

Piotr Szczypiński

Biostatystyka
Ćwiczenia laboratoryjne

część 9

PCA i LDA

PCA

Kod analizy PCA omówiony na wykładzie jest udostępniony na stronie

<http://eletek.p.lodz.pl/pms/DydaktykaBiostatystyka.html> w pliku

<http://eletek.p.lodz.pl/pms/rfl/SkryptyLab3.zip>

Zadanie 1

Uruchom program do PCA omówiony na wykładzie. Zmodyfikuj parametry analizowanego rozkładu zmieniając wariancje oraz kąt obrotu. Jak zmieniają się wartości własne jeśli zadane wariancje są równe? Jakie są wartości kowariancji (poza diagonalą) jeśli zadany kąt obrotu jest zerem lub wielokrotnością kąta prostego?

Zadanie 2

Zmodyfikuj program tak aby mógł analizować dane w przestrzeni trójwymiarowej. Przeprowadź analizę dla danych z tabeli 1 oraz danych z tabeli 2 (w pliku SkryptyLab3.zip). Oceń wizualnie uzyskane wyniki i to czy dane spełniają założenia analizy.

LDA

Kod analizy LDA omówiony na wykładzie jest udostępniony na stronie

<http://eletek.p.lodz.pl/pms/DydaktykaBiostatystyka.html> w pliku

<http://eletek.p.lodz.pl/pms/rfl/SkryptyLab3.zip>

Zadanie 3

Uruchom program do LDA omówiony na wykładzie. Zmodyfikuj parametry analizowanego rozkładu zmieniając wariancje, przesunięcia oraz kąty obrotu obydwu rozkładów. Jak zmieniają się wartości własne jeśli środki rozkładów są blisko lub daleko od siebie? Czy możliwe jest aby w przypadku analizy dwóch klas druga wartość własna była znacząco różna od zera? Oceń wizualnie uzyskane wyniki.

Zadanie 4

Zmodyfikuj kod programu w taki sposób, żeby analizowane były trzy różne klasy. Czy możliwe jest aby w przypadku analizy trzech klas druga wartość własna była znacząco różna od zera? Oceń wizualnie uzyskane wyniki.

Zadanie 5

Pobierz z internetu dane *Iris* (https://en.wikipedia.org/wiki/Iris_flower_data_set ,

<http://archive.ics.uci.edu/ml/datasets/Iris>), których użył w swojej publikacji Ronald Fisher, twórca LDA. Dane obejmują trzy klasy odmian kwiatów opisane czterema różnymi zmiennymi.

Zmodyfikuj kod programu w taki sposób, aby analizowane były trzy różne klasy w przestrzeni czterowymiarowej. Dokonaj analizy danych *Iris* oraz oceń wizualnie uzyskane wyniki.