国防科技大学学报
JOURNAL OF NATIONAL UNIVERSITY OF DEFENSE TECHNOLOGY
一九八五年第三期　总第五十一期　No.3 1985 Sum. 51

正态迫近、若干矩公式及其在非线性系统状态估计中的应用

茅　宁

提　要　本文首先推导了若干矩阵函数对向量变量的求导公式，并以此为基础对任意的随机向量，给出其分布密度函数的正态迫近表示。其次，文中给出了关于正态随机向量的若干矩公式。作为上述结果的应用，文中讨论了非线性系统状态估计的一种迫近方法。

一、问题的提出

设 $f(x)$ 为任意 n 維随机向量 $X=(X_1\cdots X_n)^T$ 的分布密度函数，则我们可以对 $f(x)$ 在任意均值向量 $\hat{X}$ 和方差矩阵 P 处做正态迫近[1]：

$$f(x)=N_n(x:\hat{X},P)\sum_{k=0}^{\infty}\sum_{i_1+i_2+\cdots+i_n=k}\frac{1}{i_1!\,i_2!\cdots i_n!}H_{i_1\cdots i_n}(x-\hat{X})\,b_{i_1\cdots i_n} \tag{1.1}$$

其中 $N_n(x:\hat{X},P)$ 表示均值为 $\hat{X}$、方差阵为 P 的 n 維正态分布密度函数，而 $H_{i_1\cdots i_n}(x)$ 为 n 維变量的 Hermite 多项式：

$$H_{i_1\cdots i_n}(x)=(-1)^{i_1+\cdots+i_n}e^{\frac{1}{2}x^TP^{-1}x}\frac{\partial^{i_1+\cdots+i_n}}{\partial x_1^{i_1}\cdots\partial x_n^{i_n}}e^{-\frac{1}{2}x^TP^{-1}x} \tag{1.2}$$

$x=(x_1\cdots x_n)^T$，此外，若记

$$G_{i_1\cdots i_n}(x)=(-1)^{i_1+\cdots+i_n}e^{\frac{1}{2}z^TPz}\frac{\partial^{i_1+i_2+\cdots+i_n}}{\partial z_1^{i_1}\cdots\partial z_n^{i_n}}e^{-\frac{1}{2}z^TPz} \tag{1.3}$$

$z=P^{-1}x$，则

$$b_{i_1\cdots i_n}=\int G_{i_1\cdots i_n}(x)f(x)\,dx \tag{1.4}$$

显然，由于上述迫近公式在结构上比较松散，使用起来是很不方便的。本文将用矩阵形式给出 $k=4$ 时的正态迫近表示，更高阶的形式也可以类似推出。然后，利用正态

本文1984年11月9日收到

迫近讨论非线性系统状态估计的一种方法。

二、有关的导数公式及正态迫近表示

首先研究一下有关向量、矩阵函数对向量变量的求导公式。为叙述方便，记任意 $n\times m$ 矩阵 A 的第 i 列元素为 $A_{\cdot i}$，第 j 行元素为 $A_{j\cdot}$，$(i=1,\cdots,m,j=1,\cdots,n)$；$A$ 的列条（column string）为

$$cs(A)=(A^T_{\cdot 1}\cdots A^T_{\cdot m})^T \tag{2.1}$$

A 的行条（row string）为

$$rs(A)=(A_{1\cdot}\cdots A_{n\cdot}) \tag{2.2}$$

它们分别为 nm 維列向量和 nm 維行向量。此外记$\otimes$为矩阵的 Kronecker 积。

假设 $P_{n\times n}$ 为对称矩阵，则我们知道如下事实

$$\frac{\partial}{\partial x}(x^TPx)=2Px \tag{2.3}$$

$$\frac{\partial}{\partial x^T}(x^TPx)=2x^TP \tag{2.4}$$

其中 x 为 n 維列向量。此外可推得以下几个求导公式（证明见附录）

1° 对于任意矩阵 $A_{n\times m}$

$$\frac{\partial}{\partial x}(Ae^{-\frac{1}{2}x^TPx})=-(Px)\otimes Ae^{-\frac{1}{2}x^TPx} \tag{2.5}$$

$$\frac{\partial}{\partial x^T}(Ae^{-\frac{1}{2}x^TPx})=-(Px)^T\otimes Ae^{-\frac{1}{2}x^TPx} \tag{2.6}$$

2° $$\frac{\partial}{\partial x}(Pxx^TP)=cs(P)(Px)^T+P\otimes(Px) \tag{2.7}$$

$$\frac{\partial}{\partial x^T}(Pxx^TP)=rs(P)\otimes(Px)+P\otimes(Px)^T \tag{2.8}$$

3° $$\frac{\partial}{\partial x^T}[(Px)\otimes P]=P\otimes P \tag{2.9}$$

4° $$\frac{\partial}{\partial x^T}[cs(P)(Px)^T]=cs(P)rs(P) \tag{2.10}$$

5° $$\frac{\partial}{\partial x^T}[P\otimes(Px)]=(P\otimes P_{\cdot 1}\cdots P\otimes P_{\cdot n}) \tag{2.11}$$

6° 设 $Y(x)$ 为 n 維列向量函数，$A(x)$ 为 $n\times n$ 矩阵函数，则

$$\frac{\partial}{\partial x^T}[Y(x)\otimes A(x)]=\frac{\partial}{\partial x^T}[Y(x)]\otimes A(x)+Y(x)\otimes\frac{\partial}{\partial x^T}[A(x)] \tag{2.12}$$

利用上述公式可方便地得到标量函数

$$f_1(x)=e^{-\frac{1}{2}x^TPx}$$

关于 n 維向量 x 的各阶偏导数，其中 $P_{n\times n}$ 为对称矩阵。例如：

1° $$\frac{\partial}{\partial x}(e^{-\frac{1}{2}x^TPx})=-Pxe^{-\frac{1}{2}x^TPx} \tag{2.13}$$

2° $$\frac{\partial}{\partial x^T}\left[\frac{\partial}{\partial x}(e^{-\frac{1}{2}x^TPx})\right]=(-P+Pxx^TP)e^{-\frac{1}{2}x^TPx} \tag{2.14}$$

