

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique



N° Réf :.....

Centre Universitaire de Mila

Institut des Sciences et de la Technologie

Département de Mathématiques et Informatique

Mémoire préparé En vue de l'obtention du diplôme de licence

En : - Filière Mathématiques Fondamentales

Thème

Méthodes de recherche linéaire

Préparé par :

- Marouf Hadjer
- Kherbache Fouzia
- Kennouche Widad
- Ghichi Samira

Encadré par : Baiche Kanzia *Grade : M.A.B*

Année universitaire : 2013/2014

Remerciements

Nous tenons à remercier tout d'abord dieu, le tout puissant pour la volonté et le courage qu'il m'a donné pour mener à terme ce travail.

Nous tenons à remercier notre encadreur monsieur **Baiche Kanzia**, qui nous en couragé en nous faisant part d'observations constructives et pour ses précieux conseils.

Nous remercions nos parents, pour tous les sacrifices qu'ils ont consentis pour nous permettre de suivre nos études dans les meilleures conditions et de nous avoir en couragé tout au long de ces années.

En fin, Nous tenons à remercier toutes les amis au département de mathématiques et information, et ailleurs, pour leur disponibilité et leur soutien.

Table des matières

Introduction Générale	2
1 Optimisation sans contrainte et avec contrainte	3
1.1 Introduction	3
1.2 Notion de base	3
1.3 Eléments d'analyse convexe	3
1.4 Optimisation sans contraintes	4
1.5 Optimisation avec contraintes	7
1.6 Classification des problème d'optimisation	11
1.7 Les méthodes des optimisations sans contraintes et avec contraintes . . .	11
2 Programmation linéaire	12
2.1 Introduction	12
2.2 Programmation linéaire	12
2.3 Les Formes d'un programme linéaire	13
2.3.1 Forme standard	13
2.3.2 Forme canonique	13
2.4 Dualité d'un programme linéaire	14
3 Les méthodes recherche linéaire	16
3.1 Introduction	16

3.2	Recherche linéaire d'ordre zéro	17
3.2.1	Recherche de Fibonacci	18
3.2.2	Recherche d'or	20
3.3	Méthode d'approximation de courbes	22
3.3.1	Méthode de newton	23
	Conclusion Générale	24
	Bibliographie	25

Introduction

L'optimisation est une branche intéressante des mathématiques, cherchant à analyser et à résoudre analytiquement ou numériquement les problèmes qui consistent à déterminer le meilleur élément d'un ensemble, au sens d'un critère quantitatif donné. Le mot "optimisation" vient du mot latin "optimum" qui signifie le meilleur. L'optimisation linéaire est une discipline qui étudie ces problèmes ; Elle est également désignée par le nom de programme linéaire, terme introduit par George Dantzig vers 1947, mais cette dénomination tend à être donnée à cause de la confusion possible avec la notion de programme informatique. L'objectif de ce mémoire est de présenter certaines méthodes de recherche linéaire. La recherche linéaire est un composant fondamental de toutes les méthodes traditionnelles d'optimisation multi-dimensionnelle. La recherche linéaire est nommée par Arthur Evans, fouilleur de Knossos en 1909, à une écriture cursive, utilisant 158 idéogrammes, 87 signes syllabiques (chacun représentant une syllabe), 11 signes de poids et de mesure, 5 signes numériques (unité, dizaine, centaine, mille, dix mille). Ce mémoire est composé d'une introduction générale, trois chapitres et une conclusion générale : le premier chapitre, nous présentons des notions sur l'optimisation sans contrainte et avec contrainte ; Le deuxième chapitre, nous abordons l'étude de la programmation linéaire, dont nous exposons les techniques et les formes d'un programme linéaire ; Le dernier chapitre, nous présentons quelques méthodes de recherche linéaire (comme méthode de Fibonacci, Newton...).

Chapitre 1

Optimisation sans contrainte et avec contrainte

1.1 Introduction

Dans ce chapitre nous allons étudier les problèmes d'optimisation.

1.2 Notion de base

1.3 Eléments d'analyse convexe

Un sous ensemble C de \mathbb{R}^n est dit convexe si : $(1 - \lambda)x + \lambda y \in C, \forall x, y \in C, \forall \lambda \in [0, 1]$ Autrement dit, si le segment de droite joignant deux points quelconques $x, y \in C : [x, y] = \{(1 - \lambda)x + \lambda y : 0 \leq \lambda \leq 1\}$ est entièrement inclu dans C .

Définition 1 : Une fonction $f : C \rightarrow \mathbb{R}$ est dit convexe si :

$$\forall t \in [0, 1], \forall x, y \in C, f((1 - t)x + ty) \leq (1 - t)f(x) + tf(y).$$

Définition 2 : Une fonction $f : C \rightarrow \mathbb{R}$ est dit strictement convexe si :

$$\forall t \in [0, 1], \forall x, y \in C, f((1 - t)x + ty) < (1 - t)f(x) + tf(y).$$

Définition 3 : Un ensemble convexe de la forme $X = \{x / Ax = b, x \geq 0\}$ est appelé un polytope convexe .

Un polytope borné sera appelé un polyédre convexe .

