

平稳序列拟合与预测

04

本章内容

01

建模步骤

02

样本自相关系数与偏自相关系数特征

03

参数估计

04

模型检验

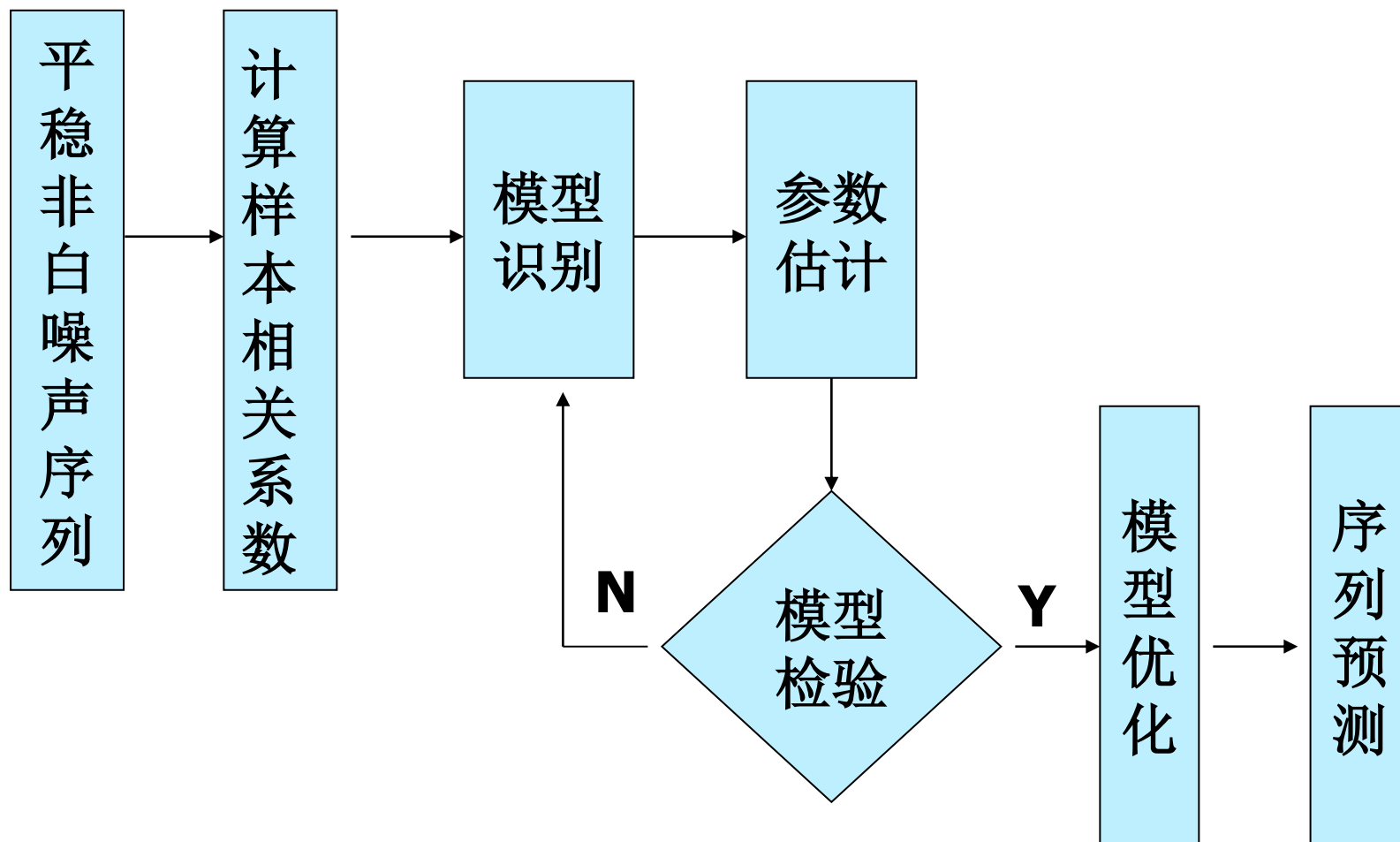
05

模型优化

06

序列预测

建模步骤



本章内容

01

建模步骤

02

样本自相关系数与偏自相关系数特征

03

参数估计

04

模型检验

05

模型优化

06

序列预测

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

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

平稳序列拟合模型识别

自相关系数	偏自相关系数	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 $\hat{\rho}_k$ 或 $\hat{\phi}_{kk}$ 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 $k \rightarrow \infty$ ， $\hat{\rho}_k$ 与 $\hat{\phi}_{kk}$ 都会衰减至零值附近作小值波动
- 当 $\hat{\rho}_k$ 或 $\hat{\phi}_{kk}$ 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？
- 这实际上没有绝对的标准，在很大程度上依靠分析人员的主观经验。但样本自相关系数和偏自相关系数的近似分布可以帮助缺乏经验的分析人员做出尽量合理的判断。

