UNIVERSIDADE FEDERAL DE GOIÁS
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DEPARTAMENTO DE MATEMÁTICA

PEDRO BONFIM DE ASSUNÇÃO FILHO

Um Algoritmo Proximal com Quase-distância

Goiânia
2015

UNIVERSIDADE FEDERAL DE GOIÁS
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA-MESTRADO

UFG

Termo de Ciência e de Autorização para Disponibilizar as Dissertações Eletrônicas à CAPES

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás–UFG a disponibilizar gratuitamente através do Programa de Pós-Graduação em Matemática-Mestrado, sem ressarcimento dos direitos autorais, de acordo com a Lei nº 9610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data.

1. Identificação do material bibliográfico: **[X] Dissertação**

2. Identificação da Dissertação

Autor(a):	Pedro Bonfim de Assunção Filho						
CPF:	009.203.491-88	E-mail:	pedro.ufg.mat@gmail.com				
Seu e-mail pode ser disponibilizado na página? [X]Sim [ ] Não							
Vínculo Empregatício do autor	Sem vínculo						
Agência de fomento:	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior					Sigla:	CAPES
País:	Brasil	UF:	GO	CNPJ:	00.889.834/0001-08		
Título:	Um Algoritmo Proximal com Quase-distância						
Palavras-chave:	Algoritmo Proximal, desigualdade de Kurdyka-Lojasiewicz						
Título em outra língua:	A proximal Algorithm with quasi-distance						
Palavras-chave em outra língua:	Proximal Algorithm, Inequality Kurdyka-Lojasiewicz						
Área de concentração:	Otimização						
Data defesa: (dd/mm/aaaa)	25/02/2015						
Programa de Pós-Graduação:	Mestrado em Matemática						
Orientador(a):	Glaydston de Carvalho Bento						
CPF:	844.079.481-91	E-mail:	glaydston@mat.ufg.br				
Co-orientador(a):	-						
CPF:	-	E-mail:	-				

3. Informações de acesso ao documento:

Liberação para disponibilização?[1] [X] total [ ] parcial

Em caso de disponibilização parcial, assinale as permissões:
[ ] Capítulos. Especifique: ______________________
[ ] Outras restrições: ______________________

Havendo concordância com a disponibilização eletrônica, torna-se imprescindível o envio do(s) arquivo(s) em formato digital PDF ou DOC da tese ou dissertação.
O Sistema da Biblioteca Digital de Teses e Dissertações garante aos autores, que os arquivos contendo eletronicamente as teses e ou dissertações, antes de sua disponibilização, receberão procedimentos de segurança, criptografia (para não permitir cópia e extração de conteúdo, permitindo apenas impressão fraca) usando o padrão do Acrobat.

Pedro Bonfim de Assunção Filho
Assinatura do(a) autor(a) Data: 09/03/2015

[1] Em caso de restrição, esta poderá ser mantida por até um ano a partir da data de defesa. A extensão deste prazo suscita justificativa junto à coordenação do curso. Todo resumo e metadados ficarão sempre disponibilizados.

PEDRO BONFIM DE ASSUNÇÃO FILHO

Um Algoritmo Proximal com Quase-distância

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
2015

Ficha catalográfica elaborada automaticamente
com os dados fornecidos pelo(a) autor(a), sob orientação do Sibi/UFG.

Assunção Filho, Pedro Bonfim de
Um Algoritmo Proximal com Quase-distância [manuscrito] / Pedro Bonfim de Assunção Filho. - 2015.
lxvi, 66 f.

Orientador: Prof. Dr. Glaydston de Carvalho Bento.
Dissertação (Mestrado) - Universidade Federal de Goiás, Instituto de Matemática e Estatística (IME) , Programa de Pós-Graduação em Matemática, Goiânia, 2015.
Bibliografia.

1. Algoritmo Proximal. 2. Desigualdade de Kurdyka-Lojasiewicz. I. Carvalho Bento, Glaydston de , orient. II. Título.

PEDRO BONFIM DE ASSUNÇÃO FILHO

UM ALGORITMO PROXIMAL COM QUASE-DISTÂNCIA

Dissertação defendida no 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, aprovada no dia 25 de fevereiro de 2015, pela Banca Examinadora constituída pelos professores:

Glaydston Carvalho Bento

Prof. Dr. Glaydston de Carvalho Bento
Instituto de Matemática e Estatística - UFG
Presidente da Banca

João Xavier da Cruz Neto

Prof. Dr. João Xavier da Cruz Neto
Departamento de Matemática - UFPI

Prof. Dr. Orizon Pereira Ferreira
Instituto de Matemática e Estatística - UFG

Pedro Bonfim de Assunção Filho

Graduou–se em Matemática na UFG - Universidade Federal de Goiás. Durante o Mestrado, na UFG - Universidade Federal de Goiás, foi bolsista da CAPES.

Aos meus pais Pedro e Maria e à minha irmã Pollyanna.

Agradecimentos

Agradeço primeiramente à Deus por me conceder vida, saúde, sabedoria na escolha dos melhores caminhos, coragem e força para não desistir e concretizar este trabalho.

Agradeço à toda minha família. Em particular, aos meus pais Pedro e Maria, pelos conselhos e motivação.

Agradeço ao Prof. Glaydston de Carvalho Bento, pelos grandes ensinamentos, orientação, apoio, paciência, atenção nas revisões e sugestões na elaboração desta dissertação.

Agradeço aos professores Dr. Orizon Pereira Ferreira e Dr. João Xavier da Cruz Neto, por aceitar o convite de participar da banca de defesa e pelas sugestões dadas para a melhoria desse trabalho.

Agradeço aos meus amigos do mestrado e Doutorado que tanto me ensinaram: Aderval, Ademir, Carlos, Donizeth, Jeferson, Ricardo, Dassael, Laredo, Marcos Túlio, Marcelo Bezerra, Tiago, Jean Carlos e Valdinês.

Agradeço à minha namorada e futura esposa, Sara Ribeiro, por todo carinho, atenção e compreensão em todo o período de elaboração dessa dissertação.

Agradeço a todos os meus professores e professoras (IME-UFG). Obrigado a todos vocês pela dedicação e grande paciência quando me tiveram como aluno.

Agradeço a CAPES pela bolsa de mestrado.

"Ninguém vai bater mais forte do que a vida. Não importa como você bate e sim o quanto aguenta apanhar e continuar lutando; o quanto pode suportar e seguir em frente. É assim que se ganha."

Sylvester Stallone,
Rocky Balboa-6.

Resumo

Assunção Filho,P.B.. **Um Algoritmo Proximal com Quase-distância**. Goiânia, 2015. 66p. Dissertação de Mestrado. Departamento de Matemática, Instituto de Matemática e Estatística, Universidade Federal de Goiás.

Neste trabalho, baseado em [1, 18], estudamos a convergência do método do ponto proximal (MPP) regularizado por uma quase-distância aplicado a um problema de otimização. A função objetivo considerada não é necessariamente convexa e satisfaz a propriedade de Kurdyka-Lojasiewicz ao redor de seus pontos críticos generalizados. Mais precisamente, mostraremos que qualquer sequência limitada, gerada pelo MPP, converge a um ponto crítico generalizado.

Palavras–chave

Algoritmo Proximal, desigualdade de Kurdyka-Lojasiewicz

Abstract

Assunção Filho,P.B.. **A Proximal Algorithm with quasi-distance**. Goiânia, 2015. 66p. MSc. Dissertation. Departamento de Matemática, Instituto de Matemática e Estatística, Universidade Federal de Goiás.

In this work, based in [1, 18], we study the convergence of method of proximal point (MPP) regularized by a quasi-distance, applied to an optimization problem. The objective function considered not is necessarily convex and satisfies the property of Kurdyka-Lojasiewicz around by their generalized critical points. More specifically, we will show that any limited sequence, generated from MPP, converge the a generalized critical point.

Keywords

Proximal Algorithm, inequality Kurdyka-Lojasiewicz

Sumário

Introdução

Considere o problema de otimização:

$$\min\{f(x) \mid x \in \mathbb{R}^n\}, \tag{0-1}$$

onde $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ é uma função própria semicontinua inferior.

O método do ponto proximal MPP é um método iterativo para resolver o Problema (0-1). Ele gera uma sequência $\{x^k\} \subset \mathbb{R}^n$, como segue: dado $x^0 \in \mathbb{R}^n$,

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

onde $\{\lambda_k\}$ é uma sequência de números positivos e o conjunto que aparece em (0-2) representa o conjunto dos minimizadores irrestritos da função $f(\cdot) + \dfrac{1}{2\lambda_k}\|\cdot - x^k\|^2$. Este método foi introduzido por Martinet[16] e Rockafellar[19] na década de 70, para otimização convexa. No caso onde 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. Essa análise pode ser encontrada, por exemplo, em [13].

Nas últimas décadas, o MPP tem sido considerado em várias outras situações como, por exemplo, em otimização não convexa. Neste trabalho, será apresentado um método do ponto proximal generalizado (MPPG) que pode ser aplicado para resolver o Problema (0-1), no caso onde a função objetivo não é necessariamente convexa, mas satisfaz a desigualdade de Kurdyka-Lojasiewicz em $\bar{x} \in \operatorname{dom} \partial f$ ver; [2, 3], isto é, existem, $\eta \in]0, +\infty]$, uma vizinhaça U de $\bar{x}$ e uma função côncava e contínua $\varphi : [0, \eta[\to \mathbb{R}_+$ tal que, $\varphi(0) = 0, \quad \varphi \in C^1(0, \eta), \quad \varphi'(s) > 0, \quad s \in (0, \eta)$ e

$$\varphi'(f(x) - f(\bar{x}))\operatorname{dist}(0, \partial f(x)) \geq 1, \quad x \in U \cap [f(\bar{x}) < f < f(\bar{x}) + \eta]. \tag{0-3}$$

Dado $x^0 \in \mathbb{R}^n$ o MPPG gera uma sequência $\{x^k\} \subset \mathbb{R}^n$ como segue:

$$x^{k+1} \in \operatorname{argmin}\{f(u) + \frac{1}{2\lambda_k} q^2(x^k, u) \mid u \in \mathbb{R}^n\}, \tag{0-4}$$

onde $q(\cdot,\cdot)$ representa uma quase distância e $\{\lambda_k\}$ é uma sequência de números positivos. Note que, a regularização em (0-4) é mais natural do ponto de vista da aplicação fornecida pelo recente progresso na modelagem de processos de decisão em economia; ver [2, 3, 18].

Esse trabalho está organizado em 3 capítulos. No primeiro temos as preliminares, onde estão as notações, definições e resultados básicos que serão úteis ao longo do trabalho. Em particular, apresentaremos alguns exemplos de quase-distância e uma breve introdução à análise não diferenciável. No segundo capítulo, definiremos a desigualdade de Kurdyka-Lojasiewicz e apresentaremos exemplos de funções que a cumprem. No terceiro e último capítulo, serão apresentados os resultados de convergência parcial e total embasados em; [1, 3, 18]. Faremos também uma análise da taxa de convergência do método como feito em [1, 3].

CAPÍTULO 1

Preliminares

Neste capítulo, apresentaremos algumas definições básicas, e resultados importantes para o desenvolvimento do trabalho. Daremos a definição de quase-distância, subdiferencial de Fréchet e subdiferencial limite, bem como algumas propriedades relevantes dos subdiferenciais e exemplos dos mesmos.

1.1 Definições e Resultados Básicos

O produto escalar de $\mathbb{R}^n$ e a sua correspondente norma são denotados por $\langle \cdot, \cdot \rangle$ e $\| \cdot \|$.

Definição 1.1 *Considere a função $f : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$.*

1) f é dita própria se o seu domínio efetivo, escrito por:

$$\mathrm{dom}(f) := \{x \in \mathbb{R}^n \mid f(x) < +\infty\},$$

for diferente do vazio;

2) f é semicontinua inferior em um ponto $\bar{x} \in \mathrm{dom}(f) \subset \mathbb{R}^n$, se para qualquer sequência $\{x^k\}_{k \in \mathbb{N}}$ em $\mathbb{R}^n$, com x^k convergindo a $\bar{x}$, tem-se:

$$\liminf_{k \to \infty} f(x^k) \geq f(\bar{x}).$$

Além disso, f é semicontínua inferior em $\mathbb{R}^n$, se é semicontínua inferior em todos os pontos de $\mathbb{R}^n$;

3) O gráfico de f é dado por:

$$\mathrm{Graf}(f) := \{(x,z) \in \mathbb{R}^n \times \mathbb{R} \mid z = f(x)\};$$

4) Se $\mathrm{dom}(f) \neq \emptyset$, o conjunto dos minimizadores globais de f, possivelmente vazio, é denotado por:

$$\arg\min f := \{x \in \mathbb{R}^n | f(x) = \inf f\};$$

5) f é dita ser coerciva se, e somente se, para cada sequência $\{x^k\}_{k\in\mathbb{N}}$ em $\mathbb{R}^n$ com $\lim_{k\to+\infty} \| x^k \| = +\infty$, tem-se:

$$\limsup_{k\to+\infty} f(x^k) = +\infty;$$

6) O epígrafo de f é dado por:

$$epi(f) := \{(x,\alpha) \in \mathbb{R}^n \times \mathbb{R} \mid f(x) \leq \alpha\};$$

7) O conjunto de nível da função é dado por:

$$\mathcal{L}_f(c) = \{x \in \mathbb{R}^n \mid f(x) \leq c\}.$$

Definição 1.2 *A função indicadora δ_C de um conjunto $C \subset \mathbb{R}^n$ é dada por:*

$$\delta_C(x) = 0 \quad se \quad x \in C, \qquad \delta_C(x) = +\infty \quad se \quad x \notin C.$$

Observação 1.3 *A função constante 0 é a indicadora de $C = \mathbb{R}^n$, enquanto que a função $+\infty$ é a indicadora de $C = \emptyset$. Obviamente $dom(\delta_C) = C$, e δ_C é própria se, e somente se, C é não-vazio.*

Vejamos alguns resultados elementares, porém de relevante importância para o decorrer deste trabalho.

Teorema 1.4 *Seja $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. Então, f é semicontínua inferior se, e somente se, $epi(f)$ é um conjunto fechado.*

Demonstração. Tome $(\bar{x},\alpha) \in \mathbb{R}^n \times \mathbb{R}$ e $\{(x^k,\alpha_k)\} \subset epi(f)$ uma sequência tal que:

$$(x^k,\alpha_k) \longrightarrow (\bar{x},\alpha).$$

Visto que f é semicontínua inferior, em particular f é semicontínua inferior em $\bar{x}$ e, consequentemente,

$$\liminf_{k\to+\infty} f(x^k) \geq f(\bar{x}). \tag{1-1}$$

Por outro lado, como $\{(x^k,\alpha_k)\} \subset epi(f)$, temos que:

$$\alpha_k \geq f(x^k), \quad k \in \mathbb{N}. \tag{1-2}$$

Tomando o limite inferior em k na última desigualdade e levando em conta que $\alpha_k \to \alpha$ quando k vai a $+\infty$, obtemos:

$$\alpha = \liminf_{k\to+\infty} \alpha_k \geq \liminf_{k\to+\infty} f(x^k) \geq f(\bar{x}),$$

onde a última desigualdade segue a partir de (1-2). Mas isto nos diz que $(\bar{x}, \alpha) \in epi(f)$ e, consequentemente, que $epi(f)$ é fechado.

