Cours n°o
Introduction à l'optimisation

Le premier cours introductif vise à présenter de manière intuitive et générale l'optimisation, ses applications et ses enjeux. Des notions plus théoriques ainsi que des questions de vocabulaire sont abordées dans la seconde partie de ce cours.

1 Qu'est-ce que l'optimisation ?

1.1 Une première définition

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.

Plus concrètement, on peut citer des exemples dans lesquels le terme *optimiser* intervient naturellement :

- optimiser une production
- optimiser des dépenses

On voit dans ces deux exemples que le même terme (*optimiser*) est utilisé pour désigner deux choses *a priori* opposées, puisque lorsque l'on cherche à *optimiser* une production, on sous-entend généralement que l'on souhaite la *maximiser*, tandis que des dépenses *optimales* désigne plus fréquemment des dépenses *minimales*. Néanmoins, on peut adopter dans un premier temps la définition suivante, plus précise : *optimiser, c'est maximiser ou minimiser*.

1.2 Pourquoi optimise-t-on ?

Avant de creuser plus avant la définition de l'optimisation, penchons-nous sur l'intérêt que l'on a à optimiser. Des raisons simples pour lesquelles on souhaite optimiser sont, par exemple, la volonté de réduire des coûts (qu'ils soient financiers, énergétiques...), de réduire le temps ou la durée d'une activité, ou bien encore augmenter les profits. Ces points sont importants, car ils vont permettre de clarifier le problème mathématique qui se cache derrière l'optimisation.

1.3 Qu'optimise-t-on ?

En effet, la première définition que nous avons proposée pour le verbe *optimiser* conduit très naturellement à la question suivante : maximiser ou minimiser quoi ? Nous allons voir que la réponse à cette question n'est pas toujours évidente.

Tout d'abord, il faut distinguer la *quantité* (une production, des dépenses, un temps) que l'on cherche à optimiser, de la *configuration* (objets produits, nature des dépenses, activités) qui permet d'atteindre cette quantité optimale. Ainsi, dans le langage courant, on pourra parler d'optimiser un trajet. Or, derrière cette expression, se cache plusieurs significations possibles : parle-t-on plutôt d'optimiser le temps de parcours ? la distance parcourue ? le coût ? l'empreinte carbone ? Il est évident que, dès lors qu'on parle de maximiser ou de minimiser *quelque chose*, il doit être possible de comparer ce *quelque* *chose* à *autre chose* en termes de *grandeur*, autrement dit, d'être capable de choisir une relation d'ordre. La quantité à optimiser doit donc naturellement être une quantité réelle (un nombre réel). Ainsi, pour revenir à l'exemple du trajet, il faudra, par exemple, parler d'optimiser un *temps* de trajet, ce qui donnera un temps de trajet optimal (c'est-à-dire, minimal) et un itinéraire optimal (c'est-à-dire, permettant d'atteindre le temps de trajet optimal). En termes mathématiques, cela revient (comme on le verra plus tard) à distinguer un minimiseur d'un minimum, ou un maximiseur d'un maximum.

Il est à noter que, suivant les contextes, l'une ou l'autre des informations est utile à l'utilisateur. Par exemple, si l'on souhaite arriver avant une certaine heure (pour un rendez-vous), il est plus utile de savoir combien de temps on peut mettre à atteindre un certain lieu (en supposant par exemple que l'itinéraire est donné par un GPS) ; dans d'autres cas, la durée du trajet importe peu, on souhaite surtout connaître le chemin à emprunter. On remarquera que connaître l'itinéraire optimal permet de connaître le temps de trajet optimal (il suffit de se chronométrer), alors que le temps de trajet ne donne aucune information sur l'itinéraire à suivre.

Enfin, on verra que, contrairement à ce que laisse entendre l'article défini *le* utilisé pour *itinéraire optimal* dans le langage courant, celui-ci n'est pas nécessairement unique.

1.4 Optimiser sous contraintes

Un pan important dans l'optimisation est celui de l'optimisation sous contraintes. Il s'agit à nouveau d'une notion très naturelle. Considérons un premier exemple : on fait des achats en ligne sur un site qui offre les frais de port pour des commandes d'un montant supérieur à 50 euros. On peut alors souhaiter valider un panier pour un montant aussi petit que possible, tout en restant au-dessus des 50 euros demandés.

