

# Optimisation

## Chapitre 1 : Exemples et définitions

Le cours polycopié est disponible sous Moodle

Joseph GERGAUD, Serge GRATTON & Daniel RUIZ

4 septembre 2023

- Documents : les slides des cours, le cours polycopié
- 7 cours, 6 TD et 4 TP
- Evaluation : Examen écrit
- Rappel du règlement intérieur de l'ENSEEIH : plus de 30% d'absences en TD ou TP à une Unité d'Enseignement  $\Rightarrow$  pas de session 2 pour cette UE
- Les TD sont à préparer. Nous ferons passer directement des étudiants au tableau en TD.

Soit  $f$  une application d'un ensemble  $E$  dans un ensemble  $F$ . Un problème d'optimisation consiste à rechercher le meilleur élément de l'ensemble  $E$ , c'est-à-dire celui qui rend la valeur de  $f$  la plus petite possible.

$$(P) \left\{ \begin{array}{l} \min f(x) \\ x \in C \subset E. \end{array} \right.$$

- Il faut donc avoir une structure d'ordre sur  $F (= \mathbb{R})$
- $E$  s'appelle l'ensemble des stratégies, des états, des paramètres, l'espace
- $C$  est l'ensemble des contraintes
- $f$  est la fonction coût, économique ou le critère, l'objectif

---

Deux questions

- **Question 1** : Existence de solution
  - Si  $C$  est fini, c'est évident
  - Si  $C$  est infini, c'est moins trivial
- **Question 2** : Calcul de la solution
  - Si  $C$  est fini mais grand, c'est "difficile"
  - Si  $C \subset \mathbb{R}^n$  et que les fonctions sont dérivables, c'est plus facile

## 1.1. Introduction

## 1.2. Exemples

1.2.1. Cas Continu et de dimension finie

1.2.2. Problèmes en nombres entiers

1.2.3. Problèmes en dimension infinie

## 1.3. Problème d'optimisation

1.3.1. Définition

1.3.2. Classification

## 1.4. Plan

## 1.5. Exercices à faire pour le TD



- Pierre de Fermat, 1601  
(Beaumont-de-Lomagne, près de Montauban) – 1665 (Castres)
- méthode, dite de maximis et minimis
- précurseur du calcul différentiel
- Principe de Fermat : la lumière se propage d'un point à un autre sur des trajectoires telles que la durée du parcours soit minimale

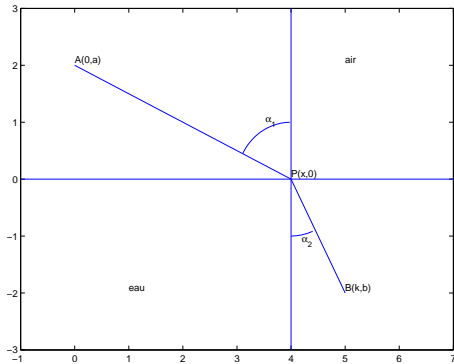


FIGURE 1 – Principe de Fermat.

On s'intéresse ici à la trajectoire d'un rayon lumineux d'un point  $A(0, a)$  vers un point  $B(k, b)$  situés dans deux milieux homogènes différents (cf. la figure 1). Nous allons grâce au principe de Fermat retrouver la loi de la réfraction. On suppose pour cela que la trajectoire d'un rayon lumineux dans un milieu homogène est un segment de droite (ce qui peut aussi se démontrer grâce au principe de Fermat via le calcul des variations qui est un problème d'optimisation en dimension infinie!).

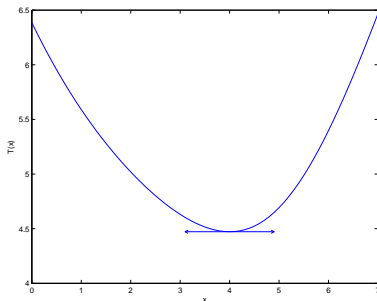
On note  $P$ , de coordonnées  $(x, 0)$ , le point d'impact du rayon lumineux sur la surface du changement de milieu et  $c_1$  et  $c_2$  les vitesses de la lumière dans l'air et dans l'eau. Le temps de parcours entre les point  $A$  et  $B$  est donc (temps = distance/vitesse)

$$T(x) = \frac{1}{c_1} \sqrt{a^2 + x^2} + \frac{1}{c_2} \sqrt{b^2 + (k - x)^2}.$$

Le problème est alors ici de trouver le point  $P$  (c'est-à-dire  $x^* \in \mathbb{R}$ ) tel que

$$T(x^*) \leq T(x) \forall x \in \mathbb{R} \iff (P) \left\{ \begin{array}{l} \text{Min} \\ x \in \mathbb{R}. \end{array} T(x) \right.$$

On peut ici tracer cette fonction (cf. la figure 2).





