Université Ibn Zohr Faculté Polydisciplinaire Ouarzazate

Modélisation
Exercices et examens corrigés

Filière Sciences Mathématiques et Informatiques (Semestre 6)

Par

Abdelkrim El mouatasim Professeur Habilité

Années Universitaires : 2015–2021

Première partie

Exercices corrigés

Exercice I.1 *Soit $X \subset \mathbb{R}^n$ et $f : X \longrightarrow \mathbb{R}$ avec $f \in C^1/X$. Démontrer que si f est une fonction convexe sur X, alors $(\nabla f(x) - \nabla f(y))^T(x-y) \geq 0$ pour toute paire de points $x, y \in X$.*

Corrigé :

Puisque f est convexe, alors l'inégalité de gradient s'applique :

$$f(x) \geq f(y) + \nabla f(y)^T(x-y) \Leftrightarrow f(x) - f(y) \geq \nabla f(y)^T(x-y)$$

$$f(y) \geq f(x) + \nabla f(x)^T(y-x) \Leftrightarrow f(y) - f(x) \geq \nabla f(x)^T(y-x)$$

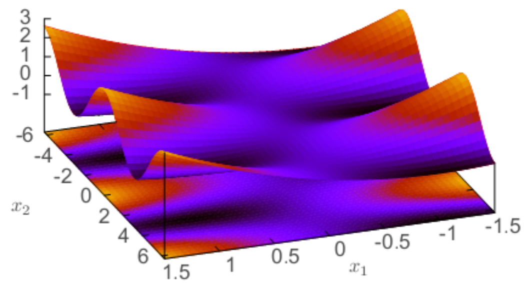
donc

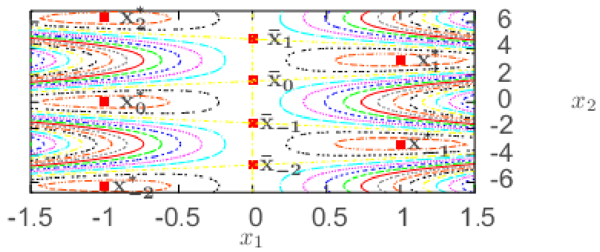
$$0 \geq \nabla f(y)^T(x-y) + \nabla f(x)^T(y-x)$$

ou

$$(\nabla f(x) - \nabla f(y))^T(x-y) \geq 0.$$

Exercice I.2 $f(x_1, x_2) = \frac{1}{2}x_1^2 + x_1 \cos x_2$,





Utilisons les cond. d'opt. pour identifier les minima locaux.

Corrigé :

$$\nabla f(x_1, x_2) = \begin{pmatrix} x_1 + \cos x_2 \\ -x_1 \sin x_2 \end{pmatrix}.$$

Le gradient s'annule pour

— $x_k^* = ((-1)^{k+1}, k\pi)^T, \quad k \in \mathbf{Z}$,

— $\bar{x}_k = (0, \frac{\pi}{2} + k\pi)^T, \quad k \in \mathbf{Z}$,

$$\nabla^2 f(x_1, x_2) = \begin{pmatrix} 1 & -\sin x_2 \\ -\sin x_2 & -x_1 \cos x_2 \end{pmatrix}.$$

$$\nabla^2 f(x_k^*) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \nabla^2 f(\bar{x}_k) = \begin{pmatrix} 1 & (-1)^{k+1} \\ (-1)^{k+1} & 0 \end{pmatrix}.$$

— x_k^* *vérifie les conditions suffisantes d'optimalité pour tout* k

— $\bar{x}_k$ *ne vérifie les conditions nécessaires d'optimalité pour aucun* k.

Exercice I.3 *Soit la fonction*

$$f(x) = \sum_{i=1}^{n} 2x_i^2 - 2\sum_{i=1}^{n-1} x_i x_{i+1}.$$

1. *Misé* f *sous forme de carrés.*
2. *Écrire* f *sous la forme* $x^T Ax$*, avec* A *symétrique.*
3. *Calculer le gradient* $\nabla f(x)$ *et le Hessien de* f*.*
4. *Donner l'approximation de gradient.*
5. *Calculer les valeurs propres de* A*.*
6. *Étudier la convexité de* f*.*

Corrigé :

1.
$$f(x) = x_1^2 + (x_1 - x_2)^2 + \cdots + (x_{n-1} - x_n)^2 + x_n^2$$

2.
$$f(x) = x^T Ax$$

avec $A = \begin{pmatrix} 2 & -1 & & \\ -1 & \ddots & & \\ & & & -1 \\ & & -1 & 2 \end{pmatrix}$

3.
$$\nabla f(x) = 2Ax$$

et

$$H(f) = 2A$$

4. *On pose* $g = \nabla f(x)$ *donc pour* $1 \leq i \leq n$

$$g_i \approx \frac{f(x_1, \ldots, x_i + h, \ldots, x_n) - f(x_1, \ldots, x_i, \ldots, x_n)}{h}$$

5.
$$\lambda_i = 2 - 2\cos\frac{2i\pi}{n+1}, \qquad \forall i \in [1, n]$$

6. *Toutes les valeurs propres de A sont positives, donc la fonction f est convexe.*

Exercice I.4 *Trouver l'optimum de*

$$f(x,y) = 4x^2 - xy + y^2 - x^3$$

Corrigé : *On a :*

$$\frac{\partial f}{\partial x} = 8x - y - 3x^2 \qquad \frac{\partial f}{\partial y} = -x + 2y$$

$$\frac{\partial^2 f}{\partial x^2} = 8 - 6x \qquad \frac{\partial^2 f}{\partial y^2} = 2 \qquad \frac{\partial^2 f}{\partial x \partial y} = -1$$

En résolvant l'équation : $\nabla f(x) = 0$. *Donc les points candidats sont :* $P_1 = (0;0;0)$ *et* $P_2 = (\frac{5}{2}; \frac{5}{4}; \frac{125}{16})$. *La matrice Hessienne est définie par :*

$$Hf(x) = \begin{pmatrix} 8-6x & -1 \\ -1 & 2 \end{pmatrix}$$

Alors

$$Hf(P_1) = \begin{pmatrix} 8 & -1 \\ -1 & 2 \end{pmatrix}$$

Comme $\triangle_1 = 8 > 0$ *et* $\triangle_2 = 15 > 0$, *donc* P_1 *est un minimum.*
Et

$$Hf(P_2) = \begin{pmatrix} -7 & -1 \\ -1 & 2 \end{pmatrix}$$

Comme $\triangle_2 = -7 < 0$ *et* $\triangle_2 = -15 < 0$, *donc* P_2 *est un point-selle.*

Exercice I.5 *Considérer le problème suivant*

$$\min f(x) = x^2 - \frac{1}{2}x$$

1. *Résoudre le problème avec la méthode de Newton en considérant le point initial* $x^0 = 3$

2. *Trouver un intervalle contenant le point où s'annule en utilisant 3 itérations de la méthode de bisection.*

Corrigé :

1. *On a :*

$$x_1 = x_k - f'(x_0)/f''(x_0) = 3 - 11/4 = 1/4$$

   et $f'(x_1) = 0$ *stop.*

Méthode de la bissection : $f'(x) = x^3 + x^2 - 3x - 3 = 0$						
x_1	x_2	x_m	$f'(x_1)$	$f'(x_2)$	$f'(x_m)$	Err.abs
1.0	2.0	1.5	-4.0	3.0	-1.875	0.5
1.5	2.0	1.75	-1.875	3.0	+0.171 87	0.25
1.5	1.75	1.625	-1.875	0.171 87	-0.943 35	0.125
1.625	1.75	1.6875	-0.943 35	0.171 87	-0.409 42	0.0625
1.6875	1.75	1.718 75	-0.409 42	0.171 87	-0.124 78	0.03125

2.

Itération 1 :

On a $f'(0)f'(1) = (-1/2)(3/4) < 0$ donc on pose $a = 0$ et $b = 1$.

Itération 2 :

$c = 1/2$ $f'(c) = 1/2 > 0$ donc on pose $a = 0$ et $b = 1/2$.

Itération 3 :

$c = 1/4$ et $f'(c) = 0$. Stop.

Exercice I.6 *Soit $f(x) = \frac{1}{4}x^4 + \frac{1}{3}x^3 - \frac{3}{2}x^2 - 3x$.*
Utilisé la méthode de la bi-section pour déterminé le optimum de f.

Corrigé :

Si x est l'optimum de f donc $f'(x) = x^3 + x^2 - 3x - 3 = 0$. Dans l'intervalle $[x_1 = 1, x_2 = 2]$ il y a une racine car est continue et $f'(1)f'(2) = -4 * 3 < 0$. On connait les racines pour ce cas : $f'(x) = (x^2 - 3)(x + 1) = 0$, on a trois racines réels : $r_1 = -1, r_2 = -\sqrt{3}, r_3 = \sqrt{3}$

1. *$x_m = \frac{x_1+x_2}{2} = 1.5$ et $f'(x_m) = -1.875$*
2. *Puisque $f'(x_m)f'(x_2) < 0$ alors $x_1 = x_m = 1.5$ et $x_2 = 2$*
3. *$x_m = \frac{x_1+x_2}{2} = 1.75$ et $f'(x_m) = 0.17187$*
4. *Puisque $f'(x_1)f'(x_m) = -1.875 * 0.17187 < 0$ alors $x_1 = 1.5$ et $x_2 = x_m = 1.75$*
5. *$x_m = \frac{x_1+x_2}{2} = 1.625$ alors $f'(x_m) = -0.94335$*
6. *Puisque $f'(x_m)f'(x_2) = -0.94335 * 0.17187 < 0$ la racine se trouve donc dans l'intervalle réduit $[x_1 = 1.625, x_2 = 1.75]$*
7. *$x_m = \frac{x_1+x_2}{2} = 1.6875$ alors $f'(x_m) = -0.40942$*
8. *Puisque $f'(x_m)f'(x_2) = -0.40942 * 0.17187 < 0$ la racine se trouve donc dans l'intervalle réduit $[x_1 = 1.6875, x_2 = 1.75]$.*
   Et ainsi de suite...

