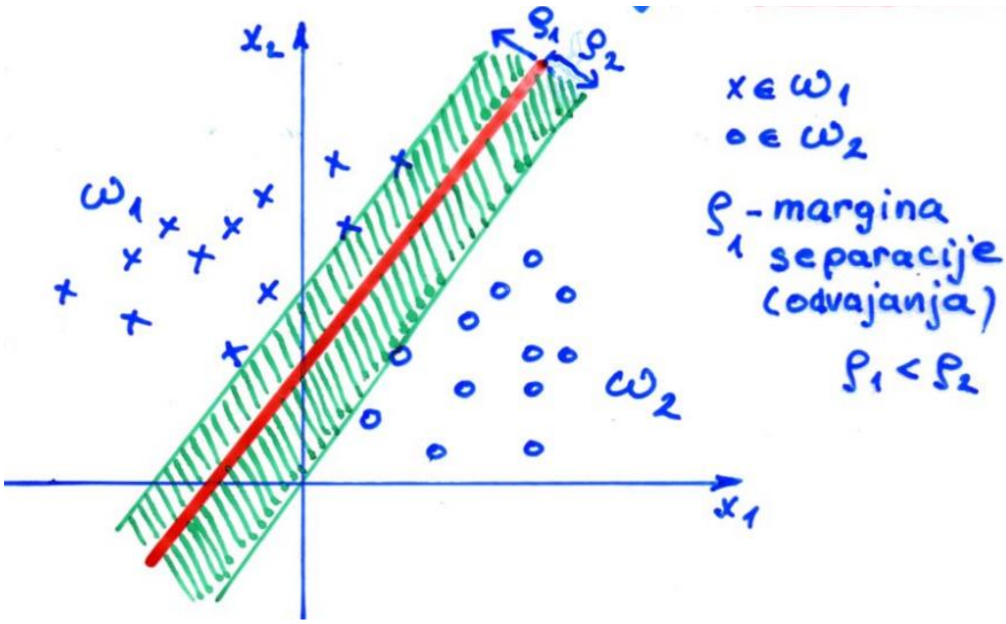
SVM – Suport Vector Machines

Strojevi s potpornim vektorima(Vapnik, 1992)

-Izvorno SVM je linearni stroj

- Osnovna zamisao SVM : konstrukcija hiperravnine kao decizijske plohe ali tako da je margina odvajanja između pozitivne „pozitivne“ i „negativne“ skupine uzoraka (za učenje) maskimalna.



Imamo skup uzoraka za učenje: $\{(\overrightarrow{x_i}, d_i)\}_{i=1}^{N}$, gdje je $\vec{x}_i$; i= 1,2,...,N ulazni vektor uzoraka za i – ti primjer.

d_i- željeni odgovor klasifikatora

Označeni uzorci:

$w_1 \rightarrow d_i = +1$

$w_2 \rightarrow d_i = -1$

Pretpostavka: razredi w_1 i w_2 su linearno separatibilni.

Jednadžba decizijske ravnine:

$\overrightarrow{w^T}\vec{x} + b = 0$

$\vec{x}$ - ulazni vektor

$\vec{w}$ - vektor težinskih koeficijenata

b – pomaknuće (w_0)

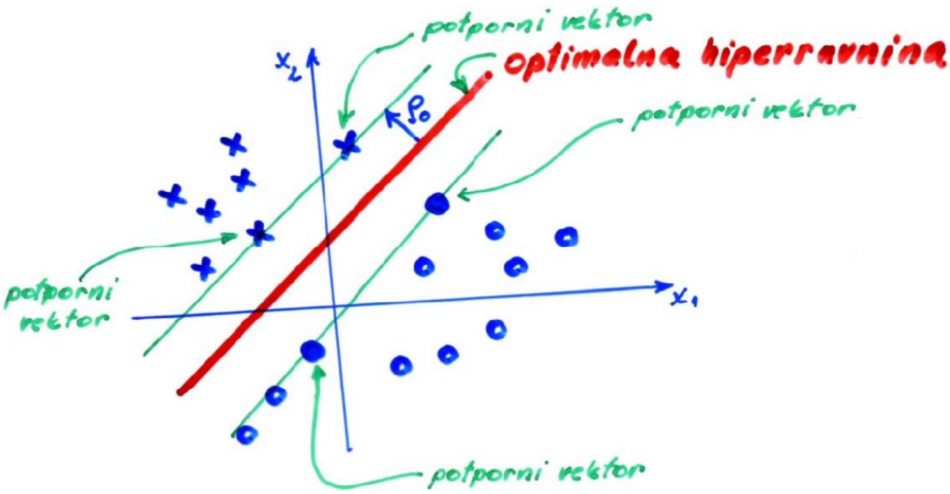
Vrijedi :

