

LEONARDO MORETO ELIAS

UMA BASE TEÓRICA PARA A CONJUGAÇÃO DE
FUNÇÕES SEMICONTÍNUAS INFERIORMENTE

LEONARDO MORETO ELIAS

UMA BASE TEÓRICA PARA A CONJUGAÇÃO DE
FUNÇÕES SEMICONTÍNUAS INFERIORMENTE

Dissertação apresentada ao Programa de Pós-Graduação em Matemática Aplicada da Universidade Federal do Paraná, como requisito parcial à obtenção do grau de Mestre em Matemática Aplicada.

Orientador: Prof. Dr. Ademir Alves Ribeiro.
Coorientador: Prof. Dr. Wilfredo Sosa Sandoval.

Curitiba
Fevereiro de 2013

UNIVERSIDADE FEDERAL DO PARANÁ
LEONARDO MORETO ELIAS

UMA BASE TEÓRICA PARA A CONJUGAÇÃO DE
FUNÇÕES SEMICONTÍNUAS INFERIORMENTE

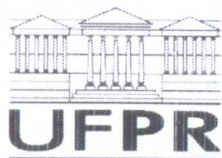
Prof. Dr. Ademir Alves Ribeiro
Orientador - Universidade Federal do Paraná - UFPR

Banca:

Prof. Dr. Wilfredo Sandoval Sosa
Coorientador - Universidade Católica de Brasília

Prof.ª Dr.ª Fernanda Raupp
Pontifícia Universidade Católica do Rio de Janeiro

Prof. Dr. Raimundo José Borges de Sampaio
Pontifícia Universidade Católica do Paraná



Ministério da Educação
Universidade Federal do Paraná
Setor de Ciências Exatas/Departamento de Matemática
Programa de Pós-Graduação em Matemática Aplicada - PPGMA

PARECER DA BANCA EXAMINADORA

Após a apresentação, a banca deliberou pela aprovação da dissertação do candidato **Leonardo Moreto Elias** devendo para tanto incorporar as sugestões feitas pelos membros da banca, no prazo estabelecido pelo regimento correspondente.

Curitiba, 28 de fevereiro de 2013.

Prof. Dr. Wilfredo Sandoval Sosa
Presidente

Profa. Dra. Fernanda Raupp
Titular

Prof. Dr. Raimundo José Borges de Sampaio
Membro

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 amigos que tanto me apoiaram e me incentivaram.

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

Aos membros da banca, professora Fernanda e professor
Raimundo. 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, 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

Nesta dissertação, apresenta-se um estudo a respeito de uma generalização das funções conjugadas de Fenchel para funções reais estendidas semicontínuas inferiormente (sci) de várias variáveis reais. Considera-se para esta, o produto interno generalizado pelas funções contínuas de \mathbb{R}^n em \mathbb{R}^n e a base teórica é a generalização dos teoremas clássicos de separação para convexos, a qual possibilita separar por funções contínuas conjuntos fechados e, em particular, garante que o epigrafo de uma função sci pode ser separado de qualquer ponto que esteja em seu complementar. Com auxílio desses resultados, verifica-se que, para funções sci, a conjugada proposta é própria e a conjugação é simétrica. Prova-se também que esta conjugada é convexa e sci, e introduz-se os Espaços Duais Conjugados que aumentam o potencial dessa teoria, pois, dependendo da função correspondente, eles podem ser de dimensão finita. Aplica-se esta generalização no desenvolvimento de uma dualidade para problemas de programação semicontínua inferior (*PSCI*). Garante-se que o dual desses problemas é de programação convexa e dependendo da função, pode ser restrito a um espaço dual conjugado de dimensão finita. Prova-se que o bidual é o próprio *PSCI* e define-se a função Lagrangeana relacionada para concluir que seu minimizador em \mathbb{R}^n é solução do problema primal e a função contínua que o maximiza é solução do problema dual.

Palavras-chave: teoremas de separação de conjuntos fechados; funções conjugadas; programação semicontínua.

Abstract

We present a study about a generalization of Fenchel conjugate functions for extended real-valued lower semicontinuous (lsc) functions of several real variables. For this, we consider the generalized inner product by continuous functions $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ and the theoretical basis is the generalization of the classical separation theorems for convex sets. They ensure the existence of continuous functions that separate two closed sets. In particular, they say that the epigraph of a lsc function can be separated from any point that it is in its complement. Thanks to these results one verifies that for lsc functions the proposed conjugate function is proper and its conjugation is symmetric. We also prove that this conjugate function is convex and lsc. We introduce the Conjugate Dual Spaces that increase the power of this theory as they can be of finite dimension in some cases. We apply this generalization to build up a duality scheme for lower semicontinuous programming (*LSCP*). We ensure that its associated dual problem is a convex programming problem and it can be restricted to a Conjugate Dual Space of finite dimension. We prove the bidual is the *LSCP* itself and present the corresponding Lagrangian function which minimizer in \mathbb{R}^n is the solution of the primal problem and the continuous function that maximizes it is the solution of the dual problem.

Keywords: separation theorems for closed sets; conjugate functions; semi-continuous programming.

Sumário

Introdução	1
1 Revisão de conceitos	3
1.1 Conceitos iniciais	3
1.2 Funções semicontínuas	6
1.3 Conjuntos de recessão	13
1.4 Separação de convexos	16
1.4.1 Projeção em conjuntos convexos e fechados	18
1.4.2 Teoremas de Separação	20
1.5 Funções conjugadas	23
1.6 Dualidade convexa	30
2 Separação de fechados	34
2.1 Uma versão simples do teorema de Seleção de Michael	34
2.2 Considerações iniciais	36
2.3 Teorema de separação para fechados	42
3 Função conjugada modificada	48
3.1 A extensão de Moreau	48
3.2 Função conjugada para funções sci	49
3.3 Espaços duais conjugados	55
4 Programação semicontínua	58
4.1 O problema <i>PSCI</i>	58
4.2 Dualidade semicontínua	61
4.3 A função Lagrangeana	63
Conclusão	66
Referências Bibliográficas	67

Introdução

A teoria de conjugação desenvolvida por Fenchel [11] associa a cada função $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função convexa dada por

$$f_F^*(a) = \sup_{x \in \mathbb{R}^n} \{\langle a, x \rangle - f(x)\}$$

com $a \in \mathbb{R}^n$, chamada de função conjugada. Esta relação possui uma importante interpretação econômica: considerando x a quantidade de produção de uma determinada empresa, $f(x)$ o custo para produzir essa quantidade e a 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. Um exemplo é que 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.

Moreau [20] foi um dos primeiros a generalizar a teoria de conjugação. Ele considerou funções $c : K_1 \times K_2 \rightarrow \mathbb{R} \cup \{+\infty\}$ em que K_1 e K_2 são conjuntos arbitrários e assim, estendeu a conjugada como

$$f^c(a) = \sup_{x \in K_2} \{c(a, x) - f(x)\}$$

com $a \in K_1$. Cotrina, Karas, Ribeiro, Sosa e Yuan [6] procuraram resolver o problema da modelagem apresentando uma conjugada na qual c é o produto interno generalizado, ou seja, $c(p, x) = \langle p(x), x \rangle$ em que $p : \mathbb{R}^n \rightarrow \mathbb{R}^n$ é uma função contínua.

Estes autores tinham o interesse em aplicar esta conjugada modificada para construir o dual de problemas de programação semicontínua. Os problemas primais associados tratam de encontrar minimizadores de funções semicontínuas inferiormente que, dependendo de certas condições de coercividade, possuem solução. No entanto, a função a ser minimizada pode não possuir um bom comportamento como o caso das funções que não são convexas. Assim, é muito difícil encontrar a solução de forma analítica ou numérica, motivando a investigar um problema dual associado que seja mais fácil de ser resolvido.

A teoria de conjugação de Fenchel foi usada como base para construir a Dualidade Convexa em [19], mas esta dualidade possui boas características para funções convexas semicontínuas inferiormente. O propósito de Cotrina, Karas, Ribeiro, Sosa e Yuan era, através de sua generalização, enfraquecer estas hipóteses para que fosse possível construir uma dualidade bem comportada para funções apenas semicontínuas inferiormente.

Em [6], foi também desenvolvido um estudo a respeito dos espaços duais conjugados que são subespaços das funções contínuas em que a conjugação modificada possui simetria. Esta análise é importante tanto na parte teórica quanto na prática, pois para muitos fins, é necessário o cálculo da conjugada apenas nestes subespaços.

O termo “generalização” dado a conjugada em [6] causou um desconforto entre muitos pesquisadores que a enxergavam apenas como uma particularidade da conjugada de Moreau. Deste modo, Sosa investigou as relações existentes entre a teoria de conjugação de Fenchel com os teoremas de separação clássicos [8]. Estes teoremas garantem a existência de hiperplanos separadores para conjuntos convexos e fechados. Em [23] e [24] é possível notar que podemos construir a teoria de conjugação a partir destes resultados.

Graças a Michael [21], que separou subconjuntos fechados particulares de \mathbb{R}^n usando funções contínuas, Sosa conseguiu generalizar em [26] os teoremas de separação estendendo os hiperplanos para funções contínuas e mostrou que a conjugada modificada em [6] provém destes teoremas.

Cotrina, Raupp e Sosa [7] desenvolveram a teoria de dualidade para a programação semicontínua e investigaram suas propriedades.

O propósito deste trabalho é fazer uma releitura de [6], [7] e [26] e relacioná-los de maneira a convencer o leitor que esta proposta da modificação da conjugada é realmente uma generalização da conjugada de Fenchel.

O primeiro capítulo faz revisão conceitual de alguns tópicos de Análise Convexa [2, 24], bem como apresenta os teoremas clássicos de separação [8] e a partir deles, constroi a conjugada de Fenchel. Por fim, resume a teoria de Dualidade para programação convexa [9, 19].

O segundo capítulo apresenta os teoremas de separação para fechados [21, 22] e através deles generaliza os teoremas de separação clássicos [26].

O terceiro capítulo apresenta a conjugada generalizada [6] e investiga suas propriedades. São introduzidos os espaços duais conjugados das funções e analisam-se alguns exemplos.

Por fim, o último capítulo expõe a aplicação da conjugada modificada na construção de uma teoria de dualidade para programação semicontínua inferior e faz uma análise da função Lagrangeana correspondente [7].

Capítulo 1

Revisão de conceitos

Esse texto é baseado em [2, 8, 9, 23, 24] e apresenta alguns conceitos iniciais de Análise Convexa que servirão de ferramenta para o desenvolvimento deste trabalho. Apresentaremos algumas definições como função estendida, funções convexas e funções semicontínuas. Estabeleceremos uma ponte entre os teoremas clássicos de separação com a teoria de conjugação de Fenchel e para finalizar, desenvolveremos um processo de dualidade relacionado com esta teoria.

1.1 Conceitos iniciais

Do ponto de vista teórico, é interessante trabalhar com funções de valores reais e definidas em todo \mathbb{R}^n , mas no contexto de otimização e de muitas aplicações nem sempre isso é possível. Por exemplo, consideremos $f_k : \mathbb{R} \rightarrow \mathbb{R}$; $f_k(x) = kx$ para $k \in \mathbb{N}$ e

$$f(x) = \sup_{k \in \mathbb{N}} f_k(x).$$

Assim, a função f atinge o valor $+\infty$ por mais que (f_k) seja de valores reais para todo k .

Outro caso é das funções convexas definidas sobre um subconjunto $K \subset \mathbb{R}^n$ que não podem ser estendidas para funções de valores reais e convexas. Como exemplo, temos $f : (0, \infty) \rightarrow \mathbb{R}$ dada por $f(x) = 1/x$. Esta função é convexa em $(0, \infty)$, mas para qualquer extensão de valores reais, a função estendida deixa de ser convexa.

Tomando esses exemplos como motivação, definiremos as funções de valores reais estendidas.

Definição 1.1 Consideremos $f : K \subset \mathbb{R}^n \rightarrow \mathbb{R}$. A função de valores reais estendida de f é definida como:

$$F(x) = \begin{cases} f(x), & x \in K \\ +\infty, & x \notin K \end{cases}.$$

Podemos estender a função impondo que f assumo $-\infty$ no complementar de K . Isso é definido de acordo com o comportamento da função f . Notemos que, de acordo com essa definição, a função estendida mantém a convexidade ou a concavidade da função f .

As regras das operações aritméticas envolvendo $+\infty$ e $-\infty$ são adotadas da seguinte forma:

(i) $x + \infty = \infty + x = +\infty$ para $-\infty < x \leq +\infty$,

(ii) $x - \infty = -\infty + x = -\infty$ para $-\infty \leq x < +\infty$,

- (iii) $x(+\infty) = (+\infty)x = +\infty$, $x(-\infty) = (-\infty)x = -\infty$ para $0 < x \leq +\infty$,
- (iv) $x(+\infty) = (+\infty)x = -\infty$, $x(-\infty) = (-\infty)x = +\infty$ para $-\infty \leq x < 0$,
- (v) $\inf \emptyset = +\infty$, $\sup \emptyset = -\infty$.

As expressões $(\infty - \infty)$, $(-\infty + \infty)$, $0(+\infty)$ e $0(-\infty)$ são consideradas indeterminadas. Devido a esse fato, devemos estender as funções de forma a evitar essa situação, incluindo no contradomínio apenas $+\infty$ ou $-\infty$.

A partir deste ponto do trabalho, será convencionado que as funções serão de valores reais estendidos, podendo atingir $\pm\infty$. Vamos definir agora alguns subconjuntos que serão utilizados na análise de convexidade e semicontinuidade das funções.

Definição 1.2

- (i) O epigrafo de uma função f , denotado por $\text{epi}(f)$, é a região que está acima do seu gráfico, ou seja,

$$\text{epi}(f) = \{(x, \lambda) \in \mathbb{R}^{n+1} \mid f(x) \leq \lambda\}.$$

- (ii) O hipografo, denotado por $\text{hip}(f)$, é a região que está abaixo do gráfico, ou seja,

$$\text{hip}(f) = \{(x, \lambda) \in \mathbb{R}^{n+1} \mid f(x) \geq \lambda\}.$$

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

$$L_f(\lambda) = \{x \in \mathbb{R}^n \mid f(x) \leq \lambda\}.$$

Podemos interpretar o conjunto de nível como uma projeção em \mathbb{R}^n de seções horizontais do $\text{epi}(f)$. A análise de suas propriedades topológicas é muito importante no contexto de minimização de funções, pois para $\lambda = \inf f$, temos que

$$L_f(\lambda) = \{x \in \mathbb{R}^n \mid f(x) = \lambda\} = \text{argmin } f.$$

Consideremos o próximo exemplo:

Exemplo 1.3 Seja $f : \mathbb{R} \rightarrow \mathbb{R} \cup \{+\infty\}$ dada por $f(x) = \begin{cases} 1/x, & x > 0 \\ +\infty, & x \leq 0 \end{cases}$.

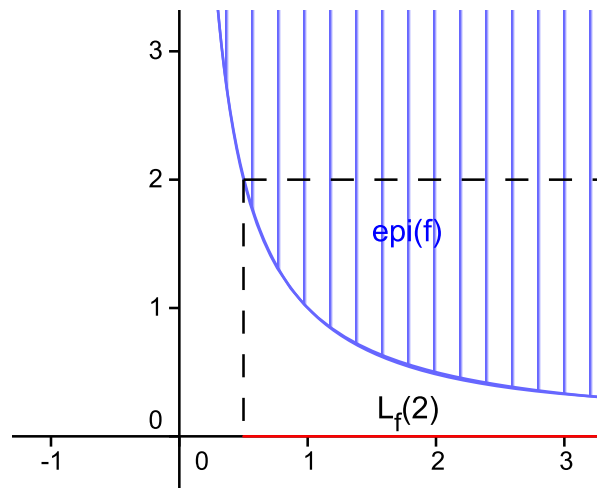


Figura 1.1: Epigrafo de f e o conjunto de nível para $\lambda = 2$.

Na Figura 1.1, notemos que o hipografo é o gráfico de f unido com a região que está abaixo dele, assim, a interseção deste com o epigrafo de f é exatamente o gráfico de f . Além disso, os pontos x para os quais $f(x) = +\infty$ não pertencem ao epigrafo de f , mas pertencem ao hipografo de f . Temos também representado o conjunto de nível de f para $\lambda = 2$.

Definição 1.4 *O domínio efetivo de uma função f , que será denotado por $\text{dom}(f)$, é a projeção do $\text{epi}(f)$ em \mathbb{R}^n , ou seja,*

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

Neste trabalho, quando estivermos nos referindo a domínio efetivo usaremos apenas o termo “domínio”. No Exemplo 1.3, o domínio de f são os números reais positivos.

Definição 1.5 *Uma função $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ é dita própria quando existe $z \in \mathbb{R}^n$ tal que $f(z) < \infty$.*

Uma função que não é própria é chamada de *imprópria*. Temos como exemplo, a função identicamente $+\infty$. Do ponto de vista prático, as funções próprias são mais interessantes e aparecem com maior frequência. Diretamente da definição, podemos perceber que f é própria quando seu domínio é não-vazio.

Lembramos que um subconjunto K de \mathbb{R}^n é convexo quando para quaisquer x e y em K , o segmento $(1 - \alpha)x + \alpha y \in K$, com $\alpha \in [0, 1]$. Assim, vamos definir funções convexas da seguinte maneira:

Definição 1.6 *Uma função f é dita convexa quando seu epigrafo é um conjunto convexo em \mathbb{R}^{n+1} .*

Note que para analisar a convexidade da função nos limitamos a trabalhar com a restrição de f apenas sobre seu domínio, pois se $f(x) = +\infty$, então $(x, \lambda) \notin \text{epi}(f)$ para nenhum $\lambda \in \mathbb{R}$. O próximo teorema relaciona essa definição com uma caracterização mais usual de convexidade.

Teorema 1.7 *Uma função f é convexa se, e somente se,*

$$f((1 - \alpha)x + \alpha y) \leq (1 - \alpha)f(x) + \alpha f(y) \quad (1.1)$$

vale para todo x e y em \mathbb{R}^n e para $\alpha \in (0, 1)$.

Demonstração. (\Rightarrow) Dados x e y em \mathbb{R}^n , se x ou y não pertencem ao $\text{dom}(f)$ então (1.1) se verifica trivialmente. Caso ambos estejam no domínio de f , notemos que $(x, f(x))$ e $(y, f(y))$ estão no $\text{epi}(f)$.

Como $\text{epi}(f)$ é convexo, então $(1 - \alpha)(x, f(x)) + \alpha(y, f(y)) \in \text{epi}(f)$ para todo $\alpha \in (0, 1)$. Além disso,

$$\begin{aligned} (1 - \alpha)(x, f(x)) + \alpha(y, f(y)) &= ((1 - \alpha)x, (1 - \alpha)f(x)) + (\alpha y, \alpha f(y)) = \\ &= ((1 - \alpha)x + \alpha y, (1 - \alpha)f(x) + \alpha f(y)). \end{aligned}$$

Portanto, $((1 - \alpha)x + \alpha y, (1 - \alpha)f(x) + \alpha f(y)) \in \text{epi}(f)$, implicando que

$$f((1 - \alpha)x + \alpha y) \leq (1 - \alpha)f(x) + \alpha f(y)$$

para todo $x, y \in \text{dom}(f)$ e $\alpha \in (0, 1)$.

(\Leftarrow) Dados $(x, \lambda_1), (y, \lambda_2) \in \text{epi}(f)$, então, por definição, $f(x) \leq \lambda_1$ e $f(y) \leq \lambda_2$. Assim, por (1.1), podemos concluir que

$$f((1 - \alpha)x + \alpha y) \leq (1 - \alpha)f(x) + \alpha f(y) \leq (1 - \alpha)\lambda_1 + \alpha\lambda_2$$

para todo $\alpha \in (0, 1)$. Portanto,

$$((1 - \alpha)x + \alpha y, (1 - \alpha)\lambda_1 + \alpha\lambda_2) = (1 - \alpha)(x, \lambda_1) + \alpha(x, \lambda_2) \in \text{epi}(f)$$

com $\alpha \in (0, 1)$ e assim, $\text{epi}(f)$ é um conjunto convexo em \mathbb{R}^{n+1} . \square

1.2 Funções semicontínuas

Da literatura, sabemos que \mathbb{R}^n é um espaço vetorial com produto interno. Deste modo, para x e y em \mathbb{R}^n , lembramos que o produto interno usual é dado por

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i$$

em que x_i e y_i correspondem respectivamente as i -ésimas coordenadas de x e de y .

A norma em \mathbb{R}^n que será considerada neste trabalho é a que provém deste produto interno, ou seja, para $x \in \mathbb{R}^n$, teremos que $\|x\| = \sqrt{\langle x, x \rangle}$. Esta norma induz a topologia cuja base são as bolas abertas $B(x_0, r) = \{x \in \mathbb{R}^n \mid \|x - x_0\| < r\}$ com $x_0 \in \mathbb{R}^n$ e $r > 0$.

Portanto, diremos que um subconjunto K é *aberto* em \mathbb{R}^n quando cada ponto dele é centro de uma bola aberta contida em K .

Definição 1.8 *Um ponto $x_0 \in \mathbb{R}^n$ é aderente a um subconjunto $K \subset \mathbb{R}^n$ quando é limite de uma sequência de pontos desse conjunto. O conjunto dos pontos aderentes de K , denotado por \overline{K} , é chamado de fecho de K .*

Notemos que o fecho de $B(x_0, r)$ é a bola fechada

$B[x_0, r] = \{x \in \mathbb{R}^n \mid \|x - x_0\| \leq r\}$ e o fecho de \overline{K} é o próprio conjunto.

Um subconjunto K de \mathbb{R}^n é dito *fechado* quando $K = \overline{K}$. Uma das propriedades interessantes é que um subconjunto é fechado se, e só se seu complementar é aberto.

Teorema 1.9 *Consideremos $(K_j)_{j \in J}$ uma família de subconjuntos fechados e convexos de \mathbb{R}^n em que J é um conjunto arbitrário de índices. Então $K = \bigcap_{j \in J} K_j$ é um conjunto convexo e fechado em \mathbb{R}^n .*

Demonstração. Dados x_1 e x_2 em K , então x_1 e x_2 pertencem a K_j para todo $j \in J$. Como cada K_j é convexo, então $(1 - \alpha)x_1 + \alpha x_2 \in K_j$ para todo $j \in J$ e para todo $\alpha \in (0, 1)$. Portanto, $(1 - \alpha)x_1 + \alpha x_2 \in K$ para $\alpha \in [0, 1]$. Dessa forma, K é convexo.

Para provar que K é fechado, considere uma sequência (x_k) contida em K que converge para um certo x_0 . Assim, (x_k) está contida em K_j para todo $j \in J$.

Como cada K_j é fechado, então $x_0 \in K_j$ para todo $j \in J$. Logo, $x_0 \in K$ e, portanto, K é fechado. \square

Antes de definir semicontinuidade, vamos revisar algumas propriedades de liminf e limsup.

Definição 1.10 Dados uma função f , $x_0 \in \mathbb{R}^n$ e $\delta > 0$, consideremos

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

À medida que δ se aproxima monotonicamente de 0, $M(x_0, \delta)$ não cresce e $m(x_0, \delta)$ não decresce. Assim, definimos:

Definição 1.11

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

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

Para qualquer sequência (δ_k) que se aproxima monotonicamente de zero, as sequências $(M(x_0, \delta_k))$ e $(m(x_0, \delta_k))$ também serão monótonas. Assim, o $\limsup_{x \rightarrow x_0} f(x)$ e o $\liminf_{x \rightarrow x_0} f(x)$ sempre existem, independente da sequência tomada, podendo admitir valores $\pm\infty$. Vamos analisar outras caracterizações dessa definição nos próximos teoremas. Demonstraremos apenas o item (i) de cada um dos três teoremas que seguem, pois a demonstração do item (ii) de cada um deles é análoga.

Teorema 1.12

(i) Se $m = \liminf_{x \rightarrow x_0} f(x)$ então existe $(x_k) \subset \mathbb{R}^n$ tal que $x_k \rightarrow x_0$ e $f(x_k) \rightarrow m$.

(ii) Se $M = \limsup_{x \rightarrow x_0} f(x)$ então existe $(x_k) \subset \mathbb{R}^n$ tal que $x_k \rightarrow x_0$ e $f(x_k) \rightarrow M$.

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

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, conseqüentemente, $m = +\infty$. 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 $k \rightarrow \infty$, a sequência $\delta_k \rightarrow 0$, $x_k \rightarrow x_0$ e $f(x_k) \rightarrow +\infty = m$.

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

$$m_k \leq f(x_k) < m_k + 1/k. \quad (1.2)$$

Fazendo $k \rightarrow \infty$, teremos $\delta_k \rightarrow 0$. Portanto, $x_k \rightarrow x_0$ e $m_k \rightarrow m$. Pela desigualdade (1.2), concluímos que $f(x_k) \rightarrow m$. \square

Teorema 1.13

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

(ii) Seja $M = \limsup_{x \rightarrow x_0} f(x)$. 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 $\lim_{\delta \rightarrow 0} m(x_0, \delta) > m'$. Por propriedade de limite, existe algum $\delta' > 0$ tal que $m(x_0, \delta') > m'$, ou seja,

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

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

A recíproca dos dois teoremas anteriores é dada a seguir.

Teorema 1.14 Dada $(x_k) \subset \mathbb{R}^n$ tal que $x_k \rightarrow x_0$ e $f(x_k) \rightarrow m$.

- (i) Se para cada $m' < m$ existe $\delta > 0$ tal que $m' < f(x)$ para todo $x \in B(x_0, \delta)$ então $m = \liminf_{x \rightarrow x_0} f(x)$.
- (ii) Se para cada $m' > m$, existe $\delta > 0$ tal que $m' > f(x)$ para todo $x \in B(x_0, \delta)$ então $m = \limsup_{x \rightarrow x_0} f(x)$.

Demonstração. (i) Como $x_k \rightarrow x_0$, para $\delta > 0$ existe k_0 tal que $x_k \in B(x_0, \delta)$ para todo $k > k_0$. Assim, $m(x_0, \delta) \leq f(x_k)$ para cada $k > k_0$.

Quando $k \rightarrow \infty$, a conservação de sinal do limite nos permite concluir que $m(x_0, \delta) \leq m$.

Por fim, fazendo δ tender a 0, obtemos $\liminf_{x \rightarrow x_0} f(x) \leq m$.

Agora, vamos supor por absurdo que $\liminf_{x \rightarrow x_0} f(x) < m$. Existe, por hipótese, $\delta > 0$ tal que $f(x) > \liminf_{x \rightarrow x_0} f(x)$ para todo $x \in B(x_0, \delta)$. Portanto,

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

Absurdo, pois contraria a definição de liminf. Concluimos assim que $m = \liminf_{x \rightarrow x_0} f(x)$.

□

Definição 1.15

- (i) Uma função f é *semicontínua inferiormente* ou *sci* em um ponto $x_0 \in \mathbb{R}^n$ quando $\liminf_{x \rightarrow x_0} f(x) \geq f(x_0)$.
- (ii) Uma função f é *semicontínua superiormente* ou *scs* em um ponto $x_0 \in \mathbb{R}^n$ quando $\limsup_{x \rightarrow x_0} f(x) \leq f(x_0)$.