样本相关系数的近似分布

- Barlett

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

- Quenouille

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

模型定阶经验方法

- 样本自相关系数和样本偏自相关系数的95 %置信区间

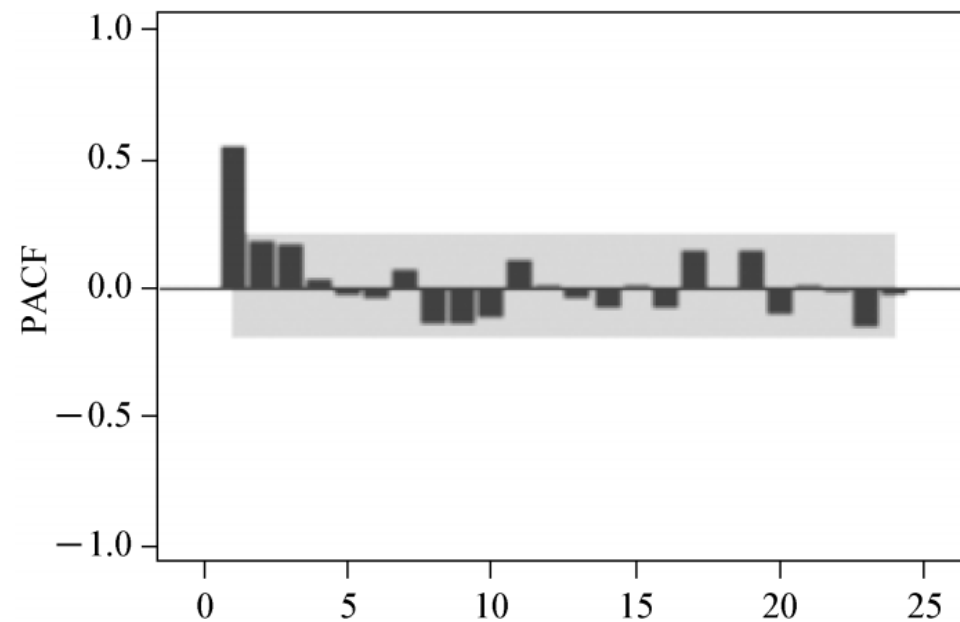
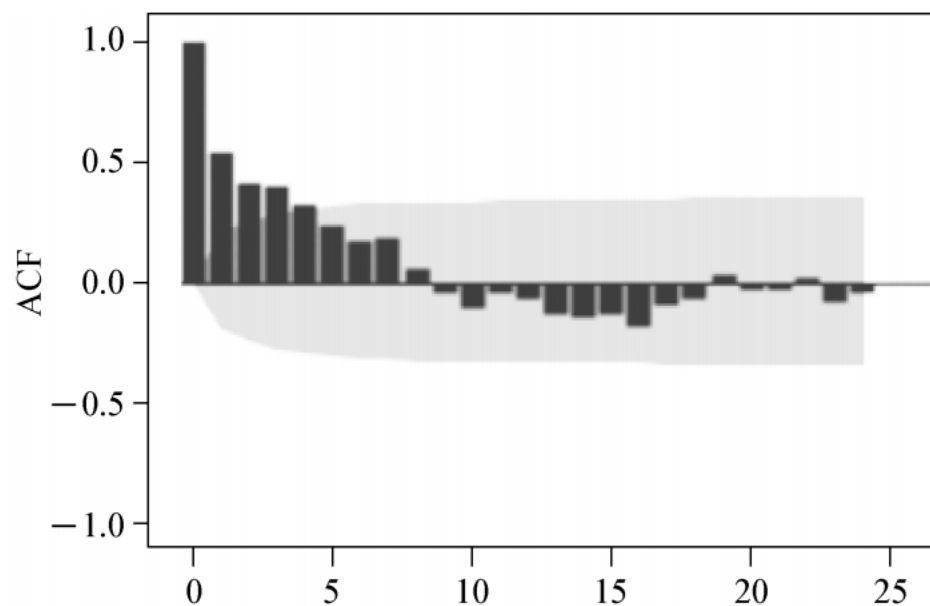
$$\Pr\left(-\frac{2}{\sqrt{n}} \leq \hat{\rho}_k \leq \frac{2}{\sqrt{n}}\right) \geq 0.95$$

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

- 模型定阶的经验方法
 - 如果样本自相关系数（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，通常视为自相关系数（偏自相关系数）截尾。截尾阶数为d。
 - 如果有超过5 %的样本自相关系数（偏自相关系数）落入2倍标准差范围之外，或者由显著非零的自相关系数（偏自相关系数）衰减为小值波动的过程比较缓慢或者非常连续，这时，通常视为自相关系数拖尾。

例4-1

- 选择合适的模型拟合1900—1998年全球7.0级以上地震年发生次数序列。
 - 在例2-6的分析中，我们已经判断该序列是平稳非白噪声序列。现在考察该序列的自相关图和偏自相关图，给该序列的拟合模型定阶

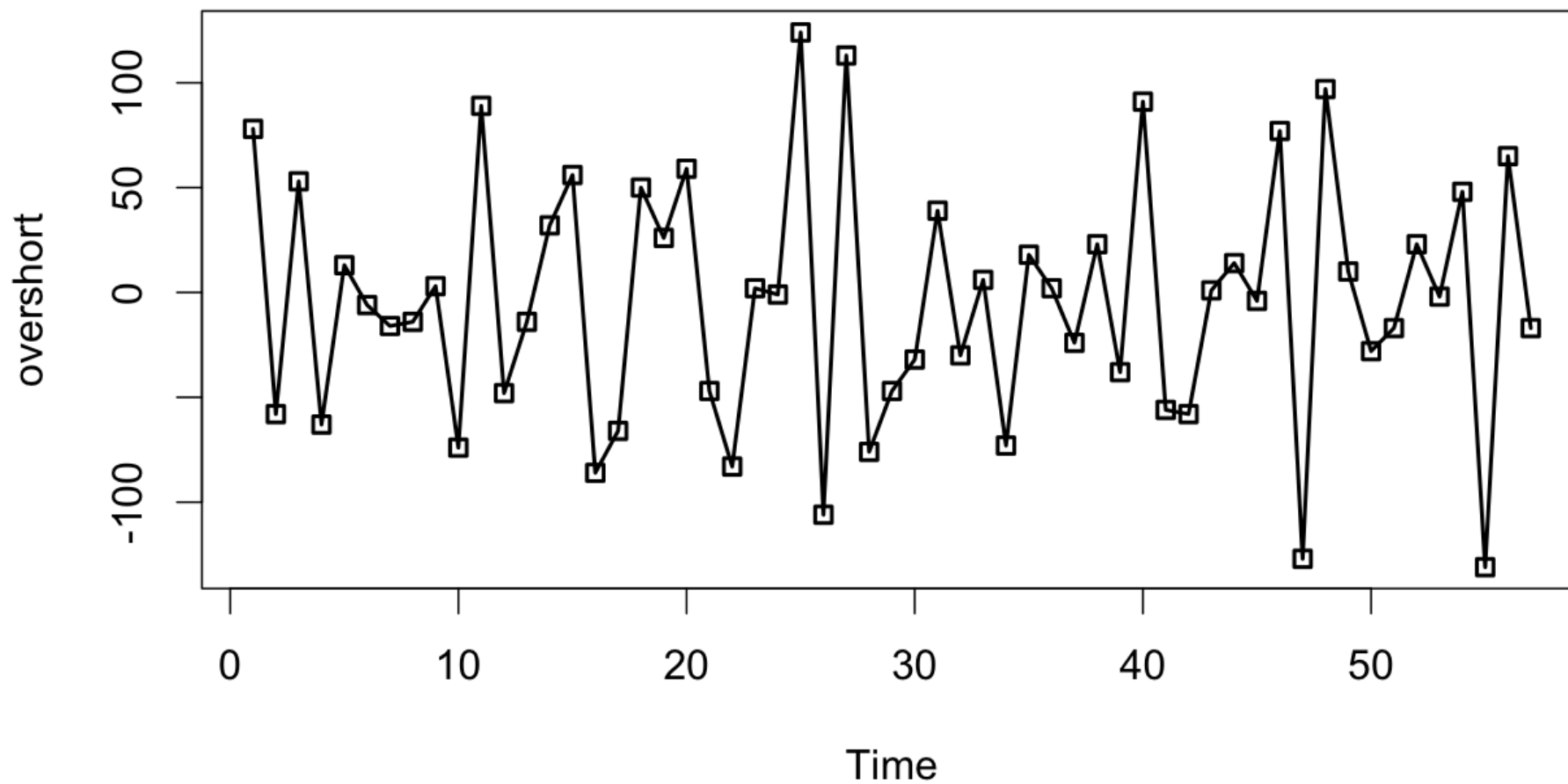


例4-1模型定阶

- 从自相关图可以看出，自相关系数是以一种有规律的方式，按指数函数轨迹衰减的，这说明自相关系数衰减到零不是一个突然截尾的过程，而是一个连续渐变的过程，这时自相关系数拖尾的典型特征，我们可以把拖尾特征形象地描述为 “坐着滑梯落水”。
- 从偏自相关图可以看出，除了1阶偏自相关系数在2倍标准差范围之外，其他阶数的偏自相关系数都在2倍标准差范围内，这是一个偏自相关系数1阶截尾的典型特征。我们可以把这种截尾特征形象地描述为 “1阶之后高台跳水”。
- 本例中，根据自相关系数拖尾，偏自相关系数1阶截尾的属性，我们可以初步确定拟合模型为 $AR(1)$ 模型。

