

Fenchel 对偶性 和 凸共轭

MODULE A7

Dualité de FENCHEL et conjuguée convexe

Sauf mention contraire, \mathcal{X} (resp. \mathcal{Y}) est un espace de HILBERT, muni du produit scalaire noté $\langle \cdot, \cdot \rangle$ et on note $\| \cdot \|$ la norme qui découle du produit scalaire.

Dans ce module, on va introduire la **conjuguée convexe** de **LEGENDRE-FENCHEL**. Cet outil joue un rôle important en **dualité min-max** en **optimisation non différentiable**.
不可微优化

1 Conjuguée convexe ou de LEGENDRE-FENCHEL

1.1 Définition et exemples

Définition 1 (Conjuguée convexe) 凸共轭

Soit $J : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ une fonction. On définit sa **conjuguée convexe** comme étant la fonction $J^* : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ donnée par

$$\forall y \in \mathcal{X}, \quad J^*(y) = \sup_{x \in \mathcal{X}} \left\{ \langle y, x \rangle - J(x) \right\}$$

On parle parfois de **conjuguée de LEGENDRE-FENCHEL**. L'application

$$J \mapsto J^*$$

est quant à elle appelée **transformée de FENCHEL**.

Commençons par **deux exemples** très classiques, qui illustrent **deux techniques différentes** permettant de déterminer **explicitement** les **valeurs** de la **conjuguée convexe**.

范数平方的凸共轭

EXEMPLE

Conjuguée convexe de la norme au carré. Considérons la fonction convexe

$$J : \begin{cases} \mathcal{X} & \rightarrow \mathbb{R} \\ x & \mapsto \frac{1}{2} \|x\|^2 \end{cases}$$

Alors on a $J^* = J$. En effet, par définition,

$$\forall y \in \mathcal{X}, \quad J^*(y) = \sup_{x \in \mathcal{X}} \left\{ \langle y, x \rangle - \frac{1}{2} \|x\|^2 \right\} = - \inf_{x \in \mathcal{X}} \left\{ -\langle y, x \rangle + \frac{1}{2} \|x\|^2 \right\}$$

On obtient le **résultat désiré** en appliquant la **règle de FERMAT** pour ce problème de minimisation convexe lisse non contraint, qui assure que le **minimum** dans le membre de droite de l'égalité précédente est atteint au point $x^* = y$ et vaut

$$-\langle y, x^* \rangle + \frac{1}{2} \|x^*\|^2 = -\frac{1}{2} \|x^*\|^2 = -\frac{1}{2} \|y\|^2$$

范数的凸共轭

EXEMPLE

Conjuguée convexe de la norme. Considérons la fonction convexe

$$J : \begin{cases} \mathcal{X} & \rightarrow \mathbb{R} \\ x & \mapsto \|x\| \end{cases}$$

Soit $y \in \mathcal{X}$. L'inégalité de CAUCHY-SCHWARZ permet de fournir la majoration suivante :

$$\forall x \in \mathcal{X}, \quad \langle y, x \rangle - \|x\| \leq \|y\| \|x\| - \|x\| = (\|y\| - 1) \|x\|$$

Ainsi, si $\|y\| \leq 1$, l'inégalité précédente assure que $\langle y, x \rangle - \|x\|$ est négatif pour tout $x \in \mathcal{X}$. On peut vérifier qu'elle s'annule pour $x = y/\|y\|$. Si $\|y\| > 1$, il suffit de poser $x_k = k y$ pour tout $k \in \mathbb{N}$. Alors

$$\langle y, x_k \rangle - \|x_k\| = \langle y, k y \rangle - \|k y\| = k \|y\|^2 - k \|y\| = k \|y\| (\|y\| - 1)$$

On voit donc $\lim_{k \rightarrow +\infty} \langle y, x_k \rangle - \|x_k\| = +\infty$

de sorte que $J^*(y) = \begin{cases} 0 & \text{si } \|y\| \leq 1 \\ +\infty & \text{sinon} \end{cases}$

Autrement dit, J^* est l'indicatrice de la boule unité fermée \mathcal{B} .
闭球 \mathcal{B} 的指示函数

1.2 Convexité de la conjuguée convexe 凸共轭的凸性

On commence par établir le résultat suivant, qui assure en particulier que, même si J n'est pas convexe, sa conjuguée convexe l'est.

Proposition 1

Soit $J : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ une fonction. Alors sa conjuguée convexe J^* est convexe et s.c.i.

DÉMONSTRATION : Il suffit de remarquer que pour tout $y \in \mathcal{X}$, la fonction

$$y \mapsto \langle y, x \rangle - J(x)$$

est une fonction convexe (car affine) ; en tant qu'enveloppe supérieure de telles fonctions, la conjuguée convexe J^* est convexe et s.c.i. ■

Le caractère propre de la conjuguée convexe requiert une hypothèse supplémentaire sur la fonction J .

Proposition 2

Soit $J : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction. Alors on a l'équivalence entre les deux énoncés suivants :

- (i) il existe $x^0 \in \mathcal{X}$ tel que $J(x^0) < +\infty$ et J admet une minorante affine ;
- (ii) sa conjuguée convexe J^* est convexe, s.c.i. et propre.

DÉMONSTRATION : On rappelle que J admet une **minorante affine** s'il existe un couple $(a, b) \in \mathcal{X} \times \mathbb{R}$ tel que

$$\forall x \in \mathcal{X}, \quad J(x) \geq \langle a, x \rangle + b$$

- **Sens direct.** En vertu de la proposition 1, il suffit de démontrer que J^* est propre. Commençons par remarquer que, puisque J n'est pas identiquement égale à $+\infty$ par hypothèse, sa **conjuguée convexe** J^* est nécessairement à valeurs dans $\mathbb{R} \cup \{+\infty\}$. Ainsi, il nous reste à démontrer qu'il existe $y_0 \in \mathcal{X}$ tel que $J^*(y_0) \neq +\infty$. Par hypothèse, on a pour tout $y \in \mathcal{X}$

$$\forall x \in \mathcal{X}, \quad \langle y, x \rangle - J(x) \leq \langle y - a, x \rangle - b$$

En particulier, si on choisit $y_0 = a$, alors la **majoration** précédente devient

$$\forall x \in \mathcal{X}, \quad \langle y_0, x \rangle - J(x) \leq -b$$

ce qui assure, en passant à la **borne supérieure**, que $J^*(y_0)$ est fini.

