(5. DZ)

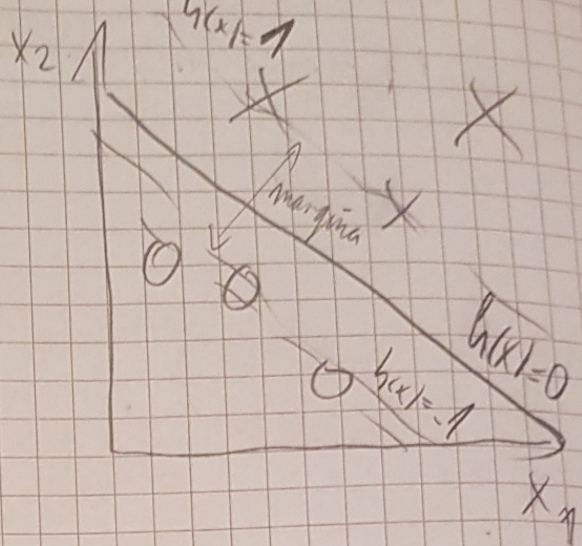
1)

a) Problem max. margine

$h(x) = w^T \phi(x) + w_0$

$y \in \{-1, 1\}$

$y = sgn(h(x))$



Ako su $\mathcal{D}$ linearno odvojivi, onda:

$$\forall (x^i, y^i) \in \mathcal{D} \Rightarrow y^i h(x^i) \geq 0$$

Prostor rješenja je beskonačan, mi tražimo rješenje max margine

Predznačna udaljenost: $d = \frac{h(x)}{\|w\|}$

Udaljenost: $\frac{y^i h(x^i)}{\|w\|}$

Margina: (udaljenost od najbližeg primjera)

$$\frac{1}{\|w\|} \min_i \{y^i h(x^i)\}$$

Mi tražimo max. marginu za w:

$$\underset{w}{\operatorname{argmax}} \left\{ \frac{1}{\|w\|} \min_i \{y^i h(x^i)\} \right\}$$

Preoblikovanje problema optimizacije:

x^i koji je najbliži:

$$y^i(w^T\phi(x)+w_0)=1$$

Za ove vrijedi onda:

$$y^i h(x^i) \geq 1$$

Znači da je margina široka $\frac{2}{\|w\|}$

Sada optimiramo:

$$\underset{w,w_0}{\operatorname{argmax}} \frac{1}{\|w\|},$$

OGRANIČENJE:

$$y^i h(x^i) \geq 1, \quad i \in 1,N$$

EKVIVALENTNO

$$\underset{w,w_0}{\operatorname{argmin}} \frac{1}{2}\|w\|^2$$

b) LAGRANŽIRANJE:

$$L(w,w_0,\alpha)=\frac{1}{2}\|w\|^2-\sum_{i=1}^{N}\alpha_i\{y^i(w^T\phi(x^i)+w_0)-1\}$$

PRIM. PROB.

Rješenje: sedlo koje je minimalno obzirom na w i w_0, a maksimalno obzirom na α.

$$\frac{\partial L}{\partial w}=-\sum\alpha_i y_i\phi(x^i)+w=0$$

$$w=\sum\alpha_i y_i \phi(x^i)$$

$$0=\sum\alpha_i y_i$$

DUALNA:

$$L(\alpha) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{N} \alpha_i \{ y^i (w^T \phi(x^i) + w_0) - 1 \}$$

$$= \frac{1}{2}\|w\|^2 - \sum_{i=1}^{N} \alpha_i (y^i (w^T \phi(x^i) + w_0)) + \sum \alpha_i$$

$$= \frac{1}{2}\|w\|^2 - \sum_{i=1}^{N} \alpha_i y^i w^T \phi(x^i) - \underbrace{\sum_{i=1}^{N} \alpha_i y_i w_0}_{0} + \sum \alpha_i$$

$$= \frac{1}{2}\|w\|^2 - \sum_{i}^{N} \alpha_i y^i \phi(x^i) \sum_{j}^{N} \alpha_j y_j \phi(x^j) + \sum \alpha_i$$

$$= \frac{1}{2} \sum_{i}^{N} \alpha_i y_i \phi(x^i) \cdot \sum_{j}^{N} \alpha_j y_j \phi(x^j) - \sum_{i}^{N} \alpha_i y^i \phi(x^i) \sum_{j}^{N} \alpha_j y_j \phi(x^j) + \sum_{i}^{N} \alpha_i$$

$$= \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i}^{N} \sum_{j}^{N} \alpha_i \alpha_j y^i y^j \phi(x^i)^T \phi(x^j)$$

