Méthodes d'optimisation

Thierry BAY - Philippe Delattre

Ressource R4.04

March 24, 2023

Introduction

- Problème d'optimisation = réduire un coût, maximiser un profit.
- Terme mathématique = min/max. d'une **fonction objectif**.
- Cette fonction :
  - Comporte des **paramètres** ;
  - Est généralement soumise à des **contraintes**.
- Grande variabilité de problèmes, au niveau :
  - Du domaine d'étude : continu, discret ;
  - Du nombre de variables ;
  - Du type de contraintes.
- "Data analysis" vs "business decisions".

Exemples de problèmes

1. **Emploi du temps -** Une université doit effectuer un certain nombre de cours à un certain nombre d'élèves, avec un certain nombre d'enseignants-chercheurs. Contraintes de précédences de cours, salles, désidératas.
   Objectif : trouver un planning cohérent.

2. **Flux de transport -** Une coopérative effectue quotidiennement le ramassage de lait.
   Objectif : effectuer la tournée la plus courte.

3. **Conception de pièces -** Une entreprise ferroviaire renouvelle son parc pour être plus éco-responsable.
   Objectif : concevoir une nouvelle forme de wagon pour réduire l'impact de la prise au vent.

D'importantes considérations

Classification

- Change d'un auteur à l'autre...
- Continuité, linéarité, bornes de la fonction de coût ?
- Variables discrètes ? Continues ?
- Mono-objectif ? Multi-objectifs ?
- Pas de contraintes ? Contraintes linéaires/quadratiques/autres ?
- Connaissances stochastiques ?
- Gradient ? Sans gradient ?

Ces caractéristiques influent sur

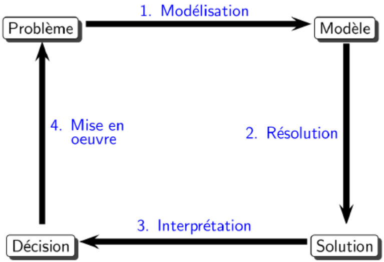
- La modélisation : puissance expressive nécessaire.
- Les techniques de résolution.
- La difficulté : efficacité des méthodes/algorithmes.

Méthodes de résolution

Nous verrons deux approches parmi "beaucoup" :

1. Approche pour l'optimisation convexe :
   - Pourquoi ? Minimum locale = minimum global.
   - Un exemple : la "programmation linéaire".
2. Approche à base de gradients :
   - Pourquoi ? Permet d'améliorer rapidement la solution lorsque l'on peut calculer la "dérivée".
   - Un exemple : la descente de gradient.
3. Méthodes efficaces, qui présentent des limitations. Donc si on applique bêtement un algorithme, il faut s'attendre à ce que ça rate !

Modélisation et choix de la méthode



Une approche en quatre étapes

1. Analyse du problème.
2. Modélisation et choix de méthode.
3. Résolution.
4. Interprétation des résultats.

Programmation linéaire et simplexe

Vocabulaire de base

Problème 1 - Maximisation

Pour une fonction $f : \mathbb{R}^n \to \mathbb{R}$ donnée, et un ensemble $M \subseteq \mathbb{R}^n$, trouver $\hat{x} \in M$ qui maximise f sur M tel que : $f(\hat{x}) \geq f(x), \quad \forall x \in M$.

Terminologie :

- f : **fonction-objectif**.
- M : **espace admissible**.
- $\hat{x}$: **maximiseur de f sur M**.
- Tout $x \in M$ est **solution réalisable ou admissible**.
- $\{x_i\}_{i=1}^n$: **variables d'optimisation ou de décision**.

Inéquations linéaires

Définition 1 - Inéquation linéaire

Une inéquation linéaire est une expression de la forme :

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n \leq b,$$

avec x_i les variables, a_i les cœfficients des variables, b une constante et n le nombre d'inconnues.

Définition 2 - Solution d'une inéquation linéaire

Solution de l'inéquation linéaire $a_1x_1 + a_2x_2 + \ldots + a_nx_n \leq b$: tout n-uplet $(y_1, \ldots, y_n)$ tel que l'inégalité $a_1y_1 + a_2y_2 + \ldots + a_ny_n \leq b$ est vraie.

