(3) GM 算法和 Tugnait 算法都是"过参数化的"，即不仅估计 $b(i)$ 本身，而且还估计 $b^2(i)$，因此只能是次最优的。相比之下，累积量算法则没有这一过参数化的缺点，因为它只估计 $b(i)$ 或者 $b^{k-1}(i)$。此外，GM 算法、Tugnait 算法和 w 切片法都只适用加性白噪声情况，而累积量方法在理论上可完全抑制加性高斯有色噪声。

MA 参数除了可以用四种线性法方程方法估计外，还有几种闭式递推估计方法，分别由 Gainnakis 与 Mendel [74], Swami 与 Mendel [180], Tugnait [188],[189] 以及 Zhang 与 Zhou [214] 提出 (限于篇幅，这里不一一介绍)。顺便指出，前四种递推同时使用相关函数和高阶累积量，因而只适用于加性白噪声，而最后一种递推只使用高阶累积量。

5.5.3 MA 阶数确定

在上面的讨论中，我们只涉及了 MA 模型的参数估计，而隐含假定 MA 阶数是已知的。实际中，这一阶数是需要在参数估计前事先确定的。

由特殊切片累积量的 BBR 公式 (5.5.21) 和 $C(q,n)$ 公式 (5.5.25) 易知

$$c_{ky}(q,0) = c_{ky}(q,0,\cdots,0) \neq 0 \tag{5.5.34a}$$

$$c_{ky}(q,n) = c_{ky}(q,n,\cdots,n) = 0, \quad \forall n > q \tag{5.5.34b}$$

上式意味着，MA 阶数 q 应该是满足式 (5.5.34b) 的最小正整数 n。这就是 Giannakis 与 Mendel [76] 的定阶方法。问题是，对于一组在高斯有色噪声中观测的短数据，样本累积量 $\hat{c}_{ky}(q,n)$ 往往表现出比较大的误差和方差，从而使得式 (5.5.34b) 的检验变得难于操作。

Zhang 与 Zhang [216] 提出了确定 MA 阶数 q 的奇异值分解方法，它具有很好的数值稳定性。这种方法的基本思想是：把 MA 阶数的估计变成矩阵秩的确定问题，其关键是构造下面的 $(q+1)\times(q+1)$ 累积量矩阵：

$$\boldsymbol{C}_{\mathrm{MA}} = \begin{bmatrix} c_{ky}(0,0) & c_{ky}(1,0) & \cdots & c_{ky}(q,0) \\ c_{ky}(1,0) & \cdots & c_{ky}(q,0) & \\ \vdots & \iddots & & \\ c_{ky}(q,0) & & & 0 \end{bmatrix} \tag{5.5.35}$$

由于对角线上的元素 $c_{ky}(q,0) \neq 0$，所以 $\boldsymbol{C}_{\mathrm{MA}}$ 显然是一个满秩矩阵，即有

$$\mathrm{rank}(\boldsymbol{C}_{\mathrm{MA}}) = q+1 \tag{5.5.36}$$

虽然 MA 的阶数估计现在变成了矩阵秩的确定，但是累积量矩阵包含了未知的阶数

q。为了使这一方法实用，我们来考虑扩展阶的累积量矩阵

$$C_{\mathrm{MA,e}} = \begin{bmatrix} c_{ky}(0,0) & c_{ky}(1,0) & \cdots & c_{ky}(q_e,0) \\ c_{ky}(1,0) & \cdots & c_{ky}(q_e,0) & \\ \vdots & ⋰ & & \\ c_{ky}(q_e,0) & & & 0 \end{bmatrix} \tag{5.5.37}$$

式中 $q_e > q$。由于 $c_{ky}(m,0) = 0, \forall m > q$，故容易验证

$$\mathrm{rank}(C_{\mathrm{MA,e}}) = \mathrm{rank}(C_{\mathrm{MA}}) = q + 1 \tag{5.5.38}$$

在实际应用中，累积量矩阵 $C_{\mathrm{MA,e}}$ 的元素用样本累积量代替，然后使用奇异值分解确定 $C_{\mathrm{MA,e}}$ 的有效秩 (它等于 $q+1$)，即可获得 q 的估计。

