Time Series Analysis

Lollins

2023 年 9 月 21 日

前言

时间序列分析的笔记，图像大多用 Python 完成，源代码放在 code 文件夹下。

Lollins

2023 年 9 月 21 日

目录

第一章　时间序列分析简介

1.1　时间序列的定义

1.1.1　随机过程中的一些基本定义

定义 1.1.1 (概率空间). $(\Omega, \mathcal{F}, \mathcal{P})$ 称为概率空间当且仅当其满足以下的条件：

1. Ω-*Simple Space* 样本空间，试验中所有可能结果的集合。

2. $\mathcal{F}$-*Sets of events* 事件集合，是 Ω 子集构成的集合，即 $\mathcal{F} \subset 2^{\Omega}$($\Omega$的幂集)，并且它需要满足以下三个特征 (也就是必须是$\sigma - field$)*:*

- $\Phi \subset \mathcal{F}$ (即必须包含不可能事件)*;*
- 如果 $A \subset \mathcal{F}$，那么 $A^c \subset \mathcal{F}$*;*
- 如果 $A_1, A_2, \ldots, A_i \in \mathcal{F}$,那么$\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$*.*

3. $\mathcal{P}$-*Probability measure* 概率测度 (或概率)，描述一次随机试验中被包含在 $\mathcal{F}$ 中的所有事件的可能性。并且它也类似的需要满足以下三个特征*:*

- $0 \leq \mathcal{P}(A) \leq 1$，即限制了总测度为 *1;*
- $\mathcal{P}(\Omega) = 1$，即包含样本空间且概率为 *1;*
- 如果 $A_1, A_2, \ldots, A_i$ 为互斥事件，那么 $\mathcal{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathcal{P}(A_i)$

定义 1.1.2 (随机过程). 随机过程是概率空间 $(\Omega, \mathcal{F}, \mathcal{P})$ 上的一族随机变量 $\{X(t), t \in T\}$，其中 t 是参数，它属于某个指标集 T，T 称为参数集。

注：依据参数集 T 可分为**离散参数过程**和**连续参数过程**，把**离散参数过程**称为**随机序列**。

1.1.2　随机序列定义

定义 1.1.3 (随机序列). 按时间顺序排列的一组随机变量

$$\ldots, X_1, X_2, \ldots, X_t, \ldots$$

定义 1.1.4 (观察值序列). 随机序列的 n 个有效观察值，称之为序列长度为 n 的观察值序列

$$x_1, x_2, \ldots, x_t$$

随机序列与观察值序列的关系

- 观察值序列是随机序列的一个实现;
- 通过分析观察值序列的性质，由观察值序列的性质来推断随机序列的性质.

随机序列的特点

- 随机序列中数据的位置与实践有关;
- 时间序列是对相关指标变量在不同时间进行观察所得到的结果。是一个观察结果，一个实现;
- 时间序列中的数据可以是一个时期内的数据也可以是一个时点上的数据。流量与存量;
- 时间序列通常存在前后时间上的相依性.

时间序列的分类

- 按所研究对象的维度，有一元时间序列和多元时间序列;
- 按观察时间的连续与否可将时间序列分为离散时间序列和连续时间序列;
- 按时间序列的统计特征分，有平稳时间序列和非平稳时间序列两类.

时间序列分析的目的

时间序列分析所讨论的内容，就是对观察到的时间序列数据进行研究，分析时间序列的统计特征和依存关系，找出系统的内在发展规律, 建立起能反映变量变化的动态模型，并将这种动态模型用于预测等应用领域.

1.2　时间序列分析方法

1.2.1　描述性时间序列分析

1.2.2　统计时间序列分析

$$\begin{cases} \text{频域分析方法} \\ \text{时域分析方法} \end{cases}$$

1.3　时间序列分析软件

Matlab，Eviews，R 和 SAS 等。

第二章　时间序列的预处理

2.1　平稳序列的定义

2.1.1　特征统计量

一、概率分布

概率分布的意义

随机变量族的统计特征完全由它们的联合分布函数或联合密度函数决定。

定义 2.1.1 (时间序列概率分布族). 对于时间序列 $\{X_t, t \in T\}$，任取正整数 m，任取 $t_1, t_2, \cdots, t_m \in T$，则 m 维随机向量 $(X_1, X_2, \cdots, X_m)'$ 的联合概率分布记为 $F_{t_1,t_2,\cdots,t_m}(x_1, x_2, \cdots, x_m)$，由这些有限维分布函数构成的全体

$$\{F_{t_1,t_2,\cdots,t_m}(x_1, x_2, \cdots, x_m), \forall m \in N, \forall t_1, t_2, \cdots, t_m \in T\} \tag{2.1}$$

就称为序列 $\{X_t\}$ 的概率分布族。

注：默认 t 为正整数。

实际应用的局限性

在实际应用中, 要得到序列的联合概率分布几乎是不可能的, 而且联合概率分布通常涉及非常复杂的数学运算, 这些原因导致我们很少直接使用联合概率分布进行时间序列分析。

