Chapitre 4

Optimisation

4.1 Sous-différentiel d'une somme et optimalité

Proposition 68. *Soient $f, g : E \to \overline{\mathbb{R}}$ deux fonctions convexes propres et semi-continues inférieurement. Alors,*

$$\forall \phi \in E^*, \quad (f+g)^*(\phi) \leq (f^* \square g^*)(\phi) \tag{4.1}$$
$$\forall x \in E, \quad \partial(f+g)(x) \supseteq \partial f(x) + \partial g(x) \tag{4.2}$$

Démonstration. Soit ϕ_0 dans E^*. Par inégalité de Young, on a pour tout point $x \in E$ et toute forme linéaire $\phi \in E^*$,

$$\begin{aligned} &f(x) + f^*(\phi) \geq \langle x|\phi\rangle \\ &g(x) + g^*(\phi_0 - \phi) \geq \langle x|\phi_0 - \phi\rangle. \end{aligned}$$

En prenant la somme de ces inégalités on obtient

$$f(x) + g(x) + f^*(\phi) + g^*(\phi_0 - \phi) \geq \langle x|\phi_0\rangle,$$

ou encore

$$f^*(\phi) + g^*(\phi_0 - \phi) \geq \langle x|\phi_0\rangle - f(x) + g(x).$$

En prenant le suprémum sur x du membre de droite et l'infimum sur ϕ du membre de gauche on obtient l'inégalité (4.1). L'inclusion (4.2) est laissée en exercice. □

Théorème 69. *Soient $f, g : E \to \overline{\mathbb{R}}$ deux fonctions convexes semi-continues inférieurement. On suppose que la condition de qualification suivante est vérifiée :*

$$\exists x_0 \in \mathrm{dom}(f) \cap \mathrm{dom}(g), \ \textit{tel que } f \textit{ est continue en } x_0 \tag{Q}$$

Alors,

$$\forall \phi \in E^*, \quad (f+g)^*(\phi) = (f^* \Box g^*)(\phi) \tag{4.3}$$
$$\forall x \in E, \quad \partial(f+g)(x) = \partial f(x) + \partial g(x) \tag{4.4}$$

De plus, l'infimum dans la définition de $f^* \Box g^*$ est atteint.

Lemme 70. *Sous les mêmes hypothèses que celle du théorème, la fonction*

$$h(u) = \inf_{x \in E} f(x) + g(x+u)$$

est convexe et continue en l'origine.

Démonstration. Soit x_0 un point de E vérifiant la condition de qualification (Q). Par continuité de f en x_0, il existe une constante K et $r > 0$ telle que $f \leq K$ sur la boule $\mathrm{B}(x_0, r) \subseteq E$, et on a donc

$$\forall u \in \mathrm{B}(0, r), \quad h(u) \leq f(x_0) + g(x_0 + u) \leq f(x_0) + K.$$

La fonction h est bornée au voisinage de l'origine, et y est donc continue. $\square$

Démonstration de la formule (4.3). On commence par démontrer la formule (4.3) lorsque $\phi = 0$. Cette formule se réécrit de la manière suivante :

$$\begin{aligned} & (f+g)^*(0) = (f^* \Box g^*)(0) \\ \iff \quad & -\inf_{x \in E} f(x) + g(x) = \inf_{\phi \in E^*} f^*(\phi) + g^*(-\phi). \end{aligned}$$

Comme on sait déjà que $(f+g)^*(0) \leq (f^* \Box g^*)(0)$, il nous suffit de démontrer que

$$h(0) = \inf_{x \in E} f(x) + g(x) \leq -\inf_{\phi \in E^*} f^*(\phi) + g^*(-\phi), \tag{4.5}$$

où h est la fonction définie dans le lemme. Par continuité de h en l'origine, il existe une forme linéaire ϕ_0 dans le sous-différentiel $\partial h(0)$, ce qui signifie que

$$\forall u \in E, h(0) + \langle \phi_0 | u - 0 \rangle \leq h(u).$$

Alors,

$$\forall u \in E, \quad h(0) \leq h(u) - \langle \phi_0 | u \rangle = \inf_{x \in E} f(x) + g(x+u) - \langle \phi_0 | u \rangle$$

ou de manière équivalente

$$\begin{aligned} \forall x, u \in E, \quad h(0) &\leq f(x) + g(x+u) - \langle \phi_0 | u \rangle \\ &= f(x) + \langle \phi_0 | x \rangle + g(x+u) - \langle \phi_O | x+u \rangle \end{aligned}$$

Ainsi, en prenant d'abord l'infimum sur $u \in E$ dans le second membre on obtient

$$\begin{aligned} \forall x \in E, \quad h(0) &\leq f(x) + \langle \phi_0 | x \rangle + \inf_{u \in E} g(x+u) - \langle \phi_0 | x+u \rangle \\ &= f(x) + \langle \phi_0 | x \rangle - g^*(\phi_0), \end{aligned}$$

puis en prenant l'infimum sur x on a $h(0) \leq -f^*(-\phi_0) - g^*(\phi_0)$, qui implique bien l'inégalité voulue (4.5). On se rend de plus compte que ϕ_0 réalise l'infimum dans la définition de $(f^* \Box g^*)(0)$.

Nous passons maintenant à la démonstration dans le cas général. Soit ϕ_0 dans E^* et soit $h = g - \phi_0$. Un calcul simple montre que $h^*(\phi) = g^*(\phi + \phi_0)$ de sorte que

$$
\begin{aligned}
(f+h)^*(0) &= \sup_{x \in E} \phi_0(x) - f(x) + g(x) = (f+g)^*(\phi_0) \\
(f^* \Box h^*)(0) &= \inf_{\phi} f^*(\phi) + h^*(-\phi) = (f^* \Box g^*)(\phi_0)
\end{aligned}
$$

En appliquant l'égalité (4.3) à f et h en $\phi = 0$, on trouve l'égalité (4.3) pour f et g en $\phi = \phi_0$. $\Box$

Démonstration de la formule (4.4). Soit x un point de E et $\phi_0 \in \partial(f+g)(x)$. Il s'agit de démontrer que ϕ_0 peut s'écrire comme la somme d'un élément de $\partial f(x)$ et d'un élément de $\partial g(x)$. Si l'on pose $h = g - \phi_0$, on voit que $0 \in \partial(f+h)(x)$ et donc que x est un minimum global de $f + h$. En appliquant la formule (4.3) (et le fait que l'infimum est atteint), on sait qu'il existe un $\phi_1 \in E^*$ tel que

$$
f(x) + h(x) = \inf_{z \in E} f(z) + g(z) = -(f+h)^*(0) = -f^*(\phi_1) - h^*(-\phi_1)
$$

c'est-à-dire $A + B = 0$ où

$$
A = f(x) + f^*(\phi_1) - \langle x|\phi_1\rangle \text{ et } B = h(x) + h^*(-\phi_1) - \langle x| - \phi_1\rangle
$$