Un calcul simple donne

$$T'(x) = \frac{x}{c_1 \sqrt{a^2 + x^2}} + \frac{-(k-x)}{c_2 \sqrt{b^2 + (k-x)^2}}$$

$$T''(x) = \frac{a^2}{c_1 (a^2 + x^2)^{3/2}} + \frac{b^2}{c_2 (b^2 + (k-x)^2)^{3/2}}$$

On en déduit le tableau de variation

$x$	$-\infty$	$x^*$	$+\infty$
$T''(x)$	+		+
$T'(x)$	-	0	+
$T(x)$	$+\infty$	$T(x^*)$	$+\infty$

La solution est donc l'unique point  $x^*$  qui vérifie  $T'(x^*) = 0$  soit

$$\begin{aligned}T'(x) &= \frac{x}{c_1 \sqrt{a^2 + x^2}} + \frac{-(k-x)}{c_2 \sqrt{b^2 + (k-x)^2}} = 0 \\ \iff \frac{x}{c_1 \sqrt{a^2 + x^2}} &= \frac{(k-x)}{c_2 \sqrt{b^2 + (k-x)^2}} \\ \iff \frac{\sin \alpha_1}{c_1} &= \frac{\sin \alpha_2}{c_2} \\ \iff n_1 \sin \alpha_1 &= n_2 \sin \alpha_2.\end{aligned}$$

**Remarque 1.2.1.** Nous retrouvons dans ce cas les lois de Descartes<sup>1</sup> ou de Snell.

**Remarque 1.2.2.** Ici une condition nécessaire et suffisante de solution de  $(P)$  est  $T'(x) = 0$ . Ceci est dû au fait que la fonction  $T$  est convexe (sa dérivée seconde est positive ou nulle en tout point de  $\mathbb{R}$ ).

---

1. Associer les noms de Fermat et Descartes est surprenant pour qui connaît les confrontations scientifiques virulentes qui les opposèrent. Les étudiants intéressés peuvent voir la vidéo ([?]) où se rendre au musée Pierre de Fermat de Beaumont de Lomagne, ville natale de P. de Fermat près de Toulouse.

**Remarque 1.2.3.** En général la condition  $T'(x) = 0$  n'est qu'une condition nécessaire. Si on prend par exemple  $f(x) = x^3$  nous avons  $f'(0) = 0$  mais 0 n'est pas un minimum de  $f$  (cf.l'exemple Figure 3).

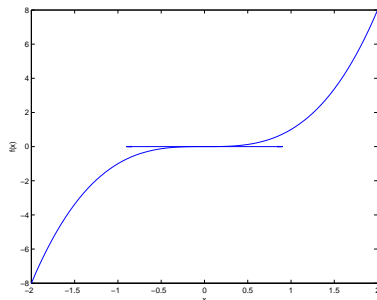


FIGURE 3 –  $f'(0) = 0$  et 0 n'est pas un minimum.

Nous verrons au chapitre 5 le cas général pour une fonction de  $\mathbb{R}^n$  à valeur dans  $\mathbb{R}$ .

Le carbone radioactif  $^{14}\text{C}$  est produit dans l'atmosphère par l'effet des rayons cosmiques sur l'azote atmosphérique. Il est oxydé en  $^{14}\text{CO}_2$  et absorbé sous cette forme par les organismes vivants qui, par suite, contiennent un certain pourcentage de carbone radioactif relativement aux carbones  $^{12}\text{C}$  et  $^{13}\text{C}$  qui sont stables. On suppose que la production de carbone  $^{14}\text{C}$  atmosphérique est demeurée constante durant les derniers millénaires.

