

第二章 平稳时间序列模型及其性质



学习目标与要求

- 了解线性差分方程及其解的结构
- 掌握自回归模型的有关概念和性质
- 掌握移动平均模型的有关概念和性质
- 掌握自回归移动平均模型的有关概念和性质



本章结构

1. 差分方程和滞后算子

2. 自回归模型的概念和性质

3. 移动平均模型的概念和性质

4. 自回归移动平均模型的概念和性质

差分运算与滞后算子

- 绪 论

如果序列值被识别为平稳的非白噪声序列, 那么该序列就蕴含着一定的相关信息. 从统计角度看, 我们就可以设法提取该序列中蕴含的有用信息, 并建立适当的统计模型来拟合该序列. 目前, 最常用的平稳序列拟合模型是自回归模型 (AR 模型)、移动平均模型 (MA 模型) 和自回归移动平均模型 (ARMA 模型). 这三类模型都属于有限参数线性模型, 它们与线性差分方程有着密切的联系, 模型的性质取决于差分方程根的性质. 因此, 在介绍这三类模型之前, 我们先简要学习线性差分方程的求解和滞后算子.



差分运算与滞后算子

■ 1. p 阶差分运算

对于一个观察值序列 $\{x_t\}$ 来讲, 相邻两时刻序列值之差称为 1 阶差分, 下面引入1 阶向后差分, 记为 ∇x_t , 即

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

对 1 阶向后差分后所得序列 $\{\nabla x_t\}$ 再进行 1 阶向后差分运算就得到2 阶向后差分, 记为 $\nabla^2 x_t$, 于是

$$\nabla^2 x_t = \nabla x_t - \nabla x_{t-1}$$

类似地, 对 $p - 1$ 阶向后差分后所得序列 $\{\nabla^{p-1} x_t\}$ 再进行 1 阶向后差分就得到 p 阶向后差分, 记为 $\nabla^p x_t$,

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



差分运算与滞后算子

■ 2. m 步差分运算

一般地, 称观察值序列 $\{x_t\}$ 相距 m 个时刻的值之差为 m 步差分, 记作: $\nabla_m x_t$, 即

$$\nabla_m x_t = x_t - x_{t-m}$$

■ 3. 滞后算子

如果算子 B 满足

$$Bx_t = x_{t-1},$$

那么称 B 为关于时间 t 的 1 步滞后算子 (lag operator), 简称滞后算子.



差分运算与滞后算子

容易证明滞后算子有如下性质:

$$(1) B^0 x_t = x_t;$$

$$(2) B^k x_t = x_{t-k}, k = 1, 2, \dots;$$

(3) 若 $\{x_t\}$ 和 $\{y_t\}$ 为任意两个序列, 且 c_1 和 c_2 为任意常数, 则有

$$B(c_1 x_t \pm c_2 y_t) = c_1 Bx_t \pm c_2 By_t;$$

(4)

$$(1 - B)^n x_t = \sum_{i=0}^n (-1)^i C_n^i B^i x_t.$$



差分运算与滞后算子

4. 差分运算与滞后算子的关系

根据差分运算和滞后算子的概念, 我们不难用滞后算子表示如下差分运算:

$$(1) \quad \nabla^n x_t = (1 - B)^n x_t = \sum_{i=0}^n (-1)^i C_n^i B^i x_t = \sum_{i=0}^n (-1)^i C_n^i x_{t-i};$$

$$(2) \quad \nabla_n x_t = (1 - B^n) x_t;$$



线性差分方程

1. 线性差分方程的概念

称如下形式的方程

$$x_t + a_1 x_{t-1} + a_2 x_{t-2} + \dots + a_n x_{t-n} = f(t) \quad (2.1)$$

为序列 $\{x_t, t=0, \pm 1, \pm 2, \dots\}$ n 阶线性差分方程, 其中 $n \geq 1; a_1, a_2, \dots, a_n$ 为实数, 且 $a_n \neq 0; f(t)$ 为 t 的已知函数.

特别地, 如果在方程 (2.1) 中 $f(t) \equiv 0$, 即

$$x_t + a_1 x_{t-1} + a_2 x_{t-2} + \dots + a_n x_{t-n} = 0 \quad (2.2)$$

那么称方程 (2.2) 为 n 阶齐次线性差分方程. 否则, 称方程 (2.1) 为 n 阶非齐次线性差分方程.



线性差分方程

2. 齐次线性差分方程解的结构

齐次线性差分方程的解的结构依赖于它的特征方程和特征根的取值情况. n 阶齐次线性差分方程 (2.2) 的特征方程为

$$\lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + \dots + a_n = 0 \quad (2.3)$$

由于 $a_n \neq 0$, 所以方程 (2.3) 最多有 n 个非零根, 我们称这 n 个非零根为 n 阶齐次线性差分方程 (2.2) 的特征根, 不妨记作 $\lambda_1, \lambda_2, \dots, \lambda_n$

根据特征根的不同情况, 齐次线性差分方程 (2.2) 的解具有不同的结构. 下面分情况讨论:



线性差分方程

(1) 当 $\lambda_1, \lambda_2, \dots, \lambda_n$ 为特征方程的互不相等的 n 个实根时, 齐次线性差分方程 (2.2) 的通解为

$$x_t = c_1 \lambda_1^t + c_2 \lambda_2^t + \dots + c_n \lambda_n^t$$

式中 c_1, c_2, \dots, c_n 为任意 n 个实常数.

(2) 当 $\lambda_1, \lambda_2, \dots, \lambda_n$ 中有相同实根时, 不妨假设 $\lambda_1 = \lambda_2 = \dots = \lambda_m$ 为 m 个相等实根, 而 $\lambda_{m+1}, \lambda_{m+2}, \dots, \lambda_n$ 为互不相等的实根, 则齐次线性差分方程 (2.2) 的通解为

$$x_t = (c_1 + c_2 t + \dots + c_m t^{m-1}) \lambda_1^t + c_{m+1} \lambda_{m+1}^t + c_{m+2} \lambda_{m+2}^t + \dots + c_n \lambda_n^t$$

式中 c_1, c_2, \dots, c_n 为任意 n 个实常数.



线性差分方程

(3) 当 $\lambda_1, \lambda_2, \dots, \lambda_n$ 中有复根时, 由于其复根必然成对共轭出现, 故而不妨假设 $\lambda_1 = re^{i\omega}$, $\lambda_2 = re^{-i\omega}$, 而 $\lambda_3, \lambda_4, \dots, \lambda_n$ 为互不相同的实根. 这时齐次线性差分方程 (2.2) 的通解为

$$x_t = r^t (c_1 e^{it\omega} + c_2 e^{-it\omega}) + c_3 \lambda_3^t + \dots + c_n \lambda_n^t$$

式中 c_1, c_2, \dots, c_n 为任意实数.

3. 非齐次线性差分方程的解

求解非齐次线性差分方程 (2.1) 需分三步进行. 第一步根据齐次线性差分方程 (2.2) 的特征根的情况确定方程 (2.2) 的通解 x'_t ; 第二步求出非齐次线性差分方程 (2.1) 的一个特解 x''_t , 所谓特解就是满足方程 (2.1) 的任意一个解; 第三步写出非齐次线性差分方程 (2.1) 的通解

$$x_t = x'_t + x''_t$$



线性差分方程

例 2.1: 求二阶非齐次线性差分方程

$$x_t - 5x_{t-1} + 6x_{t-2} = 2t - 7 \quad (2.4)$$

的通解.

解 原方程 (2.4) 对应的齐次方程

$$x_t - 5x_{t-1} + 6x_{t-2} = 0 \quad (2.5)$$

的特征方程为

$$\lambda^2 - 5\lambda + 6 = 0 \quad (2.6)$$

由方程 (2.6) 解得两个不等的特征根: $\lambda_1 = 2, \lambda_2 = 3$. 于是, 方程 (2.5) 的通解为

$$x'_t = c_1 2^t + c_2 3^t$$

其中 c_1, c_2 为任意常数. 容易观察得方程 (2.4) 的一个特解 $x''_t = t$. 故得方程 (2.4) 的通解为

$$x_t = c_1 2^t + c_2 3^t + t.$$



本章结构

1. 差分方程和滞后算子

2. 自回归模型的概念和性质

3. 移动平均模型的概念和性质

4. 自回归移动平均模型的概念和性质

自回归模型的定义

本节介绍较为简单的自回归 (autoregressive, AR) 模型以及它的统计性质. 事实上, 如果把差分方程 (2.1) 中等号右边的 $f(t)$ 换成随机噪声 ε_t , 那么就变成了随机差分方程. 从数学上来看, 下面的 AR 模型就是一种特殊的随机差分方程.

- 自回归模型的定义

设 $\{x_t, t \in T\}$ 为一个序列, 则称满足如下结构的模型为 **p 阶自回归模型**, 简记为 AR(p):

$$x_t = \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + \varepsilon_t \quad (2.7)$$

其中, $\phi_0, \phi_1, \phi_2, \dots, \phi_p$ 为 $p+1$ 个固定常数, 并要求 $\phi_p \neq 0$; ε_t 为均值为零的白噪声序列, 且 ε_t 与 x_{t-j} ($j=1, 2, \dots$) 无关.



自回归模型的定义

- 注:

(1) 在模型 AR(p) 中, 要求随机干扰项 ε_t 为零均值的白噪声序列, 即满足

$$E\varepsilon_t = 0, \text{Var}(\varepsilon_t) = \sigma^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t.$$

(2) 当期 (即现在这时刻) 的随机干扰 ε_t 与过去序列值无关, 即

$$E(x_s \varepsilon_t) = 0, \forall s < t.$$

(3) 当 $\phi_0 = 0$ 时, 自回归模型(2.7) 称为中心化 AR(p) 模型. 当 $\phi_0 \neq 0$ 时, 自回归模型 (2.7) 称为非中心化 AR(p) 模型, 此时, 令

$$\mu = \frac{\phi_0}{1 - \phi_1 - \phi_2 - \dots - \phi_p}, y_t = x_t - \mu$$

则 $\{y_t\}$ 就为中心化序列.

