第4卷 第4期　控制理论与应用　Vol.4, No.4
1987年12月　CONTROL THEORY AND APPLICATIONS　Dec., 1987

协方差矩阵的对称分块表示及其应用

罗续成

（北京遥测技术研究所）

摘　要

本文推导了协方差矩阵的对称分块表示法．作为这种表示法在控制工程中应用的实例，对B.Friedland的递推滤波中的常值偏差处理法，在一般时变偏差情况下，给出了解析证明．

一、引　　言

对协方差阵 $D\in R^{(n+m)\times(n+m)}$ 作对称分块

$$D=\begin{array}{c}\left[\begin{array}{c|c} D_1 & D_2 \\ \hline D_2^T & D_3 \end{array}\right]\begin{array}{l} n \\ m \end{array} \\ \begin{array}{cc} n \quad & \quad m \end{array}\end{array}$$

如果 D 是正定阵，那么子块 D_3 也是正定阵．不难证明如下表示形式成立

$$\begin{cases} D_2=SD_3, \\ D_1=D_c+SD_3S^T, \end{cases} \tag{1}$$

其中，$S\in R^{n\times m}$，$D_c\in R^{n\times n}$．D_3 为半正定的情况是否也能作这样的分块表示呢？

本文证明，不论为正定阵或半正定阵，凡协方差阵都可以化作（1）式这样的对称分块表示．作为这种表示法的应用实例，我们对B.Friedland的递推滤波中的常值偏差处理法[1]在一般时变偏差的情况下给出了一种解析而简单的证明．

二、协方差阵的对称分块表示

考虑随机向量 $X=(x_1,x_2,\cdots,x_n)^T$，$Y=(y_1,y_2,\cdots,y_m)^T$，它们各分量的方差和分量间的协方差有限．X的协方差阵为 $D=[d_{ij}]_{n\times n}$，X，Y的互协方差阵 $C=[c_{ij}]_{n\times m}$．用 E 表示数学期望，便有

$$d_{ij}=E[(x_i-Ex_i)(x_j-Ex_j)], \tag{2}$$

$$c_{ij}=E[(x_i-Ex_i)(y_j-Ey_j)]. \tag{3}$$

本文于1985年12月9日收到．1986年7月10日收到修改稿．

引理 1 D的秩为r的充分必要条件是：按概率1，x_1，$x_2,\cdots,x_n$间有且只有$n-r$个满足$\sum_{i=1}^{n} t_i(x_i-Ex_i)=0$的线性独立关系式。这里$r\leq n$。

这个引理参见文献〔2〕。因为据$n-r$个线性独立关系式

$$\sum_{i=1}^{n} t_i^{(j)}(x_i-Ex_i)=0, \qquad (j=r+1,\cdots,n)$$

可将$n-r$个变量表示为其余变量的线性组合，所以，如果D的秩为r，便必有$n-r$个变量可表示为其余变量的线性组合。这样我们有推论：

推论 D的秩为r的充分必要条件是：在$x_1,\cdots,x_n$中最多有r个变量按概率1线性无关，而任取$r+1$个或更多个变量则按概率1至少存在一个线性关系。

引理 2 如果$\mathrm{rank}\,D=r$，那么必有$\mathrm{rank}[D,C]=r$。

证 D的秩为r，据引理1及推论可知，对于$x_i(i=r+1,\cdots,n)$，必对应存在常数组$t_1^{(i)},\cdots,t_r^{(i)}$，满足

