

# Optimisation numérique 1: Introduction

Aleš Janka

24 septembre 2009

## 1 Modélisation et ingrédients d'une optimisation

Le processus de formalisation du problème, d'identification des objectifs et des variables à optimiser est appelé la *modélisation*. C'est une étape importante : il faut choisir un modèle ni trop simpliste ni trop coûteux à évaluer. Les ingrédients d'une optimisation, issue d'un processus de modélisation sont représentées par :

- **les variables / paramètres d'optimisation**  $\vec{u} \in U \subset V$  : l'optimum est cherché dans un ensemble  $U$  plongé dans l'espace vectoriel  $V$ . Très souvent on pose  $V = \mathbb{R}^n$ ,
- **la fonction de coût (objectif)**  $f : U \mapsto \mathbb{R}$  : une mesure quantitative de qualité d'un candidat à l'optimum,
- **les contraintes sur les variables d'optimisation** : représentés soit par une application  $\vec{c} : V \mapsto \mathbb{R}^m$  dite "vecteur de contraintes", ou par le choix de l'ensemble des solutions admissibles  $U \subset V$  (*feasible region*).

## 2 Formulation mathématique

Supposons que  $V \equiv \mathbb{R}^n$ . Le problème de minimisation dans  $\mathbb{R}^n$  avec contraintes s'écrit : trouver  $\vec{u}^* \in \mathbb{R}^n$  tel que

$$f(\vec{u}^*) = \min_{\vec{u} \in \mathbb{R}^n} f(\vec{u}) \quad \text{où} \quad \begin{cases} c_i(\vec{u}^*) = 0, & i \in \mathcal{E}, \\ c_j(\vec{u}^*) \geq 0, & j \in \mathcal{I}. \end{cases} \quad (1)$$

Ici,  $\mathcal{E}$  resp.  $\mathcal{I}$  sont des sous-ensembles d'indices  $\{1, 2, \dots, m\}$  pour dénoter les contraintes d'égalité, resp. d'inégalité. On peut alors, de manière équivalente, choisir l'ensemble des solutions admissibles

$$U \equiv \{\vec{v} \in \mathbb{R}^n : c_i(\vec{v}) = 0, \forall i \in \mathcal{E} \wedge c_j(\vec{v}) \geq 0, \forall j \in \mathcal{I}\}.$$

L'Exemple 1 illustre que, très souvent, il est nécessaire de reformuler notre problème, pour le rendre dans la forme (1).

**Exemple 1** (Minimisation de l'aire de surface, voir Fig. 1). Soit  $\Omega = (0, 1) \times (0, 1) \subset \mathbb{R}^2$  un carré unité. Soit  $u(\vec{x})$  la fonction de la hauteur,  $\vec{x} \in \Omega \mapsto \mathbb{R}$ , connue au bord de  $\Omega$ ,  $u(x, y) = u_\Gamma$  sur  $\Gamma = \partial\Omega$ . L'aire de la membrane se calcule par la formule

$$A(u) = \int_{\Omega} \sqrt{1 + \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2} dx dy = \int_{\Omega} \sqrt{1 + (\nabla u)^2} d\Omega$$

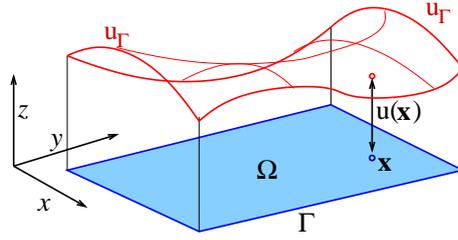


FIG. 1 – Exemple 1 : Membrane avec son bord fixé

Pour  $u_\Gamma$  donné, minimisons la surface  $A(u)$  : on cherche une fonction  $u^* \in H^1(\Omega)$  telle que

$$A(u^*) = \min_u A(u). \quad (2)$$

Pour pouvoir passer de l'espace des fonctions  $H^1(\Omega)$  à  $\mathbb{R}^n$ , il faut approcher la fonction  $u$  par une fonction  $u_h = \sum_{i=1}^n u_i \phi_i(\vec{x})$ . Ici,  $\{\phi_i(\vec{x})\}_{i=1}^n$  est une base. On définit le vecteur des paramètres d'optimisation  $\vec{u} = (u_1, \dots, u_n)^T \in \mathbb{R}^n$ . En définissant la fonction de coût  $f : \mathbb{R}^n \mapsto \mathbb{R}$  par

$$f(\vec{u}) = A\left(\sum_{i=1}^n u_i \phi_i(x, y)\right),$$

le problème (2) est alors approché par le problème de trouver  $\vec{u}^* \in U \subset \mathbb{R}^n$  tel que

$$f(\vec{u}^*) = \min_{\vec{u} \in U} f(\vec{u}). \quad (3)$$

Ici, l'ensemble  $U \subset \mathbb{R}^n$  est choisi de façon de satisfaire aux conditions de bord  $u^* = u_\Gamma$  sur  $\Gamma$ . Le problème (3) est déjà dans la forme (1).