Un autre exemple, plus formel, est celui de l'optimisation d'un (temps de) trajet en métro. Pour ce problème, on peut envisager plusieurs contraintes différentes :

- contrainte 1 : ne pas dépasser un certain nombre de correspondances ;
- contrainte 2 : n'emprunter que des lignes accessibles aux fauteuils roulants ;
- ne pas dépasser un certain coût (en restant dans une certaine zone tarifaire) ;

chaque contrainte pouvant mener à une solution différente.

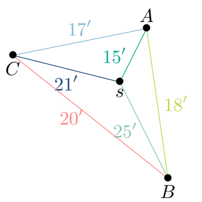
1.5 Comment trouver une solution optimale ?

Maintenant que l'on a clarifié la définition d'un problème d'optimisation, il faut s'intéresser aux méthodes de résolution. Pour cela, on va considérer deux exemples.

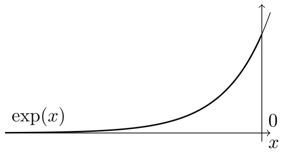
Exemple

On souhaite optimiser le temps mis pour se rendre en 3 lieux A, B et C, en partant du point s et en y retournant. Les temps de trajet pour joindre deux lieux sont donnés par le diagramme suivant.

- itinéraire 1 : $s \to A \to B \to C \to s$: 1h14
- itinéraire 2 : $s \to A \to C \to B \to s$: 1h17
- itinéraire 3 : $s \to B \to C \to A \to s$: 1h17
- itinéraire 4 : $s \to B \to A \to C \to s$: 1h21
- itinéraire 5 : $s \to C \to A \to B \to s$: 1h21
- itinéraire 6 : $s \to C \to B \to A \to s$: 1h14

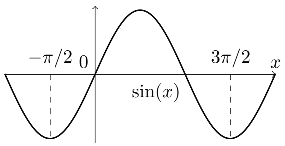


Minimiser $\exp(x)$ sous les contraintes $x \leq 0$



Ce problème ne possède pas de solution optimale. Remarque : la fonction tend vers 0 lorsque x tend vers $-\infty$, mais sans jamais l'atteindre.

Minimiser $\sin(x)$ sous les contraintes $-\pi \leq x \leq 2\pi$

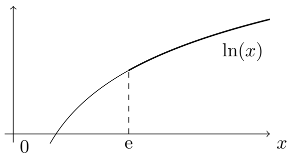


Ce problème possède deux solutions optimales,

$$x^* = -\frac{\pi}{2} \text{ et } x^* = \frac{3\pi}{2}$$

Le minimum vaut alors -1.

Minimiser $\ln(x)$ sous les contraintes $x \geq \mathrm{e}$



Ce problème possède une unique solution optimale

$$x^* = \mathrm{e}$$

Le minimum vaut alors 1.

- si x^* est solution optimale de $(\mathcal{P})$, alors x^* est solution optimale de $(\mathcal{P}')$;
- si y^* est solution optimale de $(\mathcal{P}')$, alors $|y^*|$ est solution optimale de $(\mathcal{P})$.

On peut maintenant définir l'équivalence entre un problème de minimisation et un problème de maximisation et deux problèmes de maximisation, à l'aide du problème :

$(\mathcal{D}')$ Maximiser $-g(y)$ sous les contraintes $y \in \mathcal{C}'$

Définition 7 (Problèmes équivalents (II))

- Les problèmes de minimisation $(\mathcal{P})$ et de maximisation $(\mathcal{D}')$ sont dit *équivalents* si les problèmes de minimisation $(\mathcal{P})$ et $(\mathcal{P}')$ le sont.
- Les problèmes de maximisation $(\mathcal{D})$ et $(\mathcal{D}')$ sont dit *équivalents* si les problèmes de minimisation $(\mathcal{P})$ et $(\mathcal{P}')$ le sont.

Puisque l'on peut passer aisément d'un problème de maximisation à un problème de minimisation, on considérera maintenant uniquement des problèmes de minimisation.

2.4 Existence, unicité des solutions optimales

Les notions d'existence et d'unicité éventuelle des solutions optimales d'un problème d'optimisation seront abordées de manière plus détaillée dans un cours prochain.

