线性流形上矩阵方程 $A^TXA=B$ 的 D 对称解

The D-symmetric Solutions of Matrix Equation $A^TXA = B$ on the Linear Manifold

李珍珠
LI Zhen-zhu

(湖南科技学院数学与计算科学系,湖南永州 425100)
(Department of Mathematics and Computational Science, Hunan University of Science and Engineering, Yongzhou, Hunan, 425100, China)

摘要:利用矩阵对的广义奇异值分解,给出线性流形上矩阵方程 $A^TXA=B$ 存在 D 对称解的充要条件及其通解的表达式,并导出线性流形上矩阵方程 $A^TXA=B$ 的 D 对称最小二乘解的表达式.

关键词:矩阵方程 线性流形 广义奇异值分解 D 对称矩阵 最小二乘解
中图法分类号:O241.6 **文献标识码**:A **文章编号**:1002-7378(2008)03-0174-03

Abstract: By applying the generalized singular value decomposition of a matrix pair, the necessary and sufficient conditions and the expressions for the D-symmetric solutions of matrix equation $A^TXA=B$ on the linear manifold are established. In addition, the least-squares solutions of $A^TXA=B$ with respect to D-symmetric matrix are derived.

Key words: matrix equation, linear manifold, generalized singular value decomposition, D-symmetric matrix, least-squares solution

在工程技术的许多领域广泛应用到各种特殊矩阵,D 对称矩阵就是其中的一类[1,2]. 文献[1]研究了 D 对称矩阵反问题的最小二乘解及其逼近问题,给出了最小二乘解的一般表达式;文献[2]研究了线性流形上 D 对称矩阵反问题的最小二乘解及其逼近问题. 本文主要对线性流形上矩阵方程 $A^TXA=B$ 的 D 对称解进行讨论.

1 相关定义

令 $R^{n\times m}$ 表示所有 $n\times m$ 阶实矩阵集合;$OR^{n\times n}$ 表示所有 n 阶正交矩阵全体;A^+ 表示 A 的 Moore-Penrose 广义逆;I_k 表示 k 阶单位阵;$SR^{n\times n}$ 表示所有 n 阶实对称矩阵集合,$ASR^{n\times n}$ 表示所有 n 阶反对称矩阵集合,$\forall A,B\in R^{n\times m}$,$A*B$ 表示 A 与 B 的 Hadamard 积. 在 $R^{n\times m}$ 中引入内积,设 $A,B\in R^{n\times m}$,定义 A 与 B 的内积为 $\langle A,B\rangle=\mathrm{tr}(B^TA)$,由此内积诱导的范数为 $\|A\|=\sqrt{\langle A,A\rangle}=\sqrt{\mathrm{tr}(A^TA)}$. 显然,此范数为 Frobenius 范数,$R^{n\times m}$ 构成一个完备的内积空间.

定义[1] 给定 $D=\mathrm{diag}(d_1,d_2,\cdots,d_n)>0$,若 $D^2X\in SR^{n\times n}$,则称 X 为 D 对称矩阵,D 对称矩阵的全体记为 $D^{-2}SR^{n\times n}$,即 $D^{-2}SR^{n\times n}=\{X\,|\,X=D^{-2}B,\ \forall B\in SR^{n\times n}\}$,令

$$S=\{X\in D^{-2}SR^{n\times n}\,|\,XZ=Y,Y=Y(D^2Z)^+D^2Z,\ Z^TD^2Y=Y^TD^2Z,Z,Y\in R^{n\times k}\},\qquad(1)$$

则由文献[1]知 S 是非空的线性流形.

2 D 对称解的充要条件及其通解表达式

问题 I 已知 $A\in R^{n\times m}$,$B\in R^{m\times m}$,求 $X\in S$ 使得 $A^TXA=B$.

