1 Poopćene (linearne) decizijske funkcije

- složenost granica: linearne → <u>vrlo nelinearne!</u>
- vrlo nelinearne?

Rješenje:
Poopćeni oblik (linearne) decizijske funkcije:

$d(\vec{X}) = w_1 f_1(\vec{X}) + w_2 f_2(\vec{X}) + \ldots + w_k f_k(\vec{X}) + w_{k+1} = \sum_{i=1}^{k+1} w_i f_i(\vec{X})$

$\left\{ f_i(\vec{X}) \right\}$, $i = 1, 2, \ldots, k$, $f_{k+1} = 1$

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

$\vec{X}^* = \begin{bmatrix} f_1(\vec{X}) \\ f_2(\vec{X}) \\ \vdots \\ f_k(\vec{X}) \\ 1 \end{bmatrix} \quad d(\vec{X}) = \vec{W}^T \vec{X}^*$

$d(\vec{X})$ je linearna funkcija po $\vec{X}^*$
$d(\vec{X}) = \vec{W}^T \vec{X}^*$

$\vec{X}^* = \begin{bmatrix} f_1(\vec{X}) \\ f_2(\vec{X}) \\ \vdots \\ f_k(\vec{X}) \\ 1 \end{bmatrix} \quad \vec{W} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_k \\ w_{k+1} \end{bmatrix}$

Decizijska funkcija $d(\vec{X})$ može se promatrati kao linearna funkcija u k-dimenzionalnom prostoru!
Transformacija:
$\vec{X}_{\text{n-dimenzija}} \rightarrow \vec{X}^*_{\text{k-dimenzija}}$
Vrijedi: k>n
$d(\vec{X})$ se često naziva i "virtualno linearna"

Kako izabrati $\left\{ f_i(\vec{X}) \right\}_{i=1}^{k}$?
Mogućnost: $\left\{ f_i(\vec{X}) \right\}$ $\quad f_i(\vec{X})$ 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 $\vec{X} = (x_1, x_2)^T$
  $d(\vec{X}) = w_{11} x_1^2 + w_{12} x_1 x_2 + w_{22} x_2^2 + w_1 x_1 + w_2 x_2 + w_3$
  $d(\vec{X}) = \vec{W}^T \vec{X}^*$
  $\vec{X}^* = (x_1^2, x_1 x_2, x_2^2, x_1, x_2, 1)^T$

Opći slučaj kvadratnog oblika:

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

Kvadratnu decizijsku funkciju možemo promatrati kao linearnu ("linear by virtue") funkciju s $k = \frac{(n+1)(n+2)}{2}$ varijabli

$\left\{f_1(\vec{X}), \ldots, f_n(\vec{X})\right\} = \{x_1^2, x_2^2, \ldots, x_n^2\}$

$\left\{f_{n+1}(\vec{X}), \ldots, f_{2n}(\vec{X})\right\} = \{x_1, x_2, \ldots, x_n\}$

$\left\{f_{2n+1}(\vec{X}), \ldots, f_k(\vec{X})\right\} = \{x_1x_2, x_1x_3, \ldots, x_{n-1}x_n\}$

- Decizijske dunkcije višeg reda (polinomi stupnja r>2) možemo promatrati kao linearne funkcije $\frac{(n+r)!}{r!n!}$ varijabli

n - dimenzionalnost vektora $\vec{X}$

r - stupanj polinoma

Funkcije iz skupa $\left\{f_i(\vec{X})\right\}$ određene su s $f_i(\vec{X}) = x_{p_1}^{s_1}x_{p_2}^{s_2}\ldots x_{p_r}^{s_r}$

