时间序列分析
Time Series Analysis

0 绪论

1 时间序列的预处理

2 平稳时间序列分析

3 非平稳时间序列分析

绪论

- 学习目标：
  - 理解时间序列分析中的基本概念；
  - 掌握平稳时间序列的AR、MA、ARMA模型；
  - 了解非平稳时间序列的确定性因素分解方法，掌握ARIMA模型
  - 熟练使用Python语言进行时间序列分析任务实战

绪论

- **为什么要进行时间序列分析？**
- 个人、企业和政府都需要根据历史数据（时间序列）对现象的未来发展作出预测并采取相应的决策，时间序列分析为我们提供了相应的分析工具。
- 我国每年年初都要对当年的主要经济指标作出预测，每个五年计划中要对未来五年的经济和社会发展进行预测。
- 股票经纪人要对股票市场的未来走势作出及时的预测并相应作出买入或卖出的决策。
- 企业经理人员的决策中经常需要对未来的市场供求进行预测。

绪论

- **时间序列分析问题：**
- 常用按时间顺序排列的一组随机变量$X_1, X_2, \ldots, X_t$ 来表示一个随机事件的时间序列，简记为$\{X_t\}$；用$x_1, x_2, \ldots, x_n$或$\{x_t, t = 1,2, \ldots, n\}$表示该随机序列的n个有序观察值，称之为序列长度为n的观察值序列。
- 本节课应用时间序列的目的是给定一个已被观测了的时间序列，**预测**该序列的**未来值**

1 时间序列的预处理

- 1.1 时间序列的分类
- 1.2 平稳性检验
- 1.3 纯随机性检验

1.1 时间序列的分类

- 首先对观测序列的纯随机性和平稳性进行检验，根据检验结果将序列分为不同的类型，采取不同的分析方法
- 纯随机序列又称白噪声序列，序列的各项之间没有任何相关性，是没有信息可以提取的平稳序列
- 平稳非白噪声序列的均值和方差是常数，通常建立一个线性模型来拟合该序列，借此提取有用信息
- 非平稳序列的均值和方差不稳定，一般将其变为平稳序列后进行分析，本节课主要关注其中一种经过差分后具有平稳性的时间序列，即差分平稳序列，可使用ARIMA模型分析

1.2 平稳性检验

•记时间序列$\{X_t, t \in \boldsymbol{T}\}$在任意时刻的序列值$X_t$的均值为$\mu_t$，方差为$\boldsymbol{\sigma}_t$

•任取$t, s \in \boldsymbol{T}$，定义序列$\{X_t\}$的自协方差系数：

$$\gamma(\boldsymbol{t}, \boldsymbol{s}) = \boldsymbol{E}[(\boldsymbol{X_t} - \mu_t)(\boldsymbol{X_s} - \mu_{\boldsymbol{s}})]$$

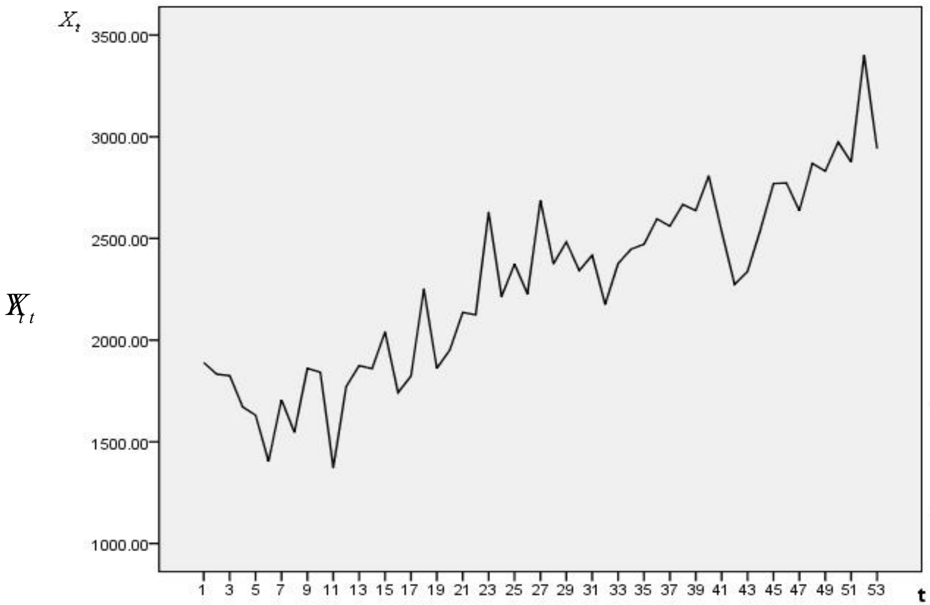
和自相关系数$\rho(\boldsymbol{t}, \boldsymbol{s}) = \frac{\gamma(\boldsymbol{t}, \boldsymbol{s})}{\boldsymbol{\sigma}_t \boldsymbol{\sigma}_s}$

