3 Conditions d'optimalité. Optimisation sans contrainte

Conformément au programme de l'Agrégation, on se limite au cas de la dimension finie, laissant en remarques des prolongements pour le cas de la dimension infinie.

Théorème 3.0.1 (Inéquation d'Euler)**.** *Soit $K \subset V$ un sous-ensemble convexe d'un espace de Hilbert V. Soit $f : K \to \mathbb{R}$ différentiable sur K. Si x est un minimum local de f sur K alors x est solution de l'inéquation d'Euler :*

$$\forall y \in K, \quad \langle \nabla f(x), y - x \rangle \geq 0.$$

Si de plus f est convexe alors x est un minimum global de f sur K.

Démonstration. Soit $y \in K$. Par hypothèse sur f, l'application

$$\varphi : [0,1] \to \mathbb{R}, \quad t \mapsto f((1-t)x + ty)$$

est dérivable sur $[0,1]$, de dérivée donnée par :

$$\forall t \in [0,1], \quad \varphi'(t) = \langle \nabla f((1-t)x + ty), y - x \rangle.$$

De plus φ admet un minimum local en $t = 0$, i.e. il existe $r \in]0,1[$ t.q. :

$$\forall t \in]0,r[, \quad \varphi(t) \geq \varphi(0) \Rightarrow \frac{1}{t}(\varphi(t) - \varphi(0)) \geq 0.$$

On en déduit :

$$\lim_{t \to 0} \frac{1}{t}(\varphi(t) - \varphi(0)) = \varphi'(0) \geq 0$$

i.e. $\langle \nabla f(x), y - x \rangle \geq 0$.

On suppose que de plus f est convexe. Alors, d'après le Théorème 2.2.3 :

$$\forall y \in K, \quad f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle \geq f(x)$$

i.e. x est un minimum global. □

3.1 Conditions d'optimalité. Optimisation sans contrainte

Sot $f : \mathbb{R}^n \to \mathbb{R}$. On considère le problème :

$$\min_{x \in \mathbb{R}^n} f(x). \tag{8}$$

On se propose d'étudier la généralisation au cas de la dimension $n \geq 2$ de la condition suffisante classique en dimension 1.

Théorème 3.1.1 (Conditions nécessaires). *Soit $x^* \in \mathbb{R}^n$ solution du problème (8).*

1. *Si f est différentiable en x^*, alors $\nabla f(x^*) = 0$ et x^* est appelé point stationnaire ou critique.*
2. *Si f est deux fois différentiable en x^*, alors la matrice $\nabla^2 f(x^*)$ est semi-définie positive.*

Démonstration. 1. On suppose f différentiable en x^* solution de (8). Soit $x \in \mathbb{R}^n$. Alors l'application

$$\varphi : \mathbb{R} \to \mathbb{R}, \quad t \mapsto f((1-t)x^* + tx) \tag{9}$$

admet un minimum en $t = 0$ et est dérivable en $t = 0$, de dérivée définie par : $\varphi'(0) = \langle \nabla f(x^*), x - x^* \rangle$. On a :

$$\forall t \in]0, \infty[, \quad \frac{\varphi(t) - \varphi(0)}{t} \geq 0$$

donc

$$\varphi'(0) = \lim_{t \to 0^+} \frac{\varphi(t) - \varphi(0)}{t} \geq 0.$$

On a de même : On a :

$$\forall t \in]-\infty, 0[, \quad \frac{\varphi(t) - \varphi(0)}{t} \leq 0$$

donc

$$\varphi'(0) = \lim_{t \to 0^-} \frac{\varphi(t) - \varphi(0)}{t} \leq 0,$$

ce qui entraîne que $\varphi'(0) = 0$, i.e. $\langle \nabla f(x^*), x - x^* \rangle = 0$. Ceci étant vrai $\forall x \in \mathbb{R}^n$, il en résulte que $\nabla f(x^*) = 0$.