Inéquations linéaires

Définition 3 - Système d'inéquations linéaires

On appelle système de m inéquations linéaires à n inconnues un système de la forme :

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n \leq b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2n}x_n \leq b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \ldots + a_{3n}x_n \leq b_3 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \ldots + a_{mn}x_n \leq b_m. \end{cases}$$

où x_j est une variable dans la colonne j, a_{ij} est le cœfficient de la variable x_j sur la ligne i, b_i est la constante de la ligne i, n est lenombre d'inconnues et m est le nombre d'inéquations.

De l'algèbre au programme linéaire

Définition 4 - Fonction linéaire

Une fonction $f : \mathbb{R}^n \to \mathbb{R}$ *est linéaire si et seulement si* $f(x+y) = f(x) + f(y)$ *et* $f(\lambda x) = \lambda f(x)$*, avec* $x, y \in \mathbb{R}^n$ *et* $\lambda \in \mathbb{R}$.

Exemple de fonctions linéaires :

- $f_1(x) = x$,
- $f_2(x) = 3x_1 - 5x_2$,
- $f_3(x) = Ax$.

Exemple de fonctions non-linéaires :

- $f_4(x) = x + 1$,
- $f_5(x) = x^2$,
- $f_6(x) = sin(x)$.

De l'algèbre au programme linéaire

Remarque 1

Toute fonction linéaire $f : \mathbb{R}^n \to \mathbb{R}$ peut s'exprimer dans la forme $f(x) = Ax$, avec $A \in \mathbb{R}^{m \times n}$ une matrice.

Fournissons quelques détails maintenant sur les notations :

- $x \in \mathbb{R}^n$: vecteur colonne $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ avec $x_1, x_2, \ldots, x_n \in \mathbb{R}$.
- x^T : transposée du vecteur x, vecteur ligne $(x_1, x_2, \ldots, x_n)$.
- $A \in \mathbb{R}^{m \times n}$ est la matrice $\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$.

De l'algèbre au programme linéaire

Définition 5 - Programme Linéaire

$$(PL)\begin{cases} \max c^T x = c_1x_1 + c_1x_1 + \ldots + c_nx_n = \sum_{i=1}^{n} c_ix_i \\ a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n \leq b_i, \quad i = 1, \ldots, m \\ x_i \geq 0, \quad i \in I \subseteq \{1, \ldots, m\} \end{cases}$$

où $x \in \mathbb{R}^n$ est le vecteur de variables inconnues.

- *$f(x) = c^T x$: **fonction-objectif/de coût/économique**.*
- *$c \in \mathbb{R}^n$: **vecteur de coût**.*
- *A et b : collectent les informations des **contraintes**.*

De l'algèbre au programme linéaire

Remarque 2 - Résolution du (PL)

- *Généralement, A n'est pas carrée ($m \neq n$). Donc ?*
- *Habituellement, A a plus de colonnes que de lignes...Hum.*
- *Un grand choix de solutions maximisant $c^T x$ existe dans l'espace admissible M.*

Remarque 3 - Maximisation et minimisation

L'opérateur de minimisation remplace parfois celui de maximisation. Le passage de l'un à l'autre requiert une simple manipulation : la maximisation de $c^T x$ devient la minimisation de $-c^T x$. Le résultat de la fonction objectif sera du coup $-f(x)$.

Exemple

Un agriculteur possède un nombre d'hectares (T), d'engrais (E) et d'insecticides (I), de maïs (x_1) ou de blé (x_2). Les cultures ont des quantités différentes d'engrais et d'insecticide (E_1, I_1, E_2, I_2), et fournissent un revenu différent (S_1, S_2). Modélisation de la situation :

Maximiser :	$S_1x_1 + S_2x_2$	: fonction objectif
Soumis à :	$x_1 + x_2 \leq T$	: limite du terrain
	$E_1x_1 + E_2x_2 \leq E$	: limite pour l'engrais
	$I_1x_1 + I_2x_2 \leq I$	: limite pour l'insecticide
	$x_1 \geq 0,\ x_2 \geq 0$	: terrain positif

Généralités sur les fonctions de $\mathbb{R}^2$

- Limitation au cas 2D (pour y voir quelque chose...).
- Un demi-plan est défini par ?*
- Idée simple : construire la zone-solution en obtenant l'intersection de tous les demi-plans, solutions des inéquations.