Pelos teoremas provados anteriormente, esta definição pode ser caracterizada da seguinte forma:

Corolário 1.16

- (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 quando 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 quando dado $M' > f(x_0)$, existir $\delta > 0$ tal que $M' > f(x)$ para todo $x \in B(x_0, \delta)$.

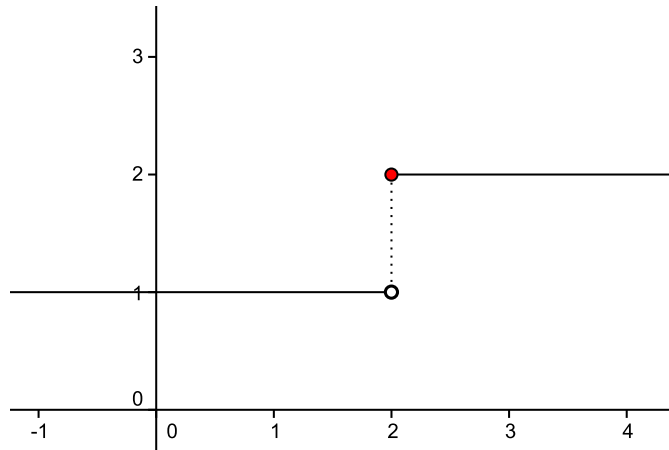
As funções semicontínuas serão o principal objeto de estudo deste trabalho. Vamos analisar alguns exemplos:

Exemplo 1.17 Consideremos $f_1(x) = \begin{cases} 1, & x < 2 \\ 2, & x \geq 2 \end{cases}$ com o gráfico apresentado na Figura 1.2.

Notemos que para $x_0 \neq 2$,

$$\liminf_{x \rightarrow x_0} f(x) = \limsup_{x \rightarrow x_0} f(x) = f(x_0)$$

e, portanto, f é scs e sci nesses pontos. Já para $x_0 = 2$, temos que $\liminf_{x \rightarrow 2} f(x) = 1 < f(2)$ e $\limsup_{x \rightarrow 2} f(x) = 2 = f(2)$. Logo, f é apenas scs em 2.

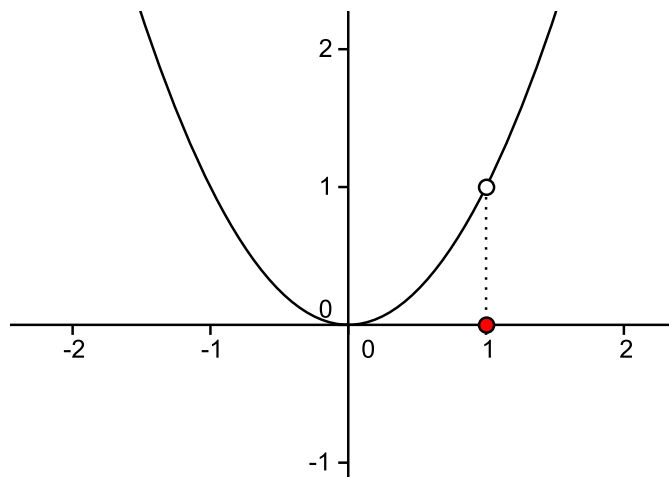
Figura 1.2: Gráfico de f_1 .

Exemplo 1.18 Seja $f_2(x) = \begin{cases} x^2, & x \neq 1 \\ 0 & x = 1 \end{cases}$ com gráfico apresentado na Figura 1.3.

Analogamente ao caso anterior, para $x_0 \neq 1$,

$$\liminf_{x \rightarrow x_0} f(x) = \limsup_{x \rightarrow x_0} f(x) = f(x_0)$$

e portanto f é scs e sci nesses pontos. Mas para $x_0 = 1$, temos que $\liminf_{x \rightarrow 1} f(x) = 0 = f(1)$ e $\limsup_{x \rightarrow 2} f(x) = 1^2 = 1 > f(1)$. Logo, f é apenas sci em 1.

Figura 1.3: Gráfico de f_2 .

O próximo teorema apresenta uma caracterização topológica da semicontinuidade. Veremos que é possível classificar a função analisando seus conjuntos de nível ou até mesmo o seu epigrafo.

Teorema 1.19 Dada a função f , são equivalentes:

- (i) f é sci em todo \mathbb{R}^n .
- (ii) $\text{epi}(f)$ é um conjunto fechado em \mathbb{R}^{n+1} .
- (iii) $L_f(\lambda)$ é fechado para todo $\lambda \in \mathbb{R}$.

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

$$f(x_k) \leq \lambda_k. \quad (1.3)$$

Para δ fixo, devido ao fato de $x_k \rightarrow x_0$, 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 (1.3), $m(x_0, \delta) \leq \lambda_k$.

Fazendo $k \rightarrow \infty$ temos, pela conservação de sinal do limite, $m(x_0, \delta) \leq \lambda_0$. Finalmente, fazendo $\delta \rightarrow 0$ e pelo fato de f ser sci em todo \mathbb{R}^n , concluímos que

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

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

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

$$L_f(\lambda) \times \{\lambda\} = (\mathbb{R}^n \times \{\lambda\}) \cap \text{epi}(f),$$

em que \times denota o produto cartesiano.

De fato,

$$(x, \lambda) \in L_f(\lambda) \times \{\lambda\} \Leftrightarrow f(x) \leq \lambda \Leftrightarrow (x, \lambda) \in \text{epi}(f) \cap (\mathbb{R}^n \times \{\lambda\}).$$

Logo, $L_f(\lambda) \times \{\lambda\}$ é fechado. Considerando $T : \mathbb{R}^n \rightarrow \mathbb{R}^{n+1}$ 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 \mathbb{R}^n . Vamos dividir em dois casos.

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

(2º caso: $\liminf_{x \rightarrow x_0} f(x) < +\infty$.) Pelo Teorema 1.12, existe (x_k) que converge para x_0 tal que $f(x_k)$ converge para $\liminf_{x \rightarrow x_0} f(x)$. Para todo

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

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 $\lambda > \liminf_{x \rightarrow x_0} f(x)$.

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

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

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

Com demonstração análoga, podemos garantir que uma função é scs quando seu hipografo é fechado ou quando o conjunto $\{x \in \mathbb{R}^n \mid f(x) \geq \lambda\}$ 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 $\{x \in \mathbb{R}^n \mid f(x) \geq \lambda\}$ são pré-imagens de conjuntos fechados e conseqüentemente serão fechados também.

A seguir, vamos analisar a semicontinuidade da seguinte função:

Exemplo 1.20 Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R} \cup \{+\infty\}$;

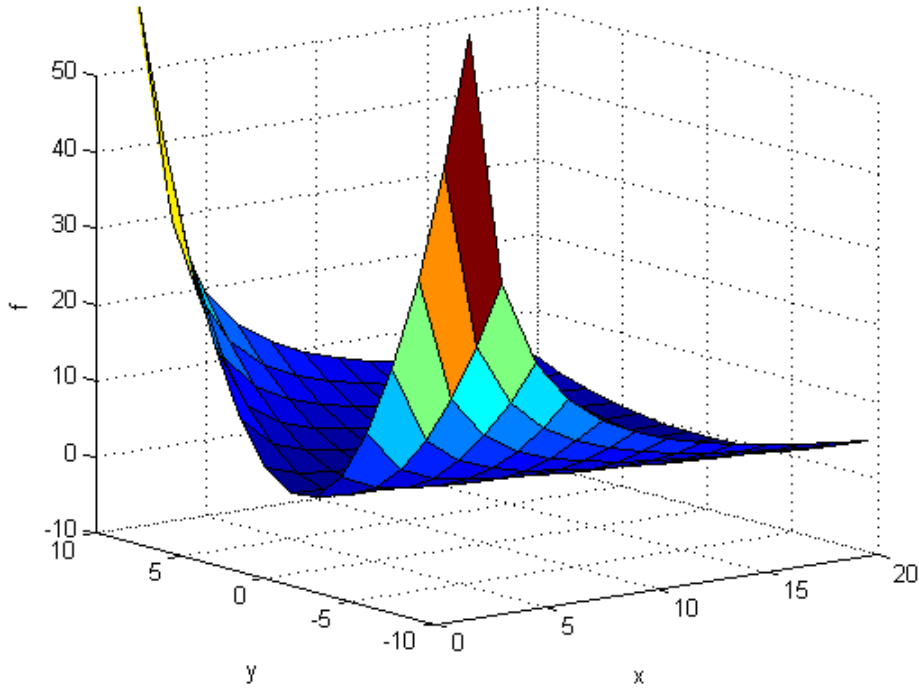
$$f(x, y) = \begin{cases} y^2/x, & x > 0 \quad e \quad y \in \mathbb{R} \\ 0, & x = 0 \quad e \quad y = 0 \\ +\infty, & \text{em outros casos} \end{cases} \quad \text{com gráfico na Figura 1.4.}$$


Figura 1.4: Gráfico de f .

Notamos inicialmente que é possível concluir que $e\text{pi}(f)$ é um conjunto convexo e assim f é convexa. Além disso, f é contínua em todos os pontos exceto na origem. De fato, para $\delta > 0$,

$$m((0, 0), \delta) = 0 \quad e \quad M((0, 0), \delta) = \infty.$$

Fazendo $\delta \rightarrow 0$, obtemos

$$\liminf_{(x,y) \rightarrow (x_0,y_0)} f(x, y) = 0 = f(x_0, y_0),$$

$$\limsup_{(x,y) \rightarrow (x_0,y_0)} f(x, y) = \infty \geq f(x_0, y_0).$$

Assim, f é sci em $(0, 0)$, mas não é scs nesse ponto. Deste modo, concluímos que f é sci e convexa em todo \mathbb{R}^n , mas não é scs apenas em $(0, 0)$.

Vamos apresentar agora dois resultados importantes de semicontinuidade. O primeiro relaciona semicontinuidade inferior com a superior e o segundo será utilizado para caracterizar as funções conjugadas.

Teorema 1.21 *A função f é sci se, e somente se, $-f$ é scs.*

Demonstração. As seguintes equivalências são válidas:

$$\begin{aligned} \liminf_{x \rightarrow x_0} f(x) \geq f(x_0) &\Leftrightarrow -\liminf_{x \rightarrow x_0} f(x) \leq -f(x_0) \Leftrightarrow \\ \Leftrightarrow \lim_{\delta \rightarrow 0} -\inf \{f(x) \mid x \in B(x_0, \delta)\} &\leq -f(x_0) \Leftrightarrow \lim_{\delta \rightarrow 0} \sup \{-f(x) \mid x \in B(x_0, \delta)\} \leq -f(x_0) \Leftrightarrow \\ &\Leftrightarrow \limsup_{x \rightarrow x_0} -f(x) \leq -f(x_0). \end{aligned}$$

Deste modo, f é sci em x_0 se, e só se, $-f$ é scs em x_0 . □

Teorema 1.22 *Seja $(f_j)_{j \in J}$ uma família de funções sci e convexas em que J é um conjunto arbitrário de índices. Então a função*

$$f(x) = \sup_{j \in J} f_j(x)$$

é sci e convexa.

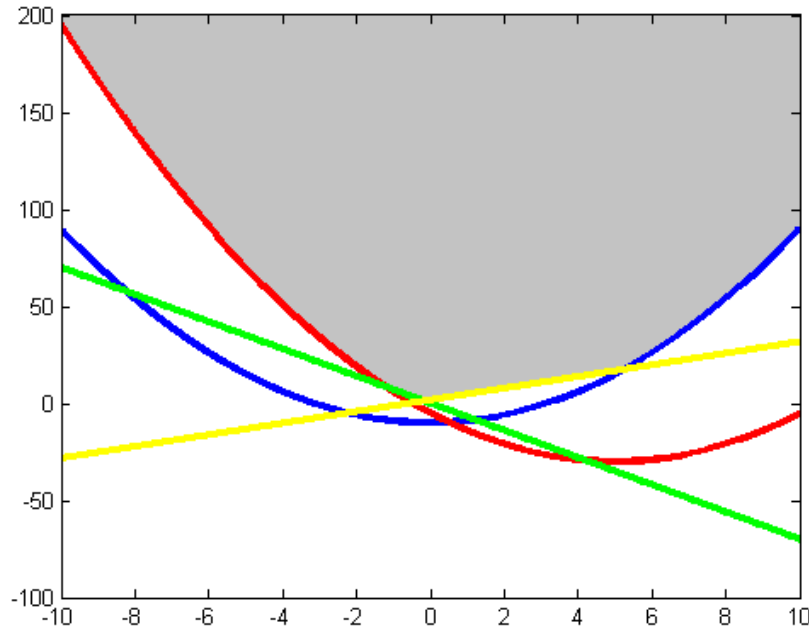


Figura 1.5: Epígrafo do supremo pontual de quatro funções convexas e sci.

Demonstração. A Figura 1.5 nos motiva a afirmar que

$$epi(f) = \bigcap_{j \in J} epi(f_j).$$

De fato, usando definição de epígrafo e de supremo, sabemos que, para cada $j \in J$,

$$(x, \lambda) \in epi(f) \Leftrightarrow f(x) \leq \lambda \Leftrightarrow f_j(x) \leq \lambda \Leftrightarrow (x, \lambda) \in epi(f_j).$$

Logo,

$$(x, \lambda) \in \text{epi}(f) \Leftrightarrow (x, \lambda) \in \bigcap_{j \in J} \text{epi}(f_j).$$

Como $\text{epi}(f_j)$ é fechado e convexo para todo j então, pelo Teorema 1.9, $\text{epi}(f)$ é fechado e convexo. Portanto, pelo Teorema 1.19, f é sci e convexa. \square

Para encerrar esta sessão, apresentaremos uma versão do Teorema de Weierstrass para funções semicontínuas inferiormente. Estamos adotando a definição de compacidade para espaços vetoriais de dimensão finita, ou seja, um conjunto é dito *compacto* quando ele é limitado e fechado em \mathbb{R}^n .

Teorema 1.23 *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função própria sci e seu domínio um subconjunto compacto em \mathbb{R}^n . Então $\inf f$ é finito e atingido para algum ponto do domínio.*

Demonstração. Seja $\bar{\lambda} = \inf f$. Como f é própria, então $\bar{\lambda} < +\infty$. Consideremos $I = (\bar{\lambda}, +\infty)$ e $(L_f(\lambda))_{\lambda \in I}$ uma família de conjuntos contida no domínio de f .

Sabemos que $L_f(\lambda)$ é não-vazio para todo $\lambda \in I$. De fato, se $L_f(\lambda)$ fosse vazio para algum λ em I , então, $f(x) > \lambda$ para todo $x \in \mathbb{R}^n$ e assim $\lambda = \bar{\lambda}$, o que é absurdo.

Como f é sci, logo $L_f(\lambda)$ é fechado em $\text{dom}(f)$. Notemos também que $L_f(\lambda) \subset L_f(\lambda')$ para $\lambda \leq \lambda'$, pois se $x \in L_f(\lambda)$, então $f(x) \leq \lambda \leq \lambda'$ e portanto, $x \in L_f(\lambda')$.

Dessa forma, para qualquer subconjunto finito $\{\lambda_1, \lambda_2, \dots, \lambda_n\} \subset I$ temos que $\bigcap_{k=1}^n L_f(\lambda_k)$ é não-vazia. Como $\text{dom}(f)$ é compacto, então $\bigcap_{\lambda \in I} L_f(\lambda) \neq \emptyset^1$.

Podemos ver que $\bigcap_{\lambda \in I} L_f(\lambda) = \text{argmin } f$, logo $\text{argmin } f$ é não vazio e compacto e, portanto, $\bar{\lambda} \in \mathbb{R}$ e f possui minimizador. \square

É possível também demonstrar que as funções scs com domínio compacto atingem seu valor máximo, o qual é finito.

1.3 Conjuntos de recessão

Os problemas de programação semicontínua inferior (*PSCI*) tratam de encontrar minimizadores de funções sci em um determinado conjunto K , ou seja, esquematicamente, queremos

$$\begin{aligned} & \text{minimizar } f(x) \\ & \text{sujeito a } x \in K, \end{aligned}$$

com $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ própria e sci e K um subconjunto não-vazio e fechado de \mathbb{R}^n .

Pelo Teorema 1.23, quando K for limitado e portanto compacto, o problema tem solução. Faremos no Capítulo 4 um estudo detalhado desses problemas e apresentaremos outras condições que garantem a existência de solução do *PSCI*. Faremos agora uma breve análise dos conjuntos de recessão e veremos alguns resultados que serão necessários para a análise dessas condições.

Definição 1.24 *Dado um subconjunto K de \mathbb{R}^n , o conjunto de recessão de K , denotado por K^∞ , é definido por*

$$K^\infty = \{d \in \mathbb{R}^n \mid \exists t_k \rightarrow +\infty, \exists (x_k) \subset K, t_k^{-1} x_k \rightarrow d\}.$$

Por convenção, $\emptyset^\infty = \{0\}$.

¹Toda família de conjuntos fechados $(F_\beta)_{\beta \in J}$ contida em um compacto de \mathbb{R}^n tal que, para qualquer subconjunto finito $\{\beta_1, \beta_2, \dots, \beta_n\} \subset J$, o conjunto $\bigcap_{k=1}^n F_{\beta_k}$ é não-vazio, então $\bigcap_{\beta \in J} F_\beta$ também é não-vazio. Para mais detalhes, ver [18].

Diretamente dessa definição, podemos deduzir o seguinte teorema.

Teorema 1.25 K^∞ é um conjunto não-vazio e fechado.

Demonstração. Primeiramente, vamos provar que a origem sempre pertence a K^∞ . De fato, se $K = \emptyset$, então, por convenção $K^\infty = \{0\}$. Caso contrário, tome $x \in K$ e considere as seqüências $x_k = x$ e $t_k = k$ para todo k . Assim, $t_k^{-1}x_k = \frac{x}{k} \rightarrow 0$ e portanto $0 \in K^\infty$. Logo, K^∞ é não vazio.

Para provar que K^∞ é fechado, considere $d_k \in K^\infty$ tal que $d_k \rightarrow d$. Dessa forma, para cada k existem $t_{k,j} \in \mathbb{R}$ e $x_{k,j} \in K$ tais que $\lim_{j \rightarrow \infty} t_{k,j} = +\infty$ e $\lim_{j \rightarrow \infty} t_{k,j}^{-1}x_{k,j} = d_k$.

Assim, é possível escolher j_k tal que

$$t_{k,j_k} > k \quad e \quad \|t_{k,j_k}^{-1}x_{k,j_k} - d_k\| < 1/k.$$

Definimos as seqüências

$$s_k = t_{k,j_k} \quad e \quad y_k = x_{k,j_k}$$

para todo $k \in \mathbb{N}$.

Deste modo, $y_k \in K$ e como $s_k > k$, temos que $\lim_{k \rightarrow \infty} s_k = +\infty$. Além disso, $\lim_{k \rightarrow \infty} s_k^{-1}y_k = d$. De fato, note que

$$\|s_k^{-1}y_k - d\| \leq \|s_k^{-1}y_k - d_k\| + \|d_k - d\| < 1/k + \|d_k - d\|.$$

Deste modo, quando $k \rightarrow \infty$, a seqüência $(s_k^{-1}y_k)$ converge para d . Portanto, $d \in K^\infty$ e assim K^∞ é fechado. \square

Esse teorema caracteriza topologicamente os conjuntos de recessão para um subconjunto arbitrário de \mathbb{R}^n . No entanto, veremos no próximo teorema que para subconjuntos não-vazios, convexos e fechados, é possível dar uma caracterização mais geométrica do conjunto K^∞ .

Teorema 1.26 Se K é um subconjunto não-vazio, convexo e fechado de \mathbb{R}^n , então

$$K^\infty = \{d \in \mathbb{R}^n \mid x + td \in K, \forall t > 0\}$$

com x sendo um elemento qualquer de K .

Demonstração. Dado $x \in K$, consideremos $d \in \{d \in \mathbb{R}^n \mid x + td \in K, \forall t > 0\}$, então, definimos a seqüência

$$t_k = k$$

para $k \in \mathbb{N}$. Deste modo, $t_k > 0$ e, portanto, $x_k = x + t_k d \in K$ para todo k . Além disso,

$$t_k^{-1}x_k = t_k^{-1}(x + t_k d) = t_k^{-1}x + d \rightarrow d.$$

Concluimos assim que $d \in K^\infty$.

Para provar a inclusão contrária, consideremos $d \in K^\infty$ e $t > 0$. Assim, por definição de conjunto de recessão, existem $t_k \rightarrow +\infty$ e $x_k \in K$ tais que $t_k^{-1}x_k \rightarrow d$. Definimos

$$d_k = t_k^{-1}(x_k - x).$$

Assim

$$\lim_{k \rightarrow \infty} d_k = \lim_{k \rightarrow \infty} (t_k^{-1}x_k - t_k^{-1}x) = d$$

e, para todo k , temos que

$$x + t_k d_k = x + t_k t_k^{-1} (x_k - x) = x_k \in K.$$

Como a sequência (t_k) é ilimitada, podemos escolher k suficientemente grande tal que $t \leq t_k$ e assim $t_k^{-1} t \leq 1$. Desde que K é convexo,

$$x + t d_k = (1 - t_k^{-1} t) x + t_k^{-1} t (x + t_k d_k) \in K$$

para todo k . Já que K é fechado, então

$$x + t d = \lim_{k \rightarrow \infty} (x + t d_k) \in K^\infty,$$

concluindo assim a demonstração. \square

Assim, para um subconjunto K que satisfaça as hipóteses do teorema, os vetores do conjunto de recessão de K são as direções das semirretas contidas em K .

Exemplo 1.27 Consideremos $f : \mathbb{R} \rightarrow \mathbb{R}; f(x) = x^2$.

Essa função é contínua, convexa e própria, portanto o $\text{epi}(f)$ é um subconjunto não-vazio, convexo e fechado em \mathbb{R}^2 . Vamos analisar o conjunto de recessão de $K = \text{epi}(f)$.

Toda semirreta contida em K é paralela ao eixo y e assim, qualquer vetor do conjunto de recessão de K é também paralelo ao eixo y . De fato, notemos que o conjunto

$$\{(d_1, d_2) \in \mathbb{R}^2 \mid d_1 = 0 \text{ e } d_2 \geq 0\} \subset K^\infty.$$

Pois, se $(d_1, d_2) \in \{(d_1, d_2) \in \mathbb{R}^2 \mid d_1 = 0 \text{ e } d_2 \geq 0\}$, então, dado $(x, \lambda) \in \text{epi}(f)$, temos que

$$(x, \lambda) + t(d_1, d_2) = (x, \lambda + t d_2) \in \text{epi}(f)$$

para todo $t > 0$, concluindo a inclusão.

Por outro lado, dado $d = (d_1, d_2) \in (\text{epi}(f))^\infty$, o vetor $(x, \lambda) + t d \in \text{epi}(f)$ para $(x, \lambda) \in \text{epi}(f)$ e $t > 0$. Assim,

$$(x + t d_1)^2 \leq \lambda + t d_2. \tag{1.4}$$

Desenvolvendo esta expressão, temos que

$$x^2 + 2t x d_1 + t^2 d_1^2 \leq \lambda + t d_2 \Rightarrow x^2/t^2 + 2x d_1/t + d_1^2 \leq \lambda/t^2 + d_2/t.$$

Quando $t \rightarrow \infty$, temos que $d_1 = 0$. Substituindo este valor em (1.4), obtemos

$$\lambda + t d_2 \geq x^2 \Rightarrow \lambda + t d_2 \geq 0.$$

Portanto, $d_2 > 0$, pois do contrário, para algum t , o valor $\lambda + t d_2$ seria negativo. Logo, todo elemento do conjunto de recessão de K é paralelo ao eixo y .

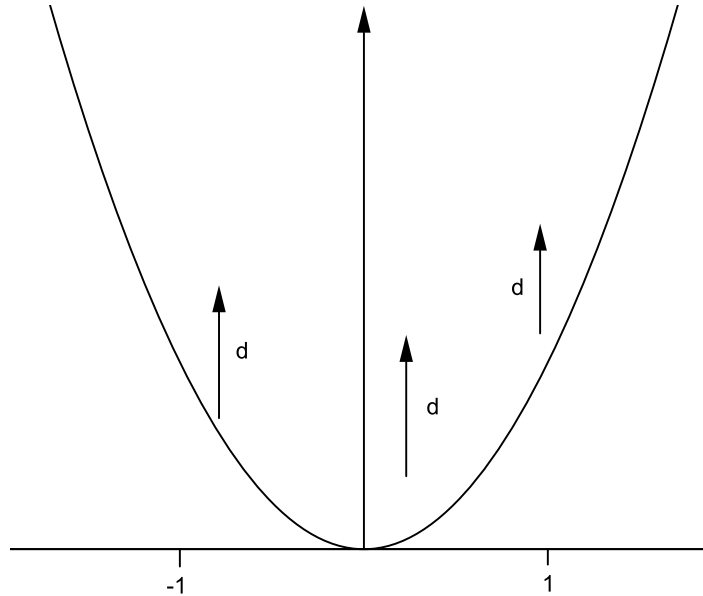


Figura 1.6: Conjunto de recessão de $\text{epi}(f)$.

Analisaremos agora alguns resultados importantes dos conjuntos de recessão.

Teorema 1.28 *Um subconjunto K de \mathbb{R}^n é limitado se, e somente se, $K^\infty = \{0\}$.*

Demonstração. (\Rightarrow) Se K é limitado, então, qualquer sequência $(x_k) \subset K$ é limitada. Portanto, para toda sequência $(t_k) \rightarrow \infty$, teremos que $\lim_{k \rightarrow \infty} t_k^{-1} x_k = 0$. Logo, 0 é o único elemento de K^∞ .

(\Leftarrow) Suponha por absurdo que K é ilimitado. Então, existe uma sequência $(x_k) \subset K$ tal que, para todo k , $x_k \neq 0$ e $t_k := \|x_k\| \rightarrow \infty$. Definimos $d_k := t_k^{-1} x_k$. Assim, $\|d_k\| = 1$ para todo k , ou seja, (d_k) é limitada e, portanto, admite subsequência (d_{k_j}) tal que $d_{k_j} \rightarrow d$ com $\|d\| = 1$ e assim, $d \in K^\infty$. Absurdo, pois $K^\infty = \{0\}$ por hipótese. \square

Teorema 1.29 *Sejam K_1 e K_2 subconjuntos de \mathbb{R}^n . Se $K_1 \subset K_2$ então $(K_1)^\infty \subset (K_2)^\infty$.*

Demonstração. Se $d \in (K_1)^\infty$ então existem $t_k \rightarrow \infty$ e $x_k \subset K_1$ tais que $t_k^{-1} x_k \rightarrow d$. Como $K_1 \subset K_2$ então $(x_k) \subset K_2$ e portanto $d \in (K_2)^\infty$. \square

Teorema 1.30 *Seja $(K_i)_{i \in I}$ uma família de subconjuntos não-vazios de \mathbb{R}^n em que I é um conjunto arbitrário de índices. Então*

$$\left(\bigcap_{i \in I} K_i \right)^\infty \subset \bigcap_{i \in I} (K_i)^\infty. \tag{1.5}$$

Demonstração. Se $d \in \left(\bigcap_{i \in I} K_i \right)^\infty$, então existem $t_k \rightarrow \infty$ e $(x_k) \subset \bigcap_{i \in I} K_i$ tais que $t_k^{-1} x_k \rightarrow d$. Portanto, $(x_k) \subset K_i$ para todo $i \in I$. Logo, $d \in (K_i)^\infty$ para todo $i \in I$ e assim $d \in \bigcap_{i \in I} (K_i)^\infty$. \square

1.4 Separação de convexos

A Teoria de Separação de Convexos foi desenvolvida por Hahn, Banach, Eidelheit e Dieudonne. No entanto, os teoremas mais citados na literatura foram obras de Hahn e Banach e constituíram uma forte ferramenta na Análise Funcional. Nesta seção, veremos os teoremas clássicos de separação que motivaram o estudo das funções conjugadas. Os hiperplanos desenvolvem um papel importante nesta teoria, pois eles possuem a propriedade de dividir um espaço em dois semi-espacos. Deste modo, faremos um resumo desse assunto.