- **Réciproque.** Si J^* est propre, alors il existe au moins un point $y_0 \in \mathcal{X}$ tel que $J^*(y_0) \in \mathbb{R}$. Par conséquent, on a

$$\forall x \in \mathcal{X}, \quad J^*(y_0) \geq \langle y_0, x \rangle - J(x) \quad \text{soit} \quad J(x) \geq \langle y_0, x \rangle - J^*(y_0)$$

ce qui assure que J admet une **minorante affine**, et ne peut donc prendre la valeur $-\infty$. Par ailleurs, si J était identiquement égale à $+\infty$, alors J^* prendrait la valeur $-\infty$, ce qui contredit le fait que J^* soit propre. ■

On en déduit alors le résultat suivant :

Corollaire 1

Soit $J : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction **convexe** et propre. On suppose qu'il existe $x^0 \in \text{dom } f$ tel que f soit continue en x^0 . Alors sa **conjuguée convexe** J^* est convexe, s.c.i. et propre.

DÉMONSTRATION : Il suffit de remarquer que, d'après la proposition 5 du module A2 : **Sous-différentiabilité** que f admet une **minorante affine**. ■

2 Propriétés de la conjuguée convexe 凸共軛的性质

On s'intéresse dans cette section aux autres **propriétés** de la conjuguée convexe.

2.1 Bi-conjuguée 二次共軛

Il est possible d'établir un grand nombre de propriétés pour la conjuguée convexe, mais il faut pour cela introduire la notion de bi-conjuguée :

Définition 2 (Bi-conjuguée convexe) 二次凸共軛

Soit $J : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ une fonction. On définit sa **bi-conjuguée convexe**, notée J^{**} , comme étant la **conjuguée convexe** de sa **conjuguée convexe** J^* .

En tant que **conjuguée convexe**, la **bi-conjuguée** présente les **mêmes propriétés** que la **conjuguée convexe** :

Corollaire 2

Soit $J : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ une fonction. Alors sa bi-conjuguée J^{**} est convexe et s.c.i.

Pour obtenir le caractère propre, il faut à nouveau ajouter une hypothèse supplémentaire sur J .

Proposition 3

Soit $J : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction propre. On suppose que J admet une minorante affine. Alors sa bi-conjuguée J^{**} est convexe, s.c.i. et propre.

DÉMONSTRATION : Laissée en exercice.

Quel lien existe-t-il entre une fonction et sa bi-conjuguée ? La proposition 4 répond à cette question dans le cas le plus général. Auparavant, démontrons le lemme suivant :

Lemme 1 (Inégalité de FENCHEL–YOUNG) *Fenchel-Young 不等式*

Soit $J : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ une fonction. Alors on a

$$\forall (x, y) \in \mathcal{X}^2, \quad J(x) + J^*(y) \geq \langle y, x \rangle$$

DÉMONSTRATION : Soit $y \in \mathcal{X}$. Par définition de la borne supérieure, on a

$$\forall x \in \mathcal{X}, \quad J^*(y) \geq \langle y, x \rangle - J(x)$$

de sorte que le résultat désiré est immédiat. ■

Une conséquence importante de l'inégalité de FENCHEL–YOUNG est la suivante : si J^* est propre, alors pour $y^0 \in \text{dom } J^*$, on a

$$\forall x \in \mathcal{X}, \quad J(x) \geq \langle y^0, x \rangle - J^*(y^0)$$

Ainsi, la conjugée convexe permet d'expliciter une minorante affine de J , tandis que, par définition de J^* , si le domaine de J est non vide, alors pour tout $x^0 \in \text{dom } J$,

$$\forall y \in \mathcal{X}, \quad J^*(y) \geq \langle y, x^0 \rangle - J(x^0)$$

Une fonction de domaine non vide et sa conjugée convexe définissent donc chacune une minorante affine pour l'autre. Une question intéressante est de savoir si elles permettent de définir une minorante affine exacte en tout point donné.

Proposition 4

Soit $J : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ une fonction. Alors on a

$$\forall x \in \mathcal{X}, \quad J^{**}(x) \leq J(x)$$

DÉMONSTRATION : Il suffit de passer à la borne supérieure sur y dans l'inégalité de FENCHEL–YOUNG. ■

En réalité, il est possible de montrer que J^{**} est la *plus grande fonction convexe* dont le graphe est située en-dessous de celui de le J . Plus précisément, la bi-conjuguée de J est l'enveloppe supérieure des minorantes affines de J :

上包络

Proposition 5

Soit $J : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ une fonction. Alors on a

$$\forall x \in \mathcal{X}, \quad J^{**}(x) = \sup \left\{ \langle a, x \rangle + b \mid \forall z \in \mathcal{X}, \quad \langle a, z \rangle + b \leq J(z) \right\}$$

DÉMONSTRATION : On va considérer deux cas de figure possibles.

