Étude des qualifications des contraintes et trajectoires centrales issues d'algorithmes de barrière logarithmique

par

Luc MARCHAND

Mémoire présenté au Département de mathématiques
en vue de l'obtention du grade de maître ès sciences (M.Sc.)

FACULTÉ DES SCIENCES
UNIVERSITÉ DE SHERBROOKE

Sherbrooke, Québec, Canada, septembre 2015

Le 15 septembre 2015

Le jury a accepté le mémoire de Monsieur Luc Marchand dans sa version finale.

Membres du jury

Professeur Jean-Pierre Dussault
Directeur de recherche
Département d'informatique

Professeur François Dubeau
Membre interne
Département de mathématique

Professeur Virginie Charette
Président-rapporteur
Département de mathématique

SOMMAIRE

Ce mémoire fait une revue du concept de qualification des contraintes de premier et deuxième ordre. Les qualifications de premier ordre les plus connues et utilisées sont présentées, ainsi que des exemples et contre-exemples qui mettent en évidence les relations entre elles. Leurs équivalences dans les problèmes convexes sont aussi mises en évidence. Quelques qualifications de second ordre, ainsi que certains contre-exemples, sont présentés. Finalement, les fondations d'un travail visant à alléger les hypothèses essentielles à l'existence de trajectoires centrales d'algorithmes de barrière logarithmique sont présentées.

REMERCIEMENTS

J'aimerais d'abord remercier mon directeur de maîtrise, Professeur Jean-Pierre Dussault, pour ses conseils, son financement et ses directions pour la complétion de ce mémoire. Il est peu dire que, sans lui, un tel document m'aurait été impossible à compléter.

Ensuite, je tiens à remercier mes parents Denis Marchand et Jeannine Therrien qui, depuis mon plus jeune âge, ont valorisé l'éducation. Ils ont toujours continué à me soutenir dans ma formation, autant financièrement que moralement, ce qui m'a permis de compléter ma formation.

Un merci particulier au département de mathématiques de l'Université de Sherbrooke pour la formation qu'ils m'ont fourni et le financement m'ayant permis de mener ce projet à termes.

Finalement, j'aimerais remercier tous mes collègues du BISOUS, en particulier Maxime Toussaint, Catherine Simard et Adam Salvail, pour leurs conseils fréquents dans la réalisation de ce mémoire ainsi que pour l'ambiance de travail exceptionnel dans le bureau.

Luc Marchand

Sherbrooke, juillet 2015

TABLE DES MATIÈRES

LISTE DES TABLEAUX

LISTE DES FIGURES

Introduction

L'optimisation non linéaire est un domaine de mieux en mieux compris par les mathématiciens. Malgré tout, même dans des espaces euclidiens de dimension finie, des situations problématiques peuvent se présenter. En effet, avec peu d'hypothèses, les problèmes peuvent facilement dégénérer, donnant lieu à des modèles difficiles à résoudre pour les algorithmes d'optimisation actuels. Parmi celles-ci, on note la non connexité de l'ensemble réalisable, la non convexité des fonctions en jeu, la présence de plusieurs minima locaux ou l'échec d'outils importants tels que les conditions de Karush-Kuhn-Tucker pour trouver des optima locaux. Les notes de Chinneck [Chi12] donnent un meilleur aperçu de certaines de ces difficultés. Les problèmes de Gilbert, Gonzaga et Karas [GGK05] donnent aussi une idée du genre de comportement que certains algorithmes peuvent avoir lorsqu'ils sont appliqués sur des problèmes non linéaires difficiles.

Une de ces situations problématiques est l'échec des conditions de Karush-Kuhn-Tucker puisqu'il s'agit d'un outil très utile pour les algorithmes d'optimisation. Certaines hypothèses ont été développées afin d'assurer que ces conditions tiennent. Ces conditions sont dénommées *qualifications des contraintes* et sont le principal sujet d'étude de ce mémoire.

Plusieurs articles ont aussi considéré le problème d'existence et de différentiabilité des trajectoires centrales dont [MZ98] et [SW98]. Ces trajectoires ont originalement été introduits par [FM90] et ne sont assurés que lorsqu'une qualification des contraintes plutôt forte est satisfaite. La liste de référence étant trop grande, on réfère le lecteur à la page web [Kra99] pour avoir plus d'information. Une excellente référence sur les méthodes utilisant les trajectoires centrales pour les problèmes linéaires et convexes est le livre de Terlaky [Ter96].

Ce mémoire est divisé en quatre chapitres. D'abord, dans le premier chapitre, une mise en évidence de la problématique nous ayant poussé à faire cette recherche est accomplie. Les conditions nécessaires et suffisantes d'optimalité de Karush-Kuhn-Tucker pour le premier et le deuxième ordre et la possibilité de leur échec en optimisation non linéaire sont explicitées. La notation utilisée tout au long du mémoire est aussi présentée dans ce chapitre.

Le second chapitre présente les différentes qualifications des contraintes de premier ordre. Il ne s'agit pas d'une présentation exhaustive de ces qualifications, puisque certaines d'entre elles sont reliées de manière quasi-évidente et que leur analyse devient moins intéressante. Les qualifications principalement utilisées dans la littérature sont toutefois explicitées, et une analyse des liens entre elles (preuves d'implication ou contre-exemples) est faite. Les équivalences dans le cas de programmes convexes sont explicitées et une extension aux multiplicateurs habituels de Karush-Kuhn-Tucker est citée, élaborant son lien avec les multiplicateurs usuels et les qualifications des contraintes présentées.

Le chapitre 3 est dédié à l'importance de l'élaboration de qualifications des contraintes de second ordre. Il y est montré que de faibles qualifications ne permettent pas d'avoir les résultats d'ordre deux aussi facilement que les résultats de premier ordre. Quelques exemples connus dans la littérature ainsi que certains développements récents sur les qualifications d'ordre deux sont présentés.

Finalement, le quatrième chapitre porte sur certaines propriétés des trajectoires centrales d'algorithmes de barrière logarithmique pour des problèmes où l'optimum ne satisfait pas à la condition de régularité. Ce dernier chapitre tente ainsi de faire le pont entre les qualifications des contraintes et les propriétés d'existence des trajectoires centrales d'algorithmes de barrière logarithmique en ajoutant progressivement des hypothèses au problème non linéaire typique. Plusieurs propriétés des trajectoires centrales sont énoncées et démontrées en exploitant une approche différente de celle empruntée par [FM90].

CHAPITRE 1

Mise en place du problème d'optimisation

Soit le programme d'optimisation non linéaire (d'où l'acronyme NLP pour *non linear program*) :

$$\left\{ \begin{array}{lll} & \min & z = f(x) \\ \text{s.à.} & g(x) & \leq 0 \\ & h(x) & = 0 \end{array} \right. \tag{NLP}$$

où $f(x) : \mathbb{R}^n \to \mathbb{R}, g(x) : \mathbb{R}^n \to \mathbb{R}^m$ et $h(x) : \mathbb{R}^n \to \mathbb{R}^p$ sont des fonctions différentiables. Nous adopterons aussi parfois l'écriture $g_i(x) : \mathbb{R}^n \to \mathbb{R}, i = 1 \dots m$ et $h_j(x) : \mathbb{R}^n \to \mathbb{R}, j = 1 \dots p$. Nous adopterons la convention que les vecteurs d'un problème dit *primal* tel que (NLP) sont des vecteurs colonnes et que les gradients sont des vecteurs lignes afin de simplifier l'écriture. Notre étude porte sur les propriétés d'un minimum local x^* de (NLP). Le point de départ de l'étude de tels problèmes est d'établir des conditions nécessaires d'optimalité, c'est-à-dire que peut-on déduire du fait que x^* est un minimum local ?

Remarque 1.0.1. Le présent mémoire contient plusieurs acronymes. Nous avons décidé d'utiliser les acronymes anglais afin de familiariser les lecteurs avec la terminologie répandue dans le domaine.

1.1 Définitions et notations

Nous rappelons ici plusieurs définitions qui seront utilisées au cours des chapitres du mémoire.

Définition 1.1.1. L'ensemble réalisable d'un problème d'optimisation sous la forme (NLP), noté E, est l'ensemble de points $x \in \mathbb{R}^n$ tels que $g(x) \leq 0$ et $h(x) = 0$.

Cette restriction de solutions possibles nous amène à vouloir définir des vecteurs qui représenteront des directions qui conservent localement la réalisabilité à partir d'un certain point réalisable.

Définition 1.1.2. Le cône tangent à l'ensemble réalisable E en un point $\bar{x}$ est l'ensemble des directions $d \in \mathbb{R}^n$ telles que soit $d = 0$ ou soit qu'il existe une suite de constantes positives réelles $\{t_k\}$ et une suite de vecteurs $\{d_k\} \subset \mathbb{R}^n$ telles que :

$$\{t_k\} \to 0 \text{ lorsque } k \to \infty,$$

$$\{d_k\} \to d \text{ lorsque } k \to \infty,$$

$$\bar{x} + t_k d_k \in E \text{ pour tout } k.$$

Ce cône est également appelé le cône tangent aux contraintes au point $\bar{x}$ et on le notera $\mathcal{T}(\bar{x})$.

Ce cône représente les directions réalisables. Puisqu'il s'agit d'un objet difficile à manipuler concrètement, nous présentons ensuite sa linéarisation, plus facile à utiliser.

Définition 1.1.3. Le cône des directions réalisables linéarisé (ou directions réalisables de premier ordre) en un point $\bar{x}$, noté $\mathcal{F}(\bar{x})$, est l'ensemble :

$$\mathcal{F}(\bar{x}) = \{d \in \mathbb{R}^n | \nabla h_j(\bar{x})d = 0, j = 0, \ldots, p, \nabla g_i(\bar{x})d \leq 0, i \text{ tels que } g_i(\bar{x}) = 0\}.$$

On peut montrer que $\mathcal{T}(\bar{x}) \subset \mathcal{F}(\bar{x})$. Ce cône est particulièrement important, puisqu'il est directement utilisé dans certains outils d'optimisation non linéaire. Il est logique pour un algorithme de minimisation de chercher itérativement des solutions de plus en plus petites. On défini donc des directions dites descendantes puisqu'aux alentours d'un minimum local, il ne devrait pas y avoir possibilité de descente.

Définition 1.1.4. Le cône des directions descendantes (ou demi-espace de descente) est l'ensemble :

$$\mathcal{D}(x) = \{d \in \mathbb{R}^n | \nabla f(x)d \leq 0\}.$$

Le cône des directions strictement descendantes est l'ensemble :

$$d(x) = \{d \in \mathbb{R}^n | \nabla f(x)d < 0\}.$$

Bien que les directions descendantes soient intéressantes, il ne faut pas oublier que notre problème est contraint. On cherche donc à restreindre l'analyse sur des directions de descente, mais aussi réalisables.

Définition 1.1.5. Le cône critique de (NLP) au point réalisable $\bar{x}$ est l'ensemble des directions descendantes réalisables linéarisé, c'est-à-dire :

$$C(\bar{x}) = \left\{ \quad d \quad \middle| \begin{array}{c} \nabla f(\bar{x})d \leq 0, \\ \nabla g_i(\bar{x})d \leq 0, i \text{ tels que } g_i(\bar{x}) = 0, \\ \nabla h_j(\bar{x})d = 0, j = 1 \ldots p \end{array} \right\}.$$

Le cône critique faible du même problème, quant à lui, suit les contraintes d'inégalités comme des contraintes d'égalités telles que la linéarisation de la fonction objectif reste constante :

$$c(\bar{x}) = \left\{ \quad d \quad \middle| \begin{array}{c} \nabla f(\bar{x})d = 0, \\ \nabla g_i(\bar{x})d = 0, i \text{ tels que } g_i(\bar{x}) = 0, \\ \nabla h_j(\bar{x})d = 0, j = 1 \dots p \end{array} \right\}.$$

On observe que le cône critique $C(\bar{x})$ est l'intersection du cône linéarisé $\mathcal{F}(\bar{x})$ et du cône des directions descendantes $\mathcal{D}(\bar{x})$.

Nous aurons aussi besoin de la définition de cône polaire (ou dual négatif) d'un ensemble.

Définition 1.1.6. Soit X un ensemble non-vide de $\mathbb{R}^n$. Le cône dual (ou dual positif) de X est donné par

$$X^* = \{u \in \mathbb{R}^n | u^T d \geq 0 \quad \forall d \in X\}.$$

Similairement, le cône polaire (ou dual négatif) de X est donné par

$$X^\circ = \{u \in \mathbb{R}^n | u^T d \leq 0 \quad \forall d \in X\}.$$

Définition 1.1.7. Le cône normal $\mathcal{N}(\bar{x})$ aux contraintes en $\bar{x}$ est le cône polaire du cône tangent, donc

$$\mathcal{N}(\bar{x}) = \mathcal{T}^\circ(\bar{x}).$$

On peut aussi montrer que $\mathcal{F}^\circ(\bar{x}) \subset \mathcal{N}(\bar{x})$. Nous rappellons maintenant ce qu'est la convexité.

Définition 1.1.8. Une fonction $f(x)$ est dite convexe si pour tous points $x_1, x_2 \in \mathbb{R}^n$ et $0 \leq \alpha \leq 1$, on a

$$\alpha f(x_1) + (1-\alpha) f(x_2) \geq f(\alpha x_1 + (1-\alpha) x_2).$$

Visuellement, une fonction est convexe lorsque tout segment de droite reliant deux points de la fonction est au dessus de la fonction elle-même.

Définition 1.1.9. Un ensemble X est dit convexe si pour tous points x_1 et x_2 dans X, $\alpha x_1 + (1-\alpha) x_2$ avec $0 \leq \alpha \leq 1$ est aussi dans X.

Intuitivement, un ensemble est convexe si, pour toute paire de points dans l'ensemble, le segment les reliant est aussi dans l'ensemble. On rappelle ici que si une fonction est convexe, ses ensembles de niveaux constants seront aussi convexes.

1.2 Conditions d'optimalités

Ayant maintenant les outils nécessaires, on peut commencer l'analyse autour de minima locaux.

Le lagrangien au problème (NLP) est la fonction

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \lambda g(x) + \mu h(x) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{j=1}^{p} \mu_j h_j(x) \tag{1.1}$$

avec $\lambda \in \mathbb{R}^m$ et $\mu \in \mathbb{R}^p$.

En optimisation non linéaire avec contraintes, les conditions nécessaires de Karush-Kuhn-Tucker (conditions de KKT) sont les plus utilisées pour définir un minimum local. Les conditions de KKT d'ordre un stipulent qu'en un minimum local x^* de (NLP) respectant *certaines hypothèses*, il existe des vecteurs $\lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^p$ tels que :

$$\begin{gathered} \nabla \mathcal{L}(x^*, \lambda, \mu) = \nabla f(x^*) + \lambda \nabla g(x^*) + \mu \nabla h(x^*) = 0 \\ \lambda \geq 0, \lambda g(x^*) = 0. \end{gathered} \tag{KKT}$$

Ces dernières équations signifient qu'en un minimum local de (NLP), le lagrangien associé au problème admet au moins un point stationnaire. Si les hypothèses de KKT sont vérifiées, alors la recherche de candidats pour des minima locaux de (NLP) peut se restreindre à trouver des points stationnaires du lagrangien. On note ici que ce ne sont pas tous les points stationnaires du lagrangien qui seront des minima locaux. Il est possible que ces points représentent des maxima locaux ou encore aucun des deux. Nos conditions

Notation	Ensemble d'indices	Type de contraintes
$h(x) = 0$	J	Contraintes d'égalité
$g(x) \leq 0$	I	Contraintes d'inégalité
$g(x^*) = 0$	$I^* \subset I$	Contraintes d'inégalité actives
$g(x^*) = 0, \lambda^* > 0$	$\hat{I} \subset I^*$	Contraintes fortement actives
$g(x^*) = 0, \lambda^* = 0$	$I_0 \subset I, I_0 = I^* \setminus \hat{I}$	Contraintes faiblement actives

Tableau 1.1 – Résumé des notations des contraintes

nécessaires (KKT) signifient que le gradient de la fonction objectif est une combinaison linéaire des gradients des contraintes actives, c'est-à-dire

$$\nabla f(x^*) = -\lambda \nabla g(x^*) - \mu \nabla h(x^*).$$

On note $\Lambda(x^*)$ l'ensemble des vecteurs (λ^*, μ^*) satisfaisant (KKT) au point x^*.

On peut interpréter KKT comme la condition que le gradient de la fonction objectif au point x^* est dans le cône normal aux contraintes $\mathcal{N}(x^*)$.

Remarque 1.2.1. Dans le résultat précédent, il est dit que *certaines hypothèses* sont requises. Un des buts de ce mémoire est de présenter ces hypothèses.

Nous notons J l'ensemble des indices des contraintes $h_j(x)$ et I l'ensemble des indices des contraintes $g_i(x)$. Nous notons aussi par I^* l'ensemble des indices de I tels que $g_i(x^*) = 0$. Les contraintes dont l'indice est dans I^* sont appelées contraintes actives au point x^* et seuls leurs multiplicateurs seront éventuellement non-nuls. Les contraintes actives dont le λ associé est non nul sont dites fortement actives. L'ensemble des indices des contraintes fortement actives est noté $\hat{I}$. Les contraintes actives dont le λ associé est nul sont dites faiblement actives. L'ensemble des indices des contraintes faiblement actives est noté I_0. Ainsi, les ensembles J et I sont définis pour un problème, alors que les indices $I^*, \hat{I}$ et I_0 sont définis en un point donné. Un résumé des indices est offert dans le Tableau 1.1.

Il serait intéressant de pouvoir distinguer la nature des points stationnaires, c'est-à-dire déterminer s'il s'agit d'un minimum local, d'une maximum local ou d'aucun des deux. Pour ce faire, nous aurons besoin des conditions de KKT de second ordre. On a qu'en un minimum local x^* du problème (NLP) satisfaisant *certaines autres hypothèses*, il existe, pour toute direction du cône critique $C(x^*)$, des multiplicateurs $\lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^p$ tels que (KKT) soient satisfaites et :

$$d^t \nabla^2_{xx} \mathcal{L}(x^*, \lambda, \mu) d \geq 0 \quad \forall d \in C(x^*). \tag{KKT2}$$

Il s'agit de la condition nécessaire d'optimalité de second ordre. Il est possible de transformer un peu cette expression. Le cône critique $C(x^*)$ n'admet que les directions de descente selon la linéarisation de la fonction objectif (d'où la condition $\nabla f(x^*)d \leq 0$). Ceci est une condition sur la fonction objectif pour le choix de directions. Il est possible d'éviter une telle condition en regardant les conséquences de se limiter aux directions de descente. En effet, si un multiplicateur λ_i lié à une contrainte g_i est strictement positif, alors tout déplacement dans une direction réalisable non tangente aux lignes de niveaux de cette contrainte entraînera une augmentation de la valeur de la fonction objectif. Dit autrement, la condition $\nabla f(x^*)d \leq 0$ peut être remplacée par une condition sur les $\lambda_i > 0$. La condition d'optimalité nécessaire de second ordre peut aussi s'écrire, pour les même multiplicateurs que ci-haut :

$$d^t \nabla^2_{xx} \mathcal{L}(x^*, \lambda, \mu) d \geq 0 \quad \forall d \text{ tel que } \nabla g_{I^*}(x^*)d = 0, \nabla h(x^*)d = 0. \tag{1.2}$$

Il est possible de renforcer cette condition en séparant les contraintes fortement actives des contraintes faiblement actives. Les conditions de premier ordre nous informent déjà partiellement sur la nature d'un point stationnaire. En effet, dans (KKT), si tous les multiplicateurs des contraintes actives sont strictement positifs (une condition appelée stricte complémentarité), on sait que les directions générées par les matrices $\nabla g_{I^*}(x)$ et

$\nabla h(x)$ ne sont pas de descente. Là où il peut y avoir confusion est pour les multiplicateurs ne respectant pas la stricte complémentarité (contraintes actives) et pour les directions tangentes aux linéarisation des contraintes. Du développement du paragraphe précédent, on sait que la restriction sur les directions $\nabla f(x^*)d \leq 0$ ne devrait en fait qu'affecter les contraintes d'inégalités dont le multiplicateur associé est strictement positif. On doit donc s'assurer d'une courbure positive du lagrangien pour les directions tangentes aux contraintes fortement actives. Dit formellement, on doit avoir $d^t\nabla^2_{xx}\mathcal{L}(x^*, \lambda, \mu)d \geq 0$ pour les directions d telles que $\nabla g_{\hat{I}}(x^*)d = 0, \nabla g_{I_0}(x^*)d \leq 0$ et $\nabla h_j(x^*)d = 0, j = 1 \dots p$.

Remarque 1.2.2. La condition (1.2) n'est pas équivalente à la formulation par le cône critique $C(x^*)$. Toutefois, son renforcement l'est.

La condition suffisante d'optimalité peut facilement être déduite de ce dernier renforcement. On a que s'il existe, pour toute direction $\nabla g_{\hat{I}}(x^*)d = 0, \nabla g_{I_0}(x^*)d \leq 0$ et $\nabla h_j(x^*)d = 0, j = 1 \dots p$ (ou $d \in C(x^*)$), des multiplicateurs $\lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^p$ tels que (KKT) soit satisfaite et :

$$d^t\nabla^2_{xx}\mathcal{L}(x^*, \lambda, \mu)d > 0 \quad \forall d \in C(x^*) \tag{1.3}$$

alors x^* est un minimum local de (NLP). Puisqu'on utilisera principalement les conditions nécessaires plutôt que suffisantes et à des fins de concision, les deux formulations équivalentes de la condition suffisante de second ordre sont dans le même énoncé. Le développement passant d'une forme à l'autre est similaire à celui fait pour les conditions nécessaires.

Remarque 1.2.3. On laisse supposer dans les développements précédents que la seconde formulation, qui prend d par rapport aux signes des λ_i, ne dépend pas de $f(x^*)$. Cette affirmation est fausse, puisque les λ_i sont fonctions continues de $\nabla f(x^*)$. Ainsi, énoncer une expression qui ne dépend que des λ_i revient indirectement à une condition sur $f(x^*)$. La distinction reste intéressante puisque les multiplicateurs λ_i sont aussi intrinsèquement liés aux contraintes $g_i(x^*)$ associées, localisant ainsi la répercussion d'une direction de descente sur les différentes contraintes actives en jeu.

