国防科技大学学报
JOURNAL OF NATIONAL UNIVERSITY OF DEFENSE TECHNOLOGY
一九八五年第二期　总第五十期　No.2 1985 Sum. 50

非线性滤波的迫近

张金槐

摘　要　本文研究动力学系统状态的估计问题。文中给出了非线性估计的两种迫近方法，一种是运用正态迫近的Gram-Charlier级数展开方法；另一种是二阶截断统计迫近方法。为了便于计算，文中给出了递推形式的近似解。人们常用的广义Kalman滤波方法及统计线性化方法仅为本文的特例。

一、问题的提出

对于线性系统的状态估计，如果仅限于线性估计，那末最小方差估计已得到较充分的讨论。但是，当考虑到一般的最优估计时，就不仅仅限于线性估计了。这种非线性估计问题较之线性估计要复杂得多，对于非线性系统就更是如此。

我们知道，对于时变线性系统的非线性估计问题，如果动力学模型噪声和观测模型噪声之一是正态的，那么可以获得状态的最小方差估计的某种表示[1]，但是，如果去除这种正态的假设，就将带来不少困难。为此，我们设法运用熟知的Gram-Charlier迫近（或者Edgeworth迫近），去获得非线性滤波的近似解。同时，将此方法运用于状态的验后密度的迫近，以获得非线性滤波的极大验后滤波的迫近解。

对于非线性系统的滤波，人们习惯于应用Taylor展开，从而运用线性化或者高阶截断的方法。但是，如果非线性函数不具有导数，例如继电系统，那么上述方法遇到了困难，为此，我们设法运用统计迫近方法。这种方法，在统计学上沒有任何新奇之处，但是要把它应用于多維向量函数的迫近，而且想在滤波公式中以递推的形式出现，那么，也还会遇到计算高阶矩的困难以及表达形式上的困难。下面，我们将在某些给定的条件下讨论上面提出的问题。可以看到，人们熟知的广义Kalman滤波以及统计线性化方法仅为我们的一个特例。

用非线性估计的方法去改善滤波的性能，将使计算工作量增大，不过，运用近代计算工具，非线性滤波的实现也还不是困难的。

二、MV估计的正态迫近

当我们进行统计滤波或估值时，常常假定观测噪声和动力学模型噪声是正态的。但

本文曾在一九八四年中南五省自动化学会年会上宣读。

是情况并非总是如此。当上述噪声之一具有正态分布时，则容易给出非线性滤波的一般公式[1]。但具体地实现滤波时，还将遇到计算上的困难。对于非线性估计，要使计算不过于复杂，同时又较之线性估计有一定程度的改善，则还有不少工作可做。这一节，我们先来讨论最小方差估计的逼近问题。

设有动力学模型

$$X_{k+1}=\Phi_{k+1,k}X_k+W_k, \tag{1}$$

其中X_k为$n\times 1$状态向量；$\{W_k\}$为噪声序列，其期望值为0；$E[W_k\cdot W_j^T]=Q_k\delta_{k,j}$，$Q_k\geqslant 0$；

设观测模型为

$$Z_k=H_kX_k+V_k, \tag{2}$$

其中Z_k为m維观测向量；$\{V_k\}$为观测噪声序列，期望值为0；$E[V_kV_j^T]=R_k\delta_{k,j}$，$R_k>0$；$\{W_k\}$与$\{V_k\}$不相关，且它们与初态$X_0$不相关，记

$$Z^k=\{Z_1,\cdots Z_k\},$$

即Z^k为直到t_k时刻的观测集。

在上述情况下，在给定Z^k之下，状态X_k的MV估计为条件数学期望

$$E[X_k/Z^k]\triangleq\hat{X}_{k/k}.$$

如果$\{W_k\}$为正态序列，则有下列结果[2]：

预测值
$$\hat{X}_{k/k-1}=\Phi_{k,k-1}\hat{X}_{k-1/k-1}, \tag{3}$$

条件期望
$$\hat{X}_{k/k}=\hat{X}_{k/k-1}+P_{k/k-1}H_K^Tg_k(\nu_k), \tag{4}$$

其中 $P_{k/k-1}$—预测误差协方差阵：

$$P_{k/k-1}=\Phi_{k,k-1}P_{k-1/k-1}\Phi_{k,k-1}^T+Q_{k-1}. \tag{5}$$

$\hat{X}_{k/k}$的误差协方差阵为

$$P_{k/k}=P_{k/k-1}-P_{k/k-1}H_k^TG_k(\nu_k)H_kP_{k/k-1}, \tag{6}$$

$\{\nu_k\}$为新息序列：

$$\nu_k=Z_k-H_k\hat{X}_{k/k-1},\ k=1,2,\cdots,$$

其期望值为0，且

$$E[\nu_k\cdot\nu_j^T]=(H_kP_{k/k-1}H_k^T+R_k)\delta_{k,j}. \tag{7}$$

$g_k(\nu_k)$为列向量：

$$g_k(\nu_k)=-\frac{\partial}{\partial\nu_k}\log p(\nu_k)\triangleq-\begin{pmatrix}\frac{\partial}{\partial\nu_k^{(1)}}\\ \vdots\\ \frac{\partial}{\partial\nu_k^{(m)}}\end{pmatrix}\log p(\nu_k), \tag{8}$$

$$\nu_k=(\nu_k^{(1)},\ \cdots,\ \nu_k^{(m)})^T,$$

$p(\nu_k)$为新息ν_k之概率密度函数.

$G_k(\nu_k)$为Fisher 信息矩阵：

$$G_k(\nu_k)=E\left[\left(\frac{\partial}{\partial\nu_k}\log p(\nu_k)\right)\left(\frac{\partial}{\partial\nu_k}\log p(\nu_k)\right)^T\right]. \tag{9}$$

在初始状态估计$\hat{X}_0=E[X_0]$之下，上述估计$\hat{X}_{k/k}$是无偏的，且有

$$E[g_k(\nu_k)]=-\int_{R^{(m)}}\left[\frac{\partial}{\partial\nu_k}\log p(\nu_k)\right]p(\nu_k)d\nu_k$$

$$=-\frac{\partial}{\partial\nu_k}\int_{R^{(m)}}p(\nu_k)d\nu_k=0. \tag{10}$$

其中$R^{(m)}$为m維欧氏空间。在具体实施滤波时，由于$g_k(\nu_k)$是未知的，因此难于付之应用。如果对$g_k(\nu_k)$不作任何假设，那么上述非线性滤波的计算是困难的。为此，我们假定：(1)$p(\nu_k)$具有对称分布；(2)对于$p(\nu_k)$的峰度作修正，则$p(\nu_k)$可以近似于正态分布。记ν_k的（协）方差矩阵为Σ_{ν_k}，$\Sigma_{\nu_k}=H_kP_{k/k-1}H_k^T+R_k$，$R_k>0$. Σ_{ν_k}为正定阵，但它不是对角阵，即是说，ν_k的分量$\nu_k^{(1)},\cdots,\nu_k^{(m)}$之间是相关的。作变换

$$\nu_k^*=\Sigma_{\nu_k}^{-\frac{1}{2}}\nu_k,$$

则ν_k^*的方差阵为$E[\nu_k^*\nu_k^{*T}]=I$，这样ν_k^*的各分量$\nu_k^{*(1)},\cdots,\nu_k^{*(m)}$之间是不相关的。记$\nu_k^*$的密度函数为$p(\nu_k^*)$，将它展开为Gram-Charlier级数

$$p(\nu_k^*)\sim\frac{1}{(\sqrt{2\pi})^m|\Sigma_{\nu_k}|^{1/2}}\exp\left\{-\frac{1}{2}\nu_k^{*T}\nu_k^*\right\}.$$

$$\sum_{k_1,\cdots,k_m=0}^{\infty}\frac{b_{k_1\cdots k_m}}{k_1!\cdots k_m!}H_{k_1}(\nu_k^{*(1)})\cdots H_{k_m}(\nu_k^{*(m)}), \tag{11}$$

其中

$$b_{k_1\cdots k_m}=\int_{R^{(m)}}p(\nu_k^*)H_{k_1}(\nu_k^{*(1)})\cdots H_{k_m}(\nu_k^{*(m)})d\nu_k^* \tag{12}$$

$k_1,\cdots,k_m$为零或正整数，每个k_i可取0，1，2，…，∞；

$H_{k_i}(x)$—Hermite多项式：

$$H_k(x)=(-1)^k e^{\frac{x^2}{2}}\frac{d^k}{dx^k}\left(e^{-\frac{x^2}{2}}\right)$$

