第15卷第2期　　华 东 交 通 大 学 学 报　　Vol.15　No.2
1998年6月　　Journal of East China Jiaotong University　　Jun.　1998

最小二乘递推算法和 Kalman 滤波算法

刘 轩 黄

(江西电力职工大学)

摘　要　就一般加权、最优加权和指数加权三种情形分别介绍了最小二乘递推(RLS)算法和 Kalman 滤波算法,且首次将状态向量的概念扩展到了状态矩阵的概念。这将使我们在某些应用中能采用比以前规模小得多的模型而丝毫不会因此带来任何误差。最后,我们还指出了 RLS 算法和 Kalman 滤波算法所存在的一些问题。

关键词　最小二乘;递推算法;滤波算法;加权;最优;指数
分类号　O 158

1　一般加权情形

考虑方程

$$Y_k^T = h_k^T X + w_k^T, \qquad k = 1, 2 \cdots \tag{1}$$

其中 $Y_k \in R^m$, $h_k \in R^n$ 均已知, $w_k \in R^m$ 为零均值白噪音序列, $X \in R^{n\times m}$ 为待估计矩阵,当获得了观测值 $y_1, y_2, \cdots, y_k$ 时,定义 X 的最小二乘估计 $\hat{X}_k$ 为最优化问题

$$\min_X \|H_k X - Y_k\|_{\Lambda_k} \tag{2}$$

的解,其中

$$H_k = [h_1, h_2, \cdots, h_k]^T \in R^{k\times n}, \tag{3}$$

$$Y_k = [y_1, y_2, \cdots, y_k]^T \in R^{k\times m}, \tag{4}$$

$$\Lambda_k = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_k) \in R^{k\times k}, \quad \lambda_1, \lambda_2, \cdots, \lambda_k > 0 \tag{5}$$

定义

$$\bar{H}_k = \Lambda_k^{1/2} H_k = \begin{bmatrix} \bar{H}_{k-1} \\ \lambda_k^{1/2} h_k^T \end{bmatrix}, \tag{6}$$

$$\bar{Y}_k = \Lambda_k^{1/2} Y_k = \begin{bmatrix} \bar{Y}_{k-1} \\ \lambda_k^{1/2} y_k^T \end{bmatrix}. \tag{7}$$

则

$$\|H_k X - Y_k\|_{\Lambda_k} = \|\bar{H}_k X - \bar{Y}_k\|.$$

收稿日期:1997-07-17　　刘轩黄,男,1935 年生,教授.

于是最优化问题式(2)等价于

$$\min_{X} \| \bar{H}_k X - \bar{Y}_k \| . \tag{8}$$

引理 1　设 A、B、C 和 D 为阶数适当的矩阵,且 A、C 和 $C^{-1}+DA^{-1}B$ 均非异,则

$$(A+BCD)^{-1}=A^{-1}-A^{-1}B(C^{-1}+DA^{-1}B)^{-1}DA^{-1}.$$

证明　可直接验证.

该引理的条件往往被忽略,然而我们将看到,正是这些条件使现行 **RLS** 算法不能给出严格的最小二乘估计.

引理 2　设 $\bar{H}_k$ 如式(6)所给,且 $\bar{H}_{k-1}$ 为列满秩,令

$$\bar{H}_k^{+}=(\bar{H}_{k-1}^{\mathrm{T}}\bar{H}_k)^{-1}\bar{H}_k^{\mathrm{T}}, \tag{9}$$

$$P_{k-1}=(\bar{H}_{k-1}^{\mathrm{T}}\bar{H}_{k-1})^{-1}, \tag{10}$$

$$b_k=P_{k-1}h_k, \tag{11}$$

$$r_k=\lambda_k^{-1}+h_k^{\mathrm{T}}b. \tag{12}$$

其中,$(\cdot)^{+}$ 为矩阵的 Moore-penrose 逆[*],则

$$\bar{H}_k^{+}=[\bar{H}_{k-1}^{+}-g_k h_k^{\mathrm{T}}\bar{H}_{k-1}^{+},\ \lambda_k^{1/2}g_k].$$