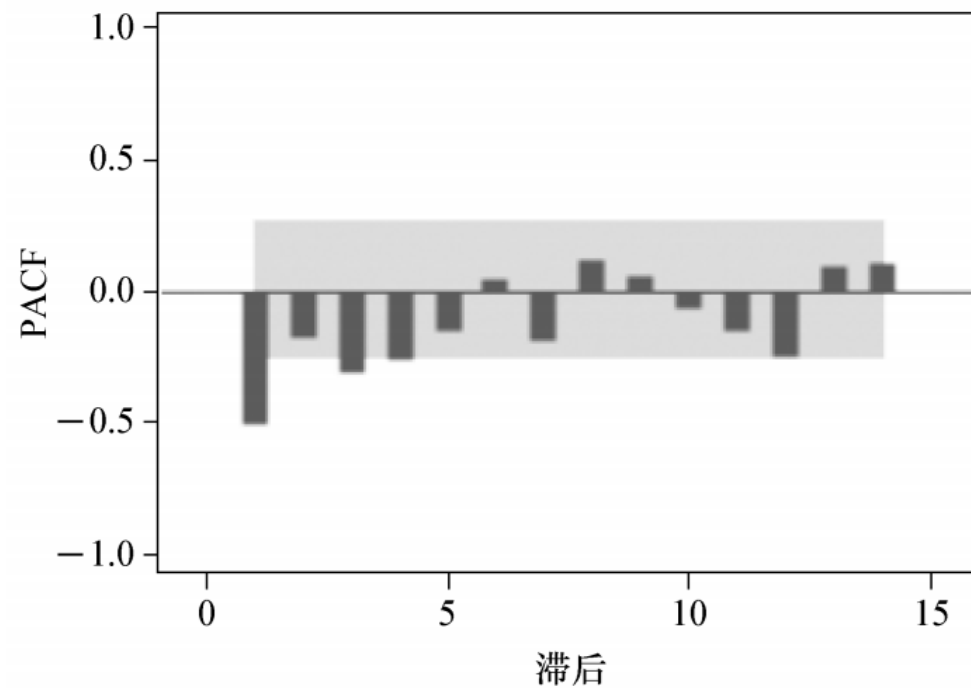
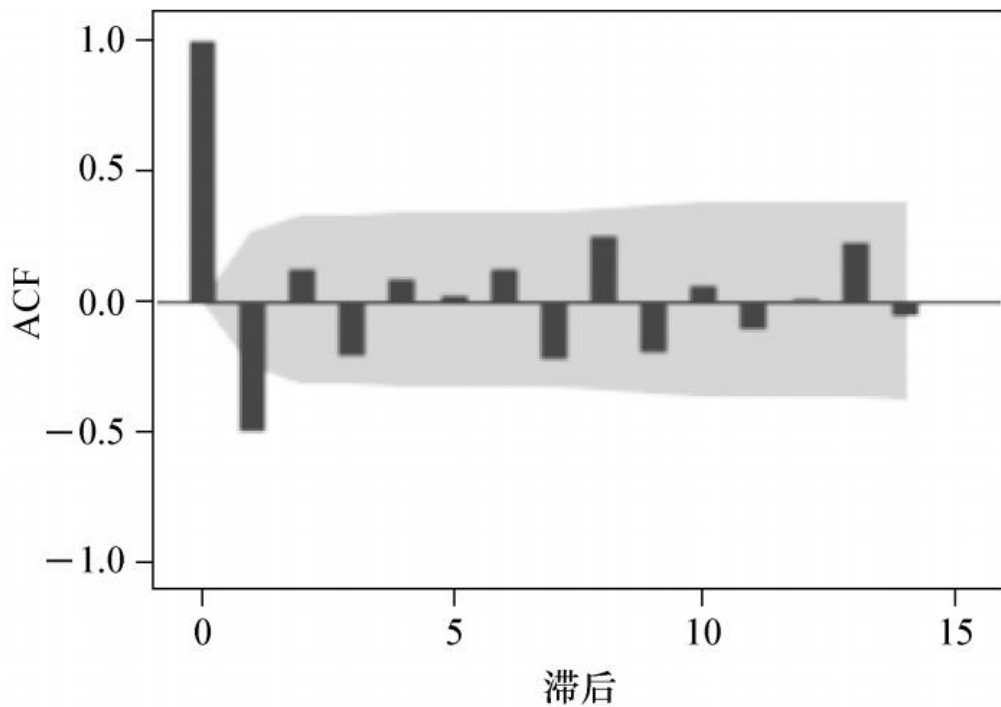
例4-2

