



HAL
open science

Contribution à l'étude de la robustesse et à la dualité en optimisation

Moussa Barro

► **To cite this version:**

Moussa Barro. Contribution à l'étude de la robustesse et à la dualité en optimisation. Analyse classique [math.CA]. Université d'Avignon, 2016. Français. NNT : 2016AVIG0416 . tel-01668600

HAL Id: tel-01668600

<https://theses.hal.science/tel-01668600>

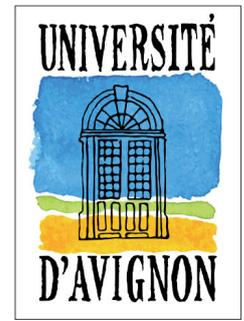
Submitted on 20 Dec 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



N° d'ordre : 03



UNIVERSITÉ D'AVIGNON
ET DES PAYS DE VAUCLUSE

Thèse

en vue de l'obtention du grade de DOCTEUR de l'Université Polytechnique de Bobo-Dioulasso (UPB)
et de l'Université d'Avignon et des Pays de Vaucluse (UAPV)

Domaine de recherche : **Mathématiques Appliquées**

Spécialité : **Optimisation**

Présentée par

Moussa BARRO

Contribution à l'étude de la robustesse et à la dualité en optimisation.

Directeurs de thèse :

Sado TRAORÉ Maître de conférences à l'Université Polytechnique de Bobo-Dioulasso (Burkina Faso)
Michel VOLLE Professeur émérite à l'Université d'Avignon et des Pays de Vaucluse (France)

Soutenue le *19 Novembre 2016*
Devant la Commission d'Examen

JURY :

Ouaténi DIALLO	Professeur des Universités, Université des Sciences, des Techniques et des Technologies de Bamako	Examineur
Stanislas OUARO	Professeur des Universités, Université Ouaga I Pr Joseph Ki-Zerbo	Président
Lionel THIBAUT	Professeur des Universités, Université Montpellier II	Examineur
Hamidou TOURÉ	Professeur des Universités, Université Ouaga I Pr Joseph Ki-Zerbo	Examineur
Michel VOLLE	Professeur émérite à l'Université d'Avignon et des Pays de Vaucluse	Directeur
Marc CILIGOT-TRAVAIN	Maître de conférences à l'Université d'Avignon et des Pays de Vaucluse	Examineur
Aboudramane GUIRO	Maître de conférences à l'Université Polytechnique de Bobo-Dioulasso	Examineur
Sado TRAORÉ	Maître de conférences à l'Université Polytechnique de Bobo-Dioulasso	Directeur

RAPPORTEURS :

Dinh THE LUC	Professeur des Universités, Université d'Avignon et des Pays de Vaucluse	France
Lionel THIBAUT	Professeur des Universités, Université Montpellier II	France
Hamidou TOURÉ	Professeur des Universités, Université Ouaga I Pr Joseph Ki-Zerbo	Burkina Faso

Thèse préparée en alternance entre l'UPB (École doctorale Sciences et Techniques /Laboratoire de mathématiques et informatique) et l'UAPV (École doctorale 536 « Sciences et agrosociétés » (Avignon)/Laboratoire de Mathématiques d'Avignon).

Table des matières

Résumé	iii
Abstract	iv
Glossaire	v
Remerciements	vi
Introduction générale	ix
0.1. Dualité et régularisation en convexité généralisée : des niveaux aux épigraphes	ix
0.2. Dualité en optimisation convexe conique à données incertaines	x
0.3. Maximisation du rayon de stabilité en optimisation continue	xi
1. Préliminaires	1
1.1. Notations ensemblistes	1
1.2. Schéma de dualité perturbative	2
1.3. Changement de variables de Charnes–Cooper	5
1.4. Autour du S-lemma	7
1.5. Formule d’Ascoli	10
I. Contribution à la dualité	12
2. Dualité et régularisation en convexité généralisée : des niveaux aux épigraphes	13
2.1. Introduction	13
2.2. Γ -régularisation de fonctions et enveloppe d’ensembles	14
2.3. Régularisations et conjugaisons par niveaux	24
2.4. Régularisations et conjugaisons épigraphiques	29
2.5. De la dualité par niveaux à la dualité épigraphique	37
2.6. Application à un problème d’optimisation : dualité par niveaux	44
3. Dualité en optimisation convexe conique à données incertaines	47
3.1. Introduction	47
3.2. Rappels d’analyse convexe	48
3.3. Pire valeur versus valeur robuste	52

3.4. Lien avec la dualité robuste forte	56
II. Contribution à l'étude de la robustesse	61
4. Rayon de stabilité en optimisation continue à données incertaines	62
4.1. Un modèle de prise de décision en environnement incertain.....	62
4.2. Premières propriétés du rayon de stabilité.....	68
4.3. Optimisation robuste et Optimisation de la robustesse	73
5. Un modèle générique de maximisation du rayon de stabilité	85
5.1. Définition du modèle	85
5.2. Étude du modèle générique dans un cadre polyédral.....	89
5.3. Étude du modèle générique dans le cas quadratique	91
6. Régression linéaire à données incertaines : méthode du rayon de stabilité ..	110
6.1. Introduction.....	110
6.2. Contrepartie robuste au sens du rayon de stabilité	111
6.3. Calcul du rayon de stabilité.....	112
6.4. Étude de la contrepartie robuste : résolution de $(\widehat{\mathcal{R}}_r^\alpha)$	119
6.5. Lien avec l'approche de l'optimisation robuste	123
6.6. Application à un problème de conception d'un réseau d'antennes	126
Conclusion et perspectives	132
Bibliographie	134

Résumé

La dualité et la robustesse sont des outils essentiels dans les processus d'aide à la décision. Cette thèse porte sur trois thèmes : dualité et régularisation en convexité généralisée, dualité en optimisation convexe conique à données incertaines, et la maximisation du rayon de robustesse en optimisation continue.

Concernant la dualité et la régularisation de fonctions numériques à valeurs réelles étendues, deux points de vue sont considérés : l'approche par les niveaux et l'approche par les épigraphes. On étend ainsi à la convexité généralisée des résultats récents concernant le passage de la dualité quasi-convexe à la dualité convexe. On applique cette théorie à un problème d'optimisation pour déduire un résultat de dualité forte.

Concernant la dualité en optimisation convexe conique à données incertaines, on considère les notions de pire valeur et de valeur robuste d'un problème d'optimisation conique à données incertaines. On donne une condition nécessaire et suffisante pour l'égalité entre la pire valeur et la valeur robuste de ce problème avec exactitude de la pire valeur. On déduit une condition suffisante permettant d'obtenir la propriété de dualité robuste forte pour ce problème.

La dernière partie de ce travail porte sur une étude du problème de maximisation du rayon de stabilité. On définit le rayon de robustesse pour un problème de décision en milieu incertain, et on établit certaines de ces propriétés analytiques (concavité et semi-continuité). La contrepartie robuste d'un problème d'optimisation à données incertaines au sens du rayon de robustesse est introduite. On montre le lien en termes d'ensemble de solutions optimales entre la contrepartie robuste au sens du rayon de robustesse et celle au sens de l'optimisation robuste d'un problème incertain d'optimisation continue. Un modèle générique du problème de maximisation du rayon de robustesse regroupant une large classe de cas pratiques est proposé. On examine ce modèle dans un cas polyédral, dans le cas de la régression linéaire puis dans un cas quadratique. Notre stratégie dans ces différents cas, consiste à expliciter le rayon de robustesse et/ou à transformer le problème de maximisation du rayon de stabilité en un programme tractable. Une application à un problème de conception d'antenne circulaire est proposée dans le cas de la régression et une application au calcul d'un estimateur robuste est proposée dans le cas quadratique.

Abstract

Duality and robustness are two important tools in decision making process. This thesis deals with three topics : duality and regularity in generalized convexity, duality for an uncertain convex conical optimization problem, maximization of the stability radius.

Concerning the duality and regularity of the extended real-valued functions, two points of view are considered : the sub-level set approach and the epigraphical approach. We then extend some recent results concerning the passage from the quasi-convex duality to convex duality to the generalized convexity. We apply this theory to an optimization problem to derive a strong duality property for this problem.

Concerning the duality for an uncertain convex conical optimization problem, we consider the notions of worst value and robust value of an uncertain convex conical optimization problem. We give a necessary and sufficient condition to obtain the equality between the robust value and the worst value with exactness for the worst value. We derive a sufficient condition to obtain a robust strong duality property for this problem.

The last part of this work is devoted to the study of the problem of maximization of the stability radius. We define the stability radius for a decision problem under data uncertainty, and prove some of its analytical properties (e.g concavity and upper semi-continuity). The robust counterpart of an uncertain optimization problem according to the stability radius is introduced. We study the relation between the solution set of this counterpart and the solution set of the robust counterpart according to the robust optimization approach. We propose a generic model of the maximization of stability radius which covers a large class of applications. We study this problem in a polyhedral case, in the case of regression and in quadratic case. In each case, we compute the stability radius and/ or transform the problem of maximization of the stability radius to a tractable problem. An application to a circular antenna design problem is given in the regression case, and an application to compute a robust estimator is provided in the quadratic case.

Glossaire

$E_{v,r}^P$	ensemble (c, P) -élémentaire
$\langle \rangle_{c,P}$	opérateur de fermeture associé à la famille d'ensembles (c, P) -élémentaires
h^{Γ_c}	Γ_c -régularisée de la fonction h
$h^{\langle \rangle_{c,P}}$	(c, P) -régularisée par niveaux de la fonction h
$h^{\langle \rangle_{\tilde{c},P}}$	(\tilde{c}, P) -régularisée épigraphique de la fonction h
$\Gamma_c(U)$	ensemble des fonctions c -régulières de U à valeurs dans $\overline{\mathbb{R}}$
$\Gamma(U)$	ensemble des fonctions convexes propres s.c.i de U à valeurs dans $\overline{\mathbb{R}}$
$\mathcal{R}_{c,P}(U)$	ensemble des sous-ensembles (c, P) -réguliers de U
$\mathcal{N}_{c,P}(U)$	ensemble des fonctions (c, P) -régulières par niveaux définies sur U dans $\overline{\mathbb{R}}$
$\mathcal{E}_{\tilde{c},P}(U)$	ensemble des fonctions (\tilde{c}, P) -épi-régulières définies sur U à valeurs dans $\overline{\mathbb{R}}$
$\mathcal{L}(E, H)$	ensemble des fonctions linéaires continues de E dans H
$\mathcal{C}(U, \mathbb{R})$	ensemble des fonctions continues de U dans \mathbb{R}
$\Delta_{c,P}$	polarité associée à l'opérateur de fermeture $\langle \rangle_{c,P}$
$\Delta_{c,P}^*$	polarité duale de la polarité $\Delta_{c,P}$
h^c	c -conjuguée de la fonction h par rapport au couplage c
h^{cc}	c -biconjuguée de la fonction h par rapport au couplage c
$\text{dom } h$	domaine effective de la fonction h
$\text{epi } h$	épigraphe de la fonction h
$\text{gr } h$	graphe de la fonction h
$\text{epi}_S g$	S -épigraphe de la fonction vectorielle g
$[g \leq_S y]$	S -tranche inférieure de niveau y de la fonction vectorielle g
$[h \leq t]$	tranche inférieure de niveau t de la fonction h
$[h \geq t]$	tranche supérieure de niveau t de la fonction h
$[h = t]$	tranche de niveau t de la fonction h
$[h < t]$	tranche inférieure stricte de niveau t de la fonction h
$[h > t]$	tranche supérieure stricte de niveau t de la fonction h
i_A	fonction indicatrice de l'ensemble A
σ_A	fonction support de l'ensemble A
ϑ_A	fonction vallée de l'ensemble A
A°	polaire de l'ensemble A
$\text{conv } A$	enveloppe convexe de l'ensemble A
$\text{Ext } A$	ensemble des points extrémaux de A
$\mathbb{B}_{\ \cdot\ }$	boule unité fermée
$\mathbb{S}_{\ \cdot\ }$	sphère unité fermée
$\Pi_{X \times Y \rightarrow Y}$	projection canonique de $X \times Y$ sur Y
I_X	identité de X
2^X	ensemble des sous-ensembles de X
$v(\mathcal{P})$	valeur du problème d'optimisation (\mathcal{P})
$\text{Argmin}(\mathcal{P})$	ensemble des solutions optimales du problème de minimisation (\mathcal{P})
$\text{Argmax}(\mathcal{P})$	ensemble des solutions optimales du problème de maximisation (\mathcal{P})
s.l.c	sous les contraintes

Remerciements

Maman chérie Sita Ouattara, je te dis merci pour toutes les valeurs que tu m'as inculquées. Merci pour ta présence dans les moments difficiles, pour tes conseils et encouragements qui m'ont permis de persévérer dans le travail. Je remercie ma famille pour leur amour et leur soutien.

Vous mes directeurs de thèse : Sado Traoré, Maître de conférences à l'Université Polytechnique de Bobo-Dioulasso et Michel Volle, Professeur émérite à l'Université d'Avignon et des pays de Vaucluse, je dis merci pour avoir cru en moi en me proposant ce projet de thèse. Je vous suis également reconnaissant pour le temps inestimable que vous m'avez consacré, vos qualités pédagogiques et scientifiques, votre franchise et votre sympathie. J'ai énormément appris avec vous et je vous adresse ma gratitude pour tout cela.

J'adresse de chaleureux remerciements à Marc Ciligot-Travain, Maître de conférences à l'Université d'Avignon et des pays de Vaucluse, pour son attention de tout instant sur mes travaux, pour ses conseils avisés et son écoute qui ont été prépondérants pour la bonne réussite de cette thèse. J'ai vraiment pris du plaisir à travailler avec vous et j'ai beaucoup appris. Je dis également merci à votre épouse et à vos adorables filles pour l'hospitalité très chaleureuse et les repas de famille que vous m'avez offert avec enthousiasme.

Je voudrais dire merci aux rapporteurs de cette thèse : Dinh THE LUC Professeur des Universités à l'Université d'Avignon et des pays de Vaucluse, Lionel THIBAUT, Professeur des Universités à l'Université Montpellier II et Hamidou TOURÉ, Professeur des Universités à l'Université Ouaga I Pr Joseph Ki-Zerbo pour l'intérêt qu'ils ont accordé à mon travail malgré leurs différentes occupations. Vos différents commentaires et remarques m'ont permis d'améliorer le document. Je vous suis reconnaissant pour tout cela. Au Professeur Dinh THE LUC, je dis encore merci pour le temps que vous m'avez accordé pour discuter de mon travail.

Je tiens à remercier les honorables membres de Jury : Ouaténi DIALLO, Stanislas OUARO, Lionel THIBAUT, Hamidou TOURÉ, Michel VOLLE, Aboudramane GUIRO et Sado TRAORÉ pour l'honneur que vous m'avez faite de votre présence dans mon jury de soutenance, pour vos différentes questions et suggestions qui ont permis d'améliorer le document. Je vous dis merci pour l'intérêt que vous avez accordé à mon travail.

Ce travail n'aurait pu être mené à bien sans l'aide de différents partenaires qui, au travers de leur soutien financiers et matériels, ont cru à mon projet de thèse et m'ont fait confiance : le gouvernement du Burkina Faso à travers la bourse nationale

et l'ambassade de France au Burkina à travers la bourse SCAC, je vous dit merci. J'associe à ces remerciements CAMPUS France pour la gestion de la bourse SCAC, je vous dis merci pour votre professionnalisme qui a rendu agréable mes séjours à Avignon.

Au Maître de conférences Joseph Bayara et au Docteur Jean Louis Zerbo, je dis grandement merci pour votre confiance, vos conseils, enseignements et soutiens. Merci d'être à la fois mes enseignants, amis et collègues !

Aux corps professoral des Universités du Burkina, en particulier les mathématiciens : Théodore Marie Yves Tapsoba, Moussa Ouattara, Hamidou Touré, Gérard Kientéga, Stanislas Ouaro, Oumar Traoré, Kaka Bernard Bonzi, Sado Traoré, Joseph Bayara, Idrissa Kaboré, Jean de Dieu Zabsonré, Aboudramane Guiro, Adama Ouédraogo, Boureima Sangaré, Ismaël Nyankini, je dis merci pour les enseignements et les conseils qui m'ont été d'un grand apport.

Je remercie les membres du Laboratoire de mathématiques d'Avignon pour l'accueil très chaleureux. En particulier je dis merci au Professeur Alberto Seeger, pour sa disponibilité à répondre avec clarté à mes questions. Aux Maîtres de conférences Mohammed Moussaoui, Samuel Amstutz, Daniel Gourion, Andréa Venturelli, je vous dis merci pour les discussions autour d'un café pour les uns, dans les bureaux pour les autres. J'ai beaucoup appris à travers ces échanges. J'associe à ces remerciements mes camarades thésards du Laboratoire de Mathématiques d'Avignon.

Je voudrais remercier Madame Evéline Traoré/Hien, l'épouse du Professeur Traoré pour m'avoir accepté comme membre à part entière de la famille Traoré. Je remercie par la même occasion vos deux adorables filles Anna et Anaïs qui ont toujours été très sympa avec moi. Tonton Barro vous dit merci les filles !

Je tiens à dire merci à ma bien aimée Mademoiselle Louise Tarnagda, pour son amour, sa patience, sa présence à mes cotés qui constituent un équilibre important dans ma vie. Je te dis grandement merci.

Je n'oublie pas mes amis qui m'ont encouragé et soutenu sur le plan professionnel et social, je voudrais remercier ici Nicolas Gil, Simon Raynal, Ako Gunn, Zakaria Yé, Harou Guindja, Eli Outoumbé, Ali Ouédraogo, Idrissa Ibrango, Cédric Somé, Souleymane Sitonoma Ouédraogo, Abdoul-Aziz Compaoré, Satafa Sanogo, Boukaré Kientéga, Pascal Somda, Wendkouni Ouédraogo, Bakary Traoré.

Enfin, je dis merci à tout ceux qui de près ou de loin ont contribué à l'aboutissement de ce travail et qui n'ont pas retrouvé leur nom ci-dessus. Je vous pris de bien vouloir m'excuser pour cet oubli.

Je dédie ce travail :

- à ma chère maman Lita Cuattara qui m'accompagne tous les jours,*
- à la mémoire de mon défunt père Liaka Barro dont j'aurais aimé qu'il soit présent en ce jour si spécial à mes côtés.*

Introduction générale

La dualité et la robustesse sont des outils dont l'importance n'est plus à démontrer en optimisation, en économie, en mécanique, en théorie des jeux, en conception, en contrôle, etc. Ce travail apporte une contribution en trois points sur ces deux notions : dualité et régularisation en convexité généralisée, dualité en optimisation convexe conique à données incertaines, maximisation du rayon de robustesse en optimisation continue.

0.1. Dualité et régularisation en convexité généralisée : des niveaux aux épigraphes

Dans l'article [73], M. Volle et ses collaborateurs ont explicité le passage de la dualité quasi-convexe à la dualité convexe dans le cas des fonctions evenly quasi-convexes et evenly convexes. On propose ici une approche générale de la méthodologie développée par ces auteurs.

Étant donné une fonction $c : U \times V \rightarrow \mathbb{R}$ dite de couplage où U et V sont deux ensembles non vides, une partie non vide P de \mathbb{R} , principalement $P_1 = \mathbb{R}_+$, $P_2 = \mathbb{R}_+^*$, $P_3 = \mathbb{R}^*$, $P_4 = \{0\}$, la notion clé est celle d'ensemble (c, P) -élémentaire :

$$(0.1) \quad E_{v,r}^P = \{u \in U \mid r - c(u, v) \in P\},$$

et la notion d'ensemble (c, P) -régulier qui en découle, à savoir les intersections d'ensembles (c, P) -élémentaires. La (c, P) -enveloppe d'une partie A de U est l'intersection des ensembles (c, P) -élémentaires (ou réguliers) contenant A , notée $\langle A \rangle_{c,P}$, généralisant notamment l'enveloppe evenly convexe et l'enveloppe coaffine dans un espace vectoriel topologique, l'enveloppe convexe fermée et l'enveloppe affine fermée dans un espace vectoriel topologique localement convexe, la fermeture topologique (l'adhérence) dans un espace métrique, ... On décrit et compare $\langle A \rangle_{c,P_i}$, $i = 1, 2, 3, 4$ (Proposition 2.2.6, Corollaire 2.2.1, Proposition 2.2.8). Le lien entre la (c, P) -enveloppe des ensembles et la Γ -régularisée des fonctions par rapport à un couplage c (ou associé à c) est précisé par l'intermédiaire des fonctions indicatrices (Proposition 2.2.9, Théorème 2.2.2, Corollaire 2.2.2).

On s'intéresse à la classe des fonctions dont les tranches inférieures larges sont fermées pour l'opérateur $\langle \cdot \rangle_{c,P}$. On définit alors la (c, P) -régularisée $h^{\langle \cdot \rangle_{c,P}}$ par niveaux

d'une fonction $h : U \rightarrow \overline{\mathbb{R}}$. On donne plusieurs expressions de $h^{\langle \rangle_{c,P}}$ (Théorème 2.3.1, Corollaire 2.3.1, Proposition 2.3.6) que l'on caractérise aussi comme la biconjuguée de h par rapport à un couple de polarités duales (Δ, Δ^*) associé à l'opérateur $\langle \rangle_{c,P}$. S'agissant de définir la régularisée épigraphique d'une fonction $h : U \rightarrow \overline{\mathbb{R}}$, on travaille naturellement avec le couplage $\check{c} : (U \times \mathbb{R}) \times (V \times \mathbb{R}) \rightarrow \mathbb{R}$,

$$(0.2) \quad \check{c}((u, r), (v, s)) = c(u, v) + rs,$$

et les ensembles élémentaires (resp. réguliers) correspondants. On définit ainsi les fonctions (\check{c}, P) -épi-régulières comme celles dont l'épigraphe est (\check{c}, P) -régulier, et la (\check{c}, P) -épi-régularisée de h , notée $h^{\langle \rangle_{\check{c}, P}}$. On obtient des formules explicites de $h^{\langle \rangle_{\check{c}, P_i}}$ pour $i = 1, 2, 3, 4$ (Théorème 2.4.1, Théorème 2.4.2, Théorème 2.4.3). On montre (Théorème 2.4.5) que la (\check{c}, P_i) -épi-régularisée de h coïncide avec la biconjuguée de h relative à un couplage γ_i approprié :

$$(0.3) \quad h^{\langle \rangle_{\check{c}, P_i}} = h^{\gamma_i \gamma_i}, \quad i = 1, 2, 3,$$

généralisant ainsi des résultats obtenus dans [73], Théorèmes 3.1 et 6.1, et dans [70] Théorèmes 4.2 et 5.2.

Concernant le passage de la régularisation par niveaux à la régularisation épigraphique, on associe à chaque fonction $h : U \rightarrow \overline{\mathbb{R}}$, la fonction $H : U \times \mathbb{R} \rightarrow \overline{\mathbb{R}}$ définie par $H(u, t) = h(u) - t$. On établit (Théorème 2.5.1) que h est (\check{c}, P) -épi-régulière si et seulement si H est (\check{c}, P) -régulière par niveaux, si et seulement si H est $(\check{\check{c}}, P)$ -épi-régulière, où

$\check{\check{c}} : (U \times \mathbb{R} \times \mathbb{R}) \times (V \times \mathbb{R} \times \mathbb{R}) \rightarrow \mathbb{R}$ est le couplage défini par :

$$(0.4) \quad \check{\check{c}}((u, t, t'), (v, s, s')) := \check{c}((u, t), (v, s)) + t's' = c(u, v) + ts + t's'.$$

Des formules sont établies entre la biconjuguée par niveaux de H et la $(\check{\check{c}}, P_i)$ -épi-régularisée de h pour $i = 1, 2, 3, 4$ (Théorème 2.5.2).

On exploite les résultats ci-dessus pour obtenir un théorème de dualité forte (Théorème 2.6.1) dans le cas des problèmes d'optimisation paramétrés (approche perturbative par niveaux). On termine par une application au cas spécifique des problèmes d'optimisation evenly quasi-convexe (Corollaire 2.6.1 et Corollaire 2.6.2).

0.2. Dualité en optimisation convexe conique à données incertaines

On considère le problème d'optimisation convexe conique incertain suivant

$$(P) \quad \inf_x f(x) \quad \text{s.l.c.} \quad g_u(x) \in -S,$$

où u désigne le paramètre incertain appartenant à U l'ensemble des incertitudes, X et Y sont deux espaces vectoriels topologiques localement convexes et séparés, $f : X \rightarrow \overline{\mathbb{R}}$ une fonction convexe propre semi-continue inférieurement, $S \subset Y$, un

cône convexe fermé non vide, et pour tout $u \in U$, la fonction $g_u : \text{dom } g_u \subset X \rightarrow Y$ est S -convexe fermée par niveaux ou par épigraphe.

On associe au problème (P) d'une part sa contrepartie robuste ([7],[8],[10]) définie par :

$$(RP) \quad \inf_x f(x) \quad \text{s.l.c.} \quad g_u(x) \in -S, \quad \forall u \in U,$$

et d'autre part, son dual optimiste ([5], [14],[47],[49]) défini par

$$(ODP) \quad \sup_{(u,\lambda)} \inf_{x \in X} \{f(x) + \lambda g_u(x)\} \quad \text{s.l.c.} \quad (u, \lambda) \in U \times S^+.$$

On introduit le problème d'optimisation

$$(Q) \quad \sup_u \inf_x \{f(x) \quad \text{s.l.c.} \quad g_u(x) \in -S\} \quad \text{s.l.c.} \quad u \in U.$$

On montre que

$$(0.5) \quad \sup(ODP) \leq \sup(Q) \leq \inf(RP).$$

Lorsque le sup (respectivement l'inf) est atteint, on dit qu'il est exact et on note max (respectivement min).

À l'aide d'exemples, on montre que les inégalités dans (0.5) peuvent être strictes.

On donne une condition nécessaire et suffisante permettant d'obtenir

$$(0.6) \quad \max(Q) = \inf(RP),$$

puis on déduit de ce résultat, une condition suffisante pour obtenir

$$(0.7) \quad \max(ODP) = \inf(RP).$$

0.3. Maximisation du rayon de stabilité en optimisation continue

Dans la plupart des modèles intervenant dans les domaines tels que la recherche opérationnelle, l'optimisation, la conception, la fiabilité, etc, les données peuvent être incertaines (incertitude due à des erreurs de mesure ou de modélisation par exemple). Dans un travail ([21]), M. Ciligot-Travain présente une notion de robustesse pouvant servir à l'étude de ces modèles incertains. Dans cette partie, on s'intéresse à cette notion de robustesse dans le cas de l'optimisation continue. Cette partie est organisée comme suit.

Dans le chapitre 4, section 4.1, on décrit de façon un peu abstraite un modèle de prise de décision en environnement incertain. On montre qu'en enrichissant un peu ce modèle, on se ramène à un problème de maximisation d'un indicateur de robustesse appelé rayon de stabilité ou de robustesse. En utilisant cette notion d'indicateur de robustesse, on définit la contrepartie robuste au sens du rayon de stabilité d'un problème d'optimisation à données incertaines.

Dans la section 4.2, on considère le rayon de robustesse défini à la section 4.1 comme fonction d'une variable de décision et d'une variable incertaine. On met un peu de structures sur l'ensemble de décision et / ou d'incertitude, afin d'étudier des propriétés de concavité et de semi-continuité supérieure de l'indicateur de robustesse. Dans la section 4.3, on considère un problème d'optimisation continue à données incertaines, on lui associe sa contrepartie robuste définie à la section 4.1 (Optimisation de la robustesse). On rappelle qu'à un problème d'optimisation à donnée incertaines est associé une contrepartie robuste basée sur la méthode du pire des cas (Optimisation robuste). On se pose la question de savoir s'il y a égalité entre l'ensemble des solutions robustes au sens du rayon de robustesse et l'ensemble des solutions robustes au sens du pire des cas. On donne une réponse à cette question en utilisant la notion d'inverse généralisée d'une fonction croissante.

Dans le chapitre 5, section 5.1, on propose un modèle générique du problème de maximisation du rayon de robustesse qui couvre un grand nombre de cas pratiques dont la contrepartie robuste au sens du rayon de robustesse d'un problème d'optimisation à données incertaines. On établit la quasi-concavité et la semi-continuité supérieure du rayon de robustesse pour ce modèle générique.

La section 5.2 porte sur l'étude du modèle générique dans un cadre polyédral. On détermine dans ce cas une formule explicite du rayon de stabilité qui montre que le modèle générique est un programme fractionnaire généralisé. On utilise ces résultats, pour expliciter la contrepartie robuste au sens du rayon de stabilité d'un programme linéaire à données incertaines.

La section 5.3 est consacré à l'étude du modèle générique dans un cadre quadratique. Dans ce cas, nous n'avons pas de formule explicite du rayon de stabilité. Cependant, en utilisant d'une part le S-lemma, et d'autre part des techniques de minimisation de fonctionnelle quadratique, on transforme le modèle qui était un problème d'optimisation semi-infinie (une infinité de contraintes) en problème d'optimisation "tractable". Par un changement de variables de type Charnes-Cooper, on transforme ce programme "tractable" en un problème de programmation semi-définie positive. On montre qu'une seconde réduction de la taille des matrices apparaissant dans des contraintes de ce dernier programme est possible et peut avoir une importance d'un point de vue numérique. On termine cette section par une illustration de nos résultats sur un problème de calcul d'estimateurs robustes.

Dans le chapitre 6, on applique la méthode du rayon de stabilité pour étudier un problème de régression à données incertaines. On montre que la contrepartie robuste au sens du rayon de stabilité de ce problème est un programme fractionnaire en déterminant une formule explicite du rayon de stabilité. On transforme ce programme fractionnaire en un problème d'optimisation convexe en utilisant les techniques de

changement de variables de Charnes-Cooper. On note également que dans certains cas, ce problème convexe se ramène à un programme linéaire par l'ajout de nouvelles variables. On illustre nos résultats sur un problème de conception d'antenne circulaire.

Préliminaires

Dans ce chapitre nous faisons un rappel de quelques notions utiles pour la suite de ce mémoire.

1.1. Notations ensemblistes

Si X et Y sont deux ensembles, on note $\Pi_{X \times Y \rightarrow X}$ et $\Pi_{X \times Y \rightarrow Y}$ les projections canoniques :

$$\Pi_{X \times Y \rightarrow X} = \begin{cases} X \times Y \rightarrow X, \\ (x, y) \rightarrow x, \end{cases} \quad \Pi_{X \times Y \rightarrow Y} = \begin{cases} X \times Y \rightarrow Y, \\ (x, y) \rightarrow y. \end{cases}$$

On note I_X l'identité de X , 2^X l'ensemble des parties de X .

Si A est une partie d'un ensemble X indiqué par le contexte, $A^c = \{x \in X \mid x \notin A\}$ est le complémentaire de A dans X .

On note

$$\begin{aligned} \overline{\mathbb{R}} &= [-\infty; +\infty]; & \mathbb{R}^* &= \mathbb{R} \setminus \{0\}; & \mathbb{R}_+ &= [0; +\infty[; & \mathbb{R}_- &=]-\infty; 0]; \\ \overline{\mathbb{R}}_+ &= [0; +\infty]; & \overline{\mathbb{R}}_- &= [-\infty; 0]; & \mathbb{R}_+^* &=]0; +\infty[; & \mathbb{R}_-^* &=]-\infty; 0[. \end{aligned}$$

Étant donné une fonction $h : U \rightarrow \overline{\mathbb{R}}$, où U est un ensemble, on utilisera les notations et définitions classiques ci-dessous :

$$\begin{aligned} \text{dom } h &:= \{u \in U \mid h(u) < +\infty\}, \text{ le domaine effectif de } h, \\ \text{epi } h &:= \{(u, t) \in U \times \mathbb{R} \mid h(u) \leq t\}, \text{ l'épigraphe de } h, \\ \text{gr } h &:= \{(u, h(u)) \mid u \in U\}, \text{ le graphe de } h \\ [h \leq t] &:= \{u \in U \mid h(u) \leq t\}, \text{ la tranche inférieure de niveau } t \text{ de } h, \\ [h \geq t] &:= \{u \in U \mid h(u) \geq t\}, \text{ la tranche supérieure de niveau } t \text{ de } h, \\ [h = t] &:= \{u \in U \mid h(u) = t\}, \text{ la tranche de niveau } t \text{ de } h, \\ [h < t] &:= \{u \in U \mid h(u) < t\}, \text{ la tranche inférieure stricte de niveau } t \text{ de } h, \\ [h > t] &:= \{u \in U \mid h(u) > t\}, \text{ la tranche supérieure stricte de niveau } t \text{ de } h. \end{aligned}$$

Soit $f : X \rightarrow \overline{\mathbb{R}}$ (resp. $f : X \rightarrow \overline{\mathbb{R}}_+$) une fonction .

Si on considère le problème de minimisation

$$(\mathcal{P}_{\min}) \quad \text{Minimiser } f(x), \quad x \in X,$$

on note $v(\mathcal{P}_{\min}) := \inf_{\overline{\mathbb{R}}} \{f(x) \mid x \in X\}$ [resp. $v(\mathcal{P}_{\min}) := \inf_{\overline{\mathbb{R}}_+} \{f(x) \mid x \in X\}$] la valeur du problème (\mathcal{P}_{\min}) et

$$\text{Argmin}(\mathcal{P}_{\min}) := \{\bar{x} \in X \mid \forall x \in X, f(\bar{x}) \leq f(x)\}$$

l'ensemble des solutions de (\mathcal{P}_{\min}) .

Si on considère le problème de maximisation

$$(\mathcal{P}_{\max}) \quad \text{Maximiser } f(x), \quad x \in X,$$

on note $v(\mathcal{P}_{\max}) := \sup_{\mathbb{R}}\{f(x) \mid x \in X\}$ [resp. $v(\mathcal{P}_{\max}) := \sup_{\mathbb{R}_+}\{f(x) \mid x \in X\}$] la valeur du problème (\mathcal{P}_{\max}) et

$$\text{Argmax}(\mathcal{P}_{\max}) := \{\bar{x} \in X \mid \forall x \in X, f(x) \leq f(\bar{x})\}$$

l'ensemble des solutions de (\mathcal{P}_{\max}) .

1.2. Schéma de dualité perturbative

En mathématique, la dualité regroupe l'ensemble des techniques permettant d'exprimer de façon équivalente un concept. Elle est par exemple utilisée en optimisation pour associer à un problème d'optimisation un autre problème d'optimisation appelé son dual dont la valeur est une minorante de celle du primal. En général, la question est de savoir quand avons nous l'égalité entre les valeurs des deux problèmes et comment passer de l'ensemble des solutions optimales de l'un à celui de l'autre. Une méthode de construction du dual est la dualité perturbative qui utilise la conjugaison de Legendre-Fenchel.

1.2.1. Conjugaison de Legendre-Fenchel

Dans cette sous-section, on note X et Y deux \mathbb{R} -espaces vectoriels mis en dualité par une forme bilinéaire $\langle \cdot, \cdot \rangle$; selon la terminologie de Bourbaki ([53]), cela signifie que les axiomes de séparation suivants sont vérifiés :

$$(1.1) \quad \begin{array}{l} \text{pour tout } x \text{ non nul dans } X, \text{ il existe } y \in Y \text{ tel que } \langle x, y \rangle \neq 0 \\ \text{pour tout } y \text{ non nul dans } Y, \text{ il existe } x \in X \text{ tel que } \langle x, y \rangle \neq 0. \end{array}$$

Une topologie localement convexe sur X (resp. sur Y) est dite compatible avec la dualité si, pour cette topologie, l'ensemble des formes linéaires continues est constitué par les fonctions $\langle \cdot, y \rangle$, $y \in Y$ (resp. les fonctions $\langle x, \cdot \rangle$, $x \in X$). Les axiomes de séparation (1.1) impliquent, d'une part qu'une telle topologie est séparée, d'autre part que chaque forme linéaire continue est représentable de la sorte d'une seule manière.

Exemple 1.2.1. L'exemple le plus usuel d'espaces vectoriels en dualité est obtenu en prenant pour X un espace vectoriel muni a priori d'une topologie localement convexe séparée et pour Y son dual topologique (ensemble des formes linéaires continue sur X).

Définition 1.2.1 (fonction affine continue).

Les fonctions affines continues par exemple sur X , s'écrivent sous la forme

$$(1.2) \quad x \mapsto \langle x, y \rangle - r, \quad \text{où } y \in Y \text{ et } r \in \mathbb{R}.$$

Soit $h : X \rightarrow \overline{\mathbb{R}}$ une fonction numérique réelle à valeurs étendues, la fonction affine continue écrite en (1.2) est une minorante de h si et seulement si

$$(1.3) \quad r \geq \sup_{x \in X} \{ \langle x, y \rangle - h(x) \}.$$

Cela conduit à s'intéresser à la fonction définie sur l'espace Y par le second membre de (1.3) :

Définition 1.2.2 (Legendre-Fenchel-Moreau).

Pour toute fonction $h : X \rightarrow \overline{\mathbb{R}}$, on appelle fonction conjuguée de h , la fonction $h^* : Y \rightarrow \overline{\mathbb{R}}$ définie pour tout $y \in Y$ par

$$(1.4) \quad h^*(y) := \sup_{x \in X} \{ \langle x, y \rangle - h(x) \}.$$

Il résulte immédiatement de (1.4), que pour toute fonction $h : X \rightarrow \overline{\mathbb{R}}$, on a

$$(1.5) \quad \forall x \in X, \forall y \in Y, \quad h^*(y) \geq \langle x, y \rangle - h(x),$$

appelée inégalité de Young-Fenchel.

Par un raisonnement analogue sur Y , on définit la conjuguée d'une fonction $k : Y \rightarrow \overline{\mathbb{R}}$, par la fonction $k^* : X \rightarrow \overline{\mathbb{R}}$ définie en tout $x \in X$ par :

$$(1.6) \quad k^*(x) := \sup_{y \in Y} \{ \langle x, y \rangle - k(y) \}.$$

Cela montre que l'on peut définir la conjuguée de la fonction h^* , notée h^{**} par

$$(1.7) \quad h^{**} : X \rightarrow \overline{\mathbb{R}}, \quad x \mapsto \sup_{y \in Y} \{ \langle x, y \rangle - h^*(y) \}.$$

La fonction h^{**} s'appelle la biconjuguée de la fonction h .

Notons que ce schéma montre clairement que pour toute fonction $h : X \rightarrow \overline{\mathbb{R}}$, on a :

- (1) $h^{***} = h^*$,
- (2) $h^{**} \leq h$ (d'après (1.7)),
- (3) si $\text{dom } h \neq \emptyset$ alors $h^*(y) > -\infty, \forall y \in Y$ (d'après (1.4)).

1.2.2. Méthode de construction du dual perturbatif [13, 62]

Considérons le problème de minimisation suivant

$$(\mathcal{P}) \quad \text{Minimiser } f(x), \quad \text{s.l.c } x \in X,$$

où X est un \mathbb{R} -espace vectoriel topologique et $f : X \rightarrow \overline{\mathbb{R}}$ est une fonction.

On se donne un autre \mathbb{R} -espace vectoriel topologique Y (appelé espace des perturbations) et une fonction $F : X \times Y \rightarrow \overline{\mathbb{R}}$ vérifiant

$$(1.8) \quad F(x, 0_Y) = f(x), \quad \forall x \in X.$$

La fonction F s'appelle *fonction de perturbation* associée au problème (\mathcal{P}) .

Il résulte immédiatement de (1.8) que le problème (\mathcal{P}) s'écrit aussi

$$(\mathcal{P}) \quad \text{Minimiser } F(x, 0_Y), \quad \text{s.l.c } x \in X.$$

On appelle *dual perturbatif* du problème (\mathcal{P}) , le programme suivant

$$(\mathcal{D}) \quad \text{Maximiser } -F^*(0_{X^*}, y^*), \quad \text{s.l.c } y^* \in Y^*,$$

où $F^* : X^* \times Y^* \rightarrow \overline{\mathbb{R}}$ est la conjuguée de Legendre-Fenchel de F .

D'après l'inégalité de Young-Fenchel (1.5), pour tout $x \in X$ et pour tout $y^* \in Y^*$, on a

$$F(x, 0_Y) + F^*(0_{X^*}, y^*) \geq \langle x, 0_{X^*} \rangle + \langle 0_Y, y^* \rangle = 0 \iff F(x, 0_Y) \geq -F^*(0_{X^*}, y^*).$$

D'où

$$(1.9) \quad -\infty \leq v(\mathcal{D}) \leq v(\mathcal{P}) \leq +\infty.$$

La propriété (1.9) s'appelle *la dualité faible*. Il existe dans la littérature plusieurs types de conditions de qualification pour obtenir l'égalité $v(\mathcal{D}) = v(\mathcal{P})$ avec le dual admettant au moins une solution optimale.

Exemple 1.2.2 (dual d'un problème conique). Supposons que le problème (\mathcal{P}) prenne la forme suivante

$$(\mathcal{P}_c) \quad \text{Minimiser } f(x), \quad \text{s.l.c } x \in C \text{ et } g(x) \in -K,$$

où $f : X \rightarrow \overline{\mathbb{R}}$ est convexe, $C \subset X$ est un convexe, K est un cône convexe fermé non vide de Y un \mathbb{R} -espace vectoriel topologique et $g : X \rightarrow Y$ est une fonction K -convexe dans le sens que

$$(1.10) \quad g(\lambda x + (1 - \lambda)x') - \lambda g(x) - (1 - \lambda)g(x') \in K, \quad \forall x, x' \in X, \forall \lambda \in]0; 1[.$$

Le problème (\mathcal{P}_c) est un problème d'optimisation convexe conique.

On associe au problème (\mathcal{P}_c) , la fonction de perturbation $F : X \times Y \rightarrow \overline{\mathbb{R}}$, définie par

$$F(x, y) = \begin{cases} f(x) & \text{si } x \in C \text{ et } y - g(x) \in K \\ +\infty & \text{sinon.} \end{cases}$$

Par un calcul élémentaire, on obtient pour $y^* \in Y^*$,

$$-F^*(0_{X^*}, y^*) = \begin{cases} \inf_{x \in C} \{f(x) + \langle g(x), y^* \rangle\} & \text{si } y^* \in K^+ \\ -\infty & \text{sinon,} \end{cases}$$

où

$$K^+ := \{y^* \in Y^* \mid \langle y, y^* \rangle \geq 0, \forall y \in K\},$$

désigne le cône polaire du cône K . Le dual perturbatif correspondant est alors défini par

$$(\mathcal{D}_c) \quad \text{Maximiser } \inf_{x \in C} \{f(x) + \langle g(x), y^* \rangle\}, \quad \text{s.l.c } y^* \in K^+.$$

1.3. Changement de variables de Charnes–Cooper

Le changement de variables de Charnes-Cooper a été introduit initialement en 1962 par A. Charnes et W.W. Cooper dans [20] pour transformer un programme de maximisation d'un quotient de fonctions affines sous des contraintes affines en un programme linéaire. Il a été généralisé au cas non affine en 1973 par S. Schaible dans [64]. Nous rappelons cette technique en reprenant les démonstrations.

1.3.1. Notations

Soient X un \mathbb{R} -espace vectoriel, $S \subset X$ et $h : S \rightarrow \mathbb{R}$ une fonction. On pose

$$\tilde{S} = \left\{ (y, t) \in X \times \mathbb{R} \mid t > 0, \frac{y}{t} \in S \right\} = \left\{ t(y, 1) \mid t > 0, y \in S \right\}.$$

et on définit $\tilde{h} : \tilde{S} \rightarrow \mathbb{R}$ par

$$\forall (y, t) \in \tilde{S}, \quad \tilde{h}(y, t) = th\left(\frac{y}{t}\right).$$

La fonction \tilde{h} est appelée *perspective (en anglais)* de h dans [45] et [46].

1.3.2. Changement de variables de Charnes-Cooper

Soient S un sous-ensemble non vide d'un \mathbb{R} -espace vectoriel X , $f : S \rightarrow \mathbb{R}$, $g : S \rightarrow \mathbb{R}$ deux fonctions. On suppose que

$$\forall x \in S, \quad g(x) > 0.$$

On considère le problème de programmation fractionnaire suivant :

$$(\mathcal{P}) \quad \text{Maximiser } \frac{f(x)}{g(x)}, \quad x \in S.$$

On pose

$$\tilde{S}_= = \tilde{S} \cap [\tilde{g} = 1].$$

et on considère le problème d'optimisation

$$(\tilde{\mathcal{P}}_=) \quad \text{Maximiser } \tilde{f}(y, t), \quad (y, t) \in \tilde{S}_=.$$

L'application $T : S \rightarrow \tilde{S}_=$ définie par

$$(C.C.T) \quad T(x) = \left(\frac{x}{g(x)}, \frac{1}{g(x)} \right).$$

s'apparente au changement de variables de Charnes-Cooper. C'est une bijection et sa bijection réciproque est définie par

$$T^{-1} : (y, t) \mapsto \frac{y}{t}.$$

On a pour tout $x \in S$ et tout $(y, t) \in \tilde{S}_=$

$$\tilde{f}(T(x)) = \frac{f(x)}{g(x)}, \quad t f(T^{-1}(y, t)) = \tilde{f}(y, t).$$

On déduit de ce qui précède le résultat suivant :

Lemme 1.3.1. *Les problèmes (\mathcal{P}) et $(\tilde{\mathcal{P}}_=)$ ont la même valeur optimal et*

$$T(\text{Argmax}(\mathcal{P})) = \text{Argmax}(\tilde{\mathcal{P}}_=).$$

Puisqu'il n'y a pas de restriction de signe sur la fonction f , alors le Lemme 1.3.1 reste vrai pour un problème de minimisation.

1.3.3. Compléments

Soient E un \mathbb{R} -espace vectoriel, $C \subset E$ et $F, G : C \rightarrow \mathbb{R}$ deux fonctions telles que $G(z) > 0, \forall z \in C$. On suppose que

$$\forall \tau > 0, \forall z \in C, \quad \tau z \in C, \quad F(\tau z) = \tau F(z) \text{ et } G(\tau z) = \tau G(z).$$

On pose

$$C_= = C \cap [G = 1], \quad C_{\leq} = C \cap [G \leq 1],$$

et on considère les problèmes d'optimisation

$$(\mathcal{P}_=) \quad \text{Maximiser } F(z), \quad z \in C_=,$$

et

$$(\mathcal{P}_{\leq}) \quad \text{Maximiser } F(z), \quad z \in C_{\leq}.$$

Lemme 1.3.2. *Si la fonction F est positive sur C alors les problèmes $(\mathcal{P}_=)$ et (\mathcal{P}_{\leq}) ont la même valeur optimale. De plus, si F n'est pas la fonction nulle,*

$$\text{Argmax}(\mathcal{P}_=) = \text{Argmax}(\mathcal{P}_{\leq}).$$

Preuve. D'abord, $C_- \subset C_\leq$ donc $v(\mathcal{P}_-) \leq v(\mathcal{P}_\leq)$.

Soit $z \in C_\leq$, alors $0 < G(z) \leq 1$ et, en posant $\tau = 1/G(z)$, $G(\tau z) = 1$, donc $\tau z \in C_-$ et, comme $\tau \geq 1$ et $F(z) \geq 0$,

$$F(z) \leq \tau F(z) = F(\tau z).$$

On a donc $F(z) \leq v(\mathcal{P}_-)$ pour tout $z \in C_\leq$ d'où $v(\mathcal{P}_\leq) \leq v(\mathcal{P}_-)$.

Si $\bar{z} \in \text{Argmax}(\mathcal{P}_-) \subset C_\leq$ et $z \in C_\leq$ alors, en reprenant ce que l'on a fait ci-dessus, en posant $\tau = 1/G(z)$, $\tau z \in C_-$ et

$$F(z) \leq F(\tau z) \leq F(\bar{z}).$$

On en déduit que $\bar{z} \in \text{Argmax}(\mathcal{P}_\leq)$. Ceci montre que

$$\text{Argmax}(\mathcal{P}_-) \subset \text{Argmax}(\mathcal{P}_\leq).$$

Si $\bar{z} \in \text{Argmax}(\mathcal{P}_\leq)$. Soit $\bar{\tau} = 1/G(\bar{z})$. On a $\bar{\tau} \geq 1$. D'autre part, $\bar{\tau}\bar{z} \in C_- \subset C_\leq$ donc

$$\bar{\tau}F(\bar{z}) = F(\bar{\tau}\bar{z}) \leq F(\bar{z}).$$

Comme F est positive et n'est pas la fonction nulle, alors $F(\bar{z}) > 0$, donc $\bar{\tau} \leq 1$. Finalement $\bar{\tau} = 1$, $\bar{z} \in C_-$ et $\bar{z} \in \text{Argmax}(\mathcal{P}_-)$, d'où

$$\text{Argmax}(\mathcal{P}_\leq) \subset \text{Argmax}(\mathcal{P}_-).$$

□

On reprend les notations de la sous-section précédente. On note

$$\tilde{S}_\leq = \tilde{S} \cap [\tilde{g} \leq 1].$$

On considère le problème

$$(\tilde{\mathcal{P}}_\leq) \quad \text{Maximiser } \tilde{f}(y, t), \quad (y, t) \in \tilde{S}_\leq.$$

On en déduit directement de ce qui précède le lemme suivant :

Lemme 1.3.3. *On a toujours $v(\tilde{\mathcal{P}}_-) = v(\tilde{\mathcal{P}}_\leq)$. De plus, si f est positive et n'est pas la fonction nulle alors*

$$\text{Argmax}(\tilde{\mathcal{P}}_-) = \text{Argmax}(\tilde{\mathcal{P}}_\leq).$$

1.4. Autour du S-lemma

1.4.1. Notations

Si $(H_1, \langle \cdot, \cdot \rangle_1)$, $(H_2, \langle \cdot, \cdot \rangle_2)$ sont deux espaces de Hilbert, on munit $H_1 \times H_2$ du produit scalaire $\langle \cdot, \cdot \rangle_{1,2} : (H_1 \times H_2) \times (H_1 \times H_2) \rightarrow \mathbb{R}$:

$$\langle (v_1, v_2), (u_1, u_2) \rangle_{1,2} = \langle v_1, u_1 \rangle_1 + \langle v_2, u_2 \rangle_2.$$

Étant donnés $A_1 : H_1 \rightarrow H_1$ linéaire auto-adjoint, $A_2 : H_2 \rightarrow H_2$ linéaire auto-adjoint, $B : H_1 \rightarrow H_2$ linéaire, on notera

$$\begin{bmatrix} A_1 & B^* \\ B & A_2 \end{bmatrix},$$

la forme quadratique définie sur $H_1 \times H_2$ par

$$(v, u) \rightarrow \langle A_1 v, v \rangle_1 + 2\langle Bv, u \rangle_2 + \langle A_2 u, u \rangle_2,$$

où B^* désigne l'adjoint de B .

Soit une fonctionnelle quadratique $f : v \rightarrow \langle Av, v \rangle + 2\langle b, v \rangle + c$ sur un espace de Hilbert $(H, \langle \cdot, \cdot \rangle)$, alors la forme quadratique associée à cette fonctionnelle quadratique est la forme quadratique définie sur $H \times \mathbb{R}$ par

$$(1.11) \quad \tilde{f}(v, s) = \langle Av, v \rangle + 2s\langle b, v \rangle + cs^2$$

On a :

$$(1.12) \quad f \geq 0 \iff \tilde{f} \geq 0.$$

Preuve. Si $\tilde{f} \geq 0$, il est immédiat en prenant $s = 1$ que l'on obtient $f \geq 0$.

Réciproquement, supposons que $f \geq 0$. Soit $s \neq 0$ et v quelconque, on a

$$s^2 f\left(\frac{v}{s}\right) = \langle Av, v \rangle + 2s\langle b, v \rangle + cs^2 = \tilde{f}(v, s) \geq 0$$

Maintenant quand $s \rightarrow 0$, $s \neq 0$, $s^2 f(v/s) \rightarrow \tilde{f}(v, 0)$ avec $s^2 f(v/s) \geq 0$ donc $\tilde{f}(v, 0) \geq 0$. □

1.4.2. Minimisation d'une fonctionnelle quadratique

Soit une fonctionnelle quadratique $v \rightarrow \langle Av, v \rangle + 2\langle b, v \rangle + c$ sur un espace de Hilbert $(H, \langle \cdot, \cdot \rangle)$. L'opérateur A est supposé linéaire, auto-adjoint et continu. On sait que (voir par exemple [17] Corollaire II.17) :

$$H = \ker A \oplus \overline{\text{Im } A}.$$

Donc l'opérateur $A|_{\overline{\text{Im } A}} : \overline{\text{Im } A} \rightarrow \text{Im } A$ est une bijection, on note

$$A^\dagger := A|_{\overline{\text{Im } A}}^{-1} : \text{Im } A \rightarrow \overline{\text{Im } A}$$

l'opérateur inverse. On suppose que $\text{Im } A$ est fermé (l'opérateur $A^\dagger : \text{Im } A \rightarrow \overline{\text{Im } A}$ est alors continu, auto-adjoint). On a :

$$\inf_{v \in H} \langle Av, v \rangle + 2\langle b, v \rangle + c = \begin{cases} -\infty & \text{si } A \not\geq 0, \\ -\infty & \text{si } A \geq 0 \text{ et } b \notin \text{Im } A, \\ -\langle A^\dagger b, b \rangle + c & \text{si } A \geq 0 \text{ et } b \in \text{Im } A. \end{cases}$$

Le résultat suivant est alors immédiat.

Proposition 1.4.1. Soit $(H, \langle \cdot, \cdot \rangle)$, un espace de Hilbert. Soient $A : H \rightarrow H$ un opérateur linéaire continu, auto-adjoint, $b \in H$ et $c \in \mathbb{R}$. Alors

$$\forall \xi \in H, \quad \langle A\xi, \xi \rangle + 2\langle b, \xi \rangle + c \geq 0 \iff \begin{cases} A \geq 0, b \in \text{Im } A \\ c - \langle A^\dagger b, b \rangle \geq 0. \end{cases}$$

1.4.3. Schur Complement Lemma

Proposition 1.4.2. Soient $(H_1, \langle \cdot, \cdot \rangle_1)$, $(H_2, \langle \cdot, \cdot \rangle_2)$ deux espaces de Hilbert de dimension finie, $P : H_1 \rightarrow H_1$ linéaire auto-adjoint, $R : H_2 \rightarrow H_2$ linéaire auto-adjoint, $Q : H_1 \rightarrow H_2$ linéaire. Alors

$$(1.13) \quad \begin{bmatrix} P & Q^* \\ Q & R \end{bmatrix} \geq 0 \iff \begin{cases} R \geq 0, \text{Im } Q \subset \text{Im } R \\ P - Q^* R^\dagger Q \geq 0. \end{cases}$$

Preuve. En effet

$$\begin{aligned} \begin{bmatrix} P & Q^* \\ Q & R \end{bmatrix} \geq 0 &\iff \forall (u, v) \in H_1 \times H_2, \quad \langle Pu, u \rangle + 2\langle Qu, v \rangle + \langle Rv, v \rangle \geq 0, \\ &\iff \forall u \in H_1, \quad R \geq 0, Qu \in \text{Im } R \text{ et } \langle Pu, u \rangle - \langle R^\dagger Qu, Qu \rangle \geq 0, \\ &\iff R \geq 0, \text{Im } Q \subset \text{Im } R \text{ et } P - Q^* R^\dagger Q \geq 0. \end{aligned}$$

□

Soient $(H, \langle \cdot, \cdot \rangle)$ un espace de Hilbert et $\|\cdot\|$ la norme associée à $\langle \cdot, \cdot \rangle$. Soient $t \in \mathbb{R}$, $y \in H$. On utilise le (1.13) avec $H_1 = \mathbb{R}$, $H_2 = H$, $P = tI_{\mathbb{R}}$ (que l'on désignera par t), $Q : s \rightarrow sy$ (que l'on désignera par y et on a alors $Q^* : v \rightarrow \langle y, v \rangle$ que l'on désignera par y^*), $R = tI_H$. Alors on obtient

$$\begin{aligned} \begin{bmatrix} t & y^* \\ y & tI_H \end{bmatrix} \geq 0 &\iff \begin{cases} tI_H \geq 0, y \in tH \\ t - \langle y, (tI_H)^\dagger y \rangle \geq 0, \end{cases} \\ &\iff [t = 0 \text{ et } y = 0] \text{ ou } \left[t > 0 \text{ et } t - \frac{1}{t} \|y\|^2 \geq 0 \right]. \end{aligned}$$

Finalement, pour tout $t \in \mathbb{R}$, $y \in H$,

$$(1.14) \quad \|y\| \leq t \iff \begin{bmatrix} t & y^* \\ y & tI_H \end{bmatrix} \geq 0.$$

Soient $t \in \mathbb{R}$, $y \in H$. On utilise le (1.13) avec $H_1 = \mathbb{R}$, $H_2 = H$, $P = tI_{\mathbb{R}}$ (que l'on désignera par t), $Q : s \rightarrow sy$ (que l'on désignera par y et on a alors $Q^* : v \rightarrow \langle y, v \rangle$ que l'on désignera par y^*), $R = I_H$. Alors on obtient

$$(1.15) \quad \|y\|^2 \leq t \iff \begin{bmatrix} t & y^* \\ y & I_H \end{bmatrix} \geq 0.$$

Soient $t \in \mathbb{R}$, $y \in H$. On vérifie sans peine que

$$(1.16) \quad \|y\|^2 \leq t \iff \left\| \left(y, \frac{t-1}{2} \right) \right\| \leq \frac{t+1}{4}.$$

1.4.4. S-lemma

Soit $(H, \langle \cdot, \cdot \rangle)$ un espace de Hilbert.

Théorème 1.4.1 (S-lemma, [39], Theorem 2.1). *Soit $s_i : H \rightarrow \mathbb{R}$, $i = 1, 2$, deux fonctionnelles quadratiques de la forme $s_i : \xi \rightarrow \frac{1}{2}\langle S_i \xi, \xi \rangle + \langle d_i, \xi \rangle + e_i$, $S_i : H \rightarrow H$ linéaire continu et auto-adjoint, $d_i \in H$ et $e_i \in \mathbb{R}$, $i = 1, 2$. Supposons qu'il existe $\xi_+ \in H$ tel que $s_2(\xi_+) > 0$. Les assertions suivantes sont équivalentes :*

- (1) $\forall \xi \in H, \quad s_2(\xi) \geq 0 \implies s_1(\xi) \geq 0,$
- (2) $\exists \lambda \geq 0 : \forall \xi \in H, \quad s_1(\xi) - \lambda s_2(\xi) \geq 0.$

De plus, s'il existe ξ_- tel que $s_1(\xi_-) < 0$, alors $\lambda > 0$.