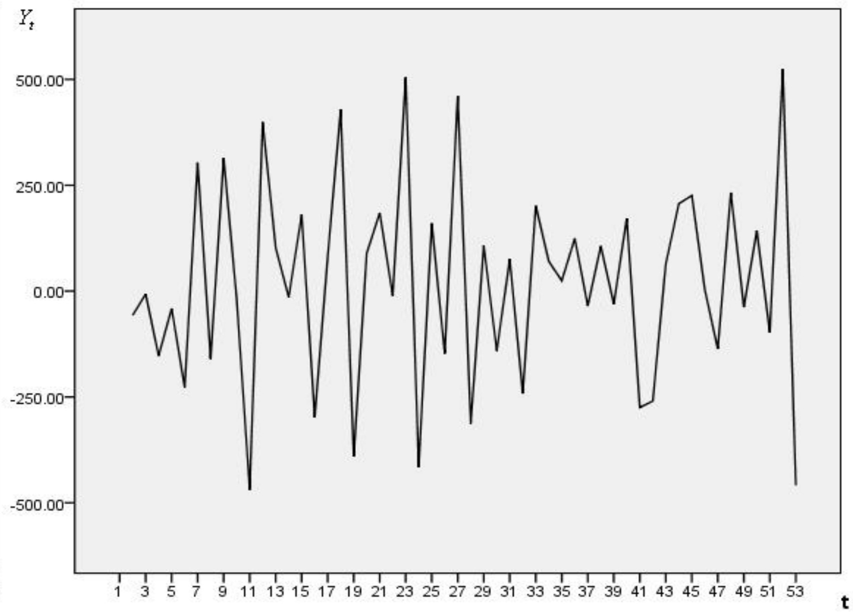
•如果时间序列$\{X_t, t \in \boldsymbol{T}\}$的均值和方差为常数，且延迟k期的序列变量的自协方差和自相关系数是相等的，则称时间序列$\{X_t, t \in \boldsymbol{T}\}$为平稳序列。

1.2 平稳性检验

- 序列平稳性的检验的三种方法：
- 时序图检验：根据平稳时间序列的均值和方差都为常数的性质，平稳序列的时序图显示该序列值始终在一个常数附近随机波动，且波动的范围有界，如果表现出明显的趋势性或者周期性，那它通常不是平稳序列
- 自相关图检验：平稳序列具有短期相关性，间隔越远的过去值对现时值的影响越小，随延迟期数的k的增加，平稳序列的自相关系数会比较快地衰减趋向于0，并在0附近随机波动。
- 单位根检验

1.2 平稳性检验





非平稳序列

平稳序列

1.3 纯随机性检验

- 如果一个序列是纯随机序列，那么理想情况下应有$\rho(t,s)=0, t \neq s$. 实际上纯随机序列的样本自相关系数在0附近波动。
- 纯随机性检验也称白噪声检验，一般构造检验统计量来检验序列的纯随机性，常用的检验统计量有Q统计量和LB统计量。

2 平稳时间序列分析

- **2.1 AR模型**
- **2.2 MA模型**
- **2.3 ARMA模型**
- **2.4 平稳时间序列建模**

2.1 AR模型（AutoregRessive model）

如果$\rho_1 \neq 0$，则X_t与X_{t-1}相关，可以用X_{t-1}预测X_t。最简单的预测为线性组合，如下模型：

$$X_t = \phi_0 + \phi_1 X_{t-1} + \varepsilon_t$$

称为一阶**自回归**模型(Autoregression model)，记作AR(1)模型。其中$\{\varepsilon_t\}$是零均值独立同分布白噪声序列，方差为σ^2，并设ε_t与$X_{t-1}, X_{t-2}, \ldots$独立。系数$|\phi_1| < 1$。更一般的定义中仅要求$\{\varepsilon_t\}$是零均值白噪声，不要求独立同分布。

这个模型与一元线性回归模型$Y_i = \phi_0 + \phi_1 x_i + \varepsilon_i$在某些方面类似，比如，$\varepsilon_t$都是起到误差或者扰动的作用。但是，在自回归模型中，自变量X_{t-1}在时刻$t-1$时作为因变量，所以自回归模型中因变量和自变量不是两个变量而是同一个变量的不同时刻。回归模型的理论结果不能直接应用到自回归模型当中。

AR(1)模型也是马尔可夫(Markov)过程：X_t在$X_{t-1}, X_{t-2}, \ldots$条件下的条件分布，只与X_{t-1}有关。已知X_{t-1}后，用$X_{t-1}, X_{t-2}, \ldots$去预测X_t，与仅用X_{t-1}去预测的效果相同。这种性质称为**马氏性**。

条件期望和条件方差：

$$E(X_t | X_{t-1}) = \phi_0 + \phi_1 X_{t-1}, \quad \mathrm{Var}(X_t | X_{t-1}) = \sigma^2$$

