第2卷第3期
1999年9月
管理科学学报
JOURNAL OF MANAGEMENT SCIENCES IN CHINA
Vol. 2 No. 3
Sep. 1999

组合预测贝叶斯模型研究[①]

曾 勇[②] 唐小我 （电子科技大学管理学院）
郑维敏 （清华大学经济管理学院）

【摘要】评述了组合预测的基本思想和组合预测贝叶斯模型的研究现状，研究了多步贝叶斯信息更新情况下针对不同偏性特征的无偏组合预测模型，讨论了结合非样本信息和样本信息的组合权重贝叶斯更新模型，扩展了贝叶斯组合预测模型的现有成果，指出了进一步研究的方向.

关键词：组合预测，模型不确定性，贝叶斯模型，无偏预测，非样本信息，收缩估计
分类号：F272；F224

0 引 言

在经济预测领域，一个有突出特点的传统就是寻求描述现实的真实模型. 当面对基于不同假设采用不同方法建立的不同模型时，通常的做法是首先采用假设检验和诊断检查的方法选择最好的模型而拒绝其他模型，然后通过修改接受的模型希望能够更加接近真实. 而 Bates 和 Granger 提出的组合预测思想[1]的基本出发点就是在大多数需要作出预测的情况下，难以获得完全的信息集，即使对于给定的信息集也难以做到最优利用[2]，即承认构造真实模型的困难，将各种单项预测看作代表不同的信息片段（pieces of information）[3]，通过信息的集成分散单个预测特有的不确定性和减少总体不确定性，从而提高预测精度. 这就如同证券组合分散风险的原理一样，组合预测相对于单项预测具有更高的预测性能和出现极端预测误差的更小风险[4,5]. 组合预测的思想也得到近来关于模型不确定性讨论的支持. 为避免通过越来越复杂的模型接近真实的努力所导致的"乐观原理"（the optimism principle）以及将相同数据段同时用于模型拟合和模型诊断检查所造成的预测精度假象，Chatfield 强调[6]，在存在模型不确定性的情况下，应考虑多个模型而非单个的"真实"模型，应考虑具有变动参数的局部模型而非固定参数的全局模型，应考虑对模型不确定性具有较强鲁棒性的模型而非对模型假设高度敏感的单个模型，应考虑采用简省原则下的信息准则、各种再抽样技术和截断评价技术等去评价模型而非采用与模型拟合具有相同数据段的假设检验和诊断检查方法.

组合预测的关键是确定组合结构和组合参数，目前主要采用的是线性组合预测，非线性组合预测也得到了一定的发展[7,8]. 由于贝叶斯方法可以清楚地表示信息的更新过程，其优越性还在于可以通过临时的或系统的、前馈的或反馈的各种介入（interventions）方便地将主观信息与模型或数据信息相结合[9]，因此，组合预测的贝叶斯方法对于适应性地更新组合权重具有重要意义.

本文综述了组合预测贝叶斯模型的研究现状，进而讨论了多步信息更新下的无偏组合预测贝叶斯模型，指出了进一步研究的方向.

① 国家杰出青年科学基金资助项目（79725005）.
② 曾 勇：硕士，教授. 通信地址：电子科技大学管理学院，邮编：610054.
本文 1998 年 11 月 23 日收到.

1 组合预测贝叶斯模型的研究现状

组合预测的贝叶斯方法主要有两种思路.一种是对组合权重给予概率解释,另一种是直接讨论单项预测对被预测变量先验分布的信息更新.前一种思路产生有别于统计组合预测模型的组合方法(即一般不解释为组合预测精度指标下的最优组合),而后一种思路则产生与各种组合预测模型形式相同的后验分布均值.

