

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE  
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA  
RECHERCHE SCIENTIFIQUE

UNIVERSITÉ A. MIRA DE BÉJAÏA

FACULTÉ DES SCIENCES EXACTES  
DÉPARTEMENT DE RECHERCHE OPÉRATIONNELLE

MÉMOIRE DE MAGISTER

EN

Mathématiques Appliquées

OPTION :

Modélisation Mathématique et Techniques de Décision

**Thème :**

**Application de la méthode dc pour  
la résolution du jeu de collusion**

*Présenté par :*

*M<sup>r</sup> L'hadi Boughani*

*Devant le jury composé de :*

<b>Président</b>	M <sup>r</sup>	Djamil	AÏSSANI	Professeur	U. Béjaïa
<b>Rapporteur</b>	M <sup>r</sup>	Mohammed Saïd	RADJEF	Professeur	U. Béjaïa
<b>Examineur</b>	M <sup>r</sup>	Mohand Ouamer	BIBI	Professeur	U. Béjaïa
<b>Examineur</b>	M <sup>r</sup>	Farid	YAICI	Professeur	U. Béjaïa

Béjaïa, 2008.

# Remerciements

Mes remerciements vont d'abord au professeur M.S. Radjef de m'avoir encadré et accompagné tout au long de ce travail et pour ses précieux conseils. Je salue ici sa grande patience.

Mes remerciements s'adressent aussi au professeur D.Aïssani d'avoir accepté de juger ce travail.

Je tiens à remercier les professeurs M.O. Bibi et F. Yaici d'avoir accepté de juger ce travail, qu'ils trouvent ici ma profonde gratitude.

Je remercie tous les membres de ma famille et mes amis pour leurs soutien et encouragements.

Que tous ceux qui ont contribué de prêt ou de loin à l'aboutissement de ce travail trouvent ici l'expression de mes vifs remerciements.

# Table des matières

<b>Introduction générale</b>	<b>1</b>
<b>1 GÉNÉRALITÉS.</b>	<b>4</b>
1.1 Définitions . . . . .	4
1.1.1 Ensembles convexes . . . . .	4
1.1.2 Fonctions convexes . . . . .	5
1.1.3 Les Formes Quadratiques . . . . .	8
1.1.4 Notions de préférence et optimum de Pareto . . . . .	14
1.2 Formulation générale des problèmes d'optimisation non linéaire .	15
1.2.1 Rappels de calcul différentiel . . . . .	16
1.2.2 Notions sur la convexité . . . . .	18
1.2.3 Résultats d'existence et d'unicité . . . . .	19
1.2.4 Conditions nécessaires d'optimalité en absence de contraintes	20
1.2.5 Conditions d'optimalité en optimisation avec contraintes .	21
1.2.6 Les conditions de Karush-Kuhn-Tucker . . . . .	24
1.2.7 Conditions suffisantes d'optimalité . . . . .	25
<b>2 Méthodes Numériques.</b>	<b>27</b>
2.1 Méthode des pénalités . . . . .	27
2.1.1 Méthodes de pénalité extérieure . . . . .	28
2.1.2 Méthodes de pénalité intérieure . . . . .	32
2.2 Méthode DC . . . . .	36
2.2.1 Les fonctions DC . . . . .	37
2.2.2 Les problèmes de Programmation DC . . . . .	39
2.2.3 Conditions d'optimalité en optimisation DC . . . . .	39
2.2.4 Dualité en optimisation DC . . . . .	42
2.2.5 Existence et bornitude des suites générées par DCA . . . .	46
<b>3 Rappels sur les jeux.</b>	<b>49</b>
3.1 Définitions . . . . .	49
3.2 Représentations . . . . .	50

3.2.1	Jeux sous forme normale (stratégique) . . . . .	50
3.3	Équilibre de Nash et fonction de meilleure réponse . . . . .	53
3.4	Jeux répétés (superjeux) . . . . .	55
3.4.1	Les stratégies dans un jeu répété . . . . .	55
3.4.2	Notion d'équilibre dans le jeu répété . . . . .	57
3.4.3	Crédibilité et sous-jeu parfait . . . . .	58
3.4.4	Profils de stratégies simples . . . . .	59
3.4.5	Le principe de déviation unique . . . . .	61
<b>4</b>	<b>Concurrence oligopolistique et modélisation.</b>	<b>63</b>
4.1	Éléments de microéconomie . . . . .	63
4.1.1	Propriétés d'un marché concurrentiel (Concurrence parfaite) . . . . .	64
4.1.2	Notions d'offre et demande globale et équilibre de marché	64
4.1.3	Notion de l'équilibre du marché . . . . .	65
4.2	Cournot et la concurrence en quantité . . . . .	66
4.3	La concurrence en prix . . . . .	70
4.4	Passage aux modèles dynamiques . . . . .	74
4.5	Collusion et modélisation . . . . .	75
4.6	Conclusion . . . . .	77
<b>5</b>	<b>Résolution du jeu de collusion.</b>	<b>78</b>
5.1	Première Approche : Algorithme(PPA)(Penalized Projection Algorithm) . . . . .	82
5.1.1	Description de l'algorithme PPA . . . . .	83
5.2	Approche basée sur DCA- Algorithme DCPA (DC Penalized Algorithm) . . . . .	86
5.2.1	Description de l'algorithme DCAP . . . . .	90
	<b>Conclusion</b>	<b>92</b>
	<b>Bibliographie</b>	<b>92</b>

# INTRODUCTION GÉNÉRALE.

Le Marché oligopolistique est l'un des champs qui a suscité beaucoup de travaux de recherches. Parmi ces travaux, on cite les développements concernant la modélisation de la concurrence oligopolistique.

Les premières théories de l'oligopole se sont développées dans un cadre statique, et l'un des modèles les plus célèbres dans ce domaine est celui de Cournot (1838)[16] dans lequel il propose une analyse de la concurrence par courbes de meilleures réponses ce qui donne naissance à la notion d'équilibre de Cournot plus tard connue sous le nom de l'équilibre Cournot-Nash. Shapiro [50] rappelle « qu'on ne peut pas analyser correctement un oligopole dynamique en utilisant des courbes de réaction à la Cournot, car il est fort improbable qu'elles représentent des réponses dynamiques optimales » (Friedman [23]). Comme tout équilibre de Nash, l'équilibre de Cournot correspond à une situation dans laquelle aucune firme n'a intérêt à modifier ses choix ou ses actions. Sur ce point, Léonard [36] montre combien les travaux de Nash (qui d'ailleurs n'avait jamais lu Cournot) ont modifié la perception de la solution proposée par Cournot. L'analyse de Cournot, très critiquée entre autres par Fellner [21] qui la jugeait dynamique, contradictoire et non réaliste, a été réinterprétée et reconstruite dans les années 60 comme une solution statique et cohérente pouvant fournir un fondement historique à l'équilibre de Nash.

Dans un duopole, cet équilibre peut être localisé comme le point d'intersection des courbes de meilleure réponse, mais en aucun cas il ne peut être pensé comme un processus dynamique de coordination. Shapiro constate que les variations conjecturales (Bowley [13]) ou la courbe de demande coudée (Sweezy [51]) correspondent aussi à une approche statique de la concurrence oligopolistique. on discutera brièvement ces modèles dans le chapitre 4.

Les jeux répétés à horizon infini sont à l'origine de la première approche dynamique de la concurrence oligopolistique. Cette approche s'est essentiellement intéressée à l'interdépendance stratégique des décisions des firmes et à leur impact sur les possibilités de collusion. Elle permet de renouveler l'analyse classique des structures de marché et des comportements collusifs, menée jusqu'alors dans un cadre statique et souvent en dehors de la théorie des jeux.

Plusieurs auteurs ont travaillé sur la modélisation d'un marché concurrentiel avec

possibilité de collusion. Parmi ces modèles on cite celui de Friedman [23, 22] dans lequel il propose de jouer pour la première fois une action qui donne un profit Pareto optimal pour les firmes et de rejouer cette dernière jusqu'à ce qu'une déviation soit enregistrée. Dès que cette dernière est identifiée, les firmes devront opter pour l'action qui infligera au déviant la pire perte et cette action n'est autre que l'équilibre de Cournot-Nash.

Ce modèle est le plus simple des modèles qui existent, mais pose deux problèmes principaux : quel serait l'objectif des firmes et comment faire la sélection du vecteur des actions menant à des profit Pareto optimaux.

Harrington [27] propose de résoudre un problème d'optimisation dont la fonction objectif serait de maximiser l'objectif de négociation de Nash sous contrainte que les actions doivent être choisies dans l'ensembles des quantités d'équilibre dans le cadre où le choix des firmes sont les quantités à produire.

Dans le chapitre 4, on présentera un modèle similaire au modèle d'Harrington pour lequel on injecte une modification à fin de remédier au problème du choix du vecteur prix qui arrange au mieux les objectifs des firmes. Dans le chapitre suivant 5, on essayera de discuter la solution du modèle construit via la méthode DC.

La programmation dc et DCA (famille des algorithmes dc) sont introduits par Pham Dinh Tao en 1986. Ces outils théoriques et algorithmiques sont intensivement développés à partir de 1993 par Le Thi Hoai An et Pham Dinh Tao (voir par exemple [8, 52] et références incluses) pour devenir maintenant classiques et de plus en plus populaires.

La plus part des problèmes d'optimisation peuvent être formulés comme des programmes DC.

En effet, en général on trouve les problèmes de la vies courante sous l'une des formes suivantes :

- $\max\{f(x) : x \in X\}$ , où  $f$  et  $X$  sont convexes,
- $\min\{f(x) = g(x) - h(x) : x \in \mathbb{R}^n\}$  où  $g$  et  $h$  sont convexes,
- $\min\{f(x) : x \in X\}$  où  $X$  est convexe et  $f$  quelconque.

Le deuxième type des modèles ci-dessus est dit programme de programmation dc, et les algorithmes traitant avec ces programmes appartiennent à une famille de programmes appelée algorithmes DCA.

Il est très facile de voire la relation qui existe entre ces trois modèles. En fait, on peut facilement constater qu'on peut réécrire le premier et le dernier modèle ci-dessus comme le deuxième modèle, d'où l'intérêt de la méthode. De plus cette dernière a montrer son efficacité par ces application nombreuses dans la littérature.

Le mémoire s'organise comme suit :

- Le premier chapitre regroupe les définitions nécessaires pour la suite de notre mémoire et les résultats principaux de l'optimisation non linéaire et les conditions d'optimalité
- Le deuxième chapitre donne les fondements théoriques de la méthode des pénalité

qu'on utilisera par la suite pour résoudre notre problème ainsi que les principaux résultats de la programmation dc et DCA sur lequel se base la conception de notre algorithme.

- Dans Le troisième chapitre, on exposera brièvement des notions utiles de la théorie de jeux nécessaire pour la compréhension du chapitre suivant.
- Le quatrième chapitre comportera le modèle du jeu de collusion à résoudre
- Le cinquième chapitre comportera la solution du modèle et l'élaboration d'algorithmes de résolution et on termine ce mémoire par une conclusion.

# 1

## GÉNÉRALITÉS.

Dans ce chapitre, nous rappellerons brièvement quelques notions qui nous seront utiles pour la suite de notre mémoire.

### 1.1 Définitions

#### 1.1.1 Ensembles convexes

**Définition 1.1** On dira qu'une partie  $C$  de  $\mathbb{R}^n$  est convexe si, pour tout  $x, y \in C$  le segment

$$[x, y] = \{x' = (1-t)x + ty, \quad t \in [0, 1]\}$$

est contenu dans  $C$ .

**Définition 1.2** Soit  $C \subset \mathbb{R}^n$ . L'enveloppe convexe de  $C$ , notée  $\text{conv}(C)$ , est l'ensemble des combinaisons convexes finies des éléments de  $C$ , c'est-à-dire

$$\text{conv}(C) = \left\{ \sum_{i=1}^n \lambda_i x^i : \lambda_i \in \mathbb{R}^+, \quad x^i \in C, \forall i = 1, 2, \dots, n \text{ et } \sum_{i=1}^n \lambda_i = 1 \right\}.$$

**Définition 1.3** Soit  $C$  un ensemble convexe de  $\mathbb{R}^n$ . On définit la variété linéaire engendrée par  $C$  comme étant l'ensemble

$$\text{aff}(C) = \left\{ \sum_{i=1}^n \lambda_i x^i, \quad \lambda_i \in \mathbb{R}, \quad x^i \in C \forall i = 1, 2, \dots, n \text{ et } \sum_{i=1}^n \lambda_i = 1 \right\}.$$

**Définition 1.4** On appelle intérieur relatif d'un ensemble convexe  $C$ , son intérieur dans  $\text{aff}(C)$ , muni de la topologie induite de celle de  $\mathbb{R}^n$ . On le note

$$\text{ir}(C) = \{x \in C, \exists r > 0 \text{ tel que } B(x, r) \cap \text{aff}(C) \subset C\},$$

où  $B(x, r) = \{y : \|x - y\| \leq r\}$  est la boule de centre  $x$  et de rayon  $r$ .

**Remarque 1.1**  $\text{ir}(C)$  est vide si et seulement si  $C$  l'est.

## 1.1.2 Fonctions convexes

Soit  $X$  un sous ensemble convexe de  $\mathbb{R}^n$  et  $Y = X^*$  son dual

**Définition 1.5** Soit  $f : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$  une fonction définie sur l'ensemble convexe  $X$ . On appelle domaine de  $f$  l'ensemble

$$\text{dom}(f) = \{x \in X : f(x) < +\infty\}.$$

**Définition 1.6** Une fonction  $f : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$  est dite propre si elle ne prend jamais la valeur  $-\infty$  et n'est pas identiquement égale à  $+\infty$ .

**Définition 1.7** L'épigraphe d'une fonction  $f : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$  est l'ensemble

$$\text{epi}(f) = \{(x, \alpha) \in X \times \mathbb{R} : f(x) \leq \alpha\}.$$

**Définition 1.8** Une fonction  $f : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$  est dite convexe si son épigraphe est un sous-ensemble convexe de  $X \times \mathbb{R}$ . Ceci équivaut à dire que l'on a l'inégalité

$$f((1 - \lambda)x + \lambda y) \leq (1 - \lambda)f(x) + \lambda f(y) \text{ pour tout } x, y \in X \text{ et tout } \lambda \in [0, 1].$$

Si l'inégalité est stricte dans l'équation précédente pour tout  $\lambda \in ]0, 1[$  et pour tout  $x, y \in C$  avec  $x \neq y$  alors la fonction  $f$  est dite strictement convexe.

**Remarque 1.2** Une fonction  $f$  est dite concave si  $-f$  est convexe

**Définition 1.9** Étant donnée une fonction convexe propre  $f$  sur l'ensemble convexe  $X$ , on dira que  $y^0 \in Y$  est un sous-gradient de  $f$  en  $x^0 \in \text{dom}(f)$  si

$$\langle y^0, x - x^0 \rangle + f(x^0) \leq f(x), \quad \forall x \in X.$$

L'ensemble de tous les sous-gradients de  $f$  au point  $x^0$  est le sous-différentiel de  $f$  au point  $x^0$  et on note  $\partial f(x^0)$ . Ainsi

$$\partial f(x^0) = \{y \in Y : f(x) \geq f(x^0) + \langle x - x^0, y \rangle, \forall x \in X\}.$$

Le domaine du sous-différentiel de la fonction  $f$  est

$$\text{dom}(\partial f) = \{x : \partial f(x) \neq \emptyset\}.$$

**Définition 1.10** Soit  $\varepsilon$  un réel positif. Un élément  $y^0$  de  $Y$  est appelé  $\varepsilon$ -sous-gradient de  $f$  au point  $x^0$  si

$$\langle y^0, x - x^0 \rangle + f(x^0) - \varepsilon \leq f(x) \quad \forall x \in X.$$

On note  $\partial_\varepsilon f(x^0)$  l'ensemble de tous les  $\varepsilon$ -sous-gradients de  $f$  au point  $x^0$ .

**Définition 1.11** Soit  $X$  un sous-ensemble convexe de  $\mathbb{R}^n$ ,  $\tilde{x} \in \mathbb{R}^n$  et  $\varepsilon \geq 0$ . L'ensemble  $N_\varepsilon(X, \tilde{x})$  défini par

$$N_\varepsilon(X, \tilde{x}) = \{d \in \mathbb{R}^n : \langle d, x - \tilde{x} \rangle \leq \varepsilon \quad \forall x \in X\}, \quad (1.1)$$

est dit cône  $\varepsilon$ -normal de  $X$  en  $\tilde{x}$ . Si  $\varepsilon = 0$ , cet ensemble définit le cône normal de  $X$  en  $\tilde{x}$ .

**Définition 1.12** Une fonction  $f : D \rightarrow \mathbb{R}$  est quasi-convexe sur l'ensemble convexe  $D$  si

$$\forall x_1, x_2 \in D, f(x_1) \geq f(x_2) \Rightarrow f(x_1) \geq f((1 - \lambda)x_1 + \lambda x_2), \quad \forall \lambda \in [0, 1],$$

autrement dit, si

$$f((1 - \lambda)x_1 + \lambda x_2) \leq \max\{f(x_1), f(x_2)\}.$$

Elle est dite quasi-concave si et seulement si

$$\forall x_1, x_2 \in D, \quad f(x_1) \geq f(x_2) \Rightarrow f(x_1) \leq f((1 - \lambda)x_1 + \lambda x_2), \quad \forall \lambda \in [0, 1],$$

autrement dit :

$$f((1 - \lambda)x_1 + \lambda x_2) \geq \min\{f(x_1), f(x_2)\}, \quad \forall x_1, x_2 \in D, \quad \forall \lambda \in [0, 1].$$

conséquence

- Une fonction convexe est quasi-convexe.
- Une fonction concave est quasi-concave.

**Définition 1.13** Une fonction  $f : X \rightarrow \mathbb{R}$  est dite semi-continue inférieurement (s.c.i) en un point  $x \in X$  si

$$\liminf_{y \rightarrow x} f(y) \geq f(x).$$

Notons par  $\Gamma_0(X)$  l'ensemble de toutes les fonctions semi-continues inférieurement, propres et convexes sur  $X$ .

**Définition 1.14** La fonction conjuguée  $f^*$  de  $f \in \Gamma_0(X)$  est une fonction appartenant à  $\Gamma_0(X)$  et définie comme suit :

$$f^*(y) = \sup\{\langle x, y \rangle - f(x) : x \in X\}.$$

**Propriété 1.1** [48]

On a les propriétés suivantes :

1.  $f^*$  est toujours convexe.
2. Si  $f$  prend la valeur  $-\infty$  alors  $f^*$  est identiquement égale à  $+\infty$ .
3. On a  $(f^*)^*(x) \leq f(x)$ ,  $\forall x \in X$ .

On a également la proposition suivante

**Proposition 1** [48]

Soit  $f : C \rightarrow \mathbb{R}$ , alors :

1.  $\partial f(x)$  est une partie convexe fermée de  $Y$ .
2.  $f \in \Gamma_0(X) \iff f^* \in \Gamma_0(Y)$ . Dans ce cas on a  $f = f^{**}$ .
3.  $y \in \partial f(x) \iff f(x) + f^*(y) = \langle x, y \rangle$ .
4. Si  $f \in \Gamma_0(X)$  alors  $y \in \partial f(x) \iff x \in \partial f^*(y)$ .
5. Si  $f \in \Gamma_0(X)$  et si  $\partial f(x)$  est réduit à un singleton  $\{y\}$ , alors  $f$  est différentiable en  $x$  et  $\nabla f(x) = y$  et réciproquement.
6.  $f(x^0) = \min\{f(x) : x \in X\} \iff 0 \in \partial f(x^0)$ .

**Fonctions convexes polyédrale**

**Définition 1.15** Une partie convexe  $C \in X$  est dite polyèdre convexe si

$$C = \bigcap_{i=1}^m \{x \mid \langle a_i, x \rangle - b_i \leq 0, a_i \in Y, b_i \in \mathbb{R}\}.$$

Une fonction  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{-\infty, +\infty\}$  est dite polyédrale si son épigraphe est un ensemble polyédral de  $\mathbb{R}^{n+1}$ .

**Remarque 1.3** : On note que toute fonction polyédrale est propre, convexe et s.c.i.

**Proposition 2** [38]

Soit  $f : X \rightarrow \mathbb{R}$  une fonction convexe. Alors  $f$  est polyédrale si et seulement si  $\text{dom}(f)$  est un ensemble convexe polyédral, et

$$f(x) = \sup\{\langle a_i, x_i - b_i \rangle : i = 1, 2, \dots, l\}, \forall x \in \text{dom}(f).$$

### Proposition 3 [38]

- Si  $f$  est convexe polyédrale alors  $f^*$  l'est aussi. De plus, si  $f$  est partout finie, alors

$$\text{dom}(f^*) = \text{conv}(K) \quad k = \{a_j : j = 1, 2, \dots, l\},$$

$$f^*(y) = \min\{y = \sum_{i=1}^l \lambda_i a_i, \lambda_i \geq 0 \text{ et } \sum_{i=1}^l \lambda_i = 1\}.$$

- Si  $f$  est polyédrale, alors  $\partial f(x)$  est une partie convexe polyédrale non vide en tout point de  $\text{dom}(f)$ .
- Si  $f_1, \dots, f_s$  sont des fonctions convexes polyédrales sur  $X$  telles que les ensembles convexes  $\text{dom}(f_i)$ ,  $i = 1, 2, \dots, s$  ont un point commun, alors

$$\partial(f_1 + \dots + f_s)(x) = \partial f_1(x) + \dots + \partial f_s(x), \quad \forall x \in X.$$

### 1.1.3 Les Formes Quadratiques

#### Définition 1.16

Une fonction réelle  $F : \mathbb{R}^n \rightarrow \mathbb{R}$ , est une forme quadratique de  $n$  variables  $x_1, x_2, \dots, x_n$  si elle s'écrit sous la forme suivante :

$$\begin{aligned} F(x) &= \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j \\ &= x^T A x \end{aligned} \tag{1.2}$$

où  $x^T = (x_1, x_2, \dots, x_n)$  un  $n$ -vecteur ligne et  $A = (a_{ij}, 1 \leq i, j \leq n)$  une matrice carrée d'ordre  $n$ .

Pour  $i \neq j$ , le coefficient du terme  $x_i x_j$  est égale à  $(a_{ij} + a_{ji})$ . En vertu de cela, la matrice  $A$  est toujours supposée symétrique. En effet, en définissant de nouveaux coefficients

$$d_{ij} = d_{ji} = \frac{a_{ij} + a_{ji}}{2}, \quad 1 \leq i, j \leq n$$

On obtient une nouvelle matrice  $D$  symétrique telle que

$$D = (d_{ij}, 1 \leq i, j \leq n), \quad \text{avec } d_{ij} + d_{ji} = a_{ij} + a_{ji}.$$

Il est clair qu'après une redéfinition des coefficients, la valeur de la forme quadratique  $F(x)$  reste inchangée pour tout point  $x$  dans  $\mathbb{R}^n$  :

$$F(x) = x^T A x = x^T D x.$$

Pour cela, il est naturel de considérer que la matrice associée à une forme quadratique est toujours symétrique.

## Gradient d'une forme quadratique

**Définition 1.17** Soit  $F : \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction réelle continûment différentiable. Le gradient de cette dernière au point  $x$  est défini par :

$$\nabla F(x) = \begin{pmatrix} \frac{\partial F}{\partial x_1} \\ \frac{\partial F}{\partial x_2} \\ \vdots \\ \frac{\partial F}{\partial x_n} \end{pmatrix} \quad (1.3)$$

Soit une forme quadratique avec une matrice  $D$  symétrique associée :

$$F(x) = x^T D x \quad (1.4)$$

En écrivant la matrice  $D$  sous forme de vecteurs colonnes

$$D = (d_1, d_2, \dots, d_n),$$

L'expression (1.4) peut se mettre sous la forme suivante :

$$F(x) = (x_1, x_2, \dots, x_j, \dots, x_n) \begin{pmatrix} d_1^T x \\ d_2^T x \\ \vdots \\ d_j^T x \\ \vdots \\ d_n^T x \end{pmatrix} = \sum_{j=1}^n x_j d_j^T x.$$

Calculons alors les dérivées partielles de  $F$  par rapport à chaque variable  $x_j$  :

$$\begin{aligned} \frac{\partial F}{\partial x_j} &= x_1 d_{1j} + x_2 d_{2j} + \dots + x_{j-1} d_{(j-1)(j)} + d_j^T x + x_j d_{jj} + \dots + x_n d_{nj} \\ &= x_1 d_{1j} + x_2 d_{2j} + \dots + x_{j-1} d_{(j-1)(j)} + x_j d_{jj} + \dots + x_n d_{nj} + d_j^T x \\ &= 2d_j^T x \end{aligned}$$

Par conséquent, le gradient de  $F(x)$  est :

$$g(x) = \nabla F(x) = 2 \begin{pmatrix} d_1^T x \\ d_2^T x \\ \vdots \\ d_j^T x \\ \vdots \\ d_n^T x \end{pmatrix} = 2Dx \quad (1.5)$$

**Définition 1.18** Soit une fonction réelle de classe  $C^2$ ,  $F : \mathbb{R}^n \rightarrow \mathbb{R}$ . Le Hessien de la fonction  $F$  est défini par :

$$\begin{aligned} \nabla^2 F(x) &= \left( \nabla \frac{\partial F}{\partial x_1}, \nabla \frac{\partial F}{\partial x_2}, \dots, \nabla \frac{\partial F}{\partial x_j}, \dots, \nabla \frac{\partial F}{\partial x_n} \right) \\ &= \begin{pmatrix} \frac{\partial^2 F}{\partial x_1^2} & \frac{\partial^2 F}{\partial x_1 x_2} & \dots & \frac{\partial^2 F}{\partial x_1 x_n} \\ \frac{\partial^2 F}{\partial x_2 x_1} & \frac{\partial^2 F}{\partial x_2^2} & \dots & \frac{\partial^2 F}{\partial x_2 x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 F}{\partial x_n x_1} & \frac{\partial^2 F}{\partial x_n x_2} & \dots & \frac{\partial^2 F}{\partial x_n^2} \end{pmatrix} \end{aligned} \quad (1.6)$$

**Définition 1.19** Soit  $F : \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction de classe  $C^1$ . La dérivée directionnelle de  $F$  dans la direction  $d$  au point  $x$  est :

$$\begin{aligned} \frac{\partial F}{\partial d} &= \lim_{t \rightarrow 0} \frac{F(x + td) - F(x)}{t} \\ &= \frac{\partial F}{\partial x_1}(x + td) \Big|_{t=0} d_1 + \dots + \frac{\partial F}{\partial x_n}(x + td) \Big|_{t=0} d_n \\ &= \nabla F(x)^T d \end{aligned}$$

### Formes quadratique définies et semis définies

Soit la forme quadratique  $F(x) = x^T D x$

**Définition 1.20**  $F(x)$  est dite définie positive si  $x^T D x > 0$ ,  $\forall x \in \mathbb{R}^n$  et  $x \neq 0$ . Elle est dite semi-définie positive ou définie non négative si  $x^T D x \geq 0$ ,  $\forall x \in \mathbb{R}^n$ .

**Définition 1.21**  $F(x)$  est dite définie négative si  $x^T D x < 0$ ,  $\forall x \in \mathbb{R}^n$  et  $x \neq 0$ . Elle est dite semi-définie négative ou définie non positive si  $x^T D x \leq 0$ ,  $\forall x \in \mathbb{R}^n$ .

### Matrices définies positives et non négatives et leurs propriétés

**Définition 1.22** Une matrice symétrique  $D$  est dite matrice définie positive (non négative) et se note  $D > 0$  ( $D \geq 0$ ), si elle est associée à une forme quadratique définie positive (non négative).

**Définition 1.23** Une forme quadratique  $F(x)$  est dite non définie si  $F(x)$  est positive pour certaines valeurs de  $x$  et négative pour d'autres.

Les matrices symétriques définies ont des propriétés très intéressantes. En voici quelques unes :

**Propriété 1.2** Soit une matrice symétrique  $D = (d_{ij}, 1 \leq i, j \leq n)$ . Si  $D$  est définie positive (non négative), alors on a :

$$d_{ii} > 0 \quad (d_{ii} \geq 0), \forall i = 1, \dots, n.$$

**Preuve 1** Soit le vecteur unitaire  $e_i = (0, 0, \dots, 1^i, \dots, 0) \neq 0$ . Puisque  $D > 0$ , on aura alors  $e_i^T D e_i > 0$ . D'autre part, calculons

$$D e_i = \begin{pmatrix} d_{1i} \\ d_{2i} \\ \vdots \\ d_{ii} \\ \vdots \\ d_{ni} \end{pmatrix} \Rightarrow e_i^T D e_i = d_{ii}.$$

Il s'ensuit que

$$d_{ii} > 0, \quad \forall i = 1, \dots, n.$$

Pour  $D \geq 0$ , on obtient donc  $d_{ii} \geq 0, \forall i = 1, \dots, n.$  ■

**Théorème 1.1** Une matrice  $D$  symétrique est définie positive si et seulement si toutes ses valeurs propres sont strictement positives.

**Preuve 2**

1. Supposons  $D$  symétrique définie positive. Soit  $\lambda_i$  une valeur propre associée au vecteur propre  $x_i$ . On a donc  $D x_i = \lambda_i x_i$ , et  $x_i^T D x_i = \lambda_i x_i^T x_i = \lambda_i \|x_i\|^2 > 0$
2. Réciproquement, si toutes les valeurs propres sont positives, on a pour tout vecteur propre  $x_i$   $x_i^T D x_i > 0$ . Ces derniers formant une base orthonormée, on écrit pour tout  $x \neq 0$

$$x^T D x = \left( \sum_{i=1}^n \alpha_i x_i \right)^T D \left( \sum_{i=1}^n \alpha_i x_i \right) = \sum_{i=1}^n \alpha_i^2 \lambda_i > 0. \quad \blacksquare$$

**Propriété 1.3** Soit une matrice  $D$  partitionnée de manière suivante

$$D = \begin{pmatrix} D_{11(m \times m)} & D_{12(k \times k)} \\ D_{21(m \times k)} & D_{22(k \times m)} \end{pmatrix}$$

avec  $m + k = n$ .

Si  $D > 0$  ( $D \geq 0$ ), alors les sous-matrices principales  $D_{11}$  et  $D_{22}$  sont aussi définies positives (non négatives). D'une manière générale, toute sous-matrice principale d'une matrice définie positive (non négative) est définie positive (non négative).

