

国际金融试验班 2019 年秋 · 时间序列

# 第 10 讲：时间序列的预测

授课人：刘 岩

武汉大学经管学院金融系

2019 年 12 月 12 日

# 本讲内容

1 条件期望与最优预测

2 线性预测

3 时间序列的预测

## 本节内容

1 条件期望与最优预测

2 线性预测

3 时间序列的预测

## 条件概率和条件期望

- 给定随机变量  $X, Y$  以及其联合密度函数  $f(x, y)$
- 给定  $X = x$ ,  $Y$  的条件概率密度可表示为

$$f(y|x) = \frac{\Pr(X = x, Y = y)}{\underbrace{\Pr(X = x)}_{\text{不严格类比}}} = \frac{f(x, y)}{\int f(x, y) dy}$$

- $\mathbb{E}(Y|X = x) = \int y f(y|x) dy$  称为  $Y$  在  $X$  上的条件期望
- 记为  $\mathbb{E}(Y|X)$ ; 可看做  $X$  的函数

## 条件期望的性质

- 全期望律 (law of total expectation):  $\mathbb{E}Y = \mathbb{E}[\mathbb{E}(Y|X)]$
- 若  $X, Y$  相互独立, 则  $\mathbb{E}(Y|X) = \mathbb{E}Y$
- 给定函数  $g(\cdot)$ ,  $\mathbb{E}[g(X)Y|X] = g(X)\mathbb{E}(Y|X)$
- $\mathbb{E}(Y|X)$  是  $X$  对  $Y$  的最小均方预测函数:  $\mathbb{E}(Y|X)$  是最小化问题

$$\min_{g(\cdot)} \mathbb{E}[Y - g(X)]^2$$

的 (唯一) 解  $\Rightarrow$  均方误差意义下,  $Y$  的最优预测为  $\mathbb{E}(Y|X)$

- 若  $Z = g(X)$  且  $g(\cdot)$  是严格单调函数, 则

$$\mathbb{E}(Y|Z) = \mathbb{E}[Y|g(X)] = \mathbb{E}(Y|X)$$

## 最优预测推导

- 对于任意的函数  $f(\cdot)$ ,  $f(X)$  对  $Y$  的均方预测误差可写为

$$\begin{aligned}
 & \mathbb{E}[Y - f(X)]^2 \\
 &= \mathbb{E}[Y - \mathbb{E}(Y|X) + \mathbb{E}(Y|X) - f(X)]^2 \\
 &= \mathbb{E}[Y - \mathbb{E}(Y|X)]^2 + 2\mathbb{E}[(Y - \mathbb{E}(Y|X))(\mathbb{E}(Y|X) - f(X))] \\
 &\quad + \mathbb{E}[\mathbb{E}(Y|X) - f(X)]^2 \\
 &= \mathbb{E}[Y - \mathbb{E}(Y|X)]^2 + 2\mathbb{E}[\underbrace{\mathbb{E}[(Y - \mathbb{E}(Y|X))(\mathbb{E}(Y|X) - f(X))|X]}_{=(\mathbb{E}(Y|X)-f(X))\mathbb{E}[(Y-\mathbb{E}(Y|X))|X]=0}] \\
 &\quad + \mathbb{E}[\mathbb{E}(Y|X) - f(X)]^2 \\
 &= \mathbb{E}[Y - \mathbb{E}(Y|X)]^2 + \mathbb{E}[\mathbb{E}(Y|X) - f(X)]^2
 \end{aligned}$$

- 故对任意  $f(\cdot)$ ,  $\mathbb{E}[Y - f(X)]^2 \geq \mathbb{E}[Y - \mathbb{E}(Y|X)]^2$

## 联合正态下的条件期望

- 给定  $X, Y$  服从二元正态分布
- 定义  $Z = (Y - \mathbb{E}Y) - \frac{\text{cov}(X, Y)}{\text{var}(X)}(X - \mathbb{E}X)$ ; 则有  $Z$  为正态分布, 且  $\mathbb{E}Z = 0$
- 可验证  $\text{cov}(Z, X) = \mathbb{E}ZX = 0$ , 故  $Z, X$  互相独立; 进一步的,  $Z, g(X)$  相互独立, 故  $\mathbb{E}Zg(X) = \mathbb{E}Z\mathbb{E}g(X) = 0$
- 由此可证明  $Y - Z = \mathbb{E}Y + \frac{\text{cov}(X, Y)}{\text{var}(X)}(X - \mathbb{E}X)$  是  $X$  对  $Y$  的最小均方预测, 故

$$\mathbb{E}(Y|X) = \mathbb{E}Y + \frac{\text{cov}(X, Y)}{\text{var}(X)}(X - \mathbb{E}X), \quad \text{进一步有}$$

$$\text{var}(Y|X) = \text{var}(Y) - \frac{[\text{cov}(X, Y)]^2}{\text{var}(X)}, \quad \text{是一个常数}$$

