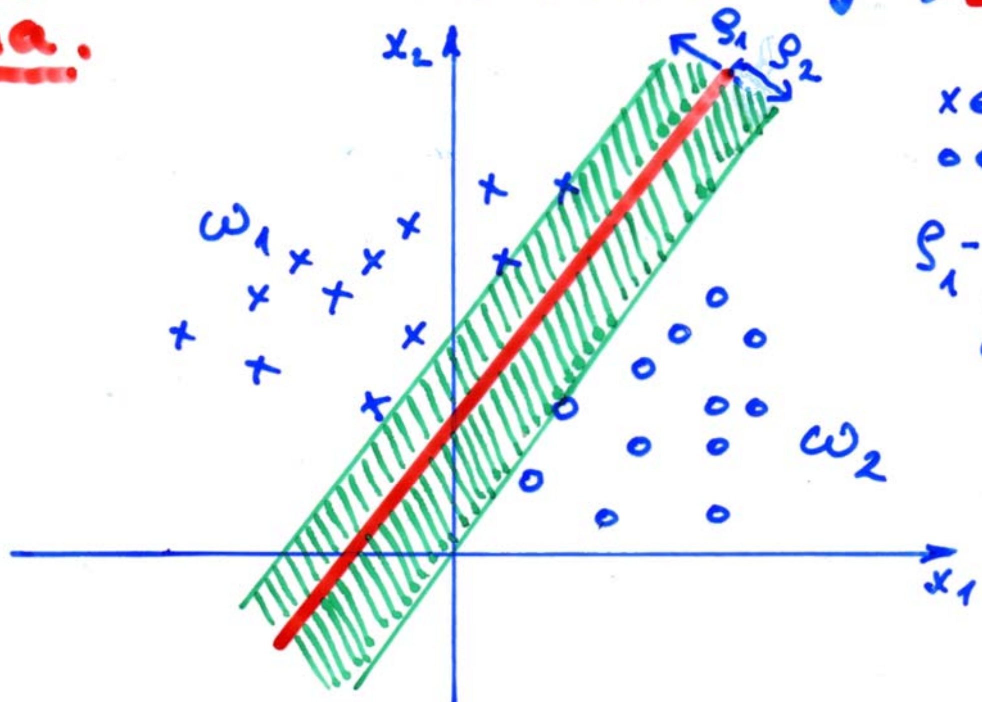
SVM - Support Vector Machines

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" i "negativne" skupine uzoraka (za učenje) maksimalna.



$x \in \omega_1$

$o \in \omega_2$

ρ_1 - margina separacije (odvajanja)

$\rho_1 < \rho_2$

Imamo skup uzoraka za učenje :

$$\{(\vec{x}_i, d_i)\}_{i=1}^{N} \quad , \text{ gdje je}$$

$\vec{x}_i\, ;\ i=1,2,\dots,N$ ulazni vektor uzoraka za i-ti primjer

d_i - željeni odgovor klasifikatora

$$\left.\begin{array}{l}\omega_1 \rightarrow d_i = +1 \\ \omega_2 \rightarrow d_i = -1\end{array}\right\} \text{ označeni uzorci}$$

Pretpostavka : razredi ω_1 i ω_2 <u>su</u> linearno separatibilni

Jednadžba decizijske ravnine :

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

$\vec{x}$ - ulazni vektor

$\vec{w}$ - vektor težinskih koeficijenata

b - pomaknuće (w_0)

