



第三章

平稳时间序列分析



本章结构

- 方法性工具
- ARMA模型
- 平稳序列建模
- 序列预测



3.1 方法性工具

- 差分运算
- 延迟算子
- 线性差分方程



差分运算

- 一阶差分

$$\nabla x_t = x_t - x_{t-1}$$

- p 阶差分

$$\nabla^p x_t = \nabla^{p-1} x_t - \nabla^{p-1} x_{t-1}$$

- k 步差分

$$\nabla_k = x_t - x_{t-k}$$



延迟算子

- 延迟算子类似于一个时间指针，当前序列值乘以一个延迟算子，就相当于把当前序列值的时间向过去拨了一个时刻
- 记**B**为延迟算子，有

$$x_{t-p} = B^p x_t, \forall p \geq 1$$



延迟算子的性质

- $B^0 = 1$
- $B(c \cdot x_t) = c \cdot B(x_t) = c \cdot x_{t-1}$, c 为任意常数
- $B(x_t \pm y_t) = x_{t-1} \pm y_{t-1}$
- $B^n x_t = x_{t-n}$
- $(1 - B)^n = \sum_{i=0}^n (-1)^i C_n^i B^i$, 其中 $C_n^i = \frac{n!}{i!(n-i)!}$



用延迟算子表示差分运算

- p 阶差分

$$\nabla^p x_t = (1 - B)^p x_t = \sum_{i=0}^p (-1)^i C_p^i x_{t-i}$$

- k 步差分

$$\nabla_k = x_t - x_{t-k} = (1 - B^k) x_t$$



线性差分方程

- 线性差分方程

$$z_t + a_1 z_{t-1} + a_2 z_{t-2} + \cdots + a_p z_{t-p} = h(t)$$

- 齐次线性差分方程

$$z_t + a_1 z_{t-1} + a_2 z_{t-2} + \cdots + a_p z_{t-p} = 0$$



齐次线性差分方程的解

- 特征方程

$$\lambda^p + a_1\lambda^{p-1} + a_2\lambda^{p-2} + \cdots + a_p = 0$$

- 特征方程的根称为特征根，记作 $\lambda_1, \lambda_2, \dots, \lambda_p$

- 齐次线性差分方程的通解

- 不相等实数根场合

$$z_t = c_1\lambda_1^t + c_2\lambda_2^t + \cdots + c_p\lambda_p^t$$

- 有相等实根场合

$$z_t = (c_1 + c_2t + \cdots + c_d t^{d-1})\lambda_1^t + c_{d+1}\lambda_{d+1}^t + \cdots + c_p\lambda_p^t$$

- 复根场合

$$z_t = r^t (c_1 e^{it\omega} + c_2 e^{-it\omega}) + c_3\lambda_3^t + \cdots + c_p\lambda_p^t$$



非齐次线性差分方程的解

- 非齐次线性差分方程的特解

- 使得非齐次线性差分方程成立的任意一个解 z_t''

$$z_t'' + a_1 z_{t-1}'' + a_2 z_{t-2}'' + \cdots + a_p z_{t-p}'' = h(t)$$

- 非齐次线性差分方程的通解

- 齐次线性差分方程的通解和非齐次线性差分方程的特解之和 z_t

$$z_t = z_t' + z_t''$$



3.2 ARMA模型的性质

- AR模型 (Auto Regression Model)
- MA模型 (Moving Average Model)
- ARMA模型 (Auto Regression Moving Average model)

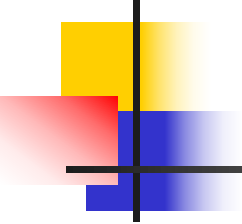


AR模型的定义

- 具有如下结构的模型称为 p 阶自回归模型，简记为 $AR(p)$

$$\begin{cases} x_t = \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + \varepsilon_t \\ \phi_p \neq 0 \\ E(\varepsilon_t) = 0, \text{Var}(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t \\ E x_s \varepsilon_t = 0, \forall s < t \end{cases}$$

- 特别当 $\phi_0 = 0$ 时，称为中心化 $AR(p)$ 模型



AR(P)序列中心化变换

- 称 $\{y_t\}$ 为中心化序列，令

$$\mu = \frac{\phi_0}{1 - \phi_1 - \dots - \phi_p}$$

$$y_t = x_t - \mu$$



自回归系数多项式

- 引进延迟算子，中心化AR(模型)又可以简记为

$$\Phi(B)x_t = \varepsilon_t$$

- 自回归系数多项式

$$\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$$



AR模型平稳性判别

- 判别原因
 - AR模型是常用的平稳序列的拟合模型之一，但并非所有的AR模型都是平稳的
- 判别方法
 - 单位根判别法
 - 平稳域判别法



例3.1:考察如下四个模型的平稳性

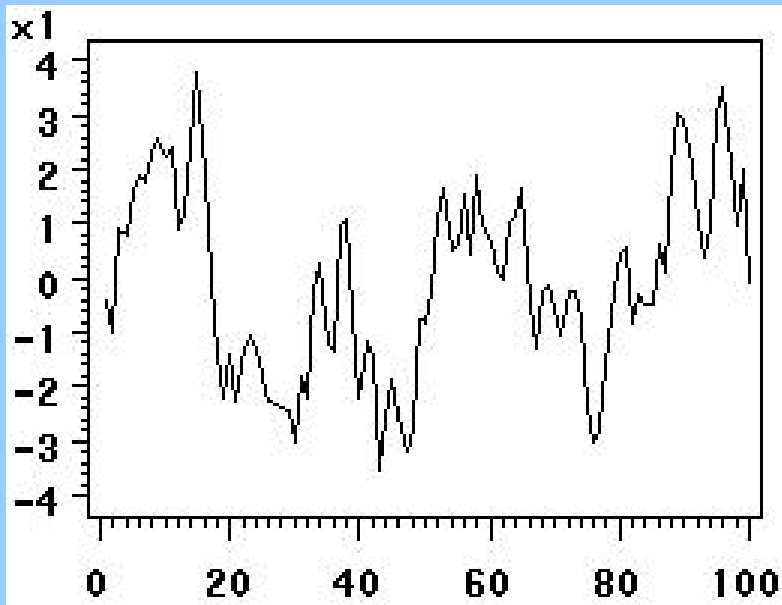
$$(1) x_t = 0.8x_{t-1} + \varepsilon_t$$

$$(2) x_t = -1.1x_{t-1} + \varepsilon_t$$

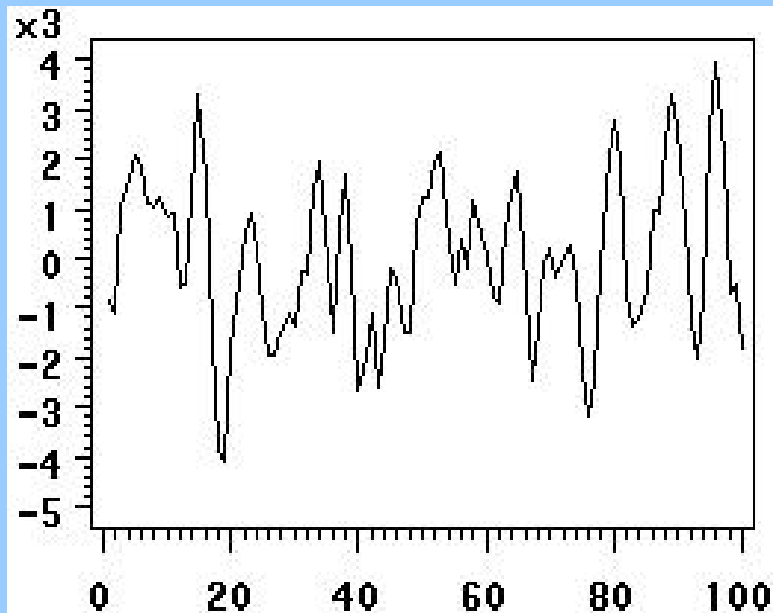
$$(3) x_t = x_{t-1} - 0.5x_{t-2} + \varepsilon_t$$

$$(4) x_t = x_{t-1} + 0.5x_{t-1} + \varepsilon_t$$

例3.1 平稳序列时序图

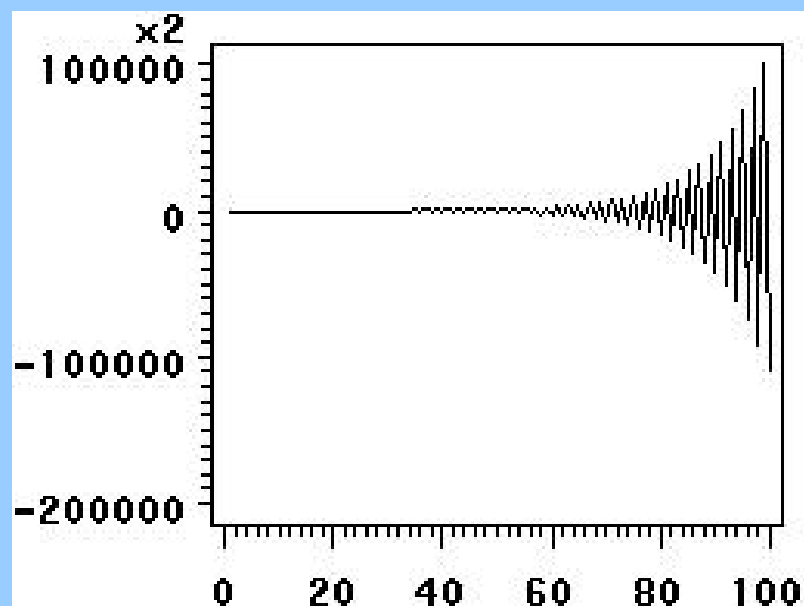


$$(1) x_t = 0.8x_{t-1} + \varepsilon_t$$

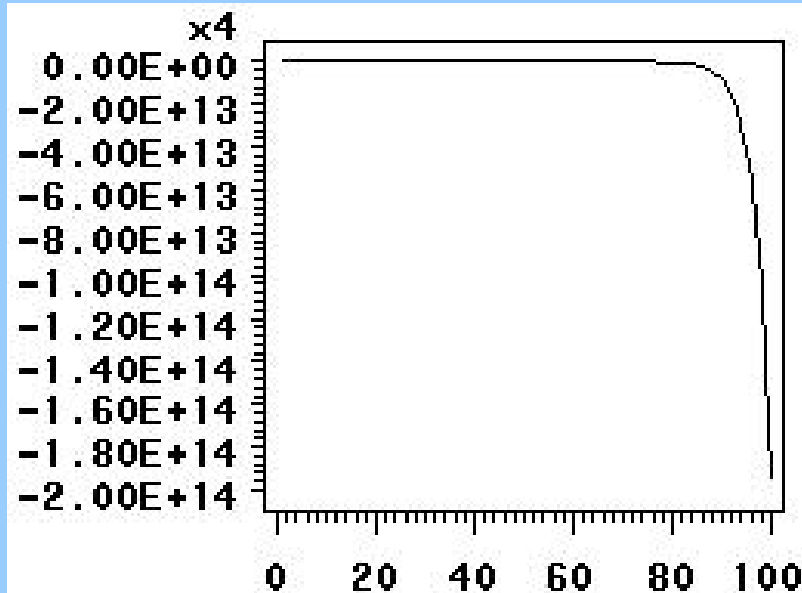


$$(3) x_t = x_{t-1} - 0.5x_{t-2} + \varepsilon_t$$

例3.1非平稳序列时序图



$$(2) x_t = -1.1x_{t-1} + \varepsilon_t$$



$$(4) x_t = x_{t-1} + 0.5x_{t-1} + \varepsilon_t$$



AR模型平稳性判别方法

■ 特征根判别

- AR(p)模型平稳的充要条件是它的p个特征根都在单位圆内
- 根据特征根和自回归系数多项式的根成倒数的性质, 等价判别条件是该模型的自回归系数多项式的根都在单位圆外

■ 平稳域判别

- 平稳域

$\{\phi_1, \phi_2, \dots, \phi_p \mid \text{单位根都在单位圆内}\}$



AR(1)模型平稳条件

- 特征根

$$\lambda = \phi$$

- 平稳域

$$|\phi| < 1$$

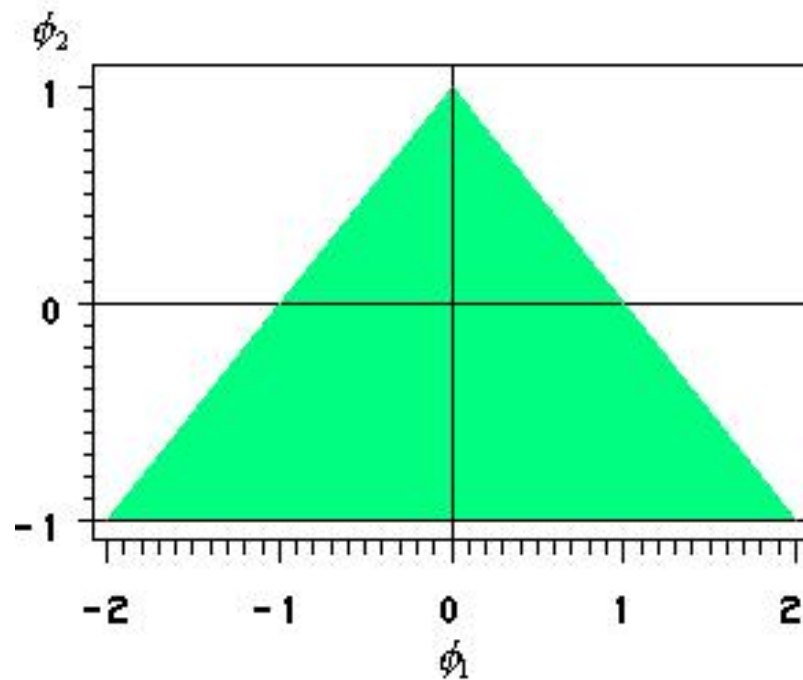
AR(2)模型平稳条件

■ 特征根

$$\lambda_1 = \frac{\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2}$$

$$\lambda_2 = \frac{\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2}$$

■ 平稳域



$$\{\phi_1, \phi_2 \mid |\phi_2| < 1, \text{ 且 } \phi_2 \pm \phi_1 < 1\}$$



例3.1 平稳性判别

模型	特征根判别	平稳域判别	结论
(1)	$\lambda_1 = 0.8$	$\phi = 0.8$	平稳
(2)	$\lambda_1 = -1.1$	$\phi = -1.1$	非平稳
(3)	$\lambda_1 = \frac{1+i}{2}$ $\lambda_2 = \frac{1-i}{2}$	$ \phi_2 = 0.5, \phi_2 + \phi_1 = 0.5, \phi_2 - \phi_1 = -1.5$	平稳
(4)	$\lambda_1 = \frac{1+\sqrt{3}}{2}$ $\lambda_2 = \frac{1-\sqrt{3}}{2}$	$ \phi_2 = 0.5, \phi_2 + \phi_1 = 1.5, \phi_2 - \phi_1 = -0.5$	非平稳



