_{kj}BD \right) = ((AC)_{ij}BD) = AC \otimes BD$$

式中 $(AC)_{ij}$ 是矩阵 AC 中第 (i, j) 元素。

定理 9.4.1 是直积的最基本性质，直积的其它许多性质都可以由它导出。更一般的结论是

$$\prod_{j=1}^n (A_j \otimes B_j) = \left(\prod_{j=1}^n A_j \right) \otimes \left(\prod_{j=1}^n B_j \right) \quad (9.4.2)$$

$$\left(\bigotimes_{j=1}^n A_j \right) \left(\bigotimes_{j=1}^n B_j \right) = \bigotimes_{j=1}^n (A_j B_j) \quad (9.4.3)$$

若 $A \in \mathbb{C}^{m \times m}, B \in \mathbb{C}^{n \times n}$ ，则根据定理 9.4.1，有

$$A \otimes B = (A \otimes I_n)(I_m \otimes B) = (I_m \otimes B)(A \otimes I_n) \quad (9.4.4)$$

这是将两个矩阵的直积转化为普通积的公式，也经常使用。

直积的几何意义

下面采用“被动描述”方法，讨论平面内绕定点的旋转运动。在这种描述下，平面向量在旋转运动下没有发生变化，而变化的是坐标基。旋转运动前的基向量为 $\{e_1, e_2\}$ ，运动后的基向量为 $\{e'_1, e'_2\}$ ，如图 9.4.1 所示。令向量 x 在前、后坐标基下的表示为

$$\mathbf{x} = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 = x'_1 \mathbf{e}'_1 + x'_2 \mathbf{e}'_2$$

由于

$$\begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \mathbf{e}'_1 \\ \mathbf{e}'_2 \end{pmatrix}$$

所以，

$$\begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{9.4.5}$$

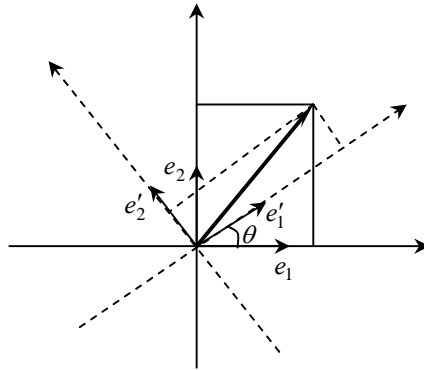


图 9.4.1 直积的几何意义

令向量 \mathbf{x}, \mathbf{y} 在基 $\{\mathbf{e}_1, \mathbf{e}_2\}, \{\mathbf{e}'_1, \mathbf{e}'_2\}$ 下的坐标分别为

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \mathbf{x}' = \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix}; \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \mathbf{y}' = \begin{pmatrix} y'_1 \\ y'_2 \end{pmatrix};$$

这两个坐标向量的张量积为

$$\mathbf{x} \otimes \mathbf{y} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \otimes \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = (x_1 y_1, x_1 y_2, x_2 y_1, x_2 y_2)^T \tag{9.4.6}$$

$$\mathbf{x}' \otimes \mathbf{y}' = \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} \otimes \begin{pmatrix} y'_1 \\ y'_2 \end{pmatrix} = (x'_1 y'_1, x'_1 y'_2, x'_2 y'_1, x'_2 y'_2)^T \tag{9.4.7}$$

考虑 $\mathbf{x} \otimes \mathbf{y}$ 与 $\mathbf{x}' \otimes \mathbf{y}'$ 之间的变化规律。因为

$$\mathbf{x}' \otimes \mathbf{y}' = \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} \otimes \begin{pmatrix} y'_1 \\ y'_2 \end{pmatrix} = \left(\begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right) \otimes \left(\begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right)$$

记 $R(\theta) = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$, 从式(9.4.1), 我们有

$$\begin{pmatrix} R(\theta) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \end{pmatrix} \otimes \begin{pmatrix} R(\theta) \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \end{pmatrix} = (R(\theta) \otimes R(\theta)) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \otimes \begin{pmatrix} y_1 \\ y_2 \end{pmatrix},$$

即

$$\mathbf{x}' \otimes \mathbf{y}' = (R(\theta) \otimes R(\theta))(\mathbf{x} \otimes \mathbf{y}) \quad (9.4.8)$$

上式表明, 在平面旋转运动的作用下, 两个向量的张量积的变化与旋转运动矩阵的张量积相关联。满足式(9.4.8)的量 $\mathbf{x} \otimes \mathbf{y}$ 通常称为二维旋转群上的 2 阶张量, 下章我们给出一般张量的概念。

直积的转置(共轭转置)

定理 9.4.2 设 $A = (a_{ij})_{m \times n}, B = (b_{ij})_{p \times q}$, 则

$$(A \otimes B)^T = A^T \otimes B^T, \quad (A \otimes B)^H = A^H \otimes B^H \quad (9.4.9)$$

证明 因为

$$(A \otimes B)^T = (a_{ij} B)^T = \begin{pmatrix} a_{11} B & \cdots & a_{1n} B \\ \vdots & & \vdots \\ a_{m1} B & \cdots & a_{mn} B \end{pmatrix}^T = \begin{pmatrix} a_{11} B^T & \cdots & a_{m1} B^T \\ \vdots & & \vdots \\ a_{1n} B^T & \cdots & a_{mn} B^T \end{pmatrix} = A^T \otimes B^T$$

同理可验证: $(A \otimes B)^H = A^H \otimes B^H$ 。

利用定理 9.4.1 和 9.4.2, 不难验证:

- (1) 若 A, B 均为对角矩阵, 则 $A \otimes B$ 也是对角矩阵;
- (2) 若 A, B 均为对称矩阵, 则 $A \otimes B$ 也是对称矩阵;
- (3) 若 A, B 均为 Hermite 矩阵, 则 $A \otimes B$ 也是 Hermite 矩阵;
- (4) 若 A, B 均为正交(酉)矩阵, 则 $A \otimes B$ 也是正交(酉)矩阵。

直积的逆

定理 9.4.3 设 A, B 分别为 m 阶和 n 阶可逆矩阵, 则 $A \otimes B$ 也为可逆矩阵, 且

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1} \quad (9.4.10)$$

证明 根据式(9.4.1), 有

$$(A \otimes B)(A^{-1} \otimes B^{-1}) = (AA^{-1} \otimes BB^{-1}) = I_m \otimes I_n = I_{mn}$$

因此,

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$$

从式(9.4.9)和(9.4.10)可以看出, 矩阵通常乘积的转置(共轭转置)与求逆的反序法则对于直积不再成立。

直积的秩

定理 9.4.4 设 $A = (a_{ij})_{m \times n}$, $B = (b_{ij})_{p \times q}$, 则

$$\text{rank}(A \otimes B) = \text{rank}(A)\text{rank}(B) \quad (9.4.11)$$

证明: 令 A_1, B_1 分别为 A 与 B 的标准形, 即

$$MAN = A_1, \quad PBQ = B_1$$

其中: M, N, P, Q 分别为 m, n, p 和 q 阶非奇异矩阵, 且

$$A_1 = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}, \quad B_1 = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

在 A_1 中, 数 1 的个数等于 $\text{rank}(A)$; 在 B_1 中, 数 1 的个数等于 $\text{rank}(B)$ 。于是, 有

$$A = M^{-1}A_1N^{-1}, \quad B = P^{-1}B_1Q^{-1}$$

因此, 根据定理 9.4.1, 有

$$A \otimes B = (M^{-1}A_1N^{-1}) \otimes (P^{-1}B_1Q^{-1}) = (M^{-1} \otimes P^{-1})(A_1 \otimes B_1)(N^{-1} \otimes Q^{-1})$$

由定理 9.4.3 知, $M^{-1} \otimes P^{-1}, N^{-1} \otimes Q^{-1}$ 均为非奇异矩阵, 所以,

$$\text{rank}(A \otimes B) = \text{rank}(A_1 \otimes B_1)$$

显然, $A_1 \otimes B_1$ 的秩为 $\text{rank}(A)\text{rank}(B)$, 所以

$$\text{rank}(A \otimes B) = \text{rank}(A)\text{rank}(B)$$

定理 9.4.4 说明矩阵直积的秩与普通积的秩有显著的差异, 普通积的秩不会超过每个因子矩阵的秩。

直积的迹

定理 9.4.5 设 A 为 m 阶矩阵, B 为 p 阶矩阵, 则

$$\text{tr}(A \otimes B) = \text{tr}(A)\text{tr}(B) \quad (9.4.12)$$

直积的行列式

定理 9.4.6 设 A 为 m 阶矩阵, B 为 p 阶矩阵, 则

$$\det(A \otimes B) = (\det(A))^p (\det(B))^m \quad (9.4.13)$$

证明 记

$$J_A = \begin{pmatrix} \lambda & & * \\ & \ddots & \\ 0 & & \lambda_m \end{pmatrix}, \quad J_B = \begin{pmatrix} \mu_1 & & * \\ & \ddots & \\ 0 & & \mu_p \end{pmatrix}$$

分别为 A, B 的 Jordan 标准形。由线性代数知, 存在可逆矩阵 P 与 Q , 使得

$$P^{-1}AP = J_A, \quad Q^{-1}BQ = J_B$$

根据式 (9.4.1), 有

$$A \otimes B = (PJ_A P^{-1}) \otimes (QJ_B Q^{-1}) = (P \otimes Q)(J_A \otimes J_B)(P \otimes Q)^{-1}$$

于是

$$\det(A \otimes B) = \det(J_A \otimes J_B)$$

显然, J_A, J_B 均为上三角矩阵, 所以 $(J_A \otimes J_B)$ 也为上三角矩阵, 故有

$$\begin{aligned} \det(A \otimes B) &= \det(J_A \otimes J_B) = \prod_{j=1}^m (\lambda_j \mu_j) \prod_{j=1}^p (\lambda_j \mu_j) \cdots \prod_{j=1}^p (\lambda_m \mu_j) \\ &= \prod_{j=1}^m (\lambda_j^p) \prod_{j=1}^p (\mu_j^m) = (\det(A))^p (\det(B))^m \end{aligned}$$

其中 $\lambda_1, \lambda_2, \dots, \lambda_m$ 为 A 的特征值, $\mu_1, \mu_2, \dots, \mu_p$ 是 B 的特征值。

9.4.2 特征值与特征向量

定理 9.4.7 设 $\lambda_1, \lambda_2, \dots, \lambda_m$ 是 $A_{m \times m}$ 的 m 个特征值, $\mu_1, \mu_2, \dots, \mu_p$ 是 $B_{p \times p}$ 的 p 个特征值, 那么 $A \otimes B$ 的 mp 个特征值为 $\lambda_i \mu_j (i=1, 2, \dots, m; j=1, 2, \dots, p)$ 。

证明 根据定理 9.4.6 的证明, 我们有 $A \otimes B = (P \otimes Q)(J_A \otimes J_B)(P \otimes Q)^{-1}$, 而

$$J_A \otimes J_B = \begin{pmatrix} \lambda_1 J_B & & * \\ & \ddots & \\ 0 & & \lambda_m J_B \end{pmatrix}$$

所以,

$$A \otimes B \sim \begin{pmatrix} \lambda_1 J_B & & * \\ & \ddots & \\ 0 & & \lambda_m J_B \end{pmatrix}$$

因此, $\lambda_i \mu_j (i=1, \dots, m; j=1, \dots, p)$ 是 $A \otimes B$ 的 mp 个特征值。

定理 9.4.8 设 A 为 m 阶矩阵, B 为 n 阶矩阵, 则 $A \otimes B \sim B \otimes A$ 。因而, $A \otimes B$ 与 $B \otimes A$ 有相同的特征值。

证明: 容易验证, 对矩阵 $A \otimes I_n$ 进行一系列“相合”变换(对矩阵的行和相应的列进行相同的初等变换, 这里是指对调矩阵的第 i 行与第 j 行, 然后再对调第 i 列与第 j 列。), 可以变成 $I_n \otimes A$, 即存在一个 mn 阶置换矩阵(有限个初等矩阵的乘积) P , 使

$$P^T (A \otimes I_n) P = I_n \otimes A$$

同理, 对矩阵 $I_m \otimes B$ 也有

$$P^T (I_m \otimes B) P = B \otimes I_m$$

不难验证: $P^T P = I$, 所以

$$\begin{aligned} P^T (A \otimes B) P &= P^T (A \otimes I_n) (I_m \otimes B) P = P^T (A \otimes I_n) P P^T (I_m \otimes B) P \\ &= (I_n \otimes A) (B \otimes I_m) = B \otimes A \end{aligned}$$

多项式的特征值

定义 9.4.2 设矩阵 $A \in C^{m \times n}$, 记

$$\underbrace{A \otimes A \otimes \dots \otimes A}_k = A^{\otimes k}$$

它是一个 $m^k \times n^k$ 矩阵。

定理 9.4.9 设 $A \in C^{m \times n}$, $B \in C^{n \times p}$, 则

$$(AB)^{\otimes k} = A^{\otimes k} B^{\otimes k} \quad (9.4.14)$$

证明 用归纳法, 当 $k=1$ 时, 显然成立, 设 $k-1$ 时定理成立, 则

$$(AB)^{\otimes k} = (AB) \otimes (AB)^{\otimes(k-1)} = (AB) \otimes A^{\otimes(k-1)} B^{\otimes(k-1)} = (A \otimes A^{\otimes(k-1)})(B \otimes B^{\otimes(k-1)}) = A^{\otimes k} B^{\otimes k}$$

定义 9.4.3 设 $f(x, y) = \sum_{i,j=0}^p a_{ij} x^i y^j$ 是变量 x, y 的复系数多项式, 对于 $A \in C^{m \times m}, B \in C^{n \times n}$ 定义 mn 阶矩阵

$$f(A; B) = \sum_{i,j=0}^p a_{ij} A^i \otimes B^j \quad (9.4.15)$$

并称它为直积多项式。

关于直积多项式的特征值, 有下面的结论。

定理 9.4.10 如果 A 和 B 的特征值分别是 $\lambda_1, \lambda_2, \dots, \lambda_m$ 和 $\mu_1, \mu_2, \dots, \mu_n$, 它们对应的特征向量分别是 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$ 和 $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$, 则直积多项式 $f(A; B)$ 的特征值是 $f(\lambda_r; \mu_s)$, 而对应 $f(\lambda_r; \mu_s)$ 的特征向量为 $\mathbf{x}_r \otimes \mathbf{y}_s$ ($r = 1, \dots, m; s = 1, \dots, n$)。

证明 由于 $A\mathbf{x}_r = \lambda_r \mathbf{x}_r$, $B\mathbf{y}_s = \mu_s \mathbf{y}_s$, 所以,

$$A^i \mathbf{x}_r = \lambda_r^i \mathbf{x}_r, \quad B^j \mathbf{y}_s = \mu_s^j \mathbf{y}_s$$

于是,

$$\begin{aligned} f(A; B)\mathbf{x}_r \otimes \mathbf{y}_s &= \left(\sum_{i,j=0}^p a_{ij} A^i \otimes B^j \right) (\mathbf{x}_r \otimes \mathbf{y}_s) \\ &= \sum_{i,j=0}^p a_{ij} (A^i \otimes B^j) (\mathbf{x}_r \otimes \mathbf{y}_s) = \sum_{i,j=0}^p a_{ij} (A^i \mathbf{x}_r \otimes B^j \mathbf{y}_s) \\ &= \sum_{i,j=0}^p a_{ij} \lambda_r^i \mu_s^j \mathbf{x}_r \otimes \mathbf{y}_s = f(\lambda_r, \mu_s) \mathbf{x}_r \otimes \mathbf{y}_s \end{aligned}$$

取 $f(x, y) = xy$, 则 $f(A; B) = A \otimes B$ 。根据定理 9.4.10, 我们有

推论 9.4.1 $A \otimes B$ 的特征值为 mn 个数 $\lambda_r \mu_s$ ($r = 1, \dots, m; s = 1, \dots, n$), 且对应 $\lambda_r \mu_s$ 的特征向量为 $\mathbf{x}_r \otimes \mathbf{y}_s$ 。

取 $f(x, y) = x + y$, 即 $f(x, y) = xy^0 + x^0 y$, 则 $f(A; B) = A \otimes I_n + I_m \otimes B$ (称为 A 与 B 的直和)。所以, 根据定理 9.4.10, 有:

推论 9.4.2 $A \otimes I_n + I_m \otimes B$ 的特征值是 $\lambda_r + \mu_s$, 其对应的特征向量是

$$\mathbf{x}_r \otimes \mathbf{y}_s, \quad r = 1, \dots, m; s = 1, \dots, n。$$

10. 张量代数

10.1 张量概述

一阶张量

令 R 是全体实数所构成的一维实空间。

设 $x = (x^1, x^2, \dots, x^n)^T$ 是 n 维列向量，其中每个分量 $x^i \in R$ ，将 n 维列向量的全体所构成 n 维列向量空间记为 R^n ；设 $y = (y_1, y_2, \dots, y_n)$ 是 n 维行向量，其中每个分量 $y_i \in R$ ，将 n 维行向量的全体所构成 n 维行向量空间记为 R_n ；并假定 R^n ， R_n 具有标准正交基。

R^n 中的元素称为一阶逆变张量，元素分量中的上标称为逆变指标； R_n 中的元素称为一阶协变张量，元素分量中的下标称为协变指标；而 R 中的元素通常称为零阶张量。

给定 R^n 中的一个协变张量 $y = (y_1, y_2, \dots, y_n)$ ，按下述方式唯一确定 R^n 上的一个线性函数：

$$f(x) = \sum_{i=1}^n y_i x^i = \langle x, y \rangle \quad (10.1.1)$$

反之， R_n 上的任一个线性函数 $f(x) = \sum_{i=1}^n a_i x^i = \langle x, a \rangle$ 也唯一确定 R_n 中的一个协变张量 $a = (a_1, a_2, \dots, a_n)$ 。因此，一个协变张量对应于 R^n 上的一个线性函数。通常，将协变张量看作是 R^n 上的线性函数，并称 R_n 是 R^n 的对偶空间，即 R^n 上的所有线性函数所构成的空间。

类似地， R^n 中的逆变张量与 R_n 上的线性函数也具有这种一一对应关系，因此逆变张量是 R_n 上的线性函数。也就是说， R^n 也是 R_n 上的所有线性函数所构成的空间，即 R^n 也是 R_n 的对偶空间。

二阶张量

n 阶方阵，按照张量的习惯记法，记为 (a_i^j) ，或简记为 a_i^j 。令 $x^i \in R^n$ ，则

$$y^j = \sum_{i=1}^n a_i^j x^i, \quad j = 1, 2, \dots, n \quad (10.1.2)$$

确定了 R^n 中的一个元素(一阶逆变张量) $y = (y^1, y^2, \dots, y^n)^T$ 。不难看出，式(10.1.2)关于 x^i 是线性的，即矩阵 (a_i^j) 确定了 R^n 到 R^n 上的一个线性变换。反之， R^n 到 R^n 上的任一个线性变换 a 也唯一确定一个矩阵 (a_i^j) 。即 R^n 到 R^n 上的线性变换所构成的空间是矩阵空间，这个矩阵空间记为 $R^n \otimes R_n$ 。在张量代数中， $R^n \otimes R_n$ 的元素称为二阶混合张量，它具有一个逆变指标和一个协变指标。

给定 $a_i^j \in R^n \otimes R_n$, 对任意的 $(x, y) \in R^n \times R_n$, 定义:

$$f(x, y) = \sum_{i,j=1}^n a_i^j x^i y_j \quad (10.1.3)$$

则它是 $R^n \times R_n$ 上的一个双线性函数, 所谓双线性函数是指对每一个变元来说都是线性的。

对于二阶张量还有逆变二阶张量 $a^{ij} \in R^n \otimes R^n$ 和二阶协变张量 $a_{ij} \in R_n \otimes R_n$ 。它们按下述方式:

$$x^i = \sum_{j=1}^n a^{ij} y_j, y = (y_1, \dots, y_n) \in R_n \quad (10.1.4)$$

$$y_i = \sum_{j=1}^n a_{ij} x^j, x = (x^1, \dots, x^n) \in R^n \quad (10.1.5)$$

分别确定了 R_n 到 R^n 上的一个线性变换和 R^n 到 R_n 上的一个线性变换。并且按照下述方式:

$$f(y, y') = \sum_{i,j=1}^n a^{ij} y_i y'_j, y = (y_1, \dots, y_n), y' = (y'_1, \dots, y'_n) \in R_n \quad (10.1.6)$$

$$f(x, x') = \sum_{i,j=1}^n a_{ij} x^i x'^j, x = (x^1, \dots, x^n), x' = (x'^1, \dots, x'^n) \in R^n \quad (10.1.7)$$

分别确定了 $R_n \times R_n$ 上的一个双线性函数和 $R^n \times R^n$ 上的一个双线性函数。

例如: 克罗内克(kronecker)符号:

$$\delta^{ij} = \delta_{ij} = \delta_j^i = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

分别定义了一个二阶逆变张量、协变张量与混合张量。

不难看出:

$$(a) \delta^{ij}(y) = \sum_{j=1}^n \delta^{ij} y_j = \begin{cases} y_i, & i = j \\ 0, & i \neq j \end{cases} \text{ 是 } R_n \text{ 到第 } i \text{ 个坐标的投影变换};$$

$$(b) \delta^{ij}(y, y') = \sum_{i,j=1}^n \delta^{ij} y_i y'_j = \sum_{k=1}^n y_k y'_k = \langle y, y' \rangle \text{ 是 } R_n \text{ 中的内积}.$$

类似地, 可以给出 δ_{ij}, δ_j^i 所确定的线性变换和双线性函数。

三阶张量

在 n 维空间中, 三阶张量有下述 4 种形式: 逆变张量 $a^{ijk} \in R^n \otimes R^n \otimes R^n$, 协变张量 $a_{ijk} \in R_n \otimes R_n \otimes R_n$, 混合张量 $a_i^{jk} \in R^n \otimes R^n \otimes R_n$, $a_{jk}^i \in R_n \otimes R_n \otimes R^n$, 其中所有指标都在 $\{1, 2, \dots, n\}$ 内取值。它们都是 n^3 个数构成三维数组。

例如: 在三维空间中, 排列符号(交错符号):

$$\varepsilon^{ijk} = \varepsilon_{ijk} = \begin{cases} 1, & (ijk) \text{ 为偶排列} \\ -1, & (ijk) \text{ 为奇排列} \\ 0, & \text{其它} \end{cases}$$

分别定义了一个三阶逆变张量与一个三阶协变张量。如果 (ijk) 是由 (123) 通过偶数次对换其中的两个元素所得到的排列, 称之为偶排列, 否则称为奇排列。例如 (231) 是偶排列, 而 (213) 是奇排列。

给定 $a^{ijk} \in R^n \otimes R^n \otimes R^n$ 。 $\forall y = (y_k) \in R_n$, 定义:

$$b^{ij} = \sum_{k=1}^n a^{ijk} y_k, i, j = 1, 2, \dots, n$$

它是一个二阶逆变张量, 即 $(b^{ij}) \in R^n \otimes R^n$ 。并且确定了 R_n 到 $R^n \otimes R^n$ 的一个线性映射。

$\forall y = (y_k), y' = (y'_j) \in R_n$, 定义:

$$c^i = \sum_{k=1}^n a^{ijk} y_k y'_j, i = 1, 2, \dots, n$$

它是一个一阶逆变张量, 即 $(c^i) \in R^n$ 。并且上式确定了 $R_n \times R_n$ 到 R^n 的一个双线性映射, 即对每一个变元来说都是线性映射。

$\forall y = (y_k), y' = (y'_j), y'' = (y''_i) \in R_n$, 定义:

$$d = \sum_{i=1}^n a^{ijk} y_k y'_j y''_i$$

它是零阶张量, 即标量。它是 $R_n \times R_n \times R_n$ 的三重线性函数。

反之, R_n 到 $R^n \times R^n$ 的一个线性映射、 $R_n \times R_n$ 到 R^n 的一个双线性映射和 $R_n \times R_n \times R_n$ 的三重线性函数也分别对应于一个三阶逆变张量。

类似地, 可以给出 $R_n \otimes R_n \otimes R_n$, $R^n \otimes R^n \otimes R_n$, $R^n \otimes R_n \otimes R_n$ 中元素的线性性质(即按照上述类似方式, 确定线性映射、双线性映射和三重线性函数的性质)。

概括地说, 可以利用三阶张量来描述各种线性关系, 尤其是双线性映射与三重线性函数。

高阶张量

按照定义三阶张量的方式, 可以给出高阶张量的概念。引进高阶张量的目的主要是利用它来描述多重线性关系。由于本节只是在 n 维欧氏空间中、并且在特定的基底下给出张量的直观概念, 帮助读者对后面章节关于抽象张量概念的理解, 所以这里不再给出欧氏空间中的高阶张量的概念。

张量运算

在张量的运算中, 涉及到张量元素求和, 按照惯例遵守下述(Einstein)约定:

(1) 所有的逆变与协变指标均取从 1 到 n 的值;

(2) 逆变与协变指标中若有一个指标重复出现，则意味着要对这个指标遍历范围 $\{1,2,\dots,n\}$ 求和，这个指标通常称为一个哑指标。如果有多个哑指标，则意味着求多重和。没有求和意义的指标称为自由指标。

根据以上约定，求和

$$x^i = a_1^i x^1 + a_2^i x^2 + \dots + a_n^i x^n = \sum_{j=1}^n a_j^i x^j$$

则可简记为 $x^i = a_j^i x^j$ ， j 是一个哑指标， i 是自由指标。

张量加法 张量的加法是两个同型张量 $x = (x_{k_1 k_2 \dots k_s}^{i_1 i_2 \dots i_r})$, $y = (y_{k_1 k_2 \dots k_s}^{i_1 i_2 \dots i_r})$ 之间的一种运算，其定义如下：

$$((x + y)_{k_1 k_2 \dots k_s}^{i_1 i_2 \dots i_r}) = (x_{k_1 k_2 \dots k_s}^{i_1 i_2 \dots i_r} + y_{k_1 k_2 \dots k_s}^{i_1 i_2 \dots i_r})$$

张量数乘 数 α 乘以张量 $x = (x_{k_1 k_2 \dots k_s}^{i_1 i_2 \dots i_r})$ 的运算定义为：

$$((\alpha x)_{k_1 k_2 \dots k_s}^{i_1 i_2 \dots i_r}) = (\alpha x_{k_1 k_2 \dots k_s}^{i_1 i_2 \dots i_r})$$

张量乘积 $x = (x_{k_1 \dots k_{s_1}}^{i_1 \dots i_{r_1}})$, $y = (y_{k_1 \dots k_{s_2}}^{i_1 \dots i_{r_2}})$ 的乘积 $x \otimes y$ 是一个 $(r_1 + r_2, s_1 + s_2)$ 型张量，其定义如下：

$$((x \otimes y)_{k_1 \dots k_{s_1+s_2}}^{i_1 \dots i_{r_1+r_2}}) = (x_{k_1 \dots k_{s_1}}^{i_1 \dots i_{r_1}} y_{k_{s_1+1} \dots k_{s_1+s_2}}^{i_{r_1+1} \dots i_{r_1+r_2}})$$

张量缩并 张量的缩并是将某个上标与某个下标对等求和得到新张量的运算，例如：在 (r, s) 型张量 $x = (x_{k_1 \dots k_s}^{i_1 \dots i_r})$ 中，第 μ 个上标与第 λ 个下标对等求和得到新张量是一个 $(r-1, s-1)$ 型张量：

$$(C_{\mu\lambda} x)_{k_1 \dots k_{s-1}}^{i_1 \dots i_{r-1}} = \sum_{j=1}^n x_{k_1 \dots k_{\lambda-1} j k_{\lambda+1} \dots k_s}^{i_1 \dots i_{\mu-1} j i_{\mu+1} \dots i_r}$$

张量内乘 利用张量的乘积和张量的缩并，可得到张量的另一种重要运算，即张量的内乘。张量的内乘给出了张量的内积，它是向量内积在张量中的推广。两个张量内乘时，若一个张量的下标中的指标与另一张量的上标中的指标有若干个相同，则表示这若干个指标遍历范围 $\{1,2,\dots,n\}$ 求和，这种演算的结果称为两个张量的内积。显然，两个张量的内积仍是一个张量，其阶数是两个张量的阶数和减去哑指标数目的两倍。

例如：内积 $c_{jr}^k = a_{ij}^k b_r^i$ 是一个 3 阶张量； $c_r^k = a_{ij}^k b_r^{ji}$ 是一个 2 阶张量； $c_r = a_{ij} b_r^{ji}$ 是一阶协变张量， $c = a_i b^i$ 是 0 阶张量即标量。

下面结合图像空间，讨论张量间的内积运算。

用张量记号，图像上的点用一阶逆变张量来表示，即点的分量指标用上标来表示，记为 $x = (x^1, x^2, x^3)^T$ 或简记成 x^i ，直线用一阶协变张量来表示，即直线的分量指标用下标来表示，记为 $l = (l_1, l_2, l_3)$ 或简记成 l_i 。

线与点的内积是 $c = l_i x^i$ ，它是一个 0 阶张量，如果 $c=0$ ，则表示点在直线上，否则点不在直线

上。

对于空间平面的两幅图像(I, I')间的单应矩阵 H , 通常用一个二阶混合张量(h_j^i)来表示, i 表示矩阵元素所在的行指标, j 表示矩阵元素所在的列指标。单应 H 将 I 上的点(x^j)变换到 I' 上的点(x'^i), 写成张量形式, 有 $(x'^i) = (h_j^i x^j)$, 简单地记作 $x'^i = h_j^i x^j$ (这意味着自由变量 i 取遍 $\{1, 2, 3\}$, 对应于点的三个分量)。如果(l'_i)是 I' 上的线, 则张量积 $l_j = l'_i h_j^i$ 意味着它是 I 上的线, 不难看出这与单应矩阵对线的作用是一致的。

下面考虑张量 ε_{ijk} 与点 x^i 的内积 $c_{jk} = \varepsilon_{ijk} x^i$, 显然有

$$\begin{aligned} c_{11} &= \varepsilon_{i11} x^i = 0, c_{12} = \varepsilon_{i12} x^i = x^3, c_{13} = \varepsilon_{i13} x^i = -x^2 \\ c_{21} &= \varepsilon_{i21} x^i = -x^3, c_{22} = \varepsilon_{i22} x^i = 0, c_{23} = \varepsilon_{i23} x^i = x^1 \\ c_{31} &= \varepsilon_{i31} x^i = x^2, c_{32} = \varepsilon_{i32} x^i = -x^1, c_{33} = \varepsilon_{i33} x^i = 0 \end{aligned}$$

所以二阶张量 $c_{jk} = \varepsilon_{ijk} x^i$ 是点 x^i 所确定的反对称矩阵, 即 $(\varepsilon_{ijk} x^i) = [x^i]_{\times}$ 。这样, 两个点 x^i, x^j 的叉积可以表示为

$$x^i \times x^j = [x^i]_{\times} x^j = \varepsilon_{ijk} x^i x^j$$

而 $l_k = \varepsilon_{ijk} x^i x^j$ 是一阶协变张量, 它表示一条(通过点 x^i, x^j)直线, 这与我们熟知的结论是一致的。

对于张量 ε^{ijk} 与线 l_i, l_j , 有类似的结论: $\varepsilon^{ijk} l_i$ 是线 l_i 所确定的反对称矩阵 $[l_i]_{\times}$; $\varepsilon^{ijk} l_i l_j$ 是两线 l_i, l_j 的交点 $l_i \times l_j$ 。

10.2 张量积

10.2.1 线性映射

向量空间

在本章中, 用 F 表示数域, 通常是指实数域 R 或复数域 C 。

所谓域 F 上的向量空间 V 是指一个集合, 在这个集合中定义了两种运算: 加法与数乘法, 并要求这两种运算满足下列条件:

- 1) 集合 V 关于加法运算构交换群, 其单位元素是零向量, 记作 0 ;
- 2) 对于 $\alpha, \beta \in F, x, y \in V$ 有
 - (a) $\alpha(x + y) = \alpha x + \beta y$,
 - (b) $(\alpha + \beta)x = \alpha x + \beta x$,
 - (c) $(\alpha\beta)x = \alpha(\beta x)$,
 - (d) $0 \cdot x = 0, 1 \cdot x = x$.

空间 V 的元素称为向量, F 的元素称为数量。

如果 V 中存在 n 个元素 a_1, \dots, a_n , 使得 V 中任意一个元素可以唯一地表示成这 n 个元素的线性组合(系数属于数域 F), 则称 V 是 n 维向量空间。这样一组向量 $\{a_1, \dots, a_n\}$ 称为空间 V 的一个基。很明显, 在给定基之后, 向量空间 V 中的每一个元素可唯一地表示成:

$$v = v^1 a_1 + v^2 a_2 + \dots + v^n a_n \quad (v^i \in F)$$

通常, 称有序数组 (v^1, \dots, v^n) 为向量在基 $\{a_1, \dots, a_n\}$ 下的坐标。这样, 向量空间 V 就等同于坐标向量空间 F^n 。

线性函数

定义 10.1.1 设 $f: V \rightarrow F$ 是 V 上的函数。如果对任意 $v_1, v_2 \in V, a^1, a^2 \in F$ 的有

$$f(a^1 v_1 + a^2 v_2) = a^1 f(v_1) + a^2 f(v_2) \quad (10.2.1)$$

则称 f 是 V 上的线性函数。

显然, 如果 f, g 是 V 上的线性函数, 则 $\alpha f + \beta g$ 仍然是 V 上的线性函数。这样, 全体 V 上的线性函数的集合构成域 F 上的向量空间 $L(V; F)$, 记作 V^* , 并它称为 V 的对偶空间。

命题 10.2.1 如果 V 为 F 上的 n 维向量空间, 则 V^* 也是 F 上的 n 维向量空间。

证明 令 $\{a_1, \dots, a_n\}$ 是 V 的一个基, 将 V 中的向量表示为基的线性组合:

$$v = \sum_{i=1}^n v^i a_i \in V, f \in V^*$$

则必有

$$f(v) = \sum_{i=1}^n v^i f(a_i) \quad (10.2.2)$$

所以, 线性函数 f 由它在基 $\{a_1, \dots, a_n\}$ 上的值 $f(a_i) (i=1, \dots, n)$ 所确定。

定义线性函数 $a^{*i} \in V^* (i=1, \dots, n)$ 使得

$$a^{*i}(a_j) = \delta_j^i = \begin{cases} 1, & j=i \\ 0, & j \neq i \end{cases} \quad (j=1, \dots, n; i=1, \dots, n) \quad (10.2.3)$$

则 $a^{*i}(a) = a^i$, 它是 a 在第 i 坐标上的投影, 所以置 $f_i = f(a_i)$, 有

$$f(v) = \sum_{i=1}^n f_i a^i = \sum_{i=1}^n f_i a^{*i}(v)$$

也就是说 f 必为

$$f = \sum_{i=1}^n f_i v^{*i} \quad (10.2.4)$$

上式(10.2.4)说明 V^* 的任意一个元素都可以用 $\{a^{*1}, \dots, a^{*n}\}$ 线性表示。不难证明这种表示是唯一的，因此 $\{a^{*1}, \dots, a^{*n}\}$ 构成 V^* 的基，所以 V^* 也是 F 上的 n 维向量空间。

在命题 10.2.1 证明中，由 V 的基 $\{a_1, \dots, a_n\}$ 构造出对偶空间 V^* 的基 $\{a^{*1}, \dots, a^{*n}\}$ 通常称为 $\{a_1, \dots, a_n\}$ 的对偶基。

例 10.2.1 $R^n = \{x = (x^1, \dots, x^n)^T\}$ ，则 $R^{n*} = R_n = \{y = (y_1, \dots, y_n)\}$ ，且

$$\{e^{*i} = (0, \dots, 0, 1, 0, \dots, 0) : i = 1, \dots, n\}$$

是

$$\{e_i = (0, \dots, 0, 1, 0, \dots, 0)^T : i = 1, \dots, n\}$$

的对偶基。

事实上，令 f 为 R^n 上的线性函数，则存在唯一的行向量 $(f_1, \dots, f_n) \in R_n$ 使得

$$f(x) = f_1 x^1 + \dots + f_n x^n$$

反之，对任一 $(f_1, \dots, f_n) \in R_n$ ，上式也确定了 R^n 上的一个线性函数。所以， R^n 上的线性函数均可以表示成

$$f = (f_1, \dots, f_n) = \sum_{i=1}^n f_i e^{*i}$$

因此， $(R^n)^* = R_n$ ，且 $\{e^{*i} : i = 1, \dots, n\}$ 是 $\{e_i : i = 1, \dots, n\}$ 的对偶基。

不难证明，实数域 R 上的 n 维向量空间 V 的对偶 V^* 与 R_n 同构。从上面的例子还可以看出，对偶关系是相互的，即 $(R^n)^{**} = R_n^* = R^n$ ，对于一般情况也有这个事实。

命题 10.2.2 如果 V 是 F 上的 n 维向量空间，则 $(V^*)^* = V$ 。

证明 定义

$$\langle v, v^* \rangle = v^*(v) \tag{10.2.5}$$

则 $\langle \cdot, \cdot \rangle$ 是定义在 $V \times V^*$ 上的 F -值函数，并且对每一个变量而言都是线性的，即：对任意

$$v, v_1, v_2 \in V, v^*, v^{*1}, v^{*2} \in V^*, \alpha, \beta \in F$$

有

$$\begin{cases} \langle \alpha v_1 + \beta v_2, v^* \rangle = \alpha \langle v_1, v^* \rangle + \beta \langle v_2, v^* \rangle \\ \langle v, \alpha v^{*1} + \beta v^{*2} \rangle = \alpha \langle v, v^{*1} \rangle + \beta \langle v, v^{*2} \rangle \end{cases} \tag{10.2.6}$$

固定式(10.2.5)中的向量 $v \in V$ ，则 $\langle v, \cdot \rangle$ 是 V^* 上的 F -值线性函数；反之， $\langle \cdot, v^* \rangle$ 也是 V 上的 F -值线性函数。设 g 是 V^* 上的 F -值线性函数，令

$$v = \sum_{i=1}^n g(a^{*i}) a_i$$

则对任意的 $v^* \in V^*$ ，有

$$\langle v, v^* \rangle = \sum_{i=1}^n g(a^{*i}) \langle a_i, v^* \rangle = \sum_{i=1}^n g(a^{*i}) \cdot v^*(a_i) = \sum_{i=1}^n g(v^*(a_i) a^{*i}) = g(v^*) \quad (10.2.7)$$

因此 V 是 V^* 上的 F -值线性函数构成的向量空间，即 $(V^*)^* = V$ 。

线性映射

定义 10.2.2 假定 V, Z 均为 F 上有限维的向量空间。如果映射 $f: V \rightarrow Z$ 满足下述条件，则称它为线性映射，或者说它是线性的：对任意的 $v_1, v_2 \in V, a^1, a^2 \in F$ 有

$$f(a^1 v_1 + a^2 v_2) = a^1 f(v_1) + a^2 f(v_2) \quad (10.2.8)$$

当 $Z=F$ 时，此定义给出的是 F -值线性函数。

从 V 到 Z 的所有线性映射构成的集合记作 $L(V; Z)$ 。在 $L(V; Z)$ 中，定义如下加法和数乘运算：对任意 $f, g \in L(V; Z), \alpha \in F$ ，令

$$\begin{cases} (f+g)(v) = f(v) + g(v) \\ (\alpha f) = \alpha f(v) \end{cases} \quad (10.2.9)$$

则 $f+g, \alpha f \in L(V; Z)$ 。显然，集合 $L(V; Z)$ 关于上述两种运算构成 F 上的向量空间。

下面考虑空间 $L(V; Z)$ 的结构。为此，令 $\{a_1, \dots, a_n\}$ 是 V 的基， $\{b_1, \dots, b_m\}$ 是 Z 的基。如果 $f \in L(V; Z)$ ，则 $f(a_i)$ 必为 $\{b_1, \dots, b_m\}$ 的线性组合，即存在 $\{f_i^\alpha \mid \alpha = 1, 2, \dots, m\} \subset F_n$ 使得

$$f(a_i) = \sum_{\alpha=1}^m f_i^\alpha b_\alpha \quad (10.2.10)$$

因此，线性映射 f 由它在基 $\{a_1, \dots, a_n\}$ 上的作用所确定。于是，给定 V 的基底 $\{a_1, \dots, a_n\}$ 和 Z 的基底 $\{b_1, \dots, b_m\}$ 后，线性映射 f 唯一确定一个 $m \times n$ 阶矩阵 (f_i^α) 。反之，给定一个 $m \times n$ 矩阵 (f_i^α) (元素属于 F)，则在这组基下唯一确定一个线性映射 f 。所以，空间 $L(V; Z)$ 和 $m \times n$ 矩阵 (f_i^α) 所成的矩阵空间是同构的。

10.2.2 多重线性映射

多重线性函数

定义 10.2.3 假定 V, W 都是 F 上有限维的向量空间。函数 $f: V \times W \rightarrow F$ 称为双线性的，如果对每一个变量它都是线性的，也就是说

$$\begin{cases} f(a^1 v_1 + a^2 v_2, w) = a^1 f(v_1, w) + a^2 f(v_2, w) \\ f(v, a^1 w_1 + a^2 w_2) = a^1 f(v, w_1) + a^2 f(v, w_2) \end{cases} \quad (10.2.11)$$

类似地, 可以定义 r 重线性函数 $f: V_1 \times V_2 \times \dots \times V_r \rightarrow F$, 其中 V_1, V_2, \dots, V_r 都是 F 上的向量空间。所有 r 重线性映射 $f: V_1 \times V_2 \times \dots \times V_r \rightarrow F$ 的全体构成的集合记为 $L(V_1 \times V_2 \times \dots \times V_r; F)$ 。

下面考虑空间 $L(V \times W; F)$ 结构。为此, 需要引进对偶空间 V^* 和 W^* 的张量积概念。

定义 10.2.4 设 $v^* \in V^*, w^* \in W^*$, 线性函数 v^* 和 w^* 的张量积 $v^* \otimes w^*$ 定义为

$$v^* \otimes w^*(v, w) = v^*(v) \cdot w^*(w) = \langle v, v^* \rangle \cdot \langle w, w^* \rangle \quad (10.2.12)$$

其中 $v \in V, w \in W$ 。

显然, 两个线性函数 v^* 和 w^* 的张量积 $v^* \otimes w^*$ 是 $V \times W$ 上的一个双线性函数。也就是说, $v^* \otimes w^* \in L(V \times W; F)$ 。

例 10.2.2 $R^n = \{x = (x^1, \dots, x^n)^T\}$, 根据例 10.2.1, $R^{n*} = R_n = \{y = (y_1, \dots, y_n)\}$ 。取

$$a^* = (a_1, \dots, a_n), b^* = (b_1, \dots, b_n) \in R_n$$

则

$$a^* \otimes b^*(x, y) = a^*(x) \cdot b^*(y) = \sum_{i=1}^n a_i x^i \cdot \sum_{j=1}^n b_j y^j = \sum_{i,j=1}^n a_i b_j x^i y^j$$

其中:

$$x = (x^1, \dots, x^n)^T, y = (y^1, \dots, y^n)^T \in R^n$$

因此, $a^* \otimes b^*$ 是 $R^n \times R^n$ 上的双线性实值函数, 即 $a^* \otimes b^* \in L(R^n \times R^n; R)$ 。

对于固定的 $v \in V, w \in W$, 式(10.2.12)中的运算 “ \otimes ” 也是 $V^* \times W^*$ 上的双线性函数。即我们有下述命题:

命题 10.1.3 $\otimes \in L(V^* \times W^*; F)$ 。

证明 对于 $v^{*1}, v^{*2} \in V^*, w^* \in W^*, \alpha, \beta \in F$, 根据张量积的定义, 有

$$\begin{aligned} & [(\alpha v^{*1} + \beta v^{*2}) \otimes w^*](v, w) \\ &= \langle v, (\alpha v^{*1} + \beta v^{*2}) \rangle \cdot \langle w, w^* \rangle \\ &= \alpha \langle v, v^{*1} \rangle \cdot \langle w, w^* \rangle + \beta \langle v, v^{*2} \rangle \cdot \langle w, w^* \rangle \\ &= [\alpha v^{*1} \otimes w^* + \beta v^{*2} \otimes w^*](v, w) \end{aligned}$$

即

$$(\alpha v^{*1} + \beta v^{*2}) \otimes w^* = \alpha v^{*1} \otimes w^* + \beta v^{*2} \otimes w^* \quad (10.2.13)$$

同理, 运算对于第二个因子也是线性的。所以 $\otimes \in L(V^* \times W^*; F)$ 。

定义 10.2.5 由形如 $v^* \otimes w^*$ 的元素张成的向量空间, 称为向量空间 V^* 和 W^* 的张量积, 并记为

$V^* \otimes W^*$ ，即

$$V^* \otimes W^* = \text{span}\{v^* \otimes w^* : v^* \in V^*, w^* \in W^*\}$$

张量积 $V^* \otimes W^*$ 的元素是形如 $v^* \otimes w^*$ 元素的有限线性组合，一般不能写成单项的形式 $v^* \otimes w^*$ ，张量积 $V^* \otimes W^*$ 中能写成单项的元素称为可分解元素。

由于 $v^* \otimes w^* \in L(V \times W; F)$ ，所以 $V^* \otimes W^*$ 是 $L(V \times W; F)$ 的子空间。下面将证明 $V^* \otimes W^* = L(V \times W; F)$ 。

定理 10.2.1 $V^* \otimes W^* = L(V \times W; F)$ ，并且它是 nm 维向量空间。

证明 令 V 和 W 的基分别为 $\{a_\alpha\}, \{b_\beta\}$ ，它们在 V^* 和 W^* 中对偶基分别为 $\{a^{*\alpha}, \dots, a^{*n}\}$ ， $\{b^{*\beta}, \dots, b^{*m}\}$ 。已经知道 $V^* \otimes W^* \subset L(V \times W; F)$ ，所以只须证明： $V^* \otimes W^* \supset L(V \times W; F)$ 。为此，令

$$f \in L(V \times W; F), (v, w) \in V \times W, v = \sum_{\alpha=1}^n x^\alpha a_\alpha, w = \sum_{\beta=1}^m y^\beta b_\beta$$

则，

$$f(v, w) = \sum_{\alpha=1}^n \sum_{\beta=1}^m x^\alpha y^\beta f(a_\alpha, b_\beta), (a^{*\alpha} \otimes b^{*\beta})(v, w) = a^{*\alpha}(v) \cdot b^{*\beta}(w) = x^\alpha y^\beta$$

所以，

$$f(v, w) = \left(\sum_{\alpha=1}^n \sum_{\beta=1}^m f(a_\alpha, b_\beta) (a^{*\alpha} \otimes b^{*\beta}) \right) (v, w)$$

即，

$$f = \sum_{\alpha=1}^n \sum_{\beta=1}^m f(a_\alpha, b_\beta) (a^{*\alpha} \otimes b^{*\beta}) \in V^* \otimes W^*$$

因此，

$$V^* \otimes W^* \supset L(V \times W; F)$$

故有 $V^* \otimes W^* = L(V \times W; F)$ 。

下面证明 $L(V \times W; F)$ 是 nm 维向量空间。为此，只须证 $\{a^{*\alpha} \otimes b^{*\beta}\}$ 线性无关。令

$$f = \sum_{\alpha, \beta} c_{\alpha\beta} (a^{*\alpha} \otimes b^{*\beta}) = 0$$

则有

$$0 = f(a_k, b_l) = \sum_{\alpha, \beta} c_{\alpha\beta} (a^{*\alpha} \otimes b^{*\beta})(a_k, b_l) = \sum_{\alpha, \beta} c_{\alpha\beta} (a^{*\alpha}(a_k) \cdot b^{*\beta}(b_l)) = c_{kl}$$

所以 $\{a^{*\alpha} \otimes b^{*\beta}\}$ 线性无关。证毕。

由于向量空间 V 和 W 分别是 V^* 和 W^* 的对偶空间，同理可以定义张量积 $V \otimes W$ ，并且同样有

$$V \otimes W = L(V^* \times W^*; F)$$

命题 10.2.5 张量积 $V^* \otimes W^*$, $V \otimes W$ 是互为对偶的, 即 $V^* \otimes W^* = (V \otimes W)^*$ 。

证明 令

$$\langle v \otimes w, v^* \otimes w^* \rangle = \langle v, v^* \rangle \cdot \langle w, w^* \rangle \quad (10.2.14)$$

特别地,

$$\begin{aligned} \langle a_\alpha \otimes b_\beta, a^{*i} \otimes b^{*j} \rangle &= \langle a_\alpha, a^{*i} \rangle \cdot \langle b_\beta, b^{*j} \rangle \\ &= \delta_\alpha^i \cdot \delta_\beta^j = \delta_{\alpha\beta}^{ij} = \begin{cases} 1, (\alpha, \beta) = (i, j) \\ 0, (\alpha, \beta) \neq (i, j) \end{cases} \end{aligned} \quad (10.2.15)$$

所以, $\{a_\alpha \otimes b_\beta\}$ 和 $\{a^{*i} \otimes b^{*j}\}$ 是互为对偶的基。因此 $V^* \otimes W^* = (V \otimes W)^*$ 。

线性函数的张量积运算可以推广到任意的多重线性函数。

设 $f \in L(V_1, \dots, V_s; F)$, $g \in L(W_1, \dots, W_r; F)$, 它们的张量积定义如下:

$$f \otimes g(v_1, \dots, v_s, w_1, \dots, w_r) = f(v_1, \dots, v_s) \cdot g(w_1, \dots, w_r) \quad (10.2.16)$$

其中: $(v_1, \dots, v_s) \in V_1 \times \dots \times V_s, (w_1, \dots, w_r) \in (W_1 \times \dots \times W_r)$ 。

显然,

$$f \otimes g : V_1 \times \dots \times V_s \times W_1 \times \dots \times W_r \rightarrow F$$

是上 F -值的 $r+s$ 重线性函数

定理 10.2.2 张量积运算 \otimes 满足结合律, 即对任意的 $\phi \in L(V_1, \dots, V_s; F)$, $g \in L(W_1, \dots, W_r; F)$ 及 $\xi \in L(Z_1, \dots, Z_t; F)$ 有

$$(\phi \otimes \varphi) \otimes \xi = \phi \otimes (\varphi \otimes \xi) \quad (10.2.17)$$

证明 只对 $r = s = t = 1$ 的情形进行证明, 一般情形的证明是类似的。设 $v \in V_1, w \in W_1, z \in Z_1$, 则

$$(\phi \otimes \varphi) \otimes \xi(v, w, z) = (\phi \otimes \varphi)(v, w) \cdot \xi(z) = \phi(v) \cdot \varphi(w) \cdot \xi(z)$$

同理

$$\phi \otimes (\varphi \otimes \xi)(v, w, z) = \phi(v) \cdot \varphi(w) \cdot \xi(z)$$

证毕。

根据上述定理, 记号 $\phi \otimes \varphi \otimes \xi$ 是有意义的, 并称它是三个元素的张量积。

形如 $v \otimes w \otimes z$ 的元素所张成的向量空间记作 $V \otimes W \otimes Z$, 并称为向量空间 V, W, Z 的张量积。这里 v, w, z 分别看作 V^*, W^*, Z^* 上的 F -值线性函数。

同理, 若 V_1, \dots, V_s , 是 F 上的向量空间, 则可定义它们的张量积 $V_1 \otimes \dots \otimes V_s$ 。若 $\{a_1^{(i)}, \dots, a_{n_i}^{(i)}\}$ 是 V_i 的基, 则 $V_1 \otimes \dots \otimes V_s$ 的基是

$$\{a_{\alpha_1}^{(1)} \otimes \dots \otimes a_{\alpha_s}^{(s)} : 1 \leq \alpha_k \leq n_k; 1 \leq k \leq s\}$$

所以,

$$\dim(V_1 \otimes \dots \otimes V_s) = \dim V_1 \cdot \dots \cdot \dim V_s$$

容易证明：

$$V_1 \otimes \dots \otimes V_s = L(V_1^*, \dots, V_s^*; F) \tag{10.2.18}$$

$$V_1^* \otimes \dots \otimes V_s^* = L(V_1, \dots, V_s; F) \tag{10.2.20}$$

且

$$(V_1 \otimes \dots \otimes V_s)^* = V_1^* \otimes \dots \otimes V_s^* \tag{10.2.21}$$

张量积运算

$$\otimes : L(V_1, \dots, V_s; F) \times L(W_1, \dots, W_r; F) \rightarrow L(V_1, \dots, V_s, W_1, \dots, W_r; F)$$

是双线性映射。

多重线性映射

定义 10.2.4 假定 V, W, Z 都是域 F 上有限维的向量空间。映射 $f : V \times W \rightarrow Z$ 称为双线性的，如果它对于每一个变量都是线性的，即

$$\begin{cases} f(a^1 v_1 + a^2 v_2, w) = a^1 f(v_1, w) + a^2 f(v_2, w) \\ f(v, a^1 w_1 + a^2 w_2) = a^1 f(v, w_1) + a^2 f(v, w_2) \end{cases} \tag{10.2.22}$$

类似地，可以定义 r 重线性映射 $f : V_1 \times V_2 \times \dots \times V_r \rightarrow Z$ ，其中 V_1, V_2, \dots, V_r 都是 F 上的向量空间。当 $Z=F$ (F 被看作的一维向量空间) 时，上面定义的双线性映射和 r 重线性映射就是 (F -值) 双线性函数和 r 重线性函数。

从 $V_1 \times V_2 \times \dots \times V_r$ 到 Z 的全体 r 重线性映射的集合记作 $L(V_1, V_2, \dots, V_r; Z)$ 。如果对于任意的 $f, g \in L(V_1, V_2, \dots, V_r; Z), \alpha \in F$ ，令

$$\begin{cases} (f + g)(v_1, v_2, \dots, v_r) = f(v_1, v_2, \dots, v_r) + g(v_1, v_2, \dots, v_r) \\ (\alpha f) = \alpha f(v_1, v_2, \dots, v_r) \end{cases} \tag{10.2.23}$$

则 $f + g, \alpha f \in L(V_1, V_2, \dots, V_r; Z)$ 。显然，集合 $L(V_1, V_2, \dots, V_r; Z)$ 关于这两种运算成为域 F 上的向量空间。

例 10.2.2 令 $V = R^n, W = R^m, Z = R^p$ ，对 $x = (x^i) \in R^n, y = (y^j) \in R^m$ ，定义：

$$f^k(x, y) = \sum_{i=1}^m \sum_{j=1}^n x^i y^j f_{ij}^k, \quad k = 1, 2, \dots, p$$

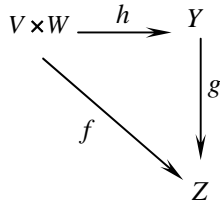
其中 $f_{ij}^k, 1 \leq i \leq m, 1 \leq j \leq n, 1 \leq k \leq p$ 是给定的实数。则不难验证： $f \in L(R^n \times R^m; R^p)$ 。

考虑空间 $L(V \times W; Z)$ 的结构。首先将 $L(V \times W; Z)$ 上的双线性映射转化为线性映射。确切地说，对于给定的向量空间 V 和 W ，构造一个只依赖于 V 和 W 的向量空间 Y 及双线性映射 $h : V \times W \rightarrow Y$ ，

使得对于任意的双线性映射 $f \in L(V \times W; Z)$ ，存在唯一的线性映射 g ，适合

$$f = g \circ h: V \times W \rightarrow Z \tag{10.2.24}$$

或者说有下面的交换图表：



下面证明所要构造的空间 Y 就是 V 和 W 的张量积。令 $Y = V \otimes W$ ，则 $\otimes: (v, w) \mapsto v \otimes w$ 是 $V \times W$ 到 Y 的双线性映射，即 $\otimes \in L(V \times W; Y)$ 。

定理 10.2.2 设 $h: V \times W \rightarrow V \otimes W$ 是张量积 \otimes 给出的双线性映射，即对于 $v \in V, w \in W$ 有

$$h(v, w) = v \otimes w \tag{10.2.25}$$

则对任意的双线性映射 $f \in L(V \times W; Z)$ ，存在唯一的线性映射 $g: V \otimes W \rightarrow Z$ ，使得

$$f = g \circ h: V \times W \rightarrow Z \tag{10.2.26}$$

证明 定义线性映射 $g: V \otimes W \rightarrow Z$ ，使得它在基底上的作用是

$$g(a_\alpha \otimes b_\beta) = f(a_\alpha, b_\beta) \tag{10.2.27}$$

显然， g 是唯一确定的。若 $v = \sum_\alpha v^\alpha a_\alpha \in V, w = \sum_\beta w^\beta b_\beta \in W$ ，则

$$\begin{aligned}
 g \circ h(v, w) &= g(v \otimes w) = \sum_{\alpha, \beta} v^\alpha w^\beta g(a_\alpha \otimes b_\beta) \\
 &= \sum_{\alpha, \beta} v^\alpha w^\beta f(a_\alpha, b_\beta) = f\left(\sum_\alpha v^\alpha a_\alpha, \sum_\beta w^\beta b_\beta\right) = f(v, w)
 \end{aligned}$$

所以 $f = g \circ h$ 。证毕。

例 10.2.3 令 $f \in L(R^n \times R^m; R^p)$ ， $x = \sum_i x^i e_i \in R^n, y = \sum_\alpha y^\alpha e_\alpha \in R^m$ ，则

$$f(x, y) = f\left(\sum_i x^i e_i, \sum_\alpha y^\alpha e_\alpha\right) = \sum_{i, \alpha} x^i y^\alpha f_{i\alpha} = \sum_k \sum_{i, \alpha} x^i y^\alpha f_{i\alpha}^k e_k$$

由于

$$h(x, y) = \sum_{i, \alpha} x^i y^\alpha e_i \otimes e_\alpha,$$

$$f(e_i, e_\alpha) = f_{i\alpha} = \begin{pmatrix} f_{i\alpha}^1 \\ \vdots \\ f_{i\alpha}^p \end{pmatrix} \in R_p, i = 1, \dots, m; \alpha = 1, \dots, n,$$

$$g(h(x, y)) = g\left(\sum_{i,\alpha} x^i y^\alpha e_i \otimes e_\alpha\right) = \sum_{i,\alpha} x^i y^\alpha g(e_i \otimes e_\alpha) = \sum_{i,\alpha} x^i y^\alpha f_{i\alpha} = \sum_k \sum_{i,\alpha} x^i y^\alpha f_{i\alpha}^k e_k,$$

所以

$$f = g \circ h$$

推论 10.2.1 向量空间 $L(V \times W; Z)$ 与 $L(V \otimes W; Z)$ 是同构的。

证明 定义映射

$$\xi: L(V \otimes W; Z) \rightarrow L(V \times W; Z)$$

使得

$$\xi(g) = g \circ h, \quad g \in L(V \otimes W; Z) \quad (10.2.28)$$

其中 h 由式(10.2.25)定义。根据定理 10.2.2, 映射 ξ 是线性的一对一满映射。因此, 它是这两个向量空间之间的同构映射。

对于 s 重线性映射, 同样也有与上面类似的结论:

定理 10.2.3 设 $h: V_1 \times \dots \times V_s \rightarrow V_1 \otimes \dots \otimes V_s$ 是张量积 \otimes 所定义的 s 重线性映射, 即

$$h(v_1, \dots, v_s) = v_1 \otimes \dots \otimes v_s \quad (10.2.29)$$

则对任意 $f \in L(V_1, \dots, V_s; Z)$ 存在唯一的线性映射 $g \in L(V_1 \otimes \dots \otimes V_s, Z)$, 使得

$$f = g \circ h: V_1 \times \dots \times V_s \rightarrow Z \quad (10.2.30)$$

从而 $L(V_1, \dots, V_s; Z)$ 与 $L(V_1 \otimes \dots \otimes V_s, Z)$ 同构。

10.3 张 量

10.3.1 张量与代数运算(Tensors and Algebraic Operation)

张量的定义

在上节讨论了张量积的一般概念, 在应用中通常情况是同一个向量空间与自身以及与其对偶空间的张量积, 这种特殊的张量积称为张量。张量的准确定义如下:

定义 10.3.1 设 V 是 F 上的 n 维向量空间，对偶空间是 V^* ，张量积

$$V_s^r = \underbrace{V \otimes \dots \otimes V}_r \otimes \underbrace{V^* \otimes \dots \otimes V^*}_s \tag{10.3.1}$$

中的元素称为 (r, s) 型张量，其中 r 是张量的逆变阶数， s 是协变阶数。

特别地， V_0^r 的元素称为 r 阶逆变张量， V_s^0 的元素称为 s 阶协变张量。同时，约定 $V_0^0 = F$ （它的元素通常又称为零阶张量）， $V_0^1 = V$ 的元素称为逆变向量， $V_1^0 = V^*$ 的元素称为协变向量。

注意：在应用时，张量积(10.3.1)中的因子 V 与 V^* 可能是交替出现的。将它们排成(10.3.1)的形式，只是为了记号的方便。

根据 10.2 节的讨论， $\dim(V_s^r) = n^{r+s}$ ，并且

$$V_s^r = L(\underbrace{V^*, \dots, V^*}_r, \underbrace{V, \dots, V}_s; F)$$

也就是说， (r, s) 型张量是定义在

$$\underbrace{V^* \times \dots \times V^*}_r \times \underbrace{V \times \dots \times V}_s$$

上的 F -值 $(r+s)$ 重线性函数。

设 $\{e_i : 1 \leq i \leq n\}$, $\{e^{*i} : 1 \leq i \leq n\}$ 分别是 V 和 V^* 中彼此对偶的基，则空间 V_s^r 的基是

$$e_{i_1} \otimes \dots \otimes e_{i_r} \otimes e^{*k_1} \otimes \dots \otimes e^{*k_s}, \quad (1 \leq i_1, \dots, i_r, k_1, \dots, k_s \leq n) \tag{10.3.2}$$

因此， (r, s) 型张量可以唯一地表示成

$$x = \sum_{\substack{i_1, \dots, i_r \\ k_1, \dots, k_s}} x_{k_1, \dots, k_s}^{i_1, \dots, i_r} e_{i_1} \otimes \dots \otimes e_{i_r} \otimes e^{*k_1} \otimes \dots \otimes e^{*k_s} \tag{10.3.3}$$

其中， $x_{k_1, \dots, k_s}^{i_1, \dots, i_r}$ 称为张量在基(10.3.2)下的分量。很明显，

$$\begin{aligned} x_{k_1, \dots, k_s}^{i_1, \dots, i_r} &= x(e^{*i_1} \otimes \dots \otimes e^{*i_r} \otimes e_{k_1} \otimes \dots \otimes e_{k_s}) \\ &= \langle e^{*i_1} \otimes \dots \otimes e^{*i_r} \otimes e_{k_1} \otimes \dots \otimes e_{k_s}, x \rangle \end{aligned} \tag{10.3.4}$$

在处理张量时，常用使用 Einstein 的求和约定，在一个单项表达式中出现重复的上、下指标，表示该式关于这个指标在它的取值范围内求和，而略去和号不写，例如，在(10.3.3)式中有 $(r+s)$ 重和号，采用这个约定则可简写成：

$$x = x_{k_1, \dots, k_s}^{i_1, \dots, i_r} e_{i_1} \otimes \dots \otimes e_{i_r} \otimes e^{*k_1} \otimes \dots \otimes e^{*k_s} \tag{10.3.5}$$

在 n 维欧氏空间 R^n 中，如果默认基，则 (r, s) 型张量 x 简记为 $x = (x_{k_1, \dots, k_s}^{i_1, \dots, i_r})$ 或者更简洁地记为 $x = x_{k_1, \dots, k_s}^{i_1, \dots, i_r}$ 。

张量变换

当向量空间 V 的基改变时, 张量的分量也是按一定规律变化的。

设 $\{\bar{e}_i : 1 \leq i \leq n\}$ 是 V 的另一个基, 相应的对偶基是 $\{\bar{e}^{*i} : 1 \leq i \leq n\}$ 。令它与原基之间的关系为

$$\bar{e}_i = a_i^j e_j \quad (10.3.6)$$

其中 $a = (a_i^j)$ 是行列式不为零的 n 阶方阵, 则对偶基之间的关系必为

$$\bar{e}^{*i} = \beta_j^i e^{*j} \quad (10.3.7)$$

其中 $\beta = (\beta_j^i)$ 是 $a = (a_i^j)$ 逆矩阵, 即

$$\alpha_i^j \beta_j^k = \beta_i^j \alpha_j^k = \delta_i^k \quad (10.3.8)$$

若张量 x 在新基下的分量记为 $\bar{x}_{k_1, \dots, k_s}^{i_1, \dots, i_r}$, 则

$$\begin{aligned} x &= \bar{x}_{k_1, \dots, k_s}^{i_1, \dots, i_r} \bar{e}_{i_1} \otimes \dots \otimes \bar{e}_{i_r} \otimes \bar{e}^{*k_1} \otimes \dots \otimes \bar{e}^{*k_s} \\ &= \bar{x}_{k_1, \dots, k_s}^{i_1, \dots, i_r} \alpha_{i_1}^{j_1} e_{j_1} \otimes \dots \otimes \alpha_{i_r}^{j_r} e_{j_r} \otimes \beta_{k_1}^{l_1} e^{*l_1} \otimes \dots \otimes \beta_{k_s}^{l_s} e^{*l_s} \\ &= \bar{x}_{k_1, \dots, k_s}^{i_1, \dots, i_r} \alpha_{i_1}^{j_1} \dots \alpha_{i_r}^{j_r} \beta_{l_1}^{k_1} \dots \beta_{l_s}^{k_s} e_{j_1} \otimes \dots \otimes e_{j_r} \otimes e^{*l_1} \otimes \dots \otimes e^{*l_s} \end{aligned}$$

所以, 张量 x 的分量在新旧基下有下述关系式:

$$x_{i_1, \dots, i_s}^{j_1, \dots, j_r} = \bar{x}_{k_1, \dots, k_s}^{i_1, \dots, i_r} \alpha_{i_1}^{j_1} \dots \alpha_{i_r}^{j_r} \beta_{l_1}^{k_1} \dots \beta_{l_s}^{k_s} \quad (10.3.9)$$

在经典的张量分析中, 变换公式(10.3.9)是定义张量的根据。

张量运算

(r, s) 型张量的空间是向量空间 V_s^r , 所以同类型的张量相加与向量的加法相同; 张量的数乘也与向量的数乘相同。即:

(1) $x, y \in V_s^r$, 则它们的和定义为:

$$(x+y)(v^{*1}, \dots, v^{*r}, v_1, \dots, v_s) = x(v^{*1}, \dots, v^{*r}, v_1, \dots, v_s) + y(v^{*1}, \dots, v^{*r}, v_1, \dots, v_s) \quad (10.3.11)$$

在取定基后, $x+y$ 的分量是 x 与 y 对应分量的和:

$$(x+y)_{k_1, \dots, k_s}^{i_1, \dots, i_r} = x_{k_1, \dots, k_s}^{i_1, \dots, i_r} + y_{k_1, \dots, k_s}^{i_1, \dots, i_r} \quad (10.3.12)$$

(2) $\alpha \in F$, $x \in V_s^r$, 则它们的乘积定义为:

$$(\alpha \cdot x)(v^{*1}, \dots, v^{*r}, v_1, \dots, v_s) = \alpha x(v^{*1}, \dots, v^{*r}, v_1, \dots, v_s) \quad (10.3.13)$$

在取定基后， αx 的分量是 x 对应分量与数 α 的积：

$$(\alpha x)_{k_1 \dots k_s}^{i_1 \dots i_r} = \alpha x_{k_1 \dots k_s}^{i_1 \dots i_r} \quad (10.3.14)$$

张量除了这两种基本运算外，还有两种特别重要的运算：乘法和缩并。

定义 10.3.2 设 x 是 (r_1, s_1) 型张量， y 是 (r_2, s_2) 型张量，它们的乘积 $x \otimes y$ 是 $(r_1 + r_2, s_1 + s_2)$ 型张量，其定义如下：

$$\begin{aligned} & x \otimes y(v_1^{*1}, \dots, v_{r_1+r_2}^{*r_1+r_2}, v_1, \dots, v_{s_1+s_2}) \\ &= x(v_1^{*1}, \dots, v_{r_1}^{*r_1}, v_1, \dots, v_{s_1}) \cdot y(v_1^{*r_1+1}, \dots, v_{r_1+r_2}^{*r_1+r_2}, v_{s_1+1}, \dots, v_{s_1+s_2}) \end{aligned} \quad (10.3.15)$$

在取定基后， $x \otimes y$ 的分量是 x 和 y 的分量的乘积，即

$$(x \otimes y)_{k_1 \dots k_{s_1+s_2}}^{i_1 \dots i_{r_1+r_2}} = x_{k_1 \dots k_{s_1}}^{i_1 \dots i_{r_1}} y_{k_{s_1+1} \dots k_{s_1+s_2}}^{i_{r_1+1} \dots i_{r_1+r_2}} \quad (10.3.16)$$

张量乘法满足分配律和结合律。

定义 10.3.3 取两个指标 λ, μ ，使 $1 \leq \lambda \leq r, 1 \leq \mu \leq s$ 。对于任意一个可分解的 (r, s) 型张量：

$$x = v_1 \otimes \dots \otimes v_r \otimes v_1^{*1} \otimes \dots \otimes v_s^{*s} \in V_s^r, \quad (10.3.12)$$

定义

$$C_{\lambda\mu}(x) = D(v_\lambda, v^{*\mu}) v_1 \otimes \dots \otimes v_\lambda \otimes \dots \otimes v_r \otimes v_1^{*1} \otimes \dots \otimes v^{*\mu} \otimes \dots \otimes v_s^{*s}, \quad (10.3.13)$$

其中“ $D(v_\lambda, v^{*\mu})$ ”记号表示去掉其中的两个因子，则 $C_{\lambda\mu}(x)$ 是一个 $(r-1, s-1)$ 型张量，即 $C_{\lambda\mu}(x) \in V_{s-1}^{r-1}$ ，将映射 $x \mapsto C_{\lambda\mu}(x)$ 作线性扩张所得到的线性映射 $C_{\lambda\mu} : V_s^r \rightarrow V_{s-1}^{r-1}$ 叫做张量的缩并。

若张量 x 用分量表示为

$$x = x_{k_1 \dots k_s}^{i_1 \dots i_r} e_{i_1} \otimes \dots \otimes e_{i_r} \otimes e^{*k_1} \otimes \dots \otimes e^{*k_s} \quad (10.3.14)$$

根据缩并的定义得

$$\begin{aligned} C_{\lambda\mu}(x) &= x_{k_1 \dots k_s}^{i_1 \dots i_r} C_{\lambda\mu}(e_{i_1} \otimes \dots \otimes e_{i_r} \otimes e^{*k_1} \otimes \dots \otimes e^{*k_s}) \\ &= x_{k_1 \dots k_{\mu-1} j k_{\mu+1} \dots k_{s-1}}^{i_1 \dots i_{\lambda-1} j i_{\lambda+1} \dots i_{r-1}} e_{i_1} \otimes \dots \otimes e_{i_{\lambda-1}} \otimes e^{*k_1} \otimes \dots \otimes e^{*k_{s-1}} \end{aligned} \quad (10.3.15)$$

因此从分量来看，缩并 $C_{\lambda\mu}$ 就是关于第 λ 个上指标与第 μ 个下指标的对等求和。缩并降低了张量的阶数，它是一个非常基本的运算。

在取定基后，张量 $x = x_{k_1 \dots k_s}^{i_1 \dots i_r}$ 缩并 $C_{\lambda\mu}(x)$ 的分量是：

$$(C_{\lambda\mu}(x))_{k_1 \dots k_{s-1}}^{i_1 \dots i_{r-1}} = \sum_{j=1}^n x_{k_1 \dots k_{\mu-1} j k_{\mu+1} \dots k_{s-1}}^{i_1 \dots i_{\lambda-1} j i_{\lambda+1} \dots i_{r-1}} \quad (10.3.16)$$

例：设 $x = \xi_i^j e_i \otimes e^{*j}$ 是 $(1, 1)$ 型张量，则 x 的缩并就是求矩阵 (ξ_i^j) 的迹 $\sum_{i=1}^n \xi_i^i$ ，所得的是一个与坐

标系选取无关的标量。

利用张量乘积和张量缩并，可得到张量的另一种重要运算，即张量内乘。张量内乘给出了张量内积，它是向量内积在张量中的推广。

例如：两个张量 $x = x_{ij}^{pq}, y = y_{st}^{uv}$ ，它们的乘积是 $(x \otimes y)_{ijst}^{pquv} = x_{ij}^{pq} \cdot y_{st}^{uv}$ ，对 $x \otimes y$ 作缩并运算 C_{13} 得到 $z_{ij}^{quv} = (C_{13}(x \otimes y))_{ij}^{quv} = x_{ij}^{kq} y_{kt}^{uv}$ ，再对 z_{ij}^{quv} 作缩并运算 C_{21} ， $a_{jt}^{qv} = (C_{21}z)_{jt}^{qv} = x_{ij}^{kq} y_{kt}^{lv}$ ，这一系列的运算所得到的结果十分类似于两个向量的内积。这样就导致张量内乘的下述定义：

若一个张量的分量下标中的指标与另一张量的分量上标中的指标有若干个相同，则表示这若干个指标遍历范围 $\{1, 2, \dots, n\}$ 求和，这种演算的结果称为两个张量的内积。显然，两个张量的内积仍是一个张量，其阶数是两个张量的阶数之和减去哑指标数目的两倍。

张量代数

令

$$T^r(V) = V_0^r = \underbrace{V \otimes \dots \otimes V}_r$$

考虑直和 $T(V) = \bigoplus_{r \geq 0} T^r(V)$ ，其元素 x 可表为下述的形式和：

$$x = \bigoplus_{r \geq 0} x^r, x^r \in T^r(V) \tag{10.3.16}$$

在和式中，除有限多项外其余各项都是零。这样， $T(V)$ 是无穷维向量空间。 $T(V)$ 中的乘法可通过张量乘法分配律定义，并且向量空间 $T(V)$ 关于这种乘法成为一个代数，称为向量空间 V 的张量代数。

同理， V^* 的张量代数是 $T(V^*) = \bigoplus_{r \geq 0} V_r^{*0}$ 。

正如上节所讨论的那样，向量空间 $T^r(V^*)$ 和 $T^r(V)$ 是彼此对偶的，它们的配合是

$$\langle v_1 \otimes \dots \otimes v_r, v^{*1} \otimes \dots \otimes v^{*r} \rangle = \langle v_1, v^{*1} \rangle \dots \langle v_r, v^{*r} \rangle \tag{10.3.17}$$

其中 $v_i \in V, v^{*i} \in V^*$ 。

10.3.2 对称与反对称张量 (Symmetric and Skew-symmetric Tensors)

对称与反对称张量

自然数 $\{1, \dots, r\}$ 的置换群记为 $\varphi(r)$ ，其中的任意一个元素(置换) σ 决定了向量空间 $T^r(V)$ 的一个自同态：设 $x \in T^r(V)$ ，则定义

$$\alpha x(v^{*1}, \dots, v^{*r}) = x(v^{*\sigma(1)}, \dots, v^{*\sigma(r)}) \tag{10.3.18}$$

其中 $v^{*j} \in V^*$ 。如果 $x = v_1 \otimes \cdots \otimes v_r$ ，则

$$\begin{aligned} \sigma x(v^{*1}, v^{*2}, \dots, v^{*r}) &= \sigma v_1 \otimes \cdots \otimes v_r(v^{*1}, \dots, v^{*r}) \\ &= \prod_{j=1}^r \langle v_j, v^{*\sigma(j)} \rangle = \prod_{j=1}^r \langle v_{\sigma^{-1}(j)}, v^{*j} \rangle = v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(r)}(v^{*1}, \dots, v^{*r}) \\ \sigma x &= v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(r)} \end{aligned} \tag{10.3.19}$$

其中 σ^{-1} 表示 σ 的逆元素。

定义 10.3.3 设 $x \in T^r(V)$ 。若对任意的 $\sigma \in \varphi(r)$ ，有

$$\sigma x = x \tag{10.3.20}$$

则称 x 是对称的 r 阶逆变张量；若对任意的 $\sigma \in \varphi(r)$ ，有

$$\sigma x = \text{sign } \sigma \cdot x \tag{10.3.21}$$

则称 x 是反对称 r 阶逆变张量。其中 $\text{sgn } \sigma$ 表示置换 σ 的符号：

$$\text{sign } \sigma = \begin{cases} 1, & \sigma \text{ 是偶置换} \\ -1, & \sigma \text{ 是奇置换} \end{cases} \tag{10.3.22}$$

定理 10.3.1 设 $x \in T^r(V)$ ，(1) x 为对称张量的充要条件是它的分量关于各指标是对称的；(2) x 为反对称的充要条件是它的分量关于各指标是反对称的。

证明 设 V 的基底是 $\{e_1, \dots, e_n\}$ 。若 x 是对称的，则对任意的 $\sigma \in \varphi(r)$ 有

$$x^{i_1 \cdots i_r} = x(e^{*i_1}, \dots, e^{*i_r}) = \sigma x(e^{*i_1}, \dots, e^{*i_r}) = x(e^{*i_{\sigma(1)}}, \dots, e^{*i_{\sigma(r)}}) = x^{i_{\sigma(1)} \cdots i_{\sigma(r)}} \tag{10.3.23}$$

反之亦然。

若 x 是反对称的，则对任意的 $\sigma \in \varphi(r)$ 有

$$x^{i_1 \cdots i_r} = x(e^{*i_1}, \dots, e^{*i_r}) = \text{sign } \sigma \cdot \sigma x(e^{*i_1}, \dots, e^{*i_r}) = \text{sign } \sigma \cdot x^{i_{\sigma(1)} \cdots i_{\sigma(r)}} \tag{10.3.24}$$

反之亦然。证毕。

张量的对称化和反对称化

对任意 $x \in T^r(V)$ ，令

$$S_r(x) = \frac{1}{r!} \sum_{\sigma \in \varphi(r)} \sigma x \tag{10.3.25}$$

$$A_r(x) = \frac{1}{r!} \sum_{\sigma \in \varphi(r)} \text{sign } \sigma \cdot \sigma x \tag{10.3.26}$$

显然， $S_r(x), A_r(x) \in T^r(V)$ ，并且映射 $S_r(x), A_r(x): T^r(V) \rightarrow T^r(V)$ 都是 $T^r(V)$ 的自同态。

下面证明 $S_r(x), A_r(x)$ 分别是对称张量和反对称张量。

设 $x \in T^r(V)$ ，则对任意的 $\tau \in \varphi(r)$ 有

$$\tau(S_r(x)) = \frac{1}{r!} \sum_{\sigma \in \varphi(r)} \tau(\sigma(x)) = S_r(x) \quad (10.3.27)$$

$$\begin{aligned} (\tau A_r(x)) &= \frac{1}{r!} \sum_{\sigma \in \varphi(r)} \text{sign } \sigma \cdot \tau(\sigma(x)) \\ &= \text{sign}(\tau) \frac{1}{r!} \sum_{\sigma \in \varphi(r)} \text{sign}(\tau \circ \sigma) \cdot \tau \circ \sigma(x) \\ &= \text{sign}(\tau) A_r(x) \end{aligned} \quad (10.3.28)$$

所以， $S_r(x), A_r(x)$ 分别是对称张量和反对称张量。

例 10.3.1 (1) $x \in T^2(R^n)$ ，考虑 $S_2(x), A_2(x)$ ；(2) $x \in T^3(R^n)$ ，考虑 $S_3(x), A_3(x)$ 。

$$(1) \quad x = \sum_{1 \leq i, j \leq n} x^{ij} e_i \otimes e_j$$

$$\sigma_1 = (ij) \rightarrow (ij), \sigma_2 = (ij) \rightarrow (ji); \text{sign} \sigma_1 = 1, \text{sign} \sigma_2 = -1$$

$$S_2(x) = \frac{1}{2}(\sigma_1 x + \sigma_2 x) = \frac{1}{2} \left(\sum_{1 \leq i, j \leq n} x^{\sigma_1(ij)} e_i \otimes e_j + \sum_{1 \leq i, j \leq n} x^{\sigma_2(ij)} e_i \otimes e_j \right) = \sum_{1 \leq i, j \leq n} \frac{x^{ij} + x^{ji}}{2} e_i \otimes e_j$$

$$A_2(x) = \frac{1}{2}(\text{sign} \sigma_1 \cdot \sigma_1 x + \text{sign} \sigma_2 \cdot \sigma_2 x)$$

$$= \frac{1}{2} \left(\sum_{1 \leq i, j \leq n} x^{\sigma_1(ij)} e_i \otimes e_j - \sum_{1 \leq i, j \leq n} x^{\sigma_2(ij)} e_i \otimes e_j \right) = \sum_{1 \leq i, j \leq n} \frac{x^{ij} - x^{ji}}{2} e_i \otimes e_j$$

$$(2) \quad x = \sum_{1 \leq i, j, k \leq n} x^{ijk} e_i \otimes e_j \otimes e_k$$

$$\sigma_1 = (ijk) \rightarrow (ijk), \text{sign} \sigma_1 = 1; \sigma_2 = (ijk) \rightarrow (jik), \text{sign} \sigma_2 = -1; \sigma_3 = (ijk) \rightarrow (jki), \text{sign} \sigma_3 = 1;$$

$$\sigma_4 = (ijk) \rightarrow (kji), \text{sign} \sigma_4 = -1; \sigma_5 = (ijk) \rightarrow (kij), \text{sign} \sigma_5 = 1; \sigma_6 = (ijk) \rightarrow (ikj), \text{sign} \sigma_6 = -1$$

$$S_3(x) = \frac{1}{6} \sum_{l=1}^6 \sigma_l x = \frac{1}{6} \sum_{1 \leq i, j, k \leq n} \sum_{l=1}^6 x^{\sigma_l(ijk)} e_i \otimes e_j \otimes e_k$$

$$A_3(x) = \frac{1}{6} \sum_{l=1}^6 \sigma_l x = \frac{1}{6} \sum_{1 \leq i, j, k \leq n} \sum_{l=1}^6 \text{sign} \sigma_l \cdot x^{\sigma_l(ijk)} e_i \otimes e_j \otimes e_k$$

定义 10.3.4 S_r, A_r 分别称为 r 阶逆变张量的对称化算子和反对称化算子。

将全体 r 阶的对称逆变张量的集合记作 $P^r(V)$ ，全体 r 阶的反对称逆变张量的集合记为 $A^r(V)$ 。

定理 10.3.2 (1) $P^r(V) = S_r(T^r(V)), A^r(V) = A_r(T^r(V))$ ；(2) $P^r(V)$ 和 $A^r(V)$ 都是 $T^r(V)$ 的线性子空间。

证明是容易的。由于对称张量在对称化算子作用下不变，反对称张量在反对称化算子作用下不

变, 所以(1)成立。由于置换 σ 在 $T^r(V)$ 上的作用是自同态, 所以对称张量的和仍是对称的, 反对称张量的和仍是反对称的, 因此 $P^r(V)$ 和 $A^r(V)$ 都是 $T^r(V)$ 的线性子空间。

上面关于对称张量和反对称张量的讨论同样可用于协变张量。全体 r 阶的对称协变张量的集合记作 $P^r(V^*)$, 全体 r 阶的反对称协变张量的集合记作 $A^r(V^*)$ 。

10.4 外代数(Grassmann 代数)

10.4.1 外积运算(Exterior Product)

外向量的外积

反对称的 r 阶逆变张量也称为 r 次外向量, 空间 $A^r(V)$ 称为 V 上的 r 次外向量空间。为陈述方便, 我们还约定 $A^1(V) = V, A^0(V) = F$ 。

外向量有一种重要的外积运算, 即两个外向量相乘得到另一个外向量, 确切定义如下:

定义 10.4.1 设 ξ 是 k 次外向量, η 是 l 次外向量, 令

$$\xi \wedge \eta = A_{k+l}(\xi \otimes \eta) \quad (10.4.1)$$

其中 A_{k+l} 是反对称化算子, 则 $\xi \wedge \eta$ 是 $(k+l)$ 次外向量, 称为外向量 ξ 和 η 的外积。

例 10.4.1 考虑两个一次外向量 $\xi, \eta \in R^n$ 的外积 $\xi \wedge \eta$: 令

$$\begin{aligned} \xi &= \sum_{i=1}^n x^i e_i, \eta = \sum_{j=1}^n y^j e_j, \text{ 根据例 10.3.1, 我们有} \\ \xi \wedge \eta &= A_2(\xi \otimes \eta) = A_2 \left(\sum_{i,j=1}^n x^i y^j e_i \otimes e_j \right) = \sum_{i,j=1}^n \frac{x^i y^j - x^j y^i}{2} e_i \otimes e_j \end{aligned}$$

这是外积 $\xi \wedge \eta$ 在张量空间 $T^2(R^n) = R^n \otimes R^n$ 中的表达。由于 2 次外向量空间 $A^2(R^n)$ 是 $T^2(R^n)$ 的子空间, 我们的目的是需要在 $T^2(R^n)$ 上来表达外积 $\xi \wedge \eta$ 。这样就需要研究向量空间 $A^2(R^n)$ 的基, 为此我们先讨论外积运算的一般规律。

定理 10.4.1 设 $\xi, \xi_1, \xi_2 \in A^k(V), \eta, \eta_1, \eta_2 \in A^l(V), \mu \in A^h(V)$, 外积满足下列运算律:

(1) 分配律: $(\xi_1 + \xi_2) \wedge \eta = \xi_1 \wedge \eta + \xi_2 \wedge \eta$;

$$\xi \wedge (\eta_1 + \eta_2) = \xi \wedge \eta_1 + \xi \wedge \eta_2;$$

(2) 反交换律: $\xi \wedge \eta = (-1)^{kl} \eta \wedge \xi$

(3) 结合律: $(\xi \wedge \eta) \wedge \mu = \eta \wedge (\xi \wedge \mu)$

证明

(1) 分配律是张量积和反对称化算子的线性性质的推论，因此是明显的。

(2) 因 $\xi \wedge \eta$ 是反对称张量，所以对任意的 $\tau \in \varphi(k+l)$ 有

$$\tau(\xi \wedge \eta) = \text{sign } \tau \cdot \xi \wedge \eta$$

取

$$\tau = \begin{pmatrix} 1 \dots \dots k & k+1 \dots \dots k+l \\ l+1 \dots l+k & 1 \dots \dots l \end{pmatrix}$$

则 $\text{sgn } \tau = (-1)^{kl}$ ，所以对任意的 v^{*1}, \dots, v^{*k+l} 有

$$\begin{aligned} \xi \wedge \eta(v^{*1}, \dots, v^{*k+l}) &= (-1)^{kl} \xi \wedge \eta(v^{*\tau(1)}, \dots, v^{*\tau(k+l)}) \\ &= \frac{(-1)^{kl}}{(k+l)!} \sum_{\sigma \in \varphi(k+l)} \text{sign } \sigma \cdot \xi(v^{*\sigma(1)}, \dots, v^{*\sigma(k)}) \cdot \eta(v^{*\sigma(k+1)}, \dots, v^{*\sigma(k+l)}) \\ &= \frac{(-1)^{kl}}{(k+l)!} \sum_{\sigma \in \varphi(k+l)} \text{sign } \sigma \cdot \eta(v^{*\sigma(k+1)}, \dots, v^{*\sigma(k+l)}) \cdot \xi(v^{*\sigma(1)}, \dots, v^{*\sigma(k)}) \\ &= \frac{(-1)^{kl}}{(k+l)!} \sum_{\sigma \in \varphi(k+l)} \text{sign } \sigma \cdot \eta(v^{*\sigma(1)}, \dots, v^{*\sigma(l)}) \cdot \xi(v^{*\sigma(l+1)}, \dots, v^{*\sigma(k+l)}) \\ &= (-1)^{kl} \eta \wedge \xi(v^{*1}, \dots, v^{*k+l}) \end{aligned}$$

(3) $(\xi \wedge \eta) \wedge \mu = A_{k+l+h}((\eta \otimes \xi) \otimes \mu) = A_{k+l+h}(\eta \otimes (\xi \otimes \mu)) = \eta \wedge (\xi \wedge \mu)$ 。 证毕。

注 根据反交换律，有

$$\xi, \eta \in A^1(V), \quad \xi \wedge \eta = -\eta \wedge \xi, \quad \xi \wedge \xi = \eta \wedge \eta = 0 \tag{10.4.3}$$