Définition 4 : Soit X un ensemble convexe non vide de \mathbb{R}^n , un point $x \in X$ dit extrémal si l'on a :

$$\forall t \in [0, 1], \forall (y, z) \in X : x = (1 - t)y + tz \Rightarrow x = y = z.$$

Théorème 1 : (point extrême)

L'ensemble des points extrêmes d'un polytope $X = \{x \in \mathbb{R}^n / Ax = b, x \geq 0\}$ correspond à l'ensemble des solutions de base réalisables.

1.4 Optimisation sans contraintes

Le problème d'optimisation sans contraintes s'écrit sous la forme suivante :

$$\begin{cases} \min f(x) \\ x \in \mathbb{R}^n \end{cases} \quad (\text{Q})$$

où f est une fonction de \mathbb{R}^n vers $\mathbb{R} \cup \{+\infty\}$.

Résultats d'existence et d'unicité

Avant d'étudier les propriétés de la solution (où des solutions) éventuelle(s) de (Q) il faut s'assurer de son/leurs existence.

Nous donnerons ensuite des résultats d'unicité.

Définition 5 :

On dit que $f : C \longrightarrow R$ est coercive si : $\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty$.

Dans le cas où $C = R^n$ les normes sont toutes équivalentes et $\|\cdot\|$ désigne une norme quelconque de R^n .

Théorème 2 : (*existence*)

Soit $f : R^n \longrightarrow R \cup \{+\infty\}$ propre, continue et coercive alors (Q) admet au moins une solution.

Toutes les fois, on n'a pas forcément unicité. Nous donnons ci-dessous un critère pour l'unicité.

Théorème 3 : (*unicité*)

Soit $f : R^n \longrightarrow R \cup \{+\infty\}$ strictement convexe alors le problème (Q) admet au plus une solution.

Donnons pour terminer un critère qu'une fonction soit strictement convexe et coercive.

Théorème 4 :

Soit f une fonction C^1 de R^n dans R . On suppose qu'il existe $\alpha > 0$ tel que $\forall (x, y) \in R^n \times R^n \quad \langle \nabla f(x) - \nabla f(y), x - y \rangle \geq \alpha \|x - y\|^2$ alors f est strictement convexe et coercive ; en particulier (Q) admet une solution unique.

On donne dans ce qui suit les conditions pour pouvoir calculer la (où les) solution(s) optimale(s) de (Q).

Conditions d'optimalité**A) Conditions nécessaires et suffisantes du premier ordre :**

Les conditions que nous allons donner sont des conditions différentielles qui dépendent de la dérivée de la fonction à minimiser.

Théorème 5 : (condition nécessaire d'optimalité du premier ordre)

Soit C une partie de R^n et $f : C \longrightarrow R$ une fonction différentiable sur C , si x^* réalise un minimum (global ou local) de f sur C alors $\nabla f(x^*) = 0$.

Définition 6 :

Un point x^* de C vérifiant $\nabla f(x^*) = 0$ est appelé "point critique" ou "point stationnaire".

Théorème 6 : (CNS du premier ordre dans le cas convexe)

Soit $f : C \longrightarrow R$ fonction différentiable et convexe sur C , un point x^* réalise un minimum global de f sur C si et seulement si $\nabla f(x^*) = 0$.

Remarque 1 : On peut raffiner le résultat précédent en ne supposant que la locale convexité de f au voisinage de x^* , c'est-à-dire en supposant que f est convexe sur une boule centré en x^* .

A ce moment là, nous pouvons affirmer que x^* est un minimum "local" de f .

Le cas où f est convexe est fréquent dans la pratique mais pas systématique. Nous allons donc donner maintenant des conditions suffisantes pour qu'un point critique réalise un minimum (ou un maximum). Ces conditions vont faire intervenir la dérivée seconde de f , ce sont des conditions du "second ordre".

B) Conditions du second ordre :

Nous commençons par une condition nécessaire permettant de préciser encore les éventuels minimax.

Théorème 7 : (Condition nécessaire du second ordre)

On suppose que x^* est un minimum (local) de f et que f est deux fois différentiables sur C , alors :

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

b) $\langle \nabla^2 f(x^*)x, x \rangle \geq 0, \forall x \in C$.

Remarque 2 :

Dans le cas où $C = R^n$, b) est équivalent à dire que la matrice Hessienne de f en x^* : $\nabla^2 f(x^*)$ est semi-d éfinie positive.

Théorème 8 : (*Condition suffisante du seconde ordre*)

Soit f deux fois dérivable sur C vérifiant $\nabla f(x^*) = 0$ et $\exists \alpha > 0$, tel que

$$\forall x \in C \langle \nabla^2 f(x^*)x, x \rangle \geq \alpha \|x\|^2 \quad (\star)$$

alors la fonction f admet un minimum "local strict" en x^* .

Remarque 3 :

Si $C = R^n$ la condition (\star) revient à dire que la matrice Hessienne $\nabla^2 f(x^*)$ est définie positive, un choix possible pour α étant alors la plus petite valeur propre, c'est une condition deité (locale) stricte au voisinage de x^* .

1.5 Optimisation avec contraintes

La forme de problème avec contraintes est :

$$\begin{cases} \min f(x) \\ x \in C \end{cases} \quad (\text{P})$$

où C est l'ensemble des contraintes.