Pour ces deux formulations des conditions de second ordre, on peut en déduire une version "faible" en utilisant les directions du cône critique faible $c(x^*)$ plutôt que les directions du cône critique fort $C(x^*)$. Sous une forme alternative, cette version faible revient à prendre les directions telles que $\nabla g_{I^*}(x^*)d = 0, \nabla h(x^*)d = 0$, c'est-à-dire la condition (1.2).

Comme il est possible que les conditions suffisantes ne soient pas respectées en un minimum local, le mieux que l'on puisse espérer d'un algorithme est de trouver un point satisfaisant les conditions nécessaires. Pour cette raison, l'essentiel des développements est basé sur la satisfaction des conditions nécessaires plutôt que suffisantes. Il est toutefois à noter que les algorithmes se comportent généralement mieux lorsqu'un minimum local satisfait aux conditions suffisantes de second ordre.

Comme dit dans l'énoncé du Théorème des conditions nécessaires (KKT), on demande *certaines autres hypothèses*. L'exemple suivant nous montre en quoi ces autres hypothèses sont essentielles.

Exemple de non-existence de multiplicateurs

Exemple 1.2.4. Soit le programme non linéaire :

$$\begin{cases} & \min \quad z = -x_1 \\ \text{s.à.} & g_1(x) = (x_1-1)^3 + x_2 \leq 0 \\ & g_2(x) = -x_1 \leq 0 \\ & g_3(x) = -x_2 \leq 0. \end{cases}$$

Une représentation graphique est disponible à la Figure 1.1. Le point optimal est $x^* = (1,0)^t$. En x^*, on remarque que les contraintes $g_1(x)$ et $g_3(x)$ sont satisfaites avec égalité. On en déduira donc que $I^* = \{1,3\}$ alors que $I = \{1,2,3\}$ puisqu'il y a trois contraintes d'inégalité. Les multiplicateurs λ_i pouvant être non nuls sont donc λ_1 et λ_3. Les conditions

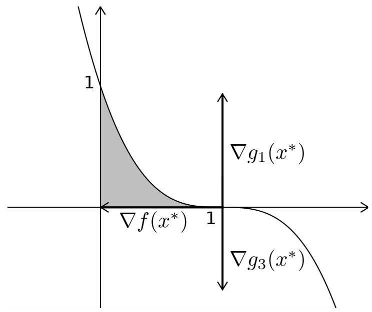


Figure 1.1 – Échec des qualifications des contraintes. Le point optimal étant situé en $(1,0)^t$, on remarque que les gradients des contraintes génèrent l'axe vertical. Notre fonction ayant son gradient vers la partie négative de l'axe horizontal, on remarque que les conditions de KKT ne sont pas vérifiées.

(KKT) s'écrivent ainsi (après avoir traité les équations $\lambda_i g_i(\bar{x}) = 0$ et ainsi, retiré $g_2(x)$ car $\lambda_2 = 0$) :

$$\begin{aligned}\nabla\mathcal{L}(\bar{x},\lambda,\mu) &= \nabla f(\bar{x}) + \lambda\nabla g(\bar{x}) = 0\\ &= \begin{pmatrix}-1 & 0\end{pmatrix} + \begin{pmatrix}\lambda_1 & \lambda_3\end{pmatrix}\begin{pmatrix}3(\bar{x}_1-1)^2 & 1\\ 0 & -1\end{pmatrix} = 0\end{aligned}$$

et en x^* nous avons

$$\nabla\mathcal{L}(x^*,\lambda,\mu) = \begin{pmatrix}-1 & 0\end{pmatrix} + \begin{pmatrix}\lambda_1 & \lambda_3\end{pmatrix}\begin{pmatrix}0 & 1\\ 0 & -1\end{pmatrix}.$$

On remarque qu'il n'existe pas de solution λ^* satisfaisant cette équation et, malgré cela, x^* est bel et bien un minimum local, même global.

Remarque 1.2.5. Dans le cas linéaire, on ne peut trouver d'exemple comme ci-dessus. En fait, l'existence des multiplicateurs sous contraintes linéaires en un minimum local ne nécessite aucune hypothèse supplémentaire.

Ainsi, on doit exiger une hypothèse supplémentaire en un optimum x^* pour que celui-ci soit un point stationnaire du lagrangien (KKT). Cette hypothèse est, en fait, qu'une *qualification des contraintes* soit respectée en x^*. Les prochains chapitres sont donc dédiés à faire le point sur plusieurs qualifications des contraintes trouvées jusqu'à présent.

CHAPITRE 2

Qualifications des contraintes de premier ordre

Cette section présente la notion d'une qualification des contraintes (QC) de premier ordre, suivi de quelques-unes d'entre elles. Plusieurs exemples sont aussi énoncés ici mettant en relation les différentes subtilités entre ces qualifications.

2.1 Définition et arborescence

L'Exemple 1.2.4 montre que, même en un minimum local x^*, il n'existe pas toujours des vecteurs de multiplicateurs de Lagrange $\lambda^* \geq 0$ et μ^* libre tels que $\nabla f(x^*) + \lambda^* \nabla g(x^*) + \mu^* \nabla h(x^*) = 0$. Une qualification des contraintes est simplement une hypothèse qui assure l'existence de ces multiplicateurs en x^* pour que (KKT) soit satisfaite. En fait, l'existence des multiplicateurs (du bon signe) en un point x^* signifie que les directions dans la linéarisation du domaine réalisable ne sont pas de descente stricte. Ainsi, pour assurer l'existence de ces multiplicateurs en un minimum local, on doit s'assurer que le domaine réalisable est bien représenté par sa linéarisation. On peut donc intuitivement voir une qualification des contraintes comme une hypothèse nous certifiant que la linéarisation des contraintes $\nabla g(x)$ et $\nabla h(x)$ agit localement de manière similaire aux contraintes $g(x)$ et $h(x)$.

Remarque 2.1.1. La linéarisation d'une fonction $g(x) : \mathbb{R}^n \to \mathbb{R}$ en x_0 consiste en l'hyperplan de dimension n passant par x_0 et ayant les mêmes dérivées directionnelles que la fonction $g(x)$ au point x_0. Formellement, la linéarisation $l_{x_0}(x)$ en x_0 d'une fonction $g(x)$ est donnée par $l_{x_0}(x) = g(x_0) + \nabla g(x_0)(x - x_0)$. Par le théorème de Taylor, cette fonction évaluée en un point $\bar{x}$ suffisamment près de x_0 devrait donner une bonne approximation de $g(\bar{x})$. Visuellement, la linéarisation de $g(x)$ en x_0 est l'hyperplan tangent aux lignes de niveau de la fonction $g(x)$ en x_0. Les propriétés de cette linéarisation dépendent grandement de la matrice $\nabla g(x_0)$. Bien que ce soit un abus de notation, nous appelons $\nabla g(x_0)$ la linéarisation de $g(x)$ en x_0 pour simplifier et alléger le discours.

En revenant à l'Exemple 1.2.4, on remarque que la linéarisation des contraintes actives

$$\nabla g_{I^*}(x^*) = \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}$$

nous laisse croire que l'axe des x_1 reste réalisable en x^*, donc que l'on pourrait se déplacer autant vers la gauche que vers la droite à partir de $(1, 0)^t$ tout en restant dans l'ensemble réalisable. Toutefois, le vrai domaine réalisable de ce problème se situe entièrement à gauche de $(1, 0)^t$ (plus formellement, dans le demi-plan $x_1 \leq 1$). Il s'agit ainsi d'une situation où le domaine réalisable est mal représenté par sa linéarisation.

Il existe évidemment plusieurs qualifications différentes pour lesquelles les multiplicateurs de Lagrange ont certaines propriétés. Dans ce qui suit, on définira certaines de ces qualifications et leurs relations entre elles.

Problème non-convexe

On considère dans cette section 2.1 un problème de la forme (NLP). Aucune hypothèse de convexité n'est supposée. Nous présenterons les simplifications et ajustements pertinents pour les problèmes convexes à la section 2.2.

Remarque 2.1.2. Les qualifications des contraintes suivantes supposent qu'il existe une boule ouverte en tout point de l'espace du problème. Dans le problème (NLP), il ne s'agit aucunement d'une contraite puisque l'espace de notre problème est $\mathbb{R}^n$, un espace ouvert.

2.1.1 LICQ

Définition 2.1.3. La qualification des contraintes d'indépendance linéaire (LICQ, *Linear Independance Constraints Qualification*) est satisfaite en un point x^* lorsque la matrice

$$\begin{pmatrix} \nabla h(x^*) \\ \nabla g_{I^*}(x^*) \end{pmatrix} \tag{2.1}$$

est de plein rang ligne, c'est-à-dire que le rang de cette matrice est le nombre de lignes de la matrice.

Le théorème suivant nous montre que (LICQ) est bel et bien une (QC).

Théorème 2.1.4. *[FM90] Soit x^* un minimum local de* (NLP)*. Si (LICQ) est satisfaite en x^*, alors il existe une et une seule paire de vecteurs $\lambda^* \in \mathbb{R}^m$ et $\mu^* \in \mathbb{R}^p$ telle que $\nabla f(x^*) + \lambda^* \nabla g(x^*) + \mu^* \nabla h(x^*) = 0$, $\lambda_i^* g_i(x^*) = 0$ et $\lambda^* \geq 0$.*

Exemple 2.1.5. Soit le programme d'optimisation non linéaire (illustré à la Figure 2.1) :

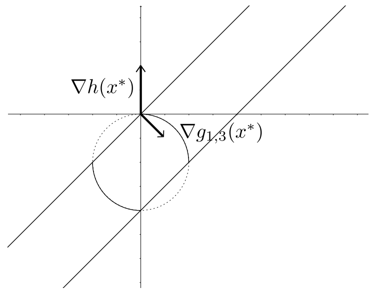


Figure 2.1 – Si le gradient de la fonction objectif est dans le sens de $\nabla h(x)$, on retrouve que la somme de λ_1 et λ_3, tous deux positifs, doit être nulle. On a donc l'unicité des multiplicateurs par l'absence des composantes $\nabla g_1(x^*)$ et $\nabla g_3(x^*)$ dans la reconstruction du gradient de la fonction objectif.

$$\left\{\begin{array}{ll} & \min \quad z = (x_1+1)^2 + (x_2-2)^2 \\ \text{s.à.} & h_1(x) = x_1^2 + (x_2+1)^2 - 1 = 0 \\ & g_1(x) = -x_1 + x_2 \leq 0 \\ & g_2(x) = x_1 - x_2 - 2 \leq 0. \end{array}\right.$$

Ce programme recherche le point de distance minimale entre le point $(-1, 2)^t$ et deux arcs de cercle. La solution de ce problème est $x^* = (0, 0)^t$. En ce point, la contrainte d'inégalité active est la contrainte d'indice 1. La matrice de la linéarisation des contraintes (2.1) en x^* est donc :

$$\begin{pmatrix} 2x_1^* & 2(x_2^*+1) \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ -1 & +1 \end{pmatrix}$$

qui est une matrice de plein rang ligne. (LICQ) étant satisfaite, on s'attend donc à avoir un unique vecteur de multiplicateurs de Lagrange. On rappelle que, puisque $g_2(x^*)$ n'est pas active, son multiplicateur associé sera nul (donc $\lambda_2 = 0$). Ainsi :

$$\begin{aligned} \nabla\mathcal{L}(x,\lambda,\mu) &= \nabla f(x) + \lambda\nabla g(x) + \mu\nabla h(x) \\ &= \begin{pmatrix} 2(x_1+1) & 2(x_2-2) \end{pmatrix} + \begin{pmatrix} \lambda_1 & 0 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} + \mu \begin{pmatrix} 2x_1 & 2(x_2+1) \end{pmatrix} \\ \nabla\mathcal{L}(x^*,\lambda,\mu) &= \begin{pmatrix} 2 & -4 \end{pmatrix} + \lambda_1 \begin{pmatrix} -1 & 1 \end{pmatrix} + \mu \begin{pmatrix} 0 & 2 \end{pmatrix} = 0. \end{aligned}$$

De la composante x_1, on déduit $\lambda_1 = 2$ et de la composante x_2 sachant λ_1, on déduit $\mu = 1$. L'ensemble des multiplicateurs Λ est donc non vide et on observe qu'il est constitué d'un singleton.

En pratique, certaines méthodes de recherche de solutions de (NLP) convergent plus rapidement lorsque x^* est un point satisfaisant (LICQ). C'est pourquoi il s'agit d'une des qualifications des contraintes les plus utilisées et connues. On remarque toutefois que (LICQ) est une qualification très forte, comme le montre l'exemple suivant.

Exemple 2.1.6. Soit le programme d'optimisation à contraintes linéaires (illustré à la Figure 2.2) :

$$\begin{cases} & \min \quad z = (x_1+1)^2 + (x_2+1)^2 \\ \text{s.à.} & g_1(x) = -0.5x_1 - x_2 \leq 0 \\ & g_2(x) = -x_1 - x_2 \leq 0 \\ & g_3(x) = -2x_1 - x_2 \leq 0. \end{cases}$$

La fonction objectif représente la distance euclidienne au point $(-1,-1)^t$. La solution de ce problème est $x^* = (0,0)^t$. Or, en x^*, trois contraintes sont actives dans un espace de dimension 2, ce qui implique que la matrice

$$\nabla g_{I^*}(x^*) = \begin{pmatrix} -0.5 & -1 \\ -1 & -1 \\ -2 & -1 \end{pmatrix}$$

ne peut être de plein rang ligne, ainsi (LICQ) n'est pas satisfaite.

On remarque aussi qu'il existe plusieurs vecteurs de multiplicateurs qui satisfont nos conditions (KKT). En particulier, puisque $\nabla f(0) = \begin{pmatrix} 2 & 2 \end{pmatrix}$, on obtient le système

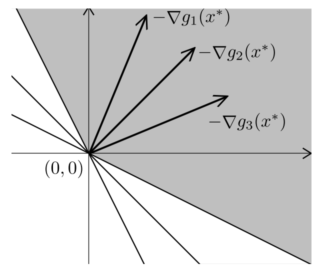


Figure 2.2 – Exemple où (LICQ) échoue. Il est évident, au point optimal $(0,0)^t$, que les gradients des contraintes actives ne peuvent pas être linéairement indépendants, puisque le nombre de contraintes actives en ce point est plus grand que la dimension du problème.

$$2 - 0.5\lambda_1 - 1\lambda_2 - 2\lambda_3 = 0$$
$$2 - \lambda_1 - \lambda_2 - \lambda_3 = 0$$
$$\lambda_i \geq 0 \quad i = 1, 2, 3$$

qui admet les solutions $0 \leq \lambda_1 \leq \frac{4}{3}, \lambda_2 = 2 - \frac{3\lambda_1}{2}$ et $\lambda_3 = \frac{\lambda_1}{2}$.

Nous venons de voir que (LICQ) peut échouer même dans un programme sous contraintes linéaires qui, comme cité à la Remarque 1.2.5, ne nécessite pas de (QC). En effet, dans le problème précédent, puisqu'il y a plus d'un multiplicateur λ^* vérifiant les conditions d'optimalité de (KKT), on pouvait s'attendre à ce que (LICQ) échoue.

Remarque 2.1.7. L'exemple en deux dimensions précédent est construit en ayant une contrainte redondante (c'est-à-dire une contrainte qui ne modifie pas l'espace réalisable). Il est à noter qu'en deux dimensions, il est difficile de faire échouer (LICQ) sans avoir une redondance de contraintes ou encore une contrainte de gradient nul. Ce n'est toutefois pas impossible à faire : ajouter une contrainte d'égalité $x_2 = -x_1 + 1$ au problème 1.2.4 est un exemple où il existe des multiplicateurs pour toute fonction objectif et où (LICQ) échoue, mais où il n'y a pas de contraintes redondantes ou de gradient nul. Dans ce cas, toutefois, l'espace réalisable est restreint à deux points et le $\mathcal{T}(x^*) = \{0\}$.

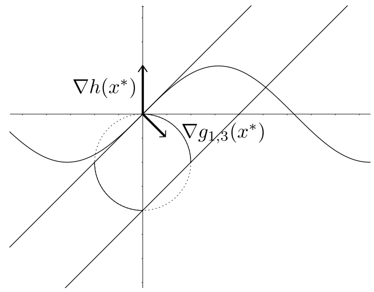


Figure 2.3 – Si le gradient de la fonction objectif est dans le sens de $\nabla h(x)$, on retrouve que la somme de λ_1 et λ_3, tous deux positifs, doit être nulle. On a donc l'unicité des multiplicateurs par l'absence des composantes $\nabla g_1(x^*)$ et $\nabla g_3(x^*)$ dans la reconstruction du gradient de la fonction objectif.

Toutefois, construire de telles situations n'est pas difficile dans des espaces de dimension plus grande que deux. À titre d'exemple, pour un problème de dimension trois, une pyramide à base carrée dont l'optimum est au sommet est un bel exemple d'échec de (LICQ) sans avoir de contrainte redondante.

On amène ici que l'unicité des multiplicateurs n'implique pas (LICQ). L'exemple suivant le montre bien.

Exemple 2.1.8. Soit le programme d'optimisation à contraintes non linéaires (illustré à la Figure 2.3) :

$$\left\{\begin{array}{rl} \min & z = x_1^2 + (x_2-2)^2 \\ \text{s.à.} & h_1(x) = x_1^2 + (x_2+1)^2 - 1 = 0 \\ & g_1(x) = -x_1 + x_2 \leq 0 \\ & g_2(x) = x_1 - x_2 - 2 \leq 0 \\ & g_3(x) = -\sin(x_1) + x_2 \leq 0. \end{array}\right.$$

Il s'agit de l'Exemple 2.1.5 auquel on a ajouté la contrainte $g_3(x)$ et modifié un peu la fonction objectif (on cherche maintenant à minimiser la distance au point $(0,2)^t$ plutôt que $(-1,2)^t$). Tout comme pour l'Exemple 2.1.6, on a trois contraintes actives à la solution $x^* = (0,0)^t$ dans un espace de dimension deux d'où on déduit que (LICQ) échoue. En effet, en x^*, on a que $\nabla g_1(x^*) = \nabla g_3(x^*)$, d'où la dépendance linéaire. Toutefois, notre lagrangien s'écrit :

$$\begin{aligned}
\nabla \mathcal{L}(x, \lambda, \mu) &= \nabla f(x) + \lambda \nabla g(x) + \mu \nabla h(x) \\
&= \begin{pmatrix} 0 & 2(x_2 - 2) \end{pmatrix} + \begin{pmatrix} \lambda_1 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 1 & -1 \\ -\cos(x_1) & 1 \end{pmatrix} + \mu \begin{pmatrix} 2x_1 & 2(x_2+1) \end{pmatrix} \\
\nabla \mathcal{L}(x^*, \lambda, \mu) &= \begin{pmatrix} 0 & -4 \end{pmatrix} + \begin{pmatrix} \lambda_1 & \lambda_3 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix} + \mu \begin{pmatrix} 0 & 2 \end{pmatrix} = 0.
\end{aligned}$$

De la composante en x_1, on déduit que $\lambda_1 + \lambda_3 = 0$. Toutefois, les conditions de (KKT) exigent que les multiplicateurs associés aux inégalités soient positifs. On en conclut $\lambda_1 = \lambda_3 = 0$. Ensuite, de la composante x_2, on trouve $\mu = 2$ et notre ensemble de multiplicateurs est un singleton malgré l'échec de (LICQ).

La non suffisance du Théorème 2.1.4 a été étudiée dans les années précédentes. L'article [Wac13] énonce une condition supplémentaire au Théorème 2.1.4 afin d'avoir aussi la suffisance.

Théorème 2.1.9. *Soit x^* un point réalisable du problème régi par les contraintes de* (NLP). *Il existe une unique paire de vecteurs $\lambda^* \in \mathbb{R}^m$ et $\mu^* \in \mathbb{R}^p$ telle que $\nabla f(x^*) + \lambda^* \nabla g(x^*) + \mu^* \nabla h(x^*) = 0$, $\lambda_i^* g_i(x^*) = 0$ et $\lambda^* \geq 0$ pour toute fonction $f \in \mathcal{C}^1$ admettant un minimum local en x^* si et seulement si (LICQ) est satisfait en x^*.*

Ce théorème donne donc des restrictions supplémentaires sur la linéarisation de la fonction $f(x)$. On note que le concept clé est dans l'expression " pour *toute* fonction $f \in \mathcal{C}^1$ admettant un minimum local en x^*". En effet, l'existence d'un vecteur de multiplicateurs pour une (ou plusieurs) fonction objectif n'est pas suffisant pour garantir (LICQ). Dans l'Exemple 2.1.8, on observe que la raison pour laquelle notre vecteur de multiplicateurs est unique repose sur le fait que $\lambda_1 + \lambda_3 = 0$. En effet, en prenant une fonction objectif telle que la première composante de $\nabla f(x)$ soit positive (disons $\xi > 0$), alors on trouve $\lambda_1 + \lambda_3 = \xi > 0$ et on a ainsi plusieurs multiplicateurs λ_1 et λ_3 possibles. Par exemple, en prenant la fonction objectif de l'Exemple 2.1.5 sujet aux contraintes de l'Exemple 2.1.8, on trouve que l'on doit avoir $\lambda_1 + \lambda_3 = 1$ et on trouve ainsi que notre ensemble de multiplicateurs n'est pas un singleton.

2.1.2 MFCQ

Définition 2.1.10. La qualification des contraintes de Mangasarian-Fromovitz (MFCQ) est satisfaite en un point x^* lorsque la matrice $\nabla h(x^*)$ est de plein rang ligne et qu'il existe un vecteur $v \in \mathbb{R}^m$ tel que $\nabla g_{I^*}(x^*)v < 0$ et $\nabla h(x^*)v = 0$.

Cette qualification signifie intuitivement qu'il existe un vecteur qui entre dans l'intérieur de la linéarisation du domaine réalisable sans rester sur la frontière des contraintes d'inégalité actives tout en restant sur les contraintes d'égalité.

Le théorème suivant montre comment exploiter cette qualification et pourquoi il s'agit d'une des qualification les plus connues dans le domaine de l'optimisation.

Théorème 2.1.11. *Soit x^* un minimum local de la fonction $f(x)$ du problème* (NLP). *On a que x^* satisfait (MFCQ) si et seulement si l'ensemble Λ des vecteurs satisfaisant $\lambda^* \in \mathbb{R}^m, \mu^* \in \mathbb{R}^p$ tels que $\lambda^* \geq 0$, $\nabla f(x^*) + \lambda^* \nabla g(x^*) + \mu^* \nabla h(x^*) = 0$ et $\lambda_i^* g_i(x^*) = 0$ est non vide et borné.*

La nécessité de ce théorème est montrée dans [FM90]. La nécessité et la suffisance sont démontrés dans [Gau77].

