

Chapitre II

Bases théoriques

Dans cette leçon, nous verrons les bases théoriques

- comment se formule le problème de l'optimisation, en particulier dans le cas continu.
- comment se caractérise

Rappel : la quasi-totalité des transparents de ce cours provient du cours d'optimisation de Max Cerf (Université Paris 6 / Ariane Espace).

1 Formulation

La formulation canonique est la suivante :

Formulation mathématique

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{sous} \quad \begin{cases} c_E(x) = 0 \\ c_I(x) \leq 0 \\ x \in X \end{cases}$$

→ formulation standard
problème noté (PO)

Notations

- x : n **variables** ou paramètres ou inconnues → vecteur de \mathbb{R}^n
- f : **critère** ou fonction **coût** ou fonction objectif → fonction de \mathbb{R}^n dans \mathbb{R}
 $x \in \mathbb{R}^n \mapsto f(x) \in \mathbb{R}$
- c_E : p **contraintes d'égalité** → fonction de \mathbb{R}^n dans \mathbb{R}^p
 $x \in \mathbb{R}^n \mapsto c_E(x) \in \mathbb{R}^p$
- c_I : q **contraintes d'inégalité** → fonction de \mathbb{R}^n dans \mathbb{R}^q
 $x \in \mathbb{R}^n \mapsto c_I(x) \in \mathbb{R}^q$
- X : ensemble convexe $X \subseteq \mathbb{R}^n$ → valeurs admissibles des variables

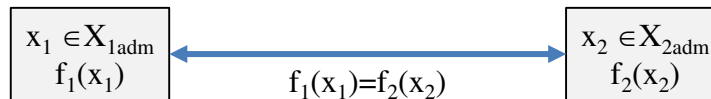
EXEMPLE : $\min_{(x,y,z) \in \mathbb{R}^3} x^2 + y^2 - 3z^3 + 3$ sous $\begin{cases} x + 2y = 3 \\ 2x - 3z \leq 4 \\ (x, y, z) \in [0, 2] \times [-3, 3] \times [1, 4] \end{cases}$

L'idée d'un problème équivalent est de transformer le problème ou les contraintes :

- pour retrouver la forme canonique d'un problème d'optimisation,
- pour transformer le problème en un problème équivalent.

Problèmes équivalents

(PO_1) et (PO_2) sont deux problèmes équivalents si on peut associer à tout point admissible x_1 de (PO_1) un point admissible x_2 de (PO_2) avec la même valeur pour le critère.



(PO_1) et (PO_2) ont alors des solutions de même coût : $f_1(x_1^*) = f_2(x_2^*)$

Transformations simples

- Changement de variable : $y = \varphi(x)$ avec φ strictement croissante sur X

- Maximisation / minimisation : $\max_x f(x) \Leftrightarrow \min_x (-f(x))$

- Contrainte inférieur / supérieur : $c(x) \geq 0 \Leftrightarrow -c(x) \leq 0$

- Variables positives : $x = x^+ - x^-$ avec $\begin{cases} x^+ \geq 0 \\ x^- \geq 0 \end{cases}$

- Variables d'écart : $c(x) \leq 0 \Leftrightarrow \begin{cases} c(x) + y = 0 \\ y \geq 0 \end{cases} \Leftrightarrow c(x) + z^2 = 0$

Afin d'adopter une première analyse mathématique du problème, on fera les hypothèses sont les suivantes :

Hypothèses

- Continuité : Fonctions continues de variables réelles
→ Optimisation continue
≠ Optimisation combinatoire, Programmation en nombres entiers
- Différentiabilité : Fonctions différentiables
→ **Méthodes à base de gradient**
≠ Méthodes sans dérivées
- Déterminisme : Les données du problème sont parfaitement connues
≠ Optimisation stochastique
- Programmation linéaire : coût linéaire et contraintes linéaires (LP)
- Programmation quadratique : coût quadratique et contraintes linéaires (QP)
- Programmation non linéaire : cas général, fonctions quelconques (**NLP**)

On reverra donc les notions de base de différentiabilités dans \mathbb{R}^n .

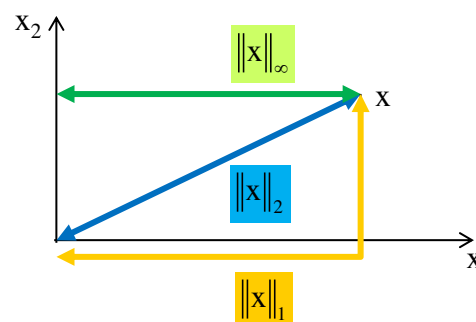
On minimise en général un critère, et ce critère consiste souvent en la minimisation d'une distance.

On rappelle la définition des différentes normes utilisées pour mesurer des distances :

Norme vectorielle sur \mathbb{R}^n

- Fonction $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$ vérifiant

$$\begin{cases} \|x\| \geq 0 \\ \|x\| = 0 \Leftrightarrow x = 0 \\ \|x + y\| \leq \|x\| + \|y\| \\ \|\alpha x\| = |\alpha| \|x\| \end{cases}$$
- Norme p : $\|x\|_p = \sqrt[p]{\sum_{i=1}^n |x_i|^p}$
- Norme ∞** : $\|x\|_\infty = \max_{i=1, \dots, n} |x_i|$
- Norme 2** = norme euclidienne



Norme matricielle

- Norme induite sur $\mathbb{R}^{m \times n}$ par la norme vectorielle $\|\cdot\|$
- Fonction $\|\cdot\|_{m \times n} : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$ définie par $\|A\|_{m \times n} = \max_{x \in \mathbb{R}^n, x \neq 0} \frac{\|Ax\|}{\|x\|}$

2 Caractérisation des solutions

2.1 Minima locaux et globaux

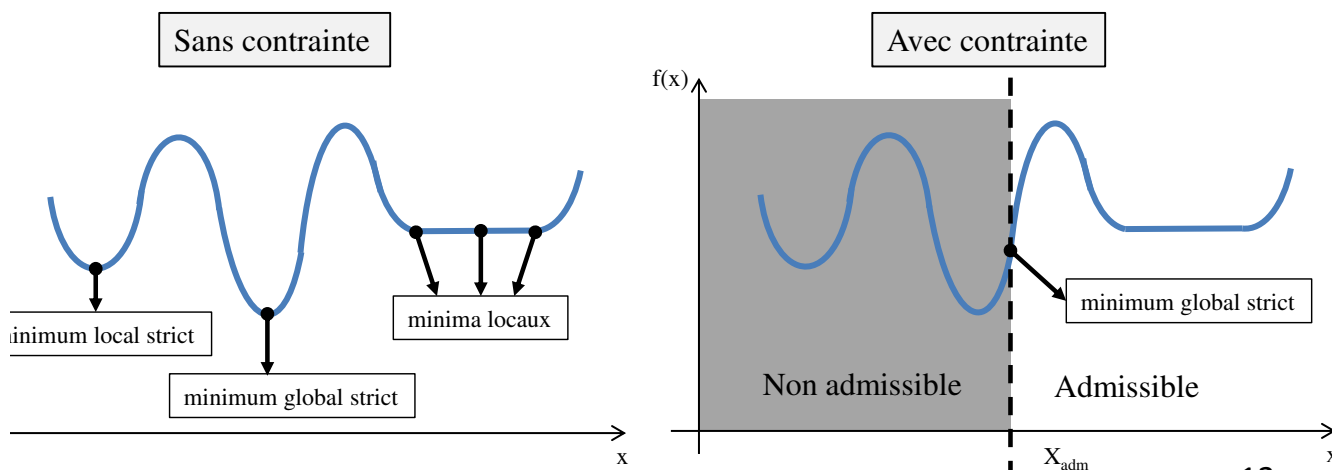
Comment caractériser un minimum local/global ?

Minimum global (= meilleure solution dans l'absolu)

- x^* minimum global de (PO) $\Leftrightarrow \forall x \in X_{\text{adm}}, f(x^*) \leq f(x)$
- x^* minimum global strict $\Leftrightarrow \forall x \in X_{\text{adm}}, f(x^*) < f(x)$ si $x \neq x^*$

Minimum local (= meilleure solution dans un voisinage)

- x^* minimum local de (PO) $\Leftrightarrow \exists \varepsilon > 0 / \forall x \in X_{\text{adm}}, \|x - x^*\| < \varepsilon, f(x^*) \leq f(x)$
- x^* minimum local strict $\Leftrightarrow \exists \varepsilon > 0 / \forall x \in X_{\text{adm}}, \|x - x^*\| < \varepsilon, f(x^*) < f(x)$ si $x \neq x^*$



2.2 Infimum

La notion d'infimum permet de définir le meilleur minimum sur un ensemble.

Borne inférieure

$$f : \mathbb{R}^n \rightarrow \mathbb{R} \quad f \text{ bornée inférieurement sur } Y \subseteq \mathbb{R}^n \\ \Leftrightarrow \exists M \in \mathbb{R} / \forall x \in Y, M \leq f(x)$$

Infimum

- Infimum de f sur Y = plus grande borne inférieure

- Notation : $\inf_Y f = \inf \{f(y), y \in Y\}$

- Propriété : $\begin{cases} \forall x \in Y, \inf_Y f \leq f(x) \\ \text{et} \\ \forall M > \inf_Y f, \exists x \in Y / f(x) < M \end{cases}$

Théorème de Weierstrass

- f atteint son infimum si f continue, Y compact : $\exists x^* \in Y / f(x^*) = \inf_Y f$
- Conditions réalisées en pratique : fonctions continues, intervalles fermés
→ Le problème (PO) admet une solution x^* .

3 Différentiabilité

On rappelle que, dans le cadre de l'optimisation continue, nous avons fait l'hypothèse de continuité sur la fonctionnelle à optimiser.

Pourquoi cette hypothèse ?

- soit on veut trouver la solution mathématiquement
alors il faut des critères indiquant que l'on est en train de passer par un minimum.
ceci peut se caractériser par des différentielles.
- soit on veut trouver la solution numériquement
alors il faut un moyen d'y aller, en suivant la "pente" de la fonctionnelle.
on rappelle que les différentielles permettent d'obtenir des informations locales sur la fonctionnelle
ce qui permet d'essayer de descendre vers le minimum.

Dans des deux cas, ceci implique une certaine forme de continuité, et de différentiabilité.

On va donc définir pour des fonctions de $\mathbb{R}^n \rightarrow \mathbb{R}$:

- les différentes différentielles (gradient, hessien), et l'intérêt direct qu'elles peuvent avoir,
- les formules de Taylor (développement limité) associées, qui permettent d'obtenir une approximation locale simplifiée de la fonctionnelle.

Par ailleurs, comme on travaille en général sur des données numériques (typiquement sur une grille d'échantillonnage), on verra comment estimer des dérivées numériques.

3.1 Dérivées partielles

Voyons les définitions des différentielles d'ordre 1 sur \mathbb{R}^n .

Différentiabilité ordre 1

f fonction continue de \mathbb{R}^n dans \mathbb{R}

Dérivée partielle

Dérivée partielle de f en x par rapport à x_i : $f_{x_i}(x) = \frac{\partial f(x)}{\partial x_i} = \lim_{s \rightarrow 0} \frac{f(x_1, \dots, x_i + s, \dots, x_n) - f(x_1, \dots, x_i, \dots, x_n)}{s}$
si la limite existe

Gradient

Gradient de f en x : $g(x) = \nabla f(x)$
 $g(x) : \mathbb{R}^n \rightarrow \mathbb{R}^n$
si toutes les dérivées partielles existent

$$g(x) = \left(\frac{\partial f(x)}{\partial x_i} \right)_{i=1, \dots, n} = \begin{pmatrix} \frac{\partial f(x)}{\partial x_1} \\ \dots \\ \frac{\partial f(x)}{\partial x_n} \end{pmatrix}$$

Dérivée directionnelle

Dérivée directionnelle de f en x dans la direction $d \in \mathbb{R}^n$:
si la limite existe
(dérivée directionnelle = produit scalaire avec le gradient) $\Rightarrow f_d(x) = g(x)^T d$

Fonction différentiable

f différentiable en $x \Leftrightarrow f$ admet une dérivée directionnelle pour tout $d \in \mathbb{R}^n$

Notes :

- remarquer que le gradient donne la valeur de la normale au point où il est calculé.
- la dérivée directionnelle est la projection de la normale dans la direction choisie.

