UNIVERSIDADE FEDERAL DO PARANÁ

LEONARDO MORETO ELIAS

UMA TEORIA DE CONJUGAÇÃO SIMPLIFICADA PARA FUNÇÕES SEMICONTÍNUAS INFERIORMENTE E UMA GENERALIZAÇÃO DA DESIGUALDADE FORTE DE FITZPATRICK

CURITIBA
2016

LEONARDO MORETO ELIAS

UMA TEORIA DE CONJUGAÇÃO SIMPLIFICADA PARA FUNÇÕES SEMICONTÍNUAS INFERIORMENTE E UMA GENERALIZAÇÃO DA DESIGUALDADE FORTE DE FITZPATRICK

Tese apresentada ao Programa de Pós-Graduação em Matemática da Universidade Federal do Paraná, como requisito parcial à obtenção do grau de Doutor em Matemática.
Orientador: Prof. Dr. Ademir Alves Ribeiro
Orientador no exterior: Prof. Dr. Juan Enrique Martínez-Legaz.

CURITIBA
JULHO DE 2016

ATA DE SESSÃO PÚBLICA DE DEFESA DE TESE PARA A OBTENÇÃO DO GRAU DE DOUTOR EM MATEMÁTICA

No dia vinte e nove de Julho de dois mil e dezesseis às 14:00 horas, na sala Anfiteatro A, Coordenação PPGMA, Centro Politécnico, UFPR, do Setor de CIÊNCIAS EXATAS da Universidade Federal do Paraná, foram instalados os trabalhos de arguição do doutorando **LEONARDO MORETO ELIAS** para a Defesa Pública de sua Tese intitulada: **"UMA CONJUGAÇÃO SIMPLIFICADA PARA FUNÇÕES SEMICONTÍNUAS INFERIORMENTE E UMA GENERALIZAÇÃO DA DESIGUALDADE FORTE DE FITZPATRICK"**. A Banca Examinadora, designada pelo Colegiado do Programa de Pós-Graduação em MATEMÁTICA da Universidade Federal do Paraná, foi constituída pelos seguintes Membros: ADEMIR ALVES RIBEIRO (UFPR), LUCAS GARCIA PEDROSO (UFPR), LUCELINA BATISTA DOS SANTOS (UFPR), ROY WILHELM PROBST (UTFPR), VALERIANO ANTUNES DE OLIVEIRA (UNESP/SJRP), WILFREDO SOSA SANDOVAL (UCB). Dando início à sessão, a presidência passou a palavra ao discente, para que o mesmo expusesse seu trabalho aos presentes. Em seguida, a presidência passou a palavra a cada um dos Examinadores, para suas respectivas arguições. O aluno respondeu a cada um dos arguidores. A presidência retomou a palavra para suas considerações finais e, depois, solicitou que os presentes e o doutorando deixassem a sala. A Banca Examinadora, então, reuniu-se sigilosamente e, após a discussão de suas avaliações, decidiu-se pela APROVAÇÃO do aluno. O doutorando foi convidado a ingressar novamente na sala, bem como os demais assistentes, após o que a presidência fez a leitura do Parecer da Banca Examinadora. Nada mais havendo a tratar a presidência deu por encerrada a sessão, da qual eu, ADEMIR ALVES RIBEIRO, lavrei a presente ata, que vai assinada por mim e pelos membros da Comissão Examinadora.

Curitiba, 29 de Julho de 2016.

Prof ADEMIR ALVES RIBEIRO
Presidente da Banca Examinadora (UFPR)

Lucas Garcia Pedroso
Prof LUCAS GARCIA PEDROSO
Avaliador Interno (UFPR)

Prof LUCELINA BATISTA DOS SANTOS
Avaliador Interno (UFPR)

Prof WILFREDO SOSA SANDOVAL
Avaliador Externo (UCB)

Prof VALERIANO ANTUNES DE OLIVEIRA
Avaliador Externo (UNESP/SJRP)

Prof ROY WILHELM PROBST
Avaliador Externo (UTFPR)

MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
Setor CIÊNCIAS EXATAS
Programa de Pós Graduação em MATEMÁTICA
Código CAPES: 40001016041P1

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em MATEMÁTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da Tese de Doutorado de **LEONARDO MORETO ELIAS**, intitulada: "**UMA CONJUGAÇÃO SIMPLIFICADA PARA FUNÇÕES SEMICONTÍNUAS INFERIORMENTE E UMA GENERALIZAÇÃO DA DESIGUALDADE FORTE DE FITZPATRICK**", após terem inquirido o aluno e realizado a avaliação do trabalho, são de parecer pela sua APROVAÇÃO.

Curitiba, 29 de Julho de 2016.

Prof ADEMIR ALVES RIBEIRO
Presidente da Banca Examinadora (UFPR)

Lucas Garcia Pedroso
Prof LUCAS GARCIA PEDROSO
Avaliador Interno (UFPR)

Lucelina
Prof LUCELINA BATISTA DOS SANTOS
Avaliador Interno (UFPR)

Prof WILFREDO SOSA SANDOVAL
Avaliador Externo (UCB)

Oliveira
Prof VALERIANO ANTUNES DE OLIVEIRA
Avaliador Externo (UNESP/SJRP)

Probst
Prof ROY WILHELM PROBST
Avaliador Externo (UTFPR)

Agradecimentos

Meus sinceros agradecimentos:

Agradeço primeiramente aos meus pais Alcimar e Maria,
aos meus irmãos e familiares,
que estiveram fortemente presente na minha caminhada.

Aos meus professores, que me passaram todo o conhecimento,
em especial a professora Elizabeth.

Aos meus amigos de graduação Ednei, Alana e Hannah,
que tanto me apoiaram e me incentivaram.

Aos meus amigos de pós-graduação Elvis, Aura, Diego,
Karla, Priscila, Paulinha, Dion, Marcos, Rodrigo e Oliver que
estiveram presente na etapa final de toda esta trajetória.

A todos os meus amigos que participaram deste longo caminho
e me deram forças pra continuar.

Aos meus orientadores Ademir e Juan Enrique que estavam
sempre dispostos a ajudar. Agradeço ao esforço e dedicação
empregados neste trabalho.

Aos membros da banca, professora Lucelina e professores Lucas,
Valeriano, Wilfredo e Roy. Obrigado por aceitarem o convite e
por fazerem
sugestões enriquecedoras.

Ao Programa de Pós-Graduação em Matemática Aplicada da
UFPR pela oportunidade e formação de qualidade propiciada.
Ao Conselho Nacional de Desenvolvimento Científico e
Tecnológico e a Fundação Araucária, pelo apoio financeiro.

A Deus pelo dom da vida, pelo equilíbrio e pela paz interior.

E, finalmente, a todos que, de alguma forma, contribuíram no
desenvolvimento deste trabalho.

" O segredo é não correr atrás das borboletas... É cuidar do jardim para que elas venham até você."

Mario Quintana

Resumo

Este trabalho engloba dois temas diferentes. O primeiro tópico trata de apresentar dois esquemas de c-conjugação para funções semicontínuas inferiormente (sci) definidas em espaços vetoriais reais de Banach cuja norma é Fréchet diferenciável fora da origem. Ambos os esquemas são baseados numa nova caracterização de funções sci via supremo pontual de um conjunto especial de funções contínuas. Para finalizar esta primeira frente do trabalho, estes esquemas são aplicados no desenvolvimento de uma teoria de dualidade. O segundo tópico trata de uma generalização da Desigualdade Forte de Fitzpatrick em espaços vetoriais de Banach reflexivos, envolvendo funções TBC. Ao final, introduz-se uma família de funções *gap* para o Problema de Inclusão Monótona Maximal e, graças à generalização proposta, é possível encontrar interessantes propriedades a respeito desta família.

Palavras-chave: *Função semicontínua inferiormente; Conjugação convexa generalizada; Desigualdade Forte de Fitzpatrick; Problema de Inclusão Monótona Maximal.*

Abstract

We present two topics. Firstly, we introduce two generalized conjugation schemes for lower semi-continuous (lsc) functions defined on a real Banach space whose norm is Fréchet differentiable off the origin. Both approaches are based upon a new characterization of lower semi-continuous functions as pointwise suprema of a special class of continuous functions. In order to conclude this part of the work, we apply these ideas for building a optimization duality theory. In the second topic, we present a generalization of the strong Fitzpatrick inequality in the context of reflexive Banach spaces, involving a TBC function. We also introduce a related family of *gap* functions for maximal monotone inclusion problems. Thanks to the proposed generalization, we find interesting properties about this family.

Keywords: *Lower semi-continuous function; Generalized convex conjugation; Strong Fitzpatrick inequality; Maximal monotone inclusion problems..*

Lista de Símbolos

f^*: Conjugada de Fenchel da função f.
$\langle \cdot, \cdot \rangle$: Produto de dualidade.
$T : X \rightrightarrows X^*$: Operador do tipo ponto-conjunto.
F_T: Função de Fitzpatrick do operador T.
$\mathbb{R}$: Espaço vetorial dos números reais.
$\mathbb{R}^n$: Espaço vetorial de todas as n-uplas de números reais.
$\overline{\mathbb{R}} : \mathbb{R} \cup +\infty \cup -\infty$.
$\approx$: Isomorfo.
$B(0, \delta)$: Bola aberta de centro na origem e raio δ.
$\overline{B(0, \delta)}$: Fecho da bola aberta de centro na origem e raio δ.
(x_k): sequência.
x_k: o $k-$ésimo elemento da sequência (x_k).
$\partial f(x)$: Subdiferencial de Fenchel da função f em x.
$\partial_F f(x)$: Fréchet subdiferencial de Fenchel da função f em x.

Sumário

Introdução

Esta tese está dividida em duas frentes de trabalho: teoria de conjugação para funções semicontínuas inferiormente (sci) e uma generalização da Desigualdade Forte de Fitzpatrick com uma aplicação na construção de funções *gap* para Problemas de Inclusão Monótona Maximal (PIM).

A seguir vamos fazer uma breve revisão bibliográfica destes dois temas e enquadrar os objetivos desse trabalho no contexto de cada um deles.

Teoria de conjugação para funções sci

A teoria de conjugação convexa desenvolvida por Fenchel [17] associa a cada função $f : E \to \mathbb{R} \cup \{+\infty\}$, definida num espaço vetorial métrico E, uma função sci convexa dada por

$$f^*(x^*) = \sup_{x \in E} \{\langle x, x^* \rangle - f(x)\}$$

em que $x^* \in E^*$, o dual de E, e $\langle \cdot, \cdot \rangle$ denota o produto de dualidade de E^* com E. A esta função chamamos de conjugada de f. Esta teoria possui aplicações na Teoria de Jogos e Economia [1, 2, 3], em Mecânica [20, 26], em Equações Diferenciais Parciais e Análise Complexa [21] e outras áreas da Matemática.

Esta teoria contribuiu também para a Otimização, pois serviu como base teórica para o desenvolvimento de um esquema de Dualidade Convexa [32]. Este esquema permite associar um problema de Minimização Convexa a um outro problema dito "dual"que, em algumas situações, pode ser mais fácil de resolver. As relações existentes entre estes problemas podem ser úteis para analisar as propriedades do problema original e, até mesmo, encontrar condições de otimalidade. De acordo com as propriedades em questão, as soluções destes problemas podem coincidir. Este esquema também é utilizado no desenvolvimento de algoritmos que encontram soluções ou, então, aproximações da solução do problema inicial.

Uma ferramenta essencial na teoria de dualidade é a noção de subgradiente [5] que, no caso de funções convexas diferenciáveis, coincide com a ideia de gradiente. Esta noção também nos permite desenvolver condições de otimalidade e algoritmos para minimizar funções.

Toda esta teoria clássica foi desenvolvida para o caso de funções convexas, porém problemas modelados de situações reais podem não possuir esta propriedade. Alguns autores optaram por trabalhar com problemas de minimização de funções semicontínuas inferiormente que, dependendo de certas condições de coercividade, possuem solução. Estes autores viram a conjugação de Fenchel como uma ferramenta para atacar estes problemas, principalmente pela sua aplicação nos esquemas de dualidade. No entanto, eles precisariam adaptar esta teoria para o caso semicontínuo.

Moreau [27] foi pioneiro em generalizar a teoria de conjugação. Ele introduziu a c-conjugada que permitiu trabalhar com esse tópico em um contexto mais geral, podendo abrir mão da convexidade. Analogamente à teoria original, a c-conjugação foi base para um esquema [23] que generaliza a Dualidade Convexa. Porém, a conjugação para as funções sci só seria estudada anos mais tarde.

Flores-Bazán e Martínez-Legaz [19], baseados num caso particular do Teorema de Seleção de Michael [25], desenvolveram uma conjugação para funções sci, bem como propuseram um subdiferencial que era não-vazio para estas funções.

Agora chamamos a atenção para uma interpretação econômica muito peculiar da conjugada: quando E é um espaço vetorial de dimensão finita, considerando x a quantidade de produção de uma determinada empresa, $f(x)$ o custo para produzir essa quantidade e x^* o preço de venda, é possível perceber que a conjugada de f modela o lucro otimizado da empresa.

No entanto, nos problemas reais, o preço de venda do produto varia de acordo com a quantidade produzida. Por exemplo, quando há uma quantidade grande de um determinado produto disponível para o mercado consumidor, o preço tende a diminuir. Já quando ocorre o contrário, ou seja, quando o produto está escasso, o preço tende a subir. Partindo dessa filosofia, a conjugada de Fenchel deveria ser modificada para que incluísse essa variação de preço na sua modelagem.

Cotrina, Karas, Ribeiro, Sosa e Yuan [11] procuraram resolver este problema da modelagem apresentando um caso particular de c-conjugada dada por

$$f^s(p) = \sup_{x \in \mathbb{R}^n} \{\langle x, p(x)\rangle - f(x)\}$$

em que $p : \mathbb{R}^n \to \mathbb{R}^n$ é uma função contínua. Estes autores perceberam que esta nova conjugada ia além de suas pretensões, pois ela se adaptou muito bem para o caso semicontínuo, permitindo assim o desenvolvimento de uma nova teoria de conjugação para funções sci.

No contexto destes estudos, uma das propostas deste trabalho de tese é elaborar uma simplificação da teoria de conjugação apresentada em [11], focando-nos em dois pontos: o fato de estarmos trabalhando com elementos do espaço dos operadores contínuos de $\mathbb{R}^n$ e a não unicidade de p no sentido de que para uma infinidade de funções p, temos a mesma função $\langle x, p(x)\rangle$ relacionada. Estas situações podem não ser viáveis tanto em questões teóricas quanto em aplicações. No decorrer da Seção 2.2, estes objetivos ficarão mais claros.

Vamos propor dois casos particulares de c-conjugada para funções sci e desenvolver, em sequência, noções de subdiferencial e esquemas de dualidade relacionados. Este trabalho foi desenvolvido sob orientação do professor Juan Enrique Martínez-Legaz do departamento de Economia da Universidade Autônoma de Barcelona, Espanha, no processo de doutorado-sanduíche realizado de setembro de 2014 até maio de 2015. Submetemos o trabalho intitulado "A simplified conjugation scheme for lower semi-continuous functions" [16] para a revista Optimization em janeiro de 2015 e o mesmo foi publicado no próximo ano.

Como os nosso objetivos iniciais foram cumpridos num tempo hábil, resolvemos expandir os conhecimentos desta tese para uma nova área. Estabelecemos um novo objetivo dentro da Teoria de Operadores.

Uma generalização da Desigualdade Forte de Fitzpatrick

Na Teoria de Operadores, muitos autores se preocuparam em encontrar funções que representassem operadores $T : X \rightrightarrows X^*$, onde a notação $X \rightrightarrows X^*$ nos diz que são aplicações do tipo ponto-conjunto, ou seja, associam elementos de um espaço vetorial de Banach X não-vazio a subconjuntos do seu dual X^*. Uma função $f : X \times X^* \to \mathbb{R} \cup \{+\infty\}$ representa o operador T quando

(i) $f(x, x^*) \geq \langle x, x^* \rangle$ para todo $(x, x^*) \in X \times X^*$;

(ii) $f(x, x^*) = \langle x, x^* \rangle$ para todo $(x, x^*) \in gra\ T = \{(x, x^*) \in X \times X^* \mid x^* \in T(x)\}$,

em que $\langle \cdot, \cdot \rangle$ denota o produto de dualidade de X^* com X. Com isso, é possível estudar propriedades do operador analisando somente tais funções.

Uma das mais conhecidas funções representantes foi exibida por Fitzpatrick [18] em 1988 e definida como

$$F_T(x, x^*) := \sup_{(y,y^*) \in gra\ T} \{\langle x - y, y^* \rangle + \langle y, x^* \rangle\}.$$

Para um operador monótono maximal, Fitzpatrick mostrou que sua função era própria, convexa, sci e minorava qualquer outra função representante com estas mesmas propriedades.

Nos anos seguintes, surgiram trabalhos relacionados a operadores monótonos em que a função de Fitzpatrick teve papel fundamental, evidenciando a sua grande relevância para a área [10, 24, 36].

Em 2009, Voisei e Zalinescu [40] exibiram uma desigualdade que mostra que quanto mais longe um ponto está do gráfico de um operador, maior será a diferença entre a função de Fitzpatrick e o produto de dualidade avaliadas neste ponto. Esta desigualdade se tornou famosa, conhecida como Desigualdade Forte de Fitzpatrick, e é dada por

$$F_T(x,x^*) - \langle x,x^*\rangle \geq \frac{1}{4} d^2_{gra\ T}(x,x^*) = \frac{1}{4} \inf_{(w,w^*)\in gra\ T} \left\{ \|w-x\|_X^2 + \|w^*-x^*\|_{X^*}^2 \right\}$$

para todo $(x,x^*) \in X \times X^*$.

Esta desigualdade tem diversas aplicações, como na análise de soluções para o Problema de Inclusão Monótona e em Problemas de Desigualdade Variacional [6]. Ainda em [6], a desigualdade é utilizada para a construção de uma *função limite de erro* (error bound). Estas funções indicam o quão longe ou perto determinado ponto está da solução do PIM.

O nosso objetivo é introduzir uma versão dessa desigualdade para funções mais gerais que $\frac{1}{4} d^2_{gra\ T}(x,x^*)$, conhecidas como funções *Twisted Bigger Conjugate* (TBC) [35]. Em sequência, usando esta generalização, queremos propor uma família de funções *gap* [Definição 3.21] para o PIM, majoradas pela função *gap* apresentada na Seção 2 de [6].

Este estudo também foi realizado em parceria com o professor Martínez-Legaz e o produto final foi submetido na revista Optimization com o título "A generalization of the strong Fitzpatrick inequality" [15] em setembro de 2015 e publicado *online* em maio de 2016.

Disposição do trabalho

Este trabalho tem âmbito teórico e estuda tópicos de minimização de funções sci e de Teoria de Operadores. Os dois primeiros capítulos são dedicados ao primeiro tema. No Capítulo 1 apresentamos a generalização proposta por Moreau, bem como a sua contribuição para a área de Otimização. Discutimos no Capítulo 2 os resultados já existentes de conjugação para funções sci e apresentamos a nossa proposta de simplificação. Definimos dois casos particulares de c-conjugação em conjunto com as teorias de subdiferenciabilidade e dualidade relacionadas.

O Capítulo 3 é dedicado ao segundo tema do trabalho. Neste, vamos apresentar a nossa generalização da Desigualdade Forte de Fitzpatrick, faremos um rápido estudo a respeito do PIM e apresentaremos a nossa família de funções gap para este problema.

No Apêndice revisamos alguns conceitos de Análise Convexa que julgamos importantes para o entendimento deste trabalho e anexamos os dois artigos resultantes desta tese.

Para um leitor que não possua familiaridade com conceitos de Análise Funcional, sugerimos que considere todos os resultados envolvidos neste trabalho restritos ao contexto de espaços de dimensão finita.

Capítulo 1

A conjugação generalizada de Moreau

Neste capítulo vamos apresentar uma generalização da teoria de conjugação de Fenchel proposta por Moreau em [27]. Vamos estudar alguns casos particulares que servirão como motivação para o tema desenvolvido no segundo capítulo deste trabalho. Apresentaremos também as contribuições dessa generalização para a área de Otimização. Os resultados apresentados neste capítulo podem ser encontrados em [23].

1.1 A c-conjugada

A teoria de conjugação de Fenchel [17] teve forte impacto em Otimização, principalmente por tratar de dualidade num âmbito bem mais geral. O primeiro passo para generalizar essa teoria é estender a definição de função conjugada.

Definição 1.1 *Dados X e Y conjuntos arbitrários não-vazios, $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$, sejam $f : X \to \overline{\mathbb{R}}$ e $\varphi : Y \to \overline{\mathbb{R}}$ funções. Consideremos $c : X \times Y \to \overline{\mathbb{R}}$ uma função dita de acoplamento. A função c-conjugada de f é definida por $f^c : Y \to \overline{\mathbb{R}}$;*

$$f^c(y) = \sup_{x \in X} \{c(x,y) - f(x)\}.$$

A função c'-conjugada $\varphi^{c'} : X \to \overline{\mathbb{R}}$ de φ é definida sob a consideração da função de acoplamento $c' : Y \times X \to \overline{\mathbb{R}}$ dada por $c'(y,x) = c(x,y)$, ou seja,

$$\varphi^{c'}(x) = \sup_{y \in Y} \{c(x,y) - \varphi(y)\}.$$

Como podemos ver, a definição de c-conjugada é muito simples e não pede nenhuma estrutura (vetorial, topológica, métrica) para os conjuntos X e Y. Para que este conceito esteja bem definido, adotaremos a convenção

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

Notemos que, dessa forma, os pontos que atingem essas indeterminações são desconsiderados no cálculo da c-conjugada (c'-conjugada).

Como a função f^c é definida em Y (e $\varphi^{c'}$ em X), vamos definir a segunda c-conjugada (e c'-conjugada) naturalmente como $f^{cc'} = (f^c)^{c'}$ (e $\varphi^{c'c} = (\varphi^{c'})^c$).

Definição 1.2 *Dado $\beta \in \overline{\mathbb{R}}$. Para $y \in Y$, uma função é dita c-elementar quando é da forma $x \in X \mapsto c(x,y) - \beta$. Analogamente, para $x \in X$, uma função é dita c'-elementar quando é da forma $y \in Y \mapsto c(x,y) - \beta \in \overline{\mathbb{R}}$.*

A definição a ser introduzida a seguir é de nossa autoria, porém pelo Corolário 6.1 de [27], coincide com a definição de ϕ_c (ou $\phi_{c'}$)-convexa da mesma referência.

Definição 1.3 *Uma função $f : X \to \overline{\mathbb{R}}$ é c-supremo em $x \in X$ quando coincide com a sua segunda c-conjugada em x ($f(x) = f^{cc'}(x)$). Da mesma maneira, uma função $\varphi : Y \to \overline{\mathbb{R}}$ é c'-supremo em $y \in Y$ quando coincide com a sua segunda c'-conjugada em y ($\varphi(y) = \varphi^{c'c}(y)$).*

Diretamente dessa definição segue que uma função é c-supremo (ou c'-supremo) quando ela coincide com a sua segunda c-conjugada (c'-conjugada) em todos os pontos. Quando se define uma teoria de c-conjugação, o objetivo é sempre identificar as funções c-supremos e é comum dizer que a teoria de c-conjugação é definida exatamente para estas funções. A seguir, verificamos que a teoria de Fenchel é um caso particular da c-conjugação.

Exemplo 1.4 (Seção A.2, Apêndice) *Para $X = E$ um espaço vetorial métrico e $Y = E^*$ seu respectivo espaço dual topológico, consideremos $c : E \times E^* \to \mathbb{R}$ definida como $c(x,x^*) = \langle x, x^* \rangle$ em que $\langle \cdot, \cdot \rangle$ denota o produto de dualidade de E^* com E. Para uma função própria $f : E \to \mathbb{R} \cup \{+\infty\}$, f^c e $f^{cc'}$ são respectivamente, a conjugada e biconjugada de Fenchel (ou clássica) e denotadas por f^* e f^{**}. Neste caso, f é c-supremo se, e somente se, f é convexa e semicontínua inferiormente (ou sci). Por isso, dizemos que a teoria de Fenchel se comporta bem para funções convexas e sci ou então que a teoria de Fenchel é uma teoria de conjugação para funções convexas e sci.*

Exemplo 1.5 ([11]) *Sejam $Y = \mathcal{C} = \{q : \mathbb{R}^n \to \mathbb{R}^n \mid q$ é um operador contínuo$\}$ e $X = \mathbb{R}^n$, consideremos $c : \mathbb{R}^n \times \mathcal{C} \to \mathbb{R}$ definida por $c(x,q) = \langle x, q(x) \rangle$ em que $\langle \cdot, \cdot \rangle$ denota o produto interno de $\mathbb{R}^n$. Para essa conjugação, uma função própria $f : X \to \mathbb{R} \cup \{+\infty\}$ é c-supremo se, e somente se f é sci. Neste caso, dizemos que esta teoria de conjugação é para funções sci.*

O próximo resultado mostra que a generalização proposta mantém algumas propriedades da conjugada de Fenchel.

Proposição 1.6 *Sejam $f : X \to \overline{\mathrm{I\!R}}$ e $\varphi : Y \to \overline{\mathrm{I\!R}}$ funções, $x \in X$ e $y \in Y$. Então*

(i) $f^c(y) \geq c(x,y) - f(x)$, $\varphi^{c'}(x) \geq c(x,y) - \varphi(y)$.

(ii) $f^{cc'}(x) \leq f(x)$, $\varphi^{c'c}(y) \leq \varphi(y)$.

(iii) $f^{cc'c}(y) = f^c(y)$, $\varphi^{c'cc'}(x) = \varphi^c(x)$.

Demonstração. Como as provas para a c-conjugada e c'-conjugada são análogas, vamos demonstrar as afirmações somente para o primeiro caso.

(i) Consequência imediata da definição.

(ii) Do item anterior, temos que, para cada x, $f^c(y) \geq c(x,y) - f(x)$ e consequentemente, $f(x) \geq c(x,y) - f^c(y)$ para todo $y \in Y$. Portanto

$$f(x) \geq \sup_{y \in Y}\{c(x,y) - f^c(y)\} = f^{cc'}(x).$$

(iii) De (ii), temos que $f^{cc'c}(y) \leq f^c(y)$. Para provar que vale a igualdade, notamos que para cada y, $f^{cc'c}(y) \geq c(x,y) - f^{cc'}(x)$ para todo x. De (ii), $f^{cc'c}(y) \geq c(x,y) - f(x)$ para todo x. Assim

$$f^{cc'c}(y) \geq \sup_{x \in X}\{c(x,y) - f(x)\} = f^c(y).$$

$\square$

Denotando por Φ_c o conjunto de todas as funções c-elementares, podemos dar a seguinte caracterização para a segunda c-conjugada.

Teorema 1.7 *Para uma função $f : X \to \overline{\mathrm{I\!R}}$, $x \in X$,*

$$f^{cc'}(x) = \sup_{g \in \Phi_c}\{g(x) \mid g \leq f\}.$$

Demonstração. Para todo $x \in X$, consideremos $s = \sup_{g \in \Phi_c}\{g(x) \mid g \leq f\}$. Definimos a função auxiliar $\overline{g}(z) = c(z,y) - f^c(y)$ para $y \in Y$. Segue do item (i) da Proposição 1.6 que $\overline{g}(z) \leq f(z)$ para todo $z \in X$. Além do mais, $\overline{g} \in \Phi_c$ e assim,

$$s \geq \overline{g}(x) = c(x,y) - f^c(y)$$

para todo $y \in Y$. Tomando o supremo em $y \in Y$, concluímos que $f^{cc'}(x) \leq s$.

Suponhamos agora por absurdo que $f^{cc'}(x) < s$. Pela propriedade de supremo, existem $y \in Y$ e $b \in \mathrm{I\!R}$ tais que $c(x,y) + b > f^{cc'}(x)$ e, para todo $z \in X$, $c(z,y) + b \leq$

$f(z)$. Equivalentemente, $c(z,y) - f(z) \leq -b$. Tomando o supremo em $z \in X$, obtemos $f^c(y) \leq -b$, logo

$$f^{cc'}(x) < c(x,y) + b \leq c(x,y) - f^c(y) \leq f^{cc'}(x).$$

Absurdo, concluindo a demonstração. □

Deste modo, podemos ver aqui uma aplicação da c-conjugada, pois esta teoria sempre está relacionada com uma caracterização de funções via supremo de um subconjunto de funções. Por exemplo, no caso de Fenchel, como as funções convexas e sci são exatamente as funções c-supremo, podemos concluir que toda função convexa e sci é o supremo de funções afins (que são exatamente as funções c-elementares para este caso).

O próximo passo é, como no caso clássico, considerar uma teoria de subdiferenciabilidade relacionada à c-conjugação, chamada c-subdiferenciabilidade.

Definição 1.8 *Dizemos que $f : X \to \overline{\mathbb{R}}$ é c-subdiferenciável em $x_0 \in$ dom f quando existe $y_0 \in Y$ tal que $(x_0, y_0) \in$ dom c e*

$$f(x) - f(x_0) \geq c(x, y_0) - c(x_0, y_0) \text{ para todo } x \in X. \tag{1.1}$$

Assim, y_0 é dito c-subgradiente de f em x_0. O conjunto de todos os c-subgradientes de f em x_0, denotado por $\partial_c f(x_0)$, é chamado de c-subdiferencial de f em x_0. Se $f(x_0) \notin \mathbb{R}$, o conjunto $\partial_c f(x_0)$ é considerado vazio.

Notamos que tomando c como no Exemplo 1.4, obtemos a definição de subdiferencial clássica [5].

De forma análoga à definição anterior, a c'-subdiferenciabilidade é definida para funções definidas em Y. Apesar de não impormos muitas condições sobre c, ainda assim é possível manter algumas propriedades do caso clássico.