Résultats d'existence et d'unicité

Commençons par donner un résultat d'existence.

Théorème 9 :

Supposons que f est continue, que C est un sous-ensemble fermé non vide de R^n et que l'une des conditions suivantes est réalisée :

a) Soit C est borné.

b) Soit f est coercive.

Alors le problème (P) admet au moins une solution.

Théorème 10 :

Sous les hypothèses du théorème 5, si f est strictement convexe et si C est convexe, alors le problème (P) admet une solution unique.

Condition d'optimalité

Tout comme dans le cas sans contraintes, nous allons établir des conditions d'optimalité du premier et du second ordre permettant de calculer les éventuels minima de f .

Conditions d'optimalité du premier ordre :

La condition nécessaire suivante est l'analogie de celle que nous avons dans le cas sans contraintes. Elle fait bien sûr intervenir l'ensemble des contraintes.

Théorème 11 : *(Condition nécessaire du premier ordre)*

Si f est une fonction différentiable et si C est un convexe fermé, alors toute solution x^* de (P) vérifie la condition nécessaire d'optimalité du premier ordre

$$\forall x \in C \quad \langle \nabla f(x^*), x - x^* \rangle \geq 0 \quad (1).$$

Théorème 12 : *(CNS du premier ordre dans le cas convexe)*

Supposons f convexe différentiable et C convexe fermé de R^n , soit x^* un élément de C . La condition (1) est nécessaire et suffisant pour que x^* soit solution de (P) . Elle caractérise donc les minima de f sur C .

On suppose maintenant que C est défini comme suit :

$$C = \{ x \in R^n : g_i(x) \leq 0, i = 1, \dots, p ; h_j(x) = 0, j = 1, \dots, q \}.$$

Les contraintes $g_i(x) \leq 0$ sont appelées "contraintes d'inégalités" et les contraintes $h_j(x) = 0$ "contraintes d'égalités".

L'obtention de conditions nécessaires d'optimalité nécessite que certaines conditions appelées conditions de qualification des contraintes soient vérifiées.

contrainte en égalité

problème P se réduit à

$\min f(x), x \in R^n, h(x) = 0$, avec $h(x) = (h_1(x), \dots, h_p(x))$ et h est continue de R^n dans R^p .

Théorème 13 : (CN du premier ordre-contraintes en égalité)

- f et h sont de classe C^1 sur R^n .
- Le problème P à une solution x^* .
- Les p vecteurs de R^n : $\nabla h_1(x^*), \dots, \nabla h_p(x^*)$ sont linéairement (et donc $p \leq n$);

Alors il existe p réels $(\lambda_1^*, \dots, \lambda_p^*)$ tels que :

$$\nabla f(x^*) + \sum_{j=1}^p \lambda_j^* \nabla h_j(x^*) = 0.$$

contrainte d'inégalité

L'ensemble de contraintes est donné par :

$$C = \{x \in R^n / g_i(x) \leq 0, \forall i = 1, \dots, m\}.$$

contraintes d'égalité et d'inégalité

Le problème P est de la forme $\min f(x), x \in C$.

$$C = \{x \in R^n / h(x) = 0, g(x) \leq 0\}.$$

Ici nous nous limiterons aux trois conditions de qualification suivantes :

★ (CQ1) C est un polyèdre convexe :

c'est le cas lorsque les fonctions g_i et h_j sont affine.

★ (CQ2) condition de Slater :

la condition de qualification des contraintes de Slater est comme suit :

1- Les fonctions g_i sont convexes et les fonctions h_j sont affine.

2- Il existe \hat{x} tel que $g_i(\hat{x}) < 0$ et $h_j(\hat{x}) = 0$ pour tout i, j .

★ (CQ3) condition de Mangasarian-Fromowitz :

la condition de qualification des contraintes de Mangasarian-Fromowitz en un point $x^* \in C$ est comme suit :

1) Les q vecteurs $\nabla h_j(x^*)$ sont linéairement indépendants.

2) Il existe \bar{d} tel que $\langle \nabla h_j(x^*), \bar{d} \rangle = 0$ pour tout j et $\langle \nabla g_i(x^*), \bar{d} \rangle < 0$ pour tout i .

On associe au problème d'optimisation suivant :

$$\begin{cases} \min f(x) \\ g_i(x) \leq 0, i = 1, \dots, p \\ h_j(x) = 0, j = 1, \dots, q \end{cases}$$

la fonction $L : \mathbb{R}^n \times [0, \infty[^p \times \mathbb{R}^q \longrightarrow \mathbb{R}$ définie par

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^p \lambda_i g_i(x) + \sum_{j=1}^q \mu_j h_j(x)$$

est appelée le " Lagrangien ".

Le théorème suivant de Karush-Kuhn-Tucker-Lagrange donne une condition nécessaire d'optimalité.

