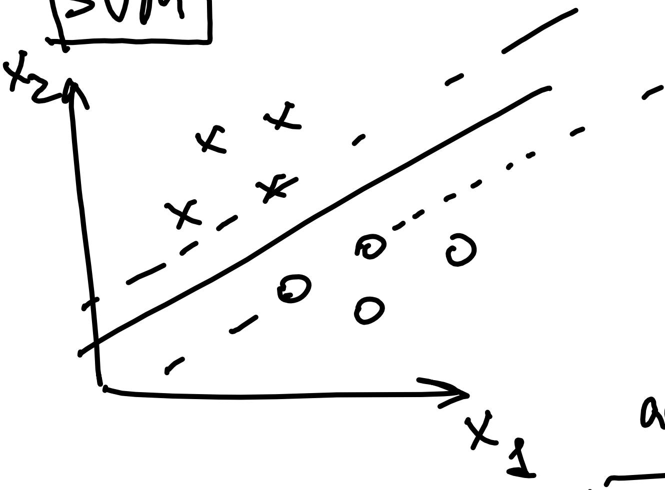
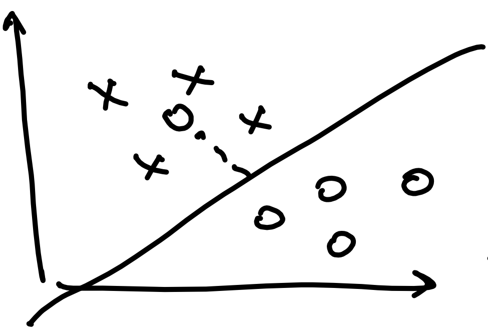
SVM



Hard Margin

$$\begin{cases} \frac{1}{2}\|w\|^2 \to \min\limits_{w,b} \\ y_i(\langle w, x_i\rangle + b) \geq 1 \end{cases}$$

$$a(x) = \operatorname{sign}(\langle w, x\rangle + b)$$



Soft Margin

$$\begin{cases} \frac{1}{2}\|w\|^2 + C \cdot \sum\limits_{i=1}^{n} \xi_i \to \min\limits_{w,b,\xi_i} \\ y_i(\langle w, x_i\rangle + b) \geq 1 - \xi_i \\ \xi_i \geq 0 \end{cases}$$

$$L = \frac{1}{2}\langle w, w\rangle + C \cdot \sum_i \xi_i - \sum_i \lambda_i \big(y_i(\langle w, x_i\rangle + b) -$$

$$- 1 + \xi_i\big) - \sum_i \mu_i \xi_i$$

$$\left\{ \sum_i \lambda_i y_i \langle w, x_i\rangle = \langle w, \sum_i \lambda_i y_i x_i\rangle \right\}.$$

K.K.T:

$$
\begin{cases}
\nabla_w L = w - \sum_i \lambda_i y_i x_i = 0 \Rightarrow \boxed{w^* = \sum_{i=1}^{n} \lambda_i y_i x_i} \\
L'_b = \sum_i \lambda_i y_i = 0 \\
L'_{\xi_i} = C - \lambda_i - \mu_i = 0 \Rightarrow \boxed{\lambda_i + \mu_i = C.} \\
\lambda_i = 0 \text{ или } y_i(\langle w, x_i \rangle + b) = 1 - \xi_i \\
\mu_i = 0 \text{ или } \xi_i = 0 \\
\lambda_i \geqslant 0,\ \mu_i \geqslant 0,\ \xi_i \geqslant 0
\end{cases}
$$

$$
g(\lambda) = \inf_w L = \frac{1}{2} \langle \sum_i \lambda_i y_i x_i, \sum_j \lambda_j y_j x_j \rangle +
$$