Reciprocamente, dado $\hat{x} \in \mathbb{R}^n$, suponhamos que exista uma sequência $\{x^k\} \subset \mathbb{R}^n$ tal que $x^k \to \hat{x}$, quando $k \to +\infty$, mas

$$\liminf_{k\to+\infty} f(x^k) < f(\hat{x}).$$

Então, existe uma subsequência $\{x^{k_j}\}$ de $\{x^k\}$ tal que:

$$L = \lim_{j\to+\infty} f(x^{k_j}) = \liminf_{j\to+\infty} f(x^j) < f(\hat{x}). \tag{1-3}$$

Agora, uma vez que, $(x^{k_j}, f(x^{k_j})) \in epi(f)$, para todo $k \in \mathbb{N}$, $(x^{k_j}, f(x^{k_j})) \longrightarrow (\hat{x}, L)$, e, por hipótese, $epi(f)$ é um conjunto fechado, então $(\hat{x}, L) \in epi(f)$, o que contradiz (1-3). Logo, $\liminf_{k\to+\infty} f(x^k) \geq f(\hat{x})$ e a prova do teorema está concluída. □

Teorema 1.5 *Se* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *é uma função coerciva, para todo* $c \in \mathbb{R}$ *tal que* $\mathcal{L}_f(c)$ *é limitado.*

Demonstração. Visto que f é coerciva, temos:

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

Assim, dado $c > 0$ existe $r > 0$ tal que,

$$\|x\| > r \Rightarrow f(x) > c.$$

Mas, isto nos diz que $\mathcal{L}_f(c) \subset B(0,r)$ (bola de centro na origem e raio r) e o resultado segue. □

Teorema 1.6 *Se* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *é coerciva e semicontínua inferiormente então* $\mathcal{L}_f(c)$ *é compacto para todo* $c \in \mathbb{R}$.

Demonstração. Suponha que $\mathcal{L}_f(c)$ não é fechado para algum $c \in \mathbb{R}$, então existe uma sequência $\{x^k\}_{k\in\mathbb{N}} \in \mathcal{L}_f(c)$ com, $\lim_{k\longrightarrow\infty} x^k = x \in \mathbb{R}^n$ e $x \notin \mathcal{L}_f(c)$. Como f é semi-contínua inferior;

$$\liminf_{k\longrightarrow\infty} f(x^k) \geq f(x) > c.$$

O que não pode acontecer, pois a sequência $\{x^k\}_{k\in\mathbb{N}} \subset \mathcal{L}_f(c)$. Logo, $\mathcal{L}_f(c)$ é fechado e pelo Teorema 1.5 é limitado. $\square$

O próximo resultado assegura que quando f é coerciva, semicontínua inferior e limitada inferiormente, temos que $\mathrm{argmin}\{f(x) \mid x \in \mathbb{R}^n\}$ é diferente do vazio e compacto. Além disso, o resultado será de grande importância, pois irá garantir a boa definição para o MPPG.

Teorema 1.7 *Seja $c \in \mathbb{R}$ tal que $\mathcal{L}_f(c) \neq \emptyset$. Se f é coerciva, semicontínua inferior e limitada inferiormente. Então:*

$$\mathrm{argmin}\{f(x) \mid x \in \mathcal{L}_f(c)\} = \mathrm{argmin}\{f(x) \mid x \in \mathbb{R}^n\}.$$

Demonstração. Seja $\inf_{x\in\mathcal{L}_f(c)} f(x) = \alpha$. Para cada $k \in \mathbb{N}$ existe $x^k \in \mathcal{L}_f(c)$ tal que, $f(x^k) \in [\alpha, \alpha + \frac{1}{k}]$. Logo, existe $\left\{x^k\right\} \subset \mathcal{L}_f(c)$ tal que, $\lim_{k\to+\infty} f(x^k) = \alpha$. Por outro lado, $\mathcal{L}_f(c)$ é compacto e assim, existe $\left\{x^{k_j}\right\}$ subsequência de $\left\{x^k\right\}$ tal que $x^{k_j} \longrightarrow \bar{x} \in \mathcal{L}_f(c)$ quando $j \longrightarrow +\infty$. Como f é semicontínua inferior em $\bar{x}$, então, $\liminf_{j\to+\infty} f(x^{k_j}) \geq f(\bar{x})$ e uma vez que

$$\alpha = \lim_{k\to+\infty} f(x^k) = \lim_{j\to+\infty} f(x^{k_j}) \geq f(\bar{x}),$$

segue que $\alpha = f(\bar{x})$ pois, pela propriedade de ínfimo $\alpha \leq f(\bar{x})$. Assim, existe $\bar{x} \in \mathcal{L}_f(c)$ tal que $\inf_{x\in\mathcal{L}_f(c)} f(x) = f(\bar{x})$. Logo,

$$f(\overline{x}) \leq f(x), \quad x \in \mathcal{L}_f(c).$$

Além disso, para $x \in \mathbb{R}^n \backslash \mathcal{L}_f(c)$, temos $f(x) \geq c$. Então, $f(\overline{x}) \leq f(x)$, para todo $x \in \mathbb{R}^n$. $\square$

1.2 Quase-distância

Nesta seção, introduziremos a definição de quase-distância e alguns exemplos. Ainda apresentaremos uma condição sobre a quase-distância, que será assumida como hipótese no Capítulo 3. Ao assumirmos essa condição demonstraremos resultados fundamentais, como por exemplo, a coercividade da função q^2, que utilizaremos como função regularizadora do MPPG.

Definição 1.8 *Seja X um conjunto. A aplicação $q : X \times X \to \mathbb{R}_+$ é dita uma quase-distância se, para todo $x, y, z \in X$, temos:*

1. $q(x,y) = q(y,x) = 0 \quad \Leftrightarrow x = y,$
2. $q(x,z) \leq q(x,y) + q(y,z).$

O par (X,q) é dito espaço quase-métrico. Observe que se $q : X \times X \to \mathbb{R}_+$ é uma aplicação simétrica, isto é,

$$q(x,y) := q(y,x), \quad x,y \ \in X,$$

então q é uma distância em X. Além disso, para cada quase-distância q, denotamos $\overline{q}$ a sua quase-distância conjugada. Definida por

$$\overline{q}(x,y) = q(y,x), \quad x,y \in X,$$

e ainda definimos a distância $\hat{q}$ dada por:

$$\hat{q}(x,y) = \max\{q(x,y), \overline{q}(x,y)\}, \quad x,y \in X,$$

a distância associada à quase-distância q.

Definição 1.9 *Considere a função $f : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$. Dizemos que f é localmente lipschitz contínua em uma vizinhaça de $\overline{x} \in dom(f)$ se existem, $\varepsilon > 0$ e $\ell > 0$ tais que:*

$$|f(x) - f(y)| \leq \ell \|x - y\|, \quad x,y \in B(\overline{x}, \varepsilon).$$

Além disso , f é dita localmente Lipschitz contínua em um subconjunto aberto $D \subset \mathbb{R}^n$ se f é localmente Lipschitz contínua em cada $\overline{x} \in D$;

Definição 1.10 *Considere a função $f : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$.*

1) f é convexa se, e somente se, para todo $x,y \in dom(f)$ e todo $t \in [0,1]$, tem-se:

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y);$$

2) f é côncava se, e somente se, $-f$ é convexa;

Nesse trabalho, consideraremos apenas quase distâncias $q : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ que satisfazem a seguinte condição:

$$\exists \ \alpha, \beta > 0 : \quad \alpha\|x - y\| \leq q(x,y) \leq \beta\|x - y\|, \quad x,y \in \mathbb{R}^n. \tag{1-4}$$

Agora, apresentaremos alguns exemplos de quase-distância:

Exemplo 1.11 *[11] Seja $q : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_+$ definida por:*

$$q(x,y) = \begin{cases} 0, & se \quad x \leq y; \\ 1, & se \quad x > y. \end{cases}$$

Note que, q é uma quase-distância em $\mathbb{R}$. Com efeito, dados $x, y \in \mathbb{R}$,

$$q(x,y) = q(y,x) = 0 \Leftrightarrow x \leq y \ \ e \ \ y \leq x \Leftrightarrow x = y.$$

Agora, dados $x, y, z \in \mathbb{R}$, é fácil ver que

$$q(x,z) \leq q(x,y) + q(y,z),$$

para qualquer posição escolhida para x, y, z em $\mathbb{R}$. Observe ainda que a quase-distância q, não verifica a condição (1-4). De fato, para $\alpha > 0$ arbitrário, existem $x = \dfrac{3}{\alpha}$ e $y = \dfrac{1}{\alpha}$, tal que:

$$q(x,y) = 1, \quad mas \quad \alpha|x-y| = 2.$$

Exemplo 1.12 *[12] Dizemos que $p : \mathbb{R}^n \to \mathbb{R}_+$, é uma norma assimétrica se, para todo $x, y \in \mathbb{R}^n$ e $\alpha \in \mathbb{R}_+$, as seguintes condições ocorrem:*

1. $p(x) = p(-x) = 0 \Leftrightarrow x = 0;$
2. $p(\alpha x) = \alpha p(x);$
3. $p(x+y) \leq p(x) + p(y).$

A partir da definição de p, é fácil ver que $q : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$, dada por:

$$q(x,y) := p(x-y), \quad x, y \in \mathbb{R}^n,$$

é uma quase-distância em $\mathbb{R}^n$.

Exemplo 1.13 *Considere $c^-, c^+ > 0$ e defina $q : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_+$ definida por*

$$q(x,y) = \begin{cases} c^+(y-x), & se \quad y - x > 0, \\ c^-(x-y), & se \quad y - x \leq 0. \end{cases}$$

Afirmamos que q é uma quase-distância em $\mathbb{R}$. De fato,

$$q(x,y) = q(y,x) = 0 \Leftrightarrow x \leq y \ \ e \ \ y \leq x \Leftrightarrow x = y.$$

Agora dados $x, y, z \in \mathbb{R}$, observe que ao considerarmos as possíveis comparações para esses números, chegamos à conclusão que a desigualdade triangular é sempre verificada.
Se $x \leq y \leq z : q(x,y) = c^+(y-x) = c^+(z-x) - c^+(z-y) \leq q(x,z) + q(z,y);$
Se $x < z < y : q(x,y) = c^+(y-x) = c^+(z-x) + c^+(z-y) = q(x,z) + q(z,y);$

Se $z < y < x : q(x,y) = c^-(x-y) = c^-(x-z) - c^-(y-z) \le q(x,z) + q(z,y)$;
Se $z < x < y : q(x,y) = c^+(y-x) = c^+(y-z) - c^+(x-z) \le q(x,z) + q(z,y)$;
Se $y < x < z : q(x,y) = c^-(x-y) = c^-(z-y) - c^-(z-x) \le q(x,z) + q(z,y)$;
Se $y < z < x : q(x,y) = c^-(x-y) = c^-(x-z) + c^-(z-y) = q(x,z) + q(z,y)$;
Do mesmo modo, a aplicação $h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$, *definida por:*

$$h(x,y) = \sum_{i=1}^{n} q_i(x_i,y_i),$$

é uma quase distância em $\mathbb{R}^n$. *Observe que agora estamos fazendo*

$$q_i(x_i,y_i) = \begin{cases} c_i^+(y_i - x_i), & se \quad y_i - x_i > 0, \\ c_i^-(x_i - y_i), & se \quad y_i - x_i \le 0, \end{cases}$$

onde x_i, y_i *representa a i-ésima componentes dos vetores* x *e* y *em* $\mathbb{R}^n$. *Além disso, teremos* n *constantes* c_i^+ *e* c_i^-. *A veracidade de* $h(x,y) = h(y,x) = 0 \quad \Leftrightarrow x = y$, *é de fácil verificação. Por outro lado, para cada* $x,y,z \in \mathbb{R}^n$, *tem-se:*

$$h(x,y) = \sum_{i=1}^{n} q_i(x_i,y_i) \le \sum_{i=1}^{n} [q_i(x_i,z_i) + q_i(z_i,y_i)] = h(x,z) + h(z,y).$$

Desse modo, h *é uma quase-distância. Para verificar a condição (1-4), basta considerarmos,* $\alpha = \min_{i=1,\ldots,n} \{c_i^-, c_i^+\}$ *e* $\beta = \max_{i=1,\ldots,n} \{c_i^-, c_i^+\}$. *Logo,*

$$\alpha \sum_{i=1}^{n} |x_i - y_i| \le h(x,y) \le \beta \sum_{i=1}^{n} |x_i - y_i|,$$

visto que a norma $\|\cdot\|_1 \left(\|x\|_1 = \sum_{i=1}^{n} |x_i| \right)$ *e a norma usual* $\|\cdot\|$ *são equivalentes.*

Exemplo 1.14 *Seja* $g : X \to \mathbb{R}_+$ *uma função injetiva, para* $x,y \in X$ *defina,* $q_g(x,y) = max\{g(x) - g(y), 0\}$. *Então,* q_g *gera uma quase distância em* X. *De fato, a veracidade de* $q(x,y) = q(y,x) = 0 \quad \Leftrightarrow x = y$, *é de fácil verificação. Vejamos agora a desigualdade triângular, para* $x,y,z \in X$ *teremos as quase-distâncias abaixo:*

$$\begin{aligned} q_g(x,z) &= \max\{g(x) - g(z), 0\}, \\ q_g(x,y) &= \max\{g(x) - g(y), 0\}, \\ q_g(y,z) &= \max\{g(y) - g(z), 0\}. \end{aligned}$$

Em qualquer combinação, sempre será válido a desigualdade:

$$q_g(x,z) \le q_g(x,y) + q_g(y,z).$$

Exemplo 1.15 *Seja $\mu > 0$ e $g : \mathbb{R}^n \to \mathbb{R}$ uma função injetiva e Lipschitz com contante ℓ. Então, para $x, y \in \mathbb{R}^n$, definimos*

$$q(x,y) = \max\{g(x) - g(y), \mu\|x-y\|\}. \tag{1-5}$$

Assim, q gera uma quase distância em $\mathbb{R}^n$. De fato, a partir de (1-5)

$$q(x,y) = q(y,x) = 0 \Leftrightarrow x = y.$$

Além disso, quando $q(x,y) = g(x) - g(y)$,

$$\begin{aligned} q(x,y) &= g(x) - g(z) + g(z) - g(y), \\ &\leq \max\{g(x) - g(z), \mu\|x-z\|\} + \max\{g(z) - g(y), \mu\|z-y\|\}, \\ &= q(x,z) + q(z,y). \end{aligned}$$

Agora, quando $q(x,y) = \mu\|x-y\|$,

$$\begin{aligned} q(x,y) &\leq \mu\|x-z\| + \mu\|z-y\|, \\ &\leq \max\{g(x) - g(z), \mu\|x-z\|\} + \max\{g(z) - g(y), \mu\|z-y\|\}, \\ &= q(x,z) + q(z,y). \end{aligned}$$

Assim,

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

e a afirmação está provada. Além disso, considerando $\alpha = \mu$ e $\beta = \ell + \mu$, temos que:

$$\alpha\|x-y\| \leq q(x,y) \leq \beta\|x-y\|, \quad x,y \in \mathbb{R}^n.$$

As quase-distâncias dos Exemplos 1.13 e 1.15 verificam a condição (1-4), e as quase-distâncias dos Exemplos 1.14 e 1.11 não verificam. Em seguida, serão apresentados alguns resultados que são obtidos considerando a condição (1-4).

