

时间序列分析

课件二

相关概念

强平稳时间序列

联合分布 Z_1, Z_2, \dots, Z_r 只依赖于 $Z_{1-i}, Z_{2-i}, \dots, Z_{r-i}$ 而不依赖于 i

弱平稳序列

- $E(z_i)$ 不依赖于 i
- $\text{Cov}(z_i, z_{i-j})$ 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
- $\text{Cov}(z_i, z_{i-j}) = 0$ 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_{-1}, \dots, Z_2, Z_1 ，有： $E(x_i | Z_{-1}, \dots, Z_2, Z_1) = x_{i-1}$

信息集举例如 $Z_1 = g_1, Z_2 = g_1 + g_2 \dots$

Random Walk

$Z_1 = g_1, Z_2 = g_1 + g_2 \dots$

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 $E(g_i | g_{-1}, \dots, g_1) = 0$ ，没有序列相关性，因为 $E(g_i | g_{i-j}) = 0$ 。

ARCH过程

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

- $g_i = \sqrt{\zeta + \alpha g_{i-1}^2} \cdot \epsilon_i$
- 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-covariance, 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, $r_t = \mu + \sum_{i=0}^{\infty} \psi_{ia} r_{t-i}$, 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)

$$r_t = \phi_0 + \phi_1 r_{t-1} + a_t$$

表示 r_t 和 r_{t-1} 的线性关系, 误差项用白噪声 a_t 表示

基本性质:

- $E(r_t) = \phi_0 + \phi_1 E(r_{t-1})$ 即 $\mu = \phi_0 + \phi_1 \mu$, 即 $\mu = \frac{\phi_0}{1 - \phi_1}$
- $Var(r_t) = \phi_1^2 Var(r_{t-1}) + \sigma_a^2$, 即 $Var(r_t) = \frac{\sigma_a^2}{1 - \phi_1^2}$
- 很容易推导出自相关函数ACF, $\rho_l = \phi_1 \rho_{l-1}$, 其中 ρ_l 表示 l 阶自相关, $\rho_0 = 1$, $\rho_1 = \phi_1$

类似的对于AR(2)

- $\mu = \frac{\phi_0}{1 - \phi_1 - \phi_2}$
- $\rho_l = \phi_1 \rho_{l-1} + \phi_2 \rho_{l-2}$ 其中 $\rho_1 = \frac{\phi_1}{1 - \phi_2}$

prove of $\rho_1 = \frac{\phi_1}{1 - \phi_2}$

对于 $r_t = \phi_0 + \phi_1 r_{t-1} + \phi_2 r_{t-2} + a_t$, 同时对 r_{t-1} 取相关系数有:

$$\rho_1 = \phi_1 \rho_0 + \phi_2 \rho_1, \text{ 即 } \rho_1 = \frac{\phi_1}{1 - \phi_2}$$

AR(p)

$$r_t = \phi_0 + \phi_1 r_{t-1} + \dots + \phi_p r_{t-p} + a_t$$

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:

$$r_t = \phi_{0,1} + \phi_{1,1} r_{t-1} + e_{1t}$$

$$r_t = \phi_{0,2} + \phi_{1,2} r_{t-1} + \phi_{2,2} r_{t-2} + e_{2t}$$

$$r_t = \phi_{0,3} + \phi_{1,3} r_{t-1} + \phi_{2,3} r_{t-2} + \phi_{3,3} r_{t-3} + e_{3t}$$

\dots

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

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

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

- $$Q(m) = T(T+2) \sum_{l=1}^m \frac{\rho_l^2}{T-l}$$

其中 ρ_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：

- $$r_t = \phi_0 - \theta_1 r_{t-1} - \theta_2 r_{t-2} - \theta_3 r_{t-3} - \dots + a_t$$
- 这其中 $\phi_i = -\theta_i$

对于上述形式的无穷阶AR模型，可以通过变换后得到：

- **MA(1)**
$$r_t = c_0 + a_t - \theta_1 a_{t-1}$$
- **MA(2)**
$$r_t = c_0 + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$$
- **MA(q)**
$$r_t = c_0 + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$
- 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**

- $$a_t = r_t + \theta_1 r_{t-1} + \theta_2 r_{t-2} \dots$$

确定MA模型的阶

使用 **ACF**， $\rho_q \neq 0$ 但 $\rho_l = 0$ 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：

- $$r_t - \phi_1 r_{t-1} = \phi_0 + a_t - \theta_1 a_{t-1}$$
- 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, $\mu = \frac{\phi_0}{1-\phi_1}$
- $E(r_t) = \sigma_a^2$ 基于 $E(r_{t-1}) = 0$ 和 $E(a_{t-1}) = 0$
- $\text{Var}(r_t) = \sigma_a^2 \frac{1-2\phi_1\theta_1 + \theta_1^2}{1-\phi_1^2}$

要求 ACF 的话, 两边乘以 r_{t-l}

- $r_{t-l} - \phi_1 r_{t-1} = a_{t-l} - \theta_1 a_{t-1}$
- 当 l 取 $1, 2, \dots$ 时, 有
 - $\gamma_1 - \phi_1 \gamma_0 = \theta_1 \sigma_a^2$
 - $\gamma_2 - \phi_1 \gamma_1 = 0$
 - \dots
- 即 $\rho_1 = \phi_1 - \frac{\theta_1 \sigma_a^2}{\gamma_0}$, $\rho_l = \phi_1 \rho_{l-1}$ 对于 $l > 1$
- 不会截断

ARMA(p,q)

- $r_t = \phi_0 + \sum_{i=1}^p \phi_i r_{t-i} + a_t + \sum_{i=1}^q \theta_i a_{t-i}$
- 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 $\frac{2}{\sqrt{T}}$ 即大于两倍 EACF 的渐进方差
- O 表示 小于 $\frac{2}{\sqrt{T}}$

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