在第一种思路下,Bunn 提出的根据单项预测历史表现最佳次数进行加权的超出表现方法(the outperformance method)[10,11]从贝叶斯观点解释就是将组合权重解释为单项预测表现最佳的概率,采用 Beta 或 Dirichlet 先验分布,以二项式或多项式过程更新信息.按照 Bunn 的思路,Firia 和 Souza 探讨了单项预测概率分布的组合问题[12],将组合权重解释为观测值属于各单项预测分布设定的总体的概率,组合权重先验分布采用 Dirichlet 分布,并且为避免多期信息更新情况下后验分布的计算复杂性问题,分布参数的更新根据单项预测似然函数进行简化计算,得到了组合预测的准贝叶斯模型. Min 和 Zellner 则将组合权重解释为单项预测模型真实描述实际过程的概率,采用单项预测之间的后验优势比率(posterior odds ratios)确定和更新权重[13].实际上,文献[12,13]的方法是更一般性的贝叶斯模型平均(Bayesian model averaging)方法的特例,该类方法可以在一定程度上处理模型不确定性问题[6,14].

在第二种思路下,Winkler 和 Bordler 首先考察了无偏单项预测的贝叶斯组合模型[15,16],在假设单项预测误差满足零均值和已知协方差阵的多元正态分布、被预测变量无信息先验分布并与单项预测误差不相关情况下,一步信息更新下的贝叶斯预测等价于 B—G 组合. Lindley 进一步考察了协方差阵未知并服从逆 Wishart 先验分布情况下一步信息更新的贝叶斯预测,得到了类似的结论[17].为消除组合预测偏差,Granger 和 Ramanathan 建议采用线性回归组合预测[18]. Bordley 考察了有偏单项预测的贝叶斯组合,在假定正态先验分布并与单项预测误差相关情况下,已知预测误差协方差阵的一步贝叶斯预测与 G—R 模型形式相同[19],但文中未考虑偏性特征. Anandalingam 和 Chen 进一步考虑了偏性特征,在假定预测偏差对先验信息无影响,并给定预测误差协方差阵和偏性特征参数情况下,一步信息更新的贝叶斯组合预测等价于偏差校正后单项预测的 B—G 组合,还等价于卡尔曼滤波[20,21].文献[22]进一步分析指出,当考虑模型预测误差信息对先验信息更新的影响作用时,无偏组合预测不等价于卡尔曼滤波,与 G—R 模型的差别是用先验预测代替了被预测变量的无条件期望值,从而可以避免 G—R 模型的一个严重问题.

在第2种思路下,Bordley 还将预测误差分布假定为对数正态分布[16],导出的贝叶斯组合预测等价于单项预测的加权几何平均. Tibiletti 推广了 Bordler 的思路,通过对单项预测进行非线性变换[7],将一般预测误差分布转化为正态分布,得出了非线性组合预测模型.高仁祥、张世英和刘豹将 Tibiletti 的一般非线性变换具体化为 Box—Cox 变换[23],所导出的贝叶斯组合预测等价为一类组合预测模型,包括加权算术平均、加权几何平均和加权调和平均等.

在下一节中,本文将第2种思路进一步推广到多步信息更新的情况,分别导出了单项预测仅存在位置偏差和同时具有尺度偏差两种情况下的无偏贝叶斯组合模型.在第3节中,本文讨论了权重的直接信息更新方法,得出了结合非样本信息和样本信息的权重贝叶斯模型.

2 多步信息更新情况下的贝叶斯无偏组合预测模型

2.1 仅存在位置偏差的贝叶斯模型

设对于被预测变量 y_t,有 n 个单项预测 f_{it}.这些单项预测可能来自时间序列模型、计量经济模型或

专家预测.若单项预测仅存在位置偏差,则偏性特征方程可表示为

$$y_t = a_i + f_{it} + u_{it} \qquad i = 1,\cdots,n \tag{1a}$$

式中 a_i 为 f_{it} 的位置偏差,$E(u_{it}) = 0$.

记 $R = (1,1,\cdots,1)'$,$f_t = (f_{1t},\cdots,f_{nt})'$,$u_t = (u_{1t},\cdots,u_{nt})'$,$\alpha = (a_1,\cdots,a_n)$,则上式可表示为

$$x_t = R \cdot y_t - f_t = \alpha + u_t \tag{1b}$$

式中 x_t 具有未知的均值 α 和精度阵 r.关于 α 和 r 的多步贝叶斯更新,有如下结论.