平稳AR模型的统计性质

- 均值
- 方差
- 协方差
- 自相关系数
- 偏自相关系数



均值

- 如果AR(p)模型满足平稳性条件，则有

$$Ex_t = E(\phi_0 + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + \varepsilon_t)$$

- 根据平稳序列均值为常数，且 $\{\varepsilon_t\}$ 为白噪声序列，有

$$Ex_t = \mu, E(\varepsilon_t) = 0, \forall t \in T$$

- 推导出

$$\mu = \frac{\phi_0}{1 - \phi_1 - \cdots - \phi_p}$$



Green函数定义

- AR模型的传递形式

$$\begin{aligned}x_t &= \frac{\varepsilon_t}{\Phi(B)} = \sum_{i=1}^p \frac{k_i}{1 - \lambda_i B} \varepsilon_t = \sum_{i=1}^p \sum_{j=0}^{\infty} k_i (\lambda_i B)^j \varepsilon_t \\ &= \sum_{j=0}^{\infty} \sum_{i=1}^p k_i \lambda_i^j \varepsilon_{t-j} \triangleq \sum_{j=0}^{\infty} G_j \varepsilon_{t-j}\end{aligned}$$

- 其中系数 $\{G_j, j = 1, 2, \dots\}$ 称为Green函数



Green函数递推公式

- 原理

$$\begin{cases} \Phi(B)x_t = \varepsilon_t \\ x_t = G(B)\varepsilon_t \end{cases} \Rightarrow \Phi(B)G(B)\varepsilon_t = \varepsilon_t$$

- 方法

- 待定系数法

- 递推公式

$$\begin{cases} G_0 = 1 \\ G_j = \sum_{k=1}^j \phi'_k G_{j-k}, \quad j = 1, 2, \dots \end{cases}, \text{其中 } \phi'_k = \begin{cases} \phi_k, & k \leq p \\ 0, & k > p \end{cases}$$



方差

- 平稳AR模型的传递形式

$$x_t = \sum_{j=0}^{\infty} G_j \varepsilon_{t-j}$$

- 两边求方差得

$$\text{Var}(x_t) = \sum_{j=0}^{\infty} G_j^2 \sigma_{\varepsilon}^2, G_j \text{为Green函数}$$



例3.2:求平稳AR(1)模型的方差

- 平稳AR(1)模型的传递形式为

$$x_t = \frac{\varepsilon_t}{1 - \phi_1 B} = \sum_{i=0}^{\infty} (\phi_1 B)^i \varepsilon_t = \sum_{i=0}^{\infty} \phi_1^i \varepsilon_{t-i}$$

- Green函数为

$$G_j = \phi_1^j, j = 0, 1, \dots$$

- 平稳AR(1)模型的方差

$$\text{Var}(x_t) = \sum_{j=0}^{\infty} G_j^2 \text{Var}(\varepsilon_t) = \sum_{j=0}^{\infty} \phi_1^{2j} \sigma_\varepsilon^2 = \frac{\sigma_\varepsilon^2}{1 - \phi_1^2}$$



协方差函数

- 在平稳AR(p)模型两边同乘 x_{t-k} , $\forall k \geq 1$, 再求期望

$$E(x_t x_{t-k}) = \phi_1 E(x_{t-1} x_{t-k}) + \cdots + \phi_p E(x_{t-p} x_{t-k}) + E(\varepsilon_t x_{t-k})$$

- 根据

$$E(\varepsilon_t x_{t-k}) = 0 \quad , \forall k \geq 1$$

- 得协方差函数的递推公式

$$Y_k = \phi_1 Y_{k-1} + \phi_2 Y_{k-2} + \cdots + \phi_p Y_{k-p}$$



例3.3:求平稳AR(1)模型的协方差

- 递推公式

$$Y_k = \phi_1 Y_{k-1} = \phi_1^k Y_0$$

- 平稳AR(1)模型的方差为

$$Y_0 = \frac{\sigma_\varepsilon^2}{1 - \phi_1^2}$$

- 协方差函数的递推公式为

$$Y_k = \phi_1^k \frac{\sigma_\varepsilon^2}{1 - \phi_1^2}, \forall k \geq 1$$



例3.4:求平稳AR(2)模型的协方差

- 平稳AR(2)模型的协方差函数递推公式为

$$\left\{ \begin{array}{l} Y_0 = \frac{1 - \phi_2}{(1 + \phi_2)(1 - \phi_1 - \phi_2)(1 + \phi_1 - \phi_2)} \sigma_\varepsilon^2 \\ Y_1 = \frac{\phi_1 Y_0}{1 - \phi_2} \\ Y_k = \phi_1 Y_{k-1} + \phi_2 Y_{k-2}, \quad k \geq 2 \end{array} \right.$$



自相关系数

- 自相关系数的定义

$$\rho_k = \frac{Y_k}{Y_0}$$

- 平稳AR(P)模型的自相关系数递推公式

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \cdots + \phi_p \rho_{k-p}$$



常用AR模型自相关系数递推公式

- AR(1)模型

$$\rho_k = \phi_1^k, k \geq 0$$

- AR(2)模型

$$\rho_k = \begin{cases} 1, & k = 0 \\ \frac{\phi_1}{1 - \phi_2} & k = 1 \\ \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} & k \geq 2 \end{cases}$$



AR模型自相关系数的性质

- 拖尾性

$$\rho(k) = \sum_{i=1}^p c_i \lambda_i^k \quad c_1, c_2, \dots, c_p \text{ 不能恒等于零}$$

- 呈复指数衰减

$$\rho(k) = \sum_{i=1}^p c_i \lambda_i^k \rightarrow 0$$



例3.5:考察如下AR模型的自相关图

$$(1) x_t = 0.8x_{t-1} + \varepsilon_t$$

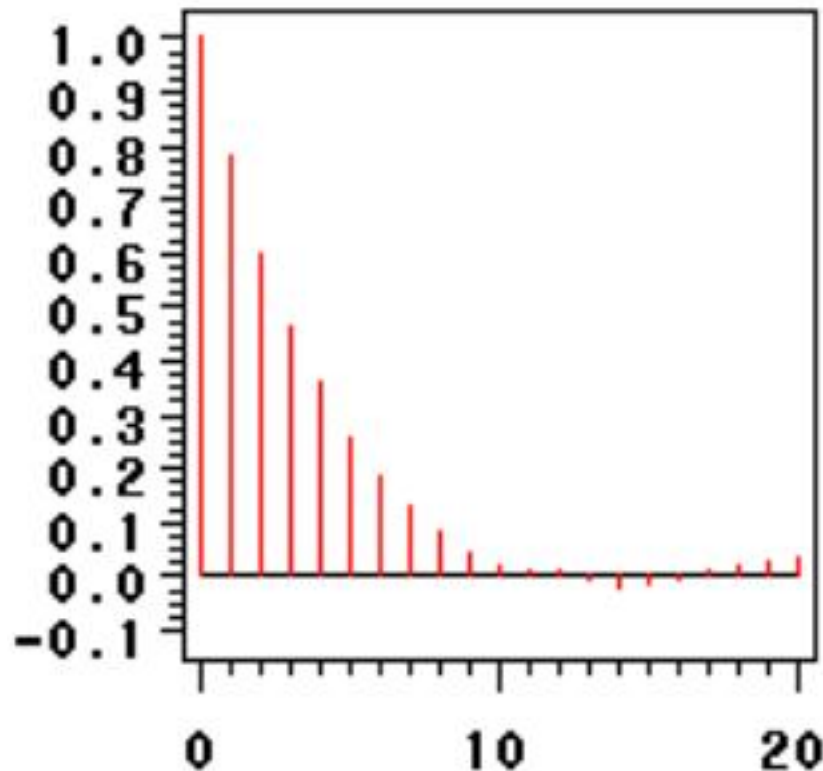
$$(2) x_t = -0.8x_{t-1} + \varepsilon_t$$

$$(3) x_t = x_{t-1} - 0.5x_{t-2} + \varepsilon_t$$

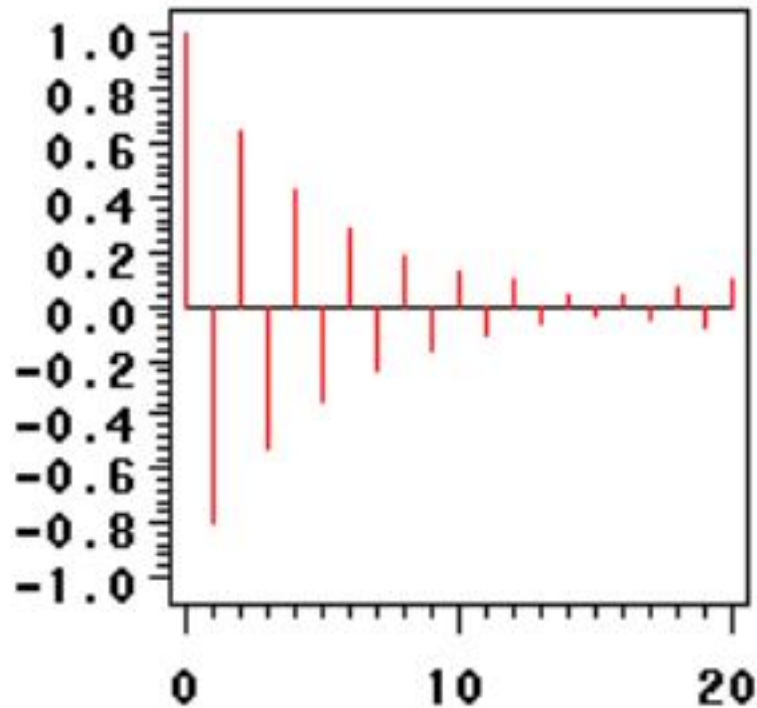
$$(4) x_t = -x_{t-1} - 0.5x_{t-2} + \varepsilon_t$$

例3.5— (1) $x_t = 0.8x_{t-1} + \varepsilon_t$

- 自相关系数按复指数单调收敛到零

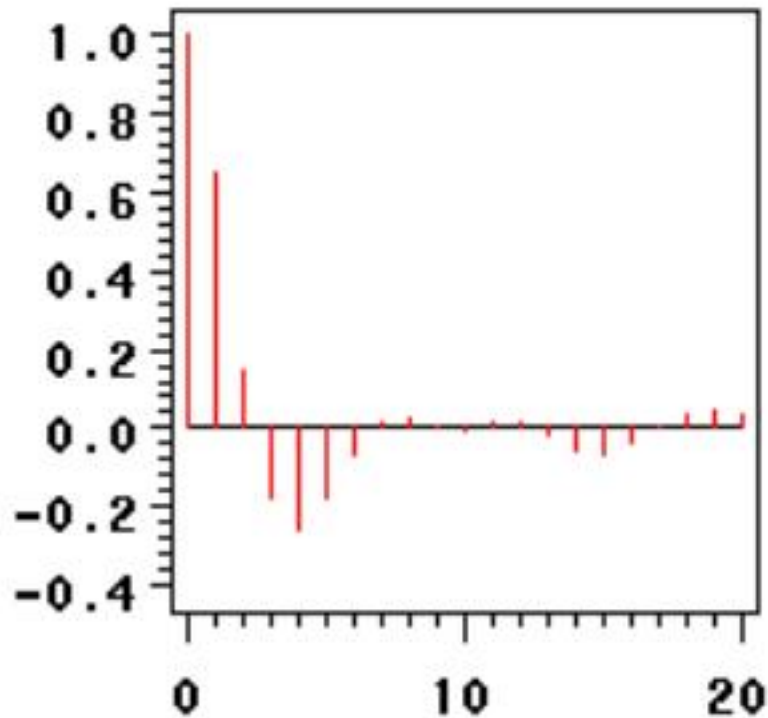


例3.5: — (2) $x_t = -0.8x_{t-1} + \varepsilon_t$



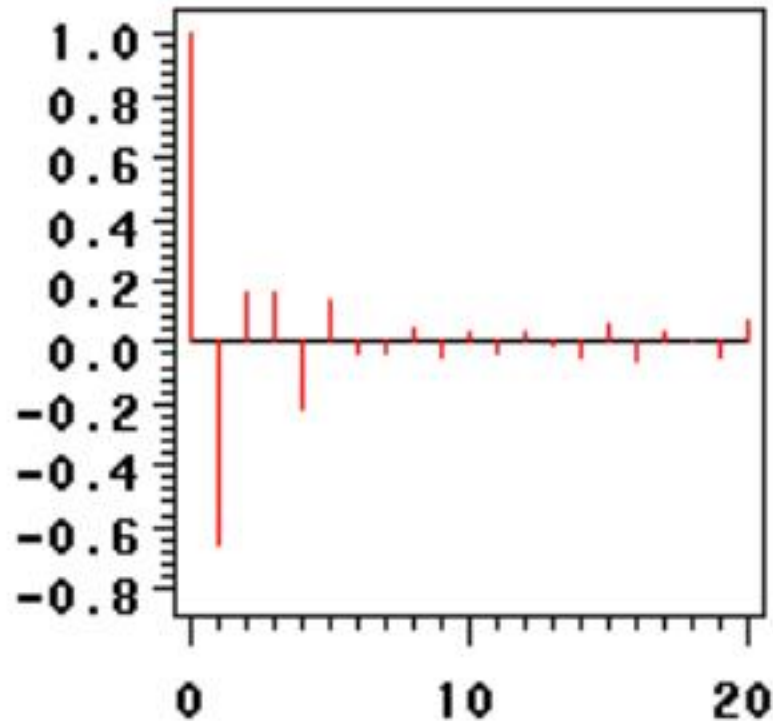
例3.5:— (3) $x_t = x_{t-1} - 0.5x_{t-2} + \varepsilon_t$

