时间序列分析

课件二

相关概念

强平稳时间序列 联合分布 $Zi_1,Zi_2...Zi_r$ 只依赖于 $Zi_1-i, Zi_2-i...Zi_r-i$ 而不依赖于 i

弱平稳序列

- $E(z_i)$ **不依赖于 i**
- $Cov(zi, z{i-j})$ **只依赖于 j**

自协方差和自相关 第 j 阶自协方差就是之后 j 期

白噪声 最特殊的弱平稳序列

- $E(z_i)=0$ 对任意 i
- $Cov(zi, z{i-j})=0$ 对任意 $j \neq\ 0$

遍历性和LLN 满足遍历性时有 **LLN**，大样本的平均数逼近于均值 μ

Martingale 给定信息集 $Z_{i-1}...,Z_2,Z_1$，有：$E(xi | Z{i-1}...,Z_2,Z1) = x{i-1}$ 信息集举例如 $Z_1 = g_1, Z_2 = g_1 + g_2 ...$

Random Walk $Z_1 = g_1, Z_2 = g_1 + g_2 ...$ 其中 g_i 是 **白噪声**

Martingale Difference Sequences 对于 g_i，满足 $E(g_i)=0$ ，同时条件期望也是0，即 $E(gi|g{i-1},...,g_1) = 0$，没有序列相关性，因为 $E(gi|g{i-j})=0$。

ARCH过程 全称autoregressive conditional heteroskedastic (ARCH) process，**ARCH(1)** 为一阶ARCH过程，性质为：

- $gi= \sqrt{\zeta+\alpha g{i-1}^2 }\cdot\ \varepsilon_i$
- 其中 ε_i 为 **i.i.d** ，均值为0，方差为1

自协方差、自相关的估计 样本自协方差，自相关

cross-covaraiance, cross-correlation

时间序列数据-传统的回归模型

模型选择依据，AIC，BIC，公式为 $log(SSR_j/n)+(j+1)C(n)/n$ **对于AIC，$C(n)=2$，对于BIC，$C(n)=log(n)$**

对于**time regression**，就是将时间 t 作为自变量，加入到回归方程中。

课件三

线性时间序列模型

线性的时间序列 r_t 可以表示成，$rt=\mu+\sum{i=0}^\infty\psiia{t-i}$，**其中 $\psi_0=0$。而 a_i 是白噪声过程** 意思是每一期的 r_t 基于白噪声的线性加和。

简单AR模型

AR(1) $r_t=\phi_0+\phi1r{t-1}+a_t$ 表示 rt 和 $r{t-1}$ 的线性关系，误差项用白噪声 a_t 表示

基本性质：

- $E(r_t)=\phi_0+\phi1E(r{t-1})$ 即 $\mu=\phi_0+\phi_1\mu$，即 $\mu = \dfrac{\phi_0}{1-\phi_1}$
- $Var(r_t)=\phi1^2Var(r{t-1})+\sigma_a^2$，即 $Var(r_t)=\dfrac{\sigma_a^2}{1-\phi_1^2}$
- 很容易推导出自相关函数**ACF**，$p_l = \phi1p{l-1}$，其中 p_l 表示 l 阶自相关，$p_0=1$，$p_1=\phi_1$

类似的对于AR(2)

- $\mu=\dfrac{\phi_0}{1-\phi_1-\phi_2}$
- $p_l=\phi1p{l-1}+\phi2p{l-2}$ 其中 $p_1=\dfrac{\phi_1}{1-\phi_2}$

prove of $p_1=\dfrac{\phi_1}{1-\phi_2}$ 对于 $r_t=\phi_0+\phi1r{t-1}+\phi2r{t-2}+at$，同时对 $r{t-1}$ 取相关系数有：$p_1=\phi_1p_0+\phi_2p_1$，即 $p_1=\dfrac{\phi_1}{1-\phi_2}$ **AR(_p)** $r_t=\phi_0+\phi1r{t-1}+...+\phipr{r-p}+a_t$ 表示 r_t 与p阶滞后项的线性组合的关系

识别AR模型

Partial Autocorrelation Function(PACF) 考虑一系列自回归： $rt = \phi{0,1} + \phi{1,1}r{t-1}+e_{1t}$ $rt = \phi{0,2} + \phi{1,2}r{t-1}+\phi{2,2}r{t-2}+e_{2t}$ $rt = \phi{0,3} + \phi{1,3}r{t-1}+\phi{2,3}r{t-2}+\phi{3,3}r{t-3}+e_{3t}$ $...$ 得到一系列**系数估计**,具有如下性质：

- $\phi_{p,p}$ 收敛到 ϕ_p，当T到正无穷时
- $\phi_{l,l}$ 收敛到0，当 $l>p$ 时

这意味着对于**AR(p)，PACF在滞后p处是截断的**

Information Criteria 对于 **AR(p)** 而言，AIC和BIC分别是

- $AIC(l) = ln(\sigma_l^2)+\dfrac{2l}{T}$
- $BIC(l) = ln(\sigma_l^2)+\dfrac{l *ln(T)}{T}$