3° $$\frac{\partial}{\partial x}\left[\frac{\partial}{\partial x^T}\left(\frac{\partial}{\partial x}(e^{-\frac{1}{2}x^TPx})\right)\right]=[(Px)\otimes P+P\otimes(Px)+cs(P)(Px)^T-(Px)\otimes(Pxx^TP)]e^{-\frac{1}{2}x^TPx} \tag{2.15}$$

4° $$\frac{\partial}{\partial x^T}\left[\frac{\partial}{\partial x}\left(\frac{\partial}{\partial x^T}\left(\frac{\partial}{\partial x}(e^{-\frac{1}{2}x^TPx})\right)\right)\right]=[P\otimes P-(Pxx^TP)\otimes P+cs(P)rs(P)+(P\otimes P_{\cdot 1}\cdots P\otimes P_{\cdot n})-(Px)^T\otimes P\otimes(Px)-(Px)^T\otimes cs(P)(Px)^T+(Pxx^TP)\otimes(Pxx^TP)-P\otimes(Pxx^TP)-(Px)rs(P)\otimes(Px)-(Px)\otimes P\otimes(Px)^T]e^{-\frac{1}{2}x^TPx} \tag{2.16}$$

根据上述偏导数公式，下面着手将(1.1)式写成紧凑的形式：

(1) 零阶项（$k=0$），$b_0=H_0=1$；

(2) 一阶项（$k=1$），记

$$H_1(x)=[H_{1,0,\cdots,0}(x),H_{0,1,0,\cdots,0}(x),\cdots,H_{0,\cdots,0,1}(x)]^T$$

则有

$$H_1(x)=(-1)^1e^{\frac{1}{2}x^TP^{-1}x}\frac{\partial}{\partial x}(e^{-\frac{1}{2}x^TP^{-1}x})=P^{-1}x \tag{2.17}$$

另记

$$b_1=[b_{1,0,\cdots,0}\cdots b_{0,\cdots,0,1}]^T,\quad G_1(x)=[G_{1,0,\cdots,0}(x),\cdots,G_{0,\cdots,0,1}(x)]^T$$

则

$$G_1(x)=(-1)^1e^{\frac{1}{2}z^TPz}\frac{\partial}{\partial z}(e^{-\frac{1}{2}z^TPz})=Pz=x$$

$$b_1=\int G_1(x)f(x)dx=E[X] \tag{2.18}$$

(3) 二阶项（$k=2$），记

$$H_2(x)=\begin{bmatrix}H_{2,0,\cdots,0}(x) & H_{1,1,0,\cdots,0}(x) & \cdots & H_{1,0,\cdots,0,1}(x)\\ \vdots & \vdots & \ddots & \vdots\\ H_{1,0,\cdots,0,1}(x) & H_{0,1,0,\cdots,0,1}(x) & \cdots & H_{0,\cdots,0,2}(x)\end{bmatrix}$$

$$G_2(x)=\begin{bmatrix}G_{2,0,\cdots,0}(x) & G_{1,1,0,\cdots,0}(x) & \cdots & G_{1,0,\cdots,0,1}(x)\\ \vdots & \vdots & \ddots & \vdots\\ G_{1,0,\cdots,0,1}(x) & G_{0,1,0,\cdots,0,1}(x) & \cdots & G_{0,\cdots,0,2}(x)\end{bmatrix}$$

则有

$$H_2(x)=(-1)^2e^{\frac{1}{2}x^TPx}\frac{\partial}{\partial x^T}\left[\frac{\partial}{\partial x}(e^{-\frac{1}{2}x^TP^{-1}x})\right]=P^{-1}xx^TP^{-1}-P^{-1} \tag{2.19}$$

$$G_2(x)=(-1)^2e^{\frac{1}{2}z^TPz}\frac{\partial}{\partial z^T}\left[\frac{\partial}{\partial z}(e^{-\frac{1}{2}z^TPz})\right]=Pzz^TP-P=xx^T-P$$

即有

$$b_2=\int G_2(x)f(x)dx=\int(xx^T-P)f(x)dx=E[X]E^T[X] \tag{2.20}$$

(4) 三阶项 ($k=3$)，记

$$H_3(x)=\begin{pmatrix} H_{3,0,\cdots,0}(x) & H_{2,1,0,\cdots,0}(x) & \cdots & H_{2,0,\cdots,0,1}(x) \\ \vdots & \vdots & \ddots & \vdots \\ H_{2,0,\cdots,0,1}(x) & H_{1,1,0,\cdots,0,1}(x) & \cdots & H_{1,0,\cdots,0,2}(x) \\ H_{2,1,0,\cdots,0}(x) & H_{1,2,0,\cdots,0}(x) & \cdots & H_{1,1,0,\cdots,0,1}(x) \\ \vdots & \vdots & \ddots & \vdots \\ H_{1,0,\cdots,0,2}(x) & H_{0,1,0,\cdots,0,2}(x) & \cdots & H_{0,\cdots,0,3}(x) \end{pmatrix}$$

$$G_3(x)=\begin{pmatrix} G_{3,0,\cdots,0}(x) & \cdots & G_{2,0,\cdots,0,1}(x) \\ \vdots & \ddots & \vdots \\ G_{2,0,\cdots,0,1}(x) & \cdots & G_{1,0,\cdots,0,2}(x) \\ \vdots & \ddots & \vdots \\ G_{1,0,\cdots,0,2}(x) & \cdots & G_{0,\cdots,0,3}(x) \end{pmatrix}$$

则有

$$\begin{aligned} H_3(x)&=(-1)^3e^{\frac{1}{2}x^TP^{-1}x}\frac{\partial}{\partial x}\left[\frac{\partial}{\partial x^T}\left(\frac{\partial}{\partial x}(e^{-\frac{1}{2}x^TP^{-1}x})\right)\right] \\ &=(P^{-1}x)\otimes(P^{-1}xx^TP^{-1})-(P^{-1}\otimes(P^{-1}x) \\ &\quad +cs(P^{-1})x^TP^{-1}+(P^{-1}x)\otimes P^{-1}) \end{aligned} \tag{2.21}$$

$$\begin{aligned} G_3(x)&=(-1)^3e^{\frac{1}{2}z^TPz}\frac{\partial}{\partial z}\left[\frac{\partial}{\partial z^T}\left(\frac{\partial}{\partial z}(e^{-\frac{1}{2}z^TPz})\right)\right] \\ &=x\otimes(xx^T)-P\otimes x-x\otimes P-cs(P)x^T \end{aligned}$$

