

Informatique & Mathématiques Appliquées
Programmation Mathématique et Application



J. Gergaud & D. Ruiz

17 avril 2008

Table des matières

1	Définition du problème	3
1	Exemples	3
1.1	Cas continu et de dimension finie	3
1.2	Problèmes en nombres entiers	10
1.3	Problème en dimension infinie	11
2	Problème d'optimisation	13
2.1	Définitions	13
2.2	Classification	15
2	Condition nécessaire, condition suffisante : cas sans contraintes et cas convexe	17
1	Introduction	17
2	Convexité des fonctionnelles	17
2.1	Ensembles convexes - fonctionnelles convexes	17
2.2	Convexité et dérivée première	19
2.3	Convexité et dérivée seconde	20
3	Extrémas de fonctionnelles	20
3.1	Prise en compte des dérivées premières	20
3.2	Prise en compte des dérivées secondes	20
3.3	Prise en compte de la convexité	21
3.4	Existence d'un minimum	21

Introduction

Optimiser, c'est rechercher parmi un ensemble C de choix possibles le meilleur (s'il existe!). Si f est une application d'un ensemble E dans F . On note le problème

$$(P) \begin{cases} \min f(x) \\ x \in C \subset E. \end{cases}$$

Il faut donc pour cela pouvoir comparer 2 choix et donc avoir une structure d'ordre sur l'ensemble F . On prendra toujours $F = \mathbf{R}$. Suivant les domaines d'applications :

- 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.

Une fois le problème bien défini, il se pose deux questions. La première est de savoir si (P) admet une solution. Si la réponse est positive, il nous faut trouver la ou les solutions. Suivant la nature de l'ensemble E les réponses sont plus ou moins faciles. Si E est fini, l'existence de solution est évidente, mais le calcul est difficile si le nombre d'éléments est grand. Par contre si $E = \mathbf{R}^n$ ou est de dimension infinie la question de l'existence de solution est moins triviale, mais si les fonctions sont dérivables il est "plus" facile de calculer la solution.

Chapitre 1

Définition du problème

1 Exemples

1.1 Cas continu et de dimension finie

Exemple 1.1.1. [Principe de Fermat] Pierre de Fermat est un juriste et mathématicien français, surnommé « le



FIG. 1.1 – Pierre de Fermat, né vers 1601, à Beaumont-de-Lomagne, près de Montauban, et mort le 12 janvier 1665 à Castres.

prince des amateurs ». On lui doit entre autre le principe de Fermat qui dit que la lumière se propage d'un point à un autre sur des trajectoires telles que la durée du parcours soit minimale. Il imagina aussi pour la solution des problèmes, une méthode, dite de maximis et minimis, qui le fait regarder comme le premier inventeur du calcul différentiel dont il est un précurseur : il est le premier à utiliser la formule (sinon le concept) du nombre dérivé!^{*}

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 1.2). 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 et à l'optimisation!).

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

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

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

$$T(x^*) \leq T(x) \forall x \in \mathbf{R} \iff (P) \begin{cases} \text{Min } f(x) \\ x \in \mathbf{R}. \end{cases}$$

^{*}http://fr.wikipedia.org/wiki/Pierre_de_Fermat.

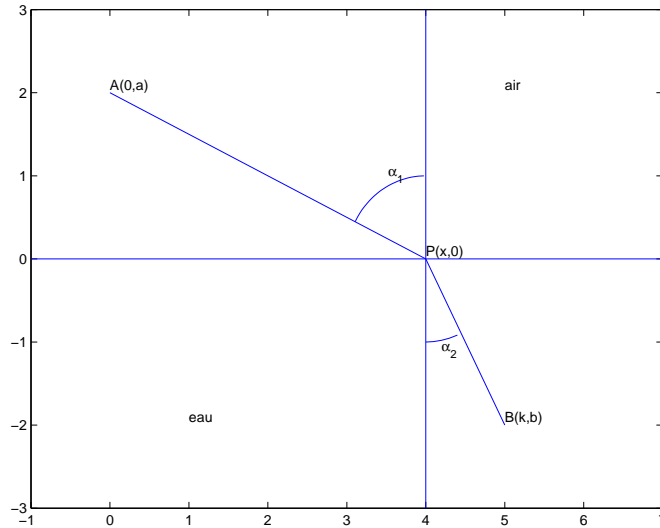
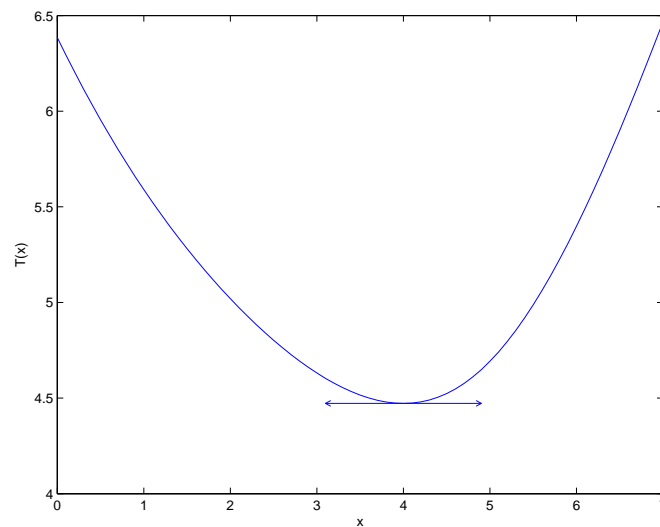


FIG. 1.2 – Principe de Fermat.

On peut ici tracer cette fonction (cf. la figure 1.3).

FIG. 1.3 – Fonction T .

Une condition nécessaire de solution de (P) est $T'(x) = 0$ (cf. la figure 1.3). Ce qui donne ici

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

Remarque 1.1.2. Nous retrouvons dans ce cas les lois de Descartes[†] ou de Snell.

[†] 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 ([3]) où se rendre au musée Pierre de Fermat de Beaumont de Lomagne, ville natale de P. de Fermat près de Toulouse.

Remarque 1.1.3. La condition $T'(x) = 0$ n'est qu'une condition nécessaire, en effet si nous considérons la fonctionnelle réelle $f(x) = x^3$ nous avons $f'(0) = 0$ mais 0 n'est pas un minimum de f (1.4).

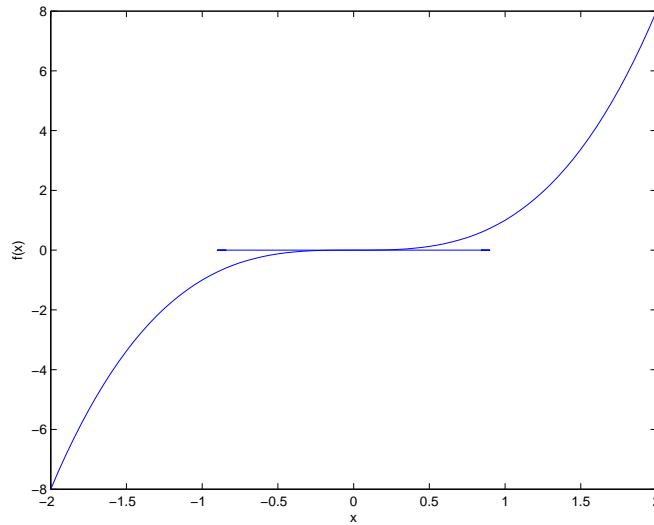


FIG. 1.4 – $f'(0) = 0$ et 0 n'est pas un minimum.

