Poopćene (linearne) decizijske funkcije

- složenost granica:

linearne $\longrightarrow$ vrlo nelinearne!

- vrlo nelinearne?

Rješenje:

Poopćeni oblik (linearne) decizijske funkcije:

$$d(\vec{x}) = w_1 f_1(\vec{x}) + w_2 f_2(\vec{x}) + \cdots + w_k f_k(\vec{x}) + w_{k+1}$$

$$= \sum_{i=1}^{k+1} w_i f_i(\vec{x})$$

$\{f_i(\vec{x})\}$, $i = 1, 2, \ldots, k$

$f_{k+1} = 1$

Pazi: $f_i(\vec{x})$

$$\vec{x}^* = \begin{bmatrix} f_1(\vec{x}) \\ f_2(\vec{x}) \\ \vdots \\ f_k(\vec{x}) \\ 1 \end{bmatrix}$$

$$\boxed{d(\vec{x}) = \vec{w}^T \vec{x}^*}$$

$d(\vec{x})$ je linearna funkcija po $\vec{x}^*$

$$d(\vec{x}) = \vec{w}^T \vec{x}^*$$

$$\vec{x}^* = \begin{bmatrix} f_1(\vec{x}) \\ f_2(\vec{x}) \\ \vdots \\ f_k(\vec{x}) \\ 1 \end{bmatrix} \qquad \vec{w} = (w_1, w_2, \dots w_k, w_{k+1})^T$$

Decizijska funkcija $d(\vec{x})$ može se promatrati kao linearna funkcija u k -dimenzionalnom prostoru!

Transformacija:

$$\vec{x} \Big/ \longrightarrow \vec{x}^* \Big/$$

n-dimenzija k-dimenzija

Vrijedi:

$$k > n$$

$d(\vec{x})$ se često naziva i "virtualno linearna"

Kako izabrati $\{f_i(\vec{x})\}_{i=1}^{k}$?

Mogućnost:

$\{ f_i(\vec{x}) \}$ $\quad f_i(\vec{x})$ u obliku polinoma

- <u>Najjednostavniji slučaj: linearna funkcija</u>

$\vec{x} = (x_1, x_2, \ldots, x_n)^T$

$f_i(\vec{x}) = x_i \qquad \vec{x} = \vec{x}^*$

$k = n$

- <u>Polinom drugog stupnja:</u>

Npr. $n = 2 \qquad \vec{x} = (x_1, x_2)^T$

$$d(\vec{x}) = w_{11} x_1^2 + w_{12} x_1 x_2 + w_{22} x_2^2 + w_1 x_1 + w_2 x_2 + w_3$$

$$d(\vec{x}) = \vec{w}^T \vec{x}^*$$

$$\vec{x}^* = (x_1^2, x_1 x_2, x_2^2, x_1, x_2, 1)^T$$

Opći slučaj kvadratnog oblika:

$$d(\vec{x}) = \underbrace{\sum_{j=1}^{n} w_{jj} x_j^2}_{n\text{-izraza}} + \underbrace{\sum_{j=1}^{n-1} \sum_{k=j+1}^{n} w_{jk} x_j x_k}_{\frac{n(n-1)}{2}\text{ izraza}} + \underbrace{\sum_{j=1}^{n} w_j x_j}_{n\text{-izraza}} + \underbrace{w_{n+1}}_{1\text{ izraz}}$$

Kvadratnu decizijsku funkciju možemo promatrati kao linearnu ("linear by virtue") funkciju s

$$k = \frac{(n+1)(n+2)}{2} \text{ varijabli}$$

$$\{ f_1(\vec{x}) \dots f_n(\vec{x}) \} = \{ x_1^2, x_2^2, \dots, x_n^2 \}$$

$$\{ f_{n+1}(\vec{x}), \dots, f_{2n}(\vec{x}) \} = \{ x_1, x_2, \dots, x_n \}$$

$$\{ f_{2n+1}(\vec{x}), \dots, f_k(\vec{x}) \} = \{ x_1 x_2, x_1 x_3, \dots, x_{n-1} x_n \}$$

- Decizijske funkcije višeg reda (polinomi stupnja $r > 2$) možemo promatrati kao linearne funkcije

$$\frac{(n+r)!}{r!\,n!} \text{ varijabli}$$

n - dimenzionalnost vektora $\vec{x}$

r - stupanj polinoma

Funkcije iz skupa $\{ f_i(\vec{x}) \}$ određene su

s :

$$f_i(\vec{x}) = x_{p_1}^{s_1} x_{p_2}^{s_2} \dots x_{p_r}^{s_r}$$

$p_1, p_2, \dots p_r = 1, 2, \dots n$

$s_1, s_2, \dots s_r = 0, 1$

Decizijsku funkciju u obliku polinoma r-tog stupnja možemo zapisati u rekurzivnom obliku:

$$\boxed{d^r(\vec{x}) = \left(\sum_{p_1=1}^{n} \sum_{p_2=p_1}^{n} \dots \sum_{p_r=p_{r-1}}^{n} w_{p_1 p_2 \dots p_r} x_{p_1} x_{p_2} \dots x_{p_r} \right) + d^{r-1}(\vec{x}) \quad ; \quad d^0(\vec{x}) = w_{n+1}}$$