$\overrightarrow{w^T}\,\vec{x}$ + b≥0   za d_i =+1

$\overrightarrow{w^T}\,\vec{x}$ + b<0   za d_i =-1

Margina: Za zadani vektor težinskih koeficijenata $\vec{w}$ i pomaknuće b udaljenost između hiperravnine i najbliže točke (uzorka) u n –dimenzionalnom prostoru naziva se MARGINA ODVAJANJA(engl. Margin of separation) i označit ćemo je s ρ.

Cilj: Naći posebnu hiperravninu za koju je margina odvajanja ρ maksimalna. Takva hiperravnina naziva se OPTIMALNA RAVNINA.

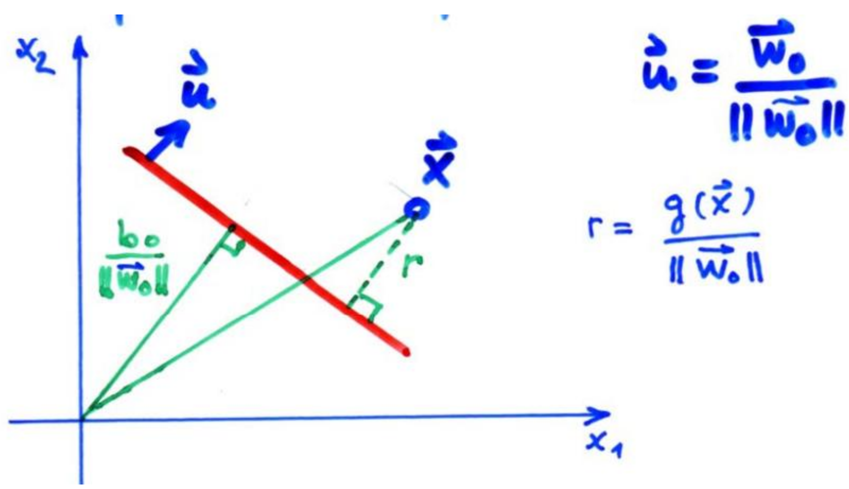


Optimalna hiperravnina: {$\vec{w}_0$, b_0} – optimalne vrijednosti

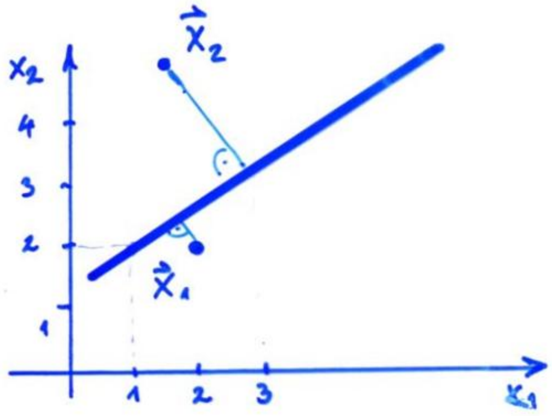
$\overrightarrow{w_0^T}\,\vec{x}$ + b_0=0

Decizijska funkcija :

g($\vec{x}$)=$\overrightarrow{w_0^T}\,\vec{x}$ + b_0 – daje mjeru udaljenosti $\vec{x}$- a od optimalne hiperravnine.



Primjer:



g($\vec{x}$)=$\overrightarrow{w_0^T}\vec{x}$ + b_0

g($\vec{x}$)= $\begin{bmatrix}-2\\3\end{bmatrix}^T * \begin{bmatrix}x_1\\x_2\end{bmatrix}$- 4

$\overrightarrow{x_1} = \begin{bmatrix}2\\2\end{bmatrix}$

g($\overrightarrow{x_1}$)= -2 - normirati r($\vec{x}_1$)= 0.54

$\overrightarrow{x_2} = \begin{bmatrix}1.5\\5\end{bmatrix}$

g($\overrightarrow{x_2}$)= 8 - normirati r($\vec{x}_2$)= 2.65

$\|\overrightarrow{w}_0\|=\sqrt{w_1^2 + w_2^2}$

Par ($\overrightarrow{w_0}, b_0$) mora zadovoljiti sljedeća ograničenja:

$\overrightarrow{w_0^T}\,\overrightarrow{x_i}$ + $b_0 \geq 1$ za d_i =+1 (1)

$\overrightarrow{w_0^T}\,\overrightarrow{x_i}$ + $b_0 < -1$ za d_i =-1 (2)

$\overrightarrow{x_i} \in \{(\overrightarrow{x_i}, d_i)\}_{i=1}^{N}$

Naravno ovo vrijedi ako su uzorci linearno odvojivi. Uvijek možemo skalirati $\overrightarrow{w_0}$ i b_0 da nejednadžbe (1) i (2) vrijede!