### 3 Quelques types de problèmes d'optimisation

On peut classifier les problèmes d'optimisation selon la forme particulière d'ingrédients de (1).

#### 3.1 Problèmes de programmation linéaire ou quadratique

On parle des problèmes de programmation linéaire, quand aussi bien  $f : V \mapsto \mathbb{R}$  et  $\vec{c} : V \mapsto \mathbb{R}^m$  sont des applications linéaires dans  $V$ . Si  $f : V \mapsto \mathbb{R}$  est une fonction quadratique de  $\vec{u}$ , alors on parle des problèmes de programmation quadratique.

**Exemple 2** (Problème du transport). La production de  $m$  usines fournit  $n$  boutiques à travers du pays. Chaque usine peut produire au maximum  $a_i$ ,  $i = 1, \dots, m$  tonnes du produit par semaine et chaque boutique écoule  $b_j$ ,  $j = 1, \dots, n$  tonnes du produit par semaine. Comment organiser la livraison du produit des usines aux boutiques de manière la plus rentable, quand on sait que le coût de transport d'une tonne de produit de l'usine  $i$  au boutique  $j$  coûte  $c_{ij}$  ?

On définit  $x_{ij}$ , la quantité du produit livré (par semaine) de l'usine  $i$  au boutique  $j$ . La facture de livraison hebdomadaire de boutique  $j$  est de  $\sum_{i=1}^m c_{ij} x_{ij}$ . La facture totale est de

$$f(\vec{u}) = \sum_{j=1}^n \sum_{i=1}^m c_{ij} x_{ij}.$$

Ici,  $\vec{u}$  dénote en fait la matrice  $\{x_{ij}\}$ . La contrainte de production se formalise par

$$\sum_{j=1}^n x_{ij} \leq a_i \quad i = 1, \dots, m,$$

la contrainte d'avoir la marchandise en boutique au moment de la vente donne

$$\sum_{i=1}^m x_{ij} \geq b_j \quad j = 1, \dots, n,$$

et le flux positif de marchandise des usines aux boutiques donne

$$x_{ij} \geq 0 \quad \forall i, j.$$

On a alors un problème typique de la programmation linéaire.

### 3.2 Optimisation continue vs. optimisation discrète

Dans l'optimisation discrète, en plus des contraintes  $\vec{c}$  on a aussi la contrainte que  $\vec{u}$  soit entier. On peut diviser les problèmes discrets en problèmes combinatoires (résolution de sudoku, par exemple), et en problèmes de programmation en nombres entiers.

En générale, les problèmes d'optimisation continue sont plus facile à résoudre que les problèmes discrets. Dans ce cours on traitera uniquement des problèmes continus.

### 3.3 Optimisation sans et avec contraintes

Si l'ensemble admissible  $U$  couvre tout un espace vectoriel,  $U \equiv V$ , on parle d'une optimisation sans contraintes. En générale, ce type de problème d'optimisation est plus facile à résoudre.

L'optimisation avec contraintes limite l'ensemble admissible  $U$  à un sous-ensemble de  $V$ . Une possible technique de résolution de ce type de problèmes est en utilisant la technique de pénalisation : on ajoute à la fonction coût  $f : V \mapsto \mathbb{R}$  (qu'on veut minimiser) un terme non-négatif, qui devient grand pour  $\vec{u} \notin U$ . Une autre technique consiste dans l'introduction des multiplicateurs de Lagrange.

### 3.4 Optimisation globale et locale

Les méthodes d'optimisation les plus efficaces procèdent par une amélioration itérative d'un candidat  $\vec{u} \in U$  à l'optimum. Elles cherchent dans le voisinage de  $\vec{u} \in U$  un "meilleur"  $\vec{v} \in U$  pour lequel on a  $f(\vec{v}) < f(\vec{u})$ . Cette technique peut néanmoins assurer seulement la convergence vers un point  $\vec{u}$  qui est localement optimal. Pour avoir la certitude que ce point est un optimum global, on doit connaître le comportement global de la fonction coût  $f$ . Soit on doit évaluer  $f(\vec{u})$  en beaucoup de points de  $U$  (ce qui est très coûteux), soit on dispose des propriétés de  $f$  qui assurent qu'un minimum local est au même temps un minimum global (par ex. la convexité de  $f(\vec{u})$ ).