自回归模型的定义

- 注:

上述变换实际上就是非中心化序列整体平移了一个常数单位, 这种整体平移对序列值之间的相关关系没有任何影响, 所以今后在分析 AR 模型的相关关系时, 都简化为中心化模型进行分析.

应用滞后算子, 中心化 AR(p) 模型可表示为

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

其中, $\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ 称为 **p** 阶自回归系数多项式.



自回归模型的定义

- 例 2.2 设某商品的价格序列为 $\{x_t\}$, 该商品的需求量为 $Q_t^d = a - bx_t$, 而供给量为 $Q_t^s = -c + dx_{t-1}$. 一种较为理想的经营策略是需求量与供给量相等, 从而得到

$$x_t = \frac{a+c}{b} - \frac{d}{b}x_{t-1}.$$

- 事实上, 商品的需求与供给还可能受到收入、偏好等其他众多非主要因素的干扰. 干扰项设为 ε_t , 将其加入模型中有

- $$x_t = \frac{a+c}{b} - \frac{d}{b}x_{t-1} + \varepsilon_t \quad (2.8)$$

- 一般假设 $\varepsilon_t \sim WN(0, \sigma^2)$, 而且 ε_t 与 x_{t-1} 不相关, 此时模型 (2.8) 就是一个非中心化 AR(1)模型.

自回归模型的定义

- 将模型 (2.8) 实施中心化变换, 化为形如

$$y_t = \phi_1 y_{t-1} + \varepsilon_t \quad (2.9)$$

的中心化 AR(1) 模型. 对模型 (2.9) 进行反复迭代运算, 得到

$$y_t = \sum_{k=0}^{\infty} \phi_1^k \varepsilon_{t-k} \quad (2.10)$$

从 (2.10) 式可见, 服从 AR(1) 过程的时间序列 $\{y_t\}$ 经过多次迭代后成为白噪声序列 $\{\varepsilon_t\}$ 的加权和. ϕ_1^k 描述了第 $t-k$ 期噪声对 y_t 的影响. 如果 $|\phi_1| < 1$, 意味着随着 k 的增加, 噪声对 y_t 的影响越来越弱, 特别是当 $k \rightarrow \infty$, 噪声对 y_t 的影响趋于零. 在下面我们将看到, 当 $|\phi_1| < 1$ 时, 服从 AR(1) 模型的时间序列 $\{y_t\}$ 是个平稳的时间序列.



序列的拟合

在研究序列的统计性质之前, 一个不错的习惯是首先拟合序列值的走向, 得到一个对数据感性的认识. 这样做有时对理论分析大有裨益.

R 提供了多种拟合序列的函数. 下面我们介绍两种常用的序列拟合方法.

- filter 拟合函数

filter() 函数可以直接拟合 AR 序列 (不论是否平稳) 和后面将讲述的 MA 序列. filter() 函数的命令格式为:

filter(a, filter= , method= , circular=)



序列的拟合

filter() 函数的命令格式为:

filter(a, filter= , method= , circular=)

该函数的参数说明:

- **a**: 随机波动序列的变量名.

- **filter**: 指定模型系数, 其中

(1) AR(p) 模型为 $filter = c(\phi_1, \phi_2, \dots, \phi_p)$

(2) MA(q) 模型为 $filter = c(1, -\theta_1, \dots, -\theta_q)$

-**method**: 指定拟合的是 AR 模型还是 MA 模型.

(1) method= “recursive” 为 AR 模型;

(2) method= “convolution” 为 MA 模型.

-**circular**: 拟合 MA 模型时专用的一个选项, circular=T 可以避免 NA 数据出现.



序列的拟合

■ arima.sim 拟合函数

arima.sim() 函数可以拟合平稳 AR 序列以及后面将讲述的 MA 序列、平稳 ARMA 序列和 ARIMA 序列. 命令格式为:

arima.sim(n, list(ar= ,ma= ,order=),sd=)

该函数的参数说明:

-n: 拟合序列的长度.

-list: 指定具体模型参数, 其中:

(1) 拟合平稳 AR(p) 模型, 要给出自回归系数: $ar = c(\phi_1, \phi_2, \dots, \phi_p)$. 如果指定拟合的 AR 模型为非平稳模型, 系统就会报错.

(2) 拟合 MA(q) 模型, 要给出移动平均系数: $ma = c(1, -\theta_1, \dots, -\theta_q)$

(3) 拟合平稳 ARMA(p,q) 模型, 需要同时给出自回归系数和移动平均系数. 如果指定拟合的 ARMA 模型为非平稳的模型, 系统会报错.



序列的拟合

- arima.sim 拟合函数

arima.sim() 函数可以拟合平稳 AR 序列以及后面将讲述的 MA 序列、平稳 ARMA 序列和 ARIMA 序列. 命令格式为:

arima.sim(n, list(ar= ,ma= ,order=),sd=)

该函数的参数说明:

(4)拟合 ARIMA(p,d,q) 模型, 需要自回归系数、移动平均系数和选项 order. order=c(p,d,q),其中 p 为自回归阶数、d 为差分阶数、q 为移动平均阶数.-sd: 指定序列的标准差, 如不特殊指定, 系统默认 sd=1.



序列的拟合

例 2.3 拟合下列 AR 模型, 并绘制时序图:

$$(1) x_t = 0.6x_{t-1} + \varepsilon_t \quad (2) x_t = x_{t-1} + \varepsilon_t$$

$$(3) x_t = 1.8x_{t-1} + \varepsilon_t \quad (4) x_t = x_{t-1} + 0.3x_{t-2} + \varepsilon_t$$

这里 $\varepsilon_t \sim WN(0, \sigma^2)$

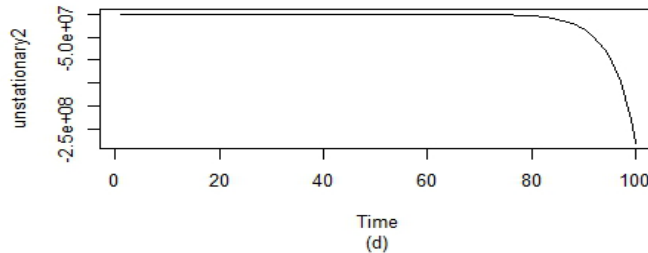
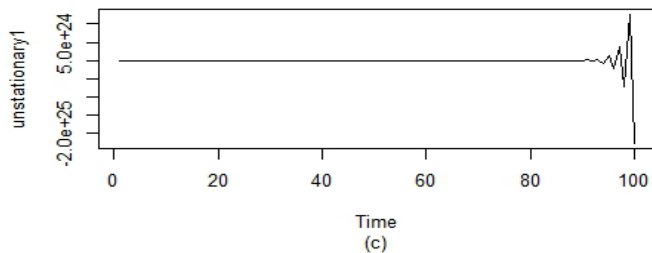
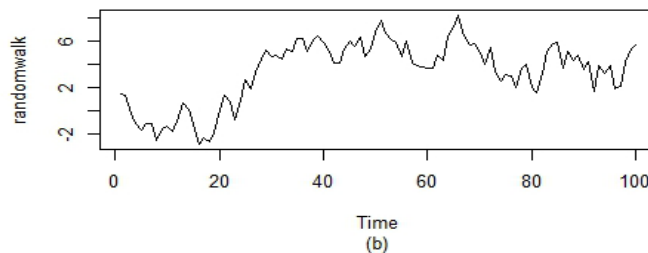
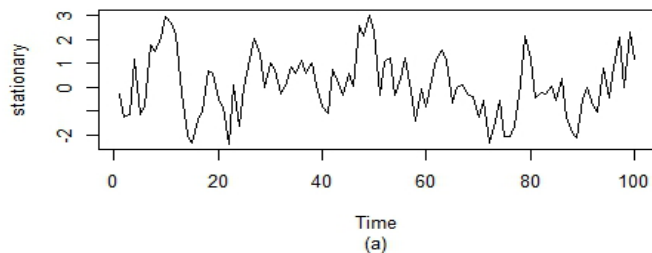
解: 我们用 R 拟合这四个序列的序列值, 并绘制时序图, 具体命令如下, 运行结果见图 2.1.

```
> stationary <- arima.sim(n=100, list(ar=0.6))  
> e <- rnorm(100)  
> randomwalk <- filter(e,filter=1, method="recursive")  
> unstationary1 <- filter(e,filter=1.8, method="recursive")  
> unstationary2 <- filter(e,filter=c(1,0.3), method="recursive")  
> par(mfrow=c(2,2))
```

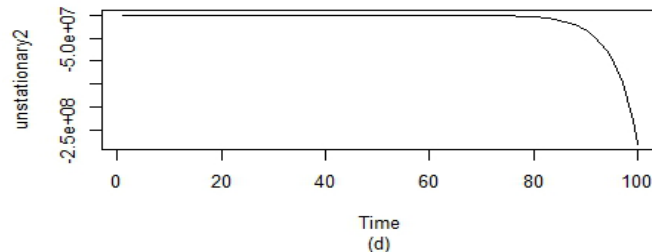
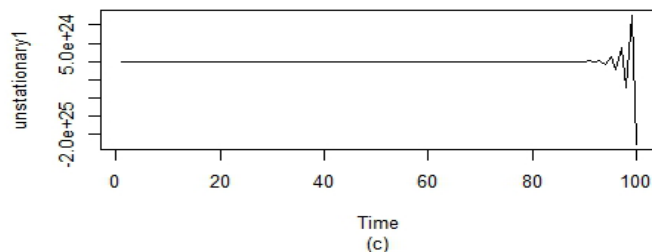
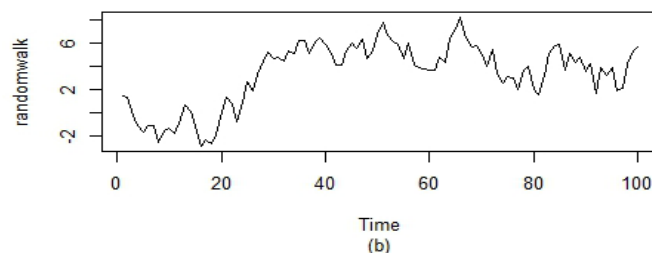
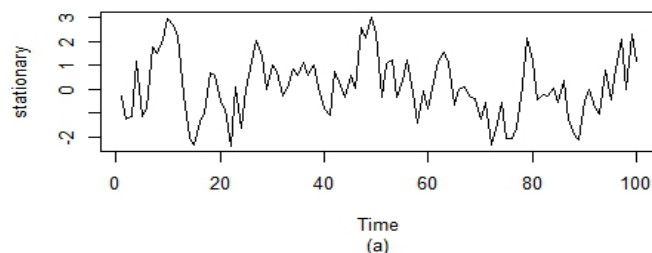


