Université Abou Bekr Belkaid, Tlemcen,
Faculté des Sciences,
Département de Mathématiques.

Année universitaire 2016/2017.
Janvier 2017.
Optimisation sans contraintes. (L3-S5)

Epreuve Finale

(durée : 02 heures)

Questions de cours [**09 pts**]
Caractérisation de la condition d'optimalité par le cône normal

1) Montrer que dans un espace vectoriel E **a** $\in$ **-A** $\Leftrightarrow$ **-a** $\in$ **A** où **a** est un vecteur non nul de E et **A** un sous-ensemble de E. (**1pt**)
2) E étant un **R**_espace de Hilbert, C une partie convexe fermée de E et **g** une fonction convexe sur C et dérivable (au sens de la dérivée directionnelle) dans un ouvert contenant C alors on définit le cône normal à C en x par $N(\mathrm{C},\mathrm{x}) = \{\mathrm{y}\in\mathrm{E} / \langle \mathrm{y} , \mathrm{x}\rangle_{\mathrm{E}} \geq \langle \mathrm{y} , \mathrm{z}\rangle_{\mathrm{E}}\ \forall \mathrm{z}\in\mathrm{C}\} = \{\mathrm{y}\in\mathrm{E} / \langle \mathrm{y} , \mathrm{z}-\mathrm{x}\rangle_{\mathrm{E}} \leq 0\ \forall \mathrm{z}\in\mathrm{C}\}$.
   a) Montrer que $\mathbf{0}_{\mathrm{E}} \in N(\mathrm{C},\mathrm{x})$ (c.àd. N(C,x) est toujours non vide). (**1pt**)
   b) Montrer que N(C,x) est une partie convexe de E. (**1pt**)
   c) Montrer que l'on peut écrire N(C,x) sous la forme $N(\mathrm{C},\mathrm{x}) = \{\mathrm{y}\in\mathrm{E} / \langle \mathrm{y} , \mathrm{z'}\rangle_{\mathrm{E}} \leq 0\ \forall \mathrm{z'}\in \mathrm{C} - \{\mathrm{x}\}\}$.(**1pt**)
   d) Montrer que si x $\in$ C et x est une solution optimale de (**P**) : $\alpha = \inf\limits_{\mathrm{y}\in\mathrm{C}} \mathrm{g}(\mathrm{y})$ alors **g**'(x) $\in$ -N(C,x). (**1pt**)
   e) Montrer que si x $\in$ C et t.q. **g**'(x) $\in$ -N(C,x) alors x est solution optimale de (**P**). (**1pt**)
   f) Montrer que si C = E alors $N(\mathrm{E},\mathrm{x}) = \{\mathbf{0}_{\mathrm{E}}\}$. (**2pts**)
   En déduire qu'on a bien l'équation d'Euler quand il s'agit d'un problème d'optimisation sans contraintes. (**1pt**)

Exercice [**11 pts**]
Dans $\mathbf{R}^{\mathbf{n}}$, $\langle\ ,\ \rangle_n$ désigne le produit scalaire euclidien et on considère la matrice carrée B d'ordre **n**.
1) Montrer que $\forall$ (x , y) $\in \mathbf{R}^{\mathbf{n}}\mathrm{x}\mathbf{R}^{\mathbf{n}}$ on a $\langle \mathrm{Bx} , \mathrm{y}\rangle_n = \langle \mathrm{x} , \mathrm{B}^{\mathrm{T}}\mathrm{y}\rangle_n$ et $\langle \mathrm{x} , \mathrm{By}\rangle_n = \langle \mathrm{B}^{\mathrm{T}}\mathrm{x} , \mathrm{y}\rangle_n$. (**2pts**)
2) On introduit alors la fonctionnelle **f** définie sur $\mathbf{R}^{\mathbf{n}}$: **f**(x) = $\frac{1}{2}\|\mathrm{Bx}\|_n^2 - \langle \mathrm{c} , \mathrm{x}\rangle_n + \mathrm{d}$ où $\|\ .\ \|_n$ désigne la norme euclidienne dans $\mathbf{R}^{\mathbf{n}}$, **c** est un vecteur de $\mathbf{R}^{\mathbf{n}}$ et d un réel.
   a) Montrer que **f** peut s'écrire sous la forme : **f**(x) = $\frac{1}{2}\langle \mathrm{B}^{\mathrm{T}}\mathrm{B}\mathrm{x} , \mathrm{x}\rangle_n - \langle \mathrm{c} , \mathrm{x}\rangle_n + \mathrm{d}$. (**1pt**)
   b) On pose A = $\mathrm{B}^{\mathrm{T}}\mathrm{B}$. Montrer alors que **f** est une fonctionnelle quadratique et coercive. (**2pts**)
   c) Montrer que A est semi-définie positive et déduire que **f** est convexe. (**1pt**)
   Rappel : Dans $\mathbf{R}^{\mathbf{n}}$, pour une fonctionnelle quadratique **f** (voir question 2)-a)), la dérivée directionnelle d'ordre 1 de **f** s'écrit comme suit : ∇ **f**(x) = grad **f**(x) = Ax – **c** et Hess**f**(x) = A.