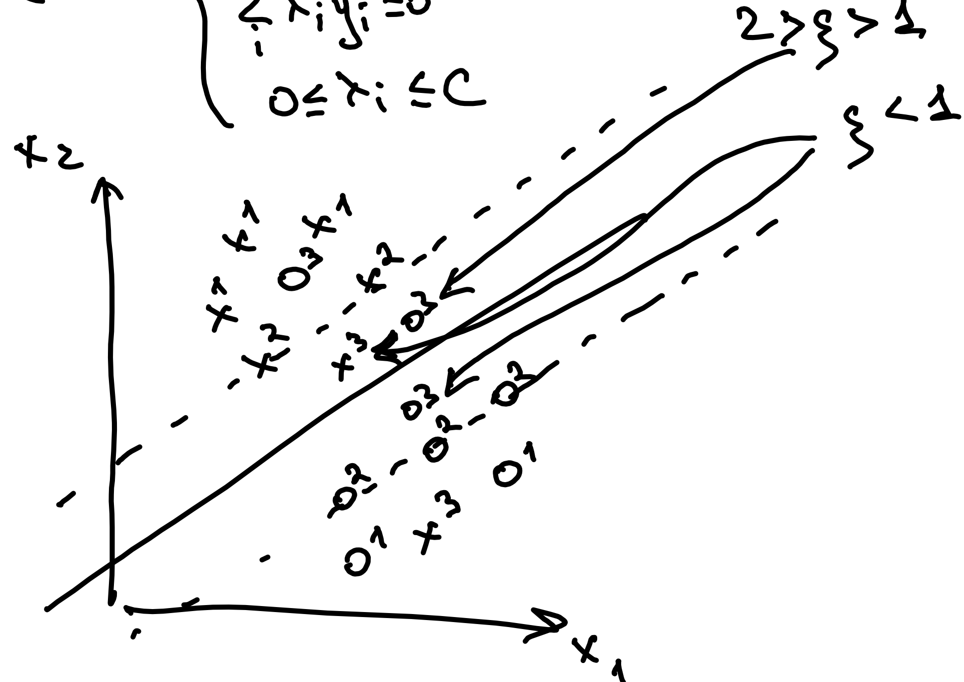
$$
+ C \cdot \sum_i \xi_i - \langle \sum_i \lambda_i y_i x_i, \sum_j y_j \lambda_j x_j \rangle - b \sum_i \lambda_i y_i
$$

$$
+ \sum_i \lambda_i - \sum_i \lambda_i \xi_i - \sum_i \mu_i \xi_i =
$$

$$
= -\frac{1}{2} \sum_i \sum_j \lambda_i \lambda_j y_i y_j \langle x_i, x_j \rangle + \sum_i \xi_i \underbrace{(C - \lambda_i - \mu_i)}_{0}
$$

$$
- b \underbrace{\sum_i \lambda_i y_i}_{0} + \sum_i \lambda_i
$$

$$(D):\begin{cases} g(\lambda) = -\frac{1}{2}\sum_i\sum_j \lambda_i\lambda_j y_i y_j \langle x_i, x_j\rangle + \sum_i \lambda_i \to \max\limits_{\lambda} \\ \sum_i \lambda_i y_i = 0 \\ 0 \le \lambda_i \le C \end{cases}$$

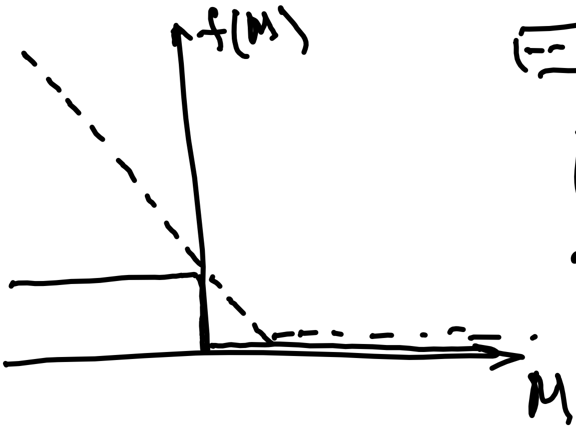


① - эталонные ($\lambda = 0$, $\xi = 0$)

