

## demo\code\6-1\_Lagrange\_interpolation.py

```
1 #-*- coding: utf-8 -*-
2 #拉格朗日插值代码
3 import pandas as pd #导入数据分析库Pandas
4 from scipy.interpolate import lagrange #导入拉格朗日插值函数
5
6 inputfile = './data/missing_data.xls' #输入数据路径,需要使用Excel格式;
7 outputfile = './tmp/missing_data_processed.xls' #输出数据路径,需要使用Excel格式
8
9 data = pd.read_excel(inputfile, header=None) #读入数据
10
11 #自定义列向量插值函数
12 #s为列向量, n为被插值的位置, k为取前后的数据个数, 默认为5
13 def ployinterp_column(s, n, k=5):
14     y = s[list(range(n-k, n)) + list(range(n+1, n+1+k))] #前后取数
15     y = y[y.notnull()] #剔除空值
16     return lagrange(y.index, list(y))(n) #插值并返回插值结果
17
18 #逐个元素判断是否需要插值
19 for i in data.columns:
20     for j in range(len(data)):
21         if (data[i].isnull())[j]: #如果为空即插值。
22             data[i][j] = ployinterp_column(data[i], j)
23
24 data.to_excel(outputfile, header=None, index=False) #输出结果
```