§ 3.4　滑动平均模型的拟合

设 $x_1,\cdots,x_n$ 为来自 $MA(q)$ 模型的样本值：

$$x_t=\varepsilon_t-\beta_1\varepsilon_{t-1}-\beta_2\varepsilon_{t-2}-\cdots-\beta_q\varepsilon_{t-q}\ ,\ t=1,\ 2,\ \cdots,\ n \qquad (3.4.1)$$

其中 $\{\varepsilon_t\}$ 为独立序列，且 $E\varepsilon_t=0$, $E\varepsilon_t^2=\sigma^2$, $E\varepsilon_t^4<+\infty$ 。

一、阶数 q 的估计

若 $MA(q)$ 模型中阶数 q 未知，首先要对其进行估计，估计的方法有自相关函数估计法、AIC 准则估计法等等。

1.　自相关函数分析法

因为一个平稳序列为 $MA(q)$ 模型的充要条件是，其自协方差函数 ρ_k 在 q 以后截尾，所以可以借助这一特征作为估计阶数 q 的依据。在实际中，是利用所知样本值，先对自协方差函数的进行估计，再对阶数 q 进行估计，估计的步骤为：

（1）首先由 $x_1,\ x_2,\ \cdots,\ x_n$ 计算均值 $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$ ；

（2）计算样本自协方差函数

$$\hat{\gamma}_k=\frac{1}{n}\sum_{j=1}^{n-k}(x_j-\bar{x})(x_{k+j}-\bar{x}) \qquad k=0,\ 1,\ 2,\ \cdots$$

（3）计算样本自相关函数

$$\hat{\rho}_k=\frac{\hat{\gamma}_k}{\hat{\gamma}_0}$$

（4）将 $(k,\ \hat{\rho}_k)$ 描在笛卡尔坐标图上，若从某个 k 以后， $\hat{\rho}_k$ 明显接近于零，则该 k 即为所求阶数 q 的估计 $\hat{q}$ 。

2.　AIC（BIC）准则法

类似于 $AR(p)$ 模型中 p 的确定方法，确定 $MA(q)$ 模型中 q 的阶数步骤为：

（1）凭经验确定 q 的上限 Q，$0\le q\le Q$

（2）再由 $x_1,\ x_2,\ \cdots,\ x_n$ 求 $\sigma^2(j)$ 的估计 $\hat{\sigma}^2(j)$

$$\hat{\sigma}^2(j)=\hat{\gamma}_0(1+\sum_{i=1}^{j}\hat{\beta}_i^2) \qquad j=1,2,\cdots,Q$$

（3） 令 $\mathrm{A(j)}=\mathrm{AIC}(j)=\ln\hat{\alpha}^2(j)=\frac{2j}{n}$

或 $\mathrm{BIC}(j)=\ln\hat{\sigma}^2(j)+\frac{j\log n}{n}$

从 $A(1)$，$A(2)$，$\cdots$ ，$A(j)$，$\cdots$ ，$A(Q)$ 中选取最小值 $\hat{q}$，即 $\hat{q}$ 满足

$$\mathrm{AIC}(\hat{q})=\min_{1\le j\le Q}\mathrm{AIC}(j)$$

由此得出的 $\hat{q}$ 即为 q 的 AIC 准则估计。

二、$MA(q)$ 模型参数的估计方法

1. 矩估计法

设对原始数据作中心化处理后所得的中心化样本值数据为 $x_1,\cdots,x_n$，计算其自相关函数 γ_k 的估计值 $\hat{\gamma}_k$，再代入滑动平均模型 $MA(q)$ 的参数与自相关函数的关系式中，即得

$$\begin{cases}\hat{\sigma}^2(1+\sum\limits_{j=1}^{q}\beta_j^2)=\hat{\gamma}_0 & \qquad (3.4.2)\\ \hat{\sigma}^2(-\hat{\beta}_k+\sum\limits_{j=1}^{q-k}\hat{\beta}_j\hat{\beta}_{j+k})=\hat{\gamma}_k \qquad k=1,\ 2,\ \cdots,\ q & \qquad (3.4.3)\end{cases}$$

从中解出 $\hat{\beta}=(\hat{\beta}_1,\ \hat{\beta}_2,\ \cdots,\ \hat{\beta}_q)$ 与 $\hat{\sigma}^2$，这样的 $\hat{\beta}$ 与 $\hat{\sigma}^2$ 即为 $\beta,\ \sigma^2$ 的矩估计。

例如对于 $q=1$ 时的 $MA(1)$ 模型，由样本值 $x_1,\cdots,x_n$ 计算出 $\hat{\gamma}_0$ 与 $\hat{\gamma}_1$，代入关系式中即得

$$\begin{cases}\hat{\sigma}^2(1+\beta_1^2)=\hat{\gamma}_0\\ \hat{\sigma}^2(-\hat{\beta}_1)=\hat{\gamma}_1\end{cases},\qquad 从而有\ \begin{aligned}\hat{\sigma}^2&=-\frac{\hat{\gamma}_1}{\beta_1}\\ \hat{\beta}_1&=(\frac{\hat{\gamma}_0}{\hat{\sigma}^2}-1)^{\frac{1}{2}}\end{aligned}$$