$$P\left[(x_i-Ex_i)=\sum_{k=1}^{r} t_k^{(i)}(x_k-Ex_k)\right]=1 \qquad (4)$$

和

$$P\left[(x_i-Ex_i)\neq\sum_{k=1}^{r} t_k^{(i)}(x_k-Ex_k)\right]=0, \qquad (5)$$

这里，P表示概率。据（2）式，D的第i行j列元素（$i=r+1,\cdots,n$；$j=1,\cdots,n$)为

$$\begin{aligned}
d_{ij}=&P\left[(x_i-Ex_i)=\sum_{k=1}^{r} t_k^{(i)}(x_k-Ex_k)\right]E\left[(x_i-Ex_i)\right.\\
&\left.(x_j-Ex_j)\middle|(x_i-Ex_i)\neq\sum_{k=1}^{r} t_k^{(i)}(x_k-Ex_k)\right]\\
&+P\left[(x_i-Ex_i)\neq\sum_{k=1}^{r} t_k^{(i)}(x_k-Ex_k)\right]E\left[(x_i-Ex_i)\right.\\
&\left.(x_j-Ex_j)\middle|(x_i-Ex_i)\neq\sum_{k=1}^{r} t_k^{(i)}(x_k-Ex_k)\right]\\
&=E\left[(x_i-Ex_i)(x_j-Ex_j)\middle|(x_i-Ex_i)=\sum_{k=1}^{r} t_k^{(i)}(x_k-Ex_k)\right]\\
&=\sum_{k=1}^{r} t_k^{(i)}E(x_k-Ex_k)(x_j-Ex_j)\\
&=\sum_{k=1}^{r} t_k^{(i)}d_{kj}.
\end{aligned} \qquad (6)$$

C 阵的第 i 行 l 列元素（$i=r+1,\cdots,n$；$l=1,\cdots,m$）为

$$C_{il}=P\left[(x_i-Ex_i)=\sum_{k=1}^{r}t_k^{(i)}(x_k-Ex_k)\right]E\left[(x_i-Ex_i)(y_l-Ey_l)\middle|(x_i-Ex_i)=\sum_{k=1}^{r}t_k^{(i)}(x_k-Ex_k)\right]$$

$$+P\left[(x_i-Ex_i)\neq\sum_{k=1}^{r}t_k^{(i)}(x_k-Ex_k)\right]E\left[(x_i-Ex_i)(y_l-Ey_l)\middle|(x_i-Ex_i)\neq\sum_{k=1}^{r}t_k^{(i)}(x_k-Ex_k)\right]$$

$$=E\left[(x_i-Ex_i)(y_l-Ey_l)\middle|(x_i=Ex_i)=\sum_{k=1}^{r}t_k^{(i)}(x_k-Ex_k)\right]$$

$$=\sum_{k=1}^{r}t_k^{(i)}E(x_k-Ex_k)(y_l-Ey_l)$$

$$=\sum_{k=1}^{r}t_k^{(i)}C_{kl}. \tag{7}$$

$$[D,C]=\begin{pmatrix} d_{11} & \cdots & d_{1n} & c_{11} & \cdots & c_{1m} \\ \vdots & & \vdots & \vdots & & \vdots \\ d_{r1} & \cdots & d_{rn} & c_{r1} & \cdots & c_{rm} \\ \sum_{k=1}^{r}t_k^{(r+1)}d_{k1} & \cdots & \sum_{k=1}^{r}t_k^{(r+1)}d_{kn} & \sum_{k=1}^{r}t_k^{(r+1)}c_{i1} & \cdots & \sum_{k=1}^{r}t_k^{(r+1)}c_{im} \\ \vdots & & \vdots & \vdots & & \vdots \\ \sum_{k=1}^{r}t_k^{(n)}d_{k1} & \cdots & \sum_{k=1}^{r}t_k^{(n)}d_{kn} & \sum_{k=1}^{r}t_k^{(n)}c_{i1} & \cdots & \sum_{k=1}^{r}t_k^{(n)}c_{im} \end{pmatrix}. \tag{8}$$

分析（8）式可见，该矩阵后 $n-r$ 行均是前 r 行的线性组合，因此它最多有 r 个线性不相关的行，$\mathrm{rank}[D,C]$不大于 r。

又因为 $\mathrm{rank}D=r$，所以 $\mathrm{rank}[D,C]$不小于 r。

综上，必有结论

$\mathrm{rank}[D,C]=r$。

定理　对任何协方差阵 D 作对称分块

$$D=\begin{array}{c}\left[\begin{array}{c|c} D_1 & D_2 \\ \hline D_2^T & D_3 \end{array}\right]\begin{array}{l} n \\ m \end{array} \\ \begin{array}{cc} n & m \end{array}\end{array} \tag{9}$$

