

Rappels et préliminaires

Le long de ce cours, on travaillera sur $\mathbb{E} = \mathbb{R}^n$ avec $n \in \mathbb{N}$.

1.1 Normes et distances

Définition 1 (Norme). Une norme sur \mathbb{R}^n est une application $N : \mathbb{R}^n \rightarrow \mathbb{R}^+$ vérifiant

1. $N(x) = 0$ ssi $x = 0$,
2. $N(\lambda x) = |\lambda|N(x)$ pour tout $\lambda \in \mathbb{R}$ et $x \in \mathbb{R}^n$,
3. $N(x + y) \leq N(x) + N(y)$ pour tout $x, y \in \mathbb{R}^n$.

Exemple 1. On définit la norme Euclidienne sur \mathbb{R}^n (notée $\|\cdot\|$ ou $\|\cdot\|_2$) par

$$\|x\| = \left(\sum_{i=1}^n x_i^2 \right)^{1/2}.$$

Plus généralement, pour $1 \leq p < \infty$, on définit la norme l_p par

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}.$$

Pour $p = \infty$

$$\|x\|_\infty = \max_{i=1, \dots, n} |x_i|.$$

Définition 2 (Métrique). Soit N une norme sur \mathbb{R}^n . On note par $d_N : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^+$ la distance induite par N , qui est définie par

$$d_N(x, y) = N(x - y), \quad \forall x, y \in \mathbb{R}^n,$$

et vérifie

1. $d_N(x, y) = 0$ ssi $x = y$,
2. $d_N(x, y) = d_N(y, x)$, $\forall x, y \in \mathbb{R}^n$,
3. $d_N(x, y) \leq d_N(x, z) + d_N(z, y)$, $\forall x, y, z \in \mathbb{R}^n$.

Définition 3 (Produit scalaire). Soient $x, y \in \mathbb{R}^n$. On note par $\langle x, y \rangle$ le produit scalaire usuel de x et y , défini par

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i.$$

On dira que x et y sont orthogonaux si $\langle x, y \rangle = 0$.

Remarque 1. On notera que $\|x\|_2 = \langle x, x \rangle^{1/2}$.

Définition 4 (Boules). Soit N une norme sur \mathbb{R}^n et $r \geq 0$. On note par $B_N(x, r) = \{y \in \mathbb{R}^n : N(x - y) \leq r\}$ la boule (par rapport à N), centrée en $x \in \mathbb{R}^n$ de rayon r . On

Définition 5 (Ensemble borné). On dira que $\Omega \subset \mathbb{R}^n$ est borné s'il existe $r \geq 0$ tel que $\Omega \subset B_N(0, R)$.

Définition 6. Soit $A, B \subset \mathbb{R}^n$ et $\lambda \in \mathbb{R}$ on définit

$$\lambda A := \{\lambda a : a \in A\}.$$

La somme de Minkowski de A et B est définie par

$$A + B := \{a + b : a \in A, b \in B\}.$$

1.2 Topologie

Définition 7 (Ensemble ouvert, Ensemble fermé). Soient N une norme sur \mathbb{R}^n . Un ensemble $U \subset \mathbb{R}^n$ est un ouvert par rapport à la norme N si pour tout $x \in U$ il existe $r > 0$ tel que $B_N(x, r) \subset U$.

Un ensemble $F \subset \mathbb{R}^n$ est fermé si son complémentaire $F^c := \mathbb{R}^n \setminus F$ est un ouvert.

Théorème 1. 1. Les ensembles \emptyset et \mathbb{R}^n sont à la fois ouverts et fermés.

2. Toute union d'ouverts est toujours un ouvert. Toute intersection finie d'ouverts est un ouvert.

3. Toute union finie de fermés est un fermé. Toute intersection de fermés est un fermé.

Proposition 1. Un ensemble $F \subset \mathbb{R}^n$ est fermé si et seulement si $\forall (x_n)_n \in F^{\mathbb{N}}$

$$\left(\lim_{n \rightarrow \infty} x_n = x \right) \Rightarrow x \in F.$$

Définition 8. Soit $\Omega \subset \mathbb{R}^n$.

