Vol. 2, No. 2, 1982

数学研究与评论
JOURNAL OF MATHEMATICAL RESEARCH & EXPOSITION

Kalman 滤波的一种改型*

程 极 泰

（上海交通大学）

对于 n 维状态 $x\in R^n$ 与不完全观测 $y\in R^m$（$m<n$）的随机离散系统，在给定干扰 $u\in R^r$，$v\in R^m$ 后恒具零均值独立高斯过程的情况下，状态与观测的系统方程在 t_k，t_{k+1} 时的形式是：

$$x_{k+1}=\Phi_k x_k+c_k+G_k u_k, \tag{1}$$

$$y_k=H_k x_k+u_k, \tag{2}$$

其中 $c_k\in R^n$ 是外加独立项．熟知有 Kalman 滤波[1]：

［1］时间修正的状态预测与误差协方差估计

$$\hat{x}_{k/k-1}=\Phi_{k-1}\hat{x}_{k-1/k-1}+c_{k-1}, \tag{3}$$

$$P_{k/k-1}=\Phi_{k-1}P_{k-1/k-1}\Phi_{K-1}^T+G_{k-1}Q_{k-1}G_{K-1}^T. \tag{4}$$

［2］观测修正的状态滤波，加权与误差协方差估计

$$\hat{x}_{k/k}=\hat{x}_{k/k-1}+A_k(y_k-H_k\hat{x}_{k/k-1}), \tag{5}$$

$$A_k=P_{k/k-1}H_K^T(H_kP_{k/k-1}H_K^T+R_k)^{-1}, \tag{6}$$

$$P_{k/k}=(I-A_kH_k)P_{k/k-1}. \tag{7}$$

由于在 Kalman 滤波(3)～(7)的二组共五个计算式中，第一组时间修正的(3)(4)二式是带有本质性的，而观测系统给出的第二组修正，即使在简单的情况下，也将带来复杂的计算．如果我们在离散模型(1)(2)中，将时间步长 $\Delta t_k=t_{k+1}-t_k$ 分得相当小，则外加随机观测干扰可以忽略，把 y 与 x 的关系看成接近连续变化的关系，就可以将动态修正(5)(6)(7)改用静态迭代的最优估计来处理．并且，相应地用类似方法估计观测误差协方差阵．从结果可以看出，误差是小的，但计算确可以简化．

先将(5)式中的权矩阵 $A_k\in R^{n\times m}$ 改用某矩阵 $B_k\in R^{m\times n}$ 能使 $B_K^TB_k$ 非奇的它的伪逆

$$B_K^+=(B_K^TB_k)^{-1}B_K^T\in R^{n\times m}. \tag{8}$$

* 1981年3月23日收到．

这时 $B_K^+ \in A\{1,2,3,4\}$，$A\{1,2,3,4\}$表示所有使

$$AXA = A$$

$$XAX = X$$

$$(AX)^{\mathrm{T}} = AX$$

$$(XA)^{\mathrm{T}} = XA$$

有唯一阵X的Moore-Penrose逆集合。表成(5)为

$$x_{k+} = \hat{x}_{k/k-1} - B_K^{\mathrm{T}}(H_k\hat{x}_{k/k-1} - y_k). \tag{9}$$

它可以从前一阶段的预测值之差按拟牛顿(Quasi-Newton)方程给出决定 B_k 的关系是

$$B_k(\hat{x}_{k/k-1} - \hat{x}_{k-1/k-2}) = (H_k\hat{x}_{k/k-1} - y_k) - (H_{k-1}\hat{x}_{k-1/k-2} - y_{k-1}), \tag{10}$$

