

Leçon 158. Matrices symétriques réelles, matrices hermitiennes

Devs :

- L'exponentielle de matrices $\exp: \mathcal{S}_n(\mathbb{R}) \rightarrow \mathcal{S}_n^{++}(\mathbb{R})$ est un homéomorphisme
- Lemme de Morse

Références :

1. Gourdon, Algèbre
2. Gourdon, Analyse
3. Grifone, Algèbre linéaire
4. Caldero, H2G2
5. Barbe-Ledoux, Probabilités
6. Objectif Agrégation
7. Hiriart, Optimisation et analyse convexe

On se donne $n \in \mathbb{N}^*$ et E un espace vectoriel de dimension finie sur $K = \mathbb{R}$ ou \mathbb{C} .

1 Généralités

1.1 Définitions et premières propriétés

Définition 1. On dit qu'une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est symétrique si $A = A^T$, et on note $\mathcal{S}_n(\mathbb{R})$ l'ensemble des matrices symétriques.

On dit qu'une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est antisymétrique si $A = -A^T$, et on note $\mathcal{A}_n(\mathbb{R})$ l'ensemble des matrices antisymétriques.

On dit qu'une matrice $H \in \mathcal{M}_n(\mathbb{C})$ est hermitienne si $A = \overline{A}^T$, et on note $\mathcal{H}_n(\mathbb{C})$ l'ensemble des matrices hermitiennes.

Exemple 2. La matrice $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ est symétrique. La matrice $\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$ est hermitienne.

Proposition 3. On a $\mathcal{M}_n(\mathbb{R}) = \mathcal{S}_n(\mathbb{R}) \oplus \mathcal{A}_n(\mathbb{R})$ et $\mathcal{H}_n(\mathbb{C}) = \mathcal{S}_n(\mathbb{R}) \oplus i\mathcal{A}_n(\mathbb{R})$.

Proposition 4. On a $\dim(\mathcal{S}_n(\mathbb{R})) = \frac{n(n+1)}{2}$, $\dim(\mathcal{A}_n(\mathbb{R})) = \frac{n(n-1)}{2}$ et $\dim(\mathcal{H}_n(\mathbb{C})) = n^2$.

Remarque 5. $\mathcal{H}_n(\mathbb{C})$ est un \mathbb{R} -espace vectoriel mais ce n'est pas un \mathbb{C} -espace vectoriel.

Définition 6. On dit qu'une matrice $A \in \mathcal{S}_n(\mathbb{R})$ est positive (resp. définie positive) si $X^T A X \geq 0$ (resp. $X^T A X > 0$) pour tout $X \in \mathbb{R}^n$. On note $\mathcal{S}_n^+(\mathbb{R})$ (resp. $\mathcal{S}_n^{++}(\mathbb{R})$) l'ensemble des matrices symétriques positives (resp. définies positives).

Définition 7. On dit qu'une matrice $H \in \mathcal{H}_n(\mathbb{R})$ est positive (resp. définie positive) si $\overline{X}^T A X \geq 0$ (resp. $\overline{X}^T A X > 0$) pour tout $X \in \mathbb{R}^n$. On note $\mathcal{S}_n^+(\mathbb{R})$ (resp. $\mathcal{S}_n^{++}(\mathbb{R})$) l'ensemble des matrices symétriques positives (resp. définies positives).

Proposition 8. Pour $A \in \mathcal{S}_n(\mathbb{R})$ ou $A \in \mathcal{H}_n(\mathbb{R})$, on a $\text{Sp}(A) \subset \mathbb{R}$.

1.2 Formes bilinéaires symétriques et hermitiennes

E et F sont deux K -espaces vectoriels.

Définition 9. Une application $\varphi: E \times F \rightarrow K$ est appelée une forme bilinéaire si pour tout $x \in E$, l'application $\varphi(x, \cdot)$ est linéaire et pour tout $y \in F$, l'application $\varphi(\cdot, y)$ est linéaire.

Définition 10. Une application $\varphi: E \times F \rightarrow \mathbb{C}$ est appelée une forme sesquilinéaire si pour tout $x \in E$, l'application $\varphi(x, \cdot)$ est linéaire et pour tout $y \in F$, l'application $\varphi(\cdot, y)$ est antilinéaire.

