

UNIVERSIDADE FEDERAL DE SANTA CATARINA
DEPARTAMENTO DE MATEMÁTICA

Asteroide Santana

CÁLCULO SUBDIFERENCIAL COM
 ε -SUBDIFERENCIAIS E LIMITES

Florianópolis

2013

Asteroide Santana

**CÁLCULO SUBDIFERENCIAL COM
 ε -SUBDIFERENCIAIS E LIMITES**

Dissertação submetida ao Curso de Pós-Graduação em Matemática Pura e Aplicada para a obtenção do Grau de Mestre em Matemática.

Orientador: Prof. Dr. Maicon Marques Alves

Florianópolis

2013

Catálogo na fonte elaborada pela biblioteca da
Universidade Federal de Santa Catarina

A ficha catalográfica é confeccionada pela Biblioteca Central.

Tamanho: 7cm x 12 cm

Fonte: Times New Roman 9,5

Maiores informações em:

<http://www.bu.ufsc.br/design/Catalogacao.html>

Asteroide Santana

**CÁLCULO SUBDIFERENCIAL COM
 ε -SUBDIFERENCIAIS E LIMITES**

Esta Dissertação foi julgada aprovada para a obtenção do Título de “Mestre em Matemática”, e aprovada em sua forma final pelo Curso de Pós-Graduação em Matemática Pura e Aplicada.

Florianópolis, 21 de fevereiro 2013.

Prof. Dr. Daniel Gonçalves
Coordenador

Banca Examinadora:

Prof. Dr. Maicon Marques Alves
Orientador

Prof. Dr. Alfredo Noel Iusem

Prof. Dr. Ruy Coimbra Charão

Prof. Dr. Danilo Royer

À minha esposa Denise, minha condição
de qualificação.

AGRADECIMENTOS

A conclusão deste trabalho representa, sem dúvida, um grande marco na minha carreira profissional. A conquista deste título superou muito as minhas expectativas, principalmente pelo nível de aproveitamento que pude alcançar. Expectativas estas que foram ligeiramente minimizadas pelo processo seletivo de ingresso no programa (tal processo mudou completamente a partir 2010 - mudou para melhor, diga-se de passagem). Sei que ainda posso ir muito além, mas tenho muito orgulho e sou extremamente grato pelo que já conquistei. Ao mesmo tempo reconheço que não teria chegado tão longe (a noção de “longitude” aqui pode variar muito, mas do ponto de vista de um guri que até os 19 anos o que melhor sabia fazer era pescar no mar de Garopaba, um título de mestre em Matemática pode parecer muito distante!) se não fosse pela contribuição de algumas pessoas. Neste sentido direciono meus sinceros agradecimentos à minha esposa Denise, que sempre deu incentivo e apoio nas minhas decisões, foi uma excelente companhia e sabiamente se comportou nos momentos em que as dificuldades consumiam parte do meu humor. Agradeço meus pais que também sempre me apoiaram e que, dentro de suas condições, nunca mediram esforços para me dar a melhor educação. Sou grato também a minha irmã, que se mostrou prestativa durante toda a minha trajetória. De maneira muito especial, agradeço o professor Maicon Marques Alves que, com sua paciência, competência e entusiasmo, certamente foi a pessoa que mais influenciou na minha formação acadêmica. Sou muito horado por ele ter sido meu orientador. Faço questão de agradecer também os professores Gustavo da Costa, Clóvis Gonzaga, Ruy Charão e Danilo Royer, que também desempenharam papéis de destaque na minha formação e, em particular, o professor Alfredo Iusem, pelo grande incentivo que me deu no dia da defesa. Agradeço a todos os colegas pelos bons momentos de descontração e cooperação nos estudos. À secretária do curso de pós-graduação, Elisa Amaral, meu muito abrigado pela atenção prestada sempre que precisei. Por fim, agradeço a CAPES pelo auxílio financeiro sem o qual esta conquista certamente não teria sido possível.

RESUMO

Neste trabalho, estudamos regras de cálculo subdiferencial para soma de funções convexas, convolução infimal, pré-composição e função marginal. Enunciamos uma série de propriedades e resultados associados a estas operações, dentre os quais merecem destaque a *fórmula de dualidade de Fenchel-Rockafellar* e o *princípio variacional de Ekeland*, cuja demonstração apresentamos tendo em vista a relevância do resultado. Tudo isto é feito com o propósito de alcançar o principal objetivo: obter regras de cálculo subdiferencial. Algumas das fórmulas estabelecidas aqui são simples e dependentes de condições de qualificação, as demais são de caráter bem geral e contemplam o conceito de ε -subdiferencial. Cálculo subdiferencial via limites também faz parte do trabalho e permite algumas aplicações como o *teorema do valor médio* para subdiferenciais e o *teorema de integração de Rockafellar*.

Palavras-chave: Análise convexa, subdiferencial, cálculo subdiferencial, conjugada de Fenchel, convolução infimal, princípio variacional de Ekeland, ε -subdiferencial.

ABSTRACT

In this work, we study subdifferential calculus rules for the sum of convex functions, infimal convolution, pre-composition and marginal function. We present some important related results such as Fenchel-Rockafellar duality formula and the Ekeland variational principle. Our main goal is to obtain subdifferential calculus rule with and also without any qualification condition. We are also concerned with limiting subdifferential calculus.

Keywords: Convex analysis, subdifferential, Fenchel conjugate, subdifferential calculus, infimal convolution, ε -subdifferential, Ekeland variational principle.

SUMÁRIO

Introdução	17
1 NOTAÇÕES E RESULTADOS PRELIMINARES	21
1.1 CONJUGADA DE FENCHEL	26
1.2 CONVOLUÇÃO INFIMAL	27
1.3 SUBDIFERENCIAIS E ε -SUBDIFERENCIAIS.....	30
2 CÁLCULO SUBDIFERENCIAL COM CONDIÇÕES DE QUALIFICAÇÃO	35
2.1 SUBDIFERENCIAL DA SOMA	35
2.2 SUBDIFERENCIAL DA CONVOLUÇÃO INFIMAL	37
3 CÁLCULO SUBDIFERENCIAL COM ε-SUBDIFERENCIAIS	41
3.1 SUBDIFERENCIAL DA CONVOLUÇÃO INFIMAL	41
3.2 SUBDIFERENCIAL DA FUNÇÃO MARGINAL	47
4 CÁLCULO SUBDIFERENCIAL COM LIMITES	51
4.1 SUBDIFERENCIAIS DA SOMA E DA PRÉ-COMPOSIÇÃO	68
APÊNDICE	71
Conclusão	77
REFERÊNCIAS	79

INTRODUÇÃO

A área de análise convexa, mais especificamente a otimização convexa, vem apresentando um desenvolvimento acentuado desde a década de 60, em grande parte impulsionado pelos trabalhos de Fenchel, Moreau e Rockafellar. Inclusive, foi a partir da década de 60 que se manifestou grande interesse com relação a análise convexa em dimensão infinita, sobretudo em espaços de Banach. Neste sentido, podemos destacar as publicações de Ekeland e Teman [9], Moreau [17], Barbu e Precupanu [1] e Rockafellar [20].

O principal objeto de estudo deste trabalho, o subdiferencial, foi introduzido neste contexto e se transformou em uma ferramenta importantíssima da otimização convexa. Só o fato de a subdiferenciação ser uma generalização da diferenciação já nos permite concluir que o subdiferencial não é menos relevante do que a derivada. Além disso, se tratando de uma generalização, é esperado que muitos resultados clássicos possam ser estendidos. E de fato podem. Exemplos são: a condição de otimalidade conhecida no cálculo como o teorema de Fermat (ver (1.16)); o teorema do valor médio (Proposição 4.3); o teorema de integração de Rockafellar (Proposição 4.2). Levando adiante esta comparação, ainda podemos notar que para uma ampla classe de funções é possível se calcular o subdiferencial mesmo em pontos onde a derivada não existe (ver Exemplo 1.2). Isto, obviamente, torna a teoria ainda mais interessante. Vale observar ainda que o subdiferencial pode ser definido em espaços de dimensão infinita. Neste trabalho ele é definido para espaços de Banach.

Condições de otimalidade podem ser obtidas usando-se subdiferenciais e, por conta disto, algoritmos eficientes vêm sendo desenvolvidos para resolver problemas tanto em dimensão finita como em dimensão infinita nas mais diversas áreas (ver [5] para um exemplo em espaço de Hilbert). Existe, inclusive, uma noção de multiplicadores de Lagrange em dimensão infinita [23]. Uma das fórmulas apresentadas nesta dissertação (Proposição 4.2) foi utilizada por Thibault a fim de estabelecer existência de multiplicadores para problemas de otimização convexa com restrições [23].

Em parte, a aplicabilidade da otimização convexa se deve ao desenvolvimento, ainda em progresso, da teoria dos subdiferenciais. No que diz respeito a tal desenvolvimento, destacam-se as regras de cálculo, cujo propósito é fornecer de forma exata o subdiferencial de

combinações de funções f_i em termos dos subdiferenciais de cada f_i (combinações neste contexto podem ser operações tais como soma, supremo, pré-composição, convolução infimal, etc). Uma vez reconhecida a importância destas operações, torna-se evidente o valor das fórmulas de cálculo subdiferencial. Não são poucas as áreas beneficiadas com as deduções de tais fórmulas. Trabalhos publicados recentemente mostram avanços no tratamento de problemas provenientes da economia, problemas inversos e de recuperação de sinais [5, 6, 13].

As primeiras fórmulas de cálculo subdiferencial que foram deduzidas são ligeiramente restritivas, como aquelas apresentadas no Capítulo 2. Na literatura costuma-se referir a tais restrições como *condições de qualificação*. Fórmulas mais gerais e relacionadas a novas operações foram objeto de estudo nas últimas cinco décadas. Conforme pode ser observado ao longo deste trabalho, os resultados obtidos até então são promissores. Sobretudo, devido ao conceito de ε -subdiferencial, introduzido por Brøndsted e Rockafellar [3] ainda na década de 60.

Esta dissertação está baseada nos artigos de Hiriart-Urruty e Phelps [12] e de Thibault [22], e tem como principal objetivo apresentar regras de cálculo subdiferencial para algumas das operações mais recorrentes em análise convexa. Obedecendo a ordem cronológica dos fatos, trabalhamos primeiro as fórmulas que exigem algum tipo de condição de qualificação. Mesmo sendo restritivas, estas ainda merecem consideração devido às suas formas simplificadas. As fórmulas mais gerais dependem de ε -subdiferenciais e são mais sofisticadas, conforme pode ser observado no Capítulo 3 e no Capítulo 4.

Como de praxe, introduzimos um capítulo, Capítulo 1, exclusivamente para fixar algumas notações e apresentar as principais definições e os resultados que consideramos preliminares para este trabalho. Dentre os resultados enunciados neste capítulo merecem destaque a *fórmula de dualidade de Fenchel-Rockafellar* (Proposição 1.4) o bem conhecido *princípio variacional de Ekeland* (Proposição 1.2), cuja demonstração fizemos questão de apresentar. Ambos são, obviamente, usados no desenvolvimento dos capítulos subsequentes. Ainda no Capítulo 1, apresentamos a conjugada de Fenchel, a operação de convolução infimal e as definições de subdiferencial e ε -subdiferencial, bem como algumas de suas propriedades mais relevantes. No Capítulo 2, demonstramos duas fórmulas clássicas de cálculo subdiferencial com condição de qualificação, a saber, o subdiferencial da soma e da convolução infimal. O Capítulo 3 é baseado no artigo de Hiriart-Urruty e Phelps [12]. Nele, apresentamos algumas fórmulas para o subdiferencial da convolução infimal e outra para o subdiferencial da função marginal. Nenhuma delas

depende de condições de qualificação. Por fim, no Capítulo 4, que é baseado no artigo de Thibault [22], lançamos mão de algumas técnicas de limite para provar um dos principais teoremas deste trabalho (Teorema 4.1). Segue deste resultado a demonstração de uma fórmula para o subdiferencial da soma e da pré-composição, sem condição de qualificação. Finalizamos o Capítulo 4 com a demonstração do teorema do valor médio e o teorema de integração de Rockafellar.

1 NOTAÇÕES E RESULTADOS PRELIMINARES

Neste capítulo, introduzimos algumas notações e uma série de definições e resultados de análise convexa que são usados ao longo do trabalho. Tais resultados são fundamentais para o desenvolvimento dos capítulos seguintes. Em particular, destacamos o bem conhecido princípio variacional de Ekeland.

Em geral, estamos interessados em funções cujo contradomínio é o conjunto dos números reais estendido. Este será denotado por $\overline{\mathbb{R}}$ e consiste no conjunto dos número reais acrescido dos elementos $+\infty$ e $-\infty$.

Seja $\mathcal{X}$ um espaço vetorial e $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ uma função. O *domínio efetivo* (ou, simplesmente, *domínio*) de f é o conjunto

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

Nestas condições, f é dita *própria* se $f(x) > -\infty$, para todo x em $\mathcal{X}$, e seu domínio é não vazio. Diz-se que f é *convexa* se a desigualdade

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

se verifica para todo λ entre 0 e 1, e para quaisquer x e y em $\mathcal{X}$ tais que o lado direito de (1.1) está bem definido. O *epigrafo* de f consiste no conjunto

$$\text{epi}(f) = \{(x, \lambda) \in \mathcal{X} \times \mathbb{R}; f(x) \leq \lambda\}. \quad (1.2)$$

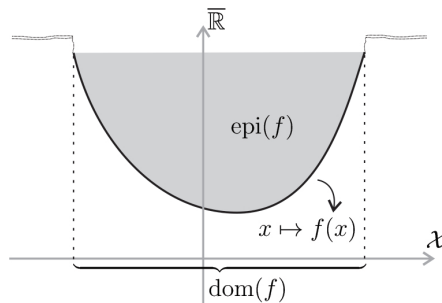


Figura 1: Representação gráfica do epigrafo.

A partir do epigrafo se obtém uma caracterização para funções

convexas, a saber, a função f é convexa se, e somente se, seu epigrafo é um conjunto convexo em $\mathcal{X} \times \mathbb{R}$ [1, Cap. 2, Proposição 1.2]. Sabe-se também que se f é convexa, então seu domínio é um subconjunto convexo de $\mathcal{X}$.

Diversas construções e demonstrações em análise convexa, em particular neste trabalho, usam técnicas envolvendo a função indicadora δ_C de um subconjunto C de $\mathcal{X}$. Por definição, $\delta_C(x) = 0$ se $x \in C$, e $\delta_C(x) = +\infty$, caso contrário.

Observação 1.0.1 *O domínio de δ_C é o próprio conjunto C e seu epigrafo é o conjunto $C \times \mathbb{R}_+$. Portanto, δ_C é convexa se, e somente se, C é um conjunto convexo.*

Outra classe de funções bastante significativa no contexto que estamos trabalhando é a das *funções semicontínuas inferiormente (s.c.i.)*.

Seja $\mathcal{X}$ um espaço de Banach e $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ uma função. Diz-se que f é s.c.i. em $x_0 \in \mathcal{X}$ se

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

A partir da definição de limite inferior¹, constata-se que a desigualdade contrária na expressão anterior é sempre válida. Portanto, f é s.c.i. em x_0 se, e somente se,

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

Diz-se que f é s.c.i. em $\mathcal{X}$ quando f é s.c.i. em todo ponto de $\mathcal{X}$. Caracterizações úteis para funções s.c.i. são dadas na proposição a seguir.

Proposição 1.1 [1, Cap. 2, Proposição 1.3] *Seja $\mathcal{X}$ um espaço de Banach e $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ uma função. São equivalentes:*

- (i) *A função f é s.c.i. em $\mathcal{X}$;*
- (ii) *Para qualquer α real, a imagem inversa do intervalo $(-\infty, \alpha]$ por f é um conjunto fechado;*
- (iii) *O epigrafo de f é fechado em $\mathcal{X} \times \mathbb{R}$.*

Segue diretamente da Proposição 1.1 que a função indicadora δ_C de um subconjunto C de $\mathcal{X}$ é s.c.i. se, e somente se, C é fechado.

¹Por definição, $\liminf_{x \rightarrow x_0} f(x) = \sup_{V \in \mathcal{V}(x_0)} \inf_{x \in V} f(x)$, onde $\mathcal{V}(x_0)$ é o conjunto das vizinhanças de x_0 .

Um resultado particularmente importante em análise convexa é o *princípio variacional de Ekeland*. Ele foi apresentado por Ivar Ekeland em 1974 [8] e desde então tem sido amplamente utilizado, não só na área de análise convexa como também em equações diferenciais parciais.

Na sequência, demonstramos o princípio variacional de Ekeland e um corolário, o qual é usado no Capítulo 4 para demonstrar um dos principais teoremas deste trabalho (Teorema 4.1). Tanto a demonstração da proposição como a do corolário foram adaptadas a partir de [15].

Proposição 1.2 (Princípio Variacional de Ekeland) *Seja $\mathcal{X}$ um espaço métrico completo munido da métrica d . Seja $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ uma função própria, s.c.i. e limitada inferiormente. Seja x_0 no domínio de f e $\varepsilon > 0$. Então, existe x_ε em $\mathcal{X}$ tal que*

- (i) $f(x_\varepsilon) \leq f(x_0) - \varepsilon d(x_\varepsilon, x_0)$;
- (ii) $f(x) > f(x_\varepsilon) - \varepsilon d(x_\varepsilon, x)$, $\forall x \neq x_\varepsilon$.

Demonstração: Para $n = 0, 1, 2, \dots$, seja

$$C_n := \{x \in \mathcal{X}; f(x) \leq f(x_n) - \varepsilon d(x, x_n)\},$$

onde (x_n) satisfaz:

- $x_{n+1} \in C_n$;
- $f(x_{n+1}) < S_n + \frac{1}{2^{n+1}}$, com $S_n = \inf \{f(x); x \in C_n\}$.

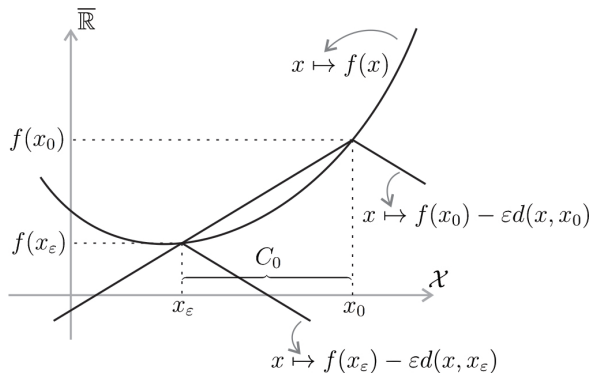


Figura 2: Interpretação geométrica do PVE.

Visto que f é limitada inferiormente, S_n é finito para todo n . Afiramos que $C_{n+1} \subset C_n$, para todo n . De fato, seja $x \in C_{n+1}$. Então,

$$\varepsilon d(x, x_{n+1}) \leq f(x_{n+1}) - f(x). \quad (1.3)$$

Como $x_{n+1} \in C_n$, temos que

$$\varepsilon d(x_{n+1}, x_n) \leq f(x_n) - f(x_{n+1}). \quad (1.4)$$

Da desigualdade triangular e de (1.3) e (1.4) obtemos

$$\varepsilon d(x, x_n) \leq \varepsilon d(x, x_{n+1}) + \varepsilon d(x_{n+1}, x_n) \leq f(x_n) - f(x).$$

Segue que $x \in C_n$. A afirmação que acabamos de provar nos permite constatar que $S_n \leq S_{n+1}$, para todo n . A partir daí podemos concluir que (x_n) é uma sequência de Cauchy. De fato, partindo de (1.4) temos

$$\varepsilon d(x_n, x_{n+1}) \leq f(x_n) - S_n \leq f(x_n) - S_{n-1} \leq \frac{1}{2^n}, \quad \forall n \geq 1.$$

Assim,

$$\sum_{n=0}^{\infty} d(x_n, x_{n+1}) < \infty \Rightarrow d(x_n, x_{n+1}) \rightarrow 0.$$

Como $\mathcal{X}$ é completo, existe $x_\varepsilon \in \mathcal{X}$ tal que x_n converge para x_ε . Vimos que $C_{n+1} \subset C_n$ para todo n , donde concluímos que $x_n \in C_0$, para todo n . Visto que f é s.c.i. e C_0 é um conjunto fechado, temos que $x_\varepsilon \in C_0$ (ver Proposição 1.1), isto é, vale (i).