**Preuve 3**

Soit  $y \in \mathbb{R}^m$ ,  $y \neq 0$ . Posons  $x = (y, 0) \in \mathbb{R}^n$ . Puisque  $D > 0$ , on aura  $x^T D x > 0$ . Explicitons cette dernière expression :

$$Dx = \begin{pmatrix} D_{11}y \\ D_{21}y \end{pmatrix} \Rightarrow x^T Dx = (y^T, 0) \begin{pmatrix} D_{11}y \\ D_{21}y \end{pmatrix} = y^T D_{11}y$$

On obtient donc

$$y^T D_{11}y > 0, \quad \forall y \neq 0, \quad y \in \mathbb{R}^m.$$

par conséquent,  $D_{11} > 0$ , et  $D_{22} > 0$ . Même démonstration pour  $D \geq 0$ . ■

**Propriété 1.4** Un élément diagonal d'une matrice symétrique  $D$  définie non négative ne peut s'annuler que si les autres éléments de la même ligne et colonne s'annulent aussi.

**Critère de Sylvester pour les formes quadratiques définies et semi-définies**

Pour caractériser une forme quadratique définie ou semi-définie, on se sert du critère de Sylvester. Pour l'établir, considérons une matrice symétrique

$$D = \begin{pmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{pmatrix}. \tag{1.7}$$

Le mineur de la matrice  $D$ , formé des lignes  $i_1, i_2, \dots, i_p$  et des colonnes  $j_1, j_2, \dots, j_p$  sera noté comme suit :

$$D \begin{pmatrix} i_1, i_2, \dots, i_p \\ j_1, j_2, \dots, j_p \end{pmatrix} = \begin{vmatrix} d_{i_1 j_1} & d_{i_1 j_2} & \cdots & d_{i_1 j_p} \\ d_{i_2 j_1} & d_{i_2 j_2} & \cdots & d_{i_2 j_p} \\ \vdots & \vdots & \ddots & \vdots \\ d_{i_p j_1} & d_{i_p j_2} & \cdots & d_{i_p j_p} \end{vmatrix}.$$

Ce mineur est dit principal si  $i_1 = j_1, i_2 = j_2, \dots, i_p = j_p$ , c'est à dire s'il est formé des lignes et des colonnes portant les mêmes numéros. Les mineurs suivants :

$$D_1 = d_{11}, \quad D_2 = \begin{vmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{vmatrix}, \dots, \quad D_n = \begin{vmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{vmatrix},$$

sont appelés mineurs principaux successifs. Nous avons alors le critère de Sylvester qui se formule comme suit :

**Théorème 1.2** (Critère de Sylvester)[35]

1. Pour que la matrice  $D$  soit définie positive ( $D > 0$ ), il est nécessaire et suffisant que les mineurs principaux successifs de  $D$  soit positifs :

$$D_1 > 0, \quad D_2 > 0, \quad \dots, \quad D_n > 0; \quad (1.8)$$

2. Pour que la matrice  $D$  soit semi-définie positive ( $D \geq 0$ ), il est nécessaire et suffisant que les mineurs principaux de  $D$  soient non négatifs :

$$D \begin{pmatrix} i_1, i_2, \dots, i_p \\ i_1, i_2, \dots, i_p \end{pmatrix} \geq 0, \quad 1 \leq i_1 < i_2 < \dots < i_p \leq n, \quad p = 1, 2, \dots, n. \quad (1.9)$$

**Remarque 1.4** La condition

$$D_1 \geq 0, \quad D_2 \geq 0, \dots, \quad D_n \geq 0$$

n'est pas suffisante pour que la matrice  $D$  soit définie non négative.

**Remarque 1.5** Pour qu'une matrice  $D$  soit définie négative ( $D < 0$ ) ou non positive ( $D \leq 0$ ), Les conditions (1.8) et (1.9) se reformulent ainsi :

1.  $D < 0 \Leftrightarrow (-1)^p D_p > 0, \quad p = 1, 2, \dots, n.$
2.  $D \leq 0 \Leftrightarrow (-1)^p D \begin{pmatrix} i_1, i_2, \dots, i_p \\ i_1, i_2, \dots, i_p \end{pmatrix} \geq 0, \quad 1 \leq i_1 < i_2 < \dots < i_p \leq n, \quad p = 1, 2, \dots, n.$

**Remarque 1.6** Le critère de Sylvester n'est valable que pour les matrices symétriques.

**Proposition 4** [35] Soit une forme quadratique  $F(x) = x^T D x$ . Alors  $F$  est convexe si et seulement si sa matrice associée  $D$  est semi-définie positive.

## Propriétés des formes quadratiques définies positives

### Propriété 1.5 [35]

Soit  $A$  une matrice symétrique  $n \times n$  et  $b \in \mathbb{R}^n$ . Considérons la forme quadratique

$$f(x) = \frac{1}{2}x^t Ax - b^t x. \quad (1.10)$$

On considère la famille de surfaces définie par

$$\gamma_c = \{x \in \mathbb{R}^n, f(x) = c\} \quad (1.11)$$

pour  $c \in \mathbb{R}$  et on définit le vecteur  $\hat{x}$  solution de

$$A\hat{x} = b.$$

Alors  $\gamma_c$  est défini comme suit :

- Si  $c < f(\hat{x})$  alors  $\gamma_c = \emptyset$ .
- Si  $c = f(\hat{x})$  alors  $\gamma_c = \{\hat{x}\}$ .
- Si  $c > f(\hat{x})$  alors  $\gamma_c$  est un ellipsoïde centré en  $\hat{x}$ .

On a le théorème suivant :

### Théorème 1 [38]

Soit  $A$  une matrice symétrique définie positive et  $b \in \mathbb{R}^n$ , et soit  $f$  la forme quadratique associée définie comme dans (1.10). Soit  $\hat{x}$  le vecteur unique vérifiant  $A\hat{x} = b$ , alors  $\hat{x}$  réalise le minimum de  $f$ , c'est à dire

$$f(\hat{x}) \leq f(x), \quad \forall x \in \mathbb{R}^n.$$

## 1.1.4 Notions de préférence et optimum de Pareto

Considérons le problème multicritère :

$$\langle X, f(X) \rangle \rightarrow \max, \quad (1.12)$$

où  $X \subset \mathbb{R}^n$  l'ensemble des décisions (actions, scénarios)

$f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}^N$

$x \in X \mapsto f(x)$  où

$$f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_N(x) \end{pmatrix}$$

avec  $f_i : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $i = 1, \dots, N$  sont les fonctions objectifs (fonctions coûts, critères).

**Notations**

Soient  $Y, Z \in \mathbb{R}^N$ , avec  $Y = (y_1, \dots, y_N)$ ,  $Z = (z_1, \dots, z_N)$  on a alors les relations suivantes :

$$Y \geq Z \iff y_i \geq z_i, \forall i = 1, \dots, N;$$

$$Y \geq Z \iff (Y \geq Z \text{ et } Y \neq Z) \iff (y_i \geq z_i, \forall i = 1, \dots, N \text{ et } \exists j \in \{1, \dots, N\} \text{ telle que } y_j > z_j);$$

$$Y > Z \iff y_i > z_i, \forall i = 1, \dots, N.$$

**Définition 1.24** Une décision  $x^* \in X$  est dite maximale de Pareto (décision efficace) du problème (1.12) s'il n'existe pas une autre décision  $x$  qui vérifie la relation :

$$f(x) \geq f(x^*) \tag{1.13}$$

Notons par  $X^p$  l'ensemble des décision maximales de Pareto on a ainsi :

$$x^p \in X^p \iff \nexists x \in X \text{ telle que } \begin{cases} f_i(x) \geq f_i(x^p), \forall i = 1, \dots, N, \\ \exists j \in \mathcal{N} = \{1, \dots, N\} \text{ telle que } f_j(x) > f_j(x^p), \end{cases}$$

$$x^p \in X^p \iff \forall x \in X : f(x) \not\geq f(x^p) \iff \forall x \in X \text{ telle que } f(x) \geq f(x^p) \text{ on a } f(x) = f(x^p)$$

**Remarque :**

$$f(x) \not\geq f(x^p) \iff \exists i \in \mathcal{N} \text{ telle que } f_i(x) < f_i(x^p) \text{ ou } f(x) = f(x^p)$$

**Définition 1.25** Un vecteur  $x \in \mathbb{R}^n$  est dit efficace s'il est un optimum de Pareto

## 1.2 Formulation générale des problèmes d'optimisation non linéaire

La forme générale d'un problème d'optimisation est la suivante :

$$\mathcal{P}_C \begin{cases} \min f(x), \\ x \in C. \end{cases} \tag{1.14}$$

où  $C = \{x \in \mathbb{R}^n : g(x) \leq 0, h(x) = 0\}$ , les fonctions  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$  et  $h : \mathbb{R}^n \rightarrow \mathbb{R}^q$  sont typiquement non-linéaires. L'équation  $g(x) \leq 0$  désigne ce que

nous appellerons des contraintes d'inégalité et l'équation  $h(x) = 0$  des contraintes d'égalité.

Nous noterons ces types de problèmes ainsi :

- ◇ (PC) problème général, avec contraintes d'inégalité et d'égalité,
- ◇ (PCE) problème avec contraintes d'égalité,
- ◇ (PCI) problème avec contraintes d'inégalité,
- ◇ (P) problème sans contraintes.

Il va de soi que la plupart des problèmes réels ou industriels ne sont pas initialement mis sous une des formes proposées. C'est pourquoi un des premiers travaux consiste en général à mettre le problème initial sous une forme standard. Par exemple, un problème donné sous la forme

$$\max_{x \in \mathbb{R}^n} g(x), \quad (1.15)$$

se mettra sous la forme standard (P) en posant  $f(x) = -g(x)$ .

## 1.2.1 Rappels de calcul différentiel

### Définition de la différentiabilité

Dans  $\mathbb{R}^n$  on note  $x$  le vecteur colonne

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

et la notation  $\|\cdot\|$  désignera, sauf indication du contraire, la norme euclidienne

$$\|x\| = \left( \sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}}.$$

Avant de donner la définition de la différentiabilité, il est important de rappeler celle de la continuité

**Définition 1.26** Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , on dit que  $f$  est continue au point  $a \in \mathbb{R}^n$  si pour tout réel  $\varepsilon > 0$  il existe  $\eta > 0$  tel que

$$\|x - a\| < \eta \Rightarrow \|f(x) - f(a)\| < \varepsilon$$

Voici maintenant la définition de la différentiabilité :

**Définition 1.27** Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  représentée dans la base canonique de  $\mathbb{R}^m$  par le vecteur

$$f(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_m(x) \end{pmatrix},$$

continue en  $a \in \mathbb{R}^n$ . On dit que  $f$  est différentiable en  $a$  s'il existe une application linéaire, notée  $f'(a)$ , telle que pour tout  $h \in \mathbb{R}^n$  on ait

$$f(a+h) = f(a) + f'(a)h + \|h\| \varepsilon(h), \quad (1.16)$$

où  $\varepsilon(\cdot)$  est une fonction continue en 0 vérifiant  $\lim_{h \rightarrow 0} \varepsilon(h) = 0$ .

On appelle  $f'(a)$  dérivée de  $f$  au point  $a$ . La notation  $f'(a)h$  doit être prise au sens “ $f'(a)$  appliquée à  $h$ ”.

### Calcul de la dérivée première

On peut d'ores et déjà donner un résultat "pratique" permettant de calculer directement la dérivée à partir du développement (1.16).

**Proposition 5** [38] Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  différentiable en  $a$ , alors

$$\lim_{t \rightarrow 0} \frac{f(a+th) - f(a)}{t} = f'(a)h.$$

La quantité  $f'(a)h$  est appelée communément dérivée directionnelle de  $f$  au point  $a$  dans la direction  $h$ . La proposition suivante fait le lien entre la matrice de  $f'(a)$  et les dérivées partielles de  $f$  au point  $a$  :

**Proposition 6** [38] Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  différentiable en  $a$ , alors on peut représenter  $f'(a)$  par sa matrice dans les bases de  $\mathbb{R}^n$  et  $\mathbb{R}^m$  et on a

$$(f'(a))_{ij} = \frac{\partial f_i(a)}{\partial x_j}$$

On appelle souvent  $f'(a)$  la matrice jacobienne de  $f$  au point  $a$ . Lorsque  $m = 1$  on adopte une notation et un nom particuliers : le gradient est le vecteur noté  $\nabla f(a)$  et défini par

$$f'(a) = \nabla f(a)^t,$$

et on a

$$f(a+h) = f(a) + \nabla f(a)^t h + \|h\| \varepsilon(h).$$

## Dérivée seconde

On se place maintenant dans le cas  $m = 1$ , soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ .

**Définition 1.28** L'application  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  est dite deux fois différentiable s'il existe une matrice symétrique  $\nabla^2 f(a)$  telle que

$$f(a+h) = f(a) + \nabla f(a)^t h + \frac{1}{2} h^t \nabla^2 f(a) h + \|h\|^2 \varepsilon(h).$$

On appelle  $\nabla^2 f(a)$  matrice hessienne de  $f$  au point  $a$ .

Comme l'énonce le théorème suivant, cette matrice s'obtient à partir des dérivées secondes de  $f$  :

**Théorème 2 [38]** Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction deux fois différentiable en un point  $a$ . Si on note  $g(x) = \nabla f(x)$  alors la matrice hessienne est définie par

$$\nabla^2 f(a) = g'(a),$$

soit

$$(\nabla^2 f(a))_{ij} = \frac{\partial^2 f(a)}{\partial x_i \partial x_j}.$$

## 1.2.2 Notions sur la convexité

En plus des définitions données dans la section 1.1.2 sur la convexité, on énonce ci-après les résultats concernant cette dernière dans le cas d'une fonction différentiable.

### Caractérisation de la convexité en terme du hessien

Dans le cas où  $f : K \subset \mathbb{R} \rightarrow \mathbb{R}$  on a le résultat suivant :

**Propriété 1.6 [38]** Si  $f : K \subset \mathbb{R} \rightarrow \mathbb{R}$  est 2 fois continûment dérivable sur  $K$  convexe alors  $f$  est convexe si et seulement si  $f''(x) \geq 0$ ,  $\forall x \in K$  et strictement convexe si et seulement si  $f''(x) > 0$ ,  $\forall x \in K$ .

Ce résultat se généralise pour  $n > 1$  : le résultat suivant fait le lien entre le hessien et la propriété de convexité :

**Théorème 3 [38]** Soit  $f : K \subset \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction deux fois différentiable, alors  $f$  est convexe si et seulement si  $\nabla^2 f(x) \geq 0$ ,  $\forall x \in K$ , et strictement convexe si et seulement si  $\nabla^2 f(x) > 0$ ,  $\forall x \in K$ .

**Propriété 1.7 [38]** Soit  $f$  une forme quadratique définie comme dans (1.10), alors  $f$  est convexe si et seulement si  $A \geq 0$ , et strictement convexe si et seulement si  $A > 0$ .

Cela provient du fait que  $\nabla^2 f(x) = A$ .

## Caractérisation de la convexité en terme du gradient

Dans le cas où la fonction  $f$  n'est supposée qu'une fois différentiable, on a le résultat suivant :

**Théorème 4** [38] Soit  $f : K \subset \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction une fois différentiable, alors  $f$  est convexe si et seulement si

$$f(y) \geq f(x) + \nabla f(x)^t(y-x), \quad \forall (x,y) \in K^2$$

. La fonction  $f$  est strictement convexe si et seulement si

$$f(y) > f(x) + \nabla f(x)^t(y-x), \quad \forall (x,y) \in K^2, \quad x \neq y$$

### 1.2.3 Résultats d'existence et d'unicité

#### Théorèmes généraux d'existence

Considérons notre problème d'optimisation (1.14), que l'on écrira pour l'occasion un peu différemment, en mettant les contraintes sous la forme  $x \in K \subset \mathbb{R}^n$  :

$$\min_{x \in K} f(x). \quad (1.17)$$

Nous allons donner deux résultats très généraux d'existence d'une solution au problème (1.17). Auparavant nous avons besoin de la définition d'un ensemble compact :

**Définition 1.29** Un ensemble  $K \subset \mathbb{R}^n$  est dit compact si, de toute suite  $\{x_k\}$ , où  $x_k \in K$ ,  $\forall k$ , on peut extraire une sous-suite convergente.

On a le théorème suivant :

**Théorème 5** [38] Un ensemble  $K \subset \mathbb{R}^n$  est compact si et seulement si il est fermé et borné.

Dans  $\mathbb{R}$ , les intervalles fermés du type  $[a,b]$  (ou des réunions de tels intervalles) sont compacts.

La notion de fermeture signifie que toute suite convergente  $\{x_k\}$ , où  $x_k \in K$ ,  $\forall k$ , doit converger vers une limite  $x \in K$ .

Voici maintenant deux résultats d'existence :

**Théorème 6** [38] Si  $f : K \subset \mathbb{R}^n \rightarrow \mathbb{R}$  est continue et si de plus  $K$  est un ensemble compact, alors le problème (1.17) admet une solution optimale  $x^* \in K$ , qui vérifie donc

$$f(x^*) \leq f(x), \quad \forall x \in K.$$

Le second résultat est moins général car il considère le cas particulier  $K = \mathbb{R}^n$  :

**Théorème 7** [38] Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction continue sur  $\mathbb{R}^n$ . Si

$$\lim_{\|x\| \rightarrow \infty} f(x) = \infty,$$

alors (1.17) admet une solution optimale  $x^*$ .

### Unicité

L'unicité résulte en général de propriétés de convexité (de  $f$  et de  $K$ ).

**Théorème 8** [38] Soit  $f : K \subset \mathbb{R}^n \rightarrow \mathbb{R}$  strictement convexe sur  $K$  convexe. Le minimum de  $f$  sur  $K$ , s'il existe, est unique.

## 1.2.4 Conditions nécessaires d'optimalité en absence de contraintes

### Conditions nécessaires

On va maintenant regarder de plus près le cas où  $K = \mathbb{R}^n$ , c'est à dire le problème sans contraintes (P). Dans le cas où  $f$  est différentiable, on a le résultat suivant :

**Théorème 9** [38] Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  différentiable et  $x^*$  vérifiant  $f(x^*) \leq f(x)$ ,  $\forall x \in \mathbb{R}^n$ , alors on a nécessairement

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

### Conditions nécessaires et suffisantes

La condition de gradient nul devient suffisante dans le cas où  $f$  est convexe :

**Théorème 10** [38] Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  convexe et différentiable. Si  $x^*$  vérifie  $\nabla f(x^*) = 0$ , alors on a  $f(x^*) \leq f(x)$ ,  $\forall x \in \mathbb{R}^n$ .

Lorsque la fonction n'est pas convexe, on ne peut donner qu'une condition suffisante d'optimalité locale. On désignera par minimum local (que l'on oppose au minimum global) un vecteur vérifiant les conditions suivantes :

**Définition 1.30** On appellera  $x^*$  minimum local de  $f$ , s'il existe  $\delta > 0$  tel que

$$f(x^*) \leq f(x), \quad \forall x, \|x - x^*\| \leq \delta.$$

Dans le cas où  $f$  est deux fois différentiable on peut alors donner le résultat suivant :

**Théorème 11** [38] Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  deux fois différentiable. Si

$$\begin{cases} \nabla f(x^*) = 0, \\ \nabla^2 f(x^*) > 0, \end{cases}$$

alors  $x^*$  est un minimum local de  $f$ .

## 1.2.5 Conditions d'optimalité en optimisation avec contraintes

### Problème avec contraintes d'égalité

On va tout d'abord s'intéresser au problème suivant, dit problème d'optimisation avec contraintes d'égalité seulement :

$$(PCE) \begin{cases} \min_{x \in \mathbb{R}^n} f(x), \\ \text{sous contraintes,} \\ h(x) = 0, \end{cases} \quad (1.18)$$

### Contraintes d'égalité linéaires

Un problème d'optimisation avec contraintes d'égalité linéaires prend la forme

$$(PCEL) \begin{cases} \min_{x \in \mathbb{R}^n} f(x), \\ \text{sous contraintes,} \\ Ax - b = 0, \end{cases} \quad (1.19)$$

où  $A$  est une matrice  $p \times n$  avec  $p < n$  et  $b \in \mathbb{R}^p$ . On notera  $S = \{x \in \mathbb{R}^n, Ax - b = 0\}$ .

Nous allons maintenant définir le concept de direction admissible dans  $S$ .

**Définition 1.31** *On dit que  $d \in \mathbb{R}^n$  est une direction admissible en  $x \in S$  s'il existe  $\alpha > 0$  tel que*

$$x + td \in S, \quad \forall t \in [-\alpha, \alpha]$$

Dans notre cas, on a  $A(x + td) - b = tAd$  puisque  $x \in S$ , et donc les directions admissibles  $d$  sont caractérisées par

$$Ad = 0. \quad (1.20)$$

On peut donc énoncer les conditions nécessaires d'optimalité pour le problème (1.19)

**Théorème 12** [38] *Soit  $x^* \in S$  solution du problème (1.19), vérifiant donc  $f(x^*) \leq f(x), \forall x \in S$ . Alors il existe nécessairement un vecteur  $\lambda \in \mathbb{R}^p$  vérifiant*

$$\nabla f(x^*) + A^t \lambda = 0. \quad (1.21)$$

*Si de plus  $A$  est de rang  $p$  alors  $\lambda$  est unique.*

## Méthode de projection

Un point stationnaire est un point vérifie la condition  $\nabla f(x) = 0$ , c'est à dire que tout point stationnaire est un point fixe de la fonction  $Proj/S(x - \nabla f(x))$ .

Si on a sous la main un point  $x \neq Proj/S(x - \nabla f(x))$  alors on peut construire une direction de descente  $d = Proj/S(x - \nabla f(x)) - x$ .

L'inconvénient de cette méthode réside dans la difficulté du calcul de la projection. Néanmoins, dans le cas de contraintes de bornes, sur les variables, alors il est possible de calculer la projection efficacement, et l'algorithme présente un certain intérêt.

## Contraintes d'égalité non-linéaires

Nous étudions maintenant le problème

$$(PCE) \begin{cases} \min_{x \in \mathbb{R}^n} f(x), \\ \text{sous contraintes,} \\ h(x) = 0, \end{cases} \quad (1.22)$$

où  $h : \mathbb{R}^n \rightarrow \mathbb{R}^p$  est différentiable.

On note comme précédemment

$$S = \{x \in \mathbb{R}^n, \quad h(x) = 0\}.$$

Le concept de direction admissible dans  $S$  ne peut pas se définir comme pour les contraintes linéaires, car pour  $x^* \in S$  il peut ne pas exister  $\alpha > 0$  et  $d \in \mathbb{R}^n$  tels que  $x^* + td \in S$ .

On doit donc définir le concept de courbe admissible. Considérons une courbe  $x(t)$  définie par

$$\begin{cases} x(t) \in S, & \forall t \in [-\alpha, \alpha], \quad \alpha > 0; \\ x(0) = x^* \end{cases}$$

Puisque  $x(t) \in S$  on a  $h_i(x(t)) = 0$  pour  $1 \leq i \leq p$  et on peut écrire que

$$\frac{\partial h_i(x(t))}{\partial t} = [\nabla h_i(x(t))]^t x'(t) = 0, \quad 1 \leq i \leq p.$$

Si on note  $y = x'(0)$  le vecteur tangent à la courbe  $x(t)$  en  $t = 0$ , on a donc

$$[\nabla h_i(x^*)]^t y = 0, \quad 1 \leq i \leq p. \quad (1.23)$$

Cela conduit à la définition suivante

**Définition 1.32** On dit que  $y \in \mathbb{R}^n$  est une direction admissible en  $x^* \in S$  s'il existe  $\alpha > 0$  et une courbe  $x(t)$  vérifiant

$$\begin{cases} x(t) \in S, & \forall t \in [-\alpha, \alpha], \quad \alpha > 0; \\ x(0) = x^* \\ x'(0) = y. \end{cases}$$

On notera alors  $y \in T(x^*)$ .

L'ensemble  $T(x^*)$  définit le plan tangent à  $S$  en  $x^*$ . L'analyse faite précédemment montre que l'on a l'implication

$$y \in T(x^*) \Rightarrow [\nabla h_i(x^*)]^t y = 0, \quad 1 \leq i \leq p$$

qui sera insuffisante pour montrer la condition nécessaire d'optimalité. Nous allons donc maintenant nous attacher à montrer sous quelles conditions la relation (1.23) est une condition suffisante d'appartenance à  $T(x^*)$ .

**Définition 1.33** On dit que  $x^*$  est un point régulier pour la contrainte  $h(x) = 0$  si

–  $h(x^*) = 0$

– Les vecteurs  $\nabla h_i(x^*)$  sont linéairement indépendants.

Si on note  $\nabla h(x^*)$  la matrice  $n \times p$

$$\nabla h(x^*) = [\nabla h_1(x^*), \dots, \nabla h_p(x^*)],$$

la condition d'indépendance linéaire des  $\nabla h_i(x^*)$  peut s'écrire

$$\text{rang} \nabla h(x^*) = p,$$

et on a donc  $[\nabla h_i(x^*)]^t x'(0) = 0$  pour toute courbe admissible  $x(t)$ .

On a la proposition suivante :

**Proposition 7** [38] Si  $x^*$  est un point régulier pour la contrainte  $h(x) = 0$ , alors

$$[\nabla h_i(x^*)]^t y = 0 \Rightarrow y \in T(x^*)$$

## Le théorème de Lagrange

**Théorème 13** [38] Soit  $x^* \in S = \{x \in \mathbb{R}^n, \quad h(x) = 0\}$  un point régulier solution du problème (1.22), vérifiant donc

$$f(x^*) \leq f(x), \quad \forall x \in S.$$

Alors il existe nécessairement un vecteur  $\lambda \in \mathbb{R}^p$  unique vérifiant

$$\nabla f(x^*) + \nabla h(x^*) \lambda = 0,$$

soit encore

$$\nabla f(x^*) + \sum_{i=1}^p \lambda_i \nabla h_i(x^*) = 0.$$

Les composantes du vecteur  $\lambda$  sont appelées multiplicateurs de Lagrange.

## 1.2.6 Les conditions de Karush-Kuhn-Tucker

### Problème avec contraintes d'inégalité

On s'intéresse maintenant au problème suivant, dit problème d'optimisation avec contraintes d'inégalité seulement :

$$(PCI) \begin{cases} \min_{x \in \mathbb{R}^n} f(x), \\ \text{sous contraintes,} \\ g(x) \leq 0, \end{cases} \quad (1.24)$$

où  $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$  est différentiable (il n'y a ici aucune condition sur  $m$ ). On notera  $K$  l'ensemble des points admissibles, c'est à dire

$$K = \{x \in \mathbb{R}^n, g(x) \leq 0\}$$

. Au point solution de (PCI) il va de soi que les contraintes actives vérifieront  $g_i(x^*) = 0$ . Cependant, puisque l'on ne sait pas à priori quelles sont ces contraintes, le passage de (PCI) à un problème du type (PCE) n'est pas direct.

**Définition 1.34** On appelle contraintes saturées en  $x^*$  l'ensemble des indices  $i$  tel que  $g_i(x^*) = 0$ , et on note

$$I(x^*) = \{i : g_i(x^*) = 0\}$$

. On note alors  $S(x^*)$  l'ensemble

$$S(x^*) = \{x \in \mathbb{R}^n, g_i(x) = 0, i \in I(x^*)\}$$

.

Le concept de direction admissible se définit comme suit :

**Définition 1.35** On dit que  $y \in \mathbb{R}^n$  est une direction admissible en  $x^* \in K$  s'il existe  $\alpha > 0$  et une courbe  $x(t)$  vérifiant

$$\begin{cases} x(t) \in K, \quad \forall t \in [-\alpha, \alpha]; \\ x(0) = x^* \\ x'(0) = y. \end{cases}$$

On notera alors  $y \in C(x^*)$ .

**Lemme 1** [38] Soit  $y \in \mathbb{R}^n$  une direction admissible en  $x^* \in K$ , alors on a nécessairement

$$[\nabla g_i(x^*)]^t y \leq 0, \quad i \in I(x^*).$$

Comme dans le cas des contraintes d'égalité, on doit définir la notion de point régulier, qui est nécessaire pour que la condition précédente soit suffisante :

**Définition 1.36** *On dit que  $x^*$  est un point régulier pour la contrainte  $g(x) \leq 0$  si*

- $g(x^*) \leq 0$
- Les vecteurs  $\{\nabla g_i(x^*)\}_{i \in I(x^*)}$  sont linéairement indépendants.

Sous l'hypothèse de régularité de  $x^*$  on aura, comme dans le cas des contraintes d'égalité

$$[\nabla g_i(x^*)]^t y \leq 0, \quad i \in I(x^*) \Rightarrow y \in C(x^*).$$

La proposition suivante permet d'effectuer le premier pas vers les conditions de Kuhn et Tucker.

**Proposition 8** [38] *Soit  $x^*$  la solution du problème (PCI). Alors il existe  $\eta > 0$  tel que*

$$\forall x \in B(x^*, \varepsilon), \quad g_i(x) < 0, \quad i \in I(x^*),$$

où on a noté  $B(x^*, \varepsilon)$  la boule de centre  $x^*$  et de rayon  $\varepsilon$ . Alors  $x^*$  est la solution du problème

$$(PCE) \begin{cases} \min_{x \in B(x^*, \varepsilon)} f(x), \\ \text{sous contraintes,} \\ g(x) = 0, \quad i \in I(x^*). \end{cases}$$

Ce résultat est uniquement dû à la continuité de  $g$ , et montre que l'on est localement ramené à un problème avec contraintes d'égalité. On peut donc maintenant énoncer le résultat principal :

**Théorème 14** [38] *Soit  $x^* \in K$  un point régulier solution du problème (PCI). Alors il existe un unique vecteur  $\mu \in \mathbb{R}^p$  tel que*

$$\nabla f(x^*) + \sum_{i=1}^m \mu_i \nabla g_i(x^*) = 0 \tag{1.25}$$

$$\mu_i \geq 0, \quad i = 1, \dots, p \tag{1.26}$$

$$\mu_i g_i(x^*) = 0, \quad i = 1, \dots, p \tag{1.27}$$

## 1.2.7 Conditions suffisantes d'optimalité

### Définition du lagrangien

Considérons le problème (PCE) avec contraintes d'égalité.