其子块必能表示为如下形式

$$D_1 = D_c + SD_3S^T, \tag{10}$$

$$D_2 = SD_3, \tag{11}$$

这里，$S \in R^{n\times m}$，$D_c \in R^{n\times n}$。

证　设 D 对应向量为 $\begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix}$，$\mathbf{X} = [x_1, \cdots, x_n]^T$，$\mathbf{Y} = [y_1, \cdots, y_m]^T$，有

$$\begin{cases} D_1 = E[(\mathbf{X} - E\mathbf{X})(\mathbf{X} - E\mathbf{X})^T], \\ D_2 = E[(\mathbf{X} - E\mathbf{X})(\mathbf{Y} - E\mathbf{Y})^T], \\ D_3 = E[(\mathbf{Y} - E\mathbf{Y})(\mathbf{Y} - E\mathbf{Y})^T]. \end{cases} \tag{12}$$

考察矩阵方程

$$SD_3 = D_2. \tag{13}$$

该方程是关于 S 阵元素的 n 组系数相同、右端项不同的 m 维线性方程组。由系数矩阵 D_3 与右端项组成的增广矩阵为

$$\left[D_3, D_2^T\right] = \left[E(\mathbf{Y} - E\mathbf{Y})(\mathbf{Y} - E\mathbf{Y})^T,\ E(\mathbf{Y} - E\mathbf{Y})(\mathbf{X} - E\mathbf{X})^T\right]. \tag{14}$$

据引理 2 得

$$\operatorname{rank}\left[D_3, D_2^T\right] = \operatorname{rank} D_3. \tag{15}$$

又据线性方程组有解的判定定理可知，(13)有解，且其范数最小的解唯一确定为

$$S = D_2 D_3^+, \tag{16}$$

D_3^+ 表示 D_3 的伪逆。进一步有

$$SD_3S^T = D_2D_3^+D_3D_3^+D_2 \tag{17}$$

存在且唯一。令

$$D_c = D_1 - SD_3S^T. \tag{18}$$

那么，矩阵 D 的子块的下述表示存在且唯一：

$$D_2 = SD_3, \tag{19}$$

$$D_1 = D_c + SD_3S^T. \tag{20}$$

定理得证。

三、Kalman 滤波中的偏差处理

考虑动力学方程和量测方程如下的离散系统

$$\begin{cases} x_1(k+1) = A_{11}(k)x_1(k) + A_{12}(k)x_2(k) + \xi(k), \\ x_2(k+1) = \qquad\qquad\qquad\quad A_{22}(k)x_2(k), \end{cases} \tag{21}$$

$$y(k) = C_{11}(k)x_1(k) + C_{12}(k)x_2(k) + \eta(k), \tag{22}$$

其中，$x_1 \in R^{n_1}$表示待求状态，$x_2 \in R^{n_2}$表示源于系统动力学模型或量测模型的偏差状态，其模型不受噪声影响。$y \in R^m$。

$$\begin{cases} C_{ov}(\xi(k),\xi(j)) = Q(k)\delta_{kj}, \\ C_{ov}(\eta(k),\eta(j)) = R(k)\delta_{kj}, \\ C_{ov}(\xi(k),\eta(j)) = 0, \end{cases} \tag{23}$$

其中 $$\delta_{kj} = \begin{cases} 1 & k=j \\ 0 & k \neq j \end{cases}$$

是Kronecker delta函数。为简化符号，下面用（+）表示$(k+1)$，用（−）表示$(k-1)$。例如：$x_1(+) \triangleq x_1(k+1)$，$x_1 \triangleq x_1(k)$，$x_1(-) \triangleq x_1(k-1)$。其余类推。

