Module B4
Optimisation sous contraintes

Dans ce module, sauf mention contraire, $\mathcal{X}$ et $\mathcal{Y}$ désignent des espaces de Hilbert, munis chacun d'un produit scalaire noté $\langle \cdot, \cdot \rangle$ et de norme associée notée $\| \cdot \|$, tandis que E désigne un espace euclidien, que l'on identifiera à $\mathbb{R}^n$ avec $n \in \mathbb{N}^*$, muni du produit scalaire usuel noté $\langle \cdot, \cdot \rangle$ également et de norme associée la norme euclidienne, notée $\| \cdot \|_2$.

L'objectif de ce module est d'introduire les notions utiles à l'étude des problèmes d'optimisation sous contraintes. On commencera par présenter le vocabulaire de l'optimisation sous contraintes, puis on s'intéressera aux résultats d'existence de solutions. Enfin, dans le cas de l'optimisation différentiable sur un **ouvert**, on énoncera les conditions d'optimalité du premier et du second ordre.

1 Solutions d'un problème d'optimisation sous contraintes

1.1 Vocabulaire de l'optimisation sous contraintes

D'après le Wiktionnaire, l'optimisation se définit comme suit :

> **Optimiser :** [...] obtenir le meilleur, selon un ensemble de critères, d'une chose ou d'une situation.

Dans le module **B1 : Minimisation d'une fonction. Conditions d'optimalité.**, on a vu l'exemple de la minimisation (et de la maximisation) d'une fonction, c'est-à-dire la recherche des éventuels minimiseurs (ou maximiseurs) d'une fonction définie sur l'espace $\mathcal{X}$.

Considérons par exemple le cas de l'optimisation d'un trajet entre le domicile d'une personne et son lieu de travail. Plusieurs interprétations de ce problème sont possibles :

- la personne souhaite connaître la distance séparant son domicile de son lieu de travail ;
- elle cherche l'itinéraire en voiture le plus court en terme de distance ;
- elle cherche l'itinéraire en voiture le moins coûteux en terme de prix du carburant ;
- elle cherche l'itinéraire à pied qui la fait passer le plus de temps dans un espace vert.

Dans ces quatre exemples, la personne recherche toujours un *chemin* (itinéraire ou courbe reliant son domicile et son lieu de travail), auquel est associé respectivement une distance (en mètres par exemple), un prix (en euros par exemple) ou une durée (en minutes par exemple). Le chemin peut être vu comme une portion de courbe ou une ligne brisée ; il s'agit donc d'un ensemble du plan $\mathbb{R}^2$. La distance, le coût ou la durée associée est une *quantité*, donc un nombre réel. On remarque aussi que les trois premiers exemples correspondent à des problèmes de minimisation, tandis que le dernier est un problème de maximisation.

Enfin, notons que, dans le premier problème, on considère tous les itinéraires possibles, c'est-à-dire tous les chemins reliant le domicile et le lieu de travail (donc *a priori*, même ceux qui passent à travers un immeuble !) : le problème est dit *non contraint*. En revanche, dans les deuxième et troisième problème, il est explicitement demandé de ne considérer que des itinéraires carrossables (et *a priori* légaux : interdiction de rouler à contre-sens ou sur un trottoir !), tandis que le dernier problème compare la durée des trajets réalisés à pied uniquement. Certains itinéraires sont donc écartés de la recherche : on dit que le problème est sous contraintes.

Définition 1 (Solution, valeur optimale)

Soit $\mathcal{C} \subset \mathcal{X}$ et $f : \mathcal{C} \to \mathbb{R}$ une fonction. On considère le problème de minimisation sous contraintes

$$\text{Minimiser } f(x) \text{ sous les contraintes } x \in \mathcal{C} \quad \text{ou} \quad \text{Minimiser } f \text{ sur } \mathcal{C} \quad (\mathcal{P}_c)$$

On dit que $x^* \in \mathcal{C}$ est une *solution (optimale)* du problème $(\mathcal{P}_c)$ si

$$\forall\, x \in \mathcal{C}, \quad f(x^*) \leq f(x)$$

Dans ce cas, la valeur $f(x^*) \in \mathbb{R}$ est appelée *valeur optimale* du problème $(\mathcal{P}_c)$.

La fonction f qui apparaît dans la définition précédente est appelée *fonction objectif* ou *critère* du problème $(\mathcal{P}_c)$. L'ensemble $\mathcal{C}$ est l'ensemble *admissible* ou *réalisable*. Tout point qui vérifie les contraintes est appelé *point admissible* ou *réalisable*. Enfin, on dit qu'un problème est réalisable si son ensemble réalisable est non vide.

Lorsque $\mathcal{C} = \mathcal{X}$, on dit que le problème $(\mathcal{P}_c)$ est *non contraint*, et ses solutions sont les minimiseurs de la fonction objectif f.

> S'il existe, on dit aussi que x^* est un minimiseur de f sur $\mathcal{C}$ et que $f(x^*)$ est le minimum de f sur $\mathcal{C}$. Cependant, dans ce cours, afin de souligner la différence fondamentale entre un problème d'optimisation non contrainte et un problème d'optimisation sous contraintes, on préférera réserver les termes minimiseurs / minimum aux problèmes non contraints.

L'introduction de contraintes peut changer les propriétés des solutions optimales (leur existence et leur nombre), ainsi que la valeur optimale, si elle existe. Mais ce n'est pas toujours le cas.

Exemple

On considère le problème sous contraintes suivant

$$\text{Minimiser } x^2 \text{ sous les contraintes } x \in \mathcal{C}$$

- Si $\mathcal{C} = \mathbb{R}$, alors ce problème admet une unique solution $x^* = 0$ et la valeur optimale atteinte vaut 0.
- Si $\mathcal{C} = [\,0\,;+\infty\,[$, alors ce problème admet une unique solution $x^* = 0$ et la valeur optimale atteinte vaut 0.
- Si $\mathcal{C} = \,]\,0\,;+\infty\,[$, alors ce problème n'admet aucune solution.
- Si $\mathcal{C} = [\,1\,;+\infty\,[$, alors ce problème admet une unique solution $x^* = 1$ et la valeur optimale atteinte vaut 1.
- Si $\mathcal{C} = \,]-\infty\,;1\,] \cup [\,1\,;+\infty\,[$, alors ce problème admet deux solutions $x^* = 1$ et $x^* = -1$ et la valeur optimale atteinte vaut 1.

