TSA

This note will be updated from time to

模型识别

Box-Jenkins 建模三步流程

1. 对于给定的ts，选取适当的ARIMAp,d,q
2. 对于确定的ARIMA，估计其参数
3. 模型拟合检验

ARMA定阶

ACF

- 定义样本的ACF

$$\hat{\rho}_k = r_k = \frac{\sum_{t=k+1}^{n}(Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^{n}(Y_t - \bar{Y})^2}$$

- 若数据近似服从MA，则应当存在截尾特征；若数据近似服从AR，则应当指数衰减
- 故通过观察数据ACF的截尾特征可以对MA模型进行定阶
- 若假设样本抽样的数据总体来自一个白噪声，则$Var(r_k) = 1/n, Cor(r_k, r_j) = 0$
- 若假设样本抽样的数据总体来自一个MA(q)，则$r_k \sim_d N(0, (1 + 2\rho_1^2 + \cdots + 2\rho_q^2)/n)$.
- **Bartlett's Approximation**
  - $\pm 1.96\sqrt{(1 + 2r_1^2 + \cdots + 2r_q^2)/n}$
  - 由于样本的抽样性质，若假设总体来自MA(q)，则该区间为$H_0: \rho_k = 0$在5%水平下的接受区间
  - 即对于一个直到q阶的ACF，若样本ACF落在这个区间内，则可以认为样本ACF反映出总体ACF在95%的统计水平下是为0的
- 若ACF衰减的很慢，也有可能指示样本数据是非平稳的

PACF

- PACF的定义
  - def1:

$$\phi_{kk} = Corr(Z_t, Z_{t-k} | Z_{t-1}, \cdots, Z_{t-k+1})$$

- PACF的求解
  - 通过解如下Yule-Walker等式：

$$\begin{pmatrix}\phi_{k1}\\ \phi_{k2}\\ \vdots\\ \phi_{kk}\end{pmatrix}=\begin{pmatrix}1&\rho_1&\cdots&\rho_{k-1}\\ \rho_1&1&\cdots&\rho_{k-2}\\ \vdots&\vdots&\ddots&\vdots\\ \rho_{k-1}&\rho_{k-2}&\cdots&1\end{pmatrix}^{-1}\begin{pmatrix}\rho_1\\ \rho_2\\ \vdots\\ \rho_k\end{pmatrix}$$

其中用r_k估计ρ_k即得到了估计值$\hat{\phi}_{kk}$

- PACF的作用
  - 对于一个来自AR(p)的过程，总体的PACF在p阶截尾
  - MA(q)过程的PACF则指数衰减
  - 与Bartlett's Approximation类似，通过$\pm 1.96\sqrt{1/n}$可以得到95%的PACF=0的接受区间

EACF

- 通过EACF可以对模型同时进行p, q的定阶

非平稳性检验

定量评估手段

- 时序图
- ACF的衰减趋势

ADF单位根检验

- 检验模型：

$$Z_t=\alpha Z_{t-1}+X_t$$

- 假设检验：

$$H_0: a=\alpha-1=0$$

说明：$\alpha=1$意味着原假设是**原序列一阶差分平稳**，$|\alpha|<1$说明原序列平稳

- 模型推导：
  - 假设$\{X_t\}$是平稳的AR(k)，故由AR(k)定义：

$$X_t=\phi_1 X_{t-1}+\cdots+\phi_k X_{t-k}+a_t$$

  - 若H_0成立，则此时：

$$X_t=Z_t-Z_{t-1}=\nabla Z_t$$

  - 将二者联立，有：

$$\nabla Z_t=aZ_{t-1}+\phi_1\nabla Z_{t-1}+\cdots+\phi_k\nabla Z_{t-k}+a_t$$

Question: 1. 这里不是a=0？为什么还要单独列出来这一项呢？ 2. 具体该如何根据特征选择mu的序列？

信息准则

AIC

BIC

参数估计

MME

CLS

MLE 与 ULS

估计量的性质

模型诊断

残差分析

残差的计算

趋势性检验

- 残差散点时序图
- 检验残差中是否还含有未被提取充分的信息

正态性检验

- 直方图
- Q-Q图
- 正态分布假设检验
  - Shapiro-Wilk 检验
    - H_0 数据是正态的
  - Jarque-Barre 检验
    - H_0 数据是正态的

残差的相关性检验

- ACF检验
- Ljung-Box检验
  - $H_0: \rho_1 = \cdots \rho_K = 0\,(K$给定)
  - 相当于检验从r_1到r_K的联合效果，联合在一起检验是否有显著相关的残差滞后项
  - 一般选择$K = 6, 12, 18 \cdots$的一系列间隔点，分别进行检验，以保证充分的残差独立

过度拟合检验

- 目的：在确定了一个ARMA(p,q)之后，我们可以通过构造 AMRA(p+1,q) 或 AMRA(p,q+1) 来确定原模型已经充分，新增模型是过度拟合（冗余）的

预测