引理 1[1] 设 D^2Z 有奇异值分解,

$$D^2Z=U\begin{pmatrix}\Sigma & 0\\ 0 & 0\end{pmatrix}V^T=U_1\Sigma V_1^T,$$

其中:$U=(U_1,U_2)\in OR^{n\times n}$,$U_1\in R^{n\times t}$;$V=(V_1,V_2)\in OR^{k\times k}$,$V_1\in R^{k\times t}$;$\Sigma=\mathrm{diag}(\sigma_1,\sigma_2,\cdots,\sigma_t)$,$\sigma_i>0$,$i$

收稿日期:2007-08-03
作者简介:李珍珠(1966-),女,教授,主要从事数值代数研究.

$=1,2,\cdots,t$.

$$Y_0=\begin{pmatrix} U_1^{\mathrm{T}}YV_1\sum^{-1} & \sum^{-1}V_1^{\mathrm{T}}Y^{\mathrm{T}}U_2 \\ U_2^{\mathrm{T}}YV_1\sum^{-1} & 0 \end{pmatrix}.$$

则(1)式中的 S 集合具有形式:

$$S=\left\{UY^0U^{\mathrm{T}}D^2+U\begin{pmatrix} 0 & 0 \\ 0 & G\end{pmatrix}U^{\mathrm{T}}D^2 \mid G\in SR^{(n-t)\times(n-t)}\right\}. \quad (2)$$

在问题 I 中,记

$$\overline{B}=B-A^{\mathrm{T}}UY_0U^2D^2A, \quad (3)$$

$$A^{\mathrm{T}}U=(\underset{t}{A_1},\underset{n-t}{A_2}), \quad (4)$$

$$U^{\mathrm{T}}D^2A=\begin{array}{l}\begin{pmatrix}\overline{A_1}\\ \overline{A_2}\end{pmatrix}\begin{array}{l} t \\ n-t\end{array}\end{array}. \quad (5)$$

设矩阵对 $(A_2^{\mathrm{T}},\overline{A_2})$ 的广义奇异值分解为:

$$A_2^{\mathrm{T}}=M\sum\nolimits_1P^{\mathrm{T}},\overline{A_2}=M\sum\nolimits_2Q^{\mathrm{T}}, \quad (6)$$

其中 $A_2\in R^{m\times(n-t)}$,$\overline{A_2}\in R^{(n-t)\times m}$,$M\in R^{(n-t)\times(n-t)}$ 非奇异,$P\in OR^{m\times m}$,$Q\in OR^{m\times m}$,

$$\sum\nolimits_1=\begin{array}{c} \begin{pmatrix} I_1 & O & O \\ O & S_1 & O \\ O & O & O_1 \\ O & O & O \end{pmatrix}\begin{array}{l} r \\ s \\ k-r-s \\ n-t-k \end{array} \\ \begin{array}{ccc} r & s & m-r-s \end{array}\end{array},$$

$$\sum\nolimits_2=\begin{array}{c} \begin{pmatrix} O_2 & O & O \\ O & S_2 & O \\ O & O & I_1 \\ O & O & O \end{pmatrix}\begin{array}{l} r \\ s \\ k-r-s \\ n-t-k \end{array} \\ \begin{array}{ccc} m-k+r & s & k-r-s \end{array}\end{array},$$

$k=\mathrm{rank}(A_2^{\mathrm{T}},\overline{A_2})$,$r=k-\mathrm{rank}(\overline{A_2})$,$s=\mathrm{rank}(A_2^{\mathrm{T}})+\mathrm{rank}(\overline{A_2})-k$.$I_1$ 和 I_2 都是单位矩阵,O_1 和 O_2 都是零矩阵,$S_1=\mathrm{diag}(\alpha_1,\alpha_2,\cdots,\alpha_s)$,$S_2=\mathrm{diag}(\beta_1,\beta_2,\cdots,\beta_s)$,并且 $1>\alpha_1\geqslant\alpha_2\geqslant\cdots\geqslant\alpha_s>0$,$0<\beta_1\leqslant\beta_2\leqslant\cdots\leqslant\beta_s<1$,$\alpha_i^2+\beta_i^2=1(i=1,2,\cdots,s)$.