R=$\frac{|g(\vec{x})|}{\|\overrightarrow{w_0}\|}$ možemo skalirati $\overrightarrow{w_0}\ i\ b_0$ tako da za najbliže(hiperravnini g($\vec{x}$)!) uzorke iz w_1 i w_2 bude:

g($\vec{x}$)=1 za w_1

g($\vec{x}$)=-1 za w_2

Za uzorke (točke u n- dimenzionalnom prostoru) iz skupa za učenje i to za one koje vrijedi

$\overrightarrow{w_0^T}\ \vec{x}$ + b_0=1   za d_i =+1

$\overrightarrow{w_0^T}\ \vec{x}$ + b_0=-1   za d_i =-1

Kažemo da su potporni vektori.(support vectors)

<u>Potporni vektori</u> su one točke koje leže najbliže decizijskoj hiperravnini i zato se najteže klasificiraju. Zbog toga oni imaju izravan utjecaj na optimalni položaj decizijske hiperravinne.

Potporni vektor $\overrightarrow{x^{(s)}}$:

g($\overrightarrow{x^{(s)}}$)= $\overrightarrow{w_0^T}\ \overrightarrow{x^{(s)}}$+ b_0=(-/+)1 za $d^{(s)}$=(-/+)1

(!)Algebarska udaljenost potpornog vektora $\overrightarrow{x^{(s)}}$ od optimalne hiperravnine je

$$r=\frac{\left|g(\overrightarrow{x^{(s)}})\right|}{\|\overrightarrow{w_0}\|}\ .$$

$$r=\begin{cases}\frac{1}{\|\overrightarrow{w_0}\|} & ako\ je\ \ d^{(s)}=+1\\ -\frac{1}{\|\overrightarrow{w_0}\|} & ako\ je\ \ d^{(s)}=-1\end{cases}$$

gdje znak + označava da $\overrightarrow{x^{(s)}}$ leži na pozitivnoj strani optimalne hiperravnine, a – predznak pokazuje da je $\overrightarrow{x^{(s)}}$ na negativnoj strani optimalne ravnine.

ρ- optimalna vrijednosti MARGINE ODVAJANJA između dva razreda koji definiraju skup uzoraka za učenje.

ρ =2r

$$\rho=\frac{2}{\|\overrightarrow{w_0}\|}$$

Iz

$$\rho=\frac{2}{\|\overrightarrow{w_0}\|}$$

slijedi da maksimiziranje margine odvajanja temelji na minimizaciji norme vektora težinskih koeficijenata $\|\overrightarrow{w_0}\|$.

Optimalna hiperravnina:

$\overrightarrow{w_0^T}\ \vec{x}$ + b_0=0

Je jedinstvena u tom smislu da vektor $\|\overrightarrow{w_0}\|$ daje maksimalnu separaciju između pozitivnih i negativnih uzoraka iz skupa za učenje.

Cilj: Razvoj djelotvorne procedure (uporabom skupa uzoraka za učenje) tako da nađemo optimalnu hiperravninu uz zadovoljenje ograničenja :

$d_i * \left(\overrightarrow{w^T}\,\overrightarrow{x_i} + \mathrm{b}\right) \geq 1\ \ za$ i=1,2,...,N.

<u>Formalno postavljen problem:</u>

- Zada je skup uzoraka za učenje $\{(\overrightarrow{x_i}, d_i)\}_{i=1}^{N}$
- Nađi optimalnu vrijednost vektora težinskih koeficijenata $\overrightarrow{w}$ i pomaknuće b tako da su zadovoljena ograničenja

$$d_i * \left(\overrightarrow{w^T}\,\overrightarrow{x_i} + \mathrm{b}\right) \geq 1\ \ za\ \text{i=1,2,...,N.}$$

a pri tomu vektor težinskih koeficijenata $\overrightarrow{w}$ minimizira kriterijsku funkciju:

$$J(\overrightarrow{w}) = \frac{1}{2}\overrightarrow{w^T}\,\overrightarrow{w}$$

$$\overrightarrow{w^T}\,\overrightarrow{w}=\|\overrightarrow{w}\|^2$$

- nelinearni optimizacijski zadatak sa skupom linearnih nejednadžbi

Optimizacijski problem riještit metodom Lagrangeovih multiplikatora.

