$$p(C_1|\bar x, D) = \int \sigma(\bar w^T \bar x)\, p(\bar w|D)\, d\bar w$$

$\bar w^T \bar x = a$ $\quad\quad$ $p(\bar w|D) \approx \mathcal{N}(\bar w|\bar \mu, \Sigma)$

$\mu_a = \bar w_{MAP}^T \bar x$

$$= \int \sigma(a)\, \mathcal{N}(a|\mu_a, \sigma_a^2)\, da \approx$$

$$\approx \int \Phi(\lambda a)\, \mathcal{N}(a|\mu_a, \sigma_a^2)\, da$$

$$\Phi(\lambda a) = \int_{-\infty}^{\lambda a} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx = \left[x = \frac{1}{\lambda} x\right] =$$

$$= \int_{-\infty}^{a} \frac{\lambda}{\sqrt{2\pi}} e^{-\lambda^2 x^2/2}\, dx = p(X \le a),$$

где $X \sim \mathcal{N}(0, \frac{1}{\lambda^2})$

$$\int \Phi(\lambda a)\, \mathcal{N}(a|\mu_a, \sigma_a^2)\, da = \int p(X \le a)\, \mathcal{N}(a|\mu_a, \sigma_a^2)\, da =$$

$= p(X \le Y)$, где $X \sim \mathcal{N}(0, \frac{1}{\lambda^2})$

$Y \sim \mathcal{N}(\mu_a, \sigma_a^2)$

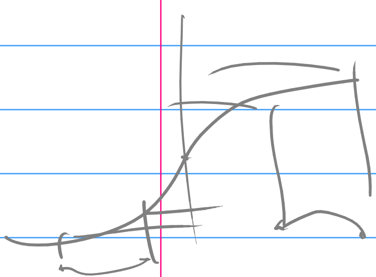
$$= p(\underbrace{X - Y}_{\sim \mathcal{N}(-\mu_a, \frac{1}{\lambda^2} + \sigma_a^2)} \le 0) = p(X - Y + \mu_a \le \mu_a)$$

$p(Z \le \mu_a)$, где $Z \sim \mathcal{N}(0, \frac{1}{\lambda^2} + \sigma_a^2)$

$$= \Phi\left(\mu_a \cdot \frac{1}{\sqrt{\frac{1}{\lambda^2} + \sigma_a^2}}\right)$$

$$\int \Phi(\lambda a)\, \mathcal{N}(a|\mu_a, \sigma_a^2)\, da = \Phi\left(\frac{\mu_a}{\sqrt{\sigma_a^2 + \frac{1}{\lambda^2}}}\right) \approx$$

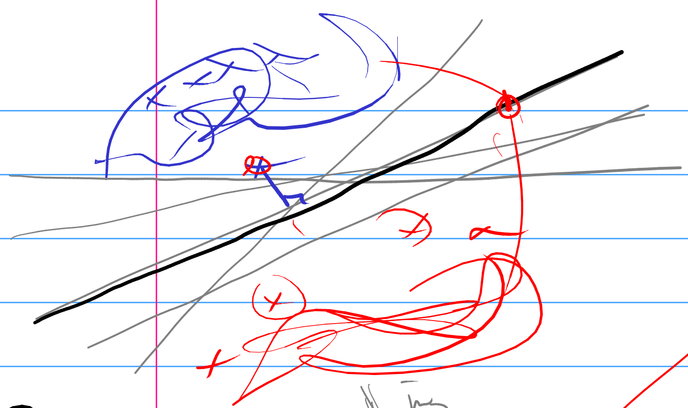
$$\approx \sigma\left(\frac{\lambda \mu_a}{\sqrt{\ }}\right) = \sigma\left(\frac{\bar w_{MAP}^T \bar x}{\sqrt{1 + \frac{\pi}{8}\sigma_a^2}}\right)$$



$a_k = \bar w_k^T \bar x$

$\bar x \to \bar w_1^T - a_1,\ a_2,\ \dots,\ \bar w_k^T - a_k \to \text{softmax} \to \left(\dots, \frac{e^{a_k}}{\sum e^{a_j}}, \dots\right)$

