Éléments d'analyse convexe

Cours de M1 Mathématiques Fondamentales

Université Paul Sabatier

Pierre Maréchal

Table des matières

Chapitre 1

Préliminaires

Ces notes de cours sont une introduction à quelques notions fondamentales en analyse convexe. Leur étude ne saurait en aucun cas remplacer une lecture approfondie des ouvrages de référence suivants :

1. D. Azé, *Éléments d'Analyse Convexe et Variationnelle*, Ellipses, 1998.

2. M. Bergounioux, *Optimisation et Contrôle des Systèmes Linéaires*, Dunod, 2001.

3. J.-B. Hiriart-Urruty, *L'Optimisation*, Collection *Que sais-je*, Presses Universitaires de France, 1996.

4. J.-B. Hiriart-Urruty & C. Lemaréchal, *Convex Analysis and Minimization Algorithms, I and II*, Springer-Verlag, 1993.

5. R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.

6. M. Willem, *Analyse Convexe et Optimisation* (troisième édition), Éditions Ciaco, Bruxelles, 1989.

1.1 Notations et définitions élémentaires

Dans toute ce qui suit, d sera un entier strictement positif. On munit $\mathbb{R}^d$ du produit scalaire euclidien usuel et de la norme correspondante :

$$\langle \mathbf{x}, \mathbf{y} \rangle := \sum_{j=1}^{d} x_j y_j, \qquad \|\mathbf{x}\| := \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}.$$

L'adhérence d'un sous-ensemble E de $\mathbb{R}^d$ est notée $\bar{E}$ ou encore $\operatorname{cl} E$. L'intérieur de E est noté $\operatorname{int} E$. On rappelle que

$$\operatorname{int} E = \{\mathbf{x} \in E \mid \exists \varepsilon > 0 \colon B(\mathbf{x}, \varepsilon) \subset E\},$$

où $B(\mathbf{x}, \varepsilon) := \{\mathbf{y} \in \mathbb{R}^d \mid \|\mathbf{y} - \mathbf{x}\| < \varepsilon\}$. L'ensemble de tous les sous-ensembles de $\mathbb{R}^d$ est noté $\mathscr{P}(\mathbb{R}^d)$.

Nous définissons maintenant quelques opérations sur les sous-ensembles de $\mathbb{R}^d$. Si $E, F \in \mathscr{P}(\mathbb{R}^d)$, la somme de E et de F est l'ensemble

$$E + F := \{\mathbf{x} + \mathbf{y} \mid \mathbf{x} \in E,\ \mathbf{y} \in F\}.$$

Si λ est un réel, l'ensemble λE est l'homothétique de E de centre $\mathbf{0}$ et de rapport λ :

$$\lambda E := \{\lambda \mathbf{x} \mid \mathbf{x} \in E\}.$$

En particulier, le symétrique de E de centre $\mathbf{0}$ est l'ensemble $-E = (-1)E$. Il faut bien sûr éviter de confondre $E - F := E + (-F)$ et

$$E \setminus F := \{\mathbf{x} \in E \mid \mathbf{x} \notin F\}.$$

Si $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$ et $\mathbf{x}' = (x'_1, \ldots, x'_{d'}) \in \mathbb{R}^{d'}$, où d' est un autre entier strictement positif, on définit

$$(\mathbf{x}, \mathbf{x}') = (x_1, \ldots, x_d, x'_1, \ldots, x'_{d'}) \in \mathbb{R}^{d+d'}.$$

Si $E \subset \mathbb{R}^d$ et $E' \subset \mathbb{R}^{d'}$, le produit cartésien de E par E' est l'ensemble

$$E \times E' := \{(\mathbf{x}, \mathbf{x}') \mid \mathbf{x} \in E,\ \mathbf{x}' \in E'\}.$$

L'intersection de deux sous-ensembles E, F de $\mathbb{R}^d$ est notée $E \cap F$, et si $(E_\alpha)_{\alpha \in A}$ est une famille de sous-ensembles de $\mathbb{R}^d$ (ici, A est un ensemble quelconque d'indices), l'intersection des E_α est l'ensemble

$$\bigcap_{\alpha \in A} E_\alpha := \left\{\mathbf{x} \in \mathbb{R}^d \mid \forall \alpha \in A,\ \mathbf{x} \in E_\alpha\right\}.$$

1.2 Sous-espaces vectoriels

Un sous-ensemble $V \subset \mathbb{R}^d$ est un *sous-espace vectoriel* si et seulement si

(a) $V \neq \emptyset$;

(b) $\forall \mathbf{x}, \mathbf{y} \in V,\ \forall \alpha, \beta \in \mathbb{R},\ \alpha \mathbf{x} + \beta \mathbf{y} \in V$.

Une combinaison linéaire d'éléments de $\mathbb{R}^d$ est un vecteur de la forme $\alpha_1 \mathbf{x}_1 + \cdots + \alpha_m \mathbf{x}_m$, où

$$m \in \mathbb{N}^*, \quad \alpha_1, \ldots, \alpha_m \in \mathbb{R} \quad \text{et} \quad \mathbf{x}_1, \ldots, \mathbf{x}_m \in \mathbb{R}^d.$$

On dit qu'un sous-ensemble U de $\mathbb{R}^d$ est stable par combinaison linéaire lorsque

$$\forall m \in \mathbb{N}^*,\ \forall \alpha_1, \ldots, \alpha_m \in \mathbb{R},\ \forall \mathbf{x}_1, \ldots, \mathbf{x}_m \in U,\ \ \alpha_1 \mathbf{x}_1 + \cdots + \alpha_m \mathbf{x}_m \in U.$$

On rappelle que V est un sous-espace vectoriel si et seulement si V est non vide et stable par combinaison linéaire. On note $\mathscr{V}(\mathbb{R}^d)$ l'ensemble de tous les sous-espaces vectoriels de $\mathbb{R}^d$. Insistons sur le fait que $\emptyset \notin \mathscr{V}(\mathbb{R}^d)$. Enfin, il est facile de voir que $\mathscr{V}(\mathbb{R}^d)$ est stable par intersection : si A est un ensemble quelconque d'indices,

$$(V_\alpha)_{\alpha\in A} \subset \mathscr{V}(\mathbb{R}^d) \Longrightarrow \bigcap_{\alpha\in A} V_\alpha \in \mathscr{V}(\mathbb{R}^d).$$

Rappelons que, si V est un sous-espace vectoriel de $\mathbb{R}^d$, alors toutes les bases de V ont le même cardinal, appelé la dimension de V. La dimension de V est notée $\dim V$.

1.3 Sous-espaces affines

Si $\mathbf{x}$ et $\mathbf{y}$ sont deux vecteurs distincts de $\mathbb{R}^d$, la droite passant par $\mathbf{x}$ et $\mathbf{y}$ est l'ensemble $\{(1-\lambda)\mathbf{x} + \lambda\mathbf{y} \mid \lambda \in \mathbb{R}\}$. Un sous-ensemble A de $\mathbb{R}^d$ est appelé un sous-espace affine lorsqu'il satisfait :

$$\forall \mathbf{x}, \mathbf{y} \in A,\ \forall \lambda \in \mathbb{R},\ \ (1-\lambda)\mathbf{x} + \lambda\mathbf{y} \in A.$$

Par exemple, $\emptyset$ et $\mathbb{R}^d$ sont des sous-espaces affines. De même, les singletons $\{\mathbf{x}\}$, les sous-espaces vectoriels, les droites, sont des sous-espaces affines. On note $\mathscr{A}(\mathbb{R}^d)$ l'ensemble de tous les sous-espaces affines de $\mathbb{R}^d$.

Théorème 1. Soit $A \in \mathscr{A}(\mathbb{R}^d)$. Alors A est un sous-espace vectoriel si et seulement si $\mathbf{0} \in A$.

Démonstration. Il est évident que, si $A \in \mathscr{V}(\mathbb{R}^d)$, alors $\mathbf{0} \in A$. Réciproquement, supposons que $\mathbf{0} \in A$. Alors,

$$\forall \mathbf{x} \in A,\ \forall \lambda \in \mathbb{R},\ \ \lambda\mathbf{x} = (1-\lambda)\mathbf{0} + \lambda\mathbf{x} \in A,$$

et donc A est stable par multiplication par les scalaires réels. Si maintenant $\mathbf{x}, \mathbf{y} \in A$, alors

$$\frac{\mathbf{x}+\mathbf{y}}{2} = \left(1 - \frac{1}{2}\right)\mathbf{x} + \frac{1}{2}\mathbf{y} \in A,$$

ce qui implique que $\mathbf{x} + \mathbf{y} \in A$. Donc A est aussi stable par addition. Ainsi, $A \in \mathscr{V}(\mathbb{R}^d)$. ∎

Théorème 2. Soient $A \in \mathscr{A}(\mathbb{R}^d)$ et $\mathbf{x}_0 \in \mathbb{R}^d$. Alors $A + \{\mathbf{x}_0\} \in \mathscr{A}(\mathbb{R}^d)$. Autrement dit, $\mathscr{A}(\mathbb{R}^d)$ est stable par translation.

Démonstration. C'est un exercice facile. ∎

Un sous-espace affine A est dit parallèle à un sous-espace affine B s'il existe $\mathbf{x}_0 \in \mathbb{R}^d$ tel que $B = A + \{\mathbf{x}_0\}$. On écrit alors $A \parallel B$. On vérifie facilement que $\parallel$ est une relation d'équivalence dans $\mathscr{A}(\mathbb{R}^d)$.

Théorème 3. Tout sous-espace affine non vide A est parallèle à un unique sous-espace vectoriel V, qui satisfait

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

Démonstration. Soit $\mathbf{y}$ un point quelconque de A. D'après le théorème 2, $A - \{\mathbf{y}\} \in \mathscr{A}(\mathbb{R}^d)$. Puisque $A - \{\mathbf{y}\}$ contient $\mathbf{0}$, le théorème 1 montre qu'en fait $A - \{\mathbf{y}\} \in \mathscr{V}(\mathbb{R}^d)$. L'existence est donc claire.
Pour montrer l'unicité, supposons que V_1, V_2 sont deux sous-espaces vectoriels parallèles à A. Par transitivité de la relation $\|$, on voit que $V_1 \parallel V_2$, et donc il existe $\mathbf{x}_0 \in \mathbb{R}^d$ tel que

$$V_2 = V_1 + \{\mathbf{x}_0\}.$$

Puisque $\mathbf{0} \in V_2$, $-\mathbf{x}_0 \in V_1$ et donc $\mathbf{x}_0 \in V_1$. On en déduit que $V_1 \supset V_1 + \{\mathbf{x}_0\} = V_2$. Par un raisonnement symétrique, $V_2 \supset V_1$, et donc $V_2 = V_1$.
Finalement, l'unicité qui vient d'être établie permet d'écrire que $V = A - \{\mathbf{y}\}$ pour tout $\mathbf{y} \in A$, et donc que

$$V = \bigcup_{\mathbf{y} \in A} \big(A - \{\mathbf{y}\}\big) = A - A. \blacksquare$$

Une *combinaison affine* d'éléments de $\mathbb{R}^d$ est un vecteur de la forme $\alpha_1 \mathbf{x}_1 + \cdots + \alpha_m \mathbf{x}_m$, où

$$m \in \mathbb{N}^*, \quad \alpha_1, \ldots, \alpha_m \in \mathbb{R}, \quad \alpha_1 + \cdots + \alpha_m = 1 \quad \text{et} \quad \mathbf{x}_1, \ldots, \mathbf{x}_m \in \mathbb{R}^d.$$

On montre que A est un sous-espace affine si et seulement si pour tout $m \in \mathbb{N}^*$, pour tout $(\alpha_1, \ldots, \alpha_m) \in \mathbb{R}^m$ tel que $\alpha_1 + \cdots + \alpha_m = 1$ et pour tous $\mathbf{x}_1, \ldots, \mathbf{x}_m \in A$,

$$\alpha_1 \mathbf{x}_1 + \cdots + \alpha_m \mathbf{x}_m \in A.$$

On remarque que $\emptyset \in \mathscr{A}(\mathbb{R}^d)$, et que $\mathscr{A}(\mathbb{R}^d)$ est stable par intersection. Enfin, la *dimension* d'un sous-espace affine A est, par définition, la dimension de l'unique sous-espace vectoriel parallèle à A si A est non vide, et zéro si A est vide. La dimension de A est notée $\dim A$.

1.4 Sous-ensembles convexes

Soient $\mathbf{x}$ et $\mathbf{y}$ deux points de $\mathbb{R}^d$. On appelle *segment joignant* $\mathbf{x}$ *et* $\mathbf{y}$ l'ensemble

$$[\mathbf{x}, \mathbf{y}] := \{(1 - \lambda)\mathbf{x} + \lambda \mathbf{y} \mid \lambda \in [0, 1]\}.$$

On dit qu'un sous-ensemble C de $\mathbb{R}^d$ est convexe lorsque C contient tout segment joignant deux de ses points, autrement dit, lorsque

$$\forall \mathbf{x}, \mathbf{y} \in C, \quad [\mathbf{x}, \mathbf{y}] \subset C.$$

On note $\mathscr{C}(\mathbb{R}^d)$ l'ensemble de tous les sous-ensembles convexes de $\mathbb{R}^d$.

Théorème 4. L'intersection d'une famille quelconque de sous-ensembles convexes est convexe.

DÉMONSTRATION. Soit $(C_\alpha)_{\alpha\in A} \subset \mathscr{C}(\mathbb{R}^d)$ (A est un ensemble quelconque d'indices), et soit $C = \cap\{C_\alpha \mid \alpha \in A\}$. Soient $\mathbf{x}, \mathbf{y} \in C$ et $\lambda \in [0,1]$. Pour tout $\alpha \in A$, $\mathbf{x}$ et $\mathbf{y}$ appartiennent à C_α, qui est convexe. On en déduit que

$$\forall \alpha \in A, \quad (1-\lambda)\mathbf{x} + \lambda\mathbf{y} \in C_\alpha.$$

Autrement dit, $(1-\lambda)\mathbf{x} + \lambda\mathbf{y} \in C$. ∎

Théorème 5. (i) Soient $C \in \mathscr{C}(\mathbb{R}^d)$ et μ un réel. Alors $\mu C \in \mathscr{C}(\mathbb{R}^d)$.

(ii) Soient $C_1, C_2 \in \mathscr{C}(\mathbb{R}^d)$. Alors $C_1 + C_2 \in \mathscr{C}(\mathbb{R}^d)$.

(iii) Soient $C_1 \in \mathbb{R}^{d_1}$ et $C_2 \in \mathbb{R}^{d_2}$, où d_1 et d_2 sont deux entiers strictement positifs. Alors $C_1 \times C_2 \in \mathscr{C}(\mathbb{R}^{d_1+d_2})$.

DÉMONSTRATION. C'est un exercice facile. ∎

On appelle *combinaison convexe* d'éléments de $\mathbb{R}^d$ tout vecteur de la forme $\alpha_1\mathbf{x}_1 + \cdots + \alpha_m\mathbf{x}_m$, où

$$m \in \mathbb{N}^*, \quad \alpha_1, \ldots, \alpha_m \in \mathbb{R}_+, \quad \alpha_1 + \cdots + \alpha_m = 1 \quad \text{et} \quad \mathbf{x}_1, \ldots, \mathbf{x}_m \in \mathbb{R}^d.$$

On appelle *simplexe de* $\mathbb{R}^m$ l'ensemble

$$\Sigma_m := \left\{(\alpha_1, \ldots, \alpha_m) \in \mathbb{R}_+^m \;\middle|\; \alpha_1 + \cdots + \alpha_m = 1\right\}.$$

On montre que $C \in \mathscr{C}(\mathbb{R}^d)$ si et seulement si C est stable par combinaison convexe, c'est-à-dire,

$$\forall m \in \mathbb{N}^*,\ \forall \alpha_1, \ldots, \alpha_m \in \mathbb{R}_+ : \alpha_1 + \cdots + \alpha_m = 1,$$
$$\forall \mathbf{x}_1, \ldots, \mathbf{x}_m \in C,\ \ \alpha_1\mathbf{x}_1 + \cdots + \alpha_m\mathbf{x}_m \in C.$$

1.5 Enveloppes

Soit $E \subset \mathbb{R}^d$. L'*enveloppe vectorielle de E* (ou *sous-espace vectoriel engendré par E*) est l'intersection de tous les sous-espaces vectoriels de $\mathbb{R}^d$ qui contiennent E. On la note $\operatorname{vect} E$:

$$\operatorname{vect} E = \bigcap \left\{V \in \mathscr{V}(\mathbb{R}^d) \;\middle|\; E \subset V\right\}.$$

L'*enveloppe affine de E* est l'intersection de tous les sous-espaces affines de $\mathbb{R}^d$ qui contiennent E. On la note $\operatorname{aff} E$:

$$\operatorname{aff} E = \bigcap \left\{A \in \mathscr{A}(\mathbb{R}^d) \;\middle|\; E \subset A\right\}.$$

L'*enveloppe convexe de* E est l'intersection de tous les sous-ensembles convexes de $\mathbb{R}^d$ qui contiennent E. On la note $\operatorname{co} E$:

$$\operatorname{co} E = \bigcap \big\{ C \in \mathscr{C}(\mathbb{R}^d) \;\big|\; E \subset C \big\}.$$

On remarque que $\operatorname{vect} \emptyset = \{\mathbf{0}\}$, alors que $\operatorname{aff} \emptyset = \operatorname{co} \emptyset = \emptyset$.

Théorème 6. Soit $E \subset \mathbb{R}^d$.

(i) Si $E \neq \emptyset$, alors $\operatorname{vect} E$ est l'ensemble de toutes les combinaisons linéaires d'éléments de E ;

(ii) $\operatorname{aff} E$ est l'ensemble de toutes les combinaisons affines d'éléments de E ;

(iii) $\operatorname{co} E$ est l'ensemble de toutes les combinaisons convexes d'éléments de E.

Démonstration. Les trois assertions se démontrent de manière analogue, et nous n'écrivons ici que la preuve de la troisième. Posons

$$C := \{\alpha_1 \mathbf{x}_1 + \cdots + \alpha_m \mathbf{x}_m \mid m \in \mathbb{N}^*,\ (\alpha_1, \ldots, \alpha_m) \in \Sigma_m,\ \mathbf{x}_1, \ldots, \mathbf{x}_m \in E\}.$$

Il est clair que $E \subset C$. Montrons que C est convexe. Soient $\mathbf{x}, \mathbf{y} \in C$ et α, β deux réels positifs tels que $\alpha + \beta = 1$. Alors,

$$\alpha \mathbf{x} + \beta \mathbf{y} = \alpha(\alpha_1 \mathbf{x}_1 + \cdots + \alpha_m \mathbf{x}_m) + \beta(\beta_1 \mathbf{y}_1 + \cdots + \beta_p \mathbf{y}_p),$$

où $m, p \in \mathbb{N}^*$, $(\alpha_1, \ldots, \alpha_m) \in \Sigma_m$, $(\beta_1, \ldots, \beta_p) \in \Sigma_p$, et

$$\mathbf{x}_1, \ldots, \mathbf{x}_m, \mathbf{y}_1, \ldots, \mathbf{y}_p \in E.$$

Il est facile de voir que $(\alpha\alpha_1, \ldots, \alpha\alpha_m, \beta\beta_1, \ldots, \beta\beta_p) \in \Sigma_{m+p}$. On en déduit que $\alpha \mathbf{x} + \beta \mathbf{y}$ est une combinaison convexe d'éléments de E, autrement dit, que $\alpha \mathbf{x} + \beta \mathbf{y} \in C$.
Puisque $E \subset C$ et C est convexe, $\operatorname{co} E \subset C$, et il reste à établir l'inclusion opposée. Soit donc $\mathbf{z} \in C$:

$$\mathbf{z} = \gamma_1 \mathbf{z}_1 + \cdots + \gamma_q \mathbf{z}_q,$$

où $q \in \mathbb{N}^*$, $(\gamma_1, \ldots, \gamma_q) \in \Sigma_q$ et $\mathbf{z}_1, \ldots, \mathbf{z}_q \in E \subset \operatorname{co} E$. On voit que $\mathbf{z}$ est une combinaison convexe d'éléments de $\operatorname{co} E$, et comme $\operatorname{co} E$ est convexe (d'après le théorème 4), il s'ensuit que $\mathbf{z} \in \operatorname{co} E$. ∎

La notion d'enveloppe affine permet de définir la dimension d'un sous-ensemble quelconque de $\mathbb{R}^d$: si E est quelconque dans $\mathscr{P}(\mathbb{R}^d)$, on appelle *dimension* de E la dimension de son enveloppe affine. On la note $\dim E$.

Chapitre 2

Propriétés topologiques et séparation

2.1 Intérieurs relatifs

Considérons, dans $\mathbb{R}^3$, le disque

$$D := \left\{(x, y, z) \in \mathbb{R}^3 \;\middle|\; z = 0,\ x^2 + y^2 \leq 1\right\}.$$

L'intérieur de D (dans la topologie usuelle) est évidemment vide. Cependant, si l'on regarde D comme un sous-ensemble du plan horizontal (c'est-à-dire, de son enveloppe affine), les points de l'intérieur de D sont ceux qui satisfont la condition $x^2 + y^2 < 1$.
De manière générale, on définit l'intérieur relatif d'un sous-ensemble E de $\mathbb{R}^d$ comme l'intérieur de E considéré comme sous-ensemble de son enveloppe affine. On le note $\operatorname{ri} E$:

$$\operatorname{ri} E := \{\mathbf{x} \in E \,|\, \exists \varepsilon > 0\colon \big(B(\mathbf{x}, \varepsilon) \cap \operatorname{aff} E\big) \subset E\}. \tag{2.1}$$

Dans l'exemple introductif précédent, on a donc :

$$\operatorname{ri} D = \left\{(x, y, z) \in \mathbb{R}^3 \;\middle|\; z = 0,\ x^2 + y^2 < 1\right\}.$$

On parle parfois de *topologie relative à l'ensemble E*. L'équation (2.1) montre que l'on a toujours l'inclusion $\operatorname{int} E \subset \operatorname{ri} E$, avec égalité lorsque $\operatorname{aff} E = \mathbb{R}^d$. Par ailleurs, on peut se convaincre aisément que la fermeture de E n'est pas affectée par passage à la topologie relative. Lorsque $\operatorname{ri} E = E$, on dit que E est *relativement ouvert*.
La notion d'intérieur relatif est intéressante surtout pour les ensembles convexes. Nous verrons notamment que si $C \in \mathscr{C}(\mathbb{R}^d)$ est non vide, alors $\operatorname{ri} C$ est aussi (convexe et) non vide.
Il est important de réaliser que le passage à l'intérieur relatif ne préserve pas, en général, l'inclusion (contrairement au passage à l'intérieur). Soient par exemple,

dans $\mathbb{R}^2$, les sous-ensembles

$$C := [0,1]^2 \quad \text{et} \quad C' := [0,1] \times \{0\}.$$

L'enveloppe affine de C est $\mathbb{R}^2$ tout entier, alors que celle de C' est la droite horizontale $\mathbb{R} \times \{0\}$. Par conséquent,

$$\operatorname{ri} C =]0,1[^2 \quad \text{et} \quad \operatorname{ri} C' =]0,1[\times \{0\}.$$

On voit donc que, bien que $C' \subset C$, $\operatorname{ri} C' \not\subset \operatorname{ri} C$.
Enfin, on appelle frontière relative d'un ensemble E l'ensemble

$$\operatorname{rbd} E := \operatorname{cl} E \setminus \operatorname{ri} E.$$

Par exemple, $\operatorname{rbd} D = \{(x,y,z) \in \mathbb{R}^3 \mid z = 0,\ x^2 + y^2 = 1\}$.

Théorème 7. Soit $C \in \mathscr{C}(\mathbb{R}^d)$ tel que $\operatorname{int} C \neq \emptyset$. Soient $\mathbf{x} \in \operatorname{int} C$ et $\mathbf{y} \in \operatorname{cl} C$. Alors

$$[\mathbf{x}, \mathbf{y}[:= \{(1-t)\mathbf{x} + t\mathbf{y} \mid t \in [0,1[\} \subset \operatorname{int} C.$$

De plus, $\operatorname{int} C = \operatorname{int} \operatorname{cl} C$ et $\operatorname{cl} C = \operatorname{cl} \operatorname{int} C$.