## 本节内容

1 条件期望与最优预测

2 线性预测

3 时间序列的预测

## 线性预测

- 给定随机变量  $Y$  及  $K$ -维随机向量  $X$  (可包含常数项), 定义  $X$  对  $Y$  的 线性预测 (linear prediction) 为  $\hat{Y} = b^\top X$ ,  $b \in \mathbb{R}^K$ , 对应的  $Y - \hat{Y}$  称为预测误差 (prediction error)
- 定义 最优均方 (optimal mean square) 线性预测为使  $\mathbb{E}[Y - \hat{Y}]^2$  最小的线性预测  $\hat{b}^\top X$ :

$$\hat{b} = \underset{b}{\operatorname{argmin}} \mathbb{E}[Y - b^\top X]^2$$

- 类似于 OLS, 上述  $\hat{b}$  的取值需满足  $\mathbb{E}[(Y - \hat{b}^\top X)X^\top] = \mathbf{0}^\top$ , 故  $\hat{b} = (\mathbb{E}[XX^\top])^{-1}\mathbb{E}[XY]$
- 最优线性预测记为  $L(Y|X)$ , 可直观理解为  $Y$  对  $X$  的投影

## 正态条件下的线性预测与最优预测

- 给定  $Y, X$  为 2-元正态分布 r.v., 期望为 0
- $X$  对  $Y$  的最优线性预测系数  $\hat{b}$  满足

$$\hat{b} = \frac{\mathbb{E}[XY]}{\mathbb{E}X^2} = \frac{\text{cov}(X, Y)}{\text{var}(X)}$$

最优线性预测  $L(Y|X) = \frac{\text{cov}(X, Y)}{\text{var}(X)} X$

- 此时  $X$  对  $Y$  的最优（均方）预测与最优（均方）线性预测相等

$$\mathbb{E}[Y|X] = L(Y|X)$$

### 本节内容

1 条件期望与最优预测

2 线性预测

3 时间序列的预测

# 时间序列的预测

- 给定时间序列  $\{Y_t\}$ , 可以考虑用  $t$  及之前的观测值, 对  $Y_{t+s}$  进行线性预测

$$\hat{Y}_{t+s|t} = L(Y_{t+s} | Y_t, Y_{t-1}, \dots), \quad s \geq 1$$

- 更一般的, 给定时间序列  $\{Y_t\}$  与  $\{X_t\}$ , 其中  $X_t \in \mathbb{R}^K$ , 可以考虑用  $t$  及之前的观测值  $\{X_{t-j}\}_{j \geq 0}$  对  $Y_{t+s}$  进行线性预测

$$\hat{Y}_{t+s|t} = L(Y_{t+s} | X_t, X_{t-1}, \dots), \quad s \geq 1$$

- 最优线性预测的具体形式, 即系数向量  $\hat{b}$  的确定, 依赖于具体模型

## AR(1) 的预测

- 给定平稳 AR(1) 过程  $X_{t+1} = \mu + \phi X_t + \varepsilon_{t+1}$ ,  $|\phi| < 1$ ,  $\{\varepsilon_t\}$  为白噪声
- $X_{t+1}$  在  $t$  的最优线性预测  $\hat{X}_{t+1|t} = \mu + \phi X_t$ 
  - 直接可验证  $\mathbb{E}[(X_{t+1} - \hat{X}_{t+1|t})X_t] = 0$
- $X_{t+2}$  在  $t$  的最优线性预测  $\hat{X}_{t+2|t} = \mu + \phi\mu + \phi^2 X_t$ 
  - 由 AR 迭代可得:  $X_{t+2} = \mu + \phi\mu + \phi^2 X_t + \phi\varepsilon_{t+1} + \varepsilon_{t+2}$
  - 由此可验证  $\mathbb{E}[(X_{t+2} - \hat{X}_{t+2|t})X_t] = 0$
- 另一角度,  $\hat{X}_{t+2|t}$  可看做  $\hat{X}_{t+2|t+1}$  在  $t$  的线性预测
  - $\hat{X}_{t+2|t+1} = \mu + \phi X_{t+1}$
  - $\hat{X}_{t+2|t} = \mu + \phi X_{t+1|t} = \mu + \phi\mu + \phi^2 X_t$

## AR(1) 预测的渐近性质

- 可验证,  $\forall s \geq 1$  有

$$\hat{X}_{t+s|t} = \mu \sum_{r=0}^{s-1} \phi^r + \phi^s X_t$$

- 由此可知,

$$\lim_{s \rightarrow \infty} \hat{X}_{t+s|t} = \frac{\mu}{1 - \phi} = \mathbb{E}X_t$$

即  $X_{t+s}$  的长期 (最优) 线性预测值收敛到其无条件期望

- 进一步计算可知

$$\lim_{s \rightarrow \infty} \mathbb{E}[X_{t+s} - \hat{X}_{t+s|t}]^2 = \frac{\sigma_\varepsilon^2}{1 - \phi^2} = \text{var}(X_t)$$

即  $X_{t+s}$  的长期均方预测误差收敛到其无条件方差

## 三点注释

- 对一般的 AR( $p$ ) 过程，同样可以利用“递归”预测的方式，计算  $\hat{X}_{t+s|t}$ 
  - 先计算  $\hat{X}_{t+s|t+s-1}$ ，再对结果中出现的  $X_{t+s-1}$  计算  $\hat{X}_{t+s-1|t+s-2}$ ，以此递推
- 对 AR( $p$ ) 过程而言，若  $\{\varepsilon_t\}$  相互独立，则有

$$\hat{X}_{t+s|t} = L(X_{t+s}|X_t, X_{t-1}, \dots) = \mathbb{E}[X_{t+s}|X_t, X_{t-1}, \dots] = \mathbb{E}_t X_{t+s}$$

- MA 过程及 ARMA 过程同样可以进行类似的预测
  - MA 的预测可借助对观测值  $X_t$  的自回归近似所得到的近似  $\hat{\varepsilon}_t$  来完成