



Uniwersytet
Bielsko-Bialski

Metoda PCA na podstawie SVD

Zadanie PCA

Sprawozdanie z ćwiczeń
Matematyka Konkretna

Data wykonania:
28.06.2025

Autor:
Bartosz Bieniek 058085

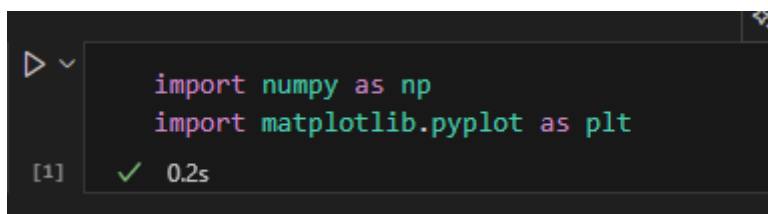
1. Cel ćwiczenia

Zadanie dotyczy obliczenia Środka, osi głównych oraz kątu obrotu danych dwuwymiarowych z pliku .csv zgodnie z wariantem zadania [Github](#).

2. Przebieg ćwiczenia

1. Import bibliotek

Zaimportowano biblioteki numpy i matplotlib.pyplot, które umożliwiły przeprowadzenie analizy danych oraz ich wizualizację. numpy posłużyło do obliczeń numerycznych i algebry liniowej, natomiast matplotlib do wykresów.



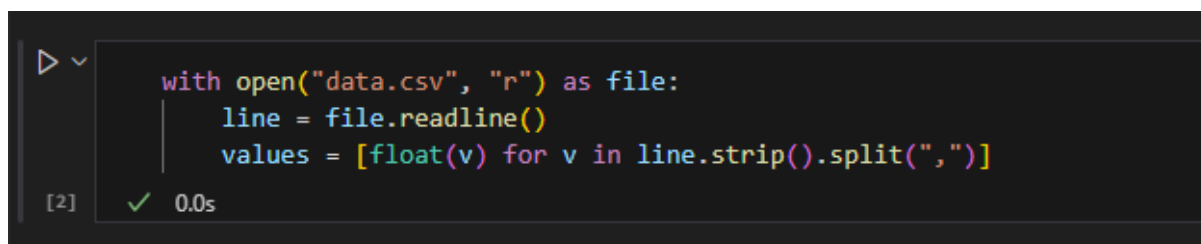
```
import numpy as np
import matplotlib.pyplot as plt
```

[1] ✓ 0.2s

Rys. 1. Import bibliotek

2. Wczytanie danych z pliku CSV

Dane wczytano z pliku data.csv, który zawierał wartości współrzędnych zapisane w jednym wierszu w formacie: $x_1, y_1, x_2, y_2, \dots$. Wszystkie wartości przekonwertowano do typu float, co umożliwiło ich dalsze wykorzystanie w analizie.



```
with open("data.csv", "r") as file:
    line = file.readline()
    values = [float(v) for v in line.strip().split(",")]
```

[2] ✓ 0.0s

Rys. 2. Wczytanie danych z pliku CSV

3. Przekształcenie do formy (n, 2)

Z wektora liczb utworzono macierz o dwóch kolumnach, gdzie każda para wartości reprezentowała współrzędne punktu w przestrzeni dwuwymiarowej. Dzięki temu dane mogły być analizowane geometrycznie.

```
[3] data = np.array(values).reshape(-1, 2)
✓ 0.0s
```

Rys. 3. Przekształcenie do formy (n, 2)

4. Obliczenie środka ciężkości

Obliczono środek ciężkości zbioru punktów jako średnią arytmetyczną współrzędnych x i y. Punkt ten stanowił nowy środek układu współrzędnych dla analizy PCA.

```
[4] center = np.mean(data, axis=0)
✓ 0.0s
```

Rys. 4. Obliczenie środka ciężkości

5. Centrowanie danych

Od każdego punktu odjęto współrzędne środka ciężkości, aby uzyskać dane względem środka układu. Operacja ta była niezbędna do prawidłowego zastosowania metody PCA opartej na dekompozycji SVD.

```
[5] centered = data - center
✓ 0.0s
```