Par inégalité de Fenchel-Young on sait de plus que A et B sont positifs. Ceci montre que A et B sont nuls. Par la caractérisation du cas d'égalité dans l'inégalité de Fenchel-Young, on en déduit que $\phi_1 \in \partial f(x)$ et $-\phi_1 \in \partial h(x)$. Comme $h = g - \phi_0$, la seconde inclusion implique $\phi_0 - \phi_1 \in \partial g(x)$. Conclusion,

$$
\phi_0 = \phi_1 + (\phi_0 - \phi_1)
$$

où $\phi_1 \in \partial f(x)$ et $\phi_0 - \phi_1 \in \partial g(x)$ comme annoncé. $\Box$

Corollaire 71. *Soient $f_1, \dots, f_n : E \to \overline{\mathbb{R}}$ des fonctions convexes semi-continues inférieurement. On suppose que la condition de qualification suivante est vérifiée :*

$$
\exists x_0 \in \mathrm{dom}(f_1) \cap \ldots \cap \mathrm{dom}(f_n), \text{ tel que } f_2, \dots, f_n \text{ sont continues en } x_0 \quad \text{(Q)}
$$

alors,

$$
\forall x \in E, \partial(f_1 + \ldots + f_n)(x) = \partial f_1(x) + \ldots + \partial f_n(x).
$$

Exemple : problème à frontière libre On s'intéresse au problème suivant, qu'on peut voir comme la discrétisation d'un problème à frontière libre :

$$\min_{x\in\mathbb{R}^n, x\geq p} \frac{1}{2}\|Gx\|^2$$

où G est une matrice définie positive. On peut réécrire le problème sous la forme suivante, où l'on a posé $C_i = \{x \in \mathbb{R}^n \mid x_i \geq p_i\}$ et $f(x) = \frac{1}{2}\|Gx\|^2$:

$$\min_{x\in\mathbb{R}^n} f(x) + \sum_{1\leq i\leq n} i_{C_i}(x)$$

Soit x^* un minimum global. Alors,

$$0 \in \partial(f + \sum_{1\leq i\leq n} i_{C_i})(x^*).$$

Comme f est continue sur $\mathbb{R}^n$, les conditions du théorème sont vérifiées et on a donc, en utilisant $\partial f(x) = \{\nabla f(x)\}$,

$$\partial(f + \sum_{1\leq i\leq n} i_{C_i})(x^*) = \nabla f(x^*) + \sum_{1\leq i\leq n} \partial i_{C_i}(x^*)$$

Nous posons $L = G^tG$, de sorte que $\nabla f(x) = Lx$. Un calcul élémentaire montre que

$$\partial i_{C_i}(x) = \begin{cases} 0 & \text{si } x_i < p_i \\ \mathbb{R}^- e_i & \text{si } x = p_i \\ \emptyset & \text{sinon} \end{cases}$$

Ainsi, on obtient $0 = x^* + \sum_{i=1}^n \lambda_i e_i$, où $\lambda_i \leq 0$ et $\lambda_i = 0$ si $f_i = g_i$. Ainsi, on voit que x^* est caractérisé par le système

$$\begin{cases} Lx^* \geq 0 \\ \forall i \in \omega, (Lx^*)_i = 0 \text{ où } \omega = \{i \in \{1,\dots,n\} \mid g_i < f_i\} \end{cases}$$

Lemme 72. *Soit E un espace vectoriel normé, ϕ_0 une forme linéaire continue sur E, $a \in \mathbb{R}$ et $C = \{x \in E \mid \langle\phi_0|x\rangle \leq a\}$. Alors,*

$$\partial i_C(x) = \begin{cases} 0 & \textit{si } \langle\phi_0|x\rangle < a \\ \mathbb{R}^+\phi_0 & \textit{si } \langle\phi_0|x\rangle = a \\ \emptyset & \textit{sinon} \end{cases}$$

Démonstration. Si $\langle\phi_0|x\rangle < a$, x est dans l'intérieur de C, donc $f^+(x,\cdot) = 0$, de sorte que $\partial i_C(x) = 0$. On suppose désormais que $\langle\phi_0|x\rangle = a$, et on se donne $\phi \in \partial i_C(x)$, c'est-à-dire que

$$\forall x \in C, \quad \langle\phi|x - x_0\rangle \leq 0.$$

Soit $H = \{v \in E \mid \langle\phi_0|v\rangle = 0\}$. On vérifie facilement que $x_0 + H \subseteq C$ de sorte que,

$$\forall v \in H, \langle\phi|v\rangle \leq 0.$$

En remplaçant v par $-v$, on obtient $\phi|_H = 0$ ou encore $\phi \in H^\perp$. Par un argument standard d'algèbre linéaire on en déduit que $\phi = \lambda\phi_0$ où $\lambda \in \mathbb{R}$. En effet, soit $w \in E \setminus H$ et $\lambda = \langle\phi|w\rangle/\langle\phi_0|w\rangle$. Alors, la forme linéaire $\phi - \lambda\phi_0$ s'annule sur l'hyperplan H et en w et est donc nulle. Il reste à déterminer le signe de λ. Pour cela, on prend un vecteur $x \in \mathrm{int}(C)$, c'est-à-dire tel que $\langle\phi_0|x - x_0\rangle < 0$. Alors,

$$\langle\phi_0|x - x_0\rangle = \lambda\langle\phi|x - x_0\rangle \leq 0, \text{ soit } \lambda \geq 0 \qquad \square$$

Exemple : projection sur un polyèdre Soit H un espace de Hilbert, et $v_1, \dots, v_n \in H$ des vecteurs, et $a_1, \dots, a_n$ des scalaires. On suppose que l'intersection des convexes $C_i = \{x \in H \mid \langle v_i|x\rangle \leq a_i\}$ a un intérieur non vide. On s'intéresse au problème de projection d'un point x_0

$$\min_{x \in \cap_i C_i} \frac{1}{2}\|x - x_0\|^2 = \min_{x \in H} f(x) + \sum_{1 \leq i \leq n} i_{C_i}(x),$$

où $f(x) = \frac{1}{2}\|x - x_0\|^2$. On peut appliquer le théorème sur la somme des sous-différentiels pour obtenir que x^* est un minimiseur si et seulement si

$$0 \in \nabla f(x) + \sum_{1 \leq i \leq n} \partial i_{C_i}(x).$$

De plus, on vérifie que

$$\partial i_{C_i}(x) = \begin{cases} 0 & \text{si } \langle v_i|x\rangle < a_i \\ \mathbb{R}^+ v_i & \text{si } \langle v_i|x\rangle = a_i \\ +\infty & \text{sinon} \end{cases}$$

Ainsi, la condition d'optimalité peut s'écrire de la façon suivante :

$$0 = x^* - x_0 + \sum_{i=0}^{n} \lambda_i v_i,$$

où $\lambda_i \geq 0$ et $\lambda_i = 0$ si $\langle v_i|x\rangle < a_i$.