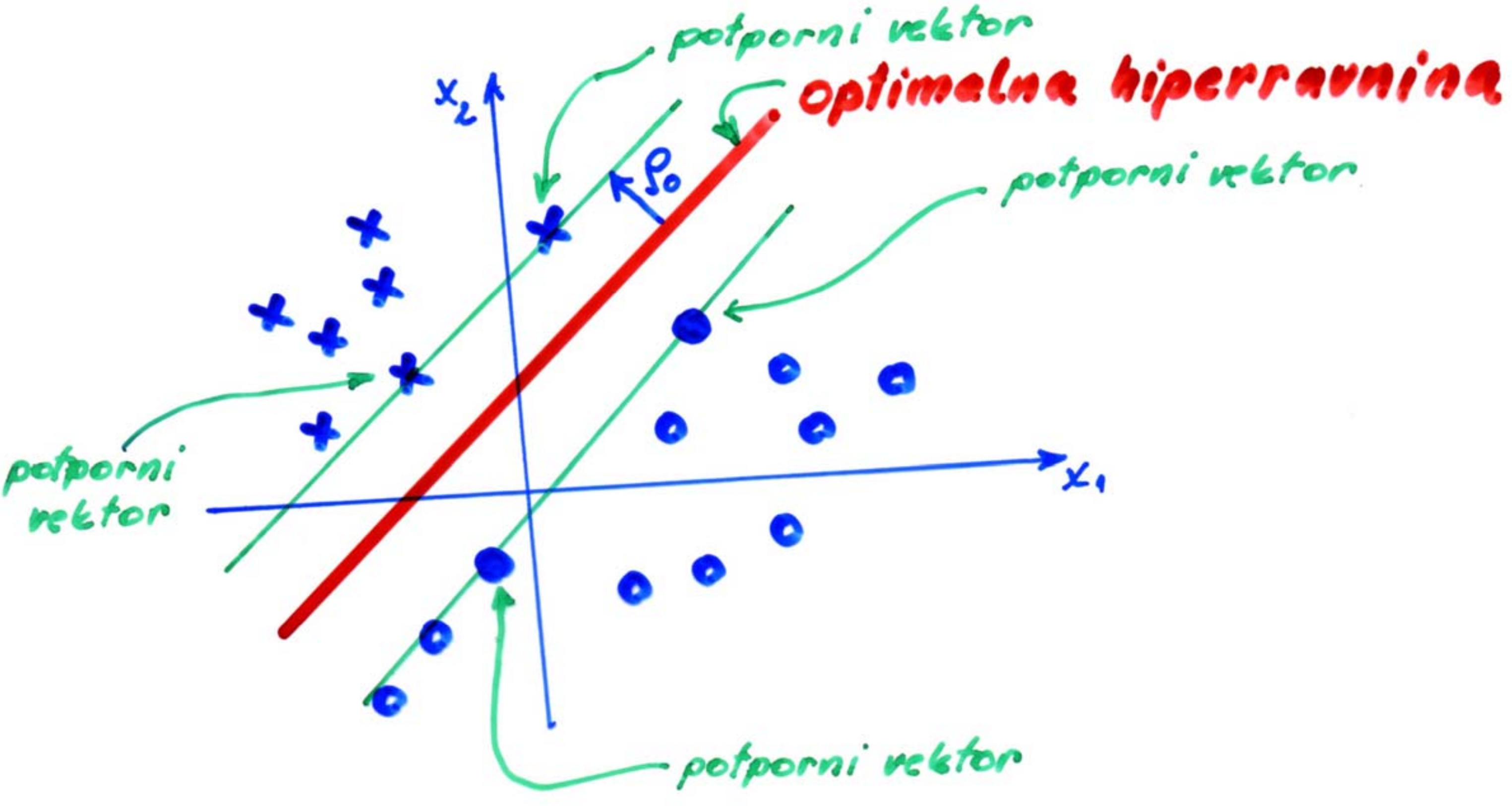
Vrijedi :

$$\vec{w}^T\vec{x} + b \geq 0 \quad \text{za} \quad d_i = +1$$

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

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.