Definição 1.31 Um hiperplano em \mathbb{R}^n , denotado por H , é um conjunto da forma

$$\{x \mid \langle a, x \rangle = b\}$$

sendo a um vetor não nulo de \mathbb{R}^n e b um escalar.

Se y é um vetor do hiperplano $H = \{x \mid \langle a, x \rangle = b\}$, então $\langle a, y \rangle = b$ e assim

$$H = \{x \mid \langle a, x \rangle = \langle a, y \rangle\}.$$

Dessa forma, podemos formular o seguinte teorema.

Teorema 1.32 Seja $H = \{x \mid \langle a, x \rangle = b\}$ e $y \in H$. Então

$$H = y + \{x \mid \langle a, x \rangle = 0\}.$$

Demonstração. De fato, seja $z = y + x$ tal que $\langle a, x \rangle = 0$. Assim

$$\langle a, z \rangle = \langle a, y \rangle + \langle a, x \rangle = \langle a, y \rangle = b,$$

e, portanto, $z \in H$.

Por outro lado, consideremos $z \in H$. A igualdade $z = y + (z - y)$ sempre é válida. Notemos que $z - y \in \{x \mid \langle a, x \rangle = 0\}$ pois $\langle a, z \rangle = \langle a, y \rangle$, concluindo assim a demonstração. \square

Deste modo, H é um conjunto afim que é paralelo ao subespaço $\{x \mid \langle a, x \rangle = 0\}$, o qual é ortogonal ao vetor a . Assim, a é chamado de *vetor normal* de H .

Os conjuntos $\{x \mid \langle a, x \rangle \geq b\}$ e $\{x \mid \langle a, x \rangle \leq b\}$ são chamados de *semi-espacos associados com o hiperplano H* . Esses conjuntos estão exemplificados na Figura 1.7.

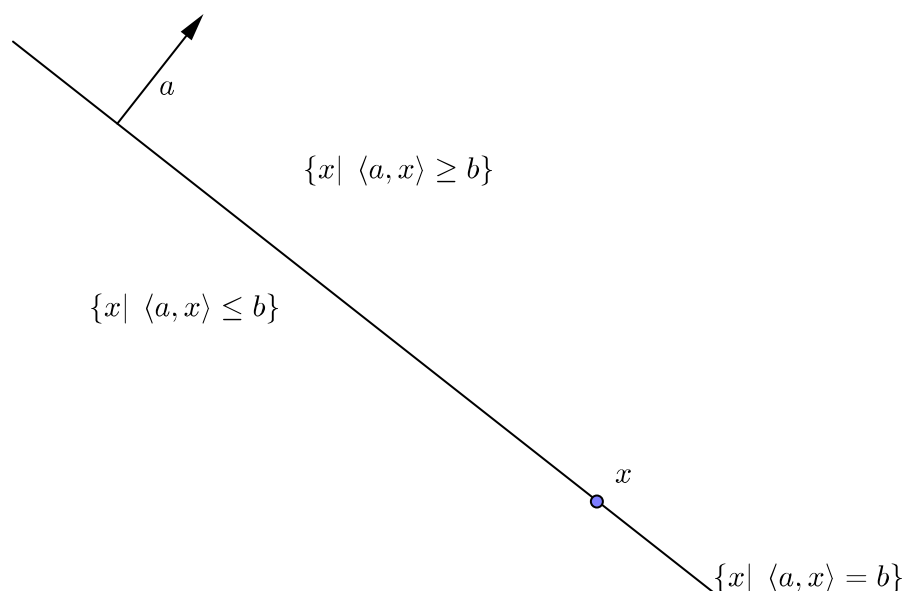


Figura 1.7: O hiperplano e seus semi-espacos associados.

Definição 1.33 Dados dois subconjuntos K_1, K_2 de \mathbb{R}^n . Dizemos que o hiperplano $H = \{x \mid \langle a, x \rangle = b\}$ separa K_1 e K_2 quando estes conjuntos pertencem a diferentes semi-espacos associados com H , ou seja, quando

$$\langle a, x_1 \rangle \leq b \leq \langle a, x_2 \rangle \quad \forall x_1 \in K_1, \forall x_2 \in K_2.$$

Dizemos que H separa estritamente K_1 e K_2 quando

$$\langle a, x_1 \rangle < b < \langle a, x_2 \rangle \quad \forall x_1 \in K_1, \forall x_2 \in K_2.$$

No caso em que K_1 é um conjunto unitário $\{x^*\}$ que está contido no fecho de K_2 , H é chamado de *hiperplano suporte* de K_2 em x^* . Deste modo, dizer que existe um hiperplano suportando K_2 em x^* significa que existe $a \neq 0$ tal que

$$\langle a, x^* \rangle \leq \langle a, x \rangle$$

para todo $x \in K_2$.

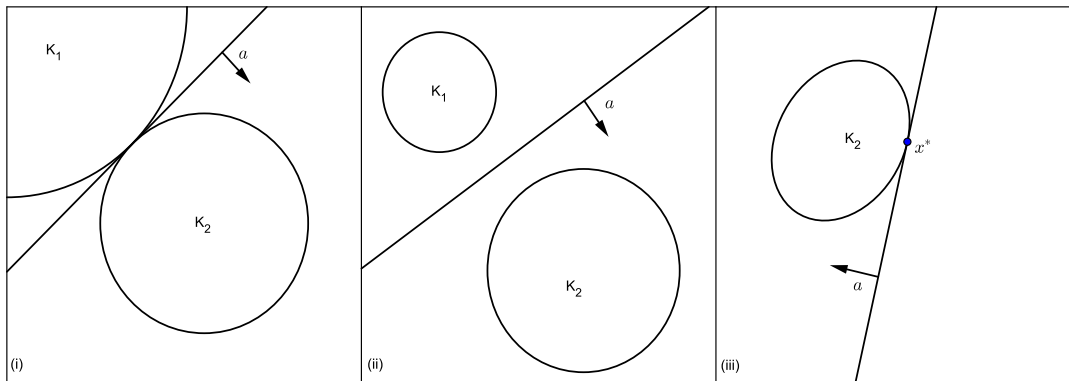


Figura 1.8: Separação de conjuntos por um hiperplano.

A existência desses hiperplanos está condicionada a algumas propriedades dos conjuntos K_1 e K_2 , como convexidade e compacidade. Antes de tratar desse tema, vamos rever alguns tópicos de projeções.

1.4.1 Projeção em conjuntos convexos e fechados

Dado um subconjunto $K \subset \mathbb{R}^n$ e um ponto $x \in \mathbb{R}^n$, vamos analisar o problema de encontrar um ponto de K que esteja mais próximo de x . Veremos que se K é fechado, então este problema tem solução. Caso K seja fechado e convexo, então a solução será única e chamada de projeção de x em K e denotada por $proj_K x$.

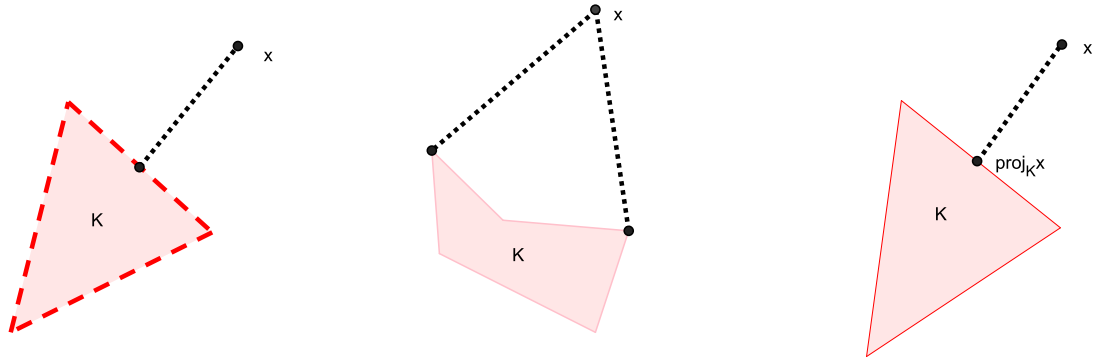


Figura 1.9: Ilustra o problema de encontrar um ponto de K mais próximo de x .

Lema 1.34 *Seja K um subconjunto não vazio e fechado de \mathbb{R}^n . Dado $x \in \mathbb{R}^n$, existe $\bar{x} \in K$ tal que*

$$\|x - \bar{x}\| \leq \|x - y\|$$

para todo $y \in K$.

Demonstração. Consideremos $\delta = \inf \{\|x - y\| \mid y \in K\}$. Então, para cada $k \in \mathbb{N}$, existe $y_k \in K$ tal que

$$\delta \leq \|x - y_k\| \leq \delta + \frac{1}{k}. \quad (1.6)$$

Em particular, $\|x - y_k\| \leq \delta + 1$ para todo k e assim, y_k é uma sequência limitada, portanto admite subsequência (y_{k_j}) convergente.

Definimos $\bar{x} := \lim_{j \rightarrow \infty} y_{k_j}$. Como K é fechado, então $\bar{x} \in K$. Além disso, $\|x - y_{k_j}\| \rightarrow \|x - \bar{x}\|$. Mas por (1.6), $\|x - y_{k_j}\| \rightarrow \delta$ e portanto $\delta = \|x - \bar{x}\|$, completando a prova. \square

Notemos que este lema garante a existência de solução do nosso problema. Como estamos trabalhando em \mathbb{R}^n e com a norma provida de um produto interno, se K é convexo, podemos caracterizar $proj_K x$ da seguinte forma.

Proposição 1.35 *Seja K um subconjunto não-vazio, fechado e convexo de \mathbb{R}^n . Dado $x \in \mathbb{R}^n$, \bar{x} é projeção de x em K se, e somente se, $\bar{x} \in K$ e*

$$\langle x - \bar{x}, y - \bar{x} \rangle \leq 0$$

para todo $y \in K$.

Demonstração. (\Rightarrow) Diretamente da definição de projeção, $\bar{x} \in K$. Além disso, dado $t \in (0, 1)$, pela convexidade de K , temos que $(1 - t)\bar{x} + ty \in K$ para todo $y \in K$. Portanto,

$$\|x - \bar{x}\| \leq \|x - (1 - t)\bar{x} - ty\| = \|x - \bar{x} + t(\bar{x} - y)\|.$$

Assim,

$$\|x - \bar{x}\|^2 \leq \|x - \bar{x} + t(\bar{x} - y)\|^2 = \|x - \bar{x}\|^2 + 2t \langle x - \bar{x}, \bar{x} - y \rangle + t^2 \|\bar{x} - y\|^2.$$

Como $t > 0$, então

$$2 \langle x - \bar{x}, -(\bar{x} - y) \rangle \leq t \|\bar{x} - y\|^2.$$

Fazendo $t \rightarrow 0$, obtemos $\langle x - \bar{x}, y - \bar{x} \rangle \leq 0$ para qualquer $y \in K$.

(\Leftarrow) Dado $y \in K$, temos que

$$\|x - \bar{x}\|^2 - \|x - y\|^2 = \langle x, x \rangle - 2 \langle x, \bar{x} \rangle + \langle \bar{x}, \bar{x} \rangle - \langle x, x \rangle + 2 \langle x, y \rangle - \langle y, y \rangle =$$

$$\begin{aligned}
&= -2 \langle x, \bar{x} \rangle + \langle \bar{x}, \bar{x} \rangle + 2 \langle x, y \rangle - \langle y, y \rangle = \langle y - \bar{x}, 2x - y - \bar{x} \rangle = \\
&= \langle y - \bar{x}, 2(x - \bar{x}) - (y - \bar{x}) \rangle = \langle y - \bar{x}, 2(x - \bar{x}) \rangle - \|y - \bar{x}\|^2 \leq 0.
\end{aligned}$$

Deste modo $\|x - \bar{x}\| \leq \|x - y\|$ para todo $y \in K$, e por definição, $\bar{x} = \text{proj}_K x$. \square

Para encerrar a discussão sobre projeções, vamos provar que para conjuntos convexos e fechados, a projeção é única.

Teorema 1.36 *Seja K um subconjunto não vazio, convexo e fechado de \mathbb{R}^n . Dado $x \in \mathbb{R}^n$, existe um único $\bar{x} \in K$ tal que*

$$\|x - \bar{x}\| \leq \|x - y\| \quad (1.7)$$

para todo $y \in K$.

Demonstração. A existência de \bar{x} é garantida pelo Lema 1.34. Para provar a unicidade, suponha que existam x_1 e x_2 em K que satisfaçam a relação (1.7). Pela Proposição 1.35, temos que

$$\langle x - x_1, x_2 - x_1 \rangle \leq 0 \quad (1.8)$$

e

$$\langle x - x_2, x_1 - x_2 \rangle \leq 0. \quad (1.9)$$

Equivalentemente a (1.9), $\langle x_2 - x, x_2 - x_1 \rangle \leq 0$. Somando esta expressão com (1.8), obtemos

$$\langle x_2 - x_1, x_2 - x_1 \rangle = \|x_2 - x_1\|^2 \leq 0.$$

Logo, $\|x_2 - x_1\| = 0$ e assim, $x_2 = x_1$.

1.4.2 Teoremas de Separação

Antes de dar início a discussão sobre os teoremas de separação, lembramos que um ponto x pertence ao interior de um subconjunto de \mathbb{R}^n quando é centro de uma bola aberta contida neste subconjunto. Quando x não pertence ao interior de um subconjunto convexo K de \mathbb{R}^n , x não pertence também ao interior de \bar{K} . Dessa maneira, é possível construir uma sequência de pontos que não estão em \bar{K} e que convergem para x .

Proposição 1.37 *Seja K um subconjunto convexo não-vazio de \mathbb{R}^n e x^* um ponto que não pertence ao interior de K . Então existe um hiperplano que passa por x^* e contém K em um dos seus semi-espacos, ou seja, existe um vetor $a \neq 0$ tal que*

$$\langle a, x^* \rangle \leq \langle a, x \rangle$$

para todo $x \in K$.

Demonstração. Como x^* não pertence ao interior de K , então existe uma sequência (x_k) convergente para x^* tal que $x_k \notin \bar{K}$ para todo k . Como K é convexo, então \bar{K} também é. Deste modo, seja \bar{x}_k a projeção de x_k em \bar{K} . Pela Proposição 1.35, temos que $\langle \bar{x}_k - x_k, y - \bar{x}_k \rangle \geq 0$ para todo $y \in \bar{K}$. Deste modo,

$$\begin{aligned}
\langle \bar{x}_k - x_k, y \rangle &\geq \langle \bar{x}_k - x_k, \bar{x}_k \rangle = \langle \bar{x}_k - x_k, \bar{x}_k - x_k \rangle + \langle \bar{x}_k - x_k, x_k \rangle = \\
&= \|\bar{x}_k - x_k\|^2 + \langle \bar{x}_k - x_k, x_k \rangle \geq \langle \bar{x}_k - x_k, x_k \rangle.
\end{aligned}$$

Definindo

$$a_k := \frac{\bar{x}_k - x_k}{\|\bar{x}_k - x_k\|},$$

obtemos

$$\langle a_k, x_k \rangle \leq \langle a_k, y \rangle \tag{1.10}$$

para todo $y \in \bar{K}$ e $k \in \mathbb{N}$. Como $\|a_k\| = 1$ para todo k , então a_k admite uma subsequência a_{k_j} que converge para um vetor não nulo a e além disso, $\langle a_{k_j}, x_{k_j} \rangle \leq \langle a_{k_j}, y \rangle$ para todo $y \in \bar{K}$ e $j \in \mathbb{N}$. Fazendo j tender para o infinito, concluímos a demonstração. \square

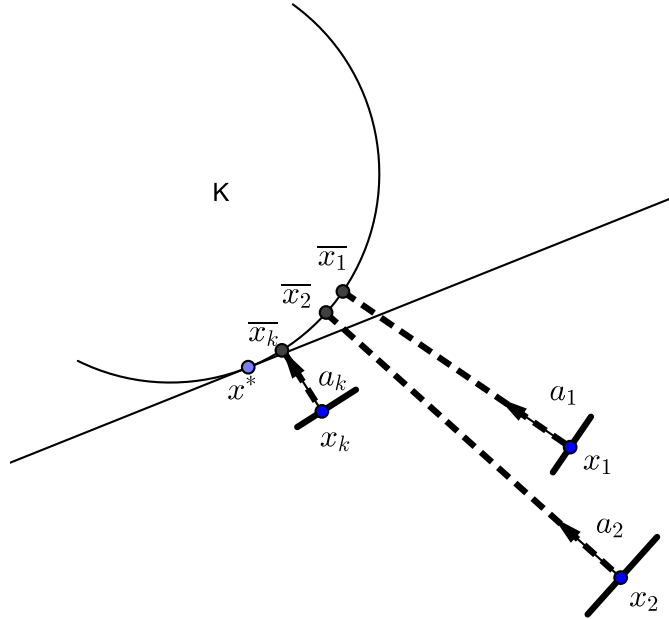


Figura 1.10: Ilustra a demonstração da Proposição 1.37 para o caso em que $x^* \in \bar{K}$.

Para o caso em que $x^* \in \bar{K}$, então o hiperplano referente à proposição suporta K em x^* . Com o auxílio desse resultado, podemos enunciar o primeiro teorema de separação.

Teorema 1.38 (Teorema de separação para conjuntos convexos) *Sejam K_1 e K_2 dois subconjuntos convexos não-vazios de \mathbb{R}^n . Se K_1 e K_2 são disjuntos, há um hiperplano que os separa, ou seja, existe um vetor $a \neq 0$ tal que*

$$\langle a, x_1 \rangle \leq \langle a, x_2 \rangle$$

para todo $x_1 \in K_1$ e $x_2 \in K_2$.

Demonstração. Considere o conjunto convexo

$$K = K_2 - K_1 = \{x \mid x = x_2 - x_1, x_1 \in K_1, x_2 \in K_2\}.$$

Como K_1 e K_2 são disjuntos, a origem não pertence a K . Assim, pela Proposição 1.37, há um vetor $a \neq 0$ tal que $0 \leq \langle a, x \rangle$ para todo $x \in K$. Portanto,

$$0 \leq \langle a, x_2 - x_1 \rangle \Rightarrow \langle a, x_1 \rangle \leq \langle a, x_2 \rangle$$

para todo $x_1 \in K_1$ e $x_2 \in K_2$. \square

Deste modo, a convexidade dos subconjuntos já garante a existência de um hiperplano que os separa. Veremos a seguir que se a diferença entre os subconjuntos é fechada, então eles podem ser separados estritamente por um hiperplano.

Proposição 1.39 *Seja K um subconjunto não-vazio, convexo e fechado de \mathbb{R}^n . Se K não contém a origem então existem $a \neq 0$ e $\alpha > 0$ tais que*

$$\alpha \leq \langle a, x \rangle$$

para todo $x \in K$.

Demonstração. Seja $\bar{z} = \text{proj}_K(0)$. Pela Proposição 1.35, temos que $0 \leq \langle \bar{z}, x - \bar{z} \rangle$ para todo $x \in K$. Assim

$$\langle \bar{z}, x \rangle \geq \langle \bar{z}, \bar{z} \rangle = \|\bar{z}\|^2.$$

Como $\bar{z} \neq 0$, pois $0 \notin K$, podemos tomar $\alpha = \|\bar{z}\|^2$ e $a = \bar{z}$ para concluir a demonstração.

□

Teorema 1.40 (Separação forte) *Sejam K_1, K_2 subconjuntos convexos não-vazios de \mathbb{R}^n tais que $K_1 \cap K_2 = \emptyset$ e $K_2 - K_1$ seja fechado. Então existem $a \in \mathbb{R}^n - \{0\}$ e $b \in \mathbb{R}$ tais que*

$$\langle a, x_1 \rangle < b < \langle a, x_2 \rangle \quad \forall x_1 \in K_1, \forall x_2 \in K_2.$$

Demonstração. Considere $K = K_2 - K_1$. Deste modo, K é um subconjunto convexo fechado que não contém a origem. Pela Proposição 1.39, existem $a \neq 0$ e $\alpha > 0$ tais que $\alpha \leq \langle a, x \rangle$ para todo $x \in K$. Logo,

$$\langle a, x_2 \rangle \geq \alpha + \langle a, x_1 \rangle$$

para todo $x_1 \in K_1$ e $x_2 \in K_2$. Portanto,

$$\inf_{x_2 \in K_2} \langle a, x_2 \rangle \geq \alpha + \sup_{x_1 \in K_1} \langle a, x_1 \rangle > \sup_{x_1 \in K_1} \langle a, x_1 \rangle.$$

Tomando

$$b = \frac{1}{2} \left(\inf_{x_2 \in K_2} \langle a, x_2 \rangle + \sup_{x_1 \in K_1} \langle a, x_1 \rangle \right),$$

concluimos a demonstração. □

Notemos que para dois subconjuntos convexos disjuntos tais que um deles é fechado e outro é compacto, então a diferença também é fechada e portanto existe um hiperplano que os separa estritamente. Este fato nos permite concluir o último resultado da seção.

Corolário 1.41 *Um subconjunto convexo fechado de \mathbb{R}^n é a interseção de semi-espacos que o contém.*

Demonstração. Seja K um convexo fechado e $x \notin K$. Como K e $\{x\}$ são convexos e além disso, K é fechado e $\{x\}$ é compacto. Então existem $a_x \neq 0$ e $b_x \in \mathbb{R}$ tais que

$$\langle a_x, x \rangle < b_x < \langle a_x, y \rangle$$

para todo $y \in K$. Assim $K \subset \{y \mid b_x \leq \langle a_x, y \rangle\}$ para qualquer $x \notin K$. Portanto

$$K \subset \bigcap_{x \notin K} \{y \mid b_x \leq \langle a_x, y \rangle\}.$$

Por outro lado, se $z \notin K$ então $z \notin \{y \mid b_x \leq \langle a_x, y \rangle\}$, logo $z \notin \bigcap_{x \notin K} \{y \mid b_x \leq \langle a_x, y \rangle\}$. Deste modo

$$K = \bigcap_{x \notin K} \{y \mid b_x \leq \langle a_x, y \rangle\}.$$

□

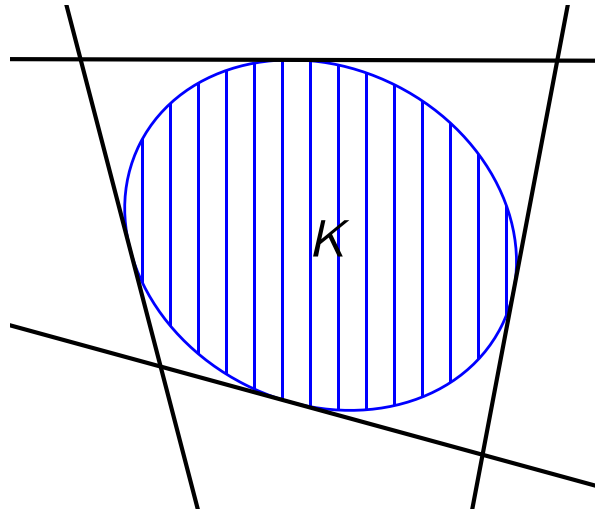


Figura 1.11: Ilustração do Corolário 1.41.

1.5 Funções conjugadas

Como vimos na seção anterior, um conjunto convexo fechado é a interseção dos semi-espacos que o contêm. Lembramos que o epigrafo de uma função própria convexa f é um subconjunto convexo e fechado de \mathbb{R}^{n+1} . Dessa forma, ele pode ser encarado como a interseção dos semi-espacos associados aos seus hiperplanos suporte. Em Análise convexa, esta outra forma de caracterizar o epigrafo de uma função nos permite entender a noção de conjugação.

Consideremos um hiperplano em \mathbb{R}^{n+1} da forma

$$H = \{(x, \lambda) \mid \langle a, x \rangle + a_{n+1}\lambda = b, (a, a_{n+1}) \in \mathbb{R}^{n+1}, b \in \mathbb{R}\}.$$

Para efeitos da ilustração gráfica, vamos considerar o eixo horizontal correspondente ao \mathbb{R}^n e o eixo vertical correspondente à \mathbb{R} . Como estamos interessados em funções próprias cujos epigrafos não contêm hiperplanos paralelos ao eixo vertical, vamos desconsiderar o caso em que $a_{n+1} = 0$. Assim, dado um hiperplano, podemos multiplicar a sua equação pelo inverso de $-a_{n+1}$, obtendo a expressão

$$\left\langle \frac{a}{-a_{n+1}}, x \right\rangle - \lambda = \frac{b}{-a_{n+1}}.$$

Sem perda de generalidade, podemos renomear $a = a/(-a_{n+1})$ e $b = b/(-a_{n+1})$. Deste modo, o hiperplano em \mathbb{R}^{n+1} será caracterizado por

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

Podemos reescrever a expressão do hiperplano como $\lambda = \langle a, x \rangle - b$. Assim, ela representa uma reta normal ao vetor $(a, 1)$ e que passa pelo ponto $(0, -b)$. Graficamente, temos:

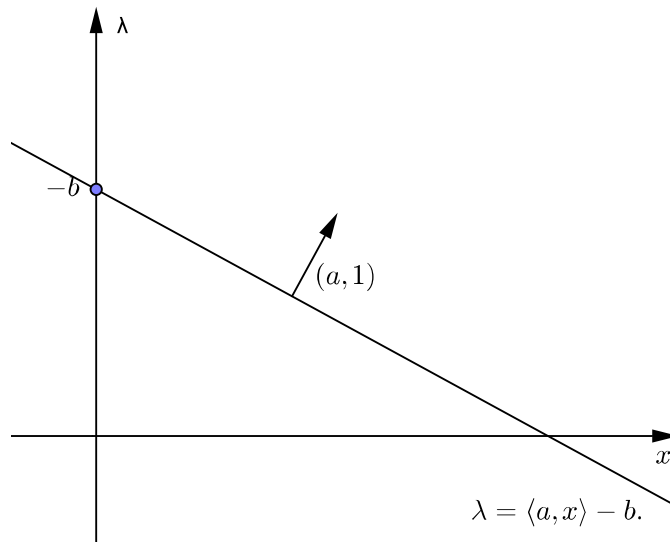


Figura 1.12: Representação gráfica do hiperplano $\lambda = \langle a, x \rangle - b$.