由于$P(\nu_k^*)$为对称的概率密度函数，于是

$b_{k_1,k_2,\dots,k_m}=0$，当$k_1,\cdots,k_m$中有一个为奇数时；

且

$$\begin{cases} b_{0,0,\cdots,0}=1, \\ b_{2,0,\cdots,0}=b_{0,2,0,\cdots,0}=\cdots=b_{0,0,\cdots,0,2}=0, \\ b_{2,2,0\cdots,0}=\alpha_{2,2,0,\cdots,0}-1, \\ \vdots \\ b_{0,0,\cdots,0,\underset{\uparrow\,i}{2},0,\cdots,0,\underset{\uparrow\,j}{2},0\cdots,0}=\alpha_{0,0,\cdots,0,\underset{\uparrow\,i}{2},0,\cdots,0,\underset{\uparrow\,j}{2},0,\cdots,0}-1, \end{cases}$$

$$\begin{cases} b_{4,0,\cdots,0}=\alpha_{4,0,\cdots,0}-3, \\ b_{0,\cdots,0,\underset{\uparrow\,i}{4},0,\cdots,0}=\alpha_{0,\cdots,0,\underset{\uparrow\,i}{4},0,\cdots,0}-3 \end{cases}$$

其中

$$\alpha_{l_1,\dots,l_m}=E[(\nu_k^{*(1)})^{l_1}\cdots(\nu_k^{*(m)})^{l_m}] \tag{13}$$

注意到 ν_k^* 的分量是不相关的，于是

$$b_{2,2,0,\cdots,0}=b_{0,2,2,\cdots,0}=\cdots=b_{0,0,\cdots,0,2,2}=0.$$

于是，我们有如下近似关系：

$$P(\nu_k^*)\cong\frac{1}{(\sqrt{2\pi})^m}\exp\left(-\frac{1}{2}\nu_k^{*T}\nu_k^*\right)\left[1+\frac{1}{4!}\sum_{i=1}^m b_{0,0,\cdots,0,4_i,0,\cdots,0}H_4(\nu_k^{*(i)})\right] \quad (14)$$

其中，$b_{0,0,\cdots,0,4_i,0,\cdots,0}$ 表示 $b_{0,0,\cdots,0,\overset{i}{4},0,\cdots,0}$，即第 i 个足码为 4，而其他足码统统为 0 的 b。

回复至 ν_k 的密度函数，注意到 $\nu_k^*=\Sigma_{\nu_k}^{-\frac{1}{2}}\nu$，于是 ν_k 的密度函数可表示为

$$\begin{aligned} p_{\nu_k}(\nu_k) &= p(\Sigma_{\nu_k}^{-\frac{1}{2}}\nu_k)\cdot\left|\frac{\partial\nu_k^*}{\partial\nu_k}\right| \\ &= p(\Sigma_{\nu_k}^{-\frac{1}{2}}\nu_k)\quad\left|\Sigma_{\nu_k}^{-\frac{1}{2}}\right| \\ &\approx\frac{1}{(\sqrt{2\pi})^m}\ \frac{1}{|\Sigma_{\nu_k}|^{1/2}}\exp\left[-\frac{1}{2}\nu_k^T\Sigma_{\nu_k}^{-1}\nu_k\right] \\ &\cdot\left[1+\frac{1}{4!}\sum_{i=1}^m b_{0,\cdots,0,4_i,0,\cdots,0}H_4((\Sigma_{\nu_k}^{-\frac{1}{2}})_{i\cdot}\ \nu)\right] \end{aligned} \quad (15)$$

其中 $(\Sigma^{-\frac{1}{2}})_{i\cdot}$ 为矩阵 $\Sigma^{-\frac{1}{2}}$ 的第 i 行组成的行向量。

至于 $g_k(\nu_k)$：

$$g_k(\nu_k)=-\frac{\partial}{\partial\nu_k}[\log p_{\nu_k}(\nu_k)],$$

而

$$\log p_{\nu_k}(\nu_k)=c_k-\frac{1}{2}\nu_k^T\Sigma_{\nu_k}^{-1}\nu_k+\log\left[1+\frac{1}{4!}\sum_{i=1}^m b_{0,\cdots,0,4_i,\cdots,0}H_4((\Sigma_{\nu_k}^{-\frac{1}{2}})_{i\cdot}\nu_k)\right]$$

其中 c_k 为与 ν_k 无关的量。于是

$$\frac{\partial\log p_{\nu_k}(\nu_k)}{\partial\nu_k}=-\Sigma_{\nu_k}^{-1}\nu_k+\mathscr{B}(\Sigma\nu_k,\nu_k)\qquad\begin{pmatrix}\mathscr{H}_4\cdot\Sigma_{\cdot 1}^{-\frac{1}{2}}\\ \cdots\cdots\cdots\\ \vdots\\ \cdots\cdots\cdots\\ \mathscr{H}_4\cdot\Sigma_{\cdot i}^{-\frac{1}{2}}\\ \cdots\cdots\cdots\\ \vdots\\ \cdots\cdots\cdots\\ \mathscr{H}_4\cdot\Sigma_{\cdot m}^{-\frac{1}{2}}\end{pmatrix} \quad (16)$$

其中

$$\mathscr{B}(\Sigma_{\nu_k},\ \nu_k)=\frac{1}{4!}H_4'(\cdot)\Big/\left[\left(1+\frac{1}{4!}\sum_{i=1}^m b_{0,\cdots,4_i,0,\cdots,0}H_4(\cdot)\right)\right] \quad (17)$$

"·"表示$(\Sigma_{\nu_k}^{-\frac{1}{2}})_{i.}\cdot\nu_k$，$H_4'(\cdot)$表示关于此量的导数，

$$\mathscr{H}_4=(b_{4,0,\cdots,0},\ b_{0,4,0,\cdots,0},\cdots,\ b_{0,\cdots,0,4})_{1\times m}$$

$\Sigma_{\cdot j}^{-\frac{1}{2}}$表示矩阵$\Sigma_{\nu_k}^{-\frac{1}{2}}$的第$j$行，而$\Sigma\nu_k=H_kP_{k/k-1}H_k^T+R_k$. 最后，条件数学期望可表示为

$$\hat{X}_{k/k}=X_{k/k-1}+P_{k/k-1}H_k^T\left[\Sigma_{\nu_k}^{-1}\nu_k-\mathscr{B}(\Sigma\nu_k,\ \nu_k)\begin{pmatrix}\mathscr{H}_4\,\Sigma_{\cdot 1}^{-1/2}\\ \vdots\\ \mathscr{H}_4\,\Sigma_{\cdot m}^{-1/2}\end{pmatrix}\right]. \tag{18}$$

要计算出 Fisher 信息量是复杂的。但是，我们知道，$\hat{X}_{k/k}$为 MV 估计，因此它的方差不会超过线性MV估计的方差。于是MV估计的误差协方差阵总有下列上界：

$$P_{k/k}\leqslant P_{k/k-1}-P_{k/k-1}H_k^T(\Sigma\nu_k)^{-1}H_kP_{k/k-1}. \tag{19}$$

推论 1 如果$p_{\nu_k}(\nu_k)\sim N(0,\ \Sigma\nu_k)$，则$b_{k_1,\cdots,k_m}$统统为0此时

$$g_k(\nu_k)=\Sigma_{\nu_k}^{-1}\cdot\nu_k,$$

而

$$\hat{X}_{k/k}=\hat{X}_{k/k-1}+P_{k/k-1}H_k^T(H_kP_{k/k-1}H_k^T+R_k)^{-1}(Z_k-H_k\hat{X}_{k/k-1}).$$

此时的MV估计即为线性MV估计，因此

$$P_{k/k}=P_{k/k-1}-P_{k/k-1}H_k^T(H_kP_{k/k-1}H_k^T+R_k)^{-1}H_kP_{k/k-1}.$$

这是人们熟知的Kalman 滤波公式，也为意料中的事。

推论 2 如果Z_k为一維观测，则

$$\hat{X}_{k/k}=\hat{X}_{k/k-1}+P_{k/k-1}H_k^T(H_kP_{k/k-1}H_k^T+\sigma_{v_k}^2)^{-1}(Z_k-H_k\hat{X}_{k/k-1})$$

$$-P_{k/k-1}H_k^T\frac{d}{d\nu_k}\left\{\log\left[1+\frac{b_4}{4!}H_4\left(\frac{\nu_k}{\sigma_{v_k}}\right)\right]\right\} \tag{20}$$

其中 $b_4=\frac{\alpha_4}{\sigma\nu_k}-3$， $\alpha_4=E[\nu_k^4]$，b_4表示了ν_k的分布的峰度。

因此(20)式右端的第二项表示了关于峰度的修正量。

具体实施滤波计算时，必须给出$\sigma_{v_k}^2$以及$b_{4,0,\cdots,0},b_{0,\cdots,0,4}$，这里$b_{4,0,\cdots,0}=\alpha_{4,0,\cdots}/\sigma_{v_k}^4-3$，$\alpha_{4,0,\cdots,0}=E[(\nu_k^{(1)})^4]$，…. 它们可以通过经典的估计方法给出。如果观测数据具有异常值，则可运用稳健（Robust）统计方法[8]给出。例如在Z_k为一維的場合，

$$\hat{\sigma}\nu_k=\underset{j=0,\cdots,k-1}{\text{median}}|Z_{k-j}-H_{k-j}^T\hat{X}_{k-j/k-j-1}|/0.6745. \tag{21}$$

