Module A6

Dualité min-max

Sauf mention contraire, $\mathcal{X}$ est un espace vectoriel réel de dimension finie, muni du produit scalaire noté $\langle \cdot, \cdot \rangle$ et on note $\| \cdot \|$ la norme euclidienne associée. Pour toutes les autres notations et définitions, le lecteur est invité à se reporter aux cours précédents.

1 Point-selle d'une fonction convexe-concave

1.1 Définition et exemples

Définition 1 (Point-selle)

Soit $\mathcal{L} : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{+\infty; -\infty\}$. On appelle *point-selle* de $\mathcal{L}$ tout point $(\bar{x}; \bar{y})$ de $\operatorname{dom} \mathcal{L}$ vérifiant l'encadrement suivant :

$$\forall \big((\bar{x}, y), (x, \bar{y})\big) \in (\operatorname{dom} \mathcal{L})^2, \qquad \mathcal{L}(\bar{x}; y) \leq \mathcal{L}(\bar{x}; \bar{y}) \leq \mathcal{L}(x; \bar{y})$$

On parle parfois de *point-col*. En anglais, on parle de *saddle point*.

Exemple

Point-selle. Considérons la fonction suivante :

$$\mathcal{L} : \left\{ \begin{array}{ccc} \mathbb{R} \times \mathbb{R} & \to & \mathbb{R} \\ (x, y) & \mapsto & x^2\, y^2 \end{array} \right.$$

Tous les points de la forme $(0, \tilde{y})$ avec $\tilde{y} \in \mathbb{R}$ sont des points-selles de $\mathcal{L}$. En effet, on a

$$\forall (x, y) \in \mathbb{R} \times \mathbb{R}, \qquad 0 = 0^2 \times y^2 \leq 0^2 \times \tilde{y}^2 \leq x^2 \times 0^2 = 0$$

Dans l'exemple qui suit, on étudie les points-selles d'une fonction dite convexe-concave (définition à suivre).

Exemple

Point-selle d'une fonction convexe-concave. Considérons la fonction

$$\mathcal{L} : \left\{ \begin{array}{ccc} \mathbb{R} \times \mathbb{R} & \to & \mathbb{R} \\ (x, y) & \mapsto & x^2 - y^2 \end{array} \right.$$

Cette fonction admet un point-selle $(0, 0)$, car

$$\forall (x, y) \in \mathbb{R} \times \mathbb{R}, \qquad -y^2 \leq 0 \leq x^2$$

Il s'agit par ailleurs de l'unique point-selle de $\mathcal{L}$, puisque tout point-selle $(\bar{x}, \bar{y})$ doit vérifier

$$0 = \bar{x}^2 - \bar{x}^2 \leq \bar{x}^2 - \bar{y}^2 \leq \bar{y}^2 - \bar{y}^2 = 0$$

impliquant que $\mathcal{L}(\bar{x}; \bar{y}) = 0$. Il s'ensuit que la définition du point-selle se lit

$$\forall\, y \in \mathbb{R}, \qquad \bar{x}^2 \leq y^2 \qquad \text{et} \qquad \forall\, x \in \mathbb{R}, \qquad \bar{y}^2 \leq x^2$$

En choisissant $x = y = 0$, il en découle que $(\bar{x}, \bar{y}) = (0, 0)$.

Dans ce cours, on va s'intéresser aux points-selles d'un cas de fonctions très particulières qui apparaîtront naturellement dans les problèmes d'optimisation.

Définition 2 (Fonction convexe-concave)

Soit $\mathcal{L} : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{+\infty; -\infty\}$. On dit que $\mathcal{L}$ est *convexe-concave* si son domaine est non vide et que

- pour tout $y \in \mathcal{Y}$, la fonction partielle $x \mapsto \mathcal{L}(x; y)$ est convexe ;
- pour tout $x \in \mathcal{X}$, la fonction partielle $y \mapsto \mathcal{L}(x; y)$ est concave.

Exercice

Fonctions affines. Montrer que les fonctions affines sont convexes-concaves.

Exemple

Fonctions convexes-concaves additivement séparables. Voici un autre exemple simple de fonctions convexes-concaves :

$$\mathcal{L}(x; y) = f(x) - g(y)$$

construit à l'aide de f et g deux fonctions convexes.