On introduit également la définition d'un problème convexe :

Définition 2 (Problème convexe)

Soit $\mathcal{C} \subset \mathcal{X}$ et $f : \mathcal{C} \to \mathbb{R}$ une fonction. On considère le problème de minimisation sous contraintes

$$\text{Minimiser } f(x) \text{ sous les contraintes } x \in \mathcal{C} \quad \text{ou} \quad \text{Minimiser } f \text{ sur } \mathcal{C} \qquad (\mathcal{P}_c)$$

On dit que le problème $(\mathcal{P}_c)$ est *convexe* si la fonction objectif est convexe et que l'ensemble admissible $\mathcal{C}$ est convexe.

> Pour la définition et les propriétés des ensembles convexes, on pourra se reporter au module **B6 : Projection sur un convexe**.

REMARQUE : Il est nécessaire de vérifier que la fonction objectif **et** l'ensemble admissible sont bien tous les deux convexes.

On peut définir de manière analogue un problème d'optimisation strictement convexe (où la fonction objectif est supposée strictement convexe). Les problèmes d'optimisation strictement convexe admettent au plus une solution :

Proposition 1

Soit $\mathcal{C} \subset \mathcal{X}$ un ensemble **convexe** et $f : \mathcal{C} \to \mathbb{R}$ une fonction **strictement convexe**. On considère le problème de minimisation sous contraintes

$$\text{Minimiser } f(x) \text{ sous les contraintes } x \in \mathcal{C} \qquad (\mathcal{P}_c)$$

Alors le problème $(\mathcal{P}_c)$ admet au plus une solution.

REMARQUE : Cette proposition est valable que l'ensemble admissible soit fermé ou non.

DÉMONSTRATION : Supposons qu'il existe deux solutions $x_1 \neq x_2$ de $(\mathcal{P}_c)$. Par définition, x_1 et x_2 sont admissibles. Puisque l'ensemble admissible est convexe, le point

$$\frac{x_1 + x_2}{2} = \frac{1}{2}\, x_1 + \left(1 - \frac{1}{2}\right) x_2$$

est également admissible. La fonction f étant strictement convexe, on a

$$f\left(\frac{x_1 + x_2}{2}\right) < \frac{f(x_1) + f(x_2)}{2} = f(x_1)$$

ce qui est absurde car x_1 est une solution du problème $(\mathcal{P}_c)$. ■

La convexité de l'ensemble admissible est essentielle (surtout s'il n'est pas ouvert, comme on le verra dans la section suivante) :

CONTRE-EXEMPLE

Minimisation d'une fonction strictement convexe sur un ensemble non convexe. Soit $\mathcal{C} = \,]-\infty\,;1\,] \cup [\,1\,;+\infty\,[$ et $f : t \mapsto t^2$ la fonction carrée. La fonction f est strictement convexe (car $f'' = 2 > 0$) tandis que $\mathcal{C}$ est un ensemble non convexe car $0 = (-1+1)/2 \notin \mathcal{C}$ alors que $-1, 1 \in \mathcal{C}$. Considérons le problème

$$\text{Minimiser } f \text{ sur } \mathcal{C} \qquad (\mathcal{P}_c)$$

Puisque la fonction carrée est strictement croissante sur $[\,0\,;+\infty\,[$, on a

$$\forall\, t \in [\,1\,;+\infty\,[\,, \qquad f(1) \leq f(t)$$

Par ailleurs, comme la fonction carrée est strictement décroissante sur $]-\infty\,;0\,]$, on a

$$\forall\, t \in \,]-\infty\,;-1\,]\,, \qquad f(-1) \leq f(t)$$

Or, $f(-1) = f(1)$. Ainsi, le problème $(\mathcal{P}_c)$ admet deux solutions -1 et 1.

On termine enfin avec la définition suivante, qui est la généralisation de la notion de suite minimisante pour une fonction :

Définition 3 (Suite minimisante)

Soit $\mathcal{C} \subset \mathcal{X}$ et $f : \mathcal{C} \to \mathbb{R}$ une fonction. On considère le problème de minimisation sous contraintes

$$\text{Minimiser } f \text{ sur } \mathcal{C} \qquad (\mathcal{P}_c)$$

Soit $(x_k)_{k\in\mathbb{N}}$ une suite de $\mathcal{C}$. On dit que $(x_k)_{k\in\mathbb{N}}$ est une *suite minimisante* pour le problème $(\mathcal{P}_c)$ si

$$\lim_{k\to+\infty} f(x_k) = \inf_{\mathcal{C}} f$$

1.2 Contraintes d'égalité, contraintes d'inégalité

Un ensemble admissible peut être un région de l'espace géométrique facile à décrire (une boule, une sphère, un demi-plan...). Il est cependant plus courant de le définir à l'aide de deux types de contraintes, à savoir les *contraintes d'égalité* :

$$\Big\{x \in \mathcal{D} \mid g(x) = a\Big\} = g^{-1}(\{a\})$$

où $g : \mathcal{D} \to \mathbb{R}$ avec $\mathcal{D} \subset \mathcal{X}$ et $a \in \mathbb{R}$, et les *contraintes d'inégalité* :

$$\Big\{x \in \mathcal{D} \mid h(x) \leq b\Big\} = h^{-1}(]-\infty\,;b\,])$$

ou

$$\Big\{x \in \mathcal{D} \mid h(x) \geq b\Big\} = h^{-1}([\,b\,;+\infty\,[)$$

avec $h : \mathcal{D} \to \mathbb{R}$ avec $\mathcal{D} \subset \mathcal{X}$ et $b \in \mathbb{R}$. Notons que, si l'on pose $\tilde{h} = -h$, alors la dernière contrainte d'inégalité s'écrit

ou

$$\Big\{x \in \mathcal{D} \mid \tilde{h}(x) \leq -b\Big\} = \tilde{h}^{-1}(]-\infty\,;-b\,])$$

si bien que les problèmes d'optimisation sous contraintes d'égalité et d'inégalité sont de la forme

$$\text{Minimiser } f(x) \text{ sous les contraintes } \begin{cases} x \in g_i^{-1}(\{a_i\}) & \text{pour } i \in [\![\,1\,;\,p\,]\!] \\ x \in h_j^{-1}(]-\infty\,;b_j\,]) & \text{pour } j \in [\![\,1\,;\,q\,]\!] \end{cases}$$