Puisque plusieurs algorithmes d'optimisation non linéaire utilisent les multiplicateurs de Lagrange comme outil de recherche de solutions, le fait que l'ensemble des multiplicateurs soit borné sous (MFCQ) est d'une grande utilité. De plus, tout comme (LICQ), cette (QC) est plutôt facile à vérifier. Le résultat que Λ soit borné et non vide implique (MFCQ) (suffisance du théorème précédent) est aussi remarquable.

Corollaire 2.1.12. *Soit un problème de la forme* (NLP). *Si (LICQ) est satisfaite en un point x^*, alors (MFCQ) est aussi satisfaite à ce point.*

Démonstration. Nous procédons par contraposition. On suppose (MFCQ) non satisfaite. Donc l'ensemble des multiplicateurs de Lagrange Λ est soit vide, soit non borné. Dans ces deux cas, l'ensemble des multiplicateurs n'est pas un singleton, et (LICQ) ne peut être satisfaite.

□

Remarque 2.1.13. L'Exemple 2.1.6 satisfait (MFCQ) sans satisfaire (LICQ). Les multiplicateurs de ce problème respectent $0.5\lambda_1 = \lambda_3$ et $2 - 1.5\lambda_1 = \lambda_2$. La condition $\lambda_i \geq 0$ nécessite donc $\lambda_1 \in [0, \frac{4}{3}]$. On remarque ainsi que l'ensemble

$$\Lambda^* = \left\{(\lambda_1, \lambda_2, \lambda_3) \quad | \quad 0 \leq \lambda_1 \leq \frac{4}{3}, \lambda_2 = 2 - \frac{3\lambda_1}{2}, \lambda_3 = \frac{\lambda_1}{2}\right\}$$

des multiplicateurs de Lagrange possibles dans les équations (KKT) est non vide, borné et de cardinalité infinie.

Une condition équivalente à (MFCQ) est donnée par l'énoncé suivant :

Proposition 2.1.14. *(MFCQ) est satisfaite en x^* si et seulement si la condition suivante est satisfaite :*

$$\begin{Bmatrix} \lambda\nabla g(x^*) + \mu\nabla h(x^*) = 0 \\ \lambda g(x^*) = 0 \\ \lambda \geq 0 \end{Bmatrix} \implies \begin{Bmatrix} \lambda = 0 \\ \mu = 0 \end{Bmatrix}. \tag{2.2}$$

On appelle la propriété (2.2) *l'indépendance linéaire positive des gradients des contraintes d'inégalités.*

Démonstration. Cette démonstration, inspirée du document de Rockafellar [Roc], provient du professeur Jean-Pierre Dussault [Dus14],

Notons q la cardinalité de I^*. Soit le programme linéaire

$$\begin{array}{cc} \min\limits_{(d,u)\in\mathbb{R}^n\times\mathbb{R}} & z = u \\ \text{s. à.} & \nabla g_{I^*}(x^*)d \leq ue \\ & \nabla h(x^*)d = 0 \end{array} \tag{2.3}$$

où $e = \Big(1, 1, \ldots, 1\Big)^t \in \mathbb{R}^q$. On observe que $(d, u) = (0, 0)$ est toujours réalisable. La condition $\lambda g(x^*) = 0$ est camouflée dans le fait que l'on ne travaille que sur les contraintes d'indice dans I^*. Le dual de ce problème est

$$\begin{array}{cc} \max\limits_{(\lambda,\mu)\in\mathbb{R}^q\times\mathbb{R}^p} & 0 \\ \text{s. à.} & \lambda\nabla g_{I^*}(x^*) + \mu\nabla h(x^*) = 0 \\ & \lambda e = 1 \\ & \lambda \geq 0. \end{array} \tag{2.4}$$

($\Rightarrow$) On suppose d'abord que (MFCQ) est satisfaite. Alors, il existe une direction $\tilde{d} \in \mathbb{R}^n$ telle que

$$\nabla g_{I^*}(x^*)\tilde{d} < 0$$

$$h(x^*)\tilde{d} = 0.$$

Posons $\nabla g_{I^*}(x^*)\tilde{d} \leq \alpha e$ où $\alpha < 0$. Ainsi, $(\tilde{\alpha}, \alpha)$ est une solution réalisable du problème primal (2.3) telle que

$$\min_{(\tilde{\alpha},\alpha)\in\mathbb{R}^n\times\mathbb{R}} \quad z = \alpha < 0.$$

La solution du primal (2.3), si elle existe, est clairement plus petite ou égale à $\alpha < 0$ alors que la solution du problème dual (2.4) est 0. On en déduit que le problème (2.3) n'est pas borné inférieurement et que le problème (2.4) n'est pas réalisable par le théorème de dualité faible. Alors il n'existe pas $\lambda \geq 0, \mu$ libre et λ non tous nuls tels que

$$\begin{gathered} \lambda \nabla g_{I^*}(x^*) + \mu \nabla h(x^*) = 0 \\ \lambda \geq 0 \end{gathered} \tag{2.5}$$

qui est équivalent au résultat (2.2).

($\Leftarrow$) On suppose d'abord que (2.2) est satisfait. Ceci signifie qu'il n'existe pas $(\lambda, \mu) \neq 0$ tels que (2.5) est satisfait ou encore qu'il n'existe pas (λ, μ) tels que (2.4) est réalisable. Ainsi, le dual est non réalisable. Par le théorème de dualité faible, on doit donc avoir le problème primal (2.3) non borné inférieurement (car réalisable en $d = 0, u = 0$), donc il existe $d \in \mathbb{R}^n, u \in \mathbb{R}$ tels que

$$\begin{gathered} u < 0 \\ \nabla g_{I^*}(x^*) d \leq ue \\ \nabla h(x^*) d = 0 \end{gathered}$$

d'où on déduit (MFCQ).

$\square$

L'exemple suivant montre une situation telle que, pour toute fonction objectif de classe $\mathcal{C}^2$ admettant $(0,0)^t$ comme minimum, il existe des multiplicateurs sans que (MFCQ) soit satisfaite. Il est donc possible d'affaiblir davantage (MFCQ), toutefois le fait que l'ensemble des multiplicateurs soit borné est perdu.

Exemple 2.1.15. Soit le programme d'optimisation à contraintes non linéaires (illustré à la Figure 2.4) :

$$\begin{cases} \quad\quad \min \quad z = (x_1+1)^2 + x_2^2 \\ \text{s.à.} \quad g_1(x) = x_1^2 - x_2 \leq 0 \\ \quad\quad g_2(x) = -x_1^2 + x_2 \leq 0 \\ \quad\quad g_3(x) = -x_1 \leq 0. \end{cases}$$

Le domaine réalisable de ce programme est la branche $x_1 \geq 0$ de la parabole. La solution de ce programme de minimisation est située au point $x^* = (0,0)^t$. On remarque ici que (MFCQ) n'est pas satisfaite puisqu'il n'y a pas d'intérieur strict des contraintes.

Si l'on calcule le lagrangien de ce problème en x^*, nous obtenons

$$\nabla \mathcal{L}(x^*, \lambda^*) = \begin{pmatrix} 2 & 0 \end{pmatrix} + \begin{pmatrix} \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 0 & 1 \\ -1 & 0 \end{pmatrix} = 0.$$

On observe donc que $\lambda_1 = \lambda_2$ et $\lambda_3 = 2$ satisfont ces équations pour toute valeur positive de $\lambda_1 \in \mathbb{R}$. Il y a donc un ensemble non borné de multiplicateurs possibles puisque les gradients des contraintes actives $\nabla g_1(x)$ et $\nabla g_2(x)$ sont linéairement dépendants et de directions opposées pour tout point réalisable.

Remarque 2.1.16. (MFCQ) réagit très mal aux contraintes d'égalité qui se dupliquent en deux contraintes d'inégalité. C'est le cas de l'exemple précédent avec les contraintes $g_1(x)$ et $g_2(x)$. En effet, si l'on définit $h_1(x) = x_1^2 - x_2 = 0$, l'ensemble réalisable reste le même en prenant les contraintes $h_1(x)$ et $g_3(x)$ et, avec une telle définition des contraintes, (MFCQ) est satisfaite en x^*.

2.1.3 CRCQ

Définition 2.1.17. La qualification des contraintes du rang constant (CRCQ) (pour « Constant Rank Constraints Qualification ») est satisfaite en un point x^* de l'ensemble

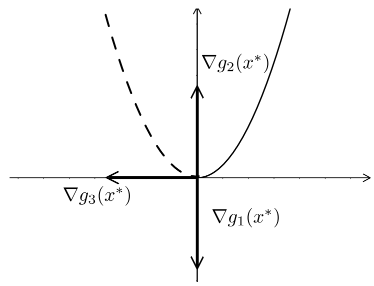


Figure 2.4 – Exemple où (MFCQ) échoue. Le domaine réalisable est la branche pleine de la parabole. La problématique vient du fait qu'il n'existe pas d'intérieur du domaine réalisable régi par les inégalités. Une forme alternative des contraintes (si l'on considère $g_1(x)$ et $g_2(x)$ comme une seule contrainte d'égalité) résout cette problématique.

réalisable E lorsqu'il existe un voisinage $X \subset E$ autour de x^* tel que pour tout sous-ensemble $K \subset I^*$ et $K' \subset J$, le rang de la matrice

$$\begin{pmatrix} \nabla h_j(\bar{x}) \\ \nabla g_i(\bar{x}) \end{pmatrix}, \quad i \in K, j \in K'$$

est constant $\forall \bar{x} \in X$.

Il s'agit bel et bien d'une qualification des contraintes (donc assurant l'existence des multiplicateurs de Lagrange). Toutefois, elle a été développée plus tard que les qualifications précédentes. Cette qualification a été amenée principalement par Spingarn [Spi83] et Janin [Jan84]. Nous avons l'implication suivante :

Proposition 2.1.18. *Soit un problème de la forme* (NLP). *Si (LICQ) est satisfaite en un point x^*, alors (CRCQ) est aussi satisfaite en ce point.*

Démonstration. (LICQ) satisfaite signifie que la matrice (2.1) est de plein rang ligne. Donc aucune ligne de la matrice n'est linéairement dépendante des autres. Puisque les fonctions $g_i(x)$ et $h_j(x)$ sont au moins de classe $\mathcal{C}^2$, alors $\nabla g_i(x)$ et $\nabla h_j(x)$ sont continues, donc il existe un voisinage X de x^* tel que la matrice

$$\begin{pmatrix} \nabla h(\bar{x}) \\ \nabla g_{I^*}(\bar{x}) \end{pmatrix}, \bar{x} \in X$$

est aussi de plein rang ligne, et on a (CRCQ).

$\square$

Remarque 2.1.19. Le fait que les contraintes $g_i(x), i \in I^*$ soient linéaires est suffisant à ce que (CRCQ) soit satisfaite en x^* puisque cela assure que la matrice composée des lignes $\nabla g_i(x), i \in I^*$ et $\nabla h_j(x^*), j \in J$ est constante sur $\mathbb{R}^n$.

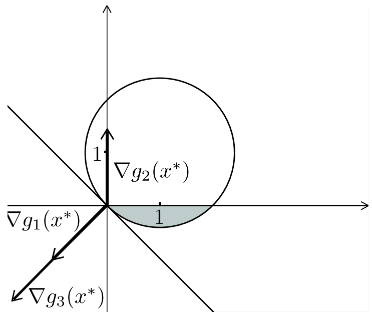
Les relations entre (CRCQ) et (MFCQ) sont plus subtiles. D'abord, remarquons que (CRCQ) est satisfaite pour l'Exemple 2.1.15, donc (CRCQ) n'implique visiblement pas (MFCQ). En effet, pour tout point de $\mathbb{R}^2$ (donc en particulier pour un voisinage de x^*), les gradients $\nabla g_1(x)$ et $\nabla g_2(x)$ sont linéairement dépendants (matrice de rang 1). De plus, les gradients $\nabla g_1(x)$ et $\nabla g_3(x)$ (ou $\nabla g_2(x)$ et $\nabla g_3(x)$) sont linéairement indépendants (de rang 2). Donc pour tout sous-ensemble des contraintes actives, la matrice des gradients des dites contraintes est de rang constant dans $\mathbb{R}^2$.

L'exemple suivant montre que (MFCQ) n'implique pas (CRCQ) non plus.

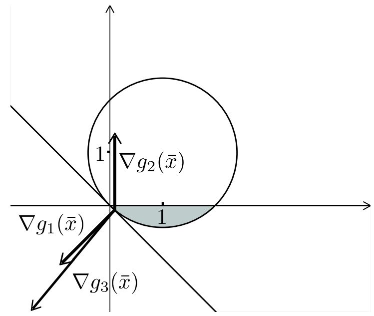
Exemple 2.1.20. Soit le programme d'optimisation à contraintes non linéaires (illustré à la Figure 2.5) :

$$\begin{cases} & \min \quad z = x_1^2 + (x_2 - 1)^2 \\ \text{s.à.} & g_1(x) = -x_1 - x_2 \leq 0 \\ & g_2(x) = x_2 \leq 0 \\ & g_3(x) = (x_1 - 1)^2 + (x_2 - 1)^2 \leq 0. \end{cases}$$

L'ensemble réalisable est borné par l'arc du cercle défini par $g_3(x) = 0$ et sous l'axe des x_1. Le minimum local de ce problème est situé en $x^* = (0, 0)^t$ où les trois contraintes sont



(a) Au point x^*



(b) En $\bar{x}$ près de x^*

Figure 2.5 – Exemple où (CRCQ) échoue. Le point optimal x^* est en $(0,0)^t$. On remarque que le long de la droite $x_2 = x_1$, les gradients des contraintes 1 et 3 sont linéairement dépendants (Figure 2.5a), mais pas ailleurs (Figure 2.5b). Puisque $x_2 = x_1$ passe par x^*, alors il n'existe pas de voisinage autour de x^* de rang constant.

actives. En ce point, le vecteur $v = (1, -1/4)^t$ nous amène dans l'intérieur du domaine réalisable linéarisé, donc (MFCQ) est satisfaite. Formellement, on voit que

$$\begin{pmatrix} \nabla g_1(x^*) \\ \nabla g_2(x^*) \\ \nabla g_3(x^*) \end{pmatrix} v = \begin{pmatrix} -1 & -1 \\ 0 & 1 \\ -2 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ -1/4 \end{pmatrix} = \begin{pmatrix} -3/4 \\ -1/4 \\ -3/2 \end{pmatrix}$$

et ainsi $\nabla g_i(x^*)v < 0, \forall i \in I^*$, d'où (MFCQ) satisfaite.

Toutefois, si l'on prend l'ensemble d'indices $K = \{1, 3\}$, alors la matrice des $\nabla g_i(x^*), i \in K$, soit

$$\begin{pmatrix} -1 & -1 \\ 2x_1 - 2 & 2x_2 - 2 \end{pmatrix}$$

n'admet pas de voisinage de rang constant autour de x^* (la matrice est de rang 1 sur la droite $x_1 = x_2$ et de rang 2 ailleurs). Ainsi, (CRCQ) n'est pas satisfaite en x^*.

Remarque 2.1.21. L'auteur [Lu11] a récemment prouvé que, dans un problème non paramétrique (tels que ceux présentés plus haut), (CRCQ) implique (MFCQ) si la forme des contraintes est modifiée d'une certaine façon. On peut voir cela dans l'Exemple 2.1.15, si l'on redéfinit $g_1(x)$ et $g_2(x)$ par une seule contrainte d'égalité, soit $h_1(x)$. Les modifications apportées ne doivent évidemment pas modifier le domaine réalisable.

L'exemple qui suit provient de [CW04] nous servira à déduire la prochaine qualification des contraintes présentée, historiquement importante.

Exemple 2.1.22. Soit le programme d'optimisation non linéaire (illustré à la Figure 2.6) :

$$\begin{cases} & \min \quad z = -x_1 \\ \text{s.à.} & g_1(x) = (x_1 - 1)^3 + x_2 \leq 0 \\ & g_2(x) = -x_1 \leq 0 \\ & g_3(x) = -x_2 \leq 0 \\ & g_4(x) = 2x_1 + x_2 - 2 \leq 0. \end{cases}$$

Il s'agit de l'Exemple 1.2.4 auquel on a ajouté une contrainte active en $x^* = (1,0)^t$. On remarque que (MFCQ) n'est pas satisfaite (les seules directions réalisables sont le long de l'axe des x_1, pas à l'intérieur strict du domaine réalisable) et (CRCQ) non plus (l'ensemble d'indices $K = \{1,3\}$ n'est pas de rang constant dans tout voisinage de x^*). On remarque toutefois qu'il existe des multiplicateurs de Lagrange (λ^*, μ^*). En effet, on trouve que $\nabla f(x^*) = (-1,0)$ et que $\nabla g_1(x^*) = (0,1), \nabla g_3(x^*) = (0,-1)$ et $\nabla g_4(x^*) = (2,1)$. En prenant $\lambda_4 = 1/2$ et $\lambda_1 - \lambda_3 = -1/2$, on obtient les équations de KKT satisfaites justifiant ainsi qu'il existe des multiplicateurs de Lagrange non bornés pour ce problème.

Remarque 2.1.23. Cet exemple nous permet d'observer une particularité remarquable de la notion de qualification des contraintes. L'ajout de la contrainte $g_4(x)$ ne modifie d'aucune manière le domaine réalisable. On pourrait même affirmer qu'elle semble inutile au problème. Toutefois, cet ajout nous permet de travailler sur les multiplicateurs de Lagrange dans la recherche du point optimal. On doit toutefois observer que cela ne nous permet pas nécessairement d'avoir une qualification forte telle que (LICQ). Vu différemment, on peut aussi en déduire que les qualifications des contraintes ne sont pas

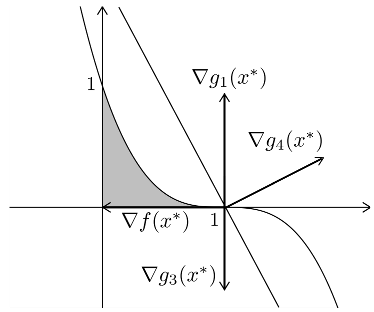


Figure 2.6 – Exemple où (MFCQ) et (CRCQ) échouent. (MFCQ) échoue car le seul vecteur réalisable est dans la direction $(-1, 0)^t$ qui n'est pas strictement réalisable ($g_3(x)$ reste satisfaite avec égalité le long de ce vecteur). De plus, (CRCQ) n'est pas satisfaite puisque pour les contraintes $g_1(x)$ et $g_3(x)$, il n'existe pas de voisinage de rang constant en x^*.

des exigences sur la structure du domaine réalisable, mais vraiment des exigences sur la structure des contraintes ou encore, sur la structure de la description analytique du domaine réalisable.

2.1.4 Kuhn-Tucker CQ

Définition 2.1.24. La qualification des contraintes de Kuhn-Tucker (KTCQ) est satisfaite en x^* si, pour tout vecteur non nul z tel que $\nabla g_i(x^*)z \leq 0$ pour tout $i \in I^*$ et $\nabla h_j(x^*)z = 0, j \in J$, z est tangent à un arc $\alpha(\theta), \theta \geq 0$ une fois différentiable tel que $\alpha(0) = x^*$ et $\alpha(\theta)$ est localement contenu dans le domaine réalisable pour $\theta \geq 0$.

Il s'agit de la première qualification des contraintes à avoir été publiée. Elle est utilisée en premier par Karush dans [Kar39], puis par Kuhn et Tucker dans [KT51] [1]. Cette (QC), bien qu'importante historiquement, est maintenant considérée de moins grande importance avec la découverte des qualifications des contraintes d'Abadie et de Guignard,

1. Karush a découvert cette qualification (et les conditions de KKT) 12 ans avant Kuhn et Tucker, au cours de sa maîtrise. Voir [Cot12] pour davantage d'informations.

des (QC) encore plus faibles que (KTCQ). De plus, (KTCQ) est difficile à vérifier, limitant son utilisation dans des algorithmes.

Remarque 2.1.25. Dans l'Exemple 1.2.4, (KTCQ) n'est pas satisfaite. En effet, il existe deux famille de vecteurs non nul z_j tels que $\nabla g_i(x^*)z_j \leq 0$ $(j = 1, 2)$; il s'agit des vecteurs $z_1 = \alpha(1, 0)^t$ et $z_2 = \beta(-1, 0)^t$, $\alpha, \beta \in \mathbb{R}^+$. Or, pour les vecteurs z_1, il n'existe pas l'arc différentiable nécessaire à ce que (KTCQ) soit satisfaite.

Remarque 2.1.26. Dans l'Exemple 2.1.22, (KTCQ) est satisfaite. On rappelle qu'il s'agit de l'Exemple 1.2.4 avec une contrainte redondante supplémentaire au point optimal. La satisfaction de (KTCQ) est expliquée par le fait que le vecteur problématique z_1 n'est plus tel que $\nabla g_i(x^*)z_1 \leq 0$ (car $\nabla g_4(x^*)z_1 > 0$). (KTCQ) est donc désormais vérifiée.

Théorème 2.1.27. *Soit un problème de la forme* (NLP). *Si (CRCQ) et/ou (MFCQ) sont satisfaites en un point x^*, alors (KTCQ) est aussi satisfaite en ce point.*

La démonstration de ce théorème se divise évidement en deux parties. La suffisance de (MFCQ) a implicitement été démontrée par Fiacco et McCormick [FM90]. Dans cet ouvrage, ils ont montré que (MFCQ) était une qualification des contraintes en construisant un arc différentiable, donc en montrant que (KTCQ) est satisfaite. Pour la suffisance de (CRCQ), on réfère à l'ouvrage original de Janin [Jan84, Proposistion 2.3].

2.1.5 CPLD

Définition 2.1.28. Soit un problème de la forme (NLP). La qualification des contraintes de dépendance linéaire positive constante (CPLD) est satisfaite en x^* si pour chaque sous-ensemble $K \subset I^*$ et $K' \subset J$ tels que les gradients des contraintes $g_i(x^*), i \in K$ et des contraintes d'égalité $h_j(x^*), j \in K'$ sont positivement linéairement dépendantes en x^*, il existe un voisinage X de x^* tel que pour tout $\tilde{x} \in X$, on a que les lignes de la matrice

$$\begin{pmatrix} \nabla h_j(\bar{x}) \\ \nabla g_i(\bar{x}) \end{pmatrix}, \quad i \in K, j \in K'$$

sont positivement linéairement dépendantes.