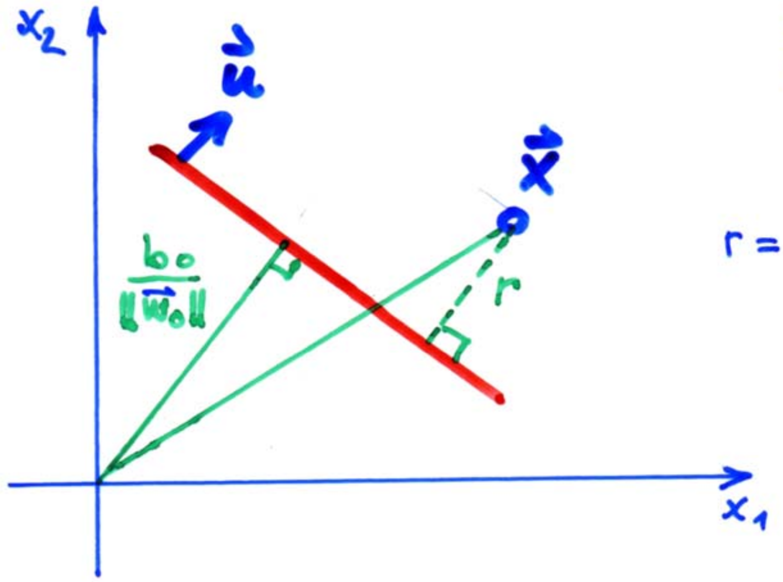
Optimalna hiperravnina: $\underbrace{\{\vec{w}_0 \,,\, b_0\}}_{\text{optimalne vrijednosti}}$

$$\vec{w}_0^T \vec{x} + b_0 = 0$$

Decizijska funkcija:

$$g(\vec{x}) = \vec{w}_0^T \vec{x} + b_0$$

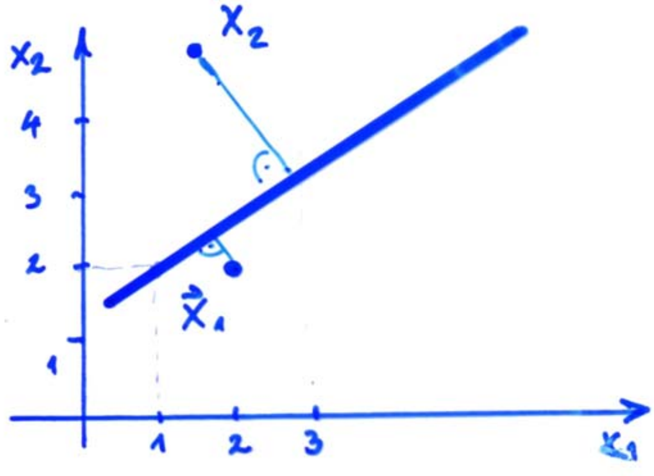
daje mjeru udaljenosti $\vec{x}$-a od optimalne hiperravnine



$$\vec{u} = \frac{\vec{w}_0}{\|\vec{w}_0\|}$$

$$r = \frac{g(\vec{x})}{\|\vec{w}_0\|}$$

Primjer:



$$g(\vec{x}) = \vec{w}_0^T \vec{x} + b_0$$

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

$$\vec{x}_1 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

$$g(\vec{x}_1) = -2 \leftarrow$$ NORMIRATI

$$r(\vec{x}_1) \doteq 0.54$$

$$\vec{x}_2 = \begin{bmatrix} 1.5 \\ 5 \end{bmatrix}$$

$$g(\vec{x}_2) = 8 \leftarrow$$ NORMIRATI

$$r(\vec{x}_2) \doteq 2.65$$

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

Par $(\vec{w}_0, b_0)$ mora zadovoljiti sljedeća ograničenja

$$\vec{w}_0^T \vec{x}_i + b_0 \geq 1 \quad \text{za } d_i = +1 \qquad (1)$$

$$\vec{w}_0^T \vec{x}_i + b_0 \leq -1 \quad \text{za } d_i = -1 \qquad (2)$$

$$\vec{x}_i \in \{(\vec{x}_i, d_i)\}_{i=1}^{N}$$

Naravno ovo vrijedi ako su uzorci linearno odvojivi.

Uvijek možemo skalirati $\vec{w}_0$ i b_0 tako da nejednadžbe (1) i (2) vrijede!

$$r = \frac{|g(\vec{x})|}{\|\vec{w}_0\|}$$

možemo skalirati $\vec{w}_0$ i b_0 tako da

za najbliže

uzorke iz ω_1 i ω_2

bude

hiperravnini $g(\vec{x})$!

$$g(\vec{x}) = 1 \quad \text{za } \omega_1$$

$$g(\vec{x}) = -1 \quad \text{za } \omega_2$$

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

$$\vec{w}_0^T \vec{x} + b_0 = 1 \quad \text{za } d_i = +1$$

i

$$\vec{w}_0^T \vec{x} + b_0 = -1 \quad \text{za } d_i = -1$$

kažemo da su <u>potporni vektori</u>.
(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 hiperravnine.