故有

$$b_3=\int G_3(x)f(x)dx=\int[x\otimes(xx^T)-P\otimes x-x\otimes P-cs(P)x^T]f(x)dx \tag{2.22}$$

(5) 四阶项 ($k=4$)，记

$$H_4(x)=\begin{pmatrix} H_{4,0,\cdots,0}(x) & \cdots & H_{3,0,\cdots,0,1}(x) & H_{3,1,0,\cdots,0}(x) & \cdots & H_{2,0,\cdots,0,2}(x) \\ \vdots & & \vdots & \vdots & & \vdots \\ H_{3,0,\cdots,0,1}(x) & \cdots & H_{2,0,\cdots,0,2}(x) & H_{2,1,0,\cdots,0,1}(x) & \cdots & H_{1,0,\cdots,0,3}(x) \\ H_{3,1,0,\cdots,0,}(x) & \cdots & H_{2,1,0,\cdots,0,1}(x) & H_{2,2,0,\cdots,0}(x) & \cdots & H_{1,1,0,\cdots,0,2}(x) \\ \vdots & & \vdots & \vdots & & \vdots \\ H_{2,0,\cdots,0,2}(x) & \cdots & H_{1,0,\cdots,0,3}(x) & H_{1,1,0,\cdots,0,2}(x) & \cdots & H_{0,\cdots,0,4}(x) \end{pmatrix}$$

$$G_4(x)=\begin{bmatrix} G_{4,0,\cdots,0}(x) & \cdots & G_{2,0,\cdots,0,2}(x) \\ \vdots & \ddots & \vdots \\ G_{2,0,\cdots,0,2}(x) & \cdots & G_{0,\cdots,0,4}(x) \end{bmatrix}$$

其中 $G_4(x)$ 中各元素 $G_{i_1,i_2,\cdots,i_n}(x)$ 的排列与 $H_4(x)$ 的排列相同。则有

$$\begin{aligned} H_4(x)=&(-1)^4e^{-\frac{1}{2}x^TP^{-1}x}\frac{\partial}{\partial x^T}\left[\frac{\partial}{\partial x}\left(\frac{\partial}{\partial x^T}\left(\frac{\partial}{\partial x}(e^{-\frac{1}{2}x^TP^{-1}x})\right)\right)\right] \\ =&(P^{-1}xx^TP^{-1})\otimes(P^{-1}xx^TP^{-1})-P^{-1}\otimes(P^{-1}xx^TP^{-1}) \\ &-(P^{-1}x)rs(P^{-1})\otimes(P^{-1}x)-(P^{-1}x)\otimes P^{-1}\otimes(P^{-1}x)^T \\ &+(P^{-1}\otimes P^{-1}_{\cdot 1}\cdots P^{-1}\otimes P^{-1}_{\cdot n})-(P^{-1}x)^T\otimes P^{-1}\otimes(P^{-1}x) \\ &+cs(P^{-1})rs(P^{-1})-(P^{-1}x)^T\otimes cs(P^{-1})(P^{-1}x)^T+P^{-1}\otimes P^{-1} \\ &-(P^{-1}xx^TP^{-1})\otimes P^{-1} \end{aligned} \tag{2.23}$$

及

$$\begin{aligned} b_4=&\int G_4(x)f(x)dx=\int[(xx^T)\otimes(xx^T)-P\otimes(xx^T)-xrs(P)\otimes x \\ &-x\otimes P\otimes x^T-x^T\otimes P\otimes x+(P\otimes P_{\cdot 1}\cdots P\otimes P_{\cdot n})+cs(P)rs(P) \\ &-x^T\otimes(cs(P)x^T)+P\otimes P-(xx^T)\otimes P]f(x)dx \end{aligned} \tag{2.24}$$

于是，在略去5阶以上项后，关于 $f(x)$ 的正态逼近公式为

$$\begin{aligned} f(x)=N_n(x;\hat{X},P)\Big[&1+H_1^T(x-\hat{X})b_1+\frac{1}{2}t_r(H_2(x-\hat{X})b_2) \\ &+\frac{1}{3!}t_r(H_3(x-\hat{X})b_3^T)+\frac{1}{4!}t_r(H_4(x-\hat{X})b_4)\Big] \end{aligned} \tag{2.25}$$

特别，若 $X\sim N_n(x;\hat{X},P)$，则易推得 $b_i=0$，$i=1,\cdots,4$.

关于上述公式的逼近精度，当 $\hat{X}$ 和 P 接近于 X 的眞实均值 $E[X]$ 和眞实方差 $V_{ar}[X]$ 时，b_i 将很小。当然，我们可类似定义 $b_j(j>4)$，此时它们可略去不计。另外，关于 $b_i(i=1,\cdots,4)$ 的计算问题，实际上是计算 X 的各阶矩。在眞实分布未知的場合，这些矩可以通过近似的方法估计出。

三、随机向量的若干矩公式及利用正态逼近的非线性滤波方法

考虑如下非线性系统

$$X_{K+1}=f(X_K,k)+W_K \tag{3.1}$$

$$Z_K=h(X_K,k)+V_K \tag{3.2}$$

其中 X_K 和 Z_K 分别为 n 維状态向量和 m 維观测向量，$\{W_K\}$、$\{V_K\}$ 为互相独立的正态白噪声序列，$E[W_K]=0$，$V_{ar}[W_K]=Q_K$, $E[V_K]=0$，$V_{ar}[V_K]=R_K$. 现要在给定观测集 $Z^K=(Z_1,\cdots,Z_K)$ 的条件下，求状态向量 X_K 的条件均值估计

$$\hat{X}_{K/K}=E[X_K|Z^K], \tag{3.3}$$

解决上述问题的途径有许多，下面讨论利用正态逼近的方法。

首先我们给出有关正态随机向量的若干矩公式。设 $X\sim N_n(x;0,P)$，则有（证明见

附录）

1° $E[(X\otimes X)(X\otimes X)^T]=P\otimes P+cs(P)rs(P)+(P\otimes P_{\cdot 1}\cdots P\otimes P_{\cdot n})$ (3.4)

2° 对任意的 $n\times n^2$ 矩阵 $A=(a_{ij})$

$$E[XX^TA(X\otimes X)]=E[X(X\otimes X)^TA^TX]$$
$$=PAcs(P)+(P\otimes rs(P))cs(A)+(rs(P)\otimes P)rs^T(A) \quad (3.5)$$