2.1 AR模型（AutoregRessive model）

即在$X_{t-1}=x_{t-1}$已知后，X_t条件的条件分布是期望为$\phi_0+\phi_1 x_{t-1}$，方差为σ^2的分布。可以证明

$$\operatorname{Var}(X_t)=\frac{\sigma^2}{1-\phi_1^2}$$

因为$|\phi_1|<1$，所以X_t在$X_{t-1}=x_{t-1}$已知条件下的条件方差小于其无条件方差，也就是说用X_{t-1}的信息去预测X_t，可以使得X_t的波动减小，能够达到预测的效果。

AR(1)模型的推广是AR(p)模型：

$$X_t=\phi_0+\phi_1 X_{t-1}+\cdots+\phi_p X_{t-p}+\varepsilon_t$$

其中$\{\varepsilon_t\}$是零均值独立同分布白噪声序列，方差为σ^2，且ε_t与$X_{t-1},X_{t-2},\ldots$独立。系数$\phi_1,\ldots,\phi_p$需要满足如下条件：方程

$$1-\phi_1 z-\cdots-\phi_p z^p=0$$

的所有复根z_*都满足$|z_*|>1$，上述方程的左边的多项式称为AR(p)模型的**特征多项式**，特征多项式的所有复根称为**特征根**，对系数的条件称为“特征根都在单位圆外”。

由$\mu=\phi_0/(1-\phi_1)$得$\phi_0=(1-\phi_1)\mu$，AR(1)模型可以写成

$$X_t=(1-\phi_1)\mu+\phi_1 X_{t-1}+\varepsilon_t$$

2.1 AR模型（AutoregRessive model）

- **AR(1)模型的自相关函数**

AR(1)模型的平稳解满足ϵ_t与$X_{t-1}, X_{t-2}, \ldots$独立。将模型写成

$$X_t - \mu = \phi_1(X_{t-1} - \mu) + \varepsilon_t \tag{4.3}$$

在(4.3)两边乘以ε_t后取期望，有

$$E[\varepsilon_t(X_t - \mu)] = \phi_1 E[\varepsilon_t(X_{t-1} - \mu)] + \sigma^2 = \sigma^2$$

在(4.3)两边乘以$X_t - \mu$后取期望，有

$$\gamma_0 = \phi_1 \gamma_1 + \sigma^2$$

在(4.3)两边乘以$X_{t-j} - \mu$后取期望，有

$$\gamma_j = \phi_1 \gamma_{j-1}, \; j = 1, 2, \ldots$$

2.1 AR模型（AutoregRessive model）

即有

$$\begin{aligned}\gamma_0 =& \sigma_X^2 = \frac{\sigma^2}{1-\phi_1^2} = \phi_1\gamma_1 + \sigma^2 \\ \gamma_j =& \phi_1\gamma_{j-1},\ j = 1, 2, \ldots\end{aligned}$$

于是ACF为

$$\rho_j = \frac{\gamma_j}{\gamma_0} = \phi_1^j,\ j = 1, 2, \ldots$$

当$0 < \phi_1 < 1$时，ACF为正的单调下降序列，以负指数速度（几何级数）下降。当$-1 < \phi_1 < 0$时，ACF为负正交替的序列，绝对值以负指数速度下降。

2.1 AR模型（AutoregRessive model）

- **偏自相关函数**

实际数据用AR模型建模时，阶数p是未知的， 确定p的问题称为**定阶**。 一般常用偏自相关函数和AIC准则。

设$X_1,\ldots,X_n,Y$为随机变量，

$$L(Y|X_1,\ldots,X_n)=\underset{\hat{Y}=b_0+b_1X_1+\cdots+b_nX_n}{\operatorname{argmin}} E(Y-\hat{Y})^2$$

称为用$X_1,\ldots,X_n$对Y的最优线性预测。 $Y-L(Y|X_1,\ldots,X_n)$与$Z-L(Z|X_1,\ldots,X_n)$ 的相关系数称为Y和Z在扣除$X_1,\ldots,X_n$影响后的**偏相关系数**。

对平稳线性时间序列， 对$n=1,2,\ldots$， 有

$$L(X_t|X_{t-1},\ldots,X_{t-n})=\phi_{n0}+\phi_{n1}X_{t-1}+\cdots+\phi_{nn}X_{t-n}$$

其中$\phi_{nj},j=0,1,\ldots,n$与t无关。 称ϕ_{nn}为时间序列$\{X_t\}$的**偏自相关系数**， $\{\phi_{nn}\}$序列称为时间序列$\{X_t\}$的**偏自相关函数**(PACF)。

ϕ_{nn}实际是X_t与X_{t-n}在扣除$X_{t-2},\ldots,X_{t-n+1}$的影响后的偏相关系数。 ϕ_{11}就是ρ_1。

ϕ_{nn}用样本进行估计， 得到的估计值$\hat{\phi}_{nn},n=1,2,\ldots$称为**样本偏自相关函数**。

2.1 AR模型（AutoregRessive model）

如果$\{X_t\}$服从如下AR(p)模型：

$$X_t = \phi_0 + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \varepsilon_t,\ \phi_p \neq 0$$