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

$$g(\vec{x}^{(s)}) = \vec{w}_0^T \vec{x}^{(s)} + b_0 = \mp 1 \quad \text{za}$$
$$d^{(s)} = \mp 1$$

(!) **<u>Algebarska</u>** udaljenost potpornog vektora $\vec{x}^{(s)}$ od optimalne hiperravnine je $r = \dfrac{g(\vec{x}^{(s)})}{\|\vec{w}_0\|}$

$$r = \begin{cases} \dfrac{1}{\|\vec{w}_0\|} & \text{ako je } d^{(s)} = +1 \\ -\dfrac{1}{\|\vec{w}_0\|} & \text{ako je } d^{(s)} = -1 \end{cases}$$

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

ρ – optimalna vrijednost MARGINE ODVAJANJA između dva razreda koji definiraju skup uzoraka za učenje

$$\rho = 2r$$

$$\boxed{\rho = \frac{2}{\|\vec{w}_0\|}}$$

Iz

$$\rho = \frac{2}{\|\vec{w}_0\|}$$

slijedi da maksimiziranje margine odvajanja temelji na minimizaciji norme vektora težinskih koeficijenata $\vec{w}_0$.

Optimalna hiperravnina:

$$\vec{w}_0^T \vec{x} + b_0 = 0$$

je jedinstvena u tom smislu da vektor $\vec{w}_0$ 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:

$$d_i \cdot (\vec{w}^T \vec{x}_i + b) \geq 1 \quad \text{za} \quad i = 1, 2, \dots, N$$

Formalno postavljen problem:

- Zadan je skup uzoraka za učenje $\{(\vec{X}_i, d_i)\}_{i=1}^{N}$
- Nađi optimalnu vrijednost vektora težinskih koeficijenata $\vec{W}$ i pomaknuće b tako da su zadovoljena ograničenja

$$d_i \cdot (\vec{W}^T \vec{X}_i + b) \geq 1 \quad \text{za}$$
$$i = 1, 2, \ldots, N$$

a pri tomu vektor težinskih koeficijenata $\vec{W}$ minimizira kriterijsku funkciju:

$$J(\vec{W}) = \frac{1}{2} \vec{W}^T \vec{W}$$

$$\vec{W}^T \cdot \vec{W} = \|\vec{W}\|^2$$

prvotni problem

- nelinearni optimizacijski zadatak sa skupom linearnih nejednadžbi

Optimizacijski problem riješiti metodom Lagrangeovih multiplikatora

Primjer :

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

Lagrangeov multiplikator

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

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

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

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

$$\frac{1}{4\lambda^2} + \frac{1}{\lambda^2} = 5 \quad / 4\lambda^2$$

$$5(1 - 4\lambda^2) = 0$$

$$\lambda_{1,2} = \pm\sqrt{\frac{1}{4}}$$

$$\lambda_1 = +\frac{1}{2} \quad ; \quad \lambda_2 = -\frac{1}{2}$$

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

$$x_1 = -1 \qquad y_1 = -2$$

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

$$x_2 = 1 \qquad y_2 = 2$$

Računamo:

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

$$F_{xx} = 2\lambda \quad ; \quad F_{yy} = 2\lambda \quad ; \quad F_{xy} = 0$$

$$d^2F = 2\lambda dx^2 + 2\lambda dy^2 = 2\lambda(dx^2 + dy^2)$$

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

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

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

$$d_i \cdot (\vec{w}^T \vec{x}_i + b) \geq 1 \quad \text{za} \quad i = 1, 2, \ldots, N$$

1. Lagrangeova funkcija:

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

λ_i - Lagrangeovi multiplikatori

a) $$\frac{\partial J(\vec{w}, b, \lambda)}{\partial \vec{w}} = \vec{0}$$

b) $$\frac{\partial J(\vec{w}, b, \lambda)}{\partial b} = 0$$

c) $$\lambda_i \left[d_i \cdot (\vec{w}^T \vec{x}_i + b) - 1 \right] = 0 \qquad i = 1, 2, \ldots, N$$

d) $$\lambda_i \geq 0 \qquad i = 1, 2, \ldots, N$$

u skladu s Karush-Kuhn-Tucker / uvjetima (KKT conditions) / (S. Theodoridis, K. Koutroumbas, 2006)