**Définition 1.37** *On appelle lagrangien associé au problème (PCE) la fonction  $L : \mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}$  définie par*

$$L(x, \lambda) = f(x) + \sum_{i=1}^p \lambda_i h_i(x) \tag{1.28}$$

Les conditions de Lagrange peuvent se reformuler à l'aide du lagrangien : soit  $x^*$  solution de (PCE). Alors il existe  $\lambda^*$  tel que

$$\nabla_x L(x^*, \lambda^*) = 0,$$

où on a noté  $\nabla_x$  le gradient partiel par rapport à la variable  $x$ . Dans la suite nous ferons l'hypothèse que  $h$  et  $f$  sont deux fois continûment différentiables.

### Condition nécessaire du second ordre pour les problèmes (PCE)

**Théorème 15** [38] Soit  $x^*$  un point régulier solution de (PCE). Alors il existe  $\lambda^*$  tel que

$$\nabla_x L(x^*, \lambda^*) = 0,$$

et de plus pour tout  $y \in T(x^*)$ ,  $y \neq 0$ , on a

$$y^t \nabla_{xx}^2 L(x^*, \lambda^*) y \geq 0.$$

### Condition suffisante du second ordre pour les problèmes (PCE)

**Théorème 16** [38] Soit  $x^* \in \mathbb{R}^n$  et  $\lambda^* \in \mathbb{R}^p$  vérifiant les conditions

$$\begin{aligned} h(x^*) &= 0, \\ \nabla f(x^*) + \sum_{i=1}^p \lambda_i^* \nabla h_i(x^*) &= 0, \\ y^t \nabla_{xx}^2 L(x^*, \lambda^*) y &> 0, \quad \forall y \in T(x^*), \quad y \neq 0. \end{aligned}$$

alors  $x^*$  est un minimum local du problème (PCE).

### Condition nécessaire du second ordre pour les problèmes (PCI)

**Théorème 17** [38] Soit  $x^* \in \mathbb{R}^n$  et  $\mu \in \mathbb{R}^m$  vérifiant les conditions

$$g(x^*) \leq 0, \tag{1.29}$$

$$\nabla f(x^*) + \sum_{i=1}^m \mu_i \nabla g_i(x^*) = 0, \tag{1.30}$$

$$\mu_i \geq 0, \quad i = 1, \dots, m, \tag{1.31}$$

$$\mu_i g_i(x^*) = 0, \quad i = 1, \dots, m, \tag{1.32}$$

$$y^t \nabla_{xx}^2 L(x^*, \lambda^*) y > 0, \quad \forall y \in T^+(x^*), \quad y \neq 0. \tag{1.33}$$

où on a noté  $T^+(x^*)$  le plan tangent en  $x^*$  à la surface

$$S^+ = \{x \in \mathbb{R}^n, \quad g_i(x^*) = 0, \quad i \in I(x^*) \quad \text{et} \quad \mu_i > 0\}.$$

Alors  $x^*$  est un minimum local du problème (PCI).

# 2

## Méthodes Numériques.

### 2.1 Méthode des pénalités

Considérons le problème de programmation non linéaire

$$(PAC) \quad \min\{f(z) : z \in C \subset \mathbb{R}^n\},$$

où  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  est une fonction continûment différentiable et  $C$  un sous-ensemble fermé de  $\mathbb{R}^n$ .

L'idée principale des méthodes des pénalités est la résolution du problème  $(PAC)$ , en construisant une suite de points  $z_i \in \mathbb{R}^n$  qui sont solutions optimales pour une série de problèmes de minimisation sans contraintes de la forme :

$$(PSC)_i \quad \min\{f(z) + P_i(z) : z \in \mathbb{R}^n\}, \quad i = 0, 1, 2, \dots$$

Les problèmes  $(PSC)_i$  sont construits de façon que la suite  $\{z_i\}$  des solutions optimales des problèmes  $(PSC)_i$  converge vers la solution optimale  $\hat{z}$  du problème  $(PAC)$ .

Il y a deux différentes approches majeures pour construire les problèmes  $(PSC)_i$ , ou plus précisément les fonctions de pénalités  $P_i(\cdot)$ .

La première approche, connue sous le nom de "méthode des pénalités extérieures", a été proposée par Courant en 1943 [43], au moment où la plupart des algorithmes de programmation non linéaire connus à ce jour n'ont pas été encore découverts. Dans la méthode des pénalités extérieures, les fonctions  $P_i(\cdot)$  sont choisies de sorte que leurs valeurs augmentent si le point sélectionné n'appartient pas à  $C$  et

$P_i(z) = 0$  pour tout  $z \in C$  ou  $P_i(z) \rightarrow 0$  quand  $i \rightarrow \infty$  pour tout  $z \in C$ .

Dans la deuxième approche, appelée "méthodes des pénalités intérieures", les fonctions de pénalités  $P_i(\cdot)$  sont choisies de sorte que les solutions optimales  $z_i$  des problèmes  $(PSC)_i$  appartiennent toutes à l'intérieur  $int(C)$  de l'ensemble  $C$  et  $f(z) + P_i(z) \rightarrow f(z)$  quand  $i \rightarrow \infty$  pour tout  $z$  appartenant à  $int(C)$ .  $f(z) + P_i(z) \rightarrow \infty$  quand le point  $z$  s'approche, de l'intérieur vers la frontière de  $C$ . C'est pour cette dernière propriété de la fonction de pénalité que les méthodes des pénalités intérieures sont aussi connues sous le nom de "méthodes des barrières". Par conséquent, le point initial  $z^0$  utilisé pour la résolution des problèmes  $(PSC)_i$  doit être choisi à l'intérieur de l'ensemble  $C$ .

Il y a plusieurs situations dans lesquelles les méthodes des pénalités intérieures et extérieures peuvent être combinées.

### 2.1.1 Méthodes de pénalité extérieure

La raison pour laquelle on parle de "méthodes" au pluriel au lieu de "méthode", est que pour chaque choix d'une famille de fonctions de pénalité  $P_i(\cdot)$ , on aura un algorithme différent. Ainsi il existe une large classe de méthodes de pénalité extérieure.

**Définition 2.1** Soit  $C$  un sous ensemble fermé de  $\mathbb{R}^n$ . Une suite de fonctions continues  $P'_i : \mathbb{R}^n \rightarrow \mathbb{R}, i = 0, 1, 2, \dots$  est dite une suite de fonctions de pénalité extérieure pour l'ensemble  $C$ , si :

$$P'_i(z) = 0, \quad \forall z \in C, \quad i = 0, 1, 2, \dots \quad (2.1)$$

$$P'_i(z) > 0, \quad \forall z \notin C, \quad i = 0, 1, 2, \dots \quad (2.2)$$

$$P'_{i+1}(z) > P'_i(z), \quad \forall z \notin C, \quad i = 0, 1, 2, \dots \quad (2.3)$$

$$P'_i(z) \rightarrow \infty \text{ quand } i \rightarrow \infty, \quad \forall z \notin C. \quad (2.4)$$

■

Considérons le problème  $(PAC)$  et soit  $P'_i(z), i = 0, 1, 2, \dots$  une suite de fonctions de pénalité extérieure pour l'ensemble  $C$ . Nous introduisons une suite de problèmes de minimisation sans contraintes  $(PSC)_i^e$ <sup>1</sup>, définis comme suit :

$$(PSC)_i^e \min\{f(z) + P'_i(z) : z \in \mathbb{R}^n\} \quad i = 0, 1, 2, \dots \quad (2.5)$$

<sup>1</sup>L'indice "e" dans  $(PSC)_i^e$  signifie extérieur

Pour assurer l'existence d'une solution pour les deux problèmes (PAC) et  $(PSC)_i^e$ , il suffit de faire l'hypothèse suivante :

**Hypothèse1** Nous allons supposer qu'il existe un point  $z' \in C$  tel que l'ensemble  $Z' = \{z : f(z) \leq f(z')\}$  soit compact.

**Proposition 9** Supposons l'hypothèse 2.1.1 vérifiée. Alors

1. Le problème (PAC) admet une solution optimale. De plus, toute solution optimale  $z^*$  du problème (PAC) appartient à l'ensemble  $Z'$ .
2. L'ensemble  $\{z : f(z) + P'_i(z) \leq f(z')\}$  est compact et inclus dans  $Z'$ .
3.  $\forall i = 0, 1, 2, \dots$ , le problème  $(PSC)_i^e$  admet une solution optimale appartenant à l'ensemble  $Z'$ .

**Lemme 2.1** Considérons les problèmes (PAC) et  $(PSC)_i^e$ , soient  $\bar{m}$  et  $\bar{m}_i$ ,  $i = 0, 1, 2, \dots$  définis comme suit

$$\bar{m} = \min\{f(z) : z \in C\}, \quad (2.6)$$

$$\bar{m}_i = \min\{f(z) + P'_i(z) : z \in \mathbb{R}^n\}. \quad (2.7)$$

Alors, on a :

$$\bar{m}_0 \leq \bar{m}_1 \leq \bar{m}_2 \leq \dots \leq \bar{m}.$$

**Preuve**

Pour  $i = 0, 1, 2, \dots$  Soit  $z_i^* \in \mathbb{R}^n$  une solution optimale du problème  $(PSC)_i^e$ , c'est-à-dire

$$\bar{m}_i = f(z_i^*) + P'_i(z_i^*) \quad i = 0, 1, 2, \dots$$

Remarquons que ces solutions optimales existent d'après la proposition 9.

Alors, d'après (2.1) et (2.3), il en résulte pour  $i = 0, 1, 2, \dots$

$$\bar{m}_i \leq f(z_{i+1}^*) + P'_i(z_{i+1}^*) \leq f(z_{i+1}^*) + P'_{i+1}(z_{i+1}^*) = \bar{m}_{i+1}.$$

Par suite, de (2.1) et (2.7), on déduit

$$\bar{m}_i \leq f(z) + P'_i(z) = f(z), \quad \forall z \in C,$$

et d'où

$$\bar{m}_i \leq \min\{f(z) : z \in C\} = \bar{m}. \quad \square$$

**Lemme 2.2** Soit  $C$  un sous ensemble fermé de  $\mathbb{R}^n$  et soit  $P'_i(\cdot)$ ,  $i = 0, 1, 2, \dots$  une suite de fonctions de pénalité extérieure de l'ensemble  $C$ . On suppose que les  $z_i$ ,  $i = 0, 1, 2, \dots$  forment une suite dans  $\mathbb{R}^n$  convergente vers un point  $\hat{z}$ . Si  $z_i \notin C$ ,  $\forall i = 0, 1, 2, \dots$  et il existe une borne  $M < \infty$  telle que  $P'_i(z_i) \leq M$  pour tout  $i = 0, 1, 2, \dots$ , alors  $\hat{z} \in C$ .

### Preuve

Supposons que  $\hat{z} \notin C$ . Par hypothèse, il existe un  $M < \infty$  telle que  $P'_i(z_i) \in ]0, M]$  pour  $i = 0, 1, 2, \dots$

Comme  $\hat{z} \notin C$ , il résulte d'après (2.4), que  $P'_i(\hat{z}) \rightarrow \infty$  quand  $i \rightarrow \infty$ .

Donc, il existe un entier  $N_1$  tel que  $P'_{N_1}(\hat{z}) > 2M$ . Et comme  $P'_{N_1}(\cdot)$  est continue, il existe une boule ouverte  $B$  de centre  $\hat{z}$  telle que

$$P'_{N_1}(z) > \frac{3M}{2}, \quad \forall z \in B. \quad (2.8)$$

Notons que  $B \cap C$  est un ensemble vide, du fait que  $P'_{N_1}(z) = 0, \forall z \in C$ . Comme  $z_i \rightarrow \hat{z}$  quand  $i \rightarrow \infty$ , il existe un entier  $N_2$  tel que  $z_i \in B$  pour tout  $i \geq N_2$ .

Soit  $N = \max\{N_1, N_2\}$ ; pour tout  $i \geq N$ ,  $z_i \in B$  et de plus, d'après (2.3) et (2.8),

$$P'_i(z_i) \geq P'_N(z_i) \geq \frac{3M}{2},$$

ce qui contredit notre hypothèse que  $P'_i(z_i) \leq M$ . Par conséquent, on a  $\hat{z} \in C$ . □

**Théorème 18** [43] Considérons la suite de problèmes  $(PSC)_i^e$  définis en (2.5) et supposons que l'hypothèse 2.1.1 est satisfaite. Si les  $z_i^*$  sont des solutions optimales pour  $(PSC)_i^e$ ,  $i = 0, 1, 2, \dots$ , alors tout point limite de la suite  $\{z_i^*\}_{i \in \mathbb{N}}$  est une solution optimale pour le problème (PAC).

### Preuve

D'après la proposition 9, les problèmes  $(PSC)_i^e$  admettent des solutions optimales  $z_i^*$  qui appartiennent toutes à l'ensemble compact  $Z'$ ,  $\forall i = 0, 1, 2, \dots$ . Donc, de la suite  $\{z_i^*\}_{i \in \mathbb{N}}$ , on peut extraire une sous-suite convergente vers un point  $\hat{z}$ .

Sans perte de généralité, on peut considérer que la suite  $\{z_i^*\}_{i \in \mathbb{N}}$  converge vers le point  $\hat{z}$ . A présent, il y a deux possibilités :

- a). il existe un entier  $N \geq 0$  tel que  $z_i^* \in C, \forall i \geq N$ .
- b). il n'existe pas un tel entier  $N$ , dans ce cas la suite  $\{z_i^*\}_{i \in \mathbb{N}}$  contient une sous-suite infinie de points  $z_i^*, i \in K \subset \{0, 1, 2, \dots\}$ , tels que  $z_i^* \notin C$ .

Supposons que c'est l'alternative a) qui est vraie. Comme l'ensemble  $C$  est fermé, alors  $\hat{z} \in C$ .

Comme pour tout  $i \geq N$ ,  $z_i^* \in C$ , alors, d'après (2.1),  $P'_i(z_i^*) = 0$ . De plus, comme  $\forall i \geq N$ ,  $z_i^*$  sont les solutions optimales respectives des problèmes  $(PSC)_i^e$  et comme

$$f(z) + P'_i(z) = f(z), \quad \forall z \in C,$$

alors d'après le lemme 2.1, on devrait avoir

$$\bar{m}_i = \bar{m} = f(z_i^*), \quad \forall i \geq N,$$

c'est-à-dire,  $\forall i \geq N$ ,  $z_i^*$  est aussi solution optimale de (PAC).  
Comme  $f(\cdot)$  est continue,  $z_i^* \rightarrow \hat{z}$  et  $f(z_i^*) = \bar{m}$ ,  $i \geq N$ , alors

$$\lim_{i \rightarrow \infty} f(z_i^*) = f(\hat{z}) = \bar{m},$$

d'où  $\hat{z}$  est une solution optimale du problème (PAC).

Maintenant, supposons que c'est l'alternative b) qui a lieu. Comme  $\lim_{i \rightarrow \infty} z_i^* = \hat{z}$  et la fonction  $f(\cdot)$  est continue, alors

$$\lim_{i \rightarrow \infty} f(z_i^*) = f(\hat{z}).$$

Donc, il existe un nombre  $M' \in ]0, \infty[$  tel que

$$|f(z_i^*)| \leq M' \quad \forall i \in K.$$

Par conséquent, comme  $\bar{m}_i = f(z_i^*) + P'_i(z_i^*) \leq \bar{m}$ ,  $\forall i \in K$ , alors, d'après le lemme 2.1, on a

$$P'_i(z_i^*) \leq \bar{m} + M', \quad \forall i \in K,$$

ce qui signifie que la suite  $\{P'_i(z_i^*)\}_{i \in K}$  est bornée par le nombre  $M = \bar{m} + M'$ .

Donc, d'après le lemme 2.2, on a  $\hat{z} \in C$  et, par suite,

$$f(\hat{z}) \geq \bar{m}. \quad (2.9)$$

Mais, pour  $i \in K$ , on a  $f(z_i^*) \leq \bar{m} - P'_i(z_i^*)$  et  $P'_i(z_i^*) > 0$ , alors quand  $i \rightarrow \infty$ , on obtient

$$f(\hat{z}) \leq \bar{m}. \quad (2.10)$$

De (2.9) et (2.10), on conclut que  $f(\hat{z}) = \bar{m}$ , c'est-à-dire  $\hat{z}$  est solution optimale du problème (PAC). □

### Remarque 2.1

Les fonctions de pénalité extérieure peuvent non seulement être utilisées pour transformer un problème de minimisation avec contraintes en une suite de problèmes de minimisation sans contraintes, mais aussi pour éliminer des contraintes que certains algorithmes n'acceptent pas.

Supposons que l'on veut résoudre un problème non linéaire de la forme suivante :  $\min\{f(z) : h(z) = 0, g(z) \leq 0\}$  et que la fonction  $h(\cdot)$  est non affine. Supposons que  $\{P'_i\}_{i \in \mathbb{N}}$  est une série de fonctions de pénalité extérieure pour l'ensemble  $\{z : h(z) = 0\}$ , alors sous certaines hypothèses, nous pouvons utiliser plusieurs méthodes pour résoudre la série de problèmes  $\min\{f(z) + P'_i(z) : g(z) \leq 0\}$ ,  $i = 0, 1, 2, \dots$  pour obtenir une solution du problème original. En enlevant uniquement quelques contraintes avec les fonctions de pénalité, nous espérons obtenir de meilleurs résultats dans nos calculs.

### Exemples [38]

- a). Soit  $g_j : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $j = 1, 2, \dots, m$  des fonctions continues et  $C = \{z \in \mathbb{R}^n : g_j(z) \leq 0, j = 1, 2, \dots, m\}$ . Fiacco et McCormack (1968) ont proposé de prendre comme fonctions de pénalité extérieure pour l'ensemble  $C$

$$P'_i(z) = \alpha_i \sum_{j=1}^m [\max\{g_j(z), 0\}]^\beta, \quad i = 0, 1, 2, \dots$$

où  $\beta \geq 1$  un réel arbitraire et la suite  $\{\alpha_i\}_{i \in \mathbb{N}}$  est strictement croissante de nombres positifs et qui tend vers  $\infty$  lorsque  $i \rightarrow \infty$ .

- b). On suppose  $h_j : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $j = 1, 2, \dots, m$  des fonctions continues et  $C = \{z \in \mathbb{R}^n : h_j(z) = 0, j = 1, 2, \dots, m\}$ . Les mêmes auteurs proposent de prendre

$$P'_i(z) = \alpha_i \|h(z)\|^\beta = \alpha_i \left[ \sum_{j=1}^m (h_j(z))^2 \right]^{\frac{\beta}{2}}, \quad i = 0, 1, 2, \dots$$

où  $\beta \geq 1$  un réel arbitraire et la suite  $\{\alpha_i\}_{i \in \mathbb{N}}$  est strictement croissante de nombres positifs et qui tend vers  $\infty$  lorsque  $i \rightarrow \infty$ .

### 2.1.2 Méthodes de pénalité intérieure

A présent, on procède à la résolution du problème (PAC) par la méthode des fonctions de pénalité intérieure.

**Hypothèse** Supposons que :

- a). L'ensemble  $C$  est fermé,  $\text{int}(C) \neq \emptyset$  et la fermeture de l'intérieur de  $C$  est égale à  $C$ . C'est-à-dire  $\text{int}(C) = C \neq \emptyset$ .
- b). L'hypothèse 2.1.1 est satisfaite.

**Définition 2.2** Soit  $C$  un sous ensemble de  $\mathbb{R}^n$ , qui satisfasse l'hypothèse 2.1.2. Une série de fonctions continues  $P_i'' : \text{int}(C) \rightarrow \mathbb{R}$ ,  $i = 0, 1, 2, \dots$  est dite une série de fonctions de pénalité intérieure pour l'ensemble  $C$ , si :

$$0 < P_{i+1}''(z) < P_i''(z), \quad \forall z \in \text{int}(C), \quad i = 0, 1, 2, \dots \quad (2.11)$$

$$P_i''(z) \rightarrow 0, \quad \text{quand } i \rightarrow \infty \text{ et } \forall z \in \text{int}(C) \quad (2.12)$$

$$P_i''(z^j) \rightarrow \infty, \quad \text{pour toute série } \{z^j\} \in \text{int}(C) \quad (2.13)$$

telle que  $z^j \rightarrow z^* \in \partial C$ ,  
quand  $j \rightarrow \infty$ ,  $i = 0, 1, 2, \dots$

$\partial C$  désigne la frontière de  $C$ .

■

Considérons la série de problèmes  $(PSC)_j^i$ <sup>2</sup>

$$(PSC)_j^i \quad \min\{f(z) + P_j''(z) : z \in \text{int}(C)\}, \quad j = 0, 1, 2, \dots,$$

où  $P_j''(\cdot)$  sont les fonctions de pénalité intérieure pour l'ensemble  $C$ .

**Lemme 2.3** Considérons les problèmes  $(PSC)_j^i$  définis ci-dessus et supposons qu'il existe un  $z'' \in \text{int}(C)$  tel que l'ensemble  $\{z : f(z) \leq f(z'') + P_0''(z'')\}$  est compact. Alors pour  $j = 0, 1, 2, \dots$  il existe  $z_j^* \in \text{int}(C)$  qui minimise  $f(z) + P_j''(z)$  sur l'ensemble  $C$ .

### Preuve

Soit  $z'' \in \text{int}(C)$  tel que l'ensemble

$$Z' = \{z : f(z) \leq f(z'') + P_0''(z'')\}$$

est compact.

Posons

$$C_j = \{z \in \text{int}(C) : f(z) + P_j''(z) \leq f(z'') + P_j''(z'')\}, \quad j = 0, 1, 2, \dots$$

On a  $\forall z \in C_j$ ,  $z \in \text{int}(C)$  et  $f(z) + P_j''(z) \leq f(z'') + P_j''(z'')$ .

$$\Rightarrow f(z) \leq f(z'') + P_j''(z'') - P_j''(z)$$

$$\Rightarrow f(z) \leq f(z'') + P_j''(z'') \quad (\text{car } P_j''(z) \geq 0).$$

<sup>2</sup>L'indice "i" dans  $(PSC)_j^i$  signifie intérieur.

Hors, d'après la relation (2.11),  $P_0''(z'') \geq P_j''(z'')$ ,  $\forall j$ .

Donc,  $f(z) \leq f(z'') + P_0''(z'')$ , d'où  $z \in Z'$ . On déduit alors, que  $C_j$  est contenu dans  $Z'$ .

Soit  $\{z_i^*\}_{i \in \mathbb{N}} \subset C_j$  telle que  $\lim_{i \rightarrow \infty} z_i^* = z^*$ . Comme  $f(\cdot)$  et  $P_j''(\cdot)$  sont des fonctions continues, alors

$$\lim_{i \rightarrow \infty} [f(z_i^*) + P_j''(z_i^*)] = f(z^*) + P_j''(z^*), \quad j = 0, 1, 2, \dots \quad (2.14)$$

On a  $\{z_i^*\}_{i \in \mathbb{N}} \subset C_j$ , c'est-à-dire :  $\forall j, z_i^* \in C_j$ . Alors,  $z_i^* \in \text{int}(C)$  et

$$f(z_i^*) + P_j''(z_i^*) \leq f(z'') + P_j''(z''), \quad j = 0, 1, 2, \dots$$

D'après (2.14), on aura

$$\lim_{i \rightarrow \infty} [f(z_i^*) + P_j''(z_i^*)] = f(z^*) + P_j''(z^*) \leq f(z'') + P_j''(z''),$$

d'où  $z^* \in C_j$  (c'est-à-dire que  $C_j$  est un ensemble fermé).

Donc  $C_j$  est compact et il existe un  $z_j^* \in C_j$  qui minimise  $f(z) + P_j''(z)$  sur l'ensemble  $C_j$ . Mais pour tout  $z \in \text{int}(C)$  et  $z \notin C_j$  on a

$$f(z) + P_j''(z) > f(z'') + P_j''(z''),$$

donc  $z_j^*$  est solution optimale pour  $(PSC)_j^i$ .

□

**Théorème 19** [43] Soit  $z_j$  une solution optimale pour le problème sans contraintes  $(PSC)_j^i$ ,  $j = 0, 1, 2, \dots$ , et supposons que l'hypothèse 2.1.2 est vérifiée. Tout point limite  $z^*$  de la suite  $\{z_j^*\}_{j \in \mathbb{N}}$  est solution optimale du problème avec contraintes (PAC).

### Preuve

Pour  $j = 0, 1, 2, \dots$ , soient  $b_j = \min\{f(z) + P_j''(z) : z \in \text{int}(C)\}$ ,  
 $b = \min\{f(z) : z \in C\}$ .

De la relation (2.11), on a

$$b_0 \geq b_1 \geq b_2 \geq \dots \geq b_j \geq b_{j+1} \geq \dots \geq b.$$

Vu que les  $b_j$  forment une suite monotone, décroissante et bornée, elle converge vers  $b' \geq b$ . Supposons que  $b' > b$ , soit  $\hat{z}$  une solution optimale du problème (PAC),

comme  $f(\cdot)$  est continue et l'hypothèse 2.1.2.a), il existe une boule ouverte  $B$ , de centre  $\hat{z}$  telle que  $B \cap \text{int}(C) \neq \emptyset$  et pour tout  $z'' \in B$ , on a

$$f(z'') < b' - \frac{b' - b}{2}.$$

Prenons  $z'' \in B \cap \text{int}(C)$ . Alors de la relation (2.12), on a  $P_j''(z'') \rightarrow 0$  quand  $j \rightarrow \infty$ . Il existe un entier  $N$  tel que pour tout  $j \geq N$ ,

$$P_j''(z'') < \frac{b' - b}{4},$$

donc pour tout  $j \geq N$ ,

$$b_j \leq f(z'') + P_j''(z'') < b' - \frac{b' - b}{4},$$

ce qui contredit le fait que  $\lim_{j \rightarrow \infty} b_j = b'$ . Donc on a :  $b' = b$ .

A présent, soit  $z^*$  un point limite de la suite  $\{z_j\}_{j \in \mathbb{N}}$ , c'est-à-dire  $\lim_{j \rightarrow \infty} z_j^* = z^*$ ,  $j \in K \subset \{0, 1, 2, \dots\}$ . Supposons que  $z^*$  n'est pas une solution optimale du problème (PAC), alors on a  $f(z^*) > b$ , et donc la suite  $\{[f(z_j) - b] + P_j''(z_j)\}_{j \in K}$  ne peut pas converger vers 0. Ce qui contredit le fait que  $(b_j - b) \rightarrow 0$  quand  $j \rightarrow \infty$ . Donc on a  $f(z^*) = b$  et  $z^*$  est solution optimale du problème (PAC). □

### Exemples [38]

- a). Si on suppose  $g_j : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $j = 1, 2, \dots, m$  des fonctions continues et  $C = \{z \in \mathbb{R}^n : g_j(z) \leq 0, j = 1, 2, \dots, m\}$ . Supposons que l'hypothèse 2.1.2 est satisfaite. Pour  $z \in \text{int}(C)$ ,  $g_j(z) < 0$ ,  $j = 1, 2, \dots, m$   
Une fonction de pénalité intérieure pour l'ensemble  $C$  peut s'écrire

$$P_i''(z) = -\alpha_i \sum_{j=1}^m \frac{1}{g_j(z)}, \quad z \in \text{int}(C), i = 0, 1, 2, \dots$$

où  $\{\alpha_i\}_{i \in \mathbb{N}}$ , une suite strictement décroissante de nombres positifs qui converge vers 0 lorsque  $i \rightarrow \infty$ .

- b). Ou bien on peut prendre

$$P_i''(z) = -\alpha_i \sum_{j=1}^m \log[-g_j(z)] \quad z \in \text{int}(C), i = 0, 1, 2, \dots$$

où  $\{\alpha_i\}_{i \in \mathbb{N}}$ , une suite strictement décroissante de nombres positifs qui converge vers 0 lorsque  $i \rightarrow \infty$ .

### Remarque 2.2

Les  $\alpha_i > 0$  sont appelés coefficients de pénalité. Pour  $\alpha_i > 0$  quelconque, notons  $\bar{z}(\alpha_i)$  un minimum de  $(PSC)_i$ <sup>3</sup>.

Le choix d'une valeur appropriée du coefficient de pénalité  $\alpha_i$  résulte d'un compromis ; d'une part,  $\alpha_i$  doit être choisi suffisamment grand pour que le point  $\bar{z}(\alpha_i)$  obtenu soit proche de l'ensemble des solutions  $C$  (autrement dit, que  $P_i(\bar{z}(\alpha_i))$  soit suffisamment faible) ; d'autre part, si  $\alpha_i$  est choisi trop grand, la fonction  $P_i(\cdot)$  peut être mal conditionnée, d'où des difficultés numériques dans la recherche de l'optimum du problème sans contraintes. Ceci explique pourquoi les méthodes de pénalités sont généralement mises en oeuvre de la façon suivante :

On commence par choisir un coefficient de pénalité  $\alpha_1$  de valeur pas trop élevée (pour éviter les difficultés numériques), puis on résout le problème sans contraintes

$$\min_z (PSC)_1.$$

Soit  $\bar{z}(\alpha_1)$  la solution obtenue. Si la quantité  $P_1(\bar{z}(\alpha_1))$  est suffisamment faible,  $\bar{z}(\alpha_1)$  est une bonne approximation de l'optimum, et les calculs sont terminés.

Dans le cas contraire, la pénalité associée à la violation des contraintes n'est pas assez élevée. On choisira donc un coefficient de pénalité  $\alpha_2 > \alpha_1$  (par exemple :  $\alpha_2 = 10 \times \alpha_1$ ) et on résout le nouveau problème sans contrainte :

$$\min_z (PSC)_2$$

On réitère le processus jusqu'à avoir une meilleure approximation de l'optimum recherché.

Remarquons qu'à chaque étape  $i$  du processus précédent, il est avantageux d'utiliser le point  $\bar{z}(\alpha_{i-1})$  obtenu, à l'étape précédente, comme solution de départ de l'algorithme d'optimisation sans contrainte choisi (gradient conjugué, méthode de quasi Newton, ...).

## 2.2 Méthode DC

Dans le domaine de la programmation non convexe, la programmation DC joue un rôle très important, grâce à son aspect théorique qu'encore ses applications pratiques assez larges. une fonction est dite DC si elle peut être représentée comme la différence de deux fonctions convexes. Les problèmes de programmation

---

<sup>3</sup>Valable pour les deux cas, pénalité extérieure et intérieure

Mathématique traitant avec des fonctions DC sont dites problèmes de programmation DC.

On présentera dans ce qui suit les principaux résultats de cette théorie, ses applications, et la méthode de résolution dans le sens de l'optimisation globale. On se restreint aux approches déterministes et à la classe des programmes DC traitants avec les fonctions DC.