   d) Montrer que si B est inversible alors A est aussi inversible. (**1pt**)
   Rappel : B inversible $\Rightarrow \mathrm{B}^{\mathrm{T}}$ inversible et son inverse est tout simplement la transposée de B^{-1} .

   e) Montrer alors que, dans ce cas (A inv.), A est aussi définie positive. (**1pt**)

   f) Dans la suite on suppose A définie positive (B inversible), montrer alors que :
   f1) le problème d'optimisation sans contraintes ($\mathbf{P}_{\mathbf{n}}$) : $\min\limits_{\mathrm{x}\in\mathrm{R}^{\mathrm{n}}} \mathrm{f}(\mathrm{x})$ admet une solution unique (sans la déterminer) notée x*. (**1pt**)

   f2) ''résoudre le problème ($\mathbf{P}_{\mathbf{n}}$)'' revient à ''résoudre le système linéaire Bx = **c'** '' où **c'** est un vecteur de $\mathbf{R}^{\mathbf{n}}$ qu'il faudra préciser (ici il faut, tout simplement, montrer l'équivalence suivante : x* solution optimale de ($\mathbf{P}_{\mathbf{n}}$) $\Leftrightarrow$ x* vecteur-solution du système linéaire Bx = **c'**). (**1pt**)

   f3) résoudre alors le problème ($\mathbf{P}_{\mathbf{n}}$) 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. (**1pt**)

Rappels utiles

Rappel1 : E étant un R-esp. de Hilbert, U ouvert de E et C convexe inclus dans U, alors si $f : E \to \mathbf{R}$ est une fois dérivable (au sens de la dérivée directionnelle) dans U, on a les résultats suivants :

- **f est convexe sur C** $\Leftrightarrow \forall x, y \in C \ \langle f'(x) - f'(y), x - y \rangle_E \geq 0$.
  $\Leftrightarrow \forall x, y \in C \ f(y) \geq f(x) + \langle f'(x), y - x \rangle_E$

- **f est strictement convexe** $\Leftrightarrow \forall x, y \in C$ avec $x \neq y \ \langle f'(x) - f'(y), x - y \rangle_E > 0$.
  $\Leftrightarrow \forall x, y \in C$ avec $x \neq y \ f(y) > f(x) + \langle f'(x), y - x \rangle_E$

Rappel2 : E étant un R-esp. de Hilbert, U ouvert de E et C convexe inclus dans U, alors si $f : E \to \mathbf{R}$ est deux fois dérivable (au sens de la dérivée directionnelle) dans U, on a les résultats suivants :

- **f est convexe sur C** $\Leftrightarrow \forall x, y \in C \ f''(x)(y - x, y - x) \geq 0$.

- Si $\forall x, y \in C$ avec $x \neq y$ on a $f''(x)(y - x, y - x) > 0$ alors **f est strictement convexe**.

Rappel3 : 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.