序列的拟合

- > ts.plot(stationary,sub="(a)")
- > ts.plot(randomwalk,sub="(b)")
- > ts.plot(unstationary1,sub="(c)")
- > ts.plot(unstationary2,sub="(d)")

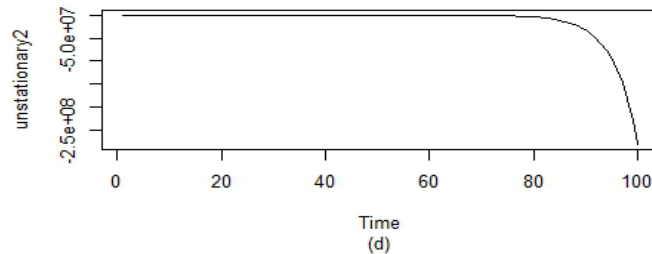
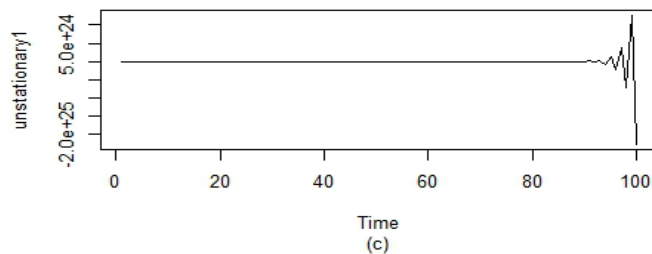
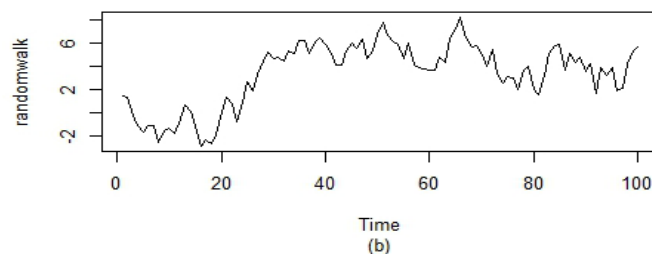
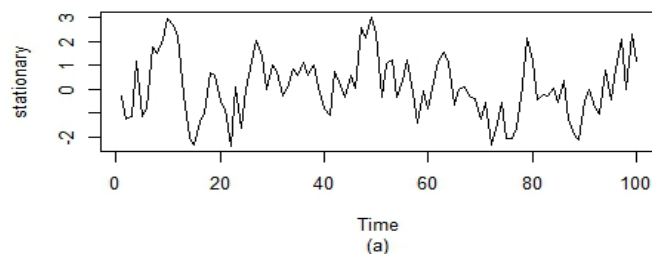


序列的拟合



题 (1) 的自回归系数 $\phi_1 = 0.6$, 所以距离时刻 t 越远的噪声对 x_t 影响越小, 时序图 2.1(a) 呈现平稳态势. 题 (2) 的自回归系数 $\phi_1 = 0.6$, 所以噪声对 x_t 的影响不随时间的推移而减弱. 我们称这种序列为随机游走. 从图 2.1(b) 可见, 随机游走序列在一定时间段具有一定的上升或者下降趋势, 也就是运动的方向有一定的持续性. 与图 2.1(a) 相比,

序列的拟合



随机游走并不是围绕某个值上下波动, 而出现了持续偏离, 因此随机游走序列不是平稳序列. 图 2.1(c) 表明题(3) 方差逐渐增大, 而图 2.4(d) 表明题 (4) 的序列值有明显的增加趋势, 故而它们都不平稳序列.

稳定性与平稳性

AR 模型是最常用的拟合平稳序列的模型之一, 但是并非所有 AR 模型都是平稳的, 甚至有些 AR 模型都不一定是稳定的. 下面我们来学习判别 AR 模型稳定性和平稳性的方法.

1. Green 函数

设 $\{x_t, t \in T\}$ 是一个序列, 如果 x_t 可表示为零均值白噪声序列 ε_t 的级数和, 即

$$x_t = G_0 \varepsilon_t + G_1 \varepsilon_{t-1} + G_2 \varepsilon_{t-2} + \dots$$

那么系数函数 $G_i (i = 0, 1, 2, \dots)$ 称为 Green 函数.

如: 前面(2.10)



稳定性与平稳性

- 下面根据待定系数法, 求中心化 AR(p) 模型的 Green 函数:
首先, 任何一个中心化 AR(p) 序列经过反复迭代运算总可以表示成

$$x_t = \sum_{k=0}^{\infty} G_k \varepsilon_{t-k} = \left(\sum_{k=0}^{\infty} G_k B^k \right) \varepsilon_t \quad (2.11)$$

(2.11) 式通常称为 AR 模型的传递形式. 将上式代入 $\Phi(B)x_t = \varepsilon_t$, 得

$$(1 - \sum_{i=0}^p \phi_i B^i) \left(\sum_{k=0}^{\infty} G_k B^k \right) \varepsilon_t = \varepsilon_t,$$

展开得

$$\left(G_0 + \sum_{k=1}^{\infty} G_k B^k - \left(\sum_{i=0}^p \sum_{k=0}^{\infty} \phi_i G_k B^{k+i} \right) \right) \varepsilon_t = \varepsilon_t$$

$$\left(G_0 + \sum_{k=1}^{\infty} G_k B^k - [G_0 \phi_1 B + (G_1 \phi_1 + G_0 \phi_2) B^2 + (G_2 \phi_1 + G_1 \phi_2 + G_0 \phi_3) B^3 + \dots] \right) \varepsilon_t = \varepsilon_t$$

$$\left(G_0 + \sum_{k=1}^{\infty} G_k B^k - [G_0 \phi_1 B + (G_1 \phi_1 + G_0 \phi_2) B^2 + (G_2 \phi_1 + G_1 \phi_2 + G_0 \phi_3) B^3 + \dots] \right) \varepsilon_t = \varepsilon_t$$



稳定性与平稳性

$$\left(G_0 + \sum_{k=1}^{\infty} G_k B^k - [G_0 \phi_1 B + (G_1 \phi_1 + G_0 \phi_2) B^2 + (G_2 \phi_1 + G_1 \phi_2 + G_0 \phi_3) B^3 + \dots] \right) \varepsilon_t = \varepsilon_t$$

$$\left(G_0 + \sum_{k=1}^{\infty} G_k B^k - \sum_{k=1}^{\infty} \left(\sum_{i=1}^k \phi_i^* G_{k-i} \right) B^k \right) \varepsilon_t = \varepsilon_t$$

$$\left(G_0 + \sum_{k=1}^{\infty} (G_k - \sum_{i=1}^k \phi_i^* G_{k-i}) B^k \right) \varepsilon_t = \varepsilon_t$$

由待定系数法得

$$G_0 = 1, G_k - \sum_{i=1}^k \phi_i^* G_{k-i} = 0, k = 1, 2, \dots$$

于是得到 Green 函数的递推计算公式

$$G_0 = 1, G_k = \sum_{i=1}^k \phi_i^* G_{k-i}, k = 1, 2, \dots$$

其中, 当 $i \leq p$ 时, $\phi_i^* = \phi_i$; 当 $i > p$ 时, $\phi_i^* = 0$.



稳定性与平稳性

- Green 函数的意义

从 AR 模型的传递形式, $x_t = \sum_{k=0}^{\infty} G_k \varepsilon_{t-k} = \left(\sum_{k=0}^{\infty} G_k B^k \right) \varepsilon_t$ 我们可以看到, Green 函数 G_k 是 $t - k$ 时刻的干扰项 ε_{t-k} 的权数, $|G_k|$ 越大, 表明过去的干扰对 t 时刻的序列值影响也越大, 说明系统的记忆性越强. 如果 $|G_k| \rightarrow 0, k \rightarrow \infty$, 那么说明过去的干扰的影响逐渐衰减; 如果当 $k \rightarrow \infty$ 时, $|G_k|$ 不收敛于零, 那么说明过去干扰的影响不随时间的推移而衰退. 这样的序列将是不平稳的.

借助于 Green 函数的概念, 我们可以考察序列如下三种类型的稳定性: 稳定 (stable); 渐近稳定 (asymptotically stable); 一致渐近稳定 (uniformly asymptotically stable).



稳定性与平稳性

- 如果存在常数 $M > 0$, 使得对于一切 k , $|G_k| \leq M$, 那么称序列

$$x_t = \sum_{k=0}^{\infty} G_k \varepsilon_{t-k}$$

是稳定的 (stable).

- 如果 $k \rightarrow \infty$ 时, $|G_k| \rightarrow 0$, 则称序列是 $x_t = \sum_{k=0}^{\infty} G_k \varepsilon_{t-k}$ 渐近稳定的 (asymptotically stable).
- 如果存在常数 $a > 0$, $b > 0$, 使得对于一切 k , $|G_k| \leq ae^{-bk}$, 则称序列 $x_t = \sum_{k=0}^{\infty} G_k \varepsilon_{t-k}$ 是一致渐近稳定的 (uniformly asymptotically stable).

稳定性与平稳性

2. AR 模型平稳性的判别

定理 2.1 设 $\{x_t, t \in T\}$ 是一个中心化 AR(p) 模型 $\Phi(B)x_t = \varepsilon_t$, 其中

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

则, $\{x_t, t \in T\}$ 平稳的充分必要条件是

$$\Phi(u) = 1 - \phi_1 u - \phi_2 u^2 - \dots - \phi_p u^p = 0$$

的根在单位圆外.