Exemple

Fonctions convexes-concaves multiplicativement séparables. Il est également possible de construire des fonctions convexes-concaves à l'aide du produit d'une fonction convexe avec une fonction concave, sous conditions que ces deux fonctions soient à valeurs positives ; on pourra citer l'exemple de la fonction

$$\mathcal{L} : \left\{ \begin{array}{ccc} \mathbb{R} \times \mathbb{R} & \to & \mathbb{R} \\ (x, y) & \mapsto & \exp(x)\, \ln(1 + y) \end{array} \right.$$

Définition 3 (Fonction convexe-concave propre)

Soit $\mathcal{L} : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{+\infty; -\infty\}$ une fonction convexe-concave. On dit que $\mathcal{L}$ est propre si les deux conditions suivantes sont vérifiées :

(i) il existe $y^0 \in \mathcal{Y}$ tel que la fonction partielle $x \mapsto \mathcal{L}(x; y^0)$ soit propre ;

(ii) il existe $x^0 \in \mathcal{X}$ tel que la fonction partielle $y \mapsto -\mathcal{L}(x^0, y)$ soit propre.

Autrement dit, le domaine de $\mathcal{L}$ est non vide.

1.2 Conditions d'existence

On se concentre ici sur le cas des fonctions convexes-concaves. Pour caractériser les points-selles (lorsqu'ils existent), il est utile d'introduire la notion de saut de dualité.

Définition 4 (Saut de dualité)

Soit $\mathcal{L} : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{+\infty; -\infty\}$ une fonction. On définit le *saut de dualité* de $\mathcal{L}$ comme étant la quantité donnée par

$$\mathcal{G} = \inf_{x\in\mathcal{X}} \sup_{y\in\mathcal{Y}} \mathcal{L}(x;y) - \sup_{y\in\mathcal{Y}} \inf_{x\in\mathcal{X}} \mathcal{L}(x;y)$$

En anglais, on parle de *duality gap*. Le saut de dualité peut prendre une valeur infinie, mais lorsqu'il est fini, il est nécessairement positif.

Proposition 1 (Dualité faible)

Soit $\mathcal{L} : \mathcal{X}\times\mathcal{Y} \to \mathbb{R}\cup\{+\infty; -\infty\}$ une fonction. On suppose qu'il existe $(x^0, y^0) \in \mathcal{X} \times \mathcal{Y}$ tel que

$$\forall\,(x,y) \in \mathcal{X} \times \mathcal{Y}, \qquad \mathcal{L}(x; y^0) > -\infty \qquad \text{et} \qquad \mathcal{L}(x^0, y) < +\infty$$

Alors
$$\sup_{y'\in\mathcal{Y}} \inf_{x\in\mathcal{X}} \mathcal{L}(x;y') \le \inf_{x'\in\mathcal{X}} \sup_{y\in\mathcal{Y}} \mathcal{L}(x';y)$$

Autrement dit, en utilisant éventuellement les règles de calculs sur la droite réelle achevée, le saut de dualité vérifie $\mathcal{G} \in [\,0\,;+\infty\,]$.

Remarque : La condition portant sur l'existence du point (x^0, y^0) se traduit, dans le cas où $\mathcal{L}$ est convexe-concave, par le fait que $\mathcal{L}$ est propre.

Démonstration : Soit $(x', y') \in \mathcal{X} \times \mathcal{Y}$. Puisque, par hypothèse, $\mathcal{L}(x'; y^0) > -\infty$, on en déduit par définition de la borne supérieure que

$$\forall\, y' \in \mathcal{Y}, \qquad \mathcal{L}(x';y') \le \sup_{y\in\mathcal{Y}} \mathcal{L}(x';y) \in \mathbb{R} \cup \{+\infty\}$$

On démontre de la même façon que

$$\forall\, x' \in \mathcal{X}, \qquad \mathcal{L}(x';y') \ge \inf_{x\in\mathcal{X}} \mathcal{L}(x;y') \in \mathbb{R} \cup \{-\infty\}$$

En particulier, on a (les deux termes ci-dessous étant possiblement infinis)

$$\forall\,(x',y') \in \mathcal{X} \times \mathcal{Y}, \qquad \inf_{x\in\mathcal{X}} \mathcal{L}(x;y') \le \sup_{y\in\mathcal{Y}} \mathcal{L}(x';y)$$

Puisque le membre de gauche ne dépend pas de x', ni celui de droite de y', on peut passer successivement à la borne supérieure, puis la borne inférieure :

$$\sup_{y'\in\mathcal{Y}} \inf_{x\in\mathcal{X}} \mathcal{L}(x;y') \le \inf_{x'\in\mathcal{X}} \sup_{y\in\mathcal{Y}} \mathcal{L}(x';y)$$

ce qui est le résultat souhaité. ■

Le saut de dualité donne une information importante sur l'existence d'un point-selle d'une fonction, comme en témoigne le résultat suivant.