Provemos (ii) pela contra positiva. Seja x em $\mathcal{X}$ tal que

$$f(x) \leq f(x_\varepsilon) - \varepsilon d(x, x_\varepsilon). \quad (1.5)$$

Devemos mostrar que $x = x_\varepsilon$. Afiramos que x_n converge a x . De fato, usando (1.5) e o fato que $x_\varepsilon \in C_n$, para todo n ,

$$\varepsilon d(x, x_n) \leq \varepsilon d(x, x_\varepsilon) + \varepsilon d(x_\varepsilon, x_n) \leq f(x_n) - f(x). \quad (1.6)$$

Portanto, $x \in C_n$, para todo n . Então, resulta da definição de C_n que $f(x) \geq S_n$, para todo n . Segue de (1.6) que

$$\varepsilon d(x, x_n) \leq f(x_n) - S_n \leq f(x_n) - S_{n-1} \leq \frac{1}{2^n}, \quad \forall n \geq 1.$$

Então, a série de termo geral $d(x, x_n)$ é finita e podemos concluir que

x_n converge a x . Por fim, segue da unicidade do limite que $x = x_\varepsilon$.

□

Sem assumir qualquer hipótese de compacidade, o corolário a seguir mostra que, se f possui um ε -minimizador², então a função

$$f_\lambda(x) := f(x) + \frac{\varepsilon}{\lambda}d(x, x_\lambda)$$

possui um minimizador global estrito no ponto x_λ , onde λ é um escalar.

Corolário 1.1 *Seja $\mathcal{X}$ um espaço métrico completo. Seja $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ uma função própria, s.c.i. e limitada inferiormente. Sejam $\varepsilon > 0$ e x_0 no domínio de f tais que*

$$f(x_0) \leq \inf_{x \in \mathcal{X}} f(x) + \varepsilon. \quad (1.7)$$

Então, para qualquer $\lambda > 0$, existe $x_\lambda \in \mathcal{X}$ tal que

$$(i) \quad f(x_\lambda) \leq f(x_0), \quad d(x_\lambda, x_0) \leq \lambda;$$

$$(ii) \quad f(x_\lambda) < f(x) + \frac{\varepsilon}{\lambda}d(x, x_\lambda), \quad \forall x \neq x_\lambda.$$

Demonstração: Com ε/λ no lugar de ε , a Proposição 1.2 garante que existe $x_\lambda \in \mathcal{X}$ tal que

$$(a) \quad f(x_\lambda) \leq f(x_0) - \frac{\varepsilon}{\lambda}d(x_\lambda, x_0);$$

$$(b) \quad f(x) > f(x_\lambda) - \frac{\varepsilon}{\lambda}d(x, x_\lambda), \quad \forall x \neq x_\lambda.$$

Portanto, (ii) segue diretamente de (b). Como ε e λ são positivos, segue de (a) que

$$f(x_\lambda) \leq f(x_0) \quad \text{e} \quad \frac{\varepsilon}{\lambda}d(x_\lambda, x_0) \leq f(x_0) - f(x_\lambda).$$

Por fim, resulta de (1.7) e da última inequação que

$$\frac{\varepsilon}{\lambda}d(x_\lambda, x_0) \leq \varepsilon \quad \Rightarrow \quad d(x_\lambda, x_0) \leq \lambda.$$

□

Daqui por diante, $\mathcal{X}$ sempre denota um espaço de Banach. O dual topológico de $\mathcal{X}$ é denotado por $\mathcal{X}^*$, e seus elementos por x^* . Além

²O ponto x_ε é dito ε -minimizador de f se $f(x_\varepsilon) \leq f(x) + \varepsilon$, para todo x .

disso, escrevemos $\langle x^*, x \rangle$ em vez de $x^*(x)$. Quando não existe risco de ambiguidades, usamos a mesma notação para as normas dos espaços primal e dual: $\|\cdot\|$. Para $(x, y) \in \mathcal{X} \times \mathcal{Y}$ e $(x^*, y^*) \in \mathcal{X}^* \times \mathcal{Y}^*$, definimos

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

Esta aplicação de dualidade é definida de maneira análoga para o produto cartesiano de três espaços.

1.1 CONJUGADA DE FENCHEL

Conforme já mencionamos, o subdiferencial de uma função é o principal objeto deste trabalho. A conjugada de Fenchel fornece uma caracterização (ver (1.18)) muito conveniente para os elementos do subdiferencial. Tal caracterização contribui para simplificar várias demonstrações que apresentamos nos próximos capítulos.

Seja $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ uma função convexa e própria. A função *conjugada* (de Fenchel) de f é definida de $\mathcal{X}^*$ em $\overline{\mathbb{R}}$ por

$$f^*(x^*) = \sup_{x \in \mathcal{X}} \langle x^*, x \rangle - f(x). \quad (1.8)$$

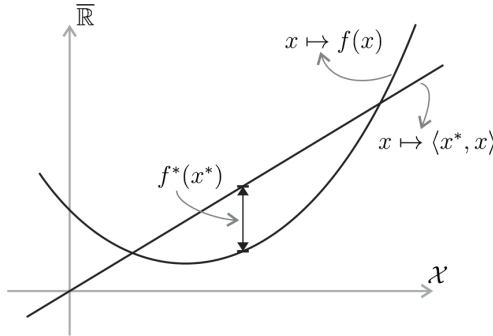


Figura 3: Interpretação geométrica da conjugada de Fenchel.

A conjugada de Fenchel é uma função convexa e s.c.i. na topologia fraca-* de $\mathcal{X}^*$ [15]. Além disso, se f é convexa, própria e s.c.i., então f^* é própria [15]. Segue diretamente da definição que

$$f^*(x^*) + f(x) \geq \langle x^*, x \rangle, \quad (1.9)$$

para todo $x \in \mathcal{X}$ e todo $x^* \in \mathcal{X}^*$. A inequação (1.9) é conhecida como *desigualdade de Fenchel-Young*. Por definição, a *biconjugada* de f é a função $f^{**} := (f^*)^*$. De maneira análoga define-se a função f^{***} . Resultados particularmente úteis com relação a conjugada de Fenchel são dados na proposição a seguir.

Proposição 1.3 [1, Cap. 2, Proposição 1.8] *Sejam $f, g : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ funções convexas e próprias. Então, cada uma das sentenças a seguir procedem.*

- (i) $f^{**}|_{\mathcal{X}} \leq f$;
- (ii) $f^{***}|_{\mathcal{X}^*} = f^*$;
- (iii) $f \leq g$ implica que $f^* \geq g^*$.

Na próxima proposição apresentamos a bem conhecida fórmula de dualidade de Fenchel-Rockafellar. Entre outras coisas ela fornece a informação de que, sob as hipóteses de (ii), o problema primal (lado esquerdo de (1.10)) pode não ter solução, mas o problema dual (lado direito de (1.10)) sempre tem solução. Neste trabalho, a fórmula de dualidade é usada para demonstrar a Proposição 1.6, a qual estabelece importantes relações entre a conjugada de Fenchel e a convolução infimal que vamos definir mais adiante.

Proposição 1.4 [2, Teorema 1.12] *Sejam $f, g : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ funções convexas e próprias. As seguintes afirmações são verdadeiras:*

- (i) $\inf_{x \in \mathcal{X}} f(x) + g(x) \geq \sup_{x^* \in \mathcal{X}^*} -f^*(-x^*) - g^*(x^*)$;
- (ii) *Se existe x_0 na interseção dos domínios de f e g tal que f é contínua em x_0 , então*

$$\inf_{x \in \mathcal{X}} f(x) + g(x) = \max_{x^* \in \mathcal{X}^*} -f^*(-x^*) - g^*(x^*). \quad (1.10)$$

1.2 CONVOLUÇÃO INFIMAL

No contexto que estamos trabalhando, a convolução infimal desperta interesse pela sua capacidade de regularização. Isto pode ser claramente observado em [5], em [2, Problema 14] e também na Proposição 3.1.

Sejam $f, g : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ funções convexas e próprias. A *convolução infimal* de f e g é a função $f \square g : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ definida por

$$(f \square g)(x) = \inf_{z \in \mathcal{X}} f(z) + g(x - z) = \inf_{x_1 + x_2 = x} f(x_1) + g(x_2). \quad (1.11)$$

A convolução é dita *exata* em um ponto x_0 de seu domínio se existe $z \in \mathcal{X}$ tal que

$$(f \square g)(x_0) = f(z) + g(x - z).$$

Além disso, f é dita *exata* em $\mathcal{X}$ se é exata em todo ponto de seu domínio. Algumas propriedades inerentes à convolução infimal são dadas na proposição a seguir.

Proposição 1.5 [4, Proposição 12.6 e Proposição 12.11] *Sejam $f, g : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ funções convexas e próprias. As seguintes afirmativas são verdadeiras:*

- (i) *O domínio de $f \square g$ é igual a soma³ dos domínios de f e de g ;*
- (ii) *Se f e g são limitadas inferiormente por uma mesma função afim⁴, então $f \square g$ é própria;*
- (iii) *$f \square g$ é convexa.*

Se $f \square g$ é exata, tem-se que a soma dos epígrafos de f e g coincide com o epígrafo de $f \square g$ (Ver Figura 4).

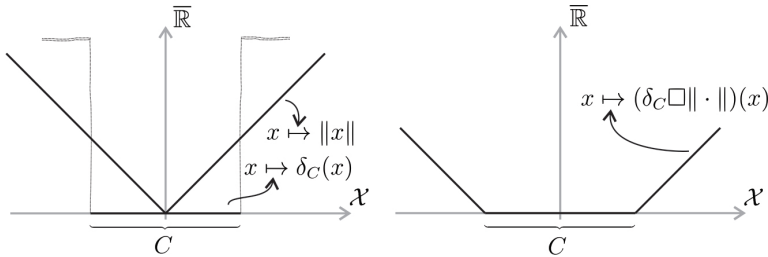


Figura 4: Interpretação geométrica da convolução infimal.

³Por definição, a soma de dois subconjuntos A e B de espaço vetorial qualquer é o conjunto $A + B := \{a + b; a \in A \text{ e } b \in B\}$.

⁴Uma função $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ é dita *afim* se f e $-f$ são convexas. Alternativamente, tem-se que f é afim se, e somente se, existem x^* em $\mathcal{X}^*$ e a em $\mathbb{R}$ tais que $f(x) = \langle x^*, x \rangle + a$, para todo $x \in \mathcal{X}$ [18, Teorema 1.5].

Seja C um subconjunto convexo de $\mathcal{X}$. Seja $f = \delta_C$ e $g = \|\cdot\|$. Então $f \square g$ fornece a distância de x ao conjunto C . De fato,

$$(f \square g)(x) = \inf_{z \in \mathcal{X}} \delta_C(z) + \|x - z\| = \inf_{z \in C} \|x - z\| = d(x, C).$$

Como g é uma função convexa, segue da Proposição 1.5, (iii), que a função distância é convexa. Cabe observar que é conveniente expressar a função distância como a convolução de duas funções. Dizemos isto porque existem fórmulas para o subdiferencial da convolução infimal (Proposição 2.1 e Teorema 3.1) e porque a função distância é usada com frequência em problemas de otimização.

A proposição a seguir estabelece relações entre a conjugada de Fenchel e a convolução infimal. Tais resultados são úteis em demonstrações que fazemos mais adiante.

Proposição 1.6 [18, Teorema 16.4] *Sejam $f, g : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ funções convexas e próprias. As seguintes afirmativas são verdadeiras:*

(i)

$$(f \square g)^* = f^* + g^*. \quad (1.12)$$

(ii) *Se existe x_0 na interseção dos domínios de f e g tal que f é contínua em x_0 , então*

$$(f + g)^* = f^* \square g^*. \quad (1.13)$$

Além disso, a convolução do lado direito de (1.13) é exata, isto é, para todo $x^ \in \mathcal{X}^*$, existe $z^* \in \mathcal{X}^*$ tal que*

$$(f^* \square g^*)(x^*) = f^*(z^*) + g^*(x^* - z^*).$$

Demonstração: Provemos (i). Partindo do lado esquerdo de (1.12), usando a definição de conjugada e, em seguida, a definição de convolução infimal, obtemos

$$\begin{aligned} (f \square g)^*(x^*) &= \sup_{x \in \mathcal{X}} [\langle x^*, x \rangle - \inf_{z \in \mathcal{X}} (f(z) + g(x - z))] \\ &= \sup_{z \in \mathcal{X}} \left[\sup_{x \in \mathcal{X}} (\langle x^*, x \rangle - f(z) - g(x - z)) \right], \quad \forall x^* \in \mathcal{X}^*. \end{aligned}$$

Após somar e subtrair z no argumento de x^* ,

$$\begin{aligned}(f \square g)^*(x^*) &= \sup_{z \in \mathcal{X}} \left[\sup_{x \in \mathcal{X}} (\langle x^*, x - z \rangle - g(x - z)) + \langle x^*, z \rangle - f(z) \right] \\ &= g^*(x^*) + \sup_{z \in \mathcal{X}} [\langle x^*, z \rangle - f(z)] \\ &= g^*(x^*) + f^*(x^*).\end{aligned}$$

Provemos (ii). A partir da definição de conjugada de Fenchel podemos escrever

$$(f + g)^*(x^*) = - \inf_{x \in \mathcal{X}} f(x) - \langle x^*, x \rangle + g(x), \quad \forall x^* \in \mathcal{X}^*. \quad (1.14)$$

Seja $h(x) := f(x) - \langle x^*, x \rangle$. As hipóteses asseguram que h é convexa e contínua em x_0 . Segue da Proposição 1.4 que

$$\begin{aligned}(f + g)^*(x^*) &= - \max_{z^* \in \mathcal{X}^*} -h^*(-z^*) - g^*(z^*) \\ &= \min_{z^* \in \mathcal{X}^*} h^*(-z^*) + g^*(z^*).\end{aligned}$$

Usando a definição de conjugada de Fenchel obtém-se que $h^*(-z^*) = f^*(x^* - z^*)$. Portanto,

$$(f + g)^*(x^*) = \min_{z^* \in \mathcal{X}^*} f^*(x^* - z^*) + g^*(z^*) = (f^* + g^*)(x^*).$$

A existência do mínimo nesta última expressão garante que a convolução é exata.

□

1.3 SUBDIFERENCIAIS E ε -SUBDIFERENCIAIS

Em áreas aplicadas, frequentemente surgem problemas de otimização cuja função objetivo não é diferenciável, mas tem a propriedade de ser convexa (por exemplo, a função objetivo pode ser uma combinação de funções lineares). Isto inviabiliza a aplicação de uma ampla classe de algoritmos. Contudo, para muitas funções, o subdiferencial está definido em pontos onde a derivada não existe. Esta característica e o fato de poder ser definido em espaços de dimensão infinita fazem do subdiferencial uma ferramenta fundamental em otimização convexa.

Seja $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ uma função convexa e própria. O *subdiferencial*

de f é o operador $\partial f : \mathcal{X} \rightarrow 2^{\mathcal{X}^*}$ dado por

$$\partial f(x) = \{x^* \in \mathcal{X}^*; \langle x^*, y - x \rangle \leq f(y) - f(x), \forall y \in \mathcal{X}\}, \quad (1.15)$$

para todo x no domínio de f . Definimos $\partial f(x) = \emptyset$ se x não está no domínio de f . O conjunto dos elementos x para os quais $\partial f(x)$ é não-vazio define o *domínio* de ∂f . Qualquer $x^* \in \partial f(x)$ é denominado *subgradiente* de f em x , e sua interpretação geométrica é dada na Figura 5: a função $y \mapsto f(x) + \langle x^*, y - x \rangle$ determina um hiperplano não-vertical que suporta o gráfico de f no ponto $(x, f(x))$.

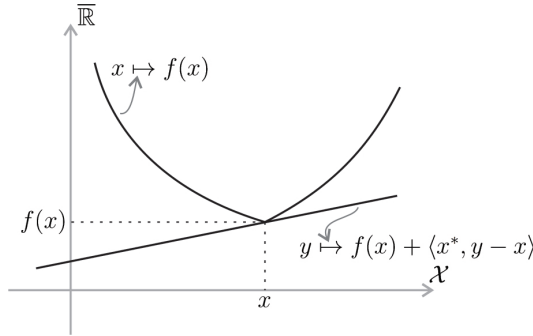


Figura 5: Interpretação geométrica do subgradiente.

Exemplo 1.1 *O subdiferencial da função norma no ponto zero coincide com a bola unitária de $\mathcal{X}^*$. De fato,*

$$x^* \in \partial \|\cdot\|(0) \Leftrightarrow \langle x^*, y \rangle \leq \|y\|, \forall y \in \mathcal{X} \Leftrightarrow x^* \in B[0, 1].$$

A interpretação geométrica sugere que a “subdiferenciação” é uma generalização da diferenciação. De fato, se f é Gâteaux diferenciável⁵ em x , então $\partial f(x) = \{\nabla f(x)\}$ [1, Cap. 2, Proposição 2.4].

Exemplo 1.2 *Consideremos a função $f = |\cdot|$ definida em $\mathbb{R}$. Para $x \neq 0$ temos que f é diferenciável, de modo que $\partial f(x) = \{-1\}$, se $x < 0$, e $\partial f(x) = \{1\}$, se $x > 0$. Para $x = 0$ temos $\partial f(0) = [-1, 1]$. Um esboço do gráfico de ∂f pode ser observado na Figura 6.*

⁵Sejam $\mathcal{X}$ um espaço métrico e $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ uma função. Seja x um ponto no interior do domínio de f . Nestas condições, f é dita *Gâteaux diferenciável* em x se a função $\langle \nabla f(x), h \rangle = \lim_{h \rightarrow 0+} \frac{f(x+h) - f(x)}{h}$ é linear e contínua na variável h .

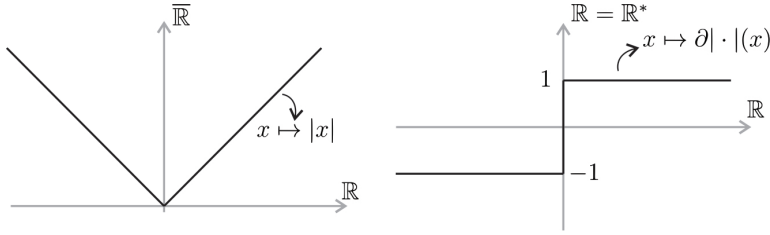


Figura 6: Gráfico do subdiferencial da função módulo.

Para qualquer $x \in \mathcal{X}$, o subdiferencial de f em x é um subconjunto convexo e fechado de $\mathcal{X}^*$ [15].