Primjer:

za $r=2$ i $n=2$:

$$d^2(\vec{x}) = \sum_{p_1=1}^{2} \sum_{p_2=p_1}^{2} w_{p_1 p_2} x_{p_1} x_{p_2} + d^1(\vec{x})$$

$$d^1(\vec{x}) = \sum_{p_1=1}^{2} w_{p_1} x_{p_1} + d^0(\vec{x})$$

$$d^1(\vec{x}) = w_1 x_1 + w_2 x_2 + w_3$$

$$d^2(\vec{x}) = w_{11}x_1^2 + w_{12}x_1x_2 + w_{22}x_2^2 + d^1(\vec{x})$$

$$d^2(\vec{x}) = w_{11}x_1^2 + w_{12}x_1x_2 + w_{22}x_2^2 + w_1x_1 +$$
$$+ w_2x_2 + w_3$$

$$\vec{W}^T = \begin{bmatrix} w_{11} \\ w_{12} \\ w_{22} \\ w_1 \\ w_2 \\ w_3 \end{bmatrix} \qquad \vec{x}^* = [x_1^2, x_1x_2, x_2^2, x_1, x_2, 1]^T$$

Primjer: $r = 3$; $n = 2$

$$d^3(\vec{x}) = \left(\sum_{p_1=1}^{2} \sum_{p_2=p_1}^{2} \sum_{p_3=p_2}^{2} w_{p_1p_2p_3} x_{p_1} x_{p_2} x_{p_3} \right)$$
$$+ d^2(\vec{x}) \ ;$$

$$d(\vec{x}) = w_{111}x_1^3 + w_{112}x_1^2x_2 + w_{122}x_1x_2^2 +$$
$$w_{222}x_2^3 + d^2(\vec{x})$$

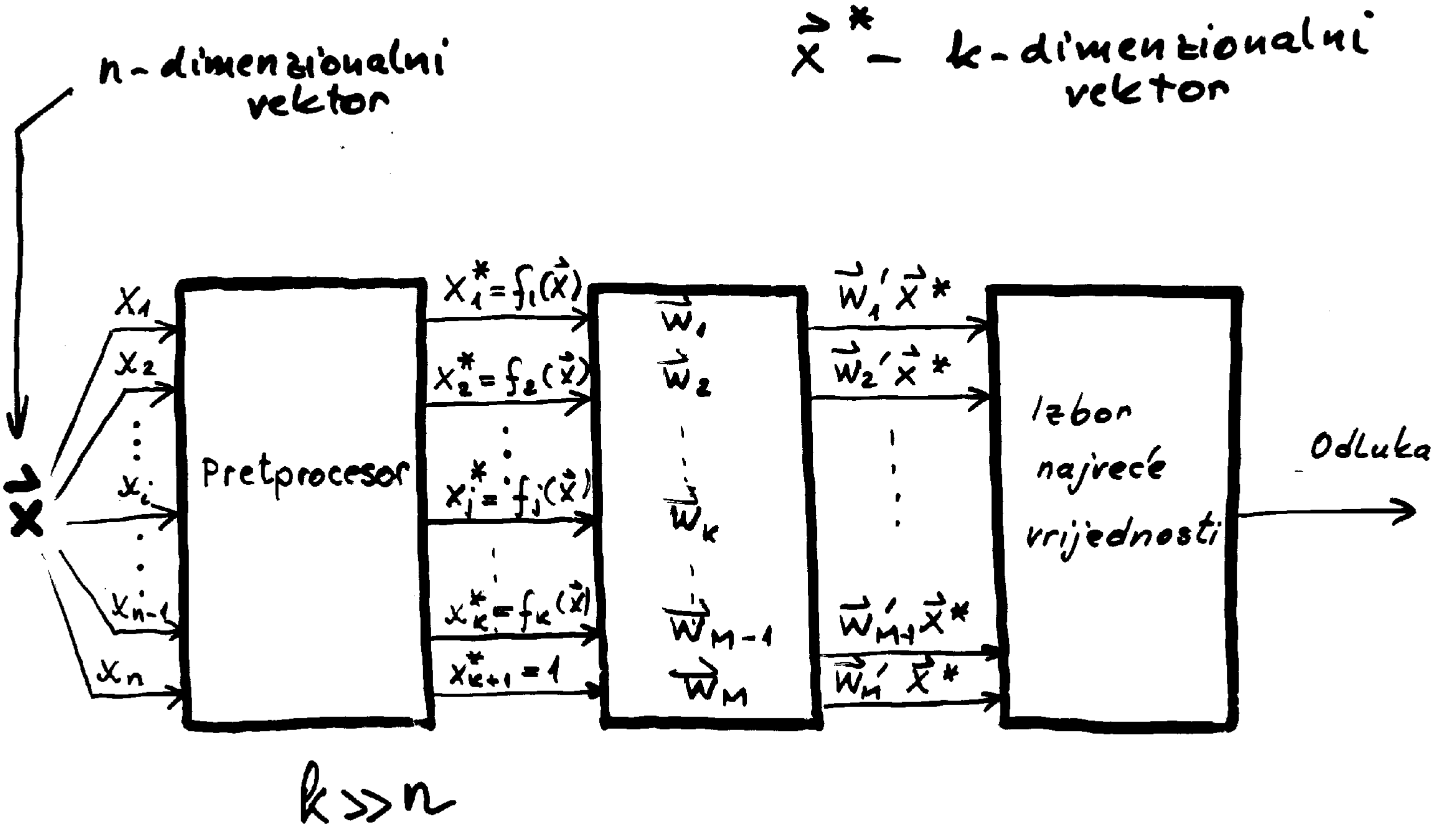
$$d(\vec{x}) = w_{111}x_1^3 + w_{112}x_1^2x_2 + w_{122}x_1x_2^2 +$$
$$w_{222}x_2^3 + w_{11}x_1^2 + w_{12}x_1x_2 + w_{22}x_2^2 +$$
$$w_1x_1 + w_2x_2 + w_3$$