On voit clairement que l'intervalle devient de plus petit
$(|x_2 - x_1|)$et que l'on se dirige vers 1.732050$(\simeq r_3 = \sqrt{3})$.
Puisque $f'(r_3) = 0$ et $f''(r_3) > 0$ donc r_3 est un minimum local de f.
On voit aussi que la méthode a certain désavantager (lenteur en particulier, et comment on s'arrête ?) : critéres d'arrêts

1. *L'erreur absolue :* $|r - x_m| \simeq \frac{|x_1 - x_2|}{2} < \epsilon_{abs}$
2. *L'erreur relative :* $\frac{|r-x_m|}{|r|} \simeq \frac{|x_1-x_2|}{|x_m|} < \epsilon_{rel}$
3. *On peut arrêter l'algorithme si* $\left|f'(x_m)\right| < \epsilon_f$

Exercice I.7 *Soit* $f(x) = -\exp(-x) - \frac{1}{2}x^2$

Utilisé la méthode de Newton pour déterminé le minimum de f.

Corrigé :

Méthode de Newton : $f'(x) = e^{-x} - x$			
n	x_n	$\|e_n\|$	$\|\frac{e_{n+1}}{e_n}\|$
0	0.000 0000	$0.5671 * 10^{+0}$	$0.1183 * 10^{+0}$
1	0.500 0000	$0.6714 * 10^{-1}$	$0.1239 * 10^{-1}$
2	0.566 3110	$0.8323 * 10^{-3}$	$0.1501 * 10^{-3}$
3	0.567 1432	$0.1250 * 10^{-6}$	$\simeq 0$
4	0.567 1433	$0.4097 * 10^{-9}$	—

On remarque la convergence très rapide de cette méthode.

Exercice I.8 *Soit le problème*

$$\min f(x,y) = \frac{1}{2}(x^2 + y^2)$$

Considérer la solution initiale $x^0 = 2, \quad y^0 = -1$. *Exécuter 3 itérations de la méthode du gradient.*

Corrigé :

1.

$$\nabla f(x,y) = (x,y)^T$$

Itération 1 : étape 0 : $X_0^T = (2,-1), \quad k = 0, \quad \delta = 10^{-2}$

étape 1 : $\nabla f(X_0)^T = (2,-1) \neq (0,0)$

étape 2 : $t_0 = \arg\min_{t\geq 0} f(X_0 - t\nabla f(X_0)) = 1$

Itération 2 :

étape 0 : $X_1^T = X_0^T - t_0 \nabla f(X_0)^T = (2,-1) - 1(2,-1) = (0,0)$

étape 1 : $\nabla f(X_1)^T = (0,0)$ *stop*

$X_*^T = X_1 = (0,0)$ *et* $f^* = f(X_*) = 0$.

Exercice I.9 *Considérons le problème (P) suivant :*

$$\max f(x_1,x_2) = \ln(x_1+1) + x_2$$
$$s.c.\ 2x_1 + x_2 \leq 3,$$
$$x_1 \geq 0, x_2 \geq 0.$$

Utilisé les conditions KKT pour résoudre le problème (P).

Corrigé :

Nous pouvons donc considérer qu'il n'y a qu'une seule contrainte avec

$$g_1(x_1, x_2) = 2x_1 + x_2,$$

et $b_1 = 3$.

Nous associons à cette contrainte un multiplicateur $u_1 \geq 0$, *et pour les contraintes de non-négativité* $u_2 \geq 0, u_3 \geq 0$, *nous avons alors les conditions suivantes :*

$$\begin{aligned}
\frac{1}{x_1^* + 1} - 2u_1 + u_2 &= 0, \\
1 - u_1 + u_3 &= 0, \\
2x_1^* + x_2^* - 3 &\leq 0, \\
u_1(2x_1^* + x_2^* - 3) &= 0, \\
-x_1 &\leq 0, \\
-x_2 &\leq 0, \\
u_2 x_1 &= 0, \\
u_3 x_2 &= 0.
\end{aligned}$$

Nous obtenons $u_1 \geq 1$. *Puisque*

$$1 - 2u_1(x_1^* + 1) + u_2 = 0$$

et $x_1^* \geq 0$, *nous en déduisons* $u_2 \geq 1$

Par conséquent, $x_1^* = 0$. *Puisque* $u_1 \neq 0$, *nous avons*

$$2x_1^* + x_2^* - 3 = 0,$$

et par conséquent $x_2^* = 3$. *Dès lors,* $u_1 = 1, u_2 = 1, u_3 = 0$. *Les conditions KKT sont donc satisfaites en un seul point :* $(0, 3)$.

Exercice I.10 *Soit (P) le problème :*

$$\begin{cases} \min \;\; f(x) = & x_1^2 + x_2^2 \\ s.c. & x_1 + 2x_2 + 3x_3 - x_4 = 6 \\ & x_1 + 3x_2 + 2x_3 + 5x_4 = 4 \\ & x \geq 0 \end{cases}$$

Utilisé la méthode de gradient réduit avec $x^0 = (2, 2, 1, 0)^t$.

Corrigé :

Ici $A = \begin{pmatrix} 1 & 2 & 3 & -1 \\ 1 & 3 & 2 & 5 \end{pmatrix}$

Choisissons $B = \{1,3\}$, $N = \{2,4\}$. *Alors* $A_B = \begin{pmatrix} 1 & 3 \\ 1 & 2 \end{pmatrix}, A_N = \begin{pmatrix} 2 & -1 \\ 3 & 5 \end{pmatrix}, x_B \begin{pmatrix} x_1 \\ x_3 \end{pmatrix}, x_N \begin{pmatrix} x_2 \\ x_4 \end{pmatrix}$

$$\begin{aligned} (A_B \ \ A_N)\begin{pmatrix} x_B \\ x_N \end{pmatrix} &= \begin{pmatrix} 1 & 3 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_3 \end{pmatrix} + \begin{pmatrix} 2 & -1 \\ 3 & 5 \end{pmatrix}\begin{pmatrix} x_2 \\ x_4 \end{pmatrix} \\ &= \begin{pmatrix} x_1 + 3x_3 + 2x_2 - x_4 \\ x_1 + 2x_3 + 3x_2 + 5x_4 \end{pmatrix} \end{aligned}$$

On exprime x_1 *en fonction de* x_2, x_4, *on obtient*

$$g(x_2, x_4) = (-5x_2 - 17x_4)^2 + x_2$$

et

$$\left(\frac{\partial g}{\partial x_2} \ \ \frac{\partial g}{\partial x_4}\right) = (52x_2 + 170x_4 \ \ 170x_2 + 578x_4).$$

Alors
$A_B^{-1} = \begin{pmatrix} -2 & 3 \\ 1 & -1 \end{pmatrix}$

Effectuons le changement de base $B' = \{2,3\}$, $N' = \{1,4\}$ $A_B^{-1}A_{\{2\}} = \begin{pmatrix} -2 & 3 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 5 \\ -1 \end{pmatrix}$ $p = 5 \neq 0$ *et donc* $A_{B'} = \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}$ *est inversible* $(A_B^{-1}A_{B'})^{-1} = \begin{pmatrix} \frac{1}{5} & 0 \\ \frac{1}{5} & 1 \end{pmatrix}$ *et* $A_{B'}^{-1} = \begin{pmatrix} \frac{1}{5} & 0 \\ \frac{1}{5} & 1 \end{pmatrix}\begin{pmatrix} -2 & 3 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} \frac{-2}{5} & \frac{3}{5} \\ \frac{3}{5} & \frac{-2}{5} \end{pmatrix}$

Pour $B = \{1,3\}$, *on a* $\begin{pmatrix} x_1 \\ x_3 \end{pmatrix} + A_B^{-1}A_N\begin{pmatrix} x_2 \\ x_4 \end{pmatrix} = A_B^{-1}b$ *soit* $\begin{pmatrix} x_1 \\ x_3 \end{pmatrix} + \begin{pmatrix} -2 & 3 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 3 & 5 \end{pmatrix}\begin{pmatrix} x_2 \\ x_4 \end{pmatrix} = \begin{pmatrix} -2 & 3 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 6 \\ 4 \end{pmatrix}$ *soit encore* $\begin{pmatrix} x_1 \\ x_3 \end{pmatrix} + \begin{pmatrix} 5 & 17 \\ -1 & -6 \end{pmatrix}\begin{pmatrix} x_2 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$

$x_3 = 0 \Rightarrow -x_2 - 6x_4 = 2$

Les pivots des colonnes A_2 *et* A_4 *(relatifs à* x_3*) sont* -1 *et* -6. *S'ils étaient nuls on aurait* $0 = 2$. *On voit que la condition de non-dégénérescence n'est pas nécessaire puisqu'ici* $A_B^{-1}b$ *a une composante (celle associée à* x_1*) nulle et pourtant on peut effectuer les changements de base* $B - \{1\} + \{2\}$ *et* $B - \{1\} + \{4\}$, *les pivots (5 et 17) étant non nuls.*

Exercice I.11

$$\begin{cases} \min \ \ f(x) = & x_1^2 + x_2^2 + x_3^2 + x_4^2 - 2x_1 - 3x_4 \\ s.c. & 2x_1 + x_2 + x_3 + 4x_4 = 7 \\ & x_1 + x_2 + 2x_3 + x_4 = 6 \\ & x \geq 0 \end{cases}$$

Utilisé la méthode de gradient réduit avec $x^0 = (2,2,1,0)^t$.

Corrigé :

Soient $x^0 = (2,2,1,0)^t$, x_1 *et* x_2 *les variables de base et* $B = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ *la base.*
On a : $\nabla f(x^0) = (2,4,2,-3)^T$

$$r(x^0) = (2,4,2,-3) - (2,4)\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 2 & 1 & 1 & 4 \\ 1 & 1 & 2 & 1 \end{pmatrix} = (0,0,-8,-1),$$

par conséquent, $d_3 = 8$ *et* $d_4 = 1$.

