金融学 2023 年秋 • 时间序列

第 7 讲：自回归模型的估计

授课人：刘　岩

武汉大学经管学院金融系

2023 年 10 月 23 日

本讲内容

本节内容

AR(1) 自回归系数的 OLS 估计

- AR(1) 过程可以使用 Yule-Walker 估计自回归系数，进而估计冲击项方差
- 而更一般的方法，是直接进行线性回归 (linear regression)：

$$X_t = \beta X_{t-1} + e_t$$

回归系数 β 可以使用一般最小二乘 (ordinary least square, OLS) 进行估计

- OLS 法：最小化残差平方和

$$\min_{\beta} \sum_{t=2}^{T} (X_t - \beta X_{t-1})^2$$

该最小化问题的解称为 OLS 估计量，记为 $\hat{\beta}$

AR(1) 自回归系数的 OLS 估计

- OLS 估计量的直接推导：最小化问题 1-阶条件

$$\sum_{t=2}^{T} X_{t-1}(X_t - \beta X_{t-1}) = 0 \Rightarrow \hat{\beta}_T = \frac{\sum_{t=2}^{T} X_{t-1}X_t}{\sum_{t=2}^{T} X_{t-1}^2}$$

- 上式中分子分母同时除以 T 可得

$$\hat{\beta}_T = \frac{\frac{1}{T}\sum_{t=2}^{T} X_{t-1}X_t}{\frac{1}{T}\sum_{t=2}^{T} X_{t-1}^2} \xrightarrow{\text{a.s.}} \frac{\sigma_X^2(1)}{\sigma_X^2(0)} = \phi$$

  故 X_t 自回归方程系数的 OLS 估计值具有一致性

- 这里的 $\hat{\beta}$ 就是之前介绍的（1-阶）偏自相关系数

OLS 估计的矩阵形式：单一解释变量

- 单一解释变量线性回归的一般写法

$$Y_t = \beta X_t + e_t$$

其向量形式为

$$\boldsymbol{Y} = \boldsymbol{X}\beta + \boldsymbol{e}$$

其中 $\boldsymbol{Y} = [Y_1, \ldots, Y_T]^\intercal$ 为被解释变量，$\boldsymbol{X} = [X_1, \ldots, X_T]^\intercal$ 解释变量，$\boldsymbol{e} = [e_1, \ldots, e_T]^\intercal$ 残差项

- 最小化残差平方和可表示为

$$\min_{\beta} \boldsymbol{e}^\intercal \boldsymbol{e} = \min_{\beta} (\boldsymbol{Y} - \boldsymbol{X}\beta)^\intercal (\boldsymbol{Y} - \boldsymbol{X}\beta)$$

- 单一解释变量不易体现矩阵形式的优越性；多变量则不同

AR(1) 的估计：冲击项方差

- 利用大数律，可以利用样本方差 $\hat{\sigma}_\varepsilon^2 = \frac{1}{T}\sum_{t=1}^{T}\varepsilon_t^2$ 来估计 σ_ε^2；但唯一的观测值为 $\{X_t\}$，而非 $\{\varepsilon_t\}$
- 利用 Yule-Walker 或者 OLS 的一致估计量 $\hat{\phi}_T$，则大样本下

$$e_t = X_t - \hat{\phi}_T X_{t-1} \xrightarrow{\text{a.s.}} X_t - \phi X_{t-1} = \varepsilon_t$$

因此，可利用冲击项 ε_t 的样本估计值 e_t 来估计 σ_ε^2:

$$\hat{\sigma}_e^2 = \frac{1}{T-1}\sum_{t=2}^{T} e_t^2 = \frac{1}{T-1}\sum_{t=2}^{T}(X_t - \hat{\phi}_T X_{t-1})^2 \xrightarrow{\text{a.s.}} \sigma_\varepsilon^2$$

本节内容

AR(p) 自回归系数的 OLS 估计

- AR(p) 的自回归系数可以用 Yule-Walker 估计，也可用 OLS 方法进行估计
- 将 X_t 写为线性回归形式，并考虑包含常数项的一般情况

$$X_t = \beta_0 + \beta_1 X_{t-1} + \cdots + \beta_p X_{t-p} + e_t$$

- 则 OLS 估计为最小化残差平方和问题的解

$$\min_{\{\beta_i\}} \sum_{t=p+1}^{T} (X_t - \beta_0 - \beta_1 X_{t-1} - \cdots - \beta_p X_{t-p})^2$$