- Exemple :

$$\begin{aligned} \text{max} \quad & 100x_1 + 250x_2 \\ \text{sujet à :} \quad & x_1 + x_2 \leq 40, \\ & 40x_1 + 120x_2 \leq 2400, \\ & 6x_1 + 12x_2 \leq 312, \\ & x_1, x_2 \geq 0. \end{aligned}$$

Généralités sur les fonctions de $\mathbb{R}^2$

- Chaque contrainte : une équation de droite.
- Équation générale d'une droite dans $\mathbb{R}^2$:

$$h(x, y) = ax + by + c = 0.$$

- Vecteur normal à la droite, noté n_h :

$$n_h = \begin{pmatrix} a \\ b \end{pmatrix}.$$

- Orienté dans la direction des valeurs croissantes de h.
- Dans l'autre sens : tourné vers les valeurs décroissantes de h.
- Utilité ? Savoir quel demi-plan il nous faut récupérer avec les inéquations.

Maintenant, l'espace de solutions !

- Construction par intersection de <u>tous</u> les *demi-espaces*, donc avec <u>toutes</u> les contraintes.
- Intersection de tous ces demi-espaces : un **polyèdre convexe**.

Définition 6 - Ensemble convexe

Un ensemble M est dit convexe si l'intégralité d'une ligne connectant deux points de M appartient à M.

Théorème 1 - Solution sur un polyèdre

Soit f une fonction linéaire définie sur un polyèdre convexe borné. Alors la fonction f atteint sa valeur maximale en ***au moins un des sommets*** *du polyèdre convexe.*

Programmation linéaire et simplexe

Exemple de construction

- Travaillons avec les 3 contraintes du problème :

$$\begin{cases} g_1(x_1, x_2) = x_1 + x_2 \\ g_2(x_1, x_2) = 40x_1 + 120x_2 \\ g_3(x_1, x_2) = 6x_1 + 12x_2. \end{cases}$$

- Nature des contraintes ?*

- $n_{g_1} =$

Exemple de construction

- Travaillons avec les 3 contraintes du problème :

$$\begin{cases} g_1(x_1, x_2) = x_1 + x_2 \\ g_2(x_1, x_2) = 40x_1 + 120x_2 \\ g_3(x_1, x_2) = 6x_1 + 12x_2. \end{cases}$$

- Nature des contraintes ?*
- $n_{g_1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.
- n_{g_1} pointe vers valeurs croissantes de g_1, ie $g(x_1, x_2) > 40$.
- Dans la direction opposée : tout point tel que $g(x_1, x_2) < 40$.
- Points dans l'espace réalisable qui nous intéresse ?*

Exemple de construction

- Même travail avec les autres contraintes, d'où 3 lignes au total :

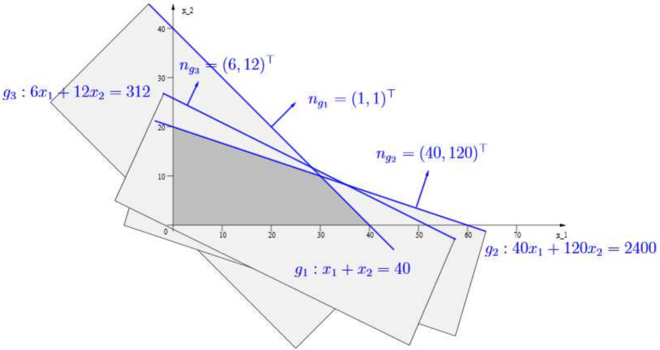
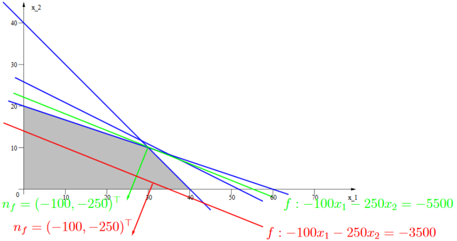


Illustration des contraintes avec les 3 fonctions g_1, g_2 et g_3.