证明: 设 $\Phi(u)=0$ 的 p 个根为 $\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \dots, \frac{1}{\lambda_p}$, 则

$$\Phi(B) = a(1 - \lambda_1 B)(1 - \lambda_2 B) \cdots (1 - \lambda_p B)$$

从而

$$x_t = \Phi^{-1}(B)\varepsilon_t = \frac{1}{a(1 - \lambda_1 B)(1 - \lambda_2 B) \cdots (1 - \lambda_p B)} \varepsilon_t$$

用待定系数法, 得

$$x_t = \sum_{k=1}^p a_k (1 - \lambda_k B)^{-1} \varepsilon_t$$



稳定性与平稳性

其中 a_k 是有限实数. 于是有

$$x_t = \sum_{k=1}^p a_k \left(\sum_{i=0}^{\infty} \lambda_k^i B^i \right) \varepsilon_t = \sum_{i=0}^{\infty} \left(\sum_{k=1}^p a_k \lambda_k^i \right) B^i \varepsilon_t = \sum_{i=0}^{\infty} \left(\sum_{k=1}^p a_k \lambda_k^i \right) \varepsilon_{t-i} \quad (2.12)$$

(2.12) 式是一个白噪声加权和. 我们指出, x_t 平稳的充分必要条件是权系数绝对收敛于零, 而权系数绝对收敛于零的充分必要条件是所有的 λ_i 的绝对值小于 1, 即它的根 $\frac{1}{\lambda_i}, i=1,2,\dots,p$ 都在单位圆之外.

推论 2.1: AR(p) 模型 $\{x_t, t \in T\}$ 平稳的充要条件是它的齐次线性差分方程 $\Phi(B)x_t = 0$, 的特征根都在单位圆内.



稳定性与平稳性

- 根据推论 2.1 知, AR(p) 模型平稳的充要条件是它的齐次线性差分方程的特征根都在单位圆内, 而特征根是由自回归系数决定的, 因此满足这个条件的自回归系数构成一个集合, 这个集合称为平稳域. 更准确地, 引入下列概念.
- 对于一个 AR(p) 模型, 我们称使其特征根都在单位圆内的 p 个系数构成的向量的集合为 AR(p) 模型的**平稳域**, 即

$$\{(\phi_1, \phi_2, \dots, \phi_p) : \text{特征根都在单位圆内}\}.$$

- 对于低阶 AR 模型用平稳域的方法判别其稳定性更为方便:

例 2.4 求 AR(1) 模型 $x_t = \phi_1 x_{t-1} + \varepsilon_t$ 的平稳域.

例 2.5 求 AR(2) 模型 $x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \varepsilon_t$ 的平稳域.



稳定性与平稳性

3. AR 模型平稳性与稳定性的关系

定理 2.2 中心化 AR(p) 模型

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + \varepsilon_t$$

平稳的充分必要条件是它渐近稳定或一致渐近稳定.

证明: 设 $\lambda_1, \lambda_2, \dots, \lambda_p$ 是中心化 AR(p) 模型的 p 个特征根, 则根据 (2.12) 式

$$x_t = \sum_{i=0}^{\infty} \left(\sum_{k=1}^p a_k \lambda_k^i \right) \varepsilon_{t-i}$$

得 Green 函数 $G_i = \sum_{k=1}^p a_k \lambda_k^i, i \geq 0.$

当 AR(p) 模型平稳时, $|\lambda_k| < 1, k = 1, 2, \dots, p.$ 记 $A = \max\{|a_k|\}, B = \max\{|\lambda_k|\}$ 得

$$|G_i| \leq pAB^i = pAe^{-(\ln B)i}, i \geq 0$$

于是得到, 平稳 AR(p) 模型是一致稳定的, 且渐近稳定的.



稳定性与平稳性

反之, 若 $AR(p)$ 模型是一致稳定的, 或渐近稳定的, 则

$$\lim |G_i| = 0$$

从而, $|\lambda_k| < 1, k = 1, 2, \dots, p$. 因此, $AR(p)$ 模型是平稳的

- ◆ 不过, 如果 $AR(p)$ 模型仅满足稳定性, 那么不一定是平稳的. 如: 随机游动序列.

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

其中 $G_i = 1$. 因而是稳定的, 但不是平稳的.

稳定性与平稳性

反之, 若 $AR(p)$ 模型是一致稳定的, 或渐近稳定的, 则

$$\lim |G_i| = 0$$

从而, $|\lambda_k| < 1, k = 1, 2, \dots, p$. 因此, $AR(p)$ 模型是平稳的

- ◆ 不过, 如果 $AR(p)$ 模型仅满足稳定性, 那么不一定是平稳的. 如: 随机游动序列.

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

其中 $G_i = 1$. 因而是稳定的, 但不是平稳的.

平稳自回归模型的统计性质

1. 均值函数

假如 AR(p) 模型 (2.7)

$$x_t = \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + \varepsilon_t$$

满足平稳性条件, 在等式两边同时取期望, 得

$$Ex_t = \phi_0 + \phi_1 Ex_{t-1} + \phi_2 Ex_{t-2} + \cdots + \phi_p Ex_{t-p} + E\varepsilon_t$$

根据平稳性条件, 得 $Ex_t = \mu, \forall t \in T$. 于是

$$\mu = \frac{\phi_0}{1 - \phi_1 - \phi_2 - \cdots - \phi_p}.$$

特别地, 对于中心化 AR(p) 模型, $Ex_t = 0, \forall t \in T$.



平稳自回归模型的统计性质

2. 方差函数

假如 AR(p) 模型是平稳的, 对其传递形式 (2.11) 式两边求方差, 有

$$\text{Var}(x_t) = \sum_{i=0}^{\infty} G_i^2 \text{Var}(\varepsilon_t) = \sigma_{\varepsilon}^2 \sum_{i=0}^{\infty} G_i^2,$$

因为 $\{x_t\}$ 平稳, 所以 $\{x_t\}$ 的方差恒为常数 $\sigma_{\varepsilon}^2 \sum_{i=0}^{\infty} G_i^2$.

3. 自协方差函数

在中心化 AR(p) 平稳模型 (2.13) 两边同时乘以 $x_{t-k}, k \geq 1$ 然后求期望, 得

$$E x_t x_{t-k} = \phi_1 E x_{t-1} x_{t-k} + \phi_2 E x_{t-2} x_{t-k} + \cdots + \phi_p E x_{t-p} x_{t-k} + E \varepsilon_t x_{t-k}$$

根据 AR(p) 模型的定义和平稳性, 得自协方差函数的递推公式:

$$\gamma(k) = \phi_1 \gamma(k-1) + \cdots + \phi_p \gamma(k-p)$$



平稳自回归模型的统计性质

4. 自相关函数

在自协方差函数递推公式 两边同时除以方差函数 $\gamma(0)$, 就得到自相关函数的递推公式:

$$\rho(k) = \phi_1 \rho(k-1) + \cdots + \phi_p \rho(k-p)$$

例 2.6 : 求平稳 AR(1) 模型 $x_t = \phi_1 x_{t-1} + \varepsilon_t$ 的方差、自协方差函数和自相关函数.

例 2.7 : 求平稳 AR(2) 模型 $x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \varepsilon_t$ 的方差、自协方差函数的递推公式以及自相关函数的递推公式.

解: 根据 Green 函数 $G_0 = 1, G_k = \sum_{i=1}^k \phi_i^* G_{k-i}, k = 1, 2, \dots$ 其中, 当 $i \leq p$ 时, $\phi_i^* = \phi_i$; 当 $i > p$ 时, $\phi_i^* = 0$. 得

$$G_0 = 1, G_1 = \phi_1 G_0 = \phi_1, G_2 = \phi_1^2 + \phi_2, \dots, G_k = \phi_1 G_{k-1} + \phi_2 G_{k-2},$$



平稳自回归模型的统计性质

例 2.7 : 求平稳 AR(2) 模型 $x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \varepsilon_t$ 的方差、自协方差函数的递推公式以及自相关函数的递推公式.

解: 根据 Green 函数 $G_0 = 1, G_k = \sum_{i=1}^k \phi_i^* G_{k-i}, k = 1, 2, \dots$ 其中, 当 $i \leq p$ 时, $\phi_i^* = \phi_i$; 当 $i > p$ 时, $\phi_i^* = 0$. 得 $G_0 = 1, G_1 = \phi_1 G_0 = \phi_1, G_2 = \phi_1^2 + \phi_2, \dots, G_k = \phi_1 G_{k-1} + \phi_2 G_{k-2}$,
在 $\gamma(k) = \phi_1 \gamma(k-1) + \dots + \phi_p \gamma(k-p)$ 式中, 取 $k = 1$ 得, $\gamma(1) = \phi_1 \gamma(0) + \phi_2 \gamma(1)$
从而

$$\gamma(1) = \frac{\phi_1 \gamma(0)}{1 - \phi_2}$$

于是得自协方差函数和自相关函数的递推公式.



平稳 AR(p) 模型的拖尾性和呈指数衰减

- ◆ 平稳 AR(p) 模型的自相关函数的两个显著性质是, 拖尾性和呈指数衰减.
- ◆ AR(p) 模型的自相关函数递推公式

$$\rho(k) = \phi_1 \rho(k-1) + \cdots + \phi_p \rho(k-p)$$

是一个 p 阶齐次差分方程. 不妨设它有 p 个互不相同的实特征根 $\lambda_1, \lambda_2, \dots, \lambda_p$, 则滞后 k 阶的自相关函数的通解为

$$\rho(k) = c_1 \lambda_1^k + c_2 \lambda_2^k + \dots + c_p \lambda_p^k$$

其中, c_1, c_2, \dots, c_p 是不全为零的任意常数. 通过上式可以看出, $\rho(k)$ 始终非零, 即不会在 k 大于某个值之后就恒为零, 这个性质称为**拖尾性**.