令

$$X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix},$$
$$E = \begin{bmatrix} I \\ 0 \end{bmatrix}, \qquad C = [C_{11} \quad C_{12}], \tag{24}$$

其中，I为单位矩阵。这样，系统可表示为

$$X(+) = AX + E\xi, \tag{25}$$
$$Y = CX + \eta. \tag{26}$$

对X可构成Kalman滤波估计器如下：

$$\hat{X} = A(-)\hat{X}(-) + K(Y - CA(-)\hat{X}(-)), \tag{27}$$

其中 $$K = PC^T(CPC^T + R)^{-1}, \tag{28}$$

或 $$K = TC^TR^{-1} \tag{29}$$

$$P = A(-)T(-)A^T(-) + EQ(-)E^T, \tag{30}$$
$$T = P - KCP. \tag{31}$$

上述公式中，$\hat{X}$表示X的线性无偏最小方差估计，K为滤波增益阵，P、T分别表示X的预测和滤波误差协方差阵。

将P、T作对称分块

$$P = \begin{array}{c} \left[\begin{array}{c|c} P_1 & P_{12} \\ \hline P_{12}^T & P_2 \end{array}\right] \begin{array}{l} n_1 \\ n_2 \end{array} \\ \begin{array}{cc} n_1 & \quad n_2 \end{array} \end{array}, \qquad T = \begin{array}{c} \left[\begin{array}{c|c} T_1 & T_{12} \\ \hline T_{12}^T & T_2 \end{array}\right] \begin{array}{l} n_1 \\ n_2 \end{array} \\ \begin{array}{cc} n_1 & \quad n_2 \end{array} \end{array}. \tag{32}$$

据前面证明的定理可知，P、T的子块可表示为

$$\begin{cases} P_{12} = UP_2, \\ P_1 = \widetilde{P}_1 + UP_2U^T, \end{cases} \tag{33}$$

$$\begin{cases} T_{12}=FT_2, \\ T_1=\widetilde{T}_1+FT_2F^T. \end{cases} \tag{34}$$

将（32）～（34）代入（27）～（31），便可用新参数代替原滤波器参数的递推关系。结果如下：

参数矩阵 $F\in R^{n_1\times n_2}$ 的递推关系

$$G(-)=A_{11}(-)F(-)+A_{12}(-), \tag{35}$$

$$H=C_{11}G(-)+C_{12}A_{22}(-), \tag{36}$$

$$F=(G(-)-\widetilde{K}_1H)A_{22}^{-1}(-), \tag{37}$$

其中，$G\in R^{n_1\times n_2}$，$H\in R^{m\times n_2}$．x_2 的滤波误差协方差阵为

$$T_2=A_{22}(-)[T_2(-)-T_2(-)H^T(R+C_{11}\widetilde{P}_1C_{11}^T+HT_2(-)H^T)^{-1}HT_2(-)]A_{22}^T(-). \tag{38}$$

滤波增益阵 $k=\begin{array}{c} \left[\begin{array}{c} k_1 \\ k_2 \end{array}\right]\begin{array}{c} n_1 \\ n_2 \end{array} \\ m \end{array}$ 有关系

$$J=C_{11}F+C_{12}, \tag{39}$$

$$K_2=T_2J^TR^{-1}, \tag{40}$$

$$K_1=\widetilde{K}_1+FK_2. \tag{41}$$

预测误差协方差阵的有关参数为

$$P_2=A_{22}(-)T_2(-)A_{22}^T(-), \tag{42}$$

$$U=G(-)A_{22}^{-1}(-), \tag{43}$$

其中，$J\in R^{m\times n_2}$，$U\in R^{n_1\times n_2}$．

$\widetilde{P}_1$、$\widetilde{K}_1$、$\widetilde{T}_1$的递推关系为

$$\widetilde{P}_1=A_{11}(-)T_1(-)A_{11}^T(-)+Q(-), \tag{44}$$

$$\widetilde{K}_1=\widetilde{P}_1C_{11}^T(C_{11}\widetilde{P}_1C_{11}^T+R)^{-1}, \tag{45}$$

