

2021

# Ateliers d'optimisation

Sameh Najeh et Ali Saada

SUP'COM

01/09/2021

## Avant-propos :

Ce nouveau cours qui entre dans le cursus de formation de SUP'COM et qui s'intitule Ateliers d'optimisation, vient combler un vide qui se trouve entre certaines matières théoriques du tronc commun et ceux des différentes options. Il s'appuie en particulier sur les éléments d'enseignement Recherche Opérationnelle et Optimisation convexe. Son contenu, qui a pour but de mettre en valeur par la pratique différents résultats rencontrés dans les deux éléments en question, est formé de six ateliers :

- Atelier Support Vector Machine (SVM),
- Atelier Régression linéaire,
- Atelier pénalisation,
- Atelier Convolutional Neural Network (CNN)
- Atelier théorie des jeux 1,
- Atelier théorie des jeux 2.

Les élèves sont tenus à faire pour les quatre premiers ateliers un rapport qui renferme un travail bibliographique et un travail de mise en œuvre avec discussion des résultats. La note<sup>1</sup> finale est la moyenne des six notes obtenues.

---

1

- Les élèves sont tenus à respecter rigoureusement la date de remise des rapports sous peine d'être pénalisés dans leurs notes. Chaque jour de retard génère un handicap d'un demi-point.
- A part la valeur du contenu, la note tiendra compte de la forme.
- Nous conseillons les élèves de travailler en groupe mais les rapports sont remis individuellement. Leurs contenus est un gage de connaissance qui peut être vérifiée par un test écrit qui peut modifier la note du rapport.

## Table des matières

Avant-propos : .....	1
Atelier 1 : Mise en œuvre de la méthode Support Vector Machine.....	3
Présentation du problème .....	3
Méthode SVM linéaire de classification : .....	3
Correction.....	6
Méthode SVM linéaire de classification : .....	6
Délivrable.....	11
Atelier 2 : Mise en œuvre de La régression linéaire .....	12
Présentation du problème .....	12
Correction.....	14
Délivrable.....	18
Atelier 3 : Mise en œuvre de la méthode de pénalisation .....	19
Présentation du problème .....	19
Correction.....	21

## Atelier 1 : Mise en œuvre de la méthode Support Vector Machine

### Présentation du problème

Notations :

Soit  $n \in \mathbb{N}^*$  et soit  $\mathbf{x} \in \mathbb{R}^n$ . Nous notons par  $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  et  $\mathbf{x}^T = (x_1, \dots, x_n)$  où pour tout  $1 \leq i \leq n$ ,  $x_i \in \mathbb{R}$  est la  $i$ -ème composante de  $\mathbf{x}$ .

Pour  $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^n \times \mathbb{R}^n$ , nous notons par  $\langle \mathbf{x} | \mathbf{y} \rangle = \sum_{i=1}^n x_i y_i = \mathbf{x}^T \mathbf{y}$  le produit scalaire canonique de  $\mathbb{R}^n$  et par norme de  $\|\mathbf{x}\|_2 = \sqrt{\langle \mathbf{x} | \mathbf{x} \rangle}$ .

Notons que la distance entre deux sous-ensembles  $E_1$  et  $E_2$  de  $\mathbb{R}^n$  est  $d = \inf_{\mathbf{x}_1 \in E_1, \mathbf{x}_2 \in E_2} \|\mathbf{x}_1 - \mathbf{x}_2\|_2$ .

Nous rappelons aussi qu'un hyperplan  $\mathcal{H}$  de  $\mathbb{R}^n$  est un espace vectoriel de dimension  $n - 1$ . D'une manière général, si  $\mathcal{H}$  est un hyperplan, alors il existe  $\mathbf{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \in \mathbb{R}^n \setminus \{0\}$  orthogonal à tous les vecteurs de  $\mathcal{H}$ , autrement dit  $\mathcal{H} = \{\mathbf{x} \in \mathbb{R}^n / \mathbf{w}^T \mathbf{x} = 0\}$ .

Soit  $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$  une base de  $\mathbb{R}^n$  et  $\mathcal{R} = (O, \mathbf{e}_1, \dots, \mathbf{e}_n)$  le repère qui lui est associé. Nous rappelons qu'un hyperplan affine de direction  $\mathcal{H}$  et passant par le point  $M$  est  $\mathcal{H}_M = \{\mathbf{x} \in \mathbb{R}^n / (\mathbf{x} - \overrightarrow{OM}) \in \mathcal{H}\}$

1. On considère pour  $n = 2$  l'hyperplan d'équation  $2x_1 - 3x_2 = 0$ . Donner  $\mathbf{w} \in \mathbb{R}^2 \setminus \{0\}$  orthogonal à tous les vecteurs de cet hyperplan.
2. Soit  $\mathcal{R} = (O, \mathbf{e}_1, \mathbf{e}_2)$  un repère orthonormé et  $K$  le point de coordonnées  $(1, 2)_{\mathcal{R}}$ . Tracer sur une même figure  $\mathbf{w}$ ,  $\mathcal{H}_O$ ,  $\mathcal{H}_K$ , et l'hyperplan affine d'équation

$$2x_1 - 3x_2 - 8 = 0$$

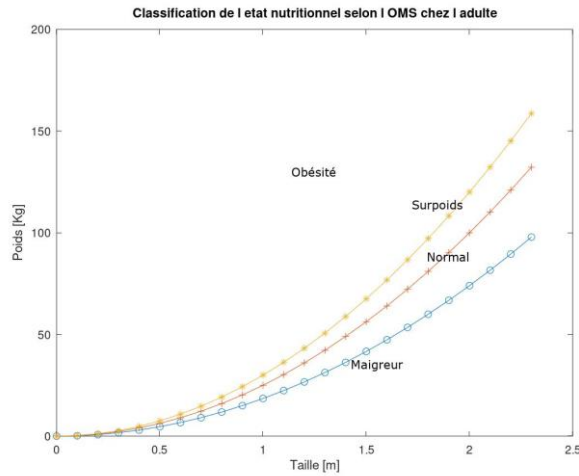
### Méthode SVM linéaire de classification :

Soit une population d'individus décrits par un nombre  $n \in \mathbb{N}^*$  de paramètres. L'objectif de la classification est de regrouper les individus selon des critères définis sur ces paramètres. Ceci revient à définir des frontières qui séparent entre les classes d'individus. Un exemple important est celui de la classification de l'état nutritionnel chez l'adulte selon l'OMS<sup>2</sup> :

Classification	IMC ( $Kg/m^2$ )	Risque
Maigreur (Dénutrition)	$< 18,5$	
Normal	$18,5 - 24,9$	
Surpoids	$25 - 29,9$	Modérément augmenté
Obésité	$> 30$	Nettement augmenté

<sup>2</sup> OMS : Organisation mondiale de la Santé <https://www.who.int/fr>

La figure ci-dessous décrit les frontières de cette classification :



Dans cette première partie nous nous proposons d'étudier la détermination de la frontière avec la méthode Support-Vector Machine (SVM) linéaire.

Soient quatre individus décrits par deux paramètres  $T$  et  $P$  selon le tableau suivant :

Individu	$T$	$P$	$IMC$
$I_1$	1,5	72	32
$I_2$	1,75	92	30,04
$I_3$	1,6	60	23,43
$I_4$	1,8	72	22,22

Les deux premiers sont hypertendus et les deux autres ne le sont pas. Nous affectons aux hypertendus la valeur 1 et aux autres la valeur  $-1$ .

