文章编号: 1001 - 2486(2003)01 - 0097 - 05

多层验前正态总体动态参数的 Bayes 融合估计*

张金槐

(国防科技大学人文与管理学院, 湖南 长沙 410073)

摘 要: 在多层验前信息下,给出了正态总体动态参数的 Bayes 融合估计。进一步, 用线模型下的递推估计方法,给出了递推解。对于实际应用中的一些问题, 提出了处置方法, 并以实例作了说明。

关键词: 多层验前; Bayes 估计

中图分类号: O212 文献标识码: A

Bayes Fusion Estimation of Variable Distribution Parameters under the Hierarchical Prior Informations of the Normal Sample

ZHANG Jin-huai

(College of Humanities and Management, National Univ. of Defense Technology, Changsha 410073, China)

Abstract: The estimation of the variable distribution parameters of the normal sample under the hierarchical prior informations is discussed. At the same time, the recursive form of the estimator of the state variables for the linear model is constructed. Furthermore, as for the application of this method, some practicle problems are considered.

Key words: hierarchical prior; Bayes estimation

1 问题的提出

对于武器装备的性能参数进行试验分析或鉴定时, 人们总是设法运用各种可以利用的验前信息, 以在较少量的现场试验条件下, 作出科学的试验结果分析。早在 20 世纪 70 年代, Goods[1] 等就提出了多层验前信息的概念, 并将其运用于 Bayes 分析。此外, 在工程实践中, 试验中所获取到的子样, 并不总是独立同分布(i. i. d) 的。例如, 每次(或阶段) 试验的条件、场景不同, 而对于试验中暴露的问题进行改进, 或对于出现的故障进行排除, 然后再进入下一次(阶段) 试验。这样, 试验中被考核的参数将是可变的。在这种情况下, 便提出了多层验前信息下的动态参数的估计问题。

关于动态分布参数的 Bayes 估计问题, J. Berger[1], C. Morris[2] 和 L. M. Berlinear[3] 等作了研究。一般地, 总是利用验前信息建立参数的动态模型, 然后, 进一步按照 Bayes 统计推断思想进行统计分析。关键问题是多层验前信息下的信息融合问题。在文献[1] 中, 对分层信息的 Bayes 分析有较一般的讨论。然而在动态参数建模以及在实际问题中的具体应用等方面, 还有不少工作要做。下面将对上述一些问题进行探索, 对于应用中的一些问题提出处置的方法。

2 模型及 Bayes 估计

设有 n 次独立试验的结果 $y_1, \cdots, y_n$, 这里 $y_i \mid \theta_i \sim N(\theta_i, \sigma_i^2)$, σ_i^2 为已知, θ_i 为分布参数, 它具有验前分布, 如 $\theta_i \mid \mu \sim N(H_i\mu, \sigma_{\theta_i}^2)$, 其中 μ 为 $p \times 1$ 超参数, H_i 为 $1 \times p$ 已知系数向量, $\sigma_{\theta_i}^2$ 为已知, 超参数 μ 的验前分布为 $N_p(\beta, \Sigma_\mu)$, 这里 β, Σ_μ 为已知。在上述假设下, 要由 $y_1, \cdots, y_n$ 去估计 $\theta_1, \cdots, \theta_n$。

记

* 收稿日期: 2003 - 02 - 27
作者简介: 张金槐(1930—), 男, 教授。

$$Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad \theta = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_n \end{pmatrix}$$

则

$$Y \mid \theta \sim N_n(\theta, R) \tag{1}$$

其中 $R = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{pmatrix}$, R 为已知。θ 为 $n \times 1$ 分布参数向量, 且有验前分布, 由

$$\theta_i \mid \mu \sim N(H_i \mu, \sigma_{\theta_i}^2), \quad i = 1, \cdots, n$$