Proposição 1.9 *Sejam $f : X \to \overline{\mathbb{R}}$ uma função, $x_0 \in X$ e $y_0 \in Y$. Se $c(x_0, y_0) \in \mathbb{R}$ então*

(i) *$y_0 \in \partial_c f(x_0)$ se, e somente se $f(x_0) + f^c(y_0) = c(x_0, y_0)$.*

(ii) *$y_0 \in \partial_c f^{cc'}(x_0)$ se, e somente se $x_0 \in \partial_{c'} f^c(y_0)$.*

(iii) *Se $\partial_c f(x_0) \neq \emptyset$ então f é c-supremo em x_0.*

(iv) *Se f é c-supremo em x_0 então $\partial_c f^{cc'}(x_0) = \partial_c f(x_0)$.*

Demonstração. (i) Notemos que

$$\begin{aligned} y_0 \in \partial_c f(x_0) &\Leftrightarrow f(x) - f(x_0) \geq c(x, y_0) - c(x_0, y_0), \, \forall x \in X \\ &\Leftrightarrow c(x_0, y_0) - f(x_0) \geq c(x, y_0) - f(x), \, \forall x \in X \\ &\Leftrightarrow c(x_0, y_0) - f(x_0) \geq f^c(y_0) \\ &\Leftrightarrow c(x_0, y_0) - f(x_0) = f^c(y_0), \end{aligned}$$

em que a última implicação segue do item (i) da Proposição 1.6.

(ii) Da Proposição 1.6, sabemos que $f^c = f^{cc'c}$. Assim

$$\begin{aligned} y_0 \in \partial_c f^{cc'}(x_0) &\Leftrightarrow f^{cc'}(x_0) + f^{cc'c}(y_0) = c(x_0, y_0) \\ &\Leftrightarrow f^{cc'}(x_0) + f^c(y_0) = c(x_0, y_0) \\ &\Leftrightarrow x_0 \in \partial_{c'} f^c(y_0). \end{aligned}$$

(iii) Se existe $y_0 \in \partial_c f(x_0)$, então por (i),

$$f(x_0) = c(x_0, y_0) - f^c(y_0) \leq f^{cc'}(x_0).$$

Pelo item (ii) da Proposição 1.6, concluímos que $f(x_0) = f^{cc'}(x_0)$.

(iv) Assumimos que $f(x_0) = f^{cc'}(x_0)$ e, pela Proposição 1.6, temos que $f^{cc'c} = f^c$. Assim

$$\begin{aligned} y_0 \in \partial_c f^{cc'}(x_0) &\Leftrightarrow f^{cc'}(x_0) + f^{cc'c}(y_0) = c(x_0, y_0) \\ &\Leftrightarrow f(x_0) + f^c(y_0) = c(x_0, y_0) \\ &\Leftrightarrow y_0 \in \partial_c f(x_0). \end{aligned}$$

□

No Capítulo 2 deste trabalho, as c-conjugadas propostas são exemplos de que os conceitos de c-subdiferenciabilidade e c-supremo podem ser distintos. No entanto, temos a seguinte equivalência.

Corolário 1.10 *Para uma função $f : X \to \overline{\mathbb{R}}$ e $x_0 \in X$, f é c-subdiferenciável em x_0 se, e somente se, f é c-supremo e o supremo $\sup_{y \in Y}\{c(x, y) - f^c(y)\}$ é atingido para algum $y_0 \in Y$.*

Combinando (iv) com (ii), temos que para uma função $f : X \to \overline{\mathbb{R}}$ c-supremo, em certo sentido, a inversa do operador subdiferencial $\partial_c f$ é $\partial_{c'} f^c$, ou seja, para quaisquer $x_0 \in X$ e $y_0 \in Y$, temos que $y_0 \in \partial_c f(x_0)$ se, e somente se $x_0 \in \partial_{c'} f^c(y_0)$.

1.2 Dualidade Generalizada

A teoria de dualidade tem o interesse de associar a um problema de minimização (primal) um outro problema, chamado de dual. Sob determinadas condições (e num certo sentido) os dois problemas são equivalentes. No entanto, às vezes, o dual é mais fácil de resolver. Nesta seção, vamos apresentar a teoria de dualidade que está relacionada à c-conjugação. Esta teoria será exibida num contexto geral e inclui casos clássicos da literatura como a Dualidade Lagrangiana [22].

O ponto de partida deste esquema é considerar, para conjuntos S e X, uma função objetivo $\phi : S \times X \to \overline{\mathbb{R}}$ e um parâmetro $x \in X$, de modo que para cada x, temos o problema de otimização $(\mathcal{P}_x)$ dado por

$$(\mathcal{P}_x) \qquad \text{minimizar } \phi(s, x);$$

Para $x_0 \in X$ fixado, esta família de problemas é dita perturbação do *Problema Primal* $(\mathcal{P}_0)$, o qual é definido no caso em que $x = x_0$,

$$(\mathcal{P}_0) \qquad \text{minimizar } \phi(s, x_0);$$

O conjunto X é dito espaço de perturbação. A função de perturbação associada é $p : X \to \overline{\mathbb{R}}$ dada por $p(x) := \inf_{s \in S} \phi(s, x)$.

Vamos considerar um conjunto Y e uma função de acoplamento $c : X \times Y \to \overline{\mathbb{R}}$. Deste modo, podemos associar $(\mathcal{P}_0)$ ao seguinte problema

$$(\mathcal{D}_c) \qquad \text{maximizar } c(x_0, y) - p^c(y);$$

Este será dito *Problema Dual* e denotado por $(\mathcal{D}_c)$. Notemos que o valor ótimo do Problema Primal é dado por $p(x_0)$ e o valor ótimo de $(\mathcal{D}_c)$ é dado por $p^{cc'}(x_0)$.

Quando X possuir uma estrutura de espaço vetorial, por simplicidade, vamos adotar $x_0 = 0$ em todo o trabalho. No próximo exemplo, vamos ver como a teoria de dualidade Lagrangiana se adapta para este esquema em que c será dada como no Exemplo 1.4.

Exemplo 1.11 *Consideremos as funções $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ e $g : \mathbb{R}^n \to \mathbb{R}^m$ no seguinte problema*

$$\begin{aligned} &\textit{minimizar} \quad f(s) \\ &\textit{sujeito a} \quad g(s) \leq 0 \end{aligned} \tag{1.2}$$

em que a desigualdade $\leq$ é no sentido de componente a componente. A maneira usual de introduzir uma família de perturbações para este problema é considerar a função objetivo

$\phi : \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$ *dada por*

$$\phi(s,x) = \begin{cases} f(s) & \textit{se } g(s) + x \le 0 \\ +\infty & \textit{caso contrário,} \end{cases}$$

e $p(x) = \inf\limits_{s \in \mathbb{R}^n} \phi(s,x)$. *Assim* $p(0)$ *é equivalente ao Problema 1.2. Notemos que* $X = \mathbb{R}^m$ *é um espaço de Hilbert. Pelo Teorema de Representação de Riesz-Fréchet [9, Teorema 5.5], temos que* $Y = (\mathbb{R}^m)^* \approx \mathbb{R}^m$ *e o produto de dualidade pode ser considerado como o produto interno em* $\mathbb{R}^m$. *Deste modo, o Dual para este problema é dado por*

$$\begin{aligned} &\textit{maximizar} \quad -p^*(y) \\ &\textit{sujeito a} \quad y \in \mathbb{R}^m, \end{aligned}$$

pois $c(0,y) = \langle 0, y \rangle = 0$. *Calculando a conjugada de* p, *vemos que*

$$\begin{aligned} -p^*(y) &= -\sup_{x \in \mathbb{R}^m} \{\langle x, y \rangle - p(x)\} \\ &= \inf_{x \in \mathbb{R}^m} \{ \inf_{s \in \mathbb{R}^n} \phi(s,x) - \langle x, y \rangle \} \\ &= \inf_{(s,x) \in \mathbb{R}^n \times \mathbb{R}^m} \{\phi(s,x) - \langle x, y \rangle\} \\ &= \inf_{s \in \mathbb{R}^n} \inf_{g(s) \le -x} \{f(s) - \langle x, y \rangle\} \\ &= \inf_{s \in \mathbb{R}^n} \{f(s) + \inf_{g(s) \le -x} \{-\langle x, y \rangle\}\} \\ &= \inf_{s \in \mathbb{R}^n} \{f(s) - \sup_{x \le -g(s)} \{\langle x, y \rangle\}\}. \end{aligned}$$

Como

$$\sup_{x \le -g(s)} \langle x, y \rangle = \begin{cases} \langle -g(s), y \rangle, \ y \ge 0 \\ +\infty, \ y < 0 \end{cases} \tag{1.3}$$

então concluimos que $-p^*(y) = \inf_{s \in \mathbb{R}^n} \{f(s) + \langle g(s), y \rangle\}$ *para* $y \ge 0$ *e* $-p^*(y) = -\infty$ *para* $y < 0$.

Recordando que a função Lagrangiana clássica para o Problema 1.2 é dada por $L(s,y) = f(s) + \langle g(s), y \rangle$ *[22], o seu problema Dual pode ser reescrito como*

$$p^{**}(0) = \sup_{y \in \mathbb{R}^m} -p^*(y) = \sup_{y \ge 0} \inf_{s \in \mathbb{R}^n} L(s,y),$$

coincidindo com a dualidade Lagrangiana.

Utilizando as ideias desenvolvidas até o momento, obtemos os seguintes resultados.

Teorema 1.12 *Dados problemas Primal* $(\mathcal{P}_0)$ *e Dual* $(\mathcal{D}_c)$, *é correto afirmar:*

(i) *se o valor ótimo de* $(\mathcal{D}_c)$ *é finito então o seu conjunto solução é exatamente* $\partial_c p^{cc'}(x_0)$.

(ii) *O valor ótimo de* $(\mathcal{D}_c)$ *não é maior que o valor ótimo de* $(\mathcal{P}_0)$. *Eles coincidem se, e somente se, a função* p *é* c*-supremo em* x_0. *Neste caso, se o valor ótimo é finito então o conjunto solução de* $(\mathcal{D}_c)$ *é* $\partial_c p(x_0)$.

(iii) *Se* $s \in S$ *e* $y \in Y$ *satisfazem* $\phi(s, x_0) = c(x_0, y) - p^c(y)$ *então eles são soluções para o problema Primal e Dual, respectivamente.*

Demonstração. (i) Se o valor ótimo de $(\mathcal{D}_c)$ é finito então há duas possibilidades. A primeira é que exista solução $y \in Y$ para o problema $(\mathcal{D}_c)$. Mas y é solução de $(\mathcal{D}_c)$ se, e só se $p^{cc'}(x_0) = c(x_0, y) - p^c(y)$ que é equivalente, pela Proposição 1.9, a $y \in \partial_c p^{cc'}(x_0)$.

A segunda possibilidade é o conjunto solução de $(\mathcal{D}_c)$ ser vazio, que é equivalente a dizer que $p^{cc'}(x_0) > c(x_0, y) - p^c(y)$ para todo $y \in Y$, ou seja, novamente pela Proposição 1.9, $\partial_c p^{cc'}(x_0) = \emptyset$.

(ii) Como $p(x_0) \geq p^{cc'}(x_0)$ e $p(x_0) = p^{cc'}(x_0)$ se, e só se p é c-supremo em x_0, obtemos o primeiro resultado. Se o valor ótimo é finito então, pelo item anterior, o conjunto solução de $(\mathcal{D}_c)$ é $\partial_c p^{cc'}(x_0)$, que é igual a $\partial_c p(x_0)$, uma vez que p é c-supremo em x_0.

(iii) É consequência direta do fato de que

$$\inf_{s \in S} \phi(s, x_0) = p(x_0) \geq p^{cc'}(x_0) = \sup_{y \in Y}\{c(x_0, y) - p^c(y)\}.$$

□

Quando a função p não é c-supremo em x_0, dizemos que há um *gap de dualidade* entre os problemas Primal e Dual. O valor desse *gap* é dado pela diferença entre os valores ótimos dos problemas Primal e Dual respectivamente. Assim, pelo item *(ii)* do teorema anterior, sabemos que este *gap* é sempre não-negativo e é nulo se, e somente se, p é c-supremo em x_0.

Capítulo 2

Conjugação para funções semicontínuas inferiormente

Neste capítulo vamos estudar alguns casos particulares de c-conjugada. A teoria relacionada foi apresentada em [11, 19] e buscava desenvolver uma conjugação que se comportasse bem para funções sci. Seguindo esta linha, vamos apresentar o nosso esquema de simplificação dessa teoria e suas consequências para a área de Otimização.

2.1 Primeiras considerações

Antes de dar início à teoria, vamos fazer algumas considerações sobre o espaço em que as funções estarão definidas neste capítulo. Reforçamos que, para um espaço métrico E e seu dual topológico E^*, $\langle \cdot, \cdot \rangle$ irá denotar o produto de dualidade de E^* com E. Quando $E = H$ um espaço de Hilbert então $H^* \approx H$ [9, Teorema 5.5] e $\langle \cdot, \cdot \rangle$ irá ser exatamente o produto interno de H. Recordamos, a seguir, a definição de Fréchet diferenciabilidade, que é uma extensão natural de diferenciabilidade para espaços de dimensão infinita.

Definição 2.1 *Para um espaço vetorial métrico E, dizemos que uma função $f : E \to \mathbb{R}$ é Fréchet diferenciável em um ponto $x \in E$ quando existe um operador linear contínuo $x^* : E \to \mathbb{R}$ tal que*

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - \langle x, x^* \rangle}{\|h\|_E} = 0$$

em que o limite a ser considerado é o limite de funções definidas em espaços métricos. Deste modo, x^* será dito gradiente de f no ponto x e denotado por $\nabla f(x)$.

Somente **neste capítulo**, $(X, \|\cdot\|)$ será considerado um espaço real de Banach cuja norma é Fréchet diferenciável fora da origem. Espaços de Hilbert são exemplos de tais espaços, bem como espaços

$$L^p(X, \mu) = \left\{ f : A \to \mathbb{R} \mid A \subset X,\ f \text{ é mensurável e } \int_A \|f(x)\|^p \, d\mu(x) < \infty \right\}$$

em que μ é uma medida em X e $1 < p < \infty$ [[28], Teorema 2.5]. Estamos interessados nestes espaços porque eles podem ser caracterizados através de propriedades do seu operador de dualidade, como veremos a seguir.

Recordamos que a notação $X \rightrightarrows X^*$ é para denotar que um operador é do tipo ponto-conjunto, ou seja, ele associa um ponto de X a um subconjunto de X^*. Quando o operador for do tipo ponto-ponto, usaremos a notação natural $X \to X^*$.

Definição 2.2 *O operador de dualidade $J : X \rightrightarrows X^*$ é definido como*

$$J(x) := \left\{x^* \in X^* \mid \|x\|^2 = \langle x, x^* \rangle = \|x^*\|_*^2\right\}.$$

O operador de dualidade é também conhecido na literatura como $J := \partial\left(\frac{1}{2}\|\cdot\|^2\right)$, ou seja, como o subdiferencial de Fenchel da função $\frac{1}{2}\|\cdot\|^2$. Como o subdiferencial de funções convexas de valores finitos é não-vazio para todo ponto, então o operador está bem definido. Além disso, diretamente da sua definição, concluímos que $J(0) = 0$ e J é contínuo em 0 referente às normas de X e X^*. A respeito dos demais pontos de X, temos o seguinte resultado.

Teorema 2.3 *[31] Uma função convexa é Fréchet diferenciável se, e somente se, seu subdiferencial é um operador do tipo ponto-ponto e contínuo (no sentido das normas de X e de X^*).*

Graças a esse resultado, para um espaço de Banach X cuja norma é Fréchet diferenciável fora da origem, temos que $J(x)$ é do tipo ponto-ponto e contínuo para todo x. Esta caracterização é fundamental para provar os resultados deste capítulo e por esse motivo, optamos por utilizar tais espaços.

Notamos que no caso em que $X = H$ um espaço de Hilbert, J coincide com o operador $I : H \to H^*$ definido como $I(x) := \langle \cdot, x \rangle$. Por abuso de linguagem, dizemos que esse é o *operador Identidade.*

Para finalizar essas primeiras considerações, vamos identificar, por abuso de linguagem, $J(x)$ como o único elemento da imagem do operador. Dessa maneira, $J(x)$ será caracterizado em X pela igualdade

$$\|x\|^2 = \langle x, J(x) \rangle = \|J(x)\|_*^2. \tag{2.1}$$

2.2 Caracterização de funções sci via supremos

Nesta seção vamos apresentar caracterizações de funções sci através de supremos de conjuntos particulares de funções. A cada uma destas caracterizações, é possível associar uma teoria de c-conjugação como vimos no capítulo anterior. Na Seção 2.3, vamos mostrar como construir esta associação para uma caracterização de nosso interesse. Estes resultados são baseados em [11, 19]. Apresentaremos também uma caracterização original que pode ser encontrada em [16].

A primeira caracterização a ser exibida faz uso de um caso particular do Teorema de Seleção de Michael [25, Teorema 3.1]. Vamos explicar sucintamente a ideia deste teorema.

Para um operador $T : X \rightrightarrows \mathbb{R}$, uma *seleção* para T é uma função contínua $h : X \to \mathbb{R}$ tal que

$$h(x) \in T(x)$$

para cada $x \in X$. Sob certas hipóteses a respeito de T e se $T(x)$ é convexo e fechado em $\mathbb{R}$ para todo $x \in X$, o teorema de Seleção de Michael garante a existência desta seleção.

Em particular, para funções $f : X \to \mathbb{R} \cup \{+\infty\}$, $g : X \to \mathbb{R} \cup \{-\infty\}$ tais que $g \leq f$, podemos considerar o operador $L(x) = \{y \in \mathbb{R} \mid g(x) \leq y \leq f(x)\}$ e deste modo, $h : X \to \mathbb{R}$ é uma seleção para L se, e somente se, h é contínua e $g(x) \leq h(x) \leq f(x)$ para todo $x \in X$. Neste caso, as hipóteses assumidas sob L são equivalente a pedir f e g como no teorema a seguir.

Teorema 2.4 *[38, p. 133] Sejam $f : X \to \mathbb{R} \cup \{+\infty\}$ semicontínua inferiormente, $g : X \to \mathbb{R} \cup \{-\infty\}$ semicontínua superiormente (scs), e $g \leq f$. Então existe uma função contínua $h : X \to \mathbb{R}$ tal que $g \leq h \leq f$.*

Através deste teorema, podemos ver que sempre é possível se aproximar inferiormente de uma função sci através de funções contínuas. Este fato é o que nos garante o próximo lema. Para o melhor entendimento de sua demonstração, recordamos que a envoltória sci de uma função $f : X \to \mathbb{R} \cup \{+\infty\}$ é uma função sci, denotada por $cl\ f$, que satisfaz

(i) $cl\ f \leq f$,

(ii) para qualquer função sci $r : X \to \mathbb{R} \cup \{+\infty\}$ que seja majorada por f, teremos que $r \leq cl\ f$.

Como resultado bem conhecido, f é sci em um ponto x se, e somente se, $cl\ f(x) = f(x)$.

Lema 2.5 *Sejam $f : X \to \mathbb{R} \cup \{+\infty\}$, $x \in X$ e $\lambda \in \mathbb{R}$ tais que $f(x) \geq \lambda$. Se f é sci em x então existe uma função contínua $h : X \to \mathbb{R}$ que satisfaz*

i) *$h \leq f$;*

ii) *$f(x) \geq h(x) \geq \lambda$.*

Demonstração. Sejam λ e x tais que $\lambda \leq f(x)$ e defina $g : X \to R \cup \{-\infty\}$ por $g(x) := \lambda$ e $g(y) := -\infty$ para $y \neq x$.

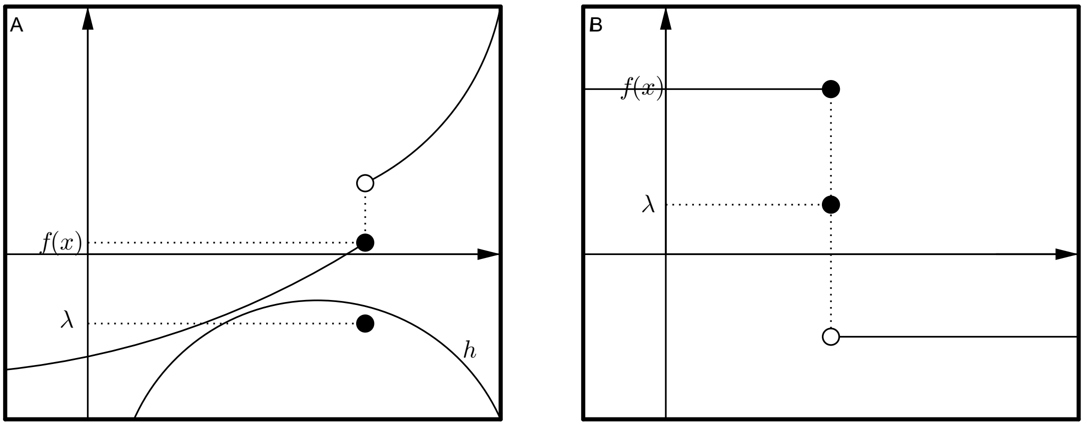

Figura 2.1: Ilustração do Lema 2.5. No quadro A, temos um exemplo de uma função f sci em x. No quadro B, vemos que a função h pode não existir caso f não seja sci.

Notemos que, como $\limsup_{y \to x} g(y) = \lambda$, podemos concluir facilmente que g é scs. Além do mais, $g(y) \leq cl\ f(y)$ para todo $y \in X$, uma vez que f é sci em x e assim $cl\ f(x) = f(x)$. Deste modo, aplicando o Teorema 2.4 para g e $cl\ f$, concluímos que existe uma função contínua $h : X \to \mathbb{R}$ tal que

$$g(y) \leq h(y) \leq cl\ f(y) \leq f(y)$$

para todo $y \in X$ e, em particular, $\lambda = g(x) \leq h(x)$ como desejado. $\square$

Observemos que para uma função f sci em x tal que $f(x) > \lambda$ é possível encontrar uma função contínua $h : E \to \mathbb{R}$ tal que $f \geq h$ e $h(x) > \lambda$, pois basta aplicar o Lema 2.5 para λ_1 tal que $f(x) \geq \lambda_1 > \lambda$.

Outro ponto importante é que, quando $f(x) \in \mathbb{R}$, podemos tomar $\lambda = f(x)$ e, consequentemente, obtemos uma função contínua h tal que $h \leq f$ e $h(x) = f(x)$. Motivados por este fato, apresentamos o primeiro resultado de caracterização para funções sci.

Teorema 2.6 *Uma função $f : X \to \mathbb{R} \cup \{+\infty\}$ é sci em $x \in X$ se, e somente se,*

$$f(x) = \sup \{h(x) \mid h : X \to \mathbb{R} \text{ é contínua, } h \leq f\}, \tag{2.2}$$

e o supremo é atingido se $x \in dom\ f$.

Demonstração. Primeiramente vamos supor que f é sci em x. É imediato que $f(x)$ é uma cota superior de $\{h(x) \mid h : X \to \mathbb{R} \text{ é contínua, } h \leq f\}$. Vamos verificar que é a menor delas.

Para todo $\lambda \in \mathbb{R}$ tal que $f(x) > \lambda$ existe, pelo Lema 2.5, uma função contínua $h : X \to \mathbb{R}$ satisfazendo $h \leq f$ e $h(x) > \lambda$. Por definição de supremo concluímos o resultado. Caso $x \in dom\ f$, podemos tomar $\lambda = f(x)$ no Lema 2.5 e concluir que o supremo é atingido.

Reciprocamente, suponha que (2.2) vale e considere a função

$$H(y) = \sup \{h(y) \mid h : X \to \mathbb{R} \text{ é contínua}, \, h \leq f\}.$$

Esta função é supremo de funções contínuas e, logo, é sci. Além do mais, H é um minorante de f e, portanto, é também um minorante de $cl f(y)$. Como, por (2.2), $H(x) = f(x)$ então $cl f(x) \geq H(x) = f(x)$, isto é, f é sci em x. □

Em [11] houve uma preocupação em reformular a caracterização anterior e extrair mais informações da função h do Teorema 2.4 com o intuito de definir uma conjugada relacionada cuja fórmula fosse mais similar a de Fenchel. No Teorema 2.8 podemos ver um complemento dessa teoria [11, Seção 4]. No entanto, é necessário primeiro garantir a continuidade da seguinte família de funções.

Proposição 2.7 *Consideremos uma função contínua $g : X \to \mathbb{R}$ e $\delta > 0$. Seja $p : X \to X^*$ definida como*

$$p(y) = \begin{cases} g(y)\frac{J(y)}{\|y\|^2}, & \|y\| > \delta \\ g(y)\frac{J(y)}{\delta^2}, & \|y\| \leq \delta \end{cases},$$

em que J satisfaz (2.1)*. Então p é contínua.*

Demonstração. Vamos verificar que p é contínua usando a continuidade de g. Notemos primeiramente que no aberto $B(0, \delta)$ p é expressa como produto de funções contínuas. Analogamente, p também é expressa como produto de funções contínuas no complementar de $\overline{B(0, \delta)}$ ($\|y\| > \delta$), que também é um aberto em X. Assim, basta verificar a continuidade de p na fronteira de $\overline{B(0, \delta)}$ ($\|y\| = \delta$).

Dado $y \in X$ tal que $\|y\| = \delta$, consideremos (y_k) uma sequência que converge para y. Queremos provar que $(p(y_k))$ converge para $p(y)$. Inicialmente, pela continuidade de g e da norma, notemos que

$$g(y_k) \longrightarrow g(y), \tag{2.3}$$

$$\|y_k\| \longrightarrow \|y\| = \delta. \tag{2.4}$$

Sejam

$$\mathbb{N}' = \{k \in \mathbb{N} | \; \|y_k\| > \delta\} \quad e \quad \mathbb{N}'' = \{k \in \mathbb{N} | \; \|y_k\| \leq \delta\}.$$

Caso o conjunto $\mathbb{N}'$ seja finito, temos que existe $k_0 \in \mathbb{N}$ tal que $k \in \mathbb{N}''$ para todo $k \geq k_0$. Assim,

$$p(y_k) = g(y_k)\frac{J(y_k)}{\delta^2}$$

para todo $k \geq k_0$. Por (2.3),

$$p(y_k) = g(y_k)\frac{J(y_k)}{\delta^2} \longrightarrow g(y)\frac{J(y)}{\delta^2} = p(y).$$

Portanto $p(y_k)$ converge para $p(y)$.

Caso o conjunto $\mathbb{N}''$ seja finito, seguindo o raciocínio análogo ao anterior, temos que existe $k_2 \in \mathbb{N}$ tal que $k \in \mathbb{N}'$ para todo $k \geq k_2$. Dessa maneira,

$$p(y_k) = g(y_k)\frac{J(y_k)}{\|y_k\|^2}$$

para todo $k \geq k_2$. Por (2.3) e (2.4),

$$p(y_k) = g(y_k)\frac{J(y_k)}{\|y_k\|^2} \longrightarrow g(y)\frac{J(y)}{\|y\|^2} = g(y)\frac{J(y)}{\delta^2} = p(y).$$

e novamente $p(y_k)$ converge para $p(y)$.

Caso ambos $\mathbb{N}'$ e $\mathbb{N}''$ sejam infinitos, consideremos as duas subsequências $(p(y_k))_{k\in\mathbb{N}'}$ e $(p(y_k))_{k\in\mathbb{N}''}$. Pelo que vimos anteriormente, podemos concluir que

$$p(y_k) \xrightarrow{\mathbb{N}'} p(y) \quad e \quad p(y_k) \xrightarrow{\mathbb{N}''} p(y).$$

Assim, dado $\epsilon > 0$ existem $k_1 \in \mathbb{N}'$ e $k_2 \in \mathbb{N}''$ tais que

$$\|p(y_k) - p(y)\| = \left\| g(y_k)\frac{J(y_k)}{\|y_k\|^2} - g(y)\frac{J(y)}{\delta^2} \right\| < \epsilon \tag{2.5}$$

para todo k em $\mathbb{N}'$ maior que k_1 e

$$\|p(y_k) - p(y)\| = \left\| g(y_k)\frac{J(y_k)}{\delta^2} - g(x)\frac{J(x)}{\delta^2} \right\| < \epsilon \tag{2.6}$$

para todo k em $\mathbb{N}''$ maior que k_2.

Tomemos $\overline{k} = max\,\{k_1, k_2\}$ e assim cada $k \geq \overline{k}$ pertence a $\mathbb{N}'$ ou a $\mathbb{N}''$. Por (2.5) e (2.6), em ambos os casos

$$\|p(y_k) - p(y)\| < \epsilon.$$

Portanto, $(p(y_k))$ converge para $p(y)$. Dessa maneira, concluímos que p é contínua. $\square$

Notamos que aqui é necessário que o espaço X seja de Banach com norma Fréchet diferenciável fora da origem, pois estamos usando a continuidade de $J(x)$, bem como o fato de ser um operador ponto-ponto.

