Sisteme de Recunoaştere a Formelor – Lab 12

Analiza Discriminanților Liniari (Linear Discriminant Analysis)

1. Obiective

În această lucrare de laborator se va studia metoda de clasificare bazată pe analiza discriminantilor liniari (Linear Discriminant Analysis – LDA). Metoda va fi implementată pentru cazul în care spațiul trăsăturilor este bidimensional, şi avem două clase. Exemplele de clasificat sunt o mulțime de puncte 2D, iar trăsăturile sunt coordonatele lor, cea verticală şi cea orizontală. Se cunoaşte că aceste puncte aparțin de două clase, şi se ştie pentru fiecare punct clasa din care face parte.

2. Fundamente teoretice

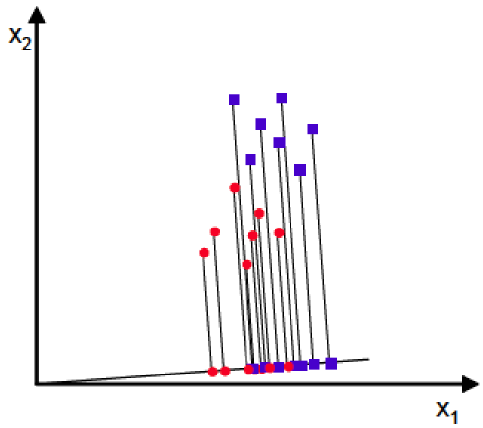
Scopul analizei discriminanților este de a clasifica obiecte (oameni, clienți, produse, etc) în două sau mai multe grupuri pe baza unei mulțimi de trăsături ce descriu obiectele (ex. sex, vârstă, venit, greutate, etc). În general, vom ataşa un obiect la unul din grupurile pre-determinate pe baza observațiilor pe care le facem cu privire la acest obiect.

Dacă presupunem că grupurile sunt separabile liniar, putem folosi modelul discriminantului liniar (LDA). Proprietatea de separabilitate liniară sugerează că grupurile pot fi separate printr-o combinație de trăsături care descriu obiectele. Dacă avem numai două trăsături, separatorii vor deveni drepte. Dacă avem trei trăsături, separatorul devine un plan, iar dacă numărul trăsăturilor este mai mare, separatorul devine un hiperplan.

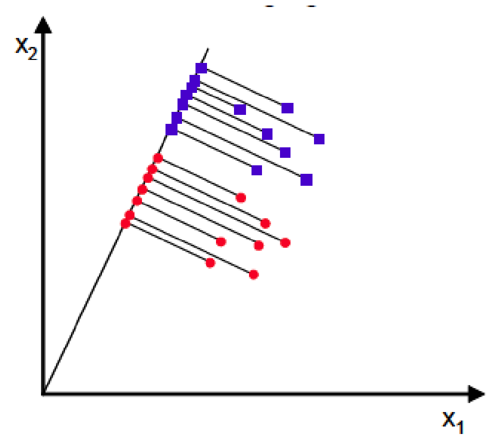
Obiectivul LDA este de a obține o **reducere a dimensiunilor** păstrând cât mai mult din informația discriminantă. Avem o mulțime de exemple D-dimensionale $\{x^{(1}, x^{(2}, \ldots, x^{(N}\}$, N_1 aparținând clasei ω_1, şi N_2 clasei ω_2, şi urmărim să obținem un scalar y prin proiecția exemplelor x pe o dreaptă:

$$\mathbf{y} = \mathbf{w}^T\mathbf{x}$$

Din toate dreptele posibile, vrem să alegem pe aceea care maximizează separabilitatea scalarilor. Acest lucru este ilustrat, pentru cazul bi-dimensional, de următoarele figuri:



Separabilitate slabă



Separabilitate bună

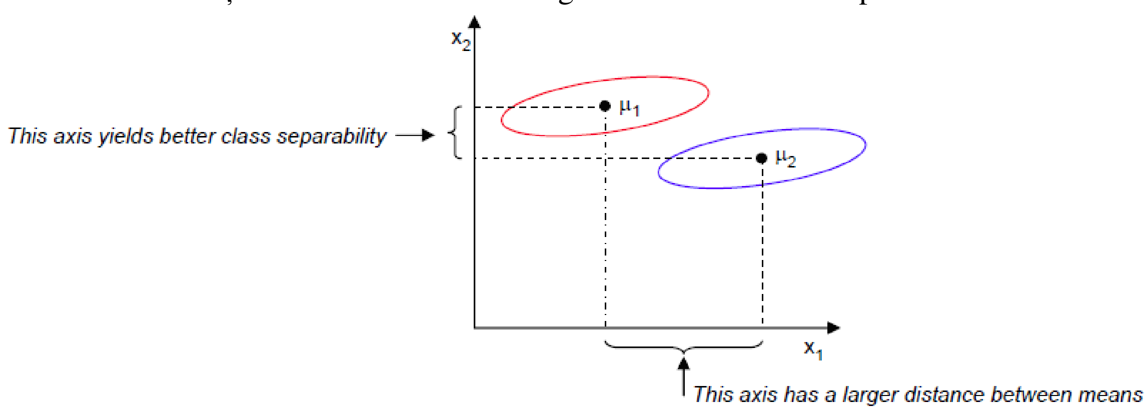
Pentru a putea găsi un vector de proiecție bun, trebuie să definim o măsură a separabilității proiecțiilor. Vectorul mediu pentru fiecare clasă în spațiul trăsăturilor **x** și **y** este:

