文章编号:1671-6833(2005)03-0086-03

多维组合预测的贝叶斯极大似然估计

陆宜清[1], 杨松华[2]

(1. 郑州牧业工程高等专科学校人文与基础科学系,河南 郑州 450011;2. 郑州大学数学系,河南 郑州 450052)

摘 要:为了实际预测的需要,提出了多维组合预测问题,即若干单项多维预测的变权组合预测.在各单项预测无偏且服从正态分布前题下,求出了 p 时刻预测向量 $\vec{X}_p$ 的先验分布密度和后验分布密度.利用主观先验信息、预测信息和样本信息,运用贝叶斯估计方法,得到了 $\vec{X}_p$ 的贝叶斯极大似然估计,其权重随时刻 p 的改变而改变.本方法充分利用了多维变量间的相关信息,进一步提高了预测的科学性和有效性,体现了对样本信息和预测信息的进一步综合应用.

关键词:多维组合预测;贝叶斯极大似然估计;协方差阵

中图分类号:O 212 **文献标识码**:A

0 引言

组合预测或预测的综合技术因能有效地提高预测精度而受到国内外预测工作者的重视,并出现了一系列的研究成果.文献[1]用贝叶斯方法给出了一种时变权组合方法,文献[2]给出了一维组合预测的贝叶斯无偏估计,文献[3,4]考虑了组合权随时间的变化(时变性)运用模糊控制理论给出了两种变权重组合预测算法.文献[5]对组合预测的贝叶斯估计方法进行了综述,文献[6]对组合预测方法进行了评述.但在进行实际预测时常常会遇到要求同时对相关的几个指标进行预测,这类预测可称之为多维预测.本文作者的目的是求出若干单项多维预测的变权组合预测,所运用的理论和方法是文献[1,2]中解决一维变权组合预测的贝叶斯极大似然估计法.变权组合预测是目前组合预测研究的一个难点,除文献[1,2]结果外,文献[3,4]运用模糊控制给出了另一方法.多维变权组合预测的理论和方法尚未见到成熟结果,笔者提出这一问题并利用贝叶斯理论给出了一种估计方法.

1 记号、前提与问题

设第 t 时刻真值为 $\vec{X}_t=(x_{t1},x_{t2},\cdots,x_{tq})^{\mathrm{T}}(t=1,2,\cdots,n,\cdots)$;$\hat{X}_t=(\hat{x}_{t1},\hat{x}_{t2},\cdots\hat{x}_{tq})$ 为观测向量($t\leqslant n$)或预测向量($t>n$);q 为待预测向量维数,q =1时为文献[1,2]中的一维问题,n 为观测向量(时刻)个数;$\vec{Y}_{tj}=(y_{tj1},y_{tj2},\cdots,y_{tjm})^{\mathrm{T}}$ 为第 t 时刻 m 种单项预测方法对第 j 个分量的拟合值($t\leqslant n$)或预测值($t>n$).在各单项预测是无偏且服从正态分布的前提下,有

$$
G_t=\begin{pmatrix}\vec{Y}_{t1}\\ \vec{Y}_{t2}\\ \vdots\\ \vec{Y}_{tq}\end{pmatrix}\sim N\left(\begin{pmatrix}x_{t1}\boldsymbol{I}\\ x_{t2}\boldsymbol{I}\\ \vdots\\ x_{tq}\boldsymbol{I}\end{pmatrix},\boldsymbol{\Sigma}_{(qm)\times(qm)}\right)
$$

其中 $\boldsymbol{I}=(1,1,\cdots,1)^{\mathrm{T}}_m$.

记 $\vec{U}_t=\begin{pmatrix}x_{t1}\boldsymbol{I}\\ x_{t2}\boldsymbol{I}\\ \vdots\\ x_{tq}\boldsymbol{I}\end{pmatrix}$,下面我们逐步求出 $\vec{X}_p(p>n)$ 的贝叶斯后验极大似然估计.