Support Vector Machines

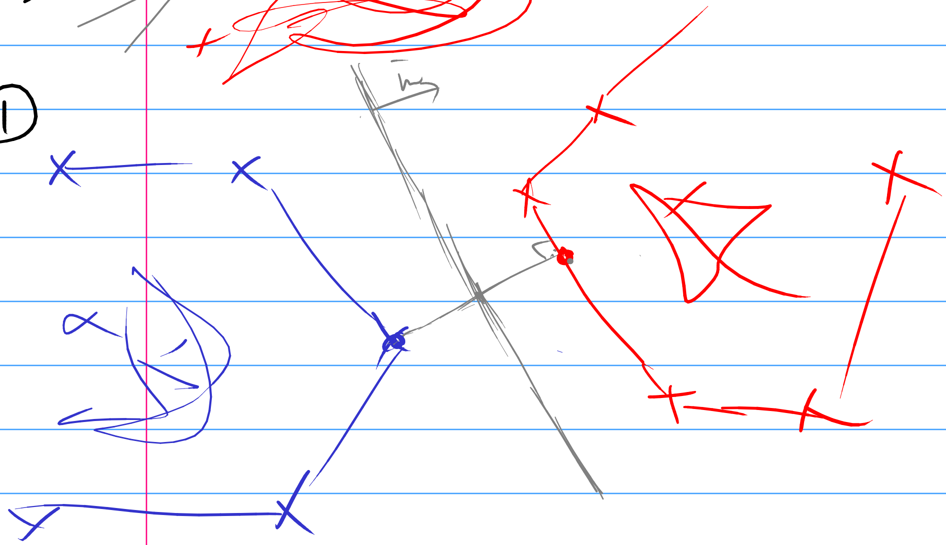


$$\min_{\bar x \in C_1} d(\bar x, \ell) \to \max$$

$$\min_{\bar x \in C_2} d(\bar x, \ell) \to \max$$

$$\mathrm{Conv}(C_1) = \left\{ \sum_{n\in C_1} \alpha_n \bar x_n \;\middle|\; \sum \alpha_n = 1,\ \alpha_n \geq 0 \right\}$$

(1)



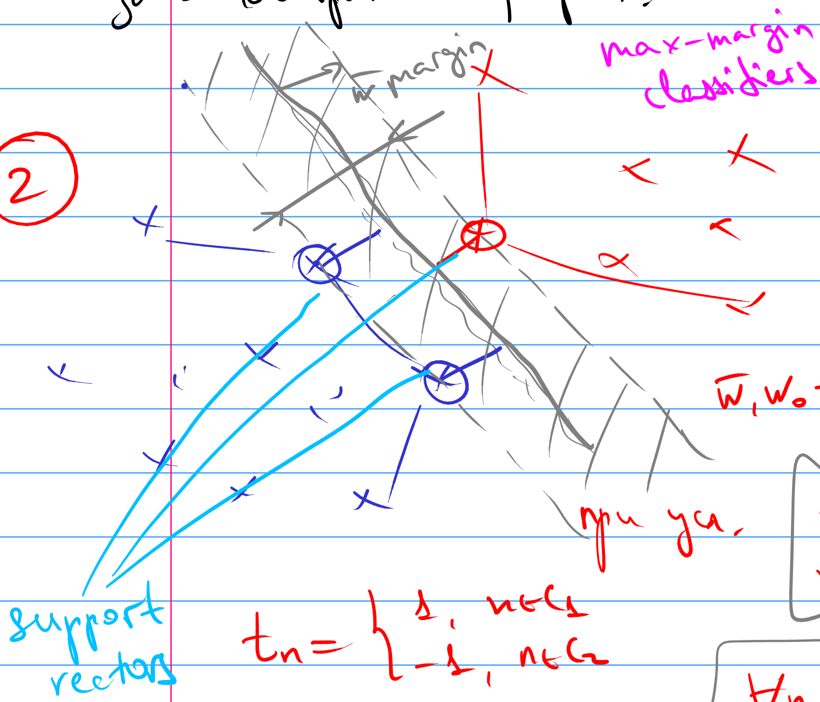
$$\min_{\alpha} \|\bar a - \bar b\|^2, \text{ где } \quad \bar a = \sum_{n\in C_1} \alpha_n \bar x_n, \quad \bar b = \sum_{n \in C_2} \alpha_n \bar x_n$$

