Exercice n°1 :

- $\mathcal{X} = \{A, T, G, C\}^N$

$$k : (x_1, \dots, x_N), (y_1, \dots, y_N) \mapsto$$

$$\sum_{(t_1,t_2,t_3) \in \{A,T,G,C\}^N} \left(\sum_{b=0}^{N-3} \mathbb{1}(x_{b+1} = t_1, \dots) \right)$$

$$\left(\quad \text{'' avec } y \right).$$

Alors $k : (x_1, \dots, x_N), (y_1, \dots, y_N) \mapsto \sum_{\substack{(t_1,t_2,t_3) \\ \in \{A,T,G,C\}^N}} \phi_{t_1,t_2,t_3}(x) \times \phi_{t_1,t_2,t_3}(y)$

$$= \langle \phi(x), \phi(y) \rangle_{\mathbb{R}^{(4^3)}}$$

où $\phi(z) = \begin{pmatrix} \phi_{A,A,A}(z) \\ \phi_{A,A,T}(z) \\ \vdots \\ \phi_{C,C,C}(z) \end{pmatrix}$.

• $\mathcal{X} = \mathbb{R}$, $\forall x, y \in \mathcal{X}$, $k(x,y) = \cos(x-y)$

$$\forall x, y, \quad \cos(x-y) = \cos(x)\cos(y) + \sin(x)\sin(y)$$
$$= \left\langle \begin{pmatrix} \cos(x) \\ \sin(x) \end{pmatrix}, \begin{pmatrix} \cos(y) \\ \sin(y) \end{pmatrix} \right\rangle_{\mathbb{R}^2}$$

donc le noyau est symmétrique positif.

• $X = \mathbb{R}$, $K : X \times X \to \mathbb{R}_+$ qui est symétrique.

C'est faux en général.

par exemple $K : \begin{cases} 0,0 \mapsto 0 \\ 1,1 \mapsto 0 \\ 0,1 \mapsto 1 \\ 1,0 \mapsto 1 \end{cases}$

Alors $\begin{pmatrix} K(0,0) & K(0,1) \\ K(1,0) & K(1,1) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

dont le déterminant vaut -1. Cette matrice symétrique a donc au moins une valeur propre négative. Elle n'est donc pas positive.

• $X = ~~\mathbb{R}~~ (-1,1)$, $\forall x,y \in X$, $K(x,y) = 1/(1-xy)$

$$\forall x,y \in (-1,1),\ K(x,y) = \frac{1}{1-xy} = \sum_{n=0}^{+\infty} (xy)^n$$

Or, $x,y \mapsto xy = \langle x,y \rangle_{\mathbb{R}}$ est symétrique positif

donc $x,y \mapsto (xy)^n$ est symétrique positif $\forall n \geq 1$ par règles de produit

et aussi, $x,y \mapsto 1$ est symétrique positif (immédiat) donc, par passage à la limite simple,

$$x, y \mapsto \sum_{i=1}^{n} (x_i y_i)^n \text{ est symétrique positif.}$$

donc K est symétrique positif.

- $\mathcal{X} = \mathcal{T}$, $\forall A, B \in \mathcal{X}$, $K(A,B) = \mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)$

Soient $A_1, \dots, A_n \in \mathcal{X}$, $\lambda_1, \dots, \lambda_n \in \mathbb{R}$.

$$\sum_{i,j} \lambda_i \lambda_j K(A_i, A_j) = \sum_{i,j} \lambda_i \lambda_j \left(\mathbb{P}(A_i \cap A_j) - \mathbb{P}(A_i)\mathbb{P}(A_j)\right)$$

$$= \sum_{i,j} \lambda_i \lambda_j \left(\mathbb{E}(\mathbb{1}_{A_i} \mathbb{1}_{A_j}) - \mathbb{E}(\mathbb{1}_{A_i})\mathbb{E}(\mathbb{1}_{A_j})\right)$$

$$= \sum_{i,j} \cancel{\mathbb{E}} \lambda_i \lambda_j \mathbb{E}(\mathbb{1}_{A_i} \mathbb{1}_{A_j}) - \sum_{i,j} \lambda_i \lambda_j \mathbb{E}(\mathbb{1}_{A_i})\mathbb{E}(\mathbb{1}_{A_j})$$