- 自相关系数呈现出“伪周期”性



例3.5:— (4) $x_t = -x_{t-1} - 0.5x_{t-2} + \varepsilon_t$

- 自相关系数不规则衰减





偏自相关系数

■ 定义

对于平稳AR(p)序列，所谓滞后k偏自相关系数就是指在给定中间k-1个随机变量 $x_{t-1}, x_{t-2}, \dots, x_{t-k+1}$ 的条件下，或者说，在剔除了中间k-1个随机变量的干扰之后， x_{t-k} 对 x_t 影响的相关度量。用数学语言描述就是

$$\rho_{x_t, x_{t-k} | x_{t-1}, \dots, x_{t-k+1}} = \frac{E[(x_t - \hat{E}x_t)(x_{t-k} - \hat{E}x_{t-k})]}{E[(x_{t-k} - \hat{E}x_{t-k})^2]}$$



偏自相关系数的截尾性

- AR(p)模型偏自相关系数P阶截尾

$$\phi_{kk} = 0, k > p$$



例3.5续:考察如下AR模型的偏自相关图

$$(1) x_t = 0.8x_{t-1} + \varepsilon_t$$

$$(2) x_t = -0.8x_{t-1} + \varepsilon_t$$

$$(3) x_t = x_{t-1} - 0.5x_{t-2} + \varepsilon_t$$

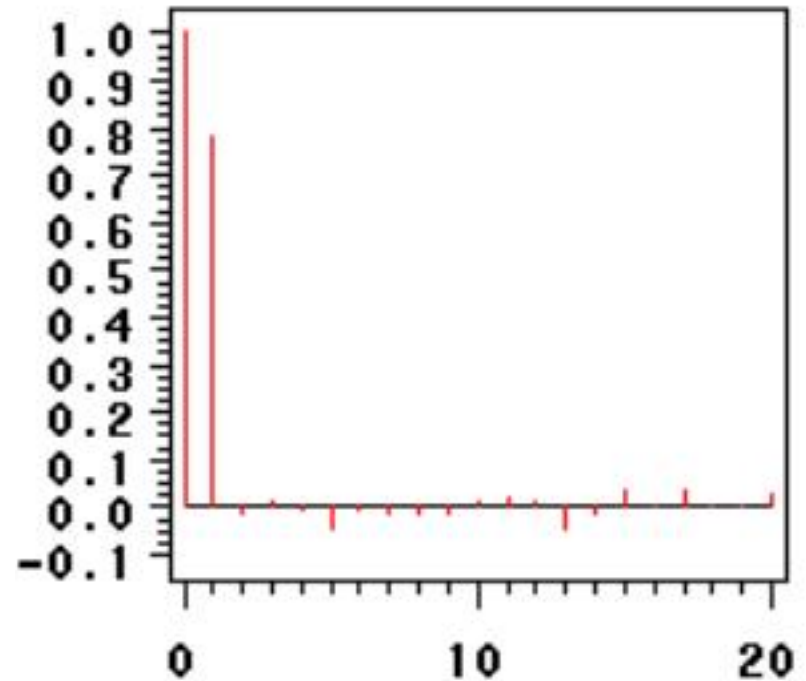
$$(4) x_t = -x_{t-1} - 0.5x_{t-2} + \varepsilon_t$$

例3.5— (1) $x_t = 0.8x_{t-1} + \varepsilon_t$

■ 理论偏自相关系数

■ 样本偏自相关图

$$\phi_{kk} = \begin{cases} 0.8 & , k = 1 \\ 0 & , k \geq 2 \end{cases}$$

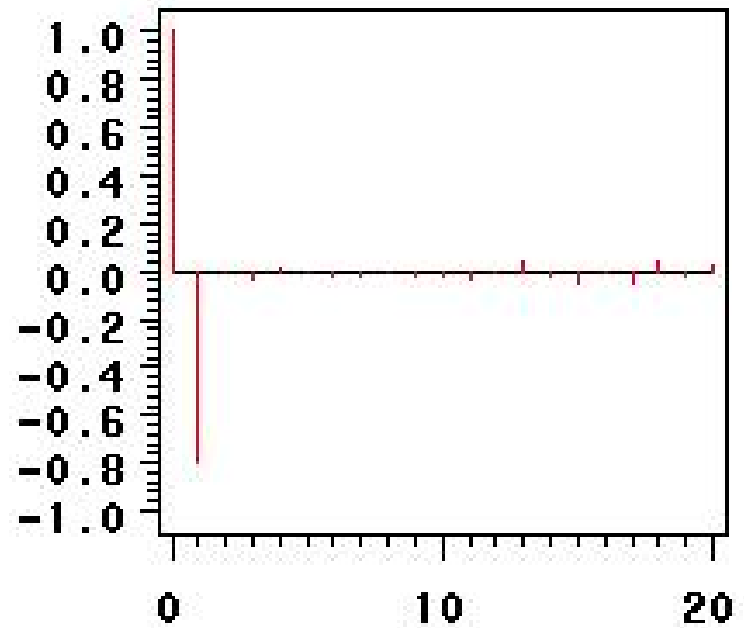


例3.5:— (2) $x_t = -0.8x_{t-1} + \varepsilon_t$

■ 理论偏自相关系数

■ 样本偏自相关图

$$\phi_{kk} = \begin{cases} -0.8 & , k = 1 \\ 0 & , k \geq 2 \end{cases}$$

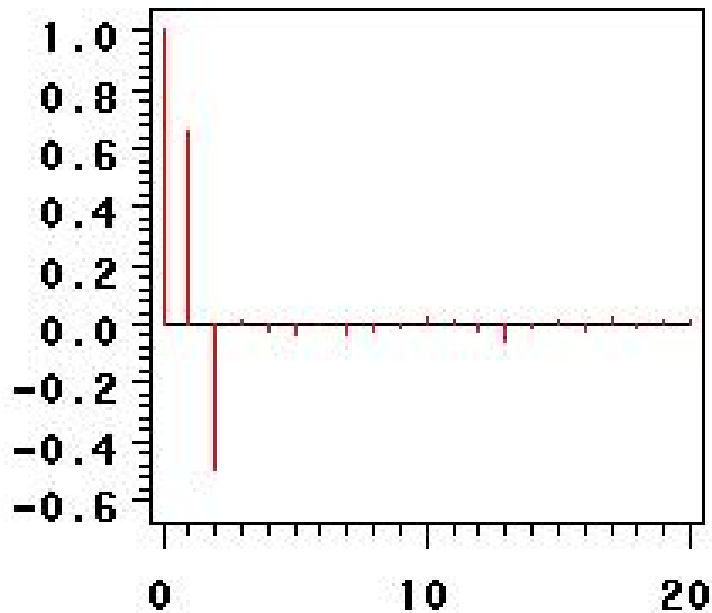


例3.5:— (3) $x_t = x_{t-1} - 0.5x_{t-2} + \varepsilon_t$

■ 理论偏自相关系数

■ 样本偏自相关图

$$\phi_{kk} = \begin{cases} \frac{2}{3} & , k = 1 \\ -0.5 & , k = 2 \\ 0 & , k \geq 3 \end{cases}$$

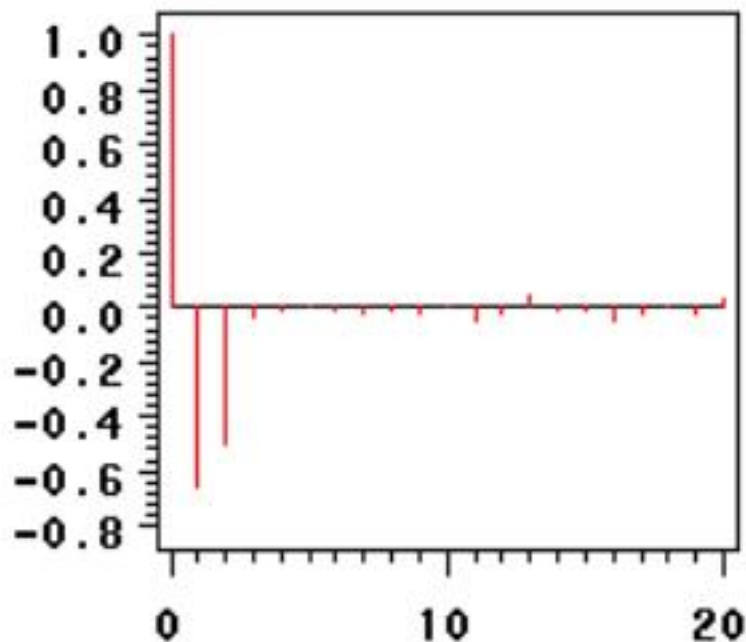


例3.5:— (4) $x_t = -x_{t-1} - 0.5x_{t-2} + \varepsilon_t$

■ 理论偏自相关系数

$$\phi_{kk} = \begin{cases} \frac{2}{3} & , k = 1 \\ -0.5 & , k = 2 \\ 0 & , k \geq 3 \end{cases}$$

■ 样本偏自相关系数图





MA模型的定义

- 具有如下结构的模型称为 q 阶自回归模型，简记为 $MA(q)$

$$\begin{cases} x_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \\ \theta_q \neq 0 \\ E(\varepsilon_t) = 0, \quad Var(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t \end{cases}$$

- 特别当 $\mu = 0$ 时，称为中心化 $MA(q)$ 模型



移动平均系数多项式

- 引进延迟算子，中心化MA模型又可以简记为

$$x_t = \dot{\cdot}(B)\varepsilon_t$$

- q 阶移动平均系数多项式

$$\dot{\cdot}(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$$



MA模型的统计性质

- 常数均值

$$\begin{aligned} E x_t &= E (\mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}) \\ &= \mu \end{aligned}$$

- 常数方差

$$\begin{aligned} Var (x_t) &= Var (\mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}) \\ &= (1 + \theta_1^2 + \cdots + \theta_q^2) \sigma_\varepsilon^2 \end{aligned}$$



MA模型的统计性质

- 自协方差函数 q 阶截尾
- 自相关系数 q 阶截尾

$$Y_k = \begin{cases} (1 + \theta_1^2 + \cdots + \theta_q^2) \sigma_\varepsilon^2, & k = 0 \\ (-\theta_k + \sum_{i=1}^{q-k} \theta_i \theta_{k+i}) \sigma_\varepsilon^2, & 1 \leq k \leq q \\ 0, & k > q \end{cases}$$

$$\rho_k = \begin{cases} 1, & k = 0 \\ \frac{-\theta_k + \sum_{i=1}^{q-k} \theta_i \theta_{k+i}}{1 + \theta_1^2 + \cdots + \theta_q^2}, & 1 \leq k \leq q \\ 0, & k > q \end{cases}$$



常用MA模型的自相关系数

■ MA(1)模型

$$\rho_k = \begin{cases} 1 & , k = 0 \\ \frac{-\theta_1}{1 + \theta_1^2} & , k = 1 \\ 0 & , k \geq 2 \end{cases}$$

■ MA(2)模型

$$\rho_k = \begin{cases} 1 & , k = 0 \\ \frac{-\theta_1 + \theta_1\theta_2}{1 + \theta_1^2 + \theta_2^2} & , k = 1 \\ \frac{-\theta_2}{1 + \theta_1^2 + \theta_2^2} & , k = 2 \\ 0 & , k \geq 3 \end{cases}$$



MA模型的统计性质

- 偏自相关系数拖尾

$$\phi_{kk} = (-\theta_1\varepsilon_{t-1} - \dots - \theta_q\varepsilon_{t-q})(-\theta_1\varepsilon_{t-k-1} - \dots - \theta_q\varepsilon_{t-k-q+1})$$

$\theta_1, \dots, \theta_q$ 不恒为零 $\Rightarrow \phi_{kk}$ 不会在有限阶之后恒为零



例3.6: 考察如下MA模型的相关性质

$$(1) x_t = \varepsilon_t - 2\varepsilon_{t-1}$$

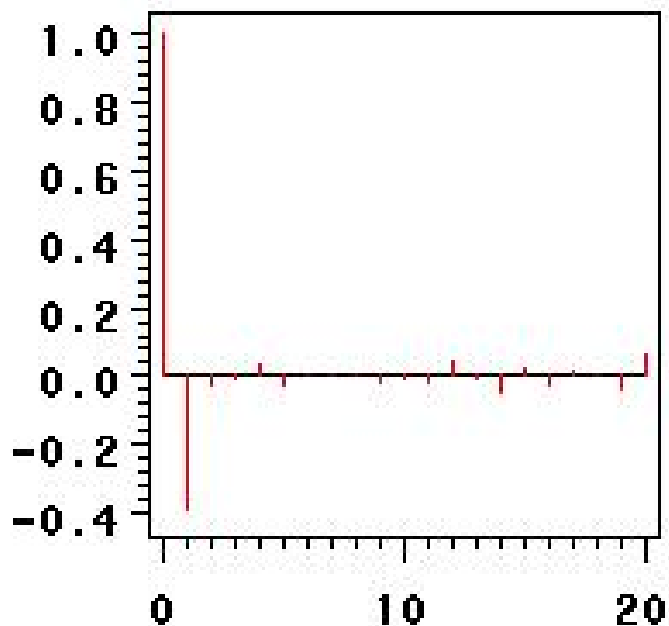
$$(2) x_t = \varepsilon_t - 0.5\varepsilon_{t-1}$$

$$(3) x_t = \varepsilon_t - \frac{4}{5}\varepsilon_{t-1} + \frac{16}{25}\varepsilon_{t-2}$$

