

UNIVERSIDADE FEDERAL DE GOIÁS  
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA

JOSÉ HENRIQUE SALAZAR DO AMARAL

**Convergência do Método do Ponto  
Proximal para Funções que Satisfazem  
a Desigualdade de Łojasiewicz**

Goiânia  
2012

JOSÉ HENRIQUE SALAZAR DO AMARAL

# **Convergência do Método do Ponto Proximal para Funções que Satisfazem a Desigualdade de Łojasiewicz**

Dissertação apresentada ao Programa de Pós-Graduação do Instituto de Matemática e Estatística da Universidade Federal de Goiás, como requisito parcial para obtenção do título de Mestre em Matemática.

**Área de concentração:** Otimização.

**Orientador:** Prof. Dr. Glaydston de Carvalho Bento

Goiânia  
2012

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador(a).

**José Henrique Salazar do Amaral**

Graduou-se em Matemática na UFG - Universidade Federal de Goiás. Durante o Mestrado, na UFG - Universidade Federal de Goiás, foi bolsista da UFG e desenvolveu um trabalho teórico na área de Otimização.

Dedico este trabalho primeiramente ao Senhor, a meus pais Felisberto Vergílio do Amaral(*in memorian*) e Maria dos Santos Salazar e a minha esposa Gabriela de Araújo Achegaua.

---

## Agradecimentos

---

Agradeço primeiramente a Deus por ter me sustentado e também porque nas dificuldades desta vida sei que Ele sempre me ouviu. Gostaria de esclarecer aqui que todas as pessoas que de alguma forma me auxiliaram, foram colocadas por Ele na minha vida.

Agradeço ao Professor Glaydston por ter aceitado me orientar, por todo o conhecimento, experiência e senso crítico que compartilhou comigo durante a realização do trabalho.

Agradeço aos professores Orizon Pereira Ferreira e João Xavier da Cruz Neto pelas contribuições para a melhoria do presente trabalho e além disso, por terem aceitado participar da Banca da defesa.

Agradeço aos meus colegas de mestrado Caike, Pablo e Alfredo pela ajuda e apoio durante todo o curso. Um agradecimento especial aos meus amigos, irmãos em Cristo Jesus, Darlan Silva Rodrigues e David Sue Alves Diniz que estiveram bastante presentes na minha vida durante o tempo em que estava fazendo este trabalho.

Agradeço a minha mãe, porque desde cedo, mesmo sem ter formação alguma soube me instruir a sempre prosseguir nos estudos.

Agradeço à minha esposa pelo carinho e compreensão.

Pela Fé Abraão, sendo chamado, obedeceu, indo para um lugar que havia de receber por herança; e saiu, sem saber para onde ía.

**Hebreus 11-8,**  
*Palavra de Deus..*

---

## Resumo

---

Henrique Salazar do Amaral, José. **Convergência do Método do Ponto Proximal para Funções que Satisfazem a Desigualdade de Łojasiewicz**. Goiânia, 2012. 53p. Dissertação de Mestrado. Instituto de Matemática e Estatística, Universidade Federal de Goiás.

Neste trabalho é feita uma análise de convergência do Método do Ponto Proximal para funções não necessariamente convexas que satisfazem a desigualdade de Łojasiewicz.

### Palavras-chave

Otimização não-convexa, Método do Ponto Proximal, Desigualdade de Łojasiewicz

---

## Abstract

---

Henrique Salazar do Amaral, José. **Convergence of the Proximal Point Method for functions that satisfy the inequality of Lojasiewicz**. Goiânia, 2012. [53](#)p. MSc. Dissertation. Instituto de Matemática e Estatística, Universidade Federal de Goiás.

This paper presents an analysis of convergence of the proximal point method for functions that satisfy the inequality of Lojasiewicz.

### Keywords

non-convex optimization, Proximal Point Method, Inequality of Lojasiewicz .



---

# Sumário

---

|                                                                                                      |    |
|------------------------------------------------------------------------------------------------------|----|
| Introdução                                                                                           | 12 |
| 1 Preliminares                                                                                       | 14 |
| 1.1 Tópicos de Análise Convexa                                                                       | 14 |
| 1.2 Subdiferencial de Fréchet e Subdiferencial Limite                                                | 30 |
| 2 Método do Ponto Proximal                                                                           | 36 |
| 2.1 Aplicação Proximal                                                                               | 36 |
| 2.2 Análise de Convergência                                                                          | 39 |
| 3 Convergência do Método do Ponto Proximal para Funções que satisfazem a Desigualdade de Łojasiewicz | 41 |
| 3.1 Resultados Auxiliares                                                                            | 41 |
| 3.2 Desigualdade de Łojasiewicz                                                                      | 46 |
| 3.3 Exemplos de funções que satisfazem a desigualdade de Łojasiewicz                                 | 46 |
| 3.4 Resultado de Convergência                                                                        | 49 |
| 3.5 Conclusão                                                                                        | 52 |
| Referências Bibliográficas                                                                           | 53 |

---

## Introdução

---

Considere o seguinte problema de otimização

$$\begin{aligned} \min f(x) \\ \text{t.q. } x \in \mathbb{R}^n, \end{aligned} \tag{0-1}$$

onde  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  é uma função própria semicontinua inferior. O método do ponto proximal, introduzido por Martinet [8] e Rockafellar [11] na década de 70, é um método iterativo utilizado para resolver o problema (0-1). O método do ponto proximal gera uma sequência  $\{x^k\} \subset \mathbb{R}^n$  como segue: dado  $x^0 \in \mathbb{R}^n$ ,

$$x^{k+1} \in \operatorname{argmin}_{x \in \mathbb{R}^n} \left\{ f(x) + \frac{1}{2\lambda_k} \|x - x^k\|^2 \right\}, \tag{0-2}$$

onde  $\{\lambda_k\}$  é uma sequência limitada de números positivos e o conjunto que aparece em (0-2) representa o conjunto dos minimizadores irrestritos da função  $f + (1/2\lambda_k)\|\cdot - x^k\|^2$ . No caso que a função  $f$  é convexa a sequência  $\{x^k\}$  é unicamente determinada e sua convergência, a uma solução do problema (0-1), se baseia na chamada Fejér convergência ao conjunto solução, quando este é não vazio; veja, por exemplo, [4].

Neste trabalho, baseado em [1], apresentamos uma análise de convergência da sequência gerada pelo Método do Ponto Proximal (0-2) para funções não convexas e não diferenciáveis que satisfazem a desigualdade de Łojasiewicz. Uma função  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  própria, semicontinua inferior é dita ter a propriedade de Łojasiewicz se:

Para qualquer ponto crítico generalizado de  $f$ , ver Seção 1.2, existe  $C, \varepsilon > 0$  e  $\theta \in [0, 1)$  tal que

$$|f(x) - f(\hat{x})|^\theta \leq C\|x^*\|, \quad x \in B(\hat{x}, \varepsilon), \quad x^* \in \partial f(x).$$

Assumindo que  $f(\hat{x}) = 0$  e levando em conta que se  $f \in C^1$  então,

$$\partial f(x) = \{\nabla f(x)\},$$

(3-7) se reduz à seguinte desigualdade

$$\|\nabla f(x)\| \geq \overline{C}|f(x)|^\theta, \quad x \in B(0, \varepsilon).$$

Podemos pensar na relação entre as funções que satisfazem a condição de Łojasiewicz e as funções que são convexas. Se essas condições fossem equivalentes, isto é, se o conjunto das funções que são convexas fosse igual ao conjunto das funções que tem a propriedade de Łojasiewicz, então não estaria sendo ampliado a classe de funções sobre as quais o método do proximal gera uma sequência convergente. Contudo, podemos ver em Łojasiewicz [7] que funções reais analíticas satisfazem a propriedade de Łojasiewicz. Assim, existem diversas funções não convexas que satisfazem a desigualdade de Łojasiewicz; por exemplo, funções da forma  $f(x) = x^p$ ,  $p$  ímpar. Em Kurdyka [6, capítulo 4, seção 4.3] é feita a construção de uma função  $C^2$  convexa que não satisfaz a propriedade de Łojasiewicz.

Este trabalho está organizado como segue. No Capítulo 1, apresentamos uma revisão de alguns tópicos de análise convexa e análise não diferenciável. No Capítulo 2, estudamos o Método do Ponto Proximal no caso que a função objetivo é convexa e provamos sua convergência usando o conceito de Féjer Convergência. Finalmente, no Capítulo 3, apresentamos uma prova de convergência da sequência gerada pelo Método do Ponto Proximal para funções não diferenciáveis que satisfazem a desigualdade de Łojasiewicz.

## Preliminares

Neste capítulo apresentaremos algumas definições e resultados básicos de análise convexa alguns dos quais podem ser encontrados, por exemplo, em [5, 9]. Além disso, apresentaremos alguns elementos de análise não diferenciável que podem ser encontrados, por exemplo, em [10].

Denotamos o produto interno e a norma (proveniente do produto interno) em  $\mathbb{R}^n$ , respectivamente, por  $\langle \cdot, \cdot \rangle$  e  $\| \cdot \|$ .

### 1.1 Tópicos de Análise Convexa

Nesta seção, introduziremos algumas propriedades fundamentais de análise convexa que serão úteis posteriormente neste trabalho. Estes resultados podem ser encontrados em [5, 9, 12].

Salvo menção explícita em contrário, ao longo desta seção,  $f$  denotará uma função  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ .

**Definição 1.1** Um conjunto  $D \subset \mathbb{R}^n$  é um conjunto convexo se, para quaisquer  $x, y \in D$  e  $\alpha \in [0, 1]$ , tem-se  $\alpha x + (1 - \alpha)y \in D$ .

**Definição 1.2** Dados  $x^i \in D$ ,  $\alpha_i \in [0, 1]$ ,  $i = 1, \dots, p$ , tais que  $\sum_{i=1}^p \alpha_i = 1$ , o ponto  $\sum_{i=1}^p \alpha_i x^i$  é chamado a combinação convexa dos pontos  $x^i \in D$ .

**Definição 1.3** Seja  $D \subset \mathbb{R}^n$  um conjunto convexo. Uma função  $f : D \rightarrow \mathbb{R}$  é dita ser convexa em  $D$  quando, para quaisquer  $x, y \in D$  e  $\alpha \in [0, 1]$ , tem-se

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

Se a desigualdade estrita ocorre para  $x \neq y$  e  $\alpha \in (0, 1)$ , então  $f$  é dita ser estritamente convexa.

**Proposição 1.4** Se  $f$  é uma função convexa e  $h : \mathbb{R}^n \rightarrow \mathbb{R}$  é uma função estritamente convexa, então,  $f + h : \mathbb{R}^n \rightarrow \mathbb{R}$  é uma função estritamente convexa.

*Prova* Segue imediatamente da última definição. □

**Proposição 1.5** *Seja  $f$  uma função estritamente convexa. Se  $f$  possui um minimizador, ele é único.*

*Prova* Suponhamos que  $f$  admita dois minimizadores  $x, x' \in \mathbb{R}^n$ ,  $x \neq x'$ , com  $f(x) = f(x') = \bar{v}$ . Visto que  $f$  é estritamente convexa, para  $\alpha \in (0, 1)$ , temos que

$$f(\alpha x + (1 - \alpha)x') < \alpha f(x) + (1 - \alpha)f(x') = \bar{v},$$

que é um absurdo. Logo, o resultado segue. □

**Definição 1.6** *O epígrafo de  $f$  é o conjunto não vazio*

$$\text{epi } f = \{(x, r) \in \mathbb{R}^n \times \mathbb{R}; r \geq f(x)\}.$$

O próximo teorema é um resultado clássico que será admitido sem demonstração, mas que será de grande utilidade nesta seção. A prova pode ser encontrada em [5, pág.76].

**Teorema 1.7** *Uma função  $f$  é convexa em  $\mathbb{R}^n$  se, e somente se, o epígrafo de  $f$  é um conjunto convexo em  $\mathbb{R}^n \times \mathbb{R}$ .*

**Definição 1.8** *Seja  $D \subset \mathbb{R}^n$  um conjunto qualquer. A envoltória convexa de  $D$ , denotado por  $\text{conv } D$ , é o conjunto de todas as combinações convexas de pontos de  $D$ .*

**Teorema 1.9** *(Desigualdade de Jensen) Se  $f$  é uma função convexa, então para todo  $m \in \mathbb{N}$ ,  $x_i \in \mathbb{R}^n$  e  $\lambda_i \geq 0$ ,  $i = 1, \dots, m$ , tais que  $\sum_{i=1}^m \lambda_i = 1$ , tem-se*

$$f\left(\sum_{i=1}^m \lambda_i x_i\right) \leq \sum_{i=1}^m \lambda_i f(x_i). \quad (1-1)$$

*Prova* Note que se  $i = 1$ , então  $\lambda_1 = 1$  e a desigualdade (1-1) ocorre trivialmente. Suponhamos agora que (1-1) ocorre para  $m = k$ ,  $k \in \mathbb{N}$ . Provaremos que (1-1) ocorre para  $m = k + 1$ . Note que se  $\lambda_{k+1} = 1$ , então  $\lambda_i = 0$ ,  $i = 1, \dots, k$  e, neste caso é imediato que (1-1) com  $m = k + 1$ . Suponhamos  $\lambda_{k+1} < 1$ . Assim, temos

$$f\left(\sum_{i=1}^{k+1} \lambda_i x_i\right) = f\left((1 - \lambda_{k+1})\left(\sum_{i=1}^k \frac{\lambda_i x_i}{1 - \lambda_{k+1}}\right) + \lambda_{k+1} x^{k+1}\right).$$

Visto que  $f$  é convexa e,  $\sum_{i=1}^{k+1} \lambda_i x^i = \sum_{i=1}^k \frac{\lambda_i}{1-\lambda_{k+1}} x^i + \lambda_{k+1} x^{k+1} = 1$ , segue que

$$f\left(\sum_{i=1}^{k+1} \lambda_i x^i\right) \leq (1-\lambda_{k+1})f\left(\sum_{i=1}^k \frac{\lambda_i}{1-\lambda_{k+1}} x^i\right) + \lambda_{k+1}f(x^{k+1}). \quad (1-2)$$

