**指数平滑法**

# 基本思想

研究时间序列的一个重要目的是预测。现实当中事物的发展都是由连续性的，事物过去的表现与现在的状态有关，现在的状态又与将来的可能表现有一定的联系。因此可以从现有数据入手通过构造某种计算方法实现对未来的预测。移动平均法正是这样一种利用已知值的某种平均值进行预测的方法。移动平均包括简单移动平均法和加权移动平均法。

# 理论铺垫

简单移动平均法是利用一定时间跨度下数据的简单平均实现对下一期值的预测，即。可见，简单移动平均认为，时间跨度内的所有数据对未来的预测贡献全部相同。然而，众所周知，事物的当前状态与其在过去时间所有点上的表现之间联系的紧密程度并不完全一致，因此这样的预测有时可能出现很大的偏差。通常，序列数据在近期的表现要比远期的表现与现实状态的联系更加紧密。因此，预测时对过去的数据应给予不同的重视程度。

加权移动平均法是对简单移动平均法的改进，通过不同的权数体现对过去状态的不同重视程度。重视程度高、与现实联系密切的时间点对应较大的权数，而重视程度低、与现实联系松散的时间点则对应较小的权数。即





不同事物的发展规律是不同的，同一种事物随时间的推移其变化规律也会发生变化。所以，权数应随不同问题、不同时间的变化而变化。通常，权数确定没有一定之规，一般可参照几种典型的具有代表性的方法设计权数。

加权移动平均法中确定合适的权数是一件比较繁琐的事情。而指数平滑法通过对权数的改进，使其再处理时简单易行，因而在实际中应用较为广泛，可带来较为理想的短期预测精度。其基本思想也是用序列过去值的加权平均数预测未来的值，且通过权数的大小体现事物发展中不同时期间与现实联系的紧密程度。

# 指数平滑法的模型

指数平滑法因权数选择和平滑方法的不同而分成多种模型形式。虽然他们都基于上述基本思想，但在具体实现上也有所差别，也有不同的适用场合。下面介绍常用的几种模型。

## 一次指数平滑法(简单指数平滑法)

一次指数平滑法是简单移动平均法的变形，模型为



其中，是时刻的一次指数平滑值，为移动步长，整理后得: 。如果令，则。其中为一次平滑模型中的平滑常数，且显然。

由 



指数平滑法以当期的平滑值作为当期的预测值 。任何预测值都是以前所有实际值的加权平均。计算时只需知道期的实际值和期的平滑值就可以得到期的预测值。

从以上推论结果可以看出，无论平滑常数取怎样的值，其随时间的变化呈一条衰减的指数函数曲线，即随时间向过去的推移，各期实际值对预测值的影响按指数规律递减，这是该方法冠以指数平滑的原因。同时，平滑常数的选择也是很重要的，它将直接影响过去各期数据对预测值的作用。当平滑常数接近1时，各期历史数据的作用将迅速衰减，近期数据作用最大。当时间序列变化剧烈时，平滑常数可选大些;当时间序列变化平缓时，平滑常数可选小些。，所有未来值的预测等于历史数据的平均值（或“平均值”），称为平均值法。，简单地将所有预测设置为最后一次观测的值，统计中称为朴素方法。

在python中可根据模型对数据的拟合情况，根据误差大小，自动选择误差最小时的平滑常数值。

简单指数平滑法适用于平稳和无季节周期的序列。当序列中存在上升趋势时，预测值往往会偏低，而存在下降趋势时则会偏高。预测往往落后于事物发展的实际趋势。

## 二次指数平滑法(线性指数平滑法)

二次指数平滑也称双重指数平滑，是对一次指数平滑值再进行一次平滑。一次指数平滑法是直接利用平滑值作为预测值，而二次指数平滑则是利用平滑值对时间序列的线性趋势进行修正，进而建立线性平滑模型进行预测。二次指数平滑法包括布朗(Brown)线性指数平滑、霍特(Holt)指数平滑等。