模型选择规则 是根据计算一系列不同阶数 l 的AIC，选择最小的AIC值的阶 k，对于BIC也是一样的。

模型诊断 AR模型的关键假设是 **误差项** 是一个白噪声过程，那么拟合出来的模型 **残差** 也应该表现出白噪声过程的性质。利用 **ACF** 和 **Ljung-Box** 统计量来检测残差和白噪声的相近程度

- $Q(m) = T(T+2)\sum\limits_{l=1}^m\dfrac{p_l^2}{T-l}$

其中 p_l 表示 **残差** 的 l 阶自相关系数

预测

课件四

简单MA模型

Infinite Order AR Model 将AR模型扩展到无穷阶，常理而言，无穷阶的AR模型是不可估计的，但是如果 **对系数加以限制** 就可行。如：

- $r_t = \phi_0 -\theta1r{t-1}-\theta1^2r{t-2}-\theta1^3r{t-3}-...+a_t$
- 这其中 $\phi_i=-\theta_1^i$

对于上述形式的无穷阶AR模型，可以通过变换后得到：

- **MA(1)** $r_t=c_0+a_t-\theta1a{t-1}$
- **MA(2)** $r_t=c_0+a_t-\theta1a{t-1}-\theta2a{t-2}$
- **MA(q)** $r_t=c_0+a_t-\theta1a{t-1}-...-\thetaqa{t-q}$
- 这其中 c_0 为常数，a_t 为白噪声

MA(1)的性质

- $E(r_t)=c_0$
- $Var(r_t)=\sigma_a^2+\theta_1^2\sigma_a^2=(1+\theta_1^2)\sigma_a^2$
- 很容易得出 **只有一阶自相关** 且 $\gamma_1 = -\theta_1\sigma_a^2$，$\gamma_l = 0$ 对于所有的 $l>1$，也就是说 **MA(1)的ACF在滞后阶数1处是截断的**

相似的推理可以推断出 **MA(2)** 的性质

从以上性质我们可以进一步推断 **MA模型是弱平稳的**，因为前两阶矩不随时间脚标变化而变化。

同时通过简单的变换我们还能得到 **MA(1)是可逆的**

- $a_t = r_t + \theta1r{t-1} + \theta2r{t-2}...$

确定MA模型的阶 使用 **ACF**，$p_q \neq 0$ 但 $p_l=0$ 对于任何 $l>q$，那么 r_t 符合一个 **MA(q)模型**

简单ARMA模型

将AR和MA结合起来，从 **ARMA(1,1)到ARMA(p,q)**

ARMA(1,1) 形如：

- $r_t - \phi1r{t-1} = \phi_0+a_t-\theta1a{t-1}$
- 其中 a_t 为白噪声，且 $\theta_1 \neq \phi_1$

基本性质（假设 r_t 是平稳的）：

- 两边同时取期望有，$\mu=\dfrac{\phi_0}{1-\phi_1}$
- $E(r_ta_t)=\sigmaa^2$ 基于 $E(r{t-1}at)=0$ 和 $E(a{t-1}a_t)=0$
- $Var(r_t)=\sigma_a^2\dfrac{1-2\phi_1\theta_1+\theta_1^2}{1-\phi_1^2}$

要求 **ACF** 的话，两边乘以 r_{t-l}

- $rtr{t-l}-\phi1r{t-1}r_{t-l}=atr{t-l}-\theta1a{t-1}r_{t-l}$
- 当 l 取 $1,2...l$ 时，有
  - $\gamma_1-\phi_1\gamma_0=\theta_1\sigma_a^2$
  - $\gamma_2-\phi_1\gamma_1=0$
  - $...$
- 即 $\rho_1=\phi_1-\dfrac{\theta_1\sigma_a^2}{\gamma_0}$，$\rho_l=\phi1\rho{l-1}$ 对于 $l>1$
- **不会截断**

ARMA(p,q)

- $r_t=\phi0+\sum\limits{i=1}^p\phiir{t-i}+at-\sum\limits{i=1}^q\thetaia{t-i}$
- **要求不能有公共根，类似于 $\phi_1 \neq \theta_1$**

确定ARMA的阶 ACF和PACF都不再有效，使用 **EACF**，**Extended autocorrelation function** 会生成一个矩阵，行为 p，列为 q，对于 **ARMA(1,1)** 而言，表现为 **矩阵上三角为0，左顶点为(1,1)**，对于 **ARMA(p,1)** 而言，**左顶点为(p,q)**，行列标从0开始

需要将 **EACF简化为O以及X符号**

- **X** 表示 **大于等于 $\dfrac{2}{\sqrt{T}}$** 即大于两倍 **EACF** 的渐进方差
- **O** 表示 **小于 $\dfrac{2}{\sqrt{T}}$**

利用Information Criteria确定阶 同样是给定范围 P 和 Q，选择 **AIC最小对应的p,q值**

同样ARMA也是 **invertible** 的，可以表示为AR的形式