平稳 AR(p) 模型的拖尾性和呈指数衰减

- ◆ AR(p) 模型的自相关函数的拖尾性质的直观解释：对于平稳AR(p)模型

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + \varepsilon_t$$

虽然表达式直接显示 x_t 受当期 ε_t 和最近 p 期的序列值 $x_{t-1}, x_{t-2}, \dots, x_{t-p}$ 的影响, 但是 x_{t-1} 也会受到 x_{t-1-p} 的影响, 以此类推, x_t 之前的每个值 $x_{t-1}, x_{t-2}, \dots, x_{t-p}, \dots$ 对 x_t 都会有影响, 这种特性表现在自相关函数上就是自相关系数的拖尾性.

- ◆ 对于平稳 AR(p) 模型而言, 其特征值 $|\lambda_i| < 1$, 所以当 $k \rightarrow \infty$ 时, $\rho(k) = c_1 \lambda_1^k + c_2 \lambda_2^k + \dots + c_p \lambda_p^k \rightarrow 0$, 且随着时间的推移是呈指数 λ^k 的速度衰减的. 这种自相关函数以指数衰减的性质就是 1.5 节中利用自相关图判断平稳序列时所说的“短期相关性”, 它是平稳序列的一个重要特征.



平稳 AR(p) 模型的拖尾性和呈指数衰减

◆ 例 2.8 观察以下四个平稳 AR 模型的自相关图:

$$(1) \quad x_t = 0.8x_{t-1} + \varepsilon_t \quad (2) \quad x_t = -0.7x_{t-1} + \varepsilon_t$$

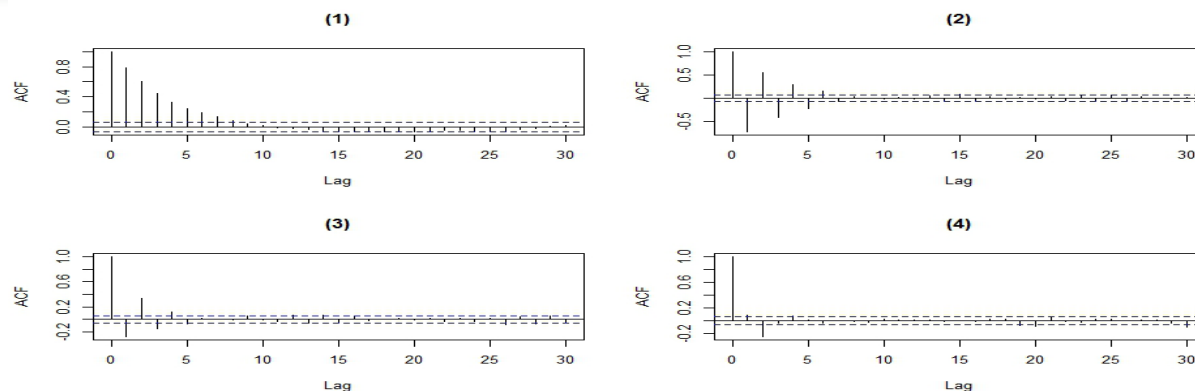
$$(3) \quad x_t = -0.2x_{t-1} + 0.3x_{t-2} + \varepsilon_t \quad (4) \quad x_t = 0.2x_{t-1} - 0.3x_{t-2} + \varepsilon_t$$

其中, ε_t 为标准正态白噪声序列.

解: 我们按如下命令绘制出四个自相关图

```
> a <- arima.sim(n=1000,list(ar=0.8))  
> b <- arima.sim(n=1000,list(ar=-0.7))  
> c <- arima.sim(n=1000,list(ar=c(-0.2,0.3)))  
> d <- arima.sim(n=1000,list(ar=c(0.2,-0.3)))  
> par(mfrow=c(2,2))  
> acf(a);acf(b);acf(c);acf(d)
```

平稳 AR(p) 模型的拖尾性和呈指数衰减



从图 2.2 看到, 这四个 AR 模型不论它们具有何种形式的特征根, 它们的自相关函数都呈现出拖尾性和指数衰减性. 只是由于特征根的不同会导致自相关函数的衰减方式也不一样: 模型(1) 的自相关函数按负指数衰减到零附近; 模型 (2) 的自相关函数呈现正负相间地衰减; 模型 (3)的自相关函数具有“伪周期”的衰减特征. 模型 (4) 的自相关函数虽然没有明显的衰减规律, 但是衰减速度非常快. 这些都是平稳模型自相关函数常见的特征.

偏自相关函数

偏自相关函数的概念反映了给定其他变量值的条件下,第 s 期与第 t 期变量的条件相关系数. 具体地, 对于平稳序列 $\{x_t\}$ 而言, 所谓**滞后 k 偏自相关函数**就是给定中间 $k-1$ 个随机变量 $x_{t-1}, x_{t-2}, \dots, x_{t-k+1}$ 的条件下, x_{t-k} 与 x_t 的相关系数, 反映了剔除中间 $k-1$ 个变量值的干扰之后, x_{t-k} 对 x_t 的纯粹相关影响的度量. 其数学表述是

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

其中, $\hat{E}x_t = E[x_t | x_{t-1}, \dots, x_{t-k+1}]$, $\hat{E}x_{t-k} = E[x_{t-k} | x_{t-1}, \dots, x_{t-k+1}]$

偏自相关函数

对于中心化平稳序列 $\{x_t\}$,用过去 k 期序列值 $x_{t-1}, x_{t-2}, \dots, x_{t-k}$ 对 x_t 作 k 阶自回归拟合, 有

$$x_t = \phi_{k1}x_{t-1} + \phi_{k2}x_{t-2} + \dots + \phi_{kk}x_{t-k} + \varepsilon_t \quad (2.18)$$

式中, ε_t 是均值为零的白噪声序列, 且对任意 $s < t, E\varepsilon_t x_s = 0$.

以 $x_{t-1}, x_{t-2}, \dots, x_{t-k+1}$ 为条件, 在 (2.18) 式两边求条件期望, 得

$$\hat{E}x_t = \phi_{k1}x_{t-1} + \phi_{k2}x_{t-2} + \dots + \phi_{k(k-1)}x_{t-k+1} + \phi_{kk}\hat{E}x_{t-k} + E(\varepsilon_t | x_{t-1}, \dots, x_{t-k+1})$$

$$\hat{E}x_t = \phi_{k1}x_{t-1} + \phi_{k2}x_{t-2} + \dots + \phi_{k(k-1)}x_{t-k+1} + \phi_{kk}\hat{E}x_{t-k}. \quad (2.19)$$

用 (2.18) 式减 (2.19) 式, 得

$$x_t - \hat{E}x_t = \phi_{kk}(x_{t-k} - \hat{E}x_{t-k}) + \varepsilon_t \quad (2.20)$$

在 (2.20) 式两边同时乘以 $x_{t-k} - \hat{E}x_{t-k}$, 然后两边同时求期望, 得

$$E[(x_t - \hat{E}x_t)(x_{t-k} - \hat{E}x_{t-k})] = \phi_{kk} E[(x_{t-k} - \hat{E}x_{t-k})^2],$$

于是

$$\phi_{kk} = \frac{E[(x_t - \hat{E}x_t)(x_{t-k} - \hat{E}x_{t-k})]}{E[(x_{t-k} - \hat{E}x_{t-k})^2]}.$$



偏自相关函数

$$\phi_{kk} = \frac{E[(x_t - \hat{E}x_t)(x_{t-k} - \hat{E}x_{t-k})]}{E[(x_{t-k} - \hat{E}x_{t-k})^2]}.$$

这表明滞后 k 偏自相关函数就是 k 阶自回归模型第 k 个回归系数 ϕ_{kk} 的值. 根据这个性质, 我们可以计算偏自相关函数的值.

现在我们构造序列 $\{x_t\}$ 的最佳 k 阶自回归拟合, 即求使得模型残差的方差

$$R(\phi_{k1}, \phi_{k2}, \dots, \phi_{kk}) = E \left[\left(x_t - \sum_{i=1}^k \phi_{ki} x_{t-i} \right)^2 \right]$$

最小的参数 $\phi_{k1}, \phi_{k2}, \dots, \phi_{kk}$.

偏自相关函数

记 $\Phi = (\phi_{k1}, \phi_{k2}, \dots, \phi_{kk})$, $\mathbf{x} = (x_{t-1}, x_{t-2}, \dots, x_{t-k})$, $\Gamma = (\gamma(1), \gamma(2), \dots, \gamma(k))$

则

$$\begin{aligned} R(\phi_{k1}, \phi_{k2}, \dots, \phi_{kk}) &= E\left[\left(x_t - \Phi \mathbf{x}^T\right)^2\right] = E\left[x_t^2 - 2x_t \Phi \mathbf{x}^T + \Phi \mathbf{x}^T \mathbf{x} \Phi^T\right] \\ &= \gamma(0) - 2\Phi \Gamma^T + \Phi \begin{bmatrix} \gamma(0) & \gamma(1) & \dots & \gamma(k-1) \\ \gamma(1) & \gamma(0) & \dots & \gamma(k-2) \\ \vdots & \vdots & & \vdots \\ \gamma(k-1) & \gamma(k-2) & \dots & \gamma(0) \end{bmatrix} \Phi^T \end{aligned}$$

$R(\phi_{k1}, \phi_{k2}, \dots, \phi_{kk})$ 关于各变元 $\phi_{k1}, \phi_{k2}, \dots, \phi_{kk}$ 求偏导, 并根据取得极值的必要条件, 得

$$\partial R / \partial \Phi = -2\Gamma^T + 2 \begin{bmatrix} \gamma(0) & \gamma(1) & \dots & \gamma(k-1) \\ \gamma(1) & \gamma(0) & \dots & \gamma(k-2) \\ \vdots & \vdots & & \vdots \\ \gamma(k-1) & \gamma(k-2) & \dots & \gamma(0) \end{bmatrix} \Phi^T = 0.$$

偏自相关函数

于是有

$$\begin{bmatrix} 1 & \rho(1) & \dots & \rho(k-1) \\ \rho(1) & 1 & \dots & \rho(k-2) \\ \vdots & \vdots & & \vdots \\ \rho(k-1) & \rho(k-2) & \dots & 1 \end{bmatrix} \Phi^T = \Psi^T. \quad (2.21)$$

其中,

$R(\phi_{k1}, \phi_{k2}, \dots, \phi_{kk})$ 关于各变元 $\phi_{k1}, \phi_{k2}, \dots, \phi_{kk}$ 求偏导, 并根据取得极值的必要条件, 得 $\Psi^T = (\rho(1), \rho(2), \dots, \rho(k))$. 我们称线性方程组 (2.21) 为 Yule-Walker 方程. 该方程组的解 $\Phi = (\phi_{k1}, \phi_{k2}, \dots, \phi_{kk})$ 的最后一个分量 ϕ_{kk} 就是滞后 k 偏自相关函数.