Proposição 1.16 *Se $q : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ é uma quase-distância que verifica (1-4), então para todo $\bar{z} \in \mathbb{R}^n$ as funções $q(\bar{z}, \cdot)$ e $q(\cdot, \bar{z})$ são Lipschitz.*

Demonstração. Para todo $x, y \in \mathbb{R}^n$, temos:

$$q(\bar{z},x) \leq q(\bar{z},y) + q(y,x) \Rightarrow q(\bar{z},x) - q(\bar{z},y) \leq q(y,x),$$

consequentemente podemos escrever:

$$|q(\bar{z},x) - q(\bar{z},y)| \leq q(y,x) \leq \max\{q(x,y), q(y,x)\} \leq q(x,y) + q(y,x),$$

por (1-4), existe $M = 2\beta > 0$ tal que,

$$|q(\bar{z},x) - q(\bar{z},y)| \leq M\|x-y\|, \quad x,y \in \mathbb{R}^n.$$

Assim, $q(\bar{z},\cdot)$ é Lipschitz. A prova para $q(\cdot,\bar{z})$ é análoga. □

Proposição 1.17 *Seja $\bar{z} \in \mathbb{R}^n$ fixo. Se q verifica (1-4), então $q^2(\bar{z},\cdot)$ e $q^2(\cdot,\bar{z})$ são funções localmente Lipschitz em $\mathbb{R}^n$.*

Demonstração. Seja $\bar{x}$ um ponto do $\mathrm{dom}(q^2(\bar{z},\cdot)) = \mathbb{R}^n$ e $\varepsilon > 0$. Se $w \in B(\bar{x},\varepsilon)$, então existe $L_{\bar{x}} > 0$ tal que:

$$|q(\bar{z},w)| \leq L_{\bar{x}}. \tag{1-6}$$

De fato,

$$|q(\bar{z},w)| = |q(\bar{z},w) - q(\bar{z},\bar{z})|,$$

e, pela Proposição 1.16, temos:

$$\begin{aligned} |q(\bar{z},w)| &\leq M\|w - \bar{x} + \bar{x} - \bar{z}\|, \\ &\leq M(\|w-\bar{x}\| + \|\bar{z}-\bar{x}\|). \end{aligned}$$

Considerando $L_{\bar{x}} = M(\varepsilon + \|\bar{z}-\bar{x}\|)$. Logo, para $x,y \in B(\bar{x},\varepsilon)$, e novamente pela Proposição 1.16, juntamente com (1-6) obtemos:

$$\begin{aligned} |q^2(\bar{z},x) - q^2(\bar{z},y)| &= |q(\bar{z},x) + q(\bar{z},y)||q(\bar{z},x) - q(\bar{z},y)| \\ &\leq 2L_{\bar{x}}M\|x-y\|. \end{aligned}$$

Como $\bar{x}$ é um ponto arbitrário de $\mathbb{R}^n$, $q^2(\bar{z},\cdot)$ é localmente Lipschitz em $\mathbb{R}^n$. Para a função $q^2(\cdot,\bar{z})$ a prova é análoga. □

Se q é uma quase-distância que verifica (1-4), então, para cada $\bar{z} \in \mathbb{R}^n$, as funções $q(\bar{z},\cdot)$, $q(\cdot,\bar{z})$, $q^2(\bar{z},\cdot)$ e $q^2(\cdot,\bar{z})$ são coercivas. De fato, como

$$\alpha\|x-y\| \leq q(x,y) \leq \beta\|x-y\|, \quad x,y \in \mathbb{R}^n,$$

daí, ao fazermos $y = \bar{z}$,

$$\alpha\|x - \bar{z}\| \leq q(x, \bar{z}) \leq \beta\|x - \bar{z}\|.$$

Tomando o limite com $\|x\| \to +\infty$, obtemos:

$$\lim_{\|x\| \to +\infty} \alpha\|x - \bar{z}\| = +\infty \Rightarrow \lim_{\|x\| \to +\infty} q(x, \bar{z}) = +\infty.$$

O mesmo pode ser feito para as demais funções.

Proposição 1.18 *Seja $\bar{z} \in \mathbb{R}^n$ e $\lambda > 0$. Se $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ limitada inferiormente e $q^2(\bar{z}, \cdot)$ é uma função coerciva, então a função $f + \frac{1}{\lambda} q^2(\bar{z}, \cdot) : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ é coerciva.*

Demonstração. Como f é limitada inferiormente, existe $\kappa \in \mathbb{R}$ tal que

$$-\infty < \kappa + \frac{1}{\lambda} q^2(\bar{z}, x) \leq \left(f + \frac{1}{\lambda} q^2(\bar{z}, \cdot) \right)(x).$$

Assim, desta desigualdade, temos,

$$\lim_{\|x\| \longrightarrow +\infty} \left(f + \frac{1}{\lambda} q^2(\bar{z}, \cdot) \right)(x) = +\infty.$$

□

Combinando o Teorema 1.7 e a Proposição 1.18, obteremos a boa definição do MPPG. Em seguida, apresentaremos uma teoria de subdiferencial que juntamente com os resultados apresentados até aqui e a desigualdade de Kurdyka-Lojasiewicz, obtemos a estrutura básica desse trabalho.

1.3 Subdiferencial

Como a função objetivo tratada nesse trabalho não é necessariamente convexa, precisaremos estabelecer uma generalização do subgradiente, justificando assim a necessidade de estudarmos os subdiferenciais de Fréchet, limite e horizonte. Tais subdiferenciais nos permitirão caracterizar os pontos críticos da função objetivo considerada. Consideraremos $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ uma função, própria e semicontínua inferior. Os fatos em seguida podem ser encontrados em [20].

Definição 1.19 *Seja $f : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$.*

(a) Para cada $\bar{x} \in dom f$, o subdiferencial de Fréchet de f em $\bar{x}$, escrito por $\hat{\partial} f(\bar{x})$, é o conjunto de vetores $v \in \mathbb{R}^n$ tal que:

$$\liminf_{\substack{y \neq \bar{x} \\ y \to \bar{x}}} \frac{1}{\| y - \bar{x} \|} \left[f(y) - f(\bar{x}) - \langle v, y - \bar{x} \rangle \right] \geq 0. \tag{1-7}$$

Se $\bar{x} \notin dom f$, então $\hat{\partial} f(\bar{x}) = \emptyset$;
(b) O subdiferencial limite de f em $\bar{x} \in \mathbb{R}^n$, escrito por, $\partial f(\bar{x})$, é definido como a seguir

$$\partial f(\bar{x}) = \{ v \in \mathbb{R}^n \mid \exists\, x^k \to \bar{x}, f(x^k) \to f(\bar{x}), v^k \in \hat{\partial} f(x^k) \to v \};$$

(c) O subdiferencial horizonte de f em $\bar{x}$, escrito por $\partial^\infty f(\bar{x})$ é definido como em (b), exceto que em vez de $v^k \to v$, temos $\lambda_k v^k \to v$, para alguma sequência $\lambda_k \downarrow 0$ (isto é $\lambda_k \to 0$ e $\lambda_k \geq 0, \quad k \in \mathbb{N}$).

Observação 1.20 *Note que para cada $\bar{x} \in dom(f)$ e $v \in \mathbb{R}^n$, a expressão:*

$$f(y) \geq f(\bar{x}) + \langle v, y - \bar{x} \rangle + o(\| y - \bar{x} \|), \tag{1-8}$$

onde $\lim\limits_{y \to \bar{x}} \dfrac{o(\|y - \bar{x}\|)}{\| y - \bar{x} \|} = 0$, é equivalente a $v \in \hat{\partial} f(\bar{x})$; ver [21, pg.142].

O seguinte resultado segue da Definição 1.19.

Proposição 1.21 *Tome $x \in dom f$. Então:*

$$\hat{\partial} f(x) \subset \partial f(x). \tag{1-9}$$

Além disso, $\hat{\partial} f(x)$ é convexo e fechado e $\partial f(x)$ é fechado.

Demonstração. Considere $v \in \hat{\partial} f(\bar{x})$. Tomando $x^k = \bar{x}$ e $v^k = v$, para todo $k \in \mathbb{N}$, segue que:

$$x^k \to \bar{x}, \quad f(x^k) \to f(\bar{x}) \quad \text{e} \quad v^k \to v,$$

onde $v^k \in \hat{\partial} f\left(x^k\right)$. Mas isto nos diz que $v \in \partial f(\bar{x})$, o que mostra a inclusão (1-9).

Agora, tome $v_1,\ v_2 \in \hat{\partial} f(\bar{x})$ e $v_t = (1-t)v_1 + tv_2$, com $t \in [0,1]$. Segue da equação (1-8) que:

$$f(y) - f(\bar{x}) - \langle v_t, y - \bar{x} \rangle + o(\| y - \bar{x} \|) = f(y) - f(\bar{x}) - \langle (1-t)v_1 + tv_2, y - \bar{x} \rangle + \\ + o(\| y - \bar{x} \|). \tag{1-10}$$

Escrevendo de maneira conveniente o produto interno em (1-10), obtemos:

$$f(y) - f(\bar{x}) - \langle v_t, y - \bar{x} \rangle + o(\| y - \bar{x} \|) = (1-t)[f(y) - f(\bar{x}) - \langle v_1, y - \bar{x} \rangle + o(\| y - \bar{x} \|)] + \\ + t[f(y) - f(\bar{x}) - \langle v_2, y - \bar{x} \rangle + o(\| y - \bar{x} \|)]. \tag{1-11}$$

Visto que $v_1, v_2 \in \hat{\partial} f(\bar{x})$, a expressão (1-11) combinada com relação (1-8) nos fornece $f(y) - f(\bar{x}) - \langle v_t, y - \bar{x}\rangle + o(\| y - \bar{x} \|) \geq 0$. Consequentemente, $v_t \in \hat{\partial} f(x)$ e o conjunto $\hat{\partial} f(x)$ é provado ser convexo.

Mostraremos agora que $\hat{\partial} f(\bar{x})$ é fechado. Tome $\left\{v^j\right\} \subset \hat{\partial} f(\bar{x})$, com $v^j \to v$. Para cada j, tem-se:

$$f(y) - f(\bar{x}) - \langle v^j, y - \bar{x}\rangle + o(\| y - \bar{x} \|) \geq 0.$$

Tomando o limite nessa última desigualdade com $j \to +\infty$, obtemos $v \in \hat{\partial} f(\bar{x})$. Logo, $\hat{\partial} f(\bar{x})$ é fechado.

Finalmente mostraremos que $\partial f(\bar{x})$ é fechado. Seja $v^j \to v$ uma sequência de pontos de $\partial f(\bar{x})$. Então, para todo $j \in \mathbb{N}$, existem sequências $\{x_k^j\} \subset \mathbb{R}^n$ e $\{v_k^j\} \subset \hat{\partial} f\left(x_k^j\right)$ com:

$$v_k^j \to v^j, \qquad x_k^j \to \bar{x} \quad \text{e} \quad f(x_k^j) \to f(\bar{x}),$$

escolhendo $\{v_k^j\} \subset \hat{\partial} f\left(x_k^j\right)$ tal que:

$$\|v_k^j - v^j\| < \frac{1}{2j}. \tag{1-12}$$

Visto que $v^j \to v$, existe $N \in \mathbb{N}$ tal que:

$$\|v^j - v\| < \frac{1}{2j}, \qquad j \geq N. \tag{1-13}$$

De (1-12) e (1-13) segue que:

$$\|v_k^j - v\| = \|v_k^j - v^j + v^j - v\| \leq \|v_k^j - v^j\| + \|v^j - v\| < \frac{1}{2j} + \frac{1}{2j} = \frac{1}{j}, \qquad j \in \mathbb{N}.$$

Mas, isto nos diz que existe uma sequência $\{v_k^j\} \subset \hat{\partial} f\left(x_k^j\right)$ talque:

$$v_k^j \to v.$$

Além disso, $x_k^j \to \bar{x}$ com $f(x_k^j) \to f(\bar{x})$. Logo, $v \in \partial f(\bar{x})$ e segue o resultado. □

O seguinte resultado, mostra que quando f é uma função convexa, os subdiferenciais de Fréchet e limite coincidem.

Proposição 1.22 *Seja $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ uma função convexa tal que $\bar{x} \in dom(f)$. Então,*

$$\hat{\partial} f(\bar{x}) = \{v \in \mathbb{R}^n \mid f(y) - f(\bar{x}) \geq \langle v, y - \bar{x}\rangle, \quad y \in \mathbb{R}^n\} = \partial f(\bar{x}). \tag{1-14}$$

Demonstração. Mostremos inicialmente a primeira igualdade em (1-14). Dado $u \in \{v \in \mathbb{R}^n \mid f(y) - f(\bar{x}) \geq +\langle v, y - \bar{x}\rangle, \quad y \in \mathbb{R}^n\}$. Da expressão (1-7) temos $u \in \hat{\partial} f(\bar{x})$. Por outro lado, se $u \in \hat{\partial} f(\bar{x})$, pela expressão (1-8),

$$f(y) \geq f(\bar{x}) + \langle u, y-\bar{x}\rangle + o(\|y-\bar{x}\|),$$

onde $\lim\limits_{y \longrightarrow \bar{x}} \dfrac{o(\|y-\bar{x}\|)}{\|y-\bar{x}\|} = 0$. Segue que para todo $x \in \mathbb{R}^n$ e para todo $t \in (0,1]$, tem-se:

$$\begin{aligned} f((1-t)\bar{x}+tx) &\geq f(\bar{x}) + \langle u, (1-t)\bar{x}+tx-\bar{x}\rangle + o(\|(1-t)\bar{x}+tx-\bar{x}\|), \\ &\geq f(\bar{x}) + t\langle u, x-\bar{x}\rangle + o(\|t(x-\bar{x})\|). \end{aligned}$$

Como f é convexa, temos:

$$(1-t)f(\bar{x}) + tf(x) \geq f((1-t)\bar{x}+tx) \geq f(\bar{x}) + t\langle u, x-\bar{x}\rangle + o(\|t(x-\bar{x})\|),$$

e, consequentemente,

$$f(\bar{x}) - tf(\bar{x}) + tf(x) \geq f(\bar{x}) + t\langle u, x-\bar{x}\rangle + o(\|t(x-\bar{x})\|).$$

Assim,

$$f(x) \geq f(\bar{x}) + \langle u, x-\bar{x}\rangle + \frac{o(\|t(x-\bar{x})\|)}{t}.$$

Fazendo $t \longrightarrow 0^+$, obtemos:

$$u \in \{v \in \mathbb{R}^n \mid f(y) - f(\bar{x}) \geq \langle v, y-\bar{x}\rangle, \quad y \in \mathbb{R}^n\}.$$

Mostremos agora a segunda igualdade em (1-14). Pela Proposição 1.21, temos a inclusão, $\{v \in \mathbb{R}^n \mid f(y) - f(\bar{x}) \geq \langle v, y-\bar{x}\rangle, \quad y \in \mathbb{R}^n\} \subset \partial f(\bar{x})$. Agora, basta mostrarmos que:

$$\partial f(\bar{x}) \subset \{v \in \mathbb{R}^n \mid f(y) - f(\bar{x}) \geq \langle v, y-\bar{x}\rangle, \quad y \in \mathbb{R}^n\}. \tag{1-15}$$

De fato, se $u \in \partial f(\bar{x})$, existem sequências $\{x^k\} \subset \mathbb{R}^n$ e $\{v^k\} \subset \hat{\partial} f(x^k)$, tal que:

$$x^k \to \bar{x} \quad f(x^k) \to f(\bar{x}) \quad \text{e} \quad v^k \to u.$$

Como $v^k \in \hat{\partial} f(x^k)$, tem-se que para todo $y \in \mathbb{R}^n$,

$$f(y) \geq f(x^k) + \langle v^k, y-x^k\rangle.$$

Fazendo $k \to +\infty$, obtemos:

$$f(y) \geq f(\bar{x}) + \langle u, y-\bar{x}\rangle.$$

Isso prova (1-15). Consequentemente fica provado (1-14) e a prova está concluída. $\square$

Em seguida apresentaremos um exemplo, mostrando que a concavidade de uma função não garante a igualdade entre os subdiferenciais Fréchet e limite.