1.5. Formule d'Ascoli

Soit (U, d) un espace métrique. Rappelons que pour tout $A \subset U$, la distance à A est l'application $d(\cdot, A) : U \rightarrow [0, +\infty]$ définie par

$$(1.17) \quad d(u, A) = \inf_{a \in A} d(u, a), \quad \forall u \in U.$$

Pour $u \in U$ et $\rho > 0$, on note $\mathbb{B}_d(u, \rho)$ la boule fermée de centre u et de rayon ρ pour la distance d définie par

$$(1.18) \quad \mathbb{B}_d(u, \rho) = \{v \in U \mid d(u, v) \leq \rho\}.$$

Lemme 1.5.1. *Pour tout $a \in U$ et $A \subset U$, on a*

$$(1.19) \quad d(a, A^c) = \sup\{\rho > 0 \mid \mathbb{B}_d(a, \rho) \subset A\},$$

où A^c désigne le complémentaire de A dans U .

Preuve. Notons $\Lambda_A := \{\rho > 0 \mid \mathbb{B}_d(u, \rho) \subset A\}$. Si $d(a, A^c) = +\infty$ alors $A = U$ et $\sup \Lambda_A = +\infty$. Supposons que $d(a, A^c) < +\infty$. Pour tout $\rho > 0$, on a

$$d(a, A^c) \leq \rho \iff \mathbb{B}_d(a, \rho) \cap A^c \neq \emptyset,$$

ce qui est équivalent à

$$(*) \quad d(a, A^c) > \rho \iff \mathbb{B}_d(a, \rho) \subset A.$$

Par conséquent, $d(a, A^c)$ est un majorant de Λ_A . Donc si $d(a, A^c) = 0$ alors $\Lambda_A = \emptyset$ et par convention $\sup \Lambda_A = 0$. Il reste donc le cas où $0 < d(a, A^c) < +\infty$. Dans ce cas, pour tout $\epsilon > 0$ tel que $d(a, A^c) - \epsilon > 0$, on a $d(a, A^c) - \epsilon \in \Lambda_A$ d'après (*). D'où $\sup \Lambda_A = d(a, A^c)$. \square

Notons que ce résultat reste vrai si on remplace la boule fermée par la boule ouverte. On suppose que $(U, \|\cdot\|)$ est un espace vectoriel normé de dual topologique U^* et $\langle \cdot, \cdot \rangle$

la forme bilinéaire standard entre U et U^* . On note $\|\cdot\|^* : U^* \rightarrow \mathbb{R}_+$ la norme duale de $\|\cdot\|$ définie par

$$\|u^*\|^* = \max_{\|u\| \leq 1} \langle u, u^* \rangle.$$

Soit $u^* \in U^*$ une forme linéaire continue non nulle et $\alpha \in \mathbb{R}$. On note

$$[\langle \cdot, u^* \rangle \leq \alpha] := \{u \in U \mid \langle u, u^* \rangle \leq \alpha\}.$$

Soit $d_{\|\cdot\|}$ la distance associée à la norme $\|\cdot\|$. D'après le lemme 1.5.1, pour tout $a \in A$, on a

$$d_{\|\cdot\|}(a, [\langle \cdot, u^* \rangle > \alpha]) = \sup\{\rho > 0 \mid \mathbb{B}_{d_{\|\cdot\|}}(a, \rho) \subset [\langle \cdot, u^* \rangle \leq \alpha]\}.$$

Par ailleurs, pour tout $\rho > 0$, on a :

$$\begin{aligned} \mathbb{B}_{d_{\|\cdot\|}}(a, \rho) \subset [\langle \cdot, u^* \rangle \leq \alpha] &\iff \langle a + u, u^* \rangle \leq \alpha, \quad \forall u \in U : \|u\| \leq \rho \\ &\iff \langle a, u^* \rangle + \rho \langle u, u^* \rangle \leq \alpha, \quad \forall u \in U : \|u\| \leq 1 \\ &\iff \langle a, u^* \rangle + \rho \max_{\|u\| \leq 1} \langle u, u^* \rangle \leq \alpha \\ &\iff \langle a, u^* \rangle + \rho \|u^*\|^* \leq \alpha. \end{aligned}$$

Il en résulte que

$$\begin{aligned} d_{\|\cdot\|}(a, [\langle \cdot, u^* \rangle > \alpha]) &= \sup\{\rho > 0 \mid \langle a, u^* \rangle + \rho \|u^*\|^* \leq \alpha\} \\ &= \begin{cases} \frac{\alpha - \langle a, u^* \rangle}{\|u^*\|^*} & \text{si } \alpha - \langle a, u^* \rangle > 0 \\ 0 & \text{si } \alpha - \langle a, u^* \rangle \leq 0. \end{cases} \end{aligned}$$

On a ainsi montré le résultat suivant connu sous le nom de la formule d'Ascoli.

Lemme 1.5.2 (formule d'Ascoli). *Pour toute forme linéaire continue non nulle $u^* \in U^*$, pour tout réel $\alpha \in \mathbb{R}$ et pour tout $a \in U$, on a :*

$$d_{\|\cdot\|}(a, [\langle \cdot, u^* \rangle > \alpha]) = \frac{(\alpha - \langle a, u^* \rangle)^+}{\|u^*\|^*}.$$

I. Contribution à la dualité

Dualité et régularisation en convexité généralisée : des niveaux aux épigraphes

2.1. Introduction

L'article [73] explicite le passage de la dualité quasi-convexe à la dualité convexe dans le cas des fonctions evenly quasi-convexes et evenly convexes. Ce travail est consacré à une approche générale de la méthodologie développée dans cet article. Étant donné une fonction $c : U \times V \rightarrow \mathbb{R}$ dite de couplage, et une partie non vide P de \mathbb{R} , principalement $P_1 = \mathbb{R}_+$, $P_2 = \mathbb{R}_+^*$, $P_3 = \mathbb{R}^*$, $P_4 = \{0\}$, la notion clé est celle d'ensemble (c, P) -élémentaire :

$$(2.1) \quad E_{v,r}^P = \{u \in U \mid r - c(u, v) \in P\},$$

et la notion d'ensemble (c, P) -régulier qui en découle, à savoir les intersections d'ensembles (c, P) -élémentaires. La (c, P) -enveloppe d'une partie A de U est l'intersection des ensembles (c, P) -élémentaires (ou réguliers) contenant A , notée $\langle A \rangle_{c,P}$, généralisant notamment l'enveloppe evenly convexe et l'enveloppe coaffine dans un espace vectoriel topologique, l'enveloppe convexe fermée et l'enveloppe affine fermée dans un espace vectoriel topologique localement convexe, la fermeture topologique (l'adhérence) dans un espace métrique, ... On décrit et compare $\langle A \rangle_{c,P_i}$, $i = 1, 2, 3, 4$ (Proposition 2.2.6, Corollaire 2.2.1, Proposition 2.2.8). Le lien entre la (c, P) -enveloppe des ensembles et la Γ -régularisée des fonctions par rapport à un couplage c (ou associé à c) est précisé par l'intermédiaire des fonctions indicatrices (Proposition 2.2.9, Théorème 2.2.2, Corollaire 2.2.2).

On s'intéresse à la classe des fonctions dont les tranches inférieures larges sont fermées pour l'opérateur $\langle \rangle_{c,P}$. On définit alors la (c, P) -régularisée $h^{\langle \rangle_{c,P}}$ par niveaux d'une fonction $h : U \rightarrow \overline{\mathbb{R}}$. On donne plusieurs expressions de $h^{\langle \rangle_{c,P}}$ (Théorème 2.3.1, Corollaire 2.3.1, Proposition 2.3.6) que l'on caractérise aussi comme la biconjugée de h par rapport à un couple de polarités duales (Δ, Δ^*) associé à l'opérateur $\langle \rangle_{c,P}$. S'agissant de définir la régularisée épigraphique d'une fonction $h : U \rightarrow \overline{\mathbb{R}}$, on travaille naturellement avec le couplage

$$\ddot{c} : (U \times \mathbb{R}) \times (V \times \mathbb{R}) \rightarrow \mathbb{R},$$

$$(2.2) \quad \ddot{c}((u, r), (v, s)) = c(u, v) + rs,$$

et les ensembles élémentaires (resp. réguliers) correspondants. On définit ainsi les fonctions (\check{c}, P) -épi-régulières comme celles dont l'épigraphe est (\check{c}, P) -régulier, et la (\check{c}, P) -épi-régularisée de h , notée $h^{(\check{c}, P)}$. On obtient des formules explicites de $h^{(\check{c}, P)}$ pour $i = 1, 2, 3, 4$ (Théorème 2.4.1, Théorème 2.4.2, Théorème 2.4.3). On montre (Théorème 2.4.5) que la (\check{c}, P_i) -épi-régularisée de h coïncide avec la biconjuguée de h relative à un couplage γ_i approprié :

$$(2.3) \quad h^{(\check{c}, P_i)} = h^{\gamma_i \gamma_i}, \quad i = 1, 2, 3,$$

généralisant ainsi des résultats obtenus dans [73], Théorèmes 3.1 et 6.1, et dans [70] Théorèmes 4.2 et 5.2.

Concernant le passage de la régularisation par niveaux à la régularisation épigraphique, on associe à chaque fonction $h : U \rightarrow \overline{\mathbb{R}}$, la fonction $H : U \times \mathbb{R} \rightarrow \overline{\mathbb{R}}$ définie par $H(u, t) = h(u) - t$. On établit (Proposition 2.5.1) que h est (\check{c}, P) -épi-régulière si et seulement si H est (\check{c}, P) -régulière par niveaux, si et seulement si H est (\check{c}, P) -épi-régulière, où $\check{c} : (U \times \mathbb{R} \times \mathbb{R}) \times (V \times \mathbb{R} \times \mathbb{R}) \rightarrow \mathbb{R}$ est le couplage défini par :

$$(2.4) \quad \check{c}((u, t, t'), (v, s, s')) := \check{c}((u, t), (v, s)) + t's' = c(u, v) + ts + t's'.$$

Des formules sont établies entre la biconjuguée par niveaux de H et la (\check{c}, P_i) -épi-régularisée de h pour $i = 1, 2, 3, 4$ (Théorème 2.5.2).

On exploite les résultats ci-dessus pour obtenir un théorème de dualité forte (Théorème 2.6.1) dans le cas des problèmes d'optimisation paramétrés (approche perturbative par niveaux). On termine par une application au cas spécifique des problèmes d'optimisation evenly quasi-convexe (Corollaire 2.6.1 et Corollaire 2.6.2).

2.2. Γ -régularisation de fonctions et enveloppe d'ensembles

On commence cette section par un rappel sur la conjugaison au sens de Moreau ([54]). Soient U, V deux ensembles non vides et $c : U \times V \rightarrow \mathbb{R}$ une fonction dite de couplage. Pour toute fonction $h : U \rightarrow \overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$, on définit sa c -conjuguée, notée h^c par :

$$(2.5) \quad h^c(v) := \sup_{u \in U} \{c(u, v) - h(u)\}, \quad \forall v \in V.$$

De façon analogue, en échangeant le rôle de U et V , on définit la c -conjuguée d'une fonction $k : V \rightarrow \overline{\mathbb{R}}$ par :

$$(2.6) \quad k^c(u) := \sup_{v \in V} \{c(u, v) - k(v)\}, \quad \forall u \in U.$$

La c -biconjuguée d'une fonction $h : U \rightarrow \overline{\mathbb{R}}$ est alors donnée par :

$$(2.7) \quad h^{cc}(u) = \sup_{v \in V} \{c(u, v) - h^c(v)\}, \quad \forall u \in U.$$

Exemple 2.2.1. Le cas le plus classique est celui où U est un espace vectoriel topologique de dual topologique V et c est la forme bilinéaire standard. On retrouve dans ce cas la conjugaison de Legendre-Fenchel.

Exemple 2.2.2. Citons quelques autres exemples de fonctions de couplages qui ont été considérés dans la littérature :

- (1) $U = V = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid x_1 > 0, \dots, x_n > 0\}$ et $c(u, v) = \min_{1 \leq i \leq n} v_i u_i$, ([63]),
- (2) (U, d) un espace métrique, $\alpha > 0$, $V = U$ et $c(u, v) = -\alpha d(u, v)$, ([41]),
- (3) (U, d) un espace métrique, $V = U \times \mathbb{R}_+^*$ et $c(u, (v, \alpha)) = -\alpha d(u, v)$, ([41]),
- (4) $U = V = \mathbb{R}^n$, $0 < \alpha \leq 1$, $\beta > 0$ et $c(u, v) = -\beta \|u - v\|^\alpha$, ([51]),
- (5) U un espace topologique, $V = \mathcal{C}(U, \mathbb{R})$, l'espace des fonctions réelles continues sur U et $c(u, v) = v(u)$, ([38],[56]).

Étant donné $A \subset U$, la fonction indicatrice de A , notée $i_A : U \rightarrow \overline{\mathbb{R}}$, est définie par :

$$i_A(u) = \begin{cases} 0 & \text{si } u \in A, \\ +\infty & \text{si } u \in U \setminus A. \end{cases}$$

La fonction vallée ([55]) de A , notée $\vartheta_A : U \rightarrow \overline{\mathbb{R}}$, est définie par :

$$\vartheta_A(u) = \begin{cases} -\infty & \text{si } u \in A, \\ +\infty & \text{si } u \in U \setminus A. \end{cases}$$

2.2.1. Γ -régularisation des fonctions

La notion de fonction affine continue conduit à la définition suivante :

Définition 2.2.1. On appelle fonction c -élémentaire définie sur U (resp. sur V), toute fonction de la forme $c(\cdot, v) - r$ (resp. $c(u, \cdot) - r$) où $v \in V$ (resp. $u \in U$) et $r \in \mathbb{R}$.

Définition 2.2.2. On appelle fonction c -régulière toute enveloppe supérieure de fonctions c -élémentaires.

Il s'agit donc des fonctions de la forme $\sup_{j \in J} \{c(\cdot, v_j) - r_j\}$, où J est un ensemble d'indices, $v_j \in V$, $r_j \in \mathbb{R}$ pour tout $j \in J$. Si $J = \emptyset$, on obtient la fonction $\{-\infty\}^U$. Par ailleurs, pour tout \bar{v} fixé dans V , on a

$$\sup_{n \geq 1} \{c(\cdot, \bar{v}) + n\} = \{+\infty\}^U.$$

Ce qui montre que la fonction $\{+\infty\}^U$ est c -régulière. Toute fonction c -élémentaire est évidemment c -régulière. On note $\Gamma_c(U)$ l'ensemble des fonctions c -régulières définies sur U à valeurs dans $\overline{\mathbb{R}}$.

Définition 2.2.3 ([54]). On appelle c -enveloppe ou Γ_c -régularisée d'une fonction $h : U \rightarrow \overline{\mathbb{R}}$, l'enveloppe supérieure des minorantes c -élémentaires de h .

On note h^{Γ_c} , la Γ_c -régularisée de h . Il résulte de la définition 2.2.2 que toute enveloppe supérieure de fonctions c -régulières est encore une fonction c -régulière. De ce fait, h^{Γ_c} est l'enveloppe supérieure des minorantes c -régulières de h , autrement dit la plus grande fonction c -régulière qui minore h , et on a $h = h^{\Gamma_c}$ si et seulement si h est c -régulière. On a aussi :

Théorème 2.2.1 ([54]). *La c -enveloppe ou Γ_c -régularisée d'une fonction $h : U \rightarrow \overline{\mathbb{R}}$ coïncide avec la c -biconjuguée de $h : h^{\Gamma_c} = h^{cc}$.*

Preuve. Pour tout $a \in U$, d'après la définition 2.2.3, on a :

$$\begin{aligned} h^{\Gamma_c}(a) &= \sup_{\substack{v \in V \\ r \in \mathbb{R}}} \{c(a, v) - r \mid c(u, v) - r \leq h(u), \forall u \in U\} \\ &= \sup_{\substack{v \in V \\ r \in \mathbb{R}}} \{c(a, v) - r \mid h^c(v) \leq r\} \\ &= \sup_{v \in V} \{c(a, v) - \inf_{r \in \mathbb{R}} \{r \mid h^c(v) \leq r\}\} \\ &= \sup_{v \in V} \{c(a, v) - h^c(v)\} \\ &= h^{cc}(a). \end{aligned}$$

□

Remarque 2.2.1. *On sait ([54]) que le théorème 2.2.1 reste vrai si la fonction c est à valeurs dans $\overline{\mathbb{R}}$. Il faut alors interpréter les conjuguées comme suit :*

$$(2.8) \quad h^c(v) = -\inf_{u \in U} \{h(u) - c(u, v)\}$$

avec les conventions usuelles

$$(2.9) \quad (+\infty) - (+\infty) = (-\infty) - (-\infty) = +\infty.$$

Nous aurons d'ailleurs à considérer des fonctions de couplage à valeurs dans $\overline{\mathbb{R}}$ dans les sections 2.3, 2.4 et 2.5.

Il existe une approche équivalente de la dualité convexe généralisée en termes de Φ -convexité [30], qui consiste à travailler avec un ensemble U et une classe de fonctions $\Phi \subset \overline{\mathbb{R}}^U$.

2.2.2. Enveloppes d'ensembles

Soit P un sous-ensemble non vide de \mathbb{R} . La définition suivante généralise la notion de demi-espaces.

Définition 2.2.4. *On appelle sous-ensemble (c, P) -élémentaire de U , tout ensemble de la forme $\{u \in U \mid r - c(u, v) \in P\}$, où $v \in V$ et $r \in \mathbb{R}$. On le note $E_{v,r}^P$.*

Si $P = \mathbb{R}$ alors $E_{v,r}^P = U$ pour tout $(v, r) \in V \times \mathbb{R}$. Dans ce cas, U est le seul sous-ensemble (c, P) -élémentaire de U . Les sous-ensembles (c, P) -élémentaires permettent de définir une notion d'enveloppe d'ensembles de la manière suivante :

Définition 2.2.5. On appelle (c, P) -enveloppe de $A \subset U$, l'intersection de tous les sous-ensembles (c, P) -élémentaires de U contenant A . La (c, P) -enveloppe de A est notée $\langle A \rangle_{c,P}$.

Remarque 2.2.2. Par définition, tout sous-ensemble (c, P) -élémentaire coïncide avec sa (c, P) -enveloppe.

Remarque 2.2.3. S'il n'existe pas de sous-ensemble (c, P) -élémentaire qui contient A alors $\langle A \rangle_{c,P} = U$.

Proposition 2.2.1. Si $P \neq \mathbb{R}$ alors $\langle \emptyset \rangle_{c,P} = \emptyset$.

Preuve. Soit $s \in \mathbb{R} \setminus P$. Supposons que $\langle \emptyset \rangle_{c,P} \neq \emptyset$. Soit $a \in \langle \emptyset \rangle_{c,P}$. Notons que a appartient à tous les ensembles (c, P) -élémentaires. On a donc

$$r - c(a, v) \in P \quad \forall (v, r) \in V \times \mathbb{R},$$

et en particulier, $s = (s + c(a, v)) - c(a, v) \in P$ ce qui est absurde. \square

Notons la caractérisation fort utile suivante du complémentaire de la (c, P) -enveloppe d'un ensemble.

Proposition 2.2.2. Pour tout $A \subset U$, et pour tout $u \in U$, on a

$$u \notin \langle A \rangle_{c,P} \iff \exists (v, r) \in V \times \mathbb{R} : A \subset E_{v,r}^P \text{ et } r - c(u, v) \notin P.$$

Preuve. Par définition de $\langle A \rangle_{c,P}$, on a : $u \notin \langle A \rangle_{c,P}$ si et seulement s'il existe un ensemble (c, P) -élémentaire de U qui contient A et ne contient pas u . \square

La proposition suivante montre que l'opérateur $\langle \cdot \rangle_{c,P}$ est un opérateur de fermeture (au sens algébrique du terme).

Proposition 2.2.3. L'opérateur $\langle \cdot \rangle_{c,P}$ vérifie les propriétés suivantes :

- (1) $A \subset \langle A \rangle_{c,P}, \quad \forall A \subset U,$
- (2) $A \subset B \subset U \implies \langle A \rangle_{c,P} \subset \langle B \rangle_{c,P},$
- (3) $\langle \langle A \rangle_{c,P} \rangle_{c,P} = \langle A \rangle_{c,P}, \quad \forall A \subset U.$

Preuve. (1) est une conséquence immédiate de la définition 2.2.5.

(2). Soit $A \subset B \subset U$. Tout sous-ensemble (c, P) -élémentaire de U contenant B contient aussi A . Alors l'intersection de tous les sous-ensembles (c, P) -élémentaires contenant B contient l'intersection de tous les sous-ensembles (c, P) -élémentaires contenant A .

(3). Il résulte de (1) que $\langle \langle A \rangle_{c,P} \rangle_{c,P} \supset \langle A \rangle_{c,P}$. Inversement, Soit $a \notin \langle A \rangle_{c,P}$. Il existe un sous-ensemble (c, P) -élémentaire E de U contenant A et ne contenant pas a . On en déduit que $a \notin E$ et $\langle A \rangle_{c,P} \subset E$. D'après la proposition 2.2.2, on a : $a \notin \langle \langle A \rangle_{c,P} \rangle_{c,P}$. \square

Définition 2.2.6. Un ensemble A est dit (c, P) -régulier si $A = \langle A \rangle_{c,P}$. On note $\mathcal{R}_{c,P}(U)$, l'ensemble des sous-ensembles (c, P) -réguliers de U .

Notons que les ensembles (c, P) -élémentaires sont (c, P) -réguliers (Remarque 2.2.2). Plus généralement, on a la proposition suivante :

Proposition 2.2.4. *Toute intersection d'ensembles (c, P) -réguliers est un ensemble (c, P) -régulier.*

Preuve. Soient $(A_j)_{j \in J}$, une famille d'ensembles (c, P) -réguliers et $A := \bigcap_{j \in J} A_j$. Soit $a \notin A$. Il existe $j \in J$ tel que $a \notin A_j$. Puisque A_j est (c, P) -régulier alors $a \notin \langle A_j \rangle_{c,P}$ et à fortiori $a \notin \langle A \rangle_{c,P}$. D'où $A \supset \langle A \rangle_{c,P}$, et finalement $A = \langle A \rangle_{c,P}$. \square

Proposition 2.2.5. *La (c, P) -enveloppe de $A \subset U$ est l'intersection de tous les ensembles (c, P) -réguliers contenant A .*

Preuve. Soit B l'intersection de tous les ensembles (c, P) -réguliers contenant A . Par définition de $\langle A \rangle_{c,P}$, on a $\langle A \rangle_{c,P} \supset B$. D'autre part $A \subset B$, d'où $\langle A \rangle_{c,P} \subset \langle B \rangle_{c,P} = B$ d'après la proposition 2.2.4. \square

Dans la suite, nous allons considérer les cas suivants pour P : $P_1 = \mathbb{R}_+$, $P_2 = \mathbb{R}_+^*$, $P_3 = \mathbb{R}^*$ et $P_4 = \{0\}$. Pour simplifier les notations, on notera $\langle \cdot \rangle_{c,i}$ au lieu de $\langle \cdot \rangle_{c,P_i}$ pour $i \in \{1, 2, 3, 4\}$. Pour $i \in \{1, 2, 3\}$, il est possible d'explicitier $\langle A \rangle_{c,i}$ comme suit :

Proposition 2.2.6. *Pour tout $A \subset U$, on a :*

$$(2.10) \quad \langle A \rangle_{c,1} = \{u \in U \mid c(u, v) \leq \sup_{a \in A} c(a, v), \forall v \in V\},$$

$$(2.11) \quad \langle A \rangle_{c,2} = \{u \in U \mid \forall v \in V, \exists a \in A \mid c(u, v) \leq c(a, v)\},$$

$$(2.12) \quad \langle A \rangle_{c,3} = \{u \in U \mid \forall v \in V, \exists a \in A \mid c(u, v) = c(a, v)\}.$$

Preuve. D'après la proposition 2.2.2, on a :

$$\begin{aligned} a \notin \langle A \rangle_{c,1} &\iff \exists (v, r) \in V \times \mathbb{R} : A \subset [c(\cdot, v) \leq r] \text{ et } r < c(a, v) \\ &\iff \exists (v, r) \in V \times \mathbb{R} : \sup_{u \in A} c(u, v) \leq r < c(a, v) \\ &\iff \exists v \in V : \sup_{u \in A} c(u, v) < c(a, v). \end{aligned}$$

Donc $a \in \langle A \rangle_{c,1} \iff \forall v \in V, c(a, v) \leq \sup_{u \in A} c(u, v)$, d'où (2.10).

$$\begin{aligned} a \notin \langle A \rangle_{c,2} &\iff \exists (v, r) \in V \times \mathbb{R} : A \subset [c(\cdot, v) < r] \text{ et } r \leq c(a, v) \\ &\iff \exists (v, r) \in V \times \mathbb{R} : c(u, v) < r \leq c(a, v), \forall u \in A \\ &\iff \exists v \in V : c(u, v) < c(a, v), \forall u \in A. \end{aligned}$$

Donc $a \in \langle A \rangle_{c,2} \iff \forall v \in V, \exists u \in A : c(a, v) \leq c(u, v)$, d'où (2.11).

$$\begin{aligned} a \notin \langle A \rangle_{c,3} &\iff \exists (v, r) \in V \times \mathbb{R} : A \subset [c(\cdot, v) \neq r] \text{ et } r = c(a, v) \\ &\iff \exists v \in V : c(u, v) \neq c(a, v), \forall u \in A. \end{aligned}$$

Donc $a \in \langle A \rangle_{c,3} \iff \forall v \in V, \exists u \in A : c(a, v) = c(u, v)$, d'où (2.12). \square

Remarque 2.2.4. Notons qu'on ne peut pas se débarrasser du paramètre réel r dans la définition de $\langle A \rangle_{c,4}$.

Examinons la situation dans le cas vectoriel : Soit U un espace vectoriel topologique de dual topologique V et c la forme bilinéaire standard. Les fonctions c -élémentaires sont alors les fonctions affines continues, et on a :

- (1) **cas de P_1** : les ensembles (c, P_1) -élémentaires sont l'ensemble vide, U , et les demi-espaces fermés. Si, de plus, U est localement convexe, alors, d'après le théorème de séparation de Hahn-Banach et (2.10), $\langle A \rangle_{c,1}$ n'est autre que l'enveloppe convexe fermée $\overline{\text{conv}}A$ de A .
- (2) **cas de P_2** : les ensembles (c, P_2) -élémentaires sont l'ensemble vide, U , et les demi-espaces ouverts. La (c, P_2) -enveloppe d'un sous-ensemble de U est son enveloppe evenly convexe ([36], [51], ...).
- (3) **cas de P_3** : les ensembles (c, P_3) -élémentaires sont l'ensemble vide, U , et les complémentaires d'hyperplans affines fermés. La (c, P_3) -enveloppe d'un sous-ensemble de U coïncide avec son enveloppe evenly coaffine, selon la terminologie introduite dans [65]. Ainsi, un sous-ensemble $A \subset U$ est (c, P_3) -régulier si et seulement si son complémentaire dans U est une réunion d'hyperplans affines fermés. Notons que ([65], Corollaire 2.2) A est evenly convexe (c'est-à-dire (c, P_2) -régulier) si et seulement si A est coaffine (c'est-à-dire (c, P_3) -régulier) et convexe ou encore si et seulement si A est (c, P_3) -régulier et connexe.
- (4) **cas de P_4** : les ensembles (c, P_4) -élémentaires sont l'ensemble vide, U et les hyperplans affines fermés. Si, de plus, U est localement convexe, il résulte de la proposition 2.2.2 et du théorème de séparation de Hahn-Banach que la (c, P_4) -enveloppe de $A \subset U$ est son enveloppe affine fermée si $A \neq \emptyset$ et l'ensemble vide si $A = \emptyset$.

Revenons au cas général.

Proposition 2.2.7. Soient P et Q deux sous-ensembles non vides de \mathbb{R} . On suppose que tout ensemble (c, P) -élémentaire est (c, Q) -régulier. On a alors

$$\langle A \rangle_{c,Q} \subset \langle A \rangle_{c,P}, \quad \forall A \subset U.$$

Preuve. Soit $a \notin \langle A \rangle_{c,P}$. Par définition, il existe un ensemble (c, P) -élémentaire E tel que $A \subset E$ et $a \notin E$. Par hypothèse, E est (c, Q) -régulier. Il résulte de la proposition 2.2.5 que $a \notin \langle A \rangle_{c,Q}$. \square

Corollaire 2.2.1. Pour tout $A \subset U$, on a :

$$(2.13) \quad \langle A \rangle_{c,1} \supset \langle A \rangle_{c,2} \supset \langle A \rangle_{c,3} \quad \text{et} \quad \langle A \rangle_{c,4} \supset \langle A \rangle_{c,3}.$$

Preuve. Soit $v \in V$ et $r \in \mathbb{R}$. Il est clair que :

$$\{u \in U \mid c(u, v) \leq r\} = \bigcap_{s > r} \{u \in U \mid c(u, v) < s\}.$$

Par conséquent les ensembles (c, P_1) -élémentaires sont (c, P_2) -réguliers et d'après la proposition 2.2.7, on a $\langle A \rangle_{c,1} \supset \langle A \rangle_{c,2}$.

On vérifie sans peine que :

$$\{u \in U \mid c(u, v) < r\} = \bigcap_{s \geq r} \{u \in U \mid c(u, v) \neq s\},$$

$$\{u \in U \mid c(u, v) = r\} = \bigcap_{s \neq r} \{u \in U \mid c(u, v) \neq s\},$$

d'où il résulte que $\langle A \rangle_{c,2} \supset \langle A \rangle_{c,3} \subset \langle A \rangle_{c,4}$. \square

On déduit du corolaire 2.2.1 la comparaison suivante entre les ensembles $\mathcal{R}_{c,P_i}(U)$:

$$(2.14) \quad \mathcal{R}_{c,P_1}(U) \subset \mathcal{R}_{c,P_2}(U) \subset \mathcal{R}_{c,P_3}(U) \text{ et } \mathcal{R}_{c,P_4}(U) \subset \mathcal{R}_{c,P_3}(U).$$

Remarque 2.2.5. Notons qu'on a :

- $P_1 \supset P_2$ et $\mathcal{R}_{c,P_1}(U) \subset \mathcal{R}_{c,P_2}(U)$.
- $P_3 \supset P_2$ et $\mathcal{R}_{c,P_3}(U) \supset \mathcal{R}_{c,P_2}(U)$.
- $\mathcal{R}_{c,P_1}(U) \subset \mathcal{R}_{c,P_3}(U)$ alors que P_1 et P_3 ne sont pas comparables au sens de l'inclusion.

En particulier,

- Dans le cas des espaces topologiques localement convexes, on retrouve ainsi le fait que tout ensemble convexe fermé est evenly convexe.
- Il a été observé dans [65] (Proposition 2.10) que dans le cas des espaces topologiques localement convexes, tout ensemble evenly convexe est evenly coafine. Notons que ce résultat reste vrai même si l'espace n'est pas localement convexe.

Proposition 2.2.8. *Supposons que le couplage c vérifie la condition suivante :*

$$(2.15) \quad \forall v \in V, \exists w \in V \mid -c(\cdot, v) = c(\cdot, w).$$

Alors on a :

$$(2.16) \quad \mathcal{R}_{c,P_4}(U) \subset \mathcal{R}_{c,P_1}(U) \subset \mathcal{R}_{c,P_2}(U) \subset \mathcal{R}_{c,P_3}(U).$$

Preuve. Notons que d'après (2.14), il suffit de montrer que $\mathcal{R}_{c,P_4}(U) \subset \mathcal{R}_{c,P_1}(U)$.

Soit $(v, r) \in V \times \mathbb{R}$. On a

$$[c(\cdot, v) = r] = [c(\cdot, v) \leq r] \cap [-c(\cdot, v) \leq -r]$$

Par hypothèse, il existe $w \in V$ tel que $-c(\cdot, v) = c(\cdot, w)$. Par conséquent

$$[c(\cdot, v) = r] = [c(\cdot, v) \leq r] \cap [c(\cdot, w) \leq -r].$$

Donc tout ensemble (c, P_4) -élémentaire est (c, P_1) -régulier. Il résulte donc de la proposition 2.2.7 que

$$\langle A \rangle_{c,1} \subset \langle A \rangle_{c,4}, \quad \forall A \subset U.$$

Par conséquent, si A est (c, P_4) -régulier alors il est (c, P_1) -régulier. \square

Exemple 2.2.3. On suppose que $U = V = \mathbb{R}^n$, et c est défini par $c(u, v) = \|u - v\|$ avec $\|\cdot\|$ la norme euclidienne de \mathbb{R}^n . Les ensembles (c, P_4) -élémentaires non triviaux sont les sphères (non convexes) tandis que les ensembles (c, P_1) -élémentaires non triviaux sont les boules fermées (convexes fermées). Dans ce cas, les ensembles $\mathcal{R}_{c, P_4}(U)$ et $\mathcal{R}_{c, P_1}(U)$ ne sont pas comparables. Notons que l'hypothèse (2.15) n'est pas satisfaite.

Proposition 2.2.9. *Pour tout $A \subset U$, on a :*

$$\langle A \rangle_{c,1} = [i_A^{\Gamma c} \leq 0].$$

Preuve. D'après (2.10), on a :

$$\begin{aligned} a \notin \langle A \rangle_{c,1} &\iff \exists v \in V : i_A^c(v) < c(a, v) \\ &\iff 0 < \sup_{v \in V} \{c(a, v) - i_A^c(v)\} \\ &\iff 0 < i_A^{\Gamma c}(a) \\ &\iff 0 \notin [i_A^{\Gamma c} \leq 0]. \end{aligned}$$

D'où $\langle A \rangle_{c,1} = [i_A^{\Gamma c} \leq 0]$. \square

Soient W un ensemble non vide et $d : U \times W \rightarrow \mathbb{R}$ une fonction de couplage. On rappelle que d'après (2.10), l'opérateur de fermeture $\langle \cdot \rangle_{d,1} : 2^U \rightarrow 2^U$ est défini par :

$$(2.17) \quad \langle A \rangle_{d,1} = \{u \in U \mid d(u, w) \leq \sup_{a \in A} d(a, w), \forall w \in W\}.$$

Le résultat suivant fait le lien entre l'enveloppe des ensembles et celle des fonctions par l'intermédiaire de fonction indicatrice.

Théorème 2.2.2. *Supposons que le couplage d vérifie la condition*

$$(2.18) \quad \forall (w, \beta) \in W \times \mathbb{R}_+^*, \quad \exists \bar{w} \in W \mid \beta d(\cdot, w) = d(\cdot, \bar{w}).$$

Alors pour tout $A \subset U$ tel que $\text{dom } i_A^d \neq \emptyset$, on a $i_A^{\Gamma d} = i_{\langle A \rangle_{d,1}}$.

Preuve. Soit $b \in U$.

(1) Supposons que $b \notin \langle A \rangle_{d,1}$. D'après (2.17), il existe $(w, \epsilon) \in W \times \mathbb{R}_+^*$ tel que :

$$d(b, w) - \sup_{a \in A} d(a, w) \geq \epsilon.$$

D'après (2.18), on a

$$\forall n \geq 1, \quad \exists w_n \in W : nd(\cdot, w) = d(\cdot, w_n).$$

Par conséquent,

$$n\epsilon \leq d(b, w_n) - \sup_{a \in A} d(a, w_n) = d(b, w_n) - i_A^d(w_n) \leq i_A^{\Gamma d}(b), \quad \forall n \geq 1.$$

D'où $i_A^{\Gamma d}(b) = +\infty$.

(2) Supposons $b \in \langle A \rangle_{d,1}$. D'après (2.17), on a

$$d(b, w) - \sup_{a \in A} d(a, w) \leq 0, \quad \forall w \in W.$$

Donc

$$i_A^{\Gamma d}(b) = \sup_{w \in W} \left\{ d(b, w) - \sup_{a \in A} d(a, w) \right\} \leq 0.$$

Soit $w \in \text{dom } i_A^d$. D'après (2.18), on a

$$\forall n \geq 1, \exists w_n \in W : \frac{1}{n}d(\cdot, w) = d(\cdot, w_n).$$

Par conséquent,

$$\frac{1}{n} \left(d(b, w) - \sup_{a \in A} d(a, w) \right) = d(b, w_n) - \sup_{a \in A} d(a, w_n) \leq i_A^{\Gamma d}(b), \quad \forall n \geq 1.$$

Donc

$$0 = \lim_{n \rightarrow +\infty} \frac{1}{n} \left(d(b, w) - \sup_{a \in A} d(a, w) \right) \leq i_A^{\Gamma d}(b),$$

d'où finalement $i_A^{\Gamma d}(b) = 0$.

□

Remarque 2.2.6. L'hypothèse (2.18) est vérifiée par les couplages (1), (3) et (5) de l'exemple 2.2.2. Elle est également vérifiée par le couplage bilinéaire standard. Elle ne l'est pas pour les couplages (2) et (4) de l'exemple 2.2.2.

On définit maintenant le couplage $\widehat{c}: U \times (V \times \mathbb{R}_+^*) \rightarrow \mathbb{R}$ associé à c comme suit :

$$(2.19) \quad \widehat{c}(u, (v, \alpha)) = \alpha c(u, v).$$

Notons que \widehat{c} vérifie l'hypothèse (2.18).

Lemme 2.2.1. *La classe des ensembles (c, P_i) -élémentaires coïncide avec la classe des ensembles (\widehat{c}, P_i) -élémentaires pour $i = 1, 2, 3, 4$.*

Preuve. Soit $i \in \{1, 2, 3, 4\}$. Soit $E := \{u \in U \mid r - c(u, v) \in P_i\}$ un ensemble (c, P_i) -élémentaire. On a $E = \{u \in U \mid r - \widehat{c}(u, (v, 1)) \in P_i\}$ ce qui montre que E est un ensemble (\widehat{c}, P_i) -élémentaire. Inversement, soit $E := \{u \in U \mid r - \widehat{c}(u, (v, \alpha)) \in P_i\}$ un ensemble (\widehat{c}, P_i) -élémentaire. On a

$$\begin{aligned} E &= \{u \in U \mid r - \alpha c(u, v) \in P_i\} \\ &= \{u \in U \mid r/\alpha - c(u, v) \in P_i\}, \quad \text{car } \alpha P_i = P_i, \quad \forall \alpha > 0. \end{aligned}$$

Ce qui montre que E est un ensemble (c, P_i) -élémentaire. □

Proposition 2.2.10. *Pour tout $i \in \{1, 2, 3, 4\}$, on a :*

$$\langle A \rangle_{\tilde{c}, i} = \langle A \rangle_{c, i}, \quad \forall A \subset U.$$

Preuve. Conséquence immédiate du lemme 2.2.1. □

Corollaire 2.2.2. *Pour tout $A \subset U$ tel que $\text{dom } i_A^c \neq \emptyset$, on a :*

$$(2.20) \quad i_A^{\Gamma_{\tilde{c}}} = i_{\langle A \rangle_{c, 1}}.$$

Preuve. Observons que

$$\text{dom } i_A^{\tilde{c}} = \{(v, \alpha) \in V \times \mathbb{R}_+^* : \sup_{u \in A} \alpha c(u, v) < +\infty\} = \text{dom } i_A^c \times \mathbb{R}_+^*.$$

Donc $\text{dom } i_A^{\tilde{c}} \neq \emptyset$. Puisque \tilde{c} vérifie (2.18), alors d'après le théorème 2.2.2, on a $i_A^{\Gamma_{\tilde{c}}} = i_{\langle A \rangle_{\tilde{c}, 1}}$. La conclusion résulte donc la proposition 2.2.10. □

On associe à c un autre couplage $\tilde{c} : U \times (V \times \mathbb{R}^*) \rightarrow \mathbb{R}$ défini par :

$$(2.21) \quad \tilde{c}(u, (v, \alpha)) = \alpha c(u, v).$$

Notons que \tilde{c} vérifie l'hypothèse (2.18) .

Proposition 2.2.11. *L'opérateur de fermeture $\langle \cdot \rangle_{\tilde{c}, 1}$ est donné par :*

$$(2.22) \quad \langle A \rangle_{\tilde{c}, 1} = \left\{ u \in U \mid \inf_{a \in A} c(a, v) \leq c(u, v) \leq \sup_{a \in A} c(a, v), \quad \forall v \in V \right\}.$$

Preuve. D'après (2.10), on a :

$$\begin{aligned} a \in \langle A \rangle_{\tilde{c}, 1} &\iff \tilde{c}(a, (v, \alpha)) \leq \sup_{u \in A} \tilde{c}(u, (v, \alpha)), \quad \forall (v, \alpha) \in V \times \mathbb{R}^* \\ &\iff \alpha c(a, v) \leq \sup_{u \in A} \alpha c(u, v), \quad \forall (v, \alpha) \in V \times \mathbb{R}^* \\ &\iff \begin{cases} \alpha c(a, v) \leq \alpha \sup_{u \in A} c(u, v), \quad \forall (v, \alpha) \in V \times \mathbb{R}_+^* \\ \text{et} \\ \alpha c(a, v) \leq \alpha \inf_{u \in A} c(u, v), \quad \forall (v, \alpha) \in V \times \mathbb{R}_-^* \end{cases} \\ &\iff \inf_{u \in A} c(u, v) \leq c(a, v) \leq \sup_{u \in A} c(u, v), \quad \forall v \in V. \end{aligned}$$

□

Notons que si pour tout $(v, \alpha) \in V \times \mathbb{R}^*$, il existe $w \in V$ tel que $\alpha c(., v) = c(., w)$ (ce qui est le cas du couplage bilinéaire standard) alors pour tout $A \subset U$, on a :

$$(2.23) \quad \langle A \rangle_{\tilde{c}, P} = \langle A \rangle_{c, P},$$

puisque dans ce cas, un ensemble est (c, P) -élémentaire si et seulement si il est (\tilde{c}, P) -élémentaire.

2.3. Régularisations et conjugaisons par niveaux

On introduit dans cette section la notion de (c, P) -régularisée par niveaux pour les fonctions. On montre que cette régularisée s'interprète comme une biconjugée relative à un couple de polarités duales. On donne enfin d'autres expressions de cette régularisée.

2.3.1. Régularisation par niveaux d'une fonction

Définition 2.3.1. Une fonction $h : U \rightarrow \overline{\mathbb{R}}$ est dite (c, P) -régulière par niveaux si ses tranches inférieures larges sont des sous-ensembles (c, P) -réguliers de U . Autrement dit :

$$\langle [h \leq r] \rangle_{c,P} = [h \leq r], \quad \forall r \in \mathbb{R}.$$

On note $\mathcal{N}_{c,P}(U)$, l'ensemble des fonctions (c, P) -régulières par niveaux définies sur U à valeurs dans $\overline{\mathbb{R}}$. Notons que $\mathcal{N}_{c,P}(U)$ contient la fonction $\{-\infty\}^U$.

Proposition 2.3.1. Soit $(h_i)_{i \in I} \subset \overline{\mathbb{R}}^U$ une famille de fonctions (c, P) -régulières par niveaux. La fonction $h := \sup_{i \in I} h_i$ est (c, P) -régulière par niveaux.

Preuve. Soit $r \in \mathbb{R}$. On a $[h \leq r] = \bigcap_{i \in I} [h_i \leq r]$. Il résulte de la proposition 2.2.5 que l'ensemble $[h \leq r]$ est (c, P) -régulier. \square

On définit la (c, P) -régularisée par niveaux d'une fonction comme suit.

Définition 2.3.2. La (c, P) -régularisée par niveaux d'une fonction $h : U \rightarrow \overline{\mathbb{R}}$, est l'enveloppe supérieure des minorantes (c, P) -régulières par niveaux de h . On la note $h^{(\cdot)}_{c,P}$.

D'après la proposition 2.3.1, $h^{(\cdot)}_{c,P}$ est la plus grande fonction (c, P) -régulière par niveaux qui minore h .

Exemple 2.3.1. Si U est un espace vectoriel topologique de dual topologique V et c la forme bilinéaire standard, les fonctions (c, P_2) -régulières par niveaux sont les fonctions evenly quasiconvexes. Si, de plus, U est localement convexe alors les fonctions (c, P_1) -régulières par niveaux sont les fonctions quasi-convexes semi-continues inférieurement.

Exemple 2.3.2. Soit U un espace métrique, $V = \mathcal{C}(U, \mathbb{R})$ l'espace des fonctions continues de U à valeur dans \mathbb{R} , $c : U \times V \rightarrow \mathbb{R}$ défini par $c(u, v) = v(u)$. Une fonction $h : U \rightarrow \mathbb{R} \cup \{+\infty\}$ est (c, P_1) -régulière par niveaux si et seulement si h est semi-continue inférieurement ([32], Corollaire 2).

Proposition 2.3.2. Toute fonction c -élémentaire est une fonction (c, P_i) -régulière par niveaux pour $i = 1, 2, 3$. Plus précisément, on a

$$\Gamma_c(U) \subset \mathcal{N}_{c,P_1}(U) \subset \mathcal{N}_{c,P_2}(U) \subset \mathcal{N}_{c,P_3}(U) \quad \text{et} \quad \mathcal{N}_{c,P_4}(U) \subset \mathcal{N}_{c,P_3}(U).$$

Preuve. Soit $h := c(\cdot, v) - r$, une fonction c -élémentaire. Pour tout $t \in \mathbb{R}$, on a

$$[h \leq t] = \{u \in U \mid t + r - c(u, v) \geq 0\},$$

qui est évidemment un ensemble (c, P_1) -élémentaire, et donc (c, P_1) -régulier. Ce qui montre que $\Gamma_c(U) \subset \mathcal{N}_{c, P_1}(U)$. Les autres inclusions sont des conséquences immédiates de (2.14). \square

Remarque 2.3.1. Les fonctions c -élémentaires ne sont pas nécessairement (c, P_4) -régulières par niveaux. Par exemple si U est un espace vectoriel topologique de dual topologique V et c la forme bilinéaire standard, on ne peut écrire un demi-espace fermé comme intersection d'hyperplans affines fermés.

Exemple 2.3.3. Soit $n \geq 1$ un entier naturel. On suppose que $U = \mathbb{R}^n = V$ et c est le produit scalaire standard de \mathbb{R}^n . Considérons les fonctions h_1 et h_2 définies par

$$h_1 : \mathbb{R}^n \rightarrow \llbracket 0, n \rrbracket := \{0, 1, \dots, n\}$$

$$h_1(x) = \begin{cases} 0 & \text{si } x = 0 \\ \max\{i \in \llbracket 1, n \rrbracket \mid x_i \neq 0\} & \text{si } x \neq 0, \end{cases}$$

$$h_2 : \mathbb{R}^n \rightarrow \llbracket 0, n \rrbracket$$

$$h_2(x) = \begin{cases} 0 & \text{si } x_i \neq 0, \forall i \in \llbracket 1, n \rrbracket \\ \max\{i \in \llbracket 1, n \rrbracket \mid x_i = 0\} & \text{sinon .} \end{cases}$$

Pour tout $r \in \mathbb{R}$, on a :

$$[h_1 \leq r] = \begin{cases} \emptyset & \text{si } r < 0 \\ \{x \in \mathbb{R}^n \mid x_{i+1} = \dots = x_n = 0\} & \text{si } i \leq r < i + 1, \quad i = 0, 1, \dots, n - 1 \\ \mathbb{R}^n & \text{si } n \leq r, \end{cases}$$

$$[h_2 \leq r] = \begin{cases} \emptyset & \text{si } r < 0 \\ \{x \in \mathbb{R}^n \mid x_{i+1} \neq 0, \dots, x_n \neq 0\} & \text{si } i \leq r < i + 1, \quad i = 0, 1, \dots, n - 1 \\ \mathbb{R}^n & \text{si } n \leq r. \end{cases}$$

Il apparaît clairement que :

- (1) h_1 est une fonction (c, P_4) -régulière par niveaux (les tranches non triviales de h_1 sont des sous-espaces vectoriels de \mathbb{R}^n). Donc $h_1 \in \mathcal{N}_{c, P_i}(U)$, $i = 1, 2, 3, 4$ d'après (2.16).
- (2) Le complémentaire d'une tranche non triviale de h_2 est une réunion d'hyperplans. Cela prouve que h_2 est une fonction (c, P_3) -régulière par niveaux. Notons que h_2 n'est pas (c, P_2) -régulière par niveaux (à fortiori elle n'est pas (c, P_1) -régulière par niveaux ni à fortiori (c, P_4) -régulière par niveaux) puisque

$$[h_2 \leq n - 1] = \{x \in \mathbb{R}^n \mid x_n \neq 0\}$$

n'est pas convexe.

Exemple 2.3.4. On prend $U = V = \mathbb{R}$, c le produit standard de \mathbb{R} . Soit la fonction $h : U \rightarrow \overline{\mathbb{R}}$ définie par :

$$h(x) = \begin{cases} 0 & \text{si } x \neq 0 \\ +\infty & \text{si } x = 0. \end{cases}$$

Pour tout $t \in \mathbb{R}$, on a :

$$[h \leq t] = \begin{cases} \emptyset & \text{si } t < 0 \\ \mathbb{R}^* & \text{si } t \geq 0. \end{cases}$$

Donc h est (c, P_3) -régulière par niveaux mais n'est pas quasi-convexe.

2.3.2. Décomposition de l'opérateur de fermeture $\langle \cdot \rangle_{c,P}$

On considère l'application $\Delta_{c,P} : 2^U \rightarrow 2^{V \times \mathbb{R}}$ définie par

$$(2.24) \quad \Delta_{c,P}(A) = \{(v, r) \in V \times \mathbb{R} \mid A \subset E_{v,r}^P\},$$

que l'on notera simplement Δ s'il n'y a pas d'ambiguïté. L'application Δ est une polarité au sens qu'elle transforme toute réunion en intersection.

Proposition 2.3.3. *Pour toute famille $(A_i)_{i \in I}$ de sous-ensembles de U . On a :*

$$\Delta \bigcup_{i \in I} A_i = \bigcap_{i \in I} \Delta A_i.$$

Preuve. Par définition de l'opérateur Δ , on a :

$$\begin{aligned} \Delta \bigcup_{i \in I} A_i &:= \left\{ (v, r) \in V \times \mathbb{R} \mid \bigcup_{i \in I} A_i \subset E_{v,r}^P \right\} \\ &= \left\{ (v, r) \in V \times \mathbb{R} \mid A_i \subset E_{v,r}^P, \quad \forall i \in I \right\} \\ &= \bigcap_{i \in I} \left\{ (v, r) \in V \times \mathbb{R} \mid A_i \subset E_{v,r}^P \right\} \\ &= \bigcap_{i \in I} \Delta A_i. \end{aligned}$$

□

À la polarité Δ , est associée la polarité duale $\Delta^* : 2^{V \times \mathbb{R}} \rightarrow 2^U$ définie par :

$$(2.25) \quad \Delta^*(B) = \bigcup \{A \in 2^U \mid B \subset \Delta(A)\}.$$

Lorsque A (resp. B) est un singleton $\{u\}$ (resp. $\{(v, r)\}$), on note $\Delta(a)$ (resp. $\Delta^*(v, r)$) au lieu de $\Delta(\{a\})$ (resp. $\Delta^*(\{(v, r)\})$).

Notons que pour tout $(v, r) \in V \times \mathbb{R}$, pour tout $u \in U$ on a

$$(2.26) \quad u \in \Delta^*(v, r) \iff (v, r) \in \Delta(u) \iff u \in E_{v,r}^P,$$

et donc $\Delta^*(v, r) = E_{v,r}^P$.

Puisque Δ^* est une polarité, on a pour tout $B \subset V \times \mathbb{R}$,

$$(2.27) \quad \Delta^*(B) = \Delta^*\left(\bigcup_{(v,r) \in B} \{(v, r)\}\right) = \bigcap_{(v,r) \in B} \Delta^*(v, r) = \bigcap_{(v,r) \in B} E_{v,r}^P.$$

L'opérateur $\langle \cdot \rangle_{c,P}$ se décompose comme suit.

Proposition 2.3.4. *Pour tout $A \subset U$, on a :*

$$(\Delta^* \circ \Delta)(A) = \langle A \rangle_{c,P}.$$

Preuve. Soit $A \subset U$. On a :

$$(\Delta^* \circ \Delta)(A) = \Delta^* (\{(v, r) \in V \times \mathbb{R} \mid A \subset E_{v,r}^P\}) = \bigcap_{A \subset E_{v,r}^P} E_{v,r}^P = \langle A \rangle_{c,P}.$$

□

2.3.3. Conjugaisons associées aux polarités Δ et Δ^* ([71, 72])

La conjuguée d'une fonction $h : U \rightarrow \overline{\mathbb{R}}$ par rapport à la polarité Δ est la fonction $h^\Delta : V \times \mathbb{R} \rightarrow \overline{\mathbb{R}}$ définie par

$$(2.28) \quad h^\Delta(v, r) := \sup_{u \notin \Delta^*(v, r)} -h(u) = \sup_{u \notin E_{v,r}^P} -h(u).$$

De même la conjuguée par rapport à Δ^* d'une fonction $k : V \times \mathbb{R} \rightarrow \overline{\mathbb{R}}$ est la fonction $k^{\Delta^*} : U \rightarrow \overline{\mathbb{R}}$ définie par

$$(2.29) \quad k^{\Delta^*}(u) := \sup_{(v, r) \notin \Delta(u)} -k(v, r) = \sup_{u \notin E_{v,r}^P} -k(v, r).$$

Ainsi, la biconjuguée par rapport aux polarités Δ , Δ^* d'une fonction $h : U \rightarrow \overline{\mathbb{R}}$ est la fonction $h^{\Delta\Delta^*} : U \rightarrow \overline{\mathbb{R}}$ définie par

$$(2.30) \quad h^{\Delta\Delta^*}(a) := \sup_{(v, r) \notin \Delta(a)} \inf_{u \notin \Delta^*(v, r)} h(u) = \sup_{a \notin E_{v,r}^P} \inf_{u \notin E_{v,r}^P} h(u).$$

On sait ([71]) que cette biconjuguée s'interprète au moyen de la fonction de couplage $\delta : U \times (V \times \mathbb{R}) \rightarrow \overline{\mathbb{R}}$ défini par

$$\delta(u, (v, r)) = \begin{cases} 0 & \text{si } u \notin E_{v,r}^P \\ -\infty & \text{si } u \in E_{v,r}^P. \end{cases}$$

Pour tout $h : U \rightarrow \overline{\mathbb{R}}$, on a ainsi $h^\Delta = h^\delta$ et $h^{\Delta\Delta^*} = h^{\delta\delta}$. De plus

Théorème 2.3.1 ([71]). *La (c, P) -régularisée par niveaux d'une fonction $h : U \rightarrow \overline{\mathbb{R}}$ coïncide avec la biconjuguée de h par rapport aux polarités Δ , Δ^* : $h^{\langle \rangle}_{c,P} = h^{\Delta\Delta^*}$.*

Corollaire 2.3.1. *Pour tout sous-ensemble A de U , on a :*

$$i_A^{\langle \rangle}_{c,P} = i_{\langle A \rangle_{c,P}} \quad \text{et} \quad \vartheta_A^{\langle \rangle}_{c,P} = \vartheta_{\langle A \rangle_{c,P}}.$$

Preuve. Soit $A \subset U$. Soit $a \in A$. Il résulte du théorème 2.3.1 que

$$i_A^{\langle \rangle}_{c,P}(a) = \sup_{a \notin E_{v,r}^P} \inf_{u \notin E_{v,r}^P} i_A(u).$$

On distingue deux cas :

- $a \notin \langle A \rangle_{c,P}$. Il existe $(v, r) \in V \times \mathbb{R}$ tel que $a \notin E_{v,r}^P$ et $A \subset E_{v,r}^P$. Par conséquent,

$$\inf_{u \notin E_{v,r}^P} i_A(u) = +\infty.$$

- $a \in \langle A \rangle_{c,P}$. Pour tout $(v,r) \in V \times \mathbb{R}$ tel que $a \notin E_{v,r}^P$ il existe $u \in A$ tel que $u \notin E_{v,r}^P$. Par conséquent

$$\inf_{u \notin E_{v,r}^P} i_A(u) = 0.$$

On en déduit que

$$\begin{aligned} i_A^{(\cdot)c,P}(a) &= \sup_{a \notin E_{v,r}^P} \inf_{u \notin E_{v,r}^P} i_A(u) \\ &= \begin{cases} 0 & \text{si } a \in \langle A \rangle_{c,P} \\ +\infty & \text{si } a \notin \langle A \rangle_{c,P} \end{cases} \\ &= i_{\langle A \rangle_{c,P}}(a). \end{aligned}$$

De même, on a

$$\begin{aligned} \vartheta_A^{(\cdot)c,P}(a) &= \sup_{a \notin E_{v,r}^P} \inf_{u \notin E_{v,r}^P} \vartheta_A(u) \\ &= \begin{cases} -\infty & \text{si } a \in \langle A \rangle_{c,P} \\ +\infty & \text{si } a \notin \langle A \rangle_{c,P} \end{cases} \\ &= \vartheta_{\langle A \rangle_{c,P}}(a). \end{aligned}$$

□

Proposition 2.3.5. *Pour toute fonction $h : U \rightarrow \overline{\mathbb{R}}$, pour tout $t \in \mathbb{R}$ on a :*

$$[h^{(\cdot)c,P} \leq t] = \bigcap_{s > t} \langle [h \leq s] \rangle_{c,P}.$$

Preuve. Soit $s > t$ et $a \notin \langle [h \leq s] \rangle_{c,P}$. Il existe $(\bar{v}, \bar{r}) \in V \times \mathbb{R}$ tel que

$$a \notin E_{\bar{v}, \bar{r}}^P \quad \text{et} \quad [h \leq s] \subset E_{\bar{v}, \bar{r}}^P.$$

On en déduit que

$$h^{(\cdot)c,P}(a) = \sup_{a \notin E_{v,r}^P} \inf_{u \notin E_{v,r}^P} h(u) \geq \inf_{u \notin E_{\bar{v}, \bar{r}}^P} h(u) \geq s > t.$$

Réciproquement, soit $a \notin [h^{(\cdot)c,P} \leq t]$. Il existe $s \in \mathbb{R}$ tel que $h^{(\cdot)c,P}(a) > s > t$. D'après le théorème 2.3.1, il existe $(v,r) \in V \times \mathbb{R}$ tel que

$$a \notin E_{v,r}^P \quad \text{et} \quad \inf_{u \notin E_{v,r}^P} h(u) > s,$$

d'où $[h \leq s] \subset E_{v,r}^P$, et finalement $a \notin \langle [h \leq s] \rangle_{c,P}$. □

2.3.4. Autres expressions de la (c,P) -régularisée par niveaux d'une fonction

Soit $h : U \rightarrow \overline{\mathbb{R}}$ et $a \in U$. Définissons les ensembles $\mathcal{I}_h(a)$ et $\mathcal{J}_h(a)$ par :

$$(2.31) \quad \mathcal{I}_h(a) = \{t \in \mathbb{R} \mid a \notin \langle [h \leq t] \rangle_{c,P}\} \quad \text{et} \quad \mathcal{J}_h(a) = \{t \in \mathbb{R} \mid a \in \langle [h \leq t] \rangle_{c,P}\}.$$

Les ensembles $\mathcal{I}_h(a)$ et $\mathcal{J}_h(a)$ sont deux intervalles (dont l'un peut être vide) disjoints de \mathbb{R} tel que $\mathcal{I}_h(a) \cup \mathcal{J}_h(a) = \mathbb{R}$ et pour tout $(r, s) \in \mathcal{I}_h(A) \times \mathcal{J}_h(a)$, on a $r \leq s$. Il en résulte que

$$(2.32) \quad \sup \mathcal{I}_h(a) = \inf \mathcal{J}_h(a).$$

Proposition 2.3.6. *Pour toute fonction $h : U \rightarrow \overline{\mathbb{R}}$, et $a \in U$, on a :*

$$h^{(\cdot)}_{c,P}(a) = \sup \{t \in \mathbb{R} : a \notin \langle [h \leq t] \rangle_{c,P}\} = \inf \{t \in \mathbb{R} : a \in \langle [h \leq t] \rangle_{c,P}\}.$$

Preuve. Soit $t \in \mathcal{I}_h(a)$. Il existe $(v, r) \in V \times \mathbb{R}$ tel que $a \notin E_{v,r}^P$ et $[h \leq t] \subset E_{v,r}^P$. On en déduit que

$$\inf_{u \notin E_{v,r}^P} h(u) \geq t, \text{ et donc } \sup_{a \notin E_{v,r}^P} \inf_{u \notin E_{v,r}^P} h(u) \geq t.$$

D'après le théorème 2.3.1, on a $h^{(\cdot)}_{c,P}(a) \geq t$. D'où $h^{(\cdot)}_{c,P}(a) \geq \sup \mathcal{I}_h(a)$. Inversement, soit $t < h^{(\cdot)}_{c,P}(a)$. Alors $a \notin [h^{(\cdot)}_{c,P} \leq t]$ et d'après la proposition 2.3.5, il existe $s > t$ tel que $a \notin \langle [h \leq s] \rangle_{c,P}$. Par conséquent, $\sup \mathcal{I}_h(a) \geq s > t$. D'où $\sup \mathcal{I}_h(a) \geq h^{(\cdot)}_{c,P}(a)$. \square

2.4. Régularisations et conjugaisons épigraphiques

2.4.1. Généralités

Pour définir la régularisée épigraphique d'une fonction $h : U \rightarrow \overline{\mathbb{R}}$, il nous faut une notion d'ensemble régulier dans l'espace produit $U \times \mathbb{R}$ contenant l'épigraphe de h . Cela nous amène à considérer le couplage $\ddot{c} : (U \times \mathbb{R}) \times (V \times \mathbb{R}) \rightarrow \mathbb{R}$ défini par :

$$(2.33) \quad \ddot{c}((u, r), (v, s)) = c(u, v) + rs.$$

Les sous-ensembles élémentaires de $U \times \mathbb{R}$ sont définis en utilisant le couplage \ddot{c} . Nous noterons

$$(2.34) \quad \ddot{E}_{v,s,t}^P := \{(u, r) \in U \times \mathbb{R} : t - \ddot{c}((u, r), (v, s)) \in P\},$$

l'ensemble (\ddot{c}, P) -élémentaire associé au triplet $(v, s, t) \in V \times \mathbb{R} \times \mathbb{R}$. L'opérateur de fermeture associé à cette famille d'ensembles élémentaires est noté par $\langle \cdot \rangle_{\ddot{c}, P}$. On a la formule suivante pour les ensembles de la forme $A \times \mathbb{R}$.

