Leçon 158 : Matrices symétriques réelles, matrices hermitiennes.

1 Définitions et premières propriétés

Définition 1. (matrice symétrique, matrice hermitienne) On dit que $A \in \mathscr{M}_n(\mathbb{R})$ (resp. $\mathscr{M}_n(\mathbb{C})$) est *symétrique* (resp. *hermitienne*) si ${}^tA = A$ (resp. $A^* := {}^t\overline{A} = A$). On note $\mathscr{S}_n(\mathbb{R})$ (resp. $\mathscr{H}_n(\mathbb{C})$) l'ensemble de ces matrices.

Exemple 2.

$$\begin{bmatrix} 1 & 7 \\ 7 & 2 \end{bmatrix} \in \mathscr{S}_2(\mathbb{R}) \text{ et } \begin{bmatrix} 1 & i \\ -i & 0 \end{bmatrix} \in \mathscr{H}_2(\mathbb{C}).$$

Proposition 3. [Gou2, page 229]

1. $\mathscr{S}_n(\mathbb{R})$ est un $\mathbb{R}$-espace vectoriel de dimension $\dfrac{n(n+1)}{2}$.
2. $\mathscr{H}_n(\mathbb{C})$ est un $\mathbb{R}$-espace vectoriel de dimension n^2.

Proposition 4. [Gou2, page 229]

1. $\mathscr{M}_n(\mathbb{R}) = \mathscr{S}_n(\mathbb{R}) \oplus \mathscr{A}_n(\mathbb{R})$.
2. $\mathscr{H}_n(\mathbb{C}) = \mathscr{S}_n(\mathbb{R}) \oplus i\mathscr{A}_n(\mathbb{R})$

où $\mathscr{A}_n(\mathbb{R}) = \{A \in \mathscr{M}_n(\mathbb{R}),\ {}^tA = -A\}$.

2 Réduction et action par congruence

2.1 Formes quadratiques, formes quadratiques hermitiennes

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie $n \in \mathbb{N}^*$, avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

Définition 5. [Gou2, page 228] (forme bilinéaire symétrique, forme sesquilinéaire hermitienne) Soit $\mathbb{K} = \mathbb{R}$ (resp. $\mathbb{C}$). Une forme $\varphi : E \times E \to \mathbb{K}$ est dite *bilinéaire symétrique* (resp. *sesquilinéaire hermitienne*) si

1. $\forall x, y, z \in E, \forall \lambda \in \mathbb{R}, \varphi(\lambda x + y, z) = \lambda\varphi(x, z) + \varphi(y, z)$.
2. $\forall x, y \in E, \varphi(x, y) = \varphi(y, x)$ (resp. $\varphi(x, y) = \overline{\varphi(y, x)}$).

Dans la suite, on fixe φ une telle forme.

Définition 6. [Gou2, page 229] (écriture matricielle) Si $\mathcal{B} = (e_1, \dots, e_n)$ est une base de E fixée, la matrice de φ dans la base $\mathcal{B}$ est

$$A = Mat_{\mathcal{B}}(\varphi) = (\varphi(e_i, e_j))_{1 \leqslant i,j \leqslant n}.$$

On a $A \in \mathscr{S}_n(\mathbb{R})$ dans le cas réel, $A \in \mathscr{H}_n(\mathbb{C})$ dans le cas complexe. Alors, pour tous $x, y \in E$, si l'on note X, Y les vecteurs colonnes de leurs coordonnées dans $\mathcal{B}$, on a $\varphi(x, y) = {}^tXAY$ si $\mathbb{K} = \mathbb{R}$, $\varphi(x, y) = X^*AY$ si $\mathbb{K} = \mathbb{C}$.

Définition 7. [Gou2, page 229] (forme quadratique, forme quadratique hermitienne) Lorsque $\mathbb{K} = \mathbb{R}$ (resp. $\mathbb{K} = \mathbb{C}$), une *forme quadratique* (resp. *forme quadratique hermitienne*) est une application de la forme $q : x \in E \mapsto \varphi(x, x) \in \mathbb{K}$, où φ est une forme bilinéaire symétrique (resp. sesquilinéaire hermitienne).