$M^{\mathrm{T}}GM$ 与 $P^{\mathrm{T}}\overline{B}Q$ 按如下形式进行分块:

$$M^{\mathrm{T}}GM=\begin{array}{c} \begin{pmatrix} G_{11} & G_{12} & G_{13} & G_{14} \\ G_{12}^{\mathrm{T}} & G_{22} & G_{23} & G_{24} \\ G_{13}^{\mathrm{T}} & G_{23}^{\mathrm{T}} & G_{33} & G_{34} \\ G_{14}^{\mathrm{T}} & G_{24}^{\mathrm{T}} & G_{34}^{\mathrm{T}} & G_{44} \end{pmatrix}\begin{array}{l} r \\ s \\ k-r-s \\ n-t-k \end{array} \\ \begin{array}{cccc} r & s & k-r-s & n-t-k \end{array}\end{array}, \quad (7)$$

$$P^{\mathrm{T}}\overline{B}Q=\begin{array}{c} \begin{pmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{pmatrix}\begin{array}{l} r \\ s \\ m-r-s \end{array} \\ \begin{array}{ccc} m-k+r & s & k-r-s \end{array}\end{array}. \quad (8)$$

定理1 已知 $A\in R^{n\times m}$,$B\in R^{m\times m}$,$\overline{B}$,$A^{\mathrm{T}}U$,$U^{\mathrm{T}}D^2A$ 分别如(3)~(5)式.矩阵对 $(A_2^{\mathrm{T}},\overline{A_2})$ 的广义奇异值分解如(6)式,$M^{\mathrm{T}}GM$ 与 $P^{\mathrm{T}}\overline{B}Q$ 分别按(7)式与(8)式进行分块,则问题 I 有解的充要条件是

$$(B_{11},B_{21},B_{31})=0,(B_{32},B_{33})=0. \quad (9)$$

当(9)式成立时,问题 I 的一般解可表示为

$$X=UY_0U^{\mathrm{T}}D^2+U\begin{pmatrix} O & O \\ O & G\end{pmatrix}U^{\mathrm{T}}D^2, \quad (10)$$

其中 $G=$

$$M^{-\mathrm{T}}\begin{pmatrix} G_{11} & B_{12}S_2^{-1} & B_{13} & G_{14} \\ S_2^{-1}B_{12}^{\mathrm{T}} & S_1^{-1}B_{22}S_2^{-1} & S_1^{-1}B_{23} & G_{24} \\ B_{13}^{\mathrm{T}} & B_{23}^{\mathrm{T}}S_1^{-1} & G_{33} & G_{34} \\ G_{14}^{\mathrm{T}} & G_{24}^{\mathrm{T}} & G_{34}^{\mathrm{T}} & G_{44}\end{pmatrix}M^{-1},$$

G_{11},G_{14},G_{24},G_{33},G_{34},G_{44} 为任意矩阵.

证明 由引理1可知,矩阵方程 $A^{\mathrm{T}}XA=B$ 若在 S 上有解,则解可以表示为(2)式,且

$$A^{\mathrm{T}}XA=A^{\mathrm{T}}\left[UY_0U^{\mathrm{T}}D^2+U\begin{pmatrix} O & O \\ O & G\end{pmatrix}U^{\mathrm{T}}D^2\right]A,$$

其中 G 满足方程

$$A_2G\,\overline{A_2}=\overline{B}. \quad (11)$$

将(6)式代入(11)式得

$$\sum\nolimits_1^{\mathrm{T}}M^{\mathrm{T}}GM\sum\nolimits_2=P^{\mathrm{T}}\overline{B}Q. \quad (12)$$

将(7)~(8)式代入(12)式,整理得:

$$\begin{pmatrix} O & G_{12}S_2 & G_{13} \\ O & S_1G_{22}S_2 & S_1G_{23} \\ O & O & O\end{pmatrix}=\begin{pmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33}\end{pmatrix}. \quad (13)$$

比较(13)式的两边可知(9)式成立,并且 $G_{12}=B_{12}S_2^{-1}$,$G_{13}=B_{13}$,$G_{22}=S_1^{-1}B_{22}S_2^{-1}$,$G_{23}=S_1^{-1}B_{23}$.

故矩阵方程 $A^{\mathrm{T}}XA=B$ 在线性流形 S 上的 D 对称解可以表示为(10)式的形式.

