

解等式约束加权线性最小二乘问题的矩阵校正方法

赵金熙.

(南京大学数学系, 210093)

MATRIX UPDATING METHOD FOR THE CONSTRAINED AND WEIGHTED LINEAR LEAST SQUARES PROBLEM

Zhao Jinxi

(Department of Mathematics, Nanjing University, 210093)

Abstract

In this paper, a direct method based on the generalized Cholesky factorization is given for the solution of the constrained and weighted linear least squares problem. Updating techniques of the system matrix are also considered.

1 引言

在实际应用中常会提出解等式约束加权线性最小二乘问题

$$\begin{aligned} & \min_{x \in R^n} (b_2 - A_2 x)^T W (b_2 - A_2 x) \\ & \text{s. t. } A_1 x = b_1, \end{aligned} \quad (1)$$

其中 $A_1 \in R^{p \times n}$, $A_2 \in R^{q \times n}$, $b_1 \in R^p$, $b_2 \in R^q$, $W \in R^{q \times q}$ 为对称正定矩阵.

对于问题(1), 目前已有多种数值求解方法, 如 Paige 利用(1)的对偶公式给出了一个向后稳定的数值方法^[5]. Gulliksson 和 Wedin^[4]利用加权 QR 分解技巧给出了解(1)的一个直接解法. 作者利用广义 Cholesky 分解构造了解(1)的矩阵分解方法^[9].

• 收稿日期: 1994-04-18.

该课题得到国家教委博士点基金、国家教委回国人员科研基金、江苏省自然科学基金资助.

另一方面,在信号处理等领域,常需对(1)中约束线性方程组

$$A_1 x = b_1$$

的数据进行增加(updating)或删除(downdating).本文利用矩阵分解方法对此问题作了详细的讨论.

2 系统方程组及矩阵分解方法

由于 $W \in R^{q \times q}$ 为对称正定矩阵,则稍加推导可以证明^[6]问题(1)等价于解线性方程组

$$\begin{bmatrix} 0 & 0 & A_1 \\ 0 & W^{-1} & A_2 \\ A_1^T & A_2^T & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ x \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix}, \quad (2)$$

其中 λ_1 是 Lagrange 乘子向量, $W^{-1}\lambda_2$ 是残量. 容易看出问题(1)有唯一解的充要条件是

$$\text{rank}(A_1) = p,$$

和

$$\text{rank} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = n.$$

在下面的讨论中,我们总假定这两个条件是满足的. 由于问题(1)等价于

$$\begin{aligned} \min_{x \in R^n} & (x^T A_2^T W A_2 x - 2b_2^T W A_2 x + b_2^T W b_2) \\ \text{s. t. } & A_1 x = b_1, \end{aligned} \quad (3)$$

故利用 Kuhn-Tucker 条件,即要确定 $x \in R^n, \lambda \in R^p$ 满足

$$\begin{cases} A_2^T W A_2 x - A_2^T W b_2 = A_1^T \lambda \\ A_1 x = b_1. \end{cases} \quad (4)$$

写成矩阵的形式,即为

$$\begin{bmatrix} A_2^T W A_2 & -A_1^T \\ A_1 & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} A_2^T W b_2 \\ b_1 \end{bmatrix}. \quad (5)$$

我们称(5)式为问题(1)的系统方程组. 对应的系数矩阵称为问题(1)的系统矩阵.

容易证明系统矩阵

$$M_1 = \begin{bmatrix} A_2^T W A_2 & -A_1^T \\ A_1 & 0 \end{bmatrix}. \quad (6)$$

的对称部分是半正定的,故常称为广义半正定矩阵. 对于系统矩阵 M_1 ,我们可以证明下列更一般的定理.

定理 1 设有矩阵

$$M = \begin{bmatrix} A_2^T W A_2 & -A_1^T \\ A_1 & C \end{bmatrix},$$

其中 $A_1 \in R^{p \times n}, A_2 \in R^{q \times n}, \text{rank}(A_1) = p, \text{rank}(A_2) = n, W \in R^{q \times q}$ 为对称正定矩阵, $C \in R^{p \times p}$ 为对称半正定矩阵,则矩阵 M 一定存在分解式

其中 $L_1 \in R^{n \times n}$ 为非奇异下三角阵, $L_2 \in R^{p \times p}$ 为非奇异下三角阵, $G_1 \in R^{p \times n}$.