另一方面，由累积量矩阵 $C_{\mathrm{MA,e}}$ 的上三角结构易知，其有效秩的确定等价于判断对角线元素的乘积不等于零，即 q 是满足条件

$$c_{ky}^{m+1}(m,0) \neq 0, \quad m = 1,2,\cdots \tag{5.5.39}$$

的最大整数 m。显然，从数值性能出发，判断式 (5.5.39) 要比判断式 (5.5.34b) 在数值上稳定得多。

总结以上讨论，MA 阶数既可通过扩展阶累积量矩阵 $C_{\mathrm{MA,e}}$ 的有效秩确定，也可由式 (5.5.39) 估计。这两种线性代数方法具有稳定的数值性能。

5.6　因果 ARMA 模型的辨识

与上一节介绍的有限冲激响应 (FIR) 系统相比，无限冲激响应 (IIR) 系统具有更广泛的代表性。对于 IIR 系统，从参数吝啬的观点看，ARMA 模型比 MA 模型和 AR 模型更加合理，因为后两种模型使用的参数太多。本节主要讨论因果 ARMA 模型的辨识，而非因果 ARMA 模型的辨识则留待下一节考查。

5.6.1　AR 参数的辨识

考查差分模型

$$\sum_{i=0}^{p} a(i)x(n-i) = \sum_{i=0}^{q} b(i)e(n-i) \tag{5.6.1a}$$

并且 ARMA(p,q) 随机过程 $\{x(n)\}$ 在加性噪声 $v(n)$ 中被观测，即

$$y(n) = x(n) + v(n) \tag{5.6.1b}$$

不失一般性，假定 $a(0)=b(0)=h(0)=1$。

关于 ARMA 模型，通常作如下假设。

(AS1) 系统传递函数 $H(z)=B(z)/A(z)=\sum_{i=0}^{\infty}h(i)z^{-i}$ 不存在任何零、极点对消，即 $a(p)\neq 0$ 和 $b(q)\neq 0$。

(AS2) 输入 $e(n)$ 是非高斯的白噪声，具有有限的非零累积量 γ_{ke}。

(AS3) 观测噪声 $v(n)$ 是高斯有色的，并且与 $e(n)$ 和 $x(n)$ 独立。

条件 (AS1) 意味着系统是因果的 (时间为负时，冲激响应恒为零)，且 ARMA(p,q) 模型不能进一步简化。注意，关于系统的零点，我们并未作任何约束，这意味着它们可以位于单位圆内和外；如果不使用 ARMA 模型的逆系统，则还允许零点在单位圆上。

若记 $c_{kx}(m,n)=c_{kx}(m,n,0,\cdots,0)$，则在 (AS1) ~ (AS3) 条件下，由 BBR 公式易得

$$\begin{aligned}\sum_{i=0}^{p}a(i)c_{kx}(m-i,n)&=\gamma_{ke}\sum_{j=0}^{\infty}h^{k-2}(j)h(j+n)\sum_{i=0}^{p}a(i)h(j+m-i)\\&=\gamma_{ke}\sum_{j=0}^{\infty}h^{k-2}(j)h(j+n)b(j+m)\end{aligned}\tag{5.6.2}$$

式中利用了冲激响应的定义式

$$\sum_{i=0}^{p}a(i)h(n-i)=\sum_{j=0}^{q}b(j)\delta(n-j)=b(n)\tag{5.6.3}$$

由于假定 $a(p)\neq 0$ 和 $b(q)\neq 0$，并注意到 $b(j)\equiv 0\ (j>q)$，即可由式 (5.6.2) 得到一组重要的法方程：

$$\sum_{i=0}^{p}a(i)c_{kx}(m-i,n)=0,\quad m>q,\forall n\tag{5.6.4}$$

这就是用高阶累积量表示的修正 Yule-Walker 方程，简称 MYW 方程。注意，MYW 方程也可以写作其他形式，例如 [183]

$$\sum_{i=0}^{p}a(i)c_{kx}(i-m,n)=0,\quad m>q,\forall n\tag{5.6.5}$$

这一法方程可仿照式 (5.6.4) 的推导得出。若这两种 MYW 方程采用相同的 (m,n) 取值范围，则它们本质上等价。

一个重要问题是，如何在 MYW 方程中取适当的 m 和 n 值，才能保证 AR 参数的解是惟一的呢？这个问题称为基于高阶累积量的 AR 参数可辨识性。不妨先看一个例子[183]。