设 $\{e_1, \dots, e_n\}$ 是 V 的一个基，根据结合律的证明，有

$$e_{i_1} \wedge \dots \wedge e_{i_r} = A_r(e_{i_1} \otimes \dots \otimes e_{i_r}), \quad 1 \leq i_1, \dots, i_r \leq n \tag{10.4.4}$$

因此，仅当 i_1, \dots, i_r ，互不相同，外向量 $e_{i_1} \wedge \dots \wedge e_{i_r}$ 才不是零。特别地，当 $r > n$ 时，指标 i_1, \dots, i_r 必有重复，故相应的外向量必然是零。

外向量空间的基

设 ξ 是 r 次外向量，在张量空间 $T^r(V)$ 分量可表成

$$\xi = \xi^{i_1 \dots i_r} e_{i_1} \otimes \dots \otimes e_{i_r} \tag{10.4.5}$$

根据反对称化算子的线性性质，有

$$\xi = A_r \xi = \xi^{i_1 \dots i_r} A_r(e_{i_1} \otimes \dots \otimes e_{i_r}) = \xi^{i_1 \dots i_r} e_{i_1} \wedge \dots \wedge e_{i_r}$$

因此，次数大于 n 的外向量都是零，即

$$A^r(V) = 0 \quad (r > n) \tag{10.4.6}$$

设 $r \leq n$ ，由定理 10.3.1， ξ 的分量 $\xi^{i_1 \dots i_r}$ 关于上指标是反对称的，所以 ξ 可表成

$$\xi = r! \sum_{i_1 < \dots < i_r} \xi^{i_1 \dots i_r} e_{i_1} \wedge \dots \wedge e_{i_r} \tag{10.4.7}$$

例 10.4.2 在例 10.4.1 中， $\xi \wedge \eta = A_2 \left(\sum_{i,j=1}^n \frac{x^i y^j - x^j y^i}{2} e_i \otimes e_j \right)$

$$= \sum_{i,j=1}^n \frac{x^i y^j - x^j y^i}{2} A_2(e_i \otimes e_j) = \sum_{1 \leq i, j \leq n} (x^i y^j - x^j y^i) e_i \wedge e_j$$

定理 10.4.2 当 $r \leq n$ 时， $\{e_{i_1} \wedge \dots \wedge e_{i_r} \mid 1 \leq i_1 < \dots < i_r \leq n\}$ 是外向量空间 $A^r(V)$ 的基。

证明 只要证明 r 次外向量组 $\{e_{i_1} \wedge \dots \wedge e_{i_r} \mid 1 \leq i_1 < \dots < i_r \leq n\}$ (一共包含 $\binom{n}{r} = \frac{n!}{r!(n-r)!}$ 个

外向量) 是线性无关的。为此，先导出 $e_{i_1} \wedge \dots \wedge e_{i_r}$ 的求值公式：设 v^{*1}, \dots, v^{*r} 是 V^* 中任意 r 个元素，则

$$\begin{aligned} & e_{i_1} \wedge \dots \wedge e_{i_r} (v^{*1}, \dots, v^{*r}) \\ &= \frac{1}{r!} \sum_{\sigma \in \varphi(r)} \text{sign } \sigma \cdot \langle e_{i_1}, v^{*\sigma(1)} \rangle \dots \langle e_{i_r}, v^{*\sigma(r)} \rangle \\ &= \frac{1}{r!} \det \begin{pmatrix} \langle e_{i_1}, v^{*1} \rangle & \dots & \langle e_{i_1}, v^{*r} \rangle \\ \langle e_{i_2}, v^{*1} \rangle & \dots & \langle e_{i_2}, v^{*r} \rangle \\ \dots & \dots & \dots \\ \langle e_{i_r}, v^{*1} \rangle & \dots & \langle e_{i_r}, v^{*r} \rangle \end{pmatrix} \end{aligned} \tag{10.4.8}$$

上式称为 $e_{i_1} \wedge \dots \wedge e_{i_r}$ 的求值公式。特别地，

$$e_{i_1} \wedge \dots \wedge e_{i_r} (e^{*j_1}, \dots, e^{*j_r}) = \frac{1}{r!} \det(\langle e_{i_\alpha}, e^{*j_\beta} \rangle) = \frac{1}{r!} \delta_{i_1 \dots i_r}^{j_1 \dots j_r}$$

其中

$$\delta_{i_1 \dots i_r}^{j_1 \dots j_r} = \begin{cases} 1, & \text{当 } i_1, \dots, i_r \text{ 互不相同, 且 } \{j_1, \dots, j_r\} \text{ 是 } \{i_1, \dots, i_r\} \text{ 的偶排列;} \\ -1, & \text{当 } i_1, \dots, i_r \text{ 互不相同, 且 } \{j_1, \dots, j_r\} \text{ 是 } \{i_1, \dots, i_r\} \text{ 的奇排列;} \\ 0, & \text{其他情形.} \end{cases} \tag{10.4.10}$$

称为广义的 Kronecker 符号。由(10.4.9)式得到

$$e_1 \wedge \cdots \wedge e_n (e^{*1}, \dots, e^{*n}) = \frac{1}{n!} \tag{10.4.11}$$

故 $e_1 \wedge \cdots \wedge e_n \neq 0$ 。

对于 $r < n$ ，如果 $\{e_{i_1} \wedge \cdots \wedge e_{i_r}, 1 \leq i_1 < \cdots < i_r \leq n\}$ 线性相关，则有不全为零的数 $a^{i_1 \cdots i_r} \in F$ ，使

$$\sum_{1 \leq i_1 < \cdots < i_r \leq n} a^{i_1 \cdots i_r} e_{i_1} \wedge \cdots \wedge e_{i_r} = 0 \tag{10.4.12}$$

不妨设其中一个不为零的数是 $a^{j_1 \cdots j_r}, 1 \leq j_1 < \cdots < j_r \leq n$ ，与它互补的指标组记为 $k_1 < \cdots < k_{n-r}$ ，也就是说 $\{j_1, \dots, j_r, k_1, \dots, k_{n-r}\}$ 恰是 $\{1, \dots, n\}$ 的一个排列。用 $e_{k_1} \wedge \cdots \wedge e_{k_{n-r}}$ 外乘 (10.4.12) 的两边，于是得到

$$a^{i_1 \cdots i_r} e_{i_1} \wedge \cdots \wedge e_{i_r} \wedge e_{k_1} \wedge \cdots \wedge e_{k_{n-r}} = \pm a^{i_1 \cdots i_r} e_1 \wedge \cdots \wedge e_n = 0$$

因 $e_1 \wedge \cdots \wedge e_n \neq 0$ ，所以

$$a^{i_1 \cdots i_r} = 0,$$

这是一个矛盾。因此 $\{e_{i_1} \wedge \cdots \wedge e_{i_r}, 1 \leq i_1 < \cdots < i_r \leq n\}$ 是线性无关的，它们构成 $A^r(V)$ 的基。证毕。

定理 10.4.2 说明， r 次外向量空间 $A^r(V)$ 是维数为 $\binom{n}{r} = \frac{n!}{r!(n-r)!}$ 的向量空间。

例 10.4.3 取 $\xi, \eta \in R^n$ ，根据例 10.4.2，我们有

$$\xi \wedge \eta = \sum_{1 \leq i, j \leq n} (x^i y^j - x^j y^i) e_i \wedge e_j = \sum_{1 \leq i, j \leq n} \det \begin{pmatrix} x^i & y^i \\ x^j & y^j \end{pmatrix} e_i \wedge e_j$$

特别地，当 $n = 2$ ， $\xi \wedge \eta = \det(\xi, \eta) e_1 \wedge e_2$ ，其系数表示向量 ξ, η 所构成的平行四边形的有向面积，如图 10.4.1 所示。一般地， $(x^i y^j - x^j y^i)$ 是 ξ, η 在子空间 $\text{span}\{e_i, e_j\}$ 的投影分量所构成的平行四边形的有向面积。

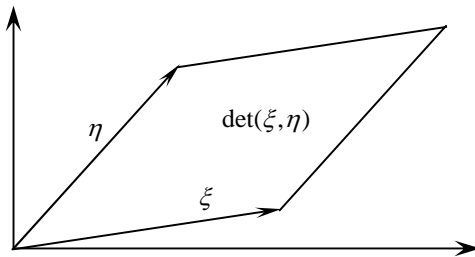


图 10.4.1 R^2 中两个向量的外积是它们所构成的平行四边形的有向面积

同理，可以考虑 $A^r(V^*)$ ，但 $A^r(V^*)$ 中元素是向量空间 V^* 上的 r 次外向量，我们通常又称它为向量空间 V 上的 r 次外形式，它是 V 上的反对称 F -值 r 重线性函数。 r 次外向量空间 $A^r(V)$ 与 r 次外形式空间 $A^r(V^*)$ 是彼此对偶的，它们之间的配合 \langle, \rangle 定义如下：

设 $v_1 \wedge \cdots \wedge v_r \in A^r(V)$, $v^{*1} \wedge \cdots \wedge v^{*r} \in A^r(V^*)$, 则

$$\langle v_1 \wedge \cdots \wedge v_r, v^{*1} \wedge \cdots \wedge v^{*r} \rangle = \det(\langle v_\alpha, v^{*\beta} \rangle) \quad (10.4.15)$$

这样， $A^r(V)$ 和 $A^r(V^*)$ 的基分别为

$$\{e_{i_1} \wedge \cdots \wedge e_{i_r}, 1 \leq i_1 < \cdots < i_r \leq n\}$$

和

$$\{e^{*j_1} \wedge \cdots \wedge e^{*j_r}, 1 \leq j_1 < \cdots < j_r \leq n\}$$

并且，有下述关系：

$$\begin{aligned} \langle e_{i_1} \wedge \cdots \wedge e_{i_r}, e^{*j_1} \wedge \cdots \wedge e^{*j_r} \rangle &= \det(\langle e_{i_\alpha}, e^{*j_\beta} \rangle) = \delta_{i_1 \cdots i_r}^{j_1 \cdots j_r} \\ &= \begin{cases} 1, & \{j_1, \cdots, j_r\} = \{i_1, \cdots, i_r\}, \\ 0, & \{j_1, \cdots, j_r\} \neq \{i_1, \cdots, i_r\}, \end{cases} \end{aligned} \quad (10.4.16)$$

因此，这两个基彼此对偶。

10.4.2 外代数 (Exterior Algebra)

用 $A(V)$ 表示形式和 $\bigoplus_{r=0}^n A^r(V)$ ，即 $A(V) = \bigoplus_{r=0}^n A^r(V)$ ，也就是说它的元素由下述形式和生成：

$$\xi = \bigoplus_{r=0}^n \xi^r, \xi^r \in A^r(V)$$

令 $\xi^r = \sum_{1 \leq j_1 < \cdots < j_r \leq n} x^{j_1 \cdots j_r} e_{j_1} \wedge \cdots \wedge e_{j_r}$ ， $A(V)$ 中的元素能被表示为

$$\xi = x^0 \cdot 1 + \sum_{1 \leq j \leq n} x^j e_j + \sum_{1 \leq j_1 < j_2 \leq n} x^{j_1 j_2} e_{j_1} \wedge e_{j_2} + \cdots + \sum_{1 \leq j_1 < \cdots < j_r \leq n} x^{j_1 \cdots j_r} e_{j_1} \wedge \cdots \wedge e_{j_r} + \cdots + x^{1 \cdots n} e_1 \wedge \cdots \wedge e_n$$

所以 $A(V)$ 是 2^n 维向量空间，它的基为

$$\{1, e_i (1 \leq i \leq n), e_{i_1} \wedge e_{i_2} (1 \leq i_1 < i_2 \leq n), \cdots, e_{i_1} \wedge \cdots \wedge e_{i_r} (1 \leq i_1 < i_2 < \cdots < i_r \leq n), \cdots, e_1 \wedge \cdots \wedge e_n\}$$

设

$$\xi = \bigoplus_{r=0}^n \xi^r, \eta = \bigoplus_{s=0}^n \eta^s \quad (10.4.17)$$

其中 $\xi^r \in A^r(V)$, $\eta^s \in A^s(V)$ ，定义 ξ 和 η 的外积是

$$\xi \wedge \eta = \bigoplus_{r,s=0}^n \xi^r \wedge \eta^s \quad (10.4.18)$$

则 $A(V)$ 关于外积成为一个代数, 称为向量空间 V 的外代数或 Grassmann 代数。

同样, 可以定义对偶空间 V^* 的外代数:

$$A(V^*) = \bigoplus_{r=0}^n A^r(V^*) \quad (10.4.19)$$

空间 $A^r(V^*)$ 的元素称为向量空间 V 上的 r 次外形式, 它是 V 上的反对称 F -值 r 重线性函数。

设 $f: V \rightarrow W$ 是向量空间 V 到 W 的线性映射, 则它诱导出外形式空间 $A^r(W^*)$ 到 $A^r(V^*)$ 的线性映射 f^* : 设 $\varphi \in A^r(W^*)$, 对任意的 $v_1, \dots, v_r \in V$, 令

$$f^*\varphi(v_1, \dots, v_r) = \varphi(f(v_1), \dots, f(v_r)). \quad (10.4.20)$$

容易证明 f^* 是线性的。由于下述定理证明了 f^* 关于外积运算是可分配的, 因此 f^* 是外代数 $A(W^*)$ 到 $A(V^*)$ 的同态。

定理 10.4.3 设 $f: V \rightarrow W$ 是线性映射, 则 f^* 关于外积运算是可分配的。即对于任意 $\varphi \in A^r(W^*)$ 和 $\psi \in A^s(W^*)$ 有

$$f^*(\varphi \wedge \psi) = f^*\varphi \wedge f^*\psi \quad (10.4.21)$$

证明 任取 $v_1, \dots, v_{r+s} \in V$, 则

$$\begin{aligned} f^*(\varphi \wedge \psi)(v_1, \dots, v_{r+s}) &= \varphi \wedge \psi(f(v_1), \dots, f(v_{r+s})) \\ &= \frac{1}{(r+s)!} \sum_{\sigma \in \gamma(r+s)} \text{sign } \sigma \cdot \varphi(f(v_{\sigma(1)}), \dots, f(v_{\sigma(r)})) \cdot \psi(f(v_{\sigma(r+1)}), \dots, f(v_{\sigma(r+s)})) \\ &= \frac{1}{(r+s)!} \sum_{\sigma \in \gamma(r+s)} \text{sign } \sigma \cdot f^*\varphi(v_{\sigma(1)}, \dots, v_{\sigma(r)}) \cdot f^*\psi(v_{\sigma(r+1)}, \dots, v_{\sigma(r+s)}) \\ &= f^*\varphi \wedge f^*\psi(v_1, \dots, v_{r+s}). \end{aligned}$$

所以,

$$f^*(\varphi \wedge \psi) = f^*\varphi \wedge f^*\psi.$$

证毕。

外代数的概念最初是由 Grassmann 为了研究线性子空间而引进的。下面给出几个很有用的命题。

定理 10.4.3 向量 $v_1, v_2, \dots, v_r \in V$. 线性相关的充要条件是

$$v_1 \wedge \dots \wedge v_r = 0. \quad (10.4.22)$$

证明 若 v_1, v_2, \dots, v_r 线性相关, 不妨设 v_r 可以表示为 v_1, \dots, v_{r-1} 的线性组合:

$$v_r = a_1 v_1 + \dots + a_{r-1} v_{r-1}.$$

因此，

$$v_1 \wedge \cdots \wedge v_{r-1} \wedge v_r = v_1 \wedge \cdots \wedge v_{r-1} \wedge (a_1 v_1 + \cdots + a_{r-1} v_{r-1}) = 0$$

若 v_1, v_2, \dots, v_r 线性无关，则可将它们扩充成 V 的一个基 $\{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$ 。由于

$$v_1 \wedge \cdots \wedge v_r \wedge v_{r+1} \wedge \cdots \wedge v_n \neq 0$$

所以

$$v_1 \wedge \cdots \wedge v_r \neq 0$$

证毕。

定理 10.4.5(Cartan) 设 $v_1, \dots, v_r; w_1, \dots, w_r$ 是 V 中两组向量，使得

$$\sum_{\alpha=1}^r v_\alpha \wedge w_\alpha = 0 \tag{10.4.23}$$

如果 v_1, \dots, v_r 线性无关，则 w_α 可表示为它们的线性组合

$$w_\alpha = \sum_{\beta=1}^r a_{\alpha\beta} v_\beta, 1 \leq \alpha \leq r \tag{10.4.24}$$

并且

$$a_{\alpha\beta} = a_{\beta\alpha} \tag{10.4.25}$$

证明 因为 v_1, \dots, v_r 是线性无关的，所以可以将它们扩充成 V 的一个基 $\{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$ 。因此，不妨假定

$$w_\alpha = \sum_{\beta=1}^r a_{\alpha\beta} v_\beta + \sum_{i=r+1}^n a_{\alpha i} v_i \tag{10.4.26}$$

代入(10.4.23)式得到

$$0 = \sum_{\alpha,\beta=1}^r a_{\alpha\beta} v_\alpha \wedge v_\beta + \sum_{\alpha=1}^r \sum_{i=r+1}^n a_{\alpha i} v_\alpha \wedge v_i = \sum_{1 \leq \alpha < \beta \leq r} (a_{\alpha\beta} - a_{\beta\alpha}) v_\alpha \wedge v_\beta + \sum_{\alpha=1}^r \sum_{i=r+1}^n a_{\alpha i} v_\alpha \wedge v_i \tag{10.4.27}$$

由于 $\{v_i \wedge v_j, 1 \leq i < j \leq n\}$ 是 $A^2(V)$ 的一个基底，因此从(10.4.27)式得到

$$a_{\alpha\beta} - a_{\beta\alpha} = 0, a_{\alpha j} = 0, j > r$$

即

$$w_\alpha = \sum_{\beta=1}^r a_{\alpha\beta} v_\beta, a_{\alpha\beta} = a_{\beta\alpha}。$$

证毕。

定理 10.4.6 设 v_1, \dots, v_r 是 V 中 r 个线性无关的向量， w 是 V 上的 p 次外向量，则存在

$$\psi_1, \dots, \psi_r \in A^{p-1}(V)$$

使 w 能表成

$$w = v_1 \wedge \psi_1 + \cdots + v_r \wedge \psi_r \tag{10.4.28}$$

充要条件是

$$v_1 \wedge \cdots \wedge v_r \wedge w = 0 \tag{10.4.29}$$

证明 当 $p+r > n$ 时, (10.4.28)和(10.4.29)两式显然成立。下面, 假定 $p+r \leq n$ 。

必要性是明显的, 只须证明充分性: 把 v_1, \dots, v_r 扩充成 V 的一个基 $\{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$, 则 w 可以表示成

$$w = v_1 \wedge \psi_1 + \cdots + v_r \wedge \psi_r + \sum_{r+1 \leq \alpha_1 < \cdots < \alpha_p \leq n} \xi^{\sigma_1 \cdots \sigma_p} v_{\alpha_1} \wedge \cdots \wedge v_{\alpha_p} = 0. \tag{10.4.30}$$

其中 $\psi_1, \dots, \psi_r \in A^{p-1}(V)$ 。代入(10.4.29)式得到

$$\sum_{r+1 \leq \alpha_1 < \cdots < \alpha_p \leq n} \xi^{\sigma_1 \cdots \sigma_p} v_1 \wedge \cdots \wedge v_r \wedge v_{\alpha_1} \wedge \cdots \wedge v_{\alpha_p} = 0. \tag{10.4.31}$$

而和号后的 $v_1 \wedge \cdots \wedge v_r \wedge v_{\alpha_1} \wedge \cdots \wedge v_{\alpha_p}$ ($r+1 \leq \alpha_1 < \cdots < \alpha_p \leq n$) 正是 $A^{p+z}(V)$ 的基的一部分, 所以从(10.4.31)式得到

$$\xi^{\sigma_1 \cdots \sigma_p} = 0 \quad (r+1 \leq \alpha_1 < \cdots < \alpha_p \leq n)$$

即

$$w = v_1 \wedge \psi_1 + \cdots + v_r \wedge \psi_r。$$

证毕。

定理 10.4.7 设 $v_\alpha, w_\alpha; v_\alpha^p, w_\alpha^p$ ($1 \leq \alpha \leq k$) 是空间 V 中两组向量。若 $\{v_\alpha, w_\alpha, 1 \leq \alpha \leq k\}$ 是线性无关的, 并且

$$\sum_{\alpha=1}^k v_\alpha \wedge w_\alpha = \sum_{\alpha=1}^k v_\alpha^p \wedge w_\alpha^p \tag{10.4.32}$$

则 v_α^p, w_α^p 都是 $v_1, \dots, v_k, w_1, \dots, w_k$ 的线性组合, 而且它们也是线性无关的。

证明 将(10.4.32)式自乘 k 次得到

$$k!(v_1 \wedge w_1 \cdots \wedge v_k \wedge w_k) = k!(v_1^p \wedge w_1^p \cdots \wedge v_k^p \wedge w_k^p) \tag{10.4.33}$$

因为 $\{v_\alpha, w_\alpha, 1 \leq \alpha \leq k\}$ 是线性无关的, 故(10.4.33)式左边 $\neq 0$ 。即 $\{v_\alpha^p, w_\alpha^p, 1 \leq \alpha \leq k\}$ 也线性无关(定理 10.4.3)。从(10.4.33)式, 还可以得到

$$v_1 \wedge w_1 \cdots \wedge v_k \wedge w_k \wedge v_\alpha^p = 0$$

即 $\{v_1, w_1, \dots, v_k, w_k, v_\alpha^p\}$ 是线性相关的, 所以 v_α^p 能表成 $v_1, \dots, v_k, w_1, \dots, w_k$ 的线性组合。上面的结论对 w_α^p 也成立。证毕。

外代数与行列式是密切相关的, 如外向量的求值公式(10.4.8)就表现为行列式。设 $v_1, \dots, v_k \in V$, 而 w_1, \dots, w_k 是它们的线性组合, 即

$$w_\alpha = \sum_{\beta=1}^k t_\alpha^\beta v_\beta, \quad (10.4.34)$$

则

$$w_1 \wedge \dots \wedge w_k = \det(t_\alpha^\beta) v_1 \wedge \dots \wedge v_k \quad (10.4.35)$$

因此, 外向量 $w_1 \wedge \dots \wedge w_k$ 和 $v_1 \wedge \dots \wedge v_k$, 只差一个行列式作为数量因子。

10.4.3 Plucker-Grassmann 坐标

在 2.1.3 节, 介绍了 3 维射影空间的直线 Plucker 坐标。3 维射影空间的直线有 4 个自由度, 在直觉上有点令人费解, 并且 Plucker 坐标的导入也不像点坐标和平面坐标那样十分自然。本节, 我们将通过外向量给出 n 维射影空间的直线、平面和超平面这些基本几何元素坐标的统一表达, 即 Plucker-Grassmann 坐标。

令 V 是 $n+1$ 维向量空间, 记 $G(n+1, k)$ 为 V 的所有 k 维线性子空间 L^k 的集合。 $G(n+1, k)$ 通常称这 Grassmann 流形, 它是 $k(n-k+1)$ 维的。在 $k=1$ 时, $G(n+1, 1)$ 是 n 维射影空间 P^n , $G(n+1, 1)$ 每个元素, 即 V 的一维线性子空间 L , 是射影空间 P^n 的一个点。 $G(n+1, 2)$ 中的元素, 即 V 的二维线性子空间 L^2 , 是射影空间 P^n 中的直线; $G(n+1, 3)$ 中的元素, 即 V 的 3 维线性子空间 L^3 , 是射影空间 P^n 中的平面; 一般地, $G(n+1, k)$ 中的元素, 即 V 的 k 维线性子空间 L^k , 是射影空间 P^n 中的 $k-1$ 维超平面。

对任意 $L^k \in G(n+1, k)$, 令 $\{v_1, v_2, \dots, v_k\}$ 是张成线性子空间 L^k 的 k 个线性无关的向量, 根据式(10.4.35)外向量:

$$\eta = v_1 \wedge v_2 \wedge \dots \wedge v_k \quad (10.4.36)$$

可以确定到相差一个非常数因子的程度, 即如果 $\{u_1, u_2, \dots, u_k\}$ 是另一个张成线性子空间 L^k 的 k 个线性无关向量组, 根据式(10.4.35)必有

$$\eta = v_1 \wedge v_2 \wedge \dots \wedge v_k = \det(t_\alpha^\beta) u_1 \wedge u_2 \wedge \dots \wedge u_k$$

很明显, L^k 在向量空间 V 中的方程可表示为 $\eta \wedge x = 0, x \in V$ 。因此, 我们定义外向量 η 是线性子空间 L^k 的坐标, 即射影空间 P^n 中 $k-1$ 维超平面的齐次坐标, 通常称它为 Plucker-Grassmann 坐标。

下面, 令 $V = R^4$, 考虑相应的 3 维射影空间 P^3 中的点、直线和平面的齐次坐标, 即考虑 $G(4, 1)$, $G(4, 2)$, $G(4, 3)$ 中元素的 Plucker-Grassmann 坐标。

令 $L \in G(4,1)$, 则存在 $x \in R^4, x \neq 0$ 使得 $L = \{\alpha x \mid \alpha \in R\}$, 按照定义, L 的 Plucker-Grassmann 坐标为 $\eta = x$, 这与 2.1 节定义的 3 维射影空间中点的齐次坐标是一致的。令 $L \in G(4,3)$, 则存在 $x, y, z \in R^4$ 使得 $L^3 = \text{span}\{x, y, z\}$, 按照定义, L^3 的 Plucker-Grassmann 坐标

$$\eta = x \wedge y \wedge z = 3! \sum_{1 \leq i < j < k \leq 4} d^{ijk} e_i \wedge e_j \wedge e_k \quad (10.4.37)$$

由于,

$$e_1 \wedge \eta = e_1 \wedge x \wedge y \wedge z = 3! d^{234} e_1 \wedge e_2 \wedge e_3 \wedge e_4$$

同时, 我们有

$$\begin{aligned} e_1 \wedge x \wedge y \wedge z &= e_1 \wedge \sum_{i=1}^4 x^i e_i \wedge \sum_{j=1}^4 y^j e_j \wedge \sum_{k=1}^4 z^k e_k \\ &= 3! \sum_{2 \leq i < j < k \leq 4} x^i y^j z^k e_1 \wedge e_i \wedge e_j \wedge e_k = 3! \sum_{2 \leq i < j < k \leq 4} \text{sign}(lijk) \cdot x^i y^j z^k e_1 \wedge e_2 \wedge e_3 \wedge e_4 \end{aligned}$$

所以,

$$d^{234} = \sum_{2 \leq i < j < k \leq 4} \text{sign}(lijk) \cdot x^i y^j z^k = \det(X(\sim 1))$$

其中 $X = (x, y, z)$, $\det(X(\sim i))$ 表示矩阵 X 删去第 i 行的 3 阶子阵行列式。同理, 我们能得到

$$d^{134} = - \sum_{1 \leq i < j < k \leq 4; i, j, k \neq 2} \text{sign}(2ijk) \cdot x^i y^j z^k = -\det(X(\sim 2)),$$

$$d^{124} = \sum_{1 \leq i < j < k \leq 4; i, j, k \neq 3} \text{sign}(3ijk) \cdot x^i y^j z^k = \det(X(\sim 3)),$$

$$d^{123} = - \sum_{1 \leq i < j < k \leq 4} \text{sign}(4ijk) \cdot x^i y^j z^k = \det(X(\sim 4))$$

于是, 3 射影空间中平面的齐次坐标为

$$\pi = (d^{234}, d^{134}, d^{124}, d^{123})^T$$

这与 2.1 节的结果是一致的。

现在考虑 3 维射影空间中直线的齐次坐标。令 $L^2 \in G(4,2)$, 则 $x, y \in R^4$ 使得 $L^2 = \text{span}\{x, y\}$, 按照定义, L^2 的 Plucker-Grassmann 坐标为

$$\eta = x \wedge y$$

根据例 10.4.2, 我们有

$$\eta = \sum_{1 \leq i < j \leq 4} p^{ij} e_i \wedge e_j \quad (10.4.38)$$

其中,

$$p^{ij} = x^i y^j - x^j y^i, 1 \leq i < j \leq 4 \quad (10.4.39)$$

下面证明，数组 $\{p^{ij} \mid 1 \leq i < j \leq 4\}$ 正是直线的 Plucker 坐标。为此，只须证明 $\{p^{ij}\}$ 满足 Plucker 方程。由于

$$\eta \wedge \eta = x \wedge y \wedge x \wedge y = 0$$

且

$$\begin{aligned} \left(\sum_{1 \leq i < j \leq 4} p^{ij} e_i \wedge e_j \right) \wedge \left(\sum_{1 \leq k < l \leq 4} p^{kl} e_k \wedge e_l \right) &= (p^{12} p^{34} - p^{13} p^{24} + p^{14} p^{23}) e_1 \wedge e_2 \wedge e_3 \wedge e_4 \\ &= (p^{12} p^{34} + p^{13} p^{42} + p^{14} p^{23}) e_1 \wedge e_2 \wedge e_3 \wedge e_4 \end{aligned}$$

所以， $\{p^{ij}\}$ 满足 Plucker 方程：

$$p^{12} p^{34} + p^{13} p^{42} + p^{14} p^{23} = 0 \quad (10.4.40)$$

另一方面，我们还可以证明，满足 Plucker 方程的数组 $\{p^{ij}\}$ 所对应的外向量

$$\eta = \sum_{1 \leq i < j \leq 4} p^{ij} e_i \wedge e_j$$

一定是可分解的，即这样的数组一定是 3 维射影空间中某直线的 Plucker 坐标。证明如下：不妨假定 $p^{12} \neq 0$ ，则

$$p^{34} = -p^{13} p^{24} / p^{12} - p^{14} p^{23} / p^{12}$$

于是，我们有

$$\begin{aligned} &\sum_{1 \leq i < j \leq 4} p^{ij} e_i \wedge e_j \\ &= e_1 \wedge (p^{12} e_2 + p^{13} e_3 + p^{14} e_4) - p^{23} / p^{12} e_3 \wedge (p^{12} e_2 + p^{13} e_3 + p^{14} e_4) - p^{24} / p^{12} e_3 \wedge (p^{12} e_2 + p^{13} e_3 + p^{14} e_4) \\ &= (e_1 - p^{23} / p^{12} e_3 - p^{24} / p^{12} e_4) \wedge (p^{12} e_2 + p^{13} e_3 + p^{14} e_4)。 \end{aligned}$$

11. 运动与结构

本章将介绍 SVD 分解在运动与结构估计的应用。

给定两幅图像 $\{I, I'\}$ ，假定两幅图像所对应的摄像机内参数矩阵是已知的，求解两个摄像机之间的欧氏运动和景物的欧氏结构。它是三维计算机视觉的主要研究内容。本章将介绍求解欧氏运动和结构问题的 SVD 分解方法，实践表明它是一种非常有效的求解方法。

给定 n 幅图像的一组点对应，并且已知 n 这幅图像分别来自一个仿射摄像机的投影，但摄像机矩阵是未知的。从这组点对应求解它们对应的空间点和仿射摄像机矩阵，通常称为仿射运动与结构问题。从仿射摄像机的一组图像点对应的获得的结构是景物的仿射结构，所以这样得到结构也称为仿射重构。应用 SVD 分解技术，非常方便地实现此问题的求解。

两幅图像的射影重构可以通过基本阵来实现，三幅图像的射影重构可以通过求解三焦张量来实现。然而，对于三幅以上图像，大多数方法是以其中的某两幅图像(或三幅图)为基准来实现射影重构，这些方法主要不足之处在于重构结果严重依赖于基准图像的选择。基于 SVD 分解的多幅图像的射影重构技术在一定程度上克服了上述方法的不足。

11.1 欧氏运动与结构

11.1.1 本质矩阵

给定两幅图像 $\{I, I'\}$ ，令第二幅图像的摄像机相对于第一幅图像的摄像机之间的运动为 (R, t) (其中 R 是摄像机的旋转， t 是摄像机的平移)。假定已知它们的内参数矩阵分别为 K, K' 。

由于已知摄像机内参数，可以对两幅图像作规范化变换：

$$m_n = K^{-1}m, m'_n = K'^{-1}m' \quad (11.1.1)$$

得到两幅新图像 $\{I_n, I'_n\}$ ，称它们为原图像的规范化图像。因为原图像间的基本矩阵为

$$F = K'^{-T}[t]_{\times}RK^{-1}$$

所以, 规范化图像 I_n, I'_n 间的极线约束方程必为

$$\mathbf{m}_n'^T [\mathbf{t}]_{\times} R \mathbf{m}_n' = 0 \quad (11.1.2)$$

方程(11.1.2)是由矩阵:

$$E = [\mathbf{t}]_{\times} R \quad (11.1.3)$$

所确定的, 通常称这个矩阵为本质矩阵。

本质矩阵描述了两幅规范化图像间的极几何, 它与基本矩阵一样也是一个秩为 2 的矩阵。由于本质矩阵消去了内参数的作用, 所以它与基本矩阵的不同之处是它仅与摄像机的运动参数有关。因此, 从本质矩阵出发可估计出摄像机的欧氏运动参数。

本质矩阵是秩 2 的, 它仅有 5 个自由度。由 8 个图像点对应, 应用(11.1.2)式可建立关于本质矩阵的线性约束方程, 因此从点对应可线性求解本质矩阵。由于方程(11.1.2)的齐次性, 只能在相差非零因子的意义下获得本质矩阵 E 。因此, 只能在相差一个非零因子的意义下建立关于运动参数 (R, \mathbf{t}) 的方程 $E = [\tilde{\mathbf{t}}]_{\times} R$, 即这里 $\tilde{\mathbf{t}}$ 与摄像机的实际平移向量 \mathbf{t} 相差一个常数因子。从图像恢复摄像机运动参数, 是指通过方程(11.1.3)求解平移向量的非零常数倍 (但需给出正确的平移方向) 和旋转矩阵 R 。

本质矩阵的性质

基于 SVD 分解的运动参数估计, 主要依赖于本质矩阵的下述性质:

命题 11.1.1 实 3 阶矩阵 E 为本质矩阵的充要条件是它有两个相同的奇异值和一个零奇异值。

证明 必要性: 因 $E = [\mathbf{t}]_{\times} R$, 对于平移向量 \mathbf{t} , 所以存在一个旋转矩阵 Q 使得 $Q\mathbf{t} = (0, 0, s)^T$, 其中 $s = \|\mathbf{t}\|$, 于是有

$$EE^T = [\mathbf{t}]_{\times} [\mathbf{t}]_{\times}^T = s^2 \mathbf{I} - \mathbf{t}\mathbf{t}^T = Q^T \text{diag}(s^2, s^2, 0) Q$$

所以 E 有两个相同的奇异值和一个零奇异值。

充分性: 假定 E 有两个相同的奇异值和一个零奇异值, 则 E 有 SVD 分解:

$$E = U \text{diag}(s^2, s^2, 0) V^T$$

其中 U, V 是正交矩阵。令

$$S = \text{diag}(s^2, s^2, 0) \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, Z = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

则 S 是反对称矩阵, Z 是旋转矩阵且

$$E = (USU^T)(UZV^T) = [\mathbf{t}]_{\times} R$$

其中

$$[t]_x = USU^T, R = UZV^T$$

所以, E 为本质矩阵。证毕。

11.1.2 欧氏运动估计

引理 11.1.1 令 $E = Udiag(\sigma, \sigma, 0)V^T$ 是本质矩阵 E 的一个奇异值分解, 则它的所有奇异值分解均可表示为

$$E = U \begin{pmatrix} W_{2 \times 2} & 0 \\ 0 & \pm 1 \end{pmatrix} diag(\sigma, \sigma, 0) \begin{pmatrix} W_{2 \times 2} & 0 \\ 0 & \pm 1 \end{pmatrix}^T V^T \quad (11.1.4)$$

其中 W 是任意的 2 阶正交矩阵。

证明 显然对任意 2 阶正交矩阵 W , 式(11.1.4)是 E 的一个奇异值分解。反之, 令

$$E = \tilde{U}diag(\sigma, \sigma, 0)\tilde{V}^T$$

是 E 的另一个奇异值分解, 则必有

$$\tilde{U}diag(\sigma, \sigma, 0)\tilde{V}^T = Udiag(\sigma, \sigma, 0)V^T$$

即,

$$\tilde{U}^T U diag(\sigma, \sigma, 0) (\tilde{V}^T V)^T = diag(\sigma, \sigma, 0)$$

令

$$\hat{U} = \tilde{U}^T U = \begin{pmatrix} A & a \\ b^T & c \end{pmatrix}, \hat{V} = \tilde{V}^T V = \begin{pmatrix} D & d \\ e^T & f \end{pmatrix}$$

并代入上式, 可推知:

$$AB^T = I_{2 \times 2}, A^T d = 0, \hat{V} b = 0$$

所以, A, B 均为可逆矩阵, 从而 $b = 0, d = 0$ 。因 \hat{U}, \hat{V} 都是正交矩阵, 于是,

$$a = 0, e = 0, c = \pm 1, f = \pm 1$$

且 A, B 为正交矩阵, $A=B$ 。因此,

$$\tilde{U} = U \begin{pmatrix} W & 0 \\ 0 & \pm 1 \end{pmatrix}, \tilde{V} = V \begin{pmatrix} W & 0 \\ 0 & \pm 1 \end{pmatrix}$$

其中 W 是 2 阶正交矩阵。证毕。

定理 11.1.1 对于给定的本质矩阵 E , 令 $E = Udiag(\sigma, \sigma, 0)V^T$ 是一个奇异值分解, 则运动参数有

四组解:

$$(1) [t]_x = USU^T, R = UZV^T$$

$$(2) [t]_x = US^T U^T, R = UZ^T V^T$$

$$(3) [t]_x = USU^T, R = UZ^T V^T$$

$$(4) [t]_x = US^T U^T, R = UZV^T$$

其中

$$S = \text{diag}(\sigma, \sigma, 0) \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, Z = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

证明 令 $E = U \text{diag}(\sigma, \sigma, 0) V^T$ 是 E 的一个奇异值分解, 由引理 11.1.1 它的所有奇异值分解均可表示为

$$E = U \begin{pmatrix} W_{2 \times 2} & 0 \\ 0 & \pm 1 \end{pmatrix} \text{diag}(\sigma, \sigma, 0) \begin{pmatrix} W_{2 \times 2} & 0 \\ 0 & \pm 1 \end{pmatrix}^T V^T$$

其中 W 是任意的 2 阶正交矩阵。令 $\hat{U} = U \begin{pmatrix} W_{2 \times 2} & 0 \\ 0 & \pm 1 \end{pmatrix}$, $\hat{V} = V \begin{pmatrix} W_{2 \times 2} & 0 \\ 0 & \pm 1 \end{pmatrix}$, 则

$$E = \hat{U} \text{diag}(\sigma, \sigma, 0) \hat{V}^T = (\hat{U} S \hat{U}^T) (\hat{U} Z \hat{V}^T)$$

当 W 为旋转矩阵时, 有 $[t]_x = \hat{U} S \hat{U}^T = USU^T$, $R = \hat{U} Z \hat{V}^T = UZV^T$, 并且从 $[t]_x = USU^T$, 可推

知 $t = U \begin{pmatrix} 0 \\ 0 \\ \sigma \end{pmatrix} = \sigma u_3$ (其中 u_3 是 U 第 3 列向量), 于是,

$$t = \sigma u_3, R = UZV^T$$

当 W 为反射矩阵时, 有 $[t]_x = \hat{U} S \hat{U}^T = US^T U^T$, $R = \hat{U} Z \hat{V}^T = UZ^T V^T$, 并且从 $[t]_x = US^T U^T$,

可推知 $t = U \begin{pmatrix} 0 \\ 0 \\ -\sigma \end{pmatrix} = -\sigma u_3$, 于是,

$$t = -\sigma u_3, R = UZ^T V^T$$

同理, 考虑 E 的反号矩阵 $-E$ 可以得到运动参数的另外两组解:

$$t = \sigma u_3, R = UZ^T V^T; \quad t = -\sigma u_3, R = UZ V^T$$

证毕。

对于实际图像, 由于数据误差所得到的本质矩阵 E 也不可避免地带有误差, 因而在一般情况下 E 不可能有零奇异值, 也没有两个相等的奇异值。这样, 不能直接利用定理 11.2.1 的计算结果作为运动参数的估计。此时, 需要用具有两个相同的奇异值和一个零奇异值的矩阵作 E 的最佳近似, 常用的方法是将 E 的最小奇异值置为零, 用另外两个奇异值的均值作为两个相等的奇异值。确切地说, 如果 $E = U \text{diag}(\sigma_1, \sigma_2, \sigma_3) V^T$, 则我们用矩阵:

$$E^* = U \text{diag}((\sigma_1 + \sigma_2)/2, (\sigma_1 + \sigma_2)/2, 0) V^T$$

作为 E 的近似。此外, 还可以证明这种近似是最小二乘意义下的最佳近似。

11.1.3 欧氏结构估计

摄像机矩阵

定理 11.1.2 给定本质矩阵 E , 令 $E = U \text{diag}(\sigma, \sigma, 0) V^T$ 是它的一个奇异值分解, 如果第一个摄像机矩阵规定为 $P = (I, 0)$, 则第二个摄像机投影矩阵的四组可能解为

$$P' = (UZV^T, \sigma u_3), (UZV^T, -\sigma u_3), (UZ^T V^T, \sigma u_3), (UZ^T V^T, -\sigma u_3) \quad (11.1.5)$$

证明 由 $[t]_x = USU^T$, 知 $[U^T t]_x = S$, 所以

$$t = U \begin{pmatrix} 0 \\ 0 \\ \sigma \end{pmatrix} = \sigma u_3, \quad t = -\sigma u_3$$

因此, 从定理 11.1.1 可知第二个摄像机投影矩阵有上述可能的四组解。证毕。

多解的几何解释 式(11.1.5)中第一组与第二组有相同的旋转矩阵, 但平移方向相反, 第三组与第四组的情况相同。第一组与第三组有相同的平移, 但具有不同的旋转矩阵, 第二组与第四组的情况相同。下面分析第一组与第三组(第二组与第四组)中旋转矩阵之间的关系。不难看出:

$$(UZ^T V^T, \sigma u_3) = (UZV^T, \sigma u_3) \begin{pmatrix} VZ^T Z^T V^T & 0 \\ 0^T & 1 \end{pmatrix}$$

由于 $VZ^T Z^T V^T = V \text{diag}(-1, -1, 1) V^T$, 因此 $VZ^T Z^T V^T$ 是绕轴 v_3 的 180 度旋转矩阵。在第二个摄像机坐标系中, 两个摄像机的基线方向是 $(UZ^T V^T)^T u_3 = VZU^T u_3 = v_3$ 。因此, 第三组摄像机是第一组摄像机绕两个摄像机的基线旋转 180 度所导致的结果, 通常称它们为一对“扭对”, 如图 11.1.1 所示。

欧氏结构

由运动参数所获得摄像机对的四个解的几何解释，可以看出只有一个解是合理的，这个合理解可以通过一个点对应的重构来确定，即这个重构点同时位于两个摄像机前方的一对摄像机是合理的，也就是说重构点在两个摄像机坐标系下的 z 坐标都大于零的摄像机对是我们所需要的解。

令 $P=(I, 0)$, $P'=(R, \sigma^{-1}t)$ 是合理的摄像机对， $\{m_j \leftrightarrow m'_j\}$ 是点对应集(这里的点对应是指图像的归一化坐标)，应用第 6.1.1 节三角原理的数值算法，就可以得到这些点对应的三维重构。值得注意的是，这里的空间点坐标是在第一个摄像机坐标系中的坐标，并且相差一个整体尺度因子 σ^{-1} ，因此重构的空间点坐标与它的真实坐标(世界坐标系中的坐标)相差一个相似变换。

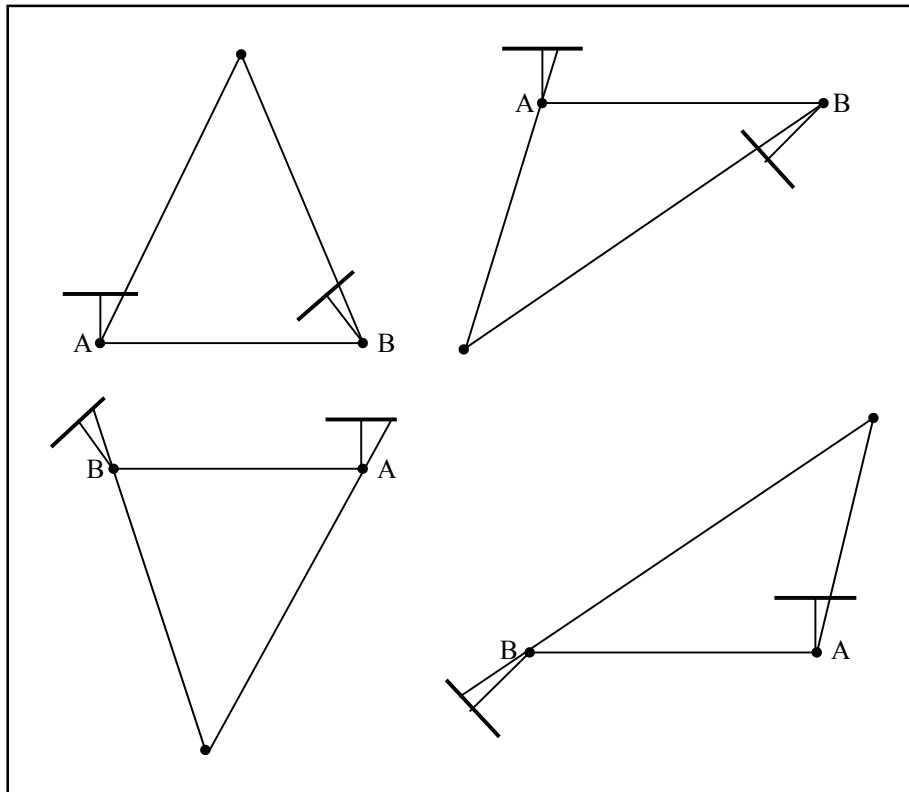


图 11.1.1：运动参数四组解的几何解释：在左右两列中第二个摄像机(B)具有相同旋转，而平移方向是相反的。在上下两行中，第二个摄像机(B)平移方向是相同的，而旋转是绕基线的 180 度旋转，通常称它们构成两个“组对”。在这四组解中只是一组重构的空间点都位于两个摄像机的前方，如图中的第一组解，它是一个在物理上可实现的解，也就是说仅有一组解是合理的。

11.2 仿射运动与结构

11.2.1 仿射摄像机

在第三章，我们讨论了光心在有限点的射影摄像机，这种摄像机简称有限摄像机。有限摄像机的特征是其摄像机矩阵的前三列构成一个可逆矩阵。光心在无穷远的摄像机，简称无穷远摄像机。因无穷远摄像机的光心在无穷远平面上，所以这种摄像机矩阵的前三列的 3×3 子矩阵是奇异的，其投影是一个平行投影。无穷远摄像机可以分为两种不同的类型：仿射摄像机（Affine Camera）和非仿射摄像机。在实践中常用的是仿射摄像机，这里介绍仿射摄像机。

定义 11.2.1 如果摄像机矩阵 P 具有下述形式：

$$P_A = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ 0 & 0 & 0 & 1 \end{pmatrix} \triangleq \begin{pmatrix} H_{2 \times 3} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{pmatrix} \quad (11.2.1)$$

则称它为仿射摄像机。

仿射变换的特征是将无穷远平面映射到无穷远平面，仿射摄像机与三维仿射变换的差别在于它将无穷远平面映射到二维平面(像平面)上的无穷远直线。这是很明显的，因为 $P_A(x, y, z, 0)^T = (x, y, 0)^T$ 。利用有限透视摄像机的术语，可以将仿射摄像机的主平面解释成是无穷远平面。由于光心在主平面上，所以主平面的图像必然也在无穷远平面上。由此，可以得到下述结论：

1. 主平面是无穷远平面的任何(射影)摄像机均为仿射摄像机；
2. 平行直线在仿射摄像机下投影仍为平行直线。这是因为平行直线相交于无穷远平面，而该交点又被映射到图像平面中的无穷远点。因此图像直线是平行的；
3. 满足 $H_{2 \times 3} \mathbf{d} = 0$ 的向量 \mathbf{d} 是平行投影的方向，而 $(\mathbf{d}^T, 0)^T$ 是仿射摄像机光心(因为 $P_A(\mathbf{d}^T, 0)^T = 0$)。

仿射摄像机的分层

从最简单的平行投影开始，可以为仿射摄像机模型建立层次关系。

正交投影(Orthographic Projection) 考虑沿 z 轴的投影，下面是它的矩阵形式表示：

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (11.2.2)$$

它是最简单的平行投影, 即将点 $(x, y, z, 1)^T$ 投影到图像点 $(x, y, 1)^T$, 即与 z 坐标无关。

为得到一般正交投影, 可用一个欧氏坐标变换右乘(11.2.2)得到:

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} R & \mathbf{t} \\ 0^T & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{r}^{1T} & t_1 \\ \mathbf{r}^{2T} & t_2 \\ 0^T & t_3 \end{pmatrix} \quad (11.2.3)$$

它有 5 个自由度。

伸缩正交投影(Scaled Orthographic Projection) 伸缩正交投影是正交投影的图像再进行一个均匀伸缩变换, 如伸缩因子为 k , 则伸缩正投影摄像机矩阵可写成下面的形式

$$P = \begin{pmatrix} k & & & \\ & k & & \\ & & & \\ & & & 1 \end{pmatrix} \begin{pmatrix} \mathbf{r}^{1T} & t_1 \\ \mathbf{r}^{2T} & t_2 \\ 0^T & 1 \end{pmatrix} \quad (11.2.4)$$

它有六个自由度。伸缩正交投影摄像机矩阵 $P=(M, \mathbf{t})$ 的特征是矩阵 M 最后一行是零, 而前两行正交并且模相等。

弱透视投影(Weak Perspective) 与有限 CCD 摄像机类似, 考虑在图像两个轴的方向上作不一致的伸缩, 如果伸缩因子分别为 f_x, f_y , 则摄像机矩阵具有下面的形式

$$P = \begin{pmatrix} f_x & & & \\ & f_y & & \\ & & & \\ & & & 1 \end{pmatrix} \begin{pmatrix} \mathbf{r}^{1T} & t_1 \\ \mathbf{r}^{2T} & t_2 \\ 0^T & 1 \end{pmatrix} \quad (11.2.5)$$

它有七个自由度。通常称它为弱透视摄像机。弱透视摄像机矩阵 $P=(M, \mathbf{t})$ 的特征是矩阵 M 最后一行是零, 而前两行正交 (但不要求如伸缩正交投影那样有相等的模)。

仿射摄像机 如果对矩阵的元素没有限制的话, 一般的仿射摄像机矩阵可以分解成:

$$P_A = \begin{pmatrix} f_x & s & 0 \\ 0 & f_y & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{r}^{1T} & t_1 \\ \mathbf{r}^{2T} & t_2 \\ 0^T & 1 \end{pmatrix} \quad (11.2.6)$$

它有八个自由度。

仿射摄像机的最一般形式可以表示为式(11.2.6), 对仿射摄像机的唯一限制是 $H_{2 \times 3}$ 的秩为 2, 它来自仿射摄像机矩阵秩等于 3 的要求。仿射摄像机综合了三个变换的效果, 一个三维空间的仿射变换、一个从三维空间到图象的正交投影和一个图象平面的仿射变换。将这些变换矩阵连接起来可得

到仿射摄像机的下述分解形式:

$$P_A = [3 \times 3 \text{ 仿射}] \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} [4 \times 4 \text{ 仿射}] \quad (11.2.7)$$

在非齐次坐标下, 仿射摄像机的投影是一个线性变换与一个平移变换的和, 即

$$\tilde{\mathbf{m}} = H_{2 \times 3} \mathbf{X} + \tilde{\mathbf{t}} \quad (11.2.8)$$

其中 $\tilde{\mathbf{t}} = (t_1, t_2)^T$ 是世界坐标原点的图像。

11.2.2 仿射运动与结构

给定 n 幅图像的一组点对应 $\{\tilde{\mathbf{m}}_{ij} \mid i=1, \dots, n; j=1, \dots, m\}$, 并且已知 n 这幅图像分别来自仿射摄像机 $P_{A_i}, i=1, 2, \dots, n$ 的投影, 但摄像机矩阵是未知的。仿射运动与结构问题, 这从这组点对应求解它们对应的空间点 $\{\tilde{\mathbf{X}}_j\}$ 和摄像机矩阵 $\{P_{A_i}\}$ 。从仿射摄像机的一组图像点对应获得的结构是景物的仿射结构, 所以这样得到结构也称为仿射重构。

空间点 $\mathbf{X} = (\tilde{\mathbf{X}}^T, 1)^T$, 在摄像机 P_A 下的图像点是

$$\begin{pmatrix} \tilde{\mathbf{m}} \\ 1 \end{pmatrix} = \begin{pmatrix} H_{2 \times 3} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{X}} \\ 1 \end{pmatrix} \quad (11.2.9)$$

由此, 可以看出仿射摄像机的每一个图像点的深度都等于 1。如果空间点和图像点均使用非齐次坐标, 则仿射摄像机的投影关系式(11.2.9)可写成

$$\tilde{\mathbf{m}} = H\tilde{\mathbf{X}} + \mathbf{t} \quad (11.2.10)$$

因此, 现在的问题是求解空间点坐标 $\{\tilde{\mathbf{X}}_j\}$ 和摄像机矩阵 $\{P_{A_i}\}$ 使得

$$\min \sum_{i=1}^n \sum_{j=1}^m d^2 \left((H_i \tilde{\mathbf{X}}_j + \mathbf{t}_i), \tilde{\mathbf{m}}_{ij} \right) \quad (11.2.11)$$

其中 m_{ij} 为第 i 幅图像上的第 j 个图像点的非齐次坐标的测量, $d(\cdot, \cdot)$ 是图像点的几何距离。式(11.2.11)表示重投影的几何误差最小。

为了简化计算, 对每幅图像坐标进行变换, 使得图像坐标的原点位于该幅图像点集

$\{\tilde{\mathbf{m}}_{ij} \mid j=1, \dots, m\}$ 的重心 $\bar{\mathbf{m}}_i = \frac{1}{m} \sum_{j=1}^m \tilde{\mathbf{m}}_{ij}$, 即对第 i 幅图像坐标进行变换 $\mathbf{m}'_{ij} = \tilde{\mathbf{m}}_{ij} - \bar{\mathbf{m}}_i$ 。由于

$$\mathbf{m}'_{ij} = (H_i \tilde{\mathbf{X}}_j + \mathbf{t}_i) - (H_i \bar{\mathbf{X}} + \mathbf{t}_i) = H_i \Delta \mathbf{X}_j$$

其中, $\Delta \mathbf{X}_j = \tilde{\mathbf{X}}_j - \bar{\mathbf{X}}, \bar{\mathbf{X}} = \frac{1}{m} \sum_{j=1}^m \tilde{\mathbf{X}}_j$ 。因此, 不妨假定世界坐标系的原点在空间点集的重心。这样, 摄像机矩阵可简化为 $P_{A_i} = H_i$ 。

在下面的讨论中, 假定图像已经作了上述坐标变换。现在所要求解的问题是: 给定 n 幅图像的一组点对应 $\{m_{ij} | i = 1, \dots, n; j = 1, \dots, m\}$, 求每幅图像的摄像机矩阵和对应点的空间坐标 $[H_i, \{\mathbf{X}_j\}]$, 并使得

$$\min \sum_{i=1}^n \sum_{j=1}^m d^2(H_i \mathbf{X}_j, m_{ij}) \quad (11.2.12)$$

记

$$W = \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1m} \\ m_{21} & m_{22} & \cdots & m_{2m} \\ \vdots & \vdots & & \vdots \\ m_{n1} & m_{n2} & \cdots & m_{nm} \end{pmatrix} \quad (11.2.13)$$

是由变换后图像测量点的坐标组成的 $2n \times m$ 矩阵, 并称它为测量矩阵。于是等式:

$$m_{ij} = H_i \mathbf{X}_j, \quad i = 1, 2, \dots, n; j = 1, 2, \dots, m$$

可写成如下更简洁的矩阵形式:

$$W = \begin{pmatrix} H_1 \\ H_2 \\ \vdots \\ H_n \end{pmatrix} (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_m) \quad (11.2.14)$$

这里, W 是已知的, H_i, \mathbf{X}_j 是需要求解的。在数据有噪声时, 方程(11.2.14)没有完全被满足的解。因此, 需要的是满足式(11.2.12)的最优解。

满足式(11.2.12)的最优解与下述问题是等价的: 在 F -范数下, 寻求 W 的一个最佳近似 \hat{W} , 并且有如下分解:

$$\hat{W} = \begin{pmatrix} \hat{m}_{11} & \hat{m}_{12} & \cdots & \hat{m}_{1m} \\ \hat{m}_{21} & \hat{m}_{22} & \cdots & \hat{m}_{2m} \\ \vdots & \vdots & & \vdots \\ \hat{m}_{n1} & \hat{m}_{n2} & \cdots & \hat{m}_{nm} \end{pmatrix} = \begin{pmatrix} \hat{H}_1 \\ \hat{H}_2 \\ \vdots \\ \hat{H}_n \end{pmatrix} (\hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2, \dots, \hat{\mathbf{X}}_m) \equiv \hat{H} \hat{\mathbf{X}}$$

这是因为

$$\|W - \hat{W}\|_F^2 = \sum_{i,j} (w_{ij} - \hat{w}_{ij})^2 = \sum_{i,j} d^2(m_{ij} - \hat{m}_{ij}) = \sum_{i,j} d^2(\hat{H}_i \hat{\mathbf{X}}_j, m_{ij})$$

考虑到 \hat{W} 是一个 $2n \times 3$ 的摄像机的运动矩阵 \hat{H} 与 $3 \times m$ 的结构矩阵^{*)} \hat{X} 的乘积，即它是一个秩 3 的矩阵，所以问题归结为在 F 范数下，寻求 W 的一个秩 3 的最佳近似矩阵 \hat{W} 。这可以通过 W 的 SVD 分解来得到。更确切地说，若 W 的 SVD 分解为 $W=UDV$ ，则 $\hat{W} = U_{2n \times 3} D_{3 \times 3} V_{m \times 3}^T$ ，其中 $U_{2n \times 3}$ 是 U 的前 3 列所构成的矩阵， $D_{3 \times 3}$ 是 D 的前 3 个对角元构成的对角矩阵， $V_{m \times 3}$ 是 V 的前 3 列所构成的矩阵。于是，摄像机的运动矩阵和空间结构可分别选择为 $H = U_{2n \times 3}$ ， $X = D_{3 \times 3} V_{m \times 3}^T$ 。

注意：解的选择不是唯一的，例如 $H = U_{2n \times 3} D_{3 \times 3}$ ， $X = V_{m \times 3}^T$ 也可以作为一组解，更一般地，对任意 3×3 非奇异矩阵 A ，若 \hat{H}, \hat{X} 是一组解，则 $H = \hat{H}A^{-1}$ ， $X = A\hat{X}$ 也是一组解，这是因为总有 $\hat{W} = (\hat{H}A^{-1})(A\hat{X}) = \hat{H}\hat{X}$ 。在非齐次坐标下可逆矩阵 A 是一个仿射变换，这个结果与仿射结构之间相差一个仿射变换的结论是一致的。

11.3 射影运动与结构

关于两幅图像，可以从二幅图像的基本矩阵 $F = [e']_x H$ 得到两个摄像机矩阵 $P = (I, \theta)$ ， $P' = (H, e')$ 。于是，由点对应 $m \leftrightarrow m'$ ，应用第 6.1 节的三角原理的数值算法就可以得到空间点 X ，这里的空间点 X 不是欧氏空间的重构，而是射影空间的重构。对于三幅以上的图像，通常都是用基本矩阵或三焦张量来实现射影重构。其主要思想是以某两幅(或三幅)图像为基础，使用基本矩阵(或三阶张量)得到它们的射影重构，再以它的射影坐标系作为基准坐标系恢复其图像所对应的摄像机矩阵并计算点对应的空间坐标。下面以基本矩阵为例来说明这种重构方法，对于三焦张量可以给出类似的方法。重构过程主要分为以下三步：

(1) 求某两幅图像的射影重构

例如从第一、二幅图像的点对应，应用这两幅图像间的基本矩阵 $F = [e_2]_x H$ ，可得到第一、二幅图像的摄像机矩阵 $P_1 = (I, \theta)$ ， $P_2 = (H, e_2)$ ，再利用三角原理计算两幅图像点对应的空间点集 $\{X\}$ 。于是，就得到这两幅图像的一个射影重构：

$$[P_1 = (I, \theta), P_2 = (H, e_2), \{X\}]$$

(2) 求其它图像在以上述两幅图像为基准的射影重构空间中的摄像机矩阵

令 $X_k^{(j)} \in \{X\}$ 是第 j 幅图像中的可见空间点，这样的空间点可以由第 1, 2, j 三幅图像间的点对应来确定，要求至少存在 6 个这样的空间点。令 $X_k^{(j)}$ 在第 j 幅图像中的投影点为 $m_k^{(j)}$ ，于是我们得到 $P_j X_k^{(j)} = s_{jk} m_k^{(j)}$ ，即 $m_k^{(j)} \times P_j X_k^{(j)} = \theta$ ，应用第 3.1.3 节的方法计算出摄像机矩阵 P_j （尽管第 3.1.3

^{*)} 由每个摄像机矩阵作为子矩阵的矩阵 \hat{M} 通常称为运动矩阵，空间点坐标所组成的矩阵通常简称为结构矩阵。

节是在欧氏坐标系下给出计算方法, 但这种方法对任意坐标系都是适用的)。

(3) 计算空间点的射影坐标

令 \mathbf{X} 是某 $k(\geq 2)$ 个图像中都出现的空间点, 从下述方程组计算空间点 \mathbf{X} 的坐标:

$$\mathbf{m}^{(j_l)} \times P_{j_l} \mathbf{X} = \mathbf{0}, \quad l = 1, 2, \dots, k$$

其中 $\mathbf{m}^{(j_l)}$, P_{j_l} 分别是 \mathbf{X} 在第 j_l 幅中的图像坐标和第 j_l 幅图像的摄像机矩阵, 它们都是已知的。

通过以上三步, 就得到了 n 幅图像的一个射影重构。这种重构技术的特点是十分简单易实现, 但其主要不足之处在于重构结果严重依赖于初始(基准)二幅图像的选择。下面所介绍的基于 SVD 分解的算法在一定程度上克服了这种重构技术的不足。

11.3.1 基本原理

上一节的 SVD 分解方法不能直接应用于(射影摄像机的)射影重构。在射影摄像机模型下, 每一个图像点的“射影深度”是未知的, 从而对应于仿射重构的矩阵分解方法中的测量矩阵含有未知深度参数, 所以不能直接使用矩阵分解方法来求解。但是, 如果已知射影深度, 则用 SVD 分解方法可以求解摄像机投影矩阵和空间结构, 并且与仿射摄像机分解方法是完全相同的。

考虑一组图像点对应 $\{\mathbf{m}_{ij} \mid i = 1, 2, \dots, n; j = 1, 2, \dots, m\}$, 则摄像机矩阵 P_i 和空间结构 \mathbf{X}_j 满足投影关系:

$$s_{ij} \mathbf{m}_{ij} = P_i \mathbf{X}_j, \quad i = 1, 2, \dots, n; j = 1, 2, \dots, m \quad (11.3.1)$$

其中 s_{ij} 是第 i 幅图像上第 j 个点的射影深度, $\mathbf{m}_{ij} = (x_{ij}, y_{ij}, 1)^T$ 是第三个分量为 1 的图像点的齐次表示并且 $(x_{ij}, y_{ij})^T$ 是由实际测量所得到的图像坐标(在下面分解算法的讨论中, 始终遵循这个约定)。我们假定每个空间点在每个摄像机中都是可见的, 因此对所有 i, j , \mathbf{m}_{ij} 都是已知的。令

$$W(s) = \begin{pmatrix} s_{11} \mathbf{m}_{12} & s_{12} \mathbf{m}_{12} & \cdots & s_{1m} \mathbf{m}_{1m} \\ s_{21} \mathbf{m}_{21} & s_{22} \mathbf{m}_{22} & \cdots & s_{2m} \mathbf{m}_{2m} \\ \vdots & \vdots & & \vdots \\ s_{n1} \mathbf{m}_{n1} & s_{n2} \mathbf{m}_{n2} & \cdots & s_{nm} \mathbf{m}_{nm} \end{pmatrix} \quad (11.3.2)$$

并称为测量矩阵。于是, 方程(11.3.2)可写成矩阵形式:

$$W(s) = (P_1^T, P_2^T, \dots, P_n^T)(X_1, X_2, \dots, X_m) \quad (11.3.3)$$

上述方程仅当每一个测量点 \mathbf{m}_{ij} 和深度 s_{ij} 完全一致时才成立。

假定已给出了正确的射影深度 s_{ij} 。对于精确的数据点，左边的矩阵 W 的秩是 4，因为它是 4 列和 4 行的两个矩阵的乘积。由于数据误差，即使有正确的射影深度，实际测量矩阵 W 也不可能为 4，因此需要用一个秩为 4 的矩阵作 W 的最佳近似。与仿射分解一样，这可以利用 SVD 分解来实现：如果 $W=UDV^T$ ，令 \hat{D} 是 D 的前 4 个对角元素所构成的对角矩阵， \hat{U}, \hat{V} 分别为 U, V 的前 4 列所构成的矩阵，则测量矩阵 W 的最佳近似是 $\hat{W} = \hat{U}\hat{D}\hat{V}^T$ 。于是，摄像机矩阵 $(P_1^T, P_2^T, \dots, P_n^T) = \hat{U}\hat{D}$ ，结构矩阵 $(X_1, X_2, \dots, X_m) = \hat{V}^T$ 。与前面的仿射重构一样，这个分解不是唯一的，事实上在(11.3.3)右边的两个矩阵之间我们可以插入一个任意的 3 维射影变换矩阵 H 和它的逆 H^{-1} ，这与两个射影重构都相差一个射影变换的结论是一致的。

11.3.2 选择射影深度

从前面的讨论可以看出，射影矩阵分解算法的主要困难在于选取正确的射影深度。首先，注意到射影深度 s_{ij} 不是唯一的。事实上，假设 $s_{ij}m_{ij} = P_i X_j$ ，用 $\alpha_i P_i$ 代替 P_i ，用 $\beta_j X_j$ 代替 X_j ，则有

$$(\alpha_i \beta_j s_{ij}) m_{ij} = (\alpha_i P_i)(\beta_j X_j)$$

换句话说，射影深度 s_{ij} 可以用一个因子 α_i 乘(11.2.1)的第 i 行和用一个因子 β_j 乘第 j 列来代替。于是，可以约定(11.2.1)的第 1 行和第一列的所有深度均为 1，也就是说测量矩阵具有下述形式：

$$W(s) = \begin{pmatrix} m_{12} & m_{12} & \cdots & m_{1m} \\ m_{21} & s_{22}m_{22} & \cdots & s_{2m}m_{2m} \\ \vdots & \vdots & & \vdots \\ m_{n1} & s_{n2}m_{n2} & \cdots & s_{nm}m_{nm} \end{pmatrix} \quad (11.3.4)$$

值得指出的是，基于矩阵分解的射影重构与基于矩阵分解的仿射重构的根本区别是在射影重构时，首先需要将测量矩阵中的射影深度一致化。一旦射影深度得到了一致化，则后面的分解方法与仿射重构一样了。所谓“一致化的射影深度” s_{ij} ，就是指在图像点无噪声的情况下，对应于 s_{ij} 的测量矩阵的秩为 4。从单个图像点来说，我们知道尺度因子(即这里的深度因子)是不确定的，可以自由地选取。但图像点的尺度因子之间是内在关联的，矩阵分解方法正是利用了这种内在关联关系来达到射影重构。下面介绍一种由基本矩阵确定一致化射影深度的方法。

假定已经求出第 1 幅图像与第 k 幅图像之间的基本矩阵 F_{1k} 和第 k 幅图像上关于第一幅图像的极点 e_{1k} 。由 $m_{1j} = P_1 X_j$, $s_{kj} m_{kj} = P_k X_j$, $j = 1, 2, \dots, m$ ，可以推出：

$$s_{kj} m_{kj} = P_k P_1^+ m_{1j} + s P_k C_1$$

其中 C_1 是第 1 个摄像机的光心。于是，消去常数因子 s 后可以得到

$$s_{kj}[P_k C_1]_x m_{kj} = [P_k C_1]_x P_k P_1^+ m_{1j}$$

由于 $F_{1k} = \alpha [P_k C_1]_x P_k P_1^+$, $[e_k]_x = \beta [P_k C_1]_x$, 所以有

$$s_{kj}[e_k]_x m_{kj} = \frac{\beta}{\alpha} F_{1k} m_{1j}$$

因此

$$s_{kj} = \frac{\beta}{\alpha} \cdot \frac{([e_k]_x m_{kj})^T F_{1k} m_{1j}}{([e_k]_x m_{kj})^T ([e_k]_x m_{kj})}, \quad j=1,2,\dots,m$$

因为 W 每一行的深度因子可以相差一个共同的尺度因子, 所以可令:

$$s_{k1} = 1, \quad \frac{\beta}{\alpha} = \frac{([e_k]_x m_{k1})^T ([e_k]_x m_{k1})}{([e_k]_x m_{kj})^T (F_{1k} [e_k]_x m_{11})}$$

这样, 就得到了确定射影深度因子的计算公式:

$$s_{kj} = \frac{([e_k]_x m_{k1})^T ([e_k]_x m_{k1})}{([e_k]_x m_{k1})^T F_{1k} m_{11}} \cdot \frac{([e_k]_x m_{kj})^T F_{1k} m_{1j}}{([e_k]_x m_{kj})^T ([e_k]_x m_{kj})}, \quad j=2,3,\dots,m; k=2,3,\dots,n \quad (11.3.5)$$

11.3.3 迭代分解算法

下面介绍一种估计深度的迭代算法, 这种算法反复使用 SVD 分解技术, 在迭代结束时也同时给出了摄像机矩阵和射影结构, 这种方法首先是由 Ushiba 和 Tomida 首先提出来的。

在图像坐标无噪声且射影深度为一致化的理想情况下, 测量矩阵 W 的秩 $\text{rank}W = 4$, 当有噪声但不大时, 对于一致化的深度, $\text{rank}W \approx 4$ 。因此, 可以定义一个代价函数:

$$J = \sum_{n=5}^N \sigma_n^2 \quad (11.3.6)$$

其中, σ_n 是 W 的第 n 个奇异值。 J 是依赖于 s_{ij} 的非负实数。于是求使 J 取最小值的 $\{s_{ij}\}$ 将给出所需要的射影深度。由于 J 不是 s_{ij} 的线性函数, 故最小化 J 是一个非线性优化问题。下面介绍一种解决此问题的共轭梯度法。

共轭梯度方法的前提是计算 J 关于 s_{ij} 的梯度。对 W 作奇异值分解: $W=UDV^T$, 则必有 $\sigma_n = u_n^T W v_n$, 其中 u_n , v_n 分别为正交矩阵 U 和 V 的第 n 列。于是,

$$\frac{\partial \sigma_n}{\partial s_{ij}} = \frac{\partial u_n^T}{\partial s_{ij}} W v_n + u_n^T \frac{\partial W}{\partial s_{ij}} v_n + u_n^T W \frac{\partial v_n}{\partial s_{ij}} = \sigma_n \frac{\partial u_n^T}{\partial s_{ij}} u_n + u_n^T \frac{\partial W}{\partial s_{ij}} v_n + \sigma_n v_n^T \frac{\partial v_n}{\partial s_{ij}}$$

由 $u_n^T u_n = v_n^T v_n = 1$, 可知

$$\frac{\partial \mathbf{v}_n^T}{\partial s_{ij}} \mathbf{v}_n = \mathbf{u}_n^T \frac{\partial \mathbf{u}_n}{\partial s_{ij}} = 0$$

因此,

$$\frac{\partial \sigma_n}{\partial s_{ij}} = \mathbf{u}_n^T \frac{\partial W}{\partial s_{ij}} \mathbf{v}_n$$

由此, 得到 J 关于 s_{ij} 的梯度:

$$\frac{\partial J}{\partial s_{ij}} = 2 \sum_{n=5}^N \sigma_n \frac{\partial \sigma_n}{\partial s_{ij}} = 2 \sum_{n=5}^N \sigma_n \mathbf{u}_n^T \frac{\partial W}{\partial s_{ij}} \mathbf{v}_n$$

迭代算法需要射影深度 s_{ij} 的一个初始估计, 这可以用其它方法如射影重构技术得到一个初始估计, 或者以所有的深度均为 1 作为初始值(这表示用仿射摄像机来近似射影摄像机)。另外, 在迭代估计射影深度时, 应以测量指标 J 作为结束循环迭代的判据。这是由于在噪声存在的条件下, 会出现某一步迭代增量变得很小而 J 却相对很大的情况, 因此以测量指标 J 作为结束循环迭代的判据比通常以迭代增量作为判据更为合理、有效。另外, 在迭代算法之前, 对图像数据作规范化处理是必要的。最后, 值得指出的是迭代算法不能保证收敛结果达到全局最小。

12. 多视点张量

第 4 章讨论了两幅图像间的极几何。极几何给出了两幅图像间的点对应关系，这个关系由基本矩阵来刻画。事实上，基本矩阵所刻画极几何关系是点对应的双重线性关系。对于多幅图像间的点对应、线对应以及点线对应也有类似的关系，即多重线性关系，刻画多重线性关系的数学工具是张量。与基本矩阵类似多幅图像的张量也是空间射影变换的不变量，也就是说，一旦获得多幅图像的张量，在理论上能得到射影意义下的摄像机矩阵。本章重点讨论三幅图像和四幅图像间的张量。为了更好地理解它们，首先用张量来描述两幅图像间的基本矩阵，然后再引进三幅图像和四幅图像间的张量。

12.1 双线性关系

给定两幅图像 (I, I') ，令 $\mathbf{x} \leftrightarrow \mathbf{x}'$ 是任一点对应，第 4.1 节的极几何指出 $l' = F\mathbf{x}$ 是图像点 \mathbf{x} 对应的极线，且图像点 \mathbf{x}' 在极线 l' ；同样， $l = F^T \mathbf{x}'$ 是 \mathbf{x}' 对应的极线，且图像点 \mathbf{x} 在极线 l 上，其中 F 是两幅图像间的基本矩阵。这个结论可等价地表述为：

$$\mathbf{x}'^T F \mathbf{x} = 0 \quad (12.1.1)$$

图像点的记坐标为 $\mathbf{x} = (x^1, x^2, x^3)^T$ ， $\mathbf{x}' = (x'^1, x'^2, x'^3)^T$ ，使用张量的语言，图像点是一阶逆变张量；基本矩阵记作 $F = (f_{ij})$ 。则式(12.1.1)可以写成：

$$f(\mathbf{x}, \mathbf{x}') = x'^i x^j f_{ij} = 0 \quad (12.1.2)$$

所以基本矩阵 $F = (f_{ij})$ 确定了图像(点)空间上的一个双线性函数。如果 $\mathbf{x} \leftrightarrow \mathbf{x}'$ 是图像空间上的一个点对应，则这个双线性函数在点对应上取零值；反之如果图像空间的一对点 $(\mathbf{x}, \mathbf{x}')$ 使这个双线性函数取零值(两幅图像的极点除外)，则必存在一个空间点 \mathbf{X} 使得它在两幅图像上的投影分别为 \mathbf{x}, \mathbf{x}' ，即 $\mathbf{x} \leftrightarrow \mathbf{x}'$ 是一个点对应。这就是所谓的两幅图像点对应的双线性关系。

如果直线的坐标记为 $l = (l_1, l_2, l_3)^T$ ， $l' = (l'_1, l'_2, l'_3)^T$ ，即直线用一阶协变张量表示， $l' = F\mathbf{x}$ 和

$l = F^T x'$ 可表示为

$$l'_i = x^j f_{ij}, l_j = x'^i f_{ij}$$

因此, 基本矩阵给出的点与极线的关系是图像(点)空间到图像(线)空间的线性映射。由于这个线性映射将 2 维图像点映射到过极点的线束, 所以这个线性映射的秩是 2。

综上所述基本矩阵是一个 2 阶协变张量, 它给出了点对应的双线性函数关系以及点与极线之间的线性映射关系。

12.1.1 基本矩阵的张量形式

根据 4.1 节的讨论, 两幅图像间的基本矩阵可表示为 $F = [e']_x P' P^+$, 其中 $\{P, P'\}$ 是两幅图像的摄像机矩阵。为了更好地理解三幅图像间的 3 重线性关系, 下面将利用摄像机矩阵 $\{P, P'\}$ 显式地表述基本矩阵的各个元素。

令两幅图像对应的摄像机矩阵 $\{P, P'\}$, 根据摄像机的投影关系空间点到图像点之间的映射为:

$$sx = PX, s'x' = P'X$$

其中 s, s' 是非零常数因子。上述方程组可改述为下面的矩阵形式:

$$\begin{pmatrix} p^1 & x^1 & 0 \\ p^2 & x^2 & 0 \\ p^3 & x^3 & 0 \\ p'^1 & 0 & x'^1 \\ p'^2 & 0 & x'^2 \\ p'^3 & 0 & x'^3 \end{pmatrix} \begin{pmatrix} X \\ -s \\ -s' \end{pmatrix} = 0 \quad (12.1.3)$$

其中 p^j, p'^j 分别表示摄像机矩阵 P, P' 的第 j 行所构成的行向量。记方程(12.1.3)的系数矩阵为

$$X = \begin{pmatrix} p^1 & x^1 & 0 \\ p^2 & x^2 & 0 \\ p^3 & x^3 & 0 \\ p'^1 & 0 & x'^1 \\ p'^2 & 0 & x'^2 \\ p'^3 & 0 & x'^3 \end{pmatrix}$$

它是一个 6×6 矩阵。由于方程(12.1.3)有非零解, 所以必有 $\det(X) = 0$ 。这个等式将给出基本矩阵各个元素的表达式。

将按第 5 列展开求 X 的行列式, 得到:

$$\det(X) = x^1 \det \begin{pmatrix} p^2 & 0 \\ p^3 & 0 \\ P' & x' \end{pmatrix} - x^2 \det \begin{pmatrix} p^1 & 0 \\ p^3 & 0 \\ P' & x' \end{pmatrix} + x^3 \det \begin{pmatrix} p^1 & 0 \\ p^2 & 0 \\ P' & x' \end{pmatrix}$$

从上式可以看出在 X 行列式的完全展开式中，仅包含形如 $x^j x'^i$ 的项。并且 $x^j x'^i$ 的系数可以按下述方式求出：分别删除矩阵 X 中包含 x^j, x'^i 的行和列得到 X 的 4×4 子矩阵，再对这个子矩阵取行列式并乘以适当的正负号。例如： $x^1 x'^1$ 的系数是：

$$f_{11} = \det \begin{pmatrix} p^2 \\ p^3 \\ p'^2 \\ p'^3 \end{pmatrix}$$

一般地， $x^j x'^i$ 的系数为：

$$f_{ji} = (-1)^{i+j} \det \begin{pmatrix} P(\sim i) \\ P'(\sim j) \end{pmatrix} \quad (12.1.4)$$

其中符号 $C(\sim k)$ 表示从矩阵 C 中删除第 k 行剩下的子矩阵。这样，就得到了两幅图像间的双线性关系：

$$x'^j x^i f_{ji} = 0$$

对于式(12.1.4)还可以给出另一种表达：

$$f_{ji} = \left(\frac{1}{4}\right) \varepsilon_{ikl} \varepsilon_{jmn} \det \begin{pmatrix} p^k \\ p^l \\ p'^m \\ p'^n \end{pmatrix} \quad (12.1.5)$$

这是不难验证的，例如：

$$\begin{aligned} f_{12} &= \left(\frac{1}{4}\right) \varepsilon_{1kl} \varepsilon_{2mn} \det \begin{pmatrix} p^k \\ p^l \\ p'^m \\ p'^n \end{pmatrix} \\ &= \left(\frac{1}{4}\right) \left(\varepsilon_{123} \varepsilon_{213} \det \begin{pmatrix} p^2 \\ p^3 \\ p'^1 \\ p'^3 \end{pmatrix} + \varepsilon_{123} \varepsilon_{231} \det \begin{pmatrix} p^2 \\ p^3 \\ p'^3 \\ p'^1 \end{pmatrix} + \varepsilon_{132} \varepsilon_{213} \det \begin{pmatrix} p^3 \\ p^2 \\ p'^1 \\ p'^3 \end{pmatrix} + \varepsilon_{132} \varepsilon_{231} \det \begin{pmatrix} p^3 \\ p^2 \\ p'^3 \\ p'^1 \end{pmatrix} \right) \end{aligned}$$

$$\begin{aligned}
 &= \left(\frac{1}{4}\right) \left(\varepsilon_{123} \varepsilon_{213} \det \begin{pmatrix} p^2 \\ p^3 \\ p'^1 \\ p'^3 \end{pmatrix} + \varepsilon_{123} (-\varepsilon_{213}) \left(-\det \begin{pmatrix} p^2 \\ p^3 \\ p'^1 \\ p'^3 \end{pmatrix} \right) + (-\varepsilon_{123}) \varepsilon_{213} \left(-\det \begin{pmatrix} p^2 \\ p^3 \\ p'^1 \\ p'^3 \end{pmatrix} \right) + (-\varepsilon_{123}) (-\varepsilon_{213}) \det \begin{pmatrix} p^2 \\ p^3 \\ p'^1 \\ p'^3 \end{pmatrix} \right) \\
 &= \varepsilon_{123} \varepsilon_{213} \det \begin{pmatrix} p^2 \\ p^3 \\ p'^1 \\ p'^3 \end{pmatrix} = (-1)^{1+2} \det \begin{pmatrix} p^2 \\ p^3 \\ p'^1 \\ p'^3 \end{pmatrix}
 \end{aligned}$$

类似地，可验证其它项。

从表达式(12.1.5)或(12.1.4)，很容易说明基本矩阵不依赖于摄像机矩阵的选择，等价地说不依赖于世界坐标系的选择。假定 $\{Q, Q'\}$ 是在另一种世界坐标系下的两幅图像对应的摄像机矩阵，由于任何两个坐标系之间变换矩阵是一个 4 阶可逆矩阵 M (两个世界坐标系之间的射影变换)，因此，有

$$P = QM, P' = Q'M$$

于是，

$$f_{ji} = \left(\frac{1}{4}\right) \varepsilon_{ikl} \varepsilon_{jmn} \det \begin{pmatrix} p^k \\ p^l \\ p'^m \\ p'^n \end{pmatrix} = \left(\frac{1}{4}\right) \varepsilon_{ikl} \varepsilon_{jmn} \det \begin{pmatrix} q^k M \\ q^l M \\ q'^m M \\ q'^n M \end{pmatrix} = \left(\frac{\det(M)}{4}\right) \varepsilon_{ikl} \varepsilon_{jmn} \det \begin{pmatrix} q^k \\ q^l \\ q'^m \\ q'^n \end{pmatrix}$$

所以，在两种坐标系下，式(12.1.5)仅相差一个不依赖于 (i, j) 的公共比例因子。因此，基本矩阵不依赖于世界坐标系的选择。

12.1.2 极点的张量形式

考虑第一幅图像的极点 $e = (e^1, e^2, e^3)^T$ ，它是第二个摄像机光心 C' 在第一幅图像上的投影，即 $PC' = se$ ，而 C' 是 $P' C' = 0$ ，因此下述方程：

$$\begin{pmatrix} p^1 & e^1 & 0 \\ p^2 & e^2 & 0 \\ p^3 & e^3 & 0 \\ p'^1 & 0 & 0 \\ p'^2 & 0 & 0 \\ p'^3 & 0 & 0 \end{pmatrix} \begin{pmatrix} C' \\ -s \\ s' \end{pmatrix} = 0 \tag{12.1.6}$$

必有形如 $\begin{pmatrix} C' \\ -s \\ s' \end{pmatrix}$ 的非零解，且 s' 为任意实数。因此，方程(12.1.6)系数矩阵的秩必为 4。所以它的所

有 5 阶行列式均为零。于是，得到：

$$\det \begin{pmatrix} p^1 & e^1 \\ p^2 & e^2 \\ P' & 0 \end{pmatrix} = 0, \quad \det \begin{pmatrix} p^1 & e^1 \\ p^3 & e^3 \\ P' & 0 \end{pmatrix} = 0, \quad \det \begin{pmatrix} p^2 & e^2 \\ p^3 & e^3 \\ P' & 0 \end{pmatrix} = 0,$$

由此，我们得到极点的张量形式：

$$e^i = \det \begin{pmatrix} p^i \\ P' \end{pmatrix} \quad (12.1.7)$$

同理，可以得到第二幅图像极点的张量形式：

$$e'^i = \det \begin{pmatrix} P \\ p'^i \end{pmatrix} \quad (12.1.8)$$

12.2 三线性关系

12.2.1 三点对应

本节考虑三幅图像 $\{I, I', I''\}$ 的三对应点的三线性关系。令

$$\mathbf{x} = (x^1, x^2, x^3)^T, \mathbf{x}' = (x'^1, x'^2, x'^3)^T, \mathbf{x}'' = (x''^1, x''^2, x''^3)^T$$

是三幅图像的一个三点对应，如图 12.2.1 所示。

假定三幅图像对应的摄像机矩阵分别为

$$P = \begin{pmatrix} p^1 \\ p^2 \\ p^3 \end{pmatrix}, P' = \begin{pmatrix} p'^1 \\ p'^2 \\ p'^3 \end{pmatrix}, P'' = \begin{pmatrix} p''^1 \\ p''^2 \\ p''^3 \end{pmatrix}$$

则有下列等式成立：

$$\begin{pmatrix} P & x & 0 & 0 \\ P' & 0 & x' & 0 \\ P'' & 0 & 0 & x'' \end{pmatrix} \begin{pmatrix} X \\ -s \\ -s' \\ -s'' \end{pmatrix} = 0 \tag{12.2.1}$$

其中 X 是点对应的空间点， s, s', s'' 是非零常数。

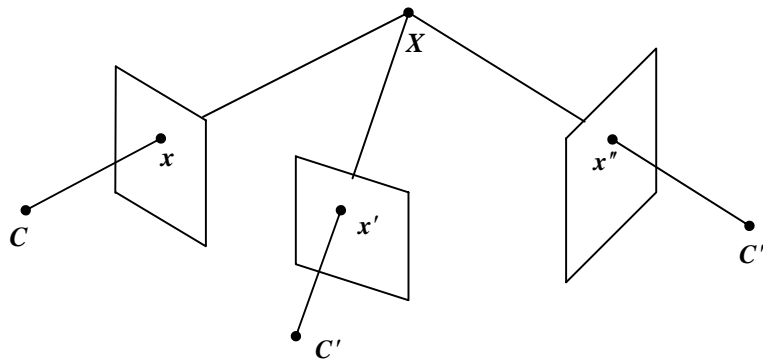


图 12.2.1：三点对应关系。

由于方程(12.2.1)有非零解，所以其系数矩阵的秩至多是 6，因此它的 7×7 子式均为零。 7×7 子矩阵因选择行的方式不同而有下列两种类型：

- I. 从两个摄像机矩阵中选取三行，而在另一个摄像机矩阵中选取一行；
- II. 从一个摄像机中选取三行，而在另两个摄像机中各选取两行。

先考虑类型 I，例如：选取第一与第二个摄像机矩阵中的三行，第三个摄像机矩阵中选取一行，得到下述 7×7 子矩阵：

$$\begin{pmatrix} P & x & 0 & 0 \\ P' & 0 & x' & 0 \\ p''^j & 0 & 0 & x''^j \end{pmatrix}$$

将它的行列式按最后一列展开，得到：

$$\det \begin{pmatrix} P & x & 0 & 0 \\ P' & 0 & x' & 0 \\ p''^j & 0 & 0 & x''^j \end{pmatrix} = x''^j \det \begin{pmatrix} P & x & 0 \\ P' & 0 & x' \end{pmatrix}$$