二、特征统计量

- **均值：** $\mu_t = EX_t = \int_{-\infty}^{\infty} x dF_t(x)$
- **方差：** $DX_t = E(X_t - \mu_t)^2 = \int_{-\infty}^{\infty} (x-\mu_t)^2 dF_t(x)$
- **自协方差：** $\gamma(t,s) = E(X_t - \mu_t)(X_s - \mu_s)$
- **自相关系数：** $\rho(t,s) = \frac{\gamma(t,s)}{\sqrt{DX_t \cdot DX_s}}$

注：$\gamma(s,t) = \gamma(t,s)$，$|\rho(t,s)| \le 1$

2.1.2　平稳时间序列的定义

一、严平稳

严平稳是一种条件比较苛刻的平稳性定义，它认为只有当序列所有的统计性质都不会随着时间的推移而发生变化时，该序列才能被认为平稳。

定义 2.1.2 (严平稳)**.** 设 $\{X_t\}$ 为一时间序列,对 $\forall$ 正整数 m,$\forall t_1, t_2, \cdots, t_m \in T$，$\forall$正整数$\tau$，有

$$F_{t_1,t_2\cdots t_m}(x_1,x_2,\cdots,x_m) = F_{t_{1+\tau},t_{2+\tau}\cdots t_{m+\tau}}(x_1,x_2,\cdots,x_m) \tag{2.2}$$

则称时间序列 $\{X_t\}$ 为严平稳时间序列。

二、宽平稳

宽平稳是使用序列的特征统计量来定义的一种平稳性。它认为序列的统计性质主要由它的低阶矩决定，所以只要保证序列低阶矩平稳（二阶），就能保证序列的主要性质近似稳定。

定义 2.1.3 (宽平稳)**.** 设 $\{X_t\}$ 为一时间序列，若满足以下条件

1. $EX_t^2 < \infty, \forall t \in T$

2. $EX_t = \mu, \mu$为常数,$\forall t \in T$

3. $\gamma(t,s) = \gamma(k, k+s-t), \quad \forall t, s, k$且$k+s-t \in T$

则称时间序列 $\{X_t\}$ 为宽平稳时间序列。

三、严平稳与宽平稳的关系

$$\begin{cases} \text{严平稳} \not\Rightarrow \text{宽平稳} \\ \text{严平稳} \not\Leftarrow \text{宽平稳} \\ \text{严平稳} + \text{二阶矩存在} \Rightarrow \text{宽平稳} \\ \text{在正态分布假设下，严平稳} \Leftrightarrow \text{宽平稳} \end{cases}$$

注：以后见到的平稳时间序列，如果没有特别注明，都是**宽平稳时间序列**。如果序列不满足平稳条件，就称为**非平稳时间序列**。

例 2.1.4 (白噪声序列 (White Noise))**.** 一个简单的随机时间序列是一具有零均值同方差的独立分布序列：$X_t = \mu_t, \mu_t\ N(0, \sigma^2)$.

解. 由于 X_t 具有相同的均值和方差，且协方差等于 0，有定义可以得到**白噪声序列是平稳的** □

例 2.1.5 (随机游走序列 (Random Walk))**.** 该随机时间序列是由随机过程 $X_t = X_{t-1} + \mu_t$ 生成，其中 μ_t 是一个白噪声序列.

解. 由 $\mu_t\ N(0, \sigma^2)$，我们可以知道该序列的均值满足

$$E(X_t) = E(X_{t-1}) + E(\mu_t) = E(X_{t-1})$$

我们设序列 X_t 的初值为常数 X_0，则

$$
\begin{aligned}
X_1 &= X_0 + \mu_1 \\
X_2 &= X_1 + \mu_2 = X_0 + \mu_1 + \mu_2 \\
\cdots &= \cdots \\
X_t &= X_0 + \sum_{i=1}^{t} \mu_i
\end{aligned}
$$

由于 X_0 为常数，μ_t 为一个白噪声，所以 $Var(X_t) = t\sigma^2$
即 X_t 的方差和时间有关而非常数，它是一个**非平稳序列**. □

2.1.3　平稳时间序列的统计性质

一、常数均值

$$EX_t = \mu, \forall t \in T \tag{2.3}$$

二、自协方差函数和自相关系数只依赖于时间的平移长度而与时间的起止点无关

- 延迟 k 自协方差函数

$$\gamma(k) = \gamma(t, t+k), \forall k\text{为整数} \tag{2.4}$$

- 延迟 k 自相关系数

$$\rho_k = \frac{\gamma(k)}{\gamma(0)} \tag{2.5}$$

由式2.5，我们可以得到**自相关系数**的如下性质：

- **规范性：** $\rho_0 = 1, \text{且} |\rho_k| \leq 1, \forall k;$
- **对称性：** $\rho_k = \rho_{-k};$

- **非负定性：**$\Gamma_m = \begin{pmatrix} \rho_0 & \rho_1 & \cdots & \rho_{m-1} \\ \rho_1 & \rho_0 & \cdots & \rho_{m-2} \\ \vdots & \vdots & \cdots & \vdots \\ \rho_{m-1} & \rho_{m-2} & \cdots & \rho_0 \end{pmatrix}$,Γ_m为非负定自协方差系数阵;
- **非唯一性：**一个平稳时间序列一定唯一决定了它的自相关函数，但一个自相关函数未必唯一对应着一个平稳时间序列.