Observação 1.3.1 *Segue diretamente da definição que $\bar{x}$ é um minimizador (global) de f se, e somente se, o funcional identicamente nulo é um subgradiente de f em $\bar{x}$. De fato,*

$$0 \in \partial f(\bar{x}) \Leftrightarrow f(\bar{x}) \leq f(y), \forall y \in \mathcal{X}. \quad (1.16)$$

Em algumas demonstrações ao longo deste trabalho nos deparamos com o subdiferencial da função indicadora de um subconjunto convexo C de $\mathcal{X}$. Por isto, é importante observar que, para qualquer $x \in C$,

$$x^* \in \partial \delta_C(x) \Leftrightarrow \langle x^*, y - x \rangle \leq 0, \forall y \in C. \quad (1.17)$$

Esta equivalência segue diretamente da definição de subdiferencial e da definição de função indicadora.

A proposição a seguir fornece uma caracterização para os elementos do subdiferencial por meio da conjugada de Fenchel.

Proposição 1.7 [1, Cap. 2, Proposição 2.1] *Seja $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ uma função convexa e própria. Para qualquer x no domínio de f vale que*

$$x^* \in \partial f(x) \Leftrightarrow f^*(x^*) + f(x) = \langle x^*, x \rangle. \quad (1.18)$$

As principais regras de cálculo que apresentamos neste trabalho dependem do conceito de ε -subdiferencial. Este foi introduzido por Brøndsted e Rockafellar em 1965 [3] e consiste numa espécie de aproximação do subdiferencial. O ε -subdiferencial apresenta propriedades adicionais que permitem a dedução de fórmulas mais gerais, por exemplo, aquela que fornece o subdiferencial da convolução infimal de duas funções convexas, próprias e limitadas inferiormente (ver Teorema 3.1).

Seja $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ uma função convexa e própria. Seja $\varepsilon \geq 0$. O

ε -subdiferencial de f é o operador $\partial_\varepsilon f : \mathcal{X} \rightarrow 2^{\mathcal{X}^*}$ dado por

$$\partial_\varepsilon f(x) = \{x^* \in \mathcal{X}^*; \langle x^*, y - x \rangle \leq f(y) - f(x) + \varepsilon, \forall y \in \mathcal{X}\}, \quad (1.19)$$

para todo x no domínio de f . Definimos $\partial_\varepsilon f(x) = \emptyset$ se x não está no domínio de f . Para qualquer $x \in \mathcal{X}$, $\partial_\varepsilon f(x)$ é um subconjunto convexo e fechado de $\mathcal{X}^*$. Para $\varepsilon = 0$ o ε -subdiferencial coincide com subdiferencial. Dados ε_1 e ε_2 não-negativos tais que $\varepsilon_1 \leq \varepsilon_2$, segue diretamente da definição que $\partial_{\varepsilon_1} f(x)$ está contido em $\partial_{\varepsilon_2} f(x)$. Em particular,

$$\partial f(x) \subset \partial_\varepsilon f(x), \quad (1.20)$$

para qualquer x no domínio de f . Seja $\bar{\varepsilon} > 0$. Suponhamos que $x^* \in \partial_\varepsilon f(x)$, para qualquer $\varepsilon \in (0, \bar{\varepsilon}]$, isto é,

$$\langle x^*, y - x \rangle \leq f(y) - f(x) + \varepsilon,$$

para todo $y \in \mathcal{X}$ e para qualquer $\varepsilon \in (0, \bar{\varepsilon}]$. Após tomar o limite com ε tendendo a zero, levando em consideração a inclusão (1.20), constatamos que

$$\partial f(x) = \bigcap_{\varepsilon > 0} \partial_\varepsilon f(x) = \bigcap_{0 < \varepsilon \leq \bar{\varepsilon}} \partial_\varepsilon f(x). \quad (1.21)$$

Os elementos do ε -subdiferencial também podem ser caracterizados por meio da conjugada de Fenchel. De fato, para qualquer x no domínio de f e $\varepsilon > 0$, tem-se

$$x^* \in \partial_\varepsilon f(x) \Leftrightarrow f^*(x^*) + f(x) \leq \langle x^*, x \rangle + \varepsilon. \quad (1.22)$$

A demonstração desta equivalência é análoga a demonstração da Proposição 1.7.

2 CÁLCULO SUBDIFERENCIAL COM CONDIÇÕES DE QUALIFICAÇÃO

Neste capítulo, apresentamos regras de cálculo para o subdiferencial da soma e da convolução infimal. As fórmulas apresentadas nesta parte do trabalho se aplicam a funções convexas e próprias, sujeitas a condições de qualificação. Mesmo sendo restritivas, as fórmulas apresentadas nas próximas duas seções merecem destaques pela estrutura simplificada. Conforme pode ser observado nos capítulos seguintes, as fórmulas que não requerem qualificação são mais sofisticadas.

2.1 SUBDIFERENCIAL DA SOMA

Visto que a derivada da soma é igual a soma das derivadas, poderia-se esperar que o subdiferencial também satisfizesse tal propriedade - afinal, subdiferenciabilidade consiste em uma generalização de diferenciabilidade [1, Cap. 2, Proposição 2.4]. Infelizmente, não é necessariamente verdade que o subdiferencial da soma é igual a soma dos subdiferenciais (ver Exemplo 2.1). A princípio, apenas podemos garantir que a soma dos subdiferenciais está contida no subdiferencial da soma, isto é, dadas $f, g : \mathcal{X} \rightarrow \mathbb{R}$ funções convexas e próprias, vale que

$$\partial f(x) + \partial g(x) \subset \partial(f + g)(x), \forall x \in \mathcal{X}. \quad (2.1)$$

Provamos esta inclusão facilmente a partir da definição de subdiferencial, e a inclusão contrária assumindo uma condição de qualificação, a saber, existe x_0 na interseção dos domínios das duas funções tal que uma delas é contínua em x_0 . Uma fórmula mais geral (que não exige qualificação) para o subdiferencial da soma de duas funções convexas e próprias é dada na Proposição 4.2.

Teorema 2.1 [1, Cap. 3, Teorema 2.6] *Sejam $f, g : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ funções convexas e próprias. Suponha que f e g satisfazem a condição de qualificação: existe x_0 na interseção dos domínios de f e g tal que f é contínua em x_0 . Então, para qualquer $x \in \mathcal{X}$,*

$$\partial(f + g)(x) = \partial f(x) + \partial g(x). \quad (2.2)$$

Demonstração: Seja $x^* \in \partial(f + g)(x)$. Por (1.18) podemos escrever

$$(f + g)^*(x^*) + (f + g)(x) = \langle x^*, x \rangle. \quad (2.3)$$

Como vale a condição de qualificação, segue da Proposição 1.6 que existe $z^* \in \mathcal{X}^*$ tal que

$$(f + g)^*(x^*) = (f^* \square g^*)(x^*) = f^*(z^*) + g^*(x^* - z^*). \quad (2.4)$$

Substituindo (2.4) em (2.3) e fazendo algumas manipulações algébricas obtemos

$$[f^*(z^*) + f(x) - \langle z^*, x \rangle] + [g^*(x^* - z^*) + g(x) - \langle x^* - z^*, x \rangle] = 0.$$

A desigualdade (1.9) garante que cada uma das parcelas entre colchetes é não-negativa. Portanto, ambas são nulas. Seque por (1.18) que $z^* \in \partial f(x)$ e $x^* - z^* \in \partial g(x)$, ou seja,

$$x^* = z^* + (x^* - z^*) \in \partial f(x) + \partial g(x).$$

Seja $x^* \in \partial f(x) + \partial g(x)$. Então, existem $x_1^* \in \partial f(x)$ e $x_2^* \in \partial g(x)$ tais que $x^* = x_1^* + x_2^*$. A definição de subdiferencial fornece as inequações

$$\langle x_1^*, u - x \rangle \leq f(u) - f(x) \quad \text{e} \quad \langle x_2^*, u - x \rangle \leq g(u) - g(x), \quad \forall u \in \mathcal{X},$$

cuja soma resulta

$$\langle x^*, u - x \rangle \leq (f + g)(u) - (f + g)(x), \quad \forall u \in \mathcal{X}.$$

Portanto, $x^* \in \partial(f + g)(x)$.

□

Conforme mostra o exemplo a seguir, a condição de qualificação exigida nas hipóteses do Teorema 2.1 é realmente necessária.

Exemplo 2.1 [15] *Consideremos os conjuntos*

$$C := \{(x, y) \in \mathbb{R}^2; y \geq x^2\} \quad \text{e} \quad D := \{(x, y) \in \mathbb{R}^2; y = 0\}.$$

Ambos são convexos e, portanto, as funções δ_C e δ_D são convexas. Vamos mostrar que

$$\partial\delta_C(0, 0) + \partial\delta_D(0, 0) \subsetneq \partial(\delta_C + \delta_D)(0, 0). \quad (2.5)$$

Por (1.17) temos que

$$(x, y) \in \partial\delta_C(0, 0) \Leftrightarrow \langle (x, y), (u, v) \rangle \leq 0 \stackrel{\times}{\Leftrightarrow} xu + yv \leq 0,$$

para todo $(u, v) \in C$ ($*$ - visto que $(\mathbb{R}^n)^*$ está identificado com $\mathbb{R}^n$ por meio da representação de Riez, a operação de dualidade entre estes espaços coincide com o produto interno usual de $\mathbb{R}^n$). Como $v \geq 0$, basta tomar $u = 0$ na última inequação para concluir que $y \leq 0$. Analogamente, ao escolher $v = 0$, constata-se que $x = 0$. Então, $\partial\delta_C(0, 0)$ consiste no semi-eixo negativo das ordenadas. Procedendo-se de maneira análoga constata-se que $\partial\delta_D(0, 0)$ coincide com o eixo das ordenadas. Portanto,

$$\partial\delta_C(0, 0) + \partial\delta_D(0, 0) = \{(x, y) \in \mathbb{R}^2; x = 0\}.$$

Por outro lado, visto que $\delta_C + \delta_D = \delta_{C \cap D} = \delta_{\{(0,0)\}}$, constata-se de (1.17) que

$$\partial(\delta_C + \delta_D)(0, 0) = \mathbb{R}^2.$$

Portanto, $\partial\delta_C(0, 0) + \partial\delta_D(0, 0)$ é subconjunto próprio de $\partial(\delta_C + \delta_D)(0, 0)$.

O Exemplo 2.1 justifica a necessidade da fórmula apresentada no Teorema 4.2.

2.2 SUBDIFERENCIAL DA CONVOLUÇÃO INFIMAL

A próxima proposição fornece uma fórmula para o subdiferencial da convolução infimal de duas funções convexas e próprias. Contudo, a validade da fórmula depende de uma condição de qualificação, a saber, a convolução deve ser exata no ponto em questão.

Proposição 2.1 [12, Capítulo 1] *Sejam $f, g : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ funções convexas e próprias. Suponha que $\partial(f \square g)(x)$ é não-vazio para algum $x \in \mathcal{X}$. Então, são equivalentes:*

(i) *A convolução de f e g é exata em x , isto é, existe $z \in \mathcal{X}$ tal que*

$$\partial(f \square g)(x) = f(z) + g(x - z).$$

(ii) *$\partial f(z) \cap \partial g(x - z) \neq \emptyset$.*

(iii) *$\partial(f \square g)(x) = \bigcup_{y \in \mathcal{X}} [\partial f(y) \cap \partial g(x - y)]$.*

Demonstração: Provemos (i) $\Leftrightarrow$ (ii) e (i) $\Leftrightarrow$ (iii). Para provar que (i) implica (ii) tomemos $x^* \in \partial(f \square g)(x)$. De (1.18) e da Proposição 1.6(i), obtemos

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

Ou ainda, usando a hipótese,

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

Reorganizando as parcelas, após somar e subtrair z no argumento de x^* , obtemos

$$[f^*(x^*) + f(z) - \langle x^*, z \rangle] + [g^*(x^*) + g(x - z) - \langle x^*, x - z \rangle] = 0.$$

Ambas as parcelas entre colchetes são não-negativas por causa da desigualdade (1.9). Então, são nulas. Segue que $x^* \in \partial f(z)$ e $x^* \in \partial g(x - z)$. Portanto, vale (ii).

Provemos que (ii) implica (i). Seja x^* na interseção dos conjuntos $\partial f(z)$ e $\partial g(x - z)$. Usando (1.18) obtemos

$$f^*(x^*) + f(z) = \langle x^*, z \rangle \text{ e } g^*(x^*) + g(x - z) = \langle x^*, x - z \rangle.$$

Após somar estas duas equações, segue da Proposição 1.6(i) e da desigualdade (1.9) que

$$f(z) + g(x - z) \leq (f \square g)(x). \quad (2.6)$$

A desigualdade contrária segue trivialmente da definição de convolução infimal. Então, vale a igualdade. Portanto, a convolução de f e g é exata em x .

Provemos que (i) implica (iii). Seja $x^* \in \partial(f \square g)(x)$. Então, (1.18) e a Proposição 1.6(i) nos permitem escrever

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

Ou ainda, usando a hipótese,

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

Depois de reorganizar as parcelas e usar a desigualdade (1.9) constatamos (como na demonstração de (i) $\Rightarrow$ (ii)) que x^* está na interseção dos conjuntos $\partial f(z)$ e $\partial g(x - z)$. Sendo assim,

$$x^* \in \bigcup_{y \in \mathcal{X}} [\partial f(y) \cap \partial g(x - y)].$$

Se o conjunto do lado direito é vazio, então a inclusão contrária vale trivialmente. Senão, existe $z \in \mathcal{X}$ tal que a interseção de $\partial f(z)$ e

$\partial g(x - z)$ é não vazia. Seja x^* nesta interseção. Por (1.18) obtemos as equações

$$f^*(x^*) + f(z) = \langle x^*, z \rangle \text{ e } g^*(x^*) + g(x - z) = \langle x^*, x - z \rangle,$$

cujas soma fornece

$$f^*(x^*) + g^*(x^*) + f(z) + g(x - z) = \langle x^*, x \rangle. \quad (2.7)$$

O fato de a interseção de $\partial f(z)$ e $\partial g(x - z)$ ser não vazia também garante (por (ii) $\Rightarrow$ (i)) que a convolução de f e g é exata em x . Levando isto em consideração na expressão anterior, juntamente com a Proposição 1.6(i), constatamos que

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

Logo, (1.18) nos permite concluir que

$$x^* \in \partial(f \square g)(x).$$

Assim, temos provado as duas inclusões que garantem a igualdade em (iii).

Por fim, resta provar que (iii) implica (i). Por hipótese, o conjunto $\partial(f \square g)(x)$ é não vazio e vale a expressão de (iii). Então, existe $z \in \mathcal{X}$ tal que a interseção de $\partial f(z)$ e $\partial g(x - z)$ é não-vazia. Segue da implicação de (ii) em (i) que a convolução de f e g é exata em x .

□

Conforme mostra o exemplo a seguir, a condição de qualificação exigida na proposição anterior é realmente necessária.

Exemplo 2.2 *Consideremos as funções $f(x) = e^x$ e $g(x) = e^{-x}$ definidas em $\mathbb{R}$. Seja $x \in \mathbb{R}$. A convolução infimal de f e g em x é a função identicamente nula. Sendo assim, tem-se*

$$\partial(f \square g)(x) = \{0\}, \quad \forall x \in \mathbb{R}.$$

Por outro lado, $\partial f(y) = \{e^y\}$ e $\partial g(x - y) = \{-e^{-(x-y)}\}$, de modo que

$$\partial f(y) \cap \partial g(x - y) = \emptyset, \quad \forall y \in \mathcal{X}.$$

Portanto, a fórmula dada na Proposição 2.1(iii) não se aplica neste caso. De fato, pode-se verificar que a convolução das funções f e g é não-exata em qualquer ponto.

O Exemplo 2.2 justifica a necessidade da fórmula apresentada no Teorema 3.1.

3 CÁLCULO SUBDIFERENCIAL COM ε -SUBDIFERENCIAIS

Depois que Brøndsted e Rockafellar apresentaram, em 1965 [3, Seção 3], o conceito de ε -subdiferencial, novas regras de cálculo subdiferencial foram desenvolvidas. Por exemplo, aquelas apresentadas nas duas seções seguintes: uma para o subdiferencial da convolução infimal; outra para o subdiferencial da função marginal (ver definição na Seção 3.2). As duas fórmulas que acabamos de mencionar são demonstradas em [12]. Neste mesmo artigo são estabelecidas outras duas fórmulas: uma para o subdiferencial da soma de duas funções convexas; outra para o subdiferencial da pré-composição de uma função convexa com um operador linear contínuo. Estas duas últimas fórmulas também estão contempladas nesta dissertação, mas são apresentadas somente no Capítulo 4. Elas são mais facilmente demonstradas quando usamos o resultado do Teorema 4.1.

3.1 SUBDIFERENCIAL DA CONVOLUÇÃO INFIMAL

Conforme mostra o Exemplo 2.2, existem casos em que a fórmula para o subdiferencial da convolução, enunciada na Proposição 2.1, não se aplica. Além disso, verificar se a convolução é exata pode ser inviável em certos casos. Isto justifica o interesse em uma fórmula que não depende de condições de qualificação. Tal fórmula é apresentada no Teorema 3.1 em função dos ε -subdiferenciais das funções envolvidas e requer apenas a existência de um minorante afim comum as duas funções em questão. O Teorema 3.1 foi demonstrado por Hiriart-Urruty e Phelps em [12]. Aliás, todo o conteúdo desta seção é desenvolvido a partir do Capítulo 1 de [12].

Teorema 3.1 *Sejam $f, g : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ funções convexas, próprias e ambas limitadas inferiormente por uma mesma função afim. Sejam $\varepsilon > 0$ e x na interseção dos domínios de f e g . Então,*

$$(i) \quad \partial(f \square g)(x) \subset \bigcup_{y \in \mathcal{X}} [\partial_\varepsilon f(y) \cap \partial_\varepsilon g(x - y)] \subset \partial_{2\varepsilon}(f \square g)(x);$$

$$(ii) \quad \partial(f \square g)(x) = \bigcap_{\varepsilon > 0} \bigcup_{y \in \mathcal{X}} [\partial_\varepsilon f(y) \cap \partial_\varepsilon g(x - y)].$$

Demonstração: É fácil constatar que (i) implica em (ii). De fato, basta tomar a intersecção em (i) sobre todos os epsilons positivos e usar (1.21). Deste modo, basta provar as duas inclusões em (i) para demonstrar o teorema.