- **Premier cas : J admet une minorante affine.** Soit $x \in \mathcal{X}$. Il existe par hypothèse $(a, b) \in \mathcal{X} \times \mathbb{R}$ tel que

$$\forall z \in \mathcal{X}, \quad \langle a, z \rangle + b \leq J(z) \quad \text{soit} \quad \langle a, z \rangle - J(z) \leq -b$$

En passant à la borne supérieure dans la seconde inégalité, on obtient que

$$J^*(a) \leq -b \quad \text{soit} \quad \langle a, x \rangle + b \leq \langle a, x \rangle - J^*(a) \leq J^{**}(x)$$

En passant à la borne supérieure en (a, b) définissant une minorante affine de J , on obtient

$$\sup \left\{ \langle a, x \rangle + b \mid \forall z \in \mathcal{X}, \quad \langle a, z \rangle + b \leq J(z) \right\} \leq J^{**}(x)$$

Démontrons à présent l'inégalité inverse. D'après l'inégalité de FENCHEL–YOUNG, pour tout $y \in \text{dom } J^*$, la fonction $z \mapsto \langle y, z \rangle - J^*(y)$ est une minorante affine de J . Par définition de la borne supérieure, on a donc pour tout $y \in \text{dom } J^*$,

$$\langle y, x \rangle - J^*(y) \leq \sup \left\{ \langle a, x \rangle + b \mid \forall z \in \mathcal{X}, \quad \langle a, z \rangle + b \leq J(z) \right\}$$

En passant à la borne supérieure sur les y , on obtient l'inégalité souhaitée.

- **Deuxième cas : J n'admet aucune minorante affine.** Dans ce cas, la conjuguée convexe J^* n'est pas propre (d'après la réciproque de la proposition 2). Puisque J^* est convexe, il s'ensuit qu'elle est identiquement égale à $+\infty$, et que sa conjuguée convexe est identiquement égale à $-\infty$. La borne supérieure d'un ensemble vide valant $-\infty$, on en déduit que l'identité annoncée est vérifiée. ■

Ainsi, lorsque J est convexe, on a le résultat central suivant :

Théorème 1 (FENCHEL–MOREAU)

Fenchel – Moreau 定理 (凸函数的二次凸共轭是凸函数本身)

Soit $J : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction convexe, s.c.i. et propre. Alors

$$J^{**} = J$$

Autrement dit, on a la relation suivante

$$\forall x \in \mathcal{X}, \quad J(x) = \sup_{y \in \mathcal{Y}} \left\{ \langle y, x \rangle - J^*(y) \right\}$$

DÉMONSTRATION : D'après la proposition précédente, il suffit de démontrer que

$$\forall x \in \mathcal{X}, \quad J(x) = \sup \left\{ \langle a, x \rangle + b \mid \forall z \in \mathcal{X}, \quad \langle a, z \rangle + b \leq J(z) \right\}$$

Autrement dit, J est l'enveloppe supérieure de ses minorantes affines ; c'est bien le cas sous les hypothèses du théorème. ■

REMARQUE : La réciproque dans le théorème de FENCHEL–MOREAU est vraie.

EXEMPLE

Conjuguée convexe de l'indicatrice de la boule unité. Considérons à nouveau la fonction convexe

$$J : \begin{cases} \mathcal{X} & \rightarrow \mathbb{R} \\ x & \mapsto \|x\| \end{cases}$$

On a montré dans un exemple précédent que J^* est l'indicatrice de boule unité fermée. Puisque J est convexe, s.c.i. et propre, on a $J^{**} = J$. Il s'ensuit donc en particulier que

$$\forall x \in \mathcal{X}, \quad \|x\| = \sup_{\|y\| \leq 1} \langle y, x \rangle$$

L'hypothèse de semi-continuité inférieure est essentielle dans le théorème de FENCHEL–MOREAU, comme le montre le contre-exemple suivant.

CONTRE-EXEMPLE

Cas d'une fonction non s.c.i. Considérons la fonction convexe définie pour tout $x \in \mathbb{R}$ par

$$J(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 & \text{si } x = 0 \\ +\infty & \text{si } x > 0 \end{cases}$$

Cette fonction n'est pas s.c.i. en 0. En effet, si on considère la suite $(1/k)_{k \in \mathbb{N}^*}$ de limite 0, on a

$$\lim_{k \rightarrow +\infty} J\left(\frac{1}{k}\right) = 0 < 1 = J(0)$$

ce qui contredit la définition de la semi-continuité inférieure. On peut vérifier (en exercice) que

$$J^* = \chi_{[0; +\infty[} \quad \text{et} \quad J^{**} = \chi_{]-\infty; 0]}$$

Ainsi, $J^{**} \neq J$.

Enfin, notons que l'on ne peut pas définir de tri-conjuguée (plus précisément, cette notion ne permet pas de définir de nouveaux objets).

Lemme 2

Soit $J_1 : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ et $J_2 : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ telles que

$$\forall x \in \mathcal{X}, \quad f(x) \leq g(x)$$

Alors

$$\forall y \in \mathcal{X}, \quad f^*(y) \geq g^*(y)$$

DÉMONSTRATION : Laissée en exercice.

Corollaire 3

Soit $J : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ une fonction. Alors

$$(J^{**})^* = J^*$$

DÉMONSTRATION : La proposition 4 appliquée à J^* assure d'une part que

$$\forall y \in \mathcal{X}, \quad (J^*)^{**}(y) \leq J^*(y)$$

D'autre part, puisque (toujours d'après la proposition 4), on a

$$\forall x \in \mathcal{X}, \quad J^{**}(x) \leq J(x)$$

le lemme précédent implique que

$$\forall y \in \mathcal{X}, \quad (J^{**})^*(y) \geq J^*(y)$$

On conclut en remarquant que $(J^*)^{**} = (J^{**})^*$. ■

2.2 Sous-différentiel de la conjuguée convexe 凸共軛的次可微性

Puisque la conjuguée convexe est toujours convexe, on peut aisément étudier son sous-différentiel.