4.2 Sous-différentiel, cône normal et théorème KKT

Définition 24. Soit K un ensemble convexe fermé dans un espace vectoriel normé E et $x \in K$. On appelle *cône normal* de K en x l'ensemble

$$N_x K = \{\phi \in E^* \mid \forall y \in K, \phi(y) \leq \phi(x)\}.$$

Lemme 73. *Soit K convexe fermé et $x \in K$. Alors, $N_x K = -\partial i_K(x)$.*

Démonstration. Par définition de la fonction indicatrice convexe,

$$\begin{aligned}\phi \in N_x K &\iff \forall y \in K, \phi(y) \le \phi(x)\\ &\iff \forall y \in K, \phi(y) + i_K(y) \le \phi(x) + i_K(x)\\ &\iff \forall y \in E, \phi(y) + i_K(y) \le \phi(x) + i_K(x)\\ &\iff \phi \in \partial i_K(x)\end{aligned}$$

□

Exemple 20. Soit $K \subseteq E$ un convexe fermé et $f : E \to \overline{\mathbb{R}}$ une fonction convexe sci. On suppose une des deux hypothèses de qualification suivante :

$$\exists x_0 \in \mathrm{int}(K) \cap \mathrm{dom}(f) \tag{Q1}$$

$$\exists x_0 \in K \text{ t.q. } f \text{ est continue en } x_0 \tag{Q2}$$

Alors, en appliquant le théorème sur la somme des sous-différentiel à $f + i_K$, on a les équivalences suivantes :

$$\begin{aligned}x^* \text{ est minimum global de } f &\iff \exists \phi \in \partial f(x) \text{ t.q. } -\phi \in \partial i_K(x)\\ &\iff \exists \phi \in \partial f(x) \text{ t.q. } \phi \in \partial N_K(x)\\ &\iff \exists \phi \in \partial f(x) \text{ t.q. } \forall y \in K, \phi(y) \le f(x).\end{aligned}$$

Proposition 74. *Soit $g : E \to \mathbb{R}$ une fonction convexe continue, et soit K l'ensemble $K = \{x \in E \mid g(x) \le 0\}$. On suppose la condition de Slater vérifiée :*

$$\exists x_0 \in K \textit{ t.q. } f(x_0) < 0 \tag{S}$$

Alors pour tout $x \in K$, le cône normal à K en x est égal au cône engendré par le sous-différentiel de g en x :

$$N_x K = \begin{cases}\{0\} \textit{ si } g(x) < 0\\ \mathbb{R}^+ \partial g(x) \textit{ si } g(x) = 0,\end{cases}$$

ou on a noté $\mathbb{R}^+ C = \{\lambda x \mid \lambda \ge 0, x \in C\}$.

Remarque 19. En d'autre termes, $\phi \in N_x K$ est équivalent à l'existence de $\lambda \ge 0$ tel que $\lambda g(x) = 0$ et $\phi \in \lambda \partial g(x)$.

Démonstration. Si $g(x) < 0$, alors i_C est localement constante au voisinage de x donc dérivable et de gradient nul, i.e. $\partial i_C(x) = \{0\}$. On suppose désormais que $g(x) = 0$, et il s'agit de démontrer que $A = B$ où $A = N_x K = \partial i_K(x)$ et $B = \mathbb{R}^+ \partial g(x)$.
Étape 1 Commençons par l'inclusion facile $B \subseteq A$: soit $\phi \in \partial g(x)$ et $\lambda \ge 0$, on a

$$\forall y \in E, g(y) \ge \langle \phi | y - x \rangle + g(x) = \langle \phi | y - x \rangle$$

de sorte que,

$$\forall y \in K, 0 \ge \lambda g(y) \ge \lambda \langle \phi | y - x \rangle, \text{ i.e. } \lambda \phi \in N_x K$$

Étape 2 On veut maintenant montrer l'inclusion réciproque. Soit $\phi \in N_x K$ et soit H le demi-espace $H := \{y \in E \mid \langle \phi | y \rangle \geq \langle \phi | x \rangle\}$. Par définition du cône normal, on sait que $C \subseteq E \setminus \mathrm{int}(H)$. En utilisant de plus $C = \{g \leq 0\}$ on obtient

$$y \in \mathrm{int}(H) \Longrightarrow y \notin C \Longrightarrow g(y) \geq g(x) = 0$$

Autrement dit, le point x est le minimum de la fonction g sur H, c'est-à-dire

$$0 \in \partial(g + i_H)(x) = \partial g(x) + \mathbb{R}^+\{-\phi\}.$$

Ainsi, il existe $\psi \in \partial g(x)$ et $\lambda \geq 0$ tels que $\lambda\phi = \psi$. Pour conclure, il suffit d'exclure le cas $\lambda = 0$: si c'était le cas on aurait $0 \in \partial g(x)$, i.e. $\min_E g = 0$ ce qui contredirait la condition de Slater $g(x_0) < 0$. □

Théorème 75 (Karush-Kuhn-Tucker)**.** *Soit $f : E \to \overline{\mathbb{R}}$ une fonction convexe sci et $g_1, \ldots, g_N : E \to \overline{\mathbb{R}}$ des fonctions convexes continues. On suppose :*

$$\exists x_0 \in \mathrm{dom}(f) \text{ t.q. } \forall i \in \{1, \ldots, N\}, g_i(x_0) < 0.$$

Alors il y a équivalence entre

(i) x est un minimum global de f sur l'ensemble $C = \{x \in E \mid g_i(x) \leq 0\}$.

(ii) il existe $\lambda_1, \ldots, \lambda_N \in \mathbb{R}_+$ tels que

$$\begin{cases} 0 \in \partial f(x) + \lambda_1 \partial g_1(x) + \ldots + \lambda_N \partial g_N(x) \\ \lambda_i g_i(x) = 0 \quad \forall i \in \{1, \ldots, N\} \end{cases}$$

Démonstration. On pose $C_i = \{x \in E \mid g_i \leq 0\}$. La condition de qualification du théorème sur la somme de sous-différentiels est vérifiée, de sorte que

$$\forall x \in E, \partial(f + i_{C_1} + \ldots + i_{C_N})(x).$$

Ainsi, x est minimiseur global de f sur $C = \bigcap_i C_i$ si et seulement si

$$0 \in \partial(f + i_{C_1} + \ldots + i_{C_N})(x) = \partial f(x) + \partial i_{C_1}(x) + \ldots + \partial i_{C_N}(x),$$

i.e. il existe $\phi \in \partial f(x)$ et $\phi_i \in \partial i_{C_i}(x)$ tels que $\phi + \phi_1 + \ldots + \phi_N = 0$. On a montré précédemment que $\phi_i \partial i_{C_i}(x) \Longleftrightarrow \exists \lambda_i \geq 0$ tel que $\lambda_i g_i(x) = 0$ et $\phi_i \in \lambda_i \partial g_i(x)$. □

Exemple 21. Soient f et $g_1, \ldots, g_N$ comme dans le théorème et on suppose f coercive pour avoir l'existence d'un minimum global. Pour $\lambda_1, \ldots, \lambda_N \in \mathbb{R}^+$, on considère le problème pénalisé

$$\min_{x \in E} f(x) + \lambda_1 g_1(x) + \ldots + \lambda_N g_N(x).$$

Alors, avec $C = \{x \in E \mid \forall i, g_i(x) \leq 0\}$ on a

$$\min_{x \in C} f(x) \geq \min_{x \in E} f(x) + \lambda_1 g_1(x) + \ldots + \lambda_N g_N(x),$$

c'est-à-dire que le minimum du problème pénalisé est toujours plus petit que le minimum du problème avec contraintes. En d'autre termes,

$$\sup_{\lambda_1, \ldots, \lambda_N \geq 0} \min_{x \in E} f(x) + \lambda_1 g_1(x) + \ldots + \lambda_N g_N(x) \leq \min_{x \in C} f(x)$$

Une question naturelle est de déterminer s'il existe des $\lambda_1, \ldots, \lambda_N \geq 0$ tels qu'on ait égalité.

