

最小二乘估值均方差计算的矩阵体积法

薛树强¹ 党亚民^{1,3} 陈 武²

(1 中国测绘科学研究院,北京市海淀区北太平路 16 号,100039)
(2 香港理工大学土地测量与地理资讯学系,香港九龙红磡)
(3 山东科技大学测绘科学与工程学院,青岛市经济技术开发区前湾港路 579 号,266510)

摘 要:研究了最小二乘估值均方差计算的矩阵体积法,该方法无需计算最小二乘估值,其数值计算的稳定性较好,可在最小二乘解算前对系统的观测结构、函数模型的准确性和观测数据质量进行评价。
关键词:最小二乘平差;均方差;数值稳定性;矩阵体积;马氏体积
中图法分类号:P207.2

最小二乘估值均方差计算在测量中有着重要应用,它综合反映了系统的观测结构、函数模型的准确性、观测数据质量等^[1-4],是最小二乘精度评定的重要内容。在理论上,最小二乘估值均方差计算与最小二乘解算过程无关;在某些实际问题中,需要在最小二乘解算前对其最小二乘估值的精度进行评定,以达到事前对系统观测结构、函数模型的准确性和观测数据质量进行评价的目的。特别地,对于病态问题,最小二乘估值均方差的经典计算公式的数值计算的稳定性严重受制于最小二乘估值数值计算的稳定性^[3],其计算结果变得很不可靠。本文基于 Hilbert 空间几何学和矩阵体积的概念,提出了一种最小二乘估值均方差计算的矩阵体积法。

1 线性最小二乘问题

记 $R(\cdot)$ 为矩阵的值域空间^[5], $D(\cdot)$ 为矩阵列向量构成的超平行体^[6], $\det(\cdot)$ 为方阵行列式^[5], $\|\cdot\|$ 为矩阵的范数^[5-7], $\|\cdot\|_F$ 为矩阵的 Frobenius 范数^[5], $\|\cdot\|_n$ 为矩阵的 Mahalano-bis 范数^[5]; $\|\cdot\|_2$ 为向量的欧氏范数^[5-7]。由 Hilbert 空间理论可知^[7],对于线性方程组系数矩阵 $A \in R^{m \times n}$ 和向量 $L \in R^{m \times 1}$,惟一存在 $\hat{L} \in R(A)$,使得:

$$\|L - \hat{L}\|_2 = \inf_{y \in R(A)} \|L - y\|_2 \quad (1)$$

式中, \hat{L} 为 L 在值域空间 $R(A)$ 中的最小二乘逼近元。由投影定理可知:

$$L = \hat{L} + V \quad (2)$$

其中, $V \in R(A)^\perp$, 且该正交分解惟一。 L 到值域空间 $R(A)$ 的投影算子 J_A 为^[1,5,7]:

$$J_A = A(A^T P A)^{-1} A^T P \quad (3)$$

其中, 正定矩阵 P 为空间具有概率意义的测度^[1,8]。对应最小二乘估计,由式(3)可知:

$$\hat{L} = A(A^T P A)^{-1} A^T P L \quad (4)$$

$$V = (I - A(A^T P A)^{-1} A^T P) L \quad (5)$$

则超定方程组 $Ax = L$ 的最小二乘解为:

$$\hat{x} = (A^T P A)^{-1} A^T P L \quad (6)$$

最小二乘单位权中误差的计算公式为^[1]:

$$\hat{\sigma}_0 = \sqrt{V^T P V / (m - n)} \quad (7)$$

最小二乘解 \hat{x} 的均方差计算公式为:

$$\hat{\sigma}_{\hat{x}} = \sqrt{\text{tr}(\hat{Q}_{\hat{x}})} \hat{\sigma}_0 \quad (8)$$

其中, $\hat{Q}_{\hat{x}} = (A^T P A)^{-1}$ 。

2 矩阵体积和广义伴随矩阵

2.1 矩阵体积的定义

对于矩阵 $A \in R^{m \times n}$, 其 p 维体积 $\text{vol}_p(A)$ 为:
$$\text{vol}_p(A) = \|C_p(A)\| \quad (9)$$