Dada uma função $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ sci e convexa e dado $a \in \mathbb{R}^n$, queremos determinar b de modo que a reta $\lambda = \langle a, x \rangle - b$ seja tangente ao gráfico de f , ou seja, que o hiperplano $H = \{(x, \lambda); \langle a, x \rangle - \lambda = b, a \in \mathbb{R}^n, b \in \mathbb{R}\}$ seja suporte do $\text{epi}(f)$. Na Figura 1.13, ilustramos essa situação.

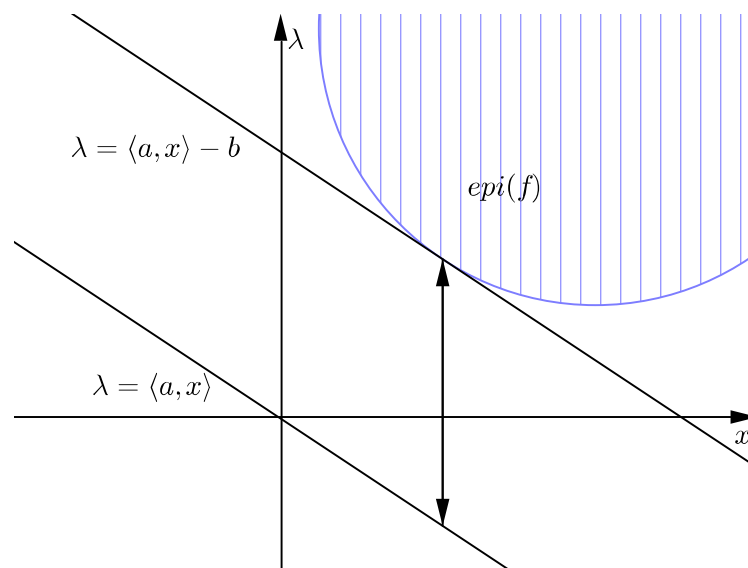


Figura 1.13: Reta $\lambda = \langle a, x \rangle - b$ tangente ao gráfico de f .

Como podemos observar pela Figura 1.13, se tomarmos $-b$ como a distância vertical do gráfico de f com a reta $\lambda = \langle a, x \rangle$, ou seja,

$$-b = \inf_{x \in \mathbb{R}^n} \{f(x) - \langle a, x \rangle\}$$

ou, equivalentemente,

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

então o hiperplano H será suporte do $\text{epi}(f)$ no ponto em que tangencia o gráfico da f . Podemos notar que b está em função de a . Essa discussão geométrica nos motiva a dar a definição de funções conjugadas.

Definição 1.42 Dada uma função $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ própria, a função conjugada de f , denotada por f_F^* , é definida como $f_F^* : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ sendo

$$f_F^*(a) = \sup_{x \in \mathbb{R}^n} \{\langle a, x \rangle - f(x)\}.$$

Notemos que f_F^* está bem definida, pois como f é própria, existe um $x \in \text{dom}(f)$ tal que $\langle a, x \rangle - f(x) > -\infty$ e portanto $f_F^*(a) > -\infty$ para todo $a \in \mathbb{R}^n$. Essa função é também conhecida na literatura como Conjugada de Fenchel ou Conjugada clássica. Notemos que esta definição é aplicada para qualquer função própria. No entanto, pelo que vimos anteriormente, para uma função f convexa e sci, o hiperplano

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

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

Para cada $x \in \mathbb{R}^n$, consideramos $f_x : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{-\infty\}$ definido como $f_x(a) = \langle a, x \rangle - f(x)$. Assim, temos que f_x é convexa e sci para todo x e, além disso, $f_F^*(a) = \sup \{f_x(a); x \in \mathbb{R}^n\}$. Pelo Teorema 1.2, a função conjugada será sempre convexa e sci. Observemos o próximo exemplo.

Exemplo 1.43 $f : \mathbb{R} \rightarrow \mathbb{R}; f(x) = e^x$.

Consideremos o caso em que $a < 0$. Assim, a expressão $ax - e^x$ cresce infinitamente quando x cresce no sentido negativo, portanto $f_F^*(a) = +\infty$.

Para o caso em que $a = 0$, temos que

$$f_F^*(0) = \sup_{x \in \mathbb{R}^n} \{-e^x\} = 0.$$

Finalmente, para o caso em que $a > 0$, consideremos $h(x) = ax - e^x$. Notemos que $h'(x) = a - e^x$, de modo que o único ponto crítico de h é $\bar{x} = \ln a$. Além disso, como $h''(x) = -e^x < 0$, pelas condições de otimalidade², \bar{x} é maximizador de h . Assim,

$$f_F^*(a) = a \ln a - e^{\ln a} = a \ln a - a.$$

Portanto

$$f_F^*(a) = \begin{cases} a \ln a - a, & a > 0 \\ 0, & a = 0 \\ +\infty, & a < 0 \end{cases}.$$

Veremos agora algumas propriedades das funções conjugadas.

Teorema 1.44 Dada uma função f própria e f^* sua conjugada. Então

- (i) $f(x) + f_F^*(a) \geq \langle a, x \rangle$ para todos x e a pertencentes a \mathbb{R}^n .
- (ii) $\inf_{x \in \mathbb{R}^n} f(x) = -f_F^*(0)$.
- (iii) se $f(x) \leq g(x)$ para todo $x \in \mathbb{R}^n$ então $f_F^*(a) \geq g_F^*(a)$ para todo $a \in \mathbb{R}^n$.

²As condições de otimalidade são condições que a função deve satisfazer para que um ponto crítico seja considerado minimizador ou maximizador. Para mais informações, ver [15]

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

(ii) Por definição de função conjugada, temos que

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

(iii) Se $f(x) \leq g(x)$ para todo $x \in \mathbb{R}^n$ então

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

para todos x e a pertencentes a \mathbb{R}^n . Logo

$$f_F^*(a) = \sup_x \{\langle a, x \rangle - f(x)\} \geq \sup_x \{\langle a, x \rangle - g(x)\} = g_F^*(a)$$

para todo $a \in \mathbb{R}^n$. □

É importante notar que os pontos em que f atinge $+\infty$ não influenciam no cálculo da conjugada de f , pois a expressão $\langle a, x \rangle - f(x)$ 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 sob o domínio de f , ou seja,

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

É possível obter a conjugada da função conjugada, denominada biconjugada e denotada por f_F^{**} . Ela é definida de maneira análoga à função conjugada, ou seja, $f_F^{**} : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$;

$$f_F^{**}(x) = (f_F^*)^*(x) = \sup_{a \in \mathbb{R}^n} \{\langle a, x \rangle - f_F^*(a)\}.$$

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

Teorema 1.45 *Seja $G = \{g : \mathbb{R}^n \rightarrow \mathbb{R} \mid g(y) = \langle a, y \rangle + b, a \in \mathbb{R}^n, b \in \mathbb{R}\}$. Considere-mos f uma função própria, então*

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

Demonstração. Dado $x \in \mathbb{R}^n$, consideremos $s = \sup_{g \in G} \{g(x) \mid g(y) \leq f(y), \forall y \in \mathbb{R}^n\}$. Definimos

$$\bar{g}(y) = \langle a, y \rangle - f_F^*(a)$$

para $a \in \mathbb{R}^n$. Segue do item (i) do Teorema 1.44 que $\bar{g}(y) \leq f(y)$ para todo $y \in \mathbb{R}^n$. Ainda, $\bar{g} \in G$ e assim,

$$s \geq \bar{g}(x) = \langle a, x \rangle - f_F^*(a)$$

para $a \in \mathbb{R}^n$. Tomando o supremo em a , temos que $f_F^{**}(x) \leq s$.

Suponha por absurdo que $f_F^{**}(x) < s$. Pela propriedade de supremo, existem $a \in \mathbb{R}^n$ e $b \in \mathbb{R}$ tais que $\langle a, x \rangle + b > f_F^{**}(x)$ e, para todo $y \in \mathbb{R}^n$,

$$\langle a, y \rangle + b \leq f(y). \tag{1.11}$$

Reescrevendo (1.11),

$$\langle a, y \rangle - f(y) \leq -b.$$

Tomando o supremo em y , obtemos $f_F^*(a) \leq -b$. Concluimos assim que

$$f_F^{**}(x) < \langle a, x \rangle + b \leq \langle a, x \rangle - f_F^*(a) \leq f_F^{**}(x).$$

Absurdo. Logo

$$f_F^{**}(x) = s = \sup_{g \in G} \{g(x) \mid g(y) \leq f(y), \forall y \in \mathbb{R}^n\}$$

para todo $x \in \mathbb{R}^n$. □

Como corolário direto, temos a seguinte desigualdade.

Corolário 1.46 *Dada uma função própria f e sua biconjugada f_F^{**} então*

$$f_F^{**}(x) \leq f(x)$$

para todo $x \in \mathbb{R}^n$.

Estamos interessados em investigar quando é satisfeita a igualdade no corolário anterior. Vamos rever o Exemplo 1.43.

A função f é definida por $f(x) = e^x$ e sua conjugada é

$$f_F^*(a) = \begin{cases} a \ln a - a, & a > 0 \\ 0, & a = 0 \\ +\infty, & a < 0 \end{cases}.$$

Vamos calcular a biconjugada de f . Deste modo, dado $x \in \mathbb{R}^n$, temos que

$$f_F^{**}(x) = \sup_{a \in \mathbb{R}} \{ax - f_F^*(a)\} = \sup_{a \geq 0} \{ax - f_F^*(a)\}.$$

Consideremos $h : (0, +\infty) \rightarrow \mathbb{R}$ definida por $h(a) = ax - a \ln a + a$. Assim,

$$h'(a) = x - \ln a.$$

Consequentemente, o único ponto crítico de h é $\bar{a} = e^x$. Como $h''(a) = -1/a < 0$, \bar{a} é maximizador de h e o valor máximo de h é e^x .

Notemos que a expressão $ax - f_F^*(a)$ vale 0 no caso em que $a = 0$. Portanto,

$$f_F^{**}(x) = \sup \{0, e^x\} = e^x = f(x)$$

para todo $x \in \mathbb{R}$.

Neste caso, a função f é convexa e contínua. Devido ao Teorema de Separação Forte, a teoria de conjugação se comporta melhor com funções convexas e sci, deste modo, é esperado que a biconjugada dessas funções tenham propriedades mais interessantes, como veremos nos próximos teoremas.

Proposição 1.47 *Se f é uma função convexa, sci e própria então a sua função conjugada f_F^* também é própria.*

Demonstração. Seja $x_0 \in \text{dom}(f)$ e $\lambda_0 < f(x_0)$, ou seja, $(x_0, \lambda_0) \notin \text{epi}(f)$. Consideremos $K_1 = \{(x_0, \lambda_0)\}$ e $K_2 = \text{epi}(f)$. Assim K_1 é não-vazio, convexo e compacto e, desde que f é própria, convexa e sci, K_2 é não-vazio, convexo e fechado. Portanto $K_2 - K_1$ é fechado. Pelo Teorema 1.40, existem $a \in \mathbb{R}^n - \{0\}$, números reais a_{n+1} e b tais que

$$\langle a, x_0 \rangle + a_{n+1} \lambda_0 < b,$$

$$b < \langle a, x \rangle + a_{n+1}\lambda$$

para todo $(x, \lambda) \in \text{epi}(f)$. Em particular, para todo $x \in \text{dom}(f)$, temos

$$b < \langle a, x \rangle + a_{n+1}f(x) \quad (1.12)$$

e logo, $\langle a, x_0 \rangle + a_{n+1}\lambda_0 < \langle a, x \rangle + a_{n+1}f(x)$. Quando $x = x_0$ temos que $a_{n+1}(f(x_0) - \lambda_0) > 0$, então $a_{n+1} > 0$. De (1.12),

$$\left\langle -\frac{a}{a_{n+1}}, x \right\rangle - f(x) < -\frac{b}{a_{n+1}}$$

para todo $x \in \text{dom}(f)$ e portanto

$$f_F^*\left(-\frac{a}{a_{n+1}}\right) = \sup_{x \in \text{dom}(f)} \left\{ \left\langle -\frac{a}{a_{n+1}}, x \right\rangle - f(x) \right\} \leq -\frac{b}{a_{n+1}} < +\infty.$$

Assim, o $\text{dom}(f_F^*)$ é não vazio. □

Observamos que f_F^* é uma função convexa e sci. Assim, quando f é convexa, sci e própria tanto f_F^* quanto f_F^{**} são próprias. Para finalizar, apresentamos o principal resultado da seção.

Teorema 1.48 (Teorema de Fenchel-Moreau) *Uma função própria f é convexa e sci se, e somente se*

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

para todo $x \in \mathbb{R}^n$.

Demonstração. (\Leftarrow) Diretamente da definição, podemos concluir que f_F^{**} é convexa e sci. Desde que $f(x) = f_F^{**}(x)$ para todo $x \in \mathbb{R}^n$, então f também é convexa e sci.

(\Rightarrow) Vamos inicialmente supor que $f(x) \geq 0$ para todo $x \in \mathbb{R}^n$. Pelo Corolário 1.46, temos que $f_F^{**}(x) \leq f(x)$ para todo $x \in \mathbb{R}^n$. Suponhamos por absurdo que existe um x_0 tal que

$$f_F^{**}(x_0) < f(x_0).$$

Claramente $f_F^{**}(x_0) < +\infty$ e $(x_0, f_F^{**}(x_0)) \notin \text{epi}(f)$. Consideremos $K_1 = \{(x_0, f_F^{**}(x_0))\}$ e $K_2 = \text{epi}(f)$. Ambos são não-vazios e convexos, K_1 é compacto e K_2 é fechado. Pelo Teorema 1.40, existem $a \in \mathbb{R}^n - \{0\}$ e números reais a_{n+1} e $b \in \mathbb{R}$ tais que

$$\langle a, x_0 \rangle + a_{n+1}f_F^{**}(x_0) < b, \quad (1.13)$$

$$b < \langle a, x \rangle + a_{n+1}\lambda \quad (1.14)$$

para todo $(x, \lambda) \in \text{epi}(f)$.

Notemos que para $(\bar{x}, \bar{\lambda}) \in \text{epi}(f)$, teremos que $(\bar{x}, \lambda) \in \text{epi}(f)$ para todo $\lambda \geq \bar{\lambda}$. Assim,

$$\frac{b}{\lambda} < \frac{\langle a, \bar{x} \rangle}{\lambda} + a_{n+1}$$

para todo λ positivo e maior ou igual que $\bar{\lambda}$. Desde modo, fazendo λ crescer infinitamente, concluimos que $a_{n+1} \geq 0$. Dado $\epsilon > 0$, sendo $f(x) \geq 0$ para todo $x \in \mathbb{R}^n$, temos que

$$\langle a, x \rangle + a_{n+1}\lambda + \epsilon f(x) > b + \epsilon f(x) \geq b$$

para todo $(x, \lambda) \in \text{epi}(f)$, ou seja

$$\langle a, x \rangle + (a_{n+1} + \epsilon)f(x) \geq b$$

para todo $x \in \text{dom}(f)$. Como $(a_{n+1} + \epsilon) > 0$, então

$$\left\langle -\frac{a}{a_{n+1} + \epsilon}, x \right\rangle - f(x) \leq -\frac{b}{a_{n+1} + \epsilon}$$

para $x \in \text{dom}(f)$. Por propriedade de supremo, temos

$$f_F^*\left(-\frac{a}{a_{n+1} + \epsilon}\right) \leq \sup_{x \in \text{dom}(f)} \left\{ \left\langle -\frac{a}{a_{n+1} + \epsilon}, x \right\rangle - f(x) \right\} \leq -\frac{b}{a_{n+1} + \epsilon}.$$

Da definição de biconjugada, obtemos

$$f_F^{**}(x_0) \geq \left\langle -\frac{a}{a_{n+1} + \epsilon}, x_0 \right\rangle - f_F^*\left(-\frac{a}{a_{n+1} + \epsilon}\right) \geq \left\langle -\frac{a}{a_{n+1} + \epsilon}, x_0 \right\rangle + \frac{b}{a_{n+1} + \epsilon}.$$

Por fim

$$\langle a, x_0 \rangle + (a_{n+1} + \epsilon)f_F^{**}(x_0) \geq b$$

para todo $\epsilon > 0$. Quando fazemos ϵ tender a 0, concluímos que

$$\langle a, x_0 \rangle + a_{n+1}f_F^{**}(x_0) \geq b.$$

Absurdo, pois contradiz (1.13). Portanto $f_F^{**}(x) = f(x)$ para todo $x \in \mathbb{R}^n$.

Vamos agora considerar f arbitrária. Pela Proposição 1.47, sabemos que $\text{dom}(f_F^*) \neq \emptyset$. Assim, consideremos $a_0 \in \text{dom}(f_F^*)$ e, para voltarmos ao caso anterior, definimos a função

$$\bar{f}(x) = f(x) - \langle a_0, x \rangle + f_F^*(a_0).$$

Notemos que \bar{f} é convexa e sci, desde que ela é expressa como soma de funções convexas e semicontínuas inferiormente. Além disso, \bar{f} é própria, pois f é própria. Como $f_F^*(a_0) = \sup_{x \in \mathbb{R}^n} \{\langle a_0, x \rangle - f(x)\}$ então $\bar{f}(x) \geq 0$ para todo $x \in \mathbb{R}^n$. Deste modo, sabemos que $\bar{f}(x) = (\bar{f})^{**}(x)$ para todo $x \in \mathbb{R}^n$.

Calculando a $(\bar{f})^*$ e a $(\bar{f})^{**}$, obtemos

$$\begin{aligned} (\bar{f})^*(a) &= \sup_{x \in \mathbb{R}^n} \{\langle a, x \rangle - \bar{f}(x)\} = \sup_{x \in \mathbb{R}^n} \{\langle a, x \rangle - f(x) + \langle a_0, x \rangle - f_F^*(a_0)\} = \\ &= \sup_{x \in \mathbb{R}^n} \{\langle a + a_0, x \rangle - f(x)\} - f_F^*(a_0) = f_F^*(a + a_0) - f_F^*(a_0). \\ (\bar{f})^{**}(x) &= \sup_{a \in \mathbb{R}^n} \{\langle a, x \rangle - (\bar{f})^*(a)\} = \sup_{a \in \mathbb{R}^n} \{\langle a, x \rangle - f_F^*(a + a_0) + f_F^*(a_0)\} = \\ &= \sup_{a \in \mathbb{R}^n} \{\langle a + a_0, x \rangle - f_F^*(a + a_0) - \langle a_0, x \rangle\} + f_F^*(a_0) \end{aligned}$$

em que a última igualdade resulta do acréscimo de $\langle a_0, x \rangle - \langle a_0, x \rangle$ no argumento do supremo. Tomando $z = a + a_0$, temos que

$$(\bar{f})^{**}(x) = \sup_{z \in \mathbb{R}^n} \{\langle z, x \rangle - f_F^*(z)\} - \langle a_0, x \rangle + f_F^*(a_0) = f_F^{**}(x) - \langle a_0, x \rangle + f_F^*(a_0).$$

Como $\bar{f}(x) = (\bar{f})^{**}(x)$ para todo $x \in \mathbb{R}^n$, então

$$f(x) - \langle a_0, x \rangle + f_F^*(a_0) = \bar{f}(x) = (\bar{f})^{**}(x) = f_F^{**}(x) - \langle a_0, x \rangle + f_F^*(a_0)$$

e assim $f(x) = f_F^{**}(x)$ para todo $x \in \mathbb{R}^n$. □

1.6 Dualidade convexa

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.

Para encerrar o capítulo, usaremos a teoria de conjugação para desenvolver uma teoria de dualidade cujas relações mais fortes são obtidas quando o problema primal é de minimização convexa. Mas mesmo em casos mais gerais, esta dualidade pode ser bastante útil.

Definição 1.49 *Uma função $\varphi : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R} \cup \{+\infty\}$ é dita perturbação de uma função $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ quando*

$$\varphi(x, 0) = f(x)$$

para todo $x \in \mathbb{R}^n$.

Notemos que para toda f existe uma função perturbação φ , pois podemos considerar $\varphi(x, y) = f(x)$ para todo $x \in \mathbb{R}^n$ e $y \in \mathbb{R}^m$. Veremos no próximo exemplo um caso mais interessante de função perturbação.

Exemplo 1.50 *Consideremos o problema de Programação Linear:*

$$\text{minimizar } \langle c, x \rangle$$

$$\text{sujeito a } x \in K = \{x \in \mathbb{R}^n \mid Ax = b \text{ e } x \geq 0\},$$

em que $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ e $A \in \mathbb{R}^{m \times n}$ são dados. A desigualdade em \mathbb{R}^n é no sentido de coordenada a coordenada, ou seja, quando dizemos que $x \geq 0$, estamos querendo dizer que cada coordenada de x é não negativa. O problema é equivalente a determinar

$$\inf_{x \in \mathbb{R}^n} f(x)$$

tal que

$$f(x) = \begin{cases} \langle c, x \rangle, & x \in K \\ +\infty, & x \notin K \end{cases}$$

Dessa forma, podemos considerar $\varphi : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R} \cup \{+\infty\}$ dada por

$$\varphi(x, y) = \begin{cases} \langle c, x \rangle, & Ax = b + y; x \geq 0 \\ +\infty, & \text{caso contrário} \end{cases}$$

Notemos que $\varphi(x, 0) = f(x)$ para todo $x \in \mathbb{R}^n$ e portanto φ é uma função perturbação de f .

Lembramos que o produto interno no espaço $\mathbb{R}^n \times \mathbb{R}^m$ para $z_1 = (x_1, y_1)$ e $z_2 = (x_2, y_2)$ com x_1 e x_2 em \mathbb{R}^n e y_1 e y_2 em \mathbb{R}^m é naturalmente estendido para

$$\langle z_1, z_2 \rangle = \langle x_1, x_2 \rangle_{\mathbb{R}^n} + \langle y_1, y_2 \rangle_{\mathbb{R}^m}.$$

Como f é própria, então φ também é. Assim, a conjugada φ_F^* é definida de maneira natural como $\varphi_F^* : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R} \cup \{+\infty\}$, sendo

$$\varphi_F^*(a_1, a_2) = \sup_{(x,y) \in \mathbb{R}^n \times \mathbb{R}^m} \{\langle a_1, x \rangle + \langle a_2, y \rangle - \varphi(x, y)\}$$

para a_1 em \mathbb{R}^n e a_2 em \mathbb{R}^m . A biconjugada φ_F^{**} é calculada como

$$\varphi_F^{**}(x, y) = \sup_{(a_1, a_2) \in \mathbb{R}^n \times \mathbb{R}^m} \{ \langle a_1, x \rangle + \langle a_2, y \rangle - \varphi_F^*(a_1, a_2) \}$$

para x em \mathbb{R}^n e y em \mathbb{R}^m .

Seguindo essa extensão, as propriedades da teoria de conjugação apresentadas na Seção 1.5 continuam válidas para a conjugação de φ .

Dada uma função própria $f : \mathbb{R}^n \rightarrow \mathbb{R}$, o problema de minimizar a f em \mathbb{R}^n é chamado de problema *primal* (P). Esquemáticamente, queremos

$$\text{minimizar } f(x)$$

$$\text{sujeito a } x \in \mathbb{R}^n,$$

ou equivalentemente, determinar o

$$\inf_{x \in \mathbb{R}^n} f(x) \quad (P).$$

Agora consideremos $\varphi : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função de perturbação de f . Para cada $y \in \mathbb{R}^m$, consideremos o problema

$$\inf_{x \in \mathbb{R}^n} \varphi(x, y),$$

denotado por P_y . Cada P_y é chamado de problema perturbado de P com respeito a função de perturbação φ . Para $y = 0$, temos o próprio problema primal.

O problema dual associado a P com respeito a φ é dado por

$$\sup_{a \in \mathbb{R}^m} -\varphi_F^*(0, a)$$

e denotado por P^* .

Voltando ao Exemplo 1.50, dado $a \in \mathbb{R}^m$, vamos calcular a $\varphi_F^*(0, a)$.

$$\begin{aligned} \varphi_F^*(0, a) &= \sup_{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m} \{ \langle (x, y), (0, a) \rangle - \varphi(x, y) \} = \sup_{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m} \{ \langle y, a \rangle - \varphi(x, y) \} = \\ &= \sup_{\substack{x \geq 0 \\ Ax = b + y}} \{ \langle y, a \rangle - \langle c, x \rangle \} = \sup_{x \geq 0} \{ \langle Ax - b, a \rangle - \langle c, x \rangle \} = \\ &= \sup_{x \geq 0} \{ \langle A^t a - c, x \rangle - \langle a, b \rangle \} = \sup_{x \geq 0} \{ \langle A^t a - c, x \rangle \} - \langle a, b \rangle. \end{aligned}$$

Portanto,

$$\varphi_F^*(0, a) = \begin{cases} -\langle a, b \rangle, & A^t a \leq c; \\ +\infty, & A^t a > c \end{cases}$$

Assim o problema dual associado à programação linear é dado por

$$\sup_{\{a \in \mathbb{R}^m; A^t a \leq c\}} \langle a, b \rangle,$$

ou seja,

$$\begin{aligned} &\text{maximizar } \langle a, b \rangle \\ &\text{sujeito a } a \in K^* = \{a \in \mathbb{R}^m; A^t a \leq c\}. \end{aligned}$$

Veremos a seguir algumas importantes relações entre os problemas primal e dual que mostram a grande relevância dessa teoria na Otimização.

Proposição 1.51 *Para qualquer par de problemas primal e dual, temos que*

$$-\varphi_F^*(0, a) \leq f(x)$$

para todo $x \in \mathbb{R}^n$ e $a \in \mathbb{R}^m$. Em particular,

$$\sup_{a \in \mathbb{R}^m} -\varphi_F^*(0, a) \leq \inf_{x \in \mathbb{R}^n} f(x).$$

Demonstração. Para $a \in \mathbb{R}^m$, temos, por definição, que

$$\varphi_F^*(0, a) = \sup_{(x,y) \in \mathbb{R}^n \times \mathbb{R}^m} \{ \langle (x, y), (0, a) \rangle - \varphi(x, y) \} = \sup_{(x,y) \in \mathbb{R}^n \times \mathbb{R}^m} \{ \langle y, a \rangle - \varphi(x, y) \}.$$

Deste modo,

$$\varphi_F^*(0, a) \geq \langle y, a \rangle - \varphi(x, y)$$

para todo $x \in \mathbb{R}^n$ e $y \in \mathbb{R}^m$. Em particular, para $y = 0$ temos que

$$-\varphi_F^*(0, a) \leq \varphi(x, 0) = f(x),$$

concluindo assim a demonstração. □

Segue diretamente o próximo corolário.