Primjer:

Određivanje vezanih ekstrema funkcije z=f(x,y) uz uvjet φ(x,y)=0 svodi se na računanje slobodnih ekstrema Lagrangeove funkcije:

$$F=f(x,y)+\lambda\,\varphi(x,y)$$

$$\frac{\partial F}{\partial x} = 0; \frac{\partial F}{\partial y} = 0 \quad \phi(\mathrm{x}, \mathrm{y}) = 0$$

Iz tog se sustava jednadžbi određuju vrijednosti za x, y i Lagrangev multiplikator λ.

- Ako je $d^2F<0$ u izračunatoj točki funkcije z=f(x,y) ima <u>maksimun</u>
- Ako je $d^2F>0$ funkcija z=f(x,y) ima <u>minimum</u>

Tražimo ekstrem funkcije :

Z=x + 2y

Uz uvjet $x^2+y^2=5$

- Lagrangeova funkcija :
  $F=x + 2y + \lambda(x^2+y^2-5)$
- Računamo :
  $F_x= \frac{\partial F}{\partial x} =1 +2x\lambda$
  $F_y= \frac{\partial F}{\partial y} =2 +2y\lambda$

- Iz sustava jednadžbi :
  $1+2\lambda x=0$
  $2+2\lambda y=0$
  $x^2+y^2=5$
  slijedi: $x =-\frac{1}{2\lambda}$; $y=-\frac{1}{\lambda}$
  uvrštavamo u 3. Jednadžbu :

$$\frac{1}{4\lambda^2}+\frac{1}{\lambda^2}=5/4\lambda^2$$
$$5(1-4\lambda^2)=0$$
$$\lambda_{1,2}=\pm\sqrt{\frac{1}{4}}$$
$$\lambda_1=+\frac{1}{2}\ ;\ \lambda_2=-\frac{1}{2}$$

za $\lambda_1=+\frac{1}{2}$ dobivamo:

$x_1=-1$ $y_1=-2$

za $\lambda_2=-\frac{1}{2}$

$x_2=1$ $y_2=2$

Računamo: $$d^2F=\frac{\partial F^2}{\partial x^2}dx^2+\ 2\frac{\partial^2 F}{\partial x\partial y}dxdy+\frac{\partial^2 F}{\partial y^2}dy^2$$

$$F_{xx}=2\lambda\ ;\ F_{yy}=2\lambda\ ;\ F_{xy}=0$$
$$d^2F=2\lambda dx^2+2\lambda dy^2=2\lambda(dx^2+dy^2)$$

za $\lambda_1=+\frac{1}{2}$ $d^2F>0$ minimum

za $\lambda_2=-\frac{1}{2}$ $d^2F<0$ maksimum funkcije f(x,y).

$$J(\overrightarrow{w})=\frac{1}{2}\overrightarrow{w^T}\ \overrightarrow{w}$$

$$d_i*\left(\overrightarrow{w^T}\ \overrightarrow{x_i}\ +\ \mathrm{b}\right)\geq 1\ \ za\ \text{i=1,2,...,N.}$$

1. Lagrangeova funkcija : $J(\overrightarrow{w},b,\lambda)=\frac{1}{2}\overrightarrow{w^T}\ \overrightarrow{w}-\sum_{i=1}^{N}\lambda_i[d_i*\left(\overrightarrow{w^T}\ \overrightarrow{x_i}\ +\ \mathrm{b}\right)-1]$

   λ_i- Lagrangeovi multiplikatori

   a) $\frac{\partial J\ (\overrightarrow{w},b,\lambda)}{\partial\overrightarrow{w}}=\overrightarrow{0}$
   b) $\frac{\partial J\ (\overrightarrow{w},b,\lambda)}{\partial b}=\overrightarrow{0}$
   c) $\lambda_i\left[d_i*\left(\overrightarrow{w^T}\ \overrightarrow{x_i}\ +\ \mathrm{b}\right)-1\right]=0\ i=1,2,\dots,N$
   d) $\lambda_i\geq 0\ i=1,2,\dots,N$
   e) $\frac{\partial J\ (\overrightarrow{w},b,\lambda)}{\partial\overrightarrow{w}}=\overrightarrow{0}$