由定理 1 知系统矩阵 M_1 满足定理的要求, 即 M_1 是定理 1 中矩阵 M 取 $C=0$ 的特殊情况. 因此 M_1 一定存在广义 Cholesky 分解式

$$M_1 = \begin{bmatrix} L_w & 0 \\ G & L_c \end{bmatrix} \begin{bmatrix} L_w^T & -G^T \\ 0 & L_c^T \end{bmatrix}, \quad (7)$$

其中 $L_w \in R^{n \times n}$ 为下三角阵, $L_c \in R^{p \times p}$ 为下三角阵, $G \in R^{p \times n}$.

若我们已经得到 M_1 的分解式(7), 则系统方程组(5)就可以通过解下三角方程组

$$\begin{bmatrix} L_w & 0 \\ G & L_c \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} A_2^T W b_2 \\ b_1 \end{bmatrix} \quad (8)$$

及上三角方程组

$$\begin{bmatrix} L_w^T & -G^T \\ 0 & L_c^T \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \quad (9)$$

方便地得到解向量 $\begin{bmatrix} x \\ \lambda \end{bmatrix}$, 从而得到问题(1)的解 x .

3 系统矩阵的分解

值得注意的是在矩阵 M_1 的分解式(7)中, 我们只要得到下三角矩阵 L_w 、 L_c 及矩阵 G 后, 分解式(7)就完全确定了. 由于

$$L_w L_w^T = A_2^T W A_2.$$

故 L_w 就是对称正定矩阵 $A_2^T W A_2$ 的 Cholesky 因子, 也就是对 $A_2^T W A_2$ 进行 Cholesky 分解即可得到下三角矩阵 L_w . 事实上, 我们也可首先对 W 作 Cholesky 分解, 即

$$W = L_2 L_2^T,$$

然后再对列满秩矩阵 $L_2^T A_2$ 进行 QR 分解, 即

$$L_2^T A_2 = Q_q \begin{bmatrix} R \\ 0 \end{bmatrix}, \quad (10)$$

其中 Q_q 为 $q \times q$ 直交矩阵, $R \in R^{n \times n}$ 为主对角元素均大于 0 的上三角矩阵, 这样

$$\begin{aligned} R^T R &= (L_2^T A_2)^T (L_2^T A_2) \\ &= A_2^T W A_2, \end{aligned}$$

取

$$L_w = R^T. \quad (11)$$

这样就得到了下三角矩阵

$$L_w = [l_{ij}] \in R^{n \times n}.$$

若令矩阵

$$A_1 = [a_{ij}] \in R^{p \times n},$$

则

$$G = [g_{ij}] = A_1 (L_w^T)^{-1},$$

其元素可用递推式得到

$$g_{ij} = (a_{ij} - \sum_{k=1}^{j-1} g_{ik}l_{jk})/l_{jj}, \quad i = 1:p, j = 1:n. \quad (12)$$

由于 G 为行满秩矩阵,故 GG^T 为正定阵,这时总存在下三角矩阵

$$L_c = [l_{ij}] \in R^{p \times p}$$

使得

$$GG^T = L_c L_c^T.$$

L_c 的元素可由下式得到,

$$t_{ij} = (\sum_{k=1}^n g_{ik}g_{jk} - \sum_{k=1}^{j-1} t_{ik}t_{jk})/t_{jj}, \quad i = 1:p, j = 1:p. \quad (13)$$

值得注意的是在表达式(12)、(13)中,求和号 $\sum_{k=1}^{j-1}$ 中若 $j=1$ 时则该和式取为 0. 下面的讨论该约定仍有效.

综上所述我们就得到了矩阵

$$L_w = [l_{ij}], \quad L_c = [t_{ij}], \quad G = [g_{ij}].$$

也就完成了系统矩阵 M_1 的矩阵分解式(7)的计算.

4 约束矩阵的 updating

在信号处理等领域,常需对约束矩阵构成的方程组

$$A_1 x = b_1$$

的数据进行增删. 这里首先考虑 $[A_1; b_1]$ 增加 k 行的情况. 这就是 updating 问题.