C'est bien beau, mais la solution ?

- On a toutes les solutions possibles ! Reste à trouver le meilleur des points !
- **Objectif** : max $f(x_1, x_2) = 100x_1 + 250x_2$.
- Augmentation des valeurs de f dans la direction de la normale.
- La figure ci-dessous illustre deux déplacements de f :



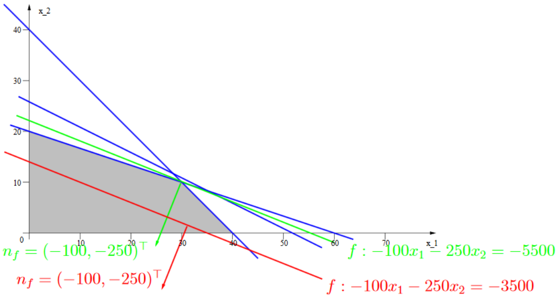
Déplacements de la fonction-objectif selon la normale (négative ici !).

Et donc ???

- Sa normale est :

$$n_f(x_1, x_2) = \begin{pmatrix} 100 \\ 250 \end{pmatrix}.$$

- Puisque nous devons maximiser f :
  - Déplacement de la droite de f dans la direction de n_f.
  - Où sont les optimums ?*



Bref, résumons...

<u>**Étapes à suivre pour la résolution graphique**</u> **:**

1. Tracer l'ensemble réalisable M_c à partir des contraintes.
2. Tracer la fonction-objectif. Dans la pratique, on la fait initialement passer par $0_{\mathbb{R}^n}$ (par $(0,0)^T$ dans $\mathbb{R}^2$).
3. Déplacer la ligne de la fonction-objectif dans la direction définie par le vecteur c pour une maximisation.
4. La solution optimale est l'intersection la plus extrême entre M_c et la ligne de la fonction-objectif.

Bidouillons le système initial...

- À partir du problème initial, on peut mettre le PL sous :

1. Sa forme **canonique** : que des inéquations <u>de même sens</u> !

$$\sum_{j=1}^{n} a_{ij}x_j \geq b_i \Rightarrow \sum_{j=1}^{n}(-a_{ij})x_j \leq -b_i$$

$$\sum_{j=1}^{n} a_{ij}x_j = b_i \Rightarrow b_i \leq \sum_{j=1}^{n} a_{ij}x_j \leq b_i$$

$$\Rightarrow \sum_{j=1}^{n} a_{ij}x_j \leq b_i \quad \text{et} \quad \sum_{j=1}^{n}(-a_{ij})x_j \leq -b_i$$

2. Sa forme **standard** : que des équations !
   - On utilisera des **variables d'écart**.
   - Elles compensent l'écart entre l'inégalité et l'égalité :

$$2x_1 - 3x_2 \leq 7 \Rightarrow 2x_1 - 3x_2 + e_1 = 7, \text{ avec } e_1 \geq 0.$$

Théorème fondamental

Théorème 2 - Théorème fondamental de prog. linéaire

Soit un (PLS), avec $M_s \neq \varnothing$. Alors :

- *Soit la fonction-objectif n'est pas bornée et il n'y a pas de solution optimale, soit le problème a une solution optimale et au moins un sommet de M_s est parmi ces solutions.*
- *Si M_s est borné, une solution optimale existe, et $x \in M_s$ est optimal si et seulement si il s'agit d'une combinaison convexe de sommets optimaux.*

Le simplexe en pratique

En principe :

1. Avec les contraintes d'un (PLC), nous avons un ensemble convexe et la **solution optimale**, si elle existe, est **l'un des sommets de l'ensemble**.
2. Pour la calculer, nous allons **partir d'un de ces sommets**.
3. Nous nous **déplaçons de sommet en sommet** le long des arêtes du polyèdre convexe jusqu'à trouver la solution optimale.