其中, $C_p(A)$ 为矩阵 A 的 p 阶复合矩阵^[4,5,9]。若

定义采用 Frobenius 范数,则称其为矩阵 \mathbf{A} 的 p 维欧氏体积^[10,11],记为 $\text{vol}_p(\mathbf{A})_F$;若定义采用 Mahalanobis 范数,称其为矩阵 \mathbf{A} 的 p 维马氏体积,记为 $\text{vol}_p(\mathbf{A})_\Omega$ 。

由 Binet-Cauchy 定理可得,若 $\mathbf{A} \in R^{m \times n}, \mathbf{B} \in R^{n \times s}$,则:

$$\mathbf{C}_p(\mathbf{AB}) = \mathbf{C}_p(\mathbf{A})\mathbf{C}_p(\mathbf{B}) \tag{10}$$

其中, $p \leq \min\{m, n\}$ 。当 $s = m, p = n$ 时, Binet-Cauchy 推广定理退化为 Binet-Cauchy 定理,即

$$\det(\mathbf{BA}) = C_n(\mathbf{B}) \cdot C_n(\mathbf{A}) \tag{11}$$

对于矩阵 $\mathbf{A} \in R^{m \times n}$, 矩阵的 n 维马氏体积满足:

$$\text{vol}_n(\mathbf{A})_\Omega = \sqrt{\det(\mathbf{A}^T \mathbf{PA})} \tag{12}$$

其中, $\Omega = C_n(\mathbf{P})$ 。式(12)可由式(10)和式(11)导出。当 $\mathbf{P} = \mathbf{I}$ 时, 矩阵的马氏体积退化为矩阵的欧氏体积, 其几何意义为超平行体 $D(\mathbf{A})$ 的体积^[10,12,13]。

2.2 矩阵欧氏体积的几何意义

矩阵的欧氏体积 $\text{vol}_p(\mathbf{A})_F$ 的几何意义为超平行体 $D(\mathbf{A})$ ^[10] 大小的度量。设矩阵 $\mathbf{A} \in R^{m \times n}$, 其 n 维欧氏体积为:

$$V_n(D(\mathbf{A})) = \text{vol}_n(\mathbf{A})_F = \sqrt{\det(\mathbf{A}^T \mathbf{A})}$$

k 维空间内的欧氏体积(与平行体 $D(\mathbf{A})$ 的“表面积”相关)满足:

$$V_k(D(\mathbf{A})) = \text{vol}_k(\mathbf{A})_F = 2^{n-k} \|D_k(\mathbf{A})\|_F$$

2.3 任意矩阵的伴随矩阵

对于矩阵 $\mathbf{A} \in R^{m \times n}$, 其 p 阶伴随矩阵定义为:

$$\mathbf{C}_p^*(\mathbf{A}) = \begin{cases} \mathbf{J}_{n1} \mathbf{C}_p(\mathbf{A}^T) \mathbf{J}_{m1}, C_n^p \text{ 为奇数}, C_m^p \text{ 为奇数} \\ \mathbf{J}_{n1} \mathbf{C}_p(\mathbf{A}^T) \mathbf{J}_{m2}, C_n^p \text{ 为奇数}, C_m^p \text{ 为偶数} \\ \mathbf{J}_{n1} \mathbf{C}_p(\mathbf{A}^T) \mathbf{J}_{m1}, C_n^p \text{ 为偶数}, C_m^p \text{ 为奇数} \\ \mathbf{J}_{n1} \mathbf{C}_p(\mathbf{A}^T) \mathbf{J}_{m2}, C_n^p \text{ 为偶数}, C_m^p \text{ 为偶数} \end{cases}$$