Rappel4 : Coercivité et optimalité.
E étant un R-esp. normé de dimension finie, alors si $f : E \to \mathbf{R}$ est **continue** et **coercive** (c'est-à-dire $\lim_{\|x\|_E \to \infty} f(x) = +\infty$), la fonction f admet, dans ce cas, au moins un minimum sur E.
Si de plus f est **strictement convexe** alors ce minimum est unique.

Université Abou Bekr Belkaid
Faculté des Sciences
Département de Mathématiques

Année universitaire 2016/2017
Janvier 2017
Optimisation sans contraintes
L3/S5

Corrigé de l'épreuve finale

Questions de cours

1) E étant $\mathbb{R}$-esp. vect., on montre alors que $a \in -A \Leftrightarrow -a \in A$ où $a \in E \setminus \{0_E\}$ et $A \subset E$, $A \neq \emptyset$

$a \in -A \Rightarrow \exists x \in A / a = -x \Rightarrow x = -a \in A$ (car $x \in A$)

$-a \in A \Rightarrow \exists x \in A / x = -a \Rightarrow a = -x$ et $x \in A \Rightarrow a \in -A$

2) E étant $\mathbb{R}$-esp. de Hilbert, C convexe fermé de E et $g : E \to \mathbb{R}$ convexe sur C et dérivable ds un ouvert contenant C. On définit le cône normal à C en x : $N(C,x) = \{y \in E / \langle y, x\rangle_E \geq \langle y, z\rangle_E \ \forall z \in C\}$
$= \{y \in E / \langle y, z-x\rangle_E \leq 0 \ \forall z \in C\}$

a) $0_E \in N(C,x)$ car $\langle 0_E, x\rangle_E = 0 = \langle 0_E, z\rangle_E \ \forall z \in C$ c.à.d. $\langle 0_E, z-x\rangle_E \leq 0 \ \forall z \in C$

b) $N(C,x) \underset{\text{convexe}}{\subset} E$ puisque $y_1, y_2 \in N(C,x)$ et $\lambda \in [0,1]$ alors on montre que $\lambda y_1 + (1-\lambda) y_2 \in N(C,x)$: $y_1, y_2 \in N(C,x) \Rightarrow \langle y_1, z-x\rangle_E \leq 0 \ \forall z \in C$ ① et $\langle y_2, z-x\rangle_E \leq 0 \ \forall z \in C$ ②

$$\left.\begin{array}{l} ① \Rightarrow \forall \lambda \in [0,1] \ \lambda \langle y_1, z-x\rangle_E \leq 0 \ \forall z \in C \\ ② \Rightarrow \forall \lambda \in [0,1] \ (1-\lambda)\langle y_2, z-x\rangle_E \leq 0 \ \forall z \in C \end{array}\right\} \Rightarrow \langle \lambda y_1 + (1-\lambda) y_2, z-x\rangle_E \leq 0 \ \forall z \in C$$

Donc $\lambda y_1 + (1-\lambda) y_2 \in N(C,x)$. C'est donc un convexe de E.

c) Posant $A = \{y \in E / \langle y, z'\rangle_E \leq 0 \ \forall z' \in C - \{x\}\}$, soit $y_0 \in N(C,x) \Rightarrow \langle y_0, z-x\rangle_E \leq 0 \ \forall z \in C$
on pose $z' = z - x$ si $z \in C \Rightarrow z' \in C - \{x\}$
Ds ce cas $\langle y_0, z-x\rangle_E = \langle y_0, z'\rangle \leq 0 \ \forall z' \in C - \{x\} \Rightarrow y_0 \in A$ Dc $N(C,x) \subset A$
inversement, $y_0 \in A \Rightarrow \langle y_0, z'\rangle_E \leq 0 \ \forall z' \in C - \{x\}$. Or $\forall z' \in C - \{x\} \ \exists z \in C / z' = z - x$
Donc $y_0 \in A \Rightarrow \langle y_0, z-x\rangle_E \leq 0 \ \forall z \in C \Rightarrow y_0 \in N(C,x)$ Dc $A \subset N(C,x)$