记

$$H = \begin{pmatrix} H_1 \\ \vdots \\ H_n \end{pmatrix}$$

于是由验前知识得, $\theta \mid \mu \sim N_n(H\mu, \Sigma)$, 其中 H, Σ 是已知的。

超参数 μ 具有验前分布 $N_p(\beta, \Sigma_\mu)$, 其中 β, Σ_μ 为已知。

在上述模型下, 要由 Y 作出 θ 的 Bayes 估计, 这是一个具有二层验前信息下的 Bayes 估计问题。此时的 Bayes 估计要计算条件期望 $E[\theta / Y]$, 并估计 $Var[\theta / Y]$。

将模型稍作变换, 令

$$y'_i = y_i / \sigma_i, \quad \theta'_i = \theta_i / \sigma_i$$

且记

$$Y' = \begin{pmatrix} y'_1 \\ \vdots \\ y'_n \end{pmatrix}, \quad \theta' = \begin{pmatrix} \theta'_1 \\ \vdots \\ \theta'_n \end{pmatrix}$$

则

$$Y' \mid \theta' \sim N_n(\theta', I_n) \tag{1'}$$

记

$$H_i / \sigma_i = H'_i, \quad H' = \begin{pmatrix} H'_i \\ \vdots \\ H'_n \end{pmatrix}$$

则有

$$\theta' \mid \mu \sim N_n(H'\mu, \Sigma')$$

超参数 μ 仍有验前分布, $\mu \sim N_p(\beta, \Sigma_\mu)$。 (3)

由式(2), 将 θ' 看作关于 μ 的观测方程, 此时, θ' 的概率密度为

$$\pi(\theta') = \int_{\Theta_\mu} p(\theta' / \mu) p(\mu) \mathrm{d}\mu, \quad \Theta_\mu \text{ 为 } \mu \text{ 的域}$$

由于 $p(\theta' / \mu) \sim N_n(H'\mu, \Sigma_{\delta'}), \quad p(\mu) \sim N_p(\beta, \Sigma_\mu)$

因此, $\pi(\theta')$ 也是正态分布, 且可知[4]

$$\pi(\theta') \sim N_n(H'\beta, \Sigma_{\delta'} + H'\Sigma_\mu H'^{\mathrm{T}})。$$

再由式(1)′, θ' 的验后分布密度 $p(\theta' / Y)$ 可由 θ' 的分布及 Y' / θ 的分布 $p(Y' / \theta')$ 确定。由于 θ' 的验前分布为共轭正态分布, 因此, 只需给出 $\theta' \mid Y'$ 的期望和方差阵就可以了。可以算得[5]

$$E[\theta' / Y'] = [(\Sigma_{\delta'} + H'\Sigma_\mu H'^{\mathrm{T}})^{-1} + I]^{-1}[(\Sigma_{\delta'} + H'\Sigma_\mu H'^{\mathrm{T}})^{-1}H'\beta + Y'] \tag{4}$$

$$Var[\theta' / Y'] = [(\Sigma_{\delta'} + H'\Sigma_\mu H'^{\mathrm{T}})^{-1} + I]^{-1} \tag{5}$$

式(4)、(5) 就是所需要的 Bayes 估计公式。

3 多层验前信息下, Bayes 融合估计的递推形式解

为了应用中的需要, 往往要给出递推形式的估计式, 即随着观测量 y_i 的出现, 及时给出 Q_i 的 Bayes 融合估计。这种情况, 在试验分析中常称之为"试试看看, 看看试试"的递推算法。这里讨论这种方法。

首先给出关于动态可变参数的验前差分形式的离散线性模型。为此, 记

$$\theta_i = H_i\mu + \delta_i$$

则

$$\theta_{i+1} = \theta_i + \Delta H_i\mu + \Delta\delta_i$$

式中
$$\Delta H_i = H_{i+1} - H_i, \quad \Delta\delta_i = \delta_{i+1} - \delta_i$$

而 $\Delta\delta_i \sim N(0, \sigma^2_{\delta_{i+1}} + \sigma^2_{\delta_i})$, 这里假定 δ_i 间独立。

又由 $\mu \sim N_p(\beta, \Sigma_\mu)$, 于是可记

$$\mu = \beta + \varepsilon_\mu$$

且 $\mathrm{Var}(\varepsilon_\mu) = \Sigma_\mu$ 为已知, $\varepsilon_\mu \sim N_p(0, \Sigma_\mu)$。于是

$$\theta_{i+1} = \theta_i + \Delta H_i\beta + \Delta H_i\varepsilon_\mu + \Delta\delta_i$$