如果令这个行列式等于零，则导致第一与第二幅图像间的双线性关系。因此，第一种选择类型不能为三幅图像之间建立任何约束关系。这种类型不是我们所关心的。

下面考虑类型 II, 例如选取第一个摄像机矩阵的三行, 而另外两个摄像机矩阵各选取两行, 就得到下述 7×7 子矩阵:

$$X_{uv} = \begin{pmatrix} P & \mathbf{x} & 0 & 0 \\ \mathbf{p}'^j & 0 & x'^j & 0 \\ \mathbf{p}'^l & 0 & x'^l & 0 \\ \mathbf{p}''^k & 0 & 0 & x''^k \\ \mathbf{p}''^m & 0 & 0 & x''^m \end{pmatrix} \quad (12.2.2)$$

其中 u, v 是自由指标, 它们分别对应于从摄像机矩阵 P', P'' 中删去的行所对应的行指标。使用处理两幅图像间的双线性关系时的同样方法, 将这个矩阵的行列式按包含 \mathbf{x} 列(即矩阵的第 5 列)进行展开, 我们可以得到:

$$\det X_{uv} = -\frac{1}{2} x^i x'^j x''^k \varepsilon_{ilm} \varepsilon_{jqu} \varepsilon_{krv} \det \begin{pmatrix} \mathbf{p}'^l \\ \mathbf{p}''^m \\ \mathbf{p}'^q \\ \mathbf{p}''^r \end{pmatrix} \quad (12.2.3)$$

引进张量:

$$t_i^{qr} = \frac{1}{2} \varepsilon_{ilm} \det \begin{pmatrix} \mathbf{p}'^l \\ \mathbf{p}''^m \\ \mathbf{p}'^q \\ \mathbf{p}''^r \end{pmatrix} \quad (12.2.4)$$

这样, 我们就得到了三点对应的三线性关系:

$$x^i x'^j x''^k \varepsilon_{jqu} \varepsilon_{krv} t_i^{qr} = 0_{uv} \quad (12.2.5)$$

由于每一个自由指标 $u(v)$ 都有 3 种不同的选择, 所以一共有 9 个形如(12.2.5)的三线性关系。然而, 在这 9 个三线性关系中仅有 4 个关于 t_i^{qr} 的约束是独立的, 其证明在第 12.2.3 节。张量 t_i^{qr} 通常称为三焦张量。

三焦张量 t_i^{qr} 是通过选取第一个摄像机矩阵的三行和另两个摄像机矩阵的两行而得到的, 我们也可通过选取其它摄像机矩阵的三行和另两个摄像机矩阵的两行类似地建立三幅图像间的三线性关系和相应的三焦张量。由于一个摄像机矩阵的三行有三种选择方式, 因此三幅图像能产生三个不同形式的三焦张量, 而每一个三焦张量可导致 9 个(点对应)三线性关系。这样, 从三幅图像能导致 27 个(点对应)三线性关系, 它们分别归结于三个不同形式的三焦张量。尽管三焦张量有不同形式, 但它们的本质是相同的, 都揭示了三对应点的同一关联关系。因此, 实际应用中只须要考虑其中的一个

三个张量形式。

值得指出的是，从表达式(12.3.4)很容易说明三焦张量不依赖于摄像机矩阵的具体形式，等价地说不依赖于世界坐标系的选择。即三焦张量在空间射影变换下是不变的，它仅与三幅图像有关。

三焦张量 t_i^{pr} 由三个 3×3 矩阵构成，一共有 27 个元素，所以除去所有矩阵的一个全局尺度因子后有 26 个独立的比值参数，然而它仅有 18 个独立的自由度。自由度的数目可用下述方法计算出来。每一个摄像机矩阵有 11 个自由度，三个摄像机矩阵共有 33 个自由度，考虑到三焦张量与世界射影坐标系的选择无关，应该减去三维射影变换的 15 个自由度，所以三焦张量有 18 个自由度。这样，一旦给定了三焦张量中的 18 个元素，从理论上来说在相差一个公共尺度因子意义下所有 27 个元素都可以确定。但在实际应用中，由于元素间的约束非常复杂，从已知的 18 个元素确定另外 9 个元素是非常困难的，同时在有噪声的情况下也是不可取的或不可能的。

12.2.2 点、线对应

点—线—线对应

下面考虑空间点 X^j 投影到第一幅图像的点 x^i ，同时投影到另两幅图像直线 l'_p 和 l''_r 上，(如图 12.2.2 所示，所形成的点—线—线对应的三线性关系。

设摄像机矩阵分别为 $P = p_j^i$, $P' = p_j'^i$, $P'' = p_j''^i$ ，则有

$$sx^i = p_j^i X^j, \quad l'_i p_j'^i X^j = 0, \quad l''_i p_j''^i X^j = 0 \quad (12.2.6)$$

不难看出， $l'_i p_j'^i$, $l''_i p_j''^i$ 分别是图像直线 l'_p 和 l''_r 的反投影平面。方程组(12.2.6)可以写成下述矩阵形式：

$$\begin{pmatrix} P & \mathbf{x} \\ l'_p \mathbf{p}'^p & 0 \\ l''_r \mathbf{p}''^r \end{pmatrix} \begin{pmatrix} X \\ -s \end{pmatrix} = 0 \quad (12.2.7)$$

由于上述方程有非零解，所以其系数矩阵的行列式等于零。而

$$\det \begin{pmatrix} P & \mathbf{x} \\ l'_p \mathbf{p}'^p & 0 \\ l''_r \mathbf{p}''^r \end{pmatrix} = -\frac{1}{2} x^i \varepsilon_{ilm} l'_p l''_r \det \begin{pmatrix} \mathbf{p}^l \\ \mathbf{p}'^p \\ \mathbf{p}''^r \end{pmatrix} = -x^i l'_p l''_r t_i^{pr}$$

所以，我们有下述点—线—线的三线性关系：

$$x^i l'_p l''_r t_i^{pr} = 0 \quad (12.2.8)$$

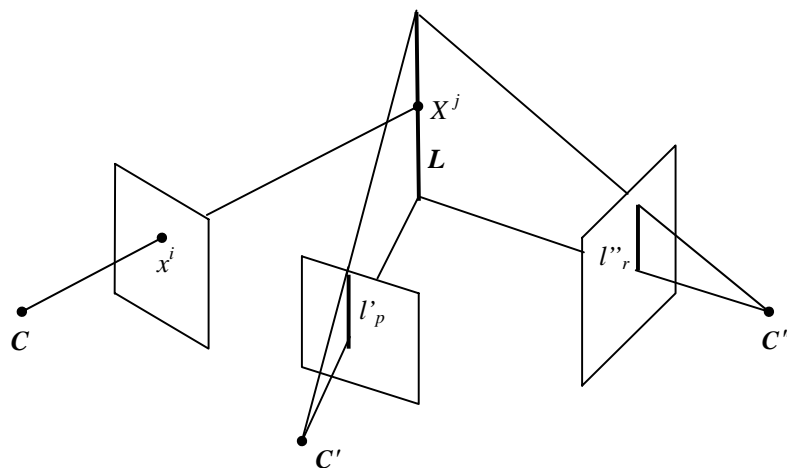


图 12.2.2: (点, 线, 线)的对应关系。

三线对应

下面考虑空间直线 L 投影到三幅图像的直线分别为 l_p , l'_q 和 l''_r , 如图 12.2.3 所示, 所形成的三线对应的三线性关系。

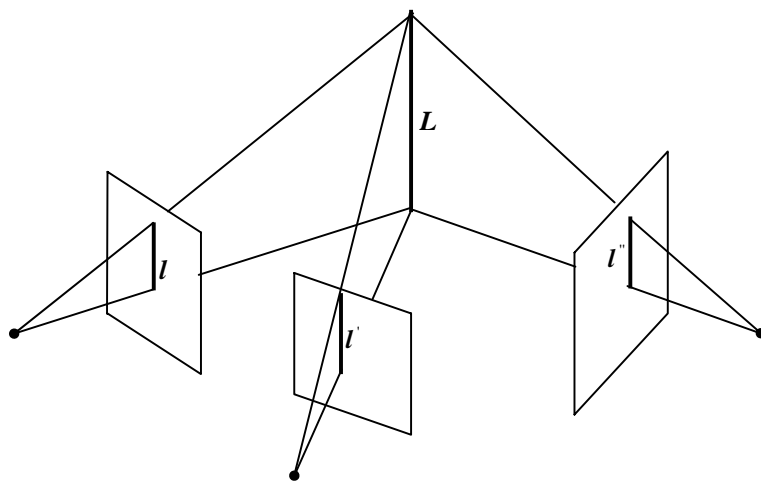


图 12.2.3: 三维空间直线 L 在三幅图像中的投影直线构成三线对应 $l \leftrightarrow l' \leftrightarrow l''$ 。

反之, 三条对应线的反向投影面必交于空间的一条直线 L 。

在直线 l_p 上一点 x^i , 则 x^i , l'_q 和 l''_r 构成(点, 线, 线)对应, 所以有

$$x^i l'_q l''_r t_i^{qr} = 0$$

于是，必有

$$l_p = l'_q l''_r t_i^{qr}$$

这个等式在相差一个常数倍的意义下相等，即等式两边向量的叉积是一个零向量，因此有下述三线性关系：

$$l_p l'_q l''_r \varepsilon^{ipw} t_i^{qr} = 0^w \tag{12.2.9}$$

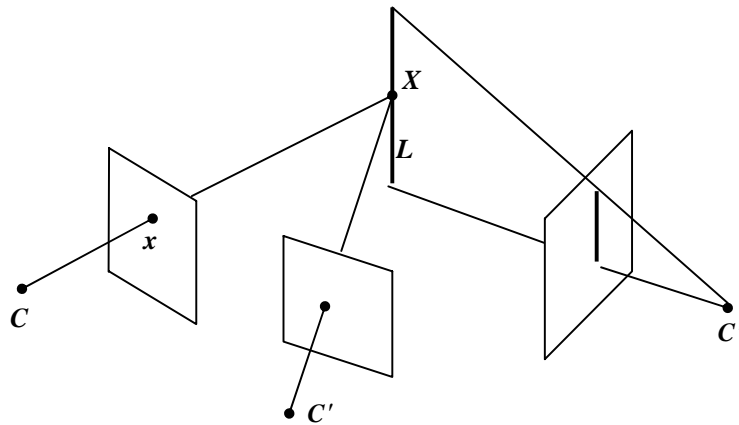


图 12.2.4：(点—点—线)的对应关系。

点—点—线对应

用上面相类似的方法，还可以建立两点一线对应的三线性关系。例如空间点映射到第一和第二幅图像上的点 x^i, x'^j ，同时映射到第三幅图像中的一条直线 l''_r 上，如图 12.2.4 所示，则对应 $x^i \leftrightarrow x'^j \leftrightarrow l''_r$ 具有下述三线性关系：

$$x^i x'^j l''_r \varepsilon_{jqk} t_i^{kr} = 0_u \tag{12.2.10}$$

12.2.3 三线性关系的独立数

本节主要考虑三线性关系的独立数，各种类型三线性关系的独立数概括在下表 12.3.1 中。下面以三点对应为例，对三线性关系作更为详细分析。

在三线性关系式 $x^i x'^j x''^k \varepsilon_{ip1} \varepsilon_{kq2} t_i^{pq} = 0$ 中，根据 r 和 s 的三种选择可以得到 9 个这样的三线性关系。从几何上看，这些三线性关系是由点—线—线对应关系在第二幅和第三幅图像上选择特殊直线

而得到。选择 $r=1,2,3$ 分别对应于图像上平行于 x 轴、 y 轴和通过图像原点的直线。例如，选择 $r=1$ ，并展开 $x'^j \varepsilon_{jpr}$ 得到一条水平直线：

$$l'_p = x'^j \varepsilon_{jpr} = (0, -x'^3, x'^2)$$

因为对任意 t ，形如 $\mathbf{y}' = (x'^1 + t, x'^2, x'^3)^T$ 的点都满足 $l'_p \mathbf{y}' = 0$ ，所以它是第二幅图像上过点 \mathbf{x}' 的一条水平直线。在第三幅图像上选择 $s=2$ ，类似地可以得到过 \mathbf{x}'' 的铅直线：

$$l''_q = x''^k \varepsilon_{kq2} = (x''^3, 0, x''^1)$$

从而，三点对应的三线性关系可展开为

$$x^i (-x'^3 (x''^3 t_i^{21} - x''^1 t_i^{23}) + x'^2 (x''^3 t_i^{31} - x''^1 t_i^{33})) = x^i x'^j x''^k \varepsilon_{ip1} \varepsilon_{kq2} t_i^{pq} = 0$$

在这 9 个三线性等式中，只有四个是独立的。这意味着所有 9 个三线性等式都是其中某 4 个的线性组合。关于这一点可以回溯到点—线—线关系 $x^i l'_j l''_k t_i^{jk} = 0$ ，其证明如下：在第三幅图像中过 \mathbf{x}'' 的直线形成一个单参数直线簇，如果 $\mathbf{m}'', \mathbf{n}''$ 是该直线簇中两条直线，则任一条过 \mathbf{x}'' 的其它直线都可以由它们的线性组合而得到：

$$l'' = \alpha \mathbf{m}'' + \beta \mathbf{n}''$$

由于(点，线，线)关系 $x^i l'_j l''_k t_i^{jk} = 0$ 关于 l'' 是线性的，所以又有

$$x^i l'_j m''_k t_i^{jk} = 0, \quad x^i l'_j n''_k t_i^{jk} = 0$$

于是，对其它任意直线 l'' 的关联关系可以由这两式的线性组合而得到。所以，对 l'' 来说仅有两个独立的关联关系。类似地，过 \mathbf{x}' 有一个单参数的直线簇，并且该关联关系关于过 \mathbf{x}' 的直线也是线性的。因此在第一幅图像上的点和第二、第三幅图像上的点之间总共有四个独立的关联关系。

表 12.2.1: 三线性等式与独立数

对应	三线性等式	独立数
线—线—线	$(l_r \varepsilon^{ris}) l'_j l''_k t_i^{jk} = 0^s$	2
点—线—线	$x^i l'_j l''_k t_i^{jk} = 0$	1
点—线—点	$x^i l'_j (x''^k \varepsilon_{kqs}) t_i^{jq} = 0_s$	2
点—点—线	$x^i (x'^j \varepsilon_{jpr}) l''_k t_i^{pk} = 0_r$	2
点—点—点	$x^i (x'^j \varepsilon_{jpr}) (x''^k \varepsilon_{kqs}) t_i^{pq} = 0_{rs}$	4

12.2.4 恢复摄像机矩阵

前几节从三个摄像机矩阵引入三幅图像间三焦张量的概念，三焦张量不依赖于空间坐标系的选择，即它在空间射影变换下是不变的。并且，我们还给出了三焦张量的三线性约束关系，因此从点对应、线对应以及点线对应可以计算三焦张量。本节将讨论如何从三焦张量在射影等价的意义下恢复出摄像机矩阵。

极线

如果直线 l' 反向投影平面 π' 是前两个摄像机的极平面(从而通过第一幅摄像机的光心 C)时，就会产生一种(点, 线, 线)对应的特殊情况，如图 12.2.5 所示。设 X 是平面 π' 的一个点，那么由 X 和 C 确定的射线在这个平面上，并且 l' 是对应于 X 的图像点 x 的极线。

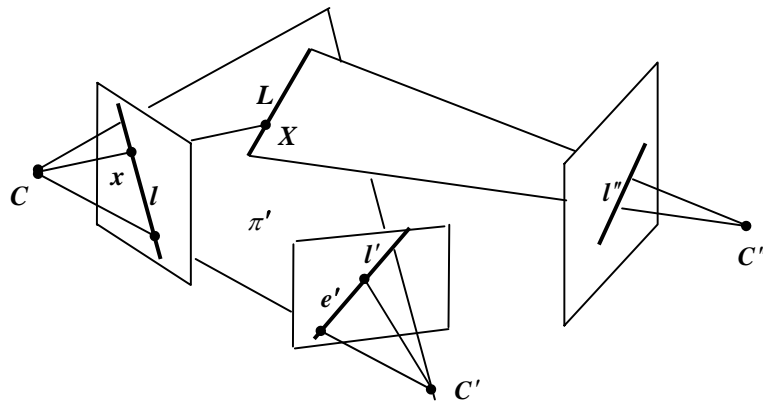


图 12.2.5: 若 l' 的反投影面 π' 是前两幅图像的一个极平面，在第三幅图像上的任意直线 l'' 都能提供一个(点—线—线)关联。

由第三幅图像直线 l'' 反向投影得到的平面 π'' 交平面 π' 于直线 L 。此外，由于 x 的反投影射线在平面 π' 上，故它必与 L 相交。这样点 x 的反投影射线和直线 l', l'' 两个反向投影平面交于一空间点 X ，从而它们是一个(点, 线, 线)对应，因此 $x^i l'_j l''_k t_i^{jk} = 0$ 。由于上式对任意的直线 l'' 都成立，从而有 $x^i l'_j t_i^{jk} = 0^k$ 。对换 l', l'' ，有类似的结论。于是，得到下述命题：

命题 12.2.1 如果 x 是第一幅图像上的一个点，它在第二幅和第三幅图像上的极线分别为 l', l'' ，则有

$$x^i l'_j t_i^{jk} = 0^k, \quad x^i l''_k t_i^{jk} = 0^j \tag{12.2.11}$$

即，对应于点 x 的极线 l', l'' 分别是矩阵 $(x^i t_i^{jk})$ 的左零向量和右零向量。

当点 \mathbf{x} 变化时，对应的极线也随之变化，而图像上的所有极线都通过极点。因此，可以通过计算不同的 \mathbf{x} 所对应极线的交点来获得极点。为了计算方便，点 \mathbf{x} 可选择为由齐次坐标表示的 3 个点： $(1, 0, 0)^T, (0, 1, 0)^T, (0, 0, 1)^T$ ，对应点 \mathbf{x} 的这三种选择所得到的矩阵 $(x^i t_i^{jk})$ 分别记为 T_1, T_2, T_3 。由此导出下述重要命题：

命题 12.2.2 第二幅图像上的极点 \mathbf{e}' 是由矩阵 T_i 的左零向量所表示的极线的公共交点。类似地，极点 \mathbf{e}'' 是由矩阵 T_i 的右零向量所表示的极线的公共交点。

基本矩阵

利用三焦张量，很容易计算出第一幅图和其它两幅图像之间的基本矩阵 F_{21}, F_{31} 。由命题 12.2.2，已经得到了第二幅图像和第三幅图像关于第一个摄像机的极点 \mathbf{e}' 和 \mathbf{e}'' 。如果找到第一幅图像与第二幅图像之间的一个单应 H_{21} ，就可以得到 F_{21} ，这是因为 $F_{21} = [\mathbf{e}']_{\times} H_{21}$ 。令 \mathbf{l}'' 是第三幅视图上的一条直线，显然， $H = (T_1 \mathbf{l}'', T_2 \mathbf{l}'', T_3 \mathbf{l}'')$ 是第一幅图像与第二幅图像之间的一个单应。为了避免退化情况发生， \mathbf{l}'' 不能是 T_i 的右零空间。为此，可取 $\mathbf{l}'' = \mathbf{e}''$ ，这是因为 \mathbf{e}'' 正交于每一个 T_i 的右零空间，从而不会发生退化情况。于是，第一幅到第二幅视图的基本矩阵由下述公式给出：

$$F_{21} = [\mathbf{e}']_{\times} [T_1, T_2, T_3] \mathbf{e}'' \quad (12.2.12)$$

类似地，也有

$$F_{31} = [\mathbf{e}'']_{\times} [T_1^T, T_2^T, T_3^T] \mathbf{e}' \quad (12.2.13)$$

恢复摄像机矩阵

正如从两幅图像的基本矩阵恢复摄像机矩阵对一样，从三焦张量恢复摄像机矩阵的三元组 $\{P, P', P''\}$ 时，也只能在射影等价的意义下得到恢复。这是容易理解的，因为三焦张量所揭示的是图像间的几何性质，与三维空间坐标无关，即三焦张量不依赖于射影坐标系的选择。因此，对三元组 $\{P, P', P''\}$ ，总可以将三个摄像机矩阵的形式为

$$P = (I, \mathbf{0}) \quad P' = (A, \mathbf{a}_4), \quad P'' = (B, \mathbf{b}_4)$$

其中 A 与 B 是 3×3 矩阵，向量 \mathbf{a}_i 和 \mathbf{b}_i 是对应摄像机矩阵的第 i 列， $i=1, \dots, 3$ 。 \mathbf{a}_4 和 \mathbf{b}_4 分别为第二和第三幅图像关于第一个摄像机的极点，即 $\mathbf{a}_4 = \mathbf{e}'$ ， $\mathbf{b}_4 = \mathbf{e}''$ 。此时，三焦张量可表示为

$$t_i^{jk} = a_i^j b_4^k - a_4^j b_i^k \quad (12.2.14)$$

由于 F_{21} 已知，见式(12.2.12)，根据第 3.4 节的讨论，第二个摄像机矩阵可取为

$$P' = ([T_1 \mathbf{e}'', T_2 \mathbf{e}'', T_3 \mathbf{e}''], \mathbf{e}') \quad (12.2.15)$$

这样，摄像机矩阵对 $\{P, P'\}$ 必有基本矩阵 F_{21} 。下面考虑第三个摄像机矩阵。

第三个摄像机的射影坐标系不能独立于前两个摄像机的射影坐标系，即三个摄像机必须有相同的射影坐标系。由于与基本矩阵 F_{31} 相容的摄像机矩阵 P'' 一般的形式是

$$P'' = ([T_1 e', T_2 e', T_3 e'] + e'' v''^T, \lambda'' e'')$$

为了得到一组与三焦张量相容的摄像机矩阵三元组，只需从这个一般形式中找到 $\{P, P', P''\}$ 具有三焦张量 t_i^{jk} 的 P'' 值。从 $P' = ([T_1 e'', T_2 e'', T_3 e''], e')$ 知 $a_i = T_i e''$ ，将它代入到(12.2.14)，并注意到 $a_4 = e'$ ，并选取 $b_4 = e''$ ，则有 $T_i = T_i e'' e''^T - e' b_i^T$ ，于是，

$$e' b_i^T = T_i (e'' e''^T - I)$$

选择适当的齐次因子使得 $\|e'\| = e'^T e' = 1$ ，并用 e'^T 左乘上式两边并转置有 $b_i = (e'' e''^T - I) T_i^T e'$ 。所以第三个摄像机有下述形式：

$$P'' = ((e'' e''^T - I) [T_1^T, T_2^T, T_3^T] e', e'') \quad (12.2.16)$$

12.3 四线性关系

12.3.1 四线性关系

本节主要考虑四幅图像间的四线性关系，关于三线性关系的推导同样适用于四线性关系。

四点对应

令 $x^i \leftrightarrow x'^j \leftrightarrow x''^k \leftrightarrow x'''^l$ 是四幅图像的一组点对应，四幅图像对应的摄像机矩阵分别为 $P = p_j^i, P' = p_j'^i, P'' = p_j''^i, P''' = p_j'''^i$ ，则有下述方程：

$$\begin{pmatrix} P & x & 0 & 0 & 0 \\ P' & 0 & x' & 0 & 0 \\ P'' & 0 & 0 & x'' & 0 \\ P''' & 0 & 0 & 0 & x''' \end{pmatrix} \begin{pmatrix} X \\ -s \\ -s' \\ -s'' \\ -s''' \end{pmatrix} = 0 \quad (12.3.1)$$

由于这个方程有非零解，所以其系数矩阵的秩至多等于 7，因此它的所有 8×8 子式均等于零。与推导三线性关系一样，如果 8×8 子式中包含某些摄像机矩阵的一行，则得到的是三线性关系或双线性关系，即此时不能为四幅图像提供任何新的约束关系。如果 8×8 子式中包含每个摄像机矩阵的两行，

用推导三线性关系的类似方法能获得下述四线性关系：

$$x^i x'^j x''^k x'''^l \varepsilon_{ipw} \varepsilon_{jqx} \varepsilon_{kry} \varepsilon_{lsz} t^{pqrs} = 0_{wxyz} \quad (12.3.2)$$

其中

$$t^{pqrst} = \det \begin{pmatrix} p^p \\ p'^q \\ p''^r \\ p'''^s \end{pmatrix} \quad (12.3.3)$$

并称它为四焦张量。

在表达式(12.3.2)中, w,x,y,z 是自由指标, 它们在 $\{1,2,3\}$ 范围内取值, 因此给定四幅图像能导致 81 个一组四对应点的四线性关系。然而, 在这些四线性关系中仅有 16 个是独立的。与三焦张量不同的是, 式(12.3.3)是唯一类型的四焦张量, 因为从选取摄像机矩阵中两行构成 8×8 子式只有一种选取方法, 即从所有的摄像机矩阵中都要选取两行才能构成 8×8 子式。

四线对应

与三线性关系类似, 也可以得到点、线对应的四线性关系。例如: 四对应线 $l_p \leftrightarrow l'_q \leftrightarrow l''_r \leftrightarrow l'''_s$ 具有下述四线性关系:

$$l_p l'_q l''_r l'''_s t^{pqrs} = 0 \quad (12.3.4)$$

与三线性不同的是, 这个四线性关系在图像直线不是同一空间直线的投影时, 也可能成立。例如: 当存在空间点在四幅图像的投影分别位于相应的图像直线时, 如图 12.3.1 所示, 也具有(12.3.4)的四线性关系, 论证在 13.4 节。由此, 可导致四线对应关于四焦张量 t^{pqrs} 更多的线性约束。

先考虑三直线对应的特殊情况。令 $l'_q \leftrightarrow l''_r \leftrightarrow l'''_s$ 是(第 2,3,4 幅图像)三对应直线, 对应的空间直线为 L , 如图 12.3.2 所示。任意给定第一幅图像上的一条直线 l_p , 则它的反投影平面必交空间直线 L 于一点 X , 因此关系式(12.3.4)成立。由于 l_p 是第一幅图像上的任意直线, 所以必有

$$l'_q l''_r l'''_s t^{pqrs} = 0^p \quad (12.3.5)$$

这就是三对应线关于 t^{pqrs} 的约束。自由指标的不同选择, 式(12.3.5)可导致 3 个线性约束。

现在回到四对应线的情况。由于在四对应线的情况下, 在每三幅图像中构成一个三对应线, 选取三幅图像有 4 种选取方式, 因此从四对应线可导致 12 个关于 t^{pqrs} 的线性约束方程。但是, 可以验证在这 12 个方程中只有 9 个是独立的。各种类型四线性关系的独立数概括在下表 12.3.1 中。

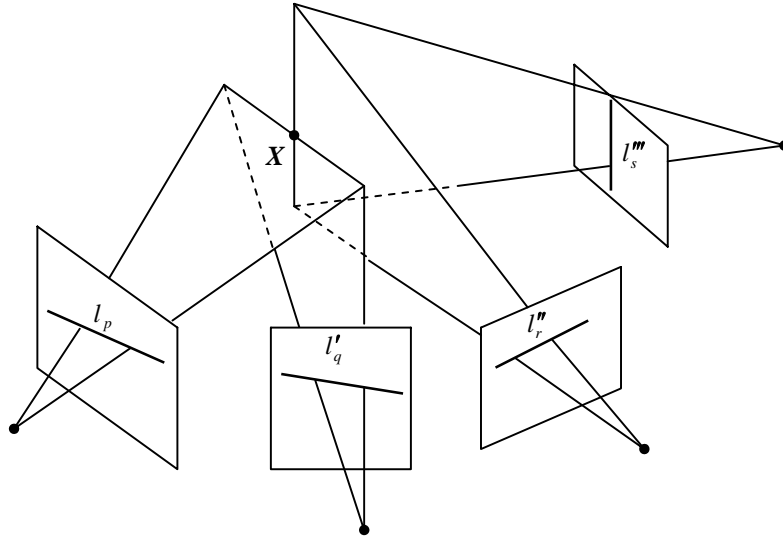


图 12.3.1: 当空间点 X 在四幅图像的投影分别位于相应的图像直线 l_p, l'_q, l''_r, l'''_s 时, 仍有四线性关系 $l_p l'_q l''_r l'''_s t^{pqrs} = 0$ 。

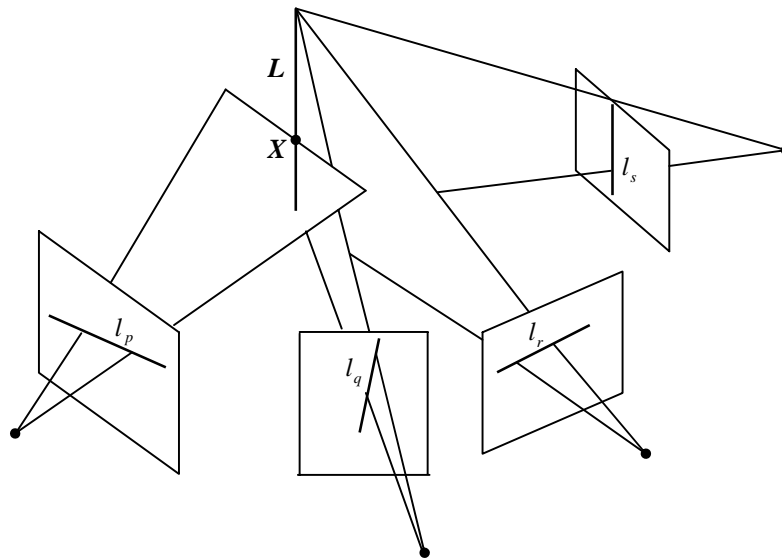


图 12.3.2: 三对应直线 $l'_q \leftrightarrow l''_r \leftrightarrow l'''_s$ 关于 t^{pqrs} 的约束: $l'_q l''_r l'''_s t^{pqrs} = 0^p$

表 12.3.1: 四线性等式与独立数

对应	关系	独立数
四点	$x^i x^j x^k x^l \varepsilon_{ipw} \varepsilon_{jqx} \varepsilon_{kry} \varepsilon_{lsz} t^{pqrst} = 0_{wxyz}$	16
三点一线	$x^i x^j x^k l_s^m \varepsilon_{ipw} \varepsilon_{jqx} \varepsilon_{kry} t^{pqrst} = 0_{wxy}$	8
两点两线	$x^i x^j l_r^m \varepsilon_{ipw} \varepsilon_{jqx} t^{pqrst} = 0_{wx}$	4
三线	$l_q^i l_r^j l_s^m t^{pqrs} = 0^p$	3
四线	$l_p^i l_r^j l_s^m t^{pqrs} = 0^s, l_q^i l_r^j l_s^m t^{pqrs} = 0^p, \dots$	9

12.3.2 四线性约束的独立数

本节主要考虑四对应点导致的 81 个约束方程的独立数。对此，有下述结论：

命题 12.3.1 给定四幅图像的一组点对应 $x^i \leftrightarrow x^j \leftrightarrow x^k \leftrightarrow x^l$ ，在它所引起的关于 t^{pqrst} 的约束方程：

$$x^i x^j x^k x^l \varepsilon_{ipw} \varepsilon_{jqx} \varepsilon_{kry} \varepsilon_{lsz} t^{pqrs} = 0, w, x, y, z = 1, 2, 3, 4 \tag{12.3.6}$$

中，仅有 16 个独立的约束。并且，如果将这个方程组写成 $At = 0$ 的形式，其中 A 是一个 81×81 矩阵， t 是由四焦张量所构成的 81 维向量，则 A 的 16 个非零奇异值均相等。

证明 方程 $At = 0$ 的系数矩阵 A 可表示为

$$A_{(wxyz)(pqrs)} = (x^i \varepsilon_{ipw})(x^j \varepsilon_{jqx})(x^k \varepsilon_{kry})(x^l \varepsilon_{lsz})$$

其中 $(wxyz)$ 表示矩阵行的组合指标，即 w, x, y, z 分别取 1, 2, 3, 4 得到矩阵的 81 个行；类似地， $(pqrs)$ 表示矩阵列的组合标。

不难验证： $(x^i \varepsilon_{ipw}), (x^j \varepsilon_{jqx}), (x^k \varepsilon_{kry}), (x^l \varepsilon_{lsz})$ 都是 3×3 反对称矩阵，而 3×3 反称矩阵必有两个相等的非零奇异值和一个零奇异值。所以，有

$$U_a^w (x^i \varepsilon_{ipw}) V_e^p = \alpha D_{ae}, U_b^x (x^j \varepsilon_{jqx}) V_f^q = \beta D_{bf}, U_c^y (x^k \varepsilon_{kry}) V_g^r = \gamma D_{cg}, U_d^z (x^l \varepsilon_{lsz}) V_h^s = \eta D_{dh}$$

其中 U_σ^r, V_σ^r 是正交矩阵， $D_{\sigma\tau}$ 是两个角元等于 1 而另一个对角元等于零的对角矩阵。所以，

$$(U_a^w U_b^x U_c^y U_d^z) A_{(wxyz)(pqrs)} (V_e^p V_f^q V_g^r V_h^s) = (\alpha\beta\gamma\eta)(D_{ae} D_{bf} D_{cg} D_{dh}) = \delta D_{(abcd)(efgh)}$$

可以验证： $U_{(abcd)}^{(wxyz)} = (U_a^w U_b^x U_c^y U_d^z), V_{(efgh)}^{(pqrs)} = (V_e^p V_f^q V_g^r V_h^s)$ 是两个 81×81 正交矩阵， $D_{(abcd)(efgh)}$ 是有 16 个都等于 1 的非零对角元素的 81×81 对角矩阵。证毕。

第三篇 模型估计

本篇提要

模型估计是众多应用学科的基本问题，计算机视觉中的模型估计通常是指变换或某种数学量的估计，例如：直线、二次曲线估计，摄像机矩阵估计，基本矩阵估计、单应矩阵估计和三焦张量估计等。模型估计方法，可分为四大类：代数方法、几何方法、鲁棒方法与统计方法。代数方法是线性的，计算简单，尽管估计结果可能较差，但它是其它方法的起点。几何方法通常都是高度非线性的，数值计算具有挑战性，但能给出较好的估计结果。代数和几何两类方法都不能从包含错误的测量数据集中有效地估计，鲁棒方法能克服这一缺点。统计方法主要有最大似然方法和贝叶斯两类方法，它们需要更精细的计算技术。

本篇具体内容以下六章构成：

第 13 章，介绍一些常用的非线性迭代优化技术，主要包括无约束和约束最小化迭代技术，利用它们才能实现模型估计的几何方法、鲁棒方法和统计方法。

第 14 章，介绍参数估计方法，主要包括最大似然方法和贝叶斯方法，同时还介绍实现这两种方法的期望最大化 (EM) 算法。

第 15 章，介绍视觉估计问题的代数方法，包括直接线性方法、因子化线性方法和归一化线性方法。代数方法是其它方法的起点。

第 16 章，介绍视觉估计问题的几何方法，基于高斯分布的最大似然方法等价于 MAH 距离最小化几何方法，因此最大似然方法也纳入本章的几何方法框架。

第 17 章，介绍视觉估计问题的鲁棒方法，主要包括应用最广泛的 RANSAC 方法、M-估计方法、最小中值方法和鲁棒最大后验方法。

第 18 章，介绍模型选择方法，主要包括模型选择 AIC、BIC 和 GRIC 信息标准，以及它们在视觉问题中的应用。

13. 迭代优化

在模型估计问题中，特别是在模型估计的几何方法、鲁棒方法和统计方法中，常常涉及到求解非线性最小化问题。非线性最小化问题通常使用迭代技术求解。本章将介绍一些常用的数值迭代方法，主要内容包括：求解无约束优化问题的变度量法、共轭梯度法和 L-M 法；对于约束优化问题，主要介绍惩罚法与乘子法。

13.1 最优性条件

13.1.1 最优性条件

非线性最优化问题可表达为

$$\begin{cases} \min f(\mathbf{x}) \\ \mathbf{x} \in \Omega \end{cases} \quad (13.1.1)$$

其中 $f(\mathbf{x})$ 是 $\mathbf{x} \in R^n$ 的非线性函数，称为代价函数(或目标函数)； Ω 是变量 \mathbf{x} 的约束区域，称为(13.1.1)的可行区域。通常，可行区域用一组(线性或非线性)函数的不等式或等式来描述，即将(13.1.1)表示为

$$\begin{cases} \min f(\mathbf{x}) \\ \text{subject to } g_i(\mathbf{x}) \leq 0, i = 1, 2, \dots, p; \\ h_j(\mathbf{x}) = 0, i = 1, 2, \dots, q \end{cases} \quad (13.1.2)$$

如果 $\Omega = R^n$ ，式(13.1.1)简记为

$$\min f(\mathbf{x}) \quad (13.1.3)$$

通常，称(13.1.2)为约束非线性优化问题，(13.1.3)为无约束非线性优化问题。

优化问题(13.1.2)或(13.1.3)的最优解通常是指它的局部最优解，所满足的必要条件和充分条件是算法设计的基础。我们先介绍无约束优化问题的最优性条件。

无约束问题

定理 13.1.1 设 $f(x)$ 在点 $x^0 \in R^n$ 可微，若存在 $p \in R^n$ 使得 $(\nabla f(x^0))^T p < 0$ ，则 p 是 $f(x)$ 在点 x^0 的下降方向。

所谓下降方向 p 是指，存在 $\delta > 0$ ，使得 $\forall t \in (0, \delta)$ 均有 $f(x^0 + tp) < f(x^0)$ 。

证明：由于 $f(x)$ 在点 x^0 可微，所以它在点 x^0 有 Taylor 展开式：

$$f(x^0 + tp) = f(x^0) + t(\nabla f(x^0))^T p + o(\|tp\|)$$

由于 $(\nabla f(x^0))^T p < 0$ ，取 $t > 0$ ，有 $t(\nabla f(x^0))^T p < 0$ ，故存在 $\delta > 0$ ，使得 $\forall t \in (0, \delta)$ 有

$$f(x^0 + tp) - f(x^0) = t(\nabla f(x^0))^T p + o(\|tp\|) < 0$$

因此，

$$f(x^0 + tp) < f(x^0)$$

证毕。

从定理 13.1.1，我们有下述推论：

推论 13.1.1 设 $f(x)$ 在点 $x^* \in R^n$ 可微，若 x^* 是(13.1.3)的局部最优解，则必有 $\nabla f(x^*) = 0$ 。

如果代价函数 $f(x)$ 是 R^n 上的可微凸函数，则 $\nabla f(x^*) = 0$ 也是 x^* 为(13.1.3)的整体最优解充分条件，即我们有

定理 13.1.2 如果函数 $f(x)$ 是 R^n 上的可微凸函数，若 $\nabla f(x^*) = 0$ ，则 x^* 是(13.1.3)的整体最优解。

证明：由于 $f(x)$ 是可微凸函数，所以 $\forall x \in R^n$ 有

$$(\nabla f(x^*))(x - x^*) \leq f(x) - f(x^*)$$

于是，根据 $\nabla f(x^*) = 0$ ，我们有 $f(x) \geq f(x^*)$ ，即 x^* 是(13.1.3)的整体最优解。

在一般情况下，由微积分学可知，局部最优解必是函数的驻点(或者说平稳点)，但反之不一定成立，因为驻点可能是极大点，也可能是鞍点。下面的定理给出了局部最优解的充分条件，它是多元微分理论的一个应用。

定理 13.1.3 设 $f(x)$ 二阶连续偏导数，在点 $x^* \in R^n$ 有 $\nabla f(x^*) = 0$ 且 $\nabla^2 f(x^*)$ 半正定，则 x^* 是 $f(x)$ 的局部最优解。若在上述条件中， $\nabla^2 f(x^*)$ 是正定的，则 x^* 是 $f(x)$ 的严格局部最优解。

约束问题

对于约束问题(13.1.2)的最优性条件比无约束问题要复杂得多。

将问题(13.1.2)的约束函数写成向量函数形式：

$$\mathbf{g}(x) = (g_1(x), g_2(x), \dots, g_p(x))^T$$

$$\mathbf{h}(\mathbf{x}) = (h_1(\mathbf{x}), h_2(\mathbf{x}), \dots, h_q(\mathbf{x}))^T$$

则(13.1.2)可简记为

$$\begin{cases} \min f(\mathbf{x}) \\ \text{subject to } \mathbf{g}(\mathbf{x}) \leq 0; \\ \mathbf{h}(\mathbf{x}) = 0 \end{cases} \quad (13.1.4)$$

先考虑等式约束问题：

$$\begin{cases} \min f(\mathbf{x}) \\ \text{subject to } \mathbf{h}(\mathbf{x}) = 0 \end{cases} \quad (13.1.5)$$

引进 Lagrange 函数

$$L(\mathbf{x}, \boldsymbol{\mu}) = f(\mathbf{x}) + \boldsymbol{\mu}^T \mathbf{h}(\mathbf{x}), (\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_q)^T), \quad (13.1.6)$$

令 \mathbf{x}^* 是(13.1.5)的局部最优解，且 $\text{rank}(\nabla \mathbf{h}(\mathbf{x}^*)) = q$ ，其中

$$\nabla \mathbf{h}(\mathbf{x}^*) = (\nabla h_1(\mathbf{x}^*), \nabla h_2(\mathbf{x}^*), \dots, \nabla h_q(\mathbf{x}^*))$$

则根据微积分学的条件极值定理，存在 $\boldsymbol{\mu}^*$ 使得

$$\nabla_{\mathbf{x}} L(\mathbf{x}^*, \boldsymbol{\mu}^*) = \nabla f(\mathbf{x}^*) + \nabla \mathbf{h}(\mathbf{x}^*) \cdot \boldsymbol{\mu}^* = 0 \quad (13.1.7)$$

即

$$-\nabla f(\mathbf{x}^*) = \sum_{j=1}^q \mu_j^* \nabla h_j(\mathbf{x}^*) \quad (13.1.8)$$

这表明：如果 \mathbf{x}^* 是(13.1.5)的局部最优解，则 $f(\mathbf{x})$ 在点 \mathbf{x}^* 的负梯度方向 $-\nabla f(\mathbf{x}^*)$ 落入到向量组 $\{\nabla h_1(\mathbf{x}^*), \nabla h_2(\mathbf{x}^*), \dots, \nabla h_q(\mathbf{x}^*)\}$ 所张成的线性子空间中。

再考虑不等式约束问题：

$$\begin{cases} \min f(\mathbf{x}) \\ \text{subject to } \mathbf{g}(\mathbf{x}) \leq 0 \end{cases} \quad (13.1.9)$$

如果点 \mathbf{x}^0 是(13.1.9)是可行解，即满足 $\mathbf{g}(\mathbf{x}^0) \leq 0$ ，在 $\mathbf{g}(\mathbf{x}^0)$ 的分量中可能有两种情况，第一种情况是使得等式 $g_i(\mathbf{x}^0) = 0$ 成立的分量，另一种是严格不等式 $g_j(\mathbf{x}^0) < 0$ 成立的分量。这种两种情况对 \mathbf{x}^0 的约束具有不同的效果：第一种情况的点 \mathbf{x}^0 是可行解集边界点， \mathbf{x}^0 非常微小的变化可能导致第一种情况不满足；对于第二种情况则不同，一般地，当 \mathbf{x}^0 有微小变化时，第二种情况仍然成立。通常，点 \mathbf{x}^0 使得第一种情况发生的分量称为点 \mathbf{x}^0 的积极约束，并记相应的分量指标集为：

$$I(\mathbf{x}^0) = \{i \mid g_i(\mathbf{x}^0) = 0\}$$

引进(13.1.9)的 Lagrange 函数:

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{x}), (\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_p)^T), \quad (13.1.10)$$

假定 \mathbf{x}^* 是(13.1.9)的局部最优解, 如果 $\{\nabla g_j(\mathbf{x}^*) \mid j \in I(\mathbf{x}^*)\}$ 线性无关, 则可以证明存在 $\boldsymbol{\lambda}^*$ 使得

$$\begin{cases} \nabla_x L(\mathbf{x}^*, \boldsymbol{\lambda}^*) = \nabla f(\mathbf{x}^*) + \sum_{j \in I(\mathbf{x}^*)} \lambda_j^* \nabla g_j(\mathbf{x}^*) = 0 \\ \boldsymbol{\lambda}^* \geq 0 \end{cases} \quad (13.1.11)$$

上式与(13.1.7)有相同的形式, 不同的是在这里乘子被要求是非负的, 因此它们的几何意义不相同。 $f(\mathbf{x})$ 在 \mathbf{x}^* 的负梯度方向 $-\nabla f(\mathbf{x}^*)$ 落入到向量组 $\{\nabla g_j(\mathbf{x}^*) \mid j \in I(\mathbf{x}^*)\}$ 所张成的锥中。

式(13.1.11)可以写成下述等价形式

$$\begin{cases} \nabla_x L(\mathbf{x}^*, \boldsymbol{\lambda}^*) = \nabla f(\mathbf{x}^*) + \nabla \mathbf{g}(\mathbf{x}^*) \cdot \boldsymbol{\lambda}^* = 0 \\ \text{diag}(\boldsymbol{\lambda}^*) \mathbf{g}(\mathbf{x}^*) = 0 \\ \boldsymbol{\lambda}^* \geq 0 \end{cases} \quad (13.1.12)$$

条件 $\text{diag}(\boldsymbol{\lambda}^*) \mathbf{g}(\mathbf{x}^*) = 0$, 通常称为互补松紧条件, 目的是使得式(13.1.12)与式(13.1.11)一致。

对一般约束问题(13.1.4), 它的 Lagrange 函数为

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{x}) + \boldsymbol{\mu}^T \mathbf{h}(\mathbf{x}) \quad (13.1.13)$$

我们有下述定理:

定理 13.1.4 设 $f(\mathbf{x}), \mathbf{g}(\mathbf{x}), \mathbf{h}(\mathbf{x})$ 在点 $\mathbf{x}^* \in R^n$ 连续可微, 且向量组

$$\{\nabla g_j(\mathbf{x}^*), \nabla h_i(\mathbf{x}^*) \mid j \in I(\mathbf{x}^*), i = 1, 2, \dots, q\}$$

线性无关。若 \mathbf{x}^* 是(13.1.4)的局部最优解, 则存在 $\boldsymbol{\mu}^*, \boldsymbol{\lambda}^*$ 使得

$$\begin{cases} \nabla_x L(\mathbf{x}^*, \boldsymbol{\lambda}^*) = \nabla f(\mathbf{x}^*) + \nabla \mathbf{g}(\mathbf{x}^*) \cdot \boldsymbol{\lambda}^* + \nabla \mathbf{h}(\mathbf{x}^*) \cdot \boldsymbol{\mu}^* = 0 \\ \text{diag}(\boldsymbol{\lambda}^*) \mathbf{g}(\mathbf{x}^*) = 0 \\ \boldsymbol{\lambda}^* \geq 0 \end{cases} \quad (13.1.14)$$

证明略。式(13.1.14)被称为 K-T 条件, \mathbf{x}^* 被称为 K-T 点, 它是一个非常著名的约束优化必要条件, 由 Kuhn 和 Tucker 在 1951 年所给出。在大多数优化算法中, 都是求解 K-T 点。

在一定的凸性条件下, K-T 条件也是充分的, 例如, 我们有

定理 13.1.4 设 $f(\mathbf{x}), \mathbf{g}(\mathbf{x}), \mathbf{h}(\mathbf{x})$ 在点 $\mathbf{x}^* \in R^n$ 连续可微, 且 $f(\mathbf{x}), \mathbf{g}(\mathbf{x})$ 是凸函数, $\mathbf{h}(\mathbf{x})$ 是线性函数。若 \mathbf{x}^* 满足 K-T 条件, 则它是(13.1.4)的整体最优解。

13.1.2 迭代格式

在上节,介绍了最优解的必要条件。通常的优化算法都是寻求满足必要条件的点(驻点或 K-T 点),从理论上可以通过求解相应的方程组而获得驻点或 K-T 点,在实际中这通常是不可行的,因为所涉及的方程组十分复杂,具有高度非线性。因此,非线性优化问题大多都是使用迭代方法求解。

迭代方法的开始是按某种方法给出函数 $f(\mathbf{x})$ 极小点 \mathbf{x}^* 的初始估计 \mathbf{x}^0 , 并称它为初始点, 然后按照某种特定的迭代规则产生一个点列 $\{\mathbf{x}^k\}$, 使得 $\{\mathbf{x}^k\}$ 收敛到极小点 \mathbf{x}^* 。一般地, 假定第 k 次迭代点为 \mathbf{x}^k , 下面分析如何产生下一次迭代点 \mathbf{x}^{k+1} :

令 $\Delta\mathbf{x}^k = \mathbf{x}^{k+1} - \mathbf{x}^k$, 则 $\Delta\mathbf{x}^k$ 是一个以 \mathbf{x}^k 为起点, \mathbf{x}^{k+1} 为终点的向量, $\mathbf{x}^{k+1} = \mathbf{x}^k + \Delta\mathbf{x}^k$ 。设 $\mathbf{p}^k \in R^n$ 是向量 $\Delta\mathbf{x}^k$ 的方向, 则存在数 $t_k > 0$ 使 $\Delta\mathbf{x}^k = t_k \mathbf{p}^k$, 所以有

$$\mathbf{x}^{k+1} = \mathbf{x}^k + t_k \mathbf{p}^k$$

对于约束优化问题, 还必须使得 $\mathbf{x}^{k+1} \in \Omega$ 。 \mathbf{p}^k 称为第 $k+1$ 次的搜索方向, t_k 为第 $k+1$ 次的迭代步长。在迭代算法中, 通常还要求函数值是下降的, 即要求

$$f(\mathbf{x}^{k+1}) < f(\mathbf{x}^k)$$

因此, 确定一个好的搜索方向和沿此方向的迭代步长, 就构成迭代算法的最重要步骤。

迭代格式

- 步骤 1: 选取初始点 \mathbf{x}^0 , $k = 0$;
- 步骤 2: 构造搜索方向 \mathbf{p}^k ;
- 步骤 3: 根据 \mathbf{p}^k , 确定搜索步长 t_k ;
- 步骤 4: 令 $\mathbf{x}^{k+1} = \mathbf{x}^k + t_k \mathbf{p}^k$;
- 步骤 5: 检验 \mathbf{x}^{k+1} 是否满足某种终止条件。终止条件满足, 停止迭代, 输出结果 \mathbf{x}^{k+1} ;
否则, 令 $k = k + 1$, 转到步骤 2。

对格式中划下线的部分赋予具体内容, 就可以得到具体的迭代算法。

13.2 一维搜索

在多变量的优化迭代格式中, 假定第 k 次迭代点 \mathbf{x}^k 和第 $k+1$ 次迭代方向 \mathbf{p}^k , 如何确定第 $k+1$ 次迭代步长 t_k ? 这是多变量非线性优化中的子问题, 此问题归结为

$$\min_{t \geq 0 (0 \leq t \leq t_{\max})} f(\mathbf{x}^k + t\mathbf{p}^k) \quad (13.2.1)$$

其中： $t \geq 0$ 还是 $0 \leq t \leq t_{\max}$ 应视原多变量优化问题是无约束优化还是约束优化问题而确定。

记 $\varphi(t) = f(\mathbf{x}^k + t\mathbf{p}^k)$ ，单变量的非线性优化问题(13.2.1)可以写成一般形式

$$\min_{t \geq 0(0 \leq t \leq t_{\max})} \varphi(t) \quad (13.2.2)$$

这类问题通常称为一维搜索问题。

求解一维搜索问题的方法，按照不同求解原则，可分为两大类：精确一维搜索和非精确一维搜索。

13.2.1 精确搜索

Newton 方法

基本思想：用函数 $\varphi(t)$ 在探索点 t_k 处的二阶 Taylor 展开式

$$g(t) = \varphi(t_k) + \varphi'(t_k)(t - t_k) + \frac{\varphi''(t_k)}{2}(t - t_k)^2$$

近似 $\varphi(t)$ ，即

$$\varphi(t) \approx g(t) = \varphi(t_k) + \varphi'(t_k)(t - t_k) + \frac{\varphi''(t_k)}{2}(t - t_k)^2$$

然后，用 $g(t)$ 的极小点作为新的探索点 t_{k+1} 。由于 $g(t)$ 是一个二次函数，可以直接得到它的极小点：

$$t_{k+1} = t_k - \frac{\varphi'(t_k)}{\varphi''(t_k)} \quad (13.2.3)$$

给定一个初始点 t_0 ，应用式(13.2.3)进行迭代计算，当 $|\varphi'(t_k)| < \varepsilon$ 时，迭代结束，输出 t_k 作为 $\varphi(t)$ 最小点的一个近似。

Newton 迭代算法：

- 1) 给定初始点 t_1 和终止控制常数 $\varepsilon > 0$ ， $k = 1$ ；
- 2) 如果 $|\varphi'(t_k)| < \varepsilon$ ，停止迭代，输出 t_k ，否则，当 $\varphi''(t_k) = 0$ 时，停止，算法失效；当 $\varphi''(t_k) \neq 0$ 时，转下一步；
- 3) 计算 $t_{k+1} = t_k - \frac{\varphi'(t_k)}{\varphi''(t_k)}$ ，如果 $|\varphi'(t_k)| < \varepsilon$ （或 $|t_{k+1} - t_k| < \varepsilon$ ），停止迭代，输出结果；否则 $k = k + 1$ ，转步 2)。

注意，Newton 算法的前提是 $\varphi''(t_k) \neq 0$ 。如果 $\varphi(t)$ 不存在二阶导数，此算法失效，此时，可用下述方法进行一维搜索。

0.618 方法

这种方法是在 $\varphi(t)$ 的单谷区间上进行一维搜索。

如果存在 $t^* \in [a, b]$ 使得 $\varphi(t)$ 在 $[a, t^*]$ 上单调减，并且在 $[t^*, b]$ 上单调增，则称 $\varphi(t)$ 在 $[a, b]$ 是单谷的或者说 $[a, b]$ 是 $\varphi(t)$ 的单谷区间。显然 t^* 是 $\varphi(t)$ 在区间 $[a, b]$ 上的最小值。

函数 $\varphi(t)$ 在 $[a, b]$ 上的最小点是 t^* ，所以在 $[a, b]$ 上选取两点 $t_1 < t_2$ ，通过计算和比较 $\varphi(t_1), \varphi(t_2)$ ，可以缩小区间 $[a, b]$ 使得最小点 t^* 在缩小后的区间上。这是因为根据单谷区间的定义，我们有

$$\varphi(t_1) \leq \varphi(t_2) \Rightarrow t^* \in [a, t_2]$$

$$\varphi(t_1) > \varphi(t_2) \Rightarrow t^* \in [t_1, b]$$

如图 13.1.1 所示。

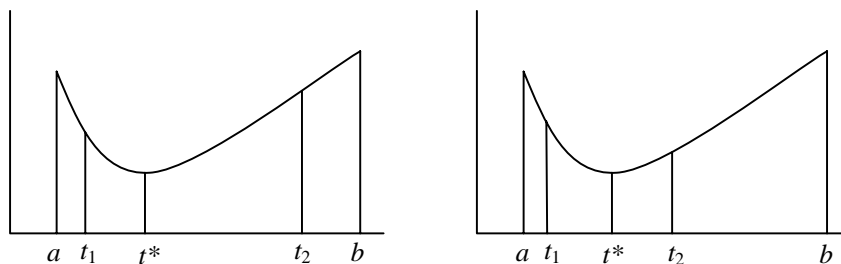


图 13.1.1：搜索单谷区间的最小点 t^*

这样，可以将对 t^* 的搜索由区间 $[a, b]$ 缩小为区间 $[a, t_2]$ (或 $[t_1, b]$)，称 t^* 所在的区间为搜索区间。继续这个过程，可使搜索区间长度达到任意小。这种方法在于搜索点的选择，不同的选择方式构成了不同的一维搜索方法。下面介绍 0.618 方法。

由于事先不知道缩小后的区间是 $[a, t_2]$ 还是 $[t_1, b]$ ，因此设置这两个区间等长，即

$$t_2 - a = b - t_1$$

令

$$\sigma = \frac{t_2 - a}{b - a} = \frac{b - t_1}{b - a} \quad (13.2.4)$$

它表示搜索区间的缩小比，则有

$$\begin{aligned} t_1 &= a + (1 - \sigma)(b - a) \\ t_2 &= a + \sigma(b - a) \end{aligned} \quad (13.2.5)$$

希望在每一搜索中，搜索区间的长度按相同的比值 σ 缩小。不妨假定新的搜索区间为 $[a, t_2]$ ，由于 $t_1 \in [a, t_2]$ 且函数值 $\varphi(t_1)$ 已知，所以希望它是下一次迭代中的探索点之一，这种选择的目的是为了

在以后的迭代中仅需要计算一个新探索点的函数值，减少计算量。假定在新一次迭代中的两个探索点为 $t'_1 < t'_2$ ，它们均在区间 $[a, t_2]$ 内，并且其中一个是 t_1 。于是，有

$$\sigma = \frac{t'_2 - a}{t_2 - a} = \frac{t_2 - t'_1}{t_2 - a}$$

结合式(13.2.5)，得到

$$t'_2 - a = t_2 - t'_1 = \sigma^2(b - a)$$

如果 $t'_1 = t_1$ ，则导致 $\sigma^2 - 2\sigma + 1 = 0$ ，此方程仅有的解是 $\sigma = 1$ ，这是不可能的。因此，必有 $t'_2 = t_1$ 。用类似的方法，可得到方程 $\sigma^2 + \sigma - 1 = 0$ 。因 $0 < \sigma < 1$ ，所以

$$\sigma = \frac{\sqrt{5}-1}{2} \approx 0.618$$

如果新的搜索区间为 $[t_1, b]$ ，关于 σ 可得到相同的结果，但此时 $t'_1 = t_2$ ， t'_2 为新的探索点。

由于 $\frac{\sqrt{5}-1}{2}$ 称为黄金分割数，所以 0.618 法通常也称为近似黄金分割法。下面是 0.618 法的具体迭代步骤：

- 1) 确定单谷区间 $[a, b]$ ，并给定控制精度 ε ；
- 2) 计算两个初始探索点

$$t_1 = a + 0.382(b - a) = b - 0.618(b - a), \quad t_2 = a + 0.618(b - a),$$

以及 $\varphi_1 = \varphi(t_1), \varphi_2 = \varphi(t_2)$ ；

- 3) 若 $\varphi_1 \leq \varphi_2$ ，转步骤 4)；否则转步骤 5)；
- 4) 若 $t_2 - a < \varepsilon$ ，停止迭代，输出 t_1 ；否则，置

$$b := t_2, t_2 := t_1, t_1 := b - 0.618(b - a), \varphi_2 := \varphi_1, \varphi_1 := \varphi(t_1),$$

转步骤 3)；

- 5) 若 $b - t_1 < \varepsilon$ ，停止迭代，输出 t_2 ；否则，置

$$a := t_1, t_1 := t_2, t_2 := a + 0.618(b - a), \varphi_1 := \varphi_2, \varphi_2 := \varphi(t_2),$$

转步骤 3)。

13.2.2 非精确搜索

在求解多元优化问题中，通常关心的是迭代点列 $\{\mathbf{x}^k\}$ 的收敛速度，如果在每次迭代中，对迭代步长进行精确一维搜索，计算量非常大，就整体而言不能加快点列 $\{\mathbf{x}^k\}$ 的收敛速度。因此，人们建立了一些求解问题(13.2.2)的非精确一维搜索方法，减少计算量。非精确一维搜索方法通常包括两个部分：给定搜索点的规则和求解符合规则的点。规则的作用是要求非精确一维搜索解 t_k 满足

$f(\mathbf{x}^{k+1}) < f(\mathbf{x}^k)$ ，但 t_k 不能过大以致 $\{\mathbf{x}^k\}$ 有很大的摆动，也不能过小以致在收敛到最优解之前而止步不前。另外，非精确一维搜索方法通常假定 $\varphi(t)$ 在 $t = 0$ 的邻域内是可微的并且有 $\varphi'(0) < 0$ 。下面介绍一种实用非精确一维搜索方法，称为 Goldstein 法。

Goldstein 法 该方法限定搜索点不过大也不过小的规则是：

$$\varphi(t_k) \leq \varphi(0) + s_1 t_k \varphi'(0) \tag{13.2.5}$$

$$\varphi(t_k) \geq \varphi(0) + s_2 t_k \varphi'(0) \tag{13.2.6}$$

其中 $0 < s_1 < s_2 < 1$ 是事先指定的两个数。

如图 13.1.2 所示，式(13.1.5)是限定搜索点 t_k 在直线 $y = \varphi(0) + s_1 t \varphi'(0)$ 的下方，式(13.1.6) 是限定搜索点 t_k 在直线 $y = \varphi(0) + s_2 t \varphi'(0)$ 的上方，同时满足这两式子的搜索点 t_k 构成 $[a, b] \cup [c, d]$ 。在 $[a, b] \cup [c, d]$ 中的点都被认为是可接受的搜索点。

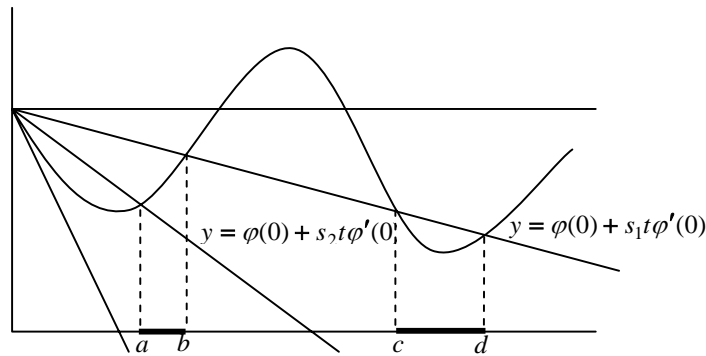


图 13.1.2: 非精确一维搜索(Goldstein 方法)

在 Goldstein 方法中，记 (a_k, b_k) 为当前搜索范围。如果当前探索点 t_k 不满足式(13.2.5)，则减小 t_k ；如果当前探索点 t_k 不满足式(13.2.6)，则增大 t_k ，直到两式同时满足为止。具体计算步骤如下：

- 1) 给定两个数 s_1, s_2 ： $0 < s_1 < s_2 < 1$ ，探索点的放大倍数 $\lambda > 1$ ，以及初始探索点 $t_0 \in (0, \infty) \setminus \{0, t_{\max}\}$ ；置： $a_0 := 0, b_0 := \infty(t_{\max}), k := 0$ ；
- 2) 计算 $\varphi(t_k)$ ，若 $\varphi(t_k) \leq \varphi(0) + s_1 t_k \varphi'(0)$ ，进行步骤 3)；否则，置 $a_{k+1} := a_k, b_{k+1} := t_k$ ，转步骤 4)；
- 3) 若 $\varphi(t_k) \geq \varphi(0) + s_2 t_k \varphi'(0)$ ，停止迭代，输出 t_k ；否则，置 $a_{k+1} := t_k, b_{k+1} := b_k$ ，若 $b_{k+1} < \infty(t_{\max})$ ，转步骤 4)；否则，置 $t_{k+1} := \lambda t_k, k+1 := k$ ，转步骤 2)；
- 4) 置 $t_{k+1} := \frac{a_k + b_k}{2}, k+1 := k$ ，转步骤 2)。

13.3 无约束优化

13.3.1 最速下降法

最速下降法假定目标函数 $f(\mathbf{x})$ 是一阶连续可微的，其基本思想是：从当前点 \mathbf{x}^k 出发，取 $f(\mathbf{x})$ 在点 \mathbf{x}^k 处下降最快的方向作为搜索方向。

由 $f(\mathbf{x})$ 在 \mathbf{x}^k 点处的一阶 Taylor 展开式

$$f(\mathbf{x}^k) - f(\mathbf{x}^k + t\mathbf{p}^k) = -t\nabla f(\mathbf{x}^k)^T \mathbf{p}^k + o(\|t\mathbf{p}^k\|)$$

可知，取 $\mathbf{p}^k = -\nabla f(\mathbf{x}^k)$ 时，函数下降最快（略去高阶无穷小不计）。于是，最速下降法的求解步骤为：

- 1) 给定初始点 \mathbf{x}^0 和终止控制常数 $\varepsilon > 0$ ，令 $k := 0$ ；
- 2) 计算 $\nabla f(\mathbf{x}^k)$ ，若 $\|\nabla f(\mathbf{x}^k)\| \leq \varepsilon$ ，停止迭代，输出 \mathbf{x}^k ，否则进行下一步；
- 3) 取 $\mathbf{p}^k = -\nabla f(\mathbf{x}^k)$ ，并进行一维搜索

$$f(\mathbf{x}^k + t_k \mathbf{p}^k) = \min_{t \geq 0} f(\mathbf{x}^k + t\mathbf{p}^k)$$

求步长 t_k ；令 $\mathbf{x}^{k+1} = \mathbf{x}^k + t_k \mathbf{p}^k$ ， $k = k + 1$ 转 2) 步。

13.3.2 Newton 法

Newton 法与一维搜索的 Newton 法类似，假定目标函数是二阶连续可微的，并假定二阶导数 $\nabla^2 f(\mathbf{x})$ 总是正定的。在每一个当前点 \mathbf{x}^k 处，以 $f(\mathbf{x})$ 的二阶近似函数

$$Q(\mathbf{x}) = f(\mathbf{x}^k) + \nabla f(\mathbf{x}^k)^T (\mathbf{x} - \mathbf{x}^k) + \frac{1}{2} (\mathbf{x} - \mathbf{x}^k)^T \nabla^2 f(\mathbf{x}^k) (\mathbf{x} - \mathbf{x}^k)$$

的极小点作为下一次的迭代点 \mathbf{x}^{k+1} 。这就是 Newton 法的基本思想。

由于 $\nabla^2 f(\mathbf{x}^k)$ 正定，所以 $Q(\mathbf{x})$ 是二次凸函数，因此 $Q(\mathbf{x})$ 的整体极小点 \mathbf{x}^{k+1} 可由下述方程

$$\nabla Q(\mathbf{x}) = \nabla f(\mathbf{x}^k) + \nabla^2 f(\mathbf{x}^k) (\mathbf{x} - \mathbf{x}^k) = 0$$

求得 $\mathbf{x}^{k+1} = \mathbf{x}^k - [\nabla^2 f(\mathbf{x}^k)]^{-1} \nabla f(\mathbf{x}^k)$ 。在 Newton 法中，取

$$\begin{cases} \mathbf{p}^k = -[\nabla^2 f(\mathbf{x}^k)]^{-1} \nabla f(\mathbf{x}^k) \\ t_k = 1 \end{cases} \quad (13.3.1)$$

Newton 法的计算步骤如下：

- 1) 选取初始点 \mathbf{x}^0 和终止控制常数 $\varepsilon > 0$ ，令 $k = 0$ ；

- 2) 计算 $\nabla f(\mathbf{x}^k)$, 若 $\|f(\mathbf{x}^k)\| \leq \varepsilon$, 停止迭代, 输出 \mathbf{x}^k , 否则进行下一步;
- 3) 取 $\mathbf{p}^k = -[\nabla^2 f(\mathbf{x}^k)]^{-1} \nabla f(\mathbf{x}^k)$, $\mathbf{x}^{k+1} = \mathbf{x}^k + \mathbf{p}^k, k = k+1$ 转第 2) 步。

对于二次凸函数, Newton 法只需要一次迭代就得到最优解。但是, 对于一般非线性优化问题, Newton 法通过有限次迭代并不能保证求得最优解, 除非目标函数在最优解附近能近似于一个二次函数并且初始值离最优解很近。如果初始点离最优解太远, Newton 法不能保证收敛, 甚至下降性也不能保证。为了克服 Newton 法的这些缺点, 人们已经给出一些 Newton 法的修正, 如 Murray 修正法等。

13.3.3 变度量法

变度量法是优化迭代问题的一类方法总称。在这类方法中, 仅使用目标函数 $f(\mathbf{x})$ 的梯度 $\nabla f(\mathbf{x})$, 不涉及二阶导数, 它们是最速下降法的改进, 其算法速度比最速下降法要快。而 Newton 法可以看作是变度量法的一个特例。在变度量方法中, 每一次迭代关于“最速下降”的度量是变化的, 因此称为变度量法。下面介绍变度量法的基本思想。

给定 n 阶对称正定矩阵 H , 在 R^n 空间中, 定义向量 \mathbf{x} 的椭球范数

$$\|\mathbf{x}\|_H = (\mathbf{x}^T H \mathbf{x})^{1/2}$$

正定矩阵 H 称为度量矩阵。在这个范数下, $f(\mathbf{x})$ 在点 $\mathbf{x}^k \in R^n$ 关于 \mathbf{p} 的方向导数应为

$$\lim_{t \rightarrow 0^+} \frac{f(\mathbf{x}^k + t\mathbf{p}) - f(\mathbf{x}^k)}{\|t\mathbf{p}\|_H} = \frac{(\nabla f(\mathbf{x}^k))^T \mathbf{p}}{\|\mathbf{p}\|_H}$$

因此, 当 $\nabla f(\mathbf{x}^k) \neq 0$ 时, 在范数 $\|\cdot\|_H$ 意义下, $f(\mathbf{x})$ 在点 $\mathbf{x}^k \in R^n$ 的最速下降方向可取为变量 $\mathbf{p} \in R^n$ 的非线性优化问题

$$\begin{cases} \min \nabla f(\mathbf{x}^k)^T \mathbf{p} \\ \text{subject to } \|\mathbf{p}\|_H^2 = 1 \end{cases} \quad (13.3.2)$$

的最优解 \mathbf{p}^* 。由微积分学知, 对应于(13.3.1)的 Lagrange 函数为

$$L(\mathbf{p}, \lambda) = \nabla f(\mathbf{x}^k)^T \mathbf{p} + \lambda(\mathbf{p}^T H \mathbf{p} - 1)$$

所以, 由 $\nabla L(\mathbf{p}^*, \lambda^*) = 0$ 和 $\nabla^2 L(\mathbf{p}^*, \lambda^*)$ 正定性, 可推知:

$$\lambda^* = \|\mathbf{H}^{-1} \nabla f(\mathbf{x}^k)^T\|_H, \mathbf{p}^* = -\frac{\mathbf{H}^{-1} \nabla f(\mathbf{x}^k)^T}{\|\mathbf{H}^{-1} \nabla f(\mathbf{x}^k)^T\|_H}$$

于是, 与 \mathbf{p}^* 方向相同的向量 $\mathbf{p}^k = -\mathbf{H}^{-1} \nabla f(\mathbf{x}^k)^T$ 是 $f(\mathbf{x})$ 在点 \mathbf{x}^k 的关于范数 $\|\cdot\|_H$ 的最速下降方向。

当 $H=I$ 时, $\mathbf{p}^k = -\nabla f(\mathbf{x}^k)^T$ 是最速下降法的迭代方向;

当 $H = \nabla^2 f(\mathbf{x}^k)$ 时, $\mathbf{p}^k = -(\nabla^2 f(\mathbf{x}^k))^{-1} \nabla f(\mathbf{x}^k)^T$ 是 Newton 法的迭代方向, 即 Newton 法的迭代方向是 Hesse 矩阵意义下的最速下降方向。

下面寻求一种迭代向量, 它能保证收敛速度快又可以避免 Hesse 矩阵的求逆。假定每次迭代的方向是

$$\mathbf{p}^k = -H_k \nabla f(\mathbf{x}^k)^T \quad (13.3.3)$$

其中正定矩阵 H_k 随迭代而发生变化, 我们要求它近似于 Hesse 矩阵的逆 $(\nabla^2 f(\mathbf{x}^k))^{-1}$, 并且只与 f 在相邻两次迭代点的梯度向量有关。为此, 设 f 二阶连续可微, 且 $\nabla^2 f(\mathbf{x})$ 可逆, 则由定理 13.1.1, 式(13.3.3)是 f 在 \mathbf{x}^k 点的下降方向。为了使 H_{k+1} 有类似于 $(\nabla^2 f(\mathbf{x}^k))^{-1}$ 的性质, 我们考虑 f 在 \mathbf{x}^{k+1} 点的 Taylor 二次逼近:

$$f(\mathbf{x}) \approx f(\mathbf{x}^{k+1}) + \nabla f(\mathbf{x}^{k+1})(\mathbf{x} - \mathbf{x}^{k+1}) + \frac{1}{2}(\mathbf{x} - \mathbf{x}^{k+1})^T \nabla^2 f(\mathbf{x}^{k+1})(\mathbf{x} - \mathbf{x}^{k+1})$$

于是, 有

$$\nabla f(\mathbf{x}) \approx \nabla f(\mathbf{x}^{k+1}) + \nabla^2 f(\mathbf{x}^{k+1})(\mathbf{x} - \mathbf{x}^{k+1})$$

令 $\mathbf{x} = \mathbf{x}^k$, 得到

$$(\nabla^2 f(\mathbf{x}^{k+1}))^{-1} \Delta \mathbf{g}^k \approx \Delta \mathbf{x}^k \quad (13.3.4)$$

其中

$$\Delta \mathbf{g}^k = \nabla f(\mathbf{x}^{k+1}) - \nabla f(\mathbf{x}^k), \Delta \mathbf{x}^k = \mathbf{x}^{k+1} - \mathbf{x}^k$$

如果 $f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} + c$, 则式(13.3.4)变为等式, 即

$$(\nabla^2 f(\mathbf{x}^{k+1}))^{-1} \Delta \mathbf{g}^k = \Delta \mathbf{x}^k \quad (\nabla^2 f(\mathbf{x}^{k+1}) = \mathbf{A}) \quad (13.3.5)$$

式(13.3.4)和(13.3.5)分别是非二次代价函数与二次代价函数的 Hesse 矩阵在点 \mathbf{x}^{k+1} 所要满足的条件, 所以为了使 H_{k+1} 有类似于 $(\nabla^2 f(\mathbf{x}^k))^{-1}$ 的性质, H_{k+1} 应该满足

$$H_{k+1} \Delta \mathbf{g}^k = \Delta \mathbf{x}^k \quad (13.3.6)$$

记 $H_{k+1} = H_k + \Delta H_k$, 则上式可改写成

$$\Delta H_k \Delta \mathbf{g}^k = \Delta \mathbf{x}^k - H_k \Delta \mathbf{g}^k \quad (13.3.7)$$

如果已知 $\mathbf{x}^k, \mathbf{x}^{k+1}, H_k$, 为了求解点 \mathbf{x}^{k+1} 的迭代方向 $\mathbf{p}^{k+1} = -H_{k+1} \nabla f(\mathbf{x}^{k+1})^T$, 只须从方程组(13.3.7)求解 ΔH_k 。由于方程组(13.2.7)是由 n^2 个变元 n 个方程所构成的方程组, 因此它的解不唯一。

求解问题(13.3.7)的不同方法就导致了不同的变度量方法, 其中最常用的是 DFP 方法和 BFGS

方法。DFP 方法是由 Davidon、Fletcher 和 Powell 所给出的，BFGS 方法是由 Broyden, Fletcher, Goldfarb 和 Shanno 的研究成果。它们都是从下述方程出发的：

$$\Delta H_k = \Delta \mathbf{x}^k \mathbf{u}_k^T - H_k \Delta \mathbf{g}^k \mathbf{v}_k^T \quad (13.3.8)$$

其中 $\mathbf{u}_k, \mathbf{v}_k$ 为待定向量。在 DFP 方法中，取

$$\mathbf{u}_k = \frac{\Delta \mathbf{x}^k}{(\Delta \mathbf{x}^k)^T \Delta \mathbf{g}^k}, \mathbf{v}_k = \frac{H_k \Delta \mathbf{g}^k}{\Delta \mathbf{g}^k{}^T H_k \Delta \mathbf{g}^k} \quad (13.3.9)$$

BFGS 方法中，取

$$\mathbf{u}_k = \frac{1}{(\Delta \mathbf{x}^k)^T \Delta \mathbf{g}^k} \left(1 + \frac{(\Delta \mathbf{g}^k)^T H_k \Delta \mathbf{g}^k}{(\Delta \mathbf{x}^k)^T \Delta \mathbf{g}^k} \right) \Delta \mathbf{x}^k, \mathbf{v}_k = \frac{\Delta \mathbf{x}^k}{(\Delta \mathbf{x}^k)^T \Delta \mathbf{g}^k} \quad (13.3.10)$$

DFP 方法

- 1) 选取初始点 \mathbf{x}^0 ，初始矩阵 $H_0 = I_n$ ，给定终止控制常数 $\varepsilon > 0$ ；
- 2) 计算 $\nabla f(\mathbf{x}^0)$ ，若 $\|\nabla f(\mathbf{x}^0)\| \leq \varepsilon$ ，停止迭代，输出 \mathbf{x}^0 ，否则进行下一步；
- 3) 令 $\mathbf{p}^0 = -H_0 \nabla f(\mathbf{x}^0)$ ， $k := 0$ ；
- 4) 一维搜索，求 t_k ： $f(\mathbf{x}^k + t_k \mathbf{p}^k) = \min_{t \geq 0} f(\mathbf{x}^k + t \mathbf{p}^k)$ ；令 $\mathbf{x}^{k+1} = \mathbf{x}^k + t_k \mathbf{p}^k$ ；
- 5) 计算 $\nabla f(\mathbf{x}^{k+1})$ ，若 $\|\nabla f(\mathbf{x}^{k+1})\| \leq \varepsilon$ ，停止迭代，输出 \mathbf{x}^{k+1} ，否则，进行下一步；
- 6) 检验迭代次数，若 $k+1 = n$ ，令 $\mathbf{x}^0 := \mathbf{x}^{k+1}$ ，转步骤 3)；否则，进行下一步；
- 7) 构造 DFP 方向：用 DFP 公式

$$H_{k+1} = H_k + \frac{\Delta \mathbf{x}^k (\Delta \mathbf{x}^k)^T}{(\Delta \mathbf{x}^k)^T \Delta \mathbf{g}^k} - \frac{H_k \Delta \mathbf{g}^k (\Delta \mathbf{g}^k)^T H_k}{(\Delta \mathbf{g}^k)^T H_k \Delta \mathbf{g}^k} \quad (13.3.11)$$

$$\Delta \mathbf{x}^k = \mathbf{x}^{k+1} - \mathbf{x}^k, \Delta \mathbf{g}^k = \nabla f(\mathbf{x}^{k+1}) - \nabla f(\mathbf{x}^k)$$

计算 H_{k+1} ，令 $\mathbf{p}^{k+1} = -H_{k+1} \nabla f(\mathbf{x}^{k+1})$ ， $k := k+1$ ，转步骤 4)。

BFGS 方法

BFGS 方法与 DFP 方法类似，所不同的是用下述 BFGS 公式：

$$H_{k+1} = H_k + \frac{\Delta \mathbf{x}^k (\Delta \mathbf{x}^k)^T}{(\Delta \mathbf{x}^k)^T \Delta \mathbf{g}^k} \left(1 + \frac{(\Delta \mathbf{g}^k)^T H_k \Delta \mathbf{g}^k}{(\Delta \mathbf{x}^k)^T \Delta \mathbf{g}^k} \right) - \frac{1}{(\Delta \mathbf{x}^k)^T \Delta \mathbf{g}^k} \left(\Delta \mathbf{x}^k (\Delta \mathbf{g}^k)^T H_k + H_k \Delta \mathbf{g}^k (\Delta \mathbf{x}^k)^T \right) \quad (13.3.12)$$

代替 DFP 公式(13.2.11)。

13.3.4 共轭方向法

共轭方向法与变度量法一样，也是一类方法的总称，也仅仅使用目标函数 $f(\mathbf{x})$ 的梯度。当初，共轭方向法是为求解二次函数的优化问题而设计的，其特点是所采用的搜索方向与二次函数系数矩阵相关的共轭方向，后来将这种方法应用于求解一般非线性优化问题。

设 A 为 n 阶实对称矩阵，对于非零向量 $\mathbf{p}, \mathbf{q} \in R^n$ ，如果 $\mathbf{p}^T A \mathbf{q} = 0$ ，则称 \mathbf{p}, \mathbf{q} 是相互共轭的。对于一组非零向量 $\mathbf{p}^j \in R^n, j=1, 2, \dots, n$ ，若 $i \neq j$ 时有 $(\mathbf{p}^i)^T A \mathbf{p}^j = 0$ ，则称它们是 A 的共轭方向组。

若 $\mathbf{p}^1, \mathbf{p}^2, \dots, \mathbf{p}^n$ 是一组共轭方向，则它们一定是线性无关的；反之，从 R^n 的任一组 n 个线性无关的向量出发，可以构造 A 的一组共轭方向 $\mathbf{p}^1, \mathbf{p}^2, \dots, \mathbf{p}^n$ 。

对于二次严格凸函数（即矩阵 A 是对称正定）最优化问题

$$\min f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T A \mathbf{x} + \mathbf{b}^T \mathbf{x} + c \quad (13.3.14)$$

若 $\mathbf{p}^1, \mathbf{p}^2, \dots, \mathbf{p}^n$ 为任一组 A 的共轭方向，则从任意初始值 \mathbf{x}^0 出发，依次沿共轭方向 $\mathbf{p}^1, \mathbf{p}^2, \dots, \mathbf{p}^n$ 进行一维精确搜索，则最多经过 n 次迭代可以得到(13.3.14)的最优解。

由于 A 的共轭方向组 $\{\mathbf{p}^1, \mathbf{p}^2, \dots, \mathbf{p}^n\}$ 选择具有很大的随意性，因此按不同方式产生一组共轭方向就导致了不同的共轭方向法。最常用的是共轭梯度法。共轭梯度法是以迭代点处的负梯度向量为基础而产生的一组共轭方向。共轭梯度法求解非线性优化问题的步骤如下：

共轭梯度法

- 1) 选取初始点 \mathbf{x}^0 ，给定终止控制常数 $\varepsilon > 0$ ；
- 2) 计算 $\nabla f(\mathbf{x}^0)$ ，若 $\|\nabla f(\mathbf{x}^0)\| \leq \varepsilon$ ，停止迭代，输出 \mathbf{x}^0 ，否则进行下一步；
- 3) 令 $\mathbf{p}^0 = -\nabla f(\mathbf{x}^0)$ ，令 $k := 0$ ；
- 4) 一维搜索，求 t_k 使 $f(\mathbf{x}^k + t_k \mathbf{p}^k) = \min_{t \geq 0} f(\mathbf{x}^k + t \mathbf{p}^k)$ 。令 $\mathbf{x}^{k+1} = \mathbf{x}^k + t_k \mathbf{p}^k$ ；
- 5) 计算 $\nabla f(\mathbf{x}^{k+1})$ ，若 $\|\nabla f(\mathbf{x}^{k+1})\| \leq \varepsilon$ ，停止迭代，输出 \mathbf{x}^{k+1} ，否则，进行下一步；
- 6) 若 $k+1 = n$ ，令 $\mathbf{x}^0 := \mathbf{x}^{k+1}$ ，转步骤 3)；否则，进行下一步；
- 7) 计算共轭梯度方向： $\mathbf{p}^{k+1} = -\nabla f(\mathbf{x}^{k+1}) + \lambda_k \mathbf{p}^k$ ，令 $k := k+1$ ，转步骤 4)，其中

$$\lambda_k = \frac{\|\nabla f(\mathbf{x}^{k+1})\|^2}{\|\nabla f(\mathbf{x}^k)\|^2}$$

13.3.5 L-M 方法

考虑函数关系 $\mathbf{Y} = f(\mathbf{X})$ ，其中 \mathbf{Y} 是 n 维向量， \mathbf{X} 是 m 维向量。通常 \mathbf{Y} 被看作是接近于真值 $\bar{\mathbf{Y}}$ 的测量向量，由于测量误差，不存在变量 \mathbf{X} 使得关系式 $\mathbf{Y} = f(\mathbf{X})$ 严格地满足。我们只能求解 \mathbf{X} 使得这个关系式最佳近似满足。确切地说，求最小化问题

$$\min_{\mathbf{X}} \|\mathbf{Y} - f(\mathbf{X})\| \quad (13.3.16)$$

Levenberg-Marquardt 方法是一阶 Newton 迭代算法的一种改进，主要目的是为大参数化问题 (13.3.16) 提供快速收敛的正则化方法。该方法可以看作是 Newton 一阶迭代方法和最速下降方法的结合产物。

先介绍求解 (13.3.16) 问题的一阶 Newton 迭代方法。给定 (13.3.16) 一个初始解 \mathbf{X}^k ，令 $\boldsymbol{\varepsilon}^k = \mathbf{Y} - f(\mathbf{X}^k)$ 。假定 f 在 \mathbf{X}^k 附近的值可由

$$f(\mathbf{X}^k + \Delta^k) = f(\mathbf{X}^k) + \mathbf{J}_{\mathbf{X}^k} \Delta^k$$

近似（一阶近似），其中 $\mathbf{J}_{\mathbf{X}^k} = \left. \frac{\partial \mathbf{Y}}{\partial \mathbf{X}} \right|_{\mathbf{X}=\mathbf{X}^k}$ 是 f 的 Jacobi 矩阵在 \mathbf{X}^k 的值。寻找下一个迭代点 $\mathbf{X}^{k+1} = \mathbf{X}^k + \Delta^k$ 使得

$$\|\mathbf{Y} - f(\mathbf{X}^{k+1})\| = \min_{\Delta^k} \|\mathbf{Y} - f(\mathbf{X}^k) - \mathbf{J}_{\mathbf{X}^k} \Delta^k\| = \min_{\Delta^k} \|\boldsymbol{\varepsilon}^k - \mathbf{J}_{\mathbf{X}^k} \Delta^k\|$$

该最小化的正规化方程是

$$\mathbf{N}^k \Delta^k = \mathbf{J}_{\mathbf{X}^k}^T \mathbf{J}_{\mathbf{X}^k} \Delta^k = \mathbf{J}_{\mathbf{X}^k}^T \boldsymbol{\varepsilon}^k, (\mathbf{N}^k = \mathbf{J}_{\mathbf{X}^k}^T \mathbf{J}_{\mathbf{X}^k}) \quad (13.3.16)$$

因此，一阶 Newton 方法的迭代点列为

$$\mathbf{X}^{k+1} = \mathbf{X}^k + \mathbf{J}_{\mathbf{X}^k}^+ (\mathbf{Y} - f(\mathbf{X}^k)) \quad (13.3.17)$$

与二阶 Newton 方法一样，一阶 Newton 方法的收敛行为也强烈地依赖于初始值 \mathbf{X}^0 。

Levenberg-Marquardt 方法，是用所谓的增量正规方程

$$\bar{\mathbf{N}} \Delta = \mathbf{J}_{\mathbf{X}} \boldsymbol{\varepsilon} \quad (16.3.18)$$

代替 Newton 方法中的正规方程 (13.3.16)，其中 $\bar{N}_{ii} = (1 + \lambda) N_{ii}$ ，当 $i \neq j$ 时， $\bar{N}_{ij} = N_{ij}$ 。

在 L-M 方法中，每一次迭代是寻找一个合适的 λ 值。开始时，通常取 $\lambda = 10^{-3}$ ，若增量正规方程 (13.3.18) 的解 Δ 导致误差 $\boldsymbol{\varepsilon}$ 减少，则接受 λ 的当前值并在下一次迭代中以 $\lambda/10$ 代替 λ 。若 (16.3.18) 的解 Δ 导致误差的增大，则以 10λ 代替 λ 并重新求解增量方程 (16.3.18)，这个过程一直到求出一个使误差 $\boldsymbol{\varepsilon}$ 下降的 Δ 为止。对不同的 λ ，重复地解增量正规方程，直到得出一个可接受的 Δ ，构成 L-M

方法的一次迭代。

L-M 算法

- 1) 选取初始点 \mathbf{X}^0 ，终止控制常数 ε ，计算 $\varepsilon^0 = \|\mathbf{Y} - f(\mathbf{X}^0)\|$ ，令 $k := 0, \lambda_0 = 10^{-3}$ ；
- 2) 计算 $\mathbf{J}_{\mathbf{X}^k} = \left. \frac{\partial \mathbf{Y}}{\partial \mathbf{X}} \right|_{\mathbf{X}=\mathbf{X}^k}$ ，构造增量正规方程： $\bar{\mathbf{N}}(\lambda^k) \Delta^k = \mathbf{J}_{\mathbf{X}^k} \varepsilon^k$ ；
- 3) 通过求解增量正规方程，得到 Δ^k 。
 - 3.1) 若 $\|\mathbf{Y} - f(\mathbf{X}^k + \Delta^k)\| < \varepsilon^k$ ，令 $\mathbf{X}^{k+1} = \mathbf{X}^k + t_k \mathbf{P}^k$ ，若 $\|\Delta^k\| < \varepsilon$ ，停止迭代，输出结果；
否则，令 $\lambda^{k+1} = \lambda^k / 10$ ，置 $k := k + 1$ ，转到第 2) 步。
 - 3.2) 若 $\|\mathbf{Y} - f(\mathbf{X}^k + \Delta^k)\| \geq \varepsilon^k$ ，则令 $\lambda^{k+1} = 10\lambda^k$ ，重新解正规方程得 Δ^k ，返回 3.1)。

