



④ 时间序列分析

知识点概要

- ◆ 时间序列分析概述
- ◆ STATA中的时间日期数据
- ◆ 时间序列数据预处理
- ◆ 时间序列修匀/平滑处理
- ◆ 时间序列标绘图
- ◆ 时滞、前导和差分
- ◆ 自相关检验
- ◆ ARIMA模型
- ◆ ARMAX模型

▣ 什么是时间序列？

- ▣ 时间序列是指某种统计指标的数值，按照时间先后顺序排列起来的数列。如：年度GDP数据，股票数据

▣ 什么是时间序列分析？

- ▣ 分析目标变量在**时间方面的发展变化规律**，并为之建立数学模型；利用建立的数学模型**预测未来一段时间内的取值**。
- ▣ 通俗来讲：如果我们去观察时间序列图，也能用语言描述出时间序列的规律，时间序列分析是用统计学/数学语言更精确的表述数据变化规律的统计分析方法。

▣ 时间序列分析的用途

- ▣ 主要用在计量经济学、宏观经济学。当然数据分析方法的应用主要看分析师的想象力（想象力依赖于对模型的理解程度，以及对业务的理解/对要解决的问题的理解）

▣ 时间序列分析预测的局限性

- ▣ 有足够长的时间序列数据，数据越多越好；
- ▣ 时间序列数据不能有缺失，如果有，分析时需要进行缺失值的填充；
- ▣ 数据序列的变动是稳定而规律的；
 - ▣ 想想为什么没怎么听说过厉害的数学家和统计学家，炒股抄的很好？股票价格变化不是稳定的，也没有明显规律！**不要学完时间序列就用来炒股。**

时间序列分析概述—时间序列的成分

□ 长期趋势

- 长期趋势是指由于某种根本性因素的影响，时间序列在较长时间内朝着一定的方向持续上升或下降，以及停留在某一水平上的倾向。它反映了事物的主要变化趋势。

□ 季节变动

- 季节变动是由于自然条件和社会条件的影响，时间序列在一年内随着季节的转变而引起的周期性变动。

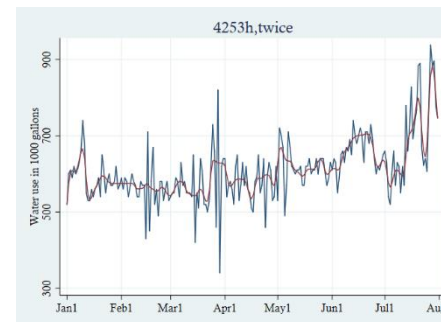
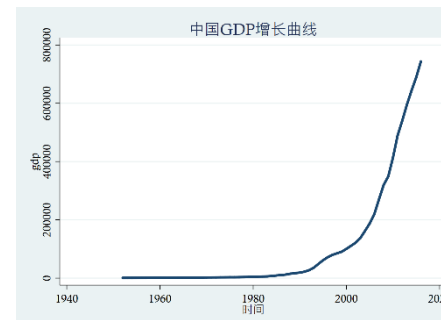
□ 循环变动（少见）

- 循环变动是以数年为周期的周期变动。它与长期趋势不同，不是朝单一方向持续发展，而是涨落相间的波浪式起伏变动。与季节变动不同，它的波动时间较长，变动周期长短不一，短则在一年以上，长则数年、数十年，上次出现以后，下次何时出现，难以预料。如：经济危机。

□ 不规则变动

- 不规则变动是指由各种偶然性因素引起的无周期变动。不规则变动又可分为突然变动和随机变动。
- 突然变动，是指诸如战争、自然灾害、地震、意外事故、方针、政策的改变所引起的变动。
- 随机变动，是指由于大量的随机因素所产生的影响。