Démonstration. Etape 1 : Nous montrons que si $\mathbf{z} := (1-t)\mathbf{x} + t\mathbf{y}$ avec $t \in]0,1[$, alors $\mathbf{z} \in \operatorname{int} C$. Soit B une boule ouverte de centre $\mathbf{x}$ telle que $B \subset C$. Soient $\mu := (t-1)/t$ et h_μ l'homothétie de centre $\mathbf{z}$ et de rapport μ. On a :

$$\begin{aligned} h_\mu(\mathbf{x}) &= \mathbf{z} + \frac{t-1}{t}(\mathbf{x} - \mathbf{z}) \\ &= (1-t)\mathbf{x} + t\mathbf{y} + \frac{t-1}{t}(t\mathbf{x} - t\mathbf{y}) \\ &= \mathbf{y}. \end{aligned}$$

Comme h_μ est linéaire et inversible, $h_\mu(B)$ est un ouvert, et ce qui précède montre que $h_\mu(B)$ contient $\mathbf{y}$. C'est donc un voisinage de $\mathbf{y}$. Puisque $\mathbf{y} \in \operatorname{cl} C$, on peut trouver un point $\mathbf{x}_0$ dans $C \cap h_\mu(B)$. Il existe alors $\mathbf{z}_0 \in B$ tel que $\mathbf{x}_0 = h_\mu(\mathbf{z}_0)$. Soit

$$\gamma := \frac{\mu}{\mu - 1} = 1 - t \in]0,1[\,,$$

et soit h_γ l'homothétie de centre $\mathbf{x}_0$ et de rapport γ. Alors,

$$\begin{aligned} h_\gamma(\mathbf{z}_0) &= \mathbf{x}_0 + (1-t)(\mathbf{z}_0 - \mathbf{x}_0) \\ &= \mathbf{z} + \frac{t-1}{t}(\mathbf{z}_0 - \mathbf{z}) + (1-t)\left(\mathbf{z}_0 - \mathbf{z} + \frac{1-t}{t}(\mathbf{z}_0 - \mathbf{z})\right) \\ &= \mathbf{z} + \frac{t-1}{t}(\mathbf{z}_0 - \mathbf{z}) + \frac{1-t}{t}(\mathbf{z}_0 - \mathbf{z}) \\ &= \mathbf{z}. \end{aligned}$$

Puisque $\mathbf{z}_0 \in B$, on déduit du calcul précédent que

$$\mathbf{z} \in h_\gamma(B) = \{\mathbf{x}_0\} + (1-t)\big(B - \{\mathbf{x}_0\}\big) = \{t\mathbf{x}_0 + (1-t)\mathbf{u} \mid \mathbf{u} \in B\} \subset C.$$

En résumé, $h_\gamma(B)$ est une boule ouverte contenant $\mathbf{z}$, contenue dans C. On a donc bien que $\mathbf{z} \in \operatorname{int} C$.

Etape 2 : Nous montrons maintenant que $\operatorname{int} C = \operatorname{int} \operatorname{cl} C$. Comme $C \subset \operatorname{cl} C$, il est clair que $\operatorname{int} C \subset \operatorname{int} \operatorname{cl} C$. Réciproquement, soit $\mathbf{z} \in \operatorname{int} \operatorname{cl} C$, soit $r > 0$ tel que $\operatorname{cl} B(\mathbf{z}, r) \subset \operatorname{cl} C$, et soit $\mathbf{x} \in \operatorname{int} C$ tel que $\mathbf{x} \neq \mathbf{z}$. Choisissons $t > 0$ tel que $t\|\mathbf{x} - \mathbf{z}\| \leq r$ (donc $\mathbf{z} - t(\mathbf{x} - \mathbf{z}) \in \operatorname{cl} B(\mathbf{z}, r)$). D'après l'étape 1,

$$\left.\begin{array}{r} \mathbf{z} - t(\mathbf{x} - \mathbf{z}) \in \operatorname{cl} C \\ \mathbf{x} \in \operatorname{int} C \end{array}\right\} \Longrightarrow [\mathbf{x}, \mathbf{z} - t(\mathbf{x} - \mathbf{z})[\subset \operatorname{int} C.$$

Puisque $\mathbf{z} \in [\mathbf{x}, \mathbf{z} - t(\mathbf{x} - \mathbf{z})[$, la dernière inclusion montre que $\mathbf{z} \in \operatorname{int} C$.

Etape 3 : Nous montrons finalement que $\operatorname{cl} C = \operatorname{cl} \operatorname{int} C$. Comme $\operatorname{int} C \subset C$, il est clair que $\operatorname{cl} \operatorname{int} C \subset \operatorname{cl} C$. Réciproquement, soit $\mathbf{x} \in \operatorname{cl} C$, et soit $\mathbf{z} \in \operatorname{int} C$. D'après l'étape 1,

$$\mathbf{x}_n := \frac{\mathbf{z}}{n} + \left(1 - \frac{1}{n}\right)\mathbf{x} \in \operatorname{int} C.$$

Comme $\mathbf{x}_n \to \mathbf{x}$ lorsque $n \to \infty$, on voit que $\mathbf{x} \in \operatorname{cl} \operatorname{int} C$. ∎

Corollaire 1. Soit $C \in \mathscr{C}(\mathbb{R}^d)$, et soient $\mathbf{x} \in \operatorname{ri} C$, $\mathbf{y} \in \operatorname{cl} C$. Alors $[\mathbf{x}, \mathbf{y}[\subset \operatorname{ri} C$.

Démonstration. Il suffit de passer à la topologie relative, et d'appliquer le théorème. ∎

Corollaire 2. Soit $C \in \mathscr{C}(\mathbb{R}^d)$ non vide. Il y a équivalence entre

(i) $\mathbf{z} \in \operatorname{ri} C$;

(ii) $\forall \mathbf{x} \in C$, $\exists \mu > 1 \colon (1 - \mu)\mathbf{x} + \mu\mathbf{z} \in C$.

Démonstration. Si $\mathbf{z} \in \operatorname{ri} C$, il est clair que tout segment inclus dans C ayant $\mathbf{z}$ comme une de ses extrémités peut être prolongé au-dela de $\mathbf{z}$ sans quitter C. C'est la condition (ii).

Réciproquement, supposons (ii). Puisque $\operatorname{ri} C \neq \emptyset$, on peut choisir un point $\mathbf{x} \in \operatorname{ri} C$. Soit alors $\mathbf{y} := (1 - \mu)\mathbf{x} + \mu\mathbf{z}$, avec $\mu > 1$, tel que $\mathbf{y} \in C$. Alors

$$\mathbf{z} = \frac{1}{\mu}\big(\mathbf{y} - (1 - \mu)\mathbf{x}\big) = \lambda\mathbf{y} + (1 - \lambda)\mathbf{x}, \quad \text{avec} \quad \lambda := \frac{1}{\mu} \in]0, 1[\,.$$

Le théorème 7 montre alors que $\mathbf{z} \in \operatorname{ri} C$. ∎

Soit $A := \{\mathbf{x}_0, \ldots, \mathbf{x}_k\} \subset \mathbb{R}^d$. On dit que A est *affinement indépendant* si l'ensemble

$$\operatorname{aff}\{\mathbf{x}_0, \ldots, \mathbf{x}_k\} = \{\mathbf{x}_0\} + \operatorname{vect}\{\mathbf{x}_1 - \mathbf{x}_0, \ldots, \mathbf{x}_k - \mathbf{x}_0\}$$

est de dimension k. Bien sûr, dans l'égalité ci-dessus, le rôle donné à $\mathbf{x}_0$ aurait pu être donné à n'importe quel autre élément de A. On peut se convaincre aisément qu'un sous-ensemble E de $\mathbb{R}^d$ est de dimension k si et seulement si il contient un ensemble $\{\mathbf{x}_0, \ldots, \mathbf{x}_k\}$ affinement indépendant.

Lemme 1. $\{\mathbf{x}_0, \ldots, \mathbf{x}_k\}$ est affinement indépendant si et seulement si le système

$$\sum_{i=0}^{k} \alpha_i \mathbf{x}_i = \mathbf{0}, \qquad \sum_{i=0}^{k} \alpha_i = 0$$

admet pour solution unique $\alpha_0 = \ldots = \alpha_k = 0$.

Démonstration. C'est un exercice facile. ∎

Lemme 2. Soit $A := \{\mathbf{x}_0, \ldots, \mathbf{x}_k\} \subset \mathbb{R}^d$ un ensemble affinement indépendant. Soient $H := \{\mathbf{u} \in \mathbb{R}^{k+1} \mid \sum_{j=0}^{k} u_j = 1\}$ et $\varphi \colon H \to \operatorname{aff} A$ définie par

$$\varphi(\mathbf{u}) = \sum_{j=0}^{k} u_j \mathbf{x}_j.$$

Alors φ est bijective et bicontinue.

Démonstration. Soient $\mathbf{u}, \mathbf{v} \in H$ tels que $\varphi(\mathbf{u}) = \varphi(\mathbf{v})$. Alors,

$$\sum_{j=0}^{k} (u_j - v_j)\mathbf{x}_j = \mathbf{0} \quad \text{et} \quad \sum_{j=0}^{k} (u_j - v_j) = 0,$$

et le lemme précédent montre que $u_j - v_j = 0$ pour tout j, c'est-à-dire, $\mathbf{u} = \mathbf{v}$. Donc φ est injective. Soit maintenant $\mathbf{x}$ quelconque dans $\operatorname{aff} A$. Alors $\mathbf{x}$ est de la forme

$$\mathbf{x} = \mathbf{x}_0 + \sum_{j=1}^{k} \alpha_j (\mathbf{x}_j - \mathbf{x}_0) = \left(1 - \sum_{j=1}^{k} \alpha_j\right) \mathbf{x}_0 + \sum_{j=1}^{k} \alpha_j \mathbf{x}_j. \tag{2.2}$$

Il est clair que $(1 - \sum_{j=1}^{k} \alpha_j, \alpha_1, \ldots, \alpha_k) \in H$, donc $\mathbf{x} \in \varphi(H)$. Outre la surjectivité de φ, l'équation (2.2) donne aussi l'application inverse de φ :

$$\varphi^{-1}\left(\mathbf{x}_0 + \sum_{j=1}^{k} \alpha_j (\mathbf{x}_j - \mathbf{x}_0)\right) = \left(1 - \sum_{j=1}^{k} \alpha_j, \alpha_1, \ldots, \alpha_k\right).$$

La continuité de φ et de φ^{-1} est alors évidente. ∎

Théorème 8. Soit $C \in \mathscr{C}(\mathbb{R}^d)$. Alors $\operatorname{cl} C, \operatorname{ri} C \in \mathscr{C}(\mathbb{R}^d)$. De plus,

$$\operatorname{aff} \operatorname{ri} C = \operatorname{aff} C = \operatorname{aff}(\operatorname{cl} C).$$

Démonstration. La convexité de $\operatorname{cl} C$ résulte immédiatement de la formule

$$\operatorname{cl} C = \bigcap_{\varepsilon > 0} \big(C + B(\mathbf{0}, \varepsilon)\big)$$

et du théorème 4, puisque $C + B(\mathbf{0}, \varepsilon)$ est convexe pour tout ε (voir le théorème 5(ii)). La convexité de $\operatorname{ri} C$ est une conséquence immédiate du corollaire 1. Montrons l'égalité $\operatorname{aff} C = \operatorname{aff} \operatorname{cl} C$. D'une part, on a

$$C \subset \operatorname{cl} C \implies \operatorname{aff} C \subset \operatorname{aff}(\operatorname{cl} C).$$

Et d'autre part, on a

$$C \subset \operatorname{aff} C \implies \operatorname{cl} C \subset \operatorname{cl}(\operatorname{aff} C) = \operatorname{aff} C \implies \operatorname{aff}(\operatorname{cl} C) \subset \operatorname{aff} C.$$

Reste à montrer l'égalité $\operatorname{aff}(\operatorname{ri} C) = \operatorname{aff} C$. L'inclusion $\operatorname{aff}(\operatorname{ri} C) \subset \operatorname{aff} C$ étant évidente, il suffit de montrer l'inclusion opposée. Soit $k := \dim C$. Il existe un ensemble $A = \{\mathbf{x}_0, \ldots, \mathbf{x}_k\} \subset C$ affinement indépendant. On considère alors

$$\Delta := \operatorname{co}\{\mathbf{x}_0, \ldots, \mathbf{x}_k\}.$$

Puisque $\Delta \subset C$ et $\dim \Delta = \dim C$, nous voyons que Δ et C ont la même enveloppe affine. Soient $H := \{\mathbf{u} \in \mathbb{R}^{k+1} \mid \sum_{j=0}^{k} u_j = 1\}$ et $\varphi\colon \mathbb{R}^{k+1} \to \operatorname{aff} C$ l'application linéaire définie par

$$\varphi(\mathbf{u}) = \sum_{j=0}^{d} u_j \mathbf{x}_j.$$

Soient encore

$$\Sigma := \left\{ \mathbf{u} \in \mathbb{R}_+^{k+1} \;\middle|\; \sum_{j=0}^{k} u_j = 1 \right\} \quad \text{et} \quad \Sigma' := \left\{ \mathbf{u} \in (\mathbb{R}_+^*)^{k+1} \;\middle|\; \sum_{j=0}^{k} u_j = 1 \right\}.$$

Remarquons que $\varphi(\Sigma') \subset \varphi(\Sigma) = \Delta \subset C \subset \operatorname{aff} C = \operatorname{aff} A$. Puisque Σ' est ouvert dans H et φ est bicontinue d'après le lemme 2, $\varphi(\Sigma')$ est ouvert dans $\operatorname{aff} C$. Il s'ensuit que $\varphi(\Sigma') \subset \operatorname{ri} C$ de sorte que $\operatorname{aff} \varphi(\Sigma') \subset \operatorname{aff}(\operatorname{ri} C)$. Or,

$$\operatorname{aff} \varphi(\Sigma') = \operatorname{aff}(\operatorname{cl} \varphi(\Sigma')) = \operatorname{aff} \varphi(\operatorname{cl} \Sigma') = \operatorname{aff} \varphi(\Sigma) = \operatorname{aff} \Delta = \operatorname{aff} C,$$

et l'inclusion $\operatorname{aff} C \subset \operatorname{aff}(\operatorname{ri} C)$ est démontrée. ∎

Il résulte immédiatement du théorème 8 que, si $C \in \mathscr{C}(\mathbb{R}^d)$, alors $\operatorname{ri} C$, C et $\operatorname{cl} C$ ont même dimension, et que si $C \in \mathscr{C}(\mathbb{R}^d)$ est non vide, alors $\operatorname{ri} C \neq \emptyset$.

Théorème 9. Soit $C \in \mathscr{C}(\mathbb{R}^d)$. Alors $\operatorname{cl}(\operatorname{ri} C) = \operatorname{cl} C$ et $\operatorname{ri}(\operatorname{cl} C) = \operatorname{ri} C$.

DÉMONSTRATION. Puisque $\operatorname{ri} C \subset C$, $\operatorname{cl} \operatorname{ri} C \subset \operatorname{cl} C$. Réciproquement, soit $\mathbf{y} \in \operatorname{cl} C$. On peut supposer C non vide. D'après le théorème 8, on peut trouver $\mathbf{x} \in \operatorname{ri} C$. Le théorème 7 montre alors que $[\mathbf{x}, \mathbf{y}[\subset \operatorname{ri} C$, et donc que $\mathbf{y} \in \operatorname{cl}(\operatorname{ri} C)$. La première égalité est établie.
Puisque $\operatorname{aff} C = \operatorname{aff}(\operatorname{cl} C)$, l'inclusion $C \subset \operatorname{cl} C$ implique l'inclusion $\operatorname{ri} C \subset \operatorname{ri}(\operatorname{cl} C)$. Réciproquement, soit $\mathbf{z} \in \operatorname{ri}(\operatorname{cl} C)$, et montrons que $\mathbf{z} \in \operatorname{ri} C$. Soit $\mathbf{x}$ un point quelconque de $\operatorname{ri} C$. On peut supposer que $\mathbf{x} \neq \mathbf{z}$, sinon $\mathbf{z} \in \operatorname{ri} C$ trivialement. Soit $\mu > 1$, et soit

$$\mathbf{y} := (1-\mu)\mathbf{x} + \mu \mathbf{z} = \mathbf{z} - (\mu - 1)(\mathbf{x} - \mathbf{z}).$$

Pour μ suffisamment proche de 1, $\mathbf{y} \in \operatorname{ri}(\operatorname{cl} C)$, donc $\mathbf{y} \in \operatorname{cl} C$. Alors, $\mathbf{z} = (1-\lambda)\mathbf{x} + \lambda \mathbf{y}$ pour un certain $\lambda \in]0, 1[$. Puisque $\mathbf{x} \in \operatorname{ri} C$ et $\mathbf{y} \in \operatorname{cl} C$, le théorème 7 montre que $\mathbf{z} \in \operatorname{ri} C$. ∎

Proposition 1. (1) Soient $C_1, C_2 \subset \mathscr{C}(\mathbb{R}^d)$ tels que $\operatorname{ri} C_1 \cap \operatorname{ri} C_2 \neq \emptyset$. Alors

$$\operatorname{cl}(C_1 \cap C_2) = \operatorname{cl} C_1 \cap \operatorname{cl} C_2 \quad \text{et} \quad \operatorname{ri}(C_1 \cap C_2) = \operatorname{ri} C_1 \cap \operatorname{ri} C_2.$$

(2) Soient $C_1 \subset \mathscr{C}(\mathbb{R}^{d_1})$ et $C_2 \subset \mathscr{C}(\mathbb{R}^{d_2})$. Alors

$$\operatorname{aff}(C_1 \times C_2) = \operatorname{aff} C_1 \times \operatorname{aff} C_2 \quad \text{et} \quad \operatorname{ri}(C_1 \times C_2) = \operatorname{ri} C_1 \times \operatorname{ri} C_2.$$

DÉMONSTRATION. En TD. ∎

Théorème 10. Soit $A \colon \mathbb{R}^n \to \mathbb{R}^m$ une application linéaire.

(1) Soit $C \in \mathscr{C}(\mathbb{R}^n)$. Alors

$$\operatorname{ri}(AC) = A(\operatorname{ri} C) \quad \text{et} \quad \operatorname{cl}(AC) \supset A(\operatorname{cl} C).$$

(2) Soit $C \in \mathscr{C}(\mathbb{R}^m)$ tel que $A^{-1}(\operatorname{ri} C) \neq \emptyset$. Alors

$$\operatorname{ri}(A^{-1}(C)) = A^{-1}(\operatorname{ri} C) \quad \text{et} \quad \operatorname{cl}(A^{-1}(C)) = A^{-1}(\operatorname{cl} C).$$

DÉMONSTRATION.

(1) Montrons tout d'abord que $A(\operatorname{cl} C) \subset \operatorname{cl}(AC)$. Si $A = 0$, c'est évident. Supposons alors que $A \neq 0$ (donc que $\|A\| > 0$). En appliquant A à l'égalité
$$\operatorname{cl} C = \bigcap_{\varepsilon > 0} \big(C + B(\mathbf{0}, \varepsilon)\big),$$
nous obtenons :
$$A \operatorname{cl} C = \bigcap_{\varepsilon > 0} \big(AC + AB(\mathbf{0}, \varepsilon)\big),$$

où
$$AB(\mathbf{0},\varepsilon) = \{A\mathbf{x} \mid \|\mathbf{x}\| < \varepsilon\} \subset B(\mathbf{0}, \|A\|\,\varepsilon).$$
Par conséquent,
$$A\,\mathrm{cl}\,C \subset \bigcap_{\varepsilon>0} \big(AC + B(\mathbf{0}, \|A\|\,\varepsilon)\big) = \bigcap_{\varepsilon'>0} \big(AC + B(\mathbf{0}, \varepsilon')\big) = \mathrm{cl}(AC).$$

Afin de montrer l'égalité du point (1), montrons tout d'abord l'inclusion $\mathrm{ri}\,AC \subset A(\mathrm{ri}\,C)$. D'après le début de la démonstration,
$$\mathrm{cl}\,A(\mathrm{ri}\,C) \supset A\big(\mathrm{cl}(\mathrm{ri}\,C)\big) = A(\mathrm{cl}\,C) \supset AC \supset A\,\mathrm{ri}\,C,$$
où l'égalité résulte du théorème 9. En prenant la fermeture, on obtient que $\mathrm{cl}\,AC = \mathrm{cl}\,A(\mathrm{ri}\,C)$. Or il est facile de voir, en utilisant à nouveau le théorème 9, que deux convexes ont la même fermeture si et seulement si ils ont même intérieur relatif. Donc $\mathrm{ri}\,AC = \mathrm{ri}\,A(\mathrm{ri}\,C)$, ce qui montre que $\mathrm{ri}\,AC \subset A(\mathrm{ri}\,C)$. Réciproquement, supposons que $\mathbf{z} \in A(\mathrm{ri}\,C)$, et montrons que $\mathbf{z} \in \mathrm{ri}\,AC$. Soit $\mathbf{x} \in AC$. Alors
$$\mathbf{x} = A\mathbf{x}' \quad \text{avec} \quad \mathbf{x}' \in C \qquad \text{et} \qquad \mathbf{z} = A\mathbf{z}' \quad \text{avec} \quad \mathbf{z}' \in \mathrm{ri}\,C.$$
D'après le corollaire 2, il existe $\mu > 1$ tel que
$$(1-\mu)\mathbf{x}' + \mu\mathbf{z}' \in C, \quad \text{donc} \quad A\big((1-\mu)\mathbf{x}' + \mu\mathbf{z}'\big) \in AC.$$
Autrement dit, quel que soit $\mathbf{x} \in AC$, il existe $\mu > 1$ tel que $(1-\mu)\mathbf{x}+\mu\mathbf{z} \in AC$. Le corollaire 2 montre alors que $\mathbf{z} \in \mathrm{ri}\,AC$.

(2) Le *graphe* de A est l'ensemble $\mathrm{gr}\,A := \{(\mathbf{x},\mathbf{y}) \in \mathbb{R}^{n+m} \mid \mathbf{y} = A\mathbf{x}\}$. C'est évidemment un sous-espace vectoriel de $\mathbb{R}^{n+m}$. On considère aussi la projection canonique
$$\begin{array}{rccc} P\colon & \mathbb{R}^{n+m} & \longrightarrow & \mathbb{R}^n \\ & (\mathbf{x},\mathbf{y}) & \longmapsto & \mathbf{x}, \end{array}$$
et on peut écrire $A^{-1}C = P(\mathrm{gr}\,A \cap (\mathbb{R}^n \times C))$. On a alors :
$$\begin{aligned} \mathrm{ri}(A^{-1}C) &= \mathrm{ri}\,P(\mathrm{gr}\,A \cap (\mathbb{R}^n \times C)) \\ &= P\big(\mathrm{ri}(\mathrm{gr}\,A \cap (\mathbb{R}^n \times C))\big) \\ &= P\big(\mathrm{ri}(\mathrm{gr}\,A) \cap \mathrm{ri}(\mathbb{R}^n \times C)\big) \\ &= P\big(\mathrm{gr}\,A \cap (\mathbb{R}^n \times \mathrm{ri}\,C)\big) \\ &= A^{-1}(\mathrm{ri}\,C), \end{aligned}$$
où la seconde égalité résulte du point (1), et la troisième de la proposition 1(1) (puisque $\mathrm{ri}(\mathrm{gr}\,A) = \mathrm{gr}\,A$ rencontre $\mathrm{ri}(\mathbb{R}^n \times C) = \mathbb{R}^n \times \mathrm{ri}\,C$, $\mathrm{ri}\,C$

étant non vide). Enfin,

$$\begin{aligned}
\mathrm{cl}(A^{-1}C) &= \mathrm{cl}\,P(\mathrm{gr}\,A \cap (\mathbb{R}^n \times C)) \\
&\supset P\big(\mathrm{cl}(\mathrm{gr}\,A \cap (\mathbb{R}^n \times C))\big) \\
&= P\big(\mathrm{cl}(\mathrm{gr}\,A) \cap \mathrm{cl}(\mathbb{R}^n \times C)\big) \\
&= P\big(\mathrm{gr}\,A \cap (\mathbb{R}^n \times \mathrm{cl}\,C)\big) \\
&= A^{-1}(\mathrm{cl}\,C),
\end{aligned}$$

où les seconde et troisième lignes résultent encore du point (1) et de la proposition 1(1). Réciproquement, $C \subset \mathrm{cl}\,C$ entraîne $A^{-1}C \subset A^{-1}(\mathrm{cl}\,C)$, le dernier ensemble étant bien sûr fermé (A est continue). On en déduit que $\mathrm{cl}\,A^{-1}C \subset A^{-1}(\mathrm{cl}\,C)$. ■

Nous donnons maintenant un exemple pour lequel l'inclusion $A\,\mathrm{cl}\,C \subset \mathrm{cl}\,AC$ est stricte. Soient $C := \{(x,y) \in \mathbb{R}^2 \,|\, x > 0,\ y \geq 1/x\}$ et

$$\begin{array}{rccc}
A\colon & \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\
& (x,y) & \longmapsto & x.
\end{array}$$

Il est facile de voir que C est fermé, que $AC =]0,\infty[$, de sorte que $\mathrm{cl}\,AC = [0,\infty[$, alors que $A\,\mathrm{cl}\,C = AC =]0,\infty[$.