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

$$\boxed{\vec{w} = \sum_{i=1}^{N} \lambda_i d_i \vec{x}_i}$$

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

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

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

$$\vec{w} = \sum_{i=1}^{\tilde{N}} \lambda_i d_i \cdot \vec{x}_i$$

$$\lambda_i \neq 0$$

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

- Budući da je skup ograničenja

$$\lambda_i [d_i \cdot (\vec{w}^T \vec{x}_i + b) - 1] = 0$$

$$i = 1, 2, \dots, N$$

potporni vektori leže u dvije hiperravnine:

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

Potporni (support) vektori su oni vektori koji leže tako da su NAJBLIŽI hiperravnini linearnog klasifikatora i određuju <u>kritične elemente skupa 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 hiperravnina.
- Rezultirajuća (optimalna) hiperravnina je neosjetljiva na broj i položaj takvih vektora
- $\vec{w}$ je eksplicitno određena, b se može dobiti iz jednog od uvjeta

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

za $\lambda_i \neq 0$.

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 $f(\vec{\Theta})$

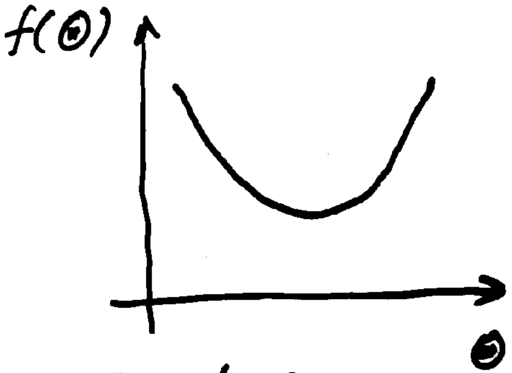
$$f : S \subseteq R^{\ell} \to R$$

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

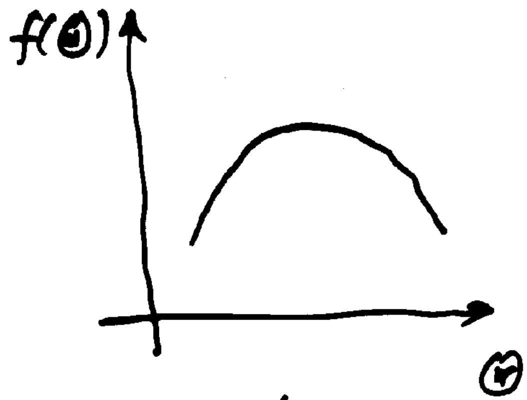
NIJE Lagrangeov multiplikator!

$$f(\lambda\vec{\Theta} + (1-\lambda)\vec{\Theta}') \le \lambda f(\vec{\Theta}) + (1-\lambda) f(\vec{\Theta}')$$

za svaki $\lambda \in [0,1]$



konveksna funkcija



konkavna funkcija

Lagrangeův dualni problem

- Optimizacijski zadatak

minimiziraj $J(\vec{w})$

uz ograničenje

$$\varphi_i(\vec{w}) \geq 0 \qquad i = 1, 2, \ldots, N$$

Lagrangeova funkcija

$$J(\vec{w}, \vec{\lambda}) = J(\vec{w}) - \sum_{i=1}^{N} \lambda_i \varphi_i(\vec{w})$$

Neka je $J^*(\vec{w}, \vec{\lambda}) = \max_{\lambda} (J(\vec{w}, \vec{\lambda}))$

Budući da je $\vec{\lambda} \geq \vec{0}$ i $\varphi_i(\vec{w}) \geq 0$

- Maksimalna vrijednost Lagrangeove funkcije ~~onda kad~~ je

$\lambda_i = 0; \; i = 1, 2, \ldots N$ ili

kada je $\varphi_i(\vec{w}) = 0$ (ili oboje)

u tom slučaju je

$$J^*(\vec{w}, \vec{\lambda}) = J(\vec{w})$$

Originalni problem je ekvivalentan sa:

$$\boxed{\min_{\vec{w}} J(\vec{w}) = \min_{\vec{w}} \max_{\vec{\lambda}} J(\vec{w}, \vec{\lambda})}$$