对于非线性系统的状态的最小方差估计，问题要复杂得多。我们将在后面讨论。这里，我们指出，如果运用线性化方法，则甚易归化为线性模型的情形。事实上，设非线性模型为

$$X_{k+1}=f(X_k,t_k)+W_k,$$
$$Z_k=h(X_k,t_k)+V_k,$$

此处关于$\{W_k\}$，$\{V_k\}$的假设同于线性模型，且f和h是可微分的函数，于是

$$X_{k+1}=f(X_k,t_k)+W_k=f(\hat{X}_{k/k},t_k)+\left(\frac{\partial f}{\partial X_k}\right)_{\hat{X}_{k/k}}(X_k-\hat{X}_{k/k})+\text{高阶项}+W_k,$$

或

$$X_{k+1}=\left(\frac{\partial f}{\partial X_k}\right)_{\hat{X}_{k/k}}X_k+W_k+f(\hat{X}_{k/k},t_k)-\left(\frac{\partial f}{\partial X_k}\right)_{\hat{X}_{k/k}}\cdot\hat{X}_{k/k}$$

$$\triangleq F_kX_k+W_k+q_k,$$

此处

$$F_k=\left(\frac{\partial f}{\partial X_k}\right)_{\hat{X}_{k/k}},\quad q_k=f(\hat{X}_{k/k},\ t_k)-F_k\hat{X}_{k/k}.$$

对于观测模型，将$h(X_k,\ t_k)$在$\hat{X}_{k/k-1}$近旁展开，取一阶项，则有

$$Z_k\cong H_kX_k+V_k+r_k,$$

此处

$$r_k=h(\hat{X}_{k/k-1},\ t_k)-H_k\hat{X}_{k/k-1},$$

$$H_k=\left(\frac{\partial h}{\partial X_k}\right)_{\hat{X}_{k/k-1}}.$$

这样，前面的讨论仍然成立，只需将新息改为

$$\nu_k=Z_k-H_k\hat{X}_{k/k-1}-r_k.$$

而在状态预测方程中，添加q_k就行了。

三、Bayes 极大验后（MAP）估计的追近

考虑下列模型

$$X_{k+1}=\Phi_{k+1,k}X_k+W_k,\tag{22}$$

其中 X_k为n维状态向量；关于W_k的假设和第二节相同。观测模型为

$$Z_k=H_kX_k+v_k\tag{23}$$

此处$\{v_k\}$为独立的观测噪声序列，Z_k为纯量，v_k与W_k互相独立，$K=1,2,\cdots$.

记X_k的MAP估计为$\hat{X}_k$，则它满足

$$p(\hat{X}_k/Z^k)=\text{Max}$$

其中 $p(X_k/Z^k)$为Z^k给定之下的 X_k的条件概率密度函数。

由 Bayes 公式，

$$p(X_k/Z^k)=p(X_k/Z^{k-1},\ Z_k)$$

$$=\frac{p(Z_k/X_k,\ Z^{k-1})\,p(X_k/Z^{k-1})}{p(Z_k/Z^{k-1})}\tag{23}$$

而$p(Z_k/X_k,Z^{k-1})=p(Z_k/X_k)$。这是由于当$X_k$给定之下，$Z_k$的随机性来自$v_k$，而$v_k$为独立观测噪声，因此与$Z^{k-1}$为独立。这样，(24)式成为

$$p(X_k/Z^k)=\frac{p(Z_k/X_k)\,p(X_k/Z^{k-1})}{p(Z_k/Z^{k-1})}\tag{25}$$

于是MAP估计为寻求X_k，使

$$L(X_k)=-\log p(X_k/Z^{k-1})-\log p(Z_k/X_k) \tag{26}$$

取极小。在下面的讨论中，假定：

(1) $p(X_k/Z^{k-1})\sim N(\hat{X}_{k/k-1},\ P_{k/k-1})$;

(2) $p(Z_k/X_k)$为对称分布，且其峰度经过修正后可以近似地用正态分布来迫近。

注意到，$Z_k=H_kX_k+v_k$，$\{v_k\}$为独立观 测噪声序列，不妨 设$E[v_k]=0$。当X_k给定之下，$\sigma^2_{Z_k}=\sigma^2_{v_k}$，$\mu_{Z_k}=E[Z_k]=H_kX_k$，由第二节的讨论，我们有

$$p(Z_k/X_k)\cong\frac{1}{\sqrt{2\pi}\sigma_{Z_k}}e^{-\frac{(Z_k-\mu_{Z_k})^2}{2\sigma^2_{Z_k}}}\left[1+\frac{b_4}{4!}H_4\left(\frac{Z_k-\mu_{Z_k}}{\sigma_{Z_k}}\right)\right] \tag{27}$$

其中

$$b_4=\frac{\alpha_4}{\sigma^4_{Z_k}}-3,\quad \alpha_4=E[\nu_k^4/X_k]$$

若

$$\left|\frac{b_4}{4!}H_4\left(\frac{Z_k-\mu_{Z_k}}{\sigma_{Z_k}}\right)\right|<1$$

则，

$$\log p(Z_k/X_k)\cong-\frac{(Z_k-\mu_{Z_k})^2}{2\sigma^2_{Z_k}}+c_k+\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n}\left[\frac{b_4}{4!}H_4(\cdot)\right]^n.$$

其中c_k为与X_k无关的量，将上式代入(26)，且令

$$L_1(X_k)=\mathrm{Min}.$$

并注意到下式：

$$p(X_k/Z^{k-1})=\frac{1}{(\sqrt{2\pi})^n}\frac{1}{|P_{k/k-1}|^{1/2}}\exp\left[-\frac{1}{2}(X_k-\hat{X}_{k/k-1})^TP_{k/k-1}(X_k-\hat{X}_{k/k-1})\right]$$

于是

$$\frac{\partial}{\partial X_k}L_1(X_k)=P^{-1}_{k/k-1}(X_k-\hat{X}_{k/k-1})+H_k^T\frac{\partial}{\partial Z_k}\log p(Z_k/X_k)=0 \tag{28}$$

而

$$\frac{\delta}{\partial Z_k}\log p(Z_k/X_k)=-\frac{Z_k-\mu_{Z_k}}{\sigma^2_{Z_k}}+\sum_{n=1}^{\infty}(-1)^{n+1}\frac{b_4}{4!}[H_4(\cdot)]'_{Z_k}$$

由(28)，即得

$$\hat{X}^{(MAP)}_{k/k}=\hat{X}_{k/k-1}-P_{k/k-1}H_k^T\frac{\partial\log p(Z_k/\hat{X}^{(MAP)}_{k/k})}{\partial Z_k}, \tag{29a}$$

或

$$\hat{X}^{(MAP)}_{k/k}=\hat{X}_{k/k-1}-\frac{P_{k/k-1}H_k^T}{\sigma_{Z_k}}\left[-\frac{Z_k-\mu_{Z_k}}{\alpha_{Z_k}}+\sum_{n=1}^{\infty}(-1)^{n+1}\frac{b_4}{4!}\frac{d}{dZ_k}H_4(\cdot)\right] \tag{29b}$$

其中

$$\mu_{Z_k}=H_k\hat{X}^{(MAP)}_{k/k},$$

$$b_4=\frac{\alpha_4}{\sigma^4_{v_k}}-3,\quad \alpha_4=E[v_k^4/X_k]$$

(29)为关于$\hat{X}_{k/k}$的非线性方程。因此，可用迭代方法求解。在迭代过程中还必须给出$\hat{X}_{k/k-1}$及$P_{k/k-1}$。我们采用W.S.Agee等人的方法[4]，将(29)中的$\hat{X}_{k/k-1}$用近似的条件期望$\bar{X}_{k/k-1}$代替，此处

$$\bar{X}_{k+1/k}=\Phi_{k+1,k}\bar{X}_{k/k} \tag{30}$$

$$\bar{X}_{k/k}=\bar{X}_{k/k-1}+\frac{P_{k/k-1}H_k^T}{\sigma_{Z_k}}\left(\frac{Z_k-H_k\bar{X}_{k/k-1}}{\sigma_{Z_k}}\right)$$
$$-\frac{P_{k/k-1}H_k^T}{\sigma_{Z_k}}\sum_{n=1}^{\infty}(-1)^{n+1}\frac{b_4}{4!}H_4'\left(\frac{Z_k-H_k\bar{X}_{k/k-1}}{\sigma_{Z_k}}\right) \tag{31}$$