Agora, como  $\frac{\lambda_i}{1-\lambda_{k+1}} \geq 0$  e  $\sum_{i=1}^k \frac{\lambda_i}{1-\lambda_{k+1}} = 1$ , por hipótese de indução, obtemos que

$$f\left(\sum_{i=1}^k \frac{\lambda_i x^i}{1-\lambda_{k+1}}\right) \leq \sum_{i=1}^k \frac{\lambda_i}{1-\lambda_{k+1}} f(x^i).$$

Logo, combinando a última desigualdade com (1-2), segue o resultado desejado.  $\square$

**Definição 1.10** Dizemos que  $f$  é localmente Lipschitz em  $z \in \mathbb{R}^n$ , com constante  $K$ , se existe algum  $\varepsilon > 0$  tal que

$$|f(x) - f(y)| \leq K\|x - y\|, \quad x, y \in B(z, \varepsilon).$$

**Teorema 1.11** Se  $f$  uma função convexa, então  $f$  é localmente Lipschitz.

*Prova* Tome  $u \in \mathbb{R}^n$  arbitrário (fixado). Afirmamos que  $f$  é limitada em uma vizinhança de  $u$ . De fato, seja  $\varepsilon > 0$  e considere a caixa

$$S := \{y \in \mathbb{R}^n : -\varepsilon \leq y_i - u_i \leq \varepsilon, i = 1, \dots, n\},$$

com vértices  $u^1, \dots, u^m$ . É conhecido que  $m = 2^n$  e  $S = \text{conv}(\{u^1, \dots, u^m\})$  (envoltória convexa de  $\{u^1, \dots, u^m\}$ ), ver [5, pág.129]. Então, para cada  $y \in S$ , existem  $\lambda_i \geq 0$ ,  $i = 1, \dots, m$ , com  $\sum_{i=1}^m \lambda_i = 1$ , tal que  $y = \sum_{i=1}^m \lambda_i u_i$ . Denote por  $M$  o máximo de  $f$  em  $\{u^1, \dots, u^m\}$ . Visto que  $f$  é convexa e  $\sum_{i=1}^m \lambda_i = 1$ , do Teorema 1.9 segue que

$$f(y) = f\left(\sum_{i=1}^m \lambda_i u_i\right) \leq \sum_{i=1}^m \lambda_i f(u_i) \leq M \sum_{i=1}^m \lambda_i = M,$$

o que prova que  $f$  é limitada superiormente por  $M$  quando restrita ao conjunto  $S$ . Por outro lado, para qualquer  $y \in S$ , existe  $x \in S$  tal que  $u = \frac{1}{2}x + \frac{1}{2}y$ . Assim, pela convexidade de  $f$ , temos que  $f(u) \leq \frac{1}{2}f(x) + \frac{1}{2}f(y)$ , de onde obtemos que

$$f(y) \geq 2f(u) - f(x).$$

Agora, visto que  $x \in S$  e  $f$  é limitada superiormente por  $M$  quando restrita ao conjunto  $S$ , da última desigualdade podemos concluir que  $f$  é limitada inferiormente por  $2f(u) - M$  quando restrita ao conjunto  $S$  e a afirmação está provada.

Consideremos a bola  $B(u, 2\delta)$  onde  $f$  é limitada e sejam  $\alpha$  e  $\beta$  tais que

$$\alpha \leq f(z) \leq \beta, \quad z \in B(u, 2\delta). \quad (1-3)$$

Dados  $y, y' \in B(u, \delta)$ ,  $y \neq y'$  é fácil ver que

$$y'' = y' + \delta \frac{y' - y}{\|y' - y\|} \in B(u, 2\delta).$$

Além disso,  $y' = \frac{\|y' - y\|}{\delta + \|y' - y\|} y'' + \frac{\delta}{\delta + \|y' - y\|} y$ . Assim, visto que  $f$  é convexa, obtemos

$$f(y') \leq \frac{\|y' - y\|}{\delta + \|y' - y\|} f(y'') + \frac{\delta}{\delta + \|y' - y\|} f(y).$$

Subtraindo  $f(y)$  de ambos os lados da última desigualdade, segue que

$$f(y') - f(y) \leq \frac{\|y' - y\|}{\delta + \|y' - y\|} [f(y'') - f(y)]. \quad (1-4)$$

Agora, visto que  $y, y' \in B(u, 2\delta)$ , de (1-3), temos que  $f(y'') \leq \beta$  e  $-f(y) \leq -\alpha$  e, consequentemente,

$$\frac{\|y' - y\|}{\delta + \|y' - y\|} [f(y'') - f(y)] \leq \frac{\|y' - y\|}{\delta} (\beta - \alpha).$$

Assim, combinando última desigualdade com (1-4), concluímos que

$$f(y') - f(y) \leq \frac{\|y' - y\|}{\delta} (\beta - \alpha).$$

Procedendo de modo análogo, é possível mostrar que

$$f(y) - f(y') \leq \frac{\|y' - y\|}{\delta} (\beta - \alpha).$$

Portanto,

$$|f(y') - f(y)| \leq \frac{\beta - \alpha}{\delta} \|y' - y\|,$$

e o resultado segue. □

O Próximo resultado é uma consequência imediata do Teorema 1.11.

**Corolário 1.12** *Se  $f$  é uma função convexa então  $f$  é contínua.*

Seja  $C \subset \mathbb{R}^n$  um conjunto não vazio e  $x \in \mathbb{R}^n$ . Uma projeção de  $x$  sobre  $C$ , é uma solução do problema de minimização

$$\min_{y \in C} \|y - x\|. \quad (1-5)$$

Afirmamos que o problema 1-5 sempre tem solução quando o conjunto  $C$  é fechado. De fato, fixemos  $y \in \mathbb{R}^n$  arbitrário e consideremos a função

$$f : \mathbb{R}^n \rightarrow \mathbb{R}_+, \quad f(x) = \|x - y\|.$$

Note que,  $f$  é contínua em  $\mathbb{R}^n$  e

$$C \cap B[y, \alpha] = \{x \in C : f(x) \leq \alpha\}, \quad (1-6)$$

é compacto (pois  $C$  é fechado e  $B[y, \alpha] = \{x \in \mathbb{R}^n; f(x) \leq \alpha\}$  é compacto). Além disso, tomando  $\alpha$  suficiente grande, se necessário, podemos admitir que  $C \cap B[y, \alpha] \neq \emptyset$ . Portanto, pelo Teorema de Weierstrass  $f$  restrita a  $C \cap B[y, \alpha]$  possui um minimizador  $x^*$ . Tome  $x \in C \setminus C \cap B[y, \alpha]$ . Note que  $f(x) > \alpha \geq f(x^*)$ . Mas isto nos diz que  $x^*$  também é solução de 1-5.

**Teorema 1.13 (Teorema da Projeção)** *Seja  $C \subset \mathbb{R}^n$  um conjunto não vazio, convexo e fechado. Então, para todo  $x \in \mathbb{R}^n$ , a projeção de  $x$  sobre  $C$ , denotada por  $P_C(x)$ , existe e é única. Além disso,  $\bar{x} = P_C(x)$  se, e somente se,*

$$\bar{x} \in C, \quad \langle x - \bar{x}, y - \bar{x} \rangle \leq 0 \quad y \in C. \quad (1-7)$$

*Prova* Tome  $x \in \mathbb{R}^n$  fixado e seja  $\bar{x}$  uma solução do problema  $\min_{y \in C} \|y - x\|$ . Como  $\bar{x} \in C$  e  $C$  é convexo,

$$(1 - \alpha)\bar{x} + \alpha y = x(\alpha) \in C, \quad y \in C \quad \alpha \in (0, 1].$$

Assim, segue que  $\|x - \bar{x}\| \leq \|x - x(\alpha)\|$  e, conseqüentemente,

$$0 \geq \|x - \bar{x}\|^2 - \|x - x(\alpha)\|^2, \quad \alpha \in (0, 1].$$

Da última desigualdade, temos que

$$0 \geq 2\langle x - \bar{x}, x(\alpha) - \bar{x} \rangle - \|x(\alpha) - \bar{x}\|^2, \quad \alpha \in (0, 1].$$



e, a partir da definição de  $x(\alpha)$ , obtemos:

$$0 \geq 2\alpha \langle x - \bar{x}, y - \bar{x} \rangle - \alpha^2 \|y - \bar{x}\|^2, \quad \alpha \in (0, 1].$$

Dividindo os dois lados da última desigualdade acima por  $2\alpha > 0$  e tomando o limite quando  $\alpha \rightarrow 0^+$ , obtemos (1-7).

Suponhamos agora que um certo  $\bar{x}$  satisfaz (1-7). Visto que

$$\langle x - \bar{x}, y - \bar{x} \rangle = \frac{1}{2} \left( \|x - \bar{x}\|^2 + \|y - \bar{x}\|^2 - \|x - y\|^2 \right),$$

de (1-7), obtemos

$$\frac{1}{2} \left( \|x - \bar{x}\|^2 - \|x - y\|^2 \right) \leq 0, \quad y \in C.$$

Assim,

$$\|x - \bar{x}\| \leq \|x - y\|, \quad y \in C.$$

Portanto,  $\bar{x}$  é uma projeção de  $x$  sobre  $C$ . Provaremos agora que a projeção é única. Seja  $\hat{x}$  uma outra possível solução do problema  $\min_{y \in C} \|y - x\|$ . Usando (1-7) para  $\bar{x}$  com  $y = \hat{x} \in C$  e para  $\hat{x}$  com  $y = \bar{x} \in C$ , temos

$$\langle x - \bar{x}, \hat{x} - \bar{x} \rangle \leq 0, \quad \langle x - \hat{x}, \bar{x} - \hat{x} \rangle \leq 0.$$

Somando as duas últimas desigualdades acima, adquirimos

$$0 \geq \langle x - \bar{x} - (x - \hat{x}), \hat{x} - \bar{x} \rangle = \|\hat{x} - \bar{x}\|^2,$$

de onde segue que  $\hat{x} = \bar{x}$ .

□

**Teorema 1.14** *Seja  $C \subset \mathbb{R}^n$  um conjunto não vazio, convexo e fechado, e  $x \notin C$ . Então, existe  $v \in \mathbb{R}^n$  tal que*

$$\langle v, x \rangle > \sup\{\langle v, y \rangle : y \in C\}. \quad (1-8)$$

*Prova* Defina  $v = x - P_C(x) \neq 0$ . Dessa forma, temos

$$\langle v, y - x + v \rangle = \langle v, y \rangle - \langle v, x \rangle + \|v\|^2, \quad y \in C.$$

Combinando a última igualdade acima com (1-7) segue que

$$\langle v, y \rangle + \|v\|^2 \leq \langle v, x \rangle, \quad y \in C,$$

e (1-8) ocorre para  $v = x - P_C(x)$ .

□

**Definição 1.15 (Hiperplano Suporte)** Dado um conjunto  $C \subset \mathbb{R}^n$  com fronteira  $\partial C \neq \emptyset$ , considere o hiperplano  $H_{v,r} = \{y \in \mathbb{R}^n; \langle v, y \rangle = \langle v, x_0 \rangle\}$  para  $v \in \mathbb{R}^n \setminus \{0\}$  e  $x_0 \in \partial C$  ambos fixados. Dizemos que  $H_{v,r}$  é suporte ao conjunto  $C$  em  $x_0$  se vale a desigualdade

$$\langle v, x \rangle \leq \langle v, x_0 \rangle, \quad x \in C.$$

**Lema 1.16** Seja  $C \subset \mathbb{R}^n$  um conjunto não vazio, convexo e fechado. Então, dado  $x \in \partial C$ , existe um hiperplano suporte a  $C$  em  $x$ .

*Prova* Como  $C$ , o fecho de  $C$  e os complementos de  $C$  e do fecho de  $C$  possuem a mesma fronteira, uma sequência pode ser encontrada tal que  $x^k$  não pertença ao fecho de  $C$ , para  $k \in \mathbb{N}$  e  $\lim_{k \rightarrow \infty} x^k = x$ . Para cada  $k$  temos, pelo Teorema 1.14, que existe algum  $s^k$  com  $\|s^k\| = 1$ , tal que

$$\langle s^k, x^k - y \rangle > 0, \quad y \in C. \quad (1-9)$$

Visto que  $\{s^k\}$  é limitada, podemos assumir, sem perda de generalidade, que  $\{s^k\}$  converge a  $s \neq 0$ . Tomando o limite em (1-9) com  $k \rightarrow +\infty$  de onde obtemos que

$$\langle s, y - x \rangle \geq 0,$$

$H_{s,x} = \{y \in \mathbb{R}^n : \langle s, y \rangle = \langle s, x \rangle\}$  é suporte ao conjunto  $C$  no ponto  $x$ .

□

**Definição 1.17** Seja  $f$  uma função convexa e  $\xi$  um vetor em  $\mathbb{R}^n$ . Dizemos que  $\xi$  é um subgradiente de  $f$  em  $x$  se

$$f(y) \geq f(x) + \langle \xi, y - x \rangle, \quad y \in \mathbb{R}^n.$$

O conjunto dos subgradientes de  $f$  em  $x$  é chamado subdiferencial de  $f$  em  $x$ , e é denotado por

$$\partial_c f(x) = \{\xi \in \mathbb{R}^n : \xi \text{ é subgradiente de } f \text{ em } x\}.$$

**Proposição 1.18** Seja  $f$  uma função convexa e  $x \in \mathbb{R}^n$ . Então,  $x \in \mathbb{R}^n$  é minimizador de  $f$  se, e somente se,  $0 \in \partial_c f(x)$ .

*Prova* A prova segue imediatamente da definição de minimizador irrestrito combinada com a definição de subdiferencial de  $f$  em  $x$ .

□

Vejamos agora alguns exemplos.

**Exemplo 1.19** Considere a função real  $f(x) = |x|$ . Decorre diretamente da definição de subdiferencial para funções convexas que  $\partial_c f(0) = [-1, 1]$ .