特别，若 A 的元素满足

$$a_{i,kn+j}=a_{i,(j-1)n+k-1},\ k=0,1,\cdots,n-2,i=1,\cdots,n,j=k+2,\cdots,n \quad (3.6)$$

则

$$E[XX^TA(X\otimes X)]=PAcs(P)+2(P\otimes rs(P))cs(A) \quad (3.7)$$

3° 对于 $n\times n$ 对称阵 A

$$E[XX^T(X^TAX)]=tr(AP)P+2PAP \quad (3.8)$$

4° 对于任意 n^2 維列向量 $Y=(y_j)$

$$E[XX^T(X\otimes X)^TY]=Y^Tcs(P)P+(I_n\otimes Y^T)(cs(P)\otimes P_{\cdot 1}\cdots cs(P)\otimes P_{\cdot n})$$
$$+\begin{pmatrix} rs(P)\otimes P_{1\cdot} \\ \vdots \\ rs(P)\otimes P_{n\cdot}\end{pmatrix}(I_n\otimes Y) \quad (3.9)$$

特别，若 Y 的元素满足

$$y_{kn+j}=y_{(j-1)n+k+1},\ k=0,1,\cdots,n-2,j=k+2,\cdots,n \quad (3.10)$$

则

$$E[XX^T(X\otimes X)^TY]=Y^Tcs(P)P+2(I_n\otimes Y^T)(cs(P)\otimes P_{\cdot 1}\cdots cs(P)\otimes P_{\cdot n}) \quad (3.11)$$

5° 对于任意的 n 維行向量 Y

$$E[XYX]=PY^T \quad (3.12)$$

利用上述结果来解决非线性系统状态估计的正态追近问题。设 $p(\cdot)$ 为分布密度函数，则由 Bayes 公式

$$p(x_K|Z^K)=p(x_K|Z^{K-1})p(Z_K|x_K)/p(Z_K|Z^{K-1}) \quad (3.13)$$

其中 $Z^{K-1}=(Z_1\cdots Z_{K-1})$，$Z^K=(Z_1,\cdots,Z_{K-1},Z_K)$。下面分别对上式右端各项进行研究：

(1) $p(x_K|Z^{K-1})$：在 $\hat{X}_{K/K-1}$ 和 $P_{K/K-1}$ 处对 $p(x_K|Z^{K-1})$ 做正态追近并略去三阶以上项，则有

$$p(x_K|Z^{K-1})=N_n(x_K;\ \hat{X}_{K/K-1},P_{K/K-1})[1+H_1^T(x_K-\hat{X}_{K/K-1})b_1$$
$$+\frac{1}{2}tr(H_2(x_K-\hat{X}_{K/K-1})b_2)]$$

其中系数

$$b_1=\int(x_K-\hat{X}_{K/K-1})p(x_K|Z^{K-1})dx_K$$

$$b_2=\int[(x_K-\hat{X}_{K/K-1})(x_K-\hat{X}_{K/K-1})^T-P_{K/K-1}]\,p(x_K|Z^{K-1})\,dx_K$$

选择适当的$\hat{X}_{K/K-1}$和$P_{K/K-1}$，使b_1和b_2为零，则有

$$\hat{X}_{K/K-1}=\int x_K p(x_K|Z^{K-1})\,dx_K=\int f(x_{K-1})\,p(x_{K-1}|Z^{K-1})\,dx_{K-1} \tag{3.14}$$

$$\begin{aligned}P_{K/K-1}&=\int(x_K-\hat{X}_{K/K-1})(x_K-\hat{X}_{K/K-1})^T p(x_K|Z^{K-1})\,dx_K\\&=Q_{K-1}+\int(f(x_{K-1})-\hat{X}_{K/K-1})(f(x_{K-1})-\hat{X}_{K/K-1})^T p(x_{K-1}|Z^{K-1})\,dx_{K-1}\end{aligned} \tag{3.15}$$

则此时有

$$p(x_K|Z^{K-1})=N_n(x_K;\ \hat{X}_{K/K-1},\ P_{K/K-1}) \tag{3.16}$$

其中$\hat{X}_{K/K-1}$和$P_{K/K-1}$分别由（3.14）和（3.15）确定。

(2) $p(Z_K|Z^{K-1})$：同样，将$p(Z_K|Z^{K-1})$在$\hat{Z}_{K/K-1}$和$C_{K/K-1}$处做正态迫近并略去三阶以上项，并适当选择$\hat{Z}_{K/K-1}$和$C_{K/K-1}$使系数b_1和b_2为零。则有

$$p(Z_K|Z^{K-1})=N_m(Z_K;\ \hat{Z}_{K/K-1},\ C_{K/K-1}) \tag{3.17}$$

其中

$$\hat{Z}_{K/K-1}=\int h(x_K)\,p(x_K|Z^{K-1})\,dx_K \tag{3.18}$$

$$C_{K/K-1}=\int(h(x_K)-\hat{Z}_{K/K-1})(h(x_K)-\hat{Z}_{K/K-1})^T p(x_K|Z^{K-1})\,dx_K+R_K \tag{3.19}$$

(3) $p(Z_K|x_K)$：将$p(Z_K|x_K)$在$\hat{Z}_{K/K-1}$和$C_{K/K-1}$做正态迫近并略去三阶以上项，则有

$$\begin{aligned}p(Z_K|x_K)=&N_m(Z_K;\ \hat{Z}_{K/K-1},C_{K/K-1})[1+H_1^T(Z_K-\hat{Z}_{K/K-1})b_1\\&+\frac{1}{2}tr(H_2(Z_K-\hat{Z}_{K/K-1})b_2)]\end{aligned}$$

其中，若记$\nu_K=Z_K-\hat{Z}_{K/K-1}$，则

$$H_1(\nu_K)=C_{K/K-1}^{-1}\nu_K$$

$$H_2(\nu_K)=-C_{K/K-1}^{-1}+C_{K/K-1}^{-1}\nu_K\nu_K^TC_{K/K-1}^{-1}$$

$$b_1=\int\nu_K p(Z_K|x_K)\,dZ_K=h(x_K)-\hat{Z}_{K/K-1}$$

$$\begin{aligned}b_2&=\int[-C_{K/K-1}+(\nu_K\nu_K^T)]\,p(Z_K|x_K)\\&=-C_{K/K-1}+R_K+(h(x_K)-\hat{Z}_{K/K-1})(h(x_K)-\hat{Z}_{K/K-1})^T\end{aligned}$$

