Optimisation - Partiel en présentiel

Durée 1h. Aucun document n'est autorisé. Les deux exercices sont indépendants.

Exercice 1.

Soient $n, m \in \mathbb{N}^*$. On définit pour tout $x \in \mathbb{R}^n$,

$$f(x) = \frac{1}{2} \|x - y\|_{n,2}^2 + \lambda \|Lx\|_{m,2}^2 ,$$

avec $L : \mathbb{R}^n \to \mathbb{R}^m$ est un opérateur linéaire, $y \in \mathbb{R}^n$ un vecteur fixé, $\lambda > 0$, $\|\cdot\|_{n,2}$ et $\|\cdot\|_{m,2}$ respectivement la norme euclidienne sur $\mathbb{R}^n$ et $\mathbb{R}^m$.

1. (a) Rappeler la définition d'une fonctionnelle quadratique.
   (b) Montrer que f est une fonctionnelle quadratique.
   (c) Que vaut $\nabla f(x)$ pour tout $x \in \mathbb{R}^n$?
   (d) Que vaut $\nabla^2 f(x)$ pour tout $x \in \mathbb{R}^n$?
2. Démontrer que f est strictement convexe.
3. Démontrer que f admet un unique minimum global.
4. Les hypothèses du théorème de convergence de la méthode de gradient à pas optimal sont-elles satisfaites ?

On considère maintenant pour tout $x \in \mathbb{R}^n$,

$$g(x) = \frac{1}{2} \|Hx - z\|_{M,2}^2 + \lambda \|Lx\|_{m,2}^2 ,$$

avec $H : \mathbb{R}^n \to \mathbb{R}^M$ un opérateur linéaire, $M \geq n$ et $z \in \mathbb{R}^M$. On suppose que H est de *rang plein*, c'est-à-dire $\mathrm{rg}(H) = n$.

5. Montrer que g est aussi strictement convexe. (*Indication : on pensera à montrer que $H^T H$ est inversible*).
6. Ecrire une fonction *gradG* qui renvoie le gradient de g et qui prend en argument un vecteur x quelconque. On supposera que H, L, λ et z ont été stockés dans des variables de même nom au préalable et sont donc utilisables comme des variables globales.
8. Les hypothèses du théorème de convergence de la méthode de gradient à pas optimal sont-elles toujours satisfaites ?

Exercice 2.

Soit f une fonction $\mathcal{C}^2$ m-fortement convexe ($m > 0$) sur $\mathbb{R}^n$, c'est-à-dire pour tous $x, h \in \mathbb{R}^n$,

$$m \left\| h \right\|_2^2 \leq \left\langle \nabla^2 f(x) h,\, h \right\rangle,$$

et telle que ∇f soit L-Lipschitzienne ($L > 0$), c'est-à-dire pour tous $x, y \in \mathbb{R}^n$,

$$\left\| \nabla f(x) - \nabla f(y) \right\|_2 \leq L \left\| x - y \right\|_2.$$

Soient $x_0 \in \mathbb{R}^n$ et $\tau > 0$. On considère la suite $(x_k)_{k \in \mathbb{N}}$ définie par récurrence par

$$\forall k \in \mathbb{N}, \quad x_{k+1} = F(x_k),$$

où pour tout $x \in \mathbb{R}^n$, $F(x) = x - \tau \nabla f(x)$. C'est la méthode de descente de gradient à pas fixe. On veut trouver une condition sur $\tau > 0$ de sorte que la suite $(x_k)_{k \in \mathbb{N}}$ converge.