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

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

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

$$d^r(\vec{X}) = \left(\sum_{p_1=1}^{n}\sum_{p_2=p_1}^{n}\cdots\sum_{p_r=p_{r-1}}^{n} w_{p_1p_2\ldots p_r}x_{p_1}x_{p_2}\ldots x_{p_r}\right) + d^{r-1}(\vec{X}) \;;$$

$$d^0(\vec{X}) = w_{n+1}$$

Primjer: za r=2 i n=2:

$$d^2(\vec{X}) = \sum_{p_1=1}^{2}\sum_{p_2=p_1}^{2} w_{p_1p_2}x_{p_1}x_{p_2} + d^1(\vec{X})$$

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

$$d^1(\vec{X}) = w_1x_1 + w_2x_2 + w_3$$

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

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

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

Primjer: za r=3 i n=2:

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

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

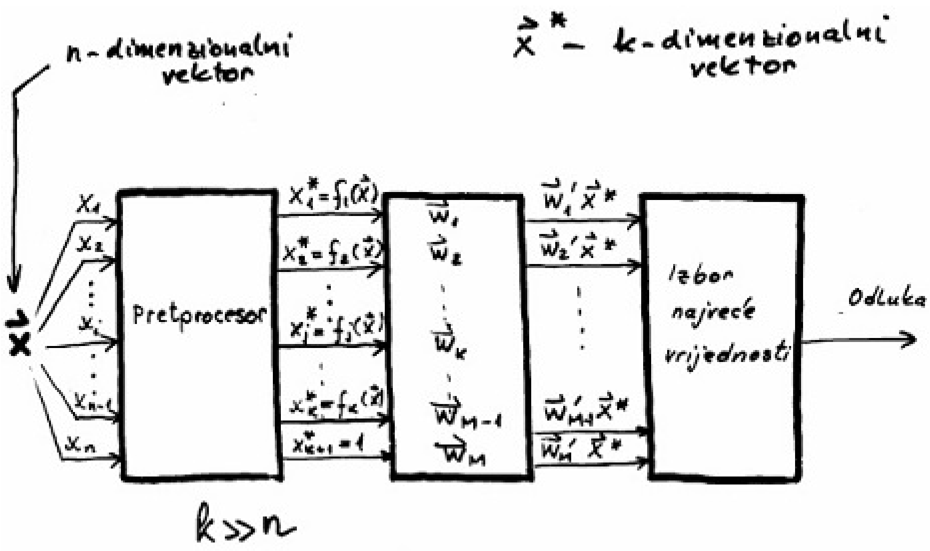
$$d^3(\vec{X}) = w_{111}x_1^3 + w_{112}x_1^2x_2 + w_{122}x_1x_2^2 + w_{222}x_2^3 + w_{11}x_1^2 + w_{12}x_1x_2 + w_{22}x_2^2 + w_1x_1 + w_2x_2 + w_3$$

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
$\vdots$							
9	10	55	220		5005		92378
10	1	66	286		8008		184756

Dimenzionalnost "lineariziranog prostora"!



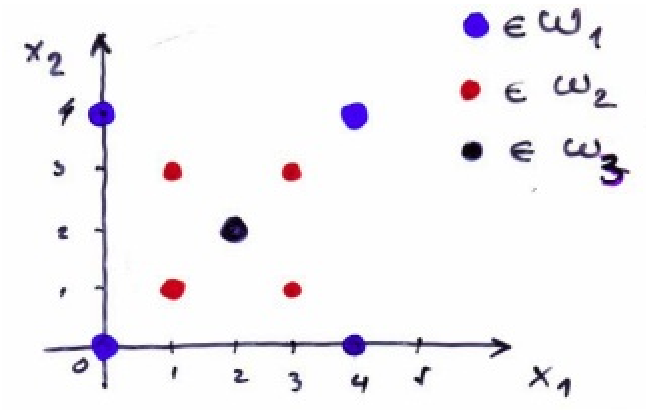
Slika 1: Blok shema sustava za raspoznavanje

<u>PRIMJER</u>
Zadan je skup uzoraka za učenje:

- $U_4^1 = \left\{(0,0)^T, (0,4)^T, (4,0)^T, (4,4)^T\right\}$
- $U_4^2 = \left\{(1,1)^T, (1,3)^T, (3,1)^T, (3,3)^T\right\}$
- $U_1^3 = \left\{(2,2)^T\right\}$

U_c^s — c-broj uzoraka u razredu, s-oznaka razreda

Zadane uzorke <u>ne možemo odijeliti linearnim</u> decizijskim funkcijama!
Pokušajmo <u>kvadratnim</u>! → r=2; n=2;
n=2

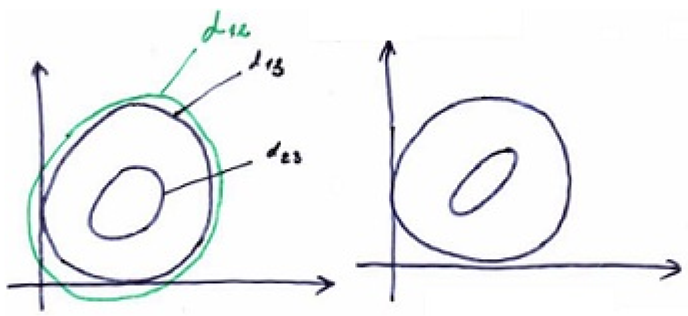


Slika 2: Prikaz uzoraka iz primjera

$k = \frac{(n+1)(n+2)}{2} = 6$ - dimenzionalni prostor
2-dim. $\rightarrow$ 6-dim.
Preslikavanje 2- u 6- dimenzionalni 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 skupa za učenje je :
$U_4^1 = \left\{(0,0,0,0,0,1)^T, (0,0,16,0,4,1)^T, (16,0,0,4,0,1)^T, (16,16,16,4,4,1)^T\right\}$
$U_4^2 = \left\{(1,1,1,1,1,1)^T, (1,3,9,1,3,1)^T, (9,3,1,3,1,1)^T, (9,9,9,3,3,1)^T\right\}$
$U_1^3 = \left\{(4,4,4,2,2,1)^T\right\}$

$$\vec{X}^* = \begin{bmatrix} x_1^2 \\ x_1 x_2 \\ x_2^2 \\ x_1 \\ x_2 \\ 1 \end{bmatrix} \qquad d(\vec{X}^*) = \vec{W}^T \vec{X}^* \qquad \vec{W} = ?$$

Upotrijebimo poopćeni algoritam perceptrona (M=3)
Dva od mogućih rezultata (u zavisnosti od izbora početnih vrijednosti):



Slika 3: Prikaz dobivenih granica

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

Poopćeni algoritam perceptrona

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

Pretpostavimo da u k-tom koraku tijekom učenja uzorak $\vec{X}(k)$ pripada razredu ω_i. Računamo M decizijskih funkcija.

Ako je $d_i[\vec{X}(k)] > d_j[\vec{X}(k)] \quad j = 1, 2, \ldots, M; \quad j \neq i$

težinski faktor se ne ugađa:

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

Pretpostavimo da je za neki l

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

težinski faktori se sada ugađaju:

$\vec{W}_i(k+1) = \vec{W}_i(k) + c\vec{X}(k)$

$\vec{W}_l(k+1) = \vec{W}_l(k) - c\vec{X}(k)$

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

c - pozitivna konstanta

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

2 Skalarno produktno jezgro (Inner-Product Kernel)

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

Odgovor: Uporabom sljedećih koraka:

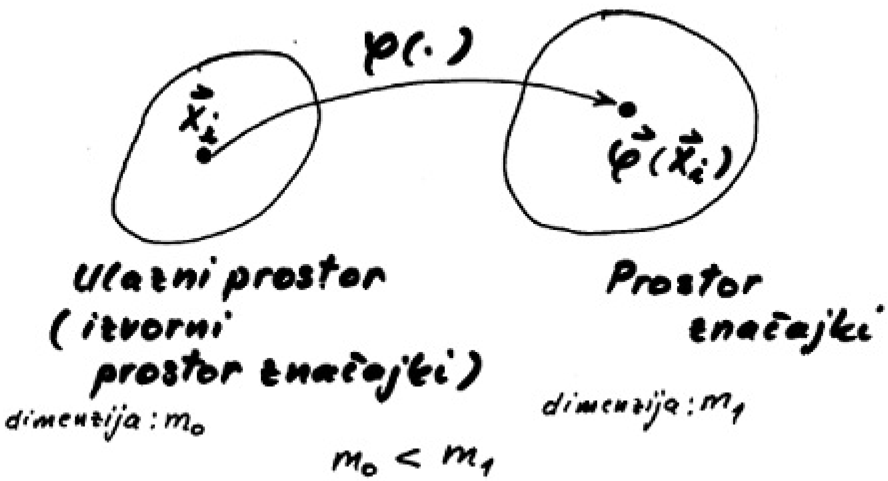
1. Nelinearnim preslikavanjem izvornog (ulaznog) vektora u prostor značajki većih dimenzija;

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

Opaska: Korak 1. temelji se na Coverovom teoremu (1965. godina) $\rightarrow$ 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;



Slika 4: Pretvorba uzoraka

- Označimo s $\vec{X}$ vektor iz ulaznog prostora; dimenzija vektora nema je m_0

- Označimo $\left\{\varphi_j(\vec{X})\right\}_{j=1}^{m_1}$ 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:
  $\sum\limits_{j=1}^{m_1} w_j\varphi_j(\vec{X}) + b = 0$, gdje $\{w_j\}_{j=1}^{m_1}$ označava skup linearnih težina (težinskih koeficijenata) a b-pomaknuće.
  Možemo zapisati:
  $\sum\limits_{j=0}^{m_1} w_j\varphi_j(\vec{X}) = 0$, 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:
$\vec{W} = \sum\limits_{i=1}^{N} \lambda_i d_i \vec{X}_i$, odnosno
$\vec{W} = \sum\limits_{i=1}^{N} \lambda_i d_i \vec{\varphi}(\vec{X}_i)$, gdje vektor $\vec{\varphi}(\vec{X}_i)$ odgovara transformiranom ulaznom vektoru $\vec{X}_i$ u i-tom slučaju (primjerku)
Uvrstimo $\vec{W} = \sum\limits_{i=1}^{N} \lambda_i d_i \vec{\varphi}(\vec{X}_i)$ u $\vec{W}^T\vec{\varphi}(\vec{X}) = 0$!
Dobivamo: $\sum\limits_{i=1}^{N} \lambda_i d_i \vec{\varphi}^T(\vec{X}_i)\vec{\varphi}(\vec{X}) = 0$
$\vec{\varphi}^T(\vec{X}_i)\vec{\varphi}(\vec{X})$ predstavlja skalarni produkt dvaju vektora iz prostora značajki koji su "inducirani" ulaznim vektorima $\vec{X}$ i $\vec{X}_i$.
$K(\vec{X}, \vec{X}_i) = \vec{\varphi}^T(\vec{X})\vec{\varphi}(\vec{X}_i) \leftarrow$ skalarno produktno jezgro
$K(\vec{X}, \vec{X}_i) = \sum\limits_{j=0}^{m_1} \varphi_j(\vec{X})\varphi_j(\vec{X}_i), \quad$ za $i = 1, 2, \ldots, N$
Jezgro $K(\vec{X}, \vec{X}_i)$ je simetrična funkcija svojih argumenata $K(\vec{X}, \vec{X}_i) = K(\vec{X}_i, \vec{X})$ 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 :
$\sum\limits_{i=1}^{N} \lambda_i d_i K(\vec{X}, \vec{X}_i) = 0$