Seja $x^* \in \partial(f \square g)(x)$. Usando (1.18) e a Proposição 1.6(i) obtemos

$$f^*(x) + g^*(x) + (f \square g)(x) = \langle x^*, x \rangle \quad (3.1)$$

Seja $\varepsilon > 0$. Por hipótese, f e g são limitadas inferiormente por uma função afim. Segue da Proposição 1.5 que $f \square g$ é própria. Assim, existe y em $\mathcal{X}$ tal que

$$f(y) + g(x - y) < (f \square g)(x) + \varepsilon.$$

Usando este fato em (3.1) resulta

$$f^*(x^*) + g^*(x^*) + f(y) + g(x - y) < \langle x^*, x \rangle + \varepsilon. \quad (3.2)$$

Portanto, (3.2) vale para todo $y \in \mathcal{X}_\varepsilon(x)$, onde

$$\mathcal{X}_\varepsilon(x) := \{y \in \mathcal{X}; f(y) + g(x - y) < (f \square g)(x) + \varepsilon\}. \quad (3.3)$$

A desigualdade (1.9) fornece

$$f^*(x^*) \geq \langle x^*, u \rangle - f(u), \quad \forall u \in \mathcal{X}, \quad \text{e} \quad g^*(x^*) \geq \langle x^*, v \rangle - g(v), \quad \forall v \in \mathcal{X}.$$

Depois de levar em consideração estas duas desigualdades na expressão (3.2) obtemos

$$f(y) - f(u) + g(x - y) - g(v) + \langle x^*, (u + v) - x \rangle < \varepsilon. \quad (3.4)$$

Esta desigualdade vale para quaisquer u e v em $\mathcal{X}$ e para todo y em $\mathcal{X}_\varepsilon(x)$. A partir de (3.4), se $v = x - y$, concluímos que $x^* \in \partial_\varepsilon f(y)$. Por outro lado, se $u = y$, concluímos que $x^* \in \partial_\varepsilon g(x - y)$. Portanto,

$$x^* \in \partial_\varepsilon f(y) \cap \partial_\varepsilon g(x - y) \subseteq \bigcup_{y \in \mathcal{X}} [\partial_\varepsilon f(y) \cap \partial_\varepsilon g(x - y)].$$

Para provar a segunda inclusão em (i), suponhamos que existe y em $\mathcal{X}$ tal que x^* está na intersecção dos conjuntos $\partial_\varepsilon f(y)$ e $\partial_\varepsilon g(x - y)$. Segue de (1.22) que

$$f^*(x^*) + f(y) \leq \langle x^*, y \rangle + \varepsilon \quad \text{e} \quad g^*(x^*) + g(x - y) \leq \langle x^*, x - y \rangle + \varepsilon.$$

Destas inequações e da Proposição 1.6(i) obtemos

$$(f \square g)^*(x^*) + f(y) + g(x - y) \leq \langle x^*, x \rangle + 2\varepsilon.$$

A partir da definição de convolução, ainda podemos minorar o lado esquerdo esta última inequação para obtermos

$$(f \square g)^*(x^*) + (f \square g)(x) \leq \langle x^*, x \rangle + 2\varepsilon.$$

Por fim, usando (1.22) novamente, $x^* \in \partial_{2\varepsilon}(f \square g)(x)$. Isto prova a segunda inclusão e completa a demonstração do teorema.

□

A partir da inequação (3.2) e da conclusão que a sucede, podemos perceber que as uniões em (i) e (ii) não precisam ser tomadas sobre todos os $y \in \mathcal{X}$. Em vez disso, basta tomar a interseção sobre os y que estão no conjunto $\mathcal{X}_\varepsilon(x)$. Além disso, por (1.21), a interseção em (ii) pode ser tomada apenas sobre os epsilons pertencentes ao intervalo $(0, \bar{\varepsilon}]$, onde $\bar{\varepsilon}$ pode ser tão pequeno quanto se queira. Estas considerações nos permitem rescrever (i) e (ii) como segue:

$$\partial(f \square g)(x) \subset \bigcup_{y \in \mathcal{X}_\varepsilon(x)} [\partial_\varepsilon f(y) \cap \partial_\varepsilon g(x - y)] \subset \partial_{2\varepsilon}(f \square g)(x); \quad (3.5)$$

$$\partial(f \square g)(x) = \bigcap_{0 < \varepsilon \leq \bar{\varepsilon}} \bigcup_{y \in \mathcal{X}_\varepsilon(x)} [\partial_\varepsilon f(y) \cap \partial_\varepsilon g(x - y)]. \quad (3.6)$$

A fórmula (3.5) pode ser simplificada ainda mais. De fato, para cada $\varepsilon \in (0, \bar{\varepsilon}]$, seja $y_\varepsilon \in \mathcal{X}_\varepsilon(x)$. Segue de (3.6) que

$$\partial(f \square g)(x) \subset \partial_\varepsilon f(y_\varepsilon) \cap \partial_\varepsilon g(x - y_\varepsilon) \subset \partial_{2\varepsilon}(f \square g)(x). \quad (3.7)$$

Basta tomar a interseção nesta última expressão sobre todos os $\varepsilon \in (0, \bar{\varepsilon}]$ e usar (1.21) para obter

$$\partial(f \square g)(x) = \bigcap_{0 < \varepsilon \leq \bar{\varepsilon}} \partial_\varepsilon f(y_\varepsilon) \cap \partial_\varepsilon g(x - y_\varepsilon). \quad (3.8)$$

Se a convolução de f e g é exata, então existe y na interseção dos conjuntos $\mathcal{X}_\varepsilon(x)$, para todo $\varepsilon > 0$. Neste caso, a fórmula (3.8) fica mais simples (sem a dependência de y em ε) e poderia substituir aquela dada na Proposição 2.1. O mesmo raciocínio pode ser aplicado à fórmula (ii) do corolário a seguir.

Em (3.8) os ε -subdiferenciais de f e g dependem do mesmo ε . O resultado a seguir mostra que é possível reescrever aquela expressão de tal forma que os ε -subdiferenciais de f e g dependam de parâmetros distintos, ainda que dependentes um do outro.

Corolário 3.1 *Sob as mesmas hipóteses do Teorema 3.1 valem as fórmulas:*

$$(i) \quad \partial(f \square g)(x) \subset \bigcup_{y \in \mathcal{X}_\varepsilon(x)} \bigcup_{\substack{\varepsilon_1, \varepsilon_2 \geq 0 \\ \varepsilon_1 + \varepsilon_2 = \varepsilon}} [\partial_{\varepsilon_1}(y) \cap \partial_{\varepsilon_2}g(x-y)] \subset \partial_\varepsilon(f \square g)(x);$$

$$(ii) \quad \partial(f \square g)(x) = \bigcap_{0 < \varepsilon \leq \bar{\varepsilon}} \bigcup_{\substack{y \in \mathcal{X}_\varepsilon(x) \\ \varepsilon_1, \varepsilon_2 \geq 0 \\ \varepsilon_1 + \varepsilon_2 = \varepsilon}} [\partial_{\varepsilon_1}(y) \cap \partial_{\varepsilon_2}g(x-y)],$$

onde $\bar{\varepsilon}$ pode ser tão pequeno quanto se queira e $\mathcal{X}_\varepsilon(x)$ é definido em (3.3).

Demonstração: Provemos (i). Sejam $x^* \in \partial(f \square g)(x)$ e $\varepsilon > 0$. Basta proceder como na demonstração do Teorema 3.1 para obter-se

$$f(y) - f(u) + g(x - y) - g(v) + \langle x^*, (u + v) - x \rangle < \varepsilon. \quad (3.9)$$

Esta desigualdade vale para quaisquer u e v em $\mathcal{X}$ e para todo y em $\mathcal{X}_\varepsilon(x)$. Fixe $y \in \mathcal{X}_\varepsilon(x)$. Seja

$$\varepsilon_1 := \sup_{u \in \mathcal{X}} \langle x^*, u - y \rangle - f(u) + f(y). \quad (3.10)$$

Segue que $\varepsilon_1 \leq \varepsilon$ (basta tomar $v = x - y$ em (3.9) para constatar isto). De (3.10) obtemos

$$\varepsilon_1 \geq \langle x^*, u - y \rangle - f(u) + f(y), \quad \forall u \in \mathcal{X}. \quad (3.11)$$

Tomando $u = y$ nesta expressão conclui-se que $\varepsilon_1 \geq 0$. Além disso, (3.11) permite concluir que

$$x^* \in \partial_{\varepsilon_1} f(y). \quad (3.12)$$

Por outro lado, somando e subtraindo y no argumento de x^* em (3.9) e reorganizando as parcelas obtemos

$$[\langle x^*, u - y \rangle - f(u) + f(y)] + [g(x - y) - g(v) + \langle x^*, v - (x - y) \rangle] < \varepsilon.$$

Esta desigualdade vale para todo u e v em $\mathcal{X}$. Após tomar o supremo em u e usar a definição de ε_1 resulta

$$\langle x^*, v - (x - y) \rangle \leq g(v) - g(x - y) + \varepsilon - \varepsilon_1, \quad \forall v \in \mathcal{X}.$$

Segue da definição de subdiferencial que

$$x^* \in \partial_{\varepsilon_2} g(x - y), \quad (3.13)$$

onde $\varepsilon_2 := \varepsilon - \varepsilon_1$. Portanto, de (3.12) e (3.13),

$$x^* \in \partial_{\varepsilon_1} f(y) \cap \partial_{\varepsilon_2} g(x - y), \quad (3.14)$$

onde $\varepsilon = \varepsilon_1 + \varepsilon_2$ e $y \in \mathcal{X}_\varepsilon(x)$. Isso prova a primeira inclusão.

Provemos a segunda inclusão. Sejam $y \in \mathcal{X}_\varepsilon(x)$ e $\varepsilon_1, \varepsilon_2$ não-negativos tais que

$$x^* \in \partial_{\varepsilon_1} f(y) \cap \partial_{\varepsilon_2} g(x - y), \quad \varepsilon = \varepsilon_1 + \varepsilon_2. \quad (3.15)$$

De (1.22) obtemos as seguintes inequações

$$\begin{aligned} f^*(x^*) + f(y) &\leq \langle x^*, y \rangle + \varepsilon_1, \\ g^*(x^*) + g(x - y) &\leq \langle x^*, x - y \rangle + \varepsilon_2, \end{aligned}$$

cuja soma fornece, depois de usar a Proposição 1.6(i),

$$(f \square g)^*(x^*) + f(y) + g(x - y) \leq \langle x^*, x \rangle + \varepsilon.$$

Da definição de convolução infimal, temos

$$(f \square g)^*(x^*) + (f \square g)(x) \leq \langle x^*, x \rangle + \varepsilon.$$

Logo, de (1.22), $x^* \in \partial_\varepsilon (f \square g)(x)$. Com isto completamos a demonstração da fórmula (i).

Para provar (ii) basta tomar a interseção sobre todos os ε pertencentes ao intervalo $(0, \bar{\varepsilon}]$ em (i) e usar (1.21).

□

Um exemplo de aplicação dos resultados apresentados anteriormente é dado na Proposição 3.1. Ele mostra que para uma escolha apropriada da função g é possível trocar o problema de minimizar uma função f pelo problema de minimizar $f \square g$. Isto pode ser muito conveniente visto que para um escolha adequada de g a função $f \square g$ é

uma regularização de f , ou seja, o segundo problema pode ser mais facilmente tratado quando comparado com o primeiro.

Proposição 3.1 *Sejam $f, g : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ funções convexas, próprias e s.c.i.. Suponha que g satisfaz*

- (i) $g(0) = \inf_{x \in \mathcal{X}} g(x)$;
- (ii) se $g(x_n) \rightarrow g(0)$, então $x_n \rightarrow 0$.

Nestas condições, $\bar{x}$ é minimizador de f se, e somente se, $\bar{x}$ é minimizador de $f \square g$ em $\mathcal{X}$.

Demonstração: Seja $\bar{x}$ um minimizador de f em $\mathcal{X}$. Após usar da definição de convolução infimal e as hipóteses, obtemos

$$\begin{aligned} (f \square g)(\bar{x}) &\leq f(\bar{x}) + g(0) \\ &= \min_{y \in \mathcal{X}} f(y) + \min_{x \in \mathcal{X}} g(x) \\ &= \min_{y \in \mathcal{X}} f(y) + \min_{y \in \mathcal{X}} g(x - y), \\ &\leq (f \square g)(x), \quad \forall x \in \mathcal{X}. \end{aligned}$$

Para provar a implicação contrária, suponha-se que $\bar{x}$ seja um minimizador de $f \square g$ em $\mathcal{X}$. Então, $0 \in \partial(f \square g)(\bar{x})$. Da fórmula (3.8), com $\bar{\varepsilon} = 1$, temos

$$0 \in \bigcap_{0 < \varepsilon \leq 1} [\partial_\varepsilon f(y_\varepsilon) \cap \partial_\varepsilon g(\bar{x} - y_\varepsilon)].$$

Em particular, para $\varepsilon = 1/n$,

$$0 \in \partial_{\frac{1}{n}} f(y_n) \cap \partial_{\frac{1}{n}} g(\bar{x} - y_n), \quad \forall n \in \mathbb{N}.$$

Isto significa que y_n e $\bar{x} - y_n$ são $1/n$ -minimizadores de f e g , respectivamente, donde obtemos

$$f(y_n) \leq \inf_{x \in \mathcal{X}} f(x) + \frac{1}{n}, \quad \forall n \in \mathbb{N}, \quad (3.16)$$

$$g(\bar{x} - y_n) \leq \inf_{x \in \mathcal{X}} g(x) + \frac{1}{n}, \quad \forall n \in \mathbb{N}. \quad (3.17)$$

Da hipótese (i) e de (3.17) resulta

$$g(0) \leq g(\bar{x} - y_n) \leq g(0) + \frac{1}{n},$$

donde concluímos que $g(\bar{x} - y_n) \rightarrow g(0)$, quando $n \rightarrow \infty$. Segue-se da hipótese (ii) que $y_n \rightarrow \bar{x}$. Por outro lado, de (3.16) temos que

$$f(y_n) \leq f(x) + \frac{1}{n}, \quad \forall x \in \mathcal{X}.$$

Por fim, basta tomar o limite inferior com n tendendo ao infinito em ambos os lados desta última inequação e usar a semicontinuidade de f para constatar que $\bar{x}$ é um minimizador de f em $\mathcal{X}$.

□

Um exemplo bem conhecido de função que possui as mesmas características de g da proposição anterior é

$$g = \frac{1}{2\gamma} \|\cdot\|^2, \quad \gamma > 0.$$

Se $\mathcal{H}$ é um espaço de Hilbert e $f : \mathcal{H} \rightarrow \overline{\mathbb{R}}$ é uma função convexa, própria e s.c.i., então a função $f_\gamma : \mathcal{H} \rightarrow \overline{\mathbb{R}}$, definida por

$$f_\gamma(x) = (f \square g)(x) = \inf_{y \in \mathcal{H}} f(y) + \frac{1}{2\gamma} \|x - y\|^2, \quad (3.18)$$

é convexa, própria, contínua e sua imagem é um subconjunto próprio de $\mathbb{R}$. Além disso, a convolução infimal em (3.18) é exata e, para cada $x \in \mathcal{H}$, o ínfimo é realizado em um único $y \in \mathcal{H}$ [4, Proposição 12.15]. Esta função é conhecida como *regularização de Moreau* e desempenha um papel importante no estudo de problemas relacionados a recuperação de sinais [5].

3.2 SUBDIFERENCIAL DA FUNÇÃO MARGINAL

Seja A um operador linear contínuo de $\mathcal{X}$ em $\mathcal{Y}$, onde $\mathcal{Y}$ é um espaço de Banach. Seja $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ uma função convexa e própria. A função marginal de f sobre A é definida por

$$Af : \mathcal{Y} \rightarrow \overline{\mathbb{R}}, \quad Af(y) = \inf \{f(x); Ax = y\}.$$

Algumas propriedades relevantes a respeito de Af estão enunciadas na proposição a seguir, cuja demonstração optamos por apresentar no apêndice visto que a mesma não tem muito a acrescentar à esta seção.

Proposição 3.2 *Seja A um operador linear contínuo de $\mathcal{X}$ em $\mathcal{Y}$ e*

A^* o operador adjunto de A . Seja $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ uma função convexa e própria. As seguintes afirmativas procedem:

- (i) O domínio de Af coincide com a imagem do domínio de f pelo operador A ;
- (ii) Af é convexa;
- (iii) Se existe uma constante c e um funcional $x_0^* \in \{A^*y^*; y^* \in \mathcal{Y}^*\}$ de modo que $\langle x_0^*, x \rangle + c \leq f(x)$, $\forall x \in \mathcal{X}$, então Af é própria.

Convém observar que a condição exigida em (iii) para garantir que Af é própria é realmente necessária. Consideremos, por exemplo, $\mathcal{X} = \mathbb{R}^2$. Seja $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ um operador linear cujo núcleo coincide com o eixo das abscissas. Seja $f(x, y) = x$. Neste caso,

$$Af(0) = \inf \{f(x, y); A(x, y) = 0\} = -\infty,$$

ou seja, Af não é própria.

O próximo teorema fornece uma fórmula para o subdiferencial da função marginal.

Teorema 3.2 [12, Teorema 4.1] *Seja A um operador linear contínuo de $\mathcal{X}$ em $\mathcal{Y}$. Seja $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ uma função convexa e própria. Suponhamos que existe uma constante c e um funcional $x_0^* \in \{A^*y^*; y^* \in \mathcal{Y}^*\}$ de modo que $\langle x_0^*, x \rangle + c \leq f(x)$, $\forall x \in \mathcal{X}$. Sejam $\varepsilon > 0$ e y no domínio de Af . Então,*

$$\partial(Af)(y) = \bigcap_{0 < \varepsilon \leq \bar{\varepsilon}} \bigcup_{\{x; Ax=y\}} \{y^* \in \mathcal{Y}^*; A^*y^* \in \partial_\varepsilon f(x)\}. \quad (3.19)$$

Demonstração: Seja $y^* \in \partial(Af)(y)$. Por (1.18) temos

$$(Af)^*(y^*) + Af(y) = \langle y^*, y \rangle. \quad (3.20)$$

Seja $\varepsilon \in (0, \bar{\varepsilon}]$. Da Proposição 3.2(i), existe x no domínio de f tal que $Ax = y$ e

$$f(x) < Af(y) + \varepsilon. \quad (3.21)$$

A soma de (3.20) e (3.21) fornece

$$(Af)^*(y^*) + f(x) < \langle y^*, Ax \rangle + \varepsilon. \quad (3.22)$$

Afirmamos que $f^* \circ A^* \leq (Af)^*$. De fato, a partir da definição de função marginal podemos constatar que $(Af \circ A)(u) \leq f(u)$, para qualquer u

em $\mathcal{X}$. Segue da Proposição 1.3(iii) que

$$f^*(u^*) \leq (Af \circ A)^*(u^*), \quad \forall u^* \in \mathcal{X}^*.$$

Em particular, esta última desigualdade vale em $u^* = A^*y^*$. Então,

$$\begin{aligned} f^*(A^*y^*) &\leq (Af \circ A)^*(A^*y^*) = \sup_{z \in \mathcal{X}} \langle A^*y^*, z \rangle - (Af \circ A)(z) \\ &= \sup_{z \in \mathcal{X}} \langle y^*, Az \rangle - Af(Az) \\ &\leq \sup_{v \in \mathcal{Y}} \langle y^*, v \rangle - Af(v) = (Af)^*(y^*). \end{aligned}$$

Assim, concluímos a afirmação. Ao levar esse fato em consideração na inequação (3.22) obtemos

$$f^*(A^*y^*) + f(x) \leq \langle y^*, Ax \rangle + \varepsilon = \langle A^*y^*, x \rangle + \varepsilon.$$

Segue de (1.22) que $A^*y^* \in \partial_\varepsilon f(x)$, para todo $\varepsilon \in (0, \bar{\varepsilon}]$ e algum $x \in \mathcal{X}$ tal que $Ax = y$. Portanto, o lado esquerdo de (3.19) está contido no lado direito.

Mostremos a inclusão contrária. Seja y^* contido no lado direito de (3.19). Então, para todo $\varepsilon \in (0, \bar{\varepsilon}]$ e para algum $x \in \mathcal{X}$ tal que $Ax = y$, temos $A^*y^* \in \partial_\varepsilon f(x)$. Assim, como $Af(y) \leq f(x)$, a definição de ε -subdiferencial e propriedades de operador auto-adjunto nos permitem escrever

$$\langle y^*, Au - y \rangle \leq f(u) - Af(y) + \varepsilon, \quad \forall u \in \mathcal{X}. \quad (3.23)$$

Seja v no domínio de Af . Então, da Proposição 3.2(i), existe $u \in \mathcal{X}$ tal que $Au = v$. Com esta mudança de variável segue de (3.23) que

$$\langle y^*, v - y \rangle \leq \inf \{f(u); Au = v\} - Af(y) + \varepsilon = Af(v) - Af(y) + \varepsilon,$$

para todo v no domínio de Af . Portanto, $y^* \in \partial_\varepsilon Af(y)$, para todo $\varepsilon \in (0, \bar{\varepsilon}]$. Assim, basta usar (1.21) para concluir a demonstração.