2 由拟合区信息确定 $\vec{X}_p$ 的先验分布

把 $\vec{G}_p$ 代入其分布密度函数得

$f(\vec{G}_p)=C\exp\{-\frac{1}{2}(\vec{G}_p-\vec{U}_p)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\vec{G}_p-\vec{U}_p)\}$,$C$ 为与 $\vec{X}_p$ 无关的常数.

$\max\limits_{\vec{X}_p}f(\vec{G}_p)$ 即 $\min\limits_{\vec{X}_p}[(\vec{G}_p-\vec{U}_p)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\vec{G}_p-\vec{U}_p)]=$

收稿日期:2005-03-15;修订日期:2005-05-13

基金项目:河南省杰出青年科学基金资助项目(0212000200)

作者简介:陆宜清(1966-),女,河南省镇平县人,郑州牧业工程高等专科学校副教授,主要从事应用数学及相关学科研究.

$\min\limits_{\vec{X}_p} \boldsymbol{Q}$.

由 $\frac{\partial \boldsymbol{Q}}{\partial x_{pj}} = \vec{\boldsymbol{O}}, j = 1,2,\cdots,q$ 得：

$$(\vec{\boldsymbol{U}}_p - \vec{\boldsymbol{G}}_p)^{\mathrm{T}} \boldsymbol{\Sigma}^{-1}\boldsymbol{L} = \vec{\boldsymbol{O}}$$

其中，$\boldsymbol{L} = \begin{pmatrix} \boldsymbol{I} & & & \\ & \boldsymbol{I} & & \\ & & \ddots & \\ & & & \boldsymbol{I} \end{pmatrix}_{(mq)\times q}$，$\vec{\boldsymbol{U}}_p = \boldsymbol{L}\vec{\boldsymbol{X}}_p$.

有 $\hat{\vec{\boldsymbol{X}}}_p = (\boldsymbol{L}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{L})^{-1}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\vec{\boldsymbol{G}}_p$（单样本极大似然估计）. 故先验分布为 $\vec{\boldsymbol{X}}_p \sim N(\hat{\vec{\boldsymbol{X}}}_p, (\boldsymbol{L}^{\mathrm{T}}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{L})^{-1})$，其中 $(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{L})^{-1}$ 是 $\vec{\boldsymbol{X}}_p$ 的先验协方差阵，$\boldsymbol{\Sigma}_1$ 可由拟合区信息确定，可仿文[1]取

$$\hat{\boldsymbol{\Sigma}}_1 = \sum_{t=1}^{n}(\vec{\boldsymbol{G}}_t - \hat{\vec{\boldsymbol{U}}}_t)(\vec{\boldsymbol{G}}_t - \hat{\vec{\boldsymbol{U}}}_t)^{\mathrm{T}}/n$$，其中 $\hat{\vec{\boldsymbol{U}}}_t = \begin{pmatrix} \hat{x}_{t1}\boldsymbol{I} \\ \vdots \\ \hat{x}_{tq}\boldsymbol{I} \end{pmatrix}$.

由此得 $\vec{\boldsymbol{X}}_p$ 近似先验分布，近似地，$\vec{\boldsymbol{X}}_p \sim N(\hat{\vec{\boldsymbol{X}}}_p, (\boldsymbol{L}^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_1^{-1}\boldsymbol{L})^{-1})$.

