Optimisation Non Linéaire Sous Contraintes

Année universitaire 2024/2025

Optimisation non linéaire avec contraintes pour Master Modélisation et techniques de décision

Se reporter à des manuels de base et à certaines recherches

Septembre 2023

Chapitre 1

Introduction et Motivation

1.1 introduction

Prérequis : Algèbre linéaire, le calcul matriciel et la géométrie affine euclidienne.

Note : Dans ce cours, on se focalise sur les techniques d'optimisations (critère d'optimalité, démarches et Algorithmes) plutôt que la modélisation qui est aussi une étape importante dans la chaîne des problèmes d'optimisation.

L'optimisation vise à résoudre des problèmes dont on cherche à trouver une solution satisfaisant sur un ensemble définis par des contraintes linéaires et non linéaires (égalités et inégalités). Elle cherche à minimiser ou à maximiser une fonction d'une caractéristique numérique. Souvent la théorie et les méthodes de résolution de ces problèmes d'extrémum partent d'un point de domaine des solutions réalisables S, en suite à base d'une stratégie bien tracée on détermine une valeur exacte ou approchée de la solution optimale.

On générale dans le domaine de la programmation Mathématique, on peut trouver les différents problèmes suivants :

- La programmation linéaire : Traite la résolution des problèmes d'optimisation pour lesquels la fonction objective et les contraintes sont affines.
- La programmation quadratique : est la minimisation d'une fonction objective quadratique sous des contraintes linéaires.
- La programmation non linéaire : est la recherche de l'optimum d'une fonction non linéaire sur un sous-ensemble convexe ou non d'un espace donné.

Le problème non linéaire s'écrit sous la forme,

$$\begin{cases} Min \quad f(x) \\ s.c. \\ \ \ h(x) = 0 \\ \ \ g(x) \leq 0 \end{cases}$$

Où : $x \in \mathbb{R}^n$, f, h et g des fonctions tel que $f : \mathbb{R}^n \to \mathbb{R}$, $\ h : \mathbb{R}^n \to \mathbb{R}^p$ et $g : \mathbb{R}^n \to \mathbb{R}^m$.

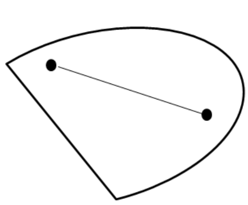

FIGURE 1.1 – Ensemble convexe

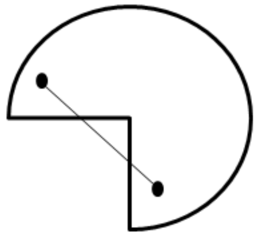

FIGURE 1.2 – Ensemble non convexe

Remarque 1.1.1. *Depuis les années cinquante, un nombre important des travaux ont été publiés traitant de problèmes très variés : continus et combinatoires, linéaires et non-linéaires, algébriques ou dans les graphes, déterministes ou stochastiques.*

Domains d'applications : la gestion et le planning industriels, établissement des projets et la planification à long terme, le domaine militaire.

Propriétés : $x, y \in E$: ensembles d'un espace euclidien.
- Addition : $x + y = (x_l + y_1, x_2 + y_2, \ldots, x_n + y_n)$,
- Multiplication : $ax = (ax_l, ax_2, \ldots, ax_n)$,
- Produit scalaire : $< x, y >= \sum_{i=1}^{n} x_i y_i$,
- Longueur d'un vecteur : $\|x\| = (x, x)^{1/2}$,
- Inégalités de normes : $\|x + y\| \leq \|x\| + \|y\|, \quad | < x, y > | \leq \|x\| \cdot \|y\|$.

1.2 Définitions et vocabulaire

Définition 1.2.1. *Une solution $x^* \in S$ est un minimum global de la fonction f sur S si,*

$$f(x^*) \leq f(x), \forall x \in S. \tag{1.1}$$

Remarque 1.2.1. *Le minimum global x^* n'est pas unique mais la valeur $f(x^*)$ l'est.*

Définition 1.2.2. *Une solution $x^* \in S$ est un minimum local de la fonction f sur S si,*

$$f(x^*) \leq f(x), \forall x \in B_\epsilon(x^*) \subseteq S. \tag{1.2}$$

Tel que : $B_\epsilon(x^*)$ est appelée une boule de rayon ϵ centrée en x^* avec,

$$B_\epsilon(x^*) = \{x \in \mathbb{R}^n, \| x - x^* \| < \epsilon\}. \tag{1.3}$$