$$= \mathbb{E}\left(\sum_{i,j} \cancel{\mathbb{1}} \lambda_i \lambda_j \mathbb{1}_{A_i} \mathbb{1}_{A_j}\right) - \sum_{i,j} \lambda_i \lambda_j \mathbb{E}(\mathbb{1}_{A_i})\mathbb{E}(\mathbb{1}_{A_j})$$

$$= \mathbb{E}\left(\left(\sum_i \lambda_i \mathbb{1}_{A_i}\right)^2\right) - \left(\mathbb{E}\left(\sum_i \lambda_i \mathbb{1}_{A_i}\right)\right)^2$$

$$= \mathrm{Var}\left(\sum_i \lambda_i \mathbb{1}_{A_i}\right)$$

$$\geq 0.$$

Donc K est symétrique positif.

• $X = \mathbb{R}_+$, $\forall x, y \in X$, $K(x,y) = \min(x,y)$

$\forall x, y \in \mathbb{R}_+$, $\min(x,y) = \int_{\mathbb{R}_+} \mathbb{1}_{[0,x]} \mathbb{1}_{[0,y]}$

$$= \langle \mathbb{1}_{[0,x]}, \mathbb{1}_{[0,y]} \rangle_{L^2(\mathbb{R}_+)}$$

• $X = \mathbb{N}_*$, $\forall x, y \in X$, $K(x,y) = \mathrm{pgcd}(x,y)$

Soit $\mathcal{P} = \{p_1, p_2, \dots\}$ l'ensemble des nombres premiers énuméré par ordre croissant.

Alors, $\forall x, y$, $\mathrm{pgcd}(x,y)$

$$= p_1^{\min(v_{p_1}(x), v_{p_1}(y))} \times p_2^{\min(v_{p_2}(x), v_{p_2}(y))} \times \dots$$

d'après la question précédente et par composition à gauche,

$x, y \mapsto \min(v_p(x), v_p(y))$ est symétrique positif pour tout nombre premier p.

donc, par passage à l'exponentielle (cf cours)

$x, y \mapsto p^{\min(v_p(x), v_p(y))}$ est symétrique positif pour tout nombre premier p.

donc, $\forall N$, $x, y \longmapsto \prod_{i=1}^{N} p_i^{\min(v_{p_i}(x), v_{p_i}(y))}$ est symétrique positif (par produit).

donc par passage à la limite simple,

$$x, y \longmapsto \text{pgcd}(x,y) \text{ est symétrique positif.}$$

• <u>$X = \mathbb{N}^*$, $\forall x, y \in X$, $K(x,y) = \dfrac{1}{\text{ppcm}(x,y)}$.</u>

$$\forall x, y \in X, \ \text{pgcd}(x,y)\,\text{ppcm}(x,y) = xy.$$

Alors $\forall x, y \in X$, $\text{ppcm}(x,y) = \dfrac{\text{pgcd}(x,y)}{xy}$.

or $x, y \mapsto \text{pgcd}(x,y)$ est symétrique positif d'après la question précédente.

Si nous prouvons que $x, y \mapsto \dfrac{1}{xy}$ est symétrique positif, nous aurons gagné.

or, $\forall x, y \in X$, $\dfrac{1}{xy} = \left\langle \dfrac{1}{x}, \dfrac{1}{y} \right\rangle_{\mathbb{R}}$

ce qui conclut la preuve.