**Exemple 3** (minimum global et local, non-unicité de solution). *Une sonde spatiale de poids  $m$  se trouve dans un système de 4 planètes de poids  $m_i$  et de position  $\vec{x}_i \in \mathbb{R}^2$ ,  $i = 1, \dots, 4$ . Trouvez la/les position(s) de la sonde  $(x, y)$  de façon que la forces de gravitation dans  $(x, y)$  sont minimales.*

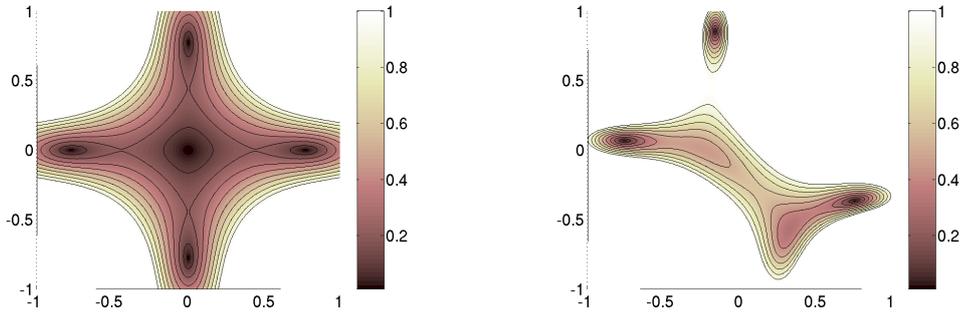


FIG. 2 – Exemple 3 : Quatre planètes sont dans les sommets du carré  $(-1, 1) \times (-1, 1)$ . On affiche ici les valeurs de  $f(\vec{u})$ , en fonction de la position de la sonde  $\vec{u} = (x, y)^T$ . A gauche le cas  $m_i = 1$ , à droite  $m_i = i, i = 1, \dots, 4$ .

C'est à dire, on trouve  $\vec{u} = (x, y)^T \in \mathbb{R}^2$  tel que

$$f(\vec{u}) = \left\| \sum_{i=1}^4 \vec{F}_i(\vec{u}) \right\|^2$$

soit minimale. Ici,  $\vec{F}_i(\vec{u})$  est la force gravitationnelle entre la sonde et la  $i$ -ème planète,

$$\vec{F}_i(\vec{u}) = (\vec{u} - \vec{x}_i) \cdot \frac{G m m_i}{\|\vec{u} - \vec{x}_i\|^3},$$

$G$  est une constante et  $\|\cdot\|$  est la norme euclidienne.

On voit sur la Fig. 2, qu'il s'agit ici d'un problème à plusieurs minima locaux.

## 4 Solution d'un problème d'optimisation

Il n'existe pas d'algorithme universel pour résoudre les problèmes d'optimisation. Plûtôt un ensemble d'algorithmes qui traitent des problèmes d'une certaine classe.

La quasi-totalité de méthodes de solution sont les méthodes itératives qui améliorent un candidat  $\vec{u} \in U$  dans son voisinage. La façon comment elles essaient améliorer  $\vec{u}$  peut varier. La majorité de méthodes utilisent la valeur  $f(\vec{u})$ , la valeur de  $\vec{c}(\vec{u})$ , voire les dérivées première et seconde de  $f$  par rapport à  $\vec{u}$ . Certaines méthodes accumulent l'information des itérations précédentes, tandis que d'autres se basent uniquement sur le pas courant.

Independamment de l'approche, une bonne méthode d'optimisation doit considérer les points suivants :

- **robustesse** : la méthode devrait fonctionner pour une large classe de problèmes, independamment des paramètres et des valeurs initiales,
- **efficacité** : elle ne devrait pas être prohibitivement chère en temps CPU et en mémoire vive,
- **précision** : le résultat numérique devrait bien approcher la vraie solution  $\vec{u}^*$  du problème de minimisation

## 4.1 Que est-ce une solution optimale ?

Supposons qu'on a un problème d'optimisation sans contraintes. Comment caractérise-t-on la solution optimale ?

**Définition 4** (minimum global). *Soit  $V \equiv \mathbb{R}^n$ ,  $f : V \mapsto \mathbb{R}$ . Le point  $\bar{u}^* \in V$  s'appelle l'optimum globale, quand*

$$f(\bar{u}^*) \leq f(\bar{u}) \quad \forall \bar{u} \in V.$$

Trouver les minima globaux serait le meilleur résultat qu'on puisse espérer. Malheureusement, pour une fonction coût  $f$  générale, c'est une tâche difficile : très souvent, on n'a pas la vision globale de  $f$ , on sait seulement évaluer  $f$  (et ses dérivées) dans quelques (peu) points  $\bar{u}_k$ ,  $k = 1, 2, \dots$ . La plupart d'algorithmes trouve seulement un minimum local.