其中

$$g_k=b_k r_k^{-1}. \tag{13}$$

证明　由式(6)、(10)~(13)和引理 1 可知

$$P_k=\left([\bar{H}_{k-1}^{\mathrm{T}}\quad \lambda_k^{1/2}h_k]\begin{bmatrix}\bar{H}_{k-1}\\ \lambda_k^{1/2}h_k^{\mathrm{T}}\end{bmatrix}\right)^{-1}=P_{k-1}-g_k b_k^{\mathrm{T}} \tag{14}$$

因此

$$\bar{H}_k^{+}=P_k\bar{H}_k^{\mathrm{T}}=[\bar{H}_{k-1}^{+}-g_k h_k^{\mathrm{T}}\bar{H}_{k-1}^{+}\quad \lambda_k^{1/2}g_k].$$

由引理 2,便知 X 的最小二乘估计为

$$\hat{X}_k=\bar{H}_k^{+}\bar{Y}_k=\hat{X}_{k-1}+g_k(y_k^{\mathrm{T}}-h_k^{\mathrm{T}}\hat{X}_{k-1}) \tag{15}$$

式(11)~(15)便构成了 RLS 算法的全部递推方程.但为启动这一算法.还必须确定初始条件,一种办法是令

$$P_{k_0}=(\bar{H}_{k_0}^{\mathrm{T}}\bar{H}_{k_0})^{-1},\qquad X_{k_0}=P_{k_0}\bar{H}_{k_0}^{\mathrm{T}}\bar{Y}_{k_0}.$$

其中,$k_0\geqslant n$ 是使 $\bar{H}_{k_0}$ 为列满秩的某一整数,但这一方法计算量太大,k_0 的合理选择也往往是困难的,因而可能破坏递推算法的适时性,本文不考虑这种初始条件的设定方法,而像通常那样认为可先已从其它途径获得了合适的 $\hat{X}_0$ 和 P_0.例如,$\hat{X}_0=0$,$P_0=cI$,其中 c 为一很大的正数.

算法 1　给定 $\hat{X}_0$ 和 P_0,令 $b_k=P_{k-1}h_k$,$r_k=\lambda_k^{-1}+h_k b_k$,则式(1)中 X 的最小二乘估计可如下递推求得.

$$\left.\begin{aligned} g_k&=b_k r_k^{-1},\\ \hat{X}_k&=\hat{X}_{k-1}+g_k(y_k^{\mathrm{T}}-h_k^{\mathrm{T}}\hat{X}_{k-1}),\qquad k=1,2,\dots\\ P_k&=P_{k-1}-g_k b_k^{\mathrm{T}}.\end{aligned}\right\} \tag{16}$$

现将上述 RLS 算法应用于系统

[*] 本文所用广义逆的记号采自文献[1]19.

$$\begin{cases} X_k = F_{k,k-1}X_{k-1} + U_{k-1}, & (17) \\ Y_k^{\mathrm{T}} = h_k^{\mathrm{T}} X_k + w_k^{\mathrm{T}} (13) & (18) \end{cases}$$

的状态估计问题，其中 $X_k \in R^{n\times m}$ 为待估计的状态矩阵，$F_{k,k-1} \in R^{n\times n}$ 为已知状态转移矩阵，U_{k-1} 为系统的输入，其余与式(1) 相同(13)

类似地，称最优化问题

$$\left.\begin{aligned} &\min_{X_j} \left\| \begin{bmatrix} h_1^{\mathrm{T}} X_1 \\ \vdots \\ h_k^{\mathrm{T}} X_k \end{bmatrix} - \begin{bmatrix} y_1^{\mathrm{T}} \\ \vdots \\ y_k^{\mathrm{T}} \end{bmatrix} \right\|_{\Lambda_k} \\ &s.t.\quad X_k = F_{k,k-1}X_{k-1} + U_{k-1} \end{aligned}\right\} \tag{19}$$

的解 $\hat{X}_{j,k}(j = 1,\cdots,k)$ 为给定 $y_1, y_2, \cdots, y_k$ 时 X_j 的最小二乘估计，特别地，称 $\hat{X}_{k,k}$ 为 X_k 的最小二乘滤波，这里

$$\Lambda_k = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_k)$$

为 $k \times k$ 正定加权阵，由式(17) 知

$$X_j = F_{j,0}X_0 + \sum_{i=0}^{j-1} F_{j,i+1}U_i, \qquad j = 1, 2, \cdots, k \tag{20}$$