1. On dit que $x \in \Omega$ est un point intérieur à Ω s'il existe $\varepsilon > 0$ tel que $B(x, \varepsilon) \subset \Omega$. L'ensemble des points intérieurs à Ω est noté $\text{int}(\Omega)$ ou $\overset{\circ}{\Omega}$.

2. On dira que Ω est ouvert si $\text{int}(\Omega) = \Omega$.

3. L'adhérence de Ω , noté $\text{cl}(\Omega)$ ou $\overline{\Omega}$ est l'ensemble $\mathbb{R}^n \setminus \text{int}(\mathbb{R}^n \setminus \Omega)$.

4. Le bord de Ω est $\partial\Omega = \overline{\Omega} \setminus \text{int}(\Omega)$.

Définition 9 (Ensemble compact). On dira que $\Omega \subset \mathbb{R}^n$ est compact s'il est fermé borné dans \mathbb{R}^n .

Théorème 2. Soit $\Omega \subset \mathbb{R}^n$ un compact. Alors de toute suite $(x_n)_n \in \Omega^{\mathbb{N}}$ on peut extraire une sous-suite convergente.

Théorème 3 (Weierstrass). Soit $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction continue avec Ω un compact. Alors f est bornée et atteint ses bornes, i.e., $\exists x_* x^* \in \Omega$ tels que

$$f(x^*) = \sup_{x \in \Omega} f(x) \text{ et } f(x_*) = \inf_{x \in \Omega} f(x).$$

1.3 Calcul différentiel

1.3.1 Fonctions dérivables, fonctions différentiables

Définition 10. Une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ est dite dérivable en un point $a \in \mathbb{R}$ s'il existe $l_a \in \mathbb{R}$ tel que

$$\frac{f(a+h) - f(a)}{h} \rightarrow l_a, \text{ lorsque } h \rightarrow 0.$$

Cela revient à dire que

$$f(a+h) = f(a) + hl_a + \varepsilon_a(h)|h|,$$

avec $\varepsilon_a(h) \rightarrow 0$ quand $h \rightarrow 0$.

Définition 11. Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$ avec U un ouvert. On dit que f est différentiable en $a \in U$ s'il existe une application linéaire $L_a : \mathbb{R}^n \rightarrow \mathbb{R}^p$ telle que, pour tout $h \in \mathbb{R}^n$ tel que $a+h \in U$

$$f(a+h) = f(a) + L_a(h) + \varepsilon_a(h)\|h\|_{\infty}.$$

De façon équivalente

$$\forall \varepsilon > 0, \exists \delta_{\varepsilon} > 0 : \|h\|_{\infty} \leq \delta_{\varepsilon} \Rightarrow \|f(a+h) - f(a) - L_a(h)\| \leq \varepsilon \|h\|_{\infty}$$

Remarque 2. Rappelons les résultats suivants:

1. Une fonction différentiable en un point $a \in U$ est continue en a .
2. La différentiabilité d'une fonction est équivalente à la différentiabilité des fonctions partielles de f .
3. Quand $n = p = 1$, les notions de différentiabilité et dérivabilité coïncident.

1.3.2 Dérivées partielles

Définition 12. Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$ avec U un ouvert et fixons $k \in \{1, \dots, n\}$. On dit que f admet une dérivée partielle par rapport à sa k ème variable en $a \in U$ si

$$\frac{f(a + te_k) - f(a)}{t}$$

admet une limite quand $t \rightarrow 0$. On note cette limite $\frac{\partial f}{\partial x_k}(a)$. Ici, $\{e_1, \dots, e_n\}$ est la base canonique de \mathbb{R}^n .