   $\frac{\partial J\ (\overrightarrow{w},b,\lambda)}{\partial\overrightarrow{w}}=\overrightarrow{w}-\sum_{i=1}^{N}\lambda_i d_i\overrightarrow{x_i}=\overrightarrow{0}$

$$\overrightarrow{w}=\sum_{i=1}^{N}\lambda_i d_i\overrightarrow{x_i}$$

$$\frac{\partial J\ (\overrightarrow{w},b,\lambda)}{\partial b}=-\sum_{i=1}^{N}\lambda_i d_i=0$$

$$\sum_{i=1}^{N} \lambda_i d_i = 0$$

- Traženi vektor $\vec{w}$ određen je s $N_S \leq N$ vektora uzoraka

$$\vec{w} = \sum_{i=1}^{N} \lambda_i d_i \vec{x_i}$$

$$\lambda_i \neq 0$$

Vektor $\vec{w}$ je optimalno rješenje!

- Budući da je skup ograničenja

$$\lambda_i \left[d_i * \left(\overrightarrow{w^T}\ \vec{x_i}\ +\ \mathrm{b} \right) - 1 \right] = 0\ i = 1,2,\dots,N$$

  Potporni vektori leže u dvije hperravnine:

$$\overrightarrow{w^T}\ \vec{x}\ +\ \mathrm{b} = \pm 1$$

  Potporni (support) vektori su oni vektori koji leže tako da su NAJBLIŽI hiperravnini linearnog klasifikatora i određuju <u>kritične elemente skupova za učenje.</u>
- Vektori $\vec{x}_i$ za koje je $\lambda_i = 0$ mogu ležati izvan pojasa odvajanja ali mogu ležati, također, na jednoj od hiperavnina.
- Rezultirajuća (optimalna) hiperravnina je neosjetljiva na broj i položaj takvih vektora
- $\vec{w}$ je eksplicitno određena, b se može dobiti iz jednog od uvjeta

$$(*)\lambda_i \left[d_i * \left(\overrightarrow{w^T}\ \vec{x_i}\ +\ \mathrm{b} \right) - 1 \right] = 0\ i = 1,2,\dots,N$$

$$\lambda_i \neq 0.$$

  U praksi, b se obično računa kao srednja vrijednost dobivena uporabom svih uvjeta tog tipa (*)

Optimalna hiperravnina linearnog klasifikatora je jedinstvena.

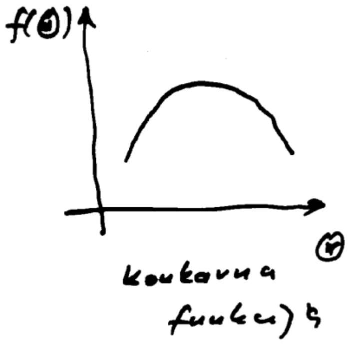
- J($\vec{w}$) je konveksna(strogo)
- Nejednadžbe su linearne (*lokalni minimum je ujedno i globalni*)

Konveksna funkcija f($\vec{\Theta}$)

F : S podskup $R^l \rightarrow R$

je konveksna u S ako za svaki $\vec{\Theta}$ i $\vec{\Theta}' \in S$ vrijedi:

$$f(\lambda\vec{\Theta} + (1-\lambda)\vec{\Theta}') \leq \lambda f(\vec{\Theta}) + (1-\lambda) f(\vec{\Theta}') \qquad \text{za svaki } \lambda \in [0,1]$$



Lagrangeov dualni problem

- Optimizacijski zadatak minimiziraj J($\vec{w}$) uz ograničenje φ_i ($\vec{w}$)≥ 0 i=1,2,...,N
  Lagrangeova funkcija J($\vec{w}$,$\vec{\lambda}$)=j($\vec{w}$)- $\sum_{i=1}^{N} \lambda_i \varphi_i(\vec{w})$
  Neka je J*($\vec{w}$,$\vec{\lambda}$)== max(J($\vec{w}$,$\vec{\lambda}$))
  Budući da je $\vec{\lambda} \geq \vec{0}\ i$ φ_i($\vec{w}$)≥ 0
- Maksimalna vrijednost Lagrangeove funkcije onda kad je λ_i=0; i=1,2,...,N ili kada je φ_i($\vec{w}$)$= 0$ (ili oboje) u tom slučaju je J*($\vec{w}$,$\vec{\lambda}$)= J($\vec{w}$)
- Ogrinalni problem je ekvivalentan sa: $\min\limits_{\vec{w}} J(\vec{w}) = \min\limits_{\vec{w}} \max\limits_{\vec{\lambda}} J(\vec{w},\vec{\lambda})$