定理 1 假定观察数据有 T 期,u_t 为正态分布,α 和 r 的先验联合分布为:给定 r,α 服从均值 α_0、精度阵 υr 的 n 维正态分布($\upsilon > 0$);而 r 的边缘分布服从自由度 λ,精度阵 τ 的 Wishart 分布($\lambda > n-1$).则 α 和 r 的后验联合分布为:给定 r,α 服从均值 α^*、精度阵 $(\upsilon + T)r$ 的 n 维正态分布;而 r 的边缘分布为自由度 $\lambda + T$,精度阵 τ^* 的 Wishart 分布.其中

$$\begin{cases} \alpha^* = \dfrac{\upsilon \alpha_0 + T\overline{X}}{\upsilon + T} & (2a) \\ \tau^* = \tau + s + \dfrac{\upsilon T}{\upsilon + T}(\alpha_0 - \overline{x})(\alpha_0 - \overline{x})' & (2b) \end{cases}$$

式中 $\overline{x} = \dfrac{1}{n}\sum_{t=1}^{T} x_t$,$s = \sum_{t=1}^{T}(x_t - \overline{x})(x_t - \overline{x})'$.

证明 由定理 1 条件可知

$$\begin{cases} p(x_1,\cdots,x_T|\alpha,r) \propto |r|^{\frac{T}{2}}\exp\left\{-\dfrac{1}{2}\sum_{t=1}^{T}(x-\alpha)'r(x-\alpha)\right\} & (3a) \\ p(\alpha|r) \propto |\upsilon r|^{\frac{1}{2}}\exp\left\{-\dfrac{1}{2}(\alpha-\alpha_0)'(\upsilon r)(\alpha-\alpha_0)\right\} & (3b) \\ p(r) \propto |r|^{(\lambda-n-1)/2}|\tau|^{\lambda/2}\exp\left\{-\dfrac{1}{2}\mathrm{tr}(\tau r)\right\} & (3c) \end{cases}$$

式中 $p()$ 为有关变量的概率密度,$\mathrm{tr}()$ 为有关矩阵的迹.

由贝叶斯公式可得 α 和 r 的联合后验密度满足:

$$p(\alpha,r|x_1,\cdots,x_T) \propto |r|^{\frac{1}{2}}\exp\left\{-\frac{\upsilon}{2}(\alpha-\alpha_0)'r(\alpha-\alpha_0)\right\} \cdot$$
$$|r|^{(\lambda-T-n-1)/2}\exp\left\{-\frac{1}{2}\mathrm{tr}(\tau r) - \frac{1}{2}\sum_{t=1}^{T}(x_t-\alpha)'r(x_t-\alpha)\right\} \tag{4a}$$

由 $$\sum_{t=1}^{T}(x-\alpha)'r(x-\alpha) = T(\alpha-\overline{x})'r(\alpha-\overline{x}) + \mathrm{tr}(sr)$$

及 $$\upsilon(\alpha-\alpha_0)'r(\alpha-\alpha_0) + T(\alpha-\overline{x})'r(\alpha-\overline{x}) =$$
$$(\upsilon+T)(\alpha-\alpha^*)'r(\alpha-\alpha^*) + \frac{\upsilon T}{\upsilon+T}(\alpha_0-\overline{x})'r(\alpha_0-\overline{x})$$

即可将式(4a)重写为

$$p(\alpha,r|x_1,\cdots,x_T) \propto |r|^{\frac{1}{2}}\exp\left\{-\frac{\upsilon+T}{2}(\alpha-\alpha^*)'r(\alpha-\alpha^*)\right\} \times$$
$$|r|^{(\lambda+T-n-1)/2}\exp\left\{-\frac{1}{2}\mathrm{tr}(\tau^* r)\right\} \tag{4b}$$

由式(4b)即可得定理 1 的结论. 证毕.

定理 1 的条件还可根据先验信息的情况作进一步简化.若对 α 无明确的先验信息,即 $\upsilon \to 0$,则相应的后验分布 $\alpha = \overline{x}$,$\tau^* = \tau + s$.若对 α 和 r 均无先验信息,则 $\upsilon \to 0$,$\tau \to 0$,$\lambda \to -1$,α 和 r 的后验联合分布为:给定 r,α 的条件后验分布为均值 $\overline{x}$、协方差阵 $(Tr)^{-1}$ 的正态分布;r 的边缘后验分布为自由度 $T-1$、精度阵 s 的 Wishart 分布.α 和 r 相应的广义先验分布为

$$p(\alpha,r) \propto |r|^{-(n-1)/2} \tag{5}$$

现讨论第 $T+1$ 期的贝叶斯组合预测. 为避免繁琐,设 y_{T+1} 的先验分布为无信息先验,对于 y_{T+1} 的先验分布为正态分布且与误差信息相关的情况,可参照文献[22]的方法推导. 由此,有如下结论.