Pour alléger les notations, on écrira plutôt

$$\text{Minimiser } f(x) \text{ sous les contraintes } \begin{cases} g_i(x) = a_i & \text{pour } i \in [\![1\,;\,p]\!] \\ h_j(x) \leq b_j & \text{pour } j \in [\![1\,;\,q]\!] \end{cases}$$

En remplaçant $g_i(x)$ par $g_i(x) - a_i$, on peut toujours remplacer une contrainte d'égalité en contrainte de nullité ; en remplaçant $h_j(x)$ par $h_j(x) - b_j$, on peut toujours remplacer une contrainte d'inégalité en contrainte de négativité ; on dit que les contraintes sont sous forme normale :

Définition 4 (Forme normale des contraintes d'égalité et d'inégalité)

Soit $\mathcal{C} \subset \mathcal{X}$. Soit $f : \mathcal{C} \to \mathbb{R}$, $g_i : \mathcal{C} \to \mathbb{R}$, avec $i \in [\![1\,;\,p]\!]$ et $h_j : \mathcal{C} \to \mathbb{R}$, avec $j \in [\![1\,;\,q]\!]$ des fonctions. On considère le problème d'optimisation sous contraintes suivant

$$\text{Minimiser } f(x) \text{ sous les contraintes } \begin{cases} g_i(x) = 0 & \text{pour } i \in [\![1\,;\,p]\!] \\ h_j(x) \leq 0 & \text{pour } j \in [\![1\,;\,q]\!] \end{cases} \qquad (\mathcal{P}_c)$$

On dit que les contraintes du problème $(\mathcal{P}_c)$ sont écrites *sous forme normale*.

> Dans le module **B5 : Théorème de Karush–Kuhn–Tucker**, les résultats sont démontrés pour des problèmes dont les contraintes sont écrites sous forme normale. La forme normale d'une contrainte d'inégalité n'étant pas toujours standard suivant les auteurs, il est essentiel de vérifier la convention choisi par l'ouvrage consulté si l'on souhaite appliquer correctement les résultats qui y figurent.

Une contrainte d'égalité peut toujours s'écrire à l'aide de deux contraintes d'inégalité. En effet, puisque

$$g(x) = 0 \qquad \Longleftrightarrow \qquad \big(g(x) \leq 0 \text{ et } -g(x) \leq 0\big)$$

en posant $h_1 = g$ et $h_2 = -g$, on a que

$$\Big\{x \in \mathcal{C} \mid g(x) = 0\Big\} = \Big\{x \in \mathcal{C} \mid h_1(x) \leq 0\Big\} \cap \Big\{x \in \mathcal{C} \mid h_2(x) \leq 0\Big\}$$

> Autrement dit, quitte à doubler le nombre de contraintes, on peut toujours remplacer des contraintes d'égalités en contraintes d'inégalités. Cette propriété sera utilisée pour donner des énoncés généralisés sans alourdir les notations, en ne distinguant pas les contraintes d'égalité des contraintes d'inégalité.

On va maintenant voir que la réciproque est vraie, à savoir qu'il est toujours possible, moyennant l'ajout d'une variable, de transformer une contrainte d'inégalité en contrainte d'égalité. On considère la contrainte d'inégalité suivante :

$$\Big\{x \in \mathcal{C} \mid h(x) \leq 0\Big\}$$

Soit x un point admissible pour cette contrainte (c'est-à-dire que $h(x) \leq 0$). Il existe donc $\varepsilon \in \mathbb{R}$ tel que

$$h(x) + \varepsilon^2 = 0$$

(il suffit de choisir $\varepsilon = \sqrt{-h(x)}$).

> La variable ε est appelée *variable d'écart*, ou encore *variable molle* (*slack variable* en anglais).

On peut donc réécrire la contrainte d'inégalité sous la forme

$$\left\{x \in \mathcal{C} \mid \exists \varepsilon \in \mathbb{R}, \quad \tilde{g}(x, \varepsilon) = h(x) + \varepsilon^2 = 0\right\}$$

Il ne s'agit pas encore d'une contrainte d'égalité à proprement parler. Pour en obtenir une, il faut maintenant s'intéresser au problème d'optimisation associé. Pour simplifier, on suppose qu'il n'y a qu'une seule contrainte d'inégalité :

$$\text{Minimiser } f(x) \text{ sous la contrainte } h(x) \leq 0 \tag{$\mathcal{P}_{ci}$}$$

Supposons que ce problème admet une solution $x^* \in \mathcal{C}$. Alors, par définition, $h(x^*) \leq 0$ et

$$\forall\, x \in \mathcal{C}, \qquad h(x) \leq 0 \quad \Longrightarrow \quad f(x^*) \leq f(x)$$

On pose $\varepsilon^* = \sqrt{-h(x^*)}$ et $\tilde{f}(x, \varepsilon) = f(x)$ pour tout $x \in \mathcal{C}$. On a par définition de ε^*

$$\tilde{g}(x^*, \varepsilon^*) = 0$$

Soit $(x, \varepsilon) \in \mathcal{C} \times \mathbb{R}$ tel que $\tilde{g}(x, \varepsilon) = 0$. Alors on a $h(x) \leq 0$; par conséquent,

$$f(x^*) = \tilde{f}(x^*, \varepsilon^*) \leq \tilde{f}(x, \varepsilon) = f(x)$$

On en déduit donc que (x^*, ε^*) est solution du problème sous contrainte d'égalité

$$\text{Minimiser } \tilde{f}(x, \varepsilon) \text{ sous la contrainte } \tilde{g}(x, \varepsilon) = 0 \tag{$\mathcal{P}_{ce}$}$$

Réciproquement, si (x^*, ε^*) est solution du problème $(\mathcal{P}_{ce})$, alors $\tilde{g}(x^*, \varepsilon^*) = 0$, ce qui implique que $h(x^*) \leq 0$; ainsi, x^* est admissible pour le problème $(\mathcal{P}_{ci})$. Par ailleurs, pour tout $x \in \mathcal{C}$ tel que $h(x) \leq 0$, on a

$$f(x^*) = \tilde{f}(x^*, \varepsilon^*) \leq \tilde{f}(x, \sqrt{-h(x)}) = f(x)$$

Par conséquent, x^* est solution du problème $(\mathcal{P}_{ci})$.