于是

$$\begin{aligned}p(Z_K|x_K)=&N_m(Z_K;\ \hat{Z}_{K/K-1},\ C_{K/K-1})[1+(h(x_K)-\hat{Z}_{K/K-1})^TC_{K/K-1}\nu_K\\&+\frac{1}{2}tr[(-C_{K/K-1}+R_K+(h(x_K)-\hat{Z}_{K/K-1})(h(x_K)-\hat{Z}_{K/K-1})^T)A(\nu_K)]]\end{aligned} \tag{3.20}$$

其中

$$A(\nu_K)=-C_{K/K-1}^{-1}+C_{K/K-1}^{-1}\nu_K\nu_K^T C_{K/K-1}^{-1} \tag{3.21}$$

综上所述，我们得到关于 $p(x_K|Z^K)$ 的正态逼近公式

$$p(x_K|Z^K)=N_n(x_K;\ \hat{X}_{K/K-1},P_{K/K-1})\rho \tag{3.22}$$

其中

$$\rho=1+(h(x_K)-\hat{Z}_{K/K-1})^T C_{K/K-1}^{-1}\nu_K+\frac{1}{2}tr[(-C_{K/K-1}+R_K$$

$$+(h(x_K)-\hat{Z}_{K/K-1})(h(x_K)-\hat{Z}_{K/K-1})^T)A(\nu_K)] \tag{3.23}$$

下面具体推导 $\hat{X}_{K/K}$ 的递推计算公式。首先将非线性函数 $h(X_K)$ 在 $\hat{X}_{K/K-1}$ 处展开至二阶项

$$h(X_K)=h(\hat{X}_{K/K-1})+h_X\tilde{X}_{K/K-1}+\frac{1}{2}h_{XX}(\tilde{X}_{K/K-1}\otimes\tilde{X}_{K/K-1})$$

其中 h_X 为 Jacobi 阵，h_{XX} 为 Hesse 阵，$\tilde{X}_{K/K-1}=X_K-\hat{X}_{K/K-1}$。于是

$$\hat{Z}_{K/K-1}=\int h(x_K)p(x_K|Z^{K-1})dx_K$$

$$=\int h(x_K)N_n(x_K;\ \hat{X}_{K/K-1},P_{K/K-1})dx_K$$

$$=h(\hat{X}_{K/K-1})+\frac{1}{2}h_{XX}cs(P_{K/K-1}) \tag{3.24}$$

$$C_{K/K-1}=\int[h_X\tilde{x}_{K/K-1}+\frac{1}{2}h_{XX}(\tilde{x}_{K/K-1}\otimes\tilde{x}_{K/K-1}-cs(P_{K/K-1}))+V_K]$$

$$\cdot\left[h_X\tilde{x}_{K/K-1}+\frac{1}{2}h_{XX}(\tilde{x}_{K/K-1}\otimes\tilde{x}_{K/K-1}-cs(P_{K/K-1}))+V_K\right]^T$$

$$\cdot N_n(\tilde{x}_{K/K-1};\ 0,P_{K/K-1})d\tilde{x}_{K/K-1}+R_K$$

$$=h_XP_{K/K-1}h_X^T+\frac{1}{4}h_{XX}[P_{K/K-1}\otimes P_{K/K-1}+(P_{K/K-1}\otimes P_{\cdot 1}(K/K-1)\cdots$$

$$P_{K/K-1}\otimes P_{\cdot n}(K/K-1))]h_{XX}^T+R_K \tag{3.25}$$

其中 $P_{\cdot i}(K/K-1)$，$i=1,\cdots,n$ 为 $P_{K/K-1}$ 的第 i 列元素。于是

$$\rho=1-\frac{1}{2}tr[(C_{K/K-1}^*+h_XP_{K/K-1}h_X^T)A(\nu_K)]+\tilde{x}_{K/K-1}^T h_X^T C_{K/K-1}^{-1}\nu_K$$

$$+\frac{1}{2}B^T(\tilde{x}_{K/K-1})h_{XX}^T C_{K/K-1}^{-1}\nu_K+\frac{1}{2}\tilde{x}_{K/K-1}^T h_X^T A(\nu_K)h_X\tilde{x}_{K/K-1}$$

$$+\frac{1}{4}\tilde{x}_{K/K-1}^T h_X^T A(\nu_K)h_{XX}B(\tilde{x}_{K/K-1})+\frac{1}{4}B^T(\tilde{x}_{K/K-1})h_{XX}^T A(\nu_K)h_X\tilde{x}_{K/K-1}$$

$$+\frac{1}{8}B^T(\tilde{x}_{K/K-1})h_{XX}^T A(\nu_K)h_{XX}B(\tilde{x}_{K/K-1}) \tag{3.26}$$

其中

$$C^*_{K/K-1}=\frac{1}{4}h_{XX}[P_{K/K-1}\otimes P_{K/K-1}+(P_{K/K-1}\otimes P_{\cdot 1}(K/K-1)\cdots$$

$$P_{K/K-1}\otimes P_{\cdot n}(K/K-1))]h^T_{XX} \tag{3.27}$$

$$B(\tilde{x}_{K/K-1})=\tilde{x}_{K/K-1}\otimes\tilde{x}_{K/K-1}-cs(P_{K/K-1}) \tag{3.28}$$

若记

$$\alpha_K=\int(x_K-\hat{X}_{K/K-1})\,p(x_K/Z^K)\,dx_K$$

考虑到奇次项积分为零，有

$$\alpha_K=\int\tilde{x}_{K/K-1}N_n(\tilde{x}_{K/K-1};\ 0,P_{K/K-1})\,\rho d\tilde{x}_{K/K-1}$$

$$=\int\tilde{x}_{K/K-1}(\tilde{x}_{K/K-1}h^T_XC^{-1}_{K/K-1}\nu_K+\frac{1}{4}\tilde{x}^T_{K/K-1}h^T_XA(\nu_K)h_{XX}B(\tilde{x}_{K/K-1})$$

$$+\frac{1}{4}B(\tilde{x}_{K/K-1})h^T_{XX}A(\nu_K)h_X\tilde{x}_{K/K-1})N_n(\tilde{x}_{K/K-1};\ 0,P_{K/K-1})\,d\tilde{x}_{K/K-1}$$