1.3.3 Jacobienne et Gradient

Définition 13. Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$ avec U un ouvert. Supposons que f admette en $a \in U$ des dérivées partielles par rapports à toutes ses variables. La matrice Jacobienne de f en a s'écrit

$$J_f(a) = \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \dots & \frac{\partial f_2}{\partial x_n}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_p}{\partial x_1}(a) & \frac{\partial f_p}{\partial x_2}(a) & \dots & \frac{\partial f_p}{\partial x_n}(a) \end{pmatrix}.$$

Si $p = 1$, on définit le gradient de f en a par

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right)^T.$$

1.3.4 Fonctions de classe C^1

Définition 14. Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$ avec U un ouvert. On dit que f est de classe C^1 sur U , et on note $f \in C^1(U; \mathbb{R}^p)$ si pour tout $a \in U$, f admette en a des dérivées partielles par rapports à toutes ses variables et toutes les applications $a \mapsto \frac{\partial f}{\partial x_k}(a)$ sont continues sur U pour tout $k = 1, \dots, n$.

On rappelle le résultat fondamental suivant.

Théorème 4. Soit $f \in C^1(U; \mathbb{R}^p)$ avec $U \subset \mathbb{R}^n$ un ouvert. Soit $a \in U$, $r > 0$ tel que $B(a, r) \subset U$. Pour tout $h \in B(a, r)$

$$f(a + h) = f(a) + J_f(a)h + \varepsilon_a(h)\|h\|_\infty,$$

avec $\varepsilon_a : B(a, r) \rightarrow \mathbb{R}^p$ telle que $\varepsilon_a(h) \rightarrow 0$ quand $h \rightarrow 0$.

Remarque 3. Rappelons que $f \in C^1$ implique que f est différentiable. La différentiabilité de f implique la continuité de f et l'existence des dérivées partielles.

Remarque 4. Rappelons que pour tout $a \in U$ et $h \in \mathbb{R}^n$

$$J_f(a)h = \langle \nabla f(a), h \rangle,$$

i.e., $\nabla f(a) = J_f^T(a)$.

1.3.5 Composition

Théorème 5. Soit $U \subset \mathbb{R}^n, V \subset \mathbb{R}^m$ deux ouverts et $f : U \rightarrow V, g : V \rightarrow \mathbb{R}^p$ deux fonctions de classe C^1 . Alors $g \circ f \in C^1(U; \mathbb{R}^p)$ et on a, pour tout $a \in U$ et $j \in \{1, \dots, n\}$

$$\frac{\partial}{\partial x_j}(g \circ f)(a) = \sum_{k=1}^m \frac{\partial g}{\partial y_k}(f(a)) \frac{\partial f_k}{\partial x_j}(a).$$

Cela s'exprime aussi de manière plus compacte comme

$$\forall a \in U, J_{g \circ f}(a) = J_g(f(a))J_f(a).$$

1.3.6 Dérivées partielles d'ordre deux:

Définition 15. Soit $f \in C^1(U; \mathbb{R})$ avec U un ouvert de \mathbb{R}^n . On dit que f admet des dérivées partielles d'ordre deux en $a \in U$ si $\partial_{x_1} f, \dots, \partial_{x_n} f$ admettent des dérivées partielles en a . On a

$$\frac{\partial^2}{\partial x_i \partial x_j} f(a) := \frac{\partial}{\partial x_i} \left(\frac{\partial}{\partial x_j} f \right) (a).$$

On appelle matrice Hessienne de f en a , notée $H_f(a)$ ou $\nabla^2 f(a)$ la matrice des dérivées partielles de f en a . On a

$$\nabla^2 f(a) := \left(\frac{\partial^2}{\partial x_i \partial x_j} f(a) \right)_{i,j=1,\dots,n}.$$

Définition 16. On dit que $f : U \rightarrow \mathbb{R}$ est de classe C^2 si f est C^1 admettant des dérivées partielles d'ordre deux et pour tout $i, j = 1, \dots, n$, l'application

$$a \in U \mapsto \frac{\partial^2}{\partial x_i \partial x_j} f(a),$$

est continue.