或

$$\widetilde{K}_1=\widetilde{T}_1C_{11}^TR^{-1}, \tag{46}$$

$$\widetilde{T}_1=\widetilde{P}_1-\widetilde{K}_1C_{11}\widetilde{P}_1. \tag{47}$$

它们分别表示假定 $x_2=0$ 时 x_1 的预测误差协方差阵，滤波增益阵，和滤波误差协方差阵。

又

$$\hat{x}_1=A_{11}(-)\hat{x}_1(-)+A_{12}(-)\hat{x}_2(-)+K_1[y-C_{11}A_{11}(-)\hat{x}_1(-)$$

$$-(C_{11}A_{12}(-)+C_{12}A_{22}(-))\hat{x}_2(-)], \tag{48}$$

$$\hat{x}_2=A_{22}(-)\hat{x}_2(-)+K_2[y-C_{11}A_{11}(-)\hat{x}_1(-)-(C_{11}G(-)$$
$$+C_{12}A_{22}(-))\hat{x}_2(-)] \tag{49}$$

可以证明

$$\hat{x}_1=\tilde{x}_1+F\hat{x}_2, \tag{50}$$

其中，$\tilde{x}_1$为假定 $x_2=0$ 条件下 x_1 的估计。

$$\tilde{x}_1=A_{11}(-)\tilde{x}_1(-)+\tilde{K}_1 r_1, \tag{51}$$

$$r_1=y-C_{11}A_{11}(-)\tilde{x}_1. \tag{52}$$

我们用归纳法证明上述结论如下：设在时刻 $k-1$（50）式成立，即

$$\hat{x}_1(-)=\tilde{x}_1+F(-)\hat{x}_2(-). \tag{53}$$

将（53）代入（48）、（49）式整理得：

$$\hat{x}_2=A_{22}(-)\hat{x}_2(-)+K_2(r_1-H\hat{x}_2(-)), \tag{54}$$

$$\hat{x}_1=A_{11}(-)\tilde{x}_1(-)+G(-)\hat{x}_2(-)+K_1(r_1-H\hat{x}_2(-)). \tag{55}$$

将（54）、（55）式代入（50）式右端得

$$\begin{aligned}
&\tilde{x}_1+F\hat{x}_2\\
&=A_{11}(-)\tilde{x}_1(-)+\tilde{K}_1 r_1+F[A_{22}(-)\hat{x}_2(-)+K_2(r_1-H\hat{x}_2(-))]\\
&=A_{11}(-)\tilde{x}_1(-)+\tilde{K}_1 r_1+(G(-)-\tilde{K}_1 H)\hat{x}_2(-)+FK_2(r_1-H\hat{x}_2(-))\\
&=A_{11}(-)\tilde{x}_1(-)+G(-)\hat{x}_2(-)+(\tilde{K}_1+FK_2)(r_1-H\hat{x}_2(-))\\
&=A_{11}(-)\tilde{x}_1(-)+G(-)\hat{x}_2(-)+K_1(r_1-H\hat{x}_2(-))\\
&=\hat{x}_1.
\end{aligned}$$

可见，只要满足　$\hat{x}_1(0)=\tilde{x}_1(0)+F(0)\hat{x}_2(0)$,　（56）

那么对所有时刻 k（50）式都成立。如果在初始时刻 x_1、x_2不存在先验的相关性，只要取初始条件

$$\begin{cases}F(0)=0,\\ \tilde{T}_1(0)=T_1(0),\\ \tilde{x}_1(0)=\hat{x}_1(0),\end{cases}$$

便满足了（56）式的要求。

（50）式正是时变偏差情况下 B. Friedland 的偏差处理法的推广。计算 X 的线性无偏最小方差估计的流程如图 1 所示。显然，除最后求 $\hat{x}_1$的相加步骤外，计算$\tilde{x}_1$和$\hat{x}_2$的过程是两个分开的小估计器。