Exemple 11. L'application $\varphi: \begin{cases} \mathcal{C}^0([0, 1], \mathbb{C})^2 \rightarrow \mathbb{C} \\ (f, g) \mapsto \int_0^1 f \overline{g} d\lambda \end{cases}$ est une forme sesquilinéaire sur $\mathcal{C}^0([0, 1], \mathbb{C})$.

Dans la suite, on supposera $F = E$, et on se donne φ une forme bilinéaire sur E , et ψ une forme sesquilinéaire sur E .

Proposition 12. Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E . Alors

$$\varphi(x, y) = \sum_{1 \leq i, j \leq n} x_i y_j \varphi(e_i, e_j) = X^T M Y,$$

où $M = \text{mat}_{\mathcal{B}}(\varphi)$ est la matrice carrée définie par $m_{ij} = \varphi(e_i, e_j)$.

Proposition 13. Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E . Alors

$$\psi(x, y) = \sum_{1 \leq i, j \leq n} \overline{x_i} y_j \psi(e_i, e_j) = \overline{X}^T H Y,$$

où $H = \text{mat}_{\mathcal{B}}(\psi)$ est la matrice carrée définie par $h_{ij} = \psi(e_i, e_j)$.

Définition 14. On dit que φ est symétrique si $\varphi(x, y) = \varphi(y, x)$ et antisymétrique si $\varphi(x, y) = -\varphi(y, x)$. On dit que ψ est hermitienne si $\varphi(x, y) = \varphi(y, \overline{x})$.

Proposition 15.

- φ est symétrique si et seulement si $\text{mat}_{\mathcal{B}}(\varphi) \in \mathcal{S}_n(\mathbb{R})$.
- φ est antisymétrique si et seulement si $\text{mat}_{\mathcal{B}}(\varphi) \in \mathcal{A}_n(\mathbb{R})$.
- ψ est hermitienne si et seulement si $\text{mat}_{\mathcal{B}}(\psi) \in \mathcal{H}_n(\mathbb{C})$.

1.3 Adjoint d'un endomorphisme

On considère un espace $(E, \langle \cdot, \cdot \rangle)$ euclidien, et $f \in \mathcal{L}(E)$.

Proposition 16. Il existe un unique endomorphisme $g \in \mathcal{L}(E)$ tel que :

$$\forall x, y \in E \quad \langle f(x), y \rangle = \langle x, g(y) \rangle$$

g est appelé adjoint de E , et on note $f^* := g$.

Remarque 17. $f \mapsto f^*$ est un endomorphisme involutif de $\mathcal{L}(E)$.

Remarque 18. Si \mathcal{B} est une base orthonormée de E , on a :

$$\text{mat}_{\mathcal{B}}(f^*) = \text{mat}_{\mathcal{B}}(f)^T$$

Proposition 19. Pour tout $f, g \in \mathcal{L}(E)$, on a :

- $(f \circ g)^* = g^* \circ f^*$
- $\text{rg}(f^*) = \text{rg}(f)$ et $\det(f^*) = \det(f)$
- $\text{Im}(f)^\perp = \text{Im}(f^*)$ et $\text{Ker}(f)^\perp = \text{Ker}(f^*)$
- $\chi_{f^*} = \chi_f$ et $\pi_{f^*} = \pi_f$

Définition 20. On dit que f est autoadjoint si $f^* = f$.

Matriciellement, cela se traduit donc par $\text{mat}_{\mathcal{B}}(f) \in \mathcal{S}_n(\mathbb{R})$.

Exemple 21. Les projecteurs sont autoadjoints.

Remarque 22. Si de plus f vérifie $\langle f(x), x \rangle \geq 0$ (resp. > 0), on dit que f est positif (resp. défini positif). Matriciellement, cela se traduit par $\text{mat}_{\mathcal{B}}(f) \in \mathcal{S}_n^+(\mathbb{R})$ (resp. $\text{mat}_{\mathcal{B}}(f) \in \mathcal{S}_n^{++}(\mathbb{R})$).