Exemple 1.1.4. Un condensateur chargé à une tension de U_0 volts se décharge sur une résistance. On mesure la tension U entre les armatures du condensateur toutes les secondes pendant un intervalle de temps de 10 secondes. Les résultats des mesures sont données dans la table (1.1).

t_i	U_i	t_i	U_i
0	100	6	15
1	75	7	10
2	55	8	10
3	40	9	5
4	30	10	5
5	20		

TAB. 1.1 – Données.

Théoriquement, la tension en fonction du temps s'écrit

$$U(t) = Ae^{-\alpha t}.$$

On désire ici estimer les valeurs des constantes A et α . Notre but est donc de trouver les valeurs de ces constantes pour que cette fonction $U(t)$ "colle" au mieux à nos données. Si on donne des valeurs à ces constantes, nous pouvons calculer les quantités appelées résidus (Fig. 1.5)

$$r_i(A, \alpha) = U_i - U(t_i) = U_i - Ae^{-\alpha t_i}.$$

Par suite nous pouvons calculer la quantité

$$f(A, \alpha) = \frac{1}{2} \sum_{i=1}^n (U_i - Ae^{-\alpha t_i})^2.$$

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 et α par les moindres carrés, c'est rechercher la valeur solution du problème d'optimisation suivant :

$$(P) \begin{cases} \text{Min} f(A, \alpha) = \frac{1}{2} \sum_{i=1}^n (U_i - Ae^{-\alpha t_i})^2 \\ (A, \alpha) \in \mathbf{R}^2. \end{cases}$$

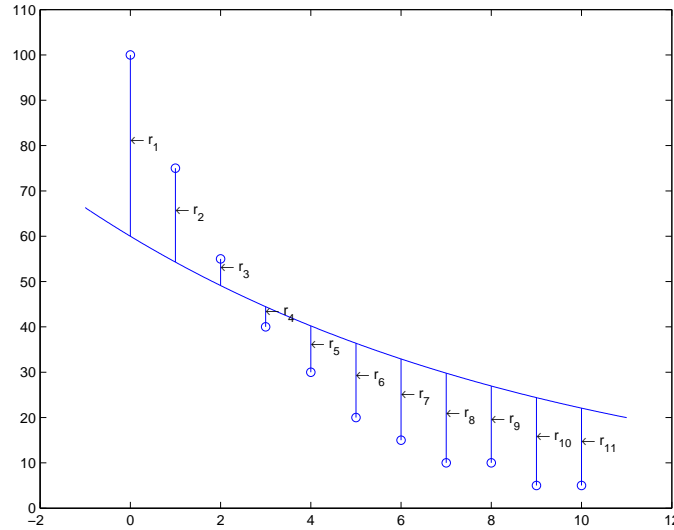


FIG. 1.5 – Critères des moindres carrés.

Attention dans le problème (P) ci-dessus, les instants t_i et les valeurs U_i sont connus. Ce sont les valeurs des paramètres que l'on cherche.

Remarque 1.1.5. • Dans l'exemple précédent on peut aussi écrire : $f(\beta) = \frac{1}{2} \| r(\beta) \|^2$ où

$$\beta = (A \quad \alpha), \quad r(\beta) = \begin{pmatrix} r_1(\beta) \\ \vdots \\ r_n(\beta) \end{pmatrix}, \quad r_i(\beta) = U_i - Ae^{-\alpha t_i} \quad \text{et } \|\cdot\| \text{ est la norme euclidienne.}$$

- 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 $f(\beta)$.
- On peut aussi prendre comme critère :
 - $f(\beta) = \|r(\beta)\|_1 = \sum_{i=1}^n |r_i(\beta)|$;
 - $f(\beta) = \|r(\beta)\|_\infty = \text{Max}_{i=1, \dots, n} |r_i(\beta)|$.

Exemple 1.1.6 (Modèle de Kaplan). On désire étudier la diffusion d'une drogue dans un organe d'un corps donné. La drogue est injectée par intraveineuse dans le sang à l'instant $t_0 = 0$. On modélise le système par un modèle à compartiments (cf. la figure 1.6).

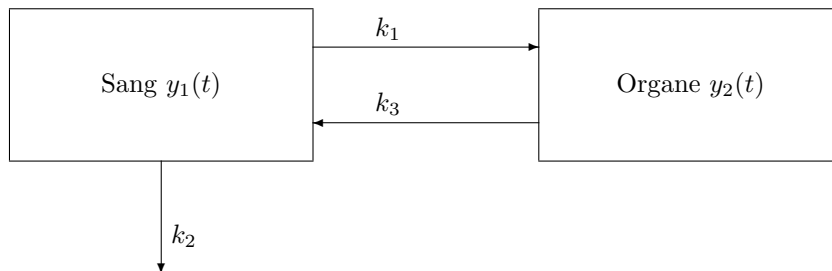


FIG. 1.6 – Modèle par compartiments.

Les concentrations dans le sang, mesurées à différents instants, sont données à la table 1.2.

t_i	y_{i1}	t_i	y_{i1}
0.25	215.6	3.00	101.2
0.50	189.2	4.00	88.0
0.75	176.0	6.00	61.6
1.00	162.8	12.00	22.0
1.50	138.6	24.00	4.4
2.00	121.0	48.00	0.0

TAB. 1.2 – Données pour l'exemple de Kaplan.

Le système d'équations différentielles décrivant le modèle est alors

$$(EDO) \begin{cases} \frac{dy_1}{dt} = \dot{y}_1(t) = -(k_1 + k_2)y_1(t) + k_3y_2(t) \\ \frac{dy_2}{dt} = \dot{y}_2(t) = k_1y_1(t) - k_3y_2(t) \\ y_1(0) = c_0 \\ y_2(0) = 0. \end{cases}$$

On désire estimer les paramètres c_0, k_1, k_2 et k_3 par les moindres carrés. Posons $\beta = (c_0, k_1, k_2, k_3)$, alors pour toute valeur de β , on peut intégrer le système d'équations différentielles ordinaires à condition initiale (EDO). Notons $(y_1(t\beta), y_2(t\beta))$ cette solution. Par suite on peut calculer les n résidus

$$r_i(\beta) = y_{i1} - y_1(t_i\beta).$$

Ces résidus sont visualisés sur la figure 1.7. Nous estimerons alors le paramètre β en résolvant le problème d'optimisation aux moindres carrés

$$(P) \begin{cases} \text{Min}f(\beta) = \frac{1}{2} \sum_{i=1}^n r_i^2(\beta) = \frac{1}{2} \|r(\beta)\|^2 \\ \beta \in \mathbf{R}^4. \end{cases}$$

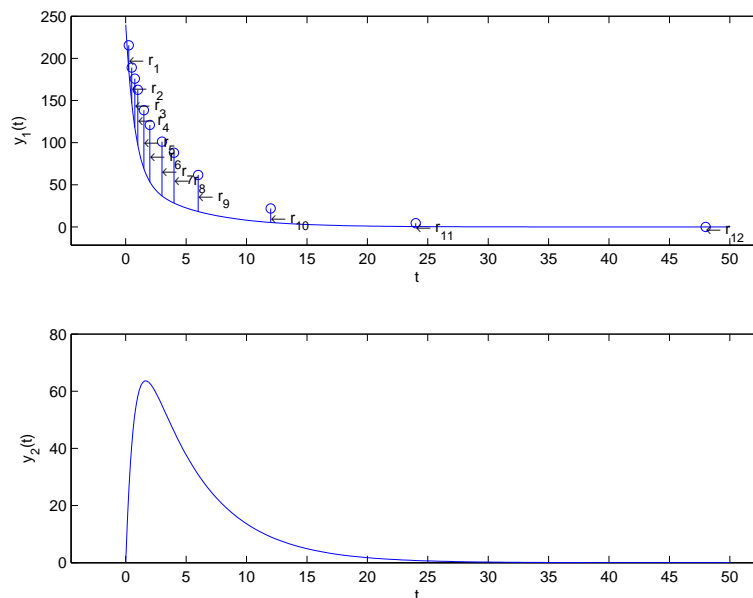


FIG. 1.7 – Critère des moindres carrés pour le modèle de Kaplan.