Corollaire 3. Soient $C_1, C_2 \in \mathscr{C}(\mathbb{R}^d)$ et $\alpha, \beta \in \mathbb{R}$. Alors l'ensemble convexe $\alpha C_1 + \beta C_2$ satisfait :

$$\mathrm{ri}(\alpha C_1 + \beta C_2) = \alpha\,\mathrm{ri}\,C_1 + \beta\,\mathrm{ri}\,C_2 \quad \text{et} \quad \mathrm{cl}(\alpha C_1 + \beta C_2) \supset \alpha\,\mathrm{cl}\,C_1 + \beta\,\mathrm{cl}\,C_2.$$

Démonstration. On considère l'application linéaire

$$\begin{array}{rccc}
A\colon & \mathbb{R}^d \times \mathbb{R}^d & \longrightarrow & \mathbb{R}^d \\
& (\mathbf{x},\mathbf{y}) & \longmapsto & \alpha\mathbf{x} + \beta\mathbf{y}.
\end{array}$$

Alors $A(C_1 \times C_2) = \alpha C_1 + \beta C_2$. D'une part,

$$\begin{aligned}
\mathrm{ri}(\alpha C_1 + \beta C_2) &= \mathrm{ri}\,A(C_1 \times C_2) \\
&= A\,\mathrm{ri}(C_1 \times C_2) \\
&= A(\mathrm{ri}\,C_1 \times \mathrm{ri}\,C_2) \\
&= \alpha\,\mathrm{ri}\,C_1 + \beta\,\mathrm{ri}\,C_2,
\end{aligned}$$

où les deuxième et troisième égalités résultent du théorème 10(1) et de la proposition 1(2), respectivement. D'autre part,

$$\begin{aligned}
\mathrm{cl}(\alpha C_1 + \beta C_2) &= \mathrm{cl}\,A(C_1 \times C_2) \\
&\supset A\,\mathrm{cl}(C_1 \times C_2) \\
&= A(\mathrm{cl}\,C_1 \times \mathrm{cl}\,C_2) \\
&= \alpha\,\mathrm{cl}\,C_1 + \beta\,\mathrm{cl}\,C_2,
\end{aligned}$$

où la deuxième égalité résulte à nouveau du théorème 10(1). ∎

Nous donnons maintenant un exemple pour lequel l'inclusion $\alpha \operatorname{cl} C_1 + \beta \operatorname{cl} C_2 \subset \operatorname{cl}(\alpha C_1 + \beta C_2)$ est stricte. Soient

$$C_1 := \left\{(x,y) \in \mathbb{R}^2 \;\middle|\; x > 0,\ y \geq 1/x\right\}, \quad C_2 := \{0\} \times \left]-\infty, 0\right]$$

et $\alpha = \beta = 1$. Puisque C_1 et C_2 sont fermés, on a :

$$\operatorname{cl} C_1 + \operatorname{cl} C_2 = C_1 + C_2 = \left]0, \infty\right[\times \mathbb{R} \subsetneq \left[0, \infty\right[\times \mathbb{R} = \operatorname{cl}(C_1 + C_2).$$

2.2 Séparation

Définitions 1.

(1) On appelle *hyperplan* de $\mathbb{R}^d$ tout sous-espace affine de $\mathbb{R}^d$ de dimension $d-1$, autrement dit, tout ensemble de la forme

$$\left\{\mathbf{x} \in \mathbb{R}^d \;\middle|\; \langle \mathbf{b}, \mathbf{x} \rangle = \beta\right\},$$

avec $\mathbf{b} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ et $\beta \in \mathbb{R}$. À un hyperplan $H = \{\mathbf{x} \in \mathbb{R}^d \mid \langle \mathbf{b}, \mathbf{x} \rangle = \beta\}$ correspondent deux *demi-espaces ouverts*

$$\left\{\mathbf{x} \in \mathbb{R}^d \;\middle|\; \langle \mathbf{b}, \mathbf{x} \rangle < \beta\right\} \quad \text{et} \quad \left\{\mathbf{x} \in \mathbb{R}^d \;\middle|\; \langle \mathbf{b}, \mathbf{x} \rangle > \beta\right\},$$

et deux *demi-espaces fermés*

$$\left\{\mathbf{x} \in \mathbb{R}^d \;\middle|\; \langle \mathbf{b}, \mathbf{x} \rangle \leq \beta\right\} \quad \text{et} \quad \left\{\mathbf{x} \in \mathbb{R}^d \;\middle|\; \langle \mathbf{b}, \mathbf{x} \rangle \geq \beta\right\}.$$

(2) On dit que $E_1, E_2 \in \mathscr{P}(\mathbb{R}^d)$ sont séparés s'il existe un hyperplan H tel que E_1 soit contenu dans l'un des demi-espaces fermés associés à H et E_2 dans l'autre, autrement dit, s'il existe $\mathbf{b} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ tel que

$$\inf_{\mathbf{x} \in E_1} \langle \mathbf{b}, \mathbf{x} \rangle \geq \sup_{\mathbf{x} \in E_2} \langle \mathbf{b}, \mathbf{x} \rangle.$$

On remarque que tout sous-ensemble contenu dans un hyperplan est séparé de lui-même par cet hyperplan.

(3) On dit que $E_1, E_2 \in \mathscr{P}(\mathbb{R}^d)$ sont séparés proprement s'il sont séparés par un hyperplan tel qu'au moins un des deux sous-ensembles n'est pas contenu dans l'hyperplan séparateur, autrement dit, s'il existe $\mathbf{b} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ tel que

$$\inf_{\mathbf{x} \in E_1} \langle \mathbf{b}, \mathbf{x} \rangle \geq \sup_{\mathbf{x} \in E_2} \langle \mathbf{b}, \mathbf{x} \rangle \quad \text{et} \quad \sup_{\mathbf{x} \in E_1} \langle \mathbf{b}, \mathbf{x} \rangle > \inf_{\mathbf{x} \in E_2} \langle \mathbf{b}, \mathbf{x} \rangle.$$

On remarque qu'il est possible d'avoir deux sous-ensembles E_1, E_2 tels que $E_1 \subset E_2$ qui sont séparés proprement.

(4) On dit que $E_1, E_2 \in \mathscr{P}(\mathbb{R}^d)$ sont séparés fortement s'il existe $\mathbf{b} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ tel que

$$\inf_{\mathbf{x}\in E_1} \langle \mathbf{b}, \mathbf{x}\rangle > \sup_{\mathbf{x}\in E_2} \langle \mathbf{b}, \mathbf{x}\rangle .$$

Il est facile de voir que E_1 et E_2 sont séparés fortement si et seulement si l'on peut trouver $\varepsilon > 0$ tel que $E_1 + B(\mathbf{0}, \varepsilon)$ et $E_2 + B(\mathbf{0}, \varepsilon)$ soient séparés.

Lemme 3. Soient $E \subset \mathbb{R}^d$ et $\mathbf{x} \in \mathbb{R}^d$. Si E et $\{\mathbf{x}\}$ sont séparés, alors $\mathbf{x} \notin \operatorname{int} E$.

Démonstration. L'hypothèse de séparation s'écrit :

$$\exists \mathbf{b} \in \mathbb{R}^d \setminus \{\mathbf{0}\} \colon \sup_{\mathbf{y}\in E} \langle \mathbf{b}, \mathbf{y}\rangle \leq \langle \mathbf{b}, \mathbf{x}\rangle \tag{2.3}$$

Supposons, en vue d'obtenir une contradiction, que $\mathbf{x} + \varepsilon\mathbf{b} \in E$ pour un certain $\varepsilon > 0$. On aurait alors

$$\langle \mathbf{b}, \mathbf{x} + \varepsilon\mathbf{b}\rangle = \langle \mathbf{b}, \mathbf{x}\rangle + \varepsilon \|\mathbf{b}\|^2 > \langle \mathbf{b}, \mathbf{x}\rangle ,$$

en contradiction avec la condition (2.3). On en déduit que $\mathbf{x} \notin \operatorname{int} E$. ∎

Nous rappelons maintenant, sans démonstration, le théorème de projection sur un convexe fermé dans un espace de Hilbert.

Théorème 11. Soit $(F, \langle\cdot,\cdot\rangle)$ un espace de Hilbert réel. On note $\|\cdot\|$ la norme associée à $\langle\cdot,\cdot\rangle$. Soient $C \subset F$ un sous-ensemble convexe fermé et $x \in F$. Alors, il existe un unique $x_0 \in C$ tel que

$$\|x - x_0\| = \inf_{y\in C} \|x - y\| .$$

Le vecteur x_0 s'appelle le projeté de x sur C, et l'on écrit : $x_0 = \operatorname{proj}_C(x)$. De plus, il y a équivalence entre

(a) $x_0 = \operatorname{proj}_C(x)$;

(b) pour tout $y \in C$, $\langle x - x_0, y - x_0\rangle \leq 0$.

Théorème 12. Soient $C \in \mathscr{C}(\mathbb{R}^d)$ et $\mathbf{x} \in \mathbb{R}^d$. Il y a équivalence entre

(a) $\mathbf{x} \notin \operatorname{cl} C$;

(b) C et $\{\mathbf{x}\}$ sont séparés fortement, autrement dit, il existe $\mathbf{b} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ tel que

$$\sup_{\mathbf{y}\in C} \langle \mathbf{b}, \mathbf{y}\rangle < \langle \mathbf{b}, \mathbf{x}\rangle .$$

Démonstration. Montrons tout d'abord que (a) implique (b). Soit $\mathbf{x}_0 := \operatorname{proj}_{\operatorname{cl} C}(\mathbf{x})$ et soit $\mathbf{b} := \mathbf{x} - \mathbf{x}_0 \neq \mathbf{0}$. D'après le théorème 11, pour tout $\mathbf{y} \in \operatorname{cl} C$,

$$\langle \mathbf{x} - \mathbf{x}_0, \mathbf{y} - \mathbf{x}_0\rangle = \langle \mathbf{b}, \mathbf{y} - \mathbf{x} + \mathbf{b}\rangle \leq 0.$$

Donc, pour tout $\mathbf{y} \in \operatorname{cl} C$, $\langle \mathbf{b}, \mathbf{y} \rangle \leq \langle \mathbf{b}, \mathbf{x} \rangle - \|\mathbf{b}\|^2$. Comme $\|\mathbf{b}\| > 0$, on obtient :

$$\sup_{\mathbf{y} \in \operatorname{cl} C} \langle \mathbf{b}, \mathbf{y} \rangle < \langle \mathbf{b}, \mathbf{x} \rangle .$$

On conclut en remarquant que $\sup_{\mathbf{y} \in \operatorname{cl} C} \langle \mathbf{b}, \mathbf{y} \rangle = \sup_{\mathbf{y} \in C} \langle \mathbf{b}, \mathbf{y} \rangle$, par continuité de $\langle \mathbf{b}, \cdot \rangle$. Réciproquement, supposons que $\mathbf{x} \in \operatorname{cl} C$. Alors, pour tout $\mathbf{b} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$,

$$\sup_{\mathbf{y} \in C} \langle \mathbf{b}, \mathbf{y} \rangle = \sup_{\mathbf{y} \in \operatorname{cl} C} \langle \mathbf{b}, \mathbf{y} \rangle \geq \langle \mathbf{b}, \mathbf{x} \rangle . \blacksquare$$

Corollaire 4. Tout sous-ensemble convexe fermé de $\mathbb{R}^d$ est l'intersection des demi-espaces fermés qui le contiennent.

Démonstration. Si $C = \emptyset$ ou $C = \mathbb{R}^d$, l'assertion est triviale (l'intersection d'une famille vide de sous-ensembles de $\mathbb{R}^d$ est, par convention, $\mathbb{R}^d$ tout entier). Soit donc C un convexe fermé non vide et différent de $\mathbb{R}^d$. Alors, $\mathbf{x} \in C^c$ si et seulement si $\{\mathbf{x}\}$ et C sont séparés fortement, et cette derniére condition équivaut à

> $\mathbf{x}$ n'appartient pas à l'intersection des demi-espaces fermés contenant C. $\blacksquare$

Corollaire 5. Soit $C \in \mathscr{C}(\mathbb{R}^d)$. Alors $\operatorname{cl} C$ est l'intersection des demi-espaces fermés qui contiennent C.

Démonstration. D'aprés le corollaire précédent, il suffit de voir que la famille des demi-espaces fermés qui contiennent C et celle des demi-espaces fermés qui contiennent $\operatorname{cl} C$ sont les mêmes. Ceci résulte immédiatement du fait que si $C \subset D$ et D est fermé, alors $\operatorname{cl} C \subset D$. $\blacksquare$

Théorème 13. Soit $C \in \mathscr{C}(\mathbb{R}^d)$ tel que $\operatorname{int} C \neq \emptyset$. Alors il y a équivalence entre

(a) $\mathbf{x} \notin \operatorname{int} C$;

(b) C et $\{\mathbf{x}\}$ sont séparés proprement, c'est à dire, il existe $\mathbf{b} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ et $\mathbf{y}_0 \in C$ tels que

$$\sup_{\mathbf{y} \in C} \langle \mathbf{b}, \mathbf{y} \rangle \leq \langle \mathbf{b}, \mathbf{x} \rangle \quad \text{et} \quad \langle \mathbf{b}, \mathbf{y}_0 \rangle < \langle \mathbf{b}, \mathbf{x} \rangle .$$

Démonstration. Le fait que (b) implique (a) résulte immédiatement du lemme 3. Montrons que (a) implique (b). D'après le théorème 9, $\operatorname{int} C = \operatorname{int} \operatorname{cl} C$, donc (a) s'écrit aussi $\mathbf{x} \notin \operatorname{int} \operatorname{cl} C$, soit encore $\mathbf{x} \in \operatorname{cl}(\operatorname{cl} C)^c$. Donc il existe une suite $(\mathbf{x}_k)_{k \in \mathbb{N}^*} \subset (\operatorname{cl} C)^c$ qui converge vers $\mathbf{x}$. D'après le théorème 12, pour chaque k, il existe $\mathbf{b}_k \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ tel que

$$\sup_{\mathbf{y} \in C} \langle \mathbf{b}_k, \mathbf{y} \rangle < \langle \mathbf{b}_k, \mathbf{x}_k \rangle. \tag{2.4}$$

En divisant $\mathbf{b}_k$ par sa norme si nécessaire, on peut supposer que $\|\mathbf{b}_k\| = 1$ pour tout k. Comme la sphère unité de $\mathbb{R}^d$ est compacte, on peut aussi supposer (en prenant une sous-suite si nécessaire) que la suite $(\mathbf{b}_k)_{k\in\mathbb{N}^*}$ converge vers un point $\mathbf{b}$ (de norme 1). On déduit de l'inégalité (2.4) que, pour tout $\mathbf{y} \in C$, $\langle \mathbf{b}, \mathbf{y}\rangle \le \langle \mathbf{b}, \mathbf{x}\rangle$ et donc, que

$$\sup_{\mathbf{y}\in C} \langle \mathbf{b}, \mathbf{y}\rangle \le \langle \mathbf{b}, \mathbf{x}\rangle .$$

Montrons par l'absurde qu'il existe $\mathbf{y}_0 \in C$ tel que $\langle \mathbf{b}, \mathbf{y}_0\rangle < \langle \mathbf{b}, \mathbf{x}\rangle$. Supposons que

$$\inf_{\mathbf{y}\in C} \langle \mathbf{b}, \mathbf{y}\rangle = \langle \mathbf{b}, \mathbf{x}\rangle .$$

Alors $C \subset H := \{\mathbf{y} \in \mathbb{R}^d \,|\, \langle \mathbf{b}, \mathbf{y}\rangle = \langle \mathbf{b}, \mathbf{x}\rangle\}$, de sorte que $\operatorname{aff} C \subset H$. Ceci est absurde puisque $\operatorname{aff} C = \mathbb{R}^d$. ∎

Proposition 2. Soit C un convexe de $\mathbb{R}^2$ relativement ouvert ne contenant pas l'origine. Alors il existe une droite vectorielle Δ telle que $\Delta \cap C = \emptyset$.

Démonstration. Si $\dim C = 0$, c'est-à-dire si $C = \emptyset$ ou C est un singleton, le résultat est trivial. Supposons que $\dim C = 1$. Alors $\operatorname{aff} C$ est une droite. Supposons tout d'abord que $\operatorname{aff} C$ contient l'origine, autrement dit, que $\operatorname{aff} C$ est une droite vectorielle. Si Δ une autre droite vectorielle, alors $\Delta \cap C = \emptyset$ car, sinon, on aurait $\Delta = \operatorname{aff} C$, Δ et $\operatorname{aff} C$ ayant en commun $\mathbf{0}$ et un point $\mathbf{x} \neq \mathbf{0}$. Supposons maintenant que $\operatorname{aff} C$ n'est pas une droite vectorielle. Alors, le choix $\Delta = \operatorname{aff} C - \operatorname{aff} C$ convient, puisque $\operatorname{aff} C - \operatorname{aff} C$ est la droite vectorielle parallèle à $\operatorname{aff} C$.
Reste à considérer le cas où $\dim C = 2$. Alors C est un ouvert qui ne contient pas l'origine. D'après le théorème 13, C et $\{\mathbf{0}\}$ sont séparés proprement par une droite Δ'. Soit D le demi-plan fermé associé à Δ' qui contient C. Puisque C est ouvert, $C \subset D$ et D est fermé, il est clair que $C \subset \operatorname{int} D$, ce qui entraîne que $\Delta' \cap C = \emptyset$. Il est alors évident que $\Delta := \Delta' - \Delta'$ convient. ∎

Théorème 14. Soit $C \in \mathscr{C}(\mathbb{R}^d)$. Il y a équivalence entre

(a) $\mathbf{x} \notin \operatorname{ri} C$;

(b) C et $\{\mathbf{x}\}$ sont séparés proprement.

Démonstration. On peut supposer, en appliquant une translation si nécessaire, que $\mathbf{x} = \mathbf{0}$.
Supposons (a). Nous allons montrer qu'il existe alors un sous-espace vectoriel de dimension $d-1$ qui ne rencontre pas $\operatorname{ri} C$, ce qui établira (b). Considérons la propriété

$$(\mathscr{P}_k) \quad \exists V_k \in \mathscr{V}(\mathbb{R}^d) :\ \dim V_k = k \text{ et } V_k \cap \operatorname{ri} C = \emptyset.$$

La condition (a) n'est autre que $\mathscr{P}_0$). Supposons que $(\mathscr{P}_k)$ soit satisfaite pour $k < d-1$, et montrons que $(\mathscr{P}_{k+1})$ est aussi satisfaite. Puisque $V_k \cap \operatorname{ri} C = \emptyset$ et $\mathbf{0} \in V_k$,

$$\mathbf{0} \notin \operatorname{ri} C - V_k \supset \operatorname{ri} C,$$

et puisque $k \leq d-2$, $V_k^\perp$ contient un sous-espace vectoriel W de dimension 2. Soit alors

$$C' := W \cap (\operatorname{ri} C - V_k).$$

L'ensemble C' ne contient pas $\mathbf{0}$. Montrons que C' est relativement ouvert. Si C' est vide, c'est trivial ($\operatorname{ri} \emptyset = \emptyset$). Si C' est non vide, alors

$$\emptyset \neq W \cap (\operatorname{ri} C - V_k) = \operatorname{ri} W \cap \operatorname{ri}(\operatorname{ri} C - V_k),$$

où l'égalité résulte du fait évident que $W = \operatorname{ri} W$ et du corollaire 3. La proposition 1(1) implique alors que

$$\begin{aligned} \operatorname{ri} C' &= \operatorname{ri}\big(W \cap (\operatorname{ri} C - V_k)\big) \\ &= \operatorname{ri} W \cap \operatorname{ri}(\operatorname{ri} C - V_k) \\ &= W \cap (\operatorname{ri} C - V_k) \\ &= C'. \end{aligned}$$

D'après la proposition 2, il existe une droite vectorielle $\Delta \subset W$ telle que $\Delta \cap C' = \emptyset$. Puisque $W \subset V_k^\perp$, $V_k + \Delta$ est un sous-espace vectoriel de dimension $k+1$. D'autre part, $V_k + \Delta$ ne rencontre pas $\operatorname{ri} C$. Supposons en effet qu'il existe $\mathbf{x}_0 \in (V_k + \Delta) \cap \operatorname{ri} C$. Alors $\mathbf{x}_0 = \mathbf{x} + \mathbf{y} \in \operatorname{ri} C$, avec $\mathbf{x} \in V_k$ et $\mathbf{y} \in \Delta$. Mais alors $\mathbf{y} = \mathbf{x}_0 - \mathbf{x} \in \operatorname{ri} C - V_k$, et comme $\mathbf{y} \in \Delta \subset W$, $\mathbf{y} \in C'$, donc $\mathbf{y} \in \Delta \cap C'$. Mais ceci est absurde, puisque $\Delta \cap C' = \emptyset$. On en déduit que $(\mathscr{P}_{k+1})$ est satisfaite.
Réciproquement, supposons que C est $\{\mathbf{0}\}$ soient séparés proprement. Alors il existe un demi-espace fermé D tel que

$$C \subset D, \quad \mathbf{0} \notin \operatorname{int} D \quad \text{et} \quad \operatorname{int} D \cap C \neq \emptyset.$$

Remarquons que $\operatorname{int} D = \operatorname{ri} D$. Soit H la frontière de D (H est l'hyperplan séparateur). Montrons par l'absurde que

$$\operatorname{ri} C \cap \operatorname{ri} D \neq \emptyset. \tag{2.5}$$

Puisque $\operatorname{ri} C \subset D$, la condition $\operatorname{ri} C \cap \operatorname{ri} D = \emptyset$ impliquerait la condition $\operatorname{ri} C \subset H$ et donc, puisque H est fermé, la condition $\operatorname{cl} \operatorname{ri} C = \operatorname{cl} C \subset H$, en contradiction avec le caractère propre de la séparation. Montrons finalement que $\mathbf{0} \notin \operatorname{ri} C$. En vertu de la condition (2.5) ci-dessus, la proposition 1(1) implique que

$$\operatorname{ri} C \cap \operatorname{ri} D = \operatorname{ri}(C \cap D) = \operatorname{ri} C,$$

et donc que $\operatorname{ri} C \subset \operatorname{ri} D = \operatorname{int} D$. Comme $\mathbf{0} \notin \operatorname{int} D$ il s'ensuit que $\mathbf{0} \notin \operatorname{ri} C$. ∎

Corollaire 6. Soient $C_1, C_2 \in \mathscr{C}(\mathbb{R}^d)$. Il y a équivalence entre

(a) $\operatorname{ri} C_1 \cap \operatorname{ri} C_2 = \emptyset$;

(b) C_1 et C_2 sont séparés proprement.