Lemme 3 (Identité de LEGENDRE-FENCHEL)

Legendre-Fenchel 等式

Soit $J : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction convexe. Soit $(x, p) \in \mathcal{X}^2$. Alors on a équivalence entre les deux énoncés suivants :

- (i) $p \in \partial J(x)$
- (ii) $J(x) + J^*(p) = \langle p, x \rangle$

DÉMONSTRATION : Soit $(x, p) \in \mathcal{X}^2$. Il suffit d'écrire les équivalences suivantes :

$$\begin{aligned} J^*(p) = \langle p, x \rangle - J(x) &\iff \forall z \in \mathcal{X}, \quad \langle p, z \rangle - J(z) \leq \langle p, x \rangle - J(x) \\ &\iff \forall z \in \mathcal{X}, \quad J(z) \geq J(x) + \langle p, z - x \rangle \end{aligned}$$

soit $J^*(p) = \langle p, x \rangle - J(x) \iff p \in \partial J(x)$

en notant que les relations apparaissant dans cette série d'équivalences impliquent nécessairement que $x \in \text{dom } J$. Dans le cas contraire, aucune des deux affirmations (i) et (ii) n'est vraie. ■

Ce résultat peut s'interpréter à la lumière des minorantes affines : il assure que, dans le cas d'une fonction convexe propre J , la conjuguée convexe permet de définir une minorante affine exacte en tout point où la fonction J admet un sous-gradient.

On a en réalité un résultat plus fort, qui relie les sous-différentiels de J et de sa conjuguée convexe dans le cas convexe :

Proposition 6 (Règle de bascule) *bascule 原则*

Soit $J : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction **convexe**. Soit $(x, p) \in \mathcal{X}^2$. Alors on a

$$p \in \partial J(x) \quad \Longleftrightarrow \quad x \in \partial(J^*)(p)$$

DÉMONSTRATION : Soit $p \in \partial J(x)$. En utilisant l'identité de **LEGENDRE-FENCHEL**, montrer que

$$\forall y \in \mathcal{X}, \quad \langle p - y, x \rangle - J^*(y) \geq \langle p, x \rangle - \sup_{y \in \mathcal{X}} \{ \langle y, x \rangle - J^*(y) \} \geq J^*(p)$$

On reconnaît là la **définition** d'un **sous-gradient** de J^* en p ; plus précisément, on a démontré que $x \in \partial(J^*)(p)$. ■

Proposition 7 (Règle de bascule dans le cas convexe, s.c.i. et propre)

Soit $J : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction **convexe**, **s.c.i.** et **propre**. Soit $(x, p) \in \mathcal{X}^2$. Alors on a

$$p \in \partial J(x) \quad \Longleftrightarrow \quad x \in \partial(J^*)(p)$$

DÉMONSTRATION : Il suffit d'appliquer la règle de la bascule à la fonction J^* et d'utiliser l'identité entre J et sa bi-conjuguée. ■

2.3 Règles de calcul

Voyons quelques **règles de calcul** pour la **conjuguée convexe**.

Proposition 8 (Produit par un scalaire et changement d'échelle)

Soit $J : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ une fonction et $\alpha \in \mathbb{R}$. Alors on a

- $\forall y \in \mathcal{X}, \quad (\alpha J)^*(y) = \alpha J^*\left(\frac{y}{\alpha}\right) \quad \text{si } \alpha > 0$
- $\forall y \in \mathcal{X}, \quad (x \mapsto J(\alpha x))^*(y) = J^*\left(\frac{y}{\alpha}\right) \quad \text{si } \alpha \neq 0$

DÉMONSTRATION : Soit $\alpha > 0$ et $y \in \mathcal{X}$.

- Puisque α est **strictement positif**, on a

$$(\alpha J)^*(y) = \sup_{x \in \mathcal{X}} \{ \langle y, x \rangle - \alpha J(x) \} = \alpha \sup_{x \in \mathcal{X}} \left\{ \left\langle \frac{y}{\alpha}, x \right\rangle - J(x) \right\} = \alpha J^*\left(\frac{y}{\alpha}\right)$$

- En effectuant le changement de variables $u = \alpha x$, on obtient

$$(x \mapsto J(\alpha x))^*(y) = \sup_{x \in \mathcal{X}} \{ \langle y, x \rangle - J(\alpha x) \} = \sup_{u \in \mathcal{X}} \left\{ \left\langle y, \frac{u}{\alpha} \right\rangle - J(u) \right\} = J^*\left(\frac{y}{\alpha}\right)$$

ce qui donne le résultat souhaité. ■

Proposition 9 (Translation)

Soit $J : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ une fonction et $x^0 \in \mathcal{X}$. Alors on a

$$\forall y \in \mathcal{X}, \quad (x \mapsto J(x - x^0))^*(y) = J^*(y) + \langle y, x^0 \rangle$$

DÉMONSTRATION : Soit $y \in \mathcal{X}$.

$$\sup_{x \in \mathcal{X}} \{ \langle y, x \rangle - J(x - x^0) \} = \sup_{x \in \mathcal{X}} \{ \langle y, x - x^0 \rangle - J(x - x^0) \} + \langle y, x^0 \rangle = J^*(y) + \langle y, x^0 \rangle$$

ce qui achève la preuve. ■

EXERCICE

Conjuguée convexe d'une **distance euclidienne**. Soit $x^0 \in \mathcal{X}$. Considérons la fonction convexe

$$J : \begin{cases} \mathcal{X} & \rightarrow \mathbb{R} \\ x & \mapsto \frac{1}{2\tau} \|x - x^0\|^2 \end{cases}$$

Montrer que $\forall y \in \mathcal{X}, \quad J^*(y) = \frac{\tau}{2} \left\| y - \frac{x^0}{\tau} \right\|^2 - \frac{1}{2\tau} \|x^0\|^2$

Pour le cas de la **somme de deux fonctions convexes**, il faut au préalable introduire la définition de l'**inf-convolution** de deux fonctions :

