Epreuve de rattrapage

(durée : 02 heures)

Exercice [**06 pts**]

Etant données trois valeurs réelles positives a, b et c $\in \mathbf{R}^+$,

1) Déterminer l'enveloppe convexe des quatre points de $\mathbf{R}^2$: $O = (0,0)^T$, $A = (-a,a)^T$, $B = (b,b)^T$ et $C= (c,-c)^T$ selon une discussion sur les valeurs positives de a, b et c. (**4pts**)

2) On remplace le point A par $A' = (-2a\ ,\ a)^T$ et on suppose que les trois valeurs a, b et c $\in \left]0, +\infty\right[$, déterminer alors **CO**({O, A', B, C}). (**1pt**)

3) On remplace, à présent, le point A par $A'' = ((-1/2)a\ ,\ a)^T$ et on suppose tjrs que a, b et c > 0, déterminer alors **CO**({O, A'', B, C}). (**1pt**)

Indication : Dans cet exercice, on pourra se contenter de réponses schématiques (figures géométriques) pour justifier les réponses données.

Problème [**14 pts**]

Dans $\mathbf{R}^n$, $\langle\ ,\ \rangle_n$ désigne le produit scalaire euclidien et on considère la matrice carrée B d'ordre **n**.

1) Montrer que $\forall\ (x\ ,\ y) \in \mathbf{R}^n x \mathbf{R}^n$ on a $\langle Bx\ ,\ y\rangle_n = \langle x\ ,\ B^T y\rangle_n$ et $\langle x\ ,\ By\rangle_n = \langle B^T x\ ,\ y\rangle_n$. (**1pt**)

2) On introduit alors la fonctionnelle **f** définie sur $\mathbf{R}^n$: $\mathbf{f}(x) = \|Bx\|_n^2 + \langle c\ ,\ x\rangle_n + d$ où $\|\ .\ \|_n$ désigne la norme euclidienne dans $\mathbf{R}^n$, **c** est un vecteur de $\mathbf{R}^n$ et d un réel.

   a) Montrer que **f** peut s'écrire sous la forme : $\mathbf{f}(x) = \langle B^T B x\ ,\ x\rangle_n + \langle c\ ,\ x\rangle_n + d$. (**0.5pt**)

   b) On pose $\mathbf{A} = B^T B$. Montrer alors que **f** est une fonctionnelle quadratique. (**0.5pt**)

   c) Après avoir déterminé $\nabla \mathbf{f}(x) = \text{grad}\ \mathbf{f}(x)$ et Hess**f**(x), montrer que **A** est semi-définie positive et déduire que **f** est convexe. (**2pts**)

   Rappel : Dans $\mathbf{R}^n$, pour une fonctionnelle quadratique **g** de la forme : $\mathbf{g}(x) = \frac{1}{2}\langle \Omega x\ ,\ x\rangle_n - \langle \gamma\ ,\ x\rangle_n + \delta$, la dérivée directionnelle d'ordre 1 de **g** s'écrit comme suit : $\nabla \mathbf{g}(x) = \text{grad}\ \mathbf{g}(x) = \Omega x - \gamma$ **?** et $\text{Hess}\mathbf{g}(x) = \Omega$.

   d) En précisant un contre-exemple, montrer qu'à ce niveau **f** ne peut être coercive si **A** est seulement semi-définie-positive. (**0.5pt**)

   e) Montrer que si B est inversible alors **A** est aussi inversible. (**0.5pt**)

   Rappel : B inversible $\Rightarrow B^T$ inversible et son inverse est tout simplement la transposée de B^{-1} .

   f) Montrer alors que, dans ce cas (**A** inv.), **A** est aussi définie positive. (**0.5pt**)

   g) En déduire alors que dans ce cas **f** est coercive. (**1pt**)

   h) Dans la suite on suppose A définie positive (B inversible), montrer alors que :

   h1) le problème d'optimisation sans contraintes ($\mathbf{P_n}$) : $\min_{x \in R^n} f(x)$ admet une solution unique (sans la déterminer) notée x*. (**1pt**)