其中$H_4'(\cdot)$为关于"·"的导数，而

$$P_{k/k}=P_{k/k-1}-\left(\psi'\cdot\frac{\nu_k}{\sigma_{Z_k}}\Big/\sigma^2_{Z_k}\right)P_{k/k-1}H_k^TH_kP_{k/k-1}, \tag{32}$$

$$\nu_k=Z_k-H_k\bar{X}_{k/k-1};$$

另外，

$$P_{k+1/k}=\Phi_{k+1,k}P_{k/k}\Phi^T_{k+1,k}+Q_k. \tag{33}$$

(32)中的ψ为

$$\psi\left(\frac{Z_k-H_k\bar{X}_{k/k-1}}{\sigma_{Z_k}}\right)=\frac{Z_k-H_k\bar{X}_{k/k-1}}{\sigma_{Z_k}}-\sum_{n=1}^{\infty}(-1)^{n+1}\frac{b_4}{4!}H_4'\left(\frac{Z_k-H_k\bar{X}_{k/k-1}}{\sigma_{Z_k}}\right) \tag{34}$$

具体迭代过程可如下进行：令$\hat{X}^{(\alpha)}_{k/k}$为迭代中的任一点，则初始点取$\hat{X}^{(0)}_{k/k}=\bar{X}_{k/k-1}$，而

$$\hat{X}^{(\alpha+1)}_{k/k}=\bar{X}_{k/k-1}+\frac{P_{k/k-1}H_k^T}{\sigma_{Z_k}}\psi\left(\frac{Z_k-H_k\hat{X}^{(\alpha)}_{k/k}}{\sigma_{Z_k}}\right). \tag{35}$$

另外，我们也可以用类似于解稳健回归中的迭代加权最小二乘方法。如果$\hat{X}^{(\alpha)}_{k/k}$为迭代序列中的任一点，则

$$\hat{X}^{(\alpha+1)}_{k/k}=\bar{X}_{k/k-1}+W_k^{(\alpha)}P^{(\alpha)}_{k/k}H_k^T/S_k^2\cdot(Z_k-H_k\bar{X}_{k/k-1})$$

其中

$$W_k^{(\alpha)}=\psi\left(\frac{Z_k-H_k\hat{X}^{(\alpha)}_{k/k}}{S_k}\right)\Big/\frac{Z_k-H_k\hat{X}^{(\alpha)}_{k/k}}{S_k} \tag{36}$$

$$P^{(\alpha)}_{k/k}=\left[P^{-1}_{k/k-1}+\frac{W_k^{(\alpha)}}{S_k^2}H_k^TH_k\right]^{-1}. \tag{37}$$

其中$S_k=\hat{\sigma}_{v_k}$如式21所示。

四、非线性滤波的统计逼近

在广义 Kalman 滤波基础上建立起来的二阶滤波，改善了滤波的性能。但是对非线性系统的滤波来说，带来了估值的偏倚。特别要指出的是，当非线性函数不具有导数（例如继电系统），那么用 Taylor 展开取二阶项这种方法将是不可能的，为此，运

用统计逼近方法。我们只需假定随机向量具有较低阶的矩就可以了。这一节，我们就来给出非线性滤波的统计逼近方法。

1. 二阶统计逼近的表示

设$f(x)$为n維状态X的非线性m維向量函数，记$f(X)$的逼近为（近似地取到二阶项）

$$f(X)\cong a_1+A_1X+A_2(X\otimes X) \tag{38}$$

其中$\otimes$为矩阵的 Kronecker 积，记

$$e=f(X)-[a_1+A_1X+A_2(X\otimes X)],$$

$$J=E[\|e\|_A^2]，此处A>0，即A为E定阵。$$

我们用$J=\mathrm{Min}$的方法确定出待定系数a_1，A_1，A_2，注意到

$$J=E[f(X)-a_1-A_1X-A_2(X\otimes X)]^TA[f(X)-a_1-A_1X-A_2(X\otimes X)],$$

令 $\partial J/\partial a_1=0$，则得

$$E[A(f(X)-a_1-A_1X-A_2(X\otimes X))]=0$$

因此

$$a_1=\widehat{f(X)}-A_1\hat{X}-A_2(X\widetilde{\otimes}\hat{X})。 \tag{39}$$

其中$\hat{\cdot}$表示$\cdot$的期望值。将a_1的表达式代入J，则有

$$J=E\{[(f(X)-\widehat{f(X)}))-A_1\tilde{X}-A_2(X\widetilde{\otimes}X)]^TA.$$
$$[(f(X)-\widehat{f(X)}))-A_1\tilde{X}-A_2(X\widetilde{\otimes}X)]\},$$

式中 $\tilde{X}=\hat{X}-X$，$X\widetilde{\otimes}X=\widehat{X\otimes X}-X\otimes X$. 注意$J$中含有$A_1$，$A_2$的项为

$$E\{-(f(X)-\widehat{f(X)}))^TA(A_1\tilde{X}+A_2(X\widetilde{\otimes}X))$$
$$-(A_1\tilde{X})^TA(f(X)-\widehat{f(X)}))-A_1\tilde{X}-A_2(X\widetilde{\otimes}X)]$$
$$-(A_2(X\widetilde{\otimes}X))^TA[(f(X)-\widehat{f(X)}))-A_1\tilde{X}-A_2(X\widetilde{\otimes}X)]\},$$

于是

$$\partial J/\partial A_1=E\{-[\tilde{X}(f(X)-\widehat{f(X)}))^TA]^T$$
$$-A(f(X)-\widehat{f(X)}))\tilde{X}^T+\frac{\partial}{\partial A_1}(\tilde{X}^TA_1^TAA_1\tilde{X})$$
$$+AA_2(X\widetilde{\otimes}X)\tilde{X}^T+[\tilde{X}(A_2(X\widetilde{\otimes}X))^TA]^T\},$$

即

$$\partial J/\partial A_1=E\{-A(f(X)-\widehat{f(X)}))\tilde{X}^T-A(f(X)-\widehat{f(X)}))\tilde{X}^T$$
$$+AA_1\tilde{X}\tilde{X}^T+(\tilde{X}\tilde{X}^TA_1^TA)^T$$
$$+AA_2(X\widetilde{\otimes}X)\tilde{X}^T+AA_2(X\widetilde{\otimes}X)\tilde{X}^T\}$$
$$=A\{-2(\widehat{f(X)}X^T-\widehat{f(X)}\widehat{X^T})+2A_1P+A_2\widehat{(X\widetilde{\otimes}X)\tilde{X}^T}\}$$

令 $\dfrac{\partial J}{\partial A_1}=0$，则得

$$A_1=\left[(\widehat{f(X)X^T}-\widehat{f(X)}\hat{X}^T)-\frac{1}{2}A_2(\widehat{X\tilde{\otimes}X)\tilde{X}^T}\right]P^{-1}. \quad (40)$$

其中

$$P=E[\tilde{X}\tilde{X}^T]. \quad (41)$$

将 A_1 易作 A_2，$\tilde{X}$ 易作 $(X\tilde{\otimes}X)$，则可得 A_2 的表达式为

$$A_2=\left[(\widehat{f(X)X^T}-\widehat{f(X)}\widehat{X^T})-\frac{1}{2}A_1\widehat{\tilde{X}(X\tilde{\otimes}X)^T}\right]P_{\otimes}^{-1}. \quad (42)$$

式中

$$P_{\otimes}=E[(X\tilde{\otimes}X)(X\tilde{\otimes}X)^T]. \quad (43)$$

由式(40)(42)，易得

$$A_1=[\widehat{f(X)X^T}-\widehat{f(X)}\widehat{X^T}]\left[I-\frac{1}{2}P_{\otimes}^{-1}\widehat{(X\tilde{\otimes}X)\tilde{X}^T}\right]P^{-1}$$

$$\cdot\left[I-\frac{1}{4}\widehat{\tilde{X}(X\tilde{\otimes}X)^T}\cdot P_{\otimes}^{-1}\cdot\widehat{(X\tilde{\otimes}X)\tilde{X}^T}\cdot P^{-1}\right]^{-1}. \quad (14)$$

如果忽略高阶矩 $\widehat{\tilde{X}(X\tilde{\otimes}X)^T}P_{\otimes}^{-1}\widehat{(X\tilde{\otimes}X)\tilde{X}^T}P^{-1}$ 不计，则

$$A_1=[\widehat{f(X)X^T}-\widehat{f(X)}\widehat{X^T}]\left[I-\frac{1}{2}P_{\otimes}^{-1}\widehat{(X\tilde{\otimes}X)\tilde{X}^T}\right]P^{-1}. \quad (45)$$

而

$$A_2=\left[(\widehat{f(X)X^T}-\widehat{f(X)}\widehat{X^T})-\frac{1}{2}A_1\widehat{\tilde{X}(X\tilde{\otimes}X)^T}\right]P_{\otimes}^{-1}. \quad (46)$$

这样，我们确定了统计追近之下的系数 a_1，A_1，A_2．此时

$$f(X)\cong a_1+A_1X+A_2(X\otimes X)$$

$$=\widehat{f(X)}+A_1(X-\hat{X})+A_2(X\otimes X-\widehat{X\otimes X}). \quad (47)$$