**Définition 5** (minimum local (faible)). *Le point  $\bar{u}^* \in V$  s'appelle le minimum local (faible) quand il existe une voisinage  $\mathcal{N}$  de  $\bar{u}^*$  telle que*

$$f(\bar{u}^*) \leq f(x) \quad \forall \bar{u} \in \mathcal{N}.$$

**Définition 6** (minimum local fort). *Le point  $\bar{u}^* \in V$  s'appelle le minimum local fort quand il existe une voisinage  $\mathcal{N}$  de  $\bar{u}^*$  telle que*

$$f(\bar{u}^*) < f(x) \quad \forall \bar{u} \in \mathcal{N} \setminus \{\bar{u}^*\}.$$

**Définition 7** (minimum local isolé). *Le point  $\bar{u}^* \in V$  s'appelle le minimum local isolé quand il existe une voisinage  $\mathcal{N}$  de  $\bar{u}^*$  telle que  $\bar{u}^*$  est le seul minimum dans  $\mathcal{N}$ .*

## 4.2 Minimum local d'un $f$ lisse

Du point de vue pratique, les définitions des minima dans la Section 4.1 ne sont pas très utiles – pour s'assurer que  $\bar{u}^*$  est bien un minimum local, il faudrait parcourir tous les  $\bar{u}$  dans la voisinage de  $\bar{u}^*$ , ce qui est trop coûteux.

Heureusement, quand  $f$  est une fonction lisse de  $\bar{u}$ , et en particulier si  $f$  est deux fois continûment différentiable, on peut écrire des **conditions d'optimalité** nécessaires (et suffisantes) pour que  $\bar{u}^*$  est un minimum local en utilisant le gradient  $\nabla f(\bar{u}^*)$  et le hessien  $\nabla^2 f(\bar{u}^*)$ .

**Théorème 8** (Développement de Taylor). *Soit  $f : \mathbb{R}^n \mapsto \mathbb{R}$  contiûment différentiable et soit  $p \in \mathbb{R}^n$ . Alors*

$$f(\bar{u} + \vec{p}) = f(u) + \nabla f(\bar{u} + t\vec{p})^T \cdot \vec{p},$$

pour un  $t \in (0, 1)$ . Si en plus  $f$  est deux fois continûment différentiable, on a que

$$\nabla f(\bar{u} + \vec{p}) = \nabla f(\bar{u}) + \int_0^1 \nabla^2 f(\bar{u} + t\vec{p}) \cdot \vec{p} dt$$

et

$$f(\bar{u} + \vec{p}) = f(u) + \nabla f(\bar{u} + \vec{p})^T \cdot \vec{p} + \frac{1}{2} \vec{p}^T \cdot \nabla^2 f(\bar{u} + t\vec{p}) \cdot \vec{p},$$

pour un  $t \in (0, 1)$ .

**Théorème 9** (condition d'optimalité nécessaire, d'ordre 1). *Si  $\bar{u}^*$  est un minimum local et si  $f$  est continûment différentiable dans un voisinage ouverte de  $\bar{u}^*$ , alors  $\nabla f(\bar{u}^*) = 0$ .*

*Démonstration.* Supposons par contradiction que  $\nabla f(\bar{u}^*) \neq 0$  et posons  $\vec{p} \equiv -\nabla f(\bar{u}^*)$ . On note que  $\vec{p}^T \cdot \nabla f(\bar{u}^*) = -\|\nabla f(\bar{u}^*)\|^2 < 0$ . Mais  $\nabla f$  est continu dans le voisinage de  $\bar{u}^*$ , alors il existe un  $T > 0$  tel que

$$\vec{p}^T \cdot \nabla f(\bar{u}^* + t\vec{p}) < 0 \quad \forall t \in [0, T].$$

Pour tout  $\tilde{t} \in (0, T]$ , on a par le Théorème de Taylor que

$$f(\bar{u}^* + \tilde{t}\vec{p}) = f(\bar{u}^*) + \tilde{t}\vec{p}^T \cdot \nabla f(\bar{u}^* + t\vec{p}) \quad \text{pour un } t \in (0, \tilde{t}).$$

On a alors que  $f(\bar{u}^* + \tilde{t}\vec{p}) < f(\bar{u}^*)$  pour tout  $\tilde{t} \in (0, T)$ , ce qui contredit l'hypothèse que  $f(\bar{u}^*)$  était un minimum local.  $\square$

**Définition 10.** *Le point  $\bar{u}^*$  est appelé le point stationnaire de  $f$  quand  $\nabla f(\bar{u}^*) = 0$ .*

Chaque minimum local doit alors être le point stationnaire de  $f$ .