13.4 约束优化

13.4.1 惩罚法

惩罚法，按照 Fiacco 和 McCormik 的说法，又称为序列无约束极小化方法。该方法是通过引入罚函数构造增广代价函数，将给定的约束优化问题代为一系列无约束优化问题，然后通过求解一系列的无约束优化问题而获得原约束优化问题的解。惩罚法有多种类型，这里仅介绍两个常用的类型：外惩罚法(又称为罚函数法)和内惩罚法(也称为障碍函数法)。

外惩罚法

将约束优化问题：

$$\begin{cases} \min f(\mathbf{x}) \\ \text{subject to } g_i(\mathbf{x}) \leq 0, i = 1, 2, \dots, p; \\ h_j(\mathbf{x}) = 0, i = 1, 2, \dots, q \end{cases} \quad (13.4.1)$$

的可行域记为

$$\Omega = \{\mathbf{x} \in R^n \mid g_i(\mathbf{x}) \leq 0, h_j(\mathbf{x}) = 0, i = 1, 2, \dots, p; j = 1, 2, \dots, q\}$$

在本小节，假定约束优化问题中所涉及的函数都是连续的。外部惩罚法的基本思想：首先选择一个充分大的数 C ，构造一个罚函数

$$b(\mathbf{x}) = \begin{cases} 0, & \mathbf{x} \in \Omega \\ C, & \mathbf{x} \notin \Omega \end{cases} \quad (13.4.2)$$

其中 C 称为惩罚因子，表示对不可行点的惩罚。然后将原问题归结为求下述无约束优化问题：

$$\min p(\mathbf{x}) \quad (13.4.3)$$

其中

$$p(\mathbf{x}) = f(\mathbf{x}) + b(\mathbf{x}) = \begin{cases} f(\mathbf{x}), & \mathbf{x} \in \Omega \\ C, & \mathbf{x} \notin \Omega \end{cases}$$

并称它为(13.4.1)的增广代价函数。

由于 C 是一个充分大的数，所以函数 $p(\mathbf{x})$ 的最优解也必然是原问题的最优解。然而，罚函数(13.4.2)在可行域 Ω 的边界具有非常大的跳跃，也就是说它不能保持原代价函数在可行域边界的连续性，而约束优化问题的最优化解通常是可行域的边界点，这使得我们无法使用无约束优化方法求解问题(13.4.3)。

为了使增广代价函数 $p(\mathbf{x})$ 在可行域边界具有原代价函数的性能，必须使 $p(\mathbf{x})$ 在可行域的边界不发生跳跃。如果选择罚函数为：

$$b_c(\mathbf{x}) = C \sum_{i=1}^p [\max(g_i(\mathbf{x}), 0)]^2 + \frac{C}{2} \sum_{i=1}^q (h_i(\mathbf{x}))^2 \quad (13.4.4)$$

则相应的增广代价函数 $p_c(\mathbf{x}) = f(\mathbf{x}) + b_c(\mathbf{x})$ 在可行域边界就具有连续性。因此，只要 C 充分大，原问题就可以归结为无约束问题：

$$\min p_c(\mathbf{x}) = f(\mathbf{x}) + b_c(\mathbf{x}) \quad (13.4.5)$$

在实际应用中，惩罚因子 C 究竟选择多才合适，这并非容易解决的问题。为此，人们作了如下改进。选择一个单调严格递增趋于正无穷的惩罚因子序列： $\{C_k \mid k = 1, 2, \dots\}$ ，此时，随着 k 的增大 $b_c(\mathbf{x})$ 对不可行点的惩罚 C_k 也将增大，并且 k 趋于正无穷时惩罚 C_k 也将趋于正无穷。这样，就求解原约束优化问题归结为求解一系列无约束优化问题：

$$\min p_{c_k}(\mathbf{x}), k = 1, 2, \dots \quad (13.4.6)$$

其中

$$p_{c_k}(\mathbf{x}) = f(\mathbf{x}) + b_{c_k}(\mathbf{x}), b_{c_k}(\mathbf{x}) = C_k \sum_{i=1}^p [\max(g_i(\mathbf{x}), 0)]^2 + \frac{C_k}{2} \sum_{i=1}^q (h_i(\mathbf{x}))^2 \quad (13.4.7)$$

对于(13.4.6)的最优解 $\{\mathbf{x}^k\}$ ，我们期望它能趋近于原问题的最优解。这就是外惩罚法的整个过程。外惩罚法名字的由来，是在算法整个过程中对可行域外部点加以惩罚（即对不可行点的惩罚）。

外惩罚法的具体计算步骤如下：

- 1) 选取初始点 \mathbf{x}^0 ，给定终止控制常数 $\varepsilon > 0$ 和惩罚因子序列 $\{C_k \mid k = 1, 2, \dots\}$ ； $k := 1$ ；
- 2) 按(13.4.7)构造惩罚函数 $b_{c_k}(\mathbf{x})$ 和增广代价函数 $p_{c_k}(\mathbf{x})$ ；
- 3) 以初始点 \mathbf{x}^{k-1} ，用无约束优化方法求 $\min p_{c_k}(\mathbf{x})$ ，得到它的最优解 \mathbf{x}^k 。若 \mathbf{x}^k 已满足终止

条件，则输出最优解 \mathbf{x}^k ；否则，令 $k := k + 1$ ，转步骤 2)。

注：在实际应用中， $\{C_k \mid k = 1, 2, \dots\}$ 可按下述递推方式产生：

$$C_{k+1} = \sigma C_k \quad (C_1 > 0, \sigma \geq 2)$$

对于终止条件有多种选取方式，如：令

$$S(\mathbf{x}) = \frac{1}{C_k} p_{c_k}(\mathbf{x})$$

终止条件可选取为 $S(\mathbf{x}) \leq \varepsilon$ 。

$$g^k = \max\{g_i(\mathbf{x}^k)\}, h^k = \max\{h_j(\mathbf{x}^k)\}$$

终止条件可选取为 $\max\{g^k, h^k\} \leq \varepsilon$ ；

内惩罚法

内部惩罚法的基本思想与外部惩罚法的思想相类似，首先在可行域边界设置一道“障碍”，充当这道“障碍”的是所谓的障碍函数，然后由构造增广价函数将原问题归结为求解无约束优化问题。障碍函数的作用是使得在解无约束优化问题时的迭代点总是可行点。

为了陈述方便起见，我们仅考虑不等式约束优化问题：

$$\begin{cases} \min f(\mathbf{x}) \\ \text{subject to } g_i(\mathbf{x}) \leq 0, i = 1, 2, \dots, p \end{cases} \quad (13.4.8)$$

它的可行域的内部是

$$\Omega^o = \{\mathbf{x} \in R^n \mid g_i(\mathbf{x}) < 0, i = 1, 2, \dots, p\}$$

当可行域的内点 $\mathbf{x} \in \Omega^o$ 趋近边界时，在 $\{g_i(\mathbf{x})\}$ 中至少存在一个函数值趋近于零，因此下述函数趋近于无穷大：

$$b(\mathbf{x}) = -\sum_{i=1}^p \frac{1}{g_i(\mathbf{x})} \quad \text{或} \quad b(\mathbf{x}) = -\sum_{i=1}^p \log(-g_i(\mathbf{x}))$$

这个函数被称为障碍函数。构造增广代价函数如下：

$$p(\mathbf{x}) = f(\mathbf{x}) + b(\mathbf{x})$$

不难看出 $p(\mathbf{x})$ 极小点总是在可行域的内部，所以，障碍函数 $b(\mathbf{x})$ 的作用是对哪些企图越出可行域的点加以惩罚。

由于我们最终要寻求原约束问题的最优解，而最优解往往在可行域的边界上，所以在迭代过程中要逐渐减小 $b(\mathbf{x})$ 的惩罚程度。为此，与外惩罚法类似，首先选取一个单调减趋于零的正惩罚因子序列 $\{D_k \mid k = 1, 2, \dots\}$ ，并对每一个 k ，构造障碍函数：

$$b_{D_k}(\mathbf{x}) = -D_k \sum_{i=1}^p \frac{1}{g_i(\mathbf{x})} \quad \text{或} \quad b_{D_k}(\mathbf{x}) = -D_k \sum_{i=1}^p \log(-g_i(\mathbf{x})) \quad (13.4.9)$$

然后构造定义在 Ω^o 上增广代价函数：

$$p_{D_k}(\mathbf{x}) = f(\mathbf{x}) + b_{D_k}(\mathbf{x}) \quad (13.4.10)$$

由 $b_{D_k}(\mathbf{x})$ 构造可知，当一个点从可行域内部趋向可行域边界时， $p_{D_k}(\mathbf{x})$ 将无限增大，因此优化问题：

$$\min p_{D_k}(\mathbf{x}) \quad (13.4.11)$$

的最优解总是在可行域内部。如果(13.4.8)的最优解在可行域内部，则当 D_k 取到适当值时， $\min p_{D_k}(\mathbf{x})$ 的最优解 \mathbf{x}^k 可以达到(13.4.8)的最优解；如果(13.4.8)的最优解在可行域边界上，因随 k 增大 $b_{D_k}(\mathbf{x})$ 的影响逐渐减弱，所以 $\{\mathbf{x}^k\}$ 将逼近(13.4.8)的最优解。于是，求解(13.4.8)归结为求解下述一系列的无约束优化问题：

$$\min p_{D_k}(\mathbf{x}), k = 1, 2, \dots \quad (13.4.12)$$

内惩罚法的具体计处步骤如下：

- 1) 选取初始点 $\mathbf{x}^0 \in \Omega^o$ ，给定终止控制常数 $\varepsilon > 0$ 和单调减趋于零的正惩罚因子序列 $\{D_k | k = 1, 2, \dots\}$ ； $k := 1$ ；
- 2) 按(13.4.9)和(13.4.10)构造惩罚函数 $b_{D_k}(\mathbf{x})$ 和增广代价函数 $p_{D_k}(\mathbf{x})$ ；
- 3) 以初始点 \mathbf{x}^{k-1} ，用无约束优化方法求 $\min p_{D_k}(\mathbf{x})$ ，得到它的最优解 \mathbf{x}^k 。若 \mathbf{x}^k 已满足终止条件，则输出最优解 \mathbf{x}^k ；否则，令 $k := k + 1$ ，转步骤 2)。

注：在内惩罚法中，初始点必须是可行域的内点。 $\{D_k | k = 1, 2, \dots\}$ 可按下述递推方式产生：

$$D_{k+1} = \sigma^{-1} D_k \quad (C_1 > 0, \sigma \geq 2)$$

终止条件可选取为 $b_{D_k}(\mathbf{x}^k) \leq \varepsilon$ 或 $g^k = \max\{|g_i(\mathbf{x}^k)|\} \leq \varepsilon$ 。

惩罚法的优点是方法结构较简单，但存在下述缺点：(1) 收敛速度慢；(2) 计算量大，每一次迭代都需要求解一个无约束优化问题；(3) 方法自身造成了数值计算上的困难，因为在求解过程中要求惩罚因子无限增大或无限减小，这可能导致相应增广代价函数的 Hessian 矩阵的严重病态，直接影响了惩罚法的效率，甚至算法失败。下一节所要介绍的求解约束优化的乘子法，它利用了惩罚法的思想，但它有效地克服了惩罚法的缺点。乘子法已被普遍认为是求解约束优化的最好方法之一。

13.4.2 乘子法

Hestenes 在 1969 年，针对等式约束优化问题提出了著名的 Hestenes 乘子法，它借用了惩罚法中构造增广代价函数的思想，并利用 Lagrange 乘子法克服了惩罚法所固有的数值困难。后来，Buys, Bertsekas, Rockafellar 等人将 Hestenes 乘子法拓广到一般约束优化问题。我们先介绍 Hestenes 乘子法。

Hestenes 乘子法

考虑等式约束优化问题：

$$\begin{cases} \min f(\mathbf{x}) \\ \text{subject to } h_i(\mathbf{x}) = 0, i = 1, 2, \dots, q \end{cases} \quad (13.4.13)$$

在惩罚法中，惩罚因子无限增大是其方法本身所固有的，现分析如下：

假定(13.4.13)中所涉及的函数都是连续可微的。惩罚法中的增广代价函数是

$$p_{c_k}(\mathbf{x}) = f(\mathbf{x}) + \frac{C_k}{2} \sum_{i=1}^q (h_i(\mathbf{x}))^2$$

令 $\min p_{c_k}(\mathbf{x}), k = 1, 2, \dots$ 的最优解为 \mathbf{x}^k ，则必有

$$\nabla p_{c_k}(\mathbf{x}^k) = \nabla f(\mathbf{x}^k) + C_k \sum_{i=1}^q h_i(\mathbf{x}^k) \nabla h_i(\mathbf{x}^k) = 0$$

因此

$$\frac{1}{C_k} \nabla f(\mathbf{x}^k) = - \sum_{i=1}^q h_i(\mathbf{x}^k) \nabla h_i(\mathbf{x}^k)$$

假定 $\mathbf{x}^k \rightarrow \mathbf{x}^* (k \rightarrow \infty)$ ，其中 \mathbf{x}^* 为(13.4.13)的最优解，于是在上式两边取极限，我们有

$$\lim_{k \rightarrow \infty} \frac{1}{C_k} \nabla f(\mathbf{x}^k) = - \sum_{i=1}^q h_i(\mathbf{x}^*) \nabla h_i(\mathbf{x}^*) = 0$$

对于约束优化问题，一般有 $\nabla f(\mathbf{x}^*) \neq 0$ ，所以 $C_k \rightarrow \infty$ 。由此可见，惩罚法的数值困难是其自身所固有的。

惩罚法的惩罚因子无限增大的本质是 $\nabla f(\mathbf{x}^*) \neq 0$ 。如果我们能够在不改变最优解 \mathbf{x}^* 的前提下，用在最优点处导数为零的函数代替原问题的代价函数，再对此函数应用惩罚法，则惩罚因子就不会无限增大了。这就是乘子法的基本思想。下面，讨论如何选择这样的函数。

根据最优性的必要条件，必存在 Lagrange 乘子 $\boldsymbol{\mu}^*$ 使得问题(13.4.13)的 Lagrange 函数

$$L(\mathbf{x}, \boldsymbol{\mu}) = f(\mathbf{x}) + \sum_{i=1}^q \mu_i h_i(\mathbf{x})$$

在点 $(\mathbf{x}^*, \boldsymbol{\mu}^*)$ 的梯度为零，即

$$\nabla L(\mathbf{x}^*, \boldsymbol{\mu}^*) = \begin{pmatrix} \nabla_{\mathbf{x}} L(\mathbf{x}^*, \boldsymbol{\mu}^*) \\ \nabla_{\boldsymbol{\mu}} L(\mathbf{x}^*, \boldsymbol{\mu}^*) \end{pmatrix} = 0$$

然而，点 $(\mathbf{x}^*, \boldsymbol{\mu}^*)$ 不是函数 $L(\mathbf{x}, \boldsymbol{\mu})$ 的极小点或极大点，而是一个鞍点，即有

$$L(\mathbf{x}^*, \boldsymbol{\mu}) \leq L(\mathbf{x}^*, \boldsymbol{\mu}^*) \leq L(\mathbf{x}, \boldsymbol{\mu}^*)$$

于是， \mathbf{x}^* 为 $L(\mathbf{x}, \boldsymbol{\mu}^*)$ 函数的极小点。因此， $L(\mathbf{x}, \boldsymbol{\mu}^*)$ 就是我们所要寻求的函数。

综上所述，问题(13.4.13)与下述问题等价：

$$\begin{cases} \min L(\mathbf{x}, \boldsymbol{\mu}^*) \\ \text{subject to } h_i(\mathbf{x}) = 0, i = 1, 2, \dots, q \end{cases} \quad (13.4.14)$$

再将惩罚法应用于上述问题，原问题(13.4.13)就转化为求解增广 Lagrange 代价函数

$$p_c(\mathbf{x}, \boldsymbol{\mu}^*) = L(\mathbf{x}, \boldsymbol{\mu}^*) + \frac{C}{2} \sum_{j=1}^q [h_j(\mathbf{x})]^2 \quad (13.4.15)$$

的无约束极小问题。

现在的困难是，在数值上如何确定 $\boldsymbol{\mu}^*$ 和 C ，特别是确定 $\boldsymbol{\mu}^*$ 。对此，Hestenes 提出了如下方法：对惩罚因子 C 选取一个无限增大的正数序列 $\{C_k\}$ ，使之随迭代次增加而增大；对乘子 $\boldsymbol{\mu}^*$ ，先给定一个初始值，然后再迭代过程中不断更新它。下面给出 Hestenes 的乘子 $\boldsymbol{\mu}^*$ 更新公式。

假定我们已有 $\boldsymbol{\mu}^k$ ，并令 $\min p_{C_k}(\mathbf{x}, \boldsymbol{\mu}^*)$ 的最优解为 \mathbf{x}^k ，则必有

$$\begin{aligned} \nabla_x p_{C_k}(\mathbf{x}^k, \boldsymbol{\mu}^k) &= \nabla_x L(\mathbf{x}^k, \boldsymbol{\mu}^k) + C_k \sum_{j=1}^q h_j(\mathbf{x}^k) \nabla h_j(\mathbf{x}^k) \\ &= \nabla f(\mathbf{x}^k) + \sum_{j=1}^q \mu_j^k \nabla h_j(\mathbf{x}^k) + C_k \sum_{j=1}^q h_j(\mathbf{x}^k) \nabla h_j(\mathbf{x}^k) \\ &= \nabla f(\mathbf{x}^k) + \sum_{j=1}^q (\mu_j^k + C_k h_j(\mathbf{x}^k)) \nabla h_j(\mathbf{x}^k) = 0 \end{aligned} \quad (13.4.16)$$

为了使 \mathbf{x}^k ， $\boldsymbol{\mu}^k$ 逼近最优解 \mathbf{x}^* 与乘子 $\boldsymbol{\mu}^*$ ，考虑 $\nabla_x L(\mathbf{x}^*, \boldsymbol{\mu}^*)$ ，我们有

$$\nabla_x L(\mathbf{x}^*, \boldsymbol{\mu}^*) = \nabla f(\mathbf{x}^*) + \sum_{j=1}^q \mu_j^* \nabla h_j(\mathbf{x}^*) \quad (13.4.17)$$

比较(13.4.16)与(13.4.17)，在下一迭代中 $\boldsymbol{\mu}^{k+1}$ 应为

$$\mu_j^{k+1} = \mu_j^k + C_k h_j(\mathbf{x}^k), j = 1, 2, \dots, q \quad (13.4.18)$$

这就是 Hestenes 的乘子 $\boldsymbol{\mu}^*$ 迭代更新公式。

Hestenes 乘子法的具体计算步骤如下：

- 1) 选取初始点 \mathbf{x}^0 与初始乘子 $\boldsymbol{\mu}^1$ ，给定终止控制常数 $\varepsilon > 0$ 和单调增趋于无穷大的正惩罚因子序列 $\{C_k \mid k = 1, 2, \dots\}$ ； $k := 1$ ；
- 2) 以初始点 \mathbf{x}^{k-1} ，用无约束优化方法求 $\min p_{C_k}(\mathbf{x}, \boldsymbol{\mu}^k)$ ，得到它的最优解 \mathbf{x}^k 。

其中：

$$p_{C_k}(\mathbf{x}, \boldsymbol{\mu}^k) = f(\mathbf{x}) + \sum_{j=1}^q \mu_j^k h_j(\mathbf{x}) + \frac{C_k}{2} \sum_{j=1}^q [h_j(\mathbf{x})]^2$$

若 \mathbf{x}^k 已满足终止条件，则输出最优解 \mathbf{x}^k ；否则，转步骤 3)；

3) 更新乘子： $\mu_j^{k+1} = \mu_j^k + C_k h_j(\mathbf{x}^k)$, $j = 1, 2, \dots, q$ ，令 $k := k + 1$ ，转步骤 2)。

注：在实际应用中， $\{C_k \mid k = 1, 2, \dots\}$ 可按下述递推方式产生：

$$C_{k+1} = \sigma C_k \quad (C_1 \in [0.1, 1], \sigma \geq 2)$$

初始乘子常取零。

对于终止条件有多种选取方式，如： $h^k = \max\{|h_j(\mathbf{x}^k)|\} \leq \varepsilon$ ，或者 $\|\mathbf{x}^{k+1} - \mathbf{x}^k\| \leq \varepsilon$ 且 $\|\boldsymbol{\mu}^{k+1} - \boldsymbol{\mu}^k\| \leq \varepsilon$ ，或者 $|f(\mathbf{x}^{k+1}) - f(\mathbf{x}^k)| \leq \varepsilon$ 。

Rockafellar 乘子法

Rockafellar 将 Hestenes 乘子法拓广到不等式约束优化问题：

$$\begin{cases} \min f(\mathbf{x}) \\ \text{subject to } g_i(\mathbf{x}) \leq 0, i = 1, 2, \dots, p \end{cases} \quad (13.4.19)$$

Rockafellar 乘子法首先引入松弛变量 $\mathbf{y} = (y_1, y_2, \dots, y_p)^T$ 将不等式约束优化(13.4.19)转化为变量 (\mathbf{x}, \mathbf{y}) 的等式约束优化：

$$\begin{cases} \min f(\mathbf{x}) \\ \text{subject to } g_i(\mathbf{x}) + y_i^2 = 0, i = 1, 2, \dots, p \end{cases} \quad (13.4.20)$$

根据 Hestenes 乘子法，(13.4.20)的增广 Lagrange 代价函数为

$$p_{C_k}(\mathbf{x}, \mathbf{y}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \sum_{i=1}^p \lambda_i (g_i(\mathbf{x}) + y_i^2) + \frac{C_k}{2} \sum_{i=1}^p (g_i(\mathbf{x}) + y_i^2)^2 \quad (13.4.21)$$

其中 $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_p)^T$ 是乘子， $\{C_k\}$ 是一个单调增趋于无穷的正惩罚因子序列。逼近最优解的迭代点列 $\{(\mathbf{x}^k, \mathbf{y}^k)\}$ 与乘子的迭代点列 $\{\boldsymbol{\lambda}^k\}$ 分别为

$(\mathbf{x}^k, \mathbf{y}^k)$ 是 $\min p_{C_k}(\mathbf{x}, \mathbf{y}, \boldsymbol{\lambda}^k)$ 的最优解，

$$\lambda_j^{k+1} = \lambda_j^k + C_k g_j(\mathbf{x}^k + (y_i^k)^2), j = 1, 2, \dots, p。$$

在 Rockafellar 乘子法中，将进一步消去松弛变量 \mathbf{y} 。为此，考虑函数 $p_{C_k}(\mathbf{x}, \mathbf{y}, \boldsymbol{\lambda})$ 关于变量 \mathbf{y} 极小，令

$$L(\mathbf{x}, \boldsymbol{\mu}) = \min_{\mathbf{y}} p_{C_k}(\mathbf{x}, \mathbf{y}, \boldsymbol{\lambda})$$

从下述方程：

$$\nabla_{\mathbf{y}} p_{C_k}(\mathbf{x}, \mathbf{y}, \boldsymbol{\lambda}) = \begin{pmatrix} y_1(\lambda_1 + C_k(g_1(\mathbf{x}) + y_1^2)) \\ y_2(\lambda_2 + C_k(g_2(\mathbf{x}) + y_2^2)) \\ \vdots \\ y_p(\lambda_p + C_k(g_p(\mathbf{x}) + y_p^2)) \end{pmatrix} = 0$$

得到

$$y_i^2 = \begin{cases} -\frac{1}{C_k}(\lambda_i + C_k(g_i(\mathbf{x}))), & \text{若 } (\lambda_i + C_k(g_i(\mathbf{x}))) < 0 \\ 0, & \text{若 } (\lambda_i + C_k(g_i(\mathbf{x}))) \geq 0 \end{cases}, i = 1, 2, \dots, p \quad (13.4.22)$$

将上式代入 $p_{C_k}(\mathbf{x}, \mathbf{y}, \boldsymbol{\lambda})$ 中的相应项，可推知：

$$\begin{aligned} & \lambda_i(g_i(\mathbf{x}) + y_i^2) + \frac{C_k}{2}(g_i(\mathbf{x}) + y_i^2)^2 \\ = & \begin{cases} -\frac{\lambda_i^2}{2C_k}, & \text{若 } (\lambda_i + C_k(g_i(\mathbf{x}))) < 0 \\ \lambda_i g_i(\mathbf{x}) + \frac{C_k}{2}(g_i(\mathbf{x}))^2, & \text{若 } (\lambda_i + C_k(g_i(\mathbf{x}))) \geq 0 \end{cases} \\ = & \frac{1}{2C_k}[(\max(0, \lambda_i + C_k(g_i(\mathbf{x}))))^2 - \lambda_i^2], i = 1, 2, \dots, p \end{aligned}$$

因此，我们有

$$L(\mathbf{x}, \boldsymbol{\mu}) = \min_{\mathbf{y}} p_{C_k}(\mathbf{x}, \mathbf{y}, \boldsymbol{\mu}) = f(\mathbf{x}) + \frac{1}{2C_k} \sum_{i=1}^p [(\max(0, \lambda_i + C_k g_i(\mathbf{x})))^2 - \lambda_i^2] \quad (13.4.23)$$

所以，迭代式(13.4.22)得到的 \mathbf{x}^k 是 $\min_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\mu}^k)$ 的最优解。于是，乘子迭代式为：

$$\lambda_i^{k+1} = \lambda_i^k + C_k g_i(\mathbf{x}^k) + (y_i^k)^2 = \max(0, \lambda_i^k + C_k g_i(\mathbf{x}^k)), i = 1, 2, \dots, p \quad (13.4.24)$$

这样，我们就得到以乘子法。现将整个计算步骤概括如下：

- 1) 选取初始点 \mathbf{x}^0 与初始乘子 $\boldsymbol{\mu}^1$ ，给定终止控制常数 $\varepsilon > 0$ 和单调增趋于无穷大的正惩罚因子序列 $\{C_k \mid k = 1, 2, \dots\}$ ； $k := 1$ ；
- 2) 以初始点 \mathbf{x}^{k-1} ，用无约束优化方法求 $\min L_{C_k}(\mathbf{x}, \boldsymbol{\mu}^k)$ ，得到它的最优解 \mathbf{x}^k 。

其中：

$$L_{C_k}(\mathbf{x}, \boldsymbol{\mu}^k) = f(\mathbf{x}) + \frac{1}{2C_k} \sum_{i=1}^p [(\max(0, \lambda_i + C_k g_i(\mathbf{x})))^2 - \lambda_i^2]$$

若 \mathbf{x}^k 已满足终止条件，则输出最优解 \mathbf{x}^k ；否则，转步骤 3)；

3) 更新乘子： $\lambda_i^{k+1} = \max(0, \lambda_i^k + C_k g_i(\mathbf{x}^k))$, $i = 1, 2, \dots, p$ ，令 $k := k + 1$ ，转步骤 2)。

在 Rockafellar 乘子法中，惩罚因子序列和终止条件的选择可参考 Hestenes 乘子法。

对于一般约束优化问题：

$$\begin{cases} \min f(\mathbf{x}) \\ \text{subject to } g_i(\mathbf{x}) \leq 0, i = 1, 2, \dots, p; \\ h_j(\mathbf{x}) = 0, j = 1, 2, \dots, q \end{cases} \quad (13.4.25)$$

它的增广代价函数序列为：

$$L_{C_k}(\mathbf{x}, \boldsymbol{\lambda}^k, \boldsymbol{\mu}^k) = f(\mathbf{x}) + \frac{1}{2C_k} \sum_{i=1}^p [(\max(0, \lambda_i^k + C_k g_i(\mathbf{x})))^2 - (\lambda_i^k)^2] + \sum_{i=1}^q \mu_i^k h_i(\mathbf{x}) + \frac{C_k}{2} \sum_{i=1}^q (h_i(\mathbf{x}))^2 \quad (13.4.26)$$

逼近(13.4.26)的最优解的点列 $\{\mathbf{x}^k\}$ 是(13.4.25)的最优解；乘子迭代公式为：

$$\lambda_i^{k+1} = \max(0, \lambda_i^k + C_k g_i(\mathbf{x}^k)), i = 1, 2, \dots, p$$

$$\mu_j^{k+1} = \mu_j^k + C_k h_j(\mathbf{x}^k), j = 1, 2, \dots, q;$$

终止条件可选取： $\sum_{i=1}^p \left| \max\left(g_i(\mathbf{x}^k), -\frac{\lambda_i^k}{C_k}\right) \right| + \sum_{j=1}^q |h_j(\mathbf{x}^k)| \leq \varepsilon$ 。

14. 参数估计

最大似然估计和贝叶斯估计分别是经典参数估计与现代参数估计中的最重要方法，它们在众多应用学科都具有广泛的应用。近年来，模型估计的贝叶斯方法已成为计算机视觉领域的热门研究课题。本章主要介绍最大似然估计和贝叶斯估计的基本数学原理，以及实现这两种估计的统计计算方法：期望/最大化（EM）算法。

14.1 最大似然估计

最大似然估计的思想来源于高斯关于误差理论的研究，当初并未受到人们的重视。直到 Fisher 于 1921 年将它作为一般估计方法提出来，经过 Fisher 自己以及许多统计学家研究发现这种估计方法相对其它方法具有更好的品质，才被人们所重视。现在，最大似然估计已成为参数估计的最重要方法之一，在众多应用学科得到广泛应用。

14.1.1 基本概念

假定 n 维随机向量 \mathbf{x} 的密度函数为 $p(\mathbf{x}; \boldsymbol{\theta})$ ，其中 $\boldsymbol{\theta}$ 是未知参数向量。现在的问题是：从 k 次独立抽样所得到的样本 $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k)$ ，估计参数向量 $\boldsymbol{\theta}$ 。由于这 k 次抽样是独立的，所以 $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k)$ 的联合密度为

$$L(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k; \boldsymbol{\theta}) = \prod_{j=1}^k p(\mathbf{x}_j; \boldsymbol{\theta}) \quad (14.1.1)$$

对于固定的 $\boldsymbol{\theta}$ ， $L(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k; \boldsymbol{\theta}) > L(\mathbf{x}'_1, \mathbf{x}'_2, \dots, \mathbf{x}'_k; \boldsymbol{\theta})$ 的意义是容易解释的，即在抽样中出现 $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k)$ 的概率大于出现 $(\mathbf{x}'_1, \mathbf{x}'_2, \dots, \mathbf{x}'_k)$ 的概率。反之，对于固定的 $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k)$ ，则 $L(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k; \boldsymbol{\theta}') > L(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k; \boldsymbol{\theta}'')$ 只表明 $\boldsymbol{\theta}$ 是 $\boldsymbol{\theta}'$ 的“可能性”比是 $\boldsymbol{\theta}''$ 的“可能性”更大一些。由于 $\boldsymbol{\theta}$ 不是随机变量，这种“可能性”不能称为概率，通常称为“似然”。对于固定的 $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k)$ ， $L(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k; \boldsymbol{\theta})$ 作为 $\boldsymbol{\theta}$ 的函数，称为似然函数，最大似然估计就是求使得似然达到最大的 $\boldsymbol{\theta}$ 值。

定义 14.1.1 令 $p(\mathbf{x}; \boldsymbol{\theta})$ 是 n 维随机向量 \mathbf{x} 的密度函数族，其中 $\boldsymbol{\theta} \in \Theta$ ， Θ 称为参数空间。对

固定 \mathbf{x} , $p(\mathbf{x}; \boldsymbol{\theta})$ 作为参数向量 $\boldsymbol{\theta}$ 的函数称为似然函数, 记为

$$L(\mathbf{x}; \boldsymbol{\theta}) = p(\mathbf{x}; \boldsymbol{\theta}), \quad \boldsymbol{\theta} \in \Theta \quad (14.1.2)$$

并称

$$\boldsymbol{\theta}^* = \arg \max_{\boldsymbol{\theta} \in \Theta} L(\mathbf{x}; \boldsymbol{\theta}) \quad (14.1.3)$$

为 $\boldsymbol{\theta}$ 的最大似然估计。

对似然函数(14.1.2)取对数, 有

$$l(\mathbf{x}; \boldsymbol{\theta}) = \log L(\mathbf{x}; \boldsymbol{\theta}) \quad (14.1.4)$$

并称它为对数似然。根据对数函数的单调性, 不难证明:

$$\boldsymbol{\theta}^* = \arg \max_{\boldsymbol{\theta} \in \Theta} l(\mathbf{x}; \boldsymbol{\theta}) \Leftrightarrow \boldsymbol{\theta}^* = \arg \max_{\boldsymbol{\theta} \in \Theta} L(\mathbf{x}; \boldsymbol{\theta})$$

即, 求最大似然等价于求最大对数似然。

引进对数似然的目的是为了更方便最大似然的计算。因为在实际问题中, $p(\mathbf{x}; \boldsymbol{\theta})$ 是同一参数空间 Θ 上若干个指数型分布族

$$p_1(\mathbf{y}_1; \boldsymbol{\theta}), p_2(\mathbf{y}_2; \boldsymbol{\theta}), \dots, p_k(\mathbf{y}_k; \boldsymbol{\theta})$$

因此, 独立抽样 $\mathbf{x} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_k)$ 的联合密度为

$$p(\mathbf{x}; \boldsymbol{\theta}) = \prod_{j=1}^k p_j(\mathbf{y}_j; \boldsymbol{\theta})$$

这样, 应用对数似然

$$l(\mathbf{x}; \boldsymbol{\theta}) = \log L(\mathbf{x}; \boldsymbol{\theta}) = \sum_{j=1}^k \log p_j(\mathbf{y}_j; \boldsymbol{\theta}) \quad (14.1.5)$$

就可以简化最大似然的求解。

例 14.1.1 设 (x_1, x_2, \dots, x_k) 是从正态总体 $N(\mu, \sigma^2)$ 中抽取的样本, 估计未知参数 μ, σ^2 。显然, 对数似然为

$$l(\mathbf{x}; \mu, \sigma^2) = -k \log \sigma - \sum_{j=1}^k (x_j - \mu)^2 / 2\sigma^2$$

因此, 由似然方程:

$$\begin{aligned} \frac{\partial l}{\partial \mu} &= \frac{1}{2\sigma^2} (\sum_{j=1}^k x_j - n\mu) = 0 \\ \frac{\partial l}{\partial \sigma^2} &= \frac{\sum_{j=1}^k (x_j - \mu)^2}{\sigma^4} - \frac{n}{2\sigma^4} = 0 \end{aligned}$$

得到 μ, σ^2 的估计值为

$$\mu^* = \sum_{j=1}^k x_j = \bar{x}, \quad \sigma^{*2} = \frac{1}{k} \sum_{j=1}^k (x_j - \bar{x})^2$$

注 在进行最大似然估计中, 应该注意参数的取值范围。例如, 在上例中, 如果限制 $\mu > 0$, 则

最大似然估计是

$$\mu^* = \max(0, \bar{x}), \sigma^{*2} = \frac{1}{k} \sum_{j=1}^k (x_j - \mu^*)^2$$

请读者自行验证。

在计算上, 最大似然估计有一个非常重要的性质: 如果 θ^* 是 θ 的最大似然, $f(\theta): \Theta \rightarrow \Phi$ 是可测函数 (如: 连续函数函数), 则 $f(\theta^*)$ 也是 $f(\theta)$ 的最大似然。这个性质被称为最大似然不变性, 应用它可简化参数函数的最大似然估计。

为了证明这一性质, 对任意 $\omega \in \Omega$, 定义 $g^{-1}(\omega) = \{\theta \mid g(\theta) = \omega\}$, 并且令

$$M(\mathbf{x}; \omega) = \max_{\theta \in g^{-1}(\omega)} L(\mathbf{x}; \theta) \quad (14.1.6)$$

式(14.1.6)通常称为诱导似然。由于 $L(\mathbf{x}; \theta^*) = \max_{\theta \in \Theta} L(\mathbf{x}; \theta)$, 且 $\theta^* \in g^{-1}(g(\theta^*))$, 所以, 对任意 $\omega \in \Omega$, 有

$$\begin{aligned} M(\mathbf{x}; g(\theta^*)) &= \max_{\theta \in g^{-1}(g(\theta^*))} L(\mathbf{x}; \theta) = L(\mathbf{x}; \theta^*) \\ &= \max_{\theta \in \Theta} L(\mathbf{x}; \theta) \geq \max_{\theta \in g^{-1}(\omega)} L(\mathbf{x}; \theta) = M(\mathbf{x}; \omega) \end{aligned}$$

这样, 我们就证明了最大似然不变性。

例 14.1.2 设 $(x_1, y_1), \dots, (x_k, y_k)$ 是从二维正态总体中抽取的样本, 这个二维正态分布的均值是零, 协方差矩阵为 $\Sigma = \sigma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, $-1 < \rho < 1$ 。求未知参数 ρ, σ^2 的最大似然估计。

由条件, 这个二维正态分布的密度为

$$p(x, y; \rho, \sigma^2) = \frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}} \exp\left(-\frac{x^2 + y^2 - 2\rho xy}{2\sigma^2(1-\rho^2)}\right)$$

引入变量:

$$\begin{cases} \alpha_1 = \frac{1}{\sigma^2(1-\rho^2)} \\ \alpha_2 = \rho\alpha_1 = \frac{\rho}{\sigma^2(1-\rho^2)} \end{cases}$$

则有 $\begin{cases} \rho = \alpha_2 / \alpha_1 \\ \frac{1}{\sigma^2\sqrt{1-\rho^2}} = \sqrt{\alpha_1^2 - \alpha_2^2} \end{cases}$ 。于是, 我们有

$$p(x, y; \alpha_1, \alpha_2) = \exp\left(\alpha_1\left(-\frac{x^2 + y^2}{2}\right) + \alpha_2 xy + \frac{1}{2} \log(\alpha_1^2 - \alpha_2^2) + \log(2\pi)\right)$$

由此, 得到似然方程:

$$\frac{2\alpha_1}{2(\alpha_1^2 - \alpha_2^2)} = \frac{1}{2k} \sum_{j=1}^k (x_j^2 + y_j^2)$$

$$\frac{2\alpha_1}{2(\alpha_1^2 - \alpha_2^2)} = \frac{1}{k} \sum_{j=1}^k (x_j y_j)$$

所以, 根据最大似然不变性, 得到 ρ, σ^2 的最大似然估计:

$$\begin{cases} \rho = \alpha_2 / \alpha_1 = \frac{\sum_{j=1}^k x_j y_j}{2 \sum_{j=1}^k (x_j^2 + y_j^2)} \\ \sigma^2 = \frac{1}{(1 - \rho^2) \sqrt{\alpha_1^2 - \alpha_2^2}} = \frac{1}{2k} \sum_{j=1}^k (x_j^2 + y_j^2) \end{cases}$$

14.1.2 相合性与渐近正态性

最大似然估计的优良性质主要表现在大样本场合, 就小样本而言还没有能说明它有什么优良性质。在大样本理论中, 最大似然估计扮演了一个十分重要的角色。本节主要介绍最大似然估计的相合性与渐近正态性。

相合性

先介绍估计量的相合性概念。估计量与样本容量是相关联的, 如果用 $\tilde{\theta}_n(x_1, x_2, \dots, x_n)$ 估计参数 θ , 不可能指望对某个 n , 均方误差 $MSE(\tilde{\theta}_n) = E[(\tilde{\theta}_n(x_1, x_2, \dots, x_n) - \theta)^2]$ 对所有的 $\theta \in \Theta$ 都任意小。但是, 当 $n \rightarrow \infty$ 时通常可实现这一点, 这就是所谓的相合性概念, 其确切的数学定义如下:

定义 14.1.2 设 $\tilde{\theta}_n(x_1, x_2, \dots, x_n)$ 是 θ 的估计, 如果当 $n \rightarrow \infty$ 时 $\tilde{\theta}_n(x_1, x_2, \dots, x_n)$ 依概率收敛于 θ , 则称 $\tilde{\theta}_n$ 是 θ 的(弱)相合估计。如果 $\tilde{\theta}_n(x_1, x_2, \dots, x_n)$ 几乎处处收敛于 θ , 则称 $\tilde{\theta}_n$ 是 θ 的强相合估计。

尽管强相合性比弱相合性的要求更强, 但在统计中通常只考虑弱相合性, 因此下面所讲的相合性均指弱相合性。相合性是估计问题的一个基本要求, 试想: 如果一个估计量无论进行多少次抽样或进行多少次观测, 都不能将未知参数估计到任意指定的精度, 那么这种估计是不能令人信服的。也就是说, 不满足相合性要求的估计通常是不被考虑的。

对于最大似然估计, 有下述定理。

定理 14.1.1 设 (x_1, x_2, \dots, x_n) 是来自 $p(\mathbf{x}; \theta)$ 的一个样本 (为简单起见, 只考虑单参数情形, 并

且参数空间是一个开区间), 令

$$l(\mathbf{y}; \theta) = \sum_{j=1}^k \log p_j(\mathbf{x}_j; \theta)$$

若 $\log p(\mathbf{x}; \theta)$ 在 Θ 上可微, 并且 $\forall \theta \neq \theta'$, 集合 $\{\mathbf{x} \mid p(\mathbf{x}; \theta) \neq p(\mathbf{x}; \theta')\}$ 有正测度, 则似然方程在 $n \rightarrow \infty$ 时以概率 1 有解, 并且此解是相合的。

证明: 因 $\{\mathbf{x} \mid p(\mathbf{x}; \theta) \neq p(\mathbf{x}; \theta')\}$ 有正测度, 根据 Jensen 不等式^{*)}, 我们有

$$E_{\theta} \left(\log \left(\frac{p(\mathbf{x}; \theta')}{p(\mathbf{x}; \theta)} \right) \right) < \log \left(E_{\theta} \left(\frac{p(\mathbf{x}; \theta')}{p(\mathbf{x}; \theta)} \right) \right) = 0$$

记参数真值为 θ_0 , 则对充分小的 $\delta > 0$, 有 $(\theta_0 - \delta, \theta_0 + \delta) \subset \Theta$, 并且

$$E_{\theta_0} \left(\log \left(\frac{p(\mathbf{x}; \theta_0 - \delta)}{p(\mathbf{x}; \theta_0)} \right) \right) < 0$$

$$E_{\theta_0} \left(\log \left(\frac{p(\mathbf{x}; \theta_0 + \delta)}{p(\mathbf{x}; \theta_0)} \right) \right) < 0$$

根据强大数定理, 在几乎处处的意义下,

$$\frac{1}{n} (l(\mathbf{x}; \theta_0 - \delta) - l(\mathbf{x}; \theta_0)) \rightarrow E_{\theta_0} \left(\log \left(\frac{p(\mathbf{x}; \theta_0 - \delta)}{p(\mathbf{x}; \theta_0)} \right) \right) < 0$$

$$\frac{1}{n} (l(\mathbf{x}; \theta_0 + \delta) - l(\mathbf{x}; \theta_0)) \rightarrow E_{\theta_0} \left(\log \left(\frac{p(\mathbf{x}; \theta_0 + \delta)}{p(\mathbf{x}; \theta_0)} \right) \right) < 0$$

又因 $l(\mathbf{x}; \theta)$ 在 $(\theta_0 - \delta, \theta_0 + \delta) \subset \Theta$ 上连续, 因而必有一个局部极大点, 记为 $\tilde{\theta}$ 。由于 $l(\mathbf{x}; \theta)$ 可微, 因此 $\tilde{\theta}$ 必是似然方程的解, 所以, 似然方程在 $n \rightarrow \infty$ 时以概率 1 有解, 并且 $|\tilde{\theta} - \theta_0| < \delta$, δ 可以是任意小的正数, 即解 $\tilde{\theta}$ 与 θ_0 是相合的。

渐近正态性

相合性是估计量在 $n \rightarrow \infty$ 时的性质。对任意有限的 n , 相合性是没有意义的, 因为它不能说明为了使 $\tilde{\theta}_n$ 达到指定精度, n 应该至少取多大。事实上, 相合估计可以不只一个, 它们之间是有差异

^{*)} Jensen 不等式: 设 f, g 为区域 Ω 上的可积函数, 且 $m \leq f \leq M, g > 0, \int_{\Omega} g d\sigma > 0$ 。若 h 是 $[m, M]$ 上的下凸函数, 则 $h \left(\frac{\int_{\Omega} f g d\sigma}{\int_{\Omega} g d\sigma} \right) \leq \frac{\int_{\Omega} g(h \circ f) d\sigma}{\int_{\Omega} g d\sigma}$; 若 h 上凸, 则取相反的不等式。

的, 这种差异通常用估计量的渐近分布的渐近方差来反映, 而最常用的渐近分布是正态分布。下面给出估计渐近正态性的数学定义。

定义 14.1.3 设 $\tilde{\theta}_n(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ 是 θ 的估计, 如果存在 $\sigma_n^2(\theta)$ 使得 $(\tilde{\theta}_n - \theta)/\sigma_n^2(\theta)$ 依测度收敛于正态分布 $N(\theta, I_k)$ (其中 k 是参数向量的维数), 则称 $\tilde{\theta}_n$ 是 θ 的渐近正态估计。并记为 $\tilde{\theta}_n \sim AN(\theta, \sigma_n^2(\theta)I_k)$ 。

对于最大似然估计, 我们有下述定理:

定理 14.1.2 设 Θ 是一个开区间, 若密度函数 $p(\mathbf{x}; \theta)$, $\theta \in \Theta$ 满足下述条件:

- (1) 在参数真值 θ_0 附近, $p(\mathbf{x}; \theta)$ 关于 θ 的一至三阶偏导数对所有 \mathbf{x} 都存在;
- (2) 在参数真值 θ_0 处, 有

$$E_{\theta_0} \left(\frac{p'_\theta(\mathbf{x}; \theta_0)}{p(\mathbf{x}; \theta_0)} \right) = 0, E_{\theta_0} \left(\frac{p''_\theta(\mathbf{x}; \theta_0)}{p(\mathbf{x}; \theta_0)} \right) = 0, I(\theta_0) = E_{\theta_0} \left(\frac{p'_\theta(\mathbf{x}; \theta_0)}{p(\mathbf{x}; \theta_0)} \right)^2 > 0$$

记 $\tilde{\theta}_n$ 为似然方程的相合解, 则在测度的意义下,

$$(\tilde{\theta}_n - \theta_0)/n^{1/2} \rightarrow N(0, I^{-1}(\theta))$$

这个定理表明, 最大似然估计的渐近方差为 $(nI(\theta))^{-1}$ 。

定理 14.1.2 可以推广到高维参数的情形。在一般情况下, 指数分布族都满足定理中的条件, 更一般的 Cramer-Rao 分布族也满足定理中的条件。因此, 人们通常都说最大似然估计具有渐近正态性。

14.1.3 混合模型

在以上两节给出了最大似然估计的基本概念以及它的一些优良性质。在 14.1.1 节, 指出最大似然估计可以通过似然方程来求解, 并且给出了两个例子。但是, 对于实际问题, 似然方程往往是高度非线性的, 不可能通过似然方程求解。本节所介绍的混合模型是计算机视觉中的常用模型, 我们将看出通过似然方程求解是十分困难的。

假定随机变量空间 \mathbf{X} 包含 K 个成份, 即 $\mathbf{X} = \mathbf{X}_1 \cup \mathbf{X}_2 \cup \dots \cup \mathbf{X}_k$, 每个成份 \mathbf{X}_k 被选择的概率为 π_k , 并且有各自的概率密度函数 $p(\mathbf{x}; \theta_k)$ 。因此, 随机变量 \mathbf{x} 的概率密度为

$$p(\mathbf{x}; \boldsymbol{\pi}, \boldsymbol{\theta}) = \sum_{k=1}^K \pi_k p(\mathbf{x}; \theta_k) \quad (14.1.7)$$

其中: $\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_K)$, $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_K)$ 是未知参数向量。

模型(14.1.7)通常称为混合模型, 因为它是密度的混合或者说是密度的加权和。给定随机变量的一组观测值: $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M)$, 我们的问题是从混合模型(14.1.7)估计未知参数向量 $\boldsymbol{\pi}, \boldsymbol{\theta}$ 的值。对于实际问题, 在估计出 $\boldsymbol{\pi}, \boldsymbol{\theta}$ 之后, 还要推断观测向量 \mathbf{x}_j 来自哪个成份 \mathbf{X}_k 。

对于观测 $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M)$, 模型(14.1.7)的似然函数为

$$L(\mathbf{X}; \boldsymbol{\pi}, \boldsymbol{\theta}) = \prod_{j=1}^M \left(\sum_{k=1}^K \pi_k p(\mathbf{x}; \boldsymbol{\theta}_k) \right) \quad (14.1.8)$$

由于(14.1.8)的对数似然方程的高度非线性, 直接求解未知参数是十分困难的。在 14.3 节, 我们将介绍求解这类问题的 EM 算法。下面给出直线拟合的概率模型作为混合模型的一个例子, 更多的应用见 17.4 节。

直线拟合

假定平面上有 K 条直线, 第 k 条直线的参数向量记为 $\mathbf{a}_k = (a_k, b_k, c_k)$, 该直线被选择的概率为 π_k 。假定随机变量是来自这 K 条直线的测量点, 随机变量空间是来自所有测量点的集合。给定第 k 条直线, 点 \mathbf{x} 由它生成的概率为 $p(\mathbf{x}; \mathbf{a}_k)$, 即点 \mathbf{x} 关于直线 \mathbf{a}_k 的条件概率密度。于是, 生成点 \mathbf{x} 的概率为

$$p(\mathbf{x}; \boldsymbol{\pi}, \mathbf{a}) = \sum_{k=1}^K \pi_k p(\mathbf{x}; \mathbf{a}_k) \quad (14.1.9)$$

图像点的测量误差通常满足零均值、 σ 方差的高斯分布, 因此, 我们假定 $p(\mathbf{x}; \mathbf{a}_k)$ 具有下述形式:

$$p(\mathbf{x}; \mathbf{a}_k) = \frac{1}{(2\pi)^{1/2} \sigma} \exp\left(-\frac{d^2(\mathbf{x}, \mathbf{a}_k)}{2\sigma^2}\right)$$

其中: $d^2(\mathbf{x}, \mathbf{a}_k)$ 是点 $\mathbf{x} = (x, y)$ 到直线 \mathbf{a}_k 之间的距离平方, 即

$$d^2(\mathbf{x}, \mathbf{a}_k) = \frac{(a_k x + b_k y + c_k)^2}{a_k^2 + b_k^2}$$

由于方差 σ 也是未知参数, 所以, 我们将生成点 \mathbf{x} 的概率(14.1.9)写成如下形式:

$$p(\mathbf{x}; \boldsymbol{\pi}, \mathbf{a}, \sigma) = \sum_{k=1}^K \pi_k p(\mathbf{x}; \mathbf{a}_k, \sigma) \quad (14.1.10)$$

其中:

$$p(\mathbf{x}; \mathbf{a}_k, \sigma) = \frac{1}{(2\pi)^{1/2} \sigma} \exp\left(-\frac{d^2(\mathbf{x}, \mathbf{a}_k)}{2\sigma^2}\right) \quad (14.1.11)$$

如果图像点集的实际测量是 \mathbf{X} , 则出现这个实际测量的概率是

$$L(\mathbf{X}; \boldsymbol{\pi}, \mathbf{a}, \sigma) = \prod_{j=1}^M \left(\sum_{k=1}^K \pi_k p(\mathbf{x}; \mathbf{a}_k, \sigma) \right) \quad (14.1.12)$$

这是在测量 \mathbf{X} 下，参数 π, \mathbf{a}, σ 的似然函数。因此，我们应该选择使得这个概率达到最大的参数值作为 π, \mathbf{a}, σ 的估计。这就导致一个最大似然估计问题。

14.2 Bayes 估计

14.2.1 贝叶斯模型

贝叶斯公式

在经典参数估计中，带有参数 $\theta \in \Theta$ 的密度族记为 $p(\mathbf{x}; \theta)$ ，它表示参数空间 Θ 中不同参数值对应于不同的密度。但是，贝叶斯学派则将参数 θ 视为随机变量，认为 $p(\mathbf{x}; \theta)$ 是随机变量取某个值时的条件分布，因此，记为条件密度的形式 $p(\mathbf{x} | \theta)$ 。随机变量 θ 的先验信息确定了先验密度，记为 $\pi(\theta)$ 。按照贝叶斯的观点，以下述方式产生样本 $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k)$ ：

- (1) 设想先验密度 $\pi(\theta)$ 产生一个观测值 θ ；
- (2) 从条件密度 $p(\mathbf{x} | \theta)$ 产生观测样本值： $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k)$ 。

此时，样本 \mathbf{X} 的联合条件密度为

$$p(\mathbf{X} | \theta) = \prod_{j=1}^n p(\mathbf{x}_j | \theta) \quad (14.2.1)$$

它包含了样本信息。不难看出，在经典参数估计中它是似然函数。使用先验密度 $\pi(\theta)$ ，可以将先验信息与样本信息综合起来。显然，样本与参数的联合密度为：

$$\pi(\mathbf{X}, \theta) = p(\mathbf{X} | \theta)\pi(\theta)$$

由于我们的目的是对参数 θ 做出估计，为此，先求出 $\pi(\mathbf{X}, \theta)$ 关于 \mathbf{X} 的边际密度：

$$m(\mathbf{X}) = \int_{\Theta} p(\mathbf{X} | \theta)\pi(\theta)d\theta \quad (14.2.2)$$

$m(\mathbf{X})$ 与 θ 无关，即它不包含参数 θ 的任何信息。因此，对估计起作用的仅是条件密度：

$$\pi(\theta | \mathbf{X}) = \frac{\pi(\mathbf{X}, \theta)}{m(\mathbf{X})} = \frac{p(\mathbf{X} | \theta)\pi(\theta)}{\int_{\Theta} p(\mathbf{X} | \theta)\pi(\theta)d\theta} \quad (14.2.3)$$

这是贝叶斯公式的密度形式。

给定样本 \mathbf{X} ， θ 的条件密度 $\pi(\theta | \mathbf{X})$ 称为 θ 的后验密度，它集中了样本与先验中有关参数 θ 的所有信息。后验密度 $\pi(\theta | \mathbf{X})$ 是通过观测后所得到的，它比先验密度 $\pi(\theta)$ 更接近于真实情况。因此，利用后验密度对参数可以作出更好的估计。

注 式(14.2.3)所给出的是 X 和 θ 都是连续型变量的贝叶斯公式。在其它情况下，贝叶斯公式如下：

- (1) X 为离散型变量：
$$\pi(\theta | X = X_i) = \frac{p(X = X_i | \theta)\pi(\theta)}{\int_{\Theta} p(X = X_i | \theta)\pi(\theta)d\theta};$$
- (2) θ 为离散型变量：
$$\pi(\theta_i | X) = \frac{p(X | \theta_i)\pi(\theta_i)}{\sum_i p(X | \theta_i)\pi(\theta_i)};$$
- (3) X, θ 均为离散型变量：
$$\pi(\theta_i | X = X_j) = \frac{p(X = X_j | \theta_i)\pi(\theta_i)}{\sum_i p(X = X_j | \theta_i)\pi(\theta_i)}.$$

若随机变量 x 的概率密度为 $f(x) = cg(x)$ ，其中 c 是与 x 无关的常数，则将它记为 $f(x) \propto g(x)$ ，并称 $g(x)$ 为 $f(x)$ 的核。如正态分布 $N(\mu, \sigma^2)$ 的核是

$$\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

在贝叶斯公式(14.2.3)中，边际密度 $m(X)$ 与参数 θ 无关，所以后验密度 $\pi(\theta | X)$ 可记为

$$\pi(\theta | X) \propto p(X | \theta)\pi(\theta) \quad (14.2.4)$$

后验密度 $\pi(\theta | X)$ 的这种表示主要是便于运算。

贝叶斯模型

- (1) 参数空间 Θ 的先验密度 $\pi(\theta)$ ；
- (2) 样本 X 的联合条件密度 $p(X | \theta)$ ，通常称为样本密度；
- (3) 先验密度 $\pi(\theta)$ 与样本密度 $p(X | \theta)$ 构成后验密度 $\pi(\theta | X)$ ；
- (4) 贝叶斯估计准则：从后验密度 $\pi(\theta | X)$ 对参数 θ 进行估计。

贝叶斯估计与经典参数估计的区别在于：经典参数估计是从密度函数族 $p(x; \theta)$ 出发，对参数进行估计，而贝叶斯估计是从后验密度 $\pi(\theta | X)$ 对参数 θ 进行估计。

下面给出贝叶斯估计的一个例子。

例 14.2.1 设 $X = (x_1, x_2, \dots, x_k)$ 是来自正体总体 $N(\theta, 1)$ 的样本， θ 的先验密度为正态分布 $N(\mu, \sigma^2)$ (当然， μ, σ^2 都是已知的)，求 θ 使得后验密度达到最大的估计。

因为

$$\pi(\theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(\theta-\mu)^2}{2\sigma^2}\right)$$

$$p(x|\theta) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-\theta)^2}{2}\right)$$

所以, 由贝叶斯公式, θ 的后验密度为

$$\pi(\theta|X) \propto \exp\left(-\frac{(\theta-\mu)^2}{2\sigma^2} - \frac{1}{2} \sum_{j=1}^n (x_j - \theta)^2\right)$$

通过简单的计算, 我们有

$$\pi(\theta|X) \propto \exp\left(-\frac{(\theta-t)^2}{2\eta^2}\right)$$

其中

$$\begin{cases} t = \frac{n\bar{x} + \mu/\sigma^2}{n + 1/\sigma^2} \\ \eta^2 = \frac{1}{1 + 1/\sigma^2} \end{cases}$$

因此, $\pi(\theta|X) \sim N(t, \eta^2)$ 。故 θ 使得后验密度达到最大的估计为

$$\hat{\theta} = t = \frac{n}{n + 1/\sigma^2} \bar{x} + \frac{1/\sigma^2}{n + 1/\sigma^2} \mu \quad (14.2.5)$$

下面, 我们解释贝叶斯估计 $\hat{\theta}$ 的意义。如果只有样本信息而毫无先验信息, 则样本均值 \bar{x} 是 θ 的估计; 另一方面, 如果只有先验信息 $N(\mu, \sigma^2)$, 则我们只能用先验密度的均值 μ 作为 θ 的估计。从式(14.2.5)可以看出, 当这两种情况都发生时, θ 的贝叶斯估计是这两种估计 \bar{x} 和 μ 的加权平均, 权的比值为 $n:1/\sigma^2$ 。这个比值是非常合理的, 对样本而言, 容量愈大, 即样本信息愈多, \bar{x} 的权值就应该更大; 对 μ 而言, 其重要性在于 σ^2 的大小, σ^2 愈大, 表示先验信息愈不肯定(因为 θ 在 μ 周围散布愈大), 反之, σ^2 很小, 就可以肯定 θ 在 μ 附近, 因此 μ 的权值应该与 σ^2 成反比。公式(14.2.5)正好体现了上述分析。

14.2.2 无信息先验密度

在无先验信息时, 如何确定先验密度? 如何进行贝叶斯分析? 贝叶斯学派对此进行了大量研究, 得到一批大量成果。这些成果统称为无信息先验密度。

贝叶斯假设

在对参数 θ 无先验信息的情况下, 人们对参数 θ 的一切可能值将“一视同仁”, 不可能对哪些值

有所偏受，即参数 θ 取参数空间内的一切值都是同等可能的。这样，就导致了贝叶斯选取先验密度的“同等无知”原则和假设：

无信息先验密度 $\pi(\theta)$ 是参数空间 Θ 上的“均匀密度”，即

$$\pi(\theta) = C, \text{ 或 } \pi(\theta) \propto 1, \theta \in \Theta \quad (14.2.6)$$

当 Θ 是无界区域时， $\pi(\theta)$ 可能不是一个正常的概率密度，而是一种广义先验密度。此时，称 $\pi(\theta)$ 为广义均匀密度。

广义先验密度 若 $\pi(\theta)$ 满足下述两个条件：

$$\int_{\Theta} \pi(\theta) d\theta = +\infty \quad (14.2.7)$$

$$\int_{\Theta} p(x|\theta)\pi(\theta)d\theta < +\infty \quad (14.2.8)$$

则称 $\pi(\theta)$ 为广义先验密度。

虽然，由式(14.2.7)， $\pi(\theta)$ 不是正常的概率密度，但根据式(14.2.8)，由式(14.2.3)所确定的后验密度 $\pi(\theta|x)$ 存在，而这个后验密度仍能为贝叶斯分析提供依据。

当 $\pi(\theta)$ 为均匀密度或广义均匀密度时，必有

$$\pi(\theta|x) \propto p(x|\theta) = L(x;\theta)$$

即似然函数是后验密度的核。故采用均匀密度或广义均匀密度作为先验密度时，最大后验估计与最大似然估计是一致的。

不变先验密度

如果一个概率密度在某个变换群下具有不变性，则称该密度是这个变换群的不变密度，简称不变密度。根据变换不变性来选择的先验密度称为不变先验密度。下面介绍三种不变密度，它们分别是位置参数、尺度参数以及位置—尺度参数的不变先验密度。

位置参数 下述形式的概率密度族称为位置参数族：

$$\{p(x-\theta) | \theta \in (-\infty, +\infty)\} \quad (14.2.9)$$

其中 θ 称为位置参数。由于位置的不变先验密度 $\pi(\theta)$ 要求对位置变换是不变的，即 $\forall C \subset R, b \in R$ ，记 $C-b = \{c-b | c \in C\}$ ，必有

$$\int_C \pi(\theta) d\theta = \int_{C-b} \pi(\theta) d\theta = \int_C \pi(\theta-b) d\theta$$

由 C 与 b 的任意性，得到

$$\pi(\theta) = \pi(\theta-b)$$

所以 $\pi(\theta)$ 是常数, 从而 $\pi(\theta) \propto 1$ 。这表明, 位置的不变先验密度服从贝叶斯假设。

尺度参数 下述形式的概率密度族称为尺度参数族:

$$\left\{ \frac{1}{\sigma} p\left(\frac{x}{\sigma}\right) \mid \sigma \in (0, +\infty) \right\} \quad (14.2.10)$$

其中 σ 称为尺度参数。由于尺度的不变先验密度 $\pi(\sigma)$ 要求对尺度变换是不变的, 即 $\forall C \subset R, b \in R^+$, 有

$$\int_C \pi(\sigma) d\sigma = \int_{bC} \pi(\sigma) d\sigma = \int_C \pi(b^{-1}\sigma) d(b^{-1}\sigma) = \int_C b^{-1} \pi(b^{-1}\sigma) d\sigma$$

由 C 任意性, 我们有

$$\pi(\sigma) = b^{-1} \pi(b^{-1}\sigma)$$

上式对一切 $\sigma > 0, b > 0$ 成立, 所以, 令 $b = \sigma$, 得到

$$\pi(\sigma) = \sigma^{-1} \pi(1) \propto \sigma^{-1} \quad (14.2.11)$$

这就是尺度的不变先验密度。

位置-尺度参数 下述形式的概率密度族称为位置-尺度参数族:

$$\left\{ \frac{1}{\sigma} p\left(\frac{x-\theta}{\sigma}\right) \mid \theta \in (-\infty, +\infty), \sigma \in (0, +\infty) \right\} \quad (14.2.12)$$

其中 θ 是位置参数, σ 是尺度参数。用前面的类似方法, 可证明位置-尺度的不变先验密度为

$$\pi(\theta, \sigma) \propto \sigma^{-1} \quad (14.2.13)$$

Jeffreys 先验密度

设 $\mathbf{x} = (x_1, x_2, \dots, x_n)$ 是来自密度函数 $p(x|\boldsymbol{\theta})$ 的一个样本, 其中 $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_k)$ 是 k 维参数向量。当 $\boldsymbol{\theta}$ 无先验信息可用时, Jeffreys 根据变换群和 (Harr) 测度理论, 用 Fisher 信息矩阵行列式的平方根来表示先验密度。这种无信息先验密度称为 Jeffreys 先验密度。Jeffreys 先验密度可用下述步骤导出:

- (1) 计算样本的对数似然函数: $l(\mathbf{x}; \boldsymbol{\theta}) = \log L(\mathbf{x}; \boldsymbol{\theta}) = \sum_{j=1}^k \log p(x_j; \boldsymbol{\theta})$;
- (2) 计算参数向量 $\boldsymbol{\theta}$ 的 Fisher 信息矩阵: $I(\boldsymbol{\theta}) = E_{\mathbf{x}|\boldsymbol{\theta}}[-\partial^2 l / \partial \theta_i \partial \theta_j]$;
- (3) $\boldsymbol{\theta}$ 的无信息先验密度为 $\pi(\boldsymbol{\theta}) \propto (\det(I(\boldsymbol{\theta})))^{1/2}$ 。

例 14.2.2 设 $\mathbf{x} = (x_1, x_2, \dots, x_n)$ 是来自正态总体 $N(\mu, \sigma^2)$ 的样本, 求参数向量 $\boldsymbol{\theta} = (\mu, \sigma)$ 的 Jeffreys 先验密度。

样本的对数似然函数为

$$l(\mathbf{x}; \mu, \sigma^2) = -\frac{1}{2} \log(2\pi) - n \log \sigma - \sum_{j=1}^k (x_j - \mu)^2 / 2\sigma^2$$

$\theta = (\mu, \sigma)$ 的 Fisher 信息矩阵为

$$I(\theta) = \begin{pmatrix} E(-\partial^2 / \partial \mu^2) & E(-\partial^2 l / \partial \mu \partial \sigma) \\ E(-\partial^2 l / \partial \mu \partial \sigma) & E(-\partial^2 l / \partial \sigma^2) \end{pmatrix} = \begin{pmatrix} n / \sigma^2 & 0 \\ 0 & 2n / \sigma^2 \end{pmatrix}$$

$$\det I(\theta) = 4n^2 / \sigma^2 \propto \sigma^{-4}$$

所以, $\theta = (\mu, \sigma)$ 的 Jeffreys 先验密度为 $\pi(\mu, \sigma) \propto \sigma^{-2}$ 。

下面给出例 14.2.2 的几个特殊情况:

- (1) 当 μ 为已知时, $I(\sigma) = E(-\partial^2 l / \partial \sigma^2) = 2n / \sigma^2$, 所以 $\pi(\sigma) \propto \sigma^{-1}, \sigma \in R^+$;
- (2) 当 σ 为已知时, $I(\mu) = E(-\partial^2 l / \partial \mu^2) = n / \sigma^2$, 所以 $\pi(\mu) \propto 1, \mu \in R$;
- (3) 当 μ 与 σ 独立时, $\pi(\mu, \sigma) \propto \sigma^{-1}, \mu \in R, \sigma \in R^+$ 。

可见, Jeffreys 先验密度表明: μ 与 σ 的无信息先验密度是不独立的。在联合密度的两种形式 σ^{-2} 与 σ^{-1} 中, Jeffreys 推荐使用 $\pi(\sigma) \propto \sigma^{-1}, \sigma \in R^+$ 。

注: 在一般情况下, 无信息先验不是唯一的, 但它们对贝叶斯统计推断 (参数估计、假设检验等) 的结果影响都很小, 很少对结果产生重大影响。所以任何无信息先验密度都是可以使用的。当今无论是在数理统计研究还实际应用中, 采用无信息先验密度愈来愈多, 即使是经典统计学者也认为无信息先验是“客观”的, 是可以接受的。

14.2.3 共轭先验密度

从例 14.2.1, 可以看出: 对于一元正态密度族 (参数是均值), 如果先验密度取正态密度, 则后验密度也是正态密度, 即先验密度与后验密度属于同类型的密度族。这个例子可以推广到多元正态分布的情形。

例 14.2.3 设 $\mathbf{x} = (x_1, x_2, \dots, x_k) \sim N_k(\boldsymbol{\mu}, \Sigma)$, 其中 $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_k)^T$ 未知, Σ 已知。令 $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k)$ 是来自 $N_k(\boldsymbol{\mu}, \Sigma)$ 的样本, 则样本 \mathbf{X} 的密度函数为

$$p(\mathbf{X} | \boldsymbol{\mu}) \propto \exp\left(-\frac{1}{2} \sum \| \mathbf{x}_j - \boldsymbol{\mu} \|_{\Sigma^{-1}}^2\right)$$

设 $\boldsymbol{\mu}$ 的先验密度为

$$\pi(\boldsymbol{\mu}) \propto \exp\left(-\frac{1}{2} \| \boldsymbol{\mu} - \boldsymbol{\mu}_0 \|_{\Sigma_0^{-1}}^2\right)$$

$\boldsymbol{\mu}_0, \Sigma_0$ 已知，则 $\boldsymbol{\mu}$ 的后验密度为

$$\pi(\boldsymbol{\mu} | \mathbf{X}) \propto \exp\left(-\frac{1}{2}\left(\|\boldsymbol{\mu} - \boldsymbol{\mu}_0\|_{\Sigma_0^{-1}}^2 + \sum_{j=1}^k \|\mathbf{x}_j - \boldsymbol{\mu}\|_{\Sigma^{-1}}^2\right)\right)$$

记 $\bar{\mathbf{x}} = \frac{1}{k} \sum_{j=1}^k \mathbf{x}_j$ ，则可以计算出

$$\pi(\boldsymbol{\mu} | \mathbf{X}) \propto \exp\left(-\frac{1}{2}\|\boldsymbol{\mu} - \boldsymbol{\mu}_k\|_{\Sigma_k^{-1}}^2\right)$$

其中

$$\boldsymbol{\mu}_k = (\Sigma_0^{-1} + k\Sigma^{-1})^{-1}(\Sigma_0^{-1}\boldsymbol{\mu}_0 + k\Sigma^{-1}\bar{\mathbf{x}})$$

$$\Sigma_k = (\Sigma_0^{-1} + k\Sigma^{-1})^{-1}$$

故后验密度 $\pi(\boldsymbol{\mu} | \mathbf{X})$ 与先验密度 $\pi(\boldsymbol{\mu})$ 同属于正态密度族。

后验密度 $\pi(\boldsymbol{\mu} | \mathbf{X})$ 与先验密度 $\pi(\boldsymbol{\mu})$ 属于同类型的密度族不是偶然的，事实上对于很多分布族都有这种性质。为了描述这样的事实，引进下述定义。

共轭先验密度 设 $P = \{p(x | \theta) | \theta \in \Theta\}$ 是以 θ 为参数（向量）的密度函数族， $\Pi = \{\pi(\theta)\}$ 是 θ 的先验密度族。如果对任意 $p \in P$ 和 $\pi \in \Pi$ ，所得到的后验密度 $\pi(\theta | \mathbf{X}) \in \Pi$ ，则称 Π 为 P 的共轭密度族，或者称 $\pi(\theta)$ 为参数（向量） θ 的共轭先验密度。

下表列出了若干常用的共轭先验分布。

常用的共轭先验分布

总体密度	共轭先验密度	后验密度
二项分布 $b(n, \theta)$	β 分布 $\beta(a, b)$	$\beta(a + x, n + b - x)$
泊松分布 $P(\lambda)$	Γ 分布 $\Gamma(\alpha, \mu)$	$\Gamma(\alpha + \bar{x}, n + \mu)$
指数分布 $E(\theta^{-1})$	逆 Γ 分布 $IG(\alpha, \mu)$	$IG(\alpha n, \mu + \bar{x})$
均匀分布 $U(0, \theta)$	Parato 分布 $P_a(\alpha, \theta_0)$	$P_a(\alpha + n, \theta_1), \theta_1 = \max\{x_1, x_2, \dots, x_n, \theta_0\}$
正态分布 $N(\mu, \sigma^2)$	$N(\mu_0, \tau^2)$	$N(\hat{\mu}, \gamma^2),$

σ^2 已知

$$\hat{\mu} = \left(\frac{n}{\sigma^2} + \frac{1}{\tau^2}\right)^{-1} \left(\frac{n\bar{x}}{\sigma^2} + \frac{\mu_0}{\tau^2}\right)$$

		$\gamma^2 = \left(\frac{n}{\sigma^2} + \frac{1}{\tau^2} \right)^{-1}$
正态分布 $N(0, \sigma^2)$ 未知参数 $\theta = \sigma^2$	逆 Γ 分布 $IG(\alpha, \mu)$	$IG\left(\alpha + \frac{n}{2}, \mu + \frac{\bar{x}}{2}\right)$
多元正态分布 $N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma}$ 已知	$N_k(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$	$N_k(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ $\boldsymbol{\mu}_k = (\boldsymbol{\Sigma}_0^{-1} + k\boldsymbol{\Sigma}^{-1})^{-1}(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_0 + k\boldsymbol{\Sigma}^{-1}\bar{\mathbf{x}})$ $\boldsymbol{\Sigma}_k = (\boldsymbol{\Sigma}_0^{-1} + k\boldsymbol{\Sigma}^{-1})^{-1}$

注： n 表示样本容量， $\bar{\mathbf{x}}$ 表示样本均值。

共轭先验分密度是对某一分布中的参数而言的。离开指定的参数及其所在的分布，讨论共轭先验分布是没有意义的。引进共轭先验密度的目的是为了便于后验密度的计算，因为后验分布 $\pi(\boldsymbol{\theta} | \mathbf{x})$ 可直接从它的分布核得到，无需计算边缘分布 $m(\mathbf{x})$ 。在其它情况下， $m(\mathbf{x})$ 与 $\pi(\boldsymbol{\theta} | \mathbf{x})$ 的计算都会遇到困难，如 $\mathbf{x} \sim N(\boldsymbol{\mu}, \sigma^2)$ 并且先验分布取柯希分布时，则后验分布只能通过数值计算得到。但值得指出的是，选用共轭先验分布时必须注意先验的合理性，因为对于贝叶斯估计而言，毕竟先验的合理性比计算方便更重要。

14.2.4 贝叶斯估计

记 $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k)$ 为来自某个密度族 $\{p(\mathbf{x} | \boldsymbol{\theta}) | \boldsymbol{\theta} \in \Theta\}$ 的样本， $\pi(\boldsymbol{\theta})$ 是给定的先验密度， $\pi(\boldsymbol{\theta} | \mathbf{x})$ 是相应的后验密度。我们的目的是从后验密度 $\pi(\boldsymbol{\theta} | \mathbf{x})$ 对参数 $\boldsymbol{\theta}$ 进行贝叶斯估计。贝叶斯估计使后验风险达到最小，而后验风险依赖于损失函数的选择，即不同类型的损失函数导致不同的后验风险，这样就导致了参数 $\boldsymbol{\theta}$ 的各种贝叶斯估计。

后验风险

令 $\boldsymbol{\delta} = \boldsymbol{\delta}(\mathbf{x})$ 是 $\boldsymbol{\theta}$ 的任一个估计， $\boldsymbol{\delta}$ 的损失函数 $L(\boldsymbol{\theta}, \boldsymbol{\delta})$ 关于后验密度 $\pi(\boldsymbol{\theta} | \mathbf{x})$ 的数学期望称为后验风险，并记为

$$R_\pi(\boldsymbol{\delta} | \mathbf{x}) = E_{\boldsymbol{\theta} | \mathbf{x}}[L(\boldsymbol{\theta}, \boldsymbol{\delta}(\mathbf{x}))]$$

$$= \begin{cases} \int_{\Theta} L(\boldsymbol{\theta}, \boldsymbol{\delta}(\mathbf{x}))\pi(\boldsymbol{\theta} | \mathbf{x})d\boldsymbol{\theta}, & \boldsymbol{\theta} \text{ 为连续型变量} \\ \sum_j L(\boldsymbol{\theta}_j, \boldsymbol{\delta}(\mathbf{x}))\pi(\boldsymbol{\theta}_j | \mathbf{x}), & \boldsymbol{\theta} \text{ 为离散型变量} \end{cases} \quad (14.2.14)$$

如果存在 $\boldsymbol{\theta}$ 的一个估计 $\boldsymbol{\delta}^* = \boldsymbol{\delta}^*(\mathbf{x})$ 使得

$$R_\pi(\boldsymbol{\delta}^* | \mathbf{x}) = \min_{\boldsymbol{\delta}} R_\pi(\boldsymbol{\delta} | \mathbf{x}) \quad (14.2.15)$$

则称 $\boldsymbol{\delta}^*(\mathbf{x})$ 为 $\boldsymbol{\theta}$ 的贝叶斯（后验）估计。

显然，对于给定的先验密度，贝叶斯估计与所选择的损失函数有关。在贝叶斯估计中，常用的损失函数有以下几种类型：

1. 平方损失函数：

$$L(\theta, \delta) = (\delta - \theta)^2 \quad (14.2.16)$$

或加权平方损失函数：

$$L(\theta, \delta) = w(\theta)(\delta - \theta)^2 \quad (14.2.17)$$

2. 绝对损失函数：

$$L(\theta, \delta) = |\delta - \theta| \quad (14.2.18)$$

3. 0-1 损失函数

$$L(\theta, \delta) = \begin{cases} 0, & |\delta - \theta| \leq \varepsilon \\ 1, & |\delta - \theta| > \varepsilon \end{cases} \quad (14.2.19)$$

其中 ε 是小正数。如果 θ 是多维参数向量，上式中的绝对值改为向量的范数。

4. 多元二次损失函数：对于多维参数向量 θ ，二次损失函数为

$$L(\theta, \delta) = (\delta - \theta)^T G (\delta - \theta) \quad (14.2.20)$$

其中 G 是正定矩阵。如果 $G = \text{diag}(g_1, g_2, \dots, g_n)$ ，则二次损失函数变为下述简单形式：

$$L(\theta, \delta) = \sum_{j=1}^n g_j (\delta_j - \theta_j)^2 \quad (14.2.21)$$

贝叶斯估计

下面，我们给出各种损失函数下的贝叶斯估计。

定理 14.2.1 在平方损失下， θ 的贝叶斯估计是后验分布 $\pi(\theta | \mathbf{x})$ 的均值，即

$$\delta^\pi(\mathbf{x}) = E[\theta | \mathbf{x}] \quad (14.2.22)$$

证明：在平方损失下， θ 的任一个估计 $\delta = \delta(\mathbf{x})$ 的后验风险为

$$E[(\delta - \theta)^2 | \mathbf{x}] = \delta^2 - 2\delta E[\theta | \mathbf{x}] + E[\theta^2 | \mathbf{x}]$$

因此，仅当 $\delta = \delta^\pi(\mathbf{x})$ 时上式达到最小。证毕。

定理 14.2.2 在加权平方损失下， θ 的贝叶斯估计为

$$\delta^\pi(\mathbf{x}) = \frac{E[w(\theta)\theta | \mathbf{x}]}{E[w(\theta) | \mathbf{x}]} \quad (14.2.23)$$

证明：与定理 14.2.1 类似。

定理 14.2.3 在绝对损失下， θ 的贝叶斯估计是后验密度 $\pi(\theta | \mathbf{x})$ 的中位数。

$$\delta^\pi(\mathbf{x}) = \text{Med}[\theta | \mathbf{x}] \quad (14.2.24)$$

证明：设 $\delta = \delta(\mathbf{x})$ 为 θ 另一估计。不妨假定 $\delta > \delta^\pi$ ，由绝对损失函数定义，我们有

$$L(\theta, \delta^\pi) - L(\theta, \delta) = \begin{cases} \delta^\pi - \delta, & \theta \leq \delta^\pi \\ 2\theta - (\delta^\pi + \delta), & \delta^\pi < \theta < \delta \\ \delta - \delta^\pi, & \theta \geq \delta \end{cases}$$

当 $\delta^\pi < \theta < \delta$ 时，必有 $2\theta - (\delta^\pi + \delta) \leq 2\delta - (\delta^\pi + \delta) = \delta - \delta^\pi$ ，所以，

$$L(\theta, \delta^\pi) - L(\theta, \delta) = \begin{cases} \delta^\pi - \delta, & \theta \leq \delta^\pi \\ \delta - \delta^\pi, & \theta > \delta^\pi \end{cases}$$

根据中位数的定义，有 $P(\theta \leq \delta^\pi | \mathbf{x}) \geq 1/2$ 且 $P(\theta > \delta^\pi | \mathbf{x}) \leq 1/2$ 。因此，我们有

$$\begin{aligned} R_\pi(\delta^\pi | \mathbf{x}) - R_\pi(\delta | \mathbf{x}) &= E_{\theta|\mathbf{x}}[L(\theta, \delta^\pi) - L(\theta, \delta)] \\ &\leq (\delta^\pi - \delta)P(\theta \leq \delta^\pi | \mathbf{x}) + (\delta - \delta^\pi)P(\theta > \delta^\pi | \mathbf{x}) \\ &\leq (\delta^\pi - \delta)/2 + (\delta - \delta^\pi)/2 = 0 \end{aligned}$$

于是， $R_\pi(\delta^\pi | \mathbf{x}) \leq R_\pi(\delta | \mathbf{x})$ 。同理可证：当 $\delta < \delta^\pi$ 时，也有 $R_\pi(\delta^\pi | \mathbf{x}) \leq R_\pi(\delta | \mathbf{x})$ 。故后验分布中位数使得后验风险达到最小。证毕。

定理 14.2.4 对于参数向量 $\theta = (\theta_1, \theta_2, \dots, \theta_n)^T$ ，在二次损失函数下， θ 的贝叶斯估计是后验密度 $\pi(\theta | \mathbf{x})$ 的均值向量，即

$$\delta^\pi(\mathbf{x}) = E[\theta | \mathbf{x}] = (E[\theta_1 | \mathbf{x}], E[\theta_2 | \mathbf{x}], \dots, E[\theta_n | \mathbf{x}])^T \quad (14.2.25)$$

证明：在二次损失函数 $L(\theta, \delta) = (\delta - \theta)^T G(\delta - \theta)$ 下， θ 的任一个估计 $\delta = \delta(\mathbf{x})$ 的后验风险为

$$\begin{aligned} E[L(\theta, \delta) | \mathbf{x}] &= E[(\delta - \theta)^T G(\delta - \theta) | \mathbf{x}] \\ &= E[((\delta - \delta^\pi) + (\delta^\pi - \theta))^T G((\delta - \delta^\pi) + (\delta^\pi - \theta)) | \mathbf{x}] \\ &= (\delta - \delta^\pi)^T G(\delta - \delta^\pi) + E[(\delta^\pi - \theta)^T G(\delta^\pi - \theta) | \mathbf{x}] \end{aligned}$$

在上述最后一个等式中，我们利用了 $E[(\delta^\pi - \theta) | \mathbf{x}] = 0$ 的性质。由于 $E[(\delta^\pi - \theta)^T G(\delta^\pi - \theta) | \mathbf{x}]$ 是常数，所以仅当 $\delta = \delta^\pi$ 使后验风险达到最小。证毕。

定理 14.2.5 在 0-1 损失下， θ 的贝叶斯估计是后验密度 $\pi(\theta | \mathbf{x})$ 的最大值点（通常称为最大后验，众数，或广义最大似然），即

$$\delta^\pi(\mathbf{x}) = \arg \max_{\theta \in \Theta} \pi(\theta | \mathbf{x}) \quad (14.2.26)$$

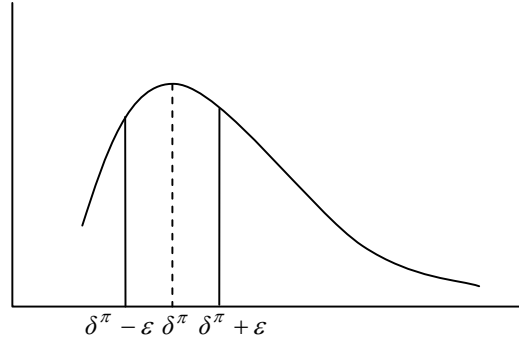


图 14.2.1 0-1 损失的贝叶斯估计

证明：由于 θ 的任一个估计 $\delta = \delta(\mathbf{x})$ 的后验风险为

$$\begin{aligned} E[L(\theta, \delta(\mathbf{x})) | \mathbf{x}] &= \int_{\Theta} L(\theta, \delta(\mathbf{x}))\pi(\theta | \mathbf{x})d\theta \\ &= \int_{\Theta} \pi(\theta | \mathbf{x})d\theta - \int_{\|\delta - \theta\| \leq \varepsilon} \pi(\theta | \mathbf{x})d\theta \\ &= 1 - \int_{\|\delta - \theta\| \leq \varepsilon} \pi(\theta | \mathbf{x})d\theta \end{aligned}$$

因此，后验风险最小当且仅当 $\int_{\|\delta - \theta\| \leq \varepsilon} \pi(\theta | \mathbf{x})d\theta$ 达到最大。当 ε 较小时， δ 取后验密度的最大值点积分 $\int_{\|\delta - \theta\| \leq \varepsilon} \pi(\theta | \mathbf{x})d\theta$ 达到最大值（如图 14.2.1 所示）。故 $\delta = \delta^\pi$ 使后验风险达到最小。

定理 14.2.1—定理 14.2.5 给出了三种贝叶斯估计：后验均值估计、后验中位数估计和最大后验估计（又称为众数估计，或广义最大似然估计），并指出这些估计在各自损失下的风险达到最小。在实际应用中，最广泛使用的是后验均值估计和最大后验估计。下表列出了常用分布的后验均值估计和最大后验估计。

常用分布的后验均值估计和最大后验估计

总体分布	先验分布	后验均值估计	最大后验估计
二项分布 $b(n, \theta)$	$\beta(a, b)$	$\frac{a+x}{a+b+n}$	$\frac{a+x-1}{a+b+n-2}$
参数 θ	1	$\frac{x}{n}$	$\frac{x+1}{n+2}$
泊松分布 $P(\lambda)$	$\Gamma(\alpha, \mu)$	$\frac{n\bar{x} + \alpha}{n + \mu}$	$\frac{n\bar{x} + \alpha - 1}{n + \mu}$
参数 λ	1	$\bar{x} + 1/n$	\bar{x}
指数分布 $E(\theta^{-1})$ 参	$I\Gamma(\alpha, \mu)$	$\frac{n\bar{x} + \lambda}{\alpha + n + 1}$	$\frac{n\bar{x} + \lambda}{\alpha + n + 1}$
数 θ			

均匀分布 $U(0, \theta)$ 参数 θ	$P_a(\alpha, \theta_0)$	$\frac{\alpha + n}{\alpha + n - 1} \max\{x_1, x_2, \dots, x_n, \theta_0\}$	$\max\{x_1, x_2, \dots, x_n, \theta_0\}$
正态分布 $N(\mu, \sigma^2)$ σ^2 已知	$N(\mu_0, \tau^2)$	$\left(\frac{n}{\sigma^2} + \frac{1}{\tau^2}\right)^{-1} \left(\frac{n\bar{x}}{\sigma^2} + \frac{\mu_0}{\tau^2}\right)$	$\left(\frac{n}{\sigma^2} + \frac{1}{\tau^2}\right)^{-1} \left(\frac{n\bar{x}}{\sigma^2} + \frac{\mu_0}{\tau^2}\right)$
正态分布 $N(0, \sigma^2)$ 未知参数 $\theta = \sigma^2$	$I\Gamma(\alpha, \mu)$	$\frac{2\mu + \sum_{j=1}^n x_j^2}{2\alpha + n - 2}$	$\frac{2\mu + \sum_{j=1}^n x_j^2}{2\alpha + n + 2}$
	$1/\theta$	$\frac{1}{n-2} \sum_{j=1}^n x_j^2$	$\frac{1}{n+2} \sum_{j=1}^n x_j^2$
多元正态分布 $N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma}$ 已知	$N_k(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$	$(\boldsymbol{\Sigma}_0^{-1} + k\boldsymbol{\Sigma}^{-1})^{-1}$ $\times (\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_0 + k\boldsymbol{\Sigma}^{-1}\bar{\mathbf{x}})$	$(\boldsymbol{\Sigma}_0^{-1} + k\boldsymbol{\Sigma}^{-1})^{-1}$ $\times (\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_0 + k\boldsymbol{\Sigma}^{-1}\bar{\mathbf{x}})$

14.3 期望/最大化 (EM) 算法

EM(Expectation/Maximization)算法是一种迭代方法，由 Dempster 于 1977 年首先提出，最初主要用于最大后验估计。当然，EM 算法也可以用于最大似然估计，因为最大后验估计与最大似然估计，在计算上完全是相同的。EM 算法不是直接对复杂的后验分布（或似然函数）求极大值，而是在观测数据的基础上添加“潜在数据”简化计算，通过一系列简单的极大化实现最大后验估计（最大似然估计）。本节将具体介绍这种算法。