Dualni problem:

$$\max_{\vec{\lambda} \geq \vec{0}} \underbrace{\min_{\vec{w}} J(\vec{w}, \vec{\lambda})}$$

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 nеživih (umjetnih) ruku
(2 tipa)

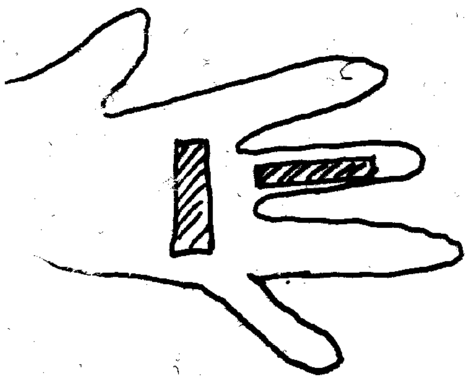
skup za ispitivanje

SVM

45 slika za učenje
40 slika za ispitivanje

žive ruke $\notin$ skupu za učenje
umjetne ruke ne pripadaju skupu za učenje

SVM s linearnim jezgrom 5% pogreške



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

$$\vec{w}^T \cdot \vec{w} = \sum_{i=1}^{N} \lambda_i d_i \vec{w}^T\vec{x}_i = \sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_i \lambda_j d_i d_j \vec{x}_i^T \cdot \vec{x}_j$$

$$J(\vec{w}, b, \lambda) = Q(\lambda)$$

$$Q(\lambda) = \sum_{i=1}^{N}\lambda_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_i \lambda_j d_i d_j \vec{x}_i^T \cdot \vec{x}$$

uz ograničenja:

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

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

<u>maksimiziraj</u> <u>$Q(\lambda)$!</u>

Dualni problem –

– Lagrangeova dualnost
(Wolfe dual representation)

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

Dualni problem

Maksimiziraj $J(\vec{w}, b, \lambda)$

uz

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

VEKTORI UZORAKA

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

ograničenja jednadžbe ne nejednadžbe !!! ...

$$\lambda_i \geq 0 \qquad i = 1, 2, \ldots, N$$

– Nalazimo Lagrangeove multiplikatore koji daju optimalno rješenje!

Neke značajke dualnog pristupa:

- Kriterijska funkcija koja se treba maksimizirati zavisi samo od ulaznih uzoraka u obliku skupa skalarnog produkta

$$\left\{ \vec{x}_i^{\,T} \vec{x}_j \right\}_{(i,j)=1}^{N}$$

Optimalno rješenje $\vec{w} = \vec{w}_0$

$$\vec{w}_0 = \sum_{i=1}^{N} \lambda_{0,i}\, d_i\, \vec{x}_i$$

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

$$b_0 = 1 - \vec{w}_0^{\,T} \vec{x}^{(s)} \quad \text{za } d^{(s)} = 1$$

(1) $$J(\vec{w}, b, \lambda) = \frac{1}{2}\vec{w}^T\vec{w} - \sum_{i=1}^{N} \lambda_i [d_i(\vec{w}^T\vec{x}_i + b) - 1]$$

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

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

$$\lambda_i \geq 0$$

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

(**) $$\max_{\lambda} \left(\sum_{i=1}^{N} \lambda_i - \frac{1}{2} \sum_{i,j}^{N} \lambda_i \lambda_j d_i d_j \vec{x}_i^T \vec{x}_j \right)$$