2 Réduction des endomorphismes symétriques**2.1 Théorème spectral**

Théorème 23. (Théorème spectral).

Soit E un espace euclidien (resp. hermitien) et $f \in \mathcal{L}(E)$ un endomorphisme autoadjoint. Il existe une base orthonormée de vecteurs propres de f .

Corollaire 24. (Théorème spectral, version matricielle).

Soit $M \in \mathcal{S}_n(\mathbb{C})$ (resp. $H \in \mathcal{H}_n(\mathbb{C})$). Alors il existe $P \in \mathcal{O}_n(\mathbb{R})$ (resp. $U \in \mathcal{U}_n(\mathbb{C})$) tel que

$$M = PDP^T \quad (\text{resp. } H = UD\bar{U}^T),$$

où D est une matrice diagonale réelle.

Corollaire 25. On a $M \in \mathcal{S}_n^+(\mathbb{R}) \iff \text{Sp}(M) \subset \mathbb{R}_+$, $M \in \mathcal{S}_n^{++}(\mathbb{R}) \iff \text{Sp}(M) \subset \mathbb{R}_+^*$.

De même, $H \in \mathcal{H}_n^+(\mathbb{C}) \iff \text{Sp}(H) \subset \mathbb{R}_+$ et $H \in \mathcal{H}_n^{++}(\mathbb{C}) \iff \text{Sp}(H) \subset \mathbb{R}_+^*$.

Proposition 26. Soit $H \in \mathcal{H}_n^+(\mathbb{C})$. Alors il existe une unique matrice $R \in \mathcal{H}_n^+(\mathbb{C})$ telle que $R^2 = H$.

Proposition 27. Soit E un espace hermitien et f, g deux endomorphismes autoadjoints tels que $f \circ g = g \circ f$. Alors f et g sont diagonalisables dans une base commune de vecteurs propres orthogonaux.

2.2 Réduction de Gauss des formes quadratiques

On se donne une forme quadratique $\Phi: E \rightarrow K$ (on rappelle que cela signifie que $\Phi(x) = \varphi(x, x)$ avec φ une forme bilinéaire symétrique).

Définition 28. Une base \mathcal{B} de E est dite Φ -orthogonale si pour tout (e, e') distincts dans \mathcal{B} on a $\varphi(e, e') = 0$.

Théorème 29. Si E est de dimension finie, il existe une base Φ -orthogonale pour E .

Théorème 30. (Sylvester).

On suppose $K = \mathbb{R}$.

Il existe une base $\mathcal{B} = (e_1, \dots, e_n)$ de E telle que $\Phi(x) = x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_r^2$, avec $r \leq n$. Autrement dit, la matrice de φ dans \mathcal{B} s'écrit sous la forme $\text{diag}(I_p, -I_{r-p}, 0)$. Le couple $(p, r-p)$ est appelé la signature de Φ .

Exemple 31. La forme quadratique réelle q définie par $q(x, y, z) = x^2 + 2y^2 + 15z^2 - 4xy + 6xz - 8yz$ s'écrit $q(x, y, z) = (x - 2y + 3z)^2 - 2(y - z)^2 + 8z^2$. Sa signature est $(2, 1)$.

2.3 Lien avec les propriétés topologiques de $\mathcal{S}_n(\mathbb{R})$

Proposition 32. Les sous-groupes orthogonaux $O_n(\mathbb{R})$ et unitaires $U_n(\mathbb{C})$ de $\text{GL}_n(\mathbb{R})$ agissent par translation à gauche sur $\mathcal{M}_n(\mathbb{R})$ et sur $\mathcal{M}_n(\mathbb{C})$.

Théorème 33. (décomposition polaire)

La multiplication matricielle induit des homéomorphismes :

$$O_n(\mathbb{R}) \times \mathcal{S}_n^{++}(\mathbb{R}) \simeq \text{GL}_n(\mathbb{R}) \quad \text{et} \quad U_n(\mathbb{C}) \times \mathcal{H}_n^{++}(\mathbb{C}) \simeq \text{GL}_n(\mathbb{C}),$$

via $(O, S) \mapsto OS$ et $(U, H) \mapsto UH$.