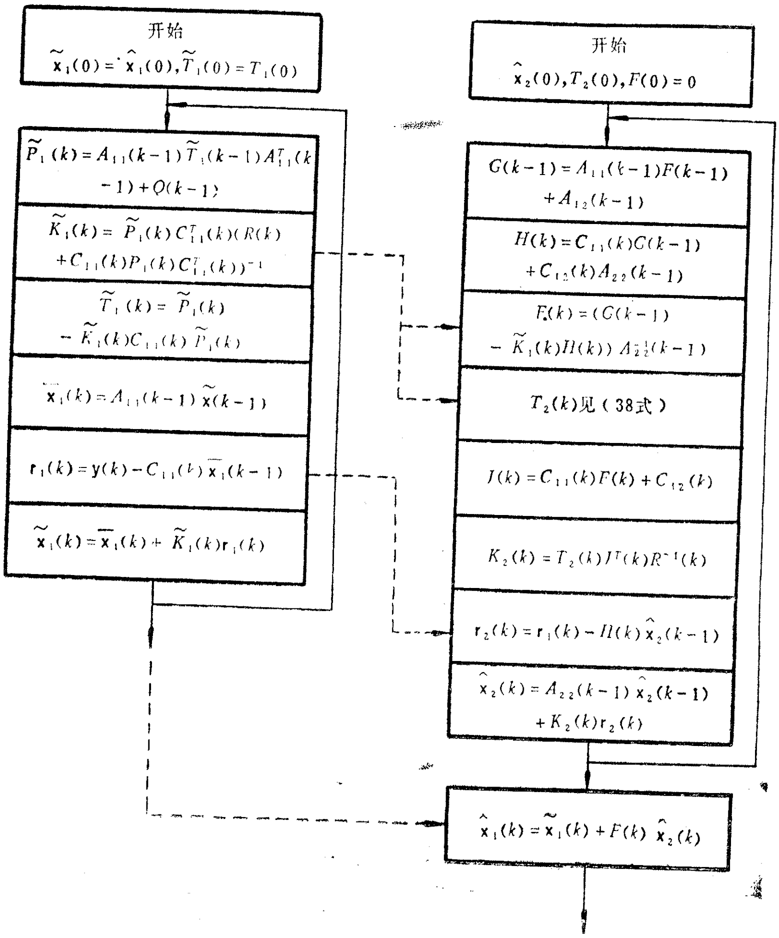

图 1

致谢 本工作得到空间技术研究院杨家墀教授的悉心指导。作者与北京控制工程研究所陈义庆同志讨论并采纳了他的宝贵意见。谨致谢忱。

参 考 文 献

〔1〕 Friedland, B., Treatment of Bias in Recursive Filtering, IEEE Trans. Autom. Control, AC-**14**,（1969）, 359-367.

〔2〕 H.克拉美， 魏宗舒等译，统计学数学 方法，上海科学技术出版社，（1966），287-288.

THE SYMMETRICAL PARTITIONING EXPRESSION OF A COVARIANCE MATRIX AND TI'S APPLICATION

Luo Xucheng

(Beijing Research Institute of Telemetry)

Abstract

The symmetrical partitioning expression of a covariance matrix is derived in this paper. The major result is: The submatrices D_1, D_2 and D_3 of a covariance matrix

$$D = \begin{array}{cc} \left[\begin{array}{c|c} D_1 & D_2 \\ \hline D_2^T & D_3 \end{array}\right] & \begin{array}{c} m \\ n \end{array} \\ \begin{array}{cc} m & n \end{array} & \end{array}$$

must be satisfied by

$$\begin{cases} D_2 = SD_3 \\ D_1 = D_c + SD_3S^T \end{cases}$$

This result was applied to prove analytically B. Friedland's treatment of bias in recursive filtering in time-varying case.