偏自相关函数

特别地, 若线性方程组 (2.21) 的系数行列式不为零, 则根据 Cramer 法则, 得

$$\phi_{kk} = \frac{D_k}{D} \quad (2.22)$$

其中,

$$D = \begin{vmatrix} 1 & \rho(1) & \cdots & \rho(k-1) \\ \rho(1) & 1 & \cdots & \rho(k-2) \\ \vdots & \vdots & & \vdots \\ \rho(k-1) & \rho(k-2) & \cdots & 1 \end{vmatrix}, D_k = \begin{vmatrix} 1 & \rho(1) & \cdots & \rho(1) \\ \rho(1) & 1 & \cdots & \rho(2) \\ \vdots & \vdots & & \vdots \\ \rho(k-1) & \rho(k-2) & \cdots & \rho(k) \end{vmatrix}$$

对于中心化平稳 AR(p) 模型而言, 当 $k > p$ 时, $\phi_{kk} = 0$, 即滞后 k 偏自相关函数为 0. 这个性质我们称为平稳 AR(p) 模型的 p 步截尾性.

平稳 AR(p) 模型的偏自相关函数的 p 步截尾性、自相关函数的拖尾性和指数衰减性是其模型识别的重要依据.



偏自相关函数

例 2.9 分别求中心化平稳 AR(1) 模型: $x_t = \phi_1 x_{t-1} + \varepsilon_t$, 和中心化平稳 AR(2) 模型: $x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \varepsilon_t$, 的偏自相关函数.

解: 根据 (2.22) 式, 我们立刻得到中心化平稳 AR(1) 模型的偏自相关函数为

$$\phi_{kk} = \begin{cases} \phi_1, k = 1; \\ 0, k \geq 2. \end{cases}$$

中心化平稳 AR(2) 模型的偏自相关函数为

$$\phi_{kk} = \begin{cases} \frac{\phi_1}{1 - \phi_2}, k = 1; \\ \phi_2, k = 2; \\ 0, k > 2. \end{cases}$$

偏自相关函数

◆ 例 2.10 观察以下四个平稳 AR 模型的偏自相关函数的截尾性.

$$(1) \quad x_t = 0.8x_{t-1} + \varepsilon_t \quad (2) \quad x_t = -0.7x_{t-1} + \varepsilon_t$$

$$(3) \quad x_t = -0.2x_{t-1} + 0.3x_{t-2} + \varepsilon_t \quad (4) \quad x_t = 0.2x_{t-1} - 0.3x_{t-2} + \varepsilon_t$$

其中, ε_t 为标准正态白噪声序列.

解: 由例 2.9 的结论容易计算出例 2.8 中模型 (1) 和 (2) 的偏自相关函数分别为

$$\phi_{kk} = \begin{cases} 0.8, k=1; \\ 0, k \geq 2. \end{cases} \quad \phi_{kk} = \begin{cases} -0.7, k=1; \\ 0, k \geq 2. \end{cases}$$

模型 (3) 和模型 (4) 的偏自相关函数分别为

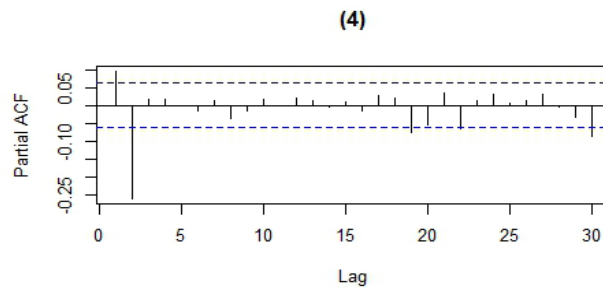
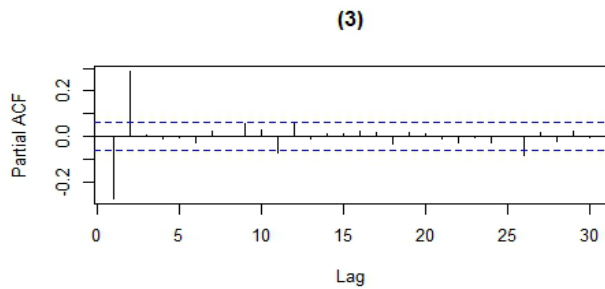
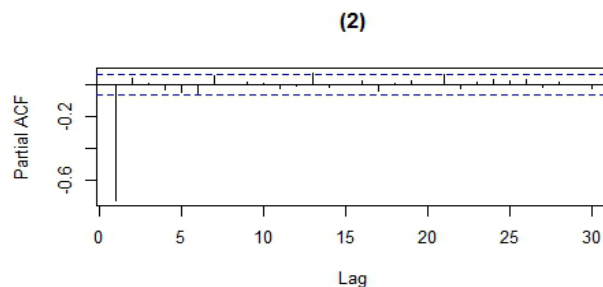
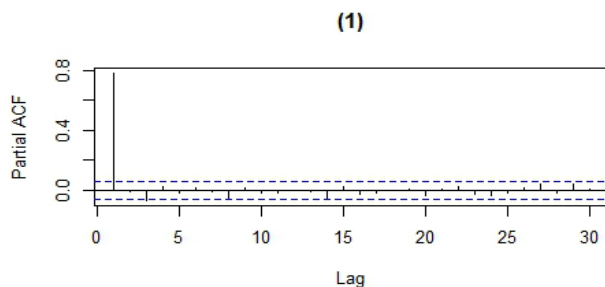
$$\phi_{kk} = \begin{cases} -2/7, k=1; \\ 0.3, k=2; \\ 0, k > 2. \end{cases} \quad \phi_{kk} = \begin{cases} 2/13, k=1; \\ -0.3, k=2; \\ 0, k > 2. \end{cases}$$



偏自相关函数

我们也可以通过绘制偏自相关函数图来观察平稳自回归模型的截尾性. 继续例 2.8 的操作,我们用下列语句绘制偏自相关函数图, 运行结果见图 2.3.

```
> pacf(a); pacf(b); pacf(c); pacf(d)
```



1. 差分方程和滞后算子

2. 自回归模型的概念和性质

3. 移动平均模型的概念和性质

4. 自回归移动平均模型的概念和性质

移动平均模型的定义

- 移动平均模型的定义

设 $\{x_t\}$ 是一个时间序列, 称满足如下结构的模型为 **q 阶移动平均模型** (qth-order moving average model), 简记为 MA(q),

$$x_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q},$$

其中, $\theta_q \neq 0$, 并且 $\{\varepsilon_t\}$ 是均值为零的白噪声序列.

- 中心化 MA(q) 模型
- q 阶移动平均系数多项式



移动平均模型的统计性质

从 MA(q) 模型的定义可以看出, $\{x_t\}$ 是由有限个白噪声的线性组合构成的, 因此 MA(q) 模型是平稳的. 下面我们来研究 MA(q) 模型的统计性质.

1. 均值函数

2. 方差函数

3. 自协方差函数

$$\begin{aligned}\gamma(k) &= E(x_t x_{t-k}) = E[(\varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q})(\varepsilon_{t-k} - \theta_1 \varepsilon_{t-k-1} - \theta_2 \varepsilon_{t-k-2} - \dots - \theta_q \varepsilon_{t-k-q})] \\ &= \begin{cases} (1 + \theta_1^2 + \dots + \theta_q^2) \sigma_\varepsilon^2, & k = 0; \\ (-\theta_k + \sum_{i=1}^{q-k} \theta_i \theta_{k+i}) \sigma_\varepsilon^2, & 1 \leq k \leq q; \\ 0, & k > q. \end{cases}\end{aligned}$$

由上式可见, MA(q) 模型的自协方差函数具有 q 阶截尾性



移动平均模型的统计性质

例 2.11 求 MA(1) 模型:

$$x_t = \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

和MA(2) 模型:

$$x_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}$$

的自相关函数.

解 :MA(1) 模型和 MA(2) 模型的自相关函数分别为

$$\rho(k) = \begin{cases} 1, k = 0; \\ -\frac{\theta_1}{1 + \theta_1^2}; k = 1; \\ 0, k \geq 2. \end{cases} \quad \text{和} \quad \rho(k) = \begin{cases} 1, k = 0; \\ \frac{-\theta_1 + \theta_2}{1 + \theta_1^2 + \theta_2^2}; k = 1; \\ \frac{-\theta_2}{1 + \theta_1^2 + \theta_2^2}; k = 2; \\ 0; k \geq 3. \end{cases}$$

移动平均模型的统计性质

● 5. 逆函数

自回归模型传递形式的实质是用过去和现在的干扰项表示当前序列值, 其系数就是 Green函数. 对于一个移动平均模型来讲, 我们也可以用现在和过去的序列值表示当前干扰项, 即

$$\varepsilon_t = I(B)x_t = \left(\sum_{i=0}^{\infty} I_i B^i \right) x_t \quad (2.26)$$

我们称 (2.26) 式为平均移动模型的**逆转形式**, 并称系数

$$I_0 = 1, I_i (i = 1, 2, \dots)$$

为**逆函数**. 然而, 并不是所有移动平均模型都可以写成逆转形式.



