

指数平滑算法

目录

SPSSAU 操作	2
SPSSAU 数据格式	2
算法	3
一次指数平滑	3
1. 初始化平滑值	3
2. 设定平滑常数	3
3. 计算平滑值	3
4. 计算预测值	3
5. 计算预测误差	3
6. 评估预测精度	3
二次指数平滑法	4
1. 初始化平滑值	4
2. 计算平滑公式	4
3. 计算趋势	4
4. 计算预测误差	4
5. 评估预测精度	4
三次指数平滑	5
1. 初始化平滑值和季节指数	5
2. 计算平滑公式	5
3. 计算趋势	5
4. 计算预测误差	6
5. 评估预测精度	6
参考文献	6

指数平滑法是一种常用的时间序列预测方法，主要用于平滑和预测数据。它通过对历史数据赋予不同的权重来计算未来值，较近的数据权重更高，较远的数据权重较低。SPSSAU 提供三种指数平滑法，分别是一次指数平滑法，二次指数平滑法和三次指数平滑法。其位于 SPSSAU-» 综合评价-» 指数平滑法。

SPSSAU 操作



该截图展示了 SPSSAU 平台的分析操作界面。左侧有一个蓝色的“开始分析”按钮。右侧上方有四个下拉菜单，分别用于设置“初始值S0”、“平滑类型”、“alpha值”和“向后预测期数”。下方是一个大的浅灰色框，用于放置数据，底部有一个“标签 [可选]”的输入框。

将时间序列数据（仅 1 项）拖拽至右侧框然后‘开始分析’即可。SPSSAU 中涉及 4 项参数，分别是初值 S0，alpha 值，平滑类型和向后预测期数。

- ✓ 初始值 S0：默认为‘自动’即算法自动寻找最优初始值（标准是均方根误差值 RMSE），可选项包括第 1 期，前 2 期平均值，前 3 期平均值，前 4 期平均值或前 5 期平均值。
- ✓ alpha 值：默认为‘自动’即算法自动寻找最优初始值（标准是均方根误差值 RMSE），可选项分别是 0.05，0.1，0.2 直到 0.9，0.95 共 11 项。
- ✓ 平滑类型：默认为‘自动’即算法自动寻找最优初始值（标准是均方根误差值 RMSE），可选项分别是一次平滑、二次平滑和三次平滑。
- ✓ 向后预测期数：默认提供向后 12 期的预测值，可自行设置向后预测期数。

SPSSAU 数据格式

	A	B	C	I
1	年份	城市交通噪声	...	
2	1986	71.10		
3	1987	72.40		
4	1988	72.40		
5	1989	72.10		
6	1990	71.40		
7	1991	72.00		
8	1992	71.60		
9		

指数平滑法通常针对数量较少的样本进行预测，如果数据带有时间项，其并不需要纳入分析项中（可将其放入‘标签’项中），但整理数据时一般需要将数据依次按时间排序好录入数据。

算法

一次指数平滑

一次指数平滑法 (Simple Exponential Smoothing) 是一种用于时间序列预测的基本方法，以下是一次指数平滑法的计算步骤：

1. 初始化平滑值

选择初始平滑值 S_1 ，通常可以取第一个实际观测值 x_1 ： $S_1 = x_1$ ，此处可由 SPSSAU 自动选择，也可下拉参数进行选择。

2. 设定平滑常数

选择平滑常数 α ，其值在 0 和 1 之间。 α 的选择会影响平滑的灵敏度，较大的 α 值会使模型对最近数据的反应更敏感。

3. 计算平滑值

对于每个时间点 t ($t > 1$)，使用以下公式计算平滑值 S_t ： $S_t = \alpha x_t + (1 - \alpha)S_{t-1}$ ，其中：

x_t 为第 t 个时间点的实际值

S_{t-1} 为前一个时间点的平滑值

4. 计算预测值

对于第 $t + 1$ 个时间点的预测值 \hat{x}_{t+1} ，使用最新的平滑值：

$$\hat{x}_{t+1} = S_t$$

5. 计算预测误差

预测误差 e_t 定义为实际值与预测值的差：

$$e_t = x_t - \hat{x}_t$$

以及绝对误差 (AE)：

$$AE = |x_t - \hat{x}_t|$$

6. 评估预测精度

可以使用均方根误差 (RMSE)、均方误差 MSE、平均绝对误差 (MAE) 或平均绝对百分比误差 MAPE 等指标来评估预测的精度：

$$\begin{aligned} \text{RMSE} &= \sqrt{\frac{1}{n} \sum_{t=1}^n e_t^2} \\ \text{MSE} &= \frac{1}{n} \sum_{t=1}^n e_t^2 \\ \text{MAE} &= \frac{1}{n} \sum_{t=1}^n |e_t| \end{aligned}$$

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^n \left| \frac{e_t}{x_t} \right|$$

其中 n 为预测的时间点数, e_t 表示误差。

二次指数平滑法

二次指数平滑法 (Double Exponential Smoothing) 在一次平滑法基础上再一次平滑, 其适用于具有一定线性趋势的数据预测, 比如人均 GDP 数据, 其计算说明如下:

1. 初始化平滑值

初始平滑值 S_1 , 通常取第一个实际观测值 x_1 : $S_1 = x_1$, 此处可由 SPSSAU 自动选择, 也可下拉参数进行选择。

2. 计算平滑公式

$$\begin{cases} S_t^{(1)} = \alpha x_t + (1 - \alpha) S_{t-1}^{(1)} \\ S_t^{(2)} = \alpha S_t^{(1)} + (1 - \alpha) S_{t-1}^{(2)} \end{cases}$$