定理 2　在 y_{T+1} 无信息先验情况下,y_{T+1} 的后验分布为自由度 $\lambda+T$、均值 m^*、方差 V^* 的 t 分布,其中

$$\begin{cases} m^* = R'(\tau^*)^{-1}\tilde{f}_{T+1}/R'(\tau^*)^{-1}R & \text{(6a)} \\ V^* = \dfrac{(\upsilon+T+1)+(\upsilon+T)(m^*R-\tilde{f}_{T+1})'(\tau^*)^{-1}(m^*R-\tilde{f}_{T+1})}{(\lambda+T-2)[R'(\tau^*)^{-1}R]} & \text{(6b)} \end{cases}$$

式中　$\tilde{f}_{T+1} = f_{T+1} + \alpha^*$,为校正偏差后的单项预测向量.

证明　由定理 2 条件和式(1)表示可知 y_{T+1} 的先验密度和 f_{T+1} 的条件密度满足

$$\begin{cases} p(y_{T+1}|\alpha,r) \propto \text{常数} & \text{(7a)} \\ p(f_{T+1}|y_{T+1},\alpha,r) \propto |r|^{\frac{1}{2}}\exp\left\{-\dfrac{1}{2}(f_{T+1}-Ry_{T+1}+\alpha)'r(f_{T+1}-Ry_{T+1}+\alpha)\right\} & \text{(7b)} \end{cases}$$

结合式(4b)由贝叶斯定理并经整理可得

$$p(y_{T+1},\alpha,r|f_{T+1},x_1,\cdots,x_T) \propto |r|^{\frac{1}{2}}\exp\left\{-\frac{\upsilon+T+1}{2}(\alpha-\alpha^{**})'r(\alpha-\alpha^{**})\right\}\times$$
$$|r|^{(\lambda+T-n)/2}\exp\left\{-\frac{1}{2}\left[\operatorname{tr}(\tau^* r)+\frac{\upsilon+T}{\upsilon+T+1}(f_{T+1}-Ry_{T+1}+\alpha^*)'r(f_{T+1}-Ry_{T+1}+\alpha^*)\right]\right\} \tag{8}$$

式中　$\alpha^{**} = \alpha^* - (f_{T+1} - Ry_{T+1} + \alpha^*)/(\upsilon+T+1)$

上式对 α 和 r 积分并由

$$\int_r |r|^{(a-n)/2}\exp\left\{-\frac{1}{2}\operatorname{tr}(\Sigma r)\right\}\mathrm{d}r = |\Sigma|^{-(a+1)/2} \tag{9}$$[24]

及

$$|A+bb'| = |A|(1+b'A^{-1}b) \tag{10}$$

可得

$$\begin{aligned} & p(y_{T+1}|f_{T+1},x_1,\cdots,x_T) \\ \propto & \left[1+\frac{\upsilon+T}{\upsilon+T+1}(f_{T+1}-Ry_{T+1}+\alpha^*)'(\tau^*)^{-1}(f_{T+1}-Ry_{T+1}+\alpha^*)\right]^{-\frac{\lambda+T+1}{2}} \\ \propto & \left\{1+\frac{(\upsilon+T)[R'(\tau^*)^{-1}R](y_{T+1}-m^*)^2}{(\upsilon+T+1)+(\upsilon+T)(m^*R-\tilde{f}_{T+1})'(\tau^*)^{-1}(m^*R-\tilde{f}_{T+1})}\right\}^{-\frac{\lambda+T+1}{2}} \end{aligned} \tag{11}$$

对照 t 分布密度即得定理 2 结论.

由于对误差协方差阵为 Σ 的 n 种单项预测的 B－G 组合为 $R'\Sigma^{-1}f_t/R'\Sigma^{-1}R$,定理 2 的结论表示贝叶斯后验分布均值为单项预测校正偏差后的 B－G 组合,其中偏差和预测误差协方差阵以及预测精度由贝叶斯更新确定.