Teorema 2.8 *Seja* $f : X \to \mathbb{R} \cup \{+\infty\}$ *e* $x \in dom\ f \backslash \{0\}$. *Então* f *é sci em* x *se, e* *somente se, existe uma aplicação contínua* $p : X \to X^*$ *tal que* $p(0) = 0$ *e*

$$f(y) \geq f(x) + \langle y, p(y)\rangle - \langle x, p(x)\rangle \ \textit{para todo } y \in X. \tag{2.7}$$

Demonstração. Se f é sci em x, então, pelo Lema 2.5, existe uma função contínua $h : X \to \mathbb{R}$ tal que $h \leq f$ e $h(x) = f(x)$. Seja $\gamma := \inf_{\|y\| \leq \|x\|} h(y)$ e defina $p : X \to X^*$ por

$$p(y) = \begin{cases} (h(y) - \gamma)\frac{J(y)}{\|y\|^2}, & \|y\| > \|x\| \\ (h(y) - \gamma)\frac{J(y)}{\|x\|^2}, & \|y\| \leq \|x\| \end{cases}.$$

Tomando $\delta = \|x\|$ na Proposição 2.7, concluímos facilmente que p é contínua. Além do mais, de (2.1), concluímos que p satisfaz

$$\langle y, p(y)\rangle = (h(y) - \gamma)\left(\min\left\{\frac{\|y\|}{\|x\|}, 1\right\}\right)^2.$$

Portanto, $\langle y, p(y)\rangle \leq h(y) - \gamma \leq f(y) - \gamma$ para todo $y \in X$, que implica $\langle y, p(y)\rangle - f(y) \leq -\gamma$ para todo $y \in X$. Além do mais, $\langle x, p(x)\rangle = h(x) - \gamma = f(x) - \gamma$. Consequentemente,

$$f(y) \geq f(x) + \langle y, p(y)\rangle - \langle x, p(x)\rangle$$

para todo $y \in X$.

Por outro lado, assuma que (2.7) vale, e defina $h_0 : X \to \mathbb{R}$ por

$$h_0(y) := f(x) + \langle y, p(y)\rangle - \langle x, p(x)\rangle. \tag{2.8}$$

Claramente $h_0 \leq f$ e, como h_0 é composição de funções contínuas, é também contínua. Logo

$$\sup\{h(x) \mid h : X \to \mathbb{R} \text{ é contínua}, h \leq f\} \geq h_0(x) = f(x).$$

Assim, pelo Teorema 2.6, concluímos que f é sci em x. □

Notemos que, quando f é sci e convexa, podemos tomar p como sendo o operador identidade, e assim recuperar a Teoria de Fenchel. Podemos observar que este teorema é um aprimoramento do Teorema 2.6 para os pontos do domínio da função. O Lema 4.5 e o Teorema 4.6 em [11] apresentam uma versão estendida deste resultado para todo o espaço.

Após essa rápida revisão desses resultados, retomamos os objetivos dessa primeira parte do trabalho apontados na Introdução. Comparando os Teoremas 2.6 e 2.8, percebemos que na segunda situação estamos trabalhando com um espaço mais complexo, uma vez que no Teorema 2.8 as funções contínuas são de $X \to X^*$ e no Teorema 2.6 são de $X \to \mathbb{R}$. Outro ponto que deve ser notado é que podemos associar uma única função h a uma infinidade de funções p. Vamos verificar isso com um exemplo simples: consideremos $X = \mathbb{R}^2$ e $q_\alpha(x_1, x_2) = (\alpha x_2, -\alpha x_1)$ para $\alpha \in \mathbb{R}$. Levando em conta a associação dada em (2.8), consideramos

$$h\,(y) := f(x) + \langle y, p(y)\rangle - \langle x, p(x)\rangle\,.$$

Como $\langle y, q_\alpha(y)\rangle = \alpha y_1 y_2 - \alpha y_1 y_2 = 0$ para todo $y \in X$, temos que $h\,(y) := f(x) + \langle y, p(y) + q_\alpha(y)\rangle - \langle x, p(x) + q_\alpha(x)\rangle$ para qualquer $\alpha \in \mathbb{R}$.

Veremos mais adiante que, em problemas de Otimização, é importante encontrar a função h que satisfaça (2.2) (quando o supremo é atingido), o que é equivalente a encontrar p que satisfaça (2.7). Deste modo, pelas razões apontadas anteriormente, podemos dizer que é "mais fácil" encontrar h do que p.

A nossa proposta é investigar o conjunto das funções do tipo $\langle \cdot, p(\cdot)\rangle$ com o objetivo de desenvolver uma caracterização de funções sci que mantenha as informações adquiridas no Teorema 2.8 mas leve em conta os dois pontos apontados anteriormente. Encontramos a seguinte equivalência.

Teorema 2.9 *Consideremos uma função $f : X \to \mathbb{R}$. As seguintes afirmações são equivalentes:*

(i) *Existe uma aplicação contínua $p : X \to X^*$ tal que $f(x) = \langle x, p(x)\rangle$ para todo $x \in X$.*

(ii) *A função f é contínua em X, Fréchet diferenciável em 0 e satisfaz $f\,(0) = 0$.*

Demonstração. $(i) \Rightarrow (ii)$ A igualdade $f\,(0) = 0$ é óbvia e a continuidade da f segue da continuidade de p. Agora notemos que

$$0 \leq \frac{|f(x) - f(0) - \langle x, p\,(0)\rangle\,|}{\|x\|} = \frac{|\langle x, p(x) - p(0)\rangle|}{\|x\|}.$$

Pela Desigualdade de Schwarz, temos que

$$\frac{|\langle x, p(x) - p(0)\rangle|}{\|x\|} \leq \frac{\|x\|\,\|p(x) - p(0)\|_*}{\|x\|} = \|p(x) - p(0)\|_*\,.$$

Como p é contínua em 0, concluímos que

$$0 \leq \lim_{x\to 0} \frac{|f(x) - f(0) - \langle x, p\,(0)\rangle\,|}{\|x\|} \leq \lim_{x\to 0} \|p(x) - p(0)\|_* = 0$$

e portanto, f é Fréchet diferenciável em 0 com $\nabla f(0) = p(0)$.

$(ii) \Rightarrow (i)$ Defina $p : X \to X^*$ por

$$p(x) = \begin{cases} \nabla f(0) + \frac{f(x) - \langle x, \nabla f(0)\rangle}{\|x\|^2} J(x), & \text{se } x \neq 0 \\ \nabla f(0), & \text{caso contrário.} \end{cases} \tag{2.9}$$

Como f é contínua, então p é composição de funções contínuas em $X \backslash \{0\}$. Portanto, resta mostrar continuidade apenas em 0. Observe que, de (2.1), obtemos

$$\begin{aligned}\lim_{x\to 0}\|p(x)-p(0)\|_* &= \lim_{x\to 0}\frac{|f(x)-\langle x,\nabla f(0)\rangle|}{\|x\|^2}\|J(x)\|_*\\ &= \lim_{x\to 0}\frac{|f(x)-\langle x,\nabla f(0)\rangle|}{\|x\|}=0.\end{aligned}$$

A última igualdade é consequência da Fréchet diferenciabilidade de f em 0. Assim, concluímos que p é contínua em X. Além do mais, novamente por (2.1), concluímos que, para $x \in X \setminus \{0\}$,

$$\begin{aligned}\langle x,p(x)\rangle &= \langle x,\nabla f(0)+\frac{f(x)-\langle x,\nabla f(0)\rangle}{\|x\|^2}J(x)\rangle\\ &= \langle x,\nabla f(0)\rangle+\frac{f(x)-\langle x,\nabla f(0)\rangle}{\|x\|^2}\langle x,J(x)\rangle = f(x).\end{aligned}$$

Como, para $x=0$, a igualdade $f(x)=\langle x,p(x)\rangle$ é óbvia, concluímos a demonstração. □

O resultado acima foi obtido de maneira independente, mas pode ser encontrado também no artigo recente [30, Teorema 3.11]. Deste ponto até o fim do capítulo, os resultados a serem apresentados são originais e podem ser encontrados em [16].

Notemos que, no resultado acima, a condição $f(0)=0$ pode ser facilmente omitida, pois para uma função f contínua em X, Fréchet diferenciável em 0, podemos aplicar o Teorema 2.9 para a função $h(x)=f(x)-f(0)$. Logo, isto é equivalente a dizer que existe $p: X\to X^*$ contínua tal que $f(x)=\langle x,p(x)\rangle+f(0)$.

Outra observação importante é que, para cada x, quando $\nabla f(0)=0$, $p(x)$ definida por (2.9) é a solução do seguinte problema de otimização

$$\begin{aligned}&\text{minimizar} && \|y^*\|_{X^*}\\ &\text{sujeito a} && \langle x,y^*\rangle = f(x).\end{aligned}$$

Isso é óbvio para $x=0$. Para $x\neq 0$, notamos que, da definição de norma dual, para qualquer y^* satisfazendo a restrição, temos que

$$\|y^*\|_*\geq \frac{|\langle x,y^*\rangle|}{\|x\|}=\frac{|f(x)|}{\|x\|^2}\|J(x)\|_*=\|p(x)\|_*.$$

Quando $\nabla f(0)\neq 0$, $p(x)$ é obtida adicionando o operador constante $T: X\to X^*$ dado por $T(y):=\langle\cdot,\nabla f(0)\rangle$ à solução do problema

$$\begin{aligned}&\text{minimizar} && \|y^*\|_{X^*}\\ &\text{sujeito a} && \langle x,y^*\rangle = f(x)-\langle x,\nabla f(0)\rangle.\end{aligned}$$

O próximo resultado é apenas uma adaptação do Teorema 2.8 através da equivalência do Teorema 2.9.

Proposição 2.10 *Seja* $f : X \to \mathbb{R} \cup \{+\infty\}$ *e* $x \in dom\ X \backslash \{0\}$. *Se* f *é sci em* x *então existe uma função contínua* $h : X \to \mathbb{R}$ *Fréchet diferenciável em* 0, *com* $\nabla h(0) = 0$, *majorada por* f *tal que*

$$f\,(x) = h\,(x)\,.$$

Demonstração. Consideremos p como no Teorema 2.8 e definamos $h : X \to \mathbb{R}$ por $h\,(y) := \langle y, p(y)\rangle + f(x) - \langle x, p(x)\rangle$. Do Teorema 2.9, implicação (i) $\Longrightarrow$ (ii), temos que h é contínua em X, Fréchet diferenciável em 0, com $\nabla h(0) = p(0)$. Do Teorema 2.8, temos que $p(0) = 0$. Além do mais, por (2.7), nós temos que $h \leq f$ e claramente $h(x) = f(x)$. □

Vamos estender agora o resultado anterior para pontos não-nulos que não estejam no domínio. Nossa estratégia para provar a extensão é fazer uso das funções de valores reais apresentadas no seguinte lema.

Lema 2.11 *Seja* $f : X \to \mathbb{R} \cup \{+\infty\}$ *e para cada* $\lambda \in \mathbb{R}$ *consideremos a função de valores reais* $f_\lambda : X \to \mathbb{R}$ *definida por* $f_\lambda(x) := \min\{f(x), \lambda\}$. *Então*

$$f(x) = \sup_{\lambda \in \mathbb{R}}\{f_\lambda(x) \mid \lambda \leq f(x)\}$$

para todo $x \in X$.

Demonstração. Notemos que $f_\lambda \leq f$ para todo $\lambda \in \mathbb{R}$. Logo,

$$f(x) \geq \sup_{\lambda \in \mathbb{R}}\{f_\lambda(x) \mid \lambda \leq f(x)\}$$

para todo $x \in X$. Para provar que a igualdade vale, vamos considerar dois casos: se $x \in dom\ f$ então consideremos $\overline{\lambda} = f(x)$ e portanto $f_{\overline{\lambda}}(x) = f(x)$. Caso $x \notin dom\ f$, temos que $f_\lambda(x) = \lambda$ para todo $\lambda \in \mathbb{R}$. Portanto $\sup_{\lambda \in \mathbb{R}}\{f_\lambda(x) \mid \lambda \leq f(x)\} = \sup_{\lambda \in \mathbb{R}} \lambda = +\infty = f(x)$. □

Proposição 2.12 *Seja* $f : X \to \mathbb{R} \cup \{+\infty\}$ *e* $x \in X \backslash \{0\}$ *tal que* $f(x) = +\infty$. *Se* f *é sci em* x *então*

$$\begin{aligned} f\,(x) = \sup\{h\,(x) \mid h : X \to \mathbb{R} \textit{ é contínua em } X \textit{ e Fréchet} \\ \textit{diferenciável em } 0,\ \nabla h(0) = 0,\ h \leq f\}. \end{aligned} \tag{2.10}$$

Demonstração. Dado um $\lambda \in \mathbb{R}$, consideremos f_λ definida como no Lema 2.11. Como f é sci em x, logo

$$cl\ f_\lambda(x) = \min\{cl\ f(x), \lambda\} = \min\{f(x), \lambda\} = f_\lambda(x).$$

Então f_λ também é sci em x e como $f_\lambda(x) < +\infty$, pela Proposição 2.10, existe uma função contínua $h_\lambda : X \to \mathbb{R}$ Fréchet diferenciável em 0, com $\nabla h_\lambda(0) = 0$, majorada por f_λ tal que $f_\lambda(x) = h_\lambda(x)$. Assim,

$$\sup_{\lambda\in\mathbb{R}}\{f_\lambda(x) \mid \lambda \le f(x)\} = \sup_{\lambda\in\mathbb{R}}\{h_\lambda(x) \mid \lambda \le f(x)\}.$$

Pelo Lema 2.11, $f(x) = \sup_{\lambda\in\mathbb{R}}\{h_\lambda(x) \mid \lambda \le f(x)\}$ e como cada h_λ é contínua, Fréchet diferenciável em 0, com $\nabla h_\lambda(0) = 0$ e $h_\lambda \le f_\lambda \le f$, concluímos o resultado desejado. □

Finalmente, usando os resultados anteriores, estabelecemos nossa caracterização para todos os pontos de X.

Teorema 2.13 *Seja $f : X \to \mathbb{R} \cup \{+\infty\}$ e $x \in X$. Então f é sci em x se, e somente se,*

$$f(x) = \sup\{h(x) \mid h : X \to \mathbb{R} \text{ é contínua em } X \text{ e Fréchet diferenciável em } 0,\ \nabla h(0) = 0,\ h \le f\}; \tag{2.11}$$

o supremo é atingido se $x \in dom\ f \backslash \{0\}$.

Demonstração. ($\Rightarrow$) Basta demonstrar o caso 0. Obviamente $f(0)$ é uma cota superior de $\{h(0) \mid h : X \to \mathbb{R}$ é contínua em X e Fréchet diferenciável em 0, $\nabla h(0) = 0,\ h \le f\}$. Vamos provar que é a menor das cotas. Assim, seja $\lambda < f(0)$ e tome $\lambda_1 \in \mathbb{R}$ tal que $f(0) > \lambda_1 > \lambda$. Pelo Lema 2.5 existe uma função contínua $h : X \to \mathbb{R}$ tal que $\lambda_1 < h(0) \le f(0)$ e $h \le f$. Consideremos $h_{\lambda_1}(y) := \min\{h(y), \lambda_1\}$ e como h é contínua então h_{λ_1} também é contínua. Logo, desde que $h(0) > \lambda_1$, $h_{\lambda_1}(y) = \lambda_1$ em uma vizinhança de 0. Portanto, h_{λ_1} é Fréchet diferenciável em 0 com $\nabla h_{\lambda_1}(0) = 0$. Além do mais, $h_{\lambda_1} \le h \le f$ e $h_{\lambda_1}(0) = \lambda_1 > \lambda$, concluindo o desejado.

($\Leftarrow$) De (2.11), concluímos que f é supremo de funções contínuas no ponto x, portanto f é sci. Da Proposição 2.10, concluímos que o supremo em (2.11) é atingido para todo $x \in dom\ f \backslash \{0\}$. □

Observemos que o supremo em (2.11) somente é atingido em 0 quando existe uma função contínua Fréchet diferenciável em 0, cujo gradiente se anula na origem, que minorize f e em 0 assumam o mesmo valor. Veremos na próxima seção que isto está relacionado com a *Fréchet subdiferenciabilidade* [ver Definição 2.24] da função na origem. De qualquer maneira, vamos ver a seguir que podemos dar liberdade para o ponto em que as funções h sejam diferenciáveis, usando argumentos de translação no Teorema 2.9.

Corolário 2.14 *Consideremos uma função contínua $f : X \to \mathbb{R}$ e $y \in X$. Então f é Fréchet diferenciável em y e satisfaz $f(y) = 0$ se, e somente se, existe uma aplicação contínua $p : X \to X^*$ tal que $f(x) = \langle x - y, p(x)\rangle$ para todo $x \in X$.*

Demonstração. Consideremos $g(x) = f(x+y)$ para todo $x \in X$. Sabemos que f é Fréchet diferenciável em y e satisfaz $f(y) = 0$ se, e somente se, g é Fréchet diferenciável em 0 e $g(0) = f(y) = 0$. Logo, pelo Teorema 2.9, existe uma aplicação contínua $q : X \to X^*$

tal que $g(x) = \langle x, q(x)\rangle$ para todo $x \in X$ e, logo, $f(x) = g(x-y) = \langle x-y, q(x-y)\rangle$. Considerando $p(x) = q(x-y)$, claramente $f(x) = \langle x-y, p(x)\rangle$ para todo $x \in X$.

Por outro lado, se $f = \langle \cdot - y, p(\cdot)\rangle$ para alguma aplicação contínua p, então $g(x) = \langle x, p(x+y)\rangle$. Como $p(\cdot+y)$ também é contínua, então pelo Teorema 2.9, concluímos o resultado. $\square$

Consequentemente podemos dar a mesma liberdade no Teorema 2.13.

Corolário 2.15 *Seja $f : X \to \mathbb{R} \cup \{+\infty\}$ e $x, y \in X$. Então f é sci em x se, e somente se,*

$$f(x) = \sup\{h(x) \mid h : X \to \mathbb{R} \textit{ é contínua em } X \textit{ e Fréchet } \textit{diferenciável em } y,\ \nabla h(y) = 0,\ h \le f\};$$

o supremo é atingido se $x \in dom\ f \backslash \{y\}$.

Demonstração. Consideremos novamente $g(z) := f(z+y)$ e notemos que, se f é sci em x então g é sci em $x-y$. Logo, pelo Teorema 2.13,

$$f(x) = g(x-y) = \sup\{r(x-y) \mid r : X \to \mathbb{R} \text{ é contínua em } X \text{ e Fréchet diferenciável em } 0,\ \nabla r(0) = 0,\ r \le g\}$$

e o supremo é atingido quando $x-y \in dom\ g\backslash\{0\}$, ou, equivalentemente, $x \in dom\ f\backslash\{y\}$.

Considerando a mudança de variável $h(z) = r(z-y)$, obtemos que

$$\begin{aligned} f(x) &= \sup\{r(x-y) \mid r : X \to \mathbb{R} \text{ é contínua em } X \text{ e Fréchet diferenciável em } 0,\ \nabla r(0) = 0,\ r \le g\} \\ &= \sup\{h(x) \mid h : X \to \mathbb{R} \text{ é contínua em } X \text{ e Fréchet diferenciável em } y,\ \nabla h(y) = 0,\ h \le f\}. \end{aligned}$$

Por outro lado, a semicontinuidade inferior de f em x segue diretamente do fato de ser o supremo pontual de funções contínuas. $\square$

A título de curiosidade, encerramos a seção com uma versão do Corolário 2.14 para espaços de Hilbert. Como o operador de dualidade coincide com o operador identidade nestes espaços, a função p definida em (2.9) é Fréchet diferenciável em $H\backslash\{0\}$, o que implica, por argumentos de translação, o seguinte resultado.

Teorema 2.16 *Seja H um espaço de Hilbert e $y \in H$. Para $f : H \to \mathbb{R}$ as seguintes afirmações são equivalentes:*

(i) *Existe uma aplicação contínua $p : H \to H$, Fréchet diferenciável em $H\backslash\{y\}$, tal que $f(x) = \langle x-y, p(x)\rangle$ para todo $x \in H$.*

(ii) *A função f é Fréchet diferenciável e satisfaz $f(y) = 0$.*

2.3 Nossos esquemas de conjugação

Nesta seção vamos apresentar dois esquemas de conjugação para funções sci associados ao Teorema 2.13. Estas construções são baseadas no Teorema 1.7. Como ambos são casos particulares de c-conjugada, todos os resultados apresentados no Capítulo 1 são válidos neste contexto. Os resultados e definições aqui apresentados são de nossa autoria, salvo menção explícita em contrário. Os mesmos podem ser encontrados em [16].

2.3.1 O primeiro esquema de conjugação

Vamos considerar o seguinte conjunto de funções:

$$\mathcal{H}_X := \{h : X \to \mathbb{R} \mid h \text{ é contínua em } X \text{ e Fréchet} \\ \text{diferenciável em } 0,\ h(0) = 0,\ \nabla h(0) = 0\}.$$

Devido ao Corolário 2.15, poderíamos considerar funções Fréchet diferenciáveis em algum outro ponto fixado distinto de 0 para desenvolver o nosso estudo. No entanto, como os resultados são análogos, por questões de simplicidade, optamos por trabalhar no ponto 0.

Vamos considerar, deste modo, a função de acoplamento $b : X \times (X^* \times \mathcal{H}_X) \to \mathbb{R}$ definida por

$$b\,(x, (x^*, h)) := \langle x, x^* \rangle + h(x).$$

Deste modo, seguindo a ideia da Definição 1.1, apresentamos a b-conjugada.

Definição 2.17 *A função b-conjugada de $f : X \to \overline{\mathbb{R}}$ é $f^b : X^* \times \mathcal{H}_X \to \overline{\mathbb{R}}$, dada por*

$$f^b(x^*, h) = \sup_{x \in X} \{\langle x, x^* \rangle + h(x) - f(x)\}\,. \tag{2.12}$$

Notamos que a função conjugada de Fenchel [Exemplo 1.4] se relaciona com a b-conjugada da seguinte maneira:

$$f^b(x^*, h) = (f - h)^*\,(x^*)\,, \qquad f^*\,(x^*) = f^b(x^*, 0). \tag{2.13}$$

Ambas relações são consequências diretas das definições destas funções.

A função b'-conjugada naturalmente é definida como em 1.1, considerando $b' : (X^* \times \mathcal{H}_X) \times X \to \mathbb{R}$ definida por

$$b'\,((x^*, h)\,, x) := b\,(x, (x^*, h))\,.$$

Seguindo a Definição 1.8, apresentamos a noção de b-subdiferenciabilidade.

Definição 2.18 *Dizemos que $f : X \to \overline{\mathbb{R}}$ é b-subdiferenciável em $x_0 \in dom(f)$ quando* *existe* $(x^*, h) \in X^* \times \mathcal{H}_X$ *tal que*

$$f(x) - f(x_0) \geq \langle x - x_0, x^* \rangle + h(x) - h(x_0) \text{ para todo } x \in X. \tag{2.14}$$

O par (x^*, h) *é dito b-subgradiente de* f *em* x_0*. O conjunto de todos os b-subgradientes de* f *em* x_0*, denotado por* $\partial_b f(x_0)$*, é chamado de b-subdiferencial de* f *em* x_0*. Se* $f(x_0) \notin \mathrm{I\!R}$*, o conjunto* $\partial_b f(x_0)$ *é considerado vazio.*

Os operadores b-subdiferencial e o subdiferencial de Fenchel ∂ se relacionam do seguinte modo:

$$(x^*, h) \in \partial_b f(x_0) \Longleftrightarrow x^* \in \partial\left(f - h\right)(x_0), \qquad x^* \in \partial f\left(x_0\right) \Longleftrightarrow (x^*, 0) \in \partial_b f(x_0). \tag{2.15}$$

Essas relações seguem diretamente das definições de tais operadores. A seguir, apresentamos um exemplo simples em que o subdiferencial de Fenchel é vazio mas o $b-$subdiferencial não.

Exemplo 2.19 *Consideramos* $f : \mathrm{I\!R} \to \mathrm{I\!R}$ *definida por* $f(x) = x^3$*. Notamos que como* f *não é convexa então não há garantia de que o* ∂f *é não-vazio. Vamos, de fato, verificar que* $\partial f(0)$ *é vazio. Para isso, vamos mostrar que para todo* $x^* \in \mathrm{I\!R}$*, existe* $x \in \mathrm{I\!R}$ *tal que*

$$f(x) < f(0) + xx^* = xx^*. \tag{2.16}$$

Primeiramente, se $x^* > 0$*, temos que*

$$\begin{aligned} x < -\sqrt{x^*} < 0 \quad &\Rightarrow \quad x(x^* - x^2) > 0 \\ &\Rightarrow \quad xx^* > x^3, \end{aligned}$$

portanto, (2.16) *é válida para todo* $x < -\sqrt{x^*}$*.*

Agora para $x^* \leq 0$*, temos que* (2.16) *vale para todo* $x < 0$*, pois*

$$0 > x \Rightarrow x(x^* - x^2) > 0 \Rightarrow xx^* > x^3.$$

Concluímos assim que $\partial f(0) = \emptyset$*. No entanto,* $\partial_b f(0) \neq \emptyset$*, pois claramente* $(0, h) \in \partial_b f(0)$ *para qualquer função contínua* h *diferenciável em 0, com* $h(0) = 0$ *e que seja majorada por* f*. Em particular,* $h = f$*.*

Definição 2.20 *A função* b'*-conjugada* $\varphi^{b'} : X \to \overline{\mathrm{I\!R}}$ *de* $\varphi : X^* \times \mathcal{H}_X \to \overline{\mathrm{I\!R}}$ *é definida como*

$$\varphi^{b'}(x) = \sup_{(x^*, h) \in X^* \times \mathcal{H}_X} \left\{ \langle x, x^* \rangle + h(x) - \varphi\left(x^*, h\right) \right\}. \tag{2.17}$$

A função b'-conjugada se relaciona com a conjugação de Fenchel de acordo com a seguinte fórmula, que segue de (2.17):

$$\varphi^{b'}(x) = \sup_{h \in \mathcal{H}_X} \left\{ h(x) + \varphi(\cdot, h)^*(x) \right\}. \tag{2.18}$$

A segunda b-conjugada de f é naturalmente $f^{bb'} := (f^b)^{b'}$. Combinando (2.18) com (2.13), concluímos que

$$f^{bb'}(x) = \sup_{h \in \mathcal{H}_X} \left\{ h(x) + (f - h)^{**}(x) \right\}. \tag{2.19}$$

Teorema 2.21 *Seja* $f : X \to \mathbb{R} \cup \{+\infty\}$. *Então*

$$f^{bb'}(x) = \sup\{g(x) \mid g : X \to \mathbb{R} \text{ é contínua em } X \text{ e Fréchet diferenciável em } 0,\ g \leq f\}.$$

Demonstração. Pelo Teorema 1.7, temos que

$$f^{bb'}(x) = \sup_{g \in \Phi_b} \{g(x) |\ g \leq f\}$$

para todo $x \in X$. Lembramos que toda função em Φ_b é da forma $\langle \cdot, x^* \rangle + h(\cdot)$, em que $x^* \in X^*$ e $h \in \mathcal{H}_X$, ou seja, toda função em Φ_b é contínua em X e Fréchet diferenciável em 0. Por outro lado, toda função g contínua em X e Fréchet diferenciável em 0 é b-elementar. De fato, basta considerar $x^* = \nabla g(0)$ e $h = g - \langle \cdot, x^* \rangle$ e assim $g(z) = \langle z, x^* \rangle + h(z)$ para todo $z \in X$. Deste modo, Φ_b é exatamente o conjunto das funções contínuas em X e Fréchet diferenciáveis em 0. □

Consequentemente, temos o seguinte corolário que afirma que as funções b-supremos são exatamente as funções sci.

Corolário 2.22 *Seja* $f : X \to \mathbb{R} \cup \{+\infty\}$ *e* $x \in X$. *Então* f *é sci em* x *se, e somente se,* $f^{bb'}(x) = f(x)$. *Além do mais, se* $x_0 \in dom\ f \backslash \{0\}$ *e* f *é sci em* x_0, *então* $\partial_b f(x_0) \neq \emptyset$.

Demonstração. Pelo Teorema 2.13, f é sci em x se, e somente se,

$$f(x) = \sup\{h(x) \mid h : X \to \mathbb{R} \text{ é contínua em } X \text{ e Fréchet diferenciável em } 0,\ \nabla h(0) = 0,\ h \leq f\}. \tag{2.20}$$

Pelo Teorema 2.21 e por propriedade de supremo referente à inclusão de conjuntos, concluímos que $f(x) \leq f^{bb'}(x)$. A desigualdade contrária segue da Proposição 1.6.

Por outro lado, se $f(x) = f^{bb'}(x)$ então f é sci em x por ser supremo de funções contínuas neste ponto.

Pelo Teorema 2.13, temos também que o supremo em (2.20) é atingido quando $x_0 \in dom\ f \backslash \{0\}$, o que é equivalente a dizer, via Corolário 1.10, que f é b-subdiferenciável em x_0. □

Notamos que, devido ao Corolário 2.22 e graças ao Teorema A.13, para uma função f convexa, própria, e sci, o supremo em (2.19) é sempre atingido para $h = 0$ pois

$$f^{**} = f = f^{bb'}.$$

Ou seja, para este conjunto de funções, $f^{bb'}$ e f^{**} coincidem.

Como podemos ver pelo Corolário 2.22, garantimos que uma função sci é b-subdiferenciável em todos os pontos do seu domínio, podendo não ser na origem. Como mencionamos anteriormente, para este caso uma função sci é b-subdiferenciável se, e somente se, existe uma função contínua Fréchet diferenciável em 0 que a minora e que assume o mesmo valor em 0. Este fato não é verdade para toda função sci como veremos no próximo exemplo.

Exemplo 2.23 *Seja* $f : \mathbb{R} \to \mathbb{R}$ *tal que*

$$f(x) = \begin{cases} x sen\frac{1}{x}, & \text{se } x \neq 0 \\ 0, & \text{caso contrário.} \end{cases}$$

Notamos inicialmente que f *é contínua em* 0*, pois*

$$-x \leq x sen\frac{1}{x} \leq x$$

para todo $x \neq 0$*, consequentemente* $\lim_{x\to 0} f(x) = 0$.