EXERCICE 1: Dérivées partielles

Calculer $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ pour les fonctions suivantes :

1. $f(x, y) = x \cos(x) \sin(x)$
2. $f(x, y) = e^{xy^2}$
3. $f(x, y) = (x^2 + y^2) \ln(x^2 + y^2)$

EXERCICE 2: Calcul du gradient

Calculer le gradient des fonctions suivantes au point p indiqué :

1. $f(x, y) = \sqrt{a^2 - x^2 - y^2}^{1/2}$ en $p = (a/2, a/2)$.
2. $g(x, y) = \ln \sqrt{1 + xy}$ en $p = (1, 1)$.
3. $h(x, y) = e^y \cos(3x + y)$ en $p = (2\pi/3, 0)$.

EXERCICE 3: Calcul des dérivées directionnelles

Calculer les dérivées des fonctions suivantes au point p et dans la direction v :

1. $f(x, y) = x + 2xy - 3y^2$ en $p = (1, 2)$ et dans la direction $v = (3, 4)$.
2. $g(x, y) = e^{xy} + y$ en $p = (1, 1)$ et dans la direction $v = (+1, -1)$.
On rappelle que $(\arctan x)' = 1/(1 + x^2)$.
3. $h(x, y) = \sqrt{x^2 + y^2}$ en $p = (0, 5)$ et dans la direction $v = (+1, -1)$.

EXERCICE 4: Maximisation d'une dérivée directionnelle

Soit la fonction $f(x, y, z) = xy^2 + z^2y$.

1. Trouver la dérivée de f au point p et dans la direction v où $p = (1, 1, 0)$ et $v = (1, 1, 2)$.
2. Déterminer la direction qui maximise (resp. minimise) la dérivée directionnelle au point $(1, 1, 0)$.
3. Quelles est la plus grande et la plus petite valeur de la dérivée directionnelle en ce point ?

EXERCICE 5: Plan tangent

Calculer le plan tangent à la surface au point indiqué.

1. $x^2 + 2xy + 2y^2 - z = 0$ au point $(1, 1, 5)$.
2. $x^2 + y^2 - z = 0$ au point $(1, 2, 5)$.
3. $(y - x^2)(y - 2x^2) - z = 0$ au point $(1, 3, 2)$.

3.2 Hessian

Voyons les définitions des différentielles d'ordre 2 sur \mathbb{R}^n .

Différentiabilité ordre 2

f fonction deux fois différentiable de \mathbb{R}^n dans \mathbb{R}

Hessien

Hessien de f en x : $H(x) = \nabla^2 f(x)$

$H(x) : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$

$$H(x) = \left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right)_{i,j=1,\dots,n} = \begin{pmatrix} \frac{\partial^2 f(x)}{\partial^2 x_1} & \dots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 f(x)}{\partial^2 x_n} \end{pmatrix}$$

Courbure

On définit pour une direction $d \in \mathbb{R}^n$ au point x la fonction φ à une variable : $\varphi(s) = f(x+sd)$
 \rightarrow variation de f dans la direction d

- Dérivée première : $\varphi'(s) = d^T \nabla f(x+sd) \rightarrow \varphi'(0) = d^T \nabla f(x)$
- Dérivée seconde : $\varphi''(s) = d^T \nabla^2 f(x+sd) d \rightarrow \varphi''(0) = d^T H(x) d$

La **courbure de f en x dans la direction d** est définie par : $\frac{d^T H(x) d}{d^T d}$
 \rightarrow normalisation de φ'' en $s=0$
 $=$ quotient de Rayleigh de $H(x)$ dans la direction d

Notes : d'après le théorème de Schwarz, si f est deux fois dérivable en un point, alors son Hessian est symétrique en ce point.

- les propriétés du hessien permettent de connaître la convexité de la surface au point où il est calculé.

EXERCICE 6: Hessian

Calculer le Hessian des fonctions suivantes :

1. $Q(x, y, z) = x^2 + 5y^2 + 4xy - 2yz$.
2. $f(x, y) = 3x^2y + 4x^3y^4 - 7x^9y^4$.

3.3 Jacobien

Soit $c = (c_1, \dots, c_m)$ la fonction construite comme le vecteur contenant l'ensemble des m fonctions c_i , chacune $\mathbb{R}^n \rightarrow \mathbb{R}$. Noter que cette fonction est de \mathbb{R}^n dans \mathbb{R}^m .

Matrice gradient

c fonction continue de \mathbb{R}^n dans \mathbb{R}^m

Gradient de c en x : $\nabla c(x) : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m}$

$$\nabla c(x) = (\nabla c_1(x), \dots, \nabla c_m(x)) = \left(\frac{\partial c_j(x)}{\partial x_i} \right)_{\substack{i=1, \dots, n \\ j=1, \dots, m}} = \begin{pmatrix} \frac{\partial c_1(x)}{\partial x_1} & \dots & \frac{\partial c_m(x)}{\partial x_1} \\ \dots & \dots & \dots \\ \frac{\partial c_1(x)}{\partial x_n} & \dots & \frac{\partial c_m(x)}{\partial x_n} \end{pmatrix}$$

Matrice jacobienne (« jacobien » = déterminant de J_c)

$$J_c(x) = \nabla c(x)^T = \begin{pmatrix} \nabla c_1(x)^T \\ \dots \\ \nabla c_m(x)^T \end{pmatrix} = \left(\frac{\partial c_i(x)}{\partial x_j} \right)_{\substack{i=1, \dots, m \\ j=1, \dots, n}} = \begin{pmatrix} \frac{\partial c_1(x)}{\partial x_1} & \dots & \frac{\partial c_1(x)}{\partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial c_m(x)}{\partial x_1} & \dots & \frac{\partial c_m(x)}{\partial x_n} \end{pmatrix}$$

Contraintes du problème (PO)

Matrice jacobienne regroupant les contraintes égalité c_E et les contraintes inégalité c_I

$$J(x) : \mathbb{R}^n \rightarrow \mathbb{R}^{(p+q) \times n} \quad J(x) = \begin{pmatrix} J_E(x) \\ J_I(x) \end{pmatrix} = \begin{pmatrix} \nabla c_E(x)^T \\ \nabla c_I(x)^T \end{pmatrix}$$

Notes :

- dans un PO, les fonctions c_i utilisés sont les contraintes du problème.
- la matrice gradient/jacobienne stocke l'ensemble des différentielles des m contraintes $\{c_i\}_{i=1 \dots m}$.
- noter que si c est une fonction unique, le Jacobien est la transposée du gradient.

EXERCICE 7: Jacobienne

Donner la matrice Jacobienne pour les fonctions F suivantes :

1. $F(x, y, z) = (xyz, x^2z)$
2. $F(x, y) = (e^{xy}, \ln x)$
3. $F(x, y, z) = (\sin(xyz), xz)$

4 Théorème de Taylor

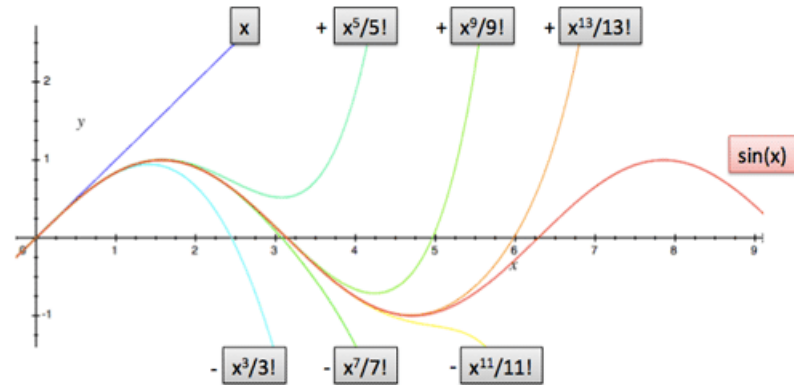
4.1 Ordre 1 et 2

Pour une fonction réelle, la formule de Taylor s'écrit $f(x+h) = \sum_{k=0}^n \frac{f^{(k)}(x)}{k!} h^k + R_n(h)$ où $\exists \epsilon \in [0, 1]$, tel que $R_n(h) = \frac{f^{(n+1)}(x+h\epsilon)}{(n+1)!} h^{n+1} = o(h^n)$.

Exemple :

L'approximation de Taylor de $\sin(x)$ en $x = 0$ est :

$$\sin(x) = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{7!}x^7 + \frac{1}{9!}x^9 - \frac{1}{11!}x^{11} + \frac{1}{13!}x^{13} + o(h^{15})$$



Dans \mathbb{R}^n , il s'exprime sous la forme suivante :

Théorème de Taylor

f fonction de \mathbb{R}^n dans \mathbb{R} : $x \in \mathbb{R}^n \mapsto f(x) \in \mathbb{R}$

$d \in \mathbb{R}^n$: déplacement à partir de x

• **Ordre 1**

f continument différentiable 1 fois au voisinage de x

$$f(x + d) = f(x) + \nabla f(x)^T d + o(\|d\|)$$

Il existe $s \in [0, 1]$ tel que : $f(x + d) = f(x) + \nabla f(x + sd)^T d$

• **Ordre 2**

f continument différentiable 2 fois au voisinage de x

$$f(x + d) = f(x) + \nabla f(x)^T d + \frac{1}{2} d^T \nabla^2 f(x) d + o(\|d\|^2)$$

Il existe $s \in [0, 1]$ tel que : $f(x + d) = f(x) + \nabla f(x)^T d + \frac{1}{2} d^T \nabla^2 f(x + sd) d$

A l'ordre 1, on approxime la fonction par un hyperplan. A l'ordre 2, on l'approxime par une hypersurface quadratique.

4.2 Modèle quadratique-linéaire

Par modèle quadratique-linéaire, on entend :

- f est une fonction quadratique (= d'ordre 2), donc le développement de Taylor s'arrête à l'ordre 2.
- les contraintes c_i sont linéaires (= d'ordre 1), donc le développement de Taylor s'arrête à l'ordre 1.