Il faut d'abord remarquer que, s'il existe, le minimum d'une fonction est **unique**. En effet, revenons à la définition et supposons qu'il existe deux minima pour une même fonction $f : X \to \mathbb{R}$. Cela revient à supposer l'existence de deux minimiseurs $x_1^* \neq x_2^*$ tels que les minima associés $f(x_1^*)$ et $f(x_2^*)$ soient différents. Dans ce cas, la définition du minimiseur x_1^* s'écrit

$$\forall\, x \in X, \quad f(x_1^*) \leq f(x)$$

et puisque $x_2^* \in X$, on a en particulier

$$f(x_1^*) \leq f(x_2^*)$$

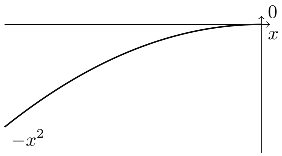
De la même manière, la définition du minimiseur x_2^* implique que

$$f(x_2^*) \leq f(x_1^*)$$

Par conséquent, on a montré que $f(x_1^*) = f(x_2^*)$, ce qui contredit notre hypothèse de départ (à savoir que $f(x_1^*) \neq f(x_2^*)$), qui est donc fausse. En revanche, les minimiseurs peuvent ne pas être uniques, comme le montrent les exemples suivants.

Exemple

Minimiser $-x^2$ sous les contraintes $x \leq 0$



Ce problème ne possède pas de solution optimale. Remarque : la fonction tend vers $-\infty$ lorsque x tend vers $-\infty$.

Dans ce premier exemple, on trouve les deux itinéraires optimaux en testant *tous* les itinéraires possibles, puis en sélectionnant celui ou ceux qui présentent le temps de parcours minimal. Une telle opération est possible sur des petits graphes, dans lesquels l'exploration de toutes les possibilités reste possible.

Exemple

Un fermier souhaite construire un enclos rectangulaire avec 10 mètres de barrière, tout en maximisant la surface de l'enclos. Soit L la longueur et $\ell \leq L$ la largeur de l'enclos. Le périmètre de l'enclos vaut $2 \times (L + \ell) \leq 10$m.

Relation entre L et ℓ : $0 \leq \ell \leq L \leq \underbrace{5 - \ell}_{=L(\ell)}$.

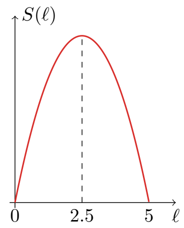
Surface max. de l'enclos pour ℓ fixé :

$$S(\ell) = L(\ell) \times \ell = (5 - \ell)\,\ell$$

avec $S'(\ell) = 5 - 2\,\ell$

Solution : $S_{\max} = 6.25\text{m}^2$

atteinte pour $\{\ell_{\max}, L_{\max}\} = \{2.5\text{m}, 2.5\text{m}\}$.



Pour ce second exemple, la résolution passe par l'étude des variations de la fonction, qui montre qu'elle présente un unique maximum, donné par le point où la dérivée s'annule. Une telle démarche n'est possible que si la fonction à optimiser est dérivable, que l'on sait calculer cette dérivée, et que l'on sait comment trouver les zéros (ou au moins un zéro) de cette dérivée.

1.6 Quelques exemples d'applications

L'optimisation n'apparaît pas uniquement dans les exemples présentés dans ce cours et dans ceux qui suivent. Dans la vie courante, elle est beaucoup utilisée dans des problèmes de recherche d'itinéraire (GPS, Google maps), où intervient par ailleurs la problématique de mise-à-jour en temps réel (si on se trompe en suivant les instructions, le système doit être capable de s'adapter et de recalculer un nouvel itinéraire optimal).

De manière moins évidente, d'autres problèmes utilisent l'optimisation pour leur résolution. On pourra citer par exemple les problèmes de restauration d'images ou de segmentation d'images, ou encore de colorisation automatique de vidéos (voir Fig. 1–3). Dans ces trois cas, le problème sous-jacent s'écrit comme la recherche d'un minimiseur d'une certaine fonction, qui peut être complexe.

1.7 Les enjeux de l'optimisation moderne

Pour être utile et utilisable, l'optimisation se doit donc de satisfaire plusieurs critères :

- il faut être capable de résoudre un grand nombre de problèmes, quitte à adapter la méthode de résolution au problème considéré ;
- la résolution numérique (c'est-à-dire à l'aide d'un ordinateur) doit être *suffisamment* rapide (*suffisamment* pour l'application considérée, de l'ordre de quelques secondes pour un GPS, alors que des minutes ou des heures de calculs peuvent être envisageables si le calcul ne doit être effectué qu'une seule fois) ;
- et être *suffisamment* précise (si le GPS ne trouve pas un itinéraire optimal, mais un itinéraire dont le temps de trajet diffère de deux minutes au plus par rapport au temps optimal, cela peut rester acceptable).