例 5.6.1 考察下面的因果最大相位系统

$$H(z)=\frac{(z-\alpha_1^{-2})(z-\alpha_1^{-1}\alpha_2^{-1})}{(z-\alpha_1)(z-\alpha_2)} \tag{5.6.6}$$

式中 $\alpha_1\neq\alpha_2$。假定使用三阶累积量，则当取 $n\neq 0$ 时， MYW 方程为

$$c_{kx}(-m,n)-\alpha_2 c_{kx}(1-m,n)=0,\quad m>1,n\neq 0 \tag{5.6.7}$$

此式表明，如将 $n=0$ 排除在外，则无论怎样选择 m 和 n 值的组合，都不可能辨识出极点 α_1。然而，若将 $n=0$ 包含在内，则 MYW 方程将与式 (5.6.7) 不同，并且能够同时辨识出极点 α_1 和 α_2。

这个例子说明，选择 m 和 n 值的组合需要小心，不能任意。那么，怎样的组合才能保证 AR 参数的惟一可辨识性能呢？下面的定理给出了这个问题的答案。

定理 5.6.1 在假设条件 (AS1) ~ (AS3) 之下，ARMA 模型 (5.6.1) 的 AR 参数可以由

$$\sum_{i=0}^{p}a(i)c_{ky}(m-i,n)=0,\quad m=q+1,\cdots,q+p;\quad n=q-p,\cdots,q \tag{5.6.8}$$

的最小二乘解惟一辨识。

证明 由假设 (AS3) 知 $c_{ky}(\tau_1,\cdots,\tau_{k-1})=c_{kx}(\tau_1,\cdots,\tau_{k-1})$，故式 (5.6.8) 可以写成下面的矩阵形式：

$$\boldsymbol{C}\boldsymbol{a}=-\boldsymbol{c} \tag{5.6.9}$$

式中

$$\boldsymbol{C}=\begin{bmatrix} c_{kx}(q+1-p,q-p) & \cdots & c_{kx}(q,q-p)\\ \vdots & \vdots & \vdots\\ c_{kx}(q+1-p,q) & \cdots & c_{kx}(q,q)\\ \vdots & \vdots & \vdots\\ c_{kx}(q,q-p) & \cdots & c_{kx}(q+p-1,q-p)\\ \vdots & \vdots & \vdots\\ c_{kx}(q,q) & \cdots & c_{kx}(q+p-1,q)\end{bmatrix}$$

和

$$\boldsymbol{a} = [a(p), a(p-1), \cdots, a(1)]^{\mathrm{T}}$$

$$\boldsymbol{c} = -\begin{bmatrix} c_{kx}(q+1, q-p) \\ \vdots \\ c_{kx}(q+1, q) \\ \vdots \\ c_{kx}(q+p, q-p) \\ \vdots \\ c_{kx}(q+p, q) \end{bmatrix}$$

于是，只要证明累积量矩阵 $\boldsymbol{C}$ 具有满列秩即可。在式 (5.6.2) 中取 $m=q$，则有

$$\begin{aligned} \sum_{i=0}^{p} a(i) c_{kx}(q-i, n) &= \gamma_{ke} \sum_{j=0}^{\infty} h^{k-2}(j) h(j+n) b(q+j) \\ &= \gamma_{ke} h^{k-2}(0) h(n) b(q) = \mu h(n) \end{aligned} \tag{5.6.10}$$

式中 $\mu = \gamma_{ke} h^{k-2}(0) b(q) \neq 0$。令 $\boldsymbol{H}$ 代表系统的 Hankel 矩阵 ①，即

$$\boldsymbol{H} = \mu\{h(q-p-1+i+j)\}_{i,j=1}^{p} \tag{5.6.11}$$

则式 (5.6.10) 可以写作

$$\boldsymbol{H} = \mu \begin{bmatrix} h(q+1-p) & h(q+2-p) & \cdots & h(q) \\ h(q+2-p) & h(q+3-p) & \cdots & h(q+1) \\ \vdots & \vdots & \vdots & \vdots \\ h(q) & h(q+1) & \cdots & h(q+p-1) \end{bmatrix}$$