Exemplo 1.23 *Seja $f:\mathbb{R}\to\mathbb{R}$ a função dada por $f(x)=-|x|$. Vejamos quem são os elementos do $\hat{\partial} f(\bar{x})$. Assuma que $\bar{x}=0$ e tome $v\in\mathbb{R}$. A partir de (1-7) temos:*

$$\liminf_{\substack{y\neq\bar{x}\\ y\to\bar{x}}} \frac{1}{|y-\bar{x}|}[-|y|-vy] = \liminf_{\substack{y\neq 0\\ y\to 0}} -1-\frac{vy}{|y|}.$$

Se $v\in\hat{\partial} f(0)$ e $y>0$, a partir da última igualdade, temos que $v\leq -1$. Por outro lado, se $v\in\hat{\partial} f(0)$ e $y<0$, temos que $v\geq 1$. Considerando a interseção desses dois casos, vemos que $\hat{\partial} f(0)=\emptyset$. Se $\bar{x}>0$, fazendo o mesmo procedimento acima,

$$\liminf_{\substack{y\neq\bar{x}\\ y\to\bar{x}}} \frac{1}{|y-\bar{x}|}[f(y)-f(\bar{x})-\langle v,y-\bar{x}\rangle] = \liminf_{\substack{y\neq\bar{x}\\ y\to\bar{x}}} \frac{-|y|-vy+(1-v)\bar{x}}{|y-\bar{x}|}.$$

Consequentemente, é fácil ver que se $v\in\hat{\partial} f(\bar{x})$ e $y>0$, então $v\leq -1$. Mas, se $v\in\hat{\partial} f(x)$ e $y<0$ temos que $v\geq -1$. Considerando novamente a interseção dos dois casos, vemos que se $\bar{x}>0$, então $\hat{\partial} f(\bar{x})=\{-1\}$. Um argumento análogo mostra que se $\bar{x}<0$, então $\hat{\partial} f(\bar{x})=\{1\}$. Resumidamente, podemos escrever:

$$\hat{\partial} f(\bar{x}) = \begin{cases} \{-1\}, & se \quad \bar{x}>0,\\ \{1\}, & se \quad \bar{x}<0,\\ \emptyset, & se \quad \bar{x}=0.\end{cases} \tag{1-16}$$

Como os elementos do conjunto $\partial f(\bar{x})$ são pontos limites de sequências em $\hat{\partial} f(\bar{x})$, chegamos ao conjunto (1-17) diretamente de (1-16).

$$\partial f(\bar{x}) = \begin{cases} \{-1\}, & se \quad \bar{x}>0,\\ \{1\}, & se \quad \bar{x}<0,\\ \{-1,1\}, & se \quad \bar{x}=0.\end{cases} \tag{1-17}$$

Lema 1.24 *Seja $\sigma(z)=\left\langle w,\frac{z}{\|z\|}\right\rangle$, $z\in\mathbb{R}^n\setminus\{0\}$. Se $\sigma(z)\geq\frac{o(\|z\|)}{\|z\|}$ para qualquer z não nulo, então $w=0$.*

Demonstração. 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 σ é contínua, existe $z_0 \in C$ tal que,

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

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

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

Mas essa última desigualdade contraria a hipótese

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

Assim, segue o resultado. □

Utilizaremos o lema acima para provar a Proposição 1.25, a qual nos mostra que quando a função f é diferenciável em um ponto $\bar{x} \in \mathbb{R}^n$ o subdiferencial de Fréchet em $\bar{x}$ coincide com o vetor $\nabla f(\bar{x})$.

Proposição 1.25 *Se $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ é diferenciável em $\bar{x}$, então $\hat{\partial} f(\bar{x}) = \{\nabla f(\bar{x})\}$.*

Demonstração. Se f é diferenciável em $\bar{x}$ temos:

$$f(\bar{x} + h) = f(\bar{x}) + \langle \nabla f(\bar{x}), h \rangle + \bar{o}(\|h\|). \tag{1-18}$$

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

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

Combinando a última desigualdade com 1-18 adquirimos:

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

Mas isso nos diz que:

$$\left\langle \nabla f(\bar{x}) - v, \frac{h}{\|h\|} \right\rangle \geq \frac{\bar{o}(\|h\|)}{\|h\|},$$

e o resultado segue utilizando o Lema 1.24, com $w = \nabla f(x) - v$. □

Observe no Exemplo 1.30 que o mesmo não ocorre com o subdiferencial limite quando f é diferenciável em $\bar{x}$, assim uma condição necessária para obtermos o resultado acima para o subdiferencial limite, é supormos f continuamente diferenciável em uma vizinhança de $\bar{x}$. Como mostra seguinte proposição.

Proposição 1.26 *Se f é uma função continuamente diferenciável em uma vizinhança de $\bar{x}$, então*

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

Demonstração. Pela Proposição 1.21, $\hat{\partial} f(\bar{x}) \subset \partial f(\bar{x})$. Consequentemente, temos $\nabla f(\bar{x}) \in \partial f(\bar{x})$. Agora tome $v \in \partial f(\bar{x})$. Assim, existem sequências $\left\{x^k\right\} \subset \mathbb{R}^n$ e $\left\{v^k\right\} \subset \hat{\partial} f(x^k)$ tais que, $x^k \longrightarrow \bar{x}$ com $f(x^k) \longrightarrow f(\bar{x})$ e $v^k \longrightarrow v$. Seja $B(\bar{x}, \varepsilon)$ uma vizinhança de $\bar{x}$ onde f é continuamente diferenciável. Visto que $x^k \longrightarrow \bar{x}$, então existe um índice $N \in \mathbb{N}$, suficientemente grande, tal que:

$$x^k \in B(\bar{x}, \varepsilon), \quad k \geq N.$$

A última inclusão combinada com a Proposição 1.25, implica que :

$$v^k = \nabla f(x^k).$$

Logo, como por hipótese temos que ∇f é contínuo em $B(\bar{x}, \varepsilon)$, segue que $v^k \longrightarrow \nabla f(\bar{x})$, e assim segue o resultado desejado. $\square$

Agora apresentaremos, um caso não convexo onde os subdiferenciais Fréchet e Limite coincidem.

Exemplo 1.27 *Seja $f : \mathbb{R} \to \mathbb{R}$, dada por $f(x) = \sqrt{|x|}$. Calculemos os subdiferenciais $\hat{\partial} f(\bar{x})$ e $\partial f(\bar{x})$. Se $\bar{x} = 0$, então para todo $v \in \mathbb{R}$.*

$$\frac{f(y) - f(\bar{x}) - \langle v, y - \bar{x}\rangle}{|y - \bar{x}|} = \frac{\sqrt{|y|} - vy}{|y|} = \frac{1}{\sqrt{|y|}} - v\frac{y}{|y|} \geq \frac{1}{\sqrt{|y|}} - |v|.$$

Aplicando o limite inferior com $|y| \to 0$ na última igualdade, temos:

$$\liminf_{|y| \to 0} \frac{\sqrt{|y|} - vy}{|y|} \geq \liminf_{|y| \to 0} \left(\frac{1}{\sqrt{|y|}} - |v|\right) = +\infty.$$

Portanto, $\hat{\partial} f(0) = (-\infty, +\infty)$. Além disso, $(-\infty, +\infty) = \hat{\partial} f(0) \subset \partial f(0) \subset (-\infty, +\infty)$, o que implica que $\partial f(0) = (-\infty, +\infty)$. Por outro lado, se $\bar{x} \neq 0$, f é de classe C^1 e

$$\hat{\partial} f(\bar{x}) = \partial f(\bar{x}) = \left\{f'(\bar{x})\right\} = \left\{\frac{1}{2\sqrt{|\bar{x}|}}\right\}.$$

Assim,

$$\hat{\partial} f(\bar{x}) = \partial f(\bar{x}) = \begin{cases} \left\{ \dfrac{1}{2\sqrt{|\bar{x}|}} \right\}, & se \quad \bar{x} \neq 0, \\ (-\infty, +\infty), & se \quad \bar{x} = 0. \end{cases} \tag{1-19}$$

A seguir daremos uma condição necessária para se ter um minimizador local, e a caracterização de ponto crítico para uma função f.

Proposição 1.28 *Uma condição necessária, mas não suficiente para $\bar{x} \in \mathbb{R}^n$ ser um minimizador local de f é:*

$$0 \in \partial f(\bar{x}). \tag{1-20}$$

Demonstração. Se $\bar{x} \in \mathbb{R}^n$ é minimizador local de f, então existe $\varepsilon > 0$ e uma bola $B(\bar{x}, \varepsilon) = B$, tal que $f(y) \geq f(\bar{x})$, para todo $y \in B$. Escrevendo essa desigualdade de maneira conveniente obtemos:

$$f(y) \geq f(\bar{x}) + \langle 0, y - \bar{x} \rangle \Rightarrow 0 \in \hat{\partial} f(\bar{x}) \subset \partial f(\bar{x}),$$

o que mostra a primeira parte da Proposição. A condição não é suficiente, pois se tomarmos a função $f : \mathbb{R} \to \mathbb{R}$, dada por $f(x) = x^3$, como f assim definida é continuamente diferenciável em $(-\delta, \delta)$ para qualquer $\delta > 0$, temos:

$$\partial f(0) = \{\nabla f(0)\}.$$

Assim, $0 \in \partial f(0)$, mas 0 não é um ponto de mínimo em $(-\delta, \delta)$. □

Um ponto $\bar{x} \in \mathbb{R}^n$ que satisfaz a inclusão 1-20 é chamado ponto crítico de f . Denotamos o conjunto de pontos críticos de f por $\text{crit}(f)$.

Definição 1.29 *(Gráfico fechado) Seja $(x^k, v^k) \in Graf(T)$, T representando o operador ponto conjunto, com (x^k, v^k) convergindo para (x, v) e $f(x^k)$ convergindo para $f(x)$ então $(x, v) \in Graf(T)$.*

Em nosso estudo, utilizaremos o subdiferencial limite ∂f, pois ele cumpre a propriedade do gráfico ser fechado. O subdiferencial de Fréchet não tem o gráfico fechado, como ilustra o exemplo abaixo.

Exemplo 1.30 *Seja $f : \mathbb{R} \to \mathbb{R}$ definida por:*

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right), & se \quad x \neq 0; \\ 0, & se \quad x = 0. \end{cases}$$

Para essa função temos que, $\hat{\partial} f(0)=\{0\}$ *e* $\partial f(0)=[-1,1]$. *De fato, como f é diferenciável em* 0, *temos pela Proposição 1.25, que* $\hat{\partial} f(0)=\{f'(0)\}=\{0\}$. *Além disso, 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*

$$x^k=\frac{1}{\alpha+2k\pi},$$

temos $x^k\to 0$ *e* $f(x^k)\to 0$. *Por outro lado,*

$$f'(x^k)=\frac{2}{\alpha+2\pi k}\sin(\alpha+2\pi k)+\cos(\alpha+2\pi k).$$

Mas isto nos diz que $f'(x^k)\to-\cos\alpha=-z\in[-1,1]\setminus\{0\}$. *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 sequências* $\{x^k\}\in\mathbb{R}$ *e* $\{v^k\}\subset\hat{\partial} f(x^k)$ *tal que* $x^k\to 0$ *e* $v^k\to z$. *Agora, pelo Teorema 1.26, temos que*

$$v^k=2x^k\sin\left(\frac{1}{x^k}\right)-\cos\left(\frac{1}{x^k}\right),\quad k\in\mathbb{N}.$$

Mas isto nos diz que

$$z=\lim_{k\to+\infty}v^k\in[-1,1],$$

que é um absurdo. Logo, $\partial f(0)=[-1,1]$. *Agora mostraremos que o subdiferencial de Fréchet, não tem o gráfico fechado. Para tanto, considere a sequência:*

$$x^k=\frac{1}{\pi+2k\pi}\quad e\quad v^k=2x^k\sin\left(\frac{1}{x^k}\right)-\cos\left(\frac{1}{x^k}\right).$$

Note que,

$$(x^k,v^k)\subset Graf\left(\hat{\partial} f\right)\quad com\quad (x^k,v^k)\longrightarrow(0,1).$$

Mas como, $\hat{\partial} f(0)=\{0\}$, *temos que* $1\notin\hat{\partial} f(0)$. *Assim o* $Graf\left(\hat{\partial} f\right)$ *não é fechado.*

Agora, daremos as definições de Cone Proximal, Cone Normal Regular (ou Fréchet) e Cone Normal (ou Limite) a um conjunto fechado $C\subset\mathbb{R}^n$, os quais denotamos por, $N_C^P(\bar{x})$, $\hat{N}_C(\bar{x})$ e $N_C(\bar{x})$ respectivamente. Além disso, apresentaremos caracterizações dos subdiferenciais apresentados na Definição 1.19 em termos dos Cones, ver; [20, 17, 21, 10]. Será enunciado algumas proposições e teoremas que podem ser encontrados nas literaturas apresentadas. Provaremos resultados úteis na demonstração da Proposição 1.44 e do Teorema 1.45, os quais são essenciais na prova do resultado principal desse trabalho.

Definição 1.31 *Sejam* $C\subset\mathbb{R}^n$ *um conjunto fechado e* $\bar{x}\in C$. *Então,*

(a) $v\in N_C^P(\bar{x})$ *se existe* $M>0$ *tal que:*

$$\langle v,y-\bar{x}\rangle\le M\|y-\bar{x}\|^2,\quad y\in C.$$

Os elementos em $N_C^P(\bar{x})$ são chamados de Normais Proximais a C em $\bar{x}$.

(b) O Cone Normal Regular (ou Cone Fréchet) ao conjunto C em $\bar{x}$, é o conjunto:

$$\hat{N}_C(\bar{x}) := \{v \in \mathbb{R}^n \mid \langle v, y-\bar{x}\rangle \leq o(\|y-\bar{x}\|), \quad y \in C \cap B(\bar{x},\delta)\}, \tag{1-21}$$

para algum $\delta > 0$, sendo $o : \mathbb{R}_+ \to \mathbb{R}_+$ tal que $\lim\limits_{\varepsilon \downarrow 0} \dfrac{o(\varepsilon)}{\varepsilon} = 0$.

(c) O Cone Normal (ou Limite) a C em $\bar{x}$ é o conjunto:

$$N_C(\bar{x}) := \{v \in \mathbb{R}^n \mid \exists x^k \xrightarrow{C} \bar{x}, v^k \longrightarrow v, \quad v^k \in N_C^P(x^k), \quad k \in \mathbb{N}\}.$$

Se $v \in N_C(\bar{x})$, v é chamado um Normal Regular ao conjunto C em $\bar{x}$.

Observação 1.32 *O conjunto $\hat{N}_C(\bar{x})$ também pode ser definido como:*

$$\hat{N}_C(\bar{x}) := \{v \in \mathbb{R}^n \mid \limsup_{y \xrightarrow{C} \bar{x}} \frac{\langle v, y-\bar{x}\rangle}{\|y-\bar{x}\|} \leq 0\}. \tag{1-22}$$

Os vetores $v \in \hat{N}_C(\bar{x})$ são chamados de vetores normais regulares a C em $\bar{x}$. A notação $y \xrightarrow{C} \bar{x}$, significa que $y \longrightarrow \bar{x}$ e $y \in C$. Pode-se mostrar que as definições (1-21) e (1-22) são equivalentes.

Proposição 1.33 *Sejam $C \subset \mathbb{R}^n$ um conjunto fechado e $\bar{x} \in C$. Então:*

$$N_C^P(\bar{x}) \subset \hat{N}_C(\bar{x}).$$

Demonstração. Seja $v \in N_C^P(\bar{x})$, assim existe $M > 0$ tal que:

$$\langle v, y-\bar{x}\rangle \leq M\|y-\bar{x}\|^2, \quad y \in C.$$

Tome, $o : \mathbb{R}_+ \to \mathbb{R}_+$ definida por $o(t) = M|t|^2$, uma vez que $\|y-\bar{x}\| \geq 0$, para todo $y \in C$, então:

$$\langle v, y-\bar{x}\rangle \leq o(\|y-\bar{x}\|), \quad y \in C \cap B(\bar{x},\delta)\backslash\{\bar{x}\},$$

para algum $\delta > 0$. Além disso, tem-se que $\lim\limits_{\|y-\bar{x}\|\downarrow 0} \dfrac{o(\|y-\bar{x}\|)}{\|y-\bar{x}\|} = 0$, portanto $v \in \hat{N}_C(\bar{x})$. □

Proposição 1.34 *Sejam $C \subset \mathbb{R}^n$ um conjunto fechado e $\bar{x} \in \mathbb{R}^n$. As seguintes afirmações são equivalentes.*

1. $\bar{v} \in N_C(\bar{x})$;
2. $\exists\, x^k \xrightarrow{C} \bar{x}, \quad \bar{v}^k \longrightarrow \bar{v}, \quad$ *tal que* $\quad \bar{v}^k \in \hat{N}_C(x^k), \quad k \in \mathbb{N}$.