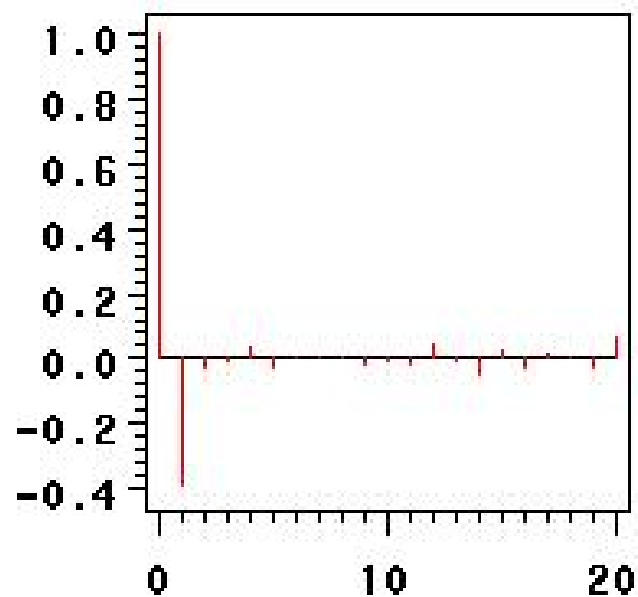
$$(4) x_t = \varepsilon_t - \frac{5}{4}\varepsilon_{t-1} + \frac{25}{16}\varepsilon_{t-2}$$

MA模型的自相关系数截尾

■ (1) $x_t = \varepsilon_t - 2\varepsilon_{t-1}$

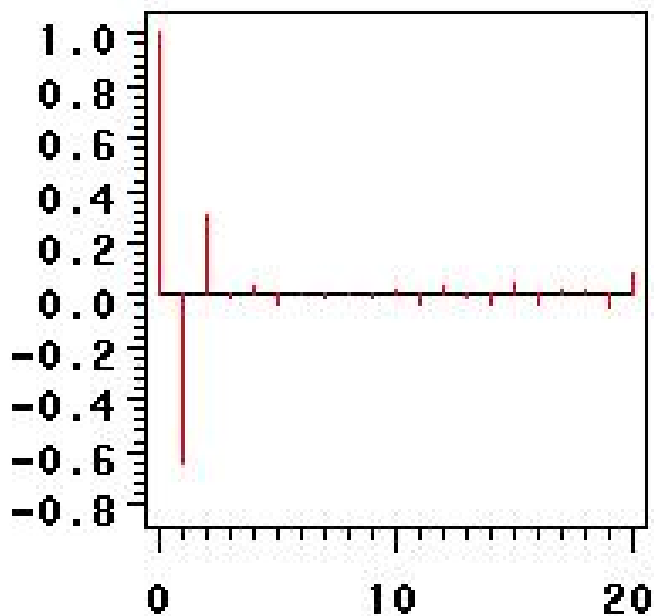


■ (2) $x_t = \varepsilon_t - 0.5\varepsilon_{t-1}$

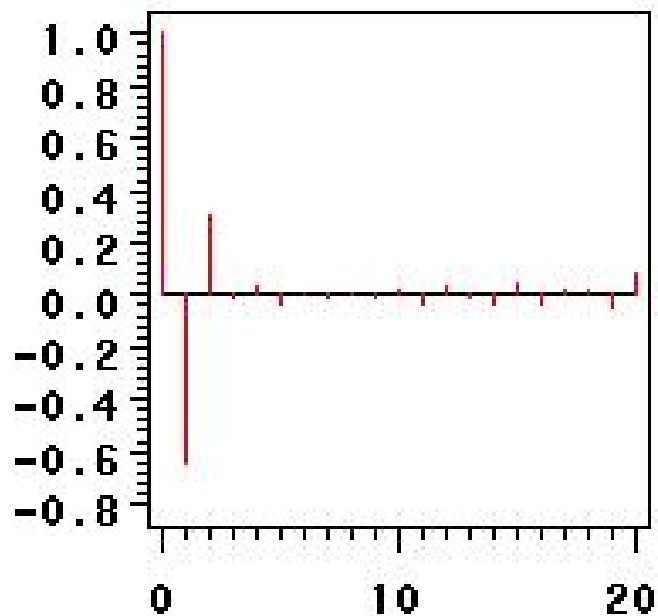


MA模型的自相关系数截尾

■ (3) $x_t = \varepsilon_t - \frac{4}{5}\varepsilon_{t-1} + \frac{16}{25}\varepsilon_{t-2}$



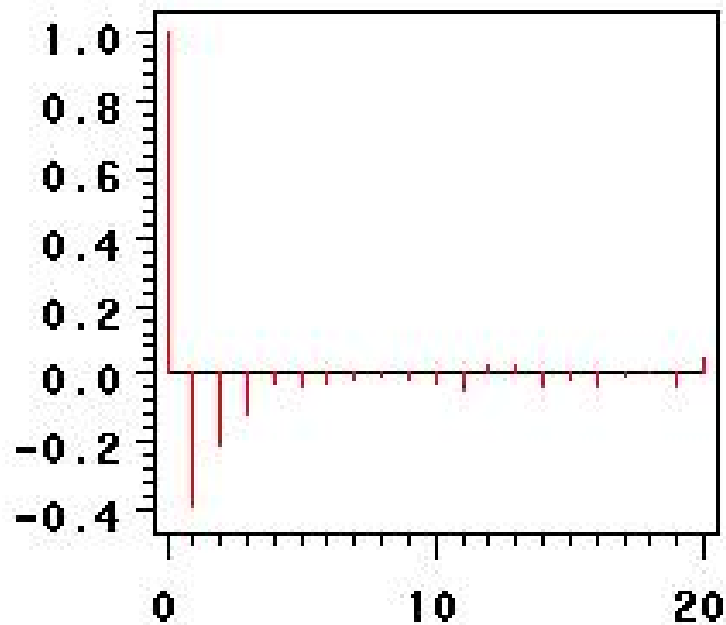
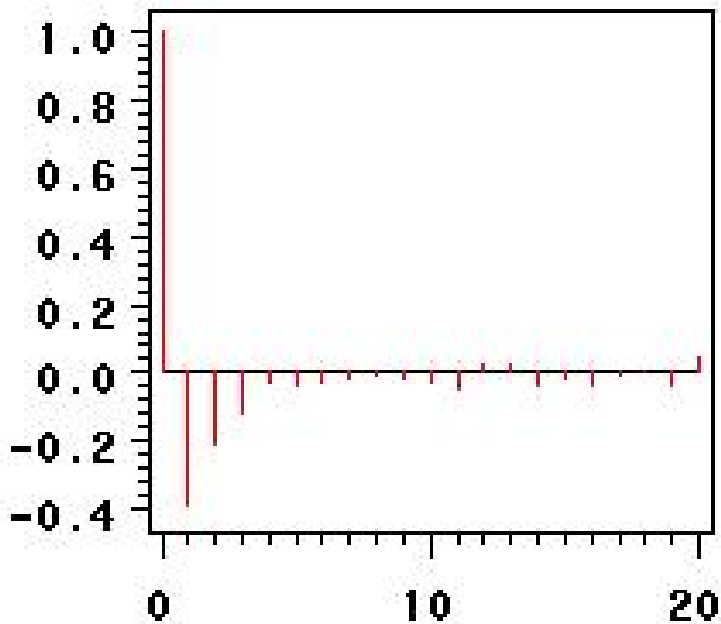
■ (4) $x_t = \varepsilon_t - \frac{5}{4}\varepsilon_{t-1} + \frac{25}{16}\varepsilon_{t-2}$



MA模型的偏自相关系数拖尾

■ (1) $x_t = \varepsilon_t - 2\varepsilon_{t-1}$

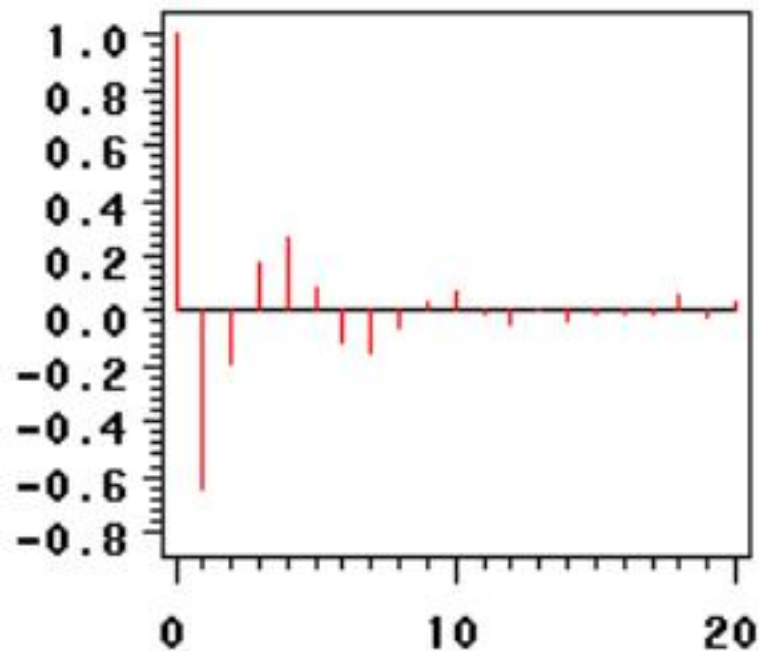
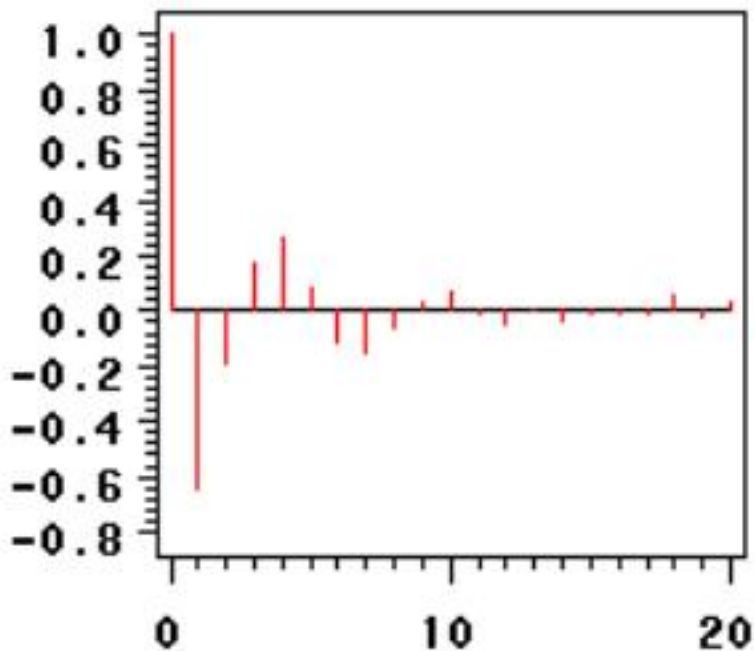
■ (2) $x_t = \varepsilon_t - 0.5\varepsilon_{t-1}$



MA模型的偏自相关系数拖尾

■ (3) $x_t = \varepsilon_t - \frac{4}{5}\varepsilon_{t-1} + \frac{16}{25}\varepsilon_{t-2}$

■ (4) $x_t = \varepsilon_t - \frac{5}{4}\varepsilon_{t-1} + \frac{25}{16}\varepsilon_{t-2}$





MA模型的可逆性

- MA模型自相关系数的不唯一性
 - 例3.6中不同的MA模型具有完全相同的自相关系数和偏自相关系数

$$(1)x_t = \varepsilon_t - 2\varepsilon_{t-1} \quad \Leftrightarrow \quad (2)x_t = \varepsilon_t - 0.5\varepsilon_{t-1}$$

$$(3)x_t = \varepsilon_t - \frac{4}{5}\varepsilon_{t-1} + \frac{16}{25}\varepsilon_{t-2} \quad \Leftrightarrow \quad (4)x_t = \varepsilon_t - \frac{5}{4}\varepsilon_{t-1} + \frac{25}{16}\varepsilon_{t-2}$$

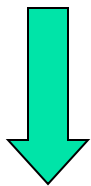


可逆的定义

- 可逆MA模型定义
 - 若一个MA模型能够表示称为收敛的AR模型形式，那么该MA模型称为可逆MA模型
- 可逆概念的重要性
 - 一个自相关系数列唯一对应一个可逆MA模型。

可逆MA(1)模型

- $x_t = \varepsilon_t - \theta\varepsilon_{t-1}$



$$\frac{x_t}{1 - \theta B} = \varepsilon_t$$

$|\theta| < 1$, 可逆

- $x_t = \varepsilon_t - \frac{1}{\theta}\varepsilon_{t-1}$



$$\frac{x_t}{1 - \frac{1}{\theta}B} = \varepsilon_t$$

$|\theta| > 1$, 可逆

$$\rho = \frac{-\theta}{1 + \theta^2}$$



MA模型的可逆条件

- MA(q)模型的可逆条件是:
 - MA(q)模型的特征根都在单位圆内

$$|\lambda_i| < 1$$

- 等价条件是移动平滑系数多项式的根都在单位圆外

$$\left| \frac{1}{\lambda_i} \right| > 1$$



逆函数的递推公式

- 原理

$$\begin{cases} x_t = \cdot(B)\varepsilon_t \\ \varepsilon_t = I(B)x_t \end{cases} \Rightarrow \cdot(B)I(B)x_t = x_t$$

- 方法

- 待定系数法

- 递推公式

$$\begin{cases} I_0 = 1 \\ I_j = \sum_{k=1}^j \theta'_k I_{j-k}, \quad j = 1, 2, \dots \end{cases}, \text{其中 } \theta'_k = \begin{cases} \theta_k, & k \leq q \\ 0, & k > q \end{cases}$$



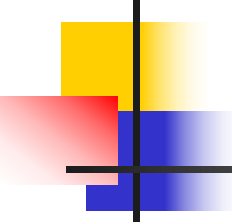
例3.6续:考察如下MA模型的可逆性

$$(1) x_t = \varepsilon_t - 2\varepsilon_{t-1}$$

$$(2) x_t = \varepsilon_t - 0.5\varepsilon_{t-1}$$

$$(3) x_t = \varepsilon_t - \frac{4}{5}\varepsilon_{t-1} + \frac{16}{25}\varepsilon_{t-2}$$

$$(4) x_t = \varepsilon_t - \frac{5}{4}\varepsilon_{t-1} + \frac{25}{16}\varepsilon_{t-2}$$