$$\mu_i = \frac{1}{N_i}\sum_{x \in \omega_i} x \quad \text{and} \quad \tilde{\mu}_i = \frac{1}{N_i}\sum_{y \in \omega_i} y = \frac{1}{N_i}\sum_{x \in \omega_i} w^T x = w^T \mu_i$$

Am putea alege distanța dintre mediile proiectate ca fiind funcția noastră obiectiv:

$$J(w) = \left|\tilde{\mu}_1 - \tilde{\mu}_2\right| = \left|w^T(\mu_1 - \mu_2)\right|$$

Din păcate, distanța dintre mediile proiectate nu este o măsură bună, deoarece nu ia în considerare deviația standard intra-clasă. Figura următoare descrie problema:



Soluția propusă de Fisher este de a maximiza o funcție ce reprezintă diferența dintre medii (centre), normalizată de o măsură a variației intra-clasă. Pentru fiecare clasă vom defini gradul de difuzie, un echivalent al varianței, ca

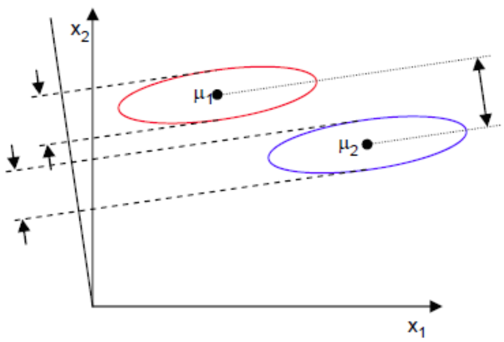
$$\tilde{s}_i^2 = \sum_{y \in \omega_i} (y - \tilde{\mu}_i)^2$$

iar cantitatea $\left(\tilde{s}_1^2 + \tilde{s}_2^2\right)$ va fi numită difuzia exemplelor proiectate.

Discriminantul liniar Fisher este definit ca funcția liniară $\mathbf{w}^T\mathbf{x}$ care maximizează funcția criteriu:

$$J(w) = \frac{|\tilde{\mu}_1 - \tilde{\mu}_2|^2}{\tilde{s}_1^2 + \tilde{s}_2^2}$$

Astfel, vom căuta o proiecţie unde exemplele din aceeaşi clasă vor fi proiectate cât mai aproape unul de celălalt, iar în acelaşi timp centrele proiectate sunt cât mai depărtate posibil.



Pentru a găsi proiecţia optimă w*, trebuie să scriem J(w) ca o funcţie explicită de w. Vom defini o măsură a difuziei în spaţiul multivariat al trăsăturilor **x**, matricile de difuzie:

$$S_i = \frac{1}{N_i} \sum_{x \in \omega_i} (x - \mu_i)(x - \mu_i)^T$$

$$S_1 + S_2 = S_W$$

Unde $\mathbf{S_W}$ este **matricea de difuzie intra-clasă**.

Difuzia proiecţiei **y** poate fi scrisă în funcţie de matricea de difuzie în spaţiul **x:**

$$\tilde{s}_i^2 = \frac{1}{N_i} \sum_{y \in \omega_i} (y - \tilde{\mu}_i)^2 = \frac{1}{N_i} \sum_{x \in \omega_i} (w^T x - w^T \mu_i)^2 = \frac{1}{N_i} \sum_{x \in \omega_i} w^T (x - \mu_i)(x - \mu_i)^T w = w^T S_i w$$

$$\tilde{s}_1^2 + \tilde{s}_2^2 = w^T S_W w$$

În mod similar, diferenţa dintre centrele proiectate poate fi scrisă în funcţie de centrele din spaţiul original al trăsăturilor:

$$(\tilde{\mu}_1 - \tilde{\mu}_2)^2 = (w^T \mu_1 - w^T \mu_2)^2 = w^T \underbrace{(\mu_1 - \mu_2)(\mu_1 - \mu_2)^T}_{S_B} w = w^T S_B w$$

Matricea $\mathbf{S_B}$ este numită **matricea de difuzie între clase**. Deoarece $\mathbf{S_B}$ este un produs exterior a doi vectori, rangul ei este maxim 1.