- 选择合适的模型拟合美国科罗拉多州某一加油站连续57天的每日盈亏序列



序列自相关图和偏自相关图

- 对序列进行ADF检验和白噪声检验，检验结果显示该序列为平稳非白噪声序列。现在考察该序列的自相关图和偏自相关图，给该序列的拟合模型定阶

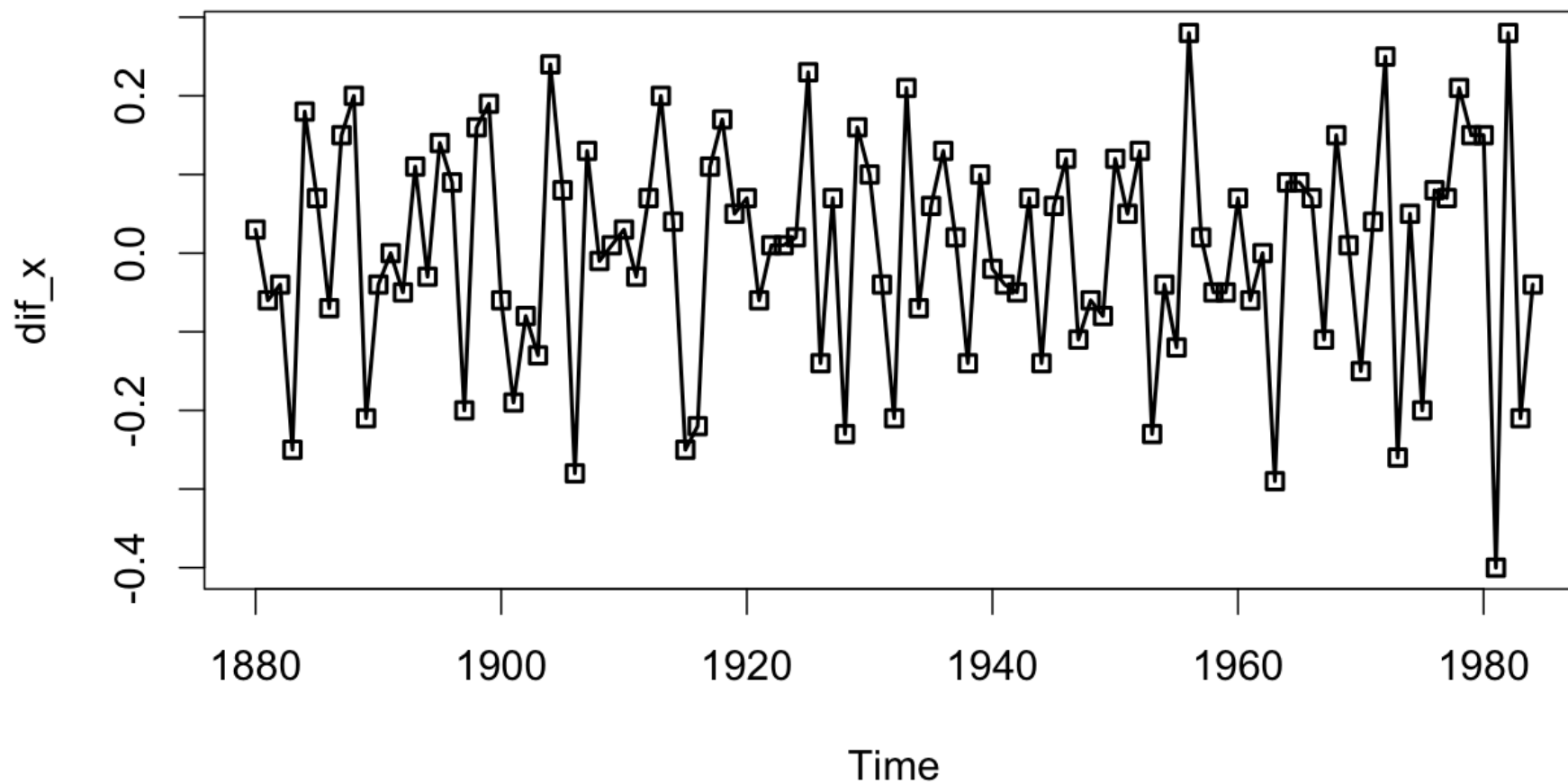


例4-2模型定阶

- 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)。

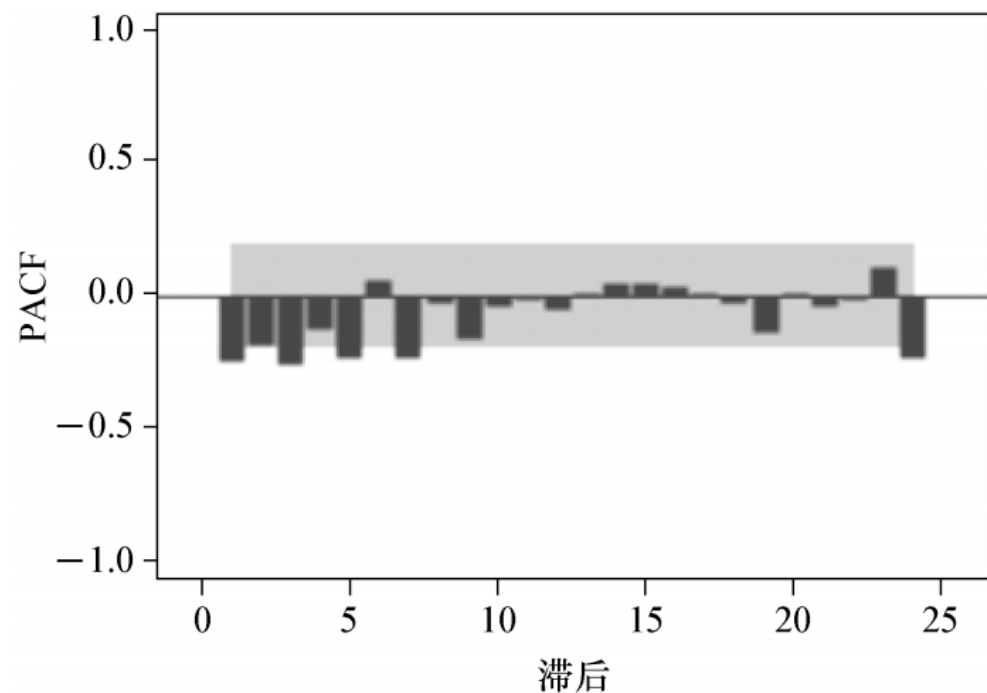
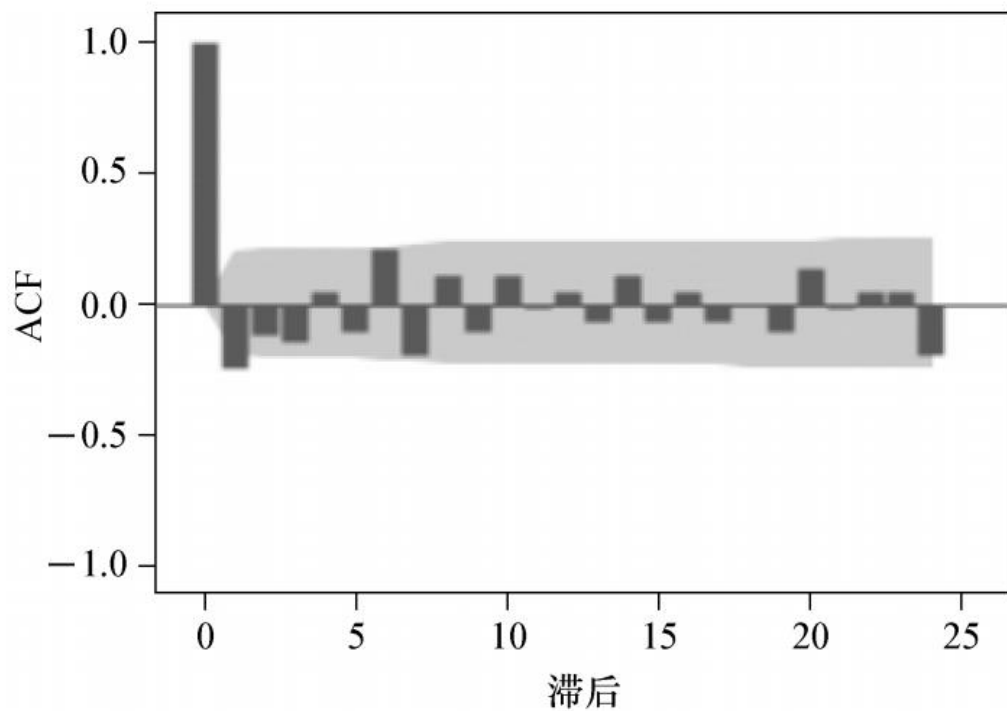
例4-3