Lemme 2.4.1. *Pour tout sous-ensemble A de U , on a :*

$$\langle A \times \mathbb{R} \rangle_{\ddot{c}, P} = \langle A \rangle_{c, P} \times \mathbb{R}.$$

Preuve. Soit $(a, r) \notin \langle A \times \mathbb{R} \rangle_{\ddot{c}, P}$. Il existe $(v, s, t) \in V \times \mathbb{R} \times \mathbb{R}$ tel que

$$A \times \mathbb{R} \subset \ddot{E}_{v,s,t}^P \quad \text{et} \quad (a, r) \notin \ddot{E}_{v,s,t}^P.$$

En particulier

$$A \subset E_{v,t-rs}^P \quad \text{et} \quad a \notin E_{v,t-rs}^P.$$

Il en résulte que, $a \notin \langle A \rangle_{c,P}$, et donc $\langle A \rangle_{c,P} \times \mathbb{R} \subset \langle A \times \mathbb{R} \rangle_{\check{c},P}$. Réciproquement, soit $(a, r) \notin \langle A \rangle_{c,P} \times \mathbb{R}$. Il est clair que $a \notin \langle A \rangle_{c,P}$. Il existe $(v, t) \in V \times \mathbb{R}$ tel que

$$A \subset E_{v,t}^P \quad \text{et} \quad a \notin E_{v,t}^P.$$

On en déduit que $A \times \mathbb{R} \subset \mathbb{E}_{v,0,t}^P$ et $(a, r) \notin \mathbb{E}_{v,0,t}^P$. Donc $(a, r) \notin \langle A \times \mathbb{R} \rangle_{\check{c},P}$. \square

Remarque 2.4.1. Il résulte du lemme 2.4.1 qu'un sous-ensemble $A \subset U$ est (c, P) -régulier si et seulement si $A \times \mathbb{R}$ est un sous-ensemble (\check{c}, P) -régulier de $U \times \mathbb{R}$.

2.4.2. Régularisation épigraphique

Définition 2.4.1. Une fonction $h : U \rightarrow \overline{\mathbb{R}}$ est dite (\check{c}, P) -épi-régulière si son épigraphe est un sous-ensemble (\check{c}, P) -régulier de $U \times \mathbb{R}$.

On note $\mathcal{E}_{\check{c},P}(U)$ le sous-ensemble de $\overline{\mathbb{R}}^U$ constitué des fonctions (\check{c}, P) -épi-régulières. Cet ensemble contient la fonction $\{-\infty\}^U$.

Proposition 2.4.1. Soit $(h_i)_{i \in I} \subset \overline{\mathbb{R}}^U$ une famille de fonctions (\check{c}, P) -épi-régulières. La fonction $h := \sup_{i \in I} h_i$ est (\check{c}, P) -épi-régulière.

Preuve. Observons que

$$\text{epi } h = \bigcap_{i \in I} \text{epi } h_i.$$

Il en résulte que $\text{epi } h$ est (\check{c}, P) -régulier d'après la proposition 2.2.4. \square

La proposition suivant donne un lien entre les ensembles $\mathcal{N}_{c,P}(U)$ et $\mathcal{E}_{\check{c},P}(U)$.

Proposition 2.4.2. Toute fonction (\check{c}, P) -épi-régulière est une fonction (c, P) -régulière par niveaux.

Preuve. Soit $h \in \mathcal{E}_{\check{c},P}(U)$ et $r \in \mathbb{R}$. Soit $a \notin [h \leq r]$. Alors $(a, r) \notin \text{epi } h = \langle \text{epi } h \rangle_{\check{c},P}$. Il existe $(v, s, t) \in V \times \mathbb{R} \times \mathbb{R}$ tel que

$$\text{epi } h \subset \ddot{E}_{v,s,t}^P \quad \text{et} \quad (a, r) \notin \ddot{E}_{v,s,t}^P.$$

En particulier,

$$[h \leq r] \subset E_{v,t-rs}^P \quad \text{et} \quad a \notin E_{v,t-rs}^P.$$

Par conséquent $[h \leq r] \supset \langle [h \leq r] \rangle_{c,P}$, et finalement $[h \leq r] = \langle [h \leq r] \rangle_{c,P}$. \square

Proposition 2.4.3. On a :

$$\Gamma_c(U) \subset \Gamma_{\widehat{c}}(U) \subset \mathcal{E}_{\check{c},P_1}(U) \subset \mathcal{E}_{\check{c},P_2}(U) \subset \mathcal{E}_{\check{c},P_3}(U) \supset \mathcal{E}_{\check{c},P_4}(U).$$

Preuve. Par définition de \widehat{c} , on a $\Gamma_c(U) \subset \Gamma_{\widehat{c}}(U)$. Soit $(v, r, \alpha) \in V \times \mathbb{R} \times \mathbb{R}_+^*$.

$$\text{epi } \{\widehat{c}(\cdot, (v, \alpha)) - r\} = \{(u, t) \in U \times \mathbb{R} \mid \alpha c(u, v) - r \leq t\} = \ddot{E}_{v, -\frac{1}{\alpha}, \frac{r}{\alpha}}^{P_1}.$$

On déduit que toute fonction \widehat{c} -élémentaire est une fonction (\check{c}, P_1) -épi-régulière. Par conséquent $\Gamma_{\widehat{c}}(U) \subset \mathcal{E}_{\check{c},P_1}(U)$. Les autres inclusions sont des conséquences directes de (2.14). \square

On définit la régularisée d'une fonction au sens des épigraphes comme suit.

Définition 2.4.2. La (\check{c}, P) -épi-régularisée d'une fonction $h : U \rightarrow \overline{\mathbb{R}}$ est l'enveloppe supérieure de ses minorantes (\check{c}, P) -épi-régulières. On la note $h^{(\check{c}, P)}$.

Remarque 2.4.2. Il résulte de la proposition 2.4.1 que $h^{(\check{c}, P)}$ est la plus grande fonction (\check{c}, P) -épi-régulière qui minore h .

Proposition 2.4.4. Pour toute fonction $h : U \rightarrow \overline{\mathbb{R}}$, on a :

$$(2.35) \quad \langle \text{epi } h \rangle_{\check{c}, P} \subset \text{epi } h^{(\check{c}, P)}.$$

Preuve. Par définition $h^{(\check{c}, P)} \leq h$. On en déduit que $\text{epi } h \subset \text{epi } h^{(\check{c}, P)}$ qui est un ensemble (\check{c}, P) -régulier par définition, d'où (2.35) en prenant la (\check{c}, P) -enveloppe. \square

L'inclusion (2.35) peut être stricte. En effet, soit $h : \mathbb{R} \rightarrow \mathbb{R} \cup \{+\infty\}$ définie par

$$h(x) = \begin{cases} \sqrt{x} & \text{si } x > 0 \\ +\infty & \text{sinon.} \end{cases}$$

On a $\langle \text{epi } h \rangle_{\check{c}, 2} = \mathbb{R}_+^* \times \mathbb{R}_+^*$ qui n'est pas un épigraphe .

Notons cependant que si $\langle \text{epi } h \rangle_{\check{c}, P}$ est un épigraphe, c'est nécessairement l'épigraphe de $h^{(\check{c}, P)}$:

Proposition 2.4.5. Pour toute fonction $h : U \rightarrow \overline{\mathbb{R}}$ telle que $\langle \text{epi } h \rangle_{\check{c}, P}$ est un épigraphe, on a

$$\langle \text{epi } h \rangle_{\check{c}, P} = \text{epi } h^{(\check{c}, P)}.$$

Preuve. Soit k la fonction définie par $\text{epi } k = \langle \text{epi } h \rangle_{\check{c}, P}$. Par définition k est (\check{c}, P) -épi-régulière et plus petite que h . Il en résulte que $k \leq h^{(\check{c}, P)}$. Soit g une minorante (\check{c}, P) -épi-régulière de h . Donc $\text{epi } g \supset \text{epi } h$ et en passant à la (\check{c}, P) -enveloppe, on obtient $\text{epi } g \supset \langle \text{epi } h \rangle_{\check{c}, P} = \text{epi } k$. D'où $g \leq k$, et finalement $k = h^{(\check{c}, P)}$. \square

Corollaire 2.4.1. Pour tout $A \subset U$, on a : $\vartheta_A^{(\check{c}, P)} = \vartheta_{\langle A \rangle_{c, P}}$.

Preuve. En effet,

$$\begin{aligned} \langle \text{epi } \vartheta_A \rangle_{\check{c}, P} &= \langle A \times \mathbb{R} \rangle_{\check{c}, P} \quad (\text{par définition de } \vartheta_A) \\ &= \langle A \rangle_{c, P} \times \mathbb{R} \quad (\text{d'après le lemme 2.4.1}) \\ &= \text{epi } \vartheta_{\langle A \rangle_{c, P}}. \end{aligned}$$

\square

Théorème 2.4.1. Soit $i \in \{1, 2\}$. Pour toute fonction $h : U \rightarrow \overline{\mathbb{R}}$, on a

$$h^{(\check{c}, i)} = \max \{ h^{\Gamma_{\check{c}}}, \vartheta_{\langle \text{dom } h \rangle_{c, i}} \} = h^{\Gamma_{\check{c}}} + i_{\langle \text{dom } h \rangle_{c, i}}.$$

Preuve. Soit $i \in \{1, 2\}$. D'après la proposition 2.4.3, $\Gamma_{\check{c}}(U) \subset \mathcal{E}_{\check{c}, P_i}(U)$. Donc pour toute fonction $h : U \rightarrow \overline{\mathbb{R}}$, $h^{\Gamma_{\check{c}}}$ est une fonction (\check{c}, P_i) -épi-régulière. Par ailleurs,

$$\text{epi } \vartheta_{\langle \text{dom } h \rangle_{c, i}} = \langle \text{dom } h \rangle_{c, i} \times \mathbb{R}$$

qui est un ensemble (\ddot{c}, P_i) -régulier d'après le lemme 2.4.1. D'après la proposition 2.4.1, la fonction $\max \{h^{\Gamma\bar{c}}, \vartheta_{\langle \text{dom } h \rangle_{c,i}}\}$ est (\ddot{c}, P_i) -épi-régulière. Il en résulte que

$$\max \{h^{\Gamma\bar{c}}, \vartheta_{\langle \text{dom } h \rangle_{c,i}}\} \leq h^{\langle \ddot{c}, i \rangle},$$

d'où

$$\text{epi } h^{\langle \ddot{c}, i \rangle} \subset (\text{epi } h^{\Gamma\bar{c}}) \cap (\langle \text{dom } h \rangle_{c,i} \times \mathbb{R}).$$

Notons que si $\text{dom } h = \emptyset$ alors l'inclusion

$$(\text{epi } h^{\Gamma\bar{c}}) \cap (\langle \text{dom } h \rangle_{c,i} \times \mathbb{R}) \subset \text{epi } h^{\langle \ddot{c}, i \rangle}$$

est immédiate. On suppose donc que $\text{dom } h \neq \emptyset$. Soit $(a, \bar{r}) \notin \text{epi } h^{\langle \ddot{c}, i \rangle}$. On distingue deux cas :

1^{er} cas : $i = 1$. Il existe $(v, s, t) \in V \times \mathbb{R} \times \mathbb{R}$ tel que

$$t - \bar{r}s - c(a, v) < 0 \quad \text{et} \quad \text{epi } h \subset \{(u, r) : t - rs - c(u, v) \geq 0\}.$$

Puisque $\text{dom } h \neq \emptyset$, alors $s \leq 0$.

- Si $s = 0$, on a

$$t - c(a, v) < 0 \quad \text{et} \quad \text{epi } h \subset \{(u, r) : t - c(u, v) \geq 0\},$$

et en particulier $a \notin \langle \text{dom } h \rangle_{c,1}$, d'où $(a, \bar{r}) \notin \text{epi } \vartheta_{\langle \text{dom } h \rangle_{c,1}}$.

- Si $s < 0$, on a

$$\frac{t}{s} - \bar{r} - \frac{1}{s}c(a, v) > 0 \quad \text{et} \quad \text{epi } h \subset \{(u, r) : \frac{t}{s} - r - \frac{1}{s}c(u, v) \leq 0\}.$$

On en déduit que

$$h(u) \geq \frac{t}{s} - \frac{1}{s}c(u, v), \quad \forall u \in U.$$

Par conséquent

$$h^{\Gamma\bar{c}}(u) \geq \frac{t}{s} - \frac{1}{s}c(u, v), \quad \forall u \in U.$$

et en particulier

$$h^{\Gamma\bar{c}}(a) \geq \frac{t}{s} - \frac{1}{s}c(a, v) > \bar{r} + \frac{1}{s}c(a, v) - \frac{1}{s}c(a, v) = \bar{r},$$

d'où $(a, \bar{r}) \notin \text{epi } h^{\Gamma\bar{c}}$.

2^e cas : $i = 2$. Il existe $(v, s, t) \in V \times \mathbb{R} \times \mathbb{R}$ tel que

$$t - \bar{r}s - c(a, v) \leq 0 \quad \text{et} \quad \text{epi } h \subset \{(u, r) : t - rs - c(u, v) > 0\}.$$

Puisque $\text{dom } h \neq \emptyset$, $s \leq 0$.

- Si $s = 0$, on a

$$t - c(a, v) \leq 0 \quad \text{et} \quad \text{epi } h \subset \{(u, r) : t - c(u, v) > 0\},$$

et en particulier $a \notin \langle \text{dom } h \rangle_{c,2}$, d'où $(a, \bar{r}) \notin \text{epi } \vartheta_{\langle \text{dom } h \rangle_{c,2}}$.

- Si $s < 0$, on a

$$\frac{t}{s} - \bar{r} - \frac{1}{s}c(a, v) \geq 0 \quad \text{et} \quad \text{epi } h \subset \{(u, r) : \frac{t}{s} - r - \frac{1}{s}c(u, v) < 0\}.$$

On en déduit que

$$h(u) > \frac{t}{s} - \frac{1}{s}c(u, v), \quad \forall u \in U.$$

Par conséquent

$$h^{\Gamma\tilde{c}}(u) > \frac{t}{s} - \frac{1}{s}c(u, v), \quad \forall u \in U.$$

et en particulier

$$h^{\Gamma\tilde{c}}(a) > \frac{t}{s} - \frac{1}{s}c(a, v) \geq \bar{r} + \frac{1}{s}c(a, v) - \frac{1}{s}c(a, v) = \bar{r},$$

d'où $(a, \bar{r}) \notin \text{epi } h^{\Gamma\tilde{c}}$.

En résumé nous avons établi que

$$\text{epi } h^{\langle \tilde{c}, i \rangle} = (\text{epi } h^{\Gamma\tilde{c}}) \cap (\langle \text{dom } h \rangle_{c,i} \times \mathbb{R}).$$

□

Corollaire 2.4.2. Soit $i \in \{1, 2\}$ et $h : U \rightarrow \overline{\mathbb{R}}$. On a :

$$h^{\langle \tilde{c}, i \rangle}(a) = h^{\tilde{c}}(a), \quad \forall a \in \langle \text{dom } h \rangle_{c,i}.$$

Preuve. D'après le théorème 2.4.1, pour tout $a \in U$, on a :

$$h^{\langle \tilde{c}, i \rangle}(a) = h^{\Gamma\tilde{c}}(a) + i_{\langle \text{dom } h \rangle_{c,i}}(a).$$

En particulier,

$$h^{\langle \tilde{c}, i \rangle}(a) = h^{\Gamma\tilde{c}}(a), \quad \forall a \in \langle \text{dom } h \rangle_{c,i}.$$

□

Proposition 2.4.6. Toute fonction \tilde{c} -élémentaire est une fonction (\tilde{c}, P_3) -épi-régulière. On a donc $\Gamma_{\tilde{c}}(U) \subset \mathcal{E}_{\tilde{c}, P_3}(U)$.

Preuve. Pour tout $(v, r, \alpha) \in V \times \mathbb{R} \times \mathbb{R}^*$, on a :

$$\begin{aligned} \text{epi } \{\tilde{c}(\cdot, (v, \alpha)) - r\} &= \begin{cases} \{(u, t) \in U \times \mathbb{R} \mid \frac{t}{\alpha} + \frac{r}{\alpha} - c(u, v) \geq 0\} & \text{si } \alpha > 0 \\ \{(u, t) \in U \times \mathbb{R} \mid \frac{t}{\alpha} + \frac{r}{\alpha} - c(u, v) \leq 0\} & \text{si } \alpha < 0 \end{cases} \\ &= \begin{cases} \{(u, t) \in U \times \mathbb{R} \mid \frac{r}{\alpha} - \tilde{c}((u, t), (v, -\frac{1}{\alpha})) \geq 0\} & \text{si } \alpha > 0 \\ \{(u, t) \in U \times \mathbb{R} \mid \frac{r}{\alpha} - \tilde{c}((u, t), (v, -\frac{1}{\alpha})) \leq 0\} & \text{si } \alpha < 0 \end{cases} \\ &= \begin{cases} \bigcap_{s > \frac{r}{\alpha}} \{(u, t) \in U \times \mathbb{R} \mid s - \tilde{c}((u, t), (v, -\frac{1}{\alpha})) \neq 0\} & \text{si } \alpha > 0 \\ \bigcap_{s < \frac{r}{\alpha}} \{(u, t) \in U \times \mathbb{R} \mid s - \tilde{c}((u, t), (v, -\frac{1}{\alpha})) \neq 0\} & \text{si } \alpha < 0. \end{cases} \end{aligned}$$

Par conséquent les fonctions \tilde{c} -élémentaires sont (\tilde{c}, P_3) -épi-régulières. On a donc $\Gamma_{\tilde{c}}(U) \subset \mathcal{E}_{\tilde{c}, P_3}(U)$ d'après la proposition 2.4.1. \square

Théorème 2.4.2. *Pour toute fonction $h : U \rightarrow \overline{\mathbb{R}}$ on a*

$$h^{(\cdot)\tilde{c},3} = \max \left\{ h^{\Gamma_{\tilde{c}}}, \vartheta_{\langle \text{dom } h \rangle_{c,3}} \right\} = h^{\Gamma_{\tilde{c}}} + i_{\langle \text{dom } h \rangle_{c,3}}.$$

Preuve. D'après la proposition 2.4.6 la fonction $h^{\Gamma_{\tilde{c}}}$ est (\tilde{c}, P_3) -épi-régulière. D'après le lemme 2.4.1 la fonction $\vartheta_{\langle \text{dom } h \rangle_{c,3}}$ est (\tilde{c}, P_3) -épi-régulière car

$$\text{epi } \vartheta_{\langle \text{dom } h \rangle_{c,3}} = \langle \text{dom } h \rangle_{c,3} \times \mathbb{R}.$$

On en déduit que la fonction $\max \left\{ h^{\Gamma_{\tilde{c}}}, \vartheta_{\langle \text{dom } h \rangle_{c,3}} \right\}$ est (\tilde{c}, P_3) -épi-régulière d'après la proposition 2.4.1. Il en résulte que

$$\max \left\{ h^{\Gamma_{\tilde{c}}}, \vartheta_{\langle \text{dom } h \rangle_{c,3}} \right\} \leq h^{(\cdot)\tilde{c},3},$$

d'où

$$\text{epi } h^{(\cdot)\tilde{c},3} \subset \text{epi } \max \left\{ h^{\Gamma_{\tilde{c}}}, \vartheta_{\langle \text{dom } h \rangle_{c,3}} \right\} = (\text{epi } h^{\Gamma_{\tilde{c}}}) \cap (\langle \text{dom } h \rangle_{c,3} \times \mathbb{R}).$$

Notons que si $\text{dom } h = \emptyset$, l'inclusion

$$(\text{epi } h^{\Gamma_{\tilde{c}}}) \cap (\langle \text{dom } h \rangle_{c,3} \times \mathbb{R}) \subset \text{epi } h^{(\cdot)\tilde{c},3}$$

est immédiate. Supposons donc que $\text{dom } h \neq \emptyset$. Soit $(a, \bar{r}) \notin \text{epi } h^{(\cdot)\tilde{c},3}$. Il existe $(v, s, t) \in V \times \mathbb{R} \times \mathbb{R}$ tel que

$$t - \bar{r}s - c(a, v) = 0 \quad \text{et} \quad \text{epi } h \subset \text{epi } h^{(\cdot)\tilde{c},3} \subset \{(u, r) \in U \times \mathbb{R} : t - rs - c(u, v) \neq 0\}.$$

Si $s = 0$ alors $a \notin \langle \text{dom } h \rangle_{c,3}$ et on a terminé. Supposons que $s \neq 0$. On a :

$$\frac{t}{s} - \frac{1}{s}c(a, v) = \bar{r} \quad \text{et} \quad \text{epi } h \subset \{(u, r) : \frac{t}{s} - \frac{1}{s}c(u, v) \neq r\}.$$

Ainsi, on a :

$$\frac{t}{s} = \bar{r} + \frac{1}{s}c(a, v) \quad \text{et} \quad \text{epi } h \subset (U \times \mathbb{R}) \setminus \text{gr } f$$

où $f := \frac{t}{s} - \tilde{c}(\cdot, (v, \frac{1}{s}))$. Donc le graphe de f est inclus dans l'hypographe stricte de h . Comme la fonction f est à valeurs réelles alors on en déduit que $(u, f(u))$ appartient à l'hypographe stricte de h pour tout $u \in U$ ce qui montre que $h > f$. Donc $h^{\Gamma_{\tilde{c}}} > f$ et en particulier

$$h^{\Gamma_{\tilde{c}}}(a) > f(a) = \frac{t}{s} - \tilde{c}(a, (v, \frac{1}{s})) = \bar{r} + \frac{1}{s}c(a, v) - \frac{1}{s}c(a, v) = \bar{r}.$$

D'où $(a, \bar{r}) \notin \text{epi } h^{\Gamma_{\tilde{c}}}$, et finalement

$$(\text{epi } h^{\Gamma_{\tilde{c}}}) \cap (\langle \text{dom } h \rangle_{c,3} \times \mathbb{R}) = \text{epi } h^{(\cdot)\tilde{c},3}.$$

\square

Corollaire 2.4.3. *Pour toute fonction $h : U \rightarrow \overline{\mathbb{R}}$, on a :*

$$h^{(\cdot)\tilde{c},3}(a) = h^{\tilde{c}\tilde{c}}(a), \quad \forall a \in \langle \text{dom } h \rangle_{c,3}.$$

Preuve. D'après le théorème 2.4.2, pour tout $a \in U$, on a :

$$h^{(\cdot)\bar{\varepsilon},3}(a) = h^{\Gamma\bar{\varepsilon}}(a) + i_{\langle \text{dom} \rangle_{c,3}}(a).$$

En particulier,

$$h^{(\cdot)\bar{\varepsilon},3}(a) = h^{\Gamma\bar{\varepsilon}}(a), \quad \forall a \in \langle \text{dom} \rangle_{c,3}.$$

□

Corollaire 2.4.4. Soit $i \in \{1, 2, 3\}$. Pour tout $A \subset U$ tel que $\text{dom } i_A^c \neq \emptyset$ on a :

$$i_A^{(\cdot)\bar{\varepsilon},i} = i_{\langle A \rangle_{c,i}}.$$

Preuve. D'après le théorème 2.2.2, on a $i_A^{\Gamma\bar{\varepsilon}} = i_{\langle A \rangle_{\bar{\varepsilon},1}}$ et $i_A^{\Gamma\bar{\varepsilon}} = i_{\langle A \rangle_{\bar{\varepsilon},1}}$. D'après le lemme 2.2.1, on a $\langle A \rangle_{\bar{\varepsilon},1} = \langle A \rangle_{c,1}$ pour tout $A \subset U$. Il résulte des théorèmes 2.4.1 et 2.4.2 que

$$\begin{aligned} i_A^{(\cdot)\bar{\varepsilon},1} &= i_A^{\Gamma\bar{\varepsilon}} + i_{\langle A \rangle_{c,1}} = i_{\langle A \rangle_{c,1}} + i_{\langle A \rangle_{c,1}} = i_{\langle A \rangle_{c,1}}, \\ i_A^{(\cdot)\bar{\varepsilon},2} &= i_A^{\Gamma\bar{\varepsilon}} + i_{\langle A \rangle_{c,2}} = i_{\langle A \rangle_{c,1}} + i_{\langle A \rangle_{c,2}} = i_{\langle A \rangle_{c,1} \cap \langle A \rangle_{c,2}} = i_{\langle A \rangle_{c,2}}, \\ i_A^{(\cdot)\bar{\varepsilon},3} &= i_A^{\Gamma\bar{\varepsilon}} + i_{\langle A \rangle_{c,3}} = i_{\langle A \rangle_{\bar{\varepsilon},1}} + i_{\langle A \rangle_{c,3}} = i_{\langle A \rangle_{\bar{\varepsilon},1} \cap \langle A \rangle_{c,3}}. \end{aligned}$$

Par ailleurs, d'après (2.12), $u \in \langle A \rangle_{c,3}$ si et seulement si pour tout $v \in V$, il existe $a \in A$ tel que $c(u, v) = c(a, v)$. Par conséquent,

$$\begin{aligned} u \in \langle A \rangle_{c,3} &\implies \inf_{a \in A} c(a, v) \leq c(u, v) \leq \sup_{a \in A} c(a, v) \\ &\implies u \in \langle A \rangle_{\bar{\varepsilon},1} \quad \text{d'après la proposition 2.2.11.} \end{aligned}$$

Donc $\langle A \rangle_{c,3} \subset \langle A \rangle_{\bar{\varepsilon},1}$, d'où $i_A^{(\cdot)\bar{\varepsilon},3} = i_{\langle A \rangle_{c,3}}$. □

Remarque 2.4.3. Dans le cas bilinéaire, il résulte du corollaire 2.4.4 que la fonction $i_{\mathbb{R}^*}$ est (c, P_3) -épi-régulière mais n'est pas quasi-convexe.

Théorème 2.4.3. Pour toute fonction $h : U \rightarrow \overline{\mathbb{R}}$, on a :

$$h^{(\cdot)\bar{\varepsilon},4} = \vartheta_{\langle \text{dom } h \rangle_{c,4}}.$$

Preuve. Si $\text{dom } h = \emptyset$, le résultat est évident. Supposons que $\text{dom } h \neq \emptyset$. Il est facile de voir que

$$\text{epi } h^{(\cdot)\bar{\varepsilon},4} \subset \langle \text{dom } h \rangle_{c,4} \times \mathbb{R}.$$

Réciproquement, soit $(a, \bar{r}) \notin \text{epi } h^{(\cdot)\bar{\varepsilon},4}$. Il existe $(v, s, t) \in V \times \mathbb{R} \times \mathbb{R}$ tel que

$$t - \bar{r}s - c(a, v) \neq 0 \quad \text{et} \quad \text{epi } h \subset \text{epi } h^{(\cdot)\bar{\varepsilon},4} \subset \{(u, r) \in U \times \mathbb{R} : t - rs - c(u, v) = 0\}.$$

On en déduit que $s = 0$ car $\text{dom } h \neq \emptyset$. Par conséquent,

$$t - c(a, v) \neq 0 \quad \text{et} \quad \text{epi } h \subset \{(u, r) \in U \times \mathbb{R} : t - c(u, v) = 0\},$$

et particulier $\text{dom } h \subset \{u \in U : t - c(u, v) = 0\}$. Donc $a \notin \langle \text{dom } h \rangle_{c,4}$. □

2.4.3. Dualité épigraphique

On montre dans cette sous-section que la (\check{c}, P_i) -épi-régularisée d'une fonction s'interprète comme une biconjugaison pour $i = 1, 2, 3$.

Soit le couplage $\xi_P : U \times (V \times \mathbb{R}) \rightarrow \overline{\mathbb{R}}$ défini par

$$(2.36) \quad \xi_P(u, (v, r)) := \vartheta_{E_{v,r}^P}(u) = \begin{cases} -\infty & \text{si } r - c(u, v) \in P \\ +\infty & \text{si } r - c(u, v) \notin P. \end{cases}$$

On introduit le couplage $\gamma_P : U \times (V \times \mathbb{R} \times \{0, 1\}) \rightarrow \overline{\mathbb{R}}$ défini par

$$(2.37) \quad \gamma_P(u, (v, r, j)) = \begin{cases} c(u, v) & \text{si } j = 0 \\ \xi_P(u, (v, r)) & \text{si } j = 1, \end{cases}$$

que l'on notera simplement γ_i si $P = P_i$ pour $i = 1, 2, 3, 4$. Par définition de γ_P , toute fonction γ_P -élémentaire est soit c -élémentaire, soit ξ_P -élémentaire.

Théorème 2.4.4. *Pour toute fonction $h : U \rightarrow \overline{\mathbb{R}}$, on a :*

$$(2.38) \quad h^{\gamma_P \gamma_P} = \max \{h^{\Gamma_c}, \vartheta_{\langle \text{dom } h \rangle_{c,P}}\} = h^{\Gamma_c} + i_{\langle \text{dom } h \rangle_{c,P}}.$$

Preuve. D'après le théorème de biconjugaison de Moreau (Remarque 2.2.1), pour toute fonction $h : U \rightarrow \overline{\mathbb{R}}$, on a

$$h^{\gamma_P \gamma_P} = h^{\Gamma_{\gamma_P}} = \max \{h^{\Gamma_c}, h^{\Gamma_{\xi_P}}\} = \max \{h^{\Gamma_c}, h^{\xi_P \xi_P}\}.$$

Par ailleurs, pour tout $(v, r) \in V \times \mathbb{R}$, on a

$$\begin{aligned} -h^{\xi_P}(v, r) &:= \inf_{u \in U} \{h(u) - \xi_P(u, (v, r))\} \\ &= \inf_{u \in U} \{h(u) - \vartheta_{E_{v,r}^P}(u)\} \\ &= \begin{cases} -\infty & \text{si } \text{dom } h \not\subset E_{v,r}^P \\ +\infty & \text{si } \text{dom } h \subset E_{v,r}^P. \end{cases} \end{aligned}$$

Donc

$$h^{\xi_P}(v, r) = \begin{cases} +\infty & \text{si } \text{dom } h \not\subset E_{v,r}^P \\ -\infty & \text{si } \text{dom } h \subset E_{v,r}^P, \end{cases}$$

et par conséquent, pour tout $u \in U$, on a

$$\begin{aligned} -h^{\xi_P \xi_P}(u) &= \inf_{(v,r) \in V \times \mathbb{R}} \{h^{\xi_P}(v, r) - \xi_P(u, (v, r))\} \\ &= \inf_{(v,r) \in V \times \mathbb{R}} \{h^{\xi_P}(v, r) - \vartheta_{E_{v,r}^P}(u)\} \\ &= \begin{cases} -\infty & \text{si } \exists (v, r) \in V \times \mathbb{R} : u \notin E_{v,r}^P \text{ et } \text{dom } h \subset E_{v,r}^P \\ +\infty & \text{sinon} \end{cases} \\ &= \begin{cases} -\infty & \text{si } u \notin \langle \text{dom } h \rangle_{c,P} \\ +\infty & \text{sinon} . \end{cases} \end{aligned}$$

D'où $h^{\xi_P \xi_P} = \vartheta_{\langle \text{dom } h \rangle_{c,P}}$. □

Théorème 2.4.5. *Soit U est un espace vectoriel topologique de dual topologique V et c la forme bilinéaire standard. Pour tout $h : U \rightarrow \overline{\mathbb{R}}$, on a*

$$(2.39) \quad h^{\langle \cdot \rangle_{\check{c}, i}} = h^{\gamma_i \gamma_i}, \quad \text{pour } i = 1, 2, 3.$$

Preuve. Rappelons que si U est un espace vectoriel topologique de dual topologique V et c la forme bilinéaire standard, on a $h^{\Gamma_c} = h^{\Gamma_\varepsilon} = h^{\Gamma_\varepsilon}$ pour tout $h : U \rightarrow \overline{\mathbb{R}}$. La conclusion résulte donc des théorèmes 2.4.1, 2.4.2 et 2.4.4. \square

2.5. De la dualité par niveaux à la dualité épigraphique

À toute fonction $h : U \rightarrow \overline{\mathbb{R}}$, on associe la fonction $H : U \times \mathbb{R} \rightarrow \overline{\mathbb{R}}$ définie par

$$H(u, t) = h(u) - t, \quad \forall (u, t) \in U \times \mathbb{R}.$$

Observons que pour tout $r \in \mathbb{R}$, on a :

$$(2.40) \quad [H \leq r] = \{(u, t) \in U \times \mathbb{R} \mid h(u) - t \leq r\} = \text{epi}(h - r).$$

On associe à \check{c} , le couplage $\check{c} : (U \times \mathbb{R} \times \mathbb{R}) \times (V \times \mathbb{R} \times \mathbb{R}) \rightarrow \mathbb{R}$ défini par

$$(2.41) \quad \check{c}((u, t, t'), (v, s, s')) := \check{c}((u, t), (v, s)) + t's' = c(u, v) + ts + t's',$$

et on note $\langle \cdot \rangle_{\check{c}, P}$ l'opérateur de fermeture associé.

Théorème 2.5.1. *Pour toute fonction $h : U \rightarrow \overline{\mathbb{R}}$, les assertions suivantes sont équivalentes :*

- (1) h est (\check{c}, P) -épi-régulière,
- (2) H est (\check{c}, P) -épi-régulière,
- (3) H est (\check{c}, P) -régulière par niveaux.

Preuve. (1) \implies (2). Observons que pour tout $(u, t, r) \in U \times \mathbb{R} \times \mathbb{R}$, on a :

$$(2.42) \quad (u, t, r) \in \text{epi } H \iff (u, r + t) \in \text{epi } h.$$

Soit $(\bar{u}, \bar{t}, \bar{r}) \notin \text{epi } H$. Alors $(\bar{u}, \bar{r} + \bar{t}) \notin \text{epi } h$ qui est (\check{c}, P) -régulière d'après (1). Il existe $(\bar{v}, \bar{s}, \bar{s}') \in V \times \mathbb{R} \times \mathbb{R}$ tel que

$$\bar{s}' - \bar{s}(\bar{r} + \bar{t}) - c(\bar{u}, \bar{v}) \notin P \quad \text{et} \quad \text{epi } h \subset \{(u, t) \in U \times \mathbb{R} : \bar{s}' - \bar{s}t - c(u, \bar{v}) \in P\}.$$

Par conséquent,

$$\bar{s}' - \check{c}((\bar{u}, \bar{t}, \bar{r}), (\bar{v}, \bar{s}, \bar{s})) \notin P \quad \text{et} \quad \text{epi } H \subset \{(u, t, r) \in U \times \mathbb{R} \times \mathbb{R} : \bar{s}' - \check{c}((u, t, r), (\bar{v}, \bar{s}, \bar{s})) \in P\},$$

et finalement, $(\bar{u}, \bar{t}, \bar{r}) \notin \langle \text{epi } H \rangle_{\check{c}, P}$.

(2) \implies (3) résulte immédiatement de la proposition 2.4.2.

(3) \implies (1). D'après (2.40), on a

$$\text{epi } h = [H \leq 0],$$

qui est un ensemble (\check{c}, P) -régulier d'après (3). \square

Le théorème 2.5.1 généralise la proposition 4.1 de [73] qui donne le résultat dans le cas evenly convexe.

2.5.1. Conjugaison associée à l'opérateur de fermeture $\langle \cdot \rangle_{\check{c}, P}$

En suivant le schéma de décomposition d'un opérateur de fermeture ([71]), on associe à l'opérateur de fermeture $\langle \cdot \rangle_{\check{c}, P}$, le couplage $\omega_P : (U \times \mathbb{R}) \times (V \times \mathbb{R} \times \mathbb{R}) \rightarrow \overline{\mathbb{R}}$ défini par :

$$(2.43) \quad \omega_P((u, t), (v, r, s)) = \begin{cases} 0 & \text{si } s - \check{c}((u, t), (v, r)) \notin P \\ -\infty & \text{si } s - \check{c}((u, t), (v, r)) \in P. \end{cases}$$

On notera simplement ω_i au lieu de ω_{P_i} pour $i \in \llbracket 1; 4 \rrbracket$. D'après le théorème 2.3.1, pour toute fonction $k : U \times \mathbb{R} \rightarrow \overline{\mathbb{R}}$, la régularisée par niveaux de k par rapport à l'opérateur $\langle \cdot \rangle_{\check{c}, P}$ coïncide avec la biconjuguée de k par rapport au couplage ω_P . Le théorème suivant exprime la (\check{c}, P) -régularisée par niveaux de H en fonction de la (\check{c}, P) -épi-régularisée de h pour $i = 1, 2, 3, 4$.

Théorème 2.5.2. *Pour toute fonction $h : U \rightarrow \overline{\mathbb{R}}$, pour tout $(u, t) \in U \times \mathbb{R}$, on a :*

$$(2.44) \quad H^{\omega_1 \omega_1}(u, t) = h^{\Gamma_{\check{c}}}(u) - t + i_{\langle \text{dom } h \rangle_{c,1}}(u) = h^{\langle \cdot \rangle_{\check{c},1}}(u) - t,$$

$$(2.45) \quad H^{\omega_2 \omega_2}(u, t) = h^{\Gamma_{\check{c}}}(u) - t + i_{\langle \text{dom } h \rangle_{c,2}}(u) = h^{\langle \cdot \rangle_{\check{c},2}}(u) - t,$$

$$(2.46) \quad H^{\omega_3 \omega_3}(u, t) = h^{\Gamma_{\check{c}}}(u) - t + i_{\langle \text{dom } h \rangle_{c,3}}(u) = h^{\langle \cdot \rangle_{\check{c},3}}(u) - t,$$

$$(2.47) \quad H^{\omega_4 \omega_4}(u, t) = \vartheta_{\langle \text{dom } h \rangle_{c,4}}(u) = h^{\langle \cdot \rangle_{\check{c},4}}(u).$$

Preuve. On procède par un calcul explicite de $H^{\omega_i \omega_i}$ pour $i \in \llbracket 1; 4 \rrbracket$. Le résultat est immédiat si $\text{dom } h = \emptyset$. On suppose donc que $\text{dom } h \neq \emptyset$.

Cas de $i = 1$. Soit $(v, r, s) \in V \times \mathbb{R} \times \mathbb{R}$. Par définition, on a :

$$\begin{aligned} -H^{\omega_1}(v, r, s) &= \inf_{\substack{u \in U \\ t \in \mathbb{R}}} \{h(u) - t \mid s - tr - c(u, v) < 0\} \\ &= \inf_{u \in U} \left\{ h(u) - \sup_{t \in \mathbb{R}} \{t \mid tr > s - c(u, v)\} \right\}. \end{aligned}$$

Par ailleurs,

$$\sup_{t \in \mathbb{R}} \{t \mid tr > s - c(u, v)\} = \begin{cases} \frac{s - c(u, v)}{r} & \text{si } r < 0 \\ +\infty & \text{si } r > 0 \\ +\infty & \text{si } r = 0 \text{ et } c(u, v) > s \\ -\infty & \text{si } r = 0 \text{ et } c(u, v) \leq s. \end{cases}$$

On en déduit que :

- Si $r < 0$ alors $-H^{\omega_1}(v, r, s) = \inf_{u \in U} \left\{ h(u) - \frac{s}{r} + \frac{1}{r}c(u, v) \right\} = -\frac{s}{r} - h^{\check{c}}\left(v, -\frac{1}{r}\right)$.
- Si $r > 0$ alors $-H^{\omega_1}(v, r, s) = \inf_{u \in U} \{h(u) - \infty\} = -\infty$ car $\text{dom } h \neq \emptyset$.

$$\bullet -H^{\omega_1}(v, 0, s) = \begin{cases} +\infty & \text{si } \text{dom } h \subset [c(\cdot, v) \leq s] \\ -\infty & \text{si } \text{dom } h \not\subset [c(\cdot, v) \leq s]. \end{cases}$$

Par conséquent,

$$H^{\omega_1}(v, r, s) = \begin{cases} \frac{s}{r} + h^{\widehat{c}}(v, -\frac{1}{r}) & \text{si } r < 0 \\ +\infty & \text{si } r > 0 \\ -\infty & \text{si } r = 0 \text{ et } \text{dom } h \subset [c(\cdot, v) \leq s] \\ +\infty & \text{si } r = 0 \text{ et } \text{dom } h \not\subset [c(\cdot, v) \leq s]. \end{cases}$$

Ainsi, pour tout $(u, t) \in U \times \mathbb{R}$, on a :

$$\begin{aligned} -H^{\omega_1 \omega_1}(u, t) &= \inf_{\substack{v \in V \\ (r, s) \in \mathbb{R}^2}} \{H^{\omega_1}(v, r, s) \mid c(u, v) + tr > s\} \\ &= \min(k(u, t), g(u)), \end{aligned}$$

où

$$\begin{aligned} k(u, t) &:= \inf_{\substack{v \in V \\ r \in \mathbb{R}_+^* \\ s \in \mathbb{R}}} \left\{ \frac{s}{r} + h^{\widehat{c}}\left(v, -\frac{1}{r}\right) \mid c(u, v) + tr > s \right\} \\ &= \inf_{\substack{v \in V \\ r' \in \mathbb{R}_-^* \\ s' \in \mathbb{R}}} \left\{ s' + h^{\widehat{c}}(v, -r') \mid r'c(u, v) + t < s' \right\} \\ &= \inf_{\substack{v \in V \\ r' \in \mathbb{R}_-^*}} \left\{ h^{\widehat{c}}(v, -r') + r'c(u, v) + t \right\} \\ &= \inf_{\substack{v \in V \\ s \in \mathbb{R}_+^*}} \left\{ h^{\widehat{c}}(v, r) - rc(u, v) \right\} + t \\ &= t - h^{\widehat{c\widehat{c}}}(u), \end{aligned}$$

et

$$\begin{aligned} g(u) &:= \inf_{\substack{v \in V \\ s \in \mathbb{R}}} \{H^{\omega_1}(v, 0, s) \mid c(u, v) > s\} \\ &= \begin{cases} -\infty & \text{si } \exists (v, s) \in V \times \mathbb{R} : u \notin E_{v, s}^{P_1} \text{ et } \text{dom } h \subset E_{v, s}^{P_1} \\ +\infty & \text{sinon} \end{cases} \\ &= \begin{cases} -\infty & \text{si } u \notin \langle \text{dom } h \rangle_{c, 1} \\ +\infty & \text{si } u \in \langle \text{dom } h \rangle_{c, 1} \end{cases} \\ &= -\vartheta_{\langle \text{dom } h \rangle_{c, 1}}(u). \end{aligned}$$

Il en résulte que

$$\begin{aligned} H^{\omega_1 \omega_1}(u, t) &= \max \left\{ h^{\widehat{c\widehat{c}}}(u) - t, \vartheta_{\langle \text{dom } h \rangle_{c, 1}}(u) \right\} \\ &= h^{\Gamma \widehat{c}}(u) - t + i_{\langle \text{dom } h \rangle_{c, 1}}(u) \\ &= h^{\langle \rangle \widehat{c}, 1}(u) - t \text{ d'après le théorème 2.4.1.} \end{aligned}$$

Cas de $i = 2$. Soit $(v, r, s) \in V \times \mathbb{R} \times \mathbb{R}$. Par définition, on a :

$$\begin{aligned} -H^{\omega_2}(v, r, s) &= \inf_{\substack{u \in U \\ t \in \mathbb{R}}} \{h(u) - t \mid s - tr - c(u, v) \leq 0\} \\ &= \inf_{u \in U} \left\{ h(u) - \sup_{t \in \mathbb{R}} \{t \mid tr \geq s - c(u, v)\} \right\}. \end{aligned}$$

Par ailleurs,

$$\sup_{t \in \mathbb{R}} \{t \mid tr \geq s - c(u, v)\} = \begin{cases} \frac{s - c(u, v)}{r} & \text{si } r < 0 \\ +\infty & \text{si } r > 0 \\ +\infty & \text{si } r = 0 \text{ et } c(u, v) \geq s \\ -\infty & \text{si } r = 0 \text{ et } c(u, v) < s. \end{cases}$$

On en déduit que :

- Si $r < 0$ alors $-H^{\omega_2}(v, r, s) = \inf_{u \in U} \left\{ h(u) - \frac{s}{r} + \frac{1}{r}c(u, v) \right\} = -\frac{s}{r} - h^{\widehat{c}}\left(v, -\frac{1}{r}\right)$.
- Si $r > 0$ alors $-H^{\omega_2}(v, r, s) = \inf_{u \in U} \{h(u) - \infty\} = -\infty$ car $\text{dom } h \neq \emptyset$.
- $-H^{\omega_2}(v, 0, s) = \begin{cases} +\infty & \text{si } \text{dom } h \subset [c(\cdot, v) < s] \\ -\infty & \text{si } \text{dom } h \not\subset [c(\cdot, v) < s]. \end{cases}$

Par conséquent,

$$H^{\omega_2}(v, r, s) = \begin{cases} \frac{s}{r} + h^{\widehat{c}}\left(v, -\frac{1}{r}\right) & \text{si } r < 0 \\ +\infty & \text{si } r > 0 \\ -\infty & \text{si } r = 0 \text{ et } \text{dom } h \subset [c(\cdot, v) < s] \\ +\infty & \text{si } r = 0 \text{ et } \text{dom } h \not\subset [c(\cdot, v) < s]. \end{cases}$$

Ainsi, pour tout $(u, t) \in U \times \mathbb{R}$, on a :

$$\begin{aligned} -H^{\omega_2 \omega_2}(u, t) &= \inf_{\substack{v \in V \\ (r, s) \in \mathbb{R}^2}} \{H^{\omega_1}(v, r, s) \mid c(u, v) + tr \geq s\} \\ &= \min(k(u, t), g(u)) \end{aligned}$$

où

$$\begin{aligned} k(u, t) &:= \inf_{\substack{v \in V \\ r \in \mathbb{R}_+^* \\ s \in \mathbb{R}}} \left\{ \frac{s}{r} + h^{\widehat{c}}\left(v, -\frac{1}{r}\right) \mid c(u, v) + tr \geq s \right\} \\ &= \inf_{\substack{v \in V \\ r' \in \mathbb{R}_+^* \\ s' \in \mathbb{R}}} \{s' + h^{\widehat{c}}(v, -r') \mid r'c(u, v) + t \leq s'\} \\ &= \inf_{\substack{v \in V \\ r' \in \mathbb{R}_+^*}} \{h^{\widehat{c}}(v, -r') + r'c(u, v) + t\} \\ &= \inf_{\substack{v \in V \\ s \in \mathbb{R}_+^*}} \{h^{\widehat{c}}(v, r) - rc(u, v)\} + t \\ &= t - h^{\widehat{c\widehat{c}}}(u), \end{aligned}$$

et

$$\begin{aligned}
g(u) &:= \inf_{\substack{v \in V \\ s \in \mathbb{R}}} \{H^{\omega_2}(v, 0, s) \mid c(u, v) \geq s\} \\
&= \begin{cases} -\infty & \text{si } \exists (v, s) \in V \times \mathbb{R} : u \notin E_{v,s}^{P_2} \text{ et } \text{dom } h \subset E_{v,s}^{P_2} \\ +\infty & \text{sinon} \end{cases} \\
&= \begin{cases} -\infty & \text{si } u \notin \langle \text{dom } h \rangle_{c,2} \\ +\infty & \text{si } u \in \langle \text{dom } h \rangle_{c,2} \end{cases} \\
&= -\vartheta_{\langle \text{dom } h \rangle_{c,2}}(u).
\end{aligned}$$

Il en résulte que

$$\begin{aligned}
H^{\omega_2 \omega_2}(u, t) &= \max \{h^{\widehat{c}}(u) - t, \vartheta_{\langle \text{dom } h \rangle_{c,2}}(u)\} \\
&= h^{\Gamma \widehat{c}}(u) - t + i_{\langle \text{dom } h \rangle_{c,2}}(u) \\
&= h^{\langle \rangle \widehat{c}, 2}(u) - t \quad \text{d'après le théorème 2.4.1.}
\end{aligned}$$

Cas de $i = 3$. Soit $(v, r, s) \in V \times \mathbb{R} \times \mathbb{R}$. Par définition, on a :

$$\begin{aligned}
-H^{\omega_3}(v, r, s) &= \inf_{\substack{u \in U \\ t \in \mathbb{R}}} \{h(u) - t \mid s - tr - c(u, v) = 0\} \\
&= \inf_{u \in U} \left\{ h(u) - \sup_{t \in \mathbb{R}} \{t \mid tr = s - c(u, v)\} \right\}.
\end{aligned}$$

Par ailleurs,

$$\sup_{t \in \mathbb{R}} \{t \mid tr = s - c(u, v)\} = \begin{cases} \frac{s - c(u, v)}{r} & \text{si } r \neq 0 \\ +\infty & \text{si } r = 0 \text{ et } c(u, v) = s \\ -\infty & \text{si } r = 0 \text{ et } c(u, v) \neq s. \end{cases}$$

On en déduit que :

- Si $r \neq 0$ alors $-H^{\omega_3}(v, r, s) = \inf_{u \in U} \{h(u) - \frac{s}{r} + \frac{1}{r}c(u, v)\} = -\frac{s}{r} - h^{\widehat{c}}(v, -\frac{1}{r})$.
- $-H^{\omega_3}(v, 0, s) = \begin{cases} +\infty & \text{si } \text{dom } h \subset [c(\cdot, v) \neq s] \\ -\infty & \text{si } \text{dom } h \not\subset [c(\cdot, v) \neq s]. \end{cases}$

Par conséquent,

$$H^{\omega_3}(v, r, s) = \begin{cases} \frac{s}{r} + h^{\widehat{c}}(v, -\frac{1}{r}) & \text{si } r \neq 0 \\ -\infty & \text{si } r = 0 \text{ et } \text{dom } h \subset [c(\cdot, v) \neq s] \\ +\infty & \text{si } r = 0 \text{ et } \text{dom } h \not\subset [c(\cdot, v) \neq s]. \end{cases}$$

Ainsi, pour tout $(u, t) \in U \times \mathbb{R}$, on a :

$$\begin{aligned}
-H^{\omega_3 \omega_3}(u, t) &= \inf_{\substack{v \in V \\ (r, s) \in \mathbb{R}^2}} \{H^{\omega_3}(v, r, s) \mid \check{c}((u, t), (v, r)) = s\} \\
&= \min(k(u, t), g(u)),
\end{aligned}$$

où

$$\begin{aligned}
k(u, t) &:= \inf_{\substack{v \in V \\ r \in \mathbb{R}^* \\ s \in \mathbb{R}}} \left\{ \frac{s}{r} + h^{\tilde{c}}\left(v, -\frac{1}{r}\right) \mid c(u, v) + tr = s \right\} \\
&= \inf_{\substack{v \in V \\ r' \in \mathbb{R}^* \\ s' \in \mathbb{R}}} \left\{ s' + h^{\tilde{c}}(v, -r') \mid r'c(u, v) + t = s' \right\} \\
&= \inf_{\substack{v \in V \\ r' \in \mathbb{R}^*}} \left\{ h^{\tilde{c}}(v, -r') + r'c(u, v) + t \right\} \\
&= \inf_{\substack{v \in V \\ s \in \mathbb{R}^*}} \left\{ h^{\tilde{c}}(v, r) - rc(u, v) \right\} + t \\
&= t - h^{\tilde{c}\tilde{c}}(u),
\end{aligned}$$

et

$$\begin{aligned}
g(u) &:= \inf_{\substack{v \in V \\ s \in \mathbb{R}}} \{ H^{\omega_3}(v, 0, s) \mid c(u, v) = s \} \\
&= \begin{cases} -\infty & \text{si } \exists (v, s) \in V \times \mathbb{R} : u \notin E_{v,s}^{P_3} \text{ et } \text{dom } h \subset E_{v,s}^{P_3} \\ +\infty & \text{sinon} \end{cases} \\
&= \begin{cases} -\infty & \text{si } u \notin \langle \text{dom } h \rangle_{c,3} \\ +\infty & \text{si } u \in \langle \text{dom } h \rangle_{c,3} \end{cases} \\
&= -\vartheta_{\langle \text{dom } h \rangle_{c,3}}(u).
\end{aligned}$$

Il en résulte que

$$\begin{aligned}
H^{\omega_3\omega_3}(u, t) &= \max \{ h^{\tilde{c}\tilde{c}}(u) - t, \vartheta_{\langle \text{dom } h \rangle_{c,3}}(u) \} \\
&= h^{\Gamma\tilde{c}}(u) - t + i_{\langle \text{dom } h \rangle_{c,3}}(u) \\
&= h^{(\cdot)\tilde{c},3}(u) - t \quad \text{d'après le théorème 2.4.2.}
\end{aligned}$$

Cas de $i = 4$. Soit $(v, r, s) \in V \times \mathbb{R} \times \mathbb{R}$. Par définition, on a :

$$\begin{aligned}
-H^{\omega_4}(v, r, s) &= \inf_{\substack{u \in U \\ t \in \mathbb{R}}} \{ h(u) - t \mid s - tr - c(u, v) \neq 0 \} \\
&= \inf_{u \in U} \left\{ h(u) - \sup_{t \in \mathbb{R}} \{ t \mid tr \neq s - c(u, v) \} \right\}.
\end{aligned}$$

Par ailleurs,

$$\sup_{t \in \mathbb{R}} \{ t \mid tr \neq s - c(u, v) \} = \begin{cases} +\infty & \text{si } r \neq 0 \\ -\infty & \text{si } r = 0 \text{ et } c(u, v) = s \\ +\infty & \text{si } r = 0 \text{ et } c(u, v) \neq s. \end{cases}$$

Donc

$$-H^{\omega_4}(v, r, s) = \begin{cases} -\infty & \text{si } r \neq 0 \\ +\infty & \text{si } r = 0 \text{ et } \text{dom } h \subset [c(\cdot, v) = s] \\ -\infty & \text{si } r = 0 \text{ et } \text{dom } h \not\subset [c(\cdot, v) = s], \end{cases}$$

$$H^{\omega_4}(v, r, s) = \begin{cases} +\infty & \text{si } r \neq 0 \\ -\infty & \text{si } r = 0 \text{ et } \text{dom } h \subset [c(\cdot, v) = s] \\ +\infty & \text{si } r = 0 \text{ et } \text{dom } h \not\subset [c(\cdot, v) = s]. \end{cases}$$

Ainsi, pour tout $(u, t) \in U \times \mathbb{R}$, on a :

$$\begin{aligned} -H^{\omega_4\omega_4}(u, t) &= \inf_{\substack{v \in V \\ s \in \mathbb{R}}} \{H^{\omega_4}(v, 0, s) \mid s \neq c(u, v)\} \\ &= \begin{cases} -\infty & \text{si } \exists (v, s) \in V \times \mathbb{R} : \text{dom } h \subset [c(\cdot, v) = s] \text{ et } c(u, v) \neq s \\ +\infty & \text{sinon} \end{cases} \\ &= \begin{cases} -\infty & \text{si } u \notin \langle \text{dom } h \rangle_{c,4} \\ +\infty & \text{sinon} \end{cases} \\ &= -\vartheta_{\langle \text{dom } h \rangle_{c,4}}. \end{aligned}$$

Il en résulte que

$$H^{\omega_4\omega_4}(u, t) = \vartheta_{\langle \text{dom } h \rangle_{c,4}} = h^{(\cdot)\varepsilon,4}(u) \quad \text{d'après le théorème 2.4.3.}$$

□

2.5.2. Autres expressions de $h^{(\cdot)\varepsilon,i}$, pour $i = 1, 2, 3, 4$

Proposition 2.5.1. Soit $i \in \{1, 2, 3, 4\}$. Pour tout $h : U \rightarrow \overline{\mathbb{R}}$ et pour tout $a \in U$, on a

$$h^{(\cdot)\varepsilon,i}(a) = \sup\{t \in \mathbb{R} \mid (a, t) \notin \langle \text{epi } h \rangle_{\varepsilon,i}\}.$$

Preuve. Soit $\bar{t} < h^{(\cdot)\varepsilon,i}(a)$. Alors $(a, \bar{t}) \notin \langle \text{epi } h \rangle_{\varepsilon,i}$ d'après la proposition 2.4.4 . Par conséquent

$$\bar{t} \leq \sup\{t \in \mathbb{R} \mid (a, t) \notin \langle \text{epi } h \rangle_{\varepsilon,i}\},$$

donc

$$h^{(\cdot)\varepsilon,i}(a) \leq \sup\{t \in \mathbb{R} \mid (a, t) \notin \langle \text{epi } h \rangle_{\varepsilon,i}\}.$$

Par ailleurs, d'après le théorème 2.5.2, on a

$$(2.48) \quad H^{\omega_i\omega_i}(u, t) = h^{(\cdot)\varepsilon,i}(u) - t, \quad \forall (a, \bar{t}) \in U \times \mathbb{R}.$$

Soit $\bar{t} \in \{t \in \mathbb{R} \mid (a, t) \notin \langle \text{epi } h \rangle_{\varepsilon,i}\}$. Alors $(a, \bar{t}) \notin \langle \text{epi } h \rangle_{\varepsilon,i}$. Donc, d'après la proposition 2.2.2 , il existe $(v, s, r) \in V \times \mathbb{R} \times \mathbb{R}$ tel que

$$(a, \bar{t}) \notin \ddot{E}_{v,s,t}^{P_i} \quad \text{et} \quad \text{epi } h \subset \ddot{E}_{v,s,t}^{P_i}.$$

Par conséquent,

$$\sup_{(a,\bar{t}) \notin \ddot{E}_{v,s,t}^{P_i}} \inf_{(u,t) \notin \ddot{E}_{v,s,t}^{P_i}} h(u) - t \geq 0.$$

Or

$$\begin{aligned} \sup_{(a,\bar{t}) \notin \ddot{E}_{v,s,t}^{P_i}} \inf_{(u,t) \notin \ddot{E}_{v,s,t}^{P_i}} h(u) - t \geq 0 &\implies H^{\omega_i \omega_i}(a, \bar{t}) \geq 0 \\ &\implies h^{(\cdot)\ddot{e},i}(a) \geq \bar{t}, \text{ d'après (2.48)}. \end{aligned}$$

D'où $h^{(\cdot)\ddot{e},i}(a) \geq \sup\{t \in \mathbb{R} \mid (a, t) \notin \langle \text{epi } h \rangle_{\ddot{e},i}\}$. \square

Remarque 2.5.1. *On peut vérifier sans peine que pour tout $i \in \{1, 2, 3, 4\}$, pour toute fonction $h : U \rightarrow \overline{\mathbb{R}}$, on a*

$$(2.49) \quad h^{(\cdot)\ddot{e},i}(a) = \inf\{t \in \mathbb{R} \mid (a, t) \in \langle \text{epi } h \rangle_{\ddot{e},i}\}, \quad \forall a \in U.$$

2.6. Application à un problème d'optimisation : dualité par niveaux

Considérons le problème de minimisation suivant :

$$(\mathcal{P}) \quad \text{Minimiser } f(x), \quad x \in X,$$

où X est un ensemble non vide et $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$ est une fonction numérique.