2. On suppose f deux fois différentiable en x^*. Soit $x \in \mathbb{R}^n$. Alors φ définie par (9) admet une dérivée seconde en $t = 0$ définie par :

$$\varphi''(0) = \langle \nabla^2 f(x^*)(x - x^*), x - x^* \rangle.$$

De plus :

$$\varphi(t) = \varphi(0) + t\varphi'(0) + \frac{t^2}{2}\varphi''(0) + o(t^2)$$

$$\underset{\varphi'(0)=0}{=} \varphi(0) + \frac{t^2}{2}\varphi''(0) + o(t^2) \geq \varphi(0)$$

Si $\varphi''(0) \neq 0$, alors $\varphi(t) \sim \varphi(0) + \frac{t^2}{2}\varphi''(0)$ et $\varphi''(0) + o(1) > 0$, donc $\varphi''(0) > 0$. On en déduit que $\varphi''(0) \geq 0$ en général, i.e.

$$\langle \nabla^2 f(x^*)(x - x^*), x - x^* \rangle \geq 0.$$

Ceci étant vrai $\forall x \in \mathbb{R}^n$, il en résulte que $\nabla^2 f(x^*)$ est semi-définie positive.

□

Remarque 8.

1. Si f est deux fois différentiable en x^*, la hessienne en x^* n'est pas en général définie positive. Par exemple, si $f(x) = x^4$, alors f atteint son minimum absolu sur $\mathbb{R}$ en $x = 0$ et $f''(0) = 0$.
2. Le Théorème 3.1.1 fournit des conditions nécessaires mais non suffisantes. En effet, si $f(x) = x^3$, alors $f'(0) = f''(0) = 0$ et f n'a pas de minimum sur $\mathbb{R}$.

Théorème 3.1.2 (Conditions suffisantes)**.** *Soit $f : \mathbb{R}^n \to \mathbb{R}$ deux fois différentiable en x^* t.q. $\nabla f(x^*) = 0$. On suppose en outre que :*
(i) $\nabla^2 f(x^*)$ est définie positive ;
(ii) f est deux fois différentiable de hessienne semi-définie positive dans un voisinage de x^*.
Alors x^* est un minimum local pour f.

Démonstration. De (ii), on déduit qu'il existe $r > 0$ t.q. f soit deux fois différentiable de hessienne semi-définie positive dans la boule ouverte $B(x^*, r)$ de rayon $r > 0$ centrée en x^*. Soit $x \in \mathbb{R}^n$, $x \neq x^*$, et soit

$$\varphi : \mathbb{R} \to \mathbb{R}, \quad \varphi(t) = f((1-t)x^* + tx).$$

Alors φ est deux fois dérivable sur I_x centré en 0 :

$$\|(1-t)x^* + tx - x^*\| < r$$

$I_x := \left]-\frac{r}{\|x-x^*\|}, \frac{r}{\|x-x^*\|}\right[$ de dérivée donnée par :

$$\varphi''(t) = \langle \nabla^2 f((1-t)x^* + tx)(x - x^*), x - x^* \rangle \geq 0, \quad \forall t \in I_x.$$

En particulier, par hypothèse sur f :

$$\varphi'(0) = \langle \nabla f(x^*), x - x^* \rangle = 0$$

$$\varphi''(0) = \langle \nabla^2 f(x^*)(x - x^*), x - x^* \rangle > 0$$

On en déduit, quitte à restreindre $r > 0$, que : $\forall t \in I_x$,

$$\varphi(t) = \varphi(0) + \frac{t^2}{2}\varphi''(0) + o(t^2) \geq \varphi(0),$$

i.e. :

$$f((1-t)x^* + tx) \geq f(x^*), \quad \forall t \in I_x.$$

On en déduit, ceci étant vrai pour tout $x \in \mathbb{R}^n \setminus \{x^*\}$, que $f \geq f(x^*)$ sur $B(x^*, r)$, i.e. que x^* est un minimum local de f. □

On obtient un résultat optimal dans le cas convexe.

Théorème 3.1.3 (Cas convexe. Condition nécessaire et suffisante.)**.** *Soit $f : \mathbb{R}^n \to \mathbb{R}$ convexe et différentiable sur $\mathbb{R}^n$. Une CNS pour que $x^* \in \mathbb{R}^n$ soit un minimum local (et donc global) de f est que x^* soit un point critique de f i.e. vérifie :*

$$\nabla f(x^*) = 0.$$

Démonstration. $\Rightarrow$ Soit $x^* \in \mathbb{R}^n$ un point critique de f sur $\mathbb{R}^n$. D'après la caractérisation des fonctions convexes du Théorème 2.2.3 :

$$\forall x \in \mathbb{R}^n, \quad f(x) \geq f(x^*) + \langle \nabla f(x^*), x - x^* \rangle = f(x^*)$$

i.e., x^* est un minimum global de f sur $\mathbb{R}^n$.