Optimalni dizajn SVM

Jezgro $K(\vec{X}, \vec{X}_i)$ i njegova ekspanzija $\sum\limits_{j=0}^{m_1} \varphi_j(\vec{X})\varphi_j(\vec{X}_i)$ 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 $\left\{\vec{X}_i, d_i\right\}_{i=1}^{N}$, nađi Lagrangeove multiplikatore $\{\lambda_i\}_{i=1}^{N}$ koji maksimiziraju funkciju:
$J(\vec{\lambda}) = \sum\limits_{i=1}^{N} \lambda_i - \frac{1}{2} \sum\limits_{i=1}^{N} \sum\limits_{j=1}^{N} \lambda_i \lambda_j d_i d_j K(\vec{X}_i, \vec{X}_j)$
uz ograničenja:
(1) $\sum\limits_{i=1}^{N} \lambda_i d_i = 0$
(2) $0 \leq \lambda_i \leq c \qquad$ 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 $b_0 \leftarrow$ optimalno pomaknuće

Jezgro?

Marcer-ov teorem(1908):

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

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

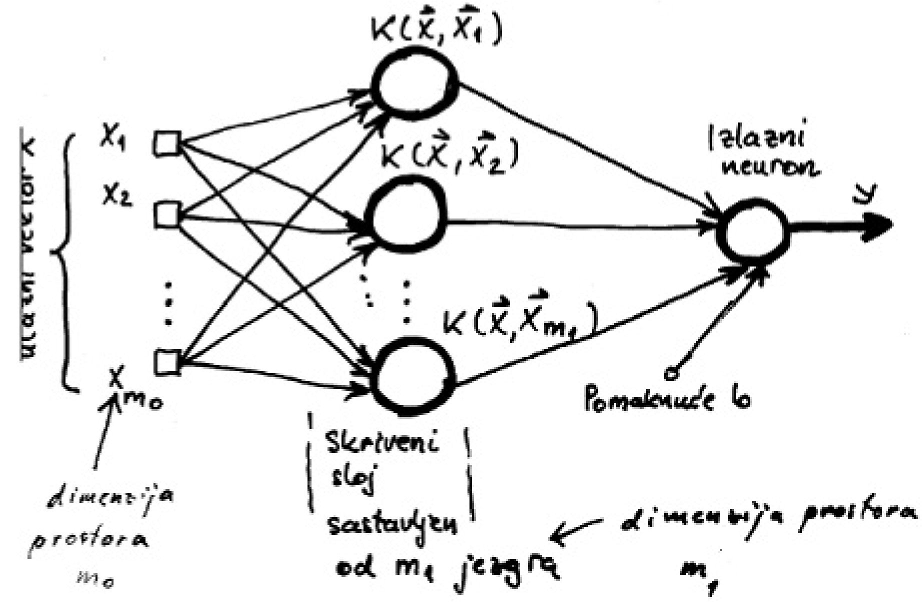
$K(\vec{X}, \vec{X}') = \sum_{i=1}^{\infty} \alpha_i \varphi_i(\vec{X}) \varphi_i(\vec{X}')$ gdje je $\alpha_i > 0$ za sve i.

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);