Problème avec contraintes égalité

$$\min_{x \in \mathbb{R}^n} f(x) \text{ sous } c(x) = 0 \rightarrow \text{contrainte actives}$$

Fonction modèle

- Application du théorème de Taylor au point $x_0 \in \mathbb{R}^n$
- **Modèle quadratique du critère** : $\hat{f}_0(x) = f(x_0) + \nabla f(x_0)^T (x - x_0) + \frac{1}{2} (x - x_0)^T \nabla^2 f(x_0) (x - x_0)$
- **Modèle linéaire des contraintes** : $\hat{c}_0(x) = c(x_0) + \nabla c(x_0)^T (x - x_0)$

Problème quadratique-linéaire local

$$\text{Au voisinage de } x_0 : \begin{cases} f(x) \approx \hat{f}_0(x) \\ c(x) \approx \hat{c}_0(x) \end{cases}$$

$$\min_{x \in \mathbb{R}^n} \hat{f}_0(x) \text{ sous } \hat{c}_0(x) = 0 \rightarrow \text{Problème quadratique-linéaire}$$

- localement « proche » du problème initial
- plus simple à résoudre

4.3 Exemples

- Fonction : $f(x) = -x^4 + 12x^3 - 47x^2 + 60x$
- Gradient : $\nabla f(x) = -4x^3 + 36x^2 - 94x + 60$
- Hessien : $\nabla^2 f(x) = -12x^2 + 72x - 94$
- Modèle quadratique en $x_0 = 3$: $f(x_0) = 0$, $\nabla f(x_0) = -6$, $\nabla^2 f(x_0) = 14$

$$\hat{f}_0(x) = -6(x - 3) + \frac{1}{2} 14(x - 3)^2 = 7x^2 - 48x + 81$$
- Modèle quadratique en $x_0 = 4$: $f(x_0) = 0$, $\nabla f(x_0) = 4$, $\nabla^2 f(x_0) = 2$

$$\hat{f}_0(x) = 4(x - 4) + \frac{1}{2} 2(x - 4)^2 = x^2 - 4x$$
- Modèle quadratique en $x_0 = 5$: $f(x_0) = 0$, $\nabla f(x_0) = -10$, $\nabla^2 f(x_0) = -34$

$$\hat{f}_0(x) = -10(x - 5) - \frac{1}{2} 34(x - 5)^2 = -17x^2 + 160x - 375$$

- Fonction : $f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2 \rightarrow$ **fonction de Rosenbrock**
 - Gradient : $\nabla f(x) = \begin{pmatrix} -400(x_2 - x_1^2)x_1 - 2(1 - x_1) \\ 200(x_2 - x_1^2) \end{pmatrix}$
 - Hessien : $\nabla^2 f(x) = \begin{pmatrix} -400(x_2 - 3x_1^2) + 2 & -400x_1 \\ -400x_1 & 200 \end{pmatrix}$
 - Modèle quadratique en $x_0 = (1, 1)$: $f(x_0) = 0$, $\nabla f(x_0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $\nabla^2 f(x_0) = \begin{pmatrix} 802 & -400 \\ -400 & 200 \end{pmatrix}$
- $$\hat{f}_0(x_1, x_2) = \frac{1}{2} \begin{pmatrix} x_1 - 1 \\ x_2 - 1 \end{pmatrix}^T \begin{pmatrix} 802 & -400 \\ -400 & 200 \end{pmatrix} \begin{pmatrix} x_1 - 1 \\ x_2 - 1 \end{pmatrix}$$
- $$\Rightarrow \hat{f}_0(x_1, x_2) = 401(x_1 - 1)^2 - 400(x_1 - 1)(x_2 - 1) + 100(x_2 - 1)^2$$

EXERCICE 8: Développement de Taylor

Calculer le développement de Taylor à l'ordre des fonctions suivante au point indiqué :

1. $f(x, y) = \ln(1 + x + 2y)$ au point $(2, 1)$.
2. $f(x, y) = x^3 + 3x^2y + 6xy^2 - 5x^2 + 3xy^2$ au point $(1, 2)$.
3. $f(x, y) = e^{x+y}$ au point $(0, 0)$.
4. $f(x, y) = \sin(xy) + \cos(xy)$ au point $(0, 0)$.
5. $f(x, y, z) = x - y^2 + xz$ au point $(1, 0, 3)$.

4.4 Lignes de niveau

Les lignes de niveau sont un moyen pratique pour représenter les "reliefs" d'une fonction (cf carte IGN).

Définition

- **Ligne (ou surface) de niveau l_0 de la fonction f**

$L_0 = \{x \in \mathbb{R}^n / f(x) = l_0\} \rightarrow$ hypersurface dans \mathbb{R}^n (sous-espace de dimension $n-1$)

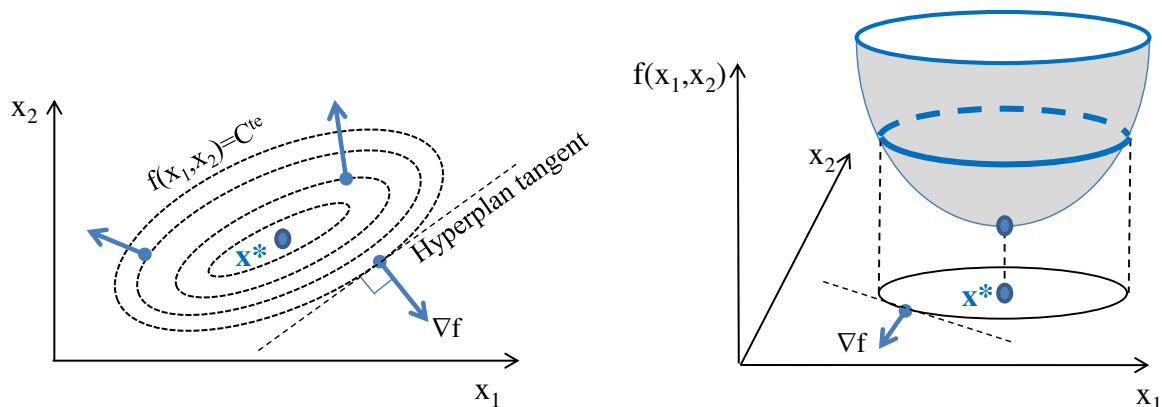
- Ligne de niveau passant par x_0

$L_0 = \{x \in \mathbb{R}^n / f(x) = f(x_0)\}$ avec $f(x) = f(x_0) + \nabla f(x_0)^T (x - x_0) + o(\|x - x_0\|)$ à l'ordre 1

Gradient

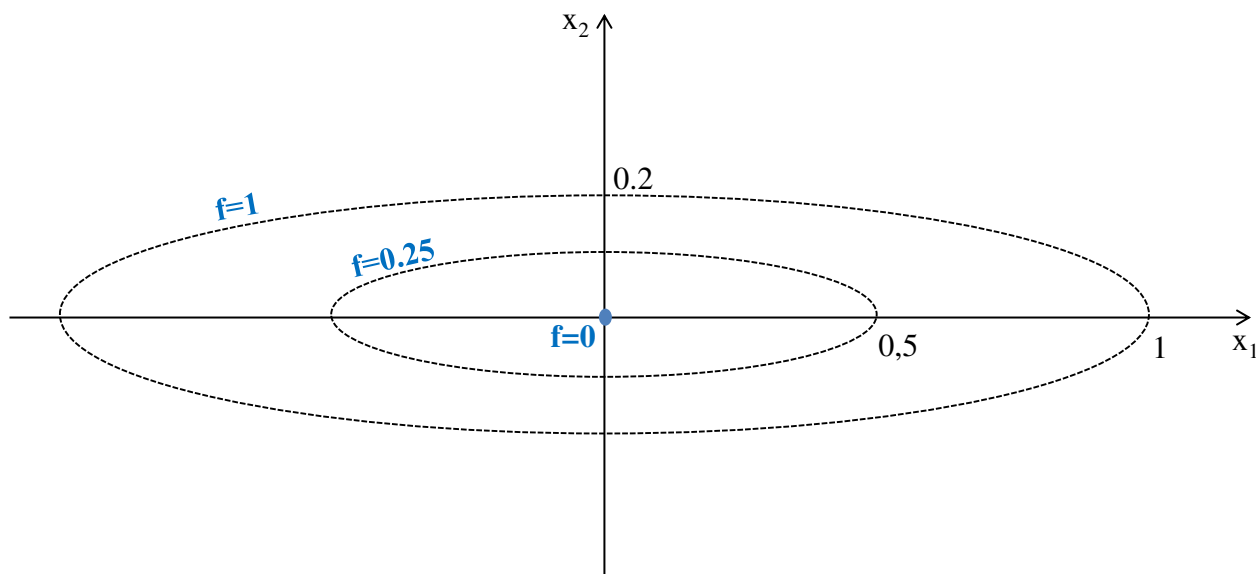
$x \in L_0 \Rightarrow f(x) = f(x_0) \Rightarrow \nabla f(x_0)^T (x - x_0) = 0 \rightarrow$ hyperplan tangent à L_0 en x_0

Le gradient de f est orthogonal aux lignes de niveau.



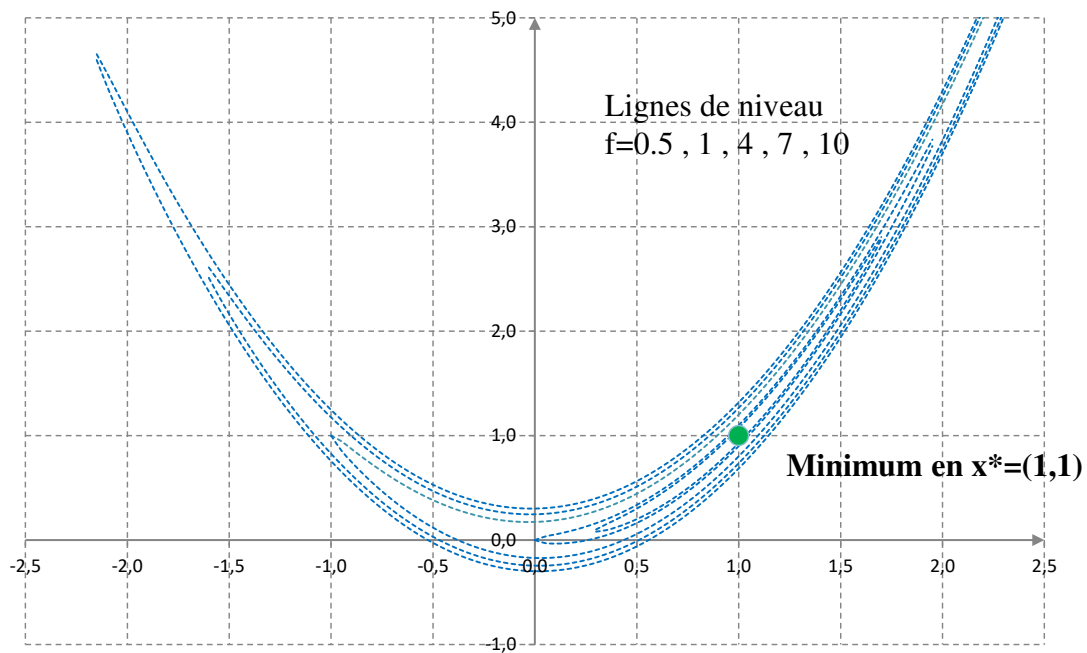
Ligne de niveau : fonction quadratique

$$f(x_1, x_2) = x_1^2 + 25x_2^2$$



Ligne de niveau : fonction de Rosenbrock

$$f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

**EXERCICE 9: Courbes de niveau**

les fonctions $f(x, y)$ suivantes pour les valeurs c indiquée, tracer les courbes de niveau $f(x, y) = c$.

1. $f(x, y) = xy$ pour $c = +1, -1, +3$.
2. $f(x, y) = e^{xy}$ pour $c = +1, -1, +3$.
3. $f(x, y) \ln(xy)$ pour $c = 0, +1, -1$.
4. $f(x, y) = (x + y)/(x - y)$ pour $c = 0, +2, -2$.
5. $f(x, y) = x^2 - y$ pour $c = 0, +1, -1$.
6. $f(x, y) = ye^x$ pour $c = 0, +1, -1$.

5 Dérivées numériques

5.1 Différences finies

Différences finies

Les expressions analytiques de $\nabla f(x)$ et $\nabla^2 f(x)$ ne sont généralement pas disponibles.

→ Evaluation par différences finies avec incrément h appliqué successivement sur chaque variable

$x \rightarrow x + he_i$ avec $e_i = (0, \dots, 0, 1, 0, \dots, 0)^T$, $i = 1$ à n

Gradient

- Différence finie simple : $\frac{\partial f}{\partial x_i}(x) = \frac{f(x + he_i) - f(x)}{h} + o(h)$
 amont si $h < 0$
 aval si $h > 0$
 → n appels fonction pour évaluer $\nabla f(x)$
- Différence finie centrée : $\frac{\partial f}{\partial x_i}(x) = \frac{f(x + he_i) - f(x - he_i)}{2h} + o(h^2)$
 plus précis
 → $2n$ appels fonction pour évaluer $\nabla f(x)$

Hessien

- Différence finie simple : $\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{f(x + he_i + he_j) - f(x + he_i) - f(x + he_j) + f(x)}{h^2} + o(h)$
 → $n(n+1)/2 + n$ appels fonction pour évaluer $\nabla^2 f(x)$

EXERCICE 10: Différences finies

On considère la fonction $f(x) = x^3$.