- 若仍然用微积分求解极值问题的 1-阶最优条件来求解，会比较繁琐

OLS 估计的矩阵形式：多解释变量

- 多解释变量线性回归的一般写法

$$Y_t = \boldsymbol{X}_t^{\top}\boldsymbol{\beta} + e_t, \qquad t = 1, \ldots, T$$

其中 $\boldsymbol{X}_t^{\top} = [1, X_{1t}, \ldots, X_{Kt}]$ 为 $K+1$ 个回归变量

- 其矩阵形式为

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{e}$$

其中

$$\boldsymbol{Y} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_T \end{bmatrix}_{T\times 1}, \quad \boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_1^{\top} \\ \vdots \\ \boldsymbol{X}_T^{\top} \end{bmatrix}_{T\times (K+1)}, \quad \boldsymbol{e} = \begin{bmatrix} e_1 \\ \vdots \\ e_T \end{bmatrix}_{T\times 1}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_K \end{bmatrix}_{(K+1)\times 1}$$

OLS 估计的矩阵形式：多解释变量

- OLS 估计量为最小化残差平方和问题的解

$$\min_{\beta} e^\top e = \min_{\beta} (Y - X\beta)^\top (Y - X\beta)$$

- **结论：**若 $\hat{\beta}$ 满足 $X^\top(Y - X\hat{\beta}) = \mathbf{0}$（零向量），则 $\hat{\beta}$ 为上述最小化问题的解；若 $X^\top X$ 可逆，则 $\hat{\beta} = (X^\top X)^{-1} X^\top Y$
  - 只需要说明，$\forall \beta \in \mathbb{R}^{K+1}$，下列不等式成立即可

$$(Y - X\hat{\beta})^\top (Y - X\hat{\beta}) \le (Y - X\beta)^\top (Y - X\beta)$$

  - 为此，只需考虑下式的展开

$$(Y - X\beta)^\top (Y - X\beta) = (Y - X\hat{\beta} + X\hat{\beta} - X\beta)^\top (Y - X\hat{\beta} + X\hat{\beta} - X\beta)$$

及 $\hat{\beta}$ 所满足的性质即可

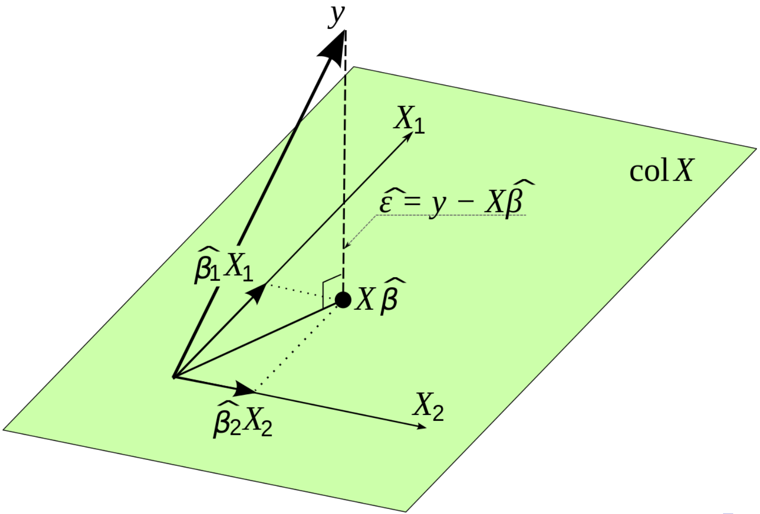
OLS 估计的矩阵形式：多解释变量

$$
\begin{aligned}
& (Y - X\beta)^\top (Y - X\beta) \\
= & (Y - X\hat{\beta} + X\hat{\beta} - X\beta)^\top (Y - X\hat{\beta} + X\hat{\beta} - X\beta) \\
= & (Y - X\hat{\beta})^\top (Y - X\hat{\beta}) + (Y - X\hat{\beta})^\top (X\hat{\beta} - X\beta) \\
& + (X\hat{\beta} - X\beta)^\top (Y - X\hat{\beta}) + (X\hat{\beta} - X\beta)^\top (X\hat{\beta} - X\beta) \\
= & (Y - X\hat{\beta})^\top (Y - X\hat{\beta}) + 2\underbrace{(Y - X\hat{\beta})^\top X}_{\equiv 0}(\hat{\beta} - \beta) \\
& + \underbrace{(\hat{\beta} - \beta)^\top X^\top X(\hat{\beta} - \beta)}_{\geq 0} \\
\geq & (Y - X\hat{\beta})^\top (Y - X\hat{\beta}).
\end{aligned}
$$