2.6.1. Dualité perturbative par niveaux

On se donne une fonction $F : X \times U \rightarrow \overline{\mathbb{R}}$ vérifiant

$$(2.50) \quad \exists a \in U : F(\cdot, a) = f(\cdot).$$

On associe à la fonction F la fonction valeur $h : U \rightarrow \overline{\mathbb{R}}$ définie par

$$(2.51) \quad h(u) = \inf_{x \in X} F(x, u).$$

On note α la valeur optimale du problème (\mathcal{P}) , c'est-à-dire $\alpha = \inf_X f$, on a

$$(2.52) \quad h(a) = \alpha.$$

Considérons le couplage $c : U \times V \rightarrow \mathbb{R}$, l'ensemble $P \subset \mathbb{R}$ et la polarité $\Delta_{c,P}$ correspondant que l'on notera simplement Δ . Le problème dual (\mathcal{D}) du problème (\mathcal{P}) associé à la perturbation F et à la polarité Δ est donné par ([71]) :

$$(\mathcal{D}) \quad \text{Maximiser } -h^\Delta(v, r), \quad a \notin E_{v,r}^P.$$

Notons β , la valeur optimale du dual. On a par définition

$$(2.53) \quad -\infty \leq \beta := \sup(\mathcal{D}) = h^{\Delta\Delta^*}(a) \leq h(a) =: \alpha = \inf(\mathcal{P}) \leq +\infty.$$

Ainsi, la dualité faible est vérifiée, c'est-à-dire que la valeur optimale du dual est inférieure ou égale à celle du primal. Le théorème suivant donne une condition nécessaire et suffisante qui garantit la dualité forte c'est-dire l'égalité entre α et β et le dual admet au moins une solution.

Théorème 2.6.1. *Les assertions suivantes sont équivalentes :*

$$(1) \quad \inf(\mathcal{P}) = \max(\mathcal{D}),$$

(2) $a \notin \langle [h < \alpha] \rangle_{c,P}$.

Preuve. (1) \implies (2). D'après (1), il existe $(\bar{v}, \bar{r}) \in V \times \mathbb{R}$ tel que

$$\bar{r} - c(a, \bar{v}) \notin P \quad \text{et} \quad \alpha = h(a) = -h^\Delta(\bar{v}, \bar{r}) := \inf_{u \in E_{\bar{v}, \bar{r}}^P} h(u).$$

On en déduit que $[h < \alpha] \subset E_{\bar{v}, \bar{r}}^P$. D'où $a \notin \langle [h < \alpha] \rangle_{c,P}$ d'après la proposition 2.2.2.

(2) \implies (1). D'après (1) et la proposition 2.2.2, il existe $(\bar{v}, \bar{r}) \in V \times \mathbb{R}$ tel que

$$a \notin E_{\bar{v}, \bar{r}}^P \quad \text{et} \quad [h < \alpha] \subset E_{\bar{v}, \bar{r}}^P.$$

Il en résulte que

$$\inf_{u \in E_{\bar{v}, \bar{r}}^P} h(u) \geq \alpha \geq \beta.$$

Or

$$\beta := \sup_{a \in E_{\bar{v}, \bar{r}}^P} -h^\Delta(v, r) = \sup_{a \in E_{\bar{v}, \bar{r}}^P} \inf_{u \in E_{\bar{v}, \bar{r}}^P} h(u).$$

Donc

$$\beta \geq \inf_{u \in E_{\bar{v}, \bar{r}}^P} h(u) = -h^\Delta(\bar{v}, \bar{r}) \geq \alpha \geq \beta.$$

D'où $\beta = -h^\Delta(\bar{v}, \bar{r}) = \alpha$. □

2.6.2. Application à la dualité evenly quasi-convexe([23],[51],[58],[59],[72])

On suppose que X et U sont deux espaces vectoriels topologiques, $V = U^*$ le dual topologique de U , $c = \langle, \rangle$ le couplage bilinéaire standard entre U et U^* .

Corollaire 2.6.1. *On suppose que la fonction $F : X \times U \rightarrow \overline{\mathbb{R}}$ est quasi-convexe et pour tout $x \in X$, $F(x, \cdot) : U \rightarrow \overline{\mathbb{R}}$ est semi-continue supérieurement. On a*

$$\inf(\mathcal{P}) = \max_{u^* \in U^*} \inf_{\substack{(x,u) \in X \times U \\ \langle u-a, u^* \rangle \geq 0}} F(x, u).$$

Preuve. La fonction h est quasi-convexe et semi-continue supérieurement puisque F est quasi-convexe et pour tout $x \in X$, $F(x, \cdot)$ est semi-continue supérieurement. Par conséquent $[h < \alpha]$ est un convexe ouvert donc evenly convexe. Comme $a \notin [h < \alpha]$, il résulte du théorème 2.6.1 (avec $P = \mathbb{R}_+^*$) que

$$\inf(\mathcal{P}) = \max(\mathcal{D}) = \max_{r - \langle a, u^* \rangle \leq 0} \inf_{r - \langle u, u^* \rangle \leq 0} h(u) = \max_{u^* \in U^*} \inf_{\langle u, u^* \rangle \geq \langle a, u^* \rangle} h(u),$$

où la dernière égalité résulte du fait que pour tout $u^* \in U^*$, la fonction $k_{u^*} : \mathbb{R} \rightarrow \overline{\mathbb{R}}$ définie par

$$k_{u^*}(r) = \inf_{r - \langle u, u^* \rangle \leq 0} h(u)$$

est croissante. □

Corollaire 2.6.2. *On suppose que la fonction $F : X \times U \rightarrow \overline{\mathbb{R}}$ est quasi-convexe et pour tout $x \in X$, $F(x, \cdot) : U \rightarrow \overline{\mathbb{R}}$ est semi-continue supérieurement. On a*

$$\inf(\mathcal{P}) = \max_{u^* \in U^*} \inf_{\substack{(x,u) \in X \times U \\ \langle u-a, u^* \rangle = 0}} F(x, u)$$

Preuve. Avec ses hypothèses, on sait que $[h < \alpha]$ est un convexe ouvert donc $(\langle \cdot, \mathbb{R}^* \rangle)$ -régulier. Puisque $a \notin [h < \alpha]$, il résulte du théorème 2.6.1 que

$$\begin{aligned} \inf(\mathcal{P}) &= \max(\mathcal{D}) \\ &= \max_{\substack{(u^*, r) \in U^* \times \mathbb{R} \\ \langle a, u^* \rangle = r}} \inf_{\substack{u \in U \\ \langle u, u^* \rangle = r}} h(u) \\ &= \max_{u^* \in U^*} \max_{\substack{r \in \mathbb{R} \\ \langle a, u^* \rangle = r}} \inf_{\substack{u \in U \\ \langle u, u^* \rangle = r}} h(u) \\ &= \max_{u^* \in U^*} \inf_{\substack{u \in U \\ \langle u-a, u^* \rangle = 0}} h(u) \\ &= \max_{u^* \in U^*} \inf_{\substack{(x,u) \in X \times U \\ \langle u-a, u^* \rangle = 0}} F(x, u) \quad \text{par définition de } h. \end{aligned}$$

□

Dualité en optimisation convexe conique à données incertaines

Notons que ce chapitre fait l'objet d'un papier accepté pour publication dans *pacific journal of optimization* ([3]).

3.1. Introduction

Considérons le problème de programmation convexe conique incertain suivant :

$$(P) \quad \inf_x f(x) \quad \text{s.l.c.} \quad g_u(x) \in -S,$$

où u désigne le paramètre incertain appartenant à U l'ensemble des incertitudes, X et Y sont deux espaces vectoriels topologiques localement convexes et séparés, $f : X \rightarrow \overline{\mathbb{R}}$ une fonction convexe propre semi-continue inférieurement, $S \subset Y$, un cône convexe fermé non vide, et pour tout $u \in U$, la fonction $g_u : \text{dom } g_u \subset X \rightarrow Y$ est S -convexe fermée par niveaux ou par épigraphe.

Au problème (P) est associée sa contrepartie robuste ([7],[8],[10],...) définie par :

$$(RP) \quad \inf_x f(x) \quad \text{s.l.c.} \quad g_u(x) \in -S, \forall u \in U.$$

On appelle **valeur robuste**, la valeur du problème (RP) c'est-à-dire $\inf(RP)$.

Étant donnée $u \in U$, (P_u) désigne l'occurrence correspondante du problème (P) :

$$(P_u) \quad \inf_x f(x) \quad \text{s.l.c.} \quad g_u(x) \in -S.$$

On considère le problème qui consiste à maximiser sur U , la fonction valeur des problèmes (P_u) :

$$(Q) \quad \sup_u \inf_x \{f(x) \quad \text{s.l.c.} \quad g_u(x) \in -S\} \quad \text{s.l.c.} \quad u \in U,$$

et on appelle **pire valeur** du problème incertain (P) , la valeur du problème (Q) , c'est-à-dire $\sup(Q)$.

Nous allons montrer que $\sup(Q) \leq \inf(RP)$ puis donner un exemple dans lequel l'inégalité est stricte (Proposition 3.3.1 et Exemple 3.3.1).

L'objectif de ce travail est d'établir une condition nécessaire et suffisante afin d'obtenir l'égalité entre la valeur robuste et la pire valeur du problème incertain (P) avec exactitude de la pire valeur. Le chapitre est organisé comme suit. La section 3.2 présente quelques éléments d'analyse convexe que nous allons utiliser par la suite. Dans

la section 3.3, on formule une condition nécessaire et suffisante pour obtenir l'égalité entre la valeur robuste et la pire valeur avec exactitude de la pire valeur (Théorème 3.3.1, Corollaire 3.3.1, Corollaire 3.3.2, Corollaire 3.3.3). Dans la section 3.4 on s'intéresse au dual optimiste (ODP) de (P). On montre que $\sup(ODP) \leq \sup(Q)$ (Proposition 3.4.1) et on donne une condition permettant d'obtenir $\sup(ODP) = \sup(Q)$ (Proposition 3.4.3). Dans le cas où la dualité robuste forte est vérifiée, on obtient $\inf(RP) = \max(Q)$ (Proposition 3.4.2). Inversement, on établit la dualité robuste forte (Théorème 3.4.1) à partir de nos résultats précédents en utilisant un principe de dualité en optimisation convexe. Finalement, on compare notre corollaire 3.4.1 au corollaire 3.1 de [49] (Remarque 3.4.4 et Proposition 3.4.4).

3.2. Rappels d'analyse convexe

Dans cette section, on rappelle quelques éléments d'analyse convexe qui nous seront utiles dans la suite de ce chapitre. Soient X un \mathbb{R} -espace vectoriel topologique localement convexe et séparé de dual topologique X^* et $\langle \cdot, \cdot \rangle$ la forme bilinéaire standard de dualité entre X et X^* . Étant donné une fonction $h : X \rightarrow \overline{\mathbb{R}}$, on note par $\text{dom } h := \{x \in X \mid h(x) < +\infty\}$ le domaine effectif de h . On dira que h est propre si $\text{dom } h \neq \emptyset$ et $-\infty \notin h(X)$. L'épigraphe de h est l'ensemble $\text{epi } h := \{(x, t) \in X \times \mathbb{R} \mid h(x) \leq t\}$. Rappelons que h est convexe si et seulement si $\text{epi } h$ est convexe, h est semi-continue inférieurement si et seulement si $\text{epi } h$ est fermé. L'ensemble de toutes les fonctions convexes propres semi-continues inférieurement à valeurs réelles étendues définies sur X est noté $\Gamma(X)$.

La conjuguée de Legendre-Fenchel de $h : X \rightarrow \overline{\mathbb{R}}$ est la fonction $h^* : X^* \rightarrow \overline{\mathbb{R}}$ définie par

$$h^*(x^*) := \sup_{x \in X} \{\langle x, x^* \rangle - h(x)\},$$

qui est convexe et semi-continue inférieurement pour la topologie faible-*. La bi-conjuguée de h est définie sur X à valeurs dans $\overline{\mathbb{R}}$ par

$$h^{**}(x) := \sup_{x^* \in X^*} \{\langle x, x^* \rangle - h^*(x^*)\}.$$

Il est clair que $h \geq h^{**}$, et que h^{**} est une fonction convexe semi-continue inférieurement. Si h est convexe propre semi-continue inférieurement alors $h = h^{**}$.

Propriété 3.2.1. *Pour toute famille de fonctions $(h_i)_{i \in I}$, définies de X à valeurs dans $\overline{\mathbb{R}}$, on a*

$$\left(\inf_{i \in I} h_i \right)^* = \sup_{i \in I} h_i^*.$$

Étant donné un sous-ensemble A de X , on note par $i_A : X \rightarrow \overline{\mathbb{R}}$ la fonction indicatrice de A définie par :

$$i_A(x) = \begin{cases} 0 & \text{si } x \in A \\ +\infty & \text{si } x \in X \setminus A, \end{cases}$$

$\sigma_A = i_A^*$ la fonction support de A , $\text{conv}A$ l'enveloppe convexe de A , \overline{A} la fermeture de A et $\overline{\text{conv}A}$ l'enveloppe convexe fermée de A . Sur l'espace X^* , on considère uniquement la topologie faible-*, et pour tout $B \subset X^*$, on note simplement \overline{B} la fermeture de B pour la topologie faible-*

Étant donné A et B deux sous-ensembles de X , on dit que A est fermé par rapport à B si $\overline{A} \cap B = A \cap B$ ([13]). De manière analogue, on dit que A est convexe fermé par rapport à B si $\overline{\text{conv}A} \cap B = A \cap B$ ([34]).

Étant donné $E \subset \overline{\mathbb{R}}$, on écrit $\min A$ (resp. $\max A$) au lieu de $\inf A$ (resp. $\sup A$) lorsque que l'infimum (resp. le supremum) est atteint, dans ce cas, on dit qu'il est exact.

Lemme 3.2.1 ([15]). *Pour toutes $h_1, h_2 \in \Gamma(X)$ telles que $\text{dom } h_1 \cap \text{dom } h_2 \neq \emptyset$, on a :*

$$(3.1) \quad \text{epi} (h_1 + h_2)^* = \overline{\text{epi } h_1^* + \text{epi } h_2^*}.$$

Lemme 3.2.2 ([13]). *Soit $(h_i)_{i \in I} \subset \Gamma(X)$, avec I un ensemble quelconque d'indices. Supposons qu'il existe $x \in X$ tel que $\sup_{i \in I} h_i(x) < +\infty$. Alors*

$$(3.2) \quad \text{epi} \left(\sup_{i \in I} h_i \right)^* = \overline{\text{conv} \left(\bigcup_{i \in I} \text{epi } h_i^* \right)}.$$

Les fermetures dans (3.1) et (3.2) sont prises par rapport à la topologie produit de la topologie faible-* sur X^* et de la topologie usuelle de \mathbb{R} . De plus, si, h_1 est finie et continue en un point de $\text{dom } h_2$, alors la fermeture n'est pas nécessaire dans (3.1) d'après le théorème de Moreau-Rockafellar ([60] Théorème 3).

Soit Y un autre espace vectoriel topologique localement convexe séparé et $S \subset Y$ un cône convexe fermé non vide. Le S -épigraphe d'une fonction $g : \text{dom } g \subset X \rightarrow Y$ est l'ensemble

$$\text{epi}_S g := \{(x, y) \in \text{dom } g \times Y \quad : \quad y - g(x) \in S\},$$

et la S -tranche de niveau $y \in Y$ est l'ensemble

$$[g \leq_S y] := \{x \in \text{dom } g \quad : \quad y - g(x) \in S\}.$$

En particulier,

$$[g \leq_S 0_Y] = \{x \in \text{dom } g \quad : \quad g(x) \in -S\} =: g^{-1}(-S).$$

Il existe dans la littérature plusieurs notions de semi-continuité inférieure des fonctions de type g ([1], [22],[27],[50], [57], ...). Nous utilisons ici celles qui dépendent

du S -épigraphe ou des S -tranches. On dira que g est S -épi-convexe fermée (ou S -convexe fermée par épigraphe) si $\text{epi}_S g$ est convexe et fermé, et g est S -convexe fermée par niveaux si $[g \leq_S y]$ est convexe et fermé pour tout $y \in Y$. Il est clair que toute fonction S -épi-convexe fermée est S -convexe fermée par niveaux.

On note

$$S^+ := \{\lambda \in Y^* \quad : \quad \langle y, \lambda \rangle \geq 0, \quad \forall y \in S\},$$

le cône polaire positif de S . Étant donné $\lambda \in S^+$, on note $\lambda g : X \rightarrow \overline{\mathbb{R}}$ la fonction définie par :

$$\lambda g(x) = \begin{cases} \langle g(x), \lambda \rangle & \text{si } x \in \text{dom } g \\ +\infty & \text{si } x \notin \text{dom } g. \end{cases}$$

Si g est définie sur tout l'espace X alors λg est la composée de λ par g notée habituellement par $\lambda \circ g$.

Considérons l'ensemble K_g défini par

$$K_g := \bigcup_{\lambda \in S^+} \text{epi}(\lambda g)^*,$$

que l'on peut voir comme le cône caractéristique du système d'inéquations

$$(3.3) \quad \{x \in \text{dom } g \quad : \quad \langle g(x), \lambda \rangle \leq 0, \quad \forall \lambda \in S^+\},$$

selon la terminologie utilisée dans [40]. Notons que sans aucune hypothèse de convexité sur g , l'ensemble K_g est toujours un cône convexe. Plus précisément, on a la propriété suivante :

Propriété 3.2.2. *Pour toute fonction $g : \text{dom } g \subset X \rightarrow Y$ tel que $g^{-1}(-S) \neq \emptyset$, on a :*

$$(1) \quad i_{g^{-1}(-S)} = \sup_{\lambda \in S^+} (\lambda g),$$

(2) K_g est un cône convexe.

Preuve. Montrons (1). Soit $x \in X$. Si $x \in g^{-1}(-S)$ alors $\langle g(x), \lambda \rangle \leq 0$ pour tout $\lambda \in S^+$, et comme $0_{Y^*} \in S^+$, on en déduit que

$$\sup_{\lambda \in S^+} (\lambda g(x)) = 0.$$

Si $x \notin g^{-1}(-S)$ alors $g(x) \notin -S$ qui est un convexe fermé. D'après le théorème de séparation de Hahn-Banach, il existe $(y^*, r) \in Y^* \times \mathbb{R}$ tel que $\langle y, y^* \rangle < r < \langle g(x), y^* \rangle$ pour tout $y \in -S$. Puisque $0_Y \in S$, il en résulte que $r > 0$ et donc $y^* \in S^+$. Donc

$$\sup_{\lambda \in S^+} (\lambda g(x)) \geq \langle g(x), ny^* \rangle > 0, \quad \forall n \geq 1.$$

Par conséquent

$$\sup_{\lambda \in S^+} (\lambda g(x)) \geq \lim_{n \rightarrow +\infty} \langle g(x), ny^* \rangle = +\infty,$$

d'où $\sup_{\lambda \in S^+}(\lambda g(x)) = +\infty$.

Montrons (2). Montrons d'abord que K_g est un cône. Soit $(x^*, r) \in K_g$ et $t > 0$. Il existe $\lambda \in S^+$ tel que $(\lambda g)^*(x^*) \leq r$ et on a

$$(\lambda g)^*(tx^*) = \sup_{x \in X} \{ \langle x, tx^* \rangle - t\lambda g(x) \} = t \sup_{x \in X} \{ \langle x, x^* \rangle - \lambda g(x) \} = t(\lambda g)^*(x^*) \leq tr.$$

Donc $t(x^*, r) \in \text{epi}(t\lambda g)^* \subset K_g$. Montrons maintenant que K_g est convexe.

Soit $(x_i^*, r_i) \in K_g$, $i = 1, 2$. On doit montrer que $(x_1^* + x_2^*, r_1 + r_2) \in K_g$. Il existe $\lambda_i \in S^+$ tel que $(x_i^*, r_i) \in \text{epi}(\lambda_i g)^*$. Pour tout $x \in \text{dom } g$, on a

$$\begin{aligned} \langle x, x_1^* + x_2^* \rangle - \langle g(x), \lambda_1 + \lambda_2 \rangle &= \langle x, x_1^* \rangle - \langle g(x), \lambda_1 \rangle + \langle x, x_2^* \rangle - \langle g(x), \lambda_2 \rangle \\ &\leq (\lambda_1 g)^*(x_1^*) + (\lambda_2 g)^*(x_2^*) \\ &\leq r_1 + r_2. \end{aligned}$$

En passant au sup sur $\text{dom } g$, on obtient $((\lambda_1 + \lambda_2)g)^*(x_1^* + x_2^*) \leq r_1 + r_2$, et donc

$$(x_1^* + x_2^*, r_1 + r_2) \in \text{epi}((\lambda_1 + \lambda_2)g)^* \subset K_g.$$

□

En présence d'hypothèse de convexité sur g , on a :

Proposition 3.2.1. *Pour toute fonction S -épi-convexe fermée $g : \text{dom } g \subset X \rightarrow Y$ tel que $g^{-1}(-S) \neq \emptyset$, on a*

$$(3.4) \quad \text{epi } \sigma_{g^{-1}(-S)} = \overline{K_g}.$$

Preuve. Introduisons la fonction $H : X \times Y \rightarrow \mathbb{R} \cup \{+\infty\}$ définie par

$$H(x, y) = i_{\text{epi}_S g}(x, -y).$$

Soit $(x^*, \lambda) \in X^* \times Y^*$. On a :

$$\begin{aligned} H^*(x^*, \lambda) &= \sup_{\substack{x \in X \\ y \in Y}} \{ \langle x, x^* \rangle + \langle y, \lambda \rangle - i_{\text{epi}_S g}(x, -y) \} \\ &= \sup_{x \in X} \left\{ \langle x, x^* \rangle + \sup_{-y-g(x) \in S} \langle y, \lambda \rangle \right\} \\ &= \sup_{x \in X} \left\{ \langle x, x^* \rangle - \langle g(x), \lambda \rangle - \inf_{y \in S} \langle y, \lambda \rangle \right\} \\ H^*(x^*, \lambda) &= \begin{cases} (\lambda g)^*(x^*) & \text{si } \lambda \in S^+ \\ +\infty & \text{sinon .} \end{cases} \end{aligned}$$

Puisque g est S -épi-convexe fermée et $g^{-1}(-S) \neq \emptyset$ alors $H \in \Gamma(X \times Y)$. En particulier, pour tout $x \in X$, on a

$$\begin{aligned}
H(x, 0_Y) &= H^{**}(x, 0_Y) \\
&= \sup_{\substack{x^* \in X^* \\ \lambda \in Y^*}} \{ \langle x, x^* \rangle - \langle 0_Y, \lambda \rangle - H^*(x^*, \lambda) \} \\
&= \sup_{\substack{x^* \in X^* \\ \lambda \in S^+}} \{ \langle x, x^* \rangle - (\lambda g)^*(x^*) \} \\
&= \sup_{\lambda \in S^+} \left\{ \sup_{x^* \in X^*} \{ \langle x, x^* \rangle - (\lambda g)^*(x^*) \} \right\} \\
&= \sup_{\lambda \in S^+} (\lambda g)^{**}(x).
\end{aligned}$$

D'autre part, pour tout $x \in X$, on a

$$H(x, 0_Y) = i_{\text{epi}_{Sg}}(x, 0_Y) = i_{g^{-1}(-S)}(x).$$

Donc $i_{g^{-1}(-S)} = \sup_{\lambda \in S^+} (\lambda g)^{**}$, et comme $g^{-1}(-S) \neq \emptyset$, d'après (3.2), on a

$$\text{epi } \sigma_{g^{-1}(-S)} = \overline{\text{conv}} \left(\bigcup_{\lambda \in S^+} \text{epi } (\lambda g)^{***} \right) = \overline{\text{conv}} \left(\bigcup_{\lambda \in S^+} \text{epi } (\lambda g)^* \right) = \overline{K_g},$$

où la dernière égalité résulte de la propriété 3.2.2 (2) et de la définition de K_g . \square

Remarque 3.2.1. *Malgré le fait que $i_{g^{-1}(-S)} = \sup_{\lambda \in S^+} (\lambda g)$, on ne peut pas appliquer directement le lemme 3.2 à la fonction $i_{g^{-1}(-S)}$ pour obtenir (3.4). La raison est que λg n'est pas nécessairement semi-continue inférieurement même si g est S -épi-convexe fermée. En effet, si g est une fonction S -épi-convexe fermée dont le domaine effectif n'est pas fermé, alors pour $\lambda = 0_{Y^*}$, on a $\lambda g = i_{\text{dom } g}$ qui n'est pas semi-continue inférieurement.*

3.3. Pire valeur versus valeur robuste

Dans cette section, nous donnons une condition nécessaire et suffisante permettant d'obtenir l'égalité entre la pire valeur et la valeur robuste du problème incertain (P) avec exactitude de la pire valeur :

$$(3.5) \quad \inf(RP) = \max_{u \in U} \inf(P_u).$$

Pour tout $u \in U$, on désigne par F_u , l'ensemble des solutions réalisables de (P_u) :

$$F_u := \{x \in \text{dom } g_u : g_u(x) \in -S\}.$$

On note F l'ensemble des solutions réalisables de (RP) :

$$F := \{x \in X : x \in \text{dom } g_u \text{ et } g_u(x) \in -S, \forall u \in U\} = \bigcap_{u \in U} F_u,$$

et on définit la fonction $p : X \rightarrow \overline{\mathbb{R}}$ par

$$p(x) := \sup_{u \in U} (f(x) + i_{F_u}(x)) = f(x) + \sup_{u \in U} i_{F_u}(x) = f(x) + i_F(x).$$

On a

$$\text{dom } p = F \cap \text{dom } f \quad \text{et} \quad \inf_{x \in X} p(x) = \inf(RP).$$

Proposition 3.3.1. *On a toujours*

$$(3.6) \quad \sup(Q) \leq \inf(RP).$$

Preuve.

$$\sup(Q) = \sup_{u \in U} \inf(P_u) = \sup_{u \in U} \inf_{x \in X} (f + i_{F_u})(x) \leq \inf_{x \in X} \sup_{u \in U} (f + i_{F_u})(x) = \inf(RP). \quad \square$$

L'exemple suivant montre que l'inégalité (3.6) peut être stricte même si f est linéaire, l'ensemble U est fini et $\inf(RP)$ est atteint.

Exemple 3.3.1. *Considérons le problème convexe conique incertain suivant :*

$$(P) \quad \text{Minimiser } x_1 + x_2 \quad \text{s.l.c.} \quad \frac{1}{2}[(2 - u_1)x_1^2 + (1 + u_2)x_2^2] \leq 1,$$

où $(u_1, u_2) = (1, 2)$ ou $(u_1, u_2) = (\frac{1}{2}, 1)$. Dans cet exemple, on a :

$$X = \mathbb{R}^2, \quad f(x_1, x_2) = x_1 + x_2, \quad Y = \mathbb{R}, \quad S = \mathbb{R}_+, \quad U = \{(1, 2), (\frac{1}{2}, 1)\},$$

$$g_u(x_1, x_2) = \frac{1}{2}[(2 - u_1)x_1^2 + (1 + u_2)x_2^2] - 1.$$

Pour tout $u \in U$, (P_u) est un programme convexe vérifiant la condition de qualification de Slater. En résolvant le système d'optimalité de Karush-Kuhn-Tucker correspondant, on obtient

$$\min(P_u) = -\sqrt{2 \left(\frac{1}{2 - u_1} + \frac{1}{1 + u_2} \right)} = \begin{cases} -\sqrt{\frac{8}{3}} & \text{si } u = (1, 2) \\ -\sqrt{\frac{7}{3}} & \text{si } u = (\frac{1}{2}, 1). \end{cases}$$

Donc

$$\max(Q) = -\sqrt{\frac{7}{3}}.$$

D'autre part, la contrepartie robuste de (P) est donnée par

$$(RP) \quad \text{Minimiser } x_1 + x_2 \quad \text{s.l.c.} \quad \begin{cases} x_1^2 + 3x_2^2 \leq 2 \\ 3x_1^2 + 4x_2^2 \leq 4. \end{cases}$$

(RP) est un programme convexe vérifiant la condition de qualification de Slater. En résolvant le système d'optimalité de Karush-Kuhn-Tucker correspondant, on obtient

$$\min(RP) = -\frac{2 + \sqrt{2}}{\sqrt{5}} > -\sqrt{\frac{7}{3}} = \max(Q).$$

Considérons « l'opposé » du problème (Q) :

$$(-Q) \quad \inf_u \sup\{-f(x) \text{ s.l.c } g_u(x) \in -S\} \quad \text{s.l.c } u \in U.$$

Une perturbation de la fonction objective de (-Q) par l'addition d'une forme linéaire continue permet de définir une fonction $q : X^* \rightarrow \overline{\mathbb{R}}$ par

$$q(x^*) := \inf_{u \in U} \sup_{x \in F_u} \{\langle x, x^* \rangle - f(x)\} = \inf_{u \in U} (f + i_{F_u})^*(x^*).$$

D'après la propriété 3.2.1, on a

$$q^* = \sup_{u \in U} (f + i_{F_u})^{**} \leq \sup_{u \in U} (f + i_{F_u}) = p,$$

donc

$$(3.7) \quad p^* \leq q^{**} \leq q.$$

Introduisons la condition de qualification suivante :

$$(CQ) \quad \begin{cases} f \in \Gamma(X) \\ F \cap \text{dom } f \neq \emptyset \text{ (c'est-à-dire } \inf(RP) < +\infty) \\ g_u \text{ est } S\text{-convexe fermée par niveaux, } \forall u \in U. \end{cases}$$

Lemme 3.3.1. *Supposons que (CQ) est vérifiée. Alors*

$$(3.8) \quad \text{epi } p^* = \overline{\text{conv}}\left(\bigcup_{u \in U} \text{epi } (f + i_{F_u})^*\right).$$

Preuve. Si la condition (CQ) est vérifiée alors $f + i_{F_u} \in \Gamma(X)$ pour tout $u \in U$, et on a

$$q^* = \sup_{u \in U} (f + i_{F_u})^{**} = \sup_{u \in U} (f + i_{F_u}) = p.$$

donc $q^{**} = p^*$ et $\text{dom } q^* = F \cap \text{dom } f \neq \emptyset$. On applique le lemme 3.2 à la fonction $q^* = p$ pour terminer. \square

Théorème 3.3.1. *Supposons que (CQ) est vérifiée. Alors pour tout $x^* \in X^*$, les affirmations suivantes sont équivalentes :*

$$(1) \quad p^*(x^*) = \min_{u \in U} \sup_{x \in F_u} \{\langle x, x^* \rangle - f(x)\},$$

$$(2) \quad \bigcup_{u \in U} \text{epi } (f + i_{F_u})^* \text{ est convexe fermé par rapport à } \{x^*\} \times \mathbb{R}.$$

Preuve. Puisque $\text{dom } p \neq \emptyset$, la fonction p^* ne prend pas la valeur $-\infty$. Soit $x^* \in X^*$. On considère d'abord le cas où $p^*(x^*) = +\infty$. Comme $p^* \leq q$, on a $q(x^*) = +\infty$ et (1) est vérifiée. D'autre part, d'après le lemme 3.3.1, on a

$$\overline{\text{conv}}\left(\bigcup_{u \in U} \text{epi } (f + i_{F_u})^*\right) \cap (\{x^*\} \times \mathbb{R}) = \text{epi } p^* \cap (\{x^*\} \times \mathbb{R}) = \emptyset, \text{ d'où (2).}$$

Par conséquent, si $p^*(x^*) = +\infty$ alors (1) et (2) sont toutes les deux vérifiées.

On suppose maintenant que $p^*(x^*) \in \mathbb{R}$.

Montrons que (2) \implies (1). D'après le lemme 3.3.1,

$$(x^*, p^*(x^*)) \in \text{epi } p^* \cap (\{x^*\} \times \mathbb{R}) = \overline{\text{conv}} \left(\bigcup_{u \in U} \text{epi } (f + i_{F_u})^* \right) \cap (\{x^*\} \times \mathbb{R}).$$

Donc, d'après (2),

$$(x^*, p^*(x^*)) \in \left(\bigcup_{u \in U} \text{epi } (f + i_{F_u})^* \right) \cap (\{x^*\} \times \mathbb{R}).$$

Par conséquent, il existe $\bar{u} \in U$ tel que

$$\inf_{u \in U} (f + i_{F_u})^*(x^*) = q(x^*) \leq (f + i_{F_{\bar{u}}})^*(x^*) \leq p^*(x^*).$$

Comme $p^*(x^*) \leq q(x^*)$, on obtient (1).

Montrons que (1) \implies (2). Soit $(x^*, r) \in \overline{\text{conv}} \left(\bigcup_{u \in U} \text{epi } (f + i_{F_u})^* \right)$. Alors $p^*(x^*) \leq r$ d'après le lemme 3.3.1. D'après (1), il existe $\bar{u} \in U$ tel que

$$p^*(x^*) = (f + i_{F_{\bar{u}}})^*(x^*)$$

et finalement $(x^*, r) \in \text{epi } (f + i_{F_{\bar{u}}})^*$. \square

Corollaire 3.3.1. *Supposons que (CQ) est vérifiée. Alors, les affirmations suivantes sont équivalentes :*

$$(1) \quad -\infty \leq \max_{u \in U} \inf(P_u) = \inf(RP) < +\infty,$$

$$(2) \quad \bigcup_{u \in U} \text{epi } (f + i_{F_u})^* \text{ est convexe fermé par rapport à } \{0_{X^*}\} \times \mathbb{R}.$$

Preuve. Notons que $-p^*(0_{X^*}) = \inf(RP)$. La conclusion résulte immédiatement du théorème 3.3.1 appliqué au point 0_{X^*} . \square

Corollaire 3.3.2. *Supposons que (CQ) est vérifiée. Alors, les affirmations suivantes sont équivalentes :*

$$(1) \quad -\infty < p^*(x^*) = \min_{u \in U} (f + i_{F_u})^*(x^*) \leq +\infty, \quad \forall x^* \in X^*,$$

$$(2) \quad \text{l'ensemble } \mathbb{A} := \bigcup_{u \in U} \text{epi } (f + i_{F_u})^* \text{ est convexe fermé.}$$

Preuve. Notons que l'ensemble \mathbb{A} est convexe fermé si et seulement si \mathbb{A} est convexe fermé par rapport à $\{x^*\} \times \mathbb{R}$ pour tout $x^* \in X^*$. La conclusion résulte donc du théorème 3.3.1. \square

On renforce maintenant la condition (CQ), en supposant que pour tout $u \in U$, la fonction g_u est S -épi-convexe fermée. On considère la nouvelle condition (CQ') :

$$(CQ') \quad \begin{cases} f \in \Gamma(X) \\ F \cap \text{dom } f \neq \emptyset \text{ (c'est-à-dire } \inf(RP) < +\infty) \\ g_u \text{ est } S\text{-épi-convexe fermée, } \forall u \in U. \end{cases}$$

Corollaire 3.3.3. *Supposons que (CQ') est vérifiée. Les assertions suivantes sont équivalentes :*

$$(1) \inf(RP) = \max(Q),$$

$$(2) \bigcup_{u \in U} \overline{(\text{epi } f^* + K_{g_u})} \text{ est convexe fermé par rapport à } \{0_{X^*}\} \times \mathbb{R}.$$

Preuve. Puisque F est non vide alors pour tout $u \in U$, $g_u^{-1}(-S) = F_u \neq \emptyset$. En appliquant la proposition 3.2.1 à la fonction g_u , on obtient

$$\text{epi } i_{F_u}^* = \overline{K_{g_u}}, \quad \forall u \in U.$$

D'après le lemme 3.2.1, on a, pour tout $u \in U$,

$$\begin{aligned} \text{epi } (f + i_{F_u})^* &= \overline{(\text{epi } f^* + \text{epi } i_{F_u}^*)} \\ &= \overline{(\text{epi } f^* + \overline{K_{g_u}})} \\ &= \overline{(\text{epi } f^* + K_{g_u})}. \end{aligned}$$

Par conséquent,

$$\bigcup_{u \in U} \text{epi } (f + i_{F_u})^* = \bigcup_{u \in U} \overline{(\text{epi } f^* + K_{g_u})}$$

et la conclusion résulte du corollaire 3.3.1. \square

3.4. Lien avec la dualité robuste forte

Pour tout $u \in U$, on associe à (P_u) , son dual Lagrangien classique (D_u) défini par :

$$(D_u) \quad \sup_{\lambda} \inf_{x \in X} \{f(x) + \lambda g_u(x)\} \quad \text{s.l.c.} \quad \lambda \in S^+.$$

Le dual optimiste du problème (P) est défini ([5], [14],[47],[49]) par

$$(ODP) \quad \sup_{(u,\lambda)} \inf_{x \in X} \{f(x) + \lambda g_u(x)\} \quad \text{s.l.c.} \quad (u, \lambda) \in U \times S^+.$$

Proposition 3.4.1. *La valeur de (ODP) est inférieure où égale à celle de (Q) :*

$$(3.9) \quad \sup(ODP) \leq \sup(Q).$$

Preuve. Par la dualité faible entre les problèmes (P_u) et (D_u) , $u \in U$, on a

$$\sup_{\lambda \in S^+} \inf_{x \in X} \{f(x) + \lambda g_u(x)\} \leq \inf(P_u).$$

En passant au sup sur U , on obtient

$$\sup(ODP) = \sup_{\substack{u \in U \\ \lambda \in S^+}} \inf_{x \in X} \{f(x) + \lambda g_u(x)\} \leq \sup_{u \in U} \inf(P_u) = \sup(Q).$$

\square

Remarque 3.4.1. *Il est intéressant de noter que, même dans le cas certain c'est-à-dire $U = \{\bar{u}\}$, il peut arriver que (voir le lemme 3.4.2 ci-dessous) :*

$$\sup_{\lambda \in S^+} \inf_{x \in X} \{f(x) + \lambda g_{\bar{u}}(x)\} = \sup(ODP) < \sup(Q) = \inf(P_{\bar{u}}).$$

Remarque 3.4.2. *En juxtaposant les propositions 3.3.1 et 3.4.1, on obtient*

$$(3.10) \quad \sup(ODP) \leq \sup(Q) \leq \inf(RP).$$

On dit que la **dualité robuste forte** est satisfaite pour le problème convexe conique incertain (P) si les valeurs de sa contrepartie robuste et de son dual optimiste coïncident avec exactitude de la valeur du dual optimiste :

$$(3.11) \quad \inf(RP) = \max(ODP).$$

La terminologie de dualité robuste forte a été introduite dans [49]. Cette propriété se trouve dans [5] et [47] sous l'expression « primal worst equals dual best ».

Proposition 3.4.2. *Si la dualité robuste forte est satisfaite pour (P) alors on a :*

$$\inf(RP) = \max(Q).$$

Preuve. En juxtaposant les propositions 3.3.1 et 3.4.1, on obtient

$$\max(ODP) = \sup(Q) = \inf(RP), \text{ d'après (3.11) .}$$

Par conséquent, il existe $(\bar{u}, \bar{\lambda}) \in U \times S^+$ tel que

$$\inf_{x \in X} \{f(x) + \bar{\lambda}g_{\bar{u}}(x)\} = \inf(RP) = \sup(Q) \geq \inf(P_{\bar{u}}) \geq \inf_{x \in X} \{f(x) + \bar{\lambda}g_{\bar{u}}(x)\},$$

où la dernière inégalité résulte de la dualité faible entre les problèmes $(P_{\bar{u}})$ et $(D_{\bar{u}})$. Donc $\inf(RP) = \inf(P_{\bar{u}})$ ce qui achève la preuve. \square

Afin d'obtenir la dualité robuste forte à partir de nos résultats précédents, nous rappelons deux résultats de dualité épigraphique en optimisation convexe.

Lemme 3.4.1 ([13] Théorème 8.3). *Soit $f \in \Gamma(X)$, $g : \text{dom } g \subset X \rightarrow Y$ une fonction S -épi-convexe fermée telles que $g^{-1}(-S) \cap \text{dom } f \neq \emptyset$. Alors les assertions suivantes sont équivalentes :*

- (1) $\inf_{g(x) \in -S} \{f(x) - \langle x, x^* \rangle\} = \max_{\lambda \in S^+} \inf_{x \in X} \{f(x) - \langle x, x^* \rangle + \lambda g(x)\}, \quad \forall x^* \in X^*,$
- (2) $\bigcup_{\lambda \in S^+} \text{epi}(f + \lambda g)^*$ est fermé.

Lemme 3.4.2 ([27] Corollaire 5). *Soit $f \in \Gamma(X)$, $g : \text{dom } g \subset X \rightarrow Y$ une fonction S -épi-convexe fermée telles que $g^{-1}(-S) \cap \text{dom } f \neq \emptyset$. Alors les assertions suivantes sont équivalentes :*

- (1) $\inf_{g(x) \in -S} f(x) = \max_{\lambda \in S^+} \inf_{x \in X} \{f(x) + \lambda g(x)\},$
- (2) $\bigcup_{\lambda \in S^+} \text{epi}(f + \lambda g)^*$ est fermé par rapport à $\{0_{X^*}\} \times \mathbb{R}$.

Revenons aux problèmes (ODP) et (Q) .

Proposition 3.4.3. *Supposons que la condition (CQ') est satisfaite et pour tout $u \in U$, l'ensemble $\bigcup_{\lambda \in S^+} \text{epi}(f + \lambda g_u)^*$ est fermé par rapport à $\{0_{X^*}\} \times \mathbb{R}$. Alors*

$$\sup(ODP) = \sup(Q).$$

Preuve. En appliquant le lemme 3.4.2, on obtient

$$\inf_{x \in F_u} f(x) = \max_{\lambda \in S^+} \inf_{x \in X} \{f(x) + \lambda g_u(x)\}, \quad \forall u \in U.$$

En passant au sup sur U , on obtient

$$\sup(Q) = \sup(ODP).$$

□

Remarque 3.4.3. *Considérons l'exemple 3.3.1. Comme la condition de qualification de Slater est satisfaite alors l'ensemble $\bigcup_{\lambda \geq 0} \text{epi}(f + \lambda g_u)^*$ est fermé pour tout $u \in U$ ([18], Remarque 4.3), et d'après la proposition 3.4.3, on a :*

$$\sup(Q) = \sup(ODP) < \inf(RP).$$

On note $\text{Argmax}(Q)$, l'ensemble des solutions optimales du problème (Q) :

$$\text{Argmax}(Q) = \{u \in U \quad : \quad \inf(P_u) = \sup(Q)\}.$$

Théorème 3.4.1. *Supposons que (CQ) est satisfaite et que*

$$(3.12) \quad \bigcup_{u \in U} \text{epi}(f + i_{F_u})^* \text{ est convexe fermé par rapport à } \{0_{X^*}\} \times \mathbb{R},$$

$$(3.13) \quad \exists \bar{u} \in \text{Argmax}(Q) : \begin{cases} g_{\bar{u}} \text{ est } S\text{-epi-convexe fermée,} \\ \bigcup_{\lambda \in S^+} \text{epi}(f + \lambda g_{\bar{u}})^* \text{ est fermé} \\ \text{par rapport à } \{0_{X^*}\} \times \mathbb{R}. \end{cases}$$

Alors la propriété de dualité robuste forte a lieu.

Preuve. On a

$$\begin{aligned} \inf(RP) &= \max(Q) \quad (\text{résulte du corollaire 3.3.1 et de (3.12)}) \\ &= \inf(P_{\bar{u}}) \quad (\text{pour un certain } \bar{u} \in \text{Argmax}(Q) \text{ vérifiant (3.13)}) \\ &= \max(D_{\bar{u}}) \quad (\text{d'après (3.13) et le lemme 3.4.2}) \\ &\leq \sup(ODP) \quad (\text{par définition du dual optimiste}). \end{aligned}$$

D'après la proposition 3.4.1, on a

$$\sup(ODP) \leq \sup(Q) = \inf(RP) = \max(D_{\bar{u}}) \leq \sup(ODP).$$

Il existe donc $(\bar{u}, \bar{\lambda}) \in U \times S^+$ tel que

$$\sup(ODP) = \inf_{x \in X} \{f(x) + \bar{\lambda} g_{\bar{u}}(x)\} = \inf(RP),$$

d'où le résultat. □

On dit que la dualité robuste forte a lieu au point $x^* \in X^*$ si

$$(3.14) \quad \inf_{x \in F} \{f(x) - \langle x, x^* \rangle\} = \max_{\substack{\lambda \in S^+ \\ u \in U}} \inf_{x \in X} \{f(x) - \langle x, x^* \rangle + \lambda g_u(x)\}.$$

Pour $x^* = 0_{X^*}$, on retrouve la propriété de dualité robuste forte. Si la propriété de dualité a lieu en tout $x^* \in X^*$ alors, on dit que (P) satisfait la propriété de **dualité robuste forte stable**.

Corollaire 3.4.1. *Supposons que la condition (CQ') est satisfaite et*

$$(3.15) \quad \bigcup_{u \in U} \overline{\text{epi } f^* + K_{g_u}} \text{ est convexe fermé,}$$

$$(3.16) \quad \forall u \in U, \quad \bigcup_{\lambda \in S^+} \text{epi } (f + \lambda g_u)^* \text{ est fermé.}$$

Alors le problème (P) satisfait la propriété de dualité robuste forte stable.

Preuve. D'après le lemme 3.1 et la proposition 3.2.1, on a :

$$\mathbb{A} := \bigcup_{u \in U} \text{epi } (f + i_{F_u})^* = \bigcup_{u \in U} \overline{\text{epi } f^* + K_{g_u}}.$$

D'après (3.15), l'ensemble \mathbb{A} est convexe fermé. Donc pour tout $x^* \in X^*$ on a

$$\begin{aligned} \inf_{x \in F} \{f(x) - \langle x, x^* \rangle\} &= \max_{u \in U} \inf_{x \in F_u} \{f(x) - \langle x, x^* \rangle\} \quad (\text{d'après le corollaire 3.3.2}) \\ &= \inf_{x \in F_{\bar{u}}} \{f(x) - \langle x, x^* \rangle\} \quad (\text{pour un certain } \bar{u} \in U) \\ &= \max_{\lambda \in S^+} \inf_{x \in X} \{f(x) - \langle x, x^* \rangle + \lambda g_{\bar{u}}(x)\} \quad (\text{d'après le lemme 3.4.2 et (3.16)}) \\ &= \inf_{x \in X} \{f(x) - \langle x, x^* \rangle + \bar{\lambda} g_{\bar{u}}(x)\} \quad (\text{pour un certain } \bar{\lambda} \in S^+) \\ &\leq \sup_{\substack{\lambda \in S^+ \\ u \in U}} \inf_{x \in X} \{f(x) - \langle x, x^* \rangle + \lambda g_u(x)\}. \end{aligned}$$

Ainsi, par la dualité faible (appliquer les propositions 3.3.1 et 3.4.1 à $f - x^*$), on a

$$\sup_{\substack{\lambda \in S^+ \\ u \in U}} \inf_{x \in X} \{f(x) - \langle x, x^* \rangle + \lambda g_u(x)\} \leq \inf_{x \in F} \{f(x) - \langle x, x^* \rangle\},$$

et nous pouvons conclure. □

Remarque 3.4.4. *Supposons que pour tout $u \in U$, la fonction $g_u : X \rightarrow Y$ est continue et S -épi-convexe. Pour tout $\lambda \in S^+$, la fonction $\lambda g_u = \lambda \circ g_u$ est convexe et continue, et par le théorème de Moreau-Rockafellar, on a*

$$\text{epi } (f + \lambda g_u)^* = \text{epi } f^* + \text{epi } (\lambda g_u)^*.$$

Dans ce cas, la condition (3.16) devient

$$(3.17) \quad \forall u \in U, \quad \text{epi } f^* + K_{g_u} \text{ est fermé.}$$

La propriété de dualité robuste forte a été établie dans [49] (Corollaire 3.1) pour des fonctions $g_u : X \rightarrow Y$ continues et S -épi-convexes sous la condition

$$(3.18) \quad \text{epi } f^* + \bigcup_{u \in U} K_{g_u} \text{ est fermé}$$

sans supposer (3.16) (qui est (3.17) dans ce cas) comme nous l'avons fait dans le corollaire 3.4.1. Cependant, notons que (3.15) est plus faible que (3.18).

Proposition 3.4.4. *La condition (3.15) est plus faible que la condition (3.18).*

Preuve. On a

$$\text{epi } f^* + \bigcup_{u \in U} K_{g_u} \subset \bigcup_{u \in U} \overline{(\text{epi } f^* + K_{g_u})} \subset \overline{\left(\text{epi } f^* + \bigcup_{u \in U} K_{g_u} \right)} \subset \overline{\text{conv} \left(\text{epi } f^* + \bigcup_{u \in U} K_{g_u} \right)}.$$

Donc si (3.18) est satisfaite, alors les inclusions ci-dessus sont des égalités et en particulier

$$\bigcup_{u \in U} \overline{(\text{epi } f^* + K_{g_u})} = \overline{\text{conv} \left(\text{epi } f^* + \bigcup_{u \in U} K_{g_u} \right)},$$

qui est un ensemble convexe et fermé. □

II. Contribution à l'étude de la robustesse

Rayon de stabilité en optimisation continue à données incertaines

4.1. Un modèle de prise de décision en environnement incertain

4.1.1. Formulation du problème

Considérons le modèle d'événements incertains suivant : tout ce que l'on sait c'est qu'un événement incertain w appartient à un ensemble W . A priori, on ne possède aucune information supplémentaire (on ne connaît pas de probabilité associée à cette incertitude) sur cet événement incertain.

On se donne un sous-ensemble $S \subset W$ ($S \in 2^W$) qui est fixé avant que l'événement incertain w ne se produise.

Une fois que l'événement incertain w s'est produit, deux situations se présentent :

- si $w \in S$ la situation est jugée «bonne» ou «satisfaisante» ;
- si $w \notin S$ la situation est considérée comme «moins bonne» ou «insatisfaisante».

On appellera S l'ensemble de satisfaction et S^c , le complémentaire de S dans W l'ensemble d'insatisfaction.

Imaginons maintenant que l'on possède un certain pouvoir de décision, un certain degré de liberté sur le choix de l'ensemble de satisfaction S (qui reste à préciser), que l'on puisse le choisir dans $\mathcal{S} \subset 2^W$. Un problème naturel est alors :

$(\mathcal{R}_{\text{init}})$ Quel est un/le «meilleur» choix pour l'ensemble de satisfaction $S \in \mathcal{S}$?

Naturellement, si S_1 et S_2 sont deux ensembles de satisfaction,

$$S_1 \supset S_2 \implies S_1 \text{ «est préférable à» } S_2.$$

Malheureusement, la relation d'inclusion a peu de chance d'être une relation d'ordre totale sur \mathcal{S} , cette propriété ne permettra pas, sauf cas très particulier, de répondre au problème $(\mathcal{R}_{\text{init}})$.

La situation telle que décrite ci-dessus est incomplète et ne permet pas d'aborder le problème $(\mathcal{R}_{\text{init}})$.

4.1.2. Paramétrisation du modèle

On reprend la situation décrite ci-dessus et on suppose qu'elle dépend d'une décision à prendre x appartenant à un ensemble D . On peut donc voir les choses de la façon suivante : suite à une décision $x \in D$, se produit un évènement incertain $w_x \in W_x$ (W_x peut donc dépendre de x). A tout $x \in D$, on associe un ensemble de satisfaction $S_x \subset W_x$. On définit ainsi une multi-application

$$S : D \rightrightarrows W = \bigcup_{x \in D} W_x,$$

qui à x associe son ensemble de satisfaction S_x , et on pose

$$\mathcal{S} = \{S_x : x \in D\} = S(D).$$

Le problème $(\mathcal{R}_{\text{init}})$ devient

$(\mathcal{R}_{\text{init}})$ *Quelle est une/la «meilleure» décision $x \in D$ à prendre ?*

On remarquera que, à priori, si les ensembles W_x sont disjoints alors il n'y a pas de comparaison possible au sens de l'inclusion entre un ensemble de satisfaction $S_{x_1} \subset W_{x_1}$ correspondant à une décision x_1 et un ensemble de satisfaction $S_{x_2} \subset W_{x_2}$ correspondant à une décision x_2 . Dans ce cas, on modélise le problème de la façon suivante. Soit $\mathcal{W} = (D, W, \pi)$ un fibré «ensembliste», c'est-à-dire que D et W sont deux ensembles et $\pi : W \rightarrow D$ est une surjection. D est la *base* du fibré, W l'*espace total*.

Pour tout $x \in D$, $W_x := \pi^{-1}(\{x\}) = \{w \in W \mid \pi(w) = x\}$ est la *fibres au dessus de x* . On se donne également $S : D \rightrightarrows W$ une multi-section de \mathcal{W} (c'est-à-dire pour tout $x \in D$, $S_x \subset W_x$).

Notons que dans cette construction, la famille $(W_x)_{x \in D}$ est une partition de W .

Notre modèle de prise de décision en environnement incertain devient : on prend une décision $x \in D$, ensuite un évènement $w_x \in W_x$ se produit et la décision est plus ou moins bonne suivant que $w_x \in S_x$ ou non.

4.1.3. Mesure ou indicateur de robustesse

4.1.3.1. Modèle non paramétré (simple)

Une manière de commencer à préciser les choses est d'associer, tout en le justifiant, à chaque ensemble de satisfaction $S \in \mathcal{S} \subset 2^W$, un nombre positif $r(S) \in [0; +\infty]$ tel qu'un ensemble S soit d'autant plus souhaitable que $r(S)$ est grand. Anticipant un peu ce qui suit, on pourra qualifier $r : \mathcal{S} \rightarrow [0, +\infty]$ de mesure ou indicateur de robustesse. C'est en fait, pour le moment, un simple indicateur de préférence.

