四、平稳序列的拟合与预测

目录

模型识别

- 样本自相关系数和偏自相关系数
  - 样本自相关系数：$\hat{\rho}_k = \dfrac{\sum\limits_{t=1}^{n-k}(x_t-\bar{x})(x_{t+k}-\bar{x})}{\sum\limits_{t=1}^{n}(x_t-\bar{x})^2}$
  - 样本偏自相关系数：$\hat{\phi}_{kk} = \dfrac{\hat{D}_k}{\hat{D}}$，其中 $\hat{D} = \begin{vmatrix} 1 & \hat{\rho}_1 & \cdots & \hat{\rho}_{k-1} \\ \hat{\rho}_1 & 1 & \cdots & \hat{\rho}_{k-2} \\ \vdots & \vdots & & \vdots \\ \hat{\rho}_{k-1} & \hat{\rho}_{k-2} & \cdots & 1 \end{vmatrix}$，

$$\hat{D}_k = \begin{vmatrix} 1 & \hat{\rho}_1 & \cdots & \hat{\rho}_1 \\ \hat{\rho}_1 & 1 & \cdots & \hat{\rho}_2 \\ \vdots & \vdots & & \vdots \\ \hat{\rho}_{k-1} & \hat{\rho}_{k-2} & \cdots & \hat{\rho}_k \end{vmatrix}$$

模型识别

- 样本自相关系数和偏自相关系数
  - $\hat{\rho}_k \to N\left(0, \dfrac{1}{n}\right)$
  - $\hat{\phi}_{kk} \to N\left(0, \dfrac{1}{n}\right)$
  - 拖尾性：按指数函数轨迹衰减的，一个连续渐变的过程
  - 截尾性：$i(i > q)$ 阶数的偏自相关系数都在 2 倍标准差范围内

参数估计

- 矩估计
  - 均值：$\hat{\mu} = \bar{x} = \dfrac{\sum x_i}{n}$
  - 方差：$\hat{\sigma}_{\varepsilon}^2 = \dfrac{\sum \varepsilon_t^2}{n}$
  - AR(2)：$\hat{\phi}_1 = \dfrac{1-\hat{\rho}_2}{1-\hat{\rho}_1^2}\hat{\rho}_1, \hat{\phi}_2 = \dfrac{\hat{\rho}_2-\hat{\rho}_1^2}{1-\hat{\rho}_1^2}$
  - MA(1)：$\hat{\theta}_1 = \dfrac{-1+\sqrt{1-4\hat{\rho}_1^2}}{2\hat{\rho}_1}$
- 极大似然估计
- 最小二乘估计

模型检验

- 模型的显著性检验
  - $H_0:\rho_1=\rho_2=\cdots=\rho_m=0$ (模型不显著)，$H_1:\exists\rho_k\neq 0$ (模型显著)
  - 检验统计量：$\mathrm{LB}=n(n+2)\sum_{k=1}^{m}\frac{\hat{\rho}_k^2}{n-k}\sim\chi^2(m)$
- 参数的显著性检验
  - $Q(\tilde{\beta})=\sum\varepsilon_t^2$
  - $H_0:\beta_j=0$，$H_1:\beta_j\neq 0$
  - 统计量：$T=\sqrt{n-m}\frac{\tilde{\beta}}{\sqrt{a_{jj}Q(\tilde{\beta})}}\sim t(n-m)$

模型优化

- AIC 准则和 SBC 准则
  - AIC 准则的指导思想：似然函数值越大越好，未知参数的个数越少越好
  - 中心化 ARMA(p,q) 的 AIC 函数：$\mathrm{AIC}=n\ln\sigma_\varepsilon^2+2(p+q+1)$
  - 非中心化 ARMA(p,q) 的 AIC 函数：$\mathrm{AIC}=n\ln\sigma_\varepsilon^2+2(p+q+2)$
  - 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
  - 中心化 ARMA(p,q) 的 SBC 函数：$\mathrm{SBC}=n\ln\sigma_\varepsilon^2+(\ln n)(p+q+1)$
  - 非中心化 ARMA(p,q) 的 SBC 函数：$\mathrm{SBC}=n\ln\sigma_\varepsilon^2+(\ln n)(p+q+2)$

序列预测

- 线性预测函数

$$\hat{x}_{t+j} = -I_1\hat{x}_{t+j-1} - I_2\hat{x}_{t+j-2} - \cdots - I_j x_t - I_{j+1}x_{t-1} - \cdots$$

- 预测方差最小原则与线性最小方差预测的性质
  - $e_t(l) = x_{t+l} - \hat{x}_t(l)$
  - $E[e_t(l)] = 0$，$\mathrm{Var}[e_t(l)] = \sum_{i=0}^{l-1} G_i^2 \sigma_\varepsilon^2$
  - 置信区间：$(\hat{x}_t(l) \pm z_{1-\alpha/2}(1 + G_1^2 + \cdots + G_{l-1}^2)^{1/2}\sigma_\varepsilon)$

序列预测

- AR(p) 与 MA(q) 序列预测
  - AR(p)： $\hat{x}_t(l) = \phi_1 \hat{x}_t(l-1) + \cdots + \phi_p \hat{x}(l-p)$
  - AR(p)： $\mathrm{Var}[e_t(l)] = \sum_{i=0}^{l-1} G_i^2 \sigma_\varepsilon^2$
  - MA(q)： $\hat{x}_t(l) = \begin{cases} \mu - \sum_{i=l}^{q} \theta_i \varepsilon_{t+l-i}, & l \leq q \\ \mu, & l > q \end{cases}$
  - $\mathrm{Var}[e_t(l)] = \sum_{i=0}^{l-1} G_i^2 \sigma_\varepsilon^2$

- ARMA(p,q) 序列预测
  - $\hat{x}_t(k) = \begin{cases} \hat{x}_t(k), & k \geq 1 \\ x_{t+k}, & k \leq 0 \end{cases}$
  - MA(q)： $\mathrm{Var}[e_t(l)] = \begin{cases} (1+\theta_1^2+\cdots+\theta_{l-1}^2)\sigma_\varepsilon^2, & l \leq q \\ (1+\theta_1^2+\cdots+\theta_q^2)\sigma_\varepsilon^2, & l > q \end{cases}$

序列预测

- 预测修正
  - $\hat{x}_{t+p}(l-p) = G_{l-p}\varepsilon_{t+p} + \cdots + G_{l-1}\varepsilon_{t+1} + \hat{x}(l)$
  - $e_{t+p}(l-p) = G_0\varepsilon_{t+l} + \cdots + G_{l-p-1}\varepsilon_{t+p+1}$
  - $\mathrm{Var}[e_{t+p}(l-p)] = \sum_{i=0}^{l-p-1} G_i^2\sigma_\varepsilon^2 = \mathrm{Var}[e_t(l-p)]$

看程序读结论

条件最小二乘估计					
参数	估计	标准误差	t 值	近似 Pr > |t|	滞后
MU	-0.0013871	0.34414	-0.00	0.9968	0
MA1,1	-0.91784	0.08919	-10.29	<.0001	1
MA1,2	-0.83200	0.11931	-6.97	<.0001	2
MA1,3	-0.59806	0.11906	-5.02	<.0001	3
MA1,4	-0.62317	0.08945	-6.97	<.0001	4

第一列是参数名称；第二列是各参数统计值；

第三列是各参数统计值的标准差；第四列是参数显著性检验的 t 统计量；

第五列是 t 统计量的 P 值；第六列是各参数对应的延迟阶数。