Exemple 1.1.7. On veut mesurer la liaison entre 2 gènes dominants, l'un contrôlant la couleur d'une fleur, rouge (R) est dominant sur blanc (b), et l'autre la taille, grand (G) est dominant sur petit (p). Dans la descendance F_2 , issu de deux populations homozygotes de phénotype $[RG]$ et $[bp]$, on a étudié $n = 3839$ plantes. On a obtenu les résultats de la table 1.3.

Phénotypes	$[RG]$	$[Rp]$	$[bG]$	$[bp]$
Effectifs observés	1997	906	904	32

TAB. 1.3 – Données de Sir R.A. Fisher.

Le problème est d'estimer, à partir de ces données le taux de recombinaison r . Ici la population F_1 est hétérozygote de génotype Rb, Gp . Nous avons donc les probabilités de la table 1.4 pour les différents gamètes possibles et les différents croisements possibles.

$\varphi : \sigma$	RG	bp	Rp	bG
	$\frac{1}{2}(1-r)$	$\frac{1}{2}(1-r)$	$\frac{1}{2}r$	$\frac{1}{2}r$
RG	$\frac{1}{4}(1-r)^2$	$\frac{1}{4}(1-r)^2$	$\frac{1}{4}r(1-r)$	$\frac{1}{4}r(1-r)$
bp	$\frac{1}{4}(1-r)^2$	$\frac{1}{4}(1-r)^2$	$\frac{1}{4}r(1-r)$	$\frac{1}{4}r(1-r)$
Rp	$\frac{1}{4}r(1-r)$	$\frac{1}{4}r(1-r)$	$\frac{1}{4}r^2$	$\frac{1}{4}r^2$
bG	$\frac{1}{4}r(1-r)$	$\frac{1}{4}r(1-r)$	$\frac{1}{4}r^2$	$\frac{1}{4}r^2$

TAB. 1.4 – Probabilités pour la descendance F_2 .

Par suite nous avons dans la population F_2 la loi suivante pour la variable aléatoire phénotype X

$$\begin{aligned} X : F_2 &\longrightarrow \{[RG], [Rp], [bG], [bp]\} \\ 1 \text{ plante} &\longmapsto \text{son phénotype,} \end{aligned}$$

$$\begin{aligned} P(X = [RG]) &= \frac{1}{4}(3 - 2r + r^2) = \frac{2 + \theta}{4} \\ P(X = [Rp]) &= \frac{1}{4}(2r - r^2) = \frac{1 - \theta}{4} \\ P(X = [bG]) &= \frac{1}{4}(2r - r^2) = \frac{1 - \theta}{4} \\ P(X = [bp]) &= \frac{1}{4}(1 - r)^2 = \frac{\theta}{4} \end{aligned}$$

où $\theta = (1 - r)^2 \in [\frac{1}{4}; 1]$.

Définissons maintenant le vecteur aléatoire de dimension 4

$$\begin{aligned} (A, B, C, D) : F_2^n &\longrightarrow \mathbf{R}^4 \\ n \text{ plantes} &\longmapsto \begin{aligned} &\text{(nb de plantes de phénotypes } [RG], \\ &\text{nb de plantes de phénotypes } [Rp], \\ &\text{nb de plantes de phénotypes } [bG], \\ &\text{nb de plantes de phénotypes } [bp]). \end{aligned} \end{aligned}$$

On suppose la population F_2 de taille infinie, donc la loi de ce vecteur aléatoire est une loi multinomiale

$$\begin{aligned} L(a, b, c, d; \theta) &= P((A, B, C, D) = (a, b, c, d)) \\ &= \frac{n!}{a!b!c!d!} P(X = [RG])^a P(X = [Rp])^b P(X = [bG])^c P(X = [bp])^d \\ &= \frac{n!}{a!b!c!d!} \left(\frac{2 + \theta}{4}\right)^a \left(\frac{1 - \theta}{4}\right)^{b+c} \left(\frac{\theta}{4}\right)^d. \end{aligned}$$

L s'appelle la vraisemblance[‡]. L'estimation de θ par le maximum de vraisemblance consiste alors à rechercher la valeur de θ solution du problème de maximisation suivant

$$(P) \begin{cases} \text{Max } L(1997, 906, 904, 32; \theta) \\ \theta \in [\frac{1}{4}; 1]. \end{cases}$$

Exemple 1.1.8. *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 1.5.*

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

TAB. 1.5 – Coûts et compositions des engrais.

Le fermier veut appliquer au moins 75 kg ha^{-1} d'azote, 25 kg ha^{-1} de phosphate et 35 kg ha^{-1} de potasse. Il ne peut appliquer le lisier qu'à un taux maximum de 8 t/heure et l'engrais qu'à un taux maximum de 0.4 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 est alors d'obtenir un coût minimum, c'est-à-dire que l'on cherche à minimiser $25x_1 + 1300x_2$. Mais nous avons aussi les contraintes suivantes :

$$\begin{array}{ll} x_1 \geq 0 & \text{non négativité de } x_1 \\ x_2 \geq 0 & \text{non négativité de } x_2 \\ 6x_1 + 250x_2 \geq 75 \times 20 = 1500 & \text{contrainte sur l'azote} \\ 1.5x_1 + 100x_2 \geq 500 & \text{contrainte sur le phosphate} \\ 4x_1 + 100x_2 \geq 700 & \text{contrainte sur la potasse} \\ (1/8)x_1 + (1/0.4)x_2 \leq 25 & \text{contrainte de temps.} \end{array}$$

En résumé nous avons le problème suivant à résoudre :

$$(P) \begin{cases} \text{Min } f(x) = 25x_1 + 1300x_2 \\ x_1 \geq 0 \\ x_2 \geq 0 \\ 6x_1 + 250x_2 \geq 75 \times 20 = 1500 \\ 1.5x_1 + 100x_2 \geq 500 \\ 4x_1 + 100x_2 \geq 700 \\ (1/8)x_1 + (1/0.4)x_2 \leq 25. \end{cases}$$

Exemple 1.1.9 (Gestion de portefeuille[2]). *La théorie de la sélection optimale de portefeuille a été développée par Harry Markowitz, prix Nobel d'économie en 1990, dans les années 1950. On considère un investisseur qui a une quantité fixée d'argent à sa disposition pour investir dans n actifs différentes (actions, stocks, ...) dont le retour est aléatoire. Pour chaque actif, on suppose connu son espérance mathématique μ_i , sa variance σ_i^2 . On suppose aussi connu pour deux actifs i et j leur coefficient de corrélation linéaire ρ_{ij} . On note x_i la proportion investie dans l'actif i . On peut donc calculer les espérance mathématique et variance résultant d'un portefeuille $x = (x_1, \dots, x_n)$*

$$\begin{aligned} E(x) &= \mu^T x \\ \text{Var}(x) &= x^T Q x, \end{aligned}$$

où Q est la matrice des covariances, $q_{ij} = \rho_{ij}\sigma_i\sigma_j$. Le portefeuille sera dit efficace si, pour une variance fixée, il a la plus grande espérance mathématique. C'est à dire s'il est solution du problème d'optimisation

$$(P) \begin{cases} \text{Max } E(x) \\ \text{Var}(x) = V \\ \sum_{i=1}^n x_i = 1 \\ x \geq 0. \end{cases}$$

[‡]likelihood en anglais.