3. Nous nous proposons de trouver les paramètres  $\mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in \mathbb{R}^2 \setminus \{\mathbf{0}\}$  et  $w_3 \in \mathbb{R}$  de la fonction frontière d'équation  $g(\mathbf{x}) = 0$ , où

$$\begin{aligned} g: \mathbb{R}^2 &\rightarrow \mathbb{R} \\ \mathbf{x} &\mapsto \mathbf{w}^T \mathbf{x} - w_3 \end{aligned}$$

et vérifie  $g(I_i) \geq 1$  pour  $i = 1$  et  $2$ ,  $g(I_i) \leq -1$  pour  $i = 3$  et  $4$  et telle que la distance entre l'ensemble des potentiellement hypertendus

$$HT_+ = \{\mathbf{x} \in \mathbb{R}^2 / g(\mathbf{x}) \geq 1\}$$

et l'ensemble des potentiellement non-hypertendus

$$HT_- = \{\mathbf{x} \in \mathbb{R}^2 / g(\mathbf{x}) \leq -1\}$$

soit **maximale**.

- Ecrire la contraintes  $g_i$  que doivent vérifier  $\mathbf{w}$  et  $w_3$  aux différents  $I_i$  pour  $i = 1, 2, 3$  et  $4$
- Montrer que la distance qui existe entre l'ensemble  $D_1 = \{\mathbf{x} \in \mathbb{R}^2 / g(\mathbf{x}) = 1\}$  et l'ensemble  $D_{-1} = \{\mathbf{x} \in \mathbb{R}^2 / g(\mathbf{x}) = -1\}$  est égale à  $\frac{2}{\|\mathbf{w}\|_2}$ .
- Calculer la distance entre  $O$ , l'origine du repère, et  $D_0 = \{\mathbf{x} \in \mathbb{R}^2 / g(\mathbf{x}) = 0\}$

- d. Dédurre de a. et de b. le problème de minimisation à résoudre pour déterminer  $W = \begin{pmatrix} \overline{w_1} \\ \overline{w_2} \\ \overline{w_3} \end{pmatrix}$ .
- e. Donner, en appliquant le théorème de Kuhn & Tucker<sup>3</sup>, le système à résoudre en explicitant les équations.
- f. Les contraintes peuvent-elles être simultanément non saturées. Justifier la réponse.
- g. Donner le nombre de contraintes maximales qui soient saturées simultanément. Justifier la réponse
- h. Donner les équations à résoudre si les contraintes 1, 2 et 3 sont saturées. (On ne demande pas de résoudre ces équations)
4. Tracer sur une même figure les points  $M_i / \overrightarrow{OM_i} = \mathbf{I}_i$ , pour  $i = 1, 2, 3$  et 4 et la frontière en supposant que les contraintes 1, 2 et 3 sont saturées.
5. **Question subsidiaire** : Donner une formulation plausible pour retrouver les frontières relatives à la classification OMS de la figure ci-dessus.

---

<sup>3</sup> **Théorème de Kuhn & Tucker.** Soit  $f \in \mathcal{C}(\mathbb{R}^n, \mathbb{R})$  et soient  $g_i \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$  pour  $i = 1, 2, \dots, p$ . On pose  $g = (g_1, g_2, \dots, g_p)^t$  et on considère  $K = \{x \in \mathbb{R}^n / g_i(x) \leq 0, i = 1, 2, \dots, p\}$  et on considère  $\bar{x} \in \mathbb{R}^n / f(\bar{x}) = \inf_{x \in K} f(x)$ . On suppose que  $f$  est différentiable en  $\bar{x}$  et que  $\{\nabla g_i(\bar{x})\}_{i \in I(\bar{x})}$  où  $I(\bar{x}) = \{i \in \{1, 2, \dots, p\} / g_i(\bar{x}) = 0\}$  est une famille libre, alors il existe  $\lambda = (\lambda_i)_{i \in I(\bar{x})}$  de réels positifs tels que

$$\nabla f(\bar{x}) + \sum_{i \in I(\bar{x})} \lambda_i \nabla g_i(\bar{x}) = 0$$

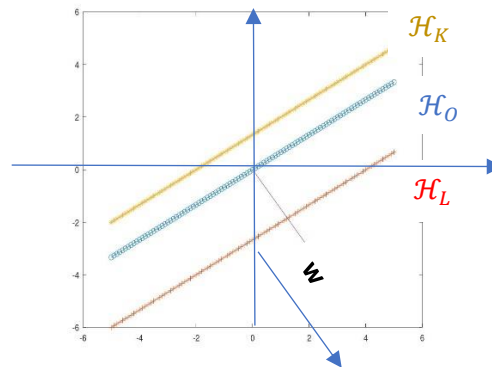
## Correction

1. On considère pour  $n = 2$  l'hyperplan d'équation  $2x_1 - 3x_2 = 0$ . Donner  $w \in \mathbb{R}^2 \setminus \{0\}$  orthogonal à tous les vecteurs de cet hyperplan.

Tout vecteur  $w = \alpha \begin{pmatrix} 2 \\ -3 \end{pmatrix}$  où  $\alpha \in \mathbb{R}^*$  est orthogonal à  $\{x \in \mathbb{R}^2 / 2x_1 - 3x_2 = 0\}$

2. Soit  $\mathcal{R} = (O, e_1, e_2)$  un repère orthonormé et  $K$  le point de coordonnées  $(1, 2)_{\mathcal{R}}$ . Tracer sur une même figure  $w$ ,  $\mathcal{H}_O$ ,  $\mathcal{H}_K$ , et l'hyperplan affine d'équation  $2x_1 - 3x_2 - 8 = 0$

$\mathcal{H}_K = \{x \in \mathbb{R}^2 / 2x_1 - 3x_2 + 4 = 0\}$ , et  $\{x \in \mathbb{R}^2 / 2x_1 - 3x_2 - 8 = 0\}$  est l'hyperplan  $\mathcal{H}_L$  où  $K$  est le point  $(4, 0)_{\mathcal{R}}$ . Nous avons alors la figure



Remarque : Les hyperplan de même direction restent parallèles

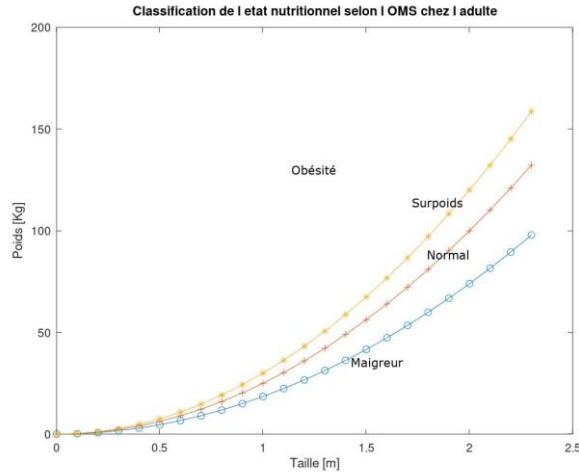
Méthode SVM linéaire de classification :

Soit une population d'individus décrits par un nombre  $n \in \mathbb{N}^*$  de paramètres. L'objectif de la classification est de regrouper les individus selon des critères définis sur ces paramètres. Ceci revient à définir des frontières qui séparent entre les classes d'individus. Un exemple important est celui de la classification de l'état nutritionnel chez l'adulte selon l'OMS<sup>4</sup> :

Classification	IMC ( $Kg/m^2$ )	Risque
Maigre (Dénutrition)	$< 18,5$	
Normal	$18,5 - 24,9$	
Surpoids	$25 - 29,9$	Modérément augmenté
Obésité	$> 30$	Nettement augmenté

La figure ci-dessous décrit les frontières de cette classification :

<sup>4</sup> OMS : Organisation mondiale de la Santé <https://www.who.int/fr>



Dans cette première partie nous nous proposons d'étudier la détermination de la frontière avec la méthode Support-Vector Machine (SVM) linéaire.