Le théorème de Karush-Kuhn-Tucker répond à cette question. Il existe $\lambda_i \in \mathbb{R}^+$ tel que

$$0 \in \partial f(x) + \lambda_1 \partial g_1(x) + \ldots + \lambda_N \partial g_N(x),$$

de sorte que

$$0 \in \partial(f + \lambda_1 g_1 + \ldots + \lambda_N g_N)(x).$$

Ainsi, x est aussi le minimum global du problème pénalisé, on parle alors de *pénalisation exacte.*

Exemple 22. On se donne un ensemble fini $y_1, \ldots, y_N \in \mathbb{R}^d$ et on s'intéresse au problème de déterminer le rayon de la plus petite sphère contenant ces points, i.e.

$$\min\{r \geq 0 \mid \exists x \in \mathbb{R}^d, \forall i \in \{1, \ldots, N\}, \|y_i - x\| \leq r\}.$$

On introduit $g_i : \mathbb{R}^{N+1} \to \mathbb{R}, (x,r) \mapsto \|y_i - x\| - r$, de sorte que $g_i(x,r)$ si et seulement si $\|y_i - x\| \leq r$. On peut alors réécrire le problème comme

$$\min_{(x,r) \in C} f(x,r),$$

où $f(x,r) = r$, $C = \bigcap_i C_i$ et $C_i = \{g_i \leq 0\}$. Les hypothèses du théorème Karush-Kuhn-Tucker sont vérifiée : tout le point (x,R) est dans l'intérieur de C si $R > \max_i \|x_i - x\|$. Un point $(x,r) \in \mathbb{R}^d \times \mathbb{R}$ résoud ce problème si et seulement si il existe $\lambda_1, \ldots, \lambda_N \geq 0$ vérifiant $\lambda_i g_i(x) = 0$ et tels que

$$0 \in \partial f(x) + \lambda_1 \partial g_1(x) + \ldots + \lambda_N \partial g_N(x).$$

On remarque d'abord que si $x = y_i$, alors $g_i(x) < 0$ (sinon, $r = 0$ et $y_1 = \ldots = y_N$), de sorte que $\lambda_i = 0$. Si $x \neq y_i$, on a $\partial g_i(x) = \{\nabla g_i(x)\}$ où

$$\nabla g_i(x) = ((x - y_i)/\left\|x - y_i\right\|, -1).$$

Ainsi,

$$0 = (0,1) + \sum_{i \mid \|x - y_i\| = r} \lambda_i((x - y_i)/\left\|x - y_i\right\|, -1),$$

En regardant la dernière coordonnée, on a $1 = \sum_i \lambda_i$. De plus, en observant les d premières coordonnées on obtient

$$0 = \sum_{i \mid \|x - y_i\| = r} \lambda_i \frac{x - y_i}{\|x - y_i\|} = \frac{1}{r} \sum_{i \mid \|x - y_i\| = r} \lambda_i(x - y_i)$$

Autrement dit, (x,r) est un minimiseur global si et seulement x est combinaison convexe des points y_i tels que $\|y_i - x\| = r$.

4.3 Théorème de Fenchel-Rockafellar et applications

Théorème 76 (Fenchel-Rockafellar). *Soit E un espace vectoriel normé, $f, g : E \to \overline{\mathbb{R}}$ deux fonctions convexes sci vérifiant l'hypothèse de qualification* (Q). *On suppose de plus* $\inf_E f + g > +\infty$. *Alors,*

$$\inf_{x \in E} f(x) + g(x) = \max_{\phi \in E^*} -f^*(-\phi) - g^*(\phi) \tag{4.6}$$

Démonstration. On a par définition de la conjuguée,

$$\inf_{x \in E} f(x) + g(x) = -(f+g)^*(0),$$

et la conclusion suit par le théorème sur la transformée de Legendre-Fenchel d'une somme (Théorème 69 et plus précisément (4.3)) appliqué en $\phi = 0$:

$$(f+g)^*(\phi) = \inf_{\psi \in E^*} f(\phi - \psi) + g(\psi).$$

□

Définition 25 (Plan de transport). Soient X et Y deux ensembles finis et $\mu : X \to \mathbb{R}$ et $\nu : Y \to \mathbb{R}$ deux mesures de probabilité sur X et Y respectivement (c'est-à-dire que $\mu, \nu \geq 0$ et $\sum_x \mu(x) = \sum_y \nu(y) = 1$). On appelle *plan de transport* entre μ et ν une matrice $\gamma : X \times Y \to \mathbb{R}$ vérifiant

$$\forall (x,y) \in X \times Y, \gamma(x,y) \geq 0 \tag{4.7}$$

$$\begin{cases} \forall x \in X,\ \sum_{y \in Y} \gamma(x,y) = \mu(x) \\ \forall y \in Y,\ \sum_{x \in X} \gamma(x,y) = \nu(y) \end{cases} \tag{4.8}$$

L'ensemble des plans de transport entre X et Y est noté $\Gamma(X,Y)$.

On se donne de plus une fonction $c : X \times Y \to \mathbb{R}$ mesurant le coût de déplacement d'une unité de masse d'un point $x \in X$ vers un point $y \in Y$. Le problème du transport optimal entre μ et ν pour le coût c consiste à minimiser

$$P = \inf_{\gamma \in \Gamma(X,Y)} \sum_{(x,y) \in X \times Y} \gamma(x,y) c(x,y) \tag{K}$$

Ce problème est appelle *problème de Kantorovich primal*. Un minimiseur de ce problème d'optimisation est appelé plan de transport optimal entre X et Y.

Théorème 77 (Kantorovich). *Le problème de Kantorovich primal est équivalent au problème de Kantorovich dual, i.e. $P = D$ où*

$$D = \sup \left\{ \sum_{x \in X} \phi(x)\mu(x) + \sum_{y \in Y} \phi(y)\nu(y) \mid \phi \in \mathbb{R}^X, \psi \in \mathbb{R}^Y, \quad \phi(x) + \psi(y) \leq c(x,y) \right\}. \tag{4.9}$$

Remarque 20. Interprétation économique du problème dual.