Exercice 2 :

1) Soit $n \geq 1$, $x_1, \dots, x_n \in \mathcal{X}$, $\alpha_1, \dots, \alpha_n \in \mathbb{R}$,

$$\sum_{i,j} \alpha_i \alpha_j k(x_i, x_j) = \sum_{i,j} \alpha_i \alpha_j \langle \phi(x_i), \phi(x_j) \rangle \text{ (par définition)}$$

$$= \langle \sum_i \alpha_i \phi(x_i), \sum_j \alpha_j \phi(x_j) \rangle \text{ (bilinéarité)}.$$

$$= \| \sum_k \alpha_k \phi(x_k) \|^2 \text{ (définition de la norme)}$$

$$\geq 0 \quad \text{(car carré tjs positif)}.$$

2) D'après le <u>théorème d'Aronszajn</u>, si k est un noyau symétrique positif, alors il existe H un espace de Hilbert et $\phi : \mathcal{X} \to H$ tq

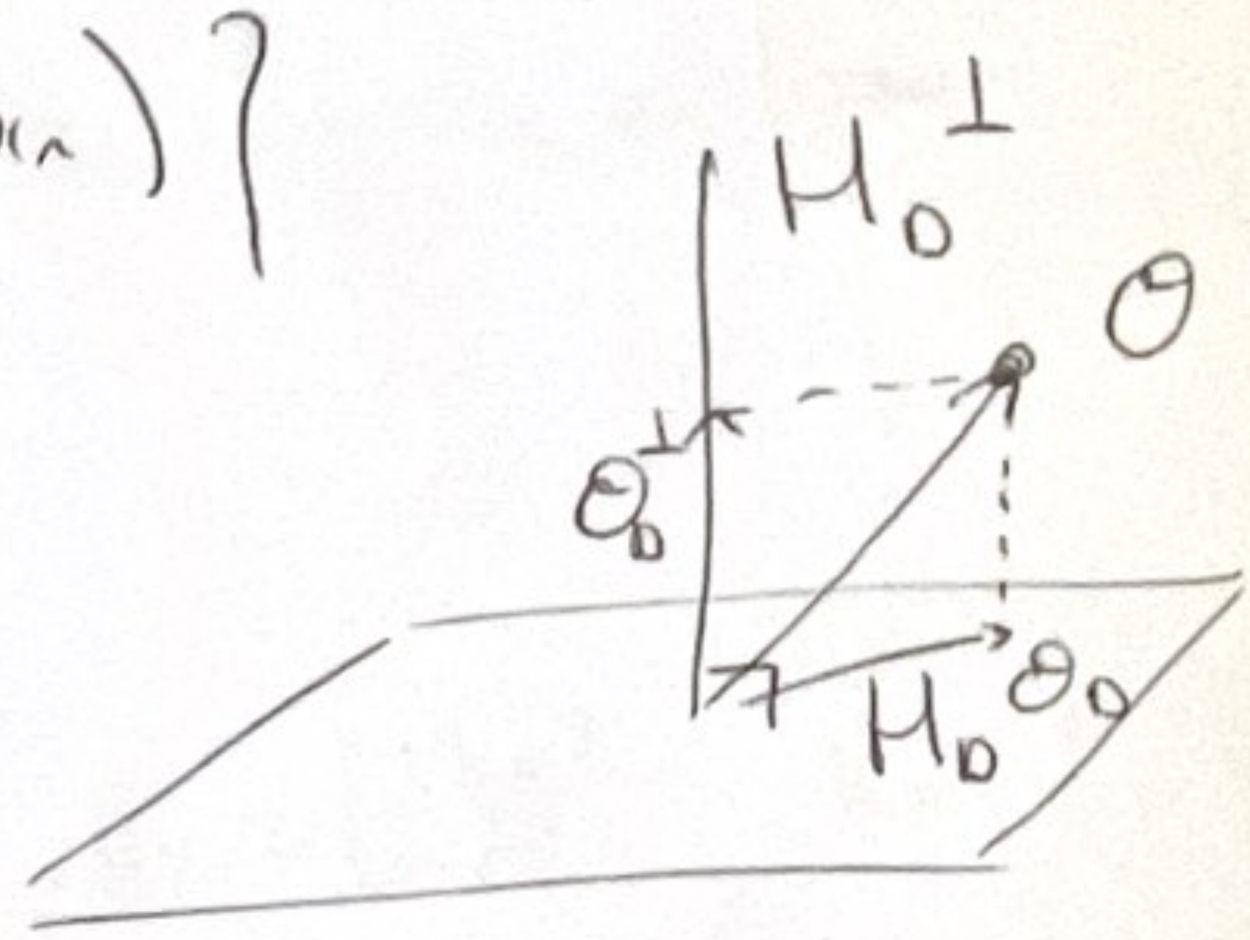
$$\forall x_1, x_2 \in \mathcal{X}, \; k(x_1, x_2) = \langle \phi(x_1), \phi(x_2) \rangle.$$

Donc, fixer un noyau pas un produit scalaire et une fonction ϕ n'est pas restrictif, il y a équivalence entre les deux.

