第6卷 第4期　　大 气 科 学　　Vol. 6, No. 4
1982年12月　　SCIENTIA ATMOSPHERICA SINICA　　Dec., 1982

自回归模式的拟合

么枕生
(南京大学气象系)

提　要

本文首先略述用自回归模式去拟合平稳时间序列的各种方法;其次,证明自回归模式的第末个参数就是偏自相关系数,所以用迭代法逐步估计出自回归模式的参数后,就可以应用 t 检验或 F 检验选定模式的阶,使自回归模式拟合的计算量大大简化. 文中举出日降水量的气候预报实例,以便说明上述方法. 最后,讨论了 t 检验使用中的问题.

一、引　言

线性随机模式可划分为自回归模式、移动平均(滑动平均)模式与自回归-移动平均混合模式[1]. 所谓自回归模式就是平稳时间序列值(离差) y_t 用一些过去值 y_{t-1}, y_{t-2}, $\cdots$ 叠加噪声 ε_t 所表示的式子. 所谓移动平均模式就是平稳时间序列值应用一系列随机噪声 ε_t, ε_{t-1}, ε_{t-2}, $\cdots$ 所表示的式子. 我们可以证明自回归模式和移动平均模式可以互相转换,二者是同等的线性随机模式. 所谓混合模式就是把自回归模式和移动平均模式混合在一起来表示平稳时间序列值 y_t. 在这三种模式中,自回归模式是统计预报和统计气候学中最重要的统计模式,因为在较高阶的自回归模式中剩余项已几成为随机噪声,而随机噪声并无相关(不能预报),平均数为零,且自回归模式可以写成差分形式,便于计算预报值,自回归参数代表周期圈,具有气候学意义.

周期性的古典概念就是认为周期性仅仅是由于谐波形成的. Yule[2] 发展了周期性的古典概念,提出自回归过程可以代表谐波序列叠加随机噪声的过程. 所谓自回归模式就是用数学符号代表的自回归过程. 在用自回归模式拟合平稳时间序列的方法中,以 Yule-Walker 方程最为有名. Jones[3]曾根据 Wiener-Kolmogoroff 线性预报理论应用周期图估计谱去估计自回归模式的参数(详见文献[1]). Box 与 Jenkins[4] 提出了系统方法去估计自回归模式的参数. 他们首先应用 Yule-Walker 方程计算出自回归参数的初期估计值,然后还要迭代订正,以便求得更好的拟合(详见文献[1]). 不过,在国际上估计自回归参数时,常用逐步迭代法进行计算,而模式阶数的选定则采用赤池[5] 所提出的方法、本文提出在用迭代法拟合自回归模式时,可逐步用 t 检验或 F 检验选定自回归模式的阶数.

二、自回归模式的建立

设 y_t 代表平稳随机过程自平均数的离差,拟合平稳时间序列(随机过程)的自回归

1981年5月20日收到修改稿.

模式可写如

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \cdots + \alpha_p y_{t-p} + \varepsilon_t \tag{1}$$

模式 (1) 可认为是过程的现在值 y_t 对于其过去值 y_{t-1}, y_{t-2}, $\cdots$, y_{t-p} 的回归. ε_t 代表外界扰动或白噪声. 这个具有参数 α_1, α_2, $\cdots$, α_p 的模式就称为 p 阶自回归模式.

自回归参数可求解下列方程组去估计:

$$\begin{aligned} \rho_1 &= \alpha_1 + \alpha_2\rho_1 + \cdots + \alpha_p\rho_{p-1} \\ \rho_2 &= \alpha_1\rho_1 + \alpha_2 + \cdots + \alpha_p\rho_{p-2} \\ &\cdots\cdots\cdots \\ \rho_p &= \alpha_1\rho_{p-1} + \alpha_2\rho_{p-2} + \cdots + \alpha_p \end{aligned} \tag{2}$$

这就是所谓 Yule-Walker 方程. 这是一个用自相关系数 ρ_1, ρ_2, $\cdots$, ρ_p 表示自回归参数 α_1, α_2, $\cdots$, α_p 相互关系的线性方程组. 在式 (2) 中用估计的自相关系数 r_k 代 ρ_k($k=1$, 2, $\cdots$, p) 时,就求得自回归模式中参数的估计值.