> Ainsi, un problème sous contraintes d'égalité et d'inégalité peut toujours s'écrire de manière équivalente sous la forme d'un problème sous contraintes d'inégalité uniquement ou d'égalité uniquement. Il est parfois utile d'utiliser l'une ou l'autre écriture, soit parce qu'un résultat se démontre plus simplement dans l'un ou l'autre cas, soit pour alléger les notations (en supprimant les contraintes d'égalité par exemple).

1.3 Problèmes équivalents

Dans le paragraphe précédent, on a vu que changer la manière dont on décrit l'ensemble admissible ne change pas les solutions ni la valeur optimale d'un problème. On a également vu qu'il est possible de transformer un problème avec des contraintes d'inégalité en un problème avec des contraintes d'égalité, tout en explicitant le lien entre les solutions des deux problèmes. Aussi, la résolution de l'un permet la résolution de l'autre. Il s'agit d'un cas particulier d'équivalence entre deux problèmes :

Définition 5 (Équivalence entre problèmes d'optimisation)

Soit $\mathcal{C} \subset \mathcal{X}$ et $\mathcal{D} \subset \mathcal{Y}$. Soit $f : \mathcal{C} \to \mathbb{R}$ et $g : \mathcal{D} \to \mathbb{R}$ deux fonctions. On dit que les problèmes de minimisation

$$\text{Minimiser } f \text{ sur } \mathcal{C} \qquad (\mathcal{P}_1)$$

et

$$\text{Minimiser } g \text{ sur } \mathcal{D} \qquad (\mathcal{P}_2)$$

sont *équivalents* s'il existe une bijection **connue** $M : \mathcal{C} \to \mathcal{D}$ telle que x^* est solution de $(\mathcal{P}_1)$ si et seulement si $M(x^*)$ est solution de $(\mathcal{P}_2)$.

> On dit également que les deux ensembles de solutions sont *équipotents* (et *isomorphes* si la bijection est linéaire).

Une conséquence immédiate de cette définition est le fait que si deux problèmes sont équivalents, alors ils admettent le même nombre de solutions. En particulier, si l'un n'admet pas de solution, alors l'autre non plus.

EXERCICE

Montrer que la notion définie dans la définition 5 est bien une relation d'équivalence, c'est-à-dire qu'elle est réflexive, symétrique et transitive.

On démontre un résultat analogue à la proposition 4 du module **B1 : Minimisation d'une fonction. Conditions d'optimalité.** :

Proposition 2

Soit $f : \mathcal{C} \to \mathbb{R}$ une fonction, $x^0 \in \mathcal{X}$ et $g : f(\mathcal{C}) \to \mathbb{R}$ une fonction strictement croissante. Alors les deux problèmes suivants sont équivalents :

$$\text{Minimiser } f(x) \text{ sous les contraintes } x \in \mathcal{C} \qquad (\mathcal{P}_1)$$

$$\text{Minimiser } g \circ f(x - x^0) \text{ sous les contraintes } x \in x^0 + \mathcal{C} \qquad (\mathcal{P}_2)$$

Attention, la stricte croissance de g est essentielle.

DÉMONSTRATION : On note $J : x \mapsto g \circ f(x - x^0)$.

- **Cas où** $(\mathcal{P}_1)$ **admet au moins une solution.** Soit x^* une solution du problème $(\mathcal{P}_1)$. Par définition, on a

  $$\forall\, x \in \mathcal{C}, \qquad f(x^*) \leq f(x)$$

  Cette inégalité peut se réécrire

  $$\forall\, x \in \mathcal{C}, \qquad f\big((x^* + x^0) - x^0\big) \leq f\big((x + x^0) - x^0\big)$$

  La stricte convexité de g assure alors que

  $$\forall\, x \in \mathcal{C}, \qquad g \circ f\big((x^* + x^0) - x^0\big) \leq g \circ f\big((x + x^0) - x^0\big)$$

  soit $$\forall\, x \in \mathcal{C}, \qquad J(x^* + x^0) \leq J(x + x^0)$$

Or, pour tout $y \in x^0 + \mathcal{C}$, il existe $x \in \mathcal{C}$ tel que $y = x + x^0$. On en déduit finalement que l'inégalité précédente s'écrit également

$$\forall\, y \in x^0 + \mathcal{C}, \qquad J(x^* + x^0) \leq J(y)$$

Autrement dit : si x^* est solution de $(\mathcal{P}_1)$, alors $M(x^*) = x^* + x^0$ est solution de $(\mathcal{P}_2)$. En particulier, en remarquant que g est inversible car strictement croissante, d'inverse g^{-1} vérifiant

$$\forall\, x, x' \in \mathcal{C}, \qquad g^{-1}(x) < g^{-1}(x') \quad \Longleftrightarrow \quad x = g \circ g^{-1}(x) < \circ g^{-1} g(x') = x'$$

car g est strictement croissante, on en déduit que g^{-1} est strictement croissante également. Ainsi, puisque

$$\forall\, x \in \mathcal{C}, \qquad f(x) = g^{-1} \circ J(x + x^0) = g^{-1} \circ J(x - (-x^0))$$

le même raisonnement que celui qui précède montre que, si y^* est solution de $(\mathcal{P}_2)$, alors $\tilde{M}(y^*) = y^* - x^0 = M^{-1}(y^*)$ est solution de $(\mathcal{P}_1)$. Ainsi, M définit une bijection entre les ensembles de solution des problèmes $(\mathcal{P}_1)$ et $(\mathcal{P}_2)$.

