Support Vector Machines (SVM)

$T = \{(x_1, y_1), \dots, (x_N, y_N)\}$ $x_i \in \mathbb{R}^D$ $y_i \in \{-1, +1\}$

$$d_i(w,b) = \frac{(w^T x_i + b)\, y_i}{\|w\|}$$ "signed" distance

2d = margin

$\min_i d_i \rightarrow \max d$ s.t. $d \le d_i \ \forall i$

$$(w^*, b^*) = \arg\max_{w,b} \overbrace{\min_i d_i}^{d}$$
$$\text{s.t. } (w^T x_i + b)\, y_i > 0 \quad \forall i$$

optimální → optimální

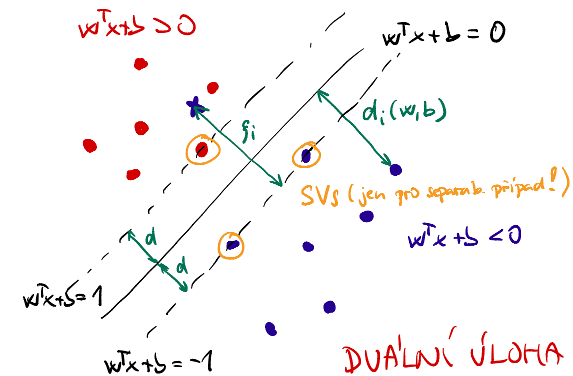
$(w, b) \rightarrow (sw, sb)$ $s > 0$

$$d_i(sw, sb) = \frac{(sw^T \cdot x_i + sb)\, y_i}{\|sw\|} = \frac{\cancel{s}(w^T x_i + b)\, y_i}{\cancel{s}\|w\|}$$

$$(w^*, b^*) = \arg\max_{w,b,d} d$$
$$\text{s.t. } \frac{(w^T x_i + b)\, y_i}{\|w\|} \ge d \quad \forall i$$

$\|w\| = 1/d$

$$(w^*, b^*) = \arg\max_{w,b} \frac{1}{\|w\|} = \arg\min_{w,b} \frac{1}{2}\|w\|^2$$
$$\text{s.t. } (w^T x_i + b)\, y_i \ge 1$$



nesep. data

$$\boxed{\begin{aligned}(w^*, b^*) &= \arg\min_{w,b} \frac{1}{2}\|w\|^2 + C \sum_i \xi_i \\ \text{s.t. } (w^T x_i + b)\, y_i &\ge 1 - \xi_i \quad \forall i \\ \xi_i &\ge 0 \quad \forall i\end{aligned}}$$

PRIMÁRNÍ ÚLOHA

přednáška →

DUÁLNÍ ÚLOHA

$$\boxed{\begin{aligned}\alpha^* &= \arg\max_{\alpha} \sum_{i=1}^{N} \alpha - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j x_i^T x_j \\ \text{s.t. } &\sum_{i=1}^{N} \alpha_i y_i = 0 \\ &0 \le \alpha_i \le C \quad \forall i\end{aligned}}$$

support vectors

$$w^* = \sum_{i \in SV} \alpha_i y_i x_i \qquad SV = \{i : 0 < \alpha_i \le C\}$$

10^{-10}

$$b^* = \frac{1}{|I|} \sum_{i \in I} \left(y_i - \sum_{j=1}^{N} \alpha_j y_j x_j^T x_i\right) \qquad I = \{i : 0 < \alpha_i < C\}$$

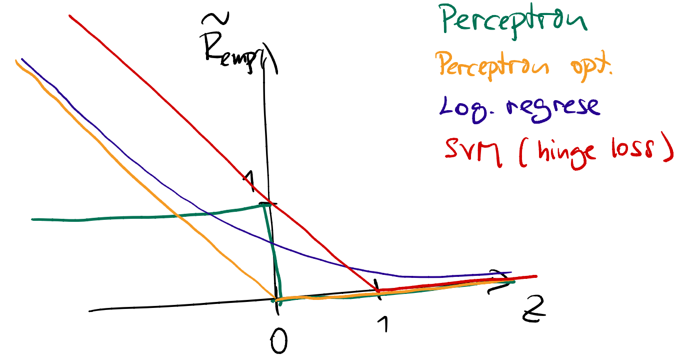
může být $I = \emptyset$!

DÚ: přečíst excercise 4 (viz text cvičení)

klasifikace:

$$q(x): \sum_{i=1}^{N} \alpha_i y_i x_i^T x + b^* \ge 0$$

$$R_{emp}(q(x)) = \frac{1}{N}\sum_{i=1}^{N} W(q(x_i), y_i)$$ empirické riziko



PERCEPTRON

$v^T x$

$$\underbrace{w^T x_i}_{z_i} \leq 0 \rightarrow \text{chyba}$$

$$\tilde{R}_p(w) = \sum_{\substack{i=1 \\ w^T x_i \leq 0}}^{N} -w^T x_i = \sum_{i=1}^{N} \max(-w^T x_i, 0)$$

LOG. REGRESE

$$E(w) = \sum_{i=1}^{N} \ln\left(1 + e^{\overbrace{-y_i w^T x_i}^{z_i}}\right)$$ diferencovatelné $\rightarrow$ GD

SVM

$$(w^*, b^*) = \arg\min_{w,b} \frac{1}{2}\|w\|^2 + C\sum_i \xi_i$$

$$\text{s.t. } (w^T x_i + b) y_i \geq 1 - \xi_i \quad \forall i$$

$$\xi_i \geq 0 \quad \forall i$$

fix: ξ_j pro $j \neq i$, w, b

optimální ξ_i ?

$$\min_{\xi_i} \xi_i + \boxed{\frac{1}{C}\left(\sum_{j\neq i} \xi_j + \frac{1}{2}\|w\|^2\right)} \text{ const}$$

$$\xi_i \geq 1 - \underbrace{(w^T x_i + b) y_i}_{z_i}$$

$$\xi_i \geq 0$$

$$\xi_i^* = \max(0, 1 - z_i)$$