Observemos que se para algum $x^* \in \mathbb{R}$,

$$\liminf_{x\to 0} \frac{f(x) - f(0) - xx^*}{|x|} < 0 \tag{2.21}$$

então para qualquer função $g : \mathbb{R} \to \mathbb{R}$ *tal que*

(i) $g(0) = f(0)$;

(ii) $g \leq f$,

teremos que

$$\frac{g(x) - g(0) - xx^*}{|x|} \leq \frac{f(x) - f(0) - xx^*}{|x|}$$

para todo x *não nulo, e logo* $\liminf_{x\to 0} \frac{g(x) - g(0) - xx^*}{|x|} < 0$*, ou seja,* g *não é Fréchet diferenciável. Deste modo, nos resta mostrar* (2.21)*, como fazemos a seguir.*

Dado $\delta > 0$*, temos que*

$$\begin{aligned}
\inf_{x\in(0,\delta)} \frac{f(x)-f(0)-xx^*}{|x|} &= \inf_{x\in(0,\delta)} \frac{x sen\frac{1}{x} - xx^*}{x} \\
&= \inf_{x\in(0,\delta)} sen\frac{1}{x} - x^* \\
&= -1 - x^*.
\end{aligned}$$

Além disso,

$$\begin{aligned}
\inf_{x\in(-\delta,0)} \frac{f(x)-f(0)-xx^*}{|x|} &= \inf_{x\in(-\delta,0)} \left(x^* - sen\frac{1}{x}\right) \\
&= x^* - \sup_{x\in(-\delta,0)} sen\frac{1}{x} \\
&= x^* - 1.
\end{aligned}$$

Sendo assim,

$$\inf_{x\in(-\delta,\delta)\setminus\{0\}} \frac{f(x)-f(0)-xx^*}{|x|} = \min\{-1-x^*, -1+x^*\}$$

para todo $\delta > 0$. *Como* $\frac{f(x)-f(0)-xx^*}{|x|}$ *não está definida para* $x = 0$, *concluímos que*

$$\liminf_{x\to 0} \frac{f(x)-f(0)-xx^*}{|x|} = \min\{-1-x^*, -1+x^*\}.$$

Caso $x^* \geq 0$ *então* $-1-x^* \leq -1+x^*$ *e* $-1-x^* < 0$. *Já se* $x^* < 0$ *então* $-1+x^* < -1-x^*$ *e* $-1+x^* < 0$. *Ambos os casos implicam* (2.21).

Portanto $\partial_b f(0) = \emptyset$ *mesmo com* f *sendo sci em* 0.

Aqui mostramos um exemplo de uma função que é b-supremo mas não é b-subdiferenciável, evidenciando que estes dois conceitos não são necessariamente equivalentes. Como podemos notar, a condição (2.21) foi fundamental para garantir que a função não fosse b-subdiferenciável em 0. Portanto as funções que não cumprem (2.21) são as candidatas a serem b-subdiferenciáveis. Veremos que elas coincidem com um conjunto de funções conhecido na literatura como Fréchet subdiferenciável.

Definição 2.24 *[33] Dizemos que* $x^* \in X^*$ *é Fréchet (ou regular, seguindo a terminologia de [33]) subgradiente de* f *em* $x_0 \in dom\ f$ *quando*

$$\liminf_{x\to x_0} \frac{f(x) - f(x_0) - \langle x - x_0, x^*\rangle}{\|x - x_0\|} \geq 0.$$

O conjunto de todos os Fréchet subgradientes é chamado de Fréchet subdiferencial de f *em* x_0. *Este conjunto será denotado por* $\partial_F f(x_0)$.

Uma função f é dita *Fréchet subdiferenciável* em um ponto x_0 do seu domínio quando $\partial_F f(x_0) \neq \emptyset$. Notemos que qualquer função Fréchet diferenciável ou Fenchel subdiferenciável são Fréchet subdiferenciáveis. Vemos no próximo teorema que a caracterização que estamos buscando para garantir b-subdiferenciabilidade na origem é exatamente a Fréchet subdiferenciabilidade na origem.

Teorema 2.25 *[33, Proposição 8.5] Sendo $f : X \to \mathbb{R} \cup \{+\infty\}$, x^* é Fréchet subgradiente de f em 0 se, e somente se, existe uma função contínua $g : X \to \mathbb{R} \cup \{+\infty\}$ Fréchet diferenciável em 0 tal que*

(i) $g \leq f$;

(ii) $g(0) = f(0)$;

(iii) $\nabla g(0) = x^*$.

Demonstração. Suponha que exista g satisfazendo as hipóteses do Teorema, então temos que

$$\begin{aligned} \liminf_{x \to 0} \frac{f(x) - f(0) - \langle x, x^* \rangle}{\|x\|} &\geq \\ \liminf_{x \to 0} \frac{g(x) - g(0) - \langle x, x^* \rangle}{\|x\|} &= \\ \lim_{x \to 0} \frac{g(x) - g(0) - \langle x, x^* \rangle}{\|x\|} &= 0. \end{aligned}$$

Portanto $x^* \in \partial_F f(0)$.

Para ver a demonstração da recíproca, consulte [33, Proposição 8.5] e [41, Proposição 3.1]. □

Corolário 2.26 *Seja $f : X \to \mathbb{R} \cup \{+\infty\}$ tal que $0 \in dom\ f$. Então a projeção de $\partial_b f(0)$ em X^* coincide com $\partial_F f(0)$.*

Demonstração. Se x^* pertence à projeção de $\partial_b f(0)$ em X^*, então existe $h \in \mathcal{H}_X$ tal que (2.14) vale para $x_0 = 0$. Considerando a função $g := x^* + h + f(0)$, concluímos que g cumpre as hipóteses do Teorema 2.25, com $\nabla g(0) = x^*$. Portanto $x^* \in \partial_F f(0)$.

Por outro lado, se $x^* \in \partial_F f(0)$ então, novamente pelo Teorema 2.25, existe g Fréchet diferenciável em 0 tal que $g \leq f$, $g(0) = f(0)$ e $\nabla g(0) = x^*$. Logo, considerando $h = -x^* + g$, concluímos facilmente que $(x^*, h) \in \partial_b f(0)$. □

Finalmente, temos o nosso resultado de b-subdiferenciabilidade na origem.

Corolário 2.27 *Seja $f : X \to \mathbb{R} \cup \{+\infty\}$ tal que $0 \in dom\ f$. Então $\partial_b f(0) \neq \emptyset$ se, e somente se, $\partial_F f(0) \neq \emptyset$.*

Para encerrar a subseção, vamos aplicar este esquema para o desenvolvimento de uma teoria de dualidade. Estes resultados seguem diretamente do que foi apresentado na Seção 1.2.

Seja S um conjunto arbitrário e $\phi : S \times X \to \overline{\mathbb{R}}$ a função objetivo. O problema de otimização a ser considerado é

$$(\mathcal{P}) \qquad \text{minimizar } \phi\left(s, 0\right); \tag{2.22}$$

Considerando $p : X \to \overline{\mathbb{R}}$, $p(x) = \inf_{s \in S} \phi(s, x)$, a função de perturbação associada, vamos definir o problema Dual de $(\mathcal{P})$ como

$$(\mathcal{D}_b) \qquad \text{maximizar } -p^b\left(x^*, h\right),$$

em que

$$-p^b\left(x^*, h\right) = \inf_{x \in X} \left\{ \inf_{s \in S} \phi\left(s, x\right) - \langle x, x^* \rangle - h(x) \right\} = \inf_{x \in X,\ s \in S} \left\{ \phi\left(s, x\right) - \langle x, x^* \rangle - h(x) \right\}.$$

Como a teoria desenvolvida aqui é um caso particular de c-conjugada, todos os resultados apresentados no Capítulo 1 são válidos neste contexto. Do item (ii) do Teorema 1.12 e do Corolário 2.22, temos o resultado de Dualidade forte.

Teorema 2.28 *Os valores ótimos de* $(\mathcal{P})$ *e* $(\mathcal{D}_b)$ *coincidem se, e somente se,* p *é sci em* 0.

Comparando os Teoremas 2.13 e 2.21, notamos que a condição de que o gradiente das funções consideradas no supremo (2.11) se anula na origem não é levada em conta na teoria de b-conjugação. Isso acontece porque essa condição não afeta o valor do supremo, porém ela reduz o conjunto das funções consideradas, podendo simplificar o seu cálculo. A partir dessa motivação, vamos definir na próxima subseção um segundo esquema de conjugação que inclui tal condição.

2.3.2 O segundo esquema de conjugação

Consideremos $d : X \times \mathcal{H}_X \to \mathbb{R}$ dada por

$$d\left(x, h\right) := h(x)$$

e $d' : \mathcal{H}_X \times X \to \mathbb{R}$, definida por $d'\left(h, x\right) := d\left(x, h\right)$. Assim definimos a d-conjugação.

Definição 2.29 *A função d-conjugada de* $f : X \to \overline{\mathbb{R}}$ *é* $f^d : \mathcal{H}_X \to \overline{\mathbb{R}}$, *definida por*

$$f^d(h) = \sup_{x \in X} \left\{ h(x) - f(x) \right\}.$$

A função d'-conjugada $\varphi^{d'} : X \to \overline{\mathbb{R}}$ de uma função $\varphi : \mathcal{H}_X \to \overline{\mathbb{R}}$ é definida por

$$\varphi^{d'}(x) = \sup_{h \in \mathcal{H}_X} \{h(x) - \varphi(h)\}.$$

Notamos que essa definição é mais simples que a da b-conjugada, no entanto, não conseguimos enxergar uma relação direta dessa definição com a conjugação de Fenchel. Este fato é um ponto negativo quando estamos trabalhando com funções convexas, pois esta propriedade não nos acrescenta nenhuma informação a respeito de sua d-conjugação.

Definição 2.30 *Dizemos que $f : X \to \overline{\mathbb{R}}$ é d-subdiferenciável em $x_0 \in dom(f)$ quando existe $h \in \mathcal{H}_X$ tal que*

$$f(x) - f(x_0) \geq h(x) - h(x_0) \text{ para todo } x \in X. \tag{2.23}$$

h é dita d-subgradiente de f em x_0. O conjunto de todos d-subgradientes de f em x_0, denotado por $\partial_d f(x_0)$, é chamado de d-subdiferencial de f em x_0. Se $f(x_0) \notin \mathbb{R}$, consideraremos $\partial_d f(x_0) := \emptyset$.

Naturalmente, a biconjugada de f é $f^{dd'} := (f^d)^{d'}$. Análogo ao Teorema 2.21, temos o seguinte resultado.

Teorema 2.31 *Seja $f : X \to \mathbb{R} \cup \{+\infty\}$. Então*

$$f^{dd'}(x) = \sup\{g(x) \mid g : X \to \mathbb{R} \text{ é contínua em } X \text{ e Fréchet diferenciável em } 0,\ \nabla g(0) = 0,\ g \leq f\}$$

para todo $x \in X$.

Notemos que este teorema é consequência direta da definição do conjunto $\mathcal{H}_X$. Como desejado, a condição $\nabla g(0) = 0$ está sendo considerada. Assim, diretamente do Teorema 2.13, concluímos o seguinte corolário.

Corolário 2.32 *Seja $f : X \to \mathbb{R} \cup \{+\infty\}$ e $x \in X$. Então f é sci em x se, e somente se, $f^{dd'}(x) = f(x)$. Além do mais, se $x_0 \in dom\ f \backslash \{0\}$ e f é sci em x_0, então $\partial_d f(x_0) \neq \emptyset$.*

Notamos que da Definição 2.30, uma função f é $d-$subdiferenciável na origem se, e somente se, existe uma função contínua diferenciável em 0, cujo gradiente se anula em 0, que seja majorada por ela e assuma o mesmo valor em 0. Diretamente do Teorema 2.25, concluímos exatamente quem são tais funções.

Corolário 2.33 *Seja $f : X \to \mathbb{R} \cup \{+\infty\}$ tal que $0 \in dom\ f$. Então $\partial_d f(0) \neq \emptyset$ se, e somente se, $0 \in \partial_F f(0)$.*

Demonstração. Se $\partial_d f(0) \neq \emptyset$, temos que existe uma função $h \in \mathcal{H}_X$ tal que $h + f(0)$ é majorada por f e assume o mesmo valor que f na origem. Como $\nabla(h + f(0))(0) = 0$, concluímos pelo Teorema 2.25 que $0 \in \partial_F f\,(0)$.

Por outro lado, se $0 \in \partial_F f\,(0)$, novamente pelo Teorema 2.25 existe uma função g Fréchet diferenciável em 0, com $\nabla g(0) = 0$ tal que $g \leq f$ e $g(0) = f(0)$. Considerando $h = g - g(0)$, podemos ver que $h \in \mathcal{H}_X$ e satisfaz (2.23). □

Para finalizar a subseção, assim como feito para b-conjugação, aplicamos a d-conjugação para desenvolver uma teoria de dualidade. Deste modo, considerando o Problema Primal como em (2.22), o problema Dual relacionado a d-conjugação é dado agora por

$$(\mathcal{D}_d) \qquad \text{maximizar } -p^d\,(h)$$

em que

$$-p^d\,(h) = \inf_{x \in X,\ s \in S} \{\phi\,(s, x) - h(x)\}\,.$$

Usando o Corolário 2.32 e o item (ii) do Teorema 1.12, obtemos uma versão similar ao Teorema 2.28.

Teorema 2.34 *Os valores ótimos de $(\mathcal{P})$ e $(\mathcal{D}_d)$ coincidem se, e somente se, p é sci em* 0.

Os resultados sobre Dualidade apresentados no Capítulo 1 são obviamente válidos para essa conjugação. Para encerrar o capítulo, vamos apresentar exemplos de aplicações de ambas teorias.

2.3.3 Exemplos

Nesta subseção apresentamos dois exemplos para as dualidades desenvolvidas anteriormente. Ambos os exemplos são para fins ilustrativos, ou seja, somente para mostrar como se aplicam as teorias de dualidade propostas.

O primeiro é um Problema Quadrático com restrições lineares. Vamos verificar que, para este caso, a função perturbação é sci em 0 e portanto $(\mathcal{P})$, $(\mathcal{D}_b)$ e $(\mathcal{D}_d)$ coincidem.

O segundo exemplo foi proposto por Duffin [13] a fim de mostrar que na Dualidade Lagrangiana [22] pode haver gap de dualidade quando a condição de Slater [22, pág. 200] não for satisfeita. Veremos que para ambas as dualidades apresentadas, também teremos gap de dualidade, ou seja, a função perturbação não é sci em 0.

Exemplo 2.35 *Consideremos o seguinte problema*

$$\begin{aligned} &\textit{minimizar} \quad -s_1^2 + s_2^2 + s_1 s_2 \\ &\textit{sujeito a} \quad s_1 + s_2 \leq 1 \\ &\qquad\qquad\quad s_1 - s_2 \leq 0 \\ &\qquad\qquad\quad s_1 \geq 0,\ \ s_2 \geq 0. \end{aligned}$$

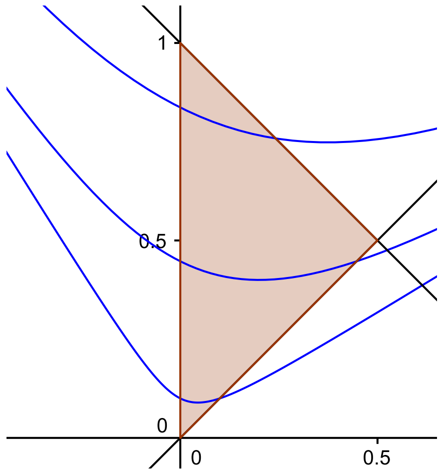

Figura 2.2: Ilustração da região viável com as curvas de nível da função objetivo do problema.

Notamos que a hessiana da função objetivo é dada por $\begin{bmatrix} -2 & 1 \\ 1 & 2 \end{bmatrix}$, *ou seja, é indefinida. Vamos introduzir a função objetivo* $\phi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R} \cup \{+\infty\}$ *da maneira usual [Exemplo 1.11].*

$$\phi(s_1, s_2, x_1, x_2) := \begin{cases} -s_1^2 + s_2^2 + s_1 s_2, & \textit{se } (s_1, s_2, x_1, x_2) \in \Omega, \\ +\infty, & \textit{caso contrário,} \end{cases}$$

em que

$$\Omega := \left\{(s_1, s_2, x_1, x_2) \in \mathbb{R}^4_+ : s_1 + s_2 + x_1 \leq 1,\ s_1 - s_2 + x_2 \leq 0,\ x_1 + x_2 \leq 1\right\}.$$

Notemos que $x_1 + x_2 \leq 1$ *é uma condição necessária para que existam* $(s_1, s_2) \in \mathbb{R}_2^+$ *tais que* $s_1 + s_2 + x_1 \leq 1$ *e* $s_1 - s_2 + x_2 \leq 0$.

Assim, é fácil ver que

$$(\mathcal{P}) \qquad \textit{minimizar } \phi\,(s, 0)$$

é equivalente ao problema dado. Como, para $(s_1, s_2, x_1, x_2) \in \Omega$, *temos*

$$\begin{aligned} \phi(s_1, s_2, x_1, x_2) &= -s_1^2 + s_2^2 + s_1 s_2 \geq -s_1^2 + (s_1 + x_2)^2 + s_1(s_1 + x_2) \\ &= s_1^2 + 3x_2 s_1 + x_2^2 \geq x_2^2, \end{aligned}$$

então a função perturbação p *associada satisfaz* $p(x_1, x_2) \geq x_2^2$ *para todo* $(x_1, x_2) \in \mathbb{R}^2$ *tal que* $x_1 \geq 0$, $x_2 \geq 0$ *e* $x_1 + x_2 \leq 1$. *Além do mais, desde que* $\phi(0, x_2, x_1, x_2) = x_2^2$, *concluímos que*

$$p(x_1, x_2) = \begin{cases} x_2^2, & \textit{se } x_1 \geq 0,\ x_2 \geq 0,\ x_1 + x_2 \leq 1, \\ +\infty, & \textit{caso contrário.} \end{cases}$$

Dessa forma, $p(0) = 0$ e, como p é sci em 0, concluímos que não há gap de dualidade para ambos problemas duais $(\mathcal{D}_b)$ e $(\mathcal{D}_d)$. De fato, no caso de $(\mathcal{D}_b)$, o par $(0, h_0)$, com $h_0(x_1, x_2) := x_2^2$, é uma solução ótima, pois $-p^b(0, h_0) = \inf_{(s,x)\in\Omega} \{\phi(s_1, s_2, x_1, x_2) - x_2^2\} = 0$ (o ínfimo é atingido, por exemplo, em $(s_1, s_2, x_1, x_2) = (0, 0, 0, 0)$), enquanto que no caso de $(\mathcal{D}_d)$, a mesma função h_0 é uma solução ótima, pois temos que $-p^d(h_0) = -p^b(0, h_0) = 0$.

Exemplo 2.36 *[13, Seção 3] Consideremos o problema*

$$\begin{array}{rl} \textit{minimizar} & e^{s_2} \\ \textit{sujeito a} & \sqrt{s_1^2 + s_2^2} \leq s_1. \end{array}$$

Notemos que e^{s_2} é convexa no entanto a condição de Slater não é satisfeita para as restrições do problema, pois não existe $(s_1, s_2) \in \mathbb{R}^2$ tais que $\sqrt{s_1^2 + s_2^2} < s_1$.

Seja a função objetivo $\phi : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ definida por

$$\phi(s_1, s_2, x) := \begin{cases} e^{s_2}, & \textit{se } \sqrt{s_1^2 + s_2^2} \leq s_1 - x, \\ +\infty, & \textit{caso contrário.} \end{cases}$$

Notemos que

$$(\mathcal{P}) \qquad \textit{minimizar } \phi(s, 0)$$

é equivalente ao problema dado.

Para cada $x \in \mathbb{R}$, vamos analisar o conjunto

$$\Omega_x = \{(s_1, s_2) \in \mathbb{R}^2 \mid \sqrt{s_1^2 + s_2^2} \leq s_1 - x\}.$$

Inicialmente se $s_1 - x < 0$ então não existe $(s_1, s_2) \in \mathbb{R}^2$ tal que $\sqrt{s_1^2 + s_2^2} \leq s_1 - x$. Logo, vamos supor que $s_1 - x \geq 0$. Deste modo, temos a seguinte equivalência:

$$\begin{aligned} \sqrt{s_1^2 + s_2^2} \leq s_1 - x \quad & \Leftrightarrow \quad s_1^2 + s_2^2 \leq s_1^2 - 2s_1x + x^2 \\ & \Leftrightarrow \quad s_2^2 \leq -2s_1x + x^2. \end{aligned}$$

Para $x = 0$, temos que $\Omega_x = \{(s_1, s_2) \in \mathbb{R}^2 \mid s_1 \geq 0 \text{ e } s_2 = 0\}$. Para o caso em que $x > 0$, desde que $s_1 \geq x > \frac{x}{2}$ então $-2s_1x + x^2 < 0$ e portanto $\Omega_x = \emptyset$. Finalmente, para $x < 0$, temos que para todo s_2, existe $s_1 \geq \frac{x}{2}$ tal que $s_2^2 \leq -2s_1x + x^2$. Logo, Ω_x é ilimitado na segunda variável.

Deste modo, desde que a função perturbação é definida como $p(x) := \inf_{s\in\mathbb{R}_+^2} \phi(s, x) =$

$\inf_{s\in\Omega_x} \phi(s,x)$, *concluímos que*

$$p(x) = \begin{cases} 1, & se\ x = 0 \\ 0, & se\ x < 0 \\ +\infty, & se\ x > 0. \end{cases}$$

Dessa forma $p(0) = 1$ *e como* p *não é sci em* 0, *há um gap de dualidade estritamente positivo para ambos problemas duais* $(\mathcal{D}_b)$ *e* $(\mathcal{D}_d)$. *Vamos verificar que os valores ótimos dos problemas* $(\mathcal{D}_b)$ *e* $(\mathcal{D}_d)$ *são ambos* 0.

Primeiro afirmamos que, para qualquer função contínua $\alpha : \mathbb{R} \to \mathbb{R}$ *com* $\alpha(0) = 0$, *temos que*

$$\inf_{x\in\mathbb{R}} \{p(x) - \alpha(x)\} \leq 0.$$

De fato, como α *é contínua em* 0 *então, para todo* $\epsilon > 0$, *existe* $\delta > 0$ *tal que*

$$|x| < \delta \Rightarrow -\alpha(x) < \epsilon.$$

Deste modo, para $-\delta < y < 0$, *temos que* $\inf_{x\in\mathbb{R}}\{p(x) - \alpha(x)\} \leq p(y) - \alpha(y) < \epsilon$ *como desejado.*

Sendo assim, como $(x^* + h)$ *e* h, *com* $x^* \in \mathbb{R}^* = \mathbb{R}$ *e* $h \in \mathcal{H}_{\mathbb{R}}$, *são ambas contínuas, concluímos que* $-p^b(x^*, h) \leq 0$ *e* $-p^d(h) \leq 0$ *para todo* $x^* \in \mathbb{R}$ *e* $h \in \mathcal{H}_{\mathbb{R}}$. *Portanto,* $\max_{(x^*,h)\in\mathbb{R}\times\mathcal{H}_{\mathbb{R}}}\{-p^b(x^*, h)\} \leq 0$ *e* $\max_{h\in\mathcal{H}_{\mathbb{R}}}\{-p^d(h)\} \leq 0$. *Considerando a função identicamente nula, concluímos que os valores ótimos de* $(\mathcal{D}_b)$ *e* $(\mathcal{D}_d)$ *são ambos* 0. *E portanto há gap de dualidade que equivale a 1.*

Deste modo, cumprimos com os nossos objetivos. Sabemos que ainda há muito a se fazer neste contexto, como por exemplo, investigar mais propriedades a respeito das teorias de subdiferenciabilidade introduzidas e possíveis aplicações do Teorema 2.13. No entanto, resolvemos fazer uma pausa neste assunto porque um outro problema chamou nosso interesse. O próximo capítulo é dedicado aos resultados encontrados a respeito deste novo problema.

Capítulo 3

Uma generalização da Desigualdade Forte de Fitzpatrick

Neste capítulo, vamos apresentar um tópico independente do que foi apresentado nos capítulos anteriores. Vamos revisar brevemente a teoria de Operadores Monótonos e apresentaremos a nossa proposta de generalização para a Desigualdade Forte de Fitzpatrick [40, Teorema 4] através de um conjunto de funções conhecidas como *Twisted Bigger Conjugate* (TBC) [35]. Para finalizar, usamos esta generalização para definir uma família de funções *gap*.

Em todo este capítulo, X será considerado um espaço de Banach reflexivo. Seguiremos denotando $\langle \cdot, \cdot \rangle$ como o produto de dualidade de X^* com X. O produto de dualidade entre os espaços $X \times X^*$ e seu dual $(X \times X^*)^* = X^* \times X$ será definido naturalmente por

$$\langle (x, x^*), (y^*, y) \rangle = \langle x, y^* \rangle + \langle y, x^* \rangle$$

para $(x, x^*) \in X \times X^*$ e $(y^*, y) \in X^* \times X$. Para mais informações a respeito de espaços reflexivos, ver [9, Seção 3.5].

3.1 Operadores Monótonos Maximais

Nesta seção vamos apresentar resultados sobre Operadores Monótonos Maximais. Tais operadores são interessantes no contexto de Otimização pois englobam o operador subdiferencial. Os resultados a seguir são baseados em [35].

Recordamos ao leitor que para um operador ponto-conjunto $T : X \rightrightarrows X^*$, o seu gráfico e o seu domínio são dados, respectivamente, por

$$gra\ T := \{(x, x^*) \in X \times X^* \mid x^* \in T(x)\}$$

e

$$dom\ T := \{x \in X \mid T(x) \neq \emptyset\}.$$

Definição 3.1 *Um operador ponto-conjunto $T : X \rightrightarrows X^*$ é dito monótono quando*

$$\langle x - y, x^* - y^* \rangle \geq 0$$

para todo $(x, x^*), (y, y^*) \in gra\ T$.

Exemplo 3.2 *Seja $S : X \to X^*$ um operador ponto-ponto linear, com dom $S = X$, tal que*

$$\langle x, Sx \rangle \geq 0$$

para todo $x \in X$. Então facilmente percebemos que S é monótono. Em particular, quando $X = H$, um espaço de Hilbert, o operador Identidade é monótono.

Exemplo 3.3 *Seja $f : X \to \mathbb{R} \cup \{+\infty\}$ uma função. Então $\partial f : X \rightrightarrows X^*$ é um operador monótono. De fato, tomando $(x, x^*), (y, y^*) \in gra\ \partial f$, temos que*

$$\begin{aligned} f(y) &\geq f(x) + \langle y - x, x^* \rangle \\ f(x) &\geq f(y) + \langle x - y, y^* \rangle . \end{aligned}$$

Logo

$$\begin{aligned} \langle x - y, x^* - y^* \rangle &= \langle x - y, x^* \rangle + \langle x - y, -y^* \rangle \\ &\geq f(x) - f(y) - f(x) + f(y) = 0. \end{aligned}$$

Caso gra $\partial f = \emptyset$ então, por vacuidade, o operador é monótono, concluindo assim o desejado.

Definição 3.4 *Um operador monótono $T : X \rightrightarrows X^*$ é dito maximal quando a seguinte condição vale: se $\langle x - y, x^* - y^* \rangle \geq 0$ para todo $(x, x^*) \in gra\ T$ então $(y, y^*) \in gra\ T$.*

É possível provar que a condição de maximalidade é equivalente a dizer que não existe outro operador monótono cujo gráfico contém propriamente o gráfico de T. A seguir, vamos ver exemplos de operadores monótonos maximais.

Exemplo 3.5 *Seja S como no Exemplo 3.2. Com o objetivo de verificar que S é maximal, consideremos $(y, y^*) \in X \times X^*$ tal que*

$$\langle x - y, x^* - y^* \rangle \geq 0$$

vale para todo $(x, x^*) \in gra\ S$. Em particular, dados $z \in X$ e $\lambda \in \mathbb{R}$, tomamos $x = y + \lambda z$, e assim obtemos que $\langle \lambda z, Sy + \lambda Sz - y^* \rangle \geq 0$, implicando $\lambda \langle z, Sy - y^* \rangle + \lambda^2 \langle z, Sz \rangle \geq 0$. Como λ é arbitrário, concluímos que $\langle z, Sy - y^* \rangle = - \langle z, Sz \rangle \leq 0$. Usando agora o fato de que z também é arbitrário, concluímos que $Sy = y^*$ como desejado.

Exemplo 3.6 *Seja $f : X \to \mathbb{R} \cup \{+\infty\}$ uma função própria, convexa e sci. Então ∂f é um operador monótono maximal. A demonstração da maximalidade não é trivial e pode ser encontrada no Teorema 18.7 de [35].*

Teorema 3.7 *Seja $T : X \rightrightarrows X^*$ um operador monótono maximal, então seu gráfico é um conjunto fechado em $X \times X^*$.*

Demonstração. Consideremos uma sequência $(x_k, x_k^*) \subset gra\ T$ que converge para (x, x^*) e vamos provar que $(x, x^*) \in gra\ T$. De fato, tomando $(y, y^*) \in gra\ T$, temos pela monotonicidade de T que $\langle x_k - y, x_k^* - y^* \rangle \geq 0$ para todo k. Passando o limite, concluímos que $\langle x - y, x^* - y^* \rangle \geq 0$. Como (y, y^*) é arbitrário em $gra\ T$, concluímos, pela maximalidade de T, que $(x, x^*) \in gra\ T$. □

Vamos associar agora os operadores à classe de Funções de Fitzpatrick. Tais funções foram introduzidas em [18] e são conhecidas como funções representativas de operadores monótonos maximais, pois elas atuam exatamente como uma representação do operador no contexto de funções convexas. Veremos mais adiante que podemos analisar propriedades dos operadores através destas funções.

