第9卷 第2期 华侨大学学报自然科学版 Vol.9 No.2
1988年5月 JOURNAL OF HUAQIAO UNIVERSITY (NATURAL SCIENCE) May 1988

逐步回归最小方差集成预报法

陈 建 伟

（应用数学系）

摘　　要

本文在不同信度下，选出多个逐步回归方程．在最小均方意义下，建立可变动最优集成预报方程，所确定的集成系数是和各时刻的预报值有关的滑动系数，它能及时地反映单个回归方程在各时刻的变化规律，从而充分地采用单个回归方程在各时刻的预报信息．本文导出整个计算过程的递推公式，便于计算机程序设计，並证明这种集成预报方程比单个回归方程提高了预报精度和可靠性．

一、前　　言

在回归分析中，由于使用了电子计算机，对每一个预报对象可以找到许多相关因子．但是，用逐步回归法只能选入其中极少数的几个因子，因而浪费了大量因子的信息，所建立的预报方程其效果也不一定很好．为了充分地利用因子的信息，以建立效果更好的预报方程，我们首先在不同的信度下建立多个逐步回归方程，然后再将这些方程集成，在预报均方误差达到最小下确定各时刻的集成系数，建立最优预报方程．将看到所求的集成系数与所对应方程在各时刻的拟合程度有关，由于集成系数随时间的不断变化来确定各回归预报方程对集成预报方程的贡献大小．我们将证明用集成预报方程来预测是无偏的，并且预测误差的方差比其它单个回归方程的预测误差的方差都小．所以集成预报方程是较优的预报方程，它综合地利用各方程的优点，充分地利用大量因子信息，从而能得到良好的预报效果．这样做或许是提高预报及控制效果的途径之一．

二、方法讨论

设预报对象 y 与相关因子 x_1，x_2，…，x_s 有 N 组独立观察值

$$(y_i;\ x_{1i},\ x_{2i},\ \cdots,\ x_{si}),\qquad i=1,\ 2,\ \cdots,\ N$$

本文1987年3月19日收到．

记

$$Y=\begin{pmatrix} y_1\\ y_2\\ \vdots\\ y_N\end{pmatrix},\quad X=\begin{pmatrix} 1 & x_{11} & x_{21} & \cdots & x_{s1}\\ 1 & x_{12} & x_{22} & \cdots & x_{s2}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 1 & x_{1N} & x_{2N} & \cdots & x_{sN}\end{pmatrix},\quad \beta=\begin{pmatrix}\beta_0\\ \beta_1\\ \vdots\\ \beta_s\end{pmatrix},\quad \varepsilon=\begin{pmatrix}\varepsilon_1\\ \varepsilon_2\\ \vdots\\ \varepsilon_N\end{pmatrix}$$

并且 Y 与 X 满足线性关系式

$$Y=X\beta+\varepsilon \tag{1}$$

其中 ε 为随机向量，且 ε 服从 $N(0,\ \sigma^2 I_N)$，σ^2 是未知量，I_N 是 N 阶单位矩阵。

用逐步回归法建立 m 个回归方程：

先给定一个较高的显著性水平 $\alpha_1(0<\alpha_1<1)$，用逐步回归法建立 y 与因子的第一个回归方程

$$\hat{y}^{(1)}=\hat{\beta}_0^{(1)}+\hat{\beta}_1^{(1)}x_1^{(1)}+\cdots+\hat{\beta}_{m_1}^{(1)}x_{m_1}^{(1)}$$

再给一个低一点的显著性水平 $\alpha_2(>\alpha_1)$，建立第二个回归方程

$$\hat{y}^{(2)}=\hat{\beta}_0^{(2)}+\hat{\beta}_1^{(2)}x_1^{(2)}+\cdots+\hat{\beta}_{m_2}^{(2)}x_{m_2}^{(2)}$$

如此继续做下去，直到在某一低的显著性水平下已挑不出因子来建立回归方程为止。设最后共建立 m 个回归方程

$$\hat{y}^{(l)}=\hat{\beta}_0^{(l)}+\hat{\beta}_1^{(l)}x_1^{(l)}+\cdots+\hat{\beta}_{m_l}^{(l)}x_{m_l}^{(l)},\qquad l=\overline{1,\ m} \tag{2}$$

其中 $x_k^{(l)}\in\{x_1,\ x_2,\ \cdots,\ x_s\}$ $k=\overline{1,m_l},\ l=\overline{1,\ m}$，且各不相同，$\sum\limits_{l=1}^{m}m_l<s$。

由此得到以上各回归方程的参数估计

$$\hat{\beta}^{(l)}=(X^{\tau(l)}X^{(l)})^{-1}X^{\tau(l)}Y,\qquad l=\overline{1,m} \tag{3}$$

其中

$$\hat{\beta}^{(l)}=\begin{pmatrix}\hat{\beta}_0^{(l)}\\ \hat{\beta}_1^{(l)}\\ \vdots\\ \hat{\beta}_{m_l}^{(l)}\end{pmatrix},\quad X^{(l)}=\begin{pmatrix} 1 & x_{11}^{(l)} & x_{21}^{(l)} & \cdots & x_{m_l 1}^{(l)}\\ 1 & x_{12}^{(l)} & x_{22}^{(l)} & \cdots & x_{m_l 2}^{(l)}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 1 & x_{1N}^{(l)} & x_{2N}^{(l)} & \cdots & x_{m_l N}^{(l)}\end{pmatrix}_{N\times(m_l+1)},\quad Y=\begin{pmatrix}y_1\\ y_2\\ \vdots\\ y_N\end{pmatrix}$$

实践证明，在较高的置信度下建立的回归方程预报效果未必最好，为了更充分地利用各回归方程在各时刻的预报能力，综合各个回归方程的预报信息，我们建立集成回归方程使各时刻预报均方误差达到最小。即

$$\hat{y}_t=\sum_{l=1}^{m}C_l(t)\,\hat{y}_t^{(l)},\qquad t=1,\ 2,\ \cdots \tag{4}$$