- 选择合适的模型拟合1880-1985全球气表平均温度改变值差分序列



序列自相关图和偏自相关图

- 对序列进行ADF检验和白噪声检验，检验结果显示该序列为平稳非白噪声序列。现在考察该序列的自相关图和偏自相关图，给该序列的拟合模型定阶



例4-3模型定阶

- 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。

本章内容

01

建模步骤

02

样本自相关系数与偏自相关系数特征

03

参数估计

04

模型检验

05

模型优化

06

序列预测

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

$$\hat{\mu} = \bar{x} = \frac{\sum_{i=1}^n x_i}{n} \qquad \hat{\sigma}_{\epsilon}^2 = \frac{\sum_{t=1}^n \epsilon_t^2}{n}$$

例4-4

- 求AR(2)模型 $x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \varepsilon_t$ 系数的矩估计
 - AR(2)模型的Yule-Walker方程

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

- 用样本自相关系数代入Yule-Walker方程，得到AR(2)模型参数的矩估计

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

例4-5

- 求MA(1)模型 $x_t = \varepsilon_t - \theta_1 \varepsilon_{t-1}$ 系数的矩估计
 - MA(1)模型的Yule-Walker方程

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

- 用样本自相关系数代入Yule-Walker方程，得到MA(1)模型参数的矩估计矩估计

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

例4-6

- 求ARMA(1,1)模型 $x_t = \phi_1 x_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$ 系数的矩估计
 - ARMA(1,1)模型的Yule-Walker方程

$$\begin{cases} \rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{(\phi_1 - \theta_1)(1 - \theta_1 \phi_1)}{1 + \theta_1^2 - 2\theta_1 \phi_1} \\ \rho_2 = \phi_1 \rho_1 \end{cases}$$

- 用样本自相关系数代入Yule-Walker方程，得到ARMA(1,1)模型参数的矩估计

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

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

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

- 似然方程组

- 似然方程组实际上是由 $p+q+1$ 个超越方程构成，需要使用迭代算法求出未知参数

$\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q, \sigma_\varepsilon^2$ 的极大似然估计值

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

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

- 令 $F_t(\tilde{\beta}) = \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} - \theta_1 \epsilon_{t-1} - \cdots - \theta_q \epsilon_{t-q}$

- 残差项为

$$\epsilon_t = x_t - F_t(\tilde{\beta})$$

- 残差平方和为

$$\begin{aligned} Q(\tilde{\beta}) &= \sum_{t=1}^n \epsilon_t^2 \\ &= \sum_{t=1}^n (x_t - \phi_1 x_{t-1} - \cdots - \phi_p x_{t-p} + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q})^2 \end{aligned}$$

- 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

最小二乘估计的特征与评价

- 由于随机扰动 $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots$ 不可观测，所以 $Q(\tilde{\beta})$ 也不是 $\tilde{\beta}$ 的显性函数，未知参数的最小二乘估计值通常也得借助迭代法求出。
- 在实际中，最常用的是条件最小二乘估计方法。
 - 它假定过去未观测到的序列值等于零，即 $x_t = 0, t \leq 0$ ，这个假定条件下进行的最小二乘估计称为条件最小二乘估计。
- 最小二乘估计方法的优点
 - 原理简单，方法普适，估计精度高

例4-1续

- 使用极大似然估计方法确定 1 9 0 0 — 1 9 9 8 年全球 7 级以上地震发生次数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：

$$x_t = 9.085 + 0.543x_{t-1} + \epsilon_t, \text{Var}(\epsilon_t) = 37.454$$

例4-2续

- 确定美国科罗拉多州某一加油站连续57天的每日盈亏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：

$$x_t = -4.409 + \epsilon_t - 0.821\epsilon_{t-1}, \text{Var}(\epsilon_t) = 2181.637$$

例4-3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：

$$x_t = 0.003 + 0.407x_{t-1} + \epsilon_t - 0.899\epsilon_{t-1}, \text{Var}(\epsilon_t) = 0.016$$

本章内容

01

建模步骤

02

样本自相关系数与偏自相关系数特征

03

参数估计

04

模型检验

05

模型优化

06

序列预测

模型检验

- 对序列进行模型拟合之后，我们还要对该拟合模型进行必要的检验。
- 检验内容
 - 模型的显著性检验
 - 确保序列中蕴含的相关信息被充分提取，拟合模型的残差序列必须是白噪声序列
 - 参数的显著性检验
 - 确保拟合模型的精简，每个保留在拟合模型中的参数必须显著非零

模型的显著性检验

- 目的
 - 检验拟合模型的有效性（对相关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

模型显著性检验的假设条件

- 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

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

例4-1续

- 检验 1 9 0 0 — 1 9 9 8 年全球 7 级以上地震发生次数序列拟合模型的显著性 ($\alpha = 0.05$)
- 残差白噪声序列检验结果

延迟阶数	纯随机性检验	
	LB 检验统计量的值	P 值
6	5.25	0.386 2
12	10.52	0.484 3
18	15.49	0.560 5

- 由于各阶延迟下 L B 统计量的 P 值都显著大于 0.05，可以认为这个拟合模型的残差序列属于白噪声序列，即该拟合模型显著有效。

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4-1续

- 检验 1 9 0 0 — 1 9 9 8 年全球 7 级以上地震发生次数序列拟合模型的参数显著性 ($\alpha = 0.05$)
- 参数显著性检验结果

参数	t 统计量的值	P 值	结论
μ	14.97	$<0.000\ 1$	显著非零
ϕ_1	6.4	$<0.000\ 1$	显著非零

例4-2续

- 对美国科罗拉多州某一加油站连续57天的每日盈亏序列拟合模型进行检验
 - 模型显著性检验

延迟阶数	LB 统计量的值	P 值	结论
6	3.14	0.678 0	拟合模型 显著有效
12	9.10	0.613 0	

- 参数显著性检验

估计方法	均值的检验		θ_1 的检验		结论
	t 统计量的值	P 值	t 统计量的值	P 值	
条件最小二乘	-3.71	<0.000 5	10.53	<0.000 1	两参数检验 均显著

例4-3续

- 对 1 8 8 0 — 1 9 8 5 年全球气表平均温度改变值差分序列拟合模型进行检验
 - 模型显著性检验

延迟阶数	LB 统计量的值	P 值	结论
6	5.11	0.276 3	拟合模型 显著有效
12	10.21	0.422 1	
18	13.27	0.652 6	

- 参数显著性检验

估计方法	θ_1 的检验		ϕ_1 的检验		结论
	t 统计量的值	P 值	t 统计量的值	P 值	
条件最小二乘	16.16	$<0.000\ 1$	3.49	0.000 7	两参数检验 均显著