Corolário 1.52 *Dada uma função própria f e seus respectivos problemas primal e dual referente a alguma perturbação φ , são válidas:*

- (i) *Se existe $\bar{x} \in \mathbb{R}^n$ tal que $f(\bar{x}) = \sup_{a \in \mathbb{R}^m} -\varphi_F^*(0, a)$ então \bar{x} é solução do problema primal.*
- (ii) *Se existe $\bar{a} \in \mathbb{R}^m$ tal que $-\varphi_F^*(0, \bar{a}) = \inf_{x \in \mathbb{R}^n} f(x)$ então \bar{a} é solução do problema dual.*
- (iii) *Se existem $\bar{x} \in \mathbb{R}^n$ e $\bar{a} \in \mathbb{R}^m$ tais que $-\varphi_F^*(0, \bar{a}) = f(\bar{x})$ então \bar{x} é solução do problema primal e \bar{a} é solução do problema dual.*

Deste modo, percebemos que as soluções desses problemas estão fortemente relacionadas. Para o caso de Programação Linear (Exemplo 1.50) que admita solução, pode se provar que sempre existe \bar{a} que satisfaz a hipótese de (iii) e portanto o valor ótimo do Primal é igual ao valor ótimo do Dual.

Agora vamos analisar o que ocorre quando repetimos o processo de dualidade. Podemos encarar a forma dual como um problema de minimização, considerando

$$\sup_{a \in \mathbb{R}^m} -\varphi_F^*(0, a) = - \inf_{a \in \mathbb{R}^m} \varphi_F^*(0, a).$$

Assim, definimos $\psi(a) = \varphi_F^*(0, a)$. Deste modo, a função $g(a, z) = \varphi_F^*(z, a)$ é uma função perturbação de ψ . Para cada $z \in \mathbb{R}^n$, podemos considerar o problema de perturbação

$$\inf_{a \in \mathbb{R}^m} g(a, z) \quad (P_z).$$

Assim, o problema dual de P^* com respeito a g , denominado por bidual e denotado por P^{**} , é caracterizado por

$$- \sup_{z \in \mathbb{R}^n} \{ -g^*(0, z) \} = \inf_{z \in \mathbb{R}^n} \{ \varphi_F^{**}(z, 0) \}$$

tal que φ_F^{**} é a biconjugada da função φ .

Quando a função de perturbação é convexa e sci, então $\varphi^{**}(x, 0) = \varphi(x, 0) = f(x)$ para todo $x \in \mathbb{R}^n$ e assim o problema bidual é idêntico ao primal, logo haverá uma simetria no processo de dualidade. Se a função original for convexa e sci, então podemos garantir que existe uma função de perturbação com as mesmas propriedades, já que a própria função pode ser considerada uma perturbação de si mesma. Deste modo, essa teoria de dualidade se comporta muito bem em problemas de minimização convexa e semicontínua.

Com o propósito de enfraquecer essas condições retirando a exigência de convexidade, a partir do próximo capítulo, faremos algo semelhante ao que foi desenvolvido no Capítulo 1 exigindo apenas a semicontinuidade das funções para obter resultados análogos.

Capítulo 2

Separação de fechados

Neste capítulo, nosso objetivo é generalizar os Teoremas de Separação para subconjuntos fechados de modo que seja possível construir sistematicamente uma teoria de conjugação para funções que sejam apenas semicontínuas. Inicialmente, vamos rever alguns tópicos já existentes na literatura sobre o assunto e usá-los como ferramenta para a construção da nova ideia proposta. O referencial teórico a respeito da generalização é [26].

2.1 Uma versão simples do teorema de Seleção de Michael

Na Topologia, existem dois importantes teoremas que separam conjuntos fechados através de funções contínuas. Um deles é o *Lema de Urysohn*, que diz que para dois subconjuntos K_1 e K_2 fechados não-vazios e disjuntos de \mathbb{R}^n , existe uma função contínua $f : \mathbb{R}^n \rightarrow [0, 1]$ que vale 0 para todo ponto de K_1 e 1 para todo ponto de K_2 . Um exemplo de função que satisfaz o lema é

$$f(x) = \frac{d(x, K_1)}{d(x, K_1) + d(x, K_2)}$$

em que $d(x, K) = \inf_{c \in K} \|x - c\|$ para $K \neq \emptyset$ é chamada de distância de ponto a conjunto. Recordamos que a função $g(x) = d(x, K)$ é contínua não negativa e

$$d(x, K) = 0 \Leftrightarrow x \in \overline{K}.$$

Assim, é fácil verificar que f satisfaz as condições do teorema. O outro teorema, que será de grande importância no desenvolvimento deste capítulo, é uma versão simples do Teorema de Seleção de Michael:

Teorema 2.1 (Versão simples do teorema de seleção de Michael) *Sejam $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função sci e $g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{-\infty\}$ uma função scs. Se $g(x) \leq f(x)$ para todo $x \in \mathbb{R}^n$, então existe uma função contínua $h : \mathbb{R}^n \rightarrow \mathbb{R}$ tal que*

$$g(x) \leq h(x) \leq f(x)$$

para todo $x \in \mathbb{R}^n$.

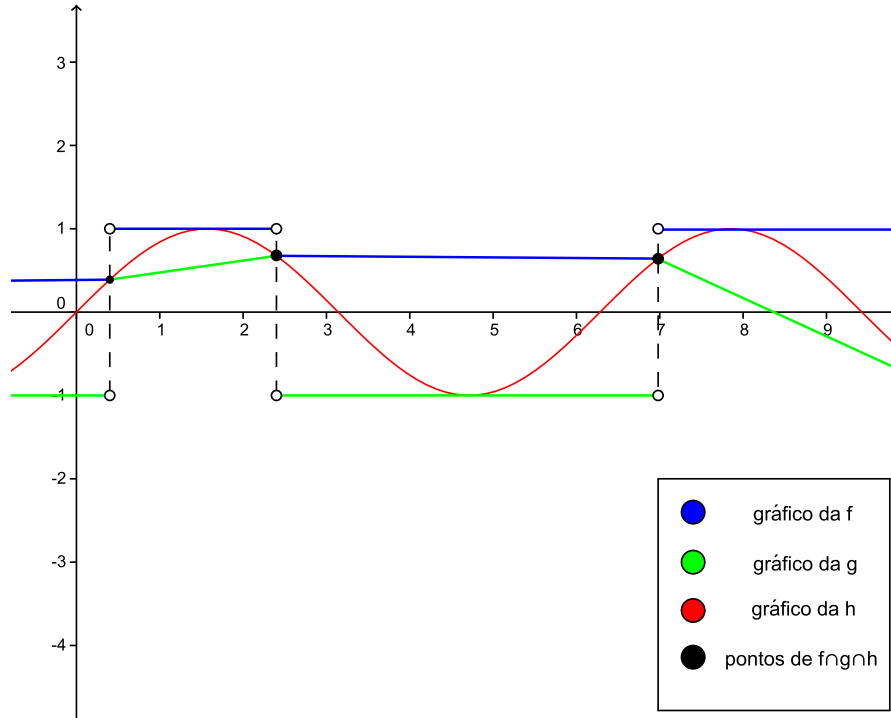


Figura 2.1: Ilustração do Teorema 2.1.

Nessa versão do Teorema de Seleção de Michael, existe a separação de dois subconjuntos especiais de \mathbb{R}^n : o epigrafo de uma função sci do hipografo de uma função scs. No entanto, o teorema não dá muitas informações sobre h . Vamos investigar este fato para um caso particular no próximo exemplo.

Exemplo 2.2 *Sejam K_1 e K_2 dois subconjuntos convexos disjuntos de \mathbb{R}^n . Consideremos $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ tal que $f(x) = \begin{cases} 0, & x \in \overline{K_1} \\ +\infty, & x \notin \overline{K_1} \end{cases}$ e $g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{-\infty\}$ tal que $g(x) = \begin{cases} 0, & x \in \overline{K_2} \\ -\infty, & x \notin \overline{K_2} \end{cases}$.*

Claramente $g(x) \leq f(x)$ para todo $x \in \mathbb{R}^n$. Além disso, podemos ver que f é sci e g é scs. De fato, para os pontos que não estão na fronteira de K_1 e de K_2 , ambas as funções são constantes e portanto contínuas. Para x_0 na fronteira de K_1 e y_0 na fronteira de K_2 , temos que:

$$\liminf_{x \rightarrow x_0} f(x) = 0 = f(x_0) \text{ e } \limsup_{x \rightarrow x_0} f(x) = +\infty > f(x_0),$$

$$\liminf_{y \rightarrow y_0} g(y) = -\infty < g(y_0) \text{ e } \limsup_{y \rightarrow y_0} g(y) = 0 = g(y_0)$$

e, portanto, f é sci e g é scs.

Pelo Teorema 2.1, existe uma função $h : \mathbb{R}^n \rightarrow \mathbb{R}$ contínua tal que

$$g(x) \leq h(x) \leq f(x) \quad (2.1)$$

para todo $x \in \mathbb{R}^n$. Como $g(x) = 0$ para todo $x \in K_2$ e $f(x) = 0$ para todo $x \in K_1$ então h satisfaz a equação (2.1) se, e somente se atende as seguintes condições:

Um ponto x pertence a fronteira de um subconjunto K de \mathbb{R}^n quando toda bola aberta centrada em x contém pontos de K e do complementar de K .

(i) $h(x) \leq 0$ para todo $x \in K_1$,

(ii) $h(x) \geq 0$ para todo $x \in K_2$.

Por outro lado, como K_1 e K_2 são convexos disjuntos, pelo Teorema 1.38, existem $a \in \mathbb{R}^n$ não nulo e $b \in \mathbb{R}$ tais que

$$\langle a, x_1 \rangle \leq b \leq \langle a, x_2 \rangle$$

para todo $x_1 \in K_1$ e $x_2 \in K_2$.

Assim, notemos que a função $h(x) = \langle a, x \rangle - b$ satisfaz as condições (i) e (ii) e portanto satisfaz o Teorema 2.1.

Esse exemplo nos dá uma ideia da expressão de h e além disso, relaciona a versão simples do teorema de Michael com os clássicos teoremas de separação. É com essa motivação que queremos estender a ideia de separar conjuntos, usando, além dos hiperplanos separadores, funções contínuas. Para obter mais informações a respeito do Lema de Urysohn e do Teorema de Seleção de Michael, veja [21, 22].

2.2 Considerações iniciais

Denotaremos por $C(\mathbb{R}^n, \mathbb{R}^m)$ o conjunto de todas as funções contínuas de \mathbb{R}^n em \mathbb{R}^m . Considerando as operações usuais de soma (+) e de multiplicação por escalar (\cdot) em $C(\mathbb{R}^n, \mathbb{R}^m)$, podemos perceber que $(C(\mathbb{R}^n, \mathbb{R}^m), +, \cdot)$ é um espaço vetorial real de dimensão infinita. Vamos agora introduzir uma topologia nesse espaço.

Definição 2.3 A bola aberta de centro em $h \in C(\mathbb{R}^n, \mathbb{R}^m)$ e raio $\epsilon > 0$ é definida por

$$B(h, \epsilon) = \left\{ f \in C(\mathbb{R}^n, \mathbb{R}^m) \mid \sup_{x \in \mathbb{R}^n} \|f(x) - h(x)\| < \epsilon \right\}.$$

A topologia a ser considerada nesse espaço será aquela cuja base sejam estas bolas. Esta topologia é conhecida como topologia da convergência uniforme. Deste modo, dizemos que uma sequência (p_k) converge em $C(\mathbb{R}^n, \mathbb{R}^m)$ para uma função p quando para todo $\epsilon > 0$ exista $k_0 \in \mathbb{N}$ tal que

$$\sup_{x \in \mathbb{R}^n} \|p_k(x) - p(x)\| < \epsilon$$

para todo $k \geq k_0$. Isso implica que $\|p_k(x) - p(x)\| < \epsilon$ para todo $x \in \mathbb{R}^n$. Sendo assim, a convergência uniforme em $C(\mathbb{R}^n, \mathbb{R}^m)$ implica na convergência pontual de funções.

Mais adiante, trabalharemos muito com o espaço $C(\mathbb{R}^n, \mathbb{R}^n)$. Dessa forma, para simplificar a notação, vamos denotar este conjunto apenas por \mathbf{C} .

De maneira natural, definiremos a bola aberta de centro em $(p, \lambda) \in \mathbf{C} \times \mathbb{R}$ e raio $\epsilon > 0$ como

$$B((p, \lambda), \epsilon) = \left\{ (q, \beta) \in \mathbf{C} \times \mathbb{R} \mid \sup_{x \in \mathbb{R}^n} \|q(x) - p(x)\| < \epsilon \text{ e } |\beta - \lambda| < \epsilon \right\}.$$

A topologia a ser considerada em $\mathbf{C} \times \mathbb{R}$ também será aquela cuja base sejam estas bolas.

Antes de enunciar o próximo teorema, vamos analisar a continuidade de um caso particular de funções.

Proposição 2.4 Consideremos $g \in C(\mathbb{R}^n, \mathbb{R})$ e $\delta > 0$. Seja $p : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definida como

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

Então p é contínua.

Demonstração. Vamos verificar que $p \in C$ 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 $\overline{B(0, \delta)}$ ($\|x\| > \delta$), que também é um aberto em \mathbb{R}^n . Assim, basta verificar a continuidade de p na fronteira de $\overline{B(0, \delta)}$ ($\|x\| = \delta$).

Dado $x \in \mathbb{R}^n$ tal que $\|x\| = \delta$, consideremos (x_k) uma sequência que converge para x . Queremos provar que $(p(x_k))$ converge para $p(x)$. Inicialmente, pela continuidade de g e da norma, notemos que

$$g(x_k) \longrightarrow g(x), \quad (2.2)$$

$$\|x_k\| \longrightarrow \|x\| = \delta. \quad (2.3)$$

Consideremos

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

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,

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

para todo $k \geq k_0$. Por (2.2),

$$p(x_k) = g(x_k) \frac{x_k}{\delta^2} \longrightarrow g(x) \frac{x}{\delta^2} = p(x).$$

Portanto $p(x_k)$ converge para $p(x)$.

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,

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

para todo $k \geq k_2$. Por (2.2) e (2.3),

$$p(x_k) = g(x_k) \frac{x_k}{\|x_k\|^2} \longrightarrow g(x) \frac{x}{\|x\|^2} = g(x) \frac{x}{\delta^2} = p(x).$$

e novamente $p(x_k)$ converge para $p(x)$.

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

$$p(x_k) \xrightarrow{\mathbb{N}'} p(x) \quad e \quad p(x_k) \xrightarrow{\mathbb{N}''} p(x).$$

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

$$\|p(x_k) - p(x)\| = \left\| g(x_k) \frac{x_k}{\|x_k\|^2} - g(x) \frac{x}{\delta^2} \right\| < \gamma \quad (2.4)$$

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

$$\|p(x_k) - p(x)\| = \left\| g(x_k) \frac{x_k}{\delta^2} - g(x) \frac{x}{\delta^2} \right\| < \gamma \quad (2.5)$$

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

Tomemos $\bar{k} = \max\{k_1, k_2\}$. Notemos que cada $k \geq \bar{k}$ pertence a \mathbb{N}' ou a \mathbb{N}'' . Por (2.4) e (2.5), em ambos os casos

$$\|p(x_k) - p(x)\| < \gamma.$$

Portanto, $(p(x_k))$ converge para $p(x)$. Dessa maneira, $p \in C$. \square

Veremos agora um importante resultado das funções contínuas. Ele caracteriza os elementos de $C(\mathbb{R}^n, \mathbb{R})$ usando funções de C .

Teorema 2.5 *Uma função $h : \mathbb{R}^n \rightarrow \mathbb{R}$ é contínua se, e somente se, para todo $\epsilon > 0$, existem $\delta \in (0, \epsilon]$ e $p \in C$ tais que*

(i) $h(x) = \langle p(x), x \rangle + h(0) - \epsilon$ para todo $x \notin \overline{B(0, \delta)}$;

(ii) $0 \leq \langle p(x), x \rangle \leq 2\epsilon$ para todo $x \in \overline{B(0, \delta)}$;

(iii) $|h(x) - \langle p(x), x \rangle - h(0)| \leq \epsilon$ para todo $x \in \overline{B(0, \delta)}$.

Demonstração. (\Rightarrow) Como h é contínua em \mathbb{R}^n , então, em particular, h é contínua em 0. Assim, para cada $\epsilon > 0$, existe um $\delta \in (0, \epsilon]$ tal que

$$h(0) - \epsilon \leq h(x) \leq h(0) + \epsilon \quad (2.6)$$

para todo $x \in \overline{B(0, \delta)}$.

Consideremos $p : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definida como

$$p(x) = \begin{cases} (h(x) - h(0) + \epsilon) \frac{x}{\|x\|^2}, & x \notin \overline{B(0, \delta)} \\ (h(x) - h(0) + \epsilon) \frac{x}{\delta^2}, & x \in \overline{B(0, \delta)} \end{cases}.$$

Se considerarmos $g(x) = h(x) - h(0) + \epsilon$, então, como h é contínua, então g também é contínua e portanto, pelo Lema 2.4, $p \in C$.

Observemos que

$$\langle p(x), x \rangle = \begin{cases} (h(x) - h(0) + \epsilon) \frac{\langle x, x \rangle}{\|x\|^2} = (h(x) - h(0) + \epsilon), & x \notin \overline{B(0, \delta)} \\ (h(x) - h(0) + \epsilon) \frac{\langle x, x \rangle}{\delta^2} = (h(x) - h(0) + \epsilon) \frac{\|x\|^2}{\delta^2}, & x \in \overline{B(0, \delta)} \end{cases}. \quad (2.7)$$

Assim:

(i) Se $x \notin \overline{B(0, \delta)}$ então $h(x) = \langle p(x), x \rangle + h(0) - \epsilon$.

(ii) Se $x \in \overline{B(0, \delta)}$ ($\|x\| \leq \delta$), então somando $(-h(0) + \epsilon)$ na equação (2.6), obtemos

$$0 \leq h(x) - h(0) + \epsilon \leq 2\epsilon,$$

logo

$$0 \leq (h(x) - h(0) + \epsilon) \frac{\|x\|^2}{\delta^2} \leq 2\epsilon \frac{\|x\|^2}{\delta^2} \leq 2\epsilon$$

pois $\|x\| \leq \delta$. Portanto, por (2.7),

$$0 \leq \langle p(x), x \rangle \leq 2\epsilon$$

para todo $x \in \overline{B(0, \delta)}$.

(iii) Para $x \in \overline{B(0, \delta)}$, novamente usando a continuidade de h em 0, temos que

$$h(x) \leq h(0) + \epsilon \Leftrightarrow -\epsilon \leq h(0) - h(x).$$

Usando o item (ii), podemos concluir que

$$-\epsilon \leq h(0) - h(x) \leq h(0) - h(x) + \langle p(x), x \rangle = h(0) - h(x) + (h(x) - h(0) + \epsilon) \frac{\|x\|^2}{\delta^2}.$$

Como $\|x\| \leq \delta$, então

$$-\epsilon \leq h(0) - h(x) + \langle p(x), x \rangle \leq h(0) - h(x) + h(x) - h(0) + \epsilon = \epsilon.$$

Assim $|h(x) - \langle p(x), x \rangle - h(0)| \leq \epsilon$ para todo $x \in \overline{B(0, \delta)}$.

(\Leftarrow)Primeiramente vamos verificar a continuidade de h em x' não nulo. Deste modo, tomemos $\epsilon \in (0, \|x'\|)$.

Por hipótese, existem $\delta \in (0, \epsilon]$ e $p \in C$ tais que

$$h(x) = \langle p(x), x \rangle + h(0) - \epsilon$$

para todo $x \notin \overline{B(0, \delta)}$.

Como $\|x'\| > \epsilon \geq \delta$, então $x' \notin \overline{B(0, \delta)}$. Assim $\overline{h(x)}$ pode ser expressa como uma composição de funções contínuas no complementar da $\overline{B(0, \delta)}$, que é aberto em \mathbb{R}^n . Isso nos permite concluir que h é contínua em x' .

Agora vamos verificar que h é contínua na origem. Por hipótese, dado $\epsilon > 0$, existe $\delta \in (0, \epsilon]$ e $p \in C$ tais que

$$|h(x) - \langle p(x), x \rangle - h(0)| \leq \epsilon$$

para todo $x \in \overline{B(0, \delta)}$. Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ definida por $f(x) = h(x) - \langle p(x), x \rangle$. Assim,

$$\|x\| \leq \delta \Rightarrow |f(x) - f(0)| = |h(x) - \langle p(x), x \rangle - h(0)| \leq \epsilon$$

e, portanto, f é contínua em 0.

Notemos que $h(x) = f(x) + \langle p(x), x \rangle$ e p é contínua. Como $B(0, \delta)$ é aberto em \mathbb{R}^n , então h pode ser expressa como composta de funções contínuas numa vizinhança de 0 e portanto é contínua em 0. \square

Com o auxílio deste resultado técnico, poderemos provar a generalização dos teoremas de separação. Para finalizar a seção, vamos analisar um caso particular do teorema de seleção de Michael para funções estritamente positivas. Para sua demonstração, será necessário o seguinte resultado de Análise.

Lema 2.6 *Toda sequência limitada em \mathbb{R}^n cujas subsequências convergem para um mesmo ponto é convergente.*

Demonstração. Seja (x_k) uma sequência limitada cujas subsequências convergem para x . Se (x_k) não converge para $x \in \mathbb{R}^n$, então existe $\epsilon > 0$ tal que o conjunto $\mathbb{N}' = \{k \in \mathbb{N}; \|x_k - x\| > \epsilon\}$ é infinito. A subsequência $(x_k)_{k \in \mathbb{N}'}$ também é limitada e portanto admite subsequência $(x_{k_j})_{k_j \in \mathbb{N}'}$ convergente. Por hipótese, $(x_{k_j})_{k_j \in \mathbb{N}'}$ converge para x , entrando em contradição com a definição de \mathbb{N}' . \square

Teorema 2.7 *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função sci e própria. Se $f(x) > 0$ para todo $x \in \mathbb{R}^n$, então existe $h \in C(\mathbb{R}^n, \mathbb{R})$ tal que*

$$0 < h(x) < f(x)$$

para todo $x \in \mathbb{R}^n$.

Demonstração. Consideremos $K_1 = \mathbb{R}^n \times (-\infty, 0]$ e $K_2 = \text{epi}(f)$. Notemos que os dois subconjuntos de \mathbb{R}^{n+1} são não-vazios, fechados e disjuntos. Para $(x, \lambda) \in \mathbb{R}^{n+1}$, definimos

$$d_1(x, \lambda) = d((x, \lambda), K_1) = \inf_{(w, \alpha) \in K_1} \{\|x - w\| + |\lambda - \alpha|\} = \inf_{\alpha \leq 0} \{|\lambda - \alpha|\} = \begin{cases} \lambda, & \lambda > 0 \\ 0, & \lambda \leq 0 \end{cases}$$

$$d_2(x, \lambda) = d((x, \lambda), K_2) = \inf_{(w, \alpha) \in K_2} \{\|x - w\| + |\lambda - \alpha|\}.$$

Consideremos $D : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ definida por

$$D(x, \lambda) = \frac{d_1(x, \lambda)}{d_1(x, \lambda) + d_2(x, \lambda)}.$$

A função D é contínua pois é composição de funções contínuas. Além disso, $D(x, \lambda) = 0$ para todo $(x, \lambda) \in K_1$ e $D(x, \lambda) = 1$ para todo $(x, \lambda) \in K_2$.

Como K_1 e K_2 são disjuntos e fechados, podemos concluir que, para $x \in \mathbb{R}^n$,

$$D(x, \lambda) = \frac{1}{2} \Leftrightarrow d_1(x, \lambda) = d_2(x, \lambda) \Rightarrow d_1(x, \lambda) = d_2(x, \lambda) \neq 0 \Rightarrow 0 < \lambda < f(x).$$

Dessa forma, definimos $h(x) = \{\lambda; D(x, \lambda) = 1/2\}$.

1º) Vamos mostrar que h está bem definida. Para isso, vamos verificar os itens que seguem:

a) Dado $x \in \mathbb{R}^n$, existe $\lambda \in \mathbb{R}$ tal que $D(x, \lambda) = 1/2$.

Consideremos, para x fixo, a função $g(t) = D(x, t)$. Notemos que g é contínua devido à continuidade da função D .

Caso $x \in \text{dom}(f)$, então $g(t) = 0$ para $t \leq 0$ e $g(t) = 1$ para $t \geq f(x)$. Portanto, pelo Teorema do Valor Intermediário³, existe $\lambda \in (0, f(x))$ tal que $D(x, \lambda) = g(\lambda) = 1/2$.

Para o caso em que $x \notin \text{dom}(f)$, consideremos

$$\bar{d} = d(x, \text{dom}(f)) = \inf_{y \in \text{dom}(f)} \{\|x - y\|\}.$$

Pelo Lema 1.34, existe $\bar{x} \in \overline{\text{dom}(f)}$ tal que $\|x - \bar{x}\| = \bar{d}$.

Vamos agora provar que existe $\bar{\lambda} > 0$ tal que $d_2(x, \bar{\lambda}) < \bar{\lambda} = d_1(x, \bar{\lambda})$.

Se $\bar{x} \in \text{dom}(f)$, então $f(\bar{x}) < +\infty$. Considerando $\bar{\lambda} = \max\{f(\bar{x}), \bar{d} + 1/2\}$, teremos que $(\bar{x}, \bar{\lambda}) \in \text{epi}(f)$. Deste modo

$$d_2(x, \bar{\lambda}) \leq \|x - \bar{x}\| + |\bar{\lambda} - \bar{\lambda}| = \bar{d} < \bar{\lambda} = d_1(x, \bar{\lambda}).$$

Caso $f(\bar{x}) = +\infty$, consideramos, a partir da definição de fecho, uma sequência (x_k) em $\text{dom}(f)$ que converge para \bar{x} . Além disso, como f é sci, temos que

$$\liminf_{y \rightarrow \bar{x}} f(y) \geq f(\bar{x}) = +\infty.$$

³O teorema do Valor Intermediário diz que para uma função contínua $f : [a, b] \rightarrow \mathbb{R}$ tal que $f(a) < d < f(b)$, existe $c \in (a, b)$ tal que $f(c) = d$.

Suponhamos que $(f(x_k))$ seja limitada. Então existe uma subsequência $(f(x_k))_{k \in \mathbb{N}'}$ que converge para $z \in \mathbb{R}$. Como $(x_k)_{k \in \mathbb{N}'}$ converge para \bar{x} , temos que, para $\delta > 0$ existe $k_0 \in \mathbb{N}'$ tal que $m(\bar{x}, \delta) \leq f(x_k)$ para todo $k > k_0$. Portanto, $m(\bar{x}, \delta) \leq z$ e assim

$$\liminf_{y \rightarrow \bar{x}} f(x) \leq z < +\infty,$$

que é absurdo. Portanto, $(f(x_k))$ é ilimitada.

Deste modo, seja $k_1 \in \mathbb{N}$ tal que

$$\|x_{k_1} - \bar{x}\| \leq 1 \text{ e } f(x_{k_1}) > \bar{d} + 1.$$

Considerando $\bar{\lambda} = f(x_{k_1})$, temos que $(x_{k_1}, \bar{\lambda}) \in \text{epi}(f)$ e que

$$d_2(x, \bar{\lambda}) \leq \|x - x_{k_1}\| + |\bar{\lambda} - \bar{\lambda}| \leq \|x - \bar{x}\| + \|\bar{x} - x_{k_1}\| \leq \bar{d} + 1 < \bar{\lambda} = d_1(x, \bar{\lambda}).$$

Assim, $g(\bar{\lambda}) > 1/2$. Novamente, pelo Teorema do Valor Intermediário, existe $\lambda \in (0, \bar{\lambda})$ tal que $D(x, \lambda) = g(\lambda) = 1/2$.

b) Dado $x \in \mathbb{R}^n$, existe um único λ tal que $D(x, \lambda) = 1/2$.

