

Basic Python - Einsum in Numpy

Hoàng-Nguyên Vũ

1 Correlation (Tương quan)

Correlation là một phép đo thống kê cho biết mức độ và hướng của mối quan hệ giữa hai biến số. Khi hai biến số có mối quan hệ tương quan, giá trị của một biến có thể ảnh hưởng đến giá trị của biến còn lại.

- **Positive Correlation (Tương quan dương):** Khi giá trị của một biến tăng thì giá trị của biến kia cũng tăng.
- **Negative Correlation (Tương quan âm):** Khi giá trị của một biến tăng thì giá trị của biến kia giảm.
- **No Correlation (Không tương quan):** Không có mối quan hệ rõ ràng giữa hai biến.

2 Correlation Coefficient (Hệ số tương quan)

Hệ số tương quan là một giá trị số từ -1 đến 1 cho biết mức độ của mối quan hệ tương quan:

- **1:** Mối quan hệ tương quan hoàn hảo dương.
- **0:** Không có mối quan hệ tương quan.
- **-1:** Mối quan hệ tương quan hoàn hảo âm.

Công thức tính hệ số tương quan Pearson (Pearson Correlation Coefficient)

$$r = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum (X - \bar{X})^2 \sum (Y - \bar{Y})^2}}$$

Trong đó:

- X và Y là hai biến số.
- \bar{X} và \bar{Y} là giá trị trung bình của X và Y .

3 Bài tập: Tính hệ số tương quan giữa hai biến số

1. Tạo hai mảng dữ liệu (array) với numpy.
2. Tính toán hệ số tương quan giữa hai mảng dữ liệu này.

Hướng dẫn:

```
1 import numpy as np
2
3 # Tạo hai mảng dữ liệu
4 X = np.array([1, 2, 3, 4, 5])
5 Y = np.array([2, 4, 6, 8, 10])
6
7 # Tính toán hệ số tương quan
8 ## Your code Here ##
9
10 print("Hệ số tương quan giữa X và Y là:", ## Your result variable here ##)
```

3.1 Bài tập mở rộng

Hãy tạo hai mảng dữ liệu mới và tính hệ số tương quan giữa chúng.

Ví dụ:

```
1 import numpy as np
2
3 # Tạo hai mảng dữ liệu mới
4 X_new = np.array([10, 20, 30, 40, 50])
5 Y_new = np.array([5, 15, 25, 35, 45])
6
7 # Tính toán hệ số tương quan
8 ## Your code Here ##
9
10
11 print("Hệ số tương quan giữa X_new và Y_new là:", ## Your code Here ##)
```

Hãy thử nghiệm với các mảng dữ liệu khác nhau và quan sát kết quả để hiểu rõ hơn về mối quan hệ tương quan giữa các biến số!