模型-预测主题-连续型预测时间序列模型-移动平均法【hxy】
1. 模型名称
2. 模型评价
2.1 模型优点
2.2 模型局限
3. 基本算法
3.1 简单移动平均法
3.2 加权移动平均法
3.3 趋势移动平均法
4. 实例
4.1 简单移动平均法
4.1.1 问题描述
4.1.2 数学解法
4.1.3 代码实现
4.2 加权移动平均法
4.2.1 问题描述
4.2.2 数学解法
4.2.3 代码实现
4.3 趋势移动平均法
4.3.1 问题描述
4.3.2 数学解法
5. 参考资料

模型-预测主题-连续型预测时间序列模型-移动平均法【hxy】

1. 模型名称

移动平均法（Moving Average）

2. 模型评价

2.1 模型优点

- 简单
- **趋势移动平均法**对于同时存在直线趋势与周期波动的序列比较有效，是一种既能反映趋势变化，又能有效分离出周期变动的方法

2.2 模型局限

- 不适用于实际数据波动较大的序列，一般较少采用此法进行预测

3. 基本算法

3.1 简单移动平均法

设观测序列为y_1，…，y_T，取移动平均的项数$N < T$。一次移动平均值计算公式为

$$M_t^{(1)}(N) = \frac{1}{N}(y_t + y_{t-1} + \cdots + y_{t-N+1}) = \frac{1}{N}\sum_{i=0}^{N-1} y_{t-i}$$

其递推公式为

$$M_t = M_{t-1} + \frac{y_t - y_{t-N}}{N}$$

由于移动平均可以平滑数据，消除周期变动和不规则变动的影响使长期趋势显示出来，因而可以用于预测：

$$\hat{y}_{t+1} = M_t$$

即以第t期移动平均数作为第$t+1$期的预测值。

3.2 加权移动平均法

设观测序列为y_1，…，y_T，取移动平均的项数$N < T$。一次移动平均值计算公式为

$$M_{tw} = \frac{w_1 y_t + w_2 y_{t-1} + \cdots + w_n y_{t-n+1}}{w_1 + w_2 + \cdots + w_n}, t \geq N$$

利用加权移动平均数来作预测：

$$\hat{y}_{t+1} = M_{tw}$$

即以第t期加权移动平均数作为第t+1期的预测值。

3.3 趋势移动平均法

一次移动平均值计算公式为

$$M_t^{(1)}(N) = \frac{1}{N}(y_t + y_{t-1} + \cdots + y_{t-N+1}) = \frac{1}{N}\sum_{i=0}^{N-1} y_{t-i}$$

二次移动平均计算公式为

$$M_t^{(2)} = \frac{M_t^{(1)} + M_{t-1}^{(1)} + \cdots + M_{t-N+1}^{(1)}}{N}$$

推出公式

$$M_t^{(2)} = M_{t-1}^{(2)} + \frac{M_t^{(1)} - M_{t-n}^{(1)}}{N}$$

根据移动平均来确定$\hat{y}_{t+T} = a_t + b_t T$中$a_t$和$b_t$的值，可得$a_t$和$b_t$的计算公式

$$a_t = 2M_t^{(1)} - M_t^{(2)}$$
$$b_t = \frac{2}{N-1}(M_t^{(1)} - M_t^{(2)})$$

4. 实例

4.1 简单移动平均法

4.1.1 问题描述

汽车配件某年1~12年月份的化油器销售量（单位：只）统计数据见下表中第2行，试用一次移动平均法预测下一年1月份的销售量。

月份	1	2	3	4	5	6	7	8	9	10	11	12	预测
实际	423	358	434	445	527	429	502	480	384	427	446		

4.1.2 数学解法

分别取$N=3, N=5$，按预测公式

$$\hat{y}_{t+1}(3)=M_t^1(3)=\frac{y_t+y_{t-1}+y_{t-2}}{3}, t=3,4,\dots,12$$
$$\hat{y}_{t+1}(5)=M_t^1(5)=\frac{y_t+y_{t-1}+y_{t-2}+y_{t-3}+y_{t-4}}{5}, t=5,6,\dots,12$$

计算3个月和5个月移动平均预测值，见下表。$N=3$时，预测的标准误差为56.5752；$N=5$时，预测的标准误差为39.8159。通过预测后，可以看到，实际数据波动较大，经移动平均后，随机波动明显减少，且N越大，波动也越小。同时，也可以看到，一次移动平均法的预测标准误差还是有些大，对于实际数据波动较大的序列，一般较少采用此法进行预测。

月份	1	2	3	4	5	6	7	8	9	10	11	12	预测
N=3				405	412	469	467	461	452	469	455	430	419
N=5						437	439	452	466	473	444	444	448

4.1.3 代码实现

[moving_average.py]

```
import numpy as np
y=np.array([423,358,434,445,527,429,426,502,480,384,427,446])
def MoveAverage(y,N):
    Mt=['*']*N
    for i in range(N+1,len(y)+2):
        M=y[i-(N+1):i-1].mean()
        Mt.append(M)
    return Mt
yt3=MoveAverage(y,3)
s3=np.sqrt(((y[3:]-yt3[3:-1])**2).mean())
yt5=MoveAverage(y,5)
s5=np.sqrt(((y[5:]-yt5[5:-1])**2).mean())
print('N=3时,预测值：',yt3,'，预测的标准误差：',s3)
print('N=5时,预测值：',yt5,'，预测的标准误差：',s5)
```