Demonstração. Ver [21, pg.131]. □

Proposição 1.35 *Sejam $C \subset \mathbb{R}^n$ um conjunto fechado e $\bar{x} \in C$. Então,*

$$N_C^P(\bar{x}) \subset \hat{N}_C(\bar{x}) \subset N_C(\bar{x}).$$

Demonstração. Como mostrado na Proposição 1.33, $N_C^P(\bar{x}) \subset \hat{N}_C(\bar{x})$. Falta mostrarmos a segunda inclusão acima. Dado $\bar{v} \in \hat{N}_C(\bar{x})$ tome as sequências $\{x^k\}$, $\{v^k\} \subset \mathbb{R}^n$ tal que $x^k = \bar{x}$ e $v^k = \bar{v}, \quad k \in \mathbb{N}$. Assim, temos que:

$$\exists x^k \xrightarrow{C} \bar{x}, \quad v^k \longrightarrow \bar{v}, \quad \text{de modo que} \quad v^k \in \hat{N}_C(x^k), \quad k \in \mathbb{N}.$$

Pela Proposição 1.34, $\bar{v} \in N_C(\bar{x})$. Como $v \in \hat{N}_C(\bar{x})$ foi escolhido de modo arbitrário segue que $\hat{N}_C(\bar{x}) \subset N_C(\bar{x})$. □

Teorema 1.36 *(Caracterização do Normal Regular em termos do Gradiente). Um vetor v é um Normal Regular ao conjunto C em $\bar{x}$ se, e somente se, existe uma função h que atinge seu máximo local relativo a C em $\bar{x}$, e é diferenciável em $\bar{x}$ com $\nabla h(\bar{x}) = v$.*

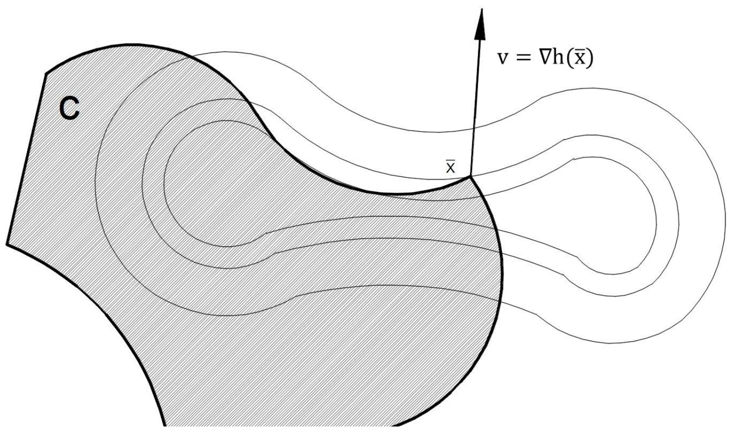

Figura 1.1: *Normal regular em termos do gradiente*

Demonstração. Se h assume o seu máximo local no conjunto C em $\bar{x}$, $h(\bar{x}) = f(\bar{x})$ e h é diferenciável em uma vizinhança de $\bar{x}$ com $\nabla h(\bar{x}) = v$, mostraremos que $v \in \hat{N}_C(\bar{x})$. Sendo h diferenciável em $\bar{x}$, temos:

$$h(x) = h(\bar{x}) + \langle \nabla h(\bar{x}), x - \bar{x} \rangle + o(|x - \bar{x}|), \tag{1-23}$$

com $\lim_{x \longrightarrow \bar{x}} \dfrac{o(|x-\bar{x}|)}{|x-\bar{x}|} = 0$, combinando a hipótese (1-23) e $h(x) \leq h(\bar{x})$, para todo x na vizinhança de $\bar{x}$, obtemos:

$$h(\bar{x}) - h(\bar{x}) = 0 \geq h(x) - h(\bar{x}) = \langle v, x - \bar{x} \rangle + o(|x - \bar{x}|), \quad x \in C \cap B(\bar{x}, \delta),$$

para algum $\delta > 0$. Logo, $v \in \hat{N}_C(\bar{x})$. Agora, se $v \in \hat{N}_C(\bar{x})$, a função

$$r \mapsto \theta_0(r) := sup\{\langle v, x - \bar{x} \rangle \mid x \in C, |x - \bar{x}| \leq r\} \leq r|v|$$

é decrescente em $[0, \infty)$ com $0 = \theta_0(0) \leq \theta_0(r) \leq o(r)$. A função

$$h_0(x) = \langle v, x - \bar{x} \rangle - \theta_0(|x - \bar{x}|),$$

é diferenciável em $\bar{x}$ com,

$$\nabla h(\bar{x}) = \nabla h_0(\bar{x}) = v \quad \text{e} \quad h(\bar{x}) = h_0(\bar{x}).$$

Mas, $h(x) < h_0(\bar{x})$ para todo $x \neq \bar{x}$ em C. Tal que, h atinge seu máximo ao longo de C unicamente em $\bar{x}$. □

Teorema 1.37 *Um vetor v pertence a $\hat{\partial} f(\bar{x})$ se, e somente se, em alguma vizinhança de $\bar{x}$ existe uma função $h \leq f$ com $h(\bar{x}) = f(\bar{x})$ tal que h é diferenciável em $\bar{x}$ com $\nabla h(\bar{x}) = v$.*

Demonstração. Primeiro suponhamos que exista uma função h diferenciável em $\bar{x}$ da seguinte forma:

$$h \leq f, \quad h(\bar{x}) = f(\bar{x}) \quad \text{e} \quad \nabla h(\bar{x}) = v. \tag{1-24}$$

Assim, temos:

$$h(x) = h(\bar{x}) + \langle \nabla h(\bar{x}), x - \bar{x} \rangle + o(|x - \bar{x}|),$$

agora utilizando o fatos em (1-24), temos:

$$f(x) - f(\bar{x}) \geq h(x) - h(\bar{x}) = \langle v, x - \bar{x} \rangle + o(|x - \bar{x}|).$$

Dessa maneira, $v \in \hat{\partial} f(\bar{x})$. Agora suponha que, dado $v \in \hat{\partial} f(\bar{x})$. Definamos:

$$h(x) = \langle v, x - \bar{x} \rangle + \theta_0(|x - \bar{x}|) + f(x),$$

onde $\theta_0 = \sup\{\langle v, x - \bar{x} \rangle \mid x \in V_\varepsilon(\bar{x}), \quad |x - \bar{x}| \leq r\} \leq r|v|$, θ_0 é não decrescente em $[0, +\infty)$, observe que $0 = \theta_0(0) \leq \theta_0(r) \leq o(r)$, além disso , $\nabla \theta_0(\bar{x}) = 0$. Logo, h é diferenciável em $\bar{x}$, $\nabla h(\bar{x}) = v$ e $h(\bar{x}) = f(\bar{x})$. □

Teorema 1.38 *Seja $f(x) = f_1(x_1) + ... + f_m(x_m)$, onde $f_i : \mathbb{R}^{n_i} \to \bar{\mathbb{R}}$ são funções semicontínuas inferiores, e $x \in \mathbb{R}^n$ é expressado como $(x_1, ..., x_m)$ com $x_i \in \mathbb{R}^{n_i}$. Então, para todo $\bar{x} = (\bar{x}_1, ..., \bar{x}_m)$ com $f(\bar{x})$ finito, temos:*

$$\begin{aligned} \hat{\partial} f(\bar{x}) &= \hat{\partial} f_1(\bar{x}_1) \times ... \times \hat{\partial} f_m(\bar{x}_m), \\ \partial f(\bar{x}) &= \partial f_1(\bar{x}_1) \times ... \times \partial f_m(\bar{x}_m). \end{aligned}$$

Demonstração. Seja $v \in \hat{\partial} f(\bar{x})$, em seguida defina:

$$h(x) = \langle v, x - \bar{x} \rangle - K(x), \quad \text{onde} \quad K(x) = \sum_{i=1}^{m} \theta_0(|x_i - \bar{x}_i|) + f_i(\bar{x}_i)$$

h é diferenciável, com $\nabla h(\bar{x}) = v$, $h < f$ e $h(\bar{x}) = f(\bar{x})$, como visto na Proposição 1.37,

$$\begin{aligned} h(x) &= \langle (v_1, ..., v_m), (x_1 - \bar{x}_1, ..., x_m - \bar{x}_m) \rangle - K(x), \\ &= \langle v_1, x_1 - \bar{x}_1 \rangle + \langle v_2, x_2 - \bar{x}_2 \rangle + ... + \langle v_m, x_m - \bar{x}_m \rangle - K(x), \end{aligned}$$

ou ainda,

$$\begin{aligned} h(x) = \langle v_1, x_1 - \bar{x}_1 \rangle - [\theta_0(|x_1 - \bar{x}_1|) + f_1(\bar{x}_1)] + ... + \langle v_m, x_m - \bar{x}_m \rangle - \\ - [\theta_0(|x_m - \bar{x}_m|) + f_m(\bar{x}_m)]. \end{aligned}$$

Logo,

$$h(x) = h_1(x_1) + h_2(x_2) + ... + h_m(x_m),$$

onde $h_i(x_i) = \langle v_i, x_i - \bar{x}_i \rangle - [\theta_0(|x_i - \bar{x}_i|) + f_i(\bar{x}_i)]$, são diferenciáveis em $\bar{x}$, com $h_i(\bar{x}_i) = f_i(\bar{x}_i)$, $1 \leq i \leq m$, assim:

$$v \in \nabla h(\bar{x}) = (v_1, ..., v_m) \in \nabla h_1(\bar{x}_1) \times ... \times h_m(\bar{x}_m) \subset \prod_{i=1}^{m} \hat{\partial} f_i(\bar{x}_i).$$

Agora se considerarmos, $v \in \hat{\partial} f_1(\bar{x}_1) \times ... \times \hat{\partial} f_m(\bar{x}_m)$, então escrevemos, $v = (v_1, ..., v_m)$, onde $v_i \in \hat{\partial} f_i(\bar{x}_i)$, $\quad 1 \leq i \leq m$, novamente pela Proposição 1.37, existem funções h_i diferenciáveis com, $h_i < f_i$, $f_i(\bar{x}_i) = h_i(\bar{x}_i)$ e $\nabla h_i(\bar{x}_i) = v_i$, dessa maneira:

$$v \in \nabla h(\bar{x}) \subset \hat{\partial} f(\bar{x}).$$

Onde, $h(x) = \langle v, x - \bar{x} \rangle - K(x)$. E assim provamos que, $\hat{\partial} f(\bar{x}) = \hat{\partial} f_1(\bar{x}_1) \times ... \times \hat{\partial} f_m(\bar{x}_m)$. Agora para a segunda parte da Proposição, tome $v \in \partial f(\bar{x})$, assim existem $x^k \longrightarrow \bar{x}$ com

$f(x^k) \longrightarrow f(\bar{x})$ e $v^k \in \hat{\partial} f(x^k)$ com $v^k \longrightarrow v$, e como já mostramos que:

$$\hat{\partial} f(\bar{x}) = \hat{\partial} f_1(\bar{x}_1) \times ... \times \hat{\partial} f_m(\bar{x}_m),$$

agora, utilizando um resultado de análise, $x^k \longrightarrow \bar{x}$ e $f(x^k) \longrightarrow f(\bar{x})$, se , e somente se, $x_i^k \longrightarrow \bar{x}_i$ e $f_i(x_i^k) \longrightarrow f_i(\bar{x}_i)$, mas precisamente uma sequência em $\mathbb{R}^n$ é convergente se, e somente se, cada uma de suas componentes é uma sequência convergente em $\mathbb{R}$. Como aqui esse fato é mais delicado, pois expressamos $\bar{x} \in \mathbb{R}^n$ como $\bar{x} = (\bar{x}_1, ..., \bar{x}_m)$ com $x_i \in \mathbb{R}^{n_i}$ o resultado ainda continua sendo satisfeito, basta observarmos que para cada $\mathbb{R}^{n_i} = \mathbb{R}_1 \times ... \times \mathbb{R}_i$. teremos esse resultado. Como, $v \in \partial f(\bar{x})$ é limite de uma sequência de pontos $v^k \in \hat{\partial} f(x^k)$, e como $\hat{\partial} f(\bar{x}) \subset \partial f(\bar{x})$, segue:

$$\partial f(\bar{x}) = \partial f_1(\bar{x}_1) \times ... \times \partial f_m(\bar{x}_m).$$

□

O próximo resultado é a Regra da Cadeia generalizada. Vamos utilizá-la na demostração do Teorema 1.40. Aqui utilizaremos o caso em que $F : \mathbb{R}^n \to \mathbb{R}^m$, é fortemente diferenciável. Introduziremos a seguinte notação para a matriz jacobiana:

$$\nabla F(x) := \left[\frac{\partial f_i}{\partial x_j}(x) \right]_{i,j=1}^{m,n} \in \mathbb{R}^{m \times n},$$

e $\nabla F(x)^*$ será sua matriz transposta.

Teorema 1.39 *(Regra da Cadeia) Suponha $f(x) = g(F(x))$. Para $g : \mathbb{R}^m \to \overline{\mathbb{R}}$ uma função pópria, semicontíniua inferior e $F : \mathbb{R}^n \to \mathbb{R}^m$ fortemente diferenciável. Então para todo $\bar{x} \in dom f = F^{-1}(dom g)$ temos:*

$$\hat{\partial} f(\bar{x}) \supset \nabla F(\bar{x})^* \hat{\partial} g(F(\bar{x})),$$

e se o único vetor, $y \in \partial^\infty g(F(\bar{x}))$ com $\nabla F(\bar{x})^* y = 0$ é $y = 0$, teremos:

$$\partial f(\bar{x}) \subset \nabla F(\bar{x})^* \partial g(F(\bar{x})).$$

Demonstração. Ver [20, pg.427]. □

Teorema 1.40 *Suponha $f = f_1 + f_2 + \cdots + f_m$, onde $f_i : \mathbb{R}^n \to \overline{\mathbb{R}}$ são funções próprias semicontínuas inferiores, e seja $\bar{x} \in dom f$, considere que a única combinação de vetores*

$v_i \in \partial^\infty f_i(\bar{x})$ *com* $v_1 + v_2 + \cdots + v_m$ *é* $v_1 = v_2 = \cdots = v_m = 0$*, então:*

$$\partial f(\bar{x}) \subset \partial f_1(\bar{x}) + \ldots + \partial f_m(\bar{x}).$$

Demonstração. Seja $F : \mathbb{R}^n \to ((\mathbb{R}^n)^m)$, uma aplicação definida por:

$$F(x) = (x, \ldots, x),$$

a qual é diferenciável. E defina a função, $g : (\mathbb{R}^n)^m \to \overline{\mathbb{R}}$ por:

$$g(x_1, \ldots, x_m) = f_1(x_1) + \ldots + f_m(x_m).$$

Então $f(x) = g(F(x))$. A Regra da Cadeia no Teorema 1.39 nos garante,

$$\partial f(\bar{x}) \subset \nabla F(\bar{x})^* \partial g(F(\bar{x})),$$

pois o único vetor $y \in \partial^\infty g(F(x))$ com $\nabla F(\bar{x})^* y = 0$ é $y = 0$. O fato das funções separáveis no Teorema 1.38 aplicado a g teremos:

$$\partial f(\bar{x}) \subset (1, \ldots, 1)(\partial f_1(\bar{x}), \ldots, \partial f_m(\bar{x})) = \partial f_1(\bar{x}) + \ldots + \partial f_m(\bar{x}).$$

O que completa a demonstração. □

Apresentaremos resultados auxiliares para a demostração do Teorema 1.44, o qual garante a limitação dos elementos do subdiferencial limite, para funções localmente Lipschitz contínua. Lembrando que quando a função é semicontínua inferior o $epi(f)$ é um conjunto fechado. Dessa maneira, não teremos problemas ao enunciar a seguinte proposição.