d) $x \in C$ et x sol. opt. de (P) : $\alpha = \inf\limits_{y \in C} g(y) \Longrightarrow g$ vérifie l'inéquation d'Euler en x c.à.d. $\forall z \in C \longrightarrow \langle g'(x), z-x\rangle_E \geq 0$
$\Rightarrow \forall z \in C \ \langle -g'(x), z-x\rangle_E \leq 0 \Rightarrow -g'(x) \in N(C,x) \underset{\text{d'après 1)}}{\Longrightarrow} g'(x) \in -N(C,x)$

e) $x \in C$ et t.q. $g'(x) \in -N(C,x)$ c.à.d. $-g'(x) \in N(C,x)$ d'après 1)
$\Rightarrow \langle -g'(x), z-x\rangle_E \leq 0 \ \forall z \in C$ d'après la déf. de $N(C,x)$
$\Rightarrow \langle g'(x), z-x\rangle_E \geq 0 \ \forall z \in C$. Ce n'est autre que l'inéquation d'Euler vérifiée pour g en x
Comme g convexe sur C (convexe fermé de E) alors x est sol. opt. de (P).

f) $C=E$ on montre alors que $N(E,x)=\{0_E\}$:

Soit $y\in N(E,x) \Rightarrow \langle y, z-x\rangle_E \leq 0 \ \forall z\in E$. Par l'absurde, on suppose que $y\neq 0_E \Rightarrow \exists z_y\in E / y=z_y-x\neq 0_E \Leftrightarrow z_y\neq x \ (y\neq 0_E)$

Dans ce cas $\langle y, z-x\rangle_E = \langle z_y-x, z-x\rangle_E \leq 0 \ \forall z\in E$

En particulier lorsque $z=z_y$ $\langle z_y-x, z_y-x\rangle_E = \|z_y-x\|_E^2 \leq 0$ et $\|z_y-x\|_E \geq 0$. On en déduit alors $\|z_y-x\|_E=0$ c.à.d. $z_y=x$ contradictoire avec $z_y\neq x$. Donc nécessairement $y=0_E$.

D'après d) et e) $x\in \mathring{C}$ est sol. opt. de $(P) \Leftrightarrow g'(x)\in -N(E,x)$

$C=E \Rightarrow x\,(\in E)$ est sol. opt. de $(P): \inf_{y\in E} g(y) \Leftrightarrow g'(x)\in -N(E,x)=\{0_E\}$
$\Rightarrow g'(x)=0_E$

(car $-0_E=0_E$)

Exercice

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

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

2) $f: \mathbb{R}^n \longrightarrow \mathbb{R}$
$x \longmapsto f(x) = \frac{1}{2}\|Bx\|_n^2 - \langle c, x\rangle_n + d = \frac{1}{2}\langle Bx, Bx\rangle_n - \langle c, x\rangle_n + d$ $\left[\begin{array}{l} c\in\mathbb{R}^n \\ d\in\mathbb{R}\end{array}\right.$

→ a) D'après 1) $\langle Bx, Bx\rangle_n = \langle B^TBx, x\rangle_n$ d'où le résultat.

→ b) On pose $A=B^TB$. f est quadratique ssi A est symétrique or $A^T=(B^TB)^T = B^T(B^T)^T = B^TB = A \Rightarrow A$ sym. $\Rightarrow f$ quadratique

0.25pt pour quelqu'un qui dit que A sym. mais ne le montre pas

Important : A ce niveau, on ne peut pas montrer que f est coercive si la matrice A est seulement semi-déf. positive. Aussi, la question de la coercivité ne devrait pas être considérée en b) mais à partir de la question e) où l'on montre d'abord que A est définie positive.

Pour cette raison si, en b), quelqu'un essaie de montrer que f est coercive il aura 0.5pt, 1pt pour quelqu'un qui ne considère pas la question et 1.5pt pour quelqu'un qui montre l'impossibilité de la coercivité de f au niveau de la quest. b) (Contre-exple : B matrice nulle qui est sym. et semi-déf. positive)

→ c) A est semi-déf. positive ssi, par définition, $\forall x\in\mathbb{R}^n$ $x^TAx = \langle Ax, x\rangle_n = \langle x, Ax\rangle_n \geq 0$ A étant sym.