Il est naturel d'imposer à r au minimum la propriété suivante :

$$(4.1) \quad \text{pour } S_1, S_2 \text{ quelconques dans } \mathcal{S}, \quad S_1 \subset S_2 \implies r(S_1) \leq r(S_2).$$

On peut maintenant aborder le problème $(\mathcal{R}_{\text{init}})$ à travers le problème d'optimisation naturel suivant :

$$(\mathcal{R}) \quad \text{Maximiser } r(S), \quad S \in \mathcal{S}.$$

Il faut noter que le choix d'un indicateur r et le passage du problème $(\mathcal{R}_{\text{init}})$ au problème (\mathcal{R}) reviennent à enrichir (de façon plus ou moins explicite) le modèle et le problème initial.

4.1.3.2. Modèle paramétré

On se place dans le cadre de la sous-section 4.1.2. On sait que dans ce cas, à chaque décision $x \in D$, correspond un ensemble de satisfaction $S_x \subset W_x$. On associe à cet ensemble de satisfaction un nombre positif $r(S_x) \in [0; +\infty]$ et on suppose que r vérifie la propriété suivante

$$(4.2) \quad \text{pour } x, y \text{ quelconques dans } D, \quad S_x \subset S_y \implies r(S_x) \leq r(S_y).$$

Le problème (\mathcal{R}) s'écrit

$$(\mathcal{R}) \quad \text{Maximiser } r(S_x), \quad x \in D.$$

4.1.4. Rayon de stabilité ou de robustesse

4.1.4.1. Le rayon de stabilité «classique»

On se donne un événement incertain **nominal** \bar{w} dans W que l'on peut voir comme l'événement «auquel on s'attend le plus». On se donne également une distance d sur W puis on définit alors l'indicateur de robustesse d'un ensemble de satisfaction $S \subset W$ (relativement à \bar{w} et d) comme étant la plus petite distance entre \bar{w} et l'ensemble d'insatisfaction S^c . Ceci donne la définition suivante :

$$(4.3) \quad r(S, \bar{w}, d) := d(\bar{w}, S^c) := \inf_{\mathbb{R}_+} \left\{ d(\bar{w}, w) \mid w \in S^c \right\},$$

avec la convention classique $\inf_{\mathbb{R}_+} \emptyset = +\infty$.

On remarquera que $r(S) = +\infty$ si et seulement si $S = W$.

Il est assez facile de voir que cet indicateur est compatible avec la relation d'inclusion (4.1). On l'appelle **rayon de stabilité ou de robustesse**.

Si pour $\rho \geq 0$, on note $\mathbb{B}(\bar{w}, \rho)$ la boule fermée de centre \bar{w} et de rayon ρ pour la distance d (on autorise $\rho = 0$),

$$\mathbb{B}(\bar{w}, \rho) = \left\{ w \in W \mid d(\bar{w}, w) \leq \rho \right\},$$

alors on peut écrire (d'après le lemme 1.5.1)

$$(4.4) \quad r(S, \bar{w}, d) = \sup_{\mathbb{R}_+} \left\{ \rho \geq 0 \mid \mathbb{B}(\bar{w}, \rho) \subset S \right\} = \sup_{\mathbb{R}_+} \left\{ \rho > 0 \mid \mathbb{B}(\bar{w}, \rho) \subset S \right\},$$

avec la convention $\sup_{\mathbb{R}_+} \emptyset = 0$.

Notons que (4.4) reste vraie si on remplace la boule fermée par la boule ouverte.

D'après (4.4), on voit que (en abusant un peu) $r(S, \bar{w}, d)$ est le rayon de la plus grosse boule centrée en \bar{w} et contenue dans S .

Définition 4.1.1 (Rayon de stabilité). *On appelle **rayon de stabilité ou de robustesse de $S \subset W$** (relativement à w et d), l'élément de $\overline{\mathbb{R}}_+$ défini par :*

$$r(S, w, d) := d(w, S^c).$$

On notera éventuellement $r(S, \bar{w}, d)$ par $r(S, \bar{w})$, $r(S)$, etc, sous-entendant éventuellement des éléments définis par le contexte.

Le problème (\mathcal{R}) devient

$$(\mathcal{R}') \quad \text{Maximiser } r(S, \bar{w}, d), \quad S \in \mathcal{S}.$$

On lui associe de façon naturelle (grâce à (4.4)) le problème

$$(\mathcal{R}'') \quad \begin{array}{l} \text{Maximiser } \rho, \\ \mathbb{B}(\bar{w}, \rho) \subset S, \\ \text{s.l.c.} \\ (\rho, S) \in \mathbb{R}_+^* \times \mathcal{S}. \end{array}$$

Il est facile de voir que $v(\mathcal{R}') = v(\mathcal{R}'')$, et s'il existe $S \in \mathcal{S}$ tel que \bar{w} appartienne à l'intérieur de S , alors

$$(4.5) \quad \text{Argmax}(\mathcal{R}') = \Pi_{\mathbb{R}_+^* \times \mathcal{S} \rightarrow \mathcal{S}}(\text{Argmax}(\mathcal{R}'')).$$

Un bref historique. A notre connaissance, la notion du rayon de stabilité est apparue dans les travaux scientifiques à partir des années 1960 ([21],[52],[74], ...). Deux décennies plus tard, dans les années 1980, elle est utilisée dans la théorie du contrôle ([44]) puis en optimisation ([76]). Une grande partie de l'approche d'un problème de décision à données incertaines par la méthode du rayon de stabilité a été valorisée par Yakov Ben-Haïm sous l'appellation "info-gap decision theory" dans les années 1990 début 2000 ([6]). Il faut noter cependant que ce dernier n'a pas suffisamment fait de connexion avec les travaux antérieurs concernant le rayon de stabilité. La connexion entre les travaux de Yakov Ben-Haïm et les travaux antérieurs a été établie partiellement par M. Sneedovich dans ([66]) et repris dans un cadre beaucoup plus large par M. Ciligot-Travain ([21]).

4.1.4.2. Une généralisation du rayon de stabilité

Considérons maintenant que l'on se donne $U :]0, +\infty[\rightrightarrows W$, alors on étend de façon naturelle la définition du rayon de stabilité de la façon suivante :

Définition 4.1.2 (Rayon de stabilité généralisé). *On appelle **rayon de stabilité de $S \subset W$** (relativement à U), l'élément de $\overline{\mathbb{R}}_+$ défini par :*

$$r(S, w, d) := \sup\{\rho > 0 \mid U(\rho) \subset S\}.$$

Le cas classique correspond au choix $U(\rho) = \mathbb{B}(\bar{w}, \rho)$.

4.1.4.3. Paramétrisation

On reprend le modèle paramétré. Pour tout $x \in D$, on se donne une distance $d_x : W_x \rightarrow \mathbb{R}_+$ et une incertitude nominale $\bar{w}_x \in W_x$. Le rayon de robustesse de x est donc défini par

$$r(S_x, \bar{w}_x, d_x) = d_x(\bar{w}_x, W_x \setminus S_x).$$

On peut étendre sans peine cette paramétrisation au modèle général de la section 4.1.4.2.

4.1.5. Rayon de robustesse signé

Soit (W, d) un espace métrique. Étant donnée $A \subset W$, on appelle **distance signée à A** et on note $\Delta(\cdot, A) : W \rightarrow \overline{\mathbb{R}}$, la fonction définie par :

$$\Delta(w, A) := d(w, A) - d(w, A^c) := \begin{cases} d(w, A) & \text{si } w \notin A, \\ -d(w, A^c) & \text{si } w \in A. \end{cases}$$

Définition 4.1.3 (Rayon de robustesse signé). *On appelle **rayon de robustesse signé de $S \subset W$** , l'élément de $\overline{\mathbb{R}}$ défini, pour $w \in W$, par*

$$\varrho(S, w, d) := \Delta(w, S^c) = -\Delta(w, S).$$

On notera éventuellement $\varrho(S, \bar{w}, d)$ par $\varrho(S, \bar{w})$, $\varrho(S)$, etc, sous-entendant éventuellement des éléments définis par le contexte.

Dans le cas du modèle paramétré, le rayon de robustesse signé se définit naturellement par

$$\varrho(S_x, w_x, d_x) := \Delta_x(w_x, W_x \setminus S_x) = -\Delta_x(w_x, S_x),$$

où $\Delta_x(\cdot, A)$ est la distance signée à $A \subset W_x$ relativement à la métrique d_x .

Remarque 4.1.1. *Le rayon de robustesse de S est la partie positive du rayon de robustesse signé de S : $r(S) = \varrho(S)^+ := \max(0, \varrho(S))$.*

La remarque 4.1.1 fournit une connexion non négligeable entre le rayon de stabilité et le rayon de stabilité signé. Par exemple s'il existe $S \in \mathcal{S}$ tel que $\varrho(S) > 0$ alors la maximisation de r sur \mathcal{S} est équivalente (même valeur et même ensemble de solutions optimales) à la maximisation de ϱ sur \mathcal{S} .

4.1.6. Contrepartie robuste au sens du rayon de stabilité

4.1.6.1. Problème d'optimisation paramétré par un paramètre incertain

On considère le problème d'optimisation paramétré suivant :

$$(\mathcal{P}) \quad \text{Minimiser } f(x, w), \quad x \in X,$$

où $w \in W$ est un paramètre «incertain», X est l'ensemble des décisions et $f : X \times W \rightarrow \overline{\mathbb{R}}$ est une fonction.

Rappelons que dans le cas qui nous intéresse, le paramètre incertain w n'est observé qu'après une prise de décision x et, a priori, on ne connaît aucune loi de probabilité régissant cette incertitude.

Il est assez naturel d'autoriser W à dépendre de x , ainsi que w . Ceci nous conduit à la généralisation qui suit.

$$(\mathcal{P}) \quad \text{Minimiser } f(x, w_x), \quad x \in X,$$

où le paramètre incertain $w_x \in W_x$. Pour ce modèle, à l'image du modèle paramétré, on pose

$$W = \bigcup_{x \in X} W_x.$$

4.1.6.2. Optimisation robuste

Dans la méthode du «pire des cas», on suppose que l'on a quand même une information sur l'incertitude. Information du type $w \in U$ où $U \subset W$ et l'on est amené à associer au problème (\mathcal{P}) , le problème

$$(\mathcal{R}_U) \quad \text{Minimiser } \sup_{w \in U} f(x, w), \quad x \in D,$$

appelé *contrepartie robuste* du problème d'optimisation à données incertaines (\mathcal{P}) . Les origines de cette approche remontent aux travaux de A. L. Soyster ([67]). Une solution «intéressante» de (\mathcal{R}_U) vérifie

$$\sup_{w \in U} f(x, w) < +\infty$$

et donc, en particulier, $\forall w \in U, x \in [f(\cdot, w) < +\infty]$.

En général, dans les applications, soit on suppose que U est l'enveloppe convexe d'un nombre fini de points, soit on se donne une incertitude nominale \bar{w} et on suppose que l'ensemble U est une boule de centre \bar{w} et de rayon ρ de l'espace métrique (W, d) . Cette approche a été mise en valeur par A. Ben-Tal et ses collaborateurs. Ces auteurs ont résumé une grande partie de leurs travaux dans le livre "robust optimization" [7].

4.1.6.3. Optimisation de la robustesse

Nous présentons une autre façon peu connue d'approcher le problème (\mathcal{P}) . On suppose que W est muni d'une métrique d . Pour ce faire, on se donne une incertitude nominale $\bar{w} \in W$ et un seuil de tolérance $\alpha \in \mathbb{R}$ tels que

$$\inf_{x \in X} f(x, \bar{w}) < \alpha.$$

Définition 4.1.4. On appelle *contrepartie robuste* du problème d'optimisation paramétré par un paramètre incertain, (\mathcal{P}) , $w \in W$, au sens du rayon de stabilité, le

problème de maximisation suivant :

$$(\mathcal{R}^\alpha) \quad \text{Maximiser } r([f(x, \cdot) > \alpha], \bar{w}, d) := d(\bar{w}, [f(x, \cdot) > \alpha]), \quad x \in X.$$

Dans la suite on parlera simplement de contrepartie robuste en omettant le terme «au sens du rayon de stabilité» s'il n'y a pas d'ambiguïté avec la contrepartie robuste au sens du «pire des cas».

Nous allons établir au chapitre 4.3 qu'il existe des relations entre les deux contreparties.

4.2. Premières propriétés du rayon de stabilité

Dans cette section, on étudie certaines propriétés du rayon de stabilité et du rayon de stabilité signé pouvant servir à l'utilisation des arguments d'analyse convexe dans le problème de maximisation du rayon de robustesse. On reprend le modèle paramétrique avec un paramètre x appartenant à un ensemble X .

On suppose ici que l'ensemble d'incertitude W , la distance définie sur cet ensemble, d , et la valeur nominale \bar{w} ne dépendent pas du paramètre ou variable de décision x .

Soit S un sous-ensemble de $X \times W$. On souhaite que le couple (x, w) appartienne à S où w est l'incertitude observée suite à la décision x . On peut voir S comme définissant une contrainte ou un ensemble de satisfaction. Pour chaque décision $x \in X$, on définit alors l'ensemble des incertitudes $w \in W$ pour lesquelles les couples (x, w) sont jugés satisfaisants c'est-à-dire $(x, w) \in S$. Cet ensemble est noté $S(x)$ et défini par

$$S(x) = \{w \in W \mid (x, w) \in S\}.$$

De même pour chaque incertitude $w \in W$, on note $S^-(w)$, l'ensemble des décisions $x \in X$ pour lesquelles les couples (x, w) sont jugés satisfaisants :

$$S^-(w) = \{x \in X \mid (x, w) \in S\}.$$

Le rayon de stabilité défini à partir de S , noté ici r_S , est la fonction de $X \times W$ à valeurs dans $\overline{\mathbb{R}}_+$ donnée, pour $x \in X$, $\bar{w} \in W$, par :

$$(4.6) \quad r_S(x, \bar{w}) := r(S(x), \bar{w}, d) = \begin{cases} d(\bar{w}, S(x)^c), \\ \text{ou} \\ \sup\{\rho > 0 \mid \mathbb{B}(\bar{w}, \rho) \subset S(x)\}. \end{cases}$$

4.2.1. Semi-continuité supérieure

On rappelle que la fonction indicatrice d'un sous ensemble $A \subset W$ est la fonction $i_A : W \rightarrow \overline{\mathbb{R}}$ définie par :

$$i_A(w) = \begin{cases} 0 & \text{si } w \in A, \\ +\infty & \text{sinon.} \end{cases}$$

Proposition 4.2.1. *Si X est un espace topologique et si pour tout $w \in W$, $S^-(w)$ est fermé alors la fonction $r_S(\cdot, \bar{w}) : X \rightarrow [0, +\infty]$ est semi-continue supérieurement pour tout $\bar{w} \in W$.*

Preuve. Soit $\bar{w} \in W$. Pour tout $x \in X$, on a

$$(4.7) \quad r_S(x, \bar{w}) = \inf_{v \in W} \left\{ d(\bar{w}, v) + i_{S^-(v)^c}(x) \right\}.$$

Par ailleurs, pour tout $v \in W$,

$$[i_{S^-(v)^c} \geq t] = \begin{cases} S^-(v) & \text{si } t > 0, \\ X & \text{si } t \leq 0. \end{cases}$$

On en déduit alors que la fonction $d(\bar{w}, v) + i_{S^-(v)^c}(\cdot)$ est semi-continue supérieurement. Donc, d'après (4.7), la fonction $r_S(\cdot, \bar{w})$ est semi-continue supérieurement en tant que borne inférieure d'une famille de telles fonctions. \square

4.2.2. Quasi-concavité

On suppose ici que X est une partie d'un espace vectoriel.

Proposition 4.2.2. *Pour tout $(x, \bar{w}) \in X \times W$, on a $r_S(x, \bar{w}) > 0$ si et seulement si \bar{w} est un point intérieur de $S(x)$.*

Preuve. w est un point intérieur de $S(x)$ si et seulement s'il existe $\rho_w > 0$ tel que $\mathbb{B}(w, \rho_w) \subset S(x)$. Ce qui est équivalent à

$$r_S(x, w) = \sup\{\rho > 0 \mid \mathbb{B}(w, \rho) \subset S(x)\} \geq \rho_w > 0.$$

\square

Proposition 4.2.3. *Si X est convexe et si $S^-(w)$ est convexe pour tout $w \in W$ alors l'application $r_S(\cdot, \bar{w}) : X \rightarrow [0, +\infty]$ est quasi-concave pour tout $\bar{w} \in W$.*

Preuve. Soit $\bar{w} \in W$. Pour tout $x \in X$, on a :

$$r_S(x, \bar{w}) = \inf_{v \in W} \left\{ d(\bar{w}, v) + i_{S^-(v)^c}(x) \right\}.$$

Par ailleurs, pour tout $v \in W$, on a

$$[i_{S^-(v)^c} \geq t] = \begin{cases} S^-(v) & \text{si } t > 0, \\ X & \text{si } t \leq 0. \end{cases}$$

Les tranches supérieures de la fonction $i_{S^-(v)^c}$ sont donc convexes. Par conséquent, la fonction $d(\bar{w}, v) + i_{S^-(v)^c}(\cdot)$ est quasi-concave sur X pour tout v dans W . On en déduit que la fonction $r_S(\cdot, \bar{w}) : X \rightarrow [0, +\infty]$ est quasi-concave en tant que borne inférieure d'une famille de fonctions quasi-concaves. \square

Corollaire 4.2.1. *On reprend les hypothèses de la proposition 4.2.3. On suppose en plus que X est un sous-ensemble d'un espace vectoriel normé D . Alors tout maximum local strict de $r_S(\cdot, \bar{w})$ sur X est un maximum global de $r_S(\cdot, \bar{w})$ sur X .*

Preuve. Soit \bar{x} un maximum local strict de $r_S(\cdot, \bar{w})$ sur X . Il existe Ω un voisinage de \bar{x} dans D tel que

$$r_S(\bar{x}, \bar{w}) > r_S(x, \bar{w}), \quad \forall x \in \Omega \cap X, \quad x \neq \bar{x}.$$

Supposons que \bar{x} n'est pas un maximum global de $r_S(\cdot, \bar{w})$ sur X . Il existe donc $x \in X$ tel que $r_S(x, \bar{w}) > r_S(\bar{x}, \bar{w})$. Puisque $r_S(\cdot, \bar{w})$ est quasi-concave d'après la proposition 4.2.3, alors on a :

$$r_S(\lambda x + (1 - \lambda)\bar{x}, \bar{w}) \geq \min \left\{ r_S(x, \bar{w}), r_S(\bar{x}, \bar{w}) \right\} = r_S(\bar{x}, \bar{w}), \quad \forall \lambda \in]0, 1[.$$

En particulier, pour λ assez petit, on a $\lambda x + (1 - \lambda)\bar{x} \in \Omega \cap X$ et $\lambda x + (1 - \lambda)\bar{x} \neq \bar{x}$, d'où la contradiction. \square

4.2.3. Concavité

Soit $(V, \|\cdot\|)$ un espace vectoriel normé de dual topologique V^* . On note $\langle \cdot, \cdot \rangle$ le crochet de dualité entre V et V^* . Pour tout $(v, v^*) \in V \times V^*$, on note $\langle v, v^* \rangle := v^*(v)$.

Rappelons que la *conjuguée de Legendre-Fenchel* d'une fonction $f : V \rightarrow \overline{\mathbb{R}}$ est la fonction $f^* : V^* \rightarrow \overline{\mathbb{R}}$ définie par :

$$f^*(v^*) = \sup_{v \in V} \left\{ \langle v, v^* \rangle - f(v) \right\}.$$

Si $C \subset V$, la *fonction support* de C , notée $\sigma_C : V^* \rightarrow \overline{\mathbb{R}}$, est définie par :

$$\sigma_C(v^*) := i_C^*(v^*) = \sup_{v \in C} \langle v, v^* \rangle.$$

La norme duale de la norme $\|\cdot\|$, notée $\|\cdot\|^*$, est définie sur V^* par :

$$\|v^*\|^* = \sup_{\|v\| \leq 1} \langle v, v^* \rangle,$$

autrement dit $\|\cdot\|^* = \sigma_B = i_B^*$.

Théorème 4.2.1. *Soit d la distance associée à la norme $\|\cdot\|$ et C un sous-ensemble convexe fermé. Alors pour tout $v \in V$, on a :*

$$(4.8) \quad \Delta(v, C) = \sup_{\|v^*\|^* = 1} \left\{ \langle v, v^* \rangle - \sigma_C(v^*) \right\}$$

Preuve. Remarquons tout d'abord que si C est vide alors l'égalité est évidente. En effet, si $C = \emptyset$ alors on a d'une part

$$\Delta(v, C) = d(v, \emptyset) = +\infty, \quad \forall v \in V,$$

et d'autre part, $i_C(v) = i_\emptyset(v) = +\infty$ pour tout $v \in V$. Par conséquent $\sigma_C(v^*) = -\infty$ pour tout $v^* \in V^*$.

On suppose donc C non vide, alors par le théorème de séparation de Hahn-Banach, on a :

$$C = \bigcap_{\|v^*\|^*=1} [\langle \cdot, v^* \rangle \leq i_C^*(v^*)].$$

En effet, par définition de σ_C , pour tout $v^* \in V^*$ et $v \in C$, on a $\langle v, v^* \rangle \leq \sigma_C(v^*)$. D'où

$$C \subset \bigcap_{\|v^*\|^*=1} [\langle \cdot, v^* \rangle \leq \sigma_C(v^*)].$$

Réciproquement, soit $v \notin C$. D'après une version du théorème de séparation de Hahn-Banach (voir par exemple [16], II-§5 proposition 4) il existe $v^* \in V^*$, tel que

$$\|v^*\|^* = 1 \quad \text{et} \quad \langle v, v^* \rangle > \sigma_C(v^*).$$

Par conséquent, $v \notin \bigcap_{\|v^*\|^*=1} [\langle \cdot, v^* \rangle \leq \sigma_C(v^*)]$. D'où l'égalité.

Soit $v \in V$. On distingue deux cas :

(a) On suppose que $v \in C$. Dans ce cas, $\Delta(v, C) = -d(v, C^c)$ et on a :

$$\begin{aligned} d(v, C^c) &= d\left(v, \bigcup_{\|v^*\|^*=1} [\langle \cdot, v^* \rangle > \sigma_C(v^*)]\right) \\ &= \inf_{\|v^*\|^*=1} d(v, [\langle \cdot, v^* \rangle > \sigma_C(v^*)]) \\ &= \inf_{\|v^*\|^*=1} (\sigma_C(v^*) - \langle v, v^* \rangle)^+ \quad (\text{formule d'Ascoli : Lemme 1.5.2}) \\ &= \inf_{\|v^*\|^*=1} \{\sigma_C(v^*) - \langle v, v^* \rangle\} \quad \text{car } v \in C. \end{aligned}$$

$$\text{Donc } -d(v, C^c) = \sup_{\|v^*\|^*=1} \{\langle v, v^* \rangle - \sigma_C(v^*)\} = \Delta(v, C).$$

(b) On suppose que $v \notin C$. Dans ce cas, $\Delta(v, C) = d(v, C)$ et on a :

$$\begin{aligned} \Delta(v, C) = d(v, C) &= \inf_{z \in C} \|v - z\| \\ &= \inf_{z \in C} \sup_{\|v^*\|^*=1} \langle v - z, v^* \rangle \\ &\geq \sup_{\|v^*\|^*=1} \left\{ \langle v, v^* \rangle - \sigma_C(v^*) \right\}. \end{aligned}$$

Puisque C est un fermé non vide et $v \notin C$ alors $d(v, C) \in]0, +\infty[$. Ainsi, pour tout $\epsilon \in]0, d(v, C)[$, la boule ouverte de centre v et de rayon $d(v, C) - \epsilon$, $\mathbb{B}_o(v, d(v, C) - \epsilon)$ ne rencontre pas C . Par conséquent, d'après une version du théorème de séparation de Hahn-Banach (voir par exemple [16], II-§5 proposition 1), il existe $z^* \in V^*$ tel que $\|z^*\|^* = 1$ et $\mathbb{B}_o(v, d(v, C) - \epsilon) \subset [\langle \cdot, z^* \rangle \geq \sigma_C(z^*)]$. On a

$$\sup_{\|v^*\|^*=1} \{\langle v, v^* \rangle - \sigma_C(v^*)\} \geq \langle v, z^* \rangle - \sigma_C(z^*) \geq 0.$$

Donc, la formule d'Ascoli (lemme 1.5.2) donne

$$\langle v, z^* \rangle - \sigma_C(z^*) = \left(\frac{\langle v, z^* \rangle - \sigma_C(z^*)}{\|z^*\|^*} \right)^+ = d(v, [\langle \cdot, z^* \rangle < \sigma_C(z^*)]).$$

Comme $\mathbb{B}_o(v, d(v, C) - \epsilon) \cap [\langle \cdot, z^* \rangle < \sigma_C(z^*)] = \emptyset$ alors

$$d(v, [\langle \cdot, z^* \rangle < \sigma_C(z^*)]) \geq d(v, C) - \epsilon.$$

Il en résulte que :

$$\sup_{\|v^*\|^*=1} \{\langle v, v^* \rangle - \sigma_C(v^*)\} \geq d(v, C) - \epsilon, \quad \forall \epsilon \in]0, d(v, C)[.$$

En faisant tendre ϵ vers 0, on obtient l'inégalité inverse

$$\Delta(v, C) \leq \sup_{\|v^*\|^*=1} \{\langle v, v^* \rangle - \sigma_C(v^*)\}.$$

□

Corollaire 4.2.2. *Supposons que X est un convexe d'un espace vectoriel topologique D , W est un convexe de V et S est un sous-ensemble convexe fermé de $D \times V$ tel que $S(x)$ soit fermé pour tout $x \in X$. Alors $\varrho_S : X \times W \rightarrow \overline{\mathbb{R}}$ est une fonction concave.*

Preuve. Soit $(x, w) \in X \times W$. On a :

$$\begin{aligned} \varrho_S(x, w) &:= d(w, S(x)^c) - d(w, S(x)) \\ &= -\Delta(w, S(x)) \\ &= -\sup_{\|v^*\|^*=1} \{\langle w, v^* \rangle - \sigma_{S(x)}(v^*)\} \\ &= -\sup_{\|v^*\|^*=1} \left\{ \langle w, v^* \rangle - \sup_{v \in S(x)} \langle v, v^* \rangle \right\} \\ &= -\sup_{\|v^*\|^*=1} \left\{ \inf_{v \in S(x)} \langle w - v, v^* \rangle \right\} \\ &= -\sup_{\|v^*\|^*=1} \left\{ \inf_{v \in V} \left\{ \langle w - v, v^* \rangle + i_{S(x)}(v) \right\} \right\}. \end{aligned}$$

Finalement

$$\varrho_S(x, w) = -\sup_{\|v^*\|^*=1} \left\{ \inf_{v \in V} \left\{ \langle w - v, v^* \rangle + i_S(x, v) \right\} \right\}.$$

Pour tout $v^* \in V^*$, la fonction $(x, w, v) \mapsto \langle w - v, v^* \rangle + i_S(x, v)$ est convexe. Donc sa fonction marginale $(x, w) \mapsto \inf_{v \in V} \{\langle w - v, v^* \rangle + i_S(x, v)\}$ est aussi convexe. Il en résulte que $(x, w) \mapsto \varrho_S(x, w)$ est concave. □

Corollaire 4.2.3. *Sous les hypothèses du corollaire 4.2.2, $r_S : X \times W \rightarrow [0, +\infty]$ est une fonction quasi-concave.*

Preuve. Puisque $r_S = \varrho_S^+$ alors pour tout $t \in \mathbb{R}$, on a :

$$[r_S \geq t] = [\varrho_S^+ \geq t] = \begin{cases} [\varrho_S \geq t] & \text{si } t \geq 0 \\ X & \text{sinon.} \end{cases}$$

D'où r_S est quasi-concave car ϱ_S est concave. \square

4.3. Optimisation robuste et Optimisation de la robustesse

Dans cette section, on établit une relation entre l'ensemble des solutions optimales de la contrepartie robuste (méthode du pire des cas) d'un problème d'optimisation à données incertaines et l'ensemble des solutions de sa contrepartie robuste au sens du rayon de robustesse (méthode du rayon de stabilité).

On considère le problème d'optimisation paramétré par un paramètre « incertain » suivant

$$(\mathcal{P}_w) \quad \text{Minimiser } f(x, w), \quad x \in X,$$

où X est un ensemble non vide, w désigne le paramètre incertain et appartient à W un ensemble non vide, et $f : X \times W \rightarrow \overline{\mathbb{R}}$ ou $\overline{\mathbb{R}}_+$, est une fonction.

Pour tout $x \in X$, on se donne une métrique d_x sur W et une incertitude nominale $\bar{w}_x \in W$.

Méthode de l'optimisation robuste : On définit la fonction "pire des cas", notée $wc : X \times \overline{\mathbb{R}}_+ \rightarrow \overline{\mathbb{R}}$ ou $\overline{\mathbb{R}}_+$, par

$$(4.9) \quad \forall (x, \alpha) \in X \times \overline{\mathbb{R}}_+, \quad wc(x, \alpha) := \sup \left\{ f(x, w) \mid w \in [d_x(\cdot, \bar{w}_x) < \alpha] \right\}.$$

La contrepartie robuste du problème paramétré incertain (\mathcal{P}_w) , $w \in W$, au sens de l'optimisation robuste est définie par :

$$(\mathcal{W}_\alpha) \quad \text{Minimiser } wc(x, \alpha), \quad x \in X.$$

Moyennant une hypothèse assez faible, on a

$$(4.10) \quad \forall (x, \alpha) \in X \times \overline{\mathbb{R}}_+, \quad wc(x, \alpha) := \sup \left\{ f(x, w) \mid w \in [d_x(\cdot, \bar{w}_x) \leq \alpha] \right\}.$$

ce qui correspond mieux à ce qui est considéré habituellement en optimisation robuste.

Les origines de cette approche remontent dans les années 1970 suite aux travaux de A. L. Soyster ([67]). Elle a été mise en valeur par A. Ben-Tal et ses collaborateurs ([7]).

Méthode de l'optimisation de la robustesse : On définit le rayon de stabilité

$r : X \times \overline{\mathbb{R}} \text{ ou } \overline{\mathbb{R}}_+ \rightarrow \overline{\mathbb{R}}_+$ par

(4.11)

$$\forall (x, \beta) \in X \times \overline{\mathbb{R}} \text{ ou } \overline{\mathbb{R}}_+, \quad r(x, \beta) := \inf \left\{ d_x(\bar{w}_x, w) \mid f(x, w) > \beta \right\} = d_x(\bar{w}_x, [f(x, \cdot) > \beta]).$$

La contrepartie robuste du problème paramétré incertain (\mathcal{P}_w) , $w \in W$, au sens de l'optimisation de la robustesse est le problème de la maximisation du rayon de stabilité :

$$(\mathcal{R}_\beta) \quad \text{Maximiser } r(x, \beta), \quad x \in X.$$

L'objectif de cette section est de chercher des relations entre l'ensemble des solutions optimales de (\mathcal{W}_α) et celui de (\mathcal{R}_β) pour des valeurs de α et β convenablement choisies et fixées.

On note $v_{wc}(\alpha)$ la valeur optimale du problème (\mathcal{W}_α) et $v_r(\beta)$ celle du problème (\mathcal{R}_β) . On obtient donc deux fonctions croissantes $v_{wc} : I \rightarrow J$ et $v_r : J \rightarrow I$, avec $I = \overline{\mathbb{R}}_+$ et $J = \overline{\mathbb{R}} \text{ ou } \overline{\mathbb{R}}_+$ définies comme suit :

$$v_{wc}(\alpha) := \inf_{x \in X} wc(x, \alpha) \quad \text{et} \quad v_r(\beta) := \sup_{x \in X} r(x, \beta).$$

Ce qui permet d'obtenir des relations entre l'ensemble des solutions optimales de (\mathcal{W}_α) et celui de (\mathcal{R}_β) tient pour l'essentiel dans la propriété suivante

$$(4.12) \quad \forall (\alpha, \beta, x) \in I \times J \times X, \quad wc(x, \alpha) \leq \beta \iff \alpha \leq r(x, \beta).$$

On détaille un peu plus ci-dessous suivant l'adoption de la définition (4.10) ou (4.9). En effet, pour tout $(\alpha, \beta, x) \in I \times J \times X$, on a :

$$(4.13a) \quad wc(x, \alpha) \leq \beta \iff \sup\{f(x, w) \mid w \in [d_x(\bar{w}_x, \cdot) < \alpha]\} \leq \beta,$$

$$(4.13b) \quad \iff [d_x(\bar{w}_x, \cdot) < \alpha] \subset [f(x, \cdot) \leq \beta],$$

$$(4.13c) \quad \iff [f(x, \cdot) > \beta] \subset [d_x(\bar{w}_x, \cdot) \geq \alpha],$$

$$(4.13d) \quad \iff \alpha \leq \inf\{d_x(\bar{w}_x, w) \mid w \in [f(x, \cdot) > \beta]\},$$

$$(4.13e) \quad \iff \alpha \leq r(x, \beta),$$

et également

$$(4.14a) \quad wc(x, \alpha) \leq \beta \iff \sup\{f(x, w) \mid w \in [d_x(\bar{w}_x, \cdot) \leq \alpha]\} \leq \beta,$$

$$(4.14b) \quad \iff [d_x(\bar{w}_x, \cdot) \leq \alpha] \subset [f(x, \cdot) \leq \beta],$$

$$(4.14c) \quad \iff [f(x, \cdot) > \beta] \subset [d_x(\bar{w}_x, \cdot) > \alpha],$$

$$(4.14d) \quad \implies \alpha \leq \inf\{d_x(\bar{w}_x, w) \mid w \in [f(x, \cdot) > \beta]\}.$$

$$(4.14e) \quad \iff \alpha \leq r(x, \beta).$$

Donc la propriété ci-dessous est toujours vraie, que l'on adopte la définition (4.10) ou (4.9).

$$(4.15) \quad \forall (\alpha, \beta, x) \in I \times J \times X, \quad wc(x, \alpha) \leq \beta \implies \alpha \leq r(x, \beta).$$

Si on adopte la définition (4.9), les équivalences (4.13) montrent directement l'équivalence (4.12). Si on adopte la définition (4.10), les implications (4.14) montrent seulement une implication (4.12). Si on suppose :

$$(4.16) \quad \forall x \in X, \quad f(x, \cdot) \text{ est semicontinue inférieurement}$$

alors

$$(4.17a) \quad \alpha \leq r(x, \beta) \iff \alpha \leq \inf\{d_x(\bar{w}_x, w) \mid w \in [f(x, \cdot) > \beta]\},$$

$$(4.17b) \quad \iff [f(x, \cdot) > \beta] \subset [d_x(\bar{w}_x, \cdot) \geq \alpha],$$

$$(4.17c) \quad \iff [d_x(\bar{w}_x, \cdot) < \alpha] \subset [f(x, \cdot) \leq \beta],$$

$$(4.17d) \quad \implies [d_x(\bar{w}_x, \cdot) \leq \alpha] \subset [f(x, \cdot) \leq \beta],$$

$$(4.17e) \quad \iff wc(x, \alpha) \leq \beta,$$

et on obtient l'équivalence (4.12).

Cette section est organisée comme suit : Dans la sous-section 4.3.1, on établit quelques propriétés des inverses généralisées inférieure et supérieure d'une fonction croissante. Ces propriétés sont appliquées à un couple de problèmes paramétrés dans la sous-section 4.3.2. La sous-section 4.3.3 termine la section en proposant des comparaisons entre l'ensemble des solutions de (\mathcal{W}_α) et celui de (\mathcal{R}_β) pour des valeurs de α et β bien choisies.

4.3.1. Inverses généralisées d'une fonction croissante

La notion d'inverse généralisée se rencontre dans divers travaux ([33], [35], [37],[48], [69], . . .).

Soient I et J deux intervalles fermés de $\overline{\mathbb{R}}$. On note $\inf I$ (resp. $\sup I$) la borne inférieure (resp. supérieure) de I , puis on définit les intervalles $I^<$, $I^>$ et $\overset{\circ}{I}$ par

$$(4.18) \quad I^< := I \setminus \{\inf I\} \quad , \quad I^> := I \setminus \{\sup I\} \quad \text{et} \quad \overset{\circ}{I} = I \setminus \{\inf I, \sup I\}.$$

On a

$$(4.19) \quad \inf_{I^<} \emptyset = \sup I \quad \text{et} \quad \sup_{I^>} \emptyset = \inf I.$$

Soit $h : I \rightarrow J$ une fonction croissante.

Définition 4.3.1 ([35],[37]). *On appelle inverse généralisée supérieure de h , la fonction $h^+ : J \rightarrow I$ définie par :*

$$\forall s \in J, \quad h^+(s) := \sup [h \leq s] = \inf [s < h].$$

Montrons la deuxième égalité.

Si $s \in J$, $t_1, t_2 \in I$ tels que $t_1 \in [h \leq s]$, $t_2 \in [s < h]$, $h(t_1) < h(t_2)$, et comme h est croissante, $t_1 < t_2$. Ceci montre que $\sup [h \leq s] \leq \inf [s < h]$.

Si $\sup [h \leq s] = \sup I$, nécessairement $\inf [s < h] = \sup I = \sup [h \leq s]$.

Si $\sup [h \leq s] < \sup I$, si $\sup [h \leq s] < \inf [s < h]$, il existe $t' \in I$ tel que

$$\sup [h \leq s] < t' < \inf [s < h].$$

On a donc $h(t') > s$ d'où $\inf [s < h] < t'$ et une contradiction. D'où $\sup [h \leq s] \geq \inf [s < h]$ et on a l'égalité cherchée.

Définition 4.3.2 ([33],[35], [37]). *On appelle inverse généralisée inférieure de h , la fonction $h^- : J \rightarrow I$ définie par :*

$$h^-(s) := \inf [s \leq h] = \sup [h < s].$$

Par définition, les fonctions h^- et h^+ sont croissantes et pour tout $(t, s) \in I \times J$, on a

$$(4.20) \quad h(t) \leq s \implies t \leq h^+(s), \quad s < h(t) \implies h^+(s) \leq t,$$

$$(4.21) \quad s \leq h(t) \implies h^-(s) \leq t, \quad h(t) < s \implies t \leq h^-(s).$$

Les propriétés (4.20) et (4.21) caractérisent $h^+(s)$ et $h^-(s)$ dans le sens suivant : si $\alpha \in I$, $s \in J$

$$(4.22) \quad \left[\forall t \in I, \quad h(t) \leq s \implies t \leq \alpha \text{ et } s < h(t) \implies \alpha \leq t \right] \implies \alpha = h^+(s).$$

$$(4.23) \quad \left[\forall t \in I, \quad s \leq h(t) \implies \alpha \leq t \text{ et } h(t) < s \implies t \leq \alpha \right] \implies \alpha = h^-(s).$$

En prenant $s = h(t)$ dans (4.21) et (4.20), on obtient

$$(4.24) \quad h^-(h(t)) \leq t \leq h^+(h(t)), \quad \forall t \in I.$$

Par ailleurs, on a toujours

$$(4.25) \quad \forall s \in J, \quad]h^-(s), h^+(s)[\subset [h = s] \subset [h^-(s), h^+(s)].$$

Proposition 4.3.1. *h^+ est continue à droite et h^- est continue à gauche.*

Preuve. Montrons la continuité à droite de h^+ . Soit $s_0 \in J$, $s_0 < \sup J$. Si $h^+(s_0) = \sup I$ alors h^+ est constante sur $[s_0; +\infty] \cap J^>$, donc continue à droite en s_0 . Supposons que $h^+(s_0) < \sup I$. Alors pour tout $t \in I$ tel que $t > h^+(s_0)$, on a $h(t) > s_0$ et il existe $s \in J$ tel que $h(t) > s > s_0$ donc $t \geq h^+(s) \geq h^+(s_0^+)$. Ceci implique $h^+(s_0) \geq h^+(s_0^+) \geq h^+(s_0)$ (prendre $t = h^+(s_0) + \epsilon$ avec ϵ assez petit puis faire tendre ϵ vers 0). D'où la continuité de h^+ à droite. La continuité à gauche de h^- s'obtient par un raisonnement analogue. \square

Proposition 4.3.2. *Soit $\varphi : I \rightarrow J$ et $\psi : J \rightarrow I$ deux fonctions croissantes. Les assertions suivantes sont équivalentes :*

(1) $\varphi^+ = \psi$ et $\psi^- = \varphi$.

(2) $\forall (t, s) \in I \times J, \quad \varphi(t) \leq s \iff t \leq \psi(s)$.

Preuve.

Montrons (1) \implies (2). Soit $(t, s) \in I \times J$. On a :

$$\begin{aligned} \varphi(t) \leq s &\implies t \leq \varphi^+(s) \quad \text{d'après (4.20)} \\ &\implies t \leq \psi(s) \quad \text{car } \varphi^+ = \psi \text{ d'après (1)} \\ &\implies \psi^-(t) \leq s \quad \text{d'après (4.21)} \\ &\implies \varphi(t) \leq s \quad \text{car } \psi^- = \varphi \text{ d'après (1)}. \end{aligned}$$

Donc $\varphi(t) \leq s \implies t \leq \psi(s) \implies \varphi(t) \leq s$, d'où (2).

Montrons (2) \implies (1). Soit $s \in J$. On a :

$$\begin{aligned} \varphi^+(s) &:= \sup\{t \in I \mid \varphi(t) \leq s\} \\ &= \sup\{t \in I \mid t \leq \psi(s)\} \quad \text{d'après (2)} \\ &= \psi(s) \quad \text{car } \psi(s) \in I. \end{aligned}$$

Soit $t \in I$. On a :

$$\begin{aligned} \psi^-(t) &:= \inf\{s \in J \mid t \leq \psi(s)\} \\ &= \inf\{s \in J \mid \varphi(t) \leq s\} \quad \text{d'après (2)} \\ &= \varphi(t) \quad \text{car } \varphi(t) \in J. \end{aligned}$$

□

En appliquant la proposition 4.3.2 avec $\varphi = h$, on obtient

$$(4.26) \quad h = (h^+)^- \iff [\forall (t, s) \in I \times J, \quad h(t) \leq s \iff t \leq h^+(s)].$$

Proposition 4.3.3. Soit $h : I \rightarrow J$ une fonction croissante et $s \in J$.

(1) Si $h^-(s) \in I^<$ et h est continue à gauche en $h^-(s)$ alors $h(h^-(s)) \leq s$,

(2) Si $h^+(s) \in I^>$ et h est continue à droite en $h^+(s)$ alors $h(h^+(s)) \geq s$.

Preuve. Si $h^-(s) \in I^<$ alors il existe une suite $(t_n)_{n \in \mathbb{N}}$ croissante d'éléments de I , convergente vers $h^-(s)$ et telle que $t_n < h^-(s)$ pour tout $n \in \mathbb{N}$. On déduit donc de la relation (4.21), que $h(t_n) < s$ pour tout $n \in \mathbb{N}$. La continuité à gauche de h en $h^-(s)$ donne

$$h(h^-(s)) = \lim_{n \rightarrow +\infty} h(t_n) \leq s, \quad \text{d'où (1).}$$

Si $h^+(s) \in I^>$ alors il existe une suite $(t_n)_{n \in \mathbb{N}}$ décroissante d'éléments de I , convergente vers $h^+(s)$ et telle que $h^+(s) < t_n$ pour tout $n \in \mathbb{N}$. On déduit donc de la relation (4.20), que $s < h(t_n)$ pour tout $n \in \mathbb{N}$. Ainsi la continuité à droite de h en $h^+(s)$ nous donne

$$h(h^+(s)) = \lim_{n \rightarrow +\infty} h(t_n) \geq s, \quad \text{d'où (2).}$$

□

4.3.2. Application à un problème d'optimisation paramétré

On suppose maintenant que la fonction h dépend d'un paramètre de décision x appartenant à un ensemble non vide X mais que les intervalles I et J ne dépendent pas de ce paramètre. La fonction h est alors définie sur $X \times I$ à valeurs dans J et pour tout $x \in X$, la fonction $h(x, \cdot)$ est croissante. Pour tout $x \in X$, on note $h^-(x, \cdot) : J \rightarrow I$ (resp. $h^+(x, \cdot) : J \rightarrow I$), l'inverse généralisée inférieure (resp. supérieure) de $h(x, \cdot)$, et $(h^+)^-(x, \cdot) : I \rightarrow J$ l'inverse généralisée inférieure de $h^+(x, \cdot)$. On considère les problèmes d'optimisations (\mathcal{W}_t) et (\mathcal{R}_s) définis respectivement par :

$$(\mathcal{W}_t) \quad \text{Minimiser } h(x, t), \quad x \in X,$$

$$(\mathcal{R}_s) \quad \text{Maximiser } h^+(x, s), \quad x \in X,$$

avec $t \in I$ et $s \in J$.

On note respectivement par $v(t)$ et $v_+(s)$ les valeurs optimales des problèmes (\mathcal{W}_t) et (\mathcal{R}_s) c'est-à-dire

$$(4.27) \quad v(t) := \inf_{x \in X} h(x, t) \quad \text{et} \quad v_+(s) := \sup_{x \in X} h^+(x, s).$$

On définit ainsi deux fonctions croissantes $v : I \rightarrow J$ et $v_+ : J \rightarrow I$ données par (4.27).

On a

$$\text{Argmin}(\mathcal{W}_t) := \{x \in X \mid h(x, t) \leq v(t)\} = [h(\cdot, t) \leq v(t)],$$

$$\text{Argmax}(\mathcal{R}_s) := \{x \in X \mid v_+(s) \leq h^+(x, s)\} = [v_+(s) \leq h^+(\cdot, s)].$$

Proposition 4.3.4. *On a les implications suivantes :*

- (1) Pour tout $t \in I$, $v_+(v(t)) \leq t \implies \text{Argmin}(\mathcal{W}_t) \subset \text{Argmax}(\mathcal{R}_{v(t)})$.
- (2) Pour tout $t \in I$, $v_+(v(t)) > t \implies \text{Argmin}(\mathcal{W}_t) \supset \text{Argmax}(\mathcal{R}_{v(t)})$.
- (3) Pour tout $s \in J$, $v(v_+(s)) \leq s \implies \text{Argmin}(\mathcal{W}_{v_+(s)}) \subset \text{Argmax}(\mathcal{R}_s)$.
- (4) Pour tout $s \in J$, $v(v_+(s)) > s \implies \text{Argmin}(\mathcal{W}_{v_+(s)}) \supset \text{Argmax}(\mathcal{R}_s)$.

Preuve. Soit $t \in I$ tel que $v_+(v(t)) \leq t$. On a

$$\begin{aligned} \text{Argmin}(\mathcal{W}_t) &= [h(\cdot, t) \leq v(t)], \quad \text{d'après la définition de } \text{Argmin}(\mathcal{W}_t) \\ &\subset [t \leq h^+(\cdot, v(t))], \quad \text{d'après la relation (4.20)} \\ &\subset [v_+(v(t)) \leq h^+(\cdot, v(t))], \quad \text{car } v_+(v(t)) \leq t \\ &= \text{Argmax}(\mathcal{R}_{v(t)}), \quad \text{d'après la définition de } \text{Argmax}(\mathcal{R}_{v(t)}), \text{ d'où (1)}. \end{aligned}$$

Soit $t \in I$ tel que $v_+(v(t)) > t$. On a

$$\begin{aligned} \text{Argmin}(\mathcal{W}_t)^c &= [h(., t) > v(t)], \\ &\subset [h^+(., v(t)) \leq t], \quad \text{d'après la relation (4.20)} \\ &\subset [h^+(., v(t)) < v_+(v(t))], \\ &= \text{Argmax}(\mathcal{R}_{v(t)})^c, \quad \text{d'où (2)}. \end{aligned}$$

On procédera de même dans les autres cas. □

Considérons l'hypothèse suivante :

$$(H) \quad \forall x \in X, \quad (h^+)^-(x, .) = h(x, .).$$

Il résulte de l'équivalence (4.26) que si l'hypothèse (H) est vérifiée alors

$$(4.28) \quad \forall (t, s) \in I \times J, \quad [h(., t) \leq s] = [t \leq h^+(., s)].$$

Proposition 4.3.5. *On suppose que (H) est vérifiée.*

$$(1) \text{ Pour tout } t \in I, v_+(v(t)) \geq t \implies \text{Argmin}(\mathcal{W}_t) \supset \text{Argmax}(\mathcal{R}_{v(t)}).$$

$$(2) \text{ Pour tout } s \in J, v(v_+(s)) \geq s \implies \text{Argmin}(\mathcal{W}_{v_+(s)}) \supset \text{Argmax}(\mathcal{R}_s).$$

Preuve. Soit $t \in I$ tel que $v_+(v(t)) \geq t$. D'après (4.28),

$$[h(., t) \leq v(t)] = [t \leq h^+(., v(t))] \supset [v_+(v(t)) \leq h^+(., v(t))], \quad \text{d'où (1)}.$$

Soit $s \in J$ tel que $v(v_+(s)) \geq s$. D'après (4.28),

$$[h(., v_+(s)) \leq v(v_+(s))] \supset [h(., v_+(s)) \leq s] = [v_+(s) \leq h^+(., s)], \quad \text{d'où (2)}. \quad \square$$

Lemme 4.3.1. *Soient $t \in I, s \in J$. Alors*

$$\text{Argmin}(\mathcal{W}_t) \neq \emptyset \implies [v(t) \leq s \implies t \leq v_+(s)],$$

donc, en particulier,

$$\text{Argmin}(\mathcal{W}_t) \neq \emptyset \implies t \leq v_+(v(t)).$$

Si on suppose de plus que (H) est vérifiée alors

$$\text{Argmax}(\mathcal{R}_s) \neq \emptyset \implies [t \leq v_+(s) \implies v(t) \leq s],$$

donc, en particulier,

$$\text{Argmax}(\mathcal{R}_s) \neq \emptyset \implies v(v_+(s)) \leq s.$$

Preuve. Soit $\bar{x} \in \text{Argmin}(\mathcal{W}_t)$, alors $v(t) \leq s \implies h(\bar{x}, t) \leq s \implies t \leq h^+(\bar{x}, s) \leq v_+(s)$.

Soit $\bar{x} \in \text{Argmax}(\mathcal{R}_s)$, alors $t \leq v_+(s) \implies t \leq h^+(\bar{x}, s) \implies v(t) \leq h(\bar{x}, t) \leq s$. □

Lemme 4.3.2. *Soient $t \in I, s \in J$.*

(1) On a

$$s < v(t) \implies v_+(s) \leq t \quad \text{et} \quad v(t) < s \implies t \leq v_+(s).$$

(2) On en déduit

$$t < v_+(s) \implies v(t) \leq s \quad \text{et} \quad v_+(s) < t \implies s \leq v(t).$$

Preuve.

$$\begin{aligned} s < v(t) &\iff \forall x \in X, s < h(x, t), \\ &\implies \forall x \in X, h^+(x, s) \leq t, \\ &\iff v_+(s) \leq t. \\ v(t) < s &\iff \exists x \in X, h(x, t) < s, \\ &\implies \exists x \in X, t \leq h^+(x, s), \\ &\implies t \leq v_+(s). \end{aligned}$$

□

Remarque 4.3.1. Il résulte du lemme 4.3.2 que

$$[v^+ \geq v_+ \text{ et } v^- \leq v_+] \text{ et } [v_+^+ \geq v \text{ et } v_+^- \leq v].$$

Autrement dit

$$(4.29) \quad v^- \leq v_+ \leq v^+ \quad \text{et} \quad v_+^- \leq v \leq v_+^+.$$

Lemme 4.3.3. Soient $t \in I, s \in J$.

(1) Si v_+ est continue à droite en s alors $v(t) \leq s \implies t \leq v_+(s)$ et on a $v_+(s) = v^+(s)$. En particulier, si v_+ est continue à droite en $v(t)$ alors $t \leq v_+(v(t))$.

(2) Suppose que (H) est vérifiée. Si la fonction v est continue à gauche en t alors $t \leq v_+(s) \implies v(t) \leq s$ et on a $v(t) = v_+^-(t)$. En particulier, si v est continue à gauche en $v_+(s)$ alors $v(v_+(s)) \leq s$.

Preuve. $v(t) \leq s \implies t \leq v_+(s)$. Si $s = \sup J$ alors pour tout $x \in X$, on a $h(x, t) \leq s$, donc $t \leq h^+(x, s) \leq v_+(s)$. Supposons que $s \in J^\succ$. Pour tout $s' \in J$ tel que $s' > s$, il existe $x' \in X$ tel que $h(x', t) \leq s'$. Donc $t \leq h^+(x', s') \leq v_+(s')$. Puisque v_+ est continue à droite en s , on en déduit que

$$t \leq \lim_{\substack{s' \rightarrow s \\ s' > s}} v_+(s') = v_+(s).$$

Par ailleurs, par définition de v^+ , on a

$$[v(t) \leq s \implies t \leq v_+(s)] \implies v^+(s) \leq v_+(s).$$

Par conséquent, $v^+(s) = v_+(s)$ d'après (4.29).

$t \leq v_+(s) \implies v(t) \leq s$. Si $t = \inf I$ alors pour tout $x \in X$, on a $t \leq h^+(x, s)$. Donc $(h^+)^-(x, t) \leq s$, et d'après (H), $h(x, t) = (h^+)^-(x, t) \leq s$, d'où $v(t) \leq s$. Supposons

que $t \in I^<$. Alors pour tout $t' \in I$ tel que $t' < t$ il existe $x' \in X$ tel que $t' \leq h^+(x', s)$. Donc $(h^+)^-(x', t') \leq s$, et d'après (H), $h(x', t') = (h^+)^-(x', t') \leq s$, d'où $v(t') \leq s$. Puisque v est continue à gauche en t , on en déduit que

$$v(t) = \lim_{\substack{t' \rightarrow t \\ t' < t}} v(t') \leq s.$$

De même, par définition de v_+ , on a

$$[t \leq v_+(s) \implies v(t) \leq s] \implies v(t) \leq v_+(s).$$

Par conséquent, $v(t) = v_+(s)$ d'après (4.29). \square

Proposition 4.3.6. *Soient $t \in I, s \in J$. On suppose que (H) est vérifiée.*

(1) *Si $\text{Argmin}(\mathcal{W}_t) \neq \emptyset$ ou si v_+ est continue à droite en $v(t)$ alors*

$$\text{Argmin}(\mathcal{W}_t) \supset \text{Argmax}(\mathcal{R}_{v(t)}).$$

(2) *Si $\text{Argmax}(\mathcal{R}_s) \neq \emptyset$ ou si v est continue à gauche en $v_+(s)$ alors*

$$\text{Argmax}(\mathcal{R}_s) \supset \text{Argmin}(\mathcal{W}_{v_+(s)}).$$

Preuve. Montrons (1). Si $\text{Argmin}(\mathcal{W}_t) \neq \emptyset$ alors $t \leq v_+(v(t))$ d'après le lemme 4.3.1. Si v_+ est continue à droite en $v(t)$ alors $t \leq v_+(v(t))$ d'après le lemme 4.3.3. On conclut donc avec la proposition 4.3.5.

Montrons (2). Si $\text{Argmax}(\mathcal{R}_s) \neq \emptyset$ alors $v(v_+(s)) \leq s$ d'après le lemme 4.3.1. Si v est continue à gauche en $v_+(s)$ alors $v(v_+(s)) \leq s$ d'après le lemme 4.3.3. La conclusion résulte de la proposition 4.3.5. \square

Lemme 4.3.4. *Soient $t \in I, s \in J$.*

(1) *Si v_+ est continue à gauche en $s \in J^<$ alors*

$$s \leq v(t) \implies v_+(s) \leq t \quad \text{et on a } v_+(s) = v^-(s).$$

En particulier, si v_+ est continue à gauche en $v(t) \in J^<$, $v_+(v(t)) \leq t$.

(2) *Si v est continue à droite en $t \in I^>$ alors*

$$v_+(s) \leq t \implies s \leq v(t) \quad \text{et on a } v(t) = v_+^+(t).$$

En particulier, si v est continue à droite en $v_+(s) \in I^>$, alors $v(v_+(s)) \geq s$.

Preuve. Montrons (1). Soit $s \in J^<$. Si $s \leq v(t)$ alors pour tout $s' \in J$ tel que $s' < s$ on a $v_+(s') \leq t$ d'après le lemme 4.3.2. Comme v_+ est continue à gauche en s , on en déduit que

$$v_+(s) = \lim_{\substack{s' \rightarrow s \\ s' < s}} v_+(s') \leq t.$$

Par conséquent,

$$v^-(s) = \inf[s \leq v] \geq v_+(s) \geq v^-(s),$$

où la dernière inégalité résulte de (4.29).

Montrons maintenant (2). Soit $t \in I^>$. Si $v_+(s) \leq t$ alors pour tout $t' \in I$ tel que $t' > t$ on a $s \leq v(t')$ d'après le lemme 4.3.2. Comme v est continue à droite en t , on en déduit que

$$s \leq \lim_{\substack{t' \rightarrow t \\ t' > t}} v(t') = v(t).$$

D'où,

$$v_+^+(t) = \sup[v_+ \leq t] \leq v(t) \leq v_+^+(t),$$

où la dernière inégalité résulte de (4.29). \square

Proposition 4.3.7. *Soient $t \in I, s \in J$.*

(1) *Si v_+ est continue à gauche en $v(t) > \min J$ alors*

$$\text{Argmin}(\mathcal{W}_t) \subset \text{Argmax}(\mathcal{R}_{v(t)}).$$

(2) *Si on suppose que (H) est vérifiée alors si v est continue à droite en $v_+(s)$ avec*

$$v_+(s) < \max I \text{ alors}$$

$$\text{Argmax}(\mathcal{R}_s) \subset \text{Argmin}(\mathcal{W}_{v_+(s)}).$$

Preuve. (1) est une conséquence du lemme 4.3.4 et de la proposition 4.3.4. (2) résulte du lemme 4.3.4 et de la proposition 4.3.5. \square

Corollaire 4.3.1. *On suppose que (H) est vérifiée.*

(1) *Soit $t \in I$ tel que $v(t) > \min J$. Si $\text{Argmin}(\mathcal{W}_t) \neq \emptyset$ et si v_+ est continue à gauche en $v(t)$ alors*

$$\text{Argmin}(\mathcal{W}_t) = \text{Argmax}(\mathcal{R}_{v(t)}).$$

(2) *Soit $s \in J$ tel que $v_+(s) < \max I$. Si $\text{Argmax}(\mathcal{R}_s) \neq \emptyset$ et si v est continue à droite en $v_+(s)$ alors*

$$\text{Argmax}(\mathcal{R}_s) = \text{Argmin}(\mathcal{W}_{v_+(s)}).$$

Preuve. Conséquence immédiate des propositions 4.3.6 et 4.3.7. \square

Corollaire 4.3.2. *On suppose que (H) est vérifiée.*

(1) *Pour tout $t \in I$ tel que v_+ soit continue en $v(t)$, on a*

$$\text{Argmin}(\mathcal{W}_t) = \text{Argmax}(\mathcal{R}_{v(t)}).$$

De plus dans ce cas $[v = v(t)] = \{t\}$.