3 由预测区信息确定 $\boldsymbol{X}_p$ 的后验分布

$\vec{\boldsymbol{X}}_p$ 的先验分布密度为 $f_{\vec{X}_p}(\vec{\boldsymbol{X}}_p) = C_1\exp\{-\frac{1}{2}(\vec{\boldsymbol{X}}_p - \hat{\vec{\boldsymbol{X}}}_p)^{\mathrm{T}}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{L}(\vec{\boldsymbol{X}}_p - \hat{\vec{\boldsymbol{X}}}_p)\}$

条件分布密度为 $f_{\vec{G}_p|\vec{X}_p}(\vec{\boldsymbol{G}}_p) = C_2\exp\{-\frac{1}{2}\cdot(\vec{\boldsymbol{G}}_p - \vec{\boldsymbol{U}}_p)^{\mathrm{T}}\boldsymbol{\Sigma}_2^{-1}(\vec{\boldsymbol{G}}_p - \vec{\boldsymbol{U}}_p)\}$

C_1, C_2 为与 $\vec{\boldsymbol{X}}_p$ 无关常数.

仿文献[1,2]可对 $\boldsymbol{\Sigma}_2$ 作如下估计

$$\hat{\boldsymbol{\Sigma}}_2 = \sum_{t=r}^{r+N-1}\boldsymbol{\Lambda}_t(\vec{\boldsymbol{G}}_t - \boldsymbol{L}\vec{\boldsymbol{H}}_t)(\vec{\boldsymbol{G}}_t - \vec{\boldsymbol{H}}_t)^{\mathrm{T}}/N,$$

N 为待选常数，$N > mq$，$\boldsymbol{\Lambda}_t = \mathrm{diag}(\lambda_{t1}\boldsymbol{E}_q, \lambda_{t2}\boldsymbol{E}_q, \lambda_{t3}\boldsymbol{E}_q, \cdots, \lambda_{tm}\boldsymbol{E}_q)$，$\boldsymbol{E}_q$ 为 q 阶单位矩阵.

$$\lambda_{tj} = \left(\frac{D_{pj}}{D_{tj}}\right)^2, D_{ij} = \max_{1\leqslant k\leqslant m} y_{ijk} - \min_{1\leqslant k\leqslant m} y_{ijk}.$$

$\vec{\boldsymbol{H}}_t = (h_{t1}, h_{t2}, \cdots, h_{tq})^{\mathrm{T}}$，$h_{ij}$ 为 m 个方法对第 i 时刻第 j 个分量预测的 m 个数的中位数. $\boldsymbol{\Lambda}_t$ 为各点各方法预测方差非齐性的校正方阵，校正为 p 时刻方差.

近似后验分布密度为

$$f_{\vec{X}_p|\vec{G}_p}(\vec{\boldsymbol{X}}_p) = C\exp\{-\frac{1}{2}(\vec{\boldsymbol{X}}_p - \hat{\vec{\boldsymbol{X}}}_p)^{\mathrm{T}}(\mathrm{L}^{\mathrm{T}}\boldsymbol{\Sigma}_1^{-1}L)(\vec{X}_p - \hat{\vec{\boldsymbol{X}}}_p)\}\cdot\exp\{-\frac{1}{2}(\vec{\boldsymbol{G}}_p - \boldsymbol{L}\vec{\boldsymbol{X}}_p)^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_2^{-1}(\vec{\boldsymbol{G}}_p - \boldsymbol{L}\vec{\boldsymbol{X}}_p)\}.$$

将 $\vec{\boldsymbol{G}}_p$ 代入后，解 $\max\limits_{\vec{X}_p} f_{\vec{X}_p|\vec{G}_p}(\vec{\boldsymbol{X}}_p)$，即

$$\min_{\vec{X}_p}\boldsymbol{Q} = \min[(\vec{\boldsymbol{X}}_p - \hat{\vec{\boldsymbol{X}}}_p)^{\mathrm{T}}(\mathrm{L}^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_1^{-1}\boldsymbol{L})(\vec{X}_p - \hat{\vec{\boldsymbol{X}}}_p) + (\vec{\boldsymbol{G}}_p - \boldsymbol{L}\vec{\boldsymbol{X}}_p)^{\mathrm{T}}\boldsymbol{\Sigma}_2^{-1}(\vec{G}_p - \boldsymbol{L}\vec{\boldsymbol{X}}_p)].$$