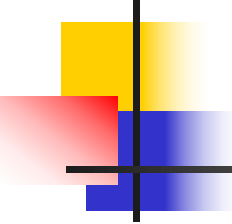
(1)—(2)

- $x_t = \varepsilon_t - 2\varepsilon_{t-1} \Rightarrow |\theta| = 2 > 1 \Rightarrow$ 不可逆
- $x_t = \varepsilon_t - 0.5\varepsilon_{t-1} \Rightarrow |\theta| = 0.5 < 1 \Rightarrow$ 可逆
- 逆函数

$$I_k = \begin{cases} 1 \\ 0.5^k, k \geq 1 \end{cases}$$

- 逆转形式

$$\varepsilon_t = \sum_{k=0}^{\infty} 0.5^k x_{t-k}$$



(3)—(4)

- $x_t = \varepsilon_t - \frac{4}{5}\varepsilon_{t-1} + \frac{16}{25}\varepsilon_{t-2} \Rightarrow \theta_2 < 1, \theta_2 \pm \theta_1 < 1 \Rightarrow$ 可逆

- $x_t = \varepsilon_t - \frac{5}{4}\varepsilon_{t-1} + \frac{25}{16}\varepsilon_{t-2} \Rightarrow \theta_2 = \frac{25}{16} > 1 \Rightarrow$ 不可逆

- 逆函数

$$I_k = \begin{cases} (-1)^n \theta_1^k, & k = 3n \text{ 或 } 3n + 1 \\ 0, & k = 3n + 2 \end{cases}, n = 0, 1, \dots$$

- 逆转形式

$$\varepsilon_t = \sum_{n=0}^{\infty} (-1)^n 0.8^{3n} x_{t-3n} + \sum_{n=0}^{\infty} (-1)^n 0.8^{3n+1} x_{t-3n-1}$$



ARMA模型的定义

- 具有如下结构的模型称为自回归移动平均模型，简记为 $ARMA(p, q)$

$$\begin{cases} x_t = \phi_0 + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q} \\ \phi_p \neq 0, \quad \theta_q \neq 0 \\ E(\varepsilon_t) = 0, \quad Var(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t \\ E x_s \varepsilon_t = 0, \forall s < t \end{cases}$$

- 特别当 $\phi_0 = 0$ 时，称为中心化 $ARMA(p, q)$ 模型



系数多项式

- 引进延迟算子，中心化 $ARMA(p, q)$ 模型又可以简记为

$$\Phi(B)x_t = \Psi(B)\varepsilon_t$$

- p 阶自回归系数多项式

$$\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$$

- q 阶移动平均系数多项式

$$\Psi(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$$



平稳条件与可逆条件

- ARMA(p,q)模型的平稳条件
 - P阶自回归系数多项式 $\phi(B) = 0$ 的根都在单位圆外
 - 即ARMA(p,q)模型的平稳性完全由其自回归部分的平稳性决定
- ARMA(p,q)模型的可逆条件
 - q阶移动平均系数多项式 $\psi(B) = 0$ 的根都在单位圆外
 - 即ARMA(p,q)模型的可逆性完全由其移动平滑部分的可逆性决定



传递形式与逆转形式

■ 传递形式

$$\begin{aligned}x_t &= \Phi^{-1}(B)\cdot\cdot(B)\varepsilon_t \\ &= \varepsilon_t + \sum_{j=1}^{\infty} G_j \varepsilon_{t-j}\end{aligned}$$

$$\begin{cases} G_0 = 1 \\ G_k = \sum_{j=1}^k \phi_j' G_{k-j} - \theta_j' \end{cases}, \quad k \geq 1$$

■ 逆转形式

$$\begin{aligned}\varepsilon_t &= \cdot\cdot^{-1}(B)\Phi(B)x_t \\ &= x_t + \sum_{j=1}^{\infty} I_j x_{t-j}\end{aligned}$$

$$\begin{cases} I_0 = 1 \\ I_k = \sum_{j=1}^k \theta_j' I_{k-j} - \phi_j' \end{cases}, \quad k \geq 1$$



ARMA(p,q)模型的统计性质

- 均值

$$Ex_t = \frac{\phi_0}{1 - \phi_1 - \dots - \phi_p}$$

- 协方差

$$\gamma(k) = \sigma_\varepsilon^2 \sum_{i=0}^{\infty} G_i G_{i+k}$$

- 自相关系数

$$\rho(k) = \frac{\gamma(k)}{\gamma(0)} = \frac{\sum_{j=0}^{\infty} G_j G_{j+k}}{\sum_{j=0}^{\infty} G_j^2}$$



ARMA模型的相关性

- 自相关系数拖尾
- 偏自相关系数拖尾

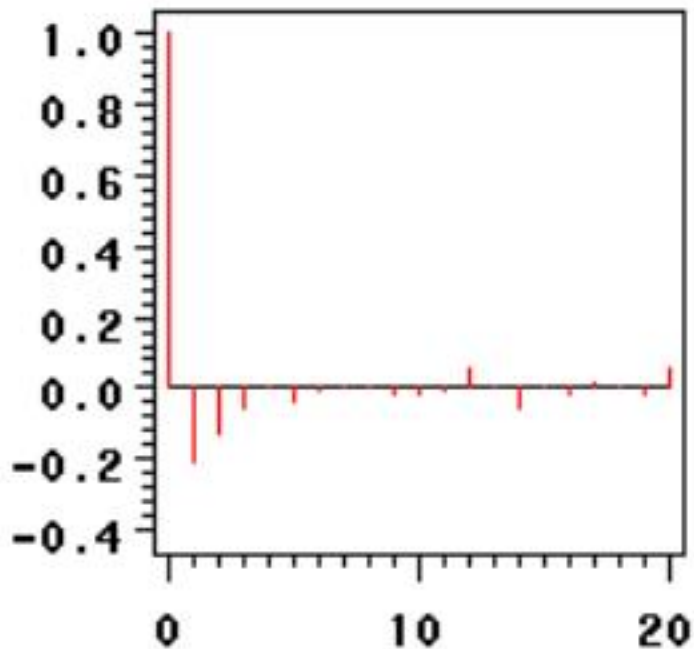


例3.7: 考察ARMA模型的相关性

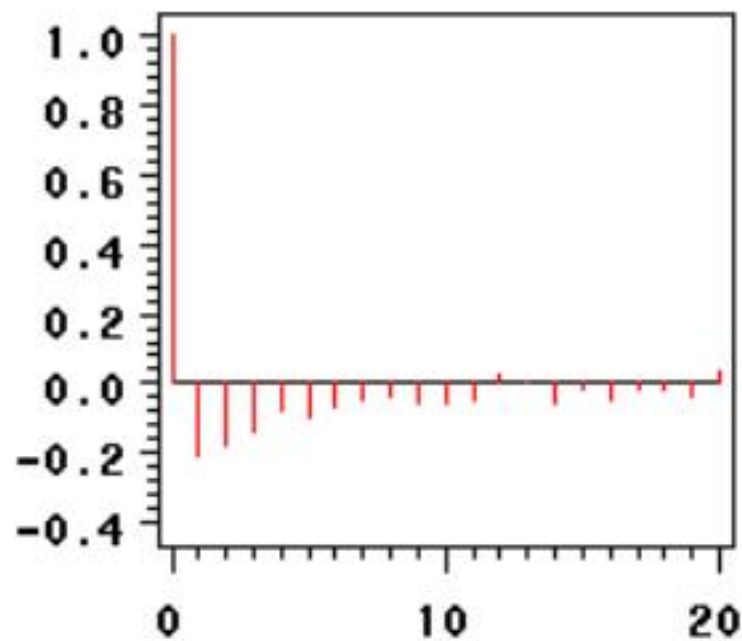
- 拟合模型ARMA(1,1): $x_t - 0.5x_{t-1} = \varepsilon_t - 0.8\varepsilon_t$
并直观地考察该模型自相关系数和偏自相关系数的性质。

自相关系数和偏自相关系数拖尾性

■ 样本自相关图



■ 样本偏自相关图





ARMA模型相关性特征

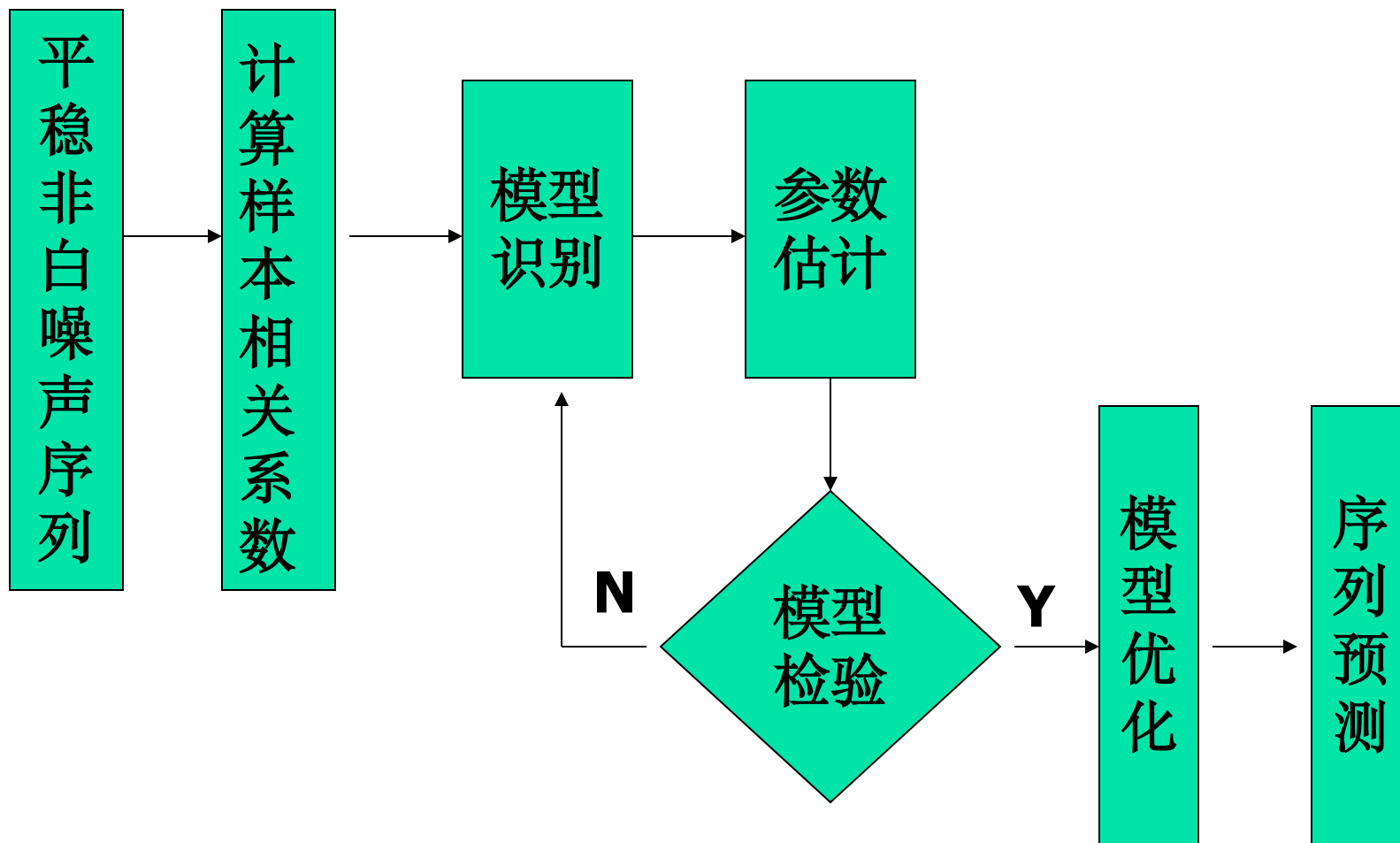
模型	自相关系数	偏自相关系数
AR(P)	拖尾	P阶截尾
MA(q)	q阶截尾	拖尾
ARMA(p,q)	拖尾	拖尾



3.3 平稳序列建模

- 建模步骤
- 模型识别
- 参数估计
- 模型检验
- 模型优化
- 序列预测

建模步骤





计算样本相关系数

■ 样本自相关系数

$$\hat{\rho}_k = \frac{\sum_{t=1}^{n-k} (x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^n (x_t - \bar{x})^2}$$

■ 样本偏自相关系数

$$\hat{\phi}_{kk} = \frac{\hat{D}_k}{\hat{D}}$$



模型识别

- 基本原则

$\hat{\alpha}$	$\hat{\phi}_{kk}$	选择模型
拖尾	P阶截尾	AR(P)
q阶截尾	拖尾	MA(q)
拖尾	拖尾	ARMA(p,q)



模型定阶的困难

- 因为由于样本的随机性，样本的相关系数不会呈现出理论截尾的完美情况，本应截尾的 $\hat{\rho}_k$ 仍会呈现出小值振荡的情况
- 由于平稳时间序列通常都具有短期相关性，随着延迟阶数 $k \rightarrow \infty$ 与 $\hat{\rho}_k$ 都会衰减至零值附近作小值波动
- ? 当 $\hat{\rho}_k$ 在延迟若干阶之后衰减为小值波动时，什么情况下该看作为相关系数截尾，什么情况下该看作为相关系数在延迟若干阶之后正常衰减到零值附近作拖尾波动呢？



样本相关系数的近似分布

- Barlett

$$\hat{\rho}_k \sim N\left(0, \frac{1}{n}\right), n \rightarrow \infty$$

- Quenouille

$$\hat{\phi}_{kk} \sim N\left(0, \frac{1}{n}\right), n \rightarrow \infty$$



模型定阶经验方法

- 95%的置信区间

$$\Pr\left(-\frac{2}{\sqrt{n}} \leq \hat{\rho}_k \leq \frac{2}{\sqrt{n}}\right) \geq 0.95$$

$$\Pr\left(-\frac{2}{\sqrt{n}} \leq \hat{\phi}_{kk} \leq \frac{2}{\sqrt{n}}\right) \geq 0.95$$