En pratique, du pivot de Gauss :

1. Au départ, seules les variables d'écart sont prises en compte.
2. On choisit la "variable entrante" qui contribue le plus.
3. On choisit la "variable sortante" qui contribue le moins.
4. On répète tant que l'un des coûts est positif.

Quelle sera la variable entrante ?

Remarque 4 - Premier critère de Dantzig : choix du pivot q

- *En principe, n'importe quelle composante de x (avec un coût positif).*
- *S'il n'existe pas de pivot q tel que $c_q > 0$, la solution optimale est trouvée et l'algorithme s'arrête.*
- *Dans le cas contraire, la stratégie la plus courante est de choisir le pivot q correspondant à **la composante la plus grande du vecteur de coûts**. Ce choix garantit la plus grande croissance de la fonction-objectif.*

Quelle sera la variable sortante ?

Remarque 5 - Second critère de Dantzig : choix du pivot p

Dans la pratique, grâce à l'application du pivot de Gauss, la stratégie pour choisir le pivot p (et donc la variable sortant de la base) est de prendre le minimum des rapports $\frac{b_i}{a_{iq}}$ pour $i = 1, \ldots, m$, avec q la colonne de la variable entrante.

Exemple

Résolvons le problème suivant :

$$
\begin{array}{ll}
\text{max} & f(x_1, x_2) = 6x_1 + 4x_2 \\
\text{Soumis à :} & 3x_1 + 9x_2 \leq 81 \\
& 4x_1 + 5x_2 \leq 55 \\
& 2x_1 + x_2 \leq 20 \\
& x_1, x_2 \geq 0
\end{array}
$$

Le problème est déjà écrit sous forme d'un (PLC). Sous sa forme standard, nous introduisons les variables d'écart afin d'obtenir des contraintes d'égalité :

$$
\begin{array}{ll}
\text{max} & f(x_1, x_2) = 6x_1 + 4x_2 \\
\text{Soumis à :} & 3x_1 + 9x_2 + e_1 = 81 \\
& 4x_1 + 5x_2 + e_2 = 55 \\
& 2x_1 + x_2 + e_3 = 20 \\
& x_1, x_2, e_1, e_2, e_3 \geq 0
\end{array}
$$

Exemple

Itération 1

Variables de la base			Variables hors-base		
e_1	e_2	e_3	x_1	x_2	b
1	0	0	3	9	81
0	1	0	4	5	55
0	0	1	2	1	20
0	0	0	6	4	$f = 0$

Variable entrante : $x_e^{(1)} = \max(6, 4) = x_1$. Variable sortante : $\min\left(\frac{81}{3}, \frac{55}{4}, \frac{20}{2}\right) \Rightarrow x_s^{(1)} = e_3$.

Exemple

Itération 2

Variables de la base			Variables hors-base		
e_1	e_2	x_1	e_3	x_2	b
1	0	0	-3/2	15/2	51
0	1	0	-2	3	15
0	0	1	1/2	1/2	10
0	0	0	-3	1	$f = 60$

Variable entrante : $x_e^{(2)} = x_2$. Variable sortante : $\min\left(\frac{51}{15/2}, \frac{15}{3}, \frac{10}{1/2}\right) \Rightarrow x_s^{(2)} = e_2$.

- Pivot de Gauss : intersection de la colonne de x_1 et de la ligne e_3.
- Élimination de tous les autres éléments de la colonne x_1.
- Pourquoi Gauss ? La colonne de b est solution à chaque itération.

Exemple

Itération 3

Variables de la base			Variables hors-base		
e_1	x_2	x_1	e_3	e_2	b
1	0	0	7/2	-5/2	27/2
0	1	0	-2/3	1/3	5
0	0	1	5/6	-1/6	15/2
0	0	0	-11/3	-1/3	$f = 65$

Tous les coûts réduits sont négatifs. Optimum atteint avec $f = 65$.