On rappelle le résultat fondamental suivant.

Théorème 6. Soit $f \in C^2(U; \mathbb{R})$. Alors pour tout $i, j = 1, \dots, n$ et $a \in U$

$$\frac{\partial^2}{\partial x_i \partial x_j} f(a) = \frac{\partial^2}{\partial x_j \partial x_i} f(a).$$

Corollaire 1. Soit $f \in C^2(U; \mathbb{R})$. Alors pour tout $a \in U$, $\nabla^2 f(a)$ est symétrique, i.e.,

$$\nabla^2 f(a) = (\nabla^2 f(a))^T.$$

1.4 Structure affine de \mathbb{R}^n

1.4.1 Plans

Définition 17 (Plan). On dit que $A \subset \mathbb{R}^n$ est un plan s'il s'écrit de la forme

$$A = a + V = \{a + v : v \in V\},$$

avec $a \in \mathbb{R}^n$ et V un sous-espace de \mathbb{R}^n .

Remarque 5. La Définition 17 nous dit qu'une partie A de \mathbb{R}^n est un plan si elle est une translation d'un sous-espace vectoriel. Évidemment \mathbb{R}^n est un plan. Par convention, \emptyset est un plan. Dans la littérature, on rencontre souvent d'autres terminologies comme, ensemble affine, espace affine, variété affine etc.

Si $A \subset \mathbb{R}^n$ est un plan non-vide, et $a \in \mathbb{R}^n$. On considère

$$V = \{x - y : x, y \in A\}.$$

Théorème 7. Tout plan non-vidé $A \subset \mathbb{R}^n$ est la translation d'un unique sous-espace vectoriel $V \subset \mathbb{R}^n$ donné par

$$V = \{x - y : x, y \in A\} := A - A.$$

De plus, pour $a \in \mathbb{R}^n$, $A = a + V$ si et seulement si $a \in A$. En particulier, A est un sous-espace si et seulement si $0 \in A$.

Notation. On notera par o , 0_n , $0_{\mathbb{R}^n}$ (ou tout simple 0 s'il n'y'a pas de confusion) le vecteur nul \mathbb{R}^n .

Définition 18 (Espace vectoriel associé et dimension). Soit $A \subset \mathbb{R}^n$ un plan non-vidé. Le sous-espace $V := A - A$ est appelé le sous-espace vectoriel associé à A (ou sous-espace caractéristique de A) et est noté $\text{lin}(A)$. La dimension de A , notée $\dim(A)$ est définie comme la dimension (au sens des espaces vectoriels) de son espace associé V .

Remarque 6. Par convention $\dim(A) = -1$ si $A = \emptyset$. On appellera points (resp. droites) les plans de dimension 0 (resp. 1).

1.4.2 Droites et segments

Définition 19 (Droite). Soient $x, y \in \mathbb{R}^n$ distincts. La droite passant par x et y , notée par $D(x, y)$ ou $\langle x, y \rangle$ est définie par

$$\langle x, y \rangle = \{(1 - \lambda)x + \lambda y : \lambda \in \mathbb{R}\}. \quad (1.1)$$

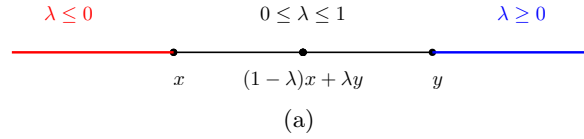


Figure 1.1: Droite passant par x et y .

D'après [Théorème 7](#) et [Remark 6](#), une droite de \mathbb{R}^n est la translation d'un sous-espace de dimension 1. En vue de (1.1), on voit que

$$\langle x, y \rangle = x + \{\lambda(y - x) : \lambda \in \mathbb{R}\},$$

i.e., $\langle x, y \rangle$ est la translation du sous-espace engendré par $y - x$.