□

4 CÁLCULO SUBDIFERENCIAL COM LIMITES

Neste capítulo, vamos demonstrar outro resultado importante da teoria de subdiferenciais (Teorema 4.1). Uma série de resultados podem ser facilmente obtidos a partir dele. Para ilustrar o que acabamos de afirmar, vamos usar os resultados do Teorema 4.1 para demonstrar três resultados bem conhecidos em análise convexa: o primeiro diz respeito a fórmula para o subdiferencial da soma de duas funções convexas e próprias, bem como da pré-composição de uma função convexa e própria com um operador linear contínuo; o segundo trata da densidade do domínio do subdiferencial no domínio da função; por fim, o teorema de integração de Rockafellar.

A demonstração do Teorema 4.1 requer os resultados dos dois lemas que passamos a demonstrar. Tanto o teorema como os dois lemas estão enunciados em [22].

Lema 4.1 *Sejam $f_1, f_2, f_3 : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ funções convexas, próprias e s.c.i.. Seja V um subespaço vetorial de $\mathcal{X}$. Suponha que a dimensão de V é finita e que $f_3(x) = +\infty$ para todo x no complementar de V . Seja S um subconjunto fechado e limitado de $\mathcal{X}$. Suponha que f_1, f_2 e f_3 são finitas em algum $\bar{x} \in S$. Para cada $(x_1, x_2, x_3) \in \mathcal{X}^3$ e $t > 0$ sejam*

$$\Delta(x_1, x_2, x_3) := \frac{1}{2}(\|x_1 - x_2\|^2 + \|x_1 - x_3\|^2 + \|x_2 - x_3\|^2);$$

$$w_t(x_1, x_2, x_3) := f_1(x_1) + f_2(x_2) + f_3(x_3) + t\Delta(x_1, x_2, x_3).$$

Nestas condições, valem as sentenças:

(i) *Para qualquer rede $(x_{1,t}, x_{2,t}, x_{3,t}) \subset S^3$ tal que*

$$w_t(x_{1,t}, x_{2,t}, x_{3,t}) - \inf_{(x_1, x_2, x_3) \in S^3} w_t(x_1, x_2, x_3) \xrightarrow[t \rightarrow \infty]{} 0,$$

tem-se

$$t\Delta(x_{1,t}, x_{2,t}, x_{3,t}) \xrightarrow[t \rightarrow \infty]{} 0; \quad (4.1)$$

(ii) $\inf_{(x_1, x_2, x_3) \in S^3} w_t(x_1, x_2, x_3) \xrightarrow[t \rightarrow \infty]{} \inf_{x \in S} [f_1(x) + f_2(x) + f_3(x)].$

Demonstração: Afirmamos que existe o ínfimo de w_t sobre S^3 . De fato, como f_1, f_2 e f_3 são convexas, próprias e s.c.i., existem funcionais x_1^*, x_2^* e x_3^* em $\mathcal{X}^*$, bem como uma constante c , tais que

$$\langle x_1^*, x_1 \rangle + \langle x_2^*, x_2 \rangle + \langle x_3^*, x_3 \rangle + c \leq f_1(x_1) + f_2(x_2) + f_3(x_3).$$

Como S é limitado, existe uma constante que minora o lado esquerdo e garante a existência do ínfimo.

Provemos (i). Para tanto, seja $(x_{1,t}, x_{2,t}, x_{3,t})$ uma rede como na hipótese. Para todo $t > 0$ temos

$$\begin{aligned} \inf_{(x_1, x_2, x_3) \in S^3} w_{\frac{t}{2}}(x_1, x_2, x_3) &\leq w_{\frac{t}{2}}(x_{1,t}, x_{2,t}, x_{3,t}) \\ &= f_1(x_{1,t}) + f_2(x_{2,t}) + f_3(x_{3,t}) \\ &\quad + t\Delta(x_{1,t}, x_{2,t}, x_{3,t}) - \frac{t}{2}\Delta(x_{1,t}, x_{2,t}, x_{3,t}) \\ &= w_t(x_{1,t}, x_{2,t}, x_{3,t}) - \frac{t}{2}\Delta(x_{1,t}, x_{2,t}, x_{3,t}). \end{aligned}$$

Como o ínfimo é finito,

$$0 \leq t\Delta(x_{1,t}, x_{2,t}, x_{3,t}) \leq 2 \left[w_t(x_{1,t}, x_{2,t}, x_{3,t}) - \inf_{(x_1, x_2, x_3) \in S^3} w_{\frac{t}{2}}(x_1, x_2, x_3) \right].$$

A partir daí, basta somar e subtrair o ínfimo de w_t no lado direito e usar a hipótese para constatar (4.1).

Para provar (ii) observamos que

$$\Delta_t(x, x, x) = f_1(x) + f_2(x) + f_3(x), \quad \forall x \in \mathcal{X}.$$

Assim,

$$\begin{aligned} \inf_{(x_1, x_2, x_3) \in S^3} w_t(x_1, x_2, x_3) &\leq \inf_{x \in S} w_t(x, x, x) \\ &= \inf_{x \in S} f_1(x) + f_2(x) + f_3(x). \end{aligned}$$

Portanto,

$$\lim_{t \rightarrow \infty} \inf_{(x_1, x_2, x_3) \in S^3} w_t(x_1, x_2, x_3) \leq \inf_{x \in S} f_1(x) + f_2(x) + f_3(x).$$

Resta provar a desigualdade contrária. Para todo n natural, existe $(x_{1,n}, x_{2,n}, x_{3,n})$ em S^3 tal que

$$w_n(x_{1,n}, x_{2,n}, x_{3,n}) < \inf_{(x_1, x_2, x_3) \in S^3} w_n(x_1, x_2, x_3) + \frac{1}{n}. \quad (4.2)$$

Resulta que

$$w_n(x_{1,n}, x_{2,n}, x_{3,n}) - \inf_{(x_1, x_2, x_3) \in S^3} w_n(x_1, x_2, x_3) \xrightarrow{n \rightarrow \infty} 0. \quad (4.3)$$

Segue de (i) que $n\Delta(x_{1,n}, x_{2,n}, x_{3,n}) \rightarrow 0$. Então,

$$\Delta(x_{1,n}, x_{2,n}, x_{3,n}) \xrightarrow{n \rightarrow \infty} 0. \quad (4.4)$$

A partir de (4.2) concluímos que $w_n(x_{1,n}, x_{2,n}, x_{3,n})$ é finito para todo n natural. Em particular, resulta que $f_3(x_{3,n})$ é finito para todo n natural. Então, a sequência $(x_{3,n})_{n \in \mathbb{N}}$ está contida no compacto $S \cap V$. Portanto, existe $(x_{3,n_k})_{k \in \mathbb{N}}$, subsequência de $(x_{3,n})_{n \in \mathbb{N}}$, e existe $\tilde{x} \in S$ de modo que (x_{3,n_k}) converge a $\tilde{x}$. Tendo em vista a definição de Δ , segue de (4.4) que as subsequências $(x_{1,n_k})_{n \in \mathbb{N}}$ e $(x_{2,n_k})_{n \in \mathbb{N}}$ também convergem a $\tilde{x}$. Por fim, como

$$\inf_{S^3} w_n - [\inf_{S^3} w_n - w_n(x_{1,n}, x_{2,n}, x_{3,n})] = w_n(x_{1,n}, x_{2,n}, x_{3,n}),$$

após usar (4.3) e a semicontinuidade das funções f_1 , f_2 e f_3 podemos escrever

$$\begin{aligned} \lim_{n \rightarrow \infty} \inf_{(x_1, x_2, x_3) \in S^3} w_n(x_1, x_2, x_3) &= \lim_{n \rightarrow \infty} w_n(x_{1,n}, x_{2,n}, x_{3,n}) \\ &\geq \liminf_{n \rightarrow \infty} f_1(x_{1,n}) + f_2(x_{2,n}) + f_3(x_{3,n}) \\ &= f_1(\tilde{x}) + f_2(\tilde{x}) + f_3(\tilde{x}) + n\Delta(\tilde{x}, \tilde{x}, \tilde{x}) \\ &\geq \inf_{x \in S} f_1(x) + f_2(x) + f_3(x). \end{aligned}$$

Isto completa a prova do lema. □

Usemos o lema anterior e o princípio variacional de Ekeland (Corolário 1.1) para demonstrar o lema a seguir.

Lema 4.2 *Sejam $f_i : \mathcal{X} \rightarrow \overline{\mathbb{R}}$, $i = 1, 2, 3$, funções convexas, próprias e s.c.i.. Seja V um subespaço vetorial de $\mathcal{X}$. Suponhamos que a dimensão de V é finita e que $f_3(x) = +\infty$ para todo x no complementar de V . Suponhamos ainda que existe $\bar{x} \in \mathcal{X}$ tal que*

$$f_1(\bar{x}) + f_2(\bar{x}) + f_3(\bar{x}) = \inf_{x \in \mathcal{X}} f_1(x) + f_2(x) + f_3(x) < +\infty. \quad (4.5)$$

Seja $\varepsilon > 0$. Então, existem x_i em $\mathcal{X}$, bem como x_i^ em $\mathcal{X}^*$, tais que $x_i^* \in \partial f_i(x_i)$ e vale cada uma das sentenças a seguir:*

$$(i) \quad \|x_i - \bar{x}\| \leq \varepsilon;$$

$$(ii) \quad |f_i(x_i) - f_i(\bar{x})| \leq \varepsilon;$$

$$(iii) \quad \|x_1^* + x_2^* + x_3^*\| \leq \varepsilon;$$

$$(iv) \quad \|x_i^*\| \max \{\|x_1 - x_2\|, \|x_1 - x_3\|, \|x_2 - x_3\|\} \leq \varepsilon.$$

Demonstração: Por meio do Lema 4.1 vamos buscar condições suficientes para aplicar o princípio variacional de Ekeland (Corolário 1.1). Com este propósito, definimos $S := B[\bar{x}, \varepsilon]$. Assim, segue do Lema 4.1, de (4.5) e do fato que $\Delta(\bar{x}, \bar{x}, \bar{x}) = 0$, que

$$\inf_{(x_1, x_2, x_3) \in S^3} w_t(x_1, x_2, x_3) \xrightarrow{t \rightarrow \infty} w_t(\bar{x}, \bar{x}, \bar{x}).$$

Então, para cada n natural existe t_n tal que

$$w_{t_n}(\bar{x}, \bar{x}, \bar{x}) < \inf_{(x_1, x_2, x_3) \in S^3} w_{t_n}(x_1, x_2, x_3) + \frac{\varepsilon^2}{n}. \quad (4.6)$$

Esta inequação nos diz que $(\bar{x}, \bar{x}, \bar{x})$ é um ε_n -minimizador ($\varepsilon_n = \varepsilon^2/n$) da função $w_{t_n}(x_1, x_2, x_3)$ em S^3 . Consideremos o espaço métrico S^3 munido da métrica induzida de $\mathcal{X}^3$:

$$d((x_1, x_2, x_3), (y_1, y_2, y_3)) = \|(x_1, x_2, x_3) - (y_1, y_2, y_3)\|,$$

Onde $\|\cdot\|$ denota a norma do espaço $\mathcal{X}^3$, definida por

$$\|(x_1, x_2, x_3)\| := \|x_1\| + \|x_2\| + \|x_3\|.$$

Nestas condições, o Corolário 1.1 garante que, fixado n natural, existe $(x_{1,n}, x_{2,n}, x_{3,n}) \in S^3$ de modo que

$$w_{t_n}(x_{1,n}, x_{2,n}, x_{3,n}) \leq w_{t_n}(\bar{x}, \bar{x}, \bar{x}), \quad (4.7)$$

com

$$\|(x_{1,n}, x_{2,n}, x_{3,n}) - (\bar{x}, \bar{x}, \bar{x})\| \leq \lambda_n, \quad (\lambda_n = \varepsilon/n). \quad (4.8)$$

Além disso, o corolário também garante que para todo $(x_1, x_2, x_3) \in S^3$, $(x_1, x_2, x_3) \neq (x_{1,n}, x_{2,n}, x_{3,n})$, vale a desigualdade

$$\begin{aligned} w_{t_n}(x_{1,n}, x_{2,n}, x_{3,n}) &< w_{t_n}(x_1, x_2, x_3) \\ &+ \frac{\varepsilon_n}{\lambda_n} \|(x_1, x_2, x_3) - (x_{1,n}, x_{2,n}, x_{3,n})\|. \end{aligned} \quad (4.9)$$

O resultado de (i) segue diretamente de (4.8) e da definição de $\|\cdot\|$.

Para provar (ii) afirmamos que

$$f_1(x_{1,n}) + f_2(x_{2,n}) + f_3(x_{3,n}) \xrightarrow{n \rightarrow \infty} f_1(\bar{x}) + f_2(\bar{x}) + f_3(\bar{x}). \quad (4.10)$$

De fato, de (4.6) e (4.7) obtemos

$$\begin{aligned} w_{t_n}(x_{1,n}, x_{2,n}, x_{3,n}) &\leq f(\bar{x}, \bar{x}, \bar{x}) \\ &\leq \inf_{(x_1, x_2, x_3) \in S^3} w_{t_n}(x_1, x_2, x_3) + \frac{\varepsilon^2}{n}, \end{aligned} \quad (4.11)$$

onde $f(x_1, x_2, x_3) := f_1(x_1) + f_2(x_2) + f_3(x_3)$. Após majorar o ínfimo da expressão anterior por $w_{t_n}(x_{1,n}, x_{2,n}, x_{3,n})$ constatamos que

$$w_{t_n}(x_{1,n}, x_{2,n}, x_{3,n}) \xrightarrow{n \rightarrow \infty} f(\bar{x}, \bar{x}, \bar{x}).$$

Por outro lado,

$$f(\bar{x}, \bar{x}, \bar{x}) \leq f(x_{1,n}, x_{2,n}, x_{3,n}) \leq w_{t_n}(x_{1,n}, x_{2,n}, x_{3,n}).$$

As duas últimas expressões combinadas confirmam a afirmação. Usamos a semicontinuidade de cada função f_i e (i) para concluir que

$$\liminf_{n \rightarrow \infty} f_i(x_{i,n}) = f_i(\bar{x}). \quad (4.12)$$

Por fim, (4.10) e (4.12) implicam (ver Proposição 4.4)

$$f_i(x_{i,n}) \xrightarrow{n \rightarrow \infty} f_i(\bar{x}).$$

Assim, basta tomar um natural suficientemente grande para concluir a prova de (ii).

Para provar (iii) consideremos a função convexa (ver Afirmação 4.1) $F_n : \mathcal{X}^3 \rightarrow \mathbb{R}$ definida por

$$\begin{aligned} F_n(x_1, x_2, x_3) &= f(x_1, x_2, x_3) + t_n \Delta(x_1, x_2, x_3) \\ &\quad + \varepsilon \| (x_1, x_2, x_3) - (x_{1,n}, x_{2,n}, x_{3,n}) \| + \delta_{S^3}(x_1, x_2, x_3). \end{aligned}$$

Segue diretamente de (4.9) que $(x_{1,n}, x_{2,n}, x_{3,n})$ é um minimizador de F_n . Portanto, como Δ é uma função contínua, o Teorema 2.1 nos

permite escrever

$$\begin{aligned} 0 \in \partial F_n(x_{1,n}, x_{2,n}, x_{3,n}) &= \partial f(x_{1,n}, x_{2,n}, x_{3,n}) + t_n \partial \Delta(x_{1,n}, x_{2,n}, x_{3,n}) \\ &\quad + \varepsilon \partial \|\cdot - (x_{1,n}, x_{2,n}, x_{3,n})\| (x_{1,n}, x_{2,n}, x_{3,n}) \\ &\quad + \partial \delta_{S^3}(x_{1,n}, x_{2,n}, x_{3,n}). \end{aligned} \quad (4.13)$$

Da Proposição 4.1(1), temos que

$$\partial f(x_{1,n}, x_{2,n}, x_{3,n}) = \partial f_1(x_{1,n}) \times \partial f_2(x_{2,n}) \times \partial f_3(x_{3,n}).$$

Também temos (Exemplo 1.1) que

$$\partial \|\cdot - (x_{1,n}, x_{2,n}, x_{3,n})\| (x_{1,n}, x_{2,n}, x_{3,n}) = \partial \|\cdot\| (0) = B_{(\mathcal{X}^*)^3}[0, 1].$$

Além disso, segue de (1.17) que $\partial \delta_{S^3}(x_{1,n}, x_{2,n}, x_{3,n}) = \{0\}$, visto que $(x_{1,n}, x_{2,n}, x_{3,n})$ está no interior de S^3 (pelo menos para n maior que 1). Deste modo, resulta de (4.13) que

$$\begin{aligned} 0 \in \partial f_1(x_{1,n}) \times \partial f_2(x_{2,n}) \times \partial f_3(x_{3,n}) \\ + t_n \partial \Delta(x_{1,n}, x_{2,n}, x_{3,n}) + \varepsilon B_{(\mathcal{X}^*)^3}[0, 1]. \end{aligned}$$

Portanto, existem $x_{i,n}^* \in \partial f_i(x_{i,n})$, tais que

$$-(x_{1,n}^*, x_{2,n}^*, x_{3,n}^*) \in t_n \partial \Delta(x_{1,n}, x_{2,n}, x_{3,n}) + \varepsilon B_{(\mathcal{X}^*)^3}[0, 1]. \quad (4.14)$$

Sejam $(u_{1,n}^*, u_{2,n}^*, u_{3,n}^*)$ em $\partial \Delta(x_{1,n}, x_{2,n}, x_{3,n})$ e $(v_{1,n}^*, v_{2,n}^*, v_{3,n}^*)$ em $B_{(\mathcal{X}^*)^3}[0, 1]$ tais que

$$-(x_{1,n}^*, x_{2,n}^*, x_{3,n}^*) = t_n(u_{1,n}^*, u_{2,n}^*, u_{3,n}^*) + \varepsilon(v_{1,n}^*, v_{2,n}^*, v_{3,n}^*). \quad (4.15)$$

Então,

$$\|x_{1,n}^* + x_{2,n}^* + x_{3,n}^*\| \leq t_n \|u_{1,n}^* + u_{2,n}^* + u_{3,n}^*\| + \varepsilon \|v_{1,n}^* + v_{2,n}^* + v_{3,n}^*\|.$$

Da Afirmação 4.1(3), a primeira parcela do lado direito é nula. Denotando a norma de $(\mathcal{X}^*)^3$ com a mesma notação usada para a norma de $\mathcal{X}^3$, obtemos

$$\begin{aligned} \|x_{1,n}^* + x_{2,n}^* + x_{3,n}^*\| &\leq \varepsilon [\|v_{1,n}^*\| + \|v_{2,n}^*\| + \|v_{3,n}^*\|] \\ &= \varepsilon \|(v_{1,n}^*, v_{2,n}^*, v_{3,n}^*)\| \leq \varepsilon. \end{aligned}$$

Portanto, a sentença (iii) está provada.