例 2.1.6 (平稳序列的线性变换). 平稳序列 $\{X_t\}$，期望 μ，自协方差函数 $\gamma(t)$，经线性变化后，还是平稳序列。

证明. 令 $Y_t = a + bX_t$，则 $EY_t = a + b\mu$，$\mathrm{Cov}(Y_s, Y_{s+t}) = b^2\mathrm{Cov}(X_t, X_{t+s}) = b^2\gamma(t)$，可见 $\{Y_t\}$ 平稳。若取

$$Y_t = \frac{X_t - \mu}{\sqrt{\gamma_0}}$$

则 $EY_t = 0, \mathrm{Var}(Y_t) = 1$，称 $\{Y_t\}$ 为 $\{X_t\}$ 的标准化序列。 □

例 2.1.7 (调和平稳序列). 设 a，b 为常数，随机变量 U 在 $(-\pi, \pi)$ 内均匀分布，则

$$X_t = bcos(at + U), t \in \mathbb{Z}$$

是平稳序列，称为调和平稳序列。

证明.

$$\begin{aligned} EX_t &= \frac{1}{2\pi}\int_{-\pi}^{\pi} b\cos(at+u)du = 0, \\ E(X_tX_s) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} b^2\cos(at+u)\cos(as+u)du \\ &= \frac{1}{2}b^2\cos((t-s)a), \end{aligned}$$

证毕！ □

2.1.4 平稳性的重大意义

- 在平稳序列场合，序列的均值等于常数，这意味着原本含有可列多个随机变量的均值序列变成了只含有一个变量的常数序列。

$$\{\mu_t, t \in T\} \Rightarrow \{\mu, t \in T\} \tag{2.6}$$

- 原本每个随机变量的均值（方差，自相关系数）只能依靠唯一的一个样本观察值去估计，现在由于平稳性，每一个统计量都将拥有大量的样本观察值。
- 这极大地减少了随机变量的个数，并增加了待估变量的样本容量。极大地简化了时序分析的难度，同时也提高了对特征统计量的估计精度。

2.2 序列的平稳性检验

检验的方法分为**图检验**与**构造检验统计量进行假设检验**。

图检验分为时序图检验和自相关图检验。

2.2.1 时序图检验

得到观察值序列后，我们绘制该序列的时序图，根据时序图的不同特征，可以把序列分为三大类。

1、序列具有明显的趋势特征

如果序列具有明显的趋势特征，那么该序列显然不具有常数均值，所以趋势序列可以判断为非平稳。

例 2.2.1. *图2.1时序图显示：序列有明显的递增趋势特征，所以是非平稳序列。*

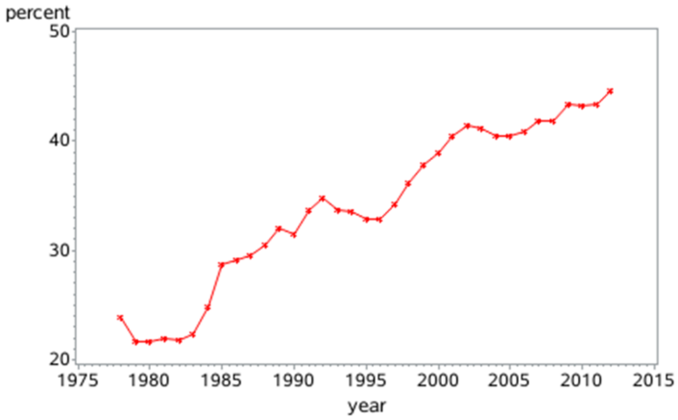


图 2.1 1978-2012 年我国第三产业占国内生产总值

2、序列具有明显的周期特征

具有周期特征的序列平稳性识别是困难的。理论上，如果周期波动的振幅和频率不随着时间的变化而变化，通常序列是平稳的，例如例2.1.7中的调和平稳序列。但是如果 $X_t = b_t cos(a_t t + U)$，振幅和频率随着时间的变化而变化，那么序列 $\{X_t\}$ 就是非平稳序列。