On a :

$$d_B = B^{-1}Nd_N = (5,-22)^T$$

et $\mu \leq \mu_{\max} = \min\limits_{j\in B, d_j<0} \{-\frac{x_j^k}{d_j}\} = \frac{1}{11}$.
On vérifie que le minimum de la fonction $f(x^0 + \mu d)$ *pour* $\mu \in [0, \frac{1}{11}]$ *est obtenu pour* $\mu = \frac{1}{11}$.
La variable x_2 *quitte alors la base pour y être remplacée par* x_3 *ou* x_4*, au choix.*

Exercice I.12 *On considère le problème d'optimisation sous contraintes suivante :*

$$\left\{\begin{array}{ll} \min & x_1^2 + x_2^2 \\ s.c. & x_1 + x_2 \geq 1 \\ & x_1 \geq 0, x_2 \geq 0 \end{array}\right. \qquad (P)$$

(P) revient à le pb d'optimisation avec contraintes d'égalité

$$\left\{\begin{array}{ll} \min & x_1^2 + x_2^2 \\ s.c. & x_1 + x_2 + x_3 = 1 \\ & x_1 \geq 0, x_2 \geq 0, x_3 \geq 0 \end{array}\right.$$

Choisissons le point initial $P_0 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$ *et* $B = \{3\}$, $N = \{1,2\}$ *pour démarré la méthode de gradient réduit.*

Corrigé :

Itération 1.
On a $x_3 = -1 + x_1 + x_2$. *On en déduit* $g(x_1,x_2) = f(x_1,x_2,-1+x_1+x_2)$, $\frac{\partial g}{\partial x_N} = (\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}) = (2x_1, 2x_2)$ *et* $r(P_0) = (4\ 0)$. *La direction de déplacement est* $\begin{pmatrix} -4 \\ 0 \\ -4 \end{pmatrix}$.

$$\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} + \alpha \begin{pmatrix} -4 \\ 0 \\ -4 \end{pmatrix} \geq 0 \Rightarrow \left\{ \begin{array}{l} \alpha \leq \frac{1}{2} \\ \alpha \leq \frac{1}{4} \end{array} \right. \Rightarrow \alpha_{\max} = \frac{1}{4}$$

Le minimum de $f(2-4\alpha, 0, 1-4\alpha)$ *sous la contrainte* $0 \leq \alpha \leq \frac{1}{4}$ *est atteint en* $\alpha_1 = \frac{1}{4}$.
Le nouveau point est $P_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. *Sa coordonnée 3 est nulle. Il faut effectuer un changement de base. L'indice 1 est l'indice dans N associé à la coordonnée de* P_1 *la plus grande : on forme la nouvelle décomposition* $B = \{1\}$ *et* $N = \{2,3\}$.

Itération 2.
On a $x_1 = 1 - x_2 + x_3$ *d'où* $g(x_2, x_3) = f(1 - x_2 + x_3, x_2, x_3)$, $\frac{\partial g}{\partial x_N} = (-\frac{\partial f}{\partial x_1} + \frac{\partial f}{\partial x_2} \ \ \frac{\partial f}{\partial x_1} + \frac{\partial f}{\partial x_3}) =$ $(-2x_1 + 2x_2 \ \ 2x_1)$ *et* $r(P_1) = (-2 \ 2)$. $d_N = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$ *et* $d_B = -2$ *donc la direction de déplacement est* $\begin{pmatrix} -2 \\ 2 \\ 0 \end{pmatrix}$.

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} -2 \\ 2 \\ 0 \end{pmatrix} \geq 0 \Rightarrow \begin{cases} \alpha \leq \frac{1}{2} \\ \alpha \leq 0 \end{cases} \Rightarrow \alpha_{\max} = \frac{1}{2}$$

Le minimum de $f(1 - 2\alpha, 2\alpha, 0)$ *sous la contrainte* $0 \leq \alpha \leq \frac{1}{2}$ *est atteint en* $\alpha_2 = \frac{1}{4}$.
Le nouveau point est $P_2 = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix}$.

Itération 3.
Le gradient réduit en P_2 *est* $r(P_2) = (0 \ \ 1)$ *donc* $d_N = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. *On arrête.*

Vérifions que les conditions de Kuhn-Tucker sont satisfaites : ici $A = (1 \ \ 1 \ \ -1)$
soient $\mu = -\frac{\partial f}{\partial x_1}(\frac{1}{2}, \frac{1}{2}, 0)A_1^{-1} = -2 \times \frac{1}{2} = -1$ *(car* $A_1 = 1$*),* $\lambda_1 = 0$, $(\lambda_2 \ \ \lambda_3) = r(P_2) = (0 \ \ 1)$
$\frac{\partial f}{\partial x}(\frac{1}{2}, \frac{1}{2}, 0) + \mu A - \lambda = 0 \Leftrightarrow (2 \times \frac{1}{2} \ \ 2 \times \frac{1}{2} \ \ 0) - (1 \ \ 1 \ \ -1) - (0 \ \ 0 \ \ 0)$, *donc vérifié.*
On aura remarqué que l'on a une écriture "transposée" des équations de Kuhn-Tucker.
$\lambda \geq 0$ *et les conditions de complémentarités* $\lambda_1 \times \frac{1}{2} = 0, \lambda_2 \times \frac{1}{2} = 0, \lambda_3 \times 0 = 0$ *sont vérifiées.*

Remarque :
En partant du même point on obtient des cheminements différents selon le choix du B initial.

Si l'on choisit $B = \{1\}$ *et* $N = \{2, 3\}$, *on passe par les 3 points* $P_1 = \begin{pmatrix} \frac{2}{5} \\ \frac{4}{5} \\ \frac{1}{5} \end{pmatrix}$, $P_2 = \begin{pmatrix} \frac{2}{5} \\ \frac{3}{5} \\ 0 \end{pmatrix}$, $P_3 = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix}$ *sans changer de base B.*

Exercice I.13 *Le tableau suivant donne l'évolution en fonction de l'année du budget publicitaire d'une entreprise, en dizaines de milliers dh.*

années	*1998*	*1999*	*2000*	*2001*	*2002*	*2003*	*2004*	*2005*
rang x_i	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
budget y_i	*2*	*2,2*	*2,5*	*3*	*3,2*	*3,5*	*3,7*	*4,2*

1. *Montrer que* $\mathbf{cov}(X, Y) = \overline{XY} - \bar{X}\bar{Y}$.
2. *Calculer le coefficient de corrélation entre ces deux séries.*
3. *Déterminer une équation de la droite* Δ *d'ajustement par la méthode des moindres carrés, commenter la qualité de cet ajustement. (On donnera les résultats à 0,001 près).*

4. *On considère que cette droite permet un ajustement de cette série statistique. Calculer à partir de quelle année le budget devrait **dépasser** 60 000 dh.*

Corrigé :

1. *On a*

$$\mathbf{cov}(X,Y) = \overline{(X-\bar{X})(Y-\bar{Y})}$$

donc

$$\begin{aligned}\mathbf{cov}(X,Y) &= \frac{\sum(X_i-\bar{X})(Y_i-\bar{Y})}{n}\\ &= \frac{\sum(X_iY_i-X_i\bar{Y}-\bar{X}Y_i+\bar{X}\bar{Y})}{n}\\ &= \frac{\sum(X_iY_i)-\bar{Y}\sum X_i-\bar{X}\sum Y_i+n\bar{X}\bar{Y})}{n}\\ &= \overline{XY}-\bar{X}\bar{Y}\end{aligned}$$

2.

$$\begin{aligned}\mathbf{r}(X,Y) &= \frac{\mathbf{cov}(X,Y)}{\sigma(X)\sigma(Y)}\\ &= \frac{1.63125}{2.29128*0.71578}\\ &= 0.99462\end{aligned}$$

3. *L'équation de la droite* $X = aY + b$ *avec :*

$$a = \frac{\mathbf{cov}(X,Y)}{\sigma(Y)^2} = 3.18389$$

et

$$b = (\bar{X}) - a(\bar{Y}) = b = -5.17109$$

donc $X = 3.18389Y - 5.17109$ *On a* $\mathbf{r}(X,Y) \approx 1$ *donc la régression entre X et Y est fort.*

4. $Y = 6$ *donc* $X = 3.18389 * 6 - 5.17109 \approx 14$
Alors à partir de l'année 2011 le budget devrait dépasser 60000dh.

Exercice I.14 *Nous considérons l'ensemble* $X = \{1,2,3,6,7,8,13,15,17\}$ *à partir duquel nous souhaitons créer 3 clusters.*

1. *Appliqué l'algorithme kmeans ont choisissons arbitrairement pour centres de cluster les éléments* $M_1 = 13,\ M_2 = 15,\ M_3 = 17$ *associés respectivement aux clusters* $C_1 = 13,\ C_2 = 15, C_3 = 17$.

2. *Appliqué l'algorithme kmeans++ ont choisissons arbitrairement pour centre de cluster l'éléments* $M_1 = 2$ *associés au clusters* $C_1 = 2$.

Corrigé :

Table 1 : Calcul des distances $D^2(x)$ *séparant chaque élément de X à chaque centre* $M_1 = 13,\ M_2 = 15,\ M_3 = 17$ *et identification des clusters* $C_1 = \{1,2,3,6,7,8,13\}$ *de centre* $M_1 = \frac{40}{7},\ C_2 = \{15\}$ *de centre* $M_2 = 15$ *et* $C_3 = 17$ *de centre* $M_3 = 17$.