Consideremos λ_1 e λ_2 distintos tais que $D(x, \lambda_1) = D(x, \lambda_2) = 1/2$. Deste modo, $\lambda_1 = d_1(x, \lambda_1) = d_2(x, \lambda_1)$ e $\lambda_2 = d_1(x, \lambda_2) = d_2(x, \lambda_2)$. Sem perda de generalidade, vamos considerar $\lambda_1 < \lambda_2$ e $\epsilon = \lambda_2 - \lambda_1$.

Por definição de ínfimo, existe $(x_0, \lambda_0) \in \text{epi}(f)$ tal que

$$\|x - x_0\| + |\lambda_1 - \lambda_0| < d_2(x, \lambda_1) + \epsilon = \lambda_1 + \epsilon = \lambda_2. \quad (2.8)$$

Assim,

$$\|x - x_0\| < \lambda_2. \quad (2.9)$$

Temos dois possíveis casos:

- Se $\lambda_0 \leq \lambda_2$ então $(x_0, \lambda_2) \in \text{epi}(f)$ e assim, $\lambda_2 = d_2(x, \lambda_2) \leq \|x - x_0\|$. Por (2.9), temos que $\lambda_2 < \lambda_2$. Absurdo.
- Se $\lambda_2 < \lambda_0$, primeiramente notemos que $|\lambda_2 - \lambda_0| = \lambda_0 - \lambda_2 \leq \lambda_0 - \lambda_1 \leq |\lambda_0 - \lambda_1|$. Como $(x_0, \lambda_0) \in \text{epi}(f)$, temos que

$$\lambda_2 = d_2(x, \lambda_2) \leq \|x - x_0\| + |\lambda_2 - \lambda_0| \leq \|x - x_0\| + |\lambda_1 - \lambda_0|.$$

Por (2.8), podemos concluir que $\lambda_2 < \lambda_2$. Absurdo.

Logo, $\lambda_1 = \lambda_2$

2º) Por fim, vamos mostrar que h é contínua.

Dado $x \in \mathbb{R}^n$, consideremos (x_k) uma sequência que converge para x . Queremos provar que $(h(x_k))$ converge para $h(x)$. Vamos provar inicialmente que $h(x_k)$ é limitada.

De fato, suponha por absurdo que não. Então existe uma subsequência (x_{k_j}) tal que

$$\lim_{j \rightarrow \infty} h(x_{k_j}) = +\infty. \quad (2.10)$$

Deste modo, existe $j_0 \in \mathbb{N}$ tal que $f(x_0) \leq h(x_{k_j})$, ou equivalentemente, $(x_0, h(x_{k_j})) \in \text{epi}(f)$ para todo $k_j > k_{j_0}$. Logo

$$h(x_{k_j}) = \lambda_{k_j} \leq \|x_{k_j} - x_0\| + |h(x_{k_j}) - h(x_{k_j})| = \|x_{k_j} - x_0\|$$

para todo $k_j > k_{j_0}$. Isso é absurdo, pois (x_{k_j}) é convergente e, portanto $(x_{k_j} - x_0)$ é limitada, contrariando (2.10).

Consideremos agora $(h(x_{k_s}))_{s \in \mathbb{N}}$ uma subsequência que converge para $z \in \mathbb{R}^n$. Então $h(x_{k_s}) = \lambda_{k_s}$ tal que

$$D(x_{k_s}, \lambda_{k_s}) = 1/2 \quad (2.11)$$

para todo $s \in \mathbb{N}$. Deste modo, pela continuidade da D , temos que

$$D(x_{k_s}, \lambda_{k_s}) \xrightarrow{s} D(x, z).$$

Mas por (2.11), temos que $D(x_{k_s}, \lambda_{k_s})$ converge para $1/2$. Portanto $D(x, z) = 1/2$ e assim $z = h(x)$. Pelo Lema 2.6, segue que h é contínua. \square

2.3 Teorema de separação para fechados

Nesta seção, vamos generalizar o produto interno usual do espaço \mathbb{R}^n de modo a induzir uma generalização dos hiperplanos. Com essas ferramentas, poderemos estender os teoremas de separação vistos na Seção 1.4 para conjuntos fechados.

Definição 2.8 A aplicação $\phi : \mathbb{C} \times \mathbb{R}^n \rightarrow \mathbb{R}$, denotada por $\langle \cdot, \cdot \rangle$ e definida como

$$\phi(p, x) = \langle p, x \rangle = \langle p(x), x \rangle$$

é a generalização do produto interno usual de \mathbb{R}^n .

Notemos que, quando consideramos F o subespaço de \mathbb{C} formado por todas as funções constantes e restringimos ϕ ao subespaço $F \times \mathbb{R}^n$, podemos concluir que o produto interno usual é, de fato, um caso particular da aplicação ϕ .

Consideremos agora a aplicação $T : \mathbb{C} \rightarrow \mathbb{R}$ definida por $T(p) = \langle p, x \rangle$ para $x \in \mathbb{R}^n$ fixo. Notemos que T é linear e contínua. De fato; a linearidade é consequência direta da linearidade do produto interno. Para provar a continuidade de T , considere (p_k) uma seqüência que converge em \mathbb{C} para uma função p . Então $p_k(x) \rightarrow p(x)$ para todo $x \in \mathbb{R}^n$. Como produto interno também é um operador contínuo em $\mathbb{R}^n \times \mathbb{R}^n$, temos que

$$\lim_{k \rightarrow \infty} T(p_k) = \lim_{k \rightarrow \infty} \langle p_k(x), x \rangle = \left\langle \lim_{k \rightarrow \infty} p_k(x), x \right\rangle = \langle p, x \rangle = T(p).$$

Deste modo, definimos o hiperplano em \mathbb{C} como

$$H(x, b) = \{p \in \mathbb{C} \mid T(p) = \langle p, x \rangle = b\}$$

para $x \in \mathbb{R}^n$ e $b \in \mathbb{R}$.

Além disso, de maneira natural, os hiperplanos de \mathbb{R}^n serão generalizados por conjuntos do tipo

$$H(p, b) = \{x \in \mathbb{R}^n \mid \langle p, x \rangle = b\}$$

com $p \in \mathbb{C}$ e $b \in \mathbb{R}$.

Notemos que quando p for uma função constante, $H(p, b)$ será um hiperplano de \mathbb{R}^n . Em nossas generalizações, os conjuntos separadores serão do tipo $H(p, b)$. Assim, de maneira análoga ao caso dos hiperplanos, vamos definir o seguinte conceito.

Definição 2.9 Dados dois subconjuntos K_1, K_2 de \mathbb{R}^n e $(p, b) \in \mathbb{C} \times \mathbb{R}$, dizemos que $H(p, b)$ separa K_1 e K_2 quando

$$\langle p, x_1 \rangle \leq b \leq \langle p, x_2 \rangle \quad \forall x_1 \in K_1, \forall x_2 \in K_2.$$

Dizemos que $H(p, b)$ separa estritamente K_1 e K_2 quando

$$\langle p, x_1 \rangle < b < \langle p, x_2 \rangle \quad \forall x_1 \in K_1, \forall x_2 \in K_2.$$

Para investigar as propriedades dos subconjuntos que garantam a existência desses $H(p, b)$, vamos nos preocupar inicialmente com um caso particular: separar o epigrafo de uma função sci f do hipografo de uma função scs g .

Pelo Teorema 2.1, já sabemos que existe uma função h contínua que separa estes conjuntos quando $g(x) \leq f(x)$ para todo $x \in \mathbb{R}^n$. Do Teorema 2.5, existe uma bola fechada de \mathbb{R}^n centrada na origem em que h pode ser expressa como $\langle p, x \rangle - b$ para alguma $p \in \mathbb{C}$ e $b \in \mathbb{R}$ para todo x fora da bola. Deste modo, veremos a seguir, que se $g(0) < f(0)$, então podemos construir esta bola com o raio menor que $|f(0) - g(0)|$ e, assim, verificar que $\langle p, x \rangle - b$ satisfaz o Teorema 2.1 em todo \mathbb{R}^n .

Teorema 2.10 Sejam $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função sci e $g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{-\infty\}$ uma função scs. Se $g(x) \leq f(x)$ para todo $x \in \mathbb{R}^n$ e $(0, g(0)) \notin \text{epi}(f)$, então existem $p \in \mathbb{C}$ e $b \in \mathbb{R}$ tais que

$$g(x) \leq \langle p, x \rangle - b \leq f(x)$$

para todo $x \in \mathbb{R}^n$.

Demonstração. Como $(0, g(0)) \notin \text{epi}(f)$, então $g(0) < f(0)$. Seja $\epsilon > 0$ e γ no intervalo $(g(0), f(0))$ suficientemente pequenos de modo que $g(0) < \gamma - \epsilon < \gamma < \gamma + \epsilon < f(0)$.

Consideremos $\bar{f} : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ tal que $\bar{f}(x) = \begin{cases} f(x), & x \neq 0 \\ \gamma, & x = 0 \end{cases}$ e $\bar{g} : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{-\infty\}$

tal que $\bar{g}(x) = \begin{cases} g(x), & x \neq 0 \\ \gamma, & x = 0 \end{cases}$.

Para $m' < \bar{f}(0) = \gamma$, temos que $f(0) > \gamma > m'$. Como f é sci, pelo Corolário 1.16, temos que existe $\delta > 0$ tal que $f(y) > m'$ para todo $y \in B(0, \delta)$. Dessa forma, $\bar{f}(y) > m'$ para todo $y \in B(0, \delta)$, ou seja, \bar{f} é sci em 0.

Para os pontos distintos de 0, \bar{f} é escrita como uma função sci. Usando raciocínio análogo, podemos concluir que \bar{g} é scs. Além disso, $\bar{g}(x) \leq \bar{f}(x)$ para todo $x \in \mathbb{R}^n$ e, portanto, pelo Teorema 2.1, existe uma função $h : \mathbb{R}^n \rightarrow \mathbb{R}$ contínua tal que $\bar{g}(x) \leq h(x) \leq \bar{f}(x)$ para todo $x \in \mathbb{R}^n$.

Notemos que $\gamma = \bar{g}(0) \leq h(0) \leq \bar{f}(0) = \gamma$. Logo $h(0) = \gamma$ e

$$g(0) < h(0) - \epsilon < h(0) < h(0) + \epsilon < f(0).$$

Pela continuidade de h em 0, temos que existe $\eta_1 \in (0, \epsilon]$ tal que

$$h(0) - \epsilon < h(x) < h(0) + \epsilon$$

para todo $x \in B(0, \eta_1)$.

Como g é scs em 0, pelo Corolário 1.16, existe η_2 tal que $g(x) < h(0) - \epsilon$ para todo $x \in B(0, \eta_2)$.

Analogamente, como f é sci em 0, pelo Corolário 1.16, existe η_3 tal que $h(0) - \epsilon < f(x)$ para todo $x \in B(0, \eta_3)$.

Assim, para $\eta = \min \{\eta_1, \eta_2, \eta_3\} \leq \epsilon$, temos que

$$g(x) < h(0) - \epsilon < h(x) < h(0) + \epsilon < f(x) \quad (2.12)$$

para todo $x \in B(0, \eta)$.

Além do mais, pelo Teorema 2.5, existe $\delta \in (0, \eta] \subset (0, \epsilon]$ e $p \in C$ tais que

$$h(x) = \langle p, x \rangle + h(0) - \eta \quad \forall x \notin \overline{B(0, \delta)}; \quad (2.13)$$

$$0 \leq \langle p, x \rangle \quad \forall x \in \overline{B(0, \delta)}; \quad (2.14)$$

$$\langle p, x \rangle - h(x) + h(0) \leq \eta \quad \forall x \in \overline{B(0, \delta)}. \quad (2.15)$$

Consideremos $b = \eta - h(0)$. Assim, para $x \notin \overline{B(0, \delta)}$, temos que

$$g(x) \leq \langle p, x \rangle - b = h(x) \leq f(x).$$

Assim, basta provar o teorema para os pontos de $\overline{B(0, \delta)}$. Deste modo, para $x \in \overline{B(0, \delta)}$, notemos que, de (2.14) e (2.15),

$$0 \leq \langle p, x \rangle \leq \eta + h(x) - h(0).$$

Assim, de (2.12),

$$0 \leq \langle p, x \rangle < \eta + \epsilon. \quad (2.16)$$

Por outro lado, como $\eta \leq \epsilon$ e novamente de (2.12),

$$g(x) < h(0) - \epsilon \leq h(0) - \eta.$$

Por fim, combinando (2.14), (2.16) e (2.12), temos que

$$g(x) < h(0) - \eta \leq \langle p, x \rangle + h(0) - \eta = \langle p, x \rangle - b < \epsilon + h(0) < f(x)$$

para todo $x \in \overline{B(0, \delta)}$, concluindo a demonstração. \square

Graças a este último teorema, podemos construir a seguir o conjunto $H(p, b)$ que separa o epigrafo de f do hipografo de g .

Para os próximos resultados da seção, consideraremos $q : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^{n+1}$ uma função contínua definida para cada $x \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}$ como $q(x, \lambda) = (p(x), -1)$ tal que p é uma função de C .

Corolário 2.11 *Sejam $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função sci e $g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{-\infty\}$ uma função scs. Se $g(x) \leq f(x)$ para todo $x \in \mathbb{R}^n$ e $(0, g(0)) \notin \text{epi}(f)$, então existe $p \in C$ de modo que $H(q, b)$ separa o epigrafo de f do hipografo de g , ou seja,*

$$\langle p, z \rangle - \lambda_1 \leq b \leq \langle p, y \rangle - \lambda_2$$

para todo $(z, \lambda_1) \in \text{epi}(f)$ e $(y, \lambda_2) \in \text{hip}(g)$.

Demonstração. Pelo Teorema 2.10, existem $p \in C$ e $b \in \mathbb{R}$ tais que

$$g(x) \leq \langle p, x \rangle - b \leq f(x)$$

para todo $x \in \mathbb{R}^n$. Assim, para $(y, \lambda_2) \in \text{hip}(g)$, temos que

$$\lambda_2 \leq g(y) \leq \langle p, y \rangle - b \Rightarrow b \leq \langle p, y \rangle - \lambda_2.$$

Por outro lado, para $(z, \lambda_1) \in \text{epi}(f)$, temos que

$$\langle p, z \rangle - b \leq f(z) \leq \lambda_1 \Rightarrow \langle p, z \rangle - \lambda_1 \leq b.$$

Combinado as expressões, obtemos

$$\langle p, z \rangle - \lambda_1 \leq b \leq \langle p, y \rangle - \lambda_2$$

para todo $(z, \lambda_1) \in \text{epi}(f)$ e $(y, \lambda_2) \in \text{hip}(g)$. \square

O próximo teorema nos mostra que quando o epigrafo de uma função sci e o hipografo de uma função scs são disjuntos então eles podem ser separados estritamente.

Teorema 2.12 *Sejam $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função sci e $g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{-\infty\}$ uma função scs tais que $g(x) < f(x)$ para todo $x \in \mathbb{R}^n$, então existem $p \in \mathbb{C}$ e $b \in \mathbb{R}$ tais que*

$$g(x) < \langle p, x \rangle - b < f(x)$$

para todo $x \in \mathbb{R}^n$.

Demonstração. Consideremos $\bar{f} : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ tal que $\bar{f}(x) = 1/2(f(x) - g(x))$. Como $-g$ é sci, segue que \bar{f} também é. Além disso, por construção, \bar{f} é estritamente positiva.

Pelo Teorema 2.7, existe $h \in C(\mathbb{R}^n, \mathbb{R})$ tal que $0 < h(x) < \bar{f}(x)$ e assim, $g(x) < g(x) + h(x) < f(x) - h(x) < f(x)$ para todo $x \in \mathbb{R}^n$.

Em particular, notemos que $(0, g(0) + h(0)) \notin \text{epi}(f - h)$. Além disso, $g + h$ é scs e $f - h$ é sci. Pelo Teorema 2.10, existem $p \in \mathbb{C}$ e $b \in \mathbb{R}$ tais que

$$g(x) < g(x) + h(x) \leq \langle p, x \rangle - b \leq f(x) - h(x) < f(x)$$

para todo $x \in \mathbb{R}^n$, concluindo o resultado. \square

Corolário 2.13 *Sejam $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função sci e $g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{-\infty\}$ uma função scs. Se $g(x) < f(x)$ para todo $x \in \mathbb{R}^n$, então existem $p \in \mathbb{C}$ e $b \in \mathbb{R}$ de modo que $H(q, b)$ separa estritamente o epigrafo de f do hipografo de g , ou seja,*

$$\langle p, z \rangle - \lambda_1 < b < \langle p, y \rangle - \lambda_2$$

para todo $(z, \lambda_1) \in \text{epi}(f)$ e $(y, \lambda_2) \in \text{hip}(g)$.

Demonstração. Pelo Teorema 2.12, existem $p \in \mathbb{C}$ e $b \in \mathbb{R}$ tais que

$$g(x) < \langle p, x \rangle - b < f(x)$$

Assim a demonstração de que $H(q, b)$ separa estritamente o $\text{epi}(f)$ do $\text{hip}(g)$ segue analogamente ao do corolário (2.11). \square

Notemos que, no corolário anterior, os conjuntos separados eram fechados sem necessariamente serem convexos. Veremos em seguida mais um caso particular de separação de conjuntos fechados: o epigrafo de uma função sci de um ponto que está no seu complementar.

Corolário 2.14 *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função sci. Se $(y, \lambda) \notin \text{epi}(f)$, então existem $p \in \mathbb{C}$ e $b \in \mathbb{R}$ de modo que $H(q, b)$ separa estritamente o epigrafo de f de $\{(y, \lambda)\}$, ou seja,*

$$\langle p, z \rangle - \lambda_1 < b < \langle p, y \rangle - \lambda$$

para todo $(z, \lambda_1) \in \text{epi}(f)$.

Demonstração. Consideremos $g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{-\infty\}$ tal que $g(x) = \begin{cases} \lambda, & x = y \\ -\infty, & x \neq y \end{cases}$. Notemos que g é scs. De fato, para $x \neq y$, g é constante e portanto contínua. Para $x = y$, temos que:

$$\liminf_{x \rightarrow y} g(x) = -\infty < g(y) \text{ e } \limsup_{x \rightarrow y} g(x) = \lambda = g(y)$$

e, portanto, g é scs. Além disso, como $(y, \lambda) \notin \text{epi}(f)$, então $\lambda < f(y)$. Logo, $g(x) < f(x)$ para todo $x \in \mathbb{R}^n$. Aplicando o Corolário 2.13, concluímos a demonstração. \square

Por fim, vamos enunciar o principal resultado do capítulo. Sua demonstração é baseada no Exemplo 2.2.

Teorema 2.15 (Teorema de separação para fechados) *Sejam K_1 e K_2 subconjuntos fechados e disjuntos de \mathbb{R}^n . Então existe $H(p, b)$ que os separa estritamente, ou seja,*

$$\langle p, x_1 \rangle < b < \langle p, x_2 \rangle$$

para todos $x_1 \in K_1$ e $x_2 \in K_2$.

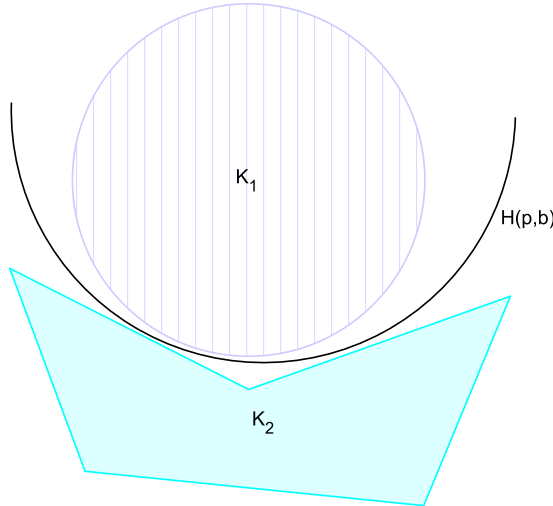


Figura 2.2: Ilustração do Teorema de Separação para Fechados.

Demonstração. Consideremos $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ sendo $f(x) = \begin{cases} 0, & x \in K_1 \\ +\infty, & x \notin K_1 \end{cases}$ e $g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{-\infty\}$ tal que $g(x) = \begin{cases} 0, & x \in K_2 \\ -\infty, & x \notin K_2 \end{cases}$. Como vimos no Exemplo 2.2, f é sci e g é scs. Além disso, desde que $K_1 \cap K_2 = \emptyset$, então $g(x) < f(x)$ para todo $x \in \mathbb{R}^n$.

Pelo Corolário 2.13, existem $p \in \mathbb{C}$ e $b \in \mathbb{R}$ tais que

$$\langle p, z \rangle - \lambda_1 < b < \langle p, y \rangle - \lambda_2$$

para todo $(z, \lambda_1) \in \text{epi}(f)$ e $(y, \lambda_2) \in \text{hip}(g)$. Como $(x_1, 0) \in \text{epi}(f)$ e $(x_2, 0) \in \text{hip}(g)$ para todo $x_1 \in K_1$ e $x_2 \in K_2$, concluímos que

$$\langle p, x_1 \rangle < b < \langle p, x_2 \rangle$$

para todo $x_1 \in K_1$ e $x_2 \in K_2$. \square

Notemos que as condições dos conjuntos serem fechados e disjuntos são essenciais para a demonstração, pois permitem a construção de funções f e g semicontínuas de modo que g seja estritamente menor que f em todos os pontos.

Consideremos K_1 e K_2 subconjuntos convexos disjuntos não-vazios de \mathbb{R}^n tais que $K_2 - K_1$ seja fechado. Quando aplicamos o Teorema 2.15 nos conjuntos $K_2 - K_1$ e $\{0\}$, podemos concluir que existe $p \in \mathbb{C}$ e $b \in \mathbb{R}$ de modo que $H(p, b)$ separa estritamente K_1 e K_2 . Dessa forma, o Teorema 2.15 é uma generalização do Teorema 1.40. Seguindo essa motivação, no próximo capítulo iremos construir uma teoria de conjugação a partir desses novos resultados.

Capítulo 3

Função conjugada modificada

Neste capítulo vamos generalizar as funções conjugadas de Fenchel a partir dos teoremas apresentados no capítulo anterior. Queremos analisar as propriedades dessa modificação e investigar o seu comportamento para funções sci. Neste contexto, vamos introduzir os espaços duais conjugados que enriquecerão a teoria de Dualidade relacionada. O referencial teórico é [6].

3.1 A extensão de Moreau

Na Literatura, a teoria de conjugação de Fenchel foi estendida por Moreau da seguinte maneira.

Definição 3.1 *Dados K_1 e K_2 conjuntos arbitrários e sejam $f : K_1 \rightarrow \mathbb{R} \cup \{+\infty\}$ e $g : K_2 \rightarrow \mathbb{R} \cup \{+\infty\}$ funções próprias. Consideremos $c : K_1 \times K_2 \rightarrow \mathbb{R} \cup \{+\infty\}$. A função c -conjugada de f é definida por $f^c : K_2 \rightarrow \mathbb{R} \cup \{+\infty\}$;*

$$f^c(y) = \sup_{x \in K_1} \{c(x, y) - f(x)\}.$$

A função c' -conjugada de g é definida por $g^{c'} : K_1 \rightarrow \mathbb{R} \cup \{+\infty\}$;

$$g^{c'}(x) = \sup_{y \in K_2} \{c(x, y) - g(y)\}.$$

Notemos que quando $K_1 = K_2 = \mathbb{R}^n$ e c é o produto interno usual de \mathbb{R}^n , as funções c -conjugada e c' -conjugada são a mesma e coincidem com a conjugada clássica. Já para qualquer outra função c , teremos uma nova teoria de conjugação. Assim, temos uma certa liberdade para estender a conjugada de Fenchel, mas nem todas essas extensões são interessantes. Vamos analisar o próximo exemplo.

Exemplo 3.2 *Sejam $K_1 = K_2 = (0, +\infty)$ e $c(x, y) = \log(xy)$.*

Consideremos $f(x) = x^2$ e vamos calcular a c -conjugada de f . Dado $y \in K_2$, temos que

$$f^c(y) = \sup_{x > 0} \{\log(xy) - x^2\}.$$

Consideremos $h : (0, +\infty) \rightarrow \mathbb{R}$ dada por $h(x) = \log(xy) - x^2$. Percebemos que

$$h'(x) = \frac{1}{x} - 2x$$

e assim, o único ponto crítico de h é $\bar{x} = \frac{1}{\sqrt{2}}$. Além disso, como $h''(x) = -\frac{1}{x^2} - 2 < 0$, temos, pelas condições de otimalidade, que o maximizador de h é \bar{x} e o valor máximo que h atinge é $\log(\frac{y}{\sqrt{2}}) - \frac{1}{2}$. Portanto, $f^c(y) = \log(\frac{y}{\sqrt{2}}) - \frac{1}{2}$ para todo $y \in K_2$.

Lembramos que a conjugação clássica preserva a convexidade da função. Já neste exemplo isso não ocorre, pois f é contínua e convexa, mas f^c não é convexa. Este fato é desinteressante no ponto de vista da otimização. Assim, devemos nos preocupar em estender a teoria de Fenchel sem perder suas principais características.

Usando como base teórica os teoremas de separação do Capítulo 2, vamos generalizar a conjugada clássica utilizando como função c o produto interno generalizado e veremos que é possível concluir resultados similares aos da Seção 1.5 exigindo condições mais fracas da função inicial. Esta nova teoria foi desenvolvida para se comportar bem em funções sci, por isso dizemos que essa generalização é a função conjugada para funções sci.

3.2 Função conjugada para funções sci

Seja $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função própria, consideremos

$$D_f = \left\{ p \in \mathbb{C} \mid \sup_{x \in \mathbb{R}^n} \{ \langle p, x \rangle - f(x) \} < +\infty \right\}.$$

Notemos que, para qualquer p em \mathbb{C} ,

$$\sup_{x \in \mathbb{R}^n} \{ \langle p, x \rangle - f(x) \} > -\infty.$$

De fato, como f é própria, então existe \bar{x} tal que $f(\bar{x}) < \lambda < +\infty$ para algum $\lambda \in \mathbb{R}$. Deste modo,

$$\sup_{x \in \mathbb{R}^n} \{ \langle p, x \rangle - f(x) \} \geq \langle p, \bar{x} \rangle - f(\bar{x}) > \langle p, \bar{x} \rangle - \lambda > -\infty.$$

Com a ajuda do Corolário 2.14, vamos verificar que D_f é não-vazio se a função f for sci.