Proposition 2 (Existence d'un point-selle)

Soit $\mathcal{L} : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{+\infty; -\infty\}$ une fonction. Soit $(\bar{x}, \bar{y}) \in \operatorname{dom} \mathcal{L}$. Alors $(\bar{x}, \bar{y})$ est un point-selle de $\mathcal{L}$ si et seulement si les trois conditions suivantes sont vérifiées :

(i) la fonction $\quad x \mapsto \sup_{y\in\mathcal{Y}} \mathcal{L}(x;y) \quad$ atteint son minimum en $\bar{x}$;

(ii) la fonction $\quad y \mapsto \inf_{x\in\mathcal{X}} \mathcal{L}(x;y) \quad$ atteint son maximum en $\bar{y}$;

(iii) son saut de dualité $\mathcal{G}$ est nul.

Dans ce cas, le point $(\bar{x}, \bar{y})$ vérifie

$$\mathcal{L}(\bar{x};\bar{y}) = \max_{y\in\mathcal{Y}} \inf_{x\in\mathcal{X}} \mathcal{L}(x;y) = \min_{x\in\mathcal{X}} \sup_{y\in\mathcal{Y}} \mathcal{L}(x;y)$$

Cette valeur est parfois appelée *valeur-selle*.

La troisième propriété est appelée *dualité forte* ; elle équivaut à

$$\sup_{y'\in\mathcal{Y}} \inf_{x\in\mathcal{X}} \mathcal{L}(x;y') = \inf_{x'\in\mathcal{X}} \sup_{y\in\mathcal{Y}} \mathcal{L}(x';y)$$

Démonstration : Démontrons séparément les deux sens de la proposition.

- **Sens direct.** On suppose que $\mathcal{L}$ admet un point-selle $(\bar{x}, \bar{y})$. Par définition, on a alors que

  $$\forall\,(x,y) \in \mathcal{X} \times \mathcal{Y}, \qquad \mathcal{L}(\bar{x};y) \le \mathcal{L}(\bar{x};\bar{y}) \le \mathcal{L}(x;\bar{y})$$

  En passant à la borne supérieure sur y dans le première inégalité, et à la borne inférieure dans la seconde, on obtient que

  $$\sup_{y\in\mathcal{Y}} \mathcal{L}(\bar{x};y) \le \mathcal{L}(\bar{x};\bar{y}) \le \inf_{x\in\mathcal{X}} \mathcal{L}(x;\bar{y})$$

  Puisque, par ailleurs,

  $$\inf_{x\in\mathcal{X}} \sup_{y\in\mathcal{Y}} \mathcal{L}(\bar{x};y) \le \sup_{y\in\mathcal{Y}} \mathcal{L}(\bar{x};y) \qquad \text{et} \qquad \inf_{x\in\mathcal{X}} \mathcal{L}(x;\bar{y}) \le \sup_{y\in\mathcal{Y}} \inf_{x\in\mathcal{X}} \mathcal{L}(x;\bar{y})$$

  on obtient l'inégalité suivante :

  $$\inf_{x\in\mathcal{X}} \sup_{y\in\mathcal{Y}} \mathcal{L}(\bar{x};y) \le \sup_{y\in\mathcal{Y}} \inf_{x\in\mathcal{X}} \mathcal{L}(x;\bar{y})$$

  Or, on a démontré plus haut que

  $$\inf_{x\in\mathcal{X}} \sup_{y\in\mathcal{Y}} \mathcal{L}(\bar{x};y) \ge \sup_{y\in\mathcal{Y}} \inf_{x\in\mathcal{X}} \mathcal{L}(x;\bar{y})$$

  On en déduit l'égalité dans cette relation. Celle-ci implique en particulier que toutes les inégalités écrites ci-dessus sont également des égalités ; on a alors notamment les deux égalités

  $$\inf_{x\in\mathcal{X}} \sup_{y\in\mathcal{Y}} \mathcal{L}(\bar{x};y) = \sup_{y\in\mathcal{Y}} \mathcal{L}(\bar{x};y) \qquad \text{et} \qquad \inf_{x\in\mathcal{X}} \mathcal{L}(x;\bar{y}) = \sup_{y\in\mathcal{Y}} \inf_{x\in\mathcal{X}} \mathcal{L}(x;\bar{y})$$

ce qui implique, puisque ces deux quantités sont égales à la valeur-selle (qui est finie par définition du point-selle), que les fonctions

$$x \mapsto \sup_{y\in\mathcal{Y}} \mathcal{L}(x;y) \qquad \text{et} \qquad y \mapsto \inf_{x\in\mathcal{X}} \mathcal{L}(x;y)$$

atteignent respectivement leur minimum et leur maximum (la première en $\bar{x}$ et la seconde en $\bar{y}$), ce qui permet d'écrire

$$\inf_{x\in\mathcal{X}} \sup_{y\in\mathcal{Y}} \mathcal{L}(x;y) = \min_{x\in\mathcal{X}} \sup_{y\in\mathcal{Y}} \mathcal{L}(x;y) \quad \text{et} \quad \sup_{y\in\mathcal{Y}} \inf_{x\in\mathcal{X}} \mathcal{L}(x;y) = \max_{y\in\mathcal{Y}} \inf_{x\in\mathcal{X}} \mathcal{L}(x;y)$$