② - опорные ($\xi = 0 \Rightarrow \mu > 0$, $0 < \lambda < C$)

③ - нарушители ($\xi > 0 \Rightarrow \mu = 0$, $\lambda = C$)

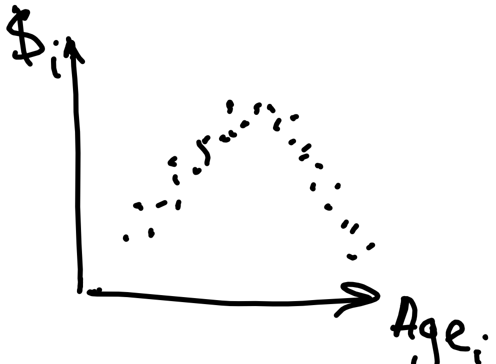
$$Q = \frac{1}{n}\sum_i \left[a(x_i) \neq y_i\right]$$

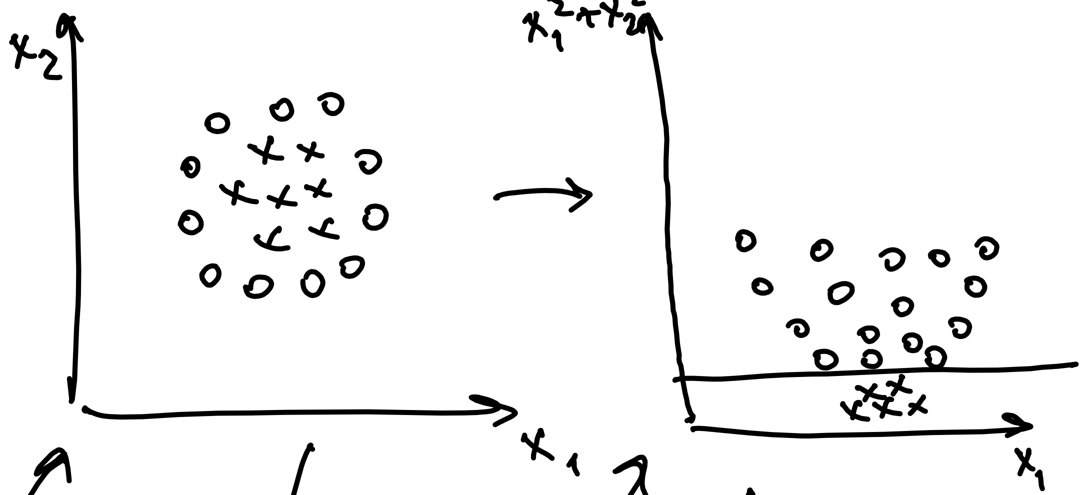


- - - $\max(0, 1-M)$

$$\begin{cases} \frac{1}{2}\|w\|^2 + C\cdot\sum_i \xi_i \to \min\limits_{w, b, \xi_i} \\ y_i\left(\langle w, x_i\rangle + b\right) \geq 1 - \xi_i \\ \xi_i \geq 0 \end{cases}$$

$$\frac{1}{2}\|w\|^2 + C\cdot\sum_i \underbrace{\max\left(0, 1 - y_i\left(\langle w, x_i\rangle + b\right)\right)}_{\text{Hinge-Loss.}}$$





$x_1^2+x_2^2$

$\mathbb{X}$ $\quad$ $\mathrm{sign}(\langle w, x\rangle)$ $\quad$ $\mathbb{H}$ $\quad$ $\mathrm{sign}(\langle w, \varphi(x)\rangle)$

$$\varphi : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \to \begin{pmatrix} x_1 \\ x_1^2 + x_2^2 \end{pmatrix}$$

$\varphi : \mathbb{X} \to \mathbb{H}$ (спрямляющее пр-во)

$$\boxed{k(x, z) = \langle \varphi(x), \varphi(z) \rangle}$$ – ядро.