Rys. 5. Centrowanie danych

6. SVD danych scentrowanych

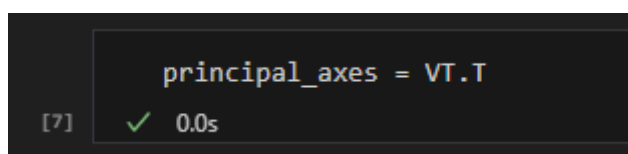
Na zcentrowanej macierzy danych wykonano dekompozycję SVD, uzyskując macierze UU, SS, VT. Wektory własne zawarte w VT odpowiadały kierunkom głównych osi PCA, a wartości SS – rozrzutowi danych w tych kierunkach.

```
[6] U, S, VT = np.linalg.svd(centered)
✓ 0.1s
```

Rys. 6. SVD danych scentrowanych

7. Główne osie (wektory PCA)

Z macierzy VTVT wyodrębniono kolumny odpowiadające wektorom własnym ukazującym główne kierunki zmienności danych. Kierunki te interpretowano jako osie główne PCA.



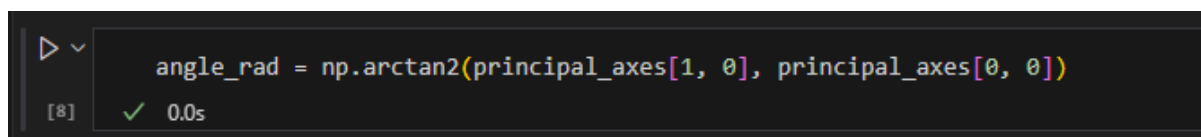
```
principal_axes = VT.T
```

[7] ✓ 0.0s

Rys. 7. Główne osie (wektory PCA)

8. Obliczenie kąta obrotu

Na podstawie pierwszego wektora własnego obliczono kąt jego nachylenia względem osi x, wykorzystując funkcję `arctan2`. Wynik wyrażony został w radianach i zinterpretowano go jako obrót układu względem głównej osi rozrzutu danych.



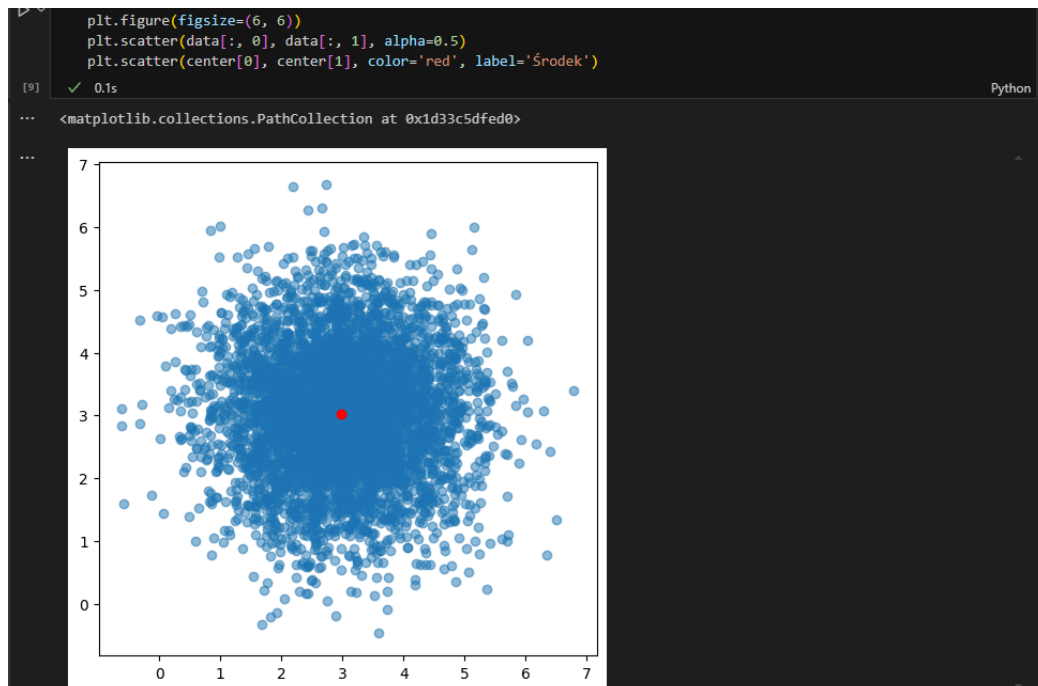
```
angle_rad = np.arctan2(principal_axes[1, 0], principal_axes[0, 0])
```