Soient quatre individus décrits par deux paramètres  $T$  et  $P$  selon le tableau suivant :

Individu	$T$	$P$
$I_1$	1,5	72
$I_2$	1,75	92
$I_3$	1,6	60
$I_4$	1,8	72

Les deux premiers sont hypertendus et les deux autres ne le sont pas. Nous affectons aux hypertendus la valeur 1 et aux autres la valeur  $-1$ .

3. Nous nous proposons de trouver les paramètres  $\mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in \mathbb{R}^2 \setminus \{\mathbf{0}\}$  et  $w_3 \in \mathbb{R}$  de la fonction frontière d'équation  $g(\mathbf{x}) = 0$ , où

$$\begin{aligned} g: \mathbb{R}^2 &\rightarrow \mathbb{R} \\ \mathbf{x} &\mapsto \mathbf{w}^T \mathbf{x} - w_3 \end{aligned}$$

et vérifie  $g(I_i) \geq 1$  pour  $i = 1$  et  $2$ ,  $g(I_i) \leq -1$  pour  $i = 3$  et  $4$  et telle que la distance entre l'ensemble des potentiellement hypertendus

$$HT_+ = \{\mathbf{x} \in \mathbb{R}^2 / g(\mathbf{x}) \geq 1\}$$

et l'ensemble des potentiellement non-hypertendus

$$HT_- = \{\mathbf{x} \in \mathbb{R}^2 / g(\mathbf{x}) \leq -1\}$$

soit maximale.

- i. Ecrire la contraintes  $g_i$  que doivent vérifier  $\mathbf{w}$  et  $w_3$  aux différents  $I_i$  pour  $i = 1, 2, 3$  et  $4$

Nous avons  $g(I_1) \geq 1$ ,  $g(I_2) \geq 1$ ,  $g(I_3) \leq -1$  et  $g(I_4) \leq -1$  ceci implique que nous avons  $\mathbf{w}^T I_1 - w_3 \geq 1$ ,  $\mathbf{w}^T I_2 - w_3 \geq 1$ ,  $\mathbf{w}^T I_3 - w_3 \leq -1$  et  $\mathbf{w}^T I_4 - w_3 \leq -1$  ou encore  $1 - \mathbf{w}^T I_1 + w_3 \leq 0$ ;  $1 - \mathbf{w}^T I_2 + w_3 \leq 0$ ,  $\mathbf{w}^T I_3 - w_3 + 1 \leq 0$  et  $\mathbf{w}^T I_4 - w_3 + 1 \leq 0$  ou encore  $g_1(\mathbf{w}, w_3) \leq 0$ ,  $g_2(\mathbf{w}, w_3) \leq 0$ ,  $g_3(\mathbf{w}, w_3) \leq 0$  et  $g_4(\mathbf{w}, w_3) \leq 0$  avec  $g_1(\mathbf{w}, w_3) = 1 - \mathbf{w}^T I_1 + w_3$ ,  $g_2(\mathbf{w}, w_3) = 1 - \mathbf{w}^T I_2 + w_3$ ,  $g_3(\mathbf{w}, w_3) = \mathbf{w}^T I_3 - w_3 + 1$  et  $g_4(\mathbf{w}, w_3) = \mathbf{w}^T I_4 - w_3 + 1$



- j. Montrer que la distance qui existe entre l'ensemble  $D_1 = \{x \in \mathbb{R}^2 / g(x) = 1\}$  et l'ensemble  $D_{-1} = \{x \in \mathbb{R}^2 / g(x) = -1\}$  est égale à  $\frac{2}{\|w\|_2}$ .

$$d(D_1, D_2) = \inf_{x_1 \in D_1, x_2 \in D_2} \|x_1 - x_2\|_2$$

Il s'agit donc de minimiser  $\|x_1 - x_2\|_2$  avec  $w^T x_1 - w_3 = 1$  et  $w^T x_2 - w_3 = -1$ . C'est un problème de minimisation avec deux contraintes d'égalité. Comme la norme est positive le problème devient à minimiser  $\frac{1}{2} \|x_1 - x_2\|_2^2$  sous les contraintes  $w^T x_1 - w_3 = 1$  et  $w^T x_2 - w_3 = -1$  ou encore, en faisant un changement de variable ( $y = x_1 - x_2$ ), de minimiser  $f(y) = \frac{1}{2} \|y\|_2^2$  sous les contraintes  $w^T x_1 - w_3 = 1$  et  $w^T (x_1 - y) - w_3 = -1 \Leftrightarrow w^T x_1 - w_3 - w^T y = -1 \Rightarrow w^T y - 2 = 0$ .

D'après le théorème de Lagrange, on sait qu'il existe  $\lambda \in \mathbb{R}$  tel que :

$$\begin{cases} y + \lambda w = 0 \\ w^T y - 2 = 0 \\ w^T x_1 - w_3 = 1 \end{cases} \Rightarrow \text{en multipliant la première équation par } w^T \text{ et en tenant compte}$$

de la deuxième on a  $\lambda = \frac{-2}{\|w\|_2^2} \Rightarrow y = \frac{2}{\|w\|_2^2} w$  et pour cette valeur, la distance est

$$\|y\|_2 = \sqrt{\langle y | y \rangle} = \sqrt{\left\langle \frac{2}{\|w\|_2^2} w \mid \frac{2}{\|w\|_2^2} w \right\rangle} = \frac{2}{\|w\|_2^2} \|w\|_2 = \frac{2}{\|w\|_2}$$

- k. Calculer la distance entre  $O$ , l'origine du repère, et  $D_0 = \{x \in \mathbb{R}^2 / g(x) = 0\}$

$d(O, D_2) = \inf_{x_1 \in D_1} \|x\|_2$ . Il s'agit donc de minimiser  $\|x\|_2$  avec  $w^T x - w_3 = 0$ , ou encore comme précédemment, Il s'agit donc de minimiser  $\frac{1}{2} \|x\|_2^2$  avec  $w^T x - w_3 = 0$ . Encore une fois par Lagrange, on sait qu'il existe  $\lambda \in \mathbb{R}$  tel que :

$$\begin{cases} x + \lambda w = 0 \\ w^T x - w_3 = 0 \end{cases}, \text{ et de la même manière on obtient } \lambda = \frac{-w_3}{\|w\|_2^2} \text{ et } \|x\|_2 = \sqrt{\langle x | x \rangle} = \sqrt{\left\langle \frac{w_3}{\|w\|_2^2} w \mid \frac{w_3}{\|w\|_2^2} w \right\rangle} = \frac{w_3}{\|w\|_2^2} \|w\|_2 = \frac{w_3}{\|w\|_2}$$

- l. Dédurre de a. et de b. le problème de minimisation à résoudre pour déterminer

$$W = \begin{pmatrix} \overline{w_1} \\ \overline{w_2} \\ \overline{w_3} \end{pmatrix}.$$

Comme le but est de maximiser la distance entre  $HT_+$  de frontière  $D_1$  et entre  $HT_-$  de frontière  $D_{-1}$  et que la distance entre ces deux frontières est égale à  $\frac{2}{\|w\|_2}$ , alors le

problème revient à minimiser  $f(w, w_3) = \frac{1}{2} \|w\|_2^2$  sous les contraintes  $g_1(w, w_3) \leq 0$ ,  $g_2(w, w_3) \leq 0$ ,  $g_3(w, w_3) \leq 0$  et  $g_4(w, w_3) \leq 0$  avec  $g_1(w, w_3) = 1 - w^T I_1 + w_3$ ,  $g_2(w, w_3) = 1 - w^T I_2 + w_3$ ,  $g_3(w, w_3) = w^T I_3 - w_3 + 1$  et  $g_4(w, w_3) = w^T I_4 - w_3 + 1$

- m. Donner, en appliquant le théorème de Kuhn & Tucker, le système à résoudre en explicitant les équations.