解得

$$\hat{\hat{\vec{\boldsymbol{X}}}}_p = (\boldsymbol{L}^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_1^{-1}\boldsymbol{L} + \boldsymbol{L}^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_2^{-1}L)^{-1}(L^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_1^{-1} + L^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_2^{-1})\vec{G}_p,$$

近似地 $\vec{\boldsymbol{X}}_p \mid \vec{\boldsymbol{G}}_p \sim N(\hat{\hat{\vec{\boldsymbol{X}}}}_p, (\boldsymbol{L}^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_1^{-1}\boldsymbol{L} + \boldsymbol{L}^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_2^{-1}\boldsymbol{L})^{-1})$，对各分量而言 $\hat{\hat{\vec{\boldsymbol{X}}}}_p$ 为一组合预测. 又

$$\hat{\hat{\vec{\boldsymbol{X}}}}_p = (\mathrm{L}^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_1^{-1}\boldsymbol{L} + \boldsymbol{L}^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_2^{-1}\boldsymbol{L})^{-1}(\mathrm{L}^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_1^{-1}\boldsymbol{L}\hat{\vec{\boldsymbol{X}}}_p + \boldsymbol{L}^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_2^{-1}\boldsymbol{L}\hat{\vec{\boldsymbol{X}}}_p^*),$$

其中 $\hat{\vec{\boldsymbol{X}}}_p = (\mathrm{L}^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_1^{-1}\boldsymbol{L})^{-1}\boldsymbol{L}^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_1^{-1}\vec{\boldsymbol{G}}_p$,

$\hat{\vec{\boldsymbol{X}}}_p^* = (\mathrm{L}^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_2^{-1}\boldsymbol{L})^{-1}\boldsymbol{L}^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_2^{-1}\vec{\boldsymbol{G}}_p$ 为两个组合预测，$\hat{\hat{\vec{\boldsymbol{X}}}}_p$ 为 $\hat{\vec{\boldsymbol{X}}}_p$ 与 $\hat{\vec{\boldsymbol{X}}}_p^*$ 的组合预测.

4 几个需要指出的问题与结果

(1) 对各分量单项预测的个数不一定相等，作这一推广时只须将 L 中的 I 改为不同维，维数与各分量单项预测方法的个数相等.

(2) 本文结论是在 $\vec{\boldsymbol{G}}_p \sim N(\vec{\boldsymbol{U}}_p, \boldsymbol{\Sigma})$ 前提下得到的.

(3) $\Sigma_1, \boldsymbol{\Sigma}_2 > 0$（正定）时，有 $(\mathrm{L}^{\mathrm{T}}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{L})^{-1} > (\mathrm{L}^{\mathrm{T}}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{L} + \boldsymbol{L}^{\mathrm{T}}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{L})^{-1}$. 即 $\hat{\hat{\vec{\boldsymbol{X}}}}_p$ 优于 $\hat{\vec{\boldsymbol{X}}}_p$.（$\hat{\vec{\boldsymbol{X}}}_p \sim N(\vec{\boldsymbol{X}}_p, (\mathrm{L}^{\mathrm{T}}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{L})^{-1})$，$\hat{\hat{\vec{\boldsymbol{X}}}}_p \sim N(\vec{\boldsymbol{X}}_p, (\mathrm{L}^{\mathrm{T}}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{L} + \boldsymbol{L}^{\mathrm{T}}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{L})^{-1})$）.