(2) *Pour tout $s \in J$ tel que v soit continue en $v_+(s)$, on a*

$$\text{Argmax}(\mathcal{R}_s) = \text{Argmin}(\mathcal{W}_{v_+(s)}).$$

De plus, dans ce cas $[v_+ = v_+(s)] = \{s\}$.

Preuve. En juxtaposant les propositions 4.3.6 et 4.3.7, on obtient

$$\text{Argmin}(\mathcal{W}_t) = \text{Argmax}(\mathcal{R}_{v(t)}) \text{ et } \text{Argmax}(\mathcal{R}_s) = \text{Argmin}(\mathcal{W}_{v_+(s)}).$$

Par ailleurs, d'après les lemmes 4.3.3 et 4.3.4, si v_+ est continue en $v(t)$ alors $v_+(v(t)) = t$. Donc si $t' \in [v = v(t)]$ alors v_+ est continue en $v(t')$ et on a $t' = v_+(v(t')) = v_+(v(t)) = t$. D'où $[v = v(t)] = \{t\}$.

De même, comme (H) est vérifiée, il résulte des lemmes 4.3.3 et 4.3.4 que si v est continue en $v_+(s)$ alors $v(v_+(s)) = s$. Donc si $s' \in [v_+ = v_+(s)]$ alors v est continue en $v_+(s')$ et on a $s' = v(v_+(s')) = v(v_+(s)) = s$. D'où $[v_+ = v_+(s)] = \{s\}$. \square

4.3.3. Optimisation robuste versus Optimisation de la robustesse

On s'intéresse maintenant au problème posé en introduction de cette section à savoir faire le lien moyennant des conditions entre les solutions optimales de (\mathcal{W}_α) et de celles de (\mathcal{R}_β) pour des valeurs de α et β convenablement choisies et fixées.

Proposition 4.3.8. *Pour tout $x \in X$, on a $wc^+(x, \cdot) = r(x, \cdot)$ et $r^-(x, \cdot) = wc(x, \cdot)$. En particulier la fonction wc vérifie (H), c'est-à-dire*

$$(wc^+)^-(x, \cdot) = wc(x, \cdot), \quad \forall x \in X.$$

Preuve. Soit $x \in X$, d'après (4.12), on a

$$\forall (\alpha, \beta) \in I \times J, \quad wc(x, \alpha) \leq \beta \iff \alpha \leq r(x, \beta).$$

Il résulte alors de la proposition 4.3.2 que $wc^+(x, \cdot) = r(x, \cdot)$ et $r^-(x, \cdot) = wc(x, \cdot)$. \square

Proposition 4.3.9. *Soit $\alpha \in I$ et $\beta \in J$.*

- (1) *Si $\text{Argmin}(\mathcal{W}_\alpha) \neq \emptyset$ alors $\text{Argmin}(\mathcal{W}_\alpha) \supset \text{Argmax}(\mathcal{R}_{v_{wc}(\alpha)})$.*
- (2) *Si $\text{Argmax}(\mathcal{R}_\beta) \neq \emptyset$ alors $\text{Argmax}(\mathcal{R}_\beta) \supset \text{Argmin}(\mathcal{W}_{v_r(\beta)})$.*
- (3) *Si $v_r(v_{wc}(\alpha)) = \alpha$ alors $\text{Argmin}(\mathcal{W}_\alpha) = \text{Argmax}(\mathcal{R}_{v_{wc}(\alpha)})$.*
- (4) *Si $v_{wc}(v_r(\beta)) = \beta$ alors $\text{Argmax}(\mathcal{R}_\beta) = \text{Argmin}(\mathcal{W}_{v_r(\beta)})$.*

Preuve. (1) et (2) résulte d'une application directe de la proposition 4.3.6. Par ailleurs, comme (H) est vérifiée par la fonction wc , il résulte de (4.28) que

$$(4.30) \quad \forall (\alpha, \beta) \in I \times J, \quad [wc(\cdot, \alpha) \leq \beta] = [\alpha \leq r(\cdot, \beta)].$$

Donc si $v_r(v_{wc}(\alpha)) = \alpha$ alors en prenant $\beta = v_{wc}(\alpha)$ dans (4.30), on obtient

$$[wc(\cdot, \alpha) \leq v_{wc}(\alpha)] = [v_r(v_{wc}(\alpha)) \leq r(\cdot, v_{wc}(\alpha))], \quad \text{d'où (3)}.$$

De même, si $v_{wc}(v_r(\beta)) = \beta$, alors en prenant $\alpha = v_r(\beta)$ dans (4.30), on obtient

$$[wc(\cdot, v_r(\beta)) \leq v_{wc}(v_r(\beta))] = [v_r(\beta) \leq r(\cdot, \beta)], \quad \text{d'où (4)}.$$

\square

Théorème 4.3.1. *On suppose que les fonctions v_{wc} et v_r sont continues.*

(1) Pour tout $\alpha \in I$ tel que $v_{wc}(\alpha) \in \overset{\circ}{J}$, on a $\text{Argmin}(\mathcal{W}_\alpha) = \text{Argmax}(\mathcal{R}_{v_{wc}(\alpha)})$.

(2) Pour tout $\beta \in J$ tel que $v_r(\beta) \in \overset{\circ}{I}$, on a $\text{Argmax}(\mathcal{R}_\beta) = \text{Argmin}(\mathcal{W}_{v_r(\beta)})$.

Preuve. Puisque l'hypothèse (H) est vérifiée par la fonction wc , la conclusion résulte du corollaire 4.3.2. \square

Remarque 4.3.2. Si $wc(x, \cdot)$ est concave et $r(x, \cdot)$ est convexe pour tout $x \in X$ alors la fonction v_{wc} est concave et la fonction v_r est convexe. Ces fonctions sont donc continues sur l'intérieur de leurs domaines.

Proposition 4.3.10. Pour tout $x \in X$, la fonction $wc(x, \cdot) : I \rightarrow J$ est continue à gauche et la fonction $r(x, \cdot) : J \rightarrow I$ est continue à droite.

Preuve. D'après la proposition 4.3.8, $wc(x, \cdot) = r^-(x, \cdot)$ qui est continue à gauche d'après la proposition 4.3.1. De même, d'après la proposition 4.3.8, $r(x, \cdot) = wc^+(x, \cdot)$ qui est continue à droite d'après la proposition 4.3.1. \square

Un modèle générique de maximisation du rayon de stabilité

Dans ce chapitre, nous proposons un modèle générique du problème de maximisation du rayon de stabilité qui prend en compte une large classe de cas pratiques. On suppose en faite que l'ensemble de satisfaction est l'image réciproque d'un ensemble par une application. Ce qui est le cas pour la contrepartie robuste au sens du rayon de stabilité d'un problème d'optimisation à données incertaines. On montre que dans un cadre polyédral, ce modèle se ramène à un programme fractionnaire généralisé et que dans un cadre quadratique, il se ramène à un problème de programmation semi-définie positive.

5.1. Définition du modèle

Soient $(W, \|\cdot\|)$, un espace vectoriel normé, F un espace vectoriel, X un ensemble non vide.

Soit $\ell \in \mathbb{N}^*$. On considère des applications \mathbb{T}_i , $i = 1, \dots, \ell$, de X dans l'ensemble des applications affines de W dans F , définies par

$$\mathbb{T}_i(x)(w) = a_i(x) + T_i(x)w, \quad \forall x \in X, \quad \forall w \in W.$$

où $a_i : X \rightarrow F$ et $T_i : X \rightarrow \mathcal{L}(W, F)$.

On considère également des multi-applications $C_i : X \rightrightarrows F$, $i = 1, \dots, \ell$.

On définit alors, pour $i \in \llbracket 1, \ell \rrbracket$, $S_i : X \rightrightarrows W$ par

$$S_i(x) = \mathbb{T}_i(x)^-(C_i(x)) := \{w \in W \mid a_i(x) + T_i(x)w \in C_i(x)\}.$$

Cette formulation permet de se ramener au cas où la valeur nominale du paramètre incertain est le vecteur nul. On s'intéresse au problème de maximisation du rayon de robustesse

$$(\mathcal{R}_g) \quad \text{Maximiser } r(x) := d(0, S_x^c), \quad x \in X,$$

avec, pour tout $x \in X$,

$$W_x = W, \quad d_x = d_{\|\cdot\|} = d, \quad \bar{w}_x = 0, \quad \text{et} \quad S_x = \bigcap_{i=1}^{\ell} S_i(x).$$

Par définition, on a

$$r(x) = d(0, S_x^c) = d(0, \bigcup_{i=1}^{\ell} S_i(x)^c) = \min_{1 \leq i \leq \ell} d(0, S_i(x)^c).$$

En posant $r_i(x) := d(0, S_i(x)^c)$, pour $i = 1, \dots, \ell$, on obtient

$$r(x) = \min_{1 \leq i \leq \ell} r_i(x)$$

Le problème (\mathcal{R}_g) s'écrit aussi

$$(\mathcal{R}_g) \quad \text{Maximiser } \min_{i=1, \dots, \ell} r_i(x), \quad x \in X.$$

Par ailleurs, d'après le lemme 1.5.1, on a

$$\begin{aligned} d(0, S_x^c) &= \sup\{\rho > 0 \mid \rho \mathbb{B}_{\|\cdot\|} \subset S_x\}, \\ &= \sup\{\rho > 0 \mid \rho \mathbb{B}_{\|\cdot\|} \subset S_i(x), \forall i = 1, \dots, \ell\}, \\ &= \sup\{\rho > 0 \mid a_i(x) + T_i(x)w \in C_i(x), \forall w \in \rho \mathbb{B}_{\|\cdot\|}, \forall i = 1, \dots, \ell\}. \end{aligned}$$

On en déduit que résoudre (\mathcal{R}_g) revient à résoudre

$$\text{Maximiser } \rho,$$

$$\begin{aligned} (\mathcal{R}'_g) \quad & \text{s.l.c. } a_i(x) + T_i(x)w \in C_i(x), \quad \forall i \in \llbracket 1, \ell \rrbracket, \forall w \in \rho \mathbb{B}_W, \\ & (\rho, x) \in \mathbb{R}_+^* \times X, \end{aligned}$$

dans le sens où (moyennant l'hypothèse « il existe $x \in X$ tel que 0 soit à l'intérieur de S_x »)

$$v(\mathcal{R}_g) = v(\mathcal{R}'_g) \quad \text{et} \quad \text{Argmax}(\mathcal{R}_g) = \Pi_{\mathbb{R}_+^* \times X \rightarrow X} \text{Argmax}(\mathcal{R}'_g).$$

Remarque 5.1.1. Si pour tout $i \in \llbracket 1, \ell \rrbracket$, $x \in X$, $C_i(x) = C_i \subset F$, le problème (\mathcal{R}'_g) peut s'écrire

$$\text{Maximiser } \rho,$$

$$\begin{aligned} (\mathcal{R}'_g) \quad & \text{s.l.c. } a_i(x) + \rho T_i(x)w \in C_i, \quad \forall i \in \llbracket 1, \ell \rrbracket, \forall w \in \mathbb{B}_W, \\ & (\rho, x) \in \mathbb{R}_+^* \times X, \end{aligned}$$

qui est un problème d'optimisation robuste classique. Cependant, d'une part, cette approche n'est pas forcément naturelle, et, d'autre part, le problème (\mathcal{R}'_g) n'entre pas (au moins directement) dans le cadre des problèmes classiquement étudiés en Optimisation robuste.

5.1.1. Réduction partielle ou complète à des espaces de type \mathbb{R}^n

Pour simplifier un peu les écritures, on suppose ici que $\ell = 1$.

Très souvent, W et F sont des espaces vectoriels de dimensions finies et on a explicitement des bijections linéaires $L_W : \mathbb{R}^n \rightarrow W$ et $L_F : \mathbb{R}^m \rightarrow F$ pour certains $n, m \in \mathbb{N}^*$. Le problème (\mathcal{R}_g) est alors équivalent au problème

Maximiser ρ ,

$$\begin{aligned} (\underline{\mathcal{R}}'_g) \quad & \text{s.l.c. } \underline{a}(x) + \underline{T}(x)y \in \underline{C}(x), \quad \forall y \in \rho \mathbb{B}_{\|\cdot\| \circ L_W}, \\ & (\rho, x) \in \mathbb{R}_+^* \times X, \end{aligned}$$

avec

$$\underline{C}(x) := L_F^{-1}(C(x)) \subset \mathbb{R}^m, \quad \underline{a} := L_F^{-1} \circ a : X \rightarrow \mathbb{R}^m, \quad \underline{T} := \begin{cases} X \rightarrow L(\mathbb{R}^n, \mathbb{R}^m), \\ x \rightarrow L_F^{-1} \circ T(x) \circ L_W. \end{cases}$$

On notera que la norme dont on munit \mathbb{R}^n , $\|\cdot\| \circ L_W$, n'est pas forcément la norme euclidienne.

Si on suppose de plus que $X \subset D$, \mathbb{R} -espace vectoriel de dimension finie et que l'on se donne explicitement une bijection linéaire $L_D : \mathbb{R}^p \rightarrow D$, $p \in \mathbb{N}^*$, alors Le problème $(\underline{\mathcal{R}}_g)$ est équivalent au problème

Maximiser ρ ,

$$\begin{aligned} (\underline{\underline{\mathcal{R}}}'_g) \quad & \text{s.l.c. } \underline{\underline{a}}(x) + \underline{\underline{T}}(z)y \in \underline{\underline{C}}(x), \quad \forall y \in \rho \mathbb{B}_{\|\cdot\| \circ L_W}, \\ & (\rho, z) \in \mathbb{R}_+^* \times \underline{\underline{X}}, \end{aligned}$$

avec

$$\begin{aligned} \underline{\underline{C}}(x) &:= L_F^{-1}(C) \subset \mathbb{R}^m, \quad \underline{\underline{X}} := L_D^{-1}(X) \subset \mathbb{R}^p, \\ \underline{\underline{a}} &:= L_F^{-1} \circ a \circ L_D : \underline{\underline{X}} \rightarrow \mathbb{R}^m \quad \underline{\underline{T}} := \begin{cases} \underline{\underline{X}} \rightarrow L(\mathbb{R}^n, \mathbb{R}^m), \\ z \rightarrow L_F^{-1} \circ T(L_D z) \circ L_W. \end{cases} \end{aligned}$$

Tout ceci montre que l'on peut se ramener, en toute généralité, à ne considérer que le cas d'espace W , F ou D de type \mathbb{R}^k , $k \in \mathbb{N}^*$. Mais cette approche n'apporte pas forcément grand chose et n'est pas forcément naturelle, sauf, éventuellement, au moment de résoudre effectivement (sur machine) le problème, une fois qu'il aura été correctement réduit (et c'est donc d'ailleurs plutôt ce dernier problème que l'on ramènera à un problème où les espaces sont de type \mathbb{R}^k).

5.1.2. Propriétés du rayon de stabilité du modèle générique

On suppose dans cette section que $X \subset D$ est un \mathbb{R} -espace vectoriel topologique (localement convexe) séparé, et que les multi-applications C_i , $i = 1, \dots, \ell$, sont constantes et on note C_i leurs valeurs.

Proposition 5.1.1. *On suppose que :*

- (1) X est fermé et C_i est fermé pour $i = 1, \dots, \ell$;

(2) Pour tout $w \in W$, les applications $\mathcal{A}_i(\cdot, w) : X \rightarrow F$, $i = 1, \dots, \ell$, définies par $\mathcal{A}_i(x, w) = a_i(x) + T_i(x)w$, sont continues.

Alors le rayon de stabilité $r : X \rightarrow [0; +\infty]$ défini par $r(x) = d(0, S(x)^c)$ est semi-continu supérieurement.

Preuve. Pour tout $w \in W$, l'ensemble $S^-(w)$ défini par

$$S^-(w) := \{x \in X \mid \mathcal{A}_i(x, w) \in C_i, i = 1, \dots, \ell\} = \bigcap_{i=1}^{\ell} \mathcal{A}_i(\cdot, w)^{-1}(C_i)$$

est fermé. Il résulte alors de la proposition 4.2.1 que le rayon de stabilité est semi-continu supérieurement. \square

Proposition 5.1.2. *On suppose que :*

- (1) X est convexe et les ensembles C_i , $i = 1, \dots, \ell$ sont convexes ;
- (2) Pour tout $w \in W$, les applications $\mathcal{A}_i(\cdot, w) : D \rightarrow F$, $i = 1, \dots, \ell$ définies par $\mathcal{A}_i(x, w) = a_i(x) + T_i(x)w$ sont affines pour $i = 1, \dots, \ell$.

Alors le rayon de stabilité $r : X \rightarrow [0; +\infty]$ défini par $r(x) = d(0, S(x)^c)$ est quasi-concave.

Preuve. Pour tout $w \in W$, l'ensemble $S^-(w)$ défini par

$$S^-(w) := \{x \in X \mid \mathcal{A}_i(x, w) \in C_i, i = 1, \dots, \ell\} = \bigcap_{i=1}^{\ell} \mathcal{A}_i(\cdot, w)^{-1}(C_i)$$

est convexe. Il résulte de la proposition 4.2.3 que le rayon de stabilité est quasi-concave. \square

Proposition 5.1.3. *On suppose que :*

- (1) X est compact et les ensembles C_i , $i = 1, \dots, \ell$ sont fermés ;
- (2) Pour tout $w \in W$, les applications $\mathcal{A}_i(\cdot, w) : X \rightarrow F$, $i = 1, \dots, \ell$, définies par $\mathcal{A}_i(x, w) = a_i(x) + T_i(x)w$, sont continues.

Alors $\text{Argmax}(\mathcal{R}_g) \neq \emptyset$.

Preuve. D étant séparé et X un compacte de D alors X est fermé ([11], § 6, Théorème 1). Il résulte alors de la proposition 5.1.1 que l'application $r : X \rightarrow [0; +\infty]$ défini par $r(x) = d(0, S(x)^c)$ est semi-continue supérieurement. X étant compact et $r : X \rightarrow [0; +\infty]$ semi-continue supérieurement alors il existe $x_0 \in X$ tel que

$$r(x_0) = \sup_{x \in X} r(x) \quad ([29], \text{Théorème 7.4.14}).$$

\square

5.2. Étude du modèle générique dans un cadre polyédral

Dans cette section, on étudie le modèle (\mathcal{R}_g) dans un cas polyédrique c'est-à-dire qu'on suppose que pour tout $x \in X$, $C(x)$ est un polyèdre (une intersection finie de demi-espaces fermés). On détermine une formule explicite du rayon de stabilité qui montre que (\mathcal{R}_g) est un problème de programmation fractionnaire généralisé avec un nombre fini de quotients. On utilise ces résultats pour expliciter la contrepartie robuste au sens du rayon de stabilité d'un programme linéaire à données incertaines.

5.2.1. Modèle polyédrique

On reprend les notations de la section précédente.

5.2.1.1. Description des données du modèle

On suppose ici que W et F sont deux espaces vectoriels normés de dual topologique respectif W^* et F^* . On note $\langle \cdot, \cdot \rangle$ la forme bilinéaire standard de dualité. Pour tout $i \in \llbracket 1, \ell \rrbracket$, pour tout $x \in X$, on suppose que l'ensemble $C_i(x)$ est un demi-espace fermé : c'est-à-dire qu'il existe $z_i(x) \in F^*$ et $\alpha_i(x) \in \mathbb{R}$ tel que

$$(5.1) \quad C_i(x) = [\langle \cdot, z_i(x) \rangle \leq \alpha_i(x)] = \{y \in F \mid \langle y, z_i(x) \rangle \leq \alpha_i(x)\}.$$

Cette hypothèse implique que S_x est un polyèdre pour tout $x \in X$.

5.2.1.2. Calcul du rayon de stabilité du modèle

Dans le cadre décrit ci-dessus, pour tout $i \in \llbracket 1, \ell \rrbracket$, par définition de $S_i(x)$, on a

$$\begin{aligned} S_i(x) &:= \{w \in W \mid a_i(x) + T_i(x)w \in C_i(x)\} \\ &= \{w \in W \mid \langle a_i(x) + T_i(x)w, z_i(x) \rangle \leq \alpha_i(x)\} \\ &= [\langle \cdot, T_i(x)^*(z_i(x)) \rangle \leq \alpha_i(x) - \langle a_i(x), z_i(x) \rangle]. \end{aligned}$$

où $T_i(x)^*$ est l'adjoint de $T_i(x)$.

Soit $\|\cdot\|$ une norme sur W , d la distance associée et $\|\cdot\|^*$ sa norme duale.

Pour tout $x \in X$, le rayon de stabilité de x est donc défini par

$$(5.2) \quad r(x) = \min_{1 \leq i \leq \ell} d(0, [\langle \cdot, T_i(x)^*(z_i(x)) \rangle > \alpha_i(x) - \langle a_i(x), z_i(x) \rangle]).$$

D'après la formule d'Ascoli (lemme 1.5.2), si $T_i(x)^*(z_i(x))$ est non nul alors on a

$$(5.3) \quad d(0, [\langle \cdot, T_i(x)^*(z_i(x)) \rangle > \alpha_i(x) - \langle a_i(x), z_i(x) \rangle]) = \frac{(\alpha_i(x) - \langle a_i(x), z_i(x) \rangle)^+}{\|T_i(x)^*(z_i(x))\|^*}.$$

On en déduit que si pour tout $i \in \llbracket 1, \ell \rrbracket$, $T_i(x)^*(z_i(x)) \neq 0$, on a

$$(5.4) \quad r(x) = \min_{1 \leq i \leq \ell} \frac{(\alpha_i(x) - \langle a_i(x), z_i(x) \rangle)^+}{\|T_i(x)^*(z_i(x))\|^*}.$$

Par conséquent, si les applications T_i , $i \in \llbracket 1, \ell \rrbracket$ vérifient

$$(5.5) \quad T_i(x)^*(z_i(x)) \neq 0, \quad \forall x \in X,$$

alors le modèle (\mathcal{R}_g) s'écrit

$$(\mathcal{R}_{\text{lin}}) \quad \text{Maximiser } \min_{1 \leq i \leq \ell} \frac{(\alpha_i(x) - \langle a_i(x), z_i(x) \rangle)^+}{\|T_i(x)^*(z_i(x))\|^*}, \quad x \in X,$$

qui est un programme fractionnaire généralisé que l'on sait résoudre efficacement dans certains cas selon la structure de X ([24],[25],[26], [42],[68],...). Par exemple si X est un compact, on peut résoudre ce problème avec l'algorithme de Crouzeix-Ferland-Schaible ([25]).

5.2.2. Contrepartie robuste d'un programme linéaire à donnée incertaines

Considérons le programme linéaire paramétrique incertain suivant :

$$(\mathcal{P}_w^{\text{lin}}) \quad \begin{array}{ll} \text{Minimiser} & \langle \bar{c} + c(w), x \rangle + \bar{d} + d(w) \\ \text{s.l.c} & \begin{cases} \bar{A}x + A(w)x \leq \bar{b} + b(w) \\ x \in X \subset \mathbb{R}^n, \end{cases} \end{array}$$

où le paramètre incertain $w \in W$ un \mathbb{R} -espace vectoriel normé, X est un polyèdre, $c \in \mathcal{L}(W, \mathbb{R}^n)$, $d \in \mathcal{L}(W, \mathbb{R})$, $A \in \mathcal{L}(W, \mathbb{R}^{m \times n})$, $b \in \mathcal{L}(W, \mathbb{R}^m)$ et $(\bar{c}, \bar{d}, \bar{A}, \bar{b}) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^{m \times n} \times \mathbb{R}^m$ est la donnée nominale.

Soit $f : X \times W \rightarrow \mathbb{R} \cup \{+\infty\}$ la fonction définie par

$$f(x, w) = \begin{cases} \langle \bar{c} + c(w), x \rangle + \bar{d} + d(w) & \text{si } \bar{A}x + A(w)x \leq \bar{b} + b(w) \\ +\infty & \text{sinon.} \end{cases}$$

Il est facile de voir que le problème $(\mathcal{P}_w^{\text{lin}})$ se réécrit comme suit

$$(\mathcal{P}_w^{\text{lin}}) \quad \text{Minimiser } f(x, w), \quad x \in X.$$

On se donne $\alpha \in \mathbb{R}$ tel que

$$(5.6) \quad \alpha > \inf_{x \in X} f(x, 0).$$

Soit $\|\cdot\|$ une norme sur W et $d_{\|\cdot\|}$ la distance associée.

La contrepartie robuste du programme $(\mathcal{P}_w^{\text{lin}})$ au sens du rayon de stabilité (définition 4.1.4) est le programme

$$(\mathcal{R}^{\text{lin}}) \quad \text{Maximiser } d_{\|\cdot\|}(0, [f(x, \cdot) > \alpha]), \quad x \in X.$$

Par définition de l'application f , pour tout $x \in X$, on a :

$$\begin{aligned} [f(x, \cdot) \leq \alpha] &:= \{w \in W \mid f(x, w) \leq \alpha\} \\ &= \left\{ w \in W \mid \begin{cases} \langle \bar{c} + c(w), x \rangle + \bar{d} + d(w) \leq \alpha \\ \bar{A}x + A(w)x \leq \bar{b} + b(w) \end{cases} \right\}. \end{aligned}$$

On note a_i (resp. \bar{a}_i) la $i^{\text{ème}}$ ligne de A (resp. \bar{A}), b_i (resp. \bar{b}_i) la $i^{\text{ème}}$ coordonnée de b (resp. \bar{b}) pour $i \in \llbracket 1, m \rrbracket$, $\bar{a}_0 = \bar{c}$, $a_0 = c$, $\bar{b}_0 = \alpha - \bar{d}$ et $b_0 = d$.

Pour $i \in \llbracket 0, m \rrbracket$, on définit $T_i : W \rightarrow \mathbb{R}^n \times \mathbb{R}$ par $T_i(w) = (a_i(w), b_i(w))$. Avec ces notations, pour tout $x \in \mathbb{R}^n$, on a

$$\begin{aligned} [f(x, \cdot) \leq \alpha] &= \bigcap_{i=0}^m \{w \in W \mid \langle T_i(w), (x, -1) \rangle \leq \bar{b}_i - \langle \bar{a}_i, x \rangle\} \\ &= \bigcap_{i=0}^m [\langle \cdot, T_i^*(x, -1) \rangle \leq \bar{b}_i - \langle \bar{a}_i, x \rangle]. \end{aligned}$$

Supposons que pour tout $i \in \llbracket 0, m \rrbracket$, l'application linéaire T_i est surjective. Ce qui implique que T_i^* est injective. Dans ce cas, il résulte de ce qui précède que

$$(5.7) \quad d_{\|\cdot\|}(0, [f(x, \cdot) > \alpha]) = \min_{0 \leq i \leq m} \frac{(\bar{b}_i - \langle \bar{a}_i, x \rangle)^+}{\|T_i^*(x, -1)\|^*}, \quad \forall x \in \mathbb{R}^n.$$

Le programme $(\mathcal{R}^{\text{lin}})$ est donc un programme fractionnaire généralisé avec des contraintes affines :

$$(\mathcal{R}^{\text{lin}}) \quad \text{Maximiser } \min_{0 \leq i \leq m} \frac{(\bar{b}_i - \langle \bar{a}_i, x \rangle)^+}{\|T_i^*(x, -1)\|^*}, \quad x \in X.$$

5.3. Étude du modèle générique dans le cas quadratique

Dans cette section, on étudie le modèle générique du problème de maximisation du rayon de stabilité décrit à la section 5.1 en se plaçant dans des espaces de Hilbert. On suppose que $\ell = 1$ et on étudiera le modèle dans deux cas particuliers de C : cas d'une boule et cas d'un paraboloid. On terminera ce chapitre par une application au calcul d'estimateur robuste au sens du rayon de stabilité avec une illustration numérique.

On suppose ici que les espaces W et F sont des espaces de Hilbert dont on note indistinctement $\langle \cdot, \cdot \rangle$ le produit scalaire et $\|\cdot\|$ la norme associée. On rappelle que l'on se donne $a : X \rightarrow F$ et $T : X \rightarrow \mathcal{L}(W, F)$ tels que

$$\mathbb{T}(x)(w) = a(x) + T(x)w, \quad \forall (x, w) \in X \times W.$$

Notations. Dans cette section, on utilisera les notations complémentaires suivantes :

- Étant donné $(H, \langle \cdot, \cdot \rangle)$ un espace de Hilbert on définit $\mathcal{G} : H \rightarrow \mathcal{L}(\mathbb{R}, H)$ l'application qui à $a \in H$ associe $\mathcal{G}_a : \mathbb{R} \rightarrow H$ définie par $\mathcal{G}_a(t) = ta$. Il est clair que pour tout $a \in H$, \mathcal{G}_a est linéaire continue et

$$\forall x \in H, a \in H, \quad \langle \mathcal{G}_a(t), x \rangle = \langle ta, x \rangle = t \langle a, x \rangle = t \mathcal{G}_a^*(x),$$

où \mathcal{G}_a^* est l'adjoint de \mathcal{G}_a . On a donc $\mathcal{G}_a^* = \langle a, \cdot \rangle$. Ainsi, on notera souvent \mathcal{G}_a par a et \mathcal{G}_a^* par a^* .

- Si $\alpha, \beta, \gamma \in \mathbb{R}$, $a \in F$, $b \in W$ et $A \in \mathcal{L}(W, F)$, on note

$$\mathcal{F}(\alpha, a, \beta, b, A, \gamma) := \begin{bmatrix} \alpha & a^* & b^* \\ a & \beta I_F & A^* \\ b & A & \gamma I_W \end{bmatrix},$$

la fonctionnelle quadratique $\mathcal{F}(\alpha, a, \beta, b, A, \gamma) : F \times W \rightarrow \mathbb{R}$ définie par

$$\mathcal{F}(\alpha, a, \beta, b, A, \gamma)(y, w) = \alpha + \beta \|y\|^2 + \gamma \|w\|^2 + 2 \langle a, y \rangle + 2 \langle b, w \rangle + 2 \langle Aw, y \rangle.$$

Ainsi, on note simplement $\mathcal{F}(\alpha, a, \beta, b, A, \gamma) \geq 0$ pour dire que

$$\alpha + \beta \|y\|^2 + \gamma \|w\|^2 + 2 \langle a, y \rangle + 2 \langle b, w \rangle + 2 \langle Aw, y \rangle \geq 0, \quad \forall (y, w) \in F \times W.$$

- Si U et V sont deux espaces vectoriels et $f : U \rightarrow V$ une application affine alors on notera f_0 la partie constante de f et L_f sa partie linéaire.

5.3.1. Étude du modèle dans le cas d'une boule

Dans cette sous-section, on se donne une application $\alpha : X \rightarrow \mathbb{R}$ et on définit $C : X \rightrightarrows F$ par

$$(5.8) \quad C(x) := \{y \in F \mid \|y\| \leq \alpha(x)\}.$$

La contrainte semi-infinie (en la variable (ρ, x))

$$(5.9) \quad a(x) + T(x)w \in C(x), \quad \forall w \in \rho \mathbb{B}_W$$

se transforme en

$$(5.10) \quad \|a(x) + T(x)w\| \leq \alpha(x), \quad \forall w \in \rho \mathbb{B}_W.$$

Le problème (\mathcal{R}_g) dévient

$$\text{Maximiser } \rho,$$

$$(\mathcal{R}_b) \quad \text{s.l.c. } \|a(x) + T(x)w\| \leq \alpha(x), \quad \forall w \in \rho \mathbb{B}_W,$$

$$(\rho, x) \in \mathbb{R}_+^* \times X.$$

Ou de manière condensée

$$(\mathcal{R}_b) \quad \text{Maximiser } \rho, \quad (\rho, x) \in \mathcal{X},$$

si on définit l'ensemble des contraintes \mathcal{X} par

$$(5.11) \quad \mathcal{X} := \left\{ (\rho, x) \in \mathbb{R}_+^* \times X \mid \|a(x) + \mathbb{T}(x)w\| \leq \alpha(x), \quad \forall w \in \rho\mathbb{B}_W \right\}.$$

5.3.1.1. Une première réduction du problème (\mathcal{R}_b)

On se propose de remplacer la contrainte semi-infinie (5.10) en la variable (ρ, x) par une contrainte en ajoutant une nouvelle variable.

Proposition 5.3.1. *Pour tout $a \in F$, $A \in \mathcal{L}(W, F)$, $\rho \in \mathbb{R}_+^*$, $\alpha \in \mathbb{R}$ on a*

$$\|a + Aw\| \leq \alpha, \quad \forall w \in \rho\mathbb{B}_W \iff \exists \lambda \geq 0 : \mathcal{F}(\alpha - \lambda\rho^2, a, \alpha, 0, A, \lambda) \geq 0.$$

Si de plus A est non nulle, alors $\lambda \in \mathbb{R}_+^$.*

Preuve. D'après (1.14), on a

$$\forall w \in \rho\mathbb{B}_W, \|a + Aw\| \leq \alpha$$

$$\iff \alpha + 2\langle y, a + Aw \rangle + \alpha\|y\|^2 \geq 0, \quad \forall w \in \rho\mathbb{B}_W, \forall y \in F$$

$$\iff [\rho^2 - \|w\|^2 \geq 0 \Rightarrow \alpha + 2\langle a, y \rangle + 2\langle A^*y, w \rangle + \alpha\|y\|^2 \geq 0], \quad \forall w \in W, \forall y \in F.$$

On déduit alors du Théorème 1.4.1 que

$$\|a + Aw\| \leq \alpha, \quad \forall w \in \rho\mathbb{B}_W$$

$$\iff \exists \lambda \geq 0 : \alpha + 2\langle a + Aw, y \rangle + \alpha\|y\|^2 - \lambda[\rho^2 - \|w\|^2] \geq 0, \quad \forall (y, w) \in F \times W$$

$$\iff \mathcal{F}(\alpha - \lambda\rho^2, a, \alpha, 0, A, \lambda) \geq 0.$$

Si $A \neq 0$, alors il existe $(y_0, w_0) \in F \times W$ tel que $\langle Aw_0, y_0 \rangle \neq 0$.

— Si $\langle Aw_0, y_0 \rangle < 0$, on a $\lim_{n \rightarrow +\infty} \langle A(nw_0), y_0 \rangle = -\infty$.

— Si $\langle Aw_0, y_0 \rangle > 0$, on a $\lim_{n \rightarrow +\infty} \langle A(-nw_0), y_0 \rangle = -\infty$.

Il existe $w_1 \in W$ tel que

$$\alpha + 2\langle a + Aw_1, y_0 \rangle + \alpha\|y_0\|^2 < 0.$$

Par conséquent $\lambda > 0$. □

On introduit le problème d'optimisation suivant :

Maximiser ρ ,

$$(\widehat{\mathcal{R}}_b) \quad \text{s.l.c.} \quad \begin{bmatrix} \alpha(x) - \lambda\rho & a(x)^* & 0 \\ a(x) & \alpha(x)I_F & \mathbb{T}(x)^* \\ 0 & \mathbb{T}(x) & \lambda I_W \end{bmatrix} \geq 0,$$

$$(\rho, x, \lambda) \in \mathbb{R}_+^* \times X \times \mathbb{R}_+^*,$$

autrement dit

$$(\widehat{\mathcal{R}}_b) \quad \text{Maximiser } \rho, \quad (\rho, x, \lambda) \in \widehat{\mathcal{X}},$$

si on définit $\widehat{\mathcal{X}}$ par

$$\widehat{\mathcal{X}} = \left\{ (\rho, x, \lambda) \in \mathbb{R}_+^* \times X \times \mathbb{R}_+^* \mid \mathcal{F}(\alpha(x) - \lambda\rho, a(x), \alpha(x), 0, \mathbb{T}(x), \lambda) \geq 0 \right\}.$$

Théorème 5.3.1. *On suppose que pour tout $x \in X$, $\mathbb{T}(x) \in \mathcal{L}(W, F) \setminus \{0\}$. On a*

$$(5.12) \quad v(\mathcal{R}_b)^2 = v(\widehat{\mathcal{R}}_b)$$

et

$$(5.13) \quad \text{Argmax}(\mathcal{R}_b) = \left\{ (\sqrt{\rho}, x) \mid (\rho, x) \in \Pi_{\mathbb{R} \times X} \text{Argmax}(\widehat{\mathcal{R}}_b) \right\}.$$

On rappelle que $\Pi_{\mathbb{R} \times X} : (\rho, x, \lambda) \in \mathbb{R} \times X \times \mathbb{R} \rightarrow (\rho, x) \in \mathbb{R} \times X$.

Preuve. D'après la proposition 5.3.1, pour tout $(\tau, x) \in \mathbb{R}_+^* \times X$, on a

$$(5.14) \quad (\tau, x) \in \mathcal{X} \iff \exists \lambda > 0, (\tau^2, x, \lambda) \in \widehat{\mathcal{X}}.$$

Montrons d'abord (5.12). On a

$$\begin{aligned} v(\mathcal{R}_b) &= \sup \{ \rho \mid (\rho, x) \in \mathcal{X} \} \\ &= \sup \{ \rho \mid (\rho^2, x, \lambda) \in \widehat{\mathcal{X}} \} \\ &= \left(\sup \{ \rho^2 \mid (\rho^2, x, \lambda) \in \widehat{\mathcal{X}} \} \right)^{1/2} \\ &= v(\widehat{\mathcal{R}}_b)^{1/2}. \end{aligned}$$

Montrons maintenant (5.13). On procède par double inclusion. Soit $(\bar{\rho}, \bar{x}) \in \text{Argmax}(\mathcal{R}_b)$.

D'après (5.14), il existe $\bar{\lambda} > 0$ tel que $(\bar{\rho}^2, \bar{x}, \bar{\lambda}) \in \widehat{\mathcal{X}}$. Il résulte donc de (5.12) que

$$\forall (\rho, x, \lambda) \in \widehat{\mathcal{X}}, \quad \rho \leq v(\widehat{\mathcal{R}}_b) = v(\mathcal{R}_b)^2 = \bar{\rho}^2.$$

Donc $(\bar{\rho}^2, \bar{x}, \bar{\lambda}) \in \text{Argmax}(\widehat{\mathcal{R}}_b)$, et par suite $(\bar{\rho}, \bar{x}) \in \left\{ (\sqrt{\rho}, x) \mid (\rho, x) \in \Pi_{\mathbb{R} \times X} \text{Argmax}(\widehat{\mathcal{R}}_b) \right\}$.

Inversement, soit $(\sqrt{\bar{\rho}}, \bar{x}) \in \left\{ (\sqrt{\rho}, x) \mid (\rho, x) \in \Pi_{\mathbb{R} \times X} \text{Argmax}(\widehat{\mathcal{R}}_b) \right\}$. Alors il existe $\bar{\lambda} \in \mathbb{R}$ tel que $(\bar{\rho}, \bar{x}, \bar{\lambda}) \in \widehat{\mathcal{X}}$. Il résulte donc de (5.14) que $(\sqrt{\bar{\rho}}, \bar{x}) \in \mathcal{X}$. D'après (5.12), on a

$$\forall (\rho, x) \in \mathcal{X}, \quad \rho \leq v(\mathcal{R}_b) = \sqrt{v(\widehat{\mathcal{R}}_b)} = \sqrt{\bar{\rho}}.$$

Donc $(\sqrt{\bar{\rho}}, \bar{x}) \in \text{Argmax}(\mathcal{R}_b)$, d'où l'inclusion inverse. \square

Remarque 5.3.1. Supposons que $X \subset D$ un espace de Hilbert, les applications $\alpha : D \rightarrow \mathbb{R}$, $a : D \rightarrow F$ et $\mathbb{T} : D \rightarrow \mathcal{L}(W, F)$ sont affines. Dans ce cas, si $X = D$ ou la contrainte $x \in X$ est conique, c'est-à-dire de la forme

$$G(x) \in K,$$

où K est un cône convexe (fermé) de H un espace de Hilbert et $G : D \rightarrow H$ une application affine (continue) alors le programme $(\widehat{\mathcal{R}}_b)$ est « proche » d'un problème de programmation convexe conique dans les espaces de Hilbert. La structure de convexité étant détruite par le terme $\lambda\rho$ dans l'expression de $\mathcal{F}(\alpha(x) - \lambda\rho, a(x), \alpha(x), T(x), \lambda)$.

Nous allons à présent effectuer un changement de variables qui nous permettra de se ramener dans certains cas (par exemple les cas de la Remarque 5.3.1) à la résolution d'un problème de programmation convexe conique. Pour cela, **on suppose que $X \subset D$ un espace de Hilbert, les applications $\alpha : D \rightarrow \mathbb{R}$, $a : D \rightarrow \mathbf{F}$ et $T : D \rightarrow \mathcal{L}(\mathbf{W}, \mathbf{F})$ sont affines.**

5.3.1.2. Application d'un changement de variables direct

On définit l'application $\varphi : \mathbb{R}_+^* \times D \times \mathbb{R}_+^* \rightarrow \mathbb{R}_+^* \times D \times \mathbb{R}_+^*$ par

$$(5.15) \quad \varphi(\rho, x, \lambda) = \left(\rho, \frac{x}{\lambda}, \frac{1}{\lambda} \right).$$

Notons que l'application φ est une involution (i.e. bijection telle que $\varphi^{-1} = \varphi$).

On pose

$$\tilde{\mathcal{X}} := \varphi(\widehat{\mathcal{X}}) = \{(\tau, y, t) \in \mathbb{R}_+^* \times D \times \mathbb{R}_+^* \mid \varphi(\tau, y, t) \in \widehat{\mathcal{X}}\}.$$

On a donc

$$\tilde{\mathcal{X}} = \left\{ (\tau, y, t) \in \mathbb{R}_+^* \times D \times \mathbb{R}_+^* \mid \frac{y}{t} \in X, \right. \\ \left. \mathcal{F}(t\alpha_0 + L_\alpha(y) - \tau, ta_0 + L_a(y), t\alpha_0 + L_\alpha(y), 0, tT_0 + L_T(y), 1) \geq 0 \right\}.$$

On introduit le problème

$$(\tilde{\mathcal{R}}_b) \quad \text{Maximiser } \tau, \quad (\tau, y, t) \in \tilde{\mathcal{X}},$$

ou, plus explicitement,

$$\text{Maximiser } \tau,$$

$$(\tilde{\mathcal{R}}_b) \quad \text{s.l.c.} \quad \begin{bmatrix} t\alpha_0 + L_\alpha(y) - \tau & [ta_0 + L_a(y)]^* & 0 \\ ta_0 + L_a(y) & (t\alpha_0 + L_\alpha(y))I_F & [tT_0 + L_T(y)]^* \\ 0 & tT_0 + L_T(y) & I_W \end{bmatrix} \geq 0$$

$$\frac{y}{t} \in X, \quad (\tau, y, t) \in \mathbb{R}_+^* \times D \times \mathbb{R}_+^*.$$

On déduit immédiatement du fait que φ est une involution la proposition suivante.

Proposition 5.3.2. *Les problèmes $(\widehat{\mathcal{R}}_b)$ et $(\tilde{\mathcal{R}}_b)$ sont liés par*

$$v(\widehat{\mathcal{R}}_b) = v(\tilde{\mathcal{R}}_b) \text{ et } \text{Argmax}(\widehat{\mathcal{R}}_b) = \varphi(\text{Argmax}(\tilde{\mathcal{R}}_b)).$$

5.3.1.3. Application d'un changement de variables de Charnes-Cooper

En opérant le changement de variables

$$\psi := \begin{cases} \mathbb{R}_+^* \times D \times \mathbb{R}_+^* \rightarrow \mathbb{R}_+^* \times D \times \mathbb{R}_+^*, \\ (\rho, x, \lambda) \rightarrow (\lambda\rho, x, \lambda), \end{cases}$$

on transforme le problème $(\widehat{\mathcal{R}}_b)$ en le problème

$$(\mathcal{P}) \quad \text{Maximiser } \frac{\tau}{s}, \quad (\tau, x, s) \in \mathcal{S},$$

où l'ensemble \mathcal{S} est défini par

$$\mathcal{S} = \left\{ (\tau, x, s) \in \mathbb{R}_+^* \times D \times \mathbb{R}_+^* \mid x \in X, \mathcal{F}(\alpha(x) - \tau, a(x), \alpha(x), 0, T(x), s) \geq 0 \right\}.$$

On définit les fonctions f et g sur \mathcal{S} par

$$\forall (\tau, x, s) \in \mathcal{S}, \quad f(\tau, x, s) = \tau \text{ et } g(\tau, x, s) = s.$$

Conformément aux notations de la section 1.3, on a

$$\begin{aligned} \tilde{\mathcal{S}} &:= \left\{ (\tau, x, s, t) \in \mathbb{R} \times D \times \mathbb{R} \times \mathbb{R} \mid t > 0, \left(\frac{\tau}{t}, \frac{x}{t}, \frac{s}{t} \right) \in \mathcal{S} \right\}, \\ &= \left\{ (\tau, x, s, t) \in \mathbb{R}_+^* \times D \times \mathbb{R}_+^* \times \mathbb{R}_+^* \mid \right. \\ &\quad \left. \frac{x}{t} \in X, \mathcal{F}(t\alpha_0 + L_\alpha(x) - \tau, ta_0 + L_a(x), t\alpha_0 + L_\alpha(x), 0, tT_0 + L_T(x), s) \geq 0 \right\}. \end{aligned}$$

Pour tout $(\tau, x, s, t) \in \tilde{\mathcal{S}}$,

$$\tilde{f}(\tau, x, s, t) = \tau \quad \text{et} \quad \tilde{g}(\tau, x, s, t) = s,$$

où \tilde{f} et \tilde{g} sont respectivement les fonctions perspectives de f et g (section 1.3).

Il en résulte que

$$\begin{aligned} \tilde{\mathcal{S}} \cap [\tilde{g} = 1] &= \left\{ (\tau, x, s, t) \in \mathbb{R}_+^* \times D \times \mathbb{R}_+^* \times \mathbb{R}_+^* \mid s = 1, \right. \\ &\quad \left. \frac{x}{t} \in X, \mathcal{F}(t\alpha_0 + L_\alpha(x) - \tau, L_a(x) + ta_0, t\alpha_0 + L_\alpha(x), 0, tT_0 + L_T(x), s) \geq 0 \right\}. \end{aligned}$$

Si on considère le problème

$$(\tilde{\mathcal{P}}_=) \quad \text{Maximiser } \tau, \quad (\tau, x, s, t) \in \tilde{\mathcal{S}} \cap [\tilde{g} = 1],$$

il est immédiat que l'on peut l'écrire

$$(\tilde{\mathcal{P}}_=) \quad \text{Maximiser } \tau, \quad (\tau, x, t) \in \tilde{\mathcal{X}}, \quad s = 1$$

qui est trivialement relié au problème $(\tilde{\mathcal{R}}_b)$.

Plus exactement, d'après le lemme 1.3.1, on a

$$v(\widehat{\mathcal{R}}_b) = v(\mathcal{P}) = v(\tilde{\mathcal{P}}_=) = v(\tilde{\mathcal{R}}_b).$$

De plus

$$\text{Argmax}(\mathcal{P}) = \mathcal{F}^{-1}(\text{Argmax}(\tilde{\mathcal{P}}_=))$$

avec $\mathcal{T}^{-1} : (\tau, x, s, t) \rightarrow (\tau/t, x/t, s/t)$, et

$$\text{Argmax}(\mathcal{P}) = \hat{\mathcal{T}}^{-1}(\text{Argmax}(\tilde{\mathcal{R}}_b))$$

avec $\hat{\mathcal{T}}^{-1} : (\tau, x, t) \rightarrow (\tau/t, x/t, 1/t)$, puis, comme $\psi^{-1} : (\tau, x, s) \rightarrow (\tau/s, x, s)$,

$$\text{Argmax}(\widehat{\mathcal{R}}_b) = \psi^{-1}(\text{Argmax}(\mathcal{P})) = \psi^{-1} \circ \hat{\mathcal{T}}^{-1}(\text{Argmax}(\tilde{\mathcal{R}}_b)) = \varphi(\text{Argmax}(\tilde{\mathcal{R}}_b)).$$

5.3.2. Étude du modèle dans le cas d'un « paraboloïde »

Dans cette sous-section, on suppose que :

- (1) $a = (\tilde{a}, \alpha) : X \rightarrow F \times \mathbb{R}$ définie par $a(x) = (\tilde{a}(x), \alpha(x))$ pour tout $x \in X$,
- (2) $T = (\tilde{T}, 2b) : X \rightarrow \mathcal{L}(W, F) \times \mathcal{L}(W, \mathbb{R})$ définie par $T(x)w = (\tilde{T}(x)w, 2\langle b(x), w \rangle)$ pour tout $(x, w) \in X \times W$,
- (3) C est le « paraboloïde » défini par $C := \{(z, t) \in F \times \mathbb{R} \mid \|z\|^2 \leq t\}$.

Pour tout $(x, w) \in X \times W$, on a l'équivalence

$$(5.16) \quad a(x) + T(x)w \in C \iff \|\tilde{a}(x) + \tilde{T}(x)w\|^2 \leq \alpha(x) + 2\langle b(x), w \rangle.$$

Dans la suite de cette sous-section, par soucis de simplification, on fait un abus en notant \tilde{a} par a et \tilde{A} par A . Le problème (\mathcal{R}_g) devient dans ce cas

Maximiser ρ ,

$$(\mathcal{R}_p) \quad \text{s.l.c.} \quad \|a(x) + T(x)w\|^2 \leq \alpha(x) + 2\langle b(x), w \rangle, \quad \forall w \in \rho\mathbb{B}_W$$

$$(\rho, x) \in \mathbb{R}_+^* \times X.$$

5.3.2.1. Une première réduction du problème (\mathcal{R}_p)

Proposition 5.3.3. *Pour tout $a \in F$, $b \in W$, $A \in \mathcal{L}(W, F)$, $\rho \in \mathbb{R}_+^*$ et $\alpha \in \mathbb{R}$ on a*

$$\|a + Aw\|^2 \leq \alpha + 2\langle b, w \rangle, \quad \forall w \in \rho\mathbb{B}_W \iff \exists \lambda \geq 0 : \mathcal{F}(\alpha - \lambda\rho^2, a, 1, b, A, \lambda) \geq 0.$$

Si de plus A est non nulle, alors $\lambda \in \mathbb{R}_+^$.*

Preuve. On a

$$\|a + Aw\|^2 \leq \alpha + 2\langle b, w \rangle, \quad \forall w \in \rho\mathbb{B}_W$$

$$\iff \alpha + 2\langle b, w \rangle - \|a + Aw\|^2 \geq 0, \quad \forall w \in \rho\mathbb{B}_W$$

$$\iff \alpha + 2\langle b, w \rangle + \min_{\xi \in F} \{2\langle a + Aw, \xi \rangle + \|\xi\|^2\} \geq 0, \quad \forall w \in \rho\mathbb{B}_W$$

$$\iff \alpha + 2\langle b, w \rangle + 2\langle a + Aw, \xi \rangle + \|\xi\|^2 \geq 0, \quad \forall w \in \rho\mathbb{B}_W, \forall \xi \in F$$

$$\iff [\rho^2 - \|w\|^2 \geq 0 \Rightarrow \alpha + 2\langle b, w \rangle + 2\langle a + Aw, \xi \rangle + \|\xi\|^2 \geq 0], \quad \forall (\xi, w) \in F \times W.$$

On déduit alors du théorème 1.4.1 que

$$\|a + Aw\|^2 \leq \alpha + 2\langle b, w \rangle, \quad \forall w \in \rho\mathbb{B}_W \iff$$

$$\exists \lambda \geq 0 : \alpha + 2\langle a + Aw, \xi \rangle + 2\langle b, w \rangle + \|\xi\|^2 - \lambda[\rho^2 - \|w\|^2] \geq 0, \quad \forall (\xi, w) \in F \times W,$$

d'où le résultat. Si de plus $A \neq 0$, alors il existe $\xi \in F$ tel que $A^*\xi + b \neq 0$. Il existe $w_0 \in W$ tel que $\langle A^*\xi + b, w_0 \rangle \neq 0$.

— Si $\langle A^*\xi + b, w_0 \rangle < 0$ alors $\lim_{n \rightarrow +\infty} \langle A^*\xi + b, nw_0 \rangle = -\infty$.

— Si $\langle A^*\xi + b, w_0 \rangle > 0$ alors $\lim_{n \rightarrow +\infty} \langle A^*\xi + b, -nw_0 \rangle = -\infty$.

Il existe $w_1 \in W$ tel que $\alpha + 2\langle a, \xi \rangle + 2\langle A^*\xi + b, w_1 \rangle + \|\xi\|^2 < 0$. Donc $\lambda > 0$. \square

On introduit le problème

$$\begin{aligned} & \text{Maximiser } \rho, \\ (\widehat{\mathcal{R}}_p) \quad & \text{s.l.c. } \begin{bmatrix} \alpha(x) - \lambda\rho & a(x)^* & b(x)^* \\ a(x) & I_F & T(x)^* \\ b(x) & T(x) & \lambda I_W \end{bmatrix} \geq 0, \\ & (\rho, x, \lambda) \in \mathbb{R}_+^* \times X \times \mathbb{R}_+^*. \end{aligned}$$

Les problèmes (\mathcal{R}_p) et $(\widehat{\mathcal{R}}_p)$ sont reliés par le résultat suivant.

Théorème 5.3.2. *Supposons que $\forall x \in X, A(x) \in \mathcal{L}(W, F) \setminus \{0\}$. On a*

$$v(\mathcal{R}_p)^2 = v(\widehat{\mathcal{R}}_p) \quad \text{et} \quad \text{Argmax}(\mathcal{R}_p) = \left\{ (\sqrt{\rho}, x) \mid (\rho, x) \in \Pi_{\mathbb{R} \times X} \text{Argmax}(\widehat{\mathcal{R}}_p) \right\}.$$

Preuve. Analogue à celle du théorème 5.3.1. \square

5.3.2.2. Changement de variables

On suppose que $X \subset D$ un espace de Hilbert, les applications $\alpha : D \rightarrow \mathbb{R}$, $a : D \rightarrow F$, $b : D \rightarrow \mathcal{L}(W, \mathbb{R})$ et $T : D \rightarrow \mathcal{L}(W, F)$ sont affines, puis on considère le problème

$$\begin{aligned} & \text{Maximiser } \tau, \\ (\tilde{\mathcal{R}}_p) \quad & \text{s.l.c. } \begin{bmatrix} t\alpha_0 + L_\alpha(y) - \tau & [ta_0 + L_a(y)]^* & [tb_0 + L_b(y)]^* \\ ta_0 + L_a(y) & tI_F & [tT_0 + L_T(y)]^* \\ tb_0 + L_b(y) & tT_0 + L_T(y) & I_W \end{bmatrix} \geq 0, \\ & \frac{y}{t} \in X, \quad (\tau, y, t) \in \mathbb{R}_+^* \times D \times \mathbb{R}_+^*. \end{aligned}$$

Les problèmes $(\widehat{\mathcal{R}}_p)$ et $(\tilde{\mathcal{R}}_p)$ sont liés par le résultat suivant.

Proposition 5.3.4. *On a*

$$v(\widehat{\mathcal{R}}_p) = v(\tilde{\mathcal{R}}_p) \quad \text{et} \quad \text{Argmax}(\widehat{\mathcal{R}}_p) = \varphi(\text{Argmax}(\tilde{\mathcal{R}}_p)).$$

Preuve. Résulte immédiatement du fait que φ est une involution. \square

5.3.3. Un exemple en dimension finie : une deuxième réduction de (\mathcal{R}_b) et (\mathcal{R}_p)

On suppose que tous les espaces de Hilbert apparaissant dans cette sous-section sont de dimension finie.