Dualni problem : $\max\limits_{\vec{\lambda} \geq \vec{0}} \min\limits_{\vec{w}} J(\vec{w},\vec{\lambda})$

Rješenje ovog dijela:

$$\vec{w} = \sum_{i=1}^{N} \lambda_i d_i \vec{x_i}$$

i

$$\sum_{i=1}^{N} \lambda_i d_i = 0$$

Detekcija živosti ruke

85 slika IR

29 živih ruku

56 neživih (umjetnih) ruku

skup za ispitivanje

(2 tipa)

SVM

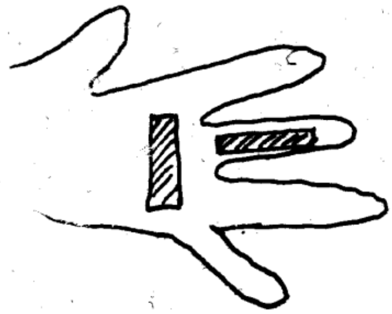
45 slika za učenje

40 slika za ispitivanje

žive ruke nije element u skupu za učenje

umjetne ruke ne pripadaju skupu za učenje

SVM s linearnim jezgrom 5% pogreške



(1) : J($\vec{w}, b, \lambda$)= $\frac{1}{2}\overrightarrow{w^T}\ \vec{w}$- $\sum_{i=1}^{N} \lambda_i d_i \overrightarrow{w^T}\ \vec{x_i}$ - b$\sum_{i=1}^{N} \lambda_i d_i$ + $\sum_{i=1}^{N} \lambda_i$

$\overrightarrow{w^T}\ \vec{w}$=$\sum_{i=1}^{N} \lambda_i d_i \overrightarrow{w^T}\ \vec{x_i}$ =$\sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j d_i d_j \overrightarrow{x_i^T}\ \vec{x_j}$

J($\vec{w}, b, \lambda$)=Q(λ)

Q(λ)= $\sum_{i=1}^{N} \lambda_i - \frac{1}{2}\sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j d_i d_j \overrightarrow{x_i^T}\ \vec{x_j}$

Uz ograničenja:

(1) $\sum_{i=1}^{N} \lambda_i d_i = 0$

(2) $\lambda_i \geq 0\ za\ i = 1,2, \dots, N$

Maksimiziraj Q(λ)!

Dualni Problem(Lagrangeova dualnost)

a) Ako prvotni problem ima optimalno rješenje tada dualni problemima također optimalno rješenje i odgovarajuća optimalna rješenja su jednaka.

Maksimiziraj J($\vec{w}, b, \lambda$) uz $\vec{w} = \sum_{i=1}^{N} \lambda_i d_i \vec{x_i}$; $\sum_{i=1}^{N} \lambda_i d_i = 0$; $\lambda_i \geq 0\ za\ i = 1,2, \dots, N$ (ograničenja samo jednadžbe, ne nejednadžbe)

- Nalazimo Langrangeove mutiplikatore koji daju optimalno rješenje!

- Neke značajke dualnog pristupa :
  - Kriterijska funkcija koja treba maksimizirati zavisi samo od ulaznih uzoraka u obliku skupa skalarnog produkta $\{\overrightarrow{x_i^T}\ \vec{x_j}\}^{N}_{(i,j)=1}$
  - Optimalno rješenje $\vec{w} = \overrightarrow{w_0}$

$$\overrightarrow{w_0} = \sum_{i=1}^{N} \lambda_{0,i} d_i \vec{x_i}$$

$$gdje\ je\ \lambda_{0,i}\ optimalni\ L.multiplikator$$

b0= 1-$\overrightarrow{w_0^T}\ \overrightarrow{x^{(s)}}\ \ za\ d^{(s)=1}$

(1) J($\vec{w}, b, \lambda$)=$\frac{1}{2}\overrightarrow{w^T}\ \vec{w}$- $\sum_{i=1}^{N} \lambda_i \left[d_i * \left(\overrightarrow{w^T}\ \vec{x_i} + \text{b}\right) - 1\right]$

(2) $\vec{w} = \sum_{i=1}^{N} \lambda_i d_i \vec{x_i}$

(3) $\sum_{i=1}^{N} \lambda_i d_i = 0$; $\lambda_i \geq 0\ za\ i = 1,2, \dots, N$