- Calculer $f'(x)$, et $f'(1)$.
- Calculer $f(1 + h)$. On développera.
- Comment aurait-on pu obtenir autrement l'expression de la question précédente ?
- Donner l'expression de $f'(1)$ par différence finie.
- On veut estimer $f'(1)$ par différence finie. On utilise un pas $h_n = 10^{-n}$, et on mesure l'erreur d'approximation avec $\varepsilon_n(x) = \frac{f(x+h_n) - f(x)}{h_n} - f'(x)$.

h_n	$\varepsilon_n(x)$	h_n	$\varepsilon_n(x)$	h_n	$\varepsilon_n(x)$	h_n	$\varepsilon_n(x)$
1	6.410e-01	5	6.000043e-05	9	3.309615e-07	13	-3.197111e-03
2	6.040100e-02	6	5.999760e-06	10	3.309615e-07	14	-3.197111e-03
3	6.004001e-03	7	6.018559e-07	11	3.309615e-07	15	4.408921e-01
4	6.000400e-04	8	4.230350e-08	12	3.556023e-04	16	-4.000000e+00

Expliquer le comportement de $\varepsilon_n(x)$ sachant que $\varepsilon = 10^{-16}$ sur la machine considérée.

EXERCICE 11: Différences finies

Une fonctionnelle f est "mesurée" pour les valeurs suivantes.

n	1	2	3	4	5	6
x_n	0.50	0.60	0.70	0.80	0.90	1.00
$f(x_n)$	0.125	0.216	0.343	0.512	0.729	1.000

n	7	8	9	10	11
x_n	1.10	1.20	1.30	1.40	1.50
$f(x_n)$	1.331	1.728	2.197	2.744	3.375

1. calculer le pas h .
2. calculer le gradient en x_6 avec la méthode des différences finies simples.
3. calculer le gradient en x_6 avec la méthode des différences finies centrées.
4. calculer le Hessien en x_6 avec la méthode des différences finies simples.
5. la fonctionnelle f est en fait la fonction $f(x) = x^3$. Calculer les erreurs d'approximation.

EXERCICE 12: Différences finies

On dispose de l'estimation suivante des valeurs d'une fonctionnelle $f(x, y)$ sur une grille.

0.512	0.648	0.800	0.968	1.152
0.576	0.729	0.900	1.089	1.296
0.640	0.810	1.000	1.210	1.440
0.704	0.891	1.100	1.331	1.584
0.768	0.972	1.200	1.452	1.728

On effectuera toutes les estimations au centre de la grille (*i.e.* au point où $f(x, y) = 1$), et on supposera que $h = 0.1$.

1. Calculer le gradient au centre de la grille par la méthode des différences finies.
2. Calculer la dérivée directionnelle dans la direction $d = (1, 1)$.
3. Calculer le Hessien.
4. Donner la formule de Taylor d'ordre 1 de f , et l'utiliser avec $\mathbf{h}' = [0.1 \quad 0.1]^T$.
5. Donner la formule de Taylor d'ordre 2 de f , et l'utiliser avec $\mathbf{h}' = [0.1 \quad 0.1]^T$.

5.2 Sources d'erreurs

Sources d'erreurs

L'évaluation d'une dérivée par différence finie génère 2 types d'erreurs :

- Erreur d'arrondi (ou de conditionnement)
- Erreur de troncature

Erreur d'arrondi

Les réels sont représentés en mémoire machine calcul avec une précision finie.

La précision machine ε_m est le plus petit réel tel que : $1 + \varepsilon_m \neq 1$

→ erreur relative $\varepsilon_m = 10^{-16}$ sur la valeur d'un réel x en double précision

→ erreur relative ε_r sur la valeur évaluée de $f(x)$

$\varepsilon_r \gg \varepsilon_m$ par cumul des erreurs au cours des opérations pour passer de x à $f(x)$

$$f_{\text{eval}}(x) = f_{\text{exact}}(x)(1 \pm \varepsilon_r) = f_{\text{exact}}(x) \pm \varepsilon_f \quad \rightarrow \quad \varepsilon_f = \text{erreur absolue sur } f$$

Erreur de troncature

L'évaluation d'une dérivée par différence finie tronque le développement de Taylor à l'ordre 1.

- $f(x_0 + h) = f(x_0) + hf'(x_0) + \frac{1}{2}h^2f''(x_0 + sh)$ avec $s \in [0,1]$
- $f(x_0 + h) = f(x_0) + hf'(x_0) \pm \frac{1}{2}h^2\varepsilon_t$ avec $\varepsilon_t < M$ majorant de $|f''(x)|$ sur $[x_0, x_0 + d]$

Erreur sur la dérivée

$$f'_{\text{eval}}(x_0) = \frac{f_{\text{eval}}(x_0 + h) - f_{\text{eval}}(x_0)}{h} \quad \text{avec}$$

- $f_{\text{eval}}(x_0) = f_{\text{exact}}(x_0) \pm \varepsilon_f \quad \rightarrow \text{arrondi sur } f_{\text{eval}}(x_0)$
- $f_{\text{eval}}(x_0 + h) = f_{\text{exact}}(x_0 + h) \pm \varepsilon_f \quad \rightarrow \text{arrondi sur } f_{\text{eval}}(x_0 + h)$
- $f_{\text{exact}}(x_0 + h) = f_{\text{exact}}(x_0) + hf'_{\text{exact}}(x_0) \pm \frac{1}{2}h^2\varepsilon_t \quad \rightarrow \text{troncature sur } f_{\text{exact}}(x_0 + h)$

En remplaçant dans l'expression de $f'_{\text{eval}}(x_0)$:

$$\Rightarrow f'_{\text{eval}}(x_0) = \frac{f_{\text{exact}}(x_0 + h) - f_{\text{exact}}(x_0) \pm 2\varepsilon_f}{h} = \frac{hf'_{\text{exact}}(x_0) \pm \frac{1}{2}h^2\varepsilon_t \pm 2\varepsilon_f}{h}$$

$$\Rightarrow f'_{\text{eval}}(x_0) = f'_{\text{exact}}(x_0) \pm \frac{h\varepsilon_t}{2} \pm \frac{2\varepsilon_f}{h}$$

L'erreur maximale sur la dérivée numérique est :

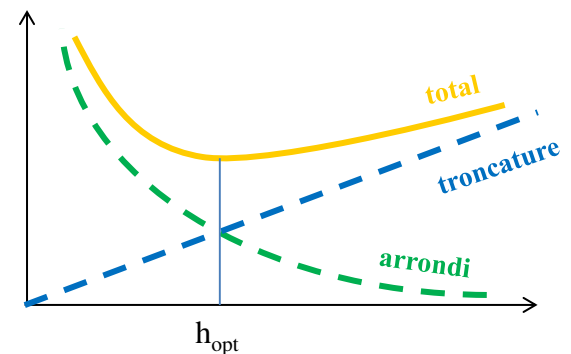
$$\boxed{\varepsilon_{f'} = \frac{h\varepsilon_t}{2} + \frac{2\varepsilon_f}{h}}$$

Incrément optimal

- On choisit l'incrément pour minimiser l'erreur : $\min_h \varepsilon_{f'} = \frac{h\varepsilon_t}{2} + \frac{2\varepsilon_f}{h}$

$$\frac{d\varepsilon_{f'}}{dh} = \frac{\varepsilon_t}{2} - \frac{2\varepsilon_f}{h^2} = 0 \Rightarrow h_{\text{opt}} = 2\sqrt{\frac{\varepsilon_f}{\varepsilon_t}}$$

$$\Rightarrow \varepsilon_{f'} = 2\sqrt{\varepsilon_f \varepsilon_t}$$



- Règle empirique** de réglage de l'incrément

En supposant que l'ordre de grandeur de la dérivée seconde est de l'ordre de 1 :

$$h_{\text{opt}} \approx \sqrt{\varepsilon_f}$$

→ incrément de l'ordre de la racine de la précision d'évaluation de f

$$\varepsilon_{f'} \approx \sqrt{\varepsilon_f}$$

→ précision sur f' de l'ordre de la racine de la précision d'évaluation de f
(2 fois moins de chiffres significatifs)

En effet,

$$\min_h \varepsilon_{f'} = 2\sqrt{\frac{\varepsilon_f}{\varepsilon_t} \frac{\varepsilon_t}{2}} + 2\varepsilon_f \frac{1}{2} \sqrt{\frac{\varepsilon_t}{\varepsilon_f}} = \sqrt{\varepsilon_t \varepsilon_f}$$

à savoir, à une constante proportionnelle près.

si h est trop grand, $o(h)$ n'est pas négligeable.

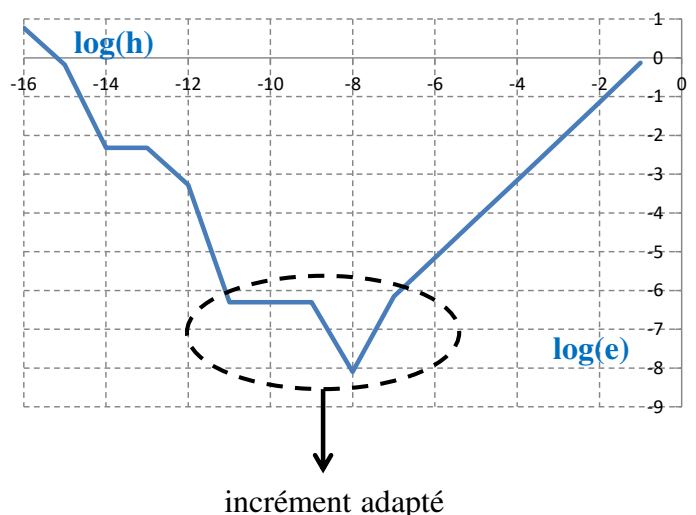
si h est trop petit, les erreurs numériques dominent $o(h)$.

Dérivée numérique

$$f(x) = x^4 + x^2 \rightarrow f'(x) = 4x^3 + 2x$$

- Dérivée en $x=1$: $f'(1) = 6$
- Dérivée numérique avec incrément $h \rightarrow$ erreur $e(h) = \left| \frac{f(x+h) - f(x)}{h} - f'(x) \right|$

h	$(f(x+h)-f(x))/h$	Erreur
1E-01	6,7410000000	7,410E-01
1E-02	6,0704010000	7,040E-02
1E-03	6,0070040010	7,004E-03
1E-04	6,0007000400	7,000E-04
1E-05	6,0000700004	7,000E-05
1E-06	6,0000069997	7,000E-06
1E-07	6,0000007007	7,007E-07
1E-08	6,0000000079	7,944E-09
1E-09	6,0000004964	4,964E-07
1E-10	6,0000004964	4,964E-07
1E-11	6,0000004964	4,964E-07
1E-12	6,0005334035	5,334E-04
1E-13	5,9952043330	-4,796E-03
1E-14	5,9952043330	-4,796E-03
1E-15	6,6613381478	6,613E-01



5.3 Dérivée complexe

On peut améliorer la précision sur la dérivée par différence finie en nombres complexes.