Démonstration. Soit $C := C_1 - C_2$. D'après le corollaire 3, $\operatorname{ri} C = \operatorname{ri} C_1 - \operatorname{ri} C_2$. Donc

$$\text{(a)} \iff \mathbf{0} \notin \operatorname{ri} C_1 - \operatorname{ri} C_2 \iff \mathbf{0} \notin \operatorname{ri} C.$$

D'autre part, pour tout $\mathbf{b} \in \mathbb{R}^d$,

$$\inf_{\mathbf{x} \in C} \langle \mathbf{b}, \mathbf{x} \rangle = \inf_{\mathbf{x}_1 \in C_1} \langle \mathbf{b}, \mathbf{x}_1 \rangle - \sup_{\mathbf{x}_2 \in C_2} \langle \mathbf{b}, \mathbf{x}_2 \rangle$$

et

$$\sup_{\mathbf{x} \in C} \langle \mathbf{b}, \mathbf{x} \rangle = \sup_{\mathbf{x}_1 \in C_1} \langle \mathbf{b}, \mathbf{x}_1 \rangle - \inf_{\mathbf{x}_2 \in C_2} \langle \mathbf{b}, \mathbf{x}_2 \rangle .$$

Donc

$$\begin{aligned}
\text{(b)} \quad &\iff \quad \begin{cases} \inf\limits_{\mathbf{x}_1 \in C_1} \langle \mathbf{b}, \mathbf{x}_1 \rangle & \geq & \sup\limits_{\mathbf{x}_2 \in C_2} \langle \mathbf{b}, \mathbf{x}_2 \rangle \\ \sup\limits_{\mathbf{x}_1 \in C_1} \langle \mathbf{b}, \mathbf{x}_1 \rangle & > & \inf\limits_{\mathbf{x}_2 \in C_2} \langle \mathbf{b}, \mathbf{x}_2 \rangle \end{cases} \\
&\iff \quad \inf_{\mathbf{x} \in C} \langle \mathbf{b}, \mathbf{x} \rangle \geq 0 \quad \text{et} \quad \sup_{\mathbf{x} \in C} \langle \mathbf{b}, \mathbf{x} \rangle > 0 \\
&\iff \quad \{\mathbf{0}\} \text{ et } C \text{ sont séparés proprement.} \blacksquare
\end{aligned}$$

Chapitre 3

Cônes

3.1 Définitions et propriétés élémentaires

Un sous-ensemble K de $\mathbb{R}^d$ est un cône si $\mathbb{R}_+^* K \subset K$, c'est-à-dire, si K est stable par multiplication par les réels strictement positifs. Un cône est donc un ensemble de demi-droites d'extrémité l'origine, contenant ou ne contenant pas l'origine. Il est faile de voir que l'intersection d'une famille quelconque de cônes [resp. cônes convexes, cônes convexes fermés] est un cône [resp. cône convexe, cône convexe fermé].

Exemple 1. Soit $C \subset \mathbb{R}^d$ un ensemble convexe fermé. On vérifie facilement que, pour tout $\mathbf{x} \in C$, l'ensemble

$$N_C(\mathbf{x}) := \left\{\mathbf{y} \in \mathbb{R}^d \;\middle|\; \forall \mathbf{z} \in C, \langle \mathbf{y}, \mathbf{z} - \mathbf{x}\rangle \leq 0\right\}$$

est un cône convexe fermé. On l'appelle le *cône normal à C en* $\mathbf{x}$. Le théorème 11 montre que, pour $\mathbf{x} \in \mathbb{R}^d$ et $\bar{\mathbf{x}} \in C$, il y a équivalence entre

(a) $\bar{\mathbf{x}} = \mathrm{proj}_C(\mathbf{x})$;

(b) $\mathbf{x} - \bar{\mathbf{x}} \in N_C(\bar{\mathbf{x}})$.

De plus, on peut montrer que

(1) si $\mathbf{x} \in \operatorname{int} C$, $N_C(\mathbf{x}) = \{\mathbf{0}\}$;

(2) si $C = E$ est un sous-espace vectoriel, alors $N_E(\mathbf{x}) = E^\perp$ pour tout $\mathbf{x} \in E$;

(3) si C est le demi-espace fermé $\{\mathbf{x} \in \mathbb{R}^d \,|\, \langle \mathbf{b}, \mathbf{x}\rangle \leq \beta\}$ (où $\mathbf{b} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ et $\beta \in \mathbb{R}$), alors $N_C(\mathbf{x}) = \mathbb{R}_+\{\mathbf{b}\}$ pour tout $\mathbf{x} \in H := \{\mathbf{y} \in \mathbb{R}^d \,|\, \langle \mathbf{b}, \mathbf{y}\rangle = \beta\}$. ∎

Proposition 3. Soit $K \in \mathscr{P}(\mathbb{R}^d)$. Il y a équivalence entre

(a) K est un cône convexe ;

(b) $K + K \subset K$ et $\mathbb{R}_+^* K \subset K$;

(c) $\{\lambda_1 \mathbf{x}_1 + \cdots + \lambda_q \mathbf{x}_q \,|\, q \in \mathbb{N}^*, \lambda_j \in \mathbb{R}_+^*, \mathbf{x}_j \in K\} \subset K$;

(d) $\{\lambda_1 \mathbf{x}_1 + \cdots + \lambda_q \mathbf{x}_q \,|\, q \in \mathbb{N}^*, \lambda_j \in \mathbb{R}_+^*, \mathbf{x}_j \in K\} = K$.

Démonstration. L'équivalence entre (b), (c) et (d) étant évidente, montrons l'équivalence entre (a) et (b). Supposons que K est un cône convexe, et montrons que K est stable par addition. Si $\mathbf{x}, \mathbf{y} \in K$, alors

$$\mathbf{x} + \mathbf{y} = 2\left(\frac{\mathbf{x}}{2} + \frac{\mathbf{y}}{2}\right) \in K,$$

puisque K est convexe et $\mathbb{R}_+^* K \subset K$. Réciproquement, supposons (b), et montrons que K est convexe. Si $\mathbf{x}, \mathbf{y} \in K$ et $\lambda \in]0,1[$, alors $(1-\lambda)\mathbf{x}$ et $\lambda\mathbf{y}$ puisque $\mathbb{R}_+^* K \subset K$, et la stabilité par addition implique alors que

$$(1-\lambda)\mathbf{x} + \lambda\mathbf{y} \in K. \blacksquare$$

Proposition 4. (1) Soient $E \in \mathscr{P}(\mathbb{R}^d)$ et

$$K := \left\{\lambda_1\mathbf{x}_1 + \cdots + \lambda_q\mathbf{x}_q \;\middle|\; q \in \mathbb{N}^*, \lambda_j \in \mathbb{R}_+^*, \mathbf{x}_j \in E\right\}.$$

Alors K est le plus petit cône convexe contenant E.

(2) Soient $C \in \mathscr{C}(\mathbb{R}^d)$ et $K := \{\lambda\mathbf{x} \,|\, \lambda \in \mathbb{R}_+^*, \mathbf{x} \in C\}$. Alors K est le plus petit cône convexe contenant C.

Démonstration.

(1) Il est facile de voir que K est un cône convexe contenant E. Soit K' un autre un cône convexe contenant E, et montrons que $K \subset K'$. Comme $E \subset K'$,

$$K \subset \left\{\lambda_1\mathbf{x}_1 + \cdots + \lambda_q\mathbf{x}_q \;\middle|\; q \in \mathbb{N}^*, \lambda_j \in \mathbb{R}_+^*, \mathbf{x}_j \in K'\right\}.$$

La proposition 3 montre que ce dernier ensemble n'est autre que K'.

(2) Il est clair que K est un cône. Montrons que K est convexe. Soient $\mathbf{x}, \mathbf{y} \in K$ et $\alpha \in]0,1[$. Alors $\mathbf{x} = \lambda_1\mathbf{x}_0$ avec $\lambda_1 > 0$ et $\mathbf{x}_0 \in C$, et $\mathbf{y} = \lambda_2\mathbf{y}_0$ avec $\lambda_2 > 0$ et $\mathbf{y}_0 \in C$. On a :

$$\begin{aligned}(1-\alpha)\mathbf{x} + \alpha\mathbf{y} &= (1-\alpha)\lambda_1\mathbf{x}_0 + \alpha\lambda_2\mathbf{y}_0 \\ &= \big((1-\alpha)\lambda_1 + \alpha\lambda_2\big) \times \\ &\qquad \left[\frac{(1-\alpha)\lambda_1}{(1-\alpha)\lambda_1 + \alpha\lambda_2}\mathbf{x}_0 + \frac{\alpha\lambda_2}{(1-\alpha)\lambda_1 + \alpha\lambda_2}\mathbf{y}_0\right].\end{aligned}$$

L'expression entre crochets est une conbinaison convexe d'éléments de C. On en déduit que $(1-\alpha)\mathbf{x} + \alpha\mathbf{y} \in K$. Soit maintenant K' un autre cône convexe contenant C, et montrons que $K \subset K'$:

$$\begin{aligned}K &= \{\lambda\mathbf{x} \,|\, \lambda \in \mathbb{R}_+^*, \mathbf{x} \in C\} \\ &\subset \{\lambda_1\mathbf{x}_1 + \cdots + \lambda_q\mathbf{x}_q \,|\, q \in \mathbb{N}^*, \lambda_j \in \mathbb{R}_+^*, \mathbf{x}_j \in C\} \\ &\subset \{\lambda_1\mathbf{x}_1 + \cdots + \lambda_q\mathbf{x}_q \,|\, q \in \mathbb{N}^*, \lambda_j \in \mathbb{R}_+^*, \mathbf{x}_j \in K'\},\end{aligned}$$

puisque $C \subset K'$. Mais la proposition 3 montre que ce dernier ensemble est égal à K'. $\blacksquare$

Le *cône convexe engendré par* $E \in \mathscr{P}(\mathbb{R}^d)$ est l'ensemble

$$\begin{aligned}\mathbb{P}\,E &:= \{\lambda_1\mathbf{x}_1 + \cdots + \lambda_q\mathbf{x}_q \,|\, q \in \mathbb{N}^*, \lambda_j \in \mathbb{R}_+, \mathbf{x}_j \in E\} \\ &= \{\mathbf{0}\} \cup \{\lambda_1\mathbf{x}_1 + \cdots + \lambda_q\mathbf{x}_q \,|\, q \in \mathbb{N}^*, \lambda_j \in \mathbb{R}_+^*, \mathbf{x}_j \in E\}.\end{aligned}$$

Théorème 15. Soit $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^d$. Alors $K := \mathbb{P}\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$ est fermé.

Démonstration. **Étape 1.** On montre tout d'abord le résultat dans le cas où $\mathbf{0} \notin \mathrm{co}\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$. Soit $(\mathbf{x}_k)_{k\in\mathbb{N}^*} \subset K$ une suite qui converge vers un point $\mathbf{x}$, et montrons que $\mathbf{x} \in K$. En notant A la matrice de taille $m \times d$ dont la j-ième ligne est égale au vecteur $\mathbf{a}_j^\top$ $(j = 1, \ldots, m)$, on remarque que

$$K = \{y_1\mathbf{a}_1 + \cdots + y_m\mathbf{a}_m \mid y_1, \ldots, y_m \in \mathbb{R}_+\} = A^\top\mathbb{R}_+^m.$$

Pour tout $k \in \mathbb{N}^*$, $\mathbf{x}_k = A^\top\mathbf{y}_k$ avec $\mathbf{y}_k \in \mathbb{R}_+^m$. Supposons, en vue d'obtenir une contradiction, que la suite $(\mathbf{y}_k)_{k\in\mathbb{N}^*}$ n'est pas bornée. En prenant une sous-suite si nécessaire, cela revient à supposer que

$$\|\mathbf{y}_k\|_1 \to \infty, \quad \text{où} \quad \|\mathbf{y}\|_1 := \sum_{j=1}^m |y_j|\,.$$

Soit $\mathbf{z}_k := \|\mathbf{y}_k\|_1^{-1}\mathbf{y}_k$. Par compacité de $\{\mathbf{u} \in \mathbb{R}^d \,|\, \|\mathbf{u}\|_1 = 1\}$, on peut supposer, en prenant à nouveau une sous-suite si nécessaire, que $(\mathbf{z}_k)_{k\in\mathbb{N}^*}$ converge vers un point $\mathbf{z} \in \mathbb{R}_+^m$ tel que $\sum_{j=1}^m z_j = 1$. Alors,

$$\sum_{j=1}^m z_j\mathbf{a}_j = A^\top\mathbf{z} = \lim_{k\to\infty} A^\top\mathbf{z}_k = \lim_{k\to\infty} \frac{1}{\|\mathbf{y}_k\|_1} A^\top\mathbf{y}_k = \mathbf{0},$$

puisque la suite des $A^\top\mathbf{y}_k = \mathbf{x}_k$ est bornée. Ceci montre que $\mathbf{0} \in \mathrm{co}\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$, en contradiction avec l'hypothèse. La suite $(\mathbf{y}_k)_{k\in\mathbb{N}^*}$ est donc bornée. En prenant encore une sous-suite si nécessaire, on peut supposer que $\mathbf{y}_k \to \mathbf{y} \in \mathbb{R}_+^m$ lorsque $k \to \infty$. Alors, la suite $(\mathbf{x}_k)_{k\in\mathbb{N}^*}$ converge vers $A^\top\mathbf{y}$, ce qui montre que

$$\mathbf{x} = A^\top\mathbf{y} \in A^\top\mathbb{R}_+^m = K.$$

Étape 2. On montre ensuite que, si $\mathbf{0} \in \mathrm{co}\{\mathbf{a}_1, \ldots, \mathbf{a}_m\}$, alors

$$K = \mathbb{P}\{\mathbf{a}_1, \ldots, \mathbf{a}_m\} = \bigcup_{i=1}^m \mathbb{P}\big(\{\mathbf{a}_1, \ldots, \mathbf{a}_m\} \setminus \{\mathbf{a}_i\}\big) =: K'.$$

L'inclusion $K' \subset K$ est évidente. Soit donc $\mathbf{x} \in K$, et montrons que $\mathbf{x} \in K'$. Il existe $\mathbf{y} \in \mathbb{R}_+^m$ tel que $\mathbf{x} = A^\top\mathbf{y}$. Soit $\boldsymbol{\lambda} \in \mathbb{R}_+^m$ tel que

$$\sum_{j=1}^m \lambda_j = 1 \quad \text{et} \quad \mathbf{0} = \sum_{j=1}^m \lambda_j\mathbf{a}_j, \tag{3.1}$$

et soit $J := \{j \in \{1, \dots, m\} \mid \lambda_j > 0\}$. Il est clair que $J \neq \emptyset$. On pose

$$\alpha := \min_{j \in J} \frac{y_j}{\lambda_j}.$$

Alors $y_j - \alpha\lambda_j \geq 0$ pour tout $j \in \{1, \dots, m\}$ et il existe $i \in \{1, \dots, m\}$ tel que $y_i - \alpha\lambda_i = 0$. En utilisant la deuxième égalité de (3.1), on voit que

$$\mathbf{x} = A^\top \mathbf{y} = \sum_{j=1}^{m} y_j \mathbf{a}_j = \sum_{j=1}^{m} (y_j - \alpha\lambda_j)\mathbf{a}_j \in \mathbb{P}\big(\{\mathbf{a}_1, \dots, \mathbf{a}_m\} \setminus \{\mathbf{a}_i\}\big).$$

Étape 3. On montre enfin le théorème par récurrence. Étant donné $m \in \mathbb{N}^*$, on considère la propriété suivante :

$$(\mathscr{P}_m) \quad \forall \mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^d, \quad \mathbb{P}\{\mathbf{a}_1, \dots, \mathbf{a}_m\} \text{ est fermé.}$$

Il est clair que $(\mathscr{P}_1)$ est satisfaite, puisque $\mathbb{R}_+\{\mathbf{a}\}$ est soit une demi-droite fermée, soit le singleton $\{\mathbf{0}\}$. Supposons $(\mathscr{P}_m)$ satisfaite, et soit

$$\{\mathbf{a}_1, \dots, \mathbf{a}_{m+1}\} \subset \mathbb{R}^d.$$

Si $\mathbf{0} \notin \operatorname{co}\{\mathbf{a}_1, \dots, \mathbf{a}_{m+1}\}$, alors $\mathbb{P}\{\mathbf{a}_1, \dots, \mathbf{a}_{m+1}\}$ est fermé d'après l'étape 1. Dans le cas contraire, l'étape 2 montre que

$$\mathbb{P}\{\mathbf{a}_1, \dots, \mathbf{a}_{m+1}\} = \bigcup_{i=1}^{m+1} \mathbb{P}\big(\{\mathbf{a}_1, \dots, \mathbf{a}_{m+1}\} \setminus \{\mathbf{a}_i\}\big).$$

Par hypothèse de récurrence, $\mathbb{P}\big(\{\mathbf{a}_1, \dots, \mathbf{a}_{m+1}\} \setminus \{\mathbf{a}_i\}\big)$ est fermé pour tout i. Il s'ensuit que $\mathbb{P}\{\mathbf{a}_1, \dots, \mathbf{a}_{m+1}\}$ est fermé, comme réunion finie de fermés. ■

3.2 Cônes polaires

Soit $K \subset \mathbb{R}^d$ un cône. Le *polaire* de K est l'ensemble

$$K^\circ := \left\{\boldsymbol{\xi} \in \mathbb{R}^d \;\middle|\; \forall \mathbf{x} \in K,\ \langle \boldsymbol{\xi}, \mathbf{x} \rangle \leq 0\right\}.$$

Proposition 5. (1) Soit K un cône. Alors K° est un cône convexe fermé.

(2) Soit K un cône convexe. Alors $K^{\circ\circ} = \operatorname{cl} K$.

Démonstration.

(1) On vérifie facilement que, pour tout $\mathbf{x} \in \mathbb{R}^d$, $\{\boldsymbol{\xi} \in \mathbb{R}^d \mid \langle \boldsymbol{\xi}, \mathbf{x} \rangle \leq 0\}$ est un cône convexe fermé (c'est soit un demi-espace fermé dont le bord contient l'origine, soit $\mathbb{R}^d$ tout entier). Le résultat découle alors immédiatement du fait que

$$K^\circ = \bigcap_{\mathbf{x} \in K} \left\{\boldsymbol{\xi} \in \mathbb{R}^d \;\middle|\; \langle \boldsymbol{\xi}, \mathbf{x} \rangle \leq 0\right\}.$$

(2) Soit $\mathbf{x} \in K$. Pour tout $\boldsymbol{\xi} \in K^\circ$, $\langle \boldsymbol{\xi}, \mathbf{x} \rangle \leq 0$. Donc $\mathbf{x} \in K^{\circ\circ}$. Donc $K \subset K^{\circ\circ}$, et puisque $K^{\circ\circ}$ est fermé, $\operatorname{cl} K \subset K^{\circ\circ}$. Afin de montrer l'inclusion opposée, supposons que $\mathbf{x} \notin \operatorname{cl} K$. D'après le théorème 12, $\{\mathbf{x}\}$ et K sont séparés fortement :

$$\exists \mathbf{b} \in \mathbb{R}^d \setminus \{\mathbf{0}\} : \quad \sup_{\mathbf{y} \in K} \langle \mathbf{b}, \mathbf{y} \rangle < \langle \mathbf{b}, \mathbf{x} \rangle. \tag{3.2}$$

Pour tout $\mathbf{y} \in K$ et tout $t > 0$, $t\mathbf{y} \in K$, de sorte que $t\langle \mathbf{b}, \mathbf{y} \rangle < \langle \mathbf{b}, \mathbf{x} \rangle$, d'après l'inégalité (3.2). En divisant par t, puis en faisant tendre t vers l'infini, on voit que

$$\sup_{\mathbf{y} \in K} \langle \mathbf{b}, \mathbf{y} \rangle \leq 0.$$

Par ailleurs, choisissons $\mathbf{z}$ quelconque dans K. Alors, pour tout $t > 0$, $t\mathbf{z} \in K$, de sorte que $\sup_{\mathbf{y} \in K} \langle \mathbf{b}, \mathbf{y} \rangle \geq t\langle \mathbf{b}, \mathbf{z} \rangle$. En faisant tendre t vers 0, on voit que

$$\sup_{\mathbf{y} \in K} \langle \mathbf{b}, \mathbf{y} \rangle \geq 0.$$

Par conséquent, $\sup_{\mathbf{y} \in K} \langle \mathbf{b}, \mathbf{y} \rangle = 0$, ce qui montre que $\mathbf{b} \in K^\circ$. Comme $\langle \mathbf{b}, \mathbf{x} \rangle > 0$, on voit que $\mathbf{x} \notin K^{\circ\circ}$. ∎

3.3 Théorèmes de Farkas et de Gordan

Le théorème suivant est souvent appelé le *lemme de Farkas*[1]. Il est essentiel en programmation linéaire, en particulier dans l'approche dite *duale*. Il intervient aussi en programmation non linéaire, dans la preuve du théorème de Karush-Kuhn-Tucker. Dans tout ce paragraphe, $\{\mathbf{e}_1, \ldots, \mathbf{e}_m\}$ désignera la base canonique de $\mathbb{R}^m$.

Théorème 16. Soit $A \in M_{m \times n}(\mathbb{R})$, et soit $\boldsymbol{\xi} \in \mathbb{R}^n$. Les assertions suivantes sont équivalentes :

(a) pour tout $\mathbf{x} \in \mathbb{R}^n$ tel que $A\mathbf{x} \leq \mathbf{0}$, $\langle \boldsymbol{\xi}, \mathbf{x} \rangle \leq \mathbf{0}$;

(b) il existe $\mathbf{y} \in \mathbb{R}^m_+$ tel que $\boldsymbol{\xi} = A^\top \mathbf{y}$.

Autrement dit, $\{\mathbf{x} \in \mathbb{R}^n \,|\, A\mathbf{x} \leq \mathbf{0}\}^\circ = A^\top \mathbb{R}^m_+$.

Démonstration. On note a_{jk} le coefficient de la j-ième ligne et de la k-ième colonne de A et $\mathbf{a}_j := (a_{j1}, \ldots, a_{jn})^\top \in \mathbb{R}^n$. Rappelons que

$$A^\top \mathbb{R}^m_+ = \{y_1 \mathbf{a}_1 + \cdots + y_m \mathbf{a}_m \mid y_1, \ldots, y_m \in \mathbb{R}_+\}.$$

Montrons tout d'abord l'inclusion $A^\top \mathbb{R}^m_+ \subset \{\mathbf{x} \in \mathbb{R}^n \,|\, A\mathbf{x} \leq \mathbf{0}\}^\circ$. Soit donc $\mathbf{z} = y_1 \mathbf{a}_1 + \cdots + y_m \mathbf{a}_m$, où $\mathbf{y} = (y_1, \ldots, y_m)^\top \in \mathbb{R}^m_+$. Pour tout $\mathbf{x} \in \mathbb{R}^n$ tel que $A\mathbf{x} \leq \mathbf{0}$,

$$\langle \mathbf{z}, \mathbf{x} \rangle = \sum_{j=1}^m y_j \, \langle \mathbf{a}_j, \mathbf{x} \rangle \leq 0,$$

[1] Julius Farkas (1847-1930) : Mathématicien et physicien hongrois.

puisque $\langle \mathbf{a}_j, \mathbf{x} \rangle = (A\mathbf{x})_j \leq 0$ pour tout j. Donc $\mathbf{z} \in \{\mathbf{x} \in \mathbb{R}^n \,|\, A\mathbf{x} \leq \mathbf{0}\}^\circ$. Réciproquement, montrons que, si $\mathbf{x}_0 \notin A^\top \mathbb{R}^m_+$, alors $\mathbf{x}_0 \notin \{\mathbf{x} \in \mathbb{R}^n \,|\, A\mathbf{x} \leq \mathbf{0}\}^\circ$. D'après le théorème 15, $A^\top \mathbb{R}^m_+$ est un cône convexe fermé. Si $\mathbf{x}_0 \notin A^\top \mathbb{R}^m_+$, on peut séparer $\{\mathbf{x}_0\}$ et $A^\top \mathbb{R}^m_+$ fortement : il existe $\mathbf{b} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ tel que

$$\sup_{\mathbf{z} \in A^\top \mathbb{R}^m_+} \langle \mathbf{b}, \mathbf{z} \rangle < \langle \mathbf{b}, \mathbf{x}_0 \rangle \, .$$

Puisque $A^\top \mathbb{R}^m_+$ est un cône, $\langle \mathbf{b}, t\mathbf{z} \rangle < \langle \mathbf{b}, \mathbf{x}_0 \rangle$ pour tout $t > 0$ et tout $\mathbf{z} \in A^\top \mathbb{R}^m_+$. En divisant la dernière inégalité par t, puis en faisant tendre t vers l'infini, on obtient :

$$\sup_{\mathbf{z} \in A^\top \mathbb{R}^m_+} \langle \mathbf{b}, \mathbf{z} \rangle \leq 0. \tag{3.3}$$

En fait, le *supremum* ci-dessus est égal à zéro, puisque $\mathbf{0} \in A^\top \mathbb{R}^m_+$, de sorte que $\langle \mathbf{b}, \mathbf{x}_0 \rangle > 0$. Maintenant, en prenant $\mathbf{z} = A^\top \mathbf{e}_1, \ldots, A^\top \mathbf{e}_m$ dans (3.3), on voit que $\langle A\mathbf{b}, \mathbf{e}_j \rangle \leq 0$ pour tout j, c'est-à-dire, $A\mathbf{b} \leq \mathbf{0}$. Puisque $\langle \mathbf{b}, \mathbf{x}_0 \rangle > 0$, ceci montre que $\mathbf{x}_0 \notin \{\mathbf{x} \in \mathbb{R}^n \,|\, A\mathbf{x} \leq \mathbf{0}\}^\circ$. ∎

Corollaire 7. Soient $A \in M_{m \times n}(\mathbb{R})$ et $\mathbf{b} \in \mathbb{R}^m$. Les assertions suivantes sont équivalentes :

(a) $\{\mathbf{x} \in \mathbb{R}^n \,|\, A\mathbf{x} \leq \mathbf{b}\} \neq \emptyset$;

(b) si $A^\top \mathbf{y} = \mathbf{0}$ avec $\mathbf{y} \in \mathbb{R}^m_+$, alors $\langle \mathbf{b}, \mathbf{y} \rangle \geq 0$.