□ 注意：并不是每个序列都包含所有4种成分



- ▣ **时间序列图**

- ▣ 通过时间序列图能够直观的观察时间序列的变化趋势。

- ▣ **多种修匀/平滑处理方法**

- ▣ 移动平均数
 - ▣ 移动中位数
 - ▣ 指数平滑法
 - ▣ 季节性/非季节性的Holt-Winters修匀

- ▣ **时间序列差分**

- ▣ **ARMA模型**：自回归移动平均模型（Autoregressive Moving Average Model, 简称：ARMA模型）
- ▣ **ARIMA模型**：差分整合移动平均自回归模型（Autoregressive Integrated Moving Average model, 简称：ARIMA模型）
- ▣ **ARMAX模型**：包含外生变量的移动平均自回归模型（Autoregressive Moving Average with eXogenous VariableS,简称：ARMAX模型）

▣ 时间日期数据类型的特点：

- ▣ (1) 计算机内部存储的时候存储的是数字类型;
- ▣ (2) 显示出来给人看到的是时间日期
- ▣ (3) 以文本格式存储的数据文件，里面的时间日期类型是字符型

▣ 时间日期数据一般需要进行的处理：

- ▣ (1) 将以文本存储的人能识别的时间日期转换成计算机能理解的数字型时间日期;
- ▣ (2) 给存储在计算机里面的数字型时间日期添加显示格式，显示时间日期;
- ▣ (3) 时间日期数据之间的相互转换

▣ 在stata中处理时间日期数据

- ▣ (1) HRFs: Human readable forms
- ▣ (2) SIF: Stata internal form
- ▣ (3) HRF-to-SIF convention
- ▣ (4) Displaying SIFs in HRF
- ▣ (5) Building SIFs from components
- ▣ (6) help datetime for more functions

- 以案例的形式介绍不同时间粒度的时间序列数据的预处理

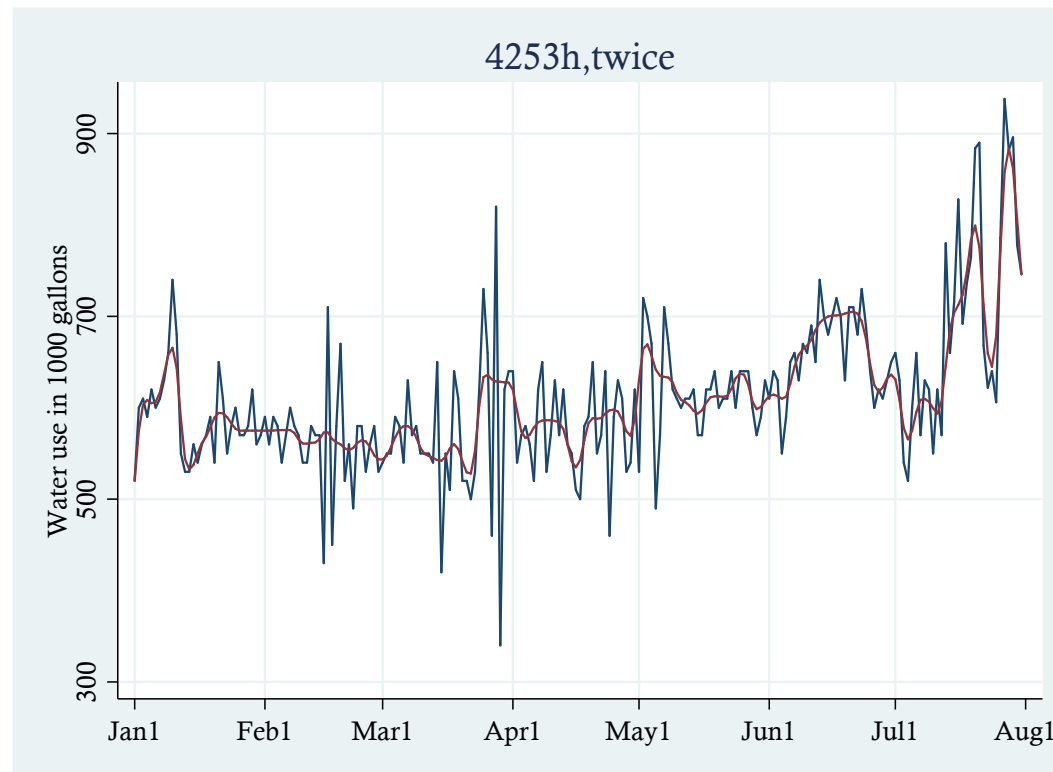
- 年度数据（中国建国以来GDP增长曲线）
- 季度数据（国内生产总值季度变化）
- 月度数据（国内商品房月度销售额）
- 周粒度数据（万科股价周变化曲线）
- 天粒度数据（万科股价日变化曲线）
- 分钟粒度数据（万科股价某日上午每5分钟变化曲线）