$\Leftarrow$ Inversement on suppose que f admet un minimum local en x^*. Alors $\nabla f(x^*) = 0$ d'après le Théorème 3.1.1. □

Dans la suite, on étudie en détails deux problèmes fondamentaux en mathématiques appliquées : la minimisation d'une fonctionnelle quadratique et la méthode des moindres carrés.

3.2 Minimisation d'une fonctionnelle quadratique. Optimisation sans contrainte

Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice carrée réelle symétrique, soit $b \in \mathbb{R}^n$ et soit $c \in \mathbb{R}$. On définit la fonctionnelle quadratique (somme d'une forme quadratique et d'une fonction affine)

$$f : \mathbb{R}^n \to R, \quad x \mapsto \frac{1}{2}\langle Ax, x \rangle - \langle b, x \rangle + c$$

et on considère le problème de minimisation :

$$\min_{x\in\mathbb{R}^n} f(x).$$

Dans l'Exemple 10, on a montré que f est deux fois différentiable sur $\mathbb{R}^n$, de différentielle définie par le gradient : $\nabla f(x) = Ax - b$, de matrice hessienne constante $\nabla^2 f(x) = A$. En particulier, f est convexe ssi A est semi-définie positive.

On en déduit que si A est semi-définie positive, alors f admet un minimum local (et donc global) s'il existe $x^* \in \mathbb{R}^n$ solution de $Ax^* = b$, ce qui est réalisé ssi $b \in \mathrm{Im}(A)$ avec, en dimension finie : $\mathrm{Im}(A) = \mathrm{Ker}(A^T)^\perp = \mathrm{Ker}(A)^\perp$.

Dans le cas général où A est symétrique réelle, A et digonalisable dans une bon, i.e. il existe $U \in \mathcal{O}_n(\mathbb{R})$ orthogonale t.q. $A = U^T DU$ avec D diagonale formée des valeurs propres de A, soit :

$$\lambda_1 \leq \cdots \leq \lambda_n.$$

On distingue plusieurs cas suivant le signe de la plus petite valeur propre λ_1.

(i) Si $\lambda_1 < 0$, soit $u_1 \in \mathbb{R}^n$ t.q. $Au_1 = \lambda_1 u_1$ et $\|u_1\| = 1$. On a : $\forall t \in \mathbb{R}$,

$$f(tu_1) = \frac{t^2}{2}\lambda_1 - t\langle b, u_1\rangle + c = \frac{\lambda_1}{2}\left(t - \frac{\langle b, u_1\rangle}{\lambda_1}\right)^2 + \frac{\langle b, u_1\rangle^2}{2\lambda_1} + c$$

avec $\lim_{t\to+\infty}\left|t - \frac{\langle b,u_1\rangle}{\lambda_1}\right| = +\infty \Rightarrow$

$$\lim_{t\to+\infty} f(tu_1) = -\infty$$

donc $\inf_{\mathbb{R}^n} f = -\infty$ et le problème (8) n'a pas de solution.

(ii) Si $\lambda_1 = 0$ et $b \notin \mathrm{Ker}(A)^\perp$, alors il existe $u \in \mathrm{Ker}(A)$ t.q. $\langle b, u\rangle \neq 0$. On a : $\forall t \in \mathbb{R}$,

$$f(tu) = -t\langle b, u\rangle + c.$$

On peut supposer, quitte à remplacer u par $-u$, que $\langle b, u\rangle > 0$. Alors

$$\lim_{t\to+\infty} f(tu) = -\infty \Rightarrow \inf_{x\in\mathbb{R}^n} f(x) = -\infty$$

et le problème (8) n'a pas de solution. Autrement dit, A étant semi-définie positive, f est convexe non minorée sur $\mathbb{R}^n$. En particulier, l'équation $Ax = b$ n'a pas de solution i.e. f n'a pas de point critique, car $b \notin (\mathrm{Ker}(A)^\perp$.