14.3.1 EM 算法

潜在数据

在实际问题中，求解最大似然估计或最大后验估计的困难主要是观测数据不能提供完全信息，或者说存在一些“潜在数据”不能被观测到。这些潜在数据是“缺损数据 (Missing Data)”也可以是一些未知参数。当然，给“潜在数据”下一个准确的数学定义是非常困难的，下面通过两个具体例子来说明。

例 14.3.1 在 14.1.3 节所介绍的混合模型 $p(\mathbf{x}; \boldsymbol{\pi}, \boldsymbol{\theta}) = \sum_{k=1}^K \pi_k p(\mathbf{x}; \boldsymbol{\theta}_k)$ 中，如果已知每个成份 X_k 被选择的概率 π_k ，则对于观测数据 $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M)$ ，求解极大对数似然：

$$\boldsymbol{\theta}^* = \arg \max_{\boldsymbol{\theta}} \sum_{j=1}^M \left(\log \sum_{k=1}^K \pi_k p(\mathbf{x}_j; \boldsymbol{\theta}_k) \right)$$

就要容易得多。如果 $p(\mathbf{x}; \boldsymbol{\theta}_k)$ 是高斯密度，则非常容易求解极大对数似然。因此，可以将 $\{\pi_k\}$ 看作是混合模型的潜在数据。

我们也可以从另一个角度来考虑混合模型，对每一个观测数据 \mathbf{x}_j 配置一个向量 K 维向量 \mathbf{y}_j 来表征该观测数据来自哪一个成份，即定义：

$$y_{jk} = \begin{cases} 1, & \mathbf{x}_j \text{ 来自第 } k \text{ 个成份} \\ 0, & \text{否则} \end{cases}$$

通常称 \mathbf{y}_j 为索引变量，并称 $\mathbf{z}_j = (\mathbf{x}_j, \mathbf{y}_j)$ 为完全数据。可以推出完全数据 \mathbf{Z} 的对数似然函数为

$$L(\mathbf{Z}; \boldsymbol{\theta}) = \sum_{j=1}^M \left(\log \sum_{k=1}^K z_{jk} p(\mathbf{x}_j; \boldsymbol{\theta}_k) \right)$$

同样，如果已知索引变量 \mathbf{Y} 的值，混合模型也是容易求解的。因此，我们也可以将索引变量 \mathbf{Y} 的值作为潜在数据。

下面给出一个将未知参数作为潜在数据的例子。

例 14.3.2 设观测数据 $(\mathbf{X}_i = (x_{i1}, x_{i2}, \dots, x_{in_i}) \mid u_i \mathbf{I}_{n_i}, \sigma \mathbf{I}_{n_i}) \sim N(u_i \mathbf{I}_{n_i}, \sigma \mathbf{I}_{n_i})$, $i = 1, 2, \dots, m$ ， $\{\mathbf{X}_i\}$ 是独立的，其中 \mathbf{I}_{n_i} 是每个元素均为 1 的 n_i 维向量，并且 $\mu_i \sim N(\beta, \tau^2)$ 。记

$$\boldsymbol{\theta} = (\beta, \sigma, \tau)^T, \mathbf{Y} = (\mu_1, \mu_2, \dots, \mu_m)^T$$

它们都是未知参数向量，但我们感兴趣的是参数 $\boldsymbol{\theta}$ 。假定参数 $\boldsymbol{\theta}$ 的先验分布为 $\pi(\boldsymbol{\theta}) \propto \tau$ ，现在的问题是从观测向量 $\mathbf{X} = \{x_{ij} \mid i = 1, 2, \dots, m; j = 1, 2, \dots, n_i\}$ 估计参数 $\boldsymbol{\theta}$ 。由贝叶斯公式，不难得到后验密度：

$$\begin{aligned} \pi(\boldsymbol{\theta}, \mathbf{Y} \mid \mathbf{X}) &\propto \pi(\beta, \sigma, \tau) \cdot \prod_{i=1}^m \pi(\mu_i \mid \beta, \tau) \cdot \prod_{i=1}^m \prod_{j=1}^{n_i} \pi(x_{ij} \mid \mu_i, \sigma) \\ &\propto \tau \cdot \frac{1}{\tau^m} \exp\left(-\frac{1}{2\tau^2} \sum_{i=1}^m (\mu_i - \beta)^2\right) \cdot \frac{1}{\sigma^n} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \mu_i)^2\right) \end{aligned}$$

其中， $n = \sum_{i=1}^m n_i$ 。将上式两边取对数，我们有

$$\log \pi(\boldsymbol{\theta}, \mathbf{Y} \mid \mathbf{X}) \propto -(m-1) \log \tau - \frac{1}{2\tau^2} \sum_{i=1}^m (\mu_i - \beta)^2 - n \log \sigma - \frac{1}{2\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \mu_i)^2$$

直接从上式求最大后验是非常困难的，其主要原因是不知道参数向量 $\mathbf{Y} = (\mu_1, \mu_2, \dots, \mu_m)^T$ 。

如果已知参数 $\mathbf{Y} = (\mu_1, \mu_2, \dots, \mu_m)^T$ ，则求解这个问题是相当容易的。因为，此时通过下述方程组：

$$\begin{cases} \frac{\partial \log \pi(\boldsymbol{\theta}, \mathbf{Y} | \mathbf{X})}{\partial \tau} \propto \frac{m-1}{\tau} - \frac{1}{\tau^3} \sum_{i=1}^m (\mu_i - \beta)^2 = 0 \\ \frac{\partial \log \pi(\boldsymbol{\theta}, \mathbf{Y} | \mathbf{X})}{\partial \sigma} \propto \frac{n}{\sigma} - \frac{1}{\sigma^3} \sum_{i=1}^m \sum_{j=1}^{n_i} (\mu_i - x_{ij})^2 = 0 \\ \frac{\partial \log \pi(\boldsymbol{\theta}, \mathbf{Y} | \mathbf{X})}{\partial \beta} \propto \frac{1}{\tau^2} \sum_{i=1}^m (\mu_i - \beta) = 0 \end{cases}$$

可以得到

$$\begin{cases} \beta = \frac{1}{m} \sum_{i=1}^m \mu_i = \bar{\mu} \\ \sigma^2 = \frac{1}{n} \sum_{i=1}^m \sum_{j=1}^{n_i} (\mu_i - x_{ij})^2 \\ \tau^2 = \frac{1}{m-1} \sum_{i=1}^m (\mu_i - \bar{\mu})^2 \end{cases}$$

因此, 在这个问题中, 未知参数 $\mathbf{Y} = (\mu_1, \mu_2, \dots, \mu_m)^T$ 是潜在数据。

EM 算法

令 \mathbf{X} 为观测数据, $\boldsymbol{\theta}$ 未知参数的后验密度 $\pi(\boldsymbol{\theta} | \mathbf{X})$, 称为观测后验密度。我们的目的是求观测后验密度 $\pi(\boldsymbol{\theta} | \mathbf{X})$ 的极大后验。记 $\pi(\boldsymbol{\theta} | \mathbf{X}, \mathbf{Y})$ 为添加潜在数据 \mathbf{Y} 后的后验分布, 并称为添加后验密度。 $\pi(\mathbf{Y} | \boldsymbol{\theta}, \mathbf{X})$ 表示给定 $\boldsymbol{\theta}$ 和观测数据 \mathbf{X} 下的潜在数据 \mathbf{Y} 的条件密度。

EM 算法是一种迭代算法, 记 $\boldsymbol{\theta}^{(j)}$ 为第 $j+1$ 次迭代开始时极大后验的估计值, 则第 $j+1$ 次迭代分为以下两步进行:

E 步: 求 $\pi(\boldsymbol{\theta} | \mathbf{X}, \mathbf{Y})$ 或 $\log(\pi(\boldsymbol{\theta} | \mathbf{X}, \mathbf{Y}))$ 关于潜在数据 \mathbf{Y} 的条件分布 $\pi(\mathbf{Y} | \boldsymbol{\theta}^{(j)}, \mathbf{X})$ 的期望:

$$\begin{aligned} Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(j)}, \mathbf{X}) &\equiv E_{\mathbf{Y}}[\log(\pi(\boldsymbol{\theta} | \mathbf{X}, \mathbf{Y})) | \boldsymbol{\theta}^{(j)}, \mathbf{X}] \\ &= \int \log(\pi(\boldsymbol{\theta} | \mathbf{X}, \mathbf{Y})) \pi(\mathbf{Y} | \boldsymbol{\theta}^{(j)}, \mathbf{X}) d\mathbf{Y} \end{aligned} \quad (14.3.1)$$

M 步: 求 $Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(j)}, \mathbf{X})$ 的极大点 $\boldsymbol{\theta}^{(j+1)}$, 即

$$\boldsymbol{\theta}^{(j+1)} = \arg \max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(j)}, \mathbf{X}) \quad (14.3.2)$$

这样就构成了 EM 算法的一次迭代。将上述 E 步与 M 步继续迭代下去, 直至收敛。

在实际应用中, EM 算法的迭代终止条件为

$$\|\boldsymbol{\theta}^{(j+1)} - \boldsymbol{\theta}^{(j)}\| < \varepsilon \quad (14.3.3)$$

或者

$$\|Q(\boldsymbol{\theta}^{(j+1)} | \boldsymbol{\theta}^{(j)}, \mathbf{X}) - Q(\boldsymbol{\theta}^{(j)} | \boldsymbol{\theta}^{(j-1)}, \mathbf{X})\| < \varepsilon \quad (14.3.4)$$

其中 ε 是控制精度。

例 14.3.2 (续) 在例 14.3.2 中, 潜在数据为 $\mathbf{Y} = (\mu_1, \mu_2, \dots, \mu_m)^T$, 下面给出它的 EM 算法。

E 步：假定 θ 的当前估计值为 $\theta^{(k)} = (\beta^{(k)}, \sigma^{(k)}, \tau^{(k)})$ 。E 步是在给定 $\mathbf{X}, \theta^{(k)}$ 下求式(14.3.2)的期望。由于正态分布是 μ_i 的共轭先验分布，所以，

$$(\mu_i | \theta^{(k)}, \mathbf{X}) \sim N(\hat{\mu}_i^{(k)}, \delta_i^{(k)})$$

其中：

$$\begin{cases} \hat{\mu}_i^{(k)} = \left(\frac{n_i}{(\sigma^{(k)})^2} + \frac{1}{(\tau^{(k)})^2} \right)^{-1} \left(\frac{n_i \bar{x}_i}{(\sigma^{(k)})^2} + \frac{\beta^{(k)}}{(\tau^{(k)})^2} \right) \\ \delta_i^{(k)} = \left(\frac{n_i}{(\sigma^{(k)})^2} + \frac{1}{(\tau^{(k)})^2} \right)^{-1} \end{cases} \quad (14.3.5)$$

因此，对任意与 μ_i 无关的量 c ，我们有

$$\begin{aligned} E[(\mu_i - c)^2 | \theta^{(k)}, \mathbf{X}] &= (E[\mu_i | \theta^{(k)}, \mathbf{X}])^2 + \text{Var}[\mu_i | \theta^{(k)}, \mathbf{X}] \\ &= (\hat{\mu}_i^{(k)} - c)^2 + (\delta_i^{(k)})^{(2)} \end{aligned}$$

在上式中，分别令 $c = \mu, x_{ij}$ ，立即得到式 (14.3.2) 在给定 $\mathbf{X}, \theta^{(k)}$ 下的期望：

$$\begin{aligned} Q(\theta | \theta^{(k)}, \mathbf{X}) &= -(m-1) \log \tau - \frac{1}{2\tau^2} \sum_{i=1}^m [(\hat{\mu}_i^{(k)} - \beta)^2 + \delta_i^{(k)}] - n \log \sigma - \frac{1}{2\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} [(\hat{\mu}_i^{(k)} - x_{ij})^2 + \delta_i^{(k)}] \end{aligned} \quad (14.3.6)$$

M 步：式(14.3.15)的极大化是简单的，通过解方程组 $\frac{\partial Q}{\partial \theta} = 0$ ，我们得到：

$$\begin{cases} \beta^{(k+1)} = \frac{1}{m} \sum_{i=1}^m \hat{\mu}_i^{(k)} \\ \sigma^{(k+1)} = \left(\frac{1}{n} \sum_{i=1}^m \sum_{j=1}^{n_i} [(\hat{\mu}_i^{(k)} - x_{ij})^2 + \delta_i^{(k)}] \right)^{1/2} \\ \tau^{(k+1)} = \left(\frac{1}{m-1} \sum_{i=1}^m [(\hat{\mu}_i^{(k)} - \beta^{(k+1)})^2 + \delta_i^{(k)}] \right)^{1/2} \end{cases} \quad (14.3.7)$$

将式(14.3.5)代入上式，我们就得到了 EM 算法的完整迭代公式。与例 14.3.2 比较，EM 算法在 M 步中的极大化方法与完全数据（即已知 $\mathbf{Y} = (\mu_1, \mu_2, \dots, \mu_m)^T$ ）情况下的极大化方法完全相同。

14.3.2 收敛性与估计精度

EM 算法是利用潜在数据将复杂的最大后验估计（或最大似然估计）转化为一系列简单的极小化问题，并通过迭代技术来实现。一个很自然的问题是 EM 算法所得到的估计序列 $\{\theta^{(j)}\}$ 一定收敛吗？如果收敛，一定收敛到最大点或局部极小点吗？作为一种估计，如何计算 EM 算法的估计精度？下面将讨论这些问题。

收敛性

关于 EM 算法的收敛性, 有下述定理。

定理 14.3.1 在 EM 算法中, 估计序列 $\{\theta^{(j)}\}$ 所对应后验密度值序列 $\pi(\theta^{(j)} | \mathbf{X})$ 是单调增的, 即

$$\pi(\theta^{(j)} | \mathbf{X}) \leq \pi(\theta^{(j+1)} | \mathbf{X}), \forall j \quad (14.3.8)$$

证明: 根据全概率公式, 我们有

$$\pi(\theta, \mathbf{Y} | \mathbf{X}) = \pi(\mathbf{Y} | \theta, \mathbf{X})\pi(\theta | \mathbf{X}) = \pi(\theta | \mathbf{X}, \mathbf{Y})\pi(\mathbf{Y} | \mathbf{X})$$

利用后一个等式, 得到

$$\log(\pi(\theta | \mathbf{X})) = \log(\pi(\theta | \mathbf{X}, \mathbf{Y})) - \log(\pi(\mathbf{Y} | \theta, \mathbf{X})) + \log(\pi(\mathbf{Y} | \mathbf{X})) \quad (14.3.9)$$

所以, 上式对 \mathbf{Y} 关于 $\pi(\mathbf{Y} | \theta^{(j)}, \mathbf{X})$ 求期望, 可得

$$\begin{aligned} \log(\pi(\theta | \mathbf{X})) &= \int [\log(\pi(\theta | \mathbf{X}, \mathbf{Y})) - \log(\pi(\mathbf{Y} | \theta, \mathbf{X})) + \log(\pi(\mathbf{Y} | \mathbf{X}))] \pi(\mathbf{Y} | \theta^{(j)}, \mathbf{X}) d\mathbf{Y} \\ &= Q(\theta | \theta^{(j)}, \mathbf{X}) - R(\theta | \theta^{(j)}, \mathbf{X}) + T(\theta^{(j)}, \mathbf{X}) \end{aligned} \quad (14.3.10)$$

其中:

$$\begin{aligned} Q(\theta | \theta^{(j)}, \mathbf{X}) &= \int \log(\pi(\theta | \mathbf{X}, \mathbf{Y})) \pi(\mathbf{Y} | \theta^{(j)}, \mathbf{X}) d\mathbf{Y} \\ R(\theta | \theta^{(j)}, \mathbf{X}) &= \int [\log(\pi(\mathbf{Y} | \theta^{(j)}, \mathbf{X}))] \pi(\mathbf{Y} | \theta^{(j)}, \mathbf{X}) d\mathbf{Y} \\ T(\theta^{(j)}, \mathbf{X}) &= \int \log(\pi(\mathbf{Y} | \mathbf{X})) \pi(\mathbf{Y} | \theta^{(j)}, \mathbf{X}) d\mathbf{Y} \end{aligned}$$

将 $\theta = \theta^{(j)}, \theta^{(j+1)}$ 分别代入(14.3.10), 我们得到

$$\begin{aligned} &\log(\pi(\theta^{(j+1)} | \mathbf{X})) - \log(\pi(\theta^{(j)} | \mathbf{X})) \\ &= [Q(\theta | \theta^{(j+1)}, \mathbf{X}) - Q(\theta | \theta^{(j)}, \mathbf{X})] - [R(\theta | \theta^{(j+1)}, \mathbf{X}) - R(\theta | \theta^{(j)}, \mathbf{X})] \end{aligned} \quad (14.3.11)$$

由 Jensen 不等式, 得到

$$\begin{aligned} R(\theta | \theta^{(j+1)}, \mathbf{X}) - R(\theta | \theta^{(j)}, \mathbf{X}) &= E_{(\mathbf{Y} | \theta^{(j)}, \mathbf{X})} \left[\log \frac{\pi(\mathbf{Y} | \theta^{(j+1)}, \mathbf{X})}{\pi(\mathbf{Y} | \theta^{(j)}, \mathbf{X})} \right] \\ &\leq \log \left\{ E_{(\mathbf{Y} | \theta^{(j)}, \mathbf{X})} \left[\frac{\pi(\mathbf{Y} | \theta^{(j+1)}, \mathbf{X})}{\pi(\mathbf{Y} | \theta^{(j)}, \mathbf{X})} \right] \right\} \leq 0 \end{aligned} \quad (14.3.12)$$

又因 $\theta^{(j+1)} = \arg \max_{\theta} Q(\theta | \theta^{(j)}, \mathbf{X})$, 所以,

$$Q(\theta | \theta^{(j+1)}, \mathbf{X}) - Q(\theta | \theta^{(j)}, \mathbf{X}) \geq 0 \quad (14.3.13)$$

从式(14.3.11-13), 我们有

$$\log(\pi(\theta^{(j+1)} | \mathbf{X})) - \log(\pi(\theta^{(j)} | \mathbf{X})) \geq 0$$

故 $\pi(\theta^{(j)} | \mathbf{X}) \leq \pi(\theta^{(j+1)} | \mathbf{X}), \forall j$ 。证毕。

定理 14.3.2 在 EM 算法中, (1) 如果 $\pi(\theta | \mathbf{X})$ 有界, 则 $\{\pi(\theta^{(j)} | \mathbf{X})\}$ 收敛;

(2) 如果 $Q(\theta|\varphi)$ 关于 θ, φ 都是连续的, 则在 $\pi(\theta|\mathbf{X})$ 满足很一般的条件下, 估计序列 $\{\theta^{(j)}\}$ 收敛到 $\pi(\theta|\mathbf{X})$ 的稳定点。

注 在定理 14.3.2 的条件下, 只能保证估计点列收敛到后验密度的稳定点, 但不能保证收敛到最大点或局部极大点, 这是所有迭代算法的共性。在实际应用中, 我们只能通过选择一个好的初始点, 或选取一系列的初始点进行 EM 迭代, 然后比较迭代结果确定后验密度的最大点。

估计精度

Louis 算法 假定 θ^* 是 EM 算法的估计结果, 则 θ^* 的渐近方差可以用 Fisher 观测信息矩阵的逆来近似:

$$I_{\theta^*}^{-1} = \left(- \frac{\partial^2 \log \pi(\theta|\mathbf{X})}{\partial \theta_k \partial \theta_l} \Big|_{\theta=\theta^*} \right)^{-1} \quad (14.3.14)$$

因此, 关键在于(14.3.14)的计算。下面是 Louis 所给的算法。

由式(14.3.9), 我们有

$$\frac{\partial^2 \log(\pi(\theta|\mathbf{X}))}{\partial \theta_k \partial \theta_l} = \frac{\partial^2 \log(\pi(\theta|\mathbf{X}, \mathbf{Y}))}{\partial \theta_k \partial \theta_l} - \frac{\partial^2 \log(\pi(\mathbf{Y}|\theta, \mathbf{X}))}{\partial \theta_k \partial \theta_l}$$

将上式对 $\pi(\mathbf{Y}|\theta, \mathbf{X})$ 求期望, 得到

$$\begin{aligned} \frac{\partial^2 \log(\pi(\theta|\mathbf{X}))}{\partial \theta_k \partial \theta_l} &= \int \frac{\partial^2 \log(\pi(\theta|\mathbf{X}, \mathbf{Y}))}{\partial \theta_k \partial \theta_l} \pi(\mathbf{Y}|\theta, \mathbf{X}) d\mathbf{Y} - \int \frac{\partial^2 \log(\pi(\mathbf{Y}|\theta, \mathbf{X}))}{\partial \theta_k \partial \theta_l} \pi(\mathbf{Y}|\theta, \mathbf{X}) d\mathbf{Y} \\ &= \frac{\partial^2 Q(\theta|\varphi, \mathbf{X})}{\partial \theta_k \partial \theta_l} \Big|_{\varphi=\theta} - \frac{\partial^2 H(\theta|\varphi, \mathbf{X})}{\partial \theta_k \partial \theta_l} \Big|_{\varphi=\theta} \end{aligned}$$

于是, 我们有

$$\left(- \frac{\partial^2 \log \pi(\theta|\mathbf{X})}{\partial \theta_k \partial \theta_l} \Big|_{\theta=\theta^*} \right) = \left(- \frac{\partial^2 Q(\theta|\varphi, \mathbf{X})}{\partial \theta_k \partial \theta_l} \Big|_{\varphi=\theta^*} \right) - \left(- \frac{\partial^2 H(\theta|\varphi, \mathbf{X})}{\partial \theta_k \partial \theta_l} \Big|_{\varphi=\theta^*} \right) \quad (14.3.15)$$

通常, $-\left(\frac{\partial^2 Q}{\partial \theta_k \partial \theta_l}\right)$, $-\left(\frac{\partial^2 H}{\partial \theta_k \partial \theta_l}\right)$ 分别称为完全信息与缺损信息, 因此式(14.3.15)给出了所谓的缺

损信息原则: 观测信息 = 完全信息 - 缺损信息。

Louis 曾经还得到了一个重要结论:

$$-\frac{\partial^2 \log \pi(\boldsymbol{\theta} | \mathbf{X})}{\partial \theta_k \partial \theta_l} = -E_{(Y|\theta, X)} \left(\frac{\partial^2 \log \pi(\boldsymbol{\theta} | \mathbf{X}, \mathbf{Y})}{\partial \theta_k \partial \theta_l} \right) - \text{Var}_{(Y|\theta, X)} \left(\frac{\partial \log \pi(\boldsymbol{\theta} | \mathbf{X}, \mathbf{Y})}{\partial \boldsymbol{\theta}} \right) \quad (14.3.16)$$

将 $\boldsymbol{\theta} = \boldsymbol{\theta}^*$ 代入上式, 取逆就得到 $\boldsymbol{\theta}^*$ 的渐近方差。

模拟 (Monte Carlo) 计算 从分布 $\log \pi(\mathbf{Y} | \boldsymbol{\theta}^*, \mathbf{X})$ 抽取 r 个样本: $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_r$, 当 r 充分大时, 则(14.3.16)右端均值可按下述近似公式进行计算:

$$\begin{aligned} E_{(Y|\theta, X)} \left(\frac{\partial^2 \log \pi(\boldsymbol{\theta} | \mathbf{X}, \mathbf{Y})}{\partial \theta_k \partial \theta_l} \right) &= \left(\int \frac{\partial^2 \log \pi(\boldsymbol{\theta} | \mathbf{X}, \mathbf{Y})}{\partial \theta_k \partial \theta_l} \pi(\mathbf{Y} | \boldsymbol{\theta}^*, \mathbf{X}) d\mathbf{Y} \right) \\ &\approx \left(\frac{1}{r} \sum_{j=1}^r \frac{\partial^2 \log \pi(\boldsymbol{\theta} | \mathbf{X}, \mathbf{y}_j)}{\partial \theta_k \partial \theta_l} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} \right) \end{aligned} \quad (14.3.17)$$

类似地, 协方差可按下述近似公式计算:

$$\begin{aligned} \text{Var}_{(Y|\theta, X)} \left(\frac{\partial \log \pi(\boldsymbol{\theta} | \mathbf{X}, \mathbf{Y})}{\partial \boldsymbol{\theta}} \right) &\approx \left(\frac{1}{r} \sum_{j=1}^r \left(\frac{\partial \log \pi(\boldsymbol{\theta} | \mathbf{X}, \mathbf{y}_j)}{\partial \theta_k} \cdot \frac{\partial \log \pi(\boldsymbol{\theta} | \mathbf{X}, \mathbf{y}_j)}{\partial \theta_l} \right) \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} \right) \\ &\quad - \left(\frac{1}{r^2} \left(\sum_{j=1}^r \frac{\partial \log \pi(\boldsymbol{\theta} | \mathbf{X}, \mathbf{y}_j)}{\partial \theta_k} \cdot \sum_{j=1}^r \frac{\partial \log \pi(\boldsymbol{\theta} | \mathbf{X}, \mathbf{y}_j)}{\partial \theta_l} \right) \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} \right) \end{aligned} \quad (14.3.18)$$

14.3.3 EM 算法的推广

GEM 算法

在 EM 算法中, M 步的极大化要求寻找 Q 函数的极小点 (见式(14.3.2))。在实际应用中, Q 函数通常都是多变量的高度非线性函数, 这使得 M 步的实现异常困难, 广义 EM (GEM) 算法就是针对这一困难而提出的。GEM 方法降低了 M 步求极大点的要求, 它在 M 步是寻找一个点 $\boldsymbol{\theta}^{(j+1)}$ 使得下式成立:

$$Q(\boldsymbol{\theta}^{(j+1)} | \boldsymbol{\theta}^{(j)}, \mathbf{X}) > Q(\boldsymbol{\theta}^{(j)} | \boldsymbol{\theta}^{(j-1)}, \mathbf{X}) \quad (14.3.17)$$

GEM 算法所产生的迭代点列 $\{\boldsymbol{\theta}^{(j)}\}$ 也能保证后验密度函数值序列 $\pi(\boldsymbol{\theta}^{(j)} | \mathbf{X})$ 是单调增的, 即

$$\pi(\boldsymbol{\theta}^{(j)} | \mathbf{X}) \leq \pi(\boldsymbol{\theta}^{(j+1)} | \mathbf{X}), \forall j$$

在一定的条件下定理 14.3.2 也成立。

Meng 和 Rubin 提出了一种特殊的 GEM 算法, 他们称之为 ECM (Expectation/Conditional

Maximization) 算法, 该算法通过一系列单变量的极大化实现 M 步。令 $\theta^{(j)} = (\theta_1^{(j)}, \theta_2^{(j)}, \dots, \theta_k^{(j)})$ 是当前的估计值, 则 ECM 利用下述一系列的极大化得到 $\theta^{(j+1)}$:

$$\begin{cases} \theta_1^{(j+1)} = \arg \max_{\theta_1} Q(\theta_1, \theta_2^{(j)}, \dots, \theta_{k-1}^{(j)}, \theta_k^{(j)}) \\ \theta_2^{(j+1)} = \arg \max_{\theta_2} Q(\theta_1^{(j+1)}, \theta_2, \dots, \theta_{k-1}^{(j)}, \theta_k^{(j)}) \\ \vdots \\ \theta_k^{(j+1)} = \arg \max_{\theta_k} Q(\theta_1^{(j+1)}, \theta_2^{(j+1)}, \dots, \theta_{k-1}^{(j+1)}, \theta_k) \end{cases} \quad (14.3.10)$$

ECM 保持了 EM 的简单性和稳定性, 它是值得推荐的算法。

MCEM (Monte Carlo EM) 算法

GEM 算法是解决 EM 算法在 M 步的极大化时可能带来的困难。实际上, EM 算法的 E 步涉及期望的积分计算, 在众多实际问题中, 也同样存在计算上的困难。MCEM 算法通过随机模拟来完成 E 步, 具体步骤如下:

E1 步: 由分布 $\pi(\mathbf{Y} | \theta^{(j)}, \mathbf{X})$ 随机抽取 r 个随机数 (向量): $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_r$;

E2 步: 计算 $\hat{Q}(\theta | \theta^{(j)}, \mathbf{X}) \equiv \frac{1}{r} \sum_{j=1}^r \log \pi(\theta | \mathbf{y}_j, \mathbf{X})$, (或 $\frac{1}{r} \sum_{j=1}^r \pi(\theta | \mathbf{y}_j, \mathbf{X})$)。

根据大数定律, 当 r 充分大时, $\hat{Q}(\theta | \theta^{(j)}, \mathbf{X})$ 就足够接近

$$Q(\theta | \theta^{(j)}, \mathbf{X}) = \int (\log \pi(\theta | \mathbf{Y}, \mathbf{X})) \pi(\mathbf{Y} | \theta^{(j)}, \mathbf{X}) d\mathbf{Y}$$

因此, 在 M 步可以对 $\hat{Q}(\theta | \theta^{(j)}, \mathbf{X})$ 极大化。

在 MCEM 算法中, 最棘手的两个问题是:

- (1) r 究竟取多大? 从逼近 $Q(\theta | \theta^{(j)}, \mathbf{X})$ 的精度考虑, 自然是愈大愈好。从计算的角度, 愈小愈好。计算效率不仅仅体现在 E2 步, 更多的是体现在 M 步, 因为 r 愈大 $\hat{Q}(\theta | \theta^{(j)}, \mathbf{X})$ 愈复杂, 从而极大化更加困难。对此, 目前还没有非常有效方法, 通常所选取的策略是在满足计算资源的要求下, 尽可能取较大的 r 。
- (2) 如何终止迭代? 即收敛性的判断问题。在 MCEM 中, 要求迭代点列 $\theta^{(j)}$ 收敛到某一点是不现实的。对此, 人们普遍使用的方法是 ECEM 进行一定的迭代次数后, 观察以后的迭代点是否在某个点附近只有小幅度的摆动, 如果是, 则认为算法已收敛。否则, 继续迭代一定的次数后, 再重新观察。

14.4 混合模型的 EM 算法

混合模型是计算机视觉中最常用的概率模型, 本节主要讨论它的 EM 算法。在混合模型的 EM 算法中, E 步的 Q 函数具有解析表达式, M 步的混合系数存在显式迭代表达式, 但 M 步中的成份密度参数的确定依赖于成份密度函数的复杂性, 但对于高斯成份密度参数可以得到显式迭代表达式。

14.4.1 一般混合模型

根据 14.1.3 节的讨论, K 个成份的混合模型由下式定义:

$$p(\mathbf{x} | \boldsymbol{\alpha}, \boldsymbol{\Theta}) = \sum_{k=1}^K \alpha_k p_k(\mathbf{x} | \boldsymbol{\theta}_k) \quad (14.4.1)$$

其中: \mathbf{x} 是 D 维随机向量;

$p_k(\mathbf{x} | \boldsymbol{\theta}_k)$ 是第 k 个成份的密度函数, $\boldsymbol{\theta}_k$ 是它的参数向量;

$\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_K)$ ($\sum_{k=1}^K \alpha_k = 1$) 是混合系数向量, α_k 表示第 k 个成份被选择的概率;

$\boldsymbol{\Theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_K)$ 是成份密度参数所构成的向量。

如果每个成份密度都是高斯的, 则称(14.4.1)为高斯混合模型。

给定观测样本 $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$, 则参数 $(\boldsymbol{\alpha}, \boldsymbol{\Theta})$ 的对数似然为

$$l(\mathbf{X}; \boldsymbol{\alpha}, \boldsymbol{\Theta}) = \log \prod_{n=1}^N p(\mathbf{x}_n | \boldsymbol{\alpha}, \boldsymbol{\Theta}) = \sum_{n=1}^N \log \left(\sum_{k=1}^K \alpha_k p_k(\mathbf{x}_n | \boldsymbol{\theta}_k) \right) \quad (14.4.2)$$

直接求解极大似然将是非常困难的。

现在, 引进潜在数据 $\mathbf{Y} = (y_1, y_2, \dots, y_N)$, 其中 $y_n \in \{1, 2, \dots, K\}$, 如果 \mathbf{x}_n 是由第 k 个成份所生成的, 则 $y_n = k$ 。 \mathbf{X} 与 \mathbf{Y} 构成完全数据:

$$\{\mathbf{X}, \mathbf{y}\} = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_N, y_N)\}$$

对于这个完全数据, 参数 $(\boldsymbol{\alpha}, \boldsymbol{\Theta})$ 的对数似然为

$$l(\mathbf{X}, \mathbf{y}; \boldsymbol{\alpha}, \boldsymbol{\Theta}) = \log P_r(\mathbf{X}, \mathbf{y} | \boldsymbol{\alpha}, \boldsymbol{\Theta}) = \log \prod_{n=1}^N P_r(\mathbf{x}_n | y_n) P_r(y_n) = \sum_{n=1}^N \log(\alpha_{y_n} p_{y_n}(\mathbf{x}_n | \boldsymbol{\theta}_{y_n})) \quad (14.4.3)$$

EM 算法的 E 步

假定第 $s+1$ 步迭代开始时, 参数的估计值是:

$$\boldsymbol{\alpha}^{(s)} = (\alpha_1^{(s)}, \alpha_2^{(s)}, \dots, \alpha_K^{(s)}), \quad \boldsymbol{\Theta}^{(s)} = (\boldsymbol{\theta}_1^{(s)}, \boldsymbol{\theta}_2^{(s)}, \dots, \boldsymbol{\theta}_K^{(s)})$$

则容易计算 $p_k(\mathbf{x}_n | \theta_k^{(s)})$ 。所以, 由贝叶斯公式, y_n 的条件密度为

$$p(y_n | \mathbf{x}_n, \mathbf{a}^{(s)}, \Theta^{(s)}) = \frac{p(\mathbf{x}_n, y_n | \mathbf{a}^{(s)}, \Theta^{(s)})}{p(\mathbf{x}_n | \mathbf{a}^{(s)}, \Theta^{(s)})} = \frac{\alpha_{y_n}^{(s)} p_{y_n}(\mathbf{x}_n | \theta_{y_n}^{(s)})}{\sum_{k=1}^K \alpha_k^{(s)} p(\mathbf{x}_n | \theta_k^{(s)})} \quad (14.4.4)$$

从而, \mathbf{y} 的联合条件密度为

$$p(\mathbf{y} | X, \mathbf{a}^{(s)}, \Theta^{(s)}) = \prod_{j=1}^N p(y_j | \mathbf{x}_j, \mathbf{a}^{(s)}, \Theta^{(s)}) \quad (14.4.5)$$

于是, $l(X, \mathbf{y}; \mathbf{a}, \Theta)$ 关于联合条件密度(14.4.5)的期望为

$$\begin{aligned} Q(\mathbf{a}, \Theta; \mathbf{a}^{(s)}, \Theta^{(s)}) &= E[l(X, \mathbf{y}; \mathbf{a}, \Theta) | X, \mathbf{a}^{(s)}, \Theta^{(s)}] \\ &= \sum_{\mathbf{y} \in \{1, 2, \dots, K\}^N} \left(\sum_{n=1}^N \log(\alpha_{y_n} p_{y_n}(\mathbf{x}_n | \theta_{y_n})) \right) \prod_{j=1}^N p(y_j | \mathbf{x}_j, \mathbf{a}^{(s)}, \Theta^{(s)}) \\ &= \sum_{y_1=1}^K \sum_{y_2=1}^K \dots \sum_{y_N=1}^K \left(\sum_{n=1}^N \log(\alpha_{y_n} p_{y_n}(\mathbf{x}_n | \theta_{y_n})) \right) \prod_{j=1}^N p(y_j | \mathbf{x}_j, \mathbf{a}^{(s)}, \Theta^{(s)}) \\ &= \sum_{y_1=1}^K \sum_{y_2=1}^K \dots \sum_{y_N=1}^K \sum_{n=1}^N \sum_{k=1}^K \delta_{k, y_n} \log(\alpha_k p_k(\mathbf{x}_n | \theta_k)) \prod_{j=1}^N p(y_j | \mathbf{x}_j, \mathbf{a}^{(s)}, \Theta^{(s)}) \\ &= \sum_{y_1=1}^K \sum_{y_2=1}^K \dots \sum_{y_N=1}^K \sum_{n=1}^N \sum_{k=1}^K \delta_{k, y_n} \log(\alpha_k p_k(\mathbf{x}_n | \theta_k)) \prod_{j=1}^N p(y_j | \mathbf{x}_j, \mathbf{a}^{(s)}, \Theta^{(s)}) \\ &= \sum_{n=1}^N \sum_{k=1}^K \log(\alpha_k p_k(\mathbf{x}_n | \theta_k)) \sum_{y_1=1}^K \sum_{y_2=1}^K \dots \sum_{y_N=1}^K \delta_{k, y_n} \prod_{j=1}^N p(y_j | \mathbf{x}_j, \mathbf{a}^{(s)}, \Theta^{(s)}) \end{aligned}$$

下面, 进一步化简上式。由于

$$\begin{aligned} &\sum_{y_1=1}^K \sum_{y_2=1}^K \dots \sum_{y_N=1}^K \delta_{k, y_n} \prod_{j=1}^N p(y_j | \mathbf{x}_j, \mathbf{a}^{(s)}, \Theta^{(s)}) \\ &= \left(\sum_{y_1=1}^K \sum_{y_2=1}^K \dots \sum_{y_{n-1}=1}^K \sum_{y_{n+1}=1}^K \dots \sum_{y_N=1}^K \prod_{j=1, j \neq n}^N p(y_j | \mathbf{x}_j, \mathbf{a}^{(s)}, \Theta^{(s)}) \right) p(k | \mathbf{x}_n, \mathbf{a}^{(s)}, \Theta^{(s)}) \\ &= \prod_{j=1, j \neq n}^N \left(\sum_{y_j=1}^K p(y_j | \mathbf{x}_j, \mathbf{a}^{(s)}, \Theta^{(s)}) \right) p(k | \mathbf{x}_n, \mathbf{a}^{(s)}, \Theta^{(s)}) \\ &= p(k | \mathbf{x}_n, \mathbf{a}^{(s)}, \Theta^{(s)}) \end{aligned}$$

在第一个等式和最后一个等式的推导过程中, 分别使用了下述等式:

$$\sum_{y_n=1}^K \delta_{k, y_n} p(y_n | \mathbf{x}_n, \mathbf{a}^{(s)}, \Theta^{(s)}) = p(k | \mathbf{x}_n, \mathbf{a}^{(s)}, \Theta^{(s)})$$

$$\sum_{y_j=1}^K p(y_j | \mathbf{x}_j, \boldsymbol{\alpha}^{(s)}, \boldsymbol{\Theta}^{(s)}) = \sum_{k=1}^K p(k | \mathbf{x}_j, \boldsymbol{\alpha}^{(s)}, \boldsymbol{\Theta}^{(s)}) = 1$$

所以，

$$\begin{aligned} Q(\boldsymbol{\alpha}, \boldsymbol{\Theta}; \boldsymbol{\alpha}^{(s)}, \boldsymbol{\Theta}^{(s)}) &= \sum_{n=1}^N \sum_{k=1}^K \log(\alpha_k p_k(\mathbf{x}_n | \boldsymbol{\theta}_k)) p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \boldsymbol{\Theta}^{(s)}) \\ &= Q_1(\boldsymbol{\alpha}; \boldsymbol{\alpha}^{(s)}, \boldsymbol{\Theta}^{(s)}) + Q_2(\boldsymbol{\Theta}; \boldsymbol{\alpha}^{(s)}, \boldsymbol{\Theta}^{(s)}) \end{aligned} \quad (14.4.6)$$

其中：

$$Q_1(\boldsymbol{\alpha}; \boldsymbol{\alpha}^{(s)}, \boldsymbol{\Theta}^{(s)}) = \sum_{n=1}^N \sum_{k=1}^K \log(\alpha_k) p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \boldsymbol{\Theta}^{(s)}) \quad (14.4.7)$$

$$Q_2(\boldsymbol{\Theta}; \boldsymbol{\alpha}^{(s)}, \boldsymbol{\Theta}^{(s)}) = \sum_{n=1}^N \sum_{k=1}^K (\log p_k(\mathbf{x}_n | \boldsymbol{\theta}_k)) p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \boldsymbol{\Theta}^{(s)}) \quad (14.4.8)$$

这样，我们就完成了 E 步。

EM 算法的 M 步

M 步是求解下述最大化问题：

$$\boldsymbol{\alpha}^{(s+1)} = \arg \max_{\sum_{i=1}^K \alpha_i} Q_1(\boldsymbol{\alpha}; \boldsymbol{\alpha}^{(s)}, \boldsymbol{\Theta}^{(s)}) \quad (14.4.9)$$

$$\boldsymbol{\Theta}^{(s+1)} = \arg \max_{\boldsymbol{\Theta}} Q_2(\boldsymbol{\Theta}; \boldsymbol{\alpha}^{(s)}, \boldsymbol{\Theta}^{(s)}) \quad (14.4.10)$$

首先，求解最大化问题(14.4.9)。为此，引进 $Q_1(\boldsymbol{\alpha}; \boldsymbol{\alpha}^{(s)}, \boldsymbol{\Theta}^{(s)})$ 的 Lagrange 乘子函数：

$$La(\boldsymbol{\alpha}, \lambda; \boldsymbol{\alpha}^{(s)}, \boldsymbol{\Theta}^{(s)}) = Q_1(\boldsymbol{\alpha}; \boldsymbol{\alpha}^{(s)}, \boldsymbol{\Theta}^{(s)}) + \lambda \left(\sum_{k=1}^K \alpha_k - 1 \right) \quad (14.4.11)$$

对上式求偏导，并令其为零，得到

$$\begin{cases} \frac{\partial La(\boldsymbol{\alpha}, \lambda; \boldsymbol{\alpha}^{(s)}, \boldsymbol{\Theta}^{(s)})}{\partial \alpha_k} = \frac{1}{\alpha_k} \sum_{n=1}^N p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \boldsymbol{\Theta}^{(s)}) + \lambda = 0, k=1, 2, \dots, K \\ \frac{\partial La(\boldsymbol{\alpha}, \lambda; \boldsymbol{\alpha}^{(s)}, \boldsymbol{\Theta}^{(s)})}{\partial \lambda} = \sum_{k=1}^K \alpha_k - 1 = 0 \end{cases}$$

解上述方程组，我们有

$$\begin{cases} \lambda = -N \\ \alpha_k^{(s+1)} = \frac{1}{N} \sum_{n=1}^N p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \boldsymbol{\Theta}^{(s)}), k=1, 2, \dots, K \end{cases} \quad (14.4.12)$$

对于最大化问题(14.4.10)，不能像求解混合参数 α_k 那样，用密度函数给出它的显式迭代公式。但是，对于高斯密度而言，仍能给出最大化问题(14.4.10)的显式迭代公式(见下一节)。因此，在一般情况下，需要使用 16 章给出的优化技术才能得到最大化问题(14.4.10)的数值解。

综合上述讨论，我们得到混合模型的 EM 算法：

给定初始值： $\alpha_i^{(0)}$ ， $\boldsymbol{\Theta}^{(0)}$ ，按下述公式进行迭代直至收敛：

$$\alpha_l^{(s+1)} = \frac{1}{N} \sum_{n=1}^N p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \Theta^{(s)}), k=1,2,\dots,K$$

$$\Theta^{(s+1)} = \arg \max_{\Theta} \sum_{n=1}^N \sum_{k=1}^K (\log p_k(\mathbf{x}_n | \boldsymbol{\theta}_k)) p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \Theta^{(s)})$$

14.4.2 高斯混合模型

对于高斯混合模型，需要讨论的仅是求解最大化问题(14.4.10)。

令高斯密度函数为

$$p_k(\mathbf{x} | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = \frac{1}{(2\pi)^{D/2} (\det \boldsymbol{\Sigma}_k)^{1/2}} \exp\left(-\frac{1}{2} \|\mathbf{x} - \boldsymbol{\mu}_k\|_{\boldsymbol{\Sigma}_k^{-1}}^2\right) \quad (14.4.13)$$

则成份密度参数向量为

$$\Theta = ((\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1), (\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2), \dots, (\boldsymbol{\mu}_K, \boldsymbol{\Sigma}_K))$$

且 Q_2 函数为

$$Q_2(\Theta; \boldsymbol{\alpha}^{(s)}, \Theta^{(s)}) = \sum_{n=1}^N \sum_{k=1}^K (\log p_k(\mathbf{x}_n | \boldsymbol{\theta}_k)) p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \Theta^{(s)})$$

$$= \sum_{n=1}^N \sum_{k=1}^K \left(-\frac{1}{2} \log(\det \boldsymbol{\Sigma}_k) - \frac{1}{2} (\mathbf{x}_n - \boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1} (\mathbf{x}_n - \boldsymbol{\mu}_k) \right) p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \Theta^{(s)}) + C$$

其中 C 是与 Θ 无关的常函数。

对 Q_2 函数关于 $\boldsymbol{\mu}_k$ 求偏导，并令其为零，有

$$\frac{\partial Q_2(\Theta; \boldsymbol{\alpha}^{(s)}, \Theta^{(s)})}{\partial \boldsymbol{\mu}_k} = \sum_{n=1}^N \boldsymbol{\Sigma}_k^{-1} (\mathbf{x}_n - \boldsymbol{\mu}_k) p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \Theta^{(s)}) = \mathbf{0}$$

于是，得到

$$\boldsymbol{\mu}_k^{(s+1)} = \frac{\sum_{n=1}^N \mathbf{x}_n p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \Theta^{(s)})}{\sum_{n=1}^N p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \Theta^{(s)})}, k=1,2,\dots,K \quad (14.4.14)$$

下面求 $\boldsymbol{\Sigma}_k^{(s+1)}$ 。为此，将 Q_2 函数改写成：

$$Q_2(\Theta; \boldsymbol{\alpha}^{(s)}, \Theta^{(s)}) = \sum_{k=1}^K \left[\frac{1}{2} \log(\det \boldsymbol{\Sigma}_k^{-1}) \sum_{n=1}^N p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \Theta^{(s)}) - \sum_{n=1}^N p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \Theta^{(s)}) \text{tr}(\boldsymbol{\Sigma}_k^{-1} M_{n,k}) \right] + C$$

其中：

$$M_{n,k} = (\mathbf{x}_n - \boldsymbol{\mu}_k^{(s+1)})(\mathbf{x}_n - \boldsymbol{\mu}_k^{(s+1)})^T$$

对 $\boldsymbol{\Sigma}_k^{-1}$ 求偏导，有

$$\begin{aligned}
\frac{\partial Q_2(\Theta; \boldsymbol{\alpha}^{(s)}, \Theta^{(s)})}{\partial \Sigma_k^{-1}} &= \frac{1}{2} (2\Sigma_k - \text{diag}\Sigma_k) \sum_{n=1}^N p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \Theta^{(s)}) - \frac{1}{2} \sum_{n=1}^N p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \Theta^{(s)}) (2M_{n,k} - \text{diag}M_{n,k})^*) \\
&= \frac{1}{2} (2\Sigma_k - \text{diag}\Sigma_k) \sum_{n=1}^N p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \Theta^{(s)}) - \frac{1}{2} \sum_{n=1}^N p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \Theta^{(s)}) (2M_{n,k} - \text{diag}M_{n,k}) \\
&= 2G - \text{diag}G
\end{aligned}$$

其中:

$$G = \frac{1}{2} \sum_{n=1}^N p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \Theta^{(s)}) (\Sigma_k - M_{n,k})$$

令 $\frac{\partial Q_2(\Theta; \boldsymbol{\alpha}^{(s)}, \Theta^{(s)})}{\partial \Sigma_k^{-1}} = \boldsymbol{\theta}$, 即 $2G - \text{diag}G = 0$ 。因此, 必有

$$G = \frac{1}{2} \sum_{n=1}^N p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \Theta^{(s)}) (\Sigma_k - M_{n,k}) = 0 \quad (14.4.15)$$

于是, 得到

$$\Sigma_k^{(s+1)} = \frac{\sum_{n=1}^N p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \Theta^{(s)}) M_{n,k}}{\sum_{n=1}^N p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \Theta^{(s)})} = \frac{\sum_{n=1}^N p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \Theta^{(s)}) (\mathbf{x}_n - \boldsymbol{\mu}_k^{(s+1)}) (\mathbf{x}_n - \boldsymbol{\mu}_k^{(s+1)})^T}{\sum_{n=1}^N p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \Theta^{(s)})} \quad (14.4.15)$$

综上所述, 我们有高斯混合模型 EM 算法的下述迭代公式:

$$\alpha_l^{(s+1)} = \frac{1}{N} \sum_{n=1}^N p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \Theta^{(s)}), k=1,2,\dots,K \quad (14.4.16)$$

$$\boldsymbol{\mu}_k^{(s+1)} = \frac{\sum_{n=1}^N \mathbf{x}_n p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \Theta^{(s)})}{\sum_{n=1}^N p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \Theta^{(s)})}, k=1,2,\dots,K \quad (14.4.17)$$

$$\Sigma_k^{(s+1)} = \frac{\sum_{n=1}^N p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \Theta^{(s)}) (\mathbf{x}_n - \boldsymbol{\mu}_k^{(s+1)}) (\mathbf{x}_n - \boldsymbol{\mu}_k^{(s+1)})^T}{\sum_{n=1}^N p(k | \mathbf{x}_n, \boldsymbol{\alpha}^{(s)}, \Theta^{(s)})}, k=1,2,\dots,K \quad (14.4.18)$$

^{*} 这里使用了等式: $\frac{\partial \log \det(A)}{\partial A} = 2A^{-1} - \text{diag}(A^{-1})$, $\frac{\partial \text{tr}(AB)}{\partial A} = 2B + B^T - \text{diag}(B)$, 见第 9 章。

15. 代数方法

计算机视觉中的估计问题，通常是指模型（或变换）以及某个数学实体的估计，这些估计问题是以实际测量数据为基础的。如果没有测量误差的话，这些估计问题在数学上通常是容易求解的。如果测量数据存在误差时，给出好的估计却是十分困难的，因为视觉中的大多数估计问题对测量误差十分敏感。在计算机视觉估计问题的方法中，可分为如下四类：代数方法（基于代数距离的最小化方法），这一类方法是线性的，所以通常又称为线性估计方法。第二类是几何方法（基于几何距离的最小化方法），这类方法是非线性的，需要迭代计算，所以通常又称为迭代估计方法。以上两类方法只适用于测量数据仅有小测量误差情况，而不适用于测量数据中包含错误数据点的情况。第三类是鲁棒方法，对于测量数据包含错误数据点的情况，这类方法也能给出较好的估计。第四类是统计方法，是基于概率框架下对模型参数进行估计，它的计算更为复杂、更为精细。在这四类方法中，线性估计方法是最基本的，由于它不需要复杂的计算，所以在实践被广泛地使用。本章主要介绍线性估计方法。

15.1 估计问题概述

15.1.1 模型

给出模型确切的数学定义是困难的。在计算机视觉中的估计问题，通常是所涉及的变换估计或其它数学实体的估计，因此本书中的模型是指变换或某个数学实体。下面列举一些在三维计算机视觉中常见的估计问题。

直线估计 给定平面点集 $X = \{(x_j, y_j) | 1 \leq j \leq n\}$ ，估计一条直线 $l = (a, b, c)$ 使得点集 X 尽可能地“接近”该直线，即尽可能地使下述等式成立：

$$ax_j + by_j + c = 0, (x_j, y_j) \in X \quad (15.1.1)$$

这一问题通常出现在图像数据的直线拟合中, 例如图像边缘点的直线拟合, 进而求两条直线的交点或求空间平行直线在图像平面的隐消点等等。

二次曲线(曲面)估计 给定平面点集 $X = \{(x_j, y_j) | 1 \leq j \leq n\}$, 估计一条二次曲线 C 使得点集 X 尽可能地“接近”该二次曲线, 即尽可能地使下述等式成立:

$$(x_j, y_j, 1) C \begin{pmatrix} x_j \\ y_j \\ 1 \end{pmatrix} = 0, \quad (x_j, y_j) \in X \quad (15.1.2)$$

这一问题通常出现在图像数据的二次曲线的图像拟合中, 因为二次曲线的图像仍是一条二次曲线, 例如: 在基于圆环点的摄像机自标定中, 由圆环点的图像点拟合出绝对二次曲线的图像曲线, 即摄像机 IAC。对于二次曲面的估计, 也可以给出类似的描述。二次曲面估计也常常出现在三维计算机视觉的有关问题中, 例如在一般射影空间估计绝对二次曲面而获得度量重构。

在上述两个例子中, 平面点集 $X = \{(x_j, y_j) | 1 \leq j \leq n\}$ 通常也说成是相应模型的点对应集。而等式(15.1.1)与等式(15.1.2)是由模型关系所确定的。

基本矩阵估计 给定两幅图像的点对应集:

$$X = \{(\tilde{\mathbf{x}}_j, \tilde{\mathbf{y}}_j) | \tilde{\mathbf{x}}_j = (x_j^1, x_j^2)^T \in R^2, \tilde{\mathbf{y}}_j = (y_j^1, y_j^2)^T, 1 \leq j \leq n\}$$

估计一个矩阵 F 使得数据点它尽可能地“满足”下述等式:

$$f(F, \tilde{\mathbf{x}}_j, \tilde{\mathbf{y}}_j) = (\tilde{\mathbf{x}}_j^T, 1)^T F \begin{pmatrix} \tilde{\mathbf{y}}_j \\ 1 \end{pmatrix} = 0, \quad (\tilde{\mathbf{x}}_j, \tilde{\mathbf{y}}_j) \in X \quad (15.1.3)$$

三焦张量估计 给定三幅图像的点对应集:

$$X = \{(\mathbf{x}_j, \mathbf{y}_j, \mathbf{z}_j) | \mathbf{x}_j = (x_j^1, x_j^2, x_j^3), \tilde{\mathbf{y}}_j = (y_j^1, y_j^2, y_j^3), \mathbf{z}_j = (z_j^1, z_j^2, z_j^3), 1 \leq j \leq n\}$$

估计一个三焦张量 t_i^{kl} 使得点对应集尽可能地“满足”下述等式:

$$x_j^i y_j^k z_j^l \varepsilon_{jkl} \varepsilon_{khl} t_i^{kl} = 0_{uvj}, \quad (\mathbf{x}_j, \mathbf{y}_j, \mathbf{z}_j) \in X \quad (15.1.4)$$

变换估计 给定 N 维空间到 M 维空间的点对应集:

$$X = \{(\tilde{\mathbf{x}}_j, \tilde{\mathbf{y}}_j) | \tilde{\mathbf{x}}_j = (x_j^1, x_j^2, \dots, x_j^N) \in R^N, \tilde{\mathbf{y}}_j = (y_j^1, y_j^2, \dots, y_j^M) \in R^M, 1 \leq j \leq n\}$$

估计一个变换 f 使得点对应集尽可能地“接近”该变换, 即尽可能地使下述等式成立:

$$\tilde{\mathbf{y}}_j - f(\tilde{\mathbf{x}}_j) = 0, \quad (\tilde{\mathbf{x}}_j, \tilde{\mathbf{y}}_j) \in X \quad (15.1.5)$$

关于变换估计, 我们特别感兴趣的是下述一些特殊变换:

1. 三维射影变换: $f = H = (h_{ij})_{4 \times 4}$, 此时式(15.1.5)转化为:

$$\begin{pmatrix} \tilde{y}_j \\ 1 \end{pmatrix} = s_j H \begin{pmatrix} \tilde{x}_j \\ 1 \end{pmatrix} = 0, (\tilde{x}_j, \tilde{y}_j) \in X$$

其中, s_j 是非零常数因子。令 $\mathbf{x}_j = \begin{pmatrix} \tilde{x}_j \\ 1 \end{pmatrix}, \mathbf{y}_j = \begin{pmatrix} \tilde{y}_j \\ 1 \end{pmatrix}$ 。消去 s_j 后, 上式可写成下面的形式:

$$g(H, \tilde{x}_j, \tilde{y}_j) = \begin{pmatrix} (\mathbf{h}^{1T} \mathbf{x}_j) - y_j^1 (\mathbf{h}^{3T} \mathbf{x}_j) \\ (\mathbf{h}^{2T} \mathbf{x}_j) - y_j^2 (\mathbf{h}^{3T} \mathbf{x}_j) \end{pmatrix} = 0, (\tilde{x}_j, \tilde{y}_j) \in X \quad (15.1.6)$$

如: 在重构中, 将不同视点下重构的三维空间点融合到同一个三维射影空间的问题就涉及到三维射影变换估计, 因为同一物体点在不同的射影坐标系下相差一个射影变换。在基于三维特征的运动分析中, 也同样涉及到三维变换的估计问题。

2. 二维射影变换: $f = H = (h_{ij})_{3 \times 3}$, 此时式(15.1.5)转化为:

$$g(H, \tilde{x}_j, \tilde{y}_j) = \begin{pmatrix} \tilde{y}_j \\ 1 \end{pmatrix} \times H \begin{pmatrix} \tilde{x}_j \\ 1 \end{pmatrix} = 0, (\tilde{x}_j, \tilde{y}_j) \in X \quad (15.1.7)$$

这种类型的估计, 主要出现在由点对应(或线对应)计算二维单应矩阵的问题中。

3. 三维射影空间到二维射影空间的投影变换: $f = P = (p_{ij})_{4 \times 3}$, 类似地, 此时式(5)转化为:

$$g(H, \tilde{x}_j, \tilde{y}_j) = \begin{pmatrix} \tilde{y}_j \\ 1 \end{pmatrix} \times P \begin{pmatrix} \tilde{x}_j \\ 1 \end{pmatrix} = 0, (\tilde{x}_j, \tilde{y}_j) \in X \quad (15.1.8)$$

这种类型的估计, 主要出现在由摄像机矩阵估计问题中。

对于抽象模型, 充当上面各个估计问题的“点对应”角色的数学量, 称为抽象模型的测量向量(或测量数据点)。

15.1.2 模型参数化

模型参数

给定模型 M , 令 $\Phi = \{\mathbf{p} \mid \mathbf{p} = (p_1, p_2, p_3, \dots, p_k) \in R^k\}$ 是 k 维向量集, 如果对每一个 $\mathbf{p} \in \Phi$ 都对应模型 M 的一个解(或实例)。反之, 对每一个模型 M 的解都对应于集 Φ 中的一个向量 \mathbf{p} , 则称 \mathbf{p} 为模型参数向量, 其分量都称为模型的参数, 集 Φ 称为模型的定义域。不难看出, 模型的定义域由模

型参数之间的约束条件所确定。模型的解集合称为模型的值域(或值空间), 对于某个确定的参数向量所对应的解也称为模型的实例。

根据参数向量与模型解之间的对应关系, 模型可由它的参数向量来表达。用参数向量来表达给定的模型, 称为该模型参数化。对于给定的模型, 其参数化不是唯一的, 在所有的参数化中, 具有最小参数数目的参数化称为模型的最小参数化。最小参数化中的参数个数由模型的自由度所确定。

给定模型 M 的参数化 $\Phi = \{p \mid p \in S \subset R^k\}$, 和一组测量向量 $X = \{x_j \mid 1 \leq j \leq n\}$, 如果测量向量是精确的, 则由模型关系可导致关于参数 p 的一组等式:

$$g(p, x_j) = 0, \quad x_j \in X \quad (15.1.10)$$

其中: g 是标量函数也可能是向量值映射, 由模型规则所确定。当测量向量精确时, 可以通过求解方程组(15.1.10)的非零解而得到参数向量 p 的值, 从而获得模型的解。这不是模型估计问题所要研究的对象。模型估计问题是指: 当测量向量不准确(或带有噪声)时, 求解参数向量 p 使得式(15.1.10)“尽可能满足”。

如果 g 关于参数向量 p 是线性的, 则模型 M 关于参数化 Φ 是线性的, 否则称为非线性的。同一个模型关于不同的参数化可能表现出不同性质, 如上一节所提到的所有模型对于给定的参数化都是线性的。但是, 对于基本矩阵关于参数化表示(15.1.13)却是非线性的。三维计算机视觉中的大多数估计问题在适当的过参数化下都是线性的。在本章, 假定所讨论的模型在给定的参数化下都是线性的。

如果模型 M 关于参数化 $\Phi = \{p \mid p \in S \subset R^k\}$ 是线性的, 则式(15.1.10)可以写成如下形式:

$$Mp = 0 \quad (15.1.11)$$

其中矩阵 M 称为测量矩阵。例如: 二次曲线估计问题中, 模型参数为对称矩阵 C 的上三角的 6 个元素构成的向量 c , 则式(15.1.11)可以表示为 $Mc = 0$, 测量矩阵 M 由下式定义:

$$C = \begin{pmatrix} x_1^2 & x_1 y_1 & y_1^2 & x_1 & y_1 & 1 \\ x_2^2 & x_2 y_2 & y_2^2 & x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n^2 & x_n y_n & y_n^2 & x_n & y_n & 1 \end{pmatrix} \quad (15.1.12)$$

最小配置数与最小配置解

确定模型所需要的测量向量的最小数目称为模型的最小配置数, 记为 n 。如果给定的 $s (\geq n)$ 个测量向量不能确定模型参数, 则称这 $s (\geq n)$ 个测量向量是一组退化配置。由 n 个非退化配置的测量

向量所确定的模型解称为最小配置解。以后, 假定所给的测量向量集都是非退化的。最小配置数与模型的自由度以及模型关系对测量向量导致参数向量的约束数有关。

例 15.1.1: 在直线估计中, 模型参数向量是非零三维向量 $l = (a, b, c)$ 。由于非零三维向量 $l = (a, b, c)$ 在相差任一非零常数因子的情况下都确定同一条直线, 因此它的定义域是

$$\Phi = \{(a, b, c) : \|(a, b, c)\|_2^2 = a^2 + b^2 + c^2 = 1\}$$

模型的值空间是平面上所有直线所构成的二维空间。由于平面直线的自由度等于 2, 所以它的最小参数化中的参数数目等于 2。例如, 取 $l = (a, -1, c)$, 则是直线的一种最小参数化, 但在这种最小参数化中, 不包含直线 $y = 0$ (即横坐标轴)。直线的任何一种最小参数化表示, 实际上是直线的一种非齐次表示。

例 15.1.2: 在二维射影变换 H 估计中, 模型参数向量是 3 阶矩阵 H 的 9 个元素, 由于二维射影变换是齐次的, 即在相差任一非零常数因子的情况下都表示同一个射影变换, 并且变换矩阵 H 是一个可逆矩阵, 因此它的定义域是 $\Phi = \{H_{3 \times 3} \mid \|H\|_F = 1, \det(H) \neq 0\}$, 模型的值空间是 3 阶可逆的全体。由于二维射影变换有 8 个自由度, 所以它的最小参数化中的参数数目等于 8。二维射影变换的最小参数化, 也是它的非齐次表示。

例 15.1.3: 在基本矩阵估计中, 与二维射影变换估计类似, 模型参数是 3 阶矩阵的 9 个元素, 与二维射影变换不同的是基本矩阵不是可逆的, 而是一个秩为 2 的矩阵。因此在基本矩阵估计中, 模型的定义域是 $\Phi = \{F_{3 \times 3} \mid \|F\|_F = 1, \det(F) = 0\}$ 。模型的值空间是秩为 2 的 3 阶矩阵的全体。基本矩阵有 7 个自由度, 因此最小参数化中的参数数目等于 7。基本矩阵的最小参数化表示是它的秩 2 参数化的非齐次表示。下面给出它的一种最小参数化表示: 由于基本矩阵 F 的秩为 2, 所以 F 的三个行向量必线性相关。因此, 在一般情况下第三行可由前两行线性表示, 同时第三列也可以由前两列线性表示, 于是基本矩阵可表示为下述形式:

$$F = \begin{pmatrix} a & b & ga + hb \\ c & d & gc + hd \\ ea + fc & eb + fd & g(ea + fc) + h(eb + fd) \end{pmatrix} \quad (15.1.13)$$

其中 $e' = (-e, -f, 1)^T$, $e = (-g, -h, 1)^T$ 是两幅图像的极点坐标。由于基本矩阵是齐次的, 可以置其中的某个元素等于 1, 这样就得到了基本矩阵的最小参数化表示。

15.2 直接线性方法

15.2.1 线性计算框架

令 \mathbf{f} 是所要估计模型 (变换) 的参数向量, 满足约束条件 $\|\mathbf{f}\|=1$ 。给定一组测量向量 $X = \{\mathbf{x}_j | 1 \leq j \leq n\}$, 它所对应的测量矩阵记为 $M_{n \times m}$, 定义:

$$d_L(\mathbf{M}, \mathbf{f}) = \|M_{n \times m} \mathbf{f}\| \quad (15.2.1)$$

为模型的代数距离, 其中 $\|\cdot\|$ 是对应向量空间的欧氏范数。

如果测量向量无误差时, 则 $d_L(\mathbf{M}, \mathbf{f}) = 0$, 因此线性算法是求方程组 $M_{n \times m} \mathbf{f} = 0$ 的非零解。当测量数据存在误差时, $d_L(\mathbf{M}, \mathbf{f}) \neq 0$, 即方程组 $M_{n \times m} \mathbf{f} = 0$ 不可能有非零解, 通常用最小二乘方法估计模型的解, 即最小二乘解。这就是直接线性方法, 有时也称为最小二乘估计。最小二乘估计是求代数距离的平方达到最小的模型参数 \mathbf{f} , 即求解下述最小化问题:

$$\begin{cases} \min \|M_{n \times m} \mathbf{f}\|^2 \\ \text{subject to } \|\mathbf{f}\|=1 \end{cases} \quad (15.2.2)$$

其中: M_{nm} 是由测量数据和所要估计的变换(模型)所确定的矩阵, 通常称为测量矩阵; \mathbf{f} 是所要估计的变换(模型)参数向量, $\|\cdot\|$ 是对应向量空间的欧氏范数。

求解最小化问题(15.2.2), 等价于在约束 $\|\mathbf{f}\|=1$ 下求下述方程组的(欧氏范数)最小二乘解:

$$M_{n \times m} \mathbf{f} = 0 \quad (15.2.3)$$

根据 SVD 分解理论, 最小化问题(15.2.2)的解是矩阵 $M_{n \times m}$ 的最小奇异值的右奇异向量(见 8.4.4 节)。因此, 直接线性方法的数值计算是非常简单的, 可通过 SVD 分解技术来实现。

15.2.2 视觉估计问题

摄像机矩阵

令 $X_j = (\tilde{X}_j, 1)^T \leftrightarrow \mathbf{x}_j = (\tilde{\mathbf{x}}_j^T, 1)^T$ 是 3D 空间到 2D 图像的一组测量点对应, 它们关联于一个摄像机矩阵:

$$P = \begin{pmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \end{pmatrix}$$

即, 当测量数据无误差时, 必有

$$\mathbf{x}_j = \lambda_j P \mathbf{X}_j$$

我们的任务是, 从这组测量点对应估计出摄像机矩阵。记

$$\mathbf{p} = (P_{11}, P_{12}, P_{13}, P_{14}, P_{21}, P_{22}, P_{23}, P_{24}, P_{31}, P_{32}, P_{33}, P_{34})^T,$$

则直接线性方法是在约束 $\|\mathbf{p}\|=1$ 条件下, 求下述方程组的最小二乘解:

$$M_{2n \times 12} \mathbf{p} = \mathbf{0} \quad (15.2.4)$$

其中:

$$M_{2n \times 12} = \begin{pmatrix} \mathbf{X}_1^T & \mathbf{0}^T & -x_1 \mathbf{X}_1^T \\ \mathbf{0}^T & \mathbf{X}_1^T & -y_1 \mathbf{X}_1^T \\ \vdots & \vdots & \vdots \\ \mathbf{X}_n^T & \mathbf{0}^T & -x_n \mathbf{X}_n^T \\ \mathbf{0}^T & \mathbf{X}_n^T & -y_n \mathbf{X}_n^T \end{pmatrix} \quad (15.2.5)$$

单应矩阵

令 $\{\mathbf{x}_j = (\tilde{\mathbf{x}}_j^T, 1)^T \leftrightarrow \mathbf{x}'_j = (\tilde{\mathbf{x}}'_j{}^T, 1)^T, j=1, 2, \dots, n\}$ 是摄像机关于空间平面的两幅图像间的一组点对应, 这组点对应关联于一个单应矩阵:

$$H = \begin{pmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{pmatrix}$$

我们的问题是, 从测量点对应 $\mathbf{x}_j \leftrightarrow \mathbf{x}'_j$ 估计单应矩阵 H 。

记 $\mathbf{h} = (H_{11}, H_{12}, H_{13}, H_{21}, H_{22}, H_{23}, H_{31}, H_{32}, H_{33})^T$, 单应矩阵估计的直接线性方法是在约束条件 $\|\mathbf{h}\|=1$ 下, 求下述方程组的最小二乘解:

$$M_{2n \times 9} \mathbf{h} = \mathbf{0} \quad (15.2.6)$$

其中 $M_{2n \times 9}$ 具有如下形式的测量矩阵:

$$M_{2n \times 9} = \begin{pmatrix} \mathbf{x}_1^T & \mathbf{0}^T & -x'_1 \mathbf{x}_1^T \\ \mathbf{0}^T & \mathbf{x}_1^T & -y'_1 \mathbf{x}_1^T \\ \vdots & \vdots & \vdots \\ \mathbf{x}_n^T & \mathbf{0}^T & -x'_n \mathbf{x}_n^T \\ \mathbf{0}^T & \mathbf{x}_n^T & -y'_n \mathbf{x}_n^T \end{pmatrix} \quad (15.2.7)$$

FOE

当摄像机作纯平移运动时, 运动前、后两幅图像的所有对应点的连线必交于一点, 该点称为 FOE 点。显然, FOE 点是平移运动的两幅图像的极点。我们的任务是从一组给定的测量点对应 $\{\mathbf{x}_j = (\tilde{\mathbf{x}}_j^T, 1)^T \leftrightarrow \mathbf{x}'_j = (\tilde{\mathbf{x}}_j'^T, 1)^T, j=1, 2, \dots, n\}$, 估计 FOE 点。

记 FOE 点为 \mathbf{e} , 直接线性方法是在约束条件 $\|\mathbf{e}\|=1$ 下, 求下述方程组的最小二乘解:

$$M_{n \times 3} \mathbf{e} = \mathbf{0} \quad (15.2.8)$$

其中:

$$M_{n \times 3} = \begin{pmatrix} y_1 - y'_1 & -x_1 + x'_1 & x_1 y'_1 - x'_1 y_1 \\ y_2 - y'_2 & -x_2 + x'_2 & x_2 y'_2 - x'_2 y_2 \\ \vdots & \vdots & \vdots \\ y_n - y'_n & -x_n + x'_n & x_n y'_n - x'_n y_n \end{pmatrix} \quad (15.2.9)$$

基本矩阵

令 $\{\mathbf{x}_j = (\tilde{\mathbf{x}}_j^T, 1)^T \leftrightarrow \mathbf{x}'_j = (\tilde{\mathbf{x}}_j'^T, 1)^T, j=1, 2, \dots, n\}$ 是两幅图像间的一组测量点对应, 根据两幅图像间的极几何, 它们必关联于一个基本矩阵:

$$F = \begin{pmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{pmatrix}$$

即, 当测量数据无误差时, 必有

$$\mathbf{x}'_j{}^T F \mathbf{x}_j = 0$$

我们的任务是从测量点对应估计出基本矩阵。基本矩阵估计的直接线性方法, 先是求解下述最小化问题:

$$\begin{cases} \min_F \sum_{j=1}^n (\mathbf{x}'_j{}^T F \mathbf{x}_j)^2 \\ \text{subject to } \|F\|_F = 1 \end{cases} \quad (15.2.10)$$

然后通过 SVD 分解技术, 以问题(15.2.10)的最小化解在 F-范数意义下的最佳秩 2 矩阵近似作为基本矩阵的估计, 这一步是必要的, 因为基本矩阵是一个秩为 2 的矩阵。记

$$\mathbf{f} = (F_{11}, F_{12}, F_{13}, F_{21}, F_{22}, F_{23}, F_{31}, F_{32}, F_{33})^T,$$

则 8 点算法等价于在约束条件 $\|\mathbf{f}\|=1$ 下, 求下述方程组的最小二乘解:

$$M_{n \times 9} \mathbf{f} = \mathbf{0} \quad (15.2.11)$$

其中 $M_{n \times 9}$ 是测量矩阵:

$$M_{n \times 9} = \begin{pmatrix} x'_1 \mathbf{x}_1^T & y'_1 \mathbf{x}_1^T & \mathbf{x}_1^T \\ x'_2 \mathbf{x}_2^T & y'_2 \mathbf{x}_2^T & \mathbf{x}_2^T \\ \vdots & \vdots & \vdots \\ x'_n \mathbf{x}_n^T & y'_n \mathbf{x}_n^T & \mathbf{x}_n^T \end{pmatrix} \quad (15.2.12)$$

因此, 基本矩阵估计的直接线性方法由下述两个步骤构成:

1. 由 SVD 分解技术, 求解矩阵 $M_{n \times m}$ 的最小奇异值的单位右奇异向量 \mathbf{f}^* ;
2. 由向量 \mathbf{f}^* 得到对应的矩阵 F^* , 并对 F^* 进行 SVD 分解 $F^* = U \text{diag}(\sigma_1, \sigma_2, \sigma_3) V$, 得到基本矩阵的估计: $F = U \text{diag}(\sigma_1, \sigma_2, 0) V$ 。

三焦张量

三幅视图之间的对应点以及对应线之间满足三线性关系, 并且这种三线性关系囊括在三焦张量之中。令 $\{x^i \leftrightarrow x'^j \leftrightarrow x''^k\}$ 是三幅图像的一个三点对应, 则有三焦张量 t_i^{qr} 的线性约束关系^[1]:

$$x^i x'^j x''^k \varepsilon_{jqu} \varepsilon_{krv} t_i^{qr} = 0_{uv} \quad (15.2.13)$$

其中张量 ε_{jqu} 的定义如下:

$$\varepsilon_{jqu} = \begin{cases} 0, & \text{若 } j, q, u \text{ 不相异,} \\ +1, & \text{若 } jqu \text{ 是 } 123 \text{ 的偶排列,} \\ -1, & \text{若 } jqu \text{ 是 } 123 \text{ 的奇排列.} \end{cases}$$

在等式(15.2.13)中每一个自由指标 $u(v)$ 都有 3 种不同的选择, 所以一共可得到 9 个三线性关系式, 但是在这 9 个三线性关系式中仅有 4 个是关于 t_i^{qr} 独立的。

令 $\mathbf{x} = (\tilde{\mathbf{x}}^T, 1)^T \leftrightarrow \mathbf{x}' = (\tilde{\mathbf{x}}'^T, 1)^T \leftrightarrow \mathbf{x}'' = (\tilde{\mathbf{x}}''^T, 1)^T$ 是三幅视图之间的一组点对应, 则由方程(15.2.13)可以得到如下四个独立的三线性关系式:

$$\begin{aligned} x(x''t_1^{13} - x''x't_1^{33} + x't_1^{31} - t_1^{11}) + y(x''t_2^{13} - x''x't_2^{33} + x't_2^{31} - t_2^{11}) + (x''t_3^{13} - x''x't_3^{33} + x't_3^{31} - t_3^{11}) &= 0 \\ x(y''t_1^{13} - y''x't_1^{33} + x't_1^{32} - t_1^{12}) + y(y''t_2^{13} - y''x't_2^{33} + x't_2^{32} - t_2^{12}) + (y''t_3^{13} - y''x't_3^{33} + x't_3^{32} - t_3^{12}) &= 0 \\ x(x''t_1^{23} - x''x't_1^{33} + y't_1^{31} - t_1^{21}) + y(x''t_1^{23} - x''x't_1^{33} + y't_1^{31} - t_1^{21}) + (x''t_1^{23} - x''x't_1^{33} + y't_1^{31} - t_1^{21}) &= 0 \\ x(y''t_1^{23} - y''y't_1^{33} + y't_1^{32} - t_1^{22}) + y(y''t_1^{23} - y''y't_1^{33} + y't_1^{32} - t_1^{22}) + (y''t_1^{23} - y''y't_1^{33} + y't_1^{32} - t_1^{22}) &= 0 \end{aligned}$$

令

$$M_{4 \times 27} = \begin{pmatrix} xC_1 & yC_1 & C_1 \\ xC_2 & yC_2 & C_2 \\ xC_3 & yC_3 & C_3 \\ xC_4 & yC_4 & C_4 \end{pmatrix} \quad (15.2.14)$$

其中:

$$C_1 = (-1 \ 0 \ x'' \ 0 \ 0 \ 0 \ x' \ 0 \ -x'y''), \quad C_2 = (0 \ -1 \ y'' \ 0 \ 0 \ 0 \ 0 \ x' \ -x'y''), \\ C_3 = (0 \ 0 \ 0 \ -1 \ 0 \ x'' \ y' \ 0 \ -y'x''), \quad C_4 = (0 \ 0 \ 0 \ 0 \ -1 \ y'' \ 0 \ y' \ -y'y''),$$

则上述四个独立的三线性关系可以写成如下的矩阵形式

$$M_{4 \times 27} \mathbf{t}_{27 \times 1} = \mathbf{0}_{4 \times 1}$$

其中: $\mathbf{t} = (t_1^{11}, t_1^{12}, t_1^{13}, t_1^{21}, t_1^{22}, t_1^{23}, t_1^{31}, t_1^{32}, t_1^{33}, t_2^{11}, t_2^{12}, \dots, t_3^{31}, t_3^{32}, t_3^{33})^T$ 。因此, 给定三幅视图的 n 组点对应, 可得到 $4n$ 个关于三焦张量独立的线性方程:

$$M_{4n \times 27} \mathbf{t} = \mathbf{0} \quad (15.2.15)$$

直接线性算法就是在约束条件 $\|\mathbf{t}\|=1$ 下求方程组(15.2.15)的最小二乘解。由于三焦张量共有 27 个参数, 至少需要三幅图像中的 7 组对应点才能在相差一个公共因子的意义下确定三焦张量。

15.3 因子化线性方法

直接线性方法的主要不足之处是测量数据的较小误差可能会导致较坏的估计结果。发生这种情况的主要原因有: (1) 测量矩阵的元素是测量数据的非线性函数(参见式(15.2.5), (15.2.7), (15.2.9), (15.2.12)), 从而放大了测量误差, 使得测量矩阵严重地偏了真实矩阵, 因此导致直接线性方法的估计结果与真实值有很大偏差。(2) 测量矩阵可能有较大的条件数, 导致估计结果的不稳定性。

本节所介绍的因子化线性方法, 旨在降低测量矩阵元素关于测量数据的非线性, 以抑制测量误差对估计结果的影响, 从而克服直接线性方法的第一个缺点。下一节将克服直接线性方法第二个缺点的归一化线性方法。

15.3.1 因子化计算框架

如果测量矩阵 $M_{n \times m}$ 中的元素是测量数据的多重线性函数, 则称相应的估计问题为多重线性估计问题, 如: 估计基本矩阵, 摄像机矩阵和单应矩阵等为二重线性估计问题; 估计三焦张量则是三重

线性估计问题。因子化线性方法是针对多重线性估计问题而设计的。

因子化线性方法不是根据直接线性方法的测量矩阵直接求解估计问题, 而是先将直接线性方法的测量矩阵分解为若干个因子矩阵的乘积, 使得因子矩阵的元素是某一幅图像的测量数据值或者是常量; 然后引进中间变量构造新测量矩阵, 并使得新测量矩阵的元素仅是某个测量数据值或者是常量; 最后求解相应线性问题的最小二乘解。因子矩阵的个数等于线性估计问题的重数, 如在估计基本矩阵因子化方法中, 有 2 个因子矩阵; 对于三焦张量, 则有 3 个因子矩阵。因子化线性方法不仅仅是一种线性方法, 更重要的是它能降低测量矩阵元素的关于测量数据的非线性, 从而有效地抑制了测量误差对估计结果的影响。具体地说, 一般估计问题的因子化方法主要由下述三个步骤所构成:

1. 分解直接线性方法中的测量矩阵 $M_{n \times m}$ 为下述形式:

$$M_{n \times m} = A^{(1)} A^{(2)} \dots A^{(k)} \quad (15.3.1)$$

其中: $A^{(j)}$ 的每一个元素或者是某个测量数据值或者是一个常量;

2. 引进中间变量: $f^{(1)} = A^{(k)} f, f^{(2)} = A^{(k-1)} f^{(1)}, \dots, f^{(k-1)} = A^{(2)} f^{(k-2)}$, 并构造新测量矩阵:

$$M_{new} = \begin{pmatrix} A^{(k)} & -I & & & & \\ & A^{(k-1)} & -I & & & \\ & & A^{(k-2)} & \ddots & & \\ & & & \ddots & -I & \\ & & & & & A^{(1)} \end{pmatrix} \quad (15.3.2)$$

3. 利用 SVD 分解技术, 在约束条件 $\|\tilde{f}\|=1$ 下, 求下述线性方程组的(欧氏范数)最小二乘解:

$$M_{new} \tilde{f} = 0 \quad (\tilde{f} = (f, f^{(1)}, \dots, f^{(k-1)})^T) \quad (15.3.3)$$

在测量数据无误差时, 由方程组(15.3.3)的最小二乘解 f^* 与方程组(15.2.3)的最小二乘解相差一个非零因子, 即两种方法的估计结果是一致的。在测量数据出现较小误差时, 测量矩阵 $M_{n \times m}$ 可能严重地偏离真实系数矩阵, 而测量矩阵 M_{new} 仍接近对应的原始真实系数矩阵, 这是为什么因子化线性方法要优于直接算法的原因。

15.3.2 视觉估计问题

摄像机矩阵

1. 将测量矩阵(15.2.5)分解为下述形式:

$$M_{2n \times 12} = \begin{pmatrix} \mathbf{X}_1^T & & & \\ & \mathbf{X}_1^T & & \\ & & \ddots & \\ & & & \mathbf{X}_n^T \\ & & & & \mathbf{X}_n^T \end{pmatrix} \begin{pmatrix} I_4 & \mathbf{0} & -x_1 I_4 \\ \mathbf{0} & I_4 & -y_1 I_4 \\ \vdots & \vdots & \vdots \\ I_4 & \mathbf{0} & -x_n I_4 \\ \mathbf{0} & I_4 & -y_n I_4 \end{pmatrix} = A_{2n \times 8n} B_{8n \times 12} \quad (15.3.4)$$

2. 引进中间变量 $\mathbf{g} = B_{8n \times 12} \mathbf{p}$ ，构造新测量矩阵：

$$M_{new} = \begin{pmatrix} B_{8n \times 12} & -I_{8n} \\ \mathbf{0}_{2n \times 12} & A_{2n \times 8n} \end{pmatrix} \quad (15.3.5)$$

3. 在约束条件 $\|\tilde{\mathbf{p}}\|=1$ 下，求下述方程组的最小二乘解：

$$M_{new} \tilde{\mathbf{p}} = \mathbf{0} \quad \left(\tilde{\mathbf{p}} = \begin{pmatrix} \mathbf{p} \\ \mathbf{g} \end{pmatrix} \right) \quad (15.3.6)$$

中间变量的几何意义 不难计算

$$\mathbf{g}_j = \begin{pmatrix} g_{8(j-1)+1} \\ g_{8(j-1)+2} \\ \vdots \\ g_{8(j-1)+8} \end{pmatrix} = \begin{pmatrix} I_4 & \mathbf{0} & -x_j I_4 \\ \mathbf{0} & I_4 & -y_j I_4 \end{pmatrix} \mathbf{p} = \begin{pmatrix} \mathbf{p}_{(1)}^T - x_j \mathbf{p}_{(3)}^T \\ \mathbf{p}_{(2)}^T - y_j \mathbf{p}_{(3)}^T \end{pmatrix} \quad (15.3.7)$$

其中 $\mathbf{p}_{(j)}$ 是摄像机矩阵 P 的第 j 行所构成的行向量。由于空间平面 $\mathbf{p}_{(1)}^T - x_j \mathbf{p}_{(3)}^T$ 是过图像点 \mathbf{x}_j 的铅直线的反投影平面，而 $\mathbf{p}_{(2)}^T - y_j \mathbf{p}_{(3)}^T$ 是过图像点 \mathbf{x}_j 的水平直线的反投影平面，所以 \mathbf{g}_j 是摄像机中心与图像点 \mathbf{x}_j 的连线，如图 15.3.1 所示。综上所述，当测量数据无误差时，必有

$$\begin{pmatrix} \mathbf{X}_j^T \\ \mathbf{X}_j^T \end{pmatrix} \mathbf{g}_j = \mathbf{0}。$$

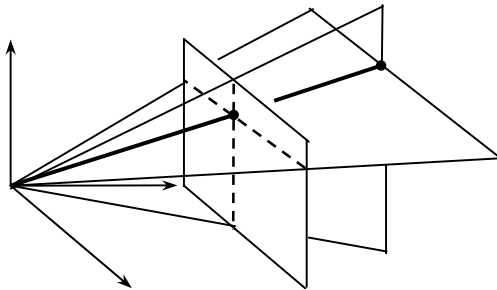


图 15.3.1: 中间变量的几何意义(摄像机矩阵)

单应矩阵

1. 将测量矩阵(15.2.7)分解成下述形式：

$$M_{2n \times 9} = \begin{pmatrix} \mathbf{x}_1^T & & & & & & & & \\ & \mathbf{x}_1^T & & & & & & & \\ & & \ddots & & & & & & \\ & & & & \mathbf{x}_n^T & & & & \\ & & & & & & & & \mathbf{x}_n^T \end{pmatrix} \begin{pmatrix} I_3 & \mathbf{0} & -x'_1 I_3 \\ \mathbf{0} & I_3 & -y'_1 I_3 \\ \vdots & \vdots & \vdots \\ I_3 & \mathbf{0} & -x'_n I_3 \\ \mathbf{0} & I_3 & -y'_n I_3 \end{pmatrix} = A_{2n \times 6n} B_{6n \times 9} \quad (15.3.8)$$

2. 引进中间变量 $\mathbf{g} = B_{6n \times 9} \mathbf{h}$ ，构造新测量矩阵：

$$M_{new} = \begin{pmatrix} B_{6n \times 9} & -I_{6n} \\ \mathbf{0}_{2n \times 9} & A_{2n \times 6n} \end{pmatrix} \quad (15.3.9)$$

3. 在约束条件 $\|\tilde{\mathbf{h}}\|=1$ 下，求下述方程组的最小二乘解：

$$M_{new} \tilde{\mathbf{h}} = \mathbf{0} \quad (\tilde{\mathbf{h}} = \begin{pmatrix} \mathbf{h} \\ \mathbf{g} \end{pmatrix}) \quad (15.3.10)$$

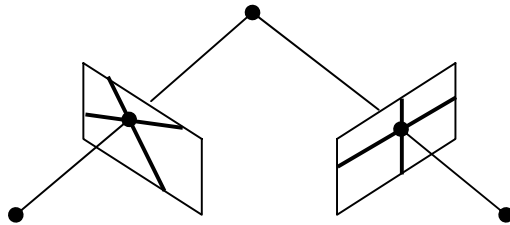


图 15.3.2: 中间变量的几何意义(单应矩阵)

中间变量的几何意义 由于

$$\begin{pmatrix} \mathbf{g}_{2(j-1)+1} \\ \mathbf{g}_{2(j-1)+2} \end{pmatrix} = \begin{pmatrix} I_3 & \mathbf{0} & -x'_j I_3 \\ \mathbf{0} & I_3 & -y'_j I_3 \end{pmatrix} \mathbf{h} = \begin{pmatrix} \mathbf{h}_{(1)}^T - x'_j \mathbf{h}_{(3)}^T \\ \mathbf{h}_{(2)}^T - y'_j \mathbf{h}_{(3)}^T \end{pmatrix} = \begin{pmatrix} 1 & 0 & -x'_j \\ 0 & -y'_j & 1 \end{pmatrix} H$$

而矩阵

$$\begin{pmatrix} 1 & 0 & -x'_j \\ 0 & -y'_j & 1 \end{pmatrix}$$

的两行分别是通过点 \mathbf{x}'_j 的铅直线与水平线，所以 $\mathbf{g}_{2(j-1)+1}, \mathbf{g}_{2(j-1)+2}$ 分别是过点 \mathbf{x}'_j 的铅直线与水平线