On suppose d'autre part que, lorsqu'un organisme meurt, ses échanges avec l'atmosphère cessent et que la radioactivité due au carbone  $^{14}\text{C}$  décroît suivant la loi exponentielle suivante :

$$A_{(A_0, \lambda)}(t) = A_0 e^{-\lambda t}$$

l'analyse les troncs (le bois est un tissu mort) de très vieux arbres *Sequoia gigantea* et *Pinus aristata*. Par un prélèvement effectué sur le tronc, on peut obtenir :

- son âge  $t$  en année, en comptant le nombre des anneaux de croissance,
- sa radioactivité  $A$  en mesurant le nombre de désintégration.

$t_i$	500	1000	2000	3000	4000	5000	6300
$A_i$	14.5	13.5	12.0	10.8	9.9	8.9	8.0

TABLE 1 – Données.

Notre but est ici de trouver les valeurs des paramètres  $A_0$  et  $\lambda$  pour que la fonction  $A_{(A_0, \lambda)}(t)$  "colle" au mieux aux données.



Ici les instants  $t_i$  et les valeurs  $A_i$ , pour  $i = 1, \dots, 7$  sont connus. Ce sont les valeurs des paramètres  $A_0$  et  $\lambda$  que l'on cherche.

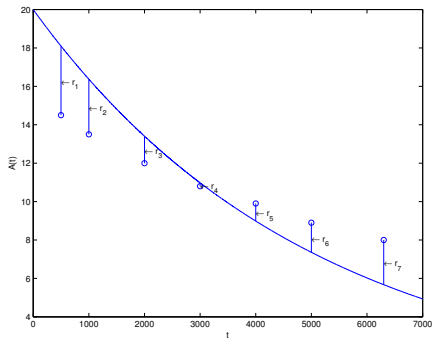


FIGURE 4 – Résidus  $r(20, 0.0002)$  pour le problème de datation par le carbone 14.

Si on donne des valeurs aux paramètres, nous pouvons calculer les quantités appelées résidus (cf.. la Fig. 4 pour les valeurs des paramètres  $A_0 = 20$  et  $\lambda = 0.0002$ )

$$r_i(A_0, \lambda) = A_i - A_{(A_0, \lambda)}(t_i) = A_i - A_0 e^{-\lambda t_i}.$$

Par suite nous pouvons calculer la quantité

$$f(A_0, \lambda) = \frac{1}{2} \sum_{i=1}^n (A_i - A_0 e^{-\lambda t_i})^2.$$

Cette quantité est la somme des carrés des longueurs des résidus.

Plus cette quantité sera faible, plus notre courbe sera proche de nos points expérimentaux.

Estimer les paramètres  $A_0$  et  $\lambda$  par les moindres carrés, c'est rechercher la valeur solution du problème d'optimisation suivant :

$$(P) \begin{cases} \text{Min } f(A_0, \lambda) = \frac{1}{2} \sum_{i=1}^n (A_i - A_0 e^{-\lambda t_i})^2 \\ (A_0, \lambda) \in \mathbb{R}^2. \end{cases}$$

#### Remarque 1.2.4.

- Dans l'exemple précédent on peut aussi écrire :  $f(\beta) = \frac{1}{2} \| r(\beta) \|^2$  où

$$\begin{aligned} r: \quad \mathbb{R}^2 &\longrightarrow \mathbb{R}^7 \\ \beta = (A_0, \lambda) &\longmapsto \begin{pmatrix} r_1(\beta) \\ \vdots \\ r_7(\beta) \end{pmatrix} \end{aligned}$$

avec

$$r_i(\beta) = A_i - A_0 e^{-\lambda t_i}$$

et où  $\| \cdot \|$  est la norme euclidienne.

- Minimiser  $f(\beta)$  est équivalent à minimiser  $\alpha f(\beta)$  avec  $\alpha > 0$ . Le terme  $\frac{1}{2}$  est mis ici afin de ne pas avoir le terme 2 lorsque l'on dérive la fonction  $f(\beta)$ .

### Définition 1.2.1

On appelle problème aux moindres carrés tout problème qui s'écrit

$$(P) \begin{cases} \text{Min } f(\beta) = \frac{1}{2} \|r(\beta)\|^2 \\ \beta \in \mathbb{R}^p. \end{cases}$$

où  $r$  est une fonction de  $\mathbb{R}^p$  à valeurs dans  $\mathbb{R}^n$ .

le problème aux moindres carrés est dit linéaire si  $r$  est une fonction affine :  $r(\beta) = y - X\beta$ .