On peut aussi s'intéresser au problème (MVO)[§] de Markowitz.

$$(MVO) \begin{cases} \text{Min } \text{Var}(x) \\ E(x) \geq R \\ \sum_{i=1}^n x_i = 1 \\ x \geq 0. \end{cases}$$

Ces deux formulations sont en fait équivalentes.

1.2 Problèmes en nombres entiers

Exemple 1.2.1 (Problème du sac à dos de Knapsack). *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 (1.6).*

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

TAB. 1.6 – 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

$$(P) \begin{cases} \text{Max } 4x_1 + 10x_2 + 15x_3 \\ 2x_1 + 5x_2 + 7x_3 \leq 16 \\ (x_1, x_2, x_3) \in \mathbf{N}^3. \end{cases}$$

Exemple 1.2.2. ([1]) *Dans un service hospitalier, les malades i attendent d'être opérés. Le malade i a besoin d'une durée d'opération D_i . D'autre part, compte tenu des disponibilités des chirurgiens, la somme des durées des opérations possibles chaque jours j de la période étudiée est connue et égale à T_j . On veut minimiser la somme des pénalités d'attente pour les différents malades. On note :*

- $x_{ij} = 1$ si le malade i est opéré le jour j ;
- $x_{ij} = 0$ si le malade i n'est pas opéré le jour j ;
- c_{ij} la pénalité du malade i s'il est opéré le jour j . c_{ij} est une fonction croissante de j .

Le problème s'écrit alors :

$$(P) \begin{cases} \text{Min } f(x) = \sum_i \sum_j c_{ij} x_{ij} \\ \sum_i D_i x_{ij} \leq T_j \quad \forall j \text{ limitation des possibilités opératoire du jour } j \\ \sum_j x_{ij} = 1 \quad \forall i \text{ Le malade } i \text{ est opéré une fois et une seule} \\ x_{ij} = 0 \text{ ou } 1 \text{ l'opération est effectuée en une fois.} \end{cases}$$

Exemple 1.2.3 (Alignement de séquences). *Soit 2 séquences CTGTATC et CTATAATCCC. On désire trouver le "meilleur" alignement possible. À chaque alignement, est associé un score (simple ici) suivant : pour chaque position on associe 0 si les 2 bases sont identiques, +1 si les deux bases sont différentes et +3 s'il y a un "trou". On effectue ensuite la somme. La figure (1.8) donne un exemple de la fonction score S .*

C	T	A	T	–	A	A	–	T	C	C	C	
–	–	C	T	G	T	A	T	C	–	–	–	
3	3	1	0	3	1	0	3	1	3	3	3	= 24

FIG. 1.8 – Exemple de calcul d'un score.

Le problème est alors de résoudre le problème d'optimisation suivant

$$(P) \begin{cases} \text{Min } S(\text{alignement}) \\ \text{pour tous les alignements possibles.} \end{cases}$$

[§]mean-variance optimization.

Remarque 1.2.4. *la difficulté est ici de construire l'ensemble de tous les alignements possibles. Ceci se fait de la façon suivante. Supposons que l'on soit à la position i , alors pour aller à la position $i + 1$, nous avons trois possibilités :*

- *avancer d'un nucléotide pour les 2 séquences ;*
- *avancer d'un nucléotide pour la séquence S_1 et mettre un "trou" pour la séquence S_2 ;*
- *avancer d'un nucléotide pour la séquence S_2 et mettre un "trou" pour la séquence S_1 .*

Nous pouvons ainsi construire un arbre permettant d'avoir tous les alignements possibles.

1.3 Problème en dimension infinie

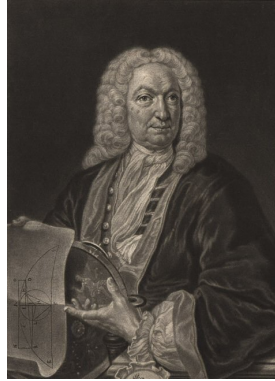


FIG. 1.9 – Jean Bernoulli 27 juillet 1667 – 1er janvier 1748.

Exemple 1.3.1 (Problème de la brachistochrone[¶]). *Le problème posé par Jean Bernoulli^{||} (cf. la figure 1.9) en 1696 est considéré comme le problème fondateur du calcul des variations. Ce problème consiste en la recherche 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. Introduisons dans le plan un système de coordonnées (x, y) pour lequel P_0 ait comme coordonnées $(0, 0)$ et P_f (x_f, y_f) , $x_f > 0$ et $y_f < 0$. Supposons que $y(\cdot)$ est la fonction qui donne l'équation de la courbe joignant les points P_0 et P_f . Les lois de la mécanique nous disent que le module de la vitesse v en $(x, y(x))$ ne dépend pas de la forme de la courbe $y(\cdot)$ sur $[0, x]$, mais seulement de l'ordonnée $y(x)$, et que cette vitesse est égale à $\sqrt{2g(-y(x))}$, où g est l'accélération gravitationnelle. Si on note s l'abscisse curviligne, le temps pour parcourir l'élément $ds = \sqrt{dx^2 + dy^2}$ est alors $ds/\sqrt{2g(-y(x))}$. Posons*

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

$$y(\cdot) \longmapsto \int_0^{x_f} \frac{\sqrt{1 + y'(x)^2}}{\sqrt{2g(-y(x))}} dx.$$

Le problème s'écrit alors

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

La solution de ce problème est visualisée à la figure 1.10.

Exemple 1.3.2 (Transfert orbital). *On désire transférer un satellite S d'une orbite initiale (celle où la fusée Ariane l'a "posé") vers l'orbite géostationnaire (cf. la figure 1.11), le moteur du satellite étant un moteur à poussée faible.*

Le satellite est considéré comme un point matériel et on note $r(t) \in \mathbf{R}^3$ la position, $v(t) \in \mathbf{R}^3$ la vitesse, $m(t)$ la masse du satellite et $T(t) \in \mathbf{R}^3$ la poussée du moteur. L'équation du mouvement, provenant des équations de Newton est alors

$$\begin{aligned} \dot{r}(t) &= v(t) \\ \dot{v}(t) &= -\frac{\mu r(t)}{\|r(t)\|^3} + \frac{T(t)}{m(t)} \\ \dot{m}(t) &= -\beta \|T(t)\|, \end{aligned}$$

[¶]Le mot brachistochrone vient du grec *brakhisto* qui signifie *le plus court* et de *chronos* qui signifie *temps*.

^{||}http://fr.wikipedia.org/wiki/Jean_Bernoulli

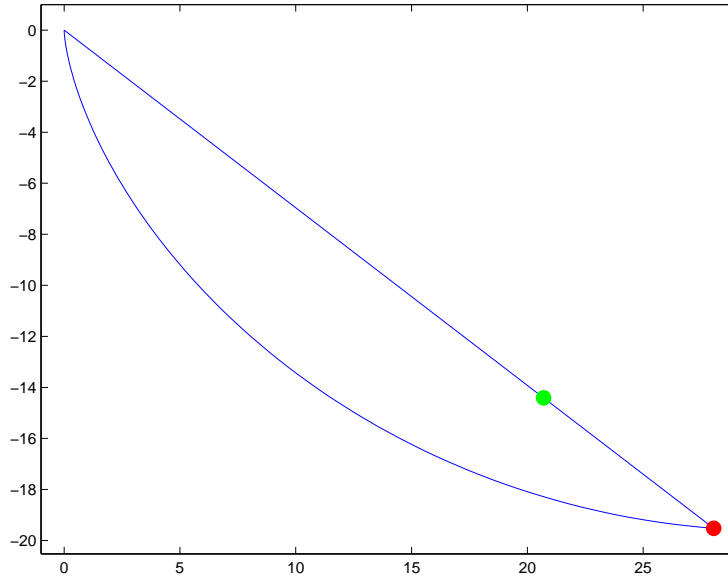


FIG. 1.10 – Brachistochrone.