或简写为

$$B_k\Delta\hat{x}_k = \Delta\tilde{y}_k\ . \tag{11}$$

于是仿照1977年Dennio与Moré的工作[2]，可以给出

定理1　给定 $B\in \mathrm{R}^{m\times n}$，$y\in \mathrm{R}^m$ 以及非零的 $s\in \mathrm{R}^n$，定义

$$\bar{B} = B + \frac{(y - Bs)s^{\mathrm{T}}}{\langle s, s\rangle}, \tag{12}$$

则$\bar{B}$是问题

$$\min_{\hat{B}}\{\|\hat{B} - B\|_F;\ \hat{B}s = y\} \tag{13}$$

的唯一解，F 表示Frobenius范数。

证明　因为，用 $y = \hat{B}s$ 代入(12)就有

$$\|\bar{B} - B\|_F = \left\|\frac{(y - Bs)s^{\mathrm{T}}}{\langle s, s\rangle}\right\|_F = \left\|(\hat{B} - B)\frac{ss^{\mathrm{T}}}{\langle s, s\rangle}\right\|_F$$

$$\leqslant \|\hat{B} - B\|_F\left\|\frac{ss^{\mathrm{T}}}{\langle s, s\rangle}\right\|_F = \|\hat{B} - B\|_F,$$

其中，因为

$$\left\|\frac{ss^{\mathrm{T}}}{\langle s, s\rangle}\right\|_F = \frac{1}{\langle s, s\rangle}\|ss^{\mathrm{T}}\|_F = \frac{1}{\langle s, s\rangle}[tr((ss^{\mathrm{T}})(ss^{\mathrm{T}}))]^{\frac{1}{2}}$$

$$= \frac{1}{\langle s, s\rangle}[\langle s, s\rangle tr(ss^{\mathrm{T}})]^{\frac{1}{2}} = \frac{1}{\langle s, s\rangle}[\langle s, s\rangle^2]^{\frac{1}{2}} = 1$$

并且由于 $f(A) = \|B - A\|_F$ 是在 $\mathrm{R}^{m\times n}$ 中严格凸，故 $\bar{B}$ 是唯一解。　〈证毕〉

现在，利用(12)的结果给出(11)中 B_k 的最优递推关系：

$$B_{k+1} = B_k + \frac{(\Delta\tilde{y}_k - B_k\Delta\hat{x}_k)\Delta\hat{x}_K^{\mathrm{T}}}{\langle\Delta\hat{x}_k, \Delta\hat{x}_k\rangle}\ . \tag{14}$$

其次，可以从(14)给出 B_{k+1}^{+} 与 B_k^{+} 的直接递推关系：

引理 1 设 $u\in\mathrm{R}^m$，$v\in\mathrm{R}^n$，且设 $B\in\mathrm{R}^{m\times n}$，$B^{\mathrm{T}}B$ 非奇，则当且仅当

$$\sigma=1+\langle v,\ B^{+}u\rangle\neq 0 \tag{15}$$

时，$B+uv^{\mathrm{T}}$ 有伪逆，且

$$(B+uv^{\mathrm{T}})^{+}=B^{+}-\frac{1}{\sigma}B^{+}uv^{\mathrm{T}}B^{+}. \tag{16}$$

证明 为使

$$(B+uv^{\mathrm{T}})(B+uv^{\mathrm{T}})^{+}(B+uv^{\mathrm{T}})=(B+uv^{\mathrm{T}}),$$

$$(B+uv^{\mathrm{T}})^{+}(B+uv^{\mathrm{T}})(B+uv^{\mathrm{T}})^{+}=(B+uv^{\mathrm{T}})^{+},$$

只要证 $(B+uv^{\mathrm{T}})^{+}(B+uv^{\mathrm{T}})=I$，现在从(16)右边得

$$\begin{aligned}
&(B^{+}-\frac{1}{\sigma}B^{+}uv^{\mathrm{T}}B^{+})(B+uv^{\mathrm{T}})\\
&=B^{+}(B+uv^{\mathrm{T}})-\frac{1}{\sigma}B^{+}uv^{\mathrm{T}}B^{+}(B+uv^{\mathrm{T}})\\
&=B^{+}B+B^{+}uv^{\mathrm{T}}-\frac{1}{\sigma}B^{+}uv^{\mathrm{T}}B^{+}B-\frac{1}{\sigma}B^{+}uv^{\mathrm{T}}B^{+}uv^{\mathrm{T}}\\
&=I+B^{+}uv^{\mathrm{T}}-\frac{1}{\sigma}B^{+}uv^{\mathrm{T}}(1+v^{\mathrm{T}}B^{+}u)\\
&=I+B^{+}uv^{\mathrm{T}}-B^{+}uv^{\mathrm{T}}=I,
\end{aligned}$$