Teorema 3.3 *Dada uma função própria $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ sci. Então $D_f \neq \emptyset$.*

Demonstração. Consideremos $y \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}$ tal que $f(y) > \lambda$. Como f é sci, então, pelo Corolário 2.14, existem $p \in \mathbb{C}$ e $b \in \mathbb{R}$ tais que

$$\langle p, z \rangle - \lambda_1 < b$$

para todo $(z, \lambda_1) \in \text{epi}(f)$. Se $x \in \text{dom}(f)$, temos que $\langle p, x \rangle - f(x) < b$. Se $x \notin \text{dom}(f)$, então $\langle p, x \rangle - f(x) = -\infty < b$. Assim,

$$\sup_{x \in \mathbb{R}^n} \{ \langle p, x \rangle - f(x) \} \leq b < +\infty,$$

permitindo concluir que $p \in D_f$. □

Com esse resultado, a função conjugada clássica será estendida para a função cujo domínio seja D_f . Utilizando a Definição 2.8, vamos generalizar a função conjugada como segue.

Definição 3.4 Consideremos a função própria $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$. A função conjugada generalizada de f , denotada por f^* , será definida por $f^* : \mathbb{C} \rightarrow \mathbb{R} \cup \{+\infty\}$;

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

Para evitar confusões, a partir deste ponto do trabalho, a conjugada clássica será referenciada por conjugada de Fenchel e a correspondente à Definição 3.4 será referenciada simplesmente por conjugada.

Exemplo 3.5 $f : \mathbb{R} \rightarrow \mathbb{R}; f(x) = x$.

Vamos calcular a conjugada de f para alguns subconjuntos de \mathbb{C} . Notemos que

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

Consideremos primeiramente o caso em que

$$p \in \{p : \mathbb{R} \rightarrow \mathbb{R} \mid p(x) = a_1x + a_0; (a_1, a_0) \in \mathbb{R}^2, a_1 < 0\}.$$

Assim $f^*(p) = \sup_{x \in \mathbb{R}^n} \{a_1x^2 + a_0x - x\}$. Seja $h(x) = a_1x^2 + a_0x - x$, então podemos concluir que $h'(x) = 2a_1x + a_0 - 1$, de modo que o único ponto crítico de h é $\bar{x} = (a_0 - 1)/2a_1$. Como $h''(x) = 2a_1 < 0$, pelas condições de otimalidade, \bar{x} é maximizador de h . Deste modo,

$$f^*(p) = \frac{a_1(a_0 - 1)^2}{4a_1^2} + \frac{a_0(a_0 - 1)}{2a_1} - \frac{a_0 - 1}{2a_1} = \frac{a_1(a_0 - 1)^2}{4a_1^2} + \frac{(a_0 - 1)^2}{2a_1}.$$

Já para o caso em que $p \in \{p : \mathbb{R} \rightarrow \mathbb{R} \mid p(x) = a_3x^3; a_3 < 0\}$, temos que $f^*(p) = \sup_{x \in \mathbb{R}^n} \{a_3x^4 - x\}$. Considerando $h(x) = a_3x^4 - x$, então $h'(x) = 4a_3x^3 - 1$. Assim, o único ponto crítico de h é $\bar{x} = 1/\sqrt[3]{4a_3}$. Desde que $h''(x) = 12a_3x^2 \leq 0$, pelas condições de otimalidade, \bar{x} é maximizador de h . Deste modo,

$$f^*(p) = \frac{a_3}{(\sqrt[3]{4a_3})^4} - \frac{1}{\sqrt[3]{4a_3}} = \frac{-3}{4\sqrt[3]{4a_3}}.$$

Para o caso particular em que $p(x) = x$ temos que

$$f^*(p) = \sup_{x \in \mathbb{R}^n} \{x^2 - x\} = +\infty.$$

Como podemos perceber, existe dificuldade em obter uma expressão geral para a conjugada devido ao fato de estarmos trabalhando em \mathbb{C} . No entanto, veremos mais adiante que para os nossos propósitos não será necessário avaliar a conjugada em todo o espaço.

Lembramos que para calcular a conjugada, assim como no caso de Fenchel, podemos nos limitar somente ao domínio da função. Assim,

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

Vamos agora estender de maneira natural algumas definições para funções cujo domínio esteja em \mathbb{C} .

Definição 3.6 Consideremos a função $g : C \rightarrow \mathbb{R} \cup \{+\infty\}$. O domínio e o epigrafo de g serão definidos respectivamente por

$$\begin{aligned} \text{dom}(g) &= \{p \in C \mid g(p) < +\infty\}; \\ \text{epi}(g) &= \{(p, \lambda) \in C \times \mathbb{R} \mid g(p) \leq \lambda\}. \end{aligned}$$

Assim como no caso em que o domínio da função está em \mathbb{R}^n , diremos que a função $g : C \rightarrow \mathbb{R} \cup \{+\infty\}$ é sci quando seu epigrafo é fechado em $C \times \mathbb{R}$ e diremos que g é convexa quando seu epigrafo é convexo. Além disso, quando g for uma função contínua, também será sci. Quando $\text{dom}(g)$ for não-vazio, diremos que g é própria.

Podemos também estender o Teorema 1.2. Logo, se a função g é o supremo pontual de funções sci e convexas então ela também será sci e convexa. Com este resultado, podemos garantir o seguinte teorema.

Teorema 3.7 Dada uma função $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ própria sci. A função conjugada f^* é própria, convexa e sci.

Demonstração. Por definição, $\text{dom}(f^*) = D_f$. Pelo Teorema 3.3, sabemos que f^* é própria. Para provar a convexidade e semicontinuidade, vamos considerar para cada $x \in \mathbb{R}^n$, a função $T_x(p) = \langle p, x \rangle - f(x)$. Lembramos que $T_x(p) = b$ descreve um hiperplano em C . Além disso, como já vimos no capítulo anterior, T_x é contínuo e linear para todo $x \in \mathbb{R}^n$, ou seja, T_x é sci e convexa. Como $f^*(p) = \sup_{x \in \mathbb{R}^n} \{T_x(p)\}$, concluímos a demonstração.

□

Vamos verificar algumas propriedades da conjugada de Fenchel que continuam valendo nesta generalização.

Proposição 3.8 Seja $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função própria. A conjugada de f possui as seguintes propriedades:

- (i) Para toda $h : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ tal que $f(x) \leq h(x)$ para todo $x \in \mathbb{R}^n$, teremos que $f^*(p) \geq h^*(p)$ para todo $p \in C$;
- (ii) $f(x) + f^*(p) \geq \langle p, x \rangle$ para todo $x \in \mathbb{R}^n$ e $p \in C$.
- (iii) $\inf_{x \in \mathbb{R}^n} f(x) = -f^*(0)$.

Demonstração. (i) Se $f(x) \leq h(x)$ para todo $x \in \mathbb{R}^n$, então

$$\langle p, x \rangle - h(x) \leq \langle p, x \rangle - f(x)$$

para todo $x \in \mathbb{R}^n$. Logo,

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

para todo $p \in C$.

(ii) O resultado segue direto da definição da conjugada e da definição de supremo.

(iii) De fato,

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

□

De acordo com a extensão dada por Moreau na Definição 3.1, podemos construir uma conjugada para funções cujo domínio esteja em C .

Definição 3.9 Considere a função própria $g : C \rightarrow \mathbb{R} \cup \{+\infty\}$. A função conjugada de g , denotada por g^* , será definida por $g^* : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$;

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

Vamos provar a semicontinuidade dessa conjugada, no entanto não será possível garantir a convexidade.

Teorema 3.10 Dada uma função própria $g : C \rightarrow \mathbb{R} \cup \{+\infty\}$, a função conjugada g^* é sci.

Demonstração. Para cada $p \in C$, consideremos a função $\phi_p(x) = \langle p, x \rangle - g(p)$. Notemos que ϕ_p é contínuo para todo p . De fato, como o produto interno de \mathbb{R}^n pode ser encarado como uma função contínua, então ϕ_p é composição de funções contínuas e portanto, também é contínua.

Como $g^*(x) = \sup_{p \in C} \{\phi_p(x)\}$, pelo Teorema 1.2, concluímos que g^* é sci. □

Não podemos seguir a ideia da demonstração do Teorema 3.7 para verificar a convexidade de ϕ_p , pois a linearidade dessa função está dependendo da linearidade da função p e muitas funções contínuas não são lineares.

Podemos concluir características similares as da Proposição 3.8 para esta conjugada. A demonstração segue analogamente.

Proposição 3.11 Seja $g : C \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função própria. A conjugada de g possui as seguintes propriedades:

- (i) Para toda $h : C \rightarrow \mathbb{R} \cup \{+\infty\}$ tal que $g(p) \leq h(p)$ para todo $p \in C$, teremos que $g^*(x) \geq h^*(x)$ para todo $x \in \mathbb{R}^n$;
- (ii) $g(p) + g^*(x) \geq \langle p, x \rangle$ para todo $x \in \mathbb{R}^n$ e $p \in C$.
- (iii) $\inf_{p \in C} g(p) = -g^*(0)$.

Com o auxílio da Definição 3.9, podemos calcular a conjugada de f^* , também denominada biconjugada e denotada por f^{**} . Deste modo, ela é definida como $f^{**} : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$;

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

Analogamente, a biconjugada da função g , denotada por g^{**} , é definida como $g^{**} : C \rightarrow \mathbb{R} \cup \{+\infty\}$;

$$g^{**}(p) = (g^*)^*(x) = \sup_{x \in \mathbb{R}^n} \{\langle p, x \rangle - g^*(x)\}.$$

Como no caso da conjugada de Fenchel, é possível caracterizar a biconjugada f^{**} como supremo de funções majoradas pela f .

Teorema 3.12 Seja $H = \{h : \mathbb{R}^n \rightarrow \mathbb{R} \mid h(y) = \langle p, y \rangle + b, p \in C, b \in \mathbb{R}\}$. Considere-mos $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função própria, então

$$f^{**}(x) = \sup_{h \in H} \{h(x) \mid h(y) \leq f(y), \forall y \in \mathbb{R}^n\}.$$

A demonstração segue analogamente ao do Teorema 1.45 se considerarmos o produto interno generalizado. Como consequência direta, temos o seguinte corolário.

Corolário 3.13 *Dada uma função própria f e sua biconjugada f^{**} então*

$$f^{**}(x) \leq f(x)$$

para todo $x \in \mathbb{R}^n$.

Podemos enunciar uma versão semelhante do Teorema 3.12 para a biconjugada de g , cuja demonstração é análoga.

Teorema 3.14 *Seja $H = \{h : \mathbb{C} \rightarrow \mathbb{R} \mid h(q) = \langle q, x \rangle + b, x \in \mathbb{R}^n, b \in \mathbb{R}\}$. Se $g : \mathbb{C} \rightarrow \mathbb{R} \cup \{+\infty\}$ é uma função própria, então*

$$g^{**}(p) = \sup_{h \in H} \{h(p) \mid h(q) \leq g(q), \forall q \in \mathbb{C}\}.$$

Com a ajuda desses resultados, podemos calcular a biconjugada do seguinte exemplo.

Exemplo 3.15 $f : \mathbb{R} \rightarrow \mathbb{R}$;

$$f(x) = \begin{cases} 0, & x \leq 0 \\ 1, & x > 0 \end{cases}.$$

Primeiramente, vamos considerar $x \leq 0$. Pelo Corolário 3.13, $f^{**}(x) \leq f(x) = 0$. Agora, tomando $q \in \mathbb{C}$ tal que $q(y) = 0$ para todo $y \in \mathbb{R}^n$, temos que

$$f^{**}(x) \geq q(x)x - f^*(q) = -f^*(0) = \inf_{y \in \mathbb{R}} f(y) = 0.$$

Logo, $f^{**}(x) = f(x) = 0$.

Para cada $x > 0$, já sabemos também que $f^{**}(x) \leq f(x) = 1$. Consideremos $q \in \mathbb{C}$ definida por

$$q(y) = \frac{-3y^3}{x^4} + \frac{4y^2}{x^3}$$

para todo $y \in \mathbb{R}^n$. Notemos que

$$q(x) = \frac{-3x^3}{x^4} + \frac{4x^2}{x^3} = \frac{-3}{x} + \frac{4}{x} = \frac{1}{x}.$$

Vamos verificar que $f^*(q) = 0$.

De fato, $f^*(q) = \sup_{y \in \mathbb{R}} \{\langle q, y \rangle - f(y)\}$. Consideremos

$$h(y) = \langle q, y \rangle - f(y) = \begin{cases} \frac{-3y^4}{x^4} + \frac{4y^3}{x^3}, & y \leq 0 \\ \frac{-3y^4}{x^4} + \frac{4y^3}{x^3} - 1, & y > 0 \end{cases}$$

Calculando a derivada de h para $y > 0$, podemos concluir que

$$h'(y) = \frac{-12y^3}{x^4} + \frac{12y^2}{x^3}$$

e assim, o único ponto crítico positivo é x . Como

$$h''(y) = \frac{-36y^2}{x^4} + \frac{24y}{x^3}$$

então $h''(x) = -12/x^2 < 0$ e portanto x é maximizador de h . Como $h(x) = 0$ e $\frac{-3y^4}{x^4} + \frac{4y^3}{x^3} \leq 0$ para todo $y \leq 0$, então $f^*(q) = 0$.

Deste modo, $f^*(q) + f(x) = q(x)x$. Assim

$$f^{**}(x) \geq q(x)x - f^*(q) = f(x) = 1$$

e, portanto $f^{**}(x) = f(x) = 1$.

Neste exemplo, temos que $f^{**} = f$ apesar da função f ser apenas sci. Como f não é convexa, pelo Teorema 1.48, $f_F^{**}(x) < f(x)$ para algum $x \in \mathbb{R}^n$ e de fato, isso ocorre para todo $x > 0$, pois

$$f_F^*(a) = \sup_{x \in \mathbb{R}^n} \{xa - f(x)\} = \begin{cases} 0, & a = 0 \\ +\infty, & a \neq 0 \end{cases}$$

Deste modo,

$$f_F^{**}(x) = \sup_{a \in \mathbb{R}^n} \{xa - f_F^*(a)\} = \sup_{a \in \text{dom}(f_F^*)} \{xa - f_F^*(a)\} = 0 < f(x)$$

para todo $x > 0$.

Veremos mais adiante que o Teorema 1.48 será mais geral quando estivermos trabalhando com a conjugada generalizada. A próxima proposição relaciona as Definições 3.4 e 3.9.

Proposição 3.16 *Sejam $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ e $g : \mathbb{C} \rightarrow \mathbb{R} \cup \{+\infty\}$ funções próprias. São equivalentes:*

(i) $f(x) = g^*(x)$ e $g(p) = g^{**}(p)$ para $x \in \mathbb{R}^n$ e $p \in \mathbb{C}$;

(ii) $g(p) = f^*(p)$ e $f(x) = f^{**}(x)$ para $x \in \mathbb{R}^n$ e $p \in \mathbb{C}$.

Demonstração. (i) \Rightarrow (ii) Como $g^*(x) = f(x)$ então $g^{**}(p) = f^*(p)$, logo

$$g(p) = g^{**}(p) = f^*(p).$$

Com isso, temos que $g^*(x) = f^{**}(x)$. Portanto

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

(ii) \Rightarrow (i) Como $f^*(p) = g(p)$ então $f^{**}(x) = g^*(x)$, logo

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

Assim, temos que $f^*(p) = g^{**}(p)$. Portanto

$$g(p) = f^*(p) = g^{**}(p).$$

□

Baseado nessa proposição, dizemos que f e g são *funções conjugadas* quando elas satisfazem (i) ou (ii). Na próxima seção, veremos quais as condições necessárias para f e f^* sejam conjugadas.

3.3 Espaços duais conjugados

Para cada subespaço vetorial S de C e para uma função própria $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$, podemos restringir a conjugada de f apenas ao subespaço S . Denotaremos essa restrição por f_S^* . O mesmo pode ser feito para a biconjugada e assim a função $f_S^{**} : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ será definida como

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

Lembramos que F denota o subespaço de C formado por todas as funções constantes e assim, conciliando com a notação anterior, f_F^* é a conjugada de Fenchel. Quando estivermos nos referindo a todo espaço C , denotaremos a conjugada apenas por f^* .

Para S, V subespaços vetoriais de C tais que $S \subset V$, temos que, por propriedade de supremo,

$$f_S^{**}(x) = \sup_{p \in S} \{\langle p, x \rangle - f^*(p)\} \leq \sup_{p \in V} \{\langle p, x \rangle - f^*(p)\} = f_V^{**}(x)$$

para todo $x \in \mathbb{R}^n$. Assim, podemos provar a seguinte proposição.

Proposição 3.17 *Seja S um subespaço vetorial de C . Então*

$$f_S^{**}(x) \leq f(x)$$

para todo $x \in \mathbb{R}^n$.

Demonstração. Como $S \subset C$ então $f_S^{**}(x) \leq f^{**}(x)$ para todo x em \mathbb{R}^n . Pelo Corolário 3.13, podemos concluir que

$$f_S^{**}(x) \leq f^{**}(x) \leq f(x)$$

para todo $x \in \mathbb{R}^n$. □

Novamente, estamos interessados em analisar quando ocorre a igualdade na proposição. Veremos a seguir que os subespaços em que a igualdade é satisfeita serão chamados de duais conjugados da função.

Definição 3.18 *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função própria, dizemos que um subespaço vetorial S de C é um espaço dual conjugado de f quando satisfaz:*

- (i) $S \supset F$
- (ii) $f(x) = f_S^{**}(x)$ para todo $x \in \mathbb{R}^n$.

Antes de analisar alguns exemplos, vamos ver o seguinte teorema.

Teorema 3.19 *Consideremos $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função própria tal que V é um espaço dual conjugado de f . Então todo subespaço S de C que contém V é dual conjugado de f . Em particular, f e f^* são funções conjugadas.*

Demonstração. Se $V \subset S$, então $f_V^{**}(x) \leq f_S^{**}(x)$ para todo $x \in \mathbb{R}^n$. Como V é um espaço dual conjugado de f , então $F \subset V$ e $f(x) = f_V^{**}(x) \leq f_S^{**}(x)$. Logo, $F \subset S$ e pela Proposição 3.17, $f(x) = f_S^{**}(x)$ para todo $x \in \mathbb{R}^n$.

Em particular, C é um espaço dual conjugado de f e portanto, f e f^* são conjugadas. □

Exemplo 3.20 Uma função $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ própria, sci e convexa.

Pelo Teorema 1.48, o subespaço F é um espaço dual conjugado de f . Já pelo Teorema 3.19, qualquer outro subespaço de C que contenha F também é dual conjugado de f . Portanto, f e f^* são conjugadas. Assim, vemos que a conjugada generalizada preservou esta importante propriedade da conjugada de Fenchel. O próximo exemplo mostra que ela vale para funções não convexas.

Exemplo 3.21 $f : \mathbb{R} \rightarrow \mathbb{R}$;

$$f(x) = \begin{cases} 0, & x \leq 0 \\ 1, & x > 0 \end{cases} .$$

Como vimos no Exemplo 3.15, F não é um espaço dual conjugado de f . No entanto, se considerarmos $P = \{p \in C; p(x) = a_3x^3 + a_2x^2 + a_0, (a_3, a_2, a_0) \in \mathbb{R}^3\}$, vemos que $f(x) = f_P^*(x)$ para todo $x \in \mathbb{R}^n$ e portanto P é um espaço dual conjugado de f . Além disso, podemos concluir que f e f^* são conjugadas, ou seja, C é um espaço dual conjugado de f . Veremos a seguir que essa propriedade vale somente para funções sci.

Teorema 3.22 Seja $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função própria. A função f é sci se, e somente se, C é um espaço dual conjugado de f .

Demonstração. (\Rightarrow) Dado $x \in \mathbb{R}^n$, vamos provar inicialmente que $\lambda < f^{**}(x)$ para todo $\lambda < f(x)$. De fato, para $\lambda < f(x)$ temos que $(x, \lambda) \notin \text{epi}(f)$ e portanto, pelo Teorema 2.14, existe $p \in C$ tal que

$$\langle p, z \rangle - \lambda_1 < \langle p, x \rangle - \lambda$$

para todo $(z, \lambda_1) \in \text{epi}(f)$. Em particular,

$$\langle p, z \rangle - f(z) < \langle p, x \rangle - \lambda$$

para todo $z \in \text{dom}(f)$. Logo

$$f^*(p) = \sup_{z \in \text{dom}(f)} \{\langle p, z \rangle - f(z)\} < \langle p, x \rangle - \lambda.$$

Portanto,

$$\lambda < \langle p, x \rangle - f^*(p) \leq f^{**}(x) \tag{3.1}$$

para todo $\lambda < f(x)$.

Agora, pelo Corolário 3.13, temos que $f^{**}(x) \leq f(x)$ para todo $x \in \mathbb{R}^n$. Supondo que $f^{**}(x) < f(x)$ e considerando

$$\lambda = \frac{f(x) + f^{**}(x)}{2},$$

temos que $\lambda < f(x)$ e assim, por (3.1),

$$f^{**}(x) > \lambda = \frac{f(x) + f^{**}(x)}{2},$$

implicando que $f^{**}(x) > f(x)$ que é um absurdo. Portanto, $f^{**}(x) = f(x)$ para todo $x \in \mathbb{R}^n$.

(\Leftarrow) Se C é um espaço dual conjugado de f , então $f(x) = f^{**}(x)$ para todo $x \in \mathbb{R}^n$. Pelo Teorema 3.10, f^{**} é sci em todo \mathbb{R}^n e portanto, f também é. \square

Dessa forma, o propósito foi cumprido. Foi construída uma teoria de conjugação que se comporta bem para funções sci, nos possibilitando construir uma teoria de dualidade para estas funções. Para encerrar, o próximo corolário apresenta um resumo dos principais resultados do capítulo.

Corolário 3.23 *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função própria. Os seguintes itens são equivalentes:*

- (i) *f é sci.*
- (ii) *$f(x) = f^{**}(x)$ para todo $x \in \mathbb{R}^n$.*
- (iii) *C é um espaço dual conjugado de f .*
- (iv) *f e f^* são funções conjugadas.*

Capítulo 4

Programação semicontínua

Neste capítulo, vamos fazer um breve estudo a respeito dos problemas de programação semicontínua. Com o auxílio dos resultados obtidos anteriormente, apresentaremos uma teoria de dualidade para estes problemas seguindo o modelo da programação convexa e obteremos resultados análogos. Por fim, analisaremos a função Lagrangeana relacionada a estes problemas. Esta teoria é construída em [7].

4.1 O problema *PSCI*

Nos problemas de programação semicontínua temos dois objetivos: minimizar uma função $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ sci num conjunto não-vazio e fechado ou maximizar uma função $\bar{f} : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{-\infty\}$ scs também num conjunto não-vazio e fechado.

Porém, esses dois objetivos são equivalentes, pois minimizar f é o mesmo que maximizar $-f$ e pelo Teorema 1.21, $-f$ é scs. Logo, podemos analisar somente um dos problemas, porque os resultados serão análogos para ambos os casos. Assim, sem perda de generalidade, vamos fazer um estudo do problema de minimizar f , que é conhecido na literatura como Programação Semicontínua Inferior e denotado por *PSCI*.

Antes de tratar do problema, vamos modificar a função f de acordo com os nossos interesses. Como o nosso propósito é minimizar a função f sci num subconjunto não-vazio e fechado K de \mathbb{R}^n , neste capítulo, cometendo um abuso de notação, vamos sempre considerar f da seguinte forma:

$$f(x) = \begin{cases} f(x), & x \in K \\ +\infty, & x \notin K \end{cases}.$$

Assim, os pontos que não estão no conjunto K são descartados nesse processo.

Seguindo essa convenção, quando nos referirmos ao domínio de f , estaremos falando do domínio da função f restrita ao conjunto K e diremos que f é própria quando existe $x \in K$ de modo que $f(x) < +\infty$. O problema *PSCI* é reformulado como

$$\text{minimizar } f(x)$$

$$\text{sujeito a } x \in \mathbb{R}^n,$$

com $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ própria e sci e K um subconjunto não-vazio e fechado de \mathbb{R}^n .

Definição 4.1 *O conjunto solução de um problema *PSCI*, denotado por $S(f, K)$, é definido como*

$$S(f, K) = \{x' \in K \mid f(x') \leq f(x) \forall x \in K\}.$$

Notemos que, de acordo com a maneira como a f está sendo considerada, os conjuntos $S(f, K)$ e $S(f, \mathbb{R}^n)$ coincidem. Além disso, é possível perceber que $S(f, K) = \bigcap_{x \in K} (K \cap L_f(f(x)))$, permitindo concluir pelo Teorema 1.19 que $S(f, K)$ é sempre fechado. A seguir, vamos expor uma inclusão que é válida para o conjunto de recessão de $S(f, K)$.

Lema 4.2 $(S(f, K))^\infty \subset \bigcap_{x \in K} (K \cap L_f(f(x)))^\infty$.

Demonstração. Notemos que, para cada $x \in K$, $S(f, K) \subset K \cap L_f(f(x))$. Portanto, $(S(f, K))^\infty \subset (K \cap L_f(f(x)))^\infty$, permitindo concluir o resultado. \square

Pelo Teorema 1.23, quando K for limitado, $S(f, K)$ é não-vazio e portanto, existe solução para *PSCI*. Muitos autores se preocuparam em encontrar condições que possibilitassem generalizar esse teorema com base na análise dos *Problemas de Equilíbrio*.

Definição 4.3 Considere um subconjunto K não-vazio, convexo e fechado em \mathbb{R}^n e $f : K \times K \rightarrow \mathbb{R}$ satisfazendo as três propriedades seguintes:

- (i) $f(x, x) = 0$ para todo $x \in K$.
- (ii) $f_x(y) = f(x, y)$ é sci e convexa para cada x fixado em \mathbb{R}^n .
- (iii) $f_y(x) = f(x, y)$ é scs para cada y fixado em \mathbb{R}^n .

O problema de equilíbrio, denotado por *PE*, consiste em encontrar $x \in K$ tal que $f(x, y) \geq 0$ para todo $y \in K$.

Notemos que para uma função f sci, convexa e própria, se considerarmos $F(x, y) = f(y) - f(x)$, então F satisfaz as três condições de um *PE* para um conjunto não-vazio, fechado e convexo K . Assim, resolvê-lo é o mesmo que achar o minimizador de f em K . Logo, o problema de minimização é um caso particular dos problemas de equilíbrio.