- Développement de f en x_0 : $f(x_0 + h) = f(x_0) + hf'(x_0) + \frac{h^2}{2}f''(x_0) + \frac{h^3}{6}f'''(x_0) + o(h^3)$

- Estimation de la dérivée avec un **incrément réel h**

$$\frac{f(x_0 + h) - f(x_0)}{h} = f'(x_0) + \frac{h}{2}f''(x_0) + \frac{h^2}{6}f'''(x_0) + o(h^2) \rightarrow \text{précision de l'ordre de } h$$

L'incrément h ne peut pas être choisi trop petit
à cause de l'erreur d'arrondi

$$\rightarrow h \approx \sqrt{\varepsilon_f}$$

- Estimation de la dérivée avec un **incrément imaginaire ih**

$$\frac{f(x_0 + ih) - f(x_0)}{h} = if'(x_0) - \frac{h}{2}f''(x_0) - i\frac{h^2}{6}f'''(x_0) + o(h^2)$$

$$\Rightarrow \operatorname{Im}\left(\frac{f(x_0 + ih) - f(x_0)}{h}\right) = f'(x_0) - \frac{h^2}{6}f'''(x_0) + o(h^2) \rightarrow \text{précision de l'ordre de } h^2$$

L'incrément h peut être choisi à la **précision machine**
car il ne porte que sur la partie imaginaire de f

$$\rightarrow h \approx \varepsilon_f$$

Dérivée complexe

- La dérivée numérique est évaluée par différence finie complexe.

$$f'(x_0) \approx \operatorname{Im}\left(\frac{f(x_0 + ih) - f(x_0)}{h}\right)$$

- Il faut que l'ensemble du logiciel soit écrit en variables complexes pour propager l'incrément imaginaire ih de $x_0 + ih$ à $f(x_0 + ih)$.
→ Programmation directe en déclarant des variables complexes
ou
→ **Surcharge d'opérateurs** pour effectuer toutes les opérations en complexes

Exemple

$$f(x) = x^2 \rightarrow f'(x) = 2x$$

- Différence finie réelle : $\frac{f(x_0 + h) - f(x_0)}{h} = \frac{(x_0 + h)^2 - (x_0)^2}{h} = 2x_0 + h \rightarrow \text{erreur} = h$
- Différence finie imaginaire : $\frac{f(x_0 + ih) - f(x_0)}{h} = \frac{(x_0 + ih)^2 - (x_0)^2}{h} = 2ix_0 - h^2$
$$\Rightarrow \operatorname{Im}\left(\frac{f(x_0 + ih) - f(x_0)}{h}\right) = 2x_0 \rightarrow \text{valeur exacte}$$

5.4 Méthode d'extrapolation

Méthode d'extrapolation

- La dérivée de f en x_0 est par définition : $f'(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}$

L'approximation par différence finie comporte une erreur de troncature en $o(h)$.

La réduction de l'incrément h introduit une erreur d'arrondi liée à la précision machine.

Pour améliorer la précision, on peut appliquer la **méthode d'extrapolation de Richardson**.

- On cherche la limite a_0 d'une fonction $A(h)$ **non définie en $h = 0$** .

$$\lim_{h \rightarrow 0} A(h) = a_0 \quad \rightarrow \text{à évaluer}$$

Le développement de Taylor à l'ordre n de A en 0 est

$$A(h) = a_0 + a_1 h + a_2 h^2 + \dots + a_n h^n + o(h^n)$$

- La méthode d'extrapolation de Richardson consiste à :
 - évaluer la fonction A en plusieurs points h_0, h_1, \dots, h_m
 - combiner les valeurs $A_k = A(h_k)$ pour éliminer les termes du développement de Taylor

Développement de Taylor

- On choisit un incrément $h > 0$ et un rapport $r : 0 < r < 1$.

La fonction A est évaluée en $m+1$ points h_0, h_1, \dots, h_m définis par : $h_k = r^k h$

Pour $r = \frac{1}{2}$ (valeur usuelle) : $h_0 = h$, $h_1 = \frac{h}{2}$, $h_2 = \frac{h}{4}$, $h_3 = \frac{h}{8}$, ...

- On calcule les valeurs de A aux points h_k : $A_{k,0} \stackrel{\text{def}}{=} A(h_k) \quad \rightarrow \text{m+1 valeurs de rang 0}$

En utilisant le développement de Taylor : $A(h) = a_0 + a_1 h + a_2 h^2 + \dots + a_n h^n + o(h^n)$

$$\text{on obtient à l'ordre } n : \begin{cases} A_{0,0} = A(r^0 h) = a_0 + a_1 h + a_2 h^2 + \dots + a_n h^n + o(h^n) \\ A_{1,0} = A(r^1 h) = a_0 + a_1 r h + a_2 r^2 h^2 + \dots + a_n r^n h^n + o(h^n) \\ A_{2,0} = A(r^2 h) = a_0 + a_1 r^2 h + a_2 r^4 h^2 + \dots + a_n r^{2n} h^n + o(h^n) \\ \vdots \\ A_{m,0} = A(r^m h) = a_0 + a_1 r^m h + a_2 r^{2m} h^2 + \dots + a_n r^{mn} h^n + o(h^n) \end{cases}$$

- Les $m+1$ valeurs $A_{k,0}$ de rang 0 sont ensuite combinées pour éliminer les termes en h .
 $\rightarrow m$ valeurs $A_{k,1}$ de rang 1

Elimination des termes en h

- On définit les m valeurs $A_{k,1}$ **de rang 1** à partir des m+1 valeurs $A_{k,0}$ de rang 0.

$$A_{k,1} \stackrel{\text{def}}{=} \frac{A_{k,0} - rA_{k-1,0}}{1-r} \quad \text{pour } k = 1 \text{ à } m$$

- Ces combinaisons gardent le terme a_0 et éliminent les termes en h du développement de Taylor.

$$\begin{cases} A_{1,1} = a_0 + b_2 h^2 + \dots + b_n h^n + o(h^n) \\ A_{2,1} = a_0 + b_2 r^2 h^2 + \dots + b_n r^n h^n + o(h^n) \\ A_{3,1} = a_0 + b_2 r^4 h^2 + \dots + b_n r^{2n} h^n + o(h^n) \\ \vdots \\ A_{m,1} = a_0 + b_2 r^{2m} h^2 + \dots + b_n r^{mn} h^n + o(h^n) \end{cases}$$

Les termes d'ordre supérieur en h^2, \dots, h^n ont les mêmes coefficients b_2, \dots, b_n .

- On élimine de la même façon les termes en $h^2 \rightarrow m-1$ valeurs $A_{k,2}$ de rang 2
 puis les termes en $h^3 \rightarrow m-2$ valeurs $A_{k,3}$ de rang 3

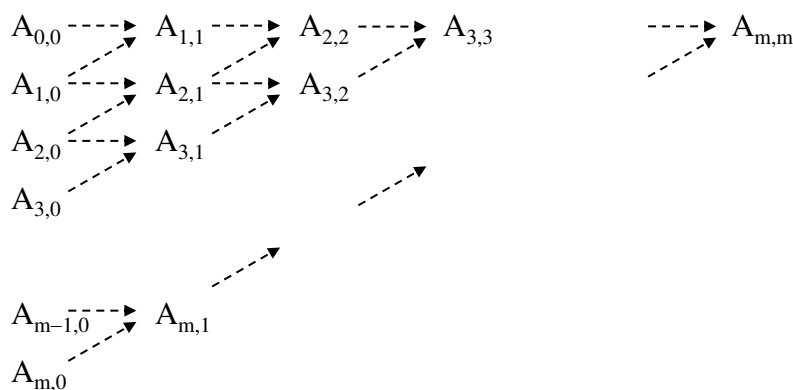
 jusqu'aux termes en $h^m \rightarrow 1$ valeur $A_{m,m}$ de rang m

Elimination des termes jusqu'à h^m

- On applique la formule de récurrence entre le rang $j-1$ et le rang j .

$$A_{k,j} \stackrel{\text{def}}{=} \frac{A_{k,j-1} - r^j A_{k-1,j-1}}{1-r^j} \quad \text{pour } k = j \text{ à } m, \text{ et } j = 1 \text{ à } m$$

- En pratique on dispose les calculs en colonnes.



- Les termes de rang j sont des approximations de a_0 à l'ordre h^j .
 Le dernier terme $A_{m,m}$ est une approximation de a_0 à l'ordre h^m : $A_{m,m} = a_0 + \alpha r^{m(m+1)} h^{m+1} + \dots$

Dérivée première et seconde

- Pour évaluer les dérivées premières et secondes de f en x_0 , on définit 2 fonctions $A(h)$ et $B(h)$.

$$A(h) = \frac{f(x_0 + rh) - f(x_0 + h)}{(r-1)h}$$

$$B(h) = 2 \frac{f(x_0 + rh) - rf(x_0 + h) + (r-1)f(x_0)}{r(r-1)h^2}$$

- Les développements de Taylor de A et B en $h=0$ sont

$$A(h) = f'(x_0) + \frac{1}{2} \frac{r^2 - 1}{r-1} hf''(x_0) + \dots + \frac{1}{n!} \frac{r^n - 1}{r-1} h^{n-1} f^{(n)}(x_0) + \dots$$

$$B(h) = f''(x_0) + \frac{1}{3} \frac{r^2 - 1}{r-1} hf'''(x_0) + \dots + \frac{2}{n!} \frac{r^{n-1} - 1}{r-1} h^{n-2} f^{(n)}(x_0) + \dots$$

- La méthode d'extrapolation de Richardson appliquée aux fonctions $A(h)$ et $B(h)$ permet d'évaluer les dérivées premières et secondes de f en x_0 .

Dérivée première et seconde

- On choisit un incrément initial h , un rapport r entre 0 et 1 et un ordre d'extrapolation m .

La fonction f est évaluée en $x_0+h_0, x_0+h_1, \dots, x_0+h_m$ avec $h_k = r^k h \rightarrow f_k = f(x_0 + h_k)$

- Les termes de rang 0 des fonctions $A(h)$ et $B(h)$ sont calculés à partir des valeurs f_k .

$$A_{k,0} = A(h_k) = \frac{f_{k+1} - f_k}{(r-1)h_k}$$

$$B_{k,0} = B(h_k) = 2 \frac{f_{k+1} - rf_k + (r-1)f(x_0)}{r(r-1)h_k^2}$$

- Les termes $A_{m,m}$ et $B_{m,m}$ de rang m sont des approximations de $f'(x_0)$ et $f''(x_0)$.
- En pratique l'ordre d'extrapolation m est limité par la précision machine ε .

$$A_{m,m} = a_0 + \alpha r^{m(m+1)} h^{m+1} + \dots \Rightarrow r^{m^2} h^m > \varepsilon \quad (\text{en supposant } a_0 \approx 1)$$

$$\Rightarrow m < \frac{\log h + \sqrt{(\log h)^2 + 4 \log r \log \varepsilon}}{-2 \log r} \quad \text{avec } r < 1$$

Exemple : $r=0.5, \varepsilon=10^{-16}, h=10^{-4} \rightarrow m < 4$
 $h=10^{-8} \rightarrow m < 3$

corrections : $h = 10^{-4} \Rightarrow m > 2.35 \geq 3, h = 10^{-8} \Rightarrow m > 3.75 \geq 4,$

Dérivée par extrapolation

$$f(x) = x^2 + \frac{1}{x} \rightarrow f'(x) = 2x - \frac{1}{x^2} \rightarrow f''(x) = 2 + \frac{2}{x^3}$$

- Dérivée première et seconde en $x=1$: $f'(1) = 1$, $f''(1) = 4$
- Incrément initial : $h = 0.1$ → ordre d'extrapolation $m = 6$

h_k	$x+h_k$	$f(x+h_k)$
0,10000000	1,10000000	2,11909091
0,05000000	1,05000000	2,05488095
0,02500000	1,02500000	2,02623476
0,01250000	1,01250000	2,01281057
0,00625000	1,00625000	2,00632788
0,00312500	1,00312500	2,00314450
0,00156250	1,00156250	2,00156738

A_0	A_1	A_2	A_3	A_4	A_5
1,2841991342	1,0074965685	1,0001968322	1,0000025180	1,0000000159	1,0000000000
1,1458478513	1,0020217662	1,0000268073	1,0000001723	1,0000000005	
1,0739348088	1,0005255470	1,0000035017	1,0000000113		
1,0372301779	1,0001340130	1,0000004476			
1,0186820955	1,0000338389				
1,0093579672					