其中 $\sum\limits_{l=1}^{m}C_l(t)=1(t=1,\ 2,\ \cdots)$，$C_l(t)(l=\overline{1,\ m})$ 为待定系数，称式（4）为集成回归预报方程，$C_l(t)(l=\overline{1,\ m})$ 称为 t 时刻的集成系数。这样建立的预报方程在 t 时刻的预报均方误差为

$$E(y_t-\hat{y}_t)^2=E\,(y_t-\sum_{l=1}^{m}C_l(t)\,\hat{y}_t^{(l)})^2=E(\sum_{l=1}^{m}C_l(t)(y_t-\hat{y}_t^{(l)}))^2$$

利用拉格朗日数乘法求解在条件

$$\varphi \equiv 1-\sum_{l=1}^{m} C_l(t)=0$$

下使上式达到极小值的 $C_l(t)$ $(l=\overline{1,m})$，令

$$\frac{\partial E(y_t-\hat{y}_t)^2}{\partial C_l(t)}+\lambda\frac{\partial \varphi}{\partial C_l(t)}=0, \qquad (l=\overline{1,m})$$

由此得

$$2E[\sum_{j=1}^{m} C_j(t)(y_t-\hat{y}_t^{(j)})(y_t-\hat{y}_t^{(l)})]=\lambda$$

$$\sum_{j=1}^{m} C_j(t)E(y_t-\hat{y}_t^{(j)})(y_t-\hat{y}_t^{(l)})=\frac{\lambda}{2}, \qquad (l=\overline{1,m})$$

记

$$b_{lj}^{(t)} \equiv E(y_t-\hat{y}_t^{(l)})(y_t-\hat{y}_t^{(j)})=E\epsilon_t^{(l)}\epsilon_t^{(j)}$$

其中

$$\epsilon_t^{(l)}=y_t-\hat{y}_t^{(l)}, \qquad (l,\ j=\overline{1,m},\ t\geqslant 1)$$

则得方程组

$$\sum_{j=1}^{m} C_j(t)b_{lj}(t)=\frac{\lambda}{2}, \qquad (l=\overline{1,m})$$

$$\sum_{j=1}^{m} C_j(t)=1$$

将前面 m 个方程写成矩阵形式，即

$$\mathbf{B}(t)\mathbf{C}(t)=\mathbf{D} \tag{5}$$

其中

$$\mathbf{B}(t)=\begin{pmatrix} b_{11}(t) & b_{12}(t) & \cdots & b_{1m}(t)\\ b_{21}(t) & b_{22}(t) & \cdots & b_{2m}(t)\\ \vdots & \vdots & \ddots & \vdots\\ b_{m1}(t) & b_{m2}(t) & \cdots & b_{mm}(t)\end{pmatrix}_{m\times m}, \quad \mathbf{C}(t)=\begin{pmatrix} C_1(t)\\ C_2(t)\\ \vdots\\ C_m(t)\end{pmatrix}_{m\times 1}, \quad \mathbf{D}=\begin{pmatrix}\lambda/2\\ \lambda/2\\ \vdots\\ \lambda/2\end{pmatrix}_{m\times 1}$$

可见式(5)的系数矩阵 $\mathbf{B}(t)$ 是对称矩阵，又由于 $E(y_t-\hat{y}_t)^2\geqslant 0$，可进一步断定 $E(y_t-\hat{y}_t)^2>0$。事实上，如果 $E(y_t-\hat{y}_t)^2=0$，那么 $p(y_t=\hat{y}_t)=1$，而在实际中预报值 $\hat{y}_t$ 与实际值 y_t 是不会相等的，因事件 $\{y_t=\hat{y}_t\}$ 是小概率事件，由实际推断原理可断定事件 $\{y_t=\hat{y}_t\}$ 实际是不会发生的。这与 $p(y_t=\hat{y}_t)=1$ 矛盾，从而得 $E(y_t-\hat{y}_t)^2\neq 0$，故 $E(y_t-\hat{y}_t)^2>0$。于是

$$\sum_{j=1}^{m}\sum_{l=1}^{m} C_j(t)C_l(t)b_{jl}(t)=\sum_{j=1}^{m}\sum_{l=1}^{m} C_j(t)C_l(t)E(y_t-\hat{y}_t^{(j)})(y_t-\hat{y}_t^{(l)})$$

$$=E[\sum_{j=1}^{m}\sum_{l=1}^{m} C_j(t)C_l(t)(y_t-\hat{y}_t^{(j)})(y_t-\hat{y}_t^{(l)})]$$

$$=E[\sum_{l=1}^{m} C_l(t)(y_t-\hat{y}_t^{(l)})]^2=E(y_t-\hat{y}_t)^2>0$$

则 $\sum_{j=1}^{m}\sum_{l=1}^{m} C_j(t)C_l(t)b_{jl}(t)$ 是正定二次型，故其系数矩阵的行列式大于零，因此方程组(5)有

唯一解。记 $B^{-1}(t)=[a_{ij}(t)]_{m\times m}$ 则

$$C_l(t)=\frac{\lambda}{2}\sum_{j=1}^{m}a_{lj}(t), \qquad (l=\overline{1,\ m})$$

由条件 $\sum_{l=1}^{m}C_l(t)=1$ 得方程组(5)的唯一解

$$\hat{\lambda}=\frac{2}{\sum_{j=1}^{m}\sum_{l=1}^{m}a_{lj}(t)}; \qquad \hat{C}_l(t)=\frac{\sum_{j=1}^{m}a_{lj}(t)}{\sum_{j=1}^{m}\sum_{i=1}^{m}a_{ij}(t)}, \qquad l=\overline{1,\ m};\ t=1,\ 2,\ \cdots \tag{6}$$

