新息预报程序设计中分配存贮单元的一种方法

陈 小 明

（数学系七八级）

在生产斗争和科学实验中，人们往往需要对某一随机序列进行分析，以达到对未来时刻的值作出预测。一般步骤是：(1)对该随机时间序列的观测值$z_1,\cdots\cdots,z_n$作相关分析，找出一般规律性，建立数学模型。(2)选择合适的预报方法。随机序列适时预报——新息预报是其中较好的方法之一。

新息预报的计算中，涉及到一个变带宽稀疏矩阵的存贮问题；若按原公式安排存贮 用PDP—11/34计算机，只能处理样本长度为一百左右的问题。本文是在陈兴钩老师指导下，给出了一种较简便的存贮方法，使我们用同样型号的计算机，可处理样本长度到500左右。这次毕业实践，我们用这种方法，处理了样本长度达300多的几组数据，已在机器上实现。面先简单介绍新息预报的计算公式，详细论证见〔1〕

设一宽平稳随机时间序列$\{z_t\}$属于ARMA（p，q）模型：

$$z_t-\varphi_1 z_{t-1}\cdots\phi_p z_{t-p}=a_t-Q_1 a_{t-1}\cdots Q_q a_{t-q}$$

$z_1,\cdots\cdots z_N$为其一个样本序列。$\hat{Z}_{N+l}$为根据$z_1,\cdots,z_N$作出的z_{N+l}的预报值。令：

$$Q=max(p、q)$$

$$y_j=\begin{cases} z_j & j\leqslant Q \\ z_j-\sum_{i=1}^{p} z_j-i\phi_i & j>Q \end{cases}$$

则有如下递推预报式：

$$\hat{Z}_{N+l}=\sum_{j=1}^{p}\phi_j\hat{Z}_{N+(l-j)}+\sum_{j=N+l-q}^{N}C_{N+l,j}\,\epsilon_j \qquad (*)$$

其中

$$\varepsilon_k=y_k-\sum_{j=n}^{k-1}C_{kj}\,\varepsilon_j$$

$$C_{kj}=\left[r_{kj}(y)-\sum_{i=n}^{j-1}C_{ji}C_{ki}r_{ii}(\varepsilon)\right]\left[r_{jj}(\varepsilon)\right]^{-1} \qquad j=n,\cdots,k-1$$

$$r_{kk}(\varepsilon)=r_{kk}(y)-\sum_{i=1}^{k-1}C_{kj}^{2}\cdot r_{jj}(\varepsilon)$$

$$r_{kj}(y)=-\frac{1}{k}\sum_{i=1}^{j}y_i\cdot y_{k-j+i}\qquad j=1,\cdots,k$$

又其中当 $k\leqslant Q$ 时 $n=1$，当 $k>Q$ 时 $n=k-q$。以上公式中当求和上限小于下限时规定和值为零，当 $l-j\leqslant 0$ 时 $\hat{z}_{N+(l-i)}=z_{N+(l-j)}$。

从预报公式（＊）不难看出，当计算 $C_{k,j}$ 从 $k=1$ 到预报起点 $k=N$ 时，并不能马上开始预报；还要进一步计算 $C_{N+1,j}$，l 为预报步长。由于样本长度为N，所以 y_{N+l} 等已不能计算出来了，从而原来的关于求 $\gamma_{k,j}(y)$ 的递推公式也不能用了。但由于我们假设 $\{z_t\}$ 是宽平稳的，而 $z_1,\cdots\cdots,z_N$ 是其样本序列，故当 $K>N$，（$N>2Q$），有

$$\gamma_{k+1,j}(y)=\gamma_{k,j-1}(y)$$

所以 $C_{N+l,j}$ 可由如下公式计算

$$C_{N+l,j}=\left[\gamma_{N,j-l}(y)-\sum_{i=N+l-q}^{N}C_{ji}C_{N+l,i}\gamma_{ji}(\varepsilon)\right]\left[\gamma_{jj}(\varepsilon)\right]^{-1}$$

$$j=N+l-q,\cdots,N$$

在整个计算过程中，我们要计算 $\gamma_{kj}(y)$，$\gamma_{ii}(\epsilon)$ 和 C_{kh} 等，分析一下公式，不难发现 $\gamma_{kj}(y)$ 和 $\gamma_{ii}(\epsilon)$ 可以用一维数组来贮存所要计算的元素。对于 C_{kj} 如按原公式用二维数组安排存贮单元，程序设计比较简单，但这样太浪费存贮空间，因此当样本序列较长而又是在内存参量不大的小型机上执行程序时，就会出现内存不够用的情况，例如：PDP—11/34计算机，一般每一终端使用的存储空间为10000。这样当样本长度到达100时，用PDP—11/34机就无法计算了。为了解决这一问题，我们进一步深入分析计算 C_{kj} 的公式，可以看到，C_{kj} 需要算的元素只需占矩阵（1）所标出的位置；它是一个变带宽稀疏矩阵。每一行至多只要算Q个元素。