- 模型定阶的经验方法

- 如果样本(偏)自相关系数在最初的d阶明显大于两倍标准差范围，而后几乎95%的自相关系数都落在2倍标准差的范围以内，而且通常由非零自相关系数衰减为小值波动的过程非常突然。这时，通常视为(偏)自相关系数截尾。截尾阶数为d。



例2.5续

- 选择合适的模型ARMA拟合1950年——1998年北京市城乡居民定期储蓄比例序列。

序列自相关图

Autocorrelations

Lag	Covariance	Correlation	-1	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	8	9	1	
0	30.725523	1.00000																						
1	21.583411	0.70246																						
2	18.293557	0.59539																						
3	14.684303	0.47792																						
4	10.080193	0.32807																						
5	10.931717	0.35579																						
6	9.318240	0.30327																						
7	8.944975	0.29113																						
8	4.927541	0.16037																						
9	1.842114	0.05995																						
10	-1.151434	-.03747																						
11	-2.369343	-.07711																						
12	-1.130247	-.03679																						

“. " marks two standard errors

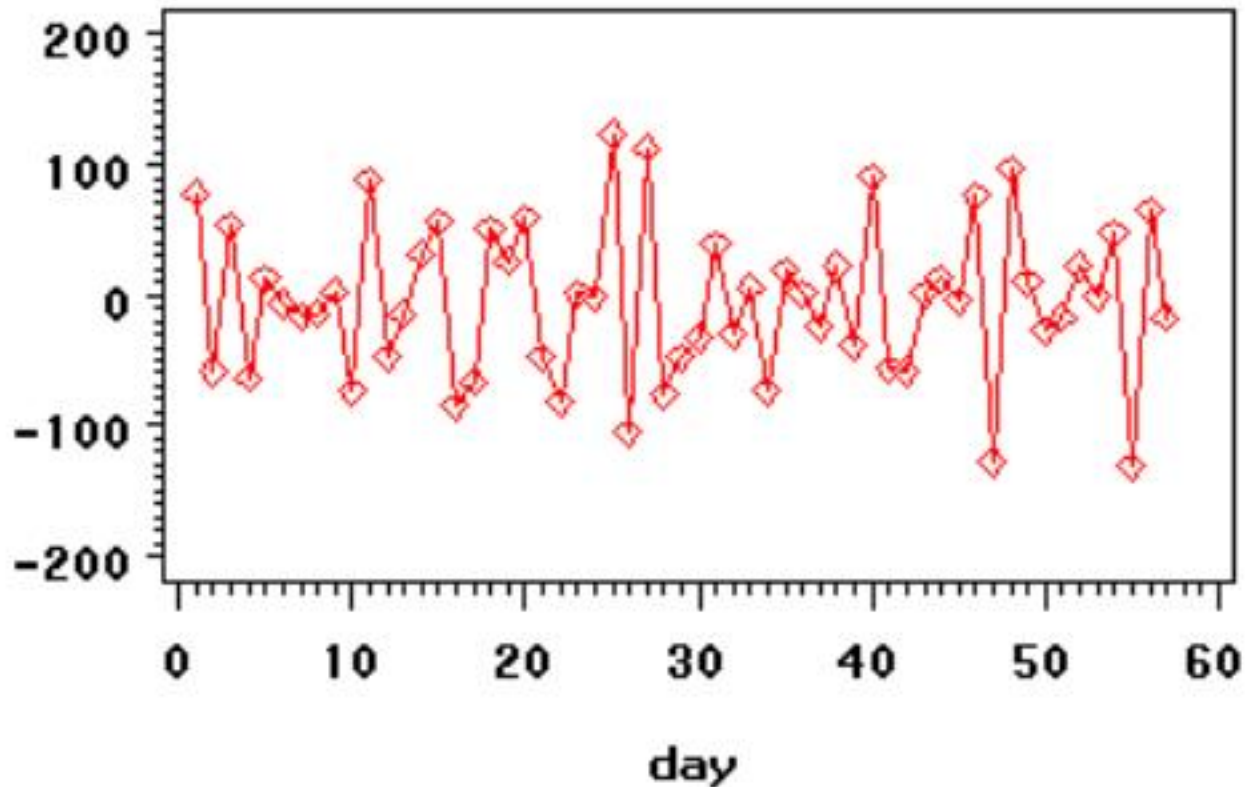


拟合模型识别

- 自相关图显示延迟3阶之后，自相关系数全部衰减到2倍标准差范围内波动，这表明序列明显地短期相关。但序列由显著非零的相关系数衰减为小值波动的过程相当连续，相当缓慢，该自相关系数可视为不截尾
- 偏自相关图显示除了延迟1阶的偏自相关系数显著大于2倍标准差之外，其它的偏自相关系数都在2倍标准差范围内作小值随机波动，而且由非零相关系数衰减为小值波动的过程非常突然，所以该偏自相关系数可视为一阶截尾
- 所以可以考虑拟合模型为AR(1)

例3.8

美国科罗拉多州某一加油站连续57天的OVERSHORT序列



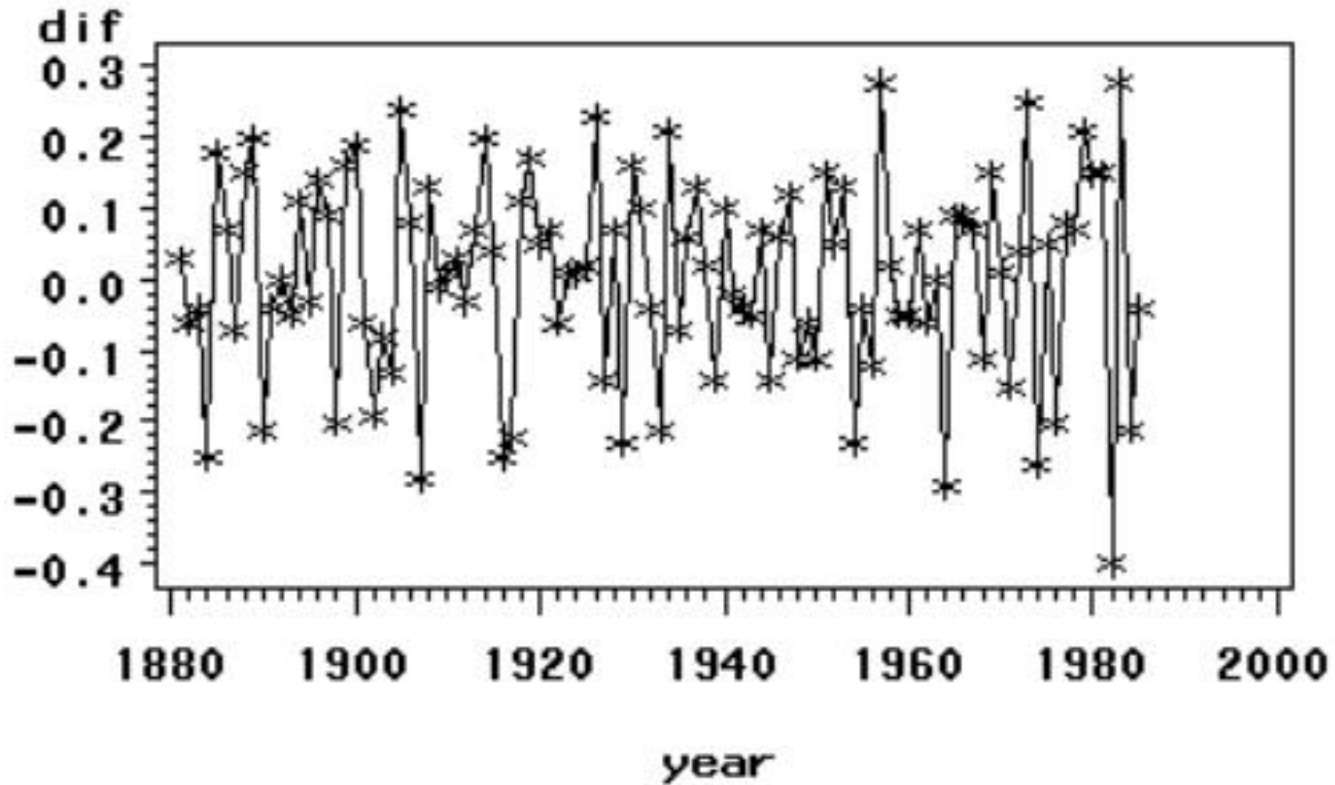


拟合模型识别

- 自相关图显示除了延迟1阶的自相关系数在2倍标准差范围之外，其它阶数的自相关系数都在2倍标准差范围内波动。根据这个特点可以判断该序列具有短期相关性，进一步确定序列平稳。同时，可以认为该序列自相关系数1阶截尾
- 偏自相关系数显示出典型非截尾的性质。
- 综合该序列自相关系数和偏自相关系数的性质，为拟合模型定阶为MA(1)

例3.9

- 1880-1985全球气表平均温度改变值差分序列



序列自相关图

Autocorrelations

Lag	Covariance	Correlation	-1	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	8	9	1
0	0.020061	1.00000											*****										
1	-0.0050732	-.25289										*****		.									
2	-0.0022778	-.11354										.***		.									
3	-0.0031786	-.15845										.***		.									
4	0.00089372	0.04455										.		*									
5	-0.0019930	-.09935										. **		.									
6	0.0043177	0.21523										.		****									
7	-0.0040567	-.20222										****		.									
8	0.0021045	0.10490										.		**									
9	-0.0021195	-.10566										. **		.									
10	0.0020664	0.10301										.		**									
11	-0.0002314	-.01154										.		.									
12	0.00086729	0.04323										.		*									

“. ” marks two standard errors



拟合模型识别

- 自相关系数显示出不截尾的性质
- 偏自相关系数也显示出不截尾的性质
- 综合该序列自相关系数和偏自相关系数的性质，可以尝试使用**ARMA(1,1)**模型拟合该序列



参数估计

- 待估参数
 - $p + q + 2$ 个未知参数

$$\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, \mu, \sigma_\varepsilon^2$$

- 常用估计方法
 - 矩估计
 - 极大似然估计
 - 最小二乘估计



矩估计

■ 原理

- 样本自相关系数估计总体自相关系数

$$\begin{cases} \rho_1(\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q) = \hat{\rho}_1 \\ \vdots \\ \rho_{p+q}(\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q) = \hat{\rho}_{p+q} \end{cases}$$

- 样本一阶均值估计总体均值，样本方差估计总体方差

$$\hat{\mu} = \bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

$$\hat{\sigma}_\varepsilon^2 = \frac{1 + \hat{\phi}_1^2 + \dots + \hat{\phi}_p^2}{1 + \hat{\theta}_1^2 + \dots + \hat{\theta}_q^2} \hat{\sigma}_x^2$$



例3.10:求AR(2)模型系数的矩估计

- AR(2)模型 $x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \varepsilon_t$
- Yule-Walker方程

$$\begin{cases} \rho_1 = \phi_1 + \phi_2 \rho_1 \\ \rho_2 = \phi_1 \rho_1 + \phi_2 \end{cases}$$

- 矩估计 (Yule-Walker方程的解)

$$\hat{\phi}_1 = \frac{1 - \hat{\rho}_2}{1 - \hat{\rho}_1^2} \hat{\rho}_1 \quad \hat{\phi}_2 = \frac{\hat{\rho}_2 - \hat{\rho}_1^2}{1 - \hat{\rho}_1^2}$$

例3.11:求MA(1)模型系数的矩估计

- MA(1)模型 $x_t = \varepsilon_t - \theta_1 \varepsilon_{t-1}$
- 方程

$$\begin{cases} \gamma_0 = (1 + \theta_1^2) \sigma_\varepsilon^2 \\ \gamma_1 = -\theta_1 \sigma_\varepsilon^2 \end{cases} \Rightarrow \rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{-\theta_1}{1 + \theta_1^2}$$

- 矩估计

$$\hat{\theta}_1 = \frac{-1 + \sqrt{1 - 4\hat{\rho}_1^2}}{2\hat{\rho}_1}$$

例3.12:求ARMA(1,1)模型系数的矩估计

- ARMA(1,1)模型 $x_t = \phi_1 x_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$

- 方程

$$\begin{cases} \rho_1 = \frac{Y_1}{Y_0} = \frac{(\phi_1 - \theta_1)(1 - \theta_1 \phi_1)}{1 + \theta_1^2 - 2\theta_1 \phi_1} \\ \rho_2 = \phi_1 \rho_1 \end{cases}$$

- 矩估计

$$\hat{\phi}_1 = \frac{\hat{\rho}_2}{\hat{\rho}_1}, \quad \hat{\theta}_1 = \begin{cases} \frac{c + \sqrt{c^2 - 4}}{2}, & c \leq -2 \\ \frac{c - \sqrt{c^2 - 4}}{2}, & c \geq 2 \end{cases}, \quad c = \frac{1 - \phi_1^2 - 2\hat{\rho}_2}{\phi_1 - \hat{\rho}_1}$$



对矩估计的评价

- 优点
 - 估计思想简单直观
 - 不需要假设总体分布
 - 计算量小（低阶模型场合）
- 缺点
 - 信息浪费严重
 - 只用到了 $p+q$ 个样本自相关系数信息，其他信息都被忽略
 - 估计精度差
- 通常矩估计方法被用作极大似然估计和最小二乘估计迭代计算的初始值



极大似然估计

■ 原理

- 在极大似然准则下，认为样本来自使该样本出现概率最大的总体。因此未知参数的极大似然估计就是使得似然函数（即联合密度函数）达到最大的参数值

$$L(\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k; x_1, \tilde{x}) = \max \{p(\tilde{x}); \beta_1, \beta_2, \dots, \beta_k\}$$



似然方程

$$\begin{cases} \frac{\partial}{\partial \sigma_\varepsilon^2} l(\tilde{\beta}; \tilde{x}) = \frac{n}{2\sigma_\varepsilon^2} - \frac{S(\tilde{\beta})}{2\sigma_\varepsilon^4} = 0 \\ \frac{\partial}{\partial \tilde{\beta}} l(\tilde{\beta}; \tilde{x}) = \frac{1}{2} \frac{\partial \ln|\Omega|}{\partial \tilde{\beta}} + \frac{1}{2\sigma_\varepsilon^2} \frac{\partial S(\tilde{\beta})}{\partial \tilde{\beta}} = 0 \end{cases}$$