根据式 (2), 用矩阵计算参数的估计值必须求逆阵, 所以一般都用逐步迭代法[6]进行计算 (详见文献 [1]):

$$\alpha_k^{(k)} = \frac{\rho_k - [\alpha_1^{(k-1)}\rho_{k-1} + \alpha_2^{(k-1)}\rho_{k-2} + \cdots + \alpha_{k-1}^{(k-1)}\rho_1]}{1 - [\alpha_1^{(k-1)}\rho_1 + \alpha_2^{(k-1)}\rho_2 + \cdots + \alpha_{k-1}^{(k-1)}\rho_{k-1}]} \tag{3}$$

$$\alpha_j^{(k)} = \alpha_j^{(k-1)} - \alpha_k^{(k)}\alpha_{k-j}^{(k-1)}$$

$$(j = 1, 2, \cdots, k-1;\ k = 1, 2, \cdots, p)$$

式中上标 (k) 代表自回归模式的阶,下标 k 代表其第 k 个参数. 例如, $\alpha_k^{(k)}$ 就是 k 阶自回归模式的第末个参数.

对于一个 k 阶自回归模式,剩余平方和的递推公式为

$$S_k = S_{k-1}\{1 - [\alpha_k^{(k)}]^2\} \tag{4}$$

均方误差为

$$\hat{\sigma}_k^2 = \frac{S_k}{(n-k-1)} \tag{5}$$

Yule-Walker 方程可以根据自回归过程的自相关函数求得,但也可以根据最小二乘方原理推导. 如命 $E[\varepsilon_t^2]$ 成为最小,则用以估计参数 α_1, α_2, $\cdots$, α_p 的正规方程可写如

$$E[y_{t-j}\varepsilon_t] = 0 \quad (j = 1, 2, \cdots, p) \tag{6}$$

此处 $E[\ ]$ 代表求数学期望的符号. 如逐步命 j 为 $1, 2\cdots p$, Yule-Walker 方程就可以求得. 因此,自回归模式的参数可以利用记录序列中的偏自相关去估计, 其结果和用式 (2) 求得的完全一样.

现在分别用 y, y_1, y_2, $\cdots$, y_p 代 y_t, y_{t-1}, y_{t-2}, $\cdots$, y_{t-p}, 式(1)就可以写成下列形式:

$$y = \alpha_1 y_1 + \alpha_2 y_2 + \cdots + \alpha_p y_p + \varepsilon \tag{7}$$

其中下标 $1, 2, \cdots, p$ 代表过去记录的号数. 于是,自回归模式中的参数就可以根据下列关系去计算:

$$\alpha_1^{(k)} = \rho_{1y\cdot 23\cdots k}\frac{\sigma_{y\cdot 23\cdots k}}{\sigma_{1\cdot 23\cdots k}}$$

$$\alpha_2^{(k)} = \rho_{2y\cdot 13\cdots k}\frac{\sigma_{y\cdot 13\cdots k}}{\sigma_{2\cdot 13\cdots k}}$$

$$\cdots\cdots\cdots \tag{8}$$

$$\alpha_k^{(k)} = \rho_{ky\cdot 12\cdots k-1}\frac{\sigma_{y\cdot 12\cdots k-1}}{\sigma_{k\cdot 12\cdots k-1}}$$

$$(k = 1,\ 2,\ \cdots,\ p)$$

此处 $\rho_{ky\cdot 12\cdots k-1}$ 是 y_k 与 y 对于 y_1, y_2, $\cdots$, y_{k-1} 的偏自相关系数 (partial autocorrelation coefficient), $\sigma^2_{y\cdot 12\cdots k-1}$ 代表剩余方差.