3、序列既没有趋势特征，也没有周期特征

如果一个序列既没有明显的趋势特征，也没有明显的周期特征，几乎是围绕着一个常数附近做有解波动，这通常是平稳序列。

2.2.2　自相关图检验

例 2.2.2. *利用图检验方法判断 1915-2004 年澳大利亚自杀率序列（每 10 万人自杀人口数）的平稳性，如图2.2所示。*

图2.2难以把握该序列的平稳性，我们可以借助自相关图的性质来进一步辅助判别。

Python 中可以用 statsmodels 库来绘制自相关图，如图2.3所示。

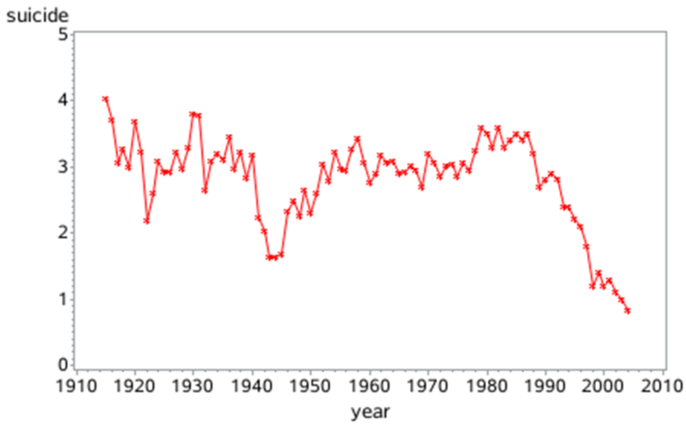


图 2.2

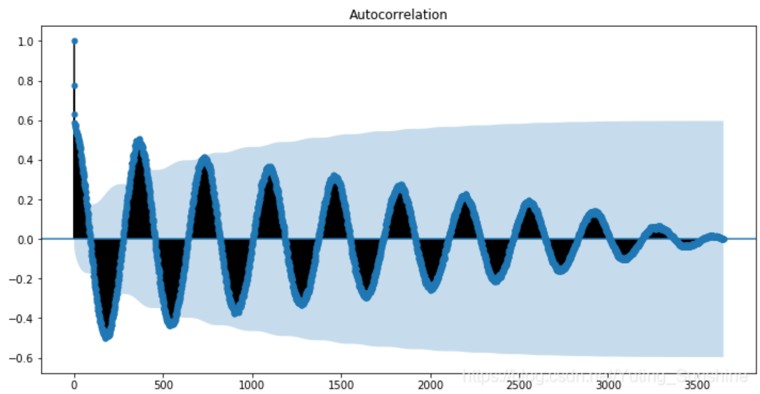
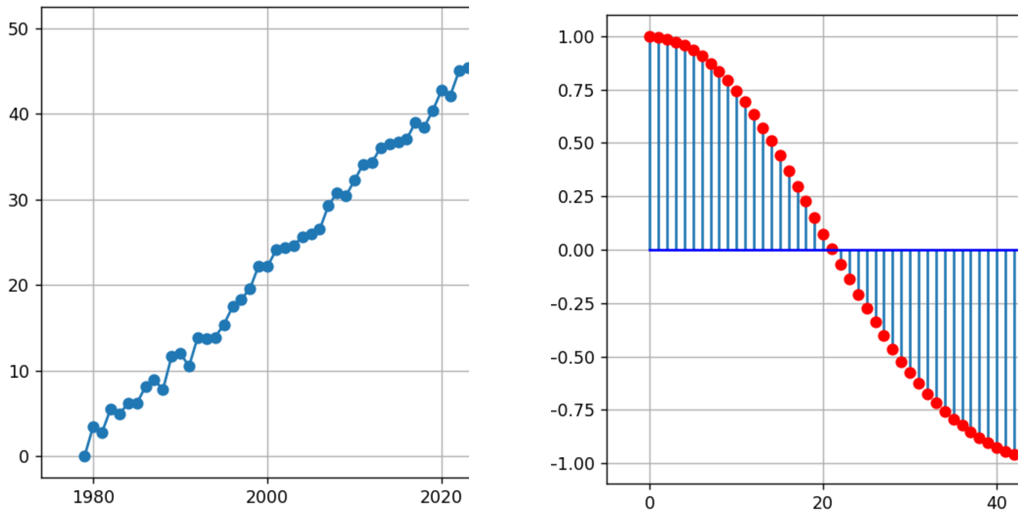


图 2.3

自相关图是一个平面二维坐标悬垂线图，横坐标表示延迟时期数，纵坐标表示自相关系数，悬垂线的长度表示自相关系数的大小。

1、序列具有明显的趋势特征

该序列的时序图与自相关图如图2.4所示，自相关图呈现明显的倒三角特征，是非平稳序列，代码为 `2_4.py`。



(a) 时序图 (b) 自相关图

图 2.4

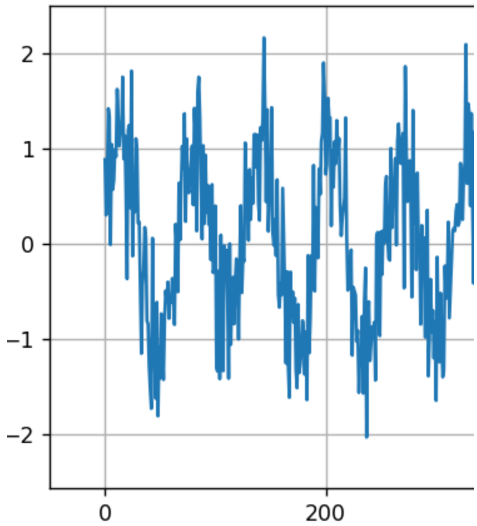
2、序列具有明显的周期特征

该序列的时序图与自相关图如图2.5所示，自相关图呈现明显的周期特征，是非平稳序列，代码为 `2_5.py`。

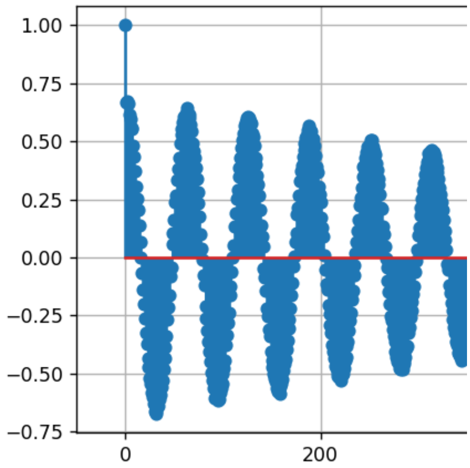
对于有周期性的序列，平稳与非平稳时间序列的自相关图的图像很相像，但是平稳时间序列的自相关图随着之后 (lag) 的增加，自相关系数减小的更快。从图2.6中可以看出，代码为 `2_5a.py`。

