Département Sciences du Numérique

Calcul différentiel - Optimisation sans contraintes - Premiers algorithmes

O. Cots, J. Gergaud, S. Gratton, D. Ruiz et E. Simon

10 octobre 2020

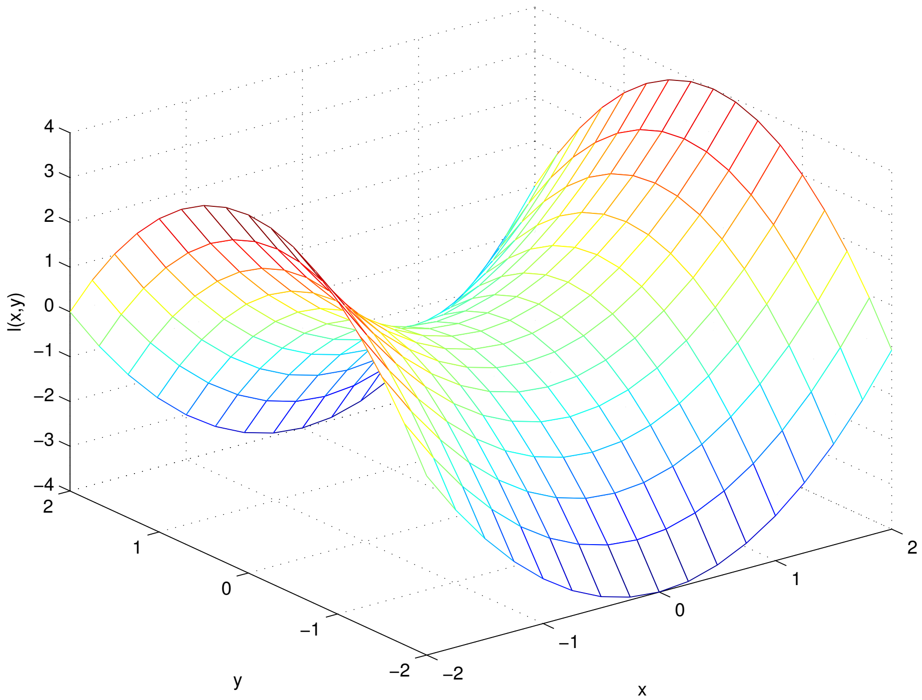

Table des matières

Chapitre 1

Introduction à la dualité

1.1 Introduction

Ce chapitre est très fortement inspiré du livre du Professeur Jean-Baptiste Hiriart-Urruty de l'Université Paul Sabatier[1].

On note dans ce chapitre le produit scalaire canonique de $\mathbb{R}^n$ $(.\,|\,.)$.

1.2 Point selle

1.2.1 Définition

Définition 1.2.1 – Point selle, point col

Soit l une application d'un produit cartésien $X \times Y$ à valeurs dans $\mathbb{R}$. Un point $(\bar{x}, \bar{y}) \in X \times Y$ est un point selle ou un point col si et seulement si pour tout point $(x, y) \in X \times Y$ on a

$$l(\bar{x}, y) \leq l(\bar{x}, \bar{y}) \leq l(x, \bar{y}) \tag{1.1}$$

Remarque 1.2.1. L'équation (1.1) est équivalente à

$$\forall (x, y) \in X \times Y,\ l(\bar{x}, y) \leq l(x, \bar{y}).$$

Il suffit en effet de faire dans cette dernière inégalité $x = \bar{x}$ (respectivement $y = \bar{y}$) et on obtient l'inégalité de droite (respectivement de gauche) de la définition d'un point selle.

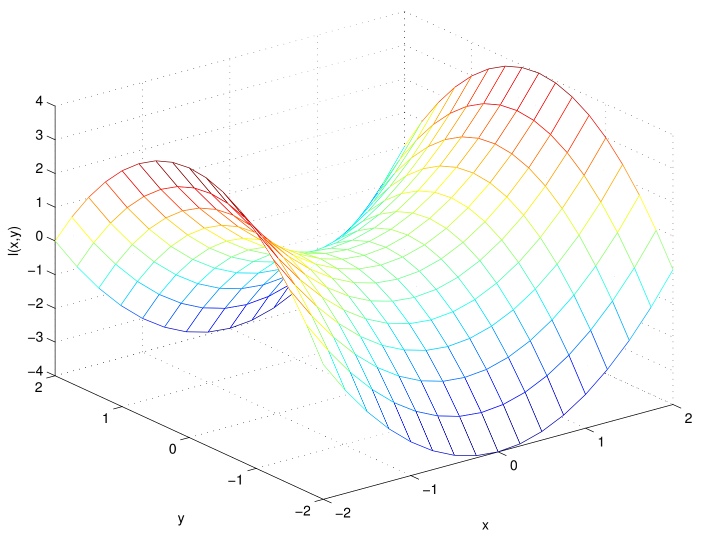

Figure 1.1 – *Exemple de point selle,* $l(x, y) = x^2 - y^2$.