On reconnait ici une sorte de mixte entre (MFCQ) et (CRCQ). Cette qualification indique que pour tout sous-ensemble des contraintes ne repectant (MFCQ), ce sous-ensemble reste positivement linéairement dépendant. Cette condition est moins restrictive que le rang constant employé par (CRCQ).

Théorème 2.1.29. *Soit un problème de la forme* (NLP)*. Si (MFCQ) est satisfaite en un point x^*, alors (CPLD) est aussi satisfaite en ce point.*

Démonstration. (MFCQ) est une hypothèse qui implique l'indépendance linéaire positive des gradients des contraintes d'inégalité actives et des contraintes d'égalité (tel que vu à la Proposition 2.1.14). Ainsi, un problème satisfaisant (MFCQ) ne contient aucun sous-ensemble de gradients de contraintes positivement linéairement dépendants, d'où (CPLD) trivialement satisfait. □

Théorème 2.1.30. *Soit un problème de la forme* (NLP)*. Si (CRCQ) est satisfaite en un point x^*, alors (CPLD) est aussi satisfaite en ce point.*

Démonstration. (CRCQ) est une hypothèse qui implique que pour tout sous-ensemble de gradients des contraintes d'inégalité actives ou d'égalité, il existe un voisinage de rang constant. En particulier, en prenant le même voisinage, le rang constant implique que si le sous-ensemble des gradients sélectionné est positivement linéairement dépendant en x^*, il n'est pas de plein rang et restera ainsi dans ce voisinage, d'où (CPLD) satisfaite. □

On apporte ici un exemple qui montre que (CPLD) n'implique ni (MFCQ), ni (CRCQ).

Exemple 2.1.31. Soit le programme d'optimisation à contraintes non linéaires (illustré à la Figure 2.7) :

$$\begin{cases} & \min \quad z = (x_1+1)^2 + (x_2+1)^2 \\ \text{s.à.} & g_1(x) = x_1 - x_2 \leq 0 \\ & g_2(x) = -x_1 + x_2 \leq 0 \\ & g_3(x) = x_1^2 - x_2 \leq 0 \\ & g_4(x) = -x_2 \leq 0. \end{cases}$$

Le domaine réalisable est le segment de droite $x_1 = x_2$ tel que $0 \leq x_1 \leq 1$. Le point optimal est en $x^* = (0,0)^t$. (MFCQ) n'est pas satisfaite en x^* puisque les contraintes $g_1(x)$ et $g_2(x)$ forment ensemble une contrainte d'égalité (comme noté dans la Remarque 2.1.16). On remarque aussi que (CRCQ) n'est pas satisfaite en x^* car si l'on prend le sous-ensemble de contraintes $g_3(x)$ et $g_4(x)$, leurs gradients sont dépendants le long de $x_1 = 0$ et indépendants ailleurs, donc de rang non constant dans tout voisinage de x^*.

Vérifions maintenant que (CPLD) est satisfaite en x^*. Nous rappelons qu'en vertu de la Proposition 2.1.14, il existe $\lambda_1, \lambda_2, \lambda_3, \lambda_4 \geq 0$ tels que :

$$\begin{pmatrix} \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 1 \\ 0 & -1 \\ 0 & -1 \end{pmatrix} = 0$$

d'où on déduit que $\lambda_1 = \lambda_2$ et $\lambda_3 = \lambda_4 = 0$. Pour vérifier (CPLD), il faut donc montrer que les gradients des contraintes dont le multiplicateur est non nul (donc $g_1(x)$ et $g_2(x)$ dans ce cas-ci) sont toujours positivement linéairement dépendants en un voisinage de x^*, ce qui est trivial puisque $\nabla g_1(x) = -\nabla g_2(x)$ pour tout $x \in \mathbb{R}^2$. Ainsi, (CPLD) est satisfaite.

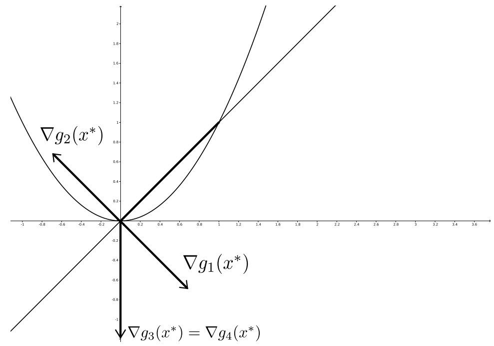


Figure 2.7 – (CPLD) satisfaite, mais (CRCQ) et (MFCQ) échouent en x^*. Ceci est dû au fait que les contraintes actives ne satisfaisant pas (CRCQ) ne sont pas les mêmes que les contraintes actives ne satisfaisant pas (MFCQ).

2.1.6 QNCQ

Définition 2.1.32. Soit le problème (NLP). Un point réalisable x^* satisfait à la qualification des contraintes de quasinormalité (QNCQ) si, pour tout ensemble $\{\lambda, \mu\} \subset \mathbb{R}^m \times \mathbb{R}^p, \lambda \geq 0$ tel que

$$\sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum_{j=1}^{p} \mu_j \nabla h_j(x^*) = 0$$

il n'existe pas de suite de points y_k telle que $\lim_{k\to\infty} y_k = x^*$ et $\forall i, j$ t.q. $\lambda_i, \mu_j \neq 0$,

$$\lambda_i g_i(y_k) > 0, \mu_j h_j(y_k) > 0.$$

De manière intuitive, cette qualification nécessite qu'en un point optimal, pour tout sous-ensemble des contraintes actives ne respectant pas (LICQ), on ne peut pas construire une suite de points qui n'est pas réalisable pour chacune de ces contraintes séparément. L'exemple suivant illustre bien cette qualification.

Exemple 2.1.33. Soit le programme non linéaire (illustré à la Figure 1.1) :

$$\begin{cases} & \min \quad z = -x_1 \\ \text{s.à.} & g_1(x) = (x_1 - 1)^3 + x_2 \leq 0 \\ & g_2(x) = -x_1 \leq 0 \\ & g_3(x) = -x_2 \leq 0. \end{cases}$$

Il s'agit de l'Exemple 1.2.4 où toutes les qualifications échouent. Ainsi, on devrait avoir (QNCQ) non satisfaite en $x^* = (1, 0)$. En effet, prenons $\lambda_1 = \lambda_3$. On obtient que

$$\sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) = 0.$$

Or, prennons la suite de points $\{y_k\}_{k\in\mathbb{N}^*}$ telle que :

$$y_k = \left(1 + \frac{1}{k}, -\frac{1}{2}\left(\frac{1}{k}\right)^3\right)^t.$$

On remarque que $g_1(y_k) = \frac{1}{2k^3} > 0$ et $g_3(y_k) = \frac{1}{2k^3} > 0$ et (QNCQ) n'est pas satisfaite.

On considère maintenant le problème :

$$\begin{cases} & \min \quad z = -x_1 \\ \text{s.à.} & g_1(x) = -|(x_1 - 1)^3| + x_2 \leq 0 \\ & g_2(x) = -x_1 \leq 0 \\ & g_3(x) = -x_2 \leq 0. \end{cases}$$

Il s'agit de l'exemple précédent avec pour seule différence le terme cubique de la première contrainte qui est en valeur absolue. Le quatrième quadrant devient donc entièrement réalisable par rapport à la contrainte $g_1(x)$. On observe alors que le problème satisfait (QNCQ). En effet, pour que $\lambda_3 g_3(y_k) > 0$, on doit avoir la deuxième composante de y_k strictement négative. Similairement, pour avoir $\lambda_1 g_1(y_k) > 0$, on doit avoir la deuxième composante de y_k strictement positive, d'où la non-existence de la dite suite.

Dans ce dernier programme, on remarque que (LICQ), (MFCQ), (CRCQ) et (CPLD) ne sont pas satisfaites, contrairement à (KTCQ).

Théorème 2.1.34. *[AMS04] Soit un problème de la forme* (NLP). *Si la qualification des contraintes (CPLD) est satisfaite en un point x^*, alors (QNCQ) est aussi satisfaite en ce point.*

On en déduit évidemment que (MFCQ) et (CRCQ) impliquent tous deux la quasinormalité. L'exemple suivant montre que la contraposée n'est toutefois pas vérifiée.

Exemple 2.1.35. [AMS04] Soit le programme d'optimisation à contraintes non linéaires :

$$\begin{array}{rl} \min & z = x_1^2 + (x_2+1)^2 \\ \text{s.à.} & h(x) = x_2 e^{x_1} = 0 \\ & g(x) = x_2 \leq 0. \end{array}$$

Le minimum de ce problème se situe en $x^* = (0,0)^t$. La matrice (2.1) s'écrit

$$\begin{pmatrix} \nabla h(x) \\ \nabla g(x) \end{pmatrix} = \begin{pmatrix} x_2 e^{x_1} & e^{x_1} \\ 0 & 1 \end{pmatrix}$$
$$\begin{pmatrix} \nabla h(x^*) \\ \nabla g(x^*) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

Le domaine réalisable de ce problème est l'axe des x_1. Ainsi, on trouve que pour $\mu = -\lambda$ avec $\lambda \geq 0$, on a $\mu \nabla h(x^*) + \lambda \nabla g(x^*) = 0$. On observe aussi que $h(x)$ et $g(x)$ ont le même signe pour tout $x \in \mathbb{R}^2$. Ainsi, $\mu h(x)$ est de signe contraire à $\lambda g(x)$ et les deux ne peuvent être tous deux positifs. Il n'existe donc pas de suite $\{y_k\}_{k \in \mathbb{N}^*}$ nécessaire pour faire échouer (QNCQ), donc cette dernière est satisfaite.

Toutefois, (MFCQ) n'est pas satisfaite en x^* car les seuls vecteurs tels que $h(x^*)v = 0$ sont aussi tels que $g(x^*)v = 0 \not< 0$, ou encore il n'existe pas de vecteur tangent à $h(x^*)$ entrant dans l'intérieur du domaine réalisable défini par $\nabla g(x^*)$. Aussi, on observe que dès que l'on quitte l'axe des x_1, $\nabla h(x)$ est linéairement indépendant à $\nabla g(x)$. Le point x^* est donc un point ne respectant pas (MFCQ) et tel que $\nabla g(x)$ et $\nabla h(x)$ ne restent pas positivement linéairement dépendants dans tout voisinage de x^*, d'où l'échec de (CPLD).

Une intuition présentée précédemment énoncait que cette qualification nécessite qu'en un point optimal, il est impossible de construire une suite de points qui n'est pas réalisable pour chacune des contraintes séparément. Pour les contraintes d'égalité, on remarque une restriction supplémentaire sur le signe de la contrainte (et du gradient de la contrainte). On peut observer ceci en remplaçant $g(x)$ de l'exemple précédent par $h_2(x) = x_2 = 0$.

Bien qu'il existe alors des suites de points non réalisables, les signes de μ_1 et μ_2 assurent que $\mu_1 h_1(x^*)$ est de signe contraire à $\mu_2 h_2(x^*)$, ce qui entraîne que (QNCQ) est satisfaite.

2.1.7 Abadie CQ et Guignard CQ

Revenons au concept de définir une condition d'optimalité. Au début de ce chapitre, on a énoncé qu'il n'y avait aucune direction réalisable de descente en un minimum local. Cette intuition s'écrit formellement par l'expression $-\nabla f(x) \in \mathcal{T}(x)^\circ$. Cette intuition indique que toute direction réalisable se retrouve à faire augmenter ou stagner la linéarisation de la fonction objectif. Les conditions de KKT, quant à elles, reconstruisent $\nabla f(x)$ à l'aide de la linéarisation des contraintes actives, c'est-à-dire $-\nabla f(x) \in \mathcal{F}(x)^\circ$. Dans cet optique, une qualification des contraintes serait une hypothèse qui assure que $\mathcal{T}(x^*)^\circ = \mathcal{F}(x^*)^\circ$. Ceci motive la qualification des contraintes de Guignard.

Définition 2.1.36. La qualification des contraintes de Guignard (GCQ) est satisfaite en un point réalisable x^* lorsque

$$\mathcal{T}(x^*)^\circ = \mathcal{F}(x^*)^\circ.$$

Théorème 2.1.37. *[Gui69] (GCQ) est une qualification des contraintes.*

Grâce à des techniques d'analyse fonctionnelle, on obtient que si les cônes $\mathcal{F}$ et $\mathcal{T}$ sont égaux, il en va de même de leur polaire. On en déduit la qualification des contraintes d'Abadie.

Définition 2.1.38. La qualification des contraintes d'Abadie (ACQ) est dite satisfaite en un point réalisable x^* lorsque

$$\mathcal{T}(x^*) = \mathcal{F}(x^*).$$

Un point satisfaisant la condition d'Abadie est dit quasirégulier.

Proposition 2.1.39. *Soit un problème de la forme* (NLP). *Si (ACQ) est satisfaite en un point x^*, alors (GCQ) est aussi satisfaite en ce point.*

Démonstration. Par propriété des cônes polaires : si C_1 et C_2 sont des cônes, alors

$$C_1 \subset C_2 \Rightarrow C_2^\circ \subset C_1^\circ.$$

Ainsi, si $\mathcal{T}(x^*) = \mathcal{F}(x^*)$, on en déduit

$$\mathcal{T}(x^*)^\circ = \mathcal{F}(x^*)^\circ.$$

□

L'exemple suivant montre qu'il est possible d'avoir (GCQ) sans avoir (ACQ). Nous détaillons cet exemple afin de montrer comment travailler avec les cônes $\mathcal{T}(x) = \mathcal{F}(x)$ et leurs polaires.

Exemple 2.1.40. [Bur12] Soit le programme d'optimisation non linéaire :

$$\begin{cases} & \min \quad z = x_1 + x_2 \\ \text{s.à.} & h(x) = x_1 x_2 = 0 \\ & g_1(x) = -x_1 \leq 0 \\ & g_2(x) = -x_2 \leq 0. \end{cases}$$

Remarque 2.1.41. La forme des contraintes de ce problème est assez connue. Les programmes affectés par ces contraintes sont appelés *Mathematical Program with Equilibrium Constraints* (MPEC), aussi parfois appelés *Mathematical Program with Complementarity Constraints* (MPCC). On peut noter deux variables x_1 et x_2 sujettes à ces contraintes par $0 \leq x_1 \perp x_2 \geq 0$.

L'ensemble réalisable est donc l'axe des x_1 positifs et l'axe des x_2 positifs. On s'intéresse au point optimal $x^* = (0,0)^t$. Le cône tangent au domaine réalisable de ce problème est composé de l'ensemble des directions d telles qu'il existe une suite $d_k \to d$ tel que pour θ_k

tendant vers 0, on a $x^* + \theta_k d_k$ réalisable. Le domaine réalisable étant deux demi-droites, on trouve

$$\mathcal{T}(x^*) = \{d | d_1 \geq 0, d_2 \geq 0, d_1 d_2 = 0\}.$$

Les directions de $\mathcal{T}(x^*)$ sont donc les directions suivant les axes positifs. Maintenant, calculons le cône linéarisé $\mathcal{F}(x^*)$. On a :

$$\begin{aligned} \nabla h(x) &= \begin{pmatrix} x_2 & x_1 \end{pmatrix} \\ \nabla h(x^*) &= \begin{pmatrix} 0 & 0 \end{pmatrix} \\ \nabla g(x) &= \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned}$$

On voit ainsi qu'en x^*, la contrainte $h(x)$ n'exerce aucune restriction supplémentaire quant aux directions du cône linéarisé. En effet, $\nabla h(x^*) d = 0$ pour tout d. Le cône linéarisé s'écrit donc :

$$\mathcal{F}(x^*) = \{d | d_1 \geq 0, d_2 \geq 0\}.$$

Autrement dit, le cône linéarisé $\mathcal{F}(x^*)$ est composé des directions contenues dans le premier quadrant.

Ainsi, on remarque $\mathcal{T}(x^*) \subset \mathcal{F}(x^*)$, mais $\mathcal{F}(x^*) \not\subset \mathcal{T}(x^*)$ d'où $\mathcal{T}(x^*) \neq \mathcal{F}(x^*)$ et (ACQ) n'est pas satisfaite.

Nous calculons à présent leur polaire selon la Définition 1.1.6. On peut récrire les directions de $\mathcal{T}(x^*)$ par $\theta_k \begin{pmatrix} 1 & 0 \end{pmatrix}$ ou $\theta_k \begin{pmatrix} 0 & 1 \end{pmatrix}$, avec θ_k positif.

$$\begin{aligned} \mathcal{T}(x^*)^\circ &= \{u | u^t d \leq 0, d = \theta_k \begin{pmatrix} 1 & 0 \end{pmatrix}\} \cup \{u | u^t d \leq 0, d = \theta_k \begin{pmatrix} 0 & 1 \end{pmatrix}\} \\ &= \{u | u_1 \theta_k \leq 0\} \cup \{u | u_2 \theta_k \leq 0\} \\ &= \{u | u_1 \leq 0, u_2 \leq 0\}. \end{aligned}$$

Le cône polaire au cône tangent $\mathcal{T}(x^*)$ est donc composé des vecteurs du troisième quadrant du plan $\mathbb{R}^2$.

Pour le cône linéarisé, on admet d avec $d_1, d_2 \geq 0$. On retrouve :

$$\begin{aligned}\mathcal{F}(x^*)^\circ &= \{u | u^t d \leq 0, d_1 \geq 0, d_2 \geq 0\} \\ &= \{u | u_1 d_1 \leq 0, u_2 d_2 \leq 0, d_1 \geq 0, d_2 \geq 0\} \\ &= \{u | u_1 \leq 0, u_2 \leq 0\}.\end{aligned}$$

On retrouve donc $\mathcal{F}(x^*)^\circ = \mathcal{T}(x^*)^\circ$ et (GCQ) est satisfaite. En effet, comme calculé, il est possible de générer n'importe quel vecteur du troisième quadrant par une combinaison linéaire positive des lignes de la matrice $\nabla g(x^*)$ (on ignore $\nabla h(x^*)$ car celui-ci est nul). Or, si la fonction admet son minimum en $x^* = (0,0)^t$, alors le gradient de $-f(x^*)$ doit nécessairement être orienté dans le troisième quadrant (ou être nul). En particulier, dans notre situation, on a :

$$\begin{aligned}\mathcal{L}(x^*, \lambda) &= \nabla f(x^*) + \lambda \nabla g(x^*) = 0 \\ &\implies \begin{pmatrix} 1 & 1 \end{pmatrix} + \begin{pmatrix} \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ -1 & 0 \\ 0 & -1 \end{pmatrix} = 0 \\ &\implies \lambda_1 \text{ libre}, \lambda_2 = 1, \lambda_3 = 1\end{aligned}$$

et il y a existence des multiplicateurs de Lagrange.

La proposition suivante montre où ces deux qualifications se situent par rapport aux autres citées précédemment.

Proposition 2.1.42. *[Ber95] Soit un problème de la forme* (NLP). *Si (QNCQ) est satisfaite en un point x^*, alors (ACQ) est aussi satisfaite en ce point.*

Par cette proposition, il est ainsi facile de voir que (CPLD), (MFCQ), (CRCQ) et (LICQ) impliquent toutes (ACQ), qui à son tour implique (GCQ). On peut aussi démontrer que (KTCQ) implique (ACQ).

Théorème 2.1.43. *[BSS93] Soit un problème de la forme* (NLP)*. Si (KTCQ) est satisfaite en un point x^*, alors (ACQ) est aussi satisfaite en ce point.*

Nous incluons ici un exemple pour lequel (ACQ) est respecté, mais pas (KTCQ). Prendre en note que cet exemple est plutôt dégénéré. Celui-ci ne s'écrit pas sous la forme (NLP) et, pour cette raison, certaines qualifications telles que (MFCQ) et (LICQ) n'ont pas les hypothèses nécessaires sur le domaine original X^0 pour être applicable (le domaine original de (NLP) est $\mathbb{R}^n$). L'exemple est simplement énoncé pour montrer que topologiquement, les qualifications d'Abadie et de Kuhn-Tucker ne sont pas équivalentes.

Exemple 2.1.44. [Pet73] On s'intéresse à un problème de minimisation sur $x \in X^0$ dont l'optimum est au point $(0,0)^t$ et où l'on définit

$$X^0 = \{x \in \mathbb{R}^2 | x_1 < 0 \text{ ou } |x_2| \geq x_1^2 \text{ et } \nexists n \in \mathbb{Z}, n \neq 0 \text{ tel que } x_2 = x_1 \arctan(\frac{1}{n\pi})\}.$$

Le demi-plan $x_1 \leq 0$ est entièrement réalisable. Dans le demi-plan $x_1 > 0$, on considère d'abord non réalisables les points situés entre les paraboles $x_1^2 = x_2$ et $-x_1^2 = x_2$. Aux points réalisables restants, nous retirons toutes les demies-droites dont une extrémité est l'origine (non incluse) et orientées vers les quadrants 1 et 4 de pentes $\pm \arctan(\frac{1}{n\pi})$. Ici, le cône linéarisé sera le domaine x^0 au complet car il n'y a pas de contraintes (donc pas de linéarisation des contraintes). On remarque donc que (ACQ) est satisfaite car les cônes tangents et linéarisés sont tous deux X^0. On remarque aussi que (KTCQ) n'est pas satisfaite car le vecteur $(1,0)^t$ ne possède pas d'arc différentiable, ce qui est nécessaire à la satisfaction de (KTCQ).

Remarque 2.1.45. On peut voir que, si l'on exige que $f(x)$ soit une fonction de classe $\mathcal{C}^1$, alors les seules fonctions admettant $x^* = (0,0)^t$ comme minimum local avec ce domaine réalisable sont les fonctions admettant x^* comme minimum local dans $\mathbb{R}^2$. Une manière analogue d'écrire cette affirmation est de dire que $\nabla f(x^*) = (0,0)$ et $\lambda, \mu = 0$. Ainsi, dans un contexte d'optimisation, cet exemple est peu intéressant puisqu'aucune qualification des contraintes n'est nécessaire si $\nabla f(x^*) = 0$. Il permet toutefois d'observer les subtilités entre (ACQ) et (KTCQ).