   h2) ''résoudre le problème ($\mathbf{P_n}$)'' revient à ''résoudre le système linéaire Bx = **c'** '' où **c'** est un vecteur de $\mathbf{R}^n$ qu'il faudra préciser (ici il faut, tout simplement, montrer l'équivalence suivante : x* solution optimale de ($\mathbf{P_n}$) $\Leftrightarrow$ x* vecteur-solution du système linéaire Bx = **c'**). (**1pt**)

   h3) Résoudre alors le problème ($\mathbf{P_n}$) en explicitant x* en fonction des données du problème initial c.àd. en fonction de la matrice B, du vecteur **c** et éventuellement du scalaire d. (**0.5pt**)

h4) Calculer la valeur optimale de ($\mathbf{P_n}$) en fonction de ces mêmes données (de **h3)**). (**1pt**)

i) **Application** : Soit à résoudre le problème d'optimisation suivant dans $\mathbf{R^2}$: ($\mathbf{P_2}$) : $\min_{x\in R^2} f(x)$
où $f(x) = f(x_1, x_2) = ((x_1 + 2x_2)^2 + (2x_1 + x_2)^2) + 2x_1 + 2x_2 + 2$ avec $x = (x_1, x_2) \in \mathbf{R^2}$.

i1) Par identification avec l'expression de la fonction **f** introduite ci-dessus (lorsque n = 2) :

$$\mathbf{f}(x) = \|Bx\|_2^2 + \langle c, x\rangle_2 + d$$

préciser la matrice carrée B d'ordre 2, le vecteur **c** de $\mathbf{R^2}$ et le scalaire $d \in \mathbf{R}$. (**2.5pts**)

i2) Après avoir vérifié que B est inversible et en utilisant les résultats précédents, déterminer l'unique solution optimale x^* du problème de minimisation ($\mathbf{P_2}$). (**1pt**)

i3) Calculer la valeur optimale de ($\mathbf{P_2}$). (**0.5pt**)

Rappels utiles

Rappel 1 : Conditions d'optimalité.
E étant un R-esp. de Hilbert, U ouvert dans E et C convexe fermé inclus dans U, alors si $f : E \to \mathbf{R}$ est **convexe** et une fois dérivable (au sens de la dérivée directionnelle) dans U, une condition nécessaire est suffisante pour que $x \in C$ soit solution optimale du problème d'optimisation : $\inf_{y\in C} f(y)$ est : $\forall y \in C \ \langle f'(x), y - x)\rangle_E \geq 0$. C'est l'inéquation d'Euler.
Dans le cas sans contraintes (C = U = E), une condition nécessaire est suffisante pour que $x \in E$ soit solution optimale du problème d'optimisation **sans** contraintes : $\inf_{y\in E} f(y)$ est : $f'(x) = 0_E$. C'est l'équation d'Euler.

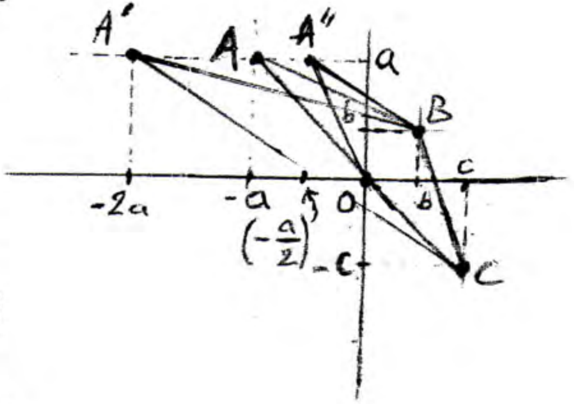
Rappel 2 : Coercivité et optimalité.
E étant un R-esp. normé de dimension finie, si $f : E \to \mathbf{R}$ est **continue** et **coercive** (c'est-à-dire $\lim_{\|x\|_E\to\infty} f(x) = +\infty$), alors la fonction f admet au moins un minimum sur E.
Si de plus f est **strictement convexe** alors ce minimum est unique.

Université Abou Bekr Belkaid, Tlemcen. Année universitaire 2017/2018
Faculté des Sciences Juin 2018.
Département de Mathématiques. Optimisation sans contraintes (L3 - S5)

Corrigé de l'épreuve de rattrapage

Exercice

a, b et $c \in \mathbb{R}^+$.