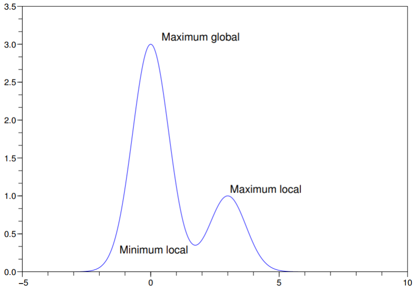
Solution optimale :

$$\begin{cases} x_1^* & = 15/2 \\ x_2^* & = 5 \end{cases}$$

Optimum : $f(x_1^*, x_2^*) = 6x_1^* + 4x_2^* = 6 * 15/2 + 4 * 5 = 65.$

Algorithmes de descente

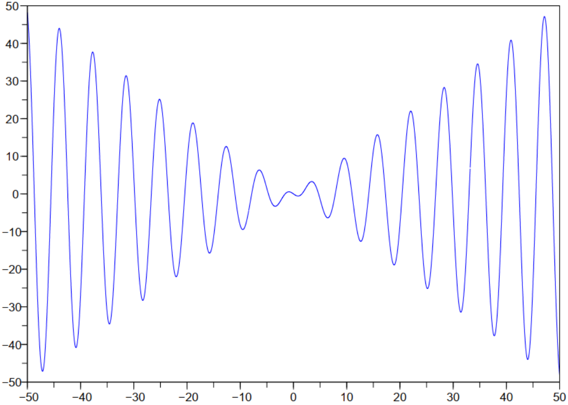
Vocabulaire



Exemple de minima et maxima locaux et globaux de $f(x) = 3e^{-x^2} + e^{-(x-3)^2}$.

- Différences avec les cas précédemment présentés ?
- **Par la suite : optimisation numérique sans contraintes !**

Vocabulaire



Infinité de minima et maxima locaux, sans extrama globaux avec $f(x) = xcos(x)$.

Vocabulaire

Définition 7 - Ensemble convexe

Soit $E \subset \mathbb{R}^n$. L'ensemble E est convexe ssi :

$$\forall (x, y) \in E^2,\ \forall \lambda \in]0, 1[,\ \lambda x + (1 - \lambda) y \in E.$$

En d'autres termes : si y et y sont deux éléments de E, le segment qui relie x à y est inclus dans E.

Définition 8 - Fonction convexe

Soit $E \subset \mathbb{R}^n$ et $f : E \to \mathbb{R}$. La fonction f est convexe ssi :

$$\forall (x, y) \in E^2,\ \forall \lambda \in]0, 1[,\ f(\lambda x + (1 - \lambda) y) \leq \lambda f(x) + (1 - \lambda) f(y).$$

Théorème 3 - Condition suffisante d'optimalité globale

Soit $E \subset \mathbb{R}^n$ un convexe et $f : E \to \mathbb{R}$ une fonctionnelle. Soit x^* un point minimum local de f. Alors :

- *Si f est convexe, alors x^* est un minimum global de f.*
- *Si f est strictement convexe, alors x^* est l'unique point minimum global de f.*

Mais le monde n'est pas uniquement convexe... :(

Objectif 1

Comprendre les méthodes numériques pour la recherche de points $x \in \mathbb{R}^n$ qui réalisent le minimum d'une fonction $f : \mathbb{R}^n \to \mathbb{R}$.

Théorème 4 - CN d'optimalité locale d'ordre 1

Soit $f : \mathbb{R}^n \to \mathbb{R}$ une application différentiable. Si $x^* \in \mathbb{R}^n$ réalise un minimum local (resp. maximum local) de f, alors :

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

Et en français ? Les zéros du gradient (de la dérivée) sont tous les points candidats en tant qu'extrema locaux potentiels. On les appelle **points critiques** ou **points stationnaires**. Parmi eux : minima locaux, maxima locaux, **points selle**.

Définition 9 - Fonction différentiable et dérivable

- *Une fonction $f : I \to \mathbb{R}$ est dérivable en $a \in I$ si le taux d'accroissement $\frac{f(x)-f(a)}{x-a}$ admet une limite quand x tend vers a. La limite est alors notée $f'(a)$.*
- *La notion de fonction différentiable est la généralisation aux fonctions de plusieurs variables.*

Définition 10 - Point critique

Soit $f : \mathbb{R}^n \to \mathbb{R}$ une application différentiable. Tout point $x \in \mathbb{R}^n$ vérifiant :

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

est appelé point ***critique*** *ou point* ***stationnaire*** *de f.*

Algorithmes de descente

Généralité

Objectif 2

À partir d'un point x_0 arbitraire, un algorithme de descente va chercher à générer une suite d'itérés $(x_k)_{k\in\mathbb{N}}$ telle que :

$$\forall k \in \mathbb{N}, \quad f(x_{k+1}) \leq f(x_k).$$

Qu'est-ce-que cette "descente" ?