3、平稳时间序列

平稳时间序列的自相关图，除了 0 阶延迟为 1，其余各阶自相关系数都很小，而且自相关系数时正时负，没有倒三角特征，也没有周期特征，如图2.7所示，代码为 `2_6.py`。



(a) 时序图



(b) 自相关图

图 2.5

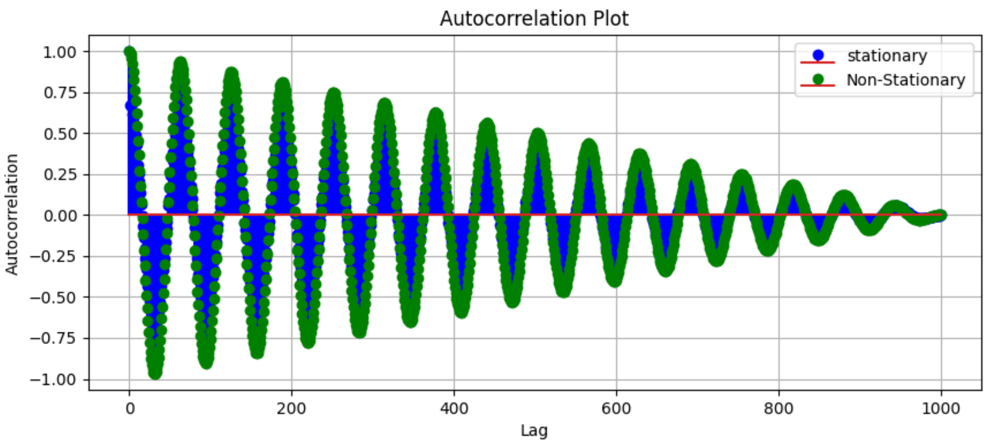
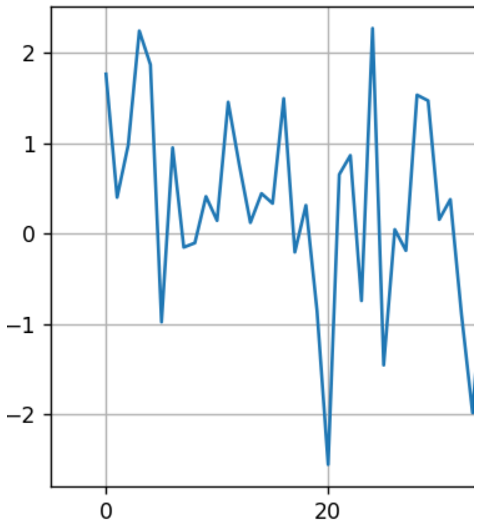
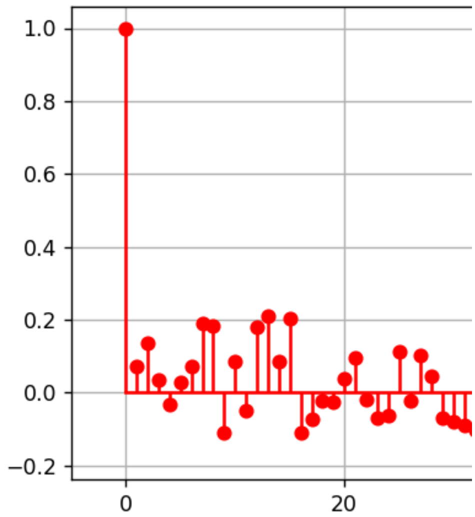


图 2.6 蓝色为平稳时间序列，绿色为非平稳



(a) 时序图



(b) 自相关图

图 2.7

2.3　纯随机性检验

拿到一个观察值序列后，首先判断它的平稳性。通过**平稳性检验**，可以将序列分为**平稳和非平稳**。

对于平稳时间序列，我们有一套非常成熟的平稳序列建模方法。但是，并**不是所有的平稳时间序列都值得建模**。

如果序列值彼此之间没有任何相关性，那就意味着该序列是一个没有记忆的序列，过去的行为对将来的发展没有丝毫影响，这种序列称为**纯随机序列**。从统计分析的角度来说，纯随机序列是没有任何分析价值的序列。

2.3.1　纯随机序列的定义

定义 2.3.1 (白噪声 (White Noise) 序列). *如果对于时间序列* $\{X_t\}$ *满足以下性质：*

1、 任取 $t \in T$*，有* $E(X_t) = \mu$*；*

2、 任取 $t \in T$，有

$$\gamma(t,s)=\begin{cases}\sigma^2, t=s\\ 0, t\neq s\end{cases} \tag{2.7}$$

称序列 $\{X_t\}$ 为纯随机序列，也称为白噪声 *(White Noise)* 序列，简记为 $X_t \sim WN(\mu,\sigma^2)$。容易证明，白噪声序列是平稳序列，而且是最简单的平稳序列。