Corollaire 34. Pour toute matrice inversible $A \in \text{GL}_n(\mathbb{R})$, on a

$$\|A\|_2 = \sqrt{\rho(A^T A)},$$

où ρ est le rayon spectral de A , donné par $\rho(A) := \max_{1 \leq i \leq n} \{|\lambda_i| : \lambda_i \in \text{Sp}(A)\}$.

Corollaire 35. Tout sous-groupe compact de $\text{GL}_n(\mathbb{R})$ qui contient $O_n(\mathbb{R})$ est $O_n(\mathbb{R})$ lui-même.

Proposition 36. Soit $D \in \mathcal{M}_n(\mathbb{R})$ diagonale, avec $D = \text{diag}(\lambda_1, \dots, \lambda_n)$. Alors $\exp(D) = \text{diag}(e^{\lambda_1}, \dots, e^{\lambda_n})$.

Développement 1 :

Théorème 37. L'exponentielle de matrices $\exp: \mathcal{S}_n(\mathbb{R}) \rightarrow \mathcal{S}_n^{++}(\mathbb{R})$ est un homéomorphisme.

Corollaire 38. Compte tenu du théorème 33, on en déduit que $\text{GL}_n(\mathbb{R}) \simeq O_n(\mathbb{R}) \times \mathbb{R}^{n(n+1)/2}$.

3 Applications et utilisations

3.1 Calcul différentiel

Théorème 39. (Lemme de Schwarz).

Soit f définie sur un ouvert $U \subset \mathbb{R}^2$ de classe \mathcal{C}^2 . Alors $D^2 f \in \mathcal{S}_n(\mathbb{R})$.

Proposition 40.

- Si f admet un minimum (resp un maximum) relatif en $a \in U$, alors la forme quadratique $Q(h) = \left[\sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(x) \right]^{[2]} = \langle Hf(h), h \rangle$ est positive (resp. négative).
- Si $Q(h) = \left[\sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(x) \right]^{[2]} = \langle Hf(h), h \rangle$ est définie positive (resp. définie négative), alors f admet un minimum (resp. un maximum) local en a .

Exemple 41. (cas de la dimension 2)

Soit $f: U \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ de classe \mathcal{C}^2 telle que $df_a = 0$ pour $a \in U$. On note $A = \begin{pmatrix} r & s \\ s & t \end{pmatrix} = Hf(x, y)$.

- Si $\det(A) > 0$ et $r > 0$, alors f admet un minimum relatif en a .

- Si $\det(A) > 0$ et $r < 0$, alors f admet un maximum relatif en a .
- Si $\det(A) < 0$, alors f n'a pas d'extremum en a .
- Si $\det(A) = 0$, on ne peut pas conclure.

Développement 2 :

Lemme 42. (Réduction différentiable des formes quadratiques)

Soit $A_0 \in \mathcal{S}_n(\mathbb{R}) \cap \text{GL}_n(\mathbb{R})$. Alors il existe un voisinage V de A_0 et une application $\Psi: V \rightarrow \text{GL}_n(\mathbb{R})$ de classe \mathcal{C}^1 tels que

$$\forall A \in V \quad A = \Psi(A)^T A_0 \Psi(A).$$

Théorème 43. (Lemme de Morse)

Soit $f: U \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^3 . On suppose que 0 est un point critique quadratique non dégénéré de f , c'est-à-dire que $Df(0) = 0$ et que la forme quadratique hessienne $D^2f(0)$ est non dégénérée, de signature $(p, n - p)$.

Alors il existe deux voisinages V et W de l'origine et un \mathcal{C}^1 -difféomorphisme $\varphi: V \rightarrow W$ tel que $\varphi(0) = 0$, et en notant $u = (u_1, \dots, u_n) =: \varphi(x)$ pour $x \in U$, on a :¹

$$f(x) - f(0) = u_1^2 + \dots + u_p^2 - u_{p+1}^2 - \dots - u_n^2.$$

3.2 Optimisation et résolution de systèmes linéaires

Exemple 44. Soit $A \in \mathcal{S}_n$. Alors $f: x \mapsto \langle x, Ax \rangle$ est convexe sur \mathbb{R}^n si et seulement si A est positive. Elle est strictement convexe si et seulement si $f \in \mathcal{S}_n^{++}$.