Théorème 14 :

Supposons que les fonctions f, g_i, h_j sont C^1 dans un voisinage de $x^ \in C$ et que les contraintes vérifient une condition de qualification. Si f a un minimum local en x^* sur C alors il existe $\lambda \in \mathbb{R}^p$ et $\mu \in \mathbb{R}^q$ tels que :*

$$\nabla_x L(x^*, \lambda, \mu) = 0, \nabla_\lambda L(x^*, \lambda, \mu) \leq 0, \nabla_\mu L(x^*, \lambda, \mu) = 0, \lambda \geq 0, \langle \lambda, \nabla_\lambda L(x^*, \lambda, \mu) \rangle = 0.$$

Les quantités λ_i et μ_j sont appelées "multiplicateurs de Karush-Kuhn-Tucker". Lorsque le problème est convexe (f et g_i convexes et h_j affine) on a la condition suffisante d'optimalité.

Théorème 15 :

Supposons que les fonctions f, g_i sont C^1 dans un voisinage de $x^* \in C$, convexes et que les fonctions h_j sont affine. S'il existe $\lambda \in \mathbb{R}^p$ et $\mu \in \mathbb{R}^q$ tels que

$$\nabla f(x^*) + \sum_{i=1}^p \lambda_i \nabla g_i(x^*) + \sum_{j=1}^q \mu_j \nabla h_j(x^*) = 0, \lambda_i \geq 0, g_i(x^*) \leq 0, \lambda_i g_i(x^*) = 0, h_j(x^*) = 0,$$

$\forall i, j$ alors f a un minimum global en x^* sur C .

Conditions d'optimalité nécessaires du deuxième ordre :

Théorème 16 :

on suppose que f, h et g sont de classe C^2 , que x^* est un minimum(local) de f sur C
Programmation linéaire

1.6 Classification des problème d'optimisation

Les problèmes d'optimisation sont classifiés selon les caractéristiques des fonctions f, g_i, h_j .

Parmi les cas particuliers les plus étudiés, on note :

- La programmation linéaire (f linéaire, g_i, h_j affine, C orthant positif).
- La programmation convexe (f, g_i convexe, h_j affine, C convexe).
- La programmation en nombres entiers (C discret).

1.7 Les méthodes des optimisations sans contraintes et avec contraintes

- Méthode du gradient projeté.
- Méthode de Newton.
- Méthode de pénalisation.
- Méthode de dualité : méthode d'Uzawa.

Chapitre 2

Programmation linéaire

2.1 Introduction

La programmation linéaire est certainement, l'un des plus importants développements dans le domaine de la recherche opérationnelle

Dans ce chapitre, nous allons parler les plus importants du programmation linéaire comme leur forme, la dualité.

2.2 Programmation linéaire

On définit un programme linéaire sous la forme :

$$(PL) \begin{cases} \min z = c.x \\ A.x = b \\ x \geq 0 \end{cases}$$

où

n =nombre de variable.

m =nombre de contraintes.

A =matrice réelle $m \times n$ (matrice des contraintes).

$c = (c_1, c_2, \dots, c_n)$ =vecteur-ligne des coûts.

$b = (b_1, b_2, \dots, b_n)^T$ =m-vecteur des seconds membres.

$z = c.x = \sum_{j=1}^n c_j.x_j$ fonction à minimiser (fonction objectif).

un programme linéaire est réalisable si l'ensemble S est non vide.

Définition 7 : L'ensemble $S = \{x \in C : g_i(x) \leq 0, h_j(x) = 0\}, i = 1, \dots, k, j = 1, \dots, m$ est appelé ensemble des solutions réalisables.

Définition 8 : Une solution réalisable qui minimise $z = c.x$ sur S est appelé solution optimale.

Définition 9 : Un programme linéaire (PL) réalisable est borné si l'objectif borné sur S .

2.3 Les Formes d'un programme linéaire

2.3.1 Forme standard

On dit qu'un programme linéaire est mis sous forme standard si toutes les contraintes (en dehors des contraintes de non-négativité) sont des égalités.

Le programme linéaire primale caractérise sa forme standard.

2.3.2 Forme canonique

La forme canonique d'un programme linéaire est représentée par :

$$\left\{ \begin{array}{l} \min z = c.x \\ Ax \leq b, (\geq b) \\ x \geq 0 \end{array} \right.$$

Remarque 4 : On peut remencer de la forme standard à la forme canonique par l'équivalence suivante :

$$Ax = b \iff Ax \leq b \text{ et } Ax \geq b$$

Remarque 5 : On utilisatant la variable d'écart, on peut se ramencer la forme standard comme suite :

$$Ax \leq b \Rightarrow Ax + y = b \text{ avec } y \geq 0$$

$$Ax \geq b \Rightarrow Ax - y = b \text{ avec } y \geq 0$$

2.4 Dualité d'un programme linéaire

Tout programme linéaire (P) standard associé un autre programme linéaire appelé dual (D) comme suite :

$$(P) \left\{ \begin{array}{l} \min z = c.x \\ Ax = b \\ x \geq 0 \end{array} \right.$$

$$(D) \left\{ \begin{array}{l} \max w = u.b \\ \text{sous les cantraites :} \\ x \geq 0 \end{array} \right.$$

Remarque 6 :

-La matrice des contraintes de (D) est la transposé de la matrice des contraintes de (P).

-Le vecteur des coûts de (P) est le vecteur des seconds membres de (D) et vice-versa .

-Le dual de dual est primal .