在第一幅图像上的对应直线(如图 15.3.2 所示)。如果点对应 $\mathbf{x}_j \leftrightarrow \mathbf{x}'_j$ 无误差的话, 则 $\mathbf{g}_{2(j-1)+1} \cdot \mathbf{g}_{2(j-1)+1}$ 必相交于点 \mathbf{x}_j , 即

$$\begin{pmatrix} \mathbf{x}_j^T & \mathbf{0}^T \\ \mathbf{0}^T & \mathbf{x}_j^T \end{pmatrix} \begin{pmatrix} \mathbf{g}_{2(j-1)+1} \\ \mathbf{g}_{2(j-1)+1} \end{pmatrix} = \mathbf{0}。$$

基本矩阵

1. 将测量矩阵(15.2.12)分解成:

$$M_{n \times 9} = A_{n \times 3n} B_{3n \times 9}, \tag{15.3.11}$$

其中:

$$A_{n \times 3n} = \begin{pmatrix} \mathbf{x}_1^{iT} & 0 & 0 & 0 & 0 \\ 0 & \mathbf{x}_2^{iT} & 0 & 0 & 0 \\ 0 & 0 & \mathbf{x}_3^{iT} & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \mathbf{x}_n^{iT} \end{pmatrix}, \quad B_{3n \times 9} = \begin{pmatrix} \mathbf{x}_1^T & 0 & 0 \\ 0 & \mathbf{x}_1^T & 0 \\ 0 & 0 & \mathbf{x}_1^T \\ \mathbf{x}_2^T & 0 & 0 \\ 0 & \mathbf{x}_2^T & 0 \\ 0 & 0 & \mathbf{x}_2^T \\ \vdots & \vdots & \vdots \\ \mathbf{x}_n^T & 0 & 0 \\ 0 & \mathbf{x}_n^T & 0 \\ 0 & 0 & \mathbf{x}_n^T \end{pmatrix};$$

2. 引进中间变量 $\mathbf{g} = B_{3n \times 9} \mathbf{f}$, 构造新测量矩阵:

$$M_{new} = \begin{pmatrix} B_{3n \times 9} & -I_{3n} \\ \mathbf{0}_{n \times 9} & A_{n \times 3n} \end{pmatrix} \tag{15.3.12}$$

3. 由 SVD 分解技术, 在约束条件下 $\|\tilde{\mathbf{f}}\|=1$ 下, 求下述方程组的最小二乘解:

$$M_{new} \tilde{\mathbf{f}} = \mathbf{0} \quad (\tilde{\mathbf{f}} = \begin{pmatrix} \mathbf{f} \\ \mathbf{g} \end{pmatrix}) \tag{15.3.13}$$

即, 方程组(15.3.13)的最小二乘解 $\tilde{\mathbf{f}}^* = \begin{pmatrix} \mathbf{f}^* \\ \mathbf{g}^* \end{pmatrix}$ 是测量矩阵 M_{new} 的最小奇异值的右奇异向量;

4. 由向量 \mathbf{f}^* 得到对应的矩阵 F^* , 并对 F^* 进行奇异值分解: $F^* = U \text{diag}(\sigma_1, \sigma_2, \sigma_3) V$, 最后得到基本矩阵的估计: $F = U \text{diag}(\sigma_1, \sigma_2, 0) V$ 。

中间变量的几何意义 记

$$\mathbf{g} = B_{3n \times 9} \mathbf{f} = \begin{pmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \\ \vdots \\ \mathbf{g}_n \end{pmatrix}, \quad \mathbf{g}_j = \begin{pmatrix} g_{3(j-1)+1} \\ g_{3(j-1)+2} \\ g_{3(j-1)+3} \end{pmatrix}, \quad j = 1, 2, \dots, n$$

由于,

$$\begin{pmatrix} g_{3(j-1)+1} \\ g_{3(j-1)+2} \\ g_{3(j-1)+3} \end{pmatrix} = \begin{pmatrix} \mathbf{x}_j^T & & \\ & \mathbf{x}_j^T & \\ & & \mathbf{x}_j^T \end{pmatrix} \mathbf{f} = F \mathbf{x}_j, \quad (15.3.14)$$

所以, 当数据无误差时, 中间变量 \mathbf{g} 是第一幅图像点在第二幅图像上的对应极线。

最小代数残差 记 \mathbf{f}_{L_s} 是方程组 (15.2.3) 在约束条件 $\|\mathbf{f}\|=1$ 下的最小二乘解, 对应的矩阵记为 F_{L_s} , 则方程组 (15.2.3) 的最小代数残差为:

$$r_{L_s} = \|M_{n \times 9} \mathbf{f}_{L_s}\|^2 = \sum_{j=1}^n (\mathbf{x}_j^T F_{L_s} \mathbf{x}_j)^2$$

令 $\hat{\mathbf{g}} = B_{3n \times 9} \mathbf{f}_{L_s}$, $\tilde{\mathbf{f}} = \frac{1}{\sqrt{1 + \|\hat{\mathbf{g}}\|^2}} \begin{pmatrix} \mathbf{f}_{L_s} \\ \hat{\mathbf{g}} \end{pmatrix}$, 则 $\|\tilde{\mathbf{f}}\|=1$, 且

$$\begin{aligned} \|M_{new} \tilde{\mathbf{f}}\|^2 &= \frac{1}{1 + \|\hat{\mathbf{g}}\|^2} \left\| \begin{pmatrix} B_{3n \times 9} & -I_{3n} \\ \mathbf{0}_{n \times 9} & A_{n \times 3n} \end{pmatrix} \begin{pmatrix} \mathbf{f}_{L_s} \\ \hat{\mathbf{g}} \end{pmatrix} \right\|^2 \\ &= \frac{1}{1 + \|\hat{\mathbf{g}}\|^2} (\|B_{3n \times 9} \mathbf{f}_{L_s} - \hat{\mathbf{g}}\|^2 + \|A_{n \times 3n} \hat{\mathbf{g}}\|^2) \\ &= \frac{1}{1 + \|\hat{\mathbf{g}}\|^2} (\|A_{n \times 3n} B_{3n \times 9} \mathbf{f}_{L_s}\|^2) \\ &= \frac{1}{1 + \|\hat{\mathbf{g}}\|^2} \sum_{j=1}^n (\mathbf{x}_j^T F_{L_s} \mathbf{x}_j)^2 < r_{L_s} \end{aligned}$$

所以, 有

$$\tilde{r}_{L_s} = \min_{\|\tilde{\mathbf{f}}\|=1} \|M_{new} \tilde{\mathbf{f}}\|^2 < r_{L_s} \quad (15.3.15)$$

这说明, 方程组 (15.3.13) 的最小代数残差总小于方程组 (15.2.3) 的最小代数残差。

三焦张量

按照因子化计算框架, 首先需要将三焦张量测量矩阵 $M_{4n \times 27}$ 分解为三个因子的乘积, 使得每个因子矩阵的每一元素是某一幅图像的测量数据或者是一个常量。令

$$U' = (1 \ 0 \ -x')^T, V' = (0 \ 1 \ -y')^T, U'' = (-1 \ 0 \ x'')^T, V'' = (0 \ -1 \ y'')^T$$

显然 U' 和 V' 分别是第二幅视图中过点 X' 的铅直线和水平线; U'' 和 V'' 分别是第三幅视图中过点 X'' 的铅直线和水平线。不难看出:

$$\begin{aligned} C_1 &= (-1 \ 0 \ x'' \ 0 \ 0 \ 0 \ x' \ 0 \ -x'x'') = U'^T \otimes U''^T, \\ C_2 &= [0 \ -1 \ y'' \ 0 \ 0 \ 0 \ 0 \ x' \ -x'y''] = U'^T \otimes V''^T, \\ C_3 &= [0 \ 0 \ 0 \ -1 \ 0 \ x'' \ y' \ 0 \ -y'x''] = V'^T \otimes U''^T, \\ C_4 &= [0 \ 0 \ 0 \ 0 \ -1 \ y'' \ 0 \ y' \ -y'y''] = V'^T \otimes V''^T \text{ *).} \end{aligned}$$

因此, (15.2.14)式可以写成:

$$M_{4 \times 27} = \begin{pmatrix} X^T \otimes U'^T \otimes U''^T \\ X^T \otimes U'^T \otimes V''^T \\ X^T \otimes V'^T \otimes U''^T \\ X^T \otimes V'^T \otimes V''^T \end{pmatrix} = (I_4 \otimes X^T) \begin{pmatrix} I_3 \otimes U'^T \otimes U''^T \\ I_3 \otimes U'^T \otimes V''^T \\ I_3 \otimes V'^T \otimes U''^T \\ I_3 \otimes V'^T \otimes V''^T \end{pmatrix}$$

由于

$$\begin{pmatrix} I_3 \otimes U'^T \otimes U''^T \\ I_3 \otimes U'^T \otimes V''^T \end{pmatrix} = (I_2 \otimes I_3 \otimes U'^T) \begin{pmatrix} I_3 \otimes U''^T \\ I_3 \otimes V''^T \end{pmatrix} = (I_6 \otimes U'^T) \begin{pmatrix} I_3 \otimes U''^T \\ I_3 \otimes V''^T \end{pmatrix}$$

和

$$\begin{pmatrix} I_3 \otimes V'^T \otimes U''^T \\ I_3 \otimes V'^T \otimes V''^T \end{pmatrix} = (I_2 \otimes I_3 \otimes V'^T) \begin{pmatrix} I_3 \otimes U''^T \\ I_3 \otimes V''^T \end{pmatrix} = (I_6 \otimes V'^T) \begin{pmatrix} I_3 \otimes U''^T \\ I_3 \otimes V''^T \end{pmatrix}$$

所以, 可得到

$$M_{4 \times 27} = (I_4 \otimes X^T) \begin{pmatrix} I_6 \otimes U'^T \\ I_6 \otimes V'^T \end{pmatrix} \begin{pmatrix} I_3 \otimes U''^T \\ I_3 \otimes V''^T \end{pmatrix} \quad (15.3.16)$$

因此, 给定三幅图像的 n 组点对应 $\{X_j \leftrightarrow X'_j \leftrightarrow X''_j\}$, 相应的三焦张量测量矩阵 $M_{4n \times 27}$ 能够被分解为:

$$M_{4n \times 27} = P_{4n \times 12n} Q_{12n \times 18n} L_{18n \times 27} \quad (15.3.17)$$

其中

$$P_{4n \times 12n} = \text{diag}(I_4 \otimes X_1^T, I_4 \otimes X_2^T, \dots, I_4 \otimes X_n^T),$$

*) \otimes 表示矩阵的直积, 见 9.4 节。

$$Q_{12n \times 18n} = \text{diag} \left(\begin{pmatrix} I_6 \otimes U_1^{nT} \\ I_6 \otimes V_1^{nT} \end{pmatrix}, \begin{pmatrix} I_6 \otimes U_2^{nT} \\ I_6 \otimes V_2^{nT} \end{pmatrix}, \dots, \begin{pmatrix} I_6 \otimes U_n^{nT} \\ I_6 \otimes V_n^{nT} \end{pmatrix} \right),$$

$$L_{18n \times 27} = \begin{pmatrix} \begin{pmatrix} I_9 \otimes U_1^{nT} \\ I_9 \otimes V_1^{nT} \end{pmatrix} \\ \begin{pmatrix} I_9 \otimes U_2^{nT} \\ I_9 \otimes V_2^{nT} \end{pmatrix} \\ \vdots \\ \begin{pmatrix} I_9 \otimes U_n^{nT} \\ I_9 \otimes V_n^{nT} \end{pmatrix} \end{pmatrix}$$

显然，因子矩阵 $P_{4n \times 12n}$ 、 $Q_{12n \times 18n}$ 和 $L_{18n \times 27}$ 的元素是图像的测量数据或者是一个常量，并且 $P_{4n \times 12n}$ ($Q_{12n \times 18n}$, $L_{18n \times 27}$) 分别仅与第一(二、三)幅视图相关。

引进中间变量 $\mathbf{h} = L_{18n \times 27} \mathbf{t}$ 和 $\mathbf{l} = Q_{12n \times 18n} \mathbf{h}$ ，构造新的测量矩阵：

$$\tilde{M}_{34n \times (30n+27)} = \begin{pmatrix} L_{18n \times 27} & -I_{18n \times 18n} & \\ & Q_{12n \times 18n} & -I_{12n \times 12n} \\ & & P_{4n \times 12n} \end{pmatrix} \quad (15.3.18)$$

则因子化算法是在约束条件 $\|\tilde{\mathbf{t}}\|=1$ 下，求下述方程组的最小二乘解：

$$\tilde{M}_{34n \times (30n+27)} \tilde{\mathbf{t}} = \mathbf{0} \quad (\tilde{\mathbf{t}} = (\mathbf{t}^T, \mathbf{h}^T, \mathbf{l}^T)^T) \quad (15.3.19)$$

因此通过对矩阵 $\tilde{M}_{34n \times (30n+27)}$ 进行 SVD 分解，在相差一个比例因子意义下得到三焦张量 t_i^{qr} 。

中间变量的几何意义 令 $\mathbf{h} = L_{18n \times 27} \mathbf{t} = (\mathbf{h}_1^T \quad \mathbf{h}_2^T \quad \dots \quad \mathbf{h}_n^T)^T$ ， $\mathbf{h}_j \in R^{18}, j=1,2,\dots,n$ ，则有

$$\mathbf{h}_j = \begin{pmatrix} (I_9 \otimes U_j^{nT}) \mathbf{t} \\ (I_9 \otimes V_j^{nT}) \mathbf{t} \end{pmatrix} = \begin{pmatrix} (I_3 \otimes U_j^{nT}) \mathbf{t}_1 \\ (I_3 \otimes U_j^{nT}) \mathbf{t}_2 \\ (I_3 \otimes U_j^{nT}) \mathbf{t}_3 \\ (I_3 \otimes V_j^{nT}) \mathbf{t}_1 \\ (I_3 \otimes V_j^{nT}) \mathbf{t}_2 \\ (I_3 \otimes V_j^{nT}) \mathbf{t}_3 \end{pmatrix} = \begin{pmatrix} (T_1 U_j^n) \\ (T_2 U_j^n) \\ (T_3 U_j^n) \\ (T_1 V_j^n) \\ (T_2 V_j^n) \\ (T_3 V_j^n) \end{pmatrix} \triangleq \begin{pmatrix} \mathbf{h}_{j1} \\ \mathbf{h}_{j2} \end{pmatrix} \quad (15.3.20)$$

其中 $\mathbf{t}_i = [t_i^{11}, t_i^{12}, t_i^{13}, t_i^{21}, t_i^{22}, t_i^{23}, t_i^{31}, t_i^{32}, t_i^{33}]^T, i=1,2,3$, $T_i = \begin{pmatrix} t_i^{11} & t_i^{12} & t_i^{13} \\ t_i^{21} & t_i^{22} & t_i^{23} \\ t_i^{31} & t_i^{32} & t_i^{33} \end{pmatrix}, i=1,2,3$ 。显然，向量 \mathbf{h}_j 是由第三幅视图

中过点 X_j'' 的铅直线 U_j'' 和水平线 V_j'' 的两个反向投影平面诱导的单元 h_{j1} 和 h_{j2} 所构成的, $j=1,2,\dots,n$ 。

令 $l = Q_{12n \times 18n} h = (l_1^T \ l_2^T \ \dots \ l_n^T)^T, l_j \in R^{12}, j=1,2,\dots,n$, 则

$$l_j = \begin{pmatrix} I_6 \otimes U_j'' \\ I_6 \otimes V_j'' \end{pmatrix} \begin{pmatrix} h_{j1} \\ h_{j2} \end{pmatrix} = \begin{pmatrix} (I_3 \otimes U_j'') h_{j1} \\ (I_3 \otimes U_j'') h_{j2} \\ (I_3 \otimes V_j'') h_{j1} \\ (I_3 \otimes V_j'') h_{j2} \end{pmatrix}, j=1,2,\dots,n. \quad (15.3.21)$$

因此, 向量 l_j 是由第二幅视图中过点 X_j'' 的铅直线 U_j'' 和水平线 V_j'' 在单元 h_{j1} 和 h_{j2} 作用下的像直线构成的, 其中 $j=1,2,\dots,n$ 。

最小代数残差 记 t_{LS} 是方程组(15.3.19)在约束条件 $\|t\|=1$ 下的最小二乘解, 则方程组(15.3.1,9)的最小代数残差为:

$$r_{LS} = \|M_{4n \times 27} t_{LS}\|^2$$

令 $\hat{h} = L_{18n \times 27} t_{LS}, \hat{l} = Q_{12n \times 18n} \hat{h}, \tilde{t} = \frac{1}{\sqrt{1 + \|\hat{h}\|^2 + \|\hat{l}\|^2}} \begin{pmatrix} t_{LS} \\ \hat{h} \\ \hat{l} \end{pmatrix}$, 则 $\|\tilde{t}\|=1$, 并且有

$$\begin{aligned} \|\tilde{M}_{34n \times (30n+27)} \tilde{t}\|^2 &= \frac{1}{1 + \|\hat{h}\|^2 + \|\hat{l}\|^2} \left\| \begin{pmatrix} L_{18n \times 27} & -I_{18n \times 18n} \\ Q_{12n \times 18n} & -I_{12n \times 12n} \\ P_{4n \times 12n} & \end{pmatrix} \begin{pmatrix} t_{LS} \\ \hat{h} \\ \hat{l} \end{pmatrix} \right\|^2 \\ &= \frac{1}{1 + \|\hat{h}\|^2 + \|\hat{l}\|^2} (\|L_{18n \times 27} t_{LS} - \hat{h}\|^2 + \|Q_{12n \times 18n} \hat{h} - \hat{l}\|^2 + \|P_{4n \times 12n} \hat{l}\|^2) \\ &\leq \frac{1}{1 + \|\hat{h}\|^2 + \|\hat{l}\|^2} \|P_{4n \times 12n} \hat{l}\|^2 \\ &= \frac{1}{1 + \|\hat{h}\|^2 + \|\hat{l}\|^2} \|P_{4n \times 12n} Q_{12n \times 18n} L_{18n \times 27} t_{LS}\|^2 \\ &= \frac{1}{1 + \|\hat{h}\|^2 + \|\hat{l}\|^2} \|M_{4n \times 27} t_{LS}\|^2 < r_{LS} \end{aligned}$$

所以, $\tilde{r}_{LS} = \min_{\|\tilde{t}\|=1} \|\tilde{M}_{34n \times (30n+27)} \tilde{t}\|^2 < r_{LS}$, 这说明方程组(15.3.19)的最小代数残差总小于方程组(15.2.15)的代数残差。

15.4 归一化线性方法

对于直接线性方法的第二个缺点, Hartley 就基本矩阵估计问题提出了归一化线性方法, 试图改善测量矩阵的条件数。尽管归一化线性方法是针对基本矩阵估计问题而提出的, 但其基本原理几乎适用于所在的线性估计问题。归一化线性方法在某种程度上提高了直接线性方法的性能。

归一化线性方法的基本思想是对测量数据作适当的归一化变换, 使得变换后的数据所对应的测量矩阵有好的条件数, 以提高数值计算的稳定性。数据的归一化变换有很多种方法, Hartley 所建议的方法是最简单的, 它由数据的位移变换和缩放变换所成。具体地说, 归一化分为以下两步:

- **位移变换:** 对测量数据作坐标位移变换, 使得原始数据的重心是新坐标系的原点;
- **伸缩变换:** 再对数据作坐标缩放变换, 使得位移变换后的数据点分布在以重心为圆心半径为 \sqrt{n} 的球内, 其中 n 是测量数据的维数。

归一化线性方法

1. 对原始测量数据作归一化变换, 并由变换后数据构造测量矩阵 M ;
2. 由直接线性方法求解对应测量矩阵 M 的估计问题;
3. 解除归一化变换得到对应原测量数据的模型 (变换) 的估计结果。

由于归一化变换是直接对测量数据的操作, 所以在归一化变换同时也增加了对测量误差的额外操作。如果归一化变换不能在本质上改善测量矩阵的条件数, 或者测量矩阵本身是良态的, 归一化变换不仅不能为估计问题带来任何好处, 反而是有害的。这是在实践中常会出现归一化线性方法算法不如直接线性方法的原因。也就是说, 在测量误差对估计问题的影响大于测量矩阵条件数的影响时, 归一化方法就不再适用。

下面给出基本矩阵估计问题的归一化线性方法, 读者不难推广到其它的估计问题中。

假定 $\{\mathbf{x}_j \leftrightarrow \mathbf{x}'_j\}$ 是一组测量点对应。

1. 构造归一化变换:

对于第一幅图像, 令

$$\bar{x} = \frac{1}{N} \sum_{j=1}^N x_j, \bar{y} = \frac{1}{N} \sum_{j=1}^N y_j, d_x = \frac{1}{\max\{|x_j - \bar{x}|\}}, d_y = \frac{1}{\max\{|y_j - \bar{y}|\}},$$

$$\text{构造矩阵 } H = \begin{pmatrix} d_x & -\bar{x} \\ & d_y & -\bar{y} \\ & & 1 \end{pmatrix};$$

对于第二幅图像, 用类似的方法构造矩阵 H' 。

2. 两幅图像分别作坐标变换:

$\hat{x}_j = Hx_j, \hat{x}'_j = H'x'_j$, 产生一组新的点对应 $\{\hat{x}_j \leftrightarrow \hat{x}'_j\}$, 并由直接线性方法求解点对应 $\{\hat{x}_j \leftrightarrow \hat{x}'_j\}$ 所对应的基本矩阵 \hat{F} 。

3. 解除归一化变换, 得到原始点对应的基本矩阵: $F = H'^T \hat{F} H$ 。

归一化仅是对图像数据预处理操作, 它可应用于任何线性估计算法之前。如果在因子化线性算法之前执行图像数据的归一化, 则可以得到归一因子化线性算法。在实际问题中, 归一因子化线性算法的估计结果往往优于归一化线性算法。

16. 几何方法

本章讨论求解估计问题的几何方法。估计问题的几何方法一般不象代数方法那样能通过矩阵奇异值分解得到估计结果，需要求解一个众多变量的非线性优化问题，一般使用迭代优化技术才能求解，其计算复杂性永远高于代数方法的计算复杂性。非线性优化迭代算法在第 13 章已给出，因此本章仅将估计问题的几何方法归结为最优化问题。另外，本章还介绍了测量误差服从一般高斯分布的模型估计的最大似然方法，它等价于 Mah 距离最小化的几何方法。测量数据误差服从各向同性高斯分布，最大似然方法等价于几何方法。因此，我们将测量误差服从高斯分布的模型估计的最大似然方法纳入几何方法的框架，并给出最大似然的残差、误差和协方差矩阵的计算方法。

16.1 几何方法

16.1.1 直线与二次曲线

直线

假定平面点集 $\Theta = \{(x_j, y_j) | 1 \leq j \leq n\}$ 是来自一条直线上点的测量，问题是要估计该条直线。由于测量点不可避免地带有测量误差，只能估计一条直线 $l = (a, b, c)$ 尽可能地使点集 Θ 尽可能地“接近”该直线。这样，就导致需要给出点“接近”的数学描述。不同描述方法可能导致不同估计结果。在数学上，通常都用距离来度量“接近”的程度。例如，上节中所介绍的代数方法是用代数距离来度量“接近”的程度的。点 $\mathbf{x} = (x, y)$ 到直线 $l = (a, b, c)$ 代数距离被定义为：

$$d_{al}(\mathbf{x}, l) = |ax + by + c| \quad (16.1.1)$$

这里要求直线参数向量(齐次坐标) $l = (a, b, c)$ 满足 $a^2 + b^2 + c^2 = 1$ ，这是因为直线的齐次坐标可以相差一个常数因子。如果没有这个要求，则点到直线的代数距离(16.1.1)就是一个不确定的量。直线估计的代数方法就是寻找一条直线使得测量点到它的代数距离平方和达到最小，即求解下述线性最小二乘问题：

$$\begin{cases} \min \sum_{j=1}^n (ax_j + by_j + c)^2 \\ \text{subject to } a^2 + b^2 + c^2 = 1 \end{cases} \quad (16.1.2)$$

如果用点 $\mathbf{x} = (x, y)$ 到直线 $l = (a, b, c)$ 欧氏距离(通常称为几何距离)来度量“接近”的程度, 就得到直线估计的几何方法, 这种方法通常也称为整体最小二乘估计。由解析几何, 点 $\mathbf{x} = (x, y)$ 到直线 $l = (a, b, c)$ 几何距离的表达式为:

$$d_G(\mathbf{x}, l) = \frac{|ax + by + c|}{\sqrt{a^2 + b^2}} \quad (16.1.3)$$

不难看出几何距离 $d_G(\mathbf{x}, l)$ 与直线 l 齐次坐标选择无关, 所以可以选择直线的齐次坐标 $l = (a, b, c)$ 使得 $a^2 + b^2 = 1$ 。这样, 点 $\mathbf{x} = (x, y)$ 到直线 $l = (a, b, c)$ 几何距离可以表达为更简洁的形式:

$$d_G(\mathbf{x}, l) = |ax + by + c|, (a^2 + b^2 = 1)$$

直线的几何估计是寻找一条直线使得测量点到它的几何距离平方和达到最小, 即求解下述线性最小化问题:

$$\begin{cases} \min \sum_{j=1}^n (ax_j + by_j + c)^2 \\ \text{subject to } a^2 + b^2 = 1 \end{cases} \quad (16.1.4)$$

如果直线 $l = (a, b, c)$ 通过点 (x_0, y_0) , 则必有 $c = -(ax_0 + by_0)$, 因此点 $\mathbf{x} = (x, y)$ 到直线 $l = (a, b, c)$ 几何距离还可以表达成

$$d_G(\mathbf{x}, l) = |a(x - x_0) + b(y - y_0)|, (a^2 + b^2 = 1)$$

于是, 测量点 (x_j, y_j) 到直线 $l = (a, b, c)$ 的几何距离平方和为

$$f(x_0, y_0) = \sum_{j=1}^n (a(x_j - x_0) + b(y_j - y_0))^2$$

令 $\bar{\mathbf{x}} = (\bar{x}, \bar{y})$ 是测量数据集 $\{(x_j, y_j)\}$ 的重心, 即

$$(\bar{x}, \bar{y}) = \left(\frac{1}{n} \sum_{j=1}^n x_j, \frac{1}{n} \sum_{j=1}^n y_j \right)$$

下面证明几何距离最小化的直线通过测量数据点集的重心, 为此只须证明: $f(x_0, y_0) \geq f(\bar{x}, \bar{y})$ 。事实上, 令

$$\mathbf{w} = (w_1, w_2, \dots, w_n)^T, \mathbf{z} = (z_1, z_2, \dots, z_n)^T$$

其中

$$w_j = a(x_j - \bar{x}) + b(y_j - \bar{y}), z_j = a(x_j - x_0) + b(y_j - y_0)$$

则

$$z = \mathbf{w} + s\mathbf{1}, \begin{pmatrix} s = a(\bar{x} - x_0) + b(\bar{y} - y_0) \\ \mathbf{1} = (1, 1, \dots, 1)^T \end{pmatrix}$$

且

$$f(\bar{x}, \bar{y}) = \|\mathbf{w}\|^2, f(x_0, y_0) = \|z\|^2 = \|\mathbf{w} + s\mathbf{1}\|^2$$

由于 $\langle \mathbf{w}, \mathbf{1} \rangle = \sum_{j=1}^n w_j = a \sum_{j=1}^n (x_j - \bar{x}) + b \sum_{j=1}^n (y_j - \bar{y}) = 0$, 所以 $\mathbf{w} \perp \mathbf{1}$, 于是我们有

$$f(x_0, y_0) = \|\mathbf{w} + s\mathbf{1}\|^2 = \|\mathbf{w}\|^2 + s^2 n \geq \|\mathbf{w}\|^2 = f(\bar{x}, \bar{y})$$

这样, 最小化问题(16.1.4)归结为下述最小化问题:

$$\begin{cases} \min \sum_{j=1}^n (a(x_j - \bar{x}) + b(y_j - \bar{y}))^2 \\ \text{subject to } a^2 + b^2 = 1 \end{cases} \quad (16.1.5)$$

这个最小化问题的解是矩阵

$$M = \begin{pmatrix} x_1 - \bar{x} & y_1 - \bar{y} \\ x_2 - \bar{x} & y_2 - \bar{y} \\ \vdots & \vdots \\ x_n - \bar{x} & y_n - \bar{y} \end{pmatrix}^T \begin{pmatrix} x_1 - \bar{x} & y_1 - \bar{y} \\ x_2 - \bar{x} & y_2 - \bar{y} \\ \vdots & \vdots \\ x_n - \bar{x} & y_n - \bar{y} \end{pmatrix}$$

的最小特征值的单位特征向量。

综合上述讨论, 可以看出直线几何估计的算法是简单的, 它存在闭合形式的解。

二次曲线

点 $\mathbf{x} = (x, y, 1)^T$ 到二次曲线 C 的几何距离被定义为

$$d_G(\mathbf{x}, C) = \min\{\|\mathbf{x} - \mathbf{y}\| \mid \mathbf{y}^T C \mathbf{y} = 0\}$$

假定点集 $X = \{(x_j, y_j, 1) \mid 1 \leq j \leq n\}$ 是来自一条二次曲线上点的测量。几何估计方法是求解一条二次曲线 C 使测量点到它的几何距离平方和达到最小, 即求解下述最小化问题:

$$\min_C \sum_{j=1}^n d_G(\mathbf{x}_j, C) \quad (16.1.6)$$

由于几何距离 $d_G(\mathbf{x}, C)$ 没有显式的解析表达, 上述最小化问题在数学上缺乏可操作性。因此, 二次曲线的几何估计常通过求解下面的最小化问题来实现:

$$\begin{cases} \min_{\{C, \hat{x}_j\}} \sum_{j=1}^n \|\hat{x}_j - \mathbf{x}_j\|^2 \\ \text{subject to } \hat{x}_j^T C \hat{x}_j = 0 \end{cases} \quad (16.1.7)$$

可见，二次曲线的几何估计比直线的几何估计要复杂得多，不仅需要优化二次曲线参数，同时还需要优化测量数据点。因此，在二次曲线的几何估计过程中，需要使用非线性最优化迭代技术。

16.1.2 几何距离最小化

类比直线与二次曲线估计的几何方法，可以给出一般模型估计的几何方法。

对于一般模型 M ，对模型参数向量 $\mathbf{p} \in \Phi$ 的实例 M_p 可以看作是 R^N 中一张曲面，这张曲面通常是某个方程(组)的解集：

$$\Psi_p = \{\mathbf{x} \mid M_p(\mathbf{x}) = 0\} \quad (16.1.8)$$

因此，模型 M 可以看作是 R^N 中某类曲面的集合。

给定模型 M 的一组测量向量 $\mathbf{X} = \{\mathbf{x}_j \mid 1 \leq j \leq n\}$ ，估计模型 M 的代数方法是寻找一个模型参数向量 \mathbf{p} 使得代数距离平方和达到最小：

$$\min_p \sum_{j=1}^n \|M_p(\mathbf{x}_j)\|^2。$$

点 $\mathbf{x} \in R^N$ 到模型 M 的几何距离被定义为点 $\mathbf{x} \in R^N$ 到曲面 Ψ_p 的欧氏距离 $d_G(M_p, \mathbf{x})$ ，即

$$d_G(\mathbf{x}, M_p) = \min\{\|\mathbf{x} - \mathbf{y}\| \mid M_p(\mathbf{y}) = 0\} \quad (16.1.9)$$

估计模型 M 的几何方法是寻找一个模型参数向量 \mathbf{p} 使得几何距离平方和最小化：

$$\begin{cases} \min_p \sum_{j=1}^n d_G^2(\mathbf{x}_j, M_p) \\ \text{subject to } \mathbf{p} \in \Phi \end{cases} \quad (16.1.10)$$

一般模型 M 的几何估计需要非常复杂的计算，没有直线的几何估计那样简单，这是因为一般模型的几何距离 $d_G(M_p, \mathbf{x})$ 往往没有解析表达式，例如二次曲线的几何距离就没有解析表达式。

在几何距离 $d_G(M_p, \mathbf{x})$ 没有解析表达式的情况下，为了使几何方法在数学上易于操作，需要将最小化问题(16.1.10)转换为另一种等价形式。如果最小化问题(16.1.10)的解为 \mathbf{p}_0 ，令 $\hat{x}_j \in \Psi_{\mathbf{p}_0}$ 使得 $\|\hat{x}_j - \mathbf{x}_j\| = d_G(M_{\mathbf{p}_0}, \mathbf{x}_j)$ ，则必有

$$\min_p \sum_{j=1}^n d_G^2(M_p, \mathbf{x}_j) = \sum_{j=1}^n \|\hat{x}_j - \mathbf{x}_j\|^2 \quad (16.1.11)$$

且

$$M_{p_0}(\hat{x}_j) = 0, j = 1, 2, \dots, n \quad (16.1.12)$$

反之, 如果 $p_0 \in \Phi$, 且 \hat{x}_j 满足式(16.1.11)和(16.1.12), 则 p_0 必为最小化问题(16.1.10)的解。因此, 几何估计等价于求解下述最小化问题:

$$\begin{cases} \min_{\{p, \hat{x}_j\}} \sum_{j=1}^n \|\hat{x}_j - x_j\|^2 \\ \text{subject to } p \in \Phi, M_p(\hat{x}_j) = 0, j = 1, 2, \dots, n \end{cases} \quad (16.1.13)$$

最小化几何距离 $\|\hat{x} - x\|$ 的向量 \hat{x} 是曲面 Ψ_p 上最接近测量 x 的点, 因曲面 Ψ_p 的非线性, 所以 \hat{x} 不能被直接地估计, 需要通过迭代技术来实现。几何方法不仅要优化模型参数, 同时还需要优化测量数据, 这使得几何方法通常是求解一个众多变量的非线性优化问题。

Sampson 近似

Sampson 在“曲面 Ψ_p 的每点附近都能被很好的线性近似”的假定下, 提出了估计点 \hat{x} 的一阶近似的思想, 将最小化代价函数化为仅与模型参数有关的函数, 得到了一种近似几何估计方法, 通常称为 Sampson 近似。下面介绍这种方法。

测量点 x 通常使得 $M_p(x) \neq 0$, 我们希望在 x 附近寻找一点 $\hat{x} = x + \Delta_x$ 使得

$$M_p(x + \Delta_x) = 0$$

为此, 在 x 附近, $M_p(x)$ 用它的一阶 Taylor 展开式来近似:

$$M_p(x + \Delta_x) = M_p(x) + J_p(x) \cdot \Delta_x \quad (16.1.14)$$

其中 $J_p(x) = \frac{\partial M_p}{\partial x}$ 是曲面 M_p 的 Jacobi 矩阵。于是, 我们有

$$M_p(x) + J_p(x)\Delta_x = 0$$

即

$$J_p(x)\Delta_x = -M_p(x) \quad (16.1.15)$$

这样, 现在所面临的问题是

$$\begin{cases} \min \|\Delta_x\|^2 \\ \text{subject to } J_p(x)\Delta_x = -M_p(x) \end{cases} \quad (16.1.16)$$

求解这种问题的标准方法是 Lagrange 乘子法。在引进 Lagrange 乘子 2λ 后, (16.1.16)化为最小化下述函数:

$$f = \Delta_x^T \Delta_x - 2\lambda^T (J_p(x)\Delta_x + M_p(x)) \quad (16.1.17)$$

先求 f 关于 Δ_x 的导数，并令其为零，我们有

$$2\Delta_x - 2\lambda^T J_p(\mathbf{x}) = 0$$

因此，

$$\Delta_x = (J_p(\mathbf{x}))^T \lambda \tag{16.1.18}$$

再求 f 关于 λ 的导数，我们有

$$J_p(\mathbf{x})\Delta_x + M_p(\mathbf{x}) = 0 \tag{16.1.19}$$

从式(16.1.18)和(16.1.19)，得到

$$J_p(\mathbf{x})(J_p(\mathbf{x}))^T \lambda = -M_p(\mathbf{x})$$

于是，有

$$\lambda = -[J_p(\mathbf{x})(J_p(\mathbf{x}))^T]^{-1}M_p(\mathbf{x})$$

将上式代入式(16.1.18)，得到

$$\Delta_x = -J_p(\mathbf{x})[J_p(\mathbf{x})(J_p(\mathbf{x}))^T]^{-1}M_p(\mathbf{x})$$

因此，

$$\|\hat{\mathbf{x}} - \mathbf{x}\|^2 = \Delta_x^T \Delta_x = M_p^T(\mathbf{x})[J_p(\mathbf{x})(J_p(\mathbf{x}))^T]^{-1}M_p(\mathbf{x}) \tag{16.1.20}$$

根据上面的讨论，一般模型的 Sampson 近似是求解下述最小化问题：

$$\min_p \sum_{j=1}^n M_p^T(\mathbf{x}_j)[J_p(\mathbf{x}_j) \cdot (J_p(\mathbf{x}_j))^T]^{-1}M_p(\mathbf{x}_j) \tag{16.1.21}$$

其中： $J_p(\mathbf{x}) = \frac{\partial M_p}{\partial \mathbf{x}}$ 是曲面 M_p 的 Jacob 矩阵。这里的代价函数仅与模型参数 p 有关。

注意：由 Sampson 近似得到测量点 \mathbf{x} 的估计 $\hat{\mathbf{x}}$ 一般不严格满足 $M_p(\hat{\mathbf{x}}) = 0$ ，因为它仅是曲面 M_p 上距测量点 \mathbf{x} 的最近点 $\bar{\mathbf{x}}$ 的一阶近似，如下图 16.1.1 所示。

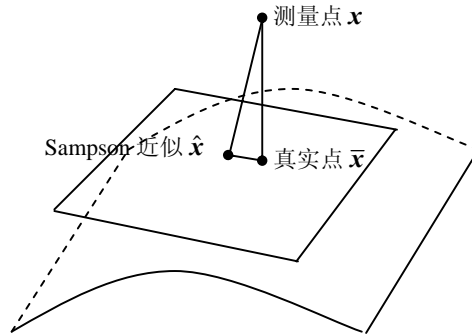


图 16.1.1 Sampson 近似

例：考虑二次曲线的 Sampson 近似。假定点集 $\{\mathbf{x}_j = (x_j, y_j, 1) \mid 1 \leq j \leq n\}$ 是来自一条二次曲线 C 的测量数据点集。由于 $M_C(\mathbf{x}) = \mathbf{x}^T C \mathbf{x}$ ，不难计算：

$$J_C(\mathbf{x}) = \left(\frac{\partial(\mathbf{x}^T C \mathbf{x})}{\partial x}, \frac{\partial(\mathbf{x}^T C \mathbf{x})}{\partial y} \right) = 2((C\mathbf{x})_1, (C\mathbf{x})_2)$$

因此，

$$\begin{aligned} \|\hat{\mathbf{x}} - \mathbf{x}\|^2 &= \Delta_x^T \Delta_x = M_C^T(\mathbf{x}) [J_C(\mathbf{x}) \cdot J_C^T(\mathbf{x})]^{-1} M_C(\mathbf{x}) \\ &= \frac{(\mathbf{x}^T C \mathbf{x})^2}{4((C\mathbf{x})_1^2 + (C\mathbf{x})_2^2)} \end{aligned}$$

所以，二次曲线的 Sampson 近似是求解下述最小化问题：

$$\min_C \sum_{j=1}^n \frac{(\mathbf{x}_j^T C \mathbf{x}_j)^2}{(C\mathbf{x}_j)_1^2 + (C\mathbf{x}_j)_2^2} \quad (16.1.22)$$

16.2 在视觉问题中的应用

16.2.1 单应矩阵

给定两幅图像间的单应矩阵：

$$H = \begin{pmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{pmatrix}$$

令 $\mathbf{x} = (x_1, x_2, 1)^T \leftrightarrow \mathbf{x}' = (x'_1, x'_2, 1)^T$ 是它的点对应，记

$$\mathbf{h} = (H_{11}, H_{12}, H_{13}, H_{21}, H_{22}, H_{23}, H_{31}, H_{32}, H_{33})^T$$

则有

$$M_H(\mathbf{X}) \equiv \begin{pmatrix} \mathbf{x}^T & \mathbf{0}^T & -x'_1 \mathbf{x}^T \\ \mathbf{0}^T & \mathbf{x}^T & -x'_2 \mathbf{x}^T \end{pmatrix} \mathbf{h} = 0 \quad (16.2.1)$$

其中 $\mathbf{X} = (x, y, x', y')^T \in R^4$ 被称为测量空间的点，它是由点对应的非齐次坐标构成的 4 维向量。实际上， $M_H(\mathbf{X})$ 是 3 维向量 $\mathbf{x}' \times H\mathbf{x}$ 的前两个分量，这里仅考虑前两个分量的理由是第三个分量为前两

个分量的线性组合。所以，单应矩阵 H 可以看作是 4 维测量空间中两个二次曲面的交集：

$$S_H = \{X \in R^4 \mid M_H(X) = 0\} \tag{16.2.2}$$

假定 $\{x_j = (x_j, y_j, 1)^T \leftrightarrow x'_j = (x'_j, y'_j, 1)^T, j=1,2,\dots,n\}$ 是来自单应矩阵 H 的一组测量对应点，即我们有单应矩阵 H 的一组测量向量 $\{X_j = (x_j, y_j, x'_j, y'_j)^T, j=1,2,\dots,n\}$ 。由这组测量向量估计单应矩阵 H 的几何方法是求解下述最小化问题：

$$\begin{cases} \min_{\{H, \hat{x}_j\}} \sum_{j=1}^n \|\hat{X}_j - X_j\|^2 \\ \text{subject to } M_H(X_j) = 0, j = 1, 2, \dots, n \end{cases} \tag{16.2.3}$$

由于

$$\|\hat{X}_j - X_j\|^2 = (\hat{x}_j - x_j)^2 + (\hat{y}_j - y_j)^2 + (\hat{x}'_j - x'_j)^2 + (\hat{y}'_j - y'_j)^2 = \|\hat{x}_j - x_j\|^2 + \|\hat{x}'_j - x'_j\|^2$$

并且

$$M_H(\hat{X}_j) = 0 \Leftrightarrow \hat{x}'_j \times H\hat{x}_j = 0$$

所以，式(16.2.3)可改写成：

$$\begin{cases} \min_{\{H, \hat{x}_j, \hat{x}'_j\}} \sum_{j=1}^n \|\hat{x}_j - x_j\|^2 + \|\hat{x}'_j - x'_j\|^2 \\ \text{subject to } \hat{x}'_j \times H\hat{x}_j = 0, j = 1, 2, \dots, n \end{cases} \tag{16.2.10}$$

因此，几何方法是从一组测量点对应 $\{x_j \leftrightarrow x'_j\}$ 确定一组与单应矩阵 H 完全相匹配的点对应 $\{\hat{x}_j \leftrightarrow \hat{x}'_j\}$ ，并使得它们的重投影 $H^{-1}\hat{x}'_j, H\hat{x}_j$ 与原始测量点之间的距离平方和最小，如图 16.2.1 所示。即几何方法等价于求解下述最小化问题：

$$\min_{\{H, \hat{x}_j, \hat{x}'_j\}} \sum_{j=1}^n (d^2(H^{-1}\hat{x}'_j, x_j) + d^2(H\hat{x}_j, x'_j)) \tag{16.2.4}$$

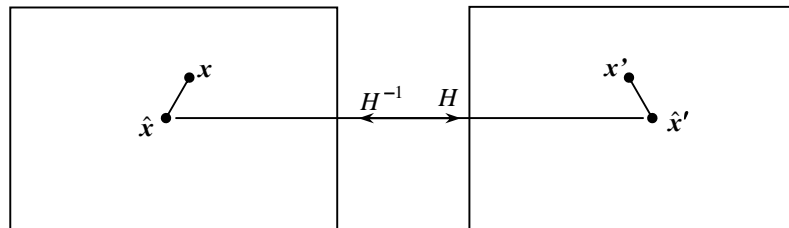


图 16.2.1 单应几何估计

式(16.2.4)通常称重投影误差最小化。重投影误差最小化方法是计算机视觉估计问题的常用方

法。上面分析表明, 对于单应矩阵估计, 几何方法与重投影误差最小化方法是等价的。

Sampson 近似

由(16.2.1)式, 容易计算出测量点 \mathbf{X} 与 Sampson 估计 $\hat{\mathbf{X}}$ 之间的距离:

$$\|\hat{\mathbf{X}} - \mathbf{X}\|^2 = \mathbf{M}_H^T(\mathbf{X})[\mathbf{J}_H(\mathbf{X})(\mathbf{J}_H(\mathbf{X}))^T]^{-1}\mathbf{M}_H(\mathbf{X})$$

其中:

$$\mathbf{J}_H(\mathbf{X}) = \begin{pmatrix} H_{11} - x'H_{31} & H_{12} - x'H_{32} & -xH_{31} - yH_{32} - H_{33} & 0 \\ H_{21} - y'H_{31} & H_{22} - y'H_{32} & 0 & -xH_{31} - yH_{32} - H_{33} \end{pmatrix}$$

$$\mathbf{M}_H(\mathbf{X}) \equiv \begin{pmatrix} \mathbf{x}^T & \mathbf{0}^T & -x'\mathbf{x}^T \\ \mathbf{0}^T & \mathbf{x}^T & -y'\mathbf{x}^T \end{pmatrix} \mathbf{h}$$

所以, 从一组测量向量 $\{\mathbf{X}_j = (x_j, y_j, x'_j, y'_j)^T, j=1, 2, \dots, n\}$, 估计单应矩阵的 Sampson 近似是求解下述最小化问题:

$$\min_H \sum_{j=1}^n \mathbf{M}_H^T(\mathbf{X}_j)[\mathbf{J}_H(\mathbf{X}_j)(\mathbf{J}_H(\mathbf{X}_j))^T]^{-1}\mathbf{M}_H(\mathbf{X}_j) \quad (16.2.4)$$

16.2.2 基本矩阵

对一个给定的基本矩阵 F 和任一点对应 $\mathbf{x} = (x, y, 1)^T \leftrightarrow \mathbf{x}' = (x', y', 1)^T$, 必有

$$\mathbf{x}'^T F \mathbf{x} = 0$$

令 $\mathbf{X} = (x, y, x', y')^T \in R^4$, 则上式可改写成下面的形式

$$g_F(\mathbf{X}) = (\mathbf{X}^T, 1)M_F \begin{pmatrix} \mathbf{X} \\ 1 \end{pmatrix} = 0$$

其中

$$M_F = \begin{pmatrix} 0 & 0 & f_{11} & f_{21} & f_{31} \\ 0 & 0 & f_{12} & f_{22} & f_{32} \\ f_{11} & f_{12} & 0 & 0 & f_{13} \\ f_{21} & f_{22} & 0 & 0 & f_{23} \\ f_{31} & f_{32} & f_{13} & f_{23} & 2f_{33} \end{pmatrix} \quad (16.2.5)$$

由基本矩阵的性质, 可以证明 $\text{rank}(M_F)=4$, 因此 $g_F(\mathbf{X})=0$ 定义了 4 维空间中的一个锥面。

令 $\{\mathbf{x}_j = (x_j, y_j, 1)^T \leftrightarrow \mathbf{x}'_j = (x'_j, y'_j, 1)^T, j=1, 2, \dots, n\}$ 是两幅图像间的一组测量点对应, 记

$\mathbf{X}_j = (x_j, y_j, x'_j, y'_j)^T$, $j=1,2,\dots,n$, 它是测量空间 R^4 的一组测量数据。于是, 基本矩阵的几何估计是求解下述最小化问题:

$$\begin{cases} \min_{\{\mathbf{F}, \hat{\mathbf{X}}_j\}} \sum_{j=1}^n \|\hat{\mathbf{X}}_j - \mathbf{X}_j\|^2 \\ \text{subject to } g_F(\hat{\mathbf{X}}_j) = 0, j = 1, 2, \dots, n \end{cases} \quad (16.2.6)$$

由于 $\|\hat{\mathbf{X}}_j - \mathbf{X}_j\|^2 = \|\hat{\mathbf{x}}_j - \mathbf{x}_j\|^2 + \|\hat{\mathbf{x}}'_j - \mathbf{x}'_j\|^2$, 且 $g_F(\hat{\mathbf{X}}_j) = 0 \Leftrightarrow \hat{\mathbf{x}}_j'^T F \hat{\mathbf{x}}_j = 0$, 式(16.2.6)可改写为

$$\begin{cases} \min_{\{\mathbf{F}, \hat{\mathbf{x}}_j, \hat{\mathbf{x}}'_j\}} \sum_{j=1}^n \|\hat{\mathbf{x}}_j - \mathbf{x}_j\|^2 + \|\hat{\mathbf{x}}'_j - \mathbf{x}'_j\|^2 \\ \text{subject to } \hat{\mathbf{x}}_j'^T F \hat{\mathbf{x}}_j = 0, j = 1, 2, \dots, n \end{cases}$$

Sampson 近似

不难计算

$$J_F(\mathbf{X}) = \frac{\partial g_F}{\partial \mathbf{X}} = ((F^T \mathbf{x}'_1), (F^T \mathbf{x}'_2), (F\mathbf{x})_1, (F\mathbf{x})_2)$$

所以

$$\|\hat{\mathbf{X}} - \mathbf{X}\|^2 = g_F^T(\mathbf{X}) [J_F(\mathbf{X}) (J_F(\mathbf{X}))^T]^{-1} g_F(\mathbf{X}) = \frac{(\mathbf{x}'^T F \mathbf{x})^2}{(F^T \mathbf{x}'_1)^2 + (F^T \mathbf{x}'_2)^2 + (F\mathbf{x})_1^2 + (F\mathbf{x})_2^2}$$

因此, 基本矩阵的 **Sampson** 近似是求解下述最小化问题:

$$\min_F \sum_{j=1}^n \frac{(\mathbf{x}'_j{}^T F \mathbf{x}_j)^2}{(F^T \mathbf{x}'_j)_1^2 + (F^T \mathbf{x}'_j)_2^2 + (F\mathbf{x}_j)_1^2 + (F\mathbf{x}_j)_2^2} \quad (16.2.7)$$

极距离最小化

根据极几何理论, 还可以给出十分类似于 **Sampson** 近似的估计方法。对于第一幅像平面上任一点 \mathbf{x} , 它在第二幅像平面上的对应点 \mathbf{x}' 必位于对应的外极线 l'_x 上。由基本矩阵的性质, 极线 $l'_x = F\mathbf{x}$, 并且不难计算 \mathbf{x}' 到极线 l'_x 距离为

$$d(\mathbf{x}', F\mathbf{x}) = \frac{|\mathbf{x}'^T F \mathbf{x}|}{\sqrt{(F\mathbf{x})_1^2 + (F\mathbf{x})_2^2}}$$

对于精确的点对应, 上述距离应等于零。类似地, 有 \mathbf{x} 到极线 l_x 距离为

$$d(\mathbf{x}, F^T \mathbf{x}') = \frac{|\mathbf{x}'^T F \mathbf{x}|}{\sqrt{(F^T \mathbf{x}')_1^2 + (F^T \mathbf{x}')_2^2}}$$

因此, 基于极距离最小化的基本矩阵估计是解下述最小化问题:

$$\min_F \sum_{j=1}^n \left(\frac{1}{(F \mathbf{x}_j)_1^2 + (F \mathbf{x}_j)_2^2} + \frac{1}{(F^T \mathbf{x}'_j)_1^2 + (F^T \mathbf{x}'_j)_2^2} \right) (\mathbf{x}'_j^T F \mathbf{x}_j)^2 \quad (16.2.8)$$

注意: 估计基本矩阵的几何方法所得到的矩阵 F 不满足基本矩阵的秩 2 约束, 为了保证秩 2 约束成立, 可通过两种方法来实现。第一种方法与线性方法相同, 对矩阵 F 在 F-范数下作秩 2 矩阵的最佳近似, 这可通过 SVD 分解技术来实现。第二种方法是对 F 作适当的参数化, 使得几何方法的估计结果是一个秩 2 矩阵, 这种方法将导致几何方法的高度非线性。下面介绍两种基本矩阵的参数化方法。

秩 2 参数化表示

极参数表示 由于基本矩阵 F 的秩为 2, 所以 F 的三个行向量必线性相关。因此, 在一般情况下第三行可由前两行线性表示, 这样基本矩阵就可表示为下述形式:

$$F = \begin{pmatrix} a & b & c \\ d & e & f \\ ga + hd & gb + he & gc + hf \end{pmatrix}$$

不难验证: 极点 $\mathbf{e}' = (-g, -h, 1)^T$ 。

在迭代算法中如果使用这种参数表示, 则其初始值可按下述方法确定: 令线性算法的解为

$$F_0 = \begin{pmatrix} f_{01} & f_{02} & f_{03} \\ f_{04} & f_{05} & f_{06} \\ f_{07} & f_{08} & f_{09} \end{pmatrix}$$

取 $a = f_{01}, b = f_{02}, c = f_{03}, d = f_{04}, e = f_{05}, f = f_{06}$ 作为其初始值, g, h 的初始值取下述最小线性二乘问题的解:

$$\min [(f_{01}g + f_{04}h - f_{07})^2 + (f_{02}g + f_{05}h - f_{08})^2 + (f_{03}g + f_{06}h - f_{09})^2]$$

这种表示方法的不足之处在于, 它可能导致算法的不稳定。这是因为第三行不总是可以由前两行线性表示。由于 F 的秩为 2, 所以第三行不能由前两行线性表示的充要条件是前两行线性相关, 而前两行线性相关等价于极点在无穷远点, 即 $\mathbf{e}' = (e'_1, e'_2, 0)^T$ 。这说明在实际问题中, 这种情况很

少会发生。另外，可以利用 8-点算法对极点的估计结果来预测是否发生这种情况。如果 8-点算法所估计的极点距图像中心非常远，则改变参数化方式，如用第一行和第三行来线性表示第二行。

两极表示 在一般情况下第三行可由前两行线性表示，同时第三列也可以由前两列线性表示，于是基本矩阵可表示为下述形式：

$$F = \begin{pmatrix} a & b & ga + hb \\ c & d & gc + hd \\ ea + fc & eb + fd & g(ea + fc) + h(eb + fd) \end{pmatrix}$$

同样，可以验证 $e' = (-e, -f, 1)^T, e = (-g, -h, 1)^T$ 。

给定代价函数，如果有若干种参数化可供选择的话，参数化的一般策略是所选择参数与求解的问题相匹配，同时用这些参数计算代价函数时有较小的计算量。例如 F 可以用 9 个参数来参数化——那么它是过参数化的，因为实际上 F 仅有 7 个自由度。如果最小化问题不是太过参数化，只要所采用的参数化与最小化问题相匹配，就不会产生坏的效应。因为一个性能好的非线性最小化算法会“注意”不必要在冗余方向移动，能有效地丢弃冗余参数组合，如 Levenberg-Marquardt 算法就很容易处理冗余参数化问题。经验发现采用最低限度参数化时常使代价函数变得更为复杂，这样陷入局部最小值的机会就可能更多。因此，在基本矩阵的参数化过程中，不必用 7 个参数表达。但是，如果采用的参数化与最小化问题不匹配，过参数化会产生坏的效应。例如，纯平移运动下的基本矩阵，是一个反对称矩阵，有两个自由度，用一个 3 参数的反对称矩阵来表示虽然过参数化，但不会影响计算结果。如果利用 9 参数或 8 参数的基本矩阵来表示，与对应的最小化问题就不相匹配，其计算结果往往都不是一个反对称矩阵，这不是所期望的结果。

16.2.3 三焦张量

给定三幅图像之间的三焦张量 t_i^{qr} ，令

$$\mathbf{x} = (x^1, x^2, 1)^T, \mathbf{x}' = (x'^1, x'^2, 1)^T, \mathbf{x}'' = (x''^1, x''^2, 1)^T$$

是三幅图像间三点对应，根据第 12 章的讨论，我们有

$$x^i x'^j x''^k \varepsilon_{jqn} \varepsilon_{krv} t_i^{qr} = 0_{uv} \tag{16.2.9}$$

其中 $x^3 = x'^3 = x''^3 = 1$ ，并且在这 9 个方程中，仅有 4 个是相互独立的。因此，三焦张量 t_i^{qr} 是 R^6 维空间中 4 个三次曲面的交集，记为 $M_{t_i^{qr}}(\mathbf{X}) = 0$ ，其中 $\mathbf{X} = (x^1, x^2, x'^1, x'^2, x''^1, x''^2)^T$ 。

令 $\{x_j \leftrightarrow x'_j \leftrightarrow x''_j, j=1,2,\dots,n\}$ 是三幅图像间的一组测量点对应，则不难看出三焦张量 t_i^{qr} 的几何估计是求解下述最小化问题：

$$\begin{cases} \min_{\{t_i^{qr}, \hat{X}_j\}} \sum_{j=1}^n \|\hat{X}_j - X_j\|^2 \\ \text{subject to } M_{t_i^{qr}}(\hat{X}_j) = 0, j = 1, 2, \dots, n \end{cases} \quad (16.2.10)$$

注意：这里没有考虑三焦张量的内在约束。如果考虑内在约束，则需要求解更复杂的优化问题。

16.2.4 FOE

FOE 估计是基本矩阵估计的特别情况，因为 FOE 是摄像机作纯平移所获取两幅图像上的极点 $\mathbf{e} = (e_1, e_2, 1)^T$ ，此时的基本矩阵为 $F = [\mathbf{e}]_{\times}$ 。根据基本矩阵估计的几何方法，FOE 估计的几何方法是从测量点对应 $\{\mathbf{x}_j \leftrightarrow \mathbf{y}_j\}$ 求解下述非线性最小化问题：

$$\begin{cases} \min_{\mathbf{e}, \{\hat{\mathbf{x}}_j \leftrightarrow \hat{\mathbf{y}}_j\}} \sum_{j=1}^n (\|\mathbf{x}_j - \hat{\mathbf{x}}_j\| + \|\mathbf{y}_j - \hat{\mathbf{y}}_j\|) \\ \text{subject to } \hat{\mathbf{y}}_j \cdot (\mathbf{e} \times \hat{\mathbf{x}}_j) = 0, \forall j \end{cases} \quad (16.2.11)$$

这是一具有 $2n+2$ 个变量的最小化问题。下面，利用 FOE 的特殊性，将它转化为两个变量的最小化问题。

假定 $(\hat{\mathbf{e}}, \{\hat{\mathbf{x}}_j \leftrightarrow \hat{\mathbf{y}}_j\})$ 是最小化问题 (16.2.11) 的解，则 $\hat{\mathbf{e}}$ 必在直线 $l_j = \hat{\mathbf{x}}_j \times \hat{\mathbf{y}}_j$ 上，并且从 $\mathbf{x}_j(\mathbf{y}_j)$ 到线 $l_j = \hat{\mathbf{x}}_j \times \hat{\mathbf{y}}_j$ 的垂足为 $\hat{\mathbf{x}}_j(\hat{\mathbf{y}}_j)$ ，因此不难证明 最小化问题(16.2.11) 与下述最小化问题等价：

$$\min_{\mathbf{e}, \lambda_j} \sum_{j=1}^M \frac{|l_j(\lambda_j, \mathbf{e}) \cdot \mathbf{x}_j| + |l_j(\lambda_j, \mathbf{e}) \cdot \mathbf{y}_j|}{\sqrt{(l_j(\lambda_j, \mathbf{e}))_1^2 + (l_j(\lambda_j, \mathbf{e}))_2^2}} \quad (16.2.12)$$

其中：

$l_j(\lambda_j, \mathbf{e}) = \mathbf{e} \times \mathbf{p}_j(\lambda_j)$ 是由两点 \mathbf{e} ， $\mathbf{p}_j(\lambda_j) = \lambda_j \mathbf{x}_j + (1 - \lambda_j) \mathbf{y}_j$ 确定的直线；

$(l_j(\lambda_j, \mathbf{e}))_k$ 是直线 $l_j(\lambda_j, \mathbf{e})$ 的第 k 个坐标分量；

$d_1(\lambda_j, \mathbf{e}) = \frac{|l_j(\lambda_j, \mathbf{e}) \cdot \mathbf{x}_j|}{\sqrt{(l_j(\lambda_j, \mathbf{e}))_1^2 + (l_j(\lambda_j, \mathbf{e}))_2^2}}$ 是点 \mathbf{x}_j 到线 $l_j(\lambda_j, \mathbf{e})$ 的距离；

$$d_2(\lambda_j, \mathbf{e}) = \frac{|\mathbf{l}_j(\lambda_j, \mathbf{e}) \cdot \mathbf{y}_j|}{\sqrt{(\mathbf{l}_j(\lambda_j, \mathbf{e}))_1^2 + (\mathbf{l}_j(\lambda_j, \mathbf{e}))_2^2}}$$

是点 \mathbf{y}_j 到线 $\mathbf{l}_j(\lambda_j, \mathbf{e})$ 的距离。

假定 $(\hat{\mathbf{e}}, \hat{\lambda}_j)$ 是最小化问题 (16.2.12) 的解, 则 $\{\hat{\mathbf{x}}_j, \hat{\mathbf{y}}_j\}$ 位于 $\mathbf{l}_j(\hat{\lambda}_j, \hat{\mathbf{e}})$ 上, 并且测量点对应 $\{\mathbf{x}_j \leftrightarrow \mathbf{y}_j\}$ 到线 $\mathbf{l}_j(\hat{\lambda}_j, \hat{\mathbf{e}})$ 的距离在 $\{\hat{\mathbf{x}}_j, \hat{\mathbf{y}}_j\}$ 达到最小值。因此, 必有

$$\begin{cases} \hat{\mathbf{x}}_j = \frac{(\mathbf{x}_j - \hat{\mathbf{e}}) \cdot (\mathbf{p}_j(\hat{\lambda}_j) - \hat{\mathbf{e}})}{\|\mathbf{p}_j(\hat{\lambda}_j) - \hat{\mathbf{e}}\|^2} (\mathbf{p}_j(\hat{\lambda}_j) - \hat{\mathbf{e}}) + \hat{\mathbf{e}} \\ \hat{\mathbf{y}}_j = \frac{(\mathbf{y}_j - \hat{\mathbf{e}}) \cdot (\mathbf{p}_j(\hat{\lambda}_j) - \hat{\mathbf{e}})}{\|\mathbf{p}_j(\hat{\lambda}_j) - \hat{\mathbf{e}}\|^2} (\mathbf{p}_j(\hat{\lambda}_j) - \hat{\mathbf{e}}) + \hat{\mathbf{e}} \end{cases} \quad (16.2.13)$$

简化的几何方法

纯平移摄像机得到的图像点对应 $\{\mathbf{x}_j \leftrightarrow \mathbf{y}_j\}$ 必满足下述序约束:

如果摄像机朝向物体运动, 则有

$$\|\mathbf{x}_j - \mathbf{e}\| > \|\mathbf{y}_j - \mathbf{e}\|;$$

否则, 如果摄像机背离物体运动, 则有

$$\|\mathbf{x}_j - \mathbf{e}\| < \|\mathbf{y}_j - \mathbf{e}\|。$$

通过比较物体在两幅图像中面积的大小, 我们可以确定是朝向物体运动, 还是背离物体运动。根据这个约束, 我们可以证明下述结论:

引理 令 $(\hat{\mathbf{e}}, \hat{\lambda}_j)$ 是最小化问题 (16.2.12) 的解。摄像机朝向物体运动, 则 $\hat{\lambda}_j = 0$; 否则, 否则, 如果摄像机背离物体运动, 则 $\hat{\lambda}_j = 1$ 。

证明 对每一个 j , 考虑下述最小化问题:

$$\min_{\lambda_j} \frac{|\mathbf{l}_j(\lambda_j, \mathbf{e}) \cdot \mathbf{x}_j| + |\mathbf{l}_j(\lambda_j, \mathbf{e}) \cdot \mathbf{y}_j|}{\sqrt{(\mathbf{l}_j(\lambda_j, \mathbf{e}))_1^2 + (\mathbf{l}_j(\lambda_j, \mathbf{e}))_2^2}} \quad (16.2.12')$$

令

$$d_1(\lambda_j) = \frac{|\mathbf{l}_j(\lambda_j, \hat{\mathbf{e}}) \cdot \mathbf{x}_j|}{\sqrt{(\mathbf{l}_j(\lambda_j, \hat{\mathbf{e}}))_1^2 + (\mathbf{l}_j(\lambda_j, \hat{\mathbf{e}}))_2^2}}$$

$$d_2(\lambda_j) = \frac{|I_j(\lambda_j, \hat{\mathbf{e}}) \cdot \mathbf{y}_j|}{\sqrt{(I_j(\lambda_j, \hat{\mathbf{e}}))_1^2 + (I_j(\lambda_j, \hat{\mathbf{e}}))_2^2}}$$

假定摄像机朝向物体运动，则

当 $\lambda_j < 0$ 时，我们有 (见图 16.2.2 (a)):

$$d_1(0) + d_2(0) = d_1(0) < d_1(\lambda_j) < d_1(\lambda_j) + d_2(\lambda_j). \quad (16.2.14)$$

当 $\lambda_j > 1$ 时，我们有 (见图 16.2.2 (b)):

$$d_1(1) + d_2(1) = d_2(1) < d_2(\lambda_j) < d_1(\lambda_j) + d_2(\lambda_j). \quad (16.2.15)$$

当 $0 \leq \lambda_j \leq 1$ 时，我们有 (见图 16.2.2 (c)):

$$2\sigma(\mathbf{x}_j \hat{\mathbf{e}} \mathbf{p}_j(\lambda_j)) = \|\hat{\mathbf{e}} - \mathbf{p}_j(\lambda_j)\| \cdot d_1(\lambda_j),$$

$$2\sigma(\mathbf{y}_j \hat{\mathbf{e}} \mathbf{p}_j(\lambda_j)) = \|\hat{\mathbf{e}} - \mathbf{p}_j(\lambda_j)\| \cdot d_2(\lambda_j),$$

其中 $\sigma(\mathbf{abc})$ 表示 $\Delta \mathbf{abc}$ 的面积。显然，

$$\sigma(\mathbf{x}_j \hat{\mathbf{e}} \mathbf{y}_j) = \sigma(\mathbf{x}_j \hat{\mathbf{e}} \mathbf{p}_j(\lambda_j)) + \sigma(\mathbf{y}_j \hat{\mathbf{e}} \mathbf{p}_j(\lambda_j)),$$

因此，有

$$d_1(\lambda_j) + d_2(\lambda_j) = \frac{2\sigma(\mathbf{x}_j \hat{\mathbf{e}} \mathbf{y}_j)}{\|\hat{\mathbf{e}} - \mathbf{p}_j(\lambda_j)\|}.$$

根据序约束，我们有

$$\|\hat{\mathbf{e}} - \mathbf{p}_j(0)\| > \|\hat{\mathbf{e}} - \mathbf{p}_j(1)\|$$

所以， $\|\hat{\mathbf{e}} - \mathbf{p}_j(\lambda_j)\|$ 在 $[0, 1]$ 上单调递减，从而 $d_1(\lambda_j) + d_2(\lambda_j)$ 在 $[0, 1]$ 上单调递增。于是，当 $0 < \lambda_j \leq 1$ 时，我们有

$$d_1(0) + d_2(0) < d_1(\lambda_j) + d_2(\lambda_j). \quad (16.2.16)$$

根据 (16.2.14, 16.2.15, 16.2.16)， $d_1(\lambda_j) + d_2(\lambda_j)$ 在 $\lambda_j = 0$ 达到最小值。这样，下述代价函数

$$f(\mathbf{e}, \lambda_j) = \sum_{j=1}^M \frac{|I_j(\lambda_j, \mathbf{e}) \cdot \mathbf{x}_j| + |I_j(\lambda_j, \mathbf{e}) \cdot \mathbf{y}_j|}{\sqrt{(I_j(\lambda_j, \mathbf{e}))_1^2 + (I_j(\lambda_j, \mathbf{e}))_2^2}}$$

必在 $\lambda_j = 0$ 达到最小值。由于 $(\hat{e}, \hat{\lambda}_j)$ 最小化问题 (16.2.12) 的解，所以 $\hat{\lambda}_j = 0$ 。类似地，可以证明当摄像机背离物体运动时，相应的最小化问题在 $\hat{\lambda}_j = 1$ 达到最小值。

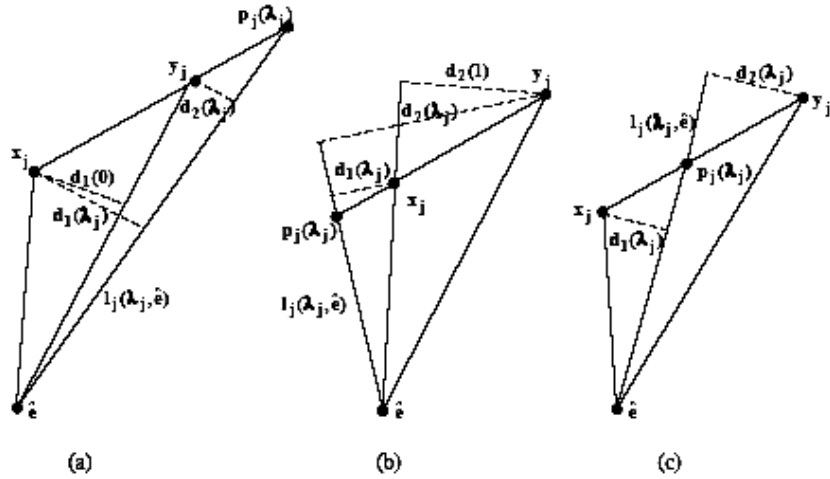


图 16.2.2. 点 $x_j(y_j)$ 到线 $l(\lambda_j, \hat{e})$ 的距离 $d_1(\lambda_j)(d_2(\lambda_j))$: (a) $\lambda_j < 0$; (b) $\lambda_j > 1$; (c) $0 \leq \lambda_j \leq 1$.

根据上述引理结论，我们有

命题 7.2.1 如果摄像机朝向物体运动，则几何方法等价于求解下述最小化问题：

$$\min_e \sum_{j=1}^M \frac{|(e \times y_j) \cdot x_j|}{\sqrt{(e \times y_j)_1^2 + (e \times y_j)_2^2}} \quad (16.2.17)$$

如果摄像机背离物体运动，则几何方法等价于求解下述最小化问题：

$$\min_e \sum_{j=1}^M \frac{|(e \times x_j) \cdot y_j|}{\sqrt{(e \times x_j)_1^2 + (e \times x_j)_2^2}} \quad (16.2.18)$$

证明 假定摄像机朝向物体运动，根据引理，最小化问题 (16.2.12) 等价于下述最小化问题：

$$\min_e \sum_{j=1}^M \frac{|l_j(0, e) \cdot x_j| + |l_j(0, e) \cdot y_j|}{\sqrt{(l_j(0, e))_1^2 + (l_j(0, e))_2^2}}. \quad (16.2.19)$$

由于 $l_j(0, e) = e \times p_j(0) = e \times y_j$ ，所以 $l_j(0, e) \cdot x_j = (e \times y_j) \cdot x_j$ ， $l_j(0, e) \cdot y_j = (e \times y_j) \cdot y_j = 0$ ，因

此 (16.2.19) 化为(16.2.17)。同理, 可证明摄像机背离物体运动的情况。

16.3 最大似然方法

16.3.1 高斯分布

令 $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ 是一个 n 维随机向量, 它的均值为 $\bar{\mathbf{x}} = E[\mathbf{x}]$, 其中 $E[\cdot]$ 表示期望值, 而且它的协方差矩阵为

$$\Sigma = E[(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}})^T] = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{pmatrix}$$

其中 $\sigma_{ij} = E[(x_i - \bar{x}_i)(x_j - \bar{x}_j)]$ 。

如果 \mathbf{x} 有概率密度

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{n/2} (\det(\Sigma))^{1/2}} \text{Exp}\left(-\frac{\|\mathbf{x} - \bar{\mathbf{x}}\|_{\Sigma}^2}{2}\right) \quad (16.3.1)$$

其中 Σ 为正定矩阵, $\|\mathbf{x} - \bar{\mathbf{x}}\|_{\Sigma} = ((\mathbf{x} - \bar{\mathbf{x}})^T \Sigma^{-1} (\mathbf{x} - \bar{\mathbf{x}}))^{1/2}$ (通常称为 **Mahalanobis (Mah)** 距离), 则称 \mathbf{x} 服从高斯分布, 或者说 x_j 服从一个联合高斯分布, 记作 $\mathbf{x} \sim N_n(\bar{\mathbf{x}}, \Sigma)$ 。可以验证 $\bar{\mathbf{x}}$, Σ 分别为这个分布的均值和协方差矩阵。

当 $\Sigma = \sigma^2 I$ (纯量矩阵) 时, 高斯密度有下面的简单形式:

$$f(\mathbf{x}) = \frac{1}{(2\pi\sigma)^{n/2}} \text{Exp}\left(-\frac{\|\mathbf{x} - \bar{\mathbf{x}}\|^2}{2\sigma^2}\right) \quad (16.3.2)$$

并称随机向量 \mathbf{x} 有各向同性的高斯分布。

坐标变换

由于 Σ 为正定矩阵, 所以它有特证值分解 $\Sigma = U^T D U$, 其中 U 是正交矩阵, $D = \text{diag}(d_1^2, d_2^2, \dots, d_n^2)$ 是对角矩阵。置 $\mathbf{y} = U\mathbf{x}$, 则 $\bar{\mathbf{y}} = U\bar{\mathbf{x}}$, 并将它们代入式(16.3.1), 我们有

$$g(\mathbf{y}) = \frac{1}{(2\pi)^{n/2} (\det(D))^{1/2}} \text{Exp}\left(-\frac{\|\mathbf{y} - \bar{\mathbf{y}}\|_D^2}{2}\right) \quad (16.3.3)$$

这表明由正交坐标变换将一般高斯密度函数变为具有对角协方差矩阵的高斯密度函数。如果在每个方向上再进行一个 d_j 的伸缩变换, 即 $\mathbf{z} = D^{-1/2} \mathbf{y}$, 则有 $\bar{\mathbf{z}} = D^{-1/2} \bar{\mathbf{y}}$, 并且

$$h(\mathbf{z}) = \frac{1}{(2\pi)^{n/2}} \text{Exp} \left(-\frac{\|\mathbf{z} - \bar{\mathbf{z}}\|^2}{2} \right) \quad (16.3.4)$$

即 $\mathbf{z} = D^{1/2} \mathbf{y}$ 是一个各向同性的高斯随机变量。对于更一般情况, 有下述命题。

命题 16.3.1 (1) 若 $\mathbf{x} \sim N_n(\bar{\mathbf{x}}, \Sigma)$, $\mathbf{y} = A\mathbf{x} + \mathbf{b}$, 其中 \mathbf{b} 是 n 元常向量, A 是秩为 m 的 $m \times n$ 矩阵, 则

$$\mathbf{y} \sim N_m(A\bar{\mathbf{x}} + \mathbf{b}, A\Sigma A^T)$$

即 \mathbf{y} 是服从均值为 $A\bar{\mathbf{x}} + \mathbf{b}$ 、协方差矩阵为 $A\Sigma A^T$ 的高斯随机向量;

(2) 若 $\mathbf{x} \sim N_n(\bar{\mathbf{x}}, \Sigma)$, 令

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}_{n_1, n_2}, \bar{\mathbf{x}} = \begin{pmatrix} \bar{\mathbf{x}}_1 \\ \bar{\mathbf{x}}_2 \end{pmatrix}_{n_1, n_2}, \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}_{n_1, n_2}, (n_1 + n_2 = n)$$

则 $\mathbf{x}_1 \sim N_{n_1}(\bar{\mathbf{x}}_1, \Sigma_{11}), \mathbf{x}_2 \sim N_{n_2}(\bar{\mathbf{x}}_2, \Sigma_{22})$

(3) 若 $\mathbf{x} \sim N_n(\bar{\mathbf{x}}, \Sigma)$, 则 x_i, x_j 相互独立的充要条件为 $\sigma_{ij} = 0$ 。又若 $\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}_{n_1, n_2}$, 则 $\mathbf{x}_1, \mathbf{x}_2$ 相互

独立的充要条件为 $\Sigma_{12} = 0$ 。

16.3.2 最大似然估计

几何方法的代价函数与测量点和被估计点之间的几何距离有关, 没有考虑测量误差(噪声)的类型。为了得到一个最优(最好)的估计, 需要为噪声建模。假定在点 \mathbf{X} 没有测量误差时, 其真实点 $\bar{\mathbf{X}}$ 准确地满足一个模型实例 \mathbf{M}_p , 即 $\mathbf{M}_p(\bar{\mathbf{X}}) = 0$ 。通常, 还假定测量误差 $\mathbf{X} - \bar{\mathbf{X}}$ 服从均值为零、协方差矩阵为 Σ 的高斯(或正态)分布。当然, 这个假设一般没有被验证。在测量数据中包含有严重的误差(错误数据, 即偏离真实非常远), 这个假设是靠不住的, 如果测量数据中不包含错误数据或者这样的数据事先被去除, 则这个假设是比较可靠的, 尽管仍没有严格论证。于是, 测量点 \mathbf{X} 的概率密度为

$$f(\mathbf{X} | \mathbf{M}_p) = \frac{1}{(2\pi)^{n/2} (\det(\Sigma))^{1/2}} \text{Exp} \left(-\frac{\|\mathbf{X} - \bar{\mathbf{X}}\|_{\Sigma}^2}{2} \right) \quad (16.3.5)$$

$f(\mathbf{X} | \mathbf{M}_p)$ 表示在给定的模型实例 \mathbf{M}_p 时得到测量 \mathbf{X} 的概率。

进一步假定每个点的测量误差是相互独立的，则受噪声干扰的测量数据 $\{\mathbf{X}_j\}$ 的概率密度是它们单个密度的积，即

$$f(\{\mathbf{X}_j\} | \mathbf{M}_p) = \prod_{j=1}^n \left(\frac{1}{(2\pi)^{n/2} (\det(\Sigma))^{1/2}} \text{Exp} \left(-\frac{\|\mathbf{X}_j - \bar{\mathbf{X}}_j\|_{\Sigma}^2}{2} \right) \right) \quad (16.3.6)$$

因此，测量数据 $\{\mathbf{X}_j\}$ 的对数似然为

$$\log(f(\{\mathbf{X}_j\} | \mathbf{M}_p)) = -\frac{1}{2} \sum_{j=1}^n \|\mathbf{X}_j - \bar{\mathbf{X}}_j\|_{\Sigma}^2 + \text{常数}$$

因此，模型的最大似然(ML)估计 $\mathbf{M}_{\hat{p}}$ 使得这个对数似然最大化，即使得下式最小化：

$$\sum_{j=1}^n \|\mathbf{X}_j - \bar{\mathbf{X}}_j\|_{\Sigma}^2 \quad (16.3.7)$$

在实际应用中，无法确定真实值 $\{\bar{\mathbf{X}}_j\}$ ，只能寻找满足模型 \mathbf{M}_p 的点 $\{\hat{\mathbf{X}}_j\}$ ，即满足 $\mathbf{M}_p(\hat{\mathbf{X}}_j) = 0$ 的点，作为真实值 $\{\bar{\mathbf{X}}_j\}$ ，因此模型的最大似然方法是求解下述最小化问题：

$$\begin{cases} \min_{\{p, \hat{\mathbf{X}}_j\}} \sum_{j=1}^n \|\mathbf{X}_j - \hat{\mathbf{X}}_j\|_{\Sigma}^2 \\ \text{subject to } \mathbf{M}_p(\hat{\mathbf{X}}_j) = 0 \end{cases} \quad (16.3.8)$$

由此可见，最大似然方法与几何方法的不同之处是用 Mah 距离代替几何距离。

如果测量误差服从各向同性高斯分布，即， $\Sigma = \sigma^2 \mathbf{I}$ ，则最大似然方法求解下述最小化问题

$$\begin{cases} \min_{\{p, \hat{\mathbf{X}}_j\}} \frac{1}{\sigma^2} \sum_{j=1}^n \|\mathbf{X}_j - \hat{\mathbf{X}}_j\|^2 \\ \text{subject to } \mathbf{M}_p(\hat{\mathbf{X}}_j) = 0 \end{cases} \quad (16.3.9)$$

此时，最大似然方法等价于几何方法。

单应矩阵 假定 $\{\mathbf{x}_j = (x_j, y_j, 1)^T \leftrightarrow \mathbf{x}'_j = (x'_j, y'_j, 1)^T, j=1, 2, \dots, n\}$ 是来自单应矩阵 \mathbf{H} 的一组测量点对应，并且所有点的测量误差是互相独立的，分别具有协方差矩阵 Σ, Σ' ，则单应矩阵 \mathbf{H} 的最大似然估计是求解下述最小化问题：

$$\begin{cases} \min_{\{\hat{x}_j, \hat{x}'_j\}} \sum_{j=1}^n \|\hat{x}_j - \mathbf{x}_j\|_{\Sigma}^2 + \|\hat{x}'_j - \mathbf{x}'_j\|_{\Sigma'}^2 \\ \text{subject to } \hat{x}'_j \times H\hat{x}_j = 0, j=1,2,\dots,n \end{cases} \quad (16.3.10)$$

基本矩阵 假定 $\{\mathbf{x}_j = (x_j, y_j, 1)^T \leftrightarrow \mathbf{x}'_j = (x'_j, y'_j, 1)^T, j=1,2,\dots,n\}$ 是来自两幅图像间的一组测量对应点，并且所有点的测量误差是互相独立的，分别具有协方差矩阵 Σ, Σ' ，则基本矩阵 F 的最大似然估计是求解下述最小化问题：

$$\begin{cases} \min_{\{\hat{x}_j, \hat{x}'_j\}} \sum_{j=1}^n \|\hat{x}_j - \mathbf{x}_j\|_{\Sigma}^2 + \|\hat{x}'_j - \mathbf{x}'_j\|_{\Sigma'}^2 \\ \text{subject to } \hat{x}'_j{}^T F \hat{x}_j = 0, j=1,2,\dots,n \end{cases} \quad (16.3.11)$$

16.3.3 残差与估计误差

上节指出，模型估计的最大似然方法是求解 Mah 距离最小化问题。当测量误差满足各向同性高斯分布时，最大似然方法与几何方法等价。在本节中，假定测量误差满足各向同性高斯分布 $N_m(0, \sigma^2 I)$ ，其中 m 测量数据的维数，即 \mathbf{X}_j 的维数。

令

$$e_{res}(\hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2, \dots, \hat{\mathbf{X}}_n) = \left(\frac{1}{nm} \sum_{j=1}^n \|\hat{\mathbf{X}}_j - \mathbf{X}_j\|^2 \right)^{1/2} \quad (16.3.12)$$

并称为均方根 (RMS) 残差。因此，最大似然估计使得 RMS 残差最小，即

$$e_{res}^{mel} = \min \{ e_{res}^2(\hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2, \dots, \hat{\mathbf{X}}_n) : M_p(\hat{\mathbf{X}}_j) = 0, p \in \Theta \} \quad (16.3.13)$$

记 $\bar{\mathbf{X}}_j$ 是测量 \mathbf{X}_j 的真值，则称

$$e_{est} = \left(\frac{1}{nm} \sum_{j=1}^n \|\hat{\mathbf{X}}_j - \bar{\mathbf{X}}_j\|^2 \right)^{1/2} \quad (16.3.14)$$

是 RMS 估计误差。

下面讨论最大似然估计的 RMS 残差和 RMS 估计误差的数学期望，即在平均的意义下给出最大似然估计的 RMS 残差和 RMS 估计误差，这是非常有意义的。因为这些值给出了评价一个具体算法的残差和估计误差的下界。

为了达到上述目的，我们需要用更抽象方式来表述模型估计问题。

估计问题由下述三个成份组成：

1. 由测量向量 $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n)$ 构成一个测量空间 R^N ，其中 $N=mn$ ；
2. 模型被视为测量空间的子集 $M \subset R^N$ 。如果测量向量 $\mathbf{X} \in M$ ，则称 \mathbf{X} 满足此模型 M 。通

常， M 是测量空间的子流形（或代数族）。模型的维数等于参数化子流形 M 的最小参数的数目 D ， R^D 称为模型的参数空间。通常称 $d=N-D$ 为模型的余维数。

3. 给定测量空间中的一个测量向量 X ，估计问题就是求解距离 X 最近且满足模型 M 的向量和相关模型参数。

为了解模型估计问题的这种抽象表述，下面给出两个例子。

单应矩阵 假定 $\{x_j = (x_j, y_j, 1)^T \leftrightarrow x'_j = (x'_j, y'_j, 1)^T, j=1, 2, \dots, n\}$ 是来自单应矩阵 H 的一组测量点对应，共有 $4n$ 个测量值，它们构成一个 $4n$ 维测量向量 $X = (x_1, y_1, x'_1, y'_1, \dots, x_n, y_n, x'_n, y'_n)^T$ ，测量空间是 R^{4n} 。

在 R^{4n} 中，并非每一个点都与单应相关联。对单应 H ，仅对所有 $1 \leq j \leq n$ 都满足 $x'_j = Hx_j$ 的点对应 $\{x_j \leftrightarrow x'_j\}$ 组成的向量 X 才能构成该模型 M_H ，它是 R^{4n} 中的子流形（确切地说它是代数族）。

考虑 M_H 的维数。为此，在第一幅图像上任取 n 个点 \hat{x}_j ，单应 H 也可以任意选择。一旦它们被选定，仅当第二幅图像上的点满足 $\hat{x}'_j = H\hat{x}_j$ 时，相应向量 \hat{X} 才能满足该模型（即 $\hat{X} \in M_H$ ）。因此， M_H 中的点需要 $2n+8$ 个参数来确定，这 $2n+8$ 个参数由 n 个点 \hat{x}_j 的 $2n$ 个坐标和 H 的 8 个独立参数所形成，所以模型 M_H 的维数是 $D = 2n + 8$ ，余维数是 $d = 4n - D = 2n - 8$ 。

给定测量向量 X ，估计单应是求解距离 X 最近且满足模型 M_H 的向量 \hat{X} 和 H ，即求解下述最小化问题：

$$\begin{cases} \min_{\{H, \hat{X}\}} \|\hat{X} - X\|^2 \\ \text{subject to } \hat{X} \in M_H \end{cases} \quad (16.3.15)$$

由于 $\|\hat{X} - X\|^2 = \sum_{j=1}^n \|\hat{x}_j - x_j\|^2 + \|\hat{x}'_j - x'_j\|^2$ ， $\hat{X} \in M_H \Leftrightarrow \hat{x}'_j \times H\hat{x}_j = 0$ ，所以上述最小化问题与几何最小化问题(16.2.10)是一致的。

基本矩阵 假定 $\{x_j = (x_j, y_j, 1)^T \leftrightarrow x'_j = (x'_j, y'_j, 1)^T, j=1, 2, \dots, n\}$ 是来自基本矩阵 F 的一组测量点对应，共有 $4n$ 个测量值，它们构成一个 $4n$ 维测量向量 $X = (x_1, y_1, x'_1, y'_1, \dots, x_n, y_n, x'_n, y'_n)^T$ ，因此 R^{4n} 是测量空间。

与单应矩阵一样，在 R^{4n} 中并非每一个点都与基本矩阵相关联。对基本矩阵 F ，仅对所有 $1 \leq j \leq n$ 都满足 $x_j'^T Fx_j = 0$ 的点对应形成的向量 X 才能构成该模型 M_F ，它也是 R^{4n} 中的代数族。考虑 M_F 的维数。为此，在第一幅图像上任取 n 个点 \hat{x}_j ，基本矩阵 F 也可以任意选择。一旦它们被选定，仅当第二幅图像上的点满足 $\hat{x}_j'^T F\hat{x}_j = 0$ 时，相应向量 \hat{X} 才能满足该模型（即 $\hat{X} \in M_F$ ）。与单应不同的是在极线 $F\hat{x}_j$ 上的点均可以作为 \hat{x}'_j ，因此 \hat{x}'_j 的第一个坐标分量可以任意选取。于是， M_F 中的点由 $3n+7$ 个参数来确定，其中 $3n$ 个参数来自于第一幅图像上 n 个点的坐标分量和第二幅图像上 n 个点的第一个坐标分量，另外的 7 个参数是基本矩阵的 7 个独立参数。所以，模型 M_F 的维数是 $D = 3n + 7$ ，余维数是 $d = 4n - D = n - 7$ 。

给定测量向量 \mathbf{X} ，估计基本矩阵是求解距离 \mathbf{X} 最近且满足模型 M_F 的向量 $\hat{\mathbf{X}}$ 和 F ，即求解下述最小化问题：

$$\begin{cases} \min_{\{F, \hat{\mathbf{X}}\}} \|\hat{\mathbf{X}} - \mathbf{X}\|^2 \\ \text{subject to } \hat{\mathbf{X}} \in M_F \end{cases} \quad (16.3.16)$$