- **Cas où $(\mathcal{P}_1)$ n'admet pas de solution.** D'après ce qui précède, le problème

$$\text{Minimiser } J \text{ sur } x^0 + \mathcal{C}$$

n'admet pas de solution non plus, car sinon, d'après le point précédent appliqué au problème $(\mathcal{P}_1)$, le problème $(\mathcal{P}_1)$ admet une solution. Donc dans ce cas, les deux ensembles de solutions sont vides et donc équipotents. ■

La notion de problèmes équivalents est très utile en pratique, car elle permet d'introduire des problèmes auxiliaires, potentiellement plus simples à résoudre ou à étudier, pour résoudre le problème initialement considéré. S'il s'agit seulement de déterminer des propriétés théoriques (comme le nombre de solutions), alors la connaissance explicite de la bijection n'est pas toujours nécessaire. En revanche, si l'on souhaite résoudre le problème initial en déterminant les solutions du problème auxiliaire, alors cette connaissance est indispensable pour pouvoir obtenir les solutions recherchées.
Dans certaines applications, il arrive que l'on ne s'intéresse qu'à une seule solution d'un problème. Aussi, il n'est pas nécessaire d'exhiber un problème auxiliaire équivalent, mais il est suffisant de trouver un problème auxiliaire dont une solution donnera (à l'application d'une fonction près) une solution du problème initial.

2 Optimisation sur un ouvert

Dans cette section, on considère les problèmes d'optimisation suivant :

$$\text{Minimiser } f \text{ sur } \mathcal{U} \tag{$\mathcal{P}$}$$

où $\mathcal{U} \subset \mathcal{X}$ est un **ouvert** et $f : \mathcal{U} \to \mathbb{R}$ une fonction différentiable (resp. deux fois différentiable). On va énoncer des conditions d'optimalité du premier et du second ordre qui découlent directement des preuves du module **B1 : Minimisation d'une fonction. Conditions d'optimalité.**

2.1 Condition nécessaire d'optimalité du premier ordre

Commençons par énoncer l'équivalent de la règle de Fermat pour un problème d'optimisation d'une fonction objectif différentiable sur un ensemble admissible **ouvert** :

Proposition 3

Soit $\mathcal{U} \subset \mathcal{X}$ un **ouvert** et $f : \mathcal{U} \to \mathbb{R}$ une fonction différentiable. Soit $x^* \in \mathcal{U}$ une solution du problème d'optimisation sous contraintes

$$\text{Minimiser } f \text{ sur } \mathcal{U}$$

Alors
$$\nabla f(x^*) = 0$$

DÉMONSTRATION : Puisque la démonstration de la proposition 12 du module **B1 : Minimisation d'une fonction. Conditions d'optimalité.** n'utilise que des arguments locaux (au voisinage de x^*), elle reste valable si on restreint f à l'ouvert $\mathcal{U}$. ■

REMARQUE : **Attention !** Il est absolument nécessaire de supposer que l'ensemble admissible est **ouvert**, comme le montre le contre-exemple suivant :

CONTRE-EXEMPLE

Cas d'un ensemble admissible non ouvert. On considère le problème d'optimisation sous contraintes suivant :

$$\text{Minimiser } f(t) = t^3 \text{ sous les contraintes } t \geq 1$$

Puisque $f : t \mapsto t^3$ est dérivable sur $\mathbb{R}$, on a

$$\forall\, t \in \mathbb{R}, \qquad f'(t) = 3\,t^2$$

Il s'ensuit que $f'(t) \geq 0$ pour tout $t \in \mathbb{R}$, avec $f'(t) = 0$ si et seulement si $t = 0$, c'est-à-dire que f admet comme unique point critique 0. On déduit par ailleurs que f est croissante. En particulier, il s'ensuit que

$$\forall\, t \in [\,1\,;+\infty\,[\,, \qquad f(1) \leq f(t)$$

Autrement dit, 1 est solution du problème considéré (et même l'unique solution). Or, $f'(1) = 3 \neq 0$. Cela est dû au fait que l'ensemble admissible n'est pas ouvert. En revanche, si on considère le problème suivant

$$\text{Minimiser } f(t) = t^3 \text{ sous les contraintes } t > 1$$

où l'ensemble admissible $]\,1\,;+\infty\,[$ est cette fois ouvert, on voit que la dérivée de f ne s'y annule toujours pas, mais que le problème n'admet pas de solution non plus (si $t^* > 1$ était solution, alors par stricte croissance de f sur $]\,1\,;+\infty\,[$, on aurait $f(t^*/2) < f(t^*)$). Ce résultat est cohérent avec la proposition 6 qui stipule que toute solution du **second** problème annule nécessairement la dérivée.

Si f n'est pas différentiable, mais GATEAUX-différentiable, alors en particulier, on peut démontrer (en reprenant cette fois la preuve de la proposition 13 du module **B1 : Minimisation d'une fonction. Conditions d'optimalité.**, qui n'utilise que des arguments locaux) que, si $x^* \in \mathcal{U}$ est une solution du problème $(\mathcal{P})$, alors

$$\forall\, v \in \mathcal{X}, \qquad f'(x^*; v) = 0$$

ou encore que
$$\forall\, x \in \mathcal{X}, \qquad f'(x^*; x - x^*) = 0$$

En particulier, puisque la dérivée partielle est homogène en son second argument, cette dernière égalité est équivalente à

$$\forall\, x \in \mathcal{X}, \qquad f'(x^*; x - x^*) \geq 0$$

Montrons qu'il est possible de ne considérer dans l'inégalité précédente les points $x \in \mathcal{U}$.