Le théorème de Kuhn et Tucker implique que si  $(\overline{w}, \overline{w_3})$  est un minimum, alors ils existent des réels positifs  $\lambda_1, \lambda_2, \lambda_3$  et  $\lambda_4$  tels que :

$$\begin{cases} \nabla f(\overline{w}, \overline{w_3}) + \sum_{i=1}^4 \lambda_i \nabla g_i(\overline{w}, \overline{w_3}) = 0 \\ \lambda_i g_i(\overline{w}, \overline{w_3}) = 0 \text{ pour } i = 1, 2, 3 \text{ et } 4 \\ g_i(\overline{w}, \overline{w_3}) \leq 0 \text{ pour } i = 1, 2, 3 \text{ et } 4 \end{cases}$$

$$\Leftrightarrow \begin{cases} \begin{pmatrix} w_1 \\ w_2 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} -I_{11} \\ -I_{12} \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} -I_{21} \\ -I_{22} \\ 1 \end{pmatrix} + \lambda_3 \begin{pmatrix} I_{31} \\ I_{32} \\ -1 \end{pmatrix} + \lambda_4 \begin{pmatrix} I_{41} \\ I_{42} \\ -1 \end{pmatrix} = 0 \\ \lambda_i g_i(\bar{w}, \bar{w}_3) = 0 \text{ pour } i = 1, 2, 3 \text{ et } 4 \\ g_i(\bar{w}, \bar{w}_3) \leq 0 \text{ pour } i = 1, 2, 3 \text{ et } 4 \end{cases}$$

où  $I_{i1}$  et  $I_{i2}$  sont les composantes de  $I_i$ .

- n. Les contraintes peuvent elles être simultanément non saturées. Justifier la réponse.

Si les contraintes sont simultanément non saturées alors, par les équations  $\lambda_i g_i(\bar{w}, \bar{w}_3) = 0$  pour  $i = 1, 2, 3$  et 4, on obtient que les multiplicateurs sont tous nuls et par conséquent que  $w_1 = w_2 = 0$  ce qui est impossible.

- o. Donner le nombre de contraintes maximales qui soient saturées simultanément. Justifier la réponse

La matrice  $[\nabla g_1(\bar{w}, \bar{w}_3), \nabla g_2(\bar{w}, \bar{w}_3), \nabla g_3(\bar{w}, \bar{w}_3), \nabla g_4(\bar{w}, \bar{w}_3)] =$   
 $\begin{bmatrix} -I_{11} & -I_{21} & I_{31} & I_{41} \\ -I_{12} & -I_{22} & I_{32} & I_{42} \\ 1 & 1 & -1 & -1 \end{bmatrix}$  est de rang égal à 3. Ceci implique qu'il existe au plus trois contraintes simultanément saturées.

- p. Donner les équations à résoudre si les contraintes 1, 2 et 3 sont saturées. (On ne demande pas de résoudre ces équations)

Comme les quatre contraintes ne peuvent être saturées simultanément. La saturation des contraintes 1, 2 et 3 implique que  $g_4(\bar{w}, \bar{w}_3) < 0$  et par conséquent  $\lambda_4 = 0$ . Les équations à résoudre sont alors :

$$\begin{aligned} w_1 - \lambda_1 I_{11} - \lambda_2 I_{21} + \lambda_3 I_{31} &= 0 \\ w_2 - \lambda_1 I_{12} - \lambda_2 I_{22} + \lambda_3 I_{32} &= 0 \\ \lambda_1 + \lambda_2 - \lambda_3 &= 0 \\ -w_1 I_{11} - w_2 I_{12} + w_3 + 1 &= 0 \\ -w_1 I_{21} - w_2 I_{22} + w_3 + 1 &= 0 \\ w_1 I_{31} + w_2 I_{32} - w_3 + 1 &= 0 \\ w_1 I_{41} + w_2 I_{42} - w_3 + 1 &< 0 \\ \lambda_1 \geq 0; \lambda_2 \geq 0; \lambda_3 \geq 0 \end{aligned}$$

4. Tracer sur une même figure les points  $M_i / \overrightarrow{OM_i} = I_i$ , pour  $i = 1, 2, 3$  et 4 et la frontière en supposant que les contraintes 1, 2 et 3 sont saturées.

1<sup>ère</sup> méthode pour le calcul de  $\bar{w}$  et de  $\bar{w}_3$  :

Les équations 4, 5 et 6 de la question précédente donnent un système

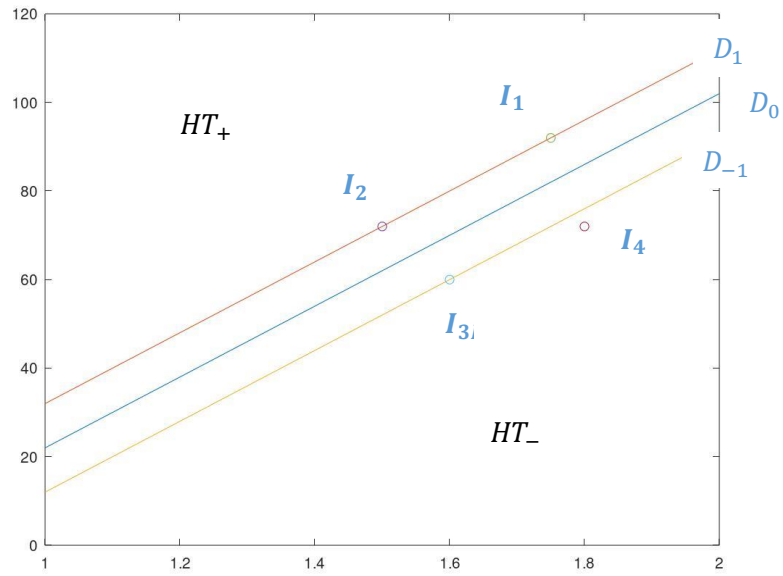
$$\begin{bmatrix} -I_{11} & -I_{12} & 1 \\ -I_{21} & -I_{22} & 1 \\ I_{31} & I_{32} & -1 \end{bmatrix} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix} \text{ qui peut être inversé pour donner } \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} -8 \\ 0.1 \\ -\frac{29}{5} \end{pmatrix}.$$

2<sup>ème</sup> méthode pour le calcul de  $\bar{w}$  et de  $\bar{w}_3$  :

Si les contraintes 1 et 2 sont saturées ceci implique que la droite  $D_1$  est de direction le vecteur  $I_2 - I_1 = \begin{pmatrix} 0,25 \\ 20 \end{pmatrix}$  où encore que  $D_1 = \{x = (x_1, x_2) \in \mathbb{R}^2 / w^T x - w_3 = 1\}$  avec  $w =$

$$\alpha \begin{pmatrix} -20 \\ 0,25 \end{pmatrix}; \quad \alpha \in \mathbb{R}^* \text{ et } w_3 = w^T I_1 - 1 = w^T I_2 - 1 = (-20 \times 1,5\alpha + 0,25 \times 72\alpha) - 1 = (-20 \times 1,75\alpha + 0,25 \times 92\alpha) - 1 = -12\alpha - 1. \quad \alpha \text{ est alors déterminé en écrivant que } g_4(w, w_3) = 0, \text{ autrement dit } w^T I_3 - w_3 = -1 \Rightarrow (-20 \times 1,6\alpha + 0,25 \times 60\alpha) + 1 + 12\alpha = -1 \Rightarrow \alpha = \frac{2}{5} \Rightarrow w = \begin{pmatrix} -8 \\ 0,1 \end{pmatrix} \text{ et } w_3 = -\frac{29}{5}$$

Ceci nous donne la figure suivante :



Remarque : Cette solution n'est pas bonne, car  $w = \begin{pmatrix} -8 \\ 0,1 \end{pmatrix}$  et  $w_3 = -\frac{29}{5}$  font que  $\lambda_2 < 0$

5. **Question subsidiaire** : Donner une formulation plausible pour retrouver les frontières relatives à la classification OMS de la figure ci-dessus.