Définition 3 (Inf-convolution) 下巻

Soit $f : \mathcal{X} \mapsto \overline{\mathbb{R}}$ et $g : \mathcal{X} \mapsto \overline{\mathbb{R}}$ deux fonctions. On définit leur **inf-convolution**, ou encore **convolution infimale**, comme la fonction notée $f \square g : \mathcal{X} \mapsto \overline{\mathbb{R}}$ définie par

$$\forall x \in \mathcal{X}, \quad f \square g(x) = \inf_{z \in \mathcal{X}} \{ f(x - z) + g(z) \}$$

Cet **opérateur généralise la régularisation de MOREAU-YOSIDA** (module **A5 : Opérateur proximal de MOREAU**). En effet, par définition de ${}^\tau J$ pour une fonction **J convexe**, on peut écrire

$$\forall x \in \mathcal{X}, \quad {}^\tau J(x) = J \square \left(z \mapsto \frac{1}{2\tau} \|z\|^2 \right)$$

L'**inf-convolution** est **commutative** ($f \square g = g \square f$) et possède en outre l'intéressante propriété de conserver la convexité.

Proposition 10 (Conjuguée convexe de l'inf-convolution)

Soit $J_1 : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ et $J_2 : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ deux fonctions. Alors on a

$$(J_1 \square J_2)^* = J_1^* + J_2^*$$

DÉMONSTRATION : Soit $y \in \mathcal{X}$. Par définition de l'**inf-convolution**, on a

$$\begin{aligned}(J_1 \square J_2)^*(y) &= \sup_{x \in \mathcal{X}} \left\{ \langle y, x \rangle - \inf_{z \in \mathcal{X}} \left\{ J_1(x - z) + J_2(z) \right\} \right\} \\ &= \sup_{x \in \mathcal{X}} \left\{ \langle y, x \rangle + \sup_{z \in \mathcal{X}} \left\{ -J_1(x - z) - J_2(z) \right\} \right\} \\ (J_1 \square J_2)^*(y) &= \sup_{(x, z) \in \mathcal{X}^2} \left\{ \langle y, x \rangle - J_1(x - z) - J_2(z) \right\}\end{aligned}$$

Réalisons le **changement** de variables $u = x - z$. On obtient alors

$$\begin{aligned}(J_1 \square J_2)^*(y) &= \sup_{(u, z) \in \mathcal{X}^2} \left\{ \langle y, u \rangle - J_1(u) + \langle y, z \rangle - J_2(z) \right\} \\ &= \sup_{u \in \mathcal{X}} \left\{ \langle y, u \rangle - J_1(u) \right\} + \sup_{z \in \mathcal{X}} \left\{ \langle y, z \rangle - J_2(z) \right\}\end{aligned}$$

On reconnaît alors l'expression de $J_1^*(y)$ et de $J_2^*(y)$. ■

Un **corollaire immédiat** de la proposition 10 dans le cas de fonctions **convexes**, **s.c.i.** et **propres** est le suivant :

Corollaire 4

Soit $J_1 : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ et $J_2 : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ deux fonctions **convexes**, **s.c.i.** et **propres**. Alors

$$J_1 + J_2 = (J_1^* \square J_2^*)^*$$

DÉMONSTRATION : Il suffit d'utiliser l'**identité** entre une fonction **convexe**, **s.c.i.** et **propre** et sa **bi-conjuguée**. ■

Il serait tentant d'en déduire immédiatement une **formule** pour la **conjuguée convexe** de la **somme** de deux **fonctions** convexes, **s.c.i.** et **propres**. Celle-ci découlerait immédiatement du **théorème de FENCHEL–MOREAU** si $(J_1^* \square J_2^*)^*$ est **propre**. L'exemple suivant montre que ce n'est pas toujours le cas, même en présence de fonctions très simples.

CONTRE-EXEMPLE

Cas d'une inf-convolution impropre. Considérons le cas des deux fonctions **convexes**, **continues** (donc **propres**) définies pour tout $x \in \mathbb{R}$ par

$$J_1(x) = x \quad \text{et} \quad J_2(x) = -x$$

Calculons leur **inf-convolution**. Par définition :

$$\forall x \in \mathbb{R}, \quad J_1 \square J_2(x) = \inf_{z \in \mathbb{R}} \left\{ x - z - z \right\} = \inf_{z \in \mathbb{R}} \left\{ x - 2z \right\} = -\infty$$

Théorème 2 (ATTOUCH–BREZIS)

On suppose que \mathcal{X} est de dimension finie. Soit $J_1 : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ et $J_2 : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ deux fonctions **convexes**, **s.c.i.** et **propres** telles que

$$0 \in \text{intr}(\text{dom } J_1 - \text{dom } J_2)$$

Alors

$$(J_1 + J_2)^* = J_1^* \square J_2^*$$

est convexe, **s.c.i.** et **propre**, et la borne inférieure est atteinte dans la définition de $J_1^* \square J_2^*$, c'est-à-dire que

$$\forall y \in \mathcal{X}, \exists x \in \mathcal{X}, \quad (J_1 + J_2)^*(y) = J_1^*(y - x) + J_2^*(x)$$

DÉMONSTRATION : Admis. La preuve se trouve dans [Bianchi].