Definição 3.8 *A função de Fitzpatrick associada com o operador monótono maximal $T : X \rightrightarrows X^*$ é a função $F_T : X \times X^* \to \mathbb{R} \cup \{+\infty\}$ definida por*

$$F_T(x, x^*) := \sup_{(y,y^*) \in gra\ T} \{\langle x - y, y^* \rangle + \langle y, x^* \rangle\}.$$

Equivalentemente,

$$F_T(x, x^*) = \langle x, x^* \rangle - \inf_{(y,y^*) \in gra\ T} \langle x - y, x^* - y^* \rangle . \tag{3.1}$$

Como o supremo de funções convexas e contínuas é convexo e sci então, diretamente da definição, observamos que F_T é uma função convexa e sci.

Exemplo 3.9 *Seja H um espaço de Hilbert, consideramos o operador identidade $I : H \to H^*$. Graças ao Teorema de Representação de Riesz-Fréchet [9, Teorema 5.5], todo funcional linear contínuo (todo elemento do dual) de um Espaço de Hilbert pode ser representado como $\langle \cdot, x^* \rangle$ para algum $x^* \in H$ em que $\langle \cdot, \cdot \rangle$ denota **neste caso** o produto interno de H. Neste contexto, recordamos que $I(x) := \langle \cdot, x \rangle$ e podemos considerar $F_I : H \times H \to \mathbb{R} \cup \{+\infty\}$.*

Calculando a função de Fitzpatrick do operador identidade, temos que

$$F_I(x, x^*) = \sup_{y \in H} \{\langle x - y, y \rangle + \langle y, x^* \rangle\} = \sup_{y \in H} \{-\langle y, y \rangle + \langle y, x^* + x \rangle\}.$$

Como estamos maximizando irrestritamente uma função côncava, concluímos que *seu maximizador é* $\overline{y} = \frac{x+x^*}{2}$ *e portanto,*

$$F_I(x, x^*) = \frac{\|x + x^*\|_H^2}{4}.$$

Vamos apresentar em seguida o primeiro resultado que relaciona as funções de Fitzpatrick com o seu operador correspondente.

Proposição 3.10 *Se um operador* $T : X \rightrightarrows X^*$ *é monótono maximal, então sua função de Fitzpatrick satisfaz*

$$F_T(x, x^*) \geq \langle x, x^* \rangle \text{ para todo } (x, x^*) \in X \times X^*, \tag{3.2}$$

com igualdade se, e só se, $(x, x^*) \in gra\ T$.

Demonstração. Primeiro tomamos $(x, x^*) \notin gra\ T$ e da maximalidade de T, existe $(y, y^*) \in gra\ T$ tal que $\langle x - y, x^* - y^* \rangle < 0$ e portanto $\inf_{(y,y^*) \in gra\ T} \langle x - y, x^* - y^* \rangle < 0$. Logo, de (3.1), concluímos que $F_T(x, x^*) > \langle x, x^* \rangle$.

O próximo passo é verificar que

$$(x, x^*) \in gra\ T \Leftrightarrow \inf_{(y,y^*) \in gra\ T} \langle x - y, x^* - y^* \rangle = 0.$$

Vamos provar somente a recíproca, uma vez que para provar o outro sentido, desde que T é monótono, basta considerar $x = y$. Deste modo, se $\inf_{(y,y^*) \in gra\ T} \langle x - y, x^* - y^* \rangle = 0$ então $\langle x - y, x^* - y^* \rangle \geq 0$ para todo $(y, y^*) \in gra\ T$. Como T é maximal, concluímos que $(x, x^*) \in gra\ T$. Novamente, de (3.1), concluímos o desejado. □

O resultado anterior mostra que a Função de Fitzpatrick caracteriza completamente o gráfico do seu operador correspondente. Deste modo, para analisar o gráfico de um operador, podemos optar entre analisar o próprio operador ou então os pontos tais que $F_T(x, x^*) = \langle x, x^* \rangle$.

O próximo teorema é uma interessante propriedade das funções de Fitzpatrick e, além do mais, será de fundamental importância para a demonstração do Teorema 3.18.

Teorema 3.11 *[8, Exercício 9.23] Seja* $T : X \rightrightarrows X^*$ *um operador monótono maximal. Então*

$$F_T(x, x^*) - \langle x, x^* \rangle + F_T(w, w^*) - \langle w, w^* \rangle \geq -\frac{1}{2} \langle x - w, x^* - w^* \rangle$$

para todo $(x, x^*), (w, w^*) \in X \times X^*$.

Demonstração. Usando a convexidade da função de Fitzpatrick e de (3.10), obtemos

$$F_T(x, x^*) - \langle x, x^* \rangle + F_T(w, w^*) - \langle w, w^* \rangle =$$

$$
\begin{aligned}
&= 2\left(\frac{1}{2}F_T(x,x^*)+\frac{1}{2}F_T(w,w^*)\right)-\langle x,x^*\rangle-\langle w,w^*\rangle\\
&\geq 2F_T\left(\frac{1}{2}(x+w),\frac{1}{2}(x^*+w^*)\right)-\langle x,x^*\rangle-\langle w,w^*\rangle\\
&\geq 2\left\langle\frac{1}{2}(x+w),\frac{1}{2}(x^*+w^*)\right\rangle-\langle x,x^*\rangle-\langle w,w^*\rangle\\
&= -\frac{1}{2}\langle x-w,x^*-w^*\rangle.
\end{aligned}
$$

□

Encerramos a seção com o nosso objeto de estudo deste capítulo. A demonstração será omitida, pois este resultado é um caso particular do Teorema 3.18, o qual será demonstrado na próxima seção.

Teorema 3.12 (Desigualdade Forte de Fitzpatrick) *[8, Teorema 9.7.2] Seja $T : X \rightrightarrows X^*$ um operador monótono maximal. Então*

$$F_T(x,x^*)-\langle x,x^*\rangle \geq \frac{1}{4}\inf_{(w,w^*)\in gra\ T}\left\{\|w-x\|_X^2+\|w^*-x^*\|_{X^*}^2\right\}$$

para todo $(x,x^*)\in X\times X^*$.

Nosso objetivo é generalizar este teorema no sentido de encontrar outras funções além da função $g(x,x^*)=\inf_{(w,w^*)\in gra\ T}\left\{\|w-x\|_X^2+\|w^*-x^*\|_{X^*}^2\right\}$ em que a desigualdade anterior continue válida.

3.2 Generalização da Desigualdade de Fitzpatrick

Recordamos que, desde que $(X\times X^*)^*\approx X\times X^*$, a conjugada de Fenchel de uma função $g : X\times X^*\to \mathbb{R}\cup\{+\infty\}$ é dada naturalmente por

$$g^*(x^*,x)=\sup_{(y,y^*)\in X\times X^*}\{\langle (y,y^*),(x,x^*)\rangle-g(y,y^*)\}.$$

Assim, apresentamos a seguinte definição.

Definição 3.13 *[35, Definição 19.14] Dizemos que uma função sci, convexa, própria $g : X\times X^*\to \mathbb{R}\cup\{+\infty\}$ é Twisted Bigger Conjugate (TBC) quando*

$$-\langle x,x^*\rangle \leq g(x,x^*)\leq g^*(-x^*,-x) \tag{3.3}$$

para todo $(x,x^*)\in X\times X^*$.

Estas funções nos permitem generalizar a desigualdade de Fitzpatrick. Vamos ver alguns exemplos.

Exemplo 3.14 *A função* $g : X \times X^* \to \mathbb{R} \cup \{+\infty\}$ *definida por* $g(x, x^*) := f(x) + f^*(-x^*)$, *em que* $f : X \to \mathbb{R} \cup \{+\infty\}$ *é uma função sci, convexa e própria, é uma função TBC. De fato, do Teorema A.13 temos que* $f^{**} = f$. *Logo,*

$$\begin{aligned} g^*(-x^*, -x) &= \sup_{(y,y^*) \in X \times X^*} \{\langle y, -x^* \rangle + \langle -x, y^* \rangle - f(y) - f^*(-y^*)\} \\ &= \sup_{y \in X} \{\langle y, -x^* \rangle - f(y)\} + \sup_{y^* \in X^*} \{\langle x, -y^* \rangle - f^*(-y^*)\} \\ &= f^*(-x^*) + f^{**}(x) \\ &= f^*(-x^*) + f(x) = g(x, x^*), \end{aligned}$$

e pelo item (i) do Teorema A.10,

$$g(x, x^*) = f(x) + f^*(-x^*) \geq -\langle x, x^* \rangle .$$

Em particular, g *definida por* $g(x, x^*) = \frac{1}{p}\|x\|_X^p + \frac{1}{q}\|x^*\|_{X^*}^q$, *em que* $p \geq 1$ *e* $\frac{1}{p} + \frac{1}{q} = 1$, *é uma função TBC.*

A seguir, vamos apresentar dois resultados simples que serão auxiliares na demonstração do Teorema 3.18.

Lema 3.15 *Se* F_T *é uma função de Fitzpatrick de um operador monótono maximal* $T : X \rightrightarrows X^*$ *então*

$$F_T^*(x^*, x) \geq F_T(x, x^*)$$

para todo $(x, x^*) \in X \times X^*$.

Demonstração. Dado $(x, x^*) \in X \times X^*$, temos que

$$\begin{aligned} F_T^*(x^*, x) &= \sup_{(y,y^*) \in X \times X^*} \{\langle (y, y^*), (x, x^*) \rangle - F_T(y, y^*)\} \\ &\geq \sup_{(y,y^*) \in gra\ T} \{\langle (y, y^*), (x, x^*) \rangle - F_T(y, y^*)\} \\ &= \sup_{(y,y^*) \in gra\ T} \{\langle (y, y^*), (x, x^*) \rangle - \langle y, y^* \rangle\} \\ &= F_T(x, x^*). \end{aligned}$$

□

O próximo resultado auxiliar retrata como os operadores se comportam em relação à translação do seu gráfico. Como veremos, a monotonicidade e a maximalidade são mantidas.

Lema 3.16 *Seja* $T : X \rightrightarrows X^*$ *um operador monótono maximal e seja* $(x, x^*) \in X \times X^*$. *Consideremos o operador* $L : X \rightrightarrows X^*$ *definido por* $L(y) = T(y + x) - x^*$. *Então*

(i) *gra* $L =$ *gra* $T - (x, x^*)$.

(ii) *L é um operador monótono maximal.*

Demonstração. (i) Notamos inicialmente que

$$\begin{aligned} (y,y^*) \in gra\ L \ &\Leftrightarrow \ y^* \in T(y+x) - x^* \\ \Leftrightarrow x^* + y^* \in T(y+x) \ &\Leftrightarrow \ (x,x^*) + (y,y^*) \in gra\ T. \end{aligned} \tag{3.4}$$

logo, $gra\ L \subset gra\ T - (x,x^*)$. Por outro lado, se $(y,y^*) \in gra\ L$ então de (3.4), $(x + y, x^* + y^*) \in gra\ T$ e como $(y,y^*) = (x+y, x^*+y^*) - (x,x^*)$, concluímos a outra inclusão.

(ii) Para verificar que L é monótono, notemos que dados $(y_1,y_1^*), (y_2,y_2^*) \in gra\ L$, existem, pelo item anterior, $(z_1,z_1^*), (z_2,z_2^*) \in gra\ T$ tais que $(y_1,y_1^*) = (z_1 - x, z_1^* - x^*)$ e $(y_2,y_2^*) = (z_2 - x, z_2^* - x^*)$. Portanto, como T é monótono,

$$\langle y_1 - y_2, y_1^* - y_2^* \rangle = \langle z_1 - z_2, z_1^* - z_2^* \rangle \geq 0.$$

Agora vamos verificar que L é maximal. Assim, dado (y,y^*) tal que $\langle z - y, z^* - y^* \rangle \geq 0$ para todo $(z,z^*) \in gra\ L$, temos por (i) que $\langle y + x - w, y^* + x^* - w^* \rangle \geq 0$, para todo $(w,w^*) \in gra\ T$. Como T é maximal, então $(x+y, x^*+y^*) \in gra\ T$ e portanto $(y,y^*) \in gra\ L$. □

Apresentamos, na sequência, uma versão do Teorema de Dualidade de Fenchel cuja demonstração pode ser encontrada em [8, Teorema 4.4.18]. Esta versão será uma ferramenta crucial para o desenvolvimento da generalização proposta.

Teorema 3.17 *Sejam $f, g : X \to \mathbb{R} \cup \{+\infty\}$ funções convexas tais que $0 \in int\ [dom\ f - dom\ g]$, em que int denota interior topológico dos conjuntos. Então*

$$\inf_{x \in X} \{f(x) + g(x)\} = \sup_{x^* \in X^*} \{-f^*(x^*) - g^*(-x^*)\} \tag{3.5}$$

e o supremo em (3.5) é atingido quando finito.

Agora exibimos a nossa contribuição. O resultado a seguir é original e pode ser encontrado em [15]. Como veremos em seguida, tal resultado generaliza a clássica Desigualdade Forte de Fitzpatrick.

Teorema 3.18 *Sejam $T : X \rightrightarrows X^*$ um operador monótono maximal e $g : X \times X^* \to \mathbb{R} \cup \{+\infty\}$ uma função TBC. Considere as seguintes condições:*

(i) *$dom\ F_T = X \times X^*$,*

(ii) *$dom\ g = X \times X^*$.*

Se uma destas hipóteses é satisfeita então

$$F_T(x,x^*) - \langle x, x^* \rangle \geq \frac{1}{2} \inf_{(w,w^*) \in gra\ T} g(w - x, w^* - x^*)$$

para todo $(x, x^*) \in X \times X^*$. Além do mais, esta desigualdade é estrita quando o ínfimo não é atingido.

Demonstração. Para $(x, x^*) \in X \times X^*$, considere o operador monótono maximal $L : X \rightrightarrows X^*$ definido por $L(y) := T(y+x)-x^*$. Pelo Lema 3.16 e se uma das hipóteses (i) ou (ii) é satisfeita, garantimos que as funções F_L e g satisfazem as hipóteses do Teorema 3.17. Sendo assim, como $F_L(y, y^*) \geq \langle y, y^* \rangle$ e $g(y, y^*) \geq -\langle y, y^* \rangle$ para todo $(y, y^*) \in X \times X^*$, temos

$$\begin{aligned} 0 &\leq \inf_{(y,y^*)\in X\times X^*} \{F_L(y, y^*) + g(y, y^*)\} \\ &= \sup_{(y^*,y)\in X^*\times X} \{-F_L^*(y^*, y) - g^*(-y^*, -y)\} \\ &\leq \sup_{(y^*,y)\in X^*\times X} \{-F_L^*(y^*, y) + \langle y, y^* \rangle\} \\ &\leq \sup_{(y^*,y)\in X^*\times X} \{-F_L(y, y^*) + \langle y, y^* \rangle\} \leq 0, \end{aligned} \tag{3.6}$$

em que as duas últimas desigualdades seguem do Lema 3.15 e de (3.3). Logo

$$\max_{(y^*,y)\in X^*\times X} \{-F_L^*(y^*, y) - g^*(-y^*, -y)\} = 0.$$

Portanto existe $(y^*, y) \in X^* \times X$ tal que $F_L^*(y^*, y) + g^*(-y^*, -y) = 0$. Novamente pelo Lema 3.15, concluímos que $F_L(y, y^*) + g^*(-y^*, -y) \leq 0$. Então, como g é uma função TBC, $F_L(y, y^*) + g(y, y^*) \leq 0$ e assim, levando em conta a primeira desigualdade em (3.6), $F_L(y, y^*) + g(y, y^*) = 0$ e (y, y^*) é um minimizador de $F_L + g$.

Como

$$0 = F_L(y, y^*) + g(y, y^*) \geq F_L(y, y^*) - \langle y, y^* \rangle \geq 0,$$

temos

$$-g(y, y^*) = F_L(y, y^*) = \langle y, y^* \rangle .$$

Segue da Proposição 3.10 que $(y, y^*) \in gra\ L$ e, considerando $(w, w^*) := (x + y, x^* + y^*)$, pelo Lema 3.16 temos que $(w, w^*) \in gra\ T$ e $-g(w - x, w^* - x^*) = \langle w - x, w^* - x^* \rangle$, e do Teorema 3.11, concluímos que

$$\begin{aligned} F_T(x, x^*) - \langle x, x^* \rangle &= F_T(w, w^*) - \langle w, w^* \rangle + F_T(x, x^*) - \langle x, x^* \rangle \\ &\geq -\frac{1}{2} \langle w - x, w^* - x^* \rangle = \frac{1}{2} g(w - x, w^* - x^*), \end{aligned}$$

que encerra a prova. □

Diretamente do Exemplo 3.14, temos o seguinte resultado.

Corolário 3.19 *Sejam $T : X \rightrightarrows X^*$ um operador monótono maximal e $f : X \to \mathbb{R} \cup \{+\infty\}$ uma função convexa, própria e sci. Considere as seguintes condições:*

(i) *$dom\ F_T = X \times X^*$,*

(ii) *dom* $f = X$ *e dom* $f^* = X^*$.

Se alguma destas condições valem então

$$F_T(x,x^*) - \langle x, x^* \rangle \geq \frac{1}{2} \inf_{(w,w^*) \in gra\ T} \{f(w-x) + f^*(x^* - w^*)\}$$

para todo $(x,x^*) \in X \times X^*$, *com desigualdade estrita quando o ínfimo não é atingido.*

Em particular, para $p \geq 1$ *e* q *tais que* $\frac{1}{p} + \frac{1}{q} = 1$, *temos*

$$F_T(x,x^*) - \langle x, x^* \rangle \geq \frac{1}{2} \inf_{(w,w^*) \in gra\ T} \left\{ \frac{1}{p}\|w - x\|^p + \frac{1}{q}\|w^* - x^*\|_*^q \right\}$$

para todo $(x,x^*) \in X \times X^*$, *com desigualdade estrita quando o ínfimo não é atingido.*

A respeito da condição *dom* $f^* = X^*$ em (ii) do Corolário 3.19, uma condição suficiente (e necessária no caso de dimensão finita) para que esta valha é que f seja supercoerciva [4, Teorema 3.4], ou seja,

$$\lim_{\|x\| \longrightarrow +\infty} \frac{f(x)}{\|x\|} = +\infty.$$

Para o caso em que $p = 2$ no corolário anterior, recuperamos o resultado já apresentado:

Corolário 3.20 (Desigualdade Forte de Fitzpatrick) *Seja* $T : X \rightrightarrows X^*$ *um operador monótono maximal. Então*

$$F_T(x,x^*) - \langle x, x^* \rangle \geq \frac{1}{4} \inf_{(w,w^*) \in gra\ T} \left\{ \|w - x\|_X^2 + \|w^* - x^*\|_{X^*}^2 \right\}$$

para todo $(x,x^*) \in X \times X^*$.

3.3 Uma nova família de funções *gap*

Nesta seção, vamos aplicar a nossa generalização da Desigualdade de Fitzpatrick na construção de uma família de funções *gap* para o *Problema de Inclusão Monótona Maximal* (PIM). Recordamos que o PIM [6, 7] é o problema de encontrar um ponto $x \in X$ para um operador monótono maximal $T : X \rightrightarrows X^*$ tal que $0 \in T(x)$. O conjunto solução (podendo ser vazio) é $T^{-1}(0) = \{x \in X \mid 0 \in T(x)\}$.

Notamos que, para uma função subdiferenciável f, resolver PIM para o operador ∂f é equivalente a encontrar o minimizador global de f. Assim PIM generaliza muitos problemas de minimização de funções e, além do mais, atua no problema sob o ponto de vista da Teoria de Operadores.

No entanto, é possível transformar o PIM em um problema de otimização fazendo uso de uma ferramenta conhecida na literatura como função *gap* [6].

Definição 3.21 *Uma função gap para o PIM é uma função $\varphi : X \to \mathbb{R} \cup \{+\infty\}$ satisfazendo as seguintes condições:*

i) $\varphi(x) \geq 0$ *para todo* $x \in X$.

ii) $\varphi(x) = 0$ *se e somente se* $x \in T^{-1}(0)$.

Em [6], podem ser encontrados alguns exemplos de funções *gap* para operadores específicos. Vamos analisar a seguir um exemplo que vale para qualquer operador monótono maximal e foi introduzido neste mesmo artigo.

Exemplo 3.22 *Dado um operador* $T : X \rightrightarrows X^*$ *e sua respectiva função de Fitzpatrick* F_T, *definimos uma função gap* G_T *para PIM como* $G_T(x) := F_T(x, 0)$, *ou, mais explicitamente,*

$$G_T(x) = \sup_{(y,y^*) \in gra\ T} \langle x - y, y^* \rangle .$$

Diretamente da Proposição 3.10, concluímos que G_T *é uma função gap.*

O seguinte resultado tem uma prova imediata e pode ser encontrado em [6, Teorema 2.1].

Proposição 3.23 *Seja* φ *uma função gap para PIM. Se PIM tem uma solução, então* $\min_{x \in X} \varphi(x) = 0$. *Por outro lado, se* $\inf_{x \in X} \varphi(x) = 0$ *e* φ *é convexa, sci e fracamente coerciva no sentido de*

$$\lim_{\|x\| \to \infty} \varphi(x) = +\infty$$

então PIM tem solução.

Este último resultado é a motivação para o estudo das funções *gap*, pois quando φ satisfaz suas hipóteses temos que seus minimizadores são exatamente os elementos do conjunto solução de PIM. Assim, podemos reformulá-lo como o problema de minimizar φ.

Antes de introduzirmos a nossa família de funções *gap*, é necessário fazer um rápido estudo sobre seguinte conjunto de funções.

Definição 3.24 *[29, Definição 10.1.1] Uma função convexa própria* $g : X \times X^* \to \mathbb{R} \cup \{+\infty\}$ *é dita Tykhonov quando ela satisfaz as seguintes condições:*

(i) *Ela possui um único minimizador* $(\overline{x}, \overline{x}^*)$.

(ii) *Se* (x_k, x_k^*) *é tal que* $g(x_k, x_k^*) \to g(\overline{x}, \overline{x}^*)$ *então* $(x_k, x_k^*) \to (\overline{x}, \overline{x}^*)$.

Exemplo 3.25 *[29, Exemplo 10.1.4] Toda função sci, convexa e própria* $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *com um único minimizador é Tykhonov. Em particular, funções estritamente convexas e sci com um minimizador.*

Exemplo 3.26 *Uma função $g : X \times X^* \to \mathbb{R}$ é dita fortemente convexa quando existe $\gamma > 0$ tal que, para todo $(x, x^*), (y, y^*) \in X \times X^*$ e $\alpha, \beta \geq 0$ com $\alpha + \beta = 1$,*

$$g(\alpha(x, x^*) + \beta(y, y^*)) \leq \alpha g(x, x^*) + \beta g(y, y^*) - \gamma\alpha\beta\|(x, x^*) - (y, y^*)\|^2_{X\times X^*}. \tag{3.7}$$

Se g é fortemente convexa e tem um minimizador então ela é Tykhonov.

De fato, recordamos que um minimizador $(\overline{x}, \overline{x}^*)$ de uma função fortemente convexa, quando existe, é único. Para provar (ii) da Definição 3.24, seja (x_n, x_n^*) uma sequência em $X \times X^*$ tal que $g(x_n, x_n^*) \to g(\overline{x}, \overline{x}^*)$. Considerando $(x, x^*) := (x_n, x_n^*)$, $(y, y^*) := (\overline{x}, \overline{x}^*)$ e $\alpha = \beta := \frac{1}{2}$ em (3.7), temos que

$$\begin{aligned} 0 &\leq \|(x_n, x_n^*) - (\overline{x}, \overline{x}^*)\|_{X\times X^*} \\ &\leq \sqrt{\frac{2}{\gamma}\left(g(x_n, x_n^*) + g(\overline{x}, \overline{x}^*) - 2g\left(\frac{1}{2}(x_n + \overline{x}, x_n^* + \overline{x}^*)\right)\right)} \\ &\leq \sqrt{\frac{2}{\gamma}\left(g(x_n, x_n^*) - g(\overline{x}, \overline{x}^*)\right)} \end{aligned}$$

em que a última desigualdade vem do fato de $(\overline{x}, \overline{x}^*)$ ser minimizador de g. Dessa forma, como $g(x_n, x_n^*) \to g(\overline{x}, \overline{x}^*)$, concluímos que $(x_n, x_n^*) \to (\overline{x}, \overline{x}^*)$.

Para mais exemplos e referências sobre funções Tykhonov, veja [12, 29]. O resultado a seguir é uma propriedade das funções TBC e será útil na demonstração do Teorema 3.30.

Proposição 3.27 *Seja $g : X \times X^* \to \mathbb{R} \cup \{+\infty\}$ uma função TBC não-negativa. Então*

$$\min_{(x,x^*)\in X\times X^*} g(x, x^*) = g(0, 0) = 0. \tag{3.8}$$

Demonstração. Por (3.3) e como g é não-negativa, temos que

$$0 \leq g(0, 0) \leq g^*(0, 0) = -\inf_{(x,x^*)\in X\times X^*} g(x, x^*) \leq 0.$$

□

Motivados pelo Teorema 3.18, finalmente vamos agora introduzir a nossa família de funções *gap*.

Definição 3.28 (Nossa família de funções *gap*) *Para uma função $g : X \times X^* \to \mathbb{R} \cup \{+\infty\}$ não-negativa, Tykhonov, TBC e $T : X \rightrightarrows X^*$, definimos $G_{T,g} : X \to \mathbb{R} \cup \{+\infty\}$ por*

$$G_{T,g}(x) := \frac{1}{2} \inf_{(w,w^*)\in gra\ T} g(w - x, w^*).$$

Notamos que $G_{T,g}$ é convexa, pois a função $(x, w, w^*) \mapsto g(w - x, w^*)$ é convexa em seus três argumentos, desde que funções TBC são, por definição, convexas. Além do mais, ela é própria quando g for finita em algum $(w, w^*) \in gra\ T$. Diretamente do Teorema 3.18, temos a seguinte propriedade.

Proposição 3.29 *Sob as hipóteses do Teorema 3.18, temos que* $G_T \geq G_{T,g}$*, e logo* $G_{T,g}$ *é finita sempre que* G_T *seja finita.*

Em [6], é possível encontrar condições sob T que garantam que G_T seja finita. Vamos verificar que esta família é composta de fato por funções *gap* para o PIM.

Teorema 3.30 *Seja* $T : X \rightrightarrows X^*$ *um operador monótono maximal e* $g : X \times X^* \to \mathbb{R} \cup \{+\infty\}$ *uma função não-negativa, Tykhonov, TBC. Então* $G_{T,g}$ *é uma função gap para o PIM.*

Demonstração. Notemos que $G_{T,g}(x) \geq 0$ para todo $x \in X$, pois g é não-negativa. Além do mais, se $G_{T,g}(x) = 0$ para algum $x \in X$ então há $(w_n, w_n^*) \subset gra\ T$ tal que $g(w_n - x, w_n^*) \longrightarrow 0$. Como g é Tykhonov, de (3.8) segue que $w_n \longrightarrow x$ e $w_n^* \longrightarrow 0$. Como $gra\ T$ é fechado pelo Teorema 3.7, deduzimos que $(x, 0) \in graT$, ou seja, $0 \in T(x)$. Por outro lado, se $0 \in T(x)$ então, usando que $(x, 0) \in gra\ T$ e da Proposição 3.27, obtemos que $0 \leq G_{T,g}(x) \leq g(0, 0) = 0$, concluindo a prova. □

Notemos que, para $x \in X$ e g como no Teorema 3.30, se existe $(w, w^*) \in gra\ T$ tal que $g(w - x, w^*) = 0$, então $G_{T,g}(x) = 0$ e logo x é uma solução para o PIM. Deste modo, as funções *gap* revelam ser uma boa ferramenta para atacar estes problemas.

Conclusão

Neste trabalho, propusemos uma simplificação da teoria de conjugação apresentada em [11] e uma generalização para a Desigualdade Forte de Fitzpatrick. Fazemos a seguir uma breve revisão das conclusões a respeito destes dois tópicos.

Inicialmente, fizemos uma revisão do conceito de c-conjugação e algumas de suas propriedades. Vimos que a c-conjugada generaliza a ideia de qualquer conjugada existente na Literatura, inclusive a que propusemos. Vimos também que a c-conjugada está sempre relacionada a uma caracterização de funções via supremo de um conjunto particular de funções.

Num segundo momento, nosso foco foi as teorias de conjugação para funções sci. Apresentamos somente as caracterizações relacionadas a elas e, através das propriedades das funções do tipo $\langle \cdot, p(\cdot) \rangle$, com p contínua, concluímos que as funções sci podem ser escritas como supremo pontual de funções contínuas, Fréchet diferenciáveis na origem, cujo gradiente é nulo, e majoradas pela própria função.

Este resultado nos permitiu construir dois esquemas de c-conjugação para funções sci. O primeiro deles não leva em conta a nulidade do gradiente na origem. Apesar de sua formulação ser mais complexa, este esquema possui a vantagem de ter uma relação clara com a teoria introduzida por Fenchel, simplificando os cálculos quando a função em questão é convexa. De qualquer modo, ambas teorias de conjugação propostas cumprem os objetivos do trabalho, simplificando o esquema apresentado em [11].

Em relação ao segundo tópico, através de uma demonstração original da Desigualdade Forte de Fitzpatrick, foi possível generalizá-la através de funções TBC. Essa generalização nos motivou a construir uma família de funções *gap* para o PIM. Essa nossa proposta teve dois fundamentos: primeiramente, definir funções *gap* para o PIM pode não ser uma tarefa fácil. Além do mais, a função *gap* introduzida em [6] [ver exemplo 3.22] está sendo bem recebida pelos pesquisadores da área. Graças a nossa generalização, introduzimos uma família de funções *gap* que é majorada pela função *gap* de [6]. Esse resultado pode vir a ser útil em aplicações. Por exemplo, nossa família de funções *gap* sempre são finitas quando a função de [6] também for.