这意味着用$X_{t-1}, X_{t-2}, \ldots$的线性组合预测X_t时，只需要用到$X_{t-1}, \ldots, X_{t-p}$，增加$X_{t-p-1}, X_{t-p-2}, \ldots$不能改进预测。这意味着$\phi_{kk} = 0, k > p$。这种性质叫做AR模型的偏自相关函数截尾性。

AR(p)序列的样本偏自相关函数$\hat{\phi}_{kk}$满足如下性质：

- $T \to \infty$时$\hat{\phi}_{pp} \to \phi_p \neq 0$。
- 对$k > p$, $\hat{\phi}_{kk} \to 0 (T \to \infty)$。
- 对$k > p$，$\hat{\phi}_{kk}$渐近方差为$\frac{1}{T}$。

这样，可以用类似对ACF的白噪声检验那样给PACF图画出$\pm\frac{2}{\sqrt{T}}$的上下界限，以此判断PACF在哪里截尾。

2.2 MA模型（Moving Average model）

移动平均模型是具有q步外不相关性质的平稳列的模型； 对于高阶的AR模型， 有些可以用低阶的MA模型更好地描述。 一般的AR模型也可以用高阶MA模型近似。

理论上，AR模型也可以是无穷阶的：

$$X_t = \phi_0 + \sum_{j=1}^{\infty} \phi_j X_{t-j} + \varepsilon_t$$

其中$\{\phi_j\}$应绝对可和。 一个特例为

$$X_t = \phi_0 - \sum_{j=1}^{\infty} (-\theta_1)^j X_{t-j} + \varepsilon_t$$

其中$0 < |\theta| < 1$。 将模型写成：

$$X_t + \sum_{j=1}^{\infty} (-\theta_1)^j X_{t-j} = \phi_0 + \varepsilon_t \qquad (*)$$

以$t-1$代入，并乘以$-\theta_1$，有

$$\sum_{j=1}^{\infty} (-\theta_1)^j X_{t-j} = -\phi_0\theta_1 - \theta_1\varepsilon_{t-1}$$

2.2 MA模型（Moving Average model）

代入到(*)式中得

$$X_t = \phi_0(1+\theta_1) + \varepsilon_t + \theta_1\varepsilon_{t-1}$$

这样的模型称为MA(1)模型。

一般地，若$\{\varepsilon_t\}$是零均值独立同分布白噪声，方差为σ^2，$|\theta_1| < 1$，令

$$X_t = \theta_0 + \varepsilon_t + \theta_1\varepsilon_{t-1}$$

易见$\{X_t\}$为线性时间序列形式的弱平稳列，称为MA(1)序列。

类似地，MA(2)序列的模型为

$$X_t = \theta_0 + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2}$$

MA(q)序列的模型为

$$X_t = \theta_0 + \varepsilon_t + \theta_1\varepsilon_{t-1} + \cdots + \theta_q\varepsilon_{t-q}$$

2.2 MA模型（Moving Average model）

- **MA模型的平稳性与自相关函数的性质**

以MA(1)为例。$X_t = \theta_0 + \varepsilon_t + \theta_1 \varepsilon_{t-1}$，其中$\{\varepsilon_t\}$是零均值独立同分布白噪声，$\theta_0, \theta_1$是任意实数，平稳性不需要特征根的条件。

易见

$$EX_t = \theta_0, \forall t, \quad \mathrm{Var}(X_t) = \sigma^2(1+\theta_1^2)$$

而

$$\gamma_1 = E[(X_t - \theta_0)(X_{t-1} - \theta_0)] = E[(\varepsilon_t + \theta_1 \varepsilon_{t-1})(\varepsilon_{t-1} + \theta_1 \varepsilon_{t-2})] = \theta_1 E\varepsilon_{t-1}^2 = \sigma^2 \theta_1$$

对$k > 1$有

$$\gamma_k = E[(\varepsilon_t + \theta_1 \varepsilon_{t-1})(\varepsilon_{t-k} + \theta_1 \varepsilon_{t-k-1})] = 0 \ (k > 1)$$

因为$k > 1$，所以$t-k-1 < t-k < t-1 < t$，求协方差时均不相关。

2.2 MA模型（Moving Average model）

所以，对于MA(1)序列，有

$$\gamma_k=\begin{cases}\sigma^2(1+\theta_1^2), & k=0\\ \sigma^2\theta_1, & k=1,\\ 0, & k>1\end{cases}$$

相应地，MA(1)的自相关函数为

$$\rho_k=\begin{cases}1, & k=0\\ \frac{\theta_1}{1+\theta_1^2}, & k=1,\\ 0, & k>1\end{cases}$$