以下进一步说明集成回归预报方程预测具有良好的性质。

定理 1　按式(6)确定集成系数的集成预报方程是无偏的，并且预测误差的方差比式(2)所建立的单个回归预报方程都小。

证明　由集成预报方程(4)得

$$E\hat{y}_t=E\left\{\sum_{l=1}^{m}\hat{C}_l(t)\hat{y}_t^{(l)}\right\}=\sum_{l=1}^{m}\hat{C}_l(t)E\hat{y}_t^{(l)}=\sum_{l=1}^{m}\hat{C}_l(t)Ey_t=Ey_t$$

故集成方程预报值 $\hat{y}_t$ 是 Ey_t 的无偏估计，从而得预测是无偏的。

其次，对单个预报方程(2)的预报误差方差

$$\mathrm{Var}(y_t-\hat{y}_t^{(l)})=E(y_t-\hat{y}_t^{(l)})^2, \qquad (l=\overline{1,\ m})$$

令

$$C_i(t)=\begin{cases}1 & i=l\\ 0 & i\neq l\end{cases}$$

则有

$$E(y_t-\hat{y}_t^{(l)})^2=E(y_t-\hat{y}_t)^2, \qquad t=1,\ 2,\ \cdots$$

由上面证明得知它不是集成预报均方误差的最小值，故按式(6)定系数的集成回归预报均方误差 $E(y_t-\hat{y}_t)^2$ 必满足

$$E(y_t-\hat{y}_t)^2<E(y_t-\hat{y}_t^{(l)})^2, \qquad (l=\overline{1,\ m})$$

从而

$$\mathrm{Var}(y_t-\hat{y}_t)=E(y_t-\hat{y}_t)^2<\mathrm{Var}(y_t-\hat{y}_t^{(l)}), \qquad (l=\overline{1,\ m})\quad t=1,\ 2,\ \cdots$$

故最小方差集成预报方程是最优预报方程。

由式(5)、(6)可见只要能求出 $E\hat{\epsilon}_t^{(l)}\ \hat{\epsilon}_t^{(j)}(l,\ j=\overline{1,\ m})$ 的值，就可由式(6)求出集成系数，建立最小方差集成预报方程(4)。

定理 2　在假设的线性模型下，则

$$E\hat{\epsilon}_t^{2(l)}=\begin{cases}\sigma^2(1-\mathbf{e}_tX^{(l)}(X^{\tau(l)}X^{(l)})^{-1}X^{\tau(l)}\mathbf{e}_t^{\tau}), & t\leqslant N\\ \sigma^2(1+X^{(l)}{}_t(X^{\tau(l)}X^{(l)})^{-1}X_t^{\tau(l)}), & t>N\end{cases} \qquad (l=\overline{1,\ m}) \tag{7}$$

$$E\hat{\epsilon}_t^{(l)}\hat{\epsilon}_t^{(j)}=\begin{cases}\sigma^2\mathbf{e}_t(I_N-X^{(l)}(X^{\tau(l)}X^{(l)})^{-1}X^{\tau(l)})(I_N-X^{(j)}(X^{\tau(l)}X^{(j)})^{-1}X^{\tau(j)})\mathbf{e}_t^{\tau}, & t\leqslant N\\ \sigma^2(1+X_t^{(l)}(X^{\tau(l)}X^{(l)})^{-1}X^{\tau(l)}X^{(j)}(X^{\tau(j)}X^{(j)})^{-1}X_t^{\tau(j)}), & t>N\end{cases}$$

$$(l,j=\overline{1,\ m}) \tag{8}$$

其中 $\mathbf{e}_1$，$\mathbf{e}_2$，…，$\mathbf{e}_N$ 都是 N 维单位向量，$\mathbf{I}_N$ 是 N 阶单位矩阵，$\mathbf{X}_t^{(l)}=(1,\ x_{1t}^{(l)},\cdots,\ x_{m_l t}^{(l)})^\tau$ 是 t 时刻 $(t=N)$ 的当前值向量。

证明　由已知假设的线性模型条件可得当 $t>N$ 时，y_t 与 $\hat{y}_t^{(l)}(l=\overline{1,m})$ 互不相关，则

$$E(y_t-Ey_t)(\hat{y}_t^{(l)}-E\hat{y}_t^{(l)})=0,\qquad (l=\overline{1,m})\tag{9a}$$

当 $t=1,2,\cdots,N$ 时，由于 $Ey_t=E\hat{y}_t^{(l)}$，$(l=\overline{1,m})$，$Ey_t^2=\sigma^2+(Ey_t)^2$，记 $(a_1^{(l)},\cdots,a_N^{(l)})$ $\triangle\mathbf{e}_t\mathbf{X}^{(l)}(\mathbf{X}^{\tau(l)}\mathbf{X}^{(l)})^{-1}\mathbf{X}^{\tau(l)}$，$(l=\overline{1,m},\ t=\overline{1,N})$，所以

$$\begin{aligned}
Ey_t\hat{y}_t^{(l)}&=E[y_t\mathbf{e}_t\mathbf{X}^{(l)}(\mathbf{X}^{\tau(l)}\mathbf{X}^{(l)})^{-1}\mathbf{X}^{\tau(l)}\mathbf{Y}]=E[y_t(a_1^{(l)},\cdots,a_N^{(l)})(y_1,y_2,\cdots,y_N)^\tau]\\
&=E[y_t\sum_{k=1}^N a_k^{(l)}y_k]=a_t^{(l)}Ey_t^2+\sum_{\substack{k=1\\k\neq t}}^N a_k^{(l)}Ey_ty_k=a_t^{(l)}(\sigma^2+(Ey_t)^2)+\sum_{\substack{k=1\\k\neq t}}^N a_k^{(l)}Ey_ty_k\\
&=a_t^{(l)}\sigma^2+Ey_t\sum_{k=1}^N a_k^{(l)}Ey_k=\sigma^2\ (a_1^{(l)},\cdots,\ a_N^{(l)})e_t^\tau\ +Ey_tE(a_1^{(l)},\cdots,\ a_N^{(l)})(y_1,\cdots,y_N)^\tau\\
&=\sigma^2\mathbf{e}_t\mathbf{X}^{(l)}(\mathbf{X}^{\tau(l)}\mathbf{X}^{(l)})^{-1}\mathbf{X}^{\tau(l)}\mathbf{e}_t^\tau+Ey_tE\hat{y}_t^{(l)}
\end{aligned}\tag{9b}$$