Proposition 1.2.2

Si l possède deux points selles $(\bar{x}_1, \bar{y}_1)$ et $(\bar{x}_2, \bar{y}_2)$, alors $l(\bar{x}_1, \bar{y}_1) = l(\bar{x}_2, \bar{y}_2)$ et $(\bar{x}_1, \bar{y}_2)$ et $(\bar{x}_2, \bar{y}_1)$ sont aussi des points selles.

► Prenons $(x, y) = (\bar{x}_2, \bar{y}_2)$ dans la définition du point selle $(\bar{x}_1, \bar{y}_1)$, on a alors les inégalités

$$l(\bar{x}_1, \bar{y}_2) \leq l(\bar{x}_1, \bar{y}_1) \leq l(\bar{x}_2, \bar{y}_1).$$

De même, en posant $(x, y) = (\bar{x}_1, \bar{y}_1)$ dans la définition du point selle $(\bar{x}_2, \bar{y}_2)$, on obtient

$$l(\bar{x}_2, \bar{y}_1) \leq l(\bar{x}_2, \bar{y}_2) \leq l(\bar{x}_1, \bar{y}_2).$$

Par suite toute ces inégalités sont des égalités. On en déduit alors que pour tout $(x, y) \in X \times Y$ on a

$$l(\bar{x}_1, y) \leq l(\bar{x}_1, \bar{y}_1) = l(\bar{x}_1, \bar{y}_2) = l(\bar{x}_2, \bar{y}_2) \leq l(x, \bar{y}_2),$$

soit que $(\bar{x}_1, \bar{y}_2)$ est un point selle. On a une démonstration identique pour le deuxième point. ■

Corollaire 1.2.3

L'ensemble des points selle est un produit cartésien $A \times B$.

► Posons

$$\begin{aligned} A &= \{x \in X, \text{ il existe } y \in Y \text{ tel que } (x, y) \text{ soit un point selle}\} \\ B &= \{y \in Y, \text{ il existe } x \in X \text{ tel que } (x, y) \text{ soit un point selle}\} \end{aligned}$$

alors l'ensemble des points selles est $A \times B$. Ceci est évident s'il n'existe pas de points selles. Dans le cas contraire, l'ensemble des points selles est de façon évidente inclus dans le produit cartésien $A \times B$. Montrons l'inclusion inverse. Soit dont $\bar{x}_1 \in A$ et $\bar{y}_2 \in B$, alors il existe $\bar{y}_1 \in Y$ et $\bar{x}_2 \in X$ tels que $(\bar{x}_1, \bar{y}_1)$ et $(\bar{x}_2, \bar{y}_2)$ soient deux points selles. La proposition précédente montre alors que $(\bar{x}_1, \bar{y}_2)$ est aussi un point selle. ■

1.2.2 Caractérisation

On définit maintenant les deux fonctions

$$\begin{array}{rcl} \varphi : X & \to & \bar{\mathbb{R}} \\ x & \mapsto & \varphi(x) = \sup_{y \in Y} l(x, y) \end{array} \qquad \begin{array}{rcl} \psi : Y & \to & \bar{\mathbb{R}} \\ y & \mapsto & \psi(y) = \inf_{x \in X} l(x, y). \end{array}$$

Proposition 1.2.4 – Dualité faible

On a toujours $\psi(y) \leq l(x, y) \leq \varphi(x)$ et donc ce qu'on appelle la dualité faible

$$\sup_{y \in Y} \inf_{x \in X} l(x, y) \leq \inf_{x \in X} Sup_{y \in Y} l(x, y). \tag{1.2}$$

Proposition 1.2.5

$(\bar{x}, \bar{y})$ est un point selle si et seulement si $\psi(\bar{y}) \geq \varphi(\bar{x})$ et on a alors $\varphi(\bar{x}) = l(\bar{x}, \bar{y}) = \psi(\bar{y})$.

► Si $(\bar{x}, \bar{y})$ est un point selle, on a pour tout $(x, y) \in X \times Y$

$$l(\bar{x}, y) \leq l(\bar{x}, \bar{y}) \leq l(x, \bar{y}).$$

l'inégalité de droite signifie que $l(\bar{x}, \bar{y}) = \min_{x \in X} l(x, \bar{y}) = \inf_{x \in X} l(x, \bar{y}) = \psi(\bar{y})$ et l'inégalité de

gauche que $l(\bar{x},\bar{y}) = \max_{y\in Y} l(\bar{x},y) = \sup_{y\in Y} l(\bar{x},y) = \varphi(\bar{x})$. Par suite on a $\varphi(\bar{x}) = \psi(\bar{y})$.
Considérons maintenant que $\psi(\bar{y}) \geq \varphi(\bar{x})$. Comme on a toujours l'inégalité inverse, on a l'égalité $\psi(\bar{y}) = \varphi(\bar{x})$. Mais $l(\bar{x},\bar{y}) \leq \varphi(\bar{x}) = \psi(\bar{y}) \leq l(\bar{x},\bar{y})$, donc ces inégalités sont des égalités. Par suite on a

$$l(\bar{x},y) \leq l(\bar{x},\bar{y}) \leq l(x,\bar{y})$$

et $(\bar{x},\bar{y})$ est un point selle. ■

Définissons maintenant les deux problèmes suivants

$$(P)\left\{\begin{array}{l}\min \varphi(x)\\ x\in X\end{array}\right. \qquad (D)\left\{\begin{array}{l}\max \psi(y)\\ y\in Y\end{array}\right.$$

et notons Φ (respectivement Ψ) l'ensemble des solutions de (P) (respectivement de (D)). on a alors le