2.2　同时存在位置偏差和尺度偏差的贝叶斯模型

若单项预测除位置偏差外,还存在尺度偏差,则偏性特征方程表示为

$$y_t = \alpha_i + \beta_i f_{it} + u_t \qquad i = 1,\cdots,n \tag{12a}$$

式中　β_i 为 f_{it} 的尺度偏差系数.

记 $F_t = [I_n, \operatorname{diag}(f_{it})]$,$\theta = (\alpha_1,\cdots,\alpha_n,\beta_1,\cdots,\beta_n)'$,则上式重写为

$$Ry_t = F_t\theta + u_t \tag{12b}$$

存在尺度偏差情况下,若将 u_t 精度阵的先验分布设定为 Wishart 分布,则不能导出 Wishart 后验分布. 为保证共轭性,将未知精度阵更强地假定为 ωr,其中仅 ω 未知,ω 服从 Gamma 先验分布. 同时为避免繁琐,θ 假定为无信息先验分布. 关于 ω 和 θ 的贝叶斯更新,有如下结论.

定理 3　设观察数据有 T 期,u_t 为均值零、协方差阵 $(\omega r)^{-1}$ 的正态分布,θ 和 ω 的先验联合分布为:给定 ω,θ 为无信息先验;ω 的边缘先验分布为 $\Gamma(a,b)(a>0,b>0)$. 则 θ 和 ω 的后验联合分布为:给定 ω,θ 服从均值 $\hat{\theta}$、精度阵 ωG_T 的正态分布,而 ω 的边缘分布为参数 a^* 和 b^* 的 Gamma 分布,其中

$$\begin{cases} \hat{\theta} = (\sum_{t=1}^{T} F_t' r F_t)^{-1} (\sum_{t=1}^{T} F_t' r R y_t) & (13a) \\ G_T = \sum_{t=1}^{T} (F_t' r F_t) & (13b) \\ a^* = a + (T-1)/2 & (13c) \\ b^* = b + \frac{1}{2} \sum_{t=1}^{T} (R y_t - F_t \hat{\theta})' r (R y_t - F_t \hat{\theta}) & (13d) \end{cases}$$

证明 由定理 1 条件可知

$$\begin{cases} p(u_1, \cdots, u_T | \theta, \omega) \propto |\omega r|^{\frac{T}{2}} \exp\{-\frac{1}{2} \sum_{t=1}^{T} (R y_t - F_t \theta)' (\omega r) (R y_t - F_t \theta)\} & (14a) \\ p(\theta, \omega) \propto \omega^{a-1} \exp(-b\omega) & (14b) \end{cases}$$

由贝叶斯公式及

$$\sum_{t=1}^{T} (R y_t - F_t \theta)' r (R y_t - F_t \theta) = \sum_{t=1}^{T} (R y_t - F_t \hat{\theta})' r (R y_t - F_t \hat{\theta}) + (\theta - \hat{\theta})' G_T (\theta - \hat{\theta})$$

可得

$$p(\theta, \omega | y_t, F_t, 1 \leqslant t \leqslant T) \propto \omega^{a^*-1} \exp(-b^* \omega) \times \omega^{\frac{1}{2}} \exp\{-\frac{\omega}{2} (\theta - \hat{\theta})' G_T (\theta - \hat{\theta})\} \quad (15)$$

由上式即得结论. 证毕.

式(13a)意味着 $\hat{\theta}$ 为加权最小二乘估计的结果,这是因为未设定 θ 的先验均值所致.

对定理 3 的先验分布也可进一步简化. 若 $a \to 0, b \to 0$, 则 θ 和 ω 的广义先验联合分布为

$$p(\theta, \omega) \propto \frac{1}{\omega} \quad (16)$$

下面讨论第 $T+1$ 期的贝叶斯组合预测. 同样,为避免繁琐,设 y_{T+1} 无信息先验,有如下结论.