- 由于 $s(\beta)$ 和 $|\Omega|$ 都不是 β 的显式表达式。因而似然方程组实际上是由 $p+q+1$ 个超越方程构成，通常需要经过复杂的迭代算法才能求出未知参数的极大似然估计值



对极大似然估计的评价

■ 优点

- 极大似然估计充分应用了每一个观察值所提供的信息，因而它的估计精度高
- 同时还具有估计的一致性、渐近正态性和渐近有效性等许多优良的统计性质

■ 缺点

- 需要假定总体分布



最小二乘估计

- 原理

- 使残差平方和达到最小的那组参数值即为最小二乘估计值

$$Q(\hat{\beta}) = \min Q(\tilde{\beta})$$

$$= \min \sum_{t=1}^n (x_t - \phi_1 x_{t-1} - \dots - \phi_p x_{t-p} - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q})^2$$



条件最小二乘估计

- 实际中最常用的参数估计方法
- 假设条件

$$x_t = 0, t > 0$$

- 残差平方和方程

$$Q(\tilde{\beta}) = \sum_{i=1}^n \varepsilon_t^2 = \sum_{i=1}^n \left[x_t - \sum_{i=1}^t \pi_i x_{t-1} \right]^2$$

- 解法
 - 迭代法



对最小二乘估计的评价

- 优点

- 最小二乘估计充分应用了每一个观察值所提供的信息，因而它的估计精度高
- 条件最小二乘估计方法使用率最高

- 缺点

- 需要假定总体分布



例2.5续

- 确定1950年——1998年北京市城乡居民定期储蓄比例序列拟合模型的口径
 - 拟合模型: AR(1)
 - 估计方法: 极大似然估计
 - 模型口径

$$x_t = 25.17 + 0.69x_{t-1} + \varepsilon_t$$

$$\text{Var}(\hat{\sigma}_\varepsilon^2) = 16.17$$



例3.8续

- 确定美国科罗拉多州某一加油站连续57天的**OVERSHORTS**序列拟合模型的口径
 - 拟合模型: $MA(1)$
 - 估计方法: 条件最小二乘估计
 - 模型口径

$$x_t = -4.40351 + (1 - 0.82303B)\varepsilon_t$$

$$Var(\hat{\sigma}_\varepsilon^2) = 2178.929$$



例3.9续

- 确定1880-1985全球气表平均温度改变值差分序列拟合模型的口径
 - 拟合模型: **ARMA(1,1)**
 - 估计方法: 条件最小二乘估计
 - 模型口径

$$x_t = 0.003 + 0.407x_{t-1} + \varepsilon_t - 0.9\varepsilon_{t-1}$$

$$\text{Var}(\hat{\sigma}_\varepsilon^2) = 0.016$$



模型检验

- 模型的显著性检验
 - 整个模型对信息的提取是否充分
- 参数的显著性检验
 - 模型结构是否最简



模型的显著性检验

- 目的
 - 检验模型的有效性（对信息的提取是否充分）
- 检验对象
 - 残差序列
- 判定原则
 - 一个好的拟合模型应该能够提取观察值序列中几乎所有的样本相关信息，即残差序列应该为白噪声序列
 - 反之，如果残差序列为非白噪声序列，那就意味着残差序列中还残留着相关信息未被提取，这就说明拟合模型不够有效



假设条件

- 原假设：残差序列为白噪声序列

$$H_0: \rho_1 = \rho_2 = \cdots = \rho_m = 0, \forall m \geq 1$$

- 备择假设：残差序列为非白噪声序列

$$H_1: \text{至少存在某个 } \rho_k \neq 0, \forall m \geq 1, k \leq m$$



检验统计量

- LB统计量

$$LB = n(n+2) \sum_{k=1}^m \left(\frac{\hat{\rho}_k^2}{n-k} \right) \sim \chi^2(m)$$



例2.5续

- 检验1950年——1998年北京市城乡居民定期储蓄比例序列拟合模型的显著性
- 残差白噪声序列检验结果

延迟阶数	LB统计量	P值	检验结论
6	5.83	0.3229	拟合模型 显著有效
12	10.28	0.5050	
18	11.38	0.8361	



参数显著性检验

- 目的

- 检验每一个未知参数是否显著非零。删除不显著参数使模型结构最精简

- 假设条件

$$H_0 : \beta_j = 0 \quad \leftrightarrow \quad H_1 : \beta_j \neq 0 \quad \forall 1 \leq j \leq m$$

- 检验统计量

$$T = \sqrt{n-m} \frac{\hat{\beta}_j - \beta_j}{\sqrt{a_{jj} Q(\tilde{\beta})}} \sim t(n-m)$$



例2.5续

- 检验1950年——1998年北京市城乡居民定期储蓄比例序列极大似然估计模型的参数是否显著
- 参数检验结果

检验参数	t统计量	P值	结论
均值	46.12	<0.0001	显著
ϕ_1	6.72	<0.0001	显著

例3.8续:对OVERSHORTS序列的拟合模型进行检验

- 残差白噪声检验

延迟阶数	LB统计量	P值	结论
6	3.15	0.6772	模型显著有效
12	9.05	0.6171	

- 参数显著性检验

检验参数	t统计量	P值	结论
均值	-3.75	<0.0004	显著
θ_1	10.60	<0.0001	显著

例3.9续:对1880-1985全球气表平均温度改变值差分序列拟合模型进行检验

- 残差白噪声检验

延迟阶数	LB统计量	P值	结论
6	5.28	0.2595	模型显著有效
12	10.30	0.4247	

- 参数显著性检验

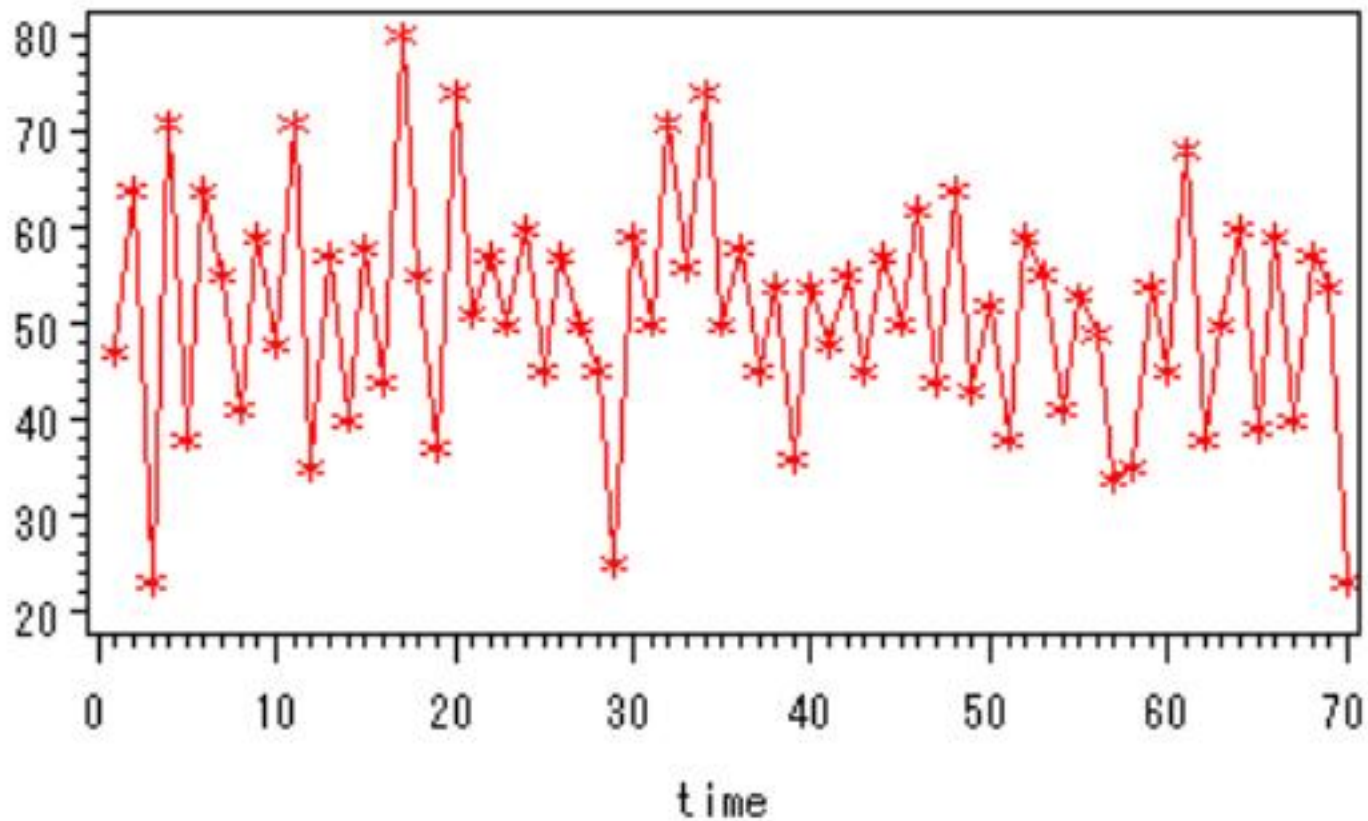
检验参数	t统计量	P值	结论
θ_1	16.34	<0.0001	显著
ϕ_1	3.5	0.0007	显著



模型优化

- 问题提出
 - 当一个拟合模型通过了检验，说明在一定的置信水平下，该模型能有效地拟合观察值序列的波动，但这种有效模型并不是唯一的。
- 优化的目的
 - 选择相对最优模型

例3.13:拟合某一化学序列



序列自相关图

Autocorrelations

Lag	Covariance	Correlation	-1	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	8	9	1
0	139.798	1.00000											*****										
1	-54.504114	-.38988									*****												
2	42.553609	0.30439										.											
3	-23.144178	-.16555												***									
4	9.886402	0.07072													*								
5	-13.565875	-.09704													**								
6	-6.578560	-.04706													*								
7	4.945082	0.03537														*							
8	-6.075359	-.04346														*							
9	-0.670493	-.00480																					
10	2.012128	0.01439																					
11	15.366178	0.10992													**								
12	-9.615079	-.06878													*								
13	20.694889	0.14803														***							
14	5.000367	0.03577													*								
15	-0.933542	-.00668																					
16	24.185609	0.17300														***							
17	-15.565443	-.11134													**								
18	2.791872	0.01997																					

"," marks two standard errors



拟合模型一

- 根据自相关系数2阶截尾，拟合MA(2)模型
- 参数估计

$$yield_t = 51.17301 + (1 - 0.32286B + 0.31009B^2)\varepsilon_t$$

- 模型检验
 - 模型显著有效
 - 三参数均显著



拟合模型二

- 根据偏自相关系数1阶截尾，拟合MA(1)模型
- 参数估计

$$yield_t = 51.26169 + \frac{\varepsilon_t}{1 + 0.42481B}$$

- 模型检验
 - 模型显著有效
 - 两参数均显著



问题

- 同一个序列可以构造两个拟合模型，两个模型都显著有效，那么到底该选择哪个模型用于统计推断呢？
- 解决办法
 - 确定适当的比较准则，构造适当的统计量，确定相对最优



AIC准则

- 最小信息量准则 (An Information Criterion)
- 指导思想
 - 似然函数值越大越好
 - 未知参数的个数越少越好
- AIC统计量

$$AIC = n \ln(\hat{\sigma}_\varepsilon^2) + 2(\text{未知参数个数})$$



SBC准则

- AIC准则的缺陷

- 在样本容量趋于无穷大时，由AIC准则选择的模型不收敛于真实模型，它通常比真实模型所含的未知参数个数要多

- SBC统计量

$$SBC = n \ln(\hat{\sigma}_\varepsilon^2) + \ln(n)(\text{未知参数})$$



例3.13续

- 用**AIC**准则和**SBC**准则评判例3.13中两个拟合模型的相对优劣

模型	AIC	SBC
MA(2)	536.4556	543.2011
AR(1)	535.7896	540.2866

- 结果
 - AR(1)优于MA(2)



序列预测

- 线性预测函数

$$x_t = \sum_{i=0}^{\infty} C_i x_{t-1-i}$$

- 预测方差最小原则

$$\text{Var}_{\hat{x}_t(l)} [e_t(l)] = \min \{ \text{Var} [e_t(l)] \}$$



序列分解

$$\begin{aligned}x_{t+l} &= (\varepsilon_{t+l} + G_1\varepsilon_{t+l-1} + \cdots + G_{l-1}\varepsilon_{t+1}) + (G_l\varepsilon_t + G_{l+1}\varepsilon_{t-1} + \cdots) \\ &= \qquad \qquad e_t(l) \qquad \qquad \qquad + \qquad \qquad \hat{x}_t(l) \\ &\qquad \qquad \qquad \downarrow \qquad \qquad \qquad \qquad \qquad \qquad \downarrow \\ &\qquad \qquad \text{预测误差} \qquad \qquad \qquad \qquad \qquad \qquad \text{预测值}\end{aligned}$$

$$E(x_{t+l} | x_t, x_{t-1}, \cdots) = \hat{x}(l)$$

$$\text{Var}(x_{t+l} | x_t, x_{t-1}, \cdots) = \text{Var}[e_t(l)]$$



误差分析

- 估计误差

$$e_t(l) = \varepsilon_{t+l} + G_1 \varepsilon_{t+l-1} + \cdots + G_{l-1} \varepsilon_{t+1}$$

- 期望

$$E[e_t(l)] = 0$$

- 方差

$$\text{Var}[e_t(l)] = \sum_{i=0}^{l-1} G_i^2 \cdot \sigma_\varepsilon^2$$



AR(p)序列的预测

- 预测值

$$\hat{x}_t(l) = \phi_1 \hat{x}_t(l-1) + \cdots + \phi_p \hat{x}_t(l-p)$$

- 预测方差

$$\text{Var}[e_t(l)] = (1 + G_1^2 + \cdots + G_{l-1}^2) \sigma_\varepsilon^2$$

- 95% 置信区间

$$\left(\hat{x}_t(l) \mp z_{1-\frac{\alpha}{2}} \cdot \left(1 + G_1^2 + \cdots + G_{l-1}^2\right)^{\frac{1}{2}} \cdot \sigma_\varepsilon \right)$$