Exemple 1 : Soit le programme primal

$$\begin{cases} \min 2x_1 - 3x_2 \\ x_1 - x_2 \leq 1 \\ 2x_1 + 3x_2 \geq 4 \\ x_1 + x_2 = 3 \\ x_1 \geq 0, x_2 \geq 0 \end{cases}$$

Le dual de programme s'écrira :

$$\begin{cases} \max u_1 + 4u_2 + 3u_3 \\ u_1 + 2u_2 + u_3 \leq 2 \\ -u_1 + 3u_2 + u_3 = -3 \\ u_1 \leq 0, u_2 \geq 0, u_3 \geq 0 \end{cases}$$

Théorème 17 : (Dualité faible)

x est solution réalisable primal (PP) et u est une solution réalisable dual de(PD), alors :

$$b.u \leq c.x$$

Théorème 18 : (Dualité forte)

Si x^* est une solution optimale pour le programme primal (PP), alors le programme dual(PD) a une solution optimale u^* , telle que :

$$b.u^* = c.x^*$$

Chapitre 3

Les méthodes recherche linéaire

3.1 Introduction

Dans ce chapitre on va consacrer à l'optimisation unidimensionnelle sans contraintes, c.-à-d, aux méthodes numériques pour résoudre le problème de type :

$$f(x) \rightarrow \min / x \in \mathbb{R} \quad (R)$$

f étant une fonction au moins continue sur l'axe ; D'habitude, on appelle ces méthodes recherche linéaire. Notre intérêt pour la recherche linéaire ne vient pas seulement du fait que dans les applications on rencontre, naturellement, des problèmes unidimensionnels. La recherche linéaire est un composant fondamentale de toutes les méthodes traditionnelles d'optimisation multi-dimensionnelle, sa technique de recherche linéaire est un brick de base fondamentale de toute méthode multi-dimensionnelle. Il ya, principalement deux méthodes de recherche linéaire savoir :

- Méthode de recherche linéaire d'ordre zéro :

Recherche fibonacci

Recherche d'or

- Méthode d'approximation de courbes :

Méthode de Newton

3.2 Recherche linéaire d'ordre zéro

Nous commençons par la recherche linéaire d'ordre zéro, c.-à-d, par des méthodes qui utilisant des valeurs de f seulement, pas ces dérivées. Les méthodes que nous sommes sur le point de développer résolvent pas le problème $f(x) \longrightarrow \min | x \in R$, tel qu'il est, mais le problème

$$f(x) \longrightarrow \min | a \leq x \leq b$$

de minimisation de l'objectif sur un segment fini donné $[a, b]$ ($-\infty < a < b < \infty$) pour assurer que le problème soit bien conditionné, nous faisons l'hypothèse suivante :

f est unimodale sur $[a, b]$, c.-à-d, possède un minimum local unique x^* sur le segment. cette hypothèse, comme on le voit facilement, implique que f strictement décroissante sur $[a, b]$ à gauche de x^* :

$$a \leq x' \leq x'' \leq x^* \Rightarrow f(x') > f(x'') \quad (2)$$

et est strictement croissante sur $[a, b]$ à droite de x^* :

$$x^* \leq x' < x'' \leq b \Rightarrow f(x') < f(x'') \quad (3)$$

En effet, si (2) étaient faux, il existerait x' et x'' tels que

$$a \leq x' < x'' \leq x^*, f(x') \leq f(x'')$$

Il suit que l'ensemble de minimiseurs de f sur $[a, x'']$ contient un minimiseur, x_* qui est différent de x^* . Comme x_* est un minimiseur de f sur $[a, x'']$ et x_* diffère de x'' , x_* est un minimiseur local de f sur $[a, b]$, alors qu'on a supposé que le minimiseur local unique de f sur $[a, b]$ est x^* ; Ceci donne la contradiction désirée. On a (3) de façon analogue.

Notez que les relations (2) et (3), à leur tour, impliquent que f est unimodal sur $[a, b]$ et même sur chaque segment $[a', b'] \subset [a, b]$ plus petit.

Etant donné que f est unimodal sur $[a, b]$, nous pouvons préciser une stratégie pour approcher x^* : choisissons deux points x^- et x^+ dans (a, b) ,

$a < x^- < x^+ < b$, et calculons les valeurs $f(x^-)$ et $f(x^+)$. On observe que si :

- Cas A : $f(x^-) \leq f(x^+)$, alors x^* se trouve à gauche de x^+ [en effet, si x^* était à droite de x^+ , on aurait $f(x^-) > f(x^+)$ d'après (2)].

- Cas B : $f(x^-) > f(x^+)$, x^* est alors à droite de x^- [raisonnement "symétrique"].

En conséquence, dans le cas A nous pouvons remplacer le “segment d’incertitude” initial $\Delta_0 = [a, b]$ par le nouveau segment d’incertitude

$\Delta_1 = [a, x+]$, et dans le cas B par le segment $\Delta_1 = [x-, b]$; Dans les deux cas le nouveau “segment d’incertitude” Δ_1 couvre x^* et est strictement plus petit que Δ_0 .