Proposition 8. [Gou2, page 229] (formules de polarisation) Si q est une forme quadratique (hermitienne) sur E, il existe une unique forme bilinéaire symétrique (resp. sesquilinéaire hermitienne) φ telle que

$$\forall x \in E, q(x) = \varphi(x, x).$$

φ s'appelle la forme polaire associée à q, et on a, pour tous $x, y \in E$,

1. $\varphi(x, y) = \dfrac{1}{4}(q(x + y) - q(x - y))$ dans le cas réel,

2. $\varphi(x,y) = \frac{1}{4}(q(x+y)+q(x-y)+q(x-iy)-q(x+iy))$ dans le cas complexe.

Exemple 9. $(x_1, x_2) \mapsto 3x_1^2 + 2x_1x_2 - x_2^2$ est une forme quadratique sur $\mathbb{R}^2$, et sa matrice dans la base canonique est

$$\begin{bmatrix} 3 & 1 \\ 1 & -1 \end{bmatrix}.$$

2.2 Réduction de Gauß et signature

Proposition 10. [Gou2, page 229] (changement de base) Deux matrices $A, A' \in \mathscr{S}_n(\mathbb{R})$ (resp. $\mathscr{H}_n(\mathbb{C})$) représentent la même forme φ dans des bases $\mathcal{B}, \mathcal{B}'$ si et seulement si elles sont *congruentes*, *i.e.* il existe $P \in GL_n(\mathbb{R})$ (resp. $GL_n(\mathbb{C})$) telle que $A' = {}^tPAP$ (resp. $A' = P^*AP$). Alors, $P = Mat_{\mathcal{B}}(\mathcal{B}')$ est la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$.

Théorème 11. [Gou2, pages 229 et 230] (réduction de Gauss) Soit $A \in \mathscr{S}_n(\mathbb{R})$ (resp. $\mathscr{H}_n(\mathbb{C})$). Il existe P inversible telle que ${}^tPAP$ (resp. P^*AP) soit de la forme

$$diag(\underbrace{1, \dots, 1}_{s \text{ fois}}, \underbrace{-1, \dots, -1}_{t \text{ fois}}, 0, \dots, 0).$$

Exemple 12. [Gou2] Dans $\mathbb{R}^3$, la forme quadratique

$$q(x,y,z) = x^2 - 2y^2 + xz + yz$$

se réduit sous la forme

$$q(x,y,z) = \left(x + \frac{z}{2}\right)^2 - \left(\sqrt{2}y - \frac{z}{2\sqrt{2}}\right)^2 - \left(\frac{z}{2\sqrt{2}}\right)^2.$$

Théorème 13. [Gou2] (loi d'inertie de Sylvester) Les entiers s et t définis ci-dessus sont entièrement déterminés par q. Le couple (s,t) s'appelle la *signature* de A et $r = s+t$ est le *rang* de A.

Proposition 14. [H2G2] Deux matrices symétriques ou hermitiennes sont congruentes si et seulement si elles ont même signature.

Définition 15. [Gou2, page 234] (matrice positive, définie positive) Une matrice $A \in \mathscr{S}_n(\mathbb{R})$ (resp. $\mathscr{H}_n(\mathbb{C})$) est dite *positive* si pour tout $x \in \mathbb{R}^n$ (resp. $\mathbb{C}^n$), ${}^txAx \geqslant 0$ (resp. $x^*Ax \geqslant 0$). On note alors $A \in \mathscr{S}_n^+(\mathbb{R})$ (resp. $A \in \mathscr{H}_n^+(\mathbb{C})$). Si de plus, l'inégalité est stricte pour tout $x \neq 0$, on dit que A est *définie positive*, et on note $A \in \mathscr{S}_n^{++}(\mathbb{R})$ (resp. $A \in \mathscr{H}_n^{++}(\mathbb{C})$).

Proposition 16. [Gou2, page 234] Une matrice $A \in \mathscr{S}_n(\mathbb{R})$ (resp. $\mathscr{H}_n(\mathbb{C})$) est positive (resp. définie positive) si et seulement si sa signature est de la forme $(s,0)$ (resp. $(n,0)$).

Proposition 17. [Gou2] Une matrice symétrique réelle est (définie) positive si et seulement si ses mineurs principaux $\Delta_k = \det(a_{i,j})_{1 \leqslant i,j \leqslant k}$, $k \in [\![1,n]\!]$ sont (strictement) positifs.