- **Réciproque.** Supposons que le saut de dualité est nul et qu'il existe deux points $\bar{x} \in \mathcal{X}$ et $\bar{y} \in \mathcal{Y}$ tels que

$$\sup_{y\in\mathcal{Y}} \mathcal{L}(\bar{x};y) = \min_{x\in\mathcal{X}} \sup_{y\in\mathcal{Y}} \mathcal{L}(x;y) \qquad \text{et} \qquad \inf_{x\in\mathcal{X}} \mathcal{L}(x;\bar{y}) = \max_{y\in\mathcal{Y}} \inf_{x\in\mathcal{X}} \mathcal{L}(x;y)$$

Alors, par définition des bornes inférieure et supérieure, on a

$$\mathcal{L}(\bar{x};\bar{y}) \leq \sup_{y\in\mathcal{Y}} \mathcal{L}(\bar{x};y) = \min_{x\in\mathcal{X}} \sup_{y\in\mathcal{Y}} \mathcal{L}(x;y) = \max_{y\in\mathcal{Y}} \inf_{x\in\mathcal{X}} \mathcal{L}(x;y) = \inf_{x\in\mathcal{X}} \mathcal{L}(x;\bar{y})$$

puis

$$\forall\, x \in \mathcal{X}, \qquad \mathcal{L}(\bar{x};\bar{y}) \leq \inf_{x\in\mathcal{X}} \mathcal{L}(x;\bar{y}) \leq \mathcal{L}(x;\bar{y})$$

On démontre de la même façon que,

$$\forall\, y \in \mathcal{Y}, \qquad \mathcal{L}(\bar{x};y) \leq \mathcal{L}(\bar{x};\bar{y})$$

ce qui assure que $(\bar{x},\bar{y})$ est un point-selle de $\mathcal{L}$. ■

L'annulation du saut de dualité ne permet pas à elle seule d'assurer l'existence d'un point-selle ; les deux conditions (i) et (ii) dans la proposition précédente sont essentielles, comme le montre l'exemple suivant.

Contre-exemple

Saut de dualité nul. Considérons la fonction convexe-concave

$$\mathcal{L} : \left\{ \begin{array}{ccc} \mathbb{R}\times\mathbb{R} & \to & \mathbb{R} \\ (x,y) & \mapsto & \exp(x) - y^2 \end{array} \right.$$

Calculons son saut de dualité : on a d'une part

$$\inf_{x\in\mathbb{R}} \sup_{y\in\mathbb{R}} \left\{ \exp(x) - y^2 \right\} = \inf_{x\in\mathbb{R}} \left\{ \exp(x) + \sup_{y\in\mathbb{R}} -y^2 \right\} = \inf_{x\in\mathbb{R}} \exp(x) = 0$$

et d'autre part

$$\sup_{y\in\mathbb{R}} \inf_{x\in\mathbb{R}} \left\{ \exp(x) - y^2 \right\} = \sup_{y\in\mathbb{R}} \left\{ \inf_{x\in\mathbb{R}} \exp(x) - y^2 \right\} = \sup_{y\in\mathbb{R}} \left\{ -y^2 \right\} = 0$$

Le lecteur peut vérifier que son saut de dualité est nul. Or, comme on le démontrera plus loin dans ce paragraphe (proposition **??**), les points-selles d'une fonction différentiable convexe-concave sont exactement ses points critiques, c'est-à-dire ici les points annulant le gradient

$$\nabla\mathcal{L}(x;y) = \begin{pmatrix} \exp(x) \\ -2\,y \end{pmatrix}$$

qui, dans ce cas précis, n'admet pas de zéro. Cela tient au fait que la première condition dans la proposition précédente n'est pas satisfaite.

Contrairement au minimum, qui permet de caractériser les minimiseurs d'une fonction, dans le sens où

$$J(x) = \min_{x\in\mathcal{X}} J(x) \qquad \Longleftrightarrow \qquad x \in \operatorname*{argmin}_{x\in\mathcal{X}} J(x)$$

la valeur-selle ne caractérise pas les points-selles du fonction, comme on peut le montrer dans l'exemple suivant.