Finalmente, resta provar (iv). De (4.11) resulta

$$w_{t_n}(x_{1,n}, x_{2,n}, x_{3,n}) - \inf_{(x_1, x_2, x_3) \in S^3} w_{t_n}(x_1, x_2, x_3) \xrightarrow{n \rightarrow \infty} 0.$$

Segue do Lema 4.1 que

$$t_n \Delta(x_{1,n}, x_{2,n}, x_{3,n}) \xrightarrow{n \rightarrow \infty} 0 \quad (4.16)$$

Da Afirmação 4.1(3) temos que

$$\|u_{i,n}^*\| \leq \|x_{1,n} - x_{2,n}\| + \|x_{1,n} - x_{3,n}\| + \|x_{2,n} - x_{3,n}\|.$$

Esta última inequação combinada com (4.15) resulta

$$\begin{aligned} \|x_{i,n}^*\| &\leq t_n(\|x_{1,n} - x_{2,n}\| + \|x_{1,n} - x_{3,n}\| + \|x_{2,n} - x_{3,n}\|) \\ &\quad + \varepsilon \|v_{i,n}^*\|. \end{aligned} \quad (4.17)$$

Seja

$$M_n := \max \{\|x_{1,n} - x_{2,n}\|, \|x_{1,n} - x_{3,n}\|, \|x_{2,n} - x_{3,n}\|\}.$$

Após multiplicar a expressão (4.17) por M_n podemos majorar as três parcelas entre parênteses por $3(M_n)^2$. Assim, podemos escrever

$$\|x_{i,n}^*\| M_n \leq 3t_n M_n^2 + \varepsilon \|v_{i,n}^*\| M_n.$$

Visto que $M_n^2/2 \leq \Delta(x_{1,n}, x_{2,n}, x_{3,n})$, segue de (4.16) que o lado direito desta última inequação converge a zero. Portanto, existe um natural n suficientemente grande de modo que

$$\|x_{i,n}^*\| M_n \leq \varepsilon.$$

Com isto concluímos a demonstração do lema. □

Finalmente, temos condições de demonstrar o Teorema 4.1. Ele fornece cinco caracterizações alternativas para o subdiferencial da soma de duas funções convexas, próprias e s.c.i., bem como para o subdiferencial da pré-composição de uma função própria, convexa e s.c.i. com um operador linear contínuo. Além disso, o Teorema 4.1 fornece suporte para demonstrar, de maneira relativamente elementar, uma série de resultados que antes eram demonstrados com técnicas mais complexas.

Teorema 4.1 *Seja A um operador linear contínuo de $\mathcal{X}$ em $\mathcal{Y}$. Sejam $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ e $g : \mathcal{Y} \rightarrow \overline{\mathbb{R}}$ funções convexas, próprias e s.c.i.. Suponhamos que existe $\bar{x}$ na interseção dos domínios de f e $g \circ A$.*

Então, $x^ \in \partial(f + g \circ A)(\bar{x})$ se, e somente se, existem redes x_t e y_t em $\mathcal{X}$, bem como x_t^* e y_t^* em $\mathcal{X}^*$, tais que $x_t^* \in \partial f(x_t)$, $y_t^* \in \partial g(y_t)$,*

$$x_t^* + y_t^* \circ A \xrightarrow{*} x^*, \quad (4.18)$$

e vale uma das seguintes afirmativas equivalentes:

- (1) (a) $(x_t, f(x_t)) \rightarrow (\bar{x}, f(\bar{x}))$, $(y_t, g(y_t)) \rightarrow (A\bar{x}, g(A\bar{x}))$,
 (b) $\langle x_t^*, x_t - \bar{x} \rangle \rightarrow 0$, $\langle y_t^*, y_t - A\bar{x} \rangle \rightarrow 0$,
 (c) $(\|x_t^*\| + \|y_t^*\|)\|y_t - Ax_t\| \rightarrow 0$;
- (2) (a) $(x_t, f(x_t)) \rightarrow (\bar{x}, f(\bar{x}))$, $(y_t, g(y_t)) \rightarrow (A\bar{x}, g(A\bar{x}))$,
 (b) $\langle x_t^*, x_t - \bar{x} \rangle \rightarrow 0$, $\langle y_t^*, y_t - A\bar{x} \rangle \rightarrow 0$;
- (3) (a) $x_t \rightarrow \bar{x}$, $y_t \rightarrow A\bar{x}$,
 (b) $f(x_t) - \langle x_t^*, x_t - \bar{x} \rangle \rightarrow f(\bar{x})$, $g(y_t) - \langle y_t^*, y_t - A\bar{x} \rangle \rightarrow g(A\bar{x})$;
- (4) (a) $x_t \rightarrow \bar{x}$, $y_t \rightarrow A\bar{x}$,
 (b) $\langle x_t^*, x_t - \bar{x} \rangle \rightarrow 0$, $\langle y_t^*, y_t - A\bar{x} \rangle \rightarrow 0$;
- (5) (a) $x_t \rightarrow \bar{x}$, $y_t \rightarrow A\bar{x}$,
 (b) $\liminf \langle x_t^*, x_t - \bar{x} \rangle \geq 0$, $\liminf \langle y_t^*, y_t - A\bar{x} \rangle \geq 0$.

Demonstração: Identifiquemos com (0) a afirmativa:

$$x^* \in \partial(f + g \circ A)(\bar{x}).$$

Além disso, mantemos a notação (1), (2), ..., (5) para as afirmativas correspondentes com a sentença adicional dada em (4.18). Faremos a demonstração provando a equivalência das afirmativas (0) a (5).

Faremos em duas partes a demonstração de que (0) implica (1). Na primeira vamos considerar o caso particular em que $\mathcal{X}$ e $\mathcal{Y}$ são iguais e A é o operador identidade sobre $\mathcal{X}$. Isto nos permite usar o resultado do Lema 4.2. Na sequência usamos o resultado do caso particular para demonstrar o caso geral. A demonstração das demais implicações são mais simples.

(0) $\implies$ (1). Inicialmente consideremos o caso em que $\mathcal{X}$ e $\mathcal{Y}$ são iguais e A é o operador identidade em $\mathcal{X}$. Então, por hipótese,

$x^* \in \partial(f+g)(\bar{x})$. Seja $\mathcal{F}$ o conjunto de todos os subespaços de dimensão finita de $\mathcal{X}$. Consideremos o conjunto $\Lambda = \mathbb{N} \times \mathcal{F}$ com a relação de ordem: $(n_1, V_1) \leq (n_2, V_2)$ se, e somente se, $n_1 \leq n_2$ e $V_1 \subset V_2$. Nestas condições Λ é um conjunto dirigido¹. De fato, sejam (n_1, V_1) e (n_2, V_2) elementos de Λ . Escolha $n = n_1 + n_2$ e V o menor subespaço de $\mathcal{X}$ contendo V_1 e V_2 . Então, $(n_1, V_1) \leq (n, V)$ e $(n_2, V_2) \leq (n, V)$. Sejam n um natural qualquer e V um subespaço de dimensão finita contendo $\bar{x}$. Visto que $x^* \in \partial(f+g)(\bar{x})$, segue da definição de subdiferencial que

$$\langle x^*, y - \bar{x} \rangle + f(\bar{x}) + g(\bar{x}) \leq f(y) + g(y), \quad \forall y \in \mathcal{X}. \quad (4.19)$$

Após somar $\delta_V(\bar{x})$ e $\delta_V(y)$ no lado esquerdo e direito, respectivamente, de (4.19), constata-se que $\bar{x}$ é um minimizador da função $f - x^* + g + \delta_V$. Seja $t := (n, V)$. Então, pelo Lema 4.2 (com $\varepsilon = 1/n$, $f_1 = f - x^*$, $f_2 = g$ e $f_3 = \delta_V$) existem x_t, y_t e z_t em $\mathcal{X}$, bem como x_t^*, y_t^* e z_t^* em $\mathcal{X}^*$, onde

$$x_t^* \in \partial f(x_t), \quad y_t^* \in \partial g(y_t) \text{ e } z_t^* \in \partial \delta_V(z_t),$$

tais que

- (i) $\max \{\|x_t - \bar{x}\|, \|y_t - \bar{x}\|, \|z_t - \bar{x}\|\} \leq \frac{1}{n}$;
- (ii) $|(f - x^*)(x_t) - (f - x^*)(\bar{x})| \leq \frac{1}{n}$, $|g(y_t) - g(\bar{x})| \leq \frac{1}{n}$;
- (iii) $\|x_t^* - x^* + y_t^* + z_t^*\| \leq \frac{1}{n}$;
- (iv) $M_t(\|x_t^* - x^*\| + \|y_t^*\| + \|z_t^*\|) \leq \frac{1}{n}$, onde $M_t := \max \{\|x_t - y_t\|, \|x_t - z_t\|, \|y_t - z_t\|\}$.

Convém observar que a existência de x_t^* no subdiferencial de f é garantida de forma implícita pelo Lema 4.2. De fato, a princípio, o lema fornece a_t^* no subdiferencial de $f - x^*$. Contudo, do Teorema 2.1,

$$\partial(f - x^*)(x_t) = \partial f(x_t) - x^*.$$

Portanto, existe tal x_t^* , o qual satisfaz

$$a_t^* = x_t^* - x^*.$$

A partir de agora, usamos os resultados de (i) a (iv) para provar (4.18) e (1)(a) a (1)(c). Começamos provando (4.18).

¹Um conjunto Λ , munido de uma relação de ordem parcial “ $\leq$ ”, é dito dirigido se para quaisquer dois elementos t_1 e t_2 de Λ , existe t em Λ tal que $t_1 \leq t$ e $t_2 \leq t$.

Seja $x \in \mathcal{X}$. Por (iii), existe b_t^* na bola unitária de $\mathcal{X}^*$ tal que

$$x_t^* + y_t^* - x^* + z_t^* = \frac{1}{n} b^*. \quad (4.20)$$

Segue de (4.20) que

$$\langle x_t^* + y_t^*, x \rangle - \langle x^*, x \rangle = \langle z_t^*, -x \rangle + \frac{1}{n} \langle b_t^*, x \rangle. \quad (4.21)$$

Sejam $\varepsilon > 0$ e $t_0 = (n_0, V_0)$ um índice de Λ tal que $\|x\| < n_0\varepsilon$ e $x \in V_0$. Seja $t = (n, V)$ um índice maior que t_0 (isto é, n é maior que n_0 e V contém V_0). Afirmamos que nestas condições a primeira parcela do lado direito de (4.21) pode ser majorada por zero. De fato, como $z_t^* \in \partial\delta_V(z_t)$, segue de (1.17) que

$$\langle z_t^*, y - z_t \rangle \leq 0, \quad \forall y \in V. \quad (4.22)$$

Assim, basta tomar $y = z_t - x$ para obter a desigualdade desejada. Depois disso, usamos propriedades da norma e o fato que $\|b_t^*\| \leq 1$ para obtermos

$$|\langle x_t^* + y_t^*, x \rangle - \langle x^*, x \rangle| \leq \frac{1}{n} \|x\|.$$

Por fim, a escolha de t_0 nos permite concluir que

$$|\langle x_t^* + y_t^*, x \rangle - \langle x^*, x \rangle| < \varepsilon.$$

Portanto, $\langle x_t^* + y_t^*, x \rangle \rightarrow \langle x^*, x \rangle$. Como x é arbitrário, constatamos que $x_t^* + y_t^*$ converge fracamente à x^* .

Com o propósito de provar (1)(a), fazemos algumas manipulações algébricas a partir de (ii) até obtermos

$$|f(x_t) - f(\bar{x})| - |\langle x^*, x_t - \bar{x} \rangle| \leq \frac{1}{n}.$$

Assim, propriedades da norma e o resultado de (i) fornecem

$$|f(x_t) - f(\bar{x})| \leq \frac{1}{n} + \frac{1}{n} \|x^*\|. \quad (4.23)$$

Sejam $\varepsilon > 0$ e $t_0 = (n_0, V_0)$ um índice de Λ tal que $1 + \|x^*\| < n_0\varepsilon$. Então, para t maior que t_0 , segue de (i) e (4.23) que

$$\max \{ \|x_t - \bar{x}\|, |f(x_t) - f(\bar{x})| \} < \varepsilon.$$

Para essa mesma escolha de t_0 , segue de (i) e (ii) que

$$\max \{ \|y_t - \bar{x}\|, |g(y_t) - g(\bar{x})| \} < \varepsilon.$$

Concluimos destas duas últimas desigualdades a demonstração do item (1)(a).

Provemos (1)(b). Para tanto, afirmamos que

$$\langle x_t^* + y_t^*, x_t - \bar{x} \rangle \rightarrow 0. \quad (4.24)$$

De fato, algumas manipulações algébricas a partir de (4.20) nos permitem escrever

$$\langle x_t^* + y_t^*, x_t - \bar{x} \rangle = \langle x^* + \frac{1}{n}b^*, x_t - \bar{x} \rangle - \langle z_t^*, z_t - \bar{x} \rangle + \langle z_t^*, z_t - x_t \rangle. \quad (4.25)$$

Seja $\varepsilon > 0$. Escolhemos um índice $t_0 = (n_0, V_0)$ de Λ tal que V_0 contém $\bar{x}$ e $\|x^*\| + 2 \leq n_0\varepsilon$. Para t maior que t_0 , tomando-se $y = \bar{x}$ em (4.22) constata-se $\langle z_t^*, \bar{x} - z_t \rangle \leq 0$. Com isto, alguns cálculos efetuados a partir de (4.25) juntamente com os resultados de (i) e (iv) fornecem

$$|\langle x_t^* + y_t^*, x_t - \bar{x} \rangle| \leq \|x^*\| \|x_t - \bar{x}\| + \|x_t - \bar{x}\| + \|z_t^*\| \|z_t - x_t\| \leq \frac{1}{n}(\|x^*\| + 2).$$

Por fim, para todo t maior que t_0 ,

$$|\langle x_t^* + y_t^*, x_t - \bar{x} \rangle| < \varepsilon.$$

Isto prova a afirmação. Por outro lado,

$$\langle x_t^* + y_t^*, \bar{x} - x_t \rangle = \langle x_t^*, \bar{x} - x_t \rangle + \langle y_t^*, y_t - x_t \rangle + \langle y_t^*, \bar{x} - y_t \rangle.$$

Depois de algumas manipulações algébricas a partir desta equação obtemos

$$\begin{aligned} \limsup_{t \in \Lambda} \langle x_t^*, x_t - \bar{x} \rangle &\leq \limsup_{t \in \Lambda} \|y_t^*\| \|y_t - x_t\| + \limsup_{t \in \Lambda} \langle y_t^*, \bar{x} - y_t \rangle \\ &\quad + \limsup_{t \in \Lambda} \langle x_t^* + y_t^*, x_t - \bar{x} \rangle. \end{aligned}$$

A primeira e a terceira parcela do lado direito desta expressão são iguais a zero por (iv) e (4.24). Resulta que

$$0 \leq \liminf_{t \in \Lambda} \langle x_t^*, \bar{x} - x_t \rangle + \limsup_{t \in \Lambda} \langle y_t^*, \bar{x} - y_t \rangle. \quad (4.26)$$

Além disso, como $x_t^* \in \partial f(x_t)$ e $y_t^* \in \partial g(y_t)$, a definição de subdiferencial fornece

$$\langle x_t^*, \bar{x} - x_t \rangle \leq f(\bar{x}) - f(x_t) \text{ e } \langle y_t^*, \bar{x} - y_t \rangle \leq g(\bar{x}) - g(y_t).$$

Ou ainda, usando (1)(a),

$$\limsup_{t \in \Lambda} \langle x_t^*, \bar{x} - x_t \rangle \leq 0 \text{ e } \limsup_{t \in \Lambda} \langle y_t^*, \bar{x} - y_t \rangle \leq 0. \quad (4.27)$$

Por fim, segue de (4.26) e (4.27) a primeira convergência em (1)(b). A segunda é provada de maneira análoga.

Resta provar (1)(c). Propriedades da norma e a definição de M_t fornecem as inequações

$$\|x_t^*\| \leq \|x_t^* - x^*\| + \|x^*\| \text{ e } \|x_t - y_t\| + \|x_t - z_t\| + \|y_t - z_t\| \leq 3M_t,$$

que multiplicadas resultam em

$$x_t^* (\|x_t - y_t\| + \|x_t - z_t\| + \|y_t - z_t\|) \leq 3M_t \|x_t^* - x^*\| + 3M_t \|x^*\|. \quad (4.28)$$

De (i) é possível majorar M_t por $2/n$ (basta partir da definição de M_t , somar e subtrair $\bar{x}$ no argumento de cada norma e usar a desigualdade triangular). Além disso, a primeira parcela do lado direito de (4.28) pode ser majorada por $3/n$ usando-se (iv). Disto resulta a inequação

$$\|x_t^*\| (\|x_t - y_t\| + \|x_t - z_t\| + \|y_t - z_t\|) \leq \frac{3}{n} (1 + 2\|x^*\|).$$

Sejam $\varepsilon > 0$ e $t_0 = (n_0, V_0)$ um índice de Λ tal que $3(1 + 2\|x^*\|) < n_0\varepsilon$. Então, para t maior que t_0 ,

$$\|x_t^*\| (\|x_t - y_t\| + \|x_t - z_t\| + \|y_t - z_t\|) < \frac{\varepsilon}{3}.$$

De forma mais direta, e também para todo t maior que t_0 , constata-se de (iv) que

$$\|y_t^*\| (\|x_t - y_t\| + \|x_t - z_t\| + \|y_t - z_t\|) < \frac{\varepsilon}{3}$$

e

$$\|z_t^*\| (\|x_t - y_t\| + \|x_t - z_t\| + \|y_t - z_t\|) < \frac{\varepsilon}{3}.$$

A soma das três últimas inequações fornece

$$(\|x_t^*\| + \|y_t^*\| + \|z_t^*\|)(\|x_t - y_t\| + \|x_t - z_t\| + \|y_t - z_t\|) < \varepsilon, \quad \forall t \geq t_0. \quad (4.29)$$

Em particular

$$(\|x_t^*\| + \|y_t^*\|)\|x_t - y_t\| < \varepsilon, \quad \forall t \geq t_0,$$

donde concluímos a demonstração do item (1)(c). Assim, completamos a demonstração do caso particular.

Consideremos o caso geral. Com o propósito de usar o resultado do caso particular, definimos as funções $\varphi, \psi_A : \mathcal{X} \times \mathcal{Y} \rightarrow \overline{\mathbb{R}}$ definidas por

$$\varphi(x, y) = f(x) + g(y) \quad \text{e} \quad \psi_A(x, y) = \delta_G(x, y),$$

onde G é o gráfico do operador A . Por hipótese $x^* \in \partial(f + g \circ A)(\bar{x})$. Afirmamos que

$$(x^*, 0) \in \partial(\varphi + \psi_A)(\bar{x}, A\bar{x}).$$

De fato, basta partir da definição de subdiferencial e usar a definição das funções φ e ψ_A para obter a desigualdade

$$\langle x^*, u - \bar{x} \rangle \leq \varphi(u, Au) - \varphi(\bar{x}, A\bar{x}),$$

válida para todo $u \in \mathcal{X}$. Ou ainda,

$$\langle (x^*, 0), (u, v) - (\bar{x}, A\bar{x}) \rangle \leq \varphi(u, v) + \psi_A(u, v) - \varphi(\bar{x}, A\bar{x}) - \psi_A(\bar{x}, A\bar{x}),$$

válida para todo $u \in \mathcal{X}$ e todo $v \in \mathcal{Y}$. Esta última desigualdade completa a prova da afirmação. O resultado do caso particular (com $(x^*, 0)$, φ e ψ_A no lugar de x^* , f e g) implica a existência de redes (x_t, y_t) e (a_t, Aa_t) em $\mathcal{X} \times \mathcal{Y}$, bem como (x_t^*, y_t^*) e (a_t^*, b_t^*) em $\mathcal{X}^* \times \mathcal{Y}^*$, com

$$(x_t^*, y_t^*) \in \partial\varphi(x_t, y_t) \quad \text{e} \quad (a_t^*, b_t^*) \in \psi_A(a_t, Aa_t)$$

tais que

$$(x_t^*, y_t^*) + (a_t^*, b_t^*) \xrightarrow{*} (x^*, 0) \quad (4.30)$$

e cada uma das sentenças a seguir procede

$$(I) \quad (a) \quad (x_t, y_t) \rightarrow (\bar{x}, A\bar{x}), \quad a_t \rightarrow \bar{x},$$

$$(b) \quad f(x_t) + g(y_t) \rightarrow f(\bar{x}) + g(A\bar{x});$$

$$(II) \quad \langle (x_t^*, y_t^*), (x_t, y_t) - (\bar{x}, A\bar{x}) \rangle \rightarrow 0$$

$$(III) \quad (\| (x_t^*, y_t^*) \| + \| (a_t^*, b_t^*) \|) \| (x_t, y_t) - (a_t, Aa_t) \| \rightarrow 0.$$

Usemos estes resultados para provar (4.18) e (1)(a) a (1)(c).