Proposition 18. [Gou2, page 231] Si $A \in \mathscr{S}_n^{++}(\mathbb{R})$ (resp. $\mathscr{H}_n^{++}(\mathbb{C})$), l'application $x \mapsto \sqrt{{}^txAx}$ (resp. $\sqrt{x^*Ax}$) définit une norme euclidienne (resp. hermitienne) sur $\mathbb{K}^n$.

3 Théorème spectral et conséquences

3.1 Endomorphismes autoadjoints

Proposition 19. [Gou2] (adjoint) Soit E un espace euclidien (resp. hermitien) de dimension finie. Pour tout $f \in \mathcal{L}(E)$, il existe un unique $f^* \in \mathcal{L}(E)$, appelé *adjoint* de f, tel que

$$\forall x, y \in E, \quad \langle f(x), y\rangle = \langle x, f^*(y)\rangle.$$

Définition 20. [Gou2] $f \in \mathcal{L}(E)$ est dit *autoadjoint* si $f^* = f$.

Proposition 21. [Gou2] Si E est euclidien (resp. hermitien), $f \in \mathcal{L}(E)$ est autoadjoint si et seulement si sa matrice dans une (ou dans toute) base orthonormée est symétrique (resp. hermitienne).

3.2 Théorème spectral

Définition 22. Une matrice $P \in \mathscr{M}_n(\mathbb{R})$ (resp. $\mathscr{M}_n(\mathbb{C})$ est dite *orthogonale* (resp. *unitaire*) si ${}^tPP = I_n$ (resp. $P^*P = I_n$).

Théorème 23. [Gou2] (théorème spectral) Soit $A \in \mathscr{S}_n(\mathbb{R})$ (resp. $\mathscr{H}_n(\mathbb{C})$). Il existe une matrice P orthogonale (resp. unitaire) telle que $P^{-1}AP$ soit diagonale réelle.

Corollaire 24. Si A est symétrique (resp. hermitienne), ses valeurs propres sont réelles, et on a

1. $A \in \mathscr{S}_n^+(\mathbb{R}) \iff \mathrm{Vp}(A) \subset \mathbb{R}_+$ (resp $A \in \mathscr{H}_n^+(\mathbb{C}) \iff \mathrm{Vp}(A) \subset \mathbb{R}_+$).
2. $A \in \mathscr{S}_n^{++}(\mathbb{R}) \iff \mathrm{Vp}(A) \subset \mathbb{R}_+^*$ (resp. $A \in \mathscr{H}_n^{++}(\mathbb{C}) \iff \mathrm{Vp}(A) \subset \mathbb{R}_+^*$).

Théorème 25. [Gou2, page 241] (orthogonalisation simultanée) Soient $A, B \in \mathscr{S}_n(\mathbb{R})$ (resp. $\mathscr{H}_n(\mathbb{C})$) telle que A soit définie positive. Alors, il existe P inversible et D diagonale réelle telles que ${}^tPAP = I_n$ et ${}^tPBP = D$ (resp. $P^*AP = I_n$ et $P^*BP = D$).

Définition 26. [Ales] (ellipsoïde) Pour $A \in \mathscr{S}_n^{++}(\mathbb{R})$, on appelle *ellipsoïde* associé à A l'ensemble

$$\Sigma_A = \{x \in \mathbb{R}^n, \quad {}^txAx \leqslant 1\}.$$

Proposition 27. Pour toute matrice $A \in \mathscr{S}_n^{++}(\mathbb{R})$, l'ellipsoïde Σ_A est compact, donc de volume $\mathrm{vol}(\Sigma_A)$ fini, et l'application $A \in \mathscr{S}_n^{++}(\mathbb{R}) \mapsto \mathrm{vol}(\Sigma_A)$ est strictement convexe.

Théorème 28. [Ales] (ellipsoïde de John-Loewner) Soit $K \subset \mathbb{R}^n$ un compact d'intérieur non vide. Alors, il existe un unique ellipsoïde contenant K, centré en 0, de volume minimal.