FIGURE 1 – Restauration d'images numériques. En haut, une image dégradée, en bas, l'image restaurée à l'aide d'une méthode basée sur l'optimisation d'une fonction objectif. *Résultats obtenus par la méthode décrite dans l'article* An Algorithm for Total Variation Minimization and Applications*, A.* CHAMBOLLE*, 2004.*

2.3 Problèmes équivalents

On va maintenant expliciter le lien entre les problèmes de minimisation et de maximisation, ce qui permettra de considérer le cas plus général des problèmes d'optimisation équivalents.

Revenons au problème de *minimisation* :

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

Soit x^* une solution optimale de $(\mathcal{P})$. Alors, par définition, $x^* \in \mathcal{C}$ et

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

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

Autrement dit, x^* est une solution optimale du problème :

$(\mathcal{D})$ Maximiser $-f(x)$ sous les contraintes $x \in \mathcal{C}$

Ainsi, pour résoudre un problème de minimisation, on peut passer par la résolution d'un problème de maximisation, puisque les problèmes $(\mathcal{P})$ et $(\mathcal{D})$ ont mêmes solutions optimales. On dit que ces deux problèmes sont *équivalents*.

Pour aborder cette notion d'équivalence de manière plus formelle, considérons deux fonctions $f : X \to \mathbb{R}$ et $g : Y \to \mathbb{R}$, deux ensembles $\mathcal{C} \subset X$ et $\mathcal{C}' \subset Y$, avec X et Y deux espaces vectoriels, ainsi que les deux problèmes de minimisation sous contraintes suivants :

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

$(\mathcal{P}')$ Minimiser $g(y)$ sous les contraintes $y \in \mathcal{C}'$

Définition 6 (Problèmes équivalents)

Les problèmes de minimisation $(\mathcal{P})$ et $(\mathcal{P}')$ sont dit *équivalents* s'il existe deux fonctions $\mathcal{M} : X \to Y$ et $\mathcal{N} : Y \to X$ telles que

$$x^* \text{ une solution optimale de } (\mathcal{P}) \Longrightarrow \mathcal{M}(x^*) \text{ une solution optimale de } (\mathcal{P}')$$

$$y^* \text{ une solution optimale de } (\mathcal{P}') \Longrightarrow \mathcal{N}(y^*) \text{ une solution optimale de } (\mathcal{P})$$

Il faut remarquer que cette définition est plus générale que celle usuellement donnée, car elle n'assure pas que l'on puisse, à partir de l'ensemble des solutions optimales d'un des deux problèmes, obtenir toutes les solutions optimales de l'autre problème. Néanmoins, si l'on est uniquement intéressé par trouver **une** solution optimale d'un problème donné (et pas toutes les solutions optimales), comme ce sera le cas dans la suite du cours, alors elle est suffisante.

EXEMPLE

Exemple de problèmes équivalents. Les deux problèmes

$(\mathcal{P})$ Minimiser $-x^2$ sous les contraintes $x \in [0, 1]$

$(\mathcal{P}')$ Minimiser $-y^2$ sous les contraintes $y \in [-1, 1]$

sont équivalents car

et on dira de manière analogue que x^* réalise le maximum de f ou que le maximum de f est atteint au point x^*, ou encore que x^* est une *solution optimale* du problème non contraint :

$$\text{Maximiser } f \text{ sur } X$$

2.2 Optimisation sous contraintes

Introduisons maintenant un ensemble de contraintes $\mathcal{C} \subset X$, et considérons le problème de *minimisation* $(\mathcal{P})$ introduit au début de ce paragraphe.

Définition 3 (Point admissible, ensemble admissible)

On dit que $x \in X$ est un *point admissible* du problème $(\mathcal{P})$ si $x \in \mathcal{C}$. On appelle alors *admissible* l'ensemble des points admissibles d'un problème d'optimisation.