**Teorema 1.20** Seja  $x \in \mathbb{R}^n$  um ponto onde  $f$  é diferenciável. Se  $f$  é uma função convexa, então  $\partial_c f(x) = \{\nabla f(x)\}$ .

*Prova* Visto que  $f$  é convexa e diferenciável em  $x \in \mathbb{R}^n$ , temos

$$f(z) \geq f(x) + \langle \nabla f(x), z - x \rangle, \quad z \in \mathbb{R}^n.$$

Portanto,  $\nabla f(x) \in \partial_c f(x)$ . Tome  $v \in \partial_c f(x)$ . Assim, para todo  $h \in \mathbb{R}^n$

$$f(x) + \langle v, h \rangle \leq f(x + h).$$

A desigualdade acima combinada com a diferenciabilidade da função  $f$  no ponto  $x$  nos mostra que

$$\langle v - \nabla f(x), h \rangle \leq o(|h|).$$

Em particular, temos que

$$\langle v - \nabla f(x), h^k \rangle \leq o(|h^k|) \tag{1-10}$$

para qualquer sequência  $\{h^k\}$  tal que  $h^k \rightarrow 0$ . Considerando a sequência  $\{h^k\}$ , com  $h^k = \frac{v - \nabla f(x)}{k \|v - \nabla f(x)\|}$ , de (1-10) obtemos que:

$$\frac{\|v - \nabla f(x)\|}{k} \leq o(1/k),$$

de onde segue o resultado desejado.

□

Quando temos duas funções  $f$  e  $g$  diferenciáveis, então

$$\nabla(f + g)(\cdot) = \nabla f(\cdot) + \nabla g(\cdot).$$

Agora provaremos que esta propriedade também vale no caso dos subdiferenciais para funções convexas, isto é,  $\partial_c(f + g)(\cdot) = \partial_c f(\cdot) + \partial_c g(\cdot)$ . Para provar este resultado estudaremos uma série de fatos que seguem abaixo.

**Definição 1.21** Dizemos que uma função  $f$  é positivamente homogênea se

$$f(\lambda x) = \lambda f(x), \quad \lambda \in \mathbb{R}, \quad x \in \mathbb{R}^n.$$

**Definição 1.22** Dizemos que uma função  $f$  é subaditiva se

$$f(x+y) \leq f(x) + f(y), \quad x, y \in \mathbb{R}^n.$$

**Definição 1.23** Sejam  $x$  e  $d$  um ponto e uma direção em  $\mathbb{R}^n$ , respectivamente. A derivada direcional de uma função  $f$  em  $x$  na direção  $v$  é definida por

$$f'(x, d) := \lim_{t \rightarrow 0^+} \frac{f(x+td) - f(x)}{t}.$$

**Proposição 1.24** Seja  $f$  uma função convexa e  $x \in \mathbb{R}^n$ . Então,

i) a derivada direcional em cada direção  $v \in \mathbb{R}^n$  existe e satisfaz

$$f'(x, v) := \inf_{t \rightarrow 0^+} \frac{f(x+tv) - f(x)}{t};$$

ii) a função  $v \mapsto f'(x; v)$  é positivamente homogênea e subaditiva em  $\mathbb{R}^n$  e, além disso, se  $K$  é a constante de Lipschitz para  $f$  em  $x$ ,

$$|f'(x; v)| \leq K\|v\|, \quad x \in \mathbb{R}^n.$$

*Prova (i)* Tome  $v \in \mathbb{R}^n$  e definamos  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  dada por

$$\varphi(t) = \frac{f(x+tv) - f(x)}{t}.$$

Seja  $\epsilon > 0$  e tome constantes  $t_1, t_2$  tais que  $0 < t_1 < t_2 < \epsilon$ . Note que

$$\varphi(t_2) - \varphi(t_1) = \frac{(t_2 - t_1)f(x) + t_1f(x + t_2v) - t_2f(x + t_1v)}{t_1t_2}.$$

que, depois de algumas manipulações algébricas, fornece

$$\varphi(t_2) - \varphi(t_1) = \frac{1}{t_1} [(1 - t_1/t_2)f(x) + t_1/t_2 f(x + t_2v) - f(x + t_1v)].$$

Agora, visto que  $0 < t_1/t_2 < 1$  e  $f$  é convexa, da última igualdade, obtemos

$$\varphi(t_2) - \varphi(t_1) \geq \frac{1}{t_1} [f((1 - t_1/t_2)x + t_1/t_2(x + t_2v)) - f(x + t_1v)] \geq 0,$$

de onde concluímos que  $\varphi$  decresce quando  $t$  tende a zero por valores positivos.

Considere agora  $0 < t < \epsilon$ . Temos que

$$\varphi(t) - \varphi(-\epsilon/2) = \frac{\frac{1}{2}f(x+tv) + \frac{1}{2}f(x) + \frac{t}{\epsilon}f(x - \frac{\epsilon}{2}v) + (1 - \frac{t}{\epsilon})f(x) - 2f(x)}{t/2}.$$

Usando que  $f$  é convexa, a última igualdade nos fornece

$$\varphi(t) - \varphi(-\epsilon/2) \geq \frac{f(x+t/2v) + f(x-t/2v) - 2f(x)}{t/2},$$

de onde obtemos

$$\varphi(t) - \varphi(-\epsilon/2) \geq \frac{1/2f(x+t/2v) + 1/2f(x-t/2v) - f(x)}{t/4}.$$

Usando novamente a convexidade de  $f$  no segundo membro da última desigualdade, temos

$$\varphi(t) - \varphi(-\epsilon/2) \geq \frac{f(x) - f(x)}{t/4} = 0,$$

o que implica que  $f$  é limitada inferiormente para  $0 < t < \epsilon$  e o resultado da proposição segue.

(ii) Seja  $\lambda > 0$ . Então,

$$f'(x; \lambda v) := \lim_{t \rightarrow 0^+} \frac{f(x + (t\lambda)v) - f(x)}{t}, \quad v \in \mathbb{R}^n.$$

Dessa forma,

$$f'(x; \lambda v) = \lambda \lim_{t \rightarrow 0^+} \frac{f(x + (t\lambda)v) - f(x)}{\lambda t}, \quad v \in \mathbb{R}^n.$$

Mas, da última igualdade, segue que

$$f'(x, \lambda v) = \lambda f'(x; v), \quad t \in \mathbb{R}, \quad v \in \mathbb{R}^n.$$

Portanto,  $f'(x; \cdot)$  é positivamente homogênea. Provaremos agora que  $f'$  é subaditiva. Sejam  $v, w \in \mathbb{R}^n$ . Assim,

$$f'(x; v + w) := \lim_{t \rightarrow 0^+} \frac{f(x + t(v + w)) - f(x)}{t}. \quad (1-11)$$

Note que, pela convexidade da função  $f$ , temos

$$f\left(\frac{1}{2}(x + 2tv) + \frac{1}{2}(x + 2tw)\right) \leq \frac{1}{2}f(x + 2tv) + \frac{1}{2}f(x + 2tw), \quad t \in \mathbb{R}. \quad (1-12)$$

Dessa forma, segue de (1-12), que

$$\lim_{t \rightarrow 0^+} \frac{f\left(\frac{1}{2}(x + 2tv) + \frac{1}{2}(x + 2tw)\right) - f(x)}{t} \leq \lim_{t \rightarrow 0^+} \frac{f(x + 2tv) - f(x)}{2t} + \lim_{t \rightarrow 0^+} \frac{f(x + 2tw) - f(x)}{2t}.$$

Agora, levando em conta que  $\frac{1}{2}(x + 2tv) + \frac{1}{2}(x + 2tw) = x + t(v + w)$ , segue, de (1-11) combinada com a última desigualdade, que

$$f'(x; v + w) \leq f'(x; v) + f'(x; w), \quad x, v, w \in \mathbb{R}^n.$$

Pelo item (i) temos que

$$f'(x, v) = \lim_{t \rightarrow 0^+} \frac{f(x + tv) - f(x)}{t}.$$

Assim, temos

$$|f'(x, v)| = \lim_{t \rightarrow 0^+} \frac{|f(x + tv) - f(x)|}{t}. \quad (1-13)$$

Note que, pela condição de Lipschitz temos

$$\frac{|f(x + tv) - f(x)|}{t} \leq \frac{K\|x + tv - x\|}{t}, \quad t \in \mathbb{R}, \quad x \in \mathbb{R}^n. \quad (1-14)$$

Agora, de (1-13) combinada com (1-14), segue que

$$f'(x; v) \leq \lim_{t \rightarrow 0^+} \frac{K\|x + tv - x\|}{t}, \quad t \in \mathbb{R}.$$

Portanto, da última desigualdade, temos

$$f'(x; v) \leq K\|v\|, \quad v \in \mathbb{R}^n.$$

□

**Exemplo 1.25** Assuma que  $f : [a, b] \rightarrow \mathbb{R}$  é uma função convexa. Dessa forma em um ponto  $c \in [a, b]$ , temos  $\partial_c f(c) = [f'_-(c), f'_+(c)]$ , onde  $f'_-(c)$  e  $f'_+(c)$ , são as derivadas laterais da função  $f$  no ponto  $c$ .

**Definição 1.26** Seja  $S$  um conjunto não vazio em  $\mathbb{R}^n$ . A função  $\sigma_S : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ , definida por

$$\sigma_S(x) := \sup \{ \langle s, x \rangle : s \in S \},$$

é chamada função suporte de  $S$ .

**Proposição 1.27** Sejam  $A, B \subset \mathbb{R}^n$  conjuntos não vazios, convexos e compactos. Então,

$$A \subset B \quad \text{se, e somente se,} \quad \sigma_A \leq \sigma_B.$$

*Prova* Veja, por exemplo, [2].

□

Usaremos a próxima definição e o Teorema de Hahn-Banach para provar a parte (i) do Teorema 1.30.

**Definição 1.28** Dizemos que  $f$  é uma função positivamente homogênea se  $f(\lambda x) = \lambda f(x)$  para todo  $\lambda \geq 0$  e subaditiva se  $f(x + y) \leq f(x) + f(y)$  para todo  $x$  e  $y$  em  $\mathbb{R}^n$ .

**Teorema 1.29 (Hahn-Banach)** Seja  $X$  um espaço vetorial real e  $Y \subset X$  um subespaço linear. Se  $p : X \rightarrow \mathbb{R}$  é um funcional subaditivo e positivamente homogêneo e  $f : Y \rightarrow \mathbb{R}$  um funcional linear tal que  $f(y) \leq p(y)$  para todo  $y \in Y$ , então existe um funcional linear  $F : X \rightarrow \mathbb{R}$  tal que

$$F(y) = f(y), \quad y \in Y$$

e

$$F(x) \leq p(x), \quad x \in X.$$

*Prova* Veja, por exemplo, [3].

□

**Teorema 1.30** Seja  $f$  uma função convexa e  $x \in \mathbb{R}^n$ . Então,

- i)  $\partial_c f(x)$  é um conjunto não vazio, convexo e compacto tal que  $\partial_c f(x) \subset B(0; L)$ , onde  $L$  é uma constante de Lipschitz de  $f$  em  $x$ ;
- ii)  $f'(x; v) = \max\{\xi \in \mathbb{R}^n : \xi \in \partial_c f(x)\}$ , para todo  $v \in \mathbb{R}^n$ .
- iii)  $\partial_c f(x) = \{\xi \in \mathbb{R}^n : f'(x; v) \geq \langle \xi, v \rangle, \quad v \in \mathbb{R}^n\}$ .

*Prova (i)* Visto que  $f$  é convexa, Teorema 1.7 nos assegura que  $\text{epi } f = \{(x, r) \in \mathbb{R}^n \times \mathbb{R} : r \geq f(x)\}$  é um subconjunto convexo de  $\mathbb{R}^n \times \mathbb{R}$ . Além disso, pelo Corolário 1.12, temos que  $f$  é contínua, o que nos permite concluir que  $\text{epi } f$  é também fechado. Dado  $x \in \mathbb{R}^n$  e  $(x, f(x)) \in \partial(\text{epi } f)$ , podemos tomar, pelo Lema 1.16, um hiperplano suporte ao  $\text{epi } f$  em  $(x, f(x))$ . Mas isto nos diz que, existe  $(v, \alpha) \in \mathbb{R}^n \times \mathbb{R}$  não nulo tal que

$$\langle v, y \rangle + \alpha r \leq \langle v, x \rangle + \alpha f(x), \quad (y, r) \in \text{epi } f. \quad (1-15)$$

Suponhamos que  $\alpha = 0$ . Assim, da última desigualdade, temos que

$$\langle v, y - x \rangle \leq 0, \quad y \in \mathbb{R}^n,$$

de onde segue que  $v = 0$ . Mas isto contradiz o fato de  $(v, \alpha) \neq 0$ . Portanto, temos  $\alpha \neq 0$ . Afirmamos que  $\alpha < 0$ . De fato, tome  $\delta > 0$  e note que  $(x, f(x) + \delta) \in \text{epi } f$ . Visto que a

desigualdade em (1-15) ocorre para todo  $(y, r) \in \text{epi } f$ , em particular, temos

$$\alpha(f(x) + \delta) \leq \alpha f(x).$$

Assim, se  $\alpha > 0$ , então  $\delta \leq 0$ , o que é uma contradição. Isto nos permite concluir que  $\alpha < 0$ .

Dividindo ambos os membros de (1-15) por  $\alpha$ , obtemos:

$$r \geq f(x) + \left\langle -\frac{v}{\alpha}, y - x \right\rangle, \quad (y, r) \in \text{epi } f.$$

Em particular, como  $(y, f(y)) \in \text{epi } f$ , segue que

$$f(y) \geq f(x) + \left\langle -\frac{v}{\alpha}, y - x \right\rangle, \quad y \in \mathbb{R}^n,$$

e segue da definição de subdiferencial que  $\partial_c f(x)$  é não vazio.