或

$$\mathbf{C}_p^*(\mathbf{A}) = \begin{cases} \mathbf{J}_{n2} \mathbf{C}_p(\mathbf{A}^T) \mathbf{J}_{m2}, C_n^p \text{ 为奇数}, C_m^p \text{ 为奇数} \\ \mathbf{J}_{n2} \mathbf{C}_p(\mathbf{A}^T) \mathbf{J}_{m1}, C_n^p \text{ 为奇数}, C_m^p \text{ 为偶数} \\ \mathbf{J}_{n2} \mathbf{C}_p(\mathbf{A}^T) \mathbf{J}_{m2}, C_n^p \text{ 为偶数}, C_m^p \text{ 为奇数} \\ \mathbf{J}_{n2} \mathbf{C}_p(\mathbf{A}^T) \mathbf{J}_{m1}, C_n^p \text{ 为偶数}, C_m^p \text{ 为偶数} \end{cases}$$

其中,

$$\mathbf{J}_{n1} = \begin{bmatrix} & & & 1 \\ & & -1 & \\ & \ddots & & \\ (-1)^{C_n^p-1} & & & \end{bmatrix} \in R^{C_n^p \times C_n^p}$$
$$\mathbf{J}_{n2} = \begin{bmatrix} & & & -1 \\ & & 1 & \\ & \ddots & & \\ (-1)^{C_n^p} & & & \end{bmatrix} \in R^{C_n^p \times C_n^p}$$

$$\mathbf{J}_{m1} = \begin{bmatrix} & & & 1 \\ & & -1 & \\ & \ddots & & \\ (-1)^{C_m^p-1} & & & \end{bmatrix} \in R^{C_m^p \times C_m^p}$$
$$\mathbf{J}_{m2} = \begin{bmatrix} & & & -1 \\ & & 1 & \\ & \ddots & & \\ (-1)^{C_m^p} & & & \end{bmatrix} \in R^{C_m^p \times C_m^p}$$

考虑到 Frobenius 范数和 Mahalanobis 范数的正交不变性:

$$\text{vol}_p(\mathbf{A})_F = \|\mathbf{C}_p^*(\mathbf{A})\|_F \tag{13}$$

$$\text{vol}_p(\mathbf{A})_\Omega = \|\mathbf{C}_p^*(\mathbf{A})\|_\Omega \tag{14}$$

由 Binet-Cauchy 定理可得:

$$\mathbf{C}_p^*(\mathbf{AB}) = \mathbf{C}_p^*(\mathbf{B})\mathbf{C}_p^*(\mathbf{A}) \tag{15}$$

若 $m = n$, 可以证明^[14]:

$$\mathbf{C}_{n-1}^*(\mathbf{A}) = \mathbf{A}^* \tag{16}$$

其中, \mathbf{A}^* 为矩阵 \mathbf{A} 的古典伴随矩阵。若 $\text{rank}(\mathbf{A}) = n$, 则 $\mathbf{A}^{-1} = \mathbf{A}^* / \det(\mathbf{A})$ 。

3 最小二乘解均方差估计的矩阵体积法

3.1 单位权中误差的矩阵体积表示

若 $\mathbf{A} \in R^{m \times n}$, 且 $m > n$, 记 $\mathbf{B} = [\mathbf{A} \ \mathbf{b}]$ 为系数矩阵 \mathbf{A} 的增广矩阵, 则:

$$\hat{\sigma}_0 = \frac{1}{\sqrt{m-n}} \frac{\text{vol}_{n+1}(\mathbf{B})_{\Omega_1}}{\text{vol}_n(\mathbf{A})_\Omega} \tag{17}$$

其中, $\Omega_1 = C_{n+1}(\mathbf{P})$ 。

对式(17)简要证明如下。由题设可得:

$$\mathbf{B}^T \mathbf{PB} = \begin{bmatrix} \mathbf{A}^T \mathbf{P} \\ \mathbf{b}^T \mathbf{P} \end{bmatrix} [\mathbf{A} \ \mathbf{b}] = \begin{bmatrix} \mathbf{A}^T \mathbf{PA} & \mathbf{A}^T \mathbf{Pb} \\ \mathbf{b}^T \mathbf{PA} & \mathbf{b}^T \mathbf{Pb} \end{bmatrix} \tag{18}$$