▣ 时间序列修匀/平滑

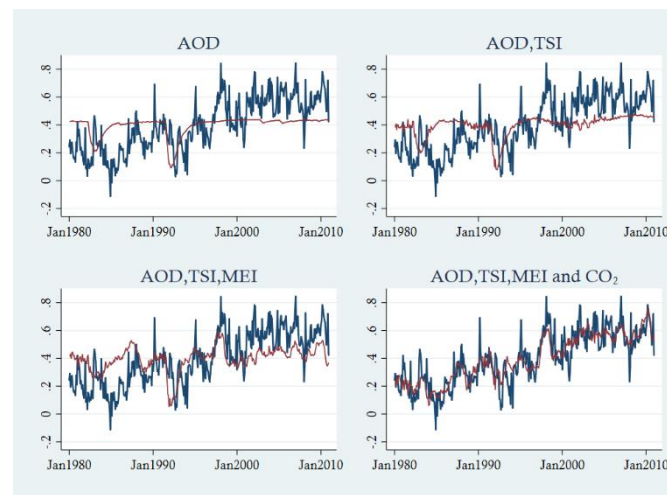
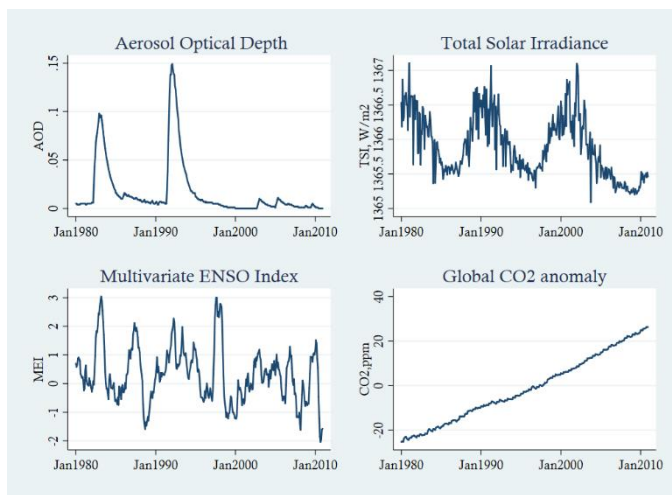
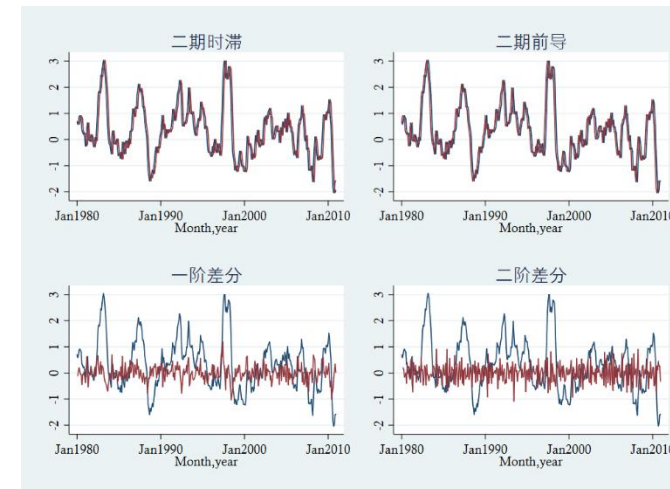
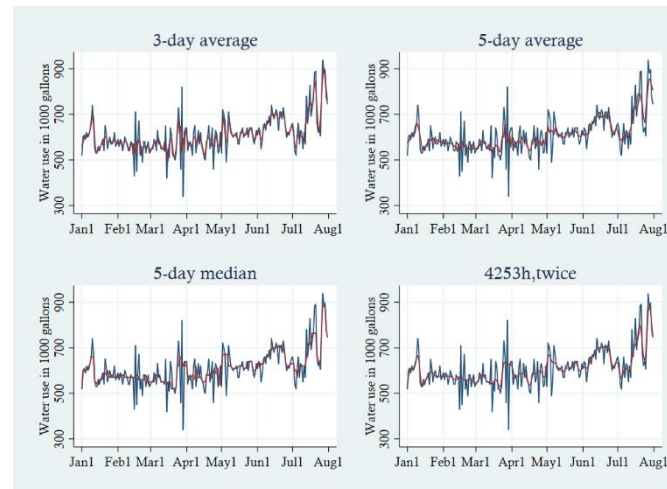