Direction de descente

- Il y a beaucoup de directions de descente...
- Mais celle qui nous intéressera est celle où la pente est la plus forte : la direction du gradient !

Définition 11 - Direction de plus forte descente

Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction différentiable et $x \in \mathbb{R}^n$. La direction de plus forte descente est donnée par :

$$d^* = -\nabla f(x).$$

Algorithmes de descente

Algorithme de descente

Principe :

1. $k = 0$.
2. Tant que le test de convergence n'est pas satisfait (cf. plus loin),
   2.1 Trouver une direction de descente d_k telle que
   $$\nabla f(x_k)^T d_k < 0$$
   2.2 Choisir un pas $s_k > 0$ à faire dans cette direction, tel que :
   $$f(x_k + s_k d_k) < f(x_k).$$
   2.3 Mettre à jour : $x_{k+1} = x_k + s_k d_k$.
   2.4 $k = k + 1$.
3. Retourner x_k.

Critères d'arrêt

- Pour l'arrêt de l'algorithme :
  1. Critère d'optimalité : $\|\nabla f(x_k)\| < \varepsilon$.
  2. Stagnation de la solution : $\|x_{k+1} - x_k\| < \varepsilon \|x_k\|$.
  3. Stagnation de la valeur : $\|f(x_{k+1}) - f(x_k)\| < \varepsilon \|f(x_k)\|$.
  4. Nombre d'itérations dépassant un seuil fixé : $k < IterMax$.
- En pratique : (1) ou [(2) et (3)] ou (4).
- Notions importantes mises de côté :
  1. Convergence locale ? Globale ?
  2. Vitesse de convergence :

$$\lim_{k \to +\infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^p} = \tau, \text{ avec} \tau \geq 0.$$

  - Si $p = 1$ et $\tau \in [0, 1] \Rightarrow$ convergence linéaire.
  - Si $p = 1$ et $\tau = 0 \Rightarrow$ convergence superlinéaire.
  - Si $p > 1$ et $\tau \geq 0 \Rightarrow$ convergence d'ordre p.

Algorithmes de descente

Descente à pas fixe et optimal

Principe :

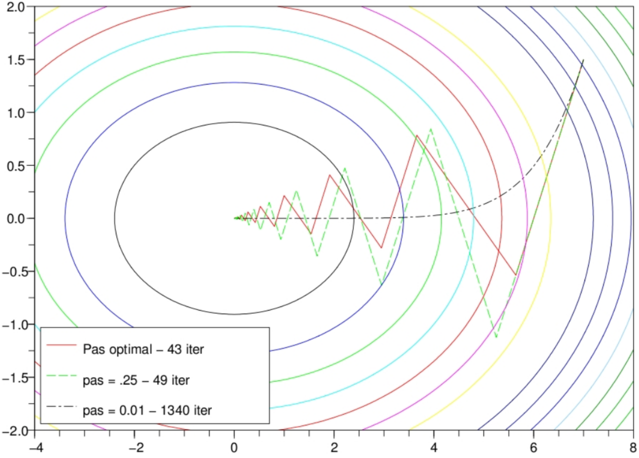
- On admet que la direction de plus profonde descente normalisée est :

$$d_k = -\frac{\nabla f(x_k)}{\|\nabla f(x_k)\|}.$$

- Il faut maintenant le "pas" le long de cette direction :
  - Choix d'un pas fixe : naïf, mais ça marche... ou pas :)
  - Choix d'un pas optimal $s_k > 0$, solution de

$$\min_{s>0} f(x_k + sd_k).$$

Descente à pas fixe et optimal



Comparaison entre pas fixe et pas optimal pour une fonction donnée.