1\.

x	$M_1 = 13$	$M_2 = 15$	$M_3 = 17$
1	144	196	256
2	121	169	225
3	100	144	196
6	49	81	121
7	36	64	100
8	25	49	81

TABLE 1 – Itération 1

x	$M_1 = \frac{40}{7}$	$M_2 = 15$	$M_3 = 17$
1	1089/49	196	256
2	676/49	169	225
3	361/49	144	196
6	4/49	81	121
7	81/49	64	100
8	256/49	49	81
13	2601/49	4	14

TABLE 2 – Itération 2

Table 2 : Calcul des distances $D^2(x)$ séparant chaque élément de X à chaque centre $M_1 = \frac{40}{7}$, $M_2 = 15$, $M_3 = 17$ et identification des clusters $C_1 = \{1, 2, 3, 6, 7, 8\}$ de centre $M_1 = 4.5$, $C_2 = \{13, 15\}$ de centre $M_2 = 14$ et $C_3 = 17$ de centre $M_3 = 17$.

x	$M_1 = 4.5$	$M_2 = 14$	$M_3 = 17$
1	12.25	169	256
2	6.25	144	225
3	2.25	121	196
6	2.25	64	121
7	6.25	49	100
8	12.25	36	81
13	72.25	1	14
15	110.25	1	4

TABLE 3 – Itération 3

Table 3 : Calcul des distances $D^2(x)$ séparant chaque élément de X à chaque centre $M_1 = 4.5$, $M_2 = 14$, $M_3 = 17$ et identification des clusters $C_1 = \{1, 2, 3, 6, 7, 8\}$ de centre $M_1 = 4.5$, $C_2 = \{13, 15\}$ de centre $M_2 = 14$ et $C_3 = 17$ de centre $M_3 = 17$.
Les centres issus du tableau 3 sont les mêmes que ceux issus du tableau 2. L'algorithme s'arrête donc et nous obtenons les 3 clustres suivants

$$\begin{aligned} C_1 = \{1, 2, 3, 6, 7, 8\} &\quad \textit{de centre } M_1 = 4.5 \\ C_2 = \{13, 15\} &\quad \textit{de centre } M_2 = 14 \\ C_3 = \{17\} &\quad \textit{de centre } M_3 = 17 \end{aligned}$$

2.

x	1	3	6	7	8	13	15	17
$D_2^2(x)$	1	1	16	25	36	121	169	225
$P(M_2 = x)$	$\frac{1}{594}$	$\frac{1}{594}$	$\frac{16}{594}$	$\frac{25}{594}$	$\frac{36}{594}$	$\frac{121}{594}$	$\frac{169}{594}$	$\frac{225}{594}$

TABLE 4 – Itération 1

Tableau 4 : Calcul des distances $D_2^2(x)$ séparant chaque élément de X au premier centre égal à 2 et identification du deuxième centre $M_2 = 17$ correspondant à l'élément de X avec la plus forte probabilité.

x	1	3	6	7	8	13	15
$D_2^2(x)$	1	1	16	25	36	121	169
$D_{17}^2(x)$	256	196	121	100	81	16	4
$P(M_3 = x)$	$\frac{1}{99}$	$\frac{1}{99}$	$\frac{16}{99}$	$\frac{25}{99}$	$\frac{36}{99}$	$\frac{16}{99}$	$\frac{4}{99}$

TABLE 5 – Itération 2

Tableau 5 : Calcul des distances $D_2^2(x)$ séparant chaque élément de X au premier centre égal à 2, $D_{17}^2(x)$ séparant chaque élément de X au deuxième centre égal à 17 et identification du troisième centre $M_3 = 8$ correspondant à l'élément de X avec la plus forte probabilité.
Tableau 6 : Calcul des distances séparant chaque élément de X à chaque centre $M_1 = 2$,

x	$M_1 = 2$	$M_2 = 17$	$M_3 = 8$
1	1	256	49
3	1	196	25
6	16	121	4
7	25	100	1
13	121	16	25
15	169	4	49

TABLE 6 – Phase 2

$M_2 = 17$ et $M_3 = 8$ et identification des clusters $C_1 = \{1,2,3\}$ de centre $M_1 = 2$, $C_2 = \{13,15,17\}$ de centre $M_2 = 15$ et $C_3 = \{6,7,8\}$ de centre $M_3 = 7$.

Exercice I.15 *Considérer le problème de produire une boîte de volume maximale avec une pièce de carton de superficie spécifiée $c > 0$. Dénotons les dimensions des côtés de cette boîte par x, y et z. Le problème peut se formuler comme suit :*

$$\begin{cases} \max & xyz \\ s.c. & yx + yz + xz = \frac{c}{2} \end{cases}$$

Déterminer les dimensions optimales en utilisant les conditions de Kuhn-Tucker (KKT).

Corrigé : *KKT :*

$$\begin{aligned} -yz + \lambda_1(y+z) - \lambda_2 = 0 \quad & (1) \\ -xz + \lambda_1(x+z) - \lambda_3 = 0 \quad & (2) \\ -xy + \lambda_1(x+y) - \lambda_4 = 0 \quad & (3) \\ xy + yz + xz = \tfrac{C}{2} \quad & (4) \\ x\lambda_2 = y\lambda_3 = z\lambda_3 = 0 \quad & (5) \end{aligned}$$

(1)+(2)+(3) :

$$-(xy + yz + zx) + 2\lambda_1(x+y+z) = \lambda_2 + \lambda_3 + \lambda_4$$

$\Rightarrow$

$$\lambda_1(x+y+z) = \frac{\lambda_2 + \lambda_3 + \lambda_4}{2} + \frac{C}{4} > 0 \Rightarrow \lambda_1 > 0$$

Si $x = 0$*, alors (3) devient* $\lambda_1 z = \lambda_3$

$$\Rightarrow \lambda_1 yz = y\lambda_3 = 0$$

$$\Rightarrow y = 0 ou z = 0$$

d'après (4) impossible.
Donc $x \neq 0$*, de même* $y \neq 0$ *et* $z \neq 0$

$$\Longrightarrow \lambda_2 = \lambda_3 = \lambda_4 = 0$$

On multiplie :

(1) par x et (2) par y et on soustrait	$\Rightarrow$	$x - y = 0$
(1) par x et (3) par z et on soustrait	$\Rightarrow$	$x - z = 0$
(2) par Y et (3) par z et on soustrait	$\Rightarrow$	$u - z = 0$

Donc (4) $\Longrightarrow x = y = z = \sqrt{\frac{C}{6}}$

Deuxième partie

Examens corrigés

Année Universitaire : 2016–2017

Examen session principale

Exercice II.1 *Soit* $f : \mathbb{R}^n \to \mathbb{R}$, $f \in C^1$ *et convexe sur* $\mathbb{R}^n$*. soit* $x^* \in \mathbb{R}^n$, $d \in \mathbb{R}^n$ *une direction, et* $\alpha > 0$ *un scalaire. Soit* $\bar{x} = x^* + \alpha d$*.*
Démontrer que si $\nabla f(\bar{x})^T d < 0$*, alors* d *est une direction de descente au point* x^**.*

Corrigé :

Utilisent l'inégalité de gradient

$$\begin{aligned} f(x^*) \geq & f(\bar{x}) + \nabla f(\bar{x})^T (x^* - \bar{x}) \\ = & f(\bar{x}) - \alpha \nabla f(\bar{x})^T d \\ > & f(\bar{x}) \end{aligned}$$

Utilisent la définition de la fonction convexe $x \in [x^*, \bar{x}]$ *est de la forme*

$$x = \theta x^* + (1 - \theta)\bar{x} \qquad (1)$$

et si $\theta \neq 1$ *on a*

$$f(x) = f(\theta x^* + (1 - \theta)\bar{x}) \leq \theta f(x^*) + (1 - \theta) f(\bar{x}) < f(x^*)$$

Également, en peut écrire (1)

$$x = \theta x^* + (1 - \theta)(x^* + \alpha d) = x^* + (1 - \theta)\alpha d$$

Donc pour $0 \leq \theta 1$*,*

$$f(x^*) > f(x^* + (1 - \theta)\alpha d)$$

$\Rightarrow$ *d est une direction de descente.*

Exercice II.2 *Modèle statistique simple*
Utiliser la méthode des moindres carrés pour estimer les paramètres de la droite de régression $Y = aX + b$ *qui approche le mieux les* $\boldsymbol{n}$ *couples* (X_i, Y_i)*.*

Corrigé :

Soit $(X_i)_{1\leq i\leq n}$ *et* Y_i *deux séries statistiques. On cherche à savoir s'il est raisonnable de prévoir* Y_i *en fonction de* X_i*, avec par exemple une relation linéaire* $Y_i = a + bX_i$*. Pour cela on minimise* $f(a,b) = \sum(a + bX_i - Y_i)^2$ *par rapport à a et b, ce qui donne les équations :*

$$\begin{cases} \frac{\partial f}{\partial a} = 0 & = & 2\sum_{i=1}^{n}(a + bX_i - Y_i) \\ \frac{\partial f}{\partial b} = 0 & = & 2\sum_{i=1}^{n} X_i(a + bX_i - Y_i) \end{cases}$$

équivalentes à :

$$\begin{cases} a\sum_{i=1}^{n} + & b\sum_{i=1}^{n} X_i & = & \sum_{i=1}^{n} Y_i \\ a\sum_{i=1}^{n} X_i + & b\sum_{i=1}^{n} X_i^2 & = & \sum_{i=1}^{n} X_iY_i \end{cases}$$

Donc

$$a = \frac{cov(X,Y)}{v(X)}, \text{ et } b = m(Y) - am(X).$$

Exercice II.3 *On considère le problème de programme non-linéaire HS48 suivant :*

$$\begin{aligned} \min z =&(x_1 - 1)^2 + (x_2 - x_3)^2 + (x_4 - x_5)^2 \\ &s.c.\ x_1 + x_2 + x_3 + x_4 + x_5 = 5 \\ &\quad x_3 - 2(x_4 + x_5) = -3 \\ &\quad x_i \geq 0,\ \ i = 1,\dots,5. \end{aligned}$$

Compléter deux itérations de la méthode du gradient réduit pour résoudre le problème HS48, on partira du point $X^0 = (2, 1.5, 0, 1.5, 0)^T$

Corrigé :

— 1^{re} *itération :*
$X^0_B = (x_1, x_4)$, $X^0_N = (x_2, x_3, x_5)$,
$\Rightarrow r_N(X^0) = (1, -\frac{9}{2}, -6)^T$, $d_N = (-1, \frac{9}{2}, 6)^T$ *et* $d_B = (-\frac{23}{4}, -\frac{15}{4})^T$
$\alpha_{max} = 0.3478 \Rightarrow \alpha_1 = 0.1807$

$$\Rightarrow X^1 = (0.9607, 1.3193, 0.8134, 0.822, 1.084)^T$$

— 2^{me} *itération :*
$X^1_B = (x_1, x_4)$, $X^1_N = (x_2, x_3, x_5)$,
$\Rightarrow d = (0.856, -0.63, -0.1509, -0.5638, -0.4884)^T$
$\Rightarrow \alpha_2 = 0.159$

$$\Rightarrow X^2 = (0.868, 1.162, 0.9794, 1.0559, 0.9338)^T$$

Examen session rattrapage

Exercice II.4 *Soit la fonction* $f = \frac{1}{2}x^T Dx - b^T x$ *où* $x, b \in \mathbb{R}^n$ *et* D *est une matrice* $n \times n$.

1. *Démontrer que si* $x^T Dx$ *est convexe sur* $\mathbb{R}^n$*, alors* f *est aussi convexe sur* $\mathbb{R}^n$.