$$a(x) = \mathrm{sign}(\langle w, \varphi(x)\rangle)$$

$$w = \sum_i \lambda_i y_i \varphi(x_i)$$

$$(D): \begin{cases} -\frac{1}{2} \sum_i \sum_j \lambda_i \lambda_j y_i y_j \overbrace{\langle \varphi(x_i), \varphi(x_j)\rangle}^{k(x_i, x_j)} \\ \ldots\ldots \end{cases}$$

$$\varphi: \underset{\in \mathbb{R}^d}{\begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix}} \to \underset{\in \mathbb{R}^{d^2}}{(x_1 \cdot x_1, x_1 \cdot x_2, \dots, x_1 \cdot x_d, x_2 \cdot x_1, \dots, x_d^2)}$$

$$K(x, z) = \langle \varphi(x), \varphi(z) \rangle =$$

$$= \sum_i \sum_j x_i x_j \cdot z_i z_j = \sum_i \left[x_i z_i \left(\sum_j x_j z_j \right) \right] \circleddash$$

$$\left\{ x_1 \cdot x_1 \cdot z_1 \cdot z_1 + x_1 \cdot x_2 \cdot z_1 \cdot z_2 + x_2 x_1 z_2 z_1 + x_2 x_2 z_2 z_2 = x_1 z_1 (\underbrace{x_1 z_1 + x_2 z_2}) + x_2 z_2 (\underbrace{x_1 z_1 + x_2 z_2}) \right.$$

$$\circleddash \underbrace{\left(\sum_j x_j z_j \right)}_{\langle x, z \rangle} \underbrace{\left(\sum_i x_i z_i \right)}_{\langle x, z \rangle} = \boxed{\langle x, z \rangle^2}$$

<u>Теорема Мерсера:</u>

$K(x, z)$ – ядро, если 1) $K(x, z) = K(z, x)$

2) $K(x, z)$ – неотр. опр $\Rightarrow$

$\forall$ конечной выборки: $\left(k(x_i, x_j)\right)_{i,j=1}^{n}$ — неотр. опр.

$\boxed{\text{Св-ва ядер:}}$ ($\alpha > 0$, k_1, k_2 — ядра)

1) $k_1 + k_2$ — ядро
2) $k_1 \cdot k_2$ — ядро.
3) $\alpha \cdot k_1$ — ядро
4) $f(x) \cdot f(z)$ — ядро (f — веществ. ф-ция)
5) $k(\varphi(x), \varphi(z))$ — ядро
6) $\lim\limits_{n\to\infty} k_n(x, z)$ — ядро

$$k(x,z) = (\langle x, z\rangle + R)^m, \quad R > 0, \ m \in \mathbb{N}_+$$

$$k(x,z) = \sum_{k=0}^{m} C_m^k \cdot R^{m-k} \langle x, z\rangle^k$$

1) $\langle x, z\rangle$ — ядро, т.к $\varphi(x) = x$
2) $\langle x, z\rangle^k$ — ядро (св-во 2)
3) $C_m^k \cdot R^{m-k} \cdot \langle x, z\rangle^k$ — ядро (по 3)
4) Σ — ядро (по 1)

$K(x,z) = (x+z)^2 \equiv, \quad x \in \mathbb{R} \ni z$

$K(x,z) \longrightarrow \mathbb{H}$?

$\quad \| $

$\langle \varphi(x), \varphi(z) \rangle$

$\equiv x^2 + 2xz + z^2 = \left\langle \underbrace{\begin{pmatrix} 1 \\ \sqrt{2}x \\ x^2 \end{pmatrix}}, \underbrace{\begin{pmatrix} z^2 \\ \sqrt{2}z \\ 1 \end{pmatrix}} \right\rangle$

$K(x,z) = (\langle x,z \rangle + 1)^2 =$

$= \langle x,z \rangle^2 + 2\langle x,z \rangle + 1 =$

$= \langle x,z \rangle \cdot \langle x,z \rangle$

Гауссово ядро: $K(x,z) = \exp\left(-\dfrac{\|x-z\|^2}{2\sigma^2}\right)$

$K(x,z) = K(\|x-z\|)$ — RBF

$K(x,z) \to \mathbb{H}$ — бесконечномерное.