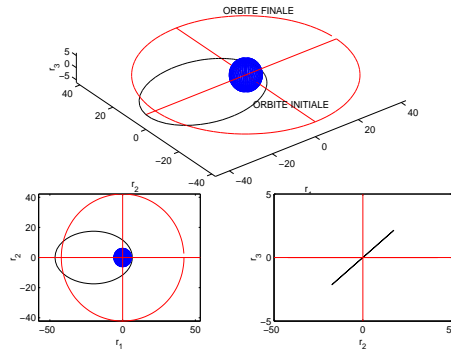


FIG. 1.11 – Transfert orbital.

où μ est la constante gravitationnelle de la Terre et $\beta = 1/g_0 Isp$ est une constante positive (g_0 est l'accélération gravitationnelle terrestre à la surface de la Terre et Isp est une constante caractéristique du moteur appelé impulsion spécifique). À l'instant initial les position, vitesse et masse du satellite sont connues et à l'instant terminal t_f le satellite doit être sur l'orbite géostationnaire à une position et vitesse $(r(t_f), v(t_f)) = (r_f, v_f)$ fixés. Bien évidemment la poussée du moteur est bornée

$$\|T(t)\| \leq T_{max}.$$

L'objectif est alors de trouver une loi de commande du moteur qui réalise le transfert et qui minimise le temps de transfert. On peut aussi s'intéresser à la maximisation de la masse finale (dans ce cas le temps de transfert doit-être fixé). Si on normalise le contrôle $u(t) = T(t)/T_{max}$ alors le problème s'écrit pour la maximisation de la masse finale (ou la minimisation de la consommation)

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

ce problème est un problème de contrôle optimal et l'inconnue est la commande, donc une fonction u , ici de $[0, t_f]$ à valeurs dans \mathbf{R}^3 .

Remarque 1.3.3. Ces problèmes d'optimisation en dimension infinie seront traités dans le cours de contrôle optimal en deuxième année majeure mathématiques appliquées.

2 Problème d'optimisation

2.1 Définitions

Définition 2.1.1 (Fonction convexe). Une fonctionnelle f de \mathbf{R}^n à valeurs dans \mathbf{R} est convexe si et seulement si elle vérifie :

$$\forall (x, y) \in \mathbf{R}^n \times \mathbf{R}^n, \forall \alpha \in [0, 1], f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y)$$

Dans la cas $n = 1$, ceci signifie que le graphe de la fonction f est toujours sous la corde (1.12)

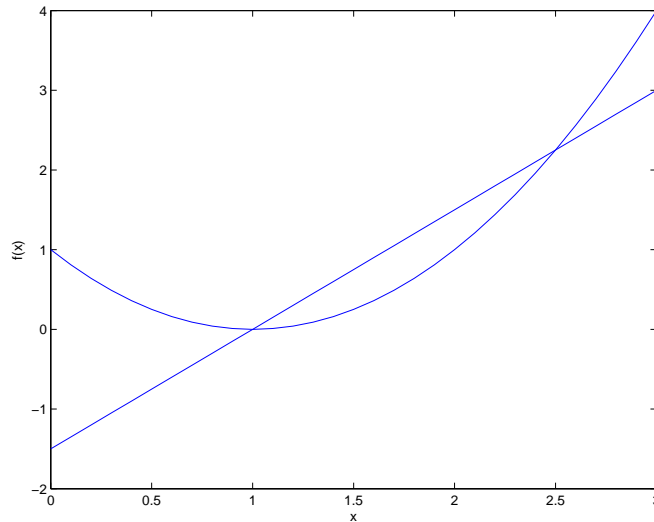


FIG. 1.12 – Fonction convexe.

Définition 2.1.2 (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 \mathbf{R}^n . On notera ce problème sous la forme suivante :

$$(P) \begin{cases} \text{Min} f(x) \\ x \in \mathbf{R}^n \end{cases}$$

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

Remarque 2.1.3. Résoudre le problème (P) revient à rechercher le point x^* de \mathbf{R}^n tel que $f(x^*) \leq f(x) \forall x \in \mathbf{R}^n$.

Remarque 2.1.4. Un problème de maximisation se ramène très facilement à un problème de minimisation :

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

Définition 2.1.5 (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 \mathbf{R}^n d'une fonctionnelle f définie sur \mathbf{R}^n . On notera ce problème sous la forme suivante :

$$(P) \begin{cases} \text{Min} f(x) \\ x \in C \subset \mathbf{R}^n \end{cases}$$

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

Remarque 2.1.6. Dans la pratique C sera défini de la façon suivante :

$$C = \{x \in \mathbf{R}^n / g_i(x) \leq 0 \quad i = 1, \dots, m \text{ et } h_l(x) = 0 \quad l = 1, \dots, p\} \quad (1.1)$$

et nous écrirons (P) sous la forme

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

Définition 2.1.7 (Optimisation non différentiable). On appelle problème d'optimisation non différentiable un problème d'optimisation où les fonctions qui interviennent ne sont pas dérivables.

Remarque 2.1.8. On ne traitera dans ce cours que des problèmes d'optimisation différentiables.

Définition 2.1.9 (Problème d'optimisation convexe). Un problème d'optimisation est dit convexe si et seulement si la fonction f est convexe et l'ensemble des contraintes C est convexe.

Remarque 2.1.10. Si C est définie par 1.1 alors C est convexe si et seulement si les fonctions g_i sont convexes et les fonctions h_l sont affines.

Définition 2.1.11 (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 :

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

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

Exemple 2.1.12. L'exemple (1.1.4) est un problème aux moindres carrés non linéaire.

Définition 2.1.13 (Problème linéaire). Un problème d'optimisation est dit linéaire si et seulement si les fonctions f , g_i , et h_l sont affines.

Exemple 2.1.14. L'exemple (1.1.8) est un problème linéaire.

Définition 2.1.15 (Optimum global, optimum local). Soit (P) un problème d'optimisation sans contraintes.

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

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

- (iii) x^* est un minimum local fort si

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

Dans le cas où $n = 1$, $\|x - x^*\|$ devient $|x - x^*|$ et par suite nous avons

$$\|x - x^*\| < \varepsilon \iff |x - x^*| < \varepsilon \iff x^* - \varepsilon < x < x^* + \varepsilon,$$

(cf. la figure 1.13).

Remarque 2.1.16. 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.

Remarque 2.1.17. 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.