Généralité. Soit $W := \mathcal{L}(E, H)$, où E et H sont deux espaces de Hilbert. On muni W du produit scalaire $\langle\langle \cdot, \cdot \rangle\rangle$ défini par

$$\langle\langle u, v \rangle\rangle := \text{Tr}(u^* \circ v), \quad \forall u, v \in L(E, H),$$

où $\text{Tr}(\cdot)$ est l'opérateur trace défini sur l'espace des endomorphismes de E . On rappelle que si $u : E \rightarrow E$ est un endomorphisme alors la trace de u est le réel défini par

$$\text{Tr}(u) = \sum_{i=1}^{\dim E} \langle u(e_i), e_i \rangle, \quad \text{où } (e_i)_{1 \leq i \leq \dim E} \text{ est une base hilbertienne de } E.$$

Rappelons les deux propriétés élémentaires (mais utiles dans les calculs qui vont suivre) du produit scalaire $\langle\langle \cdot, \cdot \rangle\rangle$:

$$(5.17) \quad \langle\langle h \langle e, \cdot \rangle, w \rangle\rangle = \langle we, h \rangle, \quad \forall (h, e, w) \in H \times E \times W,$$

$$(5.18) \quad \langle\langle h_1 \langle e_1, \cdot \rangle, h_2 \langle e_2, \cdot \rangle \rangle\rangle = \langle e_1, e_2 \rangle \langle h_1, h_2 \rangle, \quad \forall (e_1, e_2) \in E^2, \forall (h_1, h_2) \in H^2.$$

On suppose dans ce paragraphe que les applications $T \in \mathcal{L}(W, F)$ et $b \in \mathcal{L}(W, \mathbb{R})$ sont définies par :

$$(5.19) \quad \begin{aligned} Tw &= L_1(wR) \quad \text{et} \quad \langle\langle b, w \rangle\rangle = \langle L_2, wR \rangle, \quad \forall w \in \mathcal{L}(E, H), \\ \text{où } (L_1, L_2, R) &\in \mathcal{L}(H, F) \times H \times (E \setminus 0) \quad \text{est donné.} \end{aligned}$$

Proposition 5.3.5. *Soit $(\alpha, \beta, \rho, a) \in \mathbb{R}^3 \times F$. Pour tout $(\lambda, \mu) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$ tel que $\lambda = \mu \|R\|^2$ on a*

$$\begin{bmatrix} \alpha - \lambda\rho & a^* & b^* \\ a & \beta I_F & T^* \\ b & T & \lambda I_W \end{bmatrix} \geq 0 \iff \begin{bmatrix} \alpha - \mu\rho \|R\|^2 & a^* & L_2^* \\ a & \beta I_F & L_1^* \\ L_2 & L_1 & \mu I_H \end{bmatrix} \geq 0.$$

Preuve. Par définition de l'inégalité fonctionnelle quadratique, on a

$$\begin{aligned}
& \begin{bmatrix} \alpha - \lambda\rho & a^* & b^* \\ a & \beta I_F & T^* \\ b & T & \lambda I_W \end{bmatrix} \geq 0 \\
& \iff \alpha - \lambda\rho + 2\langle a, y \rangle + \beta\|y\|^2 + 2\langle T w, y \rangle + 2\langle b, w \rangle + \lambda\|w\|^2 \geq 0, \quad \forall (y, w) \in F \times W \\
& \iff \alpha - \lambda\rho + 2\langle a, y \rangle + \beta\|y\|^2 + 2\langle L_1(wR), y \rangle + 2\langle L_2, wR \rangle + \lambda\|w\|^2 \geq 0, \quad \forall (y, w) \in F \times W \\
& \iff \alpha - \lambda\rho + 2\langle a, y \rangle + \beta\|y\|^2 + 2\langle w, (L_1^*y + L_2)\langle R, \cdot \rangle \rangle + \lambda\|w\|^2 \geq 0, \quad \forall (y, w) \in F \times W \\
& \iff \alpha - \lambda\rho + 2\langle a, y \rangle + \beta\|y\|^2 + \min_{w \in W} \{2\langle w, (L_1^*y + L_2)\langle R, \cdot \rangle \rangle + \lambda\|w\|^2\} \geq 0, \quad \forall y \in F.
\end{aligned}$$

Soit y fixé dans F . La fonctionnelle (strictement convexe car $\lambda > 0$) quadratique

$$J_y : W \rightarrow \mathbb{R}, \quad w \mapsto 2\langle w, (L_1^*y + L_2)\langle R, \cdot \rangle \rangle + \lambda\|w\|^2,$$

atteint son minimum sur W en $\bar{w} = -(L_1^*y + L_2)\langle R, \cdot \rangle/\lambda$.

Soit $V := \{z\langle R, \cdot \rangle : z \in H\}$. Il est clair que $\bar{w} \in V \subset W$. Par conséquent,

$$\min_{w \in W} J_y(w) = \min_{v \in V} J_y(v) = \min_{z \in H} J_y(z\langle R, \cdot \rangle).$$

Il en résulte que

$$\begin{aligned}
& \begin{bmatrix} \alpha - \lambda\rho & a^* & b^* \\ a & \beta I_F & T^* \\ b & T & \lambda I_W \end{bmatrix} \geq 0 \\
& \iff \alpha - \lambda\rho + 2\langle a, y \rangle + \beta\|y\|^2 + \min_{w \in W} J_y(w) \geq 0, \quad \forall y \in F \\
& \iff \alpha - \lambda\rho + 2\langle a, y \rangle + \beta\|y\|^2 + \min_{z \in H} J_y(z\langle R, \cdot \rangle) \geq 0, \quad \forall y \in F \\
& \iff \alpha - \lambda\rho + 2\langle a, y \rangle + \beta\|y\|^2 + J_y(z\langle R, \cdot \rangle) \geq 0, \quad \forall (y, z) \in F \times H \\
& \iff \alpha - \lambda\rho + 2\langle a, y \rangle + \beta\|y\|^2 + 2\|R\|^2\langle z, L_1^*y + L_2 \rangle + \lambda\|R\|^2\|z\|^2 \geq 0, \quad \forall (y, z) \in F \times H \\
& \iff \alpha - \mu\rho\|R\|^2 + 2\langle a, y \rangle + \beta\|y\|^2 + 2\|R\|^2\langle z, L_1^*y + L_2 \rangle + \mu\|R\|^4\|z\|^2 \geq 0, \quad \forall (y, z) \in F \times H \\
& \iff \alpha - \mu\rho\|R\|^2 + 2\langle a, y \rangle + \beta\|y\|^2 + 2\langle \xi, L_1^*y + L_2 \rangle + \mu\|\xi\|^2 \geq 0, \quad \forall (y, \xi) \in F \times H.
\end{aligned}$$

Notons que la dernière équivalence résulte du fait que $R \neq 0$ qui nous permet d'effectuer le changement de variable $\xi = \|R\|^2 z$. \square

Corollaire 5.3.1. Soit $(\alpha, \beta, \rho, a) \in \mathbb{R}^3 \times F$. On a l'équivalence suivante :

$$\exists \lambda > 0 : \begin{bmatrix} \alpha - \lambda\rho & a^* & b^* \\ a & \beta I_F & A^* \\ b & T & \lambda I_W \end{bmatrix} \geq 0 \iff \exists \lambda > 0 : \begin{bmatrix} \alpha - \lambda\rho\|R\|^2 & a^* & L_2^* \\ a & \beta I_F & L_1^* \\ L_2 & L_1 & \lambda I_H \end{bmatrix} \geq 0.$$

Proposition 5.3.6. Soit $(\alpha, \beta, a) \in \mathbb{R}^2 \times F$. Pour tout $(\lambda, \mu, \rho) \in \mathbb{R}_+^3$ tel que $\lambda\mu\rho = \|R\|^2$, on a

$$\begin{bmatrix} \alpha - \lambda\rho & a^* & b^* \\ a & \beta I_F & T^* \\ b & T & \lambda I_W \end{bmatrix} \geq 0 \iff \begin{bmatrix} \alpha - \mu\rho\|L_2\|^2 & (a - \mu\rho L_1 L_2)^* & R^* \\ a - \mu\rho L_1 L_2 & \beta I_F - \mu\rho L_1 L_1^* & 0 \\ R & 0 & \mu I_E \end{bmatrix} \geq 0.$$

Preuve. On sait que

$$\begin{bmatrix} \alpha - \lambda\rho & a^* & b^* \\ a & \beta I_F & T^* \\ b & T & \lambda I_W \end{bmatrix} \geq 0$$

$$\begin{aligned} &\Leftrightarrow \alpha - \lambda\rho + 2\langle a, y \rangle + \beta\|y\|^2 + \min_{z \in H} J_y(z\langle R, \cdot \rangle) \geq 0, \quad \forall y \in F \\ &\Leftrightarrow \alpha - \lambda\rho + 2\langle a, y \rangle + \beta\|y\|^2 - \frac{1}{\lambda}\|R\|^2\|L_1^*y + L_2\|^2 \geq 0, \quad \forall y \in F \\ &\Leftrightarrow \alpha - \frac{1}{\mu}\|R\|^2 + 2\langle a, y \rangle + \beta\|y\|^2 - \mu\rho\|L_1^*y + L_2\|^2 \geq 0, \quad \forall y \in F \\ &\Leftrightarrow \alpha - \frac{1}{\mu}\|R\|^2 + 2\langle a, y \rangle + \beta\|y\|^2 - \mu\rho(\|L_1^*y\|^2 + 2\langle L_2, L_1^*y \rangle + \|L_2\|^2) \geq 0, \quad \forall y \in F \\ &\Leftrightarrow \alpha - \mu\rho\|L_2\|^2 + 2\langle a, y \rangle + \beta\|y\|^2 - \mu\rho\|L_1^*y\|^2 - 2\mu\rho\langle L_2, L_1^*y \rangle + \\ &\quad \min_{\xi \in E} \{2\langle R, \xi \rangle + \mu\|\xi\|^2\} \geq 0, \quad \forall y \in F \\ &\Leftrightarrow \alpha - \mu\rho\|L_2\|^2 + 2\langle a, y \rangle + \beta\|y\|^2 - \mu\rho\|L_1^*y\|^2 - 2\mu\rho\langle L_2, L_1^*y \rangle + \\ &\quad 2\langle R, \xi \rangle + \mu\|\xi\|^2 \geq 0, \quad \forall (y, \xi) \in F \times E. \end{aligned}$$

D'où le résultat. □

Corollaire 5.3.2. Soit $(\alpha, \beta, a) \in \mathbb{R}^2 \times F$. Pour tout $\rho > 0$, on a

$$\exists \lambda > 0 : \begin{bmatrix} \alpha - \lambda\rho & a^* & b^* \\ a & \beta I_F & T^* \\ b & T & \lambda I_W \end{bmatrix} \geq 0 \Leftrightarrow \exists \lambda > 0 : \begin{bmatrix} \alpha - \lambda\rho\|L_2\|^2 & (a - \lambda\rho L_1 L_2)^* & R^* \\ a - \lambda\rho L_1 L_2 & \beta I_F - \lambda\rho L_1 L_1^* & 0 \\ R & 0 & \lambda I_E \end{bmatrix} \geq 0.$$

Modèle de la boule. On reprend le modèle de la boule puis on suppose que $W := \mathcal{L}(E, H)$ muni du produit scalaire $\langle \cdot, \cdot \rangle$ avec E et H deux espaces de Hilbert. On suppose en plus que l'application T est telle que

$$(5.20) \quad \exists (L, R) \in \mathcal{L}(H, F) \times (E \setminus \{0\}) : Tw = L(wR) \quad \forall w \in W.$$

Cette structuration est l'une des plus utilisée en optimisation robuste . Elle a été explicitement considérée dans [7] (chapitre 6). En pratique, on prend $L : D \rightarrow \mathcal{L}(H, F)$ et $R : D \rightarrow E$ des applications affines non nulles dont l'une au moins est constante. Dans ce cas, on redéfinit le problème $(\widehat{\mathcal{R}}_b)$ comme suit.

— **Cas de R constante :**

Maximiser ρ ,

$$(\widehat{\mathcal{R}}_b) \quad \text{s.l.c.} \quad \begin{bmatrix} \alpha(x) - \lambda\rho\|R\|^2 & a(x)^* & 0 \\ a(x) & \alpha(x)I_F & L(x)^* \\ 0 & L(x) & \lambda I_H \end{bmatrix} \geq 0,$$

$$(\rho, x, \lambda) \in \mathbb{R}_+^* \times X \times \mathbb{R}_+^*,$$

— **Cas de L constante :**

Maximiser ρ ,

$$(\widehat{\mathcal{R}}_b) \quad \text{s.l.c.} \quad \begin{bmatrix} \alpha(x) & a(x)^* & R(x)^* \\ a(x) & \alpha(x)I_F - \lambda\rho LL^* & 0 \\ R(x) & 0 & \lambda I_E \end{bmatrix} \geq 0,$$

$$(\rho, x, \lambda) \in \mathbb{R}_+^* \times X \times \mathbb{R}_+^*.$$

— Si L et R sont constantes on prend un des deux problèmes.

Il résulte des propositions 5.3.5 et 5.3.6 que le théorème 5.3.1 reste vrai. Par ailleurs, le problème $(\tilde{\mathcal{R}}_b)$ devient

— **Cas de R constante :**

Maximiser τ ,

$$(\tilde{\mathcal{R}}_b) \quad \text{s.l.c.} \quad \begin{bmatrix} t\alpha_0 + L_\alpha(y) - \tau\|R\|^2 & [ta_0 + L_a(y)]^* & 0 \\ ta_0 + L_a(y) & (t\alpha_0 + L_\alpha(y))I_F & [tL_0 + L_L(y)]^* \\ 0 & tL_0 + L_L(y) & I_H \end{bmatrix} \geq 0,$$

$$\frac{y}{t} \in X, (\tau, y, t) \in \mathbb{R}_+^* \times D \times \mathbb{R}_+^*,$$

— **Cas de L constante :**

Maximiser τ ,

$$(\tilde{\mathcal{R}}_b) \quad \text{s.l.c.} \quad \begin{bmatrix} t\alpha_0 + L_\alpha(y) & [ta_0 + L_a(y)]^T & [tR_0 + L_R(y)]^* \\ ta_0 + L_a(y) & (t\alpha_0 + L_\alpha(y))I_F - \tau LL^* & 0 \\ tR_0 + L_R(y) & 0 & I_E \end{bmatrix} \geq 0,$$

$$\frac{y}{t} \in X, (\tau, y, t) \in \mathbb{R}_+^* \times D \times \mathbb{R}_+^*.$$

La proposition 5.3.2 reste vraie.

Modèle du paraboloïde. On se place dans le cas du paraboloïde et on suppose que $W = \mathcal{L}(E, H)$ est muni du produit scalaire $\langle \cdot, \cdot \rangle$, les applications $T \in \mathcal{L}(W, F)$ et $b \in \mathcal{L}(W, \mathbb{R})$ sont définies par

$$Tw = L_1(wR) \quad \text{et} \quad \langle b, w \rangle = \langle L_2, wR \rangle, \quad \forall w \in \mathcal{L}(E, H).$$

$$\text{avec } (L_1, L_2, R) \in \mathcal{L}(H, F) \times H \times (E \setminus 0) \text{ donné}$$

On redéfinit dans ce cas le problème $(\widehat{\mathcal{R}}_p)$ comme suit.

— **Cas de R constante :**

$$\text{Maximiser } \rho,$$

$$(\widehat{\mathcal{R}}_p) \quad \text{s.l.c.} \quad \begin{bmatrix} \alpha(x) - \lambda \rho \|R\|^2 & a(x)^* & L_2(x)^* \\ a(x) & \alpha(x)I_F & L_1(x)^* \\ L_2(x) & L_1(x) & \lambda I_H \end{bmatrix} \geq 0,$$

$$(\rho, x, \lambda) \in \mathbb{R}_+^* \times X \times \mathbb{R}_+^*,$$

— **Cas de $L = (L_1, L_2)$ constante**

$$\text{Maximiser } \rho,$$

$$(\widehat{\mathcal{R}}_p) \quad \text{s.l.c.} \quad \begin{bmatrix} \alpha(x) - \lambda \rho \|L_2\|^2 & (a(x) - \lambda \rho L_1 L_2)^* & R(x)^* \\ a(x) - \lambda \rho L_1 L_2 & \alpha(x)I_F - \lambda \rho L_1 L_1^* & 0 \\ R(x) & 0 & \lambda I_E \end{bmatrix} \geq 0,$$

$$(\rho, x, \lambda) \in \mathbb{R}_+^* \times X \times \mathbb{R}_+^*.$$

Il résulte des propositions 5.3.5 et 5.3.6 que le théorème 5.3.2 reste vrai. Le problème $(\tilde{\mathcal{R}}_p)$ devient :

— **Cas de R constante :**

$(\tilde{\mathcal{R}}_p)$

$$\text{Maximiser } \tau,$$

$$\text{s.l.c.} \quad \begin{bmatrix} t\alpha_0 + L_\alpha(y) - \tau \|R\|^2 & (ta_0 + L_\alpha(y))^* & (L_{2_0} + L_{L_2}(y))^* \\ ta_0 + L_\alpha(y) & (t\alpha_0 + L_\alpha(y))I_F & (tL_{1_0} + L_{L_1}(y))^* \\ L_{2_0} + L_{L_2}(y) & tL_{1_0} + L_{L_1}(y) & I_H \end{bmatrix} \geq 0,$$

$$\frac{y}{t} \in X, \quad (\tau, y, t) \in \mathbb{R}_+^* \times D \times \mathbb{R}_+^*.$$

— **Cas de $L = (L_1, L_2)$ constante :**

Maximiser τ ,

$$\text{s.l.c.} \quad \begin{bmatrix} t\alpha_0 + L_\alpha(y) - \tau\|L_2\|^2 & (t\alpha_0 + L_\alpha(y) - \tau L_1 L_2)^* & (tR_0 + L_R(y))^* \\ ta_0 + L_a(y) - \tau L_1 L_2 & (t\alpha_0 + L_\alpha(y))I_F - \tau L_1 L_1^* & 0 \\ tR_0 + L_R(y) & 0 & I_E \end{bmatrix} \geq 0,$$

$$\frac{y}{t} \in X, (\tau, y, t) \in \mathbb{R}_+^* \times D \times \mathbb{R}_+^*,$$

et la proposition 5.3.4 reste vraie.

Commentaire : Importance de la seconde réduction. On sait que si E et H sont de dimension finie alors $\mathcal{L}(E, H)$ est de dimension finie et on a

$$\dim \mathcal{L}(E, H) = \dim E \times \dim H.$$

Par conséquent,

$$(5.21) \quad \mathcal{F}(t\alpha_0 + L_\alpha(y) - \tau, ta_0 + L_a(y), t\alpha_0 + L_\alpha(y), tT_0 + L_T(y), 1) \geq 0$$

est une «linear matrix inequality» dans laquelle les matrices symétriques sont de taille $1 + \dim F + \dim E \times \dim H$. Les redéfinitions ci-dessus montrent qu'on peut réduire la taille des matrices de (5.21) à $1 + \dim F + \dim H$ si R est constante et à $1 + \dim F + \dim E$ si L est constante. Ce qui peut avoir un intérêt non négligeable du point de vue numérique. Par exemple si $\dim E$ est assez petit devant $\dim E \times \dim H$ dans le cas de R constante.

5.3.4. Application : Calcul d'un estimateur robuste

On reprend ici un exemple de problème d'estimation à données incertaines de [7] que nous approchons par la méthode du rayon de stabilité.

5.3.4.1. Description du problème

Considérons la situation suivante : On dispose d'une observation bruitée

$$(5.22) \quad u = (I_p + w)z + \xi,$$

d'un signal z qui est le résultat du passage d'un signal inconnu v à travers un filtre linéaire :

$$z = Av.$$

Où $A \in \mathbb{R}^{p \times q}$ est connue. Les observations contiennent deux types d'erreurs :

- le biais wz dépendant linéairement de z où la matrice de biais $w \in \mathbb{R}^{p \times p}$ est incertaine,

— le bruit aléatoire $\xi \in \mathbb{R}^p$, de moyenne nulle et de matrice de covariance Σ connue.

L'objectif est d'estimer une fonction linéaire f du signal d'entrée v : c'est-à-dire la quantité $\langle f, v \rangle := f^T v$. On se restreint aux estimateurs linéaires en u : $\widehat{f} = x^T u$, où $x \in \mathbb{R}^p$ est le vecteur poids à déterminer. En clair on cherche x solution du problème d'approximation suivant :

$$(5.23) \quad f^T v \cong \widehat{f} = x^T u = x^T [(I_p + w)z + \xi] = x^T [(I_p + w)Av + \xi].$$

Par ailleurs, on sait que le signal d'entrée v appartient à une ellipse de \mathbb{R}^q . On suppose donc qu'il existe une matrice symétrique définie positive $Q \in \mathbb{R}^{q \times q}$ et un réel $r > 0$ tel que tout signal d'entrée v vérifie $v^T Q v \leq r$.

5.3.4.2. Gestion du paramètre aléatoire

On élimine le paramètre aléatoire ξ en considérant la moyenne quadratique du résidu $f^T v - \widehat{f}$. On définit ainsi une fonction M par

$$(5.24) \quad M(x, w, v) = \sqrt{E\left([x^T[(I_p + w)Av + \xi] - f^T v]^2\right)},$$

où $E(\cdot)$ désigne l'espérance mathématique. Comme la moyenne de ξ est nulle et $E(\xi\xi^T) = \Sigma$ est connue, alors on a

$$M(x, w, v) = \sqrt{\left([A^T(I_p + w^T)x - f]^T v\right)^2 + x^T \Sigma x}.$$

Le problème revient à trouver x tel que $M(x, w, v)$ soit la plus petite possible pour tout $v \in V_Q := \{v \in \mathbb{R}^q : v^T Q v \leq r\}$. Ce qui est équivalent au problème d'optimisation incertain suivant

$$(\mathcal{P}_w) \quad \text{Minimiser } \max_{v \in V_Q} M(x, w, v), \quad x \in \mathbb{R}^p.$$

Observons que

$$\begin{aligned}
\max_{v \in V_Q} M(x, w, v) &= \max_{v \in V_Q} \sqrt{\left([A^T(I_p + w^T)x - f]^T v \right)^2 + x^T \Sigma x} \\
&= \sqrt{\max_{v \in V_Q} \left([A^T(I_p + w^T)x - f]^T v \right)^2 + x^T \Sigma x} \\
&= \sqrt{\max_{v \in V_Q} \left([A^T(I_p + w^T)x - f]^T v \right)^2 + x^T \Sigma x} \\
&= \sqrt{\left(\max_{v \in V_Q} [A^T(I_p + w^T)x - f]^T v \right)^2 + x^T \Sigma x} \\
&= \sqrt{r^2 \|Q^{-1/2} A^T(I_p + w^T)x - Q^{-1/2} f\|^2 + x^T \Sigma x} \\
&= \left\| \begin{bmatrix} r(Q^{-1/2} A^T(I_p + w^T)x - Q^{-1/2} f) \\ \Sigma^{1/2} x \end{bmatrix} \right\|.
\end{aligned}$$

Ainsi, le programme (\mathcal{P}_w) dévient

$$(\mathcal{P}_w) \quad \text{Minimiser} \quad \left\| \begin{bmatrix} r(Q^{-1/2} A^T(I_p + w^T)x - Q^{-1/2} f) \\ \Sigma^{1/2} x \end{bmatrix} \right\|, \quad x \in \mathbb{R}^p.$$

5.3.4.3. Gestion du paramètre incertain : contrepartie robuste

Pour approcher le problème incertain (\mathcal{P}_w) , on utilise la technique du rayon de stabilité. On se donne $\alpha \in \mathbb{R}$ tel que

$$(5.25) \quad \alpha > \inf_{x \in \mathbb{R}^p} \left\| \begin{bmatrix} r(Q^{-1/2} A^T x - Q^{-1/2} f) \\ \Sigma^{1/2} x \end{bmatrix} \right\|,$$

et on muni $\mathbb{R}^{p \times q}$ de la norme de Frobenius notée $\|\cdot\|$. La contrepartie robuste au sens du rayon de stabilité est donc définie par

$$(\mathcal{R}^\alpha) \quad \text{Maximiser } \rho, \quad (\rho, x) \in \mathcal{X},$$

où l'ensemble des solutions réalisables \mathcal{X} est donné par

$$\mathcal{X} = \left\{ (\rho, x) \in \mathbb{R} \times \mathbb{R}^p \mid \rho > 0 \text{ et } \left\| \begin{bmatrix} r(Q^{-1/2} A^T(I_p + w)x - Q^{-1/2} f) \\ \Sigma^{1/2} x \end{bmatrix} \right\| \leq \alpha, \quad \forall w \in \rho \mathbb{B}_{\|\cdot\|} \right\}.$$

Observons que

$$\begin{bmatrix} r(Q^{-1/2} A^T(I_p + w)x - Q^{-1/2} f) \\ \Sigma^{1/2} x \end{bmatrix} = \begin{bmatrix} rQ^{-1/2} A^T x - rQ^{-1/2} f \\ \Sigma^{1/2} x \end{bmatrix} + \begin{bmatrix} rQ^{-1/2} A^T \\ 0 \end{bmatrix} w x.$$

On en déduit que (\mathcal{R}^α) est un cas particulier du modèle (\mathcal{R}_b) en dimension finie (cas structuré avec L constante). Il résulte donc de ce qui précède que résoudre

(\mathcal{R}^α) revient à résoudre le programme suivant :

$$\begin{aligned} & \text{Maximiser } \tau, \\ (\mathcal{R}_{\text{ex}}) \quad & \text{s.l.c. } \Lambda(\tau, y, t) \geq 0, \\ & (\tau, y, t) \in \mathbb{R}_+ \times \mathbb{R}^p \times \mathbb{R}_+, \end{aligned}$$

où

$$\Lambda(\tau, y, t) := \begin{bmatrix} t\alpha & -tr f^T Q^{-1/2} + r y^T A Q^{-1/2} & y^T \Sigma^{1/2} & y^T \\ -tr Q^{-1/2} f + r Q^{-1/2} A^T y & t\alpha I_q - \tau r^2 Q^{-1/2} A^T A Q^{-1/2} & 0 & 0 \\ \Sigma^{1/2} y & 0 & t\alpha I_p & 0 \\ y & 0 & 0 & I_p \end{bmatrix}.$$

Le programme (\mathcal{R}_{ex}) est un problème d'optimisation convexe conique classique (programme semi-défini positif). Notons que nous avons remplacé les contraintes $\tau > 0$ et $t > 0$ par les contraintes $\tau \geq 0$ et $t \geq 0$, ce qui n'est pas gênant d'après [72], Corollaire 2 ou [2], Proposition 11.3.4 dans le cas de la dimension finie.

5.3.4.4. Illustration numérique : un problème de chaleur

Une mince plaque métallique homogène occupe le carré $D = [0; 1] \times [0; 1]$ dans le plan. Au temps $t = 0$, on la chauffe jusqu'à une température $T(0, \cdot) : D \rightarrow \mathbb{R}$ telle que

$$\int_D T^2(0, x, y) dx dy \leq T_0^2,$$

avec $T_0 > 0$ une constante donnée, puis on la laisse se refroidir en maintenant à 0° la température le long du périmètre de la plaque.

À un instant 2τ donné, on mesure la température $T(2\tau, x, y)$ à chaque point de la grille

$$(5.26) \quad \Gamma = \{(u_i, u_j) \mid 1 \leq i, j \leq N\}, \quad \text{avec } u_k = \frac{2k-1}{2N}, \quad \text{pour tout } k \in \llbracket 1, N \rrbracket.$$

Le vecteur u des mesures est obtenu à partir du vecteur $z = \{T(2\tau, u_i, u_j) \mid 1 \leq i, j \leq N\}$ selon l'équation (5.22), où les variables aléatoire ξ_{ij} suivent la loi normale centrée d'écart-type σ donné ($\xi_{ij} \rightsquigarrow \mathcal{N}(0, \sigma)$) et sont indépendantes.

Étant données les mesures, on cherche à estimer la température au centre de la plaque au temps τ c'est-à-dire qu'on cherche une estimation de $T(\tau, \frac{1}{2}, \frac{1}{2})$.

Il est bien connue en physique que l'évolution dans le temps de la température T d'une plaque métallique homogène occupant un domaine Ω du plan, sans source de chaleur dans le domaine et échangeant faiblement la température à travers le bord

du domaine est gouvernée par l'équation de la chaleur :

$$(5.27) \quad \frac{\partial}{\partial t} T = \frac{\partial^2}{\partial x^2} T + \frac{\partial^2}{\partial y^2} T.$$

Dans le cas, où $\Omega = D$ avec la condition au bord de Dirichlet homogène, la solution de l'équation (5.27) est donnée par

$$(5.28) \quad T(t, x, y) = \sum_{k,l=1}^{+\infty} a_{kl} \exp(-\pi^2(k^2 + l^2)t) \sin(\pi kx) \sin(\pi ly),$$

où les coefficients a_{kl} peuvent être obtenus en utilisant la température initiale et la base orthogonale $\phi_{kl}(x, y) = \sin(\pi kx) \sin(\pi ly)$ de $L^2(D)$:

$$(5.29) \quad a_{kl} = 4 \int_D T(0, x, y) \phi_{kl}(x, y) dx dy.$$

En posant $v_{kl} = a_{kl} \exp(-\pi^2(k^2 + l^2)\tau)$, le problème revient à trouver une estimation de

$$(5.30) \quad T(\tau, \frac{1}{2}, \frac{1}{2}) = \sum_{k,l=1}^{+\infty} v_{kl} \phi_{kl}(\frac{1}{2}, \frac{1}{2}),$$

étant donné une observation $u = (I + w)z + \xi$ avec $z = \{T(2\tau, u_i, u_j) \mid 1 \leq i, j \leq N\}$, $\xi = \{\xi_{ij} \sim \mathcal{N}(0, \sigma) \mid 1 \leq i, j \leq N\}$ et les ξ_{ij} sont indépendantes.

Observons que $a_{kl} = \exp(\pi^2(k^2 + l^2)\tau) v_{kl}$ et

$$(5.31) \quad \sum_{k,l=1}^{+\infty} v_{kl}^2 \exp(2\pi^2(k^2 + l^2)\tau) = \sum_{k,l=1}^{+\infty} a_{kl}^2 = 4 \int_D T^2(0, x, y) dx dy \leq 4T_0^2.$$

Il en résulte que $|v_{kl}| \leq 2T_0 \exp(-\pi^2(k^2 + l^2)\tau)$.

Étant donné $\epsilon > 0$, on détermine un entier naturel non nul L tel que

$$(5.32) \quad \sum_{k,l:k^2+l^2>L^2} \exp(-\pi^2(k^2 + l^2)\tau) \leq \frac{\epsilon}{2T_0}.$$

Ainsi, en remplaçant par 0 les coefficients v_{kl} pour $k^2 + l^2 > L^2$, on change la température au temps τ d'au plus ϵ . Par conséquent, en choisissant ϵ assez petit, on peut supposer pour des raisons pratiques que $v_{kl} = 0$ pour $k^2 + l^2 > L^2$. Ce qui rend finie la dimension de notre problème. Plus précisément, on a : étant donnés les paramètres N, σ, T_0, L et une observation $u = (I + w)z + \xi$ avec $z = \{T(2\tau, u_i, u_j) \mid 1 \leq i, j \leq N\}$, $\xi = \{\xi_{ij} \sim \mathcal{N}(0, \sigma) \mid 1 \leq i, j \leq N\}$ et les ξ_{ij} sont indépendantes, trouver une estimation de la quantité

$$(5.33) \quad \sum_{k,l:k^2+l^2 \leq L^2} v_{kl} \phi_{kl}(\frac{1}{2}, \frac{1}{2}).$$

On doit donc trouver les coefficients $x_{ij} \in \mathbb{R}$ tel que

$$(5.34) \quad T(\tau, \frac{1}{2}, \frac{1}{2}) \approx \widehat{T} = \sum_{i,j} x_{ij} u_{ij}.$$

On retrouve bien un cas particulier du modèle d'estimateur robuste proposé ci-dessus avec

$$Q = \text{diag}((\exp(2\pi^2(k^2 + l^2)\tau))_{k,l \geq 1: k^2+l^2 \leq L^2}), \quad f = (\phi_{kl}(\frac{1}{2}, \frac{1}{2}))_{k,l \geq 1: k^2+l^2 \leq L^2}$$

$$A = (\exp(-\pi^2(k^2 + l^2)\tau)\phi_{kl}(u_i, u_j))_{\substack{1 \leq i,j \leq N \\ k,l \geq 1: k^2+l^2 \leq L^2}}, \quad \Sigma = \sigma^2 I_q, \quad r = 2T_0.$$

On peut résoudre le programme (\mathcal{R}_{ex}) avec cvxopt. Par exemple, si on suppose que τ est la moitié du temps au bout duquel tous les coefficients de Fourier de $T(t, \dots)$ sont multipliés par un facteur inférieur à 10^{-1} alors on a

$$\tau = \frac{\ln 10}{4\pi^2} \approx 0,0583,$$

et si $T_0 = 1000$, $N = 4$, $\epsilon = 10^{-15}$, on obtient $L = 8$. Dans ce cas, pour $\alpha = 63,9$ et $\sigma = 0,1$ on obtient $v(\mathcal{R}_{\text{ex}}) = 0,1$ et le vecteur poids x est

$$x = \begin{pmatrix} 0,11566261 & 0,2793136 & 0,2793136 & 0,11566261 \\ 0,2793136 & 0,67451424 & 0,67451424 & 0,2793136 \\ 0,2793136 & 0,67451424 & 0,67451424 & 0,2793136 \\ 0,11566261 & 0,2793136 & 0,2793136 & 0,11566261 \end{pmatrix}.$$

Régression linéaire à données incertaines : méthode du rayon de stabilité

6.1. Introduction

L'équation basique $Ax = b$ où $A \in \mathbb{R}^{m \times n}$ et $b \in \mathbb{R}^m$ sont donnés et $x \in X \subset \mathbb{R}^n$ est l'inconnue, n'a pas toujours de solution. C'est le cas par exemple si b n'est pas dans l'image de A .

En général, on se contente de trouver $x \in X$ tel que Ax soit « le plus proche possible » de b . C'est-à-dire qu'on se donne une norme $\|\cdot\|$ sur \mathbb{R}^m et on cherche $x \in X$ tel que $\|Ax - b\|$ soit la plus petite possible. En clair, on cherche à résoudre le problème d'optimisation suivant :

$$(\mathcal{P}_r) \quad \text{Minimiser } \|Ax - b\|, \quad x \in X.$$

Le programme (\mathcal{P}_r) est un problème classique connu sous le nom de problème de régression linéaire sous contraintes. Dans les applications, il arrive très souvent que la donnée (A, b) de (\mathcal{P}_r) soit incertaine. On sait simplement qu'elle appartient à un sous-ensemble de $\mathbb{R}^{m \times n} \times \mathbb{R}^m$. Le programme (\mathcal{P}_r) est alors un problème d'optimisation à données incertaines. Dans la suite, (\mathcal{P}_r) désigne sa version incertaine. L'approche du problème (\mathcal{P}_r) par la méthodologie de l'optimisation robuste a fait l'objet de nombreux travaux ([7],[31],[43],[75], ...). On propose dans ce chapitre une nouvelle approche «robuste» du problème (\mathcal{P}_r) basée sur la méthodologie du rayon de stabilité.

Le chapitre est organisé comme suit. La section 6.2 formule la contrepartie robuste de (\mathcal{P}_r) au sens du rayon de stabilité. Dans la section 6.3, on détermine une formule plus «explicite» du rayon de robustesse moyennant des hypothèses que nous précisons. La section 6.4 étudie la contrepartie robuste du problème (\mathcal{P}_r) qui est un programme fractionnaire. On transforme ce programme fractionnaire en un programme convexe et, on montre que dans certains cas, ce programme convexe se ramène à un programme linéaire. Dans la section 6.5, on fait une petite comparaison entre la contrepartie robuste de (\mathcal{P}_r) au sens de l'optimisation robuste et celle au sens du rayon de stabilité. La section 6.6 termine le chapitre par une application à un problème de conception d'antenne circulaire.

6.2. Contrepartie robuste au sens du rayon de stabilité

Notations. On rappelle que $\mathbb{R}^{m \times n}$ désigne l'espace des matrices à m lignes et n colonnes. On munit $\mathbb{R}^{m \times n}$ (resp. \mathbb{R}^m) du produit scalaire de Frobenius (resp. canonique) noté $\langle\langle \cdot, \cdot \rangle\rangle$ (resp. $\langle \cdot, \cdot \rangle$). Pour tout $t \in \mathbb{R}$, on note $t^+ := \max(t, 0)$. Pour tout $x \in \mathbb{R}^m$ et $y \in \mathbb{R}^n$, $(x; y)$ est le vecteur de \mathbb{R}^{m+n} obtenu en concaténant les vecteurs x et y . Pour tout $A = (a_{ij}) \in \mathbb{R}^{m \times n}$, $a_{i \cdot}$ est la $i^{\text{ème}}$ ligne de A , $a_{\cdot j}$ sa $j^{\text{ème}}$ colonne, et $\text{vect}(A)$ est le vecteur de \mathbb{R}^{mn} défini par

$$\text{vect}(A) := (a_{\cdot 1}; \dots; a_{\cdot n}).$$

Pour tout $(A, b) \in \mathbb{R}^{m \times n} \times \mathbb{R}^m$, $[A \ b]$ est la matrice de $\mathbb{R}^{m \times (n+1)}$ obtenue en plaçant b à la suite de la dernière colonne de A .

Contrepartie robuste. On suppose que les données **A** et **b** du problème (\mathcal{P}_r) dépendent d'un paramètre incertain $w \in W$ un espace vectoriel normé de dimension finie dont on note $\|\cdot\|$ sa norme. On considère ainsi des applications

$$A : W \rightarrow \mathbb{R}^{m \times n}, \quad b : W \rightarrow \mathbb{R}^m$$

qui sont supposées linéaires, puis pour tout $w \in W$, on obtient le programme paramétré suivant

$$(\mathcal{P}_r^w) \quad \text{Minimiser } \|A(w)x - b(w)\|, \quad x \in X.$$

Soit \bar{w} l'incertitude nominale. On note $\bar{A} = A(\bar{w})$ et $\bar{b} = b(\bar{w})$. Soit $\alpha \in \mathbb{R}$ tel que

$$(6.1) \quad \alpha > \inf_{x \in X} \|\bar{A}x - \bar{b}\|.$$

D'après la définition 4.1.4, la contrepartie robuste au sens du rayon de stabilité du problème (\mathcal{P}_r^w) , $w \in W$, est le programme

$$(\mathcal{R}_r^\alpha) \quad \text{Maximiser } r_\alpha(x, \bar{w}) := d_{\|\cdot\|}(\bar{w}, [f(x, \cdot) > \alpha]), \quad x \in X,$$

où la fonction $f : X \times W \rightarrow \mathbb{R}$ est définie par

$$(6.2) \quad f(x, w) = \|A(w)x - b(w)\|.$$

Lien avec le modèle (\mathcal{R}_g) . On a, pour $x \in X$,

$$d_{\|\cdot\|}(\bar{w}, [f(x, \cdot) > \alpha]) = d_{\|\cdot\|}(0, [f(x, \cdot) > \alpha] - \bar{w})$$

et

$$\begin{aligned} [f(x, \cdot) \leq \alpha] - \bar{w} &= \{w \in W \mid \|\bar{A}x - \bar{b} + A(w)x - b(w)\| \leq \alpha\}, \\ &= \{w \in W \mid \|[\bar{A} \ \bar{b}](x; -1) + [A(w) \ b(w)](x; -1)\| \leq \alpha\}. \end{aligned}$$

Si on pose, pour $z \in \mathbb{R}^{n+1}$,

$$L := \begin{cases} W \rightarrow \mathbb{R}^{m \times (n+1)}, \\ w \rightarrow [A(w) \ b(w)], \end{cases} \quad V_z := \begin{cases} \mathbb{R}^{m \times (n+1)} \rightarrow \mathbb{R}^m, \\ B \rightarrow Bz, \end{cases}$$

alors, pour $w \in W$, $x \in X$,

$$\|[\bar{A} \ \bar{b}](x; -1) + [A(w) \ b(w)](x; -1)\| \leq \alpha \iff \|[\bar{A} \ \bar{b}](x; -1) + (V_{(x; -1)} \circ L)(w)\| \leq \alpha.$$

Puis, en posant, pour $z \in \mathbb{R}^{n+1}$

$$a(z) = [\bar{A} \ \bar{b}]z, \quad T(z) = V_z \circ L,$$

on a

$$\|[\bar{A} \ \bar{b}](x; -1) + [A(w) \ b(w)](x; -1)\| \leq \alpha \iff \|a((x; -1)) + T((x; -1))w\| \leq \alpha.$$

Finalement

$$[f(x, \cdot) \leq \alpha] - \bar{w} = \{w \in W \mid a((x; -1)) + T((x; -1))w \in \alpha \mathbb{B}_{\|\cdot\|}\},$$

ce qui montre clairement que (\mathcal{R}_r^α) est bien un cas particulier du modèle générique (\mathcal{R}_g) proposé au chapitre 5.

6.3. Calcul du rayon de stabilité

Le but de cette section est de «calculer» le rayon de stabilité de l'ensemble $[f(x, \cdot) \leq \alpha]$ apparaissant dans la contrepartie robuste (au sens du rayon de stabilité) du problème de régression robuste considéré dans la section précédente.

Pour la clarté du raisonnement et pour faciliter le lien avec les autres parties de notre travail, nous nous plaçons dans un cadre légèrement plus général.

Soit $T : W \rightarrow F$ un opérateur linéaire non nul d'adjoint T^* , $\alpha \in \mathbb{R}_+^*$ et $a \in F$. On définit S_α par

$$S_\alpha := \{w \in W \mid a + Tw \in \alpha \mathbb{B}_{\|\cdot\|}\}.$$

Il s'agit donc d'établir une formule pour $d_{\|\cdot\|, \|\cdot\|}(0, S_\alpha)$.

6.3.1. Une première réduction

Quelques outils d'Analyse convexe. Soient $(W, \|\cdot\|)$ et $(F, \|\cdot\|)$ deux espaces vectoriels normés de dimension finie avec $F \neq \{0\}$, de dual topologique respectif $(W^*, \|\cdot\|)$ et $(F^*, \|\cdot\|)$. On note $\langle \cdot, \cdot \rangle$ la forme bilinéaire standard de dualité entre W et W^* et $\langle \cdot, \cdot \rangle$ celle entre F et F^* . En clair, pour tout $(w, w^*) \in W \times W^*$ (resp. $((y, y^*) \in F \times F^*)$), on note $\langle w, w^* \rangle := w^*(w)$ (resp. $\langle y, y^* \rangle := y^*(y)$). Notons que si l'espace est de Hilbert, on l'identifie à son dual par le théorème de représentation de Riesz-Fréchet ([17], Théorème V.5). Dans ce cas, le crochet de dualité standard coïncide avec le produit scalaire de l'espace et on utilise la même notation pour le produit scalaire et le crochet de dualité.

Soit S un sous-ensemble de F . On appelle *enveloppe convexe de S* , notée $\text{conv}S$, le plus petit convexe contenant S . On appelle *polaire de S* , noté S° , le sous-ensemble de F^* défini par

$$(6.3) \quad S^\circ := \{y^* \in F^* \mid \langle y, y^* \rangle \leq 1, \forall y \in S\} = \bigcap_{y \in S} [\langle y, \cdot \rangle \leq 1].$$

Le *bipolaire de S* , noté $S^{\circ\circ}$, est défini par $S^{\circ\circ} = (S^\circ)^\circ$. On vérifie que :

$$(6.4) \quad S^\circ = (\text{conv}S)^\circ.$$

Théorème 6.3.1 ([61], §14). *Si C est un convexe fermé contenant l'origine alors*

$$C^{\circ\circ} = C.$$

Par ailleurs, par définition (norme duale et polaire), on a $(\mathbb{B}_{\|\cdot\|})^\circ = \mathbb{B}_{\|\cdot\|^*}$. On obtient, d'après le Théorème 6.3.1 et (6.4) que

$$(6.5) \quad \mathbb{B}_{\|\cdot\|} = \mathbb{B}_{\|\cdot\|}^{\circ\circ} = \mathbb{B}_{\|\cdot\|^*}^\circ = (\text{conv}(\mathbb{S}_{\|\cdot\|^*}))^\circ = \mathbb{S}_{\|\cdot\|^*}^\circ,$$

où $\mathbb{S}_{\|\cdot\|^*} = \{y^* \in F^* \mid \|y^*\|^* = 1\}$.

On peut vérifier que pour tout $S \subset F$, pour tout $\alpha > 0$, on a $(\alpha S)^\circ = \frac{1}{\alpha} S^\circ$, et donc $(\alpha S)^{\circ\circ} = \alpha S^{\circ\circ}$. Par conséquent, pour tout $\alpha \in \mathbb{R}_+^*$, on a, d'après (6.3) et (6.5)

$$(6.6) \quad \alpha \mathbb{B}_{\|\cdot\|} = \alpha \mathbb{S}_{\|\cdot\|^*}^\circ = \left(\frac{1}{\alpha} \mathbb{S}_{\|\cdot\|^*}\right)^\circ = \bigcap_{y^* \in \mathbb{S}_{\|\cdot\|^*}} [\langle \cdot, y^* \rangle \leq \alpha].$$

Une première formule pour le rayon de stabilité.

Théorème 6.3.2. *Supposons que l'opérateur T est surjectif. Alors on a*

$$d_{\|\cdot\|, \|\cdot\|}(0, S_\alpha^c) = \inf_{y^* \in \mathbb{S}_{\|\cdot\|^*}} \frac{(\alpha - \langle a, y^* \rangle)^+}{\|T^* y^*\|^*}.$$

Preuve. On a, d'après (6.6),

$$S_\alpha = \{w \in W \mid a + Tw \in \alpha \mathbb{B}_{\|\cdot\|}\} = \bigcap_{y^* \in \mathbb{S}_{\|\cdot\|^*}} [\langle \cdot, T^* y^* \rangle \leq \alpha - \langle a, y^* \rangle],$$

$$\begin{aligned} d_{\|\cdot\|, \|\cdot\|}(0, S_\alpha^c) &= d_{\|\cdot\|, \|\cdot\|}\left(0, \bigcup_{y^* \in \mathbb{S}_{\|\cdot\|^*}} [\langle \cdot, T^* y^* \rangle > \alpha - \langle a, y^* \rangle]\right), \\ &= \inf_{y^* \in \mathbb{S}_{\|\cdot\|^*}} d_{\|\cdot\|, \|\cdot\|}\left(0, [\langle \cdot, T^* y^* \rangle > \alpha - \langle a, y^* \rangle]\right). \end{aligned}$$

On conclut en utilisant la formule d'Ascoli et le fait que, comme T est surjectif, T^* est injectif et, comme $y^* \neq 0$ (puisque $\|y^*\|^* = 1$), $T^* y^* \neq 0$. \square

Considérons l'hypothèse suivante :

$$(H_1) \quad \text{Il existe } N(T) \text{ tel que : } \forall y^* \in F^*, \quad \|T^* y^*\|^* = N(T) \|y^*\|^*.$$

Théorème 6.3.3. *Supposons que (H_I) est vérifiée et que $T \in \mathcal{L}(W, F)$ est surjectif. Alors on a*

$$(6.7) \quad d_{\|\cdot\|}(0, S_\alpha^c) = \frac{(\alpha - \|a\|)^+}{N(T)}.$$

Preuve. À l'aide de l'hypothèse (H_I) , il résulte du théorème 6.3.2 que

$$d_{\|\cdot\|}(0, S_\alpha^c) = \frac{1}{N(T)} \inf_{y^* \in \mathbb{S}_{\|\cdot\|}^*} (\alpha - \langle a, y^* \rangle)^+ = \frac{1}{N(T)} \left(\alpha - \sup_{y^* \in \mathbb{S}_{\|\cdot\|}^*} \langle a, y^* \rangle \right)^+ = \frac{(\alpha - \|a\|)^+}{N(T)}.$$

□

6.3.2. Retour à la régression robuste dans le cas non structuré

Dans le cas de la régression robuste considérée plus haut, on a

$$F = \mathbb{R}^m, \quad a = \bar{A}z \in \mathbb{R}^m, \quad T = V_z \circ L : W \rightarrow \mathbb{R}^m$$

où $\bar{A} \in \mathbb{R}^{m \times q}$ (avec $q = n + 1$), $z \in \mathbb{R}^q$, $z \neq 0$. Dans un premier cas qui nous intéresse, le cas non structuré, on suppose que

$$W = \mathbb{R}^{m \times q}, \quad L = I_{\mathbb{R}^{m \times q}} \text{ et donc } T = V_z.$$

On a

$$S_\alpha = \{A \in \mathbb{R}^{m \times q} \mid \bar{A}z + Az \in \alpha \mathbb{B}_{\|\cdot\|}\}.$$

Si $A \in \mathbb{R}^{m \times q}$, $y \in \mathbb{R}^m$,

$$\langle V_z A, y \rangle = \langle Az, y \rangle = \sum_{i=1}^m \sum_{j=1}^q a_{ij} z_j y_i = \langle\langle A, yz^T \rangle\rangle.$$

Ceci montre que $V_z^* : \mathbb{R}^m \rightarrow \mathbb{R}^{m \times q}$ est donné par

$$\forall y \in \mathbb{R}^m, \quad T^* y = V_z^* y = yz^T.$$

On peut noter que l'on a donc

$$\|T^* y\|^* = \|V_z^* y\|^* = \sup_{\|A\| \leq 1} \langle y, Az \rangle.$$

Dans le cas qui nous intéresse, la propriété (H_I) devient :

$$(H'_I) \quad \text{Il existe } N : \mathbb{R}^q \rightarrow \mathbb{R}_+ \text{ tel que : } \forall z \in \mathbb{R}^q, \forall y \in \mathbb{R}^m, \quad \|yz^T\|^* = N(z) \|y\|^*.$$

En fait, par abus, on emploie la même notation N dans $N(z)$ et $N(V_z \circ L)$.

Nous allons voir que dans deux cas «classiques» de définitions «liées» des normes $\|\cdot\|$ et $\|\cdot\|$ respectivement sur $\mathbb{R}^{m \times q}$ et \mathbb{R}^m , la propriété (H'_I) est vérifiée et donc la formule (6.7) permettant le calcul du rayon de stabilité est vraie.

Norme subordonnée. Soit N_q une norme sur \mathbb{R}^q . On suppose que $\| \cdot \|$ est subordonnée aux normes $\| \cdot \|$ et N_q , c'est-à-dire qu'elle est définie, pour $A \in \mathbb{R}^{m \times q}$, par

$$(6.8) \quad \| \| A \| \| = \max_{N_q(z) \leq 1} \| Az \|.$$

Lemme 6.3.1. *Supposons que $\| \cdot \|$ est donnée par (6.8). On a*

$$(6.9) \quad \forall y \in \mathbb{R}^m, \forall z \in \mathbb{R}^q, \quad \| \| yz^T \| \| ^* = N_q(z) \| y \| ^*.$$

Preuve. Soit $y \in \mathbb{R}^p$ et $z \in \mathbb{R}^q$. Observons que

$$\| \| yz^T \| \| = \max_{N_q(u) \leq 1} \| yz^T u \| = \max_{N_q(u) \leq 1} |\langle z, u \rangle| \| y \| = N_q^*(z) \| y \|.$$

On en déduit que

$$\| \| yz^T \| \| ^* := \max_{w \neq 0} \frac{\langle y, wz \rangle}{\| \| w \| \|} \geq \max_{\substack{v \neq 0 \\ u \neq 0}} \frac{\langle y, vu^T z \rangle}{\| \| vu^T \| \|} = \max_{\substack{v \neq 0 \\ u \neq 0}} \frac{\langle y, v \rangle}{\| v \|} \frac{\langle u, z \rangle}{N_q^*(u)} = N_q(z) \| y \| ^*.$$

Par ailleurs, on a

$$\| \| yz^T \| \| ^* = \max_{w \neq 0} \frac{\langle y, wz \rangle}{\| \| w \| \|} \leq \max_{w \neq 0} \frac{\| y \| ^* \| wz \|}{\| \| w \| \|} \leq \max_{w \neq 0} \frac{\| y \| ^* \| \| w \| \| N_q(z)}{\| \| w \| \|} = N_q(z) \| y \| ^*.$$

□

Dans ce cas, on voit que (H'_1) est vérifiée avec $N = N_q$.

Normes composées. On rappelle qu'une norme θ sur \mathbb{R}^p est dite absolue si et seulement si $\theta(x_1, \dots, x_p) = \theta(|x_1|, \dots, |x_p|)$, $\forall x = (x_1, \dots, x_p) \in \mathbb{R}^p$. Par ailleurs, si θ est une norme absolue alors θ^* est une norme absolue ([4]). On utilisera le lemme suivant.

Lemme 6.3.2 ([19]). *Soient $\theta_1, \dots, \theta_k$ des normes sur $\mathbb{R}^{n_1}, \dots, \mathbb{R}^{n_k}$ et $\tilde{\theta}$ une norme absolue sur \mathbb{R}^k . Si θ est la norme sur $\mathbb{R}^{n_1} \times \dots \times \mathbb{R}^{n_k}$ définie par*

$$\forall (u_1, \dots, u_k) \in \mathbb{R}^{n_1} \times \dots \times \mathbb{R}^{n_k}, \quad \theta(u_1, \dots, u_k) = \tilde{\theta}(\theta_1(u_1), \dots, \theta_k(u_k))$$

alors sa norme duale, θ^ , est donnée par*

$$\forall (u_1, \dots, u_k) \in \mathbb{R}^{n_1} \times \dots \times \mathbb{R}^{n_k}, \quad \theta^*(u_1, \dots, u_k) = \tilde{\theta}^*(\theta_1^*(u_1), \dots, \theta_k^*(u_k)).$$

Soit N_q une norme sur \mathbb{R}^q . On définit la norme $\| \cdot \|$ sur $\mathbb{R}^{m \times q}$ de deux manières différentes comme suit.

(1) **Ligne-colonne :** On suppose que N_q est absolue et on définit $\| \cdot \|$ par

$$(6.10) \quad \| \| w \| \| = N_q(\| w_{.1} \|, \dots, \| w_{.q} \|), \quad \forall w \in \mathbb{R}^{m \times q}.$$

(2) **Colonne-ligne :** On suppose que $\| \cdot \|$ est absolue et on définit $\| \cdot \|$ par

$$(6.11) \quad \| \| w \| \| = \| (N_q(w_{1.}), \dots, N_q(w_{p.})) \|, \quad \forall w \in \mathbb{R}^{m \times q}.$$

Lemme 6.3.3. *Supposons que $\|\cdot\|$ est donnée par (6.10) ou (6.11). On a*

$$(6.12) \quad \forall y \in \mathbb{R}^m, \forall z \in \mathbb{R}^q, \quad \|\|yz^T\|\|^* = N_q^*(z)\|y\|^*.$$

Preuve. Dans ces deux cas, ceci résulte immédiatement du lemme 6.3.2. \square

Dans ce cas, on voit que (H'_I) est vérifiée avec $N = N_q^*$.

6.3.3. Une seconde réduction

Compléments d'Analyse convexe. On complète légèrement ce qui a été rappelé au paragraphe 6.3.1 de la sous-section 6.3.1.

Soit C un sous-ensemble convexe de F . On dit que $y \in C$ est un *point extrémal* de C si $C \setminus \{y\}$ est encore un convexe. On notera $\text{Ext}(C)$ l'ensemble des points extrémaux de l'ensemble C .

Nous allons nous servir du résultat suivant

Théorème 6.3.4 ([61], Cor. 18.5.1). *Si C est un convexe fermé borné alors*

$$C = \text{conv}(\text{Ext}(C)).$$

On reprend le raisonnement qui a été fait dans le paragraphe 6.3.1 de la sous-section précédente mais en utilisant cette fois ci que

$$\mathbb{B}_{\|\cdot\|^*} = \text{conv}(\text{Ext}(\mathbb{B}_{\|\cdot\|^*})).$$

On en déduit que

$$\alpha \mathbb{B}_{\|\cdot\|^*} = \bigcap_{y^* \in \text{Ext}(\mathbb{B}_{\|\cdot\|^*})} [\langle \cdot, y^* \rangle \leq \alpha].$$

Une seconde formule pour le rayon de stabilité.

Théorème 6.3.5. *Supposons que l'opérateur T est surjectif. Alors on a*

$$d_{\|\cdot\|, \|\cdot\|^*}(0, S_\alpha^c) = \inf_{y^* \in \text{Ext}(\mathbb{B}_{\|\cdot\|^*})} \frac{(\alpha - \langle a, y^* \rangle)^+}{\|\|T^*y^*\|\|^*}.$$

Considérons l'hypothèse

$$(H_{II}) \quad \text{Il existe } N(T) \text{ tel que : } \forall y^* \in \text{Ext}(\mathbb{B}_{\|\cdot\|^*}), \quad \|\|T^*y^*\|\|^* = N(T).$$

qui est impliquée (donc plus faible) par l'hypothèse (H_I) . On a

Théorème 6.3.6. *Supposons que (H_{II}) est vérifiée et que $T \in \mathcal{L}(W, F)$ est surjectif. Alors on a*

$$(6.13) \quad d_{\|\cdot\|, \|\cdot\|^*}(0, S_\alpha^c) = \frac{(\alpha - \|a\|)^+}{N(T)}.$$

6.3.4. Retour à la régression robuste dans le cas non structuré

On se replace dans le cadre de la sous-section 6.3.2. Nous supposons que \mathbb{R}^m est muni d'une d -norme ($d \in [1, m]$, voir la définition ci-dessous).

Comme précédemment, dans le cas qui nous intéresse, la propriété (H_{II}) devient :

$$(H'_{II}) \quad \text{Il existe } N : \mathbb{R}^q \rightarrow \mathbb{R}_+ \text{ tel que : } \forall z \in \mathbb{R}^q, \forall y \in \text{Ext}(\mathbb{B}_{\|\cdot\|_*}), \quad \|\|yz^T\|\|^* = N(z).$$

L'hypothèse (H'_{II}) reste impliquée par l'hypothèse (H'_I) .

Lemme 6.3.4. *Si (H'_{II}) est vérifiée alors N est une norme sur \mathbb{R}^q .*

Preuve. Par définition de N , on a

$$\begin{aligned} N(z) = 0 &\iff yz^T = 0, \quad \forall y \in \text{Ext}(\mathbb{B}_{\|\cdot\|_*}), \\ &\iff yz^T = 0, \quad \forall y \in \mathbb{B}_{\|\cdot\|_*}, \\ &\iff yz^T = 0, \quad \forall y \in \mathbb{R}^m, \\ &\iff z = 0. \end{aligned}$$

Pour tout $\lambda \in \mathbb{R}$, $z \in \mathbb{R}^q$, on a

$$\begin{aligned} N(\lambda z) &= \|\| \lambda yz^T \|\|^* \quad \text{pour tout } y \in \text{Ext}(\mathbb{B}_{\|\cdot\|_*}), \\ &= |\lambda| \|\| yz^T \|\|^*, \\ &= |\lambda| N(z). \end{aligned}$$

Pour tout $u \in \mathbb{R}^q$, $v \in \mathbb{R}^q$, on a

$$\begin{aligned} N(u+v) &= \|\| y(u+v)^T \|\|^* \quad \text{pour tout } y \in \text{Ext}(\mathbb{B}_{\|\cdot\|_*}), \\ &= \|\| yu^T + yv^T \|\|^*, \\ &\leq \|\| yu^T \|\|^* + \|\| yv^T \|\|^* \\ &= N(u) + N(v). \end{aligned}$$

□

Nous allons voir que dans deux cas de définition de la norme $\|\|\cdot\|\|$ sur $\mathbb{R}^{m \times q}$, la propriété (H'_{II}) est vérifiée et donc la formule (6.13) permettant le calcul du rayon de stabilité est vraie.

d-norme. Soit $d \in [1, n]$. La d -norme, notée $\|\cdot\|_{(d)}$, et introduite en Optimisation robuste dans [12], est une norme sur \mathbb{R}^n , définie, pour $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, par

$$(6.14) \quad \|x\|_{(d)} = \begin{cases} \sum_{i=1}^{\lfloor d \rfloor} |x_{(i)}| + (d - \lfloor d \rfloor) |x_{(\lfloor d \rfloor + 1)}| & \text{si } d \in [1, n[, \\ \sum_{i=1}^n |x_{(i)}| & \text{si } d = n, \end{cases}$$

où $|x_{(i)}|$ est la $i^{\text{ème}}$ plus grande coordonnée de $|x| := (|x_1|, \dots, |x_n|)$ et $[d]$ désigne la partie entière de d . Il est clair que $\|\cdot\|_{(1)} = \|\cdot\|_\infty$ et $\|\cdot\|_{(n)} = \|\cdot\|_1$.