(iii) Si $\lambda_1 = 0$ et $b \in (\mathrm{Ker}(A))^\perp$, alors l'équation $Ax = b$ admet une infinité de solutions. Plus précisément, si $x^* \in \mathbb{R}^n$ vérifie $Ax^* = b$, alors

$$Ax = b \iff A(x - x^*) = 0 \iff x \in x^* + \mathrm{Ker}(A).$$

Soit alors $x \in \mathbb{R}^n$ et soit $x = x' + x''$ la décomposition de x dans la somme directe $\mathbb{R}^n = \mathrm{Ker}(A) \oplus \mathrm{Ker}(A)^\perp$. On a

$$\langle Ax, x\rangle = \langle Ax'', x\rangle = \langle x'' Ax\rangle = \langle Ax'', x''\rangle$$

et

$$b \in \mathrm{Ker}(A)^\perp \Rightarrow \langle b, x\rangle = \langle b, x''\rangle.$$

Donc $f(x) = f(x'')$ et on en déduit que $\inf_{x\in\mathbb{R}^n} f(x) = \inf_{x\in(\mathrm{Ker}(A))^\perp} f(x)$. Or $\lambda_1 = 0 \Rightarrow \nabla^2 f$ est semi-définie positive, donc f est convexe et tout minimum local est un minimum global. De plus, l'équation $\nabla f(x) = 0 \iff Ax = b$ admet une unique solution x^* dans $(\mathrm{Ker}(A))^\perp$. On en déduit que les solutions du problème (8) sont les vecteurs de la forme $x^* + x$, $x \in \mathrm{Ker}(A)$ où x^* est l'unique solution de

$$Ax^* = b \quad \text{et} \quad x^* \in (\mathrm{Ker}(A))^\perp.$$

Le calcul montre directement que :

$$\inf_{x\in\mathbb{R}^n} f(x) = f(x^*) = -\frac{1}{2}\langle b, x^*\rangle + c.$$

(iv) Si $\lambda_1 > 0$, alors $\mathrm{Im}(A) = \mathbb{R}^n$ et l'équation $Ax = b$ admt une unique solution x^*. De plus A est définie positive donc f est (strictement) convexe et $x^* = A^{-1}b$ est l'unique solution du problème (8). On a

$$\inf_{xin\mathbb{R}n} f(x) = f(x^*) = -\frac{1}{2}\langle b, A^{-1}b\rangle + c.$$

3.3 La méthode des moindres carrés

On pourra se reporter à [7] pour des compléments sur cette méthode.

Soit $A \in \mathcal{M}_{m,n}(\mathbb{R})$ une matrice de taille $m \times n$ (m lignes et n colonnes) avec $m > n$. Soit $b \in \mathbb{R}^m$. On considère le problème :

$$\min_{x\in\mathbb{R}^n} \|Ax - b\|^2.$$

Le problème à n inconnues et m équations $Ax = b$ est mal posé en général. Il revient à minimiser la distance entre b et $\mathrm{Im}(A) = \mathrm{Ker}(A^T)^\perp$, i.e., puisque

$\mathrm{Ker}(A^T)^\perp$ est fermé en dimension finie, à chercher le projeté de b sur $\mathrm{Ker}(A^T)^\perp$.

Soit l'application :

$$f : \mathbb{R}^n \to \mathbb{R}, \quad x \mapsto f(x) = \frac{1}{2}\|Ax - b\|^2.$$

Un calcul direct montre que

$$\begin{aligned}\forall x \in \mathbb{R}^n, \quad f(x) &= \frac{1}{2}\|Ax\|^2 - \langle Ax, b\rangle + \frac{1}{2}\|b\|^2 \\ &= \frac{1}{2}\langle A^T Ax, x\rangle - \langle A^T b, x\rangle + \frac{1}{2}\|b\|^2\end{aligned}$$

et on est ainsi ramené au problème de la Section 3.2 avec $A^T A$ symétrique semi-définie positive ($\lambda_1 \geq 0$ avec les notations de la Section 3.2.)

