SPRAWOZDANIE

Zajęcia: Matematyka Konkretna

Prowadzący: prof. dr hab. inż. Vasyl Martsenyuk

|  |  |
| --- | --- |
| Ćwiczenie Nr 4  Data 17.03.2024  Temat: "Analiza głównych składowych"  Wariant 9 | Dominik Katana  Informatyka  II stopień, stacjonarne,  II semestr, gr.1 |

Link do repozytorium:

<https://github.com/Dominowy/MK>

1. Polecenie: wariant 9 zadania

Zadanie dotyczy obliczenia środka, osi głównych oraz kątu obrotu danych dwuwymiarowych z pliku .csv zgodnie z wariantem zadania

2. Opis programu opracowanego (kody źródłowe, zrzuty ekranu)

import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

from sklearn.decomposition import PCA

*# Wczytanie danych z pliku CSV bez nagłówków kolumn*

try:

    df = pd.read\_csv('9.csv', header=None)

    print(df.head())  *# opcjonalne: wyświetlenie pierwszych kilku wierszy, aby sprawdzić strukturę danych*

except FileNotFoundError:

    print("Nie można znaleźć pliku dane.csv")

    exit()

except Exception as e:

    print(f"Wystąpił problem podczas wczytywania danych: {e}")

    exit()

*# Transpozycja danych*

df = df.transpose()

*# Użycie PCA do analizy głównych składowych*

pca = PCA()

pca.fit(df)

*# Środek danych (średnia)*

center = np.mean(df, axis=0)

*# Wektory własne (osie główne)*

eigenvectors = pca.components\_

*# Wizualizacja danych i osi głównych*

plt.figure(figsize=(8, 6))

plt.scatter(df.iloc[:, 0], df.iloc[:, 1], alpha=0.5, label='Dane')

*# Przypisanie wektorów własnych do zmiennych*

arrow1 = eigenvectors[:, 0]

arrow2 = eigenvectors[:, 1]

*# Wykres strzałkowy osi głównych*

plt.quiver(\*center, \*arrow1, color='r', scale=3, label='Oś główna 1')

plt.quiver(\*center, \*arrow2, color='g', scale=3, label='Oś główna 2')

plt.title('PCA - Analiza Głównych Składowych')

plt.xlabel('x')

plt.ylabel('y')

plt.legend()

plt.grid(True)

plt.axis('equal')

plt.show()

Obraz zawierający tekst, diagram, Wykres, zrzut ekranu

Opis wygenerowany automatycznie

3. Wnioski

W analizie dwuwymiarowych danych z pliku CSV za pomocą PCA udało się zidentyfikować główne składowe oraz ich geometrię, w tym środek danych, osie główne oraz kąt obrotu. Wyniki te pozwalają lepiej zrozumieć strukturę danych oraz potencjalnie wykorzystać je do dalszych analiz czy modelowania danych.