2.2 Problème convexe

Les qualifications présentées dans la section précédente sont applicables pour tout problème de la forme (NLP). On s'intéresse à présent sur un problème de la forme (NLP) où les fonctions f et g_i sont convexes et h_i linéaires. Cette section présente une qualification connue et utilisée dans le cas des problèmes d'optimization convexes. Ses relations avec les autres qualifications ainsi que certaines propriétés des qualifications ci-haut dans le cas convexe sont aussi présentées.

2.2.1 Condition de Slater

Définition 2.2.1. La condition de Slater (SLCQ) est respectée pour le problème (NLP) avec $g(x)$ convexes et $h(x)$ linéaires s'il existe un point $\bar{x}$ dans l'intérieur relatif du domaine réalisable, c'est-à-dire

$$\begin{aligned} h(\bar{x}) &= 0 \\ g(\bar{x}) &< 0. \end{aligned}$$

Remarque 2.2.2. On exige que les contraintes h soient linéaires (ou équivalentes à des contraintes linéaires) puisque dans le cas contraire, un domaine réalisable non vide ne peut pas être convexe.

Remarque 2.2.3. Il est possible d'affaiblir davantage cette qualification des contraintes. En effet, on peut séparer les $g_i(x)$ linéaires des $g_i(x)$ non linéaires. La qualification de

Slater s'écrirait donc :

$$h(\bar{x}) = 0$$
$$g_i(\bar{x}) < 0 \text{ pour } g_i \text{ non linéaires}$$
$$g_i(\bar{x}) \leq 0 \text{ pour } g_i \text{ linéaires.}$$

Bien que cet affaiblissment soit intéressante en soi, sa structure analytique s'en retrouve complexifiée et non essentielle pour l'analyse faite dans ce mémoire.

Cette (QC) se situe à une position relative comparable à celle de (LICQ) par rapport aux autres qualifications.

Théorème 2.2.4. *Soit un problème de la forme* (NLP). *Si (SLCQ) est satisfaite pour le problème, alors (MFCQ) est satisfaite en tout point réalisable* $\bar{x}$.

Démonstration. Soit y un point strictement réalisable. On a l'existence du point y en vertu de la condition de Slater. Ainsi, les points $\bar{x} + v$ avec $v = \kappa(y - \bar{x}), \kappa \in [0, 1]$ sont tous réalisables en vertu de la convexité, c'est-à-dire $g(\bar{x} + v) \leq 0$. En fait, on peut même développer une boule ouverte $B(y)$ autour de y dont tous les points sont réalisables. Ainsi, toujours en vertu de la convexité, si on note $\xi \in B(y) \cap \{x | h_j(x) = 0\}$ et $\zeta = \gamma\bar{x} + (1-\gamma)\xi$ pour $0 \leq \gamma \leq 1$, on a $g(\zeta) \leq 0$. Cet ensemble de points ζ forme un cône tronqué dont $\bar{x} + v$ est dans son intérieur relatif, ce qui permet de déduire que $g(\bar{x} + v) < 0$.

Nous rappelons une propriété intéressante des fonctions convexes, appelée *Inégalité du gradient*, c'est-à-dire pour une fonction f convexe, on a

$$f(x) \geq f(y) + \nabla f(y)(x - y).$$

Cette propriété suit d'un développement de Taylor tronqué à l'ordre deux. Étant un résultat classique d'analyse convexe, la démonstration est ici omise.

Il reste à appliquer cette propriété au problème précédent. En vertu de l'inégalité du gradient, on obtient :

$$g(\bar{x}+v) \geq g(\bar{x}) + \nabla g(\bar{x})v$$
$$g(\bar{x}+v) - g(\bar{x}) \geq \nabla g(\bar{x})v.$$

Or, $g(\bar{x}+v) < 0$ et $g(\bar{x}) = 0$ car g est une contrainte active. On en déduit que

$$0 > g(\bar{x}+v) > \nabla g(\bar{x})v$$

d'où le résultat. □

Remarque 2.2.5. Le théorème précédent est faux si on écrit Slater sous sa forme relaxée présentée dans la Remarque 2.2.3. Pour s'en convaincre, il suffit de prendre un problème similaire à l'Exemple 2.1.15, mais en utilisant des contraintes linéaires (voir l'Exemple 2.2.11). L'échec vient de la possibilité d'avoir une contrainte d'égalité divisée en deux contraintes d'inégalités. Dans ces situations où il est connu que (MFCQ) échoue, la première forme de la condition de Slater échoue aussi, alors que sa deuxième forme est satisfaite.

Ce résultat prend son sens lorsque l'on voit (MFCQ) comme une hypothèse exigeant l'existence d'un vecteur passant de x^* à un autre point dans l'intérieur relatif du domaine réalisable linéarisé. On remarque toutefois que la suffisance n'est pas vérifiée puisque (SLCQ) nécessite que les fonctions g soient convexes et que les fonctions h soient linéaires, alors que ce n'est pas le cas pour (MFCQ) (ou LICQ). L'exemple suivant montre que (SLCQ) n'implique pas (LICQ), nous permettant ainsi de bien situer (SLCQ).

Exemple 2.2.6. [Pet73] Soit le programme non linéaire :

$$\begin{cases} & \min \quad z = -x \\ \text{s.à.} & g_1(x) = x^2 + x \leq 0 \\ & g_2(x) = x^2 + 2x \leq 0. \end{cases}$$

Le minimum de ce programme est en $x^* = 0$. On remarque que (SLCQ) est satisfaite puisque g_1 et g_2 sont toutes deux convexes (donc l'ensemble réalisable l'est aussi) et que le point $x = \frac{-1}{2}$ est dans l'intérieur relatif de celles-ci. Toutefois, deux contraintes sont actives au point optimal dans un espace de dimension un, d'où l'impossibilité que (LICQ) soit satisfaite.

On peut toutefois observer que (LICQ) avec convexité implique (SLCQ) et similairement pour (MFCQ). La démonstration de ce dernier est très similaire à celle du Théorème 2.2.4.

La Figure 2.8 fait ainsi les liens entre les différentes qualifications développées jusqu'à maintenant pour un problème qui n'est pas nécessairement convexe. Nous verrons plus tard pourquoi (SLCQ) n'est pas en relation avec (CRCQ).

Relations entre les qualifications

Cette section cherche à identifier quelles qualifications des contraintes sont équivalentes dans le cas convexe.

Parlons d'abord de (KTCQ) et d'(ACQ).

Proposition 2.2.7. *En programmation convexe, la condition d'Abadie et la qualification des contraintes de Kuhn-Tucker sont équivalentes.*

Démonstration. Dans le cas convexe, l'arc différentiable est tel que pour tout point d'$\alpha(\theta)$, on peut relier ce point à x^* par une droite. Donc les arcs α peuvent tous être pris linéaires. (KTCQ) s'écrit donc : pour tout z tel que $\nabla g_{i\in I^*}(x^*)z \leq 0$ et $\nabla h(x^*)z = 0$, alors z est tangent à un vecteur réalisable, d'où on trouve l'équivalence avec (ACQ). $\square$

Ensuite, dans le même optique, on parlera d'(ACQ)et de (GCQ) en observant comment

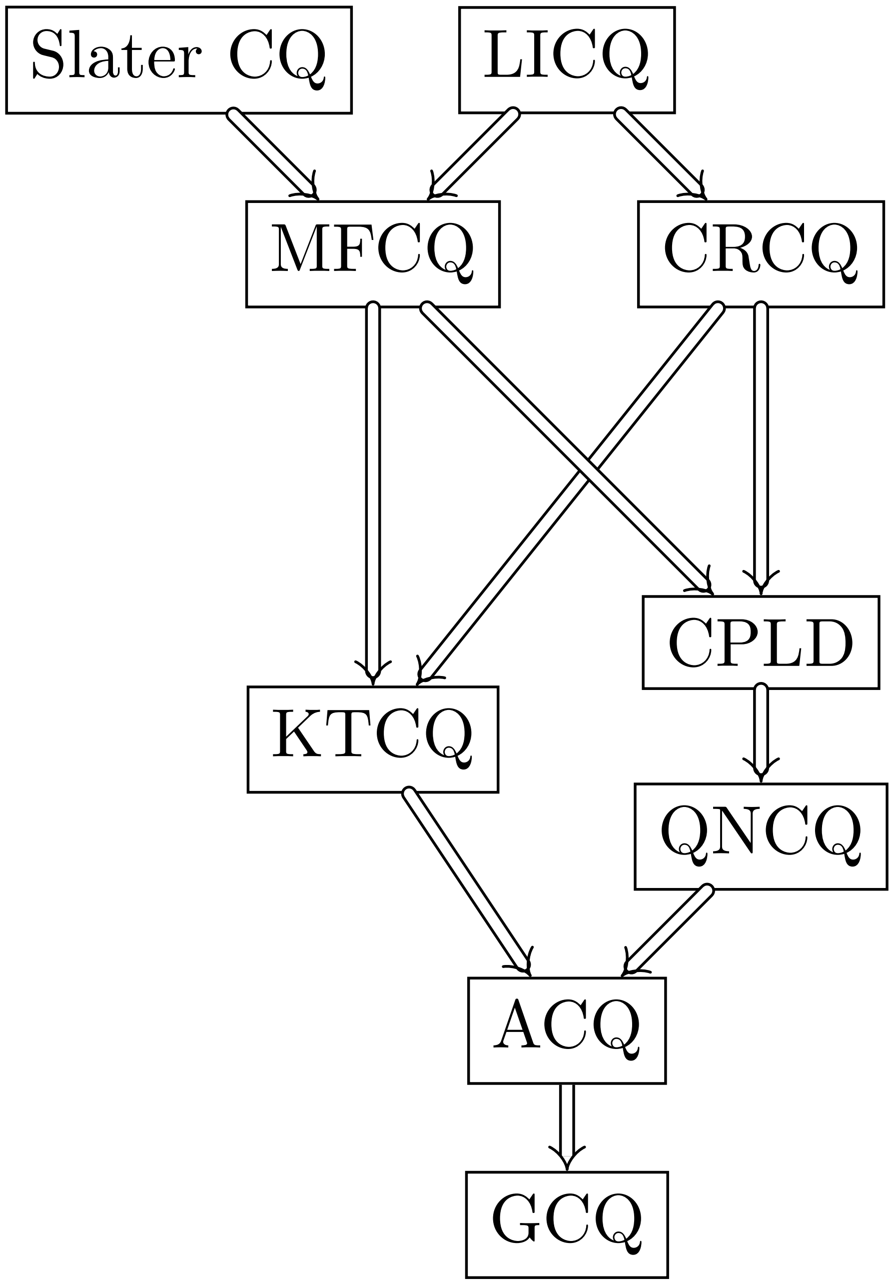


Figure 2.8 – Arbre d'implications des qualifications des contraintes de premier ordre présentées pour un problème pas nécessairement convexe.

les cônes réagissent entre eux. Les problèmes convexes admettent certaines particularités intéressantes au niveau des qualifications des contraintes. En guise d'exemple, il est possible de remarquer une équivalence sur les cônes et leur polaire, admettant ainsi une équivalence entre (ACQ) et (GCQ).

Théorème 2.2.8. *Si l'ensemble réalisable est convexe, alors $\bar{\mathcal{T}} = \mathcal{T}$, c'est-à-dire que $\mathcal{T}$ est un ensemble fermé.*

Corollaire 2.2.9. *En programmation convexe, la condition d'Abadie est équivalente à la condition de Guignard.*

Par la suite, on fait remarquer que (LICQ) et (MFCQ) ne sont pas équivalentes, de même que (LICQ) et (CRCQ). L'exemple suivant le montre bien.

Exemple 2.2.10. Soit le problème d'optimisation non linéaire de $\mathbb{R}^3$:

$$\begin{cases} & \min \quad z = -x_3 \\ \text{s.à.} & g_{1,2}(x) = (x_1 \pm 1)^2 - x_3 - 2 \geq 0 \\ & g_{3,4}(x) = (x_2 \pm 1)^2 - x_3 - 2 \geq 0 \\ & g_5(x) = -x_3 \leq 0. \end{cases}$$

Il s'agit de l'intérieur d'une sorte de pyramide à base carrée dont les murs sont de forme parabolique. Ce problème admet sa solution optimale en $(0,0,1)^t$. L'ensemble des multiplicateurs de ce problème s'écrit $\Lambda((0,0,1)) = \{(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5) | \lambda_1 = \lambda_2, \lambda_3 = \lambda_4, \lambda_5 = 0\}$.

On remarque d'abord qu'au point optimal, quatre contraintes se rencontrent dans un espace de trois dimensions, rendant (LICQ) certainement insatisfaite. On remarque toutefois que le rang de la matrice jacobienne est de trois sur un voisinage de x^*, d'où (CRCQ) est satisfaite. On observe aussi que le vecteur $v = (0,0,1)$ est tel que $\nabla g_i v < 0, i = 1,2,3,4$ d'où (MFCQ) est satisfaite.

Le problème précédent est concave. Il est toutefois très aisé de montrer que le problème convexe associé, soit :

$$\begin{cases} & \min \quad z = x_3 \\ \text{s.à} & -(x_1 \pm 1)^2 + x_3 + 2 \leq 0 \\ & -(x_2 \pm 1)^2 + x_3 + 2 \leq 0 \\ & x_3 \leq 0 \end{cases}$$

possède les même propriétés en son point optimal $x^* = (0, 0, -1)$.

De même, on note que la pyramide à base carrée dans $\mathbb{R}^3$ est telle que (MFCQ) et (CRCQ) sont respectées, mais pas (LICQ).

On observera que (MFCQ) et (CRCQ) ne sont toujours pas reliés dans le cas convexe. De plus, suite à l'architecture de l'arbre des qualifications que nous avons déjà développé, démontrer ce résultat nous permettra aussi de justifier que :

— (CRCQ) $\not\Leftrightarrow$ (MFCQ),

— (CRCQ) $\not\Leftrightarrow$ (CPLD) et,

— (MFCQ) $\not\Leftrightarrow$ (CPLD).

Exemple 2.2.11. Prenons le problème d'optimisation linéaire :

$$\begin{cases} & \min \quad z = (x_1 + 1)^2 + (x_2 + 1)^2 \\ \text{s.à.} & g_1(x) = x_1 - x_2 \leq 0 \\ & g_2(x) = -x_1 + x_2 \leq 0 \\ & g_3(x) = -x_1 \leq 0. \end{cases}$$

Il s'agit d'une variante à contraintes linéaires du problème 2.1.15. Le constat est toujours le même : (MFCQ) échoue car une contrainte d'égalité est écrite sous la forme de deux contraintes d'inégalité, laissant l'intérieur strict du domaine réalisable vide par rapport aux contraintes d'inégalité. On remarque aussi, tout comme dans l'Exemple 2.1.15, que (CRCQ) est satisfaite.

Remarque 2.2.12. Pour une forme alternative des contraintes, (MFCQ) pourrait être respectée, tel que mis en évidence par les résultats de [Lu11].

Le contre-exemple 2.1.20 est un problème convexe, justifiant que (MFCQ) n'implique pas (CRCQ) dans le cas convexe.

Le contre-exemple convexe suivant, provenant de [BO02], justifie qu'(ACQ) n'implique pas la quasinormalité (QNCQ). On en déduit donc qu'(ACQ) n'implique pas (CPLD) non plus.

Exemple 2.2.13.

$$\begin{cases} & \min \quad z = x_1^2 + (x_2 - 1)^2 \\ \text{s.à.} & g_1(x) = (x_1 - 1)^2 + x_2^2 - 1 \leq 0 \\ & g_2(x) = (x_1 + 1)^2 + x_2^2 - 1 \leq 0 \\ & h(x) = x_2 \leq 0. \end{cases}$$

(ACQ) est satisfaite, car le cône linéarisé est le vecteur $(0,0)^t$ et le cône tangent aussi. En fait, le point $x = (0,0)^t$ est le seul point réalisable. On remarque toutefois que (QNCQ) n'est pas satisfaite car pour les contraintes g_1 et g_2, la suite de point $\{y\}_k = (0, 1/k)$ n'est pas réalisable pour chacune des contraintes et tend vers x^*.

Il reste à vérifier si (QNCQ) et (CPLD) sont équivalentes. Au meilleur de nos connaissances, aucun contre-exemple ni démontration n'ont été trouvés.

L'arbre des qualifications des contraintes ainsi développé dans le cas convexe est illustré dans la Figure 2.9.

Le livre de Rockafellar [Roc70] est une excellente source d'informations sur les problème convexes. D'autres références pertinentes incluent le livre de Bertsekas [Ber95] et les notes de Nemirovski [Nem13].

2.3 Conditions approximées et qualifications strictes

Comme il a été vu précédemment, les exemples qui ne satisfont pas aux qualifications des contraintes sont souvent des cas isolés, un peu exceptionnels en soi, où le voisinage

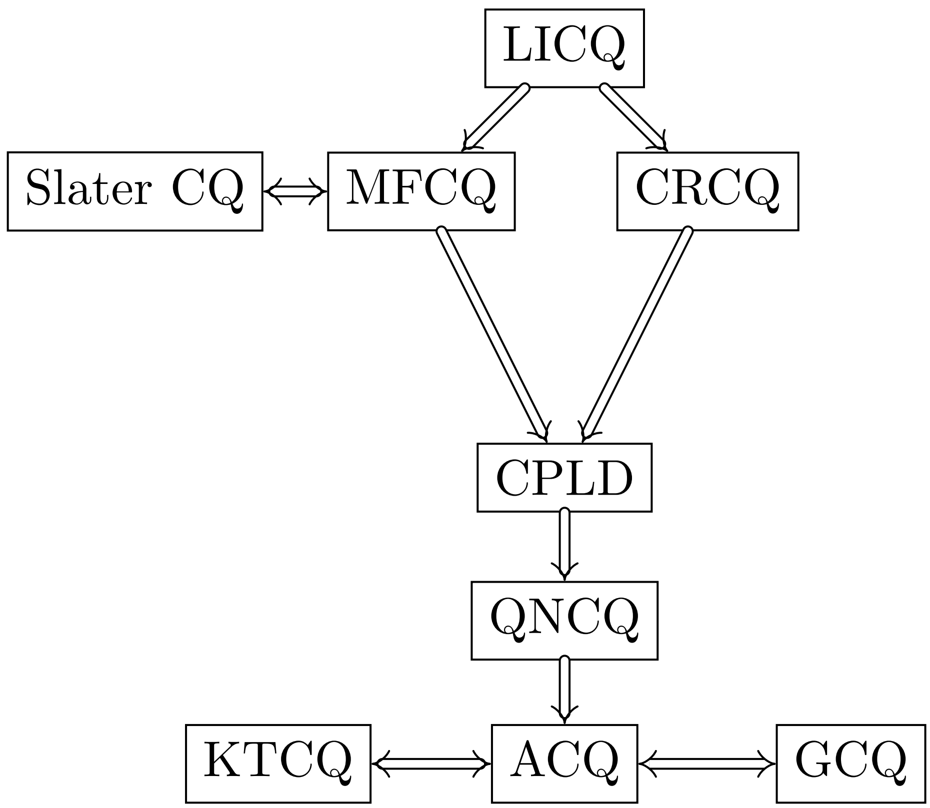


Figure 2.9 – Arbre d'implications des qualifications des contraintes de premier ordre présentées pour un problème convexe.

du point optimal est irrégulier. Ceci nous amène à vouloir définir des conditions d'optimalité qui dépendent non seulement du point optimal, mais aussi de son voisinage. Nous présentons dans cette section une de ces conditions et ses relations avec les différentes qualifications présentées précédemment.

Condition de KKT approximée

La section suivante portera sur une relaxation des conditions d'optimalité de (KKT).

Définition 2.3.1. [AHM11] [AHRS15] Soit un problème de le forme (NLP). La condition nécessaire d'optimalité de KKT approximée (AKKT) est satisfaite en un point x^* minimum local s'il existe des suites de points $\{x^k \in \mathbb{R}^n\}$, $\{\lambda^k \in \mathbb{R}^m\}$ et $\{\mu^k \in \mathbb{R}^p\}$ telles que :

$$\begin{aligned}
&\lim_{k\to\infty} x^k = x^*,\\
&\lim_{k\to\infty} \nabla f(x^k) + \lambda^k \nabla g(x^k) + \mu^k \nabla h(x^k) = 0,\\
&\lim_{k\to\infty} \lambda^k g(x^k) = 0, \lambda^k \geq 0.
\end{aligned}$$

Cela signifie qu'il existe des suites de points (pour le problème primal et le problème dual) telles que leur limite satisfont aux conditions de KKT. Le résultat suivant est des plus intéressants dans notre objectif de définir des conditions d'optimalité.

Théorème 2.3.2. *[AHM11] AKKT est une condition nécessaire d'optimalité ne nécessitant aucune qualification des contraintes. Autrement dit, si x^* est un minimum local de* (NLP)*, alors AKKT est satisfaite.*

Exemple 2.3.3. Revenons à l'Exemple 1.2.4. Puisque le point $(1, 0)^t$ est un minimum local, nous devrions normalement avoir AKKT satisfaite. En effet, on rappelle que notre lagrangien s'écrit :

$$\begin{aligned}\nabla\mathcal{L}(\bar{x},\lambda,\mu) &= \nabla f(\bar{x}) + \sum_{i=1}^{m}\lambda_i \nabla g_i(\bar{x}) \\ &= \begin{pmatrix}-1 & 0\end{pmatrix} + \begin{pmatrix}\lambda_1 & \lambda_3\end{pmatrix}\begin{pmatrix}3(\bar{x}_1-1)^2 & 1 \\ 0 & -1\end{pmatrix} = 0\end{aligned}$$

d'où on tire les équations :

$$\begin{aligned}-1+3\lambda_1(x_1-1)^2 &= 0 \\ \lambda_1-\lambda_3 &= 0.\end{aligned}$$

De la deuxième équation, on déduit $\lambda_1=\lambda_3$. Ainsi, en prenant par exemple la suite :

$$\{x^k\}_{k\in\mathbb{N}} = \left(1-\frac{1}{2}^k, \frac{1}{2}(x_1^k-1)^3\right)^t \tag{2.6}$$

on obtient une suite satisfaisant AKKT, ayant comme suite de multiplicateurs associés :

$$\{\lambda^k\}_{k\in\mathbb{N}} = \left(\frac{1}{3(x_1^k-1)^2}, 0, \frac{1}{3(x_1^k-1)^2}\right). \tag{2.7}$$

La suite de points x^k utilisée ici est une suite de points réalisables. Toutefois, la réalisabilité des points de la suite n'est pas obligatoire.

Qualifications des contraintes strictes

Maintenant, la relation entre AKKT et KKT sera étudiée à l'aide de qualifications des contraintes.