本章内容

01

建模步骤

02

样本自相关系数与偏自相关系数特征

03

参数估计

04

模型检验

05

模型优化

06

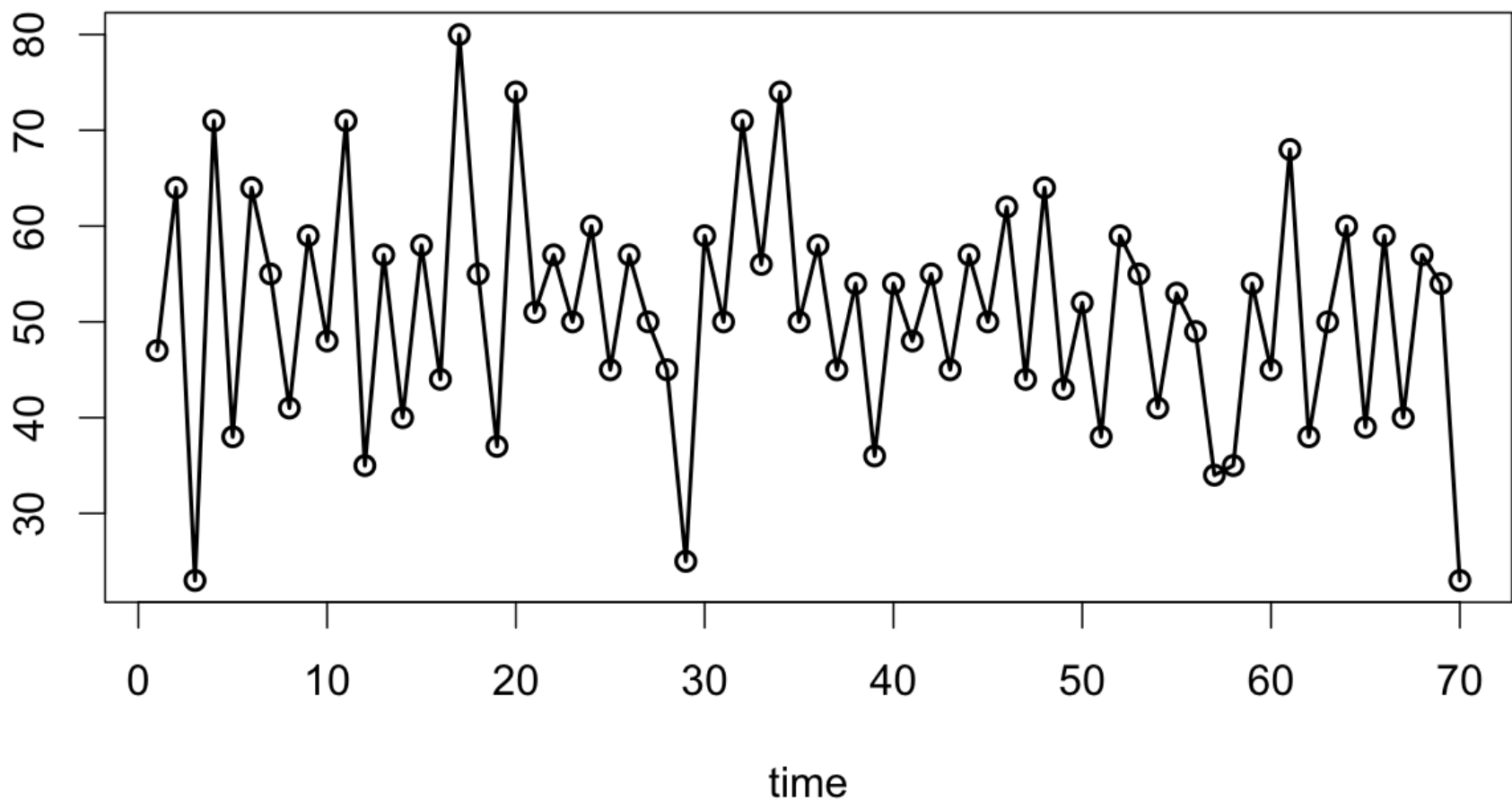
序列预测

模型优化

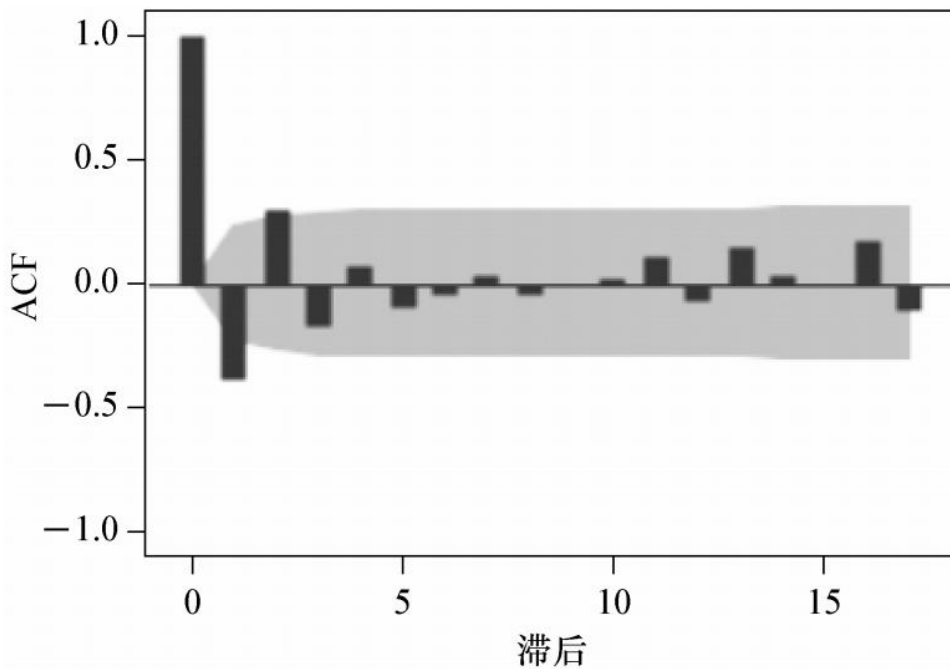
- 问题提出
 - 当一个拟合模型通过了检验，说明在一定的置信水平下，该模型能有效地拟合观察值序列的波动，但这种有效模型并不是唯一的。
- 优化的目的
 - 选择相对最优模型

例4-7

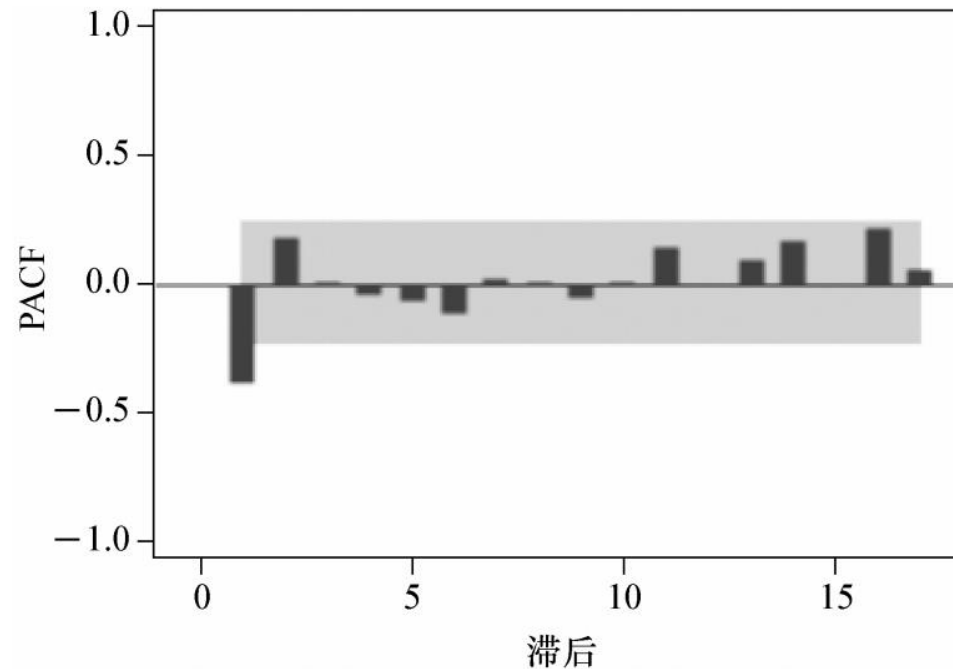
- 等时间间隔连续读取 70 个某次化学反应的过程数据，构成一时间序列。预处理显示该序列为平稳非白噪声序列。



