

国际金融试验班 2018 年秋 · 时间序列

第 3 讲：时间序列数据基本特征

授课人：刘 岩

武汉大学经管学院金融系

2018 年 9 月 28 日

本讲内容

① 基本统计特征

② 趋势与季节性

本节内容

① 基本统计特征

② 趋势与季节性

时间序列数据主要统计特征

与普通数据序列一样，时间序列数据 $\{X_t\}$ 主要统计特征包括：

- 期望： $\mathbb{E}X_t$
- 方差、标准差： $\text{var}(X_t), \sigma_X$
- 以及两个变量间的协方差： $\text{cov}(X_t, Y_t)$ 及相关系数 ρ_{XY}

但与普通数据不同，时间序列数据的主要统计特征还包括

- 自协方差 (autocovariance): $\text{cov}(X_t, X_{t-k}), k \geq 1$
- 对应的，自相关系数 (autocorrelation):

$$\rho_k = \frac{\text{cov}(X_t, X_{t-k})}{\sigma_X^2}$$

- 注意，上述各阶矩和 t 无关 $\Rightarrow$ 平稳性假设

矩的估计

- 给定时间序列数据 $\{X_t\}_{t=1}^T$, 各阶矩的估计 (estimate) 如下
- 期望 (均值): $\hat{\mu}_X = \frac{1}{T} \sum_t X_t$; $\sum_t$ 表示对 $t = 1, \dots, T$ 求和
- 方差:

$$\hat{\sigma}_X^2 = \begin{cases} \frac{1}{T-1} \sum_t (X_t - \hat{\mu}_X)^2, & \text{unbiased} \\ \frac{1}{T} \sum_t (X_t - \hat{\mu}_X)^2, & \text{biased} \end{cases}$$

大样本之下, 无偏 (unbiased) 估计与有偏 (biased) 估计等价

- 协方差:

$$\hat{\sigma}_{XY} = \begin{cases} \frac{1}{T-1} \sum_t (X_t - \hat{\mu}_X)(Y_t - \hat{\mu}_Y), & \text{unbiased} \\ \frac{1}{T} \sum_t (X_t - \hat{\mu}_X)(Y_t - \hat{\mu}_Y), & \text{biased} \end{cases}$$

- 相关系数: $\hat{\rho}_{XY} = \hat{\sigma}_{XY} / (\hat{\sigma}_X \hat{\sigma}_Y)$

矩的估计

- 自协方差：

$$\hat{\sigma}_k^2 = \frac{1}{T-k} \sum_{t=k+1}^T (X_t - \hat{\mu}_X)(X_{t-k} - \hat{\mu}_X)$$

注意求和中 t 的取值范围

- 自相关系数：

$$\hat{\rho}_k = \frac{\hat{\sigma}_k^2}{\hat{\sigma}_X^2}$$

练习 估计值 $\hat{\rho}_{XY}, \hat{\rho}_k$ 是否仍然介于 ± 1 之间？

- 在上下文区分明确时，也可以略去 “ $\hat{\cdot}$ ” 符号
- 但有时为明确样本量，也会增加 T 下标，如 $\hat{\mu}_{X,T}$

矩估计的理论基础

给定时间序列数据 $\{X_t\}_{t=1}^T$, 上述矩估计的理论基础在于 大数定律 (law of large number)

定理 1

当 X_t 满足特定条件时, 样本均值收敛到总体均值

$$\lim_{T \rightarrow \infty} \hat{\mu}_{X,T} = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T X_t = \mathbb{E}X_t$$

大数定律：注释

- **特定条件**：极端情况如 $\{X_t\}$ iid；本课程中，可以理解为“平稳性条件”——严格说法是“严平稳遍历性条件”
- **收敛**：随机变量的收敛概念有 4 种，依分布收敛 $\xrightarrow{d}$ ，依概率收敛 $\xrightarrow{P}$ ，矩收敛 $\xrightarrow{m}$ ，概率-1 收敛（几乎处处收敛） $\xrightarrow{\text{a.s.}}$
 - 假设有 r.v. 序列 $\{Z_t\}_{t=1}^{\infty}$ 和 r.v. Z ，则称

① $Z_t \xrightarrow{d} Z$ ，若 $\lim_{t \rightarrow \infty} F_{Z_t}(z) = F_Z(z)$ 对几乎所有 z 成立

② $Z_t \xrightarrow{P} Z$ ，若 $\forall \varepsilon > 0$ ， $\lim_{t \rightarrow \infty} \mathbb{P}(|Z_t - Z| \geq \varepsilon) = 0$

③ $Z_t \xrightarrow{m} Z$ ，若 $\lim_{t \rightarrow \infty} \mathbb{E}[(Z_t - Z)^2] = 0$

④ $Z_t \xrightarrow{\text{a.s.}} Z$ ，若 $\mathbb{P}(\lim_{t \rightarrow \infty} Z_t = Z) = 1$

- 4 种收敛的关系

$$\left\{ \begin{array}{c} \xrightarrow{\text{a.s.}} \\ \xrightarrow{m} \end{array} \right\} \Rightarrow \xrightarrow{P} \Rightarrow \xrightarrow{d}$$