### 2.2.1 Les fonctions DC

Afin d'étendre la programmation convexe à la résolution de la plupart des problèmes d'optimisation non convexes de la vie courante tout en continuant à utiliser son arsenal théorique et numérique, une nouvelle classe de fonctions fut introduite : la classe des fonctions D.C :

Soit  $C$  un ensemble convexe de  $X$ .

**Définition 2.3** Une fonction  $f : C \subset X \rightarrow R \cup \{+\infty\}$  est dite DC sur  $C$  si elle s'écrit

$$f(x) = g(x) - h(x), \text{ pour tout } x \in C, \quad (2.15)$$

où  $g$  et  $h$  sont des fonctions convexes sur  $C$ .

On note  $DC(C)$  l'ensemble des fonctions DC sur  $C$ , c'est également l'espace vectoriel engendré par  $Conv(C)$ . Mise à part sa structure d'espace vectoriel,  $DC(C)$  est également stable par rapport aux opérations usuelles utilisées en optimisation.

Si  $C = \mathbb{R}^n$ , alors  $f$  est simplement dite une fonction DC.

Une représentation de la forme (2.15) est dite une décomposition DC de  $f$ .

**Définition 2.4** Soit  $C$  un ouvert convexe,  $f$  est dite localement DC sur  $C$  si tout point  $x^0$  de  $C$  admet un voisinage  $V(x^0)$  noté  $V$ , tel que pour tout  $x$  dans  $V$

$$f(x) = g_V(x) - h_V(x) \text{ pour tout } g_V, h_V \in Conv(V).$$

Cette définition peut être réécrite comme suit :

**Définition 2.5** On dit qu'une fonction  $f$  est localement DC, si, pour chaque  $x_0 \in \mathbb{R}^n$ ,  $\exists \epsilon > 0$  tel que  $f$  est DC sur la boule :  $B(x_0, \epsilon) = \{x \in \mathbb{R}^n : \|x - x_0\| \leq \epsilon\}$ .

Un résultat important concernant la reconnaissance des fonctions DC vient de Hartman :

**Proposition 10** [28]

Chaque fonction localement DC est DC.

Ce résultat clé nous mène aux conséquences importantes suivantes :

**Proposition 11 [26]**

- a). Chaque fonction  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  dont les dérivées partielles sont continues est DC.
- b). Soit  $C$  un sous ensemble compact convexe de  $\mathbb{R}^n$ . Alors chaque fonction continue sur  $C$  est la limite d'une séquence de fonctions DC qui convergent uniformément sur  $C$  ; c'est à dire,  $\forall c : C \rightarrow \mathbb{R}$  continue et  $\forall \varepsilon > 0$ , il existe une fonction DC  $f : C \rightarrow \mathbb{R}$  telle que  $|c(x) - f(x)| \leq \varepsilon, \forall x \in C$
- c). Soit  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  DC et  $g : \mathbb{R} \rightarrow \mathbb{R}$  convexe. Alors, la fonction composée :  $g \circ f$  est DC.

Une sous-classe importante de fonctions de  $\text{Conv}(C)$  est la classe de fonctions  $C^2$  sur  $\mathbb{R}^n$ . Cette classe est particulièrement importante car les fonctions objectif de la plupart des problèmes d'optimisation de la vie courante appartiennent à cette sous-classe.

**Proposition 12 [44]**

Les deux propositions suivantes sont équivalentes

- $f$  est dans  $L_{C^2}(C)$  (le cône convexe des fonctions localement enveloppe supérieure d'une famille de fonctions de  $C^2(C)$ .)
- $f$  est localement DC sur  $C$  avec  $h_V$  une forme quadratique convexe.

On a le théorème suivant,

**Théorème 20 [44]**

Soit  $C$  un ouvert convexe de  $X$ .

- Toute fonction localement DC sur  $C$  est DC sur  $C$ . En particulier toute fonction de classe  $C^2$  sur  $C$  est DC sur  $C$ .
- On a la chaîne d'inclusions

$$\text{Conv}(C) \subset L_{C^2}(C) \subset \text{DC}(C).$$

- Si  $C$  est en plus compact, alors  $\text{DC}(C)$  est un sous-espace vectoriel dense dans l'ensemble des fonctions continues sur  $C$  muni de la norme de la convergence uniforme sur  $C$ .

## 2.2.2 Les problèmes de Programmation DC

Les problèmes traitant avec les fonctions DC sont des problèmes de Programmation DC. La forme générale d'un problème de programmation DC est donnée comme suit :

$$\min\{f_0(x) : x \in X, f_i(x) \leq 0, i = 1, \dots, m\}, \quad (2.16)$$

où  $f_i = g_i - h_i$ ,  $i = 1, \dots, m$  sont des fonctions DC ( $g_i, h_i$ ,  $i = 1, \dots, m$  sont des fonctions convexes) et  $X$  est un sous ensemble convexe fermé de  $\mathbb{R}^n$ .

Les problèmes

$$\min\{c(x) : x \in X, \psi(x) \leq 0\}, \quad (2.17)$$

où  $c$  est une fonction linéaire,  $X$  est un sous ensemble convexe fermé de  $\mathbb{R}^n$  et  $\psi$  est une fonction concave est dit programme DC canonique.

Chaque programme de type (2.16) peut être transformé en un programme canonique (2.17) comme suit :

En utilisant une variable additionnelle  $x_{n+1}$ , le programme (2.16) est réécrit comme suit :

$$\min\{x_{n+1} : f_0(x) - x_{n+1} \leq 0, x \in X, f_i(x) \leq 0, i = 1, \dots, m\}, \quad (2.18)$$

Définissons, à présent la fonction DC  $f(x, x_{n+1}) = \max\{(f_0(x) - x_{n+1}), f_1(x), \dots, f_m(x)\}$  et soit  $f = g - h$  une décomposition DC de  $f$ . En utilisant maintenant une autre variable additionnelle  $x_{n+2}$  et en posant :

$$z = (x, x_{n+1}, x_{n+2}) \in \mathbb{R}^{n+2} \quad (2.19)$$

$$Z = \{z \in \mathbb{R}^{n+2}, x \in X, g(x, x_{n+1}) - x_{n+2} \leq 0\} \quad (2.20)$$

$$\psi(z) = x_{n+2} - h(x, x_{n+1}) \quad (2.21)$$

on obtient le programme DC canonique suivant :

$$\min\{c(z) = x_{n+1} : z \in Z, \psi(z) \leq 0\},$$

Pour l'établissement des conditions d'optimalité et le développement des méthodes de résolution du problème (2.17), la fonction linéaire  $c$  est remplacée par une fonction convexe. Ainsi, après un tel changement, le programme (2.17) est dit programme DC canonique généralisé.

## 2.2.3 Conditions d'optimalité en optimisation DC

On commencera cette section par le rappel du concept de réciprocity de Tikhonov, initialement introduit pour l'optimisation non convexe et les problèmes variationnels mal posés. Considérons la paire de programmes donnée par :

$$w_{\delta}^* = \inf\{w(z) : z \in Z, \psi(z) \leq \delta\} \quad (2.22)$$

$$\psi_{\eta}^* = \inf\{\psi(z) : z \in Z, w(z) \leq \eta\} \quad (2.23)$$

où  $Z \subseteq \mathbb{R}^n$ ,  $w$  et  $\psi$  sont des fonctions finies dans  $\mathbb{R}^1$ . Soit  $\Omega_{\delta}^*$  et  $\Psi_{\eta}^*$  les ensembles des solutions optimales des problèmes (2.22) et (2.23) respectivement.

**Définition 2.6** On dit que les problèmes (2.22) et (2.23) sont réciproques si  $\Omega_{\delta}^* = \Psi_{\eta}^*$ ; dans ce cas, on dit aussi que le principe de réciprocité est vérifié.

**Proposition 13** [54]

Si  $Z = \mathbb{R}^p$ ,  $w(z) = \|z\|$  et  $\eta = w_{\delta}^*$  alors le principe de réciprocité est vérifié pour n'importe quelle fonction  $\psi(z)$  vérifiant  $\{z \in Z : \psi(z) \leq \Omega_{\delta}^*\} \neq \emptyset$ .

On va énoncer à présent un résultat clé à partir duquel on dérive les conditions d'optimalité

**Proposition 14** [57]

Supposons que dans les problèmes (2.22) et (2.23) la condition suivante est vérifiée :  $\delta = \psi_{\eta}^*$  et  $\eta = w_{\delta}^*$ . Alors, le principe de réciprocité est vérifié pour n'importe quelle ensemble  $Z$  et des fonctions  $w$  et  $\psi$  quelconques.

On donne, par la suite quelques conditions d'optimalité des problèmes d'optimisation traitant avec les fonctions DC, dérivant immédiatement de la proposition (14).

On commence par introduire les ensembles robustes :

**Définition 2.7** Soit  $Z \in \mathbb{R}^p$ ,  $f : \mathbb{R}^p \rightarrow \mathbb{R}$  et  $\delta \in \mathbb{R}$ . L'ensemble  $S(Z, f, \delta) = \{z \in Z : f(z) \leq \delta\}$  est dit robuste si  $S(Z, f, \delta) = cl(\{z \in Z : f(z) < \delta\})$ , où  $cl(S)$  est la fermeture de l'ensemble  $S$ .

**Lemme 2.4** [57]

Si dans les problèmes (2.22) et (2.23) la fonction  $\psi$  est convexe, et si la condition suivante est remplie :

$$\exists z^0 \in Z \text{ telle que } w(z^0) < w^* \quad (2.24)$$

alors l'ensemble  $S(Z, w, \eta)$  est robuste pour n'importe quel  $\eta \geq w^*$ .

La condition d'optimalité suivante concernant l'optimisation DC canonique généralisée est un corollaire de la proposition 14. Notons que le programme DC canonique défini précédemment est un cas particulier du problème (2.22), où  $\delta = 0$ , et les fonctions  $w$  et  $\psi$  sont convexes.

**Proposition 15** [54]

Supposons que dans le programme DC canonique généralisé

$$w^* = \inf\{w(z) : z \in Z, \psi(z) \leq 0\}, \quad (2.25)$$

l'ensemble  $S(Z, w, 0) = \{z \in Z : w(z) \leq 0\}$  est robuste et la condition (2.24) est remplie. Alors un point réalisable  $z^*$  est solution optimale de (2.25) si et seulement si

$$0 = \inf\{\psi(z) : z \in Z, w(z) \leq w(z^*)\}. \quad (2.26)$$

Une classe importante des programmes d'optimisation DC est la suivante :

$$w^* = \inf\{g(x) - h(x) : x \in X\}, \quad (2.27)$$

où  $g$  et  $h$  sont des fonctions convexes de  $\mathbb{R}^n$ ,  $X$  est un sous-ensemble convexe et fermé de  $\mathbb{R}^n$ . La proposition suivante donne une condition d'optimalité du problème (2.27) :

**Proposition 16** [44]

Supposons que le problème (2.27) admet une solution. Alors un point  $x^* \in X$  est solution optimale du problème (2.27) si et seulement si

$$\exists t^* \in \mathbb{R} \text{ telle que : } 0 = \inf\{-h(x) + t : x \in X, g(x) - t \leq g(x^*) - t^*\}. \quad (2.28)$$

En se basant sur la condition d'optimalité ci-dessus, on peut développer un algorithme de résolution des problèmes de type (2.27). En utilisant les notations du chapitre précédent on obtient la condition d'optimalité géométrique suivante pour le programme (2.27) :

**Proposition 17** [44]

Un point  $x^* \in X$  est une solution optimale du problème (2.27) si et seulement si

$$\text{epi}(\bar{g}|_X) \subset \text{epi}(h|_X) \quad (2.29)$$

où  $\bar{g}(x) = g(x) - (g(x^*) - h(x^*))$

La condition (2.29) peut être réécrite comme suit :

**Proposition 18** [44]

Un point  $x^* \in X$  est une solution optimale du problème (2.27) si et seulement si

$$\partial_\varepsilon h(x^*) \subset \partial_\varepsilon g(x^*), \quad \forall \varepsilon \geq 0. \quad (2.30)$$

**Remarque 2.3** a). Pour le cas où  $X = \mathbb{R}^n$ , la condition (2.30) devient

$$\partial_\varepsilon h(x^*) \subset \partial_\varepsilon g(x^*) \quad \forall \varepsilon > 0. \quad (2.31)$$

C'est un des résultats les plus utilisés dans l'optimisation DC et il a été démontré par Hiriart Urruty [9, 10, 11]

b). Si  $g \equiv 0$ , le problème (2.27) est dit problème de Programmation concave. En utilisant la fonction indicatrice de  $X$ , ce problème peut être réécrit comme suit :

$$\inf\{I_X(x) - h(x) : x \in \mathbb{R}^n\}, \quad (2.32)$$

et la condition (2.30) devient

$$\partial_\varepsilon h(x^*) \subset N_\varepsilon(X, x^*) \quad \forall \varepsilon \geq 0, \quad (2.33)$$

où  $N_\varepsilon(X, x^*)$  est le cône  $\varepsilon$ -normal de  $X$  en  $x^*$  défini dans (1.1). Il a été démontré [37] que la condition (2.33) est équivalente aux conditions suivantes :

$$h(x^*) = \sup\{I_X^*(y) - h^*(y) : y \in \mathbb{R}^n\}, \quad (2.34)$$

$$\partial h(z) \subset N_\varepsilon(X, z) \quad \forall z \in \{x \in X : h(x) = h(x^*)\} \quad (2.35)$$

L'équation (2.34) est établie en se basant sur le dual du problème (2.32).

## 2.2.4 Dualité en optimisation DC

Considérons le problème d'optimisation

$$(P_{dc}) \quad \alpha = \inf\{f(x) = g(x) - h(x) : x \in \mathbb{R}^n\}, \quad (2.36)$$

où  $g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  et  $h : \mathbb{R}^n \rightarrow \mathbb{R}$  sont deux fonctions convexes. Notons que le problème de la forme  $\inf\{f(x) = g(x) - h(x) : x \in X\}$  peut être réécrit de la forme (2.36) en posant  $g = g + I_X$  où  $I_X$  est la fonction indicatrice de  $X$ .

Soient  $g^*$  et  $h^*$  les fonctions conjuguées de  $g$  et  $h$  respectivement, et soit

$$\mathcal{Y} = \{y \in \mathbb{R}^n : h^*(y) < +\infty\}, \quad (2.37)$$

alors, le problème de programmation suivant

$$(D_{dc}) \quad \alpha = \inf\{h^*(y) - g^*(y) : y \in \mathcal{Y}\}, \quad (2.38)$$

est dit programme dual du programme (2.36) selon Fenchel-Rockafellar.

Avant d'établir la relation entre les problèmes (2.36) et (2.38), on donnera quelques formules pour le calcul des fonctions conjuguées.

**Lemme 2.5** [44]

a). Soit  $\bar{x} \in \mathbb{R}^n$  et  $\bar{y} \in \partial h(\bar{x})$ . Alors,

$$h^*(\bar{y}) = \langle \bar{y}, \bar{x} \rangle - h(\bar{x}). \quad (2.39)$$

b). Soit  $g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  une fonction quelconque, et soit  $h : \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction convexe et soit  $f = g - h$ . Alors, la fonction conjuguée  $f^*$  de  $f$  est donnée par

$$f^*(z) = \sup\{g^*(y+z) - h^*(y) : y \in \mathcal{Y}\}. \quad (2.40)$$

En utilisant le lemme 2.5, on obtient le résultat suivant qui établit la relation entre les problèmes  $(P_{dc})$  (2.36) et  $(D_{dc})$  (2.38).

**Théorème 21** [44]

(i) Si le problème (2.36) admet une solution optimale alors il s'ensuit que

$$\inf\{f(x) = g(x) - h(x) : x \in \mathbb{R}^n\} = \inf\{h^*(y) - g^*(y) : y \in \mathcal{Y}\}. \quad (2.41)$$

(ii) Le dual du programme (2.38) s'identifie par le programme (2.36)

En examinant le programme (2.38), et du fait que  $h^*$  et  $g^*$  sont convexes, on constate que ce programme est lui même un programme DC de type (2.36). On peut remarquer aussi une parfaite symétrie entre le programme DC et son dual qu'on résume dans la proposition suivante :

**Proposition 19** [44]

(i) Si  $x^*$  est une solution optimale du problème (2.36) alors chaque  $y^* \in \partial h(x^*)$  est une solution optimale du problème (2.38).

(ii) Si  $y^*$  est une solution optimale du problème (2.38) alors chaque  $x^* \in \partial g^*(y^*)$  est une solution optimale du problème (2.36).

On observe la parfaite symétrie entre le problème primal  $(P_{dc})$  et le problème dual  $(D_{dc})$ ; il va de soi que les résultats établis pour l'un se transposent à l'autre.

On a alors les résultats suivants :

Soient  $S_P$  et  $S_D$  les ensembles de solutions des problèmes  $(P_{dc})$  et  $(D_{dc})$  respectivement, et posons

$$P_l = \{x \in S_P \mid \partial h(x^*) \subset \partial g(x^*)\} \text{ et } D_l = \{y \in S_D \mid \partial g^*(y^*) \subset \partial h^*(y^*)\}$$

Le théorème suivant est celui à partir duquel DCA est bâti.

**Théorème 22** [44, 26]

i) *Transport de minima globaux :*

$$\cup\{\partial h(x) \mid x \in S_P\} \subset S_D \subset \text{dom}(h^*).$$

*La première inclusion se transforme en égalité si  $g^*$  est sous-différentiable sur  $S_D$  (en particulier si  $S_D \subset \text{ri}(\text{dom}(g^*))$ ) ou si  $g^*$  est sous-différentiable sur  $\text{dom}(h^*)$  et dans ce dernier cas  $S_D \subset (\text{dom}(\partial g^*) \cap \text{dom}(\partial h^*))$ .*

ii) *Si  $x^*$  est un minimum local de  $g - h$ , alors  $x^* \in P_l$ . L'implication inverse est vraie si  $h$  est une fonction convexe polyédrale.*

iii) *Soit  $x^*$  un point critique ( $x^*$  est un point critique de  $g - h$  si  $\partial h(x^*) \cap \partial g(x^*)$  est non vide.) de  $g - h$  et  $y^* \in \partial g(x^*) \cap \partial h(x^*)$ . Soit  $U$  un voisinage de  $x^*$  tel que  $U \cap \text{dom}(g) \subset \text{dom}(\partial h)$ . Si pour tout  $x \in U \cap \text{dom}(g)$  il existe un  $y \in \partial h(x)$  tel que  $h^*(y) - g^*(y) \geq h^*(y^*) - g^*(y^*)$ , alors  $x^*$  est un minimum local de  $g - h$ . Plus précisément,*

$$g(x) - h(x) \geq g(x^*) - h(x^*), \text{ pour tout } x \in U \setminus \text{dom}(g). \quad (2.42)$$

iv) *Transport de minima locaux : Soit  $x^* \in \text{dom}(\partial h)$  un minimum local de  $g - h$  et soit  $y^* \in \partial h(x^*)$ . Sous l'hypothèse*

$$y^* \in \text{int}(\text{dom}(g^*)) \text{ et } \partial g^*(y^*) \subset U, \quad (2.43)$$

*par exemple si  $g^*$  est différentiable en  $y^*$ , alors  $y^*$  est un minimum local de  $h^* - g^*$ .*

*Notons que ce théorème admet sa forme duale grâce à la symétrie observée dans la section précédente entre le problème  $(P_{dc})$  et le problème  $(D_{dc})$ . Notons également que tous ces résultats concernent les composantes DC  $g$  et  $h$  et non la fonction  $f$  elle-même.*

**Proposition 20** *Soit  $f = g - h$  avec  $g, h \in \Gamma_0(\mathcal{X})$  vérifiant  $\text{dom}(g) \subset \text{dom}(h)$  et  $\text{ir}(\text{dom}(g)) \cap \text{ir}(\text{dom}(h)) \neq \emptyset$ . Soit  $x^0 \in \text{dom}(g)$  (resp.  $\text{dom}(h)$ ) un point où  $g$  (resp.  $h$ ) est continue. Si  $f$  est convexe, alors*

(i)  *$h$  est continue en  $x^0$*

(ii)  *$\partial h(x^0) = \partial g(x^0) * \partial h(x^0)$  ( $A * B = \{x \in \mathcal{X} : x + B \subset A\}$ )*

(iii)  *$0 \in \partial f(x^0) \iff \partial h(x^0) \subset \partial g(x^0)$*

$$\emptyset \neq \partial h(x^*) \subset \partial g(x^*) \quad (\text{resp. } \emptyset \neq \partial g^*(y^*) \subset \partial h^*(y^*)), \quad (2.44)$$

$$\partial h(x^*) \cap \partial g(x^*) \neq \emptyset \quad (\text{resp. } \partial g^*(y^*) \cap \partial h^*(y^*) \neq \emptyset) \quad (2.45)$$

La condition (2.44) est dite conditions de Kuhn-Tucker généralisées pour  $(P_{dc})$ . Elle coïncide avec (2.45) si  $h$  est différentiable en  $x^*$ . Un point  $x^*$  vérifiant (2.45) est dit point critique de  $g - h$ .

Pour les programmes DC polyédraux, une condition nécessaire et suffisante pour qu'un point  $x^*$  soit un minimum local de  $g - h$  est

$$\partial h(x^*) \subset \partial g(x^*).$$

### Algorithme d'optimisation DC : DCA

La construction des DCA repose sur la génération de deux suites  $\{x^k\}_k$  et  $\{y^k\}_k$  qui sont améliorées à chaque itération de sorte que leur limite respective  $x^*$  et  $y^*$  soient candidates pour être les optima locaux du problème primal et du problème dual respectivement. Ces deux suites sont liées par la dualité et vérifient les propriétés suivantes :

- Les suites  $\{g(x^k) - h(x^k)\}$  et  $\{g^*(y^k) - h^*(y^k)\}$  décroissent à chaque itération.
- Si  $(g - h)(x^{k+1}) = (g - h)(x^k)$  l'algorithme s'arrête à l'itération  $k + 1$ , et le point  $x^k$  (resp.  $y^k$ ) est un point critique de  $g - h$  (resp.  $h^* - g^*$ ),
- sinon toute valeur d'adhérence  $x^*$  de la suite  $\{x^k\}_k$  (resp.  $y^*$  de la suite  $\{y^k\}_k$ ) est un point critique de  $g - h$  (resp.  $h^* - g^*$ ).

Ces suites sont générées de la manière suivante :  $x^{k+1}$  (resp.  $y^k$ ) est solution du problème convexe  $(P_k)$  (resp.  $(D_k)$ ) défini par

$$(P_k) \quad \alpha_k = \inf\{g(x) - [h(x^k) + \langle x - x^k, y^k \rangle] \mid x \in X\}, \quad (2.46)$$

$$(D_k) \quad \inf\{h^*(y) - [g^*(y^{k-1}) + \langle x^k, y - y^{k-1} \rangle] \mid y \in Y\}. \quad (2.47)$$

Comme on peut le voir  $(P_k)$  (resp.  $(D_k)$ ) est obtenu de  $(P_{dc})$  (resp.  $(D_{dc})$ ) en remplaçant  $h$  (resp.  $g^*$ ) par sa minorante affine définie par  $y^k \in \partial h(x^k)$  (resp.  $x^k \in \partial g^*(y^{k-1})$ ), DCA conduit au schéma

$$\begin{array}{ccc} x^k & \longrightarrow & y^k \in \partial h(x^k) \\ & \swarrow & \\ x^{k+1} \in \partial g^*(y^k) & \longrightarrow & y^{k+1} \in \partial h(x^{k+1}) \end{array}$$

Ce schéma correspond à la forme simplifiée de DCA, et se traduit par l'algorithme ci-dessous.

### Algorithme DCA simplifié

0.  $x^0$  donné.

1. Pour chaque  $k$ ,  $x^k$  étant connu, déterminer  $y^k \in \partial h(x^k)$ .

2. Trouver  $x^{k+1} \in \partial g^*(y^k)$ .

3. Si test d'arrêt vérifié Stop; sinon  $k \leftarrow k+1$  et aller en 1.

Dans la forme complète de DCA on impose le choix suivant :

$$x^{k+1} \in \arg \min \{g(x) - h(x) : x \in \partial g^*(y^k)\} \quad (2.48)$$

$$y^k \in \arg \min \{h^*(y) - g^*(y) : y \in \partial h^*(x^k)\} \quad (2.49)$$

Les problèmes (2.48) et (2.49) sont équivalents aux problèmes respectifs

$$x^{k+1} \in \arg \min \{\langle x, y^k \rangle - h(x) : x \in \partial g^*(y^k)\} \quad (2.50)$$

$$y^k \in \arg \min \{\langle x^k, y \rangle - g^*(y) : y \in \partial h^*(x^k)\} \quad (2.51)$$

Les problèmes (2.50) et (2.51) sont des problèmes de minimisation concave. Ainsi DCA assure l'appartenance  $(x^\infty, y^\infty) \in P_l \times D_l$ .

Malgré leur apparente simplicité par rapport aux problèmes d'origine ( $P_{dc}$ ) et ( $D_{dc}$ ), ces problèmes restent difficiles à résoudre. En effet il s'agit de problèmes de minimisation concave comme nous l'avons déjà signalé plus haut, de plus le travail se fait sur  $\partial g^*(y^k)$  (resp.  $\partial h(x^k)$ ). Ces difficultés rendent en pratique l'utilisation de ce schéma presque impossible, excepté pour des cas bien particuliers pour lesquels la résolution de (2.50) et (2.51) est une tâche facile.

### 2.2.5 Existence et bornitude des suites générées par DCA

On dira que DCA est bien défini si on peut construire les suites  $\{x^k\}_k$  et  $\{y^k\}_k$  à partir d'un point arbitraire  $x^0 \in X$ .

**Lemme 2.6** [2] Les propositions suivantes sont équivalentes

- Les suites  $\{x^k\}_k$  et  $\{y^k\}_k$  sont bien définies
- $\text{dom}(\partial g) \subset \text{dom}(\partial h)$  et  $\text{dom}(\partial h^*) \subset \text{dom}(\partial g^*)$

Le lemme suivant établit les conditions pour lesquelles les suites générées par DCA sont bornées.

**Lemme 2.7** [2] Si  $(g - h)$  est coercive <sup>4</sup>. Alors on a

- a). La suite  $\{x^k\}_k$  est bornée,
- b). Si  $\{x^k\}_k \subset \text{int}(\text{dom}(h))$  alors la suite  $\{y^k\}_k$  est aussi bornée.

Par dualité, si  $(h^* - g^*)$  est coercive alors on a

- a). La suite  $\{y^k\}_k$  est bornée,
- b). Si  $\{y^k\}_k \subset \text{int}(\text{dom}(g^*))$  alors la suite  $\{x^k\}_k$  est aussi bornée.

Le théorème suivant établit la convergence de DCA simplifié.

1. On suppose les suites  $\{x^k\}_k$  et  $\{y^k\}_k$  bien définies. Alors on a

$$(g - h)(x^{k+1}) \leq (h^* - g^*)(y^k) - \delta_1 \leq (g - h)(x^k) - \delta_2, \quad (2.52)$$

avec

$$\begin{aligned} \delta_1 &= \max\left\{\frac{\rho_2}{2}\|dx^k\|^2, \frac{\rho_2^*}{2}\|dy^k\|^2\right\}, \\ \delta_2 &= \max\left\{\frac{\rho_1 + \rho_2}{2}\|dx^k\|^2, \frac{\rho_1^*}{2}\|dy^{k-1}\|^2 + \frac{\rho_2}{2}\|dx^k\|^2, \frac{\rho_1^*}{2}\|dy^{k-1}\|^2 + \frac{\rho_2^*}{2}\|dy^k\|^2\right\}. \end{aligned}$$

L'égalité  $(g - h)(x^{k+1}) = (g - h)(x^k)$  a lieu si et seulement si  $x^k \in \partial g^*(y^k)$ ,  $y^k \in \partial h(x^{k+1})$  et  $(\rho_1 + \rho_2)dx^k = \rho_1^*dy^{k-1} = \rho_2^*dy^k = 0$ .

Dans ce cas on a

- $(g - h)(x^{k+1}) = (h^* - g^*)(y^k)$  et  $x^k, x^{k+1}$  sont des points critiques de la fonction  $(g - h)$  tels que

$$y^k \in (\partial g(x^k) \cap \partial h(x^k)) \text{ et } y^k \in (\partial g(x^{k+1}) \cap \partial h(x^{k+1})). \quad (2.53)$$

- $y^k$  est un point critique de  $(h^* - g^*)$  tel que

$$[x^k, x^{k+1}] \subset (\partial g^*(y^k) \cap \partial h^*(y^k)).$$

- $x^{k+1} = x^k$  si  $\rho(g) + \rho(h) > 0$ ,  $y^k = y^{k-1}$  si  $\rho(g^*) > 0$  et  $y^k = y^{k+1}$  si  $\rho(h^*) > 0$ .

2. Si  $\alpha$  est fini alors les suites décroissantes  $\{(g - h)(x^k)\}_k$  et  $\{(h^* - g^*)(y^k)\}_k$  sont convergentes et ont la même limite  $\beta \geq \alpha$ .

---

<sup>4</sup>une fonction  $\psi$  est coercive si  $\lim_{\|x\| \rightarrow +\infty} \psi(x) = +\infty$ .

Si  $\rho(g) + \rho(h) > 0$ , alors  $\lim_{k \rightarrow +\infty} \{x^{k+1} - x^k\} = 0$ .

Si  $\rho(g^*) + \rho(h^*) > 0$ , alors  $\lim_{k \rightarrow +\infty} \{y^{k+1} - y^k\} = 0$ .

On a en plus

$$\lim_{k \rightarrow +\infty} \{g(x^k) + g^*(y^k) - \langle x^k, y^k \rangle\} = 0$$

$$\lim_{k \rightarrow +\infty} \{h(x^{k+1}) + h^*(y^k) - \langle x^{k+1}, y^k \rangle\} = 0$$

3. Si  $\alpha$  est fini et si les suites  $\{x^k\}_k$  et  $\{y^k\}_k$  sont bornées alors pour toutes valeur d'adhérence  $x^*$  de  $\{x^k\}_k$  (resp.  $y^*$  de  $\{y^k\}_k$ ) il existe une valeur d'adhérence  $y^*$  de  $\{y^k\}_k$  (resp.  $x^*$  de  $\{x^k\}_k$ ) telle que

(a)  $y^* \in (\partial g(x^*) \cap \partial h(x^*))$  et  $g(x^*) - h(x^*) = \beta$ ;

(b)  $x^* \in (\partial g^*(y^*) \cap \partial h^*(y^*))$  et  $h^*(y^*) - g^*(y^*) = \alpha$ .