令

$$\bar{A}_1 = \begin{bmatrix} A_1 \\ Z \end{bmatrix}, \quad Z = [z_{ij}] \in R^{k \times n}, \quad (14)$$

这时当然也有

$$\bar{b}_1 = \begin{bmatrix} b_1 \\ s \end{bmatrix}, \quad s \in R^k.$$

这里同样要求

$$\text{rank}k(\bar{A}_1) = p + k < n.$$

也就是说我们若已经得到了 M_1 的分解式(7),在增加 k 组观测数据后,要利用已经得到的信息计算

$$\bar{M}_1 = \begin{bmatrix} A_2^T W A_2 & -\bar{A}_1^T \\ \bar{A}_1 & 0 \end{bmatrix} \quad (15)$$

的矩阵分解式. 令

$$\bar{G} = \bar{A}_1 (L_w^T)^{-1} = \begin{bmatrix} A_1 (L_w^T)^{-1} \\ Z (L_w^T)^{-1} \end{bmatrix} = \begin{bmatrix} G \\ H \end{bmatrix} \quad (16)$$

即

$$H = Z(L_w^T)^{-1} = [h_{ij}] \in R^{k \times n}.$$

这时矩阵 H 的元素可由下式得到:

$$h_{ij} = (z_{ij} - \sum_{\mu=1}^{j-1} h_{i\mu} l_{\mu j}) / l_{jj} \quad i = 1:k, j = 1:n.$$

由于 \bar{A}_1 为行满秩矩阵, 故 $\bar{G}\bar{G}^T$ 为正定矩阵. 下面考虑 $\bar{G}\bar{G}^T$ 的 Cholesky 分解.

$$\bar{G}\bar{G}^T = \begin{bmatrix} G \\ H \end{bmatrix} [G^T, H^T] = \begin{bmatrix} GG^T & GH^T \\ HG^T & HH^T \end{bmatrix}. \quad (17)$$

在这里 GG^T 已有 Cholesky 因子 L_c , 即

$$L_c L_c^T = GG^T.$$

而 HH^T 为对称正定矩阵. 设 $\bar{G}\bar{G}^T$ 有 Cholesky 分解

$$\bar{G}\bar{G}^T = \begin{bmatrix} L_c & 0 \\ F & L_e \end{bmatrix} \begin{bmatrix} L_c^T & F^T \\ 0 & L_e^T \end{bmatrix}. \quad (18)$$

这样取

$$F = HG^T (L_c^T)^{-1}.$$

令

$$F = [f_{ij}], \quad HG^T = [s_{ij}].$$

则可得到 F 元素的递推关系式:

$$f_{ij} = (s_{ij} - \sum_{\mu=1}^{j-1} f_{i\mu} l_{\mu j}) / l_{jj} \quad (19)$$

$$i = 1:k, \quad j = 1:n.$$

又由(18)式得

$$HH^T = FF^T + L_e L_e^T,$$

设下三角矩阵 $L_e = [l_{ij}^*] \in R^{k \times k}$, 且

$$HH^T = [h_{ij}^*] \in R^{k \times k}.$$

则

$$l_{ij}^* = (h_{ij}^* - \sum_{\mu=1}^i f_{i\mu} f_{\mu j} - \sum_{\mu=1}^{j-1} l_{i\mu}^* l_{\mu j}^*) / l_{jj}^* \quad (20)$$

$$i = 1:k, \quad j = 1:i.$$

因此我们有

$$\bar{G}\bar{G}^T = L_c L_c^T,$$

其中

$$L_c = \begin{bmatrix} L_c & 0 \\ F & L_e \end{bmatrix}.$$

这样最终得到 M_1 的 updating 矩阵 \bar{M}_1 的分解式

$$\bar{M}_1 = \begin{bmatrix} A_2^T W A_2 & -\bar{A}_1^T \\ \bar{A}_1 & 0 \end{bmatrix} \quad (21)$$

$$= \begin{bmatrix} L_w & 0 \\ \bar{G} & L_c \end{bmatrix} \begin{bmatrix} L_w^T & -\bar{G}^T \\ 0 & L_c^T \end{bmatrix}.$$