**Remarque 1.2.5.** On peut aussi prendre comme critère :

- $f(\beta) = \|r(\beta)\|_1 = \sum_{i=1}^n |r_i(\beta)|$ ;
- $f(\beta) = \|r(\beta)\|_\infty = \text{Max}_{i=1,\dots,n} |r_i(\beta)|$ .

Mais ce n'est plus un critère aux moindres carrés. La résolution du problème d'optimisation est alors plus compliquée car nous n'avons plus la dérivabilité.



**Exercice 1.2.1. Régression linéaire simple**

Soit  $n$  points expérimentaux  $M_i = (x_i, y_i)$  pour  $i = 1, \dots, n$ . On considère le modèle suivant :  $y(x, \beta) = \beta_0 + \beta_1 x$ .

1. On veut estimer les paramètres par les moindres carrés. Écrire le problème sous la forme :

$$\begin{cases} \text{Min} & f(\beta) = \frac{1}{2} \|r(\beta)\|^2 = \frac{1}{2} \|y - X\beta\|^2 \\ & \beta \in \mathbb{R}^p \end{cases}$$

On donnera les valeurs de  $X$  et de  $y$  et à quoi correspond  $\beta$ .

2. On souhaite maintenant trouver la meilleure droite au sens des moindres carrés qui passe par l'origine. Écrire le problème d'optimisation.

**Exercice 1.2.2** (Courbe étalon). On considère le modèle suivant :

$$y(x, \beta) = \beta_2 + \frac{\beta_1 - \beta_2}{(1 + \exp(\beta_3 + \beta_4 x))^{\beta_5}}. \quad (1)$$

Dose en ng/.1 ml	Réponse en c.p.m.			
0	2868	2785	2849	2805
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
100	131	135	134	133

TABLE 2 – Données pour un dosage de Cortisol

On désire estimer les paramètres par les moindres carrés (attention, il y a pour chaque dose 4 observations de  $y$ ). On notera  $(x_i)_{i=1,\dots,16}$  (respectivement  $(y_{i,j})_{i=1,\dots,16;j=1,\dots,4}$ ) les éléments de la première colonne (respectivement des 4 dernières colonnes) de la table 2 et  $r_{i,j}(\beta)$  le résidu liés au point  $(x_i, y_{i,j})$ .

1. Écrire le résidu lié au point  $(0.04, 2378)$ .
2. 1) Quelle est la dimension du vecteur des paramètres  $\beta$ .  
2) Quel est le nombre de points  $n$ ?
3. Écrire le problème d'optimisation des paramètres par les moindres carrés.

Un fermier désire déterminer les quantités de lisier de porc et d'engrais composé à étendre sur 20 ha de prairie de façon à optimiser le coût total de la fertilisation. Le coût et la composition du lisier et de l'engrais sont donnés à la table 3.

	coût (par tonne)	composition chimique ( $kg t^{-1}$ )		
		azote	phosphate	potasse
lisier	25 francs	6	1.5	4
engrais	1300 francs	250	100	100

TABLE 3 – Coûts et compositions des engrais

Le fermier veut appliquer au moins  $75\text{ }kg ha^{-1}$  d'azote,  $25\text{ }kg ha^{-1}$  de phosphate et  $35\text{ }kg ha^{-1}$  de potasse. Il ne peut appliquer le lisier qu'à un taux maximum de  $8\text{ t/heure}$  et l'engrais qu'à un taux maximum de  $0.4\text{ t/heure}$ . Il ne peut de plus consacrer pour ce travail qu'un maximum de 25 heures.