Définition 4 (Solution optimale d'un problème de minimisation)

On dit que le point $x^* \in X$ est une *solution optimale* du problème de minimisation $(\mathcal{P})$ si

1. x^* est un point admissible du problème d'optimisation $(\mathcal{P})$
2. pour tout x point admissible de $(\mathcal{P})$

$$f(x^*) \leq f(x)$$

S'il existe, on dit aussi que x^* réalise le minimum de f sur $\mathcal{C}$ ou que le minimum de f sur $\mathcal{C}$ est atteint au point x^*. On définit de même les solutions optimales d'un problème de *maximisation* sous contraintes, en inversant le sens de l'inégalité dans la définition 4.

Définition 5 (Solution optimale d'un problème de maximisation)

On dit que le point $x^* \in X$ est une *solution optimale* du problème de maximisation $(\mathcal{P})$ si

1. x^* est un point admissible du problème d'optimisation $(\mathcal{P})$
2. pour tout x point admissible de $(\mathcal{P})$

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

Il faut remarquer que la définition d'un point admissible (resp. d'un ensemble admissible) d'un problème d'optimisation ne dépend pas de la nature exacte du problème (c'est-à-dire qu'elle est la même que le problème soit un problème de minimisation ou de maximisation), alors que celle de solution optimale en dépend (puisque le sens de l'inégalité change).

FIGURE 2 – Segmentation d'images. En haut, une image numérique (une vue nocturne de l'Europe de l'ouest), en bas, un masque permettant de classer les pixels de l'image dans deux ensembles différents (terre en noir *versus* mer/océan en blanc). *Résultats obtenus par la méthode décrite dans l'article* Active Contours Without Edges*, T.F.* CHAN *and L.A.* VESE*, 2001.*

FIGURE 3 – Colorisation automatique d'images et de vidéos. En haut à gauche, une image en niveaux de gris (sans couleur). En haut à droite, une palette de couleurs donnée par une image similaire à la première image, en couleur. En bas, la première image colorisée à l'aide de la seconde, sans intervention d'un opérateur humain. *Résultats obtenus par la méthode décrite dans l'article* Inertial Alternating Generalized Forward-Backward Splitting for Image Colorization, *P.* TAN, *F.* PIERRE, *and M.* NIKOLOVA, *2018.*

Dans ce module, on verra que, pour certaines classes de problèmes d'optimisation (linéaire, quadratique, convexe lisse,...), on peut proposer une méthode de résolution adaptée. Cependant, il faut garder à l'esprit que l'on ne sait pas résoudre un grand nombre de problèmes d'optimisation à l'heure actuelle.

2 Un peu de vocabulaire

Dans cette section, on va introduire des termes de vocabulaire qui nous permettrons d'aborder rigoureusement l'optimisation d'un point de vue mathématique.

Commençons par les problèmes de *minimisation*. Ils s'écrivent généralement sous la forme suivante :

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

où

- $f : X \to \mathbb{R}$ est la *fonction objectif*, ou le *critère* à optimiser ;
- X est un espace vectoriel ;
- $\mathcal{C} \subset X$ est l'ensemble des *contraintes*.

Notons que la fonction objectif f est à valeurs réelles, mais qu'elle peut être définie sur un espace vectoriel de dimension quelconque.

2.1 Minimum, maximum, minimiseurs, maximiseurs

Considérons tout d'abord le cas d'un problème *non contraint*, c'est-à-dire celui où l'ensemble des contraintes $\mathcal{C}$ est l'espace tout entier X.

Définition 1 (Minimiseur, minimum)

Soit $f : X \to \mathbb{R}$. On dit que $x^* \in X$ est un *minimiseur* de f si

$$\forall x \in X, \quad f(x^*) \le f(x)$$

Dans ce cas, la valeur $f(x^*) \in \mathbb{R}$ est appelée *minimum* de f.

S'il existe, on dit aussi que x^* réalise le minimum de f ou que le minimum de f est atteint au point x^*. On peut également dire que x^* est une *solution optimale* du problème non contraint :

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

De la même manière, on peut définir

Définition 2 (Maximiseur, maximum)

Soit $f : X \to \mathbb{R}$. On dit que $x^* \in X$ est un *maximiseur* de f si

$$\forall x \in X, \quad f(x^*) \ge f(x)$$

Dans ce cas, la valeur $f(x^*) \in \mathbb{R}$ est appelée *maximum* de f.