Proposition 4 (Inéquation variationnelle)

Soit $\mathcal{U} \subset \mathcal{X}$ un **ouvert** et $f : \mathcal{U} \to \mathbb{R}$ une fonction Gateaux-différentiable. Soit $x^* \in \mathcal{U}$ une solution du problème d'optimisation sous contraintes

$$\text{Minimiser } f \text{ sur } \mathcal{U}$$

Alors

$$\forall\, x \in \mathcal{U}, \qquad f'(x^*; x - x^*) \geq 0 \tag{IV}$$

L'inégalité dans la proposition 4 s'appelle une *inéquation variationnelle.*

Démonstration : Soit $x \in \mathcal{U}$. Puisque $\mathcal{U}$ est ouvert, on a $x^* + h\,(x - x^*) \in \mathcal{U}$ pour h assez petit et donc

$$f(x^*) \leq f(x^* + h\,(x - x^*)) = f(x^*) + h\, f'(x^*; x - x^*) + o(h)$$

Si $h > 0$, alors on obtient après simplification

$$0 \leq f'(x^*; x - x^*) + \varepsilon(h)$$

avec $\varepsilon(h) \to 0$ lorsque $h \to 0$, si bien qu'en passant à la limite lorsque $h \to 0$, on obtient finalement

$$0 \leq f'(x^*; x - x^*) \;\blacksquare$$

> Le résultat de la proposition 4 semble moins intéressant que le fait que $f'(x^*; v)$ s'annule pour tout $v \in \mathcal{X}$. Cependant, on verra que, dans le cas où $\mathcal{U}$ n'est plus un ouvert, la nullité de la différentielle (de Gateaux aussi bien que de Fréchet) n'est plus assurée en une solution du problème $(\mathcal{P})$, mais que l'inéquation variationnelle reste vérifiée.

Si la fonction objectif f est strictement convexe, on peut montrer qu'elle admet au plus un point critique. Ainsi, on a le résultat suivant :

Proposition 5

Soit $\mathcal{U} \subset \mathcal{X}$ un ensemble **ouvert** et $f : \mathcal{X} \to \mathbb{R}$ une fonction **strictement convexe**. On considère le problème de minimisation sous contraintes

$$\text{Minimiser } f(x) \text{ sous les contraintes } x \in \mathcal{U} \tag{$\mathcal{P}_c$}$$

Alors le problème $(\mathcal{P}_c)$ admet au plus une solution.

Démonstration : Si x^* est solution de $(\mathcal{P}_c)$, il est nécessairement point critique de la restriction de f sur $\mathcal{U}$ (proposition 6). Or, f admet au plus un minimiseur, donc au plus un point critique, et en particulier au plus un point critique dans $\mathcal{U}$. Par conséquent, le problème $(\mathcal{P}_c)$ admet au plus une solution. $\blacksquare$

REMARQUE : Ce résultat ne nécessite pas la convexité de l'ensemble admissible $\mathcal{U}$ lorsque $\mathcal{U}$ est **ouvert**, mais f doit être définie sur $\mathcal{X}$ (en réalité, il suffit que f soit définit sur un convexe ouvert contenant $\mathcal{U}$).

2.2 Optimisation différentiable convexe

La condition nécessaire et suffisante d'optimalité du premier ordre (règle de FERMAT) pour la minimisation d'une fonction convexe se généralise aux cas de la minimisation d'une fonction convexe sur un ouvert.

Proposition 6

Soit $\mathcal{U} \subset \mathcal{X}$ un **ouvert** et $f : \mathcal{X} \to \mathbb{R}$ une fonction différentiable **convexe**. On a équivalence entre les deux énoncés suivants :

(i) $x^* \in \mathcal{U}$ une solution du problème d'optimisation sous contraintes

$$\text{Minimiser } f \text{ sur } \mathcal{U}$$

(ii) $x^* \in \mathcal{U}$ est un point critique de f, c'est-à-dire

$$\nabla f(x^*) = 0$$

DÉMONSTRATION : Puisque le sens direct a déjà démontré dans la proposition 6, on se contente de démontrer la réciproque. On suppose donc que $x^* \in \mathcal{U}$ et que $\nabla f(x^*) = 0$. Par convexité de f, on a

$$\forall\, x \in \mathcal{X}, \qquad f(x) \geq f(x^*) + \langle \nabla f(x^*), x - x^* \rangle = f(x^*)$$

Cette minoration reste valable si l'on se restreint aux points $x \in \mathcal{U}$. On en déduit que x^* est une solution de $(\mathcal{P}_c)$. ■

Si f est GATEAUX-différentiable seulement, on a un résultat analogue : x^* est une solution du problème $(\mathcal{P})$ si et seulement si

$$\forall\, v \in \mathcal{X}, \qquad f'(x^*; v) = 0$$

Par ailleurs, on peut démontrer que la réciproque dans la proposition 4 est vraie :

Proposition 7 (Inéquation variationnelle dans le cas convexe)

Soit $\mathcal{U} \subset \mathcal{X}$ un **ouvert** et $f : \mathcal{X} \to \mathbb{R}$ une fonction **convexe** GATEAUX-différentiable. On a équivalence entre les deux énoncés suivants :

(i) $x^* \in \mathcal{U}$ une solution du problème d'optimisation sous contraintes

$$\text{Minimiser } f \text{ sur } \mathcal{U}$$

(ii) x^* est une solution du problème d'inéquation variationnelle suivant :

$$\text{Trouver } x^* \in \mathcal{U} \quad \text{tel que} \quad \forall\, x \in \mathcal{U}, \qquad f'(x^*; x - x^*) \geq 0 \quad (\mathcal{P}_{\mathrm{IV}})$$

DÉMONSTRATION : La réciproque est à nouveau une conséquence de la caractérisation d'une fonction convexe par ses dérivées partielles :

$$f(x^* + t\,(x - x^*)) \geq f(x^*) + t\, f'(x^*, x - x^*)$$

appliquée au point $t = 1$. ■

> Il est possible de démontrer des résultats analogues à ceux énoncés dans ce paragraphe lorsque la fonction objectif f n'est plus définie sur l'espace $\mathcal{X}$, mais un ensemble **ouvert convexe**.