$$\begin{bmatrix}
C_{2,1} & & & & & & & \\
C_{3,1} & C_{3,2} & & & & & & \\
\vdots & \vdots & \ddots & & & & & \\
C_{Q,1} & C_{Q,2} & \cdots & C_{Q,Q-1} & & & & \\
 & C_{Q+1,Q+1-\beta} & \cdots & C_{Q+1,Q} & & & & \\
 & & \ddots & & \ddots & & & \\
 & & & C_{N,N-\beta} & \cdots & C_{N,N-1} & & \\
 & & & & C_{N+1,N+1-\beta} & \cdots & C_{N+1,N} & \\
 & & & & & \ddots & \vdots & \\
 & & & & & & C_{N+\ell,N} &
\end{bmatrix}_{(N+\ell)\times N}$$

（1）

其运算的规律是以三角阵的形式向下推移，因此，也可用一维数组来存放。但这样做的规律性不直观。我们采用矩阵(2)的存放方式来存放(1)所要计算的元素。

这样，所有涉及到 C_{kj} 的公式都要改变，但规律清楚，新的公式如下：

$$\varepsilon_k=y_k-\sum_{j=m}^{k_1}C_{kj}\epsilon_{j+m}$$

$$C_{kj}=\left[\gamma_{k,j+m}(y)-\sum_{i=n}^{j-1}C_{j+m,i+m}C_{k,i}\gamma_{i+m,i+m}(\epsilon)\right]\left[\gamma_{j+m,j+m}(\varepsilon)\right]^{-1}$$

$$j=n,\cdots k_1$$

$$\gamma_{kk}(\varepsilon)=\gamma_{kk}(y)-\sum_{j=n}^{k_1}C_{kj}^2\cdot\gamma_{j+mj+m}(\epsilon)$$

这里分几种情况：

i）当$k\leqslant Q$时：$n=1$，$m=0$　$k_1=k-1$　$m_1=0$

ii）当$Q<k\leqslant N$时：$n=Q-q+1$　$m=k-Q-1$　$k_1=Q$

$$m_1=\begin{cases}k-j-m & j+m>Q+1\\ m & j+m\leqslant Q+1\end{cases}$$

iii）当$k>N$，求Ckj的公式还要变为：

$$C_{N+l,j}=\left[\gamma_{Nj+m-l}(y)-\sum_{l=n}^{j-1}C_{j+m,i+m}C_{N+l,i}\gamma_{i+m,i+m}(\epsilon)\right]\left[\gamma_{j+m,j+n}(\epsilon)\right]^{-1}$$

$$j=n,\cdots k_1$$

其中　$n=Q-q+1$　$m=N+l-Q-1$　$m_1=Q-j+1$　$k_1=Q-1+1$

预报公式如：

$$\hat{z}_{N+l}=\sum_{j=N+l-q}^{N}C_{N+l,i}\varepsilon_j+\sum_{j=1}^{P}\phi_j\hat{z}_{N+(l-j)}$$

其中　$i_1=j-N-l+Q+1$

矩阵(1)是(N+1)×N维矩阵，矩阵(2)是(N+1)×Q矩阵，而Q与N相比一般说来要小得多；特别是N较大时，有Q<<N，所以这样处理就可节省许多存储单元。同样是PDP—11/34机，用这种方法没计程序，样本个数为400～500时也能算。我们这些想法已在PDP—11/34机上实现，计算了样本长序为312的季节性模型的气象预报问题，取得了有益的结果。

$$\begin{pmatrix}
C_{2,1} & & & & \\
C_{3,1} & C_{3,2} & & & \\
\vdots & \vdots & \ddots & & \\
C_{Q,1} & C_{Q,2} & \cdots & C_{Q,Q-1} & \\
 & C_{Q+1,Q+1-\beta} & \cdots & C_{Q+1,Q} & \\
 & \vdots & & \vdots & \\
 & C_{N,N-\beta} & \cdots & C_{N,N-1} & \\
 & C_{N+1,N+1-\beta} & \cdots & C_{N+1,N} & \\
 & \vdots & & \vdots & \\
 & C_{N+l,N+l-\beta} & \cdots & C_{N+l,N} &
\end{pmatrix}_{(N+l)\times Q}$$

（2）

参考文献

〔1〕 陈兴钧，线性随机序列分析（一）（讲义），1981。

〔2〕 中国科学院应用数学研究所，线性随机序列分析（讲义），1976。