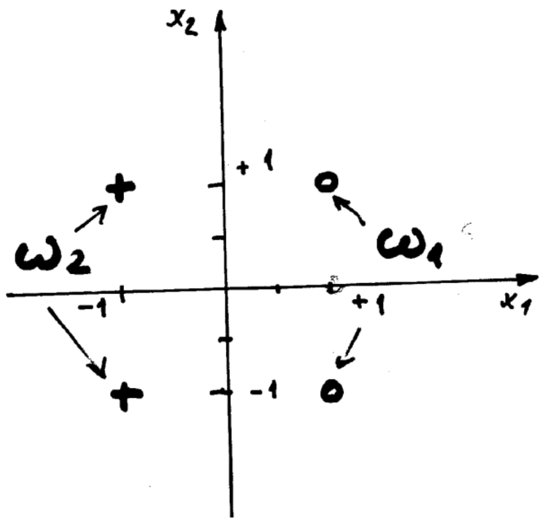
Zamjenom (2) i (3) u (1) i nakon uređivanja dobiva se :

(**) $max(\sum_{i=1}^{N} \lambda_i - \frac{1}{2}\sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j d_i d_j \overrightarrow{x_i^T}\ \vec{x_j})$

λ uz uvjet: $\sum_{i=1}^{N} \lambda_i d_i = 0$; $\lambda_i \geq 0$

(4) Optimalni Lagrangeovi multiplikatori se računaju optimiziranjem (MAKSIMIZIRANJEM) izraza (**) a optimalna se hiperravnina dobiva $\vec{w} = \sum_{i=1}^{N} \lambda_i d_i \vec{x_i}$ (optimalni Lagrangeovi multiplikatori)

Primjer:



W_1: $[1,1]^T$

$[1,-1]^T$

W_2 : $[-1,1]^T$

$[-1,-1]^T$

Vizualno: optimalna ravnina je : $w_1=1$, $w_2=0$, $b=0$ -> $g(\vec{x})=x_1$

$$g(\vec{x}) = \overrightarrow{w^T}\,\vec{x} + b = 0$$

$$w_1 * x_1 + w_2 * x_2 + b = 0$$

Ograničenja (linearne nejednadžbe):

$$w_1 + w_2 + b - 1 \geq 0$$

$$w_1 - w_2 + b - 1 \geq 0$$

$$w_1 - w_2 - b - 1 \geq 0$$

$$w_1 + w_2 - b - 1 \geq 0$$

npr. $d_i * \left(\overrightarrow{w^T}\,\overrightarrow{x_i} + b\right) \geq 1 \; za \; w_2 \;\; d_i = -1$

$(-1)*(-1*w_1+1*w_2+b)-1 \geq 0$

J($\vec{w}, b, \lambda$)=$\frac{1}{2}\overrightarrow{w^T}\ \vec{w}$- $\sum_{i=1}^{N}\lambda_i\left[d_i * \left(\overrightarrow{w^T}\ \vec{x_i}\ +\ \text{b}\right) - 1\right]$

J($\vec{w}, b, \lambda$)=$\frac{w_1^2+w_2^2}{2} - \lambda_1(w_1 + w_2 - b - 1)$- $\lambda_2(w_1 - w_2 + b - 1) - \ \lambda_3(w_1 - w_2 - b - 1) -$

$$\lambda_4(w_1 + w_2 - b - 1)$$

KKT uvjeti su zadani sa:

$$\frac{\partial J}{\partial w_1} = 0 \rightarrow w_1 = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 \ (1)$$

$$\frac{\partial J}{\partial w_2} = 0 \rightarrow w_1 = \lambda_1 - \lambda_2 - \lambda_3 + \lambda_4 \ (2)$$

$$\frac{\partial J}{\partial b} = 0 \rightarrow w_1 = \lambda_1 + \lambda_2 - \lambda_3 - \lambda_4 \ (3)$$

$$\lambda_1(w_1 + w_2 - b - 1) = 0 \ (4)$$

$$\lambda_2(w_1 - w_2 + b - 1) = 0 \ (5)$$

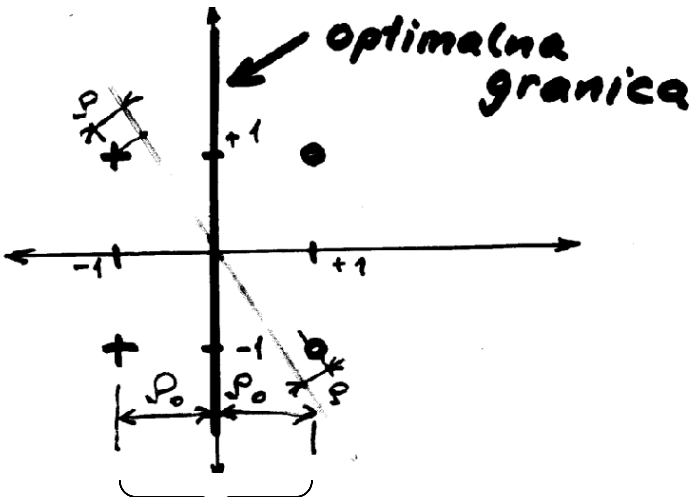
$$\lambda_3(w_1 - w_2 - b - 1) = 0 \ (6)$$