Définition 20 (Segments). Soient $x, y \in \mathbb{R}^n$ distincts. Le segment fermé $[x, y]$, ouvert (x, y) et les segments semi-ouverts $(x, y]$ et $[x, y)$ sont définis par

$$\begin{aligned} [x, y] &= \{(1 - \lambda)x + \lambda y : 0 \leq \lambda \leq 1\}, \\ [x, y) &= \{(1 - \lambda)x + \lambda y : 0 \leq \lambda < 1\}, \\ (x, y] &= \{(1 - \lambda)x + \lambda y : 0 < \lambda \leq 1\}, \\ (x, y) &= \{(1 - \lambda)x + \lambda y : 0 < \lambda < 1\}. \end{aligned}$$

1.4.3 Combinaisons affines

Définition 21. Une combinaison affine des points $x_1, \dots, x_m \in \mathbb{R}^n$ est une combinaison linéaire de la forme $\sum_{i=1}^m \lambda_i x_i$ avec $\sum_{i=1}^m \lambda_i = 1$.

Exemple 2. La combinaison affine de deux points $x \neq y \in \mathbb{R}^n$ est la droite passant par x et y .

Théorème 8. Une partie non vide A de \mathbb{R}^n est un plan si et seulement si elle contient tous les points de la forme $(1 - \lambda)x + \lambda y$ pour tout $x, y \in A$ et $\lambda \in \mathbb{R}$.

Preuve. Laissée en exercice. □

Plus généralement, on a

Théorème 9. Une partie non vide A de \mathbb{R}^n est un plan si et seulement si elle contient toutes les combinaisons affines de points de A .

1.4.4 Espace affine engendré ou Enveloppe affine

Définition 22. Soit $A \subset \mathbb{R}^n$. On appelle sous-espace affine engendré par A , note $\text{aff}(A)$, l'intersection de tout les plan contenant A .

Le résultat suivant décrit $\text{aff}(A)$ en terme des combinaisons affines de points de A .

Théorème 10. Soit A une partie non-vide de \mathbb{R}^n .

1. L'enveloppe affine de A est l'ensemble de toutes le combinaisons affine de point de A , i.e.,

$$\text{aff}(A) = \left\{ \sum_{i=1}^m \lambda_i x_i : m \geq 1, x_1, \dots, x_m \in A, \text{ et } \sum_{i=1}^m \lambda_i = 1 \right\}.$$

2. $\text{aff}(A)$ est le plus petit sous-espace affine contenant A .
3. Pour $x_1, \dots, x_m \in \mathbb{R}^n$:

$$\text{aff}(\{x_1, \dots, x_m\}) = \left\{ \sum_{i=1}^m \lambda_i x_i : \sum_{i=1}^m \lambda_i = 1 \right\}$$

1.4.5 Indépendance affine

Définition 23. On dit que des points $x_1, \dots, x_m \in \mathbb{R}^n$ avec $m \geq 2$, sont affinement dépendants s'il existe des scalaires $\lambda_1, \dots, \lambda_m$ non tous nuls tel que

$$\sum_{i=1}^m \lambda_i x_i = 0 \text{ et } \sum_{i=1}^m \lambda_i = 0.$$

On dira que la famille $\{x_1, \dots, x_m\}$ est affinement indépendante si elle n'est pas affinement dépendante.

1.5 Vecteurs et valeurs propres

Définition 24. Soit $A \in \mathbb{R}^{n \times n}$ et $0 \neq v \in \mathbb{R}^n$. On dit que v est un vecteur propre de A s'il existe $\lambda \in \mathbb{C}$ tel que $Av = \lambda v$. On appelle λ la valeur propre associée à v . L'ensemble des valeurs propres de A est appelé spectre de A et est souvent noté par $\sigma(A)$.