Cijena koju plaćamo za linearizaciju:

– broj koeficijenata (težinskih faktora) za funkciju r-tog stupnja i različite n:

n \ r	1	2	3	…	6	10
1	2	3	4		7	11
2	3	6	10		28	66
3	4	10	20		84	286
⋮						
9	10	55	220		5005	92378
10	11	66	286		8008	184756

Dimenzionalnost "lineariziranog prostora"!



Blok-shema sustava za raspoznavanje

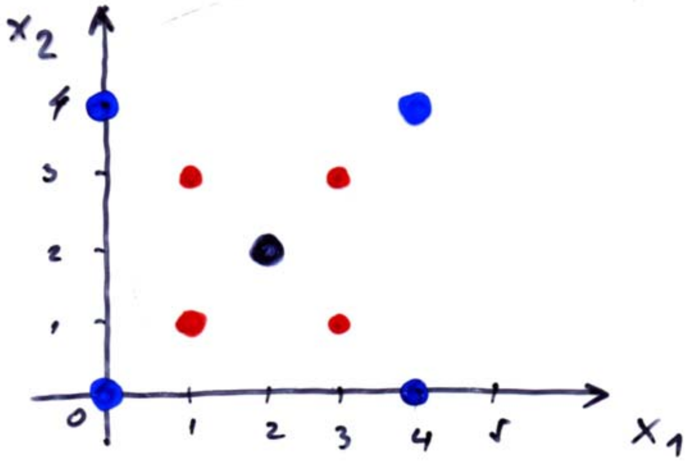
PRIMJER

Zadan je skup uzoraka za učenje:

● $U_4^1 = \{(0,0)^T, (0,4)^T, (4,0)^T, (4,4)^T\}$

● $U_4^2 = \{(1,1)^T, (1,3)^T, (3,1)^T, (3,3)^T\}$;

● $U_1^3 = \{(2,2)^T\}$



● $\in \omega_1$

● $\in \omega_2$

● $\in \omega_3$

U_e^s ← oznaka razreda; broj uzoraka u raz.

Zadane uzorke ne možemo odijeliti linearnim dec. funkcijama!

Pokušajmo kvadratnim! $\Rightarrow$ $r=2$; $n=2$

$$K = \frac{(n+1)(n+2)}{2} = 6 \text{ — dimenzionalni prostor}$$

$\underline{n=2}$ 2-dim. $\longrightarrow$ 6-dim.

Preslikavanje 2- u 6-dimenz. prostor funkcijama

$f_1 = x_1^2 \quad f_2 = x_1 x_2 \quad f_3 = x_2^2 \quad f_4 = x_1 \quad f_5 = x_2 \quad f_6 = 1$

Preslikan skup za učenje je:

$U_4^1 = \{(0,0,0,0,0,1)^T, (0,0,16,0,4,1)^T, (16,0,0,4,0,1)^T,$
$(16,16,16,4,4,1)^T\}$

$U_4^2 = \{(1,1,1,1,1,1)^T, (1,3,9,1,3,1)^T, (9,3,1,3,1,1)^T,$
$(9,9,9,3,3,1)^T\}$

$U_1^3 = \{(4,4,4,2,2,1)^T\}$.

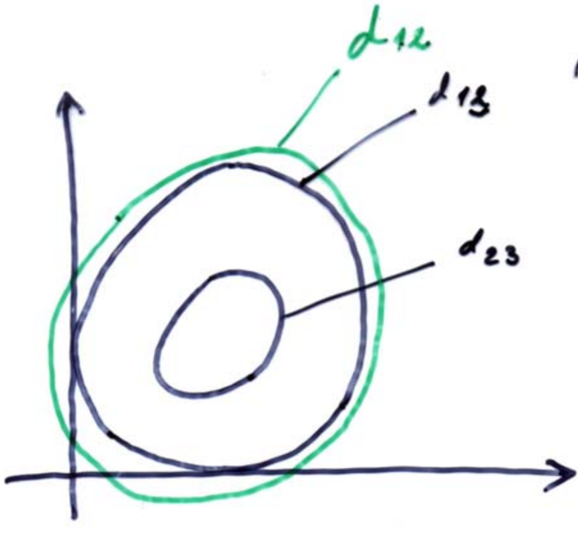
$$\vec{X}^* = \begin{bmatrix} x_1^2 \\ x_1 \cdot x_2 \\ x_2^2 \\ x_1 \\ x_2 \\ 1 \end{bmatrix}$$