么枕生[7] 提出用偏相关系数的 t 检验建立多重回归的方法 (所谓偏相关筛选), 其结果和用逐步回归是同等的,但计算量大大减少. 同样,自回归模式的参数可用式 (8) 去估计,且模式的阶数可用偏自相关系数的 t 检验去选定:

$$t = \frac{r_{ky\cdot 12\cdots k-1}}{\sqrt{1 - r^2_{ky\cdot 12\cdots k-1}}}\sqrt{\nu} \tag{9}$$

此处自由度 $\nu = n - 1 - k$, n 为样本大小, k 为自回归模式的阶数; $r_{ky\cdot 12\cdots k-1}$ 代表偏自相关系数的估计值.

如果自回归模式的阶数高,则应用式 (8) 逐步计算仍较麻烦. 但是,很有趣的是作者对自回归模式证明了下列关系:

$$\begin{aligned}&\sigma^2_{y\cdot 12\cdots k-1} = \sigma^2_{1\cdot 23\cdots k} = \sigma^2_{k\cdot 12\cdots k-1}\\&\sigma^2_{y\cdot 23\cdots k} \neq \sigma^2_{y\cdot 12\cdots k-1} = \sigma^2_{1\cdot 23\cdots k}\\&\sigma^2_{y\cdot 13\cdots k} \neq \sigma^2_{1\cdot y2\cdots k-1} = \sigma^2_{2\cdot 13\cdots k}\\&\cdots\cdots\cdots\cdots\end{aligned} \tag{10}$$

因此,根据式 (10), 方程式 (8) 的第末个方程可以简单表示为

$$\alpha_k^{(k)} = \rho_{ky\cdot 12\cdots k-1} \quad (k = 1, 2, \cdots, p) \tag{11}$$

这个关系式 (11) 说明: k 阶自回归模式的第末个参数 $\alpha_k^{(k)}$ 就是一个偏自相关系数.

y 对于 y_1, y_2, $\cdots$, y_k 的剩余方差为

$$\begin{aligned}\sigma^2_{y\cdot 12\cdots k} = \sigma^2_y(1 - \rho^2_{1y})(1 - \rho^2_{2y\cdot 1})(1 - \rho^2_{3y\cdot 12})\\ \cdots(1 - \rho^2_{ky\cdot 12\cdots k-1})\end{aligned} \tag{12}$$

此外, 式 (4) 可用递推方法从简单自回归模式的剩余方差去证明. 现在由式 (4) 更用递退可求得下列结果:

$$\begin{aligned}\sigma^2_{y\cdot 12\cdots k} &= \sigma^2_{y\cdot 12\cdots k-1}\{1 - [\alpha_k^{(k)}]^2\}\\&= \sigma^2_{y\cdot 12\cdots k-2}\{1 - [\alpha_{k-1}^{(k-1)}]^2\}\{1 - [\alpha_k^{(k)}]^2\}\\&= \sigma^2_{y\cdot 12\cdots k-3}\{1 - [\alpha_{k-2}^{(k-2)}]^2\}\{1 - [\alpha_{k-1}^{(k-1)}]^2\}\{1 - [\alpha_k^{(k)}]^2\}\\&\cdots\cdots\cdots\\&= \sigma^2_y\{1 - [\alpha_1^{(1)}]^2\}\{1 - [\alpha_2^{(2)}]^2\}\{1 - [\alpha_3^{(3)}]^2\}\cdots\{1 - [\alpha_k^{(k)}]^2\}\end{aligned} \tag{13}$$

因此,我们对比式 (12) 与 (13) 也可证明式 (11) 是正确的, 只是以前的作者们并没有发现或证明这个能使自回归模式拟合手续大大简化且具有统计学意义的事实.