Recebemos uma resposta bastante positiva e animadora da revista que publicou o nosso segundo artigo referente a nossa família de funções *gap*. Isso nos motiva a seguir estudando esta família em trabalhos futuros, principalmente investigando as vantagens que elas trazem para os problemas de Desigualdade Variacional.

Referências Bibliográficas

[1] J. P. Aubin. *Mathematical Methods of Game and Economic Theory*, North-Holland, 1979.

[2] J. P. Aubin. *Optima and Equilibria*, Springer, 1993.

[3] J. P. Aubin e I. Ekeland. *Applied Nonlinear Analysis*, Wiley, 1984.

[4] H. H. Bauschke, J. M. Borwein, P. L. Combettes. *Essential smoothness, essential strict convexity, and Legendre functions in Banach spaces*, Commun. Contemp. Math., 615–647, 2001.

[5] D. P. Bertsekas, A. Nedic e A. E. Ozdaglar. *Convex Analysis and Optimization*, Athena Scientific, Belmont, Massachusetts, 2003.

[6] J. Borwein e J. Dutta. *Maximal monotone inclusions and Fitzpatrick functions,* edição especial do J Optim Theory Appl. on Nondifferentiable Optimization and Nonsmooth Analysis, 2015. Doi:10.1007/s10957-015-0813-x. Versão online em 30 de Setembro.

[7] J. M. Borwein e A. Lewis. *Convex Analysis and Nonlinear Optmization: Theory and Examples,* Springer, 2000.

[8] J. M. Borwein e J. D. Vanderwerff. *Convex Functions: Constructions, Characterizations and Counterexamples,* Cambridge University Press, 2010.

[9] H. Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer, 2010.

[10] R. S. Burachik e B. F. Svaiter. *Maximal monotone operators, convex functions and a special family of enlargements,* Set-Valued Analysis, vol. 10, pp. 297-316, 2002.

[11] J. Cotrina, E. W. Karas, A. A. Ribeiro, W. Sosa e Y. J. Yun. *Moreau-Fenchel conjugation for lower semi-continuous functions*, Optimization 60, no. 8-9, 1045–1057, 2011.

[12] A. L. Dontchev e T. Zolezzi. *Well-Posed Optimization Problems,* Lecture Notes in Mathematics 1543, Springer, 1993.

[13] R. J. Duffin. *Convex analysis treated by linear programming*, Mathemathical Programming 4, 125–143, 1973.

[14] I. Ekeland e R. Témam. *Convex Analysis and Variational Problems,* Society for Industrial and Applied Mathematics, Philadelphia, EUA, 1999.

[15] L. M. Elias e J. E. Martínez-Legaz. *A generalization of the strong Fitzpatrick inequality*, Optimization 2016. Doi:10.1080/02331934.2016.1179738. Versão online em 07 de Maio.

[16] L. M. Elias e J. E. Martínez-Legaz. *A simplified conjugation scheme for lower semicontinuous functions*, Optimization, vol. 65, 1–13, 2015.

[17] W. Fenchel. *On conjugate convex functions*, Canad. J. Math I, pg. 73-77, 1949.

[18] S. Fitzpatrick. *Representing monotone operators by convex functions,* Workshop Miniconference on Functional Analysis and Optimization, Proc. Centre Math. Anal. Austral. Nat. Univ., 20, Austral. Nat. Univ., Canberra, pág. 59-65, 1988.

[19] F. Flores-Bazán e J. E. Martínez-Legaz. *Simplified global optimality conditions in generalized conjugation theory*, Generalized convexity, generalized monotonicity: recent results, Dordrecht: Kluwer; pg. 305-329, 1998.

[20] P. Germain. *Functional concepts in continuum mechanics*, Meccanica 33, 1998.

[21] L. Hörmander. *Notions of Convexity*, Birkhäuser, 1994.

[22] A. Izmailov e M. V. Solodov. *Condições de Otimalidade, Elementos de Análise Convexa e de Dualidade*, Volume 1, Instituto de Matemática Pura e Aplicada, Rio de Janeiro, Brasil, 2009.

[23] J. E. Martínez-Legaz. *Generalized convex duality and its economic applications*, Handbook of generalized convexity and generalized monotonicity, 237–292, Springer, 2005.

[24] J.E. Martínez-Legaz e B. F. Svaiter. *Monotone operators representable by l.s.c. convex functions,* SetValued Analysis, vol. 13, pp. 21-46, 2005.

[25] E. Michael. *Continuous selection I*, Annals of Mathematics, v. 63, pg. 361-382, 1956.

[26] J. J. Moreau. *Applications of convex analysis to the treatment of elastoplastic systems*, Applications of methods of Functional Analysis to Problems in Mechanics, Sympos. IUTAM/IMU (Germain, P. and Nayroles, B., eds.), Springer, 1976.

[27] J.J. Moreau. *Inf-convolution, sous-additivité, convexité des fonctions numériques*, Journal of Mathematics, v. 63, pg 361-382, 1956.

[28] I. E. Leonard e K. Sundaresan. *Geometry of Lebesgue-Bochner function spaces—smoothness,* Trans. Amer. Math. Soc. 198, 229–251, 1974.

[29] R. Lucchetti. *Convexity and well-posed problems,* Springer, 2006.

[30] J. P. Penot. *Conjugacies adapted to lower semicontinuous functions*, Optimization 64, no. 3, 473–494, 2015.

[31] R. R. Phelps. *Convex functions, monotone operators and differentiability,* Springer, 1993.

[32] R.T. Rockafellar. *Conjugate duality and optimization*, SIAM 1974.

[33] R. T. Rockafellar, R. J.-B. Wets. *Variational analysis,* Springer, 1998.

[34] A. Rubinov, *Abstract convexity and global optimization,* Kluwer Academic Publishers, 2000.

[35] S. Simons. *From Hahn-Banach to Monotonicity,* Springer, 2008.

[36] S. Simons e C. Zalinescu. *A new proof for Rockafellar's characterization of maximal monotone operators,* Proceedings of the American Mathematical Society, vol. 132, pp. 2969-2972, 2004.

[37] I. Singer. *Abstract convex analysis,* John Wiley & Sons, 1997.

[38] K. R. Stromberg. *Introduction to classical real analysis*, Wadsworth International, 1981.

[39] W. Sosa. *Representation of continuous functions and its applications*, J. Optim. Theory Appl. 159, no. 3, 795–804, 2013.

[40] M. D. Voisei e C. Zalinescu. *Strongly-representable monotone operators*, J. Convex Anal. 16, 1011-1033, 2009.

[41] E. Zafer. *Extension and separation of vector valued functions*, Turkish Journal of Mathematics 21 : 423-430, 1997.

APÊNDICES

Apêndice A

Resultados auxiliares

Vamos apresentar alguns conceitos de Análise Convexa que serviram de ferramenta para o desenvolvimento deste trabalho. Vamos exibir propriedades de funções semicontínuas e um resumo sobre conjugação de Fenchel. Os resultados apresentados são baseados em [5, 14].

A.1 Funções semicontínuas inferiormente

Consideremos E um espaço vetorial métrico, $f : E \to \mathbb{R} \cup \{+\infty\} \cup -\infty$ uma função de valores reais estendida, $x_0 \in E$ e $\delta > 0$. Definimos assim

$$M(x_0, \delta) = sup\,\{f(x) \mid x \in B(x_0, \delta)\} \quad e \quad m(x_0, \delta) = inf\,\{f(x) \mid x \in B(x_0, \delta)\}\,.$$

À medida que δ se aproxima monotonicamente de 0, $M(x_0, \delta)$ não cresce e $m(x_0, \delta)$ não decresce. Logo, podemos definir os seguintes limites

$$\liminf_{x \to x_0} f(x) = \lim_{\delta \to 0} m(x_0, \delta) = \sup_{\delta > 0} m(x_0, \delta).$$

$$\limsup_{x \to x_0} f(x) = \lim_{\delta \to 0} M(x_0, \delta) = \inf_{\delta > 0} M(x_0, \delta).$$

Definição A.1

(i) *Uma função f é semicontínua inferiormente ou sci em um ponto $x_0 \in E$ quando* $\liminf_{x \to x_0} f(x) \geq f(x_0)$.

(ii) *Uma função f é semicontínua superiormente ou scs em um ponto $x_0 \in E$ quando* $\limsup_{x \to x_0} f(x) \leq f(x_0)$.

Em sequência, recordamos algumas definições.

Definição A.2

(i) *O epigrafo de uma função f, denotado por epi f, é a região que está acima do seu gráfico, ou seja,*

$$\text{epi } f = \{(x, \lambda) \in E \times \mathbb{R} \mid f(x) \leq \lambda\}.$$

(ii) *O hipografo de uma função f, denotado por hip f, é a região que está abaixo do seu gráfico, ou seja,*

$$\text{hip } f = \{(x, \lambda) \in E \times \mathbb{R} \mid f(x) \geq \lambda\}.$$

(iii) *O conjunto de nível de uma função f para um escalar λ, denotado por $L_f(\lambda)$, é definido como*

$$L_f(\lambda) = \{x \in E \mid f(x) \leq \lambda\}.$$

Definição A.3 *O domínio de uma função f, denotado por dom f, será considerado como a projeção do epi f em E, ou seja,*

$$\text{dom } f = \{x \in E \mid \exists \lambda \in \mathbb{R},\ (x, \lambda) \in \text{epi } f\} = \{x \in E \mid f(x) < +\infty\}.$$

Definição A.4 *Uma função $f : E \to \mathbb{R} \cup \{+\infty\}$ é dita própria quando existe $z \in E$ tal que $f(z) < \infty$.*

O próximo teorema apresenta uma caracterização topológica da semicontinuidade. A topologia considerada é a induzida pela métrica do espaço. Veremos que é possível classificar a função analisando seus conjuntos de nível ou até mesmo o seu epigrafo. Antes de enunciá-lo, vamos apresentar um lema que será útil na sua demonstração.

Lema A.5 *Se $m = \liminf\limits_{x \to x_0} f(x)$ então existe $(x_k) \subset E$ tal que $x_k \to x_0$ e $f(x_k) \to m$.*

Demonstração. Consideremos (δ_k) uma sequência de números positivos que converge para 0 e $m_k = m(x_0, \delta_k)$. Por definição, $m_k \to m$.

Se $m_{k_0} = +\infty$ para algum $k_0 \in \mathbb{N}$, então, como m_k é não decrescente, teremos $m_k = +\infty$ para todo $k > k_0$ e, consequentemente, $m = +\infty$. Além disso, para cada $k > k_0$, como m_k é ínfimo do conjunto $\{f(x) \mid x \in B(x_0, \delta_k)\}$, podemos escolher $x_k \in B(x_0, \delta_k)$ tal que $f(x_k) = +\infty$. Quando $k \to \infty$, a sequência $\delta_k \to 0$, $x_k \to x_0$ e $f(x_k) \to +\infty = m$.

No caso em que $m_k < +\infty$ para todo k, novamente pelo fato de m_k ser ínfimo do conjunto $\{f(x) \mid x \in B(x_0, \delta_k)\}$, podemos escolher $x_k \in B(x_0, \delta_k)$ tal que

$$m_k \leq f(x_k) < m_k + 1/k. \tag{A.1}$$

Fazendo $k \to \infty$, teremos $\delta_k \to 0$, $x_k \to x_0$ e $m_k \to m$. Pela desigualdade $(A.1)$, concluímos que $f(x_k) \to m$. □

Teorema A.6 *Dada a função f, são equivalentes:*

(i) *f é sci em todo E.*

(ii) *epi f é um conjunto fechado em $E \times \mathbb{R}$.*

(iii) *$L_f(\lambda)$ é fechado para todo $\lambda \in \mathbb{R}$.*

Demonstração. $(i) \Rightarrow (ii)$ Consideremos $(x_k, \lambda_k)_{k \in \mathbb{N}} \subset epi\ f$ tal que $(x_k, \lambda_k) \to (x_0, \lambda_0)$. Assim $x_k \to x_0$ e $\lambda_k \to \lambda_0$. Por definição de epigrafo, para todo k, tem-se

$$f(x_k) \leq \lambda_k. \tag{A.2}$$

Para δ fixo, devido ao fato de $x_k \to x_0$, existe k_0 tal que $x_k \in B(x_0, \delta)$ para todo $k > k_0$. Assim, para cada $k > k_0$, $m(x_0, \delta) \leq f(x_k)$ e, pela desigualdade $(A.2)$, $m(x_0, \delta) \leq \lambda_k$.

Fazendo $k \to \infty$ temos, pela conservação de sinal do limite, $m(x_0, \delta) \leq \lambda_0$. Finalmente, fazendo $\delta \to 0$ e pelo fato de f ser sci em todo E, concluímos que

$$f(x_0) \leq \liminf_{x \to x_0} f(x) \leq \lambda_0.$$

Assim $(x_0, \lambda_0) \in epi\ f$ e, portanto, $epi\ f$ é fechado.

$(ii) \Rightarrow (iii)$ Para $\lambda \in \mathbb{R}$, notemos que

$$L_f(\lambda) \times \{\lambda\} = (E \times \{\lambda\}) \cap epi\ f,$$

em que $\times$ denota o produto cartesiano.

De fato,

$$(x, \lambda) \in L_f(\lambda) \times \{\lambda\} \Leftrightarrow f(x) \leq \lambda \Leftrightarrow (x, \lambda) \in epi\ f \cap (E \times \{\lambda\}).$$

Logo, $L_f(\lambda) \times \{\lambda\}$ é fechado. Considerando $T : E \to E \times \mathbb{R}$ dado por $T(x) = (x, \lambda)$, temos que T é contínuo e a pré-imagem de $L_f(\lambda) \times \{\lambda\}$ por T é $L_f(\lambda)$, permitindo concluir que $L_f(\lambda)$ é fechado.

$(iii) \Rightarrow (i)$ Seja x_0 um ponto arbitrário de E. Vamos dividir em dois casos.

1º caso: $\liminf_{x \to x_0} f(x) = +\infty$. Trivialmente $\liminf_{x \to x_0} f(x) \geq f(x_0)$ e portanto, f é sci em x_0.

2º caso: $\liminf_{x \to x_0} f(x) < +\infty$. Pelo Lema A.5, existe (x_k) que converge para x_0 tal que $f(x_k)$ converge para $\liminf_{x \to x_0} f(x)$. Para todo

$$\lambda > \liminf_{x \to x_0} f(x) = \lim_{k \to \infty} f(x_k)$$

existe, por propriedade de limite, k_0 tal que $f(x_k) < \lambda$ para todo $k > k_0$. Assim, a subsequência $(x_{k_0+k})_{k \in \mathbb{N}}$ está contida em $L_f(\lambda)$. Como $L_f(\lambda)$, por hipótese, é fechado então $x_0 \in L_f(\lambda)$. Logo $f(x_0) < \lambda$ para todo $\lambda > \liminf_{x \to x_0} f(x)$.

Agora consideremos a sequência $(\liminf_{x\to x_0} f(x) + 1/k)$. Pelo que já foi provado,

$$f(x_0) < \liminf_{x\to x_0} f(x) + 1/k$$

para todo $k \in \mathbb{N}$. Fazendo $k \to \infty$, pela conservação de sinal do limite, obtemos $f(x_0) \leq \liminf_{x\to x_0} f(x)$ e portanto, f é sci em x_0. Como x_0 é um ponto arbitrário, então f é sci em todo E. □

Com demonstração análoga, podemos garantir que uma função é scs quando seu hipografo é fechado ou quando o conjunto $\{x \in E \mid f(x) \geq \lambda\}$ é fechado para todo número real λ.

Notemos que, pela linearidade do limite, podemos concluir que se f e g são sci (scs) então $cf + g$ também é sci (scs) para todo real $c > 0$. Podemos concluir também, por esse último resultado, que toda função contínua é scs e sci, pois os conjuntos $L_f(\lambda)$ e $\{x \in E \mid f(x) \geq \lambda\}$ são pré-imagens de conjuntos fechados e consequentemente serão fechados também.

A seguir, vemos um resultado que possibilita uma interessante caracterização de semicontinuidade.

Lema A.7

(i) *Seja* $m = \liminf_{x\to x_0} f(x)$. *Dado* $m' < m$, *existe* $\delta > 0$ *tal que* $m' < f(x)$ *para todo* $x \in B(x_0, \delta)$.

(ii) *Seja* $M = \limsup_{x\to x_0} f(x)$. *Dado* $M' > M$, *existe* $\delta > 0$ *tal que* $M' > f(x)$ *para todo* $x \in B(x_0, \delta)$.

Demonstração. (i) Dado m' tal que $m > m'$, temos, por definição, que $\lim_{\delta\to 0} m(x_0, \delta) > m'$. Por propriedade de limite, existe algum $\delta' > 0$ tal que $m(x_0, \delta') > m'$, ou seja,

$$inf\left\{f(x) \mid x \in B(x_0, \delta')\right\} > m'.$$

Logo, $f(x) > m'$ para todo $x \in B(x_0, \delta')$.

A demonstração do segundo item é análoga à do primeiro. □

Teorema A.8

(i) *Se* $f(x_0) = -\infty$, *então* f *é sci em* x_0. *Caso* $f(x_0) > -\infty$, *então* f *é sci em* x_0 *se, e só se, dado* $m' < f(x_0)$, *existir* $\delta > 0$ *tal que* $f(x) > m'$ *para todo* $x \in B(x_0, \delta)$.

(ii) *Se* $f(x_0) = +\infty$, *então* f *é scs em* x_0. *Caso* $f(x_0) < \infty$, *então* f *é scs em* x_0 *se, e só se, dado* $M' > f(x_0)$, *existir* $\delta > 0$ *tal que* $M' > f(x)$ *para todo* $x \in B(x_0, \delta)$.

Demonstração. (i) Se $f(x_0) = -\infty$, não há nada a se provar. Caso $f(x_0) > -\infty$ e f seja sci em x_0, então dado $m' < f(x_0)$, diretamente da definição de semicontinuidade, concluímos que $\liminf\limits_{x \to x_0} f(x) > m'$. Logo, pelo Lema A.7, concluímos o desejado.

Para provar a recíproca, suponhamos por absurdo que $\liminf\limits_{x \to x_0} f(x) < f(x_0)$. Assim existe, por hipótese, $\delta > 0$ tal que $f(x) > \liminf\limits_{x \to x_0} f(x)$ para todo $x \in B(x_0, \delta)$. Portanto,

$$m(x_0, \delta) > \liminf_{x \to x_0} f(x).$$

Absurdo, pois contraria a definição de liminf.

A demonstração do segundo item é novamente análoga à do primeiro. □

A.2 A conjugada de Fenchel

Nesta seção, vamos tratar da teoria de conjugação de Fenchel (ou dita clássica).

Consideremos E^* o espaço dual topológico do espaço vetorial métrico E, ou seja, o espaço de todos os funcionais lineares contínuos de E e denotamos o produto de dualidade de E^* com E por $\langle \cdot, \cdot \rangle$. Assim, definimos:

Definição A.9 *Dada uma função $f : E \to \mathbb{R} \cup \{+\infty\}$ própria, a função conjugada de f, denotada por f^*, é definida como $f^* : E^* \to \mathbb{R} \cup \{+\infty\}$ sendo*

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

Notemos que f^* está bem definida, pois como f é própria, existe um $x \in dom\ f$ tal que $\langle x, x^* \rangle - f(x) > -\infty$ e portanto $f^*(x^*) > -\infty$ para todo $x^* \in E^*$. Essa função é também conhecida na literatura como Conjugada de Fenchel ou Conjugada clássica.

Para efeitos da ilustração gráfica, considerando $E = \mathbb{R}^n$, para uma função f própria, convexa e sci, o hiperplano

$$H = \{(x, \lambda) \mid \langle x, x^* \rangle - \lambda = f^*(x^*), x^* \in \mathbb{R}^n\}$$

suporta o epigrafo de f no ponto em que tangencia o gráfico da f.

Como podemos observar pela Figura A.1, se tomarmos $-b$ como a distância vertical do gráfico de f com a reta $\lambda = \langle x, x^* \rangle$, ou seja,

$$-b = \inf_{x \in E} \{f(x) - \langle x, x^* \rangle\}$$

ou, equivalentemente,

$$b = \sup_{x \in E} \{\langle x, x^* \rangle - f(x)\},$$

então o hiperplano H será suporte do *epi* f no ponto em que tangencia o gráfico da f.

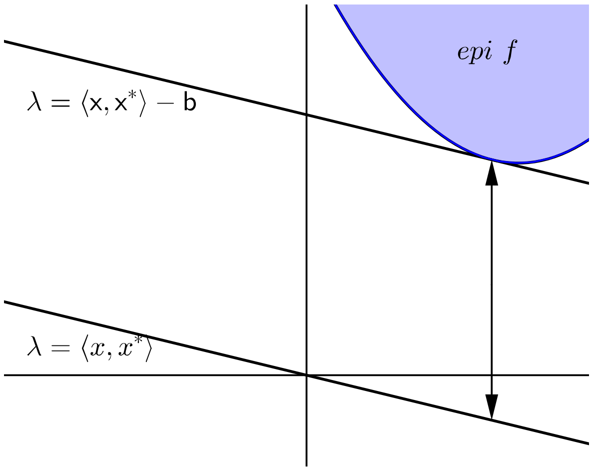

Figura A.1: Reta $\lambda = \langle x, x^* \rangle - b$ tangente ao gráfico de f.

Para cada $x \in E$, consideramos $f_x : E^* \to \mathbb{R} \cup \{-\infty\}$ definido como $f_x(x^*) = \langle x, x^* \rangle - f(x)$. Assim, temos que f_x é convexa e contínua para todo x e, além disso, $f^*(x^*) = \sup \{f_x(x^*);\ x \in E\}$. Logo, como f^* é o supremo de funções convexas e contínuas, podemos concluir que a função conjugada será sempre convexa e sci. Veremos agora outras propriedades.

Teorema A.10 *Sejam uma função f própria e f^* sua conjugada. Então*

(i) $f(x) + f^*(x^*) \geq \langle x, x^* \rangle$ *para todos x em E e x^* em E^*;*

(ii) $\inf_{x \in E} f(x) = -f^*(0)$;

(iii) *se $f(x) \leq g(x)$ para todo $x \in E$ então $f^*(x^*) \geq g^*(x^*)$ para todo $x^* \in E^*$.*

Demonstração. (i) É consequência direta da definição de função conjugada.

(ii) Por definição de função conjugada, temos que

$$-f^*(0) = -\sup_{x \in E} \{\langle x, 0 \rangle - f(x)\} = \inf_{x \in E} \{f(x)\}.$$

(iii) Se $f(x) \leq g(x)$ para todo $x \in E$ então

$$\langle x, x^* \rangle - f(x) \geq \langle x, x^* \rangle - g(x)$$

para todos $x \in E$ e $x^* \in E^*$. Logo

$$f^*(x^*) = \sup_{x \in E} \{\langle x, x^* \rangle - f(x)\} \geq \sup_{x \in E} \{\langle x, x^* \rangle - g(x)\} = g^*(x^*)$$

para todo $x^* \in E^*$. □

É importante notar que os pontos em que f atinge $+\infty$ não influenciam no cálculo da conjugada de f, pois a expressão $\langle x, x^* \rangle - f(x)$ vale $-\infty$ nesses pontos e, como f é própria, quando tomarmos o supremo, estes pontos serão descartados. Logo, quando for conveniente, podemos tomar o supremo apenas sobre o domínio de f, ou seja,

$$f^*(x^*) = \sup_{x \in dom\ f} \{\langle x, x^* \rangle - f(x)\}.$$

De maneira análoga ao que foi feito antes, podemos definir conjugadas de funções do tipo $g : E^* \to \mathbb{R} \cup \{+\infty\}$ como $g^* : E \to \mathbb{R} \cup \{+\infty\}$ tal que

$$g^*(x) = \sup_{x^* \in E^*} \{\langle x, x^* \rangle - g(x^*)\}.$$

Assim a conjugada da função conjugada, denominada biconjugada e denotada por f^{**}, é dada naturalmente por

$$f^{**}(x) = (f^*)^*(x) = \sup_{x \in E^*} \{\langle x, x^* \rangle - f^*(x^*)\}.$$

O próximo teorema mostra que é possível caracterizar a biconjugada f^{**} como supremo de funções majoradas por f.

Teorema A.11 *Seja $G = \{g : E \to \mathbb{R} \mid g(y) = \langle y, x^* \rangle + b,\ x^* \in E^*,\ b \in \mathbb{R}\}$. Consideremos f uma função própria, então*

$$f^{**}(x) = \sup_{g \in G} \{g(x) \mid g(y) \leq f(y),\ \forall y \in E\}.$$

Demonstração. Demonstração análoga à do Teorema 1.7. □

Como corolário direto, temos a seguinte desigualdade.

Corolário A.12 *Dada uma função própria f e sua biconjugada f^{**} então*

$$f^{**}(x) \leq f(x) \tag{A.3}$$

para todo $x \in E$.

A igualdade em $(A.3)$ é satisfeita para o seguinte conjunto de funções.

Teorema A.13 (Teorema de Fenchel-Moreau) *Uma função própria f é convexa e sci se, e somente se*

$$f^{**}(x) = f(x)$$

para todo $x \in E$.

Demonstração. Ver demonstração em [9, Teorema 1.11]. □

Apêndice B

Artigo: *A simplified conjugation scheme for lower semi-continuous functions*

Taylor & Francis
Taylor & Francis Group

Optimization
A Journal of Mathematical Programming and Operations Research

ISSN: 0233-1934 (Print) 1029-4945 (Online) Journal homepage: http://www.tandfonline.com/loi/gopt20

A simplified conjugation scheme for lower semi-continuous functions

Leonardo M. Elias & Juan E. Martínez-Legaz

To cite this article: Leonardo M. Elias & Juan E. Martínez-Legaz (2016) A simplified conjugation scheme for lower semi-continuous functions, Optimization, 65:4, 751-763, DOI: 10.1080/02331934.2015.1080700

To link to this article: http://dx.doi.org/10.1080/02331934.2015.1080700

Published online: 01 Sep 2015.

Optimization, 2016
Vol. 65, No. 4, 751–763, http://dx.doi.org/10.1080/02331934.2015.1080700

A simplified conjugation scheme for lower semi-continuous functions

Leonardo M. Elias[ab1] and Juan E. Martínez-Legaz[cd*]

[a] *Programa de Pós-Graduação em Matemática, Universidade Federal do Paraná, Curitiba, Brazil;* [b] *CAPES Foundation, Ministry of Education of Brazil, Brasília, Brazil;* [c] *Departament d'Economia i d'Història Econòmica, Universitat Autònoma de Barcelona, Bellaterra, Spain;* [d] *MOVE (Markets, Organizations and Votes in Economics), Bellaterra, Spain*

(*Received 9 January 2015; accepted 25 July 2015*)

We present two generalized conjugation schemes for lower semi-continuous functions defined on a real Banach space whose norm is Fréchet differentiable off the origin, and sketch their applications to optimization duality theory. Both approaches are based upon a new characterization of lower semi-continuous functions as pointwise suprema of a special class of continuous functions.

Keywords: lower semi-continuous function; generalized convex conjugation; optimization duality theory

1. Introduction

This paper elaborates on a conjugation theory for lower semi-continuous functions introduced in the recent paper [1], in which, for functions defined on $\mathbb{R}^n$, the authors present a conjugation scheme such that the second conjugate of any function coincides with the pointwise supremum of its minorants of the type $x \longmapsto \langle x, p(x)\rangle + r$, with $\langle \cdot, \cdot \rangle$, $p : \mathbb{R}^n \to \mathbb{R}^n$ and r denoting the Euclidean scalar product, a continuous mapping, and a constant, respectively.[1, Proposition 3.2(iii)] To simplify that conjugation scheme, we first characterize this special class of functions: In the case of a real Banach space X whose norm is Fréchet differentiable off the origin, a function admits such a representation if, and only if, it is continuous on X and Fréchet differentiable at the origin (Theorem 4). Thanks to this characterization, we can simplify the conjugation theory of [1] by considering a 'dual' space consisting of real-valued functions on X rather than continuous mappings $p : X \to X$, the latter being, in general, much more complex mathematical objects than the former. Notice that a continuous mapping $p : X \to X$ yields a unique function of the type $x \longmapsto \langle x, p(x)\rangle$, but there are in general infinitely many such continuous mappings p yielding the same function. Consequently, dealing with functions $X \to \mathbb{R}$ rather than mappings $X \to X$, we drastically reduce the dimension of the 'dual' space. We present two alternative conjugation schemes. In the first one, the 'dual' variables are pairs consisting of continuous linear functionals and general continuous real-valued functions; the

*Corresponding author. Email: JuanEnrique.Martinez.Legaz@uab.cat

[1] This work was developed during a visit of this author to the Departament d'Economia i d'Història Econòmica of the Universitat Autònoma de Barcelona.

consideration of continuous linear functionals allows us to relate our conjugation operator to classical convex conjugation. The second scheme is simpler, in that the arguments of the conjugate functions are just continuous real-valued functions; however, since the dual space X^* does not appear in the picture, this second scheme is not directly related to Fenchel conjugation. We briefly sketch the application of our conjugation schemes to optimization duality theorem.

The rest of the paper is organized as follows. Section 2 presents the fundamental results on which our conjugation approaches are based. One of its main results states that every lower semi-continuous function is the pointwise supremum of a collection of continuous functions which are Fréchet differentiable at the origin and whose gradients at that point are equal to 0. In Section 3 we develop our first conjugation scheme and show its relationship with classical convex conjugation as well as its application to optimization duality theory. Section 4 presents our second approach to generalized conjugation for lower semi-continuous functions and the duality theory based on it. In Section 5 we illustrate both duality schemes by means of simple examples.