Proposição 1.41 *Sejam* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *semicontínua inferior,* $\bar{x} \in dom(f)$ *e* $\bar{v} \in \mathbb{R}^n$*. Então, para algum* $M > 0$ *e para todo* $(y, \alpha) \in epi(f)$

$$(\bar{v}, -1)\left[(y, \alpha) - (\bar{x}, f(\bar{x}))\right] \leq M\left[\|y - \bar{x}\|^2 + |\alpha - f(\bar{x})|^2\right], \tag{1-25}$$

se, e somente se, $(\bar{v}, -1) \in N^P_{epi(f)}(\bar{x}, f(\bar{x}))$.

Demonstração. Primeiramente consideremos, $(\bar{v}, -1) \in N^P_{epi(f)}(\bar{x}, f(\bar{x}))$. Então, existe $M > 0$ tal que:

$$\begin{aligned}(\bar{v}, -1)\left[(y, \alpha) - (\bar{x}, f(\bar{x}))\right] &\leq M\|(y, \alpha) - (\bar{x}, f(\bar{x}))\|^2, \\ &= M\|(y - \bar{x}, \alpha - f(\bar{x}))\|^2, \quad (y, \alpha) \in epi(f).\end{aligned}$$

Como as normas em $\mathbb{R}^p$, $p \in \mathbb{N}$, são equivalentes, vamos considerar em $\mathbb{R}^n \times \mathbb{R}$ a norma do máximo e sem perda de generalidade, podemos supor que:

$$\max\{\|y-\bar{x}\|, |\alpha - f(\bar{x})|\} = \|y-\bar{x}\|.$$

Logo,

$$\begin{aligned}(\bar{v},-1)\,[(y,\alpha)-(\bar{x},f(\bar{x}))] \le M\|(y-\bar{x}\|^2 &\le M\|(y-\bar{x}\|^2 + M|\alpha - f(\bar{x})|^2,\\ &= M\left[\|(y-\bar{x}\|^2 + |\alpha - f(\bar{x})|^2\right].\end{aligned}$$

Reciprocamente, vamos supor que $\bar{v}$ verifica (1-25). Considerando a norma da soma em $\mathbb{R}^n \times \mathbb{R}$. Temos:

$$\begin{aligned}(v,-1)\,[(y,\alpha)-(\bar{x},f(\bar{x}))] &\le M\left[\|y-\bar{x}\|^2 + |\alpha - f(\bar{x})|^2\right],\\ &\le M\|(y,\alpha)-(\bar{x},f(\bar{x}))\|^2\end{aligned}$$

para todo $(y,\alpha) \in \mathrm{epi}(f)$. Portanto, $(\bar{v},-1) \in N^P_{epi(f)}(\bar{x},f(\bar{x}))$. □

Teorema 1.42 *Para $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ e para qualquer ponto $\bar{x}$ em que o valor de f é finito, temos:*

$$\begin{aligned}\hat{\partial} f(\bar{x}) &= \left\{v \mid (v,-1) \in \hat{N}_{epi(f)}(\bar{x},f(\bar{x}))\right\},\\ \partial f(\bar{x}) &= \left\{v \mid (v,-1) \in N_{epi(f)}(\bar{x},f(\bar{x}))\right\},\\ \partial^\infty f(\bar{x}) &\subset \left\{v \mid (v,0) \in N_{epi(f)}(\bar{x},f(\bar{x}))\right\}.\end{aligned}$$

Na última relação teremos uma igualdade, quando f é localmente semicontínua inferior em $\bar{x}$.

Demonstração. Ver, [20, pg.305] □

Lema 1.43 *Seja f localmente Lipschitz em $\bar{x}$ com constate de Lipschitz $\ell \ge 0$. Dado $(v,-\lambda) \in N_{epi(f)}(\bar{x},f(\bar{x}))$ de forma arbitrária temos que:*

$$\|v\| \le \lambda\ell. \tag{1-26}$$

Demonstração. Se tomarmos $(v,-\lambda) \in N_{epi(f)}(\bar{x},f(\bar{x}))$ de forma arbitrária, existem

$$(x^k,\alpha_k) \xrightarrow{epi(f)} (\bar{x},f(\bar{x})) \quad \text{e} \quad (v^k,-\lambda_k) \longrightarrow (v,-\lambda),$$

com $(v^k, -\lambda_k) \in N^P_{epi(f)}(x^k, \alpha_k), \quad k \in \mathbb{N}$. Em particular, existe k_0 suficientemente grande tal que:

$$(v^k, -\lambda_k) \in N^P_{epi(f)}(x^k, \alpha_k), \quad k \geq k_0,$$

para k suficientemente grande, tal que x^k está na vizinhança onde f é Lipschitz. Então, existe $M_k > 0$ tal que

$$(v^k, -\lambda_k)\left[(y,\alpha) - (x^k, \alpha_k)\right] \leq M_k \|(y,\alpha) - (x^k, \alpha_k)\|^2, \quad k \geq k_0,$$

para qualquer $(y,\alpha) \in epi(f)$. Considere $\|\cdot\|$ como a norma do máximo. Sem perda de generalidade, podemos supor que:

$$\|(y,\alpha) - (x^k, \alpha_k)\| = \|y - x^k\|, \quad k \geq k_0.$$

Desse modo,

$$\begin{aligned}(v^k, -\lambda_k)\left[(y,\alpha) - (x^k, \alpha_k)\right] &\leq M_k \|y - x^k\|^2, \\ &\leq M_k \left[\|y - x^k\|^2 + |\alpha - \alpha_k|^2\right], \quad (y,\alpha) \in \mathrm{epi}(f) \quad \text{e} \quad k \geq k_0.\end{aligned}$$

Uma vez que, para qualquer $x \in \mathbb{R}^n$ tem-se que $(x, f(x)) \in epi(f)$, então para todo $\beta \geq 0$ segue que $(x, [f(x) + \beta]) \in epi(f)$. Além disso, temos que $\alpha_k \geq f(x^k)$. Consequentemente, existe $\beta_k \geq 0$ tal que $\alpha_k = f(x^k) + \beta_k$. Logo,

$$(v^k, -\lambda_k)\left[(x, [f(x) + \beta]) - (x^k, \alpha_k)\right] \leq M_k \left[\|x - x^k\|^2 + |f(x) + \beta - f(x^k) - \beta_k|^2\right],$$

ou seja,

$$v^k(x - x^k) - \lambda_k \left[f(x) + \beta - f(x^k) - \beta_k\right] \leq M_k \left[\|x - x^k\|^2 + |f(x) + \beta - f(x^k) - \beta_k|^2\right], \tag{1-27}$$

para todo $x \in \mathbb{R}^n$ e para todo $\beta \geq 0$. Em particular, para $x = x^k$ e $\beta > \beta_k$ temos que

$$-\lambda_k(\beta - \beta_k) \leq M_k |\beta - \beta_k|^2 = M_k(\beta - \beta_k)^2.$$

Como $\beta > \beta_k$, então, $-\lambda_k \leq M_k(\beta - \beta_k)$. Ao fazer $\beta_k \longrightarrow \beta$, temos que $-\lambda_k \leq 0$ ou seja, $\lambda_k \geq 0$ para todo $k \in \mathbb{N}$ suficientemente grande. Ou mais, $\lambda_k \geq 0$ para todo $k \in \mathbb{N}$ pois, em momento algum se fez necessária a hipótese de f ser Lipschitiziana. Uma vez que $\lambda_k \longrightarrow \lambda$, segue que $\lambda \geq 0$ e desse modo, mostramos também que se $(v, -\lambda) \in N_{\mathrm{epi}(f)}(\bar{x}, f(\bar{x}))$, então, $\lambda \geq 0$. Agora, se tomarmos $\beta = \beta_k$ e x na vizinhança de $\bar{x}$, onde f é Lipschitz, tem-se na expressão (1-27) que,

$$v^k(x-x^k)-\lambda_k\left[f(x)-f(x^k)\right]\leq M_k\left[\|x-x^k\|^2+|f(x)-f(x^k)|^2\right].$$

Logo,

$$\begin{aligned}
v^k(x-x^k) &\leq M_k\left[\|x-x^k\|^2+|f(x)-f(x^k)|^2\right]+\lambda_k\left[f(x)-f(x^k)\right],\\
&\leq M_k\left[\|x-x^k\|^2+|f(x)-f(x^k)|^2\right]+\lambda_k|f(x)-f(x^k)|,\\
&\leq M_k\left[\|x-x^k\|^2+\ell^2\|x-x^k\|^2\right]+\lambda_k\ell\|x-x^k\|.
\end{aligned}$$

Como x foi escolhido de modo arbitrário na vizinhança onde f é Lipschitziana, podemos então, fazer $x=x^k+rv^k$, com $r>0$ suficientemente pequeno. Desse modo,

$$\begin{aligned}
v^k(x^k+rv^k-x^k) &\leq M_k\left[\|x^k+rv^k-x^k\|^2+\ell^2\|x^k+rv^k-x^k\|^2\right]+\lambda_k\ell\|x^k+rv^k-x^k\|,\\
&\leq M_k\left[\|rv^k\|^2+\ell^2\|rv^k\|^2\right]+\lambda_k\ell\|rv^k\|.
\end{aligned}$$

Logo,

$$r\|v^k\|^2\leq M_k\left(r^2\|v^k\|^2+\ell^2r^2\|v^k\|^2\right)+\lambda_k r\ell\|v^k\|,$$

e isto implica que,

$$\|v^k\|^2\leq M_k r\|v^k\|^2+M_k\ell^2 r\|v^k\|^2+\lambda_k\ell\|v^k\|.$$

Faça $r\longrightarrow 0$. Logo, $\|v^k\|^2\leq\lambda_k\ell\|v^k\|$. Note que se $\|v^k\|=0,\quad k\in\mathbb{N}$ teremos, $\|v^k\|\leq\lambda_k\ell$. Suponha que $\|v^k\|>0$, então, também vale que $\|v^k\|\leq\lambda_k\ell$. Fazendo $k\longrightarrow+\infty$ tem-se que $\|v\|\leq\lambda\ell$, o que prova (1-26) □

Proposição 1.44 *Seja f localmente Lipschitz em $\bar{x}$ com constante de Lipschitz $\ell\geq 0$. Então:*

i) $\|v\|\leq\ell,\quad v\in\partial f(\bar{x})$,

ii) $\partial^\infty f(\bar{x})=\{0\}$.

Demonstração. Do Teorema 1.42 temos que, $\partial f(\bar{x})=\left\{v\mid(v,-1)\in N_{epi(f)}(\bar{x},f(\bar{x}))\right\}$. Dado $(v,-\lambda)\in N_{epi(f)}(\bar{x},f(\bar{x}))$ de forma arbitrária concluímos pelo Lema 1-26, que $\|v\|\leq\lambda\ell$. Então, dado $v\in\partial f(\bar{x})$ segue que, $(v,-1)\in N_{epi(f)}(\bar{x},f(\bar{x}))$ e, consequentemente, $\|v\|\leq\ell$. O que prova o item *i*).

Agora para o item *ii*) sabemos que $\partial^\infty f(\bar{x})=\left\{v\mid(v,0)\in N_{epi(f)}(\bar{x},f(\bar{x}))\right\}$. Dado $v\in\partial^\infty f(\bar{x})$ arbitrariamente, segue que $(v,0)\in N_{epi(f)}(\bar{x},f(\bar{x}))$. Portanto, $\|v\|\leq 0$ e isto implica que $\|v\| = 0$. Logo, $\partial^\infty f(\bar{x}) = \{0\}$. □

O próximo resultado, será utilizado para darmos uma caracterização de otimalidade em termos dos sub-diferenciais da função objetivo e a quase distância. Mais precisamente, iremos utilizá-lo no Capítulo 3.

Teorema 1.45 *Se f_1 é localmente Lipschitz contínua em $\bar{x}$, f_2 é própria semi-contínua inferior com $f_2(\bar{x})$ finito, então:*

$$\partial(f_1 + f_2)(\bar{x}) \subset \partial f_1(\bar{x}) + \partial f_2(\bar{x}).$$

Demonstração. Adicionando a função indicadora δ em uma vizinhança de $\bar{x}$, por exemplo $\overline{B(\bar{x},\varepsilon)}$ (bola fechada de centro $\bar{x}$ e raio ε) com $\varepsilon > 0$ com a função f_1. Teremos que $f_1 + \delta_{\overline{B(\bar{x},\varepsilon)}}$ será semicontinua inferior e própria, aplicando o Teorema 1.40 juntamente com item *ii*) da Proposição 1.44, para obter a inlusão. □

No Teorema 1.45 acima, é verificado com uma igualdade, quando consideramos uma das $f_i, \quad i \in \{1,2\}$ diferenciável.

Teorema 1.46 *Se $f_i : \mathbb{R}^n \longrightarrow \mathbb{R}$, $i = 1,2$, Lipschitz contínuas em torno de $\bar{x}$. Se $f_i \geq 0$, $i = 1,2$, então:*

$$\partial(f_1.f_2)(\bar{x}) \subset f_2(\bar{x})\partial f_1(\bar{x}) + f_1(\bar{x})\partial f_2(\bar{x}).$$

Demonstração. Ver [17, pg.1263] □

CAPÍTULO 2

Desigualdade de Kurdyka-Lojasiewicz

Neste capítulo, apresentaremos a desigualdade de Kurdyka-Lojasiewicz extraída de [3].

Definição 2.1 *Seja $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ uma função própria, semicontínua inferior. Diz-se que f satisfaz a propriedade de Kurdyka-Lojasiewicz em $\bar{x} \in dom\,\partial f$, se existem, $\eta \in]0, +\infty]$, uma vizinhaça U de $\bar{x}$ e uma função côncava e contínua $\varphi : [0, \eta[\to \mathbb{R}_+$ tal que:*

$$\varphi(0) = 0, \quad \varphi \in C^1(0, \eta), \quad \varphi'(s) > 0, \quad s \in (0, \eta); \tag{2-1}$$

$$\varphi'(f(x) - f(\bar{x}))dist(0, \partial f(x)) \geq 1, \quad x \in U \cap [f(\bar{x}) < f < f(\bar{x}) + \eta], \tag{2-2}$$

- $dist(0, \partial f(x)) := inf\{\|v\| \mid v \in \partial f(x)\}$,
- $[\eta_1 < f < \eta_2] := \{x \in \mathbb{R}^n \mid \eta_1 < f(x) < \eta_2\}, \quad \eta_1 < \eta_2$.

Dizemos que uma função é KL, se satisfaz a desigualdade de Kurdyka-Lojasiewicz em cada ponto do $dom(\partial f)$.

Observação 2.2 *Lojasiewicz [15] provou em 1963, que funções reais analíticas satisfazem uma desigualdade do tipo acima, isto é, ele mostrou que se $f : U \to \mathbb{R}$ é uma função analítica, U um subconjunto aberto de $\mathbb{R}^n$ e $\bar{x} \in U$ um ponto crítico de f, então existem $\theta \in \left[\frac{1}{2}, 1\right)$, $C > 0$ e uma vizinhança W de $\bar{x}$ tal que:*

$$\forall\, x \in W, \;\; |f(x) - f(\bar{x})|^\theta \leq C\|\nabla f(x)\|. \tag{2-3}$$

Note que a desigualdade (2-3) é um caso particular de (2-2) com $\varphi(s) = \frac{C}{1-\theta}s^{1-\theta}$. De fato, para todo $x \in U \cap [f(\bar{x}) < f < f(\bar{x}) + \eta]$, temos

$$\varphi'(f(x) - f(\bar{x}))dist(0, \partial f(x)) = \frac{C\|\nabla f(x)\|}{(f(x) - f(\bar{x}))^\theta} \geq 1.$$

Observação 2.3 *Note que a análise de convergência do método do ponto proximal proposto em [1], é feita para funções que satisfazem a desigualdade (2-3) acima. Neste contexto, o método proposto em [18] é uma generalização de [1].*

Uma função própria semicontínua inferior verifica a desigualdade KL em qualquer ponto que não é crítico, como mostra o seguinte lema.