这就验证了MA(1)序列是弱平稳列。 MA(1)的自相关函数在$k>1$后为零的性质叫做MA序列的自相关函数截尾性。

对于MA(q)序列

$$X_t=\theta_0+\varepsilon_t+\theta_1\varepsilon_{t-1}+\cdots+\theta_q\varepsilon_{t-q}$$

易见

$$EX_t=\theta_0,\quad \mathrm{Var}(X_t)=\sigma^2(1+\theta_1^2+\cdots+\theta_q^2)$$

其自相关函数ρ_k也满足q后截尾性， 即$\rho_k=0, \forall k>q$。 如果$\theta_q\neq 0$， 则$\rho_q\neq 0$。 这样， MA(q)序列的两个时间点的观测X_s和X_t当$|s-t|>q$时不相关， 代表了一种特殊的“有限记忆”的模型。 MA序列的自相关函数截尾性也是在模型识别和定阶时的重要依据。

2.3 ARMA模型

AR模型有偏自相关函数截尾性质； MA模型有相关函数截尾性质。 有些因果线性时间序列有与AR和MA类似的表现， 但是不能在低阶实现偏自相关函数截尾或者相关函数截尾。

ARMA模型结合了AR和MA模型， 在对数据拟合优度相近的情况下往往可以得到更简单的模型， 而且不要求偏自相关函数截尾也不要求相关函数截尾。

ARMA(1,1)模型为

$$X_t = \phi_0 + \phi_1 X_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

或

$$X_t - \phi_1 X_{t-1} = \phi_0 + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

其中$|\phi_1| < 1$, $|\theta_1| < 1$, $-\phi_1 \neq \theta_1$。 $\{\varepsilon_t\}$是独立同分布零均值白噪声列, ε_t与$X_{t-1}, X_{t-2}, \dots$独立。

一般的ARMA(p, q)类似。

AR(p)可以看成ARMA(p, 0), MA(q)可以看成是ARMA(0, q)。

2.4 平稳时间序列建模

- 某个时间序列经过预处理被判定为平稳非白噪声序列，就可以利用ARMA模型进行建模。计算平稳非白噪声序列$\{X_t\}$的自相关系数的性质，选择合适的模型。
- **自相关函数(ACF)**描述时间序列观测值与其过去的观测值之间的线性相关性。
- **偏自相关函数(PACF)**描述在给定中间观测值的条件下时间序列观测值与其过去的观测值之间的线性相关性。