$$d(\vec{x}^*) = \vec{W}^T \cdot \vec{X}^*$$

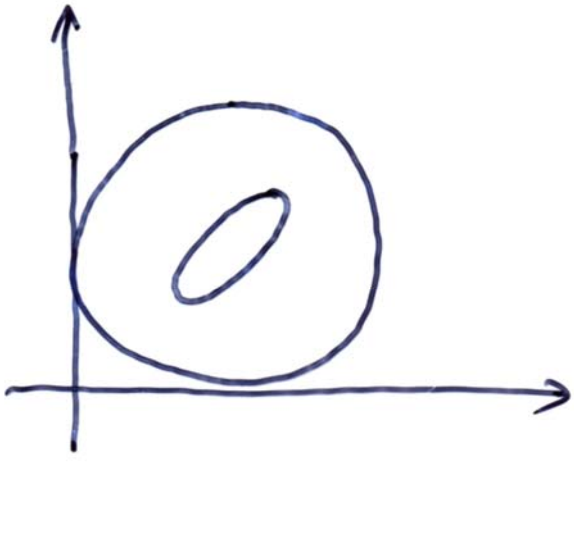
$$\vec{W} = ?$$

Upotrijebimo poopćeni algoritam perceptrona ($M=3$).

Dva od mogućih rezultata (u zavisnosti od izbora početnih vrijednosti :



Decizijske se granice dobivaju nakon nešto više od 2000 koraka poopćenog algoritma perceptrona



Poopćeni algoritam perceptrona

M razreda : $\omega_1, \omega_2, \dots, \omega_M$

Pretpostavimo da u k-tom koraku tijekom učenja uzorak $\vec{x}(k)$ pripada razredu ω_i.

Računamo M decizijskih funkcija.

Ako je

$$d_i[\vec{x}(k)] > d_j[\vec{x}(k)] \quad j = 1, 2, \dots, M; \; j \neq i$$

težinski vektor se ne ugađa:

$$\vec{w}_j(k+1) = \vec{w}_j(k), \quad j = 1, 2, \dots, M$$

Pretpostavimo da je za neki l

$$d_i[\vec{x}(k)] \leq d_l[\vec{x}(k)]$$

težinski vektori se sada ugađaju:

$$\vec{w}_i(k+1) = \vec{w}_i(k) + c\,\vec{x}(k)$$

$$\vec{w}_l(k+1) = \vec{w}_l(k) - c\,\vec{x}(k)$$

$$\vec{w}_j(k+1) = \vec{w}_j(k) \;; \quad j = 1, 2, \dots, M; \; j \neq i; \; j \neq l.$$

c - pozitivna konstanta

$\vec{w}_i(1)$ su proizvoljni početni vektori $i = 1, 2, \dots, M$.

Skalarno produktno jezgro
(Inner-Product Kernel)

– Kako naći optimalnu hiperravninu za linearno neodvojive razrede uporabom SVM?

Odgovor:

Uporabom sljedećih koraka:

1. Nelinearnim preslikavanjem izvornog (ulaznog) vektora u prostor značajki većih dimenzija;

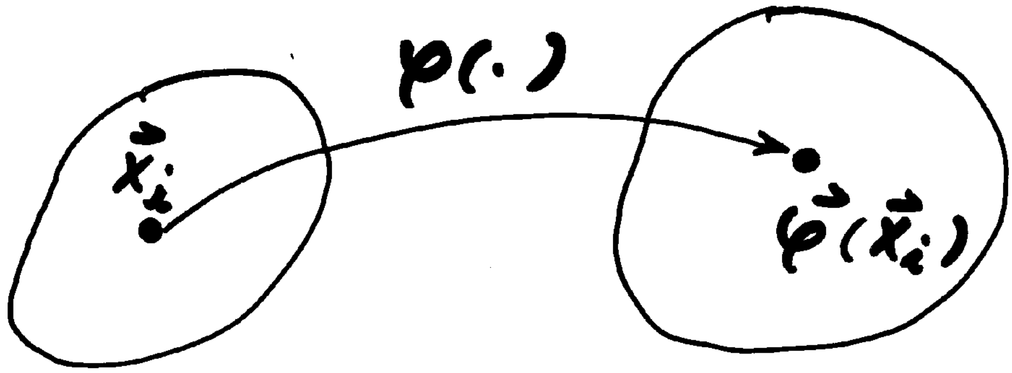
2. Konstrukcijom optimalne hiperravnine za odvajanje vektora značajki dobivenih u 1. koraku;