定理 4 在 y_{T+1} 无信息先验情况下,y_{T+1} 的后验分布为自由度 $2a+T-1$、均值 m^*、方差 V^* 的 t 分布,其中

$$\begin{cases} m^* = R' V_{T+1}^{-1} (F_{T+1} \hat{\theta}) / R' V_{T+1}^{-1} R & (17a) \\ V^* = 2b^{**} / (2a + T - 3)(R' V_{T+1}^{-1} R) & (17b) \end{cases}$$

式中 $V_{T+1} = r^{-1} + F_{T+1} G_T^{-1} F_{T+1}'$, $b^{**} = b^* + \frac{1}{2} (m^* R - F_{T+1} \hat{\theta})' V_{T+1}^{-1} (m^* R - F_{T+1} \hat{\theta})$

证明 由式(12b)可知

$$p(f_{T+1} | y_{T+1}, \theta, \omega) \propto \omega^{\frac{1}{2}} \exp\{-\frac{\omega}{2} (R y_{T+1} - F_{T+1} \theta)' r (R y_{T+1} - F_{T+1} \hat{\theta})\} \quad (18)$$

由定理 4 条件和贝叶斯公式并经整理得

$$\begin{aligned} & p(y_{T+1}, \theta, \omega | f_{T+1}; y_t, f_t, 1 \leqslant t \leqslant T) \propto \omega^{a^*} \exp(-b^* \omega) \times \\ & \exp(-\frac{\omega}{2} [(\theta - \theta^*)' G_{T+1} (\theta - \theta^*) + (R y_{T+1} - F_{T+1} \hat{\theta})' V_{T+1}^{-1} (R y_{T+1} - F_{T+1} \hat{\theta}]\} \end{aligned} \quad (19)$$

式中 $\theta^* = \hat{\theta} + G_{T+1}^{-1} F_{T+1}' r (R y_{T+1} - F_{T+1} \hat{\theta})$, $G_{T+1} = G_T + F_{T+1}' r F_{T+1}$.

上式对 θ 积分并经整理可得

$$\begin{aligned} & p(y_{T+1}, \omega | f_{T+1}; y_t, f_t, 1 \leqslant t \leqslant T) \\ & \propto \omega^{\frac{2a^*-1}{2}} \exp\{-\frac{\omega}{2} [2b^{**} + (R' V_{T+1}^{-1} R)(y_{T+1} - m^*)^2]\} \end{aligned} \quad (20)$$

上式对 ω 积分可得

$$p(y_{T+1}, \omega | f_{T+1}; y_t, f_t, 1 \leqslant t \leqslant T) \propto [1 + \frac{R' V_{T+1}^{-1} R}{2b^{**}} (y_{T+1} - m^{*2})]^{-\frac{2a^*+1}{2}} \quad (21)$$

对照 t 分布即得定理 4 结论.

由于 $F_{T+1}\hat{\theta} = (\hat{a}_1 + \hat{\beta}_1 f_{1t}, \cdots, \hat{a}_n + \hat{\beta}_n f_{nt})'$,因此定理 4 意味着贝叶斯后验均值为同时校正位置和尺

度偏差后的 B — G 组合.

3 结合非样本信息和样本信息的组合权重贝叶斯更新模型

前面讨论的模型未直接涉及组合权重的贝叶斯更新问题.另一方面,经验研究表明[4,25~27],在变动的环境中,简单的组合预测模型如简单平均组合相对于复杂的组合预测模型常常表现出更良好的性能.因此,结合简单组合模型的鲁棒性与单项预测历史表现的贝叶斯方法值得进一步考虑.下面,将简单组合的权重作为非样本信息,通过贝叶斯方式结合样本信息来讨论组合权重的直接更新问题.

将组合预测模型一般性地表示为 $f_{ct} + w_0 + \sum_{i=1}^{n} w_i f_{it}$,考虑到随机误差以及组合预测所引入的误差序列自相关问题[28],采用处理误差自相关的简单有效方法[29],将 y_{t-1} 引入组合中,则组合预测方程表示为

$$y_t = f_t' w + \varepsilon_t \tag{22}$$

式中 $f_t = (1, f_{1t}, \cdots, f_{nt}, y_{t-1})'$, $w = (w_0, w_1, \cdots, w_n, w_{n+1})'$, ε_t 为随机误差.