# 3

## *Rappels sur les jeux.*

### **Introduction**

*Dans ce chapitre, on donne quelques définitions relatives à la théorie des jeux nécessaires pour la suite de notre mémoire.*

### **3.1 Définitions**

**Définition 3.1** *Un "jeu" est une situation où des joueurs sont conduits à faire des choix stratégiques parmi un certain nombre d'actions possibles, et dans un cadre défini à l'avance qui seront les "règles du jeu", le résultat de ces choix constituant une "issue du jeu", à laquelle est associé un "gain" (ou paiement), positif ou négatif, pour chacun des participants.*

**Remarque 3.1** *Un joueur peut être une personne, un groupe de personnes, une société, une région, un parti politique, un pays ou la Nature.*

**Définition 3.2** *"jeux coopératifs et non coopératifs" : un jeu est dit coopératif lorsque les joueurs peuvent communiquer librement entre eux et passer des accords (par ex. sous forme d'un contrat). Ils forment alors une coalition et recherchent l'intérêt général suivi d'un partage des gains entre tous les joueurs. Dans un jeu non coopératif, les joueurs (qui ne communiquent pas ou ne peuvent pas communiquer entre eux) agissent selon le principe de rationalité économique :*

chacun cherche à prendre les meilleures décisions pour lui-même (c'est à dire cherche à maximiser ses gains individuels).

**Définition 3.3** "jeux à somme nulle et non nulle" : un jeu est dit à "somme nulle" lorsque la somme des gains des joueurs est constante ou autrement dit : ce que l'un gagne est nécessairement perdu par un autre (échecs, poker...). Les jeux de société sont souvent des jeux à somme nulle mais les situations réelles sont souvent mieux décrites par les jeux non coopératifs à somme non nulle car certaines issues sont profitables pour tous, ou dommageables pour tous (vie politique, situations d'affaires...).

Les jeux à somme nulle sont parfois appelés "jeux antagonistes".

**Définition 3.4** "jeux compétitifs et non compétitifs" : un jeu non compétitif est à l'opposé d'un jeu compétitif tel que par définition, lorsque toute couple de stratégie (dans le cas d'un jeu à deux joueurs) est tel qu'il fait perdre ou gagner simultanément à tous les joueurs un gain donné (quand je perds quelque chose tu perds quelque chose, quand je gagne quelque chose tu gagnes aussi quelque chose).

## 3.2 Représentations

Il existe différentes manières de formaliser la théorie des jeux et de la décision et ce d'autant plus suivant le type de situations dont il s'agit. Ainsi, nous distinguons :

- 1. Les "formes extensives" qui sont des formes synoptiques (arbre, branche, feuille) utiles à une compréhension simple des stratégies possibles et où l'issue d'un jeu est assimilée à une feuille dans laquelle nous retrouvons le vecteur des gains (ou "payements") respectif des joueurs. Ce genre de représentation devient compliquée (longue à dessiner) lors de jeux répétitifs.
- 2. Les "formes normales" qui permettent de réduire considérablement la taille et le temps de représentation graphique d'un jeu sous forme d'un tableau (matrice) de gains (ou "payements").

Nous allons par la suite nous restreindre à la représentation d'un jeu sous forme normale.

### 3.2.1 Jeux sous forme normale (stratégique)

**Définition 3.5** Une stratégie pure du joueur  $i$  est un plan d'action qui prescrit une action de ce joueur pour chaque fois qu'il est susceptible de jouer. On note par  $X_i$  l'ensemble des stratégies pures du joueur  $i$  et par  $x_i$  une stratégie pure de ce joueur.

Ainsi, On appelle jeu sous forme normale à  $N$  joueurs le triplet

$$\langle \mathcal{F}, \{X_i\}_{i \in \mathcal{F}}, \{g_i\}_{i \in \mathcal{F}} \rangle, \quad (3.1)$$

où

$\mathcal{F} = \{1, 2, \dots, N\}$  est l'ensemble des joueurs,

$N \geq 2$  est le nombre de joueurs.

$X_i$  est l'ensemble des actions (stratégies) du joueur "i", ( $i \in \mathcal{F}$ ).

$g_i : X = \prod_{i=1}^N X_i \longrightarrow \mathbb{R}$  : fonction utilité ou de gain du joueur "i". L'utilité du joueur  $i$  dépend à la fois de son action et des actions des autres joueurs. Soit  $x = (x_1, \dots, x_n) \in X$ . L'utilité du joueur  $i$  sera noté  $g_i(x) = g_i(x_i; x_{-i})$  avec  $x_{-i} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$ .

### **Exemple : Dilemme du prisonnier**

Deux individus (Bonnie et Clyde) sont arrêtés par la police pour la complicité dans un vol à main armée et ils sont enfermés dans deux cellules séparées sans possibilité de communication.

Chaque individu est interrogé séparément et il a le choix entre nier d'avoir commis le vol ou avouer l'avoir commis avec son complice.

Nous avons donc un jeu non-coopératif avec  $N = 2$  joueurs,  $\mathcal{F} = \{1, 2\} = \{\text{Bonnie}, \text{Clyde}\}$ .

L'ensemble de stratégies de chaque joueur est  $X_1 = X_2 = \{\text{nier}, \text{avouer}\}$ . Il y a donc 4 résultats possibles du jeu

$$X = \{ x = (\text{nier}, \text{nier}), y = (\text{nier}, \text{avouer}), z = (\text{avouer}, \text{nier}), t = (\text{avouer}, \text{avouer}) \}.$$

Les gains des individus représentent leur situation qui résulte des années de prisons auxquelles ils sont condamnés en fonction de leurs aveux et ils sont négativement liés avec ces années.

- Si Bonnie et Clyde avouent tous les deux leur crime, ils sont condamnés à 8 ans de prison.
- S'ils le nient tous les deux, ils auront 1 année de prison du fait d'absence de preuves accablantes.
- Si l'un seul avoue, il est relâché en récompense de sa coopération et l'autre est condamné à 10 ans de prison.

Nous avons donc les gains (symétriques) suivants :

- $g_1(x) = g_2(x) = -1$ ,
- $g_1(y) = g_2(y) = -10$ ,
- $g_1(z) = g_2(z) = 0$ ,
- $g_1(t) = g_2(t) = -8$ .

		Clyde	
		Avouer	Nier
Bonnie	Avouer	(-1,-1)	(-10,0)
	Nier	(0,-10)	(-8,-8)

Nous pouvons alors représenter ce jeu sous forme normale, sous la forme d'un tableau :

Les stratégies de Bonnie sont représentées en lignes et celles de Clyde en colonnes. Les gains donnent d'abord celui du joueur qui est en ligne (Bonnie) et ensuite, celui du joueur en colonne (Clyde) en fonction des choix possibles de chaque joueur.

A quel type de solution ce type de situation peut-il déboucher ? La situation que nous pouvons prédire doit être une situation à partir de laquelle aucun des joueurs n'a intérêt à dévier. Cela traduit bien une idée d'équilibre. Mais on peut parfois prévoir le résultat d'un jeu sans se préoccuper directement de l'équilibre. C'est aussi le cas de ce jeu. Pour observer cela, regardons les choix possibles de Bonnie en fonction des stratégies de Clyde :

- ◊ Si Clyde choisit de nier, Bonnie obtient  $-1$  en niant et  $0$  en avouant ;
- Si Clyde choisit d'avouer, Bonnie obtient  $-10$  en niant et  $-8$  en avouant.

Donc quelque soit le choix de Clyde, Bonnie a intérêt à avouer. Ce type de stratégie optimale indépendamment des choix de l'autre joueur s'appelle une stratégie dominante. On a alors les définitions suivantes :

**Définition 3.6** Une stratégie  $x_i \in X_i$  est (strictement) dominée pour le joueur  $i$  s'il existe une stratégie  $x'_i$  telle que pour tous les profils  $x_{-i}$

$$g_i(x'_i, x_{-i}) > g_i(x_i, x_{-i})$$

**Définition 3.7** Une stratégie  $x_i$  est faiblement dominée pour le joueur  $i$  s'il existe une stratégie  $x'_i$  telle que pour tous les profils  $x_{-i}$

$$g_i(x'_i, x_{-i}) \geq g_i(x_i, x_{-i})$$

Donc, Bonnie a une stratégie dominante qui est avouer. On peut donc imaginer qu'elle va la choisir et dénoncer son complice par la même occasion.

Si l'on raisonne de la même manière pour Clyde, on observe aussi que avouer est une stratégie dominante pour lui. Donc on peut raisonnablement prévoir qu'il va avouer aussi.

Le résultat de ce jeu sera donc (avouer, avouer) qui conduit en fin de compte à la situation la pire pour les deux joueurs.

L'existence des stratégies dominantes nous simplifie beaucoup l'analyse des interactions. Mais peu de situations correspondent à ce type de stratégies fortes. Pour ces cas plus généraux, nous devons utiliser un concept d'équilibre plus riche.

### 3.3 Équilibre de Nash et fonction de meilleure réponse

**Définition 3.8** La fonction de meilleure réponse du joueur  $i$  est la fonction  $R_i$  qui associe à chaque combinaison de stratégies des autres joueurs  $x_{-i}$  les stratégies du joueur  $i$  qui maximise son utilité :

$$R_i(x_{-i}) = \{x_i \in X_i \quad t.q. \quad g_i(x_i; x_{-i}) \geq g_i(x'_i; x_{-i}) \quad \text{pour tout } x'_i \in X_i\}.$$

En d'autre terme

$$R_i(x_{-i}) = \{x_i \in X_i \quad t.q. \quad g_i(x_i; x_{-i}) = \max_{x'_i \in X_i} g_i(x'_i; x_{-i})\}.$$

**Définition 3.9** Un équilibre de Nash dans le jeu non coopératif est une situation pour laquelle aucun joueur n'a intérêt à dévier de sa stratégie d'équilibre étant donné que les autres joueurs ont adopté leurs stratégie d'équilibre

**Définition 3.10** Considérons le jeu noncoopératif(3.1).  $x^* = (x_i^*, x_{-i}^*)$  est dite équilibre de Nash si et seulement si

$$g_i(x_i^*) \geq g_i(x_i, x_{-i}^*), \quad \forall x_i \in X_i, \quad \forall i \in N. \quad (3.2)$$

**Définition 3.11** Un équilibre de Nash est un profil  $x^* = (x_i^*; x_{-i}^*)$  tel que la stratégie du joueur  $i$  est une meilleure réponse ; c'est à dire

$$x_i^* \in R_i(x_{-i}^*) \quad \text{pour tout } i \in N$$

**Remarque 3.2** Un jeu (en stratégies pures) peut avoir plusieurs équilibres de Nash, mais il peut aussi n'en avoir aucun

**Théorème 23** Théorème de Nash[47]  
supposons que dans le jeu (3.1) :

- (a)  $\forall i \in N$ ,  $X_i$  est un compact convexe de  $\mathbb{R}^n$ ,
  - (b)  $\forall i \in N$ ,  $g_i : X \rightarrow \mathbb{R}$  est continue,
  - (c)  $\forall i \in N$ ,  $\forall x_{-i} \in X_{-i}$  l'application  $x_i \mapsto g_i(x_i; x_{-i})$  est concave,
- alors : (3.1) admet une situation d'équilibre de Nash, c'est à dire :

$$\exists x^* \in X \text{ tel que } g_i(x^*) \geq g_i(x_i, x_{-i}^*), \quad \forall x_i \in X_i, \quad \forall i \in N.$$

**Définition 3.12** – Une règle de décision du joueur  $i$  est une correspondance

$$c_i : X_{-i} \rightarrow X_i$$

$$x_{-i} \mapsto c_i(x_{-i}) \subset X_i$$

- Supposons que chaque joueur ait choisi sa stratégie du jeu suivant la règle de décision  $c_i : X_{-i} \rightarrow X_i$ . Une situation du jeu  $x \in X$  est dite cohérente par rapport aux  $(c_i)_{i \in N}$  si  $x_i \in c_i(x_{-i})$ ,  $\forall i \in N$

Posons

$$c_i : X \rightarrow X$$

$$x \mapsto c(x) = \prod_{i \in N} c_i(x_{-i}),$$

alors  $x$  cohérente  $\iff x \in c(x)$ .

- $c$  est dite semi continue supérieurement si  $\forall \varepsilon > 0$ ,  $\exists v(x^*)$  tel que  $c(x) \subset c(x^*) + \varepsilon B$ , où  $B = \{y \in Y : \|y\| = 1\}$
- $c$  est dite semi continue inférieurement si  $\forall \{x_n\} \subset c$  convergeant vers  $x^*$ , et  $\forall y \in c(x^*)$ ,  $\exists \{y_n\} \in c(x_n)$  telle que  $\lim_{n \rightarrow \infty} y_n = y^*$ .
- $c$  est continue en  $x$  si elle est semi continue inférieurement et supérieurement en  $x$ .

**Théorème 24** Théorème de Kakutani[47]

Soit  $K \subset \mathbb{R}^n$  convexe et compact, et  $D$  une correspondance semi continue supérieurement tel que :  $\forall x \in K$   $D(x)$  convexe et fermée, alors  $\exists x^* \in K : x^* \in D(x^*)$ .

Si maintenant on définit la fonction  $R(x) = (R_1(x_{-1}), \dots, R_N(x_{-N}))$  des fonctions de meilleure réponse définies ci-haut qui est bien une correspondance, alors d'après la définition (3.11) et le théorème de Kakutani ci-dessus, si  $X$  est convexe et compact et si  $R$  semi continue supérieurement tel que  $\forall x \in X$   $R(x)$  convexe et fermée, alors  $\exists x^* \in X$  tel que  $x^*$  est un équilibre de Nash.

## 3.4 Jeux répétés (superjeux)

**Définition 3.13** *Un jeu répété est un jeu ordinaire réitéré plusieurs fois de suite. Par jeu ordinaire on entend un jeu statique (qui se joue une seule fois), appelé jeu constituant, dans lequel les joueurs choisissent simultanément leurs stratégies.*

**Définition 3.14** *Un jeu répété est la donnée de :*

- un jeu constituant  $\langle \mathcal{F}, \{X_i\}_{i \in \mathcal{F}}, \{g_i\}_{i \in \mathcal{F}} \rangle$ , défini comme dans (3.1),
- nombre de périodes  $T$ ,
- Un vecteur  $\Delta = (\delta_1, \dots, \delta_N)$  des coefficients d'actualisation des joueurs. Chaque  $\delta_i$  reflète à combien un joueur valorise son gain de la période courante dans la période suivante,  $0 < \delta_i \leq 1$  pour tout  $i \in \mathcal{F}$ . Dans les jeux de marchandage, ce coefficient est généralement identique pour tous les joueurs (les entreprises) et est proche de 1.

### 3.4.1 Les stratégies dans un jeu répété

Nous considérons le jeu répété dans lequel les joueurs font face à chaque période au jeu constituant défini comme précédemment. La répétition de ce jeu permet aux joueurs de conditionner leurs choix présents et futurs sur les choix passés. Cette interdépendance temporelle peut conduire à des solutions plus coopératives ou plus agressives que celles observées dans le jeu constituant. La répétition élargit donc l'ensemble des solutions.

Nous définissons  $x_i(t)$  comme étant l'action du joueur  $i$  à la date  $t$ .  $x(t) = (x_1(t), \dots, x_n(t)) \in X$  est alors le profil d'actions sélectionnés par les  $n$  joueurs à la date  $t$  et  $h_t = (x(0), x(1), \dots, x(t-1))$  l'histoire du jeu à cette même date. L'histoire du jeu correspond à l'ensemble des actions que les joueurs ont choisi entre la période initiale 0 et la période  $t-1$ . Nous avons alors  $h_t \in X^t$  où  $X^t = \prod_{t \text{ fois}} X$ .

dans un jeu répété, le profil d'action sélectionné par les joueurs à chaque période crée un nouveau sous-jeu qui est entièrement défini par l'histoire à cette date.

**Définition 3.15** *Dans un jeu répété, une stratégie pour le joueur  $i$  consiste en séquence de règles de décision, une par période. Cette stratégie est notée  $\sigma_i = (\sigma_{i(0)}, \sigma_{i(1)}, \sigma_{i(2)}, \dots, \sigma_{i(t)}, \dots)$  où  $\sigma_{i(t)}$  est la règle de décision du joueur  $i$  à la date  $t$ .*

La règle de décisions  $\sigma_{i(t)}$  est une application de  $X^t$  vers  $X_i$  qui spécifie pour chaque histoire possible du jeu  $h_t$  une action ou une conduite à tenir à la date  $t$ . Si  $\Sigma_{i(t)}$  est l'ensemble des règles de décisions du joueur  $i$  à la date  $t$ , l'ensemble des stratégies du joueur  $i$  dans le jeu répété est alors  $\Sigma_i = X_i \times \Sigma_{i(1)} \times \Sigma_{i(2)} \times \dots$ . Nous notons  $\sigma = (\sigma_1, \dots, \sigma_n)$  le profil de stratégies du jeu répété et  $\sigma_t =$

$(\sigma_{1(t)}, \dots, \sigma_{n(t)})$  le profil des règles de décisions à la période  $t$ . On a  $\sigma \in \Sigma$  où  $\Sigma = \prod \Sigma_i$ .

A ce stade, nous pouvons distinguer plusieurs classes de stratégies selon la place que l'histoire occupe dans les règles de décisions des joueurs.

### les stratégies en boucle ouverte

le joueur ne tient pas compte de l'histoire du jeu. En fait cela revient à dire que le joueur choisit initialement une séquence d'action (une par période) et qu'il applique ce programme quel que soit le comportement de ses rivaux. Le comportement d'un tel joueur peut être qualifié de myope. Toute la dimension dynamique du jeu répété est éliminée. Une stratégie en boucle ouverte s'écrit  $\sigma_i = \{(x_{i(t)})_{t=0,1,\dots}\}$  ou encore  $\sigma_i = (\sigma_{i(0)}, \sigma_{i(1)}, \dots, \sigma_{i(t)}, \dots)$  avec  $\sigma_{i(t)}[h_i] = \sigma_{i(t)}[h'_i]$  pour toutes les histoires  $h_i \neq h'_i$  dans  $X$ .

### La stratégie en boucle fermée

Les joueurs tiennent compte de l'histoire du jeu. Initialement, chaque joueur adopte une séquence de règles de décision, une règle par période. A la période  $t$ , il observe l'histoire du jeu  $h_t$  et joue l'action prescrite par sa règle de décision  $\sigma_{i(t)}$ . Nous pouvons alors avoir  $\sigma_{i(t)}[h_t] \neq \sigma_{i(t)}[h'_t]$  pour toutes les histoires  $h_t, h'_t \in X$ . Les stratégies en boucle fermée sont une représentation générale de l'ensemble des stratégies d'un jeu répété et notamment celles en boucle ouverte.

Dans un jeu répété, les règles de décision peuvent prendre des formes extrêmement complexes, puisqu'elles sont supposées spécifier la conduite à tenir dans n'importe quel sous-jeu et face à n'importe quelle histoire. Abreu [17, 18] a proposé une représentation des stratégies sous une forme plus simple. L'ensemble des stratégies d'un jeu répété comportant  $N$  joueurs sera réduit à seulement  $(N + 1)$  règles d'actions qui sont elles mêmes définies par  $(N + 1)$  sentiers d'actions. La détermination des équilibres est ainsi plus facile.

**Définition 3.16** Un sentier est une séquence infinie de profils d'actions  $\{x(t)\}_{t=0}^{\infty}$  où  $x(t)$  est le profil d'actions sélectionné par les joueurs à la date  $t$ .

Dans un jeu répété, les joueurs choisissent à chaque période leurs actions selon des règles de décisions définies préalablement. La séquence des profils d'actions sélectionnés au cours du jeu détermine un sentier unique dans le jeu complet. Ce sentier est une succession de déplacements des joueurs d'un point de décision à un autre et d'un sous-jeu à un autre.

### exemple de sentier

Soit  $x^* = (x_1^*, \dots, x_n^*)$  l'équilibre de Nash du jeu constituant. La séquence  $\{x^*, x^*, \dots\}$  ( $\{x(t) = x^* \forall t\}$ ) est un sentier stationnaire dans lequel les  $n$  joueurs répètent à chaque période l'équilibre de Nash du jeu constituant.

Un jeu infiniment répété admet une infinité de sentiers. Même dans un jeu répété fini, le nombre de sentiers fini devient rapidement élevé. Par exemple, dans le jeu du dilemme du prisonnier répété dix fois, il existe plus d'un million de sentiers possibles ( $4^{10}$  sentiers possibles).

**Définition 3.17** Soit  $\sigma = (\sigma(0), \sigma(1), \dots, \sigma(t), \dots) \in \Sigma$  un profil de stratégies où  $\sigma(t) = (\sigma_1(t), \dots, \sigma_n(t))$  sont les règles de décision des  $n$  joueurs à la date  $t$ . Le profil de stratégies  $\sigma$  génère un sentier noté  $Q[\sigma] \in X^\infty$  défini par  $Q[\sigma] = \{q[\sigma](t)\}_{t=0}^\infty$  telle que :

$$\begin{aligned} q[\sigma](0) &= x(0) = (x_1(0), \dots, x_n(0)) \\ q[\sigma](t) &= \sigma(t)(q[\sigma](0), q[\sigma](1), \dots, q[\sigma](t-1)) \quad \text{pour tout } t \geq 1. \end{aligned}$$

Un profil de stratégies spécifie un profil d'actions initial  $x(0) \in X$  que les joueurs choisissent à la période  $t = 0$ . Puis, à la date  $t = 1$ , ils choisissent un profil d'actions  $q[\sigma](1)$  conformément à leurs règles de décisions  $\sigma(1)$  et en se basant sur l'histoire du jeu (c'est à dire sur les choix observés à la date  $t = 0$ ). A la période  $t$ , l'histoire du jeu est  $h_t = (q[\sigma](0), q[\sigma](1), \dots, q[\sigma](t-1))$  et  $q[\sigma](t)$  est le profil d'actions spécifiées par les règles  $\sigma(t)$ .

La valeur présente des gains que génère le sentier  $Q[\sigma]$  est égale pour le joueur  $i$  à :

$$G_i(t) = \sum_{t=0}^\infty \delta^t g_i(q[\sigma](t)). \quad (3.3)$$

où  $\delta$  est le facteur d'actualisation commun à tous les joueurs.

Nous pouvons aussi noter cette valeur  $G_i(Q[\sigma])$ .  $G_i$  est la fonction gain du jeu répété avec actualisation.

### 3.4.2 Notion d'équilibre dans le jeu répété

On considère les restrictions suivantes pour le jeu constituant :

**Hypothèse1** : le nombre de joueurs est fini et égal à  $N$ .

**Hypothèse2** : l'ensemble des stratégies  $X_i$  est compact et convexe.

Les deux hypothèses suivantes concernent les fonctions gains des joueurs

**Hypothèse3** : la fonction de gain du joueur  $i$ ,  $g_i(x) \in \mathbb{R}$ , est définie, continue et bornée pour tout  $x \in X$  et pour tout  $i \in N$ . Notons que  $g_i$  est une application de  $X$  vers  $\mathbb{R}$

**Hypothèse 4 :**  $g_i(t_i, x_{-i})$  est quasi concave en  $t_i$  pour tout  $x_{-i} \in X$  et pour tout  $i \in N$ .

Le vecteur des fonctions individuelles de gains s'écrit  $g = (g_1, \dots, g_n) \in \mathbb{R}^N$ . Le jeu constituant caractérisé par l'ensemble des joueurs  $N$ , l'ensemble des stratégies  $X$  et le vecteur de gains  $g$  et qui satisfait les hypothèses 1 à 4 sera alors noté  $\Gamma = (\mathcal{F}, X, g)$ .

Nous supposons que le jeu  $\Gamma = (\mathcal{F}, X, g)$  est en information complète. Chaque joueur connaît l'ensemble des stratégies et les fonctions de gains des autres joueurs. De plus, cette information complète sur  $\mathcal{F}$ ,  $X$  et  $g$  est une connaissance commune : chaque joueur sait que les autres joueurs connaissent  $\mathcal{F}$ ,  $X$  et  $g$ , de même chaque joueur sait que les autres joueurs savent qu'il connaît  $\mathcal{F}$ ,  $X$  et  $g$  et ainsi de suite. Sous les hypothèses 1 à 4 citées précédemment (continuité et quasi-concavité des fonctions gains d'une part, convexité et compacité des ensemble de stratégies d'autre part), l'existence d'un équilibre de Nash dans le jeu constituant est assurée. Le théorème de l'existence d'un équilibre de Nash est le résultat central dans la théorie des jeux statiques ; dans la théorie des jeux répétés, l'équivalent de ce théorème est le théorème du Folk qui sous les mêmes hypothèses garantit l'existence d'un équilibre de Nash.

Si on note le jeu infiniment répété avec actualisation par  $\Gamma^\infty = (\Sigma, G, \mathcal{F}, \delta)$  où  $G = (G_1, \dots, G_n)$  est le vecteurs des gains des joueurs dans le jeu répété.

**Définition 3.18** Dans le jeu  $\Gamma^\infty = (\Sigma, G, N, \delta)$ , le profil de stratégies  $\sigma$  est un équilibre de Nash si et seulement si pour tout  $i \in N$  et tout  $\sigma_i \in \Sigma_i$  on :

$$G_i(\sigma) \geq G_i(\sigma_i, \sigma_{-i}). \quad (3.4)$$

Dans un jeu répété, l'équilibre de Nash est un concept insatisfaisant car il n'est pas assez exigeant sur les profils d'actions prescrits dans les sous-jeux hors équilibre. C'est pourquoi les équilibres de Nash d'un jeu dynamique sont toujours soumis au critère de sous-jeu parfait.

### 3.4.3 Crédibilité et sous-jeu parfait

La notion d'équilibre de sous-jeu parfait à été définie par Selten [46]. Elle permet d'éliminer les équilibres de Nash qui spécifient sur certains sous-jeux hors équilibre des actions non crédibles.

Un profil d'actions est crédible si chaque joueur a intérêt à individuellement à s'y conformer étant donné que les autres joueurs font de même. La crédibilité dans un jeu répété est étroitement liée aux punitions que les joueurs sont censés mettre en oeuvre lorsque l'un d'eux dévie du sentier d'équilibre. Ces punitions sont crédibles s'il est bien dans l'intérêt des joueurs de les appliquer lorsqu'une déviation est observée.

De manière formelle, un profil de stratégies  $\sigma$  est un équilibre parfait si c'est un équilibre de Nash sur le sentier d'équilibre, mais aussi sur tous les sentiers hors équilibre ou sur tous les sous-jeux possibles. Un sous-jeu étant entièrement défini (ou localisé) par une date  $t$  et une histoire  $h_t$ , nous appelons  $\sigma^{h_t} = (\sigma_i^{h_t}; \sigma_{-i}^{h_t})$  le profil de stratégies induit par  $\sigma$  sur le sous-jeu qui commence en  $t$  après l'histoire  $h_t$  et  $\Sigma_i^{h_t}$  l'ensemble des stratégies du joueur  $i$  dans ce même sous-jeu.

**Définition 3.19** Le profil de stratégies sigma est un équilibre parfait pour le jeu répété  $\Gamma^\infty = (\Sigma, G, N, \delta)$  si et seulement si :

1.  $\sigma$  est un équilibre de Nash sur l'ensemble du jeu  $\Gamma^\infty$  :

$$G_i(\sigma) \geq G_i(\sigma_i, \sigma_{-i}) \text{ pour tout } i \in N \text{ et tout } \sigma_i \in \Sigma_i.$$

2.  $\sigma^{h_t}$  est un équilibre de Nash pour tout  $t$  et toute histoire  $h_t$  :

$$G_i(\sigma^{h_t}) \geq G_i(\sigma_i, \sigma_{-i}^{h_t}) \text{ pour tout } i \in N, \text{ tout } t = 1, \dots, h_t \text{ et tout } \sigma_i \in \Sigma_i^{h_t}.$$

Un profil de stratégies d'équilibre parfait du jeu  $\Gamma^\infty$  peut être pensé en terme de sentiers et de règles spécifiant un sentier initial  $Q^0$ , un sentier  $Q^1$  en cas de déviation du sentier  $Q^0$ , un sentier  $Q^2$  en cas de déviation du sentier  $Q^1$ , ..., un sentier  $Q^t$  en cas de déviation du sentier en cours  $Q^{t-1}$ . On mesure la complexité que peut prendre un tel profil si le sentier prescrit après la  $t^{\text{ième}}$  déviation dépend de l'histoire entière du jeu, de la date de cette déviation, de son ampleur ou de l'identité du coupable. La recherche d'un équilibre parfait nécessite de prendre en compte toutes les histoires possibles et toutes les séquences de déviations finies ou infinies. Cette tâche dépasse de loin les capacités de calcul du théoricien des jeux et encore plus celles des joueurs en présence du jeu.

Le mérite d'Abreu [17, 18] était de proposer des profils de stratégies simples qui facilitent la détermination des équilibres parfaits et qui fournissent une représentation assez juste des comportements stratégiques des joueurs dans le jeu répété.

### 3.4.4 Profils de stratégies simples

L'idée d'Abreu est de restreindre le nombre de sentiers possible du jeu à seulement  $N + 1$  sentiers. Les joueurs se déplacent sur ces sentiers selon les règles explicitées dans la définition suivante :

**Définition 3.20** Un profil de stratégies simples (PSS) est un ensemble de règles de décision entièrement définies par  $N + 1$  sentiers  $(Q^0, \dots, Q^N)$  et qui vérifient les conditions suivantes :

1. De commencer sur le sentier  $Q^0$  à la date  $t = 0$ .

2. De suivre  $Q^0$  aux dates suivantes si aucun joueur n'a dévié unilatéralement de ce sentier.
3. De passer sur le sentier  $Q^j$  dès que le joueur  $j$  dévie unilatéralement de sentier en cours  $Q^i$  ( $i = 0, 1, \dots, N$ ).
4. de rester sur le sentier  $Q^i$  si  $Q^i$  est le sentier en cours et si aucune déviation n'est observée ou si deux joueurs ou plus ont dévié simultanément de  $Q^i$ .

$Q^0$  est le sentier d'équilibre que les joueurs veulent soutenir initialement et  $Q^i$  est le sentier de punition qui s'impose à tous les joueurs lorsque le joueur  $i$  dévie. Les règles de décision d'un PSS sont indépendantes de l'histoire passée, mis à part l'identité du dernier déviant. Elles sont fondées simplement sur la comparaison entre les actions prescrites sur le sentier en cours et les actions réellement sélectionnées. Si à la période précédente tous les joueurs se sont conformés aux actions prescrites, les joueurs restent sur le sentier en cours. Si un des joueurs a dévié du sentier en cours, la règle suggère de démarrer (ou de redémarrer) la punition à l'encontre de ce joueur.