反过来,若(9)式成立,取

$$G=M^{-1}\begin{pmatrix} O & B_{12}S_2^{-1} & B_{13} & O \\ S_2^{-1}B_{12}^{\mathrm{T}} & S_1^{-1}B_{22}S_2^{-1} & S_1^{-1}B_{23} & O \\ B_{13}^{\mathrm{T}} & B_{23}^{\mathrm{T}}S_1^{-1} & O & O \\ O & O & O & O\end{pmatrix}M,$$

则由引理1可知,

$$X=UY_0U^{\mathrm{T}}D^2+U\begin{pmatrix} O & O \\ O & G\end{pmatrix}U^{\mathrm{T}}D^2\in S.$$

又因为

$$A^{\mathrm{T}}XA=A^{\mathrm{T}}DY_0U^{\mathrm{T}}D^2A+A^{\mathrm{T}}U\begin{pmatrix} O & O \\ O & G\end{pmatrix}U^{\mathrm{T}}D^2A=A^{\mathrm{T}}DY_0U^{\mathrm{T}}D^2A+A_2G\,\overline{A_2}=A^{\mathrm{T}}DY_0U^{\mathrm{T}}D^2A+P\sum\nolimits_1^{\mathrm{T}}M^{\mathrm{T}}GM\sum\nolimits_2Q^{\mathrm{T}}=A^{\mathrm{T}}DY_0U^{\mathrm{T}}D^2A+\overline{B}=B.$$

因此,X 是矩阵方程 $A^{\mathrm{T}}XA=B$ 在线性流形 S 上的 D 对称解.

3 D 对称最小二乘解

问题 Ⅱ 已知 $A\in R^{n\times m}$,$B\in R^{m\times m}$,求 $X\in S$ 使

得$\|A^{T}XA-B\|=\min$.

引理2 设$B_{22}\in R^{s\times s}$,$S_1=\mathrm{diag}(\alpha_1,\alpha_2,\cdots,\alpha_s)>0$,$S_2=\mathrm{diag}(\beta_1,\beta_2,\cdots,\beta_s)>0$,则$\|S_1G_{22}S_2-B_{22}\|=\min$在$SR^{s\times s}$中存在唯一解

$$\hat{G}_{22}=S_1^{-1}[\Phi*(B_{22}S_1^{-1}S_2+S_1^{-1}S_2B_{22}^{T})]S_1^{-1},\tag{14}$$

其中$\Phi=(\varphi_{ij})\in R^{s\times s}$,$\varphi_{ij}=\dfrac{(\alpha_i\alpha_j)^2}{(\alpha_i\beta_j)^2+(\alpha_j\beta_i)^2}$,$1\leqslant i,j\leqslant s$.

证明 对于$G_{22}=(g_{ij})\in SR^{s\times s}$和$B_{22}=(b_{ij})\in R^{s\times s}$,展开

$$\|S_1G_{22}S_2-B_{22}\|^2=\sum_{i=1}^{s}(\alpha_ig_{ii}\beta_i-b_{ii})^2+\sum_{1\leqslant i<j\leqslant s}g_{ij}^2[(\alpha_i\beta_j-b_{ij})^2+(\alpha_j\beta_i-b_{ji})^2].\tag{15}$$

在(15)式中用微分法求驻点得到

$$g_{ij}=\frac{b_{ij}\alpha_i\beta_j+b_{ji}\alpha_j\beta_i}{\alpha_i^2\beta_j^2+\alpha_j^2\beta_i^2},1\leqslant i,j\leqslant s.$$

因此(14)式成立.