式中 ε_μ 与 $\Delta\delta_i$ 独立, 且

$$\Delta H_i\varepsilon_\mu + \Delta\delta_i \sim N(0, \Delta H_i\Sigma_\mu\Delta H_i^{\mathrm{T}} + \sigma^2_{\Delta\delta_i})$$

最后, 记

$$\theta_{i+1} = \theta_i + \Delta H_i\beta + \Delta w_i = \theta_i + \alpha_i + \Delta w_i \tag{6}$$

式中
$$\Delta w_i \sim N(0, \sigma^2_{\Delta w_i}), \quad \alpha_i = \Delta H_i\beta,$$
$$\sigma^2_{\Delta w_i} = \Delta H_i\Sigma_\mu\Delta H_i^{\mathrm{T}} + \sigma^2_{\Delta\delta_i}$$

(6) 式就是所需要的基于二层验前建立起来的关于 θ_i 的验前离散动力学模型。

此外, 在试验中, 我们获得的观测量是 y_i, 可将其记为

$$y_i = \theta_i + v_i \tag{7}$$

由于 $y_i|_{\theta_i} \sim N(\theta_i, \sigma_i^2)$, 因此上式中 $v_i \sim N(0, \sigma_i^2)$, σ_i^2 为已知。

在 θ_i 的动态递推过程中, 引入符号:

$\hat{\theta}_{i/i}$ —— i 时刻观测后, θ_i 的 MV 估计, $P_{i/i}$ —— $\hat{\theta}_i$ 的估计误差的方差, $\hat{\theta}_{i/i-1}$ —— 由 $i-1$ 时刻的估计预报 i 时刻的估计值(预报值), $P_{i/i-1}$ —— 预报估值误差方差。

利用验前建立的 θ_i 的动力学模型预报, 有

$$\hat{\theta}_{i/i-1} = \hat{\theta}_{i-1/i-1} + \alpha_{i-1} \tag{8}$$

式中
$$\alpha_{i-1} = \Delta H_{i-1}\beta$$

此时
$$P_{i/i-1} = P_{i-1/i-1} + \sigma^2_{\Delta W_{i-1}} \tag{9}$$

在 i 时刻 θ_i 的估计, 可由预报估计和当前观测量进行融合而成, 此时, θ_i 的融合 MV 估计为

$$\hat{\theta}_{i/i} = \hat{\theta}_{i/i-1} + K_i(y_i - \hat{\theta}_{i/i-1}) \tag{10}$$

式中
$$K_i = P_{i/i-1}(P_{i-1/i-1} + \sigma_i^2)^{-1} \tag{11}$$

而 i 时刻 θ_i 的估计误差方差可表示为

$$P_{i/i} = (1 - K_i)P_{i/i-1} \tag{12}$$

式(8) ~ (12) 就是所需要的递推估计公式。其起始运算条件为$(\hat{\theta}_{0/0}, P_{0/0})$, 它由验前信息确定。如果"0"时刻无验前信息可利用, 则可以任取 $\hat{\theta}_{0/0}$, 而 $P_{0/0}^{-1} = 0$。有关计算实现问题, 可参阅文献[5]。

4 工程应用中的实现问题

上面的 Bayes 递推公式是在给定模型以及模型的统计特性均为已知的前提下给出的。在工程应用

中,情况往往比较复杂。例如,每次试验时,试验的条件和场景不同,此时观测量 y_i 的随机误差的方差 σ_i^2 并不确切地知道。此外,动态参数 θ_i 的模型 $\theta_i = H_i\mu + \delta_i$,它是验前建立的,而在运用具有超参数的线性动力学模型 $\theta_{i+1} = \theta_i + \alpha_i + \Delta w_i$ 时,认为模型是完全正确的,且 $\sigma_{\Delta w_i}^2$ 也是完全确定的,这在实际工程问题中并不完全如此,因此在具体实现 Bayes 估计时,必须注意这些应用中的问题。