2.3 Conditions d'optimalité du second ordre

On suppose à présent que $\mathcal{X} = \mathbb{R}^n$ est un espace euclidien et que la fonction objectif f est deux fois différentiable. La condition nécessaire d'optimalité du second ordre est à nouveau également valable si l'ensemble admissible est **ouvert** :

Proposition 8 (Condition nécessaire d'optimalité du second ordre)

Soit $U \subset \mathbb{R}^n$ un **ouvert** et $f : U \to \mathbb{R}$ une fonction deux fois différentiable. Soit $x^* \in U$ une solution du problème d'optimisation sous contraintes

$$\text{Minimiser } f \text{ sur } U$$

Alors $\operatorname{Hess} f(x^*)$ est semi-définie positive.

Démonstration : Il suffit d'adapter la preuve de la proposition 16 du module **B1 : Minimisation d'une fonction. Conditions d'optimalité.**

3 Optimisation sur un fermé

Dans cette section, on considère les problèmes d'optimisation suivant :

$$\text{Minimiser } f \text{ sur } \mathcal{C} \qquad (\mathcal{P})$$

où $\mathcal{C} \subset \mathcal{X}$ est un **fermé** et $f : \mathcal{C} \to \mathbb{R}$ une fonction. On va établir des résultats d'existence et d'unicité de solutions, puis on va s'intéresser à une première caractérisation au premier ordre des solutions dans le cas convexe.

3.1 Résultats d'existence en dimension infinie

On commence par rappeler le théorème de Weierstrass, reformulé dans un contexte d'optimisation sous contraintes :

Théorème 1 (Weierstrass)

Soit $\mathcal{K} \subset \mathcal{X}$ un ensemble **compact** non vide et $f : \mathcal{K} \to \mathbb{R}$ une fonction **continue**. On considère le problème de minimisation sous contraintes

$$\text{Minimiser } f(x) \text{ sous les contraintes } x \in \mathcal{K} \qquad (\mathcal{P}_c)$$

Alors le problème $(\mathcal{P}_c)$ admet une solution.

On mentionne dès à présent un résultat très spécifique. Il s'agit d'un résultat qui sera démontré dans le module **B6 : Projection sur un convexe** :

Proposition 9 (Projection sur un convexe fermé non vide)

Soit $\mathcal{C} \subset \mathcal{X}$ un ensemble **convexe fermé** non vide et $x^0 \in \mathcal{X}$. On considère le problème de minimisation sous contraintes

$$\text{Minimiser } f(x) = \frac{1}{2} \|x - x^0\|^2 \text{ sous les contraintes } x \in \mathcal{C} \qquad (\mathcal{P}_c)$$

Alors le problème $(\mathcal{P}_c)$ admet une unique solution.

Démonstration : Démontrée dans le module **B6 : Projection sur un convexe**.

En dimension finie, le résultat d'existence de la proposition précédente est une conséquence de la proposition 10, puisque f est continue (car différentiable) et infinie à l'infini. L'unicité est quant à elle une conséquence de la stricte convexité u problème.

3.2 Résultat d'existence en dimension finie

L'équivalent de l'existence d'un minimiseur pour une fonction continue sur un espace euclidien et infinie à l'infini est donné dans la proposition suivante :

Proposition 10

Soit $\mathcal{C} \subset E$ un ensemble **fermé non vide** et $f : E \to \mathbb{R}$ une fonction continue et **infinie à l'infini**. On considère le problème de minimisation sous contraintes

$$\text{Minimiser } f(x) \text{ sous les contraintes } x \in \mathcal{C} \qquad (\mathcal{P}_c)$$

Alors le problème $(\mathcal{P}_c)$ admet une solution.

Remarque : En dimension finie, le théorème 1 et la proposition 10 sont des résultats très proches : ils font tous les deux l'hypothèse que l'ensemble admissible est fermé et non vide, tandis que la fonction objectif doit être continue. Dans la proposition 10, l'hypothèse de compacité de l'ensemble admissible (c'est-à-dire l'hypothèse selon laquelle il doit être borné) est remplacé par le caractère infini à l'infini de la fonction objectif. On notera que, dans la proposition 9, l'ensemble admissible est également fermé et non vide, et la fonction objectif continue et infinie à l'infini.

Démonstration : Puisque f est continue et infinie à l'infini, la preuve de la proposition 9 du module **B1 : Minimisation d'une fonction. Conditions d'optimalité.** assure qu'il existe $x_0 \in E$ tel que l'ensemble

$$\Big\{x \in E \mid f(x) \leq f(x_0)\Big\}$$

soit borné. Il s'ensuit que, par intersection, l'ensemble

$$K = \Big\{x \in \mathcal{C} \mid f(x) \leq f(x_0)\Big\} = \Big\{x \in E \mid f(x) \leq f(x_0)\Big\} \cap \mathcal{C}$$

est fermé et borné, donc compact (car E est de dimension finie). Par conséquent, puisque f est continue, il existe $x^* \in K$ tel que

$$\forall\, x \in K, \qquad f(x^*) \leq f(x)$$

Par ailleurs, si $x \in \mathcal{C}$ mais $x \notin K$, alors par définition de K,

$$f(x) > f(x_0) \geq f(x^*)$$

Finalement, on a montré que x^* est solution de $(\mathcal{P}_c)$. ■

Attention ! L'hypothèse de fermeture de l'ensemble admissible est très importante :

CONTRE-EXEMPLE

Minimisation d'une fonction infinie à l'infini sur un ensemble ouvert. Soit $\mathcal{C} = \left]\, 0\,; +\infty \,\right[$ et $f : t \mapsto t^2$ la fonction carrée. La fonction f est continue et infinie à l'infini tandis que $\mathcal{C}$ est un ensemble ouvert. Considérons le problème

$$\text{Minimiser } f \text{ sur } \mathcal{C} \tag{$\mathcal{P}_c$}$$

Si le problème $(\mathcal{P}_c)$ admet une solution t^*, alors t^* est admissible par définition : on a donc $t^* > 0$. Or, la fonction carrée est strictement croissante sur $\left[\, 0\,; +\infty \,\right[$, donc $f(t^*/2) < f(t^*)$ avec $t^*/2$ admissible : on aboutit à une contradiction. Il s'ensuit que le problème $(\mathcal{P}_c)$ n'admet aucune solution.