移动平均模型的统计性质

例 2.12 考察下列两个 MA(1)

$$x_t = \varepsilon_t - \theta \varepsilon_{t-1} \quad (2.27)$$

和

$$x_t = \varepsilon_t - \frac{1}{\theta} \varepsilon_{t-1} \quad (2.28)$$

易见模型 (2.27) 和模型 (2.28) 的自相关函数相等. 将它们分别写成自相关模型形式

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

和

$$\frac{x_t}{1 - (1/\theta)B} = \varepsilon_t.$$

容易看出, 如果 $|\theta| < 1$, 那么 $\sum_{i=0}^{\infty} \theta^i B^i = \frac{1}{1 - \theta B}$, 即模型 (2.27) 具有逆转形式

$\varepsilon_t = \sum_{i=0}^{\infty} \theta^i B^i x_t$, 而此时 $\sum_{i=0}^{\infty} \theta^{-i} B^i$ 发散, 模型 (2.28) 不具有逆转形式.



移动平均模型的统计性质

- 一般地, 当一个 $MA(q)$ 模型具有逆转形式, 我们也称该模型为**可逆的**; 否则, 称该模型为**不可逆的**. 通常情况下, 不同的 $MA(q)$ 模型可以有相同的自相关函数, 但是对于可逆的 $MA(q)$ 模型来讲, 其自相关函数与该模型是一一对应的. 下面我们分析移动平均模型可逆的条件.

- 将 $MA(q)$ 模型 (2.24) 表示为 $\varepsilon_t = \frac{x_t}{\Theta(B)}$, (2.29)

其中, $\Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$, 为 q 阶移动平均系数多项式. 设 $1/\lambda_i, i=1, 2, \dots, q$ 是系数多项式 $\Theta(B)$ 的 q 个根, 则 (2.29) 式可以表示为

$$\varepsilon_t = \frac{x_t}{(1 - \lambda_1 B)(1 - \lambda_2 B) \dots (1 - \lambda_q B)}.$$

容易看出, 模型 (2.29) 具有逆转形式当且仅当 $|\lambda_i| < 1, i=1, 2, \dots, q$, 即系数多项式 $\Theta(B)$ 的 q 个根 $1/\lambda_i, i=1, 2, \dots, q$ 在单位圆外. 这个条件我们称为 $MA(q)$ 模型的**可逆性条件**.



移动平均模型的统计性质

■ 例 2.13 :写出 MA(2) 模型

$$x_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}$$

的可逆性条件.

- 当 MA(q) 模型 (2.24) 满足可逆性条件时, 它可以写成逆转形式 (2.26) 式. 我们将 (2.26) 式代入模型 (2.24), 得

$$\Theta(B)I(B)x_t = x_t$$

将上式展开得

$$\left(1 - \sum_{k=1}^q \theta_k B^k\right) \left(1 + \sum_{i=1}^{\infty} \theta_i B^i\right) x_t = x_t$$

由待定系数法得逆函数的递推公式为

$$\begin{cases} I_0 = 1; \\ I_i = \sum_{k=1}^i \tilde{\theta}_k I_{i-k}, i \geq 1. \end{cases} \quad \text{其中} \quad \tilde{\theta}_k = \begin{cases} \theta_k, k \leq q; \\ 0, k > q. \end{cases}$$



移动平均模型的统计性质

■ 例 2.14 : 判断模型

$$x_t = \varepsilon_t - 0.8\varepsilon_{t-1} + 0.64\varepsilon_{t-2}$$

的可逆性. 如果可逆, 那么写出该模型的逆转形式.

解: 根据例 2.13 , 可知该模型可逆. 再根据逆函数的递推公式,

$$\begin{cases} I_0 = 1; \\ I_i = \sum_{k=1}^i \tilde{\theta}_k I_{i-k}, i \geq 1. \end{cases} \quad \tilde{\theta}_k = \begin{cases} \theta_k, k \leq q; \\ 0, k > q. \end{cases}$$

得逆函数为
$$I_k = \begin{cases} (-1)^n 0.8^k, k = 3n \text{ 或 } 3n+1, n=0,1,\dots; \\ 0, k=3n+2. \end{cases}$$

从而得该模型的逆转形式.



移动平均模型的统计性质

6. 偏自相关函数的拖尾性

对于可逆的 $MA(q)$ 模型而言, 其逆转形式实质上是个 $AR(\infty)$ 模型. 于是根据 AR 模型偏自相关函数的截尾性知, 可逆的 $MA(q)$ 模型偏自相关函数 ∞ 截尾, 即其偏自相关函数具有拖尾性.

■ 例 2.15 求 $MA(1)$ 模型:

$$x_t = \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

的偏自相关函数的表达式.

解 由偏自相关函数的求法可知, 延迟 k 阶偏自相关函数是 Yule-Walker 方程 (2.21) 的解的最后一个分量 ϕ_{kk} , 于是根据 (2.22) 式有

$$\phi_{11} = \rho_1 = \frac{-\theta_1}{1+\theta_1^2}; \phi_{22} = \frac{-\rho_1^2}{1-\rho_1^2} = \frac{-\theta_1^2}{1+\theta_1^2+\theta_1^4}; \dots; \phi_{kk} = \frac{-\theta_1^{2k}}{1+\theta_1^2+\theta_1^4+\dots+\theta_1^{2k}}$$

从上述 $MA(1)$ 模型的偏自相关函数的表达式, 我们看到其偏自相关函数具有拖尾性. 另外, 从 $MA(q)$ 模型偏自相关函数图, 也可观察到其偏自相关函数具有拖尾性.



移动平均模型的统计性质

$$\phi_{11}=\rho_1=\frac{-\theta_1}{1+\theta_1^2}; \phi_{22}=\frac{-\rho_1^2}{1-\rho_1^2}=\frac{-\theta_1^2}{1+\theta_1^2+\theta_1^4}; \dots; \phi_{kk}=\frac{-\theta_1^2}{1+\theta_1^2+\theta_1^4+\dots+\theta_1^{2k}}$$

从上述 MA(1) 模型的偏自相关函数的表达式, 我们看到其偏自相关函数具有拖尾性. 另外, 从 MA(q) 模型偏自相关函数图, 也可观察到其偏自相关函数具有拖尾性.

- **例 2.16** 绘制下列 MA(1), MA(2) 模型的偏自相关函数图, 并观察 MA(1), MA(2) 模型偏自相关函数的拖尾性.

$$(1) x_t = \varepsilon_t - 0.5\varepsilon_{t-1}; (2) x_t = \varepsilon_t - 0.25\varepsilon_{t-1} + 0.5\varepsilon_{t-2}$$

其中, ε_t 是标准正态白噪声序列.



移动平均模型的统计性质

例 2.16 绘制下列 MA(1), MA(2) 模型的偏自相关函数图, 并观察 MA(1), MA(2) 模型偏自相关函数的拖尾性.

$$(1) x_t = \varepsilon_t - 0.5\varepsilon_{t-1}; (2) x_t = \varepsilon_t - 0.25\varepsilon_{t-1} + 0.5\varepsilon_{t-2}$$

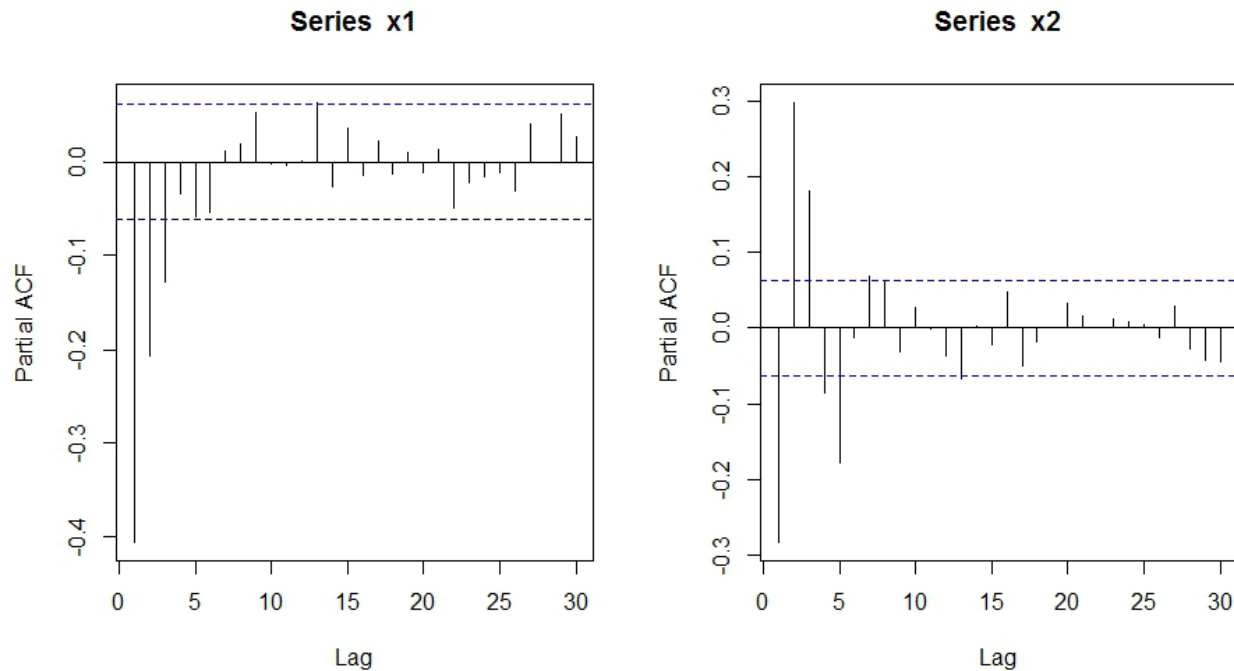
其中, ε_t 是标准正态白噪声序列.

解: 应用下列 R 语言命令绘制 MA(1), MA(2) 模型的偏自相关函数, 运行结果如 2.4所示.

```
> x1 <- arima.sim(n=1000,list(ma=-0.5))  
> x2 <- arima.sim(n=1000,list(ma=c(-0.25,0.5)))  
> par(mfrow=c(1,2))  
> pacf(x1)  
> pacf(x2)
```



移动平均模型的统计性质



由图 2.4 可见, 模型 (1) 和模型 (2) 均具有拖尾性.