Em [10] demonstrou-se a existência de solução do *PE* caso K fosse compacto. Já em [3] e [5], os autores demonstraram a existência de solução do *PE* substituindo a compacidade de K por condições de coercividade de f . Em [14], fez-se uso de aproximações assintóticas do conjunto K para obter novas condições de coercividade de f , dentre elas, o desenvolvimento da condição *P5*, dada a seguir:

Definição 4.4 Uma função f cumpre a condição *P5* quando para toda sequência $(x_k) \subset K / \{0\}$ satisfazendo os seguintes itens:

- (i) $\lim_{k \rightarrow \infty} \|x_k\| = +\infty$;
- (ii) $\left(\frac{x_k}{\|x_k\|}\right)$ converge fracamente⁴ para um ponto x tal que $f(y, y + x) \leq 0$ para todo $y \in K$;

Existe $(u_k) \subset K$ tal que, para todo k suficientemente grande,

$$\|u_k\| < \|x_k\| \quad e \quad f(x_k, u_k) \leq 0.$$

⁴Em \mathbb{R}^n , uma sequência (x_k) converge fracamente para x quando, para cada $a \in \mathbb{R}^n$, $\langle a, x_k \rangle$ converge para $\langle a, x \rangle$

Provou-se que $P5$ é condição necessária e suficiente para garantir a existência de solução do PE . Em [13], foram apresentadas condições equivalentes à $P5$. Com relação ao $PSCI$, em [12], a propriedade $P5$ foi adaptada em forma de duas condições equivalentes:

Definição 4.5 (i) *Uma função f cumpre a condição $P5'$ quando, para toda sequência $(x_k) \subset K$ tal que $\lim_{k \rightarrow \infty} \|x_k\| = +\infty$, existe $u \in K$ tal que para todo k suficientemente grande tem-se que $f(u) \leq f(x_k)$.*

(ii) *Uma função f cumpre a condição $P5''$ quando, para toda sequência $(x_k) \subset K$ tal que $\lim_{k \rightarrow \infty} \|x_k\| = +\infty$, existe $(u_k) \in K$ tal que para todo k suficientemente grande tem-se que $u_k \in L_f(x_k) \cap B(0, \|x_k\|)$.*

Finalmente, em [7], fazendo uso de técnicas de recessão, chegou-se a seguinte condição.

Definição 4.6 *Dizemos que uma função f cumpre a condição A quando, para toda sequência $(x_k) \subset \mathbb{R}^n$ tal que $\lim_{k \rightarrow \infty} \|x_k\| = +\infty$, existe $k_0 \in \mathbb{N}$ tal que*

$$L_f(f(x_{k_0})) \cap B(0, \|x_{k_0}\|) \neq \emptyset.$$

Quando K é limitado, qualquer sequência que satisfaça as hipóteses da condição A não pode estar contida em K . Dessa maneira, existe k_0 tal que $f(x_{k_0}) = +\infty$. Portanto, $L_f(f(x_{k_0})) = \mathbb{R}^n$ e a função f relacionada cumpre trivialmente a condição. Vamos agora provar que a condição A é, de fato, necessária e suficiente para garantir existência de solução do $PSCI$.

Teorema 4.7 *Dado um problema $PSCI$, são equivalentes:*

- (i)** *A é satisfeita.*
- (ii)** *$S(f, K)$ é não-vazio.*

Demonstração. (\Rightarrow) Como f é própria, então existe $\bar{x} \in K$ tal que $f(\bar{x}) < +\infty$. Assim, para cada $k \in \mathbb{N}$, consideremos

$$f_k(x) = \begin{cases} f(x), & \|x\| \leq \|\bar{x}\| + k \\ +\infty, & \|x\| > \|\bar{x}\| + k \end{cases}.$$

Cada f_k é própria, então existe $\bar{y} \in K$ tal que $f_k(\bar{y}) = f(\bar{y}) < +\infty$ e $\|\bar{y}\| \leq \|\bar{x}\| + k$.

Notemos que

$$K \cap L_{f_k}(f_k(\bar{y})) = \{x \in K \mid f_k(x) \leq f_k(\bar{y}) = f(\bar{y}) < +\infty\} \subset \overline{B(0, \|\bar{x}\| + k)},$$

ou seja, $K \cap L_{f_k}(f_k(\bar{y}))$ é limitado. Pelo Teorema 1.28, $(K \cap L_{f_k}(f_k(\bar{y})))^\infty = \{0\}$ então

$$(S(f_k, K))^\infty \subset \bigcap_{x \in K} (K \cap L_{f_k}(f_k(x)))^\infty = \{0\} \subset (S(f_k, K))^\infty \Rightarrow (S(f_k, K))^\infty = \{0\}.$$

Deste modo, $S(f_k, K)$ é limitado e assim, compacto.

Como f_k é sci, pois seu epigrafo é fechado, tem-se que $\text{epi}(f_k) = \text{epi}(f) \cap (\overline{B(0, \|\bar{x}\| + k)} \times \mathbb{R})$, e que seu domínio é compacto, então, pelo Teorema 1.23, $S(f_k, K)$ é não-vazio.

Lembramos que a função norma é contínua, então podemos considerar

$$x_k = \text{Min}(k) = \{x \in S(f_k, K) \mid \|x\| \leq \|y\| \text{ para todo } y \in S(f_k, K)\}.$$

Como f_k é própria, então $x_k \in \text{dom}(f)$ e $\|x_k\| \leq \|\bar{x}\| + k$. Além disso, o conjunto $\text{Min}(k)$ não é necessariamente unitário, mas as propriedades que veremos dessa sequência independem da escolha do elemento deste conjunto.

Suponha que (x_k) não seja limitada, então existe subsequência $(x_k)_{k \in \mathbb{N}'}$ tal que $x_k \xrightarrow{\mathbb{N}'} +\infty$.

Como A é satisfeita por hipótese, então existe $k_0 \in \mathbb{N}'$ tal que $L_f(f(x_{k_0})) \cap B(0, \|x_{k_0}\|) \neq \emptyset$.

Para $z \in L_f(f(x_{k_0})) \cap B(0, \|x_{k_0}\|)$, temos que

$$\|z\| < \|x_{k_0}\| \leq \|\bar{x}\| + k_0 \text{ e } f(z) \leq f(x_{k_0}) = f_{k_0}(x_{k_0}) < +\infty.$$

Portanto, $z \in \text{dom}(f) \subset K$ e $f_{k_0}(z) = f(z) \leq f_{k_0}(x_{k_0})$, implicando que $z \in S(f_{k_0}, K)$. Absurdo, pois x_{k_0} é o elemento que possui menor norma em $S(f_{k_0}, K)$.

Deste modo, (x_k) é limitada e assim, existe uma subsequência $(x_k)_{k \in \mathbb{N}'}$ convergente para x' . Vamos provar que $x' \in S(f, X)$.

De fato, notemos que $f_{k+1}(x) \leq f_k(x)$ para todo $k \in \mathbb{N}'$. Assim,

$$f(x_k) = f_k(x_k) \leq f_k(x) \leq f_{k-i}(x) \tag{4.1}$$

para todo $x \in \mathbb{R}^n$ e natural i menor que k . Isso nos permite afirmar que $f(x') \leq f_k(x)$ para todo $k \in \mathbb{N}'$ e $x \in \mathbb{R}^n$.

Pois, se para algum x existir $k' \in \mathbb{N}'$ tal que $f_{k'}(x) < f(x')$ então, pelo Corolário 1.16, existe $\delta > 0$ tal que $f(y) > f_{k'}(x)$ para todo $y \in B(x', \delta)$. Em particular, $f(x_k) > f_{k'}(x)$ para k em \mathbb{N}' e maior que k' , contrariando (4.1).

Por fim, para cada $x \in \mathbb{R}^n$, existe um $k \in \mathbb{N}'$ tal que $\|x\| \leq \|\bar{x}\| + k$ implicando que

$$f(x') \leq f_k(x) = f(x),$$

concluindo a primeira parte da demonstração.

(\Leftarrow) Consideremos uma sequência $(x_k) \subset \mathbb{R}^n$ tal que $\lim_{k \rightarrow \infty} \|x_k\| = +\infty$. Como $S(f, K)$ é não-vazio, tomando $z \in S(f, K)$ e k_0 suficientemente grande de modo que $\|x_{k_0}\| > \|z\|$, concluímos a demonstração. \square

Assim, qualquer outra condição de existência de solução do *PSCI* é equivalente a condição A e além disso, o Teorema 4.7 generaliza o Teorema de Weierstrass para funções sci. Na próxima seção, vamos construir a teoria de dualidade para problemas *PSCI* que tenham solução, ou seja, para funções que satisfaçam a condição A .

4.2 Dualidade semicontínua

Dado um *PSCI*, para a dualidade que será proposta consideraremos $\varphi : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R} \cup \{+\infty\}$ uma função de perturbação de f quando φ for própria sci e $\varphi(x, 0) = f(x)$ para todo $x \in \mathbb{R}^n$. Notemos que ela sempre existe, pois podemos considerar $\varphi(x, y) = f(x)$ para todo $x \in \mathbb{R}^n$.

Para simplificar a notação, a partir deste ponto do trabalho, vamos considerar $C_n = C(\mathbb{R}^n, \mathbb{R}^n)$ e $C_m = C(\mathbb{R}^m, \mathbb{R}^m)$. Deste modo,

$$C(\mathbb{R}^n \times \mathbb{R}^m, \mathbb{R}^n \times \mathbb{R}^m) \approx C_n \times C_m = \{(p_1, p_2) \mid p_1 \in C_n, p_2 \in C_m\}.$$

Podemos também estender a noção de produto interno generalizado usando elementos de $C_n \times C_m$ da seguinte forma:

$$\langle (p_1, p_2), (x, y) \rangle = \langle p_1, x \rangle_{C_n} + \langle p_2, y \rangle_{C_m} = \langle p_1(x), x \rangle + \langle p_2(y), y \rangle,$$

para $x \in \mathbb{R}^n$ e $y \in \mathbb{R}^m$.

Assim, a função conjugada φ^* será calculada como

$$\varphi^*(p_1, p_2) = \sup_{(x,y) \in \mathbb{R}^n \times \mathbb{R}^m} \{ \langle p_1, x \rangle + \langle p_2, y \rangle - \varphi(x, y) \}$$

para p_1 em C_n e p_2 em C_m e a biconjugada φ^{**} será dada por

$$\varphi^{**}(x, y) = \sup_{(p_1, p_2) \in C_n \times C_m} \{ \langle p_1, x \rangle + \langle p_2, y \rangle - \varphi^*(p_1, p_2) \}$$

para x em \mathbb{R}^n e y em \mathbb{R}^m .

Com essa generalização, podemos recuperar os resultados obtidos no capítulo anterior. Dessa maneira, pelo Teorema 3.22, o espaço $C_n \times C_m$ é o dual conjugado de φ .

Vamos agora apresentar uma versão da Proposição 1.51 para a conjugada modificada. Este resultado é a principal motivação para generalizar o processo apresentado na Seção 1.6.

Teorema 4.8 *Para todo $p \in C_m$ e $x \in \mathbb{R}^n$, temos que*

$$-\varphi^*(0, p) \leq f(x)$$

Em particular,

$$\sup_{p \in C_m} -\varphi^*(0, p) \leq \inf_{x \in \mathbb{R}^n} f(x).$$

Demonstração. Pela Proposição 3.8, para p em C_m , observemos que

$$\varphi^*(0, p) \geq \langle 0, x \rangle + \langle p, y \rangle - \varphi(x, y)$$

para todo x e y em \mathbb{R}^n . Tomando $y = 0$, concluímos a demonstração. \square

Como vimos no Capítulo 3, a conjugação modificada é simétrica para funções sci. Deste modo, generalizamos a teoria de dualidade para o PSCI da seguinte maneira.

Definição 4.9 *O problema dual associado a PSCI com respeito à φ , denotado por DPSCI, é definido como*

$$\begin{aligned} & \text{minimizar } \varphi^*(0, p) \\ & \text{sujeito a } p \in C_m. \end{aligned}$$

Notemos que, do Teorema 3.7, a função $G(p) = \varphi^*(0, p)$ é convexa. De fato, dados p_1 e p_2 em C_m e $t \in [0, 1]$, temos, pela convexidade da função conjugada que

$$\begin{aligned} G((1-t)p_1 + tp_2) &= \varphi^*(0, (1-t)p_1 + tp_2) = \varphi^*((1-t)(0, p_1) + t(0, p_2)) \leq \\ &\leq (1-t)\varphi^*(0, p_1) + t\varphi^*(0, p_2) = (1-t)G(p_1) + tG(p_2). \end{aligned}$$

Assim, o DPSCI é um problema de programação convexa, o qual é bem visto aos olhos da Otimização. Vamos agora analisar algumas propriedades da dualidade convexa que continuam valendo nessa nova teoria. Lembramos que

$$\inf_{p \in C_m} \varphi^*(0, p) = - \sup_{p \in C_m} -\varphi^*(0, p).$$

Do Teorema 4.8, segue diretamente o próximo corolário.

Corolário 4.10 *Dados PSCI e DPSCI referente a alguma perturbação φ , são válidas:*

- (i) *Se existe $\bar{x} \in \mathbb{R}^n$ tal que $f(\bar{x}) = \sup_{p \in C_m} -\varphi^*(0, p)$ então \bar{x} é solução do problema primal.*
- (ii) *Se existe $\bar{p} \in C_m$ tal que $-\varphi^*(0, \bar{p}) = \inf_{x \in \mathbb{R}^n} f(x)$ então \bar{p} é solução do problema dual.*
- (iii) *Se existem $\bar{x} \in \mathbb{R}^n$ e $\bar{p} \in C_m$ tais que $-\varphi^*(0, \bar{p}) = f(\bar{x})$ então \bar{x} é solução para DPSCI e \bar{p} é solução para PSCI.*

Como DPSCI é um problema de programação convexa então é possível construir o seu dual de forma análoga ao que foi feito para desenvolver o bidual na Seção 1.6. Assim sendo, consideremos $w(p) = \varphi^*(0, p)$. A função $\bar{\varphi}(p, p_1) = \varphi^*(p_1, p)$ é uma função perturbação de w .

Deste modo, o problema dual de DPSCI, denominado por bidual e denotado por DDPSCI, é caracterizado por

$$-\sup_{x \in \mathbb{R}^n} \{-\bar{\varphi}^*(0, x)\} = \inf_{x \in \mathbb{R}^n} \{\varphi^{**}(x, 0)\}.$$

Como φ é sci, então, do Teorema 3.22, $\varphi^{**}(x, 0) = \varphi(x, 0) = f(x)$ para todo $x \in \mathbb{R}^n$. Logo, DDPSCI é o próprio PSCI.

Portanto, a dualidade desenvolvida é bem comportada. No entanto, percebemos que o DPSCI é um problema de minimização em C_m , o qual possui dimensão infinita. Em alguns casos, como problemas de minimização quadrática, é possível substituir C_m por algum de seus subespaços de dimensão finita, devido ao fato de que as funções quadráticas possuem espaços duais conjugados de dimensão finita. Para ver mais detalhes, consulte [7].

4.3 A função Lagrangeana

Vamos agora introduzir a função Lagrangeana referente ao PSCI. Através dessa função, podemos encontrar uma nova condição para que o valor ótimo dos problemas primal e dual coincida.

Definição 4.11 *Dada uma função $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ própria, definimos a função Lagrangeana $L : \mathbb{R}^n \times C_m \rightarrow \mathbb{R} \cup \{-\infty, +\infty\}$ relacionada a função perturbação φ como*

$$L(x, p) = \inf_{y \in \mathbb{R}^n} \{\varphi(x, y) - \langle p, y \rangle\}.$$

Com o auxílio das propriedades da conjugada generalizada, podemos garantir duas proposições. A primeira relaciona o supremo em C_m da função Lagrangeana com a f e a segunda relaciona o ínfimo em \mathbb{R}^n da Lagrangeana com a conjugada de φ .

Proposição 4.12 *Para cada $x \in \mathbb{R}^n$,*

$$f(x) = \sup_{p \in C_m} L(x, p)$$

Demonstração. Para cada $x \in \mathbb{R}^n$, consideremos $\theta_x(y) = \varphi(x, y)$. Vamos dividir a demonstração em dois casos:

(1º caso: θ_x é imprópria.) Assim $\theta_x(y) = +\infty$ para todo $y \in \mathbb{R}^n$ e, em particular, $f(x) = \varphi(x, 0) = \theta_x(0) = +\infty$. Por outro lado, $L(x, p) = \inf_{y \in \mathbb{R}^n} \{\theta_x(y) - \langle p, y \rangle\} = +\infty$ para todo $p \in C_m$. Logo,

$$\sup_{p \in C_m} L(x, p) = +\infty = f(x).$$

(2º caso: θ_x é própria.) Notemos que θ_x é sci. De fato, para cada $y_0 \in \mathbb{R}^n$, consideremos $m' < \theta_x(y_0) = \varphi(x, y_0)$. Pela semicontinuidade de φ , existe $\delta > 0$ tal que $\varphi(z, y) > m'$ para todo $(z, y) \in B((x, y_0), \delta)$. Em particular, $\theta_x(y) = \varphi(x, y) > m'$ para todo $y \in B(y_0, \delta)$. Pelo Corolário 1.16, θ_x é sci em todo $y_0 \in \mathbb{R}^n$.

Agora, observemos que $L(x, p) = -\sup_{y \in \mathbb{R}^n} \{\langle p, y \rangle - \theta_x(y)\}$. Deste modo,

$$\sup_{p \in C_m} L(x, p) = \sup_{p \in C_m} \left\{ - \sup_{y \in \mathbb{R}^n} \{ \langle p, y \rangle - \theta_x(y) \} \right\} = \sup_{p \in C_m} -\theta_x^*(p) = \sup_{p \in C_m} \{ \langle p, 0 \rangle - \theta_x^*(p) \} = \theta_x^{**}(0).$$

Pelo Teorema 3.22, podemos concluir que

$$\sup_{p \in C_m} L(x, p) = \theta_x^{**}(0) = \theta_x(0) = \varphi(x, 0) = f(x).$$

□

Proposição 4.13 Para cada $p \in C_m$,

$$\inf_{x \in \mathbb{R}^n} L(x, p) = -\varphi^*(0, p).$$

Demonstração. Dado $p \in C_m$, temos que

$$\begin{aligned} \inf_{x \in \mathbb{R}^n} L(x, p) &= \inf_{x \in \mathbb{R}^n} \left\{ \inf_{y \in \mathbb{R}^n} \{ \varphi(x, y) - \langle p, y \rangle \} \right\} = \inf_{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n} \{ \varphi(x, y) - \langle (0, p), (x, y) \rangle \} = \\ &= - \sup_{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n} \{ \langle (0, p), (x, y) \rangle - \varphi(x, y) \} = -\varphi^*(0, p). \end{aligned}$$

□

Combinando estes dois últimos resultados, temos que $-\varphi^*(0, p) \leq L(x, p) \leq f(x)$ para todo $x \in \mathbb{R}^n$ e $p \in C_m$. Assim, podemos concluir o seguinte teorema:

Teorema 4.14 Dado um problema PSCI, uma função de perturbação φ e sua função Lagrangeana relacionada L . O ponto $(\bar{x}, \bar{p}) \in \mathbb{R}^n \times C_m$ satisfaz a equação

$$\inf_{x \in \mathbb{R}^n} L(x, \bar{p}) = L(\bar{x}, \bar{p}) = \sup_{p \in C_m} L(\bar{x}, p) \tag{4.2}$$

se, e somente se, $f(\bar{x}) + \varphi^*(0, \bar{p}) = 0$.

Demonstração. (\Rightarrow) Notemos que

$$f(\bar{x}) = \sup_{p \in C_m} L(\bar{x}, p) = \inf_{x \in \mathbb{R}^n} L(x, \bar{p}) = -\varphi^*(0, \bar{p}).$$

Logo, $f(\bar{x}) + \varphi^*(0, \bar{p}) = 0$.

(\Leftarrow) Sabemos que $\inf_{x \in \mathbb{R}^n} L(x, \bar{p}) \leq L(\bar{x}, \bar{p}) \leq \sup_{p \in C_m} L(\bar{x}, p)$. Por outro lado, temos que

$$\sup_{p \in C_m} L(\bar{x}, p) = f(\bar{x}) = -\varphi^*(0, \bar{p}) = \inf_{x \in \mathbb{R}^n} L(x, \bar{p}).$$

Portanto

$$\inf_{x \in \mathbb{R}^n} L(x, \bar{p}) = L(\bar{x}, \bar{p}) = \sup_{p \in C_m} L(\bar{x}, p)$$

Com este último resultado, podemos considerar a seguinte relação para as soluções do par de problemas primal e dual. □

Corolário 4.15 *Dados PSCI e DPSCI com respeito à função de perturbação φ e L a função Lagrangeana relacionada. Se o ponto (\bar{x}, \bar{p}) satisfaz a equação (4.2) então*

- (i) \bar{x} é solução do PSCI.
- (ii) \bar{p} é solução do DPSCI.

Dessa maneira, se o valor ótimo de um par de problemas primal e dual coincide e a solução \bar{p} do problema dual é conhecida, \bar{x} é solução do problema primal se, e somente se, (\bar{x}, \bar{p}) cumpre (4.2). Deste modo, podemos usar esta equivalência para resolver o problema primal, pois, para alguns casos, a função Lagrangeana possui uma boa estrutura numérica.

Conclusão

Neste trabalho, o nosso propósito foi relacionar [6], [7] e [26] de uma maneira clara e sistêmica. Estas três referências tratam de generalizar resultados da Literatura. Assim, apresentamos os conceitos clássicos no processo construtivo em que eles estavam inseridos e seguimos o mesmo roteiro para apresentar esta relação.

Inicialmente, fizemos uma revisão dos teoremas de separação de convexos em \mathbb{R}^n . Vimos que, para dois subconjuntos convexos tais que a diferença deles é fechada, existe um hiperplano que os separa estritamente. Graças a esse resultado, o epigrafo de uma função convexa e *sci* pode ser entendido como a interseção dos semi-espacos que o contêm. Essa possibilidade nos permitiu definir a conjugada clássica e garantir que, para estas funções, a biconjugada é igual a própria função.

No contexto de otimização, mostramos que a teoria de conjugação de Fenchel pode ser aplicada na construção de uma dualidade para Programação Convexa.

Analisamos uma particularidade do Teorema de Seleção de Michael a qual separa com funções contínuas o epigrafo de uma função *sci* do hipografo de uma função *scs*. Provamos que esta função contínua descreve a equação de um hiperplano generalizado e, considerando os subconjuntos fechados como epigrafo e hipografo de funções *sci* e *scs* respectivamente, foi possível generalizar os teoremas clássicos de separação para subconjuntos fechados.

Consequentemente, generalizamos de maneira natural a teoria de conjugação de Fenchel usando o produto interno generalizado, pois a Separação de Fechados garantiu que essa conjugada proposta seria própria para funções *sci*. Observamos que essa generalização manteve certas propriedades da conjugação clássica como convexidade e semicontinuidade, no entanto, o domínio da função conjugada passou a ser o espaço dos operadores contínuos.

Para simplificar os cálculos, introduzimos os espaços duais conjugados que são subespaços de C cuja conjugação de uma função restrita a eles é simétrica. Com isso, recuperamos a conjugada de Fenchel para o caso das funções *sci* e convexas. Além disso, vimos exemplos de espaços duais conjugados de dimensão finita. O principal resultado obtido foi que a biconjugada generalizada de funções *sci* é a própria função.

Por fim, analisamos os problemas de Programação Semicontínua e vimos algumas condições existentes na literatura que garantem a existência de solução. Aplicamos a conjugada generalizada na construção de uma teoria de dualidade para estes problemas e verificamos que o Dual do *PSCI* é um problema de Minimização Convexa e o bidual é o próprio problema original. Definimos a função Lagrangeana relacionada e concluímos que o seu minimizador em \mathbb{R}^n é solução do problema Primal e o seu maximizador em C é solução do problema Dual.

Referências Bibliográficas

- [1] A. Auslender e M. Teboulle. *Asymptotic cones and functions in Optimization and Variational Inequalities*. Springer Monographs in Mathematics, 2003.
- [2] D. P. Bertsekas, A. Nedic e A. E. Ozdaglar. *Convex Analysis and Optimization*. Athena Scientific, Belmont, Massachusetts, 2003.
- [3] E. Blum e W. Oettli. *From optimization and variational inequalities to equilibrium problems*, Math. Student 63, 1994, pg. 123-145.
- [4] H. Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer, 2010.
- [5] H. Brezis, L. Nirenberg e G. Stampacchia. *A remark on Ky Fan's minimax principle*, Boll. Un. Mat. Ital 6, 1972, pg. 293-300.
- [6] J. Cotrina, E. W. Karas, A. A. Ribeiro, W. Sosa e Y. J. Yun. *Moreau-Fenchel conjugation for lower semi-continuous functions*, Optimization, 2011, pg. 1-13.
- [7] J. Cotrina, F. M. P. Raupp e W. Sosa. *Semi-continuous quadratic programming: existence conditions and duality scheme*. Artigo em processo de submissão.
- [8] J.P. Crouzeix, A. Keraghel e W. Sosa. *Programación Matemática Diferenciable*. Universidad Nacional de Ingeniería, Lima, Peru, 2011.
- [9] I. Ekeland e R. Témam. *Convex Analysis and Variational Problems*. Society for Industrial and Applied Mathematics, Philadelphia, EUA, 1999.
- [10] K. Fan. *A minimax inequality and applications*, O. Shisha, Inequality III, Academic Press, New York, 1972, pg. 103-113.
- [11] W. Fenchel. *On conjugate convex functions*. Canad. J. Math I, 1949, pg. 73-77.
- [12] A. N. Iusem, G. Kassay e W. Sosa. *An existence results for equilibrium problems with some surjectivity consequences*, Journal of Convex Analysis, v. 16, 2009, pg. 807-826.
- [13] A. N. Iusem, G. Kassay e W. Sosa *On certain conditions for the existence of solutions of equilibrium problems*, Mathematical Programming, v. 116, 2009, pg. 259-273.
- [14] A. N. Iusem, W. Sosa *New existence results for equilibrium problems*, J. of Nonlinear An. Theory Methods and Applications, v. 52, 2003, pg. 621-635.
- [15] A. Izmailov e M. V. Solodov. *Condições de Otimalidade, Elementos de Análise Convexa e de Dualidade*. Instituto de Matemática Pura e Aplicada, Rio de Janeiro, Brasil, 2009.

- [16] E. W. Karas e A. A. Ribeiro. *Um curso de Otimização*, Notas de aula, UFPR, Curitiba, Brasil, 2011.
- [17] E. L. Lima. *Curso de Análise volume 2*. Instituto de Matemática Pura e Aplicada, Rio de Janeiro, Brasil, 1981.
- [18] E. L. Lima. *Espaços Métricos*. 4^a ed. Instituto de Matemática Pura e Aplicada, Rio de Janeiro, Brasil, 2007.
- [19] J. E. Martínez-Legaz. *Generalized convex duality and its economic applications*, Handbook of Generalized Convexity and Generalized Monotonicity, v.76, 2005, pg. 237-292.
- [20] J.J. Moreau *Inf-convolution, sous-additivité, convexité des fonctions numériques*. Journal of Mathematics, v. 63, pg 361-382, 1956.
- [21] E. Michael. *Continuos selection I*. Annals of Mathematics, v. 63, pg. 361-382, 1956.
- [22] J. V. Mill. *The Infinite-Dimensional Topology of Function Spaces*. Elsevier Science B. V., Amsterdam, 2001.
- [23] M. R. Pinheiro. *Conjugação e Dualidade em Programação Convexa*. Workin paper 25, Universidade Nova de Lisboa, Portugal, 1984.
- [24] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, Nova Jersey, 1970.
- [25] R. T. Rockafellar. e R. J. B. Wets *Variational Analysis*. Springer, 1998.
- [26] W. Sosa. *Representation of continuous functions and its applications*. Artigo em processo de submissão.