从而得

$$E(y_t-Ey_t)(\hat{y}_t^{(l)}-E\hat{y}_t^{(l)})=\sigma^2\mathbf{X}^{(l)}(\mathbf{X}^{\tau(l)}\mathbf{X}^{(l)})^{-1}\mathbf{X}^{\tau(l)}\mathbf{e}_t^\tau,\qquad (l=\overline{1,m},\ t=\overline{1,N})$$

故

$$\begin{aligned}
E\hat{\epsilon}_t^{2(l)}&=E(y_t-\hat{y}_t^{(l)})^2=E(y_t-Ey_t)^2+E(\hat{y}_t^{(l)}-E\hat{y}_t^{(l)})^2-2E(y_t-Ey_t)(\hat{y}_t^{(l)}-E\hat{y}_t^{(l)})\\
&=\begin{cases}\sigma^2+\sigma^2\mathbf{e}_t\mathbf{X}^{(l)}(\mathbf{X}^{\tau(l)}\mathbf{X}^{(l)})^{-1}\mathbf{X}^{\tau(l)}\mathbf{e}_t^\tau-2\sigma^2\mathbf{e}_t\mathbf{X}^{(l)}\ (\mathbf{X}^{\tau(l)}\mathbf{X}^{(l)})^{-1}\mathbf{X}^{\tau(l)}\mathbf{e}_t^\tau, & t\leqslant N\\ \sigma^2+\sigma^2\mathbf{X}_t^{(l)}\ (\mathbf{X}^{\tau(l)}\mathbf{X}^{(l)})^{-1}\mathbf{X}_t^{\tau(l)}, & t>N\end{cases}\\
&\qquad (l=\overline{1,m})
\end{aligned}\tag{9c}$$

即式(7)得证。又由于当 $t\leqslant N$ 时

$$\begin{aligned}
E\hat{y}_t^{(l)}\hat{y}_t^{(j)}&=E[\mathbf{e}_t\mathbf{X}^{(l)}\ (\mathbf{X}^{\tau(l)}\mathbf{X}^{(l)})^{-1}\mathbf{X}^{\tau(l)}\mathbf{Y}\mathbf{e}_t\mathbf{X}^{(j)}(\mathbf{X}^{\tau(j)}\mathbf{X}^{(j)})^{-1}\mathbf{X}^{\tau(j)}\mathbf{Y}]\\
&=E[(a_1^{(l)},\ \cdots,a_N^{(l)})\mathbf{Y}(a_1^{(j)},\ \cdots,a_N^{(j)})\mathbf{Y}]\\
&=E[\sum_{k=1}^N a_k^{(l)}y_k\sum_{i=1}^N a_i^{(j)}y_i]=\sum_{k=1}^N a_k^{(l)}a_k^{(j)}Ey_k^2\ +\sum_{k\neq i}^N a_k^{(l)}a_i^{(j)}Ey_kEy_i\\
&=\sigma^2\sum_{k=1}^N a_k^{(l)}a_k^{(j)}+\sum_{k=1}^N a_k^{(l)}a_k^{(j)}(Ey_k)^2+\sum_{k\neq i}^N a_k^{(l)}a_i^{(j)}Ey_kEy_j\\
&=\sigma^2\sum_{k=1}^N a_k^{(l)}a_k^{(j)}+E(\sum_{k=1}^N a_k^{(l)}y_k)E\ (\sum_{i=1}^N a_i^{(j)}y_i)\\
&=\sigma^2(a_1^{(l)},\ \cdots,\ a_N^{(l)})(a_1^{(j)},\ \cdots,\ a_N^{(j)})^\tau+E[(a_1^{(l)},\ \cdots,\ a_N^{(l)})Y]E[(a_1^{(j)},\ \cdots,\ a_N^{(j)})Y]\\
&=\sigma^2\mathbf{e}_t\mathbf{X}^{(l)}(\mathbf{X}^{\tau(l)}\mathbf{X}^{(l)})^{-1}\mathbf{X}^{\tau(l)}\mathbf{X}^{(j)}(\mathbf{X}^{\tau(j)}\mathbf{X}^{(j)})^{-1}\mathbf{X}^{\tau(j)}\mathbf{e}_t^\tau+E\hat{y}_t^{(l)}E\hat{y}_t^{(j)}
\end{aligned}\tag{9d}$$

同理可推出当 $t>N$ 时

$$E\hat{y}_t^{(l)}y_t^{(j)}=\sigma^2\mathbf{X}_t^{(l)}(\mathbf{X}^{\tau(l)}\mathbf{X}^{(l)})^{-1}\mathbf{X}^{\tau(l)}\mathbf{X}^{(j)}(\mathbf{X}^{\tau(j)}\mathbf{X}^{(j)})^{-1}\mathbf{X}_t^{\tau(j)}+E\hat{y}_t^{(l)}E\hat{y}_t^{(j)}\tag{9e}$$