2.3.2　纯随机序列的性质

1、纯随机性

白噪声序列具有以下性质：

$$\gamma(k)=0, \forall\ k\neq 0 \tag{2.8}$$

这说明白噪声序列之间没有任何相关联系，这种“没有记忆”的序列就是**白噪声序列**。

如果序列值之间呈现出某种显著的相关分析：

$$\gamma(k)\neq 0, \exists\ k\neq 0 \tag{2.9}$$

说明**该序列不是纯随机序列**，该序列间隔 k 期之间存在一定程度的相互影响。

2、方差齐性

所谓的方差齐性，是指序列中每一个变量的方差都相等，即

$$DX_t=\gamma(0)=\sigma^2 \tag{2.10}$$

如果序列不满足方差齐性，就称该序列具有**异方差性质**。

2.3.3　纯随机性检验

如果一个序列是纯随机的，那么它的序列值之间应该没有任何相关关系，即式2.8所示。

Bartlett 证明，如果一个时间序列是纯随机的，得到一个观察期数为 n 的观察序列 $\{x_t,\ t=1,2,...,n\}$，那么该序列的延迟非零期的样本自相关系数将近服从均值为 0，方差为序列观察期倒数的正态分布，即

$$\hat{\rho_k} \overset{\bullet}{\sim} N(0,\frac{1}{n}),\ \forall k\neq 0 \tag{2.11}$$

式中，n 为序列观察期数。

根据 Bartlett 定理，我们可以构造检验统计量来检验序列的纯随机性。

1、假设条件

原假设：延迟期数小于或等于 m 期的序列值之间相互独立，即序列是延迟到 m 期都是平稳，纯随机的。备择假设：延迟期数小于或等于 m 期的序列值之间有相关性。该假设条件用数学语言描述为：

$$\begin{aligned} &H_0:\rho_1=\rho_2=\cdots=\rho_m=0,\forall m\geqslant 1\\ &H_1:\text{至少存在某个}\rho_k\neq 0,\forall m\geqslant 1,k\leqslant m \end{aligned} \tag{2.12}$$

2、检验统计量

(1)Q 统计量

Box 和 Pierce 推导出了 Q 统计量：

$$Q=n\sum_{k=1}^{m}\hat{\rho}_k^2 \overset{\bullet}{\sim} \chi^2(m) \tag{2.13}$$

n 为序列观察期数，m 为指定延迟期数。

(2)LB 统计量

实际应用中，人们发现 Q 统计量在大样本场合检验效果很好，但在小样本场合不太准确。为了弥补这一缺陷，Ljung 和 Box 又推导了 LB 统计量

$$LB = n(n+2)\sum_{k=1}^{m}\left(\frac{\hat{\rho}_k^2}{n-k}\right) \overset{\bullet}{\sim} \chi^2(m) \tag{2.14}$$

2.3.4 Eviews 差分指令

对序列 x_t 做差分，即

$$y_t = x_t - x_{t-1} \tag{2.15}$$

指令如下为 genr y = D(x)，同理，二阶差分为 genr y = D(D(x))。

第三章 ARMA 模型的性质

3.1 Wold 分解定理

3.1.1 Wold 分解定理的内容

定义 3.1.1 (Wold 分解定理). *对于任意一个离散平稳时间序列 $\{x_t\}$, 它都可以分解为两个不相关的平稳序列之和, 其中一个为确定性的 (deterministic), 另一个为随机性的 (stochastic), 不妨记作:*

$$x_t = V_t + \xi_t \tag{3.1}$$

确定性序列 $\{V_t\}$ 代表序列的当期波动可以由其历史信息预测的部分。Wold 证明，平稳时间序列的确定性部分一定可以表达为历史序列值的线性组合：

$$V_t = \sum_{j=1}^{\infty} \phi_j x_{t-j} \tag{3.2}$$

随机序列 $\{\xi_t\}$ 代表序列的当期波动不能由其历史信息预测的部分。Wold 证明，这部分信息可以等价表达为：

$$\xi_t = \sum_{j=0}^{\infty} \theta_j \varepsilon_{t-j} \tag{3.3}$$

式中，$\theta_0 = 1, \sum_{j=0}^{\infty} \theta_j^2 < \infty$，$\{\epsilon_t\}$ 称为新息过程 (innovation process)，是每个过程新加入的随机信息。ε_t 为白噪声序列，序列值之间相互独立，不可预测，通常假定 $\{\epsilon_t\} \overset{i.i.d}{\sim} N(0, \sigma_\varepsilon^2)$

具有式3.2结构的模型是 1927 年 Yalu 提出的自回归 (autoregression) 模型，简称 AR 模型。具有式3.3结构的模型是 1931 年 Walker 提出的移动平均 (moving average) 模型，简称 MA 模型。

3.2　AR 模型

3.2.1　差分的定义

- 一阶差分：$\nabla x_t = x_t - x_{t-1}$
- p 阶差分：$\nabla^p x_t = \nabla^{p-1} x_t - \nabla^{p-1} x_{t-1}$
- k 步差分：$\nabla_k x_t = x_t - x_{t-k}$

定义 B 为一个延迟算子，即 $B^p = x_{t-p}$，则我们有

- 一阶差分：$\nabla x_t = (1-B)x_t$
- 二阶差分：$\nabla^p x_t = (1-B)^p x_t$
- k 步差分：$\nabla_k x_t = (1-B^k)x_t$

