

## 平稳过程

宽平稳过程是指前两阶矩，即期望和自相关函数，不随时间变化的过程，即  $E[X(t)]$  为常数，且  $Cov[X(s), X(t)]$  是  $t - s$  的函数。

严平稳过程是指任意有限维分布不随时间变化的过程，即对所有的  $n$  以及所有的  $s, t_1, \dots, t_n$ ,  $(X(t_1), \dots, X(t_n))$  和  $(X(t_1 + s), \dots, X(t_n + s))$  的分布相同。由于所有的有限维分布均不随时间变化，严平稳过程的一切概率性质也不随时间改变。特别地，期望和自相关函数也不随时间变化。

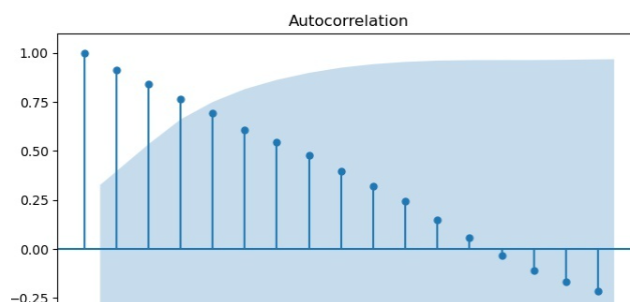
## 平稳检验

一般自然的时间序列很难有严平稳过程，所以主要使用宽平稳检验

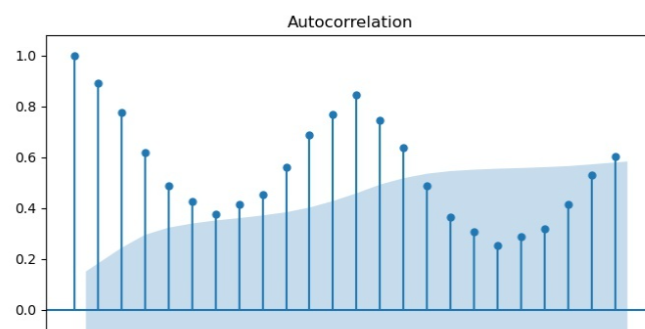
### 利用自相关性检验

平稳序列通常具有短期相关性，即随着延迟期数  $k$  的增加，平稳序列的自相关系数会很快地衰减向零，而非平稳序列的自相关系数的衰减速度会比较慢

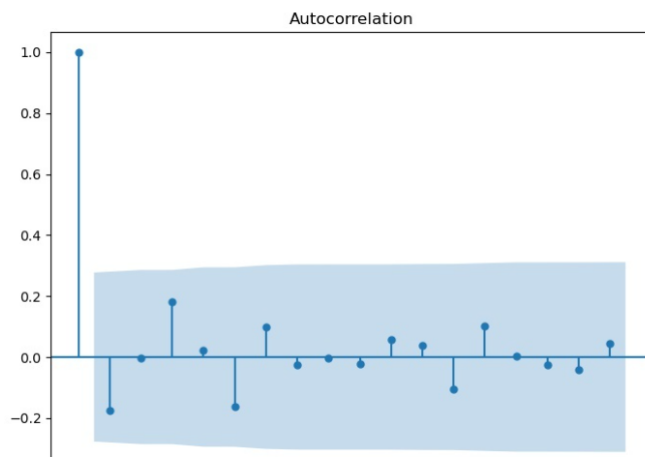
所以可采用自相关系数图进行检验，具体可利用如 statsmodels 的 acf 和 pacf 图



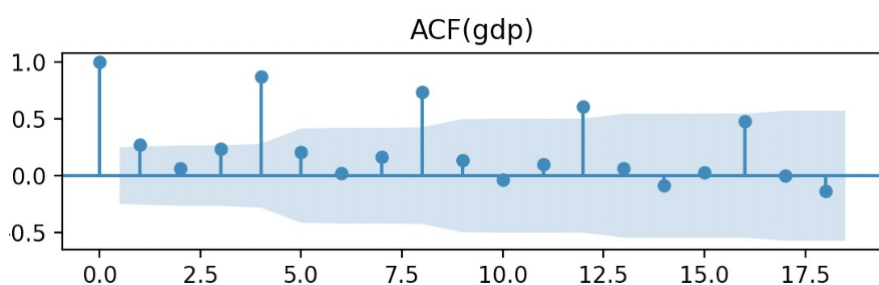
在很长的延迟期内，自相关系数一直为正，然后为负，呈现出三角对称性，这是具有单调趋势的非平稳序列的一种典型的自相关图形式



同时还呈现出明显的正弦波动规律，这是具有周期变化规律的非平稳序列的典型特征



该序列的自相关系数一直比较小，可以认为该序列一直在零轴附近波动，这是随机性较强的平稳序列通常具有的自相关图



可以看到存在一定的周期性，滞后4、8、12等自相关系数较大下降较慢，差分后自相关可能下降，差分后可能是平稳的

## 利用定义与统计量检验

根据宽平稳定义，直接将序列前后拆分成2个序列，分别计算这2个序列的均值、方差，对比看是否差异明显

<p>white noise sample:</p> <p>mean1=-0.038644,</p> <p>mean2=-0.040484variance1=1.006416,</p> <p>variance2=0.996734</p>	<p>random walk sample:</p> <p>mean1=5.506570,</p> <p>mean2=8.490356variance1=53.911003,</p> <p>variance2=126.866920</p>
------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------