设 ε_t 服从均值零、方差 ω^{-1} 的正态分布,且序列无关. w 和 ω 的先验联合分布为:给定 ω, w 服从均值 w_0、协方差阵 $(\omega\tau)^{-1}$ 的多元正态分布;而 ω 的边缘分布为 $\Gamma(a,b)$. w_0 可以表示关于组合权重的非样本信息,如 $w_0 = (0, \frac{1}{n}, \cdots, \frac{1}{n}, 0)'$ 即为简单平均组合.若有 T 期观测值,并记 $N = n+2$,则

$$p(\omega, w) \propto \omega^{a-1}\exp(-b\omega) \times \omega^{\frac{N}{2}}\exp\{-\frac{\omega}{2}(w - w_0)'\tau(w - w_0)\} \tag{23}$$

$$p(\varepsilon_1, \cdots, \varepsilon_T | \omega, w) \propto \omega^{\frac{T}{2}}\exp\{-\frac{\omega}{2}\sum_{t=1}^{T}(y_t - f_t' w)^2\} \tag{24}$$

由贝叶斯定理并经整理可得

$$p(\omega, w | y_t, f_t, 1 \leqslant t \leqslant T) \propto$$

$$\omega^{\hat{a}-1}\exp(-\hat{b}\omega) \times \omega^{\frac{N}{2}}\exp\{-\frac{\omega}{2}(w - \hat{w})'(\tau + \sum_{t=1}^{T} f_t f_t')(w - \hat{w})\} \tag{25}$$

式中

$$\begin{cases} \hat{a} = a + \dfrac{T}{2} & (26\text{a}) \\ \hat{b} = b + \dfrac{1}{2}[w_0'\tau w_0 + \sum_t y_t^2 - \hat{w}'(\tau + \sum_t f_t f_t')\hat{w}] & (26\text{b}) \\ \hat{w} = (\tau + \sum_t f_t f_t')^{-1}(\tau w_0 + \sum_t f_t y_t) & (26\text{c}) \end{cases}$$

可见, w 和 ω 的后验联合分布仍为正态——Gamma 分布.式(26c)还可重写为

$$\hat{w} = (\tau + \sum_t f_t f_t')^{-1}[\tau w_0 + (\sum_t f_t f_t')w_1)] \tag{27}$$

式中 $w_1 = (\sum_t f_t f_t')^{-1}(\sum_t f_t y_t)$,为 w 的 OLS 估计.

式(27)表明,组合权重向量为非样本估计与样本估计的矩阵加权平均.为进一步简化,还可取 $\tau = \gamma(\sum_t f_t f_t')$, $\gamma > 0$,则式(27)简化为

$$\hat{w} = \frac{\gamma}{\gamma+1}w_0 + \frac{1}{\gamma+1}w_1 \tag{28}$$

从式(28)可知, γ 反映了对非样本信息的信任度及相应的重视程度.

4 结束语

本文评述了组合预测的基本思想和组合预测贝叶斯模型的研究现状，导出了单项预测仅存在位置偏差和同时存在尺度偏差两种情况下多步贝叶斯信息更新的无偏组合预测模型，还给出了结合非样本信息和样本信息的组合权重贝叶斯更新模型，进一步扩展了贝叶斯组合预测模型的现有成果. 除本文涉及的基本思路外，组合预测贝叶斯模型还可从其他方向加以发展. 如建立组合权重的随机时变模型[9]，采用贝叶斯方法进行适应性预测，在这方面文献[27]已作了有益的尝试. 此外，在结合非样本信息和样本信息方面，各种收缩估计量(如斯坦规则估计量和经验贝叶斯估计量)具有优良的性能[30]，值得应用于组合权重的估计.

参考文献

1 Bates J M, Granger C W J. The combioation of forecasts. Operatiooal Research Quarterly, 1969; 20(4): 451～468

2 Granger C W J, Newbold P. Forecasting economic time series. 2nd ed. Orlando: Acadmic Press, 1986

3 Wiokler R L. Combiniog forecasts: a philosophical basis and some current issues. Int. J. Forecastiog, 1989; 5: 605～609

4 Makridakis S, Winkler R L. Averages of forecasts: some empirical results. Managemeot Science, 1983; 29(9): 987～996

5 West C T. System-based weights versus series-specific weights in the combination of forecasts. J. Forecasting, 1996; 15(5): 369～383