Opaska: Korak 1. temelji na Coverovom teoremu (1965. godina) → višedimenzionalni prostor može biti transformiran u novi prostor značajki u kojem su uzorci linearno separatibilni (s visokom vjerojatnosti) ako su zadovoljena dva uvjeta:

a) transformacija mora biti NELINEARNA;

b) dimenzionalnost prostora mora biti DOVOLJNO VELIKA;

1. KORAK



Ulazni prostor (izvorni prostor značajki)
dimenzija : m_0

Prostor značajki
dimenzija : m_1

$m_0 < m_1$

- Označimo s $\vec{x}$ vektor iz ulaznog prostora; dimenzija vektora neka je m_0
- Označimo $\{\varphi_j(\vec{x})\}_{j=1}^{m_1}$ skup nelinearnih transformacija iz ulaznog prostora u prostor značajki
- Pretpostavlja se da je $\varphi_j(\vec{x})$ definiran apriorno za sve j.

- za tako zadani skup nelinearnih transformacija možemo definirati decizijsku ravninu:

$$\sum_{j=1}^{m_1} w_j \varphi_j(\vec{x}) + b = 0 ,$$

gdje $\{w_j\}_{j=1}^{m_1}$ označava skup linearnih težina (težinskih koeficijenata) a b - pomaknuće.

Možemo zapisati:

$$\sum_{j=0}^{m_1} w_j \varphi_j(\vec{x}) = 0$$

gdje je $\varphi_0(\vec{x}) = 1$ za sve $\vec{X}$ tako da $\underline{w_0 \text{ označava } b.}$

Definirajmo vektor:

$$\vec{\varphi}(\vec{x}) = [\varphi_0(\vec{x}), \varphi_1(\vec{x}), \ldots, \varphi_{m_1}(\vec{x})]^T$$

gdje je $\varphi_0(\vec{x}) = 1$ za sve $\vec{x}$.

Vektor $\vec{\varphi}(\vec{x})$ predstavlja sliku "induciranu" u prostoru značajki zbog ulaznog vektora $\vec{x}$.

Vrijedi

$$\vec{w}^T \vec{\varphi}(\vec{x}) = 0$$

U prostoru značajki zahtijevamo "linearnu" separatibilnost !

SVM nam je dao rezultat:

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

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

gdje vektor $\vec{\varphi}(\vec{x}_i)$ odgovara transformiranom ulaznom vektoru $\vec{x}_i$ u i-tom slučaju (primjerku)

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

u $\vec{w}^T \vec{\varphi}(\vec{x}) = 0$!

Dobivamo:

$$\sum_{i=1}^{N} \lambda_i d_i \vec{\varphi}^T(\vec{x}_i) \cdot \vec{\varphi}(\vec{x}) = 0$$

$\boxed{\vec{\varphi}^T(\vec{x}_i) \cdot \vec{\varphi}(\vec{x})}$ predstavlja skalarni produkt dvaju vektora iz prostora značajki koji su "inducirani" ulaznim vektorima $\vec{x}$ i $\vec{x}_i$.

$$\boxed{K(\vec{x}, \vec{x}_i) = \vec{\varphi}^T(\vec{x}) \vec{\varphi}(\vec{x}_i)}$$

skalarno produktno jezgro

$$K(\vec{x}, \vec{x}_i) = \sum_{j=0}^{m_1} \varphi_j(\vec{x}) \varphi_j(\vec{x}_i)$$

za $i = 1, 2, \ldots, N$

Jezgro $K(\vec{x}, \vec{x}_i)$ je simetrična funkcija svojih argumenata

$$K(\vec{x}, \vec{x}_i) = K(\vec{x}_i, \vec{x}) \text{ za sve } i$$

Najvažnije:

Možemo koristiti jezgro $K(\vec{x}, \vec{x}_i)$ za konstrukciju optimalne hiperravnine u prostoru značajki:

Optimalna hiperravnina je definirana s:

$$\boxed{\sum_{i=1}^{N} \lambda_i d_i K(\vec{x}, \vec{x}_i) = 0}$$

Optimalni dizajn SVM

Jezgro $K(\vec{x}, \vec{x}_i)$ i njegova ekspanzija $\sum_{j=0}^{m_1} \varphi_j(\vec{x})\varphi_j(\vec{x}_i)$ nam omogućava konstrukciju decizijske ravnine koja je nelinearna u ulaznom (izvornom prostoru) značajki, ali je njena slika u (transformiranom) prostoru značajki linearna.

Dualni oblik za optimizaciju (SVM):

- zadan je skup uzoraka za učenje $\{\vec{x}_i, d_i\}_{i=1}^{N}$, nađi Lagrangeove multiplikatore $\{\lambda_i\}_{i=1}^{N}$ koji maksimiziraju funkciju:

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

$$\cdot K(\vec{x}_i, \vec{x}_j)$$

uz ograničenja:

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

$$(2) \quad 0 \leq \lambda_i \leq C \quad \text{za } i = 1, 2, \ldots N,$$

gdje je C user-specific pozitivni parametar

Kad nađemo optimalne vrijednosti Lagrangeovih multiplikatora $\lambda_{0,i}$ dobivamo $\vec{w}_0$

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

Prva komponenta vektora $\vec{w}_0$ — predstavlja pomaknuće

optimalno pomaknuće $\longrightarrow b_0$

Jezgro?