Remarque 21. Commençons par montrer comment retrouver ce résultat de manière informelle par la méthode des multiplicateurs de Lagrange. On écrit $i_{\geq 0}$ la fonction indicatrice des γ vérifiant la contrainte (4.7) et i_μ et i_ν celles associée aux deux contraintes (4.8)

$$i_{\geq 0}(\gamma) = \sup_{\sigma:X\times Y\to\mathbb{R},\sigma\geq 0} - \sum_{(x,y)\in X\times Y} \gamma(x,y)\sigma(x,y)$$

$$i_\mu(\gamma) = \sup_{\phi:X\to\mathbb{R}} \sum_{x\in X} \phi(x) \left(\mu(x) - \sum_{y\in Y} \gamma(x,y)\right)$$

$$i_\nu(\gamma) = \sup_{\psi:Y\to\mathbb{R}} \sum_{y\in Y} \psi(y) \left(\nu(y) - \sum_{x\in X} \gamma(x,y)\right)$$

Alors,

$$\begin{aligned} P &= \inf_{\gamma\in\Gamma(X,Y)} \sum_{(x,y)\in X\times Y} \gamma(x,y)c(x,y) \\ &= \inf_{\gamma} \sum_{x,y} \gamma(x,y)c(x,y) + i_{\geq 0}(\gamma) + i_\mu(\gamma) + i_\nu(\gamma) \\ &= \inf_{\gamma} \sup_{\sigma\geq 0,\phi,\psi} \sum_{x,y} \gamma(x,y)(c(x,y) - \sigma(x,y) - \mu(x) - \nu(y)) + \sum_x \phi(x)\mu(x) + \sum_y \psi(y)\gamma(y) \end{aligned}$$

Pour trouver le problème dual, on inverse l'infimum et le supremum (cette opération sera justifiée par Fenchel-Rockafellar),

$$D := \sup_{\sigma\geq 0,\phi,\psi} \inf_{\gamma} \sum_{x,y} \gamma(x,y)(c(x,y) - \sigma(x,y) - \mu(x) - \nu(y)) + \sum_x \phi(x)\mu(x) + \sum_y \psi(y)\nu(y).$$

Ensuite, on écrit les conditions nécessaires d'optimalité du problème de minimisation interne $\inf_\gamma$, les valeurs de σ, ϕ, ψ étant fixées :

$$c(x,y) = \sigma(x,y) + \phi(x) + \psi(y)$$

En utilisant ces conditions dans la formule précédente, on voit que les termes dans la double somme $\sum_{x,y}$ s'annulent, de sorte que

$$D = \sup_{\sigma\geq 0,\phi,\psi} \sum_x \phi(x)\mu(x) + \sum_y \psi(y)\gamma(y),$$

sous la contrainte $c(x,y) = \sigma(x,y) + \phi(x) + \psi(y)$. On peut supprimer la variable σ et se retrouver avec l'équation (4.9).

Démonstration du théorème 77. Le problème de Kantorovich revient à minimiser la somme $f + g$ sur l'espace $E = \mathbb{R}^{X\times Y}$, où

$$f(\gamma) = \begin{cases} \sum_{(x,y)\in X\times Y} \gamma(x,y)c(x,y) & \text{si } \gamma \geq 0 \\ +\infty & \text{sinon} \end{cases}, \tag{4.10}$$

$$g(\gamma) = i_\mu(\gamma) + i_\nu(\gamma) = \begin{cases} 0 & \text{si (4.8) est verifiee} \\ +\infty & \text{sinon} \end{cases} \tag{4.11}$$

Dans les calculs suivant, on munit l'espace E de la structure Euclidienne canonique, et on l'identifie avec son dual. On a :

$$\begin{aligned} f^*(\pi) &= \sup_{\gamma\in E} \langle \pi|\gamma\rangle - f(\gamma) \\ &= \sup_{\gamma\in E, \gamma\geq 0} \sum_{(x,y)\in X\times Y} (\pi(x,y) - c(x,y))\gamma(x,y) \\ &= \begin{cases} 0 & \text{si } \forall (x,y)\in X\times Y \text{ t.q. } \pi(x,y)\leq c(x,y) \\ +\infty & \text{sinon} \end{cases} \end{aligned}$$

Passons maintenant au calcul de la conjuguée de g :

$$\begin{aligned} g^*(\pi) &= \sup_{\gamma\in E} \langle \pi|\gamma\rangle - g(\gamma) \\ &= \sup \left\{ \sum_{(x,y)\in X\times Y} \pi(x,y)\gamma(x,y) \mid \forall x\in X, \sum_{y\in Y}\gamma(x,y) = \mu_x \text{ et } \forall y\in Y, \sum_{x\in X}\gamma(x,y) = \nu_y \right\} \end{aligned}$$

Par la remarque 21, on s'attend à ce que $\pi(x,y) = \phi(x) + \phi(y)$ où $\phi\in\mathbb{R}^X$ et $\psi\in\mathbb{R}^Y$. Supposons dans un premier temps que c'est bien le cas. Alors, pour tout γ vérifiant les conditions de marge,

$$\begin{aligned} \sum_{(x,y)\in X\times Y} \pi(x,y)\gamma(x,y) &= \sum_{x\in X}\phi(x)\sum_{y\in Y}\gamma(x,y) + \sum_{y\in Y}\psi(y)\sum_{x\in X}\gamma(x,y) \\ &= \sum_{x\in X}\phi(x)\mu(x) + \sum_{y\in Y}\phi(y)\nu(y) \end{aligned}$$

Supposons maintenant que $\pi(x,y)$ n'est pas de la forme $\phi(x)+\psi(y)$. Ceci signifie qu'il existe $x_0\neq x_1\in X$ tels que $\pi(x_0,\cdot) - \pi(x_1,\cdot)$ n'est pas constant sur Y, i.e.

$$\exists y_0\neq y_1\in y \text{ t.q. } \pi(x_0,y_0) - \pi(x_1,y_0) \neq \pi(x_0,y_1) - \pi(x_1,y_1)$$

On supposera (par exemple) que $\pi(x_0,y_0)+\pi(x_1,y_1) > \pi(x_1,y_0)+\pi(x_0,y_1)$. Soit $\delta\in\mathbb{R}^{X\times Y}$ défini par $\delta(x_0,y_0) = \delta(x_1,y_1) = 1$ et $\delta(x_0,y_1) = \delta(x_1,y_0) = -1$. Alors, si γ vérifie les conditions de marge, $\gamma + t\delta$ les vérifie aussi. De plus,

$$\sum_{(x,y)\in X\times Y}\pi(x,y)(\gamma+t\delta)(x,y) = \sum_{(x,y)\in X\times Y}\pi(x,y)\gamma(x,y) + t(\pi(x_0,y_0)+\pi(x_1,y_1)-\pi(x_1,y_0)-\pi(x_0,y_1)),$$

de sorte qu'en faisant tendre t vers $+\infty$, on voit que le supremum dans la définition de g^* vaut $+\infty$. En conclusion,

$$g^*(\pi) = \begin{cases} \sum_{x\in X}\phi(x)\nu(x) + \sum_{y\in Y}\phi(y)\nu(y) & \text{si } \exists\phi\in\mathbb{R}^X, \psi\in\mathbb{R}^Y,\ \ \pi(x,y) = \phi(x)+\psi(y) \\ +\infty & \text{sinon} \end{cases}$$

Pour conclure, il suffit donc de vérifier qu'on est bien dans les conditions d'application du théorème de Fenchel-Rockafellar, et en particulier que la la condition de qualification (Q) est vérifiée. Soit $\pi_0(x,y) = \mu(x)\nu(y)$, $\pi_0 \in E$. Ce π_0 satisfait les conditions de marge (4.8), de sorte que $\pi \in \mathrm{dom}(g)$. De plus, $\pi \geq \min \mu \cdot \min \nu > 0$ de sorte que π est à l'intérieur du domaine de f, et donc en dimension finie, f est continue en π. Ainsi,

$$P = \sup_{\pi} -f^*(\pi) - g^*(-\pi) = \sup_{\phi,\psi|\phi+\psi\leq c} \sum_{x\in X} \mu(x)\phi(x) + \sum_{y\in Y} \nu(y)\psi(y). \qquad \square$$