2. *Démontrer que si tout les valeurs propre de D sont positive, alors f est convexe sur $\mathbb{R}^n$.*

Corrigé :

1. *La fonction $g(x) = -b^T x$ est linéaire donc convexe. En effet $-b^T(\alpha x + (1-\alpha)y) = -\alpha b^T x - (1-\alpha)b^T y$*
   Puisque $x^T Dx$ est convexe, alors $f(x)$ est une combinaison linéaire de deux fonction convexe avec un poids $\frac{1}{2}$ pour $x^T Dx$ et 1 pour $-b^T x$. Donc f est convexe.
2. *Si tout les valeurs propre de D positive donc D est semi défini positive $\Rightarrow x^T Dx$ est convexe d'après Q_1 f est convexe.*

Exercice II.5 1. *Montrer que en partant du point $X^0 = \begin{pmatrix} \alpha \\ 0 \end{pmatrix}$ ($\alpha \neq 0$ quelconque), la méthode de Newton converge.*

2. *Appliquer la méthode de Newton avec le pas optimale en partant de $X^0 = \begin{pmatrix} \alpha \\ 0 \end{pmatrix}$ pour calculer X^1. Que constate-t-on ?*

Corrigé :

1. $X^1 = (\frac{2}{3}\alpha, 0)^T$, $X^2 = ((\frac{2}{3})^2\alpha, 0)^T$
   $\Rightarrow X^k = ((\frac{2}{3})^k\alpha, 0)^T \xrightarrow{k\to\infty} 0$
2. *Direction de Newton en X^0 est $-(\frac{1}{3}\alpha, 0)^T$*
   $X^1(t) = X^0 - t(\frac{1}{3}\alpha, 0)^T$
   $\Rightarrow g(t) = f(X(t)) = \dfrac{(\alpha - \frac{t}{3}\alpha)^4}{4} \geq 0$ *$g(t)$ atteint son minimum en $t = 3$ $X^1 = X(3) = (0,0)^T$ minimum atteint.*

Exercice II.6 *On considère le problème d'optimisation non-linéaire (P1) suivant :*

$$\begin{aligned}\min z =& (x_1 - 1)^2 + (x_2 - x_3)^2 + (x_4 - x_5)^2 \\ & s.c.\ x_1 + x_2 + x_3 + x_4 + x_5 = 5 \\ & \quad x_2 - 2(x_4 + x_5) = -3 \\ & \quad x_i \geq 0, \ \ i = 1, \ldots, 5.\end{aligned}$$

Soit $X^0 = (2, 0, 1.5, 0, 1.5)^T$ le point initial, pour résoudre le problème (P1) par la méthode du gradient réduit.

1. *Donner les bases possible.*
2. *Compléter une itération de l'algorithme. Donner la base de deuxième itération.*
3. *Vérifier que le point $X^* = (1, 1, 1, 1, 1)^T$ vérifie les conditions de Kuhn-Tucker.*

Corrigé :

1. $B_1 = \{x_1, x_5\}$ *et* $B_2 = \{x_3, x_5\}$
2. *Si* $B = B_1$ *on trouve* $X^1 = (0.961, 0.834, 1.319, 1.084, 0.822)^T$ *et* $B = B_1$
   Si $B = B_2$ *on trouve* $X^1 = (2.108, 0.648, 0.42, 0.648, 1.176)^T$ *et* $B = B_2$
3. $d_N = 0 \Rightarrow X^*$ *vérifier les conditions de KKT.*

Année Universitaire : 2017–2018

Examen session principale

Exercice II.7 *Soit $X \in \mathbb{R}^n$ est un ensemble convexe et si la fonction $f : X \longrightarrow \mathbb{R}$ est strictement convexe sur X. On pose qu'il existe un point minimum local x^*.*

1. *Démontrer que x^* est un minimum global.*
2. *Démontrer que le minimum est unique.*

Corrigé :

1. *La preuve se fait par contradiction en supposant qu'il existe un point $\hat{x} \in X$ tel que $f(\hat{x}) < f(\bar{x})$. Puisque f est convexe, $f(\theta\hat{x} + (1-\theta)\bar{x}) \leq \theta f(\hat{x}) + (1-\theta)f(\bar{x}) < \theta f(\bar{x}) + (1-\theta)f(\bar{x}) = f(\bar{x})$ pour tout $\theta \in (0,1]$. Or pour $\theta > 0$ suffisamment petit,*

$$x(\theta) = \theta\hat{x} + (1-\theta)\bar{x} \in B\epsilon(\bar{x}) \cap X.$$

   Ainsi $f(x(\theta)) < f(\bar{x})$ où $x(\theta) \in B\epsilon(\bar{x}) \cap X$, contredisant que $\bar{x}$ est un minimum local de f sur X.

2. *Soit donc $x^* \in K$ tel que $f(x^*) \leq f(x), \quad \forall x \in K$. Supposons qu'il existe $y^* \neq x^*$ tel que $f(y^*) \leq f(x), \quad \forall x \in K$. Formons pour $\lambda \in \,]0,1[$ le vecteur*

$$u = \lambda y^* + (1-\lambda)x^*.$$

   D'après la stricte convexité de f et puisque nécessairement $f(y^*) = f(x^*)$ on a

$$f(u) < \lambda f(y^*) + (1-\lambda)f(x^*) = f(x^*),$$

   ce qui contredit le fait que x^* soit un minimum. On a donc $x^* = y^*$.

Exercice II.8 *Considérer le problème de déterminer le point de la droite $a_1x + a_2y = b$ le plus proche de l'origine. Nous faisons l'hypothèse que $a_1 \neq 0$ ou $a_2 \neq 0$.*

1. *Formuler ce problème comme un problème d'optimisation.*
2. *Écrire les conditions d'optimalité KKT pour ce problème.*
3. *Déterminer une solution optimale de ce problème.*

4. *Les conditions KKT sont-elle suffisantes pour ce problème ? Justifier votre réponse.*

Corrigé :

1.
$$\begin{aligned} \min z =& x^2 + y^2 \\ & s.c.\ a_1x + a_2y = b \end{aligned}$$

2.
$$\begin{aligned} L(x, y, z) = x^2 + y^2 - \lambda(a_1x + a_2y - b) \\ 2x - a_1\lambda = 0 \\ 2y - a_2\lambda = 0 \\ a_1x + a_2y = b(*) \end{aligned}$$

3. $x = \frac{a_1\lambda}{2}$ *et* $y = \frac{a_2\lambda}{2}$ *d'après* $(*)$ $\lambda = \frac{2b}{a_1^2 + a_2^2}$ *donc* $x = \frac{a_1 b}{a_1^2 + a_2^2}$ *et* $y = \frac{a_2 b}{a_1^2 + a_2^2}$
4. *On a le problème d'optimisation convexe donc les conditions de KKT sont suffisantes.*

Exercice II.9 *On considère le programme nonlinéaire (P) suivant :*

$$\begin{aligned} \min z = x_1^2 - x_1x_2 + 6x_2^2 - 12x_2 + 2x_3^2 \\ s.c.\ x_1 - x_2 = -3 \\ x_1 - x_3 = 1 \\ x_1,\ x_2, x_3 \geq 0. \end{aligned}$$

1. *Résoudre le problème (P) avec la méthode du gradient réduit, on partira du point* $X^0 = \begin{pmatrix} 2 \\ 5 \\ 1 \end{pmatrix}$
2. *Vérifier que le point trouver vérifie les conditions de Kuhn-Tucker (KT).*

Corrigé :

1. $d^0 = (-49, -49, -49)^T$ *et* $\alpha^0 = \alpha_{\max} = \frac{1}{49}$
donc $X^1 = X^0 + \alpha^0 d^0 = (1, 4, 0)^T$ *et* $d_N = 0$ *stop.*
2. *on a* $\frac{\partial f}{\partial x} + UA - \lambda = (-2, 35, 0) + (2, -35, 33) - (0, 0, 33) = (0, 0, 0)$, $\lambda_i X_i^1 = 0, \lambda_i \geq 0$
donc X^1 *vérifie les conditions de Kuhn-Tucker (KT).*

Examen session rattrapage

Exercice II.10 *Modèle statistique multiple*
Considérer le problème d'estimer les paramètres de la régression

$$Y = \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3$$

*qui approche le mieux les **n** triples* (X_1^i, X_2^i, Y^i).

1. *Quelle méthode que on peut utiliser ?*
2. *Formuler ce problème comme un problème d'optimisation.*
3. *Écrire les conditions d'optimalité pour ce problème.*

Corrigé :

1. *La méthode des moindres carrés.*
2. $F(\alpha_1, \alpha_2, \alpha_3) = \min \sum_{i=1}^{n} (y_i - (\alpha_1 x_i^1 + \alpha_2 x_i^2 + \alpha_3 x_i^3))^2$.
3. $\dfrac{\partial F(\alpha}{\partial \alpha_i} = 0, \quad i = 1, 2, 3.$

Exercice II.11 *On considère le problème (P) d'optimisation sous contraintes suivante :*

$$\begin{aligned} &\min z = x_1^2 + x_2^2 \\ &\quad s.c.\ x_1 + x_2 \geq 1 \\ &\qquad x_1,\ x_2 \geq 0. \end{aligned}$$

1. *Résoudre le problème (P) avec la méthode du gradient réduit, on partira du point* $X^0 = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$
2. *Vérifier que le point trouver vérifie les conditions de Kuhn-Tucker (KT).*

Corrigé :

1. *Iteration 1 :*
   $X^0 = (0, 2, 1)^T$ *donc* $B = \{3\}$ *et* $N = \{1, 2\} \Rightarrow d^0 = (0, -4, -4)^T$ *et* $\alpha^0 = \alpha_{\max} = \dfrac{1}{4}$
   donc $X^1 = X^0 + \alpha^0 d^0 = (0, 1, 0)^T$
   Iteration 2 :
   $X^2 = (0.5, 0.5, 0)^T$ *et* $d_N = 0$ *stop.*
2. *on a* $\dfrac{\partial f}{\partial x} + UA - \lambda = (0, 0, 0)$, $\lambda_i X_i^2 = 0, \lambda_i \geq 0$ *donc* X^2 *vérifie les conditions de Kuhn-Tucker (KT).*

Année Universitaire : 2018–2019

Examen session principale

Exercice II.12 *Soit la fonction* $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ *suivante* $f(x, y) = x^3 + y^2 + 2xy$.
Soit un ensemble $X \in \mathbb{R}^2$ *suivant :* $X = \{(x, y) : x \geq 1, y \leq -1\}$.