OLS 估计的几何视角

- 将 $\boldsymbol{X}$ 看做 $\mathbb{R}^T$ 中的 $K+1$ 个列向量，则 $\boldsymbol{X\beta}$ 为这 $K+1$ 个向量张成的子空间 $\mathcal{X}$ 中的一个向量
- $(\boldsymbol{Y}-\boldsymbol{X\beta})^\intercal(\boldsymbol{Y}-\boldsymbol{X\beta})$ 为 $\boldsymbol{Y}$ 到 $\boldsymbol{X\beta}$ 的距离
- 最小二乘估计就等价于寻找 $\boldsymbol{\beta}$ 使得 $\boldsymbol{Y}$ 到 $\mathcal{X}$ 距离最短
- 这又等价于使得残差向量 $\boldsymbol{e}=\boldsymbol{Y}-\boldsymbol{X\beta}$ 与子空间 $\mathcal{X}$ 垂直，即 $\boldsymbol{X}^\intercal\boldsymbol{e}=\boldsymbol{0}$
- OLS 估计可看做被解释变量对解释变量的投影 (projection)，$(\boldsymbol{X}^\intercal\boldsymbol{X})^{-1}\boldsymbol{X}^\intercal$ 称为投影矩阵

OLS 估计的几何视角


y
X_1
$\text{col}\, X$
$\widehat{\varepsilon} = y - X\widehat{\beta}$
$\widehat{\beta}_1 X_1$
$X\,\widehat{\beta}$
X_2
$\widehat{\beta}_2 X_2$

本节内容

AR(p) 自回归系数 OLS 估计的一致性

- 将含有常数项的一般 AR(p) 模型写为 OLS 回归的矩阵形式

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$$

其中

$$\boldsymbol{Y} = \begin{bmatrix} X_{p+1} \\ \vdots \\ X_T \end{bmatrix}, \quad \boldsymbol{X} = \begin{bmatrix} 1 & X_p & \cdots & X_1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{T-1} & \cdots & X_{T-p} \end{bmatrix}, \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_{p+1} \\ \vdots \\ \varepsilon_T \end{bmatrix}$$

- 注意希腊字母 epsilon ε 和 E 的花体 $\mathcal{E}$ 的区别

AR(p) 自回归系数 OLS 估计的一致性

- 若 AR(p) 模型不包括常数项，则自回归系数的 OLS 估计为

$$
\begin{aligned}
\hat{\beta}_T &= (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{Y} \\
&= \left(\frac{1}{T}\boldsymbol{X}^\top \boldsymbol{X}\right)^{-1} \frac{1}{T}\boldsymbol{X}^\top \boldsymbol{Y} \\
&\xrightarrow{\text{a.s.}} \begin{bmatrix} \sigma_X^2(0) & \cdots & \sigma_X^2(p-1) \\ \vdots & \ddots & \vdots \\ \sigma_X^2(p-1) & \cdots & \sigma_X^2(0) \end{bmatrix}^{-1} \begin{bmatrix} \sigma_X^2(1) \\ \vdots \\ \sigma_X^2(p) \end{bmatrix} = \begin{bmatrix} \phi_1 \\ \vdots \\ \phi_p \end{bmatrix}
\end{aligned}
$$