现在更举例证明如下: 当 $k = 1$, 则

$$\alpha_1^{(1)} = \rho_{1y} = \rho_1;$$

当 $k = 2$, 则

$$\alpha_2^{(2)} = \rho_{2y\cdot 1} = \frac{\rho_{2y} - \rho_{12}\rho_{1y}}{\sqrt{(1-\rho_{12}^2)(1-\rho_{1y}^2)}} = \frac{\rho_2 - \alpha_1^{(1)}\rho_1}{1-\alpha_1^{(1)}\rho_1};$$

当 $k=3$，则

$$\begin{aligned}\alpha_3^{(3)} = \rho_{3y\cdot 12} &= \frac{\rho_{3y\cdot 1} - \rho_{23\cdot 1}\rho_{2y\cdot 1}}{\sqrt{(1-\rho_{23\cdot 1}^2)(1-\rho_{2y\cdot 1}^2)}}\\ &= \frac{\rho_1(\rho_1^2-\rho_2)+\rho_2(\rho_1\rho_2-\rho_1)+\rho_3(1-\rho_1^2)}{(1-\rho_1^2)-\rho_1(\rho_1-\rho_1\rho_2)+\rho_2(\rho_1^2-\rho_2)}\\ &= \frac{\rho_3-[\alpha_1^{(2)}\rho_2+\alpha_2^{(2)}\rho_1]}{1-[\alpha_1^{(2)}\rho_1+\alpha_2^{(2)}\rho_2]}\end{aligned}$$

因为式 (11) 两边都要迭代或递推计算，所以当 k 为任何值时都可以这种方式证明式 (11) 是正确的.

式 (11) 的唯一重要意义就是可以把迭代公式 (3) 计算出的 $\alpha_k^{(k)}$ 估计值，当做自相关系数估计值，更根据式 (9) 用 t 检验去选定自回归模式的阶:

$$t = \frac{\hat{\alpha}_k^{(k)}}{\sqrt{1-[\hat{\alpha}_k^{(k)}]^2}}\sqrt{\nu} \quad (k=1,2,\cdots,p) \tag{14}$$

此处自由度 $\nu = n-1-k$，n 为计算后延 k 偏自相关系数时所用的记录对数. 如果在迭代手续的每一步中，第末个参数 $\alpha_k^{(k)}$ 已经估计出，自回归模式的阶数根据预定的显著性水平就容易选定.

我们知道

$$F = \frac{r_{y\cdot 12\cdots k}^2 - r_{y\cdot 12\cdots k-1}^2}{1-r_{y\cdot 12\cdots k}^2}(n-k-1) \tag{15}$$

遵守 F 分布，具有自由度 1 与 $n-k-1$，其中 $r_{y\cdot 12\cdots k}$ 代表多重相关系数. 因为

$$\begin{aligned}1-r_{y\cdot 12\cdots k}^2 &= (1-r_{y\cdot 12\cdots k-1}^2)(1-r_{ky\cdot 12\cdots k-1}^2)\\ &= (1-r_{y\cdot 12\cdots k-1}^2)\{1-[\hat{\alpha}_k^{(k)}]^2\},\end{aligned} \tag{16}$$

而

$$\frac{1-r_{y\cdot 12\cdots k}^2}{1-r_{y\cdot 12\cdots k-1}^2} = 1-[\hat{\alpha}_k^{(k)}]^2, \tag{17}$$

所以

$$F = \frac{[\hat{\alpha}_k^{(k)}]^2}{1-[\hat{\alpha}_k^{(k)}]^2}(n-k-1) \tag{18}$$

因此，自回归模式的阶也可以很容易用 F 检验去确定. 同时，我们根据式 (14) 与 (18) 也证明了 F 分布与 t 分布的有名关系:

$$F(1, n-k-1) = t^2(n-k-1) \tag{19}$$

如何选定自回归模式的阶是国际上目前还没有很好完全解决的问题. Box 与 Jenkins 提出用自相关函数与偏自相关函数做为选定自回归模式阶数的标准. 所谓偏自相关函数 (partial autocorrelation function) α_{kk} 就是在 k 阶自回归模式中，用 Yule-Walker 方程求解的第末个参数 $\alpha_k^{(k)}$. 因为 α_{kk} 可以用 $\rho_1, \rho_2, \cdots, \rho_k$ 的函数表示，所以他们命名 α_{kk} 为偏自相关函数. 例如，

$$\alpha_{22}=\frac{\begin{vmatrix}1 & \rho_1\\ \rho_1 & \rho_2\end{vmatrix}}{\begin{vmatrix}1 & \rho_1\\ \rho_1 & 1\end{vmatrix}}=\frac{\rho_2-\rho_1^2}{1-\rho_1^2}$$