1. *Démontrer que X est un ensemble convexe ou identifier un contre exemple pour démontrer qu'il ne l'est pas.*
2. *Démontrer que f est convexe sur X ou identifier un contre exemple pour démontrer qu'il ne l'est pas.*

Corrigé :

1. *Soit* $X = \{(x, y) : x \geq 1, y \leq -1\}$, *et* $(x^1, y^1) \in X,\ (x^2, y^2) \in X$
   on a
$$\alpha x^1 + (1-\alpha)x^2 \geq \alpha + (1-\alpha) = 1$$
$$\alpha y^1 + (1-\alpha)y^2 \leq -\alpha - (1-\alpha) = -1$$
   donc X est un ensemble convexe.

2. *Soit* $f(x, y) = x^3 + y^2 + 2xy$
   $\Rightarrow$
$$\nabla f(x, y) = \begin{bmatrix} 3x^2 + 2y \\ 2y + 2x \end{bmatrix}, \quad \nabla^2 f(x, y) = \begin{bmatrix} 6x & 2 \\ 2 & 2 \end{bmatrix}$$
   Mineurs principaux de $\nabla^2 f(x, y)$ *sont :*
$$a_{11} = 6x,\ \ a_{22} = 2,\ \ det(\nabla^2 f(x, y)) = 12x - 4$$
   puisque $x \geq 0$ *donc*
$$6x \geq 0,\ \ 12x - 4 \geq 0$$
   $\Longrightarrow \nabla^2 f(x, y)$ *est semi-défini positive sur X,*
   $\Longrightarrow f$ *est convexe sur X.*

Exercice II.13 *Une firme aéronautique fabrique des avions qu'elle vend sur deux marchés étrangers. Soit q_1 le nombre d'avions vendus sur le premier marché et q_2 le nombre d'avions vendus sur le deuxième marché. Les fonctions de demande dans les deux marchés respectifs sont :*

$$p_1 = 60 - 2q_1, \quad p_2 = 80 - 4q_2$$

où p_1 et p_2 sont les deux prix de vente. La fonction de coût total de la firme est

$$C = 50 + 40q$$

où q es le nombre total d'avions produits.
Il faut trouver le nombre d'avions que la firme doit vendre sur chaque marché pour maximiser son bénéfice.

Corrigé :

Comme $q = q_1 + q_2$, la fonction de coût devient :

$$C = 50 + 40q = 50 + 40q_1 + 40q_2$$

Le revenu total R s'obtient en multipliant le prix par la quantité sur chaque marché :

$$R = p_1q_1 + p_2q_2 = 60q_1 - 2q_1^2 + 80q_2 - 4q_2^2$$

On obtient le bénéfice B en calculant la différence entre le revenu et le coût :

$$B = R - C = 20q_1 - 2q_1^2 + 40q_2 - 4q_2^2 - 50.$$

Première condition d'optimalité :

$$\frac{\partial B}{\partial q} = 0 \Rightarrow (q_1; q_2) = (5, 5).$$

Deuxième condition d'optimalité :
La matrice Hessienne est :

$$H = \begin{bmatrix} -4 & 0 \\ 0 & -8 \end{bmatrix}$$

Par conséquent :

$$\delta_1 = -4 < 0, \quad delta_2 = 32 > 0$$

Donc (q_1, q_2) est un maximum. Le bénéfice maximum réalisé $B^* = 100$.
Quant aux prix, ils s'élèvent respectivement à :

$$p_1 = 50, \quad p_2 = 50$$

Exercice II.14 *On considère le problème (P) d'optimisation sous contraintes suivant :*

$$\begin{aligned} &\min z = x_1^2 + x_2^2 \\ &\quad s.c.\ x_1 + x_2 \geq 1 \\ &\qquad x_1,\ x_2 \geq 0. \end{aligned}$$

1. *Transformer le problème (P) vers un problème d'optimisation avec contraintes d'égalité (P1).*
2. *Déterminer le point X^0 de problème (P1) associer au sommet $A = (1, 0)^t$ de problème (P).*
3. *Résoudre le problème (P1) avec la méthode du gradient réduit, on partira du point X^0. (**1 itération**)*
4. *Vérifier que le point trouver vérifie les conditions de Kuhn-Tucker (KT).*

Corrigé :

On considère le problème (P) d'optimisation sous contraintes suivant :

1.

$$\min z = x_1^2 + x_2^2$$
$$s.c.\ x_1 + x_2 - x_3 = 1$$
$$x_1,\ x_2,\ x_3 \geq 0.$$

2.

$A = (1, 0)^t$ est un sommet de problème (P) donc $x_1 = 1, x_2 = 0$, donc $X^0 = (1, 0, x_3)$ d'après la contrainte $x_3 = 0$. $\Longrightarrow X^0 = (1, 0, 0)$.

3.

On a $x_1 = 1 - x_2 + x_3$ d'où $g(x_2, x_3) = f(1 - x_2 + x_3, x_2, x_3)$,

$$\frac{\partial g}{\partial x_N} = (-\frac{\partial f}{\partial x_1} + \frac{\partial f}{\partial x_2}\ \ \frac{\partial f}{\partial x_1} + \frac{\partial f}{\partial x_3}) = (-2x_1 + 2x_2\ \ 2x_1)$$

et $r(P_1) = (-2\ 2)$. $d_N = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$ et $d_B = -2$ donc la direction de déplacement est $\begin{pmatrix} -2 \\ 2 \\ 0 \end{pmatrix}$.

$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} -2 \\ 2 \\ 0 \end{pmatrix} \geq 0 \Rightarrow \begin{cases} \alpha \leq \frac{1}{2} \\ \alpha \leq 0 \end{cases} \Rightarrow \alpha_{\max} = \frac{1}{2}$

Le minimum de $f(1 - 2\alpha, 2\alpha, 0)$ sous la contrainte $0 \leq \alpha \leq \frac{1}{2}$ est atteint en $\alpha_2 = \frac{1}{4}$.

Le nouveau point est $X^1 = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix}$.

4.

Le gradient réduit en X^1 est $r(X^1) = (0\ \ 1)$ donc

$$d_N = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

$\Longrightarrow$ le point courant X^1 vérifie les conditions de Kuhn-Tucker (d'après le théorème).

Examen session rattrapage

Exercice II.15 *Un fabricant de postes de télévision produit q poste par semaine à un coût total de*

$$C = 6q^2 + 80q + 5000.$$

C'est un monopoleur et son prix s'exprime par la relation $p = 1080 - 4q$.

1. *Modélisé le problème pour maximiser le bénéfice.*
2. *Trouver le bénéfice net maximum.*
3. *Montrer que ce maximum est un maximum global.*

Corrigé :

1. *Le revenu total R s'obtient en multipliant le prix par la quantité :*

$$R = (pq = 1080 - 4q)q = 1080q - 4q^2$$

On obtient le bénéfice B en calculant la différence entre le revenu et le coût :

$$B = R - C = 1080q - 4q^2 - (6q^2 + 80q + 5000) = -10q^2 + 1000q - 5000$$

Donc la modélisation optimal de problème est :

$$\max B(q)$$

2. *Première condition d'optimalité :*

$$\frac{\partial B}{\partial q} = 0 \Rightarrow q = 50.$$

Sa dérivée seconde est :

$$B^{''} = -20 < 0$$

Par conséquent, la fonction de bénéfice présente un maximum en $x = 50$.

3. *Ce maximum est un maximum global car la fonction est une parabole (concave).*

Exercice II.16

1. *Utiliser la méthode des moindres carrés pour estimer les paramètres de la droite* $y = ax + b$ *qui approche le mieux les* $\boldsymbol{n}$ *couples* (x_i, y_i).
2. *Trouver la droite* $f(x) = ax + b$ *qui approche le mieux les mesures :*

$$f(0) = 0, \quad f(1) = 1, \quad f(3) = 2, \quad f(4) = 5.$$

Corrigé :

1. *Soit $(x_i)_{1\leq i\leq n}$ et y_i deux séries statistiques. On cherche à savoir s'il est raisonnable de prévoir y_i en fonction de x_i, avec par exemple une relation linéaire $y_i = ax_i + b$. Pour cela on minimise*
$$f(a,b) = \sum (ax_i + b - y_i)^2$$
par rapport à a et b, ce qui donne les équations :
$$\begin{cases} \frac{\partial f}{\partial a} = 0 & = & 2\sum_{i=1}^{n} x_i(ax_i + b - y_i) \\ \frac{\partial f}{\partial b} = 0 & = & 2\sum_{i=1}^{n}(ax_i + b - y_i) \end{cases}$$
équivalentes à :
$$\begin{cases} b\sum_{i=1}^{n} x_i + & a\sum_{i=1}^{n} x_i^2 & = & \sum_{i=1}^{n} x_i y_i \\ \sum_{i=1}^{n} b + & a\sum_{i=1}^{n} x_i & = & \sum_{i=1}^{n} y_i \end{cases}$$
On exprime b dans la seconde équation
$$b = \frac{\sum_{i=1}^{n} y_i}{n} - a\frac{\sum_{i=1}^{n} x_i}{n}$$
$\Longrightarrow b = E(y) - aE(x)$
On remplace b par sa valeur dans la 1^{er} équation pour obtenir a et on multiple par $\frac{1}{n}$ numérateur et dénominateur, on obtient
$$a = \frac{E(xy) - E(x)E(y)}{\sigma(x)^2}$$
tel que :
- $E(x)$ est la moyenne arithmétique de x
- $\sigma(x)$ est l'écart type de x.
2. *D'après la question 1*
$$a = \frac{\frac{27}{4} - 4}{10} = \frac{11}{10}$$
et
$$b = 2 - \frac{11}{10} \times 2 = -\frac{2}{10}$$
donc
$$f(t) = \frac{11}{10}t - \frac{2}{10}$$

Exercice II.17 *On considère le programme nonlinéaire (P1) suivant :*

$$\min z = x_1^2 + 2x_2^2 - x_1x_3 + 6x_3^2 - 12x_3$$
$$s.c.\ x_1 - x_3 = -3$$
$$x_1 - x_2 = 1$$
$$x_1,\ x_2, x_3 \geq 0.$$

1. *Résoudre le problème (P1) avec la méthode du gradient réduit, on partira du point* $X^0 = \begin{pmatrix} 2 \\ 1 \\ 5 \end{pmatrix}$
2. *Vérifier que le point trouver vérifie les conditions de Kuhn-Tucker (KT).*

Corrigé :

1.