Théorème 1.2.6

Une condition nécessaire et suffisante pour que l'on ait un point selle est que les problème (P) et (D) aient des solutions $\bar{x}$ et $\bar{y}$ et que $\varphi(\bar{x}) = \psi(\bar{y})$. L'ensemble des points selles est alors le produit cartésien $\Phi\times\Psi$. Et on a alors la relation

$$\sup_{y\in Y}\inf_{x\in X} l(x,y) = \inf_{x\in X} Sup_{y\in Y} l(x,y). \tag{1.3}$$

► Si $(\bar{x},\bar{y})$ est un point selle alors $\varphi(\bar{x}) = \psi(\bar{y})$. Donc $\psi(y) \leq \varphi(\bar{x}) = \psi(\bar{y})$ et $\bar{y}$ est une solution de (D). De même $\bar{x}$ est solution de (P).
Réciproquement, supposons maintenant que l'on ait $\bar{x}\in\Phi$ et $\bar{y}\in\Psi$ et que $\varphi(\bar{x}) = \psi(\bar{y})$, alors la proposition 1.2.2 permet de conclure que $(\bar{x},\bar{y})$ est un point selle. ■

1.2.3 Point selle de lagrangien

Revenons maintenant à nos problèmes d'optimisation

$$(P)\left\{\begin{array}{l}\min f(x)\\ h(x)=0\\ g(x)\leq 0\\ x\in\mathbb{R}^n,\end{array}\right.$$

et étudions le cas où l est le Lagrangien

$$\begin{array}{rccc} L: & \mathbb{R}^n\times\mathbb{R}^p\times(\mathbb{R}^+)^m & \longrightarrow & \mathbb{R}\\ & (x,\lambda,\mu) & \longmapsto & L(x,\lambda,\mu),\end{array}$$

avec

- $X = \mathbb{R}^n$;
- $Y = \mathbb{R}^p\times(\mathbb{R}^+)^m$.

Dans ce cas on a la caractérisation suivante d'un point selle

Théorème 1.2.7

Le point $(\bar{x},\bar{\lambda},\bar{\mu})\in\mathbb{R}^n\times\mathbb{R}^p\times(\mathbb{R}^+)^m$ est un point selle si et seulement si

1. *$\bar{x}$ est solution de*
$$\left\{\begin{array}{l}\min L(x,\bar{\lambda},\bar{\mu})\\ x\in\mathbb{R}^n;\end{array}\right.$$
2. *$h(\bar{x}) = 0$ et $g(\bar{x})\leq 0$;*
3. *$\bar{\mu}\geq 0$;*

4. $\bar{\mu}_j g_j(\bar{x}) = 0$ *pour tout* $j = 1, \ldots, m$.

► Soit $(\bar{x}, \bar{\lambda}, \bar{\mu}) \in \mathbb{R}^n \times \mathbb{R}^p \times (\mathbb{R}^+)^m$ un point selle du Lagrangien, alors

$$\forall (x, \lambda, \mu) \in \mathbb{R}^n \times \mathbb{R}^p \times (\mathbb{R}^+)^m, L(\bar{x}, \lambda, \mu) \leq L(\bar{x}, \bar{\lambda}, \bar{\mu}) \leq L(x, \bar{\lambda}, \bar{\mu}).$$

L'inégalité de droite est (*1*). La première inégalité donne

$$f(\bar{x}) + (\lambda | h(\bar{x})) + (\mu | g(\bar{x})) \leq f(\bar{x}) + (\bar{\lambda} | h(\bar{x})) + (\bar{\mu} | g(\bar{x})).$$

Montrons par l'aburde que (*2*) est vérifiée. Si $h(\bar{x}) \neq 0$, prenons $\mu = 0$ et $\lambda = kh(\bar{x})$, alors l'inégalité implique que pour tout k on doit avoir $k||h(\bar{x})||^2 \leq (\bar{\lambda} | h(\bar{x}))$, ce qui est impossible.
Si maintenant il existe j_0 tel que $g_{j_0}(\bar{x}) > 0$ alors il suffit de prendre $\mu_j = \bar{\mu}_j$ pour $j \neq j_0$ et $\mu_{j_0} = kg_{j_0}(\bar{x})$ pour obtenir la contradiction. Quant-à (*3*), raisonnons toujours par l'absurde et supposons qu'il existe j_0 tel que $\bar{\mu}_{j_0} g_{j_0}(\bar{x}) < 0$, alors il suffit de poser $\mu_j = \bar{\mu}_j$ pour $j \neq j_0, \mu_{j_0} = 0$ et on obtient l'inégalité $0 \leq \bar{\mu}_{j_0} g_{j_0}(\bar{x})$.
Démontrons maintenant la réciproque. Soit donc un point qui vérifie les trois conditions. (*1*) implique immédiatement l'inégalité de droite dans la définition d'un point selle. Quant-aux conditions (*2*), (*3*) et (*4*), elles impliquent que $L(\bar{x}, \bar{\lambda}, \bar{\mu}) = f(\bar{x})$ et que

$$\forall (\lambda, \mu) \in \mathbb{R}^p \times (\mathbb{R}^+)^m, L(\bar{x}, \lambda, \mu) \leq f(\bar{x}).$$

■

Théorème 1.2.8

Si $(\bar{x}, \bar{\lambda}, \bar{\mu})$ *est un point selle du Lagrangien, alors* $\bar{x}$ *est une solution du problème d'optimisation* (P).