при усл.

$$\boxed{\sum_{n\in C_1} \alpha_n = 1, \quad \sum_{n\in C_2} \alpha_n = 1, \quad \forall n\ \alpha_n \geq 0}$$

зад. квадратичн. программ.

(2)



$$d(\bar x, \ell) = \frac{|\bar w^T \bar x + w_0|}{\|\bar w\|}$$

$$\bar w, w_0 = \arg\max_{\bar w, w_0} \min_n \frac{|\bar w^T \bar x_n + w_0|}{\|\bar w\|}$$

при усл.

$$\boxed{\begin{array}{ll} \forall n \in C_1 & \bar w^T \bar x_n + w_0 > 0 \\ \forall n \in C_2 & \bar w^T \bar x_n + w_0 < 0 \end{array}}$$

$$t_n = \begin{cases} 1, & n \in C_1 \\ -1, & n \in C_2 \end{cases}$$

$$\boxed{\forall n \quad t_n(\bar w^T \bar x_n + w_0) > 0}$$

$$\bar w, w_0 = \arg\max_{\bar w, w_0} \min_n \frac{t_n(\bar w^T \bar x_n + w_0)}{\|\bar w\|}$$

$\|\bar w\| = 1$ $\quad \cancel{\max_{\bar w, w_0} \left\{ \min_n t_n(\bar w^T \bar x_n + w_0) \right\}}, \quad \forall n\ t_n(\bar w^T \bar x_n + w_0) > 0 \quad [w_1^2 + \dots + w_d^2 = 1]$

fix $\|\bar{w}\|$: $\underbrace{\min_n t_n(\bar{w}^T\bar{x}_n + w_0) = 1}$

$$\max_{\bar{w}, w_0} \frac{1}{\|\bar{w}\|} \quad \text{при усл.} \quad \forall n \ t_n(\bar{w}^T\bar{x}_n + w_0) \geq 1$$

$$= \min_{\bar{w}, w_0} \|\bar{w}\|^2$$

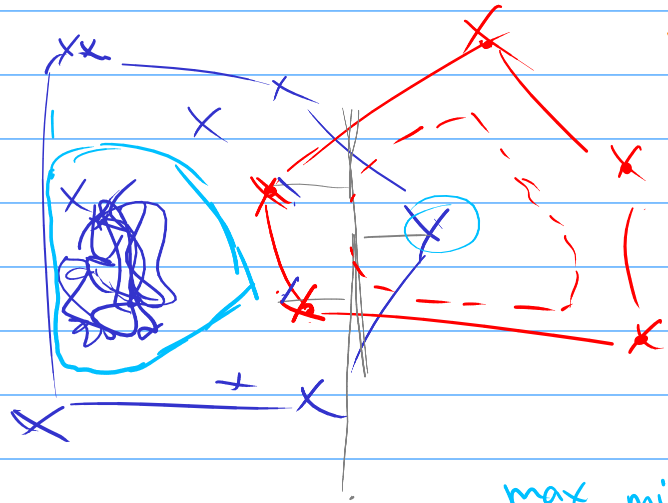
квадр. программирование

Владимир Вапник — VC-dimension

Алексей Червоненкис

$\boxed{\text{Accuracy}} \to \max$

(3)