Les règles de décision proposées par Abreu sont symétriques ( $\sigma_i(t) = \sigma_j(t)$  pour tout  $i \neq j$ ) et stationnaire à partir de la date  $t = 1$  ( $\sigma_i(t) = \sigma_i(1)$  pour tout  $t$ ). L'ensemble d'information pertinent pour mettre en oeuvre ces règles se limite au sentier en cours et au profil d'actions observé à la date ou période précédente.

De manière formelle, Si le sentier  $Q^i$  est défini par  $Q^i = \{q^i(\tau)\}_{\tau=0}^{\infty}$  pour tout  $i = 0, \dots, n$ , alors un PSS  $\sigma = (\sigma_0, \dots, \sigma_t, \dots)$  caractérisé par les  $N + 1$  sentiers  $(Q^0, \dots, Q^N)$  satisfait :

- (i) En  $t = 0$ ,  $\sigma(0) = q^0(0)$ .
- (ii) En  $t \geq 1$ , si  $\sigma(t-1)[h_{t-1}] = q^i(\tau-1)$  et si  $x(t-1) = q^i(\tau-1)$  pour  $i = 0, \dots, n$  et  $\tau \leq t$ , alors  $\sigma(t)[h_{t-1}, q^i(\tau-1)] = q^i(\tau)$  (règle de décision en l'absence de déviation à la période précédente).
- (iii) En  $t \geq 1$ , si  $\sigma(t-1)[h_{t-1}] = q^i(\tau-1)$  et si  $\exists j, k \in \mathcal{F}$ ,  $j \neq k$  tels que  $x_j(t-1) \neq q_j^i(\tau-1)$  et  $x_k(t-1) \neq q_k^i(\tau-1)$  alors  $\sigma(t)[h_{t-1}, x(t-1)] = q^i(\tau)$  (règle de décision en cas de déviations multiple à la période précédente)
- (iv) Si  $\exists i$  tel que  $x_i(t-1) \neq \sigma(t-1)[h_{t-1}]$  et  $x_i(t-1) = \sigma(t-1)[h_{t-1}]$  pour tout  $j \neq i$  alors  $\sigma(t)[h_{t-1}, x(t)] = q^i(0)$  (règle de décision en cas de déviation unilatérale à la période précédente)

En ii), il est dit que si à la date  $t - 1$ , les joueurs sont depuis  $(\tau - 1)$  périodes sur le sentier  $Q^i$  (ils ont donc déjà joué  $q(0), q(1), \dots, q(\tau - 1)$ ) et que chaque joueur se conforme à la  $\tau^{\text{ième}}$  action de ce sentier en cours, l'ensemble des histoires à la date  $t$  se réduit à  $(N + 1)$  histoires caractéristiques. La première histoire correspond à la situation dans laquelle personne n'a dévié ou plus de deux joueurs ont dévié du sentier en cours à la date précédente. La deuxième histoire correspond à une

déviations du joueur 1 dans la période précédente, ..., la  $(N + 1)^{\text{ième}}$  histoire à une déviation du joueur  $N$  à la période précédente.

Le PSS  $\sigma(Q^0, \dots, Q^N)$  défini par  $(N + 1)$  sentiers et  $(N + 1)$  histoires caractéristiques permet de dériver  $N$  autres profils de stratégies simples. Pour tout  $i \in F$ , soit  $\sigma^i(Q^1, \dots, Q^N) = \sigma(Q^i, Q^1, \dots, Q^N)$  le profil de stratégies simples induit par une déviation du joueur  $i$  à la période précédente. Il est identique au PSS  $\sigma(Q^0, \dots, Q^N)$  sauf dans le sentier qu'il prescrit initialement ( $Q^i$  au lieu de  $Q^0$ ). Un PSS  $\sigma(Q^0, \dots, Q^N)$  est alors parfait si ce profil et les  $N$  profils induits sont des équilibres de Nash. Nous devons donc vérifier que pour tout  $i$  et tout  $\sigma^i(Q^1, \dots, Q^N)$ , aucun joueur ne peut obtenir un meilleur gain en déviant de sa stratégie. Cette tâche est facilitée par un résultat central de la théorie des jeux répétés selon lequel si aucune déviation d'une période n'est profitable alors aucune séquence finie ou infinie de déviations n'est profitable.

### 3.4.5 Le principe de déviation unique

Considérons le gain et le coût d'une déviation d'une période. Une telle déviation consiste pour un joueur à ne pas respecter le sentier en cours puis à se conformer à sa punition. Si le sentier en cours est  $Q^j$  ( $j = 0, \dots, N$ ) et ce depuis  $\tau$  périodes, le joueur  $i$  obtient un gain actualisé à  $G_i(Q^j; \tau) = \sum_{x=0}^{\infty} \delta^x g_i(q^j(\tau + x))$  en ne déviant pas.

En revanche, s'il dévie pendant une période du profil d'actions  $q^j(\tau)$ , puis se conforme dès la période suivante à sa stratégie d'équilibre son gain actualisé est au plus égal à  $\max_{t_i} g_i(t_i; q_{-i}^j(\tau)) + \delta G_i(Q^i)$ .

En effet après la déviation, le jeu s'engage sur le sentier  $Q^i$ , destiné à punir le joueur  $i$ . Cette punition a une valeur présente de  $G_i(Q^i) = \sum_{x=0}^{\infty} \delta^x g_i(q^i(x))$  pour le joueur  $i$ .

Posons  $g_i^d(q) = \max_{t_i} g_i(t_i; q_{-i})$  le gain maximum que peut obtenir le joueur  $i$  lorsque les autres joueurs choisissent  $q_{-i}$ . Étant donné le sentier  $Q^j$  en cours depuis  $\tau$  périodes, aucune déviation d'une période n'est profitable si pour tout  $i$  on a :

$$G_i(Q^j; \tau) \geq g_i^d(q^j(\tau)) + \delta G_i(Q^i). \quad (3.5)$$

Après réorganisation, cette inégalité peut s'écrire sous la forme suivante :

$$g_i^d(q^j(\tau)) - g_i(q^j(\tau)) \leq \delta [G_i(Q^j; \tau + 1) - G_i(Q^i)], \quad (3.6)$$

avec  $G_i(Q^j; \tau + 1) = \sum_{x=0}^{\infty} \delta^x g_i(q^j(\tau + x + 1))$ .

Le terme de gauche dans l'inégalité (3.6) est le gain net d'une déviation d'une période et le terme de droite le coût d'opportunité de cette déviation. Si ce coût

est supérieur au gain attendu de la déviation pour tous les joueurs et pour tous les sentiers, alors aucune déviation d'une période n'est profitable.

**Proposition 21** [18]

Le profil de stratégies simples  $\sigma(Q^0, \dots, Q^N)$  est un équilibre de sous-jeu parfait si et seulement si pour tout  $i \in \mathcal{F}$  et  $\tau = 0, 1, \dots$  :

$$g_i^d(q^j(\tau)) - g_i(q^j(\tau)) \leq \delta[G_i(Q^j; \tau + 1) - G_i(Q^j)], \quad (3.7)$$

Une condition nécessaire et suffisante pour avoir un équilibre parfait en stratégies simples est que les déviations d'une période soient non profitables pour les  $N$  joueurs le long des  $N + 1$  sentiers possibles.

Cette proposition est fondamentale en théorie des jeux répétés. Sans ce résultat, toute caractérisation des équilibres parfaits serait quasi impossible.

Nous définissons  $\Sigma^P \subset \Sigma$  comme l'ensemble des profils de stratégies d'équilibres parfaits du jeu infiniment répété  $\Gamma^\infty$ .

$M^P = \{Q[\sigma] / \sigma \in \Sigma^P\}$  est alors l'ensemble des sentiers générés par des profils de stratégies d'équilibre parfaits ( $M^P \subset X^\infty$ ). Si  $Q \in M^P$ ,  $Q$  est alors appelé un sentier d'équilibre parfait.

**Lemme 3.1** [18]

Si  $\Gamma^\infty$  satisfait les hypothèses 1 à 4 alors  $\Sigma^P$  est non vide.

Suivant un facteur d'actualisation donné, la caractérisation de l'ensemble des équilibres parfaits  $\Sigma^P$  ou de l'ensemble des sentiers d'équilibres parfaits  $M^P$  passe par la recherche des punitions les plus sévères possibles. En effet, une fois connues les pires punitions pour chacun des joueurs, on peut déterminer l'ensemble des sentiers d'équilibres  $Q^0$  que les joueurs peuvent suivre sans être incités à dévier.

**Définition 3.21** Un code pénal simple  $(Q^1, \dots, Q^n)$  est un  $n$ -vecteur de profils de stratégies simples  $\{\sigma^1(Q^1, \dots, Q^n), \dots, \sigma^n(Q^1, \dots, Q^n)\}$  où le profil simple  $\sigma^i$  décrit la conduite à tenir pour punir le joueur  $i$ .

**Définition 3.22** Un code pénal optimal est un  $n$ -vecteur de profils de stratégies  $\{\underline{\sigma}^1, \dots, \underline{\sigma}^n\}$  tel que :

$$G_i[\underline{\sigma}^i] = \min\{G_i[\sigma] / \sigma \in \Sigma^P\} \text{ pour tout } i = 1, \dots, n. \quad (3.8)$$

Le profil  $\underline{\sigma}^i$  est le pire équilibre parfait en terme de gains pour le joueur  $i$ . Il prescrit la punition la plus sévère que peut se voir infliger le joueur  $i$  de manière crédible. Notons  $\underline{G}_i = G_i[\underline{\sigma}^i]$ . A la suite d'une déviation, aucun joueur ne peut être tenu à un gain actualisé inférieur à  $\underline{G}_i$ .

Il est claire que l'équilibre Nash est un code optimale. ainsi la pire punition à infliger à un déviant consiste à sélectionnant la stratégie de Nash.

# 4

## *Concurrence oligopolistique et modélisation.*

### **Introduction**

*Dans ce chapitre, on discutera des principaux modèles de la concurrence oligopolistique dans un marché d'un bien homogène. On donnera une brève synthèse des modèles basés sur la concurrence en quantité, pour ensuite passer à la concurrence en prix et on terminera le chapitre par l'élaboration d'un modèle pour la sélection du vecteur prix réalisant l'optimum pour les entreprises dans le cas de possibilité de collusion (sorte de coopération pour réduire la concurrence et faire plus de gains).*

### **4.1 Éléments de microéconomie**

*La micro-économie étudie le comportement des agents économiques individuels et l'agrégation de leurs actions dans différents cadres institutionnels. Cette définition sommaire introduit quatre catégories de base :*

- L'agent individuel qui est traditionnellement un consommateur ou une entreprise.*
- Le comportement de l'agent, qui est guidé par un objectif consistant en la maximisation de l'utilité pour le consommateur et la maximisation du profit pour l'entreprise.*

- Le cadre institutionnel qui décrit les options dont disposent chaque agent.
- La méthode d'analyse de l'agrégation des comportements des différents acteurs dans un cadre donné par le biais de la notion d'équilibre.

### 4.1.1 Propriétés d'un marché concurrentiel (Concurrence parfaite)

On distingue plusieurs propriétés :

- ◇ **L'homogénéité du produit** : Les biens sont parfaitement identiques, ce qui fait comme conséquence :
  - Chaque consommateur est prêt à acheter le bien chez n'importe quel producteur.
  - Aucun agent ne peut imposer son prix (de vente ou d'achat) sur le marché.
- ◇ **La libre entrée** : ce qui a comme conséquence le fait que des profits positifs attirent de nouvelles firmes (nouveaux concurrents).
- ◇ **La transparence** : tous les agents sont parfaitement informés sur les prix auxquels s'effectuent les transactions. Comme conséquence, les transactions s'effectuent à un prix unique : le prix du marché.

Comme conséquence à ces propriétés, la détermination du prix du marché est relatif au comportements des agents.

### 4.1.2 Notions d'offre et demande globale et équilibre de marché

Considérons un marché concurrentiel de  $n$  consommateurs et  $N$  firmes.

#### Demande globale[33]

La demande individuelle du consommateur  $i$  est notée :  $x^i(p)$ ,  $i = 1 \dots n$ .

À un prix donné  $p$ , la quantité totale demandée sur le marché est égale à la somme des quantités demandées par chaque consommateur :

$$D(p) = \sum_{i=1}^n x^i(p), \quad D'(p) \leq 0.$$

La supposition que  $D'(p) \leq 0$  est due au fait que si le prix augmente, la demande du consommateur diminue, ce qui fait que  $D(p)$  est une fonction décroissante en  $p$ .

### Offre globale[33]

Pour un prix  $p$  et une offre de chaque firme notée  $q_j = F(p)$ ,  $j = 1 \dots N$ ., les quantités totales offertes sur le marché sont alors données par l'offre globale :

$$O(p) = \sum_{j=1}^N q_j, \quad O'(p) \geq 0.$$

De même, la supposition  $O'(p) \geq 0$  revient au fait que l'offre augmentera sûrement avec l'augmentation du prix, ce qui fait d'elle une fonction croissante en  $p$ .

### 4.1.3 Notion de l'équilibre du marché

L'équilibre d'un marché concurrentiel résulte de la confrontation de l'offre et de la demande (figure 4.1). Cet équilibre se définit par :

- un prix  $p^*$  pour le bien,
- une liste des firmes actives choisies à partir de la liste de toutes les firmes potentiellement actives ; pour chaque firme, un plan de production tel que
  - chaque firme maximise son profit en prenant le prix  $p^*$  comme une donnée,
  - pour chaque firme active, ce profit maximal est non-négatif,
  - chaque firme inactive ferait au mieux des profits non-positifs si elle décidait de devenir active,
  - l'offre totale des firmes actives, qui est la somme de leur plan de production au prix  $p^*$ , est exactement égale à la demande de marché à ce prix. c'est-à-dire que  $D(p^*) = O(p^*) = Q^*$ .

À l'équilibre, il n'y aura aucun demandeur dans l'impossibilité de se procurer la quantité qu'il veut, et aucun offreur ne se trouvera dans l'impossibilité de vendre ce qu'il entend vendre : personne ne sera "insatisfait" relativement à ses actions.

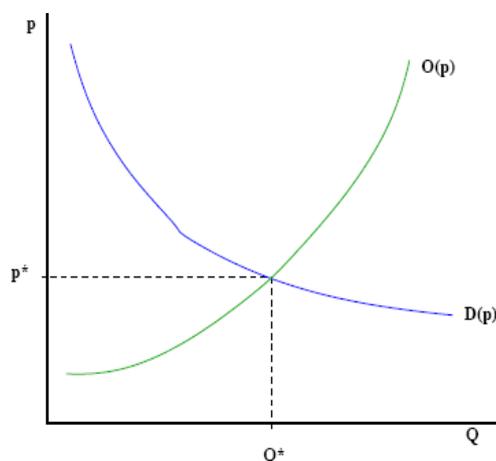


FIG. 4.1 – L'équilibre du marché concurrentiel

## 4.2 Cournot et la concurrence en quantité

Considérons un marché où  $N$  firmes produisent et vendent un bien homogène. La production, supposée égale à sa vente, de la  $i^{\text{ème}}$  firme est notée  $q_i$ .  $Q = \sum_{i \in \mathbb{F}} q_i$  représente la production industrielle, c-à-d la production totale de toutes les firmes, où  $\mathbb{F} = \{1, \dots, N\}$ .

Soit  $p$  le prix du marché, alors la fonction demande inverse de ce marché est  $p = f(Q)$ .

On fait l'hypothèse que  $f$  est deux fois continûment différentiable, décroissante et coupe les deux axes des abscisses et des ordonnées.

Chaque firme a une fonction de coût production,  $c_i(q_i)$ , supposée deux fois continûment différentiable, non négative et convexe dont la première dérivée est positive.

Ainsi, la fonction profit de la  $i^{\text{ème}}$  firme est définie en fonction de  $q = (q_1, \dots, q_N)$  comme suit :

$$\Pi_i(q) = q_i f(Q) - c_i(q_i), \quad i = 1, \dots, N. \quad (4.1)$$

L'objectif de la  $i^{\text{ème}}$  firme est la maximisation de son profit étant donnée la production  $q_{-i} = (q_1, \dots, q_{i-1}, q_{i+1}, \dots, q_N)$  de ses rivaux, c'est à dire :

$$\max_{q_i} \Pi_i(q_i, q_{-i}). \quad (4.2)$$

La condition d'optimalité d'ordre 1 pour le problème (4.2) serait :

$$\frac{\partial \Pi_i}{\partial q_i} = f(Q) + q_i \frac{\partial f}{\partial q_i}(Q) - c'_i(q_i) = 0. \quad (4.3)$$

La condition du deuxième ordre est

$$\frac{\partial^2 \Pi_i}{\partial q_i^2} < 0. \quad (4.4)$$

**Application** Pour un peu apprécier la solution du problème ci-dessus on considère l'application suivante :

– La fonction de coût de la firme  $i$  est supposée de la forme

$$c_i(q_i) = c_i \cdot q_i, \quad i = 1, \dots, N,$$

– La fonction demande inverse sera de la forme

$$f(Q) = \alpha - \beta \cdot Q, \quad (4.5)$$

où  $\alpha$  et  $\beta$  sont des nombres positifs.

Notons que de telles fonctions sont les plus utilisées dans les applications microéconomiques.

La fonction profit devient dans ce cas

$$\Pi_i(q) = q_i \cdot \alpha - q_i \cdot \beta \cdot Q - c_i \cdot q_i, \quad i = 1, \dots, N. \quad (4.6)$$

La condition d'optimalité du premier ordre s'écrit donc

$$\frac{\partial \Pi_i}{\partial q_i} = \alpha - 2\beta q_i - \beta \sum_{j=1, j \neq i}^N q_j - c_i = 0,$$

ce qui donne

$$q_i^* = \frac{\alpha - c_i - \beta \sum_{j=1, j \neq i}^N q_j}{2\beta}.$$

On remarque que  $q_i^*$  est fonction des quantités des autres firmes rivales à la firme

$i$ , c'est à dire  $q_i^* = R_i(q_{-i})$  où  $R_i(q_{-i}) = \frac{\alpha - c_i - \beta \sum_{j=1, j \neq i}^N q_j}{2\beta}$ .

Cette fonction fut appelée par Cournot la fonction de réaction de la firme  $i$ . Elle nous donne la meilleure réaction pour la firme  $i$  pour chaque niveau de production concurrent, c'est pour ça qu'on l'appelle aussi fonction de meilleure de réaction.

On remarque aussi que la condition du second ordre donne

$$\frac{\partial^2 \Pi_i}{\partial q_i^2} = -2\beta < 0.$$

Étant donné qu'il existe  $N$  firmes sur le marché, on distinguera  $N$  fonctions de réaction. la solution optimale relativement à tous les problèmes des firmes est un vecteur  $R(q)$  des fonctions  $R_i(q_{-i})$ , c'est à dire  $R(q) = (R_1(q_{-1}), \dots, R_N(q_{-N}))$

L'équilibre de ce marché doit être une situation telle qu'une fois atteinte, aucune firme ne doit avoir envie de s'éloigner de cet état ; aucune firme ne doit avoir la possibilité d'améliorer son profit en produisant différemment étant donné que les autres firmes produisent leurs quantités d'équilibre.

Soit  $q^c$  un tel vecteur. Nous devons avoir dans ce cas  $q_i^c = R_i(q_{-i}^c)$  : la quantité qui maximise le profit de la firme  $i$  étant donnée la production d'équilibre rivale, pour tout  $i = 1, \dots, N$ . Un tel vecteur  $q^c$  est dit équilibre de Cournot. Ce dernier représente le point d'intersection entre les différentes courbes des fonctions  $q_i = R_i(q_{-i})$  ; c'est donc la solution du système d'équations  $q_i^c = R_i(q_{-i}^c)$ ,  $i = 1, \dots, N$ . C'est un point fixe de la fonction vectorielle  $R(q)$  ; c'est à dire  $q^c = R(q^c)$ , ce qui définit bien, d'après le théorème du point fixe, un équilibre de Nash dans un jeu non coopératif.

Dans notre cas,  $q^c$  vérifie le système d'équations suivant :

$$q_i^c = \frac{\alpha - c_i - \beta \sum_{j=1, j \neq i}^N q_j^c}{2\beta}$$

Une première caractéristique de l'équilibre est qu'aucune firme ne peut améliorer son profit en déviant de sa stratégie d'équilibre étant donné que les autres firmes adoptent leurs stratégies d'équilibre. Une deuxième propriété de l'équilibre de Cournot est donné dans le théorème suivant :

**Théorème 25** [22] Un oligopole satisfaisant les hypothèses suivantes admet un équilibre de Cournot unique :

1. **hypothèse 1** : La fonction demande inverse du marché,  $f(Q)$ , est définie et continue pour tout  $Q \geq 0$ . De plus,  $\exists \bar{Q}$  telle que pour tout  $Q \geq \bar{Q}$ ,  $f(Q) = 0$  et pour tout  $Q < \bar{Q}$   $f(Q) > 0$ . Finalement,  $f(0) = \bar{p} < \infty$ , et  $\forall 0 < Q < \bar{Q}$ ,  $f$  admet une dérivée première négative et dérivée seconde continue.

2. **hypothèse 2** :

(a) La fonction coût de la  $i^{\text{ème}}$  firme,  $c_i(q_i)$ , est définie et continue pour  $q_i \geq 0$  avec  $c_i(0) \geq 0$ ,

(b)  $c_i(q_i)$  est strictement croissante, c'est à dire :  $C_i'(q_i) > 0$ ,  $\forall q_i \geq 0 \forall i \in \mathbb{F}$

(c)  $C_i''(q_i)$  existe et continue pour tout  $q_i \geq 0$

3. **hypothèse 3** : Pour tout  $q_i > 0$  et  $Q < \bar{Q}$ , on a  $f'(Q) + q_i f''(Q) < 0$  et  $f'(Q) - c_i''(Q) < 0$ ,  $\forall i = 1, \dots, N$

Une dernière propriété de l'équilibre de Cournot est donnée par le théorème suivant :

**Théorème 26** [22] Si  $q^c > 0$ , alors  $\exists q^* \geq 0$  telle que  $\Pi_i(q^*) > \Pi_i(q^c)$ .

Ce théorème établit que si l'équilibre de Cournot est un point intérieur à l'ensemble des stratégies (FIG.4.2), alors il ne donne pas des Profits Paréto optimaux. Les conditions du théorème assure que l'équation (4.3) soit satisfaite, ce qui donne

$$\frac{\partial \Pi_i}{\partial q_j} = q_i^c \frac{\partial f(Q^c)}{\partial q_j} < 0, \quad \forall i, j \quad i \neq j$$

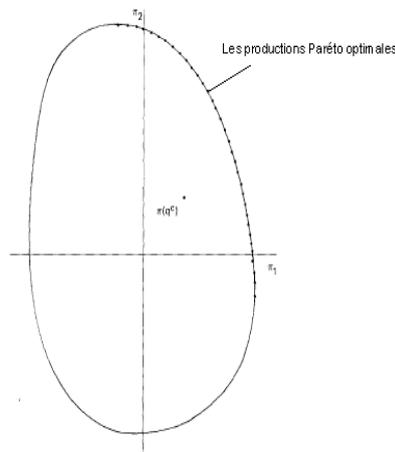


FIG. 4.2 – Productions Pareto Optimales

Ainsi, il est clairement possible de croître tous les  $q_i$  simultanément d'un petit montant  $\varepsilon$  pour lesquels les profits des firmes augmentent. Dans le langage de la théorie des jeux non coopératifs, si le gain de chaque joueur est une fonction continûment différentiable relativement aux stratégies des joueurs, l'équilibre non coopératif est un point intérieur de l'ensemble des stratégies, et le jacobien du système des gains évalué à l'équilibre est non singulier, alors l'équilibre est non Pareto optimal, la raison pour laquelle les praticiens regardent ce dernier comme désagréable.

### Bertrand et la concurrence en prix

En révisant les travaux de Cournot, Bertrand cite deux défauts : premièrement, les entreprises, à cause de la non optimalité de l'équilibre se feront certainement la collusion pour faire de meilleurs profits et deuxièmement, même si les firmes ne devraient pas coopérer, elles devront faire des choix de prix en plus des choix des quantités de production.

Le deuxième point soulevé est très important, car il est très difficile de concevoir la manière dont devront être déterminés les prix, si ces derniers ne sont pas fixés par les firmes.

Bertrand, sous l'hypothèse que les produits sont homogènes, pointe sur le fait que les consommateurs feront leurs achats chez les firmes qui vont offrir le coût le plus faible. Il stipule qu'il ne peut pas y avoir d'équilibre à n'importe quel prix en dessus de zéro. c'est seulement le cas où tous les prix sont nuls qu'aucune firme n'a intérêt à altérer son choix de prix. Par contre, si elles décide de choisir des prix positifs, chacune essaiera de choisir un prix qui rivalisera au choix des autres firmes. En effet, une petite diminution du prix de la firme d'un petit montant  $\varepsilon$ , sous l'hypothèse que les consommateurs vont acheter chez l'entreprise offrant le plus faible prix, fera gagner cette dernière une grande partie des parts des autres firmes du marché.

### Variation conjecturelles de Bowley

Par l'exemple de deux firmes et au moyen de graphes, Cournot discute la stabilité de son équilibre. Il stipule que chaque réaction d'une firme est une fonction de la stratégie de l'autre firme et il montre ainsi la convergence de cette suite de fonctions vers les meilleures réponses ainsi à son équilibre  $p^c$ .

A base de cette idée, Bowley (1924) réécrit l'équation (4.3) comme suit :

$$\frac{\partial \Pi_i}{\partial q_i} = f(Q) - c'_i(q_i) + q_i f'(Q) \frac{\partial q_j}{\partial q_i} = 0, \quad j \neq i. \quad (4.7)$$

Le terme  $f'(Q) \frac{\partial q_j}{\partial q_i} = 0$  est dit variation conjecturelle de Bowley. Donc d'après Bowley, la politique  $q$  d'une firme suit est affectée par la politique suivie par les autres firmes du marché.

Grâce à cette hypothèse, et en résolvant de nouveau le système, un nouveau équilibre peut être défini.

## 4.3 La concurrence en prix

D'après Bertrand, les firmes, en plus du choix de la quantité, elles doivent aussi faire face au problème majeur que pose la fixation du prix du bien sur le marché. On commence ces modèles par celui de Sweezy qui est basé sur une représentation particulière des courbes de la demande inverse relatives aux firmes dites courbes  $DD'$  et  $dd'$  de Chamberlin.

## Les hypothèses de Sweezy et Stackelberg

Chamberlin présente une nouvelle représentation des courbes de demandes qu'il appela les courbes  $DD'$  et  $dd'$ .

La courbe  $DD'$  représente la demande auxquelles fait face une firme en fonction de son prix sachant que les derniers changements en politique de prix étaient au même taux de changement que  $p_i$ .

La courbe  $dd'$ , quand à elle, représente la demande auxquelles fait face une firme en fonction de son prix sachant que les taux de changement des prix sont constants.

Sweezy, en 1939, grâce à ces célèbres courbes de Chamberlin, propose un nouveau équilibre pour l'oligopole correspondant à l'intersection des deux courbes. Pour mieux comprendre la nature de l'équilibre de Sweezy, supposons un vecteur prix  $p^*$  dominant. La firme est présumée penser que si elle croît son prix, elle devra le faire relativement à  $dd'$ , ainsi, toutes les autres firmes maintiendront leurs prix constant. Si par ailleurs elle décroît son prix, les autres la suivront par le même montant, ce qui fait qu'elle doit bouger le long de  $DD'$  lorsqu'elle pensera décroître son prix.

Sous ces conditions, la firme ne pourra pas faire de bénéfices positifs après le changement de prix, si :

$$\frac{\partial \Pi_i(p^*)}{\partial q_i} \leq 0 \leq \sum_{j=1}^N \frac{\partial \Pi_i(p^*)}{\partial q_j}. \quad (4.8)$$

Ainsi, tout prix vérifiant l'équation (4.8) est un équilibre de Sweezy.

Friedman[23] fait la remarque qu'un tel équilibre est insensé dans le cas statique que si  $p^*$  qui vérifie l'équation (4.8) est annoncé avant qu'il prenne effet sur le marché et ainsi permettre au autres firmes de réagir, mais ce dernier aura un sens dans le cas dynamique. Il révèle alors un point d'insatisfaction.

Von Stackelberg, par le biais d'un exemple de duopole, a imaginé une situation où une des deux firmes a une idée précise du comportement de son concurrent : elle connaît parfaitement sa fonction de réaction et elle l'intègre dans son processus de décision. On appelle alors cette firme le leader ou le meneur.

Suite à sa décision de production, son concurrent réagit en maximisant son profit et donc en suivant sa fonction de meilleure réaction ; elle se contente de "suivre" le comportement du leader et pour cette raison, on l'appelle le suiveur.

Dans ce cas, le suiveur considère que ses décisions n'ont aucun impact sur le comportement du meneur. Il est donc le seul à avoir des conjectures de Cournot. Le duopole de Cournot correspond donc à une situation où les deux firmes ont un comportement de suiveur.

Si la firme 1 est le meneur, son problème est le suivant :

$$\begin{cases} \max_{p_1} \Pi_1(p_1, p_2), \\ \text{sous contrainte,} \\ p_2 = p_2^*(p_1), \end{cases}$$

où  $p_2^*(p_1)$  est la fonction de réaction du suiveur.

Le meneur essaie donc d'atteindre le niveau le plus élevé de profit tout en respectant la fonction de réaction du suiveur.

Une fois que le meneur détermine sa stratégie  $p_1^*$ , le suiveur détermine la sienne par le biais de sa fonction de meilleure réponse  $R_2(p_1^*)$ .

Friedman, propose un modèle sur le choix des prix basé sur la notion d'équilibre de Cournot. Il a fait une étude similaire à celle de Cournot en considérant cette fois que les firmes font leurs décisions de prix et propose un équilibre de Cournot en prix :

Le prix de la  $i^{\text{ème}}$  firme est  $p_i$ , ainsi le vecteur des prix est  $p = (p_1, \dots, p_N)$ . La fonction de demande inverse pour la firme  $i$  est  $q_i = F_i(p)$ , avec  $F_i$  deux fois continûment différentiable, en plus  $\frac{\partial F_i}{\partial p_i} < 0$  et  $\frac{\partial F_i}{\partial p_j} > 0$  pour tout  $j \neq i$ . Ces hypothèses assurent que les ventes de la firme  $i$  décroissent lorsque elle augmente son prix, et augmente lorsque les firmes adverses augmentent aussi leurs prix. De plus, si chacune d'entre elles augmentent leurs prix d'un même montant, elle auront une diminution des ventes. Finalement, on suppose aussi qu'il existe des prix très grands pour lesquels les ventes sont nulles.