Contre-exemple

Valeur-selle. Considérons à nouveau la fonction convexe-concave

$$\mathcal{L} : \left\{ \begin{array}{ccc} \mathbb{R}\times\mathbb{R} & \to & \mathbb{R} \\ (x,y) & \mapsto & x^2 - y^2 \end{array} \right.$$

dont on a vu que le point $(0,0)$ est l'unique point-selle. Sa valeur-selle vaut donc $\mathcal{L}(0,0) = 0$. Or, il est aisé de vérifier que tous les points $(x,x) \in \mathbb{R}^2$ prennent la même valeur par $\mathcal{L}$, mais ne sont pas des points-selles si $x \neq 0$.

Un corollaire du résultat précédent découle immédiatement, lorsque les fonctions convexes en jeu sont coercives.

Corollaire 1

Soit $\mathcal{L} : \mathcal{X}\times\mathcal{Y} \to \mathbb{R}\cup\{+\infty;-\infty\}$ une fonction. On suppose son domaine fermé et non vide. On suppose que les trois conditions suivantes sont vérifiées :

(i) la fonction partielle $x \mapsto \mathcal{L}(x;y)$ est s.c.i. et coercive pour tout $y \in \mathcal{Y}$;

(ii) la fonction partielle $y \mapsto \mathcal{L}(x;y)$ est s.c.i. et coercive pour tout $x \in \mathcal{X}$;

(iii) son saut de dualité $\mathcal{G}$ est nul ;

Alors $\mathcal{L}$ admet un point-selle.

Démonstration : Il s'agit d'une conséquence de la proposition 6 du module B1 et de la proposition **??**. ■

Voici une autre propriété des points-selles, qui repose sur le calcul différentiel.

Proposition 3

Soit $\mathcal{L} : \mathcal{X}\times\mathcal{Y} \to \mathbb{R}\cup\{+\infty;-\infty\}$ une fonction. On suppose que $(\bar{x},\bar{y}) \in \mathcal{X}\times\mathcal{Y}$ est un point-selle de $\mathcal{L}$. Alors on a

$$0 \in \partial_x\mathcal{L}(\bar{x};\bar{y}) \qquad \text{et} \qquad 0 \in \partial_y(-\mathcal{L})(\bar{x};\bar{y})$$

Rappelons qu'en général,

$$\partial_y(-\mathcal{L})(x;y) \neq -\partial_y\mathcal{L}(x;y)$$

(il suffit de reprendre l'exemple de la fonction $y \mapsto -|y|$).

Démonstration : Remarquons (comme cela a déjà été fait plus haut) que $\mathcal{L}$ admet un point-selle $(\bar{x}, \bar{y})$ si et seulement si $\bar{x}$ est un minimiseur de la fonction partielle $x \mapsto \mathcal{L}(x; \bar{y})$ et $\bar{y}$ est un minimiseur de la fonction partielle $y \mapsto -\mathcal{L}(\bar{x}; y)$. La règle de Fermat s'écrit dans ce cas

$$0 \in \partial\big(x \mapsto \mathcal{L}(x; \bar{y})\big)(\bar{x}) \qquad \text{et} \qquad 0 \in \partial\big(y \mapsto -\mathcal{L}(\bar{x}; y)\big)(\bar{y})$$

On obtient le résultat désiré en reconnaissant les sous-différentiels partiels de $\mathcal{L}$. ■

Dans le cas d'une fonction convexe-concave, la propriété précédente caractérise les points-selles :

Proposition 4 (Caractérisation d'un point-selle)

Soit $\mathcal{L} : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{+\infty; -\infty\}$ une fonction **convexe-concave** propre. Alors on a l'équivalence entre deux affirmations suivantes :

(i) le point $(\bar{x}, \bar{y}) \in \mathcal{X} \times \mathcal{Y}$ est un point-selle de $\mathcal{L}$;

(ii) le point $(\bar{x}, \bar{y}) \in \mathcal{X} \times \mathcal{Y}$ vérifie

$$0 \in \partial_x \mathcal{L}(\bar{x}; \bar{y}) \qquad \text{et} \qquad 0 \in \partial_y(-\mathcal{L})(\bar{x}; \bar{y})$$

et le saut de dualité de $\mathcal{L}$ est nul.