2. 运用统计追近的二阶截断滤波

非线性系统的二阶截断滤波在正态的場合下已有较多的讨论。如果除去正态性的假设，一般的非线性滤波将遇到计算上的困难。这里将一般地给出统计追近之下的二阶滤波公式。在特殊情况下，我们可以给出递推滤波公式。

考虑离散非线性系统

$$X_{k+1}=f(X_k,\ t_k)+W_k, \quad (48)$$

$$Z_k=h(X_k,\ t_k)+V_k. \quad (49)$$

式中 $\{W_K\}$，$\{V_K\}$ 为零均值的白噪声向量序列。

$$E[W_kW_j^T]=Q_k\delta_{k,j},\ (Q_k\geqslant 0);$$

$$E[V_kV_j^T]=R_k\delta_{k,j}.\ \ (R_k>0);$$

$$E[W_kV_j^T]=0,\ \forall k,j.$$

初态X_0与$\{W_k\}$，$\{V_k\}$独立，记

$$\hat{X}_{i/k}=E[X_i/Z^k]\triangleq E^k[X_i] \tag{50}$$

$$P_{i/k}=E^k[(X_i-\hat{X}_{i/k})(X_i-\hat{X}_{i/k})^T] \tag{51}$$

首先给出预报问题的解，对于预报一步估计

$$\begin{aligned}\hat{X}_{k+1/k}&=E^k[f(X_k,t_k)+W_k]\\&=E^k[f(X_k,t_k)]\triangleq\widehat{f(X_k,t_k)}.\end{aligned} \tag{52}$$

$$\begin{aligned}P_{k+1/k}&=E^k[(f(X_k,t_k)+W_k-\hat{X}_{k+1/k})(f(X_k,t_k)+W_k-\hat{X}_{k+1/k})^T]\\&=E^k[f(X_k,t_k)f^T(X_k,t_k)]+E^k[W_kW_k^T]-\hat{X}_{k+1/k}\hat{X}_{k+1/k}^T.\end{aligned} \tag{53}$$

注意到

$$\begin{aligned}f(X_k,t_k)&=\hat{f}(X_k,t_k)+A_{1f}(\hat{X}_{k/k},t_k)(X_k-\hat{X}_{k/k})\\&\quad+A_{2f}(\hat{X}_{k/k},t_k)(X_k\otimes X_k-\widehat{X_k\otimes X_k}),\end{aligned}$$

其中$A_{1f}(\hat{X}_{k/k},t_k)$，$A_{2f}(\hat{X}_{k/k},t_k)$由(45)，(46)式给出，只需注意$\hat{\ }$是给定$Z^k$之下的条件期望值就可以了，于是

$$\begin{aligned}P_{k+1/k}&=A_{1f}(t_k)E^k[(X_k-\hat{X}_{k/k})(X_k-\hat{X}_{k/k})^T]A_{1f}^T(t_k)\\&\quad+A_{1f}(t_k)E^k[(X_k-\hat{X}_{k/k})(X_k\otimes X_k-\widehat{X_k\otimes X_k})^T]A_{2f}^T(t_k)\\&\quad+A_{2f}(t_k)E^k[(X_k\otimes X_k-\widehat{X_k\otimes X_k})(X_k-\hat{X}_{k/k})^T]A_{1f}^T(t_k)\\&\quad+A_{2f}(t_k)E^k[(X_k\otimes X_k-\widehat{X_k\otimes X_k})(X_k\otimes X_k-\widehat{X_k\otimes X_k})^T]A_{2f}^T(t_k)+Q_k\\&=A_{1f}(t_k)P_{k/k}A_{1f}^T(t_k)+A_{1f}(t_k)E^k[(X_k-\hat{X}_{k/k})(X_k\otimes X_k)^T]A_{2f}^T(t_k)\\&\quad+A_{2f}(t_k)E^k[(X_k\otimes X_k-\widehat{X_k\otimes X_k})X_k^T]A_{1f}^T(t_k)\\&\quad+A_{2f}(t_k)\{E^k[X_k\otimes X_k)(X_k\otimes X_k)^T]-E^k[(X_k\otimes X_k)(\widehat{X_k\otimes X_k})^T]\\&\quad-E^k[(\widehat{X_k\otimes X_k})(X_k\otimes X_k)^T]+E^k[(\widehat{X_k\otimes X_k})(\widehat{X_k\otimes X_k})^T]\}A_{2f}(t_k)\\&\quad+Q_k,\end{aligned}$$

注意上式右边之最后第二项，{ } 中的项为

$$E^k[(X_k\otimes X_k)(X_k\otimes X_k)^T]-(\widehat{X_k\otimes X_k})(\widehat{X_k\otimes X_k})^T. \tag{54}$$

由熟知的不等式

$$E[F\cdot F^T]\geqslant E[F]E[F^T],$$

其中F为随机向量。因此(54)式必大于或等于零。

因而忽略高阶矩时，应当将式(54)的两项同时略去。此时，

$$P_{k+1/k}=A_{1f}(t_k)P_{k/k}A_{1f}^T(t_k)+Q_k$$

$$+A_{1f}(t_k)[\widehat{X_k(X_k\otimes X_k)^T}-\hat{X}_{k/k}(\widehat{X_k\otimes X_k})]A_{2f}^T(t_k)$$

$$+A_{2f}(t_k)](\widehat{(X_k\otimes X_k)X_k^T}-(\widehat{X_k\otimes X_k})\hat{X}_k^T]A_{1f}^T(t_k) \tag{55}$$

式(52)和(55)就是我们所需要的预报方程。

下面给出 $\hat{X}_{k+1/k+1}$ 及其误差协方差阵 $P_{k+1/k+1}$ 的表达式。

将 $\hat{X}_{k+1/k+1}$ 写成如下形式：

$$\hat{X}_{k+1/k+1}=\hat{X}_{k+1/k}+\mathscr{K}_{k+1}[Z_{k+1}-\hat{h}(X_{k+1},t_{k+1})]. \tag{56}$$

记

$$\tilde{X}_{k+1/k}=X_{k+1}-\hat{X}_{k+1/k},\quad \tilde{X}_{k+1/k+1}=X_{k+1}-\hat{X}_{k+1/k+1}.$$

注意到

$$\tilde{X}_{k+1/k+1}=X_{k+1}-\hat{X}_{k+1/k}-\mathscr{K}_{k+1}[Z_{k+1}-\hat{h}(X_{k+1},\ t_{k+1})]$$
$$=\tilde{X}_{k+1/k}-\mathscr{K}_{k+1}[h(X_{k+1},\ t_{k+1})-\hat{h}(X_{k+1},\ t_{k+1})]-\mathscr{K}_{k+1}V_{k+1},$$

将 $h(X_{k+1},\ t_{k+1})$ 在 $\hat{X}_{k+1/k}$ 近旁作二阶统计逼近，则有

$$h(X_{k+1},t_{k+1})=\hat{h}(X_{k+1},\ t_{k+1})+A_{1h}(\hat{X}_{k+1/k},t_{k+1})(X_{k+1}-\hat{X}_{k+1/k})$$
$$+A_{2h}(\hat{X}_{k+1/k},\ t_{k+1})(X_{k+1}\otimes X_{k+1}-\widehat{X_{k+1}\otimes X_{k+1}}).$$

其中 $\hat{\ }$ 表示条件期望 $E[\cdot|Z^k]$。

因此

$$\tilde{X}_{k+1/k+1}=\tilde{X}_{k+1/k}-\mathscr{K}_{k+1}A_{1h}(\hat{X}_{k+1/k},\ t_{k+1})\tilde{X}_{k+1/k}$$
$$-\mathscr{K}_{k+1}A_{2h}(\hat{X}_{k+1/k},\ t_{k+1})(X_{k+1}\otimes X_{k+1}-\widehat{X_{k+1}\otimes X_{k+1}})-\mathscr{K}_{k+1}V_{k+1}. \tag{57}$$

下面来确定 $\mathscr{K}_{k+1}$。为此，注意

$$P_{k+1/k+1}=E^{k+1}\{[(I-\mathscr{K}_{k+1}A_{1h})\tilde{X}_{k+1/k}$$
$$-\mathscr{K}_{k+1}A_{2h}(X_{k+1}\tilde{\otimes}X_{k+1})-\mathscr{K}_{k+1}V_{k+1}]$$
$$\cdot[(I-\mathscr{K}_{k+1}A_{1h})\tilde{X}_{k+1/k}-\mathscr{K}_{k+1}A_{2h}(X_{k+1}\tilde{\otimes}X_{k+1})-\mathscr{K}_{k+1}V_{k+1}]^T\},$$

其中 $$X_{k+1}\tilde{\otimes}X_{k+1}\triangleq X_{k+1}\otimes X_{k+1}-\widehat{X_{k+1}\otimes X_{k+1}}^{k},$$