Définition 2.3.4. Soit le problème (NLP). Une qualification des contraintes stricte (QCS) est une hypothèse qui assure que AKKT implique KKT.

Quelques nuances existent entre les conditions d'optimalité utilisant AKKT et celles utilisant KKT. Le Tableau 2.1 résume ces conditions nécessaires d'optimalité appliquées au problème (NLP).

Tableau 2.1 – Conditions nécessaires d'optimalité

x^* min local	$\Rightarrow$	AKKT	AKKT	$\not\Rightarrow$	x^* min local
x^* min local + QC	$\Rightarrow$	KKT	KKT	$\not\Rightarrow$	x^* min local (avec/sans QC)
AKKT + QCS	$\Rightarrow$	KKT	AKKT + QC	$\not\Rightarrow$	KKT

La dernière ligne du Tableau 2.1 nous indique que certains points satisfaisant les conditions d'AKKT mais sujets à une qualification des contraintes non stricte ne satisfont pas les conditions de KKT. Ainsi, l'ensemble des points satisfaisant AKKT est plus grand que l'ensemble des points respectant KKT. Toutefois, sous une (QC), tous les minima locaux impliquent KKT. Donc les points supplémentaires satisfaisant AKKT et sujets à une qualification non stricte sont des points stationnaires qui ne sont pas des minima (donc ce sont des maxima ou des points d'inflexion, ce qui est indésirable dans notre étude). On doit chercher ces points parmi ceux satisfaisant AKKT, mais qui ne sont pas sous une qualification stricte. L'exemple suivant nous montre un de ces points.

Exemple 2.3.5. Soit le programme non linéaire :

$$\begin{cases} & \min \quad z = -x_1 \\ \text{s.à.} & h(x) = x_1 x_2 = 0 \\ & g_1(x) = -x_1 - x_2 \leq 0. \end{cases}$$

Il s'agit d'une version modifiée du problème (2.1.40). Le point $(0,0)^t$ est ici un maximum local. En fait, tous les points réalisables $x_1 = 0, x_2 \geq 0$ sont des maxima locaux. Toutefois, parmi ces maxima, seul $(0,0)^t$ n'est pas un minimum, puisque la direction $d = (1,0)^t$ est réalisable et de descente. Pour ce problème, notre lagrangien s'écrit :

$$\begin{pmatrix} -1 & 0 \end{pmatrix} + \begin{pmatrix} \mu & \lambda \end{pmatrix} \begin{pmatrix} x_2 & x_1 \\ -1 & -1 \end{pmatrix} = 0$$

qui se réduit aux deux équations :

$$\begin{cases} -1 + \mu x_2 - \lambda &= 0 \\ \mu x_1 - \lambda &= 0. \end{cases}$$

On considère une suite :

$$\{x^k\}_{k\in\mathbb{N}} = \left(0, \frac{1}{k}\right)^t. \tag{2.8}$$

De la deuxième équation, on déduit $\lambda = 0$. Puis, de la première équation, on a $\mu = \frac{1}{x_2}$. Ainsi, la suite 2.8 satisfait aux conditions (AKKT) et tend vers $(0,0)^t$, un point qui n'est pas un minimum local et qui ne satisfait pas les conditions d'optimalité de (KKT). Ceci démontre qu'il existe des points satisfaisant (AKKT) et respectant une qualification de contraintes qui ne sont pas des minima, alors que les conditions de (KKT) ne proposent pas ces points "indésirables".

Les qualifications (QNCQ), (ACQ) et (GCQ) ne sont pas strictes. Les qualifications (LICQ), (MFCQ), (CRCQ) et (CPLD) sont strictes [AMRS15].

Conclusion du chapitre

Les qualifications présentées dans ce chapitre sont les qualifications des contraintes de premier ordre ayant le plus d'intérêt. Plusieurs autres ont été développées dans la littérature. Parmi celles-ci, on note la seconde qualification d'Abadie, ou encore la qualification d'Arrow-Hurwitz-Uzawa [Pet73]. Ces deux qualifications n'ont pas été détaillées dans ce mémoire puisqu'elles ne consistent qu'à démontrer que pour un certain cône C tel que $\mathcal{F} \subset C$, on a $C = \mathcal{T}$. Dit autrement, elles impliquent toutes la première qualification d'Abadie (ACQ) présentée dans ce chapitre. Il existe aussi des versions relaxées de (CPLD) (notée RCPLD [AHMS12]) et (CRCQ) (notée RCRCQ [MS11]). Celles-ci ne sont pas explicitées ici puisqu'elles ont été moins développées dans la littérature et sont moins impliquées dans le développements d'algorithmes ou dans la compréhension de la géométrie de l'ensemble réalisable.

Ces résultats et exemples proviennent de différentes sources. Certains exemples sont des exemples classiques du domaine, comme l'Exemple 1.2.4 ou encore la pyramide de $\mathbb{R}^3$ 2.2.10. D'autres sont des exemples originaux, les plus intéressants étant 2.1.8, 2.1.33 et 2.3.5. Les exemples accompagnés d'une référence sont extraits de la littérature.

CHAPITRE 3

Qualifications des contraintes de second ordre

Nous revenons maintenant aux conditions nécessaires de (KKT) de second ordre (KKT2), introduits à la page 10, que nous rappelons ici :

Rappel 3.0.6. Les conditions nécessaires de second ordre de (KKT) stipulent qu'en un minimum local x^* satisfaisant *certaines hypothèses*, il existe des vecteurs $\lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^p$ tels que :

$$\begin{gathered} \nabla \mathcal{L}(x^*, \lambda, \mu) = \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum_{j=1}^{p} \mu_j \nabla h_j(x^*) = 0 \\ \lambda_i \geq 0, \lambda_i g_i(x^*) = 0 \quad i = 1, \ldots, m, j = 1, \ldots, p. \end{gathered} \tag{KKT}$$

De plus,

$$d^t \nabla^2_{xx} \mathcal{L}(x^*, \lambda, \mu) d \geq 0 \quad \forall d \in C(x^*). \tag{KKT2}$$

Il a été vu que quelques hypothèses supplémentaires, les qualifications des contraintes de premier ordre, sont essentielles pour avoir le résultat (KKT). On peut alors se demander si ces hypothèses sont suffisamment fortes pour garantir les conditions nécessaires de second ordre (KKT2) en un minimum local.

3.1 Décomposition et définitions

Il est possible que les multiplicateurs λ_i ou μ_j soient nuls même si leur contrainte associée est active. Il est aussi vraisemblable que la matrice (2.1) ne soit pas de rang n, laissant ainsi certaines directions dans lesquels les multiplicateurs λ, μ ne donnent pas d'information. C'est dans ces situations qu'il nous faut aller vérifier les conditions d'optimalité de second ordre.

L'équation (KKT2) est simple à interpréter si l'on est sous une qualification forte comme (LICQ). En effet, l'unicité des multiplicateurs sous (LICQ) implique aussi l'unicité de la matrice $\nabla^2_{xx}\mathcal{L}(x^*, \lambda, \mu)$. D'autre part, si (LICQ) échoue, alors l'ensemble des multiplicateurs Λ ne sera vraisemblablement pas un singleton et on se retrouve alors avec plusieurs matrices $\nabla^2_{xx}\mathcal{L}(x^*, \lambda, \mu)$ qui satisfont aux conditions d'optimalité de premier ordre. C'est à ce moment qu'il nous sera nécessaire de clarifier nos exigences en ce qui a trait à l'équation (KKT2).

Définition 3.1.1. La condition nécessaire (forte) d'optimalité de second ordre est satisfaite en x^* pour les multiplicateurs (λ, μ) si :

— les conditions de (KKT) sont satisfaites et

— $\forall d \in C(x^*), d^t \nabla^2 \mathcal{L}(x^*, \lambda^*, \mu^*)_{xx} d \geq 0.$

Il est à noter que la condition d'optimalité de second ordre est intrinsèquement reliée aux multiplicateurs avec lesquels on la teste. Ceci signifie que la condition d'optimalité de second ordre peut être satisfaite et non satisfaite simultanément en x^* ; il faut préciser quels multiplicateurs la rendent satisfaite et quels multiplicateurs la rendent non satisfaite. Pour alléger le texte, on adopte la convention que la condition de second ordre échoue en x^* sans préciser de multiplicateurs si elle échoue pour tous les multiplicateurs satisfaisant (KKT). À cet effet, nous expliciterons lors de la présentation des qualifications et des exemples si les conditions des Définitions 3.1.1 et 3.1.2 sont satisfaites pour un certain multiplicateur (λ, μ) ou pour l'ensemble des multiplicateurs satisfaisant (KKT).

Définition 3.1.2. La condition faible d'optimalité de second ordre est respectée en x^* si, pour les multiplicateurs (λ, μ) associés aux conditions de premier ordre, on a :

— les conditions de (KKT) sont satisfaites et

— $\forall d \in c(x^*), d^t \nabla^2 \mathcal{L}(x^*, \lambda^*, \mu^*)_{xx} d \geq 0.$

La principale différence entre la condition nécessaire forte et la condition nécessaire faible est l'information retenue des contraintes d'inégalités faiblement actives. La condition faible considère uniquement les directions tangentes aux inégalités faiblement actives, alors que la condition forte considère les directions du demi-espace de descente de ces contraintes. On en déduit trivialement la proposition suivante.

Proposition 3.1.3. *[AMS07] La condition de stricte complémentarité assure que les cônes $C(x^*) = c(x^*)$ ou encore que les conditions 3.1.1 et 3.1.2 sont équivalentes.*

Voyons maintenant un exemple (initialement proposé par Anitescu) montrant que l'on a effectivement besoin de qualifications des contraintes de second-ordre plus restrictives que les qualifications de premier ordre pour déduire une condition nécessaire de second ordre en un minimum local.

3.2 Exemple d'Anitescu

Exemple 3.2.1. [Ani99]

On construit d'abord les matrices $Q_k, k = 0, \ldots 3$ définies par :

$$Q_k = \begin{pmatrix} \cos \frac{k\pi}{4} & -\sin \frac{k\pi}{4} \\ \sin \frac{k\pi}{4} & \cos \frac{k\pi}{4} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} \cos \frac{k\pi}{4} & \sin \frac{k\pi}{4} \\ -\sin \frac{k\pi}{4} & \cos \frac{k\pi}{4} \end{pmatrix}.$$

En fait, il s'agit des quatre états possibles de la matrice $\begin{pmatrix}1 & 0\\ 0 & -2\end{pmatrix}$ sujette aux rotations d'angles multiples de $\frac{\pi}{4}$. Les explicitant, on obtient :

$$\begin{aligned}
Q_0 &= \begin{pmatrix}1 & 0\\ 0 & -2\end{pmatrix},\\
Q_1 &= \frac{1}{2}\begin{pmatrix}-1 & 3\\ 3 & -1\end{pmatrix},\\
Q_2 &= \begin{pmatrix}-2 & 0\\ 0 & 1\end{pmatrix},\\
Q_3 &= \frac{1}{2}\begin{pmatrix}-1 & -3\\ -3 & -1\end{pmatrix}.
\end{aligned}$$

On considère le programme d'optimisation non linéaire :

$$\begin{cases} & \min \quad x_3 \\ \text{s.à.} & g_k(x) = \begin{pmatrix}x_1 & x_2\end{pmatrix} Q_k \begin{pmatrix}x_1 & x_2\end{pmatrix}^t - x_3 \leq 0 \quad k = 0,\ldots,3. \end{cases}$$

En explicitant, on trouve les contraintes :

$$\begin{aligned}
g_0(x) &= x_1^2 - 2x_2^2 - x_3\\
g_1(x) &= \frac{-1}{2}(x_1^2 - 6x_1x_2 + x_2^2) - x_3\\
g_2(x) &= -2x_1^2 + x_2^2 - x_3\\
g_3(x) &= \frac{-1}{2}(x_1^2 + 6x_1x_2 + x_2^2) - x_3
\end{aligned}$$

et leur gradients :

$$\begin{aligned}
\nabla g_0(x) &= \begin{pmatrix}2x_1 & -4x_2 & -1\end{pmatrix}\\
\nabla g_1(x) &= \begin{pmatrix}-x_1 + 3x_2 & -x_2 + 3x_1 & -1\end{pmatrix}\\
\nabla g_2(x) &= \begin{pmatrix}-4x_1 & 2x_2 & -1\end{pmatrix}\\
\nabla g_3(x) &= \begin{pmatrix}-x_1 - 3x_2 & -x_2^2 - 3x_1 & -1\end{pmatrix}
\end{aligned}$$

d'où, en $x^* = (0,0,0)^t$, $\nabla g_0(x^*) = \nabla g_1(x^*) = \nabla g_2(x^*) = \nabla g_3(x^*) = (0,0,-1)$. Puisque $\nabla f(x) = (0,0,1)$, on trouve que $\sum\limits_{i=0\ldots3} \lambda_i = 1$. Ainsi, Λ étant un ensemble borné non vide, (MFCQ) est satisfaite. On observe que le cône critique de ce problème est donné par les vecteurs de $\mathbb{R}^3$ tel que $x_3 = 0$.

Un petit calcul nous amène au résultat que :

$$\nabla_{xx}\mathcal{L}(x,\lambda) = 0_{3\times 3} + \begin{pmatrix} \sum\limits_{i=0\ldots 3} \lambda_i Q_i & 0 \\ 0 & 0 \end{pmatrix}.$$

Supposons que pour un choix de $\lambda \in \Lambda$, le hessien du lagrangien $\nabla_{xx}\mathcal{L}(x,\lambda)$ est semi-défini positif sur les directions du cône critique. Donc :

$$\begin{pmatrix} x_1 & x_2 & 0 \end{pmatrix} \begin{pmatrix} \sum\limits_{i=0\ldots 3} \lambda_i Q_i & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 & x_2 & 0 \end{pmatrix}^t \geq 0.$$

Cette condition est équivalente à montrer que

$$\begin{pmatrix} x_1 & x_2 \end{pmatrix} \sum_{i=0\ldots 3} \lambda_i Q_i \begin{pmatrix} x_1 & x_2 \end{pmatrix}^t \geq 0.$$

Or la matrice de la somme des $\lambda_i Q_i$ s'explicite :

$$\begin{pmatrix} \lambda_0 - 2\lambda_1 - \frac{1}{2}(\lambda_1 + \lambda_3) & \frac{3}{2}(\lambda_1 - \lambda_3) \\ \frac{3}{2}(\lambda_1 - \lambda_3) & -2\lambda_0 + \lambda_2 - \frac{1}{2}(\lambda_1 + \lambda_3) \end{pmatrix}.$$

La trace de cette matrice est $-(\lambda_0 + \lambda_1 + \lambda_2 + \lambda_3) = -1$ en vertu de nos conditions de (KKT) de premier ordre. Or, la trace d'une matrice est la somme de ses valeurs propres. La trace est négative, alors au moins une valeur propre est négative, et le hessien du lagrangien n'est pas semi-défini positif. Il s'agit donc d'un problème pour lequel (MFCQ) est satisfaite, mais les conditions nécessaires de second ordre échouent pour tout λ respectant nos conditions de premier ordre en un minimum local.

Ainsi, cet exemple nous montre que, pour assurer les conditions nécessaires de second ordre, les qualifications des contraintes de premier ordre ne sont pas suffisamment fortes. Nous étudierons donc des qualifications suffisamment fortes pour garantir les conditions nécessaires de second ordre.

3.3 Définitions

Définition 3.3.1. Soit le problème (NLP) où les fonctions f, g, h sont deux fois différentiables. Une qualification des contraintes de second ordre est une hypothèse permettant de démontrer que, en un minimum local x^* de $f(x)$ contraint par $g(x)$ et $h(x)$, les conditions de (KKT) et de (KKT2) sont satisfaites.

Le suite du chapitre présente quelques qualifications du second ordre connues, soit les qualifications (LICQ), (MMF), (CRCQ) et (SOKTCQ).

3.4 LICQ

Théorème 3.4.1. *La qualification des contraintes de l'indépendance linéaire (définition 2.1.3) est une qualification des contraintes de second ordre.*

Démonstration. L'unicité des λ, μ assure que le hessien du lagrangien est unique et donne l'intuition que la série de Taylor tronquée à l'ordre deux nous amènerait au résultat. L'essentiel de la preuve revient donc à développer une série de Taylor sur le lagrangien. Voir Nocedal et Wright pour la démonstration complète [NW06] (p.332). □

L'Exemple 3.2.1 montre bien que (MFCQ) n'est pas une qualification de second ordre. Nous ajoutons donc quelques conditions pour avoir une variante de (MFCQ) assurant les conditions nécessaires de second ordre en un minimum local.

3.5 MMF

Définition 3.5.1. Soit le problème (NLP), avec les fonctions f, g, h de classe $\mathcal{C}^2$ et x^* un point réalisable. La qualification des contraintes de second ordre de Magassarian-Fromovitz modifiée (MMF) est dite satisfaite en x^* lorsque :

MMF-1 (MFCQ) est satisfaite ;

MMF-2 le rang de la matrice (2.1) est $p + q - 1$ où q est la cardinalité de I^* ;

MMF-3 pour des multiplicateurs (λ, μ) donnés, au plus un des λ_i est nul ($||I_0|| \leq 1$).

Remarque 3.5.2. L'item (MMF-3) n'est pas inclus dans la définition initiale de (MMF), introduite par Baccari [Bac04]. Toutefois, la démonstration que (MMF) est une qualification de second ordre a besoin de cette hypothèse. Il n'y a toutefois aucune preuve, au meilleur de nos connaissances, que (MMF-3) est essentielle. À elles seules, les hypothèses (MMF-1) et (MMF-2) n'ont pas été démontrées comme étant une qualification des contraintes d'ordre deux.

Voici un exemple qui démontre que l'hypothèse (MMF-2) est essentielle à ce que (MMF) soit une qualification d'ordre deux. Cet exemple a été construit par Baccari.

Exemple 3.5.3. [Bac04] On considère le problème :

$$\left\{\begin{array}{ll} & \min \quad x_3 \\ \text{s.à.} & g_1(x) = 2\sqrt{3}x_1x_2 - 2x_2^2 - x_3 \leq 0 \\ & g_2(x) = x_2^2 - 3x_1^2 - x_3 \leq 0 \\ & g_3(x) = -2\sqrt{3}x_1x_2 - 2x_2^2 - x_3 \leq 0. \end{array}\right.$$

Le lagrangien de ce problème s'écrit :

$$\mathcal{L}(x, \lambda) = x_3 + \lambda_1(2\sqrt{3}x_1x_2 - 2x_2^2 - x_3) + \lambda_2(x_2^2 - 3x_1^2 - x_3) + \lambda_3(-2\sqrt{3}x_1x_2 - 2x_2^2 - x_3).$$

On cherche les points stationnaires du lagrangien. Ainsi, le gradient du lagrangien s'écrit :

$$\begin{aligned}\nabla\mathcal{L}(x^*.\lambda) = (0,0,1) &+ \lambda_1(2\sqrt{3}x_2, 2\sqrt{3}x_1 - 4x_2, -1) \\ &+ \lambda_2(-6x_1, 2x_2, -1) \\ &+ \lambda_3(-2\sqrt{3}x_2, -2\sqrt{3}x_1 - 4x_2, -1)\end{aligned}$$

d'où on déduit, pour les points stationnaires du lagrangien, la forme :

$$0 = \lambda_1 2\sqrt{3}x_2 - \lambda_2 6x_1 - \lambda_3 2\sqrt{3}x_2 \tag{3.1}$$

$$0 = \lambda_1(2\sqrt{3}x_1 - 4x_2) + \lambda_2 2x_2 + \lambda_3(-2\sqrt{3}x_1 - 4x_2) \tag{3.2}$$

$$0 = 1 - \lambda_1 - \lambda_2 - \lambda_3. \tag{3.3}$$

On observe à présent que $x^* = (0,0,0)^t$ est le minimum global de ce problème. À ce point, les équations (3.1) et (3.2) sont satisfaites pour toute valeur de $\lambda_1, \lambda_2 \in \mathbb{R}$. Ainsi, l'ensemble des multiplicateurs de Lagrange à ce minimum local est :

$$\Lambda(x^*) = \{\lambda \in \mathbb{R}^3 | \lambda_i \geq 0, \sum \lambda_i = 1, i = 1,2,3\}.$$

On remarque que (MMF-1), ou (MFCQ), est satisfaite en x^*. En effet, il existe $v = (v_1, v_2, v_3)^t$ tel que $\nabla g_i(x)v < 0, i = 1,2,3$ puisque

$$\begin{aligned}\nabla g_1(x^*)v &= -v_3 \\ \nabla g_2(x^*)v &= -v_3 \\ \nabla g_3(x^*)v &= -v_3.\end{aligned}$$

On peut donc prendre $v = (0,0,1)^t$ pour montrer que (MFCQ) est satisfaite.

On cherche maintenant à vérifier la condition nécessaire d'optimalité de second ordre (KKT2) :

$$d^t \nabla^2 \mathcal{L}(x^*, \lambda) d \geq 0 \quad , \forall d \in C(x^*).$$

Pour ce faire, calculons le hessien du lagrangien

$$\nabla^2\mathcal{L}(x^*,\lambda) = \begin{pmatrix} -6\lambda_2 & 2\sqrt{3}(\lambda_1-\lambda_3) & 0 \\ 2\sqrt{3}(\lambda_1-\lambda_3) & 2(\lambda_2-2\lambda_1-2\lambda_3) & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

ainsi que le cône critique en x^*

$$\begin{aligned} C(x^*) &= \{d|\nabla f(x^*)^T d \le 0, \nabla g_i(x^*)^T d \le 0, i=1,2,3\} \\ &= \left\{ \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} |d_3 = 0 \right\}. \end{aligned}$$

Choisissons $d_1 = (1,0,0)^t$ et $d_2 = (0,1,0)^t$ comme directions. On remarque que $d_1, d_2 \in C(x^*)$. Nous vérifions maintenant s'il existe des multiplicateurs de Lagrange $\lambda_1, \lambda_2, \lambda_3$ tels que la condition de second ordre est satisfaite pour ces deux directions.

Pour d_1, la condition (KKT2) s'écrit :

$$\begin{aligned} d_1^t\nabla^2\mathcal{L}(x^*,\lambda)d_1 &= \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} -6\lambda_2 & 2\sqrt{3}(\lambda_1-\lambda_3) & 0 \\ 2\sqrt{3}(\lambda_1-\lambda_3) & 2(\lambda_2-2\lambda_1-2\lambda_3) & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \\ &= -6\lambda_2. \end{aligned}$$

D'où $d_1^t\nabla^2\mathcal{L}(x^*,\lambda)d_1 \ge 0$ si et seulement si $\lambda_2 \le 0$.