Théorème 78 (Fenchel-Rockafellar)**.** *Soit E un espace vectoriel normé, $\Lambda : E \to F$ une application linéaire continue et $f : E \to \overline{\mathbb{R}}$ et $g : F \to \overline{\mathbb{R}}$ deux fonctions convexes sci vérifiant l'hypothèse de qualification*

$$\exists x \in E, \ \text{t.q. } x \in \mathrm{dom}(f) \text{ et } g \text{ continue en } \Lambda x \tag{Q'}$$

On suppose de plus que $\inf_E f + g \circ \Lambda > -\infty$. Pour $\psi \in F^*$, on pose $\Lambda^*\psi = \psi \circ \Lambda \in E^*$. Alors,

$$\inf_{x\in E} f(x) + g(\Lambda x) = \max_{\psi\in F^*} -f^*(-\Lambda^*\psi) - g^*(\psi) \tag{4.12}$$

Démonstration. On va chercher à appliquer la précédente version du théorème de Fenchel-Rockafellar sur l'espace produit $E \times F$. On pose

$$\begin{aligned} &f_1 : E \times F \to \mathbb{R}, \quad (x,y) \mapsto \begin{cases} f(x) \text{ si } y = \Lambda x \\ 0 \text{ sinon} \end{cases} \\ &g_1 : E \times F \to \mathbb{R}, \quad (x,y) \mapsto g(y) \end{aligned}$$

On a alors facilement que $\inf_E f + g \circ \Lambda = \inf_{E\times F} f_1 + g_1$. Il s'agit maintenant d'appliquer la précédente version du théorème de Fenchel-Rockafellar à ce deuxième problème. On identifie $(E \times F)^*$ à $E^* \times F^*$. Pour $\phi \in E^*$ et $\psi \in F^*$, on a

$$\begin{aligned} f_1^*(\phi,\psi) &= \sup_{x\in E, y\in F} \langle \phi|x\rangle + \langle \psi|y\rangle - f_1(x) = \sup_{x,y|\Lambda x=y} \langle \phi|x\rangle + \langle \psi|y\rangle - f(x) \\ &= \sup_{x,y|\Lambda x=y} \langle \phi|x\rangle + \langle \psi|\Lambda x\rangle - f(x) = \sup_{x,y|\Lambda x=y} \langle \phi + \Lambda^*\psi|x\rangle - f(x) \\ &= f^*(\phi + \Lambda^*\psi) \end{aligned}$$

$$\begin{aligned} g_1^*(\phi,\psi) &= \sup_{x\in E, y\in F} \langle \phi|x\rangle + \langle \psi|y\rangle - g_1(x) = \sup_{x\in E, y\in F} \langle \phi|x\rangle + \langle \psi|y\rangle - g(x) \\ &= \begin{cases} +\infty & \text{si } \psi \neq 0 \\ g^*(\psi) & \text{sinon} \end{cases} \end{aligned}$$

Ainsi, par thèorème de Fenchel-Rockafellar et en utilisant les calculs précédents,

$$\begin{aligned} \inf_E f + g \circ \Lambda = \inf_{E\times F} f_1 + g_1 &= -\sup_{(\phi,\psi)\in E^*\times F^*} f_1^*(-\phi,-\psi) + g_1^*(\phi,\psi) \\ &= -\sup_{\psi\in F^*} f(-\Lambda^*\psi) + g(\psi) \end{aligned} \qquad \square$$

Exemple 23. On considère le problème de minimisation sur $\mathbb{R}^n$, où A est une matrice à m lignes et n colonnes, $x_0 \in \mathbb{R}^n$:

$$\min_{x\in\mathbb{R}^n} \frac{1}{2}\|x-x_0\|_2^2 + \|Ax\|_1 = \min_{x\in\mathbb{R}^n} f(x) + g(Ax).$$

Comme f et g sont continues, on peut appliquer le théorème de Fenchel-Rockafellar. Calculons maintenant les conjuguées de f et g :

$$\begin{aligned} f^*(y) &= \sup_{x\in\mathbb{R}^n} \langle x|y\rangle - \frac{1}{2}\|x-x_0\|^2 \\ &= \langle x_0|y\rangle + \sup_{x\in\mathbb{R}^n} \langle x-x_0|y\rangle - \frac{1}{2}\|x-x_0\|_2^2 \\ &= \langle x_0|y\rangle + \frac{1}{2}\|y\|_2^2 = \frac{1}{2}\|y+x_0\|_2^2 - \frac{1}{2}\|x_0\|_2^2, \end{aligned}$$

où l'on a utilisé $(\frac{1}{2}\|.\|_2^2)^* = \frac{1}{2}\|.\|_2^2$. De plus,

$$g^*(y) = \sup_{x\in\mathbb{R}^m} \sum_i x_i y_i - \sum_i |x_i| = \sum_{1\le i\le m} h^*(y_i),$$

où l'on a posé $h : x \in \mathbb{R} \to |x|$. De plus,

$$h^*(r) = \sup_{s\in\mathbb{R}} rs - |s| = \begin{cases} 0 & \text{si } r \in [-1,1] \\ +\infty & \text{sinon} \end{cases}$$

Ainsi, g^* est la fonction indicatrice de $[-1,1]^m$ (qui est la boule unité pour $\|.\|_\infty$). Ainsi,

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \frac{1}{2}\|x-x_0\|_2^2 + \|Ax\|_1 &= \min_{x\in\mathbb{R}^n} f(x) + g(Ax) \\ &= \max_{y\in\mathbb{R}^n} -f^*(-A^t y) - g^*(y) \\ &= \min_{y\in[-1,1]^n} \frac{1}{2}\left\|A^t y - x_0\right\|_2^2 - \frac{1}{2}\|x_0\|^2 \end{aligned}$$

4.4 Algorithme du point proximal

Dans cette dernière section, on suppose que E est un espace de Hilbert, que l'on identifiera à son dual. On rappelle qu'une fonction $f : E \to \overline{\mathbb{R}}$ est dite 0-coercive si $\lim_{\|x\|\to\infty} f(x) = +\infty$.

Définition 26. Soit $f : E \to \overline{\mathbb{R}}$ une fonction convexe propre et semi-continue inférieurement. Pour $\gamma > 0$, on définit l'opérateur proximal de f par

$$\mathrm{prox}_\gamma f(x) = \arg\min_{y\in E} \frac{1}{2\gamma}\|x-y\|^2 + f(y).$$

Exemple 24. Si i_C où C est un ensemble convexe fermé de H, $\operatorname{prox}_\gamma i_C = p_C$. L'opérateur proximal généralise la fonction projection sur un ensemble convexe, et possède un certain nombre de ses propriétés.

Proposition 79. *Soit $f : E \to \overline{\mathbb{R}}$ est convexe propre, semi-continue inférieurement et $\gamma > 0$. Alors,*

(i) Le problème de minimisation suivant admet un unique minimiseur.

$$\min_{y\in E} \frac{1}{2\gamma} \|x-y\|^2 + f(y),$$

et l''opérateur proximal de f est donc bien défini.