结果：

```
>>>
= RESTART: /Users/xinyuanhe/Desktop/预测主题-连续型预测时间序列模型-移动平均法【hxy】/m
oving_average.py
N=3时,预测值:  ['*', '*', '*', 405.0, 412.3333333333333, 468.6666666666667, 467.0
, 460.6666666666667, 452.3333333333333, 469.3333333333333, 455.3333333333333, 43
0.3333333333333, 419.0] , 预测的标准误差:  56.57519850976887
N=5时,预测值:  ['*', '*', '*', '*', '*', 437.4, 438.6, 452.2, 465.8, 472.8, 444.2
, 443.8, 447.8] , 预测的标准误差:  39.815861878689226
```

4.2 加权移动平均法

4.2.1 问题描述

我国1979-1988年原煤产量如图表，试用加权移动平均法预测1989年的产量：

年份	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988
产量	6.35	6.20	6.22	6.66	7.15	7.89	8.72	8.94	9.28	9.80

4.2.2 数学解法

取$w_1=3, w_2=2, w_1=1$，预测公式：

$$\hat{y}_{t+1}=\frac{3y_t+2y_{t-1}+y_{t-2}}{3+2+1}$$

三年加权移动平均预测值,其结果列于表中：

年份	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988
产量	6.35	6.20	6.22	6.66	7.15	7.89	8.72	8.94	9.28	9.80
预测值	-	-	-	6.24	6.44	6.83	7.44	8.18	8.69	9.07
相对误差（%）	-	-	-	6.31	9.93	13.43	14.68	8.50	6.36	7.45

1989年产量预测值为

$$\hat{y}_{1989}=\frac{3\times 9.80+2\times 9.28+1\times 8.94}{6}=9.48$$

这个预测值偏低，可以修正，其方法是，计算总的平均误差

$$(1-\frac{\sum \hat{y}_t}{\sum y_t})\times 100\%=(1-\frac{52.89}{58.44})\times 100\%=9.50\%$$

由于总预测的平均值比实际值低9.50%，所以将1989年的预测值修正为

$$\frac{9.48}{1-9.5\%}=10.48$$

在加权移动平均法中，w_t的选择一般原则：近期数据的权数越大。

4.2.3 代码实现

weighting_moving_average.py

```
import numpy as np
y=np.array([6.35,6.20,6.22,6.66,7.15,7.89,8.72,8.94,9.28,9.80])
def WeightMoveAverage(y,N):
    Mt=['*']*N
    for i in range(N,len(y)+1):
        M=0
        Sum=0
        for j in range(N,0,-1):
            M+=j*y[i-N+j-1]
            Sum+=j
        Mt.append(M/Sum)
    return Mt
yt3=WeightMoveAverage(y,3)
s3=np.sqrt(((y[3:]-yt3[3:-1])**2).mean())
print('N=3时,预测值：',yt3,'，预测的标准误差：',s3)
```

结果：

```
>>>
= RESTART: /Users/xinyuanhe/Desktop/预测主题-连续型预测时间序列模型-移动平均法【hxy】/weighting_moving_average.py
N=3时,预测值： ['*', '*', '*', 6.235, 6.436666666666667, 6.831666666666667, 7.438333333333333, 8.181666666666667, 8.691666666666668, 9.073333333333332, 9.483333333333333]，预测的标准误差： 0.8367845202519782
```

4.3 趋势移动平均法

4.3.1 问题描述

我国1965-1985年的发电总量如表。试预测1986年和1987年的发电总量。

1973	74	75	76	77	78	79	80	81	82	83	84	85
1668	1688	1958	2031	2234	2566	2820	3006	3093	3277	3514	3770	4107

4.3.2 数学解法

由散点图可以看出，发电量基本呈直线上升趋势，可用趋势移动平均放来预测。

一次移动平均结果为：

1979	1980	1981	1982	1983	1984	1985
2216.2	2435.8	2625.0	2832.7	3046.0	3246.7	3461.2

取$N=6$，分别计算

$$M_t^{(1)} = \frac{4107+3770+3514+3277+3093+3066}{6} = 3461.2$$
$$M_t^{(2)} = \frac{3461.2+3246.7+3046.0+2832.7+2625.0+2435.8}{6} = 2941.2$$

再由公式得

$$a_t = 2M_t^{(1)} - M_t^{(2)} = 2 \times 3461.2 - 2941.2 = 3981.2$$
$$b_t = \frac{2}{N-1}(M_t^{(1)} - M_t^{(2)}) = \frac{2}{5}(3461.2 - 2941.2) = 208$$

于是得$t = 1985$时，直线预测模型为

$$\hat{y}_{t+T} = 3981.2 + 208T$$

预测1986年和1987年发电总量为

$$\hat{y}_{1986} = 3981.2 + 208 = 4189.2$$
$$\hat{y}_{1987} = 3981.2 + 208 \times 2 = 4397.2$$

5. 参考资料

1. 时间序列移动平均法-百度文库
2. 移动平均法-参考代码