白噪声序列均值和方差略有不同，但大致在同一水平线上；随机游走序列的均值和方差差异就比较大，因此为非平稳序列

## 假设检验

平稳性的假设检验方法当前主流为单位根检验，检验序列中是否存在单位根，若存在，则为非平稳序列，不存在则为平稳序列

### (A)DF检验

ADF检验（Augmented Dickey-Fuller Testing）是最常用的单位根检验方法之一

迪基（Dickey）和弗勒（Fuller）1979年基于非平稳序列的基本特征将其大致归为三类并提出DF检验：

- (1) 当序列基本走势呈现无规则上升或下降并反复时，将其归为无漂移项自回归过程；
- (2) 当序列基本走势呈现明显的随时间递增或递减且趋势并不太陡峭时，将其归为带漂移项自回归过程；
- (3) 当序列基本走势随时间快速递增时，则将其归为带趋势项回归过程。

对应检验回归式为：

(i) 无漂移项自回归过程：  $Y_t = \rho Y_{t-1} + \varepsilon_t, (t = 1, 2, \dots, n), Y_0 = 0$

(ii) 带漂移项自回归过程：  $Y_t = \mu + \rho Y_{t-1} + \varepsilon_t, (t = 1, 2, \dots, n), Y_0 = 0$

(iii) 带漂移项和趋势项自回归过程：  $Y_t = \mu + \beta t + \rho Y_{t-1} + \varepsilon_t, (t = 1, 2, \dots, n), Y_0 = 0$

其中  $\mu$  是常数项， $\beta_t$  是时间趋势项， $\varepsilon_t$  为白噪声无自相关性。

- 原假设  $H_0 : \rho = 1$  (存在单位根，时间序列是非平稳的)
- 备择假设  $H_1 : \rho < 1$  (不存在单位根，时间序列是平稳的--不含截距项和趋势项平稳/含截距项平稳/含截距项和趋势平稳)

若检验统计量大于临界值 (p值大于显著性水平  $\alpha$ )，不能拒绝原假设，序列是非平稳的；

若检验统计量小于临界值 (p值小于显著性水平  $\alpha$ )，拒绝原假设，认为序列是平稳的。

DF的检验公式为一阶自回归过程，为了能适用于高阶自回归过程的平稳性检验，迪基等1984年对DF检验进行了一定的修正，引入了更高阶的滞后项

一般使用 `arch` 或 `statsmodels` 包进行检验

`arch`包中ADF检验可指定trend为'n' (不含截距项和时间趋势项) 'c' (含截距项) 'ct' (含截距项和时间趋势项) 'ctt' (含截距项和时间趋势项和二次型时间趋势项) 分别对应不同平稳类型的检验。  
(滞后期lags默认为AIC最小)

## 其余假设检验

有其他统计学家也给出了相应的假设检验方法，可见：

Method/Model	Package/Module (function/class)
Augmented Dickey-Fuller test	statsmodels.tsa.stattools (adfuller) ; arch.unitroot (ADF)
Phillip-Perron test	arch.unitroot (PhillipsPerron)
Dickey-Fuller GLS Test	arch.unitroot (DFGLS)
KPSS test	statsmodels.tsa.stattools (kpss) ; arch.unitroot (KPSS)
Zivot-Andrew test	statsmodels.tsa.stattools (zivot_andrews) ; arch.unitroot (ZivotAndrews)
Variance Ratio test	arch.unitroot (VarianceRatio)

## 小波变换检验

小波变换就是把一个波形分解成N个低频部分和M个高频部分的和

利用全局波动指标与局部波动指标的比值来描述两者的差异

典型利用 `pywt` 进行检验的例子

```
from pywt import wavedec

def stable_periodicity_classifier(data):
    """
    return: 0 平稳数据, 1 周期数据, 2 非周期数据
    """
    boolean_is_stable = is_stable(data)
```

```

    boolen_is_periodicity = is_periodicity(data)
    if boolen_is_stable:
        return 0
    else:
        if boolen_is_periodicity[0]:
            return 1
        else:
            return 2

def is_stable(data, n_threshold=1.1, show_pic=False):
    raw_data_std = np.std(data, ddof=1)
    # 一维离散信号的小波变换
    coeffs = wavedec(data, 'db4', level=2)
    cA2, cD2, cD1 = coeffs
    # cD2标准差
    cD2_std = np.std(cD2, ddof=1)

    # 全局波动指标与局部波动指标的比值来描述两者的差异
    n = raw_data_std / cD2_std

    if n < n_threshold:
        return True
    else:
        return False

```

## 协整检验

协整性基本上可以认为是两时间序列的线性组合是平稳的，即这两者是协整的。

大多数股票价格序列不具平稳性，但是经常可以发现一对股票的市场相对价值是平稳的，如果你买入其中一支，卖空另外一支。如果存在这样的情况，那么这两个独立的时间序列被称为具有协整性。

根据上述讨论，如果我们要判断两个时间序列{x}和{y}是否具有协整关系，需要以下几个步骤：

1. 检验x和y的平稳性，若都不平稳则进行step2
2. 对x和y进行差分直到平稳，若差分阶数dx=dy，则进行step3
3. 进行协整检验（可利用 statsmodels 的 coint 函数），若通过协整检验则进行step4
4. 确定满足协整关系的线性组合系数（线性回归）

平稳检验部分可参考前文。