证：$\exists \mathrm{J} \in R_{q\times q}$ 可逆，使 $\boldsymbol{J}^{\mathrm{T}}L^{\mathrm{T}}\boldsymbol{\Sigma}_1^{-1}\boldsymbol{L}\boldsymbol{J} = \boldsymbol{\Lambda}_1 = \mathrm{diag}(\lambda_{11}, \lambda_{12}, \cdots, \lambda_{1q}) > 0$

$\boldsymbol{J}^{\mathrm{T}}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{\Sigma}_2^{-1}\boldsymbol{L}\boldsymbol{J} = \boldsymbol{\Lambda}_2 = \mathrm{diag}(\lambda_{21}, \lambda_{22}, \cdots, \lambda_{2q}) > 0$.（$L$ 列满秩）

$$\forall \vec{C} \in R_q, \vec{C} \neq \vec{0}, \vec{C}^T[(L^T \Sigma_1^{-1} L)^{-1} - (L^T \Sigma_1^{-1} L + L^T \Sigma_2^{-1} L)^{-1}] \vec{C} = \vec{C}^T J [\Lambda_1^{-1} - (\Lambda_1 + \Lambda_2)^{-1}] J^T \vec{C} = \vec{C}^T J \mathrm{diag}\left(\frac{1}{\lambda_{11}} - \frac{1}{\lambda_{11} + \lambda_{21}}, \cdots, \frac{1}{\lambda_{1q}} - \frac{1}{\lambda_{1q} + \lambda_{2q}}\right) J^T \vec{C} > 0$$

其中：$\left(\frac{1}{\lambda_{1i}} - \frac{1}{\lambda_{1i} + \lambda_{2i}} > 0, i = 1, \cdots, q\right)$. 故上述结论成立.

近年来,组合预测理论和方法的研究得到数学工作者和数量经济工作者的广泛关注,产生了组合预测的多种求取方法.但过去已出现的求组合权的方法存在两个共同缺陷:其一是未考虑主观先验信息,其二是未充分提取各预测方法的正确的预测信息.

笔者提出了多维组合预测问题,并在各单项模型无系统偏差和尺度偏差、拟合向量和预测向量均服从正态分布前提下,充分利用了主观先验信息、预测信息和样本信息,运用贝叶斯估计方法求出了待预测向量 $\vec{X}_p$ 的极大似然估计,及多指标组合预测的计算公式,对上述两个问题进行了一定程度的解决和改进,进一步提高了预测的科学性和有效性.

参考文献:

[1] 张新育.组合预测的贝叶斯极大似然估计法[J].预测,1998,17(5):46～47.

[2] 张新育,杨松华,成立社,等.无系统偏差和尺度偏差组合预测的贝叶斯极大似然估计[J].预测,1999,18(1):68～69.

[3] 唐小我,王 景,曹长修.一种新的模糊自适应变权组合预测算法[J].电子科技大学学报,1997,16(3):25～27.

[4] 曹长修,王 景,唐小我.一种模糊变权重组合预测方法——FVW法的研究[J].预测,1996,15(5):41～43.

[5] 张新育,孙甫照,杨松华.组合预测的贝叶斯估计方法[J].郑州大学学报(工学版),2002,23(4):87～90.

[6] 唐 纪,王 景.组合预测方法评述[J].预测,1999,18(2):42～43.

Bayes Maximum Likelihood Estimate in Multidimensional Combinational Forecast

LU Yi-qing, YANG Song-hua

(1.Department of Humanities & Basic Sciences, Zhengzhou Engineering College of Animal Husbandry, Zhengzhou 450011, China; 2.Department of Mathematics, Zhengzhou University, Zhengzhou 450052, China)

Abstract: In order to meet with the actual forecast, this paper presents the question of multidimensional combinational forecast or the varied weight combinational forecast with several monomial dimensional forecasts. It finds the prior distribution density and posterior distribution density of the forecast vector at time p when all monomial forecasts are free from deviation and are in agreement with normal distribution, and then by employing subjective prioriy information, forecast information and sample information and by using the Bayes estimate method, we obtain the Bayes maximumlikelihood estimate of $\vec{X}_p$ in which the weight vary with the time p. This method makes good use of the information about the multidimensional vector, further improves the scientificalness and effectiveness of forecast and reflects comprehensive application of sample information and forecast information.

Key words: multidimensional combinational forecast; Bayes maximumlikelihood estimate; covariance matrix