3 Conjuguée convexe d'une fonction convexe

3.1 Identité de MOREAU

Proposition 11 (Identité de MOREAU) Moreau 等式

Soit $J : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction **convexe**, **propre** et **s.c.i.** Alors on a

$$\forall x \in \mathcal{X}, \quad x = \text{prox}_J(x) + \text{prox}_{J^*}(x)$$

DÉMONSTRATION : Soit $x^0 \in \mathcal{X}$. Posons $x^+ = \text{prox}_J(x^0)$ et $p = x^0 - x^+$. Nous allons montrer que $p = \text{prox}_{J^*}(x^0)$. Pour cela, utilisons la caractérisation du point proximal (proposition 5 du module **A5 : Opérateur proximal de MOREAU**), qui assure que

$$p = x^0 - x^+ \in \partial J(x^+)$$

La règle de bascule (proposition 6), on a

$$x^0 - p = x^+ \in \partial(J^*)(p)$$

ce qui, par la caractérisation du point proximal, assure que $p = \text{prox}_{J^*}(x^0)$. ■

Corollaire 5 (Identité de MOREAU généralisée)

Soit $J : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction **convexe**, **propre** et **s.c.i.** et $\tau > 0$. Alors on a

$$\forall x \in \mathcal{X}, \quad x = \text{prox}_{\tau J}(x) + \tau \text{prox}_{J^*/\tau}\left(\frac{x}{\tau}\right)$$

DÉMONSTRATION : En utilisant l'identité de MOREAU, on en revient à calculer $\text{prox}_{(\tau J)^*}$. Soit $x^0 \in \mathcal{X}$. D'après la proposition 8, on a

$$\begin{aligned} \text{prox}_{(\tau J)^*}(x^0) &= \underset{x \in \mathcal{X}}{\text{argmin}} \left\{ \tau J^*\left(\frac{x}{\tau}\right) + \frac{1}{2} \|x - x^0\|^2 \right\} \\ &= \underset{x \in \mathcal{X}}{\text{argmin}} \left\{ \frac{1}{\tau} J^*\left(\frac{x}{\tau}\right) + \frac{1}{2} \left\| \frac{x}{\tau} - \frac{x^0}{\tau} \right\|^2 \right\} \end{aligned}$$

On effectue alors un changement de variables. ■

3.2 Forte convexité et régularité

Lemme 4

Soit $J : \mathcal{X} \rightarrow \mathbb{R}$ une fonction **convexe** et continue. Soit $\alpha > 0$. Alors il y a équivalence entre les deux énoncés suivants :

- (i) J^* est fortement convexe de module α .
- (ii) La fonction $x \mapsto \frac{1}{2\alpha} \|x\|^2 - J(x)$ est convexe.

DÉMONSTRATION : Posons

$$f : \begin{cases} \mathcal{X} & \rightarrow \mathbb{R} \\ x & \mapsto \frac{1}{2\alpha} \|x\|^2 - J(x) \end{cases}$$

- (i) \implies (ii). Soit $(x, z) \in \mathcal{X}^2$ et $y \in \partial J(x)$ et $p \in \partial J(z)$. La fonction J^* étant supposée **fortement convexe**, on a

$$J^*(p) \geq J^*(y) + \langle x, p - y \rangle + \frac{\alpha}{2} \|p - y\|^2$$

D'après le lemme 3 on a donc

$$\langle p, z \rangle - J(z) \geq \langle y, x \rangle - J(x) + \langle x, p - y \rangle + \frac{\alpha}{2} \|p - y\|^2$$

soit, après réarrangement des termes et simplification,

$$\frac{1}{2\alpha} \|z\|^2 - J(z) \geq \frac{1}{2\alpha} \|x\|^2 - J(x) + \left\langle \frac{x}{\alpha} - y, z - x \right\rangle + \frac{1}{2\alpha} \|x - z + \alpha(p - y)\|^2$$

Ainsi, puisque le terme quadratique est positif, on a établi pour tout $x \in \mathcal{X}$ l'existence d'un vecteur $s = x/\alpha - y$ tel que

$$\forall z \in \mathcal{X}, \quad \frac{1}{2\alpha} \|z\|^2 - J(z) \geq \frac{1}{2\alpha} \|x\|^2 - J(x) + \langle s, z - x \rangle$$

En inversant les rôles de x et z , puis en multipliant la première inégalité par $\lambda \in [0; 1]$ et la seconde obtenue par $(1 - \lambda)$ et les sommant les deux relations que l'on obtient, on montre que la fonction f est convexe.

- (ii) \implies (i). Soit $\lambda \in [0; 1]$. Commençons par noter que, puisque f est **convexe**, on a pour tout $(x_1, x_2) \in \mathcal{X}^2$, en réarrangeant les termes dans la définition de la convexité,

$$-\lambda J(x_1) - (1 - \lambda) J(x_2) \geq -\frac{\lambda(1 - \lambda)}{2\alpha} \|x_1 - x_2\|^2 - J(\lambda x_1 + (1 - \lambda) x_2)$$

Soit $(y_1, y_2) \in \mathcal{X}^2$. Par définition de la conjuguée convexe, on a pour tout $(x_1, x_2) \in \mathcal{X}^2$

$$\lambda J^*(y_1) + (1 - \lambda) J^*(y_2) \geq \lambda \langle y_1, x_1 \rangle - \lambda J(x_1) + (1 - \lambda) \langle y_2, x_2 \rangle - (1 - \lambda) J(x_2)$$

En utilisant ce qui précède, on en déduit que

$$\lambda J^*(y_1) + (1 - \lambda) J^*(y_2) \geq \lambda \langle y_1, x_1 \rangle + (1 - \lambda) \langle y_2, x_2 \rangle - \frac{\lambda(1 - \lambda)}{2\alpha} \|x_1 - x_2\|^2 - J(\lambda x_1 + (1 - \lambda) x_2)$$

On peut vérifier que, pour tout $(x_1, x_2) \in \mathcal{X}^2$, on a

$$\begin{aligned} \lambda \langle y_1, x_1 \rangle + (1 - \lambda) \langle y_2, x_2 \rangle - \frac{\lambda(1 - \lambda)}{2\alpha} \|x_1 - x_2\|^2 \\ = \langle \lambda y_1 + (1 - \lambda) y_2, \lambda x_1 + (1 - \lambda) x_2 \rangle + \frac{\alpha}{2} \lambda(1 - \lambda) \|y_1 - y_2\|^2 \\ - \frac{\lambda(1 - \lambda)}{2\alpha} \|(x_1 - x_2) - \alpha(y_1 - y_2)\|^2 \end{aligned}$$