Marcer-ov teorem (1908):

Neka je $K(\vec{x}, \vec{x}')$ kontinuirano simetrično jezgro koje je definirano na zatvorenom intervalu $\vec{a} \leq \vec{x} \leq \vec{b}$, $\vec{a} \leq \vec{x}' \leq \vec{b}$.

Jezgro $K(\vec{x}, \vec{x}')$ može se razviti kao:

$$K(\vec{x}, \vec{x}') = \sum_{i=1}^{\infty} \alpha_i \varphi_i(\vec{x}) \varphi_i(\vec{x}')$$

gdje je $\alpha_i > 0$ za sve i.

Taj razvoj je valjan ako vrijedi:

$$\int_b^a \int_b^a K(\vec{x}, \vec{x}') \psi(\vec{x}) \psi(\vec{x}') \, d\vec{x} \, d\vec{x}' \geq 0$$

i za sve $\psi(\cdot)$ za koje vrijedi:

$$\int_b^a \psi^2(\vec{x}) \, d\vec{x} < \infty;$$

[$\varphi_i(\vec{x})$ - svojstvena funkcija (engl. eigenfunction);
α_i - svojstvena vrijednost (engl. eigenvalue);

- funkcije $\varphi_i(\vec{X})$ se zovu svojstvene funkcije

α_i - svojstvene vrijednosti

Jezgra:

$$K(\vec{X}, \vec{X_i}) \quad i = 1, 2, \ldots N$$

- polinomske funkcije

$$(\vec{X}^T \vec{X_i} + 1)^p$$

- radijalne bazne funkcije

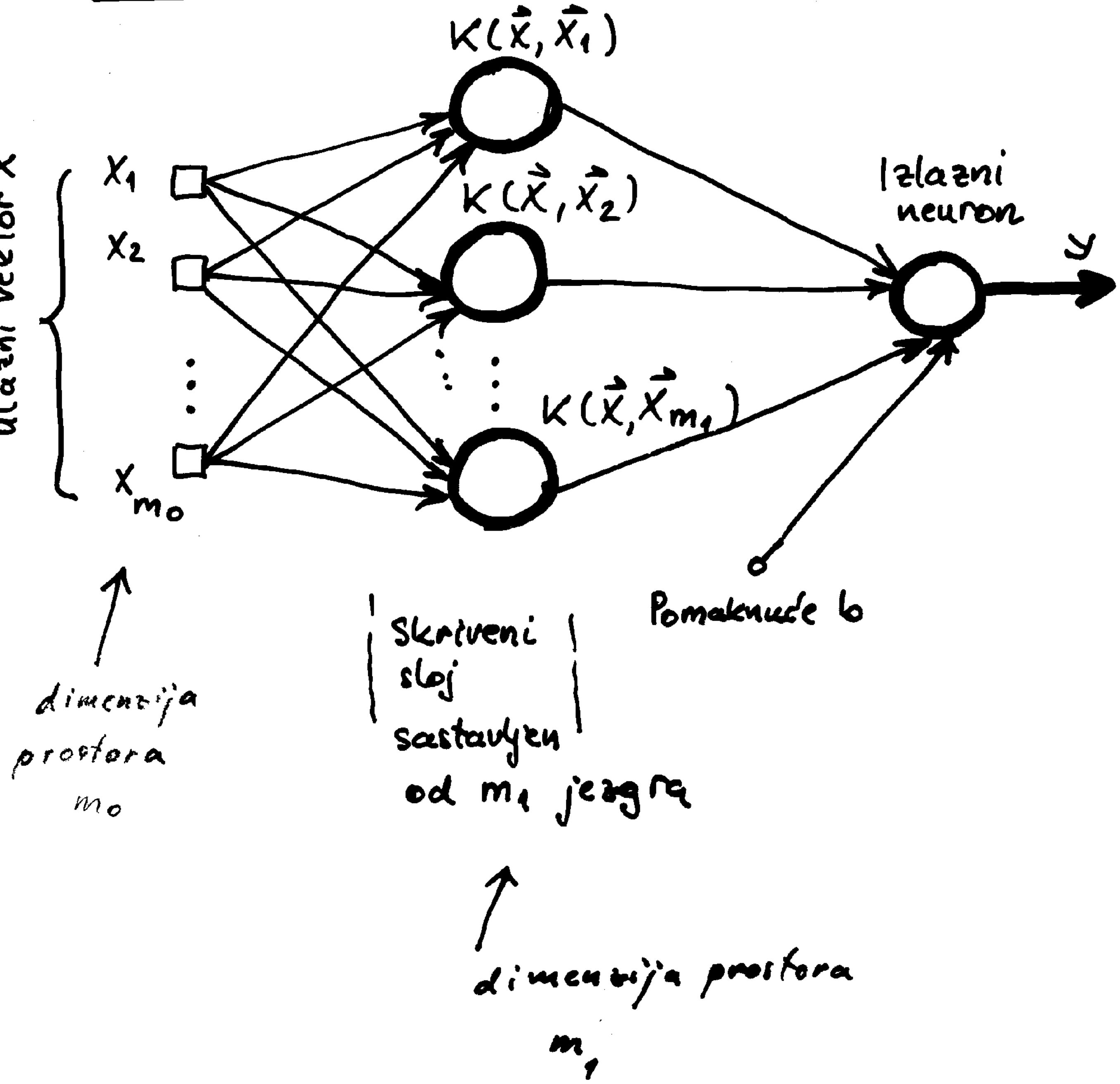
$$\exp\left(-\frac{1}{2\sigma^2} \| \vec{X} - \vec{X_i} \|^2\right)$$