On se rappelle alors que $\langle Ax, x\rangle_n = \langle B^TBx, x\rangle_n = \langle Bx, Bx\rangle_n = \|Bx\|_n^2 \geq 0 \ \forall x\in\mathbb{R}^n$ 0.5pt

On en déduit que f est convexe puisque $\mathrm{Hess} f(x) = A$ est semi-définie positive 0.5pt

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

$$A^{-1} = (B^TB)^{-1} = B^{-1}(B^T)^{-1} = B^{-1}(B^{-1})^T$$

→ e) 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.

→ f) A est définie-positive $\Rightarrow f$ strict. convexe

f1) Pour que (P_n) : $\min_{x \in \mathbb{R}^n} f(x)$ admette 1 sol. opt. unique, il suffit que f soit continue (% à la norme euclidienne par ex.), coercive et strictement convexe : f étant somme et composée de fonctions continues elle est de continue puisque $x \in \mathbb{R}^n \mapsto Bx \in \mathbb{R}^n$ est 1 application lin. cont. ($\|Bx\|_n \leq \|B\| \, \|x\|_n$ Dc $x \mapsto Bx$ lipschitz.), $y \in \mathbb{R}^n \mapsto \|y\|_n \in \mathbb{R}^+$ cont. et $t \mapsto t^2$ cont. ; f coercive puisque

$$f(x) = \frac{1}{2}\|x\|_A^2 - \langle c, x\rangle_n + d \geqslant \underbrace{\frac{1}{2}\|x\|_A^2 - \|c\|_n \alpha \|x\|_A + d}_{= \|x\|_A\left(\frac{1}{2}\|x\|_A - \|c\|_n \alpha\right) + d} \longrightarrow +\infty \quad \text{qd } \|x\|_A \to \infty$$

où $\|x\|_A^2 = \langle Ax, x\rangle$

et α est la constante d'équivalence entre $\|\cdot\|_A$ et $\|\cdot\|_n$.

f2) 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

$f'(x^*) = 0_{\mathbb{R}^n} \Leftrightarrow \nabla f(x^*) = 0_{\mathbb{R}^n} \Leftrightarrow Ax^* - c = 0_n \Leftrightarrow Ax^* = c$

$\Leftrightarrow B^T B x^* = c \Leftrightarrow Bx^* = (B^T)^{-1} c = (B^{-1})^T c = c'$

Inversement si x^* sol. de $Bx = c' = (B^{-1})^T c$ c.à.d. si $Bx^* = c'$ alors c'est équivalent d'écrire $Ax^* - c = 0_n \Leftrightarrow \nabla f(x^*) = 0_n$
$\Rightarrow x^*$ sol. opt. de (P_n) car f convexe puisqu'elle est strict. conv.

f3) $Bx^* = (B^{-1})^T c \Leftrightarrow x^* = B^{-1}(B^{-1})^T c$ (0.25 seulement pour $x^* = A^{-1}c = (B^TB)^{-1}c$)

Remarques concernant la quest. f1) : Il est très facile de m.q. l'application lin. $x \mapsto \langle c, x\rangle_n$ de $\mathbb{R}^n$ ds $\mathbb{R}$ est cont. (% à la norme euclid. $\|\cdot\|_n$) puisque $|\langle c, x\rangle_n| \leq \|c\|_n \|x\|_n \Rightarrow \forall x, y \in \mathbb{R}^n \; |\langle c, x\rangle_n - \langle c, y\rangle_n| = |\langle c, x-y\rangle_n| \leq \|c\|_n \|x-y\|_n$ ce qui prouve qu'elle est lipschitzienne et par suite continue.
0.25 pt seulement pour quelqu'un qui rappelle ds f1) les conditions d'existence et d'unicité de la sol. opt. x^* de (P_n) : f continue, coercive et strict. convexe