

Ćwiczenie 2

KLASYFIKATOR FISHERA

Zakres pracy

W ramach ćwiczenia należy do dostarczonego interfejsu, służącego do wprowadzania zbiorów uczących w przestrzeni dwuwymiarowej, dodać możliwość wyznaczania obszarów decyzyjnych przy użyciu klasyfikatora Fishera. W układzie współrzędnych należy wykreślić wektor dyskryminacyjny \mathbf{d} oraz prostą rozdzielającą klasy. Klasyfikator Fishera jest algorytmem dwuklasowym.

Informacje pomocnicze

Wektor dyskryminacyjny wyznaczamy ze wzoru:

$$\mathbf{d} = \Sigma^{-1} \Delta$$

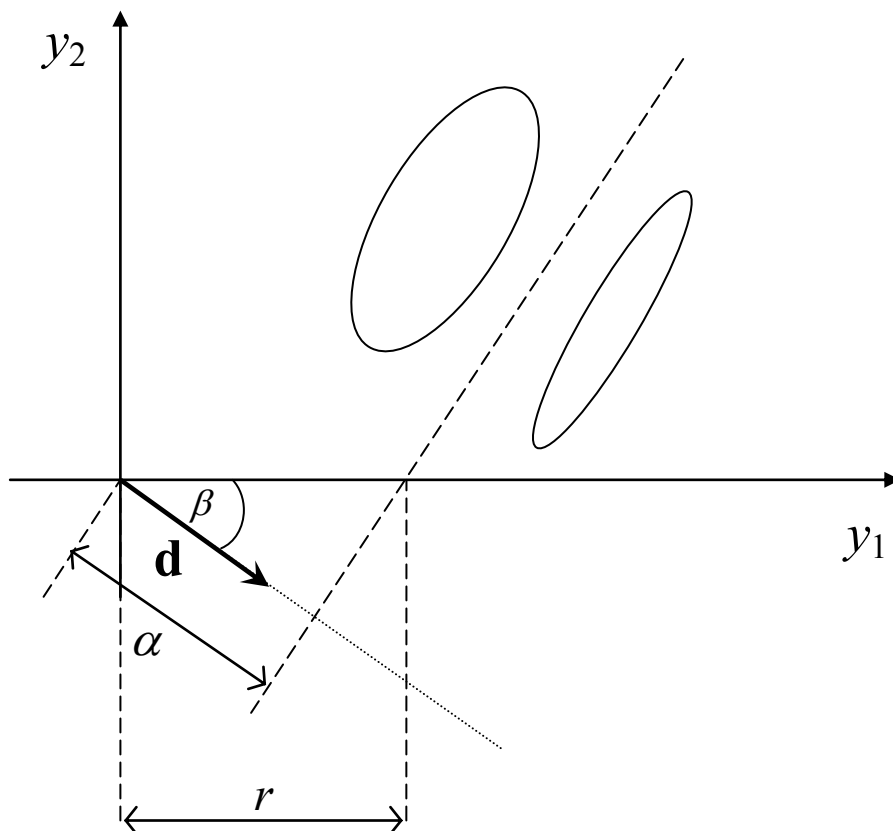
gdzie

$$\Delta = \mu_1 - \mu_2,$$

$$\Sigma = \Sigma_1 P(x_1) + \Sigma_2 P(x_2),$$

przy czym μ_i oznacza wektor średni i -tej klasy, Σ_i – macierz kowariancji i -tej klasy, $P(x_i)$ – prawdopodobieństwo *a priori* i -tej klasy.

Wektor dyskryminacyjny jest zaczepiony w początku układu współrzędnych. Linia decyzyjna jest prostopadła do wektora \mathbf{d} i przechodzi przez prostą wyznaczaną przez ten wektor w odległości α od punktu $(0,0)$.



Rys. 1. Wektor dyskryminacyjny \mathbf{d} i prosta rozdzielająca klasy

Wektor \mathbf{d} należy znormalizować:

$$\mathbf{d} = \frac{\mathbf{d}}{\|\mathbf{d}\|}$$

gdzie symbol $\|\cdot\|$ oznacza normę, czyli długość wektora. Ponieważ istotny jest tylko kierunek wektora \mathbf{d} , a nie jego zwrot, dla uproszczenia obliczeń warto wektor \mathbf{d} ustawiać zawsze w pierwszej lub czwartej ćwiartce układu współrzędnych.

Po wyznaczeniu \mathbf{d} obliczamy dwie wartości α ze wzoru

$$\alpha_{1,2} = \frac{\sigma_2^2 m_1 - \sigma_1^2 m_2 \pm \sigma_1 \sigma_2 \sqrt{(m_2 - m_1)^2 + 2(\sigma_2^2 - \sigma_1^2) \ln \frac{\sigma_2 P(x_1)}{\sigma_1 P(x_2)}}}{\sigma_2^2 - \sigma_1^2}$$

gdzie

$$m_i = \mathbf{d}^T \boldsymbol{\mu}_i,$$

$$\sigma_i^2 = \mathbf{d}^T \boldsymbol{\Sigma}_i \mathbf{d},$$

a następnie wybieramy tę wartość, która spełnia warunek $m_1 < \alpha < m_2$.

Przydatne informacje:

- znając wektor \mathbf{d} łatwo wyznaczymy współczynnik kierunkowy a prostej, na której leży \mathbf{d} , a także kąt β ,
- współczynnik kierunkowy prostej prostopadłej do \mathbf{d} jest równy $-1/a$,
- znając kąt β oraz odległość α wyznaczymy wartość r , która umożliwi obliczenie wyrazu wolnego prostej rozdzielającej klasy (rys. 1).

Wskazówki implementacyjne

Do obliczeń wykorzystujemy bibliotekę OpenCV. Podstawowym typem danych jest Mat – macierz.

Tworzenie macierzy \mathbf{A} o dwóch wierszach i trzech kolumnach, zawierającej liczby typu float i inicjowanej zerami:

```
Mat A = Mat::zeros(2, 3, CV_32F);
```

Dzięki przeciążeniu operatorów macierze można do siebie dodawać, odejmować, mnożyć przez liczbę itp.

Odwołania do elementów macierzy (nr wiersza, nr kolumny, indeksowanie od 0):

```
float a = A.at<float>(1,2);  
A.at<float>(1,2) = a;
```

Transpozycja macierzy:

```
Mat B = A.t();
```

Odwracanie macierzy:

```
Mat B = A.inv();
```

Wyznaczenie normy wektora:

```
float dNorm = norm(d);
```

Wydzielenie z macierzy kolumn z zakresu $\langle a; b \rangle$:

```
Mat B = A.colRange(Range(a,b));
```

Następująca funkcja wyliczy macierz kowariancji **S** i wektor średni **m** na podstawie macierzy **A**, która zawiera dane uczące w kolejnych kolumnach:

```
calcCovarMatrix(A, S, m, CV_COVAR_NORMAL|CV_COVAR_COLS|CV_COVAR_SCALE,  
                CV_32F);
```