1) Si $a = b = c = 0$ alors $CO(\{0, A, B, C\}) = \{0\} = \{(0,0)^T\}$

Si $b = 0$ et ($a \neq 0$ et $c \neq 0$) alors $CO(\{0, A, B, C\}) = [A, C]$

$b = 0$ et ($a = 0$ et $c \neq 0$) alors $CO(\{0, A, B, C\}) = [0, C]$

$b = 0$ et ($a \neq 0$ et $c = 0$) alors $CO(\{0, A, B, C\}) = [0, A]$

Si $b \neq 0$ et ($a = 0$ et $c = 0$) alors $CO(\{0, A, B, C\}) = [0, B]$

$b \neq 0$ et ($a \neq 0$ et $c = 0$) alors $CO(\{0, A, B, C\}) = (0AB)$

$b \neq 0$ et ($a = 0$ et $c \neq 0$) alors $CO(\{0, A, B, C\}) = (0BC)$

$b \neq 0$ et ($a \neq 0$ et $c \neq 0$) alors $CO(\{0, A, B, C\}) = (ABC)$

2) $A' = (-2a, a)^T$ remplace $A \Rightarrow CO(\{0, A', B, C\}) = (A'BC)$

3) $A'' = \left(\frac{-1}{2}a, a\right)^T$ remplace $A \Rightarrow CO(\{0, A'', B, C\}) = (0A''BC)$

Problème

1) $\forall x, y \in \mathbb{R}^n \ \langle Bx, y\rangle_n = (Bx)^T y = x^T(B^T y) = \langle x, B^T y\rangle_n$

$\forall x, y \in \mathbb{R}^n \ \langle x, By\rangle_n = x^T(By) = x^T (B^T)^T y = (B^T x)^T y = \langle B^T x, y\rangle_n$

où l'on rappelle que $(CD)^T = D^T C^T$

2) $f : \mathbb{R}^n \longrightarrow \mathbb{R}$

$x \longmapsto f(x) = \|Bx\|_n^2 + \langle c, x\rangle_n + d = \langle Bx, Bx\rangle_n + \langle c, x\rangle_n + d$ $\quad c \in \mathbb{R}^n$, $d \in \mathbb{R}$

→ a) D'après 1) $\langle Bx, Bx\rangle_n = \langle B^T B x, x\rangle_n$ d'où le résultat.

→ b) On pose $A = B^T B$. f est quadratique ssi A est symétrique

or $A^T = (B^T B)^T = B^T (B^T)^T = B^T B = A \Rightarrow A$ sym. $\Rightarrow f$ quadratique

→ c) A ce niveau on pourra écrire f sous la forme :

$f(x) = \frac{1}{2} \langle 2B^T B x, x\rangle_n - \langle -c, x\rangle_n + d$. On pose alors $\Omega = 2B^T B$ et $\gamma = -c$

D'après le rappel $\nabla f(x) = \Omega x - \gamma = 2B^T B x + c = 2Ax + c$ et $\mathrm{Hess} f(x) = 2A$

A est semi-déf. positive ssi, par définition, $\forall x \in \mathbb{R}^n \ x^T A x = \langle Ax, x\rangle_n$

On se rappelle alors que $\langle Ax, x\rangle_n = \langle B^T B x, x\rangle_n = \langle Bx, Bx\rangle_n$ $= \langle x, Ax\rangle_n \geq 0$

On en déduit que f est convexe $= \|Bx\|_n^2 \geq 0 \ \ \forall x \in \mathbb{R}^n$ $\quad A$ étant sym.

puisque $\mathrm{Hess} f(x) = 2A$ est semi-définie positive

→ d) Si A est la matrice nulle (symétrique et semi-déf. positive) alors f est une fonction affine sur $\mathbb{R}^n$: $f(x) = \langle c, x\rangle_n + d$ qui ne peut être coercive puisque $f(x) = f(x_1, \dots, x_n) = \sum_{i=1}^{n} c_i x_i + d$ $c_i \in \mathbb{R}\ \forall i = \overline{1,n}$
$\|x\|_n \to \infty \Rightarrow \exists$ au moins $x_{i_0} \to \pm\infty$. Supposons $x_{i_0} \to +\infty$ et lui correspond $c_{i_0} < 0$
Ds ce cas $f(x) \to -\infty$ et $x_{j_0} \to -\infty$ et lui correspond $c_{j_0} > 0$ alors ds ce cas aussi $f(x) \to -\infty$