## Délivrable

Un rapport contenant :

- 1-** Une bibliographie sur la méthode
- 2-** La Programmation de la méthode
- 3-** Une discussion sur les résultats de la programmation

## Atelier 2 : Mise en œuvre de La régression linéaire

### Présentation du problème

Soit  $p \in \mathbb{N}^*$  et soit  $(x, y) \in \mathbb{R}^p \times \mathbb{R}^p$ . On note par  $\langle x, y \rangle = \sum_{i=1}^p x_i y_i$  le produit scalaire canonique sur  $\mathbb{R}^p$  et par  $\|x\| = \left(\sum_{i=1}^p x_i^2\right)^{\frac{1}{2}}$  la norme euclidienne correspondante.

Soit  $(m, n) \in \mathbb{N}^* \times \mathbb{N}^*$  et  $\mathcal{M}_{n \times m}(\mathbb{R})$  l'ensemble des matrices à  $n$  lignes et  $m$  colonnes à coefficients réels. Soit  $A \in \mathcal{M}_{n \times m}(\mathbb{R})$ . On note par  $A^T \in \mathcal{M}_{m \times n}(\mathbb{R})$  la transposée de  $A$  et on définit son noyau par  $\text{Ker} A = \{x \in \mathbb{R}^m / Ax = 0_{\mathbb{R}^n}\}$

1. Exemple 1 : Dans le plan  $(s, t)$ , on cherche la droite d'équation  $t = \alpha + \beta s$  qui passe par les points  $(0,1), (1,9), (3,9)$  et  $(4,21)$ 
  - a) Montrer que si cette droite existait, le vecteur  $\begin{bmatrix} \alpha \\ \beta \end{bmatrix}$  serait solution d'un système  $Ax = b$ . On donnera explicitement la matrice  $A$  et le vecteur  $b$ .
  - b) Montrer qu'une telle droite n'existe pas.

Dans la suite du problème on va trouver la droite qui passe le « plus près » possible de ces quatre points, au sens de la norme euclidienne.

2. Exemple 2 : On cherche maintenant à déterminer les coefficients  $\alpha, \beta$  et  $\gamma$  d'une fonction linéaire  $T$  de  $\mathbb{R}^3$  dans  $\mathbb{R}$ , dont on ne connaît la valeur qu'en deux points :  $T(1,1,1) = 3$  et  $T(0,1,1) = 2$ .
  - a) Montrer que les coefficients  $\alpha, \beta$  et  $\gamma$  s'ils existent, satisfont un système linéaire  $Ax = b$ . On donnera explicitement la matrice  $A$  et le vecteur  $b$ .
  - b) Montrer qu'il existe une infinité de solutions au système  $Ax = b$ .

Dans la suite du problème, on va trouver les coefficients  $\alpha, \beta$  et  $\gamma$  qui donnent un vecteur  $x$  de norme euclidienne minimale.

On considère maintenant une matrice  $A \in \mathcal{M}_{n \times m}(\mathbb{R})$  et  $b \in \mathbb{R}^n$ , et on veut résoudre dans un sens aussi « satisfaisant » que possible le système linéaire

$$Ax = b$$

avec  $x \in \mathbb{R}^m$  lorsque  $m \neq n$  ou lorsque  $m = n$  mais que  $A$  n'est pas inversible. Soit  $f$  la fonction définie de  $\mathbb{R}^m$  dans  $\mathbb{R}$  par  $f(x) = \|Ax - b\|^2$ . On cherche à minimiser  $f$ , c'est-à-dire à trouver  $\bar{x} \in \mathbb{R}^m$  tel que

$$f(\bar{x}) = \min_{x \in \mathbb{R}^m} f(x) \quad (1)$$

3. Soit  $E$  un sous-espace vectoriel de  $\mathbb{R}^m$  tel que  $\mathbb{R}^m = E + \text{Ker} A$ .
  - a) Montrer que  $f(z) \rightarrow +\infty$  lorsque  $\|z\| \rightarrow +\infty$  et  $z \in E$ ,
  - b) Montrer que  $f$  est strictement convexe de  $E$  dans  $\mathbb{R}$ .
  - c) En déduire qu'il existe un unique  $\bar{z} \in E$  tel que  $f(\bar{z}) \leq f(z) \forall z \in E$
4. Soit  $X_b = \{\bar{z} + y, y \in \text{Ker} A\}$ , où  $\bar{z}$  est défini à la question précédente. Montrer que  $X_b$  est égal à l'ensemble des solutions du problème de minimisation (1)
5. Montrer que  $x \in X_b \Leftrightarrow A^T Ax = A^T b$ .

On appelle système d'équations normales, le système  $A^T Ax = A^T b$ .

6. Ecrire les équations normales dans le cas de l'exemple de la question 1. et en déduire l'équation de la droite obtenue par moindres carrés, i.e., par résolution de (1). Tracer les quatre points donnés à la question 1. Et la droite obtenue sur un graphique.
7. Ecrire les équations normales dans le cas de l'exemple de la question 2. et vérifier que le système obtenu n'est pas inversible.
8. Pour  $y \in \text{Ker}A$ , on pose  $g(y) = \|y + \bar{z}\|^2$ , où  $\bar{z}$  est la solution définie à la question 3. Montrer qu'il existe un unique  $\bar{y} \in \text{Ker}A$ , tel que  $g(\bar{y}) \leq g(y) \forall y \in \text{Ker}A$ . En déduire qu'il existe un unique  $\bar{x} \in X_b$  tel que  $\|\bar{x}\|^2 \leq \|x\|^2 \forall x \in X_b$ . On appelle  $\bar{x}$  pseudo-solution de (1).
9. Calculer  $\bar{x}$  dans le cas des questions 1. Et 2.

Dans la suite du problème, on considère, pour  $\varepsilon > 0$  fixé, une version pénalisée du système (1). On introduit la fonction  $f_\varepsilon(x) = \|x\|^2 + \frac{1}{\varepsilon} \|A^T Ax - A^T b\|^2$ , et on cherche à trouver  $x_\varepsilon$  solution du problème de minimisation :

$$f_\varepsilon(x_\varepsilon) \leq f_\varepsilon(x), \forall x_\varepsilon \in \mathbb{R}^m \quad (2)$$

10. Montrer que le problème (2) possède une unique solution  $x_\varepsilon$ .
11. Calculer  $\nabla f_\varepsilon(x)$  et en déduire l'équation satisfaite par  $x_\varepsilon$ .
12. Montrer que  $x_\varepsilon$  converge vers  $\bar{x}$  quand  $\varepsilon \rightarrow 0$ .