### 布朗线性指数平滑

布朗线性指数平滑的一次平滑公式为



布朗线性指数平滑的二次平滑公式为



式中，为一次指数平滑值，为二次指数平滑值。

由两个平滑值计算线性平滑模型的两个参数: , ,从而得到线性指数平滑模型



该式即为布朗单一参数线性指数平滑的预测模型，也称为线性平滑模型。为预测的超前期数。为超前期的预测值。当时，由于和是平滑初始值，需事先给定。布朗线性指数平滑适用于有线性趋势的时间序列。

### 霍特指数平滑

霍特指数平滑与布朗线性指数平滑有相同的原理，也适用于有线性趋势的趋势。差别在于，霍特指数平滑不直接应用二次指数平滑，而是分别对原序列数据和序列的趋势进行平滑。模型的一般形式为



其中：



为数据的平滑值。为趋势的平滑值，取决于相邻两个平滑值之差、上期趋势平滑值以及(模型初始参数)。为预测的超前期数。当时，由于和是平滑初始值，也需事先给定。霍特指数平滑模型也适用于有趋势的时间序列。

对于长期预测，使用Holt方法的预测在未来会无限期地增加或减少。 在这种情况下，我们使用具有阻尼参数的阻尼趋势方法来防止预测“失控”。





## 三次指数平滑法

三次指数平滑也称三重指数平滑，与二次指数平滑类似，也不直接将平滑值作为预测值，而是服务于模型建立。三次指数平滑包括布朗三次指数平滑、温特(Winter)线性和季节性指数平滑。

### 布朗三次指数平滑

布朗三次指数平滑是对二次指数平滑值再进行一次平滑，并用于估计二次多项式参数。其一般模型为：



为预测的超前期数。由上式可知，布朗三次指数模型并非一个关于的线性模型，而是类似于二次多项式的曲线模型，可表现时间序列的曲线变化趋势。其中



，，依次为一次、二次、三次指数平滑值。各次平滑形势分别为



布朗三次指数平滑模型适用于有非线性趋势存在的序列。

### 温特线性和季节性指数平滑

温特线性和季节性指数平滑模型的一般形式为



上式中包含三种成分，他们分别是水平性()、趋势性()和季节性()。为季节周期长度，为季节调整因子，分别为模型的三个参数。其中



公式三中，由于是序列的平滑值，即平均水平，不包含季节性。因而为序列第期的实际值与同期平滑值的比，体现了序列第期的季节特征，也称为季节函数。公式三采用以参数为权重计算加权季节指数，同时用加权上一期的季节调整因子，得到第期的季节调整因子。公式一中的目的就是消除序列观察值中的季节性波动。

公式二通过对序列水平值得调整估计序列的趋势性。用参数加权水平增量，再用加权上一期的趋势性，得到第期的趋势性。

温特线性和季节性指数平滑模型适用于同时具有趋势性和依赖于水平的季节性的时间序列，且只适用于短期预测。

# 参数优化与模型选择方法

|  |  |  |
| --- | --- | --- |
|  | Trend | Seasonality |
| SES | × | × |
| Halt’s Method | √ | × |
| Halt-Winters’s Method | √ | √ |

参数优化的方法是最小化误差平方和或最大化似然函数。

似然函数是一种关于统计模型参数的函数。给定输出时，关于参数的似然函数(在数值上)等于给定参数后变量的概率：



拟合程度的好坏通常采用SSE(Summary of the Squared Errors)或MSE(Mean of the Squared Errors)进行度量。SSE是平滑值与实际值之间误差平方的和，MSE是SSE的平均值。即,是平滑值的个数。

表示平滑值，表示实际值，的计算公式如下：



模型选择可以根据信息量准则，常用的有 AIC 和 BIC等。

AIC 即 Akaike information criterion， 定义为



其中 是似然函数， 是参数数量。用 AIC 选择模型时要求似然函数大，同时对参数数量作了惩罚，在似然函数相近的情况下选择复杂度低的模型。BIC 即 Bayesian information criterion，定义为



其中是样本数量。当 时， ，因此当样本量较大时 BIC 对模型复杂度的惩罚比 AIC 更严厉。