1. Утв: $\forall x_1, \dots, x_n$ — попарно разл.

$$G = \left(\exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)\right)_{i,j=1}^{n}$$

G — невырожденна $\forall n$, при $\sigma > 0$.

(2): $K(x,z) = \langle \varphi(x), \varphi(z) \rangle$

$$\exp\left(-\frac{\|x-z\|^2}{2\sigma^2}\right) = \exp\left(\frac{-\|x\|^2-\|z\|^2}{2\sigma^2}\right) \cdot$$

$$\cdot \exp\left(\frac{\langle x,z \rangle}{\sigma^2}\right) =$$

$$= \exp\left(-\frac{\|x\|^2}{2\sigma^2}\right) \cdot \exp\left(-\frac{\|z\|^2}{2\sigma^2}\right) \cdot \sum_{k=0}^{+\infty} \frac{\langle x,z \rangle^k}{k!\,\sigma^{2k}} =$$

$$= \sum_{k=0}^{+\infty} \langle \quad\quad$$

Утв: $\forall x_1, \dots, x_n$ — попарно разл. для $\exists!$ которое соотв. Гауссовому ядру $\exists$ линейный классиф., который безошибочно разделяет выборку.

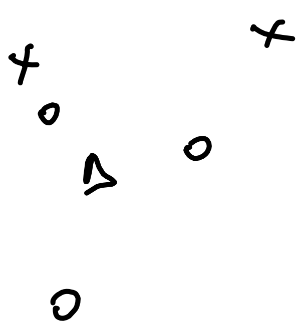
Семинар

Soft-Margin

$$
\begin{cases}
\frac{1}{2}\|w\|^2 + C \cdot \sum_i \xi_i \to \min\limits_{w, b, \xi_i} \\
\xi_i \geq 0 \\
y_i (\langle w, x_i \rangle + b) \geq 1 - \xi_i
\end{cases}
$$

$$Q = \frac{1}{n} \sum_i \ln(1 + \exp(-y_i \langle w, x_i \rangle)) + \frac{\lambda}{2}\|w\|^2$$

$$a(x) = \arg\max_c \sum_{i \in N_x(k)} \sum_{c=1}^{d} [y_i = c] \cdot w_i$$



$c \in \{1, 2, \ldots, C\}$

$$G = \sum_{k=1}^{K} p_k(1-p_k) \qquad H = -\sum_{k=1}^{K} p_k \log_2 p_k$$

$$p_k = \frac{1}{|R|} \sum_{i:(x_i,y_i)\in R} [y_i = k] \qquad G(R)$$

$$Q = G(R) - \frac{|R_l|}{|R|} \cdot G(R_l) - \frac{|R_r|}{|R|} \cdot G(R_r)$$

$$\|w\|^2 = \sum_{i=1}^{n} \lambda_i \qquad \text{K.K.T:} \begin{cases} w = \sum_i \lambda_i y_i x_i \\ \sum_i \lambda_i y_i = 0 \end{cases}$$

$$b = y_i - \langle w, x_i \rangle$$

$$0 = b \sum_{i=1}^{n} \lambda_i y_i = \sum_{i=1}^{n} \lambda_i y_i (y_i - \langle w, x_i \rangle) =$$

$$= \sum_{i=1}^{n} \lambda_i - \underbrace{\sum_{i=1}^{n} \lambda_i y_i \langle w, x_i \rangle}_{\langle w, \underbrace{\sum_i \lambda_i y_i x_i}_{w} \rangle} = \sum_i \lambda_i - \|w\|^2$$