► Si $(\bar{x}, \bar{\lambda}, \bar{\mu})$ est un point selle, alors $\bar{x}$ vérifie les contraintes et on a $f(\bar{x}) = L(\bar{x}, \bar{\lambda}, \bar{\mu}) \leq L(x, \bar{\lambda}, \bar{\mu})$, pour tout x. Soit maintenant $x \in C$, on a

$$L(x, \bar{\lambda}, \bar{\mu}) = f(x) + (\bar{\mu} | g(x)) \leq f(x),$$

car $\mu_j g_j(x) \leq 0$ pour tout j. Par suite $\bar{x}$ est une solution de (P). ■

Remarque 1.2.2. La différence entre $KKT(\bar{x}, \bar{\lambda}, \bar{\mu})$ et la caractérisation d'un point selle du Lagrangien réside dans le fait que pour (KKT) on a $\nabla_x L(\bar{x}, \bar{\lambda}, \bar{\mu}) = 0$ alors que pour la caractérisation du point selle on a $\bar{x}$ est solution de $\min_x L(x, \bar{\lambda}, \bar{\mu})$. Par suite si le Lagrangien est convexe en x et si les fonctions sont C^1 on a l'équivalence entre $KKT(\bar{x}, \bar{\lambda}, \bar{\mu})$ et $(\bar{x}, \bar{\lambda}, \bar{\mu})$ point selle. On retrouve donc ici dans le cas convexe la condition suffisante de solution.

1.2.4 Dualité

Nous allons ici considérer les problèmes primal et dual associé au Lagrangien.

$$(P) \left\{ \begin{array}{l} \min \varphi(x) \\ x \in \mathbb{R}^n \end{array} \right. \quad (D) \left\{ \begin{array}{l} \max \psi(\lambda, \mu) \\ (\lambda, \mu) \in \mathbb{R}^p \times (\mathbb{R}^+)^m \end{array} \right.$$

avec

$$\begin{array}{rcl} \varphi : \mathbb{R}^n & \to & \bar{\mathbb{R}} \\ x & \mapsto & \varphi(x) = \sup_{(\lambda,\mu) \in \mathbb{R}^p \times (\mathbb{R}^+)^m} L(x, \lambda, \mu) \end{array} \qquad \begin{array}{rcl} \psi : \mathbb{R}^p \times (\mathbb{R}^+)^m & \to & \bar{\mathbb{R}} \\ (\lambda, \mu) & \mapsto & \psi(\lambda, \mu) = \inf_{x \in \mathbb{R}^n} L(x, \lambda, \mu). \end{array}$$

Concernant le problème primal on a

$$\varphi(x) = \begin{cases} +\infty & \text{si } h(x) \neq 0 \text{ ou } \exists j_0, g_{j_0}(x) > 0 \\ f(x) & \text{si } x \text{ vérifie les contraintes.} \end{cases}$$

Par suite le problème primal est équivalent au problème de départ.

Définition 1.2.9 – Problème dual

On appelle problème dual le problème (D).

Définition 1.2.10 – Saut de dualité

On appelle saut le dualité la quantité $f(x^*) - \psi(\lambda^*, \mu^*)$ où x^* est une solution du problème (P) et (λ^*, μ^*) est une solution du problème dual (D).

Remarque 1.2.3. Le problème dual est toujours un problème concave. En effet ψ est définie sur un domaine convexe et ψ est définie comme l'inf de fonctions affines et est donc concave.

On rappelle ici que (x^*, λ^*, μ^*) est un point selle du Lagrangien si et seulement si x^* est une solution du problème primal (P), (λ^*, μ^*) est une solution du problème dual (D) et si le saut de dualité est nul.

Théorème 1.2.11 – Théorème de dualité fort

Si (P) est un problème convexe dans le sens où f et g sont convexes et h est affine, si les fonctions f, g et h sont C^1, si l'ensemble des contraintes C est non vide et si (x^*, λ^*, μ^*) vérifie les conditions de (KKT) alors (x^*, λ^*, μ^*) est un point selle du Lagrangien, x^* est une solution du problème primal (P), (λ^*, μ^*) est une solution du problème dual (D) et le saut de dualité est nul.