差分算子有如下性质：

1、　$B^0 = 1$；

2、　常数的任意结束延迟仍然等于常数，即 $B^p c = c$；

3、　对于任意常数，有 $B^p(cx_t) = cx_{t-p}$

4、　对任意两个序列 $\{x_t\},\{y_t\}$ 有 $B(x_t \pm y_t) = x_{t-1} \pm y_{t-1}$

关于差分方程的解法，可以参考周义仓老师的《差分方程及其应用》第二章的内容。

3.2.2　AR 模型的定义

定义 3.2.1. 具有如下结构的模型称为 p 阶自回归模型，简记为 *AR(p)*:

$$\begin{cases} x_t = \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \ldots + \phi_\mu x_{t-\rho} + \varepsilon_t \\ \phi_p \neq 0 \\ E(\varepsilon_t) = 0, \mathrm{Var}(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t \\ E(x_s \varepsilon_t) = 0, \forall s < t \end{cases} \tag{3.4}$$

AR(p) 模型有三个限制条件：

1、 $\phi_p \neq 0$。保证模型的最高阶数为 p。

2、 $E(\varepsilon_t) = 0,\ \mathrm{Var}(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t$。保证随机干扰项 $\{\varepsilon\}$ 为零均值白噪声序列。

3、 $E(x_s \varepsilon_t) = 0, \forall s < t$。保证当期的随机干扰项与过去的序列值无关。

当 $\phi_0 = 0$，自回归模型又称为中性化 AR(p) 模型。非中心化 AR(p) 序列都可以通过下面的变换转化为中心化 AR(p) 模型。令

$$\mu = \frac{\phi_0}{1 - \phi_1 - \cdots - \phi_p}, y_t = x_t - \mu \tag{3.5}$$

则 $\{y_t\}$ 为 $\{x_t\}$ 的中心化序列。中心化序列就是将整个非中心化序列平移一个常数，这个整体的移动对序列值之间的相关关系没有任何影响。

引入了延迟算子，中心化 AR(p) 模型又可以简记为

$$\Phi(B) x_t = \varepsilon_t \tag{3.6}$$

式中，$\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p$，称为 p 阶自回归系数多项式。

3.2.3　AR 模型的平稳性判别

并非所有的 AR 模型都是平稳的，图示法只是一种粗糙的直观判别方法，我们有两种准确的平稳性判断方法：特征根判别法和平稳域判别。

1、特征根判别法

任一中心化 AR(p) 模型 $\Phi(B)=\varepsilon_t$ 都可以视为一个非齐次线性差分方程

$$x_t-\phi_1 x_{t-1}-\phi_2 x_{t-2}-...-\phi_\rho x_{t-p}=\varepsilon_t \tag{3.7}$$

$\{x_t\}$ 平稳要求差分方程的特征根都在单位圆内，即

$$|\lambda_i|<1,\ i=1,2,...,p \tag{3.8}$$

2、平稳域判别法

从参数向量 $(\phi_1,\phi_2,...,\phi_p)'$ 角度考虑，从 p 维欧氏空间取一个子集，满足：

$$\{\phi_1,\phi_2,\cdots,\phi_p|\text{ 特征根都在单位圆内}\}$$

称为 AR(p) 模型的平稳域。

对于低阶 AR 模型，用平稳域的方法判别模型的平稳性通常更为简便。

(1)AR(1) 模型的平稳域

AR(1) 模型为：$x_t=\phi_1 x_{t-1}+\varepsilon_t$，根据特征根判别法，可以求得 AR(1) 模型的平稳域为 $\{\phi_1|-1<\phi_1<1\}$

(2)AR(2) 模型的平稳域

AR(2) 模型为：$x_t=\phi_1 x_{t-1}+\phi_2 x_{t-2}+\varepsilon_t$，根据特征根判别法，可以求得 AR(2) 模型的平稳域为 $\{\phi_1,\phi_2|\ |\phi_2|<1,\phi_2\pm\phi_1<1\}$

3.2.4　平稳 AR(p) 模型性质

1、均值

对于 AR 模型满足方程3.4，两边同时取期望

$$Ex_t=E(\phi_0+\phi_1 x_{t-1}+\phi_2 x_{t-2}+\cdots+\phi_p x_{t-p}+\varepsilon_t)$$

根据平稳时间序列的常数均值性质，有 $Ex_t = \mu \ (\forall \ t \in T)$，上式等价于

$$(1 - \phi_1 - \cdots - \phi_p)_\mu = \phi_0$$
$$\Rightarrow \mu = \frac{\phi_0}{1 - \phi_1 - \cdots - \phi_p}$$

特别的，对于中心化 AR(p) 模型，$Ex_t = 0$

2、方差

为求平稳 AR(p) 模型的方差，需要借助 Green 函数的帮助。

定义 3.2.2 (Green 函数)**.** 假设 $\{x_t\}$ 为任意阶数的平稳 AR 模型，那么一定存在一个常数序列 $\{G_j\}(j = 0, 1, 2...)$，使得 $\{x_t\}$ 可以等价为纯随机序列 $\{\varepsilon_t\}$ 的线性组合，即

$$\begin{aligned} x_t &= \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + \varepsilon_t \\ &= G_0 \boldsymbol{\varepsilon}_t + G_1 \boldsymbol{\varepsilon}_{t-1} + G_2 \boldsymbol{\varepsilon}_{t-2} \cdots \end{aligned} \tag{3.9}$$