1. En utilisant un développement de Taylor approprié (on indiquera bien lequel), démontrer que pour tous $x, y \in \mathbb{R}$,
$$\langle x - y,\, \nabla f(x) - \nabla f(y) \rangle \geq m \left\| x - y \right\|_2^2.$$

2. En développant la norme, en déduire que pour tous $x, y \in \mathbb{R}^n$,
$$\left\| F(x) - F(y) \right\|_2^2 \leq (1 + \tau^2 L^2 - 2m\tau) \left\| x - y \right\|_2^2.$$

3. En déduire que si $0 < \tau < \frac{2m}{L^2}$, alors il existe $0 \leq c < 1$ tel que pour tout $k \geq 1$,
$$\left\| x_{k+1} - x_k \right\|_2 \leq c \left\| x_k - x_{k-1} \right\|_2.$$

4. En déduire que $(x_k)_{k \in \mathbb{N}}$ converge vers un $x_* \in \mathbb{R}^n$. (*Indication : On cherchera à montrer que c'est une suite de Cauchy*).

5. Que représente ce x_* ? Justifier.

6. Commenter la vitesse de convergence. Justifier.

Exercice 1 : 9 points ⚠ Barème indicatif

1) a) Fonctionnelle quadratique :

0,5 $f : x \in \mathbb{R}^n \longmapsto \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle + c$

avec $A \in S_n(\mathbb{R})$, $b \in \mathbb{R}^n$, $c \in \mathbb{R}$.

b) On a $f(x) = \frac{1}{2}\langle x, x\rangle - \langle y, x\rangle + \lambda\langle L^TLx, x\rangle - \frac{1}{2}\|y\|_2^2$

0,5 $= \frac{1}{2}\langle \underbrace{(I_n + 2\lambda L^TL)}_{=A}x, x\rangle - \langle \underbrace{y}_{=b}, x\rangle + \underbrace{\left(-\frac{1}{2}\|y\|_2^2\right)}_{=c}$

La matrice A est bien symétrique donc f est une fonctionnelle quadratique.

0,5 c) D'après le cours, on a $\nabla f(x) = Ax - b$

0,5 d) D'après le cours, on a $D^2 f(x) = A$

2) Il suffit de montrer que A est définie positive vu 1) d).

Soit $h \in \mathbb{R}^n \setminus \{0\}$. On a :

$\langle Ah, h\rangle = \|h\|_2^2 + 2\lambda\langle L^TLh, h\rangle$

$$= \|h\|_2^2 + 2\lambda \underbrace{\|Lh\|_2^2}_{\geq 0}$$

1

$$\geq \|h\|_2^2 > 0 \text{ i.e. } A \text{ est déf. positive}$$

donc f est strictement convexe (même 1 fortement convexe)

3) Comme A est déf positive, $\nabla f(x) = 0$ admet une unique solution i.e. f a un
1 unique point critique qui est un minimum global puisque f est (strictement) convexe.

4) Il faut vérifier : $\forall x, h \in \mathbb{R}^n$

$$m\|h\|_2^2 \underset{(1)}{\leq} \langle \nabla^2 f(x) h, h \rangle \underset{(2)}{\leq} M\|h\|_2^2 .$$

(1) est vrai avec $m = 1$ (cf 2)).

(2) il suffit de prendre

$$M = 1 + 2\lambda \max_{i \in [\![1,n]\!]} \underbrace{\lambda_i(L^T L)}_{\geq 0},$$

1,5

i-ème valeur propre de $L^T L$

Les hypothèses sont donc bien vérifiées.

5) g est aussi une fonctionnelle quadratique avec cette fois $A = H^T H + 2\lambda L^T L$.
Vérifions que A est déf positive. Soit $h \in \mathbb{R}^n \setminus \{0\}$

alors $\langle Ah, h \rangle \geq \|Hh\|_2^2$

Si $\|Hh\|_2^2 = 0$ alors $Hh = 0$ ie $\dim \operatorname{Ker} H \geq 1$.
Or d'après le théorème du rang $\dim \operatorname{Ker} H = 0$

1,5 car $rg(H) = n$. D'où $\langle Ah, h \rangle > 0$ ie
A déf positive.
Donc g est strictement convexe.