利用本节给出的矩公式，并注意到 $h^T_XA(\nu_K)h_{XX}$ 满足条件（3.6），则

$$\alpha_K=P_{K/K-1}h^T_XC^{-1}_{K/K-1}\nu_K+(P_{K/K-1}\otimes rs(P_{K/K-1}))cs(h^T_XA(\nu_K))h_{XX}$$

而由 α_K 的定义，显然

$$\alpha_K=\hat{X}_{K/K}-\hat{X}_{K/K-1}$$

于是

$$\hat{X}_{K/K}=\hat{X}_{K/K-1}+P_{K/K-1}h^T_XC^{-1}_{K/K-1}\nu_K+(P_{K/K-1}\otimes rs(P_{K/K-1}))$$
$$\cdot cs(h^T_XA(\nu_K)h_{XX}) \tag{3.29}$$

在另一方面，估值误差方差阵为

$$P_{K/K}=\int(x_K-\hat{X}_{K/K})(x_K-\hat{X}_{K/K})^Tp(x_K|Z^K)\,dx_K$$

$$=\int\tilde{x}_{K/K-1}\tilde{x}^T_{K/K-1}p(x_K|Z^K)\,dx_K-\alpha_K\alpha^T_K$$

同样注意到关于 $\tilde{x}_{K/K-1}$ 的奇数项积分为零并略去六阶以上项可得

$$P_{K/K}=P_{K/K-1}-P_{K/K-1}h^T_XC^{-1}_{K/K-1}h_XP_{K/K-1}+D(P_{K/K-1}) \tag{3.30}$$

其中

$$D(P_{K/K-1})=[I_n\otimes(h^T_{XX}C^{-1}_{K/K-1}\nu_K)^T][cs(P_{K/K-1})\otimes P_{\cdot 1}(K/K-1)\cdots$$

$$cs(P_{K/K-1})\otimes P_{\cdot n}(K/K-1)]-\frac{1}{2}tr[C^*_{KK/K-1}A(\nu_K)] \tag{3.31}$$

以上我们已经得到了 $\hat{X}_{K/K}$ 和 $P_{K/K}$ 的计算公式，特别若观测方程是线性的，即

$$h(x_K)=H_Kx_K$$

则易推得此时

$$\hat{X}_{K/K}=\hat{X}_{K/K-1}+P_{K/K-1}H^T_KC^{-1}_{K/K-1}\nu_K$$
$$P_{K/K}=P_{K/K-1}-P_{K/K-1}H^T_KC^{-1}_{K/K-1}H_KP_{K/K-1}$$

$$C_{K/K-1}=H_K P_{K/K-1} H_K^T + R_K$$

这正是最佳线性滤波的结论。

为了完成滤波的递推运算，剩下的工作是求出$\hat{X}_{K+1/K}$和$P_{K+1/K}$. 同样，将非线性函数$f(X_K)$在$\hat{X}_{K/K}$处展开并取二阶项

$$f(X_K)=f(\hat{X}_{K/K})+f_X\tilde{X}_{K/K}+\frac{1}{2}f_{XX}(\tilde{X}_{K/K}\otimes\tilde{X}_{K/K})$$

其中f_X和f_{XX}分别是 Jacobi 阵和 Hesse 阵，$\tilde{X}_{K/K}=X_K-\hat{X}_{K/K}$.则由 (3.14)、(3.15)

$$\hat{X}_{K+1/K}=\int f(x_K)\,p(x_K|Z^K)\,dx_K$$
$$=f(\hat{X}_{K/K})+\frac{1}{2}f_{XX}cs(P_{K/K}) \tag{3.32}$$

$$P_{K+1/K}=\int(f(x_K)-\hat{X}_{K+1/K})(f(x_K)-\hat{X}_{K+1/K})^T p(x_K|Z^K)\,dx_K+Q_K$$
$$=f_X P_{K/K} f_X^T+Q_K+U_K \tag{3.33}$$

其中

$$U_K=\int\Big[\frac{1}{2}f_{XX}B(\tilde{x}_{K/K})\tilde{x}_{K/K}^T f_X^T+\frac{1}{2}f_X\tilde{x}_{K/K}B^T(\tilde{x}_{K/K})f_{XX}^T$$
$$+\frac{1}{4}f_{XX}B(\tilde{x}_{K/K})B^T(\tilde{x}_{K/K})f_{XX}^T\Big]p(x_K|Z^K)\,dx_K$$

$$B(\tilde{x}_{K/K})=\tilde{x}_{K/K}\otimes\tilde{x}_{K/K}-cs(P_{K/K})$$

可以把U_K化成关于$\tilde{x}_{K/K-1}$的积分而求出其具体形式，但这样做结果比较复杂。为简单起见，认为$p(x_K|Z^K)\doteq N_n(x_K;\ \hat{X}_{K/K},P_{K/K})$，则

$$U_K\doteq\frac{1}{4}\int f_{XX}B(\tilde{x}_{K/K})B^T(\tilde{x}_{K/K})f_{XX}^T N_n(\tilde{x}_{K/K};\ 0,P_{K/K})\,d\tilde{x}_{K/K}$$

$$=\frac{1}{4}f_{XX}[P_{K/K}\otimes P_{K/K}+(P_{K/K}\otimes P_{.1}(K/K)\cdots P_{K/K}\otimes P_{.n}(K/K))]f_{XX}^T \tag{3.34}$$

其中$P_{.i}(K/K)$为$P_{K/K}$的第i列元素，$i=1,\cdots,n$. 显然，若状态方程是线性的，即

$$f(X_K)=\phi_{K+1,K}X_K$$

则有

$$\hat{X}_{K+1/K}=\phi_{K+1,K}\hat{X}_{K/K}$$
$$P_{K+1/K}=\phi_{K+1,K}P_{K/K}\phi_{K+1,K}^T+Q_K$$