Proposition 45. Si $f: E \rightarrow \mathbb{R}$ est convexe, alors tout minimum local est global.

Théorème 46. Soit C un convexe non vide de E et $f: C \rightarrow \mathbb{R}$ une application strictement convexe. Alors il existe au plus un minimum de f sur C .

Lemme 47. (Inégalité de Kantorovitch)

Soit $A \in \mathcal{S}_n^{++}(\mathbb{R})$. On note λ_1 et λ_n la plus petite (respectivement la plus grande) valeur propre de A . Alors :

$$\forall x \in \mathbb{R}^n \quad \langle Ax, x \rangle \langle A^{-1}x, x \rangle \leq \frac{1}{4} \left(\sqrt{c(A)} + \sqrt{\frac{1}{c(A)}} \right)^2 \|x\|^4,$$

où $c(A) \stackrel{\text{def}}{=} \frac{\lambda_1}{\lambda_n}$ désigne le contenu de A .

Théorème 48. Soit $A \in \mathcal{S}_n^{++}(\mathbb{R})$ et $b \in \mathbb{R}^n$. On cherche à minimiser $f: x \mapsto \frac{1}{2} \langle Ax, x \rangle + \langle b, x \rangle$ quand x parcourt \mathbb{R}^n .

Il existe une unique solution \bar{x} à ce problème, et elle est caractérisée par $\nabla f(\bar{x}) = 0$.

De plus, la suite définie par $\begin{cases} x_0 \in \mathbb{R}^n \\ \forall k \in \mathbb{N} \quad x_{k+1} = x_k + t_k d_k \end{cases}$ converge vers \bar{x} , où $t_k = -\nabla f(x_k)$ est l'unique réel minimisant la fonction $t \mapsto f(x_k + t d_k)$.

3.3 Vecteur gaussiens

Définition 49. Une variable aléatoire $X = (X_1, \dots, X_d): (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ est dite gaussienne (on parle de vecteur gaussien) si pour tout $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{R}^d$, il existe $\mu \in \mathbb{R}$ et $\sigma > 0$ tel que

$$\langle \alpha, X \rangle = \sum_{1 \leq i \leq d} \alpha_i X_i \sim \mathcal{N}(\mu, \sigma^2).$$

Proposition 50. Un vecteur gaussien $X = (X_1, \dots, X_d)$ est entièrement caractérisé par le vecteur des espérances $m = (\mathbb{E}[X_1], \dots, \mathbb{E}[X_d])$ et la matrice de variance-covariance $\Gamma = (\text{Cov}(X_i, X_j))_{1 \leq i, j \leq n}$

On dit que $X \sim \mathcal{N}(m, \Gamma)$.

Exemple 51. Si G_1, \dots, G_n sont des variables aléatoires indépendantes de loi normale centrée réduite $\mathcal{N}(0, 1)$, alors $G = (G_1, \dots, G_n)$ est un vecteur gaussien, de loi $\mathcal{N}(0, I_n)$.

Sa loi a pour densité $f(x) = (2\pi)^{-d/2} \exp(-\|x\|^2/2)$ par rapport à la mesure de Lebesgue dx sur \mathbb{R}^d .

Proposition 52. La matrice de variance-covariance Γ est symétrique définie positive. Aussi, il existe toujours une matrice carrée A telle que $\Gamma = A^t A$.

Théorème 53. Soit $X = (X_1, \dots, X_d)$ un vecteur gaussien centré, de matrice de variance-covariance égale à $\Gamma = A^t A$. Alors X a la même loi que AG , où G suit une loi $\mathcal{N}(0, I_d)$.

Théorème 54. Soit $X = (X_1, \dots, X_d)$ un vecteur gaussien de matrice de covariance Γ . Si les composantes de X sont deux à deux non corrélées (i.e si Γ est diagonale), alors la famille (X_1, \dots, X_d) est mutuellement indépendante.