3) Notons $H_D = \mathrm{Vect}\{\phi(x_1), \dots, \phi(x_n)\}$

Soit $\theta \in H$, $\theta = \theta_D + \theta_D^\perp$

$$\underset{H_D}{\theta_D} \in \quad \underset{H_D^\perp}{\theta_D^\perp} \in$$



Notons $M_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} \log\left(1 + e^{-y_i\langle\theta, \phi(x_i)\rangle}\right) + \|\theta\|^2$

Alors $M_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} \log\left(1 + e^{-y_i\langle\theta_D + \theta_D^\perp, \phi(x_i)\rangle}\right) + \|\theta_D + \theta_D^\perp\|^2$

$$= \frac{1}{n}\sum_{i=1}^{n} \log\left(1 + e^{\underbrace{-y_i\left(\langle\theta_D, \phi(x_i)\rangle + \langle\theta_D^\perp, \phi(x_i)\rangle\right)}_{\text{par linéarité}}}\right) + \underbrace{\|\theta_D\|^2 + \|\theta_D^\perp\|^2}_{\text{par Pythagore}}$$

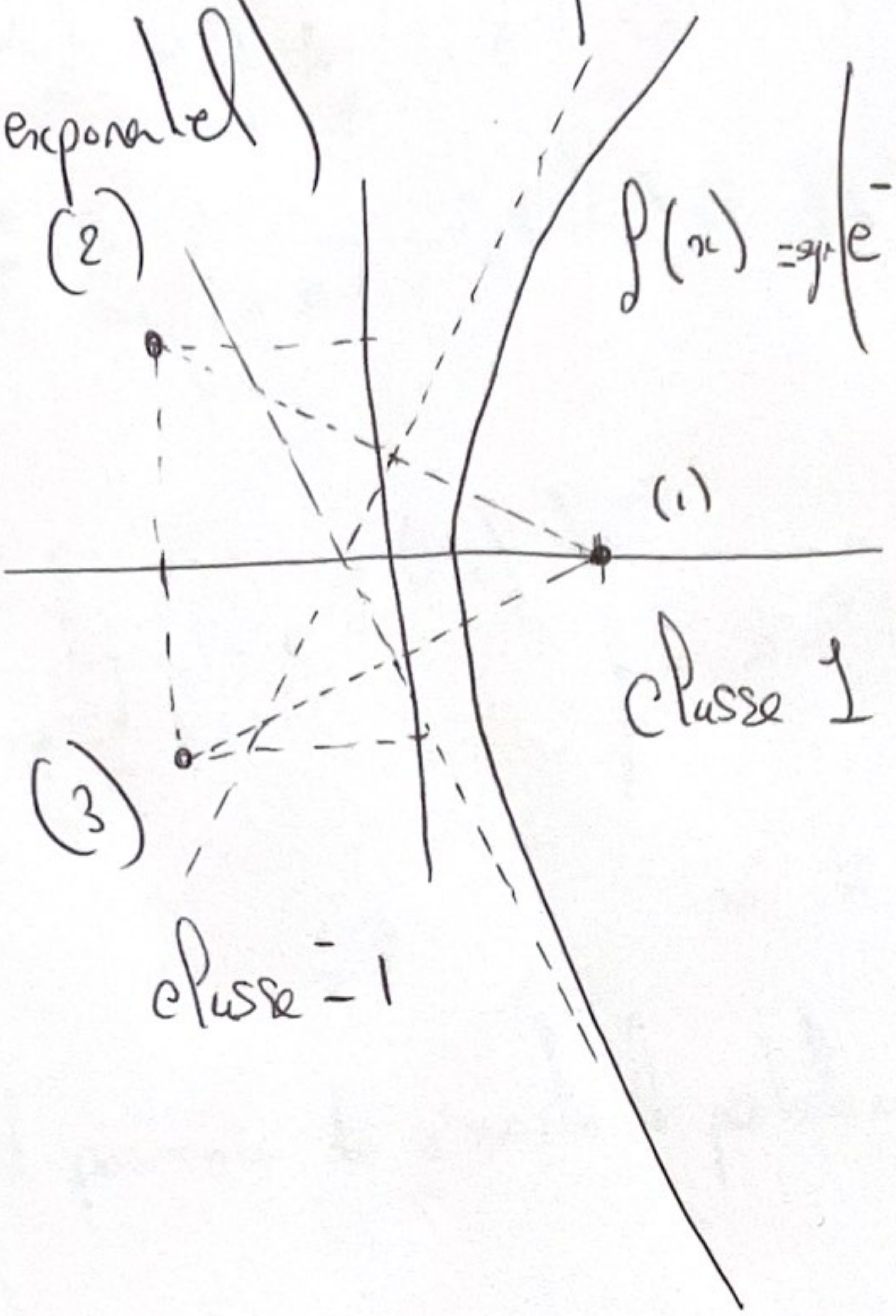
or, $\forall i, \langle\theta_D^\perp, \phi(x_i)\rangle = 0$ car $\theta_D^\perp \in \mathrm{Vect}(\phi(x_1), \dots, \phi(x_n))^\perp$.