故由式(8a)至(8e)得

$$E\hat{\epsilon}_t^{(l)}\hat{\epsilon}_t^{(j)}=E(y_t-\hat{y}_t^{(l)})(y_t-\hat{y}_t^{(j)})\ =Ey_t^2+\ E\hat{y}_t^{(l)}\hat{y}_t^{(j)}-\ E\hat{y}_t^{(l)}y_t-\ Ey_t\hat{y}_t^{(j)}$$

$$
=\begin{cases}\sigma^2-\sigma^2 e_t X^{(l)}(X^{\tau(l)}X^{(l)})^{-1}X^{\tau(l)}e_t^{\tau}-\sigma^2 e_t X^{(j)}(X^{\tau(j)}X^{(j)})^{-1}X^{\tau(j)}e_t^{\tau} \\ \qquad +\sigma^2 e_t X^{(l)}(X^{\tau(l)}X^{(l)})^{-1}X^{\tau(l)}X^{(j)}(X^{\tau(j)}X^{(j)})^{-1}X^{(j)}e_t^{\tau}, & t\leqslant N \\ \sigma^2+\sigma^2 X_t^{(l)}(X^{\tau(l)}X^{(l)})^{-1}X^{\tau(l)}X^{(j)}(X^{\tau(j)}X^{(j)})^{-1}X_t^{\tau(j)}, & t>N\end{cases}
$$

$$
=\begin{cases}\sigma^2 e_t(I_N-X^{(l)}(X^{\tau(l)}X^{(l)})^{-1}X^{\tau(l)})(I_N-X^{(j)}(X^{\tau(j)}X^{(j)})^{-1}X^{\tau(j)})e_t^{\tau}, & t\leqslant N \\ \sigma^2(1+X_t^{(l)}(X^{\tau(l)}X^{(l)})^{-1}X^{\tau(l)}X^{(j)}(X^{\tau(j)}X^{(j)})^{-1}X_t^{\tau(j)}), & t>N\end{cases}
$$

$$(l,\ j=\overline{1,\ m},\ l\neq j)$$

即式(8)得证。

由定理2可见，每个 $E\hat{\varepsilon}_t^{(l)}\hat{\varepsilon}_t^{(j)}(l,\ j=\overline{1,\ m})$ 都与 σ^2 成比例，按式(6)求集成系数时可消去 σ^2，因此集成系数与 σ^2 无关，即由已知因子的观察值和 $t>N$ 时刻的当前值就能求得 $E\hat{\varepsilon}_t^{(l)}\hat{\varepsilon}_t^{(j)}$ $(l,\ j=\overline{1,\ m},\ t=1,\ 2,\ \cdots)$，从而可建立最小方差集成预报方程。

三、用 Cholesky 分解法求集成系数

从以上可看到，要求集成系数首先应求逆矩阵 $(X^{\tau(l)}X^{(l)})^{-1}(l=\overline{1,\ m})$，为了避免求逆阵的繁琐运算，便于计算机程序设计，我们采用 Cholesky 分解法分别计算逆矩阵 $(X^{\tau(l)}X^{(l)})^{-1}$ 及集成系数。

1. 逆矩阵 $(X^{\tau(l)}X^{(l)})^{-1}$ 的计算设计

因 $X^{(l)}(l=\overline{1,\ m})$ 是满秩的，则 $X^{\tau(l)}X^{(l)}\ (l=\overline{1,m})$ 是正定对称阵，据 Cholesky 分解法可将 $(X^{\tau(l)}X^{(l)})$ 分解为 $R^{(l)}R^{*(l)}$ ($R^{(l)}$ 为下三角阵；$R^{*(l)}$ 为上三角阵)，即

$$X^{\tau(l)}X^{(l)}\triangleq(d_{ij}^{(l)})_{(m_l+1)\times(m_l+1)}=R^{(l)}R^{*(l)},\qquad (l=\overline{1,\ m})$$

其中

$$
R^{(l)}=\begin{pmatrix} r_{11}(l) & \cdots & \cdots & 0 \\ r_{21}(l) & r_{22}(l) & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ r_{(m_l+1)1}(l) & r_{(m_l+1)2}(l) & \cdots & r_{(m_l+1)(m_l+1)}(l)\end{pmatrix}_{(m_l+1)\times(m_l+1)},
$$

$$
R^{*(l)}=\begin{pmatrix} 1 & \dfrac{r_{21}(l)}{r_{11}(l)} & \cdots & \dfrac{r_{(m_l+1)1}(l)}{r_{11}(l)} \\ \vdots & 1 & \cdots & \dfrac{r_{(m_l+1)2}(l)}{r_{22}(l)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & 1\end{pmatrix}_{(m_l+1)\times(m_l+1)}
$$

且

$$
\begin{cases} r_{ji}(l)=d_{ij}^{(l)}-\displaystyle\sum_{k=1}^{i-1}\frac{r_{ik}(l)r_{jk}(l)}{r_{kk}(l)}, & (i<j) \\ r_{ii}(l)=d_{ii}^{(l)}-\displaystyle\sum_{k=1}^{i-1}\frac{r_{ik}^2(l)}{r_{kk}(l)}, & (i=j)\end{cases}\qquad l=\overline{1,\ m}\qquad (10)
$$

令

$$(X^{\tau(l)}X^{(l)})^{-1}\triangleq(A_{ij}^{(l)})_{(m_l+1)\times(m_l+1)}=(A_1^{(l)},\ \cdots,\ A_{m_l+1}^{(l)}),\qquad l=\overline{1,\ m}$$

由于

$$(X^{\tau(l)}X^{(l)})(X^{\tau(l)}X^{(l)})^{-1}=(e_1^{(l)},\cdots,e_{m_l+1}^{(l)})$$

其中，$e_1^{(l)},\cdots,e_{m_l+1}^{(l)}$，是$(m_l+1)$维单位向量，所以

$$(X^{\tau(l)}X^{(l)})A_r^{(l)}=e_r^{(l)},\qquad (r=1,\cdots,m_l+1) \tag{11}$$

故只要计算出各 $A_r^{(l)}$，$(X^{\tau(l)}X^{(l)})^{-1}$ 也就计算出来了。

而 $A_r^{(l)}$ 正是方程组

$$X^{\tau(l)}X^{(l)})A_r^{(l)}=e_r^{(l)},\qquad (r=1,\cdots,m_l+1;\quad l=\overline{1,m})$$

的解，由式(10)得

$$(R^{(l)}R^{*(l)})A_r^{(l)}=e_r^{(l)},\qquad (r=1,\cdots,m_l+1)$$

为了计算 $A_r^{(l)}(r=\overline{1,m_l+1};l=\overline{1,m})$，利用二次回代.其步骤为：对每个 l 及 $r(r=\overline{1,m_l+1};$ $l=\overline{1,m})$