[8] ✓ 0.0s

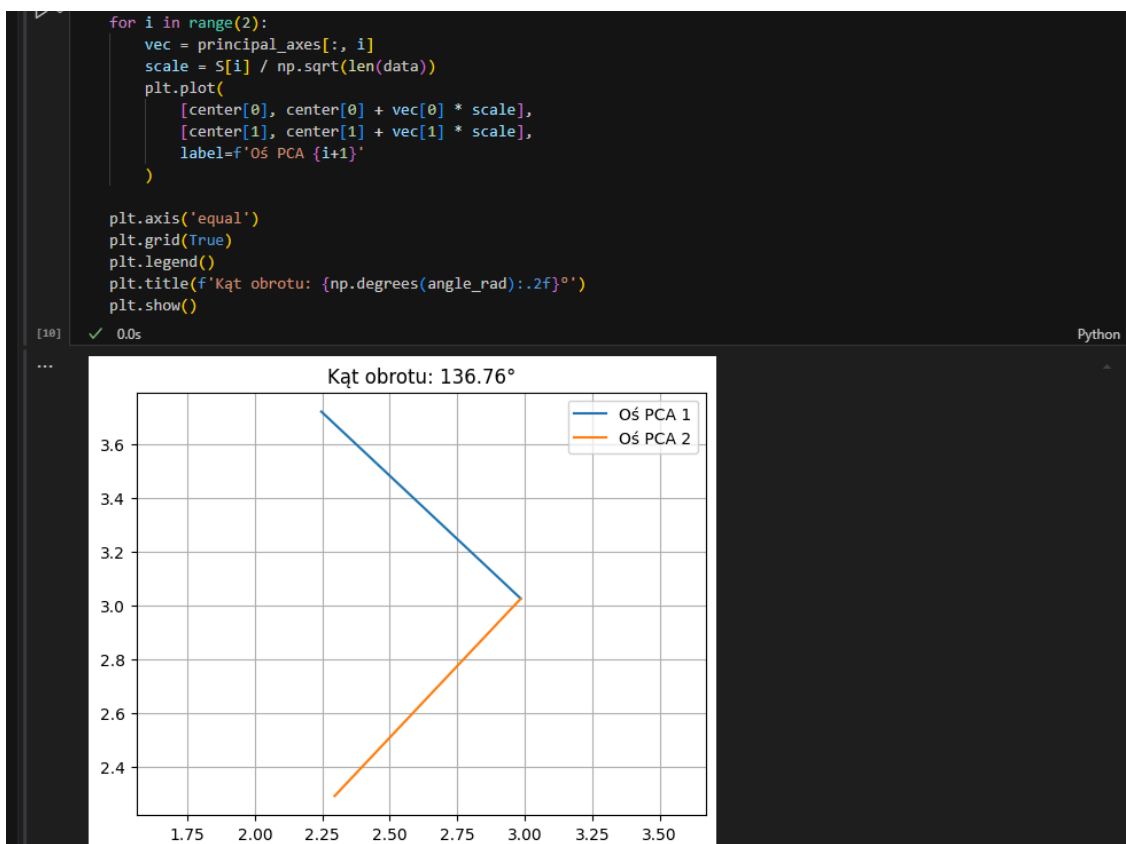
Rys. 8. Obliczenie kąta obrotu

9. Wizualizacja

Na wykresie przedstawiono dane Źródłowe, Środek ciężkości oraz osie główne PCA. Osie te pokazano w postaci wektorów wychodzących ze Środka, skalowanych zgodnie z rozrzutem danych w danym kierunku. Wizualizacja umożliwiła intuicyjne zrozumienie geometrii PCA.



Rys. 9. Wizualizacja



Rys. 10. Wizualizacja

3. Wnioski

Na podstawie analizy PCA wyznaczono główne kierunki rozrzutu danych, które odpowiadają kierunkom największej i najmniejszej wariancji w zbiorze punktów.

Obliczony kąt obrotu głównej osi PCA względem osi układu współrzędnych potwierdził, że dane są silnie zorientowane wzdłuż określonego kierunku.