Reduced convex hull

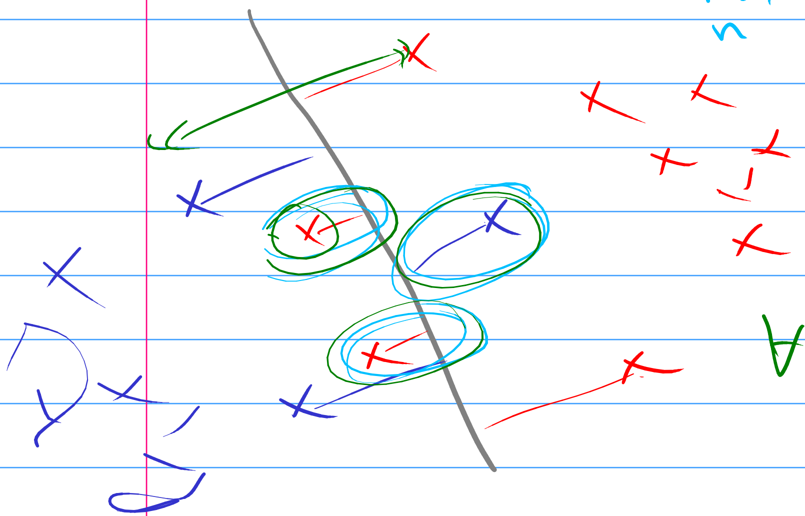
$$\left\{ \sum \alpha_n \bar{x}_n \;\middle|\; \sum \alpha_n = 1, \ 0 \leq \alpha_n \leq A \right\}$$

$$\min \|\bar{a} - \bar{b}\|^2$$

$$\bar{a} \in \mathrm{Conv}_A(C_1)$$

$$\bar{b} \in \mathrm{Conv}_A(C_2)$$

$$\max_n \min(\bar{w}^T\bar{x}_n + w_0, 0)$$



$$\min_{\bar{w}, w_0} \left(\|\bar{w}\|^2 + C \cdot \sum_n z_n \right)$$

при условии

$$\forall n \quad t_n(\bar{w}^T\bar{x}_n + w_0) + z_n \geq 1$$

$$z_n \geq 0$$

slack

(4)

$$\min_{\bar{w}, w_0} \frac{1}{2}\|\bar{w}\|^2 \quad \text{при усл.} \quad \forall n \ t_n(\bar{w}^T\bar{x}_n + w_0) \geq 1$$

$$L(\bar{w}, w_0, \bar{\alpha}) = \frac{1}{2}\|\bar{w}\|^2 - \sum_n \alpha_n \left(t_n(\bar{w}^T\bar{x}_n + w_0) - 1 \right), \quad \forall n \ \alpha_n \geq 0$$

$$\nabla_{\bar{w}} L = \bar{w} - \sum_n \alpha_n t_n \bar{x}_n = 0 \Rightarrow \bar{w} = \sum_n \alpha_n t_n \bar{x}_n$$

$$\frac{\partial L}{\partial w_0} = -\sum_n \alpha_n t_n \Rightarrow \sum_n \alpha_n t_n = 0$$

$$L(\bar{\alpha}) = \frac{1}{2}\left(\sum_n \alpha_n t_n \bar{x}_n\right)^T \left(\sum_n \alpha_n t_n \bar{x}_n\right) - \sum_n \alpha_n \left(t_n \left(\left(\sum_m \alpha_m t_m \bar{x}_m\right)^T \bar{x}_n\right) - 1\right)$$

$$= \frac{1}{2} \sum_{n,m} \alpha_n \alpha_m t_n t_m \bar{x}_n^T \bar{x}_m - \sum_{n,m} \alpha_n \alpha_m t_n t_m \bar{x}_m^T \bar{x}_n + \sum_n \alpha_n$$