6 Chatfield C. Model uncertainty aod forecast accuracy. J. Forecastiog, 1996; 15(7): 495～508

7 Tibiletti L. A noo-linear combioatioo of experts' forecasts: a Bayesian approach. J. Forecasting, 1994; 13(1): 21～27

8 Donaldson R G, Kamstra M. Forecast combining with neural networks. J. Forecasting, 1996; 15(1): 46～61

9 West M, Harrison J. Baysesian Forecasting and Dynamic Models. New York: Sprioger-Verlag, 1989

10 Buno D W. A Bayesiao approach to the lioear combination of forecasts. Operational Research Quarterly, 1975; 26: 325～329

11 Bunn D W. A simplificatioo of the matrix beta distribution for combining estimator. J. Operational Research Society, 1978; 29: 1013～1016

12 Fana A E, Souza R C. A re-evaluatioo of the quasi-Bayes approach to the lioear combioation of forecasts. J. Forecasting, 1995; 14(6): 553～542

13 Min C K, Zellner A. Bayesian aod oon-Bayesian methods for combiniog models and forecasts with application to forecasting international growth rate. J. Econometrics, 1993; 56: 89～118

14 Draper D. Assessmeot and propagatioo of model uncertaointy. J. Royal Statistics Society, Series B, 1995; 57: 479～488

15 Wiokler R T. Combioing probability distributioo from dependent ioformation sources. Maoagement Science, 1981; 27: 479～488

16 Bordley R F. The combination of forecasts: a Bayesian approach. J. Operational Research Society, 1982; 33(2): 171～174

17 Liodley D. Recoociliatioo of probability distribution. Operations Research, 1983; 31: 866～880

18 Granger C W J, Ramanathan R. Improved methods of combining information. J. Forecasting, 1984; 3: 197～204

19 Bordley R F. Linear combioatioo of forecasts with ao ioterceept: a Bayesian approach. J. Forecasting, 1986; 5(4): 243～249

20 Anandalingam G, Chen L. Linear combinatioo of forecasts: A general Bayesian model. J. Forecasting, 1989; 8(3): 199～214

21 Aoaodaliogam G, Cheo L. Bayesiao forecast combioatioo and Kalmao filtering. Int. J. System Science, 1989; 20(8): 1499～1507

22 曾 勇，唐小我. 几种无偏组合预测模型的分布. 数量经济技术经济研究，1996; (11): 41～45

23 高仁祥，张世英，刘 豹. 组合预测贝叶斯方法研究. 系统工程学报，1996; 11(1): 28～35

24 言茂松. 贝叶斯风险决策工程. 北京:清华大学出版社,1989

25 Kang H. Ustable weights in the combination of forecasts. Management Science,1986;32(6):683~695

26 Holden K,Peel D A. Unbiasedness,efficiency and the combination of economic forecasts. J. Forecasting,1989;8(3):175~188

27 Sessions D N,Chatterjee S. The combining of forecasting using recursive techniques with non-stationary weights. J. Forecasing,1989;8(3):239~251

28 Diebold F X. Serial correlation and the combination of forecasts. J. Business & Economic Statistics,1988;6:105~111

29 [美]克里夫·W·J·格兰杰著. 商业与经济预测. 何军等译. 上海:上海远东出版社,1992

30 Judge G G,Bock M E. The statistical implications of pre-test and stein-rule estimators in econometrics. Amsterdam: North-Holland,1978

Bayesian Models of Combining Forecasts

Zeng Yong, Tang Xiaowo

Management College, University of Electronic Science and Technology of China

Zheng Weimin

School of Economics and Management, Tsinghua University

Abstract In this paper, the basic ideas of combination forecasting and the current situation of studies on the Bayesian models of combining forecasts are reviewed. One of the reasearch procadures is extended to the case of multi-step Bayesian updation of information, and two Bayesian models of unbiased combining forecasts are conducted under different assumptions of characteristics of individual forecast bias. Furthermore, a Bayesian model of combining weights is developed with non-sample and sample information being combined. Finally, areas of further research are pointed out.

Keywords: combination forecasting, model uncertainty, Bayesian models, unbiased forecasts, non-sample information, shrinkage estimators