## Correction

Soit  $p \in \mathbb{N}^*$  et soit  $(x, y) \in \mathbb{R}^p \times \mathbb{R}^p$ . On note par  $\langle x, y \rangle = \sum_{i=1}^p x_i y_i$  le produit scalaire canonique sur  $\mathbb{R}^p$  et par  $\|x\| = (\sum_{i=1}^p x_i^2)^{\frac{1}{2}}$  la norme euclidienne correspondante.

Soit  $(m, n) \in \mathbb{N}^* \times \mathbb{N}^*$  et  $\mathcal{M}_{n \times m}(\mathbb{R})$  l'ensemble des matrices à  $n$  lignes et  $m$  colonnes à coefficients réels. Soit  $A \in \mathcal{M}_{n \times m}(\mathbb{R})$ . On note par  $A^T \in \mathcal{M}_{m \times n}(\mathbb{R})$  la transposée de  $A$  et on définit son noyau par  $\text{Ker} A = \{x \in \mathbb{R}^m / Ax = 0_{\mathbb{R}^n}\}$

1. Exemple 1 : Dans le plan  $(s, t)$ , on cherche la droite d'équation  $t = \alpha + \beta s$  qui passe par les points  $(0, 1)$ ,  $(1, 9)$ ,  $(3, 9)$  et  $(4, 21)$

- c) Montrer que si cette droite existait, le vecteur  $\begin{bmatrix} \alpha \\ \beta \end{bmatrix}$  serait solution d'un système  $Ax = b$ . On donnera explicitement la matrice  $A$  et le vecteur  $b$ .

Le vecteur  $\begin{bmatrix} \alpha \\ \beta \end{bmatrix}$  vérifie  $\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 1 \\ 9 \\ 9 \\ 21 \end{bmatrix}$ . Si on prenait les deux premiers points

Montrer qu'une telle droite n'existe pas.

$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 1 \\ 9 \end{bmatrix} \Leftrightarrow \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 9 \end{bmatrix} = \begin{bmatrix} 1 \\ 8 \end{bmatrix} \Rightarrow$  en considérant le troisième point  $9 = 1 + 8 * 3 = 25$  impossible. Donc il n'existe pas de droite qui passe par tous ces points

Dans la suite du problème on va trouver la droite qui passe le « plus près » possible de ces quatre points, au sens de la norme euclidienne.

2. Exemple 2 : On cherche maintenant à déterminer les coefficients  $\alpha, \beta$  et  $\gamma$  d'une fonction linéaire  $T$  de  $\mathbb{R}^3$  dans  $\mathbb{R}$ , dont on ne connaît la valeur qu'en deux points :  $T(1, 1, 1) = 3$  et  $T(0, 1, 1) = 2$ .

- c) Montrer que les coefficients  $\alpha, \beta$  et  $\gamma$  s'ils existent, satisfont un système linéaire  $Ax = b$ . On donnera explicitement la matrice  $A$  et le vecteur  $b$ .

Si elle une telle application linéaire existait on aurait  $\begin{bmatrix} \alpha & \beta & \gamma \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 3$  et

$$\begin{bmatrix} \alpha & \beta & \gamma \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = 2 \Rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

- d) Montrer qu'il existe une infinité de solutions au système  $Ax = b$ .

$$\begin{cases} \alpha + \beta + \gamma = 3 \\ \beta + \gamma = 2 \end{cases} \Rightarrow (\alpha, \beta, \gamma) \in \{(1, 0, 2) + \beta(0, 1, -1), \beta \in \mathbb{R}\}$$

Dans la suite du problème, on va trouver les coefficients  $\alpha, \beta$  et  $\gamma$  qui donnent un vecteur  $x$  de norme euclidienne minimale.

On considère maintenant une matrice non nulle  $A \in \mathcal{M}_{n \times m}(\mathbb{R})$  et  $b \in \mathbb{R}^n$ , et on veut résoudre dans un sens aussi « satisfaisant » que possible le système linéaire

$$Ax = b$$

avec  $x \in \mathbb{R}^m$  lorsque  $m \neq n$  ou lorsque  $m = n$  mais que  $A$  n'est pas inversible. Soit  $f$  la fonction définie de  $\mathbb{R}^m$  dans  $\mathbb{R}$  par  $f(x) = \|Ax - b\|^2$ . On cherche à minimiser  $f$ , c'est-à-dire à trouver  $\bar{x} \in \mathbb{R}^m$  tel que

$$f(\bar{x}) = \min_{x \in \mathbb{R}^m} f(x) \quad (1)$$

3. Soit  $E$  un sous-espace vectoriel de  $\mathbb{R}^m$  tel que  $\mathbb{R}^m = E + \text{Ker}A$ .

d) Montrer que  $f(z) \rightarrow +\infty$  lorsque  $\|z\| \rightarrow +\infty$  et  $z \in E$ ,

$$f(z) = \|Az - b\|^2 = \langle Az - b, Az - b \rangle = \langle Az, Az \rangle - 2\langle Az, b \rangle + \langle b, b \rangle \\ = \langle A^T A z, z \rangle - 2\langle Az, b \rangle + \langle b, b \rangle$$

$A^T A$  est symétrique définie positive. Si  $z \notin \text{Ker}A$  on a :

$$f(z) = \langle A^T A z, z \rangle - 2\langle Az, b \rangle + \langle b, b \rangle = \|z\|^2 \left( \langle A^T A \frac{z}{\|z\|}, \frac{z}{\|z\|} \rangle - 2 \langle A \frac{z}{\|z\|}, b \rangle + \frac{1}{\|z\|^2} \langle b, b \rangle \right) \xrightarrow{\|z\| \rightarrow +\infty} +\infty$$

e) Montrer que  $f$  est strictement convexe de  $E$  dans  $\mathbb{R}$ .

$A^T A$  est symétrique définie positive donc  $f$  est strictement convexe sur  $E$  (cours)

f) En déduire qu'il existe un unique  $\bar{z} \in E$  tel que  $f(\bar{z}) \leq f(z) \forall z \in E$

a) Et b) impliquent c) (cours)

4. Soit  $X_b = \{\bar{z} + y, y \in \text{Ker}A\}$ , où  $\bar{z}$  est défini à la question précédente. Montrer que  $X_b$  est égal à l'ensemble des solutions du problème de minimisation (1)

Soit  $x = \bar{z} + y$ , avec  $y \in \text{Ker}A \Rightarrow f(x) = \|A(\bar{z} + y) - b\|^2 = \|A\bar{z} - b\|^2 \leq \|Az - b\|^2 \forall z \in E \Rightarrow f(x) \leq \|A(z + y) - b\|^2 \forall (z, y) \in E \times \text{Ker}A = \mathbb{R}^m \Rightarrow x$  est solution de (1).  
Réciproquement, soit  $x \in E \times \text{Ker}A = \mathbb{R}^m$  alors il existe un couple  $(\tilde{z}, \tilde{y}) \in E \times \text{Ker}A / x = \tilde{z} + \tilde{y}$ .  $x$  est solution de (1)  $\Rightarrow f(x) = f(\tilde{z}) \leq f(z) \forall z \in E \subset \mathbb{R}^m \Rightarrow \tilde{z} = \bar{z}$

5. Montrer que  $x \in X_b \Leftrightarrow A^T A x = A^T b$ .

On a vu que  $f(z) = \langle A^T A z, z \rangle - 2\langle Az, b \rangle + \langle b, b \rangle = \langle A^T A z, z \rangle - 2\langle A^T b, z \rangle + \langle b, b \rangle$ , avec  $A^T A$  est symétrique définie positive sur  $E$ . On sait par le cours que le minimum de  $f$  est unique et est atteint en  $\bar{z}$  tel que  $\nabla f(\bar{z}) = 2(A^T A \bar{z} - A^T b) = 0$ .