Vamos agora provar a convexidade do conjunto  $\partial_c f(x)$ . Tome  $z, w \in \partial_c f(x)$  e  $t \in [0, 1]$ . Então,

$$(1-t)[f(y) - f(x) - \langle w, y - x \rangle] + t[f(y) - f(x) - \langle z, y - x \rangle] \geq 0, \quad y \in \mathbb{R}^n.$$

Assim,

$$f(y) - f(x) - \langle tz + (1-t)w, y - x \rangle \geq 0, \quad y \in \mathbb{R}^n,$$

donde segue que  $tz + (1-t)w \in \partial_c f(x)$ . Logo, da arbitrariedade de  $z, w \in \partial_c f(x)$  e  $t \in [0, 1]$ , concluimos que  $\partial_c f(x)$  é um conjunto convexo.

Para provar que  $\partial_c f(x)$  é fechado, tome  $\{s^k\} \subset \partial_c f(x)$  com  $s^k \rightarrow s$ . Dessa forma, segue que

$$f(y) \geq f(x) + \langle s^k, y - x \rangle, \quad y \in \mathbb{R}^n.$$

Visto que  $s^k \rightarrow s$ , tomando o limite na última desigualdade, obtemos

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

Mas isto nos diz que  $s \in \partial_c f(x)$ . Portanto,  $\partial_c f(x)$  é fechado.

Tomando  $\xi \in \partial_c f(x)$ , pelo item ii), temos

$$\|\xi\|^2 = \langle \xi, \xi \rangle \leq |f'(x; \xi)|.$$

Mas última desigualdade combinada com ii) da Proposição 1.24 segue que

$$\|\xi\|^2 \leq K\|\xi\|, \quad \xi \in \partial_c f(x).$$



Logo,

$$\|\xi\| \leq K, \quad \xi \in \partial_c f(x).$$

o que completa a prova do item *iii*).

(ii) Pela definição de subdiferencial, deduzimos que para cada  $v \in \mathbb{R}^n$

$$\varphi(t) = \frac{f(x+tv) - f(x)}{t} \geq \frac{\langle \xi, tv \rangle}{t}, \quad \xi \in \partial_c f(x). \quad (1-16)$$

Então,

$$f'(x;v) \geq \max\{\langle \xi, v \rangle; \xi \in \partial_c f(x)\}.$$

Por outro lado, se existe  $v_1 \in \mathbb{R}^n$  tal que  $f'(x, v_1) > \max\{\langle \xi, v_1 \rangle; \xi \in \partial_c f(x)\}$ , pelo Teorema de Hahn-Banach temos que existe  $\xi_1 \in \mathbb{R}^n$  tal que

$$f'(x;v) \geq \langle \xi_1, v \rangle, \quad e \quad f'(x;v_1) = \langle \xi_1, v_1 \rangle, \quad v \in \mathbb{R}^n. \quad (1-17)$$

Combinando (1-16) com a Proposição 1.24 segue que

$$f(x+tv) - f(x) = t\varphi(t) \geq tf'(x;v), \quad t > 0, v \in \mathbb{R}^n.$$

Consequentemente, a partir da desigualdade em (1-17), obtemos

$$f(x+tv) - f(x) \geq \langle \xi_1, tv \rangle, \quad t > 0, v \in \mathbb{R}^n,$$

de onde segue que  $\xi_1 \in \partial_c f(x)$  e, em particular, que

$$f'(x;v_1) > \max\{\langle \xi, v_1 \rangle; \xi \in \partial_c f(x)\} \geq \langle \xi_1, v_1 \rangle = f'(x;v_1),$$

o que é um absurdo. Portanto,

$$f'(x;v) = \max\{\langle \xi, v \rangle; \xi \in \partial_c f(x)\},$$

e o item ii) está provado.

(iii) Sejam  $V = \{\xi \in \mathbb{R}^n : f'(x;v) \geq \langle \xi, v \rangle, \quad v \in \mathbb{R}^n\}$  e  $\xi \in V$  um elemento arbitrário. Então, segue do item *ii*), que

$$\langle \xi, y \rangle \leq f'(x; y), \quad y \in \mathbb{R}^n. \quad (1-18)$$

Por outro lado,

$$f'(x; y) = \lim_{t \rightarrow 0} \frac{f(x+ty) - f(x)}{t} = \lim_{t \rightarrow 0} \frac{f((1-t)x + t(x+y)) - f(x)}{t}, \quad t \leq 1.$$

Mas a última igualdade combinada com a convexidade da função  $f$  segue que

$$f'(x; y) \leq \lim_{t \rightarrow 0} \frac{(1-t)f(x) + tf(x+y) - f(x)}{t} = f(x+y) - f(x), \quad t \leq 1. \quad (1-19)$$

Combinando (1-18) com (1-19) segue que

$$\langle \xi, y \rangle \leq f(x+y) - f(x), \quad y \in \mathbb{R}^n.$$

Fazendo  $y = z - x$  na última desigualdade obtemos que

$$f(z) \geq f(x) + \langle \xi, z - x \rangle, \quad z \in \mathbb{R}^n.$$

Portanto, pela definição de  $\partial_c f(x)$ ,  $\xi \in \partial_c f(x)$ .

Agora, se  $\xi \in \partial_c f(x)$  temos

$$f(x+ty) - f(x) \geq \langle \xi, x+ty - x \rangle, \quad y \in \mathbb{R}^n, \quad t \leq 1.$$

Dividindo ambos os lados da desigualdade acima por  $t$  segue que

$$\frac{f(x+ty) - f(x)}{t} \geq \frac{\langle \xi, x+ty - x \rangle}{t} = \langle \xi, y \rangle, \quad y \in \mathbb{R}^n.$$

Tomando o limite com  $t \rightarrow 0$  da última desigualdade temos

$$f'(x; y) \geq \langle \xi, y \rangle, \quad y \in \mathbb{R}^n,$$

o que nos diz que  $\xi \in V$ . Assim,  $V = \partial_c f(x)$ .

□

**Corolário 1.31** Para qualquer  $v \in \mathbb{R}^n$ , temos

$$f'(x; v) = \sigma_{\partial_c f(x)}(v).$$

*Prova* Segue imediatamente da definição de função suporte combinada com o Teorema 1.30.

□

**Teorema 1.32** Sejam  $f, g$  funções convexas. Então, para cada  $x \in \mathbb{R}^n$ ,

$$\sigma_{\partial_c f(x) + \partial_c g(x)}(v) = \sigma_{\partial_c f(x)}(v) + \sigma_{\partial_c g(x)}(v), \quad v \in \mathbb{R}^n.$$

*Prova* Tome  $v \in \mathbb{R}^n$ . Visto que  $f$  e  $g$  são convexas, pelo item (iii) do Teorema 1.30 podemos concluir que existem  $\tilde{s}_1 \in \partial_c f(x)$  e  $\tilde{s}_2 \in \partial_c g(x)$  tal que

$$\sigma_{\partial_c f(x)}(v) = \langle \tilde{s}_1, v \rangle, \quad \sigma_{\partial_c g(x)}(v) = \langle \tilde{s}_2, v \rangle. \quad (1-20)$$

Além disso, da Definição 1.26 segue que

$$\sigma_{\partial_c f(x) + \partial_c g(x)}(v) \geq \langle s_1 + s_2, v \rangle, \quad s_1 \in \partial_c f(x), \quad s_2 \in \partial_c g(x).$$

Assim, de (1-20), obtemos que

$$\sigma_{\partial_c f(x) + \partial_c g(x)}(v) \geq \sigma_{\partial_c f(x)}(v) + \sigma_{\partial_c g(x)}(v). \quad (1-21)$$

Por outro lado,

$$\sigma_{\partial_c f(x)}(v) + \sigma_{\partial_c g(x)}(v) \geq \langle s_1 + s_2, v \rangle, \quad s_1 \in \partial_c f(x), \quad s_2 \in \partial_c g(x),$$

o que implica que

$$\sigma_{\partial_c f(x)}(v) + \sigma_{\partial_c g(x)}(v) \geq \sigma_{\partial_c f(x) + \partial_c g(x)}(v).$$

Portanto, o resultado segue da última desigualdade combinada com a desigualdade (1-21).  $\square$

**Teorema 1.33** *Sejam  $f, g$  funções convexas e  $x \in \mathbb{R}^n$ . Então,*

$$\partial_c(f + g)(x) = \partial_c f(x) + \partial_c g(x).$$

*Prova* Tome  $v \in \mathbb{R}^n$ . Visto que  $f$  e  $g$  são convexas, pelo Teorema 1.30, segue que  $\partial_c f(x)$  e  $\partial_c g(x)$  são conjuntos não vazios, convexos e compactos tais que

$$\sigma_{\partial_c f(x)}v = f'(x, v), \quad \sigma_{\partial_c g(x)}v = g'(x, v). \quad (1-22)$$

Em particular,  $\partial_c f(x) + \partial_c g(x)$  é um conjunto não vazio, convexo e compacto e, pelo Teorema 1.32,

$$\sigma_{\partial_c f(x) + \partial_c g(x)}v = \sigma_{\partial_c f(x)}v + \sigma_{\partial_c g(x)}v. \quad (1-23)$$

Por outro lado, visto que  $(f + g)'(x, v) = f'(x, v) + g'(x, v)$ , usando o item (i) do Teorema 1.30 junto com (1-22), obtemos

$$\sigma_{\partial_c(f+g)(x)}v = \sigma_{\partial_c f(x)}v + \sigma_{\partial_c g(x)}v.$$

Combinando última igualdade com (1-23) junto com a arbitrariedade de  $v$ , o resultado segue da Proposição 1.27.

□

## 1.2 Subdiferencial de Fréchet e Subdiferencial Limite

Na presente seção, estudaremos um pouco sobre o Subdiferencial limite e o Subdiferencial de Fréchet. Essas definições são interessantes pois permitem a caracterização dos pontos críticos, quando a função objetivo é não necessariamente convexa. O subdiferencial limite é empregado neste contexto, uma vez que, o mesmo possui gráfico fechado. E o fato do subdiferencial limite possuir gráfico fechado é fundamental neste trabalho. Além disso, temos provado também propriedades operacionais que são essenciais na prova do principal resultado do trabalho. Após a leitura dessa seção é possível concluir que o subdiferencial limite e o subdiferencial de Fréchet são extensões naturais do conceito de gradiente para funções diferenciáveis, veja os Teoremas 1.38 e 1.39.

**Definição 1.34** *Seja  $f$  uma função própria semicontínua inferior. O subdiferencial de Fréchet de  $f$  em  $x$ , denotado por  $\hat{\partial}f(x)$ , é definido como segue*

$$\hat{\partial}f(x) = \begin{cases} \{x^* \in \mathbb{R}^n : \liminf_{y \rightarrow x; y \neq x} \frac{1}{\|x - y\|} (f(y) - f(x) - \langle x^*, y - x \rangle) \geq 0\}, & \text{se } x \in \text{dom} f, \\ \emptyset, & \text{se } x \notin \text{dom} f. \end{cases}$$

Em outras palavras,  $x^* \in \hat{\partial}f(x)$  se, e somente se,

$$f(y) \geq f(x) + \langle x^*, y - x \rangle + o(\|y - x\|),$$

onde  $\frac{o(\|y - x\|)}{\|y - x\|} \rightarrow 0$  quando  $y \rightarrow x$ .

O subdiferencial limite de  $f$  em  $x \in \mathbb{R}^n$ , denotado por  $\partial f(x)$ , é definido como segue

$$\partial f(x) = \{x^* \in \mathbb{R}^n \mid \exists x_n \rightarrow x, f(x_n) \rightarrow f(x), x_n^* \in \hat{\partial}f(x_n); x_n^* \rightarrow x^*\}.$$

**Observação 1.35** *Note que  $\hat{\partial}f(x) \subset \partial f(x)$  (esta inclusão pode ser própria; veja Exemplo 1.40) e  $\hat{\partial}f(x)$  é convexo. Além disso,  $\partial f(x)$  é fechado. De fato, seja  $\{z^j\} \subset \partial f(x)$  tal que  $z^j \rightarrow z$ . Visto que  $z^j \in \partial f(x)$ , então existem sequências  $\{x^n\} \subset \mathbb{R}^n$  e  $\{x_n^*\} \subset \hat{\partial}f(x)$  tais que  $x^n \rightarrow x$  com  $f(x^n) \rightarrow f(x)$  e  $x_n^* \rightarrow z^j$ ,  $j = 1, 2, \dots$ . Para cada  $j \in \mathbb{N}$ , escolha*

$x_*^{n_j} \in \hat{\partial}f(x^{n_j})$  tal que

$$\|x_*^{n_j} - z^j\| < \frac{1}{2j}. \quad (1-24)$$

Por outro lado, visto que  $z^j \rightarrow z$ , existe  $N \in \mathbb{N}$  tal que

$$\|z^j - z\| < \frac{1}{2j}, \quad j \geq N. \quad (1-25)$$

De (1-24) e (1-25) segue que

$$\|x_*^{n_j} - z\| \leq \|x_*^{n_j} - z^j\| + \|z^j - z\| < \frac{1}{j}, \quad j \geq N.$$

Mas isto nos diz que existe uma sequência  $\{x_*^{n_j}\} \subset \hat{\partial}f(x^{n_j})$  tal que

$$x_*^{n_j} \rightarrow z.$$

Como  $x^{n_j} \rightarrow x$  com  $f(x^{n_j}) \rightarrow f(x)$ , segue que  $z \in \partial f(x)$ , o que conclui a prova.

**Observação 1.36** Uma condição necessária para  $x \in \mathbb{R}^n$  ser um minimizador local de  $f$  é

$$0 \in \partial f(x). \quad (1-26)$$

A menos que  $f$  seja convexa a condição acima não é suficiente (ver Exemplo 1.40).

Um ponto  $x \in \mathbb{R}^n$  que satisfaz a inclusão (1-26) é chamado ponto crítico de  $f$ . Denotamos o conjunto dos pontos críticos de  $f$  por  $\text{crit } f$ .