Démonstration. Remarquons tout d'abord que (a) équivaut à la condition $\mathbf{b} \in A\mathbb{R}^n + \mathbb{R}^m_+$. Or, en notant $\mathbf{a}^j$ la j-ième colonne de A,

$$\begin{aligned} A\mathbb{R}^n &= \{x_1 \mathbf{a}^1 + \cdots + x_n \mathbf{a}^n \,|\, x_1, \ldots, x_n \in \mathbb{R}\} \\ &= \{(s_1 - t_1)\mathbf{a}^1 + \cdots + (s_n - t_n)\mathbf{a}^n \,|\, s_1, \ldots, s_n, t_1, \ldots, t_n \in \mathbb{R}_+\} \\ &= \mathbb{P}\{\mathbf{a}^1, \ldots, \mathbf{a}^n, -\mathbf{a}^1, \ldots, -\mathbf{a}^n\}. \end{aligned}$$

Il s'ensuit que

$$\begin{aligned} A\mathbb{R}^n + \mathbb{R}^m_+ &= \mathbb{P}\{\mathbf{a}^1, \ldots, \mathbf{a}^n, -\mathbf{a}^1, \ldots, -\mathbf{a}^n, \mathbf{e}^1, \ldots, \mathbf{e}^m\} \\ &= \big[A; -A; I_m\big] \mathbb{R}^m_+, \end{aligned}$$

où I_m désigne la matrice identité de taille $m \times m$, et où $C^\top := \big[A; -A; I_m\big] \in M_{m \times (2n+m)}(\mathbb{R})$. On remarque au passage que, d'après le théorème 15, $A\mathbb{R}^n + \mathbb{R}^m_+$ est un cône convexe fermé. Le théorème 16 montre que

$$A\mathbb{R}^n + \mathbb{R}^m_+ = \{\mathbf{y} \in \mathbb{R}^m \mid C\mathbf{y} \leq \mathbf{0}\}^\circ = \big\{\mathbf{y} \in \mathbb{R}^m_- \;\big|\; A^\top \mathbf{y} = 0\big\}^\circ,$$

où la deuxième égalité résulte du fait que $C\mathbf{y} \leq \mathbf{0}$ si et seulement si $A^\top \mathbf{y} \leq \mathbf{0}$, $-A^\top \mathbf{y} \leq \mathbf{0}$ et $\mathbf{y} \leq \mathbf{0}$ (c'est-à-dire, $\mathbf{y} \in \mathbb{R}^m_-$ et $A^\top \mathbf{y} = \mathbf{0}$). Finalement, nous avons :

$$\begin{aligned}
\text{(a)} \quad &\iff \quad \mathbf{b} \in A\mathbb{R}^n + \mathbb{R}^m_+ \\
&\iff \quad \forall \mathbf{y} \in \mathbb{R}^m_- \cap \ker A^\top,\ \langle \mathbf{b}, \mathbf{y} \rangle \leq 0 \\
&\iff \quad \forall \mathbf{y} \in \mathbb{R}^m_+ \cap \ker A^\top,\ \langle \mathbf{b}, \mathbf{y} \rangle \geq 0 \\
&\iff \quad \text{(b)}. \ \blacksquare
\end{aligned}$$

Corollaire 8. Soient $A \in M_{m\times n}(\mathbb{R})$, $C \in M_{p\times n}(\mathbb{R})$, $\mathbf{b} \in \mathbb{R}^m$ et $\mathbf{c} \in \mathbb{R}^p$. Les assertions suivantes sont équivalentes :

(a) $\{\mathbf{x} \in \mathbb{R}^n \,|\, A\mathbf{x} \leq \mathbf{b},\ C\mathbf{x} = \mathbf{c}\} \neq \emptyset$;

(b) si $A^\top \mathbf{u} + C^\top \mathbf{t} = \mathbf{0}$ avec $\mathbf{u} \in \mathbb{R}^m_+$ et $\mathbf{t} \in \mathbb{R}^p$, alors $\langle \mathbf{b}, \mathbf{u} \rangle + \langle \mathbf{c}, \mathbf{t} \rangle \geq 0$.

Démonstration. Puisque $C\mathbf{x} = \mathbf{c}$ si et seulement si $C\mathbf{x} \leq \mathbf{c}$ et $-C\mathbf{x} \leq -\mathbf{c}$,

$$\{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} \leq \mathbf{b},\ C\mathbf{x} = \mathbf{c}\} = \{\mathbf{x} \in \mathbb{R}^n \mid D\mathbf{x} \leq \mathbf{d}\},$$

$$\text{où} \quad D := \left[A^\top; C^\top; -C^\top\right]^\top \quad \text{et} \quad \mathbf{d} := \left(\mathbf{b}^\top; \mathbf{c}^\top; -\mathbf{c}^\top\right)^\top.$$

D'après le corollaire 7, (a) équivaut à l'implication

$$\left.\begin{array}{r} \left[A^\top; C^\top; -C^\top\right] \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \\ \mathbf{w} \end{pmatrix} = \mathbf{0} \\ \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \\ \mathbf{w} \end{pmatrix} \in \mathbb{R}^{m+2p}_+ \end{array}\right\} \Longrightarrow \left\langle \begin{pmatrix} \mathbf{b} \\ \mathbf{c} \\ -\mathbf{c} \end{pmatrix}, \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \\ \mathbf{w} \end{pmatrix} \right\rangle \geq 0.$$

En posant $\mathbf{t} := \mathbf{v} - \mathbf{w}$, et en remarquant que $\mathbb{R}^p_+ - \mathbb{R}^p_+ = \mathbb{R}^p$, l'implication ci-dessus peut encore s'écrire

$$\left.\begin{array}{l} A^\top \mathbf{u} + C^\top \mathbf{t} = \mathbf{0} \\ \mathbf{u} \in \mathbb{R}^m_+,\ \mathbf{t} \in \mathbb{R}^p \end{array}\right\} \Longrightarrow \langle \mathbf{b}, \mathbf{u} \rangle + \langle \mathbf{c}, \mathbf{t} \rangle \geq 0. \ \blacksquare$$

Théorème 17. [Gordan] Soit $A \in M_{m\times n}(\mathbb{R})$. Les assertions suivantes sont équivalentes :

(a) $\{\mathbf{x} \in \mathbb{R}^n \,|\, A\mathbf{x} < \mathbf{0}\} \neq \emptyset$;

(b) $\ker A^\top \cap \mathbb{R}^m_+ = \{\mathbf{0}\}$.

Démonstration. Supposons qu'il existe $\mathbf{x} \in \mathbb{R}^n$ tel que $A\mathbf{x} < \mathbf{0}$. Alors il existe $\delta > 0$ tel que $A\mathbf{x} \leq -\delta\mathbf{e}$, où $\mathbf{e} := (1, \ldots, 1)^\top$. Donc

$$\{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} \leq -\delta\mathbf{e}\} \neq \emptyset.$$

D'après le corollaire 7, tout $\mathbf{y} \in \mathbb{R}_+^m$ tel que $A^\top \mathbf{y} = \mathbf{0}$ satisfait :

$$-\delta \langle \mathbf{e}, \mathbf{y} \rangle = -\delta \sum_{j=1}^{m} y_j \geq 0.$$

Donc $\ker A^\top \cap \mathbb{R}_+^m = \{\mathbf{0}\}$. Réciproquement, supposons que $\ker A^\top \cap \mathbb{R}_+^m = \{\mathbf{0}\}$: si $A^\top \mathbf{y} = \mathbf{0}$ avec $\mathbf{y} \in \mathbb{R}_+^m$, alors $\mathbf{y} = \mathbf{0}$, de sorte que $\langle -\mathbf{e}, \mathbf{y} \rangle \geq 0$. Le corollaire 7 montre alors que

$$\{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} \leq -\mathbf{e}\} \neq \emptyset,$$

donc que $\{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} < \mathbf{0}\} \neq \emptyset$. ∎

Chapitre 4

Fonctions convexes

4.1 Généralités

Une fonction $f\colon \mathbb{R}^d \to \mathbb{R}$ est dite *convexe* lorsque

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^d, \quad \forall \lambda \in [0,1]\,, \quad f\big((1-\lambda)\mathbf{x} + \lambda \mathbf{y}\big) \leq (1-\lambda) f(\mathbf{x}) + \lambda f(\mathbf{y}). \quad (4.1)$$

Il est facile de voir que f est convexe si et seulement si son *épigraphe*

$$\operatorname{epi} f := \big\{(\mathbf{x}, \alpha) \in \mathbb{R}^{d+1} \;\big|\; f(\mathbf{x}) \leq \alpha\big\}$$

est un sous-ensemble convexe. Cette caractérisation des fonctions convexes à valeurs dans $\mathbb{R}$ permet d'étendre la définition aux fonctions à valeurs dans la droite réelle étendue $\bar{\mathbb{R}} := [-\infty, \infty]$: on dit que $f\colon \mathbb{R}^d \to \bar{\mathbb{R}}$ est convexe lorsque son épigraphe est convexe. La définition ci-dessus de l'épigraphe d'une fonction reste bien sûr valide pour les fonctions à valeurs dans $\bar{\mathbb{R}}$, alors que la définition (4.1) pose problème, l'expression $\infty - \infty$ étant non définie. Dans toute la suite, on considère des fonctions à valeurs dans $\bar{\mathbb{R}}$. L'analogue de la condition (4.1) pour de telles fonctions fait l'objet du théorème suivant.

Théorème 18. Soit $f\colon \mathbb{R}^d \to \bar{\mathbb{R}}$. Il y a équivalence entre

(a) f est convexe ;

(b) pour tous $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ et tout $\lambda \in \,]0,1[$, si $f(\mathbf{x}) < \alpha \in \mathbb{R}$ et $f(\mathbf{y}) < \beta \in \mathbb{R}$ alors $f\big((1-\lambda)\mathbf{x} + \lambda\mathbf{y}\big) < (1-\lambda)\alpha + \lambda\beta$;

(c) pour tous $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ et tout $\lambda \in \,]0,1[$, si $f(\mathbf{x}) < \alpha \in \mathbb{R}$ et $f(\mathbf{y}) < \beta \in \mathbb{R}$ alors $f\big((1-\lambda)\mathbf{x} + \lambda\mathbf{y}\big) \leq (1-\lambda)\alpha + \lambda\beta$.

Démonstration. Supposons que f est convexe, c'est-à-dire, que $\operatorname{epi} f$ est convexe, et soient $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ et $\lambda \in \,]0,1[$. Si $f(\mathbf{x}) < \alpha \in \mathbb{R}$ et $f(\mathbf{y}) < \beta \in \mathbb{R}$, alors deux cas peuvent se présenter : soit $f(\mathbf{x}), f(\mathbf{y}) \in \mathbb{R}$, et alors $(\mathbf{x}, f(\mathbf{x})), (\mathbf{y}, f(\mathbf{y})) \in \operatorname{epi} f$, soit $f(\mathbf{x})$ ou (inclusif) $f(\mathbf{y})$ vaut $-\infty$. Dans le premier cas, la convexité de $\operatorname{epi} f$ implique que

$$(1-\lambda)(\mathbf{x}, f(\mathbf{x})) + \lambda(\mathbf{y}, f(\mathbf{y})) \in \operatorname{epi} f,$$

c'est-à-dire, $((1-\lambda)\mathbf{x}+\lambda\mathbf{y},(1-\lambda)f(\mathbf{x})+\lambda f(\mathbf{y})) \in \operatorname{epi} f$, et il s'ensuit que

$$f((1-\lambda)\mathbf{x}+\lambda\mathbf{y}) \leq (1-\lambda)f(\mathbf{x})+\lambda f(\mathbf{y}) < (1-\lambda)\alpha+\lambda\beta.$$

Dans le deuxième cas, avec par exemple $f(\mathbf{x})=-\infty$ et $f(\mathbf{y}) \in \mathbb{R}$, la convexité de $\operatorname{epi} f$ implique que

$$\forall a \in \mathbb{R}, \quad (1-\lambda)(\mathbf{x},a)+\lambda(\mathbf{y},f(\mathbf{y})) \in \operatorname{epi} f,$$

c'est-à-dire, $((1-\lambda)\mathbf{x}+\lambda\mathbf{y},(1-\lambda)a+\lambda f(\mathbf{y})) \in \operatorname{epi} f$ pour tout $a \in \mathbb{R}$; il est clair que cette dernière condition implique que $f((1-\lambda)\mathbf{x}+\lambda\mathbf{y})=-\infty < (1-\lambda)\alpha+\lambda\beta$. Le fait que (b) implique (c) est évident. Supposons la condition (c) satisfaite, et montrons (a). Soient $(\mathbf{x},\alpha),(\mathbf{y},\beta) \in \operatorname{epi} f$. On doit montrer que $((1-\lambda)\mathbf{x}+\lambda\mathbf{y},(1-\lambda)\alpha+\lambda\beta) \in \operatorname{epi} f$, c'est-à-dire, que

$$f\big((1-\lambda)\mathbf{x}+\lambda\mathbf{y}\big) \leq (1-\lambda)\alpha+\lambda\beta.$$

Cette dernière inégalité est fournie par (c) dans le cas où $f(\mathbf{x})<\alpha$ et $f(\mathbf{y})<\beta$, et reste vraie par passage à la limite dans les cas où une inégalité au moins est large. ■

On appelle *fonction indicatrice* d'un sous-ensemble E de $\mathbb{R}^d$ la fonction définie par

$$\delta(\mathbf{x}|E)=\begin{cases} 0 & \text{si } \mathbf{x}\in E, \\ \infty & \text{sinon.} \end{cases}$$

On vérifie aisément que $\delta(\mathbf{x}|E)$ est convexe si et seulement si E est convexe. Une fonction $f\colon \mathbb{R}^d \to \bar{\mathbb{R}}$ est dite *convexe sur* C, où $C \in \mathscr{C}(\mathbb{R}^d)$, lorsque la fonction qui coïncide avec f sur C et vaut ∞ sur son complémentaire est convexe. Ceci s'exprime de manière concise en disant que $f+\delta(\cdot|C)$ est convexe. La proposition suivante donne une caractérisation des fonctions d'une variable qui sont convexes sur un intervalle ouvert. Sa démonstration est laissée en exercice.

Proposition 6. [Critère des pentes croissantes] Soit $\varphi\colon \mathbb{R} \to \bar{\mathbb{R}}$ une fonction et I un intervalle de $\mathbb{R}$ sur lequel φ est finie. Alors, il y a équivalence entre :

(a) φ est convexe sur I ;

(b) pour tous $x,y,z \in I$ tels que $x<y<z$,
$$\frac{\varphi(y)-\varphi(x)}{y-x} \leq \frac{\varphi(z)-\varphi(x)}{z-x};$$

(c) pour tous $x,y,z \in I$ tels que $x<y<z$,
$$\frac{\varphi(y)-\varphi(x)}{y-x} \leq \frac{\varphi(z)-\varphi(y)}{z-y};$$

(d) pour tous $x,y,z \in I$ tels que $x<y<z$,
$$\frac{\varphi(z)-\varphi(x)}{z-x} \leq \frac{\varphi(z)-\varphi(y)}{z-y}.$$

On appelle *domaine effectif* (ou simplement *domaine*) d'une fonction convexe $f\colon \mathbb{R}^d \to \bar{\mathbb{R}}$ l'ensemble

$$\operatorname{dom} f := \left\{\mathbf{x} \in \mathbb{R}^d \;\middle|\; f(\mathbf{x}) < \infty\right\}.$$

Il est facile de voir que $\operatorname{dom} f$ est un sous-ensemble convexe de $\mathbb{R}^d$ et que, si P désigne la projection canonique de $\mathbb{R}^{d+1}$ sur $\mathbb{R}^d$, alors $\operatorname{dom} f = P \operatorname{epi} f$.
Les fonctions constantes $f \equiv \infty$ et $f \equiv -\infty$ sont des cas particuliers de fonctions convexes, d'épigraphes respectifs $\emptyset$ et $\mathbb{R}^{d+1}$, et donc de domaines respectifs $\emptyset$ et $\mathbb{R}^d$.
On dira qu'une droite de $\mathbb{R}^{d+1} = \mathbb{R}^d \times \mathbb{R}$ est verticale si elle est de la forme $\{\mathbf{x}\} \times \mathbb{R}$ avec $\mathbf{x} \in \mathbb{R}^d$. On remarque que $\operatorname{epi} f$ contient une droite verticale si et seulement si elle prend la valeur $-\infty$.
Une fonction convexe $f\colon \mathbb{R}^d \to \bar{\mathbb{R}}$ est dite *propre* si $f \not\equiv \infty$ et f ne prend pas la valeur $-\infty$, autrement dit, si $\operatorname{epi} f \neq \emptyset$ et $\operatorname{epi} f$ ne contient pas de droite verticale. Dans le cas contraire, f est dite *impropre*.
On remarque que, si f est convexe propre, alors on peut à nouveau écrire sans ambiguïté la condition

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^d, \quad \forall \lambda \in [0,1], \quad f\big((1-\lambda)\mathbf{x} + \lambda\mathbf{y}\big) \leq (1-\lambda) f(\mathbf{x}) + \lambda f(\mathbf{y}).$$

Si l'égalité est stricte dès que $\mathbf{x}, \mathbf{y} \in \operatorname{dom} f$, $\mathbf{x} \neq \mathbf{y}$ et $\lambda \in \,]0,1[$, on dit que f est *strictement convexe*.

Théorème 19. Soit f une fonction convexe impropre. Alors $f(\mathbf{x}) = -\infty$ pour tout $\mathbf{x} \in \operatorname{ri} \operatorname{dom} f$.

Démonstration. Si $f \equiv \infty$, alors $\operatorname{dom} f = \emptyset$ et il n'y a rien à montrer. Supposons donc que f prenne la valeur $-\infty$ en un certain $\mathbf{x}$ (de sorte que $\mathbf{x} \in \operatorname{dom} f$). Soit $\mathbf{z} \in \operatorname{ri} \operatorname{dom} f$. D'après le corollaire 2, il existe $\mu > 1$ tel que $\mathbf{y} := (1-\mu)\mathbf{x} + \mu\mathbf{z} \in \operatorname{dom} f$. Alors,

$$\mathbf{z} = \frac{1}{\mu}\big((\mu-1)\mathbf{x} + \mathbf{y}\big) = (1-\lambda)\mathbf{x} + \lambda\mathbf{y} \quad \text{avec} \quad \lambda := \frac{1}{\mu} \in \,]0,1[\,.$$

D'après le théorème 18, si $f(\mathbf{x}) < \alpha \in \mathbb{R}$ et $f(\mathbf{y}) < \beta \in \mathbb{R}$, alors

$$f(\mathbf{z}) < (1-\lambda)\alpha + \lambda\beta.$$

Comme $f(\mathbf{y}) < \infty$, on peut fixer β (dans $]f(\mathbf{y}), \infty[$), et comme $f(\mathbf{x}) = -\infty$, on peut faire tendre α vers $-\infty$, ce qui montre que nécessairement $f(\mathbf{z}) = -\infty$. ∎

Théorème 20. Soit f une fonction convexe sur $\mathbb{R}^d$. Alors

$$\operatorname{ri} \operatorname{epi} f = \left\{(\mathbf{x}, \xi) \in \mathbb{R}^{d+1} \;\middle|\; \mathbf{x} \in \operatorname{ri} \operatorname{dom} f,\ \xi > f(\mathbf{x})\right\}.$$

DÉMONSTRATION. Soit P la projection canonique de $\mathbb{R}^{d+1}$ sur $\mathbb{R}^d$. Puisque $\operatorname{dom} f = P \operatorname{epi} f$, le théorème 10 montre que

$$\operatorname{ri}\operatorname{dom} f = P \operatorname{ri}\operatorname{epi} f. \tag{4.2}$$

Pour tout $\mathbf{x} \in \mathbb{R}^d$, $P^{-1}\{\mathbf{x}\} = \{\mathbf{x}\} \times \mathbb{R}$ est une droite (verticale), de sorte que $\operatorname{ri} P^{-1}\{\mathbf{x}\} = P^{-1}\{\mathbf{x}\}$. Par conséquent, si $\mathbf{x} \in \operatorname{ri}\operatorname{dom} f$, $\operatorname{ri} P^{-1}\{\mathbf{x}\} \cap \operatorname{ri}\operatorname{epi} f$ est non vide, et il résulte alors de la proposition 1 que

$$\operatorname{ri} P^{-1}\{\mathbf{x}\} \cap \operatorname{ri}\operatorname{epi} f = \operatorname{ri}(P^{-1}\{\mathbf{x}\} \cap \operatorname{epi} f).$$

Or, $P^{-1}\{\mathbf{x}\} \cap \operatorname{epi} f = \{\mathbf{x}\} \times \{\alpha \in \mathbb{R} \,|\, \alpha \geq f(\mathbf{x})\}$. Nous avons donc montré que

$$\forall \mathbf{x} \in \operatorname{ri}\operatorname{dom} f, \quad P^{-1}\{\mathbf{x}\} \cap \operatorname{ri}\operatorname{epi} f = \{\mathbf{x}\} \times \{\alpha \in \mathbb{R} \,|\, \alpha > f(\mathbf{x})\}. \tag{4.3}$$

Maintenant, soit $(\mathbf{z}, \gamma) \in \operatorname{ri}\operatorname{epi} f$. Alors $\mathbf{z} \in \operatorname{ri}\operatorname{dom} f$ d'après (4.2) et $\gamma > f(\mathbf{z})$ d'après (4.3). Réciproquement, si $\mathbf{z} \in \operatorname{ri}\operatorname{dom} f$ et si $\gamma > f(\mathbf{z})$, (4.3) entraîne $(\mathbf{z}, \gamma) \in \operatorname{ri}\operatorname{epi} f$. ■

4.2 Fonctions convexes *sci*

Une fonction f de $\mathbb{R}^d$ dans $[-\infty, +\infty]$ est dite *semi-continue inférieurement* en un point $\mathbf{x}$ si l'on a

$$f(\mathbf{x}) \leq \liminf_{\mathbf{y} \to \mathbf{x}} f(\mathbf{y}) := \lim_{\varepsilon \downarrow 0} \inf \{f(\mathbf{y}) \mid 0 < \|\mathbf{y} - \mathbf{x}\| < \varepsilon\},$$

c'est-à-dire si pour toute suite $(\mathbf{x}_k)_{k \in \mathbb{N}^*}$ convergeant vers $\mathbf{x}$,

$$f(\mathbf{x}) \leq \liminf_{k \to \infty} f(\mathbf{x}_k) := \lim_{k \to \infty} \inf \{f(\mathbf{x}_j) \mid j \geq k\}.$$

Une fonction f ayant cette propriété en tout point $\mathbf{x} \in \mathbb{R}^d$ est dite semi-continue inférieurement (*sci*).