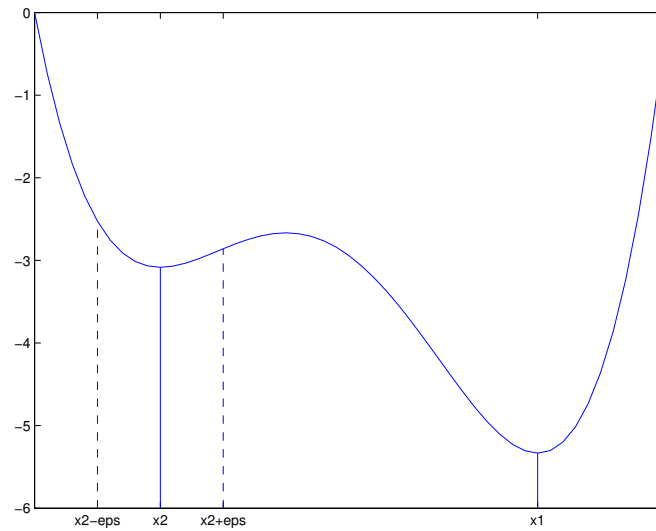


FIG. 1.13 – x^2 est un minimum local fort, x^1 est un minimum global

2.2 Classification

Considérons le problème d'optimisation suivant :

$$(P) \begin{cases} \text{Min } f(x) \\ x \in C \subset E \end{cases}$$

Suivant la nature des ensembles C et E et de la fonction f nous avons différents types de problème d'optimisation. La figure donne une classification des problèmes d'optimisation (nous n'étudierons dans ce cours que les parties en bleu).

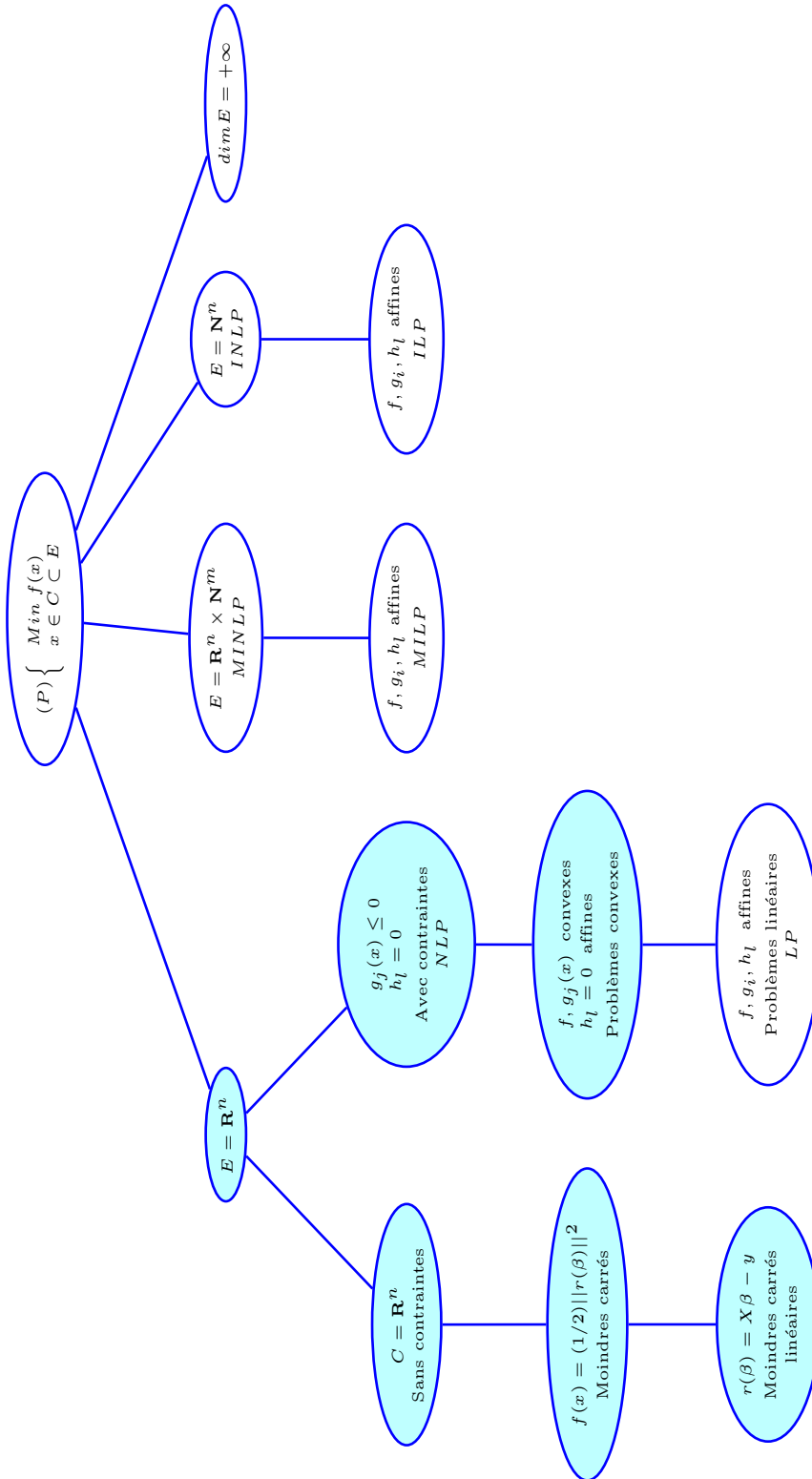


FIG. 1.14 – Classification des problèmes d'optimisation, on a en bleu ce qui sera vu en cours.

Chapitre 2

Condition nécessaire, condition suffisante : cas sans contraintes et cas convexe

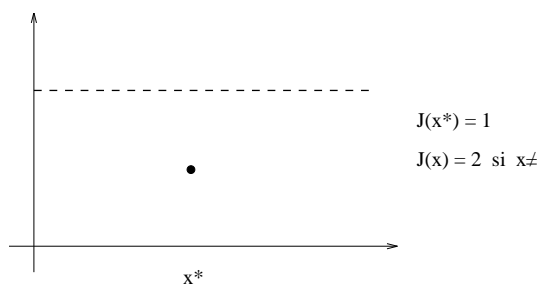
1 Introduction

Dans ce chapitre une fonctionnelle est toute fonction à valeurs dans \mathbf{R} . L'objet de ce chapitre est d'étudier, dans le cas différentiable, les conditions nécessaires et les conditions suffisantes pour les problèmes d'optimisation sans contrainte et les problèmes convexes. On suppose connu ici l'algèbre linéaire et le calcul différentiel.

Remarque 1.0.1. *On ne peut rien dire quant au comportement de f en x^* . A moins que f n'ait certaines propriétés, de continuité en particulier, on pourra difficilement trouver un algorithme permettant de déterminer un minimum de f si, par exemple*

$$\lim_{\lambda \rightarrow 0} |f(x^*) - f(x^* + \lambda h)| \neq 0$$

où $\lambda \in \mathbf{R}$, et h est un vecteur arbitraire.



Exemple 1.0.2.

Si l'on excepte certaines méthodes dites exploratoires, un grand nombre d'algorithmes de minimisation de fonctionnelles nécessitent que les dérivées partielles premières de f soient continues, et beaucoup de ces algorithmes requièrent même que les dérivées partielles secondes de f soient continues.

2 Convexité des fonctionnelles

2.1 Ensembles convexes - fonctionnelles convexes

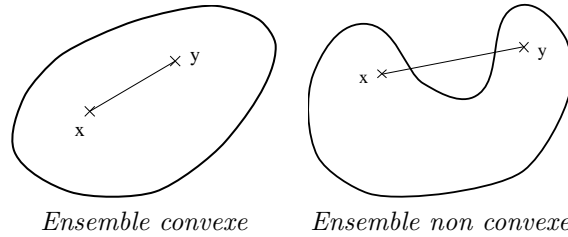
On indique dans ce paragraphe quelques propriétés de base d'une classe très importante de fonctionnelles.