模型	自相关系数（ACF）	偏自相关系数（PACF）
AR(p)	拖尾	p阶截尾
MA(q)	q阶截尾	拖尾
ARMA(p,q)	拖尾	拖尾

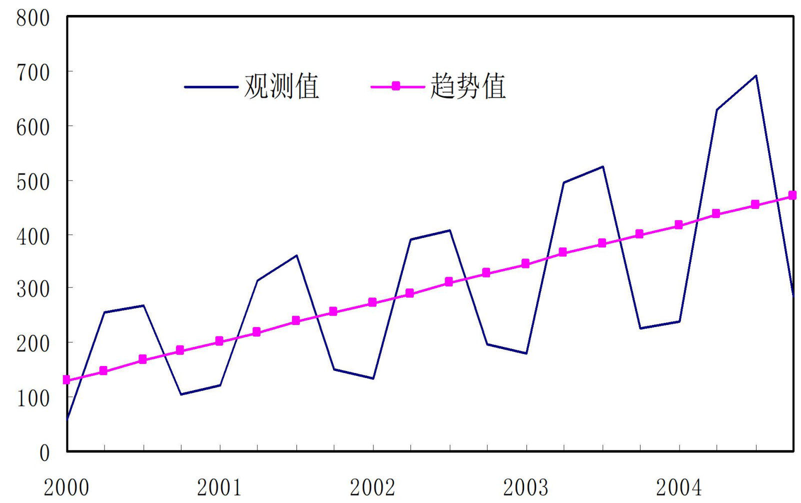
3 非平稳时间序列分析

- 3.1 确定性因素分解
- 3.2 ARIMA模型

3.1 确定性因素分解

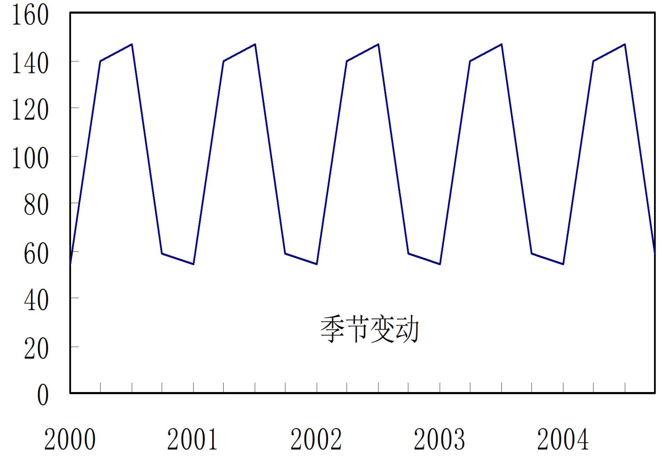
长期趋势：现象在较长时期内持续发展变化的一种趋向或状态

可以分为线性趋势和非线性趋势



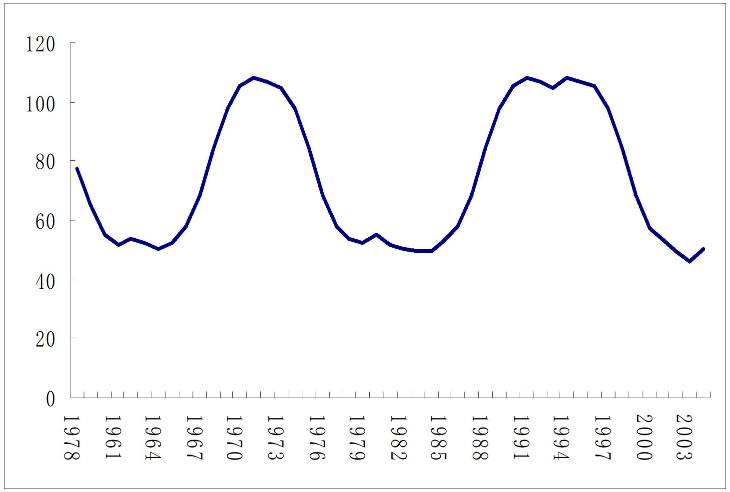
3.1 确定性因素分解

- **季节变动：**由于季节的变化引起的现象发展水平的规则变动。季节变动产生的原因主要有两个：
  - 自然因素；
  - 人为因素： 法律、习俗、制度等
- “季节变动”也用来指周期小于一年的规则变动，例如24小时内的交通流量。



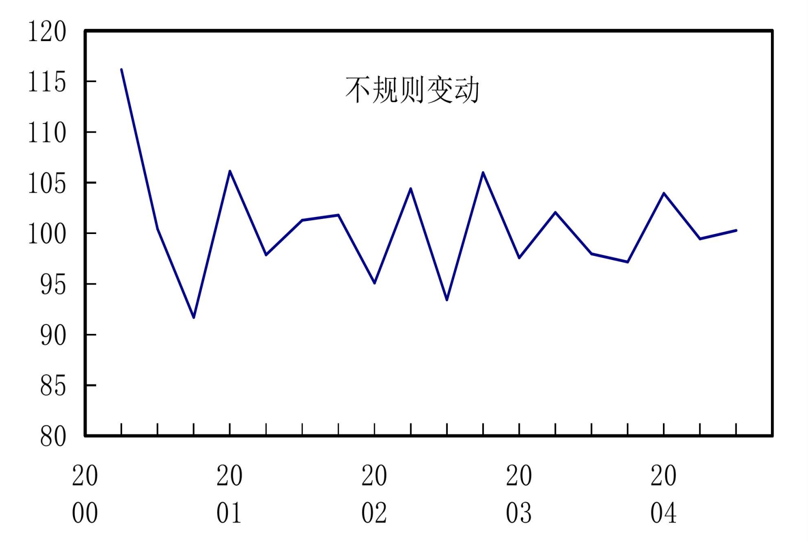
3.1 确定性因素分解

- **循环变动：**以若干年为周期、不具严格规则的周期性连续变动。
- 与长期趋势不同，它不是朝着单一方向的持续运动，而是涨落相间的波浪式起伏变化；
- 与季节变动也不同，它的波动时间较长，变动的周期长短不一，变动的规则性和稳定性较差。



3.1 确定性因素分解

- **随机变动（不规则变动）：** 由于众多偶然因素对时间序列造成的影响。
- 随机变动是不可预测的。



3.1 确定性因素分解

- 时间序列的组成成分之间可能是乘法或加法的关系，因此，时间序列可用多种模型进行分解，常见的有加法模型、乘法模型和加乘混合模型。

- **加法模型**假设时间序列中每一个指标数值都是长期趋势、季节变动、循环变动和不规则变动四种成分的总和，在加法模型中，四种成分之间是相互独立的。某种成分的变动并不影响其他成分的变动。各个成分都用绝对量表示，并且具有相同的量纲。

$$Y_t = T_t + S_t + C_t + I_t$$

3.1 确定性因素分解

- **乘法模型**是假设时间序列中每一个指标数值都是长期趋势、季节变动、循环变动和不规则变动四种成分的乘积。在乘法模型中, 四种成分之间保持着相互依存的关系。一般而言，长期趋势成分用绝对量表示，具有和时间序列本身相同的量纲，其它成分则用相对量表示。