实际上,因为 $\rho_1=\alpha_1^{(1)}$, 所以 $\alpha_{22}=\alpha_2^{(2)}$. 根据 Box 与 Jenkins, 如自相关函数(即各种后延的自相关系数)逐渐减小,而偏自相关函数 α_{kk} 当 k 小于或等于 p 时为非零值,而当 k 大于 p 时成为零的话,则自回归模式成为 p 阶. 但是,在实际应用中,α_{kk} 值不但计算麻烦,标准含混,不易掌握,而且这个标准既嫌低,也缺乏统计学意义.

有些作者用系数矩阵为正定的做为标准去确定自回归模式的阶数[6]. 不过,系数矩阵是正定的,这仅仅保证了自回归模式中各个参数不完全为零而已. 如阶数很大,这个方法也是不但计算麻烦,而且标准嫌低,并无统计学意义.

赤池 (1969) 为了选定自回归模式的阶,提出在每一步应计算最后预报误差 (Final Prediction Error):

$$FPE_k=\hat{\sigma}_k^2[1+(k+1)/n] \quad (k=1,2,\cdots,p) \tag{20}$$

此处 $\hat{\sigma}_k^2$ 应根据式 (5) 计算. 不过,赤池的方法仅仅是相对的标准,并没有统计学意义,而且计算也嫌麻烦.

如应用独立的记录,根据自回归系数的估计量计算预报量时,则 FPE_k 就是一步预报方差的估计量. Jones[8] 为了估计平均值的方差,用高阶自回归模式拟合时间序列时,就应用赤池方法选定自回归模式的阶. 他根据经验指出: 为了拟合记录,阶数最高常可计算到 25 或 50,于是就以 FPE_k 成为最低的阶数选做最佳拟合. 不过作者认为: 如选定自回归模式阶数的标准只是拟合优度,盲目追求最佳拟合,不做显著性检验时,则预报效果是会很差的.

我们最好把式 (3) 与式 (11) 结合起来,在每一步迭代计算出第末个自回归参数的估计值以后,就根据 t 检验或 F 检验去选定自回归模式的阶. 此外,在 $y_1, y_2, \cdots, y_p$ 中的任何一个,都可以用 t 检验或 F 检验从模式中去掉,就像在前向筛选手续中所做的一样.

三、实 际 拟 合

为了说明对于自回归模式的拟合方法,下列计算用做上述讨论的实际例子. 上海 (1921—1950) 6 月 30 年降水记录可以认为是平稳时间序列,用这个记录序列所计算的自相关系数样本值为

$$r_1=-0.3390,\ r_2=-0.1369,\ r_3=0.2483,$$
$$r_4=-0.03928,\ r_5=-0.1163,\ r_6=0.07049.$$

第一步,根据式 (3), (14) 与 (18) 求得

$$\hat{\alpha}_1^{(1)}=-0.3390,$$
$$t=\frac{-0.3390}{\sqrt{1-0.3390^2}}\sqrt{29-1-1}=-1.8724,$$
$$-t>t_{.95}(27);$$

且

$$F = 3.5059$$
$$F > F_{.90}(1,27).$$

因此，如用单尾检验，选定 0.10 的显著性水平时，则式 (3) 的计算还应继续.

第二步，根据式 (3)，(14) 与 (18) 求得

$$\hat{a}_1^{(2)} = -0.4354, \quad \hat{a}_2^{(2)} = -0.2845$$

$$t = \frac{-0.2845}{\sqrt{1-0.2845^2}}\sqrt{28-1-2} = -1.4838$$

$$-t > t_{.90}(25);$$

$$F = 2.2017$$

$$F > F_{.80}(1,25).$$

因此，我们根据参数的显著性检验，所拟合的 2 阶自回归模式

$$Y_t = 307.66 - 0.4354Y_{t-1} - 0.2845Y_{t-2} \qquad \text{(I)}$$

可以做为上海 6 月降水的预报方程，其中 Y_t 代表上海当年 6 月降水量，用做预报量；Y_{t-1} 与 Y_{t-2} 代表前一年与前两年 6 月降水量，用做预报因子.