定理2 符号与定理1相同.问题Ⅰ的解为

$$X=UY_0U^{T}D^2+U\begin{pmatrix}O&O\\O&\hat{G}\end{pmatrix}U^{T}D^2,\tag{16}$$

其中$\hat{G}=$

$$M^{-T}\begin{pmatrix}G_{11}&B_{12}S_1^{-1}&B_{13}&G_{14}\\S_1^{-1}B_{12}^{T}&\hat{G}_{22}&S_1^{-1}B_{23}&G_{24}\\B_{13}^{T}&B_{23}^{T}S_1^{-1}&G_{33}&G_{34}\\G_{14}^{T}&G_{24}^{T}&G_{34}^{T}&G_{44}\end{pmatrix}M^{-1},\hat{G}_{22}=$$

$S_1^{-1}[\Phi*(B_{22}S_1^{-1}S_2+S_1^{-1}S_2B_{22}^{T})]S_1^{-1}$,$\Phi=(\varphi_{ij})\in R^{s\times s}$,$\varphi_{ij}=\dfrac{(\alpha_i\beta_j)^2}{(\alpha_i\beta_j)^2+(\alpha_j\beta_i)^2}$,$1\leqslant i,j\leqslant s$. $G_{14},G_{24},G_{33},G_{34},G_{44}$为任意矩阵.

证明 设$X\in S$,则由引理1知

$$X=UY_0U^{T}D^2+U\begin{pmatrix}O&O\\O&G\end{pmatrix}U^{T}D^2,$$

而且

$$\|A^{T}XA-B\|^2=\|A^{T}UY_0U^{T}D^2A+A^{T}U\begin{pmatrix}O&O\\O&G\end{pmatrix}U^{T}D^2A-B\|^2=\|A_2G\overline{A_2}-\overline{B}\|^2=$$

$$\|P\sum\nolimits_1^{T}M^{T}GM\sum\nolimits_2Q^{T}-\overline{B}\|^2=\|\sum\nolimits_1^{T}M^{T}GM\sum\nolimits_2-P^{T}\overline{B}Q\|^2=$$

$$\left\|\begin{pmatrix}-B_{11}&G_{12}S_2-B_{12}&G_{13}-B_{13}\\-B_{21}&S_1G_{22}S_2-B_{22}&S_1G_{23}-B_{23}\\-B_{31}&-B_{32}&-B_{33}\end{pmatrix}\right\|^2=$$

$\|G_{12}S_2-B_{12}\|^2+\|G_{13}-B_{13}\|^2+\|S_1G_{22}S_2-B_{22}\|^2+\|S_1G_{23}-B_{23}\|^2+\|B_{11}\|^2+\|B_{21}\|^2+\|B_{31}\|^2+\|B_{32}\|^2+\|B_{33}\|^2$.

因此$\|A^{T}XA-B\|=\min$当且仅当$\|G_{12}S_2-B_{12}\|=\min$,$\|G_{13}-B_{13}\|=\min$,$\|S_1G_{23}-B_{23}\|=\min$,$\|S_1G_{22}S_2-B_{22}\|=\min$.

显然$G_{12}=B_{12}S_2^{-1}$,$G_{13}=B_{13}$,$G_{23}=S_1^{-1}B_{23}$.由引理2知$\hat{G}_{22}$为(14)式.

因此问题Ⅰ的一般解为(16)式.

参考文献:

[1] 张忠志,周富照,胡锡炎.D对称矩阵反问题的最小二乘解[J].湖南大学学学报,2001,28(5):6-10.
[2] 易学军,张忠志,周富照.线性流形上D对称矩阵反问题的最小二乘解[J].数学理论与应用,2002,22(1):93-97.

(责任编辑:韦廷宗)

(上接第173页)

[3] 赵林城.一类离散分布参数的经验Bayes估计的收敛速度[J].数学研究与评论.1981,1:59-69.
[4] 陈家清,刘次华.线性指数分布参数的经验贝叶斯估计[J].华中科技大学学报,2006,10:122-124.
[5] Joag-Dev K,Proschan F.Negative association of random variables with applications[J].Ann Statist,1983,11:286-295.
[6] 魏立力,张文修.线性指数危险率模型的贝叶斯判别分析[J].数学物理学报,2003,23A(4):436-443.
[7] Pan J M.On the convergence rates in the central limit theorem for negatively associated sequences[J].Chinese Journal of Applied Probability and Statistics,1997,13(2):183-192.
[8] 陈家清,刘次华.负相伴样本线性指数分布参数的经验Bayes双侧检验问题[J].数学理论与应用,2006(3):42-47.

(责任编辑:尹 闯)