3.3 Théorème de Courant-Fischer et quadriques

Théorème 29. [Cia, page 12] (théorème de Courant-Fischer) Soit $A \in \mathscr{M}_n(\mathbb{R})$ (resp. $A \in \mathscr{M}_n(\mathbb{C})$) une matrice symétrique (resp. hermitienne), et soient $\lambda_1 \leqslant \ldots \leqslant \lambda_n$ ses valeurs propres (réelles), avec d'éventuelles répétitions. Alors, pour tout $1 \leqslant k \leqslant n$, si l'on note G_k l'ensemble des sous-espaces vectoriels de $\mathbb{R}^n$ (resp. $\mathbb{C}^n$) de dimension k,

$$\lambda_k = \min_{W \in G_k} \max_{x \in W \setminus \{0\}} \frac{\langle Ax, x\rangle}{\|x\|_2^2} = \max_{W \in G_{k-1}} \min_{x \perp W, x \neq 0} \frac{\langle Ax, x\rangle}{\|x\|_2^2}.$$

Corollaire 30. [Ser, page 33] Soient $H \in \mathscr{H}_{n-1}(\mathbb{C})$, $x \in \mathbb{C}^{n-1}$, et $a \in \mathbb{C}$. Notons $\lambda_1 \leqslant \ldots \leqslant \lambda_{n-1}$ les valeurs propres de A, et $\mu_1 \leqslant \ldots \mu_n$ celles de la matrice

$$\widetilde{H} = \begin{bmatrix} H & x \\ x^* & a \end{bmatrix}.$$

Alors, $\mu_1 \leqslant \lambda_1 \leqslant \ldots \leqslant \mu_{n-1} \leqslant \lambda_{n-1} \leqslant \mu_n$. dessin

Corollaire 31. [Cia] Notons $\lambda_1(A) \leqslant \ldots \leqslant \lambda_n(A)$ les valeurs propres d'une matrice A symétrique ou hermitienne. Alors, pour tout $k \in [\![1, n]\!]$, pour toutes matrices A, B symétriques ou hermitiennes,

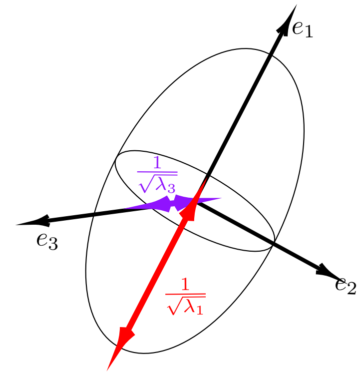
$$|\lambda_k(A) - \lambda_k(B)| \leqslant \|B - A\|_2,$$

où $\|\cdot\|_2$ est la norme subordonnée à la norme euclidienne (resp. hermitienne) usuelle sur $\mathbb{K}^n$.

Corollaire 32. [Rou] (directions principales d'une quadrique) Soit q une forme quadratique définie positive sur $\mathbb{R}^n$, et soit A sa matrice dans la base canonique, de valeurs propres $0 < \lambda_1 \leqslant \ldots \leqslant \lambda_n$. Soit $\{e_1, \ldots, e_n\}$ une base orthonormée de vecteurs propres de A. Considérons la quadrique

$$\begin{aligned} Q &= \{x \in \mathbb{R}^n, \quad q(x) = 1\} \\ &= \{x \in \mathbb{R}^n, \quad {}^t x A x = 1\}. \end{aligned}$$

Les extrema de la norme euclidienne $\|\cdot\|$ sur Q sont $\frac{1}{\sqrt{\lambda_1}}$ et $\frac{1}{\sqrt{\lambda_n}}$, atteints respectivement en $\frac{e_1}{\sqrt{\lambda_1}}$ et $\frac{e_n}{\sqrt{\lambda_n}}$.



3.4 Homéomorphismes remarquables

Théorème 33. L'exponentielle matricielle induit un homéomorphisme entre $\mathscr{S}_n(\mathbb{R})$ et $\mathscr{S}_n^{++}(\mathbb{R})$, ainsi qu'entre $\mathscr{H}_n(\mathbb{C})$ et $\mathscr{H}_n^{++}(\mathbb{C})$.

Lemme 34. Pour toute matrice $A \in \mathscr{S}_n^{+(+)}(\mathbb{R})$ (resp. $\mathscr{H}_n^{+(+)}(\mathbb{C})$), il existe une unique matrice $B \in \mathscr{S}_n^{+(+)}(\mathbb{R})$ (resp. $\mathscr{H}_n^{+(+)}(\mathbb{C})$) telle que $B^2 = A$.