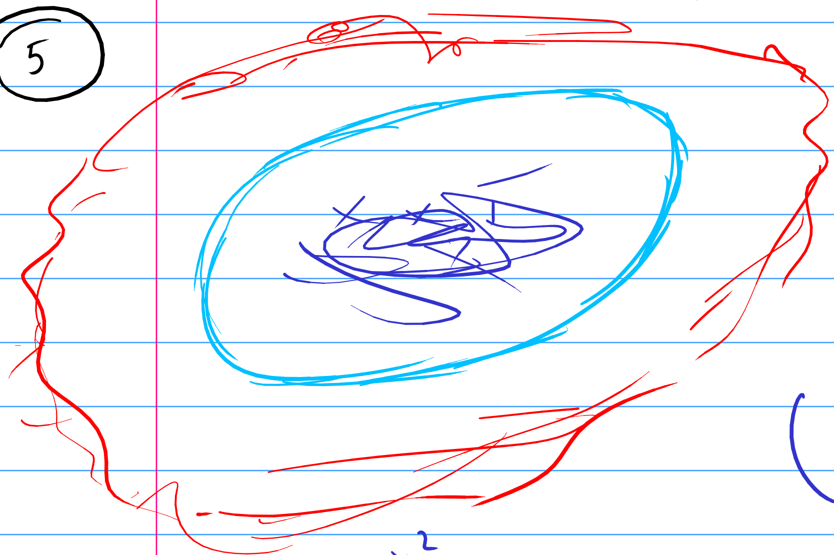
$$\boxed{L(\bar{\alpha}) = \sum_n \alpha_n - \frac{1}{2} \sum_n \sum_m \alpha_n \alpha_m t_n t_m (\bar{x}_n^T \bar{x}_m) \qquad \text{מקס׳ ע״מ} \quad \forall n\ \alpha_n \geq 0, \quad \sum_n \alpha_n t_n = 0}$$

$$y(\bar{x}) = \bar{w}^T \bar{x} + w_0 = \sum_n \alpha_n t_n (\bar{x}_n^T \bar{x}) + w_0$$

Karush-Kuhn-Tucker:

$$\begin{cases} \forall n & \alpha_n \geq 0 \\ \forall n & t_n y(\bar{x}_n) - 1 \geq 0 \\ \forall n & \alpha_n (t_n y(\bar{x}_n) - 1) = 0 \end{cases}$$

5



$$y(\bar{x}) = \bar{x}^T A \bar{x} + \bar{b}^T \bar{x} + c, \quad \bar{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$y(\bar{x}) = w_0 + w_1 x_1 + w_2 x_2 + w_3 x_1^2 + w_4 x_1 x_2 + w_5 x_2^2$$

$$\underset{\mathbb{R}^2}{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}} \longmapsto \underset{\mathbb{R}^5}{\begin{pmatrix} x_1^2 \\ x_1 x_2 \\ x_2^2 \\ x_1 \\ x_2 \end{pmatrix}}$$

$$\underset{\mathbb{R}^d}{\begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix}} \longmapsto \underset{\mathbb{R}^{\frac{d(d+1)}{2}}}{\begin{pmatrix} x_1^2 \\ \cdots \\ x_d \end{pmatrix}}$$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \longmapsto \begin{pmatrix} x_1^2 \\ x_1 x_2 \\ x_2^2 \end{pmatrix}$$

$$\begin{pmatrix} x_1^2 \\ \sqrt{2} x_1 x_2 \\ x_2^2 \end{pmatrix}^T \begin{pmatrix} y_1^2 \\ \sqrt{2} y_1 y_2 \\ y_2^2 \end{pmatrix} = x_1^2 y_1^2 + 2 x_1 x_2 y_1 y_2 + x_2^2 y_2^2 =$$

$$= (x_1 y_1 + x_2 y_2)^2 = \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^T \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right)^2$$

$$\begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix} \xrightarrow{\varphi} \begin{pmatrix} x_1^k \\ \vdots \\ x_d^k \end{pmatrix}$$

$\mathbb{R}^d$

$$\varphi(\bar{x})^T \varphi(\bar{y}) = \underbrace{(\bar{x}^T \bar{y})^k}_{\text{kernel}}$$

$\boxed{\text{kernel trick}}$

$$L(\bar{\alpha}) = \sum \alpha_n - \frac{1}{2} \sum_{n,m} \alpha_n \alpha_m t_n t_m k(\bar{x}_n, \bar{x}_m)$$