于是最优化问题式(19) 等价于

$$\left\| \begin{bmatrix} h_1^{\mathrm{T}} F_{1,0} \\ \vdots \\ h_k^{\mathrm{T}} F_{k,0} \end{bmatrix} X_0 - \begin{bmatrix} \bar{y}_1^{\mathrm{T}} \\ \vdots \\ \bar{y}_k^{\mathrm{T}} \end{bmatrix} \right\|_{\Lambda_k} \tag{21}$$

其中

$$\bar{y}_j^{\mathrm{T}} = y_j^{\mathrm{T}} - h_j^{\mathrm{T}} \sum_{i=0}^{j-1} F_{j,i+1} U_i \ (13) \tag{22}$$

按算法 1，若给定 $\hat{X}_{0,0}$ 和 $\hat{P}_{0,0}$，并令

$$b_{0,k} = P_{0,k-1} F_{k,0}^{\mathrm{T}} h_k, \tag{23}$$

则 X_0 的最小二乘估计可如下递推地求得

$$\begin{cases} g_{0,k} = b_{0,k} r_k^{-1}, & (24) \\ \hat{X}_{0,k} = \hat{X}_{0,k-1} + g_{0,k}(y_k^{\mathrm{T}} - h_k^{\mathrm{T}} F_{k,0} \hat{X}_{0,k-1}), & (25) \\ P_{0,k} = P_{0,k-1} - g_{0,k} b_{0,k}^{\mathrm{T}} & (26) \end{cases}$$

再令

$$\hat{X}_{k,l} = F_{k,0}\hat{X}_{0,l} + \sum_{i=0}^{k-1} F_{k,i+l} U_i = F_{k,0}\hat{X}_{0,l} D_k, \tag{27}$$

$$P_{k,l} = F_{k,0} P_{0,l} F_{k,0}^{\mathrm{T}} (13) \tag{28}$$

其中 $l = k, k-1$，$D_k = \sum_{i=0}^{k-1} F_{k,i+l} U_i$，以及

$$g_k = F_{k,0} g_{0,k}, \qquad b_k = P_{k,k-1} h_k \tag{29}$$

则式(24) 左乘 $F_{k,0}$，并注意到式(23)、(28) 和(29)，便有

$$g_k = b_k r_k^{-1} (13) \tag{30}$$

式(25) 左乘 $F_{k,0}$，两边再加上 D_k 并利用式(29) 和(22)，便有

$$F_{k,0}\hat{X}_{0,k} + D_k = F_{k,0}\hat{X}_{0,k-1} + D_k + g_k(Y_k^{\mathrm{T}} - h_k^{\mathrm{T}} D_k - h^{\mathrm{T}k} F_{k,0} \hat{X}_{k,k-1}) (13)$$

由式(27) 可知

$$\hat{\boldsymbol{X}}_{k,k} = \hat{\boldsymbol{X}}_{k,k-1} + \boldsymbol{g}_k(\boldsymbol{Y}_k^{\mathrm{T}} - \boldsymbol{h}_k^{\mathrm{T}}\hat{\boldsymbol{X}}_{k,k-1})⒀ \tag{31}$$

其中
$$\hat{\boldsymbol{X}}_{k,k-1} = \boldsymbol{F}_{k,0}\hat{\boldsymbol{X}}_{0,k-1} + \boldsymbol{D}_k = \boldsymbol{F}_{k,k-1}\hat{\boldsymbol{X}}_{k-1,k-1} + \boldsymbol{U}_{k-1}, \tag{32}$$

最后,用 $\boldsymbol{F}_{k,0}$ 和 $\boldsymbol{F}_{k,0}^{\mathrm{T}}$ 分别左乘和右乘式(26),同时利用式(28),便有

$$\boldsymbol{P}_{k,k} = \boldsymbol{P}_{k,k-1} - \boldsymbol{g}_k\boldsymbol{b}_k^{\mathrm{T}}, \tag{33}$$

其中,
$$\boldsymbol{P}_{k,k-1} = \boldsymbol{F}_{k,0}\boldsymbol{P}_{0,k-1}\boldsymbol{F}_{k,0}^{\mathrm{T}} = \boldsymbol{F}_{k,k-1}\boldsymbol{P}_{k-1,k-1}\boldsymbol{F}_{k,k-1}^{\mathrm{T}}, \tag{34}$$