这个常数序列 $\{G_j\}$ 就称为 *Green* 函数。

基于 Green 函数，我们可以求出 AR(p) 模型的方差为

$$Var(x_t) = \left(G_0^2 + G_1^2 + G_2^2 + \cdots\right) \sigma_\varepsilon^2 \tag{3.10}$$

其中 G_j 满足 (递推可求出)

$$\begin{cases} G_0 = 1 \\ G_j = \sum\limits_{k=1}^{j} \phi'_k G_{j-k}, \ j = 1, 2, \cdots \end{cases} \text{式中} \phi'_k = \begin{cases} \phi_k, k \le p \\ 0, k > p \end{cases} \tag{3.11}$$

3、协方差函数

把方程3.4中式 1 乘 x_{t-k}，再求期望，得

$$E(x_t x_{t-k}) = \phi_1 E(x_{t-1} x_{t-k}) + \cdots + \phi_p E(x_{t-p} x_{t-k}) + E(\varepsilon_t x_{t-k})$$

再利用 $E(\varepsilon_t x_{t-k}) = 0$，得

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \cdots + \phi_p \gamma_{k-p} \quad , \forall k \ge 1 \tag{3.12}$$

4、自相关系数

根据自相关系数的定义 $\rho_k = \frac{\gamma_k}{\gamma_0}$，我们可以得到

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \cdots + \phi_p \rho_{k-p} \tag{3.13}$$

AR 模型的自相关系数的表达式实际上是一个齐次差分方程, 它的通解形式

$$\rho_k = \sum_{i=1}^{p} c_i \lambda_i^k \qquad |\lambda_i| < 1, 且 c_1, \cdots, c_p 为任意实数$$

根据自相关系数通解的形式，我们可以推出如下性质：

- 呈指数衰减：$|\lambda_i| < 1 \quad \Rightarrow \quad \rho_k = \sum_{i=1}^{p} c_i \lambda_i^k \to 0$
- 拖尾性：$c_1, \cdots, c_p$为任意实数 $\Rightarrow \ \rho_k = \sum_{i=1}^{p} c_i \lambda_i^k$不会恒等于零,$\forall k >$ 某个常数

5、偏自相关系数

定义 3.2.3 (偏自相关系数). 对于平稳 *AR(p)* 模型，滞后 k 期的偏自相关系数就是指定中间的 *k-1* 个随机变量 $x_{t-1}, x_{t-2}, ..., x_{t-k+1}$ 的条件下，x_t 对 x_{t-k} 的相关度量，用数学语言描述为

$$\rho_{x_t, x_{t-k}|x_{t-1}, \cdots, x_{t-k+1}} = \frac{E[(x_t - \hat{E}x_t)(x_{t-k} - \hat{E}x_{t-k})]}{E[(x_{t-k} - \hat{E}x_{t-k})^2]} \tag{3.14}$$

对于方程 $x_t = \phi_{k1} x_{t-1} + \phi_{k2} x_{t-2} + \cdots \phi_{k(k-1)} x_{t-k+1} + \phi_{kk} x_{t-k}$ 等号两边同时乘 x_{t-l}，再取期望，除以方差，得

$$\rho_l = \phi_{k1} \rho_{l-1} + \phi_{k2} \rho_{l-2} + \cdots + \phi_{kk} \rho_{l-k} \quad , \forall l > 0$$

取前 k 个方程，构成 Yule-Walker 方程组，即

$$\begin{cases} \rho_1 = \phi_{k1} \rho_0 + \phi_{k2} \rho_1 + \cdots + \phi_{kk} \rho_{k-1} \\ \rho_2 = \phi_{k1} \rho_1 + \phi_{k2} \rho_0 + \cdots + \phi_{kk} \rho_{k-2} \\ \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \rho_k = \phi_{k1} \rho_{k-1} + \phi_{k2} \rho_{k-2} + \cdots + \phi_{kk} \rho_0 \end{cases} \tag{3.15}$$

写成矩阵形式，即

$$\begin{pmatrix} 1 & \rho_1 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \cdots & \rho_{k-2} \\ \vdots & \vdots & \cdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \cdots & 1 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{pmatrix} = \begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{pmatrix} \tag{3.16}$$

根据 Cramer 法则，我们有

$$\phi_{kk} = \frac{D_k}{D} \tag{3.17}$$

其中，

$$D = \begin{vmatrix} 1 & \rho_1 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \cdots & \rho_{k-2} \\ \vdots & \vdots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \cdots & 1 \end{vmatrix}, \quad D_k = \begin{vmatrix} 1 & \rho_1 & \cdots & \rho_1 \\ \rho_1 & 1 & \cdots & \rho_2 \\ \vdots & \vdots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \cdots & \rho_k \end{vmatrix}$$

3.3　MA 模型

3.4　ARMA 模型

参考文献

[1] 詹姆斯.D. 汉密尔顿, 时间序列分析（上册）, 中国人民大学出版社

[2] 王黎明, 王连等, 应用时间序列分析（第二版）复旦大学出版社,2019

[3] 易丹辉, 时间序列分析方法与应用（第二版）, 中国人民大学出版社，2018

[4] 史代敏, 谢小燕, 应用时间序列分析, 高等教育出版社,2016