Exemple 2. [Transformée de Laplace] Soit μ une mesure de probabilité sur $\mathbb{R}^d$. La transformée de Laplace de μ est la fonction définie sur $\mathbb{R}^d$ par

$$G_\mu(\boldsymbol{\xi}) := \int e^{\langle \mathbf{x}, \boldsymbol{\xi} \rangle} \, \mathrm{d}\mu(\mathbf{x}).$$

Nous allons montrer que G_μ est semi-continue inférieurement. Soit $(\boldsymbol{\xi}_k)_{k \in \mathbb{N}^*}$ une suite qui converge vers $\boldsymbol{\xi}$. Considérons les fonctions mesurables positives

$$g(\mathbf{x}) := e^{\langle \mathbf{x}, \boldsymbol{\xi} \rangle} \quad \text{et} \quad g_k(\mathbf{x}) := e^{\langle \mathbf{x}, \boldsymbol{\xi}_k \rangle}, \quad k \in \mathbb{N}^*.$$

La continuité de l'application $\boldsymbol{\xi} \mapsto e^{\langle \mathbf{x}, \boldsymbol{\xi} \rangle}$ implique que

$$\liminf_{k \to \infty} g_k(\mathbf{x}) = \lim_{k \to \infty} g_k(\mathbf{x}) = g(\mathbf{x}).$$

D'après le lemme de Fatou, on a alors

$$\int g(\mathbf{x})\,\mathrm{d}\mu(\mathbf{x}) \leq \liminf_{k\to\infty} \int g_k(\mathbf{x})\,\mathrm{d}\mu(\mathbf{x}),$$

qui est précisément l'inégalité cherchée. ∎

Dans le paragraphe précédent, la convexité d'une fonction f sur $\mathbb{R}^d$ était définie *via* la convexité de son épigraphe dans $\mathbb{R}^{d+1}$. Nous établissons maintenant le lien entre la semi-continuité inférieure de f et la topologie de $\operatorname{epi} f$.

Théorème 21. Soit $f\colon \mathbb{R}^d \to \bar{\mathbb{R}}$. Il y a équivalence entre

(a) f est *sci* ;

(b) $\operatorname{epi} f$ est un ensemble fermé de $\mathbb{R}^{d+1}$.

(c) pour tout $\alpha \in \mathbb{R}$, l'ensemble de niveau $\operatorname{lev}_\alpha(f) := \{\mathbf{x} \in \mathbb{R}^d \mid f(\mathbf{x}) \leq \alpha\}$ est fermé.

Démonstration. Montrons tout d'abord que (a) implique (b). Considérons une suite $(\mathbf{x}_k, \alpha_k)_{k\in\mathbb{N}^*} \subset \operatorname{epi} f$ qui converge vers un point $(\mathbf{x}, \alpha)$. Alors,

$$\alpha = \lim_{k\to\infty} \alpha_k = \liminf_{k\to\infty} \alpha_k \geq \liminf_{k\to\infty} f(\mathbf{x}_k) \geq f(\mathbf{x}).$$

Donc $(\mathbf{x}, \alpha) \in \operatorname{epi} f$. Montrons ensuite que (b) implique (c). Soit $\alpha \in \mathbb{R}$ tel que $\operatorname{lev}_\alpha(f)$ soit non vide. Soit $(\mathbf{x}_k)_{k\in\mathbb{N}^*} \subset \operatorname{lev}_\alpha(f)$ une suite qui converge vers un point $\mathbf{x}$. Alors,

$$\forall k \in \mathbb{N}^*, \quad (\mathbf{x}_k, \alpha) \in \operatorname{epi} f.$$

La condition (b) implique que $(\mathbf{x}, \alpha) \in \operatorname{epi} f$, c'est-à-dire, que $\mathbf{x} \in \operatorname{lev}_\alpha(f)$. Montrons enfin, par contraposition, que (c) implique (a). Supposons donc qu'il existe un point $\mathbf{x}$ tel que f ne soit pas *sci* en $\mathbf{x}$. Alors, il existe une suite $(\mathbf{x}_k)_{k\in\mathbb{N}^*}$ telle que

$$\mathbf{x}_k \to \mathbf{x} \quad \text{lorsque} \quad k \to \infty \quad \text{et} \quad f(\mathbf{x}) > \liminf_{k\to\infty} f(\mathbf{x}_k).$$

Donc, il existe une sous-suite $(\mathbf{x}_{k_l})_{l\in\mathbb{N}^*}$ et un réel β tels que

$$\mathbf{x}_{k_l} \to \mathbf{x} \quad \text{lorsque} \quad l \to \infty \quad \text{et} \quad f(\mathbf{x}_{k_l}) \to \beta < f(\mathbf{x}).$$

Soit $\alpha \in \,]\beta, f(\mathbf{x})[$. Pour l suffisamment grand, disons $l \geq L$, $f(\mathbf{x}_{k_l}) \leq \alpha < f(\mathbf{x})$. Donc $\operatorname{lev}_\alpha(f)$ contient la suite convergente $(\mathbf{x}_{k_l})_{l\in\mathbb{N}^*}$ et non sa limite. ∎

Le théorème 21 et le corollaire 4 montrent que, si f est convexe et *sci*, son épigraphe est l'intersection des demi-espaces fermés qui le contiennent. Parmi ces demi-espaces, il est clair qu'il ne peut y avoir aucun demi-espace *inférieur*, c'est-à-dire, du type

$$\left\{(\mathbf{x}, \alpha) \in \mathbb{R}^{d+1} \;\middle|\; \alpha \leq \langle \mathbf{b}, \mathbf{x}\rangle - \beta\right\} \quad \text{avec} \quad \mathbf{b} \neq \mathbf{0}.$$

Nous allons voir que, si f est convexe propre *sci*, on peut retirer de la famille des demi-espaces fermés qui contiennent $\operatorname{epi} f$ tout ceux dont le bord est vertical sans changer l'intersection. Un hyperplan de $\mathbb{R}^{d+1}$ est dit *vertical* s'il est de la forme

$$\bigl\{(\mathbf{x},\alpha)\in\mathbb{R}^{d+1}\;\big|\;\langle\mathbf{b},\mathbf{x}\rangle=\beta\bigr\}\quad\text{avec}\quad \mathbf{b}\neq\mathbf{0}.$$

Autrement dit, il suffit pour obtenir $\operatorname{epi} f$ de considérer les demi-espaces fermés qui sont des épigraphes de fonctions affines minorant f (appelées aussi *minorantes affines* de f).

Théorème 22. Soit $f\colon\mathbb{R}^d\to\,]-\infty,+\infty]$ une fonction convexe, propre et *sci*, et soit $(\mathbf{x}_0,y_0)\notin\operatorname{epi} f$. Alors il existe une fonction affine dont le graphe sépare $\operatorname{epi} f$ et $\{(\mathbf{x}_0,y_0)\}$ fortement.

Démonstration. D'après le théorème 21, $\operatorname{epi} f$ est un ensemble convexe fermé. Le théorème 12 fournit alors un hyperplan H séparant $\operatorname{epi} f$ et $\{(\mathbf{x}_0,y_0)\}$ fortement. Cela signifie qu'il existe $(\boldsymbol{\xi},\eta)\in\mathbb{R}^{d+1}\setminus\{(\mathbf{0},0)\}$ et $c\in\mathbb{R}$ tels que

$$\sup_{(\mathbf{x},y)\in\operatorname{epi} f}\{\langle\boldsymbol{\xi},\mathbf{x}\rangle+\eta y-c\}<0<\langle\boldsymbol{\xi},\mathbf{x}_0\rangle+\eta y_0-c. \tag{4.4}$$

L'hyperplan séparateur est ici l'ensemble $H=\{(\mathbf{x},y)\in\mathbb{R}^{d+1}\,|\,\langle\boldsymbol{\xi},\mathbf{x}\rangle+\eta y=c\}$. Si $\eta\neq0$, H est le graphe de la fonction affine $\mathbf{x}\mapsto(c-\langle\boldsymbol{\xi},\mathbf{x}\rangle)/\eta$, et le résultat est établi. Si $\eta=0$, H est alors un hyperplan vertical, et (4.4) peut s'écrire

$$\sup_{\mathbf{x}\in\operatorname{dom} f}\{\langle\boldsymbol{\xi},\mathbf{x}\rangle-c\}<0<\langle\boldsymbol{\xi},\mathbf{x}_0\rangle-c. \tag{4.5}$$

D'autre part, f admet au moins une minorante affine m puisqu'elle est convexe propre *sci*. Dans le cas contraire, en effet, l'ensemble convexe fermé $\operatorname{epi} f$ serait l'intersection de demi-espaces fermés de frontière verticale et, par conséquent serait vide ou contiendrait une droite verticale. On peut donc écrire que, quel que soit $\mathbf{x}\in\mathbb{R}^d$,

$$m(\mathbf{x}):=\langle\boldsymbol{\xi}_1,\mathbf{x}\rangle-c_1\leq f(\mathbf{x}), \tag{4.6}$$

Pour tout $\alpha>0$, posons

$$m_\alpha(\mathbf{x}):=\alpha\bigl(\langle\boldsymbol{\xi},\mathbf{x}\rangle-c\bigr)+\bigl(\langle\boldsymbol{\xi}_1,\mathbf{x}\rangle-c_1\bigr).$$

Puisque $\langle\boldsymbol{\xi},\mathbf{x}_0\rangle-c>0$, il est clair que, pour α suffisamment grand, $m_\alpha(\mathbf{x}_0)>y_0$. Pour un tel α, on a alors :

$$\begin{aligned}
&\sup_{(\mathbf{x},y)\in\operatorname{epi} f}\{m_\alpha(\mathbf{x})-y\}\\
&\qquad\leq\ \alpha\sup_{\mathbf{x}\in\operatorname{dom} f}\{\langle\boldsymbol{\xi},\mathbf{x}\rangle-c\}+\sup_{(\mathbf{x},y)\in\operatorname{epi} f}\{\langle\boldsymbol{\xi}_1,\mathbf{x}\rangle-c_1-y\}\\
&\qquad<\ 0\\
&\qquad<\ m_\alpha(\mathbf{x}_0)-y_0,
\end{aligned}$$

où la première inégalité stricte provient de (4.5) et (4.6). Le graphe de l'application affine m_α est l'hyperplan non vertical H_α d'équation $m_\alpha(\mathbf{x}) - y = 0$, et les inégalités ci-dessus montrent que H_α sépare les ensembles $\operatorname{epi} f$ et $\{(\mathbf{x}_0, y_0)\}$ fortement. ∎

Corollaire 9. Toute fonction convexe propre *sci* sur $\mathbb{R}^d$ est égale à l'enveloppe supérieure de ses minorantes affines.

Démonstration. Soit h l'enveloppe supérieure des minorantes affines de f. Il est clair que $h \le f$. Supposons qu'il existe $\mathbf{x}_0 \in \mathbb{R}^d$ tel que $h(\mathbf{x}_0) < f(\mathbf{x}_0)$. Alors, $(\mathbf{x}_0, h(\mathbf{x}_0)) \notin \operatorname{epi} f$, et le théorème 22 fournit une fonction affine m dont le graphe sépare $\operatorname{epi} f$ et $(\mathbf{x}_0, h(\mathbf{x}_0))$ fortement. Par conséquent, m minore f et $m(\mathbf{x}_0) > h(\mathbf{x}_0)$, en contradiction avec la définition de h. ∎

Nous allons maintenant définir la notion de *régularisée sci* d'une fonction f. Remarquons tout d'abord qu'un sous-ensemble E de $\mathbb{R}^{d+1} = \mathbb{R}^d \times \mathbb{R}$ est l'épigraphe d'une fonction $f\colon \mathbb{R}^d \to \bar{\mathbb{R}}$ si et seulement si E est la réunion d'ensembles de la forme

$$\{\mathbf{x}\} \times \mathbb{R} \quad \text{ou} \quad \{\mathbf{x}\} \times [f(\mathbf{x}), \infty[\,.$$

Autrement dit, il y a équivalence entre

(a) E est un épigraphe ;

(b) pour tout $\mathbf{x} \in \mathbb{R}^d$, $(\{\mathbf{x}\} \times \mathbb{R}) \cap E$ est fermé et satisfait :

$$\left.\begin{aligned} (\mathbf{x}, \alpha) \in E \\ \beta > \alpha \end{aligned}\right\} \Longrightarrow (\mathbf{x}, \beta) \in E.$$

Lemme 4. Si $E \in \mathscr{P}(\mathbb{R}^{d+1})$ est un épigraphe, alors $\operatorname{cl} E$ est un épigraphe.

Démonstration. Nous allons utiliser la caractérisation ci-dessus des épigraphes. Supposons que E soit un épigraphe. Il est clair que, pour tout $\mathbf{x} \in \mathbb{R}^d$, $(\{\mathbf{x}\} \times \mathbb{R}) \cap \operatorname{cl} E$ est fermé, comme intersection de deux fermés. Soit $(\mathbf{x}, \alpha) \in \operatorname{cl} E$ et $\beta > \alpha$. Alors, $(\mathbf{x}, \alpha)$ est la limite d'une suite $(\mathbf{x}_k, \alpha_k)_{k \in \mathbb{N}^*} \subset E$. Comme E est un épigraphe, $(\mathbf{x}_k, \alpha_k + \beta - \alpha) \in E$ pour tout k. Mais $(\mathbf{x}_k, \alpha_k + \beta - \alpha) \to (\mathbf{x}, \beta)$ lorsque $k \to \infty$, ce qui montre que $(\mathbf{x}, \beta) \in \operatorname{cl} E$. ∎

Soit $f\colon \mathbb{R}^d \to \bar{\mathbb{R}}$. D'après le lemme précédent, $\operatorname{cl}(\operatorname{epi} f)$ est l'épigraphe d'une fonction, que l'on appelle la *régularisée sci* de f. On la note $\operatorname{lsc} f$. La fonction f est *sci* si et seulement si f et $\operatorname{lsc} f$ coïncident. En effet,

$$\begin{aligned} f \text{ sci} &\iff \operatorname{epi} f \text{ fermé} \\ &\iff \operatorname{epi} f = \operatorname{cl}(\operatorname{epi} f) = \operatorname{epi}(\operatorname{lsc} f) \\ &\iff f = \operatorname{lsc} f. \end{aligned}$$

Théorème 23. Soit $f\colon \mathbb{R}^d \to \bar{\mathbb{R}}$. Alors,

$$\operatorname{lsc} f = \sup\left\{h\colon \mathbb{R}^d \to \bar{\mathbb{R}} \;\middle|\; h \text{ est } sci \text{ et } h \le f\right\}.$$

Démonstration. Remarquons tout d'abord que $h \leq f$ équivaut à $\operatorname{epi} f \subset \operatorname{epi} h$, et que si $(h_\alpha)_{\alpha\in A}$ est une famille de fonctions (A est un ensemble quelconque d'indices), alors

$$\operatorname{epi} \sup_{\alpha\in A} h_\alpha = \bigcap_{\alpha\in A} \operatorname{epi} h_\alpha.$$

Maintenant, $\operatorname{epi}(\operatorname{lsc} f)$ est, par définition, l'intersection de la famille des fermés qui contiennent $\operatorname{epi} f$. Comme $\operatorname{epi}(\operatorname{lsc} f)$ est un épigraphe fermé, on peut restreindre la famille aux fermés qui sont des épigraphes. On a donc :

$$\operatorname{epi}(\operatorname{lsc} f) = \bigcap \{\operatorname{epi} h \mid h \text{ est } sci \text{ et } h \leq f\}. \blacksquare$$

Corollaire 10. Soit $f\colon \mathbb{R}^d \to \bar{\mathbb{R}}$. Alors, pour tout $\mathbf{x}$ dans $\mathbb{R}^d$,

$$(\operatorname{lsc} f)(\mathbf{x}) = \min\left\{f(\mathbf{x}), \liminf_{\mathbf{y}\to\mathbf{x}} f(\mathbf{y})\right\}$$

Démonstration. Soit $\mathbf{x} \in \mathbb{R}^d$. Puisque $\operatorname{lsc} f$ est *sci* et $\operatorname{lsc} f$ est majorée par f,

$$(\operatorname{lsc} f)(\mathbf{x}) \leq \liminf_{\mathbf{y}\to\mathbf{x}} (\operatorname{lsc} f)(\mathbf{y}) \leq \liminf_{\mathbf{y}\to\mathbf{x}} f(\mathbf{y}),$$

de sorte que

$$(\operatorname{lsc} f)(\mathbf{x}) \leq \mu := \min\left\{f(\mathbf{x}), \liminf_{\mathbf{y}\to\mathbf{x}} f(\mathbf{y})\right\}$$

Supposons, en vue d'obtenir une contradiction, que $(\operatorname{lsc} f)(\mathbf{x}) < \mu$. Alors, il existe $\lambda \in \mathbb{R}$ tel que $(\operatorname{lsc} f)(\mathbf{x}) < \lambda < \mu$. La condition $\lambda < \liminf_{\mathbf{y}\to\mathbf{x}} f(\mathbf{y})$ entraîne l'existence d'un $\varepsilon > 0$ tel que

$$\inf\{f(\mathbf{y}) \mid 0 < \|\mathbf{y} - \mathbf{x}\| < \varepsilon\} > \lambda.$$

Puisque $f(\mathbf{x}) \geq \mu > \lambda$, on voit que $f(\mathbf{y}) > \lambda$ pour tout $\mathbf{y} \in B(\mathbf{x}, \varepsilon)$. Ainsi,

$$\forall (\mathbf{y}, t) \in \Omega := B(\mathbf{x}, \varepsilon) \times]-\infty, \lambda[, \quad f(\mathbf{y}) > t.$$

On remarque que $f(\mathbf{y}) > t$ s'écrit aussi $(\mathbf{y}, t) \notin \operatorname{epi} f$. L'ensemble ouvert Ω est alors un voisinage de $(\mathbf{x}, (\operatorname{lsc} f)(\mathbf{x}))$ qui ne rencontre pas $\operatorname{epi} f$, en contradiction avec le fait que

$$\big(\mathbf{x}, (\operatorname{lsc} f)(\mathbf{x})\big) \in \operatorname{epi}(\operatorname{lsc} f) = \operatorname{cl}(\operatorname{epi} f). \blacksquare$$

4.3 Fermeture d'une fonction

Étant donnée une fonction $f\colon \mathbb{R}^d \to \bar{\mathbb{R}}$, on définit la *fermeture convexe* (ou simplement *fermeture*, lorsque le contexte est suffisamment clair) de f comme l'enveloppe supérieure des minorantes affines de f. On note $\mathrm{cl}\, f$ la fermeture de f :

$$\mathrm{cl}\, f := \sup\left\{h\colon \mathbb{R}^d \to \bar{\mathbb{R}} \;\middle|\; h \text{ est affine et } h \leq f\right\}.$$

Il est clair que, quelle que soit f, la fonction $\mathrm{cl}\, f$ est convexe *sci*. On dit que f est *convexe fermée* (ou simplement *fermée*) lorsque $f = \mathrm{cl}\, f$. Si $f \equiv \infty$, alors $\mathrm{cl}\, f \equiv \infty$, et si f prend la valeur $-\infty$, alors $\mathrm{cl}\, f \equiv -\infty$.

Théorème 24. Soit $f\colon \mathbb{R}^d \to \bar{\mathbb{R}}$ une fonction convexe propre. Alors $\mathrm{cl}\, f = \mathrm{lsc}\, f$, c'est-à-dire, $\mathrm{epi}(\mathrm{cl}\, f) = \mathrm{cl}(\mathrm{epi}\, f)$.

Démonstration. D'après le corollaire 9, $\mathrm{lsc}\, f$ est égale à l'enveloppe supérieure de ses minorantes affines. Il suffit donc de voir que, si f est convexe propre, alors la famille $\mathscr{M}(f)$ des minorantes affines de f et la famille $\mathscr{M}(\mathrm{lsc}\, f)$ coïncident. D'une part, l'inclusion $\mathscr{M}(\mathrm{lsc}\, f) \subset \mathscr{M}(f)$ résulte immédiatement de l'inégalité $\mathrm{lsc}\, f \leq f$. D'autre part, soit $h \in \mathscr{M}(f)$. Alors h est *sci* et $h \leq f$. D'après le théorème 23, $h \leq \mathrm{lsc}\, f$. Donc $h \in \mathscr{M}(\mathrm{lsc}\, f)$. ∎

Nous allons voir que, si f est convexe propre, alors $\mathrm{cl}\, f$ et f ne peuvent différer que sur la frontière relative de $\mathrm{dom}\, f$.

Théorème 25. Soit $f\colon \mathbb{R}^d \to \bar{\mathbb{R}}$ une fonction convexe propre. Alors, pour tout $\mathbf{x} \in \mathrm{ri}(\mathrm{dom}\, f)$, $f(\mathbf{x}) = (\mathrm{cl}\, f)(\mathbf{x})$.