Para provar (4.18), afirmamos que

$$a_t^* = -b_t^* \circ A. \quad (4.31)$$

De fato, como $(a_t^*, b_t^*) \in \partial\psi_A(a_t, Aa_t)$, segue de (1.17) que

$$\langle (a_t^*, b_t^*), (u, Au) \rangle \leq 0 \Rightarrow \langle a_t^* + b_t^* \circ A, u \rangle \leq 0, \forall u \in \mathcal{X}.$$

Isto mostra que $a_t^* + b_t^* \circ A$ é o funcional identicamente nulo e prova a afirmação. De (4.31) e de (4.30) resulta que

$$x_t^* - b_t^* \circ A \xrightarrow{*} x^* \quad \text{e} \quad y_t^* + b_t^* \xrightarrow{*} 0.$$

Logo,

$$x_t^* + y_t^* \circ A = (x_t^* - b_t^* \circ A) + (y_t^* + b_t^*) \circ A \xrightarrow{*} x^*.$$

Para provar (1)(a) usamos a semicontinuidade de f e g , juntamente com o resultado de (I)(a), para constatar que

$$\liminf_{t \in \Lambda} f(x_t) = f(\bar{x}) \quad \text{e} \quad \liminf_{t \in \Lambda} g(y_t) = g(A\bar{x}).$$

Destas duas equações e de (I)(b) resulta que (Proposição 4.4)

$$f(x_t) \rightarrow f(\bar{x}) \quad \text{e} \quad g(y_t) \rightarrow g(A\bar{x}).$$

Isto completa a prova do item (1)(a).

Provemos o item (1)(b). Visto que $(x_t^*, y_t^*) \in \partial\varphi(x_t, y_t)$, temos que (Afirmação 4.1-(1)) $x_t^* \in \partial f(x_t)$ e $y_t^* \in \partial g(y_t)$, isto é,

$$\langle x_t^*, u - x_t \rangle \leq f(u) - f(x_t), \quad \forall u \in \mathcal{X};$$

$$\langle y_t^*, v - y_t \rangle \leq g(v) - g(y_t), \quad \forall v \in \mathcal{Y}.$$

Se $u = \bar{x}$ e $v = A\bar{x}$ segue de (1)(a) que

$$\limsup_{t \in \Lambda} \langle x_t^*, \bar{x} - x_t \rangle \leq 0 \quad \text{e} \quad \limsup_{t \in \Lambda} \langle y_t^*, A\bar{x} - y_t \rangle \leq 0. \quad (4.32)$$

Por outro lado, como

$$\langle (x_t^*, y_t^*), (\bar{x}, A\bar{x}) - (x_t, y_t) \rangle = \langle x_t^*, \bar{x} - x_t \rangle + \langle y_t^*, A\bar{x} - y_t \rangle$$

temos

$$\begin{aligned} \limsup_{t \in \Lambda} \langle x_t^*, x_t - \bar{x} \rangle &\leq \limsup_{t \in \Lambda} \langle y_t^*, A\bar{x} - y_t \rangle \\ &\quad + \limsup_{t \in \Lambda} \langle (x_y^*, y_t^*), (x_t, y_t) - (\bar{x}, A\bar{x}) \rangle. \end{aligned}$$

O resultado de (II) garante que esta última parcela é nula. Segue que

$$0 \leq \liminf_{t \in \Lambda} \langle x_t^*, \bar{x} - x_t \rangle + \limsup_{t \in \Lambda} \langle y_t^*, A\bar{x} - y_t \rangle. \quad (4.33)$$

De (4.32) e (4.33) concluímos que

$$\langle x_t^*, \bar{x} - x_t \rangle \xrightarrow[t \in \Lambda]{} 0.$$

De maneira análoga constata-se que

$$\langle y_t^*, A\bar{x} - y_t \rangle \xrightarrow[t \in \Lambda]{} 0.$$

Por fim, resta provar (1)(c). Faremos isto usando (III). Para tanto, majoramos $\|y_t - Ax_t\|$. Depois de somar e subtrair Aa_t , propriedades da norma nos permitem escrever

$$\begin{aligned} \|y_t - Ax_t\| &\leq \|y_t - Aa_t\| + \|A\|\|a_t - x_t\| \\ &\leq (1 + \|A\|)(\|y_t - Aa_t\| + \|x_t - a_t\|) \\ &= (1 + \|A\|)\|(x_t, y_t) - (a_t, Aa_t)\|. \end{aligned}$$

Ao multiplicar ambos os lados desta inequação pela norma de (x_t^*, y_t^*) obtemos

$$(\|x_t^*\| + \|y_t^*\|)\|y_t - Ax_t\| \leq (1 + \|A\|)\|(x_t^*, y_t^*)\|\|(x_t, y_t) - (a_t, Aa_t)\|.$$

Seja $\varepsilon > 0$. Por (III) existe t_0 em Λ tal que

$$\|(x_t^*, y_t^*)\|\|(x_t, y_t) - (a_t, Aa_t)\| < \frac{\varepsilon}{1 + \|A\|}, \quad \forall t \geq t_0.$$

Basta combinar as duas últimas inequações para constatar que

$$(\|x_t^*\| + \|y_t^*\|)\|y_t - Ax_t\| < \varepsilon, \quad \forall t \geq t_0.$$

Assim, concluímos a prova de (1)(c) e completamos a demonstração do caso geral, isto é, constatamos que (0) implica (1).

As implicações de (1) em (2) e de (2) em (3) são verificadas de forma imediata.

(3) $\implies$ (0). Seja $u \in \mathcal{X}$. Temos

$$\langle x_t^* + y_t^* \circ A, u - \bar{x} \rangle = \langle x_t^*, u - x_t \rangle + \langle x_t^*, x_t - \bar{x} \rangle + \langle y_t^*, Au - y_t \rangle + \langle y_t^*, y_t - A\bar{x} \rangle.$$

Como $x_t^* \in \partial f(x_t)$ e $y_t^* \in \partial g(y_t)$, a definição de subdiferencial fornece uma majoração para a primeira e terceira parcela do lado direito desta equação, de modo que podemos escrever

$$\begin{aligned} \langle x_t^* + y_t^* \circ A, u - \bar{x} \rangle &\leq f(u) - [f(x_t) - \langle x_t^*, x_t - \bar{x} \rangle] \\ &\quad + g(Au) - [g(y_t) - \langle y_t^*, y_t - A\bar{x} \rangle]. \end{aligned} \quad (4.34)$$

Assim, segue de (4.18) e (3)(b) que

$$\begin{aligned} \langle x^*, u - \bar{x} \rangle &\leq f(u) - f(\bar{x}) + g(Au) - g(A\bar{x}) \\ &= (f + g \circ A)(u) - (f + g \circ A)(\bar{x}). \end{aligned}$$

Como $u \in \mathcal{X}$ é arbitrário, concluímos que $x^* \in \partial(f + g \circ A)(\bar{x})$.

As implicações de (1) em (4) e de (4) em (5) são verificadas de forma imediata.

(5) $\implies$ (0). Seja $u \in \mathcal{X}$. As hipóteses disponíveis nos permitem obter (4.34) da mesma maneira que foi obtida anteriormente. Depois de reorganizar as parcelas de (4.34) obtemos

$$\begin{aligned} f(u) + g(Au) - \liminf_{t \in \Lambda} \langle x_t^* + y_t^* \circ A, u - \bar{x} \rangle &\geq \liminf_{t \in \Lambda} \langle x_t^*, \bar{x} - x_t \rangle \\ &\quad + \liminf_{t \in \Lambda} \langle y_t^*, A\bar{x} - y_t \rangle + \liminf_{t \in \Lambda} f(x_t) + \liminf_{t \in \Lambda} g(y_t). \end{aligned}$$

Ou ainda, de (5)(a), (5)(b) e da semicontinuidade das funções f e g obtemos

$$\begin{aligned} f(u) + g(Au) - \langle x^*, u - \bar{x} \rangle &\geq f(\bar{x}) + g(A\bar{x}) \\ \Leftrightarrow \langle x^*, u - \bar{x} \rangle &\leq (f + g \circ A)(u) - (f + g \circ A)(\bar{x}). \end{aligned}$$

Como $u \in \mathcal{X}$ é arbitrário, concluímos que $x^* \in \partial(f + g \circ A)(\bar{x})$. Com isto, completamos a demonstração do teorema.

□

As próximas duas proposições são exemplos de aplicações do teorema que acabamos de demonstrar.

Dado $\bar{x}$ no domínio de uma função f (convexa, própria e s.c.i.),

possivelmente se tem $\partial f(\bar{x})$ vazio. Contudo, sempre é possível encontrar-se x , tão próximo de $\bar{x}$ quanto se queira, tal que $\partial f(x)$ é não-vazio. Isto é o que garante a proposição a seguir.

Proposição 4.1 [22, Proposição 2] *Seja $f : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ uma função convexa, própria e s.c.i.. Seja $\varepsilon > 0$. Dado $\bar{x}$ no domínio de f , existe $x \in \mathcal{X}$ tal que $\partial f(x)$ é não-vazio e*

$$\|x - \bar{x}\| + |f(x) - f(\bar{x})| < \varepsilon.$$

Demonstração: Seja $S := \{\bar{x}\}$. Então,

$$f(\bar{x}) + \delta_S(\bar{x}) \leq f(x) + \delta_S(x), \quad \forall x \in \mathcal{X}.$$

Segue que $\bar{x}$ é um minimizador da função $f + \delta_S$. Isto implica que $0 \in \partial(f + \delta_S)(\bar{x})$. Nestas condições, a forma (2) do Teorema 4.1 garante a existência de uma rede $x_t \in \mathcal{X}$ tal que

$$x_t \rightarrow \bar{x} \text{ e } f(x_t) \rightarrow f(\bar{x}).$$

Basta tomar t suficientemente grande para concluir o resultado.

□

A prova original do teorema a seguir é datada de 1966 e foi apresentada por Rockafellar [19, Teorema 3]. A demonstração apresentada neste trabalho é ligeiramente mais simples do que a original.

Proposição 4.2 (Teorema de Integração de Rockafellar) [22, Proposição 3] *Sejam $f, g : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ funções convexas, próprias e s.c.i.. Suponha que $\partial f(x)$ está contido em $\partial g(x)$ para todo $x \in \mathcal{X}$. Então,*

$$f = g + \text{constante}.$$

Demonstração: O caso unidimensional pode ser provado usando-se derivadas laterais [18, Teorema 24.2]. Provemos o caso geral. A Proposição 4.1 nos garante que $\partial f(a)$ é não-vazio para algum a no domínio de f . Sem perda de generalidade, suponhamos que $a = 0$. Seja $b \in \mathcal{X}$ e $V_b = \mathbb{R}b$ o subespaço unidimensional gerado por b . Denotemos por A_b o operador imersão de V_b em $\mathcal{X}$, isto é,

$$A_b u = u, \quad \forall u \in V_b.$$

Nestas condições, afirmamos que

$$\partial(f \circ A_b)(u) \subset \partial(g \circ A_b)(u), \quad \forall u \in V_b.$$

De fato, seja $u \in V_b$. A inclusão vale trivialmente se $\partial(f \circ A_b)(u)$ é vazio. Seja $u^* \in \partial(f \circ A_b)(u)$. Pela forma (4) do Teorema 4.1 (com f e u no lugar de g e $\bar{x}$), existem redes $x_t \in \mathcal{X}$ e $x_t^* \in \mathcal{X}^*$ tais que

$$x_t^* \in \partial f(x_t); \quad x_t^* \circ A_b \xrightarrow{*} u^*; \quad \langle x_t^*, x_t - A_b u \rangle \rightarrow 0; \quad x_t \rightarrow A_b u = u. \quad (4.35)$$

Visto que $\partial f(x)$ está contido em $\partial g(x)$, para todo $x \in \mathcal{X}$, temos que (4.35) também vale com g no lugar de f . Segue novamente pela forma (4) do Teorema 4.1, desta vez usando a implicação contrária, que $u^* \in \partial(g \circ A_b)(u)$. Completamos a prova da afirmação. Ao combinar o resultado do caso unidimensional com a afirmação, constatamos que existe uma constante c tal que

$$(f \circ A_b)(0) = (g \circ A_b)(0) + c \quad \text{e} \quad (f \circ A_b)(b) = (g \circ A_b)(b) + c.$$

Após subtrair a primeira equação da segunda, resulta

$$f(b) + g(0) = g(b) + f(0). \quad (4.36)$$

Visto que o vetor nulo está na interseção dos domínios de f e g (caso contrário $\partial g(0)$ seria vazio e, conseqüentemente, $\partial f(0)$ também seria vazio), concluímos de (4.36) que b está no domínio de f se, e somente se, b está no domínio de g , isto é, f e g possuem o mesmo domínio. Além disso, a escolha de $b \in \mathcal{X}$ foi arbitrária. Portanto,

$$f = g + [f(0) - g(0)].$$

□

4.1 SUBDIFERENCIAIS DA SOMA E DA PRÉ-COMPOSIÇÃO

Em análise convexa, frequentemente surge o interesse em minimizar a soma de duas funções. Sendo assim, é muito conveniente ter uma fórmula para subdiferencial da soma, em particular, de duas funções convexas, próprias e s.c.i.. O mesmo se aplica à pré-composição de uma função (convexa, própria e s.c.i.) com um operador linear contínuo. Na Seção 2.1 apresentamos uma fórmula para o subdiferencial da soma que depende de uma condição de qualificação. Nesta seção apresentamos o

Teorema 4.2, que contempla tanto a soma como a pré-composição e não requer qualquer tipo de qualificação. Originalmente este resultado foi apresentado por Hiriart-Urruty e Phelps em [12] na forma de dois teoremas, a saber, Teorema 2.1 e Teorema 3.1. Para o caso da soma, outras duas demonstrações foram publicadas por Hiriart-Urruty, Moussaoui, Seeger, e Volle em [11]. Contudo, a demonstração que apresentamos neste trabalho, devida a Thibault [22], segue como uma aplicação do Teorema 4.1 e é mais simples do que as três mencionadas anteriormente.

Teorema 4.2 [22, Proposição 1] *Seja A um operador linear contínuo de $\mathcal{X}$ em $\mathcal{Y}$. Sejam $f : \mathcal{X} \rightarrow \mathbb{R}$ e $g : \mathcal{Y} \rightarrow \mathbb{R}$ funções convexas, próprias e s.c.i.. Se existe $\bar{x}$ no domínio de f tal que $A\bar{x}$ está no domínio de g , então,*

$$\partial(f + g \circ A)(\bar{x}) = \bigcap_{\varepsilon > 0} \overline{\partial_\varepsilon f(\bar{x}) + \partial_\varepsilon g(A\bar{x}) \circ A}^{w^*}, \quad (4.37)$$

onde

$$\partial_\varepsilon g(A\bar{x}) \circ A = \{y^* \circ A; y^* \in \partial_\varepsilon g(A\bar{x})\}.$$

Demonstração: Seja $x^* \in \partial(f + g \circ A)(\bar{x})$. Da forma (3) do Teorema 4.1 obtemos redes $x_t \in \mathcal{X}$ e $y_t \in \mathcal{Y}$, bem como $x_t^* \in \mathcal{X}^*$ e $y_t^* \in \mathcal{Y}^*$, tais que

$$x_t^* \in \partial f(x_t) \text{ e } y_t^* \in \partial g(y_t); \quad (4.38)$$

vale (4.18); valem as sentenças (3)(a) e (3)(b). Seja $\varepsilon > 0$. Segue de (3)(b) que existe t_0 tal que, para todo $t \geq t_0$, valem as desigualdades:

$$f(x_t) - \langle x_t^*, x_t - \bar{x} \rangle \geq f(\bar{x}) - \varepsilon \quad (4.39)$$

$$g(y_t) - \langle y_t^*, y_t - A\bar{x} \rangle \geq g(A\bar{x}) - \varepsilon. \quad (4.40)$$

Por outro lado, a partir de (4.38) obtemos, para todo $x \in \mathcal{X}$,

$$\langle x_t^*, x - x_t \rangle \leq f(x) - f(x_t). \quad (4.41)$$

De (4.39) e (4.41) resulta

$$\begin{aligned} \langle x_t^*, x - \bar{x} \rangle &= \langle x_t^*, x - x_t \rangle + \langle x_t^*, x_t - \bar{x} \rangle \\ &\leq f(x) - f(x_t) + \langle x_t^*, x_t - \bar{x} \rangle \\ &\leq f(x) - f(\bar{x}) + \varepsilon, \quad \forall x \in \mathcal{X}. \end{aligned}$$

Portanto, $x_t^* \in \partial_\varepsilon f(\bar{x})$. Procedendo-se de maneira análoga constata-se que $y_t^* \in \partial_\varepsilon g(A\bar{x})$. Assim, concluímos que

$$x_t^* + y_t^* \circ A \in \partial_\varepsilon f(\bar{x}) + \partial_\varepsilon g(A\bar{x}) \circ A, \quad \forall t \geq t_0.$$

Por fim, considerando a convergência em (4.18), segue que

$$x^* \in \overline{\partial_\varepsilon f(\bar{x}) + \partial_\varepsilon g(A\bar{x}) \circ A}^{w^*}.$$

Resta provarmos a inclusão contrária. Sejam $\varepsilon > 0$ e $x^* \in \partial_\varepsilon f(\bar{x}) + \partial_\varepsilon g(A\bar{x}) \circ A$. Então, existem $x_1^* \in \partial_\varepsilon f(\bar{x})$ e $x_2^* \in \partial_\varepsilon g(A\bar{x})$ tais que

$$x_1^* + x_2^* \circ A = x^*.$$

A partir da definição de ε -subdiferencial obtemos, para todo $x \in \mathcal{X}$,

$$\langle x_1^*, x - \bar{x} \rangle \leq f(x) - f(\bar{x}) + \varepsilon \quad \text{e} \quad \langle x_2^*, Ax - A\bar{x} \rangle \leq g(Ax) - g(A\bar{x}) + \varepsilon.$$

Após somar estas duas inequações concluímos que

$$x^* = x_1^* + x_2^* \circ A \in \partial_{2\varepsilon}(f + g \circ A)(x^*).$$

Assim, como o ε -subdiferencial é um conjunto fechado na topologia fraca-*, constatamos que

$$\overline{\partial_\varepsilon f(\bar{x}) + \partial_\varepsilon g(A\bar{x})}^{w^*} \subset \overline{\partial_{2\varepsilon}(f + g \circ A)(\bar{x})}^{w^*} = \partial_{2\varepsilon}(f + g \circ A)(\bar{x}).$$

Tomando a interseção sobre todos os $\varepsilon > 0$ e usando (1.21) concluímos a demonstração do teorema.