这样,将式(29) ~(34) 结合起来,便可得下述递推滤波算法⒀

算法 2　对系统(17) 和(18),若给定 $\hat{\boldsymbol{X}}_{0,0}$,$\boldsymbol{P}_{0,0}$,并令

$$\begin{cases} \hat{\boldsymbol{X}}_{k,k-1} = \boldsymbol{F}_{k,k-1}\hat{\boldsymbol{X}}_{k-1,k-1} + \boldsymbol{U}_{k-1}, \\ \boldsymbol{P}_{k,k-1} = \boldsymbol{F}_{k,k-1}\boldsymbol{P}_{k-1,k-1}\boldsymbol{F}_{k,k-1}^{\mathrm{T}}, \\ \boldsymbol{b}_k = \boldsymbol{P}_{k,k-1}\boldsymbol{h}_k, \quad r_k = \lambda_k^{-1} + \boldsymbol{h}_k^{\mathrm{T}}\boldsymbol{b}_k⒀ \end{cases}$$

则系统状态矩阵 $\boldsymbol{X}_k$ 的最小二乘滤波可知下递推求得

$$\left.\begin{aligned} &\boldsymbol{g}_k = \boldsymbol{b}_k r_k^{-1}, \\ &\hat{\boldsymbol{X}}_{k,k} = \hat{\boldsymbol{X}}_{k,k-1} + \boldsymbol{g}_k(\boldsymbol{Y}_k^{\mathrm{T}} - \boldsymbol{h}_k^{\mathrm{T}}\hat{\boldsymbol{X}}_{k,k-1}), \\ &\boldsymbol{P}_{k,k} = \boldsymbol{P}_{k,k-1} - \boldsymbol{g}_k\boldsymbol{b}_k^{\mathrm{T}}⒀ \end{aligned}\right\} \tag{35}$$

2　最优加权

在式(17)、(18) 中令 $m = 1$,便有系统

$$\begin{cases} \boldsymbol{X}_k = \boldsymbol{F}_{k,k-1}\boldsymbol{X}_{k-1} + \boldsymbol{u}_{k-1}, & (36) \\ \boldsymbol{y}_k = \boldsymbol{h}_k^{T}\boldsymbol{X}_k + \boldsymbol{w}_k⒀ & (37) \end{cases}$$

其中,$\boldsymbol{X}_k \in R^n$ 为状态向量;$\boldsymbol{F}_{k,k-1} \in R^{n\times n}$,$\boldsymbol{u}_{k-1} \in R^n$ 为输入;$\boldsymbol{y}_k \in R^n$,$\boldsymbol{w}_k \in R$,设 $\boldsymbol{w}_k \sim N(0, \sigma_k^2)$ 为与 $\boldsymbol{X}_0 \sim N(E(\boldsymbol{X}_0), Var(\boldsymbol{X}_0))$ 无关的 Gauss 白噪音⒀则在算法 2 中只要将初始条件设为 $\hat{\boldsymbol{X}}_{0,0} = E(\boldsymbol{X}_0)$,$\boldsymbol{P}_{0,0} = Var(\boldsymbol{X}_0)$,并取最优加权 $\lambda_k = \sigma_k^2$,就可求得 $\boldsymbol{X}_k$ 的最优滤波⒀即有

算法 3　在系统(36) 和(37) 中给定 $\hat{\boldsymbol{X}}_{0,0} = E(\boldsymbol{X}_0)$,$\boldsymbol{P}_{0,0} = Var(\boldsymbol{X}_0)$,并令

$$\begin{cases} \hat{\boldsymbol{X}}_{k,k-1} = \boldsymbol{F}_{k,k-1}\hat{\boldsymbol{X}}_{k-1,k-1} + \boldsymbol{u}_{k-1}, \\ \boldsymbol{P}_{k,k-1} = \boldsymbol{F}_{k,k-1}\boldsymbol{P}_{k-1,k-1}\boldsymbol{F}_{k,k-1}^{\mathrm{T}}, \\ \boldsymbol{b}_k = \boldsymbol{P}_{k,k-1}\boldsymbol{h}_k, \quad r_k = \sigma_k^2 + \boldsymbol{h}_k^{\mathrm{T}}\boldsymbol{b}_k⒀ \end{cases}$$