Appelons  $x_1$  (respectivement  $x_2$ ) la quantité en tonnes de lisier, (respectivement d'engrais) étendu. Le problème à résoudre est alors :

$$(P) \left\{ \begin{array}{ll} \text{Min } f(x) = 25x_1 + 1300x_2 & \text{coût à minimiser} \\ 6x_1 + 250x_2 \geq 75 \times 20 = 1500 & \text{contrainte sur l'azote} \\ 1.5x_1 + 100x_2 \geq 500 & \text{contrainte sur le phosphate} \\ 4x_1 + 100x_2 \geq 700 & \text{contrainte sur la potasse} \\ (1/8)x_1 + (1/0.4)x_2 \leq 25 & \text{contrainte de temps} \\ x_1 \geq 0 & \text{contrainte à ne pas oublier} \\ x_2 \geq 0. & \text{contrainte à ne pas oublier} \end{array} \right.$$

Un alpiniste veut mettre dans son sac à dos un maximum de 16 kg de ravitaillement. Il peut choisir un certain nombre d'unités de trois produits différents. Le poids unitaire en kilogrammes et la valeur énergétique unitaire des ces produits sont connus et donnés dans la table (4).

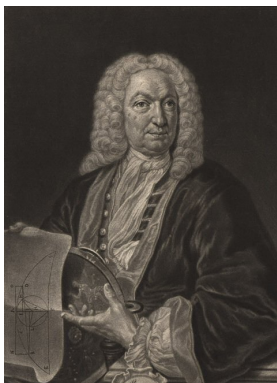
Produits	I	II	III
Poids	2	5	7
Valeurs	4	10	15

TABLE 4 – Poids unitaires et valeurs énergétiques unitaires.

Le problème pour l'alpiniste est de savoir ce qu'il doit emporter pour avoir une valeur totale en calories maximale sans dépasser les 16 kg.

Si nous notons  $x_1$ ,  $x_2$  et  $x_3$  les nombres d'unités à emporter des articles I, II et III, le problème s'écrit

$$(P) \begin{cases} \text{Max} & 4x_1 + 10x_2 + 15x_3 \\ & 2x_1 + 5x_2 + 7x_3 \leq 16 \\ & (x_1, x_2, x_3) \in \mathbb{N}^3 \end{cases} \quad \text{Les variables sont entières.}$$



Jean Bernoulli 27 juillet 1667 – 1er janvier 1748 lança un défi : Trouver dans un plan vertical du chemin reliant 2 points  $P_0$  et  $P_f$  de ce plan, suivant lequel un corps  $M$  entraîné par son propre poids effectuera le trajet de  $P_0$  à  $P_f$  en un temps minimum. On suppose qu'il n'y a pas de frottement.

La solution fut trouvée par 5 personnes :

- Les frères Bernoulli
- Leibniz
- L'Hôpital
- Un anonyme dont Bernoulli dit en lisant sa contribution : "On reconnaît le lion à sa griffe". Il s'agissait de Newton !

- $v = \sqrt{2g(-y(x))}$
- $ds = \sqrt{dx^2 + dy^2}$  et  
 $t = ds / \sqrt{2g(-y(x))}$

$$T : C^1([0, x_f], \mathbb{R}) \rightarrow \mathbb{R}$$

$$T(y(.)) = \int_0^{x_f} \frac{\sqrt{1 + y'(x)^2}}{\sqrt{2g(-y(x))}} dx$$

- Le problème d'optimisation est ici en **dimension infinie**, l'inconnue  $y$  est une fonction dans  $C^1([0, x_f], \mathbb{R})$

$$(P) \begin{cases} \text{Min } T(y(.)) \\ y(0) = 0 \\ y(x_f) = y_f. \end{cases}$$

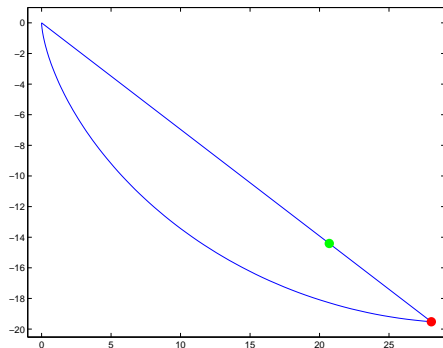
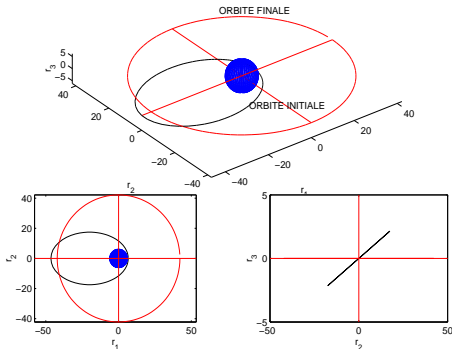


FIGURE 5 – Cycloïde inversée.





