Université Abou Bekr Belkaid Tlemcen
Département de mathématiques
L3 mathématiques 2021/2022

# Solution de l'examen final

## Optimisation sans contraintes

**Exercice 1 (6 points)**:
On définit la fonction $J : \mathbb{R}^n \to \mathbb{R}$ par

$$J(x) := \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle,$$

où $A \in \mathcal{M}_n(\mathbb{R})$ est une matrice symétrique définie positive et $b \in \mathbb{R}^n$. On considère la méthode du gradient à pas optimal pour la minimisation de $J$.
- Donner l'expression de $x_{k+1}$ obtenu par cette méthode.
- Montrer que le pas $\theta_k$ s'écrit dans ce cas:

$$\theta_k = \frac{||d_k||^2}{\langle Ad_k, d_k\rangle}.$$

- Rappelons que le problème de minimisation de $J$ admet une unique solution $\bar{x} \in \mathbb{R}^n$ caractérisée comme étant l'unique solution de l'équation d'Euler $\nabla J(x) = 0$. Montrer que

$$\langle A^{-1}d_k, d_k\rangle = 2(J(x_k) - J(\bar{x})).$$

**Solution**:

- L'algorithme de la méthode du gradient à pas optimal s'écrit:

$$\begin{cases} x_{k+1} = x_k + \theta_k d_k \\ d_k = -Ax_k + b \\ \theta_k \text{ est tel que } J(x_k + \theta_k d_k) = \min_{\theta\in\mathbb{R}} J(x_k + \theta d_k) \end{cases}$$

(1 point).
- Dans l'algorithme du gradient à pas optimal, deux directions successives sont orthogonales:

$$\langle d_{k+1}, d_k\rangle = 0.$$

En utilisant la formule de $x_{k+1}$ et de $d_k$ on obtient:

$$\begin{aligned} 0 &= \langle -Ax_{k+1} + b, d_k\rangle \\ &= \langle -A(x_k + \theta_k d_k) + b, d_k\rangle \\ &= \langle -Ax_k + b - \theta_k Ad_k, d_k\rangle \\ &= \langle d_k - \theta_k Ad_k, d_k\rangle \\ &= ||d_k||^2 - \theta_k\langle Ad_k, d_k\rangle \end{aligned}$$

on suppose que $d_k \neq 0$ et donc $\langle Ad_k, d_k \rangle \neq 0$ (car $A$ est définie positive). Ainsi, $\theta_k$ s'exprime

$$\theta_k = \frac{||d_k||^2}{\langle Ad_k, d_k \rangle}.$$

(2 points)

- La solution du problème de minimisation de $J$ est $\bar{x} = A^{-1}b$. Un simple calcul montre que

$$\begin{aligned}
\langle A^{-1}d_k, d_k \rangle &= \langle A^{-1}(Ax_k - b), Ax_k - b \rangle \\
&= \langle x_k - A^{-1}b, Ax_k - b \rangle \\
&= \langle Ax_k, x_k \rangle - \langle b, x_k \rangle - \langle A^{-1}b, Ax_k \rangle + \langle A^{-1}b, b \rangle \\
&= 2\left(\tfrac{1}{2}\langle Ax_k, x_k \rangle - \langle b, x_k \rangle + \tfrac{1}{2}\langle A^{-1}b, b \rangle\right) \\
&= 2(J(x_k) - J(\bar{x}))
\end{aligned}$$

car

$$\begin{aligned}
J(\bar{x}) &= \tfrac{1}{2}\langle A\bar{x}, \bar{x} \rangle - \langle b, \bar{x} \rangle \\
&= \tfrac{1}{2}\langle b, \bar{x} \rangle - \langle b, \bar{x} \rangle \\
&= -\tfrac{1}{2}\langle b, \bar{x} \rangle \\
&= -\tfrac{1}{2}\langle b, A^{-1}b \rangle.
\end{aligned}$$

(3 points)

**Exercice 2 (5 points)**:

On considère le problème $\min_{\mathbb{R}^2} f(x)$ où $f(x) = \frac{1}{2}\langle Ax, x \rangle$ avec

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

- Donner la solution de ce problème (qu'on note $\bar{x}$).