这里 $\hat{\ }^k$ 表示 $E[\cdot|Z^k]$，为了计算出 $P_{k+1/k+1}$，注意下列关系式：

$$E^k\{S(X_{k+1})u(Z_{k+1})\}=E^k\{E^{k+1}[S(X_{k+1})]u(Z_{k+1})\}. \tag{58}$$

在上式中，令

$$S(X_{k+1})=(X_{k+1}-\hat{X}_{k+1/k+1})(X_{k+1}-\hat{X}_{k+1/k+1})^T,$$
$$u(Z_{k+1})=1,$$

则

$$E^k\{\tilde{X}_{k+1/k+1}\cdot\tilde{X}_{k+1/k+1}^T\}=E^k\{P_{k+1/k+1}\}=P_{k+1/k+1}.$$

于是

$$P_{k+1/k+1}=(I-\mathscr{K}_{k+1}A_{1h})P_{k+1/k}(I-\mathscr{K}_{k+1}A_{1h})^T$$
$$+(I-\mathscr{K}_{k+1}A_{1h})E^k\{\tilde{X}_{k+1/k}V_{k+1}^T\}\mathscr{K}_{k+1}^T$$

$$
\begin{aligned}
&+(I-\mathscr{K}_{k+1}A_{1h})E^{k}\{\tilde{X}_{k+1/k}\cdot(\widetilde{X_{k+1}\otimes X_{k+1}})^{T}\}A_{2h}^{T}\mathscr{K}_{k+1}^{T}\\
&+\mathscr{K}_{k+1}R_{k+1}\mathscr{K}_{k+1}^{T}\\
&-\mathscr{K}_{k+1}E^{k}\{V_{k+1}\,(\widetilde{X_{k+1}\otimes X_{k+1}})^{T}\}A_{2h}^{T}\mathscr{K}_{k+1}\\
&-\mathscr{K}_{k+1}A_{2h}E^{k}\{(\widetilde{X_{k+1}\otimes X_{k+1}})\tilde{X}_{k+1/k}^{T}\}(I-\mathscr{K}_{k+1}A_{1h})^{T}\\
&+\mathscr{K}_{k+1}A_{2h}E^{k}\{(\widetilde{X_{k+1}\otimes X_{k+1}})V_{k+1}^{T}\}\mathscr{K}_{k+1}^{T}\\
&+\mathscr{K}_{k+1}A_{2h}E^{k}\{(\widetilde{X_{k+1}\otimes X_{k+1}})(\widetilde{X_{k+1}\otimes X_{k+1}})^{T}\}A_{2h}^{T}\mathscr{K}_{k+1}^{T};
\end{aligned}
\tag{59}
$$

记

$$
\begin{aligned}
E_{1}&=E^{k}\{\tilde{X}_{k+1/k}\,(\widetilde{X_{k+1}\otimes X_{k+1}})^{T}\}\\
&=E^{k}\{\tilde{X}_{k+1/k}(X_{k+1}\otimes X_{k+1}-\widehat{X_{k+1}\otimes X_{k+1}}^{k})^{T}\}\\
&=E^{k}\{\tilde{X}_{k+1/k}(X_{k+1}\otimes X_{k+1})^{T}\}\\
&=E^{k}\{(X_{k+1}-\hat{X}_{k+1/k})(X_{k+1}\otimes X_{k+1})^{T}\}\\
&=\widehat{X_{k+1}(X_{k+1}\otimes X_{k+1})^{T}}^{k}-\hat{X}_{k+1/k}\widehat{(X_{k+1}\otimes X_{k+1})^{T}}^{k};
\end{aligned}
$$

$$
E_{2}=E^{k}\{V_{k+1}(X_{k+1}\otimes X_{k+1}-\widehat{X_{k+1}\otimes X_{k+1}}^{k})^{T}\}=0;
$$

$$
\begin{aligned}
E_{3}&=E^{k}\{[X_{k+1}\otimes X_{k+1}-(\widehat{X_{k+1}\otimes X_{k+1}})^{k}][X_{k+1}\otimes X_{k+1}-(\widehat{X_{k+1}\otimes X_{k+1}})^{k}]^{T}\}\\
&=\widehat{(X_{k+1}\otimes X_{k+1})(X_{k+1}\otimes X_{k+1})^{T}}^{k}-\widehat{(X_{k+1}\otimes X_{k+1})}^{k}\widehat{(X_{k+1}\otimes X_{k+1})}^{k\,T}.
\end{aligned}
$$

将上述结果代入(59)式，则得

$$
\begin{aligned}
P_{k+1/k+1}=&(I-\mathscr{K}_{k+1}A_{1h})P_{k+1/k}(I-\mathscr{K}_{k+1}A_{1h})^{T}\\
&-(I-\mathscr{K}_{k+1}A_{1h})E_{1}A_{2h}^{T}\mathscr{K}_{k+1}^{T}+\mathscr{K}_{k+1}R_{k+1}\mathscr{K}_{k+1}^{T}\\
&-\mathscr{K}_{k+1}A_{2h}E_{1}^{T}(I-\mathscr{K}_{k+1}A_{1h})^{T}\\
&+\mathscr{K}_{k+1}A_{2h}E_{3}A_{2h}^{T}\mathscr{K}_{k+1}^{T}.
\end{aligned}
\tag{60}
$$

令 $\mathrm{tr}P_{k+1/k+1}=\mathrm{Min}$. 则由式

$$\partial\mathrm{tr}P_{k+1/k+1}/\partial\mathscr{K}_{k+1}=0,$$ 可得

$$
\begin{aligned}
P_{k+1/k}A_{1h}^{T}+E_{1}A_{2h}^{T}=&\mathscr{K}_{k+1}(A_{1h}P_{k+1/k}A_{1h}^{T}\\
&+A_{1h}E_{1}A_{2h}^{T}+A_{2h}E_{1}^{T}A_{1h}^{T}+A_{2h}E_{3}A_{2h}^{T}+R_{k+1}),
\end{aligned}
$$

于是

$$
\begin{aligned}
\mathscr{K}_{k+1}=&(P_{k+1/k}A_{1h}^{T}+E_{1}A_{2h}^{T})(A_{1h}P_{k+1/k}A_{1h}^{T}+A_{1h}E_{1}A_{2h}^{T}\\
&+A_{2h}E_{1}^{T}A_{1h}^{T}+A_{2h}E_{3}A_{2h}^{T}+R_{k+1})^{-1}.
\end{aligned}
\tag{61}
$$

这样，我们获得了统计逼近下的非线性滤波公式(52)，(55)，(56)、(60)，(61)。
如果忽略四阶矩，则可取 $E_{3}\cong 0$.
如果统计逼近仅采用线性逼近，则统计逼近滤波公式成为

$$(*)\begin{cases}\hat{X}_{k+1/k+1}=\hat{X}_{k+1/k}+\mathscr{K}_{k+1}[Z_{k+1}-\hat{h}(X_{k+1},t_{k+1})],\\ \hat{X}_{k+1/k}=\hat{f}(X_k,t_k),\\ P_{k+1/k}=A_{1f}(t_k)P_{k/k}A_{1f}^T(t_k)+Q_k,\\ \mathscr{K}_{k+1}=P_{k+1/k}A_{1h}^T(t_{k+1})[A_{1h}(t_{k+1})P_{k+1/k}A_{1h}^T(t_{k+1})+R_{k+1}]^{-1},\\ P_{k+1/k+1}=(I-\mathscr{K}_{k+1}A_{1h}(t_{k+1}))P_{k+1/k}(I-\mathscr{K}_{k+1}A_{1h}(t_{k+1}))^T+\mathscr{K}_{k+1}R_{k+1}\mathscr{K}_{k+1}^T.\end{cases}$$

这就是熟知的统计线性化滤波公式。上式中之$A_{1h}(t_{k+1})$及$A_{1h}(t_k)$由下式表示：

$$A_{1h}(t_{k+1})=[\widehat{h(X_{k+1/k}(X_{k+1/k}^T}^{k}-\hat{h}(X_{k+1/k})\hat{X}_{k+1/k}^T]\cdot P_{k+1/k}^{-1},\tag{62}$$

$$A_{1f}(t_k)=[\widehat{f(X_{k/k})X_{k/k}^T}^{k}-\hat{f}(X_{k/k})\hat{X}_{k/k}^T]P_{k/k-1}^{-1}.\tag{63}$$

统计追近滤波的实现，要求去计算出有关的条件均值。在实际问题中将遇到计算高維积分的困难。即使在正态的場合，例如

$$X_{k+1/k}\sim N(\hat{X}_{k+1/k},\ P_{k+1/k}),$$
$$X_{k/k}\sim N(\hat{X}_{k/k},\ P_{k/k}),$$

要去计算(62)，(63)中的条件期望值也还不是轻而易举的。某些实际应用中涉及的期望值计算参见文后的附录。

对于应用中常遇到的 非线性連续-离散系统 的滤波，用统计追近方法是不困难的。事实上，设系统为

$$\begin{cases}\dot{X}(t)=f(X(t),t)+W(t),\\ Z_k=h(X_k,\ t_k)+V_k.\end{cases}\tag{64}$$