$$= \begin{bmatrix} \tilde{a}^{\mathrm{T}} & 0^{\mathrm{T}} & \cdots & 0^{\mathrm{T}} \\ 0^{\mathrm{T}} & \tilde{a}^{\mathrm{T}} & \cdots & 0^{\mathrm{T}} \\ \vdots & \vdots & \vdots & \vdots \\ 0^{\mathrm{T}} & 0^{\mathrm{T}} & \cdots & \tilde{a}^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} c_{kx}(q+1-p, q-p) & \cdots & c_{kx}(q, q-p) \\ \vdots & \vdots & \vdots \\ c_{kx}(q+1-p, q) & \cdots & c_{kx}(q, q) \\ \vdots & \vdots & \vdots \\ c_{kx}(q, q-p) & \cdots & c_{kx}(q+p-1, q-p) \\ \vdots & \vdots & \vdots \\ c_{kx}(q, q) & \cdots & c_{kx}(q+p-1, q) \end{bmatrix} \tag{5.6.12a}$$

或简记为

$$\boldsymbol{H} = \boldsymbol{A}\boldsymbol{C} \tag{5.6.12b}$$

① 斜对角线上的元素相等的矩阵称为 Hankel 矩阵。

由于 $\tilde{a}^{\mathrm{T}} \neq \mathbf{0}^{\mathrm{T}}$，所以矩阵 $\boldsymbol{A}$ 是非奇异或满秩的。因此，如果 $\boldsymbol{H}$ 是满列秩的话，则 $\boldsymbol{C}$ 便一定是满列秩的。由文献 [71, p.244] 知，若 ARMA(p,q) 模型无零、极点对消，则其 Hankel 矩阵 $\boldsymbol{H}$ 是满列秩的。由于定理中的 ARMA 模型满足这一条件，故 $\boldsymbol{H}$ 满列秩，这使得矩阵 $\boldsymbol{C}$ 也是满列秩的，从而 AR 参数是惟一可辨识的。 ■

这里指出，除了上述基于矩阵理论的证明 [75],[76] 外，也可利用系统函数理论来证明定理 5.6.1。

原理上，AR 阶数 p 的确定就是求累积量矩阵 $\boldsymbol{C}$ 的秩。但是，由于 $\boldsymbol{C}$ 本身用到未知的阶数 p 和 q，故必须改造 $\boldsymbol{C}$ 的结构，以使得新的累积量不再含未知的 p 和 q，但却仍然具有秩 p。考虑到与 AR 参数估计的总体最小二乘法之间的联系，改造的结果可以叙述如下。

定理 5.6.2 定义 $M_2(N_2-N_1+1)\times M_2$ 扩展阶累积量矩阵

$$\boldsymbol{C}_{\mathrm{e}}=\begin{bmatrix} c_{kx}(M_1,N_1) & \cdots & c_{kx}(M_1+M_2-1,N_1) \\ \vdots & \vdots & \vdots \\ c_{kx}(M_1,N_2) & \cdots & c_{kx}(M_1+M_2-1,N_2) \\ \vdots & \vdots & \vdots \\ c_{kx}(M_1+M_2,N_1) & \cdots & c_{kx}(M_1+2M_2-1,N_1) \\ \vdots & \vdots & \vdots \\ c_{kx}(M_1+M_2,N_2) & \cdots & c_{kx}(M_1+2M_2-1,N_2) \end{bmatrix} \tag{5.6.13}$$

式中，取 $M_1 \geqslant q+1-p$，$M_2 \geqslant p$，$N_1 \leqslant q-p$ 和 $N_2 \geqslant q$。当且仅当 ARMA(p,q) 模型无零、极点对消时，矩阵 $\boldsymbol{C}_{\mathrm{e}}$ 具有秩 p。

证明 仿照证明式 (5.6.12b) 的步骤，可得

$$\boldsymbol{H}_{\mathrm{e}}=\boldsymbol{A}_{\mathrm{e}}\boldsymbol{C}_{\mathrm{e}} \tag{5.6.14}$$

式中 $\boldsymbol{H}_{\mathrm{e}}$ 表示系统的 Hankel 矩阵，即 $\boldsymbol{H}_{\mathrm{e}}=\{h(M_1+i+j-2)\}_{i,j}^{M_2}$，并且 $M_2\times M_2(N_2-N_1+1)$ 矩阵 $\boldsymbol{A}_{\mathrm{e}}$ 类似于式 (5.6.12b) 中矩阵 $\boldsymbol{A}$ 的定义，只不过其中的向量 $\tilde{\boldsymbol{a}}^{\mathrm{T}}$ 现在被 $\tilde{\boldsymbol{a}}_{\mathrm{e}}^{\mathrm{T}}=[0,\cdots,0,\tilde{\boldsymbol{a}}^{\mathrm{T}},0,\cdots,0]$ 所代替。