其中，引用了 $B^{+}=(B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}$，$B^{+}B=I$。 〈证毕〉

定理 2 设 $B_k^{\mathrm{T}}B_k$，$B_{k+1}^{\mathrm{T}}B_{k+1}$ 非奇，且 B_k，B_{k+1} 是合于(14)式的 $m\times n$ 阵，则

$$B_{k+1}^{+}=B_k^{+}+\frac{(\Delta\hat{x}_k-B_k^{+}\Delta\tilde{y}_k)(\Delta\hat{x}_k)^{\mathrm{T}}B_k^{+}}{\langle\Delta\hat{x}_k,B_k^{+}\Delta\tilde{y}_k\rangle}. \tag{17}$$

证明 从引理 1，用(16)，可将(14)写成

$$B_{k+1}^{+}=B_k^{+}-\frac{1}{\sigma}B_k^{+}\frac{(\Delta\tilde{y}_k-B_k\Delta\hat{x}_k)(\Delta\hat{x}_k)^{\mathrm{T}}}{\langle\Delta\hat{x}_k,\Delta\hat{x}_k\rangle}B_k^{+}, \tag{18}$$

其中

$$\sigma=1+(\Delta\hat{x}_k)^{\mathrm{T}}B_k^{+}\left(\frac{\Delta\tilde{y}_k-B_k\Delta\hat{x}_k}{\langle\Delta\hat{x}_k,\Delta\hat{x}_k\rangle}\right)=\frac{(\Delta\hat{x}_k)^{\mathrm{T}}B_k^{+}(\Delta\tilde{y}_k)}{\langle\Delta\hat{x}_k,\Delta\hat{x}_k\rangle},$$

故有

$$\begin{aligned}
B_{k+1}^{+}&=B_k^{+}+\frac{(\Delta\hat{x}_k-B_k^{+}\Delta\tilde{y}_k)(\Delta\hat{x}_k)^{\mathrm{T}}B_k^{+}}{\langle\Delta\hat{x}_k,\Delta\hat{x}_k\rangle}\ \frac{\langle\Delta\hat{x}_k,\Delta\hat{x}_k\rangle}{(\Delta\hat{x}_k)^{\mathrm{T}}B_k^{+}(\Delta\tilde{y}_k)}\\
&=B_k^{+}+\frac{(\Delta\hat{x}_k-B_k^{+}\Delta\tilde{y}_k)(\Delta\hat{x}_k)^{\mathrm{T}}B_K^{+}}{\langle\Delta\hat{x}_k,B_k^{+}\Delta\tilde{y}_k\rangle}.
\end{aligned}$$

〈证毕〉

我们将 A_k 权转用 B_K^{+} 代替，计算有递推直接结果(17)，相应的误差协方差修正还可按 Frobenius 范数最优准则线性解出，因为现在若令

$$\Delta P_k = P_{k+} - P_{k/k-1} = B_k^{\dagger} s_k^{\dagger}. \tag{19}$$

设有 s_k 使 $s_k^T s_k$ 非奇，且 $s_k^{\dagger} = (s_k^T s_k)^{-1} s_k^T$，则应从

$$\Delta P_k s_k = B_k^{\dagger} \tag{20}$$

容易看出，若取 $B_k^{\dagger}$ 是 A_k，则$s_k^{\dagger}$ 就是严格地为

$$-H_k P_{k/k-1}.$$

反之，若令

$$s_k^{\dagger} = -H_k P_{k/k-1} \tag{21}$$

或其它 $s_k^{\dagger}$ 阵，相应的 ΔP_k 最优范数(确定性)解是可以仿1980年 Salene 与 Tewarson 的方法给出[3]：

定理 3 使误差协方差阵增量 ΔP_k 满足(20)，在 $s_k^T s_k$ 非奇时的 Frobenius 范数最优解是 P_{k+} 满足 Riccati 代数方程形式：

$$P_{k+} = P_{k/k-1} + B_k^{\dagger} s_k^{\dagger} + (B_K^{\dagger} s_k^{\dagger})^T - (s_k^T)^+ (B_k^{\dagger})^T s_k s_k^{\dagger}, \tag{22}$$