Soit $A^T Ax = 0$. Alors :

$$\langle A^T Ax, x\rangle = \|Ax\|^2 = 0 \Rightarrow Ax = 0$$

donc $\mathrm{Ker}(A^T A) \subset \mathrm{Ker}(A)$. Comme il est immédiat que $\mathrm{Ker}(A) \subset \mathrm{Ker}(A^T A)$, on en déduit que $\mathrm{Ker}(A^T A) = \mathrm{Ker}(A)$. De plus :

$$A^T b \in \mathrm{Im}(A^T) = \mathrm{Ker}(A)^\perp = \mathrm{Ker}(A^T A)^\perp$$

D'après le Théorème du rang en dimension finie : $n = \dim(\mathrm{Ker}(A)) + \dim(\mathrm{Im}(A))$. Si $\dim(\mathrm{Im}(A)) < n$, alors $\mathrm{Ker}(A^T A) = \mathrm{Ker}(A) \neq \{0\}$ et l'ensemble des solutions du problème (8) est $x^* + \mathrm{Ker}(A)$, où x^* est l'unique solution de

$$A^T Ax^* = A^T b, \quad x^* \in \mathrm{Ker}(A)^T.$$

Si $\dim(\mathrm{Im}(A)) = n$, alors $\mathrm{Ker}(A^T A) = \mathrm{Ker}(A) = \{0\}$, i.e. $A^T A$ est carrée d'ordre n, injective donc inversible et le problème (8) admet pour unique solution la solution de $A^T Ax^* = A^T b$.

Remarque 9 (Pseudo-inverse). Si A est injective alors $A^T A$ est inversible et on peut définir la pseudo-inverse ou inverse généralisée de A en posant $A^\dagger := (A^T A)^{-1} A^T$.

Avec cette définition :

$$A^T Ax^* = A^T b \iff x^* = A^\dagger b.$$

On vérifie immédiatement que $A^\dagger$ est l'inverse à gauche de A.

Dans le cas général où A n'est pas nécessairement injective, $A^T A$ est symétrique donc se décompose sous la forme $A^T A = U^T \Sigma^2 U$ où $U \in \mathcal{O}_n(\mathbb{R})$

est orthogonale d'ordre n, Σ est diagonale, de valeurs propres $\sigma_i \geq 0$. On en déduit ([8], 1.2.7) : $(AU^T)^T AU^T = \Sigma^2$, i.e., en notant $c_j \in \mathbb{R}^m$ le $j$ème vecteur colonne de AU^T, $j \in [[1,n]]$, $c_i^T c_j = \sigma_j^2 \delta_{ij}$, $\forall i, j \in [[1,n]]$. Alors : $\|c_j\| = \sigma_j$, $\forall j \in [[1,n]]$. On pose

$$v_j = \sigma_j^{-1} c_j \quad \text{si} \quad \sigma_j \neq 0,$$

et on complète la famille (v_j) en une base orthonormée encore notée (v_j) de $\mathbb{R}^m$. Soit $V \in \mathcal{O}_m(\mathbb{R})$ la matrice orthogonale de vecteurs colonnes les v_j, $j \in [[1,m]]$ et soit $\Sigma' := \begin{pmatrix} \Sigma \\ 0 \end{pmatrix}$ de taille $m \times n$. On a : $\forall i \in [[1,m]]$, $\forall j \in [[1,n]]$,

$$(V\Sigma')_{ij} = \sum_{k=1}^{m} V_{ik}\Sigma'_{kj} = \sum_{k=1}^{n} V_{ik}\Sigma_{kj} = V_{ij}\sigma_j = \sigma_j (v_j)_i = (\sigma_j v_j)_i = (c_j)_i$$

i.e. : $V\Sigma' = AU^T$. Les éléments diagonaux de Σ sont appelés les valeurs propres singulières de A et $A = V\Sigma' U$ est la décomposition en valeurs propres singulières de A. Si A est injective, alors Σ est inversible et on a

$$A^\dagger = (U^T \Sigma^2 U)^{-1} U^T \Sigma'^T V^T = U^T \Sigma^{-2} U U^T \Sigma'^T V^T = U^T \begin{pmatrix} \Sigma^{-1} & 0 \end{pmatrix} V^T$$

Soit $A = QR$ la décomposition QR de A avec

$$R = \begin{pmatrix} R' \\ 0 \end{pmatrix} \quad \text{et } R' \text{ inversible si } A \text{ est injective}$$

On a $A^T A = R^T R = R'^T R' \Rightarrow$:

$$A^\dagger = (R^T R)^{-1} (QR)^T = R'^{-1} R'^{-T} R^T Q^T = \begin{pmatrix} R'^{-1} & 0 \end{pmatrix} Q^T$$

On vérifie directement que l'opération de pseudo-inversion est involutive et commute avec la transposition et la conjugaison.