✍ **Exercice 1.2.1**. On considère le problème suivant

$$(P_0) \left\{ \begin{array}{l} \min x_1^2 + x_2^2 \\ 2x_1 + x_2 \leq -4 \end{array} \right.$$

1. Visualiser les contraintes et les courbes de niveau de f.
2. Le problème est-il convexe ? strictement convexe ?
3. A-t-on l'hypothèse de qualification des contraintes ?
4. Caractériser la solution et résoudre le problème (P_0).
5. Écrire et résoudre le problème dual. □

1.3 Analyse post-optimale

1.3.1 Introduction

On considère dans cette section, pour simplifier la présentation le cas où le problème d'optimisation ne contient que des contraintes d'inégalité. On définie alors la famille de problèmes d'optimisation

$$(P_y) \left\{ \begin{array}{l} \min \ f(x) \\ g(x) \leq y. \end{array} \right.$$

On notera C_y l'ensemble des contraintes du problème (P_y).

Définition 1.3.1 – Fonction de perturbation, fonction valeur

On appelle fonction de perturbation, ou fonction valeur, associée à la famille de problèmes (P_y) la fonction v de $\mathbb{R}^m$ à valeurs dans $\bar{\mathbb{R}}$ qui à y associe

$$v(y) = \inf_{x \in C_y} f(x). \tag{1.4}$$

Si $C_y = \emptyset$, on pose $v(y) = +\infty$.

Proposition 1.3.2

La fonction de de perturbation possède la propriété de monotonie suivante : si $y \leq y'$ alors $C_y \subset C_{y'}$ et $v(y') \leq v(y)$.

1.3.2 Point selle et fonction valeur

Théorème 1.3.3

Si (P_0) admet une solution x^* alors (x^*, μ^*) est un point selle du Lagrangien si et seulement si l'hyperplan d'équation $z = v(0) - (\mu^*|y)$ est un hyperplan d'appui en $y = 0$ au graphe de le fonction de perturbation v (cf la figure 1.2), c'est-à-dire si et seulement si

$$\forall y \in \mathbb{R}^n, v(0) - (\mu^*|y) \leq v(y) \tag{1.5}$$

► 1. Implication
Soit (x^*, μ^*) un point selle, alors $\mu^* \geq 0$ et $L(x^*, \mu^*) \leq L(x, \mu^*), \forall x \in \mathbb{R}^n$. Mais comme $(\mu^*|g(x^*)) = 0$, on a $f(x^*) = v(0) = L(x^*, \mu^*) \leq f(x) + (\mu^*|g(x)), \forall x \in \mathbb{R}^n$. Donc, $\forall x \in \mathbb{R}^n, f(x^*) - (\mu^*|g(x)) \leq f(x)$.
Soit maintenant $y \in \mathbb{R}^m$. Alors, pour tout $x \in C_y$, on a $f(x) \geq v(0) - (\mu^*|g(x)) \geq v(0) - (\mu^*|y)$ car x vérifie les contraintes $g(x) \leq y$ et $\mu^* \geq 0$ (si $C_y = \emptyset$ alors l'inégalité est évidente).

2. Réciproque
Supposons que pour tout $y \in \mathbb{R}^m$, on ait $v(y) \geq v(0) - (\mu^*|y)$. Montrons tout d'abord que $\mu^* \geq 0$. Dans le cas contraire, il existe j_0 tel que $\mu^*_{j_0} < 0$ et donc en posant $y_{j_0} > 0$ et $y_j = 0$ si $j \neq j_0$, nous aurons $v(y) \geq v(0) - \mu^*_{j_0} y_{j_0} \to +\infty$ lorsque $y_{j_0} \to +\infty$. Or, si $y \geq 0$, on a $v(y) \leq v(0) = f(x^*) \in \mathbb{R}$ (x^* existe par hypothèse). D'où la contradiction.
Maintenant par définition de v on a pour tout $x \in \mathbb{R}^n$, $f(x) \geq v(g(x)) \geq v(0) - (\mu^*|g(x))$. Par suite pour tout $x \in \mathbb{R}^n$, $f(x) + (\mu^*|g(x)) \geq v(0) = f(x^*)$. En particulier, pour $x = x^*$ on a $(\mu^*|g(x^*)) \geq 0$. Mais comme $\mu^* \geq 0$ et $g(x^*) \leq 0$, on obtient $(\mu^*|g(x^*)) = 0$. De plus, on a $f(x) + (\mu^*|g(x)) \geq f(x^*) = f(x^*) + (\mu^*|g(x^*))$ pour tout $x \in \mathbb{R}^n$.
En conclusion (x^*, μ^*) est un point selle.

■

Théorème 1.3.4

Si v est dérivable en 0 et que (P) admet un point selle (x^*, μ^*) alors

$$\mu^*_j = -\frac{\partial v}{\partial y_j}(0) \tag{1.6}$$

► Le problème (P) admet un point selle donc l'hyperplan $z = v(0) - (\mu^*|y)$ est un hyperplan d'appui au graphe de v en $y = 0$. Mais v est dérivable en 0, donc l'hyperplan d'appui, s'il existe s'écrit