(1)　先解方程组　$R^{(l)}T_r^{(l)}=e_r^{(l)}$

由此得中间解

$$T_r^{(l)}=(t_{r1}^{(l)},\cdots,t_{r(m_l+1)}^{(l)})^{\tau}$$

其求解公式

$$t_{ri}^{(l)}=(\delta(i,r)-\sum_{k=1}^{i-1}r_{ik}(l)t_{rk}^{(l)})/r_{ii}(l) \tag{12}$$

其中

$$\delta(i,r)=\begin{cases}1 & i=r\\ 0 & i\neq r\end{cases}\qquad (i=1,2,\cdots,m_l+1)$$

(2)　回代解方程组　$R^{*(l)}A_r^{(l)}=T_r^{(l)}$

得　$A_r^{(l)}=(A_{1r}^{(l)},\cdots,A_{m_l+1,r}^{(l)})^{\tau}$ 的求解公式

$$A_{ir}^{(l)}=t_{ri}^{(l)}-\sum_{k=1}^{m_l+1-i}\frac{r_{k+i,i}(l)A_{k+i,r}^{(l)}}{r_{ii}(l)},\qquad (i=m_l+1,m_l,\cdots,1) \tag{13}$$

约定以上和式当上限小于下限时，其和式为零。

综合式(11)、(12)、(13)能计算 $(X^{\tau(l)}X^{(l)})^{-1}=(A_{ij}^{(l)})_{(m_l+1)\times(m_l+1)}$，$(l=\overline{1,m})$

2. 集成系数的计算设计

类似于1，我们同样用 Cholesky 分解法将正定对称阵 $B(t)$分解为下三角阵 $L(t)$ 和上三角阵 $U(t)$的乘积，即

$$B(t)=L(t)U(t)$$

$$L(t)=\begin{pmatrix} l_{11}(t) & \cdots & \cdots & 0\\ l_{21}(t) & l_{22}(t) & \cdots & \vdots\\ \vdots & \vdots & \ddots & \vdots\\ l_{m1}(t) & l_{m2}(t) & \cdots & l_{mm}(t)\end{pmatrix}_{m\times m},\quad U(t)=\begin{pmatrix} 1 & \frac{l_{21}(t)}{l_{11}(t)} & \cdots & \frac{l_{m1}(t)}{l_{11}(t)}\\ \cdots & 1 & \cdots & \frac{l_{m2}(t)}{l_{22}(t)}\\ \vdots & \vdots & \ddots & \vdots\\ 0 & \cdots & \cdots & 1\end{pmatrix}_{m\times m}$$

且

$$\begin{cases} l_{ji}(t)=b_{ij}(t)-\sum\limits_{k=1}^{i-1}\dfrac{l_{ik}(t)l_{jk}(t)}{l_{kk}(t)}, & (i<j) \\ l_{ii}(t)=b_{ii}(t)-\sum\limits_{k=1}^{i-1}\dfrac{l_{ik}^2(t)}{l_{kk}(t)}, & (i=j) \end{cases} \tag{14}$$

所以方程(5)化为

$$L(t)U(t)(C_1(t),\ \cdots,\ C_m(t))=(\frac{\lambda}{2},\ \cdots,\ \frac{\lambda}{2})^\tau$$

为了求出集成系数$C_l(t)$ ($l=\overline{1,m}$, $t>0$)，作变换化方程为

$$L(t)U(t)\ (\frac{C_1(t)}{\lambda},\ \cdots,\ \frac{C_m(t)}{\lambda})^\tau=(\frac{1}{2},\ \frac{1}{2},\ \cdots,\ \frac{1}{2})^\tau$$

其中利用了$\lambda\neq 0$的性质。

仍采用二次回代，其步骤为：

（1）先解方程组　$L(t)A(t)=(\frac{1}{2},\ \cdots,\ \frac{1}{2})^\tau$

由此得中间解

$$A(t)=(a_1(t),\ \cdots,\ a_m(t))^\tau$$

$$a_i(t)=(\frac{1}{2}-\sum_{k=1}^{i-1}l_{ik}(l)a_k(t))/l_{ii}(t),\qquad (i=1,\ \cdots,\ m) \tag{15}$$

（2）回代解方程组　$U(t)(\frac{C_1(t)}{\lambda},\ \cdots,\ \frac{C_m(t)}{\lambda})=A(t)$

得$C(t)=(C_1(t),\ \cdots,\ C_m(t))^\tau$的求解公式

$$C_j(t)=\lambda a_j(t)-\sum_{k=1}^{m-j}\frac{l_{k+j,j}(t)C_{k+j}(t)}{l_{jj}(t)},\qquad (j=m,\ m-1,\ \cdots,\ 1) \tag{16}$$

约定以上和式当上限小于下限时，其和式为零。令

$$C_j^*(t)=\frac{C_j(t)}{\lambda},\qquad (j=1,\ \cdots,\ m)$$

由式(16)得

$$C_j^*(t)=a_j(t)-\sum_{k=1}^{m-j}\frac{l_{k+j,j}(t)C_{k+j}^*(t)}{l_{jj}(t)},\qquad (j=m,\ \cdots,\ 1) \tag{17}$$

因为 $\sum\limits_{j=1}^{m}C_j(t)=1$，故

$$\lambda=1\Big/\sum_{j=1}^{m}C_j^*(t)$$

从而得

$$C_j(t)=C_j^*(t)\Big/\sum_{j=1}^{m}C_j^*(t),\qquad (j=1,\ 2,\ \cdots,\ m) \tag{18}$$

综合式(14)、(15)、(16)、(17)正是求集成系数的全部递推公式。

由上可以看出，式(10)、(11)、(12)、(13)和式(14)、(15)、(17)、(18)有相同的计算格式，可在计算机中一起实现其计算过程，从而计算出 $(X^{\tau(l)}X^{(l)})^{-1}(l=\overline{1,m})$及集成系数