其中$W(t)$—零均值白噪声过程，

$$E[W(t)W^T(\tau)]=Q(t)\delta(t-\tau).$$

其中，$\delta(t)$是Diracδ—函数。

此时的统计追近二阶滤波只需给出状态预报及其误差协方差阵的关系式就行了。而状态估计及其误差协方差阵同于前面的结果。为此，记

$$\hat{X}(t/t_{k-1})=E^{k-1}[X(t)],$$
$$P(t/t_{k-1})=E^{k-1}[(X(t)-\hat{X}(t/t_{k-1}))(X(t)-\hat{X}(t/t_{k-1}))^T].$$

于是

$$\frac{d\hat{X}(t/t_{k-1})}{dt}=E[f(X(t),t)/Z^{k-1}]=\hat{f}(X(t),t).\tag{65}$$

而

$$\frac{d}{dt}P(t/t_{k-1})=E^{k-1}\left\{\frac{d}{dt}[(X(t)-\hat{X}(t/t_{k-1}))(X(t)-\hat{X}(t/t_{k-1}))^T]\right\}$$

$$=E^{k-1}\left\{\frac{d}{dt}[X(t)X^T(t)-X(t)\hat{X}(t/t_{k-1})\right.$$

$$\left.-\hat{X}(t/t_{k-1})X^T(t)+\hat{X}(t/t_{k-1})\hat{X}^T(t/t_{k-1})]\right\}.$$

注意到式(64)，则有

$$\frac{d}{dt}(X(t)X^T(t))=\dot{X}(t)X^T(t)+X(t)\dot{X}^T(t)$$

$$=[f(X(t),t)+W(t)]X^T(t)+X(t)[f^T(X(t),t)+W(t)]$$

取期望值，则有

$$E^{k-1}\left[\frac{d}{dt}(X(t)X^T(t))\right]=\widehat{f(X(t),t)X^T(t)}^{k-1}+\widehat{X(t)f^T(X(t),t)}^{k-1};$$

而

$$\frac{d}{dt}[X(t)\hat{X}^T(t/t_{k-1})]=\dot{X}(t)\hat{X}^T(t/t_{k-1})+X(t)\dot{\hat{X}}^T(t/t_{k-1})$$

$$=[f(X(t),t)+W(t)]\hat{X}^T(t/t_{k-1})+X(t)\hat{f}^T(X(t),t),$$

$$E^{k-1}\left[\frac{d}{dt}(X(t)\hat{X}(t|t_{k-1}))\right]=\hat{f}(X(t),t)\hat{X}^T(t|t_{k-1})$$

$$+\hat{X}(t|t_{k-1})\hat{f}^T(X)t(,t);$$

其中 $\hat{\cdot}$——在Z^{k-1}给定之下，“·”的期望值。

类似地，可以得到

$$E^{k-1}\left[\frac{d}{dt}(\hat{X}(t|t_{k-1})X^T(t)\right]=\hat{f}(X(t),t)\hat{X}^T(t|t_{k-1})+\hat{X}(t|t_{k-1})\hat{f}^T(X(t),t);$$

$$E^{k-1}\left[\frac{d}{dt}(\hat{X}(t|t_{k-1})\hat{X}^T(t|t_{k-1})\right]=\hat{f}(X(t),t)\hat{X}^T(t|t_{k-1})+\hat{X}(t|t_{k-1})\hat{f}^T(X(t),t)$$

于是

$$\frac{d}{dt}P(t|t_{k-1})=\widehat{f(X(t),t)X^T(t)}-\hat{f}(X(t),t)\hat{X}^T(t|t_{k-1})$$

$$+\widehat{X(t)f^T(X(t),t)}-\hat{X}(t|t_{k-1})\hat{f}^T(X(t),t)+Q(t). \qquad (66)$$

方程(65)和 (66) 可以我们所需要的用于计算 $\hat{X}_{k/k-1}$ 及 $P_{k/k-1}$ 的一般关系式，其初始条件为给定 $\hat{X}_{k-1/k-1}$ 及 $P_{k-1/k-1}$.

如果将 $f(X(t),t)$在$\hat{X}_{k-1/k-1}$近旁应用二阶统计追近，则可以获得$\frac{d}{dt}P(t|t_{k-1})$ 的近似表达式。如果 $f(X(t),t)$ 仅用线性统计追近，则易知

$$\frac{d}{dt}P(t|t_{k-1})=A_{1f}(t_{k-1})P(t|t_{k-1})+P(t|t_{k-1})A_{1f}^T(t_{k-1})+Q(t). \qquad (67)$$

其初始条件为

$$P(t|t_{k-1})\Big|_{t=t_{k-1}}=P_{k-1/k-1}.$$

在此情况下，滤波计算中的预报估计 $\hat{X}_{k/k-1}$ 及 $P_{k/k-1}$ 由解方程 (65) 及 (67) 获得，其他公式同于公式(＊)。

附　录　关于某些积分的计算

1. 记 $X=(X_1X_2\cdots X_n)^T$，$X\sim N(\hat{X},P)$，$P>0$，

$$g(X)=X_1^{\nu_1}X_2^{\nu_2}\cdots X_n^{\nu_n},\qquad \nu_1,\cdots,\nu_n\text{为正整数}.$$

计算　$E[g(X)]$.

$$\hat{g}=\frac{1}{(2\pi)^{\frac{n}{2}}|P|^{\frac{1}{2}}}\int_{-\infty}^{+\infty}g(X)e^{-\frac{1}{2}(X-\hat{X})^TP^{-1}(X-\hat{X})}dX$$

$$=\frac{1}{(2\pi)^{\frac{n}{2}}|P|^{\frac{1}{2}}}\int_{-\infty}^{+\infty}g(r+\hat{X})e^{-\frac{1}{2}r^TP^{-1}r}dr. \tag{1}$$

其中　$r=X-\hat{X}$.

由于 $P>0$，因此存在正交阵 T，使

$$T^TP^{-1}T=D=\begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix}.$$

其中 $\lambda_1,\cdots,\lambda_n$ 为 P 的特征根，且 $\lambda_i>0$，$i=1,\cdots,n$. 作变换 $S=T^{-1}r$，此变换的Jacobi行列式 $|T^{-1}|=1$. 于是

$$\hat{g}=c\int_{-\infty}^{+\infty}g(TS+\hat{X})e^{-\frac{1}{2}S^TDS}dS$$

$$=c\int_{-\infty}^{+\infty}\cdots\int g(TS+\hat{X})e^{-\frac{1}{2}(\lambda_1S_1^2+\cdots+\lambda_nS_n^2)}dS_1dS_2\cdots dS_n,$$

其中

$$c=1/[(2\pi)^{\frac{n}{2}}|P|^{\frac{1}{2}}].$$

记

$$T=(T_{ij}),\ i,j=1,\cdots,n,$$

$$S=\begin{pmatrix}S_1\\ \vdots\\ S_n\end{pmatrix},\qquad \hat{X}=\begin{pmatrix}\hat{X}_1\\ \vdots\\ \hat{X}_n\end{pmatrix},$$

则

$$g(TS+\hat{X})=\left(\sum_{i=1}^nT_{1i}S_i+\hat{X}_1\right)^{\nu_1}\cdots\left(\sum_{i=1}^nT_{n_i}S_i+\hat{X}_n\right)^{\nu_n},$$

$$\hat{g}=c\int_{-\infty}^{+\infty}\cdots\int\left(\sum_{i=1}^nT_{1i}S_i+\hat{X}_1\right)^{\nu_1}\cdots\left(\sum_{i=1}^nT_{n_i}S_i+\hat{X}_n\right)^{\nu_n}\exp\left[-\frac{1}{2}\sum_{i=1}^n\lambda_iS_i^2\right]dS_1\cdots dS_n,$$

由于

$$\left(\sum_{i=1}^{n} T_{ki}S_i + \hat{X}_k\right)^{\nu_k} = \sum_{\substack{m_1,\cdots,m_{n+1}\\ m_1+\cdots+m_{n+1}=\nu_k}} (T_{k1}S_1)^{m_1}(T_{k2}S_2)^{m_2}\cdots(T_{kn}S_n)^{m_n}(\hat{X}_k)^{m_{n+1}}$$

$$= \sum_{\substack{m_1,\cdots,m_{n+1}\\ m_1+\cdots+m_{n+1}=\nu_k}} (T_{k1}^{m_1}\cdots T_{kn}^{m_n}\hat{X}_k^{m_{n+1}})\, S_1^{m_1}\cdots S_n^{m_n},$$