→ $f'(1) = 1$ approximations de $f'(1)$ **Dérivée par extrapolation**

$$f(x) = x^2 + \frac{1}{x} \rightarrow f'(x) = 2x - \frac{1}{x^2} \rightarrow f''(x) = 2 + \frac{2}{x^3}$$

- Dérivée première et seconde en $x=1$: $f'(1) = 1$, $f''(1) = 4$
- Incrément initial : $h = 0.1$ → ordre d'extrapolation $m = 6$

h_k	$x+h_k$	$f(x+h_k)$
0,10000000	1,10000000	2,11909091
0,05000000	1,05000000	2,05488095
0,02500000	1,02500000	2,02623476
0,01250000	1,01250000	2,01281057
0,00625000	1,00625000	2,00632788
0,00312500	1,00312500	2,00314450
0,00156250	1,00156250	2,00156738

B_0	B_1	B_2	B_3	B_4	B_5
3,7316017316	3,9850068631	3,9996063356	3,9999949640	3,9999999683	4,0000000000
3,8583042973	3,9959564675	3,9999463855	3,9999996555	3,9999999990	
3,9271303824	3,9989489060	3,9999929968	3,9999999775		
3,9630396442	3,9997319741	3,9999991049			
3,9813858091	3,9999323222				
3,9906590657					

→ $f''(1) = 4$ approximations de $f''(1)$

EXERCICE 13: Extrapolation de Richardson

On voudrait estimer la dérivée première de $f(x) = x^3$ en $x = 1$ par la méthode de Richardson, avec un pas $h = 0.2$, un rapport $r = 0.5$,

1. Quel est la valeur qui devrait être trouvée ?
2. En combien de valeurs de x faut-il faire le calcul pour un ordre d'extrapolation de 3 ?
3. Calculer alors les valeurs de h_k , $x + h_k$ et $f(x + h_k)$.
4. Rappeler l'expression des termes d'ordre 0 pour l'estimation de la dérivée première avec la méthode de Richardson.
5. Calculer les termes d'ordre 0.
6. Donner alors la formule de récurrence pour les ordres suivants.
7. En déduire l'interpolation de la dérivée, et conclure.
8. Comparer avec l'estimation par différence finie.

6 Matrice

Valeurs et vecteurs propres

Une matrice $A \in \mathbb{R}^{n \times n}$ admet la valeur propre $\sigma \in \mathbb{R}$
 s'il existe un vecteur non nul $v \in \mathbb{R}^n$ tel que : $Av = \sigma v$
 v est un vecteur propre associé à la valeur propre σ .

Matrices particulières

- A non singulière \Leftrightarrow Aucune valeur propre de A n'est nulle.

- | | | |
|--|---|--|
| <ul style="list-style-type: none"> • A symétrique | \Leftrightarrow
\Rightarrow
\Rightarrow | $A^T = A$
A admet n valeurs propres réelles (distinctes ou non)
A admet une base orthonormée de vecteurs propres |
|--|---|--|

- A orthogonale $\Leftrightarrow AA^T = A^T A = I$

- | | | |
|---|--|---|
| <ul style="list-style-type: none"> • A semi-définie positive • A définie positive | \Leftrightarrow
\Leftrightarrow | $v^T A v > 0$ pour tout $v \in \mathbb{R}^n$
$v^T A v \geq 0$ pour tout $v \in \mathbb{R}^n$ |
|---|--|---|

- A symétrique définie positive \Rightarrow A admet n valeurs propres réelles positives (distinctes ou non)

6.1 Permutation

Permutation de colonnes

$A \in \mathbb{R}^{m \times n}$, matrice à m lignes, n colonnes

$$A = \begin{matrix} & \begin{matrix} 1 & \dots & j & \dots & k & \dots & n \end{matrix} \\ \begin{matrix} 1 \\ \vdots \\ m \end{matrix} & \begin{pmatrix} \times & \dots & \times & \dots & \times & \dots & \times \\ \vdots & & \vdots & & \vdots & & \vdots \\ \times & \dots & \times & \dots & \times & \dots & \times \end{pmatrix} \end{matrix}$$

$E \in \mathbb{R}^{n \times n}$, matrice de permutation des colonnes j et k
 $E^T = E$

$$E = \begin{matrix} & \begin{matrix} 1 & \dots & j & \dots & k & \dots & n \end{matrix} \\ \begin{matrix} 1 \\ \vdots \\ j \\ \vdots \\ k \\ \vdots \\ n \end{matrix} & \begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \end{matrix}$$



$$AE = \begin{matrix} & \begin{matrix} 1 & \dots & j & \dots & k & \dots & n \end{matrix} \\ \begin{matrix} 1 \\ \vdots \\ m \end{matrix} & \begin{pmatrix} \times & \dots & \times & \dots & \times & \dots & \times \\ \vdots & & \vdots & & \vdots & & \vdots \\ \times & \dots & \times & \dots & \times & \dots & \times \end{pmatrix} \end{matrix}$$

6.2 Factorisation

Factorisation de matrice

$A \in \mathbb{R}^{m \times n}$, matrice à m lignes, n colonnes avec $m < n$

A de rang plein : $\text{rang}(A) = m < n$

$$A = \begin{matrix} & \begin{matrix} 1 & \dots & j & \dots & k & \dots & n \end{matrix} \\ \begin{matrix} 1 \\ \vdots \\ m \end{matrix} & \begin{pmatrix} \times & \dots & \times & \dots & \times & \dots & \times \\ \vdots & & \vdots & & \vdots & & \vdots \\ \times & \dots & \times & \dots & \times & \dots & \times \end{pmatrix} \end{matrix}$$

3 types de factorisations sont utiles dans les algorithmes d'optimisation.

- Factorisation **LU** → Pour réduire le problème (variables dépendantes et indépendantes)
Pour construire une base de l'espace nul
- Factorisation **QR** → Pour réduire le problème (variables dépendantes et indépendantes)
Pour construire une base orthogonale de l'espace nul
- Factorisation **LL^T** ou **LDL^T** → Pour une matrice définie positive
Pour rendre le hessien défini positif

La factorisation LL^T est dite de Cholesky.

Factorisation LU

$A \in \mathbb{R}^{m \times n}$, matrice à m lignes, n colonnes

$$A = \begin{matrix} & \begin{matrix} 1 & \dots & j & \dots & k & \dots & n \end{matrix} \\ \begin{matrix} 1 \\ \vdots \\ m \end{matrix} & \begin{pmatrix} \times & \dots & \times & \dots & \times & \dots & \times \\ \vdots & & \vdots & & \vdots & & \vdots \\ \times & \dots & \times & \dots & \times & \dots & \times \end{pmatrix} \end{matrix}$$

- Matrice carrée $n \times n$

$$\boxed{AE = LU}$$

L $n \times n$ triangulaire inférieure

U $n \times n$ triangulaire supérieure

- Matrice rectangulaire $m \times n$, $m < n$ → Factorisation de A^T

$$EA^T = LU = \begin{pmatrix} L_1 \\ L_2 \end{pmatrix} U$$

L_1 $m \times m$ triangulaire inférieure

L_2 $(n-m) \times m$ pleine

U $m \times m$ triangulaire supérieure

- Base de l'espace nul : $Z = E^T \begin{pmatrix} L_1^{-T} L_2^T \\ -I \end{pmatrix} U^{-T}$

Méthode de factorisation LU

→ Méthode d'élimination de Gauss (ou **méthode du pivot de Gauss**)

--

Factorisation QR

$A \in \mathbb{R}^{m \times n}$, matrice à m lignes, n colonnes

$$A = \begin{matrix} & \begin{matrix} 1 & \dots & j & \dots & k & \dots & n \end{matrix} \\ \begin{matrix} 1 \\ \vdots \\ m \end{matrix} & \begin{pmatrix} \times & \dots & \times & \dots & \times & \dots & \times \\ \vdots & & \vdots & & \vdots & & \vdots \\ \times & \dots & \times & \dots & \times & \dots & \times \end{pmatrix} \end{matrix}$$

- Matrice rectangulaire $m \times n$

$$\boxed{AE = QR}$$

Q $m \times m$ orthogonale

$$\rightarrow QQ^T = I$$

R $m \times n$ triangulaire supérieure

- Base de l'espace nul → Factorisation de A^T

Q_1 $n \times m$ orthogonale

$$EA^T = QR = (Q_1 \ Q_2) R$$

Q_2 $n \times (n-m)$ orthogonale

$$\rightarrow Z = E^T (Q_2)$$

R $n \times m$ triangulaire supérieure

Méthode de factorisation QR

→ **Méthode de Householder** ou méthode de Givens

--

Factorisation LL^T

$A \in \mathbb{R}^{n \times n}$, matrice carrée $n \times n$ symétrique définie positive

$$\boxed{E^T A E = L L^T} \quad \text{avec } L \text{ matrice } n \times n \text{ triangulaire inférieure}$$

- Lien avec la factorisation QR

$$A E = Q R \Rightarrow E^T A^T A E = R^T Q^T Q R = R^T R \quad \text{car } Q \text{ orthogonale}$$

$$\rightarrow \begin{cases} R = L \\ Q = A E R^{-1} \end{cases}$$

Méthode de factorisation LL^T ou LDL^T

- **Méthode de Cholesky**

→ Permet de vérifier que la matrice A est bien définie positive

- Méthode de Cholesky modifiée

→ Permet de rendre la matrice A définie positive en la modifiant au cours de la factorisation

--

6.3 Système linéaire**Système linéaire**

$$A x = b \quad \text{avec} \quad \begin{cases} A \in \mathbb{R}^{m \times n} & \rightarrow \text{matrice de rang } r : \text{rang}(A) = r \\ b \in \mathbb{R}^m & \rightarrow m \text{ équations} \\ x \in \mathbb{R}^n & \rightarrow n \text{ inconnues} \end{cases}$$

Le rang de A est la dimension du sous-espace engendré par A (image de A).

$$\text{Im}(A) = \{y = A x, x \in \mathbb{R}^n\} \quad \rightarrow r = \dim(\text{Im}(A)) \leq m, n$$

Solutions possibles

- Pas de solution : système **incompatible** ($m > n$: plus d'équations que d'inconnues)
- Solution unique : système **non singulier** ($m = n$: autant d'équations que d'inconnues)
- Infinité de solutions : système **sous-déterminé** ($m < n$: moins d'équations que d'inconnues)

Problème d'optimisation

Contraintes linéaires $A x = b$ → système sous-déterminé (**$m < n$**)
 → $n - m$ inconnues «libres» permettant de minimiser le critère

--

Contraintes redondantes

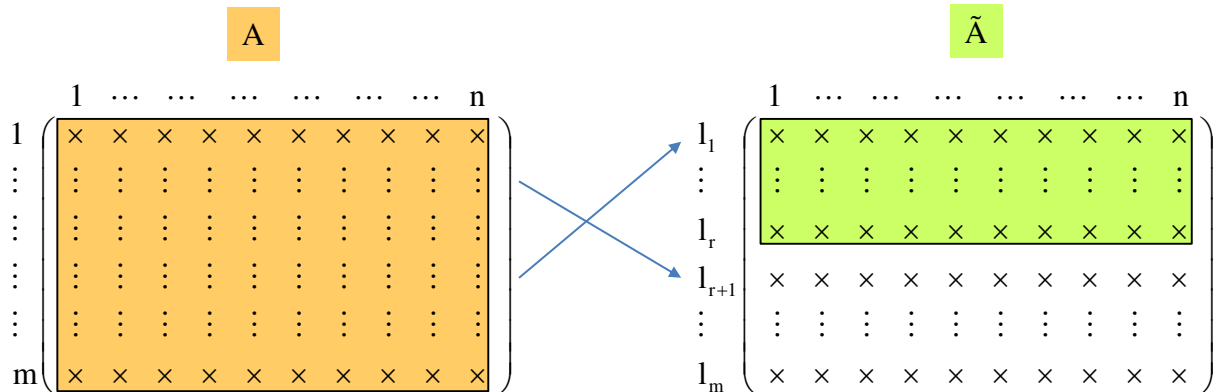
Pour un système sous-déterminé ($m < n$), si A est de rang déficient : $\text{rang}(A) = r < m$, on peut extraire de A une sous-matrice $\tilde{A} \in \mathbb{R}^{r \times n}$ de rang plein, telle que

$$\tilde{A}x = \tilde{b} \Leftrightarrow Ax = b$$

\tilde{A} est composée des lignes l_1, \dots, l_r de A .