Ainsi, pour  $x \in X_b$ ,  $x = \bar{z} + y$  avec  $y \in \text{Ker}A \Rightarrow A^T A x - A^T b = A^T A(\bar{z} + y) - A^T b = A^T A \bar{z} - A^T b = 0$

Réciproquement soit  $x \in \mathbb{R}^m \Rightarrow \exists (z, y) \in E \times \text{Ker}A / x = z + y$  et  $A^T A x - A^T b = 0 \Rightarrow A^T A z - A^T b = 0 \Rightarrow z = \bar{z}$  par unicité de  $\bar{z}$

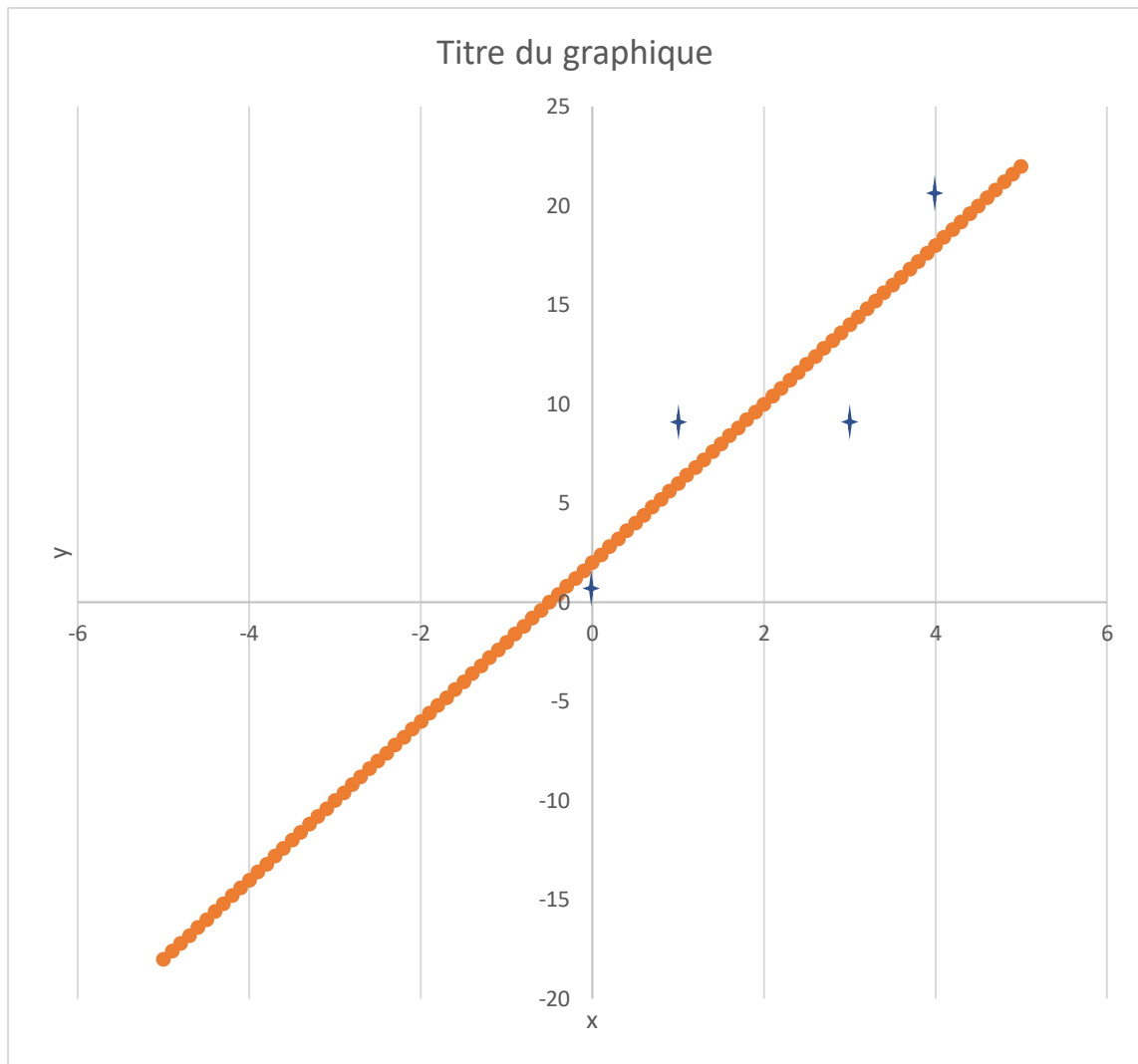
On appelle système d'équations normales, le système  $A^T A x = A^T b$ .

6. Ecrire les équations normales dans le cas de l'exemple de la question 1. et en déduire l'équation de la droite obtenue par moindres carrés, i.e., par résolution de (1). Tracer les quatre points donnés à la question 1. Et la droite obtenue sur un graphique.

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 9 \\ 9 \\ 21 \end{bmatrix} = \begin{bmatrix} 40 \\ 120 \end{bmatrix} \\ \Rightarrow \begin{bmatrix} 4 & 8 \\ 8 & 26 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 40 \\ 120 \end{bmatrix} \Rightarrow \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \frac{1}{40} \begin{bmatrix} 26 & -8 \\ -8 & 4 \end{bmatrix} \begin{bmatrix} 40 \\ 120 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

Et la droite est d'équation  $t = 2 + 4s$





7. Ecrire les équations normales dans le cas de l'exemple de la question 2. et vérifier que le système obtenu n'est pas inversible.

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \\ 5 \end{bmatrix}$$

Les lignes 2 et 3 sont identiques donc le système n'est pas inversible

8. Pour  $y \in \text{Ker} A$ , on pose  $g(y) = \|y + \bar{z}\|^2$ , où  $\bar{z}$  est la solution définie à la question 3. Montrer qu'il existe un unique  $\bar{y} \in \text{Ker} A$ , tel que  $g(\bar{y}) \leq g(y) \forall y \in \text{Ker} A$ . En déduire qu'il existe un unique  $\bar{x} \in X_b$  tel que  $\|\bar{x}\|^2 \leq \|x\|^2 \forall x \in X_b$ . On appelle  $\bar{x}$  pseudo-solution de (1).

$$\text{Ker} A = \left\{ y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \in \mathbb{R}^3 / Ay = 0 \right\} \Rightarrow \begin{cases} y_1 + y_2 + y_3 = 0 \\ y_2 + y_3 = 0 \end{cases} \Rightarrow \text{Ker} A = \left\{ y = t \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, t \in \mathbb{R} \right\} \Rightarrow f(y) = 2|t|^2 \text{ donc le minimum est obtenu pour } y = 0. \text{ Soit } E = \text{verct} \left( \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right). \text{ On a } \mathbb{R}^3 = E + \text{Ker} A.$$

9. Calculer  $\bar{x}$  dans le cas des questions 1. Et 2.

Dans la suite du problème, on considère, pour  $\varepsilon > 0$  fixé, une version pénalisée du système (1). On introduit la fonction  $f_\varepsilon(x) = \|x\|^2 + \frac{1}{\varepsilon} \|A^T Ax - A^T b\|^2$ , et on cherche à trouver  $x_\varepsilon$  solution du problème de minimisation :

$$f_\varepsilon(x_\varepsilon) \leq f_\varepsilon(x), \forall x_\varepsilon \in \mathbb{R}^m \quad (2)$$