Définition 2.1.1. Ensembles convexes

L'ensemble D_0 est dit **convexe** si et seulement si

$$\forall x \in D_0, \forall y \in D_0, \forall \alpha \in [0, 1] \subset \mathbf{R} \text{ on a } \alpha x + (1 - \alpha)y \in D_0 .$$

Illustration graphique 2.1.2.



Ensemble convexe

Ensemble non convexe

autrement dit, si $x \in D_0$ et $y \in D_0$, alors le segment qui joint ces deux points est également contenu dans D_0 , le segment $[x, y]$ étant défini par

$$z \in [x, y] \iff \exists \alpha \in [0, 1] \text{ t.q. } z = \alpha x + (1 - \alpha)y .$$

Remarque 2.1.3. La notion d'ensemble convexe correspond en fait à une propriété de régularité du domaine D_0 considéré.

Définition 2.1.4. Fonctionnelles convexes

Une fonctionnelle $f : D_0 \subset E \rightarrow \mathbf{R}$ est **convexe** sur le domaine convexe $D_0 \subset E$ (E espace vectoriel normé) si

$$\forall x, y \in D_0, \forall \alpha \in]0, 1[, \quad f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y) .$$

La fonctionnelle f est **strictement convexe** sur le domaine convexe D_0 si

$$\forall x, y \in D_0, x \neq y, \forall \alpha \in]0, 1[, \quad f(\alpha x + (1 - \alpha)y) < \alpha f(x) + (1 - \alpha)f(y) .$$

La fonctionnelle f est **uniformément convexe** sur le domaine convexe D_0 si il existe une constante $c > 0$ telle que

$$\forall x, y \in D_0, \forall \alpha \in]0, 1[, \\ \alpha f(x) + (1 - \alpha)f(y) - f(\alpha x + (1 - \alpha)y) \geq c \alpha (1 - \alpha) \|x - y\|_E^2 .$$

Remarque 2.1.5. (i) Il est clair que la convexité uniforme entraîne la convexité stricte qui à son tour entraîne la convexité.

(ii) La convexité indique une certaine régularité de la fonctionnelle. En dimension finie, par exemple, la convexité peut induire des propriétés de continuité (c.f. proposition suivante).

Proposition 2.1.6. Soit $f : D_0 \subset \mathbf{R}^n \rightarrow \mathbf{R}$ une fonctionnelle convexe sur l'ouvert convexe $D_0 \subset \mathbf{R}^n$. Alors f est continue sur D_0 .

Les définitions de base de la convexité (large, stricte, ou uniforme) peuvent parfois s'avérer d'un emploi peu commode. Le but des paragraphes qui suivent est de mettre en avant des propriétés qui s'y rapportent, exploitant la différentiabilité d'une fonctionnelle, et plus faciles à manipuler.

2.2 Convexité et dérivée première

Théorème 2.2.1. Caractérisation de la convexité à l'aide de la dérivée première

On suppose que la fonctionnelle $f : \Omega \subset E \rightarrow \mathbf{R}$ est dérivable (au sens de Gâteaux ou au sens de Fréchet) sur un sous-ensemble convexe $D_0 \subset \Omega$. On a alors :

(i) f est convexe sur D_0 si et seulement si

$$\forall x, y \in D_0, \quad f(y) - f(x) \geq f'(x) \cdot (y - x) .$$

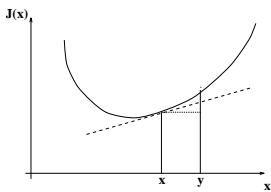
(ii) f est strictement convexe sur D_0 si et seulement si

$$\forall x, y \in D_0, x \neq y, \quad f(y) - f(x) > f'(x) \cdot (y - x) .$$

(iii) La fonctionnelle f est uniformément convexe sur D_0 si et seulement si il existe une constante $c > 0$ telle que

$$\forall x, y \in D_0, \quad f(y) - f(x) \geq f'(x) \cdot (y - x) + c \|y - x\|_E^2 .$$

Illustration graphique 2.2.2.



L'interprétation géométrique de

$$\forall x, y \in D_0, \quad f(y) - f(x) \geq f'(x) \cdot (y - x) ,$$

est que le graphe de la fonctionnelle convexe f est toujours au dessus de son plan tangent en un point quelconque du domaine D_0 .

Définition 2.2.3. Soit une fonctionnelle $f : \Omega \subset E \rightarrow \mathbf{R}$ dérivable (au sens de Gâteaux ou au sens de Fréchet) sur l'ouvert Ω .

L'application dérivée $f' : \Omega \subset E \rightarrow \mathcal{L}(E, \mathbf{R})$ est dite **monotone sur le sous-ensemble** $D_0 \subset \Omega$ si et seulement si

$$\forall x, y \in D_0, \quad (f'(y) - f'(x)) \cdot (y - x) \geq 0 .$$

L'application dérivée f' est dite **strictement monotone sur le sous-ensemble** $D_0 \subset \Omega$ si et seulement si

$$\forall x, y \in D_0, x \neq y, \quad (f'(y) - f'(x)) \cdot (y - x) > 0 .$$

L'application dérivée f' est dite **fortement monotone sur le sous-ensemble** $D_0 \subset \Omega$ si et seulement si il existe une constante $c > 0$ telle que

$$\forall x, y \in D_0, \quad (f'(y) - f'(x)) \cdot (y - x) \geq 2c \|y - x\|_E^2 .$$

Proposition 2.2.4. Relations entre convexité et monotonie de la dérivée première

On suppose que la fonctionnelle $f : \Omega \subset E \rightarrow \mathbf{R}$ est dérivable (au sens de Gâteaux ou au sens de Fréchet) sur Ω . On a alors :

(i) La fonctionnelle f est convexe sur le sous-ensemble convexe $D_0 \subset \Omega$ si et seulement si l'application dérivée f' est monotone sur D_0 .

(ii) La fonctionnelle f est strictement convexe sur le sous-ensemble convexe $D_0 \subset \Omega$ si et seulement si l'application dérivée f' est strictement monotone sur D_0 .

(iii) La fonctionnelle f est uniformément convexe sur le sous-ensemble convexe $D_0 \subset \Omega$ si et seulement si l'application dérivée f' est fortement monotone sur D_0 (la constante $c > 0$ intervenant dans la définition de la convexité uniforme correspondant à la constante $c > 0$ introduite dans la définition de la forte monotonie de la dérivée).

2.3 Convexité et dérivée seconde

Théorème 2.3.1. Relations entre convexité et positivité de la dérivée seconde

On suppose que la fonctionnelle $f : \Omega \subset E \rightarrow \mathbf{R}$ est deux fois dérivable dans un ouvert Ω de l'espace vectoriel normé E , et soit D_0 une partie convexe de Ω .

(i) La fonctionnelle f est convexe sur le sous-ensemble convexe $D_0 \subset \Omega$ si et seulement si

$$\forall x, y \in D_0, \quad f''(x)(y - x, y - x) \geq 0.$$

(ii) Si

$$\forall x, y \in D_0, \quad x \neq y, \quad f''(x)(y - x, y - x) > 0,$$

alors la fonctionnelle f est strictement convexe sur D_0 .

(iii) La fonctionnelle f est uniformément convexe sur le sous-ensemble convexe $D_0 \subset \Omega$ si et seulement si il existe une constante $c > 0$ telle que

$$\forall x, y \in D_0, \quad f''(x)(y - x, y - x) \geq 2c \|y - x\|_E^2.$$



La condition (2) ci-dessus n'est qu'une condition suffisante, la réciproque étant inexacte.