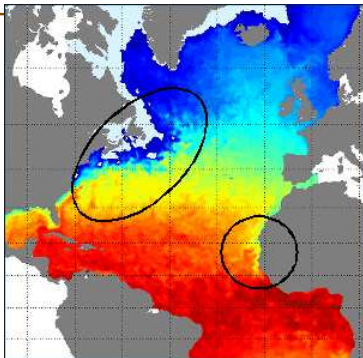
- Orbite de départ LEO :  $P = 11.625$  Mm,  $e = 0.75$  and  $i = 7^\circ$
- Orbite finale : orbite géostationnaire
- Masse initiale :  $m_0 = 1500$  kg
- Poussée :  $T_{\max} = 0.1$  N



Problème de contrôle optimal, l'inconnue est la poussée du moteur à chaque instant, donc une fonction  $u$  de  $\mathbb{R}$  à valeur dans  $\mathbb{R}^3$ .

$$(P) \left\{ \begin{array}{l} \text{Min } t_f = \int_0^{t_f} dt \quad t_f \text{ libre} \\ \dot{r}(t) = v(t) \quad pp. \\ \dot{v}(t) = -\frac{\mu r(t)}{\|r(t)\|^3} + \frac{T_{\max}}{m(t)} u(t) \\ \dot{m}(t) = -\beta T_{\max} \|u(t)\| \\ (r(t), v(t), m(t)) \in A \\ \|u(t)\| \leq 1 \\ r(0), v(0), m(0) \text{ fixé} \\ r(t_f), v(t_f) \text{ fixé,} \end{array} \right.$$

$$(P) \left\{ \begin{array}{l} \text{Min } \int_0^{t_f} \|u(t)\| dt \quad t_f \text{ fixé} \\ \dot{r}(t) = v(t) \quad pp. \\ \dot{v}(t) = -\frac{\mu r(t)}{\|r(t)\|^3} + \frac{T_{\max}}{m(t)} u(t) \\ \dot{m}(t) = -\beta T_{\max} \|u(t)\| \\ (r(t), v(t), m(t)) \in A \\ \|u(t)\| \leq 1 \\ r(0), v(0), m(0) \text{ fixé} \\ r(t_f), v(t_f) \text{ fixé,} \end{array} \right.$$



température des océans

## ■ Applications

- Prévision Météorologique
- Sismique, pétrole
- Fusion nucléaire
- Médecine
- Agronomie
- ...

## ■ Méthode

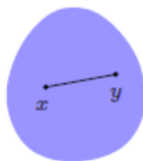
- Connaissance à priori sur l'état initial d'un système : ébauche
- Des observations ( $10^5$  données)
- Un modèle mathématique
- Trouver l'état initial ( $10^7$  inconnues) le plus proche de l'ébauche et qui "colle" au mieux aux données
- Problème d'optimisation de très grandes tailles
- Difficultés d'ordres mathématiques, numériques et informatiques

- Exemple traité en parcours HPC % Big Data en 3<sup>e</sup> année (parcours commun avec l'École Nationale de la Météorologie)

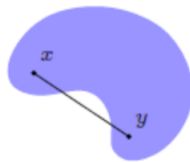
- Voir les autres exemple en génétique, affectation de ressource, ... dans le cours polycopié.
- En Intelligence artificiel : la phase d'apprentissage dans un réseaux de neurones n'est pas autre chose que de résoudre un problème d'optimisation (voir TD pour un cas simple).

**Définition 1.3.1 – Ensemble convexe**

Un sous ensemble  $C$  d'un espace vectoriel est dit convexe si pour tout  $(x, y) \in C^2$  le segment  $[x, y] = \{\alpha x + (1 - \alpha)y, \alpha \in [0, 1]\}$  est inclus dans  $C$ .



convexe



non convexe

FIGURE 6 – Ensemble convexe et non convexe.