序列的样本自相关图和偏自相关图



根据自相关图 的特征，可能有人会认为自相关系数 2 阶截尾，那么可以对序列拟合 MA(2)模型。



根据偏自相关图 的特征，可能有人会认为偏自相关系数 1 阶截尾，那么可以对序列拟合 AR(1)模型。

拟合模型

- 拟合模型一：根据自相关系数2阶截尾，拟合MA(2)模型

$$x_t = 51.173 + \epsilon_t - 0.323\epsilon_{t-1} + 0.31\epsilon_{t-2}, \text{Var}(\sigma_\epsilon^2) = 119.5653$$

- 拟合模型二：根据自相关系数2阶截尾，拟合AR(1)模型

$$x_t = 73.038 - 0.4248x_{t-1} + \epsilon_t, \text{Var}(\sigma_\epsilon^2) = 120.0735$$

- 模型检验
 - 这两个模型均显著有效
 - 这两个模型的所有参数均显著非零

模型优化标准

- 同一个序列可以构造两个甚至多个拟合模型，每个模型都显著有效，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4-7

- 用AIC准则和SBC准则评判例4-7中两个拟合模型的相对优劣

模型	AIC	SBC
MA(2)	536.455 6	543.201 1
AR(1)	535.789 6	540.286 6

- 结果
 - 最小信息量检验显示，无论是使用 AIC准则还是使用SBC准则，AR(1)模型都要优于 MA(2)模型，所以本例中 AR(1)模型是相对最优模型。

本章内容

01

建模步骤

02

样本自相关系数与偏自相关系数特征

03

参数估计

04

模型检验

05

模型优化

06

序列预测

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

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

线性预测函数

- 根据平稳ARMA模型的可逆性，可以用AR结构表达任意一个平稳ARMA模型

$$\sum_{j=0}^{\infty} I_j x_{t-j} = \varepsilon_t$$

- 其中： I_j ($j=0, 1, 2, \dots$) 为逆函数
- 这意味着使用递推法，基于现有的序列观察值可以预测未来任意时刻的序列值

$$\hat{x}_{t+1} = -I_1 x_t - I_2 x_{t-1} - I_3 x_{t-2} - \dots$$

$$\hat{x}_{t+2} = -I_1 \hat{x}_{t+1} - I_2 x_t - I_3 x_{t-1} - \dots$$

$$\hat{x}_{t+3} = -I_1 \hat{x}_{t+2} - I_2 \hat{x}_{t+1} - I_3 x_t - \dots$$

\vdots

$$\hat{x}_{t+l} = -I_1 \hat{x}_{t+l-1} - I_2 \hat{x}_{t+l-2} - I_3 \hat{x}_{t+l-3} - \dots$$

例4-8

- 假设序列 $\{x_t\}$ 可以用ARMA(1,1)模型 $x_t = 0.8x_t + \varepsilon_t - 0.2\varepsilon_{t-1}$ 拟合, 请确定该序列未来 2 期预测值中第t期和第t-1期序列值的权重。

- 根据拟合模型结构, 求出逆函数

$$I_0 = 1$$

$$I_1 = \theta_1 - \phi_1 = 0.2 - 0.8 = -0.6$$

$$I_2 = \theta_1 I_1 = -0.2 \times 0.6 = -0.12$$

$$I_3 = \theta_1 I_2 = -0.2 \times 0.12 = -0.024$$

- 未来两期递推公式

$$\hat{x}_{t+1} = 0.6x_t + 0.12x_{t-1} + 0.024x_{t-2} - \dots$$

$$\hat{x}_{t+2} = 0.6\hat{x}_{t+1} + 0.12x_t + 0.024x_{t-1} - \dots$$

$$= 0.6(0.6x_t + 0.12x_{t-1} + 0.024x_{t-2} - \dots) + 0.12x_t + 0.024x_{t-1} - \dots$$

$$= (0.36 + 0.12)x_t + (0.072 + 0.024)x_{t-1} + \dots$$

预测方差最小原则

- 预测误差

$$\begin{aligned} e_t(l) &= x_{t+l} - \hat{x}_t(l) \\ &= \sum_{i=0}^{\infty} G_i \epsilon_{t+l-i} - \sum_{i=0}^{\infty} W_i \epsilon_{t-i} = \sum_{i=0}^{l-1} G_i \epsilon_{t+l-i} + \sum_{i=0}^{\infty} (G_{l+i} - W_i) \epsilon_{t-i} \end{aligned}$$

- 预测方差

$$\text{Var}[e_t(l)] = \left[\sum_{i=0}^{l-1} G_i^2 + \sum_{i=0}^{\infty} (G_{l+i} - W_i)^2 \right] \sigma_{\epsilon}^2 \geq \sum_{i=0}^{l-1} G_i^2 \sigma_{\epsilon}^2$$

- 根据预测方差最小原则，得

$$W_i = G_{l+i}, \quad i=0, 1, 2, \dots$$

预测序列分解

$$\begin{aligned} x_{t+l} &= (\varepsilon_{t+l} + G_1 \varepsilon_{t+l-1} + \cdots + G_{l-1} \varepsilon_{t+1}) + (G_l \varepsilon_t + G_{l+1} \varepsilon_{t-1} + \cdots) \\ &= e_t(l) + \hat{x}_t(l) \end{aligned}$$

↓
预测误差
↓

↓
预测值
↓

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

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

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

- 预测方差

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

- 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4-8

- 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（万元）分别为：

101, 96, 97.2

- 请确定该超市第二季度每月销售额的95%的置信区间

例4-8解

- 预测值的计算
4月： $\hat{x}_3(1) = 10 + 0.6x_3 + 0.3x_2 = 97.12$
5月： $\hat{x}_3(2) = 10 + 0.6\hat{x}_3(1) + 0.3x_3 = 97.432$
6月： $\hat{x}_3(3) = 10 + 0.6\hat{x}_3(2) + 0.3\hat{x}_3(1) = 97.5952$
- 预测方差的计算
 $\text{Var}[e_3(1)] = G_0^2 \sigma_\epsilon^2 = 36$
 $\text{Var}[e_3(2)] = (G_0^2 + G_1^2) \sigma_\epsilon^2 = 48.96$
 $\text{Var}[e_3(3)] = (G_0^2 + G_1^2 + G_2^2) \sigma_\epsilon^2 = 64.6416$
- 预测值95%置信区间