为所求 $MA(1)$ 的参数矩估计。

2. 极大似然估计法

若 $\{\varepsilon_t\}$ 为独立且同分布序列，$\varepsilon_{\mathrm{t}}\sim N(0,\ \sigma^2)$ $\quad t=1,\ 2,\ \cdots$，而

$$x_t=\varepsilon_t-\sum_{k=1}^{q}\beta_k\varepsilon_{t-k}$$

由概率论知识可知，$x_1,\cdots,x_n$ 亦服从正态分布。即

$x=(x_1,\ \cdots,\ x_n)\sim N(0,\ \Gamma_n)$，其概率密度函数，即似然函数为：

$$L(\beta,\sigma^2)=f(x_1,x_2,\cdots,x_n;\beta,\sigma^2)=\frac{1}{(2\pi)^{\frac{n}{2}}|\Gamma_n|^{\frac{1}{2}}}\exp(-x'\Gamma_n^{-1}x)$$ 其中 Γ_n 为 $x_1,\cdots,x_n$ 的协方差阵，对上式两端取对数，即得

$$\ln L(\beta,\ \sigma^2)=\frac{n}{2}\ln(2\pi)-\frac{1}{2}\ln(\Gamma_n)-\frac{1}{2}x'\Gamma_n^{-1}x$$

再求 $\hat{\beta},\ \hat{\sigma}^2$ 使 $\ln L(\hat{\beta},\ \hat{\sigma}^2)=\max\{\ln L(\beta,\ \sigma^2)\}$，由此方法求出的则此 $\hat{\beta},\ \hat{\sigma}^2$ 即为 β 与 σ^2 极大似然估计。

但实际上求解上述似然函数的极大值，即求 β 与 σ^2 的严格的极大似然估计非常困难，所以在实际应用中经常采用近似的极大似然估计法

3. 近似极大似然估计法

设由（4.1）式中 $\{\varepsilon_t\}$的初值 ε_0，ε_1，$\varepsilon_2,\cdots$，ε_{-q+1}被给定，不妨假设它们近似为 0，这样，由（4.1）式和样本值 $x_1,\cdots,x_n$，可迭代计算得出 $\varepsilon_1,\cdots,\varepsilon_n$ 的值：

$$\varepsilon_k=x_k+\beta_1\varepsilon_{k-1}+\cdots+\beta_q\varepsilon_{k-q},\quad k=1,\ 2,\ \cdots,\ n \qquad (4.4)$$

而 $(\varepsilon_1,\cdots,\varepsilon_n)$ 可认为近似服从正态分布 $N(0,\ \sigma^2 I)$，因而 $x_1,\cdots,x_n$ 亦服从正态分布，故 $(\varepsilon_1,\cdots,\varepsilon_n)$ 的近似概率密度为

$$L(\beta,\sigma^2)=f(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n;\beta,\sigma^2)=\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n\exp\{-\frac{1}{2\sigma^2}\sum_{k=1}^{n}\varepsilon_k^2\}$$

对上式取对数，即得

$$\ln L(\beta,\sigma^2)\approx-\frac{n}{2}\ln(2\pi)-\frac{n}{2}\ln\sigma^2-\frac{1}{2\sigma^2}\sum_{k=1}^{n}\varepsilon_k^2 \qquad (3.4.5)$$

其中 ε_k 由假定及（4.4）式迭代而得，即 $\varepsilon_k=x_k+\beta_1\varepsilon_{k-1}+\cdots+\beta_q\varepsilon_{k-q}$，易见，它们不仅与 $x_1,\cdots,x_n$ 有关，亦与参数 β 有关，为确定参数 β 与 σ^2，记 $\varepsilon_1,\cdots,\varepsilon_n$ 的平方和为

$$S(\beta)=\sum_{k=1}^{n}\varepsilon_k^2=\sum_{k=1}^{n}(x_k+\beta_1\varepsilon_{k-1}+\cdots+\beta_q\varepsilon_{k-q})^2$$

现在，再求估计值 $\hat{\beta}=(\hat{\beta}_1,\ \hat{\beta}_2,\ \cdots,\ \hat{\beta}_q)$ 使 $S(\beta)$ 达到最小，即

$$S(\hat{\beta}) = \min\{S(\beta)\}$$

即为β的最小平方和估计，亦为β的近似极大似然估计。

由似然方程$\ln L(\beta, \sigma^2) \approx -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}S(\beta)$可知， $S(\beta)$满足方程：

$$\frac{\partial}{\partial\beta}[\ln L(\beta, \sigma^2] = -\frac{1}{2\sigma^2}\frac{\partial}{\partial\beta}[S(\beta)] = 0$$

$$\frac{\partial}{\partial\sigma^2}[\ln L(\beta, \sigma^2)] = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}S(\beta) = 0$$