$$Y_t = T_t \times S_t \times C_t \times I_t$$

3.1 确定性因素分解

- **加乘混合模型**，比如

$$Y_t = T_t \times C_t \times S_t + I_t$$

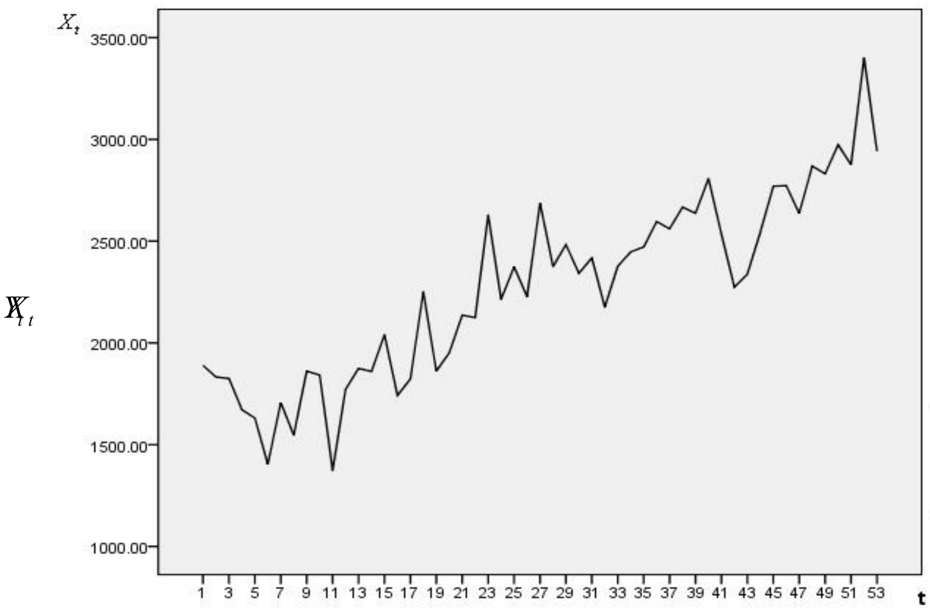
$$Y_t = S_t + T_t \times C_t \times I_t$$

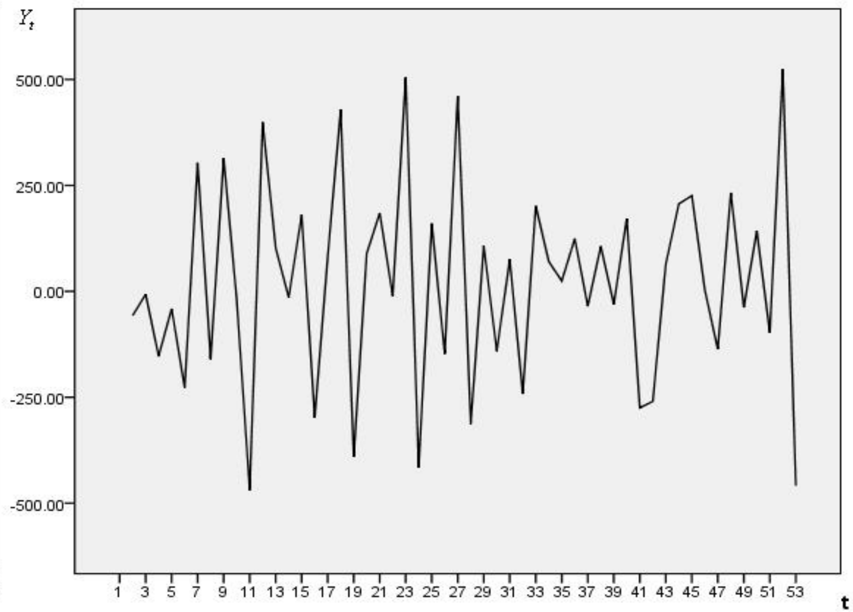
- 时间序列分解模型的选取需要考虑到现象变化的规律和数据本身的特征，如果季节变动（循环变动、不规则变动）依赖于长期趋势的变化，则宜选用乘法模型或加乘混合模型，否则可以考虑加法模型。

3.2 ARIMA模型

- **差分运算**：相距k期的两个序列值之间的减法运算称为k阶差分运算
- **ARIMA模型**：差分运算具有强大的确定性信息提取能力，许多非平稳序列差分后会显示出平稳序列的性质，这种序列称为差分平稳序列，差分后的序列可用ARIMA序列拟合，ARIMA模型实质就是差分运算与ARMA序列的组合。
- **对于有趋势性时间序列通常采用ARIMA模型进行分析。**
- 对于有季节性的时间序列可以采用乘积季节ARIMA模型进行预测。由于这类模型比较复杂，本课程不做介绍。