($l=1, \overline{m}$, $t=1, 2, \cdots$)，这样就能减少运算次数，节省了工作单元，给计算提供了方便和可行方法。

四、预报实例及效果比较

实例：根据福建省晋江地区气象台 1959—1984 共24年的资料，预报福建省惠安县 1985 年 8 月份降雨量。

预报对象用 y 表示，利用相关系数检验法选用40个因子，它们与 y 的相关系数见表 1。先后在二种不同信度下用逐步回归法分别建立二个预报方程。即

先给定 $\alpha=0.01$ 下，用逐步回归法建立第 1 个回归预报方程

$$\hat{y}^{(1)}=337.72284-1.63485x_{79}-16.85645x_{452}$$

$$\triangleq 337.72284-1.63485x_1^{(1)}-16.85645x_2^{(1)}$$

再给信度 $\alpha=0.025$ 下，用逐步回归法建立第 2 个回归预报方程

$$\hat{y}^{(2)}=-2024.48584-5.35665x_{390}+11.66806x_{433}+10.90646x_{580}$$

$$\triangleq -2024.48584-5.35665x_1^{(2)}+11.66806x_2^{(2)}+10.90646x_3^{(2)}$$

然后利用式(6)或式(14)、(15)、(17)、(18)建立最小方差集成预报方程

$$\hat{y}_t=C_1(t)\hat{y}_t^{(1)}+C_2(t)\hat{y}_t^{(2)},\qquad t=1, 2, \cdots$$

其中 $C_i(t)(i=1,2)$ 按式(6)或式(14)、(15)、(17)、(18)确定，具体值见表 2。并且有 $E(y_t-\hat{y}_t)^2<E(y_t-\hat{y}_t^{(i)})^2$ ($i=1, 2, t=1, 2, \cdots, 26$)。这说明最小方差集成方程比单个逐步回归方程好，集成预报 y 的1985年的值114.1613mm，而实际上1985年 y 之值为116mm，预报误差仅1.8387mm，回报历史值准确率（按回报误差小于 y 的15%，即 57.65mm 为衡量标准）也有所提高，与 $\hat{y}^{(1)}$ 和 $\hat{y}^{(2)}$ 的具体比较见表2、3。因此，无论从集成预报方程在各时刻的预报误差方差、预报效果或回报率来看，本文方法都比单个回归方程好。

表 1

因　子	x_{22}	x_{55}	x_{83}	x_{155}	x_{204}	x_{231}	x_{247}	x_{265}	x_{273}	x_{275}
相关系数	0.363	0.362	0.536	0.379	0.374	−0.348	0.351	−0.447	0.455	0.435
因　子	x_{279}	x_{287}	x_{307}	x_{327}	x_{337}	x_{362}	x_{372}	x_{376}	x_{377}	x_{390}
相关系数	−0.388	0.451	−0.392	0.347	0.418	−0.408	−416	−0.499	0.357	−0.544
因　子	x_{402}	x_{430}	x_{433}	x_{442}	x_{443}	x_{448}	x_{452}	x_{476}	x_{477}	x_{490}
相关系数	−0.394	−0.383	0.600	0.347	0.354	−0.360	−0.673	−0.581	−0.343	0.388
因　子	x_{529}	x_{550}	x_{551}	x_{567}	x_{580}	x_{583}	x_{584}	x_{602}	x_{625}	x_{628}
相关系数	−0.382	0.429	0.354	−0.353	0.398	−0.381	0.465	−0.385	0.424	0.431

表2

年份	y 实测值	$\hat{y}^{(1)}$ 回报值	$\hat{y}^{(2)}$ 回报值	$\hat{y}$ 回报值	$y-\hat{y}^{(1)}$	评定$\hat{y}^{(1)}$	$y-\hat{y}^{(2)}$	评定$\hat{y}^{(2)}$	$y-\hat{y}$	评定$\hat{y}$
1960	308.6	287.1535	340.5223	336.8980	−21.4465	✓	31.9223	✓	28.298	✓
61	199.7	73.9226	179.8261	147.3641	−125.7774	×	−19.8739	✓	−52.3359	✓
62	158.0	68.0197	112.8751	148.1706	−90.9803	×	−46.1249	✓	−10.8295	✓
63	15.6	−12.5434	24.6688	23.1840	−2.81434	✓	9.0688	✓	7.584	✓
64	45.9	116.1367	90.8691	88.4138	70.2367	×	44.9691	✓	42.5138	✓
65	155.4	164.1428	131.6196	151.1615	8.7428	✓	−23.7804	✓	−4.2385	✓
66	118.6	194.7320	96.2291	135.1722	76.1320	×	−22.3709	✓	16.5722	✓
67	50.6	19.4646	21.9976	22.9021	−31.1354	✓	−28.6024	✓	−27.6979	✓
68	155.1	119.1145	167.0101	166.334	−35.9855	✓	11.9104	✓	11.2340	✓
69	73	118.0986	90.6904	90.6298	45.0986	✓	17.6904	✓	17.6298	✓
70	52.3	117.2811	83.4099	108.9691	64.9811	×	31.1099	✓	56.6691	✓
71	11.6	41.8621	61.5969	44.0120	30.2621	✓	49.9969	✓	32.4120	✓
72	395.9	270.2971	277.3936	280.4483	−125.6029	×	−118.5064	×	−115.4517	×
73	77.6	134.1551	109.4314	97.1126	56.5551	✓	31.8314	✓	19.5126	✓
74	114.4	128.2871	130.8436	130.3618	13.8871	✓	16.4436	✓	15.9618	✓
75	72.7	55.3904	111.7053	93.2328	−17.3096	✓	39.0053	✓	20.5329	✓
76	212.2	234.9493	210.0309	218.1710	22.7493	✓	−2.1691	✓	5.9710	✓
77	145.8	131.2124	165.6547	117.6470	−14.5876	✓	19.8547	✓	−28.1530	✓
78	151.3	135.4454	109.2383	133.6413	−15.8546	✓	−42.0617	✓	−17.6587	✓
79	69.4	97.5344	71.3298	48.4836	28.1344	✓	1.9298	✓	−20.9164	✓
80	192.5	152.3019	190.5140	190.5140	−40.1981	✓	−1.9860	✓	−1.9850	✓
81	49.4	88.6888	65.8165	75.5198	39.2881	✓	16.4165	✓	26.1198	✓
82	55.2	116.3002	72.9040	52.6740	61.1002	×	17.7040	✓	−2.526	✓
83	89.7	135.2820	167.7717	170.1950	45.5820	✓	78.0717	×	80.4954	×
84	115.5	99.7707	3.0490	98.5704	−15.7293	✓	−112.4510	×	−16.9296	✓
预报 1985	116 (实际值)	109.7707 (预报值)	125.4760 (预报值)	114.1613 (预报值)	−6.5137 (误差)	回报率 72%	9.476 (误差)	回报率 88%	−1.8387 (误差)	回报率 92%