DUALNI:

$\max(L)$

uz:

$\alpha_i \geq 0, \quad i = 1 \ldots N$

$\sum_{i}^{N} \alpha_i y^i = 0$

KKT:

$\alpha_i \geq 0$

$y^i h(x^i) - 1 \geq 0$

$\alpha_i (y^i h(x^i) - 1) = 0 \Rightarrow$ POTPORNI VEKTORI

c) ~~Mijem~~ Mijenja se broj varijabli ciljne funkcije

U Prim. $N+1$ var, a dual N var

ako $n \gg N$, transformacija se isplati

d)

PRIM.

$$\boxed{L(w, w_0, \alpha) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{N} \alpha_i \{y^i (w^T \phi(x^i) + w_0) - 1\}}$$

DUAL:

$$\boxed{\tilde{L}(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j \phi(x^i)^T \phi(x^j)}$$

~~e) Neparametarski je model jer~~

PRIM. MODEL

$$h(x) = w^T \phi(x) + w_0$$

DUAL MODEL

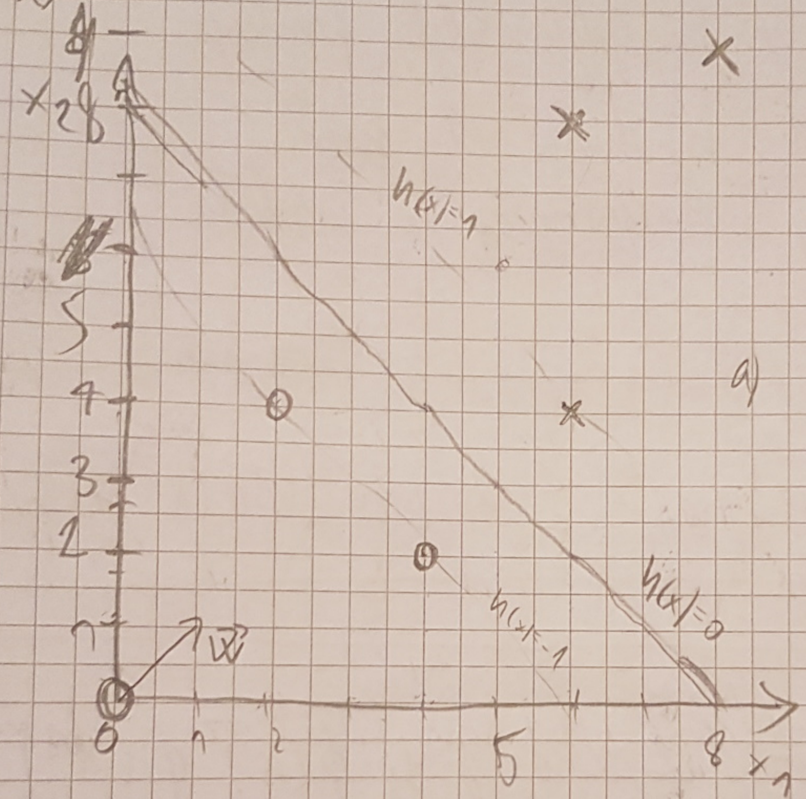
$$h(x) = \sum_{i=1}^{N} \alpha_i y_i \phi(x)^T \phi(x^i) + w_0$$

e) Dualni SVM je neparametarski model jer ne koristimo parametre w, nego α. Broj parametara ovisi o broju primjera za učenje.

2.