A seguir provamos algumas propriedades dos subdiferenciais de Fréchet e limite, que são essenciais na prova do resultado principal deste trabalho. Para isso, usaremos o lema abaixo.

**Lema 1.37** Seja  $\varphi(z) = \langle w, \frac{z}{\|z\|} \rangle$ ,  $z \in \mathbb{R}^n \setminus \{0\}$ . Se  $\varphi(z) \geq \frac{o(\|z\|)}{\|z\|}$  para qualquer  $z$  não nulo, então  $w = 0$ .

*Prova* Suponha por contradição que  $w$  seja não nulo. Suponha ainda que,  $C \subset \mathbb{R}^n \setminus \{0\}$  é um conjunto compacto. Como a função  $\varphi$  é contínua, existe  $z_0 \in C$  tal que

$$\varphi(z_0) \leq \varphi(z), \quad z \in C.$$

Note que  $\varphi(z_0) < 0$ . Assim, visto que  $\lim_{z \rightarrow 0} \frac{o(\|z\|)}{\|z\|} = 0$ , existe  $\delta > 0$  tal que, para  $z \in B(0, \delta) \setminus \{0\}$ , temos que

$$\varphi(z_0) < \frac{o(\|z\|)}{\|z\|}.$$

Mas a última desigualdade contraria a hipótese do lema

$$\varphi(z) \geq \frac{o(\|z\|)}{\|z\|}, \quad \forall z \in \mathbb{R}^n \setminus \{0\}.$$

Portanto, segue o resultado. □

O próximo teorema nos diz que se uma função  $f$  é diferenciável em um ponto  $x \in \mathbb{R}^n$  então o subdiferencial de Fréchet no ponto  $x$  é um conjunto que contém somente o vetor  $\nabla f(x)$ .

**Teorema 1.38** *Se  $f$  é diferenciável em um ponto  $x \in \mathbb{R}^n$ , então  $\hat{\partial}f(x) = \{\nabla f(x)\}$ .*

*Prova* Sendo  $f$  é diferenciável em  $x$  temos que

$$f(x+h) = f(x) + \langle \nabla f(x), h \rangle + \bar{o}(\|h\|). \quad (1-27)$$

Assim, pela definição de  $\hat{\partial}f(x)$  segue-se que  $\nabla f(x) \in \hat{\partial}f(x)$ . Tome  $v \in \hat{\partial}f(x)$  arbitrário. Pela definição de  $\hat{\partial}f(x)$ ,

$$f(x+h) \geq f(x) + \langle v, h \rangle + o^*(\|h\|).$$

Combinando a última desigualdade com (1-27) adquirimos

$$\langle \nabla f(x), h \rangle + \bar{o}(\|h\|) \geq \langle v, h \rangle + o^*(\|h\|), \quad h \neq 0.$$

Mas, isto nos diz que

$$\langle \nabla f(x) - v, \frac{h}{\|h\|} \rangle \geq \frac{o(\|h\|)}{\|h\|}. \quad (1-28)$$

E o resultado desejado segue do Lema 1.37 com  $w = \nabla f(x) - v$ . □

**Teorema 1.39** *Se  $f$  é uma função continuamente diferenciável em uma vizinhança de  $x$ , então*

$$\partial f(x) = \{\nabla f(x)\}.$$

*Prova* Pela Observação 1.35 temos que  $\nabla f(x) \in \partial f(x)$ . Tome  $v \in \partial f(x)$ . Então, existem sequências  $\{x^n\} \subset \mathbb{R}^n$  e  $\{x_*^n\} \subset \hat{\partial}f(x^n)$  tais que  $x^n \rightarrow x$  com  $f(x^n) \rightarrow f(x)$  e  $x_*^n \rightarrow v$ . Seja  $B(x, \epsilon)$  uma vizinhança de  $x$  onde  $f$  é continuamente diferenciável. Visto que  $x^n \rightarrow x$ , então existe um índice  $N$  suficientemente grande tal que

$$x^n \in B(x, \epsilon), \quad n \geq N.$$

Última inclusão combinada com Teorema 1.38, implica que

$$x_*^n = \nabla f(x^n).$$

Logo, visto que  $\nabla f$  é contínuo em  $B(x, \varepsilon)$ , segue que  $x_*^n \rightarrow \nabla f(x)$  de onde segue o resultado desejado.  $\square$

O exemplo abaixo nos mostra que a hipótese de continuidade do campo gradiente  $\nabla f$ , assumida no último teorema, é realmente necessária.

**Exemplo 1.40** Consideremos a função  $f$  abaixo

$$f(x) = \begin{cases} x^2 \sin(1/x), & x \neq 0, \\ 0, & x = 0. \end{cases} \quad (1-29)$$

Afirmarmos que  $\hat{\partial}f(0) = \{0\}$  e  $\partial f(0) = [-1, 1]$ . A primeira afirmação desse exemplo segue do Teorema 1.38.

Como  $\hat{\partial}f(0) \subset \partial f(0)$ , temos  $0 \in \partial f(0)$ . Seja  $z \in [-1, 1] \setminus \{0\}$ . Então,  $z = \cos \alpha$ .

Note que se

$$x_n = \frac{1}{\alpha + 2\pi n},$$

temos  $x_n \rightarrow 0$  e  $f(x_n) \rightarrow 0$ . Por outro lado,

$$f'(x_n) = \frac{2}{\alpha + 2\pi n} \sin(\alpha + 2\pi n) - \cos(\alpha + 2\pi n).$$

Mas isto nos diz que  $f'(x_n) \rightarrow -\cos \alpha = -z \in [-1, 1] \setminus \{0\}$ . Portanto,  $[-1, 1] \setminus \{0\} \subset \partial f(0)$ .

Agora tome  $z \notin [-1, 1]$  e assuma que  $z \in \partial f(0)$ . Então, existem seqüências  $\{x_n\} \subset \mathbb{R}$  e  $\{x_*^n\} \subset \hat{\partial}f(x_n)$  tal que  $x_n \rightarrow 0$  e  $x_*^n \rightarrow z$ . Agora, pelo Teorema 1.38, temos que

$$x_*^n = 2x_n \sin(1/x_n) - \cos(1/x_n), \quad n \in \mathbb{N}.$$

Mas isto nos diz que

$$z = \lim_{n \rightarrow +\infty} x_*^n \in [-1, 1],$$

que é um absurdo. Logo,  $\partial f(0) = [-1, 1]$ .

**Teorema 1.41** Sejam  $f, g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ ,  $\varphi = f + g$  e  $x \in \text{dom} f$ . Se  $f$  é continuamente diferenciável em uma vizinhança de  $x$ , então

$$i) \quad \hat{\partial}\varphi(x) \subset \hat{\partial}g(x) + \{\nabla f(x)\};$$

$$ii) \partial\varphi(x) \subset \partial g(x) + \{\nabla f(x)\}.$$

*Prova* Seja  $x^* \in \hat{\partial}\varphi(x)$ . Então,

$$\varphi(y) \geq \varphi(x) + \langle x^*, y - x \rangle + \bar{o}(\|y - x\|),$$

que, combinado com a definição da função  $\varphi$ , implica que

$$f(y) + g(y) \geq f(x) + g(x) + \langle x^*, y - x \rangle + \bar{o}(\|y - x\|).$$

Sendo  $f$  diferenciável em  $x$ , temos que

$$f(y) = f(x) + \langle \nabla f(x), y - x \rangle + o^*(\|y - x\|).$$

Combinando última igualdade com última desigualdade, obtemos

$$\langle \nabla f(x), y - x \rangle + g(y) + o^*(\|y - x\|) \geq g(x) + \langle x^*, y - x \rangle + \bar{o}(\|y - x\|),$$

e, conseqüentemente,

$$g(y) \geq g(x) + \langle x^* - \nabla f(x), y - x \rangle + o(\|y - x\|),$$

com  $o(\|y - x\|) = \bar{o}(\|y - x\|) - o^*(\|y - x\|)$ . Logo, da última desigualdade,  $x^* - \nabla f(x) \in \hat{\partial}g(x)$  de onde segue o item i).

Agora, tome  $x^* \in \partial\varphi(x)$ . Então, pela definição de subdiferencial limite temos que existem sequências  $\{x^n\} \subset \mathbb{R}^n$  e  $\{x_*^n\} \subset \hat{\partial}\varphi(x^n)$  tais que  $x^n \rightarrow x$  com  $\varphi(x^n) \rightarrow \varphi(x)$  e  $x_*^n \rightarrow x^*$ . Como  $f$  é contínua em  $x$ , temos que

$$f(x^n) \rightarrow f(x)$$

quando  $n \rightarrow +\infty$ . Além disso, visto que  $g = \varphi - f$ , temos que

$$\lim_{n \rightarrow \infty} g(x^n) = \lim_{n \rightarrow \infty} \varphi(x^n) - \lim_{n \rightarrow \infty} f(x^n) = \varphi(x) - f(x) = g(x).$$

Por outro lado, visto que

$$x_*^n \in \hat{\partial}\varphi(x^n), \quad n \in \mathbb{N},$$

do item i), temos que existe  $v^n \in \hat{\partial}g(x^n)$  tal que

$$x_*^n = v^n + \nabla f(x^n), \quad n \in \mathbb{N}. \quad (1-30)$$



Tomando o limite em (1-30), obtemos que  $v^n \rightarrow x^* - \nabla f(x)$ . Logo, pela definição de subdiferencial limite de  $g$  em  $x$ , concluímos que

$$x^* - \nabla f(x) \in \partial g(x),$$

de onde segue o item ii) e a prova está finalizada.

□

## Método do Ponto Proximal

Nesse capítulo estudamos o Método do Ponto Proximal no caso que a função objetivo é convexa. A função objetivo agora poderá ser diferenciável ou não. Esse resultado de convergência é baseado no conceito de Féjer Convergência.

### 2.1 Aplicação Proximal

**Definição 2.1** *Seja  $f$  uma função e  $\lambda > 0$ . O operador  $P_\lambda : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , dada por*

$$P_\lambda(w) = \operatorname{argmin}_{x \in \mathbb{R}^n} \left\{ f(x) + \frac{\lambda}{2} \|x - w\|^2 \right\},$$

*é chamada aplicação proximal.*

**Definição 2.2** *Seja  $f$  uma função contínua. Diz-se que  $f$  é 1-coerciva se*

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

**Teorema 2.3** *Seja  $f$  uma função contínua. Se  $f$  é 1-coerciva, então  $f$  possui pelo menos um mínimo global.*

*Prova* Visto que  $f$  é 1-coerciva, temos que  $\lim_{\|x\| \rightarrow +\infty} \frac{f(x)}{\|x\|} = +\infty$ . Consequentemente, dado  $M > 0$  existe um número  $r > 0$  tal que, para  $\|x\| > r$  tem-se

$$\frac{f(x)}{\|x\|} > M \Leftrightarrow f(x) > M\|x\| > Mr.$$

Mas isto nos diz que

$$L_f(Mr) := \{x \in \mathbb{R}^n : f(x) \leq Mr\} \subset B(0, r) := \{x \in \mathbb{R}^n : \|x\| \leq r\}.$$

Visto que  $f$  é convexa, do Corolário (1.12), temos que  $f$  é contínua e consequentemente,  $L_f(Mr)$  é fechado. Portanto, da última inclusão,  $L_f(Mr)$  é um conjunto compacto.

Assim, pelo Teorema de Weierstrass,  $f$  possui pelo menos um mínimo quando restrita a  $L_f(Mr)$ . Seja  $y \in \mathbb{R}^n$  um minimizador de  $f$  restrita a  $L_f(Mr)$ . Afirmamos que  $y$  é um minimizador irrestrito de  $f$ . De fato, visto que  $y$  é um minimizador de  $f$  quando restrita ao conjunto  $L_f(Mr)$ , temos que  $f(y) \leq f(x)$ , para todo  $x \in L_f(Mr)$ . Por outro lado, para todo  $x \notin L_f(Mr)$ ,

$$f(x) > Mr \geq f(y),$$

e o resultado segue. □

**Proposição 2.4** *O operador aplicação proximal está bem definido e satisfaz*

$$\|P_\lambda(w_1) - P_\lambda(w_2)\|^2 \leq \langle w_1 - w_2, P_\lambda(w_1) - P_\lambda(w_2) \rangle, \quad w_1, w_2 \in \mathbb{R}^n. \quad (2-1)$$

*Prova* Tome  $w \in \mathbb{R}^n$ . Da Definição 2.1 temos que

$$P_\lambda(w) = \operatorname{argmin}_{x \in \mathbb{R}^n} \{f(x) + \frac{\lambda}{2} \|x - w\|^2\}.$$

Visto que  $f$  é convexa, pelo item (ii) do Teorema 1.30, segue que existe  $s \in \mathbb{R}^n$  tal que

$$f(x) \geq f(w) + \langle s, x - w \rangle,$$

e, consequentemente,

$$f(x) + \frac{\lambda}{2} \|x - w\|^2 \geq f(w) + \langle s, x - w \rangle + \frac{\lambda}{2} \|x - w\|^2.$$

Dividindo ambos os membros da última por  $\|x - w\|$ , obtemos

$$\frac{f(x)}{\|x - w\|} + \frac{\lambda}{2} \|x - w\| \geq \frac{f(w)}{\|x - w\|} + \langle s, \frac{x - w}{\|x - w\|} \rangle + \lambda \|w - x\|,$$

e pela desigualdade de Cauchy Schwarz, esta última desigualdade se reduz a

$$\frac{f(x)}{\|x - w\|} + \frac{\lambda}{2} \|x - w\| \geq \frac{f(w)}{\|x - w\|} - \|s\| + \lambda \|w - x\|$$

Sendo  $f(w)$  e  $\|w\|$  constantes e  $\lambda > 0$ , fazendo  $\|x - w\|$  ir a  $+\infty$  em ambos os membros da última desigualdade, segue que  $f(x) + \frac{\lambda}{2} \|x - w\|^2$  é 1-coerciva. Assim, como  $f$  é convexa e  $\lambda \|w - x\|^2$  ( $\lambda > 0$ ) é estritamente convexa, a primeira parte da proposição segue do Teorema 2.3 combinado com as Proposições 1.4 e 1.5. Para provar a segunda parte, note

que, pela definição do operador  $P_\lambda$  e da Proposição 1.18,

$$0 \in \partial_c(f(\cdot) + \frac{\lambda}{2}\|\cdot - w\|^2)(P_\lambda(w)), \quad w \in \mathbb{R}^n.$$

Usando Teorema 1.33 com  $g = \frac{\lambda}{2}\|\cdot - w\|^2$  junto com Teorema 1.20 ( $g$  é diferenciável), da última inclusão, obtemos

$$\lambda(w - P_\lambda(w)) \in \partial_c f(P_\lambda(w)). \quad (2-2)$$

Assim, dados  $w_1, w_2 \in \mathbb{R}^n$ , temos que

$$\lambda(w_1 - P_\lambda(w_1)) \in \partial_c f(P_\lambda(w_1)) \text{ e } \lambda(w_2 - P_\lambda(w_2)) \in \partial_c f(P_\lambda(w_2)),$$

de onde obtemos a seguinte desigualdade

$$0 \leq \langle \lambda(w_1 - P_\lambda(w_1)) - \lambda(w_2 - P_\lambda(w_2)), P_\lambda(w_1) - P_\lambda(w_2) \rangle.$$

Levando em conta que  $\lambda > 0$ , a segunda parte da proposição segue de simples manipulações algébricas efetuadas na última desigualdade.