Les lignes l_{r+1}, \dots, l_m sont combinaisons linéaires des lignes l_1, \dots, l_r .

Elles correspondent à des contraintes redondantes et peuvent être éliminées de la résolution.

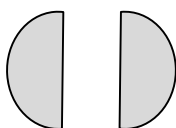


7 Convexité

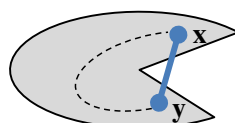
Ensemble convexe

$$X \subseteq \mathbb{R}^n \text{ convexe} \Leftrightarrow \forall x, y \in X, \forall \lambda \in [0, 1], \lambda x + (1 - \lambda)y \in X$$

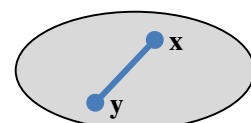
Interprétation géométrique : Segment inclus dans X



Non connexe



Non convexe



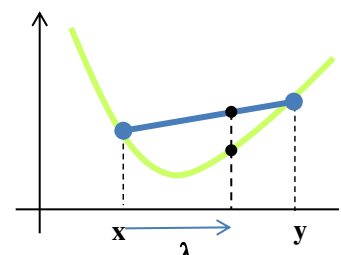
Convexe

Fonction convexe

f fonction de \mathbb{R}^n dans \mathbb{R}

- f convexe $\Leftrightarrow \forall x, y \in \mathbb{R}^n, \forall \lambda \in [0, 1], f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$
- f strictement convexe $\Leftrightarrow \forall x, y \in \mathbb{R}^n, \forall \lambda \in]0, 1[, f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y)$
- f concave $\Leftrightarrow -f$ convexe

Interprétation géométrique : Sécante au dessus de la courbe
Concavité vers le haut

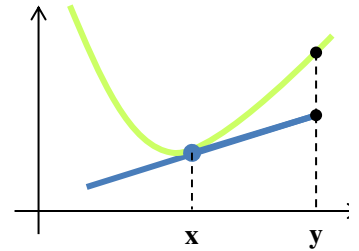


Convexité et gradient

f fonction différentiable de $X \subseteq \mathbb{R}^n$ dans \mathbb{R} , X ensemble convexe ouvert

- f convexe $\Leftrightarrow \forall x, y \in X, f(y) - f(x) \geq (y - x)^T g(x)$
- f strictement convexe $\Leftrightarrow \forall x, y \in X, f(y) - f(x) > (y - x)^T g(x)$

Interprétation géométrique : Tangente au dessous de la courbe

**Convexité et hessien**

f fonction deux fois différentiable de $X \subseteq \mathbb{R}^n$ dans \mathbb{R} , X ensemble convexe ouvert

- f convexe $\Leftrightarrow \forall x \in \mathbb{R}^n, H(x)$ **semi-définie positive**
- f strictement convexe $\Leftrightarrow \forall x \in \mathbb{R}^n, H(x)$ définie positive

Matrice définie positive

$A \in \mathbb{R}^{n \times n}$

- A définie positive $\Leftrightarrow \forall d \in \mathbb{R}^n, d^T A d > 0$
- A semi-définie positive $\Leftrightarrow \forall d \in \mathbb{R}^n, d^T A d \geq 0$

Comment vérifier la convexité d'une fonction à partir de son Hessien ?

Soit $H_f(x)$ le Hessien de f en x .

Pour résumer :

- f est strictement concave en p ssi $H_f(p)$ définie négative,
- f est concave en p ssi $H_f(p)$ définie semi-définie négative,
- f est strictement convexe en p ssi $H_f(p)$ définie positive,
- f est convexe en p ssi $H_f(p)$ définie semi-définie positive.

f est concave (resp. convexe) sur un ensemble convexe D ssi $\forall x \in D, f(x)$ est concave (resp. convexe).

Quels sont les propriétés mathématiques qui peuvent être utilisés ?

- un Hessien est **symétrique**,
- une matrice est symétrique définie positive (resp. négative) a ses valeurs propres réelles positives (resp. négative).
- dans la base de ses vecteurs propres, une matrice est diagonale.

Regardons tout d'abord comment vérifier si une matrice est (semi-) définie-positive/négative dans le cas d'une matrice diagonale A .

$$\text{Soit } A = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}.$$

Sa forme quadratique associée est $Q(x) = xAx^T = \lambda_1 x_1^2 + \dots + \lambda_n x_n^2$.

En conséquence,

- A est définie positive ($\Leftrightarrow \forall x, Q(x) > 0$) ssi $\forall i, \lambda_i > 0$.
- A est semi-définie positive ($\Leftrightarrow \forall x, Q(x) \geq 0$) ssi $\forall i, \lambda_i \geq 0$.
- A est définie négative ($\Leftrightarrow \forall x, Q(x) < 0$) ssi $\forall i, \lambda_i < 0$.
- A est semi-définie négative ($\Leftrightarrow \forall x, Q(x) \leq 0$) ssi $\forall i, \lambda_i \leq 0$.
- indéfinie ssi $\lambda_i < 0$ ou $\lambda_i > 0$

Comment faire dans le cas général d'une matrice symétrique ?

Soit une matrice symétrique $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{12} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{bmatrix}$.

Soit $D_1 = a_{11}$, $D_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix}$, $D_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix}$, ..., $D_n = |A|$.

Alors, on peut montrer qu'il existe un changement de variables $Tx = z$ telle que la forme quadratique $Q(x) = xAx^t$ s'écrit :

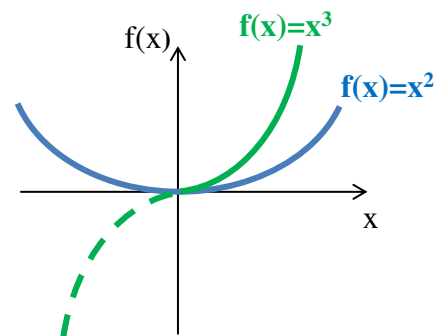
$$Q(z) = D_1 z_1^2 + \frac{D_2}{D_1} z_2^2 + \frac{D_3}{D_2} z_3^2 + \dots + \frac{D_n}{D_{n-1}} z_n^2$$

En conséquence,

- A est définie positive ($\Leftrightarrow \forall z, Q(z) > 0$) ssi $\forall i, D_i > 0$
- A est semi-définie positive ($\Leftrightarrow \forall z, Q(z) \geq 0$) si $D_n = 0$ et $\forall i < n, D_i > 0$.
- A est définie négative ($\Leftrightarrow \forall z, Q(z) < 0$) ssi $\forall i, (-1)^i D_i > 0$
à savoir tous les rapport D_n/D_{n-1} sont négatifs.
- A est semi-définie négative ($\Leftrightarrow \forall z, Q(z) \leq 0$) si $D_n = 0$ et $\forall i, (-1)^i D_i > 0$

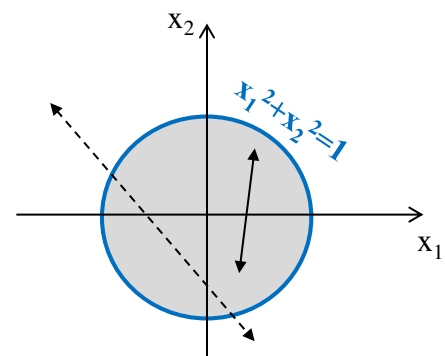
Fonction convexe

- Fonction : $f(x) = x^2$
 $f''(x) = 2 \rightarrow$ convexe sur \mathbb{R}
- Fonction : $f(x) = x^3$
 $f''(x) = 6x \rightarrow$ convexe sur \mathbb{R}^+
 \rightarrow non convexe sur \mathbb{R}



Ensemble convexe

- Ensemble : $X = \{(x_1, x_2) / x_1^2 + x_2^2 \leq 1\}$
 \rightarrow convexe
- Ensemble : $X = \{(x_1, x_2) / x_1^2 + x_2^2 \geq 1\}$
 \rightarrow non convexe



EXERCICE 14: Convexité

Déterminer les domaines du plan où les fonctions suivantes convexes ou concaves.

1. $f(x, y) = (x - 1)^2 + xy^2$.
2. $g(x, y) = \frac{x^3}{3} - 4xy + 12x + y^2$.
3. $h(x, y) = e^{-x} + e^{-y}$.
4. $k(x, y) = e^{xy}$.
5. $l(x, y) = \ln \sqrt{xy}$.

8 Conditionnement

8.1 Rappel

Rappel sur la norme d'une matrice :

- La norme induite sur les matrices carrées de taille $n \times n$ (= ensemble $\mathcal{M}_n(\mathbb{R})$) est définie par :

$$\|A\| = \sup \{ \|Ax\|; x \in \mathbb{R}^n, \|x\| = 1 \}, \forall A \in \mathcal{M}_n(\mathbb{R})$$

- cette norme vérifie les propriétés suivantes :

$$\begin{aligned} \diamond \|Ax\| &\leq \|A\| \|x\|, \forall x \in \mathbb{R}^n \\ \diamond \|A\| &= \max \left\{ \frac{\|Ax\|}{\|x\|}; x \in \mathbb{R}^n \setminus \{0\} \right\} \end{aligned}$$

- si $A \in \mathcal{M}_n(\mathbb{R})$ est inversible, alors le rayon spectral est défini comme :

$$\rho(A) = \max \{ |\lambda|; \lambda \in \mathbb{C}, \lambda \text{ valeur propre de } A \}, \text{ notée } \sigma_n$$

- la norme $\|\cdot\|_2$ sur $\mathcal{M}_n(\mathbb{R})$ est définie par :

$$\|A\|_2 = (\rho(A^t A))^{\frac{1}{2}}$$

En particulier, si A est symétrique $\|A\|_2 = \rho(A) = \sigma_n$.

- $\rho((A^t A)^{-1}) = 1/\sigma_1$ où σ_1 est la plus petite valeur propre de A .

8.2 Définition

On appelle conditionnement d'une matrice le scalaire permettant d'estimer l'erreur relative qui sera commise lorsque que l'on introduit des petites perturbations lors de l'inversion d'une matrice.

Conditionnement d'une matrice

A matrice symétrique semi-définie positive

Valeurs propres de A : $\sigma_1 \geq \dots \geq \sigma_n \rightarrow \|A\|_2 = \sigma_1$

Nombre de conditionnement de A :

$$\kappa(A) = \|A\|_2 \|A^{-1}\|_2 = \frac{\sigma_1}{\sigma_n} \geq 1$$

Conditionnement d'une fonction

f fonction deux fois différentiable

Conditionnement de f en x = nombre de conditionnement de $A=H(x)$

Interprétation

Vecteur propre d_k associé à la valeur propre σ_k : $A_k d_k = \sigma_k d_k$

Courbure de f en x dans la direction d_k : $\frac{d_k^T H(x) d_k}{d_k^T d_k} = \sigma_k$

Théorème de Rayleigh-Ritz

d_1 = direction de plus forte courbure (courbure = σ_1)

d_n = direction de plus faible courbure (courbure = σ_n)

L'interprétation géométrique est importante :

- on a vu que $d^T H_f(x) d / d^T d$ représentait la courbure de f en x dans la direction d.

- si on prend maintenant un vecteur propre d_k de H_f , on a $H_f(x)d_k = \lambda_k d_k$. En multipliant des deux côtés par d_k^T , on obtient que λ_k est la courbe de H_f dans la direction d_k .