2. Preliminaries

Throughout this paper, we will assume that $(X, \|\cdot\|)$ is a real Banach space whose norm is Fréchet differentiable off the origin. Hilbert spaces are examples of such spaces; moreover, if $(X, \|\cdot\|)$ is of that type, μ is a measure on X, and $p \in (1, \infty)$, then the norm of $L^p(X, \mu)$ has that differentiability property, too [2, Theorem 3.1]. We will denote by $(X^*, \|\cdot\|_*)$ the dual space of X, and by $\langle \cdot, \cdot \rangle : X \times X^* \to \mathbb{R}$ the duality pairing. We recall that the duality mapping $J : X \rightrightarrows X^*$ is $J := \partial \left(\frac{1}{2} \|\cdot\|^2 \right)$ or, equivalently,

$$J(x) := \left\{ x^* \in X^* | \; \|x\|^2 = \langle x, x^* \rangle = \|x^*\|_*^2 \right\} \quad (x \in X).$$

Since a convex function is Fréchet differentiable precisely when its subdifferential is single-valued and norm-to-norm continuous (see e.g. [3, Introduction]), our differentiability assumption on $\|\cdot\|$ is equivalent to the single-valuedness and norm-to-norm continuity of J. For the sake of notational simplicity, we will identify the singleton $J(x)$ with its element; under this identification, $J(x)$ is characterized by the equalities

$$\|x\|^2 = \langle x, J(x) \rangle = \|J(x)\|_*^2 . \tag{1}$$

The following classical theorem will play a key role in the generalized conjugation theory we will develop.

Theorem 1 [4, p.133] *If E is a metric space, $f : E \to \mathbb{R} \cup \{+\infty\}$ is lsc, $g : E \to \mathbb{R} \cup \{-\infty\}$ is usc, and $g \leq f$, then there exists a continuous function $h : E \to \mathbb{R}$ such that $g \leq h \leq f$.*

The lsc hull (that is, the largest semi-continuous minorant) of a function f will be denoted clf. As is well known, f is lsc at a point x if, and only if, $clf(x) = f(x)$.

COROLLARY 2 *Let E be a metric space, $f : E \to \mathbb{R} \cup \{+\infty\}$, and $x \in E$. Then f is lsc at x if, and only if,*

$$f(x) = \sup \{h(x) \mid h : E \to \mathbb{R} \text{ is continuous, } h \leq f\}, \tag{2}$$

and the supremum is attained if $x \in dom(f)$.

Proof Let as assume that f is lsc at x. We will prove that the inequality

$$f(x) \geq \sup \{h(x) \mid h : E \to \mathbb{R} \text{ is continuous, } h \leq f\}$$

holds with the equal sign and the supremum in the right-hand side of this inequality is attained if $x \in dom(f)$, by showing that, for every real number $\lambda \leq f(x)$, there exists a continuous function $h_\lambda : E \to \mathbb{R}$ satisfying $h_\lambda \leq f$ and $h_\lambda(x) \geq \lambda$. Define $g_\lambda : E \to \mathbb{R} \cup \{-\infty\}$ by $g_\lambda(x) := \lambda$ and $g_\lambda(y) := -\infty$ for $y \neq x$. Note that g_λ is usc and satisfies $g_\lambda(y) \leq clf(y)$ for all $y \in E$, since we have $clf(x) = f(x)$ by the lower semi-continuity of f at x. Hence, by Theorem 1, there exists a continuous function $h : E \to \mathbb{R}$ such that

$$g_\lambda(y) \leq h_\lambda(y) \leq clf(y) \quad \text{ for all } y \in E.$$

Therefore $h_\lambda \leq f$ and $\lambda = g_\lambda(x) \leq h_\lambda(x)$.

Conversely, assume that (2) holds. Since the function

$$y \longmapsto \sup \{h(y) \mid h : E \to \mathbb{R} \text{ is continuous, } h \leq f\}$$

is an lsc minorant of f, it is a minorant of $clf(y)$, too. Hence, by (2), we have

$$clf(x) = f(x),$$

that is, f is lsc at x. □

The following theorem is a complement to the Fenchel-Moreau conjugation theory developed in [1, Section 4]. It can be regarded as a generalized subdifferentiability result.

THEOREM 3 *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ and $x \in dom(f) \setminus \{0\}$. Then f is lsc at x if, and only if, there exists a continuous mapping $p : X \to X^*$ satisfying $p(0) = 0$ and*

$$f(y) \geq f(x) + \langle y, p(y)\rangle - \langle x, p(x)\rangle \quad \text{for all } y \in X. \tag{3}$$

Proof If f is lsc at x, then, by Corollary 2, there exists a continuous function $h : X \to \mathbb{R}$ such that $h \leq f$ and $h(x) = f(x)$. Let $\delta := \inf_{\|y\| \leq \|x\|} h(y)$. We define $p : X \to X^*$ by

$$p(y) := \frac{h(y) - \delta}{(\max\{\|y\|, \|x\|\})^2} J(y).$$

Clearly, p is continuous and satisfies

$$\langle y, p(y)\rangle = (h(y) - \delta)\left(\min\left\{\frac{\|y\|}{\|x\|}, 1\right\}\right)^2.$$

Therefore, $\langle y, p(y)\rangle \leq h(y) - \delta \leq f(y) - \delta$ for all $y \in X$, which implies $\langle y, p(y)\rangle - f(y) \leq -\delta$ for all $y \in X$. Furthermore, $\langle x, p(x)\rangle = h(x) - \delta = f(x) - \delta$. Consequently, (3) holds.

Conversely, assume that (3) holds, and define $h_0 : E \to \mathbb{R}$ by

$$h_0(y) := f(x) + \langle y, p(y)\rangle - \langle x, p(x)\rangle .$$

Since h_0 is continuous and, by (3), we have $h_0 \leq f$, it turns out that

$$\sup\{h(x) \,|\, h : E \to \mathbb{R} \text{ is continuous}, h \leq f\} \geq h_0(x) = f(x) .$$

Hence, (2) holds, and therefore the lower semi-continuity of f at x follows from Corollary 2. □

Theorem 3 shows the role of functions of the type $x \longmapsto \langle x, p(x)\rangle$, with p continuous. In fact, the generalized conjugation theory developed in [1] is built on such functions. Thanks to the following characterization, in the next section we will present a new, simplified version of that theory.

Theorem 4 *Let $f : X \to \mathbb{R}$ be a function. The following statements are equivalent:*

(i) *There exists a continuous mapping $p : X \to X^*$ such that $f(x) = \langle x, p(x)\rangle$ for every $x \in X$.*
(ii) *The function f is continuous on X and Fréchet differentiable at 0, and satisfies $f(0) = 0$.*

Proof $(i) \Rightarrow (ii)$ Continuity and the equality $f(0) = 0$ are obvious, whereas Fréchet differentiability follows from the continuity of p, given that

$$0 \leq \frac{|f(x) - f(0) - \langle x, p(0)\rangle|}{\|x\|} = \frac{|\langle x, p(x) - p(0)\rangle|}{\|x\|} \leq \|p(x) - p(0)\|_* .$$

$(ii) \Rightarrow (i)$ Define $p : X \to X^*$ by

$$p(x) = \begin{cases} \nabla f(0) + \frac{f(x) - \langle x, \nabla f(0)\rangle}{\|x\|^2} J(x), & \text{if } x \neq 0 \\ \nabla f(0), & \text{otherwise.} \end{cases} \tag{4}$$

Since f is continuous, so is p on $X\backslash\{0\}$. Continuity at 0 also holds, in view of (1):

$$\begin{aligned} \lim_{x\to 0} \|p(x) - p(0)\|_* &= \lim_{x\to 0} \frac{|f(x) - \langle x, \nabla f(0)\rangle|}{\|x\|^2} \|J(x)\|_* \\ &= \lim_{x\to 0} \frac{|f(x) - \langle x, \nabla f(0)\rangle|}{\|x\|} = 0. \end{aligned}$$

For $x \in X \setminus \{0\}$, using again (1) we obtain

$$\begin{aligned} \langle x, p(x)\rangle &= \langle x, \nabla f(0) + \frac{f(x) - \langle x, \nabla f(0)\rangle}{\|x\|^2} J(x)\rangle \\ &= \langle x, \nabla f(0)\rangle + \frac{f(x) - \langle x, \nabla f(0)\rangle}{\|x\|^2} \langle x, J(x)\rangle = f(x), \end{aligned}$$

whereas the equality $\langle x, p(x)\rangle = f(x)$ also holds, obviously, for $x = 0$. □

While preparing a revised version of this paper, we have learned that a statement equivalent to Theorem 4 can be found in Theorem 3.11 of the very recent article [5].

Corollary 5 *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ and $x \in X$. Then f is lsc at x if, and only if,*

$$f(x) = \sup\{h(x) | h : X \to \mathbb{R} \text{ is continuous on } X \text{ and Fréchet differentiable at } 0,\ \nabla h(0) = 0,\ h \leq f\}; \tag{5}$$

the supremum is attained if $x \in dom(f) \setminus \{0\}$.

Proof The 'if' statement follows from the corresponding statement of Corollary 2.

Conversely, assume that f is lsc at x. In view of Corollary 2, in order to prove (5) we may assume, without loss of generality, that f is continuous on X. Moreover, since $f = \sup_{\lambda \in \mathbb{R}} f_\lambda$, with

$$f_\lambda := \min\{f, \lambda\},$$

we can further assume that f is finite valued.

Let us consider first the case when $x = 0$. By the continuity of f at 0, if $\lambda < f(0)$, the function f_λ takes the constant value λ on a neighbourhood of 0, and hence f_λ is Fréchet differentiable with $\nabla f_\lambda(0) = 0$. Therefore, (5) follows from the equality $f(0) = \sup\{f_\lambda(0) | \lambda < f(0)\}$.

If $x \neq 0$, then, taking p as in Theorem 3 and defining $h : X \to \mathbb{R}$ by $h(y) := \langle y, p(y)\rangle + f(x) - \langle x, p(x)\rangle$, from Theorem 4, implication (i) $\Longrightarrow$ (ii), it follows that h is continuous on X, Fréchet differentiable at 0, and satisfies $\nabla h(0) = 0$; moreover, by (3), we have $h \leq f$. Therefore, since $h(x) = f(x)$, this proves that the supremum in (5) is attained at this precise h and that equality (5) holds. □

Remark 6 Notice that, when $\nabla f(0) = 0$, the mapping p defined by (4) is pointwise norm minimizing, that is, for every $x \in X$ the least norm solution of the linear equation $\langle x, y^*\rangle = f(x)$ is precisely $p(x)$. This is obvious for $x = 0$, whereas for $x \neq 0$ the assertion follows by observing that, for any solution y^*, in view of (1) one has

$$\|p(x)\|_* = \frac{|f(x)|}{\|x\|^2}\|J(x)\|_* = \frac{|\langle x, y^*\rangle|}{\|x\|} \leq \|y^*\|_*.$$

When $\nabla f(0) \neq 0$, the mapping p of (4) is obtained by adding the constant mapping $x \longmapsto \nabla f(0)$ to the pointwise norm minimizing mapping corresponding to $g := f - \nabla f(0)$.

In view of Theorem 4, Corollary 5 may be regarded as a refinement of [6, Theorem 6.1] (see also [5, Theorem 3.11]), which characterizes lsc functions on a Hilbert space as pointwise suprema of functions of the type $x \mapsto \langle x, p(x)\rangle + \beta$, with $p : X \to X^*$ continuous and β constant. Our set of minorants h is smaller, since their gradients vanish at the origin, a condition which is absent in the statement of [6, Theorem 6.1] (as well as in that of [5, Theorem 3.11]); indeed, in the formulation of [6, Theorem 6.1], it would read $p(0) = 0$.

It is worth pointing out that the special role that the origin plays in Theorem 4 and (5) can be equally played by any other point $y \in X$, as the next result state. However, for the sake of simplicity, in the next sections we will state all of our results only for the case when $y := 0$.

Corollary 7 *Let $f : X \to \mathbb{R}$ be a continuous function and $y \in X$. Then f is Fréchet differentiable at y and satisfies $f(y) = 0$ if, and only if, there exists a continuous mapping $p : X \to X^*$ such that $f(x) = \langle x - y, p(x)\rangle$ for every $x \in X$.*

Corollary 8 *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ and $x, y \in X$. Then f is lsc at x if, and only if,*

$$f(x) = \sup\{h(x) | h : X \to \mathbb{R} \text{ is continuous on } X \text{ and Fréchet differentiable at } y,\ \nabla h(y) = 0,\ h \leq f\};$$

the supremum is attained if $x \in dom\,(f) \setminus \{y\}$.

The proofs of the two latter corollaries easily follow from Theorem 4 and Corollary 5, replacing f by $g := f(\cdot + y)$ in their statements.

In the case when X is a Hilbert space, after identifying it with its dual by means of the Riesz–Fréchet Representation Theorem, it turns out that J is the identity mapping. Therefore, by a straightforward modification of the proof of Theorem 4, we get the following result, which provides two characterizations of everywhere differentiable functions on Hilbert spaces.

Theorem 9 *If H is a real Hilbert space and $y \in H$, for $f : H \to \mathbb{R}$ the following statements are equivalent:*

(i) *There exists a continuous mapping $p : H \to H$, Fréchet differentiable on $H \setminus \{y\}$, such that $f(x) = \langle x - y, p(x) \rangle$ for every $x \in H$.*
(ii) *The function f is Fréchet differentiable and satisfies $f(y) = 0$.*

3. The first conjugation scheme

In this section, we will develop our first conjugation theory for lsc functions. We will follow the approach to generalized conjugation theory presented in [7, Section 2].

We introduce the following set of functions

$$\mathcal{H}_X := \big\{h : X \to \mathbb{R} | h \text{ is continuous on } X \text{ and Fréchet differentiable at } 0, h(0) = 0,\ \nabla h(0) = 0\}$$

and the coupling function $c : X \times (X^* \times \mathcal{H}_X) \to \mathbb{R}$ defined by

$$c\left(x, \left(x^*, h\right)\right) := \left\langle x, x^* \right\rangle + h(x).$$

The c-conjugate function of $f : X \to \overline{\mathbb{R}}$ is $f^c : X^* \times \mathcal{H}_X \to \overline{\mathbb{R}}$, given by

$$f^c(x^*, h) = \sup_{x \in X} \left\{\left\langle x, x^* \right\rangle + h(x) - f(x)\right\}. \tag{6}$$

Notice that the classical Fenchel conjugation operator * is related to c-conjugation by the following formulas:

$$f^c(x^*, h) = (f - h)^* \left(x^*\right), \quad f^* \left(x^*\right) = f^c(x^*, 0). \tag{7}$$

The 'inverse' conjugation operator corresponds to the coupling function $c' : (X^* \times \mathcal{H}_X) \times X \to \mathbb{R}$ defined by

$$c'\left(\left(x^*, h\right), x\right) := c\left(x, \left(x^*, h\right)\right).$$

We say that $f : X \to \overline{\mathbb{R}}$ is c-subdifferentiable at $x \in dom(f)$ if there exists $(x^*, h) \in X^* \times \mathcal{H}_X$ such that

$$f(y) - f(x) \geq \left\langle y - x, x^* \right\rangle + h(y) - h(x) \quad \text{for all } y \in X. \tag{8}$$

Then (x^*, h) is said to be a c-subgradient of f at x. The set of all c-subgradients of f at x, denoted $\partial_c f(x)$, is called the c-subdifferential of f at x. If $f(x) \notin \mathbb{R}$, we set $\partial_c f(x) := \emptyset$. The relationship between the c-subdifferential and the classical Fenchel subdifferential ∂ is as follows:

$$\left(x^*, h\right) \in \partial_c f(x) \Longleftrightarrow x^* \in \partial\,(f-h)\,(x)\,, \quad x^* \in \partial f\,(x) \Longleftrightarrow \left(x^*, 0\right) \in \partial_c f(x). \quad (9)$$

The c'-conjugate function $\varphi^{c'} : X \to \overline{\mathbb{R}}$ of $\varphi : X^* \times \mathcal{H}_X \to \overline{\mathbb{R}}$ is defined by

$$\varphi^{c'}(x) = \sup_{(x^*,h)\in X^*\times\mathcal{H}_X} \left\{\left\langle x, x^*\right\rangle + h(x) - \varphi\left(x^*, h\right)\right\}. \quad (10)$$

This c'-conjugate function relates to Fenchel conjugation according to the following formula, which easily follows from (10):

$$\varphi^{c'}(x) = \sup_{h\in\mathcal{H}_X} \left\{h(x) + \varphi\,(\cdot, h)^*\,(x)\right\}. \quad (11)$$

The biconjugate of f is $f^{cc'} := (f^c)^{c'}$. Combining (11) with (7), one gets

$$f^{cc'}(x) = \sup_{h\in\mathcal{H}_X} \left\{h(x) + (f-h)^{**}\,(x)\right\}. \quad (12)$$

For a full description of the generalized conjugation framework we are using, we refer to [7]; more details on abstract convexity can be found in [8,9].

The following result is an easy consequence of [7, Proposition 6.2].

Theorem 10 *Let $f : X \to \mathbb{R}\cup\{+\infty\}$. Then*

$$f^{cc'} = \sup\left\{g : X \to \mathbb{R} | g \text{ is continuous on } X \textit{and Fréchet} \right.$$
$$\left. \text{differentiable at } 0,\, g \le f\right\}.$$

Combining Theorem 10 with Corollary 5, we obtain the following result.

Corollary 11 *Let $f : X \to \mathbb{R}\cup\{+\infty\}$ and $x \in X$. Then f is lsc at x if, and only if, $f^{cc'}(x) = f(x)$. Moreover, if $x \in dom\,(f)\setminus\{0\}$ and f is lsc at x, then $\partial_c f(x) \neq \emptyset$.*

Combining Corollary 11 with (12), and using the well known fact that every lsc proper convex function f coincides with its second Fenchel conjugate f^{**}, one immediately sees that the supremum in (12) is attained at $h = 0$.

The next proposition relates the c-subdifferential at the origin to the Fréchet subdifferential. We recall that $x^* \in X^*$ is a Fréchet (or regular, following the terminology of [10]) subgradient of f at $x \in dom\,(f)$ if

$$\liminf_{y\to x} \frac{f\,(y) - f\,(x) - \langle y - x, x^*\rangle}{\|y - x\|} \ge 0.$$

The set of all such Fréchet subgradients is called the Fréchet subdifferential of f at x; we will denote this set by $\partial_F\,(x)$.

Proposition 12 *Let $f : X \to \mathbb{R}\cup\{+\infty\}$ be such that $0 \in dom\,(f)$. Then the projection of $\partial_c f(0)$ onto X^* coincides with $\partial_F\,(0)$.*

Proof If x^* belongs to the projection of $\partial_c f(0)$ onto X^*, then there exists $h \in \mathcal{H}_X$ such that (8) holds for $x = 0$, which means that the function $g := x^* + h + f(0)$ is a minorant of f. Since $g(0) = f(0)$ and $\nabla g(0) = x^*$, by [10, Proposition 8.5] we have $x^* \in \partial_F f(0)$.

Conversely, let $x^* \in \partial_F f(0)$. Then, again by [10, Proposition 8.5], on some closed neighborhood V of 0 there is a smooth function $g \leq f$ with $g(0) = f(0)$ and $\nabla g(0) = x^*$. We clearly have $g \leq clf$. We extend g to the whole of X by setting $g(y) := -\infty$ for $y \in X \setminus V$. This preserves the upper semi-continuity of g as well as the inequality $g \leq clf$. Hence, by Theorem 1, there exists a continuous function $h : X \to \mathbb{R}$ such that $g \leq h \leq clf$. We now use the Tietze extension theorem to prove the existence of a nonnegative continuous extension $j : X \to \mathbb{R}$ of the restriction of $h - g$ to V. Thus $\widetilde{g} := h - j$ is a continuous extension of the restriction of g to V, and we have $\widetilde{g} \leq h \leq clf \leq f$ as well as $\widetilde{g}(0) = g(0) = f(0)$ and $\nabla \widetilde{g}(0) = \nabla g(0) = x^*$. Therefore, the function $h := \widetilde{g} - x^* - f(0)$ belongs to $\mathcal{H}_X$. Since $f \geq \widetilde{g} = f(0) + x^* + h$, it follows that $(x^*, h) \in \partial_c f(0)$, which shows that x^* belongs to the projection of $\partial_c f(0)$ onto X^*. □

Corollary 13 *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be such that $0 \in dom(f)$. Then $\partial_c f(0) \neq \emptyset$ if, and only if, $\partial_F(0) \neq \emptyset$.*

In particular, from Corollary 13 it follows that if f is Fenchel subdifferentiable or Fréchet differentiable at the origin, then $\partial_c f(0) \neq \emptyset$.

We will now apply this generalized conjugation scheme to duality theory, following the approach of [7, Section 3].

Let S be an arbitrary set, and let $\phi : S \times X \to \overline{\mathbb{R}}$ be a perturbed objective function. The optimization problem under consideration is

$$(\mathcal{P}) \quad \text{minimize } \phi(s, 0); \tag{13}$$

the perturbed problem corresponding to the perturbation parameter $x \in X$ is

$$(\mathcal{P}_x) \quad \text{minimize } \phi(s, x).$$

The associated perturbation function is $p : X \to \overline{\mathbb{R}}$, defined by

$$p(x) := \inf_{s \in S} \phi(s, x).$$

The dual problem to $(\mathcal{P})$ is

$$(\mathcal{D}_c) \quad \text{maximize } -p^c\left(x^*, h\right).$$

A straightforward computation shows that the dual objective function is given by

$$-p^c\left(x^*, h\right) = \inf_{x \in X,\ s \in S} \left\{\phi(s, x) - \langle x, x^* \rangle - h(x)\right\}.$$

From [7, Theorem 6.7] and Corollary 11, the following result follows.

Theorem 14 *The optimal values of $(\mathcal{P})$ and $(\mathcal{D}_c)$ coincide if, and only if, p is lsc at* 0.

The c-Lagrangian function $L_c : S \times X^* \times \mathcal{H}_X \to \overline{\mathbb{R}}$ is given by

$$L_c(s, x^*, h) = -\phi_s^c\left(x^*, h\right),$$

with $\phi_s : X \to \overline{\mathbb{R}}$ denoting the partial mapping $\phi_s(x) := \phi(s, x)$. We thus have

$$L_c(s, x^*, h) = \inf_{x\in X} \left\{\phi(s, x) - \langle x, x^*\rangle - h(x)\right\}.$$

For a full description of generalized convex duality, we refer to [7, Section 3].

4. The second conjugation scheme

In this section, we will present a conjugation scheme different from the one developed in Section 3. It has the advantage of using a simpler 'dual' space, namely, $\mathcal{H}_X$ instead of $X^* \times \mathcal{H}_X$, but its disadvantage is that, unlike in the case of (7), (9), (11) and (12), there is no direct relationship between this new conjugation operator and classical convex conjugation.

We set $d : X \times \mathcal{H}_X \to \mathbb{R}$

$$d(x, h) := h(x)$$

and $d' : \mathcal{H}_X \times X \to \mathbb{R}$

$$d'(h, x) := d(x, h).$$

The d-conjugate function of $f : X \to \overline{\mathbb{R}}$ is $f^d : \mathcal{H}_X \to \overline{\mathbb{R}}$, defined by

$$f^d(h) = \sup_{x\in X} \{h(x) - f(x)\}.$$

We say that $f : X \to \overline{\mathbb{R}}$ is d-subdifferentiable at $x \in dom(f)$ if there exists $h \in \mathcal{H}_X$ such that

$$f(y) - f(x) \geq h(y) - h(x) \quad \text{for all } y \in X.$$

Then h is said to be a d-subgradient of f at x. The set of all d -subgradients of f at x, denoted $\partial_d f(x)$, is called the d-subdifferential of f at x. If $f(x) \notin \mathbb{R}$, we set $\partial_d f(x) := \emptyset$.

The d'-conjugate function $\varphi^{d'} : X \to \overline{\mathbb{R}}$ of $\varphi : \mathcal{H}_X \to \overline{\mathbb{R}}$ is defined by

$$\varphi^{d'}(x) = \sup_{h\in\mathcal{H}_X} \{h(x) - \varphi(h)\}.$$

We say that $\varphi : \mathcal{H}_X \to \overline{\mathbb{R}}$ is d'-subdifferentiable at $h \in dom(\varphi)$ if there exists $x \in X$ such that

$$\varphi(j) - \varphi(h) \geq j(x) - h(x) \quad \text{for all } j \in \mathcal{H}_X.$$

Then x is said to be a d'-subgradient of φ at h. The set of all d'-subgradients of φ at h, denoted $\partial_{d'}\varphi(h)$, is called the d'-subdifferential of φ at h. If $\varphi(h) \notin \mathbb{R}$, we set $\partial_{c'}\varphi(h) := \emptyset$.

The biconjugate of f is $f^{dd'} := (f^d)^{d'}$. The following result is an immediate consequence of [7, Proposition 6.2].

Theorem 15 *Let $f : X \to \mathbb{R} \cup \{+\infty\}$. Then*

$$f^{dd'} = \sup\left\{g : X \to \mathbb{R} | g \textit{ is continuous on } X \textit{ and Fréchet differentiable at } 0,\ \nabla g(0) = 0,\ g \leq f\right\};$$

Corollary 16 *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ and $x \in X$. Then f is lsc at x if, and only if, $f^{dd'}(x) = f(x)$. Moreover, if $x \in dom(f) \setminus \{0\}$ and f is lsc at x, then $\partial_d f(x) \neq \emptyset$.*

We now apply this conjugation scheme to construct a duality theory. We will use the same approach as in Section 3, but with the coupling function d instead of c. So, the dual problem to (13) is now

$$(\mathcal{D}_d) \qquad \text{maximize } -p^d(h).$$

The dual objective function turns out to be

$$-p^d(h) = \inf_{x\in X,\ s\in S} \{\phi(s,x) - h(x)\}.$$

Similarly to Theorem 14, using Corollary 16 we now obtain the following duality theorem.

THEOREM 17 *The optimal values of* $(\mathcal{P})$ *and* $(\mathcal{D}_d)$ *coincide if, and only if,* p *is lsc at* 0.

The d-Lagrangian function $L_d : S \times \mathcal{H}_X \to \overline{\mathbb{R}}$ is given by

$$L_d(s,h) = -\phi_s^d(h).$$

A straightforward computation yields

$$L_d(s,h) = \inf_{x\in X} \{\phi(s,x) - h(x)\}.$$

For more details on the relationship between (P) and (D_d) and the way the d-Lagrangian function links these two optimization problems, we again refer to [7, Section 3].

5. Examples

In this section, we will use two optimization problems to illustrate the duality schemes proposed in the preceding sections. The first one will be an indefinite quadratic problem $(\mathcal{P})$ with linear constraints, in which the perturbation function is lower semi-continuous at 0 and hence the optimal values of $(\mathcal{P})$, $(\mathcal{D}_c)$ and $(\mathcal{D}_d)$ coincide. The second one will be an example given by Duffin [11] in order to show that a duality gap may occur in convex duality when the Slater condition does not hold; since in classical convex duality the absence of a duality gap is equivalent to the lower semi-continuity of the perturbation function at the origin, our dual problems will also exhibit a duality gap in this example.

Example 18 Consider the problem

$$\begin{aligned} &\text{minimize } -s_1^2 + s_2^2 + s_1 s_2 \\ &\text{subject to } s_1 + s_2 \le 1 \\ &\qquad\qquad s_1 - s_2 \le 0 \\ &\qquad\qquad s_1 \ge 0,\ s_2 \ge 0, \end{aligned}$$

and let $\phi : \mathbb{R}^4 \to \mathbb{R} \cup \{+\infty\}$ be the function defined by

$$\phi(s_1,s_2,x_1,x_2) := \begin{cases} -s_1^2 + s_2^2 + s_1 s_2, & \text{if } (s_1,s_2,x_1,x_2) \in \Omega, \\ +\infty, & \text{otherwise}, \end{cases}$$

where

$$\Omega := \left\{(s_1,s_2,x_1,x_2) \in \mathbb{R}_+^4 : s_1 + s_2 + x_1 \le 1,\ s_1 - s_2 + x_2 \le 0,\ x_1 + x_2 \le 1\right\}.$$

Note that

$$(\mathcal{P}) \qquad \text{minimize } \phi\,(s, 0)$$

is equivalent to the given problem. Since, for $(s_1, s_2, x_1, x_2) \in \Omega$, one has

$$\begin{aligned}\phi(s_1, s_2, x_1, x_2) &= -s_1^2 + s_2^2 + s_1 s_2 \geq -s_1^2 + (s_1 + x_2)^2 + s_1(s_1 + x_2)\\ &= s_1^2 + 3x_2 s_1 + x_2^2 \geq x_2^2,\end{aligned}$$

the associated perturbation function p satisfies

$$p(x_1, x_2) \geq x_2^2 = \phi(0, x_2, x_1, x_2) \geq p(x_1, x_2)$$

for every $(x_1, x_2) \in \mathbb{R}^2$ such that $x_1 \geq 0$, $x_2 \geq 0$ and $x_1 + x_2 \leq 1$; hence

$$p(x_1, x_2) = \begin{cases} x_2^2, & \text{if } x_1 \geq 0,\ x_2 \geq 0,\ x_1 + x_2 \leq 1,\\ +\infty, & \text{otherwise.}\end{cases}$$

Therefore, $p(0) = 0$ and, since p is lsc, we conclude that the duality gap must be 0 for both dual problems $(\mathcal{D}_c)$ and $(\mathcal{D}_d)$. Indeed, in the case of $(\mathcal{D}_c)$, the pair $(0, h_0)$, with $h_0(x_1, x_2) := x_2^2$, is an optimal solution, since $-p^c\,(0, h_0) = \inf\left\{\phi(s_1, s_2, x_1, x_2) - x_2^2\right\} = 0$ (the infimum is attained, e.g. at $(s_1, s_2, x_1, x_2) := (0, 0, 0, 0)$), whereas in the case of $(\mathcal{D}_d)$, the same function h_0 is an optimal solution, since we also have $-p^d\,(h_0) = -p^c\,(0, h_0) = 0$.