表3

序号	年份	集成系数 $C_1(t)$	集成系数 $C_2(t)$	均方预测误差 $E(y_t-\hat{y}_t^{(1)})^2$	均方预测误差 $E(y_t-\hat{y}_t^{(2)})^2$	集成后均方预测误差 $E(y_t-\hat{y}_t)^2$
1	1960	0.06791	0.93209	$0.7264\sigma^2$	$0.2174\sigma^2$	$0.2146\sigma^2$
2	61	0.36635	0.63365	$0.8722\sigma^2$	$0.822\sigma^2$	$0.7814\sigma^2$
3	62	−0.78687	1.78687	$0.8675\sigma^2$	$0.4243\sigma^2$	$0.3176\sigma^2$
4	63	0.0399	0.9601	$0.8160\sigma^2$	$0.5944\sigma^2$	$0.5940\sigma^2$
5	64	−0.09717	1.09717	$0.9341\sigma^2$	$0.8751\sigma^2$	$0.874\sigma^2$
6	65	0.60086	0.39914	$0.8942\sigma^2$	$0.8989\sigma^2$	$0.8905\sigma^2$
7	66	0.39535	0.60465	$0.8719\sigma^2$	$0.8674\sigma^2$	$0.8640\sigma^2$
8	67	−0.35714	1.35714	$0.8765\sigma^2$	$0.8501\sigma^2$	$0.8481\sigma^2$
9	68	0.01307	0.98663	$0.9579\sigma^2$	$0.8536\sigma^2$	$0.8530\sigma^2$
10	69	−0.00221	1.00221	$0.9318\sigma^2$	$0.7004\sigma^2$	$0.7002\sigma^2$
11	70	0.75460	0.2454	$0.9362\sigma^2$	$0.9398\sigma^2$	$0.9317\sigma^2$
12	71	0.89106	0.10894	$0.8820\sigma^2$	$0.9072\sigma^2$	$0.8720\sigma^2$
13	72	−0.43045	1.43045	$0.7758\sigma^2$	$0.4277\sigma^2$	$0.3930\sigma^2$
14	73	−0.49826	1.49826	$0.9551\sigma^2$	$0.5544\sigma^2$	$0.5043\sigma^2$
15	74	0.18845	0.81155	$0.9540\sigma^2$	$0.913\sigma^2$	$0.91066\sigma^2$
16	75	0.32803	0.67198	$0.7830\sigma^2$	$0.2516\sigma^2$	$0.0854\sigma^2$
17	76	0.32667	0.67333	$0.8571\sigma^2$	$0.557\sigma^2$	$0.4646\sigma^2$
18	77	1.39386	−0.39386	$0.4273\sigma^2$	$0.82031\sigma^2$	$0.3925\sigma^2$
19	78	0.93116	0.06884	$0.9159\sigma^2$	$0.9397\sigma^2$	$0.9157\sigma^2$
20	79	1.87184	−0.87184	$0.9488\sigma^2$	$0.5412\sigma^2$	$0.4283\sigma^2$
21	80	0.00063	0.99937	$0.9427\sigma^2$	$0.4203\sigma^2$	$0.4202\sigma^2$
22	81	0.42424	0.57576	$0.9515\sigma^2$	$0.9391\sigma^2$	$0.9244\sigma^2$
23	82	−0.46617	1.46617	$0.9459\sigma^2$	$0.5032\sigma^2$	$0.35721\sigma^2$
24	83	−0.0746	1.0746	$0.9400\sigma^2$	$0.9286\sigma^2$	$0.9285\sigma^2$
25	84	0.98759	0.01241	$0.3654\sigma^2$	$0.9684\sigma^2$	$0.3652\sigma^2$
26	预报 1985	0.72044	0.27956	$1.0590\sigma^2$	$1.1902\sigma^2$	$0.8511\sigma^2$

参 考 文 献

〔1〕国家测绘总局测绘研究所译，近代最小二乘法（译文集），测绘出版社，（1978）.
〔2〕中国科学院数学研究所数理统计组编，回归分析法，科学出版社，（1975）.
〔3〕吴绍敏、陈治典，关于分布自由的逐步回归集成预报法，华侨大学学报(自然科学版),7,3（1986）.
〔4〕安鸿志等，时间序列的分析与应用，科学出版社，（1983）.

A Composite Stepwise Regression Prediction Method with Minimum Variance

Chen Jianwei

Abstract

This paper presents a composite stepwise regression prediction method which is more precise and reliable in prediction than the single stepwise regression equation.

Under different confidence level, several stepwise regression equations are selected. In a sense of minimum mean square, a variable optimal composite prediction equation is established. The composite coefficient established is a variable coefficient in connection with predicted values at different time. It reflects timely the variation of single stepwise regression equations with time and thus makes full use of their predicted informations.

A series of recurrence formulae in the whole computation are derived for use in programming.