由于 $\text{rank}(\mathbf{A}) = n$, 在式(18)两边同乘以矩阵

$$\mathbf{C} = \begin{bmatrix} \mathbf{I}_{n \times n} & 0 \\ -\mathbf{b}^T \mathbf{PA} (\mathbf{A}^T \mathbf{PA})^{-1} & 1 \end{bmatrix}, \text{得:}$$

$$\mathbf{CB}^T \mathbf{PB} = \begin{bmatrix} \mathbf{A}^T \mathbf{PA} & \mathbf{A}^T \mathbf{Pb} \\ 0 & \mathbf{b}^T \mathbf{Pb} - \mathbf{b}^T \mathbf{PA} (\mathbf{A}^T \mathbf{PA})^{-1} \mathbf{A}^T \mathbf{Pb} \end{bmatrix}$$
$$\det(\mathbf{C}) = 1 \tag{19}$$

由式(19)可得:

$$\det(\mathbf{CB}^T \mathbf{PB}) = \det(\mathbf{A}^T \mathbf{PA}) (\mathbf{b}^T (\mathbf{P} - \mathbf{PA} (\mathbf{A}^T \mathbf{PA})^{-1} \mathbf{A}^T \mathbf{P}) \mathbf{b}) \tag{20}$$

由式(20)可得:

$$\frac{\det(\mathbf{B}^T \mathbf{PB})}{\det(\mathbf{A}^T \mathbf{PA})} = \mathbf{b}^T (\mathbf{P} - \mathbf{PA} (\mathbf{A}^T \mathbf{PA})^{-1} \mathbf{A}^T \mathbf{P}) \mathbf{b} =$$

$$\mathbf{V}^T \mathbf{P} \mathbf{V}$$

(21)

由式(7)、式(19)和式(21)即可得到式(17)。

若 $\mathbf{P}=\mathbf{I}$, 则 $\frac{\text{vol}_{n+1}(\mathbf{B})_{\Omega_1}}{\text{vol}_n(\mathbf{A})_{\Omega}}$ 的几何意义为超平行体 $D(\mathbf{B})$ 到底面 $D(\mathbf{A})$ 上的高, 即向量 \mathbf{b} 的终点到底面 $D(\mathbf{A})$ 的欧氏距离。

3.2 最小二乘解均方差的矩阵体积表示

由式(12)至式(14)可得权逆阵 \mathbf{Q}_x 的广义伴随矩阵为:

$$\mathbf{Q}_x = (\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} \frac{C_{n-1}^*(\mathbf{A})^T C_{n-1}^*(\mathbf{P}) C_{n-1}^*(\mathbf{A})}{\text{vol}_n^2(\mathbf{A})_{\Omega}}$$

(22)

结合式(10)可得:

$$\hat{\sigma}_x = \frac{\|C_{n-1}^*(\mathbf{A})\|_{\Omega_2}}{\text{vol}_n(\mathbf{A})_{\Omega}} \hat{\sigma}_0$$

(23)

其中, $\Omega_2=C_{n-1}^*(\mathbf{P})$ 。由式(14)和式(23)可得:

$$\hat{\sigma}_x = \frac{\text{vol}_{n-1}(\mathbf{A})_{\Omega_2}}{\text{vol}_n(\mathbf{A})_{\Omega}} \hat{\sigma}_0$$

(24)

将式(22)代入式(24)可得:

$$\hat{\sigma}_x = \frac{\text{vol}_{n-1}(\mathbf{A})_{\Omega_2} \text{vol}_{n+1}(\mathbf{B})_{\Omega_1}}{\sqrt{m-n} \text{vol}_n^2(\mathbf{A})_{\Omega}}$$

(25)

式中, $\text{vol}_{n-1}(\mathbf{A})_{\Omega_2}/\text{vol}_n(\mathbf{A})_{\Omega}$ 只与系数矩阵和权阵有关, 反映了系统的观测结构, 在导航定位领域, 它与定位图形的 GDOP 值相关^[15], 其几何意义为超平行体 $D(\mathbf{A})$ 在 $n-1$ 维空间中的马氏体积(与其在 $n-1$ 维空间中的“表面积”相关)与其在 n 维空间中的马氏体积的比值。