充分性：由矩阵秩的性质知 $\mathrm{rank}(\boldsymbol{H}_{\mathrm{e}})\leqslant\min[M_2,\mathrm{rank}(\boldsymbol{C}_e)]$。据文献 [71, p.244] 知，当 ARMA(p,q) 模型无零、极点对消时，Hankel 矩阵满秩，即 $\mathrm{rank}(\boldsymbol{H}_{\mathrm{e}})=p$。又由于 $\boldsymbol{A}_{\mathrm{e}}$ 具有满行秩 M_2(这里 $M_2\geqslant p$)，所以矩阵 $\boldsymbol{C}_{\mathrm{e}}$ 的秩不可能小于 p。但是，由 MYW 方程易看出，矩阵 $\boldsymbol{C}_{\mathrm{e}}$ 第 $p+1$ 列是其左边 p 列的线性组合，而第 $p+2$ 列又是其左边 p 列的线性组合，等等。这表明，在矩阵 $\boldsymbol{C}_{\mathrm{e}}$ 中最多只有 p 个独立的列，该矩阵的秩不可能大于 p。因此，$\mathrm{rank}(\boldsymbol{C}_{\mathrm{e}})=p$。

必要性: 假定 ARMA(p,q) 模型存在零、极点对消。此时，MYW 方程变为

$$\sum_{i=0}^{\tilde{p}} a(i) c_{kx}(m-i,n)=0, \quad m>p, \forall n$$

其中 $\tilde{p}<p$。这说明，在矩阵 C_e 中第 $\tilde{p}+1$ 列是其左边 $\tilde{p}$ 列的线性组合，而第 $\tilde{p}+2$ 列又是其左边 $\tilde{p}$ 列的线性组合，等等。换言之，矩阵 C_e 中最多只有 $\tilde{p}$ 列是线性独立的，故其秩最大为 $\tilde{p}\ (<p)$，不可能等于 p。 ■

定理 5.6.2 是由 Giannakis 与 Mendel [76] 建立的。这一定理告诉我们，若用观测数据的样本累积量 $\hat{c}_{ky}(m,n)$ 代替信号的真实累积量 $c_{kx}(m,n)$，则样本累积量矩阵 C_e 的有效秩将等于 p。

将上述结果加以整理，可得到因果 ARMA 模型 AR 阶数确定和 AR 参数估计的奇异值分解 – 总体最小二乘 (SVD-TLS) 算法如下。

算法 5.6.1 (基于累积量的 SVD-TLS 算法)

步骤 1 构造 $M_2(N_2-N_1+1)\times M_2$ 扩展阶样本累积量矩阵

$$C_e=\begin{bmatrix} \hat{c}_{kx}(M_1+M_2-1,N_1) & \cdots & \hat{c}_{kx}(M_1,N_1) \\ \vdots & \vdots & \vdots \\ \hat{c}_{kx}(M_1+M_2-1,N_2) & \cdots & \hat{c}_{kx}(M_1,N_2) \\ \vdots & \vdots & \vdots \\ \hat{c}_{kx}(M_1+2M_2-1,N_1 & \cdots & \hat{c}_{kx}(M_1+M_2,N_1) \\ \vdots & \vdots & \vdots \\ \hat{c}_{kx}(M_1+2M_2-1,N_2 & \cdots & \hat{c}_{kx}(M_1+M_2,N_2) \end{bmatrix} \tag{5.6.15}$$

并计算其奇异值分解 $C_e=U\Sigma V^{\mathrm{T}}$，存储矩阵 V。

步骤 2 确定矩阵 C_e 的有效秩，给出 AR 阶数估计 p。

步骤 3 计算 $(p+1)\times(p+1)$ 矩阵

$$S^{(p)}=\sum_{j=1}^{p}\sum_{i=1}^{M_2-p}\sigma_j^2 v_j^i (v_j^i)^{\mathrm{T}} \tag{5.6.16}$$