Soit $\alpha > 0$ et soit la suite $(x_k)_{k \in \mathbb{N}}$ définie par $x_{k+1} = x_k - \alpha A x_k$.

- Pour quelles valeurs de $\alpha$ la suite $(x_k)$ converge vers $\bar{x}$ pour tout $x_0 \in \mathbb{R}^2$?
- Quel est le pas optimal $\bar{\alpha}$?

**Solution** :

- Les valeurs propres de la matrice $A$ sont $\lambda_1 = 1$ et $\lambda_2 = 3$. La forme quadratique $f$ est strictement convexe et coercive, elle admet alors un unique minimum qui est solution de

$$\nabla f(x) = 0 \Leftrightarrow Ax = 0.$$

Comme $A$ est inversible, on obtient $\bar{x} = (0, 0)^T$. (2 points)

- Il s'agit de la suite engendrée par la méthode du gradient à pas fixe, elle

converge pour toute valeur de $\alpha \in ]0, \frac{2}{3}[$. (1.5 point)
- Le pas optimal de cette méthode est donné par:

$$\bar{\alpha} = \frac{2}{\lambda_1 + \lambda_2} = \frac{1}{2} \quad \text{(1.5 point)}$$

**Exercice 3 (9 points)**:
Déterminer et préciser la nature des points critiques des fonctions suivantes:
$f(x, y) = x^3 + y^3 - 3axy, \; a \in \mathbb{R}$
$g(x, y) = x^2 - \cos(y)$
$h(x, y) = y^2 + xy\ln(x)$.

**Solution:**
1) le gradient le hessien de $f$ sont donnés par:

$$\nabla f(x, y) := \begin{pmatrix} 3x^2 - 3ay \\ 3y^2 - 3ax \end{pmatrix} \quad \nabla^2 f(x, y) = \begin{pmatrix} 6x & -3a \\ -3a & 6y \end{pmatrix}$$

Soit $a \neq 0$, la fonction $f$ admet deux points critiques $p_0 = (0, 0)$ et $p_a = (a, a)$. La matrice hessienne de $f$ au point $p_0$ admet deux valeurs propres de signe différent. Le point $p_0$ est donc un point selle. La hessienne calculée au point $p_a$ admet deux valeurs propres $\lambda_1 = 3a$ et $\lambda_2 = 9a$. On déduit que si $a > 0$, $p_a$ est un minimum et si $a < 0$, $p_a$ est un maximum (3 points).
2) le gradient et le hessien de $g$ sont donnés par:

$$\nabla g(x, y) := \begin{pmatrix} 2x \\ \sin(y) \end{pmatrix} \quad \nabla^2 g(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & \cos(y) \end{pmatrix}$$

Les points critiques de $g$ sont de la forme $p_k = (0, k\pi)$, $k \in \mathbb{Z}$. Le calcul de $\nabla^2 g$ au point $p_k$ donne

$$\nabla^2 g(p_k) = \begin{pmatrix} 2 & 0 \\ 0 & (-1)^k \end{pmatrix}$$

donc $p_k$ est un minimum si $k$ est pair et c'est un point selle si $k$ est impair (3 points).
3) la fonction $h$ est définie pour tout $x > 0$ et elle est de classe $C^2$ sur son domaine de définition. le gradient le hessien de $h$ sont donnés par:

$$\nabla h(x, y) := \begin{pmatrix} y\ln(x) + y \\ 2y + x\ln(x) \end{pmatrix} \quad \nabla^2 h(x, y) = \begin{pmatrix} \frac{y}{x} & \ln(x) + 1 \\ \ln(x) + 1 & 2 \end{pmatrix}$$

Il existe deux points critiques $(1, 0)$ et $(\frac{1}{e}, \frac{1}{2e})$. Le premier est un point selle et le deuxième est un minimum (3 points).