La norme duale de la d -norme a été explicitée dans [12] (Proposition 2). Plus exactement, pour tout $d \in [1, n]$, on a

$$(6.15) \quad \forall x \in \mathbb{R}^n, \quad \|x\|_{(d)}^* = \max \left\{ \|x\|_\infty, \|x\|_1/d \right\}.$$

Lemme 6.3.5. *Pour tout $d \in [1, n]$, $x \in \mathbb{R}^n$, on a*

$$(6.16) \quad x \in \text{Ext } \mathbb{B}_{\|\cdot\|_{(d)}^*} \implies \|x\|_\infty = 1 = \|x\|_1/d.$$

Preuve. Posons $\Lambda = \{1; -1\}^n$ et $\Omega = \{e_1, \dots, e_n, -e_1, \dots, -e_n\}$ avec e_i le $i^{\text{ème}}$ vecteur de la base canonique de \mathbb{R}^n . D'après (6.15), on a

$$\mathbb{B}_{\|\cdot\|_{(d)}^*} = \left(\bigcap_{e \in \Lambda} [\langle e, \cdot \rangle \leq d] \right) \cap \left(\bigcap_{e \in \Omega} [\langle e, \cdot \rangle \leq 1] \right),$$

qui est un ensemble polyédral. Par conséquent, $x \in \text{Ext } \mathbb{B}_{\|\cdot\|_{(d)}^*}$ si et seulement s'il existe k vecteurs $e^1, \dots, e^k \in \Lambda$ et $n - k$ vecteurs $e^{k+1}, \dots, e^n \in \Omega$ tels que : la famille $\{e^1, \dots, e^n\}$ est libre, $\langle e^i, x \rangle = d$ pour $i \in \llbracket 1, k \rrbracket$ et $\langle e^i, x \rangle = 1$ pour $i \in \llbracket k+1, n \rrbracket$. On distingue trois cas possibles pour k :

- $k = 0$. On a $x \in \Lambda$ et $d \geq \langle x, x \rangle = n$. Par conséquent, $d = n$ et $\|x\|_\infty = 1 = \|x\|_1/d$.
- $k = n$. On a $x \in \text{Ext}(d\mathbb{B}_{\|\cdot\|_1}) = d\text{Ext } \mathbb{B}_{\|\cdot\|_1}$. Comme $x \in \mathbb{B}_{\|\cdot\|_\infty}$ et $d \in [1; n]$ alors $d = 1$ et $x \in \text{Ext } \mathbb{B}_{\|\cdot\|_1}$. D'où $\|x\|_\infty = 1 = \|x\|_1/d$.
- $1 \leq k < n$. On a $\langle e^1, x \rangle = d$ et $\langle e^n, x \rangle = 1$. En appliquant l'inégalité de Hölder, on obtient $d \leq \|x\|_1 \|e^1\|_\infty = \|x\|_1$ et $1 \leq \|e^n\|_1 \|x\|_\infty = \|x\|_\infty$. D'où $\|x\|_\infty = 1 = \|x\|_1/d$. □

Soit N_m une norme absolue sur \mathbb{R}^m et N_q une norme sur \mathbb{R}^q . On définit la norme $\|\|\cdot\|\|$ sur $\mathbb{R}^{m \times q}$ par

$$(6.17) \quad \forall w \in \mathbb{R}^{m \times q}, \quad \|\|w\|\| = N_m(N_q(w_1), \dots, N_q(w_m)).$$

On sait que pour tout $y \in \mathbb{R}^m$, $z \in \mathbb{R}^q$, on a

$$(6.18) \quad \|\|yz^T\|\|^* = N_m^*(y)N_q^*(z).$$

Soit $d' \in [1, m]$. On suppose que $N_m = \|\cdot\|_{(d')}$. Dans ce cas, d'après le lemme 6.3.5, pour tout $y \in \text{Ext } \mathbb{B}_{\|\cdot\|_{(d')}^*}$, $z \in \mathbb{R}^q$, on a

$$(6.19) \quad \|\|yz^T\|\|^* = N_m^*(y)N_q^*(z) = \max \left\{ 1, \frac{d}{d'} \right\} N_q^*(z).$$

L'hypothèse (H'_{II}) est donc vérifiée.

Soit $d' \in [1, mq]$, on définit la d' -norme $\|\cdot\|_{(d')}$ sur $\mathbb{R}^{m \times q}$ comme suit :

$$(6.20) \quad \forall w \in \mathbb{R}^{m \times q}, \quad \|\|w\|\|_{(d')} = \|\text{vect}(w)\|_{(d')}.$$

On déduit immédiatement de (6.15) que

$$(6.21) \quad \forall y \in \mathbb{R}^m, \forall z \in \mathbb{R}^q, \quad \|\|yz^T\|\|_{(d')}^* = \max \left\{ \|y\|_\infty \|z\|_\infty, \|y\|_1 \|z\|_1 / d' \right\}, \quad .$$

En particulier, d'après le lemme 6.3.5, on a

$$(6.22) \quad \forall y \in \text{Ext } \mathbb{B}_{\|\cdot\|_{(d')}^*}, \forall z \in \mathbb{R}^q, \quad \|\|yz^T\|\|_{(d')}^* = \max \left\{ \|z\|_\infty, \frac{d}{d'} \|z\|_1 \right\}.$$

L'hypothèse (H'_{II}) est donc vérifiée.

Ces deux derniers exemples montrent que l'hypothèse (H'_{II}) peut être vérifiée sans que la norme $\|\| \cdot \| \|$ soit directement fonction de $\|\cdot\|$.

6.3.5. Application au calcul du rayon de stabilité dans le cas de la régression robuste

On applique maintenant les résultats précédents pour déterminer une formule explicite du rayon de stabilité dans le cas de la régression robuste vu dans la section 6.2.

On suppose dans la suite que l'hypothèse (H'_{II}) est vérifiée. On obtient donc

$$(6.23) \quad r_\alpha(x, \bar{w}) = \frac{(\alpha - \|\bar{A}x - \bar{b}\|)^+}{N(x, -1)}, \quad \forall x \in \mathbb{R}^n,$$

où $N(x, -1) = N((x; -1))$. Dans ce cas, le programme (\mathcal{R}_r^α) dévient

$$(\widehat{\mathcal{R}}_r^\alpha) \quad \text{Maximiser } \frac{(\alpha - \|\bar{A}x - \bar{b}\|)^+}{N(x, -1)}, \quad x \in X,$$

qui est un programme fractionnaire.

6.4. Étude de la contrepartie robuste : résolution de $(\widehat{\mathcal{R}}_r^\alpha)$

Notons tout d'abord que si X est un compacte, alors l'algorithme de Dinkelbach permet de résoudre le problème $(\widehat{\mathcal{R}}_r^\alpha)$ ([26], [28]). Cependant, dans les applications il arrive très souvent que X soit l'espace tout entier (régression sans contrainte par exemple), un polyèdre (régression sous contraintes affines) ou plus généralement l'image réciproque d'un cône convexe par une application affine.

6.4.1. Transformation du programme $(\widehat{\mathcal{R}}_r^\alpha)$ en un programme convexe

On considère le programme $(\widehat{\mathcal{R}}_r^\alpha)$ et on suppose que X est l'image réciproque d'un cône convexe par une application affine. Plus précisément, on suppose que

$$(6.24) \quad X := \mathcal{G}^-(K) = \{x \in \mathbb{R}^n \mid \mathcal{G}(x) \in K\},$$

où $\mathcal{G} : \mathbb{R}^n \rightarrow \mathbb{R}^p$ est une fonction affine, et $K \subset \mathbb{R}^p$ est un cône convexe fermé. On réécrit alors le problème $(\widehat{\mathcal{R}}_r^\alpha)$ comme suit

$$(\widehat{\mathcal{R}}_r^\alpha) \quad \text{Maximiser } \frac{(\alpha - \|\bar{A}x - \bar{b}\|)^+}{N(x, -1)}, \quad \mathcal{G}(x) \in K.$$

Notons f (resp. g) le numérateur (resp. dénominateur) de la fonction objectif de $(\widehat{\mathcal{R}}_r^\alpha)$:

$$\forall x \in \mathbb{R}^n, \quad f(x) = (\alpha - \|\bar{A}x - \bar{b}\|)^+ \quad \text{et} \quad g(x) = N(x, -1).$$

On définit \tilde{X} par

$$\tilde{X} := \{(y, t) \in \mathbb{R}^n \times \mathbb{R} \mid t > 0, \mathcal{G}(y/t) \in K\}.$$

On note \tilde{f} (resp. \tilde{g}) la perspective de f (resp. g) (voir l'annexe 1.3 pour la définition de la perspective d'une fonction), puis on considère le problème

$$(\tilde{\mathcal{R}}_r^\alpha) \quad \text{Maximiser } \tilde{f}(y, t), \quad (y, t) \in \tilde{X}_\leq,$$

où l'ensemble \tilde{X}_\leq des solutions admissibles de $(\tilde{\mathcal{R}}_r^\alpha)$ est défini par

$$\tilde{X}_\leq := \tilde{X} \cap [\tilde{g} \leq 1].$$

D'après (6.1), $\alpha > \inf_{x \in X} \|\bar{A}x - \bar{b}\|$. Il existe donc $x \in X$ tel que $\alpha - \|\bar{A}x - \bar{b}\| > 0$. Par conséquent, la fonction f est positive et non identiquement nulle. On déduit immédiate du lemme 1.3.3, la proposition suivante.

Proposition 6.4.1. *Les problèmes $(\widehat{\mathcal{R}}_r^\alpha)$ et $(\tilde{\mathcal{R}}_r^\alpha)$ sont liés par*

$$v(\widehat{\mathcal{R}}_r^\alpha) = v(\tilde{\mathcal{R}}_r^\alpha) \quad \text{et} \quad \text{Argmax}(\tilde{\mathcal{R}}_r^\alpha) = \psi(\text{Argmax}(\widehat{\mathcal{R}}_r^\alpha)),$$

où $\psi : X \rightarrow \tilde{X}_= := \tilde{X} \cap [\tilde{g} = 1]$ est la transformation de Charnes-Cooper définie par

$$\forall x \in X, \quad \psi(x) = \left(\frac{x}{N(x, -1)}, \frac{1}{N(x, -1)} \right).$$

Observons que par définition des perspectives \tilde{f} et \tilde{g} , on a

$$\forall (y, t) \in \tilde{X}, \quad \tilde{f}(y, t) = t(\alpha - \|\bar{A}y - t\bar{b}\|/t)^+ = (t\alpha - \|\bar{A}y - t\bar{b}\|)^+,$$

$$\forall (y, t) \in \tilde{X}, \quad \tilde{g}(y, t) = tN(y/t, -1) = N(y, -t), \quad \text{car } N \text{ est une norme.}$$

Le problème $(\tilde{\mathcal{R}}_r^\alpha)$ s'écrit alors de manière explicite comme suit

$$(\tilde{\mathcal{R}}_r^\alpha) \quad \begin{array}{l} \text{Maximiser} \quad (t\alpha - \|\bar{A}y - t\bar{b}\|)^+ \\ \\ \text{s.l.c} \quad \begin{cases} \mathcal{G}(y/t) \in K, \\ N(y, -t) \leq 1, \\ (y, t) \in \mathbb{R}^n \times \mathbb{R}_+^*. \end{cases} \end{array}$$

Puisque $\mathcal{G} : \mathbb{R}^n \rightarrow \mathbb{R}^p$ est affine, il existe $G \in \mathbb{R}^{p \times n}$ et $G_0 \in \mathbb{R}^p$ tels que

$$\mathcal{G}(x) = Gx + G_0, \quad \forall x \in \mathbb{R}^n.$$

Par conséquent, pour tout $(y, t) \in \mathbb{R}^n \times \mathbb{R}_+^*$, on a

$$\mathcal{G}(y/t) \in K \iff Gy + tG_0 \in K, \quad \text{car } K \text{ est un c\^one.}$$

Par ailleurs, d'apr\es (6.1), il existe $x \in X$ tel que $\tilde{f}(\psi(x)) > 0$. On peut donc r\ecrire le programme $(\tilde{\mathcal{R}}_r^\alpha)$ comme suit

$$\begin{aligned} & \text{Maximiser} \quad t\alpha - \|\bar{A}y - t\bar{b}\| \\ (\tilde{\mathcal{R}}_r^\alpha) \quad & \text{s.l.c} \quad \begin{cases} Gy + tG_0 \in K, \\ N(y, -t) \leq 1, \\ (y, t) \in \mathbb{R}^n \times \mathbb{R}_+^*, \end{cases} \end{aligned}$$

ou encore, en reportant le crit\ere dans les contraintes, on obtient

$$\begin{aligned} & \text{Maximiser} \quad \tau \\ (\tilde{\mathcal{R}}_r^\alpha) \quad & \text{s.l.c} \quad \begin{cases} \tau - t\alpha + \|\bar{A}y - t\bar{b}\| \leq 0, \\ Gy + tG_0 \in K, \\ N(y, -t) \leq 1, \\ (\tau, y, t) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}_+^*, \end{cases} \end{aligned}$$

qui est \evidemment un programme d'optimisation convexe.

Remarque 6.4.1. *On sait que maximiser une fonction convexe sur un sous-ensemble d'un espace vectoriel topologique localement convexe revient \a maximiser la m\eme fonction sur l'enveloppe convexe ferm\ee de ce sous-ensemble ([72], Corollaire 2 ou [2], Proposition 11.3.4 dans le cas de la dimension finie). On peut donc (sans perdre de g\en\eralit\e) remplacer la contrainte $t > 0$ par $t \geq 0$.*

Rappelons qu'une fonction $h : \mathbb{R}^n \rightarrow \mathbb{R}$ est dite poly\edrique si son \epigraphe est un poly\edre. Il en r\esulte que si les normes $\|\cdot\|$ et N sont poly\edriques et K un c\one poly\edral alors $(\tilde{\mathcal{R}}_r^\alpha)$ se ram\ene sans peine \a un programme lin\eaire. Les d -normes sont des exemples de normes poly\edriques.

6.4.2. Un exemple où $(\tilde{\mathcal{R}}_r^\alpha)$ se ramène à un programme linéaire

Soit $d \in [1, m]$ et $d' \in [1, m(n+1)]$. On suppose que $\|\cdot\| = \|\cdot\|_{(d)}$ et $\|\cdot\| = \|\cdot\|_{(d)}$. Il résulte de (6.22) que la norme N définie par (H'_{II}) est donnée par

$$N(z) = \max \left\{ \|z\|_\infty, \frac{d}{d'} \|z\|_1 \right\}, \quad \forall z \in \mathbb{R}^{n+1}.$$

Dans ce cas, le programme $(\tilde{\mathcal{R}}_r^\alpha)$ dévient

$$(\tilde{\mathcal{R}}_r^\alpha) \quad \begin{array}{l} \text{Maximiser} \quad \tau \\ \\ \text{s.l.c} \quad \left\{ \begin{array}{l} \tau - t\alpha + \|\bar{A}y - t\bar{b}\|_{(d)} \leq 0 \\ Gy + tG_0 \in K, \\ \|(y, -t)\|_\infty \leq 1, \\ d\|(y, -t)\|_1 \leq d', \\ (\tau, y, t) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}_+. \end{array} \right. \end{array}$$

On suppose que $K = \mathbb{R}_+^p$ puis on introduit le programme linéaire suivant

$$(\mathcal{P}_{\text{lin}}) \quad \begin{array}{l} \text{Maximiser} \quad \tau \\ \\ \text{s.l.c} \quad \left\{ \begin{array}{l} \tau - t\alpha + rd + \langle \mathbb{1}_m, u \rangle \leq 0 \\ -t\bar{b} + \bar{A}y - r\mathbb{1}_m - u \leq 0 \\ t\bar{b} - \bar{A}y - r\mathbb{1}_m - u \leq 0 \\ -Gy - tG_0 \leq 0 \\ y - \mathbb{1}_n \leq 0 \\ y + \mathbb{1}_n \leq 0 \\ t - 1 \leq 0 \\ td + snd + d\langle \mathbb{1}_m, v \rangle - d' \leq 0 \\ -y - s\mathbb{1}_m - v \leq 0 \\ y - s\mathbb{1}_m - v \leq 0 \\ (\tau, y, t, r, u, s, v) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^m \times \mathbb{R}_+ \times \mathbb{R}^m. \end{array} \right. \end{array}$$

où $\mathbb{1}_p$ désigne le vecteur de \mathbb{R}^p dont toutes les coordonnées sont égales à 1.

Rappelons la propriété suivante de la d -norme sur \mathbb{R}^p .

Lemme 6.4.1 ([12]). *Soit $d \in [1, p]$. Pour tout $\lambda \in \mathbb{R}$, $z \in \mathbb{R}^p$, on a :*

$$(6.25) \quad \|z\|_{(d)} \leq \lambda \iff \exists r \in \mathbb{R}_+, u \in \mathbb{R}_+^p : \begin{cases} rd + \langle \mathbb{1}_p, u \rangle \leq \lambda \\ z - r\mathbb{1}_p - u \leq 0 \\ -z - r\mathbb{1}_p - u \leq 0. \end{cases}$$

On déduit de (6.25) que (τ, y, t) est une solution admissible de $(\tilde{\mathcal{R}}_r^\alpha)$ si et seulement s'il existe $(r, u, s, v) \in \mathbb{R}_+ \times \mathbb{R}^m \times \mathbb{R}_+ \times \mathbb{R}^m$ tel que (τ, y, t, r, u, s, v) est une solution admissible de (\mathcal{P}_{lin}) . La proposition suivante est alors immédiate.

Proposition 6.4.2. *Les programmes $(\tilde{\mathcal{R}}_r^\alpha)$ et (\mathcal{P}_{lin}) sont liés par*

$$v(\mathcal{P}_{lin}) = v(\tilde{\mathcal{R}}_r^\alpha) \quad \text{et} \quad \text{Argmax}(\tilde{\mathcal{R}}_r^\alpha) = \Pi_{\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}_+} \text{Argmax}(\mathcal{P}_{lin}).$$

6.5. Lien avec l'approche de l'optimisation robuste

Dans cette section, on compare l'ensemble solution de la contrepartie robuste de (\mathcal{P}_r) au sens de l'optimisation robuste et celui au sens du rayon de robustesse. On reprend les données et les hypothèses de la section 6.4.

6.5.1. Contrepartie robuste au sens de l'optimisation robuste de (\mathcal{P}_r)

On se donne $\beta \in \overline{\mathbb{R}}_+$. On associe à (\mathcal{P}_r) «the worst case lost function» $wc_\beta : \mathbb{R}^n \times W \rightarrow \mathbb{R}$ définie par

$$(6.26) \quad wc_\beta(x, \bar{w}) := \sup_{\|w\| \leq \beta} \|\bar{A}x - \bar{b} + w(x; -1)\|.$$

La contrepartie robuste de (\mathcal{P}_r) au sens de l'optimisation robuste est définie par

$$(\mathcal{P}_r^\beta) \quad \text{Minimiser } wc_\beta(x, \bar{w}), \quad x \in X.$$

Par définition de wc_β , on a

$$\begin{aligned} wc_\beta(x, \bar{w}) &= \sup_{\|w\| \leq \beta} \|\bar{A}x - \bar{b} + w(x; -1)\| \\ &= \sup_{\|w\| \leq \beta} \sup_{\|y\|^* \leq 1} \langle \bar{A}x - \bar{b} + w(x; -1), y \rangle \\ &= \sup_{\|y\|^* \leq 1} \left\{ \langle \bar{A}x - \bar{b}, y \rangle + \beta \sup_{\|w\| \leq 1} \langle w, y(x; -1)^T \rangle \right\} \\ &= \sup_{\|y\|^* \leq 1} \left\{ \langle \bar{A}x - \bar{b}, y \rangle + \beta \| \|y(x; -1)^T\|^* \right\} \\ &= \sup_{y \in B_{\| \cdot \|^*}} \left\{ \langle \bar{A}x - \bar{b}, y \rangle + \beta \| \|y(x; -1)^T\|^* \right\}. \end{aligned}$$

Il résulte donc de (H'_{II}) que

$$wc_\beta(x, \bar{w}) = \|\bar{A}x - \bar{b}\| + \beta N(x, -1), \quad \forall x \in \mathbb{R}^n.$$

Le problème (\mathcal{P}_r^β) s'écrit alors

$$(\mathcal{P}_r^\beta) \quad \text{Minimiser } \|\bar{A}x - \bar{b}\| + \beta N(x, -1), \quad x \in X.$$

Rappelons que dans ce cas, la contrepartie robuste de (\mathcal{P}_r) au sens du rayon de stabilité est donnée par

$$(\widehat{\mathcal{R}}_r^\alpha) \quad \text{Maximiser } \frac{(\alpha - \|\bar{A}x - \bar{b}\|)^+}{N(x, -1)}, \quad x \in X.$$

6.5.2. Comparaison de $(\widehat{\mathcal{P}}_r^\alpha)$ et (\mathcal{P}_r^β)

À priori, s'il n'y a pas de relation entre α et β , il en est de même pour les problèmes $(\widehat{\mathcal{P}}_r^\alpha)$ et (\mathcal{P}_r^β) . La question est de savoir comment choisir α et β afin de pouvoir comparer les ensembles des solutions optimales des deux problèmes? Pour répondre à cette question, posons

$$\alpha_0 = \inf_{x \in X} \|\bar{A}x - \bar{b}\| \quad , \quad I = [\alpha_0, +\infty] \quad \text{et} \quad J = [0; +\infty],$$

puis définissons les fonctions $v_r : I \rightarrow J$ et $v_{wc} : J \rightarrow I$ par

$$v_r(\alpha) = \sup_{x \in X} \frac{\alpha - \|\bar{A}x - \bar{b}\|}{N(x, -1)}, \quad \forall \alpha \in I,$$

$$v_{wc}(\beta) = \inf_{x \in X} \|\bar{A}x - \bar{b}\| + \beta N(x, -1), \quad \forall \beta \in \mathbb{R}_+^*.$$

Proposition 6.5.1. *Les fonctions v_r et v_{wc} vérifient les propriétés suivantes :*

- (1) $v_r(\alpha_0) = 0$, $v_r(+\infty) = +\infty$, $v_{wc}(0) = \alpha_0$ et $v_{wc}(+\infty) = +\infty$,
- (2) les fonctions v_r et v_{wc} sont croissantes,
- (3) la fonction v_r est convexe et la fonction v_{wc} est concave,
- (4) les fonctions v_r et v_{wc} sont continues.

Preuve. Par définition de α_0 , on a $v_r(\alpha_0) \leq 0$. Si cette inégalité était stricte alors il existerait $\eta < 0$ tel que $\alpha_0 - \|\bar{A}x - \bar{b}\| < \eta$, $\forall x \in X$, et donc $\alpha_0 \leq \eta + \alpha_0$. Par conséquent $0 \leq \eta < 0$ ce qui est absurde. D'où $v_r(\alpha_0) = 0$.

Il est clair que $v_r(+\infty) = +\infty$, $v_{wc}(+\infty) = +\infty$ et $v_{wc}(0) = \alpha_0$. D'où (1).

Soient $\alpha_1, \alpha_2 \in I$ tel que $\alpha_1 \leq \alpha_2$. Comme N est une norme, on a

$$\frac{\alpha_1 - \|\bar{A}x - \bar{b}\|}{N(x, -1)} \leq \frac{\alpha_2 - \|\bar{A}x - \bar{b}\|}{N(x, -1)}, \quad \forall x \in X.$$

En passant au sup sur X , on obtient $v_r(\alpha_1) \leq v_r(\alpha_2)$. D'où la croissance de la fonction v_r . Par un raisonnement analogue, on montre que v_{wc} est croissante. D'où (2).

La fonction v_r est convexe en tant que borne supérieure de fonctions convexes et la fonction v_{wc} est concave en tant que borne inférieure de fonctions concaves. D'où (3). D'après (3), les fonctions v_r et v_{wc} sont continues sur l'intérieur de leurs domaines. D'où (4). \square

Notons que pour tout $\alpha \in I$, $v_r(\alpha)$ est la valeur du problème $(\widehat{\mathcal{P}}_r^\alpha)$. De même, pour tout $\beta \in J$, $v_{wc}(\beta)$ est la valeur du problème (\mathcal{P}_r^β) .

Proposition 6.5.2. *Les fonction v_r et v_{wc} vérifient les relations suivantes :*

- (1) $\forall \alpha \in I$, $v_{wc}(v_r(\alpha)) = \alpha$,
- (2) $\forall \beta \in J$, $v_r(v_{wc}(\beta)) = \beta$.

Preuve. (1). Soit $\alpha \in I$. Si $\alpha = \alpha_0$ ou $\alpha = +\infty$ alors le résultat est évident. Supposons que $\alpha_0 < \alpha < +\infty$. Par définition de v_r , on a

$$v_r(\alpha) \geq \frac{\alpha - \|\bar{A}x - \bar{b}\|}{N(x, -1)}, \quad \forall x \in X.$$

Par conséquent,

$$\|\bar{A}x - \bar{b}\| + v_r(\alpha)N(x, -1) \geq \alpha, \quad \forall x \in X.$$

Donc

$$v_{wc}(v_r(\alpha)) = \inf_{x \in X} \|\bar{A}x - \bar{b}\| + v_r(\alpha)N(x, -1) \geq \alpha.$$

Si $v_r(\alpha) = 0$ alors $v_{wc}(v_r(\alpha)) = v_{wc}(0) = \alpha_0 < \alpha$ absurde. Donc $v_r(\alpha) > 0$.

Pour tout $\beta \in]0, v_r(\alpha)[$, il existe $x_\beta \in X$ tel que

$$\beta < \frac{\alpha - \|\bar{A}x_\beta - \bar{b}\|}{N(x_\beta, -1)}.$$

Or,

$$\begin{aligned} \beta < \frac{\alpha - \|\bar{A}x_\beta - \bar{b}\|}{N(x_\beta, -1)} &\implies \beta N(x_\beta, -1) + \|\bar{A}x_\beta - \bar{b}\| < \alpha \\ &\implies v_{wc}(\beta) < \alpha. \end{aligned}$$

Par conséquent, pour tout $\beta \in]0, v_r(\alpha)[$, on a $v_{wc}(\beta) < \alpha$. Comme v_{wc} est continue à gauche en $v_r(\alpha)$, il en résulte que

$$v_{wc}(v_r(\alpha)) = \lim_{\substack{\beta \rightarrow v_r(\alpha) \\ \beta < v_r(\alpha)}} v_{wc}(\beta) \leq \alpha.$$

(2). Soit $\beta \in J$. Si $\beta = 0$ ou $\beta = +\infty$ alors le résultat est immédiat. On suppose donc que $0 < \beta < +\infty$. Par définition de v_{wc} , on a

$$v_{wc}(\beta) \leq \|\bar{A}x - \bar{b}\| + \beta N(x, -1), \quad \forall x \in X.$$

Donc

$$\frac{v_{wc}(\beta) - \|\bar{A}x - \bar{b}\|}{N(x, -1)} \leq \beta, \quad \forall x \in X, \quad \text{d'où } v_r(v_{wc}(\beta)) \leq \beta.$$

Si $v_{wc}(\beta) = +\infty$ alors $v_r(v_{wc}(\beta)) = v_r(+\infty) = +\infty > \beta$ absurde. Donc $v_{wc}(\beta) < +\infty$.

Pour tout $\alpha \in]v_{wc}(\beta), +\infty[$, il existe $x_\alpha \in X$ tel que

$$\|\bar{A}x_\alpha - \bar{b}\| + \beta N(x_\alpha, -1) < \alpha.$$

Or

$$\begin{aligned} \|\bar{A}x_\alpha - \bar{b}\| + \beta N(x_\alpha, -1) < \alpha &\implies \beta < \frac{\alpha - \|\bar{A}x_\alpha - \bar{b}\|}{N(x_\alpha, -1)} \\ &\implies \beta < v_r(\alpha). \end{aligned}$$

Donc pour tout $\alpha \in]v_{wc}(\beta), +\infty[$, on a $\beta < v_r(\alpha)$. Comme v_r est continue à droite en $v_{wc}(\beta)$, on en déduit que

$$\beta \leq \lim_{\substack{\alpha \rightarrow v_{wc}(\beta) \\ \alpha > v_{wc}(\beta)}} v_r(\alpha) = v_r(v_{wc}(\beta)).$$

□

Remarque 6.5.1. *La proposition 6.5.2 montre que les fonctions v_{wc} et v_r sont des bijections réciproques l'une de l'autre.*

Proposition 6.5.3. *On a :*

- (1) $\forall \beta \in J, \quad \text{Argmin}(\mathcal{P}_r^\beta) = \text{Argmax}(\widehat{\mathcal{R}}_r^{v_{wc}(\beta)});$
- (2) $\forall \alpha \in I, \quad \text{Argmax}(\widehat{\mathcal{R}}_r^\alpha) = \text{Argmin}(\mathcal{P}_r^{v_r(\alpha)}).$

Preuve. Juxtaposer les propositions 6.5.2 et 4.3.9. □

Remarque 6.5.2. *On peut déduire de la proposition 6.5.2 une preuve directe de la proposition 6.5.3 sans utiliser la proposition 4.3.9.*

6.6. Application à un problème de conception d'un réseau d'antennes

6.6.1. Description du problème

Une antenne est un dispositif électromagnétique permettant de recevoir et / ou d'émettre une onde électromagnétique. Une caractéristique importante d'une antenne monochromatique est son diagramme $Z(\delta)$ qui est une fonction complexe de la direction δ . Par exemple le module de $Z(\delta)$, $|Z(\delta)|$ est responsable de la densité directionnelle de l'énergie envoyée par l'antenne dans la direction δ , l'argument de $Z(\delta)$, $\arg Z(\delta)$ correspond à la phase initiale de l'onde propagée dans la direction δ . On peut amplifier le signal et décaler la phase initiale en multipliant le diagramme par un nombre complexe constant ([9]).

Le diagramme d'une antenne complexe (antenne composée de plusieurs antennes élémentaires) est la somme des digrammes des antennes élémentaires qui la composent.

Un problème type en conception de réseau d'antennes est le suivant : étant données n antennes de diagrammes $Z_1(\cdot), \dots, Z_n(\cdot)$, trouver des nombres complexes z_1, \dots, z_n tels que le diagramme du réseau

$$Z(\cdot) = \sum_{j=1}^n z_j Z_j(\cdot)$$

soit le plus proche possible d'un diagramme cible $Z_*(\cdot)$ donné. Autrement dit,

(AD) trouver $(z_1, \dots, z_n) \in \mathbb{C}^n$ tel que $\left| \sum_{j=1}^n z_j Z_j(\cdot) - Z_*(\cdot) \right|$ soit minimum.

Considérons un réseau d'antennes composé d'anneaux centrés à l'origine dans le plan XY. Le diagramme d'un tel anneau est une fonction réelle qui dépend uniquement que de l'angle d'altitude θ (angle entre la direction de l'onde et le plan XY) dont une expression analytique est

$$Z_k(\theta) = \frac{1}{2} \int_0^{2\pi} \cos(2\pi k \cos(\theta) \cos(\phi)) d\phi,$$

où k est le quotient du rayon de l'anneau par la longueur d'onde λ ([10]). On suppose qu'il y a n anneaux de rayons respectifs $1, \dots, n$ dans notre réseau. Notre objectif est de déterminer $(x_1, \dots, x_n) \in \mathbb{R}^n$ tel que

$$Z(\theta) = \sum_{j=1}^n x_j Z_{k_j}(\theta) \cong Z_*(\theta), \quad \forall \theta \in [0^\circ, 90^\circ].$$

En utilisant l'approximation de Tschebyshev (approximation uniforme), on obtient le problème d'optimisation suivant :

$$(AT) \quad \text{Minimiser } \sup_{\theta \in [0^\circ, 90^\circ]} \left| \sum_{j=1}^n x_j Z_{k_j}(\theta) - Z_*(\theta) \right|, \quad x \in \mathbb{R}^n.$$

Une approximation tractable de (AT) est

$$(ATT) \quad \text{Minimiser } \max_{1 \leq i \leq m} \left| \sum_{j=1}^n x_j Z_{k_j}(\theta_i) - Z_*(\theta_i) \right|, \quad x \in \mathbb{R}^n,$$

où $(\theta_1, \dots, \theta_m)$ est une subdivision de l'intervalle $[0^\circ, 90^\circ]$.

Posons

$$A = (a_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \quad \text{avec } a_{ij} = Z_{k_j}(\theta_i), \quad (i, j) \in \llbracket 1, m \rrbracket \times \llbracket 1, n \rrbracket,$$

$$b = (b_i)_{1 \leq i \leq m} \quad \text{avec } b_i = Z_*(\theta_i), \quad i \in \llbracket 1, m \rrbracket.$$

Le problème (ATT) s'écrit alors comme suit :

$$(ATT) \quad \text{Minimiser } \|Ax - b\|_\infty, \quad x \in \mathbb{R}^n,$$

qui est un problème de régression linéaire en norme infinie que l'on peut transformer sans peine en un programme linéaire.

6.6.2. Illustration numérique

Pour résoudre numérique le problème (ATT), on suppose que $n = 40$, $\lambda = 10$ et on définit la fonction cible $Z_* : [0^\circ, 90^\circ] \rightarrow \mathbb{R}$ par

$$Z_*(\theta) = \begin{cases} 0 & \text{si } 0^\circ \leq \theta \leq 70^\circ \\ \frac{0,95\theta - 66,5}{7} & \text{si } 70^\circ \leq \theta \leq 77^\circ \\ 0,95 & \text{si } 77^\circ \leq \theta \leq 90^\circ. \end{cases}$$

Avec ses données, une résolution numérique de (ATT) donne le diagramme nominal de la figure 1.

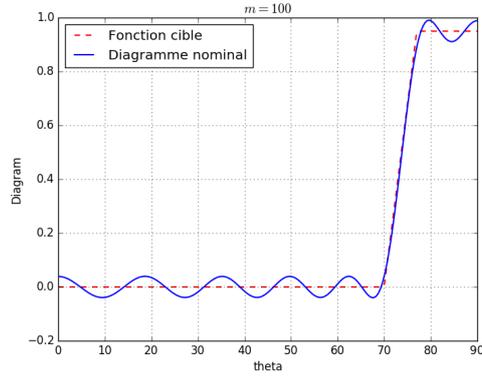


FIGURE 1. Diagramme nominal

Notons que les poids x_j , $j = 1, \dots, n$, correspondent à certaines caractéristiques physiques du dispositif et en tant que telles ne peuvent pas être utilisés avec leurs valeurs exactes. Il est donc important de savoir ce qui se passe si ces poids sont affectés par une erreur d'utilisation. Par exemple, si on remplace x_j par $(1 + \xi_j)x_j$ où les ξ_j sont indépendantes et suivent la loi uniforme sur $[-\epsilon, \epsilon]$ avec ϵ un réel positif donné.

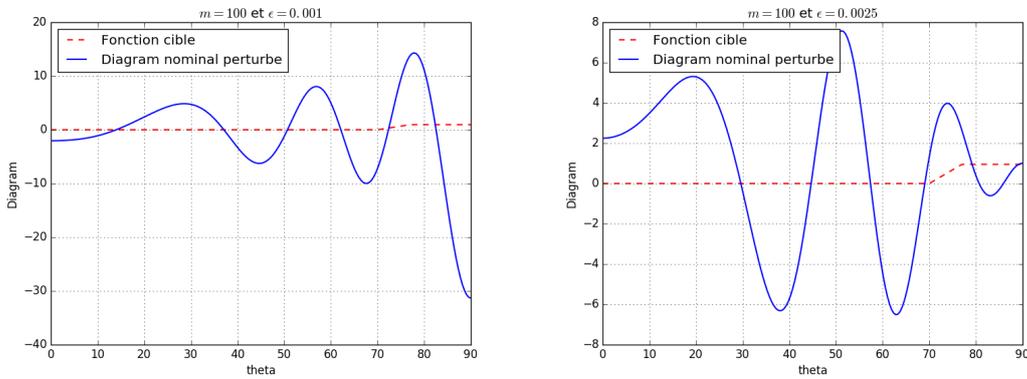


FIGURE 2. Diagramme nominal perturbé

On observe numériquement que même une petite erreur d'utilisation de ces coefficients peut avoir une conséquence désastreuse sur le diagramme du réseau.

6.6.3. Approche par la méthode du rayon de robustesse

Soit $\alpha \in \mathbb{R}$ tel que

$$(6.27) \quad \alpha > \inf_{x \in \mathbb{R}^n} \max_{1 \leq i \leq m} \left| \sum_{j=1}^n x_j Z_{k_j}(\theta_i) - Z_*(\theta_i) \right|$$

et $\| \cdot \|$ la norme définie sur $\mathbb{R}^{m \times (n+1)}$ par

$$\|B\| = \max_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n+1}} |b_{ij}|.$$

D'après le lemme 6.3.2, on a

$$(6.28) \quad \|yz^T\|^* = \|y\|_1 \|z\|_1, \quad \forall (y, z) \in \mathbb{R}^m \times \mathbb{R}^{n+1}.$$

En particulier l'hypothèse (H'_{II}) est vérifiée et la norme $N = \| \cdot \|_1$. Par conséquent, de ce qui précède, on sait que résoudre la contrepartie robuste au sens du rayon de robustesse du problème (AT), revient à résoudre le problème suivant :

$$(AT_r^\alpha) \quad \begin{array}{l} \text{Maximiser } \tau \\ \text{s.l.c } \left\{ \begin{array}{l} \tau - t\alpha + \max_{1 \leq i \leq m} \left| \sum_{j=1}^n y_j Z_{k_j}(\theta_i) - tZ_*(\theta_i) \right| \leq 0, \\ \|(y, -t)\|_1 \leq 1, \\ (\tau, t, y) \in \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}^n. \end{array} \right. \end{array}$$

Par ailleurs, on a d'une part

$$\tau - t\alpha + \max_{1 \leq i \leq m} \left| \sum_{j=1}^n y_j Z_{k_j}(\theta_i) - tZ_*(\theta_i) \right| \leq 0$$

si et seulement si

$$\left\{ \begin{array}{l} \tau - (\alpha + Z_*(\theta_i))t + \sum_{j=1}^n y_j Z_{k_j}(\theta_i) \leq 0, \quad i = 1, \dots, m \\ \tau - (\alpha - Z_*(\theta_i))t - \sum_{j=1}^n y_j Z_{k_j}(\theta_i) \leq 0, \quad i = 1, \dots, m. \end{array} \right.$$

D'autre part, d'après (6.25), on a

$$\|(y, -t)\|_1 \leq 1 \quad \text{et } t \geq 0$$

si et seulement s'il existe $r \in \mathbb{R}$ et $u \in \mathbb{R}^n$ tels que

$$\left\{ \begin{array}{l} t + nr + \sum_{j=1}^n u_j \leq 1 \\ r + y_j + u_j \geq 0, \quad j = 1, \dots, n \\ r - y_j + u_j \geq 0, \quad j = 1, \dots, n \\ u_j \geq 0, \quad j = 1, \dots, n \\ t \geq 0 \\ r \geq 0. \end{array} \right.$$

Par conséquent, le programme (\tilde{AT}_r^α) est équivalent au programme linéaire (AT_{lin}) dans le sens que

$$(6.29) \quad v(\tilde{AT}_r^\alpha) = v(AT_{\text{lin}}) \quad \text{et} \quad \text{Argmax}(\tilde{AT}_r^\alpha) = \Pi_{\mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}^n}(\text{Argmax}(AT_{\text{lin}})),$$

où le programme (AT_{lin}) est défini par

$$\begin{array}{l} \text{Maximiser} \quad \tau \\ \\ (AT_{\text{lin}}) \quad \text{s.l.c} \quad \left\{ \begin{array}{l} \tau - (\alpha + Z_*(\theta_i))t + \sum_{j=1}^n y_j Z_{k_j}(\theta_i) \leq 0, \quad i = 1, \dots, m \\ \tau - (\alpha - Z_*(\theta_i))t - \sum_{j=1}^n y_j Z_{k_j}(\theta_i) \leq 0, \quad i = 1, \dots, m \\ t + nr + \sum_{j=1}^n u_j \leq 1 \\ r + y_j + u_j \geq 0, \quad j = 1, \dots, n \\ r - y_j + u_j \geq 0, \quad j = 1, \dots, n \\ u_j \geq 0, \quad j = 1, \dots, n \\ (\tau, t, y, r, u) \in \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}_+ \times \mathbb{R}^n. \end{array} \right. \end{array}$$

En résolvant numériquement le programme (AT_{lin}) , on obtient le diagramme robuste suivant :

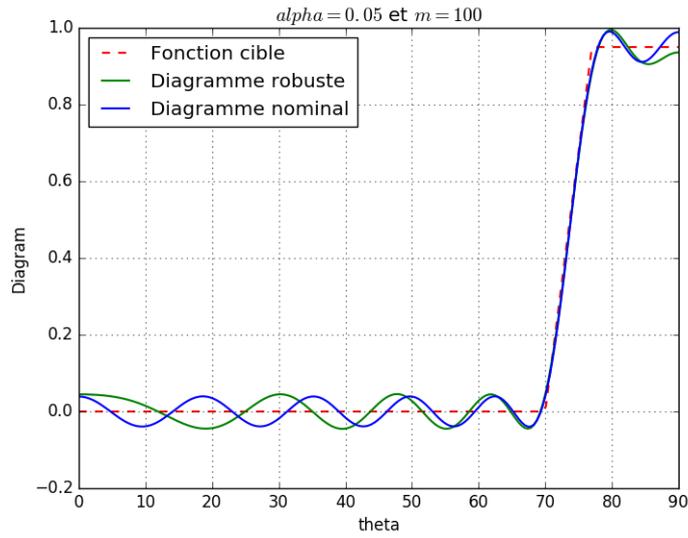


FIGURE 3. Diagramme Robuste

Notons que par construction de la contrepartie robuste au sens du rayon de stabilité, le diagramme robuste reste stable pour toute perturbation de taille $\epsilon \in [0, r_{\max}[$ où r_{\max} désigne la valeur optimale de la contrepartie robuste. Par exemple, pour $m = 100$ et $\alpha = 0,05$, on trouve $r_{\max} = 0,005254$. Dans ce cas, pour tout $\epsilon \in [0, 0.00525[$, le diagramme robuste reste stable.

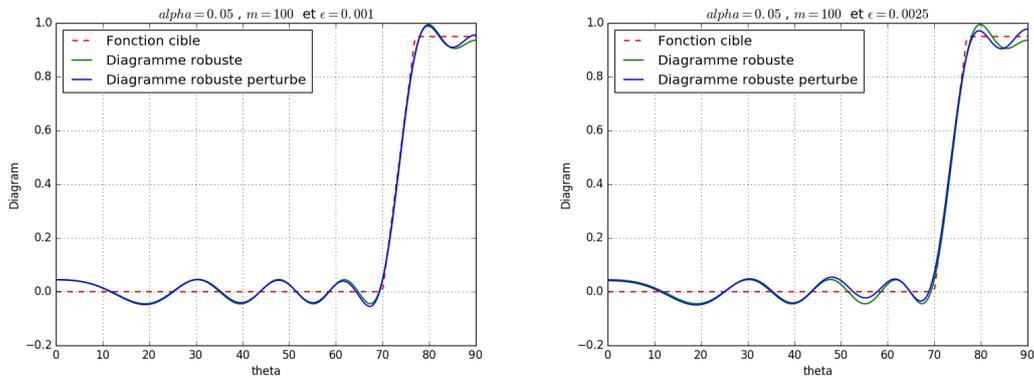


FIGURE 4. Diagramme robuste perturbé

Conclusion et perspectives

Dans cette thèse, nous avons abordé deux questions essentielles dans les processus d'aide à la décision : la dualité et la robustesse.

Dans la dualité, nous avons d'une part établi l'égalité entre la pire valeur et la valeur robuste d'un problème d'optimisation convexe conique à données incertaines et, d'autre part, nous avons proposé un schéma général permettant le passage de la dualité par niveaux à la dualité par épigraphe. Le point de vue considéré est celui de fonction $c : U \times V \rightarrow \mathbb{R}$ dite de couplage et d'ensembles élémentaires associés

$$E_{v,r}^P = \{u \in U \mid r - c(u, v) \in P\},$$

où P est un sous-ensemble non vide de \mathbb{R} , $v \in V$ et $r \in \mathbb{R}$.

En optimisation de la robustesse, en s'inspirant d'un travail fait par Marc Ciligot-Travian ([21]), nous avons introduit l'utilisation du rayon de stabilité dans l'étude des problèmes de décision en milieu incertain. Nous avons prouvé certaines propriétés (quasi-concavité, s.c.s,...) de cet indicateur de robustesse et établi des expressions explicites du rayon de stabilité dans certains cas. Nous avons également proposé un modèle générique regroupant une large classe de cas pratiques. Nous avons ensuite établi des résultats intéressants dans un cadre polyédral, dans le cas de la régression et dans un cadre quadratique pour ce modèle. Une connexion entre les solutions de la contrepartie robuste au sens de l'optimisation robuste et les solutions de la contrepartie robuste au sens du rayon de stabilité d'un problème d'optimisation à données incertaines est établie. Deux applications, une sur un problème de conception d'antennes et l'autre sur un problème de calcul d'estimateur robuste sont données.

En perspective pour la première partie, on projette utiliser cette théorie dans l'étude d'un problème d'optimisation à données incertaines d'une part et d'autre part on aimerait examiner une approche qui consiste à fixer le paramètre $r \in \{-1, 0, 1\}$ dans la définition d'ensemble élémentaire. Notons que cette approche nous permettra de récupérer entre autre le théorème du bi-polaire.

Comme perspective pour la deuxième partie, nous souhaitons travailler en collaboration avec des ingénieurs mécaniciens sur l'utilisation du rayon de stabilité dans la

conception de certaines pièces métalliques. Nous envisageons également utiliser la méthode du rayon de robustesse sur des problèmes de contrôle optimal à données incertaines.

Bibliographie

- [1] M. Aït Mansour, A. Metrane, and M. Théra. Lower semicontinuous regularization for vector-valued mappings. *J. Global Optim.*, 35 : 283–309, 2006.
- [2] G. Allaire. *Analyse numérique et optimisation : Une introduction à la modélisation mathématique et à la simulation numérique French*. Les Éditions de l'École Polytechnique, 2 edition, 2012.
- [3] M. Barro, A. Ouédraogo, and S. Traoré. On uncertain conical convex optimization problems. *accepted, to appear in Pacific Journal of Optimization*.
- [4] F. L. Bauer, J. Stoer, and C. Witzgall. Absolute and monotonic norms. *Numerische Mathematik*, 3(1) : 257–264, 1961.
- [5] A. Beck and A. Ben-Tal. Duality in robust optimization : primal worst equals dual best. *Oper. Res. Lett.*, 37 : 1–6, 2009.
- [6] Y. Ben-Haim. *Information-gap decision theory*. Series on Decision and Risk. Academic Press Inc., San Diego, CA, 2001. Decisions under severe uncertainty.
- [7] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust optimization*. Princeton Series in Applied Mathematics. Princeton University Press, Princeton, NJ, 2009.
- [8] A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Math. Oper. Res.*, 23, 1998.
- [9] A. Ben-Tal and A. Nemirovski. *Lectures on Modern Convex Optimization : Analysis, Algorithms, and Engineering Applications*. MPS-SIAM Series on Optimization. Society for Industrial Mathematics, 2001.
- [10] A. Ben-Tal and A. Nemirovski. Robuste optimization methodology and applications. *Math. Programm.*, 92 : 458–480, 2002.
- [11] C. Berge. *Espaces topologiques, fonctions multivoques*. 1959.
- [12] D. Bertsimas, D. Pachamanova, and S. Melvyn. Robust linear optimization under general norms. *Operations Research Letters*, 32 : 510–516, 2004.
- [13] R. I. Boţ. *Conjugate duality in convex optimization*. Springer-Verlag, Berlin Heidelberg, 2010.
- [14] R. I. Boţ, V. Jeyakumar, and L. Y. Li. Robust duality in parametric convex optimization. *Set-Valued Var. Anal.*, 21 : 177–189, 2013.
- [15] R. I. Boţ and G. Wanka. A weaker regularity condition for subdifferential calculus and fenchel duality in infinite dimensional spaces. *Nonlinear Anal.*, 64 : 2787–2804, 2006.
- [16] N. Bourbaki. *Espaces vectoriels topologiques*. Éléments de mathématique. Masson, Paris, 1981.
- [17] H. Brezis. *Analyse fonctionnelle*. Collection Mathématiques Appliquées pour la Maîtrise. Masson, Paris, 1983.

- [18] R. S. Burachik and V. Jeyakumar. A dual condition for the convex subdifferential sum formula with applications. *J. Convex Anal.*, 12 : 279–290, 2005.
- [19] E. Carrizosa and J. Fliege. Generalized goal programming : polynomial methods and applications. *Mathematical Programming*, 93(2) : 281–303, 2002.
- [20] A. Charnes and W. Cooper. Programming with linear fractional functionals. *Naval Research logistic*, 9 : 181–186, 1962.
- [21] M. Ciligot-Travain. Sur une notion de robustesse. *Ann. I.S.U.P.*, 56(2-3) : 37–48, 2012.
- [22] C. Combari, M. Laghdir, and L. Thibault. Sous-différentiels de fonctions convexes composées. *Ann. Math. Qué.*, 18 : 119–148, 1994.
- [23] J. P. Crouzeix. *Contributions à l'étude des fonctions quasiconvexes*. Thesis. University of Clermont-Ferrand, France, 1977.
- [24] J.-P. Crouzeix, J. Ferland, and S. Schaible. Duality in generalized linear fractional programming. *Mathematical Programming*, 27 : 342–354, 1983.
- [25] J.-P. Crouzeix, J. Ferland, and S. Schaible. An algorithm for generalized fractional programs. *JOTA*, 47(1) : 35–49, 1985.
- [26] J.-P. Crouzeix and J. A. Ferland. Algorithms for generalized fractional programming. *Mathematical Programming*, 52 : 191–207, 1991.
- [27] N. Dinh, G. Vallet, and M. Volle. Functional inequalities and theorems of the alternative involving composite functions. *J. Global Optim.*, 59 : 837–863, 2014.
- [28] W. Dinkelbach. On nonlinear fractional programming. *Management Science*, 13(7) : 492–498, 1967.
- [29] J. Dixmier. *General topology*. Springer-Verlag, 1 edition, 1984.
- [30] S. Dolecki and S. Kurcyusz. On ϕ -convexity in extremal problems. *SIAM J. Control Optim.*, 16 : 277–300, 1978.
- [31] L. El Ghaoui and H. Lebret. Robust solutions to least-squares problems with uncertain data. *SIAM J. Matrix Anal. Appl.*, 18(4) : 1035–1064, 1997.
- [32] L. M. Elias and J. E. Martínez-Legaz. A simplified conjugation scheme for lower semi-continuous functions. *Optimization*, 65(4) : 751–763, 2016.
- [33] P. Embrechts and M. Hofert. A note on generalized inverses. *Math. Methods Oper. Res.*, 77(3) : 423–432, 2013.
- [34] E. Ernst and M. Volle. Zero duality gap and attainment with possibly non-convex data. *J. Convex Anal.*, 23(2) : 615–629, 2016.
- [35] N. Falkner and G. Teschl. On the substitution rule for Lebesgue-Stieltjes integrals. *Expo. Math.*, 30(4) : 412–418, 2012.
- [36] W. Fenchel. A remark on convex sets and polarity. *Comm. Sém. Math. Univ. Lund[Medd. Lunds Univ. Mat. Sem.]*, pages 82–89, 1952.
- [37] C. Feng, H. Wang, X. Tu, and J. Kowalski. A note on generalized inverses of distribution function and quantile transformation. *Applied Mathematics*, 3(12A) : 2098–2100, 2012.
- [38] F. Flores-Bazán. On a notion of subdifferentiability for non-convex functions. *Optimization*, 33(1) : 1–8, 1995.
- [39] A. L. Fradkov and V. A. Jakubovič. The S -procedure and the duality relation in convex quadratic programming problems. *Vestnik Leningrad. Univ.*, (Vol. 6, n°2) : 101–109, 1979.

- [40] M. Goberna and M. López. *Linear Semi-Infinite Optimization*. Wiley Series in Mathematical Methods in Practice. Wiley, 1998.
- [41] S. Guillaume and M. Volle. Level set relaxation, epigraphical relaxation and conditioning in optimization. *Positivity*, 19 : 769–795, 2015.
- [42] N. Hadjisavvas, S. Komlósi, and S. Schaible, editors. *Handbook of generalized convexity and generalized monotonicity*, volume 76 of *Nonconvex Optimization and its Applications*. Springer-Verlag, New York, 2005.
- [43] H. A. Hindi and S. P. Boyd. Robust solutions to l_1 , l_2 , and l_∞ uncertain linear approximation problems using convex optimization. *Proceedings of the American Control Conference*, 6 : 3487–3491, 1998.
- [44] D. Hindrichsen and A. Pritchard. Stability radii of linear systems. *Systems and Control Letters*, 7 : 1–10, 1986.
- [45] C. L. Jean-Baptiste Hiriart-Urruty. *Fundamentals of Convex Analysis*. Grundlehren Text Editions. Springer-Verlag Berlin Heidelberg, 1 edition, 2001.
- [46] C. L. a. Jean-Baptiste Hiriart-Urruty. *Convex Analysis and Minimization Algorithms I : Fundamentals*. Grundlehren der mathematischen Wissenschaften 305. Springer-Verlag Berlin Heidelberg, 1 edition, 1993.
- [47] V. Jeyakumar and G. Y. Li. Strong duality in robust convex programming : complete characterizations. *SIAM J. Optim.*, 20 : 3384–3407, 2010.
- [48] E. P. Klement, R. Mesiar, and E. Pap. Quasi- and pseudo-inverses of monotone functions, and the construction of t-norms. *Fuzzy Sets and Systems*, 104(1) : 3–13, 1999.
- [49] G. Y. Li, V. Jeyakumar, and G. M. Lee. Robust conjugate duality for convex optimization under uncertainty with application to data classification. *Nonlinear Anal.*, 74 : 2327–2341, 2011.
- [50] D. T. Luc. *Theory of vector Optimization*. Springer-Verlag, Heidelberg, 1989.
- [51] J. Martínez-Legaz. *Generalized Convex Duality and its Economic Applications*. Nonconvex Optimization and Its Application, Handbook of generalized convexity and generalized monotonicity. Springer, New York, 2005. 76.
- [52] W. Milne and R. Reynolds. Fifth-order methods for the numerical solution of ordinary differential equations. *ACM*, 9(1) : 64–70, 1962.
- [53] J. J. Moreau. Fonctionnelles convexes. *Séminaire Jean Leray, Collège de France*, 1966.
- [54] J. J. Moreau. Inf-convolution, sous-additivité, convexité des fonctions numériques. *J. Math. Pures Appl.*, 49 : 109–154, 1970.
- [55] J. P. Penot. What is quasiconvex analysis? *Optimization*, 47 : 35–110, 2000.
- [56] J. P. Penot. Conjugacies adapted to lower semicontinuous functions. *Optimization*, 64(3) : 473–494, 2015.
- [57] J. P. Penot and M. Théra. Semi-continuous mappings in general topology. *Arch. Math.*, 38 : 158–166, 1982.
- [58] J. P. Penot and M. Volle. On quasi-convex duality. *Math. Oper. Res.*, 15 : 4597–625, 1990.
- [59] J. P. Penot and M. Volle. Surrogate programming and multipliers in quasi-convex programming. *SIAM J. Control Optim*, 42(6) : 1994–2003, 2004.

- [60] R. T. Rockafellar. Extension of fenchel's duality for convex functions. *Duke Math. J.*, 33 : 81–90, 1966.
- [61] R. T. Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.
- [62] R. T. Rockafellar. *Conjugate duality and optimization*. Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1974. Lectures given at the Johns Hopkins University, Baltimore, Md., June, 1973, Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, No. 16.
- [63] A. Rubinov. *Abstract Convexity and Global Optimization*. Nonconvex Optimization and Its Application. Springer US, 2000. 44.
- [64] S. Schiabile. Parameter-free convex equivalent and dual programs of fractional programming problems. *Operation Research*, 18 : 187–196, 1974.
- [65] I. Singer. *Abstract convex analysis*. Canadian Mathematics Series of Monographs and Texts. A wiley Interscience, 1997.
- [66] M. Sniedovich. The mighty maximin. *Working paper No. M.S-02-08*, 2008.
- [67] A. Soyster. Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research*, 21 : 1154–1157, 1973.
- [68] I. M. Stancu-Minasian. *Fractional Programming : Theory, Methods and Applications*. Mathematics and Its Applications 409. Springer Netherlands, 1 edition, 1997.
- [69] H. Tuy. Convex programs with an additional reverse convex constraint. *J. Optim. Theory Appl.*, 52(3) : 463–486, 1987.
- [70] J. Vicente-Pérez and M. Volle. Evenly convexity, subdifferentiability, and Γ -regularization in general topological vector spaces. *J. Math. Anal*, 429 : 956–968, 2015.
- [71] M. Volle. Conjugaison par tranches. *Annali di Matematica pura ed applicata*, CXXXIX(IV) : 279–312, 1985.
- [72] M. Volle. Conjugaison par tranche et dualité de toland. *Optimization*, 18(5) : 633–642, 1987.
- [73] M. Volle, J. E. Martínez-Legaz, and J. Vicente-Pérez. Duality for closed convex functions and evenly convex functions. *JOTA*, 167(3) : 985–997, 2015.
- [74] H. Wilf. Maximally stable numerical integration. *SIAM*, 8(3) : 537–540, 1960.
- [75] H. Xu and C. Caramanis. Robust regression and lasso. *IEEE TRANSACTIONS ON INFORMATION THEORY*, 56(7) : 3561–3574, 2010.
- [76] S. Zlobec. Characterizing optimality in mathematical programming models. *Acta Applicandae Mathematicae*, 12 : 113–180, 1988.