Soit $A \in \mathbb{R}^{n \times n}$, en général $\sigma(A) \subset \mathbb{C}$. Si A est symétrique, alors ses valeurs propres sont réelles. On note alors

$$\lambda_{\max}(A) := \lambda_1(A) \geq \dots \geq \lambda_{\min}(A) := \lambda_n(A),$$

les valeurs propres de A où λ_{\max} (respectivement λ_{\min}) est la plus grande (respectivement la plus petite) valeur propre de A .

Théorème 11 (Décomposition spectrale). Soit $A \in \mathbb{S}^n$. Il existe alors $U \in O(n)$ (i.e., $U^t U = U U^t = \text{Id}$) et une matrice diagonale $D = \text{diag}(\alpha_1, \dots, \alpha_n)$ tels que $U^t A U = D$.

Remarque 7. Les colonnes de U représentent les vecteurs propres de A tandis que les éléments diagonaux de D sont les valeurs propres associés.

Corollaire 2. Soit $A \in \mathbb{S}_n$, alors $\text{trace}(A) = \sum_{i=1}^n \lambda_i(A)$ et $\det(A) = \prod_{i=1}^n \lambda_i(A)$.

1.6 Classification des matrices

On rappelle ici quelques résultats concernant la classifications de matrices qui seront utiles pour les conditions d'optimalités dans la ??.

Définition 25. Soit $A \in \mathbb{S}^n$.

1. A est dite semi-définie positive, et on note $A \succeq 0$, si $\langle Ax, x \rangle \geq 0$ pour tout $x \in \mathbb{R}^n$.
2. A est dite définie positive, et on note $A \succ 0$, si $\langle Ax, x \rangle > 0$ pour tout $x \in \mathbb{R}^n \setminus \{0\}$.
3. A est dite semi-définie négative, et on note $A \preceq 0$, si $\langle Ax, x \rangle \leq 0$ pour tout $x \in \mathbb{R}^n$.
4. A est dite définie négative, et on note $A \prec 0$, si $\langle Ax, x \rangle < 0$ pour tout $x \in \mathbb{R}^n \setminus \{0\}$.
5. A est dite non-définie s'il existe $x, y \in \mathbb{R}^n$ tels que

$$\langle Ax, x \rangle < 0 \text{ et } \langle Ay, y \rangle > 0.$$

- Proposition 2.**
1. Soit $A \in \mathbb{S}_+^n$ (respectivement dans \mathbb{S}_{++}^n). Alors les éléments diagonaux de A sont positifs (respectivement strictement positifs).
 2. Soit $A \in \mathbb{S}_-^n$ (respectivement dans \mathbb{S}_{--}^n). Alors les éléments diagonaux de A sont négatifs (respectivement strictement négatifs).

Preuve. Exercice. □

Pour les matrices non-définies, on a le résultat suivant.

Proposition 3. Soit $A \in \mathbb{S}^n$. Si A possède des éléments diagonaux strictement négatifs et positifs, alors A est non-définie.

Preuve. Exercice. □

Le résultat suivant permet de classifier les matrices à partir de l'information spectrale.

Théorème 12. Soit $A \in \mathbb{S}^n$. Alors

1. $A \in \mathbb{S}_{++}^n$ ssi $\sigma(A) \subset \mathbb{R}_+^*$.
2. $A \in \mathbb{S}_+^n$ ssi $\sigma(A) \subset \mathbb{R}_+$.
3. $A \in \mathbb{S}_{--}^n$ ssi $\sigma(A) \subset \mathbb{R}_-^*$.

4. $A \in \mathbb{S}_-^n$ ssi $\sigma(A) \subset \mathbb{R}_-$.

5. $A \in \mathbb{S}_\pm^n$ ssi il existe $\lambda, \mu \in \sigma(A)$ tels que $\lambda > 0$ et $\mu < 0$.

Preuve. Laissée en exercice (utiliser [Théorème 11](#)).

□