1. 差分方程和滞后算子

2. 自回归模型的概念和性质

3. 移动平均模型的概念和性质

4. 自回归移动平均模型的概念和性质

自回归移动平均模型的定义

- 设 $\{x_t, t \in T\}$ 是一个时间序列, 称满足如下结构的模型为自回归移动平均 (autoregressive moving average, ARMA) 模型, 简记为 ARMA(p,q),

$$x_t = \phi_0 + \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}, \quad (2.30)$$

其中, $\phi_p \neq 0, \theta_q \neq 0$, 并且 $\{\varepsilon_t\}$ 是均值为零的白噪声序列, 且 ε_t 与 $x_{t-j}, j=1, 2, \dots$ 无关, 即 $\forall s < t, E(x_s \varepsilon_t) = 0$.

若 $\phi_0 = 0$, 该模型称为 中心化 ARMA(p,q) 模型. 由于模型 (2.30) 总可以中心化, 而且中心化后并不影响序列值之间的相关关系, 所以以下研究的自回归移动平均模型, 如果不做特殊约定, 我们自动默认为中心化自回归移动平均模型.



自回归移动平均模型的定义

- 借助于延迟算子, ARMA(p,q) 模型可简记为

$$\Phi(B)x_t = \Theta(B)\varepsilon_t, \quad (2.31)$$

其中,

$\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$, 为 p 阶自回归系数多项式;

$\Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$, 为 q 阶移动平均系数多项式.

- 注: (1) 模型中要求 $\Phi(B)$ 与 $\Theta(B)$ 没有公共因子.

2) 当 $q=0$ 时, ARMA(p,q) 模型就退化成了 AR(p) 模型; 当 $p=0$ 时, ARMA(p,q) 模型就退化成了 MA(q) 模型. 因此, AR(p) 模型和 MA(q) 模型是 ARMA(p,q) 模型的特例, 它们都统称为 ARMA 模型. ARMA(p,q) 模型的统计性质由 AR(p) 模型和 MA(q) 模型的统计性质共同决定

平稳性与可逆性

- 平稳性分析

对于 ARMA(p,q) 模型 (2.31) 来讲, 我们记

$$y_t = \Theta(B)\varepsilon_t,$$

则序列 $\{y_t\}$ 是均值为零, 方差为固定常数的平稳序列. 于是 ARMA(p,q) 模型也可表示为

$$\Phi(B)x_t = y_t.$$

因此, 类似于 AR(p) 模型的平稳性分析, 我们可推得 ARMA(p,q) 模型的平稳性条件是

$$\Phi(B)=0$$

的根都在单位圆外. 可见, ARMA(p,q) 模型的平稳性完全由其自回归部分的平稳性所决定.



平稳性与可逆性

- 可逆性分析

同样地,对于 ARMA(p,q) 模型 (2.31) 来讲,它的可逆性完全由其移动平均部分决定,即ARMA(p,q) 模型的可逆条件是

$$\Theta(B)=0$$

的根都在单位圆外.

综上所述,当 $\Phi(B)=0$ 和 $\Theta(B)=0$ 的根都在单位圆外时, ARMA(p,q) 模型是一个平稳可逆模型.



Green 函数与逆函数

- Green 函数

对于平稳可逆ARMA(p, q)模型(2.31), 它具有如下传递形式

$$x_t = \Phi(B)^{-1} \Theta(B) \varepsilon_t = \sum_{i=0}^{\infty} G_i \varepsilon_{t-i}$$

其中 $G_i (i \geq 0)$ 就是 Green 函数.

通过待定系数法, 可以推得ARMA(p, q)模型 (2.31) 的Green 函数的递推公式

$$\begin{cases} G_0 = 1; \\ G_k = \sum_{i=1}^k \phi'_i G_{k-i} - \theta'_k, \quad k \geq 1, \end{cases}$$

式中

$$\phi'_i = \begin{cases} \phi_i, & 1 \leq i \leq p; \\ 0, & i > p, \end{cases} \quad \text{且} \quad \theta'_k = \begin{cases} \theta_k, & 1 \leq k \leq q; \\ 0, & k > q. \end{cases}$$



Green 函数与逆函数

- 逆函数

对于平稳可逆ARMA(p, q)模型(2.31), 它具有如下逆转形式

$$\varepsilon_t = \Theta(B)^{-1} \Phi(B) x_t = \sum_{i=0}^{\infty} I_i x_{t-i}$$

其中 $I_i (i \geq 0)$ 就是逆函数.

通过待定系数法, 可以推得ARMA(p, q)模型 (2.31) 的逆函数的递推公式

$$\begin{cases} I_0 = 1; \\ I_k = \sum_{i=1}^k \theta'_i I_{k-i} - \phi'_k, \quad k \geq 1, \end{cases}$$

式中

$$\phi'_i = \begin{cases} \phi_i, & 1 \leq i \leq p; \\ 0, & i > p, \end{cases} \quad \text{且} \quad \theta'_k = \begin{cases} \theta_k, & 1 \leq k \leq q; \\ 0, & k > q. \end{cases}$$

ARMA(p,q) 模型的统计性质

- 均值

在平稳可逆 ARMA(p,q) 模型 (2.31),

$$x_t = \phi_0 + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}$$

两边求均值, 得

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

- 自协方差函数

$$\begin{aligned} \gamma(k) &= E(x_t x_{t+k}) = E \left[\left(\sum_{i=0}^{\infty} G_i \varepsilon_{t-i} \right) \left(\sum_{j=0}^{\infty} G_j \varepsilon_{t+k-j} \right) \right] = E \left(\sum_{i=0}^{\infty} G_i \sum_{j=0}^{\infty} G_j \varepsilon_{t-i} \varepsilon_{t+k-j} \right) \\ &= \sigma_{\varepsilon}^2 \sum_{i=0}^{\infty} G_i G_{i+k}. \end{aligned}$$

ARMA(p,q) 模型的统计性质

- 自协方差函数

$$\begin{aligned}\gamma(k) &= E(x_t x_{t+k}) = E \left[\left(\sum_{i=0}^{\infty} G_i \varepsilon_{t-i} \right) \left(\sum_{j=0}^{\infty} G_j \varepsilon_{t+k-j} \right) \right] = E \left(\sum_{i=0}^{\infty} G_i \sum_{j=0}^{\infty} G_j \varepsilon_{t-i} \varepsilon_{t+k-j} \right) \\ &= \sigma_{\varepsilon}^2 \sum_{i=0}^{\infty} G_i G_{i+k}.\end{aligned}$$

- 自相关函数

$$\rho(k) = \frac{\gamma(k)}{\gamma(0)} = \frac{\sum_{i=0}^{\infty} G_i G_{i+k}}{\sum_{i=0}^{\infty} G_i^2}.$$

由上式我们看出,ARMA(p,q) 模型的自相关函数拖尾.这是由于ARMA(p,q) 模型可以转化为无穷阶移动平均模型. 同样地,ARMA(p,q) 模型也可以转化为无穷阶自回归模型, 因此,ARMA(p,q)模型的偏自相关函数也拖尾.

ARMA(p,q) 模型的统计性质

- 例 2.17 绘制 ARMA(1,2) 模型:

$$x_t = 0.8x_{t-1} + \varepsilon_t - 0.8\varepsilon_{t-1} + 0.64\varepsilon_{t-2}$$

的自相关函数图和偏自相关函数图, 并观察它们的拖尾性, 其中 $\{\varepsilon_t\}$ 为标准正态白噪声序列.

解: 用下列 R 语言命令, 分别绘制自相关函数图和偏自相关函数图.

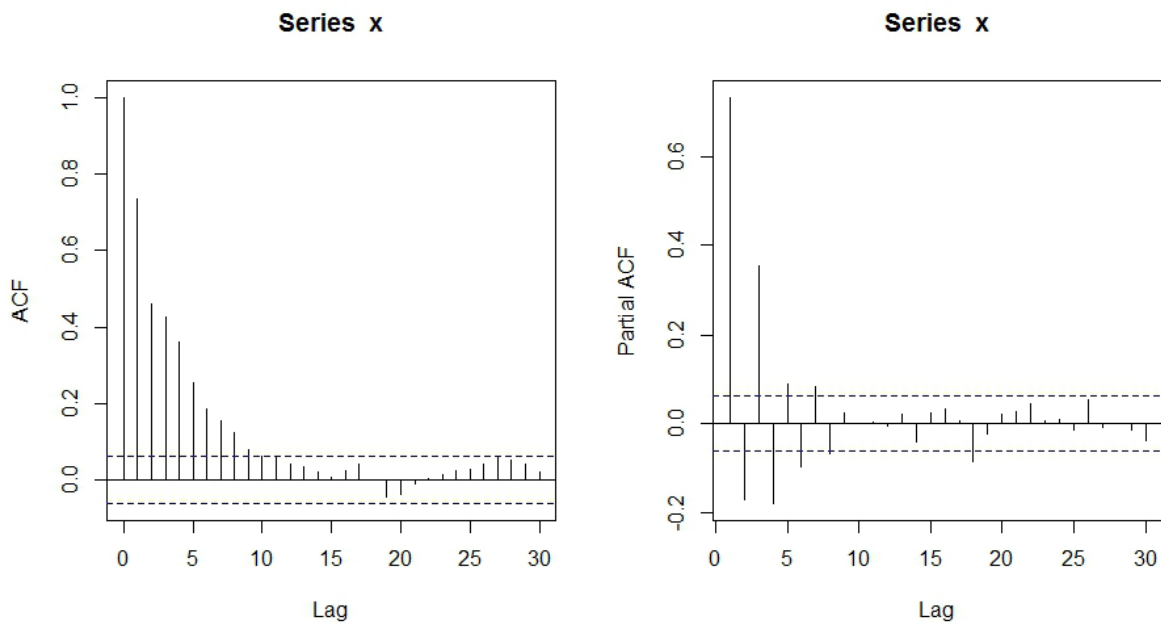
```
> x <- arima.sim(n=1000,list(ar=0.8,ma=c(0.8,-0.64)))  
> par(mfrow=c(1,2))  
> acf(x)  
> pacf(x)
```

运行结果如下图所示.



ARMA(p,q) 模型的统计性质

- 运行结果如下图所示.



由上可见, 该模型自相关函数图和偏自相关函数图都具有拖尾性.