Puisque, l’objectif, étant unimodal sur le segment initial $\Delta_0 = [a, b]$, est unimodal également sur le segment plus petit $\Delta_1 \subset \Delta_0$, nous pouvons réitérer ce procédé choisir deux points dans Δ_1 , calculer les valeurs de l’objectif en ces points, comparez les résultats et remplacez Δ_0 par un plus petit segment Δ_2 , contenant la solution désirée x^* , et ainsi de suite.

3.2.1 Recherche de Fibonacci

La recherche de Fibonacci peut être employée quand nous savons à l’avance le nombre $N > 2$ d’évaluations de fonction que nous allons exécuter.

Etant donné N , on considère la suite des $N + 1$ premiers nombres entiers de Fibonacci $F_0, F_1, F_2, \dots, F_n$ définis par la récurrence

$$F_0 = F_1 = 1; F_k = F_{k-1} + F_{k-2}$$

La méthode que nous allons utiliser est suivante : étant donné

$\Delta_0 = [a, b]$, on pose

$d_0 = |a - b|$, on choisit les deux premiers points x_1^- et x_1^+ de recherche à la distance

$$d_1 = \frac{F_{N-1}}{F_N} d_0$$

de l’extrémité droite et de l’extrémité gauche de Δ_0 respectivement

(comme $F_N/F_{N-1} = (F_{N-1} + F_{N-2})/F_{N-1} = 1 + F_{N-2}/F_{N-1} < 2$, nous avons $d_1 > d_0/2$, de sorte que $x_1^- < x_1^+$).

La longueur du nouveau segment Δ_1 d’incertitude est alors d_1 .

En suite on réitère l’étape ci-dessus, avec N remplacé $N - 1$. Ainsi, maintenant nous devrions évaluer f en deux points x_1^-, x_1^+ du segment Δ_1 placés à la distance

$$d_2 = \frac{F_{N-2}}{F_{N-1}} d_1 \left[= \frac{F_{N-2} F_{N-1}}{F_{N-1} F_N} d_0 = \frac{F_{N-2}}{F_N} d_0 \right] \quad (4)$$

des bouts droit et gauche de Δ_1 . Le fait crucial (qui résulte des propriétés arithmétiques des nombres de Fibonacci) est que un de ces deux points où f devrait être calculé est déjà traité – celui parmi les deux points précédents qui appartient à l'intérieur de Δ_1 .

En effet, supposons, sans perte de généralité, que $\Delta_1 = [a, x_1^+]$

(le cas $\Delta_1 = [x_1^-, b]$ est complètement analogue), de sorte que $x_1^- \in \text{int } \Delta_1$.

Nous avons $x_1^- - a = (b - d_1) - a = (b - a) - d_1 = d_0 - d_1 = d_0 \left(1 - \frac{F_{N-1}}{F_N}\right) = d_0 \frac{F_{N-2}}{F_N} = d_2$
[comme $F_N = F_{N-1} + F_{N-2}$ et $d_2 = \frac{F_{N-2}}{F_N} d_0$]

Ainsi, seulement un des deux points exigés de Δ_1 est réellement “nouveau”, et l'autre vient de l'étape précédente ; Par conséquent, afin de mettre à jour Δ_1 vers Δ_2 nous avons besoin d'une seule évaluation de fonction. Après cette nouvelle évaluation de fonction, nous pouvons remplacer Δ_1 vers Δ_2 . Pour traiter Δ_2 , nous agissons exactement comme ci-dessus, mais avec N remplace 2 par $N - 2$; Ici nous devons évaluer f aux deux points de Δ_2 à la distance

$d_3 = \frac{F_{N-3}}{F_{N-2}} d_2 \left[= \frac{F_{N-3}}{F_N} d_0, \text{ see (4)} \right]$ des extrémités du segment, et à nouveau, un de ces point est déjà traité.

Au bout des itérations nous venons au segment Δ_{N-1} qui couvre x^* ; La longueur du segment est

$$d_{N-1} = \frac{F_1}{F_N} d_0 = \frac{b-a}{F_N},$$

et le nombre total d'évaluations de f requis pour obtenir ce segment est N (nous avons besoin de 2 évaluations de f pour passer de Δ_0 vers Δ_1 , et chacune des $N - 1$ mises à jour suivantes $\Delta_t \mapsto \Delta_{t+1}$ nécessite une évaluation de f).

Si on prend comme approximation de x^* n'importe quel point x^N du segment Δ_{N-1} , nous avons

$$|x^N - x^*| \leq |\Delta_N| = \frac{b-a}{F_N} \tag{5}$$

Pour comparer (5) avec l'évaluation de précision (8) de notre méthode initiale – peu sophistiquée – notez que

$$F_t = 1/\lambda + 2 [(\lambda + 1)\lambda t + (-1)t\lambda - t], \lambda = 1 + \sqrt{5}/2 > 1 \tag{6}$$

En consequence, de (6) nous obtenons

$$|x^N - x^*| \leq (\lambda + 2/\lambda + 1)\lambda^{-N} |b - a| (1 + o(1)) \quad (7)$$

où on note $o(1)$ une fonction de N qui converge vers 0 quand $N \rightarrow \infty$).