Définition 1.2.3. *Le gradient d'une fonction $f : \mathbb{R}^n \to \mathbb{R}$, différentiable est un vecteur de $\mathbb{R}^n$ s'écrivant,*

$$\nabla f(x) = (\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots, \frac{\partial f}{\partial x_n})^T, \;\; x \in \mathbb{R}^n$$

Si les dérivées secondes existent et sont continues, alors la matrice Hessienne s'écrit,

$$\nabla^2 f(x) = \begin{pmatrix} f''_{x_1x_1} & f''_{x_1x_2} & \cdots & f''_{x_1x_n} \\ f''_{x_2x_1} & f''_{x_2x_2} & \cdots & f''_{x_2x_n} \\ . & . & & . \\ . & . & \cdots & . \\ . & . & & . \\ f''_{x_nx_1} & f''_{x_nx_2} & \cdots & f''_{x_nx_n} \end{pmatrix}.$$

Exemple 1.2.1. *Soit $f(x) = e^{x_2} - (x_1^2 - 3x_2)^2$. Le gradient et la matrice Hessienne sont :*

$$\nabla f(x) = (-4x_1(x_1^2 - 3x_2), e^{x_2} + 6(x_1^2 - 3x_2))^T,$$

$$\nabla^2 f(x) = \begin{pmatrix} -12x_1^2 + 12x_2 & 12x_1 \\ 12x_1 & e^(x_2) - 18 \end{pmatrix}.$$

Définition 1.2.4. *Soit S un ouvert de $\mathbb{R}^n$ et $a \in S$. On dit que f est differentiable en a s'il existe une forme linéaire $d_a f$ telle que :*

$$f(a+h) = f(a) + d_a f(h) + \parallel h \parallel \varepsilon(h) =< \nabla f(a), h > + \parallel h \parallel \varepsilon(h),$$

avec $\varepsilon(h) \to 0$ et $\parallel h \parallel \to 0$.

Définition 1.2.5. *la courbe de niveau k d'une fonction f de deux variables $(x, y) \to f(x, y)$, est l'ensemble des points du plan (x, y) qui vérifient l'équation $f(x, y) = k$.*

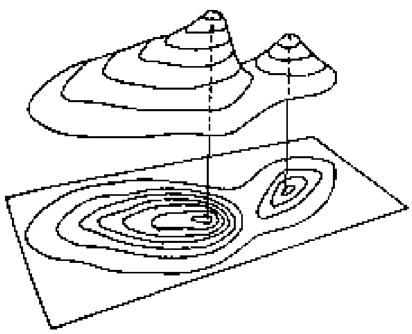

FIGURE 1.3 – courbes de niveau

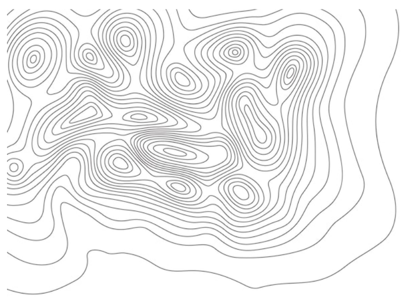

FIGURE 1.4 – Courbes de niveau

Proposition 1.2.1. *Le vecteur gradient $\nabla f(x_0, y_0)$ est orthogonal à la ligne de niveau de f passant au point (x_0, y_0). Il indique la direction de plus forte pente à partir de ce point.*

Proposition 1.2.2. *L'équation de la tangente à la ligne de niveau de f en (x_0, y_0) est*

$$\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) = 0. \tag{1.4}$$

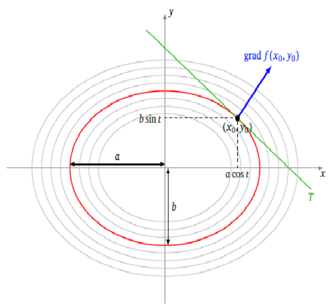

FIGURE 1.5 – courbes de niveau

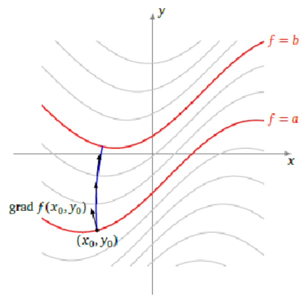

FIGURE 1.6 – la plus grande pente

Définition 1.2.6. *Une fonction $f : \mathbb{R}^n \to \mathbb{R}$ est coercive si $f(x) \to +\infty$ lorsque $\| x \| \to +\infty$*

Remarque 1.2.2. *A une matrice symétrique carrée, Alors la fonction*

$$f(x) = x^T Ax + b^T x + c \tag{1.5}$$

est coercive, si et seulement si A est une matrice définie positive.