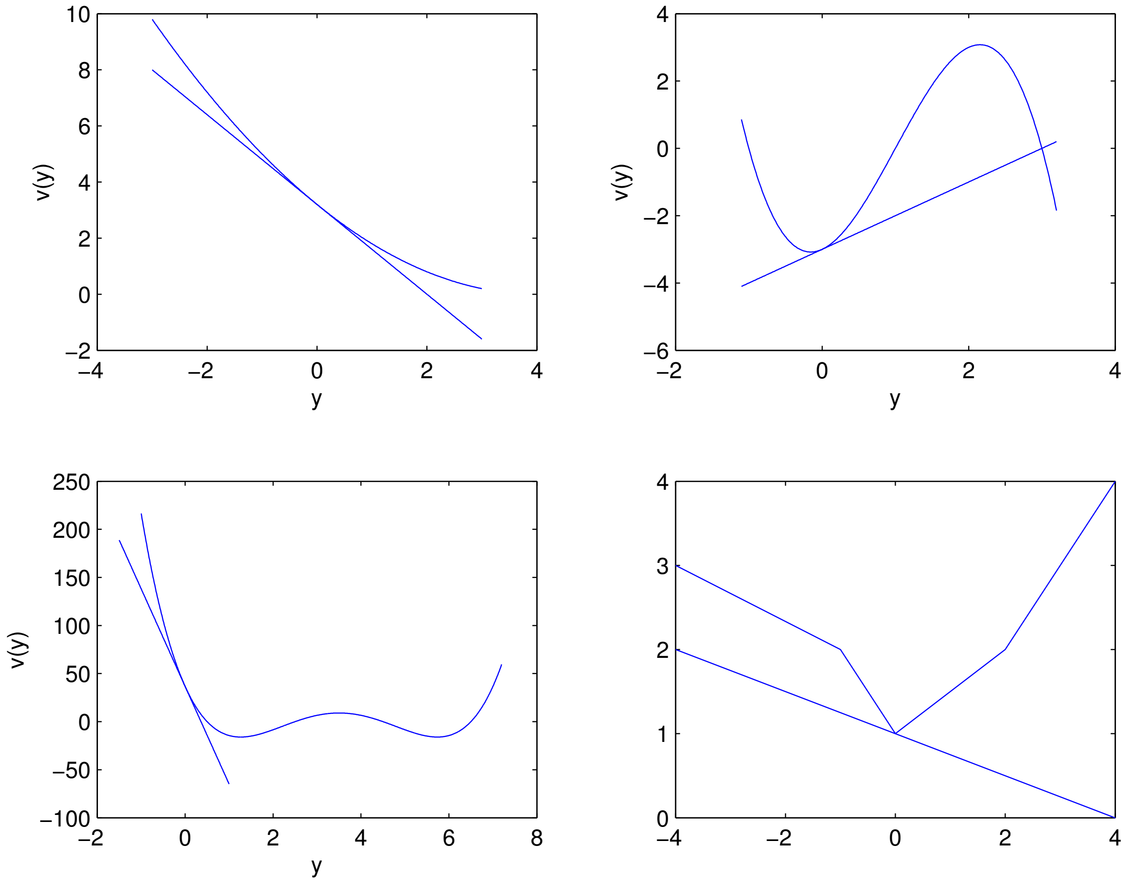

Figure 1.2 – *Exemples d'existence et de non existence d'hyperplan d'appui en* $y = 0$ *dans le cas où* $n = 1$*, à gauche, on a un hyperplan d'appui en 0, à droite en haut il n'y a pas d'hyperplan d'appui en 0, à droite en bas, il y a une infinité d'hyperplans d'appui en 0.*

$z = v(0) + v'(0).y = v(0) + (\nabla v(0)|y)$. D'où le résultat. ■

En conclusion, l'opposé du multiplicateur de lagrange μ_j^* est la dérivée de la fonction valeur par rapport à la perturbation y_j. Ceci nous permet localement de connaître le gain sur la fonction coût si on relâche la contrainte g_j.

✍ **Exercice 1.3.1**. On reprend l'exemple de l'exercice 1.2.1 que l'on perturbe :

$$(P_y) \begin{cases} \min x_1^2 + x_2^2 \\ 2x_1 + x_2 + 4 \leq y \end{cases}$$

1. Résoudre en fonction de y (P_y).

2. Calculer $\Phi(y)$ la valeur de la fonction f à la solution du problème (P_y).

3. Vérifier que $\Phi'(0) = -\mu^*$ où μ^* est la solution du problème dual associé au problème (P_0). □

1.4 Application à la programmation linéaire

1.4.1 Définitions

Exemple 1.4.1. Un brasseur produit de la bière blonde et de la bière brune. Trois ingrédients interviennent dans leurs fabrication : le maïs, le houblon et le malt, mais dans des proportions différentes. La table 1.1 donne les quantités nécessaires pour fabriquer un baril de bière. On suppose que les quantité

	bière blonde	bière brune
maïs (kg)	2.5	7.5
houblon (g)	125	125
malt (kg)	17.5	10

Table 1.1 – *Quantités nécessaires pour fabriquer un baril de bière.*

stockées par le brasseur sont 240 kgs de maïs, 5 kgs de houblon et 595 kgs de malt. Enfin, on suppose que le prix de vente d'un baril de bière blonde et d'un baril de bière brune sont respectivement de 65€et 115€. On suppose que tous les produits et bénéfices sont divisibles, le brasseur peut utiliser un demi kg de maïs, vendre un quart de baril de bière blonde...