donc $\boxed{M_n(\theta) = M_n(\theta) + \|\theta_D^\perp\|^2}$

Une solution optimale (si elle existe) a donc une composante dans l'orthogonal de H_D nulle. On peut se restreindre à chercher dans H_D.

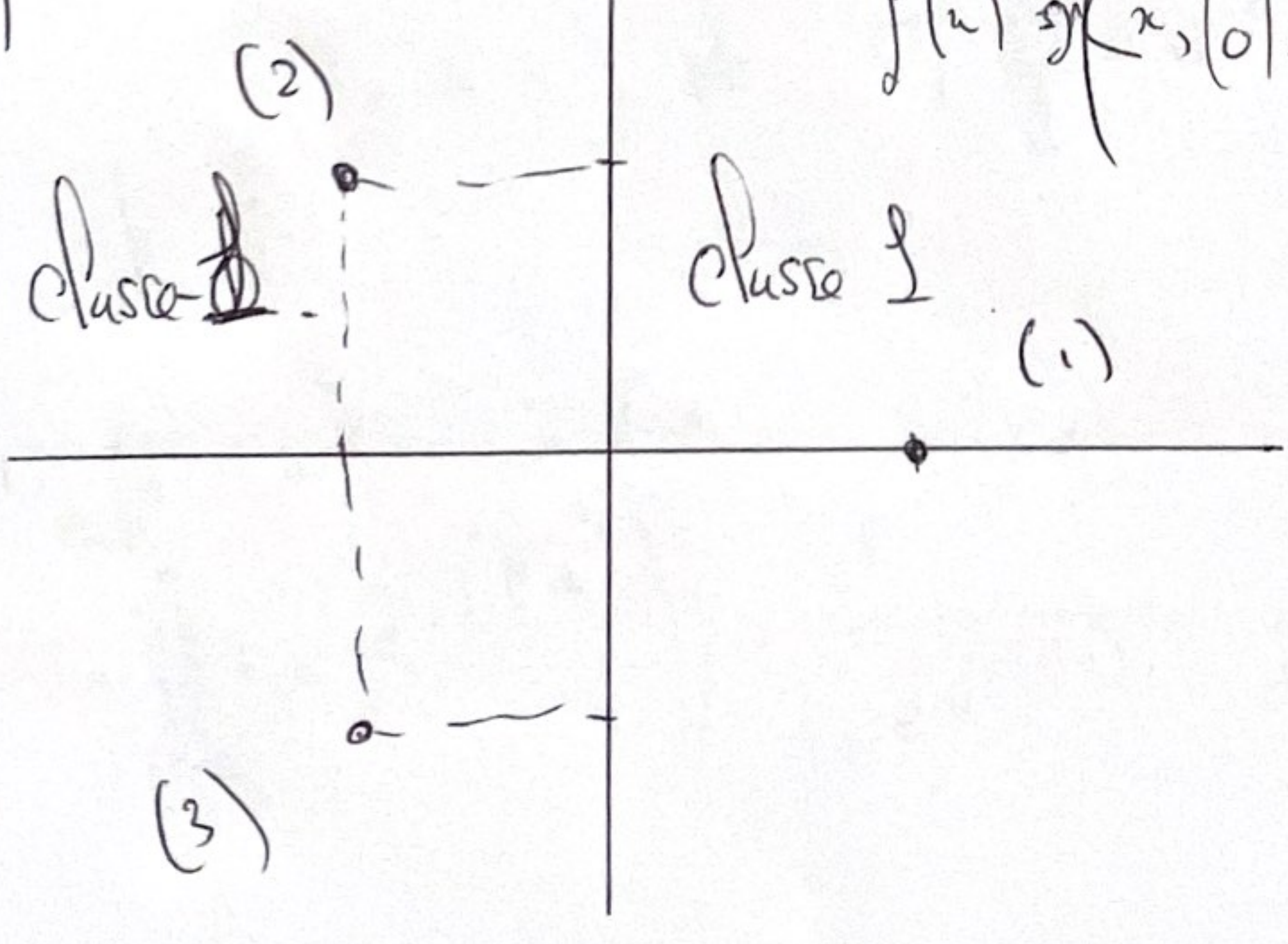
Sur H_D, le problème est fortement convexe (grâce au terme $\lambda\|\theta\|^2$). Il admet donc une unique solution.