$$\lambda_4(w_1 + w_2 - b - 1) = 0 \ (7)$$

$$\lambda_1, \lambda_2, \lambda_3, \lambda_4 \geq 0$$

7 jednadžbi -> 7 nepoznanica

Znam rješenje s maksimalnom marginom (za ovaj jednostavan slučaj): w_1=1 ;w_2=0 ;b=0



maksimalna margina razdvajanja

g($\vec{x}$)=x_1=0

Uvrstimo w_1=1 ; w_2=0 i b=0 u jednadžbe:

(1) $\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 1$
(2) $\lambda_1 + \lambda_4 - \lambda_2 - \lambda_3 = 0$
(3) $\lambda_1 + \lambda_2 - \lambda_3 - \lambda_4 = 0$

Sustav linearnih jednadžbi: 3 jednadžbe i 4 nepoznanice! Očito -> više od jednog rješenja!

Međutim, svako od rješenja vodi do JEDINSTVENE (OPTIMALNE) LINIJE ODVAJANJA!

Na primjer:

(1)+(3) $2\lambda_1$+2 λ_2=1 -> $\lambda 1=\frac{1-2\lambda_2}{2}$

(1)+(2) $2\lambda_1$+2 λ_4=1

(2)+(3) $2\lambda_1$-2 λ_3=0 -> $\lambda_1= \lambda_3$

Uzmimo $\lambda_1 = \frac{1}{8} => \lambda_3 = \frac{1}{8}$

$$\lambda_2 = \frac{1-2\lambda_1}{2} = \frac{\left(1-2\cdot\frac{1}{8}\right)}{2} = \frac{\frac{3}{4}}{\frac{2}{1}} = \frac{3}{8}$$

$$\lambda_4 = \frac{1-2\lambda_1}{2} = \frac{3}{8}$$

$$\vec{w} = \sum_{i=1}^{N=4} \lambda_i d_i \vec{x_i} = \lambda_1 \begin{bmatrix}1\\1\end{bmatrix} + \lambda_2 \begin{bmatrix}1\\-1\end{bmatrix} - \lambda_3 \begin{bmatrix}-1\\1\end{bmatrix} - \lambda_4 \begin{bmatrix}-1\\-1\end{bmatrix}$$

$$\vec{w} = \frac{1}{8}\begin{bmatrix}1\\1\end{bmatrix} + \frac{3}{8}\begin{bmatrix}1\\-1\end{bmatrix} - \frac{1}{8}\begin{bmatrix}-1\\1\end{bmatrix} - \frac{3}{8}\begin{bmatrix}-1\\-1\end{bmatrix}$$

$$\vec{w} = \begin{bmatrix}1\\0\end{bmatrix}$$

$\lambda_1 = \frac{1}{4}$ $\lambda_3 = \frac{1}{4}$ $\lambda_4 = \frac{1}{4}$ $\lambda_2 = \frac{1}{4}$

$\vec{w} = \begin{bmatrix}1\\0\end{bmatrix}$ $g(\vec{x}) = x_1 = 0$

SVM za M>2 razreda?

Podsjetimo se:

za svaki od razreda tražimo optimalnu decizijsku funkciju $g_i(\vec{x})$ $i = 1,2, \dots M$

tako da

$g_i(\vec{x}) > g(\vec{x})$ za svaki $j \neq i$

ako je $\vec{x} \epsilon \omega_i$

za SVM tražimo decizijsku funkciju $g_i(\vec{x}) = 0$ takva da bude optimalna hiperravnina koja odvaja razred ω_i od svih ostalih

$g_i(\vec{x}) > 0$ za $\vec{x} \epsilon \omega_i$

$g_i(\vec{x}) < 0$ inače

Klasifikacijsko pravilo

Dodijeli $\vec{x}$ u ω_i ako

$i = \arg maks\{g_k(\vec{x})\}$.