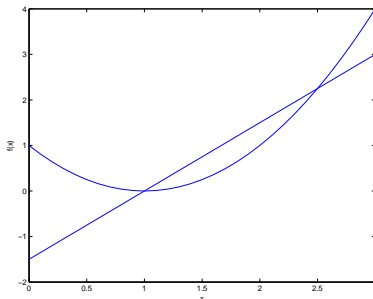
**Définition 1.3.2 – Fonction convexe**

Une fonction  $f$  de  $C \subset E$  à valeurs dans  $\mathbb{R}$ ,  $E$  espace vectoriel, est convexe si et seulement si elle vérifie :

- 1)  $C$  est convexe ;
- 2)

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

Dans le cas  $n = 1$ , ceci signifie que le graphe de la fonction  $f$  est toujours sous la corde.



**Définition 1.3.3 – Problème d'optimisation sans contraintes**

On appelle problème d'optimisation sans contraintes en dimension finie tout problème  $(P)$  consistant en la recherche d'un minimum d'une fonctionnelle  $f$  définie sur  $\mathbb{R}^n$ . On notera ce problème sous la forme suivante :

$$(P) \left\{ \begin{array}{l} \text{Min } f(x) \\ x \in \mathbb{R}^n \end{array} \right.$$

où  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  sera donnée.

**Remarque 1.3.1.** Résoudre le problème  $(P)$  revient à rechercher le point  $x^*$  de  $\mathbb{R}^n$  tel que  $f(x^*) \leq f(x) \forall x \in \mathbb{R}^n$ .

**Remarque 1.3.2.** Un problème de maximisation se ramène très facilement à un problème de minimisation :

$$\text{Max } f(x) \iff \text{Min } (-f(x))$$

**Définition 1.3.4 – Problème d'optimisation avec contraintes**

On appelle problème d'optimisation avec contraintes tout problème  $(P)$  consistant en la recherche d'un minimum sur un ensemble  $C$  inclus dans  $\mathbb{R}^n$  d'une fonctionnelle  $f$  définie sur  $\mathbb{R}^n$ . On notera ce problème sous la forme suivante :

$$(P) \left\{ \begin{array}{l} \text{Min } f(x) \\ x \in C \subset \mathbb{R}^n \end{array} \right.$$

où  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  sera donnée.

**Remarque 1.3.3.** Dans la pratique  $C$  sera défini de la façon suivante :

$$C = \{x \in \mathbb{R}^n / g_i(x) \leq 0 \quad i = 1, \dots, m \text{ et } h_l(x) = 0 \quad l = 1, \dots, p\} \quad (2)$$

et nous écrirons  $(P)$  sous la forme

$$(P) \left\{ \begin{array}{l} \text{Min } f(x) \\ g_i(x) \leq 0 \quad i = 1, \dots, m \\ h_l(x) = 0 \quad l = 1, \dots, p \end{array} \right.$$

### Définition 1.3.5 – Optimisation non différentiable

On appelle problème d'optimisation non différentiable un problème d'optimisation où les fonctions qui interviennent ne sont pas dérivables.

**Remarque 1.3.4.** On ne traitera dans ce cours que des problèmes d'optimisation différentiables.

### Définition 1.3.6 – Problème d'optimisation convexe

Un problème d'optimisation est dit convexe si et seulement si la fonction  $f$  est convexe et l'ensemble des contrainte  $C$  est convexe.

**Remarque 1.3.5.** Si  $C$  est définie par (2) et si les fonctions  $g_i$  sont convexes et les fonctions  $h_l$  sont affines, alors  $C$  est convexe. Attention, la réciproque est fausse.



**Définition 1.3.7 – Problème aux moindres carrés**

On appelle problème aux moindres carrés un problème d'optimisation sans contraintes où la fonctionnelle  $f$  est de la forme suivante :

$$f(\beta) = \frac{1}{2} \|r(\beta)\|^2 = \frac{1}{2} (r(\beta) | r(\beta)) = \frac{1}{2} \sum_{i=1}^n r_i^2(\beta)$$

Le problème est dit aux moindres carrés linéaires si la fonction  $r$  est affine :

$$\begin{aligned} r : \mathbb{R}^p &\longrightarrow \mathbb{R}^n \\ \beta &\longmapsto y - X\beta \end{aligned}$$

où  $X$  matrice de type  $(n, p)$  et  $y$  un élément de  $\mathbb{R}^n$ .