Note that every continuous function $r : \mathbb{R}^2 \to \mathbb{R}^2$ satisfying $\langle x, r(x)\rangle = h_0(x)$ for all $x \in \mathbb{R}^2$ is an optimal solution to the dual problem corresponding to the coupling function considered in [1] (defined following the standard duality scheme of [7, Section 3]). One can take, for instance, $r(x_1, x_2) := (0, x_2)$, but, clearly, for any continuous mapping $w : \mathbb{R}^2 \to \mathbb{R}^2$ satisfying $\langle x, w(x)\rangle = 0$ for all $x \in \mathbb{R}^2$, the function $r + w$ is an optimal solution, too. Thus, by means of this example we also illustrate the fact that each optimal solution of either $(\mathcal{D}_c)$ or $(\mathcal{D}_d)$ is associated, in a natural way, to infinitely many optimal solutions to the dual problem arising according to the conjugation scheme in [1]; therefore, the optimal solution sets (as well as the feasible sets) of our dual problems are substantially smaller than those of the equivalent (from the viewpoint of optimal values) dual problem arising from [1].

Example 19 [11, Section 3] Consider the problem

$$\begin{aligned}&\text{minimize} \quad e^{s_2}\\ &\text{subject to} \quad \sqrt{s_1^2 + s_2^2} \leq s_1,\end{aligned}$$

and let $\phi : \mathbb{R}^3 \to \mathbb{R} \cup \{+\infty\}$ be the function defined by

$$\phi(s_1, s_2, x) := \begin{cases} e^{s_2}, & \text{if } \sqrt{s_1^2 + s_2^2} \leq s_1 - x,\\ +\infty, & \text{otherwise;}\end{cases}$$

clearly,

$$(\mathcal{P}) \qquad \text{minimize } \phi\,(s, 0)$$

is equivalent to the given problem. One can easily check that the associated perturbation function p is given by

$$p(x) = \begin{cases} 1, & \text{if } x = 0 \\ 0, & \text{if } x < 0 \\ +\infty, & \text{if } x > 0. \end{cases}$$

Therefore, $p(0) = 1$ and, since p is not lsc at 0, there exists a strictly positive duality gap for both dual problems $(\mathcal{D}_c)$ and $(\mathcal{D}_d)$. We shall verify that the optimal values of $(\mathcal{D}_c)$ and $(\mathcal{D}_d)$ are both 0, thus showing that the duality gap is equal to 1. Indeed, for every $(x^*, h) \in \mathbb{R} \times \mathcal{H}_{\mathbb{R}}$ we have

$$\begin{aligned} -p^c(x^*, h) &= \inf_{x\in\mathbb{R}} \left\{ p(x) - xx^* - h(x) \right\} \leq \inf_{x<0} \left\{ p(x) - xx^* - h(x) \right\} \\ &= \inf_{x<0} \left\{ -xx^* - h(x) \right\} \leq 0, \end{aligned}$$

the last inequality being a consequence of the continuity of h at 0; moreover,

$$-p^d(h) = -p^c(0, h) \leq 0.$$

Since, for $h_0 \equiv 0$, we have

$$-p^d(h_0) = -p^c(0, h_0) = \inf_{x\in\mathbb{R}} p(x) = 0,$$

we conclude that $(0, h_0)$ and h_0 are optimal solutions to $(\mathcal{D}_c)$ and $(\mathcal{D}_d)$, respectively, and that the optimal values of these dual problems are both 0.

Acknowledgements

We are grateful to Wilfredo Sosa and two anonymous referees for their valuable comments and suggestions. The author gratefully acknowledges the financial support received from CAPES and the warm hospitality of the second author as well as of the members from the host institution.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by CAPES [grant number BEX 3843/14-9]; MICINN of Spain [grant number MTM2011-29064-C03-01]; Australian Research Council's Discovery Projects funding scheme [project number DP140103213].

ORCID

Martínez-Legaz http://orcid.org/0000-0002-6845-6202

References

[1] Cotrina J, Karas EW, Ribeiro AA, Sosa W, Yun YJ. Moreau–Fenchel conjugation for lower semi-continuous functions. Optimization. 2011;60:1045–1057.

[2] Leonard IE, Sundaresan K. Geometry of Lebesgue-Bochner function spaces – smoothness. Trans. Amer. Math. Soc. 1974;198:229–251.

[3] Phelps RR. Convex functions, monotone operators and differentiability. Berlin: Springer; 1993.
[4] Stromberg KR. Introduction to classical real analysis. Belmont, California: Wadsworth International; 1981.
[5] Penot J-P. Conjugacies adapted to lower semicontinuous functions. Optimization. 2015;64:473–494.
[6] Sosa W. Representation of continuous functions and its applications. J. Optim. Theory Appl. 2013;159:795–804.
[7] Martínez-Legaz JE. Generalized convex duality and its economic applications. In: Hadjisavvas N, Komlósi S, Schaible S, editor. Handbook of generalized convexity and generalized monotonicity. New York (NY): Springer; 2005. p. 237–292. (Nonconvex optimization and and its applications; vol. 76).
[8] Rubinov A. Abstract convexity and global optimization. Dordrecht, Boston: Kluwer Academic; 2000.
[9] Singer I. Abstract convex analysis. New York (NY): Wiley; 1997.
[10] Rockafellar RT, Wets RJ-B. Variational analysis. Berlin: Springer; 1998.
[11] Duffin RJ. Convex analysis treated by linear programming. Math. Program. 1973;4:125–143.

Apêndice C

Artigo: *A generalization of the strong Fitzpatrick inequality*

Taylor & Francis
Taylor & Francis Group

Optimization

A Journal of Mathematical Programming and Operations Research

ISSN: 0233-1934 (Print) 1029-4945 (Online) Journal homepage: http://www.tandfonline.com/loi/gopt20

A generalization of the strong Fitzpatrick inequality

Leonardo M. Elias & Juan E. Martínez-Legaz

To cite this article: Leonardo M. Elias & Juan E. Martínez-Legaz (2016): A generalization of the strong Fitzpatrick inequality, Optimization, DOI: 10.1080/02331934.2016.1179738

To link to this article: http://dx.doi.org/10.1080/02331934.2016.1179738

OPTIMIZATION, 2016
http://dx.doi.org/10.1080/02331934.2016.1179738

Taylor & Francis
Taylor & Francis Group

A generalization of the strong Fitzpatrick inequality

Leonardo M. Elias[a,b] and Juan E. Martínez-Legaz[c,d,e]

[a]Programa de Pós-Graduação em Matemática, Universidade Federal do Paraná, Curitiba, Brazil; [b]CAPES Foundation, Ministry of Education of Brazil, Brasília, Brazil; [c]Departament d'Economia i d'Història Econòmica, Universitat Autònoma de Barcelona, Bellaterra, Spain; [d]BGSMath, Barcelona, Spain; [e]MOVE (Markets, Organizations and Votes in Economics), Bellaterra (Cerdanyola del Vallès), Spain

ABSTRACT
We present a generalization of the strong Fitzpatrick inequality in the context of reflexive Banach spaces, involving a twisted bigger conjugate function. We also introduce a related family of gap functions for maximal monotone inclusion problems.

ARTICLE HISTORY
Received 29 September 2015
Accepted 12 April 2016

KEYWORDS
Strong Fitzpatrick inequality; gap functions; maximal monotone operators; twisted bigger conjugate functions; Tykhonov well-posed functions

1. Introduction

The main result of this mote is a generalization of the strong Fitzpatrick inequality ([1, Theorem 4], [2, Theorem 9.7.2]) in the context of reflexive Banach spaces. In our generalized inequality, a twisted bigger conjugate function,[3, Definition 19.14] defined on the product of the space with its dual plays the role that the norm on this product plays in the case of the classical strong inequality. Related to the new inequality, we introduce a new family of gap functions, parameterized by Tykhonov well-posed twisted bigger conjugate functions, for a general maximal monotone inclusion problem (MIP), that is, the problem of finding a zero of a maximal monotone operator. By means of a gap function one can reformulate the maximal MIP as a convex optimization problem. We compare our new family with a gap function recently proposed in [4].

The rest of the paper is organized as follows. Section 2 deals with the fundamental notions on maximal monotone operators and Fitzpatrick functions that are used in the paper. In Section 3, we present our generalization of the strong Fitzpatrick inequality. Section 4 introduces a related family of gap functions for maximal MIPs.

2. Preliminaries

Let $(X, \|\cdot\|)$ be a real Banach space with dual $(X^*, \|\cdot\|_*)$. We will denote by $\langle\cdot,\cdot\rangle : X \times X^* \to \mathbb{R}$ the duality pairing between these spaces. The duality pairing between the space $X \times X^*$ and $X^* \times X$ is the function on $(X \times X^*) \times X^* \times X$, also denoted by $\langle\cdot,\cdot\rangle$, given by

$$\langle (x, x^*), (y^*, y)\rangle = \langle x, y^*\rangle + \langle y, x^*\rangle . \tag{1}$$

We recall that $X^* \times X$ can be canonically identified with a subspace of the dual space $(X \times X^*)^* = X^* \times X^{**}$ of $X \times X^*$; under this identification, one has $(X \times X^*)^* = X^* \times X$ if and only if the space

CONTACT Juan E. Martínez-Legaz JuanEnrique.Martinez.Legaz@uab.cat

is reflexive, that is, if X is canonically identified with the whole of X^{**}. These and other basic facts of functional analysis used in this paper can be found, e.g. in [5].

Let $T : X \rightrightarrows X^*$ be a set-valued operator. The graph and the domain of T are given, respectively, by

$$gph\, T := \{(x, x^*) \in X \times X^* \mid x^* \in T(x)\}$$

and

$$dom\, T := \{x \in X \mid T(x) \neq \emptyset\}.$$

Recall that T is said to be monotone if

$$\langle x - y, x^* - y^* \rangle \geq 0$$

whenever $x, y \in X$ and $(x, x^*), (y, y^*) \in gph\, T$.

A monotone map $T : X \rightrightarrows X^*$ is said to be maximal monotone if there is no monotone map whose graph properly contains the graph of T. The Fitzpatrick function associated with a maximal monotone operator T is the lsc convex extended real-valued function on $X \times X^*$ defined by

$$F_T(x, x^*) := \sup_{(y,y^*)\in gph\, T} \{\langle x - y, y^* \rangle + \langle y, x^* \rangle\}.$$

Equivalently,

$$F_T(x, x^*) = \langle x, x^* \rangle - \inf_{(y,y^*)\in gph\, T} \langle x - y, x^* - y^* \rangle . \tag{2}$$

Since T is maximal monotone, we can easily see that the Fitzpatrick function satisfies

$$F_T(x, x^*) \geq \langle x, x^* \rangle \quad \text{for all } (x, x^*) \in X \times X^*, \tag{3}$$

with equality if and only if $(x, x^*) \in gph\, T$.

Given an lsc proper convex function $f : X \to \mathbb{R} \cup \{+\infty\}$, its Fenchel conjugate $f^* : X^* \to \mathbb{R} \cup \{+\infty\}$ is defined by

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

Therefore, in view of (1), the restriction of the Fenchel–Moreau conjugate of $g : X \times X^* \to \mathbb{R} \cup \{+\infty\}$ to $X^* \times X$ is given by

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

Note that the conjugate F_T^* of the Fitzpatrick function F_T associated with a maximal monotone operator T satisfies

$$\begin{aligned} F_T^*(x^*, x) &\geq \sup_{(y,y^*)\in gra\, T} \{\langle (y, y^*), (x, x^*) \rangle - F_T(y, y^*)\} \\ &\geq \sup_{(y,y^*)\in gra\, T} \{\langle (y, y^*), (x, x^*) \rangle - \langle y, y^* \rangle\} \\ &= F_T(x, x^*) \end{aligned} \tag{4}$$

for all $(x, x^*) \in X \times X^*$.

Theorem 1: (see [2, Exercise 9.23]) *Let $T : X \rightrightarrows X^*$ be maximal monotone. Then, for every $(x, x^*), (w, w^*) \in X \times X^*$, one has*

$$F_T(x, x^*) - \langle x, x^* \rangle + F_T(w, w^*) - \langle w, w^* \rangle \geq -\frac{1}{2} \langle x - w, x^* - w^* \rangle .$$

Proof: Using the convexity of the Fitzpatrick function and the fact that it is bounded below by the duality product, we obtain

$$\begin{aligned}
&F_T(x, x^*) - \langle x, x^* \rangle + F_T(w, w^*) - \langle w, w^* \rangle \\
&= 2 \left(\frac{1}{2} F_T(x, x^*) + \frac{1}{2} F_T(w, w^*) \right) - \langle x, x^* \rangle - \langle w, w^* \rangle \\
&\geq 2F_T \left(\frac{1}{2} (x + w), \frac{1}{2} (x^* + w^*) \right) - \langle x, x^* \rangle - \langle w, w^* \rangle \\
&= -\frac{1}{2} \langle x - w, x^* - w^* \rangle .
\end{aligned}$$

□

We conclude this section enunciating a version of Fenchel duality theorem. It will play an essential role in the proof of the main result of this work. We recall that the domain of a function $f : X \to \mathbb{R} \cup \{+\infty\}$ is the set $dom\, f := f^{-1} (\mathbb{R})$.

Theorem 2: (see [2, Theorem 4.4.18] *Let $f, g : X \to \mathbb{R} \cup \{+\infty\}$ be convex and such that $0 \in int \left(dom\, f - dom\, g \right)$, where int denotes interior. Then*

$$\inf_{x \in X} \{f(x) + g(x)\} = \sup_{x^* \in X^*} \{-f^*(x^*) - g^*(- x^*)\} \tag{5}$$

and the supremum in (5) is attained if finite.

3. Main result

We say that an lsc proper convex function $g : X \times X^* \to \mathbb{R} \cup \{+\infty\}$ is a twisted bigger conjugate function (TBC-function in short) [3, Definition 19.14] if

$$- \langle x, x^* \rangle \leq g(x, x^*) \leq g^*(- x^*, -x) \tag{6}$$

for all $(x, x^*) \in X \times X^*$.

Example 3: The function $g : X \times X^* \to \mathbb{R} \cup \{+\infty\}$ defined by $g(x, x^*) := f(x) + f^*(- x^*)$, where $f : X \to \mathbb{R} \cup \{+\infty\}$ is an lsc proper convex function, is a TBC-function. In particular, so is the function g defined by $g(x, x^*) := \frac{1}{p} \|x\|^p + \frac{1}{q} \|x^*\|_*^q$, where $p \geq 1$ and $\frac{1}{p} + \frac{1}{q} = 1$.

The folowing proposition will be useful in the next section.

Proposition 4: *Let $g : X \times X^* \to \mathbb{R} \cup \{+\infty\}$ be a non-negative TBC-function. Then*

$$\min_{(x,x^*) \in X \times X^*} g(x, x^*) = g \left(0, 0 \right) = 0. \tag{7}$$

Proof: By (6) and the non-negativity of g, we have

$$0 \leq g \left(0, 0 \right) \leq g^* \left(0, 0 \right) = - \inf_{(x,x^*) \in X \times X^*} g(x, x^*) \leq 0.$$

□

Theorem 5: *Let X be reflexive, $T : X \rightrightarrows X^*$ be maximal monotone, and $g : X \times X^* \to \mathbb{R} \cup \{+\infty\}$ be a TBC-function. Consider the following conditions:*

(i) *$dom\, F_T = X \times X^*$,*
(ii) *$dom\, g = X \times X^*$.*

If any one of them holds then, for every $(x, x^*) \in X \times X^*$, one has

$$F_T(x, x^*) - \langle x, x^* \rangle \geq \frac{1}{2} \inf_{(w,w^*) \in gph\, T} g(w - x, w^* - x^*),$$

with strict inequality whenever the infimum is not attained.

Proof: For $(x, x^*) \in X \times X^*$, consider the maximal monotone operator $L : X \rightrightarrows X^*$ defined by $L(y) := T(y + x) - x^*$. Since

$$gphL = gphT - (x, x^*), \tag{8}$$

if (i) holds then $dom\, F_L = X \times X^*$. Thus, if one of the conditions (i) or (ii) holds, we have $dom\, F_L - dom\, g = X \times X^*$, and, hence, the convex functions F_L and g satisfy the assumptions of Theorem 2. Therefore, since $F_L(x, x^*) \geq \langle x, x^* \rangle$ and $g(x, x^*) \geq - \langle x, x^* \rangle$ for every $(x, x^*) \in X \times X^*$, we have

$$\begin{aligned} 0 &\leq \inf_{(x,x^*) \in X \times X^*} \left\{ F_L(x, x^*) + g(x, x^*) \right\} \\ &= \sup_{(y^*,y) \in X^* \times X} \left\{ -F_L^*(y^*, y) - g^*(-y^*, -y) \right\} \\ &\leq \sup_{(y^*,y) \in X^* \times X} \left\{ -F_L^*(y^*, y) + \langle y, y^* \rangle \right\} \\ &\leq \sup_{(y^*,y) \in X^* \times X} \left\{ -F_L(y, y^*) + \langle y, y^* \rangle \right\} \leq 0, \end{aligned} \tag{9}$$

the last two inequalities following from (4) and (3). Thereby,

$$\max_{(y^*,y) \in X^* \times X} \left\{ -F_L^*(y^*, y) - g^*(-y^*, -y) \right\} = 0.$$

Thus, there exists $(y^*, y) \in X^* \times X$ such that $F_L^*(y^*, y) + g^*(-y^*, -y) = 0$. By (4), we obtain that

$$F_L(y, y^*) + g^*(-y^*, -y) \leq 0.$$

So, since g is a TBC-function, we conclude that

$$F_L(y, y^*) + g(y, y^*) \leq 0$$

and thus, in view of the first inequality in (9), $F_L(y, y^*) + g(y, y^*) = 0$ and (y, y^*) is a minimizer of $F_L + g$. Since

$$0 = F_L(y, y^*) + g(y, y^*) \geq F_L(y, y^*) - \langle y, y^* \rangle \geq 0,$$

we have

$$-g(y, y^*) = F_L(y, y^*) = \langle y, y^* \rangle.$$

Thus, $(y, y^*) \in gph\, L$ and, setting $(w, w^*) := (x + y, x^* + y^*)$, by (8) we obtain that $(w, w^*) \in gph\, T$ and $-g(w - x, w^* - x^*) = \langle w - x, w^* - x^* \rangle$, and from Theorem 1, we conclude that

$$\begin{aligned} F_T(x, x^*) - \langle x, x^* \rangle &= F_T(w, w^*) - \langle w, w^* \rangle + F_T(x, x^*) - \langle x, x^* \rangle \\ &\geq -\frac{1}{2} \langle w - x, w^* - x^* \rangle = \frac{1}{2} g(w - x, w^* - x^*), \end{aligned}$$

which ends the proof. □

Since the function g of Example 3 is a TBC-function and its domain is $X \times X^*$, applying Theorem 5 to this case we obtain the following result:

Corollary 6: *Let X be reflexive, $T : X \rightrightarrows X^*$ be maximal monotone, and $f : X \to \mathbb{R} \cup \{+\infty\}$ be convex, proper and lsc. Consider the following conditions:*

(i) *$dom\, F_T = X \times X^*$,*
(ii) *$dom f = X$ and $dom f^* = X^*$.*

If any one of them holds then, for every $(x, x^*) \in X \times X^*$, one has

$$F_T(x,x^*) - \langle x, x^* \rangle \geq \frac{1}{2} \inf_{(w,w^*) \in gph\, T} \left\{ f\left(w - x\right) + f^*\left(x^* - w^*\right) \right\},$$

with strict inequality whenever the infimum is not attained.

In particular, for every $p \geq 1$ and q such that $\frac{1}{p} + \frac{1}{q} = 1$, one has

$$F_T(x,x^*) - \langle x, x^* \rangle \geq \frac{1}{2} \inf_{(w,w^*) \in gph\, T} \left\{ \frac{1}{p} \|w - x\|^p + \frac{1}{q} \left\| w^* - x^* \right\|_*^q \right\}$$

for all $(x, x^*) \in X \times X^*$, with strict inequality whenever the infimum is not attained.

Concerning the condition *dom* $f^* = X^*$ in (ii) of Corollary 6, we recall that a sufficient (and necessary in the finite-dimensional case) condition for it to hold is f to be supercoercive [6, Theorem 3.4]:

$$\lim_{\|x\| \longrightarrow +\infty} \frac{f\left(x\right)}{\|x\|} = +\infty.$$

Setting $p = 2$ in Corollary 6, we obtain the following known result:

Corollary 7: (Strong Fitzpatrick Inequality) [1, Theorem 4] (see also [2, Theorem 9.7.2]) *Let X be reflexive and $T : X \rightrightarrows X^*$ be maximal monotone. Then, for every $(x, x^*) \in X \times X^*$, one has*

$$F_T(x,x^*) - \langle x, x^* \rangle \geq \frac{1}{4} \inf_{(w,w^*) \in gph\, T} \left\{ \|w - x\|^2 + \left\| w^* - x^* \right\|_*^2 \right\}.$$

4. A new family of gap functions

In this section, we shall consider the so called maximal MIP [4,7]: Given a maximal monotone operator $T : X \rightrightarrows X^*$, find a point $x \in X$ such that

$$0 \in T(x).$$

The (possibly empty) solution set of MIP is $T^{-1}(0) = \{x \in X \mid 0 \in T(x)\}$.

A gap function for MIP is a function $\varphi : X \to \mathbb{R} \cup \{+\infty\}$ satisfying the following conditions:

(i) $\varphi(x) \geq 0$ for all $x \in X$.
(ii) $\varphi(x) = 0$ if and only if $x \in T^{-1}(0)$.

Thereby, if the infimal value of φ is 0 and it is attained, the minimizers of φ are exactly the elements in the solution set of MIP. Thus, we can reformulate MIP as the optimization problem consisting in minimizing φ. The following result has an immediate proof.

Proposition 8: (see [4, Theorem 2.1]) *Let φ be a gap function for MIP. If MIP has a solution, then* $\min_{x \in X} \varphi\left(x\right) = 0$. *Conversely, if X is reflexive, $\inf_{x \in X} \varphi\left(x\right) = 0$, and φ is weakly lsc (in particular, if it is convex and lsc) and weakly coercive in the sense that*

$$\lim_{\|x\| \to \infty} \varphi(x) = +\infty$$

then MIP has a solution.

In [4], Borwein and Dutta present a gap function G_T for MIP, associated with the Fitzpatrick function F_T of T. It is defined by $G_T(x) := F_T(x, 0)$, or, more explicitly,

$$G_T(x) = \sup_{(y,y^*)\in gph\, T} \langle x - y, y^* \rangle .$$

It is easy to see that G_T is indeed a gap function, and Borwein and Dutta found that it has good properties.

For $g : X \times X^* \to \mathbb{R} \cup \{+\infty\}$, we define $G_{T,g} : X \to \mathbb{R} \cup \{+\infty\}$ by

$$G_{T,g}(x) := \frac{1}{2} \inf_{(w,w^*)\in gph\, T} g(w - x, w^*).$$

Note that $G_{T,g}$ is convex whenever g is convex, since in this case the function $(x, w, w^*) \mapsto g(w - x, w^*)$ is convex, jointly in its three arguments. Moreover, it is proper as long as g is finite at some $(w, w^*) \in gph\, T$.

The following proposition is an immediate consequence of Theorem 5.

Proposition 9: *Under the assumptions of Theorem 5, one has $G_T \geq G_{T,g}$, and, hence, $G_{T,g}$ is finite whenever G_T is finite.*

We refer to [4] for conditions on T ensuring that G_T is finite.

Definition 10: [8, Definition 10.1.1] Let E be a metric space. An lsc function $\varphi : E \to \mathbb{R} \cup \{+\infty\}$ is said to be Tykhonov well-posed if it satisfies the following conditions:

(i) It has a unique global minimizer $\overline{x}$.
(ii) Every sequence x_n such that $\varphi(x_n) \to \varphi(\overline{x})$ satisfies $x_n \to \overline{x}$.

Example 11: [8, Example 10.1.4] If X is a finite dimensional vector space, then every lsc proper convex function $f : X \to \mathbb{R} \cup \{+\infty\}$ with a unique minimum point is Tykhonov well-posed. In particular, so is every lsc strictly convex function defined on X.

Example 12: A function $f : X \to \mathbb{R} \cup \{+\infty\}$ is said to be strongly convex if there exists $\gamma > 0$ such that, for every $x, y \in X$ and $\alpha, \beta \geq 0$ with $\alpha + \beta = 1$, one has

$$f(\alpha x + \beta y) \leq \alpha f(x) + \beta f(y) - \gamma\alpha\beta \|x - y\|^2. \tag{10}$$

If f is strongly convex and has a minimizer, then it is Tykhonov well-posed. Indeed, it is well known that the minimizer $\overline{x}$, if it exists, must be unique. Notice that, if X is reflexive, every lsc strongly convex function has a minimizer, as it is weakly coercive. Let x_n be a sequence satisfying $f(x_n) \to f(\overline{x})$. Setting $x := x_n$, $y := \overline{x}$ and $\alpha = \beta := \frac{1}{2}$ in (10), we get

$$\begin{aligned} 0 \leq \|x_n - \overline{x}\| &\leq \sqrt{\frac{2}{\gamma}\left(f(x_n) + f(\overline{x}) - 2f\left(\frac{1}{2}(x_n + \overline{x})\right)\right)} \\ &\leq \sqrt{\frac{2}{\gamma}\left(f(x_n) - f(\overline{x})\right)} \to 0, \end{aligned}$$

which shows that $x_n \to \overline{x}$.

For more examples and references about Tykhonov well-posedness, see [8,9].

Theorem 13: *Let $T : X \rightrightarrows X^*$ be maximal monotone and $g : X \times X^* \to \mathbb{R} \cup \{+\infty\}$ be a Tykhonov well-posed function satisfying (7). Then $G_{T,g}$ is a gap function for MIP.*

Proof: Note that $G_{T,g}(x) \geq 0$ for all $x \in X$ since g is non-negative. Moreover, if $G_{T,g}(x) = 0$ for some $x \in X$ then there exist $(w_n, w_n^*) \subset gph\, T$ such that $g(w_n - x, w_n^*) \longrightarrow 0$. Since g is Tykhonov well-posed, from (7) it follows that $w_n \longrightarrow x$ and $w_n^* \longrightarrow 0$. Since $gph\, T$ is closed, we deduce that $(x, 0) \in gph\, T$, that is, $0 \in T\left(x\right)$. Conversely, if $0 \in T\left(x\right)$ then, using that $(x, 0) \in gph\, T$ and (7), we obtain $0 \leq G_{T,g}(x) \leq g(0, 0) = 0$, concluding the proof. □

Remark: For x and g as in Theorem 13, if there exist $(w, w^*) \in gph\, T$ such that $g(w - x, w^*) = 0$, then $G_{T,g}(x) = 0$ and, hence, x is a solution to MIP.

Corollary 14: *Let $T : X \rightrightarrows X^*$ be maximal monotone and $g : X \times X^* \to \mathbb{R} \cup \{+\infty\}$ be a non-negative Tykhonov well-posed TBC-function. Then, $G_{T,g}$ is a gap function for MIP.*

Proof: Combine Theorem 13 with Proposition 4. □

Acknowledgements

We are very grateful to an anonymous referee, whose useful comments have helped us to improve the presentation. This author acknowledges financial support from the Spanish Ministry of Economy and Competitiveness, through Grant MTM2014-59179-C2-2-P and the Severo Ochoa Programme for Centres of Excellence in R&D (SEV-2015-0563), and under Australian Research Council's Discovery Projects funding scheme (project number DP140103213).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by CAPES [grant number BEX 3843/14-9]; the Spanish Ministry of Economy and Competitiveness [grant number MTM2014-59179-C2-2-P]; the Severo Ochoa Programme for Centres of Excellence in R&D [SEV-2015-0563]; Australian Research Council's Discovery Projects funding scheme [project number DP140103213].

ORCID

Juan E. Martínez-Legazc http://orcid.org/0000-0002-6845-6202

References

[1] Voisei MD, Zalinescu C. Strongly-representable monotone operators. J. Convex Anal. 2009;16:1011–1033.
[2] Borwein JM, Vanderwerff JD. Convex functions: constructions, characterizations and counterexamples. Cambridge: Cambridge University Press; 2010.
[3] Simons S. From Hahn–Banach to monotonicity. Berlin: Springer; 2008.
[4] Borwein J, Dutta J. Maximal monotone inclusions and Fitzpatrick functions. J. Optim. Theory Appl. 2015. doi:10.1007/s10957-015-0813-x. First online: 30 September.
[5] Brezis H. Functional analysis, sobolev spaces and partial differential equations. New York (NY): Springer; 2011.
[6] Bauschke HH, Borwein JM, Combettes PL. Essential smoothness, essential strict convexity, and Legendre functions in Banach spaces. Commun. Contemp. Math. 2001;3:615–647.
[7] Borwein JM, Lewis A. Convex analysis and nonlinear optmization: theory and examples. New York (NY): Springer; 2000.
[8] Lucchetti R. Convexity and well-posed problems. New York (NY): Springer; 2006.
[9] Dontchev AL. Well-posed optimization problems. Vol. 1543, Lecture notes in mathematics. Berlin: Springer; 1993.