则 $\boldsymbol{X}_k$ 的最小方差滤波可按式(35) 递推求得⒀

很明显,若在式(36) 中增加与 $\boldsymbol{w}_k$ 无关的 r 维动态 Gauss 白噪音 $\boldsymbol{v}_k \sim N(0, \boldsymbol{s}_k)$,$\boldsymbol{s}_k = Var(\boldsymbol{v}_k) \in R^{r\times r}$ 为正定方差阵,即考虑系统

$$\left.\begin{aligned} &\boldsymbol{X}_k = \boldsymbol{F}_{k,k-1}\boldsymbol{X}_{k-1} + \boldsymbol{u}_{k-1} + \boldsymbol{G}_k\boldsymbol{v}_k, \\ &\boldsymbol{y}_k = \boldsymbol{h}_k^{\mathrm{T}}\boldsymbol{X}_k + \boldsymbol{w}_k⒀ \end{aligned}\right\} \tag{38}$$

其中 $\boldsymbol{G}_k \in R^{n\times r}$,则只要将算法 3 中的 $\boldsymbol{P}_{k,k-1} = \boldsymbol{F}_{k,k-1}\boldsymbol{P}_{k-1,k-1}\boldsymbol{F}_{k,k-1}^{\mathrm{T}}$ 修改为 $\boldsymbol{P}_{k,k-1} = \boldsymbol{F}_{k,k-1}\boldsymbol{P}_{k-1,k-1}\boldsymbol{F}_{k,k-1}^{\mathrm{T}} + \boldsymbol{G}_k\boldsymbol{s}_k\boldsymbol{G}_k^{\mathrm{T}}$,就可获得式(38) 中 $\boldsymbol{X}_k$ 的最优滤波,即有著名的 Kalman 滤波算法:

算法 4　令　$\hat{\boldsymbol{X}}_{0,0} = E(\boldsymbol{X}_0)$,　$\boldsymbol{P}_{0,0} = Var(\boldsymbol{X}_0)$,

$$
\begin{aligned}
&\hat{\boldsymbol{X}}_{k,k-1} = \boldsymbol{F}_{k,k-1}\hat{\boldsymbol{X}}_{k-1,k-1} + \boldsymbol{U}_{k-1},\\
&\boldsymbol{P}_{k,k-1} = \boldsymbol{F}_{k,k-1}\boldsymbol{P}_{k-1,k-1} + \boldsymbol{F}^{\mathrm{T}}_{k,k-1} + \boldsymbol{G}_k S_k \boldsymbol{G}_k^{\mathrm{T}},\\
&\boldsymbol{b}_k = \boldsymbol{P}_{k,k-1}\boldsymbol{h}_k,\quad r_k = \sigma_k^2 + \boldsymbol{h}_k^{\mathrm{T}}\boldsymbol{b}_k
\end{aligned}
$$⒀

则系统(38) 中 X_k 的最小方差滤波可按式(35) 递推地求得⒀

3 指数加权情形

在算法 4 中,为获得最小方差估计,必须首先知道噪音的统计特性,但在许多实际应用中是难以做到的,这时人们为使估计算法具有更好的适应性,往往宁可舍弃滤波的最优性,一种常采用的增强适应性的办法是引进指数加权,于是在式(2) 和(9) 中,加权阵 Λ_k 可选为

$$\Lambda_k = \mathrm{diag}(\lambda^{-1},\ \lambda^{-2},\ \cdots,\ \lambda,\ 1) \in R^{k\times k} \tag{39}$$

其中 $\lambda\in(0,1]$ 称为遗忘因子,这时的最小二乘估计称为渐消记忆最小二乘估计⒀

类似于前面的推导,对应于算法 1,我们有

算法 5　给定 $\hat{\boldsymbol{X}}_0$ 和 $\boldsymbol{P}_0$,并令 $\boldsymbol{b}_k = \boldsymbol{P}_{k-1}\boldsymbol{h}_k$, $r_k = \lambda + \boldsymbol{h}_k^{\mathrm{T}}\boldsymbol{b}_k$,则式(1) 中 $\boldsymbol{X}$ 的渐消记忆最小二乘估计可如下递推地求得