1

a)



~~$h(x) = -x+7$~~

$$\boxed{h(x) = x_1 + x_2 - 8}$$

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

$$w_0 = -8$$

b) $d = \frac{2}{\|w\|} = 2$

$\alpha^T = [0 \; 0 \; \tfrac{1}{2} \; \tfrac{1}{2} \; 0 \; 0]$

c)

$$w = \sum_{i=1}^{6} \alpha_i y^i x^i$$

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = -\alpha_3 \begin{bmatrix} 4 \\ 2 \end{bmatrix} + \alpha_4 \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

$$\boxed{\alpha_3 = \tfrac{1}{2}, \; \alpha_4 = \tfrac{1}{2}}$$

$$\alpha_3 = \frac{2-1}{2} = \frac{1}{2}$$

$$1 = -4\alpha_3 + 6\alpha_4$$

$$1 = -2\alpha_3 + 4\alpha_4$$

$$2\alpha_3 = 4\alpha_4 - 1$$

$$1 = -4\,\frac{4\alpha_4 - 1}{2} + 6\alpha_4$$

$$1 = -8\alpha_4 + 2 + 6\alpha_4$$

$$\boxed{\alpha_4 = \tfrac{1}{2}}$$

$$2\alpha_4 = 1$$

d) $\cancel{w_0 = \frac{1}{|S|} \sum_{i \in S} \left(y^i - \sum_j \alpha_j y^j \phi(x^j)^T \phi(x^i) \right)}$

$\cancel{= \frac{1}{2} \sum_i^2 \left(y^i - \sum_i \alpha_i y^i \right.}$

e) $x^7 = (5, 6)$

$$h(x^7) = w^T x + w_0 = \sum_{i=1}^{N} \alpha_i y_i x^T x^i + w_0$$

$$= -\frac{1}{2} [5 \ 6] \begin{bmatrix} 4 \\ 2 \end{bmatrix} + \frac{1}{2} [5 \ 6] \begin{bmatrix} 6 \\ 4 \end{bmatrix} - 8$$

$$= -\frac{1}{2} 32 + \frac{1}{2} 56 - 8 = -16 + 28 - 8$$

$$= -16 + 20 \boxed{= 4} > 0$$

25

28

x^7 je POZITIVAN PRIMJER!

3.

~~Problem~~ Klase su najčešće linearno neodvojive. Možemo ih preslikati u višu dimenziju tako da možda postanu lin. odvojivi, ali onda riskiramo prenaučenost.
Rješenje je da proširimo SVM da ponekad dopušta pogrešne klasifikacije, pomoću meke margine.

x_2

b)

$$y^i(w^T\phi(x^i)+w_0)\geq 1-\xi_i$$

→ UDALJENOST OD PRAVE STRANE

uz $\xi_i \geq 0$

$$\underset{w,w_0}{\text{argmin}} \left\{ \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}\xi_i \right\}$$

$y^i h(x^i) \geq 1-\xi_i$

L2 — kazna krivih primjera

$$L(w,w_0,\xi,\alpha,\beta)=\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}\xi_i - \sum_{i}^{N}\alpha_i(y_i(w^T\phi(x^i))-1+\xi_i)-\sum_{i}^{N}\beta\xi_i$$

$\alpha_i \geq 0$ $\quad i=1 \dots N$

$\beta_i \geq 0$ $\quad y^i h(x^i) \geq 1-\xi_i$

$\xi_i \geq 0$

c)

$$\tilde{L}(\alpha)=\sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum_{i}^{N}\sum_{j}^{N}\alpha_i\alpha_j y_i y_j \phi(x^i)^T\phi(x^j)$$

$\alpha_i \geq 0$

$\beta_i \geq 0$

$\alpha_i \leq C$

$$\boxed{\begin{array}{c} 0\leq \alpha_i \leq C \\ \sum_{i}^{N}\alpha_i y^i = 0 \end{array}}$$

OGRANIČENJA

4. CILJNA FJA JE $\frac{1}{2}\|w\|^2 + C\sum_{i}^{N} \xi_i$

a) Za Primjere na ispravnoj strani vrijedi

$$y^i h(x^i) \geq 1$$

i njih ne kažnjavamo.

Primjere s krive strane kažnjavamo $\xi_i = |y^i - h(x^i)|$

$$\xi_i = 1 - y^i h(x^i)$$

$$L(h(x), y) = \max(0, 1 - y h(x)) \quad \text{HINGE LOSS}$$

$$\boxed{E(h/D) = \sum_{i=1}^{N} \max(0, 1 - y^i h(x^i)) + \frac{\lambda}{2}\|w\|^2}$$

b) $L = \cancel{\max}\ 1 - y h(x) = 1 + 4 = \boxed{5}$