4) (noyau exponentiel)



$$f(x) = \operatorname{sgn}\left(e^{-\|x - \binom{1}{0}\|^2} - e^{-\|x - \binom{-1}{1}\|^2} - e^{-\|x - \binom{-1}{-1}\|^2} \right)$$

(noyau linéaire)



$$f(x) = \operatorname{sgn}\left(\langle x, \binom{1}{0} \rangle - \langle x, \binom{-1}{1} \rangle - \langle x, \binom{-1}{-1} \rangle \right)$$

$$= \operatorname{sgn}\left(\langle x, \binom{3}{0} \rangle \right)$$

5) Changement de variables $\theta = \alpha_1 \phi(x_1) + \dots + \alpha_n \phi(x_n)$

$$M_n(\alpha_1, \dots, \alpha_n) = \frac{1}{n} \sum_{i=1}^{n} \log\left(1 + e^{-y_i \left(\sum_j \alpha_j \phi(x_j), \phi(x_i)\right)}\right) + \lambda \left\| \sum_j \alpha_j \phi(x_j) \right\|^2$$

$$= \frac{1}{n} \sum_{i=1}^{n} \log\left(1 + e^{-y_i \sum_j \alpha_j (\phi(x_j), \phi(x_i))}\right) + \lambda \left(\sum_i \alpha_i \phi(x_i), \sum_j \alpha_j \phi(x_j) \right)$$

$$= \frac{1}{n} \sum_{i=1}^{n} \log\left(1 + e^{-y_i (K\alpha)_i}\right) + \lambda \alpha^T K \alpha$$

où $K = \left(k(x_i, x_j)\right)_{i,j}$.

6) Il est possible de résoudre le problème par descente de gradient (stochastique).

Il faut être capable de calculer

$$\nabla_\alpha (\lambda \alpha^T K \alpha) \text{ et } \nabla_\alpha \left(\log\left(1 + e^{-y_i (K\alpha)_i}\right)\right) \forall i$$

ce qui donne

- $\nabla_\alpha (\lambda \alpha^T K \alpha) = 2\lambda K \alpha$
- $\nabla_\alpha \left(\log\left(1 + e^{-y_i (K\alpha)_i}\right)\right) = \dfrac{-e^{-y_i (K\alpha)_i}}{1 + e^{-y_i (K\alpha)_i}} \, y_i \begin{pmatrix} (\phi(x_1), \phi(x_i)) \\ \vdots \\ (\phi(x_n), \phi(x_i)) \end{pmatrix}$