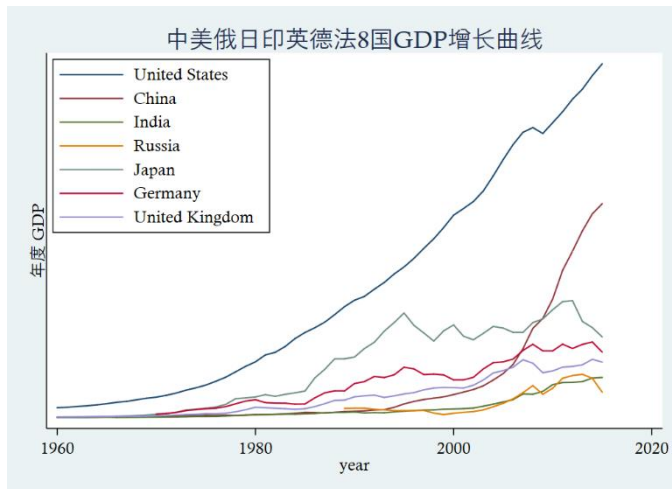
- ▣ 通过建立近似函数尝试抓住数据中的主要模式，去除噪音、结构细节或瞬时现象，来处理时间序列的方法集合。
- ▣ 在平滑过程中，时间序列原始数据点被修改，由噪音产生的数据点被降低或被提升，从而得到一个更平滑的时间序列。
- ▣ 平滑处理可以抚平短期波动，从而反映出长期趋势或周期。