4 应用及其数值算例

算例 1 利用文献[16]中的 Hilbert 矩阵(为病态矩阵)构造系数矩阵 \mathbf{A} 为:

$$\mathbf{A} = \begin{bmatrix} \text{Hilb}(3) \\ \frac{87}{157} \text{Hilb}(3) \end{bmatrix}$$

其中, Hilb(3)表示 3 阶 Hilbert 矩阵, 并取 4 位有效数字, 则:

$$\mathbf{A} = \begin{bmatrix} 1.000\ 0 & 0.500\ 0 & 0.333\ 3 \\ 0.500\ 0 & 0.333\ 3 & 0.250\ 0 \\ 0.333\ 3 & 0.250\ 0 & 0.200\ 0 \\ 0.554\ 1 & 0.277\ 1 & 0.184\ 7 \\ 0.277\ 1 & 0.184\ 7 & 0.138\ 5 \\ 0.184\ 7 & 0.138\ 5 & 0.110\ 8 \end{bmatrix}$$

解向量 \mathbf{x} 的真值取为 $\mathbf{x}=[1\ 1\ 1]^T$, 则:
 $\tilde{\mathbf{b}}=\mathbf{A}\mathbf{x}=[1.833\ 3, 1.083\ 3, 0.783\ 3, 1.015\ 9, 0.600\ 3, 0.434]^T$

对 $\tilde{\mathbf{b}}$ 施加扰动:

$$\Delta \mathbf{b} = [8.6, 3.4, 6.8, 0.5, 3.6, 5]^T \cdot 10^{-3}$$

构造观测向量:

$$\mathbf{b} = \tilde{\mathbf{b}} + \Delta \mathbf{b} = [1.841\ 9, 1.086\ 7, 0.790\ 1, 1.016\ 4, 0.603\ 9, 0.439]^T$$

式(9)和式(25)的计算结果分别记为 σ_x 和 σ'_x , 则:

$$\Delta \sigma_x = |\hat{\sigma}_x - \hat{\sigma}'_x| < 10^{-10}$$

算例 1 表明, 当不存在数值计算稳定性问题或数值稳定性良好时, 式(8)和式(25)的解算结果具有一致性。

算例 2 类似于算例 1, 分别使用 4 阶、5 阶、6 阶 Hilbert 矩阵构造病态方程组, 采用双精度和单精度对病态方程组进行最小二乘解算, 并对式(8)、式(25)进行计算, 其结果如表 1 所示。

算例 2 表明, 式(25)的数值稳定性较好, 对于严重病态问题, 其效果更明显, 而式(8)的计算结果失真较为严重。算例分析表明, 均方误差数值计算稳定性的改善在很大程度上取决于单位权中误差因子数值计算稳定性的提高, 相比于经典计算公式, 其计算量较小, 且避免了矩阵求逆, 符合数值计算稳定性原则^[17]。此外, 本算例直接利用矩阵体积的定义进行计算, 其计算量较大, 矩阵体积计算方法的改进有望进一步提高其数值稳定性。

表 1 双精度和单精度解算结果对比

Tab. 1 Comparison of the Results Computed with Double Precision and Single Precision

	列维数	双精度解算	单精度解算	单双精度结果差
3	式(8)	0.784 7	0.824 9	0.040 2
	式(25)	0.784 7	0.783 3	0.001 4
	结果差	10 ⁻¹⁰ 级	0.041 6	
4	式(8)	33.102 8	1 540.6	1 507.5
	式(25)	33.102 8	29.073 4	4.029 4
	结果差	10 ⁻⁸ 级	1 511.5	
5	式(8)	67.489 1	2 530.3	2 462.9
	式(25)	67.489 1	17.892 0	49.597 1
	结果差	10 ⁻⁷ 级	2 512.4	
6	式(8)	86.548 9	2 700.1	2 613.6
	式(25)	86.548 9	0.978 1	85.570 8
	结果差	10 ⁻⁶ 级	2 699.1	