(ii) Le point $p = \operatorname{prox}_\gamma f(x)$ est caractérisé par la relation $x \in (\mathrm{id} + \gamma\partial f)(p)$.

(iii) Le point x est un minimiseur global de la fonction f sur E si et seulement s'il est point fixe de l'opérateur proximal, i.e. $x = \operatorname{prox}_\gamma f(x)$.

Démonstration. Soit $g = \frac{1}{2\gamma} \|x - \cdot\|^2$ et $h = f + g$ (i) La fonction h est convexe propre, semi-continue inférieurement pour la topologie forte, et elle est 0-coercive (i.e. $\lim_{\|y\|\to\infty} h(y) = +\infty$). Soit $r > \inf h$: le sous niveau $K = \{h \leq r\}$ est convexe, fortement fermé et donc faiblement fermé. Comme de plus K est borné, K est donc faiblement compact. La fonction f est convexe et fortement semi-continue inférieurement et donc faiblement semi-continue inférieurement. Elle atteint donc son minimum (global) sur K. Par stricte convexité de $\|x - \cdot\|^2$, h est strictement convexe et le minimum est donc unique.

(ii) En utilisant le théorème sur la somme des sous-différentiels, on voit que $p = \operatorname{prox}_\gamma f(x)$ si et seulement si

$$0 \in \partial(f+g)(p) = \partial f(p) + \frac{1}{\gamma}\{p-x\} \Longleftrightarrow x \in (\mathrm{id} + \gamma\partial f)(p).$$

(iii) Le point x est un minimiseur global de f si et seulement si

$$0 \in \partial f(x) \Longleftrightarrow x \in (\mathrm{id} + \gamma\partial f)(x) \Longleftrightarrow x = \operatorname{prox}_\gamma f(x).$$ □

Exemple 25. Soit $h : \mathbb{R} \to \mathbb{R}, x \mapsto |x|$. On a

$$\partial h(y) = \begin{cases} -1 & \text{si } y < 0 \\ [-1,1] & \text{si } y = 0 \\ 1 & \text{si } y > 0 \end{cases} \quad \Longrightarrow \quad (\mathrm{id} + \gamma\partial h)(y) = \begin{cases} y-\gamma & \text{si } y < 0 \\ [-\gamma,\gamma] & \text{si } y = 0 \\ y+\gamma & \text{si } y > 0 \end{cases}.$$

Par la caractérisation donnée dans la proposition précédente, on obtient

$$\operatorname{prox}_\gamma h(x) = \begin{cases} x-\gamma & \text{si } x \geq \gamma \\ 0 & \text{si } -\gamma \leq x \leq \gamma \\ x+\gamma & \text{si } x \leq -\gamma. \end{cases}$$

On note $R_\gamma = \operatorname{prox}_\gamma h$. Cette fonction est appelée *opérateur de seuillage doux* (ou *soft thresholding*) en statistique et en traitement d'image. Soit maintenant $E = \mathbb{R}^n$ muni de la norme euclidienne et $f(x) = \|x\|_1 = \sum_{1\leq i\leq n} |x_i|$. Alors,

$$\min_{y\in\mathbb{R}^n} \frac{1}{2\gamma}\|x-y\|_2^2 + f(y) = \min_{y\in\mathbb{R}^n} \sum_{1\leq i\leq n} \frac{1}{2\gamma}(x_i-y_i)^2 + |y_i|,$$

de sorte que $\operatorname{prox}_\gamma f(x) = (R_\gamma(x_1), \dots, R_\gamma(x_n))$.

La caractérisation donnée en (iii) du minimum global comme un minimiseur de f invite à utiliser l'algorithme de point fixe pour résoudre numériquement le problème d'optimisation :

$$\begin{cases} x_0 \in E \\ x_{n+1} = \operatorname{prox}_\gamma f(x_n) \end{cases} \tag{PPA}$$

Cet algorithme est appelé algorihtme du point proximal et a été introduit Martinet (puis généralisé par Rockafellar) dans les années 1960. Avant de pouvoir montrer la convergence de cet algorithme, nous avons besoin de plus d'informations sur l'opérateur proximal.

Définition 27. Un opérateur $F : E \to E$ est dit *fermement non-expansif* s'il vérifie une des conditions équivalentes suivantes
(i) $\forall x, y \in E,\ \|F(x) - F(y)\|^2 \leq \langle F(x) - F(y)|x-y\rangle$
(ii) $\forall x, y \in E,\ \|F(x) - F(y)\|^2 \leq \|x-y\|^2 - \|(x-F(x)) - (y-F(y))\|^2$

Pour voir l'équivalence entre ces deux conditions, il suffit de remarquer que

$$\|(x-F(x)) - (y-F(y))\|^2 = \|x-y\|^2 + \|F(x)-F(y)\|^2 - 2\langle x-y|F(x)-F(y)\rangle$$

Proposition 80. *Soit $f : E \to \overline{\mathbb{R}}$ une fonction convexe, propre et semi-continue inférieurement et $\gamma > 0$. Alors*
(i) Le point $p = \operatorname{prox}_\gamma f(x)$ est caractérisé par l'inégalité

$$\forall q \in E,\ \frac{1}{\gamma}\langle x-p|q-p\rangle \leq f(q) - f(p)$$

(ii) L'opérateur $F : x \mapsto \operatorname{prox}_\gamma f(x)$ est fermement non-expansif.

Démonstration. (i) Le point $p = \operatorname{prox}_\lambda f(x)$ est caractérisé par $x \in (\mathrm{id} + \gamma\partial f)(p)$ ou de manière équivalente par $\frac{1}{\gamma}(x-p) \in \partial f(p)$. C'est-à-dire,

$$\forall q \in E,\ f(q) \geq f(p) + \frac{1}{\gamma}\langle x-p|q-p\rangle.$$

(ii) Soient x_1, x_2 et $p_i = \operatorname{prox}_\gamma f(x_i)$. Appliquons l'inégalité de (i) en prenant d'abord $x = x_1$, $p = p_1$ et $q = p_2$ puis en inversant les rôles :

$$f(p_2) - f(p_1) \geq \frac{1}{\gamma}\langle x_1 - p_1|p_2 - p_1\rangle$$

$$f(p_1) - f(p_2) \geq \frac{1}{\gamma}\langle x_2 - p_2|p_1 - p_2\rangle$$

En additionnant ces inégalités puis en multipliant par γ, on obtient

$$\|p_2 - p_1\|^2 \leq \langle p_1 - p_2 | x_1 - x_2 \rangle. \qquad \square$$

Remarquons que le caractère seulement 1-Lipschitz de $F = \text{prox}_\gamma f$ ne permet pas d'appliquer le théorème du point fixe contractant (il faudrait que F soit k-Lipschitz avec $k < 1$). Cependant, la propriété de de non-expansivité ferme permet d'obtenir un théorème de convergence.

Théorème 81. *Soit $f : E \to \overline{\mathbb{R}}$ une fonction convexe semi-continue inférieurement et* 0*-coercive. Alors, la suite de points défini par l''algorithme du point proximal* (PPA) *converge faiblement vers un minimum global de f.*

La démonstration de ce théorème provient des trois lemmes suivants, où l'on a posé $F = \text{prox}_\gamma f$. Les deux premiers lemmes sont valables pour toute opérateur fermement non expansif F.