Nous voyons que le taux de convergence pour la recherche de Fibonacci est

$$\lambda^{-1} = 2/1\sqrt{5} = 0.61803\dots$$

qui est bien meilleur que le taux $\sqrt{2/3} = 0.81649\dots$ donné par

$$|x^* - x^k| \leq \left(\frac{2}{3}\right)^{\lfloor k/2 \rfloor} |b - a| \quad (8)$$

On peut montrer que la recherche de Fibonacci est une méthode optimale (dans un certain sens précis) d'ordre zéro, en termes de précision garantie après N évaluations de fonction. Malgré ces bonnes propriétés théoriques, la méthode n'est pas très commode du point de vue pratique : nous devrions choisir à l'avance le nombre d'évaluations de fonction à exécuter (c.-à-d., pour ajuster la méthode à une certaine précision, choisie à l'avance), ce qui est parfois assez désagréable. La méthode de recherche d'or que nous sommes sur le point de présenter est exempte de cette imperfection et, en même temps, pour des N pas trop petits, aussi efficace que la recherche de Fibonacci originale.

L'idée de la méthode de recherche d'or est très simple : à l'étape K de recherche de la recherche de Fibonacci à N pas, nous choisissons deux points de recherche dans le segment $K - 1$, et chacun de ces points divise le segment (entre l'extrémité plus proche et la plus éloignée) en rapport

$$[1 - F_{N-K}/F_{N-K+1}] : [F_{N-K}/F_{N-K+1}].$$

3.2.2 Recherche d'or

Soit $\lambda = (1 + \sqrt{5})/2$ (aussi appelé le “nombre d'or”), dans l'implémentation de recherche d'or nous choisissons à chaque étape les points de recherche x_t^- et x_t^+ pour diviser le segment précédent de l'incertitude $\Delta_{t-1} = [a_{t-1}, b_{t-1}]$ dans le rapport $1/\lambda$;

$$x_t^- = \frac{\lambda}{1+\lambda}a_{t-1} + \frac{1}{1+\lambda}b_{t-1}; x_t^+ = \frac{1}{1+\lambda}a_{t-1} + \frac{\lambda}{1+\lambda}b_{t-1} \quad (9)$$

On voit facilement que pour $t \geq 2$, un des points de recherche exigés pour mettre à jour Δ_{t-1} vers Δ_t est déjà traité en cours de la mise à jour de Δ_{t-2} vers Δ_{t-1} . Pour le vérifier, il suffit de considérer le cas quand $\Delta_{t-1} = [\alpha, \beta]$ et $\Delta_{t-1} = [\alpha, x_{t-1}^+]$ (le cas “symétrique”

$\Delta_{t-1} = [x_{t-1}^-, \beta]$ est complètement analogue). Notons $d = \beta - \alpha$, nous avons

$$x_{t-1}^- = \alpha + \frac{1}{1+\lambda}d, x_{t-1}^+ = \alpha + \frac{\lambda}{1+\lambda}d \quad (10)$$

Maintenant, nous sommes dans la situation $\Delta_{t-1} = [\alpha, x_{t-1}^+]$ de sorte que le second des deux points de recherche requis pour mettre à jour Δ_{t-1} vers Δ_t soit

$$x_{t-1}^+ = \alpha + \frac{\lambda}{1+\lambda} (x_{t-1}^+ - \alpha) = \alpha + \frac{\lambda^2}{(1+\lambda)^2}d$$

(voyez la deuxième égalité dans (10)). La dernière quantité, dues à la première égalité dans (10) et à l'équation caractéristique $\lambda^2 = 1 + \lambda$ qui donne λ , n'est rien d'autre que

x_{t-1}^- :

$$\lambda^2 = 1 + \lambda \Leftrightarrow \frac{1}{1+\lambda} = \frac{\lambda^2}{(1+\lambda)^2}$$

Ainsi, dans la recherche d'or chaque mise à jour $\Delta_{t-1} \mapsto \Delta_t$, excepté la toute première, exige une évaluation de fonction. La longueur du segment d'incertitude est réduite par chaque mise à jour par le facteur

$$\frac{\lambda}{1+\lambda} = \frac{1}{\lambda}.$$

c.-à-d.,

$$|\Delta_t| = \lambda^{-t} (b - a)$$

Après $N \geq 2$ évaluations de fonction (après $t = N - 1$ étapes de recherche d'or) nous pouvons approcher X^* par le point x^N du segment Δ_{N-1} , est l'imprécision sera bornée par :

$$|x^N - x^*| \leq |\Delta_{N-1}| \leq \lambda^{1-N} (b - a) \quad (11)$$

Ainsi, nous observons une convergence linéaire avec le même taux $\lambda - 1 = 0.61803...$ que pour la recherche de Fibonacci, mais maintenant la méthode est "stationnaire" – nous pouvons exécuter autant de pas que nous le souhaitons.

3.3 Méthode d'approximation de courbes

Les méthodes de recherche linéaire considérées jusqu'ici possèdent, sous l'hypothèse d'unimodalité, l'excellente propriété de convergence linéaire globale. Pouvons-nous espérer quelque chose de mieux ?