Posons $v = x_1 - x_2$. On a donc, en injectant l'égalité précédente dans l'inégalité qui précède,

$$\begin{aligned} \lambda J^*(y_1) + (1 - \lambda) J^*(y_2) &\geq \langle \lambda y_1 + (1 - \lambda) y_2, x_2 + \lambda v \rangle - J(x_2 + \lambda v) \\ &\quad - \frac{\lambda(1 - \lambda)}{2\alpha} \|v - \alpha(y_1 - y_2)\|^2 + \frac{\alpha}{2} \lambda(1 - \lambda) \|y_1 - y_2\|^2 \end{aligned}$$

Cette minoration est valable pour tout $(x_2, v) \in \mathcal{X}^2$. Aussi, on peut passer à la borne supérieure sur $x_2 \in \mathcal{X}$, ce qui donne

$$\begin{aligned} \lambda J^*(y_1) + (1 - \lambda) J^*(y_2) &\geq J^*(\lambda y_1 + (1 - \lambda) y_2) \\ &\quad + \frac{\alpha}{2} \lambda(1 - \lambda) \|y_1 - y_2\|^2 - \frac{\lambda(1 - \lambda)}{2\alpha} \|v - \alpha(y_1 - y_2)\|^2 \end{aligned}$$

puis sur $v \in \mathcal{X}$, ce qui donne finalement

$$\lambda J^*(y_1) + (1 - \lambda) J^*(y_2) \geq J^*(\lambda y_1 + (1 - \lambda) y_2) + \frac{\alpha}{2} \lambda(1 - \lambda) \|y_1 - y_2\|^2$$

Autrement dit, on a démontré la forte convexité de module α de J^* . ■

Notons que ce lemme [4](#) s'écrit également sous la forme suivante : si $J : \mathcal{X} \rightarrow \mathbb{R}$ une fonction **convexe** et continue, alors la fonction

$$y \mapsto J^*(y) - \frac{\alpha}{2} \|y\|^2$$

est convexe si et seulement si la fonction

$$x \mapsto \frac{1}{2\alpha} \|x\|^2 - J(x)$$

est convexe.

Proposition 12

Soit $J : \mathcal{X} \rightarrow \mathbb{R}$ une fonction **convexe** et L -régulière. Alors J^* est fortement convexe de module $1/L$.

DÉMONSTRATION : Le lemme de BAILLON-HADDAD (lemme 1 du module [A4 : Fonctions régulières](#)) assure que

$$\forall (x, x') \in \mathcal{X}^2, \quad \langle \nabla J(x) - \nabla J(x'), x - x' \rangle \geq \frac{1}{L} \|\nabla J(x) - \nabla J(x')\|^2$$

On peut alors vérifier que $\text{Id} - (1/L) \nabla J$ est lui-même fermement non-expansif, c'est-à-dire que, quels que soit $(x, x') \in \mathcal{X}^2$,

$$\left\langle x - \frac{1}{L} \nabla J(x) - x' + \frac{1}{L} \nabla J(x'), x - x' \right\rangle \geq \left\| x - \frac{1}{L} \nabla J(x) - x' + \frac{1}{L} \nabla J(x') \right\|^2$$

(il suffit de développer le membre de droite). Puisque L est strictement positif, on en déduit en divisant par L que

$$\langle Lx - \nabla J(x) - Lx' + \nabla J(x'), x - x' \rangle \geq 0$$

Autrement dit, la fonction différentiable $x \mapsto (L/2) \|x\|^2 - J(x)$ est convexe. D'après le lemme [4](#), on en déduit que J^* est fortement convexe de module $(1/L)$. ■

Proposition 13

Soit $J : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction **convexe**, s.c.i. et propre. Supposons que J est fortement convexe de module α . Alors J^* est $(1/\alpha)$ -régulière.

DÉMONSTRATION : On démontre successivement que J^* est différentiable puis que son gradient est lipschitzien.

- **J^* est différentiable.** Soit $p \in \text{dom } J^*$ et $x \in \partial J^*(y)$. La règle de bascule affirme que $p \in \partial J(x)$. Puisque J est fortement convexe, on a

$$\forall z \in \mathcal{X}, \quad J(z) \geq J(x) + \langle p, z - x \rangle + \frac{\alpha}{2} \|x - z\|^2$$

Soit $y \in \mathcal{X}$. On a alors

$$\begin{aligned} J^*(y) &\leq \sup_{z \in \mathcal{X}} \left\{ \langle y, z \rangle - J(z) \right\} \\ &\leq \sup_{z \in \mathcal{X}} \left\{ \langle y, z \rangle - J(x) - \langle p, z - x \rangle - \frac{\alpha}{2} \|x - z\|^2 \right\} \\ J^*(y) &\leq \langle y, x \rangle - J(x) + \sup_{z \in \mathcal{X}} \left\{ \langle y - p, z - x \rangle - \frac{\alpha}{2} \|x - z\|^2 \right\} \end{aligned}$$

On peut vérifier que la borne supérieure apparaissant dans la dernière ligne est atteinte ; il s'ensuit que

$$\begin{aligned} J^*(y) &\leq \langle y, x \rangle - J(x) + \frac{1}{2\alpha} \|y - p\|^2 \\ &\leq \langle p, x \rangle - J(x) + \langle y - p, x \rangle + \frac{1}{2\alpha} \|y - p\|^2 \\ J^*(y) &\leq J^*(p) + \langle y - p, x \rangle + \frac{1}{2\alpha} \|y - p\|^2 \end{aligned}$$

où la dernière ligne est obtenue en appliquant la définition de la conjuguée convexe. On en déduit d'une part que J^* est bien définie en tout point $y \in \mathcal{X}$, puisqu'elle y prend une valeur finie. D'autre part, la convexité de cette dernière assure par ailleurs que

$$J^*(p) + \langle y - p, x \rangle \leq J^*(y)$$

On a ainsi établi que

$$\forall y \in \mathcal{X}, \quad 0 \leq J^*(y) - J^*(p) - \langle y - p, x \rangle \leq \frac{1}{2\alpha} \|y - p\|^2$$

démontrant ainsi que J^* est différentiable en p , de gradient $\nabla J^*(p) = x$.