第三步，我们还有必要检查是否一个 3 阶自回归模式可能拟合上海 6 月降水记录. 这个计算结果列如下：

$$\hat{a}_1^{(3)} = -0.4031, \quad \hat{a}_2^{(3)} = -0.2351, \quad \hat{a}_3^{(3)} = 0.1134,$$

$$t = \frac{0.1134}{\sqrt{1-0.1134^2}}\sqrt{23} = 0.5474$$

$$t > t_{.70}(23);$$

且

$$F = 0.2996$$
$$F < F_{.50}(1, 23).$$

这个 3 阶自回归模式的形式为

$$Y_t = 274.81 - 0.4031Y_{t-1} - 0.2351Y_{t-2} + 0.1134Y_{t-3} \qquad \text{(II)}$$

如单尾检验应用 0.10 的显著性水平，这个模式是不能采用的.

根据 Box 与 Jenkins 鉴定模式的技术，甚至还可能对上海 6 月的降水预报，提出 4 阶模式：

$$Y_t = 249.00 - 0.4126Y_{t-1} - 0.2155Y_{t-2} + 0.1471Y_{t-3} + 0.08352Y_{t-4}, \qquad \text{(III)}$$

因为偏自相关函数 α_{44} 初步估计出为 0.08352，而这个数确实与零并不相同. 在这里我们必须指出：模式的阶数愈高，参数的初期估计值和订正值差别就愈小，所以一个高阶自回归模式的参数就不必去订正[1].

为了深入分析，现在把上海 6 月 30 年降水记录 (1921—1950) 和另外 10 年记录 (1951—1960) 分别代入模式 (I)、(II) 与 (III) 中，自原记录计算结果，求得这三个模式的剩余平方和与误差平方和如下：

模 式	I	II	III
剩余平方和	180757.56	178613.18	173606.11
误差平方和	38301.93	39747.46	41134.00

模式 III 的剩余平方和为最小，而这个模式的误差平方和为最大．模式 I 的情况完全相反．因此，这些计算结果说明：模式 III 虽然由于拟合历史记录最好，给出参数的有效估计值，但对于预报因子的未来值并非有效．在实际上，模式 III 最末参数的 t 检验为

$$t=\frac{0.08352}{\sqrt{1-0.08352^2}}\sqrt{21}=0.3841,$$

而 $t<t_{.70}$．根据参数的显著性检验，模式 III 是绝对不能接受的．

如应用赤池的最后误差去选定自回归模式的阶，则根据式(4)、(5)与(20)计算结果如下：

$$FPE_1=7976.62,\quad FPE_2=7840.09,\quad FPE_3=8280.47,\quad FPE_4=8803.14$$

因为模式 I 的最后误差最小，所以应采这个模式去描写上海 6 月的降水时间序列．这个由赤池方法所求得的结论，是和由 t 检验所求得的结论完全相同的，只是赤池方法的计算多少嫌繁，且缺乏统计学的显著性意义罢了．

四、用 t 检验建立自回归模式的问题

根据 Nordø，在建立回归模式时，时间序列内部的自相关性可以减少样本检验统计量的自由度数[9]．换句话说，时间序列的内部相关性，根据式(14)可把选进因子的标准相对提高．因此，在向前选进因子的各步中，不良因子是不会选进的．一旦零假设 $H_0:\alpha_k^k=\rho_{ky\cdot12\cdots k-1}=0$ 在预定的显著性水平可以授受，计算这个偏自相关系数的随机变量就可以看做是独立的．如样大($n>30$)，则根据中心极限定理，此时 t 检验所需求的条件(正态随机独立)是完全符合的，用 t 检验确定自回归模式的阶数必然在统计学意义上是正确的．