证明 由于$(B_k^{\dagger} s_k^{\dagger}) s_k = B_k^{\dagger}$，所以 $B_k^{\dagger} s_k^{\dagger}$ 是(20)的 ΔP_k 一解，这与(19)的设定一致，即

$$(B_k^{\dagger} s_k^{\dagger}) s_k s_k^{\dagger} = B_k^{\dagger} s_k^{\dagger} \Rightarrow s_k^{\dagger} s_k s_k^{\dagger} = s_k^{\dagger}.$$

所以，(20)的一般解是

$$\Delta P_k = B_k^{\dagger} s_k^{\dagger} + z(I - s_k s_K^{\dagger}),\quad z \in \mathrm{R}^{n\times n}. \tag{23}$$

再从 ΔP_k 的对称性要求，就有(20)式的另一形式

$$s_k^T \Delta P_k = (B_k^{\dagger})^T. \tag{24}$$

所以，(23)中的 z 应满足

$$s_k^T z(I - s_k s_k^{\dagger}) = (B_k^{\dagger})^T - s_k^T B_k^{\dagger} s_k^{\dagger}. \tag{25}$$

由于

$$B_k^{\dagger} s_k^{\dagger} s_k = B_k^{\dagger} \Rightarrow (B_k^{\dagger})^T = (s_k^T)(s_k^{\dagger})^T (B_k^{\dagger})^T,$$

$$s_k^T B_k^{\dagger} s_k^{\dagger} = (s_k^{\dagger} s_k)^T s_k^T B_k^{\dagger} s_k^{\dagger},$$

此中

$$s_k^{\dagger} = (s_k^T s_k)^{-1} s_k^T.$$

所以(25)的右端是

$$\begin{aligned}&(B_k^{\dagger})^T - s_k^T B_k^{\dagger} s_k^{\dagger}\\ &= (s_k^T)(s_k^{\dagger})^T((B_k^{\dagger})^T - s_k^T (B_k^{\dagger}) s_k^{\dagger})(I - s_k s_k^{\dagger})^+ (I - s_k s_k^{\dagger}).\end{aligned}$$

与(25)左端比较，可知 z 解是

$$z = (s_k^T)^+ ((B_k^{\dagger})^T - s_k^T B_k^{\dagger} s_k^{\dagger})(I - s_k s_k^{\dagger})^+ + (I - s_k s_k^{\dagger}) Y,\quad Y \in \mathrm{R}^{n\times n}. \tag{26}$$

代入(23)得

$$\begin{aligned}\Delta P_k = {}& B_k^{\dagger} s_k^{\dagger} + (B_k^{\dagger} s_k^{\dagger})^T - (s_k^T)^+ (B_k^{\dagger})^T s_k s_k^{\dagger}\\ &+ (I - s_k s_k^{\dagger}) Y (I - s_k s_k^{\dagger}).\end{aligned} \tag{27}$$

今限制取Y使

$$(I-s_k s_k^+)Y(I-s_k s_k^+)$$

是对称的，(27)给出的 ΔP_k 就是对称的，因为从(20)，(24)各有

$$(\Delta P_k)s_k = B_k^+ \Rightarrow s_k^{\mathrm T}B_k^+ = s_k^{\mathrm T}(\Delta P_k)s_k,$$

$$s_k^{\mathrm T}\Delta P_k = (B_k^+)^{\mathrm T} \Rightarrow (B_k^+)^{\mathrm T}s_k = s_k^{\mathrm T}\Delta P_k s_k.$$