uz uvjet

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

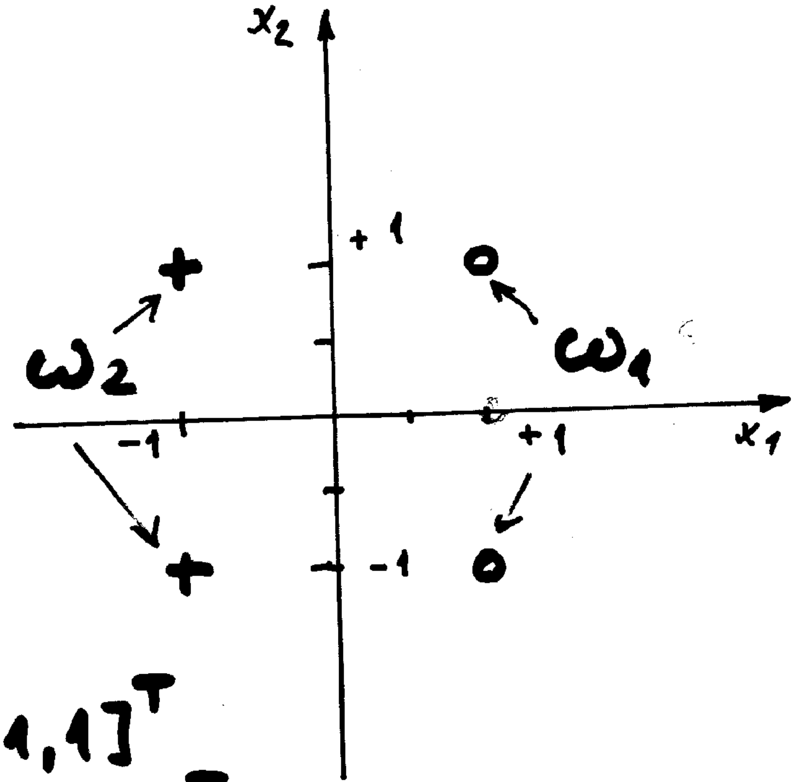
$$\lambda_i \geq 0$$

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:



ω_1 : $[1,1]^T$
$[1,-1]^T$

ω_2 : $[-1,1]^T$
$[-1,-1]^T$

vizualno:
optimalna ravnina
je : $w_1 = 1$
$w_2 = 0$
$b = 0$

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

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

$$w_1 \cdot x_1 + w_2 \cdot x_2 + b = 0$$

Ograničenja (linearne nejednadžbe):

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

pr.

$$d_i \cdot (\vec{w}^T\vec{x}_i + b) \geq 1$$

za ω_2 $d_i = -1$

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

$$J(\vec{w}, b, \lambda) = \frac{1}{2}\vec{w}^T \cdot \vec{w} - \sum_{i=1}^{4} \lambda_i [d_i \cdot (\vec{w}^T \vec{x}_i + b) - 1]$$

$$J(\vec{w}, b, \lambda) = \frac{w_1^2 + w_2^2}{2} - \lambda_1 (w_1 + w_2 + b - 1)$$
$$- \lambda_2 (w_1 - w_2 + b - 1)$$
$$- \lambda_3 (w_1 - w_2 - b - 1)$$
$$- \lambda_4 (w_1 + w_2 - b - 1)$$

KKT uvjeti su zadani sa:

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

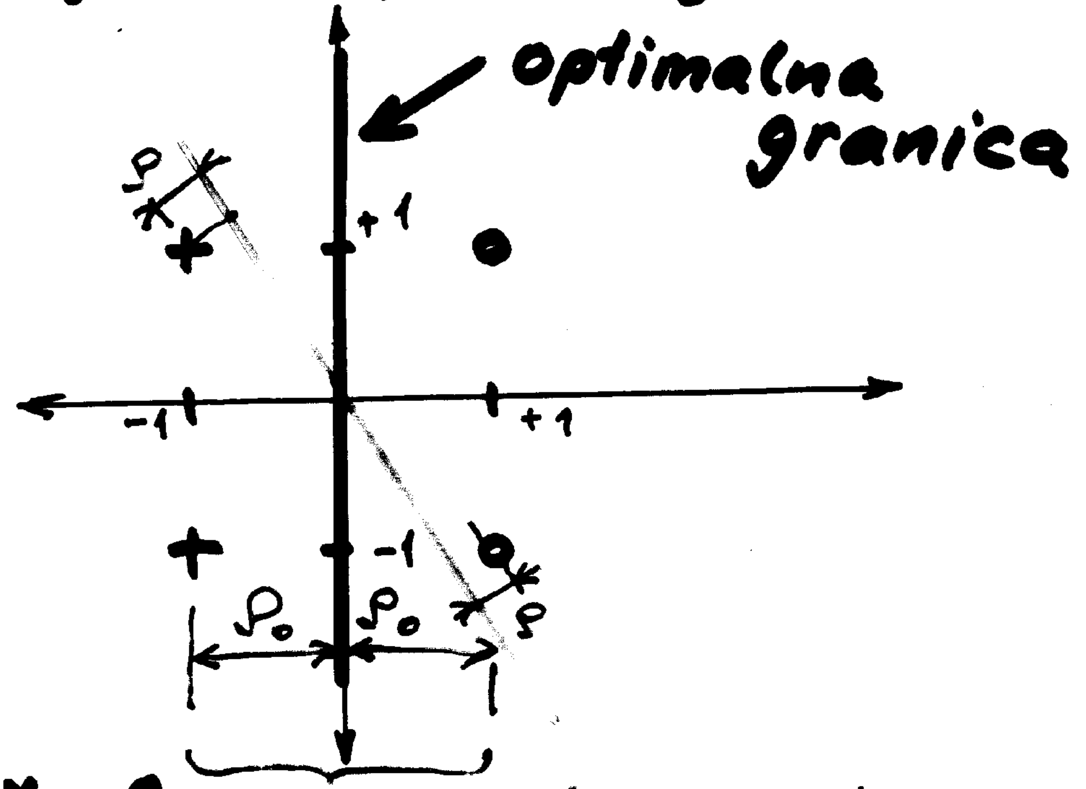
$$\frac{\partial J}{\partial w_2} = 0 \Rightarrow \quad w_2 = \lambda_1 + \lambda_4 - \lambda_2 - \lambda_3 \quad (2)$$

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

$$\lambda_1 (w_1 + w_2 + b - 1) = 0 \quad (4)$$
$$\lambda_2 (w_1 - w_2 + b - 1) = 0 \quad (5)$$
$$\lambda_3 (w_1 - w_2 - b - 1) = 0 \quad (6)$$
$$\lambda_4 (w_1 + w_2 - b - 1) = 0 \quad (7)$$
$$\lambda_1, \lambda_2, \lambda_3, \lambda_4 \geq 0$$

7 jednadžbi $\rightarrow$ 7 nepoznanica

Znamo rješenje s maksimalnom marginom (za ovaj jednostavan slučaj): $w_1 = 1 \quad w_2 = 0 \quad b = 0$



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

maksimalna margina razdvajanja

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 jednadž. i 4 nepoznanice!

Očito → više od jednog rješenja!

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

Na primjer:

① + ③ $\quad 2\lambda_1 + 2\lambda_2 = 1 \quad \Rightarrow \quad \lambda_1 = \frac{1 - 2\lambda_2}{2}$

① + ② $\quad 2\lambda_1 + 2\lambda_4 = 1$

② + ③ $\quad 2\lambda_1 - 2\lambda_3 = 0 \quad \Rightarrow \quad \lambda_1 = \lambda_3$

uzmimo $\lambda_1 = \frac{1}{8} \Rightarrow \lambda_3 = \frac{1}{8}$

$$\lambda_2 = \frac{1 - 2\lambda_1}{2} = \frac{1 - 2\cdot\frac{1}{8}}{2} = \frac{\frac{3}{4}}{\frac{2}{1}} = \frac{3}{8}$$

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

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

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

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

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

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

SVM za $M>2$ razreda?

Podsjetimo se:

- za svaki od razreda tražimo optimalnu decizijsku funkciju

$$g_i(\vec{x}) \quad i=1,2,\dots M$$

tako da

$$g_i(\vec{x}) > g(\vec{x}) \quad \forall j \neq i$$

$$\text{ako je } \vec{x} \in \omega_i$$

<u>za SVM tražimo</u>

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

$$g_i(\vec{x}) > 0 \quad \text{za } \vec{x} \in \omega_i$$

$$g_i(\vec{x}) < 0 \quad \text{inače}$$

Klasifikacijsko pravilo

dodijeli $\vec{x}$ u ω_i

ako

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