□

**Observação 2.5** *Em termos do operador aplicação proximal, a iterada de ordem  $k + 1$  do método de ponto proximal (0-2) pode ser reescrita como segue*

$$x^{k+1} = P_{\lambda_k}(x^k) := P_k(x^k). \quad (2-3)$$

De agora em diante o conjunto  $X^*$  denotará o conjunto solução do problema (0-1).

**Lema 2.6** *Suponha que  $X^*$  é não vazio. Então,  $w \in X^*$  se, e somente se,  $P_\lambda(w) = w$ .*

*Prova* A prova segue imediatamente da definição de  $P_\lambda$  combinada (2-2).

□

**Observação 2.7** *Note que a desigualdade (2-1) é equivalente a*

$$\|P_\lambda(w_1) - P_\lambda(w_2)\|^2 \leq \|w_1 - w_2\|^2 - \|(I - P_\lambda)(w_1) - (I - P_\lambda)(w_2)\|^2.$$

## 2.2 Análise de Convergência

Apresentaremos agora uma prova de convergência da sequência definida em (0-2) no caso que a função  $f$  é convexa mas não necessariamente diferenciável. Em particular teremos a convergência da sequência gerada por (0-2) para o caso em que a função  $f$  é convexa e continuamente diferenciável.

**Definição 2.8** Dizemos que uma sequência  $\{y^k\} \subset \mathbb{R}^n$  é Féjer convergente a um conjunto  $U \subset \mathbb{R}^n$ , com respeito a norma Euclidiana, se

$$\|y^{k+1} - u\| \leq \|y^k - u\|, \quad u \in U.$$

**Proposição 2.9** Se  $\{y^k\}$  é Féjer convergente a um conjunto  $U \neq \emptyset$ , então  $\{y^k\}$  é limitada. Se um ponto de acumulação  $y$  de  $\{y^k\}$  pertence a  $U$ , então  $y = \lim_{k \rightarrow \infty} y^k$ .

*Prova* Sendo  $\{y^k\}_{k \in \mathbb{N}}$  Féjer convergente ao conjunto  $U$  temos que

$$\|y^k - u\| \leq \|y^0 - u\|$$

para todo  $k \in \mathbb{N}$  e todo  $u \in U$ . Mas isto é equivalente a dizer que  $y^k \in B(u, \|y^0 - u\|)$ , onde  $B(u, \|y^0 - u\|)$  é a bola de centro  $u$  e raio  $\|y^0 - u\|$ . Portanto,  $\{y^k\}_{k \in \mathbb{N}}$  é limitada. Seja  $y \in U$  um ponto de acumulação da sequência  $\{y^k\}_{k \in \mathbb{N}}$ , e  $\{y^{j_k}\}$  subsequência de  $\{y^k\}_{k \in \mathbb{N}}$  tal que  $\lim_{k \rightarrow \infty} y^{j_k} = y$ . Da definição de Féjer Convergência temos que a sequência  $\{\|y^{k+1} - y\|\}$  é decrescente e não negativa. Assim,  $\{\|y^{k+1} - y\|\}$  é monótona limitada e possui uma subsequência  $(\{\|y^{j_k+1} - y\|\})$  que converge a 0. Assim é imediato que  $\|y^{k+1} - y\| \rightarrow 0$ . Portanto,  $\lim_{k \rightarrow \infty} y^k = y$ . □

O próximo teorema desse capítulo, estabelece condições suficientes para que a sequência gerada por (0-2) convirja a um minimizador da função objetivo  $f$ . Para tal prova, usaremos a proposição abaixo.

**Proposição 2.10** Seja  $\{x^k\}_{k \in \mathbb{N}}$  a sequência gerada por (2-3) e assumamos que  $X^* \neq \emptyset$ . Então, valem as seguintes afirmações:

- i)  $\{x^k\}_{k \in \mathbb{N}}$  é Féjer convergente ao conjunto  $X^*$ ;
- ii)  $\{x^k\}_{k \in \mathbb{N}}$  é limitada;
- iii)  $\lim_{k \rightarrow +\infty} (x^{k+1} - x^k) = 0$ .

*Prova* Da segunda parte da Proposição 2.4 combinada com a Observação 2.7, temos que

$$\|P_k(w_1) - P_k(w_2)\|^2 \leq \|w_1 - w_2\|^2 - \|(I - P_k)(w_1) - (I - P_k)(w_2)\|^2.$$

Fazendo  $w_2 = x^*$ ,  $w_1 = x^k$ ,  $P_k(x^k) = x^{k+1}$  e usando que  $P_k(x^*) = x^*$  (ver Lema 2.6), obtemos que

$$\|x^{k+1} - x^*\|^2 \leq \|x^k - x^*\|^2 - \|x^k - x^{k+1}\|^2. \quad (2-4)$$

De (2-4) segue  $i)$ . E de  $i)$  combinada com a Proposição 2.9 segue  $ii)$ . Note também que, a partir de (2-4),  $\{\|x^k - x^*\|\}$  é uma sequência decrescente e não negativa. Em particular, temos que  $\{\|x^k - x^*\|\}$  é uma sequência convergente. Portanto, usando novamente a desigualdade (2-4), obtemos o item  $iii)$  o que conclui a prova da proposição.  $\square$

**Teorema 2.11** *Se  $f$  é uma função convexa não necessariamente diferenciável e  $X^* \neq \emptyset$ , então a sequência gerada por (2-3) converge a um ponto  $x^* \in X^*$ .*

*Prova* Pela definição de  $P_k(x^k)$  em (2-3), segue de (2-2) com  $w = x^k$  e  $\lambda = \lambda_k$  que

$$2\lambda_k(x^k - x^{k+1}) \in \partial_c f(x^{k+1}). \quad (2-5)$$

Agora, visto que  $X^* \neq \emptyset$ , pela Proposição 2.10, segue que a sequência  $\{x^k\}_{k \in \mathbb{N}}$  tem pelo menos um ponto de acumulação. Seja  $x^*$  um ponto de acumulação de  $\{x^k\}_{k \in \mathbb{N}}$  e  $x^{j_k}$  uma subsequência de  $\{x^k\}_{k \in \mathbb{N}}$  tal que  $\lim_{k \rightarrow \infty} x^{j_k} = x^*$ . Então, de (2-5) obtemos que

$$2\lambda_{j_k}(x^{j_k} - x^{j_k+1}) \in \partial_c f(x^{j_k+1}).$$

Pela definição de subdiferencial para funções convexas temos,

$$f(x) - f(x^{j_k}) \geq \lambda_{j_k} \langle x^{j_k} - x^{j_k+1}, x - x^{j_k+1} \rangle, \quad x \in \mathbb{R}^n,$$

e o resultado desejado segue por combinando os seguintes fatos:  $\{\lambda_k\}$  é limitada,  $\{x^{j_k}\}$  converge a  $x^*$  e  $f$  é contínua ( $f$  é convexa).  $\square$

O próximo resultado é uma consequência imediata do último teorema.

**Corolário 2.12** *Se  $f$  é uma função convexa continuamente diferenciável e  $X^* \neq \emptyset$ , então a sequência gerada por (2-3) converge a um ponto  $x^* \in X^*$ .*

## Convergência do Método do Ponto Proximal para Funções que satisfazem a Desigualdade de Łojasiewicz

Neste capítulo apresentamos a desigualdade de Łojasiewicz que, neste trabalho, representa uma condição suficiente sobre a função objetivo para que a sequência gerada pelo método Proximal seja convergente. Uma prova de convergência da sequência gerada pelo Método do Ponto Proximal para funções não diferenciáveis que satisfazem a desigualdade de Łojasiewicz é apresentada.

### 3.1 Resultados Auxiliares

Em todo o trabalho, a menos que seja mencionado o contrário explicitamente,  $f$  representa uma função  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ . O conjunto dos pontos onde  $f$  assume valor finito é chamado domínio efetivo de  $f$  e o denotamos por  $\text{dom} f$ . Quando  $\text{dom} f \neq \emptyset$ , então  $f$  é dita ser própria. Dado  $c \in \mathbb{R}$ , denotamos por

$$L_f(c) := \{x \in \mathbb{R}^n \mid f(x) \leq c\},$$

o conjunto de nível de  $f$  associado a  $c$ .

**Definição 3.1** *Seja  $x \in \text{dom} f$ . Uma função  $f$  é dita ser semicontínua inferior em  $x$  se, para toda sequência  $\{x^k\}$  que converge a  $x$ , tem-se*

$$\liminf_{k \rightarrow +\infty} f(x^k) \geq f(x).$$

O Próximo Teorema é uma adaptação do Teorema 2.3 do capítulo 3. Este resultado será usado para garantir a boa definição da sequência gerada pelo Método do Ponto Proximal no contexto deste capítulo.

**Teorema 3.2** *Se  $f$  é uma função própria semicontínua inferior e 1-coerciva, então  $f$  possui pelo menos um mínimo global.*

*Prova* A prova deste resultado é análoga a prova do Teorema 2.3. □

Seja  $f$  uma função própria semicontínua inferior, considere a sequência  $\{x^k\}$  gerada pelo Método de Ponto Proximal (0-2) e assumamos que  $f$  é limitada inferiormente. Neste caso,

$$f_k := f + (1/2\lambda_k)\|\cdot - x^k\|^2,$$

é semicontínua inferior e 1-coerciva e, consequentemente, a partir do Teorema 3.2, temos que a sequência  $\{x^k\}$  está bem definida. Visto que  $f$  é própria, sem perda de generalidade, podemos assumir que a sequência inteira  $\{x^k\} \subset \text{dom} f$ . Além disso, pela definição da sequência  $\{x^k\}$ , segue imediatamente da Observação 1.36 combinada com Teoremas 1.41 e 1.39, que

$$x^{k+1} = x^k - \lambda_k g^{k+1}, \quad (3-1)$$

onde  $g^{k+1} \in \partial f(x^{k+1})$ . Note que, da última igualdade, se a sequência  $\{x^k\}$  é finita, então ela pára em um ponto crítico da função  $f$ . De agora em diante assumimos que  $f$  é própria, semicontínua e limitada inferiormente, a sequência  $\{x^k\}$  representa a sequência gerada pelo Método de Ponto Proximal (0-2) com  $x^0 \in \text{dom} f$ ,  $\{x^k\}$  é infinita e  $\lambda_k \in (\lambda_-, \lambda_+)$  para todo  $k \in \mathbb{N}$ , onde  $\lambda_-$  e  $\lambda_+$  são parâmetros positivos fixados com  $\lambda_- < \lambda_+ < +\infty$ .

**Hipótese H1.** A função  $f$  restrita ao seu domínio efetivo é uma função contínua.

O seguinte resultado reúne alguns fatos elementares relacionados à sequência  $\{x^k\}$  gerada por (0-2).

**Proposição 3.3** *Seja  $\omega(x^0)$  o conjunto dos pontos limites da sequência  $\{x^k\}$ . Então,*

- i) *A sequência  $\{f(x^k)\}$  é decrescente;*
- ii)  *$\sum_{k=0}^{+\infty} \|x^{k+1} - x^k\|^2 < +\infty$ ;*
- iii) *Se  $f$  satisfaz H1, então  $\omega(x^0) \subset \text{crit} f$ .  
Se, em adição, a sequência  $\{x^k\}$  é limitada, então*
- iv)  *$\omega(x^0)$  é não vazio, conexo, compacto, e*

$$d(x^k, \omega(x^0)) \rightarrow 0, \quad \text{quando } k \rightarrow +\infty.$$

- v) *Se  $f$  satisfaz H1, então  $f$  é finita e constante em  $\omega(x^0)$ .*

*Prova* (i) Note que pela definição de  $x^{k+1}$  temos

$$f_k(x^{k+1}) \leq f_k(x^k).$$



Mas isto é equivalente a dizer que

$$f(x^{k+1}) + \frac{1}{2\lambda_k} \|x^{k+1} - x^k\|^2 \leq f(x^k). \quad (3-2)$$

Logo segue da última desigualdade que  $f(x^{k+1}) \leq f(x^k)$ , e o item *i*) está provado.