În sfârşit, putem exprima criteriul Fisher în funcţie de $\mathbf{S_W}$ şi $\mathbf{S_B}$ ca

$$J(w) = \frac{w^T S_B w}{w^T S_W w}$$

Pentru a găsi maximul lui J(w) vom căuta valorile nule ale derivatei.

$$\frac{d}{dw}[J(w)] = \frac{d}{dw}\left[\frac{w^T S_B w}{w^T S_W w}\right] = 0 \Rightarrow$$

$$\Rightarrow \left[w^T S_W w\right]\frac{d\left[w^T S_B w\right]}{dw} - \left[w^T S_B w\right]\frac{d\left[w^T S_W w\right]}{dw} = 0 \Rightarrow$$

$$\Rightarrow \left[w^T S_W w\right] 2 S_B w - \left[w^T S_B w\right] 2 S_W w = 0$$

Împărţind cu $\mathbf{w^T S_W w}$:

$$\frac{\left[w^T S_W w\right]}{\left[w^T S_W w\right]} S_B w - \frac{\left[w^T S_B w\right]}{\left[w^T S_W w\right]} S_W w = 0 \Rightarrow$$

$$\Rightarrow S_B w - J S_W w = 0 \Rightarrow$$

$$\Rightarrow S_W^{-1} S_B w - J w = 0$$

Rezolvând problema generalizată a valorilor proprii ($S_W^{-1} S_B w = Jw$) se obţine:

$$\boxed{w^* = \underset{w}{\operatorname{argmax}}\left\{\frac{w^T S_B w}{w^T S_W w}\right\} = S_W^{-1}(\mu_1 - \mu_2)}$$

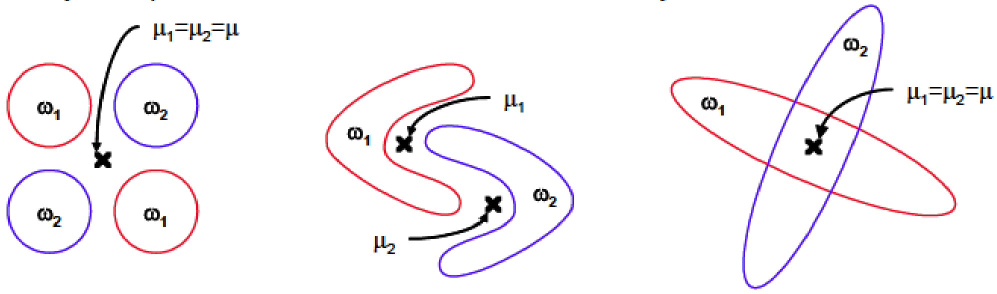
<u>Limitările LDA:</u>

- LDA produce cel mult C-1 proiecţii (C este numărul claselor)

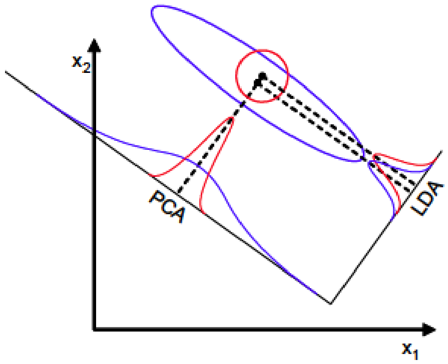
- Dacă eroarea de clasificare stabileşte că e nevoie de mai multe trăsături, alte metode trebuie folosite pentru a oferi aceste trăsături;

- LDA este o metodă parametrică, deoarece presupune verosimilităţi Gaussiene unimodale.

- Dacă distribuţiile sunt mult diferite de modelul Gaussian, proiecţiile LDA nu vor fi capabile să menţină structura complexă a datelor, necesară pentru clasificare.



- LDA va eşua dacă informaţia discriminantă nu este în medie, ci în varianţa datelor.