Démonstration. On note comme d'habitude P la projection canonique de $\mathbb{R}^{d+1}$ sur $\mathbb{R}^d$. Soit $\mathbf{x} \in \mathrm{ri}(\mathrm{dom}\, f)$. Nous allons montrer que

$$P^{-1}\{\mathbf{x}\} \cap \mathrm{epi}\, f = P^{-1}\{\mathbf{x}\} \cap \mathrm{epi}(\mathrm{cl}\, f), \tag{4.7}$$

ce qui établira le théorème. D'après le théorème 20,

$$P^{-1}\{\mathbf{x}\} \cap \mathrm{ri}(\mathrm{epi}\, f) = \{\mathbf{x}\} \times \{\alpha \in \mathbb{R} \mid \alpha > f(\mathbf{x})\} \neq \emptyset.$$

Puisque $P^{-1}\{\mathbf{x}\} = \mathrm{ri}(P^{-1}\{\mathbf{x}\}) = \mathrm{cl}(P^{-1}\{\mathbf{x}\})$,

$$\mathrm{cl}(P^{-1}\{\mathbf{x}\} \cap \mathrm{epi}\, f) = \mathrm{cl}(P^{-1}\{\mathbf{x}\}) \cap \mathrm{cl}(\mathrm{epi}\, f) = P^{-1}\{\mathbf{x}\} \cap \mathrm{epi}(\mathrm{cl}\, f),$$

où la première égalité résulte de la proposition 1 et la deuxième du théorème 24. L'équation(4.7) découle alors de ce que l'ensemble $P^{-1}\{\mathbf{x}\} \cap \mathrm{epi}\, f$, qui n'est autre que $\{\mathbf{x}\} \times [f(\mathbf{x}), \infty[$, est fermé. ∎

4.4 Fonctions concaves

Une fonction $g\colon \mathbb{R}^d \to \bar{\mathbb{R}}$ est dite *concave* lorsque $-g$ est convexe. On conçoit aisément que tout énoncé concernant les fonctions convexes admet une formulation analogue pour les fonctions concaves. L'objectif de ce paragraphe est de donner quelques définitions, notations et assertions relatives aux fonctions concaves, qui résultent directement de leurs analogues convexes.
L'*hypographe* d'une fonction quelconque $g\colon \mathbb{R}^d \to \bar{\mathbb{R}}$ est l'ensemble

$$\operatorname{hypo} g := \left\{(\mathbf{x}, \alpha) \in \mathbb{R}^{d+1} \;\middle|\; \alpha \leq g(\mathbf{x})\right\}.$$

Une fonction $g\colon \mathbb{R}^d \to \bar{\mathbb{R}}$ est concave si et seulement si son hypographe est convexe. Dans ce cas, son *domaine effectif* (ou simplement *domaine*, si le contexte est suffisamment clair) est l'ensemble

$$\operatorname{dom} g := \{\mathbf{x} \in \mathbb{R}^d \,|\, g(\mathbf{x}) > -\infty\}.$$

Les fonctions de $\mathbb{R}^d$ qui sont à la fois convexes et concaves sont les fonctions affines et les fonctions constantes $\varphi \equiv \infty$ et $\varphi \equiv -\infty$. Si φ est affine, alors $\operatorname{dom} \varphi = \mathbb{R}^d$, que φ soit considérée comme une fonction convexe ou comme une fonction concave. Par contre, si $\varphi \equiv \infty$, $\operatorname{dom} \varphi = \emptyset$ si φ est considérée comme une fonction convexe et $\operatorname{dom} \varphi = \mathbb{R}^d$ si φ est considérée comme une fonction concave. En général, le contexte ne laisse aucune ambiguïté sur le sens de l'expression $\operatorname{dom} \varphi$.
Une fonction g est dite *concave propre* lorsque $-g$ est convexe propre, c'est-à-dire, lorsqu'elle ne prend pas la valeur ∞ et qu'elle n'est pas identiquement égale à $-\infty$.
Par ailleurs, une fonction $g\colon \mathbb{R}^d \to \bar{\mathbb{R}}$ (convexe ou non) est dite *semi-continue supérieurement* en un point $\mathbf{x}$ si $-g$ est *sci* en $\mathbf{x}$, c'est-à-dire, si

$$g(\mathbf{x}) \geq \limsup_{\mathbf{y} \to \mathbf{x}} g(\mathbf{y}) := \lim_{\varepsilon \downarrow 0} \sup \left\{g(\mathbf{y}) \;|\; 0 < \|\mathbf{y} - \mathbf{x}\| < \varepsilon\right\},$$

ou encore, si pour toute suite $(\mathbf{x}_k)_{k \in \mathbb{N}^*}$ convergeant vers $\mathbf{x}$,

$$g(\mathbf{x}) \geq \limsup_{k \to \infty} g(\mathbf{x}_k) := \lim_{k \to \infty} \sup \left\{g(\mathbf{x}_j) \;|\; j \geq k\right\}.$$

Une fonction g ayant cette propriété en tout point $\mathbf{x} \in \mathbb{R}^d$ est dite semi-continue supérieurement (*scs*).
Une fonction $g\colon \mathbb{R}^d \to \bar{\mathbb{R}}$ est *scs* si et seulement si son hypographe est fermé. La régularisée *scs* d'une fonction g est la fonction $\operatorname{usc} g$ dont l'hypographe est la fermeture de $\operatorname{hypo} g$. La clôture $\operatorname{cl} g$ de g est l'enveloppe inférieure de ses majorantes affines. Si g est concave propre, alors $\operatorname{usc} g = \operatorname{cl} g$.
Enfin, si f et g sont respectivement convexe propre et concave propre, $f - g$ est bien définie, convexe, et

$$\operatorname{dom}(f - g) = \operatorname{dom} f \cap \operatorname{dom} g.$$

Chapitre 5

Théorie différentielle et dualité

5.1 Fonctions positivement homogènes

Définition 1. Une fonction $f\colon \mathbb{R}^d \to \bar{\mathbb{R}}$ est dite *positivement homogène de degré* $a \in \mathbb{R}$ lorsque

$$\forall \mathbf{x} \in \mathbb{R}^d, \quad \forall t > 0, \quad f(t\mathbf{x}) = t^a f(\mathbf{x}).$$

Une fonction positivement homogène de degré 1 est dite *positivement homogène*.

Remarque 1. Si $f\colon \mathbb{R}^d \to]-\infty, \infty]$ est positivement homogène de degré $a \neq 0$, alors $|f(\mathbf{0})| \in \{0, \infty\}$. En effet, pour tout $t > 0$,

$$f(\mathbf{0}) = f(t\mathbf{0}) = t^a f(\mathbf{0}). \blacksquare$$

Proposition 7. Soit $f\colon \mathbb{R}^d \to]-\infty, \infty]$ une fonction positivement homogène. Il y a équivalence entre

(a) f est convexe ;

(b) pour tous $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$, $f(\mathbf{x} + \mathbf{y}) \leq f(\mathbf{x}) + f(\mathbf{y})$;

(c) pour tout $m \geq 2$ et tous $\mathbf{x}_1, \dots, \mathbf{x}_m$,

$$f(\mathbf{x}_1 + \cdots + \mathbf{x}_m) \leq f(\mathbf{x}_1) + \cdots + f(\mathbf{x}_m).$$

En particulier, si f est convexe propre et positivement homogène, alors

$$\forall \mathbf{x} \in \mathbb{R}^d, \quad f(\mathbf{x}) + f(-\mathbf{x}) \geq 0.$$

Démonstration. L'équivalence entre (b) et (c) est évidente. Pour voir que (a) implique (b), il suffit d'écrire que

$$f(\mathbf{x} + \mathbf{y}) = f\left(2\frac{\mathbf{x} + \mathbf{y}}{2}\right) = 2f\left(\frac{\mathbf{x}}{2} + \frac{\mathbf{y}}{2}\right) \leq 2\left(\frac{f(\mathbf{x})}{2} + \frac{f(\mathbf{y})}{2}\right) = f(\mathbf{x}) + f(\mathbf{y}).$$

Réciproquement, supposons la condition (b) satisfaite. Soient $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ et soit $\lambda \in]0,1[$. Si $f(\mathbf{x}) < \alpha \in \mathbb{R}$ et $f(\mathbf{y}) < \beta \in \mathbb{R}$, alors

$$\begin{aligned} f\big((1-\lambda)\mathbf{x} + \lambda\mathbf{y}\big) &\leq f\big((1-\lambda)\mathbf{x}\big) + f\big(\lambda\mathbf{y}\big) \\ &= (1-\lambda)f(\mathbf{x}) + \lambda f(\mathbf{y}) \\ &< (1-\lambda)\alpha + \lambda\beta, \end{aligned}$$

et le théorème 18 montre que f est convexe. Supposons enfin que f est convexe propre et positivement homogène. Alors, pour tout $\mathbf{x} \in \mathbb{R}^d$,

$$f(\mathbf{x}) + f(-\mathbf{x}) \geq f(\mathbf{x} - \mathbf{x}) = f(\mathbf{0}) \in \{0, \infty\},$$

où l'on a fait usage de la remarque 1. ■

Proposition 8. Soit $f \colon \mathbb{R}^d \to]-\infty, \infty]$ une fonction convexe positivement homogène. Soient V un sous-espace vectoriel de $\mathbb{R}^d$ et $\{\mathbf{e}_1, \ldots, \mathbf{e}_m\}$ une base de V. Il y a équivalence entre

(a) f est linéaire sur V ;
(b) pour tout $\mathbf{x} \in V$, $f(-\mathbf{x}) = -f(\mathbf{x})$;
(c) pour tout $j \in \{1, \ldots, m\}$, $f(-\mathbf{e}_j) = -f(\mathbf{e}_j)$.

Démonstration. Il est évident que (a) implique (b) et que (b) implique (c). Supposons donc (c). Alors, pour tout $j \in \{1, \ldots, m\}$, $f(\mathbf{e}_j) \in \mathbb{R}$ (puisque la valeur $-\infty$ est interdite). Montrons tout d'abord que

$$\forall \lambda \in \mathbb{R}, \quad \forall j \in \{1, \ldots, m\}, \quad f(\lambda\mathbf{e}_j) = \lambda f(\mathbf{e}_j).$$

Pour $\lambda > 0$, c'est la positive homogénéité. Pour $\lambda = 0$, on a :

$$f(0\mathbf{e}_j) = f(\mathbf{0}) = f(\mathbf{e}_j - \mathbf{e}_j) \leq f(\mathbf{e}_j) + f(-\mathbf{e}_j) = f(\mathbf{e}_j) - f(\mathbf{e}_j) = 0,$$

où l'inégalité provient de la proposition 7. Puisque $f(\mathbf{0}) \in \{0, \infty\}$ d'après la remarque 1, on voit que nécessairement $f(0\mathbf{e}_j) = 0 = 0f(\mathbf{e}_j)$. Pour $\lambda < 0$, on a :

$$f(\lambda\mathbf{e}_j) = f(-|\lambda|\,\mathbf{e}_j) = |\lambda|\, f(-\mathbf{e}_j) = -|\lambda|\, f(\mathbf{e}_j) = \lambda f(\mathbf{e}_j).$$

Pour montrer (a), nous devons montrer que pour tout $\mathbf{x} = \lambda_1\mathbf{e}_1 + \cdots + \lambda_m\mathbf{e}_m$,

$$f(\mathbf{x}) = \lambda_1 f(\mathbf{e}_1) + \cdots + \lambda_m f(\mathbf{e}_m) = f(\lambda_1\mathbf{e}_1) + \cdots + f(\lambda_m\mathbf{e}_m).$$

Mais ceci découle des inégalités suivantes :

$$\begin{aligned} f(\lambda_1\mathbf{e}_1) + \cdots + f(\lambda_m\mathbf{e}_m) &\geq f(\mathbf{x}) \\ &\geq -f(-\mathbf{x}) \\ &\geq -f(-\lambda_1\mathbf{e}_1) - \cdots - f(-\lambda_m\mathbf{e}_m) \\ &= f(\lambda_1\mathbf{e}_1) + \cdots + f(\lambda_m\mathbf{e}_m), \end{aligned}$$

qui résultent toutes de la proposition 7. ■

5.2 Dérivées directionnelles

Soient $f\colon \mathbb{R}^d \to \bar{\mathbb{R}}$ une fonction quelconque et $\mathbf{x} \in \mathbb{R}^d$ tel que $f(\mathbf{x}) \in \mathbb{R}$. Étant donné $\mathbf{y} \in \mathbb{R}^d$, le quotient

$$\frac{f(\mathbf{x}+\lambda\mathbf{y}) - f(\mathbf{x})}{\lambda} \tag{5.1}$$

est bien défini pour tout $\lambda \in \mathbb{R}_+^*$. Si ce quotient tend vers une limite, dans $\bar{\mathbb{R}}$, lorsque $\lambda \downarrow 0$, la limite s'appelle la *dérivée directionnelle* de f en $\mathbf{x}$ de direction $\mathbf{y}$. On la note $f'(\mathbf{x};\mathbf{y})$.

Théorème 26. Soit $f\colon \mathbb{R}^d \to \bar{\mathbb{R}}$ une fonction quelconque, et soit $\mathbf{x} \in \mathbb{R}^d$ tel que $f(\mathbf{x}) \in \mathbb{R}$. Si f est différentiable en $\mathbf{x}$, alors le quotient (5.1) tend vers une limite pour tout $\mathbf{y} \in \mathbb{R}^d$, c'est-à-dire, $f'(\mathbf{x};\cdot)$ est bien définie sur $\mathbb{R}^d$. De plus, $f'(\mathbf{x};\cdot)$ est une fonction linéaire, et

$$f'(\mathbf{x};\cdot) = \langle \nabla f(\mathbf{x}), \cdot \rangle .$$

Démonstration. Soit $\langle \boldsymbol{\xi}, \cdot\rangle$ la différentielle de f en $\mathbf{x}$. Alors, pour tout $\mathbf{z} \in \mathbb{R}^d$, $f(\mathbf{z}) = f(\mathbf{x}) + \langle \boldsymbol{\xi}, \mathbf{z}-\mathbf{x}\rangle + o(\|\mathbf{z}-\mathbf{x}\|)$. Autrement dit,

$$\frac{f(\mathbf{z}) - f(\mathbf{x}) - \langle \boldsymbol{\xi}, \mathbf{z}-\mathbf{x}\rangle}{\|\mathbf{z}-\mathbf{x}\|} \to 0 \quad \text{lorsque} \quad \mathbf{z} \to \mathbf{x}.$$

En remplaçant $\mathbf{z}$ par $\mathbf{x}+\lambda\mathbf{y}$, avec $\lambda > 0$ et $\mathbf{y} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$, on obtient :

$$\frac{f(\mathbf{x}+\lambda\mathbf{y}) - f(\mathbf{x}) - \langle \boldsymbol{\xi}, \lambda\mathbf{y}\rangle}{\lambda\,\|\mathbf{y}\|} \to 0 \quad \text{lorsque} \quad \lambda \downarrow 0.$$

Ceci montre que la limite de $\lambda^{-1}(f(\mathbf{x}+\lambda\mathbf{y}) - f(\mathbf{x}))$ existe et vaut $\langle \boldsymbol{\xi}, \mathbf{y}\rangle$, et que $f'(\mathbf{x};\cdot)$ est linéaire. Finalement, en notant $(\mathbf{e}_k)_{k\in\{1,\dots,d\}}$ la base canonique de $\mathbb{R}^d$,

$$\xi_k = \langle \boldsymbol{\xi}, \mathbf{e}_k\rangle = \lim_{\lambda\downarrow 0} \frac{f(\mathbf{x}+\lambda\mathbf{e}_k) - f(\mathbf{x})}{\lambda} = f_k'(\mathbf{x}),$$

où $f_k'(\mathbf{x})$ désigne la dérivée partielle de f par rapport à la k-ième variable. Donc $\boldsymbol{\xi} = \nabla f(\mathbf{x})$. ■

Nous verrons plus bas que la limite existe aussi lorsque f est non différentiable mais convexe. Pour le montrer, nous aurons besoin des quelques lemmes qui suivent.

Lemme 5. Soient $E \subset \mathbb{R}^{d+1}$ l'épigraphe d'une fonction $f\colon \mathbb{R}^d \to \bar{\mathbb{R}}$ et t un réel strictement positif. Alors tE est l'épigraphe de la fonction φ_t définie par

$$\varphi_t(\mathbf{x}) = tf\left(\frac{\mathbf{x}}{t}\right).$$

Démonstration. On a :

$$\begin{aligned}(\mathbf{x},\alpha)\in\operatorname{epi}\varphi_t &\iff \alpha\geq\varphi_t(\mathbf{x})=tf\left(\frac{\mathbf{x}}{t}\right)\\ &\iff \frac{\alpha}{t}\geq f\left(\frac{\mathbf{x}}{t}\right)\\ &\iff \frac{1}{t}(\mathbf{x},\alpha)\in\operatorname{epi}f.\ \blacksquare\end{aligned}$$

La fonction φ_t est aussi notée ft, et l'on parle parfois de *multiplication scalaire à droite*, ou de *multiplication épigraphique*.

Lemme 6. La limite ponctuelle d'une suite de fonctions convexes de $\mathbb{R}^d$ dans $\bar{\mathbb{R}}$ est convexe.

Démonstration. Soit $(f_k)_{k\in\mathbb{N}^*}$ une suite de fonctions convexes qui converge ponctuellement vers une fonction f. Soient $\mathbf{x},\mathbf{y}\in\mathbb{R}^d$, $\lambda\in\,]0,1[$ et α,β deux réels tels que $\alpha>f(\mathbf{x})$ et $\beta>f(\mathbf{y})$. Pour k suffisamment grand, $\alpha>f_k(\mathbf{x})$ et $\beta>f_k(\mathbf{y})$, de sorte que, d'après le théorème 18,

$$f_k\big((1-\lambda)\mathbf{x}+\lambda\mathbf{y}\big)\leq(1-\lambda)\alpha+\lambda\beta.$$

Mais l'inégalité reste vraie par passage à la limite, et le même théorème montre que f est convexe. $\blacksquare$

Lemme 7. Soit $C\subset\mathbb{R}^d$ un convexe qui contient l'origine. Alors l'application de $\mathbb{R}_+^*$ dans $\mathscr{C}(\mathbb{R}^d)$ qui à λ fait correspondre λC est croissante. Autrement dit,

$$\lambda_1>\lambda_2\Longrightarrow\lambda_1C\supset\lambda_2C.$$

Démonstration. Soit $\mathbf{x}\in\lambda_2C$. Alors $\mathbf{x}=\lambda_2\mathbf{x}_0$ avec $\mathbf{x}_0\in C$. Puisque $\mathbf{0}\in C$ et $\lambda_2/\lambda_1\in\,]0,1[$,

$$\mathbf{x}=\lambda_1\frac{\lambda_2}{\lambda_1}\mathbf{x}_0=\lambda_1\left(\left(1-\frac{\lambda_2}{\lambda_1}\right)\mathbf{0}+\frac{\lambda_2}{\lambda_1}\mathbf{x}_0\right)\in\lambda_1C.\ \blacksquare$$

Théorème 27. Soient $f\colon\mathbb{R}^d\to\bar{\mathbb{R}}$ une fonction convexe et $\mathbf{x}\in\mathbb{R}^d$ tel que $f(\mathbf{x})\in\mathbb{R}$. Alors, pour tout $\mathbf{y}\in\mathbb{R}^d$, la fonction

$$\lambda\mapsto\frac{f(\mathbf{x}+\lambda\mathbf{y})-f(\mathbf{x})}{\lambda},\quad\lambda\in\mathbb{R}_+^*$$

décroît lorsque λ décroît, de sorte que sa limite lorsque $\lambda\downarrow0$ existe (dans $\bar{\mathbb{R}}$). Autrement dit,

$$f'(\mathbf{x};\mathbf{y})=\lim_{\lambda\downarrow0}\frac{f(\mathbf{x}+\lambda\mathbf{y})-f(\mathbf{x})}{\lambda}=\inf_{\lambda\in\mathbb{R}_+^*}\frac{f(\mathbf{x}+\lambda\mathbf{y})-f(\mathbf{x})}{\lambda}.$$

Démonstration. Considérons la fonction $h_{\mathbf{x}}(\mathbf{y}) := f(\mathbf{x}+\mathbf{y}) - f(\mathbf{x})$. L'épigraphe de $h_{\mathbf{x}}$ s'obtient par translation de vecteur $(-\mathbf{x}, -f(\mathbf{x}))$ de l'épigraphe de f. On voit aussi que $h_{\mathbf{x}}(\mathbf{0}) = 0$, de sorte que $(\mathbf{0}, 0) \in \operatorname{epi} h_{\mathbf{x}}$. On considère alors la multiplication épigraphique de $h_{\mathbf{x}}$ par λ^{-1}, avec $\lambda \in \mathbb{R}_+^*$:

$$h_{\mathbf{x}}\lambda^{-1}(\mathbf{y}) = \frac{f(\mathbf{x}+\lambda\mathbf{y}) - f(\mathbf{x})}{\lambda}.$$

Comme $(\mathbf{0}, 0) \in \operatorname{epi} h_{\mathbf{x}}$, le lemme 7 montre que $\lambda^{-1} \operatorname{epi} h_{\mathbf{x}} = \operatorname{epi}(h_{\mathbf{x}}\lambda^{-1})$ croît lorsque λ décroît. À la croissance de $\operatorname{epi}(h_{\mathbf{x}}\lambda^{-1})$ correspond bien sûr la décroissance de $h_{\mathbf{x}}\lambda^{-1}(\mathbf{y})$. ∎

Théorème 28. Soient $f\colon \mathbb{R}^d \to \bar{\mathbb{R}}$ une fonction convexe et $\mathbf{x} \in \mathbb{R}^d$ tel que $f(\mathbf{x}) \in \mathbb{R}$. Alors $f'(\mathbf{x}; \cdot)$ est convexe et positivement homogène. De plus,

$$f'(\mathbf{x}; \mathbf{0}) = 0 \quad \text{et} \quad -f'(\mathbf{x}; -\mathbf{y}) \leq f'(\mathbf{x}; \mathbf{y}) \quad \text{pour tout} \quad \mathbf{y} \in \mathbb{R}^d.$$

Démonstration. La convexité de $f'(\mathbf{x}; \cdot)$ est une conséquence immédiate du lemme 6 puisque, pour tout $\lambda > 0$, $h_{\mathbf{x}}\lambda^{-1}(\mathbf{y})$ est une fonction convexe de $\mathbf{y}$. Montrons maintenant la positive homogénéité de $f'(\mathbf{x}; \cdot)$. Pour tout $t > 0$,

$$h_{\mathbf{x}}\lambda^{-1}(t\mathbf{y}) = t\frac{f(\mathbf{x}+t\lambda\mathbf{y}) - f(\mathbf{x})}{t\lambda} = t\frac{f(\mathbf{x}+\lambda'\mathbf{y}) - f(\mathbf{x})}{\lambda'},$$

où $\lambda' = t\lambda$. En faisant tendre λ vers zéro dans l'équation ci-dessus, on obtient :

$$f'(\mathbf{x}; t\mathbf{y}) = tf'(\mathbf{x}; \mathbf{y}).$$

Maintenant, pour tout $\lambda > 0$, $h_{\mathbf{x}}\lambda^{-1}(\mathbf{0}) = 0$. Il s'ensuit que $f'(\mathbf{x}; \mathbf{0}) = 0$. Soit enfin $\mathbf{y} \in \mathbb{R}^d$, et montrons que $-f'(\mathbf{x}; -\mathbf{y}) \leq f'(\mathbf{x}; \mathbf{y})$. La convexité de $f'(\mathbf{x}; \cdot)$ implique, *via* le théorème 18, que

$$\left.\begin{array}{r} \mu_1 > f'(\mathbf{x}; -\mathbf{y}) \\ \mu_2 > f'(\mathbf{x}; \mathbf{y}) \end{array}\right\} \Rightarrow \frac{\mu_1}{2} + \frac{\mu_2}{2} \geq f'\left(\mathbf{x}; -\frac{\mathbf{y}}{2} + \frac{\mathbf{y}}{2}\right) = f'(\mathbf{x}; \mathbf{0}) = 0 \Rightarrow \mu_1 \geq -\mu_2.$$

Par conséquent,

$$\begin{array}{rcl} f'(\mathbf{x}; -\mathbf{y}) & = & \inf\{\mu_1 \,|\, \mu_1 > f'(\mathbf{x}; -\mathbf{y})\} \\ & \geq & \sup\{-\mu_2 \,|\, \mu_2 > f'(\mathbf{x}; \mathbf{y})\} \\ & = & -\inf\{\mu_2 \,|\, \mu_2 > f'(\mathbf{x}; \mathbf{y})\} \\ & = & -f'(\mathbf{x}; \mathbf{y}). \ \blacksquare \end{array}$$

5.3 Sous-différentiabilité

Soient $f\colon \mathbb{R}^d \to \bar{\mathbb{R}}$ et $\mathbf{x} \in \mathbb{R}^d$ un point tel que $f(\mathbf{x}) \in \mathbb{R}$. Une minorante affine m de f est dite *exacte* en $\mathbf{x}$ si $m(\mathbf{x}) = f(\mathbf{x})$. La fonction f est dite *Sous-différentiable* en $\mathbf{x}$ si elle admet une minorante affine exacte en $\mathbf{x}$, autrement dit, s'il existe $\boldsymbol{\xi} \in \mathbb{R}^d$ tel que

$$\forall \mathbf{y} \in \mathbb{R}^d, \quad f(\mathbf{y}) \geq m(\mathbf{y}) := f(\mathbf{x}) + \langle \boldsymbol{\xi}, \mathbf{y} - \mathbf{x} \rangle. \tag{5.2}$$

Un tel vecteur $\boldsymbol{\xi}$ est appelé un *sous-gradient* de f en $\mathbf{x}$ (ou une sous-dérivée, si $d = 1$). L'inégalité (5.2) est appelée *l'inégalité du sous-gradient*. L'ensemble de tous les sous-gradients de f en $\mathbf{x}$ est appelé le sous-différentiel de f en $\mathbf{x}$, et on le note $\partial f(\mathbf{x})$.

Proposition 9. Soient $f\colon \mathbb{R}^d \to \bar{\mathbb{R}}$ et $\mathbf{x} \in \mathbb{R}^d$. Alors $\partial f(\mathbf{x})$ est un ensemble convexe fermé (éventuellement vide).