这也是我们所熟悉的最佳线性滤波公式。

上述关于非线性滤波的迫近公式归纳如下：

$$\hat{X}_{K+1/K}=f(\hat{X}_{K/K})+\frac{1}{2}f_{XX}cs(P_{K/K}) \tag{I}$$

$$\hat{X}_{K+1/K+1}=\hat{X}_{K+1/K}+P_{K+1/K}h_X^T C_{K+1/K}^{-1}\nu_{K+1}+(P_{K+1/K}$$
$$\otimes rs(P_{K+1/K}))cs(h_X^T A(\nu_{K+1})h_{XX}) \tag{II}$$

$$P_{K+1/K}=f_X P_{K/K} f_X^T+Q_K+U_K \tag{Ⅲ}$$

$$P_{K+1/K+1}=P_{K+1/K}-P_{K+1/K}h_X^T C_{K+1/K}^{-1} h_X P_{K+1/K}+D(P_{K+1/K}) \tag{Ⅳ}$$

其中 $C_{K+1/K}$，$A(\nu_{K+1})$，$U_K, D(P_{K+1/K})$ 分别由（3.25）、（3.21）、（3.34）和（3.31）所决定。

本文仅从理论上对非线性滤波的正态追近方法进行了探讨。由于结果中考虑到了 $P_{K/K}$ 和 $P_{K+1/K}$ 的平方项，增加了计算量，但这些计算（如矩阵的拉直、Kronecker 积等）在计算机上是很容易实现的。虽然本文沒有涉及具体的计算，但正如其它非线性滤波方法一样，本文提出的算法不可能一致地优于其它方法，其适用性是有一定条件的。

附　　录

一、若干导数公式的证明

1° 注意到

$$\frac{\partial}{\partial x_i}[e^{-\frac{1}{2}x^TPx}]=-\sum_{j=1}^{n}P_{ij}x_j e^{-\frac{1}{2}x^TPx}=-(Px)_i e^{-\frac{1}{2}x^TPx}$$

则

$$\frac{\partial}{\partial x}(Ae^{-\frac{1}{2}x^TPx})=\begin{pmatrix}\frac{\partial}{\partial x_1}(Ae^{-\frac{1}{2}x^TPx})\\ \vdots \\ \frac{\partial}{\partial x_n}(Ae^{-\frac{1}{2}x^TPx})\end{pmatrix}=-(Px)\otimes Ae^{-\frac{1}{2}x^TPx}$$

同理可证（2.6）。

2° 注意到 Pxx^TP 的第 i 行（$i=1,\cdots,n$）为 $\sum_{j=1}^{n}P_{ij}x_jx^TP$，则

$$\frac{\partial}{\partial x_k}\left[\sum_{j=1}^{n}P_{ij}x_jx^TP\right]=P_{ik}x^TP+(Px)_iP_k, \qquad k=1,\cdots,n$$

卽有

$$\frac{\partial}{\partial x_K}(Pxx^TP)=P_{\cdot K}x^TP+P_{K\cdot}\otimes(PX)$$

从而

$$\frac{\partial}{\partial x}(Pxx^TP)=\begin{pmatrix}P_{\cdot 1}x^TP\\ \vdots \\ P_{\cdot n}x^TP\end{pmatrix}+\begin{pmatrix}P_{1\cdot}\otimes(Px)\\ \vdots \\ P_{n\cdot}\otimes(Px)\end{pmatrix}=cs(P)(Px)^T+P\otimes(Px)$$

同理可证（2.8）.

3° 因为 $(Px)\otimes P$ 的第 i 块元素（$i=1,\cdots,n$）为 $\sum_{j=1}^{n}P_{ij}x_jP$，而

$$\frac{\partial}{\partial x_k}\left(\sum_{j=1}^{n} P_{ij}x_jP\right)=P_{ik}P \qquad k=1,\cdots,n$$

故有

$$\frac{\partial}{\partial x^T}[(Px)\otimes P]=(P_{.1}\otimes P\cdots P_{.n}\otimes P)=P\otimes P$$

4° 由于

$$cs(P)(Px)^T=\begin{pmatrix}\sum_{j=1}^{n} P_{.1}x_j(P_{1j}\cdots P_{nj})\\ \vdots \\ \sum_{j=1}^{n} P_{.n}x_j(P_{1j}\cdots P_{nj})\end{pmatrix}$$

则

$$\frac{\partial}{\partial x_i}[cs(P)(Px)^T]=\begin{bmatrix}P_{.1}P_{i.}\\ \vdots \\ P_{.n}P_{i.}\end{bmatrix}=cs(P)P_{i.}$$

故

$$\frac{\partial}{\partial x^T}[cs(P)(Px)^T]=(cs(P)P_{1.}\cdots cs(P)P_{n.})=cs(P)rs(P)$$

5° 因为

$$\frac{\partial}{\partial x_i}[P\otimes(Px)]=\begin{bmatrix}P_{11}P_{.i} & \cdots & P_{1n}P_{.i}\\ \vdots & \ddots & \vdots \\ P_{1n}P_{.i} & \cdots & P_{nn}P_{.i}\end{bmatrix}=P\otimes P_{.i}$$

故

$$\frac{\partial}{\partial x^T}[P\otimes(Px)]=(P\otimes P_{.1}\cdots P\otimes P_{.n})$$

6° $$\frac{\partial}{\partial x^T}[Y\otimes A]=\begin{pmatrix}\frac{\partial y_1}{\partial x_1}A+y_1\frac{\partial A}{\partial x_1}\cdots\frac{\partial y_1}{\partial x_n}A+y_1\frac{\partial A}{\partial x_n}\\ \vdots \quad \ddots \quad \vdots \\ \frac{\partial y_n}{\partial x_1}A+y_n\frac{\partial A}{\partial x_1}\cdots\frac{\partial y_n}{\partial x_n}A+y_n\frac{\partial A}{\partial x_n}\end{pmatrix}=\frac{\partial Y}{\partial x^T}\otimes A+Y\otimes\frac{\partial A}{\partial x^T}$$