Théorème 35. [H2G2] (décomposition polaire) La multiplication matricielle induit des homéomorphismes

$$O_n(\mathbb{R}) \times \mathscr{S}_n^{++}(\mathbb{R}) \to GL_n(\mathbb{R})$$

et

$$U_n(\mathbb{C}) \times \mathscr{H}_n^{++}(\mathbb{C}) \to GL_n(\mathbb{C}).$$

Corollaire 36. [H2G2] Tout sous-groupe compact de $GL_n(\mathbb{R})$ contenant $O_n(\mathbb{R})$ est égal à $O_n(\mathbb{R})$.

4 Calcul différentiel, nature des points critiques

Définition 37. [Cia, page 146] (matrice hessienne) Soient $U \subset \mathbb{R}^n$ un ouvert et $f : U \to \mathbb{R}$ de classe C^2. La *matrice hessienne* de f en un point $a \in U$ est la matrice

$$D^2 f(a) = (\partial_j f_i(a))_{1 \leqslant 1, j \leqslant n}.$$

Proposition 38. Par le lemme de Schwarz, on a $D^2 f(a) \in \mathscr{S}_n(\mathbb{R})$.

Proposition 39. [Cia] On suppose U convexe et f de classe C^2 sur U.

1. f est convexe si et seulement si pour tout $a \in U$, $D^2 f(a) \in \mathscr{S}_n^+(\mathbb{R})$.
2. Si pour tout $a \in U$, $D^2 f(a) \in \mathscr{S}_n^{++}(\mathbb{R})$, alors f est strictement convexe.

Exemple 40. Soient $A \in \mathscr{S}_n^{++}(\mathbb{R})$, $b \in \mathbb{R}^n$. La fonctionnelle

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

admet A pour hessienne en tout point, donc est strictement convexe sur $\mathbb{R}^n$.

Proposition 41. [Rou] (lemme de Morse) Soient $U \subset \mathbb{R}^n$ ouvert contenant 0, $f : U \to \mathbb{R}$ de classe C^3. On suppose que 0 est un point critique quadratique non dégénéré de f : $d_f(0) = 0$, et $D^2 f(0)$ est non dégénérée, de signature $(p, n-p)$. Alors, il existe un difféomorphisme de classe C^1 entre deux voisinages de 0 dans $\mathbb{R}^n$, noté $\varphi = (\varphi_1, \dots, \varphi_n) : V \to W$, tel que

$$\forall x \in V, f(x) - f(0) = \varphi_1(x)^2 + \cdots + \varphi_p(x)^2 - \varphi_{p+1}(x)^2 - \cdots - \varphi_n(x)^2.$$

dessins

5 Résolution de systèmes linéaires

5.1 Méthode du gradient à pas fixe

Soient $A \in \mathscr{S}_n^{++}(\mathbb{R})$ de valeurs propres $0 < \lambda_1 \leqslant \ldots \leqslant \lambda_n$, $b \in \mathbb{R}^n$, et

$$J : x \in \mathbb{R}^n \mapsto \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle.$$

Lemme 42.

1. Pour tous $u, v \in \mathbb{R}^n$, $\langle \nabla_J(v) - \nabla_J(u), v - u\rangle \geqslant \lambda_1 \|v - u\|^2$.
2. Pour tous $u, v \in \mathbb{R}^n$, $\|\nabla_J(v) - \nabla_J(u)\| = \|A(v-u)\| \leqslant \lambda_n \|v - u\|$.

Théorème 43. [Cia, page 191] (convergence de la méthode du gradient à pas fixe) La fonctionnelle J admet un unique minimiseur $u^* \in \mathbb{R}^n$, caractérisé par $Au^* = b$, et pour tout $\tau \in \left]0, \frac{2\lambda_1}{\lambda_n^2}\right[$, la suite $(u_k)_{k \in \mathbb{N}}$ définie par

$$\begin{cases} u_0 \in \mathbb{R}^n, \\ \forall k \in \mathbb{N}, u_{k+1} = u_k - \tau \nabla_J(u_k), \end{cases}$$

converge vers u^* à vitesse géométrique : il existe $C = C(A) > 0$ telle que pour tout $k \in \mathbb{N}$,

$$\|u_{k+1} - u^*\| \leqslant C \|u_k - u^*\|.$$

5.2 Factorisations LU et de Cholesky

Théorème 44. [Cia, page 83] (factorisation LU) Soit $A \in \mathscr{S}_n^{++}(\mathbb{R})$. Alors, il existe un unique couple (L, U) de matrices tel que U soit triangulaire supérieure, L triangulaire inférieure avec des 1 sur la diagonale, et $A = LU$.