$$
\left.
\begin{aligned}
&\boldsymbol{g}_k = \boldsymbol{b}_k r_k^{-1},\\
&\hat{\boldsymbol{X}}_k = \hat{\boldsymbol{X}}_{k-1} + \boldsymbol{g}_k(\boldsymbol{y}_k^{\mathrm{T}} - \boldsymbol{h}_k^{\mathrm{T}}\hat{\boldsymbol{X}}_{k-1}),\\
&\boldsymbol{P}_k = (\boldsymbol{P}_{k-1} - \boldsymbol{g}_k\boldsymbol{b}_k^{\mathrm{T}})\lambda^{1}
\end{aligned}
\right\} \tag{40}
$$⒀

同样,对应于算法 2,我们有

算法 6　对于系统(17) 和(18),若给定 $\hat{\boldsymbol{X}}_{0,0}$,$\boldsymbol{P}_{0,0}$,并令

$$
\left\{
\begin{aligned}
&\hat{\boldsymbol{X}}_{k,k-1} = \boldsymbol{F}_{k,k-1}\hat{\boldsymbol{X}}_{k-1,k-1} + \boldsymbol{U}_{k-1},\\
&\boldsymbol{P}_{k,k-1} = \boldsymbol{F}_{k,k-1}\boldsymbol{P}_{k-1,k-1}\boldsymbol{F}^{\mathrm{T}}_{k,k-1},\\
&\boldsymbol{b}_k = \boldsymbol{P}_{k,k-1}\boldsymbol{h}_k,\quad r_k = \lambda + \boldsymbol{h}_k^{\mathrm{T}}\boldsymbol{b}_k
\end{aligned}
\right.
$$⒀

则系统状态矩阵 $\boldsymbol{X}_k$ 的渐消记忆最小二乘滤波可如下递推地求得

$$
\left.
\begin{aligned}
&\boldsymbol{g}_k = \boldsymbol{b}_k r_k^{-1},\\
&\hat{\boldsymbol{X}}_{k,k} = \hat{\boldsymbol{X}}_{k,k-1} + \boldsymbol{g}_k(\boldsymbol{y}_k^{\mathrm{T}} - \boldsymbol{h}_k^{\mathrm{T}}\hat{\boldsymbol{X}}_{k,k-1}),\\
&\boldsymbol{P}_{k,k} = (\boldsymbol{P}_{k,k-1} - \boldsymbol{g}_k\boldsymbol{b}_k^{\mathrm{T}})\lambda^{1}
\end{aligned}
\right\} \tag{41}
$$⒀

渐消记忆滤波算法以前也有人做过深刻的研究,但这时所得算法 6 似乎是一新结果⒀

4 存在问题

由前述推导可知算法 1 是算法 2～6 的基础,因此我们着重分析算法 1⒀

在推导算法 1 时为能利用引理 1 和引理 2,我们做了 $\boldsymbol{H}_{k-1}$ 为列满秩的重要假设,这等于设

$$\boldsymbol{R}_{k-1} = \boldsymbol{H}^{\mathrm{T}}_{k-1}\boldsymbol{H}_{k-1} \tag{42}$$

为非异阵,在此假设下,递推关系(16)是严格成立的[*]。但这一假设并不总成立,例如当 $\boldsymbol{H}_{k-1}$ 的行数少于其列数即 $k<n$ 时,$\boldsymbol{H}_{k-1}$ 不是列满秩的,从而算法 1 所得不是最优化问题(2)或(8)的解

$$\hat{\boldsymbol{X}}_k=(\boldsymbol{H}_k^{\mathrm{T}}\Lambda_k\boldsymbol{H}_k)^{-1}\boldsymbol{H}_k^{\mathrm{T}}\Lambda_k\boldsymbol{Y}_k。\tag{43}$$

而是最优化问题

$$\min_{\boldsymbol{X}}\|\boldsymbol{H}_k\boldsymbol{X}-\boldsymbol{Y}_k\|^2_{\Lambda_k}+\|\boldsymbol{X}-\hat{\boldsymbol{X}}_0\|^2_{\boldsymbol{P}_0}\tag{44}$$