例3.14

- 已知某超市月销售额近似服从**AR(2)**模型
(单位: 万元/每月)

$$x_t = 10 + 0.6x_{t-1} + 0.3x_{t-2} + \varepsilon_t, \varepsilon_t \sim N(0, 36)$$

- 今年第一季度该超市月销售额分别为:
101, 96, 97.2万元
- 请确定该超市第二季度每月销售额的**95%**的置信区间



例3.14解： 预测值计算

- 四月份

$$\hat{x}_3(1) = 10 + 0.6x_3 + 0.3x_2 = 97.12$$

- 五月份

$$\hat{x}_3(2) = 10 + 0.6\hat{x}_3(1) + 0.3x_3 = 97.432$$

- 六月份

$$\hat{x}_3(3) = 10 + 0.6\hat{x}_3(2) + 0.3\hat{x}_3(1) = 97.5952$$



例3.14解： 预测方差的计算

- GREEN函数

$$G_0 = 1$$

$$G_1 = \phi_1 G_0 = 0.6$$

$$G_2 = \phi_1 G_1 + \phi_2 G_0 = 0.36 + 0.3 = 0.66$$

- 方差

$$\text{Var}[e_3(1)] = G_0^2 \sigma_\varepsilon^2 = 36$$

$$\text{Var}[e_3(2)] = (G_0^2 + G_1^2) \sigma_\varepsilon^2 = 48.96$$

$$\text{Var}[e_3(3)] = (G_0^2 + G_1^2 + G_2^2) \sigma_\varepsilon^2 = 64.6416$$



例3.14解： 置信区间

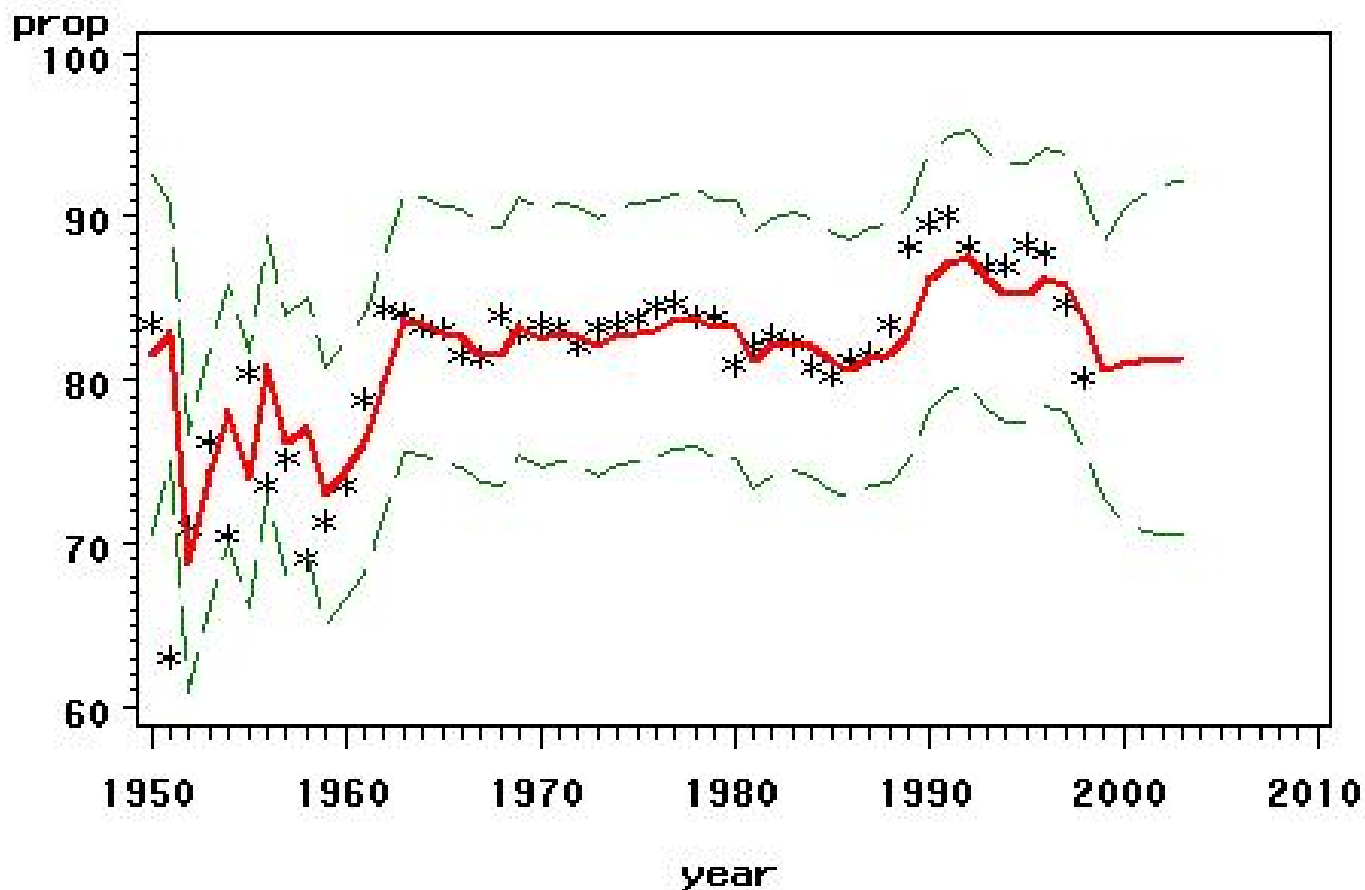
- 公式

$$(\hat{x}_3(l) - 1.96\sqrt{\text{Var}[e_3(l)]}, \hat{x}_3(l) + 1.96\sqrt{\text{Var}[e_3(l)]})$$

- 估计结果

预测时期	95% 置信区间
四月份	(85.36, 108.88)
五月份	(83.72, 111.15)
六月份	(81.84, 113.35)

例2.5:北京市城乡居民定期储蓄比例序列 拟合与预测图





MA(q)序列的预测

- 预测值

$$\hat{x}_t(l) = \begin{cases} \mu - \sum_{i=l}^q \theta_i \varepsilon_{t+l-i} & , l \leq q \\ \mu & , l > q \end{cases}$$

- 预测方差

$$\text{Var}[e_t(l)] = \begin{cases} (1 + \theta_1^2 + \cdots + \theta_{l-1}^2) \sigma_\varepsilon^2 & , l \leq q \\ (1 + \theta_1^2 + \cdots + \theta_q^2) \sigma_\varepsilon^2 & , l > q \end{cases}$$



例3.15

- 已知某地区每年常驻人口数量近似服从MA(3)模型 (单位: 万) :

$$x_t = 100 + \varepsilon_t - 0.8\varepsilon_{t-1} + 0.6\varepsilon_{t-2} - 0.2\varepsilon_{t-3}$$

- 最近3年的常驻人口数量及一步预测数量如下:

年份	统计人数	预测人数
2002	104	110
2003	108	100
2004	105	109

- 预测未来5年该地区常住人口的95%置信区间



例3.15解： 随机扰动项的计算

$$\varepsilon_{t-2} = x_{2002} - \hat{x}_{2001}(1) = 104 - 110 = -6$$

$$\varepsilon_{t-1} = x_{2003} - \hat{x}_{2002}(1) = 108 - 100 = 8$$

$$\varepsilon_t = x_{2004} - \hat{x}_{2003}(1) = 105 - 109 = -4$$



例3.15解： 估计值的计算

$$\hat{x}_t(1) = 100 - 0.8\varepsilon_t + 0.6\varepsilon_{t-1} - 0.2\varepsilon_{t-2} = 109.2$$

$$\hat{x}_t(2) = 100 + 0.6\varepsilon_t - 0.2\varepsilon_{t-1} = 96$$

$$\hat{x}_t(3) = 100 - 0.2\varepsilon_t = 100.8$$

$$\hat{x}_t(4) = 100$$

$$\hat{x}_t(5) = 100$$



例3.15解： 预测方差的计算

$$\text{Var}[e_t(1)] = \sigma_\varepsilon^2 = 25$$

$$\text{Var}[e_t(2)] = (1 + \theta_1^2)\sigma_\varepsilon^2 = 41$$

$$\text{Var}[e_t(3)] = (1 + \theta_1^2 + \theta_2^2)\sigma_\varepsilon^2 = 50$$

$$\text{Var}[e_t(4)] = (1 + \theta_1^2 + \theta_2^2 + \theta_3^2)\sigma_\varepsilon^2 = 51$$

$$\text{Var}[e_t(5)] = (1 + \theta_1^2 + \theta_2^2 + \theta_3^2)\sigma_\varepsilon^2 = 51$$



例3.15解： 置信区间的计算

预测年份	95% 置信区间
2005	(99, 119)
2006	(83, 109)
2007	(87, 115)
2008	(86, 114)
2009	(86, 114)



ARMA(p,q)序列预测

- 预测值

$$\hat{x}_t(k) = \begin{cases} \hat{x}_t(k) & , k \geq 1 \\ x_{t+k} & , k \leq 0 \end{cases}$$

- 预测方差

$$\text{Var}[e_t(l)] = (G_0^1 + G_1^2 + \cdots + G_{l-1}^2) \sigma_\varepsilon^2$$



例3.16

- 已知模型为:

$$x_t = 0.8x_{t-1} + \varepsilon_t - 0.6\varepsilon_{t-1} \quad \sigma_\varepsilon^2 = 0.0025$$

- 且 $x_{100} = 0.3$ $\varepsilon_{100} = 0.01$

- 预测未来3期序列值的95%的置信区间。



例3.16解：估计值的计算

$$\hat{x}_{100}(1) = 0.8x_{100} - 0.6\varepsilon_{100} = 0.234$$

$$\hat{x}_{100}(2) = 0.8\hat{x}_{100}(1) = 0.1872$$

$$\hat{x}_{100}(3) = 0.8\hat{x}_{100}(2) = 0.14976$$



例3.16解： 预测方差的计算

- Green函数

$$G_0 = 1$$

$$G_1 = \phi_1 G_0 - \theta_1 = 0.2$$

$$G_2 = \phi_1 G_1 = 0.16$$

- 方差

$$\text{Var}[e_{100}(1)] = G_0^2 \sigma_\varepsilon^2 = 0.0025$$

$$\text{Var}[e_{100}(2)] = (G_0^2 + G_1^2) \sigma_\varepsilon^2 = 0.0026$$

$$\text{Var}[e_{100}(3)] = (G_0^2 + G_1^2 + G_2^2) \sigma_\varepsilon^2 = 0.002664$$



例3.16解： 置信区间的计算

时期	95% 置信区间
101	(0.136, 0.332)
102	(0.087, 0.287)
103	(-0.049, 0.251)



修正预测

■ 定义

- 所谓的修正预测就是研究如何利用新的信息去获得精度更高的预测值

■ 方法

- 在新的信息量比较大时——把新信息加入到旧的信息中，重新拟合模型
- 在新的信息量很小时——不重新拟合模型，只是将新的信息加入以修正预测值，提高预测精度



修正预测原理

- 在旧信息的基础上, x_{t+l} 的预测值为

$$\hat{x}_t(l) = G_l \varepsilon_t + G_1 \varepsilon_{t-1} + \dots$$

- 假设新获得一个观察值 x_{t+1} , 则

- x_{t+l} 的修正预测值为

$$\hat{x}_{t+1}(l-1) = G_{l-1} \varepsilon_{t+1} + G_l \varepsilon_t + G_{l+1} \varepsilon_{t-1} + \dots = G_{l-1} \varepsilon_{t+1} + \hat{x}_t(l)$$

- 修正预测误差为

$$e_{t+1}(l-1) = G_0 \varepsilon_{t+l} + \dots + G_{l-2} \varepsilon_{t+2}$$

- 预测方差为

$$\text{Var}[e_{t+1}(l-1)] = (G_0^2 + \dots + G_{l-2}^2) \sigma_\varepsilon^2$$



一般情况

- 假设新获得 p 个观察值 x_{t+1}, \dots, x_{t+p} , 则

- x_{t+l} 的修正预测值为

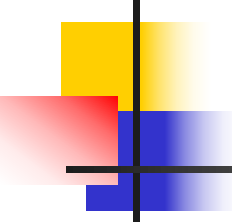
$$\hat{x}_{t+p}(l-p) = G_{l-p}\varepsilon_{t+p} + \dots + G_{l-1}\varepsilon_{t+1} + \hat{x}_t(l)$$

- 修正预测误差为

$$e_{t+p}(l-p) = G_0\varepsilon_{t+l} + \dots + G_{l-p-1}\varepsilon_{t+p+1}$$

- 预测方差为

$$\text{Var}[e_{t+p}(l-p)] = (G_0^2 + \dots + G_{l-p-1}^2)\sigma_\varepsilon^2$$



例3.14续:假如四月份的真实销售额为100万元,
求二季度后两个月销售额的修正预测值

- 计算四月份的预测误差

$$\varepsilon_4 = x_4 - \hat{x}_3(1) = 100 - 97.12 = 2.88$$

- 计算修正预测值

$$\hat{x}_4(1) = G_1 \varepsilon_4 + \hat{x}_3(2) = 99.16$$

$$\hat{x}_4(2) = G_2 \varepsilon_4 + \hat{x}_3(3) = 99.50$$

- 计算修正方差

$$\text{Var}[e_4(1)] = \text{Var}[e_3(1)] = G_0^2 \sigma_\varepsilon^2 = 36$$

$$\text{Var}[e_4(2)] = \text{Var}[e_3(2)] = (G_0^2 + G_1^2) \sigma_\varepsilon^2 = 48.96$$



修正置信区间

预测时期	修正前置信区间	修正后置信区间
四月份	(85.36, 108.88)	
五月份	(83.72, 111.15)	(87.40, 110.92)
六月份	(81.84, 113.35)	(85.79, 113.21)