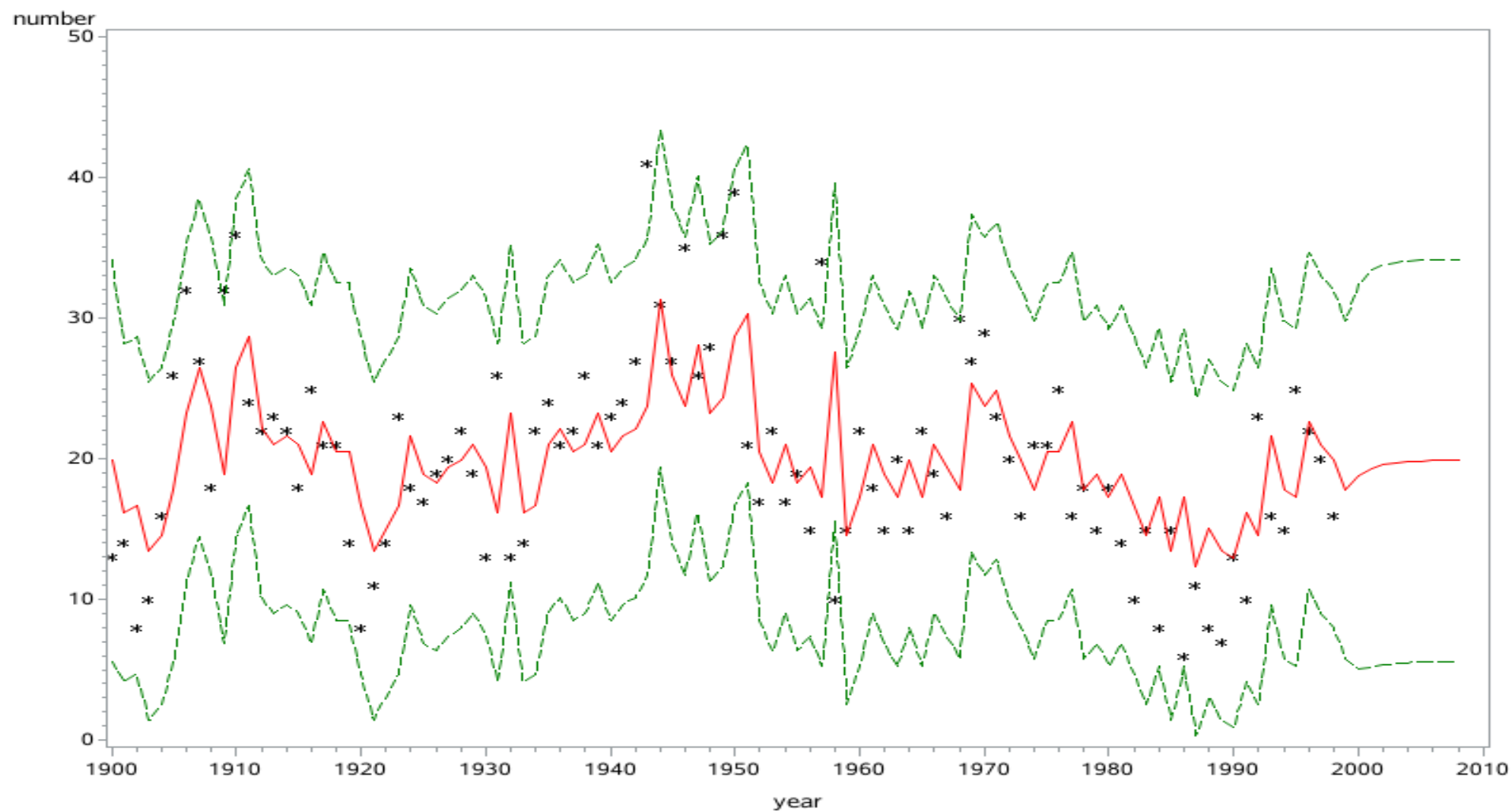
预测时期	95%的置信区间
4月	(85.36, 108.88)
5月	(83.72, 111.15)
6月	(81.84, 113.35)

例4-1续

- 根据1900-1998年全球 7 + 级地震发生次数的观察值，预测1999-2008年全球 7 + 级地震发生次数。

预测年份	预测值	标准差	95%置信下限	95%置信上限
1999	17.777 0	6.120 0	5.782 1	29.771 9
2000	18.742 4	6.964 8	5.091 7	32.393 1
2001	19.266 8	7.195 2	5.164 6	33.369 1
2002	19.551 7	7.261 8	5.319 0	33.784 5
2003	19.706 5	7.281 3	5.435 5	33.977 6
2004	19.790 6	7.287 1	5.508 3	34.073 0
2005	19.836 3	7.288 8	5.550 6	34.122 0
2006	19.861 1	7.289 3	5.574 4	34.147 8
2007	19.874 6	7.289 4	5.587 6	34.161 6
2008	19.881 9	7.289 4	5.594 9	34.169 0

例4-1续:地震序列拟合与预测效果图



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4-10

- 已知某地区每年常驻人口数量近似服从MA(3)模型（单位：万人）：

$$x_t = 100 + \varepsilon_t - 0.8\varepsilon_{t-1} + 0.6\varepsilon_{t-2} - 0.2\varepsilon_{t-3}, \sigma_{\varepsilon}^2 = 25$$

- 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4-10解

- 预测方差的计算

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

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

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

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

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

- 预测值95%置信区间的计算

预测年份	95%的置信区间
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

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

例4-11

- 已知ARMA(1,1)模型为：

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

- 且

$$x_{100} = 0.3 \quad \varepsilon_{100} = 0.01$$

- 预测未来3期序列值的95%的置信区间。

例4-11解

- 预测值的计算

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

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

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

- 预测方差的计算

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

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

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

例4-11解

- 预测值95%置信区间的计算

预测时期	95%的置信区间
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

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

- 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4-9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

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

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

修正结果

- 修正预测值

预测时期	预测值 $\hat{x}_3(l)$	新获得观察值	修正预测 $\hat{x}_4(l-1)$
1	97.12	100	
2	97.43		$\hat{x}_4(1)=G_1 \epsilon_4 + \hat{x}_3(2)=99.16$
3	97.60		$\hat{x}_4(2)=G_2 \epsilon_4 + \hat{x}_3(3)=99.50$

- 修正预测方差

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

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

- 修正置信区间

预测时期	修正前的置信区间	修正后的置信区间
4 月	(85.36, 108.88)	
5 月	(83.72, 111.15)	(87.40, 110.92)
6 月	(81.84, 113.35)	(85.79, 113.21)

THANKS

04