Lemme 82. *Soit $F : E \to E$ un opérateur fermement non-expansif admettant un point fixe. Alors, la suite définie par $x_{n+1} = F(x_n)$ est bornée et*

$$\lim_{n\to\infty} \|x_n - x_{n+1}\| = 0.$$

Démonstration. Soit c un point fixe de F. La non-expansivité ferme donne :

$$\|Fx_n - Fc\|^2 \leq \|x_n - c\|^2 - \|(x_n - F(x_n)) - (c - F(c))\|^2.$$

En utilisant $Fx_n = x_{n+1}$ et $F(c) = c$, cette inégalité implique

$$\|x_{n+1} - c\|^2 \leq \|x_n - c\|^2 - \|x_n - x_{n+1}\|^2$$

La suite $\|x_n - c\|_{n\geq 1}$ est décroissante et minorée donc convergente, de sorte que

$$\|x_n - x_{n+1}\|^2 \leq \|x_n - c\|^2 - \|x_{n+1} - c\|^2 \xrightarrow[n\to+\infty]{} 0 \qquad \square$$

Lemme 83. *Soit $F : E \to E$ un opérateur non-expansif admettant un point fixe. On suppose que tout point d'accumulation faible de la suite $(x_n)_{n\geq 1}$ est un point fixe de F. Alors $(x_n)_{n\geq 1}$ converge faiblement vers un point fixe de F.*

Démonstration. Soient c_1 et c_2 deux points d'accumulation faibles de $(x_n)_{n\geq 0}$, qui par hypothèses sont des points fixes de F. On a remarqué au cours de la démonstration du lemme précédent que $\|x_{n+1} - c_i\|^2 \leq \|x_n - c_i\|^2$, ce qui implique que la suite

$$\|x_n\|^2 - 2\langle x_n | c_i \rangle = \|x_n - c_i\|^2 - \|c_i\|^2,$$

est décroissante donc convergente. En soustrayant ces suites pour $i = 1$ et 2,

$$(\|x_n\|^2 - 2\langle x_n | c_1 \rangle) - (\|x_n\|^2 - 2\langle x_n | c_2 \rangle) = 2\langle x_n | c_1 - c_2 \rangle$$

est convergente. Comme c_1 et c_2 sont des points d'accumulation (faibles) de x_n, on a donc en passant à la limite des sous-suites correspondantes

$$\langle c_1 | c_1 - c_2 \rangle = \langle c_2 | c_1 - c_2 \rangle$$

ce qui implique que $\|c_1 - c_2\|^2 = 0$, soit $c_1 = c_2$. La suite $(x_n)_{n\geq 1}$ est bornée et a un unique point d'accumulation faible : elle est donc faiblement convergente. $\square$

Lemme 84. *Soit $F = \mathrm{prox}_\gamma f$ où f est convexe, propre, sci et 0-coercive. Alors tout point d'accumulation faible de la suite $(x_n)_{n\geq 0}$ est un point fixe de F.*

Démonstration. L'application F admet un point fixe car la fonction f admet un minimum global sur f. Par la caractérisation de $x_{n+1} = \mathrm{prox}_\gamma f(x_n)$ en terme de sous-différentiel, on a

$$x_n \in (\mathrm{id} + \frac{1}{\lambda}\partial f)(x_{n+1}) = x_{n+1} + \frac{1}{\lambda}\partial f(x_{n+1})$$

Autrement dit,

$$\forall x \in E,\ f(x) \geq f(x_{n+1}) + \frac{1}{\gamma}\langle x - x_{n+1} | x_n - x_{n-1}\rangle$$

Soit (x_{n_k}) une sous-suite faiblement convergente de la suite (x_n). On a

$$\begin{aligned}\forall x \in E, f(x) &\geq f(x_{n_k+1}) + \frac{1}{\gamma}\langle x - x_{n_k+1} | x_{n_k} - x_{n_k+1}\rangle \\ &\geq f(x_{n_k+1}) - \frac{1}{\gamma}\left\|x - x_{n_k+1}\right\| \left\|x_{n_k} - x_{n_k+1}\right\|\end{aligned}$$

Or, on sait par un des lemmes précédents que $\lim_{n\to+\infty}\|x_{n_k} - x_{n_k+1}\| = 0$, de sorte qu'en passant à la limite faible (et en utilisant la semi-continuité faible de f), $x_{n_k} \to \bar{x} \in E$, on a

$$\forall x \in E, f(x) \geq f(\bar{x}),$$

ce qui implique que $\bar{x}$ est un minimum global de f et donc un point fixe de l'opérateur F. □

Algorithme forward-backward On suppose que la fonction f minimisée peut être décomposée sous la forme $f = f_1 + f_2$ où $f_1 : E \to \overline{\mathbb{R}}$ et $f_2 : E \to \mathbb{R}$ sont deux fonctions convexes semi-continues inférieurement, propres et 0-coercives. On demande de plus que f_2 soit différentiable.

Proposition 85. *Les assertions suivantes sont équivalentes*

(i) $x \in E$ est un minimum global de $f = f_1 + f_2$;

(ii) x est un point fixe de l'opérateur $F : x \mapsto \mathrm{prox}_{\gamma f_1}(x - \gamma\nabla f_2(x))$.

Démonstration. $x = \arg\min_E f$ si et seulement si

$$\begin{aligned}&0 \in \partial(f_1 + f_2)(x) = \partial f_1(x) + \{\nabla f_2(x)\} \\ &\iff -\nabla f_2(x) \in \partial f_1(x) \\ &\iff x - \gamma\nabla f_2(x) \in (\mathrm{id} + \gamma\partial f_1)(x) \\ &\iff x = \mathrm{prox}_\gamma f_1(x - \gamma\nabla f_2(x))\end{aligned}$$

□

Théorème 86. *Soit $f = f_1 + f_2$, où $f_1 : E \to \overline{\mathbb{R}}$ et $f_2 : E \to \mathbb{R}$ sont deux fonctions convexes semi-continues inférieurement, propres et 0-coercives. On demande de plus que f_2 soit différentiable et $x \in E \mapsto \nabla f_2(x)$ est L-Lipschitz. Soit $\gamma \leq 1/L$, et*

$$\begin{cases} x_0 \in E \\ x_{n+1} = \operatorname{prox}_\gamma f_1(x_n - \gamma f_2(x_n)). \end{cases}$$

Alors, la suite (x_n) converge faiblement vers un minimiseur de f.

Exemple 26. Soit $E = \mathbb{R}^n$, $f_1(x) = \|x\|_1$ et $f_2(x) = \frac{1}{2}\|Ax - b\|_2^2$. Alors, $\nabla f_2(x) = A^T(Ax - b)$, et $\operatorname{prox}_\gamma f_1 = R_\gamma$ est l'opérateur de seuillage doux. Dans ce contexte, l'algorithme décrit par le théorème est donné par $x_0 \in E$ et

$$x_{n+1} = R_\gamma(x_n - \gamma A^T(Ax_n - b)),$$

et est connu sous le nom de *iterative shrinking-thresholding algorithm* (ISTA) en statistique, traitement d'image et peut-être plus généralement être appliqué pour résoudre des problèmes inverses linéaires.