Lema 2.4 *Sejam $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ uma função própria e semicontínua inferior e $\bar{x} \in$ dom(∂f) tal que $0 \notin \partial f(\bar{x})$. Então, a desigualdade de Kurdyka-Lojasiewicz é verificada em $\bar{x}$.*

Demonstração. Visto que $\bar{x} \notin \text{crit}(f)$ e $\partial f(\bar{x})$ é um conjunto fechado, tem-se:

$$\delta := \text{dist}(0, \partial f(\bar{x})) > 0.$$

Considere $\varphi(t) = \dfrac{t}{\delta}$, $U := B(\bar{x}, \frac{\delta}{2})$, $\eta := \frac{\delta}{2}$ e note que para cada $x \in \text{dom}(f)$,

$$\varphi'(f(x) - f(\bar{x}))\text{dist}(0, \partial f(x)) = \frac{\text{dist}(0, \partial f(x))}{\delta}. \tag{2-4}$$

Agora, para cada $x \in U \cap [f(\bar{x}) - \eta < f < f(\bar{x}) + \eta]$ arbitrário, note que:

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

Afirmamos que, para cada x verificando a última desigualdade, tem-se:

$$\text{dist}(0, \partial f(x)) \geq \delta. \tag{2-5}$$

Deixe-nos supor, por contradição que isto não ocorre. Então, existem as sequências $\{(x^k, v^k)\} \subset \text{Graf}(\partial f)$ e $\{\delta_k\} \subset \mathbb{R}_{++}$ tais que:

$$\|x^k - \bar{x}\| + |f(x^k) - f(\bar{x})| < \delta_k, \qquad \|v^k\| \leq \delta_k,$$

com $\{\delta_k\}$ convergindo a zero. Assim, como

$$\lim_{k\to+\infty} (x^k, v^k) = (\bar{x}, 0), \qquad \lim_{k\to+\infty} f(x^k) = f(\bar{x}),$$

e ∂f é uma aplicação fechada, segue-se que $\bar{x} \in \text{crit}(f)$, o que é um absurdo. Portanto, o resultado do lema segue-se por combinado (2-4) com (2-5) . $\square$

As funções convexas não satisfazem, necessariamente, a desigualdade KL; ver [8, pg.30]. A seguir apresentaremos uma classe de funções convexas que satisfazem a referida desigualdade.

Exemplo 2.5 *Considere uma função fortemente convexa f, isto é, satisfaz*

$$f(y) \geq f(x) + \langle v, y - x \rangle + \lambda \|y - x\|^p, \quad p \geq 1, \lambda > 0 \tag{2-6}$$

para todo $x, y \in \mathbb{R}^n$, $v \in \partial f(x)$. *Então,* f *satisfaz a desigualdade KL no* $dom(f)$ *para*

$$\varphi(s) = p\lambda^{\frac{-1}{p}} s^{\frac{1}{p}}.$$

Com efeito, como f *é coerciva e estritamente convexa,* $\operatorname{argmin}(f)$ *é diferente do vazio e unicamente determinado. Seja* $\{\bar{x}\} = \operatorname{argmin}(f)$. *Como* f *satisfaz* (2-6) *para todo* $x, y \in \mathbb{R}^n$, *em particular, isso ocorre para* $\bar{x}$ *e* $x \in U \cap [f(\bar{x}) < f < +\infty]$, *onde* U *é uma vizinhança de* $\bar{x}$ *e* $\eta = +\infty$. *Logo, apartir de* (2-6) *com* $y = \bar{x}$ *segue-se que*

$$f(\bar{x}) \geq f(x) + \langle v, \bar{x} - x \rangle + \lambda \|\bar{x} - x\|^p$$

e, consequentemente,

$$f(\bar{x}) - f(x) \geq \langle v, \bar{x} - x \rangle + \lambda \|\bar{x} - x\|^p \geq \langle v, \bar{x} - x \rangle .$$

Assim,

$$f(x) - f(\bar{x}) \leq \langle v, x - \bar{x} \rangle \leq \|v\| \|\bar{x} - x\| \tag{2-7}$$

e, fazendo o mesmo processo tem-se:

$$f(x) - f(\bar{x}) \geq -\|v\| \|x - \bar{x}\|. \tag{2-8}$$

Pelas desigualdades (2-7) *e* (2-8), *temos:*

$$|f(x) - f(\bar{x})| \leq \|v\| \|x - \bar{x}\|. \tag{2-9}$$

Verifiquemos a seguinte afirmação:

$$\langle v, x - \bar{x} \rangle \geq 0, \quad x \in U \cap [f(\bar{x}) < f < +\infty] . \tag{2-10}$$

De (2-6) *temos que*

$$f(\bar{x}) \geq f(x) + \langle v, \bar{x} - x \rangle + \lambda \|\bar{x} - x\|^p.$$

Consequentemente, $f(\bar{x}) \geq f(x) + \langle v, \bar{x} - x \rangle$, *pois* $\lambda \|\bar{x} - x\|^p \geq 0$. *Logo,*

$$0 \leq f(x) - f(\bar{x}) \leq \langle v, x - \bar{x} \rangle .$$

Da afirmação (2-10), *juntamente com* (2-6) *obtemos:*

$$f(x) - f(\bar{x}) \geq \lambda \|\bar{x} - x\|^p.$$

Daí,

$$\lambda^{\frac{-1}{p}}(f(x)-f(\bar{x}))^{\frac{1}{p}} \geq \|x-\bar{x}\|. \tag{2-11}$$

De (2-11) *e* (2-9) *temos:*

$$|f(x)-f(\bar{x})| \leq \lambda^{\frac{-1}{p}}\|v\|\,|f(x)-f(\bar{x})|^{\frac{1}{p}}.$$

Portanto,

$$|f(x)-f(\bar{x})|^{1-\frac{1}{p}} \leq \lambda^{\frac{-1}{p}}\|v\|.$$

Assim segue o resultado para φ *explicitada acima.*

Definição 2.6 *Dizemos que uma função* $f:\mathbb{R}^n \to \mathbb{R}$ *de classe* C^2 *é uma função de Morse se cada ponto crítico* $\bar{x}$ *de* f *é não degenerado, isto é, se o hessiano* $\nabla^2 f(\bar{x})$ *de* f *em* $\bar{x}$ *tem todos os seus autovalores diferentes de zero.*

O próximo lema será utilizado para mostrar que as funções de Morse satisfazem a desigualdade de Kurdyka-Lojasiewicz.

Lema 2.7 *Seja* $T:\mathbb{R}^n \to \mathbb{R}^n$ *uma transformação linear invertível. Então existe* $M>0$ *tal que,*

$$M\|x\| \leq \|Tx\|, \quad x \in \mathbb{R}^n.$$

Demonstração. Considere, $M = \frac{1}{\|T^{-1}\|}$. Para todo $x \in \mathbb{R}^n$, temos

$$\|x\| = \|T^{-1}(Tx)\| \leq \|T^{-1}\|\|Tx\| = \frac{1}{M}\|Tx\| \Rightarrow M\|x\| \leq \|Tx\|$$

□

Exemplo 2.8 *Sejam* $f:\mathbb{R}^n \to \mathbb{R}$ *uma função de Morse,* $\bar{x}$ *um ponto crítico de* f *e* $U = B(\bar{x},\delta)$*, tal que não exista outro ponto crítico em* U *(Isto é sempre possível, como pode ser visto no resultado em [14, pg.156]). Usando a fórmula de Taylor em* f *e em* ∇f*, obtemos:*

$$f(x)-f(\bar{x}) = \langle \nabla^2 f(\bar{x})(x-\bar{x}), x-\bar{x}\rangle + R(x)\|x-\bar{x}\|^2, \quad \textit{onde} \quad \lim_{x\to\bar{x}} R(x) = 0,$$

e

$$\nabla f(x) = \nabla^2 f(\bar{x})(x-\bar{x}) + R_1(x)\|x-\bar{x}\|, \quad \textit{onde} \quad \lim_{x\to\bar{x}} R_1(x) = 0.$$

Como f *é uma função de Morse, o hessiano* $\nabla^2 f(\bar{x})$ *de* f *em* $\bar{x}$ *tem todos os seus autovalores diferentes de zero, garantido que* $\nabla^2 f(\bar{x})$ *seja invertível, dai, pelo Lema 2.7,* *existe* $M > 0$ *tal que* $\|\nabla^2 f(\bar{x})(x-\bar{x})\| \geq M\|x-\bar{x}\|$. *Além disso, existem* $\varepsilon_1 > 0$ *e* $\delta_1 > 0$ *tal que para* $\|x-\bar{x}\| < \delta_1, \|R(x)\| \leq \varepsilon_1$, *e existe* $\delta_2 > 0$ *tal que para* $\|x-\bar{x}\| < \delta_2$, $\|R_1(x)\| < \frac{M}{2}$. *Tome* $\delta_3 = min\{\delta, \delta_1, \delta_2\}$. *Daí,*

$$\begin{aligned} |f(x) - f(\bar{x})| &\leq \|\nabla^2 f(\bar{x})\|\|x-\bar{x}\|^2 + \|R(x)\|\|x-\bar{x}\|^2, \\ &\leq (\|\nabla^2 f(\bar{x})\| + \|R(x)\|)\|x-\bar{x}\|^2, \\ &\leq (\|\nabla^2 f(\bar{x})\| + \varepsilon_1)\|x-\bar{x}\|^2, \end{aligned}$$

e portanto,

$$|f(x) - f(\bar{x})| \leq M_1\|x-\bar{x}\|^2, \quad onde \quad M_1 = \|\nabla^2 f(\bar{x})\| + \varepsilon_1.$$

Tem-se ainda que,

$$\begin{aligned} \|\nabla f(x)\| &\geq \|\nabla^2 f(\bar{x})(x-\bar{x})\| - \|R_1(x)\|\|x-\bar{x}\|, \\ &\geq M\|x-\bar{x}\| - \frac{M}{2}\|x-\bar{x}\|, \\ &= M_2\|x-\bar{x}\|, \quad onde \quad M_2 = \frac{M}{2}. \end{aligned}$$

Na definição 2.1, tome $U = B(\bar{x}, \delta_3)$, $\eta = \delta_3$ *e* $\varphi(s) = 2\frac{\sqrt{M_1 s}}{M_2}$. *Logo, para todo* $x \in U \cap [0 < f(x) - f(\bar{x}) < \eta]$,

$$\varphi'(f(x) - f(\bar{x}))dist(0, \partial f(x)) = \frac{\sqrt{M_1}}{M_2\sqrt{(f(x) - f(\bar{x}))}}\|\nabla f(x)\| \geq \frac{\|\nabla f(x)\|}{M_2\|x-\bar{x}\|} \geq 1.$$

Exemplo 2.9 *As funções reais analíticas verifica a desigualdade de Lojasiewicz. De fato, seja* f *uma função real analítica na vizinhança de um ponto* $\bar{x} \in \mathbb{R}$ *tal que* $f'(\bar{x}) = 0$. *Como* f *é analítica em uma vizinhança de* $\bar{x} \in \mathbb{R}$, *podemos escrever,*

$$f(x) = f(\bar{x}) + \sum_{k=0}^{\infty} p_k (x-\bar{x})^k, \tag{2-12}$$

onde p_k *para todo* $k \in \mathbb{N}$, *são coeficientes reais. Derivando e usando o fato de* $f'(\bar{x}) = 0$, *temos:*

$$f'(x) = \sum_{k=1}^{\infty} kp_k (x-\bar{x})^{k-1}. \tag{2-13}$$

Afirmamos que existem constantes $k_0 \in \mathbb{N}$, $\delta > 0$, c_1, c_2 *tal que, se* $|x-\bar{x}| < \delta$, *então*

$$|f(x) - f(\bar{x})| \leq c_1|x-\bar{x}|^{k_0}, \tag{2-14}$$

e

$$|f'(x)| \geq c_2|x-\bar{x}|^{k_0-1}. \tag{2-15}$$

Primeiro mostremos (2-14)*. Defina* $A = \{k \in \mathbb{N} : p_k \neq 0\}$*. Seja* k_0 *o menor elemento de A. Segue de* (2-12) *que,*

$$f(x) = f(\bar{x}) + p_{k_0}(x-\bar{x})^{k_0} + p_{k_0+1}(x-\bar{x})^{k_0+1} + \cdots.$$

Que pode ser escrito como:

$$\frac{f(x)-f(\bar{x})}{(x-\bar{x})^{k_0}} = p_{k_0} + s(x),$$

onde, $s(x) = p_{k_0+1}(x-\bar{x}) + p_{k_0+2}(x-\bar{x})^2 + \cdots$*. Como* $\lim_{x\to 0} s(x) = 0$*, existe* $\delta_1 > 0$ *tal que*

$$|x-\bar{x}| < \delta_1 \Longrightarrow |s(x)| \leq 1.$$

logo, para $|x\bar{x}| < \delta_1$*,*

$$\frac{|f(x)-f(\bar{x})|}{|x-\bar{x}|^{k_0}} \leq |p_{k_0}| + 1 = c_1,$$

ou seja,

$$|f(x)-f(\bar{x})| \leq c_1|x-\bar{x}|^{k_0}.$$

Agora provemos (2-15)*.Usando* (2-13) *obtemos,*

$$f'(x) = g(x) + r(x),$$

onde, $g(x) = k_0 p_{k_0}(x-\bar{x})^{k_0-1}$ *e* $r(x) = (k_0+1)p_{k_0+1}(x-\bar{x})^{k_0} + \cdots$*. Como* $\lim_{x\to 0} \dfrac{r(x)}{(x-\bar{x})^{k_0-1}} = 0$*, existe* $\delta_2 > 0$ *tal que para* $|x-\bar{x}| < \delta_2$*,* $|r(x)| \leq \dfrac{k_0|p_{k_0}|}{2}|x-\bar{x}|^{k_0-1}$*. Logo, para* $|x-\bar{x}| < \delta_2$*,*

$$\begin{aligned} |f'(x)| \geq |g(x)| - |r(x)| &\geq k_0|p_{k_0}||x-\bar{x}|^{k_0-1} - \frac{k_0|p_{k_0}|}{2}|x-\bar{x}|^{k_0-1}, \\ &= \frac{2k_0|p_{k_0}||x-\bar{x}|^{k_0-1} - k_0|p_{k_0}||x-\bar{x}|^{k_0-1}}{2}, \\ &= c_2|x-\bar{x}|^{k_0-1}, \quad c_2 = \frac{k_0|p_{k_0}|}{2}. \end{aligned}$$

Tomando $\delta = \{\delta_1, \delta_2\}$*, tem-se que para* $|x-\bar{x}| < \delta$*, vale* (2-14) *e* (2-15) *simultaneamente. Daí,*

$$\left(\frac{1}{c_1}\right)^{\frac{1}{k_0}} |f(x)-f(\bar{x})|^{\frac{1}{k_0}} \leq |x-\bar{x}| \leq \left(\frac{1}{c_2}\right)^{\frac{1}{k_0-1}} |f'(x)|^{\frac{1}{k_0-1}},$$

e, consequentemente,

$$c_2 \left(\frac{1}{c_1}\right)^{\frac{k_0-1}{k_0}} |f(x)-f(\bar{x})|^{\frac{k_0-1}{k_0}} \leq |f'(x)|.$$

Portanto, podemos escrever a expressão acima da seguinte forma:

$$|f(x)-f(\bar{x})|^{\frac{k_0-1}{k_0}} \leq \frac{c_1^{1-\frac{1}{k_0}}}{c_2}|f'(x)|.$$

O resultado segue tomando, $c = \dfrac{c_1^{1-\frac{1}{k_0}}}{c_2}$ *e* $\theta = 1 - \dfrac{1}{k_0}$.