由于 $\|\hat{\mathbf{X}} - \mathbf{X}\|^2 = \sum_{j=1}^n \|\hat{\mathbf{x}}_j - \mathbf{x}_j\|^2 + \|\hat{\mathbf{x}}'_j - \mathbf{x}'_j\|^2$ ， $\hat{\mathbf{X}} \in M_F \Leftrightarrow \hat{\mathbf{x}}_j^T F \hat{\mathbf{x}}_j = 0$ ，所以最小化问题(16.3.16)与几何最小化问题(16.2.6)是一致的。

从上面两个例子，可以看出在一般抽象模型中有两类参数，一类与测量空间有关，这类参数随测量空间的维数增加而增多，另一类参数不随测量空间的维数变化而发生变化，仅与估计对象有关。后一类参数称为模型的本质参数。

根据一般抽象模型的表述，一般估计问题涉及到一个从参数空间 R^D 到测量空间 R^N 一个映射 g 。令 $\bar{\mathbf{X}} \in R^N$ ，且存在一个参数向量 $\bar{\Theta} \in R^D$ 使得 $g(\bar{\Theta}) = \bar{\mathbf{X}}$ 。例如，在单应估计中，

$$\bar{\Theta} = (\bar{x}_1, \bar{y}_1, \dots, \bar{x}_n, \bar{y}_n, \bar{H})^T,$$

$$\bar{\mathbf{X}} = g(\bar{\Theta}) = (\bar{x}_1, \bar{y}_1, \dots, \bar{x}_n, \bar{y}_n, \bar{h}^{1T} \bar{\mathbf{x}}_1 / \bar{h}^{3T} \bar{\mathbf{x}}_1, \bar{h}^{2T} \bar{\mathbf{x}}_1 / \bar{h}^{3T} \bar{\mathbf{x}}_1, \dots, \bar{h}^{1T} \bar{\mathbf{x}}_n / \bar{h}^{3T} \bar{\mathbf{x}}_n, \bar{h}^{2T} \bar{\mathbf{x}}_n / \bar{h}^{3T} \bar{\mathbf{x}}_n)^T.$$

当参数向量 Θ 在 $\bar{\Theta}$ 周围变化时，映射 g 的值 $g(\Theta)$ 在测量空间中形成过点 $g(\bar{\Theta}) = \bar{\mathbf{X}}$ 的曲面片 S_M ，如图 16.3.1 所示。它是模型 M 的子集，其维数等于模型的维数。这样，通过模型参数可以将模型进行函数化。最大似然估计就是在曲面片 S_M 上靠近真值 $\bar{\mathbf{X}}$ 的周围寻找一点 $\hat{\mathbf{X}}$ 使得它距离测量向量 \mathbf{X} 最近，将最大似然估计算法所返回的向量 $\hat{\mathbf{X}}$ 称为校正向量，由于假定了测量向量服从各向同性高斯分布 $N(\bar{\mathbf{X}}, \sigma^2 I_N)$ ，最大似然算法与几何最小化算法等价，所以校正向量通常也称为几何校正向量。

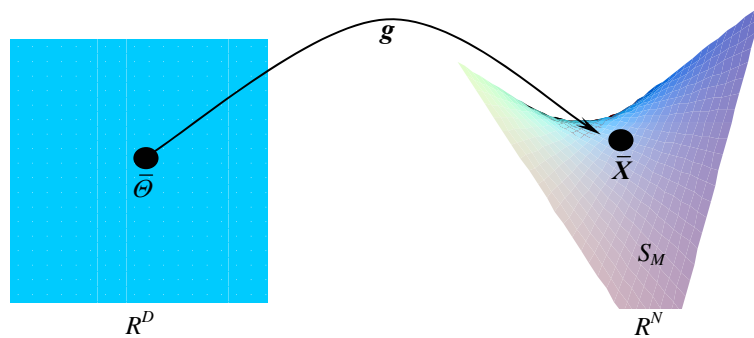


图 16.3.1 模型的函数化

现在假定曲面片 S_M 在 $\bar{\mathbf{X}}$ 的邻域内可由一个平面很好地近似，至少在 $\bar{\mathbf{X}}$ 周围处于 σ 一邻域内是如

此。在这个假定下，校正向量 \hat{X} 是测量向量到点 \bar{X} 的切平面上的垂足，即误差向量 $\hat{X} - \bar{X}$ 是向量 $X - \bar{X}$ 在点 \bar{X} 的切平面上的投影，而残差向量 $\hat{X} - X$ 是 $X - \bar{X}$ 在点 \bar{X} 的法平面上的投影，如图 16.3.2 所示，其中 \bar{X} 是未知的。但是，我们知道 $X - \bar{X} \sim N(0, \sigma^2 I_N)$ ，因切平面的维数等于模型的维数 D ，法平面的维数等于模型的余维数 $N - D$ ，所以，根据命题 16.3.1，通过正交变换知 $X - \bar{X} \sim N(0, \sigma^2 I_{N-D})$ ， $\hat{X} - \bar{X} \sim N(0, \sigma^2 I_D)$ 。

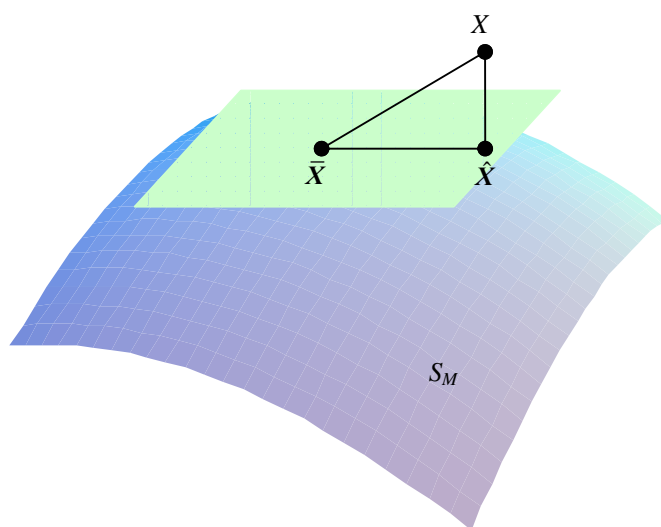


图 16.3.2 切平面局部近似 S_M 。 X ， \hat{X} 和 \bar{X} 分别是测量值，估计值和真值

根据上面的讨论，我们有下述命题：

命题 16.3.1 如果测量向量 $X \sim N(\bar{X}, \sigma^2 I_N)$ ，模型维数为 D ，则

(1) 最大似然估计的 RMS 残差期望值是

$$\varepsilon_{res} = E(\|\hat{X} - X\|^2 / N)^{1/2} = (1 - D / N)^{1/2} \sigma \tag{16.3.16}$$

(2) 最大似然估计的 RMS 估计误差期望值是

$$\varepsilon_{est} = E(\|\hat{X} - \bar{X}\|^2 / N)^{1/2} = (D / N)^{1/2} \sigma \tag{16.3.17}$$

其中， \hat{X} ， \bar{X} 分别是估计值和真值。

现在，将上述命题应用到单应矩阵、基本矩阵和摄像机矩阵估计问题。

单应矩阵 对于单应矩阵估计， $N = 4n$ ， $D = 2n + 8$ ，其中是点对应数，所以最大似然估计的 RMS

残差和估计误差期望值分别是

$$\varepsilon_{res} = \left(\frac{n-4}{2n} \right)^{1/2} \sigma \quad (16.3.18)$$

$$\varepsilon_{res} = \left(\frac{n+4}{2n} \right)^{1/2} \sigma \quad (16.3.19)$$

可以看出，残差与估计误差的渐近值均为 $\sigma/\sqrt{2}$ 。考评一个具体算法，如果在模拟实验中，当匹配点增加时，残差与估计误差都稳定在 $\sigma/\sqrt{2}$ 附近，可认为它是一种好算法，至少在最大似然的意义下是如此。另一方面，当 $n=4$ 时残差为零，这是自然的，因为对任意 4 对匹配点，不论测量误差有多大，总能找到一个单应满足这些匹配点。因此，我们不能仅从残差的指标来考评算法的优劣，试想还有什么比残差为零会更好呢？

如果测量误差仅发生在一幅图像上，例如第二幅图像，即第一幅图像上的测量是精确的，则此时 $N=2n, D=8$ ，因给定第一幅图像上的 n 个点 x_j （它是精确的不需要校正），任意选择一个单应 H ，由 $\hat{x}'_j = H\hat{x}_j$ 构成的向量 $2n$ 维向量 X 总满足模型，所以 $D=8$ 。因此，

$$\varepsilon_{res} = (1-4/n)^{1/2} \sigma \quad (16.3.20)$$

$$\varepsilon_{res} = (4/n)^{1/2} \sigma \quad (16.3.21)$$

由上式，残差随匹配点增加而递增地趋于标准差，这是自然的。估计误差随匹配点增加而递减地趋于 0，这也是自然的，因为第一幅图像上的测量是精确的，匹配点增加导致单应估计越来越准确，校正点将趋于真值。

基本矩阵 对于基本矩阵估计， $N=4n, D=3n+7$ ，其中 n 是点对数，所以最大似然估计的 RMS 残差和估计误差期望值分别是

$$\varepsilon_{res} = \left(1 - \frac{3n+7}{4n} \right)^{1/2} \sigma \quad (16.3.22)$$

$$\varepsilon_{res} = \left(\frac{3n+7}{4n} \right)^{1/2} \sigma \quad (16.3.23)$$

可以看出，估计误差的渐近值均为 $\sqrt{3}\sigma/2$ ，大于单应矩阵的估计误差的渐近值，这说明基本矩阵估计对数据噪声比单应估计更敏感，实践也支持这个结论。当 $n=7$ 时残差为零，这是自然的，因为对任意 7 对匹配点，不论测量误差有多大，总能找到一个基本矩阵满足这些匹配点。

摄像机矩阵 假定三维空间点没有测量误差，这在实践中是靠得住的，因为标定物上的空间点通常已经过精确测量。对应图像点的测量误差满足零均值方差 σ 的各向同性分布。在这个估计问题

中，测量空间的维数是 $N = 2n$ ，模型参数是摄像机矩阵的 11 个独立参数，即 $D = 11$ 。根据命题 16.3.1，最大似然估计的 RMS 残差和估计误差期望值分别是

$$\varepsilon_{res} = (1 - 11/2n)^{1/2} \sigma \quad (16.3.22)$$

$$\varepsilon_{est} = (11/2n)^{1/2} \sigma \quad (16.3.23)$$

16.3.4 (本质) 参数的协方差

上节我们对最大似然估计的残差与估计误差进行了分析，但在实际应用中，更关心的是估计对象的准确度，如单应矩阵、基本矩阵、摄像机矩阵等估计的准确度。在视觉中估计对象通常涉及到的是变换。变换估计的不确定性，不像残差与估计误差那样仅涉及到点对应的数目及匹配的准确性，它还和点配置有关。用接近于退化配置数据所得到的估计是非常不可靠的，例如用接近一条直线的数据，不论使用什么方法都很难对单应做出好的估计，因为此时缺少与直线垂直维度的信息。变换估计的不确定性通常由描述它的参数向量的协方差矩阵（简称变换的协方差矩阵）来衡量，下面介绍如何通过测量向量的协方差矩阵计算变换的协方差矩阵。

模型参数中的本质参数是描述变换的参数，而另外的参数是由测量向量的部分分量所构成的，如果能求出模型参数的协方差矩阵，我们就可以分解出变换的协方差矩阵。先考虑本质参数数目是变换的最小化参数化数目，即我们暂不考虑过参数化的情况。此时，模型函数化映射 $g: R^D \rightarrow R^N$ 的值域 $g(R^D) = M$ ，因此 $\dim M = \dim g(R^D) = D$ 。最大似然估计算法从测量向量 X 不仅返回校正向量 \hat{X} 同时也返回了相应的参数向量 $\hat{\Theta}$ ，下面利用映射来描述这个过程。返回校正向量 \hat{X} 的过程确定了一个映射 $mel: R^N \rightarrow M$ ，即

$$\hat{X} = mel(X) \equiv \arg \min \{ \|X - Y\|_X : Y \in S_M \} \quad (16.3.24)$$

其中 S 是 X 的协方差矩阵。假定 g 可逆（实际上，只需局部可逆），并记为 $g^{-1}: S_M \rightarrow R^D$ ，其中 $S_M \subset M$ 是包含均值 \bar{X} 的曲面片，则复合映射 $g^{-1} \circ mel: R^N \rightarrow R^D$ ，

$$\hat{\Theta} = g^{-1} \circ mel(X) \quad (16.3.25)$$

实现了最大似然估计算法返回参数向量 $\hat{\Theta}$ 的过程。根据式(16.3.25)，我们可以由 X 的协方差矩阵对 Θ 的协方差矩阵进行估计。

仿射变换 考虑 g 为仿射变换的情况，此时 $g^{-1} \circ mel$ 也是一个仿射变换。为了说明这一点，令 $g(\Theta) = A(\Theta - \bar{\Theta}) + g(\bar{\Theta})$ ，其中 $\bar{X} = g(\bar{\Theta})$ 是 X 的均值。因 $\dim g(R^D) = D$ ，所以 A 是列满秩的。最大似然估计 \hat{X} 最小化 $\|X - \hat{X}\|_X = \|X - g(\hat{\Theta})\|_X$ ，由于

$$\|X - g(\hat{\Theta})\|_Y = \|(X - \bar{X}) - A(\hat{\Theta} - \bar{\Theta})\|_Y = \|\Sigma^{-1/2}(X - \bar{X}) - \Sigma^{-1/2}A(\hat{\Theta} - \bar{\Theta})\|$$

所以，当 $\hat{\Theta} - \bar{\Theta} = ((\Sigma^{-1/2}A)^T \Sigma^{-1/2}A)^{-1} (\Sigma^{-1/2}A)^T \Sigma^{-1/2}(X - \bar{X}) = \underbrace{(A^T \Sigma^{-1}A)^{-1} A^T \Sigma^{-1}}_W (X - \bar{X})$ 时，上式达到

最小。于是，我们有

$$g^{-1} \circ mel(X) = \hat{\Theta} = W(X - \bar{X}) + \bar{\Theta} = W(X - \bar{X}) + g^{-1}(\bar{X}) = W(X - \bar{X}) + g^{-1} \circ eml(\bar{X})$$

即， $g^{-1} \circ mel$ 是仿射变换。所以，

$$E[\hat{\Theta}] = \bar{\Theta} = g^{-1}(\bar{X}) \quad (16.3.26)$$

$$\text{Cov}[\hat{\Theta}] = W \Sigma W^T = ((A^T \Sigma^{-1}A)^{-1} A^T \Sigma^{-1}) \Sigma ((A^T \Sigma^{-1}A)^{-1} A^T \Sigma^{-1})^T = (A^T \Sigma^{-1}A)^{-1} \quad (16.3.27)$$

总结上述讨论，我们有下述命题。

命题 16.3.2 假定测量向量 X 是一个均值为 \bar{X} 、协方差为 Σ 的随机变量。若模型函数化映射 $g: R^D \rightarrow R^N$ 是一个仿射映射 $g(\Theta) = A(\Theta - \bar{\Theta}) + g(\bar{\Theta})$ ，且 $\dim g(R^D) = D$ ，则最大似然估计算法返回的参数向量 $\hat{\Theta}$ 是均值为 $g^{-1}(\bar{X})$ 、协方差为 $(A^T \Sigma^{-1}A)^{-1}$ 的随机变量。

非线性变换 对于非线性情况不像仿射变换那样，很难对最大似然算法返回的参数向量的协方差矩阵作准确估计，一种通常使用的方法是用仿射变换对模型函数化映射作局部近似，从而给出协方差矩阵的一阶近似估计。设 $g: R^D \rightarrow R^N$ 是可微映射，它在点 $\bar{\Theta}$ 处的 Jacobian 矩阵记为 J ，即 $J = \frac{\partial g}{\partial \Theta}(\bar{\Theta})$ 。如果 $\text{rank}(J) = D$ ，则 g 在 $\bar{\Theta}$ 的邻域内是一一映射，即局部可逆。此时，可以使用仿射变换 $g_a(\Theta) = J(\Theta - \bar{\Theta}) + g(\bar{\Theta})$ 在该邻域对 g 作一阶近似。于是，应用命题 16.3.2，我们就得到下面的命题。

命题 16.3.3 假定测量向量 X 是一个均值为 \bar{X} 、协方差为 Σ 的随机变量。若模型函数化映射 g 是可微的且局部可逆的，则最大似然估计算法返回的参数向量 $\hat{\Theta}$ 的均值和协方差的一阶近似估计分别是 $g^{-1}(\bar{X})$ 和 $(J^T \Sigma^{-1}J)^{-1}$ 。

注 我们并不知道测量向量的真值 \bar{X} ，因而也无法知道 $\bar{\Theta}$ ，即无法确定映射 g 的 Jacobian 矩阵在点 $\hat{\Theta}$ 的值。实际应用中，通常是以最大似然估计返回的 $\hat{\Theta}$ 代替 $\bar{\Theta}$ 得到 Jacobian 矩阵在点 $\bar{\Theta}$ 的近似值，这样做是基于数据真值与被估计值非常接近的假设，以致协方差矩阵没有发生本质的变化。

过参数化

现在，考虑过参数化的情形，我们的目的是将命题 16.3.3 推广到过参数化。在这种情形下， $g: R^D \rightarrow R^N$ 不是局部可逆的。例如，在估计单应时使用 9 维向量，即单应矩阵的 9 个元素，则映射 g 就不是局部可逆的，因为对任意非零常数，等式 $g(p, h) = g(p, sh)$ 均成立，它们表示测量空间

中的同一个点，这里模型参数 $\Theta = (\mathbf{p}, \mathbf{h})$ 被分解为测量点坐标分量 \mathbf{p} 和单应参数 \mathbf{h} 。在过参数化情形下，映射 $g: R^D \rightarrow R^N$ 的 Jacobian 矩阵 J 不是列满秩的，它的秩 $d < D$ ，因此 D 阶方阵 $J^T \Sigma^{-1} J$ 不可逆，因而我们不能直接应用命题 16.3.3。

如果没有进一步约束，则最大似然估计算法返回的本质参数，如单应参数 $\hat{\mathbf{h}}$ ，可以相差任意非零常数因子而无界的变化，即本质参数有无穷大的方差，这显然是不合理的。因此，对参数向量 \mathbf{P} 进行进一步约束是必要的。例如在单应中，通常使用的约束是 $\|\mathbf{h}\|=1$ ，在这样的约束下 $S_D = \{(\mathbf{p}, \mathbf{h}) : \|\mathbf{h}\|=1\}$ 是参数空间中的一个 $2n+8$ 维柱面。一般地，我们可假定被估计的参数向量 \mathbf{P} 在参数空间 R^D 中的某个 d 维流形 S_D 上， g 在这个流形上的限制 $g|_{S_D}$ 是局部可逆的，且 $g(S_D) = S_M$ 。由于 $g|_{S_D}$ 是局部可逆的，因此在流形 S_D 的点 $\bar{\Theta}$ 邻域 $U_{\bar{\Theta}}$ 内， $g|_{S_D}$ 实现从 $U_{\bar{\Theta}}$ 到 $O_{\bar{x}} = g|_{S_D}(U_{\bar{\Theta}}) \subset S_M$ 的一一映射，逆映射仍记为 g^{-1} 。定义可微映射 $h: R^d \rightarrow R^D$ ，它将 R^d 在点 \bar{x} 邻域 $W_{\bar{x}}$ 一一映射到流形 S_D 在点 $\bar{\Theta}$ 的邻域 $U_{\bar{\Theta}}$ ，且 $h(\bar{x}) = \bar{\Theta}$ 。如图 16.3.3 所示。

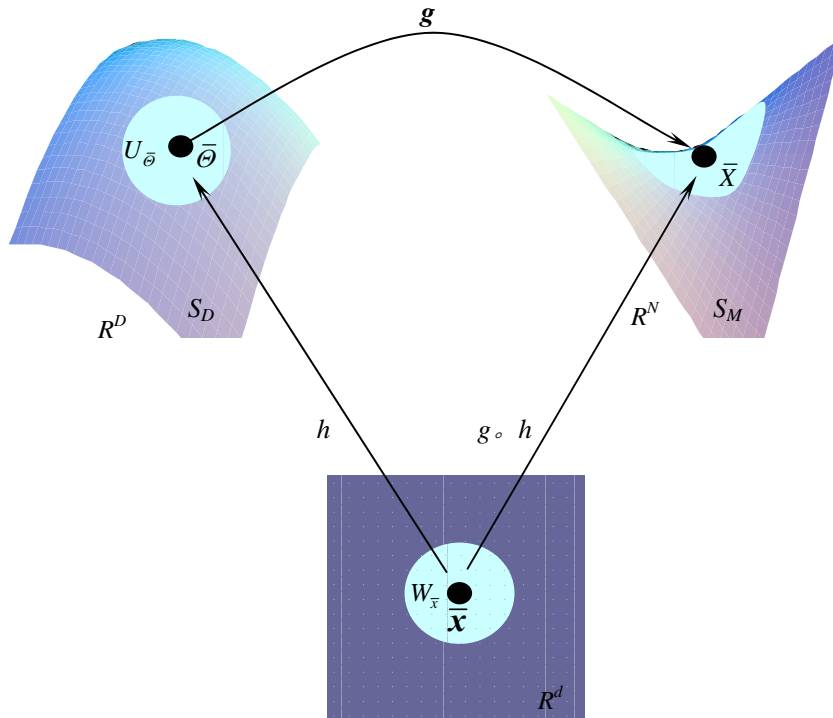


图 16.3.3 参数空间中约束流形 S_D ，映射 h 将 R^d 的邻域 $W_{\bar{x}}$ 一一映到 S_D 点 $\bar{\Theta}$ 的邻域 $U_{\bar{\Theta}}$ 且 $h(\bar{x}) = \bar{\Theta}$

复合映射 $g \circ h$ 在邻域 $W_{\bar{x}}$ 内是可逆的，并且满足命题 16.3.3 的条件，于是令 g, h 的 Jacobian 矩阵分别为 J_g, J_h ，则 $g \circ h$ 的 Jacobian 矩阵 $J_{g \circ h} = J_g J_h$ ，且在 $W_{\bar{x}}$ 内具有协方差矩阵

$(J_h^T J_g^T \Sigma^{-1} J_g J_h)^{-1}$ ，其中 Σ 测量向量 \mathbf{X} 的协方差矩阵。再通过映射 h 将它传递到 $U_{\bar{p}}$ 上，所以参数 Θ 具有协方差矩阵：

$$\text{Cov}[\Theta] = J_h (J_h^T J_g^T \Sigma^{-1} J_g J_h)^{-1} J_h^T \tag{16.3.28}$$

现在的问题是如何确定 J_h ，即如何定义 h 。为此考虑流形 S_D 在点 $\bar{\Theta}$ 切空间 $T_{\bar{\Theta}}$ ：因 $\dim(S_D) = d$ ，所以切空间 $T_{\bar{\Theta}}$ 也是 d 维的。令 $\{t_1, t_2, \dots, t_d\}$ 是 $T_{\bar{p}}$ 的一组基向量，如图 16.3.4 所示，则切空间 $T_{\bar{\Theta}}$ 可表示为 $T_{\bar{\Theta}} = \{T\mathbf{x} + \bar{\Theta} : \mathbf{x} \in R^d\}$ ， $T = (t_1, t_2, \dots, t_d)$ 。定义 $h: R^d \rightarrow R^D, h(\mathbf{x}) = \pi_{N, S_D}(T\mathbf{x} + \bar{P})$ ，其中 π_{N, S_D} 表示切空间 $T_{\bar{p}}$ 沿法空间 $N_{\bar{p}}$ 到流形 S_D 上的投影，则 h 为满足上述条件，且 $J_h = T$ 。注意到对任意 d 阶可逆矩阵 B ，我们有

$$(J_h B)((J_h B)^T J_g^T \Sigma^{-1} J_g (J_h B))^{-1} (J_h B)^T = J_h (J_h^T J_g^T \Sigma^{-1} J_g J_h)^{-1} J_h^T$$

所以，式(16.3.28)中的矩阵 J_h 可以选择流形 S_D 切空间 $T_{\bar{p}}$ 的任一组基向量所构成的 $D \times d$ 矩阵。

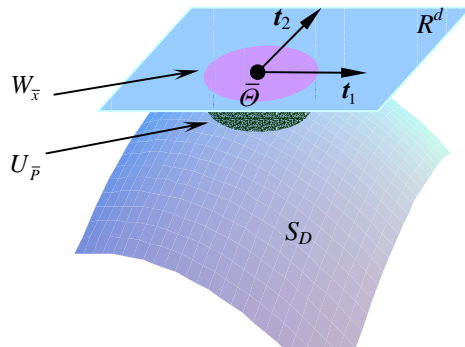


图 16.3.4 切空间 $T_{\bar{p}}$ 沿法空间 $N_{\bar{p}}$ 到流形 S_D 上的投影 π_{N, S_D} 实现映射 h ，它的 Jacobian 矩阵的列向量张成切空间 $T_{\bar{p}}$ 。

总结上述讨论，我们有下述命题。

命题 16.3.4 (过参数化) 假定测量空间 R^N 的向量 \mathbf{X} 是一个均值为 $\bar{\mathbf{X}}$ 、协方差为 Σ 的随机变量，并且模型函数化映射 g 是可微的。令 S_D 是嵌入参数空间 R^D 过点 $\bar{\Theta}$ 的 d 维光滑流形，使得 g 在 S_D 上的点 $\bar{\Theta}$ 是局部可逆的。若 $D \times d$ 矩阵 T 的列向量张成 S_D 的切空间 $T_{\bar{\Theta}}$ ，则最大似然估计算法返回的参数向量 $\hat{\Theta}$ 的协方差矩阵的一阶近似估计是

$$\text{Cov}[\hat{\Theta}] = T(T^T J_g^T \Sigma^{-1} J_g T)^{-1} T^T \tag{16.3.29}$$

上述命题给出了参数在任何约束下的协方差矩阵一阶近似估计，也就是说参数约束的不同形式可导致矩阵 T 的不同选择。其中，有一种约束非常重要，它能简化协方差矩阵的计算。这种重要的约束是“约束曲面 S_D 局部地正交于 g 的 Jacobian 矩阵 J_g 零空间”。例如，约束 $\|\Theta\|=1$ ，它是齐次参数中的一种常用的约束，就满足“曲面 S_D 局部地正交于 g 的 Jacobian 矩阵 J_g 零空间”这一条件。在约束 $\|\Theta\|=1$ 下，曲面 S_D 是一个单位球面，任何一点的切平面都垂直该参数向量，另一方面，函数 $g(\Theta)$ 不因齐次因子而变化，即径向向量 $\overline{O\Theta}$ 是 J_g 零空间，即 J_g 的零向量是曲面 S_D 在 Θ 的法向量，因此曲面 S_D 局部地正交于 g 的 Jacobian 矩阵 J_g 零空间。如果两个矩阵 A, B 有相同的左零空间，即 $\{\mathbf{y} \mid \mathbf{y}^T A = 0\} = \{\mathbf{y} \mid \mathbf{y}^T B = 0\}$ ，则必有 $A^+ = B(B^T A B)^{-1} B^T$ 。因此，在上述约束下，式(16.3.29)右端可简化为 $T(T^T J_g^T \Sigma^{-1} J_g T)^{-1} T^T = (J_g^T \Sigma^{-1} J_g)^+$ 。于是，我们得到下述命题。

命题 16.3.5 假定测量空间 R^N 的向量 X 是一个均值为 \bar{X} 、协方差为 Σ 的随机变量，并且模型函数化映射 g 可微。令 S_D 是嵌入参数空间 R^D 过点 $\bar{\Theta}$ 的 d 维光滑流形，使得在 S_D 上 g 在点 $\bar{\Theta}$ 是局部可逆的，并且曲面 S_D 局部地正交于 g 的 Jacobian 矩阵 J_g 零空间，则最大似然估计算法返回的参数向量 $\hat{\Theta}$ 的协方差矩阵的一阶近似估计为

$$\text{Cov}[\hat{\Theta}] = (J_g^T \Sigma^{-1} J_g)^+ \tag{16.3.30}$$

应用举例

上面讨论了参数协方差矩阵估计问题，在数学上有一定的难度。为了帮助理解与应用，我们介绍如何计算摄像机矩阵的协方差矩阵。假定三维空间点没有测量误差，即空间点已经过精确测量。每个图像点的测量误差均满足零均值协方差矩阵 Σ_i 的高斯分布。在估计问题中，有 n 个空间点到图像点的对应：

$$\left\{ M_j = \begin{pmatrix} \tilde{M}_j \\ 1 \end{pmatrix} \leftrightarrow m_j = \begin{pmatrix} \tilde{m}_j \\ 1 \end{pmatrix}, j = 1, 2, \dots, n \right\}$$

因此，测量向量 $X = (\tilde{m}_1^T, \tilde{m}_2^T, \dots, \tilde{m}_n^T)^T$ 服从均值 \bar{X} 、协方差矩阵 $\Sigma = \text{diag}(\Sigma_1, \Sigma_2, \dots, \Sigma_n)$ 的 $2n$ 维高斯分布，令 \hat{X}, \hat{P} 是最大似然估计算法的返回值，其中 \hat{X} 是测量向量的校正， \hat{P} 是摄像机矩阵的估计值。我们的目的是计算 \hat{P} 的协方差矩阵，下面分摄像机为仿射的和射影的两种情况来考虑此问题。

仿射摄像机 因仿射摄像机矩阵形式是 $P = \begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix}$ ，记 $p = (p^1, p^2)^T$ ，其中 p^1, p^2 是 P 的前两行。于是，映射 g 由下式确定：

$$g(\mathbf{p}) = \begin{pmatrix} \begin{pmatrix} \mathbf{M}_1^T & \\ & \mathbf{M}_1^T \end{pmatrix} \mathbf{p} \\ \begin{pmatrix} \mathbf{M}_2^T & \\ & \mathbf{M}_2^T \end{pmatrix} \mathbf{p} \\ \vdots \\ \begin{pmatrix} \mathbf{M}_n^T & \\ & \mathbf{M}_n^T \end{pmatrix} \mathbf{p} \end{pmatrix} = \begin{pmatrix} I_2 \otimes \mathbf{M}_1^T \\ I_2 \otimes \mathbf{M}_2^T \\ \vdots \\ I_2 \otimes \mathbf{M}_n^T \end{pmatrix} \mathbf{p} \quad (16.3.31)$$

它是从参数空间到测量空间的 8 维仿射变换，它的 Jacobian 矩阵为

$$\frac{\partial g}{\partial \mathbf{p}} = \begin{pmatrix} I_2 \otimes \mathbf{M}_1^T \\ I_2 \otimes \mathbf{M}_2^T \\ \vdots \\ I_2 \otimes \mathbf{M}_n^T \end{pmatrix}$$

根据命题 16.3.2， $\hat{\mathbf{P}}$ 的协方差矩阵为

$$\begin{aligned} \text{Cov}[\hat{\mathbf{P}}] &= \begin{pmatrix} I_2 \otimes \mathbf{M}_1^T \\ I_2 \otimes \mathbf{M}_2^T \\ \vdots \\ I_2 \otimes \mathbf{M}_n^T \end{pmatrix}^T \begin{pmatrix} \Sigma_1^{-1} & & & \\ & \Sigma_2^{-1} & & \\ & & \ddots & \\ & & & \Sigma_n^{-1} \end{pmatrix} \begin{pmatrix} I_2 \otimes \mathbf{M}_1^T \\ I_2 \otimes \mathbf{M}_2^T \\ \vdots \\ I_2 \otimes \mathbf{M}_n^T \end{pmatrix}^{-1} \\ &= \left(\sum_{j=1}^n (I_2 \otimes \mathbf{M}_j) \Sigma_j^{-1} (I_2 \otimes \mathbf{M}_n^T) \right)^{-1} \end{aligned} \quad (16.3.32)$$

射影摄像机 射影摄像机矩阵形式是 $\mathbf{P} = \begin{pmatrix} \mathbf{p}^1 \\ \mathbf{p}^2 \\ \mathbf{p}^3 \end{pmatrix}$ ，记 $\mathbf{p} = (\mathbf{p}^1, \mathbf{p}^2, \mathbf{p}^3)^T$ 。于是，映射 g 由下式确定：

$$g(\mathbf{p}) = \begin{pmatrix} \frac{1}{\mathbf{p}^3 \mathbf{M}_1} \begin{pmatrix} \mathbf{M}_1^T & \\ & \mathbf{M}_1^T \end{pmatrix} \begin{pmatrix} \mathbf{p}^{1T} \\ \mathbf{p}^{2T} \end{pmatrix} \\ \frac{1}{\mathbf{p}^3 \mathbf{M}_2} \begin{pmatrix} \mathbf{M}_2^T & \\ & \mathbf{M}_2^T \end{pmatrix} \begin{pmatrix} \mathbf{p}^{1T} \\ \mathbf{p}^{2T} \end{pmatrix} \\ \vdots \\ \frac{1}{\mathbf{p}^3 \mathbf{M}_n} \begin{pmatrix} \mathbf{M}_n^T & \\ & \mathbf{M}_n^T \end{pmatrix} \begin{pmatrix} \mathbf{p}^{1T} \\ \mathbf{p}^{2T} \end{pmatrix} \end{pmatrix} = \begin{pmatrix} \frac{1}{\mathbf{p}^3 \mathbf{M}_1} I_2 \otimes \mathbf{M}_1^T \\ \frac{1}{\mathbf{p}^3 \mathbf{M}_2} I_2 \otimes \mathbf{M}_2^T \\ \vdots \\ \frac{1}{\mathbf{p}^3 \mathbf{M}_n} I_2 \otimes \mathbf{M}_n^T \end{pmatrix} \begin{pmatrix} \mathbf{p}^{1T} \\ \mathbf{p}^{2T} \end{pmatrix} \quad (16.3.33)$$

它是从参数空间到测量空间的 12 维非线性变换。在约束 $\|\mathbf{p}\|=1$ 下，我们可以根据命题 16.3.5 计算 $\hat{\mathbf{P}}$ 的协方差矩阵。因 g 的 Jacobian 矩阵为

$$\frac{\partial g}{\partial \mathbf{p}} = \begin{pmatrix} \frac{1}{\mathbf{p}^3 \mathbf{M}_1} I_2 \otimes \mathbf{M}_1^T & -\frac{1}{(\mathbf{p}^3 \mathbf{M}_1)^2} \begin{pmatrix} \mathbf{p}^1 \mathbf{M}_1 \\ \mathbf{p}^2 \mathbf{M}_1 \end{pmatrix} \mathbf{M}_1^T \\ \frac{1}{\mathbf{p}^3 \mathbf{M}_2} I_2 \otimes \mathbf{M}_2^T & -\frac{1}{(\mathbf{p}^3 \mathbf{M}_2)^2} \begin{pmatrix} \mathbf{p}^1 \mathbf{M}_2 \\ \mathbf{p}^2 \mathbf{M}_2 \end{pmatrix} \mathbf{M}_2^T \\ \vdots & \vdots \\ \frac{1}{\mathbf{p}^3 \mathbf{M}_n} I_2 \otimes \mathbf{M}_n^T & -\frac{1}{(\mathbf{p}^3 \mathbf{M}_n)^2} \begin{pmatrix} \mathbf{p}^1 \mathbf{M}_n \\ \mathbf{p}^2 \mathbf{M}_n \end{pmatrix} \mathbf{M}_n^T \end{pmatrix}$$

所以， $\hat{\mathbf{P}}$ 点的 Jacobian 矩阵为

$$\frac{\partial g}{\partial \mathbf{p}}(\hat{\mathbf{p}}) = \begin{pmatrix} \frac{1}{\hat{\mathbf{p}}^3 \mathbf{M}_1} I_2 \otimes \mathbf{M}_1^T & -\frac{1}{(\hat{\mathbf{p}}^3 \mathbf{M}_1)^2} \begin{pmatrix} \hat{\mathbf{p}}^1 \mathbf{M}_1 \\ \hat{\mathbf{p}}^2 \mathbf{M}_1 \end{pmatrix} \mathbf{M}_1^T \\ \frac{1}{\hat{\mathbf{p}}^3 \mathbf{M}_2} I_2 \otimes \mathbf{M}_2^T & -\frac{1}{(\hat{\mathbf{p}}^3 \mathbf{M}_2)^2} \begin{pmatrix} \hat{\mathbf{p}}^1 \mathbf{M}_2 \\ \hat{\mathbf{p}}^2 \mathbf{M}_2 \end{pmatrix} \mathbf{M}_2^T \\ \vdots & \vdots \\ \frac{1}{\hat{\mathbf{p}}^3 \mathbf{M}_n} I_2 \otimes \mathbf{M}_n^T & -\frac{1}{(\hat{\mathbf{p}}^3 \mathbf{M}_n)^2} \begin{pmatrix} \hat{\mathbf{p}}^1 \mathbf{M}_n \\ \hat{\mathbf{p}}^2 \mathbf{M}_n \end{pmatrix} \mathbf{M}_n^T \end{pmatrix} = \begin{pmatrix} \frac{1}{\hat{w}_1} I_2 \otimes \mathbf{M}_1^T & -\frac{1}{\hat{w}_1} \hat{\mathbf{m}}_1 \mathbf{M}_1^T \\ \frac{1}{\hat{w}_2} I_2 \otimes \mathbf{M}_2^T & -\frac{1}{\hat{w}_2} \hat{\mathbf{m}}_2 \mathbf{M}_2^T \\ \vdots & \vdots \\ \frac{1}{\hat{w}_n} I_2 \otimes \mathbf{M}_n^T & -\frac{1}{\hat{w}_n} \hat{\mathbf{m}}_n \mathbf{M}_n^T \end{pmatrix}$$

其中 $\hat{w}_j \begin{pmatrix} \hat{\mathbf{m}}_j \\ 1 \end{pmatrix} = \hat{\mathbf{P}} \mathbf{M}_j$ 。根据命题 16.3.5， $\hat{\mathbf{P}}$ 的协方差矩阵为

$$\begin{aligned} \text{Cov}[\hat{\mathbf{P}}] &= \begin{pmatrix} \frac{1}{\hat{w}_1} I_2 \otimes \mathbf{M}_1^T & -\frac{1}{\hat{w}_1} \hat{\mathbf{m}}_1 \mathbf{M}_1^T \\ \frac{1}{\hat{w}_2} I_2 \otimes \mathbf{M}_2^T & -\frac{1}{\hat{w}_2} \hat{\mathbf{m}}_2 \mathbf{M}_2^T \\ \vdots & \vdots \\ \frac{1}{\hat{w}_n} I_2 \otimes \mathbf{M}_n^T & -\frac{1}{\hat{w}_n} \hat{\mathbf{m}}_n \mathbf{M}_n^T \end{pmatrix} \begin{pmatrix} \Sigma_1^{-1} & & & \\ & \Sigma_2^{-1} & & \\ & & \ddots & \\ & & & \Sigma_n^{-1} \end{pmatrix} \begin{pmatrix} \frac{1}{\hat{w}_1} I_2 \otimes \mathbf{M}_1^T & -\frac{1}{\hat{w}_1} \hat{\mathbf{m}}_1 \mathbf{M}_1^T \\ \frac{1}{\hat{w}_2} I_2 \otimes \mathbf{M}_2^T & -\frac{1}{\hat{w}_2} \hat{\mathbf{m}}_2 \mathbf{M}_2^T \\ \vdots & \vdots \\ \frac{1}{\hat{w}_n} I_2 \otimes \mathbf{M}_n^T & -\frac{1}{\hat{w}_n} \hat{\mathbf{m}}_n \mathbf{M}_n^T \end{pmatrix}^+ \\ &= \begin{pmatrix} \sum_{j=1}^n \frac{1}{\hat{w}_j^2} (I_2 \otimes \mathbf{M}_j) \Sigma_j^{-1} (I_2 \otimes \mathbf{M}_j^T) & -\sum_{j=1}^n \frac{1}{\hat{w}_j^2} (I_2 \otimes \mathbf{M}_j) \Sigma_j^{-1} \hat{\mathbf{m}}_j \mathbf{M}_j^T \\ -\sum_{j=1}^n \frac{1}{\hat{w}_j^2} \mathbf{M}_j \hat{\mathbf{m}}_j^T \Sigma_j^{-1} (I_2 \otimes \mathbf{M}_j^T) & \sum_{j=1}^n \frac{1}{\hat{w}_j^2} \mathbf{M}_j \hat{\mathbf{m}}_j^T \Sigma_j^{-1} \hat{\mathbf{m}}_j \mathbf{M}_j^T \end{pmatrix}^+ \end{aligned} \quad (16.3.34)$$

如果三维空间点 $\tilde{\mathbf{M}}_j$ 测量误差也满足一个零均值、协方差矩阵 A_j 的高斯分布，则模型的测量向

量为 $X = (\tilde{M}_1^T, \tilde{M}_2^T, \dots, \tilde{M}_n^T, \tilde{m}_1^T, \tilde{m}_2^T, \dots, \tilde{m}_n^T)^T$ ，显然它服从一个均值为 \bar{X} 、协方差矩阵为 $\Phi = \text{diag}(\Lambda_1, \Lambda_2, \dots, \Lambda_n, \Sigma_1, \Sigma_2, \dots, \Sigma_n)$ 的 $5n$ 维高斯分布。令 \hat{X}, \hat{P} 是最大似然估计算法的返回值，其中 \hat{X} 是测量向量的校正， \hat{P} 是摄像机矩阵的估计值。下面仅考虑摄像机为仿射情况。不难看出，映射 g 由下式确定：

$$g(\Psi) = \begin{pmatrix} I_{3n} & 0 \\ 0 & M \end{pmatrix} \Psi, \quad M = \begin{pmatrix} I_2 \otimes M_1^T \\ I_2 \otimes M_2^T \\ \vdots \\ I_2 \otimes M_n^T \end{pmatrix} \quad (16.3.35)$$

其中 $\Psi = (\tilde{M}_1^T, \tilde{M}_2^T, \dots, \tilde{M}_n^T, p^T)^T$ 是参数向量。 g 是从参数空间到测量空间的 $3n+8$ 维非线性变换，下面计算它的 Jacobian 矩阵。因

$$\frac{\partial(I_{3n} | 0)\Psi}{\partial\Psi} = (I_{3n} | 0),$$

$$\frac{\partial(0 | M)\Psi}{\partial\Psi} = \left(\underbrace{\begin{pmatrix} A & & \\ & A & \\ & & \ddots \\ & & & A \end{pmatrix}}_D \middle| \underbrace{\begin{pmatrix} I_2 \otimes M_1^T \\ I_2 \otimes M_2^T \\ \vdots \\ I_2 \otimes M_n^T \end{pmatrix}}_M \right) = (D | M),$$

其中 A 是摄像机矩阵左上角的 2×3 子矩阵。所以，映射 g 的 Jacobian 矩阵为

$$\frac{\partial g}{\partial \Theta} = \begin{pmatrix} I_{3n} & 0 \\ D & M \end{pmatrix}$$

$$\text{令 } \hat{D} = \begin{pmatrix} \hat{A} & & \\ & \hat{A} & \\ & & \ddots \\ & & & \hat{A} \end{pmatrix}, \hat{M} = \begin{pmatrix} I_2 \otimes \hat{M}_1^T \\ I_2 \otimes \hat{M}_2^T \\ \vdots \\ I_2 \otimes \hat{M}_n^T \end{pmatrix}, \Lambda = \text{diag}(\Lambda_1, \Lambda_2, \dots, \Lambda_n), \Sigma = \text{diag}(\Sigma_1, \Sigma_2, \dots, \Sigma_n), \text{ 则根据命题}$$

16.3.3， $\hat{\Psi}$ 的协方差矩阵为

$$\text{Cov}[\hat{\Psi}] = \left(\begin{pmatrix} I_{3n} & 0 \\ \hat{D} & \hat{M} \end{pmatrix}^T \begin{pmatrix} \Lambda^{-1} & \\ & \Sigma^{-1} \end{pmatrix} \begin{pmatrix} I_{3n} & 0 \\ \hat{D} & \hat{M} \end{pmatrix} \right)^{-1} = \begin{pmatrix} \Lambda^{-1} + \hat{D}^T \Sigma^{-1} \hat{D} & \hat{D}^T \Sigma^{-1} \hat{M} \\ \hat{M}^T \Sigma^{-1} \hat{D} & \hat{M}^T \Sigma^{-1} \hat{M} \end{pmatrix}^{-1} \quad (16.3.36)$$

它的右下角 8×8 子矩阵是摄像机的协方差矩阵。

17. 鲁棒方法

前面两章给出的求解估计问题的代数方法与几何方法，均假定测量数据点误差来源仅发生在对该点的位置测量，并且服从高斯分布。在实际情况中，测量点误差不可能都服从高斯分布，存在一些测量数据点的数值由于种种原因而严重偏离真值，这些测量数据点对于高斯(或其它类型)误差分布来说是外点(outlier)有时也称为错误点。例如，在估计基本矩阵时错误匹配点就是外点。一旦测量数据存在外点，即使外点的数目很少，如果使用代数方法和几何方法，则得到的估计结果也可能与真实值有极大的偏差。因此，我们必须使用某种方法对外点加以识别，或者说从测量数据中确定出内点(inlier)有时也称为有效点，然后再使用前两章所描述的方法在这些内点中进行模型估计，这就是所谓的鲁棒估计 (Robust Estimation)。在本章，着重介绍 RANSAC 方法和最大后验 RANSAC 方法，另个还简要介绍另两种鲁棒方法：M-估计与最小中值估计。

17.1 RANSAC 估计

RANSAC(RANdom SAMple Consensus)方法是由 Fischler 和 Bolles 于 1981 年所引入的鲁棒方法。最初它被用于三点确定摄像机姿态的估计，现在无论在计算机视觉领域还是在其它学科的估计问题中都有广泛的应用。对于处理大比例的外点，RANSAC 是十分有效的。我们先从直线估计的简单例子来说明 RANSAC 的基本思想。

17.1.1 直线的 RANSAC 估计

在几何上，鲁棒估计一条直线可描述为：给定一组二维测量数据点，寻找一条直线使得测量点到该直线的几何距离的平方和达到最小，即该直线最小化测量点到直线的几何距离平方和，并且使得内点偏离该直线的距离小于 t 个单位。因此，这个问题有两个要求：

1. 用一条直线拟合测量数据点；
2. 根据阈值 t 将测量数据分为内点与外点；

其中，阈值 t 是根据测量噪声而设置的，具体做法将在下文中讨论。

RANSAC 的思想是比较简单的, 主要有以下几步:

1. 随机选择两点(确定一条直线所需要的最小点集); 由这两个点确定一条线 l ;
2. 根据阈值 t , 确定与直线 l 的几何距离小于 t 的数据点集 $S(l)$, 并称它为直线 l 的一致集;
3. 重复若干次随机选择, 得到直线 l_1, l_2, \dots, l_n 和相应的一致集 $S(l_1), S(l_2), \dots, S(l_n)$;
4. 使用几何距离, 求最大一致集的最佳拟合直线, 作为数据点的最佳匹配直线。

如果随机选择的两点中存在外点, 则这两点所确定的直线一般不会有最大的一致集, 所以根据一致集的大小对所估计的直线进行评价有利于获得更好的拟合直线, 如图 17.1.1 所示。正如 Fischler 和 Bolles 所指出: RANSAC 与通常的数据最佳拟合技术相反, 不是用尽可能多的数据点去获得模型的估计, 而是用尽可能少的可行数据并尽量地扩大一致性数据集。

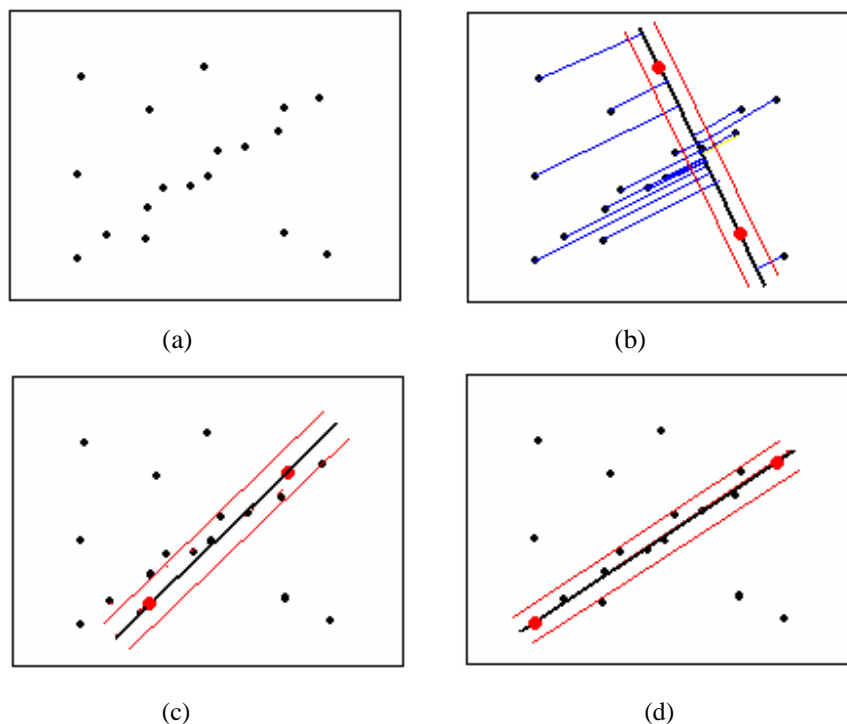


图 17.1.1: 直线的 RANSAC 估计。(a)是来自一条直线的 18 个测量数据点, 在给定的距离阈值下有 7 个外点; (b)是两个外点所确定的直线, 它的一致集仅含有三个数据点; (c)是由一个内点与一个外点所确定的直线, 它的一致集含有 8 个数据点; (d)是 RANSAC 最终估计的直线, 它具有最大的一致集, 并且所有的外点都被去除。

17.1.2 RANSAC

根据上节估计直线的思想, 对于一般模型 M 和给定的测量数据点集 D , RANSAC 估计模型参

数 p 的一般步骤如下:

1. 确定求解模型 M , 即确定模型参数 p , 所需要的最小数据点的个数 n 。由 n 个数据点组成的子集称为模型 M 的一个样本;
2. 从数据点集 D 中随机地抽取一个样本 J , 由该样本计算模型的一个实例 $M_p(J)$, 确定与 $M_p(J)$ 之间几何距离 $<$ 阈值 t 的数据点所构成的集合, 并记为 $S(M_p(J))$, 称为实例 $M_p(J)$ 的一致集;
3. 如果在一致集 $S(M_p(J))$ 中数据点的个数 $\#S(M_p(J)) >$ 阈值 T , 则用 $S(M_p(J))$ 重新估计模型 M , 并输出结果; 如果 $\#S(M_p(J)) <$ 阈值 T , 返回到步骤 2);
4. 经过 K 次随机抽样, 选择最大的一致集 $S(M_p(J))$, 用 $S(M_p(J))$ 重新估计模型 M , 并输出结果。

下面对 RANSAC 的细节作一些必要说明。

抽样次数

样本由从测量数据集中均匀随机抽取的子集所构成, 每个样本所包含数据点的个数 n 是确定模型参数所需要数据点的最小数目, 例如: 直线最少需要两个数据点才能确定, 即 $n = 2$; 圆最少需要 3 个数据点, 即 $n = 3$ 。至于为什么要选择最小数目, 下文将给出解释。为了陈述方便, 称不包含外点的样本为好样本, 否则称为坏样本。

在执行 RANSAC 时, 通常没有必要尝试每一种可能的抽样, 实际上尝试每一种可能的抽样在计算上也是不可行的。只要选择足够多的抽样次数 K , 保证至少能得一个好样本就可以了。假定数据点集中含有内点的比例是 w , 那么一个样本为好样本的概率 $p = w^n$ 。于是, 为了得到一个好样本需要抽样次数 K 的期望为

$$E(K) = 1 \cdot p_1 + 2 \cdot p_2 + 3 \cdot p_3 + \dots$$

其中, p_j 为在 j 次抽样中得到一个好样本的概率, 显然,

$$p_j = (1 - w^n)^{j-1} w^n$$

因此,

$$\begin{aligned} E(K) &= w^n + 2(1 - w^n)w^n + 3(1 - w^n)^2 w^n + \dots \\ &= w^n (1 + 2(1 - w^n) + 3(1 - w^n)^2 + \dots) \\ &= w^n \left(\frac{x}{1-x} \right)' \Bigg|_{x=1-w^n} = w^{-n} \end{aligned}$$

所以, 为了保证得到一个好样本, 抽样次数 K 应大于 w^{-n} 。这很自然地联系到 K 的标准差 $SD[K]$,

即抽样次数 $K = E[K] + SD[K]$ ，就能保证得到一个好样本。不难计算

$$SD[K] = \frac{\sqrt{1-w^n}}{w^n}$$

因此，

$$K = \frac{1 + \sqrt{1-w^n}}{w^n} \tag{17.1.1}$$

处理这个问题的另一种方法，是使得在 K 次抽样中所有样本均为坏样本的概率非常小，以保证获得一个好样本的概率非常大。记 z 为在 K 次抽样中所有样本均为坏样本的概率，则

$$z = (1-w^n)^K$$

所以，

$$K = \frac{\log z}{\log(1-w^n)} \tag{17.1.2}$$

换句话说，在 $K = \frac{\log z}{\log(1-w^n)}$ 抽样中得到一个好样本的概率为 $1-z$ 。

表 17.1.1: 样本所含数据点个数 n 与抽样次数 K 的一些对应值 ($w = 0.45$, $z = 0.02$)

n	2	3	4	5	7	9	13	16	20
K	18	41	94	210	1045	5168	126076	1.38×10^6	3.37×10^7

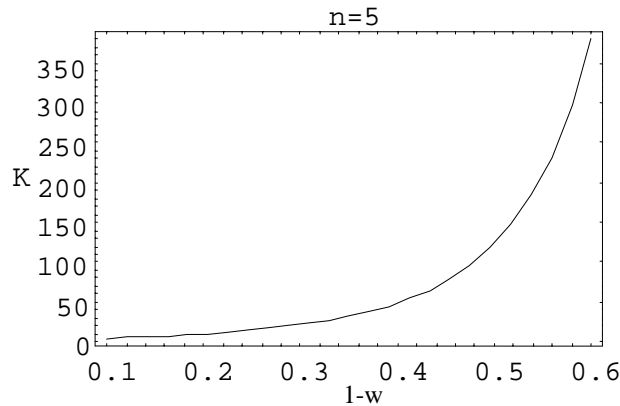


图 17.1.2: 抽样次数 K 与外点的比例 $(1-w)$ 之间的变化关系

取 $w = 0.45$ (内点的比例为 45%)， $z = 0.02$ (以 0.98 的概率获得一个好样本)，表(17.1.1)给出了样

本所含数据点个数 n 与抽样次数 K 的一些对应值。可以看出，随 n 增加抽样次数 K 将急剧增加，因此所需要数据点总数也将急剧增加，这就是为什么在执行 RANSAC 时需要对模型进行最小参数化使得样本由尽可能少的数据组成的原因。

值得说明的是，抽样次数 K 与外点的比例不是正相关的，抽样的计算代价即使在外点数目很大的时候也是可以接受的。图 17.1.2 给出 $n = 5$ 时，抽样次数 K 与外点的比例 $(1-w)$ 之间的变化关系。

距离阈值

如果我们希望所选取的阈值 t 使得内点被接受的概率是 α ，则需要通过由内点到模型之间几何距离的概率分布来计算距离阈值 t ，这是非常困难的。在实际中，距离阈值通常靠经验选取。但是，如果测量误差服从具有零均值和 σ 标准方差的高斯分布，则可以计算 t 的值。因为在这种情况下点到模型几何距离的平方 d^2 是高斯变量的平方和，因此它服从一个自由度为 m 的 χ_m^2 分布^{*)}。例如：在直线估计中，仅测量点到直线的几何距离 d ，根据假定 $d \sim N(0, \sigma)$ ，所以 $d^2 \sim \chi_1^2$ 。在单应估计中，数据点 (x, x') 到模型之间的几何距离是

$$d^2 = d_1^2 + d_2^2 = d_1^2(x, H^{-1}\hat{x}) + d_2^2(x, H\hat{x}),$$

由假定 $(d_1, d_2) \sim N(0, \sigma I_2)$ ，所以 $d^2 = d_1^2 + d_2^2 \sim \chi_2^2$ 。类似地，对于基本矩阵 F 有 $d^2 \sim \chi_1^2$ ，而对于三焦张量 T_i^{jk} 有 $d^2 \sim \chi_3^2$ 。由随机变量 χ_m^2 小于 k^2 的概率由分布 $F_m(k^2) = \int_0^{k^2} \chi_m^2(\xi) d\xi$ 来计算 (χ_m^2 分布见 17.1.4 节)，所以，

$$t^2 = F_m^{-1}(\alpha)\sigma^2 \tag{17.1.3}$$

于是，我们可以将内点与外点的划分为：

$$\begin{cases} \text{内点, } d^2 < F_m^{-1}(\alpha)\sigma^2 \\ \text{外点, } d^2 \geq F_m^{-1}(\alpha)\sigma^2 \end{cases} \tag{17.1.4}$$

通常 α 取 0.95，即数据点为内点的概率是 95%，它的意义是一个内点被排斥的错误概率仅是次数的 5%。表 17.1.2 给出了一些计算机视觉问题的距离阈值。

表 17.1.2 内点被接受的概率 $\alpha = 0.95$ ，距离阈值 $t^2 = F_m^{-1}(\alpha)\sigma^2$

m	模型	t^2
1	直线、基本矩阵	$3.84 \sigma^2$
2	单应、摄像机矩阵	$5.99 \sigma^2$

^{*)} 参考 17.1.4 节。

3	三焦张量	$7.81 \sigma^2$
---	------	-----------------

终止阈值

终止阈值是难以设置的问题。经验的做法是：给出内点比例 w 的一个估计值 ε ，如果一致集大小相当于数据集的内点规模则终止。由于很难给出内点比例 w 的一个准确估计，所以经验做法往往不能获得较好的估计结果。由于终止阈值仅仅是用来终止 RANSAC 的抽样，所以通常的做法是：初始时，给出内点比例 w 的一个最保守的估计，然后在抽样过程中不断地修正它，并利用公式(17.1.2)估计为了得到一个好样本所需要的抽样次数 K ，一旦当前的抽样次数达到或超过这个估计值 K 时，就终止抽样，结束 RANSAC 的抽样过程。在初始时，抽样次数 K 的估计值可能是非常大的，但随着抽样过程对内点比例 w 的更新，抽样次数的估计值将迅速的减小。这种终止抽样的方法是自适应的，详尽的算法如下所述。

自适应算法(终止 RANSAC 抽样):

- 1) 对内点比例作最保守估计 $w = w_0$ (如 $w_0 = 0.1$ ，这意味着在数据点集中可能有 90% 的外点，这确实是一个保守估计)，应用公式(17.1.2)，得到抽样次数 K 的初始值 K_0 ；
- 2) 抽样并更新 w_0, K_0 ：令当前抽样的一致集所含数据点占整个数据点的比例为 w ，若 $w > w_0$ ，则更新 $w_0 := w$ ，并且应用公式(17.1.2)更新抽样数 K_0 ；否则，保持原来的 w_0, K_0 ；
- 3) 如果抽样次数已达到或超过 K_0 ，则终止抽样；否则，返回步骤 2)。

不难看出，抽样次数在更新过程中是单调下降的，所以抽样过程必终止。注意：自适应算法同时还保证了有足够多次抽样，是一种值得推荐的自适应算法。

最终估计

RANSAC 方法将数据分为内点(最大一致集)和外点(剩下的数据)两个不相交的子集，同时给出模型的估计 M_{p_0} ，它由最大一致集所对应的样本计算出来。RANSAC 的最后一步是用所有的内点(最大一致集中的数据点)重新估计模型，该估计要涉及到代数方法或几何方法，最好使用几何方法(为了简化计算也可以应用 Sampson 几何近似)，它们需要迭代最小化，而 M_{p_0} 可作为最小化的初始点。

这个过程的一个缺点是内点与外点的分类变得不明确。这是因为将距离阈值应用于当前最大一致集所估计的 M_p 时，很可能有些点变为内点。解决这个问题的方法是：由内点得到模型估计 M_p ；由 M_p 应用(17.1.4)重新划分内点与外点；继续这个过程直至内点集收敛。

最后，值得指出的是，RANSAC 方法被实践证明是一种非常有效的鲁棒性方法。在可能的情况下，建议读者使用这种方法。

17.1.3 基本矩阵的 RANSAC 估计

作为一个例子，将 **RANSAC** 用于基本矩阵的估计。我们只需要

1. 定义样本；
2. 给定样本，求该样本所对应的基本矩阵；
3. 由最大一致集重新估计基本矩阵。

定义样本 令 $D = \{m_j \leftrightarrow m'_j \mid 1 \leq j \leq n\}$ 是给定的两幅图像的点对应集。由于确定基本矩阵的最低限度需要 7 个点对应，所以定义点对应集 D 中的 7 个对应点构成的子集为一个样本，所有样本的集合简称为样本集。

求样本 k 所对应的基本矩阵 F_k 由于样本 k 中仅有 7 个点对应，根据基本矩阵的 7-点算法，可能得到基本矩阵的 3 个解 $F_k^{(1)}, F_k^{(2)}, F_k^{(3)}$ ，将一致集最大的解作为该次抽样的基本矩阵。具体方法如下：

根据下述公式计算 $F_k^{(1)}, F_k^{(2)}, F_k^{(3)}$ 的一致集：

$$S(F_k^{(j)}) = \{(m \leftrightarrow m') \in D \mid d^2(m', F_k^{(j)}m) + d^2(m, (F_k^{(j)})^T m') < t^2\} \quad (17.1.5)$$

其中， d 表示点到直线的距离， t 是距离阈值。求 j_0 使得

$$\#S(F_k^{(j_0)}) = \max\{\#S(F_k^{(1)}), \#S(F_k^{(2)}), \#S(F_k^{(3)})\}$$

令 $F_k = F_k^{(j_0)}$ 为该次抽样的基本矩阵， $S(F_k) = S(F_k^{(j_0)})$ 为该次抽样的一致集。

由一致集重新估计基本矩阵 令 $S(F_{k_0})$ 是最大的一致集，则 F_{k_0} 是最低限度子集所确定的基本矩阵，将它作为初值，用 Sampson 近似方法重新估计基本矩阵：

$$\min \sum_{(m_j \leftrightarrow m'_j) \in S(F_{k_0})} \left(\frac{1}{(Fm_j)_1^2 + (Fm_j)_2^2 + (F^T m'_j)_1^2 + (F^T m'_j)_2^2} \right) \times (m'_j F m_j)^2$$

基本矩阵的自动估计

RANSAC 方法可以从有误匹配的点对应集中估计基本矩阵，这使得从两幅图像自动估计基本矩阵成为可能。基本矩阵自动估计的步骤如下：

1. 自动提取两幅图像的特征点集并建立初始“匹配对集”；
2. RANSAC 去除错误匹配对：
 - 计算当前抽样所确定的基本矩阵 F ，和它的一致点集 $S(F)$ ；
 - 如果当前的一致集大于原先的一致集，则保持当前的一致集 $S(F)$ 和相应的基本矩阵 F ，并删去原先的一致集和相应的基本矩阵；

- 由自适应算法终止抽样过程，获得最大一致集，最大一致集中的匹配对(内点)是正确匹配对。

3. 由最大一致集(即正确匹配对)重新估计基本矩阵。

为了自动估计基本矩阵，首先需要从两幅图像自动建立一个“点对应集”，可以容忍这个点对应集包含大量的误匹配，因为在 RANSAC 估计方法中，理论上只要存在 8 个“好”的点对应就可以估计出基本矩阵。

17.1.4 χ_n^2 分布

若随机变量 x 的密度函数为

$$p_{\chi_n^2}(x) = \begin{cases} \frac{1}{2^{n/2} \Gamma(n/2)} e^{-x/2} x^{n/2-1}, & 0 < x < \infty \\ 0, & \text{其它} \end{cases} \quad (17.1.6)$$

则称它服从自由度为 n 的 χ_n^2 (卡方)分布，并记作 $x \sim \chi_n^2$ 。 χ_n^2 变量 x 的分布函数(简称 χ_n^2 分布)为

$$F_n(x) = \int_0^x p_{\chi_n^2}(t) dt = 1 - \frac{\Gamma(n/2, x/2)}{\Gamma(n/2)} \quad (17.1.7)$$

其中： $\Gamma(\cdot)$ 为 Gamma 函数，即 $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$ ；而 $\Gamma(\cdot, \cdot)$ 为不完全的 Gamma 函数，即

$$\Gamma(a, z) = \int_z^\infty t^{a-1} e^{-t} dt。$$

图 17.1.4 给出了 χ_n^2 密度函数与分布的图示。表 17.1.3 给出了 χ_n^2 分布逆 $x = F_n^{-1}(P)$ 的若干对应值。

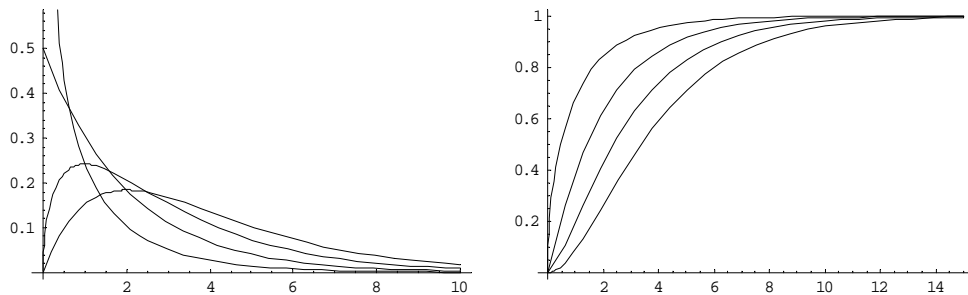


图 17.1.4 给出了 χ_n^2 密度函数(左图)与分布(右图)。在横轴为 4 处，由下至上分别表示自由度 $n=1,2,3,4$ 的情形。

表 17.1.3: χ_n^2 分布逆 ($x = F_n^{-1}(P)$)

	n=1	n=2	n=3	n=4	n=5	n=6
P=0.5	0.45	1.39	2.37	3.35	4.35	5.34
P=0.8	1.64	3.21	4.64	5.99	7.29	8.56
P=0.9	2.71	4.61	6.25	7.78	9.24	10.64
P=0.95	3.84	5.99	7.81	9.49	11.07	12.59
P=0.99	6.63	9.21	11.34	13.28	15.09	16.81

对于高斯随机向量与 χ_n^2 分布之间的关系, 我们有下述命题:

命题 5.1.1 若 n 维随机向量 $\mathbf{x} \sim N(\boldsymbol{\mu}, \Sigma)$, 且 \mathbf{x} 的分量相互独立, 即协方差矩阵 Σ 可逆, 则随机变量 $y = (\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu}) \sim \chi_n^2$, 特别地, n 维标准高斯随机向量的范数平方服从 χ_n^2 分布。

17.2 M-估计

在代数方法与几何方法中, 在形式上它们都是最小化下述函数:

$$\sum_j r_j^2(\mathbf{x}_j, \mathbf{p}) \quad (17.2.1)$$

其中 \mathbf{p} 为模型参数, $r_i(\mathbf{x}_i, \mathbf{p})$ 是模型 M_p 在数据点 \mathbf{x}_i 的残差(在代数方法中, $r_i(\mathbf{x}_i, \mathbf{p})$ 是代数残差, 即测量点到估计点的代数距离; 在几何方法中, $r_i(\mathbf{x}_i, \mathbf{p})$ 是几何残差, 即测量点到估计点的几何距离)。一旦测量数据中有外点(错误数据点), 这些方法是不鲁棒的, 因为它们在最小化过程中内点与外点不加区别而同等对待。

在 M-估计方法中, 用残差的函数代替残差。在形式上, M-估计方法是最小化下述函数:

$$\sum_j \rho(r_i(\mathbf{x}_i, \mathbf{p}); \sigma) \quad (17.2.2)$$

在最小化过程中, 函数 $\rho(t, \sigma)$ 必须具有抑制外点的作用, 使得外点对最小化几乎不起作用。这样的函数, 我们称为 M-估计子。例如, Huber 给出的估计子是:

$$\rho(t; \sigma) = \begin{cases} t^2, & |t| < \sigma \\ \sigma^2, & |t| \geq \sigma \end{cases} \quad (17.2.3)$$

它的图像如图 17.2.1 所示。这个估计子, 在残差较小时取值为残差的平方; 当残差较大时取一个常值 σ 的平方。如果将 σ 看作是 RANSAC 的距离阈值, 则 Huber 估计子在內点取其残差的平方, 外

点取常数，即距离阈值的平方。在外点取常值是有一定的道理的，通常认为外点服从均匀分布，而均匀分布的对数似然为常数。

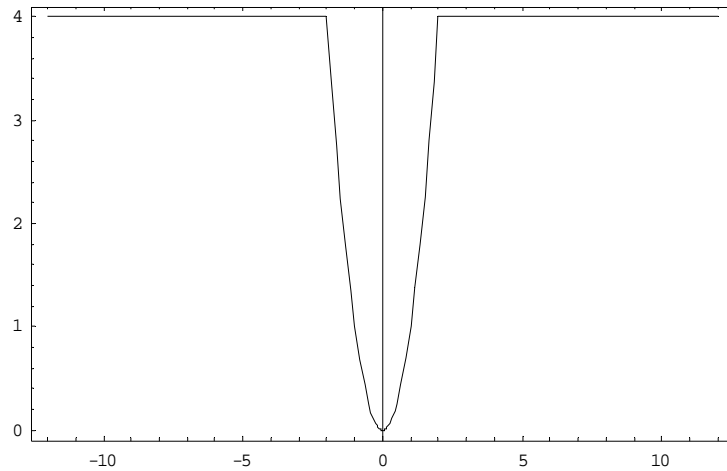


图 17.2.1: Huber 估计子($\sigma = 2$)

一般地，M-估计子 $\rho(t, \sigma)$ 在以 $t = 0$ 为中心长度为 2σ 的区间上看上去象 t^2 ，而在其它部分很平坦。M-估计子有各种各样的选择，较普遍的选择是

$$\rho(t; \sigma) = \frac{t^2}{t^2 + \sigma^2} \quad (17.2.4)$$

参数 σ 用于控制平坦区域，以抑制外点在最小化过程中的作用。图(17.2.2)是在不同参数 σ 下函数 $\rho(t, \sigma)$ 的图形。

M-估计方法有两个问题较难处理。第一个问题是最小化是非线性的，需要迭代求解且容易陷入局部极小。因此，这种方法需要一个好的初始点。解决这个问题的一般策略是，如同 RANSAC 那样进行抽样，用一个好样本对应的模型解作为迭代初始点。第二个问题是参数 σ 的选择，它既不能过大也不能过小(如图 17.2.3)。这是 M-估计中最棘手的问题。一种被认为比较好的策略是在每一步迭代中给出参数 σ 的估计，通常取残差绝对值的常数倍，即

$$\sigma^{(n)} = 1.4826 \text{Median}\{|r_j^{(n)}(\mathbf{x}_j, \mathbf{p}^{(n)})|\} \quad (17.2.5)$$

在 M-估计中，另一种常用的估计子是

$$\rho(t; \sigma) = \begin{cases} t^2, & |t| < \sigma \\ \sigma / |t|, & \sigma \leq |t| < 3\sigma \\ 0, & |t| \geq 3\sigma \end{cases} \quad (17.2.6)$$

其中 σ 残差的标准差。这种估计子与前面的差别在于：当残差很大时，这个数据被认为是错误的，因而在估计中不起任何作用。

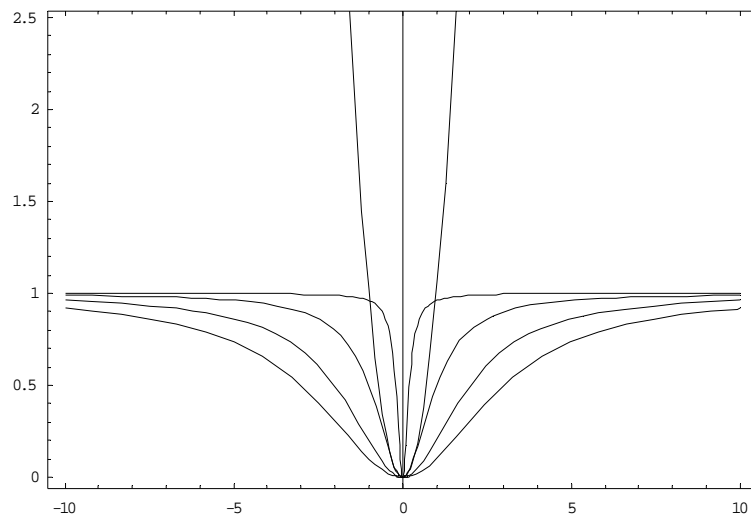


图 17.2.2: 函数 $\rho(t, \sigma)$ 的图形。在横轴 $t = 5$ 处，至上而下分别对应于 $\sigma = 0.2, 1, 2, 3$ 的情形；中间是 t^2 的图形。

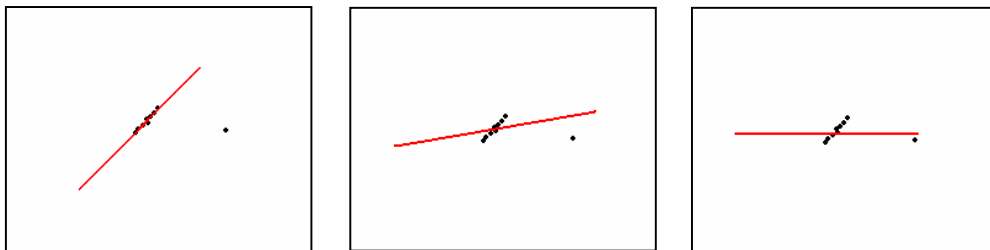


图 17.2.3: 参数 σ 对估计的影响。左: σ 近似正确值；中: σ 太小使得所有数据对优化都不敏感；右: σ 太大使得错误数据对优化很敏感。

17.3 最小中值估计

最小中值方法是通过求解下述最小化问题来估计模型参数：

$$\min \text{median}\{|r_j^2(\mathbf{x}_j, \mathbf{p})|\} \quad (17.3.1)$$

这种方法给出数据残差平方中值达到最小的模型参数。关于这种方法的计算过程无需多作说明，因为其计算过程与 RANSAC 方法十分类似，也是通过抽样来实现，所不同的是不再将数据分为内点与外点，最小化在整个数据集上进行。这种方法的优点是不需要内点与外点的先验知识，也就是说不需要划分内点与外点的距离阈值的先验知识，经验表明这种方法对错误数据也是很鲁棒的。由于在整个数据集上取中值，当数据集有大于 50% 的错误数据点时，这种方法不可能给出正确估计结果，这是最小中值方法相对 SANSAC 的主要不足之处。

当数据测量误差服从高斯分布时，最小中值法的效率非常差(关于估计方法的效率通常定义为给定方法给出的方差与理论上能达到的方差之比)。为了弥补这个不足，可以在执行最小中值方法之后，再对估计结果进一步进行加权最优化，其具体步骤如下：

首先计算鲁棒标准方差(robust standard deviation)：

$$\hat{\sigma} = 1.4826(1 + 5(N - n)\sqrt{\text{Med}}) \quad (17.3.2)$$

其中 Med 是最小中值；常数 1.4826 是为了保证达到(高斯误差分布)最大似然估计的同样效率的校正系数； N ， n 分别是整体数据点的个数与样本所含数据点的个数， $5/(N-n)$ 是为了补偿小数据集的效应(显然，仅有 n 个数据点时，总存在一个解使得模型能完全拟合这 n 数据点，这样计算方差将失去意义)。

再根据 $\hat{\sigma}$ ，为每一个数据点匹配一个权值：

$$w_j = \begin{cases} 1, & r_j^2(\mathbf{x}_j, \mathbf{p}) < (2.5\hat{\sigma})^2 \\ 0, & r_j^2(\mathbf{x}_j, \mathbf{p}) \geq (2.5\hat{\sigma})^2 \end{cases} \quad (17.3.3)$$

即，残差平方大于 $(2.5\hat{\sigma})^2$ 的那些数据点被认为是错误数据而不再考虑。

最后通过求解下述问题优化模型参数 \mathbf{p} ：

$$\min_{\mathbf{p}} \sum_j w_j r_j^2(\mathbf{x}_j, \mathbf{p}) \quad (17.3.4)$$

17.4 鲁棒最大后验估计

17.4.1 鲁棒最大后验估计

本节主要介绍模型鲁棒估计的 Bayes 方法。为了得到模型参数 θ 的最优估计 $\hat{\theta}$ ，我们引入最大后验 (MAP) 估计：

$$\theta = \arg \max_{\theta} \Pr(\theta | \mathcal{X}, \mathcal{M}, \mathcal{I}) = \arg \max_{\theta} \frac{\Pr(\mathcal{X} | \theta, \mathcal{M}, \mathcal{I}) \Pr(\theta | \mathcal{M}, \mathcal{I})}{\Pr(\mathcal{X} | \mathcal{M}, \mathcal{I})} \quad (17.4.1)$$

其中， $\hat{\theta}$ 为真实值 $\bar{\theta}$ 的估计。 \mathcal{X} 为测量数据集合， \mathcal{I} 为给定的信息。(17.4.1) 式中分子的第一项是似然项，第二项是先验项，分母为证据项。当 \mathcal{M} 固定时证据项为常数，所以在 MAP 估计中可以被省略。在下一章，将会看到证据项对于模型选择是非常重要的。

模型参数 θ 与第 16.3.3 节中的参数不尽相同，有必要对它进行解释。这里参数 θ 包括三个部分 α ， β 和 γ ：

α 部分是模型的本质参数，它与 16.3.3 节中的相应参数的意义相同，但这里使用最小化参数的数目，例如：对于单应矩阵， α 包括 8 个参数；对于基本矩阵， α 包括 7 个参数等等。

β 部分对应于测量数据坐标分量，这里与 16.3.3 节中的相应参数的不同点在于它包括测量数据的所有坐标分量，例如使用 n 个点对应估计单应或基本矩阵时， β 包括 $4n$ 个参数，但在 16.3.3 节中，单应相应的参数个数是 $2n$ 而基本矩阵是 $3n$ 。 β 的估计值 $\hat{\beta}$ 是测量数据的校正。

γ 部分是新引进的参数，是区别内点和外点的指标参数，其中参数个数等于测量数据集所包含数据点的个数。如第 i 个测量数据点是内点，则 $\gamma_i = 1$ ，否则 $\gamma_i = 0$ 。估计值 $\hat{\gamma}$ 对应于测量数据的内点与外点的一个划分。

参数集合 α 与 β 和 γ 不同之处在于参数集合 α 中的参数个数是固定的，不随测量数据的增加而增多，但估计精度随着测量数据的增加而提高；参数集合 β 的数量随测量数据量的增加而增加，它的估计精度不会随测量数据的增加而提高。 γ 与 β 类似。 β 和 γ 称为潜在参数，因为一旦它们被估计通常可以线性地估计出 α 。

在实际估计问题中有三种情况，第一种情况是对 α ， β 和 γ 都兴趣，例如在基本矩阵估计中，通常应用估计的 α 计算摄像机视点（或者摄像机运动），利用估计的 β 和 γ 计算场景的 3D 结构。第二种情况是对 α 和 β 兴趣，例如不存在外点时。最后一种情况是仅对 α 兴趣，例如曲线拟合。这三种情况很容易在 Bayes 框架下通过对参数进行边缘化(对该参数积分，去掉该参数，如： $\Pr(x|\mathcal{I}) = \int_{-\infty}^{\infty} \Pr(x,y|\mathcal{I}) \delta y$)加以解决，即通过边缘分布加以解决。这里，我们主要讨论第一种最为一般的情况。

在参数集 Θ 的上述分解下，(17.4.1)式中的似然项可表达成如下的形式：

$$\Pr(\mathcal{X}|\Theta, \mathcal{M}, \mathcal{I}) = \Pr(\mathcal{X}|\alpha, \beta, \gamma, \mathcal{M}, \mathcal{I}) = \Pr(\mathcal{X}|\beta, \gamma, \mathcal{M}, \mathcal{I}) \quad (17.4.2)$$

这表明似然项只与潜在参数 β 和 γ 有关。先验项也可以写成如下的形式

$$\Pr(\Theta|\mathcal{M}, \mathcal{I}) = \Pr(\alpha, \beta, \gamma|\mathcal{M}, \mathcal{I}) = \Pr(\beta, \gamma|\alpha, \mathcal{M}, \mathcal{I})\Pr(\alpha|\mathcal{M}, \mathcal{I}) \quad (17.4.3)$$

因此，式(17.4.1)可改写成：

$$\Theta = \arg \max_{\Theta} \Pr(\mathcal{X}|\beta, \gamma, \mathcal{M}, \mathcal{I})\Pr(\beta, \gamma|\alpha, \mathcal{M}, \mathcal{I})\Pr(\alpha|\mathcal{M}, \mathcal{I}) \quad (17.4.5)$$

上式是最一般情况下最大后验估计的形式。在下一节，我们将导出一个具体的形式，当然这种形式仍是非常普遍的，适用于大多数视觉中的实际问题。

17.4.2 似然项与先验项的具体形式

本节主要给出似然项关于内点与外点分布的混合模型和确定先验项的方法。在本节中，假定内点的服从高斯分布，外点服从均匀分布。这些假设尽管在理论上没有证明，但计算机视觉中大多数估计问题的实验结果都支持这一假设。

假定测量数据集由 n 个测量数据点所组成并且每个测量点是 D 维的。 β_j 的估计值记为 $\hat{\beta}_j = (\hat{x}_1^j, \hat{x}_2^j, \dots, \hat{x}_D^j)$ ， $j=1, 2, \dots, n$ 。例如估计基本矩阵时，测量点是点对应，它是 4 维的，即此时 $D=4$ 。根据第 16.3.2 节的讨论我们知道，对内点的高斯误差假设等价于普遍使用的高斯最小二乘原则，所选择的估计将误差平方和 $\sum_{j=1}^n e_j^2$ 最小化，其中

$$e_j^2 = \sum_{i=1}^D (\hat{x}_i^j - x_i^j)^2 \quad (17.4.6)$$

即，假设 $d_j = \sqrt{e_j^2}$ 服从各向同性的高斯分布且误差相互独立，所以

$$\Pr(\mathcal{X}|\beta, \mathcal{M}, \mathcal{I}) = \left(\frac{1}{\sqrt{2\pi}\sigma} \right)^{nD} \prod_{i=1 \dots n} \exp\left(-\frac{e_j^2}{2\sigma^2} \right), \quad (17.4.7)$$

下面考虑外点的分布。在一般情况下，测量数据位于 R^D 的有界区域，并服从密度函数为 $1/v$ 的均匀分布，其中 v 是该有界区域的体积(其单位与 σ 相同)。如果每一个测量数据，满足 $0 \leq x_i \leq L$ 则 $v = L^D$ 。当使用对数似然处理问题时，我们关心的是数值的幂指数，小的扰动不会对最后的 MAP 估计产生大的影响。对于更具体的两视点问题，幂指数相同但 v 稍有变化，这是因为来自第二个视点的特征点位置受到了搜索区域的限制。假设第二个视点的特征点搜寻区域是一个 $S \times S$ 的矩形区域，则匹配中位于第一幅图像上的特征点是图像区域的均匀分布，位于第二幅图像上的特征点在对

应的 $S \times S$ 区域中选择，所以此时外点均匀分布为 $1/v = 1/L^2 S^2$ 。

如果第 i 个测量数据点是内点，则 $\gamma_i = 1$ ，否则 $\gamma_i = 0$ 。所以结合式(17.4.7)，似然项可表示为

$$\Pr(\mathcal{X}|\beta, \gamma, \mathcal{M}, \mathcal{I}) = \prod_{i=1 \dots n} \underbrace{\left(\gamma_i \left(\frac{1}{\sqrt{2\pi}\sigma} \right)^D \exp\left(-\frac{e_i^2}{2\sigma^2}\right) + (1-\gamma_i) \frac{1}{v} \right)}_{\psi(e_i)} = \prod_{i=1 \dots n} \psi(e_i) \quad (17.4.9)$$

关于 β 和 γ 最大化(17.4.7)，得到鲁棒最大似然估计 (MLE)，相对应的对数似然的最大值记为

$$\mathcal{L} = \log\left(\Pr(\mathcal{X}|\hat{\beta}, \hat{\gamma}, \mathcal{M}, \mathcal{I})\right) \quad (17.4.9)$$

到目前为止，上述推导与最大似然方法没有什么不同，只是这里引进了外点分布。在考虑先验项时，就会出现有趣的区别。可以看出，对于两个先验项，第一项 $\Pr(\beta, \gamma|\alpha, \mathcal{M}, \mathcal{I})$ 是给定本质参数 α 和测量数据后的 β 和 γ 的先验知识，例如在基本矩阵估计中是匹配的先验知识。第二项是本质参数的先验知识 $\Pr(\alpha|\mathcal{M}, \mathcal{I})$ 。假设 β 与 γ 相互独立，这在大多数情况下是合理的，例如匹配点是否是外点仅取决于测量时的错误，而不取决于匹配规则（除非在摄像机镜头上有污点），因此

$$\Pr(\Theta|\mathcal{M}, \mathcal{I}) = \Pr(\beta|\alpha, \mathcal{M}, \mathcal{I}) \Pr(\gamma|\alpha, \mathcal{M}, \mathcal{I}) \Pr(\alpha|\mathcal{M}, \mathcal{I}) \quad (17.4.10)$$

为了定义第一项 $\Pr(\beta|\alpha, \mathcal{M}, \mathcal{I})$ ，需要定义 β 在由 α 决定的流形 S_α 上的分布（例如基本矩阵的本质参数 α 确定一个 4 维空间的锥面）。第一个假设是与每个测量点相关联的 β_j （流形 S_α 上一点）参数相互独立，则贝叶斯先验是假设 β_j 在 S_α 上服从均匀分布， $\Pr(\beta|\alpha, \mathcal{M}, \mathcal{I}) = 1/c$ 。若 S_α 相当平滑，也就是说在流形上的自然测度是 Lebesgue 测度，则 c 是曲面面积。当流形 S_α 有界时，比如椭圆或线段，Bayes 估计将会导致有趣的新结果，即 MAP 估计与流形包含的面积相关。如果几乎没有什么观察数据时，此种假设将会导致所估计的流形缩小(例如椭圆的一部分弧长变短)。一般情况下，流形扩展到 R^D 边界时， c 可近似为 R^D 的横截面面积，此曲面面积为 $c = \int_{R^D} \partial\beta_i = O(L^d)$ 。两视点问题， c 的取值稍有不同，例如对于 2D 关系（单应和仿射单应）， $c \approx L^2$ ；对于 3D 关系（基本矩阵和仿射基本矩阵）， $c \approx L^2 \cdot S$ 。一般地，对于贝叶斯模型估计和选择，没有必要获得 c 的精确数值。

对于指标参数集 γ ，假设对于所有的 j ， γ_j 是相互独立的，一个测量点为内点的先验概率为常数 ξ ，即

$$\Pr\{\gamma_j = 1\} = \xi, \forall j \quad (17.4.10)$$

其中， ξ 是内点的先验期望，用于定义 $\Pr(\gamma|\alpha, \mathcal{M}, \mathcal{I})$ 。

除非有重要的附加信息，否则对于大量的测量数据， α 上的先验知识无关紧要，可以假设它在

后验概率的峰值附近相对平滑, 因此对估计结果几乎不产生影响。一种可能的选择先验方法是借助积分几何方法得到 α 的一个先验密度, 此先验密度在某种变换群下保持不变, 如欧氏变换群, 射影变换群等。例如, 平面内直线 α , 对欧氏变换不变的先验密度分布为

$$\Pr(\alpha|\mathcal{M}, \mathcal{I}) \propto (u^2 + v^2)^{-3/2} \quad (17.4.10)$$