On appelle x_1 et x_2 les quantités de bière blonde et de bière brune produite. Le problème est alors de résoudre le problème d'optimisation suivant

$$(P) \begin{cases} \max\ 65x_1 + 115x_2 & \text{maximisation du profit} \\ 2.5x_1 + 7.5x_2 \leq 240 & \text{contrainte sur le maïs} \\ 0.125x_1 + 0.125x_2 \leq 5 & \text{contrainte dur le houblon} \\ 17.5x_1 + 10x_2 \leq 595 & \text{contrainte sur le malt} \\ x_1 \geq 0, x_2 \geq 0. \end{cases}$$

□

Définition 1.4.1 – Problème de Programmation Linéaire

On appelle problème de programmation linéaire tout problème d'optimisation où toutes les fonctions qui interviennent sont affines

$$(PL) \begin{cases} \max\ (c|x) \\ A_1 x \leq b_1 \\ A_2 x = b_2 \end{cases}$$

Définition 1.4.2 – Problème sous forme standard

Un problème de programmation linéaire est dit sous forme standard s'il s'écrit

$$(PL)_S \begin{cases} \max\ (c|x) \\ Ax = b \\ x \geq 0. \end{cases}$$

Définition 1.4.3 – Problème sous forme canonique

Un problème de programmation linéaire est dit sous forme canonique s'il s'écrit

$$(PL)_C \begin{cases} \max\ (c|x) \\ Ax \leq b \\ x \geq 0. \end{cases}$$

Théorème 1.4.4

Il y a équivalence entre les trois formulations d'un problème linéaire.

► Tout d'abord, les formulations sous forme standard ou canonique sont des cas particuliers de la formulation générale. Il nous suffit donc de montrer que tout problème de programmation linéaire général s'écrit sous forme standard.
Pour chaque contraintes d'inégalité $a_{i1}x_1 + \cdots + a_{in}x_n \leq b_i$, on introduit une variable d'écart $z_i \geq 0$ telle que $a_{i1}x_1 + \cdots + a_{in}x_n + z_i = b_i$. Et pour tout x_j sans contrainte de positivité $x_j \geq 0$, on introduit deux variables $u_j \geq 0$ et $v_j \geq 0$ avec $x_j = u_j - v_j$. ■

1.4.2 Existence de solution

On considère ici le problème sous sa forme standard.

Théorème 1.4.5

Si l'ensemble des contraintes C n'est pas vide et que le problème de programmation linéaire $(PL)_S$ est majoré, alors il admet une solution.

► ■

✍ **Exercice 1.4.2**. Donnez un exemple, dans le cas non linéaire où le théorème ci-dessus est faux. □

Définition 1.4.6 – Point extrême

Soit C un convexe de $\mathbb{R}^n$ non vide. On appelle point extrême de C tout point $x \in C$ qui ne peut s'écrire comme le milieu d'un segment de C :

$$\forall (x_1, x_2) \in C^2, x_1 \neq x_2, x \neq \frac{1}{2}(x_1 + x_2).$$

Remarque 1.4.1. Il existe des convexe sans point extrême :

- la boule ouverte $B(0,1)$ ne possède aucun point extrême ;
- Le demi-plan fermé $= \{x \in \mathbb{R}^2, x_2 \geq 0\}$ ne possède aucun points extrêmes.

Définition 1.4.7 – Sommet

Si C est le polyèdre convexe $C = \{x \in \mathbb{R}^n, Ax \leq b\}$, un point extrême s'appelle aussi un sommet.

Exemple 1.4.3.

□

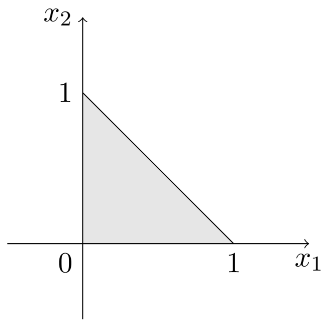

FIGURE 1.3 – *Polyèdre convexe* $C = \{x \in \mathbb{R}^2, x_1 \geq 0, x_2 \geq 0, x_1 + x_2 \leq 1\}$.

Théorème 1.4.8

L'ensemble des solutions d'un problème de programmation linéaire est convexe.

► Évident. ■

Théorème 1.4.9

Si le problème de programmation linéaire admet une solution, alors il existe une solution qui est un sommet du polyèdre convexe C.

Remarque 1.4.2. L'algorithme du simplexe [a] génère une suite de sommets qui fait croître la fonction coût. Comme le nombre de sommet est fini, l'algorithme s'arrête.

a. Dantzig (November 8, 1914 – May 13, 2005), 1947. cette même année John von Neumann (December 28, 1903 – February 8, 1957) inventa la théorie de la dualité en programmation linéaire !