υπό για $\forall n \ \alpha_n \geq 0, \ \sum \alpha_n t_n = 0$

$$\boxed{y(\bar{x}) = \sum_n \alpha_n t_n k(\bar{x}_n, \bar{x}) + w_0}$$

$$\begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix} \xrightarrow{\varphi} \begin{pmatrix} x_1^k \\ \vdots \\ x_d^k \\ x_1^{k-1} \\ \vdots \\ \vdots \\ x_d \end{pmatrix}$$

$$\varphi(\bar{x})^T \varphi(\bar{y}) = (\bar{x}^T \bar{y})^k + (\bar{x}^T \bar{y})^{k-1} + \ldots$$

$$= (\bar{x}^T \bar{y} + 1)^k - 1$$

$$k(\bar{u}, \bar{v}) = e^{-\frac{\lambda}{2} \|\bar{u} - \bar{v}\|^2}$$

6 ν-SVM

$$\frac{1}{2} \bar{w}^T \bar{w} + C \cdot \sum_n z_n, \quad z_n \geq 0, \quad t_n(\bar{w}^T \bar{x}_n + w_0) + z_n \geq 1$$

$$\frac{1}{2} \bar{w}^T \bar{w} + \frac{1}{N} \sum_{n=1}^{N} z_n - \nu \rho, \quad \text{υπό για} \quad z_n \geq 0, \ \rho \geq 0$$

$$t_n(\bar{w}^T \bar{x}_n + w_0) + z_n \geq \rho$$

$$\geq \frac{\rho}{\|\bar{w}\|}$$

$$L(\bar{w}, w_0, \bar{z}, \rho, \alpha, \beta, \delta) = \frac{1}{2} \bar{w}^T \bar{w} + \frac{1}{N} \sum_n z_n - \nu \rho$$

$$- \sum_{n=1}^{N} \left(\alpha_n \left(t_n(\bar{w}^T \bar{x}_n + w_0) + z_n - \rho \right) + \beta_n z_n \right) + \delta \rho$$

$$\boxed{\begin{aligned} \bar{w} &= \sum_n \alpha_n t_n \bar{x}_n \\ \sum_n \alpha_n t_n &= 0 \end{aligned}}$$

$$\frac{\partial L}{\partial z_n} = \frac{1}{N} - \alpha_n - \beta_n = 0$$

$$\forall n \ \boxed{\alpha_n + \beta_n = \frac{1}{N}}$$

$\alpha_n \geq 0, \ \beta_n \geq 0,$

$\delta \geq 0$

$$\frac{\partial L}{\partial \xi} = -\nu + \delta + \sum_n \alpha_n$$

$$\boxed{\sum_n \alpha_n + \delta = \nu}$$

если $\alpha_n > 0$, то

$$t_n(\bar{w}^T \bar{x}_n + w_0) + z_n = \rho$$

если $z_n > 0$, то $\beta_n = 0$

$$\frac{1}{2}\left(\sum_n \alpha_n t_n \bar{x}_n\right)^T\left(\sum_m \alpha_m t_m \bar{x}_m\right) - \sum_n \alpha_n t_n \left(\left(\sum_m \alpha_m t_m \bar{x}_m\right)^T x_n\right)$$

$$L(\alpha) = -\frac{1}{2}\sum_{n,m} \alpha_n \alpha_m t_n t_m (\bar{x}_n^T \bar{x}_m) \quad \text{"} k(\bar{x}_n, \bar{x}_m)$$

при усл. $0 \le \alpha_n \le \frac{1}{N}$, $\sum \alpha_n t_n = 0$, $\sum_n \alpha_n \le \nu$

если $\boxed{z_n > 0} \Rightarrow \beta_n = 0 \Rightarrow \alpha_n = \frac{1}{N}$

но $\sum_n \alpha_n \le \nu$

$\Rightarrow$ таких n, что $z_n > 0$, не больше νN