因而，估计$S(\hat{\beta})$应满足等式：

$$\frac{\partial}{\partial\beta}[S(\hat{\beta})] = 0, \qquad \hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_n)$$

$S(\hat{\beta})$确定以后，σ^2的估计相应得出，即为

$$\hat{\sigma}^2 = \frac{S(\hat{\beta})}{n}$$

然而，这样的估计 $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2, \cdots, \hat{\beta}_n)$依然难以求出，实际上仍采用数值计算方法求$S(\beta)$的极小值。

4. 自回归逼近拟合方法

由于上述方法在实际中难以采用，所以我们讨论一种实际中便于使用的方法。我们注意到，对于可逆的MA（q)模型而言，它也可视作一个$AR(+\infty)$模型，因而在实际中，可以利用适当高阶的AR模型来近似代替低阶的MA模型，这样的方法称为自回归逼近拟合方法。具体做法是：

第一步：利用原始数据$x_1, \cdots, x_n$作自回归模型$AR(p)$的拟合，即：

(1) 用 AIC 准则确定阶数p值，或直接取较大的p值。一般地，当样本容量n不太大时，取$p \approx n/10$，当样本容量n较大时，取$p >> q$；

(2)在确定阶数p后，再利用尤尔—沃克方法估计AR（p)中参数

$\hat{\alpha}_1$，$\hat{\alpha}_2, \cdots$，$\hat{\alpha}_n$，使其满足：

$$x_t = \hat{\alpha}_1 x_{t-1} + \hat{\alpha}_2 x_{t-2} + \cdots + \hat{\alpha}_p x_{t-p} + \hat{\varepsilon}_t$$

第二步： 由上述估计的AR（p)模型递推计算残差列$\{\hat{\varepsilon}_t\}$，即：

$$\hat{\varepsilon}_t = x_t - \hat{\alpha}_1 x_{t-1} - \hat{\alpha}_2 x_{t-2} - \cdots - \hat{\alpha}_p x_{t-p} \ , \ t = p+1, p+2, \cdots, n$$

第三步： 再视残差列$\{\hat{\varepsilon}_t\}$为独立序列，利用线性回归模型：

$$x_t = \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \cdots + \beta_q \varepsilon_{t-q} + \varepsilon_t \ , \ t = p+1, p+2, \cdots, n$$

确定出参数$\hat{\beta}_1$，$\hat{\beta}_2, \cdots$，$\hat{\beta}_n$，即令

$$x = \begin{bmatrix} x_{p+q+1} \\ x_{p+q+2} \\ \vdots \\ x_n \end{bmatrix}, \quad \hat{E} = \begin{bmatrix} \hat{\varepsilon}_{p+q} & \hat{\varepsilon}_{p+q-1} & \cdots & \hat{\varepsilon}_{p+1} \\ \hat{\varepsilon}_{p+q+1} & \hat{\varepsilon}_{p+q} & \cdots & \hat{\varepsilon}_{p+1} \\ \vdots & \vdots & \cdots & \vdots \\ \hat{\varepsilon}_{n-1} & \hat{\varepsilon}_{n-2} & \cdots & \hat{\varepsilon}_{n-q} \end{bmatrix}, \quad \hat{\varepsilon} = \begin{bmatrix} \varepsilon_{p+q+1} \\ \varepsilon_{p+q+2} \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

则上式成为线性模型 $x = \hat{E}\beta + \varepsilon \ , \ \hat{\varepsilon} \sim N(0, \sigma^2 I)$

于是可得β的最小二乘估计为

$$\hat{\beta} = (\hat{E}^{\tau}\hat{E})^{-1}\hat{E}^{\tau}x$$

上述方法所得的β的估计即为自回归逼近拟合估计。由于此估计不涉及非线性代数方程，所以有明显的表达式，适宜在实际中使用。

三、拟合模型的检验

$MA(q)$模型的检验方法与$AR(p)$模型的检验方法一样，通常也采用检验$\{\varepsilon_t\}$是否独立序列的方法：如

（1）首先由初值ε_0，ε_{-1}，$\cdots$，$\varepsilon_{-\hat{q}+1}$与$MA(q)$模型

$$x_t = \varepsilon_t - \beta_1 \varepsilon_{t-1} - \beta_2 \varepsilon_{t-2} - \cdots - \beta_q \varepsilon_{t-\hat{q}} \ , \ t = 1, \ 2, \ \cdots, \ n$$

计算$\hat{\varepsilon}_1$，$\hat{\varepsilon}_2$，$\cdots$，$\hat{\varepsilon}_n$，其中 $\hat{q}$为q估计值，；

（2）检验$\hat{\varepsilon}_1$，$\hat{\varepsilon}_2$，$\cdots$，$\hat{\varepsilon}_n$是否独立序列，若是，则$MA(q)$模型为真，否则不是$MA(q)$ 模型。

也可采用拟合优度检验方法与峰度偏度检验方法作$\{\varepsilon_t\}$的正态性检验。