On sera donc intéressé par les vecteurs prix pour lesquels la demande est positive. Soit alors

$$\mathbf{A}_i = \{p > 0 : F_i(p) > 0\} \quad (4.9)$$

et soit  $a_i = \{p \in \text{cl}(\mathbf{A}_i) : F_i(p) = 0\}$ .

Les hypothèses gouvernant la demande sont comme suit :

**Hypothèse(4) :**

- $F_i$  définie, continue et bornée  $\forall p \geq 0$ . En plus, elle est deux fois continûment différentiable pour tout  $p > 0$ , excepté pour  $p \in \bigcup_j a_j$ .
- Si  $p \in a_j$ ,  $j = i_1, \dots, i_k$  et  $p \notin a_j$ ,  $j = i_{k+1}, \dots, i_N$  alors les dérivées secondes de  $F_i$  existent au point  $p$  relativement aux indices  $i_{k+1}, \dots, i_N$ .
- Toutes les dérivées sont bornées et, si  $F_i(p) = 0$ ,  $p \notin \text{cl}(\mathbf{A}_i)$  toutes les dérivées des  $F_j$ ,  $j = 1 \dots, N$  au point  $p$  sont nulles.
- Pour  $p \in (\mathbf{A}_i \cap \mathbf{A}_j)$ ,  $j \neq i$ ,  $\frac{\partial F_i(p)}{\partial p_j} > 0$ , et pour  $p \in (\mathbf{A}_i \cap \mathbf{A}_{i_1} \cap \dots \cap \mathbf{A}_{i_k})$ ,  $\frac{\partial^2 F_i(p)}{\partial p_i^2} + \sum_{j=1}^k \frac{\partial^2 F_i(p)}{\partial p_i \partial p_j} < 0$

–  $\mathbf{A} = \prod_{i=1}^N \mathbf{A}_i$  est borné.

La fonction profit dans ce cas est donnée par la formule :

$$\Pi_i(p) = p_i F_i(p) - c_i(F_i(p)), \quad i = 1, \dots, N. \quad (4.10)$$

Une caractéristique essentielle de l'hypothèse sur le système de demande est l'unicité du maximum  $p^*$  pour la fonction  $\Pi_i$ . Ainsi,  $\exists p^* \in \mathbf{A}$  telle que :  $F_i(p^*) = 0$ ,  $\forall i$  et pour tout  $p \neq p^*$ ,  $p \in \mathbf{A}$ ,  $F_i(p^*) > 0$  pour au moins une firme.

Un équilibre au sens de Cournot peut être défini[22].

$p^c$  est un équilibre de Cournot en prix si et seulement si

$$\Pi_i(p^c) \geq \Pi_i(p_i, p_{-i}^c), \quad \forall p_i \geq 0, \quad \text{et} \quad \forall i = 1, \dots, N. \quad (4.11)$$

Une fois encore, la propriété fondamentale de cet équilibre est que quand toutes les firmes  $j \neq i$  choisissent leurs stratégies  $p_j^c$ , la firme  $i$  ne pourra faire mieux que de choisir  $p_i^c$  et ceci pour tout  $i$ , ce qui définit bien un équilibre non coopératif de Nash.

Un sous ensemble particulier de l'ensemble  $\mathbf{A}_i$  est d'un intérêt particulier est :

$$\mathbf{A}_i^* = \{p \in cl(\mathbf{A}_i) : p_i \geq c'_i(F_i(p)), \quad i = 1, \dots, N\}.$$

C'est l'ensemble des vecteurs prix de  $cl(\mathbf{A}_i)$  pour lesquels le prix de la firme est au moins plus grand que sont coût marginal. Ainsi, pour tout  $p \notin \mathbf{A}_i^*$ , le revenu total de la firme est inférieur à son coût variable total si ses ventes sont aussi positives. Alors, à l'équilibre,  $p^c \in \mathbf{A}_i^*$  ou  $p_i^c = p_i^*$ ,  $\forall i$

Posons  $\mathbf{A}^* = \prod_{i=1}^N \mathbf{A}_i^*$ ,  $\bar{\mathbf{A}}^* = \prod_{i=1}^N int(\mathbf{A}_i^*)$ .

Une dernière considération importante à prendre en compte est le prix pour lequel les ventes de la  $i^{\text{ème}}$  firme est nulle, soit  $p_i^+$ . Ce prix satisfait alors  $(p_i^+, 0) \in a_i$ .

On a alors les hypothèses suivantes pour la fonction profit :

– **Hypothèse(5)** :  $\forall p \in int(\mathbf{A}_i^*)$ ,  $\frac{\partial^2 \Pi_i(p)}{\partial p_i^2} < 0$ . cette hypothèse annonce la concavité de la fonction profit relativement à  $p_i$  dans la région où le prix de la firme n'est pas inférieur à son coût marginal. Elle est utiliser pour montrer l'existence d'un équilibre de Cournot.

– **Hypothèse(6)** :  $\forall p \in \bar{\mathbf{A}}^*$  :  $\frac{\partial^2 \Pi_i(p)}{\partial p_i^2} + \sum_{j=1}^N \frac{\partial^2 \Pi_i(p)}{\partial p_i \partial p_j} < 0$ . Cette hypothèse est utilisée pour assurer que l'équilibre est un point fixe d'une fonction pour ensuite montrer l'unicité.

– **Hypothèse(7)** :  $p_i^0 > c'_i(F_i(p_i^0, 0))$ . Cette hypothèse annonce que, peu importe quels prix sont choisis par les autres firmes, le prix le plus profitable pour la firme  $i$  est celui pour lequel ses ventes sont strictement positives. C'est une condition pour laquelle le plus petit prix pour lequel sa demande est nulle est un prix en dessus de son coût marginal à un niveau de production nul.

*On donne ci-après quelques théorèmes qui régissent le conditionnement précédent*

**Théorème 27 [22]**

*Sous les hypothèses 2, 3 et 5, le modèle (4.11) admet un équilibre de Cournot.*

**Théorème 28 [22]**

*Sous les hypothèses 2,4, 5 et 7, les prix et les niveaux de productions sont strictement positifs pour n'importe quelle firme et pour n'importe quel équilibre de Cournot.*

**Théorème 29 [22]**

*Sous les hypothèses 2,4, 5 et 6, l'équilibre pour ce modèle est unique. Si de plus l'hypothèse 7 est vérifiée, alors tous les niveaux de production et tous les prix sont positifs.*

**Théorème 30 [20]**

*Soit  $p^c$  un équilibre de Cournot pour le marché satisfaisant les hypothèses 2 et 4, avec  $p^c \prec p^*$ . Alors, le profit obtenu à l'équilibre n'est pas Paréto optimal.*

*Ce dernier théorème montre que dans la plupart des marchés oligopolistiques, l'équilibre Cournot est Paréto dominé.*

## 4.4 Passage aux modèles dynamiques

*Les modèles dynamiques de d'oligopole sont des modèles où la structure temps apparaît et les équilibres sont tous de type non coopératif. Par structure temps, on sous entend la prise en charge des histoires passés du marché. On distingue 4 types de modèles dynamiques :*

- Le premier type appelé modèles à fonction de réaction traditionnelles : ils dérivent leurs inspiration à partir de la discussion de la stabilité des modèles de Cournot 1927. Stackelberg 1934 et Fellner 1949 ont suivit cette ligne d'investigation, mais sans formaliser le rôle du temps, mais le plus important point soulevé, et celui dans les modèles à différenciation de produits où les comportements relatif aux choix des firmes sont des fonctions continues,  $p_{i,t} = \Psi_i(p_{t-1})$  donnant le prix de la firme  $i$  à la période  $t$  comme une fonction des choix de la période précédente  $t - 1$ .*
- Le deuxième type est les modèles où les changements de prix sont coûteux développés par Marshak et Selten (1978). Ils considèrent les firmes qui font face à un choix de prix, maximisant un flot escomptés de profits en prenant en compte qu'un changement de politique de prix de période en période peut être coûteux.*

- Le troisième type sont les modèles appelés "jeux résumés" pour lesquels les mêmes auteurs *Marshak et Selten* montrent l'existence d'équilibres coopératifs menant à des profits *Paréto* optimaux.
- Finalement, on trouve les modèles à structure temporelle dépendante où le profit serait dépendant de la période précédente et courante.

Le plus célèbre de ces modèles est celui proposé par *Friedman* [23] connu sous le nom de stratégie de déclic ou *Grim trigger*. C'est un modèles menant à des équilibres coopératifs. Dans ces modèles, les firmes cherchent à maximiser un flot de profits actualisés, sans associer des coûts aux changements de prix.

Soit  $p^*$  un prix qui prescrit un profit plus élevé que celui prescrit par l'équilibre Cournot  $p^c$ . Ainsi, le principe est très simple, chaque firme  $i$  choisit dans la période  $t$  le prix  $p_i^*$  si toutes les firmes  $j \neq i$  ont choisit à la période  $t - 1$  les prix  $p_j^*$ . Si, par ailleurs, à la période  $t - 1$  il existe une firme qui a dévié de  $p_j^*$  alors la firme  $i$  va choisir  $p_i^c$ .

On décrit alors le mode de comportement suivant :

$$\begin{cases} p_{i,1} = p_i^*, \\ p_{i,t} = \begin{cases} p_i^*, & \text{si } p_{j,\tau} = p_j^*, \quad j = 1 \dots, N, \quad \tau = 1, 2 \dots, \quad \forall t \geq 2. \\ p_i^c, & \text{ailleurs.} \end{cases} \end{cases} \quad \forall i = 1 \dots, N; \quad (4.12)$$

C'est un mode qui utilise les PSS définis dans le chapitre précédent.

## 4.5 Collusion et modélisation

Les paragraphes précédents révèlent plusieurs points critiques régissant la concurrence oligopolistique :

- \* Il y a plusieurs équilibre dans le jeu,
- \* l'équilibre de Cournot-Nash s'avère ne pas répondre aux exigences des firmes qui souhaitent réaliser le profit le plus grand possible.

En effet la théorie des jeux répétés montre que l'équilibre Cournot-Nash est *Paréto* dominé, en plus *Friedman* dans son modèle dynamique basé sur l'équilibre de Cournot en prix montre que les interactions répétés entre les firmes peuvent mener à des équilibres *Paréto* optimaux.

En fait l'ensemble des équilibres dans le jeu répété est défini par [27]

$$\Omega = \{q \in X : \pi_i(q) \geq (1 - \delta)\pi_i^*(q_{-i}) + \delta\pi_i^c, \quad \forall i \in \mathbb{F}\}$$

En général, il y a un nombre de questions techniques associées à l'ensemble  $\Omega$ , telle que s'il est vide ou pas, les relations entre ses éléments  $q$  est la stratégie de Nash  $q^c$ , les relations entre les profits associés  $\pi_i(q)$  et le profit de Nash  $\pi_i^c$ , et la dépendance de  $\delta$ . Le résultat simple suivant dresse ces questions.

**Proposition 22** [27]

soit  $\delta \in [0, 1]$  arbitraire. Supposons que l'équilibre de Cournot-Nash est unique, pour tout  $q \in \Omega_\delta$ , on a les résultats suivants

- (a)  $\pi_i^*(q_{-i}) - \pi_i^c \geq \pi_i(q) - \pi_i^c \geq (1 - \delta)(\pi_i^*(q_{-i}) - \pi_i^c) \geq 0$  ;
- (b)  $\pi_i(q) > (1 - \delta)\pi_i^*(q_{-i}) + \delta\pi_i^c \Rightarrow \pi_i(q) > \pi_i^c$  ;
- (c) Excepté pour  $\delta = 1$ ,

$$\pi_i^*(q_{-i}) > \pi_i^c \Rightarrow \pi_i(q) > \pi_i^c$$

de plus,  $\Omega_0 = \{q^c\}$  et

$$\Omega_1 = \{q = (q_i) \in X : \pi_i(q_i, q_{-i}) \geq \pi_i^c, \forall i \in \mathbb{F}\}$$

Finalement, pour tout  $0 \leq \delta_1 \leq \delta_2 \leq 1$ ,  $\Omega_{\delta_1} \subseteq \Omega_{\delta_2}$

On a alors la proposition suivante :

**Proposition 23** [27]

Si il existe  $q \in X$  tel que

$$\nabla \Psi_{\delta,i}(q^c)(q - q^c) > 0, \forall i \in \mathbb{F}, \quad (4.13)$$

où,  $\Psi_{\delta,i}(q) = \pi_i(q) - (1 - \delta)\pi_i^*(q_{-i})$ , alors une amélioration de Pareto existe.

Par amélioration de Pareto on entend un vecteur quantité qui Pareto domine le vecteur d'équilibre Cournot.

De tous les modèles discutés précédemment, celui de Friedman basé sur les profils de stratégies simples définie par (4.12) est le plus important à cause de sa structure simple. Il consiste à sélectionner la stratégie la plus profitable pour les entreprises et la rejouer indéfiniment tant qu'une déviation n'est enregistrée. Une fois une firme dévie du sentier collusif  $(p^*, p^*, \dots)$ , les firmes réagiront de façon à infliger au déviant la pire punition et ce en optant pour l'équilibre de Cournot-Nash.

Les accords étant interdits entre les firmes, se pose alors le problème : étant donné que les firmes sont conscientes que la collusion est bénéfique et qu'elles supportent cette dernière et que les accords sont interdits, elles devraient sélectionner un équilibre qui arrangera au mieux les objectifs des firmes ; l'objectif de chaque firme étant guidé par la maximisation de son propre profit. Comment devra être

la sélection ? Rappelons que le modèle de Friedman de la section (4.4) considère cette situation et pose le problème du choix de  $p^*$ .

Le premier modèle proposé consiste en la maximisation du profit joint sur l'ensemble des stratégies des firmes. Cette approche fut critiquée par plusieurs auteurs sur deux côtés : le premier point est que la solution optimale de ce problème peut tomber sur un point qui n'est pas un équilibre, et donc que les firmes ne voudront pas implementer, et le deuxième est qu'il n'est pas toujours vrai que les firmes maximisent leurs profits joints tels que la somme des profits dans une situation où les accords sont interdits.

Harrington, dans [27], propose comme solution à ce problème la résolution du programme suivant :

$$\max_{q \in \Omega} \prod_{i \in \mathbb{F}} (\pi_i(q) - \pi_i^c), \quad (4.14)$$

où  $\pi_i^c = \pi_i(q^c)$  est le profit prescrit par l'équilibre de Cournot.

Rappelons bien que la fonction  $\pi_i(q)$  est définie comme dans (4.1), avec les fonctions  $c_i(q_i)$  sont supposées convexes pour tout  $i = 1 \dots, N$ .

On remarque bien que le travail d'Harrington est basé sur le modèle de Cournot et la concurrence en quantité. Il consiste à chercher le vecteur quantité qui maximiserait l'objectif de négociation de Nash dans l'ensemble des quantités d'équilibre dans le jeu répété.

On suivra le travail d'Harrington et on essaiera de trouver le vecteur prix qui arrangera au mieux les objectifs des firmes.

On se propose alors de modifier le problème ci-dessus, en considérant la même analyse que Friedman a fait pour le modèle de Bertrand dans la section (4.2).

Au fait Friedman a considéré le fait que dans le modèle de Cournot à la fonction demande inverse  $p = f_i(Q)$  on peut établir que  $q_i = F_i(p)$ .

A base de cette idée, on se propose alors de résoudre le problème d'optimisation suivant :

$$\max_{p \in \Omega} \prod_{i \in \mathbb{F}} (\pi_i(F_i(p)) - \pi_i^c), \quad (4.15)$$

où  $\pi_i^c = \pi_i(p^c)$  est le profit prescrit par l'équilibre de Cournot en prix.

$$\Omega = \{p \in X^I : \pi_i(F(p)) \geq (1 - \delta)\pi_i^*(F_{-i}(p)) + \delta\pi_i^c, \forall i \in \mathbb{F}\}$$

où  $F(p) = (F_1(p), \dots, F_N(p))$  et  $F_{-i}(p) = (F_1(p), \dots, F_{i-1}(p), F_{i+1}(p), \dots, F_N(p))$

## 4.6 Conclusion

Le chapitre suivant discutera la solution de ce problème via l'optimisation et ce dans le cas où la fonction demande inverse et du coût total sont linéaires.

# 5

## Résolution du jeu de collusion.

### Introduction

Comme on l'a vu dans le chapitre précédent, on souhaite résoudre le problème de négociation de Nash sous contrainte que les stratégies de la solution doivent être des prix d'équilibre.

Considérons donc  $N$  firmes qui se font la concurrence d'un bien homogène où chaque firme doit prendre une décision de prix.

Le profit de la  $i$ ème firme étant défini comme dans (4.10), l'objectif de négociation de Nash relatif à ces firmes consiste en la maximisation des déviations positives qu'engendrent l'altération des stratégies des firmes de leurs stratégies d'équilibres de Cournot-Nash vers d'autres stratégies d'équilibres engendrant de meilleurs profits pour les firmes, c'est-à-dire

$$\max_{p \in \Omega} \prod_{i \in \mathbb{F}} (\Pi_i(p) - \Pi_i^c), \quad (5.1)$$

où  $\mathbb{F} = \{1, \dots, N\}$  est l'ensemble des indices des firmes.

$\Pi_i^c$  est le profit engendré par l'équilibre de Nash pour la  $i$ ème firme dans le jeu statique.

$p = (p_1, \dots, p_N)$  est le vecteur des stratégies des firmes où  $p_i \in X_i$  est la stratégie de la  $i$ ème firme, et  $X_i$  est l'ensemble des stratégies de la  $i$ ème firme.

$p \in X$ , où  $X = \prod_{i \in \mathbb{F}} X_i$

Considérons l'ensemble  $\Omega$  des quantités d'équilibre dans le jeu répété défini comme suit[22] :

$$\Omega = \{p \in X : \Pi_i(p) \geq (1 - \delta)\Pi_i^*(p_{-i}) + \delta\Pi_i^c, \forall i \in \mathbb{F}\}, \quad (5.2)$$

où  $\Pi_i^*(p_{-i})$  est le profit maximum que peut réaliser la firme  $i$  étant donné la stratégie  $p_{-i}$  des autres firmes, c'est à dire, c'est le profit correspondant à la stratégie  $p_i^*$  telle que

$$\Pi(p_i^*, p_{-i}) = \Pi_i^*(p_{-i}) = \max_{p_i \in X_i} \Pi_i(p_i, p_{-i}). \quad (5.3)$$

Nous allons voir la résolution du problème (5.1) dans le cas où la fonction  $F_i(p)$  dans la fonction profit est linéaire, de la forme

$$F_i(p) = \alpha_i - \langle \beta^i, p \rangle, \quad (5.4)$$

où  $\alpha_i \in \mathbb{R}_+^*$ ,  $\beta^i = (\beta_1^i, \dots, \beta_N^i) \in \mathbb{R}_+^* \times \dots \times \mathbb{R}_+^*$

La fonction coût est aussi linéaire et elle est égale à :

$$c_i(F_i(p)) = k_i \cdot F_i(p), \quad (5.5)$$

où  $k_i$  est le coût fixe relatif à la firme  $i$ ,

### Fonction profit

La fonction profit devient

$$\Pi_i(p) = (\alpha_i - \langle \beta^i, p \rangle)(p_i - k_i). \quad (5.6)$$

En posant  $z_i = p_i - k_i$ , la fonction profit est réécrite comme suit :

$$\Pi_i(p) = (\alpha_i - \langle \beta^i, z + k \rangle)z_i$$

où  $z = (z_1, \dots, z_N)$  et  $k = (k_1, \dots, k_N)$

Après changement de variable de  $t_i = \alpha_i - \langle \beta^i, k \rangle$  on obtient

$$\Pi_i(p) = -\beta_i^i z_i^2 - \sum_{j=1}^N \beta_j^i z_j z_i + t_i z_i,$$

ou bien

$$\Pi_i(p) = z^t A_i z + t_i z_i,$$

où

$$A_i = \begin{pmatrix} 0 & \dots & 0 & -\beta_1^i & 0 & \dots & 0 \\ 0 & \dots & 0 & -\beta_2^i & 0 & \dots & 0 \\ 0 & \dots & 0 & -\beta_3^i & 0 & \dots & 0 \\ \vdots & & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & -\beta_N^i & 0 & \dots & 0 \end{pmatrix}.$$

En posant

$$D_i = \begin{pmatrix} 0 & \dots & 0 & \beta_1^i & 0 & \dots & 0 \\ 0 & \dots & 0 & \beta_2^i & 0 & \dots & 0 \\ 0 & \dots & 0 & \beta_3^i & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ \beta_1^i & \dots & \beta_{i-1}^i & 2\beta_i^i & \beta_{i+1}^i & \dots & \beta_N^i \\ 0 & \dots & 0 & \beta_{i+1}^i & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ 0 & \dots & 0 & \beta_N^i & 0 & \dots & 0 \end{pmatrix},$$

la fonction profit est finalement écrite

$$\Pi_i(p) = \frac{-1}{2} z^t D_i z + t_i z_i. \quad (5.7)$$

Cette matrice n'est pas définie. En effet considérons par exemple le mineur principal suivant

$$D \begin{pmatrix} 1 & i \\ 1 & i \end{pmatrix} = \begin{vmatrix} 0 & \beta_1^i \\ \beta_1^i & 2\beta_i^i \end{vmatrix} = -(\beta_1^i)^2 < 0.$$

### Calcul de $\Pi_i^*(p_{-i})$

La condition d'optimalité du premier ordre pour le problème (5.3) nous donne

$$\frac{\partial \Pi(p_i, p_{-i})}{\partial p_i} = -2\beta_i^i p_i^* - \langle \beta_{-i}^i, p_{-i} \rangle + (\alpha_i + k_i) = 0$$

d'où

$$p_i^* = \frac{\langle (\alpha_i + k_i) \beta_i^i \rangle - \langle \beta_{-i}^i, p_{-i} \rangle}{2\beta_i^i} \quad (5.8)$$

En remplaçant (5.8) dans la fonction profit (5.6) et en appliquant le changement de variable  $z = p - k$ , on obtient

$$\Pi_i^*(p_{-i}) = \frac{1}{4\beta_i^i} (\langle \beta_{-i}^i, z_{-i} \rangle - t_i)^2. \quad (5.9)$$

Si on définit le vecteur  $\beta^i = (0, \beta_{-i}^i)$  la fonction (5.9) peut être réécrite en fonction de  $z$  comme suit :

$$\Pi_i^*(p_{-i}) = \frac{1}{4\beta_i^i} (\langle \beta^i, z \rangle - t_i)^2. \quad (5.10)$$

Sous forme matricielle

$$\Pi_i^*(p_{-i}) = \frac{1}{4\beta_i^i} \left( \frac{1}{2} z^t L_i z - 2t_i \langle \beta^i, z \rangle + t_i^2 \right), \quad (5.11)$$

où  $L_i$  est la matrice carrée symétrique définie positive suivante

$$L_i = \begin{pmatrix} (\beta_1^i)^2 & \beta_1^i \beta_2^i & \dots & \beta_1^i \beta_{i-1}^i & 0 & \beta_1^i \beta_{i+1}^i & \dots & \beta_1^i \beta_N^i \\ \beta_1^i \beta_2^i & (\beta_1^i)^2 & \dots & \beta_2^i \beta_{i-1}^i & 0 & \beta_2^i \beta_{i+1}^i & \dots & \beta_2^i \beta_N^i \\ \vdots & \vdots \\ \beta_1^i \beta_{i-1}^i & \beta_2^i \beta_{i-1}^i & \dots & (\beta_{i-1}^i)^2 & 0 & \beta_{i-1}^i \beta_{i+1}^i & \dots & \beta_{i-1}^i \beta_N^i \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \beta_1^i \beta_{i+1}^i & \beta_2^i \beta_{i+1}^i & \dots & \beta_{i+1}^i \beta_{i-1}^i & 0 & (\beta_{i+1}^i)^2 & \dots & \beta_{i+1}^i \beta_N^i \\ \vdots & \vdots \\ \beta_1^i \beta_N^i & \beta_2^i \beta_N^i & \dots & \beta_{i-1}^i \beta_N^i & 0 & \beta_{i+1}^i \beta_N^i & \dots & (\beta_N^i)^2 \end{pmatrix},$$

et l'ensemble  $\Omega$  est réécrit comme suit :

$$\Omega = \{z \in \mathbb{R}^N : \frac{1}{2} z^t (D_i + (\frac{1-\delta}{4\beta_i^i}) L_i) z - (t_i z_i + (\frac{(1-\delta)t_i}{2\beta_i^i}) \langle \beta^i, z \rangle) + \frac{(1-\delta)t_i^2}{4\beta_i^i} + \delta \Pi_i^c \leq 0, \quad i = 1, \dots, N\},$$

et en posant

$$\begin{cases} Q_i = D_i + (\frac{1-\delta}{4\beta_i^i}) L_i, \\ b_i = (\frac{(1-\delta)t_i}{2\beta_i^i}) \beta^i + (t_i, 0), \\ c_i = \frac{(1-\delta)t_i^2}{4\beta_i^i} + \delta \Pi_i^c, \end{cases} \quad (5.12)$$

où  $(t_i, 0) = (0, \dots, 0, t_i, 0, \dots, 0)^t$ .

On obtient finalement la forme simplifiée de  $\Omega$

$$\Omega = \{z \in \mathbb{R}^N : \frac{1}{2} z^t Q_i z - \langle b_i, z \rangle + c_i \leq 0, \quad i = 1, \dots, N\}$$

### Caractérisation de l'ensemble $X$

Comme on l'a signalé en haut,  $X = \prod_{i \in \mathbb{F}} X_i$ , où chaque  $X_i$  est l'ensemble des stratégies

de la firme  $i$  considéré convexe. En plus, on suppose que les firmes choisiront :

- Un prix positif;
- Le prix choisit devra donné une vente positive ; c'est à dire que les firmes n'auront pas à choisir des prix pour lesquels elles ne vendront rien.

Soit alors  $p_i^+$  le plus petit prix pour lequel les ventes de la firme  $i$  sont nulles.

$X_i$  peut être donc représenté comme suit :

$$X_i = [0, p_i^+] \quad (5.13)$$

et l'ensemble  $X$  sera donc

$$X = [0, p_1^+] \times \dots \times [0, p_N^+] = \prod_{i \in \mathbf{F}} [0, p_i^+]. \quad (5.14)$$

Suivant le changement de variable  $z = p - k$  on transforme l'ensemble comme suit :

$$X = \{z \in \mathbb{R}^N : -k \leq z \leq z^+\}, \quad (5.15)$$

où  $z^+ = p^+ - k$ .

Le programme du jeu de collusion devient alors

$$(P) \left\{ \begin{array}{l} \min - \sum_{i=1}^N \log(-\frac{1}{2}z^t D_i z + t_i z_i - \Pi_i^c) \\ \text{sous contraintes} \\ \frac{1}{2}z^t Q_i z - \langle b_i, z \rangle + c_i \leq 0, \quad i=1, \dots, N, \\ z \in X. \end{array} \right.$$

On discutera deux approches de résolution du programme (P) basées sur le principe de pénalité.

Dans la 1<sup>ère</sup> approche, on construira un algorithme basé sur une pénalité intérieure et le schéma de l'algorithme dit algorithme du gradient projeté. Dans la 2<sup>ème</sup> approche, on va essayer de résoudre le programme (P) après utilisation d'une pénalité extérieure pour le système des contraintes, en utilisant le schéma D.C.A.

## 5.1 Première Approche : Algorithme(PPA)(Penalized Projection Algorithm)

Si on applique une pénalité intérieure pour les contraintes  $\frac{1}{2}z^t Q_i z - \langle b_i, z \rangle + c_i \leq 0$  du problème, On obtient le problème suivant :

$$\hat{P} \min_{z \in X} - \sum_{i=1}^N \log(-\frac{1}{2}z^t D_i z + t_i z_i - \Pi_i^c) - U_j \sum_{i=1}^N \log(-\frac{1}{2}z^t Q_i z + \langle b_i, z \rangle - c_i),$$

où  $\{U_j\}$  sont choisis de telle sorte qu'ils convergent vers 0 quand  $j \rightarrow +\infty$

Ainsi, on obtient un programme de minimisation d'une fonction non linéaire sous contraintes linéaires.

Ce programme peut être réécrit comme suit :

$$\min_{z \in X} - \sum_{i=1}^N \log[-(-1)^{U_j} (\frac{1}{2} z^t D_i z - t_i z_i + \Pi_i^c) (\frac{1}{2} z^t Q_i z - \langle b_i, z \rangle + c_i)^{U_j}].$$

Soit

$$f_i(z) = -(-1)^{U_j} (\frac{1}{2} z^t D_i z - t_i z_i + \Pi_i^c) (\frac{1}{2} z^t Q_i z - \langle b_i, z \rangle + c_i)^{U_j}$$

$f_i$  est différentiable, car elle s'écrit comme le produit de deux fonctions différentiables :

$$\nabla f_i(z) = -(-1)^{U_j} [(D_i z + t_i e_i) ((\frac{1}{2} z^t Q_i z - \langle b_i, z \rangle + c_i)^{U_j}) + (Q_i z - b_i) (\frac{1}{2} z^t Q_i z - \langle b_i, z \rangle + c_i)^{(U_j-1)} (\frac{1}{2} z^t D_i z - t_i z_i + \Pi_i^c)], \quad (\alpha)$$

avec  $e_i = (0, \dots, 0, 1, 0, \dots, 0)$

Le log  $f_i$  est aussi différentiable et on a :

$$\frac{\partial \log f_i(z)}{\partial z} = \frac{1}{f_i(z)} \nabla f_i(z).$$

Les  $U_j$  sont choisis telle que

$$\lim_{j \rightarrow \infty} U_j = 0. \quad (5.16)$$

### 5.1.1 Description de l'algorithme PPA

Choisir  $U_1 > 0$  et résoudre le problème

$$\min_{z \in X} - \sum_{i=1}^N \log(-(-1)^{U_j} \frac{1}{2} z^t D_i z - t_i z_i + \Pi_i^c) (\frac{1}{2} z^t Q_i z - \langle b_i, z \rangle + c_i)^{U_j}$$

avec la méthode du gradient réduit.