- **∇J^* est lipschitzien.** Soit $(p, q) \in \mathcal{X}^2$. D'après ce qui précède et la règle de bascule, on a

$$p \in \partial J(\nabla J^*(p)) \quad \text{et} \quad y \in \partial J(\nabla J^*(y))$$

Posons $x = \nabla J^*(p)$ et $z = \nabla J^*(y)$. Grâce à la **forte convexité** de J , le corollaire 1 du module **A2 : Sous-différentiabilité** écrit pour les couples (x, z) puis (z, x) donne, après sommation de deux inégalités obtenues,

$$\langle p - y, x - z \rangle \geq \alpha \|x - z\|^2$$

soit, en remplaçant x et z par leur définition,

$$\|p - y\| \|\nabla J^*(p) - \nabla J^*(y)\| \geq \langle p - y, \nabla J^*(p) - \nabla J^*(y) \rangle \geq \alpha \|\nabla J^*(p) - \nabla J^*(y)\|^2$$

où le premier membre est obtenu à l'aide de l'inégalité de CAUCHY-SCHWARZ. On obtient l'inégalité souhaitée en divisant par $\|\nabla J^*(p) - \nabla J^*(y)\|$ lorsque ce terme est non nul. ■

Corollaire 6

Soit $J : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction **convexe**, s.c.i. et propre. Alors on a l'équivalence entre les deux énoncés suivants :

- J est fortement convexe de module $1/L$
- J^* est L -régulière

De même, on a l'équivalence entre les deux énoncés suivants :

- J est L -régulière
- J^* est fortement convexe de module $1/L$

On voit donc la forte convexité peut toujours être interprétée comme une forme de régularité, à condition de passer dans le domaine dual.

DÉMONSTRATION : La preuve de ce résultat est une conséquence immédiate des deux propositions précédentes, dès lors que l'on a l'identité $J = J^{**}$. ■

Terminons ce module avec la démonstration d'un résultat admis dans le module A5, à savoir la régularité de la régularisation de MOREAU–YOSIDA. On commence par énoncer le lemme suivant :

Lemme 5

Soit $J : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction **convexe** et **propre**. Soit $\tau > 0$. Alors on a

$$\forall y \in \mathcal{X}, \quad ({}^\tau J)^*(y) = J^*(y) + \frac{\tau}{2} \|y\|^2$$

En particulier, on en déduit que $({}^\tau J)^*$ est une fonction **fortement convexe**, de module $1/\tau$.

DÉMONSTRATION : Il s'agit d'une conséquence directe de la proposition [10](#). ■

Proposition 14

Soit $J : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ une fonction **convexe**, **s.c.i.** et **propre**. Soit $\tau > 0$. Alors ${}^\tau J$ est une fonction **différentiable**. Son **gradient** est donné par

$$\forall x \in \mathcal{X}, \quad \nabla({}^\tau J)(x) = \frac{1}{\tau} (x - \text{prox}_{\tau J}(x))$$

DÉMONSTRATION : Démontrons séparément les deux affirmations.

- **${}^\tau J$ est différentiable.** Il s'agit d'une conséquence du lemme de BAILLON–HADDAD et de la proposition [13](#), qui assurent que $({}^\tau J)^{**}$ est τ -régulière. Or, ${}^\tau J$ étant convexe, s.c.i. et propre, on en déduit d'après le théorème de FENCHEL–MOREAU que

$$({}^\tau J)^{**} = {}^\tau J$$

- **Gradient de ${}^\tau J$.** On applique l'identité de MOREAU généralisée à J , ce qui donne après réarrangement des termes

$$\forall x \in \mathcal{X}, \quad \text{prox}_{J^*/\tau} \left(\frac{x}{\tau} \right) = \frac{1}{\tau} (x - \text{prox}_{\tau J}(x))$$

Intéressons-nous au membre de gauche. La caractérisation du point proximal assure que

$$p = \text{prox}_{J^*/\tau} \left(\frac{x}{\tau} \right) \iff x - \tau p \in \partial(J^*)(p)$$

Par ailleurs, en différentiant l'expression apparaissant dans le lemme de BAILLON–HADDAD, on obtient que

$$\forall p \in \mathcal{X}, \quad \partial(({}^\tau J)^*)(p) = \partial(J^*)(p) + \tau p$$

Ainsi, il en découle que

$$x - \tau p \in \partial(J^*)(p) \iff x \in \partial(({}^\tau J)^*)(p)$$

On conclut en utilisant la règle de bascule. ■

Pour aller plus loin

Conjuguée convexe. La conjugée convexe peut être utilisée pour introduire de la dualité dans un problème d'optimisation. Généralement, il est difficile de calculer la conjugée convexe d'une fonction objectif. Aussi, on se contente dans la grande majorité des cas d'une partie de la fonction objectif, voire de plusieurs parties (on introduit alors autant de variables duales). Il est à noter que, de la même manière que pour le lagrangien (ou lagrangien augmenté), la fonction de couplage associée est concave en la variable duale (même s'il y en a plusieurs, car elle est alors séparable en ces variables).

Dualité régularité/forte convexité. Puisque la régularité se traduit par la forte convexité dans le domaine dual (et *vice-versa*), il apparaît qu'en utilisant la conjugée convexe, il est possible de transformer un problème régulier (resp. fortement convexe) en un problème fortement convexe (resp. régulier). Or, certains algorithmes requièrent l'une ou l'autre de ces deux hypothèses, il peut donc être avantageux de considérer le problème primal plutôt que dual (ou l'inverse).