DÉMONSTRATION. Montrons tout d'abord la convexité. Soient $\boldsymbol{\xi}, \boldsymbol{\eta} \in \partial f(\mathbf{x})$ et $\lambda \in [0, 1]$. Pour tout $\mathbf{y} \in \mathbb{R}^d$,

$$\begin{aligned} f(\mathbf{y}) &\geq f(\mathbf{x}) + \langle \boldsymbol{\xi}, \mathbf{y} - \mathbf{x} \rangle, \\ f(\mathbf{y}) &\geq f(\mathbf{x}) + \langle \boldsymbol{\eta}, \mathbf{y} - \mathbf{x} \rangle. \end{aligned}$$

En multipliant la première inégalité par $1 - \lambda$ et la seconde par λ, on obtient :

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \langle (1 - \lambda)\boldsymbol{\xi} + \lambda \boldsymbol{\eta}, \mathbf{y} - \mathbf{x} \rangle,$$

ce qui montre que $(1-\lambda)\boldsymbol{\xi} + \lambda\boldsymbol{\eta}$ appartient à $\partial f(\mathbf{x})$. Reste à montrer que $\partial f(\mathbf{x})$ est fermé. Soit $(\boldsymbol{\xi}_k)_{k \in \mathbb{N}^*} \subset \partial f(\mathbf{x})$ une suite qui converge vers un point $\boldsymbol{\xi}$. Pour tout $\mathbf{y} \in \mathbb{R}^d$ et tout $k \in \mathbb{N}^*$,

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \boldsymbol{\xi}_k, \mathbf{y} - \mathbf{x} \rangle.$$

Puisque $\langle \cdot, \mathbf{y} - \mathbf{x} \rangle$ est continue, on peut passer à la limite dans l'inégalité ci-dessus, ce qui montre que $\boldsymbol{\xi} \in \partial f(\mathbf{x})$. ∎

Remarquons qu'une fonction convexe impropre n'est sous-différentiable en aucun point. Pour une fonction convexe propre, $\partial f(\mathbf{x}) = \emptyset$ pour tout $\mathbf{x} \notin \operatorname{dom} f$.

Théorème 29. Soit $f\colon \mathbb{R}^d \to]-\infty, \infty]$ une fonction convexe propre. Pour tout $\mathbf{x} \in \operatorname{ri}(\operatorname{dom} f)$, $\partial f(\mathbf{x}) \neq \emptyset$.

DÉMONSTRATION. Soit $\mathbf{x} \in \operatorname{ri}(\operatorname{dom} f)$. D'après le théorème 20, $(\mathbf{x}, f(\mathbf{x})) \in \operatorname{epi} f \setminus \operatorname{ri}(\operatorname{epi} f)$. D'après le théorème 14, il existe un hyperplan séparant $\operatorname{epi} f$ et $\{(\mathbf{x}, f(\mathbf{x}))\}$ proprement. Supposons, en vue d'obtenir une contradiction, que H soit vertical. Notons comme d'habitude P la projection canonique de $\mathbb{R}^{d+1}$ sur $\mathbb{R}^d$. Alors PH est un hyperplan de $\mathbb{R}^d$ séparant

$$P(\operatorname{epi} f) = \operatorname{dom} f \quad \text{et} \quad P\{(\mathbf{x}, f(\mathbf{x}))\} = \{\mathbf{x}\}$$

proprement. Le théorème 14 implique alors que $\mathbf{x} \notin \mathrm{ri}(\mathrm{dom}\, f)$, en contradiction avec l'hypothèse. L'hyperplan H est donc le graphe d'une minorante affine m de f. Montrons enfin que cette minorante est exacte en $\mathbf{x}$. Puisque m minore f, $(\mathbf{x}, f(\mathbf{x}))$ appartient au demi-espace supérieur fermé associé à H. Mais puisque H sépare $(\mathbf{x}, f(\mathbf{x}))$ et epi f, $(\mathbf{x}, f(\mathbf{x}))$ appartient aussi au demi-espace fermé opposé. Par conséquent, $(\mathbf{x}, f(\mathbf{x})) \in H$. ▪

Théorème 30. Soit $f\colon \mathbb{R}^d \to \bar{\mathbb{R}}$ une fonction convexe, et soit $\mathbf{x} \in \mathbb{R}^d$ tel que $f(\mathbf{x}) \in \mathbb{R}$. Il y a équivalence entre

(a) $\boldsymbol{\xi} \in \partial f(\mathbf{x})$;

(b) $f'(\mathbf{x}; \mathbf{y}) \geq \langle \boldsymbol{\xi}, \mathbf{y} \rangle$ pour tout $\mathbf{y} \in \mathbb{R}^d$.

Démonstration. on a :

$$\begin{aligned}
\text{(a)} \quad &\iff \quad \forall \mathbf{z} \in \mathbb{R}^d,\ f(\mathbf{z}) \geq f(\mathbf{x}) + \langle \boldsymbol{\xi}, \mathbf{z} - \mathbf{x} \rangle \\
&\iff \quad \forall \mathbf{y} \in \mathbb{R}^d,\ \forall \lambda > 0,\ f(\mathbf{x} + \lambda \mathbf{y}) \geq f(\mathbf{x}) + \langle \boldsymbol{\xi}, \lambda \mathbf{y} \rangle \\
&\iff \quad \forall \mathbf{y} \in \mathbb{R}^d,\ \forall \lambda > 0,\ \frac{f(\mathbf{x} + \lambda \mathbf{y}) - f(\mathbf{x})}{\lambda} \geq \langle \boldsymbol{\xi}, \mathbf{y} \rangle \\
&\iff \quad \forall \mathbf{y} \in \mathbb{R}^d,\ \inf_{\lambda \in \mathbb{R}_+^*} \frac{f(\mathbf{x} + \lambda \mathbf{y}) - f(\mathbf{x})}{\lambda} \geq \langle \boldsymbol{\xi}, \mathbf{y} \rangle \\
&\iff \quad \text{(b)},
\end{aligned}$$

où la dernière équivalence résulte du théorème 27. ▪

Corollaire 11. Soit $f\colon \mathbb{R}^d \to]-\infty, \infty]$ une fonction convexe, et soit $\mathbf{x} \in \mathbb{R}^d$ tel que $f(\mathbf{x}) \in \mathbb{R}$. Si f est différentiable en $\mathbf{x}$, alors $\partial f(\mathbf{x}) = \{\nabla f(\mathbf{x})\}$.

Démonstration. D'après le théorème 26, $f'(\mathbf{x}; \mathbf{y})$ existe pour tout $\mathbf{y}$, et

$$f'(\mathbf{x}; \mathbf{y}) = \langle \nabla f(\mathbf{x}), \mathbf{y} \rangle .$$

Le théorème 30 montre alors que $\nabla f(\mathbf{x}) \in \partial f(\mathbf{x})$. D'autre part,

$$\begin{aligned}
\boldsymbol{\xi} \in \partial f(\mathbf{x}) \quad &\iff \quad \forall \mathbf{y} \in \mathbb{R}^d, \langle \nabla f(\mathbf{x}), \mathbf{y} \rangle \geq \langle \boldsymbol{\xi}, \mathbf{y} \rangle \\
&\iff \quad \nabla f(\mathbf{x}) = \boldsymbol{\xi},
\end{aligned}$$

où la deuxième équivalence résulte encore du théorème 30. Donc $\nabla f(\mathbf{x})$ est l'unique sous-gradient de f en $\mathbf{x}$. ▪

La réciproque du corollaire précédent est vraie : si $\partial f(\mathbf{x})$ est le singleton $\{\boldsymbol{\xi}\}$, alors f est différentiable en $\mathbf{x}$ et $\nabla f(\mathbf{x}) = \boldsymbol{\xi}$. La démonstration de ce résultat est omise dans ce cours.

Proposition 10. Soient $f\colon \mathbb{R}^d \to \bar{\mathbb{R}}$ et C un ouvert convexe de $\mathbb{R}^d$. On suppose que f est finie et différentiable sur C. Alors, il y a équivalence entre

(a) f est convexe sur C ;
(b) pour tous $\mathbf{x}, \mathbf{y} \in C$, $f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle$.

Démonstration. Le fait que (a) implique (b) résulte du corollaire 11 et de l'inégalité du sous-différentiel (appliquée à la fonction $f + \delta(\cdot|C)$). Réciproquement, supposons la condition (b) satisfaite, et soient

$$\mathbf{y}, \mathbf{z} \in C, \quad \lambda \in]0,1[\quad \text{et} \quad \mathbf{x} := (1-\lambda)\mathbf{y} + \lambda\mathbf{z}.$$

La condition (b) implique que

$$\begin{array}{rrcl} & f(\mathbf{y}) & \geq & f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \\ \text{et} & f(\mathbf{z}) & \geq & f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{z} - \mathbf{x} \rangle, \end{array}$$

d'où l'on tire que

$$(1-\lambda)f(\mathbf{y}) + \lambda f(\mathbf{z}) \geq f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), (1-\lambda)\mathbf{y} + \lambda\mathbf{z} - \mathbf{x} \rangle = f(\mathbf{x}). \blacksquare$$

Théorème 31. Soient $f\colon \mathbb{R}^d \to \bar{\mathbb{R}}$ et C un ouvert convexe de $\mathbb{R}^d$. On suppose que f est finie et deux fois différentiable sur C. Alors, il y a équivalence entre

(a) f est convexe sur C ;
(b) pour tout $\mathbf{x} \in C$, $\nabla^2 f(\mathbf{x})$ est semi-définie positive.

Démonstration. Montrons tout d'abord le théorème dans le cas où $d = 1$. Soient $\varphi\colon \mathbb{R} \to \bar{\mathbb{R}}$ et $I \subset \mathbb{R}$ un intervalle ouvert sur lequel φ est finie et deux fois dérivable. Supposons φ convexe sur I, et soit $t \in I$. D'après la proposition précédente, pour tout h tel que $t + h \in I$,

$$\varphi(t+h) \geq \varphi(t) + \varphi'(t)h. \tag{5.3}$$

Or, $\varphi(t+h) = \varphi(t) + \varphi'(t)h + \varphi''(t)h^2/2 + o(h^2)$, et l'équation (5.3) montre alors que, pour tout h tel que $t + h \in I$,

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

Cette dernière inégalité implique bien sûr que $\varphi''(t) \geq 0$. Réciproquement, supposons que, pour tout $t \in I$, $\varphi''(t) \geq 0$. Alors, φ' est croissante sur I. Pour tous $x, y, z \in I$ tels que $x < y < z$, nous avons :

$$\frac{\varphi(y) - \varphi(x)}{y - x} \leq \sup_{t \in]x,y[} \varphi'(t) \leq \inf_{u \in]y,z[} \varphi'(u) \leq \frac{\varphi(z) - \varphi(y)}{z - y}.$$

Ici, les première et troisième inégalités résultent du théorème des accroissements finis, et la deuxième de la croissance de φ'. La proposition 6 montre alors que φ est convexe sur I.

Montrons maintenant le cas général. Supposons (a), et soit $\mathbf{x} \in C$. Pour tout $\mathbf{y} \in \mathbb{R}^d$, la fonction $\varphi(t) := f(\mathbf{x} + t\mathbf{y})$ est convexe sur l'intevalle ouvert

$$I_{\mathbf{x},\mathbf{y}} := \{t \in \mathbb{R} \mid \mathbf{x} + t\mathbf{y} \in C\} .$$

On vérifie facilement que φ est aussi deux fois dérivable sur $I_{\mathbf{x},\mathbf{y}}$, de dérivée seconde $\varphi''(t) = \langle \nabla^2 f(\mathbf{x} + t\mathbf{y})\mathbf{y}, \mathbf{y}\rangle$. D'après la première partie de la preuve, $\varphi''(t) \geq 0$ pour tout $t \in I_{\mathbf{x},\mathbf{y}}$. En particulier,

$$\varphi''(0) = \left\langle \nabla^2 f(\mathbf{x})\mathbf{y}, \mathbf{y}\right\rangle \geq 0.$$

Cette dernière inégalité étant satisfaite pour tout $\mathbf{y} \in \mathbb{R}^d$, il s'ensuit que $\nabla^2 f(\mathbf{x})$ est semi-définie positive. Réciproquement, supposons (b) et montrons que f est convexe sur C. Soient $\mathbf{x}, \mathbf{z} \in C$ et $\lambda \in [0, 1]$. Soient

$$\mathbf{y} := \mathbf{z} - \mathbf{x} \quad \text{et} \quad \varphi_{\mathbf{x},\mathbf{y}}(\lambda) := f(\mathbf{x} + \lambda\mathbf{y}).$$

Pour tout $\lambda \in I_{\mathbf{x},\mathbf{y}} := \{t \in \mathbb{R} \,|\, \mathbf{x} + t\mathbf{y} \in C\}$,

$$\varphi''_{\mathbf{x},\mathbf{y}}(\lambda) = \left\langle \nabla^2 f(\mathbf{x} + \lambda\mathbf{y})\mathbf{y}, \mathbf{y}\right\rangle \geq 0.$$

La première partie de la preuve implique que $\varphi_{\mathbf{x},\mathbf{y}}$ est convexe sur $I_{\mathbf{x},\mathbf{y}} \supset [0, 1]$. Donc

$$\begin{aligned} f((1-\lambda)\mathbf{x} + \lambda\mathbf{z}) &= \varphi_{\mathbf{x},\mathbf{y}}(\lambda) \\ &\leq (1-\lambda)\varphi_{\mathbf{x},\mathbf{y}}(0) + \lambda\varphi_{\mathbf{x},\mathbf{y}}(1) \\ &= (1-\lambda)f(\mathbf{x}) + \lambda f(\mathbf{z}). \blacksquare \end{aligned}$$

5.4 Conjugaison convexe

Soit f une fonction de $\mathbb{R}^d$ dans $\bar{\mathbb{R}}$. La *conjuguée convexe* de f est la fonction $f^\star$ définie sur $\mathbb{R}^d$ par

$$f^\star(\boldsymbol{\xi}) = \sup_{\mathbf{x} \in \mathbb{R}^d} \{\langle \mathbf{x}, \boldsymbol{\xi}\rangle - f(\mathbf{x})\} .$$

C'est l'enveloppe supérieure de la famille des fonctions affines

$$l_{\mathbf{x}}(\boldsymbol{\xi}) = \langle \mathbf{x}, \boldsymbol{\xi}\rangle - f(\mathbf{x}).$$

L'épigraphe de $f^\star$ est donc l'intersection des épigraphes des fonctions $l_{\mathbf{x}}$, c'est-à-dire l'intersection d'une famille de demi-espaces (supérieurs) fermés de $\mathbb{R}^{d+1}$. L'ensemble epi $f^\star$ est donc convexe fermé, ce qui montre que $f^\star$ est une fonction convexe *sci*. Bien sûr, on peut étendre la famille des $l_{\mathbf{x}}$ ci-dessus aux fonctions affines $\langle \mathbf{x}, \cdot\rangle - \alpha$, avec $\alpha \geq f(\mathbf{x})$, sans que le *supremum* ne soit modifié.
La biconjuguée de f est la conjuguée convexe de $f^\star$, c'est-à-dire la fonction $f^{\star\star}$ de $\mathbb{R}^d$ dans $\bar{\mathbb{R}}$ définie par

$$f^{\star\star}(\mathbf{x}) = \sup_{\boldsymbol{\xi} \in \mathbb{R}^d} \{\langle \mathbf{x}, \boldsymbol{\xi}\rangle - f^\star(\boldsymbol{\xi})\} .$$

D'après ce qui précède, epi $f^{\star\star}$ est l'intersection des épigraphes des fonctions affines

$$l(\cdot) = \langle \cdot, \boldsymbol{\xi} \rangle - \alpha, \quad \text{avec} \quad \alpha \geq f^{\star}(\boldsymbol{\xi}).$$

Soit $\mathscr{L} := \{l \colon \mathbf{x} \mapsto \langle \boldsymbol{\xi}, \mathbf{x} \rangle - \alpha \,|\, \alpha \geq f^{\star}(\boldsymbol{\xi})\}$. Par définition de la conjuguée convexe, $\alpha \geq f^{\star}(\boldsymbol{\xi})$ si et seulement si

$$\forall \mathbf{x} \in \mathbb{R}^d, \quad \alpha \geq \langle \boldsymbol{\xi}, \mathbf{x} \rangle - f(\mathbf{x}),$$

de sorte que $\mathscr{L}$ est en fait l'ensemble des minorantes affines de f. Par conséquent, $f^{\star\star}$ n'est autre que l'enveloppe supérieure des minorantes affines de f. Autrement dit, $f^{\star\star} = \operatorname{cl} f$, et si f est fermée (c'est-à-dire convexe, propre et *sci* ou identiquement égale à $+\infty$), alors $f^{\star\star} = f$.

Théorème 32. Soit $f \colon \mathbb{R}^d \to \bar{\mathbb{R}}$. Il y a équivalence entre

(a) $\boldsymbol{\xi} \in \partial f(\mathbf{x})$;

(b) $f(\mathbf{x}) + f^{\star}(\boldsymbol{\xi}) = \langle \boldsymbol{\xi}, \mathbf{x} \rangle$.

Si de plus $f^{\star\star} = f$, les conditions (a) et (b) sont encore équivalentes à

(c) $\mathbf{x} \in \partial f^{\star}(\boldsymbol{\xi})$.

Démonstration. Montrons tout d'abord l'équivalence entre (a) et (b) :

$$\begin{aligned}
\boldsymbol{\xi} \in \partial f(\mathbf{x}) \quad &\Longleftrightarrow \quad \forall \mathbf{z} \in \mathbb{R}^d, \quad f(\mathbf{z}) \geq f(\mathbf{x}) + \langle \boldsymbol{\xi}, \mathbf{z} - \mathbf{x} \rangle \\
&\Longleftrightarrow \quad \forall \mathbf{z} \in \mathbb{R}^d, \quad \langle \boldsymbol{\xi}, \mathbf{x} \rangle - f(\mathbf{x}) \geq \langle \boldsymbol{\xi}, \mathbf{z} \rangle - f(\mathbf{z}) \\
&\Longleftrightarrow \quad \langle \boldsymbol{\xi}, \mathbf{x} \rangle - f(\mathbf{x}) = \sup_{\mathbf{z} \in \mathbb{R}^d} \{\langle \boldsymbol{\xi}, \mathbf{z} \rangle - f(\mathbf{z})\} = f^{\star}(\boldsymbol{\xi}).
\end{aligned}$$

Si maintenant $f^{\star\star} = f$, (b) s'écrit $f^{\star}(\boldsymbol{\xi}) + f^{\star\star}(\mathbf{x}) = \langle \mathbf{x}, \boldsymbol{\xi} \rangle$, et la première partie du théorème montre alors que les assertions (b) et (c) sont équivalentes. ∎

La conjuguée concave d'une fonction $g \colon \mathbb{R}^d \to \bar{\mathbb{R}}$ est définie sur $\mathbb{R}^d$ par

$$g_{\star}(\boldsymbol{\xi}) = \inf_{\mathbf{x} \in \mathbb{R}^d} \{\langle \mathbf{x}, \boldsymbol{\xi} \rangle - g(\mathbf{x})\}.$$

Laissons au lecteur le soin de transposer à la notion de conjuguée concave tout ce qui a été établi à propos de la conjuguée convexe.
Le théorème suivant, qui porte le nom du mathématicien allemand Werner Fenchel (1905-1988), est la pierre angulaire de la théorie de la dualité en analyse convexe.

Théorème 33. [Fenchel] Soient f et g des fonctions de $\mathbb{R}^d$ dans $\bar{\mathbb{R}}$ respectivement convexe propre et concave propre telles que

$$\operatorname{ri} \operatorname{dom} f \cap \operatorname{ri} \operatorname{dom} g \neq \emptyset. \tag{5.4}$$

On a alors

$$\eta := \inf_{\mathbf{x} \in \mathbb{R}^d} \big\{ f(\mathbf{x}) - g(\mathbf{x}) \big\} = \sup_{\boldsymbol{\xi} \in \mathbb{R}^d} \big\{ g_{\star}(\boldsymbol{\xi}) - f^{\star}(\boldsymbol{\xi}) \big\},$$

et le *supremum* est atteint (c'est donc un *maximum*).

Démonstration. Par définition de la conjugaison, on a, pour tout $\mathbf{x} \in \mathbb{R}^d$,

$$\begin{aligned} f^\star(\boldsymbol{\xi}) &\geq \langle \mathbf{x}, \boldsymbol{\xi} \rangle - f(\mathbf{x}), \\ g_\star(\boldsymbol{\xi}) &\leq \langle \mathbf{x}, \boldsymbol{\xi} \rangle - g(\mathbf{x}). \end{aligned}$$

On en déduit que $f(\mathbf{x}) - g(\mathbf{x}) \geq g_\star(\boldsymbol{\xi}) - f^\star(\boldsymbol{\xi})$ pour tout $\mathbf{x} \in \mathbb{R}^d$, ce qui montre que

$$\eta = \inf_{\mathbf{x} \in \mathbb{R}^d} \big\{ f(\mathbf{x}) - g(\mathbf{x}) \big\} \geq \sup_{\boldsymbol{\xi} \in \mathbb{R}^d} \big\{ g_\star(\boldsymbol{\xi}) - f^\star(\boldsymbol{\xi}) \big\}. \tag{5.5}$$

Si $\eta = -\infty$, l'égalité est claire puisque, dans ce cas, $g_\star(\boldsymbol{\xi}) - f^\star(\boldsymbol{\xi}) = -\infty$ pour tout $\boldsymbol{\xi} \in \mathbb{R}^d$. Supposons donc $\eta > -\infty$. Comme $\operatorname{dom}(f-g) = \operatorname{dom} f \cap \operatorname{dom} g \neq \emptyset$ (cf. la condition (5.4)) on a aussi $\eta < \infty$, de sorte que $\eta \in \mathbb{R}$. D'après le théorème 20,

$$\begin{aligned} \operatorname{ri} \operatorname{epi} f &= \big\{ (\mathbf{x}, \alpha) \in \mathbb{R}^{d+1} \;\big|\; \mathbf{x} \in \operatorname{ri} \operatorname{dom} f,\ \alpha > f(\mathbf{x}) \big\} \\ \text{et} \quad \operatorname{ri} \operatorname{hypo}(g + \eta) &= \big\{ (\mathbf{x}, \alpha) \in \mathbb{R}^{d+1} \;\big|\; \mathbf{x} \in \operatorname{ri} \operatorname{dom} g,\ \alpha < g(\mathbf{x}) + \eta \big\}. \end{aligned}$$

Puisque $f(\mathbf{x}) - g(\mathbf{x}) \geq \eta$ pour tout $\mathbf{x} \in \mathbb{R}^d$, on ne peut pas avoir simultanément $\alpha > f(\mathbf{x})$ et $\alpha < g(\mathbf{x}) + \eta$, et donc

$$\operatorname{ri} \operatorname{epi} f \cap \operatorname{ri} \operatorname{hypo}(g + \eta) = \emptyset.$$

D'après le corollaire 6, il existe un hyperplan H séparant $\operatorname{epi} f$ et $\operatorname{hypo}(g + \eta)$ proprement. Or, H ne peut pas être vertical. En effet, si tel était le cas, la projection de H sur $\mathbb{R}^d$ serait un hyperplan séparant $\operatorname{dom} f$ et $\operatorname{dom} g$ proprement. D'après le corollaire 6, ceci contredirait l'hypothèse (5.4). L'hyperplan H est donc le graphe d'une fonction affine minorant f et majorant $g + \eta$: il existe un vecteur $\boldsymbol{\xi}_0$ et un réel ζ tel que

$$\forall \mathbf{x} \in \mathbb{R}^d, \quad f(\mathbf{x}) \geq \langle \boldsymbol{\xi}_0, \mathbf{x} \rangle - \zeta \geq g(\mathbf{x}) + \eta.$$

Par conséquent, nous avons

$$\begin{aligned} \zeta &\geq \sup_{\mathbf{x} \in \mathbb{R}^d} \big\{ \langle \boldsymbol{\xi}_0, \mathbf{x} \rangle - f(\mathbf{x}) \big\} = f^\star(\boldsymbol{\xi}_0) \\ \text{et} \quad \eta + \zeta &\leq \inf_{\mathbf{x} \in \mathbb{R}^d} \big\{ \langle \boldsymbol{\xi}_0, \mathbf{x} \rangle - g(\mathbf{x}) \big\} = g_\star(\boldsymbol{\xi}_0). \end{aligned}$$

En retranchant la première inégalité à la deuxième, nous obtenons

$$g_\star(\boldsymbol{\xi}_0) - f^\star(\boldsymbol{\xi}_0) \geq \eta \geq \sup_{\boldsymbol{\xi} \in \mathbb{R}^d} \big\{ g_\star(\boldsymbol{\xi}) - f^\star(\boldsymbol{\xi}) \big\}$$

(cf. (5.5)), d'où le résultat. ∎