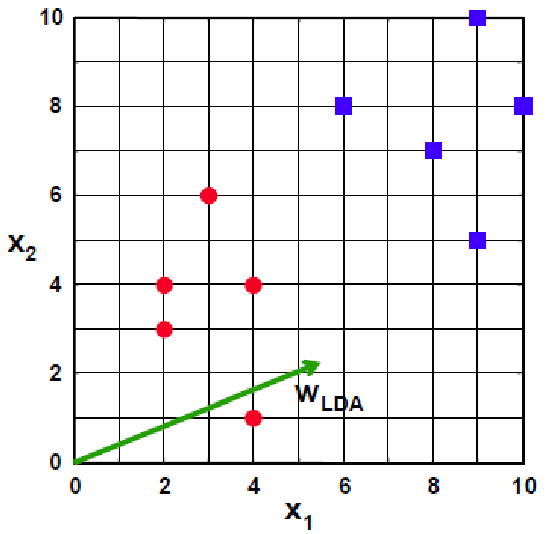


3. Exemplu numeric

Să se calculeze proiecția discriminant liniar pentru următorul set de date bidimensional.

$X_1=(x_1,x_2)=\{(4,1),(2,4),(2,3),(3,6),(4,4)\}$

$X_2=(x_1,x_2)=\{(9,10),(6,8),(9,5),(8,7),(10,8)\}$



Statisticile claselor sunt:

$$S_1=\begin{bmatrix}0.80 & -0.40\\ -0.40 & 2.60\end{bmatrix};\ S_2=\begin{bmatrix}1.84 & -0.04\\ -0.04 & 2.64\end{bmatrix}$$

$$\mu_1^T=[3.00 \quad 3.60];\qquad \mu_2^T=[8.40 \quad 7.60]$$

Difuziile intra-clasă şi între clase sunt:

$$S_B=\begin{bmatrix}29.16 & 21.60\\ 21.60 & 16.00\end{bmatrix};\ S_W=\begin{bmatrix}2.64 & -0.44\\ -0.44 & 5.28\end{bmatrix}$$

Proiecția LDA este obținută ca o soluție la problema generalizată a valorilor proprii:

$$S_W^{-1}S_B v=\lambda v \Rightarrow \left|S_W^{-1}S_B-\lambda I\right|=0 \Rightarrow \begin{vmatrix}11.89-\lambda & 8.81\\ 5.08 & 3.76-\lambda\end{vmatrix}=0 \Rightarrow \lambda=15.65$$

$$\begin{bmatrix}11.89 & 8.81\\ 5.08 & 3.76\end{bmatrix}\begin{bmatrix}v_1\\ v_2\end{bmatrix}=15.65\begin{bmatrix}v_1\\ v_2\end{bmatrix}\Rightarrow\begin{bmatrix}v_1\\ v_2\end{bmatrix}=\begin{bmatrix}0.91\\ 0.39\end{bmatrix}$$

Sau direct ca

$$w^* = S_W^{-1}(\mu_1 - \mu_2) = [-0.91 \quad -0.39]^T$$

4. Exerciții

Se dă o imagine (8 bit/pixel) conținând o mulțime de puncte 2D ce aparțin de două clase, roşu şi albastru. Paleta imaginii este modificată în aşa fel încât pe poziția 1 avem culoarea roşie, iar pe poziția 2 avem culoarea albastră.

Aplicați algoritmul LDA (soluția Fisher) pentru a găsi vectorul proiecției liniare care maximizează separabilitatea scalarilor (trăsăturile proiectate). Desenați dreapta care este definită de componentele vectorului de proiecție.

Note şi detalii de implementare (în DIBLook):

- Dacă vectorul de proiecție este $w^*=[w_{x1}, w_{x2}]^T$ unde x_1 şi x_2 sunt coordonatele punctelor 2D, atunci ecuația dreptei care defineşte această proiecție, în spațiul trăsăturilor, este $x_2=(w_{x2}/w_{x1})*x_1$. Această dreaptă trece prin originea imaginii (0, 0) şi are panta w_{x2}/w_{x1} .
- Dacă $w_{x2}/w_{x1} < 0$ atunci dreapta este situată în al doilea cadran trigonometric, şi dacă ea trece prin origine nu se poate desena pe imagine. În acest caz se poate considera o dreaptă care trece nu prin origine, ci prin (dwWidth-1, 0). Ecuația dreptei devine $x_2=(w_{x2}/w_{x1})*(x_1\text{-dwWidth}+1)$. O soluție care merge în ambele cazuri este să se deseneze o dreaptă ce trece prin centrul imaginii, cu ecuația $x_2=(w_{x2}/w_{x1})*(x_1\text{-dwWidth}/2)+\text{dwHeight}/2$.
- În ecuațiile de mai sus considerăm x_1 pe axa orizontală, şi x_2 pe axa verticală.

5. Bibliografie

[1] Ricardo Gutierrez-Osuna, *Introduction to Pattern Analysis*, Texas A&M University.
[2] Kardi Teknomo, *Linear Discriminant Analysis – tutorial*, http://people.revoledu.com/kardi/tutorial/LDA/LDA.html