5 约束矩阵的 downdating

如果在约束矩阵 A_1 中删去 k 行(即删去 k 个约束方程). 不失一般性,我们假设删去约束方程组

$$A_1 x = b_1$$

的后 k 个方程,即令

$$A_1 = \begin{bmatrix} \tilde{A}_1 \\ Z \end{bmatrix}, \quad \tilde{A}_1 \in R^{(p-k) \times n}, \quad Z \in R^{k \times n}.$$

也就是要讨论在得到矩阵分解式(7)后如何得到其 downdating 问题

$$\tilde{M}_2 = \begin{bmatrix} A_2^T W A_2 & -\tilde{A}_1^T \\ \tilde{A}_1 & 0 \end{bmatrix}$$

的矩阵分解式. 令

$$G = A_1(L_w^T)^{-1} = \begin{bmatrix} \tilde{G} \\ H \end{bmatrix}, \quad \tilde{G} \in R^{(p-k) \times n}, \quad H \in R^{k \times n}.$$

这时我们有

$$\tilde{G} = \tilde{A}_1(L_w^T)^{-1}, \quad H = Z(L_w^T)^{-1}.$$

由于

$$\begin{aligned} GG^T &= \begin{bmatrix} \tilde{G} \\ H \end{bmatrix} [\tilde{G}^T \ H^T] = \begin{bmatrix} \tilde{G}\tilde{G}^T & \tilde{G}H^T \\ H\tilde{G}^T & HH^T \end{bmatrix} \\ &= L_c L_c^T. \end{aligned}$$

令下三角矩阵 $L_c \in R^{p \times p}$ 有形式

$$L_c = \begin{bmatrix} L_\beta & 0 \\ G_\alpha & L_\gamma \end{bmatrix},$$

其中 $L_\beta \in R^{(p-k) \times (p-k)}$ 为下三角阵. 这样就有

$$\tilde{G}\tilde{G}^T = L_\beta L_\beta^T,$$

这就是说 L_β 即为 $\tilde{G}\tilde{G}^T$ 的 Cholesky 因子,故取

$$L_i = L_\beta.$$

这样就有

$$\begin{aligned} \tilde{M}_2 &= \begin{bmatrix} A_2^T W A_2 & -\tilde{A}_1 \\ \tilde{A}_1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} L_w & 0 \\ \tilde{G} & L_i \end{bmatrix} \begin{bmatrix} L_w^T & -\tilde{G}^T \\ 0 & L_i^T \end{bmatrix}. \end{aligned} \tag{22}$$

这就是说,当约束方程组删去 k 个方程时,在系统矩阵的分解式中, $\tilde{G} \in R^{(p-k) \times n}$ 即为分解式(7)中矩阵 G 的前 $(p-k)$ 行元素, L_i 即取 L_c 的左上角 $(p-k) \times (p-k)$ 主矩阵.

参 考 文 献

- 1 Barlow, J. L. and Handy, S. L. , The direct solution of weighted and equality constrained least squares problems, *SIAM. J. Sci. Statist, Comput.* , 9(1988), 704—716.
- 2 Björck, Å. , Iterative refinement of linear least squares solutions I, *BIT* , 7(1967), 257—278.
- 3 Golub, G. H. and Van Loan, C. F. , *Matrix Computation*, Second Edition, The John Hopkins University Press, Baltimore, MD, 1989.
- 4 Gulliksson, M. E. and Wedin, P. Å. , Modifying the QR decomposition to weighted and constrained linear least squares, *SIAM J. Matrix Anal. Appl.* , 13(1992), 1298—1313.
- 5 Paige, C. C. , Fast numerically stable computations for generalized linear least squares problems, *SIAM J. Numer. Anal. Appl.* , 16(1979), 165—171.
- 6 Gulliksson, M. E. , Backward error analysis for the constrained and weighted linear least squares problem when using the weighted QR decomposition, *SIAM J. Matrix Anal. Appl.* 16(1995), 675—687.
- 7 Wei, M. (魏木生), Algebraic property of the rank deficient equality constrained least squares and weighted least squares problems, *LAA* , 161(1992), 27—43.
- 8 Zhao, J. (赵金熙), A modified conjugate direction method for computing the pseudoinverse, *JCM* , 12(1994), 185—194.
- 9 赵金熙, 一类广义半正定线性方程组的直接解法, *高等学校计算数学学报* 18(1996)1.