于是

$$\hat{g}=\int_{-\infty}^{+\infty}\!\!\cdots\!\int c\left[\prod_{k=1}^{n}\sum_{\substack{m_1,\cdots,m_{n+1}\\ m_1+\cdots+m_{n+1}=\nu_k}} (T_{k1}^{m_1}\cdots T_{kn}^{m_n}\hat{X}_k^{m_{n+1}})\, S_1^{m_1}\cdots S_n^{m_n}\right]$$

$$\cdot \exp\left[-\frac{1}{2}\sum_{i=1}^{n}\lambda_i S_i^2\right]dS_1\cdots dS_n$$

$$=c\int_{-\infty}^{+\infty}\!\!\cdots\!\int \sum_{\substack{m_1^{(1)},\cdots,m_{n+1}^{(1)}\\ m_1^{(1)}+\cdots+m_{n+1}^{(1)}=\nu_1}}\cdots\sum_{\substack{m_1^{(n)},\cdots,m_{n+1}^{(n)}\\ m_1^{(n)}+\cdots+m_{n+1}^{(n)}=\nu_n}}\left[\prod_{i=1}^{n} T_{k1}^{m_1^{(i)}}\cdots T_{kn}^{m_n^{(i)}}\hat{X}_i^{m_{n+1}^{(i)}}\right]$$

$$\cdot\; S_1^{\sum_1^n m_i^{(1)}}\cdots S_n^{\sum_1^n m_i^{(n)}}\exp\left[-\frac{1}{2}\sum_{i=1}^{n}\lambda_i S_i^2\right]dS_1\cdots dS_n. \tag{2}$$

因此，只需计算下列积分

$$I=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} S_1^{M^{(1)}}\cdots S_n^{M^{(n)}}\exp\left[-\frac{1}{2}\sum_{i=1}^{n}\lambda_i S_i^2\right]dS_1\cdots dS_n.$$

易知

$$I=\prod_{i=1}^{n}\left(\int_{-\infty}^{+\infty} S_i^{M^{(i)}}\exp\left[-\frac{1}{2}\lambda_i S_i^2\right]dS_i\right)$$

$$=\prod_{i=1}^{n}\left(\frac{(M^{(i)}-1)!}{\lambda_i^{M^{(i)}-1/2}}\sqrt{2\pi}\right). \tag{3}$$

将(3)代入(2)，即得 $g(x)$ 的期望值为

$$\hat{g}=c\sum_{\substack{m_1^{(1)},\cdots,m_{n+1}^{(1)},\\ m_1^{(1)}+\cdots+m_{n+1}^{(1)}=\nu_1}}\cdots\sum_{\substack{m_1^{(n)},\cdots,m_{n+1}^{(n)},\\ m_1^{(n)}+\cdots+m_{n+1}^{(n)}=\nu_n}}\prod_{i=1}^{n}\left(T_{k1}^{m_1^{(i)}}\cdots T_{kn}^{m_n^{(i)}}\hat{X}_i^{m_{n+1}^{(i)}}\right)$$

$$\cdot\left[\prod_{j=1}^{n}\left(\frac{(M^{(j)}-1)!}{\lambda_j^{M^{(j)}-1/2}}\sqrt{2\pi}\right)\right]. \tag{4}$$

2. 如果 $g(X)=e^{-\nu(X_1+\cdots+X_n)}X_1^{\nu_1}\cdots X_n^{\nu_n}$，计算 $E[g(X)]$.

用前面同样的方法，可知

$$\hat{g}=c\int_{-\infty}^{+\infty}\cdots\int g(TS+\hat{X})\exp\left[-\frac{1}{2}\sum_{i=1}^{n}\lambda_i S_i^2\right]dS_1\cdots dS_n$$

$$=c\int_{-\infty}^{+\infty}\cdots\int e^{-\nu\left(\left[\sum_1^n T_{1i}S_i+\hat{X}_1\right)+\cdots+\left(\sum_1^n T_{ni}S_i+\hat{X}_n\right)\right]}$$

$$\cdot\left(\sum_1^n T_{1i}S_i+\hat{X}_1\right)^{\nu_1}\cdots\left(\sum_1^n T_{ni}S_i+\hat{X}_n\right)^{\nu_n}e^{-\frac{1}{2}\sum_1^n\lambda_i S_i^2}dS_1\cdots dS_n$$

$$=c\int_{-\infty}^{+\infty}\cdots\int e^{-\frac{\lambda_1}{2}[(S_1+\mu_1)^2-\mu_1^2]}\cdots e^{-\frac{\lambda_n}{2}[(S_n+\mu_n)^2-\mu_n^2]}$$

$$\cdot\left(\sum_1^n T_{1i}S_i+\hat{X}_1\right)^{\nu_1}\cdots\left(\sum_1^n T_{ni}S_i+\hat{X}_n\right)^{\nu_n}dS_1\cdots dS_n$$

$$=ce^{\sum_1^n\lambda_i\mu_i^2}\int_{-\infty}^{+\infty}\cdots\int e^{-\frac{\lambda_1}{2}(S_1+\mu_1)^2}\cdots e^{-\frac{\lambda_n}{2}(S_n+\mu_n)^2}$$

$$\cdot\left(\sum_1^n T_{1i}S_i+\hat{X}_1\right)^{\nu_1}\cdots\left(\sum_1^n T_{ni}S_i+\hat{X}_n\right)^{\nu_n}dS_1\cdots dS_n,$$

其中

$$\mu_i=\nu\frac{T_i}{\lambda_i},\ i=1,\cdots,n,\ \lambda_i>0.$$

作变换

$$S_i+\mu_i=v_i,\ i=1,\cdots,n,$$

则

$$\hat{g}=ce^{\sum_1^n\lambda_i\mu_i^2}\int_{-\infty}^{+\infty}\cdots\int\exp\left[-\frac{1}{2}\sum_1^n\lambda_i v_i^2\right]\cdot\left(\sum_1^n T_{1i}v_i+\hat{X}_1-\sum_1^n T_{1i}\mu_i\right)^{\nu_1}$$

$$\cdots\left(\sum_1^n T_{ni}v_i+\hat{X}_n-\sum_1^n T_{ni}\mu_i\right)^{\nu_n}dv_1\cdots dv_n. \tag{5}$$

于是又回复至形式 1 的积分。

示例 设有垂直下落的飞行器，记$X_1(t)$为t时刻的高度，$X_2(t)$为t时刻的速度（向下）的大小，m为质量，C_D为常值阻力系数，A为计算阻力的参考面积，则有如下的运动方程：

$$\begin{cases}\dot{X}_1(t)=-X_2(t).\\ \dot{X}_2(t)=-\dfrac{C_DA_\rho^2}{2m}X_2(t).\end{cases}$$

其中ρ为t时刻的大气密度，

应用近似公式

$$\rho=\rho_0e^{-\nu X_1(t)},$$

其中ρ_0为地面的大气密度，ν为常量，记

$$X_3 = C_D A \rho_0 / (2m),$$

于是飞行器的运动方程可写为

$$\begin{cases} \dot{X}_1 = -X_2 \triangleq f_1(x), \\ \dot{X}_2 = -e^{-\nu X_1} X_2^2 X_3 \triangleq f_2(X). \\ \dot{X}_3 = 0 \triangleq f_3(X) \end{cases}$$

其中

$$X = (X_1 \quad X_2 \quad X_3)^T,$$

则

$$\dot{X}(t) = f(x(t)) = \begin{pmatrix} X_2 \\ -e^{-\nu X_2} X_2^2 X_3 \\ 0 \end{pmatrix}$$

这样，运用统计迫近方法估计高度时，将涉及积分 2 的计算。

一般地，再入飞行器的跟踪问题将遇到附录 1,2 中的积分计算，可参见文献 [7] 的第十七章。

参考文献

[1] C. J. Masreliez, Approximation Non-Gaussian Filtering With Linear State and Observation Relations. IEEE Tran. Auto. Contr, Vol. AC-20, 1975, 107-110.

[2] C. J. Masreliez & R. D. Martin, Robust Estimation for the Linear Model and Robustifying the Kalman Filter. IEEE Tran. Auto. contr, Vol. AC-22, 1977, 361-371.

[3] P. J. Huber, Robust Statistics, John wiley, 1981.

[4] W. S. Agee, Application of Robust Filtering and Smoothing to Tracking Data, T.R.No.71, 1979.

[5] R. Henrikson, The Truncated Second-order Nonlinear Filter Revisited, IEEE Vol. AC-27, 1982, 247-251.

[6] A. Gelb, Applied OPtimal Estimation, The M.I.T.Press, 1974.

[7] 张金槐，常兆诚，飞行器试验统计学（下册），国防科技大学出版，1984年.

Approximation of Nonlinear Filtering

Zhang Jinhuai

Abstract

This paper deals with the problems of estimating state variable for dynamic System (either linear or non-linear) .If we limit our problems in linear estimating algorithm, then the minimum variance estimator is suffic-

iently optimum. But in general, the optimal estimator may not be the linear case. For this purpose, We discuss the approximations of the non-linear filtering problems. we present the methods for two types of approximations, the first, Gram-Charlier approximation; and the second, the statistical second order approximation.

In order to facilitate calculation, in this paper the approximate solution in recurrence form is given. The generalized Kalman filtering method and statistical lineanzation method are used only as the particular cases.