Maintenant, pour la direction d_2, la condition (KKT2) s'écrit :

$$\begin{aligned} d_2^t\nabla^2\mathcal{L}d_2 &= \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} -6\lambda_2 & 2\sqrt{3}(\lambda_1-\lambda_3) & 0 \\ 2\sqrt{3}(\lambda_1-\lambda_3) & 2(\lambda_2-2\lambda_1-2\lambda_3) & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \\ &= 2(\lambda_2-2\lambda_1-2\lambda_3). \end{aligned}$$

Donc $d_2^t\nabla^2\mathcal{L}(x^*.\lambda)d_2 \ge 0$ si et seulement si $\lambda_2 \ge 2\lambda_1 + 2\lambda_3$, d'où $\lambda_2 > 0$ car $\sum \lambda_i = 1, i = 1,2,3$.

Chacune de ces directions requièrent des valeurs de λ_2 incompatibles. Donc, il n'existe aucun λ_2 respectant $d^t\nabla^2\mathcal{L}(x^*,\lambda)d \ge 0$ à la fois pour d_1 et d_2, donc la condition nécessaire

(KKT2) n'est pas satisfaite malgré que l'on se trouve en un point qui est un minimum local.

Ainsi, on arrive aux conclusions suivantes :

— le problème a un minimum global en $x^* = (0, 0, 0)^t$,

— (MFCQ) est satisfaite en x^*,

— Il n'existe pas de multiplicateurs de (KKT) tels que la condition de second ordre est satisfaite pour toute direction du cône critique.

Ainsi, (MFCQ) n'implique pas qu'il existe des multiplicateurs de Lagrange tels que (KKT2) sont satisfaites ou, dit autrement, (MFCQ) n'est pas une qualification des contraintes de second ordre.

Remarque 3.5.4. Comme vu dans l'Exemple 3.5.3, la difficulté survient lorsque nos multiplicateurs sont multiples. Ainsi, l'ensemble des directions dans le cône tangent aux contraintes qui respectent la condition $d^t\nabla^2\mathcal{L}(x^*, \lambda^*, \mu^*)d \geq 0$ peut varier selon les valeurs de λ et μ choisies

Notons que (MMF) reste une (QC) moins restrictive que (LICQ). Toutefois, (LICQ) n'implique pas (MMF) à cause de l'hypothèse (MMF-3). (LICQ) et la stricte complémentarité assurent que (MMF) est satisfaite.

On note aussi que (MMF) est une qualification des contraintes de second ordre certifiant que la condition d'optimalité de second ordre est satisfaite pour *un certain multiplicateur* (et non pas pour l'ensemble des multiplicateurs).

3.6 CRCQ

Théorème 3.6.1. *La qualification (CRCQ) telle que vue à la Section 2.1.3 est une qualification des contraintes de second ordre.*

Démonstration. Voir [AES10] pour les détails de la démonstration. L'idée de la preuve réside sur le fait que (CRCQ) implique la qualification de second ordre de Kuhn et Tucker. Ainsi il est possible, avec les conditions de (CRCQ), de construire un arc deux fois différentiable respectant les conditions de (SOKTCQ) (défini dans la section suivante). □

Remarque 3.6.2. Il est démontré dans [AES10] que la condition de second ordre (KKT2) sera satisfaite *pour tout* vecteur de multiplicateurs de Lagrange respectant (KKT) si (CRCQ) est satisfaite.

Remarque 3.6.3. Nous rappelons ici que (LICQ) implique (CRCQ). Ainsi, par le Théorème 3.6.1, il est maintenant trivial que (LICQ) soit une qualification de second ordre.

3.7 SOKTCQ

Cette qualification des contraintes a été développée par Kuhn et Tucker dans [KT51]. Il s'agit, tout comme (KTCQ) pour les conditions de premier ordre, de la première qualification des contraintes de second ordre à avoir été développée.

Définition 3.7.1. Soit le problème (NLP), avec les fonctions f, g, h de classe $\mathcal{C}^2$ et x^* un point réalisable. La qualification des contraintes de second ordre de Kuhn et Tucker (SOKTCQ) est satisfaite au point x^* si pour tout vecteur non-nul y tel que $\nabla g_i(x^*)y = 0, \quad i \in I^*$ et $h_j(x^*)y = 0, j = 1 \ldots p$, alors y est tangent à un arc $\alpha(\theta)$ deux fois différentiable tel que $\alpha(0) = x^*$ et qu'il existe $\epsilon > 0$ tel que $g_i(\alpha(\theta)) = 0 (i \in I^*)$ et $h_j(\alpha(\theta)) = 0, j = 1 \ldots p$ pour tout $\theta \in [0, \epsilon]$.

Intuitivement, cette (QC) signifie que pour tout vecteur non nul y conservant la linéarisation des contraintes actives en x^*, il existe un arc deux fois différentiable conservant ces contraintes actives et tangent à y.

On retrouve ici le successeur logique à (KTCQ), à la différence que l'on exige un arc deux fois différentiable qui conserve les contraintes actives.

Cette condition peut être affaiblie davantage, comme le fait remarquer l'auteur de [Fle87]. L'exemple suivant, pris de ce même ouvrage, illustre bien la situation.

Exemple 3.7.2. Soit le problème d'optimisation non linéaire :

$$\begin{cases} & \min \quad z = x_3 - \frac{1}{2}x_1^2 \\ \text{s.à.} & g_1(x) = x_3 + x_2 + x_1^2 \geq 0 \\ & g_2(x) = x_3 - x_2 + x_1^2 \geq 0 \\ & g_3(x) = x_3 \geq 0. \end{cases}$$

Le point $\bar{x} = (0,0,0)^t$ admet les conditions d'optimalité de premier ordre. En effet, nos équations de (KKT) s'écrivent (on modifie les équations pour avoir des contraintes $\leq$) :

$$\begin{pmatrix} -x_1 & 0 & 1 \end{pmatrix} + \begin{pmatrix} \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix} \begin{pmatrix} -2x_1 & -1 & -1 \\ -2x_1 & 1 & -1 \\ 0 & 0 & -1 \end{pmatrix} = 0.$$

On en déduit que toutes les combinaisons convexes des vecteurs $\lambda^1 = (\frac{1}{2}, \frac{1}{2}, 0)$ et $\lambda^2 = (0,0,1)$ sont des multiplicateurs admissibles. On remarque que $\bar{x}$ n'est pas un minimum local. En effet, la direction $d = (1,0,0)^t$ est une direction de descente réalisable en $\bar{x}$.

On observe ici que le seul point qui admet g_1, g_2 et g_3 comme contraintes actives est $\bar{x} = (0,0,0)^t$. Ainsi, la qualification de Kuhn-Tucker de second ordre va certainement échouer en ce point (car tous les points de l'arc deux fois différentiable $\alpha(\theta)$ doivent admettre g_1, g_2 et g_3 comme contraintes actives). En utilisant une qualification des contraintes de second ordre moins forte, on peut remarquer que les conditions de second ordre ne sont pas respectées en ce point. En effet,

$$\nabla^2_{xx}\mathcal{L}(\bar{x},\lambda) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} -2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} -2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \lambda_3 \; O_{3x3}.$$

La matrice $\nabla^2_{xx}\mathcal{L}(\bar{x},\lambda)$ est donc semi-définie négative (deux valeurs propres nulles et une valeur propre strictement négative, disons $-\xi, \xi > 0$). Il reste à vérifier que les directions du cône critique sont de courbure négative. La direction de descente $d = (1,0,0)^t$ est effectivement une direction du cône critique. Or :

$$d^t\nabla^2_{xx}\mathcal{L}(\bar{x},\lambda)d = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} -\xi & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = -\xi < 0.$$

La qualification de second ordre que Fletcher amenait avec cet exemple est similaire à (SOKTCQ), à l'exception que l'on force uniquement l'arc différentiable à garder actives les contraintes fortement actives. Formellement, plutôt que de forcer $g_i(\alpha(\theta)) = 0$ pour tout $i \in I^*$, on se restreint à forcer $g_i(\alpha(\theta)) = 0$ pour tout $i \in \hat{I}$.

3.8 Qualifications de type Abadie

Récemment, Andreani a publié un article [ABHS14] très intéressant portant sur des qualifications des contraintes de second ordre dites de type Abadie. J'en fais une revue rapide dans la cette section.

Rappel 3.8.1. L'ensemble des indices des contraintes fortement actives est $\hat{I}$ et l'ensemble des indices des contraintes faiblement actives est I_0.

Théorème 3.8.2. *[ABHS14, Th' 3.1] Soit un problème de la forme* (NLP)*. Si le système d'égalités*

$$g_i(x^*) = 0, \quad i \in I^*$$
$$h_j(x^*) = 0, \quad j = 1, \ldots, m$$

satisfait à la qualification des contraintes d'Abadie (ACQ), alors la condition nécessaire de second ordre faible est satisfaite en x^*.

Noter que ce résultat ne signifie pas qu'(ACQ) est une qualification des contraintes de second ordre. En effet, elle implique seulement la condition de second ordre faible. Grâce à ce résultat, on peut déduire le corrolaire suivant.

Corollaire 3.8.3. *[ABHS14, Cor. 3.1] Soit un problème sous la forme* (NLP) *avec uniquement des égalités (donc $I = \emptyset$). Si x^* est un minimum local satisfaisant (ACQ), alors la condition d'optimalité nécessaire forte de second ordre est satisfaite pour tout multiplicateur satisfaisant* (KKT)*.*

Ainsi, dans le cas d'un problème avec uniquement des égalités, la qualification d'Abadie est une qualification des contraintes de second ordre.

Dans un problème avec des contraintes d'inégalité, on ne peut pas arriver avec un résultat aussi fort. En effet, on sait que (MFCQ) implique (ACQ) et que (MFCQ) n'est pas une qualification des contraintes de second ordre. On peut toutefois décortiquer les contraintes d'inégalité pour avoir une hypothèse assez faible, comme présentée dans le livre de Bazaraa, Sherali et Shetty, qui donne la qualification de second ordre suivante.

Théorème 3.8.4. *[BSS93] Soit x^* un minimum local de* (NLP) *et $(\lambda, \mu) \in \Lambda$ des multiplicateurs* (KKT) *associés. Si le système d'inéquations :*

$$\begin{aligned} g_i(x^*) &= 0 \quad i \in \hat{I} \\ g_i(x^*) &\leq 0 \quad i \in I_0 \\ h_j(x^*) &= 0 \quad j = 1, \dots m \end{aligned}$$

satisfait à la qualification d'Abadie, alors la condition nécessaire forte de second-ordre est satisfaite pour le multiplicateur (λ, μ).

On en déduit évidemment que si tous les multiplicateurs satisfaisant (KKT) sont strictement positifs (donc la stricte complémentarité est satisfaite), alors ce dernier théorème prend uniquement en hypothèse que le système d'équations des plans tangents aux contraintes actives satisfasse (ACQ).

Conclusion

Ce chapitre est un survol des qualifications de second ordre et de leurs subtilités. Les démarches des Exemples 3.2.1 présenté par Anitescu et 3.5.3 présenté par Baccari sont détaillées davantage que dans leur articles originaux. Certains questionnements et subtilités sont relevés, mettant en place les connaissances nécessaires pour continuer un travail plus extensif sur les qualifications de second ordre.

CHAPITRE 4

Vers la différentiabilité de trajectoires centrales

Ce chapitre introduit quelques propriétés des trajectoires centrales issues d'algorithmes de barrière logarithmique sous différentes hypothèses.

L'idée initiale de ce chapitre était de démontrer l'existence, l'unicité et la différentiabilité de ces trajectoires. Ces résultats sont déjà démontrés si (LICQ) est satisfaite [FM90]. Intuitivement, nous pensons que (MFCQ) devrait être une qualification des contraintes suffisamment forte pour impliquer l'existence et l'unicité des trajectoires centrales primales et duales. L'intérêt d'utiliser (MFCQ) est qu'elle assure que l'ensemble des multiplicateurs optimaux est borné, empêchant ainsi les trajectoires centrales duales de diverger.

Bien que le résultat n'a pas été démontré, ce chapitre fait office des différents développements accomplis dans le but d'arriver à une preuve. Plus précisément, nous avons étudié des problèmes respectant certaines hypothèses, sans toutefois satisfaire (LICQ). Parmi ces hypothèses, on retrouve, entre autres, la convexité, la compacité du domaine réalisable et dual et les problèmes quadratiques-linéaires.

Considérons un problème de la forme (NLP) avec seulement des contraintes d'inégalité. De plus, on considère la stricte complémentarité ainsi que les conditions suffisantes d'optimalité de second ordre satisfaites.

Notre intérêt actuel est de trouver des particularités et des propriétés souhaitables aux trajectoires centrales issues de l'algorithme de barrière logarithmique lorsque (LICQ) n'est pas satisfaite.

4.1 Non-inversibilité de la matrice Jacobienne

Cette section cherche d'abord à identifier les difficultés à l'existence des trajectoires centrales lorsque (LICQ) n'est pas satisfaite.

L'approche habituelle d'étude de la trajectoire centrale repose sur l'inversibilité de la matrice jacobienne du lagrangien du problème de barrière logarithmique afin d'utiliser le théorème des fonctions implicites. Cette approche a été proposée, à l'origine, par Fiacco et McCormick [FM90].

Or, si (LICQ) n'est pas satisfaite, on peut séparer :

$$Q\nabla g_{I^*}(x^*) = \begin{pmatrix} A' \\ 0 \end{pmatrix} = \begin{pmatrix} \nabla g_{I_+^*}(x^*) \\ \nabla g_{I_0^*}(x^*) \end{pmatrix}$$

avec A' une matrice de plein rang, et Q est une matrice orthonormale.
La Jacobienne du système original s'écrit donc :

$$\begin{pmatrix} \nabla^2 f(x^*) + \sum \lambda_i \nabla^2 g_i(x^*) & \nabla g_{I_+^*}(x^*)^t & \nabla g_{I_0^*}(x^*)^t & \nabla g_J(x^*)^t \\ \Lambda \nabla g_{I_+^*}(x^*) & 0 & 0 & 0 \\ \Lambda \nabla g_{I_0^*}(x^*) & 0 & 0 & 0 \\ 0 & 0 & 0 & g_J(x^*) \end{pmatrix}$$

où on retrouve la troisième colonne nulle.

Ainsi, lorsque (LICQ) n'est pas satisfaite, la jacobienne du problème de barrière est singulière, nous empêchant d'utiliser le théorème des fonctions implicites.

4.2 Algorithme de descente

Cette section démontre que la trajectoire centrale issue de la barrière logarithmique, sous certaines conditions raisonnables, est un algorithme de descente.

Proposition 4.2.1. *Sous les hypothèses nécessaires à l'existence d'une unique trajectoire centrale, l'algorithme qui consiste à suivre la trajectoire centrale est un algorithme non croissant sur $f(x)$. De plus, si les fonctions $g(x)$ sont analytiques, alors l'algorithme est décroissant (ou algorithme de descente).*

Démonstration. Notre problème de barrière logarithmique s'écrit :

$$\min f(x) - \rho \sum_i \log(-g(x)). \tag{4.1}$$

On prend $\rho_0 > \rho_1$ deux paramètres de notre problème de barrière et on en déduit les minimiseurs $x(\rho_0)$ et $x(\rho_1)$ des problèmes associés. Puisque la trajectoire centrale existe et est unique, ces minimiseurs sont aussi uniques. Intuitivement, $x(\rho_1)$ est plus "près" de la solution optimale que $x(\rho_0)$ dans le sens qu'il est plus avancé sur le trajectoire centrale dans la direction du minimum x^* de $f(x)$ contraint par $g(x)$.

En vertu de l'optimalité des solutions $x(\rho_0)$ et $x(\rho_1)$, on déduit des problèmes de barrière logarithmique à partir des deux paramètres ρ_0 et ρ_1 :

$$f(x(\rho_0)) - \rho_0 \sum_i \log(-g_i(x(\rho_0))) \leq f(\bar{x}) - \rho_0 \sum_i \log(-g_i(\bar{x}))$$
$$f(x(\rho_1)) - \rho_1 \sum_i \log(-g_i(x(\rho_1))) \leq f(\tilde{x}) - \rho_1 \sum_i \log(-g_i(\tilde{x}))$$

pour tout $\bar{x}, \tilde{x}$ réalisables. À cet effet, on peut d'abord prendre $\bar{x} = x(\rho_1)$ pour obtenir :

$$
\begin{aligned}
f(x(\rho_0)) - \rho_0 \sum_i \log(-g_i(x(\rho_0))) &\leq f(x(\rho_1)) - \rho_0 \sum_i \log(-g_i(x(\rho_1))) \\
f(x(\rho_0)) - f(x(\rho_1)) &\leq \rho_0 \sum_i \left(\log(-g_i(x(\rho_0))) - \log(-g_i(x(\rho_1)))\right) \\
\sum_i \left(\log(-g_i(x(\rho_1))) - \log(-g_i(x(\rho_0)))\right) &\leq -\frac{1}{\rho_0} \left(f(x(\rho_0)) - f(x(\rho_1))\right). \qquad (4.2)
\end{aligned}
$$

Puis, en prenant $\tilde{x} = x(\rho_0)$, on obtient similairement :

$$
\begin{aligned}
f(x(\rho_1)) - \rho_1 \sum_i \log(-g_i(x(\rho_1))) &\leq f(x(\rho_0)) - \rho_1 \sum_i \log(-g_i(x(\rho_0))) \\
f(x(\rho_1)) - f(x(\rho_0)) &\leq \rho_1 \left(\sum_i \log(-g_i(x(\rho_1))) - \sum_i \log(-g_i(x(\rho_0)))\right). \qquad (4.3)
\end{aligned}
$$

Ainsi, en substituant dans (4.3) le terme de (4.2), on obtient :

$$
\begin{aligned}
f(x(\rho_1)) - f(x(\rho_0)) &\leq \rho_1 \left(\sum_i \log(-g_i(x(\rho_1))) - \sum_i \log(g_i(x(\rho_0)))\right) \\
&\leq \frac{\rho_1}{\rho_0}(f(x(\rho_1)) - f(x(\rho_0))) \\
(f(x(\rho_1)) - f(x(\rho_0))) \left(1 - \frac{\rho_1}{\rho_0}\right) &\leq 0. \qquad (4.4)
\end{aligned}
$$

Puisque $\rho_1 < \rho_0$, on a que $1 - \frac{\rho_1}{\rho_0}$ est positif. L'équation (4.4) se réduit donc à :

$$
f(x(\rho_1)) - f(x(\rho_0)) \leq 0
$$

d'où le résultat que, le long de la trajectoire lorsque ρ décroît vers zéro, la fonction $f(x)$ est non croissante.

Maintenant, pour montrer la décroissance, il faut s'intéresser à la situation où (4.4) est satisfaite avec égalité. En revenant dans les démarches, on observe que cette situation survient lorsque (4.2) et (4.3) sont toutes deux satisfaites avec égalité. Ceci signifie que $x(\rho_0)$ est solution du problème barrière de paramètre ρ_1 et réciproquement, que $x(\rho_1)$ est solution du problème barrière de paramètre ρ_0. De là, on déduit :

$$\sum_i \log(-g(x(\rho_0))) = \sum_i \log(-g(x(\rho_1))).$$

On y distingue alors deux cas :

— $x(\rho_1) = x(\rho_0)$, où encore que de ρ_0 à ρ_1, on ne change pas de x

— $\sum_i \log(-g(x(\rho)))$ est constante sur une trajectoire reliant deux points distincts $x(\rho_0)$ et $x(\rho_1)$.

Le deuxième cas signifie que la trajectoire centrale n'est pas uniquement déterminée. Par exemple, pour $\rho_0 < \rho < \rho_1$, on peut prendre $x(\rho)$ n'importe quel point de la portion de trajectoire entre ρ_0 et ρ_1. Ayant supposé l'unicité des trajectoires centrales (donc l'unicité des solutions de (4.1)), ce cas est impossible.

Le premier cas, quant à lui, signifie que la fonction de barrière est constante sur l'intervalle ρ_0 à ρ_1. Puisque cette fonction est analytique en ρ, elle restera constante pour tout $0 < \rho$, c'est-à-dire que l'algorithme restera sur le point $x(\rho_0)$ pour tout ρ et, ainsi, le point $x(\rho)$ est un minimum local de la fonction $f(x)$ non contraint. En effet, le gradient de l'équation (4.1) s'écrit :

$$\nabla f(x(\rho)) + \rho \sum_i \frac{1}{g_i((\rho))} \nabla g_i(x(\rho))$$

qui doit être nul en $(x(\rho), \rho)$ pour que $x(\rho)$ soit un point stationnaire de (4.1). Puisque $\sum_i \log(-g(x(\rho)))$ est constant et analytique entre ρ_0 et ρ_1, on a que $\rho \sum_i \frac{1}{g_i((\rho))} \nabla g_i(x(\rho)) = 0$ et ainsi, on doit avoir $\nabla f(x(\rho)) = 0$.

$\square$

4.3 Analyse du cône normal $\mathcal{N}(x^*)$

Une qualification des contraintes est, en quelques sortes, une garantie que la linéarisation des contraintes représente bien la structure locale des contraintes. Dans le cas où il y a plusieurs vecteurs de multiplicateurs satisfaisant (KKT), il est possible de déduire des relations intéressantes entre l'orientation de la linéarisation de l'objectif par rapport au cône tangent (ou normal) aux contraintes et la stricte complémentarité des multiplicateurs de Lagrange.