Si la solution obtenue est une bonne approximation de la solution de (P)

S'arrêter

Sinon Choisir  $U_2 < U_1$  et recommencer l'itération.

On peut choisir  $U_j = \frac{1}{2^{p_j+1}}$ , ainsi  $-(-1)^{U_j} = -(-1)^{\frac{1}{2^{p_j+1}}} = -(-1) = 1$   
on a

$$\nabla f_i(z) = (D_i z + t_i e_i) \left( \frac{1}{2} z^t Q_i z - \langle b_i, z \rangle + c_i \right)^{\frac{1}{2^{p+1}}} + \frac{1}{2^{p+1}} (Q_i z - b_i) \left( \frac{1}{2} z^t Q_i z - \langle b_i, z \rangle + c_i \right)^{\frac{1}{2^{p+1}} - 1}$$

$$\begin{aligned} \nabla f_i(z) &= \left( \frac{1}{2} z^t Q_i z - \langle b_i, z \rangle + c_i \right)^{\frac{1}{2^{p+1}}} \left[ (D_i z + t_i e_i) + \frac{Q_i z - b_i}{\frac{1}{2} z^t Q_i z - \langle b_i, z \rangle + c_i} \right] \\ &= \left( \frac{1}{2} z^t Q_i z - \langle b_i, z \rangle + c_i \right)^{\frac{1}{2^{p+1}}} \left[ \frac{(D_i z + t_i e_i) \left( \frac{1}{2} z^t Q_i z - \langle b_i, z \rangle + c_i \right) + (Q_i z - b_i)}{\frac{1}{2} z^t Q_i z - \langle b_i, z \rangle + c_i} \right] \end{aligned}$$

**Enoncé de l'algorithme :**

– Etape (0) : Choisir  $p > 1, \varepsilon > 0, z^0$ , poser  $k = 0$  ;

– Etape 1 : Calculer pour  $i = \overline{1, N}$  :

$$a_i^k = \frac{1}{2} (z^k)^t Q_i z^k - \langle b_i, z^k \rangle + c_i,$$

$$b_i^k = \frac{1}{2} (z^k)^t D_i z^k - t_i z_i^k + \Pi_i^c,$$

– calculer :

$$p^k = - \sum_{i=1}^N (a_i^k)^{\frac{1}{2^{p+1}}} \left( \frac{(D_i z^k - t^0) a_i^k + (Q_i z^k - b_i)}{a_i^k} \right),$$

$$\nabla^k = \frac{p^k}{\sum_{i=1}^N b_i^k (Q_i^k)^{\frac{1}{2^{p+1}}}}$$

– Etape2 :

– Si  $\nabla^k \leq \varepsilon$  alors s'arrêter ;  $z^k$  solution.

– Sinon  $Z^{k+1} = \text{Proj}/X \left\{ z^k - \frac{\nabla^k}{\left| \sum_{i=1}^N b_i^k (Q_i^k)^{\frac{1}{2^{p+1}}} \right|} \right\}$

– Poser  $k = k + 1, p = 10p$  et aller en (1)

– Etape3 : Fin.

*Calcul de*

$$Proj/X \left\{ z^k - \frac{\nabla^k}{|\sum_i^N b_i^k (Q_i^k)^{\frac{1}{2p+1}}|} \right\}$$

$$X = \{z \in \mathbb{R}^N, -c \leq z \leq z^+\}$$

*La projection d'un vecteur y sur un ensemble E est définie par :*

$$Proj/E(y) = \min_{x \in E} (\|y - x\|)^2$$

*Dans notre cas,*

$$Proj/X(y) = \begin{cases} y & \text{si } -c \leq y \leq z^+ \\ 0 & \text{sinon.} \end{cases}$$

*L'algorithme peut être réécrit comme suit :*

---

– Etape (0) : Choisir  $p > 1, \varepsilon > 0, z^0$ , poser  $k = 0$ ;

– Etape 1 : Calculer pour  $i = \overline{1, N}$  :

$$a_i^k = \frac{1}{2}(z^k)^t Q_i z^k - \langle b_i, z^k \rangle + c_i; \quad b_i^k = \frac{1}{2}(z^k)^t D_i z^k - t_i z_i^k + \Pi_i^c;$$

– calculer :

$$p^k = - \sum_{i=1}^N (a_i^k)^{\frac{1}{2p+1}} \left( \frac{(D_i z^k - t^0) a_i^k + (Q_i z^k - b_i)}{a_i^k} \right),$$

$$\nabla^k = \frac{p^k}{\sum_{i=1}^N b_i^k (Q_i^k)^{\frac{1}{2p+1}}}$$

– Etape2 :

– Si  $\nabla^k \leq \varepsilon$  alors s'arrêter ;  $z^k$  solution.

– Sinon

– Pour  $i = 1$  à  $N$  faire

– si  $-c_i \leq z_i^k \leq z_i^+$  alors

$$z_i^{k+1} = z_i^k - \frac{\nabla_i^k}{|\sum_{i=1}^N \log(a_i^k (b_i^k)^{\frac{1}{2p+1}})|},$$

– sinon  $z_i^k = 0$ ;

– Poser  $k = k + 1$  et aller en (1)

– Etape3 : Fin.

---

## 5.2 Approche basée sur DCA- Algorithme DCPA (DC Penalized Algorithm)

*Dans cette section, on va essayer, de transformer le programme (P) de telle sorte que DCA soit applicable.*

*L'ensemble X peut être réécrit de la manière suivante*

$$X = \{z \in \mathbb{R}^N : z \geq -c \text{ et } z \leq z^+\}.$$

*En rajoutant une variable d'écart  $e_1 \in \mathbb{R}^N$  pour la première inégalité, et une autre*

variable d'écart  $e_2$  pour la deuxième inégalité de  $X$  et en posant

$$A = \begin{pmatrix} -I_N & I_N & 0_{N \times N} \\ I_N & 0_{N \times N} & I_N \end{pmatrix}$$

où  $I_N$  est la matrice identité d'ordre  $N$  et  $0_{N \times N}$  la matrice carrée nulle d'ordre  $N$ , et

$x = \begin{pmatrix} z \\ e_1 \\ e_2 \end{pmatrix}$ , et  $B = \begin{pmatrix} c \\ z^+ \end{pmatrix}$ . L'ensemble  $X$  est maintenant réécrit sous forme

matricielle comme suit

$$X = \{x \in \mathbb{R}^{3N} : Ax - B = 0\}$$

En augmentant les matrices  $D_i$  et  $Q_i$  comme suit

$$D_i := \begin{pmatrix} D_i & 0_{N \times N} & 0_{N \times N} \end{pmatrix}$$

$$Q_i := \begin{pmatrix} Q_i & 0_{N \times N} & 0_{N \times N} \end{pmatrix},$$

et les vecteurs  $b_i$  par

$$b_i := \begin{pmatrix} b_i \\ 0_N \\ 0_N \end{pmatrix},$$

le programme (P) se réécrit comme suit

$$(P) \begin{cases} \min_{x \in \mathbb{R}^{3N}} -\sum_{i=1}^N \log(-\frac{1}{2}x^t D_i x + t_i x_i - \Pi_i^c) \\ \text{sous contraintes} \\ \frac{1}{2}x^t Q_i x - \langle b_i, x \rangle + c_i \leq 0, \quad i=1, \dots, N, \\ Ax - B = 0 \end{cases}$$

En ramenant le système des contraintes dans la fonction objectif par pénalité extérieure on obtient le programme auxiliaire suivant :

$$\min_{x \in \mathbb{R}^{3N}} -\sum_{i=1}^N \log(-\frac{1}{2}x^t D_i x + t_i x_i - \Pi_i^c) + \alpha^k \left[ \sum_{i=1}^N \max\{0, \frac{1}{2}x^t Q_i x - \langle b_i, x \rangle + c_i\} + \sum_{i=1}^N (A_i x - B_i)^2 \right] \quad (5.17)$$

$\alpha^k$  est le coefficient de pénalité et il doit être choisit telle que :

$$\lim_{k \rightarrow \infty} \alpha^k = \infty. \quad (5.18)$$

Les  $A_i$  représentent les vecteurs lignes de la matrice  $A$  et les  $B_i$  les composantes du vecteur  $B$ .

La fonction objectif de ce programme et non convexe Soient alors

$$\begin{aligned} f_1(x) &= -\sum_{i=1}^N \log\left(-\frac{1}{2}x^t D_i x + t_i x_i - \Pi_i^c\right) \\ f_2(x) &= \alpha^k \sum_{i=1}^N \max\left\{0, \frac{1}{2}x^t Q_i x - \langle b_i, x \rangle + c_i\right\} \\ f_3(x) &= \alpha^k \sum_{i=1}^N (A_i x - B_i)^2 \end{aligned}$$

La fonction  $f_3$  est une forme quadratique définie positive car  $\forall x \in \mathbb{R}^{3N}$ ,  $x \neq 0$   $\sum_{i=1}^N (A_i x - B_i)^2 \geq 0$ , d'où la convexité de  $f_3$

$f_2$  n'est pas convexe, mais elle s'écrit comme le maximum entre 0 et une fonction quadratique non définie. Pour cette dernière on va essayer de trouver une décomposition dc. On alors la proposition suivante

**Proposition 24 [26]**

Si  $f$  est une fonction dc à valeurs finies sur  $C$  alors la fonction  $f^+ = \max\{0, f\}$  est une fonction DC sur  $C$ .

En effet, si  $f = g - h$  alors il suffit d'écrire  $f^+ = \max\{0, g - h\} = \max\{g, h\} - h$ , on obtient ainsi une décomposition dc pour  $f^+$  en posant  $g_1 = \max\{g, h\}$  et  $h_1 = h$

Pour trouver une décomposition dc pour la fonction  $f_2$  on se servira de la décomposition dc pour la forme quadratique  $\frac{1}{2}x^t Q_i x - \langle b_i, x \rangle + c_i$ .

Pour obtenir une décomposition dc pour cette dernière, on se servira de la décomposition proposée par Pham Dinh Tao et Le Thi Hoai An [52].

Soit la forme quadratique suivante

$$f(x) = \frac{1}{2}x^t Q x + \langle b, x \rangle + c_i,$$

où  $Q$  est une matrice symétrique non définie.

Une décomposition dc pour  $f(x)$  s'écrit sous la forme  $f(x) = g(x) - h(x)$ , où

$$\begin{aligned} g(x) &= \left[\frac{1}{2}x^t Q x + \langle b, x \rangle + c_i + \rho x^t I x\right], \\ h(x) &= [\rho x^t I x]. \end{aligned}$$

et  $\rho$  est choisit comme suit :  $\rho \geq -\lambda_1$ , où  $\lambda_1$  est la plus petite valeur propre de la matrice  $Q$  et  $I$  est la matrice unité de même dimension que  $Q$ .

Dans notre cas, la forme quadratique dans la fonction  $f_2$  est non définie à cause de la matrice  $Q_i = D_i + (\frac{1-\delta}{4\beta_i})L_i$ , qui est non définie car elle contient la matrice  $D_i$  non définie.

En posant  $\rho_i \geq -\lambda_1^i$ , avec les  $\lambda_1^i$  est la plus petite valeur propre de la matrice  $D_i$ , et en posant

$$g_i(x) = \frac{1}{2}x^t Q_i x + \rho_i x^t I x - \langle b_i, x \rangle + c_i, \quad (5.19)$$

$$h_i(x) = \rho_i x^t I x, \quad (5.20)$$

on obtient une décomposition dc pour la fonction  $f_2$ .

En ce qui concerne la fonction non convexe  $f_1$ , il est très difficile de trouver une décomposition dc. Pour remédier à ce problème on propose de remplacer cette dernière par une approximation convexe représentée par le développement de Taylor d'ordre 1.

L'approximation de Taylor d'ordre 1 de la fonction  $f_1$  au voisinage d'un point  $x^0$  est donné par l'équation

$$f_1(x) \simeq - \sum_{i=1}^N [\log((x^0)^t D_i x^0 + t_i x_i^0 - \Pi_i^c) + \frac{1}{\log((x^0)^t D_i x^0 + t_i x_i^0 - \Pi_i^c)} (D_i x^0 + t_i e_1)^t (x - x^0)].$$

Posons

$$\begin{aligned} l_1^i(x^0) &= \log((x^0)^t D_i x^0 + t_i x_i^0 - \Pi_i^c), \\ l_2^i(x^0) &= \frac{1}{l_1^i} (D_i x^0 + t_i e_1)^t. \end{aligned}$$

On a  $l_1^i(x^0) \in \mathbb{R}$  et  $l_2^i(x^0) \in \mathbb{R}^{3N}$ .  $f_1$  est réécrite comme suit :

$$f_1(x) \simeq - \sum_{i=1}^N [l_1^i(x^0) + l_2^i(x^0)(x - x^0)].$$

Finalement, on obtient le programme dc suivant :

$$\min_{x \in \mathbb{R}^{3N}} g(x) - h(x), \quad (5.21)$$

où

$$\begin{aligned}
g(x) &= - \sum_{i=1}^N \left( [l_1^i(x^0) + l_2^i(x^0)(x - x^0)] + \alpha^k (A_i x - B_i)^2 \right) \\
&+ \alpha^k \sum_{i=1}^N \max \left\{ \frac{1}{2} x^t Q_i x + \rho_i x^t I x - \langle b_i, x \rangle + c_i, \rho_i x^t I x \right\} \\
h(x) &= \alpha^k \sum_{i=1}^N \rho_i x^t I x.
\end{aligned}$$

### 5.2.1 Description de l'algorithme DCAP

D'après le schéma des DCA, l'étape essentielle consiste à :

1. commencer par un point initial  $x^0$
2. répéter le processus

$$\begin{cases} y^k \in \partial h(x^k), \\ x^{k+1} \in \partial g^*(y^k), \end{cases}$$

jusqu'à ce que  $\|x^{k+1} - x^k\| \leq \epsilon$ .

$h$  est une fonction linéaire homogène, donc son sous différentiel se réduit à l'élément unique  $\nabla h(x)$ , c'est à dire  $\partial h(x^k) = \{\nabla h(x^k)\} = \{2\alpha^k x^k \sum_{i=1}^N \rho_i\}$ . La condition  $y^k \in \partial h(x^k)$  est remplacée par  $y^k = 2\alpha^k x^k \sum_{i=1}^N \rho_i$  la deuxième tache  $x^{k+1} \in \partial g^*(y^k)$  est remplacée par la tache

$$x^{k+1} = \arg \min_{x \in \mathbb{R}^{3N}} \{g(x) - [h(x^k) + \langle x - x^k, y^k \rangle]\}$$

ce qui donne

$$x^{k+1} = \arg \min_{x \in \mathbb{R}^{3N}} \left\{ - \sum_{i=1}^N (l_1^i(x^0) + l_2^i(x^0)(x - x^0)) + \alpha^k [(A_i x - B_i)^2 + \max \left\{ \frac{1}{2} x^t Q_i x + \rho_i x^t I x - \langle b_i, x \rangle + c_i, \rho_i x^t I x \right\}] \right\} - [\alpha^k \sum_{i=1}^N \rho_i (x^k)^t I x^k + \langle x - x^k, y^k \rangle].$$

En posant  $\eta = \sum_{i=1}^N \rho_i$  et en prenant en compte  $y^k = 2\alpha^k x^k \sum_{i=1}^N \rho_i = 2\eta x^k$  on obtient le programme suivant :

$$x^{k+1} = \arg \min_{x \in \mathbb{R}^{3N}} \left\{ - \sum_{i=1}^N (l_1^i(x^0) + l_2^i(x^0)(x - x^0)) + \alpha^k [(A_i x - B_i)^2 + \max \left\{ \frac{1}{2} x^t Q_i x + \rho_i x^t I x - \langle b_i, x \rangle + c_i, \rho_i x^t I x \right\}] \right\} - [\alpha^k \eta (x^k)^t I x^k + \langle x - x^k, 2\eta x^k \rangle]$$

L'algorithme final devient alors :

**Énonce de l'algorithme :**

- 
- (0) Choisir  $\alpha^{(0)} > 0, x^1, \theta, \varepsilon \geq 0$ , poser  $k = 1, I_1 = \emptyset, I_2 = \emptyset$ .
- (1) Calculer  $\eta = \sum_{i=1}^N \rho_i, \nabla_i^k = A_i x^k - B_i$ .  
calculer  $\Delta_{i1}^k = l_2^i(x^0) - 2\rho_i - 2\eta\alpha^k$ , et  $\Delta_{i2}^k = l_2^i(x^0) + 2\rho_i - 2\eta\alpha^k$  (2) Pour  $i = 1, \dots, N$  faire  
Si  $\frac{1}{2}\Delta^t Q_i \Delta \geq \langle b_i, \Delta \rangle - c_i \geq 0$  alors

$$I_1 := I_1 \cup \{i\};$$

Sinon

$$I_2 := I_2 \cup \{i\}.$$

- (3) Calculer

$$x^{k+1} = x^k + \theta \left[ 2\alpha^k \sum_{i=1}^N \nabla_i^k + \alpha^k \left( \sum_{i \in I_1} \Delta_{i1}^k I + Q_i \right) x^k + \left( \sum_{i \in I_2} \Delta_{i2}^k I + Q_i \right) x^k \right]$$

- (4) Si  $\|x^{k+1} - x^k\| \leq \varepsilon$  alors s'arrêter, retourner  $x^{k+1}$ ,  
Sinon poser  $\alpha^{(k+1)} = 10\alpha^{(k)}, k = k + 1$  et retourner (1).
- (5) Fin.
-

## Conclusion

*Nous avons discuté dans ce mémoire une solution du jeu de collusion pour la détermination des actions des firmes représentant les prix concurrentiels. Le modèle proposé est basé essentiellement sur le modèle d'Harrington présenté dans [30] et le modèle de Friedman présenté dans le chapitre 4 section 4.3 pour lequel on discute la solution et on propose deux algorithmes de résolution dont l'un est basé sur la méthode dc, dans le cas d'une fonction de demande inverse et d'une fonction coût linéaires.*

*Même avec l'hypothèse de linéarité des fonctions de demande et de coût, qui ne reste qu'une supposition théorique, le programme formulé est très complexe. Ceci étant l'implémentation des algorithmes proposés nous avancera plus sur la complexité du travail. Donc, dans le même voie d'étude on peut imaginer les perspectives suivantes :*

- L'implémentation des algorithmes PPA et DCPA.*
- La réalisation d'une même investigation dans le cadre de différent type de fonctions de demande inverse.*
- Le prolongement de l'étude au cas de différenciation de produit ; un cas qui correspondra dans notre modèle au cas où l'action de chaque firme  $p_i$  est un vecteur.*

*D'autre part, il serait peut être intéressant d'étudier la relation qui existerait entre la théorie des jeux coopératifs et la collusion puisque cette dernière est considérée comme une négociation secrète.*

## Bibliographie

- [1] Greif A., Milgrom P., and Weingast B.R. *Coordination, commitment and enforcement : The case of merchant guild*. JOURNAL of Political Economy 102, pages 745–776, 1994.
- [2] François Bertrand AKOA. *Approches De Points Intérieurs Et De La Programmation Dc En Optimisation Non Convexe. Codes Et Simulations Numériques Industrielles. PhD thesis, Janvier 2005.*
- [3] Le Thi Hoai An. *Contribution à l'optimisation nonconvexe et l'optimisation globale : Théorie, algorithmes et applications*. Université de Rouen, 6-1997.
- [4] Le Thi Hoai An and Pham Dinh Tao. *The dc (difference of convex functions) programming and dca revisited with dc models of real world nonconvex optimization problems*. Annals of Operations Research, 133 :23–46, 2005.
- [5] Le Thi Hoai An and Pham Dinh Tao. *Dc programming : Theory, algorithms and applications. the state of art*. the Proceedings of The First International Workshop on Global Constrained Optimization and Constraint Satisfaction (Cocos' 02), Valbonne-Sophia Antipolis, France, page 28 pages, October 2-4, 2002.
- [6] Le Thi Hoai An, Pham Dinh Tao, and Le Dung Muu. *A combined dc optimization-ellipsoidal branch-and-bound algorithm for solving nonconvex quadratic programming problems*. JOURNAL of Optimization 2, page 28, 1998.
- [7] Le Thi Hoai An, Pham Dinh Tao, and Le Dung Muu. *Exact penalty in d.c. programming*. Vietnam Journal of Mathematics, 1999.
- [8] Le Thi Hoai An and Pham Dizin Tao. *D.c. programming approach for solving the multidimensional scaling problem. nonconvex optimizations and its applications*. Kluwer Academic Publishers, pages 231–276, 2001.
- [9] Hiriart-Urruty J. B. *Generalized differentiability, duality, and optimization for problems dealing with differences of convex functions*. Lecture Notes in Economics and Mathematical Systems, Springer Verlag, Berlin, Germany, 256 :37–69, 1985.

- [10] *Hiriart-Urruty J. B. Conditions for global optimality, handbook of global optimization.* Edited by R. Horst and P. Pardalos, Kluwer Academic Publishers, Dordrecht, Holland, pages 1–26, 1995.
- [11] *Hiriart-Urruty J. B. From convex optimization to non convex optimization, part 1 : Necessary and sufficient conditions for global optimality, nonsmooth optimization and related topics.* Edited by F. Clarke, Plenum Press, New York, New York, 1989.
- [12] *B.D. Bernheim and M.D. Whinston. Multimarket contact and collusive behavior.* RAND JOURNAL of Economics, 21 :1–26, 1990.
- [13] *A. Bowley. The mathematical groundwork of economics, oxford.* Oxford University Press, 1924.
- [14] *Stekalovski A. C. On problems of global extremum in nonconvex extremal problems.* Izvestya Vuzov, Series Matematika, 8 :74–80, 1990 (in Russian).
- [15] *Stekalovski A. C. On conditions for global extremum in a nonconvex minimization problem.* Izvestya Vuzov, Series Matematika, 8 :74–80, 1991 (in Russian).
- [16] *A.-A. Cournot. Recherches sur les Principes Mathématiques de la Théorie des Richesses.* Hachette, Paris., 1938.
- [17] *Abreu D. Extremal equilibria of oligopolistic supergames.* JOURNAL of Economics Theory 39, pages 191–225, 1986.
- [18] *Abreu D. On the theory of infinitely repeated games with discounting.* Econometrica 56, pages 383–396, 1988.
- [19] *G. Debreu. A Social Equilibrium Existence Theorem. Preceeding of the National Academy of Sciences 38, 1952.*
- [20] *P. Dubey. Inefficiency of nash equilibria.* Mathematics of Operation Research, 11 :1–8, 1986.
- [21] *W. Fellner. Competition among fews.* New-York, A.Knopf, 1949.
- [22] *James Friedman, K. J. Arrow, and M.D. Intriligator. Oligopoly Theory Handbook of Mathematical Economics, voL II., North-Holland Publishing Company,, 1982.*
- [23] *J.W. Friedman. Oligopoly and the theory of games.* RAND JOURNAL of Economics, 21 :1–26, 1990.
- [24] *I. Glicksberg. A further generalization of the kakutani fixed point theorem with application to nash equilibrium points.* Proceeding of the American Mathematical Society 3, pages 170–174, 1952.
- [25] *B. Guerrien. La Théorie des jeux, volume 49. Economica, rue Héricart, 75015,Paris, 1995.*

- [26] *Tuy H. Dc optimization :theory, methods and algorithms, handbook of global optimization.* Edited by Horst R. and Pardalos P., Kluwer Academic Publishers, Dordrecht, Holland, pages 149–216, 1995.
- [27] *J.E. Harrington, B.F. Hobbs, J.S. Pang, A. Liu, and G. Roch. Collusive game solution via optimization.* Springer-Verlag, pages 1–30, 2005.
- [28] *Hartman and P. On functions representable as a difference of convex functions.* Pacific Journals of Mathematics, 9 :707–713, 1959.
- [29] *K. Huynh and D. Besancenot. Économie industrielle : repères, cours, application. Amphi Economie, 2004.*
- [30] *Jr J.E. Harrington. Collusion in multiproduct oligopoly games under a finite horizon.* International Economic Review, 28 :1–14, 1987.
- [31] *Nash J.F. The bargaining problem.* Econometrica 28, pages 155–162, 1950.
- [32] *Nash J.F. Noncooperative games.* Annals of Mathematics 54, pages 289–295, 1951.
- [33] *David M. Kreps. A Course in Microeconomic Theory. Harvester Wheatsheaf, New York, 1990.*
- [34] *Bruce W. Lamar. A method for converting a class of univariate functions into d.c. functions.* J. of Global Optimization, 15(1) :55–71, 1999.
- [35] *David C. Lay. Linear Algebra and its application. Addison Wesley, July 18, 2002.*
- [36] *R. Leonard. Reading cournot, reading nash : the creation and stabilisation of the nash equilibrium.* Economic JOURNAL 104, pages 492–511, 1994.
- [37] *Dür M., Horst R., and Locatelli M. Necessary and sufficient global optimality conditions for convex maximization revisited.* Journal of Mathematical Analysis and Applications, 27 :637–649, 1998.
- [38] *MINOUX M. Programmation mathématique : Théorie et algorithmes, vol. 1. Dunod, Paris, France, 1983.*
- [39] *T. Marshak and R. Selten. General equilibrium with price-making firms, lectures notes in economics and mathematical systems.* Springer-Verlag, Berlin, 1974.
- [40] *J. Von Neumann and O. Morgenstern. Theory of games and economic behavior.* Princeton University Press, 1967.
- [41] *M.J. Osborne and A. Rubinstein. Bargaining and markets.* Academic Press, San Diego, 1990.
- [42] *Pierre Picard. Eléments de microéconomie. Montchrestien, Paris, 1992.*
- [43] *E. Polak. Computational Methods In Optimization, A Unified Approach, volume 77. Academic Press, New York, USA, 1971.*

- [44] Horst R. and Thoai N.V. *Dc programming : Overview*. Journal of Optimization Theory and Applications, 103 :1–43, 1999.
- [45] Horst R., Pardalos P.M., and Thoai N.V. *Introduction to Global Optimization*. Kluwer Academic Publishers, Dordrecht, Holland, 1995.
- [46] Selten R. *Reexamination of the perfectness concept of equilibrium points in extensive games*. International JOURNAL of GAME Theory 4, pages 25–55, 1975.
- [47] M.S. Radjef. *Cours en post graduation sur la théorie des jeux et l'optimisation multicritère*. Technical report, Université A.Mira de Béjaia, 2007.
- [48] R.T. Rockafellar. *Convex Analysis*. Princeton, New Jersey, 1993.
- [49] R. Schmalensee. *Competitive advantage and collusive optima*. International JOURNAL of Industrial Organisation 5, pages 351–367, 1987.
- [50] C. Shapiro. *Theories of Oligopoly Behaviour in Handbook of Industrial Organization*, Schmalensee, R. & R. Willig (Eds.). Amsterdam, North-Holland, 1989.
- [51] P. Sweezy. *Demand under conditions of oligopoly*. Oxford University Press, pages 568–573, 1939.
- [52] Pham Dinh Tao and Le Thi Hoai An. *Convex analysis approach to dc programming : Theory, algorithm and applications*. Acta Mathematica Vietnamica, 33, 1996.
- [53] Pham Dinh Tao and Le Thi Hoai An. *Dc optimization algorithms for solving the trust region subproblem*. SIAM J. Optimization, 8(2) :476–505, 1998.
- [54] Tikhonov. *On reciprocity principle*. Soviet Mathematic Doklady, 22 :100–103, 1980.
- [55] J. Tirole. *The Theory of Industrial Organization*. MIT Press, (Cambridge, Mass.), 1988.
- [56] A. Ulph. *Recent advances in oligopoly theory from a game theory perspective*. JOURNAL of economie Sur-veys, 1(2), 1987.
- [57] Thoai N. V. *On tikhonov's reciprocity principle and optimality conditions in dc optimization*. Journal of Mathematical Analysis and Applications, 25 :673–678, 1998.
- [58] Hal Varian. *Introduction à la Microéconomie*. De Boeck Université, Bruxelles, 1994.
- [59] L. Walras. *Theorie Mathematique De La Richesse Sociale Et Autres Ecrits D'Economie Pure*. Economica, 1999.

## Résumé

Dans ce mémoire, on a proposé un modèle pour le jeu de collusion caractérisé par la concurrence entre  $N$  firmes sur la vente d'un bien homogène avec possibilité de collusion, dans le cas où les stratégies des firmes consistent à faire des décisions de prix. Le modèle proposé se base essentiellement sur le modèle de J.E. Harrington et celui de Friedman qui donne un vecteur d'équilibre de Cournot en prix pour le modèle de base.

On a obtenu un modèle d'optimisation qu'on a tenté de résoudre. Deux algorithmes ont été alors proposés, un basé sur le principe de pénalité intérieure et de la méthode du gradient projeté et le deuxième obtenu par une transformation du modèle initial par le biais d'une pénalité extérieure et la méthode dc, un algorithme basé sur DCA simplifié est alors proposé.

**Mots clés :** Jeu de collusion, Équilibre de Cournot, Jeux répétés, Optimisation, Formes quadratiques, Méthodes de pénalité, Méthode DC, Algorithme DCA.

## ملخص

في هذه المذكرة نقترح نموذجاً رياضياً للعبة التواطؤ الكائنة بين بعض الشركات المنتجة لـ (منتجات) متماثل و المتنافسة فيما بينها في سوق مشتركة قصد تحقيق أكبر قدر من الأرباح مع إمكانية التواطؤ النموذج المقترح يعتمد أساساً في تكوينه على نموذج هارنغتون الذي يعطي كميات الإنتاج المتألية اعتماداً على نظرية الألعاب المتكررة والذي نجري فيه تعديلاً بحيث يعطينا استراتيجيات الشركات المتضمنة في أسعار المنتجات في السوق في ختام هذه المذكرة تم اقتراح نظام خوارزمي لحل النموذج المقترح والذي يعتمد في تفصيله على طريقة الـ DC مفاتيح: نظرية الألعاب، توازن ناش، توازن كورنو، طريقة الـ DC

## Abstract

In this work, we propose a model to the collusive game characterized by competition between  $N$  firms with possibility of collusion, where the firms are price choosers. The model we propose is essentially based on the J.E. Harrington's model where the firms are quantities choosers and the model of Friedman which generate a price vector Cournot equilibrium for the basic model.

Then we discussed the solution of the proposed model via optimization methods. We propose two algorithms, one based on the internal penalty method and the projected gradient algorithm, and the other based on the external penalty method and the simplified DCA.

**Key words :** Collusive game, Cournot equilibrium, Repeated game, Optimization, Quadratic forms, Penalty methods, DC method, Simplified DCA.