Remarque : Ces deux sous-différentiels partiels sont défini comme les sous-différentiels des deux fonctions convexes

$$x \mapsto \mathcal{L}(x; \bar{y}) \qquad \text{et} \qquad y \mapsto -\mathcal{L}(\bar{x}; y)$$

Démonstration : Le sens direct a déjà été démontré, la réciproque provient de la réciproque dans la règle de Fermat dans le cas convexe. ■

Corollaire 2

Soit $\mathcal{L} : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ une fonction. On suppose qu'elle admet la décomposition suivante :

$$\forall\, (x, y) \in \mathcal{X} \times \mathcal{Y}, \qquad \mathcal{L}(x; y) = f(x) + h(x, y)$$

avec $h : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ une fonction continûment différentiable. On suppose que $(\bar{x}, \bar{y}) \in \mathcal{X} \times \mathcal{Y}$ est un point-selle de $\mathcal{L}$. Alors on a

$$0 \in \partial\mathcal{L}(\bar{x}; \bar{y})$$

Si, de plus, $\mathcal{L}$ est **convexe-concave**, alors la réciproque est vraie.

Démonstration : Il suffit de rappeler que (cf. proposition 17 du module A2)

$$\forall\, (x, y) \in \operatorname{dom} \mathcal{L}, \qquad \partial\mathcal{L}(x; y) = \partial_x \mathcal{L}(x; y) \times \partial_z \mathcal{L}(x; y)$$

Par ailleurs, on a $\partial_y \mathcal{L}(x; y) = \{\nabla \mathcal{L}(x; y)\}$, de sorte que, dans ce cas précis,

$$\partial_y(-\mathcal{L})(x; y) = -\partial_y \mathcal{L}(x; y)$$

On peut alors appliquer les propositions ?? et ??. ■

En particulier, ce résultat assure que les points-selles d'une fonction convexe-concave différentiable sont exactement ses points critiques.

> Attention : dans le cas d'une fonction non différentiable, la caractérisation des points-selles d'une fonction convexe-concave $\mathcal{L}$ n'en fait pas nécessairement des points critiques de $\mathcal{L}$, qui sont généralement distincts des points critiques partiels.

Contre-exemple

Points-selles et points critiques d'une fonction non différentiable. Considérons la fonction convexe-concave suivante :

$$\mathcal{L} : \left\{ \begin{array}{ccc} \mathbb{R} \times \mathbb{R} & \to & \mathbb{R} \\ (x, y) & \mapsto & -|y| \end{array} \right.$$

Il est aisé de vérifier que ses points-selles sont les points

$$\Big\{ (\bar{x}, 0) \in \mathbb{R}^2 \mid \bar{x} \in \mathbb{R} \Big\}$$

Intéressons-nous donc à ses points critiques. Commençons par remarquer que, pour tout $(x, y) \in \mathbb{R} \times \mathbb{R}^*$, la fonction $\mathcal{L}$ est différentiable au voisinage de (x, y). Ainsi, on a

$$\partial\mathcal{L}(x; y) = \hat{\partial}\mathcal{L}(x; y) = \Big\{ \nabla\mathcal{L}(x; y) \Big\} = \begin{cases} \{(0, \;\; 1)\} & \text{si } y < 0 \\ \{(0, -1)\} & \text{si } y > 0 \end{cases}$$

On en déduit donc qu'aucun de ces points n'est un point critique de $\mathcal{L}$. Étudions à présent les autres points, c'est-à-dire les points-selles de $\mathcal{L}$. D'après ce qui précède, on a par passage à la limite

$$\Big\{ (0, 1), (0, -1) \Big\} \in \partial\mathcal{L}(x; 0)$$

En réalité, on peut démontrer que $(0, 0) \in \partial\mathcal{L}(x; 0)$ si et seulement s'il existe une suite $((x_k, 0))_{k \in \mathbb{N}}$ de $\mathbb{R}^2$ telle que $(x_k)_{k \in \mathbb{N}}$ tend vers 0 et que

$$\forall\, k \in \mathbb{N}, \exists (p_k, q_k) \in \mathbb{R}^2, \quad (p_k, q_k) \in \hat{\partial}\mathcal{L}(x_k, 0) \qquad \text{avec} \qquad \lim_{k \to +\infty} (p_k, q_k) = (0, 0)$$

Or, les (p_k, q_k) doivent vérifier par définition

$$-|y| \geq p_k\,(x_k - x) + o(|x - x_k|) + o(|y|)$$

En faisant tendre x vers x_k, cette inégalité se lit

$$-|y| \geq |y|\, \varepsilon(|y|)$$

avec $\varepsilon(t) \to 0$ lorsque $t \to 0$. Si $y \neq 0$, alors on peut simplifier par $|y|$, puis faire tendre y vers 0. On aboutit alors à une contradiction. On démontre de la sorte que les points-selles de $\mathcal{L}$ ne sont pas des points critiques de $\mathcal{L}$.