- dvorazinski perceptron

$$\tanh(\beta_0 \vec{X}^T \vec{X_i} + \beta_1)$$

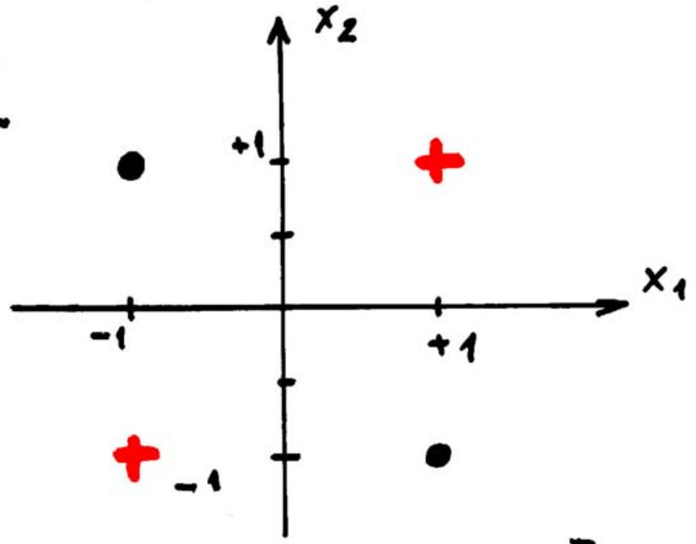
<u>Pozor</u>: za jezgro dvorazinskog perceptrona Marcerov teorem je zadovoljen za samo neke vrijednosti β_0 i β_1 !

Arhitektura SVM



Primjer: XOR (Exclusive OR)

Ulazni vektor $\vec{x}$	željeni odgovor d
$(-1, -1)$	-1
$(-1, +1)$	$+1$
$(+1, -1)$	$+1$
$(+1, +1)$	-1



$\bullet \in \omega_1$

$+ \in \omega_2$

$$\{([-1,-1]^T, d_1=-1), ([-1,+1]^T, d_2=+1), ([+1,-1]^T, d_3=+1), ([+1,+1]^T, d_4=-1)\}$$

Izaberimo jezgro $K(\vec{x}, \vec{x}_i) = (1 + \vec{x}^T \vec{x}_i)^2$

$$\vec{x} = [x_1, x_2]^T$$

$$\vec{x}_i = [x_{i1}, x_{i2}]^T$$

$$K(\vec{x}, \vec{x}_i) = \vec{\varphi}^T(\vec{x})\vec{\varphi}(\vec{x}_i)$$

Jezgro $K(\vec{x}, \vec{x}_i)$:

$$K(\vec{x}, \vec{x}_i) = 1 + x_1^2 x_{i1}^2 + 2 x_1 x_2 x_{i1} x_{i2} + x_2^2 x_{i2}^2 + 2 x_1 x_{i1} + 2 x_2 x_{i2}$$

Slika ulaznog vektora $\vec{x}$ u prostoru značajki je:

$$\vec{\varphi}(\vec{x}) = [1, x_1^2, \sqrt{2} x_1 x_2, x_2^2, \sqrt{2} x_1, \sqrt{2} x_2]^T$$

$$\vec{\varphi}(\vec{x}_i) = [1, x_{i1}^2, \sqrt{2}\, x_{i1} x_{i2}, x_{i2}^2, \sqrt{2}\, x_{i1}, \sqrt{2}\, x_{i2}]^T \quad i = 1, 2, 3, 4$$

$K(\vec{x}_i, \vec{x}_j)$ – javlja nam se u izrazu za SVM

(vidi Lpd 17-18)

$$\underline{\mathbf{K}} = \{K(\vec{x}_i, \vec{x}_j)\}_{(i,j)=1}^{N}$$

$K(\vec{x}_i, \vec{x}_j)$ možemo promatrati kao ij-ti element simetrične <u>matrice **K**</u>

$$\vec{\varphi}(\vec{x}_1) = [1, 1, \sqrt{2}, 1, -\sqrt{2}, -\sqrt{2}]^T$$

$\vec{x}_1 = (-1, -1)$

$$K(\vec{x}_1, \vec{x}_1) = \vec{\varphi}^T(\vec{x}_1) \cdot \varphi(\vec{x}_1)$$

$$K(\vec{x}_1, \vec{x}_1) = 9$$

Dualni oblik za optimizaciju (SVM):

$$\{\vec{X}_i, d_i\}_{i=1}^{N}$$

Treba naći Lagrangeove multiplikatore

$\{\lambda_i\}_{i=1}^{N}$ koji maksimiziraju funkciju

$$J(\vec{\lambda}) = \sum_{i=1}^{N} \lambda_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j d_i d_j K(\vec{X}_i, \vec{X}_j)$$

uz ograničenja:

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

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

C - pozitivna konstanta

Npr.

$$K(\vec{X}_1, \vec{X}_4)$$

$$\vec{\varphi}(\vec{X}_1) = [1, +1, \sqrt{2}, 1, -\sqrt{2}, -\sqrt{2}]^T$$

$$\vec{\varphi}(\vec{X}_4) = [1, 1, \sqrt{2}, 1, \sqrt{2}, \sqrt{2}]^T$$

$$\vec{X}_4 = [1, 1]^T$$

$$\vec{\varphi}^T(\vec{X}_1) \cdot \varphi(\vec{X}_4) = 1$$

$$\mathbf{K} = \begin{array}{c|c} i= & j= \;\; 1 \quad\; 2 \quad\; 3 \quad\; 4 \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} & \begin{bmatrix} 9 & 1 & 1 & 1 \\ 1 & 9 & 1 & 1 \\ 1 & 1 & 9 & 1 \\ 1 & 1 & 1 & 9 \end{bmatrix} \end{array}$$

Funkcija (za dualni problem):
(cfr. svm 17)

$$J(\vec{\lambda}) = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 -$$

$$- \frac{1}{2}\Big(9\lambda_1^2 - 2\lambda_1\lambda_2 - 2\lambda_1\lambda_3 + 2\lambda_1\lambda_4 + 9\lambda_2^2 + 2\lambda_2\lambda_3 - 2\lambda_2\lambda_4 + 9\lambda_3^2 - 2\lambda_3\lambda_4 + 9\lambda_4^2\Big)$$

Optimiziranje funkcije $J(\vec{\lambda})$ u odnosu na Lagrangeove multiplikatore daje skup jednadžbi:

$$9\lambda_1 - \lambda_2 - \lambda_3 + \lambda_4 = 1$$

$$-\lambda_1 + 9\lambda_2 + \lambda_3 - \lambda_4 = 1$$

$$-\lambda_1 + \lambda_2 + 9\lambda_3 - \lambda_4 = 1$$

$$\lambda_1 - \lambda_2 - \lambda_3 + 9\lambda_4 = 1$$

Optimalne vrijednosti Lagrangeovih multiplikatora su

$$\lambda_{0,1} = \lambda_{0,2} = \lambda_{0,3} = \lambda_{0,4} = \frac{1}{8}$$

- sva četiri vektora u EXOR problemu su potporni vektori

$$J_0(\vec{\lambda}) = \frac{1}{4}$$

$$\frac{1}{2} \|\vec{w}_0\|^2 = \frac{1}{4} \qquad \|\vec{w}_0\| = \frac{1}{\sqrt{2}}$$

Nadimo $\vec{W}_0$

$$\vec{W}_0 = \sum_{i=1}^{N} \lambda_{0,i} d_i \varphi(\vec{x}_i)$$

$$\vec{W}_0 = \frac{1}{8}\left[-\varphi(\vec{x}_1) + \varphi(\vec{x}_2) + \varphi(\vec{x}_3) - \varphi(\vec{x}_4)\right]$$

$$\vec{W}_0 = \frac{1}{8}\left[-\begin{bmatrix}1\\1\\\sqrt{2}\\1\\-\sqrt{2}\\-\sqrt{2}\end{bmatrix} + \begin{bmatrix}1\\1\\-\sqrt{2}\\1\\-\sqrt{2}\\\sqrt{2}\end{bmatrix} + \begin{bmatrix}1\\1\\-\sqrt{2}\\1\\\sqrt{2}\\-\sqrt{2}\end{bmatrix} - \begin{bmatrix}1\\1\\\sqrt{2}\\1\\\sqrt{2}\\\sqrt{2}\end{bmatrix}\right]$$

$$\vec{W}_0 = \begin{bmatrix}0\\0\\-1/\sqrt{2}\\0\\0\\0\end{bmatrix}$$

Pomaknuće

$b = w_0 = 0$

Optimalna hiperravnina je

$$\vec{W}_0^T \vec{\varphi}(\vec{x}) = 0$$

$$\left[0,\ 0,\ \frac{-1}{\sqrt{2}},\ 0,\ 0,\ 0\right]\begin{bmatrix}1\\ x_1^2\\ \sqrt{2}\,x_1x_2\\ x_2^2\\ \sqrt{2}\,x_1\\ \sqrt{2}\,x_2\end{bmatrix}=0$$

$$\boxed{-x_1x_2=0}$$

za $x_1 = x_2 = -1$ i $x_1 \cdot x_2 = +1$

izlaz $y = -1$

za $x_1 = -1$ i $x_2 = +1$ te za

$x_1 = +1$ i $x_2 = -1$

$y = +1$