的解

$$\boldsymbol{X}_k=(\boldsymbol{P}_0^{-1}+\boldsymbol{H}_k^{\mathrm{T}}\Lambda_k\boldsymbol{H}_k)^{-1}(\boldsymbol{P}_0^{-1}\hat{\boldsymbol{X}}_0+\boldsymbol{H}_k^{\mathrm{T}}\Lambda_k\boldsymbol{Y}_k)。\tag{45}$$

因此,严格地说,算法 1 并不能求得 X 的最小二乘估计,特别当 $\boldsymbol{H}_k^{\mathrm{T}}\Lambda_k\boldsymbol{H}_k$ 为病态矩阵时,$\boldsymbol{D}_0^{-1}$ 和 $\hat{\boldsymbol{X}}_0$ 所产生的摄动可能会使式(45)中的 $\boldsymbol{X}_k$ 严重偏离式(43)所给最小二乘估计 $\hat{\boldsymbol{X}}_k$,但一般随着 k 的增大,$\boldsymbol{P}_0^{-1}$ 和 $\hat{\boldsymbol{X}}_0$ 对 $\hat{\boldsymbol{X}}_k$ 的不良影响会相对减弱,这从统计学的原理来看是容易理解的,正因如此,本文所有算法仍具有广泛的应用价值。

文献[2]中指出,虽然理论上任何初始估计都可采用,但实际上初始估计设置不好会引起强烈的暂态过程,因此,一般都有必要选择好的初始估计,然而这往往是难以做到的。

另外,将式(45)和(43)相比,似乎可将初始条件设为 $\hat{\boldsymbol{X}}_0=0,\boldsymbol{P}_0=cI$,其中常数 c 越大越好,然而实际情况远非如此简单,而应根据具体情况小心选择 c 的值,若选择不当,则可能产生与一个野值相似的不良后果,而要在递推中抵偿其影响,可能要费相当长的时间,这对于一些工程应用是难以接受的。

Kalman 滤波算法(例如算法 4)也存在类似问题。

此外,对离散线性随机系统,在应用 Kalman 滤波公式时,需用到初值 $\boldsymbol{X}_0$ 的期望 $E(\boldsymbol{X}_0)$ 和方差阵 $Var(\boldsymbol{X}_0)$,但许多实际问题中,这些初始统计特性往往是不知道的,所以这时应用 Gauss-Markov 估计更理想,但若人为地取 $E(\boldsymbol{X}_0)=0,Var(\boldsymbol{X}_0)=cI$ 并将 c 取得相当大,则滤波方程的解 $\boldsymbol{X}_{k,k}(c)$ 不一定再是无偏估计,更谈不上是最小方差估计,滤波误差的协方差阵所满足的方程的解 $\boldsymbol{P}_{k,k}(c)$ 也不再是它的估计误差协方差阵。

可以断言,Kalman 滤波算法所给不是最优化问题(19)或(21)的解,而若令 $\boldsymbol{P}_{0,0}=cI$ 且 $c\rightarrow\infty$ 时,虽然 $\boldsymbol{X}_{k,k}(c)$ 趋向于 $\boldsymbol{X}_k$ 的 Gauss-Markov 估计,但同时也存在着数值上的困难,作者也曾以具体例子说明,c 太大可能导致 Kalman 滤波精度的降低[3]。

---

[*] 许多人认为只有 R_{k-1} 非异时式(16)才成立,但作者已证明,R_{k-1} 非异是式(16)成立的充分条件而非必要条件。

参 考 文 献

1 Ben-Israel; T N E Greville, Generalized Inverses: Theory and Applications. John wiley & sons, New York, 1974, 218～226

2 Goodwin G C, K S Sin. Adaptive filtering Prediction and Control, Prentice-Hall, Inc, 1984, 139～148

3 D Liu, X Liu Optimal state Estimation without the Requirement of a Priori statistics information of the initial state, IEEE, Trans. 1994, (39): 2087～2091

The Algorithms of Recursive Least-Square and Kalman Filtering

Liu Xuanhuang

(Jiangxi Electrical Power University for Staff)

Abstract This paper presents the well-known RLS and Kalman filtering algorithms. The concept of state vector is first extended to the new concept of state matrix, which makes it possible to use models with much lower order in some applications without causing any additional error in the models. At the end of this paper, some problems in the RLS and Kalman filtering algorithms are mentioned.

Key words least-square; recursive algorithm; filtering algorithm; weight; optima; index