**Remarque 1.3.6.** L'exemple *datation par le carbone 14* (page 12) est un problème aux moindres carrés non linéaire.

**Définition 1.3.8 – Problème linéaire**

Un problème d'optimisation est dit linéaire si et seulement si les fonctions  $f$ ,  $g_i$ , et  $h_i$  sont affines.

$$(P_L) \begin{cases} \text{Min } \langle c, x \rangle \\ Ax \leq b \\ Cx = d \end{cases}$$

**Remarque 1.3.7.** L'exemple de page 19 est un problème linéaire.

**Définition 1.3.9 – Optimum global, optimum local**

Soit  $(P)$  un problème d'optimisation sans contraintes.

- 1)  $x^*$  est un minimum global  $\iff x^*$  est la solution de  $(P)$
- 2)  $x^*$  est un minimum local faible  $\iff$  il existe  $\varepsilon > 0$  tel que  $x^*$  est la solution de  $(P')$  où

$$(P') \left\{ \begin{array}{l} \text{Min } f(x) \\ \|x - x^*\| < \varepsilon \end{array} \right.$$

- 3)  $x^*$  est un minimum local fort si

$$\forall \mathbf{x} \in B(\mathbf{x}^*, \varepsilon) = \{\mathbf{x} \in \mathbb{R}^n / \|\mathbf{x} - \mathbf{x}^*\| < \varepsilon\}, \mathbf{x} \neq \mathbf{x}^*, \quad f(\mathbf{x}^*) < f(\mathbf{x}) .$$

Dans le cas où  $n = 1$   $\|x - x^*\| < \varepsilon \iff |x - x^*| < \varepsilon \iff x^* - \varepsilon < x < x^* + \varepsilon$ ,

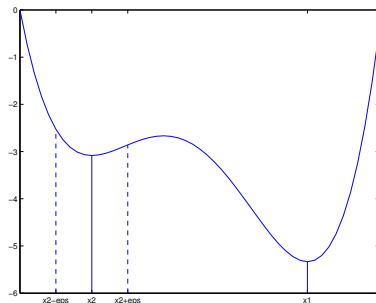
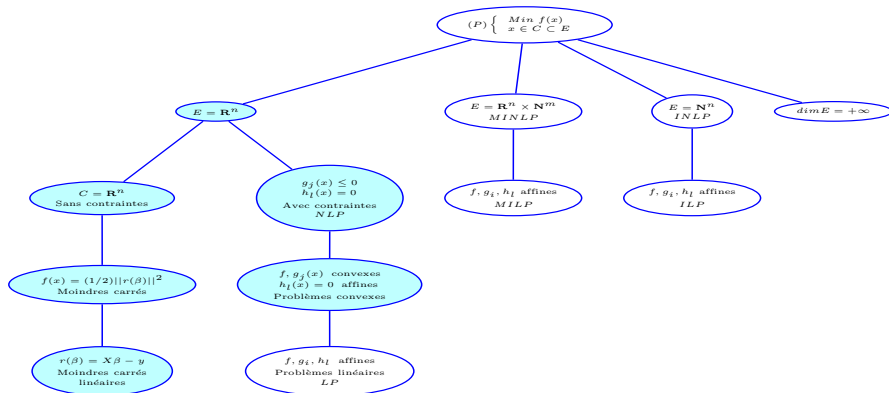


FIGURE 8 –  $x^2$  est un minimum local fort,  $x^1$  est un minimum global

### Remarque 1.3.8.

- On dit que  $x^*$  est un minimum alors que c'est  $f(x^*)$  qui est un minimum. Il s'agit d'un abus de langage que nous emploierons systématiquement.
- On appelle optimisation globale la recherche d'un optimum global
- Un algorithme globalement convergent est lui un algorithme qui converge vers un minimum local quel que soit le point de départ.



- i) Compléments de mathématiques :
  - Formes quadratiques ;
  - dérivée ;
  - convexité.
- ii) Existence de solution
- iii) Condition nécessaire, condition suffisante
- iv) Problèmes aux moindres carrés
  - Problèmes aux moindres carrés linéaires ;
  - Algorithme de Newton ;
  - Algorithme de Gauß–Newton.

- Voir sous Gitlab <https://gitlab.irit.fr/toc/etu-n7/optimisation> le TD1
- Vous cherchez les exercices pour le TD
- On fait directement passer des étudiants au tableau