Définition 1.2.7. *Soit A, une matrice symétrique réelle de taille n. On appelle* ***mineurs principaux successifs*** *les déterminants des n matrices $A_p = (a_{ij})_{(1 \leq i,j \leq p)}$, pour $p = 1, \dots, n$.*

Théorème 1.2.1. *(Critère de Sylvester)*

1. *Une forme quadratique A est définie positive (semi-définie positive) si et seulement si tous les mineurs principaux successifs de la matrice A sont positifs (≥ 0).*
2. *Une forme quadratique A est définie négative (semi-définie négative) si et seulement si tous les mineurs principaux successifs de la matrice A alternent en signe, le premier étant négatif (≤ 0).*
3. ***Sinon**. La forme quadratique change de signe.*

Remarque 1.2.3. *(Autre forme de calcul)*

1. Toute matrice symétrique est diagonalisable.

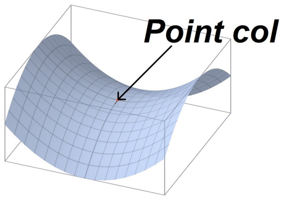

FIGURE 1.7 – Point selle

2. Une matrice est positive (resp. définie positive), si et seulement si, toutes ses valeurs propres sont positives (resp. strictement positive).

N.B : En optimisation convexe, un point de selle est un point où la courbe est plate, mais pas nécessairement un minimum ou un maximum. Cela signifie que la dérivée de la fonction est nulle, mais que la courbe ne change pas de direction à cet endroit. Les points de selle sont importants dans certaines méthodes d'optimisation, car ils peuvent poser des problèmes lors de la recherche de minimums ou de maximums.

Définition 1.2.8. *Pour tout $x \in \mathbb{R}^n$, on appelle la dérivée directionnelle de f en x de direction h,*

$$d_x = \frac{\partial f}{\partial h} =< \nabla f(x), h >; h \in \mathbb{R}^n.$$

Et on note $g(t) = f(x + th)$ tel que, $g(0) = f(x)$.

Lemme 1.2.1. *(Farkas) Soient c et a_i pour $i = \{1, \ldots, k\}$ des vecteurs de $\mathbb{R}^n$. On fait l'hypothèse suivante,*

$$\forall v \in \mathbb{R}^n, si < a_i, v >\geq 0, alors < c, v >\geq 0 \tag{1.6}$$

Donc, il existe des réels $\lambda_i \geq 0$ tel que,

$$c = \sum_{i=1}^{k} \lambda_i a_i. \tag{1.7}$$

1.3 Rappels de differentiabilité

1. Si f est continue sur S, de classe C^1 en x. on a

$$f(x + h) = f(x) + d_x f(h) + \parallel h \parallel \varepsilon(h),$$

avec $\varepsilon(h) \to 0$ et $\parallel h \parallel \to 0$.
2. (Formule de Taylor Mac-Laurin) Si f est continue sur S, de classe C^1 sur $]x, x + \theta h[\subset S$,

alors

$$f(x+h) = f(x) + d_{x+\theta h}f(h) = f(x) + < \nabla f(x+\theta h), h > .$$

3. (Formule de Taylor-Lagrange) Si f est continue sur S, de classe C^2 en x. on a

$$f(x+h) = f(x) + d_x f(h) + \frac{1}{2} d_x^2 f(h,h) + \| h \|^2 \varepsilon(h),$$

avec $\varepsilon(h) \to 0$ et $\| h \| \to 0$.

4. (Formule de Taylor-Mac-Laurin) Si f est continue sur S, de classe C^2 en $]x, x+\theta h[\subset S$. on a

$$f(x+h) = f(x) + d_{x+\theta h}f(h) + \frac{1}{2} d_{x+\theta h}^2 f(h,h) = f(x) + < \nabla f(x+\theta h), h > + \frac{1}{2} < \nabla f(x+\theta h) h, h > .$$