(ii) Agora note que de (3-2) obtem-se que

$$\|x^k - x^{k+1}\|^2 \leq 2\lambda_k(f(x^k) - f(x^{k+1})).$$

Da última desigualdade somando ambos os lados com  $k$  variando de 0 até  $N_0$  temos

$$\sum_{k=0}^{N_0} \|x^k - x^{k+1}\|^2 \leq \sum_{k=0}^{N_0} 2\lambda_k(f(x^k) - f(x^{k+1})).$$

Mas a última desigualdade combinada com a limitação superior da sequência  $\{\lambda_k\}$  nos diz que

$$\sum_{k=0}^{N_0} \|x^k - x^{k+1}\|^2 \leq 2\lambda_+(f(x^0) - f(x^{N_0})).$$

Como a função  $f$  é limitada inferiormente temos que

$$\sum_{k=0}^{N_0} \|x^k - x^{k+1}\|^2 \leq 2\lambda_+(f(x^0) - \inf_{x \in \mathbb{R}^n} f(x)) < +\infty.$$

Tomando o limite com  $N_0 \rightarrow +\infty$ , *ii*) está provado.

(iii) Para provar o item *iii*), assumamos que  $\omega(x^0) \neq \emptyset$  (quando  $\omega(x^0) = \emptyset$  é imediato que  $\omega(x^0) \subset \text{crit } f$ ). Como  $\omega(x^0) \neq \emptyset$  podemos tomar  $x \in \omega(x^0)$ . Então, existe uma subsequência  $\{x^{k_j}\}$  de  $\{x^k\}$  tal que  $x^{k_j} \rightarrow x$ . Por outro lado, de (3-1) temos que

$$g^{k+1} = \frac{1}{\lambda_k}(x^k - x^{k+1}). \quad (3-3)$$

Combinando *ii*) e a limitação superior de  $\{\lambda_k\}$  com (3-3) segue que  $\{g^k\}$  converge a 0. Podemos ver que  $(x, 0)$  é ponto de acumulação da sequência  $\{(x^{k_j}, g^{k_j})\}$ . Portanto, visto que gráfico de  $\partial f$  é fechado, segue que  $(x, 0) \in \text{graf } \partial f$  e, consequentemente, que  $0 \in \partial f(x)$ , isto é,  $x \in \text{crit } f$ .

(iv) É imediato que  $\omega(x^0) \neq \emptyset$  pois toda sequência limitada possui ponto de acumulação. Seja  $\bar{y}$  um ponto de acumulação de  $\omega(x^0)$ , e  $\{y^j\} \subset \omega(x^0)$  uma sequência com  $y^j \rightarrow \bar{y}$ . Agora, para cada  $j \in \mathbb{N}$ , existe  $x^{k_j}$ , tal que

$$\|x^{k_j} - y^j\| \leq \frac{1}{j}.$$

Assim

$$\|x^{k_j} - \bar{y}\| \leq \|x^{k_j} - y^j\| + \|y^j - \bar{y}\|. \quad (3-4)$$

Observe que para  $j$  suficientemente grande temos

$$\|x^{k_j} - y^j\| \leq \frac{1}{2j} \quad (3-5)$$

e

$$\|y^j - \bar{y}\| \leq \frac{1}{2j}. \quad (3-6)$$

Logo, combinando as desigualdades (3-5) e (3-6) com (3-4) e tomando  $j$  suficientemente grande

$$\|x^{k_j} - \bar{y}\| \leq \frac{1}{j}.$$

A última desigualdade acima nos diz que  $\{x^{k_j}\}$  é uma subsequência de  $\{x^k\}$  que converge a  $\bar{y}$ . Portanto  $\bar{y} \in \omega(x^0)$ , isto é,  $\omega(x^0)$  é fechado. Suponha por contradição que  $\omega(x^0)$  seja ilimitado, então existe  $y^j \in \omega(x^0)$  com  $\lim_{j \rightarrow \infty} y^j = \infty$ . Note que

$$\|x^{k_j}\| + \|x^{k_j} - y^j\| \geq \|y^j\|.$$

Mas isto nos diz que  $\lim_{j \rightarrow \infty} \|x^{k_j}\| = +\infty$ , o que contraria o fato de  $\{x^k\}$  ser limitada. Logo, é absurdo supor que  $\omega(x^0)$  é ilimitado. Portanto  $\omega(x^0)$  é compacto.

A prova da conexidade do conjunto  $\omega(x^0)$  será feita por contradição. Suponha que  $\omega(x^0) = A \cup B$  onde  $A$  e  $B$  são conjuntos compactos, disjuntos, e não vazios. Sejam  $\xi = \text{dist}(A, B)$  e defina  $\varepsilon = \frac{\xi}{3} > 0$ . Sejam

$$U = \{x \in \mathbb{R}^n; \text{dist}(x, A) < \varepsilon\}$$

e

$$V = \{x \in \mathbb{R}^n; \text{dist}(x, B) < \varepsilon\}.$$

Defina os seguintes subconjuntos de  $\mathbb{N}$  abaixo

$$\mathbb{N}_U = \{k \in \mathbb{N}; x^k \in U\} \quad e \quad \mathbb{N}_V = \{k \in \mathbb{N}; x^k \in V\}.$$

Em outras palavras  $\mathbb{N}_U$  e  $\mathbb{N}_V$  formam respectivamente o conjuntos dos índices das subsequências de  $\{x^k\}_{k \in \mathbb{N}}$  que convergem para algum ponto de  $A$  e o conjunto dos índices das subsequências de  $\{x^k\}_{k \in \mathbb{N}}$  que convergem para algum ponto de  $B$ . Estes conjuntos são ambos infinitos, disjuntos, limitados inferiormente e  $\mathbb{N} = \mathbb{N}_U \cup \mathbb{N}_V$ . Podemos supor, sem perda de generalidade, que  $\min \mathbb{N}_U < \min \mathbb{N}_V$ . Afirmamos que, sempre existirá um índice  $j \in \mathbb{N}_U$  tal que  $j+1 \in \mathbb{N}_V$ , uma vez que, caso não exista tal índice, o conjunto

$\mathbb{N}_V$  seria vazio. Agora a partir do índice  $j+1 \in \mathbb{N}_V$  existirá um índice  $j_m \geq j+1$  tal que  $j_m+1 \in \mathbb{N}_U$ , pois caso não exista,  $\mathbb{N}_U$  seria finito. Dessa forma existe uma subsequência de  $\{x^k\}_{k \in \mathbb{N}}$ ,  $\{x^{k_j}\}$ , tal que para algum  $j_0$ ,  $x^{k_{j_0}} \in U$  e  $x^{k_{j_0}+1} \in V$ . Mas isto nos diz que

$$\|x^{k_{j_0}+1} - x^{k_{j_0}}\| > \varepsilon.$$

Considere  $\{x^k\}_{k \geq k_{j_0}+1}$ . Novamente existe uma subsequência  $\{x^{k_m}\}_{k_m \geq k_{j_0}+1}$  tal que para algum  $j_1$ ,  $x^{k_{j_1}} \in V$  e  $x^{k_{j_1}+1} \in U$ . Logo,

$$\|x^{k_{j_1}+1} - x^{k_{j_1}}\| > \varepsilon.$$

Repetindo esse processo infinitas vezes, obtemos uma infinidade de termos  $\{x^{k_{j_m}}\} \subset \{x^k\}$  tais que

$$\|x^{k_{j_m}+1} - x^{k_{j_m}}\| > \varepsilon, \quad m = 0, 1, \dots,$$

onde

$$k_{j_0} < k_{j_1} < \dots < k_{j_p} < \dots,$$

com  $k_{j_i} \in \mathbb{N}$ . Formando assim uma subsequência de  $\{\|x^{k+1} - x^k\|\}$  que não tende a 0. O que é absurdo uma vez que  $\|x^{k+1} - x^k\| \rightarrow 0$  (item *ii*)) obriga todas as suas subsequências tenderem para 0.

Vamos supor que exista uma subsequência  $\{x^{k_j}\}$  de  $\{x^k\}$  tal que  $\text{dist}(x^{k_j}, \omega(x^0)) \rightarrow l > 0$  quando  $j \rightarrow \infty$ . Mas isto nos diz que existem infinitos índices da sequência  $\{x^k\}$  que estão a uma distância positiva do conjunto  $\omega(x^0)$ . Estes termos  $\{x^{k_j}\}$ , quando agrupados, formam uma subsequência de  $\{x^k\}$ , que é limitada. Portanto  $\{x^{k_j}\}$  possui um ponto de acumulação (ponto de acumulação de  $\{x^k\}$ ) que não pertence a  $\omega(x^0)$ , o que é uma contradição. Portanto  $\text{dist}(x^k, \omega(x^0)) \rightarrow 0$  quando  $k \rightarrow +\infty$ .

(v) Assumindo que  $\{x^k\}$  é limitada, obtemos que  $\omega(x^0) \neq \emptyset$ . Tome  $x^* \in \omega(x^0)$  arbitrário e  $\{x^{k_j}\}$  uma subsequência de  $\{x^k\}$  que converge a  $x^*$ . Visto que  $f$  satisfaz a hipótese **H1** temos que

$$f(x^{k_j}) \rightarrow f(x^*),$$

como  $\{f(x^k)\}$  é decrescente (isto segue do item *i*), a última desigualdade implica que

$$\lim_{k \rightarrow +\infty} f(x^k) = f(x^*).$$

Assim, dado que  $x^0 \in \text{dom} f$  e  $\{f(x^k)\}$  é monótona, da última desigualdade obtemos que  $f(x^*) < +\infty$ , o que prova a primeira parte do item v). Agora, assumindo que  $f$  satisfaz a hipótese **H1**, e levando em conta que  $\{x^k\} \subset \text{dom} f$  e  $x^* \in \text{dom} f$ , a segunda parte do item v) segue.

□

## 3.2 Desigualdade de Łojasiewicz

Assuma que  $f$  satisfaz a hipótese **H1**. Dizemos que a função  $f$  satisfaz a propriedade de Łojasiewicz se:

Para qualquer ponto  $\hat{x} \in \text{crit} f$ , existe  $C, \varepsilon > 0$  e  $\theta \in [0, 1)$  tal que

$$|f(x) - f(\hat{x})|^\theta \leq C \|x^*\|, \quad x \in B(\hat{x}, \varepsilon), \quad x^* \in \partial f(x). \quad (3-7)$$

**Observação 3.4** *Assumindo que  $f(\hat{x}) = 0$  e levando em conta que se  $f$  é continuamente diferenciável então*

$$\partial f(x) = \{\nabla f(x)\},$$

(3-7) se reduz à seguinte desigualdade

$$\|\nabla f(x)\| \geq \overline{C} |f(x)|^\theta, \quad x \in B(0, \varepsilon).$$

**Observação 3.5** *Quando  $\theta = 0$  usamos a convenção  $0^0 = 0$ . Neste caso, se  $|f(x) - f(\hat{x})|^0 = 0$ , então  $f(x) = f(\hat{x})$ .*

A seguir apresentamos alguns exemplos de classes de funções que satisfazem a desigualdade de Łojasiewicz.

## 3.3 Exemplos de funções que satisfazem a desigualdade de Łojasiewicz

**Exemplo 3.6** *As funções reais analíticas tem a propriedade de Łojasiewicz, veja [7]. Faremos aqui uma demonstração deste fato para o  $n = 1$ , isto é,  $f : \mathbb{R} \rightarrow \mathbb{R}$ . Considere uma função real  $f$  analítica em uma vizinhança,  $B(\bar{x}, \delta)$ , de um ponto crítico  $\bar{x}$ . Então temos*

$$f(x) - f(\bar{x}) = \sum_{k=p_0}^{+\infty} \frac{f^{(k)}(\bar{x})}{k!} (x - \bar{x})^k, \quad p_0 \geq 2, \quad \frac{f^{(p_0)}(\bar{x})}{p_0!} \neq 0, \quad x \in B(\bar{x}, \delta).$$

E pela última igualdade acima, derivando em ambos os lados obtemos

$$f'(x) = \sum_{k=p_0}^{+\infty} k \frac{f^{(k)}(\bar{x})}{k!} (x - \bar{x})^{k-1}, \quad p_0 \geq 2, \quad \frac{f^{(p_0)}(\bar{x})}{p_0!} \neq 0, \quad x \in B(\bar{x}, \delta).$$

Dessa forma, temos que existe  $0 < \tilde{\delta} < \min\{\delta, 1\}$  suficientemente pequeno tal que

$$|f(x) - f(\bar{x})|^\theta \approx \left| \frac{f^{(p_0)}(\bar{x})}{p_0!} (x - \bar{x})^{p_0} \right|^\theta, \quad x \in B(\bar{x}, \tilde{\delta}), \quad (3-8)$$

e

$$|f'(x)| \approx \left| p_0 \frac{f^{(p_0)}(\bar{x})}{p_0!} (x - \bar{x})^{p_0-1} \right|, \quad x \in B(\bar{x}, \tilde{\delta}). \quad (3-9)$$

Agora, de (3-8) e (3-9), segue que

$$\frac{|f(x) - f(\bar{x})|^\theta}{|f'(x)|} \approx \frac{|x - \bar{x}|^{p_0\theta - p_0 + 1}}{|p_0|}, \quad x \in B(\bar{x}, \tilde{\delta}).$$

Tome  $\theta \in \mathbb{R}$  tal que

$$1 - \frac{1}{p_0} < \theta < 1.$$

Note que  $(p_0\theta - p_0 + 1) \in (0, 1)$ . Assim, temos

$$\frac{|f(x) - f(\bar{x})|^\theta}{|f'(x)|} \leq 1.$$

Mas isto nos diz que

$$|f(x) - f(\bar{x})|^\theta \leq |f'(x)|.$$

E a propriedade de Łojasiewicz é verificada tomando  $C = 1$ .