5 结 语

通过本文的研究, 可得: ① 最小二乘解均方差的矩阵体积表示法形式简单, 几何意义明确; ② 最小二乘解算前, 可以对其解的精度进行评定; ③ 最小二乘单位权中误差方差计算的矩阵体积法可避免矩阵求逆, 其数值稳定性较好。

本文研究结果有望将最小二乘平差步骤简化为两步:① 对系统观测结构、函数模型和观测数据的质量进行评定;② 根据评定结果,对算法、函数模型或观测数据的权重进行调整后再平差。

参 考 文 献

[1] 黄维彬. 近代平差理论及其应用[M]. 北京:解放军出版社, 1991: 4-5, 56-57, 45-87

[2] 杨元喜. 自适应抗差最小二乘估计[J]. 测绘学报, 1996,25(3):206-211

[3] 王振杰. 大地测量中不适定问题的正则化解法研究[D]. 武汉:中国科学院测量与地球物理研究所, 2003

[4] 陶本藻,姚宜斌. 可靠性分析与数据探测[J]. 武汉大学学报·信息科学版, 2002, 27(6):607-610

[5] 张贤达. 矩阵分析与应用[M]. 北京:清华大学出版社, 2004: 44-58, 146-147, 344-345, 637-639

[6] 张贤科,许甫华. 高等代数学[M]. 北京:清华大学出版社, 1998: 40-47, 253-254

[7] 黄振友,杨建新. 泛函分析[M]. 北京:科学出版社,2003: 23-30

[8] Fremlin D H. Measure Theory[M]. New York: Cambridge University Press, 1974

[9] Atzema E J. Beyond Monge’s Theorem; A Gener-

alization of the Pythagorean Theorem[J]. Mathematics Magazine,2000,73(4): 293-294

[10] Ben-Israel A, Greville T N E. Generalized Inverses: Theory and Applications[M]. 2nd ed. New York: Springer-Verlag, 2003: 28-35

[11] 薛树强. 矩阵体积及其在网形设计中的应用[D]. 北京:中国测绘科学研究院, 2007

[12] 王孝青,党亚民,薛树强. 一种病态问题诊断的数值指标——矩阵向量正交度[J]. 武汉大学学报·信息科学版,2009,34(2):244-247

[13] 卢秀山,冯遵德. 观测结构的度量[J]. 武汉大学学报·信息科学版, 2008, 33(8):831-833

[14] 杜忠复,徐玉卿. 复合矩阵与伴随矩阵的关系及应用[J]. 工科数学, 1999(3): 156-158

[15] 郑作亚,黄诚,冯初刚,等. 4 颗卫星情况的几何优化法修正[J]. 天文学报, 2003,44(3):9-11

[16] 路浩. Vandermonde 方程 Hilbert 方程及 Vandermonde 矩阵 Hilbert 矩阵逆的快速与并行算法[J]. 计算数学, 1993,15(4):410-419

[17] 李庆扬,王能超,易大义. 数值分析[M]. 4 版. 武汉:华中科技大学出版社, 2006: 7-18

第一作者简介:薛树强,硕士,主要从事大地测量研究。
E-mail:xuesq@casm. ac. cn

Matrix Volume Method of the Mean Square Deviation
Computation of Least Square Solution

XUE Shuqiang¹ DANG Yamin^{1,3} CHEN Wu²

(1 Chinese Academy of Surveying and Mapping,16 Beitaping Road, Haidian District, Beijing 100039, China)

(2 Department of Land Survey and Geo-informatics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China)

(3 School of Survey Science and Engineering, Shandong University of Science and Technology,579 Qianwangang Road, Economic &. Technical Development Zone, Qingdao 266510, China)

Abstract: A simple and stable mean square deviation computation method without solving the linear equations is presented. The observing structure, the quality of observations and the accuracy of the function model can be evaluated before applying least square adjustment by using the proposed method in this method.

Key words: least square adjustment; mean square deviation; numerical stability; matrix volume; Mahalanobis volume

About the first author: XUE Shuqiang, master, majors in geodesy.
E-mail: xuesq@casm. ac. cn