**Théorème 11** (condition d'optimalité nécessaire, d'ordre 2). *Soit  $\bar{u}^*$  un minimum local de  $f$  et soit  $\nabla^2 f$  continu dans un voisinage ouverte de  $\bar{u}^*$ . Alors  $\nabla f(\bar{u}^*) = 0$  et  $\nabla^2 f(\bar{u}^*)$  est une matrice définie positive.*

*Démonstration.* Théorème 9 certifie que  $\nabla f(\bar{u}^*) = 0$ . Supposons par contradiction que  $\nabla^2 f(\bar{u}^*)$  n'est pas définie positive. On peut alors choisir un vecteur  $\vec{p} \neq 0$  tel que  $\vec{p}^T \cdot \nabla^2 f(\bar{u}^*) \cdot \vec{p} < 0$ . A cause de la continuité de  $\nabla^2 f$  dans le voisinage de  $\bar{u}^*$ , il existe un  $T > 0$  tel que

$$\vec{p}^T \cdot \nabla^2 f(\bar{u}^* + t\vec{p}) \cdot \vec{p} < 0 \quad \forall t \in [0, T].$$

En utilisant le Théorème 8, on obtient pour tout  $\tilde{t} \in (0, T]$  et un  $t \in (0, \tilde{t})$  que

$$f(\bar{u}^* + \tilde{t}\vec{p}) = f(\bar{u}^*) + \tilde{t}\vec{p}^T \cdot \nabla f(\bar{u}^*) + \frac{1}{2}\tilde{t}^2 \vec{p}^T \cdot \nabla^2 f(\bar{u}^* + t\vec{p}) \cdot \vec{p} < f(\bar{u}^*)$$

Ceci montre que  $\bar{u}^*$  n'était pas le minimum local, ce qui est une contradiction.  $\square$

**Théorème 12** (condition d'optimalité suffisante). *Supposons que  $\nabla^2 f(\bar{u}^*)$  est continu dans un voisinage ouverte de  $\bar{u}^*$ , que  $\nabla f(\bar{u}^*) = 0$  et que le hessien  $\nabla^2 f(\bar{u}^*)$  est défini positif. Alors  $\bar{u}^*$  est un minimum local fort de  $f$ .*

*Démonstration.* Le hessien est continu dans le voisinage  $\bar{u}^*$  et il est défini positif dans  $\bar{u}^*$ . Alors il existe un voisinage de  $\bar{u}^*$  de rayon  $r > 0$ , dans laquelle le hessien reste défini positif. Prenons  $\vec{p} \in \mathbb{R}^n \setminus \{0\}$ , arbitraire, tel que  $\|\vec{p}\| < r$ . On a alors  $\vec{p}^T \cdot \nabla^2 f(\bar{u}^* + t\vec{p}) \cdot \vec{p} > 0$  pour tout  $t \in [0, 1]$ . Au même temps, le développement de Taylor dit

$$f(\bar{u}^* + \vec{p}) = f(\bar{u}^*) + \vec{p}^T \cdot \nabla f(\bar{u}^*) + \frac{1}{2}\vec{p}^T \cdot \nabla^2 f(\bar{u}^* + t\vec{p}) \cdot \vec{p},$$

pour un  $t \in (0, 1)$ . Comme  $\nabla f(\bar{u}^*) = 0$  et  $\vec{p}^T \cdot \nabla^2 f(\bar{u}^* + t\vec{p}) \cdot \vec{p} > 0$ , on conclut que

$$f(\bar{u}^* + \vec{p}) > f(\bar{u}^*).$$

$\square$

**Théorème 13** (minimum global). *Soit  $f : \mathbb{R}^n \mapsto \mathbb{R}$  convexe. Tout minimum local  $\bar{u}^*$  de  $f$  est un minimum global. Si de plus  $f$  est différentiable, alors tout point stationnaire  $\bar{u}^*$  est un minimum global de  $f$ .*

## Références

- [1] Jorge Nocedal and Stephen J. Wright : Numerical Optimization, Springer Series in Operations Research, Springer 1999.
- [2] Page web du Proséminaire : <http://perso.unifr.ch/ales.janka/numeroptim>