→ e) B inversible $\Rightarrow A = B^T B$ inversible puisque $A^{-1} = (B^T B)^{-1} = B^{-1}(B^T)^{-1} = B^{-1}(B^{-1})^T$
puisqu'en général, le produit de 2 matrices inversibles est inversible

→ f) A étant inversible et semi-définie positive alors ttes ses valeurs propres sont réelles et strict. positives car si l'une des valeurs propres est nulle alors $\det(A) = 0$ et A serait singulière. On en conclut alors que A est déf.-positive

→ g) A étant définie-positive, f est alors coercive puisque, ds ce cas, f peut s'écrire sous la forme : $f(x) = \frac{1}{2}\|x\|_\Omega^2 - \langle \delta, x\rangle_n + d \geq \frac{1}{2}\|x\|_\Omega^2 - \alpha\|\delta\|_n \|x\|_n + d$
où $\|x\|_\Omega^2 = \langle \Omega x, x\rangle_n$ $(\Omega = 2A,\ \delta = -c)$ $= \|x\|_\Omega\left(\frac{1}{2}\|x\|_\Omega - \alpha\|\delta\|_n\right) + d \to +\infty$ qd $\|x\|_A \to \infty$
et α est l'une des constantes d'équivalence entre $\|\cdot\|_A$ et $\|\cdot\|_n$

→ h) La matrice A étant définie-positive $\Rightarrow f$ strict. convexe

h1) Pour que (P_n) : $\min_{x \in \mathbb{R}^n} f(x)$ admette une sol. opt. unique, il suffit que f soit continue (par rapport à la norme euclidienne $\|\cdot\|_n$ par exple), coercive et strict. convexe (Voir Rappel 2). Ce qui est le cas pour f :

On a déjà vu que f était coercive et strict. convexe. Pour la continuité de f il suffit de remarquer que c'est une somme de fonctions continues : $x \mapsto \|Bx\|_n^2$ (polynôme du 2nd degré à n variables), $x \mapsto \langle c, x\rangle_n$ (application ou forme linéaire sur $\mathbb{R}^n$) et $x \mapsto d$ (fonction constante)

h2) On note par x^* la sol. opt. unique de (P_n) et on montre que x^* sol. unique de $Bx = c'$. En effet, x^* vérifie l'équation d'Euler
$f'(x^*) = 0_{\mathbb{R}^n} \Leftrightarrow \nabla f(x^*) = 0_{\mathbb{R}^n} \Leftrightarrow 2Ax^* + c = 0_n \Leftrightarrow 2Ax^* = -c$
$\Leftrightarrow 2B^T B x^* = -c \Leftrightarrow Bx^* = -\frac{1}{2}(B^T)^{-1}c = \left(-\frac{1}{2}\right)(B^{-1})^T c = c'$
Inversement si x^* sol. de $Bx = c'$ c.à.d. si $Bx^* = c'$
alors c'est équivalent d'écrire $2Ax^* + c = 0_n \Leftrightarrow \nabla f(x^*) = 0_n$
$\Rightarrow x^*$ sol. opt. de (P_n) car f convexe puisqu'elle est strict. conv.

h3) $Bx^* = \left(-\frac{1}{2}\right)(B^{-1})^T c \Leftrightarrow x^* = \left(-\frac{1}{2}\right)B^{-1}(B^{-1})^T c$

h4) La valeur optimale est $f(x^*) = \|Bx^*\|_n^2 + \langle c, x^*\rangle_n + d$

$$\boxed{\begin{aligned} &= \left\|\left(-\tfrac{1}{2}\right) B B^{-1}(B^{-1})^T c\right\|_n^2 + \left\langle c, \left(-\tfrac{1}{2}\right) B^{-1}(B^{-1})^T c\right\rangle_n + d \\ &= \tfrac{1}{4}\left\|(B^{-1})^T c\right\|_n^2 - \tfrac{1}{2}\left\langle c, B^{-1}(B^{-1})^T c\right\rangle_n + d \end{aligned}}$$