(27)中第三项内应有

$$s_k^{\mathrm T}B_k^+ = (B_k^+)^{\mathrm T}s_k.$$

如果将(27)缩写为

$$\Delta P_{k,Y} = B_k^+ s_k^+ + z_Y(I-s_k s_k^+), \tag{28}$$

它就给出

$$\|\Delta P_{k,Y}\|_F^2 = \|B_k^+ s_k^+\|_F^2 + \|z_Y\|_F^2. \tag{29}$$

再从 $\|z_Y\|_F = \|z_Y^{\mathrm T}\|_F$ 就得

$$\|z_Y\|_F^2 = \|(s_k^+)((B_k^+)^{\mathrm T} - s_k^{\mathrm T}B_k^+ s_k^+)(I-s_k s_k^+)\|_F^2$$

$$+\|(I-s_k s_k^+)Y(I-s_k s_k^+)\|_F^2. \tag{30}$$

从(29)、(30)得知问题

$$\min\{\|\Delta P_k\|_F : (\Delta P_k)s_k = B_k^+,\ (\Delta P_k)^{\mathrm T} = \Delta P_k\} \tag{31}$$

的唯一解是

$$\Delta P_k = B_k^+ s_k^+ + (B_k^+ s_k^+)^{\mathrm T} - (s_k^{\mathrm T})^+(B_k^+)^{\mathrm T}s_k s_k^+, \tag{32}$$

它就是(22)式。　　〈证毕〉

定理 4　若在(31)中取 $s_K^+ = -H_k P_{k/k-1}$（即(21)式），则协方差修正(22)式与真实协方差修正之差可从 P_{k+} 与 $P_{k/k}$ 估出，其中

$$P_{k/k} = P_{k/k-1} - B_K^+ H_k P_{k/k-1} - P_{k/k-1}H_K^{\mathrm T}(B_K^+)^{\mathrm T}$$

$$+ B_k^+(H_k P_{k/k-1}H_k^{\mathrm T} + R_k)(B_k^+)^{\mathrm T} \tag{33}$$

证明　因为

$$P_{k/k} = E[(x_k - x_{k+})(x_k - x_{k+})^{\mathrm T}] = (I - B_k^+ H_k)P_{k/k-1}(I - B_k^+ H_k)^{\mathrm T} + B_k^+ R_k (B_k^+)^{\mathrm T}$$

$$= P_{k/k-1} - B_k^+ H_k P_{k/k-1} - P_{k/k-1}H_K^{\mathrm T}(B_K^+)^{\mathrm T}$$

$$+ B_k^+(H_k P_{k/k-1}H_k^{\mathrm T} + R_k)(B_k^+)^{\mathrm T},$$

其中首三项是主要部分，它与 P_{k+} 的首三项完全相同，只是最末项变动。　〈证毕〉

参考文献

[1] Barham, P. M. and Humphries, D. E., Derivation of the kalman Filtering Equations from Elementary Statistical principles, A. D. —704306, 1970.

[2] Dennis JR J. D. and Morê, Jorge J., Quasi-Newton methods, Motivation and Theory, *SIAM Review*, Vol. 19, No. 1, 1977, pp. 46~89.

[3] Salane, D. E. and Tewarson, R. P., A unified Derivation of symmetric Quasi-Newton update Formulas, *J. Inst. Maths. Applics.* Vol. 25, No. 1, 1980, pp. 29~36.

A Modification of Kalman's Filter

By Cheng Chitai (程枢泰)

Abstract

At the filter formula of the kalman recursive algorithm, optimal weight A_k is the important factor. We choose this factor that may minimize the error covarience, but calculate A_k to be rather complex.

In this paper, we use optimal weight B_k^+ in sense of minimizing the Frobenious norm with constraint that the weight update must satisfy the quasi-Newton linear equation at instant time. We also calculate the correspond. optimal error covariance estimate P_{k+} by means of minimizing the Frobenious norm with constraint that P_{k+} satisfies the linear relation with $P_{k/k-1}$.

From above results, we can compare the P_{k+} to the $P_{k/k}$ if the incomplete observe equation is linear with added white noise. We can find that Ricatti equation of error covarience has the important role at the linear filter whatever we use A_k or B_K^+ as the optimal weight.