Exemple 13 (Régression linéaire)**.** On pourra se reporter à [9], Chapitre 2, pour la régression linéaire sans contraintes et à [9], Chapitre 3, pour la régression linéaire avec contraintes.

On considère u nuage de points $(M_i)_{i \in [[1,m]]}$, $M_i = (t_i, x_i)$, $\forall i \in [[1,m]]$, résutant en pratique de mesures. Conformément à une modélisation théorique préalable, on peut s'attendre à ce qu'ils se concentrent autour d'une droite restant à déterminer. On est alors ramené au problème de minimisation :

$$\min_{(\alpha,\beta) \in \mathbb{R}^2} \underbrace{\sum_{i=1}^{m} |x_i - \alpha t_i - \beta|^2}_{f(\alpha,\beta)},$$

où $f(\alpha,\beta)$ est la somme des distances des points $(M_i)_{i\in[[1,m]]}$ à la droite du plan $x = \alpha t + \beta$. En posant :

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad A = \begin{pmatrix} t_1 & 1 \\ \vdots & \vdots \\ t_n & 1 \end{pmatrix}$$

on réécrit le problème sous la forme :

$$\min_{(\alpha,\beta)\in\mathbb{R}^2} \left\| x - A \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \right\|^2 \iff \min_{(\alpha,\beta)\in\mathbb{R}^2} \left\| A \begin{pmatrix} \alpha \\ \beta \end{pmatrix} - x \right\|^2$$

où $\|\cdot\|$ désgne la norme de $\mathbb{R}^m$. On est donc ramené à résoudre un problème de moindres carrés de matrice $A = (\ t \quad u\ )$ où u désigne le vecteur de $\mathbb{R}^m$ de composantes $u_i = 1$, $i \in [[1,m]]$. De l'égalité :

$$A^T A = \left(\begin{matrix} t^T \\ u^T \end{matrix} \right)(\ t \quad u\ ) = \begin{pmatrix} \|t\|^2 & t^T u \\ t^T u & \|u\|^2 \end{pmatrix}$$

il résulte que

$$\det(A^T A) = \|t\|^2\|u\|^2 - (t^T u)^2 \geq 0$$

avec

$$\det(A^T A) = 0 \iff \|t\|\|u\| = |t^T u|$$

ce qui est réalisé ssi les vecteurs t et u sont colinéaires. On en déduit : si $t_1 = \cdots = t_n$, i.e. alors les points M_i sont alignés sur la droite $t = t_1$. Sinon, $\det(A^T A) > 0$, alors $\min_{(\alpha,\beta)} f(\alpha,\beta) = f(\alpha^*,\beta^*)$ avec (α^*,β^*) unique solution de

$$A^T A \begin{pmatrix} \alpha^* \\ \beta^* \end{pmatrix} = A^T x$$

donné par les formules :

$$\alpha^* = \frac{mt^T x - (t^T u)(t^T x)}{m\|t\|^2 - (t^T u)^2}, \quad \beta^* = \frac{\|t\|^2(u^T x) - (t^T u)(t^T x)}{m\|t\|^2 - (t^T u)^2}.$$

4 Conditions d'optimalité. Optimisation avec contraintes

Bibliographie

[1] J.P. Demailly, Analyse Numérique et Equations Différentielles, P.U.Grenoble, Paris, 2006.

[2] H. Brezis, Analyse fonctionnelle : théorie et applications, Masson, Paris, 1983.

[3] P. Ciarlet, Introduction à l'analyse numérique matricielle et à l'optimisaion, Masson, Dunod, Paris.

[4] P. Lascaux, R. Théodor, Analyse numérique matricielle appliquée à l'art de l'ingénieur, Dunod, Paris.

[5] J.-B. Hiriart-Urruty, Convex Analysis and Minimization Algorithms I, Springer-Verlag, 1996.

[6] C. Zuily, H. Queffélec, Analyse pour l'Agrégation, 3ème Edition, Dunod, 2007.

[7] G. Allaire, S.M. Kaber, Numerical Linear Algebra, Texts in Applied Mathematics, vol. 55, Springer, 2008.

[8] P. Lascaux, R. Théodor, Analyse numérique matricielle appliquée à l'art de l'ingénieur. Tome 1. Dunod, Paris, 1986.

[9] M. Bergounioux, Optimisation et contrôle des systèmes linéaires, Dunod, 2001.