→ i) Application: (P_2) : $\min_{x \in \mathbb{R}^2} f(x)$

où $f(x) = f(x_1, x_2) =$

$$= \left[(x_1 + 2x_2)^2 + (2x_1 + x_2)^2\right] + 2x_1 + 2x_2 + 2 \text{ avec } x = (x_1, x_2)^T \in \mathbb{R}^2$$

$$= \left\|\begin{pmatrix} x_1 + 2x_2 \\ 2x_1 + x_2 \end{pmatrix}\right\|_2^2 + 2x_1 + 2x_2 + 2 = \left\|\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right\|_2^2 + \left\langle \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right\rangle_2 + 2$$

i1) Il est clair alors que, par identification avec l'expression de la fction f lorsque $n=2$, la matrice $B = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$, le vecteur $c = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$ et $d = 2$

i2) $\det B = 1 - 4 = -3 \neq 0$ 0.25 $\Rightarrow B$ inversible $B^{-1} = \begin{pmatrix} a & b \\ \alpha & \gamma \end{pmatrix}$ $B^{-1}B = I_2 \Leftrightarrow \begin{pmatrix} a & b \\ \alpha & \gamma \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

$\Rightarrow \begin{cases} a + 2b = 1 \\ 2a + b = 0 \end{cases}$ et $\begin{cases} \alpha + 2\gamma = 0 \\ 2\alpha + \gamma = 1 \end{cases}$

$0 + 3b = 2$ $\qquad$ $0 + 3\gamma = -1$

$b = 2/3 \Rightarrow a = -b/2 = -\frac{1}{2}(2/3) = -1/3$

$\gamma = -1/3 \Rightarrow \alpha = -2\gamma = -2(-1/3) = 2/3$

$B^{-1} = \begin{pmatrix} -1/3 & 2/3 \\ 2/3 & -1/3 \end{pmatrix}$ Vérif. $B^{-1}B = \begin{pmatrix} -1/3 & 2/3 \\ 2/3 & -1/3 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

D'après h3) $x^* = \left(\frac{-1}{2}\right)B^{-1}(B^{-1})^T c = \left(\frac{-1}{2}\right)\begin{pmatrix} -1/3 & 2/3 \\ 2/3 & -1/3 \end{pmatrix}\begin{pmatrix} -1/3 & 2/3 \\ 2/3 & -1/3 \end{pmatrix}\begin{pmatrix} 2 \\ 2 \end{pmatrix} = \left(\frac{-1}{2}\right)\begin{pmatrix} -1/3 & 2/3 \\ 2/3 & -1/3 \end{pmatrix}^2\begin{pmatrix} +2 \\ +2 \end{pmatrix} = \begin{bmatrix} 5/9 & -4/9 \\ -4/9 & 5/9 \end{bmatrix}\begin{pmatrix} 2 \\ 2 \end{pmatrix}\left(-\frac{1}{2}\right)$

Donc $x^* = (-1/9, -1/9)^T$

i3) D'après h4) $f(x^*) = f(-1/9, -1/9) = \frac{1}{4}\left\|(B^{-1})^T c\right\|_2^2 + \langle c, x^*\rangle_2 + d$ où $d = 2$

Donc $f(x^*) = \frac{1}{4}\left\|\begin{pmatrix} -1/3 & 2/3 \\ 2/3 & -1/3 \end{pmatrix}\begin{pmatrix} +2 \\ +2 \end{pmatrix}\right\|_2^2 + \left\langle \begin{pmatrix} +2 \\ +2 \end{pmatrix}, \begin{pmatrix} -1/9 \\ -1/9 \end{pmatrix}\right\rangle_2 + 2 = \frac{1}{4}\left\|\begin{pmatrix} +2/3 \\ +2/3 \end{pmatrix}\right\|_2^2 - \left(\frac{2}{9} + \frac{2}{9}\right) + 2$

$= \frac{1}{4}\left\langle \begin{pmatrix} +2/3 \\ +2/3 \end{pmatrix}, \begin{pmatrix} +2/3 \\ +2/3 \end{pmatrix}\right\rangle_2 - \frac{4}{9} + 2 = \frac{1}{4}\left(\frac{4}{9} + \frac{4}{9}\right) - \frac{4}{9} + 2 = \frac{2}{9} - \frac{4}{9} + 2 = -\frac{2}{9} + 2$

$= 16/9$

$B^{-1} = \begin{bmatrix} -1/3 & 2/3 \\ 2/3 & -1/3 \end{bmatrix}$ et $c = (-2, -2)^T$