  由 Yule-Walker 方程可知，该 OLS 估计满足一致性

- 一致性的关键：$\mathbb{E}[\varepsilon_t X_{t-i}] = 0,\ \forall i \geq 1$

线性回归 OLS 估计的一致性：一般情形

- 估计量 $\hat{\boldsymbol{\beta}}_T$ 统计推断的基础是一致性：$\hat{\boldsymbol{\beta}}_T \xrightarrow{\text{a.s.}} \boldsymbol{\beta}_0$
- 假设模型系数真实值为 $\boldsymbol{\beta} = \boldsymbol{\beta}_0$，则 $\boldsymbol{\beta}$ 的 OLS 估计可以写为

$$
\begin{aligned}
\hat{\boldsymbol{\beta}}_T &= (\boldsymbol{X}^\intercal \boldsymbol{X})^{-1} \boldsymbol{X}^\intercal \boldsymbol{Y} \\
&= (\boldsymbol{X}^\intercal \boldsymbol{X})^{-1} \boldsymbol{X}^\intercal (\boldsymbol{X}\boldsymbol{\beta}_0 + \boldsymbol{\varepsilon}) \\
&= \boldsymbol{\beta}_0 + (\boldsymbol{X}^\intercal \boldsymbol{X})^{-1} \boldsymbol{X}^\intercal \boldsymbol{\varepsilon} \\
&= \boldsymbol{\beta}_0 + \left(\frac{1}{T} \boldsymbol{X}^\intercal \boldsymbol{X} \right)^{-1} \frac{1}{T} \boldsymbol{X}^\intercal \boldsymbol{\varepsilon} \\
&= \boldsymbol{\beta}_0 + \left(\frac{1}{T} \sum_{t=1}^{T} \boldsymbol{X}_t \boldsymbol{X}_t^\intercal \right)^{-1} \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{X}_t \varepsilon_t
\end{aligned}
$$

- 上述表达式同样适用于包含常数项的一般形式 AR(p) 模型

线性回归 OLS 估计的一致性：一般情形

- OLS 回归的标准假设：
  1. $\{X_t\}$ 平稳，故交叉 2-阶矩矩阵 $M_{K\times K} = \mathbb{E}X_t X_t^\intercal$ 存在且可逆，即 M^{-1} 存在
  2. $\{\varepsilon_t\}$ 为 iid 序列，且给定 X_t 时的条件均值为 0，即 $\mathbb{E}[\varepsilon_t|X_t] = 0$；由此知 $\mathbb{E}X_t\varepsilon_t = \mathbf{0}_{K\times 1}$
- 在此组假设之下，应用大数定律

$$
\begin{aligned}
\left(\frac{1}{T}\sum_{t=1}^{T} X_t X_t^\intercal\right)^{-1} \frac{1}{T}\sum_{t=1}^{T} X_t\varepsilon_t &\xrightarrow{\text{a.s.}} M_{K\times K}^{-1}\mathbf{0}_{K\times 1} \\
&= \mathbf{0}_{K\times 1}
\end{aligned}
$$

故 $\hat{\beta}_T \xrightarrow{\text{a.s.}} \beta_0$，即 OLS 估计量具有一致性

OLS 估计对应的总体表达式

- 给定回归模型 $Y_t = \boldsymbol{X}_t^\top \boldsymbol{\beta} + \varepsilon_t$
- 对给定样本 $\{Y_t, X_t\}_{t=1}^T$，$\boldsymbol{\beta}$ 的 OLS 是最小化残差平方和 $\min_{\beta} \sum_t \varepsilon_t^2 = \min_{\beta} \sum_t (Y_t - \boldsymbol{X}_t^\top \boldsymbol{\beta})^2$ 的解
- 给定线性回归基础假设 $\mathbb{E}[\varepsilon_t | \boldsymbol{X}_t] = 0$，可以在回归模型两端乘以 $\boldsymbol{X}_t$ 并取期望：

$$\mathbb{E}[\boldsymbol{X}_t Y_t] = \mathbb{E}[\boldsymbol{X}_t \boldsymbol{X}_t^\top]\boldsymbol{\beta} + \mathbb{E}[\boldsymbol{X}_t \varepsilon_t] = \mathbb{E}[\boldsymbol{X}_t \boldsymbol{X}_t^\top]\boldsymbol{\beta}$$

$$\Longrightarrow \boldsymbol{\beta} = (\mathbb{E}[\boldsymbol{X}_t \boldsymbol{X}_t^\top])^{-1} \mathbb{E}[\boldsymbol{X}_t Y_t]$$

- 上式对应于 OLS 样本估计表达式

$$\hat{\boldsymbol{\beta}}_T = (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{Y} = \left(\frac{1}{T} \sum_{t=1}^T \boldsymbol{X}_t \boldsymbol{X}_t^\top \right)^{-1} \frac{1}{T} \sum_{t=1}^T \boldsymbol{X}_t Y_t$$