6) On a $\nabla g(x) = (H^T H + 2\lambda L^T L)x - H^T z$

```
def gradG(x):
    A = dot(transpose(H), H) + 2*λ*dot(transpose(L), L)
    b = dot(transpose(H), z)
    return dot(A, x) - b
```

1

7) Les hypothèses sont également satisfaites avec

$$m = \min_{i \in [\![1,n]\!]} \underbrace{\lambda_i(H^T H)}_{>0} > 0$$

1

et $M = \max \lambda_i(H^T H) + 2\lambda \max \lambda_i(L^T L)$

Exercice 2 6 points

1) On a $\forall x, y \in \mathbb{R}^n$, Taylor Maclaurin ($f$ $\mathcal{C}^2$)
$\exists z \in]x,y[$ /
$$\nabla f(x) = \nabla f(y) + \nabla^2 f(z)(x-y)$$

Donc $\langle x-y, \nabla f(x) - \nabla f(y)\rangle = \langle x-y, \nabla^2 f(z)(x-y)\rangle$

1 $\geq m \|x-y\|_2^2$

2) On a $\|F(x) - F(y)\|_2^2 = \|x-y\|_2^2 + \tau^2 \|\nabla f(x) - \nabla f(y)\|_2^2$
$- 2\tau \langle x-y, \nabla f(x) - \nabla f(y)\rangle$

1 $\leq (1 + \tau^2 L^2 - 2m\tau) \|x-y\|_2^2$

par L-Lip de ∇f et 1).

3) Soit $g(\tau) = 1 + \tau^2 L^2 - 2m\tau$. On a
$g(0) = g\left(\frac{2m}{L^2}\right) = 1$ et g minimale en $\frac{m}{L^2}$
et valant $1 - \frac{m^2}{L^2} \geq 0$. Donc pour tout

$0 < \tau < \frac{2m}{\ldots}$ $0 < g(\tau) < 1$.

Soit m tel τ et $c = \sqrt{1+\tau^2 L^2 - 2m\tau}$

alors pour $k \geq 1$:

$$\|x_{k+1} - x_k\|_2 = \|F(x_k) - F(x_{k-1})\|_2$$

1

$$\leq c \|x_k - x_{k-1}\|_2$$

d'après 2)

4) Soit $k \geq 1$ et $n \geq 1$ alors

$$\|x_{k+n} - x_k\|_2 \leq \|x_{k+n} - x_{k+n-1}\|_2 + \dots + \|x_{k+1} - x_k\|_2$$

1,5

$$\leq \left(\sum_{i=1}^{n} c^i\right) \|x_k - x_{k-1}\|_2$$

$$\leq c^{k-1} \left(\sum_{i=1}^{n} c^i\right) \|x_1 - x_0\|_2$$

$$\leq K c^k \quad \text{avec } K = \frac{\|x_1 - x_0\|_2}{1-c}$$

Comme $c^k \xrightarrow[k \to +\infty]{} 0 \quad (0 \leq c < 1)$

on en déduit que $\forall \varepsilon > 0, \exists k_0 \in \mathbb{N}$ tq $\forall k \geq k_0$, $\forall n \in \mathbb{N}$, $\|x_{k+n} - x_k\|_2 \leq \varepsilon$

i.e $(x_k)_{k\in\mathbb{N}}$ est de Cauchy et donc $x_k \xrightarrow[k\to+\infty]{} x_* \in \mathbb{R}^n$.

5) Comme $\forall k$, $x_{k+1} = x_k - \tau \nabla f(x_k)$
par passage à la limite et C^0 de ∇f on a

1 $\nabla f(x^*) = 0$ i.e x^* est l'unique minimum global de f (car f fortement cvx).

6) On a $\|x_{k+r} - x_k\|_2 \leq K c^k$

0,5 $\underset{r\to+\infty}{\Rightarrow}$ $\|x_k - x_*\|_2 \leq K c^k$

d'où $(x_k)_{k\in\mathbb{N}}$ CV linéairement vers x^*.