Naturellement, oui : on aimerait bien avoir une méthode de convergence superlinéaire. Si l'objectif se comporte "bien", autrement dit, est assez régulier, nous avons de bonnes chances d'accélérer la convergence, au moins sur la phase finale, en utilisant l'approximation de courbe, c.-à-d., en approchant l'objectif par une fonction simple dont le minimum peut être trouvé de façon explicite. Par exemple, on peut approcher f par un polynôme, en choisissant les coefficients du polynôme afin de l'adapter aux valeurs observées (et à celles des dérivées, si elles sont disponibles) de f en des itérations "les plus prometteuses". Une itération d'un algorithme "pur" d'approximation de courbe est suivante :

- Au début de l'itération, nous avons un certain ensemble de "points de travail" où nous avons déjà calculé les valeurs et, probablement, certaines dérivées de l'objectif avec ces données, nous calculons le polynôme d'approximation courant p qui devrait avoir les mêmes valeurs et les mêmes dérivées aux points de travail que ceux de l'objectif .

- Après avoir calculé le polynôme p , nous trouvons analytiquement son minimiseur et le prenons comme le nouveau point de recherche .

- Nous calculons la valeur (probablement, les dérivées) de l'objectif en ce point de recherche et mettons à jour l'ensemble de points de travail, en ajoutant le dernier point de recherche (ainsi que l'information sur l'objectif en ce point) et en excluant de cet ensemble le "plus mauvais" des points de travail précédents et on boucle. L'idée sous-jacente est très simple : si nous sommes capables d'obtenir la convergence de cette méthode, les points de travail seront éventuellement à une petite distance du minimiseur de f .

Si f est assez lisse, l'erreur qu'on commet en approchant f par p dans le d -voisinage des points de travail sera de l'ordre de d^{n+1} , q étant le degré de p , et l'erreur de l'approximation de f' par p' sera de l'ordre de d^q . En conséquence, nous pouvons espérer que la distance entre le minimiseur de p (c.-à-d., le zéro de p') et le minimiseur de f (le zéro de f') sera

de l'ordre de d^p ,

Naturellement, ce qui est dit n'est rien de plus qu'une idée très approximative. Voyons une réalisation standard de cette idée.

3.3.1 Méthode de newton

Supposons que nous résolvons le problème (R) avec l'objectif f deux fois continûment différentiable, et que étant donné x , nous pouvons calculer $f(x)$, $f'(x)$ et $f''(x)$. Sous ces hypothèses nous pouvons appliquer au problème la Méthode suivante de Newton :

Algorithme[Méthode de newton unidimensionnelle]

Initialisation : choisir le point initial x_0

Etape t : étant donné l'itération précédente x_{t-1}

– Calculer $f(x_{t-1})$, $f'(x_{t-1})$ et $f''(x_{t-1})$ et approcher f autour de x_{t-1} par son développement de Taylor du second ordre :

$$p(x) = f(x_{t-1}) + f'(x_{t-1})(x - x_{t-1}) + 1/2f''(x_{t-1})(x - x_{t-1})^2$$

– Choisir comme x_t le minimiseur de la fonction quadratique :

$$x_t = x_{t-1} - f'(x_{t-1})/f''(x_{t-1}), \text{ remplacer } t \text{ avec } t + 1 \text{ et boucler.}$$

La méthode de Newton, si initialisée près d'un minimiseur local non-dégénéré x^* de f (c.-à-d., près d'un point x^* satisfaisant la condition suffisante d'optimalité du second ordre : $f'(x^*) = 0$, $f''(x^*) > 0$) converge vers x^* quadratiquement.

Proposition 1 : [Convergence quadratique locale de la Méthode de Newton]

Soit $x^* \in R$ un minimiseur local non-dégénéré de la fonction régulière f , c.-à-d., un point tels que f est trois fois continûment différentiable dans un voisinage de x^* avec $f'(x^*) = 0$, $f''(x^*) > 0$. Alors les itération de Newton convergent vers x^* quadratiquement, à condition que le point de départ x_0 soit assez proche de x^* .

Conclusion

L'intérêt de recherche linéaire est fait dans les applications recontre, naturellement, les problèmes undimensionnelles, elle est composant fandemental de tout les méthode traditionnelle d'optimisation sans contrainte. Donc on conclure que la recherche linéaire donne la solution de tous les problèmes d'optimisation est presque exacte puisque l'erreur d'approximation est très petit.

Bibliographie

- [1] M. AZI, Cours optimisation, Centre universitaire de mila, 2013-2014.
- [2] M. Minoux, Programmation mathématique : Théorie et algorithmes, T1, Dunod, 1983.
- [3] M. Bierlaire, Introduction à l'optimisation différentielle : presses polytechniques et universitaires ronrandes, Paris, France, 2006.
- [4] J. Frédéric Bonnans, Jean Charles Gilbert, Claude Lemaréchal, Claudia Sagastizábal : "Méthodes Numériques d'optimisation", Institut National de Recherche en Informatique et Automatique Projet Promath, Rocquencourt 78153 Le Chesnay, France, (1995).
- [5] A. Keraghel, Programmation mathématique différentiable, Université Ferhat Abbas, Sétif Algérie.
- [6] L .Assma, "procédure améliorante d'une méthode projective en programmation linéaire", mémoire de magister, Université Hadaik-Skikda, Algérie, (2007).