**10. Montrer que le problème (2) possède une unique solution  $x_\varepsilon$ .**

On a  $f_\varepsilon$  est continue,  $\lim_{\|x\| \rightarrow +\infty} f_\varepsilon(x) = +\infty$  et  $f_\varepsilon$  strictement convexe, d'où l'existence et l'unicité

**11. Calculer  $\nabla f_\varepsilon(x)$  et en déduire l'équation satisfaite par  $x_\varepsilon$ .**

$$\begin{aligned} f_\varepsilon(x+h) &= \|x+h\|^2 + \frac{1}{\varepsilon} \|A^T A(x+h) - A^T b\|^2 \\ &= \|x\|^2 + 2\langle x, h \rangle + \|h\|^2 + \frac{1}{\varepsilon} \|A^T Ax - A^T b + A^T Ah\|^2 \\ &= f_\varepsilon(x) + 2\langle x, h \rangle + \frac{2}{\varepsilon} \langle A^T A(A^T Ax - A^T b), h \rangle + \|h\|^2 + \frac{1}{\varepsilon} \|A^T Ah\|^2 \end{aligned}$$

$$\text{Et } \nabla f_\varepsilon(x) = 2x + \frac{2}{\varepsilon} A^T A(A^T Ax - A^T b)$$

**12. Montrer que  $x_\varepsilon$  converge vers  $\bar{x}$  quand  $\varepsilon \rightarrow 0$ .**

$$x_\varepsilon \text{ est tel que } 2x_\varepsilon + \frac{2}{\varepsilon} A^T A(A^T Ax_\varepsilon - A^T b) = 0 \Rightarrow 2\varepsilon x_\varepsilon + 2A^T A(A^T Ax_\varepsilon - A^T b) = 0$$

## Délivrable

Un rapport contenant :

- 1-** Une bibliographie sur la méthode
- 2-** La Programmation de la méthode
- 3-** Une discussion sur les résultats de la programmation

## Atelier 3 : Mise en œuvre de la méthode de pénalisation

### Présentation du problème

Soit  $f \in \mathcal{C}(\mathbb{R}^n, \mathbb{R})$  strictement convexe vérifiant  $\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty$  et soit  $K$  un sous ensemble non vide, convexe et fermé de  $\mathbb{R}^n$ . Soit  $\varphi \in \mathcal{C}(\mathbb{R}^n, [0, +\infty])$  telle que  $\varphi(x) = 0$  si et seulement si  $x \in K$ . Pour  $k \in \mathbb{N}$ , on définit la fonction

$$\begin{aligned} f_k: \mathbb{R}^n &\rightarrow \mathbb{R} \\ x &\mapsto f(x) + k\varphi(x) \end{aligned}$$

1. Montrer qu'il existe au moins un élément  $\bar{x}_k \in \mathbb{R}^n$  tel que  $f_k(\bar{x}_k) = \inf_{x \in \mathbb{R}^n} f_k(x)$  et qu'il existe un unique élément  $\bar{x}_K \in K$  tel que  $f(\bar{x}_K) = \inf_{x \in K} f(x)$ .

2. Montrer que pour tout  $k \in \mathbb{N}$ ,

$$f(\bar{x}_k) \leq f_k(\bar{x}_k) \leq f(\bar{x}_K)$$

3. En déduire qu'il existe une sous-suite  $(\bar{x}_{k_i})_{i \in \mathbb{N}}$  et  $\tilde{x} \in K$  tels que  $\bar{x}_{k_i} \xrightarrow{i \rightarrow +\infty} \tilde{x}$ .

4. Montrer que  $\tilde{x} = \bar{x}_K$ . En déduire que toute la suite  $(\bar{x}_k)_{k \in \mathbb{N}}$  converge vers  $\bar{x}_K$ .

5. Déduire de ces questions un algorithme (dit "de pénalisation") de résolution du problème de minimisation :

$$\begin{cases} \text{Trouver} & \bar{x}_K \in K \\ f(\bar{x}_K) \leq f(x) & \forall x \in K \end{cases}$$

en donnant un exemple de fonction  $\varphi$ .



## Correction

Soit  $f \in \mathcal{C}(\mathbb{R}^n, \mathbb{R})$  strictement convexe vérifiant  $\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty$  et soit  $K$  un sous ensemble non vide, convexe et fermé de  $\mathbb{R}^n$ . Soit  $\varphi \in \mathcal{C}(\mathbb{R}^n, [0, +\infty])$  telle que  $\varphi(x) = 0$  si et seulement si  $x \in K$ . Pour  $k \in \mathbb{N}$ , on définit la fonction

$$\begin{aligned} f_k: \mathbb{R}^n &\rightarrow \mathbb{R} \\ x &\mapsto f(x) + k\varphi(x) \end{aligned}$$

1. Montrer qu'il existe au moins un élément  $\bar{x}_k \in \mathbb{R}^n$  tel que  $f_k(\bar{x}_k) = \inf_{x \in \mathbb{R}^n} f_k(x)$  et qu'il existe un unique élément  $\bar{x}_K \in K$  tel que  $f(\bar{x}_K) = \inf_{x \in K} f(x)$ .

- Sur  $\mathbb{R}^n$ ,
  - $f_k$  est continue,
  - $\lim_{\|x\| \rightarrow +\infty} f_k(x) = +\infty$ ,

Donc pour tout  $k \in \mathbb{N}$ , il existe au moins un  $\bar{x}_k \in \mathbb{R}^n$  tel que  $f_k(\bar{x}_k) = \inf_{x \in \mathbb{R}^n} f_k(x)$

- On a  $\forall x \in K, f_k(x) = f(x)$ . Donc :
  - $f_k$  est continue, strictement convexe sur  $K$ ,
  - $\lim_{\|x\| \rightarrow +\infty} f_k(x) = +\infty$ ,
  - $K$  un sous ensemble non vide, convexe et fermé,

Donc pour tout  $k \in \mathbb{N}$ , il existe un unique  $\bar{x}_K \in K$  tel que  $f(\bar{x}_K) = \inf_{x \in K} f(x)$

2. Montrer que pour tout  $k \in \mathbb{N}$ ,

$$f(\bar{x}_k) \leq f_k(\bar{x}_k) \leq f(\bar{x}_K)$$

Par construction  $f(\bar{x}_k) \leq f(\bar{x}_k) + k\varphi(\bar{x}_k) = f_k(\bar{x}_k)$

$\bar{x}_K \in K$  donc  $f(\bar{x}_K) = f_k(\bar{x}_K) \geq \inf_{x \in \mathbb{R}^n} f_k(x) = f_k(\bar{x}_k)$

3. En déduire qu'il existe une sous-suite  $(\bar{x}_{k_i})_{i \in \mathbb{N}}$  et  $\tilde{x} \in K$  tels que  $\bar{x}_{k_i} \xrightarrow{i \rightarrow +\infty} \tilde{x}$ .

4. Montrer que  $\tilde{x} = \bar{x}_K$ . En déduire que toute la suite  $(\bar{x}_k)_{k \in \mathbb{N}}$  converge vers  $\bar{x}_K$ .

5. Déduire de ces questions un algorithme (dit "de pénalisation") de résolution du problème de minimisation :

$$\begin{cases} \text{Trouver} & \bar{x}_K \in K \\ f(\bar{x}_K) \leq f(x) & \forall x \in K \end{cases}$$

en donnant un exemple de fonction  $\varphi$ .