现在我们把 Nordø 对于回归所求得的结果搬到自回归中．如自回归序列成为简单马尔科夫过程，则剩余方差的估计自由度可求得近似为

$$\nu=n-\left[\frac{1+R}{1-R}+p\frac{1+\rho_1R}{1-\rho_1R}\right]+\frac{2}{n}\left[\frac{R(1-R^n)}{(1-R)^n}+p\rho_1R\ \frac{1-\rho_1^nR^n}{(1-\rho_1R)^2}\right]\qquad(21)$$

此处 ρ_j 为 y_j 的 1 阶总体自相关系数 ($j=0,1,2\cdots,p$)，R 为 1 阶总体剩余自相关系数．现在我们的问题是如何估计 R．经过麻烦的推导，可求得计算 R 的一般公式．如 $p=2$，则

$$R=\frac{(1+\alpha_1^2+\alpha_2^2)\rho_1+\alpha_1\alpha_2(1+\rho_2)-\alpha_1(1+\rho_2)-\alpha_2(\rho_1+\rho_3)}{1+\alpha_1^2+\alpha_2^2+2\rho_1(\alpha_1\alpha_2-\alpha_1)-2\alpha_2\rho_2}\qquad(22)$$

用建立模式 I 所用的一些资料，根据式(22)可求得

$$\hat{R}=\frac{0.0261987}{0.8134314}=0.03221,$$

于是,根据式 (21) 求得

$$\nu = 28 - 3.02437$$

这说明一旦 $\alpha_3^3 = \rho_{3y\cdot 12}$ 在预定的显著性水平显著等于零时，则模式 I 中剩余方差的自由度数已不受时间序列中自相关性的影响。

我们可以对任何时间序列都应用 t 检验建立自回归模式。 如 α_k^k 逐渐变小 ($k = 1, 2, \cdots, p$)，则前几个系数如 $\alpha_1^1, \alpha_2^2, \cdots$，实际不必检验,根据直观就能决定式(3)是否应当继续计算下去;只有当 α_k^k 接近于零时,才最有必要采用 t 检验，去确定自回归模式的阶数.

参 考 文 献

[1] 么枕生，随机模式，1980年中长期水文气象预报文集，长江流域规划办公室.

[2] Yule, G. U., On a method of investigating periodicities in disturbed series. with special reference to Wolfer's Sunspot numbers, *Trans. Roy. Soc.* (A) 236, 267, 1927.

[3] Jones, R. H., Prediction of multivariate time series, *J. Appl. Meteor.*, 3, 285—289, 1964.

[4] Box, G. P. and G. M. Jenkins, Time Series Analysis, forecasting and control, Holden-Day, San Francisco, 1970.

[5] Akaike, H., Fitting autoregressions for prediction, *Ann. Inst. Statis. Math., Jokyo*, 21, 243—247, 1969.

[6] 王宗皓、李麦村等，天气预报中的概率统计方法，科学出版社，1974.

[7] 么枕生，正交回归与偏相关筛选，南京大学学报，1977年 第一期；一个用偏相关筛选建立多元回归方程的方法，气象科技资料1977年 第5期.

[8] Jones, R. H., Estimating the variance of time averages, *J. Appl. Meteor.*, 14, 159—163, 1975.

[9] Nordø, J., Significance of statistical relations derived from geophysical data, *Tellus*, 18. 39—53, 1966.

THE FITTING OF AUTOREGRESSIVE MODEL

Yao Zhen-sheng

(Department of Meteorology, Nanjing University)

Abstract

The methods for fitting the autoregressive model to the stationary time series are briefly reviewed. It is shown that the last parameter of the autoregressive model used for climatic prediction is just the partial autocorrelation coefficient, and therefore the order of the model can be chosen by means of the test or F test. The fitting of the autoregressive model can thus be performed by simple procedure, if the autoregressive parameters are estimated by the recurrence formula. An illustrating example used to forecast the monthly rainfall has been given. Finally, the test of hypothesis for the partial autocorrelation is examined.