Itération 1 :

Puisque le point initial est $X^0 = \begin{pmatrix} 2 \\ 1 \\ 5 \end{pmatrix}$ *donc* $B = \{1, 3\}$, $N = \{2\}$,

$A_N = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$ *et* $A_B = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$ *donc* $A_B^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}$

on sait que

$$r_N(X) = 4x_2 - (2x_1 - x_3, -x_1 + 12x_3 - 12)A_B^{-1}A_N$$

$\Longrightarrow r_N(X^0) = 49$ *donc* $d_N = -49$

$\Longrightarrow d_B = -A_B^{-1}A_N d_N^T = \begin{pmatrix} -49 \\ -49 \end{pmatrix}$

donc $d^0 = \begin{pmatrix} -49 \\ -49 \\ -49 \end{pmatrix}$ *et* $\alpha_{\max} = \frac{1}{49}$.

On calcule le pas optimale α^0

$$\alpha^0 = \arg \min_{0 \le \alpha \le \alpha_{\max}} f(X^0 + \alpha d^0)$$

On pose $\phi(\alpha) = f(X^0 + \alpha d^0)$

pour $\phi'(\alpha^*) = 0$ *on a* $\alpha^* = \frac{3855}{33614} > \alpha_{\max}$

donc $\alpha^0 = \alpha_{\max} = \frac{1}{49}$.

$\Longrightarrow X^1 = X^0 + \alpha^0 d^0 = \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix}$

Itération 2 :

On a $r_N(X^1) = 33$ *et* $x_2 = 0$ *donc* $d_N = 0$ *stop*

X^1 *est un solution optimal.*

2.

On a $\frac{\partial f}{partialx}(X^1) = (-2, 35, 0)^T$, *on pose* $u = -\frac{\partial f}{partialx_B}(X^1)A_B^{-1} = (35, -33)$ *et* $\lambda = r = (0, 0, 33)^T$

alors

$$\frac{\partial f}{partialx}(X^1) + uA - \lambda = 0$$

or $\lambda \ge 0,\ \ \lambda_i.x_i = 0,\ \ AX^1 = b$

donc X^1 *vérifie les conditions de KKT.*

Année Universitaire : 2019–2020

Examen session principale

Exercice II.18 *Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, une fonction différentiable. Soit $x_0 \in \mathbb{R}^2$ et $d \in \mathbb{R}^2$, une direction non nulle telle que $\|\nabla f(x_0) + d\| \leq \|\nabla f(x_0)\|$.*

1. *Montrer que d est une direction de descente de f en x_0.*
2. *Trouver alors le pas optimal $\rho \in \mathbb{R}$ minimisant la fonction $\phi(\rho) = \|\nabla f(x_0) + \rho d\|$.*
3. *Exhiber une direction de descente qui n'appartient pas à vect$(-\nabla f(x_0))$ lorsque $\nabla f(x_0) \neq 0$.*

Corrigé :

1. *Puisque $\|\nabla f(x_0) + d\| \leq \|\nabla f(x_0)\|$, on élève chacun des membres de l'inégalité au carré et on les développe. On obtient*
$$\|\nabla f(x_0)\|^2 + 2 < \nabla f(x_0), d >_2 + \|d\|^2 \leq \|\nabla f(x_0)\|^2$$
et par conséquent, $< \nabla f(x_0), d >_2 \leq -\frac{1}{2}\|d\|^2 < 0$.
Par conséquent,
$$\frac{f(x_0 + \epsilon d) - f(x_0)}{\epsilon} =< \nabla f(x_0), d >_2 + o(\epsilon) \leq -\frac{1}{2}\|d\|^2 + o(\epsilon).$$
Par conséquent, le second membre est strictement négatif si ϵ est assez petit et la conclusion s'ensuit.
2. *Pour trouver le pas optimal, remarquons que le problème $\min_{\mathbb{R}} \phi$ est équivalent au problème $\min_{\mathbb{R}} \phi^2$. Or, $\phi^2(\rho) = \|\nabla f(x_0)\|^2 + 2\rho < \nabla f(x_0), d >_2 + \rho^2\|d\|^2$, donc le minimum est atteint en $\rho^* = -\frac{<\nabla f(x_0), d>_2}{\|d\|^2}$ et vaut $\min_{\mathbb{R}} \phi^2 = \|\nabla f(x_0)\|^2 - \frac{<\nabla f(x_0), d>_2^2}{\|d\|^2}$.*
3. *En particulier, tout vecteur de la forme $d = -\nabla f(x_0) + \epsilon u$, où u est un vecteur unitaire et $\epsilon \in]0, \|\nabla f(x_0)\|[$ est une direction de descente.*

Exercice II.19 *Soit $f : \mathbb{R}^4 \longrightarrow \mathbb{R}$, la fonction définie par*
$$f(x_1, x_2, x_3, x_4) = (1 + x_4)^3(x_1^2 + x_2^2 + x_3^2) + x_4^2.$$

1. *Déterminer les points critiques de f.*

2. *Montrer que $f(0_{\mathbb{R}^4})$ est un minimum local. Est-ce un minimum global ?*

Corrigé :

1. *Les points critiques de f sont les solutions de l'équation $\nabla f(x_1, x_2, x_3, x_4) = 0_{\mathbb{R}^4}$. Or,*

$$\nabla f(x_1, x_2, x_3, x_4) = 0_{\mathbb{R}^4} \Leftrightarrow \left\{ \begin{array}{l} 2(1+x_4)^3 x_1 = 0 \\ 2(1+x_4)^3 x_2 = 0 \\ 2(1+x_4)^3 x_3 = 0 \\ 3(1+x_4)^2(x_1^2+x_2^2+x_3^2) + 2x_4 = 0. \end{array} \right.$$

Notons que $x_4 \neq -1$ sinon, la dernière équation fournit une contradiction. On en déduit que nécessairement $x_1 = x_2 = x_3 = 0$. Il vient $x_4 = 0$. L'unique point critique de f est donc $0_{\mathbb{R}^4}$.

2. *Évaluons la matrice hessienne de f en $0_{\mathbb{R}^4}$. Notons $X = (x_1, x_2, x_3, x_4)$. On a*

$$\mathrm{Hess} f(X) = \begin{pmatrix} 2(1+x_4)^3 & 0 & 0 & 6x_1(1+x_4)^2 \\ 0 & 2(1+x_4)^3 & 0 & 6x_2(1+x_4)^2 \\ 0 & 0 & 2(1+x_4)^3 & 6x_3(1+x_4)^2 \\ 6x_1(1+x_4)^2 & 6x_2(1+x_4)^2 & 6x_3(1+x_4)^2 & 2+6(1+x_4)(x_1^2+x_2^2+x_3^2) \end{pmatrix}$$

et par conséquent, $\mathrm{Hess} f(0_{\mathbb{R}^4}) = 2I_4$ est définie positive. On en déduit que $f(0_{\mathbb{R}^4}) = 0$ est un minimum local. Il n'est en revanche pas global car

$$f(1,0,0,t) = (1+t)^3 + t^2 \xrightarrow[t \to -\infty]{} -\infty$$

Exercice II.20 *Soit $N \in \mathbb{N}^*$. On considère un nuage de points $\{(t_i, x_i)\}_{1 \leq i \leq N}$, et on cherche à mettre en ?uvre une régression parabolique, autrement dit, on recherche la parabole P d'équation $y = at^2 + bt + c$, où a, b et c sont trois réels à déterminer, telle que la somme sur tous les indices i variant de 1 à N du carré de la distance du point (t_i, x_i) au point de même abscisse sur P soit minimale.*

1. *Écrire ce problème comme un problème de minimisation quadratique, c'est-à-dire un problème de la forme*

$$\min_{X \in \mathbb{R}^N} J(X) = \frac{1}{2} < AX, X > - < b, X >, \qquad (Q)$$

avec $A \in M_n(R)$, $b \in \mathbb{R}^n$. On devra donc expliciter n, A et b.
On utilisera la notation $S_k = \sum_{i=1}^{N} t_i^k$.

2. *Discuter de l'existence des solutions d'un tel problème.*
3. *On suppose que la matrice A est définie positive. Démontrer que (Q) possède une unique solution.*

Corrigé :

1. *Le problème s'écrit*

$$\min_{(a,b,c)\in\mathbb{R}^3} f(a,b,c) = \sum_{i=1}^{N}(x_i - at_i^2 - bt_i - c)^2.$$

Écrivons $J(X) = ||MX - k||^2$ *avec* $X = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$, $M = \begin{pmatrix} t_1^2 & t_1 & 1 \\ \vdots & \vdots & \vdots \\ t_N^2 & t_N & 1 \end{pmatrix}$

et $k = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}$. *D'après la méthode des moindres carrés, on a*

$$J(X) = \frac{1}{2} < Ax, X > - < b, X >$$

avec $n = 3$, $A = M^T M \in M_3(\mathbb{R})$ *et* $b = M^T k \in \mathbb{R}^3$. *On calcule* $A = \begin{pmatrix} S_4 & S_3 & S_2 \\ S_3 & S_2 & S_1 \\ S_2 & S_1 & N \end{pmatrix}$.