Proposition 4.3.1. *Soit x^* minimum local d'un problème de la forme* (NLP) *tel que (GCQ) est satisfaite en x^*. Si la stricte complémentarité est satisfaite pour un certain multiplicateur $\lambda^* \in \Lambda$, alors l'opposé du gradient de la fonction objectif est orienté vers l'intérieur relatif du cône normal en x^* :*

$$-\nabla f(x^*) \in ri(\mathcal{N}(x^*)).$$

Démonstration. (Voir [DM89] pour l'équivalent linéaire de cette proposition)

Puisque la qualification des contraintes de Guignard est admise, le cône normal aux contraintes en x^* est le même que le dual du cône tangent linéarisé. Nous travaillons donc avec le polaire du cône tangent linéarisé :

$$N(x^*) = \{y \in \mathbb{R}^m | y^T v \leq 0 \ \forall v \text{ tel que } \nabla g_{I^*}(x^*)v \leq 0\}.$$

Or, si $\nabla g_{I^*}(x^*)v \leq 0$, alors pour tout $\alpha \geq 0$ vecteur ligne de $\mathbb{R}^m$, $\alpha \nabla g_{I^*}(x^*)v \leq 0$. On en déduit donc que, peut importe le α utilisé, le vecteur $\alpha \nabla g_{I^*}(x^*)$ est dans le cône normal $N(x^*)$. En particulier, pour $\alpha = \lambda^*$ tous positifs et non nuls et $v \neq 0$, on a

$$\lambda^* \nabla g_{I^*}(x^*)v < 0.$$

Puisque $-\nabla f(x^*) = \lambda^* \nabla g_{I^*}(x^*)$, on obtient que

$$-\nabla f(x^*) \in ri(N(x^*)).$$

□

Proposition 4.3.2. *Soit x^* minimum local d'un problème de la forme* (NLP) *tel que (GCQ) est satisfaite en x^*. Si l'opposé du gradient de la fonction objectif est orienté vers l'intérieur relatif du cône normal en x^* :*

$$-\nabla f(x^*) \in ri(\mathcal{N}(x^*))$$

alors il existe un multiplicateur strictement complémentaire $0 < \lambda \in \Lambda$.

Démonstration. Posons $v = -\nabla f(x^*) \in ri(\mathcal{N}(x^*))$. La qualification de Guignard étant satisfaite, on a $\mathcal{N}(x^*) = N(x^*)$. Puisque v est dans l'intérieur relatif de $N(x^*)$, on en déduit qu'il existe $\epsilon > 0$ tel que pour tout v' tel que $||v - v'|| < \epsilon$, alors $v' \in ri(N)$.

Pour fins de contradiction, supposons maintenant que la stricte complémentarité n'est pas satisfaite en x^* pour tous les $\lambda \in \Lambda$ tels que (KKT) sont satisfaites. Ainsi, pour ces λ, il y a au moins un des $\lambda_i, i \in I^*$ qui est nul. En particulier, il y a une composante k qui est nulle pour tous $\lambda \in \Lambda$. Autrement, en choisissant $\lambda^1, \lambda^2 \in \Lambda$ tels que $\lambda^1_k = 0$ et $\lambda^2_k \neq 0$, on trouvera que $\hat{\lambda} = \frac{1}{2}(\lambda^1 + \lambda^2)$ est dans Λ et sa composante $\hat{\lambda}_k$ est strictement positive.

Puisque (KKT) sont satisfaites, alors on a que :

$$v = -\nabla f(x^*) = \sum_{i \neq k} \lambda_i \nabla g_i(x^*) + \lambda_k \nabla g_k(x^*).$$

Ce dernier terme λ_k est nul comme discuté ci-haut. Posons le vecteur $v' = v + \epsilon_v \nabla g_k(x^*)$ en s'assurant que $\epsilon_v ||\nabla g_k(x^*)|| < \epsilon$. On en déduit que $v' \in ri(N(x^*))$. De là, on construit aussi $v'' = v - \epsilon_v \nabla g_k(x^*)$. On en déduit aussi, pour les mêmes raisons, que $v'' \in ri(N(x^*))$.

Mais $v'' \in ri(N(x^*))$ signifie, par définition de $N(x^*)$, qu'il existe $\alpha \geq 0$ non tous nuls tel que $\alpha \nabla g_{I^*}(x^*) = v''$ d'où on déduit :

$$\sum_{i \neq k} \lambda_i \nabla g_i(x^*) - \epsilon_v \nabla g_k(x^*) = v'' = \sum_{i \neq k} \alpha_i \nabla g_i(x^*) + \alpha_k \nabla g_k(x^*)$$

$$\sum_{i \neq k} (\lambda_i - \alpha_i) \nabla g_i(x^*) = (\alpha_k + \epsilon_v) \nabla g_k(x^*).$$

Donc $\nabla g_k(x^*)$ est linéairement dépendant des $\nabla g_{i \neq k}(x^*)$.

Maintenant construisons un multiplicateur μ tel que $\mu_i > 0$ (donc $\mu_k > 0$) pour le vecteur v. Il est important que les autres multiplicateurs $\mu_i, i \neq k$ restent positifs. On écrit donc v :

$$\begin{aligned} v &= \sum_{i \neq k} \lambda_i \nabla g(x^*) + 0 \nabla g_k(x^*) \\ &= \sum_{i \neq k} (\lambda_i - (\lambda_i - \alpha_i)) \nabla g_i(x^*) + (0 + \alpha_k + \epsilon_v) \nabla g_k(x^*) \\ &= \sum_{i \neq k} \alpha_i \nabla g_i(x^*) + (\alpha_k + \epsilon_v) \nabla g_k(x^*). \end{aligned}$$

Donc $\mu_i = \alpha_i$, $\mu_k = \alpha_k + \epsilon_v$ est un multiplicateur satisfaisant (KKT) et il est strictement positif, d'où l'existence d'un multiplicateur respectant la stricte complémentarité, ce qui établit la contradiction.

□

Proposition 4.3.3. *Soit x^* minimum local d'un problème de la forme* (NLP) *tel que (LICQ) est satisfaite en x^*. On a que $-\nabla f(x^*) \in ri(\mathcal{N}(x^*))$ si et seulement si $\lambda \in \Lambda$ est strictement complémentaire.*

Démonstration. Trivial par (KKT) en x^* et l'unicité de λ. □

Proposition 4.3.4. *Soit x^* minimum local d'un problème de la forme* (NLP) *tel que (LICQ) n'est pas satisfaite en x^*, mais (MFCQ) l'est. Si Λ est un singleton, alors $\lambda \in \Lambda$ n'est pas strictement complémentaire.*

Démonstration. (LICQ) non satisfaite implique qu'il existe c_i, $i \in I^*$ non tous nuls tels que :

$$\sum_{i \in I^*} c_i \nabla g_i(x^*) = 0.$$

Par la Proposition 2.1.14 à la page 24, (MFCQ) satisfaite nous oblige à avoir au moins un de ces c_i négatifs. Or, Λ est un singleton si les λ satisfaisant (KKT) sont uniques.

On suppose à tort la stricte complémentarité pour l'unique $\lambda \in \Lambda$. Soit k l'indice tel que $\frac{\lambda_j}{c_j} = \delta_j, j \in I^*$ est minimal. Prenons $0 < \epsilon < \delta_k$. On a alors $\lambda_j - \epsilon c_j$ positif pour tout $j \in I^*$ et ainsi :

$$\nabla f(x^*) + \lambda \nabla g(x^*) - \epsilon c \nabla g(x^*) = 0$$
$$\nabla f(x^*) + (\lambda - \epsilon c) \nabla g(x^*) = 0$$

d'où on déduit $\lambda - \epsilon c \in \Lambda$, une contradiction à l'unicité de $\lambda \in \Lambda$.

□

4.4 Unicité de la trajectoire centrale pour des problèmes convexes

On s'attaque maintenant à la situation où le problème primal est convexe. Les problèmes quadratiques linéaires définis positif satisfont cette hypothèse.

Monteiro et Zhou [MZ98] ont fait une analyse du problème d'existence et d'unicité des trajectoires centrales dans le cas convexe avec des fonctions $f(x)$ et $g(x)$ analytiques. Les résultats qui suivent sont des adaptations de ces résultats sous des hypothèses un peu différentes puisque l'analycité des fonctions est une hypothèse assez forte.

Lemme 4.4.1. *Soit le problème de barrière logarithmique associé à un problème convexe sans contrainte linéaire :*

$$\min_x f(x) - \rho \sum_{i=1}^{m} \log(-g_i(x)).$$

On considère que l'espace réalisable est fermé et borné (compact), d'intérieur non vide (donc Slater est satisfaite) et on exige la stricte complémentarité. Alors pour tout $\rho > 0$, la solution de ce problème, notée $x(\rho)$, est uniquement déterminée.

Démonstration. On rappelle ici que Slater est équivalent à (MFCQ) dans le cas où tout est convexe. À cet effet, on en déduit que l'ensemble des multiplicateurs de Lagrange est non-vide et borné.

L'existence d'au moins une solution au problème de barrière logarithmique sous ces conditions est montré dans l'article de Monteiro ([MZ98] Th' 1). Celui-ci nécessite que l'hypothèse de Slater soit satisfaite et que l'intérieur du problème dual soit non vide, ce qui est garanti par la stricte complémentarité. Le problème de barrière n'est pas bien défini si la condition de Slater n'est pas satisfaite, car la non-existence de point strictement réalisable pour les inégalités implique la non-réalisabilité du problème barrière primal (les seuls points primal-réalisables sont tels que la barrière logarithmique est à l'infini à ces points). Il reste à montrer qu'il existe au plus une trajectoire centrale.

Dans ces conditions, il est possible de reformuler le problème de barrière logarithmique. Posons $V(\rho) = f(x(\rho))$ et $\mathcal{A}(x(\rho)) = g_{I^*}(x(\rho))$. Notre ensemble réalisable étant convexe et compact, on peut écrire :

$$\begin{cases} & \min_{g_J(x)<0} \; -\sum_{i \notin I^*} \log(-g_i(x)) \\ \text{s.à.} & f(x) = V(\rho) \\ & g_{I^*}(x) = \mathcal{A}(x(\rho)). \end{cases}$$

On suppose ainsi qu'il y a deux solutions à ce problème, disons $\bar{x}$ et $\tilde{x}$. On en déduit :

$$\sum_{i \notin I^*} \log(-g_i(\bar{x})) = \sum_{i \notin I^*} \log(-g_i(\tilde{x})).$$

Toutefois, puisque $g_i(x)$ est convexe, on trouve $-\sum_{i \notin I^*} \log(-g_i(x))$ strictement convexe, donc on ne peut avoir qu'une seule solution au problème précédent. Ainsi, $\bar{x} = \tilde{x}$. $\square$

Remarque 4.4.2. Monteiro et Zhou [MZ98] ont montré l'unicité en ne supposant ni les conditions suffisantes d'optimalité, ni la compacité de l'ensemble primal réalisable. Toutefois, ces affaiblissements ont obligé les auteurs à utiliser des fonctions $f(x)$ et $g(x)$ analytiques, une condition assez restrictive.

Corollaire 4.4.3. *Sous les mêmes hypothèses qu'au Lemme 4.4.1 et en ajoutant la compacité de l'espace réalisable dual, la trajectoire centrale duale est unique.*

Démonstration. Le problème dual auquel on s'attaque est :

$$\begin{cases} & \max_{\lambda>0} \; f(x) + g(x)^t\lambda + \rho\sum_i \log(\lambda_i) \\ \text{s.à.} & \nabla f(x) + \nabla g(x)^t\lambda = 0. \end{cases}$$

On évoque ici qu'il faudrait travailler sur les différentes conditions pour que l'ensemble dual soit compact afin que l'analyse soit complète.

Puisque l'espace réalisable dual est compact, on peut faire le même genre de transformation qu'au lemme précédent, ce qui donne la forme suivante :

$$\begin{cases} & \max_{\lambda_{I^*}>0} \; \sum_{i\in I^*} \log(\lambda_i) \\ \text{s.à.} & \nabla f(x) + \lambda\nabla g(x) = 0 \\ & f(x) + \lambda g(x) = \mathcal{D}(\rho) \\ & \lambda_J = \mathcal{B}(\rho) \end{cases}$$

où $\mathcal{B}(\rho) = \lambda(x(\rho))$ et $\mathcal{D}(\rho) = f(x(\rho)) + \lambda g(x(\rho))$. Bien entendu, dans le problème dual, ce sont les multiplicateurs des contraintes inactives qui tendent vers zéro à la solution optimale (plutôt que ce soient les contraintes actives qui tendent vers zéro). On comprend donc pourquoi le dual cherche à maximiser un problème sur les contraintes actives en fixant les contraintes inactives sur différents ensembles de niveau, dépendamment du paramètre ρ.

L'argument pour l'unicité de la solution est le même qu'au lemme précédent. En supposant deux solutions à ce dernier problème, on se retrouve avec une contradiction à la concavité stricte de la fonction log ce qui justifie qu'il existe au plus un seul maximum local, donc global.

□

Conclusion

La Proposition 4.2.1 a déjà été démontrée [Hua71, p.13], quoique la démonstration présentée ici est, au meilleur de nos connaissances, originale. Il serait possible, en utilisant les travaux de Pierre Huard sur la méthode des centres, d'affaiblir davantage les hypothèses pour obtenir une descente stricte. L'analyse du cône normal est aussi composée de démonstrations originales. Toutefois, les résultats présentés ont sans doute déjà été démontrés, possiblement autrement, puisqu'il s'agit d'analyse mathématique assez élémentaire. La démonstration utilisant la compacité 4.4.1 est aussi une démonstration originale, ainsi que son corollaire sur le dual. Beaucoup de travaux restent à être complétés à ce sujet, commençant par la différentiabilité des trajectoires centrales et l'analyse de leur point de convergence.

t

CONCLUSION

L'objectif principal était de travailler sur les qualifications des contraintes de premier et deuxième ordre et de comprendre les subtilités entre celles-ci. Pour ce faire, nous avons révisé leurs démonstrations, exploré certains chemins alternatifs de démonstration et construit des exemples et contre-exemples reliant les différentes qualifications de premier ordre. Un travail similaire a été réalisé pour les qualifications de second ordre.

Un second objectif de la maîtrise était de parvenir à élaborer des propriétés d'unicité et de différentiabilité de trajectoires centrales pour les problèmes non linéaires respectant MFCQ, mais violant (LICQ). Bien que plusieurs résultats ont été développés dans ce sens, dont l'existence des trajectoires duales et primales sous MFCQ et la convexité du problème, leur différentiabilité reste une question ouverte. Une hypothèse de convexité a dû être ajoutée pour obtenir ces résultats, ce qui en restreint l'applicabilité. Nous croyons toujours que, sous (MFCQ) et les conditions suffisantes de second ordre, la trajectoire centrale sera, localement près du minimum local, unique et différentiable.

Au cours de cette maîtrise, plusieurs problématiques ont été mises en évidence mais n'ont pas été abordées. Ce qui suit est une courte liste de celles qui ont le plus retenu notre attention. D'abord, terminer l'analyse de différentiabilité des trajectoires primales et duales dans le cas convexe lorsque (MFCQ) est respectée serait une suite immédiate aux travaux présentés. Démontrer l'existence et l'unicité locale des trajectoire dans le cas non convexe serait aussi intéressant puisque la convexité est une hypothèse très forte. Puis, on pourrait étudier le cas où (CRCQ) est satisfaite, tenter d'en déduire des propriétés et, possiblement, faire le pont avec la propriété intéressante de la Remarque 2.1.21.

Ensuite, il serait intéressant de s'attarder davantage aux qualifications de second ordre et faire un pont entre les conditions de type Abadie et les conditions de type (MFCQ), telles que (MMF). Ce pont expliquerait sûrement pourquoi les conditions de type Abadie sont suffisantes à la satisfaction des conditions nécessaires de second ordre *pour tous les multiplicateurs de premier ordre*, contrairement aux conditions de type (MFCQ) qui ne promettent la satisfaction des conditions de second ordre que pour *un certain multiplicateur* parmi l'infinité (bornée) de multiplicateurs possibles.

Finalement, une extension logique de l'étude des qualifications de premier et second ordre de ce mémoire serait le développement de qualifications de second ordre particulières aux problèmes de complémentarité (MPCC). Des qualifications de premier ordre adaptées à ces problèmes ont déjà été développées. Il serait toutefois intéressant d'analyser la possibilité d'adapter des qualifications des contraintes de deuxième ordre pour les problèmes de type (MPCC).

Bibliographie

[ABHS14] R. Andreani, R. Behling, G. Haeser, and P.J.S. Silva. On second order optimality conditions in nonlinear optimization. *Optimization-Online*, 2014.

[AES10] R. Andreani, C.E. Echagüe, and M.L. Schuverdt. Constant-rank condition and second-order constraint qualification. *Journal of Optimization Theory and Applications*, 146(2) :255–266, 2010.

[AHM11] R. Andreani, H. Haeser, and J.M. Martinez. On sequential optimality conditions for smooth constrained optimization. *Optimization*, 60 :627–641, 2011.

[AHMS12] R. Andreani, G. Haeser, M.L.Schuverdt, and P.J.S. Silva. A relaxed constant positive linear dependence constraint qualification and applications. *Mathematical Programming*, 135(1-2) :255–273, 2012.

[AHRS15] R. Andreani, H. Haeser, A. Ramos, and P.J.S. Silva. On second-order sequential optimality conditions for nonlinear optimization ans applications. *Optimization-online*, 2015.

[AMRS15] R. Andreani, J.M. Martínez, A. Ramos, and P.J.S. Silva. A cone-continuity constraint qualification and algorithmic consequences. *Optimization-online*, 2015.

[AMS04] R. Andreani, J. M. Martínez, and M. L. Schuverdt. The Constant Positive Linear Dependence Condition of Qi and Wei implies the Quasinormality

Constraint Qualification. *Journal of Optimization Theory and Applications*, (125) :473–485, 2004.

[AMS07] R. Andreani, J.M. Martinez, and M.L. Schuverdt. On second order optimality conditions for nonlinear programming. *Optimization*, 56(5-6) :529–542, 2007.

[Ani99] M. Anitescu. Degenerate nonlinear programming with a quadratic growth condition. *SIAM J. on Optimization*, 10(4) :1116–1135, 1999.

[Bac04] A. Baccari. On the Classical Necessary Second-Order Optimality Conditions. *Journal of Optimization Theory and Applications*, 123(1) :213–221, 2004.

[Ber95] D.P. Bertsekas. *Nonlinear Programming.* Athena Scientific, 1995.

[BO02] D.P. Bertsekas and A.E. Ozdaglar. Pseudonormality and a Lagrange multiplier theory for constrained optimization. *Journal of Optimization Theory and Applications*, (114) :187–343, 2002.

[BSS93] M.S. Bazaraa, H.D. Sherali, and C.M. Shetty. *Nonlinear Programming Theory and Algorithms.* John Wiley, New York, 1993.

[Bur12] J. Burke. Constraint Qualifications in Nonlinear Programming. Notes de cours Math 516, Université de Washington, Printemps 2012.

[Chi12] J. W. Chinneck. Practical optimization : A gentle introduction, chapter 16. Notes de cours, Université de Carleton, 2012.

[Cot12] R. W. Cottle. William Karush and the KKT theorem. *Doc. Math., J. DMV*, Extra Vol. :255–269, 2012.

[CW04] A.C. Chiang and K. Wainwright. *Fundamental Methods of Mathematical Economics.* McGraw-Hill/Irwin, 2004.

[DM89] J.-P. Dussault and P. Marcotte. Conditions de régularité géométrique pour les inéquations variationnelles. *RAIRO - Operations Research - Recherche Opérationnelle*, 23(1) :1–16, 1989.

[Dus14] J.-P. Dussault. Communication privée, Été 2014.

[Fle87] R. Fletcher. *Practical Methods of Optimization.* Wiley-Interscience, Chichester, 1987.

[FM90] A. V. Fiacco and G. P. McCormick. *Nonlinear Programming : Sequential Unconstrained Minimization Techniques.* Classics in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1990.

[Gau77] J. Gauvin. A necessary and sufficient regularity condition to have bounded multipliers in nonconvex programming. *Mathematical Programming*, 12(1) :136–138, 1977.

[GGK05] J.C. Gilbert, C.C. Gonzaga, and E.W. Karas. Examples of ill-behaved central paths in convex optimization. *Math. Program.*, 103(1) :63–94, 2005.

[Gui69] M. Guignard. Generalized Kuhn–Tucker Conditions for mathematical programming problems in a Banach space. *SIAM Journal on Control*, 7(2) :232–241, 1969.

[Hua71] P. Huard. Tour d'horizon : programmation non linéaire. *Revue française d'informatique and de recherche opérationnelle, série rouge*, 5(1) :3–48, 1971.

[Jan84] R. Janin. Directional Derivative of the Marginal Function in Nonlinear Programming. *Math. Program. Study*, 21 :110–126, 1984.

[Kar39] W. Karush. Minima of Functions of Several Variables with Inequalities as Side Conditions. Mémoire de maîtrise, Univesité de Chicago, 1939.

[Kra99] E. Kranich. List of references : Interior-point methods, 1999.

[KT51] H. W. Kuhn and A. W. Tucker. Nonlinear Programming. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 481–492. University of California Press, 1951.

[Lu11] S. Lu. Implications of the Constant Rank Constraint Qualification. *Math. Program.*, 126(2) :365–392, 2011.

[MS11] L. Minchenko and S. Stakhovski. On relaxed constant rank regularity condition in mathematical programming. *Optimization*, 60(4) :429–440, 2011.

[MSS11] M. C. Maciel, S. A. Santos, and G. N. Sottosanto. On second-order optimality conditions for vector optimization. *Journal of Optimization Theory and Applications*, 149(2) :332–351, 2011.

[MZ98] R.D.C. Monteiro and F. Zou. On the existence and convergence of the central path for convex programming and some duality results. *Computational Optimization and Applications*, 10(1) :51–77, 1998.

[Nem13] A. Nemirovski. Lectures on modern convex optimization. Automne 2013, Georgia Institute of Technology, 2013.

[NW06] J. Nocedal and S. J. Wright. *Numerical Optimization, 2nd edition.* Springer, New York, 2006.

[Pet73] D. Peterson. A review of constraint qualifications in finite-dimensional spaces. *SIAM Review*, 15(3) :639–654, 1973.

[Roc] R.T. Rockafellar. Lagrange Multipliers and Optimality. *SIAM Review*, 35(2) :183–238.

[Roc70] R.T. Rockafellar. *Convex Analysis.* Princeton University Press, 1970.

[Spi83] J.E. Spingarn. Partial inverse of a monotone operator. *Applied Mathematics and Optimization*, 10(1) :247–265, 1983.

[SW98] J. Stoer and M. Wechs. Infeasible-interior-point paths for sufficient linear complementarity problems and their analyticity. *Math. Program.*, 83 :407–423, 1998.

[Ter96] T. Terlaky. *Interior Point Methods of Mathematical Programming.* Applied Optimization. Springer, 1996.

[Wac13] G. Wachsmuth. On LICQ and the Uniqueness of Lagrange Multipliers. *Operations Research Letters*, 41(1) :78–80, 2013.