

数据科学 4 Scipy

SciPy是基于NumPy开发的高级模块，它提供了许多数学算法和函数的实现，用于解决科学计算中的一些标准问题。例如数值积分和微分方程求解，扩展的矩阵计算，最优化，概率分布和统计函数，甚至包括信号处理等。

作为标准科学计算程序库，SciPy类似于Matlab的工具箱，它是Python科学计算程序的核心包，它用于有效地计算NumPy矩阵，与NumPy矩阵协同工作。

SciPy库由一些特定功能的子模块构成，如下表所示：

模块	功能
cluster	矢量化 / K-均值
constants	物理和数学常数
fftpack	傅里叶变换
integrate	积分程序
interpolate	插值
io	数据输入输出
linalg	线性代数程序
ndimage	n维图像包
odr	正交距离回归
optimize	优化
signal	信号处理
sparse	稀疏矩阵
spatial	空间数据结构和算法
special	任何特殊数学函数
stats	统计

以上子模块全依赖于NumPy且相互独立，导入NumPy和这些SciPy模块的标准方式如下，示例代码：

```
import numpy as np
import scipy
from scipy import stats
```

验证下scipy是否安装成功

```
face = scipy.misc.face()
```

```
face.shape
face.dtype

import matplotlib.pyplot as plt
plt.gray()
plt.imshow(face)
plt.show()
```

生成矩阵和矩阵计算

```
a = scipy.mat([1, 2, 3])
b = scipy.mat([4, 5, 6])
c = a + b
d = a * 4
e = a - b
```

统计均值、方差、中位数、最大值和最小值

```
# 均值
scipy.mean(num_friends)

# 方差
scipy.var(num_friends)

# 中位数
scipy.median(num_friends)

# 最大值
np.max(num_friends)

# 最小值
np.min(num_friends)
```