▣ 时间序列修匀工具及对应的stata命令

- ▣ `tssmooth ma`：移动平均数过滤器，可处理加权或者未加权数据
- ▣ `tssmooth nl`：移动中位数修匀，非线性过滤器
- ▣ `tssmooth exponential`：单指数过滤器
- ▣ `tssmooth dexponential`：双指数过滤器
- ▣ `tssmooth hwinters`：非季节性的holt-winters修匀
- ▣ `tssmooth shwinters`：季节性的holt-winters修匀



时间序列标绘图



□ 时滞:

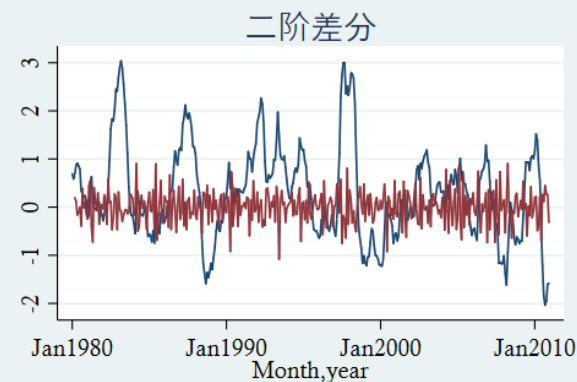
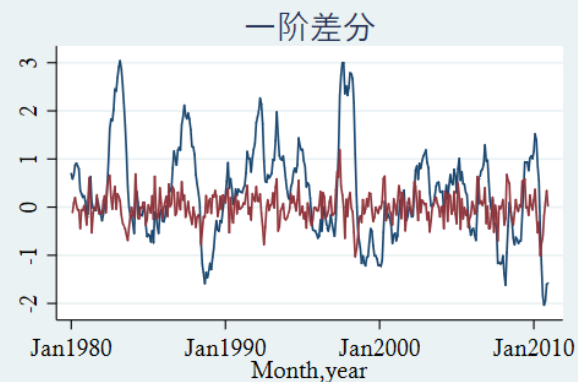
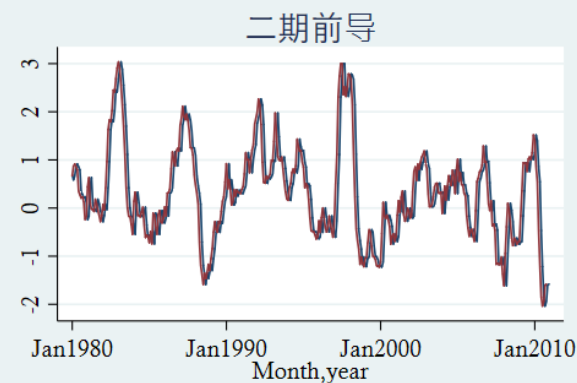
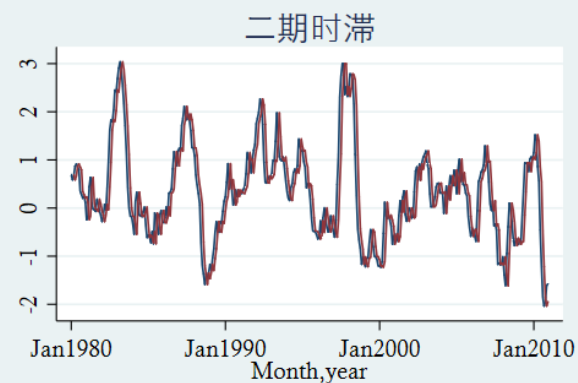
- 时滞变量就是将过去的历史值，往后进行平移，产生的当前值和过去的某个值相同有一种时间滞留的感觉。

□ 前导:

- 前导变量是取当前值稍后的数据，向前进行平移，产生前导变量。

□ 差分:

- 差分：用时间序列原始序列与时间序列的某个特定时滞序列相减，得到的新序列



- ▣ **自回归模型 (Autoregressive model, 简称AR模型) :**
 - ▣ 用同一时间序列的之前各期, 也就是用 $x_1 \sim x_{t-1}$ 来预测本期 x_t 的表现, 并假设它们为线性关系。
 - ▣ 因为这是从回归分析中的线性回归发展而来, 因为用 x 预测 x (自己) , 所以叫做自回归。
- ▣ **移动平均模型 (Moving Average Models, 简称MA):**
 - ▣ 时间序列可以由同期值与过去值给予不同的权重来解释, 这称为移动平均模型。移动平均模型在任何条件下都平稳。
- ▣ **自回归移动平均模型 (Autoregressive moving average model, 简称: ARMA模型) :**
 - ▣ 是研究时间序列的重要方法, 由自回归模型 (简称AR模型) 与移动平均模型 (简称MA模型) 为基础 “混合” 构成。
- ▣ **差分整合移动平均自回归模型 (Autoregressive Integrated Moving Average model,简称ARIMA模型) :**
 - ▣ 又称整合移动平均自回归模型 (移动也可称作滑动) 。
- ▣ **包含外生变量的移动平均自回归模型 (Autoregressive Moving Average with eXogenous VariableS,简称: ARMAX模型) :**
 - ▣ 可以纳入自变量的ARIMA模型。