二、若干矩公式的证明

1° 注意到 $(X\otimes X)(X\otimes X)^T$ 的第 ij 块元素 $(i,j=1,\cdots,n)$ 为 $X_iX_jXX^T$，而

$$E[X_iX_jXX^T]=P_{ij}P+P_{.i}P_{j.}+P_{.j}P_{i.}$$

则

$$E[(X\otimes X)(X\otimes X)^T]=\begin{pmatrix} P_{11}P & \cdots & P_{1n}P \\ \vdots & \ddots & \vdots \\ P_{1n}P & \cdots & P_{nn}P \end{pmatrix}+\begin{pmatrix} P_{.1}P_{1.} & \cdots & P_{.1}P_{n.} \\ \vdots & \ddots & \vdots \\ P_{.n}P_{1.} & \cdots & P_{.n}P_{n.} \end{pmatrix}$$

$$+\begin{pmatrix} P_{.1}P_{.1} & \cdots & P_{.n}P_{1.} \\ \vdots & \ddots & \vdots \\ P_{.1}P_{n.} & \cdots & P_{.n}P_{n.} \end{pmatrix}$$

$$=P\otimes P+cs(P)rs(P)+(P\otimes P_{.1}\cdots P\otimes P_{.n})$$

2° 注意到$XX^TA(X\otimes X)$ 的第 i 行元素 ($i=1,\cdots,n$) 为 $X_iX^TA(X\otimes X)$，且

$$E[X_iX^TA(X\otimes X)]=\sum_{j=1}^{n}a_{j.}E[X_iX_j(X\otimes X)]$$

$$=\sum_{j=1}^{n}a_{j.}[P_{ij}cs(P)+P_{.i}\otimes P_{.j}+P_{.j}\otimes P_{.i}]$$

其中

$$\sum_{j=1}^{n}a_{j.}P_{ij}cs(P)=\sum_{j=1}^{n}(P_{ij}a_{j1}\cdots P_{ij}a_{jn^2})cs(P)=\{PAcs(P)\}_i$$

$$\sum_{j=1}^{n}a_{j.}(P_{.j}\otimes P_{.i})=\sum_{j=1}^{n}(P_{j.}\otimes P_{i.})a_{j.}^T=\{(rs(P)\otimes P)rs^T(A)\}_i$$

$$\sum_{j=1}^{n}a_{j.}(P_{.i}\otimes P_{.j})=\sum_{j=1}^{n}(a_{1j}P_{i1}P_{j1}+\cdots+a_{1,n^2-n+j}P_{in}P_{j1})$$

$$+\sum_{j=1}^{n}(a_{2j}P_{i1}P_{j2}+\cdots+a_{2,n^2-n+j}P_{in}P_{j2})+\cdots$$

$$+\sum_{j=1}^{n}(a_{nj}P_{i1}P_{jn}+\cdots+a_{n,n^2-n+j}P_{in}P_{jn})$$

$$=(P_{i.}\otimes rs(P))cs(A)=\{(P\otimes rs(P))cs(A)\}_i$$

故 (3.5) 成立。特别若A满足 (3.6)，则

$$\sum_{j=1}^{n}a_{j.}(P_{.j}\otimes P_{.i})=\sum_{j=1}^{n}a_{j.}(P_{.i}\otimes P_{.j})\qquad i=1,\cdots,n$$

也即有

$$(P\otimes rs(P))cs(A)=(rs(P)\otimes P)rs^T(A).$$

3° 对于任意 i, j ($i,j=1,\cdots,n$)

$$E[\{XX^T(X^TAX)\}_{ij}]=\sum_{k=1}^{n}\sum_{l=1}^{n}E[X_kX_lX_iX_j]a_{kl}$$

$$=\sum_{k,l=1}^{n}a_{kl}(P_{kl}P_{ij}+P_{ki}P_{lj}+P_{kj}P_{li})$$

$$=\sum_{k,l=1}^{n}a_{kl}P_{kl}P_{ij}+2\sum_{k,l=1}^{n}P_{ki}P_{lj}$$

$$=tr(AP)P_{ij}+2P_{i.}(a_{.1}\cdots a_{.n})\begin{pmatrix} P_{1j} \\ \vdots \\ P_{nj} \end{pmatrix}$$

$$=tr(AP)P_{ij}+2(PAP)_{ij}$$

4° $E[XX^T(X\otimes X)^TY]=E[Y^T(X\otimes X)XX^T]$

$$=\sum_{j=1}^{n}E[(y_jX_1X_j+\cdots+y_{n^2-n+j}X_nX_j)XX^T]$$

$$=\sum_{j=1}^{n}[(y_jP_{1j}+\cdots+y_{n^2-n+j}P_{nj})P+(y_jP_{.1}P_{j.}+\cdots+y_{n^2-n+j}P_{.n}P_{j.})$$

$$+(y_jP_{.j}P_{1.}+\cdots+y_{n^2-n+j}P_{.j}P_{n.})]$$

$$=Y^Tcs(P)P+\sum_{i=1}^{n}(Y^T(cs(P)\otimes P_{.i}))+\sum_{i=1}^{n}(rs(P)\otimes P_{i.})Y$$

$$=Y^Tcs(P)P+(I_n\otimes Y^T)(cs(P)\otimes P_{.1}\cdots cs(P)\otimes P_{.n})$$

$$+\begin{pmatrix}rs(P)\otimes P_{1.}\\ \vdots\\ rs(P)\otimes P_{n.}\end{pmatrix}(I_n\otimes Y)$$

特别若Y满足条件件（3.10），则此时

$$y_j(P_{.j}\otimes P_{.i})=y_i(P_{.i}\otimes P_{.j})$$

即

$$(I_n\otimes Y^T)(cs(P)\otimes P_{.1}\cdots cs(P)\otimes P_{.n})=\begin{pmatrix}rs(P)\otimes P_{1.}\\ \vdots\\ rs(P)\otimes P_{n.}\end{pmatrix}(I_n\otimes Y)$$

5° 事实上

$$E[XYX]=E\begin{bmatrix}\sum_{j=1}^{n}y_jX_1X_j\\ \vdots\\ \sum_{j=1}^{n}y_jX_nX_j\end{bmatrix}=\begin{pmatrix}P_{1.}Y^T\\ \vdots\\ P_{n.}Y^T\end{pmatrix}=PY^T$$

参 考 文 献

[1] 辻・高田・神谷，观测予测误差の有理式を补正项にもつ离散形非线形フィルタ，システムと制御，15—8，1971.

[2] C. J. Masreliez, Approximation Non-Gaussian Filtering with Linear State and Observation Relations, IEEE Trans. AC—20, pp. 107—110, 1975.

[3] 张尧庭，方开泰，多元统计分析引论，科学出版社，1982.

Normal Approximation, Moment Formulas with Application to Nonlinear Filtering

Mao Ning

Abstract

In this paper some partial derivatives of a matrix function with respect to a vector variable are derived first. By using of these derivatives, the representation of the normal approximation for any distribution density function of a random vector is given. Next some moment formulas of normal random vector are developed. An approximation method of state estimation for a nonlinear system is discussed as an application of the above results.