其中 $v_j^i=[v(i,j),v(i+1,j),\cdots,v(i+p,j)]^{\mathrm{T}}$，而 $v(i,j)$ 是矩阵 V 的第 i 行、第 j 列元素。

步骤 4 求矩阵 $S^{(p)}$ 的逆矩阵 $S^{-(p)}$, AR 参数的总体最小二乘估计由下式给出：

$$\hat{a}(i)=S^{-(p)}(i+1,1)/S^{-(p)}(1,1) \tag{5.6.17}$$

其中 $S^{-(p)}(i,1)$ 是逆矩阵 $S^{-(p)}$ 的第 i 行、第 1 列元素。

应用 SVD-TLS 算法求解 MYW 方程，可以明显提高 AR 参数的估计精度，这一点对于使用高阶累积量进行系统辨识显得尤其重要，因为短数据的样本高阶累积量存在比较大的估计误差。

5.6.2　MA 阶数确定

如果因果 ARMA 过程的 AR 部分已经辨识，则可以利用已知的 AR 阶数和 AR 参数对原观测过程进行"滤波"，得到

$$\tilde{y}(n)=\sum_{i=0}^{p}a(i)y(n-i) \tag{5.6.18}$$

称 $\{\tilde{y}(n)\}$ 为"残差时间序列"。将式 (5.6.1b) 代入上式，则有

$$\begin{aligned}\tilde{y}(n)&=\sum_{i=0}^{p}a(i)x(n-i)+\sum_{i=0}^{p}a(i)v(n-i)\\&=\sum_{j=0}^{q}b(j)e(n-j)+\tilde{v}(n)\end{aligned} \tag{5.6.19a}$$

在得到第二个等式的第一项时，代入了式 (5.6.1a)；并且

$$\tilde{v}(n)=\sum_{i=0}^{p}a(i)v(n-i) \tag{5.6.19b}$$

仍然为高斯有色噪声。

式 (5.6.19) 表明，残差时间序列 $\{\tilde{y}(n)\}$ 是一在高斯有色噪声中观测的 MA(q) 过程。因此，只要将 5.5 节介绍的 FIR 系统参数估计的 RC 方法或累积量方法应用于残差时间序列，即可获得因果 ARMA 过程的 MA 参数估计。

将因果 ARMA 过程的 MA 参数估计转变为残差时间序列的纯 FIR 系统参数辨识的方法称为"残差时间序列方法"，是 Giannakis 与 Mendel [76] 提出的。他们还同时提出了估计 ARMA 模型的 MA 阶数的方法，即用满足

$$c_{k\tilde{y}}(m,0,\cdots,0)\neq 0 \tag{5.6.20}$$

的最大整数 m 作为 ARMA 模型的 MA 阶数。更一般地，所有适用于纯 MA 模型的阶数确定方法都可应用于残差时间序列，从而获得 ARMA 模型的 MA 阶数估计。

现在考虑无须构造残差时间序列，即可直接进行 ARMA 模型 MA 阶数确定的

方法。为此，定义拟合误差函数

$$f_k(m,n)=\sum_{i=0}^{p}a(i)c_{ky}(m-i,n)=\sum_{i=0}^{p}a(i)c_{kx}(m-i,n) \tag{5.6.21}$$

式中使用了 $c_{ky}(m,n)=c_{kx}(m,n)$ 的结果。将式 (5.6.2) 代入式 (5.6.21)，立即得

$$f_k(m,n)=\gamma_{ke}\sum_{j=0}^{\infty}h^{k-2}(j)h(j+n)b(j+m) \tag{5.6.22}$$

显然有

$$f_k(q,0)=\gamma_{ke}h^{k-1}(0)b(q)=\gamma_{ke}b(q)\neq 0 \tag{5.6.23}$$

因为 $h(0)=1$ 和 $b(q)\neq 0$。

另外一方面，MYW 方程也可以使用拟合误差函数表示为

$$f_k(m,n)=0,\quad m>q,\forall n \tag{5.6.24}$$

式 (5.6.23) 和式 (5.6.24) 一起启示了一种确定 ARMA 模型 MA 阶数 q 的方法，即 q 是使得

$$f_k(m)\overset{\text{def}}{=}f_k(m,0)=\sum_{i=0}^{p}a(i)c_{ky}(m-i,0,\cdots,0)\neq 0 \tag{5.6.25}$$