Théorème 45. [Cia, page 83] (factorisation de Cholesky) Soit $A \in \mathscr{S}_n^{++}(\mathbb{R})$. Il existe une unique matrice B triangulaire supérieure à diagonale > 0 telle que $A = {}^t BB$.

Proposition 46. **[Cia]** La méthode de Cholesky nécessite de l'ordre de $\frac{n^3}{6}$ additions, $\frac{n^3}{6}$ multiplications, $\frac{n^2}{2}$ divisions, et n extractions de racines carrées, ce qui la rend plus avantageuse que la méthode du pivot de Gauß ($\frac{n^3}{3}$ additions, $\frac{n^3}{3}$ multiplications, $\frac{n^2}{2}$ divisions) pour résoudre $Ax = b$ avec $A \in \mathscr{S}_n^{++}(\mathbb{R})$.

Remarques

— Pour la réduction de Gauß, la forme énoncée dans [Gou2] est sur un corps quelconque ! Mais dans $\mathbb{R}$ ou $\mathbb{C}$, tout le monde est un carré ou l'opposé d'un carré. Ainsi, la forme réduite $\sum_{i=1}^{r} \lambda_i |x_i|^2$ devient $\sum_{i=1}^{s} \lambda_i |x_i|^2 - \sum_{i=s=1}^{s+t} \lambda_i |x_i|^2$ avec $s+t = r$. Cette remarque est faite au moment d'énoncer la loi d'inertie de Sylvester dans [Gou2], mais on peut la faire avant pour la leçon.

— Une norme $\|\cdot\|$ sur $\mathbb{K}^n$ est euclidienne si et seulement si elle vérifie l'indentité du parallélogramme :

$$\forall x, y \in \mathbb{K}^n, \quad \|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right).$$

— $\mathscr{S}_n^{++}(\mathbb{R})$ (resp. $\mathscr{H}_n^{++}(\mathbb{C})$) est l'orbite de I_n pour l'action de congruence.

— [H2G2] En dimension 1, la décomposition polaire dans le cas complexe exprime l'unicité de l'écriture d'un nombre complexe $z \neq 0$ sous la forme $\rho e^{i\theta}$ avec $\rho > 0$ et $\theta \in \mathbb{R}$ modulo 2π, d'où le terme "décomposition polaire".

— Interprétation géométrique du théorème de Courant-Fischer : l'entrelacement des longueurs des demi-axes principaux lorsqu'on coupe un ellipsoïde par un plan (la section obtenue est une ellipse).

— [Cia] Dans la méthode du gradient, pour un pas τ fixé, la constante C optimale vaut $\|\mathrm{Id} - \tau A\|_2 = \max\{|1-\tau\lambda_1|, |1-\tau\lambda_n|\}$. On en déduit que le meilleur pas possible est $\tau_{opt} = \frac{2}{\lambda_1 + \lambda_n}$. dessin

— L'hypothèse minimale pour que A admette une décomposition LU est que tous ses mineurs principaux soient inversibles. Si $A \in \mathscr{S}_n^{++}(\mathbb{R})$, cette condition est vérifiée.

Références

[Ales] Michel Alessandri, *Thèmes de géométrie*, Dunod, 1999.

[Cia] Philippe Ciarlet, *Introduction à l'analyse numérique matricielle et à l'optimisation*, Dunod, 2006.

[Gou2] Xavier Gourdon, *Les maths en tête, Algèbre*, ellipses, 1999.

[H2G2] Philippe Caldero, Jérôme Germoni, *Histoires hédonistes de groupes et de géométries, tome 1*, Calvage & Mounet, 2013.

[Rou] François Rouvière, *Petit guide de calcul différentiel*, Cassini, 2009.

[Ser] Denis Serre, *Les matrices*, Dunod, 2001.