其中 $\alpha = (u, v)$, 对应的直线方程是 $ux + vy = 1$, 它等价于定义直线到原点的距离为一个均匀先验分布。形式为 $x \sin \omega + y \cos \omega = \rho$ 的直线集合 $G(r, \omega)$ 的测度被叫做线集合的密度, 它对欧氏变换不变的微分形式是 $dG = dp \wedge d\omega$, 式(17.4.10)可由此式导出。这种结果可以扩展到更复杂的高位流形, 但这已经超出我们讨论的范围。在这里, 假设当 $n \rightarrow \infty$ 时, 无论 $\Pr(\alpha|\mathcal{M}, \mathcal{I})$ 选取什么形式, 对结果几乎没有影响。因此, 模型的后验概率可以表达为

$$\Pr(\Theta|\mathcal{X}, \mathcal{M}, \mathcal{I}) \propto \prod_{i=1 \dots n} (\psi(e_i) \Pr(\beta_i|\alpha, \mathcal{M}, \mathcal{I})) \Pr(\gamma|\alpha, \mathcal{M}, \mathcal{I}) \quad (17.4.11)$$

注 当数据量很大时, 先验知识将变得不那么重要了, 此时最大似然估计(MLE)等同于最大后验估计(MAP), 最大后验估计不会在很大程度上改变参数估计的结果。然而, 对于模型选择问题(详见第 18 章), 情况却不是这样, 最大后验估计方法可以得到一个逻辑相容估计, 这与最大似然估计有很大的差别。

17.4.3 最大化边缘后验

对于大多数视觉估计问题, 给定 $\hat{\alpha}$, $\hat{\beta}$ 和 $\hat{\gamma}$ 可以被确定, 因此最优算法只需要在 α 上寻找即可。首先指出给定 $\hat{\alpha}$ 时, 确定 $\hat{\beta}$ 的方法。对于一个内点, 在高斯分布的假设下, $\hat{\beta}$ 的 MAP 估计在流形 $S_{\hat{\alpha}}$ 上, 并且该点到测量点的欧氏距离最小, 即使 e_i^2 最小, 由此我们可以由 $S_{\hat{\alpha}}$ 计算 $\hat{\beta}$ 。对于两视点匹配情况, 这等价于在射影重建中经常使用的最小化重投影误差。因此, Bayes 公式直接揭示出在什么条件和假设下可以使用最小化投影误差。在 $S_{\hat{\alpha}}$ 为代数簇的情况下, 最小化重投影误差等价于正交回归(orthogonal regression), 也叫整体最小二乘(total least squares)。显然, 在 $S_{\hat{\alpha}}$ 为线性流形时, 我们可以通过参数 β 边缘化的边缘后验最大化就得到一个稳定的 $\hat{\alpha}$ 。例如, 上一章中关于直线的几何估计方法。

如果流形 $S_{\hat{\alpha}}$ 是非线性的, 那么是否有可能在某些情况下, 最大化边缘后验能得到一个更稳定的估计值 $\hat{\alpha}$? Kanatani 注意到最小化几何距离将导致流形的有偏估计(曲率偏高, 即曲线被紧缩, 如线段的情况下, 估计值会比真实的线段短)并提出了重归一方法来进行校正。有些非 Bayes 统计方法把无偏性作为估计的重要准则, 直观上看这是很有必要的, 可是在最小化偏差时将导致变量的增加。关于无偏性的一个普遍问题是, 即使在一个近似无偏方法中, 也不能做到一次同时估计几个参数。

例如, 对 $\theta_1, \dots, \theta_j$ 进行无偏估计将导致对 θ_{j+1} 有过大的有偏估计。因此, 如果只需要估计 α , 忽略 β 是有意义的。可是, 如果需要同时估计 α , β 和 γ , 我们就不可能获得所有参数的无偏估计, 并且不知道最大化边缘后验与最大化后验哪个更合适。下面我们将会看到, 除测量数据集中在流形 S_α 的高曲率区域的情况外, 应用最大化边缘后验与最大化后验, 对于估计 α 和 β 差别不大。同时, 我们还会看到在确定 γ 时最大化边缘后验与最大化似然有很大差别。

边缘后验 (α, γ) 的边缘后验是后验(17.4.11)对 $\beta = (\beta_1, \dots, \beta_n)$ 的积分, 其形式为:

$$\int_{\beta} \Pr(\mathcal{O} | \mathcal{X}, \mathcal{M}, \mathcal{I}) \partial\beta \propto \left(\prod_{i=1 \dots n} \int_{\beta_i} \psi(e_i) \Pr(\beta_i | \alpha, \mathcal{M}, \mathcal{I}) \partial\beta_i \right) \Pr(\gamma | \alpha, \mathcal{M}, \mathcal{I}) \quad (17.4.12)$$

其中:

$$\Psi(e_i) = \gamma_i \left(\frac{1}{\sqrt{2\pi}\sigma} \right)^D \exp\left(-\frac{e_i^2}{2\sigma^2}\right) + (1-\gamma_i) \frac{1}{v}$$

记 $\Pr(\beta_i | \alpha, \mathcal{M}, \mathcal{I}) = 1/c$ 。为了计算积分 $\int_{\beta_i} \psi(e_i) \Pr(\beta_i | \alpha, \mathcal{M}, \mathcal{I}) \partial\beta_i$, 分别考虑均匀分布项和高斯分布项。首先, 很容易计算均匀分布项的积分:

$$\int_{\beta_i} \left((1-\gamma_i) \frac{1}{v} \right) \frac{1}{c} \partial\beta_i = (1-\gamma_i) \frac{1}{v} \quad (17.4.13)$$

再计算高斯分布项的积分: 我们所使用的关键假定是测量数据到流形 S_α 的垂足, 记为 $\mathbf{x}(\alpha(\beta_i))$, 的周围是局部近似平坦的, 该点的切平面记为 $(\Pi(\alpha(\beta_i)))$ 。在这种情况下, 余维数为 $D-d$, $d = \dim(S_\alpha)$ 。根据毕达格拉斯定理, 平方误差 e_i^2 可分解为两部分, 即与流形垂直的部分 e_\perp^2 和平行的部分 e_\parallel^2 , 使得 $e_i^2 = e_\perp^2 + e_\parallel^2$ 。应用恒等式

$$(2\pi\sigma^2)^{-n/2} \int_{\mathbb{R}^n} \exp(-\mathbf{x}^\top \mathbf{x} / 2\sigma^2) d\mathbf{x} = 1 \quad (17.4.14)$$

我们得到

$$\int_{\beta_i} \left(\gamma_i \left(\frac{1}{\sqrt{2\pi}\sigma^2} \right)^D \exp\left(-\frac{e_\perp^2 + e_\parallel^2}{2\sigma^2}\right) \right) \frac{1}{c} \partial\beta_i \approx \gamma_i \left(\frac{(\sqrt{2\pi}\sigma^2)^{d-D}}{c} \exp\left(-\frac{e_\perp^2}{2\sigma^2}\right) \right) \quad (17.4.15)$$

从上式可以看出, 当流形局部线性时, 先验项中平行于流形 S_α 的部分被积分掉了, 只有正交项起作用。在下一章中, 我们还会看到此积分对模型选择同样重要。因此, 后验(17.4.11)关于 (α, γ) 的边缘后验为

$$\int_{\beta} \Pr(\Theta | \mathcal{X}, \mathcal{M}, \mathcal{I}) \partial \beta \propto \left(\prod_{i=1 \dots n} \gamma_i \left(\frac{(2\pi\sigma^2)^{(d-D)/2}}{c} \exp\left(-\frac{e_i^2}{2\sigma^2}\right) \right) + \frac{1-\gamma_i}{v} \right) P_r(\gamma | \alpha, \mathcal{M}, \mathcal{I}) \quad (17.4.16)$$

如果 α 的先验是局部足够平坦的, 我们可以通过最大化这个边缘后验得到 (α, γ) 的 MAP 估计。计算方法是直接使用 13 章的非线性优方法, 或者将 γ 作为潜在参数应用 14 章的 EM 方法。

最大化边缘后验与最大化似然的差异 比较式(17.4.7)与(17.4.16), 不难发现最大边缘后验与最大化似然是有区别的。下面就这两种方法区分内点与外点的阈值差异, 来说明它们之间的不同。如果 β 是流形 S_{α} 上距测量数据最近的点集, 则指标参数集 γ 能被确定。

参考式(17.4.7), 由于内点的先验期望为 ξ , 所以第 i 个测量点为内点与外点的概率分别是

$$P_r(i \in \text{inlier}) = k \cdot \xi \cdot \left(\frac{1}{\sqrt{2\pi\sigma}} \right)^D \exp\left(-\frac{e_i^2}{2\sigma^2}\right), P_r(i \in \text{outlier}) = k \cdot \frac{1-\xi}{v}$$

其中 k 是一个常数, 使得 $P_r(i \in \text{inlier}) + P_r(i \in \text{outlier}) = 1$ 。如果第 i 个测量点为内点, 必有 $P_r(i \in \text{inlier}) > P_r(i \in \text{outlier})$, 因此,

$$\frac{e_i^2}{\sigma^2} < -2 \log \left(\frac{1-\xi}{\xi} \cdot \frac{(2\pi\sigma^2)^{D/2}}{v} \right) \quad (17.4.18)$$

所以, 对于最大化似然的情况,

$$\hat{\gamma}_i = \begin{cases} 1, & \frac{e_i^2}{\sigma^2} < T_{mle} \\ 0, & \frac{e_i^2}{\sigma^2} \geq T_{mle} \end{cases} \quad (17.4.19)$$

其中

$$T_{mle} = -2 \log \left(\frac{1-\xi}{\xi} \cdot \frac{(2\pi\sigma^2)^{D/2}}{v} \right) = 2 \log \left(\frac{\xi}{1-\xi} \right) + D \cdot \log \left(\frac{v^{2/D}}{2\pi\sigma^2} \right) \quad (17.4.20)$$

对于最大边缘后验的情况, 通过类似的分析, 我们得到

$$T_{map} = -2 \log \left(\frac{1-\xi}{\xi} \cdot \frac{c}{v} \cdot (2\pi\sigma^2)^{(D-d)/2} \right) = 2 \log \left(\frac{\xi}{1-\xi} \right) + (D-d) \cdot \log \left(\frac{(v/c)^{2/D-d}}{2\pi\sigma^2} \right) \quad (17.4.21)$$

特别地, 对于 2D 关系 (单应和仿射单应),

$$T_{map} = 2 \log \left(\frac{\xi}{1-\xi} \right) + 2(D-d) \cdot \log \left(\frac{S^2}{2\pi\sigma^2} \right) \quad (17.4.21a)$$

对于 3D 关系（基本矩阵和仿射基本矩阵），

$$T_{map} = 2 \log \left(\frac{\xi}{1-\xi} \right) + (D-d) \cdot \log \left(\frac{S^2}{2\pi\sigma^2} \right) \quad (17.4.21b)$$

式(17.4.21a)与(17.4.21b)的第 2 项因子 $\lambda_1 = \log \left(\frac{S^2}{2\pi\sigma^2} \right)$ ，在模型选择的 GRIC 标准中具有非常重要的作用，见 18.4.1 节。

在最大化似然和最大化后验两种情况下，测量点是内点还是外点的决策规则都依赖于误差平方 e_i^2 。尽管如此，当在最大似然估计时，阈值依赖于数据的维数 D ；在最大化边缘后验时，阈值依赖于流形的余维数 $r=D-d$ 。下面作进一步分析，不难计算：

$$T_{mle} - T_{map} = \log \left(\frac{c^2}{(2\pi\sigma^2)^d} \right) \quad (17.4.22)$$

对于视觉估计问题，总是有 $c = \int_{R^d} \partial\beta_i \approx L^d > S^d$ ，例如，在单应矩阵、基本矩阵、三焦张量估计中， L^2 是第一幅图像的大小， S 是在其它图像搜索对应点时所使用的窗口尺寸。因此，对于视觉估计问题总是有 $c^2 > S^{2d} > (2\pi\sigma^2)^d$ ，所以 $T_{mle} > T_{map}$ 。这说明最大似然算法决策一个测量点是内点还是外点时，总是偏向于内点。所以，在视觉估计问题中，外点存在时，我们更倾向于使用最大化边缘后验算法或最大化后验算法，而不是最大似然估计算法。

7.4.3 最大后验一致抽样算法（MAPSAC）

在本节我们将后验概率模型与 RANSAC 方法结合起来，给出一个高效估计算法，称为最大后验一致抽样算法（MAPSAC）。RANSAC 算法已被实践证明是一个非常成功的鲁棒估计算法，在此，将它扩展到寻找最大后验的模型估计方法中。对 RANSAC 的思考可以得出，事实上它是对如下损失函数最小化

$$C = \sum_{i=1}^n \rho \left(\frac{e_i^2}{\sigma^2} \right) \quad (17.4.23)$$

其中

$$\rho\left(\frac{e_i^2}{\sigma^2}\right) = \begin{cases} 0, & \frac{e_i^2}{\sigma^2} < T \\ s > 0, & \frac{e_i^2}{\sigma^2} \geq T \end{cases} \quad (17.4.24)$$

RANSAC 的一个不足是考虑内点时如果阈值 T 设置太高，估计的性能将会很差。换句话说，此时内点之间将没有什么区别，而对外点的惩罚是一个常数。因此，阈值 T 越高，有相同值 C 的估计性能越差。例如，若 T 足够大，则损失将会相同，因为每个测量点都将是内点。

为了克服 RANSAC 的上述不足，MAPSAC 算法使用最小化后验对数或边缘后验对数而不是最小化 C 。在给定 α 的情况下，后验对数或边缘后验对数是容易计算的。给定 α 时，在对 β 边缘化的同时，优化对内点与外点的选择，可以通过下述最小化来完成：

$$C_{map} = \sum_{i=1}^n \rho_{map}\left(\frac{e_i^2}{\sigma^2}\right) \quad (17.4.25)$$

其中

$$\rho_{map}\left(\frac{e_i^2}{\sigma^2}\right) = \begin{cases} \frac{e_i^2}{\sigma^2}, & \frac{e_i^2}{\sigma^2} < T_{map} \\ T_{map}, & \frac{e_i^2}{\sigma^2} \geq T_{map} \end{cases} \quad (17.4.26)$$

阈值 T_{map} 由(17.4.21)式定义。

从上式可以看出，对于每个外点的惩罚相同，而对内点却根据它们拟合数据的好坏程度加以区别。这种方法对所有鲁棒估计都大有裨益，并且不会增加额外的计算。实际上，在 M-estimators 方法中我们也使用了类似的方法。

在 MAPSAC 算法过程中，也可以对 γ 边缘化，因为下述积分是容易计算的：

$$\int_{[0,1]} \left(\gamma \left(\frac{(\sqrt{2\pi}\sigma^2)^{d-D}}{c} \exp\left(-\frac{e_i^2}{\sigma^2}\right) \right) + (1-\gamma) \frac{1}{v} \right) d\gamma = \frac{1}{2} \left(\frac{(\sqrt{2\pi}\sigma^2)^{d-D}}{c} \exp\left(-\frac{e_i^2}{\sigma^2}\right) + \frac{1}{v} \right) \quad (17.4.27)$$

虽然计算代价和复杂性有所增加，但 Torr 指出，此时 α 的精度有所提高。如果实际问题在估计 α 同时也注重 γ 的估计（如三维视中的 3D 重构问题，需要识别外点以获得正确的运动和结构），对 γ 边缘化是不可取的。

MAPSAC 算法的计算步骤与 RANSAC 算法是类似的，一般由下述几个步骤构成：

1. 确定求解模型 \mathcal{M} ，即确定模型参数 $\Theta = (\alpha, \beta, \gamma)$ 所需要的最小数据点的个数 n 。由 n 个数据点组成的子集称为模型 \mathcal{M} 的一个样本。

2. 从数据点集 \mathcal{X} 随机地抽取一个样本 J , 并由该样本估计 Θ 。
3. 最大化后验 $\Pr(\Theta|\mathcal{X}, \mathcal{M}, \mathcal{I}) \propto \prod_{i=1 \dots n} (\psi(e_i) \Pr(\beta_i|\alpha, \mathcal{M}, \mathcal{I})) \Pr(\gamma|\alpha, \mathcal{M}, \mathcal{I}) \Pr(\alpha|\mathcal{M}, \mathcal{I})$ 。通常可使用下述两种方法进行近似计算:
 - A. 通过最小化(17.4.25) (即, 最大化后验) 得到估计 Θ_j ;
 - B. 应用 EM 算法, 先对 γ 边缘化, 再直接最大化后验得到估计 $(\alpha, \beta)_j$ (此时, 算法仅关心 α 估计精度)。
4. 经过 K 次随机抽样或达到抽样次数 (确定抽样次数的规则与 RANSAC 方法相同), 在所有抽样中选择一个最好的估计 Θ_j , 即它最大化后验 $\Pr(\Theta_j|\mathcal{X}, \mathcal{M}, \mathcal{I}) \geq \Pr(\Theta_j|\mathcal{X}, \mathcal{M}, \mathcal{I}), \forall J$, 或者说它最小化损失 $C_{map}(\Theta_j) \leq C_{map}(\Theta_j), \forall J$ 。
5. 以 Θ_j 为初始值, 使用非线性优化技术 (如梯度下降法或 LM 法), 最大化后验 $\Pr(\Theta|\mathcal{X}, \mathcal{M}, \mathcal{I})$ 或最小化 $C_{map}(\Theta)$ 得到 Θ 最终估计。

18. 模型选择

在计算机视觉领域有两大问题，一是寻找数据的正确的表示方法，二是利用表示方法作决策和形成对现实世界的假设。这两大问题都涉及到模型选择，因此模型选择技术是计算机视觉算法中的关键技术。本章以运动模型为案例，探讨视觉模型选择问题。主要内容包括似然比、AIC 标准、BIC 标准以及 GRIC 标准在视觉模型选择中的应用。

18.1 似然比检验

18.1.1 基本运动模型

符号和术语

3D 场景中的一个点被投影到第一和第二幅图像上的齐次坐标向量分别为 \mathbf{x}^1 和 \mathbf{x}^2 。第 i 个点在第 j 幅图像上的非齐次坐标为 (x_i^j, y_i^j) 。两个视点下的第 i 个点对应表示为向量 $\mathbf{X}_i = (x_i^1, y_i^1, x_i^2, y_i^2)^\top$ 。在本章中，还会涉及簇拟合(fitting generic varieties)的更一般的情况，为了方便，我们需要更一般的范畴，即 \mathbf{X}_i 表示一个 D 维的观察数据， $\mathbf{X}_i = (x_i^1, \dots, x_i^D)^\top$ ，所有观察到的两个视点之间的点匹配的集合记作 $\mathcal{X} \subset R^D$ 。无噪(真实)数据，记作 $\bar{\mathbf{X}}$ ，估计值为 $\hat{\mathbf{X}}$ ，有噪数据为 \mathbf{X} 。给定 \mathbf{y} 之后 \mathbf{x} 的密度函数为 $\Pr(\mathbf{x}|\mathbf{y})$ 。 \mathcal{M} 是数据的假设模型，由噪声模型和以 $\Theta = \{\alpha, \beta, \gamma\}$ 为参数的两视点关系 \mathcal{R} 组成，其中 $\Theta = \{\alpha, \beta, \gamma\}$ 的意义与 17.4 节相同。一个代数簇(variety)定义为在 R^D 中满足多项式方程组 $g_q(x_q) = 0, q = 1 \dots Q$ 的所有点的集合 \mathcal{V} ， \mathcal{R} 将诱导出 R^D 中的一个 d 维代数簇 \mathcal{V} ，其参数是 α ， $r = D - d$ 称为代数簇 \mathcal{V} 的余维数。如果 g_q 的函数形式不是多项式的，则称 \mathcal{V} 是一个流形，这种区别对本章内容来说是无关紧要的。

基本运动模型

在刚体运动的假设下，本章以两视点中的点运动模型作为模型选择的案例。每个运动模型 \mathcal{M} ，由噪声模型(可能包含外点模型)和由参数集合 Θ 描述的关系 \mathcal{R} 组成，用来定义 Q 个两个视点的图像坐标的隐函数，即 $g_q(x_i^1, y_i^1, x_i^2, y_i^2, \Theta) = 0$ ，即 $g(\mathbf{X}_i, \Theta) = 0, i = 1 \dots n, q = 1 \dots Q$ ，其中 $\mathbf{0}$ 是一个零向量。

$g(\mathbf{X}_i, \Theta) = 0$ 是关系的最一般情况, 它不一定是多项式形式, 所以, S 可能是一个流形而不是代数簇。

本章作为模型选择的案例主要是以下四种两视点关系:

1. \mathcal{R}_1 : 基本矩阵 F 。假设 3D 对象经过刚体运动后, 被观察的空间特征点在第一幅图像上的图像齐次坐标 $\{\mathbf{x}_i^1\}$ 转化为第二幅图像上的图像齐次坐标 $\{\mathbf{x}_i^2\}$, 这个点对应集合满足 $\mathbf{x}_i^{2T} F \mathbf{x}_i^1 = 0$, F 是一个秩为 2 的 3×3 矩阵。基本矩阵刻划了整个的极几何, 它包含了从点对应得到的摄像机运动和摄像机内参数的所有信息。
2. \mathcal{R}_2 : 仿射基本矩阵 F_A 。它是在 Mundy 和 Zisserman 提出的仿射摄像机下的极几何关系, 仿射基本矩阵 F_A 是线性的, 它也是一个秩为 2 的 3×3 矩阵, 但它的左上角为一个 2×2 的零子阵。与基本矩阵一样, 对于点对应集合满足 $\mathbf{x}_i^{2T} F_A \mathbf{x}_i^1 = 0$ 。在一般透视摄像机下, 如果 3D 对象距摄像机非常远接近平行投影, 则两幅图像点对应也非常近似地满足这种关系。
3. \mathcal{R}_3 : 单应矩阵 H 。假设摄像机作以光心为中心的旋转运动, 或一个空间平面上的 3D 点作刚体运动, 两幅图像的点对应就满足单应关系, 即 $\mathbf{x}^2 = H \mathbf{x}^1$ 。单应是一个满秩的 3×3 矩阵。
4. \mathcal{R}_4 : 仿射单应 H_A 。在仿射摄像机的模型下, 空间平面上的 3D 点作刚体运动, 则两幅图像上的点对应就满足仿射单应关系, $\mathbf{x}^2 = H_A \mathbf{x}^1$ 。在一般透视摄像机下, 如果所有点都在一个很远的平面上, 或者焦距很长并且摄像机是旋转运动, 则两幅图像的点对应也能非常近似地满足仿射单应关系。

上面四种运动模型关系, 可以分两类: 第一类是能够恢复摄像机位置和结构的 3D 关系, 基本矩阵和仿射基本矩阵属于 3D 关系; 第二类是不能恢复 3D 结构的 2D 关系, 单应与仿射单应则属于 2D 关系。

模型复杂度 传统的模型选择方法是定义一个损失函数来惩罚参数较多的模型。那么, 每个模型究竟有多少参数呢? 基本矩阵有 7 个自由度, 单应有 8 个, 因此基本矩阵模型泛化能力更强; 同样, 仿射基本矩阵有 4 个自由度, 仿射单应有 6 个, 所以仿射基本矩阵泛化能力更强。注意: 模型自由度的数目正好是第 16.3.3 节所指出的模型本质参数的数目。上面看上去很含糊的结论可以通过考虑定义在数据集合上的模型流形的维数得到。

将两幅图像上的对应点 \mathbf{x}^1 和 \mathbf{x}^2 的非齐次坐标并起来, 得到度量空间 R^4 中的一点 \mathbf{X}_i , 度量空间 R^4 可以理解为连接的图像空间。于是, 由一个刚体运动诱导的图像对应便与 R^4 空间中的一个代数簇 S 联系起来。两幅图像间的基本矩阵和仿射基本矩阵分别是维数为 3 的二次簇和维数为 3 的线性簇, 同样, 单应和仿射单应分别是维数为 2 的 2 次代数簇和维数为 2 的线性簇。可以笼统地说, 基本矩阵描述了 R^4 空间中的一个 3 维曲面, 而单应描述了此空间中的一个 2 维曲面。基本矩阵(仿射的或是一般的)曲面上的每一点就是两幅图像的一对匹配点, 有 3 个自由度, 对应于 3 维场景中的一点, 也就是说可以用 3 维坐标系统来刻划 S 。同样, 单应(或仿射单应)曲面上的每一点也描述了一

对点对应，它有两个自由度。从观测数据估计模型等价于用该模型对应的流形拟合观测数据，在这个意义上，模型估计与流形拟合是一回事。因此，在后面我们常常用术语“流形拟合”来代替“模型估计”。

表 18.1.1 给出了摄像机模型和上述四种运动模型的性质，其中参数数目是指从 n 个测量数据在没外点的情况下估计模型时所涉及的参数数目（见第 16.3.3 节），在本章中 AIC、BIC 和 GRIC 中所涉及的模型参数数目都是指这个数目。

表 18.1.1：一些视觉模型的性质。n 是估计模型时所使用的点对应数目。

模型(M)	数据维数 (D)	流形维数 (d)	流形余维数 (r=D-d)	本质参数 数目(k)	参数数目 (p=dn+k)
$P_A = \begin{pmatrix} * & * & * & * \\ * & * & * & * \\ 0 & 0 & 0 & 1 \end{pmatrix}$	5	3	2	8	3n+8
$P = \begin{pmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{pmatrix}$	5	3	2	11	3n+11
$F_A = \begin{pmatrix} 0 & 0 & * \\ 0 & 0 & * \\ * & * & * \end{pmatrix}$	4	3	1	4	3n+4
$F = \begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * \end{pmatrix}$	4	3	1	7	3n+7
$H_A = \begin{pmatrix} * & * & * \\ * & * & * \\ 0 & 0 & 1 \end{pmatrix}$	4	2	2	6	2n+6
$H = \begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * \end{pmatrix}$	4	2	2	8	2n+8

18.1.2 似然比检验

在 2D 情况下，如果我们不知道数据是来自一条曲线、一条直线或是一个点，在两视点匹配情况下数据是来自单应还是基本矩阵，在估计时应该使用什么模型呢？在没有其它任何息的情况下，

我们只能对每一种可能的模型都进行尝试，跟随的问题是如何根据尝试结果选择正确的模型呢？在直觉上，应该选择具有最小平均误差平方和（SSE）的模型作为正确的模型，但实际情况与我们的直觉相冲突，因为最一般模型往往能提供最小的 SSE。因此，选择具有最小 SSE 的模型往往导致选择最一般模型。表 18.1.1 给出了对 100 个数据集合(每个数据集合包含 100 个人造的匹配)关于平均误差平方和的实验结果。随机产生的点对应与 F, F_A, H 和 H_A 限制一致，即一般模型，正交投影，摄像机旋转或者正交平面运动，其中点对应坐标上的噪声服从零均值标准差为 $\sigma = 1$ 的高斯分布。在每个数据集上都对任何一个模型进行估计，并记录 SSE 的平均值，如表 18.1.1 所示。可以看到，选择具有最小 SSE 的模型往往导致选择结果是最一般的模型 F ，因此我们需要更复杂的模型选择方法，即统计推理的方法。

表 18.1.1：100 个点对应 100 次实验的平均误差平方和，每个坐标上的噪声方差为 1，括号中的数字是正确模型的期望值^{*)}

SSE	Point Motion			
	General (F)	Orthographic (F_A)	Rotation (H)	Affinity (H_A)
Fundamental F	93.074 (93)	87.0370	80.6162	78.378
Affine F_A	978.350	96.448 (96)	806.389	85.875
Homography H	4986.881	4834.735	193.964 (192)	189.132
Affinity H_A	4993.045	4967.894	1023.118	191.643 (194)

统计推理的大部分理论虽然只有 70 年的历史，但却被叫做经典理论。它主要包括两个分支，估计理论和假设检验。假设检验理论主要来自 Neyman 和 Pearson 的工作。就总体而言，为了通过假设检验来对模型进行比较，需要通过比较描述数据的零假设模型 \mathcal{M}_1 与备选假设模型 \mathcal{M}_2 来进行检验。除特殊情况外，模型 \mathcal{M}_1 的参数 θ_1 包含模型 \mathcal{M}_2 的参数 θ_2 （两个模型具有嵌套结构）即 $p_1 > p_2$ ， p_i 为模型 i 的参数个数，进行检验的一般方法是应用似然比检验(likelihood ratio test)。在高斯分布情况下，它导致卡方检验。首先，对两个模型而言，必须获得参数的 MLE 估计 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ ，然后由它们获得检验统计量

$$LLR(\mathcal{X}) = 2 \log \left(\frac{\Pr(\mathcal{X} | \hat{\theta}_1, \mathcal{M}_1)}{\Pr(\mathcal{X} | \hat{\theta}_2, \mathcal{M}_2)} \right) = 2(\mathcal{L}_1 - \mathcal{L}_2) \tag{18.1.1}$$

其中 $\mathcal{L}_i = \log(\Pr(\mathcal{X} | \theta_i, \mathcal{M}_i))$ 。 $LLR(\mathcal{X})$ 渐进服从自由度为 $p_1 - p_2$ 的 χ^2 分布。如果 $LLR(\mathcal{X})$ 小于某

^{*)} 本章中的所有实验数据均来源于文献：P.H.S. Torr, Bayes Model Estimation and Selection for Epipolar Geometry and Generic Manifold Fitting, IJCV 50(1): 35-61,2002

个阈值(由一个显著性水平 α 决定)则接受 M_2 , 否则拒绝 M_2 , 即检验

$$LLR(\mathcal{X}) = 2(\mathcal{L}_1 - \mathcal{L}_2) < \chi^2(\alpha, p_1 - p_2) \quad (18.1.2)$$

如果 M_2 成立, 则过拟合的机会是所选择的 α ; 如果 M_1 成立, 则欠拟合的机会是一个未知的 β 。 $1 - \beta$ 的值可用来描述检验的能力。 检验能力明显与 α 的选择和数据的分布有关。 例如, 对于给定 α 的情况, 所有的匹配视差均很小时的检验能力会明显小于视差大的情况。 理想情况下, 应选择 α 使过拟合的机会 (α) 与欠拟合机会 (β) 都很小, 即检验能力很高。

在统计假设检验的 Neyman-Pearson 理论中, 只考虑用拒绝正确假设和接受错误假设的概率来定义一个决策的代价。 此方法的缺点是不能应用于多个模型都合适的情况, 如对于多决策问题, 检验过程很难(经常是不能处理的)对一定数目的相互依赖的显著性水平进行选择。 而且, 假设检验过程只适合嵌套的两个模型(一个模型的参数是另一个模型参数的子集), 不适合于视觉运动模型选择问题。 正如上面看到的那样, 最大似然比方法总是导致选择最一般的模型, 因此, 我们需要更一般的归纳推理方法处理模型选择的复杂性。 这导致发展了各种各样的模型选择的扩展推理理论, 它们大致可以以下几类:

- I. **预测方法。** 该方法是选择最有可能最小化未来观测数据的模型。 一般情况下, 它使用一种信息标准为每个模型计分, 每个模型的得分表示其预测能力。 例如 AIC 标准(见下节)通常被用来为模型计分。 但是, AIC 不是渐进一致的, 有过高估计模型参数数量的趋势。 还有一些与预测方法关相的其他方法, 如交叉确认和 bootstrap 方法, 它们渐进地等价于 AIC, 但对于小样本问题会产生与 AIC 不同的重要结果。
- II. **Bayes 方法。** 如果将 Bayes 规则作为公理, 可以精确计算每个模型的后验概率, 则 Bayes 方法可能是最一般和最有效的模型选择方法。 并且, Bayes 方法包含上面的预测方法, 因为最优 Bayes 估计同样能最小化预测误差。 它的主要问题是, 在计算 Bayes 后验概率时需要在模型的高维参数空间上求积分。 一般情况下, 被积函数在参数空间不解析并且参数空间维数非常高, 因此积分不易计算。 于是, 人们寻找计算积分的近似方法, 导致不同的模型选择信息标准。 事实上, 其他所有非似然的模型选择信息标准都可以看作是对 Bayes 方法的近似, 即其他方法的最好结果也只是 Bayes 方法的近似结果, 如 BIC 信息标准。
- III. **最小描述长度方法(Minimum description length methods)。** 该方法是 Solomonoff 思想的实现, Solomonoff 认为对于数据压缩和模式选择, 用最简单或是最紧凑的理论描述数据能起到最好的预测效果。 数据的复杂性与模型结合在一起考虑可以简化 Kologmorov 复杂性(编码数据和模型的最短 Turing machine 的长度), 因此能用来导出基于复杂性(非

随机量)的模型的先验概率。然而, Kolmogorov 复杂性是不可能精确估计的, 只能近似确定, 如 Rissanen 所提出的 MDL 方法, Wallace 和 Boulton 的 MML 方法。一种基于不可精确计算的理论似乎有点不能令人满意。

- IV. 结构风险最小化原则 (SRM)。SRM 是选择能最小化经验风险(误差)和信任区间(在参数集合上)的模型。

在本章中, 我们着重介绍 AIC 标准、BIC 标准和 GRIC 标准及其在视觉模型选择中的应用。

18.2 AIC 与模型选择

18.2.1 AIC 标准

Akaike 在 1973 年为模型选择建立了一个信息标准, 称为 AIC 标准, 并应用于时间序列的自动回归模型。AIC 标准选择对未来数据具有最小期望预测误差的模型作为最优模型。下面, 我们简要给出 AIC 标准的推导过程。

假定 \mathcal{X} 是当前观测数据, 它的对数似然的负 2 倍记为

$$\mathcal{L}(\mathcal{X}|\theta, \mathcal{M}) = -2 \log(\Pr(\mathcal{X}|\theta, \mathcal{M})) \quad (18.2.1)$$

在高斯误差模型中它是残差(residual)平方和, 这里我们仍称它为残差。令 $\hat{\theta}$ 是 θ 最大似然估计, 则它使残差最小化:

$$\mathcal{L}(\mathcal{X}|\hat{\theta}, \mathcal{M}) = \min \mathcal{L}(\mathcal{X}|\theta, \mathcal{M}) \quad (18.2.2)$$

一个好的模型不但要对当前观测数据的残差达到最小, 同时也应该对未来观测的残差尽可能小, 即好模型应该有预测能力。为此, 我们考虑

$$\mathcal{L}(\mathcal{X}^f|\hat{\theta}, \mathcal{M}) = -2 \log \Pr(\mathcal{X}^f|\hat{\theta}, \mathcal{M}) \quad (18.2.3)$$

其中 \mathcal{X}^f 是未来观测数据。由于 \mathcal{X}^f , $\hat{\theta}$ (当前观测数据的函数) 都是随机变量, 我们可以为模型赋予一个确定的值:

$$\mathcal{I}(\mathcal{M}) = E^f[E[\mathcal{L}(\mathcal{X}^f|\hat{\theta}, \mathcal{M})]] \quad (18.2.4)$$

其中 E^f, E 分别为未来观测数据和当前观测数的期望。 $\mathcal{I}(\mathcal{M})$ 称为模型的期望残差, 它的值越小表示模型越好。直接计算 $\mathcal{I}(\mathcal{M})$ 值是非常困难的, AIC 是它的近似值, 观测数据规模越大 AIC 值越接近于 $\mathcal{I}(\mathcal{M})$ 。下面, 我们导出 AIC。

假定模型真值为 $\Pr(\mathcal{X}|\bar{\theta}, \mathcal{M})$ 。记 $\hat{\theta} = \bar{\theta} + \Delta\hat{\theta}$ ，并在 $\bar{\theta}$ 邻域内对 $\log \Pr(\mathcal{X}^f | \hat{\theta}, \mathcal{M})$ 进行 TAYLOY 展开，我们有

$$\log \Pr(\mathcal{X}^f | \hat{\theta}, \mathcal{M}) = \log \Pr(\mathcal{X}^f | \bar{\theta}, \mathcal{M}) + l^{fT} \Delta\hat{\theta} - \frac{1}{2} \Delta\hat{\theta}^T L^f \Delta\hat{\theta} + O(\Delta\hat{\theta})^3 \quad (18.2.5)$$

其中

$$l^f = \left. \frac{\partial \log \Pr(\mathcal{X}^f | \theta, \mathcal{M})}{\partial \theta} \right|_{\theta=\bar{\theta}}, \quad L^f = \left. \frac{\partial^2 \log \Pr(\mathcal{X}^f | \theta, \mathcal{M})}{\partial \theta^2} \right|_{\theta=\bar{\theta}}$$

忽略高阶项，并注意 $\Delta\hat{\theta}$ 不依赖于 \mathcal{X}^f ，我们能得到

$$\mathcal{I}(\mathcal{M}) = -2E^f[\log \Pr(\mathcal{X}^f | \bar{\theta}, \mathcal{M})] - 2E^f[l^f] \cdot E[\Delta\hat{\theta}] + E[\Delta\hat{\theta}^T E^f[L^f] \Delta\hat{\theta}] \quad (18.2.6)$$

因 \mathcal{X}^f 和 \mathcal{X} 有相同的分布，所以

$$E^f[\log \Pr(\mathcal{X}^f | \bar{\theta}, \mathcal{M})] = E[\log \Pr(\mathcal{X} | \bar{\theta}, \mathcal{M})] \quad (18.2.7)$$

由于

$$E^f[l^f] = 0, E^f[L^f] = J \text{ (Fisher 信息矩阵)} \quad (18.2.8)$$

将(18.2.7)，(18.2.,8)代入(18.2.6)，我们有

$$\mathcal{I}(\mathcal{M}) = -2E[\log \Pr(\mathcal{X} | \bar{\theta}, \mathcal{M})] + E[\Delta\hat{\theta}^T J \Delta\hat{\theta}] \quad (18.2.9)$$

在 $\hat{\theta}$ 邻域内对 $\log \Pr(\mathcal{X} | \hat{\theta}, \mathcal{M})$ 进行 TAYLOY 展开，我们有

$$\log \Pr(\mathcal{X} | \bar{\theta}, \mathcal{M}) = \log \Pr(\mathcal{X} | \hat{\theta}, \mathcal{M}) - l^T \Delta\hat{\theta} - \frac{1}{2} \Delta\hat{\theta}^T L \Delta\hat{\theta} + O(\Delta\hat{\theta})^3 \quad (18.2.10)$$

其中

$$l = \left. \frac{\partial \log \Pr(\mathcal{X} | \theta, \mathcal{M})}{\partial \theta} \right|_{\theta=\hat{\theta}}, \quad L = \left. \frac{\partial^2 \log \Pr(\mathcal{X} | \theta, \mathcal{M})}{\partial \theta^2} \right|_{\theta=\hat{\theta}}$$

令 $\bar{L} = \left. \frac{\partial^2 \log \Pr(\mathcal{X} | \theta, \mathcal{M})}{\partial \theta^2} \right|_{\theta=\bar{\theta}}$ ，则必有 $L = \bar{L} + O(\Delta\hat{\theta})$ ，将它代入(18.2.10)，忽略高阶项，并注意

$l^T \Delta\hat{\theta} = 0$ （因 $\hat{\theta}$ 最大化似然 $\Pr(\mathcal{X} | \theta, \mathcal{M})$ ），我们有

$$\log \Pr(\mathcal{X} | \bar{\theta}, \mathcal{M}) = \log \Pr(\mathcal{X} | \hat{\theta}, \mathcal{M}) - \frac{1}{2} \Delta\hat{\theta}^T \bar{L} \Delta\hat{\theta} \quad (18.2.11)$$

根据大数定律，当观测数据规模充分越大时，

$$\log \Pr(\mathcal{X} | \bar{\theta}, \mathcal{M}) \approx \log \Pr(\mathcal{X} | \hat{\theta}, \mathcal{M}) - \frac{1}{2} \Delta \hat{\theta}^T J \Delta \hat{\theta} \quad (18.2.12)$$

将上式代入(18.2.9)，我们有

$$\mathcal{I}(\mathcal{M}) = -2E[\log \Pr(\mathcal{X} | \hat{\theta}, \mathcal{M})] + 2E[\Delta \hat{\theta}^T J \Delta \hat{\theta}] \quad (18.2.13)$$

因 $\Delta \hat{\theta}^T J \Delta \hat{\theta}$ 服从自由度 p (模型的参数数目) 的分布，所以有 $E[\Delta \hat{\theta}^T J \Delta \hat{\theta}] = p$ 。因此，当观测数据规模充分越大时，

$$\mathcal{I}(\mathcal{M}) \approx -2E[\log \Pr(\mathcal{X} | \hat{\theta}, \mathcal{M})] + 2p \quad (18.2.14)$$

令

$$AIC(\mathcal{M}) = -2 \log \Pr(\mathcal{X} | \hat{\theta}, \mathcal{M}) + 2p \quad (18.2.15)$$

根据中心极限定理，当观测数据规模趋于无穷大时， $AIC(\mathcal{M})$ 趋于 $\mathcal{I}(\mathcal{M})$ 。(18.2.15) 被称为模型的 AIC 信息标准，并简记为：

$$AIC = -2\mathcal{L} + 2p \quad (18.2.16)$$

其中， p 是模型的参数数量， \mathcal{L} 是观测数据的最大似然对数。

可以看出，AIC 有两项，第一项对应数据拟合的恶劣程度，在高斯误差模型中它等价于（测量值与估计值）残差(residual)平方和，第二项为模型复杂度的惩罚。当有多个竞争模型时，模型参数值由最大似然进行估计，AIC 为每个模型记分，具有最小 AIC 值的模型作为被接受的选择。此过程叫做最小 AIC 过程，具有最小 AIC 值的模型称为最小 AIC 估计(MAICE)，被选择为最佳模型。因此，最佳模型具有最高的信息容量和最低的复杂度。最小 AIC 方法的优点在于它的简单性，当模型的最大似然估计已知时 AIC 值很容易计算，不涉及查表。并且，此方法可以指定任意的模型可以接受的显著性水平，对于进行比较的模型不需要是嵌套的或排序的。

18.2.2 用 AIC 选择模型

AIC 方法假定最好的模型应该最小化未来数据的 SSE。考虑用一个 d 维流形拟合 D 维数据点的情况，此时余维数为 $D-d$ 。Kanatani 指出，在这种情况下 AIC 为

$$AIC = -2\mathcal{L} + 2(dn + k) \quad (18.2.17)$$

二维情况下 $D=2$ ，拟合线模型时 $d=1$ ，点模型时 $d=0$ 。

观察方程(18.2.17)，第一项为变量残差的平方和，描述的是数据与模型匹配的恶劣程度，后两项描述模型的简约程度：第二项是对模型维度的惩罚项，模型的维度越高，惩罚越大；最后一项是对模型本质参数数量进行惩罚，用来惩罚多参数模型。

下面用一个简单的例子加以说明，考虑图 18.2.1 中的二维数据拟合的案例。假设测量点来自点

模型或线模型，其 x, y 坐标有零均值单位标准差的高斯噪声。使用哪一个模型会更好呢？用最小欧氏距离平方和分别拟合一个点和一条直线，则最优拟合点(为数据点集的重心)必在最优拟合线上(见直线估计的几何方法，第 16 章)。设点到线模型的距离平方和为 $\sum e_{\perp}^2$ ，点到点模型的距离平方和为 $\sum e_p^2$ ，则 $\sum e_p^2 = \sum e_{\perp}^2 + \sum e_{\parallel}^2$ ，其中 $\sum e_{\parallel}^2$ 为到点的平行距离平方和，如图 1 所示。可以看出，除非数据都精确到重心点上，否则 $\sum e_{\perp}^2$ 总小于 $\sum e_p^2$ 。如果被估计的模型是一条直线，因线模型维数为 1，余维数也为 1，参数有两个自由度，所以它的 AIC 为

$$AIC(line) = \sum e_{\perp}^2 + 2n + 4 \tag{18.2.18}$$

即使数据量很大，模型的自由度(即本质参数数量)不影响 AIC 值，因为它对应的惩罚项是一个常数。起决定作用的是模型维数两倍与数据规模的乘积，它与数据集的内在自由度相同，是对估计点到真值偏差(估计误差)平方和的补偿，即对每个数据点的平行误差平方和的补偿。对于点模型，我们有

$$AIC(point) = \sum e_{\perp}^2 + \sum e_{\parallel}^2 + 4 \tag{18.2.19}$$

因此，如果 $AIC(point) \leq AIC(line)$ ，即 $\sum e_{\parallel}^2 \leq 2n$ ，则点模型更优。因此，此算法等价于沿着直线进行检验。

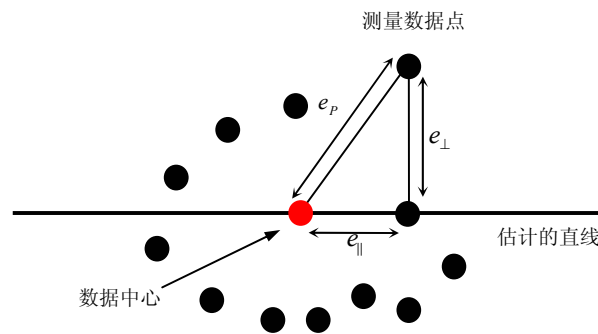


图 18.2.1 点模型与线模型的 AIC

AIC 的实验结果

表 3 给出了来自 100 个数据点，每个数据点的测量误差服从零均值标准差为 1 的高斯分布，进行 100 次实验的平均 SSE，其中括号中的数字是正确模型的期望值。 F, F_A, H 和 H_A 的 $2(nd + k)$ 值分别为

$$2p_F = 2(100 \cdot 3 + 7) = 614, \quad 2p_{F_A} = 2(100 \cdot 3 + 4) = 608,$$

$$2p_H = 2(100 \cdot 2 + 8) = 416, \quad 2p_{H_A} = 2(100 \cdot 2 + 6) = 412,$$

因此从表 18.2.1 可计算出各模型的 AIC 值，结果在表 18.2.2 中。可以看出，虽然 F 与 F_A 的区分效果不如 F 与 H ，可是具有最低 AIC 的模型等与正确模型一致。

表 18.2.1: 100 个点对应 100 次实验的平均误差平方和，每个坐标上的噪声方差为 1，括号中的数字是正确模型的期望值

SSE	Point Motion			
	General F	Orthographic F_A	Rotation H	Affinity H_A
Fundamental F	93.074 (93)	87.0370	80.6162	78.378
Affine F_A	978.350	96.448 (96)	806.389	85.875
Homography H	4986.881	4834.735	193.964 (192)	189.132
Affinity H_A	4993.045	4967.894	1023.118	191.643 (194)

表 18.2.2: 100 个点对应 100 次实验的平均 AIC，可以看出 AIC 趋于过拟合

AIC	Point Motion			
	General F	Orthographic F_A	Rotation H	Affinity H_A
Fundamental F	707.07	701.04	694.62	692.38
Affine F_A	1586.35	704.45	1414.39	683.86
Homography H	5402.88	5240.74	609.96	605.13
Affinity H_A	5405.05	5379.89	1435.12	603.64

表 18.2.3: 由自由度为 $|p_1 - p_2|$ 的 χ^2 分布的显著水平和临界值 $2|p_1 - p_2|$ 给出的 AIC 标准的显著水平

$ p_1 - p_2 $	1	2	3	4	5	6	8
α	0.156	0.135	0.111	0.091	0.074	0.051	0.042
$ p_1 - p_2 $	10	12	14	20			
α	0.029	0.020	0.014	0.005			

一般情况下，AIC 过低估计数据的维数而过高估计运动模型的本质参数数量，其原因可在前面介绍过的似然比检验中找到。用 AIC 对两个模型 M_1 与 M_2 进行比较，具有最小 AIC 的模型会被接受，即如果 $AIC_2 - AIC_1 < 0$ ，或者如果

$$2(\mathcal{L}_1 - \mathcal{L}_2) < 2(p_1 - p_2) \quad (18.2.20)$$

则接受模型 \mathcal{M}_2 , 可以看出它与(18.1.2)是等价的。 AIC 标准的显著水平可由自由度为 $|p_1 - p_2|$ 的 χ^2 分布的显著水平和临界值 $2|p_1 - p_2|$ 给出。当 $|p_1 - p_2| = 2$ 时, H 和 H_A 的本质参数数量的不同将导致 $\alpha = 0.135$, 或者说以 13.5% 的机会产生过拟合; 当 $|p_1 - p_2| = 3$ 时, F 与 F_A 的参数数量的本质不同导致 $\alpha = 0.11$, 或者说以 11% 的机会产生过拟合, 一些典型值在表 18.2.3 中给出。这可以在实验中证实, 从表 18.2.3 可以找到。例如, 第 4 列的单应与仿射单应的 AIC (对应于 $|p_1 - p_2| = 2$ 的情况), 此数据由仿射单应产生的, 但是它们的平均 AIC 值却几乎差相同。当随机产生的数据偶然处于退化配置 (generate configuration) 时, 欠拟合就会发生。考虑将一般运动模型分解为一个平面运动加上一个平移运动模型, β 的大小依赖于平行方向的误差平方和, AIC 的惩罚因子 2 意味着平行方向的误差平方和的平均值应该小于 2.0σ 个像素, 此时的一般运动才是单应关系。

随着 $p_1 - p_2$ 数值的增加 α 变得越来越小(当 $p_1 - p_2 = 20$ 时, α 小于 0.005), 过拟合的机会减少。但是, 当数据由一个光滑曲面产生时不是这种情况, 表 18.2.4 中的平面数据就是这样产生的, 对于一般运动和正交运动, 产生的数据位于一个光滑曲面上。一般地, 对于光滑曲面, 平行方向的误差会减小, 此时 AIC 也会出现欠拟合情况。

表 18.2.4: 100 次实验中每个模型被选择的次数, 可以看出 AIC 趋于过拟合

Model Selection	Point Motion			
	General F	Orthographic F_A	Rotation H	Affinity H_A
Fundamental F	99	12	0	0
Affine F_A	1	88	0	0
Homography H	0	0	98	15
Affinity H_A	0	0	2	85

上述实验案例表明 AIC 的关键不足之处在于, 它对每个参数只是指定惩罚因子 2——Akaike 称它为幻数 (magic), 这是渐进不一致的, 随着观察数据的增加, 错误模型将会渐进地被选择。认识到 AIC 过拟合的趋势, Leontaritis 和 Billings 建议用幻数 4 代替幻数 2, 可是对于小样本情况这会导致欠拟合, 因为检验能力只随观察数据数量的增加而增加。在模型选择中用一个常数因子来惩罚似然可能是合适的, 但用一个常数因子对参数数量进行惩罚, 而不考虑这些参数的来源与它们的不确定性就显得有些令人费解。实际上, Akaike 本人也明确声明, 只有当 Bayes 分析不能进行时 AIC 才有用——统计分析的目的是选择更好的模型, 从这个角度看, Bayes 模型总是优于其它非

Bayes 方法。然而，当对问题的深入 Bayes 分析不可行时，对数似然的客观性作为随机模型的一种评价为我们提供了模型选择的坚实基础。

18.3 BIC 与模型选择

18.3.1 Bayes 证据

本节给出模型比较的 Bayes 方法。Bayes 方法与上节方法有很大的差别，“一个流行的荒诞说法却是，Bayes 方法与传统方法的不同仅仅是它包含任意的很难确定的主观先验概率”。下面给出如何应用后验分布来估计每个模型的似然，它涉及到在整个参数空间上的积分。

给定可以描述数据 \mathcal{X} 的模型为 $\mathcal{M}_1 \dots \mathcal{M}_K$ ，则 Bayes 规则为

$$\Pr(\mathcal{M}_j | \mathcal{X}, \mathcal{I}) = \frac{\Pr(\mathcal{X} | \mathcal{M}_j, \mathcal{I}) \Pr(\mathcal{M}_j | \mathcal{I})}{\Pr(\mathcal{X} | \mathcal{I})} \quad (18.3.1)$$

其中， \mathcal{I} 是关于现实世界的先验信息假设。例如，在两视点问题的情况下，使用四个基本运动模型是合适的，因为模型 F, F_A, H 和 H_A 能描述大部分情况。当然，运动模型还可能包括其它运动模型，如纯平移运动。给定的模型 \mathcal{M}_i 是正确模型的后验概率为

$$\Pr(\mathcal{M}_i | \mathcal{X}, \mathcal{I}) = \frac{\Pr(\mathcal{X} | \mathcal{M}_i, \mathcal{I}) \Pr(\mathcal{M}_i | \mathcal{I})}{\sum_{j=1}^{j=K} \Pr(\mathcal{X} | \mathcal{M}_j, \mathcal{I}) \Pr(\mathcal{M}_j | \mathcal{I})} \quad (18.3.2)$$

不妨假定 $\sum_{j=1}^K \Pr(\mathcal{M}_j | \mathcal{X}, \mathcal{I}) = 1$ 。数据对 \mathcal{M}_i 的相对于 \mathcal{M}_j 的支持程度可由后验比来度量，即

$$o_{ij} = \frac{\Pr(\mathcal{M}_i | \mathcal{X}, \mathcal{I})}{\Pr(\mathcal{M}_j | \mathcal{X}, \mathcal{I})} = \frac{\Pr(\mathcal{X} | \mathcal{M}_i, \mathcal{I})}{\Pr(\mathcal{X} | \mathcal{M}_j, \mathcal{I})} \cdot \frac{\Pr(\mathcal{M}_i | \mathcal{I})}{\Pr(\mathcal{M}_j | \mathcal{I})} \quad (18.3.3)$$

其中 $B_{ij} = \frac{\Pr(\mathcal{X} | \mathcal{M}_i, \mathcal{I})}{\Pr(\mathcal{X} | \mathcal{M}_j, \mathcal{I})}$ 称为 Bayes 因子。可以看出分子与分母都是两项的乘积：一是给定模型 \mathcal{M}_i 时

数据 \mathcal{X} 的概率，称为模型 \mathcal{M}_i 的证据；另一个是模型 \mathcal{M}_j 的先验概率。Bayes 因子似乎很像带有附加先验概率的模型比较的似然比方法，它们之间的关键差别在于，这里的每个模型证据实际上是似然在模型参数空间上的积分：

$$\Pr(\mathcal{X} | \mathcal{M}_j, \mathcal{I}) = \int \Pr(\mathcal{X} | \mathcal{M}_j, \theta_j, \mathcal{I}) \Pr(\theta_j | \mathcal{M}_j, \mathcal{I}) d\theta_j \quad (18.3.4)$$

即，

$$\text{Evidence} = \int \text{likelihood} \times \text{prior} d\theta_j \quad (18.3.5)$$

其中 θ_j 是第 j 个模型的参数， $\Pr(\theta_j | \mathcal{M}_j, \mathcal{I})$ 是模型参数的先验分布。由定义， $\int \Pr(\theta_j | \mathcal{M}_j, \mathcal{I}) d\theta_j = 1$ 。(18.3.3)式的右边第一项是由(18.3.5)式给出的两个积分的比，所以(18.3.3)式可重写为

$$\text{posterior odds} = \text{Bayes Factor} \times \text{prior odds} \quad (18.3.6)$$

第二项是先验比，这里设为 1，即我们对两个模型没有任何的先验偏好，即 $\Pr(\mathcal{M}_i) = \Pr(\mathcal{M}_j)$ ，也就是说所有模型有相同的先验概率。方程(18.3.5)是一个关键的方程，它描述了一个给定模型的证据，因此值得仔细研究。如果你接受 Bayes 规则，则逻辑表明它是计算每个模型后验概率的唯一方法。

Bayes 方法的不足是，模型的选择依赖于先验概率，或者我们早就这么认为，“你也是这么认为，它越可能存在”。可是，这会往往使我们将后验分布(我们对这个世界的知识陈述)与现实世界相混淆(Jaynes 关于思想投影谬误)。这里描述的所有方法都作了隐含的先验分布假设，即使最大似然方法也等价于假设参数有均匀的先验概率分布。如果我问你：图 18.2.1 中的数据是来自一个点还是来自一条直线的观测，你会通过问我的先验看法而做出精确判断，如问我希望这条直线多长，希望误差分布是什么样的问题。你会根据我告诉你：“我希望这条线 1.5 个像素长还是 100 个像素长”，而作出不同的决策。Bayes 公式只是强迫承认我们自己的假设，并且把假设放在判断之前，而不是从数学上揭示它们。事实上，没有必要在 Bayes 方法中使用比先验概率可能性小的更复杂的模型。而且，简单性的思想来源于这样的事实，拥有更多参数的模型将会有更广的先验分布，因此应该比简单模型有更大的似然。

如何利用这些信息来选择合适的模型还需要 Bayes 决策理论这个工具。我们有一个潜在的行动集合 $a_1 \dots a_k \in \mathcal{A}$ ，每个行动对应于从模型集 $\{\mathcal{M}_1 \dots \mathcal{M}_k\}$ 中选择一个模型。为了确定一个行动过程，在给定现实世界的一个状态 $M \in \mathcal{M}$ 的情况下，为每个行动 $a \in \mathcal{A}$ 定义一个表示行动效益的损失函数 $u: \mathcal{A} \times \mathcal{M} \rightarrow \mathbb{R}$ 。最优行动是使期望效益最大的行动： $\hat{a} = \max_a \sum_j u(a, \mathcal{M}_j) \Pr(\mathcal{M}_j | \mathcal{X}, \mathcal{I})$ 。最简单的情况是 0—1 损失函数，当模型正确是，回报是 1，否则是 0。这种损失函数的模型选择方法将最大化后验概率，它是本章所采用的损失函数，当然根据实际情况，其他方法可能也很有效。

18.3.2 BIC 标准

最理想的方法是直接确定证据(18.3.5)，对于实际问题往往是要计算高维空间上的积分，例如，

在一般运动模型选择问题中，需要在 $k+nd$ 个连续参数和 n 个离散参数构成的参数空间上计算积分：

$$\Pr(\mathcal{D}|\mathcal{M}_j, \mathcal{I}) = \int_{\alpha} \int_{\beta} \int_{\gamma} \prod_{i=1 \dots n} (\psi(e_i) \Pr(\beta_i | \alpha, \mathcal{M}, \mathcal{I})) \Pr(\gamma | \alpha, \mathcal{M}, \mathcal{I}) \Pr(\alpha | \mathcal{M}, \mathcal{I}) d\alpha d\beta d\gamma \quad (18.3.7)$$

显然十分困难。所以，通过直接计算证据来选择模型的方法是不可行的。

下面介绍一种计算证据的近似方法，它将导致 BIC 标准。我们想要为每个模型计算证据

$$\Pr(\mathcal{X}|\mathcal{M}_j, \mathcal{I}) = \int_{\Theta_j} \Pr(\mathcal{X}|\mathcal{M}_j, \Theta_j, \mathcal{I}) \Pr(\Theta_j|\mathcal{M}_j, \mathcal{I}) d\Theta_j \quad (18.3.8)$$

为此，记对数后验为

$$\phi(\Theta) = \log(\Pr(\mathcal{X}|\mathcal{M}, \Theta, \mathcal{I})) \Pr(\Theta|\mathcal{M}, \mathcal{I}) \quad (18.3.9)$$

令 $\hat{\Theta}$ 为参数 Θ 的 MAP 估计，它最大化后验概率。记 $\phi(\Theta)$ 的 Hessian 矩阵为 $\nabla^2 \phi$ ，则 $\hat{\Theta}$ 的协方差矩阵可由 Hessian 矩阵的逆 $\Lambda = [\nabla^2 \phi]^{-1}$ 计算出来。因此，Bayes 证据为

$$\Pr(\mathcal{X}|\mathcal{M}, \mathcal{I}) = \int \exp(-\phi(\Theta)) d\Theta \quad (18.3.10)$$

对 $\phi(\Theta)$ 在 $\hat{\Theta}$ 点 Taylor 展开，使用推导 AIC 的类似方法，我们能得到

$$\begin{aligned} \Pr(\mathcal{X}|\mathcal{M}, \mathcal{I}) &\approx \exp(-\phi(\hat{\Theta})) \int \exp\left(-\frac{1}{2}(\hat{\Theta} - \Theta)^{\top} \Lambda^{-1} (\hat{\Theta} - \Theta)\right) d\Theta \\ &= (2\pi)^{p/2} \det \Lambda^{1/2} \exp(-\phi(\hat{\Theta})) \end{aligned}$$

其中 p 是系统中所有自由参数的数目。上面的最后一步是多变量高斯积分的 Laplace 近似的标准结果。于是，证据对数为

$$\log(\Pr(\mathcal{X}|\mathcal{M}, \mathcal{I})) \approx \log(\Pr(\mathcal{X}|\mathcal{M}, \hat{\Theta}, \mathcal{I})) + \frac{p}{2} \log 2\pi + \frac{1}{2} \log \det(\Lambda) + \log(\Pr(\hat{\Theta}|\mathcal{M}, \mathcal{I})) \quad (18.3.11)$$

它包含四项：似然项，参数数目项，确定性项以及先验项。

从(18.3.11)式，可以得到多种计算量小的信息标准，其中最简单的是 Schwarz 所给出的方法，他将先验分布近似为正态分布。计算机视觉中大多数问题的参数都服从高斯先验分布。设先验分布为高斯分布，其均值为 Θ_{prior} ，协方差为 Λ_{prior} ，则有

$$\Pr(\hat{\Theta}|\mathcal{M}, \mathcal{I}) = (2\pi)^{-p/2} \det \Lambda_{prior}^{-1/2} \exp\left(-\frac{1}{2}(\Theta_{prior} - \hat{\Theta})^{\top} \Lambda_{prior}^{-1} (\Theta_{prior} - \hat{\Theta})\right) \quad (18.3.12)$$

于是, 我们得到

$$\log(\Pr(\mathcal{X}|\mathcal{M}, \mathcal{I})) \approx \mathcal{L} - \frac{1}{2}(\boldsymbol{\Theta}_{prior} - \hat{\boldsymbol{\Theta}})^T \Lambda_{prior}^{-1} (\boldsymbol{\Theta}_{prior} - \hat{\boldsymbol{\Theta}}) + \frac{1}{2} \log \frac{\det(\Lambda)}{\det(\Lambda_{prior})}$$

其中 $\Lambda = \Lambda_{prior} \Lambda_L$, Λ_L 是对数似然在 $\hat{\boldsymbol{\Theta}}$ 点的协方差。因此,

$$\log(\Pr(\mathcal{X}|\mathcal{M}, \mathcal{I})) \approx \mathcal{L} - \frac{1}{2}(\boldsymbol{\Theta}_{prior} - \hat{\boldsymbol{\Theta}})^T \Lambda_{prior}^{-1} (\boldsymbol{\Theta}_{prior} - \hat{\boldsymbol{\Theta}}) + \frac{1}{2} \log \det(\nabla^2 \mathcal{L}) \quad (18.3.13)$$

如果假设先验分布很扩散 (即几乎没有, 或者没有先验知识, 对于高斯分布而言标准差非常大), 先验先验知识非常不确定, Schwarz 建议忽略涉及先验的第二项。同时, Schwarz 还建议通过 Hessian 矩阵的渐进期望值近似协方差的逆, 即 $\frac{1}{2} \log \det(\nabla^2 \mathcal{L}) \approx \frac{p}{2} \log N$ 。于是, 得到证据对数的近似:

$$\log(\Pr(\mathcal{X}|\mathcal{M}, \mathcal{I})) \approx \log \Pr(\mathcal{X}|\hat{\boldsymbol{\Theta}}, \mathcal{I}) - \frac{2}{p} \log N = \mathcal{L} - \frac{p}{2} \log N \quad (18.3.14)$$

其中 $p=dn+k$ 为系统中所有参数的数量, N 为所有观察数据的数量。(18.3.14)结果的负二倍:

$$\text{BIC}(\mathcal{M}) = -2\mathcal{L} + p \log N \quad (18.3.15)$$

称为 Bayes 信息标准或 BIC 标准。事实上, BIC 标准不能看作是 Bayes 的, 因为它忽略了先验项, 称它为 Schwarz 信息准标可能更贴切。

18.3.3 用 BIC 选择模型

假定第 j 个模型的 BIC 值

$$\text{BIC}(\mathcal{M}_j) = -2\mathcal{L}_j + p_j \log N \approx \log \text{Evidence} + \text{constant} \quad (18.3.15)$$

其中 N 为观察数据的个数。用 BIC 标准选择模型的决策是, 具有最小 BIC 的模型是最可能的模型。假设模型集 $\{\mathcal{M}_j\}$ 上的先验是均匀分布, 则每个模型的后验概率可近似为

$$\Pr(\mathcal{M}_j|\mathcal{X}, \mathcal{I}) = \frac{\exp(-\text{BIC}(\mathcal{M}_j))}{\sum_{i=1}^{i=K} \exp(-\text{BIC}(\mathcal{M}_i))} \quad (18.3.16)$$

由于 BIC 的形式简单且易于计算, 对解决模型选择问题很有吸引力。但对于计算机视觉中的模型选择问题, BIC 标准偏向选择低维模型, 存在一致地欠拟合现象。这可能是 BIC 对 Bayes 证据不良近似的结果, 因为对于视觉问题, 在式(18.3.4)中包含大量的潜在参数。下面给出 BIC 的一些测试

结果。

表 18.3.1: 100 次实验中每个模型被选择的次数, 可以看出 BIC 趋于过拟合

Model Selection	Point Motion			
	General F	Orthographic F_A	Rotation H	Affinity H_A
Fundamental F	69	1	0	0
Affine F_A	17	73	0	0
Homography H	14	25	96	2
Affinity H_A	0	1	4	98

考虑 100 个匹配, 观察数据的为 $N=400$ 的情况, 此时基本矩阵参数数量为 $p=307$, 单应应为 $p=208$, 因此 BIC 对基本矩阵的惩罚项为 $307\ln 400$, 而对单应的惩罚项为 $208\ln 400$, 两者相差 593.15。所以, 在两视点情况下单应会是非常糟糕的拟合, BIC 对它的惩罚往往比对基本矩阵的惩罚要小, 这种情况只在当基线很长和透视效果很明显时才能发生。当考虑具有相同维数代数簇模型选择时, 情况会更糟, 见表 18.3.2。表 18.3.1 给出了 BIC 模型选择的结果。可以看出, 对于基本矩阵来说, BIC 对维数存在欠拟合, 偏向低维模型, 这是因为基本矩阵估计在平行方向的误差很高。单应和仿射单应都没有表现出这样明显的特性, 因为它们的维数相同并且在四个模型中最低。理论上, BIC 没有令人满意的结果是正常的, 因为它省略了先验知识, 不是一种真正的 Bayes 方法。

表 18.3.2: 100 个点对应 100 次实验的平均 BIC

BIC	Point Motion			
	General F	Orthographic F_A	Rotation H	Affinity H_A
Fundamental F	1932.45	1926.42	1920.00	1917.76
Affine F_A	2799.76	1917.86	2627.80	1907.29
Homography H	6233.10	6080.96	1440.18	1435.35
Affinity H_A	6227.29	6202.13	2256.36	1425.88

18.4 GRIC 与模型选择

18.4.1 鲁棒最小二乘模型的 GRIC 标准

对视觉模型选择问题, 不论是使用 AIC 还是 BIC 标准都不能得到有效的结果, 主要原因可能是

视觉模型有它的特殊性，如它们大多是高斯最小二乘问题，有重多的潜在参数。AIC 或 BIC 是在一般情况下对期望残差或对 Bayes 证据的近似，没有考虑视觉模型的特点，对参数数 p 惩罚不论是本质参数还是潜在参数都一视同仁，使用相同的惩罚因子，这看上去显然是不合理的。我们回忆视觉模型的参数数量，它由两部分组成 $p=k+nd$ ，其中 k 是本质参数数量； n 是测量数据点的数量； d 是模型流形的维数。 d 等于每个测量点对应的潜在参数的自由度，因此 nd 正好是与所有测量点对应的潜在参数的自由度。这样，在视觉模型两部分参数中，一部分对应本质参数，另一部分正好对应潜在参数。于是，可以构造对参数惩罚的更一般的形式：

$$GRIC = -2\mathcal{L} + \lambda_1 nd + \lambda_2 k \quad (18.4.1)$$

可能对视觉模型选择更合适，因为我们可通过选择惩罚因子 λ_1, λ_2 用来纠正错误拟合。对 λ 的不同估计将导致不同的 GRIC。

Torr 首先针对运动模型选择问题，通过选择 $\lambda_1 = \log(4) = 1.4, \lambda_2 = \log N$ ，给出一种特别形式的 GRIC：

$$GRIC = -2\mathcal{L} + nd \log(4) + k \log(N) \quad (18.4.2)$$

其中 $N=4n$ (n 为对应点的个数)。在匹配独立的假设下，每个匹配的最优估计 $\hat{\mathbf{X}}_i$ (自由度为 d 的潜在参数的估计， d 为模型流形的维数) 只受到匹配 \mathbf{X} 的 4 个有噪坐标分量的影响，由匹配的分块对角协方差矩阵刻画。从直觉上，这是一种好方法，随着测量点的增加， nd 将增大，并且对应的潜在参数的估计精度不会得到提高从而误差平方和也将增大，因此对 nd 惩罚不能随数据规模的增大而增大，应该是一个常数。对本质参数数目 k 却不是这样，随着测量点的增加， k 不变但对应的本质参数的估计精度有所提高，所以模型越复杂误差平方将越小，因此需要一个随数据规模变化而变化的惩罚因子对它进行惩罚。实验也证明这是一种较好的信息标准。

下面是 Torr 给出的另一种估计惩罚因子 λ_1 方法，它将导致另一种 GRIC。在模型流形上，只要大部分的点没有分布在特别高的曲率附近，根据式(17.4.11)和(17.4.16)，对 γ 取期望值，我们有

$$\Pr(\mathcal{X} | \mathcal{M}_j, \mathcal{I}) = \int_{\alpha} \prod_{i=1 \dots n} \left(\gamma \left(\frac{(\sqrt{2\pi\sigma^2})^{d-D}}{c} \exp\left(-\frac{e_i^2(\alpha)}{2\sigma^2}\right) \right) + (1-\gamma) \frac{1}{v} \right) \Pr(\alpha | \mathcal{X}, \mathcal{I}) d\alpha \quad (18.4.3)$$

其中， $e_i^2(\alpha)$ 是用参数 α 估计的流形的第 i 个误差。这是对视觉估计特殊形式的关键近似，与前面提到的 Bayes 证据估计有根本的区别。设 $\hat{\alpha}$ 为 MAP 估计，定义

$$\mathcal{L}_{MAP} = \sum_{i=1 \dots n} \log \left(\gamma \left(\frac{(\sqrt{2\pi}\sigma^2)^{d-D}}{c} \exp \left(-\frac{e_i^2(\hat{\alpha})}{2\sigma^2} \right) \right) + (1-\gamma) \frac{1}{v} \right) \quad (18.4.3)$$

因为 $\int_{\alpha} \partial \alpha = 1$, $-\log(\Pr(\mathcal{X} | \mathcal{M}_j, \mathcal{I}))$ 的最小值为 $-\mathcal{L}_{MAP}$ 。当模型不可辨识(协方差矩阵接近无穷)时, 在所有的 α 上, 后验分布有相同的值。因此, 如果模型不可辨识时, 可以把它作为 **GRIC** 的值。下面, 针对后验分布, 给出当模型不可辨识情况下的一种新的 **GRIC** 标准。

在理论上, 证据的负对数可能是无界的。通过类似与 **BIC** 的分析, 对于很大的 n , 给出证据(后验分布的容量)的一个粗糙的估计, 这次我们只应用到 α 。和以前的 **GRIC** 标准相比, 这将会产生一个更一般的公式:

$$GRIC = -2\mathcal{L}_{MAP} + k \log n \quad (18.4.4)$$

其中 k 为 α 中参数的数量。这看上去和 **BIC** 很相似, 所以应该仔细考虑这个统计量。如果不存在外点, 则有

$$GRIC = \sum_i \frac{e_i^2}{\sigma^2} + \lambda_1 n d + \lambda_2 k + \text{constant} \quad (18.4.5)$$

其中

$$\lambda_1 = \log \left(\frac{S^2}{2\pi\sigma^2} \right), \quad \lambda_2 = \log(n) \quad (18.4.6)$$

式(18.4.6) 中的 λ_1 请参考式(17.4.21a, 17.4.21b)。在式(18.4.5)中包括三项: 一个标准误差平方和项, 一个正比于 β 的规模的流形维数的惩罚项和一个正比于 α 规模的权重项。与前面的 **GRIC** 标准不同之处在于这里的 λ_1 依赖于图像上视差的期望分布。分布越稀疏 λ_1 越大, 如搜寻区域的规模 $S=20$, $\sigma=2.0$, 则 $\lambda_1=2.8$, 这比 **AIC** 的两项惩罚因子都大; 如果 $S=10$, $\sigma=2.0$, 则 $\lambda_1=1.38$ 。如此分析预示着在搜寻区域, 内点均匀的分布在视差上, 如此近似只是为简化起见。但实际情况并非总是如此, 因此在其它情况下, 应该根据 β_i 的精确分布重新估计 λ_1 。当有外点存在时, 式(18.4.4)可以在 γ 上最大化得到近似估计

$$GRIC = \sum_i \rho_{map} \left(\frac{e_i^2}{\sigma^2} \right) + \lambda_1 n d + \lambda_2 k + \text{constant} \quad (18.4.6)$$

其中

$$\rho_{map} \left(\frac{e^2}{\sigma^2} \right) = \begin{cases} e^2/\sigma^2, & e^2/\sigma^2 < T_{map} \\ T_{map}, & e^2/\sigma^2 \geq T_{map} \end{cases} \quad (18.4.7)$$

其中 T_{map} 由(17.4.21)式给出。

18.4.2 用 GRIC 选择模型

尽管我们已用了很大的篇幅介绍了模型选择问题，由于它的复杂性读者或许没有明白要点是什么。下面用运动模型来简要陈述模型选择的要点。运动模型的重要的目标是应用模型来引导两幅图像的点对应，算法细节比较复杂，但思想却很简单：

1. 用交叉相关(cross correlation)产生一个匹配集合。
2. 鲁棒估计每个模型。
3. 计算每个模型的证据。
4. 应用最佳模型重新引导匹配。

如果使用了错误模型，要么欠拟合发生(结果使匹配减少)，或者过拟合发生(结果里增加了好多伪匹配)发生。

人造数据

三个人造数据库，每个库包含 100 个匹配集合，每个集合有 100 个人造匹配，其中有 10—30% 的外点，它们随机地与 F, F_A , H 或 H_A 一致。每个模型选择算法都在每个匹配集合上运行一遍，被选择的模型与现实情况进行比较。结果见表 18.4.1，可以得出结论：外点的存在使得模型选择过程确定性降级。

表 18.3.1：在具有外点的 100 次实验中，使用 GRIC 每个模型被选择的次数

Model Selection	Point Motion			
	General F	Orthographic F_A	Rotation H	Affinity H_A
Fundamental F	97	1	1	0
Affine F_A	1	95	0	2
Homography H	2	2	99	2
Affinity H_A	0	2	0	96

真实数据

在所用的案例中，角点通过 HARRIS 探测器获得，匹配过程在一个方形搜索窗口中进行。点对应的误差标准差 σ 应用 TORR 鲁棒方法给出估计。模型比较检测与一般情况相比要简单得多，这是因为为了那个模型是真实的，需要知道真实情况是定性的。也就是说，如果是一般运动切实可辨识的透视效果，则 F 是真实模型；如果摄像机围绕光心旋转，则 H 是真实模型；如果场景很远接近

正交，则 F_A 是真实模型；如果所有点都在一个很远的平面上，则 H_A 是真实模型；同样，如果焦距很长并且摄像机是旋转运动，则 H_A 比较合适。图 18.4.2 和 18.4.2 各有 2 对图像，每对图像对应一种模型，模型选择算法使用 GRIC 标准。实验中给出选择结果同时，也指出最佳模型重新引导匹配的结果。

真实数据实验的 GRIC 值在表 18.4.2 中，可以看出使用 GRIC 标准，所有的选择都对应于真实模型。

表 18.3.2: 图像的 GRIC 值，黑体数字对应于最小 GRIC 值的模型

GRIC	n	Point Motion			
		General F	Orthographic F_A	Rotation H	Affinity H_A
House	80	596	618	852	955
Buggy	167	921	890	940	1150
Football	565	3183	3321	2658	2691
Cup	142	801	836	587	562

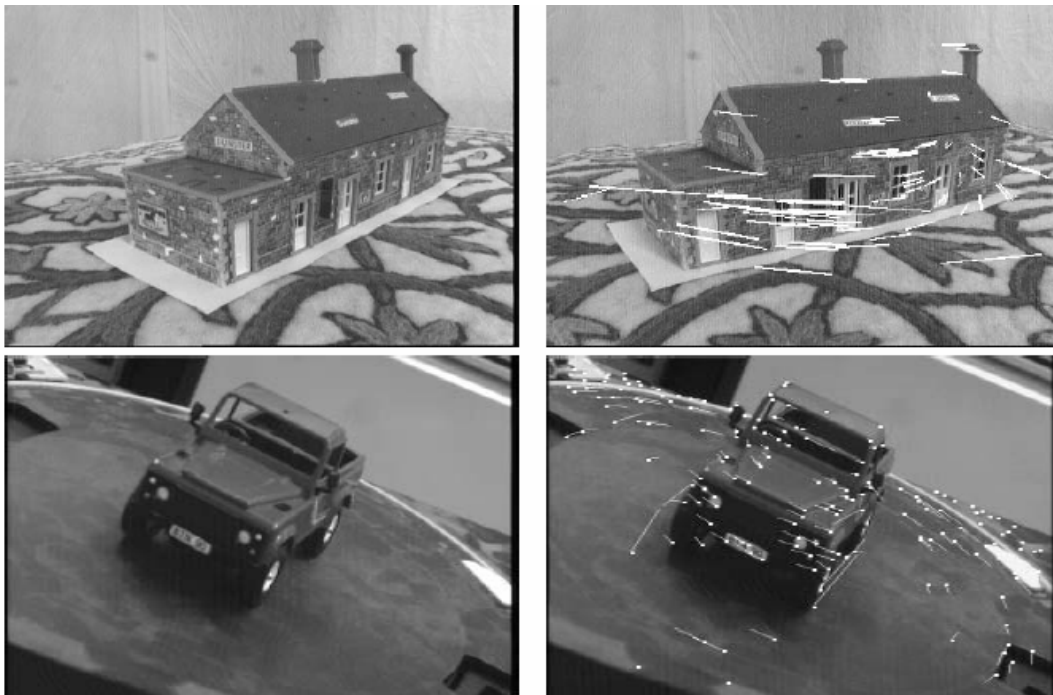


图 18.4.2 第一行：房屋，摄像机对准房屋模型旋转和平移，选择的模型是基本矩阵。第二行：童车，童车在桌面的旋转，正交运动不能产生透视结构，选择的模型是仿基本矩阵

房屋 图 18.4.2 中的前两幅图像是一个场景，摄像机对准房屋模型旋转和平移。此场景是一般运动，因为旋转和平移幅度都很大，并且整体透视结构可以恢复。在本案例中应用单应引导匹配，结果匹配点都在地平面上而没有在房屋上。

童车 第二对图像是两视点下的童车在桌面的旋转，正交运动不能产生透视结构。因此正确的模型(或者说真实的模型)是 F_A 。拟合一个单应结果是匹配点只在旋转桌面上，而拟合一个基本矩阵产生 5 个误匹配。

足球比赛 (图 18.4.4) 在此案例中，摄像机镜头旋转，正确的模型是单应，而仿射单应所得到后验概率也很接近。在这个例子中拟合一个三维模型会产生很多误匹配，原因是图像中涉及部分的错误点对应。

杯子 (图 18.4.4) 本案例中摄像机进行了扭动而且有轻微的缩放。仿射单应比较合适，可单应也可以。同样，如果三维模型被拟合的话，就会产生误匹配，这虽然与约束一致，可实际上是错误的。



图 18.4.3 第一行：足球比赛，摄像机镜头旋转，正确的模型是单应，选择的模型也是单应。第二行：杯子，摄像机进行了扭动而且有轻微的缩放。仿射单应比较合适，单应也可以，选择的模型是仿射单应

参考书目与文献

1. 矩阵论 (第二版), 程云鹏主编, 西北工业大学出版社, 西安, 1999
 2. 矩阵分析与应用, 张贤达著, 清华大学出版社, 北京, 2005
 3. 数学规划, 郑汉鼎, 刁在筠编著, 山东教育出版社, 济南, 1997
 4. 概率论与数理统计, 陈希孺编著, 科学出版社, 北京, 2005
 5. 高等数理统计, 茆诗松, 王静龙, 濮晓龙编著, 高等教育出版社, 北京, 2003
 6. 梅向明、刘增贤、林向岩, 高等几何, 高等教育出版社, 北京, 1983
 7. 微分几何, 陈省身, 陈维桓著, 北京大学出版社, 北京, 1999
 8. 积分几何与几何概率, 吴大任译, 南开大学出版社, 天津, 1986
 9. Algebraic Projective Geometry, J. G. Semple and G. T. Kneebone, Oxford University Press, Oxford , 1979.
 10. Multiple View Geometry in Computer Vision, R. Hartley and A. Zisserman, Cambridge University Press, Cambridge , 2000
 11. Faugeras and Luong, The Geometry of Multiple Images, MIT Press 2001.
 12. Computer Vision: A Modern Approach, D. A. Forsyth and J. Ponce, Pearson Education Inc. 2003
 13. Statistical Optimization for Geometric Computation: Theory and Practice, K. Kanatani, ELSEVIER,1996
-
14. 吴福朝, 李华, 胡占义, 基于主动视觉系统的摄像机自定标方法研究, 自动化学报, 第 27 卷, 第 6 期, pp.752-762, 2001
 15. 吴福朝, 胡占义, 摄像机自定标的线性理论与算法, 计算机学报, 第 24 卷, 第 11 期, pp. 1121-1135, 2001
 16. 吴福朝, 胡占义, 关于 P5P 问题的研究, 软件学报, 第 12 卷, 第 5 期, pp.768-775,2001.
 17. 吴福朝, 胡占义, 摄像机未标定的 P5P 问题研究, 计算机学报, 第 24 卷, 第 11 期, pp. 1221-1226, 2001.
 18. 吴福朝, 阮宗才, 胡占义, 非线性摄像机自标定, 计算机学报, 第 25 卷, 第 3 期, pp.276-283, 2002.
 19. 吴福朝, 胡占义, 线性确定无穷远平面的单应矩阵和摄像机自标定, 自动化学报, 第 28 卷, 第 4 期, pp.488-496, 2002
 20. 吴福朝, 胡占义, 多平面多视点单应矩阵间的约束, 自动化学报, 第 28 卷, 第 5 期, pp.690-699,

- 2002.
21. 吴福朝, 胡占义, 由二次曲线确定摄像机方位的线性算法, 计算机学报, 第 25 卷, 第 11 期, pp.1157-1164, 2002.
 22. 吴福朝, 胡占义, PnP 问题的线性求解算法, 软件学报, 第 14 卷, 第 3 期, pp.682-688, 2003.
 23. 吴福朝, 王光辉, 胡占义, 基于矩形的确定摄像机内参数与位置的线性方法, 软件学报, 第 14 卷, 第 3 期, pp. 703-712, 2003.
 24. 吴福朝, 胡占义, 仿射重构的充要条件, 技术报告(未发表).
 25. 吴福朝, 胡占义, 度量重构的充要条件, 技术报告(未发表).
 26. F. C. Wu and Z. Y. Hu, 5-point and 4-point Algorithm to Determine the Fundamental Matrix, 自动化学报, 第 29 卷, 第 2 期, pp.175-180, 2003.
 27. 雷成, 吴福朝, 胡占义, Kruppa 方程与摄像机自定标, 自动化学报, 第 27 卷, 第 5 期, pp.621-630, 2001.
 28. 孙凤梅, 吴福朝, 胡占义, 由平行平面的投影确定无穷远平面的单应矩阵, 软件学报, 第 14 卷, 第 5 期, pp.936-946, 2003.
 29. 祝海江, 吴福朝, 基于一组对应消失线的度量重构, 软件学报, 第 15 卷, 第 5 期, pp.666-677, 2004.
 30. 张淮峰, 吴福朝, 胡占义, 基于仿射点对应的分层重构, 计算机学报, 第 28 卷, 第 8 期, pp.1267-1276, 2005
 31. 张淮峰, 吴福朝, 胡占义, 三视校正的理论及鲁棒性算法, 软件学报, 15(5):676-688, 2004.
 32. 胡占义, 王光辉, 吴福朝, 基于平面与直线的仿射重建, 计算机学报, 第 26 卷, 第 6 期, pp.722-728, 2003.
 33. F. C. Wu, Z. Y. Hu, and F. Q. Duan, 8-Point Algorithm Revisited: Factorized 8-Point Algorithm, International Conference on Computer Vision, ICCV'2005, Vol.I, pp.488-494, 2005.
 34. F. C. Wu, Z. Y. Hu, and H. J. Zhu, Camera Calibration with Moving One Dimensional Objects, Pattern Recognition, Vol.38, No.5, pp.355-265, 2005
 35. F. C. Wu, F. Q. Duan, and Z. Y. Hu, An Affine Invariant of Parallelograms and Its Application to Camera Calibration and 3D Reconstruction, The 9th European Conference on Computer Vision, LNCS 3952, pp.191-204, 2006
 36. F. C. Wu, Z. Y. Hu, The LLE and a linear mapping, Pattern Recognition, Vol.39, No.9, pp.1799-1084, 2006
 37. F. C. Wu, L. Wang, Z. Y. Hu, Manuscript Title: FOE Estimation: Can Image Measurement Errors Be Totally "Corrected" by the Geometric Method? Pattern Recognition, In press, 2006
 38. Z. Y. Hu and F. C. Wu, A Note on the Number of Solutions of the Non-Coplanar P4P problem, IEEE

- Transactions on Pattern Analysis and Machine Intelligence, Vol.24, No.4, pp.550-555, 2002.
39. Y. H. Wu, G. H. Wang, F. C. Wu, and Z. Y. Hu. Euclidean Reconstruction of a Circular Truncated Cone only from its Uncalibrated Contours. *Image and Vision Computing*, Vol. 24, Iss. 8, pp. 810-818, 2006.
 40. Y. H. Wu, X. Li, F. C. Wu, and Z. Y. Hu. Coplanar Circles, Quasi-Affine Invariance and Calibration. *Image and Vision Computing*, Vol. 24, Iss. 4, pp. 319-326, 2006.
 41. Z. Y. Hu, Y. H. Wu, F. C. Wu and S. D. Ma, The Number of Independent Kruppa Constraints from N Images, *Journal of Computer Science and Technology*, Vol.21, No.2, pp.209-217, 2006.
 42. Z. Y. Hu, F. C. Wu, and G. H. Wang, Impossibility of Affine Reconstruction by a Translating Camera from Two Perspective Images, *Pattern Recognition Letters*, Vol. 24, No. 16, pp. 2909-2911, 2003
 43. B. W. Zhang, Y. F. Li and F. C. Wu, Planar Pattern for Automatic Camera Calibration, *Optical Engineering*, Vol.42. No.6, pp.1542-1549, 2003.
 44. G.H. Wang, Z.Y.Hu, F. C. Wu, and H. T. Tsui, Projector-Camera Based System for Fast Object Modeling, In Proc. of IEEE International Workshop on Projector-Camera Systems (in conjunction with ICCV), 2003.
 45. G. H. Wang, Z. Y. Hu, F. C. Wu, and H. T. Tsui, Implementation and Experimental Study on Fast Object Modeling Based on Multiple Structured Stripes, *Optics and Lasers in Engineering*, Vol.42, pp.627-638, 2004.
 46. G. H. Wang, H. T. Tsui, and Z. Y. Hu, F. C. Wu, Camera Calibration and 3D Reconstruction from a Single View Based on Scene Constraints, *Image and Vision Computing*, Vol.23, No.2, pp.311-323, 2005.
 47. G. H. Wang, Z. Y. Hu, and F. C. Wu, Single View Based Measurement on Space Planes, *Journal of Computer Science and Technology*, Vol.19, No.3, pp.374-382, 2004.
 48. Y. H. Wu, H. J. Zhu, Z.Y. Hu, and F. C. Wu, Camera Calibration from the Quasi-Affine Invariance of Two Parallel Circles. *The 8th European Conference on Computer Vision*, pp. 190-202, 2004
 49. Fishler M. A, and Bolles R. C, Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, *Communications of the ACM*, Vol.24, No.6, pp.381-395, 1981
 50. H. C. Longuet-Higgins, A computer algorithm for reconstructing from two projections. *Nature*, 293:133-135, 1981.
 51. P. H. S. Torr. Bayesian model estimation and selection for epipolar geometry and generic manifold fitting. In *IJCV*, 50(1): 35-61, 2002.