3.2 ARIMA模型





该序列有增长的趋势，首先对其进行一阶差分，原序列及一阶差分后序列如图所示。

3.2 ARIMA模型

随机游动模型的金融意义一般p_t是对数价格，则ε_t是零均值的对数收益率。实际的对数收益率常常是非零的，正数居多。所以，模型可以推广为

$$p_t = \mu + p_{t-1} + \varepsilon_t,\ t = 1, 2, \ldots$$

其中$\{\varepsilon_t\}$仍为零均值独立同分布白噪声列。常数μ并不代表均值，而是对数价格p_t的增长速度，称为模型的**漂移**(drift)。设初始价格为p_0，则

$$\begin{aligned}
p_1 =& p_0 + \mu + \varepsilon_1 \\
p_2 =& p_0 + 2\mu + \varepsilon_1 + \varepsilon_2 \\
& \cdots\cdots \\
p_t =& p_0 + t\mu + \varepsilon_1 + \cdots + \varepsilon_t
\end{aligned}$$

于是

$$E(p_t|p_0) = p_0 + \mu t, \quad \mathrm{Var}(p_t|p_0) = \sigma^2 t$$

所以带漂移的随机游动与不带漂移的随机游动相比，其条件方差不变，但是条件均值多了一个随t线性增长（若$\mu > 0$）的μt项。

这时，实际的序列的图形将沿着$y = p_0 + \mu t$这条直线附近变化。如果$\mu = 0$，则图形也有缓慢的水平变化但是没有这样的固定趋势。

3.2 ARIMA模型

所以带漂移的随机游动与不带漂移的随机游动相比， 其条件方差不变， 但是条件均值多了一个随t线性增长（若$\mu > 0$）的μt项。

这时，实际的序列的图形将沿着$y = p_0 + \mu t$这条直线附近变化。 如果$\mu = 0$，则图形也有缓慢的水平变化但是没有这样的固定趋势。

带漂移的随机游动p_t， 可以分解为两部分：

$$p_t = (p_0 + \mu t) + p_t^*$$

其中$p_t^* = \sum_{j=1}^{t} \varepsilon_t$是从0出发的不带漂移的随机游动， $p_0 + \mu t$是一个非随机的线性趋势。

3.2 ARIMA模型

将带漂移的随机游动模型中的白噪声替换成一个ARMA平稳列，其主要的性质仍能保留。即

$$Y_t = Y_{t-1} + \mu + X_t$$

其中$\{X_t\}$是零均值平稳可逆ARMA(p, q)平稳列。这时有

$$Y_t = Y_0 + \mu t + \sum_{j=1}^{t} X_j$$

于是

$$E(Y_t|Y_0) = Y_0 + \mu t, \quad \mathrm{Var}(Y_t|Y_0) = \text{随}t\text{增大而趋于}\infty$$

当$\mu = 0$时，$\{Y_t\}$的均值固定。Y_t的条件方差为t的线性函数。X_{t-j}对Y_t的影响不衰减，是永久有影响的。这些表现与带漂移的随机游动基本相同。

称$\{Y_t\}$服从ARIMA$(p, 1, q)$模型，是非平稳的。$\{Y_t\}$不能通过减去任何的非随机趋势变成平稳。但是，差分运算

$$\Delta Y_t = (1 - B)Y_t = Y_t - Y_{t-1} = \mu + X_t$$

将ARIMA$(p, 1, q)$序列$\{Y_t\}$转化成平稳可逆的ARMA(p, q)序列。

3.2 ARIMA模型

设零均值平稳可逆ARMA(p,q)序列$\{X_t\}$的模型为

$$P(B)X_t = Q(B)\varepsilon_t$$

其中$\{\varepsilon_t\}$为零均值独立同分布白噪声列，$P(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$，
$Q(z) = 1 + \theta_1 z + \cdots + \theta_q z^q$。因为$(1-B)Y_t = \mu + X_t$，所以

$$P(B)(1-B)Y_t = P(B)(\mu + X_t) = P(1)\mu + P(B)X_t = P(1)\mu + Q(B)\varepsilon_t$$

记$\tilde{P}(z) = P(z)(1-z)$, $\phi_0 = P(1)\mu$，则Y_t的模型可以写成

$$\tilde{P}(B)Y_t = \phi_0 + Q(B)\varepsilon_t$$

这个模型形式上与一个ARMA(p,q)模型相同，但是其$p+1$个特征根中有一个等于1，其余特征根才在单位圆外。

对ARIMA模型建模，只要计算Y_t的差分，然后对差分建立ARMA模型即可。

有些序列需要二阶差分才能平稳，二阶差分即

$$\begin{aligned}\Delta^2 Y_t =&(1-B)^2 Y_t = (1-B)(Y_t - Y_{t-1}) = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2})\\ =&Y_t - 2Y_{t-1} + Y_{t-2}\end{aligned}$$

3.2 ARIMA模型

如果$\xi_t = a + bt$，则

$$(1 - B)\xi_t = b$$

即差分还可以消除线性趋势。

如果 $\xi_t = a + bt + ct^2$，则

$$(1 - B)^2 = 2c$$

即二阶差分可以消除二次多项式趋势。