(1) 关于 θ_i 的验前动力学模型中 $\sigma_{\Delta w_i}^2$ 的确定问题

在模型(6)中,Δw_i 是验前建立的 θ_i 的动力学模型的动态噪声,它用以补偿模型的不确定性。在试验过程中,$\sigma_{\Delta w_i}^2$ 随不同的试验而改变,使 θ_i 的估计成为自适应的估计。因此,$\sigma_{\Delta w_i}^2$ 应在递推运算中估计出来,在文献[5]中有这方面的讨论。我们有如下公式:

$$\sigma_{\Delta w_{i-1}}^2 + P_{i-1/i-1} + \sigma_i^2 - y_{i/i-1}^2 = 0 \tag{13}$$

其中

$$y_{i/i-1} = y_i - \theta_{i/i-1}$$

由(13)式给出 $\sigma_{\Delta w_{i-1}}^2$ 的估计。

(2) 关于 σ_i^2 为未知的情形

此时,必须运用观测量进行估计。关于 σ_i^2 的估计,有多种方法,为便于运算,我们给出 σ_i^2 和 θ_i 同时递推估计方法[5]。

在 K 时刻,σ_K^2 的估计为

$$\sigma_K^2 = (y_{K/K-1})^2 - P_{K/K-1} \tag{14}$$

式中

$$y_{K/K-1} = y_K - \theta_{K-1/K-1} - \alpha_{K-1}$$
$$P_{K/K-1} = P_{K-1/K-1} + \sigma_{\Delta w_{K-1}}^2$$
$$K = 1, 2, \cdots, n$$

关于 σ_i^2 的估计,还可以给出在 N 步之内近似常量的情况,记 N 步之内的新信息为 $v_i = y_{i/i-1} = y_i - \theta_{i/i-1}(i = 1, \cdots, N)$。则 σ_i^2 的估计为[5]

$$\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left[(v_i - v)^2 - \frac{N-1}{N}P_{i/i-1}\right]$$

这样,运用式(13),即可获得 $\sigma_{\Delta w_{i-1}}^2$ 的估计为

$$\sigma_{\Delta w_{i-1}}^2 = y_{i/i-1}^2 - P_{i-1/i-1} - \sigma^2$$

5 在飞行器试验分析中的应用

前面,讨论了二层验前信息的模型及 Bayes 估计问题,对于更多层验前的情况,原则上说,没有本质的区别,只是对于条件分布的计算必须应用多层的方法。对于应用来说,二层验前是应用最广的。

关于二层验前,第一层次是建立关于动态参数的模型。上面讨论了线性回归模型的情况。由于工程实际问题常常运用线性方法,因此,线性回归有广泛的应用。第一层验前信息可以看作动态参数在试验过程中的变化规律问题,这在实际应用中也是首先要考虑的重要问题。而这种变化规律,在试验之前应该是明确的。特别是可以运用多种途径(理论的或技术的)确定出来。第二层先验是关于超参数(模型中待定的参数)。由 Bayes 观点,它是随机的,它的统计特性由验前信息给出,这也符合常规的想法。一般地,超参数(μ)的分布特性可以由多种验前信息确定。事实上,回归方程中未知参数的特性可以由线性模型未知参数辩识方法,给出它的估计和方差,从而进一步确定出它的分布。在这二层先验确定之下,再由现场试验的信息(观测子样)作出关于动态参数的 Bayes 估计。由于运用了较充分的信息,因此较之经典的试验分析方法(仅仅用现场信息建模和估计的方法)有较好的估计性能。

例 惯导系统测量装置的系统误差估计问题——工具误差分离问题。惯性导航系统的测量装置是安装在陀螺平台上的三个互相垂直的轴向加速度表,它们敏感测出非引力加速度的大小。三轴稳定平

台在飞行器上建立起一个惯性坐标系。如果加速度表本身不能精确地测量沿敏感轴的视加速度, 或者加速度表的敏感轴由于陀螺平台产生漂移而偏离其定位方向, 则最终将引起再入飞行器的落点偏差。因此提高落点精度的关键在于提高陀螺, 加速度表及惯性平台的精度。这就是制导系统仪表的误差分离问题。