$S_t^{(1)}$ 表示一次平滑法, $S_t^{(2)}$ 表示二次平滑法。上式中 x_t 表示第 t 期真实值, α 值为误差值的校正值。

3. 计算趋势

二次平滑法时, 其应用于线性趋势数据, 其实质上是寻找出一个线性回归趋势线, 数学上如下述公式: $\hat{y}_{t+T} = a_t + b_t \times T$, T 表示时间序列往后第 T 期。 a_t 和 b_t 为线性回归趋势线的参数值, 其计算公式分别如下:

$$\begin{cases} a_t = 2S_t^{(1)} - S_t^{(2)} \\ b_t = \frac{\alpha}{1 - \alpha} (S_t^{(1)} - S_t^{(2)}) \end{cases}$$

时间序列时期外的预测值 (也可称为预测值), 其公式如下:

$$\hat{y}_{t+T} = a_t + b_t \times T, T = 1, 2, \dots$$

时间序列时期内的预测值 (也可称为拟合值), 其公式如下:

$$\hat{y}_{t+1} = a_t + b_t$$

4. 计算预测误差

预测误差 e_t 定义为实际值与预测值的差:

$$e_t = x_t - \hat{x}_t$$

以及绝对误差 (AE):

$$\text{AE} = |x_t - \hat{x}_t|$$

5. 评估预测精度

可以使用均方根误差 (RMSE)、均方误差 MSE、平均绝对误差 (MAE) 或平均绝对百分比误差 MAPE 等指标来评估预测的精度:

$$\begin{aligned} \text{RMSE} &= \sqrt{\frac{1}{n} \sum_{t=1}^n e_t^2} \\ \text{MSE} &= \frac{1}{n} \sum_{t=1}^n e_t^2 \\ \text{MAE} &= \frac{1}{n} \sum_{t=1}^n |e_t| \\ \text{MAPE} &= \frac{1}{n} \sum_{t=1}^n \left| \frac{e_t}{x_t} \right| \end{aligned}$$

其中 n 为预测的时间点数, e_t 表示误差。

三次指数平滑

三次指数平滑法 (Triple Exponential Smoothing) 三次平滑法是在二次平滑法基础上再一次平滑, 其适用于具有曲线增长趋势的时间数据预测, 比如人均 GDP 数据, 其计算说明如下:

1. 初始化平滑值和季节指数

初始平滑值 S_1 , 通常取第一个实际观测值 x_1 : $S_1 = x_1$, 此处可由 SPSSAU 自动选择, 也可下拉参数进行选择。

2. 计算平滑公式

$$\begin{cases} S_t^{(1)} = \alpha y_t + (1 - \alpha) S_{t-1}^{(1)} \\ S_t^{(2)} = \alpha S_t^{(1)} + (1 - \alpha) S_{t-1}^{(2)} \\ S_t^{(3)} = \alpha S_t^{(2)} + (1 - \alpha) S_{t-1}^{(3)} \end{cases}$$

类似二次平滑法的计算, 上式中, $S_t^{(1)}$ 、 $S_t^{(2)}$ 和 $S_t^{(3)}$ 分别表示一次平滑、二次平滑和三次平滑计算数据, y_t 表示第 t 期真实值, α 值为误差值的校正值。

3. 计算趋势

三次平滑法时, 其应用于曲线趋势数据, 其实质上是寻找出二次曲线趋势线, 数学上如下述公式:

$$\hat{y}_{t+T} = a_t + b_t \times T + c_t \times T^2$$

T 表示时间序列往后第 T 期, a_t 、 b_t 和 c_t 分别为曲线回归趋势线的参数值, 其计算公式分别如下:

$$\begin{cases} a_t = 3S_t^{(1)} - 3S_t^{(2)} + S_t^{(3)} \\ b_t = \frac{\alpha}{2(1-\alpha^2)} [(6-5\alpha)S_t^{(1)} - 2(5-4\alpha)S_t^{(2)} + (4-3\alpha)S_t^{(3)}] \\ c_t = \frac{\alpha^2}{(1-\alpha)^2} [S_t^{(1)} - 2S_t^{(2)} + S_t^{(3)}] \end{cases}$$

时间序列时期外的预测值（也可称为预测值），其公式如下：

$$\hat{y}_{t+T} = a_t + b_t \times T + c_t \times T^2, T = 1, 2, \dots$$

时间序列时期内的预测值（也可称为拟合值），其计算公式如下：

$$\hat{y}_{t+1} = a_t + b_t + c_t$$

4. 计算预测误差

预测误差 e_t 定义为实际值与预测值的差：

$$e_t = x_t - \hat{x}_t$$

以及绝对误差（AE）：

$$AE = |x_t - \hat{x}_t|$$

5. 评估预测精度

可以使用均方根误差（RMSE）、均方误差 MSE、平均绝对误差（MAE）或平均绝对百分比误差 MAPE 等指标来评估预测的精度：

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n e_t^2}$$

$$MSE = \frac{1}{n} \sum_{t=1}^n e_t^2$$

$$MAE = \frac{1}{n} \sum_{t=1}^n |e_t|$$

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{e_t}{x_t} \right|$$

其中 n 为预测的时间点数， e_t 表示误差。

参考文献

【1】 The SPSSAU project (2024). SPSSAU. (Version 24.0) [Online Application Software].

Retrieved from <https://www.spssau.com>.

【2】 周俊,马世澎. SPSSAU 科研数据分析方法与应用.第 1 版[M]. 电子工业出版社,2024.