Rappel : une valeur propre λ_k est telle que $Ax_k = \lambda_k x_k$ pour un vecteur propre $x_k \neq 0$.

Donc $(A - \lambda_k I).x_k = 0$, et comme $x_k \neq 0$, il faut donc que $\det(A - \lambda_k I) = 0$.

- l'ensemble des valeurs propres $\{\lambda_k\}$ contient les solutions de cette équation.
- le vecteur propre x_k associé à la valeur propre λ_k est obtenu en résolvant l'équation $(A - \lambda_k I).x_k = 0$ (n équations à n inconnues). La solution x_k est définie à une constante proportionnelle près.

Exemple :

Matrice 2x2

- Inverse : $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \rightarrow A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ avec $\det(A) = ad - bc$
- Valeurs propres : $\det(A - \sigma I) = 0 \Rightarrow (a - \sigma)(d - \sigma) - bc = 0 \Rightarrow \sigma^2 - (a + d)\sigma + \det(A) = 0$

Conditionnement

- $A = \begin{pmatrix} 0.1 & 1 \\ 0.20002 & 2 \end{pmatrix}$
- Valeurs propres : $\sigma^2 - 2.1\sigma - 0.00002 = 0 \Rightarrow \begin{cases} \sigma_1 = 2.10001 \\ \sigma_2 = -0.00001 \end{cases}$
- Conditionnement : $\begin{cases} \|A\|_2 = \sigma_1 \\ \|A^{-1}\|_2 = 1/\sigma_2 \end{cases} \Rightarrow \kappa(A) = \frac{\sigma_1}{\sigma_2} = 210001$

Donc la matrice A est sensible aux perturbations.

8.3 Préconditionnement

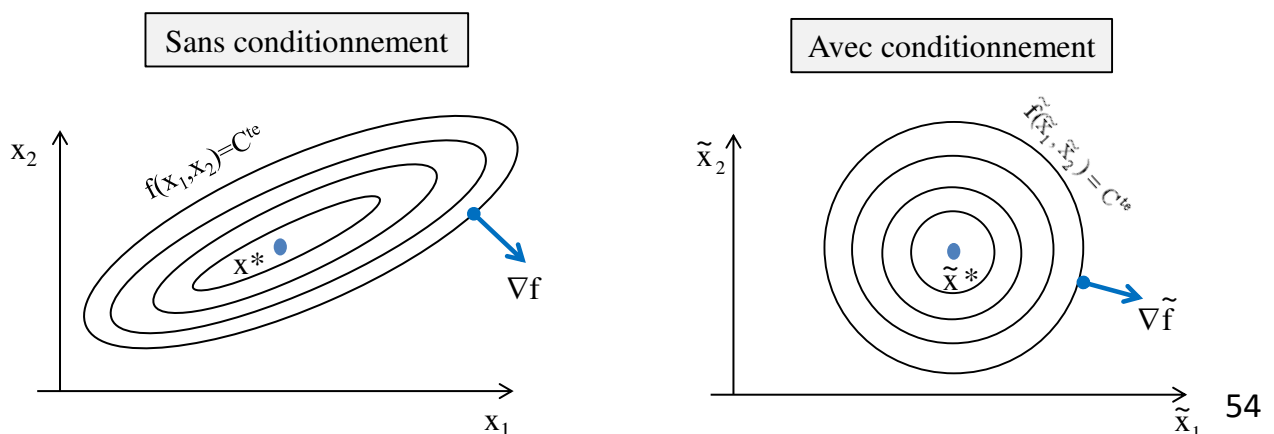
Le preconditionnement permet de modifier le conditionnement d'une matrice afin de stabiliser le comportement numérique des algorithmes.

Changement de variable

- Variable : $\tilde{x} = Mx$ ($M \in \mathbb{R}^{n \times n}$ inversible = matrice de preconditionnement)
- Fonction : $\tilde{f}(\tilde{x}) = f(x) = f(M^{-1}\tilde{x})$
- Gradient : $\nabla \tilde{f}(\tilde{x}) = M^{-T} \nabla f(M^{-1}\tilde{x}) \Rightarrow \tilde{g}(\tilde{x}) = M^{-T} g(x)$
- Hessien : $\nabla^2 \tilde{f}(\tilde{x}) = M^{-T} \nabla^2 f(M^{-1}\tilde{x}) M^{-1} \Rightarrow \tilde{H}(\tilde{x}) = M^{-T} H(x) M^{-1}$

Préconditionnement de f

- Factorisation de Cholesky (si $H(x)$ définie positive) : $H(x) = LL^T$
- Conditionnement optimal (minimal) de f en x pour : $\tilde{x} = L^T x \Rightarrow \tilde{H}(\tilde{x}) = I \Rightarrow \kappa(\tilde{H}) = 1$



Autrement dit, le conditionnement optimal est celui qui fait en sorte que les courbures soient sensiblement identiques dans toutes les directions.

8.4 Système linéaire perturbé

On regarde l'effet d'une perturbation sur le second membre d'une équation linéaire.

Perturbation du second membre

- Système non perturbé : $Ax = b \rightarrow \text{solution } x^*$
- Système perturbé au 2nd membre : $A(x + \Delta x_b) = b + \Delta b \rightarrow \text{solution } x^* + \Delta x_b$
- Système perturbé au 1^{er} membre : $(A + \Delta A)(x + \Delta x_A) = b \rightarrow \text{solution } x^* + \Delta x_A$

$$\begin{cases} Ax^* = b \\ A(x^* + \Delta x_b) = b + \Delta b \\ (A + \Delta A)(x^* + \Delta x_A) = b \end{cases} \Rightarrow \begin{cases} Ax^* = b \\ A \cdot \Delta x_b = \Delta b \\ A \cdot \Delta x_A + \Delta A \cdot x^* = 0 \end{cases} \Rightarrow \begin{cases} b = Ax^* \\ \Delta x_b = A^{-1} \Delta b \\ \Delta x_A = -A^{-1} \cdot \Delta A \cdot x^* \end{cases}$$

- Majoration de la perturbation

$$b = Ax^* \Rightarrow \|b\| \leq \|A\| \cdot \|x^*\| \Rightarrow \frac{1}{\|x^*\|} \leq \frac{\|A\|}{\|b\|}$$

$$\begin{cases} \Delta x_b = A^{-1} \Delta b \\ \Delta x_A = -A^{-1} \cdot \Delta A \cdot x^* \end{cases} \Rightarrow \begin{cases} \|\Delta x_b\| \leq \|A^{-1}\| \cdot \|\Delta b\| \\ \|\Delta x_A\| \leq \|A^{-1}\| \cdot \|\Delta A\| \cdot \|x^*\| \end{cases} \Rightarrow \begin{cases} \frac{\|\Delta x_b\|}{\|x^*\|} \leq \|A\| \cdot \|A^{-1}\| \cdot \frac{\|\Delta b\|}{\|b\|} \\ \frac{\|\Delta x_A\|}{\|x^*\|} \leq \|A\| \cdot \|A^{-1}\| \cdot \frac{\|\Delta A\|}{\|A\|} \end{cases}$$

- Amplification maximale de la perturbation : $\kappa(A) = \|A\| \cdot \|A^{-1}\|$

On retrouve bien que l'amplification maximale de la perturbation est égale au conditionnement.

Exemple : de perturbation d'un système $Ax = b$

- $\delta a = +0.001$ sur le terme a_{11} de la matrice A ,
- $\delta b = +0.01$ sur le terme b_1 du vecteur b .

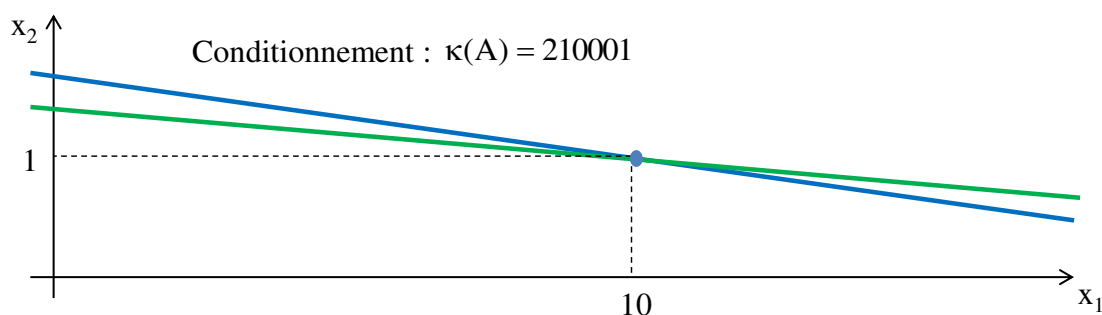
Système perturbé 2x2

- Système non perturbé

$$\begin{cases} 0.1x_1 + x_2 = 2 \\ 0.20002x_1 + 2x_2 = 4.0002 \end{cases} \Leftrightarrow \begin{pmatrix} 0.1 & 1 \\ 0.20002 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 4.0002 \end{pmatrix} \Rightarrow \begin{cases} x_1 = 10 \\ x_2 = 1 \end{cases}$$

- Perturbation δA : $A = \begin{pmatrix} 0.101 & 1 \\ 0.20002 & 2 \end{pmatrix} \Rightarrow \begin{cases} x_1 = -0.101 \\ x_2 = 2.010 \end{cases}$

- Perturbation δb : $b = \begin{pmatrix} 2.01 \\ 4.0002 \end{pmatrix} \Rightarrow \begin{cases} x_1 = -990 \\ x_2 = 101.01 \end{cases}$



8.5 Mise à l'échelle

Donc, un préconditionnement permet d'améliorer le comportement des algorithmes. Ceci se fait en général avec des mises à l'échelle.

Principe

- Des valeurs numériques trop différentes sont sources de blocage des algorithmes.
Exemple : $(1 + 10^{-20}) - 1 = 1 - 1 = 0$ au lieu de 10^{-20} avec 16 chiffres significatifs
- Pour réduire les erreurs numériques, il faut que les différentes valeurs utilisées dans les calcul aient des ordres de grandeur comparables.
- Méthode de mise à l'échelle : transformation affine $X' = \alpha X + \beta$

Quantités à mettre à l'échelle

- Variables : $x \rightarrow x' \approx 1$ (\rightarrow déplacement sur toutes les composantes de x)
- Critère : $f \rightarrow f' \approx 1$ (\rightarrow tests d'arrêt sur variation de f)
- Contraintes : $c \rightarrow c' \approx 1$ (\rightarrow contraintes de poids équivalents)
- Jacobien : $\|\nabla c\| \rightarrow \|\nabla c'\| \approx 1$ (\rightarrow directions admissibles)
- Hessien : $\|\nabla^2 L\| \rightarrow \|\nabla^2 L'\| \approx 1$ (\rightarrow courbure, conditionnement)

Difficultés

- On ne peut pas simultanément mettre toutes les quantités à l'échelle \rightarrow choix expérimental
- Le facteur d'échelle dépend du point $x \rightarrow$ à adapter au cours des itérations
(mise à l'échelle dynamique)

9 Conclusion

Nous avons seulement les notions de bases utilisées dans les méthodes d'optimisation pour des fonctions différentiables.

En particulier,

- la formulation de base du problème de l'optimisation, ainsi que la caractérisation d'une solution.
- la différentiabilité pour des fonctions multivariées de \mathbb{R}^n ,
- le théorème de Taylor permettant de définir une estimation polynomiale locale d'une surface (ordre 1 et 2),
- le problème de calcul des dérivées numériques (estimation, erreurs),
- un rappel sur quelques propriétés matricielles,
- comment caractériser la convexité pour une fonction multivariée,
- les problèmes liés au conditionnement des matrices lors d'un calcul numérique

Dans les chapitres suivants, nous utiliserons ces outils afin de faire des recherches d'optimum.