**Exemplo 3.7** Uma outra ampla classe de funções que satisfazem a desigualdade de Łojasiewicz é dada pelas funções de Morse (funções cujos pontos críticos são não degenerados). Note que se  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  é uma função de Morse e  $\bar{x} \in \text{crit } f$ , então  $\nabla^2 f(\bar{x})$  é invertível. Afirmamos que existe  $c > 0$ , tal que

$$\|\nabla^2 f(\bar{x})x\| \geq c\|x\|, \quad x \in \mathbb{R}^n.$$

De fato,

$$\|x\| = \|\nabla^2 f(\bar{x})^{-1} \circ \nabla^2 f(\bar{x})x\| \leq \|\nabla^2 f(\bar{x})^{-1}\| \|\nabla^2 f(\bar{x})x\|.$$

Combinando a última desigualdade acima com a fato de  $\|\nabla^2 f(\bar{x})^{-1}\| = \frac{1}{\|\nabla^2 f(\bar{x})\|}$ , a afirmação segue tomando  $c = \|\nabla^2 f(\bar{x})\|$ . Em particular, podemos assegurar a existência

de  $c > 0$  tal que

$$\|\nabla^2 f(\bar{x})(x - \bar{x})\| \geq c\|x - \bar{x}\|, \quad x \in \mathbb{R}^n.$$

Sendo  $f$  uma função de Morse existe uma vizinhança de  $\bar{x}$ ,  $B(\bar{x}, \bar{\delta})$ , tal que  $B(\bar{x}, \bar{\delta}) \cap \text{crit } f = \{\bar{x}\}$ . Consideremos os desenvolvimentos em série de Taylor de  $f$  e de  $\nabla f$  abaixo

$$f(x) - f(\bar{x}) = \langle \nabla^2 f(\bar{x})(x - \bar{x}), x - \bar{x} \rangle + o^*(\|x - \bar{x}\|^2), \quad (3-10)$$

e

$$\nabla f(x) = \nabla^2 f(\bar{x})(x - \bar{x}) + \bar{o}(\|x - \bar{x}\|), \quad (3-11)$$

onde  $\lim_{x \rightarrow \bar{x}} \frac{o^*(\|x - \bar{x}\|^2)}{\|x - \bar{x}\|^2} = 0$  e  $\lim_{x \rightarrow \bar{x}} \frac{\bar{o}(\|x - \bar{x}\|)}{\|x - \bar{x}\|} = 0$ . Dessa forma, dados  $\varepsilon_1 > 0$  e  $\varepsilon_2 = \frac{c}{2}$ , existem  $\delta_1 > 0$  e  $\delta_2 > 0$  tais que

$$o^*(\|x - \bar{x}\|^2) < \varepsilon_1 \|x - \bar{x}\|^2, \quad x \in B(\bar{x}, \delta_1) \quad (3-12)$$

e

$$\bar{o}(\|x - \bar{x}\|) < \frac{c}{2} \|x - \bar{x}\|, \quad x \in B(\bar{x}, \delta_2). \quad (3-13)$$

Sejam  $\delta = \min\{\bar{\delta}, \delta_1, \delta_2\}$  e  $x \in B(\bar{x}, \delta)$ . Assim, da igualdade (3-10) combinada com a desigualdade (3-12), segue que

$$|f(x) - f(\bar{x})| \leq K\|x - \bar{x}\|^2, \quad K = \|\nabla^2 f(\bar{x})\| + \varepsilon_1, \quad x \in B(\bar{x}, \delta). \quad (3-14)$$

Agora, de (3-11), temos que

$$\|\nabla f(x)\| \geq \|\nabla^2 f(\bar{x})(x - \bar{x})\| - \bar{o}(\|x - \bar{x}\|), \quad x \in B(\bar{x}, \delta),$$

e, da última desigualdade acima combinada com (3-13), segue que

$$\|\nabla f(x)\| \geq \frac{c}{2} \|x - \bar{x}\|. \quad (3-15)$$

Assim, de (3-14) e (3-15) obtemos

$$|f(x) - f(\bar{x})|^\theta \leq C\|\nabla f(x)\|, \quad x \in B(\bar{x}, \delta),$$

onde  $\theta = \frac{1}{2}$  e  $C = \frac{2\sqrt{K}}{c}$ . Portanto, da arbitrariedade de  $\bar{x} \in \text{crit } f$ , concluímos que  $f$  tem a propriedade de Łojasiewicz.

**Lema 3.8** Suponhamos que  $f$  satisfaz a propriedade de Łojasiewicz.

i) Se  $K$  é um conjunto compacto, então existe  $C, \varepsilon > 0$  e  $\theta \in [0, 1)$  tal que para todo

$x \in \mathbb{R}^n$  com  $d(x, K) \leq \varepsilon$ , temos que

$$|f(x) - f(\hat{x})|^\theta \leq C\|x^*\|, \quad x^* \in \partial f(x). \quad (3-16)$$

ii) Se  $K$  é um subconjunto conexo do conjunto dos pontos críticos de  $f$ , então  $f$  é constante em  $K$ .

*Prova* Visto que  $K$  é compacto, então existe um conjunto finito de índices  $I = \{1, \dots, p\}$  tal que  $x^i \in K$  e  $K \subset \cup_{i=1}^p B_i$ ,  $i \in I$ , onde  $B_i$  representa a bola aberta  $B(x^i, \varepsilon)$  com centro em  $x^i$  e raio  $\varepsilon_i$ . Sem perda de generalidade podemos assumir que os  $\varepsilon_i$ 's são tais que a desigualdade de Lojasiewicz ocorre com constantes  $C_i, \theta_i$ , para  $i \in I$ , isto é, para cada  $i \in I$  e  $x \in B_i$ , temos

$$|f(x) - f(x^i)|^{\theta_i} \leq C_i\|x^*\|, \quad x^* \in \partial f(x).$$

Em particular, temos que  $f$  é localmente constante (contínua) no conjunto conexo  $K$  e, portanto, constante. Agora, escolhendo  $\varepsilon > 0$  suficientemente pequeno, obtemos que

$$\{x \in \mathbb{R}^n | d(x, K) \leq \varepsilon\} \subset \cup_{i=1}^p B(x^i, \varepsilon_i),$$

e o item ii) segue tomando  $C = \max_{i \in I} C_i$  e  $\theta = \max_{i \in I} \theta_i$ . O item i) é uma consequência imediata do item ii).

□

### 3.4 Resultado de Convergência

**Lema 3.9** *Seja  $\{a_k\}$  uma sequência de números positivos tal que*

$$\sum_{k=1}^{+\infty} a_k^2 / a_{k-1} < +\infty.$$

*Então,  $\sum_{k=1}^{+\infty} a_k < +\infty$ .*

*Prova* Seja  $j \in \mathbb{N}$  fixado. Então temos pela Desigualdade de Cauchy-Schwarz

$$\sum_{k=1}^j a_k = \sum_{k=1}^j \frac{a_k}{\sqrt{a_{k-1}}} \sqrt{a_{k-1}} \leq \left( \sum_{k=1}^j \frac{a_k^2}{a_{k-1}} \right)^{1/2} \left( \sum_{k=1}^j a_{k-1} \right)^{1/2}.$$

Assim somando em ambos os lados  $a_0$ , vem que

$$\sum_{k=1}^{j+1} a_{k-1} \leq a_0 + \left( \sum_{k=1}^j \frac{a_k^2}{a_{k-1}} \right)^{1/2} \left( \sum_{k=1}^{j+1} a_{k-1} \right)^{1/2}.$$

O resultado segue dividindo ambos os lados da última desigualdade por  $\left( \sum_{k=1}^{j+1} a_{k-1} \right)^{1/2}$  e lembrando que

$$\frac{a_0}{\left( \sum_{k=1}^j a_{k-1} \right)^{1/2}} \leq \sqrt{a_0}.$$

Assim segue que

$$\left( \sum_{k=1}^{j+1} a_{k-1} \right)^{1/2} \leq \sqrt{a_0} + \left( \sum_{k=1}^j \frac{a_k^2}{a_{k-1}} \right)^{1/2}.$$

Desta última desigualdade temos que a finitude do somatório da esquerda depende apenas da finitude de  $\left( \sum_{k=1}^j \frac{a_k^2}{a_{k-1}} \right)^{1/2}$ , que é finito por hipótese. □

**Teorema 3.10** *Assuma que  $f$  satisfaz a hipótese **H1** e a propriedade de Lojasiewicz. Se a sequência  $\{x^k\}$  é limitada, então*

$$\sum_{k=0}^{+\infty} \|x^{k+1} - x^k\| < +\infty. \quad (3-17)$$

*Em particular, a sequência  $\{x^k\}$  converge a algum ponto crítico de  $f$ .*

*Prova* Podemos assumir, sem perda de generalizade, que  $\{f(x^k)\}$  converge para 0 (basta considerar  $f - \inf_{k \leq 0} f(x^k)$  como sendo a função  $f$ ). Assim, pelo item i) da Proposição 3.3, temos que a sequência  $\{f(x^k)\}$  é positiva e decresce estritamente para 0.

Consideremos a função  $\beta : \mathbb{R}_{++} \rightarrow \mathbb{R}$  dada por  $\beta(s) = -s^{1-\theta}$ . Claramente  $\beta$  é uma função convexa. Assim, temos que

$$\beta(f(x^k)) - \beta(f(x^{k+1})) \leq \beta'(f(x^k))(f(x^k) - f(x^{k+1})), \quad k \geq 0,$$

o que implica que

$$f(x^k)^{1-\theta} - f(x^{k+1})^{1-\theta} \leq (1-\theta)f(x^k)^{-\theta}(f(x^k) - f(x^{k+1})), \quad k \geq 0.$$



Agora, pela definição da sequência  $\{x^k\}$ , temos que

$$f(x^k) - f(x^{k+1}) \geq \frac{1}{2\lambda_k} \|x^{k+1} - x^k\|^2, \quad k \geq 0.$$

Assim, combinando as duas últimas desigualdades, obtemos

$$f(x^k)^{1-\theta} - f(x^{k+1})^{1-\theta} \geq (1-\theta)f(x^k)^{-\theta} \frac{1}{2\lambda_k} \|x^{k+1} - x^k\|^2. \quad (3-18)$$

Tome  $\bar{x} \in \omega(x^0)$ . Visto que  $f$  satisfaz a hipótese **H1**, e levando em conta que  $\{x^k\} \subset \text{dom} f$ , temos que a sequência  $\{f(x^k)\}$  converge a  $f(\bar{x}) = 0$ . Por outro lado, do item *iii*) da Proposição 3.3,  $\omega(x^0) \subset \text{crit} f$ . Portanto, dos itens *iii*) e *iv*) da Proposição 3.3 e Lema 3.8 (tomando  $K = \omega(x^0)$ ) podemos garantir que existe  $N_0 \in \mathbb{N}$ ,  $C > 0$  e  $\theta \in [0, 1)$  tal que, para todo  $k \geq N_0$ ,

$$|f(x^k)|^\theta \leq C \|g^k\| = \frac{C}{\lambda_k} \|x^k - x^{k-1}\|, \quad g^k \in \partial f(x^k), \quad (3-19)$$

onde a última igualdade segue de (3-1). Das desigualdades (3-18) e (3-19) segue, imediatamente, que

$$f(x^k)^{1-\theta} - f(x^{k+1})^{1-\theta} \geq \frac{(1-\theta)}{2C} \frac{\|x^{k+1} - x^k\|^2}{\|x^k - x^{k-1}\|}, \quad k \geq N_0.$$

Tomando  $M = \frac{2C}{(1-\theta)} > 0$ , da última desigualdade, segue imediatamente que

$$\frac{\|x^{k+1} - x^k\|^2}{\|x^k - x^{k-1}\|} \leq M(f(x^k)^{1-\theta} - f(x^{k+1})^{1-\theta}), \quad k \geq N_0. \quad (3-20)$$

Assim, para todo  $n \geq N_0$ , temos que

$$\sum_{k=N_0}^n \frac{\|x^{k+1} - x^k\|^2}{\|x^k - x^{k-1}\|} \leq M f(x^{N_0})^{1-\theta},$$

o que implica que

$$\sum_{k=0}^{+\infty} \frac{\|x^{k+1} - x^k\|^2}{\|x^k - x^{k-1}\|} < +\infty.$$

Logo, (3-17) segue do Lema 3.9 fazendo  $a_k = \|x^{k+1} - x^k\|$ . A última parte segue de (3-17) combinado com item *iii*) da Proposição 3.3.

□

## 3.5 Conclusão

Apresetamos uma análise de convergência da sequência gerada pelo Método de Ponto Proximal para funções não necessariamente convexas e não diferenciáveis que satisfazem a desigualdade de Lojasiewicz.

---

## Referências Bibliográficas

---

- [1] ATTOUCH, H; BOLTE, J; NIRENBERG, L. **On the convergence of the proximal algorithm for nonsmooth functions involving analytic features.** *Mathematical Programming*, 116:5–16, 2009.
- [2] CLARKE, F. H. **Optimization and Nonsmooth Analysis.** SIAM, New York, 1983.
- [3] FRIEDMAN, A. **Foundations of Modern Analysis.** Dover Publications, New York, 1982.
- [4] IUSEM, A. N. **Proximal point methods in optimization.** Colóquio Brasileiro de Matemática, IMPA-RJ, 1995.
- [5] IZMAILOV, A; SOLODOV, M. **Otimização Vol.1 Condições de Otimalidade, Elementos de Análise Convexa e de Dualidade.** IMPA-RJ, Rio de Janeiro, 2009.
- [6] KURDYKA, K. **On the gradients of functions definable in o-minimal structures.** *Annales de l'institute Fourier*, 3:769–783, 1998.
- [7] LOJASIEWICZ, S. **"une propriété topologique des sous-ensembles analytiques réels."**, in: *Les Équations aux dérivées partielles.* Éditions du centre National de la Recherche Scientifique, Paris 1963, 1:87–89, 1963.
- [8] MARTINET, B. **Régulation, d'inéquations variationnelles par approximations successives.** *Revue Française d'Informatique et de Recherche Operationelle*, Paris, 1970.
- [9] MÄKELÄ, M. M; NEITANMÄKI, P. **Nonsmooth optimization: analysis and algorithms with applications to optimal control.** World Scientific, Finland, 1992.
- [10] ROCKAFELLAR, R. T; WETS. **Variational Analysis.** Springer, Paris, 1998.
- [11] ROCKAFELLAR, R. T. **Monotone operators and the proximal point method.** *Journal of Control Optimization*, 14:144–159, 1976.
- [12] TIEL, J. V. **Convex Analysis: An Introductory Text.** John Wiley and Sons, Royal Netherlands Meteorological Institute, 1984.