3 Extrémas de fonctionnelles

3.1 Prise en compte des dérivées premières

Théorème 3.1.1. Condition nécessaire d'extrémum local

Soit Ω un ouvert d'un espace vectoriel normé E , et soit la fonctionnelle $f : \Omega \subset E \rightarrow \mathbf{R}$. Si la fonctionnelle f admet un extrémum local en un point $x \in \Omega$, et si f est différentiable en ce point, alors

$$f'(x) = 0.$$

Remarque 3.1.2. • La relation $f'(x) = 0$ est parfois appelée équation d'Euler.

• Le fait que Ω soit ouvert est bien évidemment essentiel.

Théorème 3.1.3. Condition nécessaire de minimum sur un ensemble convexe

Soit $f : \Omega \subset E \rightarrow \mathbf{R}$ une fonctionnelle définie sur un ouvert Ω d'un espace vectoriel normé E , et soit D_0 une partie convexe de Ω . Si la fonctionnelle f est différentiable au point $x \in D_0$, et si elle admet en x un minimum par rapport à l'ensemble D_0 , alors

$$f'(x) \cdot (y - x) \geq 0 \text{ pour tout } y \in D_0.$$

Remarque 3.1.4. • Si l'ensemble D_0 est un sous-espace vectoriel de E , la condition précédente devient $f'(x) \cdot h = 0$, pour tout $h \in D_0$.

• En particulier, si $D_0 = E$ on retrouve la condition d'Euler énoncée au théorème 3.1.1.

3.2 Prise en compte des dérivées secondes

Définition 3.2.1. Soit une fonctionnelle $f : \Omega \subset E \rightarrow \mathbf{R}$ deux fois différentiable au point $x \in \Omega$.

La dérivée seconde de l'application f au point x , $f''(x) \in \mathcal{L}_2(E \times E, \mathbf{R})$, est dite **semi-définie positive** au point $x \in \Omega$ si et seulement si

$$\forall h \in E, \quad f''(x)(h, h) \geq 0.$$

La dérivée seconde de l'application f au point x , $f''(x) \in \mathcal{L}_2(E \times E, \mathbf{R})$, est dite **définie positive** au point $x \in \Omega$ si et seulement si

$$\forall h \in E, \quad h \neq 0, \quad f''(x)(h, h) > 0.$$

La dérivée seconde de l'application f au point x , $f''(x) \in \mathcal{L}_2(E \times E, \mathbf{R})$, est dite **uniformément définie positive**, ou encore **elliptique**, au point $x \in \Omega$ si et seulement si il existe une constante $c > 0$ telle que

$$\forall h \in E, \quad f''(x)(h, h) \geq c \|h\|_E^2.$$

Remarque 3.2.2. Si E est de dimension finie alors les notions de définie positivité et d'uniformément définie positivité de $f''(x)$ coïncident. Dans ce cas $f''(x)$ est semi-définie positive (respectivement définie positive) si et seulement si les valeurs propres de la matrice hessienne $\nabla^2 f(x)$ sont toutes positives ou nulles (respectivement strictement positive).

Théorème 3.2.3. Condition nécessaire de minimum local

Soit Ω un ouvert d'un espace vectoriel normé E , et soit $f : \Omega \subset E \rightarrow \mathbf{R}$ une fonctionnelle dérivable dans Ω , deux fois différentiable au point $x \in \Omega$. Si la fonctionnelle f admet un minimum local au point x , alors l'application dérivée seconde de f au point x ($f''(x) \in \mathcal{L}_2(E \times E, \mathbf{R})$) est semi-définie positive.

Remarque 3.2.4. Il n'y a pas de réciproque au théorème précédent ; considérer par exemple la fonction $f(x) = x^3$ sur \mathbf{R} .

Théorème 3.2.5. Conditions suffisantes de minimum local

Soit Ω un ouvert d'un espace vectoriel normé E , et soit $f : \Omega \subset E \rightarrow \mathbf{R}$ une fonctionnelle dérivable dans Ω telle que $f'(x) = 0$ en un point $x \in \Omega$.

- (i) Si la fonctionnelle f est deux fois différentiable au point $x \in \Omega$, et si l'application dérivée seconde de f au point x ($f''(x) \in \mathcal{L}_2(E \times E, \mathbf{R})$) est elliptique, alors la fonctionnelle f admet un minimum local fort en x .
- (ii) Si la fonctionnelle f est deux fois dérivable dans Ω , et s'il existe une boule $\mathcal{B} \subset \Omega$ centrée en x telle que, pour tout $v \in \mathcal{B}$, l'application $f''(v) \in \mathcal{L}_2(E \times E, \mathbf{R})$ est semi-définie positive, alors la fonctionnelle f admet un minimum local en x .

Remarque 3.2.6. Il n'existe pas de réciproque aux deux assertions du théorème précédent.

3.3 Prise en compte de la convexité

Théorème 3.3.1. Soit D_0 une partie convexe d'un espace vectoriel normé E .

- (i) Si une fonctionnelle convexe $f : D_0 \subset E \rightarrow \mathbf{R}$ admet un minimum local en un point de D_0 , elle y admet en fait un minimum global, c'est à dire par rapport à tout l'ensemble D_0 .
- (ii) Une fonctionnelle $f : D_0 \subset E \rightarrow \mathbf{R}$ strictement convexe admet au plus un minimum, et c'est un minimum fort (s'il existe).
- (iii) Soit $f : \Omega \subset E \rightarrow \mathbf{R}$ une fonctionnelle définie sur un ouvert $\Omega \subset E$ contenant D_0 , convexe sur D_0 , et différentiable en un point $x \in D_0$. Alors, la fonctionnelle f admet un minimum en x par rapport à l'ensemble D_0 si et seulement si

$$f'(x) \cdot (y - x) \geq 0 \text{ pour tout } y \in D_0 .$$

- (iv) Si de plus l'ensemble D_0 est ouvert, la condition précédente équivaut à l'équation d'Euler $f'(x) = 0$.

Remarques :

- Au théorème 3.1.3, la condition (3) avait été énoncée comme condition nécessaire seulement, mais ceci sans supposer f convexe.
- Les relations " $f'(x) \cdot (y - x) \geq 0$ pour tout $y \in D_0$ " sont fréquemment appelées *inéquations d'Euler*.

3.4 Existence d'un minimum

Définition 3.4.1. Une fonctionnelle $f : E \rightarrow \mathbf{R}$ définie sur un espace vectoriel normé E est dite **coercive** si

$$\lim_{\|x\|_E \rightarrow +\infty} f(x) = +\infty .$$

Théorème 3.4.2. Soit D_0 une partie non vide fermée de \mathbf{R}^n et $f : \mathbf{R}^n \rightarrow \mathbf{R}$ une fonctionnelle continue, coercive si l'ensemble D_0 est non borné. Alors il existe au moins un élément x tel que

$$x \in D_0 \text{ et } f(x) = \inf_{v \in D_0} f(v) .$$

Remarque : Ce résultat valable en dimension finie peut s'étendre en dimension infinie dans le cas des espaces de Hilbert séparables, avec néanmoins des hypothèses supplémentaires. Nous ne l'énoncerons pas ici, les espaces de Hilbert faisant l'objet de développements ultérieurs.

Bibliographie

- [1] Carpentier. Cours de 3^{ème} année enseiht, filière informatique et mathématiques appliquées. INPT-ENSEEIH, 1983.
- [2] Gerard Cornuejols and Reha Tütüncü. *Optimization Methods in Finance*. Cambridge University Press, 2007.
- [3] J.M. Devos. Fermat "le premier homme du monde". France 3, IREM de Toulouse, CRDP Midi-Pyrénées, 1995. Casette vidéo.