FOUTE
TAPE
JORDAN
M1 IA.

TD 4 Support Vector Machine.

3. Cas non linéairement separable

$y(x) = w^T\phi(x) + b$; (PRIMAL) : $\text{Min } \frac{1}{2}\|w\|^2$

$s\text{-}a \quad t^i(w^T\phi(x^i)+b) \geq 1 \quad i=1,\dots N.$

$w \in \mathbb{R}^n, \; b \in \mathbb{R}$

(DUAL) : $\text{Max } e^T\lambda - \frac{1}{2}\lambda^T Q \lambda$

$s\text{-}a \quad \lambda^T t = 0$

$\lambda \geq 0$

3) Conditions d'optimalité de K-K-T sur le primal.

Posons $f(w) = \frac{1}{2}\|w\|^2$ et $g_i(w) = 1 - t^i(w^T\phi(x^i)+b)$.

1*) $\nabla_w f(w) + \sum_{i=1}^{N} \lambda_i \nabla_w g_i(w) = 0$

on a : $\nabla_w f(w) = \frac{\partial f(w)}{\partial w_k} = \frac{1}{2}\frac{\partial w_k^2}{\partial w_k} = w_k$, $k=1,\dots N.$

$\nabla_w g_i(w) = \frac{\partial(1 - t^i(w^T\phi(x^i)+b)}{\partial w_k}$ or $w^T\phi(x^i) = \sum_{i=1}^{N} w_i \phi(x_i)$

$\frac{\partial w^T\phi(x^i)}{\partial w_k} = \phi(x_i)$

$\nabla_w g_i(w) = -t^i\phi(x_i)$

donc $w_k - \sum_{i=1}^{N} \lambda_i t^i \phi(x^i) = 0$, $k=1,\dots N.$

$W - \sum_{i=1}^{N} \lambda_i t^i \phi(x^i) = 0$

2°) $\nabla_b q(w) + \sum_{i=1}^{N} \lambda_i \nabla_b g_i(w) = 0$

$\nabla_b q(w) = 0$, $\nabla_b g_i(w) = \frac{\partial(1 - t^i(w^T\phi(x^i)+b))}{\partial b}$

$= -t^i$

donc $-\sum_{i=1}^{N} \lambda_i t^i = 0$

3°) $\lambda_i g_i(w) = 0$

4) Montrons que si λ solution du DUAL alors $W = \sum_{i=1}^{N} \lambda_i t^i \phi(x^i)$

d'après la question precedente, on a:

$$W - \sum_{i=1}^{N} \lambda_i t^i \phi(x^i) = 0 \Rightarrow W = \sum_{i=1}^{N} \lambda_i t^i \phi(x^i).$$

5) Montrons que $y(x) = \sum_{i=1}^{N} \lambda_i t^i k(x, x^i) + b$.

$$y(x) = w^T\phi(x_i) + b$$
$$= \left[\sum_{i=1}^{N} \lambda_i t^i \phi(x_i)\right]^T \phi(x_j) + b$$
$$= \sum_{i=1}^{N} \lambda_i t^i \phi^T(x_i)\phi(x_j) + b \quad \text{or} \quad k(x_i, x_j) = \phi^T(x^i)\phi(x_j)$$
$$y(x) = \sum_{i=1}^{N} \lambda_i t^i k(x^i, x^j) + b.$$

6) Montrons que si $\lambda_i > 0$ alors $t^i y(x^i) = 1$.

on a: $\lambda_i g_i(w) = 0$, si $\lambda_i > 0 \Rightarrow g_i(w) = 0$

$$g_i(w) = 0 \Rightarrow 1 - t^i(\underbrace{w^T\phi(x^i) + b}_{y(x^i)}) = 0$$
$$\Rightarrow 1 - t^i y(x^i) = 0$$
$$\Rightarrow t^i y(x^i) = 1.$$

FOLITE
TAPE
JORDAN
M1 IA

7) Deduisons que si $\lambda_i > 0$ alors $y(x^i) = t^i$

si $\lambda_i > 0$, $t^i y(x^i) = 1$.

si $t^i = 1 \Rightarrow y(x^i) = 1$

si $t^i = -1 \Rightarrow y(x^i) = -1$

donc $y(x^i) = t^i$

8) Deduisons que si λ est solution optimale du DUAL, alors b optimal est $b = \frac{1}{|S|} \sum_{i \in S} (t^i - \sum_{j \in S} \lambda_j t^j K(x^i, x^j))$.

$$y(x^i) = \sum_{j=1}^{M} \lambda_j t^j K(x^i, x^j) + b = t^i$$

$$\Rightarrow \quad b = t^i - \sum_{j=1}^{M} \lambda_j t^j K(x^i, x^j)$$

$$\sum_{i \in S} b = \sum_{i \in S} (t^i - \sum_{j=1}^{M} \lambda_j t^j K(x^i, x^j))$$

$$b = \frac{1}{|S|} \sum_{i \in S} (t^i - \sum_{j \in S} \lambda_j t^j K(x^i, x^j))$$