首先是建立工具误差的结构模型, 所采用的模型可表示为[6]

$$\Delta W_K = S_K C + \varepsilon_K$$

其中 ΔW_K 是t_K 时刻遥测视速度和外测视速度偏差, S_K 为已知的环境函数, C 是待估的未知参数, ε_K 为模型的随机误差。在工程应用中, 常将 X, Y, Z 轴向在t_K 时刻的遥外测视速度偏差记为

$$\Delta W_X(t_K) = (S_{X_1}(t_K) \cdots S_{X_m}(t_K))\begin{pmatrix} C_1 \\ \vdots \\ C_m \end{pmatrix} + \varepsilon_X(t_K)$$

$$\Delta W_Y(t_K) = (S_{Y_1}(t_K) \cdots S_{Y_m}(t_K))\begin{pmatrix} C_1 \\ \vdots \\ C_m \end{pmatrix} + \varepsilon_Y(t_K)$$

$$\Delta W_Z(t_K) = (S_{Z_1}(t_K) \cdots S_{Z_m}(t_K))\begin{pmatrix} C_1 \\ \vdots \\ C_m \end{pmatrix} + \varepsilon_Z(t_K)$$

如果写成向量形式, 则有

$$\Delta W_K = S_K \cdot C + \varepsilon_K, \quad K = 1, \cdots, n$$

其中

$$\Delta W_K = \begin{pmatrix} \Delta W_X(t_K) \\ \Delta W_Y(t_K) \\ \Delta W_Z(t_K) \end{pmatrix}, \varepsilon_K = \begin{pmatrix} \varepsilon_X(t_K) \\ \varepsilon_Y(t_K) \\ \varepsilon_Z(t_K) \end{pmatrix}, S_K = \begin{pmatrix} S_{X_1}(t_K) \cdots S_{X_m}(t_K) \\ S_{Y_1}(t_K) \cdots S_{Y_m}(t_K) \\ S_{Z_1}(t_K) \cdots S_{Z_m}(t_K) \end{pmatrix}$$

S_K 称为环境函数矩阵, 它是已知的。式(15) 的线性回归方程的形式(结构) 在试验之前就已经建立。它由惯性制导平台系统的机理建立起来。它就是第一层验前信息。至于 C, 它是超参数(向量), 可以历次试验信息或地面环境试验或仿真试验等手段将 C 估计出来, 并分析其误差(方差)。一般说来, 由验前信息, 可以确认 $C \sim N_m(\bar{C}, \Sigma_C)$。$\bar{C}$, Σ_C 是已知的, 由验前信息确定, 这就是前面所说的第二层先验。在现场飞行试验之后, 由上述模型, 当获得观测 $\Delta W_K(K = 1, \cdots, n)$, 可以由二层验前作出 $\theta_K(S_K C)$ 的 Bayes 估计, 由此再作出 C 的估计(如 C 的L. S 的估计)。在飞行器试验分析中, 误差系数向量 C 的估计是重要的, 由它还可以分析由此造成的远程火箭主级关机点运动参数的偏差或射程偏差。

参 考 文 献:

[1] Berger J. Statistical Decision Theory [M]. 2nd Edi., Springer-Verlag, New York, Inc, 1985.

[2] Morris C. Parameter Empivical Bayes Inference Theory and Applications [J]. J. Amer. Statist Assoc., 78, 1983: 47- 65.

[3] Berlinear L M. A Decision-theoretic Structure for Robust Bayesian Analysis with Applications to the Estimation of Multivariate Normal Mean[J]. In Bayesian Statist., Ⅱ. North-Holand, Amsterdam, 1984.

[4] 张尧庭, 陈汉峰. 贝叶斯统计推断[M]. 北京: 科学出版社, 1991.

[5] 张金槐. 线性模型参数估计及其改进[M]. 长沙: 国防科技大学出版社, 1992.

[6] 张金槐, 贾沛然, 任萱, 何力. 远程火箭精度分析与评估[M]. 长沙: 国防科技大学出版社, 1995.