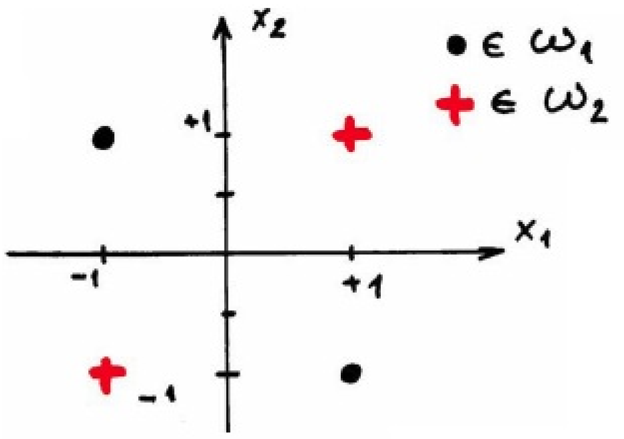
- funkcije $\varphi_i(\vec{X})$ se zovu svojstvene funkcije
  α_i - svojstvene vrijednosti
  Jezgra: $K(\vec{X}, \vec{X}_i)$ $i = 1, 2, \ldots, N$
- polinomske funkcije
  $(\vec{X}^T \vec{X}_i + 1)^p$
- radijalne bazne funkcije
  $\exp(-\frac{1}{2\sigma^2} \left\| \vec{X} - \vec{X}_i \right\|^2)$
- dvorazinski perceptron
  $\tanh(\beta_0 \vec{X}^T \vec{X}_i + \beta_1)$
  qunderlinePozor: za jezgro dvorazinskog perceptrona Marcerov teorem je zadovoljen za samo neke vrijednosti β_0 i β_1!



Slika 5: 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



Slika 6: Uzorci za primjer

$\left\{([-1,-1]^T, d_1=-1), ([-1,+1]^T, d_2=+1), ([+1,-1]^T, d_3=+1), ([+1,+1]^T, d_4=-1)\right\}$
Izaberimo jezgro $K(\vec{X},\vec{X}_i) = (1+\vec{X}^T\vec{X}_i)^2$
$\vec{X} = [x_1, x_2]^T$
$\vec{X}i = [x_{i1}, x_{i2}]^T$
Jezgro $K(\vec{X},\vec{X}_i) = 1 + x_1^2 x_{i1}^2 + 2x_1x_2x_{i1}x_{i2} + x_2^2x_{i2}^2 + 2x_1x_{i1} + 2x_2x_{i2}$
Slika ulaznog vektora $\vec{X}$ u prostoru značajki je :
$\vec{\varphi}(\vec{X}) = [1, x_1^2, \sqrt{2}x_1x_2, x_2^2, \sqrt{2}x_1, \sqrt{2}x_2]^T$
$\vec{\varphi}(\vec{X}_i) = [1, x_{i1}^2, \sqrt{2}x_{i1}x_{i2}, x_{i2}^2, \sqrt{2}x_{i1}, \sqrt{2}x_{i2}]^T \qquad i = 1,2,3,4$
$K(\vec{X}_i,\vec{X}_j)$ - javlja nam se u izrazu za SVM
$\mathbf{K} = \left\{K(\vec{X}_i,\vec{X}_j)\right\}_{(i,j)=1}^{N}$
$K(\vec{X}_i,\vec{X}_j)$ možemo promatrati kao ij-ti element simetrične matrice $\mathbf{K}$
$\vec{\varphi}(\vec{X}_1) = [1, 1, \sqrt{2}, 1, -\sqrt{2}, -\sqrt{2}]^T$
$\vec{X}_1 = (-1,-1) \qquad K(\vec{X}_1,\vec{X}_1) = \vec{\varphi}^T(\vec{X}_1)\vec{\varphi}(\vec{X}_1)$
$K(\vec{X}_1,\vec{X}_1) = 9$
Dualni oblik za optimizaciju (SVM):
$\left\{\vec{X}_i, d_i\right\}_{i=1}^{N}$
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) $\sum_{i=1}^{N}\lambda_i d_i = 0$

(2) $0 \leq \lambda_i \leq c \quad$ za $i = 1, 2, \ldots, N$, c - pozitivna konstanta
Npr.
$K(\vec{X_1}, \vec{X_4})$
$\vec{\varphi}(\vec{X_1}) = [1, 1, \sqrt{2}, 1, -\sqrt{2}, -\sqrt{2}]^T$
$\vec{\varphi}(\vec{X_4}) = [1, 1, \sqrt{2}, 1, \sqrt{2}, \sqrt{2}]^T$
$\vec{X_4} = [1, 1]^T$
$\vec{\varphi}^T(\vec{X_1})\vec{\varphi}(\vec{X_4}) = 1$

$$\mathbf{K} = \begin{bmatrix} 9 & 1 & 1 & 1 \\ 1 & 9 & 1 & 1 \\ 1 & 1 & 9 & 1 \\ 1 & 1 & 1 & 9 \end{bmatrix}$$