- 大数定律中“收敛”可以达到 $\xrightarrow{\text{a.s.}}$

平稳性的关键性质

定理 2

若 $\{X_t\}$ 是平稳序列, 对任意的 $k+1$ -元可测函数 f , 定义

$$Y_t = f(X_t, \dots, X_{t-k})$$

则 $\{Y_t\}$ 也是平稳序列。

随机收敛的性质

定理 3

假设 $X_t \xrightarrow{\text{a.s.}} X$, $Y_t \xrightarrow{\text{a.s.}} Y$, 则对任意的可测函数 f , 有

$$f(X_t, Y_t) \xrightarrow{\text{a.s.}} f(X, Y).$$

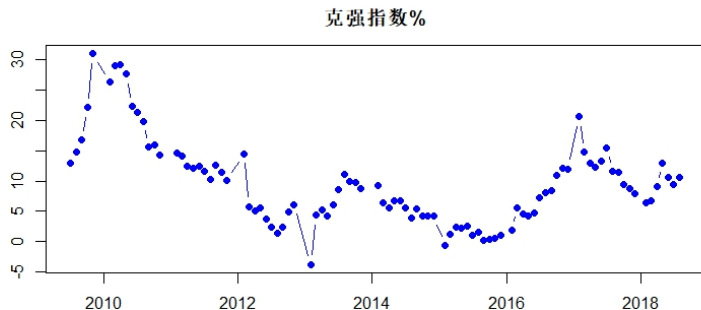
本节内容

① 基本统计特征

② 趋势与季节性

时间序列的趋势性

- 时间序列关心数据在时间上的变动

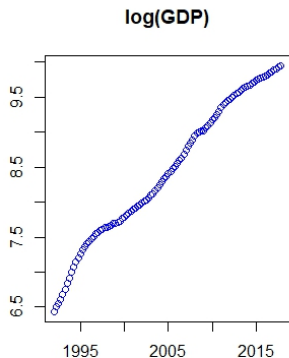
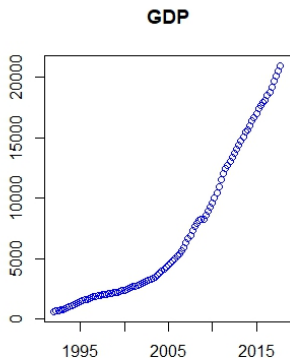


$$\hat{\mu} = 9.53, \hat{\sigma} = 6.98, \hat{\rho} = 0.92$$

- 有趋势 vs. 无趋势

趋势的常见类型

- 指数趋势 vs. 线性趋势



- 两者之间差一个对数变换: $X_t \Rightarrow \log(X_t)$

线性趋势调整：基本方法

- 原始数据 $\{X_t\}_{t=1}^T$ 包含线性趋势 $X_t = \alpha t + \tilde{X}_t$, 而 $\tilde{X}_t$ 无趋势
- 两端取差分

$$\Delta X_t = X_t - X_{t-1} = \alpha + \tilde{X}_t - \tilde{X}_{t-1} = \alpha + \Delta \tilde{X}_t$$

其中 Δ 表示差分算子, 则 ΔX_t 不包含趋势

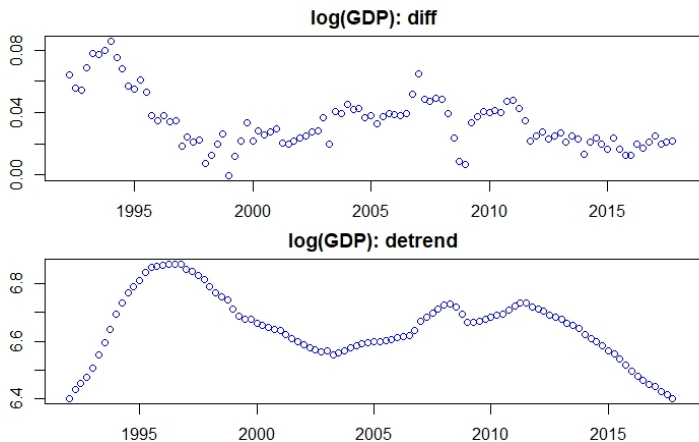
- 或者, 如果要求 $\tilde{X}_1 = \tilde{X}_T$, 则有

$$\alpha = \frac{X_T - X_1}{T - 1}$$

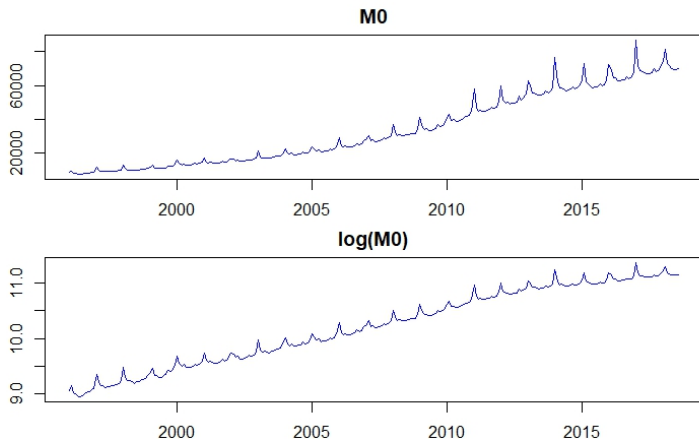
从而可以推算 $\tilde{X}_t$

- 类似的过程称为去趋势 (detrend)

去趋势示例



季节性示例



需要对原始数据进行季节调整 (seasonal adjustment)