1.4.3 Dualité

✍ **Exercice 1.4.4**. On considère le problème de programmation linéaire sous forme standard. En posant $X = \{x \in \mathbb{R}^n, x \geq 0\}$, écrire le problème dual et vérifier que le dual du dual est le problème primal. □

✍ **Exercice 1.4.5**. On considère le problème de programmation linéaire sous forme canonique. En posant $X = \{x \in \mathbb{R}^n, x \geq 0\}$, écrire le problème dual et vérifier que le dual du dual est le problème primal. □

Théorème 1.4.10

On considère le problème de programmation linéaire sous forme standard et son dual

$$(P) \begin{cases} \max\ (c|x) \\ Ax = b \\ x \geq 0 \end{cases} \qquad (D) \begin{cases} \min\ (b|\lambda) \\ A^T\lambda \geq c, \end{cases}$$

et on note respectivement C_P et C_D l'ensemble des contraintes des problèmes primal et dual. On a alors les différents cas suivants

1. *$C_P \neq \emptyset$ et $C_D \neq \emptyset$, alors (P) et (D) ont des solutions x^* et λ^*, (x^*, λ^*) est un point selle et*

$(c|x^*) = (b|\lambda^*)$. *Si* $x^* \in C_P, \lambda^* \in C_D$ *et si* $(c|x^*) = (b|\lambda^*)$, *alors* x^* *est une solution de* (P) *et* λ^* *est une solution de* (D).

2. $C_P = \emptyset$ et $C_D = \emptyset$.
3. $C_P = \emptyset$ et $C_D \neq \emptyset$, *alors* $\inf_{A^T\lambda \geq c}(b|\lambda) = -\infty$.
4. $C_P \neq \emptyset$ et $C_D = \emptyset$, *alors* $\sup_{Ax=b, x\geq 0}(c|x) = +\infty$.

► Remarquons tout d'abord que (P) est un problème convexe et que l'on a toujours l'hypothèse de qualification des contraintes. Par suite, une condition nécessaire et suffisant de solution est (KKT) et on a une solution si et seulement si on a un point selle.

1. Le théorème faible de dualité implique $(\psi(y) \leq \varphi(x))$ que pour tout $x \in X$ et $\lambda \in \mathbb{R}^n$, $-(b|\lambda) \leq -(c|x)$. Par suite $(c|x)$ est majoré et (P) admet une solution. D'où le résultat.
2.
3. Dans ce cas C_D n'a pas de solution, il est donc non minoré.
4. Dans ce cas C_P n'a pas de solution, il est donc non majoré.

■

Théorème 1.4.11

On considère le problème de programmation linéaire sous forme canonique et son dual

$$(P)\begin{cases} \max\ (c|x) \\ Ax \leq b \\ x \geq 0 \end{cases} \qquad (D)\begin{cases} \min\ (b|\mu) \\ A^T\mu \geq c \\ \mu \geq 0. \end{cases}$$

Si (P) *a une solution* x^* *(et donc* (D) *a une solution* μ^**), alors :*

1. *si* $\mu_j > 0$ *alors* $a_{i1}x_1^* + \cdots a_{in}x_n^* = b_i$;
2. *si* $x_j^* > 0$, *alors* $a_{1j}\mu_1^* + \cdots a_{mj}\mu_m^* = c_j$.

► Cela provient tout simplement des conditions de complémentarité de (KKT). ■

Théorème 1.4.12

On considère un problème de programmation linéaire sous forme canonique

$$(P)\begin{cases} \max\ (c|x) \\ Ax \leq b_0 \\ x \geq 0 \end{cases}$$

On suppose qu'il existe $\varepsilon > 0$ *tel que pour tout* $b \in B(b_0, \varepsilon)$ *le problème perturbé*

$$(P)\begin{cases} \max\ (c|x) \\ Ax \leq b \\ x \geq 0 \end{cases}$$

admet une solution $x^*(b)$, *alors si la fonction valeur* $v(b) = (c|x^*(b))$ *est dérivable, on a* $v'(b_0) = \mu^*$.

► Il suffit d'appliquer le théorème 1.3.2. ■

Remarque 1.4.3. Si pour tout i, on a pas en même temps $\mu_i^* = 0$ et $(Ax - b_0)_i = 0$ alors $v'(b_0)$ existe.

✍ **Exercice 1.4.6**. On s'intéresse ici au problème du brasseur 1.4.1

1. Résoudre graphiquement le problème.

2. on pose $b_0 = \begin{pmatrix} 240 & 5 & 595 \end{pmatrix}^T$. Calculer sans résoudre le problème dual μ^*.

3. On prend maintenant $b \in \mathbb{R}^3$ suffisamment proche de b_0. Calculer la solution du problème de programmation linéaire

$$(P) \begin{cases} \max \ 65x_1 + 115x_2 \\ 2.5x_1 + 7.5x_2 \leq b_1 \\ 0.125x_1 + 0.125x_2 \leq b_2 \\ 17.5x_1 + 10x_2 \leq b_3 \\ x_1 \geq 0, x_2 \geq 0. \end{cases}$$

en fonction de b. En déduire la fonction valeur $v(b)$ et vérifiez que $v'(b_0) = \mu^*$.

4. En déduire combien le brasseur gagnera en plus si :

1. quantité stocké en houblon passe à 6 kg ;
2. quantité stocké en malt passe à 596 kg.

□

Bibliographie

---

[1] Jean-Baptiste Hiriart-Urruty. *L'Optimisation*. Que sais-je. Presses Universitaires de France, 1996. ISBN : 2 13 047981 2. ↩ 1.