Funkcija (za dualni problem):
$J(\vec{\lambda}) = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 - \frac{1}{2}(9\lambda_1^2 - 2\lambda_1\lambda_2 - 2\lambda_1\lambda_3 + 2\lambda_1\lambda_4 + 9\lambda_2^2 + 2\lambda_2\lambda_3 - 2\lambda_2\lambda_4 + 9\lambda_3^2 - 2\lambda_3\lambda_4 + 9\lambda_4^2)$
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 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
$J_0(\vec{\lambda}) = \frac{1}{4}$
$\frac{1}{2}\left\|\vec{W_0}\right\|^2 = \frac{1}{4} \quad \rightarrow \quad \left\|\vec{W_0}\right\| = \frac{1}{\sqrt{2}}$
Nađimo $\vec{W_0}$
$\vec{W_0} = \sum_{i=1}^{N} \lambda_{0,i} d_i \varphi(\vec{X_i})$
$\vec{W_0} = \frac{1}{8}\left[-\varphi(\vec{X_1}) + \varphi(\vec{X_2}) + \varphi(\vec{X_3}) - \varphi(\vec{X_4})\right]$

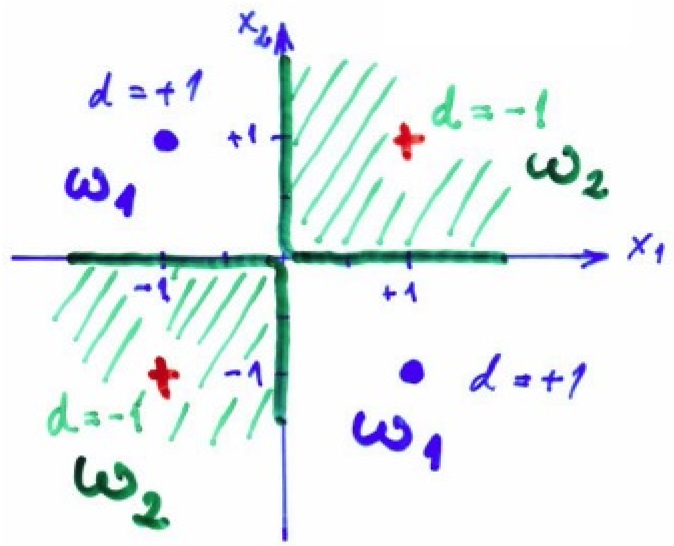
$$\vec{W_0} = \frac{1}{8}\left[-\begin{bmatrix} 1 \\ 1 \\ \sqrt{2} \\ 1 \\ -\sqrt{2} \\ -\sqrt{2} \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ -\sqrt{2} \\ 1 \\ -\sqrt{2} \\ \sqrt{2} \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ -\sqrt{2} \\ 1 \\ \sqrt{2} \\ -\sqrt{2} \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ \sqrt{2} \\ 1 \\ \sqrt{2} \\ \sqrt{2} \end{bmatrix}\right]$$

$$\vec{W_0} = \begin{bmatrix} 0 \\ 0 \\ -\frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad \text{Pomaknuće } b = w_0 = 0$$

Optimalna hiperravnina je $\vec{W_0}^T \vec{\varphi}(\vec{X}) = 0$

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

$-x_1x_2 = 0$
za $x_1 = x_2 = -1$ i $x_1 = x_2 = 1$, izlaz y=-1
za $x_1 = -1$ i $x_2 = 1$ te za $x_1 = 1$ i $x_2 = -1$, y=1



Slika 7: Prikaz rješenja