2. *Ce problème est équivalent au problème de minimiser la distance euclidienne de* k *au sous espace vectoriel (de dimension finie)* $Im(M)$. *C'est donc un problème de projection orthogonale, et il admet une solution.*

3. *Dans ce cas, on sait que* $Hess J(X) = A$ *qui est définie positive. Par conséquent,* J *est strictement convexe, et* J *possède au plus un minimum dans* $\mathbb{R}^3$. *Comme on a vu qu'elle en possède au moins un, on conclut à l'existence et l'unicité.*

Exercice II.21 *On souhaite résoudre numériquement le problème d'optimisation*

$$\min_{(x,y)\in\mathbb{R}^2} f(x,y) = x^4 + y^4 - xy,$$

à l'aide d'une méthode de gradient à pas constant, noté $\rho > 0$. *On appelle* $(x_k, y_k)_{k\in\mathbb{N}}$ *la suite des itérés obtenus.*

1. *Donner la relation de récurrence satisfaite par la suite* $(x_k, y_k)_{k\in\mathbb{N}}$.
2. *Quel critère d'arrêt numérique proposez-vous pour cet algorithme ?*

Corrigé :

1. f *est* C^∞ *car polynomiale et on a* $\nabla f(x,y) = (4x^3 - y, 4y^3 - x)$. *La relation de récurrence dans la méthode du gradient à pas constant s'écrit* $X_{k+1} = X_k - \rho\nabla f(X_k)$, *qui s'écrit ici*

$$\begin{cases} x_{k+1} = x_k - \rho(4x_k^3 - y_k) \\ y_{k+1} = y_k - \rho(4y_k^3 - x_k). \end{cases}$$

2. *On fixe* $tol > 0$. *Un critère raisonnable est de stopper l'algorithme lorsque*

$$|x_{k+1} - x_k| + |y_{k+1} - y_k| < tol$$

ou encore lorsque

$$|4x_k^3 - y_k| + |4y_k^3 - x_k| < tol.$$

Examen session rattrapage

Exercice II.22 *On définit la famille des* $\{u_i\}_{i\in\{0,\ldots,N+1\}}$ *par* $u_i = ih$*, avec* $h = \frac{1}{N+1}$ *et* $N \in \mathbb{N}^*$ *donné. On se donne un nuage de points de* $\mathbb{R}^2$ $(u_i, x_i)_{i\in\{0,\ldots,N+1\}}$*. On suppose par ailleurs que* $x_0 = 0$ *et* $x_{N+1} = 1$*. Posons* $x = (x_1, \ldots, x_N)$*. On appelle* $f(x)$*, la longueur de la courbe affine par morceaux passant par les points* (u_i, x_i)*.*

1. *Calculer la distance entre deux points* $A_{i+1} = (u_{i+1}, x_{i+1})$ *et* $A_i = (u_i, x_i)$*.*
2. *Montrer que pour tout* $x \in \mathbb{R}^N$*, on a*
$$f(x) = h\sum_{i=0}^{N}\sqrt{1 + (\frac{x_{i+1} - x_i}{h})^2}.$$
3. *Montrer que* f *est coercive (*$f(x) \geq \|x\|_\infty$*), et en déduire l'existence de la solution pour le problème* $\min f(x)$*.*
4. *Étudier la convexité de* f*, et en déduire l'unicité de la solution pour le problème* $\min f(x)$*.*

Corrigé :

1. $d(A_{i+1}, A_i) = \|(u_{i+1}, x_{i+1}) - (u_i, x_i)\|_2 = \sqrt{(u_{i+1} - u_i)^2 - (x_{i+1} - x_i)^2}$
2. $f(x) = \sum_{i=0}^{N}\sqrt{(u_{i+1} - u_i)^2 + (x_{i+1} - x_i)^2} = \sum_{i=0}^{N}\sqrt{h^2 - (x_{i+1} - x_i)^2} = h\sum_{i=0}^{N}\sqrt{1 + (\frac{x_{i+1}-x_i}{h})^2}$
3. *Soit* $k \in [1, N]$ *et* $x = (x_1, \ldots, x_N) \in \mathbb{R}^N$*. On a successivement*
$$\begin{aligned} f(x) = \; & h\sum_{i=0}^{N}\sqrt{1 + (\tfrac{x_{i+1}-x_i}{h})^2} \geq h\sum_{i=0}^{k-1}\sqrt{1 + (\tfrac{x_{i+1}-x_i}{h})^2} \\ \geq \; & h\sum_{i=0}^{k-1}\sqrt{(\tfrac{x_{i+1}-x_i}{h})^2} \geq h\sum_{i=0}^{k-1}|\tfrac{x_{i+1}-x_i}{h})| \\ \geq \; & \sum_{i=0}^{k-1}|x_{i+1} - x_i| \geq |\sum_{i=0}^{k-1}(x_{i+1} - x_i)| \geq |x_k - x_0| \geq |x_k|, \end{aligned}$$
d'après l'inégalité triangulaire f *est donc coercive.*
Puisque f *est continue, coercive sur* $\mathbb{R}^N$ *qui est fermé, le problème* $\min f$ *possède une solution.*
4. *Montrons que* f *est strictement convexe. Pour cela, remarquons d'abord que la fonction* $g : x \mapsto \sqrt{1 + x^2}$ *est strictement convexe car sa dérivée seconde est donnée par* $g''(x) = 1/(1 + x^2)^{3/2}$ *et est donc strictement positive sauf en 0.*
Soit $(x, y) \in (\mathbb{R}^N)^2$ *tel que* $x \neq y$*. ce qui signifie qu'il existe au moins un indice* $j \in [1, N]$ *tel que* $x_j \neq y_j$*. Par conséquent, il existe un rang* $k \in [0, N]$ *tel que* $x_{k+1} - x_k \neq y_{k+1} - y_k$*. En effet, si ce n'était pas le cas, puisque* $x_0 = y_0 = 0$*, on en déduirait successivement que* $x_1 = y_1$*,* $x_2 = y_2$ *et ainsi de suite jusqu'a* $x_N = y_N$*, et on aurait donc* $x = y$*. Par conséquent, pour tout* $t \in]0, 1]$*, on a*
$$g(t(x_{k+1} - x_k) + (1 - t)(y_{k+1} - y_k)) < tg(x_{k+1} - x_k) + (1 - t)g(y_{k+1} - y_k).$$

Soit $t \in]0,1]$. En utilisant la stricte convexité de g, il vient :

$$\begin{aligned}
f(tx+(1-t)y) = \;& h\sum_{i=0}^{N} g(t(\tfrac{x_{i+1}-x_i}{h})+(1-t)(\tfrac{y_{i+1}-y_i}{h}))\\
= \;& h\sum_{i=0, i\neq k}^{N} g(t(\tfrac{x_{i+1}-x_i}{h})+(1-t)(\tfrac{y_{i+1}-y_i}{h}))\\
& +hg(t(\tfrac{x_{k+1}-x_k}{h})+(1-t)(\tfrac{y_{k+1}-y_k}{h}))\\
< \;& h\sum_{i=0, i\neq k}^{N} (tg(\tfrac{x_{i+1}-x_i}{h})+(1-t)g(\tfrac{y_{i+1}-y_i}{h}))\\
& +h(tg(\tfrac{x_{k+1}-x_k}{h})+(1-t)g(\tfrac{y_{k+1}-y_k}{h}))\\
& = h\sum_{i=0}^{N}(tg(\tfrac{x_{i+1}-x_i}{h})+(1-t)g(\tfrac{y_{i+1}-y_i}{h}))\\
& = tf(x)+(1-t)f(y).
\end{aligned}$$

On en déduit que f est strictement convexe. On en déduit que le problème $\min f(x)$ *admet au plus une solution.*
En combine avec le résultat d'existence obtenu précédemment, il vient que ce problème possède une unique solution.

Exercice II.23 *On se donne $n \in \mathbb{N}^*$, $b \in \mathbb{R}^n$ un vecteur et $A \in M_n(\mathbb{R})$ une matrice symétrique et semi-définie positive. On considère la fonction $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ définie par*

$$f(x) = \epsilon\|x\|^2 + \frac{1}{2} < Ax, x > - < b, x >, \quad x \in \mathbb{R}^n.$$

I) $\epsilon > 0$,

1. *Montrer que f est une fonction quadratique et calculer ∇f et $\nabla^2 f$.*
2. *Montrer que f est fortement convexe et en déduire l'existence et l'unicité d'un point de minimum de f sur $\mathbb{R}^n$; on va noter par x^* ce point de minimum.*
3. *Écrire l'équation qui permet de trouver x^*.*
4. *Trouver x^* dans le cas particulier*

$$n = 2, \;\; \epsilon = \frac{1}{100}, \;\; b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \;\; A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

II) $\epsilon = 0$, on va noter dans la suite par u^* l'unique point de minimum de f sur $\mathbb{R}^n$. Nous appliquons ***la méthode de gradient à pas optimal*** *qui consiste à construire une suite $\{u^{(k)}\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ par la relation de récurrence vue en cours, avec $u^{(0)}$ donnée.*

1. *Montrer qu'on a*

$$u^{(k+1)} - u^* = u^{(k)} - u^* - \rho_k A(u^{(k)} - u^*), \;\; \forall k \in \mathbb{N},$$

   avec $\rho_k \in R$ le pas optimal.

2. *Programmer cet algorithme de gradient à pas optimal sous Python.*

Corrigé :

I) $\epsilon > 0$,

1. *On a*
$$f(x) = \epsilon||x||^2 + \frac{1}{2} < Ax, x > - < b, x >= \epsilon < I_n x, x > + \frac{1}{2} < Ax, x > - < b, x >$$
forme quadratique,
$$\nabla f(x) = 2\epsilon I_n x + Ax - b, \quad \forall x \in \mathbb{R}^n$$
$$\nabla^2 f(x) = 2\epsilon I_n + A, \quad \forall x \in \mathbb{R}^n$$

2. *Si* $w \in \mathbb{R}^n$ *arbitraire alors*
$$< \nabla^2 f(x)w, w >= 2\epsilon < I_n w, w > + < Aw, w >\geq 2\epsilon||w||_2 > 0$$
donc f *est fortement convexe.*
Alors d'après le théorème on a l'existence et l'unicité d'un point minimum x^*.

3. x^* *satisfait l'équation d'Euler* $\nabla f(x^*) = 0$ *donc*
$$2\epsilon x^* + Ax^* = b$$

4. $$\frac{2}{100}\begin{pmatrix} x_1^* \\ x_2^* \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x_1^* \\ x_2^* \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$
donc $x_1^* = 0$ *et* $x_2^* = 50$.

II) $\epsilon = 0$,

1. u^* *satisfait* $Au^* = b$ *donc*
$$u^{(k+1)} - u^* = u^{(k)} - u^* - \rho_k \nabla J(u^{(k)}) = u^{(k)} - u^* - \rho_k(Au^{(k)} - b) = u^{(k)} - u^* - \rho_k A(u^{(k)} - u^*).$$