□

Existe uma versão do Teorema do Valor Médio para subdiferenciais. O Lema seguinte contempla o caso particular onde $\mathcal{X} = \mathbb{R}$ e na sequência demonstramos o caso geral (Proposição 4.3).

Lema 4.3 (Teorema do Valor Médio (em $\mathbb{R}$)) [10, Teorema 2.3.3]
Seja $\varphi : \mathbb{R} \rightarrow \overline{\mathbb{R}}$ uma função convexa e própria. Sejam a e b pontos do domínio de φ . Então, existe t_0 entre a e b de modo que

$$\varphi(b) - \varphi(a) \in \partial\varphi(t_0).$$

Demonstração: Sem perda de generalidade, suponhamos que $a = 0$ e $b = 1$. A convexidade de φ garante que o interior do intervalo $(0, 1)$ está contido no interior do domínio de φ . Segue, por um resultado clássico de análise convexa, que φ é contínua no intervalo $(0, 1)$ [14, Seção 6, Teorema 10]. Seja $\psi : [0, 1] \rightarrow \mathbb{R}$ a função definida por

$$\psi(t) = \varphi(t) - \varphi(0) - t[\varphi(1) - \varphi(0)].$$

A convexidade e continuidade de φ no intervalo $(0, 1)$ implicam em ψ ser convexa e continua. A continuidade nos garante que ψ possui um minimizador, e a convexidade nos garante que ψ assume seu valor mínimo em algum ponto entre 0 e 1, pois $\psi(0) = 0 = \psi(1)$. Seja t_0 o minimizador de ψ . Então,

$$0 \in \partial\psi(t_0) = \partial\varphi(t_0) - [\varphi(1) - \varphi(0)] \Rightarrow \varphi(1) - \varphi(0) \in \partial\varphi(t_0).$$

□

Proposição 4.3 (Teorema do Valor Médio) [12, Exemplo 3.2] *Seja $f : \mathcal{X} \rightarrow \mathbb{R}$ uma função convexa, própria e s.c.i.. Sejam a e b pontos distintos no domínio de f . Então, existe c entre a e b tal que*

$$f(b) - f(a) \in \bigcap_{\varepsilon > 0} \overline{\langle \partial_\varepsilon f(c), b - a \rangle}^{w^*} = \bigcap_{\varepsilon > 0} \overline{\{ \langle y^*, b - a \rangle; y^* \in \partial_\varepsilon f(c) \}}^{w^*}. \quad (4.42)$$

Demonstração: Sem perda de generalidade, suponhamos $a = 0$. Seja A o operador imersão de $\mathbb{R}$ em $\mathcal{X}$ definido por $At = bt$. Seja $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ a função definida por $\varphi := f \circ A$. Pela proposição anterior, existe t_0 entre 0 e 1 tal que

$$\varphi(1) - \varphi(0) \in \partial\varphi(t_0). \quad (4.43)$$

Pelo Teorema 4.2

$$\partial\varphi(t_0) = \partial(f \circ A)(t_0) = \bigcap_{\varepsilon > 0} \overline{\langle y^*, b \rangle; y^* \in \partial_\varepsilon f(At_0)}^{w^*}. \quad (4.44)$$

Como $\varphi(1) = f(b)$ e $\varphi(0) = f(0)$, segue de (4.43) e (4.44), com $c = At_0$, que

$$f(b) - f(0) \in \bigcap_{\varepsilon > 0} \overline{\langle y^*, b - 0 \rangle; y^* \in \partial_\varepsilon f(c)}^{w^*}. \quad (4.45)$$

□

APÊNDICE

Demonstração da Proposição 3.2

(i) Seja $y \in \text{dom}(Af)$. Então, $\inf \{f(x); Ax = y\} < +\infty$, isto é, $\{f(x); Ax = y\}$ é não-vazio. Portanto, existe $x \in \text{dom}(f)$ tal que

$Ax = y$, ou seja, $y \in \{Ax; x \in \text{dom}(f)\}$. A inclusão contrária segue trivialmente da definição.

(ii) Sejam $u, v \in \text{dom}(Af) = \{Ax; x \in \text{dom}(f)\}$ e λ entre 0 e 1. Sejam x_1 e x_2 no domínio de f tais que $Ax_1 = u$ e $Ax_2 = v$. Neste caso, $A((1 - \lambda)x_1 + \lambda x_2) = (1 - \lambda)u + \lambda v$. Portanto,

$$\begin{aligned} Af((1 - \lambda)u + \lambda v) &= \inf \{f(x); Ax = (1 - \lambda)u + \lambda v\} \\ &\leq f((1 - \lambda)x_1 + \lambda x_2) \\ &\leq (1 - \lambda)f(x_1) + \lambda f(x_2), \end{aligned}$$

para quaisquer x_1 e x_2 tais que $Ax_1 = u$ e $Ax_2 = v$. Assim,

$$\begin{aligned} Af((1 - \lambda)u + \lambda v) &\leq (1 - \lambda) \inf \{f(x_1); Ax_1 = u\} + \lambda f(x_2) \\ &\leq (1 - \lambda)Af(u) + \lambda f(x_2). \end{aligned}$$

Logo,

$$\begin{aligned} Af((1 - \lambda)u + \lambda v) &\leq (1 - \lambda)Af(u) + \lambda \inf \{f(x_2); Ax_2 = v\} \\ &\leq (1 - \lambda)Af(u) + \lambda Af(v). \end{aligned}$$

(iii) Como f é própria, segue de (i) que o domínio de Af é não-vazio. Mostremos que $Af > -\infty$. Por hipótese existe $y^* \in \mathcal{Y}^*$ tal que $A^*y^* = x_0^*$. Da definição de função marginal e da inequação fornecida na hipótese obtemos

$$\begin{aligned} Af(y) &\geq \inf \{\langle x_0^*, x \rangle + c; Ax = y\} \\ &= \inf \{\langle A^*y^*, x \rangle + c; Ax = y\} \\ &= \inf \{\langle y^*, Ax \rangle + c; Ax = y\} \\ &= \langle y^*, y \rangle + c > -\infty, \quad \forall y \in \mathcal{Y}. \end{aligned}$$

□

Proposição 4.4 *Sejam (x_t) e (y_t) redes. Suponhamos que $\liminf x_t = x$ e $\liminf y_t = y$. Se $(x_t + y_t) \rightarrow x + y$, então $x_t \rightarrow x$ e $y_t \rightarrow y$.*

Demonstração: Sejam $a = \limsup x_t$ e $b = \limsup y_t$. Então, $a \geq x$ e $b \geq y$. Suponha que a é estritamente maior que x (o que implica em $a + b$ ser estritamente maior que $x + y$). Por definição de limite superior, existem sub-redes (x_{t_α}) e (y_{t_α}) tais que $x_{t_\alpha} \rightarrow a$ e $y_{t_\alpha} \rightarrow b$. A partir da hipótese obtemos

$$x + y = \lim (x_t + y_t) = \lim (x_{t_\alpha} + y_{t_\alpha}) = \lim x_{t_\alpha} + \lim y_{t_\alpha} = a + b.$$

Contradição! Então, $a = x$. Segue que o limite de x_t é x e

$$\lim y_t = \lim (y_t + x_t) - \lim x_t = x + y - x = y.$$

□

Afirmção 4.1 (Detalhes do Lema 4.2) *Sejam $f_i : \mathcal{X} \rightarrow \overline{\mathbb{R}}$, $i = 1, 2, 3$, funções convexas. Sejam $f, \Delta, \|\cdot\| : \mathcal{X}^3 \rightarrow \overline{\mathbb{R}}$ as funções definidas por*

$$\begin{aligned} f(x_1, x_2, x_3) &= f_1(x_1) + f_2(x_2) + f_3(x_3), \\ \Delta(x_1, x_2, x_3) &= \frac{1}{2}(\|x_1 - x_2\|^2 + \|x_1 - x_3\|^2 + \|x_2 - x_3\|^2), \\ \|\cdot\| \|(x_1, x_2, x_3)\| &= \|x_1\| + \|x_2\| + \|x_3\|. \end{aligned}$$

As seguintes afirmativas são verdadeiras:

(1) Se (x_1, x_2, x_3) está no domínio de f , então

$$\partial f(x_1, x_2, x_3) = \partial f_1(x_1) \times \partial f_2(x_2) \times \partial f_3(x_3);$$

(2) A função Δ é convexa;

(3) Se $(u_1^*, u_2^*, u_3^*) \in \partial \Delta(u_1, u_2, u_3)$, então

$$\begin{aligned} (a) \quad & u_1^* + u_2^* + u_3^* = 0; \\ (b) \quad & \|u_i^*\| \leq \|u_1 - u_2\| + \|u_1 - u_3\| + \|u_2 - u_3\|. \end{aligned}$$

Demonstração:

(1) Seja $(x_1^*, x_2^*, x_3^*) \in \partial f(x_1, x_2, x_3)$. Da definição de subdiferencial obtemos, para qualquer $(y_1, y_2, y_3) \in \mathcal{X}^3$,

$$\begin{aligned} \langle x_1^*, y_1 - x_1 \rangle + \langle x_2^*, y_2 - x_2 \rangle + \langle x_3^*, y_3 - x_3 \rangle &\leq \\ &\leq f_1(y_1) + f_2(y_2) + f_3(y_3) - [f_1(x_1) + f_2(x_2) + f_3(x_3)]. \end{aligned}$$

Se $y_2 = x_2$ e $y_3 = x_3$ resulta

$$\langle x_1^*, y_1 - x_1 \rangle \leq f_1(y_1) - f_1(x_1), \quad \forall y_1 \in \mathcal{X},$$

donde concluímos que $x_1^* \in \partial f_1(x_1)$. De maneira análoga, se $y_1 = x_1$ e $y_3 = x_3$, resulta que $x_2^* \in \partial f_2(x_2)$ e, se $y_1 = x_1$ e $y_2 = x_2$, concluímos que $x_3^* \in \partial f_3(x_3)$. Portanto,

$$(x_1^*, x_2^*, x_3^*) \in \partial f_1(x_1) \times \partial f_2(x_2) \times \partial f_3(x_3).$$

Isso mostra que o lado esquerdo está contido no lado direito. Para provar a inclusão contrária basta somar as inequações resultantes da definição de subdiferencial.

(2) Segue da definição de função convexa e do fato que $x \mapsto \|x\|^2$ é convexa.

(3) Da definição de subdiferencial temos

$$\begin{aligned} \langle (u_1^*, u_2^*, u_3^*), (y_1, y_2, y_3) - (u_1, u_2, u_3) \rangle \\ \leq \Delta(y_1, y_2, y_3) - \Delta(u_1, u_2, u_3), \end{aligned} \quad (4.46)$$

para qualquer $(y_1, y_2, y_3) \in \mathcal{X}^3$. Se $y_1 = y_2 = y_3 = z$ resulta

$$\langle u_1^* + u_2^* + u_3^*, z \rangle \leq \langle u_1^*, u_1 \rangle + \langle u_2^*, u_2 \rangle + \langle u_3^*, u_3 \rangle - \Delta(u_1, u_2, u_3),$$

para todo $z \in \mathcal{X}$. Como o lado direito é constante, só podemos ter $u_1^* + u_2^* + u_3^* = 0$.

Provemos (4)(b). Sejam $r > 0$ e $v \in \mathcal{X}$ tal que $\|v\| = 1$. Para $y_1 = u_1 + rv$, $y_2 = u_2$ e $y_3 = u_3$ resulta de (4.46) que

$$\begin{aligned} \langle u_1^*, rv \rangle &\leq \frac{1}{2} (\|u_1 + rv - u_2\|^2 - \|u_1 - u_2\|^2) \\ &\quad + \frac{1}{2} (\|u_1 + rv - u_3\|^2 - \|u_1 - u_3\|^2) \\ &= \frac{1}{2} (\|u_1 + rv - u_2\| - \|u_1 - u_2\|) (\|u_1 + rv - u_2\| + \|u_1 - u_2\|) \\ &\quad + \frac{1}{2} (\|u_1 + rv - u_3\| - \|u_1 - u_3\|) (\|u_1 + rv - u_3\| + \|u_1 - u_3\|). \end{aligned}$$

A desigualdade triangular e o fato que $\|v\| = 1$ nos permitem majorar o lado direito desta última inequação e obter

$$\begin{aligned} \langle u_1^*, rv \rangle &\leq \frac{1}{2} (\|u_1 - u_2\| + r - \|u_1 - u_2\|) (\|u_1 - u_2\| + r + \|u_1 - u_2\|) \\ &\quad + \frac{1}{2} (\|u_1 - u_3\| + r - \|u_1 - u_3\|) (\|u_1 - u_3\| + r + \|u_1 - u_3\|) \\ &= \frac{1}{2} r (\|u_1 - u_2\| + r + \|u_1 - u_2\|) \\ &\quad + \frac{1}{2} r (\|u_1 - u_3\| + r + \|u_1 - u_3\|). \end{aligned}$$

Depois de dividir esta última inequação por r e tomar o limite com r

tendendo a zero podemos escrever

$$\langle u_1^*, v \rangle \leq \|u_1 - u_2\| + \|u_1 - u_3\| + \|u_2 - u_3\|.$$

Visto que esta desigualdade vale para qualquer v tal que $\|v\| = 1$, concluimos que

$$\|u_1^*\| \leq \|u_1 - u_2\| + \|u_1 - u_3\| + \|u_2 - u_3\|. \quad (4.47)$$

De maneira análoga conclui-se que

$$\|u_2^*\|, \|u_3^*\| \leq \|u_1 - u_2\| + \|u_1 - u_3\| + \|u_2 - u_3\|. \quad (4.48)$$

□

CONCLUSÃO

Conforme já mencionado, o ε -subdiferencial foi introduzido ainda na década de 60, mas foi só em 1993 que regras de cálculo subdiferencial sem condições de qualificação foram apresentadas pela primeira vez. Entre as fórmulas estabelecidas por Hiriart-Urruty e Phelps [12] naquela ocasião, estão aquelas que apresentamos no Capítulo 3. Fórmulas para o subdiferencial da soma e da pré-composição também são encontradas em [12] e até chegaram a fazer parte deste trabalho tal com são apresentadas no artigo. Contudo, optamos por demonstrá-las usando técnicas mais recentes, no caso, usando limites, conforme pode ser observado no Capítulo 4. Afinal, não seria exagero dizer que estas são as regras de cálculo mais importantes deste trabalho.

É preciso admitir que a demonstração do Teorema 4.1 não é elementar. Tampouco o lema que o precede. Em todo caso, as caracterizações estabelecidas no Teorema 4.1 são tão ricas que permitem Thibault obter a fórmula para o subdiferencial da soma e da pré-composição (as mesmas introduzidas por Hiriart-Urruty e Phelps em [12]) como um corolário. O mesmo se aplica à Proposição 4.1 e à Proposição 4.2. Ainda como uma aplicação do Teorema 4.1, Thibault [22] provou um resultado muito bem conhecido da análise convexa, a saber, o fato de que o operador subdiferencial é monótono maximal. Também em espaços de Banach, outras demonstrações deste resultado foram publicadas usando diferentes ferramentas. Entre elas vale citar a original, devida a Rockafellar [21], e outra muito mais simples apresentada em 2008 por Marques Alves e Svaiter [16]. Mesmo reconhecendo a importância do assunto, optamos por não abordar a teoria maximal acreditando que se desvia um pouco do objetivo deste trabalho. Maiores informações podem ser encontradas nos dois artigos citados.

Cogitamos abordar o conceito de subdiferencial mais profundamente antes de passar ao estudo das regras de cálculo. Contudo, nos contivemos em fazer isso para evitar que o trabalho se tornasse demasiadamente extenso. Alguém que por ventura venha a se interessar por mais informações sobre o assunto pode ver as referências [1, 10, 18].

Enfim, além de ter contribuído significativamente para a formação do autor, acreditamos que este trabalho fornece uma boa visão geral das principais ferramentas utilizadas em análise convexa, sobretudo em dimensão infinita.

REFERÊNCIAS

- [1] BARBU, V.; PRECUPANU, T. *Convexity and Optimization in Banach Spaces*. Dordrecht: Springer, 1986.
- [2] BREZIS, H. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. New York: Springer, 2011.
- [3] BRØNDSTED, A.; ROCKAFELLAR, R. T. On the subdifferentiability of convex functions. *Proc. Amer. Math. Soc.*, p. 605–611, 1965.
- [4] BAUSCHKE, H. H.; COMBETTES, P. L. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. New York: Springer, 2011.
- [5] COMBETTES, P. L.; WAJS, V. R. Signal recovery by proximal forward-backward splitting. *Multiscale Model. Simul.*, p. 1168–1200 (electronic), 2005.
- [6] COMBETTES, P. L.; PESQUET, J.-C. Proximal splitting methods in signal processing. In: *Fixed-point algorithms for inverse problems in science and engineering*. [S.l.]: Springer, 2011.
- [7] EKELAND, I.; TÉMAM, R. *Convex Analysis and Variational Problems*. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 1999.
- [8] EKELAND, I. On the variational principle. *J. Math. Anal. Appl.*, p. 324–353, 1974.
- [9] EKELAND, I.; TEMAM, R. *Analyse convexe et problèmes variationnels*. [S.l.]: Dunod, 1974.
- [10] HIRIART-URRUTY, J.-B.; LEMARÉCHAL, C. *Fundamentals of Convex Analysis*. Berlin: Springer-Verlag, 2001.
- [11] HIRIART-URRUTY, J.-B. et al. Subdifferential calculus without qualification conditions, using approximate subdifferentials: a survey. *Nonlinear Anal.*, p. 1727–1754, 1995.
- [12] HIRIART-URRUTY, J.-B.; PHELPS, R. R. Subdifferential calculus using ϵ -subdifferentials. *J. Funct. Anal.*, p. 154–166, 1993.

- [13] IOFFE, A. D. Variational analysis and mathematical economics 1: Subdifferential calculus and the second theorem of welfare economics. In: *Advances in mathematical economics. Volume 12.* [S.l.]: Springer, 2009. p. 71–95.
- [14] LIMA, E. L. *Análise Real, volume 2.* 4. ed. Rio de Janeiro: IMPA, 2009.
- [15] MARQUES ALVES, M. Notas de aula da disciplina de análise convexa. Programa de pós-graduação do departamento de Matemática da UFSC. 2012.
- [16] MARQUES ALVES, M.; SVAITER, B. F. A new proof for maximal monotonicity of subdifferential operators. *J. Convex Anal.*, p. 345–348, 2008.
- [17] MOREAU, J. J. *Fonctionnelles convexes.* Collège de France, Paris, 1966.
- [18] ROCKAFELLAR, R. T. *Convex Analysis.* Princeton, NJ: Princeton university press, 1970.
- [19] ROCKAFELLAR, R. T. Characterization of the subdifferentials of convex functions. *Pacific J. Math.*, p. 497–510, 1966.
- [20] ROCKAFELLAR, R. T. *Conjugate duality and optimization.* [S.l.]: Society for Industrial and Applied Mathematics, 1974.
- [21] ROCKAFELLAR, R. T. On the maximal monotonicity of subdifferential mappings. *Pacific J. Math.*, p. 209–216, 1970.
- [22] THIBAUT, L. Limiting convex subdifferential calculus with applications to integration and maximal monotonicity of subdifferential. *Amer. Math. Soc.*, p. 279–289, 2000.
- [23] THIBAUT, L. Sequential convex subdifferential calculus and sequential Lagrange multipliers. *SIAM J. Control Optim.*, p. 1434–1444, 1997.