2 Dualité min-max en optimisation convexe

Dans cette section, on s'intéresse au problème de minimisation suivant

$$\min_{x\in\mathcal{X}} J(x) \tag{$\mathcal{P}$}$$

où $J : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ est une fonction.

2.1 Principe

Le principe de la dualité min-max est le suivant : on cherche à représenter la fonction objectif J du problème $(\mathcal{P})$ comme le supremum d'une famille de fonctions $\mathcal{L}(\cdot; y)$ paramétrée par $y \in \mathcal{Y}$:

$$\forall\, x \in \mathcal{X}, \qquad J(x) = \sup_{y\in\mathcal{Y}} \mathcal{L}(x; y)$$

de sorte que le problème de minimisation s'écrive comme un problème de minimisation-maximisation :

$$\min_{x\in\mathcal{X}} \sup_{y\in\mathcal{Y}} \mathcal{L}(x; y)$$

La fonction $\mathcal{L}$ est, dans le cas général, appelée *fonction de couplage*. On montrera un peu plus loin qu'il existe un lien entre les minimiseurs de J et les points-selles de la fonction de couplage $\mathcal{L}$ sous certaines conditions.

2.2 Problème dual

La réécriture du problème d'optimisation $(\mathcal{P})$ comme un problème de minimisation-maximisation conduit naturellement à considérer le problème miroir, à savoir le problème de maximisation-minimisation suivant :

$$\max_{y\in\mathcal{Y}} \inf_{x\in\mathcal{X}} \mathcal{L}(x; y) \tag{$\mathcal{D}$}$$

où l'on a simplement inversé l'ordre des deux optimisations. Ce problème peut être interprété comme la maximisation de la fonction

$$E : y \mapsto \inf_{x\in\mathcal{X}} \mathcal{L}(x; y)$$

On parle alors pour le problème $(\mathcal{D})$ de *problème dual*, dont la fonction objectif est également qualifiée de *duale* ; par opposition, le problème $(\mathcal{P})$ et sa fonction objectif J sont qualifiés de *primaux*. On verra dans la suite sous quelles conditions les problèmes primaux et duaux peuvent être liés, et comment on peut exploiter ce lien pour résoudre le problème primal.

On peut d'ores-et-déjà remarquer que

$$\sup_{y\in\mathcal{Y}} E(y) = -\inf_{y\in\mathcal{Y}} -E(y) \qquad \text{et} \qquad \arg\max_{y\in\mathcal{Y}} E(y) = \arg\min_{y\in\mathcal{Y}} -E(y)$$

On dit alors que le problème dual est **convexe** lorsque le problème de minimisation équivalent l'est, autrement dit lorsque E est concave.

Notons également que, pour un problème primal donné, il est possible de définir plusieurs problèmes duaux associés à différentes fonctions de couplage. Ainsi, dans la suite, lorsque l'on parlera *du* problème dual, on sous-entendra toujours que la fonction de couplage est connue.

Remarque : Attention, même si le problème de minimisation primal est bien posé (c'est-à-dire qu'il admet au moins une solution, un minimiseur de J), ce n'est pas nécessairement le cas du problème dual.

2.3 Convexité des problèmes primal et dual

Les propriétés de convexité (partielle) de la fonction de couplage peuvent entraîner la convexité des problèmes primal et dual.

Proposition 5

Soit $\mathcal{L} : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{+\infty; -\infty\}$. Si la fonction partielle $x \mapsto \mathcal{L}(x; y)$ est **convexe** pour tout $y \in \mathcal{Y}$, alors

$$x \mapsto \sup_{y\in\mathcal{Y}} \mathcal{L}(x; y)$$

est une fonction convexe.

Démonstration : La fonction considérée dans la proposition peut être vue comme l'exemple supérieure de la famille de fonctions

$$x \mapsto \mathcal{L}(x; y)$$

convexe pour tout $y \in \mathcal{Y}$. Elle est donc convexe en vertu de la proposition 10 du module A1. ■

On peut de même démontrer le résultat analogue suivant :

Corollaire 3

Soit $\mathcal{L} : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{+\infty; -\infty\}$. Si la fonction partielle $y \mapsto \mathcal{L}(x; y)$ est concave pour tout $x \in \mathcal{X}$, alors

$$y \mapsto \inf_{x\in\mathcal{X}} \mathcal{L}(x; y)$$

est une fonction concave.