3.3 Inéquation variationnelle

On suppose maintenant que f est **convexe** et GATEAUX-différentiable sur $\mathcal{X}$. Il est essentiel de retenir que la règle de FERMAT (c'est-à-dire la condition nécessaire et suffisante d'optimalité du premier ordre) **n'est plus valable** lorsque l'ensemble admissible n'est pas ouvert.

CONTRE-EXEMPLE

Minimisation d'une fonction convexe sur un ensemble convexe fermé. Soit $\mathcal{C} = \left[\, 1\,; +\infty \,\right[$ et $f : t \mapsto t^2$ la fonction carrée. La fonction f est convexe (car $f'' = 2 \geq 0$) tandis que $\mathcal{C}$ est un ensemble convexe et **fermé** comme intervalle réel fermé. Considérons le problème

$$\text{Minimiser } f \text{ sur } \mathcal{C} \tag{$\mathcal{P}_c$}$$

Puisque la fonction carrée est strictement croissante sur $\left[\, 0\,; +\infty \,\right[$, on a

$$\forall\, t \in \left[\, 1\,; +\infty \,\right[, \qquad f(1) \leq f(t)$$

de sorte que le problème $(\mathcal{P}_c)$ admet comme (unique) solution 1. Or, on a

$$f'(1) = 2$$

donc 1 n'est pas un point critique de f sur $\left[\, 1\,; +\infty \,\right[$. On peut d'ailleurs vérifier que f n'admet aucun point critique sur cet ensemble.

En fait, soit la ou les solutions sont à l'intérieur (topologique) de l'ensemble admissible, et dans ce cas la différentielle de GATEAUX s'y annule ; soit la ou les solutions (s'ils existent) sont sur le bord, et dans ce cas la différentielle ne s'y annule pas forcément. Avec les outils introduits dans ce cours, il n'y a donc pas de généralisation possible de la règle de FERMAT (que ce soit dans le cas convexe ou non).

> Grâce à la théorie de la *sous-différentiabilité*, il est possible de généraliser la notion de dérivée aux fonctions non dérivables (et de même la notion de différentielle et de gradient aux fonctions non différentiables). Certains choix de sous-différentiels préservent la règle de FERMAT. En réécrivant les problèmes d'optimisation sous contraintes sous la forme d'un problème équivalent d'optimisation non contrainte et non différentiable, on peut alors écrire une règle de FERMAT pour de tels problèmes.

En revanche, on a vu que, si le problème est convexe et que l'ensemble admissible est **ouvert**, alors le problème de minimisation $(\mathcal{P}_c)$ est équivalent à résoudre le problème d'inéquation variationnelle $(\mathcal{P}_{\text{IV}})$. C'est également le cas lorsque l'ensemble admissible n'est pas ouvert :

Proposition 11 (Inéquation variationnelle dans le cas convexe)

Soit $\mathcal{C} \subset \mathcal{X}$ un **convexe** et $f : \mathcal{X} \to \mathbb{R}$ une fonction **convexe** GATEAUX-différentiable. On a équivalence entre les deux énoncés suivants :

(i) $x^* \in \mathcal{C}$ une solution du problème d'optimisation sous contraintes

$$\text{Minimiser } f \text{ sur } \mathcal{C}$$

(ii) x^* est une solution du problème d'inéquation variationnelle suivant :

$$\text{Trouver } x^* \in \mathcal{C} \quad \text{tel que} \quad \forall\, x \in \mathcal{C}, \qquad f'(x^*; x - x^*) \geq 0 \qquad (\mathcal{P}_{\text{IV}})$$

DÉMONSTRATION :

- **Sens direct.** Si x^* est solution de $(\mathcal{P}_c)$, alors

$$\forall\, x \in \mathcal{C}, \qquad f(x^*) \leq f(x)$$

Soit $x \in \mathcal{C}$. Puisque C est convexe, pour tout $t \in \,]\,0\,;1\,]$, on a $t\,x + (1-t)\,x^* \in \mathcal{C}$, de sorte que

$$f(x^*) \leq f(t\,x + (1-t)\,x^*)$$

Or, $t\,x + (1-t)\,x^* = x^* + t\,(x - x^*)$. Il s'ensuit que

$$\frac{f(x^* + t\,(x - x^*)) - f(x^*)}{t} \geq 0$$

En faisant tendre t vers 0, on en déduit que

$$\forall\, x \in \mathcal{C}, \qquad f'(x^*; x - x^*) \geq 0$$

- **Réciproque.** On suppose que

$$\forall\, x \in \mathcal{C}, \qquad f'(x^*; x - x^*) \geq 0$$

Puisque $x = x^* + 1 \times (x - x^*)$, on a

$$\forall\, x \in \mathcal{C}, \qquad f(x) = f(x^* + 1 \times (x - x^*)) \geq f(x^*) + 1 \times f'(x^*; x - x^*) \geq f(x^*)$$

Autrement dit, x^* est solution de $(\mathcal{P}_c)$. ■

Si f est différentiable, alors le problème $(\mathcal{P}_{\text{IV}})$ s'écrit

$$\text{Trouver } x^* \in \mathcal{C} \quad \text{tel que} \quad \forall\, x \in \mathcal{C}, \qquad \langle \nabla f(x^*), x - x^* \rangle \geq 0$$

Si f admet un point critique x^* dans $\mathcal{C}$, alors x^* vérifie cette inégalité et est donc solution du problème $(\mathcal{P}_c)$. La réciproque est fausse, comme le montre le contre-exemple suivant :

CONTRE-EXEMPLE

Minimisation d'une fonction convexe sur un ensemble convexe fermé. Soit $\mathcal{C} = [\,1\,;+\infty\,[$ et $f : t \mapsto t^2$ la fonction carrée. Considérons le problème

$$\text{Minimiser } f \text{ sur } \mathcal{C} \qquad (\mathcal{P}_c)$$

On a déjà vu que le problème $(\mathcal{P}_c)$ admet comme (unique) solution 1 et que $f'(1) = 2$. On peut vérifier qu'on a bien

$$\forall\, t \in [\,1\,;+\infty\,[\,, \qquad f'(1)(t-1) = t - 1 \geq 0$$

Notons en revanche que $f'(1)(t-1) \neq 0$ pour tout $t \geq 1$.