成立的最大整数 m。虽然这一定阶方法在理论上有吸引力，但是在实际中却难于采用，因为当数据比较短时，估计值 $\hat{f}_k(m)$ 会存在很大的误差。

为了克服这一困难，Zhang 与 Zhang [216] 提出了一种简单而有效的方法。其基本思想是：引入一拟合误差矩阵，将 MA 阶数的确定转变为该矩阵的有效秩确定，而后者可以使用奇异值分解这样一种数值稳定的方法来实现。拟合误差矩阵的构造如下：

$$\boldsymbol{F}=\begin{bmatrix} f_k(0) & f_k(1) & \cdots & f_k(q) \\ f_k(1) & \cdots & f_k(q) & \\ \vdots & \iddots & & \\ f_k(q) & & & 0 \end{bmatrix} \tag{5.6.26}$$

这是一个 Hankel 矩阵，也是一个上三角矩阵。显而易见，

$$\det(\boldsymbol{F})=\prod_{i=1}^{q+1}f(i,q+2-i)=f_k^{q+1}(q)\neq 0 \tag{5.6.27}$$

这就是说，$\operatorname{rank}(\boldsymbol{F})=q+1$。由于 MA 阶数 q 是待确定的，所以有必要对拟合误差矩阵 $\boldsymbol{F}$ 加以改造，使得它的元素不再包含未知的 q，但却仍然具有秩 $q+1$。下面的

扩展阶拟合误差矩阵提供了这样的解：

$$\boldsymbol{F}_e = \begin{bmatrix} f_k(0) & f_k(1) & \cdots & f_k(q_e) \\ f_k(1) & \cdots & f_k(q_e) & \\ \vdots & ⋰ & & \\ f_k(q_e) & & & 0 \end{bmatrix}, \quad q_e > q \tag{5.6.28}$$

应用式 (5.6.23) 和式 (5.6.24)，容易证明

$$\text{rank}(\boldsymbol{F}_e) = \text{rank}(\boldsymbol{F}) = q + 1 \tag{5.6.29}$$

虽然 q 是未知的，但我们不难选择扩展阶 $q_e > q$。

式 (5.6.29) 表明，若矩阵 $\boldsymbol{F}_e$ 的元素用其估计值 $\hat{f}_k(m)$ 代替，则利用奇异值分解 (SVD)，可以确定矩阵 $\boldsymbol{F}_e$ 的有效秩。另一方面，由矩阵 $\boldsymbol{F}_e$ 的上三角结构知，其有效秩的确定等价于试验

$$\hat{f}_k^{m+1}(m) \neq 0 \tag{5.6.30}$$

使上式近似成立的最大整数 m 即是 MA 阶数 q 的估计值。这一试验被称为对角元素乘积 (PODE) 试验。显然，PODE 试验 (5.6.30) 比直接试验式 (5.6.25) 合理得多，因为可以期望前者提供的数值稳定性比后者的数值稳定性明显好。

Zhang 与 Zhang [216] 在仿真中发现，单独使用 SVD 或 PODE 试验确定 ARMA 模型的 MA 阶数，有时会出现过定阶或欠定阶，并建议两种方法综合起来应用。具体操作是：以 SVD 确定的阶数 M 为参考，若 $f_k^{M+1}(M)$ 和 $f_k^{M+2}(M+1)$ 明显不近似为零，则此 M 值为欠定，应该对阶数 $M' = M+1$ 试验式 (5.6.30)；反之，若 $f_k^{M+1}(M)$ 和 $f_k^{M+2}(M+1)$ 均明显近似为零，则此 M 值为超定，应对阶数 $M' = M-1$ 试验式 (5.6.30)。只有当 $f_k^{M+1}(M)$ 明显不等于零，而 $f_k^{M+2}(M+1)$ 明显接近零时，M 值是合适的。

5.6.3 MA 参数估计

前面已指出，对残差时间序列应用 5.5 节介绍的纯 MA 参数估计的 RC 方法或高阶累积量方法，都可以获得 ARMA 模型的 MA 参数的估计值，下面对这一估计问题作专门的讨论。

1. 残差时间序列累积量法

事实上，在不直接构造残差时间序列的情况下，仍然能够获得 ARMA 模型的 MA 参数估计。