Ainsi, même si le problème primal n'est pas convexe, on peut lui associer un problème dual convexe. Par ailleurs, si J n'est pas convexe, alors il n'est pas possible de le représenter à l'aide d'une fonction de couplage convexe-concave (par exemple). En revanche, si $\mathcal{L}$ est une fonction convexe-concave, alors les problèmes primal et dual sont des problèmes de minimisation convexe.

2.4 Lien entre minimiseurs de J et points-selles de $\mathcal{L}$

Dans ce paragraphe, on va s'intéresser au lien qui peut exister entre les minimiseurs de J (en supposant qu'ils existent) et les points-selles de la fonction de couplage $\mathcal{L}$.

On suppose qu'il existe une fonction de couplage $\mathcal{L}$ telle que

$$\forall\, x \in \mathcal{X}, \qquad J(x) = \sup_{y\in\mathcal{Y}} \mathcal{L}(x; y)$$

La proposition **??** donne alors une condition nécessaire et suffisante d'optimalité, sous des hypothèses d'existence de minimiseurs et de maximiseurs pour les problèmes primal et dual, respectivement. On peut traduire ce résultat dans le cadre de l'optimisation :

Corollaire 4

Soit $J : \mathcal{X} \mapsto \mathbb{R} \cup \{+\infty\}$ une fonction s.c.i. On suppose qu'il existe une fonction s.c.i. $\mathcal{L} : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{+\infty; -\infty\}$ telle que

$$\forall\, x \in \mathcal{X}, \qquad J(x) = \sup_{y \in \mathcal{Y}} \mathcal{L}(x; y)$$

Soit $(x^*, y^*) \in \operatorname{dom} \mathcal{L}$. Alors (x^*, y^*) est un point-selle de $\mathcal{L}$ si et seulement si les trois conditions suivantes sont vérifiées :

(i) le problème primal $(\mathcal{P})$ admet comme solution le point x^* ;

(ii) le problème dual $(\mathcal{D})$ admet comme solution le point y^* ;

(iii) les problèmes primal et dual admettent la même valeur optimale, c'est-à-dire que

$$\min_{x \in \mathcal{X}} J(x) = \max_{y \in \mathcal{Y}} E(y)$$

On dit alors que x^* est une *solution primale* et y^* une *solution duale*. Les points $x \in \mathcal{X}$ (resp. $y \in \mathcal{Y}$) sont parfois appelés *variables primales* (resp. *variables duales*). Dans la dualité min-max, la dualité forte se traduit donc par le fait que problèmes primal et dual ont même valeur optimale.

Le corollaire **??** est un résultat central en optimisation, car il implique que, si les problèmes primal et dual ont même valeur optimale, alors la résolution du problème primal est équivalent à la recherche d'un point-selle.

> Si la fonction de couplage $\mathcal{L}$ admet un point-selle, alors le problème dual peut être un problème de minimisation équivalent au problème $(\mathcal{P})$. Il arrive que, pour certains problèmes d'optimisation, le problème dual présente des propriétés plus intéressantes que le problème primal, le rendant plus facile à résoudre. On peut alors être amené à résoudre préférentiellement le problème dual au lieu du problème primal. On parle alors de *résolution par dualité*. Si le problème dual est un problème convexe, tous les résultats de cette partie du cours sont valables, de même que les méthodes de résolution présentées.

Pour aller plus loin

La question centrale dans la dualité min-max réside donc dans le choix de la fonction de couplage. Dans les deux modules qui suivent, on va considérer deux choix différents, qui seront adaptés à deux classes de problèmes d'optimisation différentes.

Dualité de Lagrange. Dans le module A7, on introduit une fonction de couplage appelée *fonction lagrangienne* ou simplement *lagrangien*, qui permet de résoudre par dualité min-max les problèmes d'optimisation convexe lisse sous contraintes lisses. Le résultat central traduisant le lien entre points-selles du lagrangien est connu sous le nom de conditions de Karush, Kuhn et Tucker (ou conditions KKT).

Dualité de Fenchel. Pour les problèmes convexes non différentiables ou sous contraintes non différentiables, on peut considérer d'autres fonctions de couplage, construite à partir de la conjuguée convexe dite de Fenchel. On étudiera en détails les propriétés de cette conjuguée dans le module A8.

Méthodes d'éclatements primaux. La dualité min-max intervient dans de nombreux algorithmes en optimisation convexe. Elle permet de transformer un problème en un problème équivalent (module B4) ou d'interpréter les itérations d'un algorithme sur le problème primal en des itérations d'un autre algorithme sur un problème dual associé (module B6).

Méthodes d'éclatements primaux-duaux. La recherche d'un point-selle est au cœur de certains algorithmes, connus sous le nom de *méthodes primales-duales* (module B6).