Introduziremos alguns resultados e propriedades dos conjuntos semi-algébricos, os quais darão base para apresentarmos as funções semi-algébricas. Tais funções verificam a desigualdade de Kurdyka-Lojasiewicz, como veremos a seguir. Mais detalhes dessa teoria podem ser encontrados em [5, 4, 9, 2].

Definição 2.10 *Um subconjunto S de $\mathbb{R}^n$ é dito semi-algébrico básico se existem $f, g_1, \ldots, g_k$, funções polinomiais definidas em $\mathbb{R}^n$, tais que:*

$$S = \{x \in \mathbb{R}^n \mid f(x) = 0\} \cap \left(\bigcap_{i=1}^{k} \{x \in \mathbb{R}^n \mid g_i(x) > 0\}\right).$$

Dizemos que $f, g_1, \ldots, g_k$ definem o conjunto S. Se f é a função polinomial identicamentenula, apenas $g_1, \ldots, g_k$ definem S.

Definição 2.11 *Um conjunto semi-algébrico é a reunião finita de conjuntos semi-algébricos básicos. Ou seja, se $S \subset \mathbb{R}^n$ é semi-algébrico existem f_i, g_{i_j} funções polinomiais definidas em $\mathbb{R}^n$, tais que:*

$$S = \bigcup_{i=1}^{k} \left(\{x \in \mathbb{R}^n \mid f_i(x) = 0\} \cap \left(\bigcap_{j=1}^{s_i} \{x \in \mathbb{R}^n \mid g_{i_j}(x) > 0\} \right) \right).$$

Definição 2.12 *Uma função $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ é dita semi-algébrica se seu gráfico, isto é, Graf$(f) = \left\{(x,\lambda) \in \mathbb{R}^{n+1} \mid f(x) = \lambda\right\}$ é um subconjunto semi-algébrico de $\mathbb{R}^{n+1}$.*

Observação 2.13 *Toda aplicação polinomial $p : \mathbb{R}^m \to \mathbb{R}^n$ é uma aplicação semi-algébrica. De fato,*

$$Graf(p) = \{(x,y) \in \mathbb{R}^m \times \mathbb{R}^n \mid y = p(x)\} = q^{-1}(0),$$

onde, $q(x,y) = y - p(x)$ é uma função polinomial.

Teorema 2.14 *(Tarski-Seidenberg). Considere a aplicação projeção* $\pi : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m$ *definida por* $\pi(x,y) = x$*. Então, para todo subconjunto semi-algébrico* S *de* $\mathbb{R}^{m+n}$ *tem-se que* $\pi(S)$ *é um subconjunto semi-algébrico de* $\mathbb{R}^m$

Demonstração. Ver [5, pg.26] □

Vejamos algumas propriedades de conjuntos semi-algébricos:

a) Seja $S \subset \mathbb{R}^n$ um conjunto semi-algébrico. Então, os conjuntos $\overline{S}$, $\mathrm{int}(S)$ e $\mathrm{fr}(S)$, que são respectivamente fecho, interior e fronteira do conjunto S, são conjuntos semi-algébricos;
b) Sejam $A,B,C \subset \mathbb{R}^n$ conjuntos semi-algébricos. Então, $A \cup B$, $A \cap B$, $A - B$, $B - A$, A^c e B^c são conjuntos semi-algébricos;
c) Sejam $A \subset \mathbb{R}^n$ e $B \subset \mathbb{R}^m$ conjuntos semi-algébricos. Então, $A \times B \subset \mathbb{R}^n \times \mathbb{R}^m$ é um conjunto semi-algébrico.

Proposição 2.15 *Sejam* S *um subconjunto semi-algébrico não-vazio de* $\mathbb{R}^m$*,* $g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ *uma função polinomial real e* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *definida por*

$$f(x) = sup\{g(x,y) \mid y \in S\}.$$

Então, f *é uma função semi-algébrica.*

Demonstração. Para isso, basta mostrarmos que

$$\mathrm{Graf}(f) = \{(x,\lambda) \in \mathbb{R}^n \times \mathbb{R} \mid \forall y \in S, g(x,y) \leq \lambda\} \cap \{(x,\mu) \in \mathbb{R}^n \times \mathbb{R} \mid f(x) \geq \mu\}$$

é um conjunto semi-algébrico. De fato, usando as propriedades dos conjuntos semi-algébricos, vemos que

$$\begin{aligned} \Omega &= \{(x,\lambda,y) \in \mathbb{R}^n \times \mathbb{R} \times S \mid g(x,y) > \lambda\}, \\ &= \{(x,\lambda,y) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^m \mid g(x,y) > \lambda\} \cap (\mathbb{R}^n \times \mathbb{R} \times S), \end{aligned}$$

é um conjunto semi-algébrico. Defina,$\pi : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^n \times \mathbb{R}$ pondo,

$$\pi(x,\lambda,y) = (x,\lambda).$$

Pelo Teorema 2.14 o conjunto, $\pi(\Omega) = \{(x,\lambda) \in \mathbb{R}^n \times \mathbb{R} \mid \exists y \in S, g(x,y) > \lambda\}$ é semi-algébrico, logo o seu complementar,

$$\pi(\Omega)^c = \{(x,\lambda) \in \mathbb{R}^n \times \mathbb{R} \mid \forall y \in S, g(x,y) \leq \lambda\},$$

goza da mesma propriedade. Analogamente, mostra-se que o conjunto

$$\{(x,\mu) \in \mathbb{R}^n \times \mathbb{R} \mid \forall y \in S, g(x,y) \geq \mu\} = \{(x,\mu) \in \mathbb{R}^n \times \mathbb{R} \mid f(x) \geq \mu\},$$

é semi-algébrico. Portanto, Graf(f) é um conjunto semi-algébrico. $\square$

O seguinte resultado apresenta uma classe de funções que verificam a desigualdade de Kurdyka-Lojasiewicz, pode ser encontrado em [6, 7]

Teorema 2.16 *Seja $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, uma função própria semicontínua inferior. Se f é semi-algébricas, então ela satisfaz a propriedade KL em qualquer ponto do dom(f).*

CAPÍTULO 3

Algoritmo Proximal Generalizado

Neste capítulo, apresentaremos o método do ponto proximal generalizado MPPG. Nesse contexto, a função regularizadora é uma quase-distância e a função objetivo verifica a desigualdade (2-2). Esse método, será usado para resolver o Problema (0-1). Descreveremos as hipóteses assumidas sobre a quase-distância e a função objetivo, além disso, provaremos a convergência parcial e total de uma sequência limitada gerada pelo MPPG, a um ponto crítico generalizado.

3.1 Um Algoritmo Proximal com Quase-distância

O Algorítmo gera uma sequência $\{x^k\}_{k\in\mathbb{N}}$ via a seguinte recursão.
1.Tome $x^0 \in \mathrm{dom}(f)$.
2. Se x^k é crítico **PARE**,
3.Dado x^k, encontrar x^{k+1} tal que:

$$x^{k+1} \in \operatorname{argmin}\{f(u) + \frac{1}{2\lambda_k} q^2(x^k, u) \mid u \in \mathbb{R}^n\}, \tag{3-1}$$

onde $\{\lambda_k\}_{k\in\mathbb{N}}$ é uma sequência positiva tal que $0 < \lambda_- \leq \lambda_k \leq \lambda_+$.

As seguintes hipóteses são assumidas para o estudo:
$(\mathcal{H}_1)$ $-\infty < \inf_{x\in\mathbb{R}^n} f(x)$;
$(\mathcal{H}_2)$ q satisfaz a condição:(1-4);
$(\mathcal{H}_3)$ A restrição de f em seu domínio é uma função contínua.

3.2 Análise de Convergência

3.2.1 Convergência Parcial

Proposição 3.1 *Se $(\mathcal{H}_1)$ e $(\mathcal{H}_2)$ são verificadas, então:*
i) $\{x^k\}$ está bem definida,

Em adição, se $\{x^k\}_{k\in\mathbb{N}}$ *cumpre* (3-1) *segue:*
$ii)$ *A sequência* $\{f(x^k)\}$ *é decrescente e limitada, logo convergente.*
$iii)$ $\sum_{k=0}^{+\infty} q^2(x^k,x^{k+1}) < +\infty$*, em particular,*

$$\lim_{k\to+\infty} q(x^k,x^{k+1}) = 0.$$

Demonstração. O item (i) segue do Teorema 1.7. Do processo interativo (3-1),

$$f_k(x^{k+1}) \le f_k(x^k), \quad k\in\mathbb{N},$$

onde, $f_k = f(\cdot) + \frac{1}{2\lambda_k}q^2(x^k,\cdot)$. dessa desigualdade segue que:

$$f(x^{k+1}) + \frac{1}{2\lambda_k}q^2(x^k,x^{k+1}) \le f(x^k) + \frac{1}{2\lambda_k}q^2(x^k,x^k), \tag{3-2}$$

ou simplesmente,

$$f(x^{k+1}) \le f(x^k), \quad k\in\mathbb{N}.$$

Então, $\{f(x^k)\}$ é não-crescente. Como f é limitada inferiormente. Logo, $\{f(x^k)\}$ é convergente. Isso prova (ii). Agora note que da desigualdade (3-2) , obtem-se:

$$q^2(x^k,x^{k+1}) \le 2\lambda_k[f(x^k) - f(x^{k+1})]. \tag{3-3}$$

Fazendo o somatório em ambos os lados de (3-3), com k variando de 0 até N dado, obtemos:

$$\sum_{k=0}^{N} q^2(x^k,x^{k+1}) \le \sum_{k=0}^{N} 2\lambda_k[f(x^k) - f(x^{k+1})]. \tag{3-4}$$

Combinando (3-4) com a limitação superior da sequência λ_k,

$$\sum_{k=0}^{N} q^2(x^k,x^{k+1}) \le 2\lambda_+[f(x^0) - f(x^N)]. \tag{3-5}$$

Por (3-5) e pela limitação inferior de f chegamos a:

$$\sum_{k=0}^{N} q^2(x^k,x^{k+1}) \le 2\lambda_+[f(x^0) - \inf f(x)_{x\in\mathbb{R}^n}] < +\infty, \tag{3-6}$$

tomando o limite em (3-6) com $N\to+\infty$, prova-se (iii). Em particular, da convergência de $\sum_{k=0}^{+\infty} q^2(x^k,x^{k+1})$, temos:

$$\lim_{k\to+\infty} q(x^k,x^{k+1}) = 0.$$

$\square$

O processo iterativo (3-1) pode ser caracterizado em termos do subdiderencial-limite.

Proposição 3.2 *Se* $(\mathcal{H}_1)$ *e* $(\mathcal{H}_2)$ *são verificadas, então existem* $w^{k+1} \in \partial f(x^{k+1})$ *e* $v^{k+1} \in \partial(q(x^k,\cdot))(x^{k+1})$ *tal que:*

$$0 = w^{k+1} + \frac{q(x^k, x^{k+1})}{\lambda_k} v^{k+1}. \tag{3-7}$$

Demonstração. Como a interada x^{k+1} satisfaz a condição de otimalidade de primeira ordem da função $f(\cdot) + \frac{1}{2\lambda_k} q^2(x^k,\cdot)$. Pela Proposição 1.28, teremos:

$$0 \in \partial\left(f + \frac{1}{2\lambda_k} q^2(x^k,\cdot)\right)(x^{k+1}).$$

Por outro lado, desde que (1-4) é verificada, a Proposição 1.17 garante que para $\bar{z} = x^k$, a função $\frac{1}{2\lambda_k} q^2(x^k,\cdot)$ é localmente lipschitziana em x^{k+1}, que acrecido à semicontinuidade inferior de f à Proposição 1.45 resulta em:

$$0 \in \partial f(x^{k+1}) + \frac{1}{2\lambda_k}\partial\left(q^2(x^k,\cdot)\right)(x^{k+1}), \quad k \in \mathbb{N}. \tag{3-8}$$

Além disso, considerando $f_1 = f_2 = q(x^k,\cdot) \geq 0$ e $\bar{x} = x^{k+1}$ na Proposição 1.46, resulta que:

$$\partial\left(q^2(x^k,\cdot)\right)(x^{k+1}) \subset 2q(x^k, x^{k+1})\partial\left(q(x^k,\cdot)\right)(x^{k+1}), \ \ k \in \mathbb{N}. \tag{3-9}$$

Assim, combinando (3-8) com (3-9), obtêm-se que:

$$0 \in \partial f(x^{k+1}) + \frac{1}{\lambda_k} q(x^k, x^{k+1})\partial\left(q(x^k,\cdot)\right)(x^{k+1}).$$

Portanto, existem $w^{k+1} \in \partial f(x^{k+1})$ e $v^{k+1} \in \partial(q(x^k,\cdot))(x^{k+1})$ tais que (3-7) é verificada. $\square$

Note que, no caso onde $q = \|\cdot\|$, (3-7) reduz-se à condição de otimalidade clássica, dada por:

$$\frac{(x^k - x^{k+1})}{\lambda_k} \in \partial f(x^{k+1}).$$

Lema 3.3 *Sejam* $\bar{z}, \bar{x} \in \mathbb{R}^n$ *e* q *uma quase-distância que verifica* $(\mathcal{H}_2)$*, então existe uma constante* $M > 0$ *que independe de* $\bar{z}, \bar{x}$ *tal que:*

$$\|v\| \leq M, \quad v \in \partial(q(\bar{z},\cdot))(\bar{x}). \tag{3-10}$$

Demonstração. De $(\mathcal{H}_2)$ e Proposição 1.16, tem-se que $q(\bar{z},\cdot)$ é uma função Lipschitziana em $\mathbb{R}^n$, com constante de Lipschitz $2\beta > 0$. Em particular, $q(\bar{z},\cdot)$ é uma função localmente lipschitiziana em $\bar{x} \in \mathbb{R}^n$. Note que esta constante independe de $\bar{z}$. Assim, a Proposição 1.44 implica que:

$$\|v\| \leq 2\beta, \quad v \in \partial(q(\bar{z},\cdot))(\bar{x}),$$

e, visto que 2β também independe de $\bar{x}$, o resultado desejado segue com $M = 2\beta > 0$. □

Proposição 3.4 *Seja $\{x^k\}_{k\in\mathbb{N}}$ uma sequência limitada gerada pelo processo em 3-1 e denotemos por $\omega(x^0)$ o conjunto dos seus pontos de acumulação. Suponha que f satisfaz $(\mathcal{H}_3)$. Então, qualquer ponto de acumulação da sequência $\{x^k\}_{k\in\mathbb{N}}$ é um ponto crítico limite de f.*

Demonstração. Tome $\bar{x} \in \omega(x^0)$. Dessa maneira, existe uma subsequência $\{x^{k_j}\}$ de $\{x^k\}$ tal que, $x^{k_j} \to \bar{x}$, devemos mostrar que $0 \in \partial f(\bar{x})$. De fato, da Proposição 3.2, tem-se que existem $w^{k_j+1} \in \partial f(x^{k_j+1})$ e $v^{k_j+1} \in \partial(q(x^{k_j},.))(x^{k_j+1})$ tais que

$$0 = w^{k_j+1} + \frac{q(x^{k_j},x^{k_j+1})}{\lambda_{k_j}} v^{k_j+1}.$$

Assim, tem-se que:

$$\|w^{k_j+1}\| \leq \frac{q(x^{k_j},x^{k_j+1})}{\lambda_{k_j}} \|v^{k_j+1}\| \leq \frac{L}{\lambda_-} q(x^{k_j},x^{k_j+1}),$$

onde a última desigualdade é uma consequência do Lema 3.3, logo, consideramos o limite quando $j \longrightarrow +\infty$ e pela Proposição 3.1,

$$\lim_{j\longrightarrow+\infty} \|w^{k_j}\| = 0.$$

Assim, $(\bar{x},0)$ é ponto de acumulação da sequência $\{(x^{k_j},w^{k_j})\}$ e, como o graf∂f é fechado, $\{(x^{k_j},w^{k_j})\} \subset \text{graf}\partial f(x^{k_j})$. Logo, $(\bar{x},0) \in \text{graf}\partial f(\bar{x})$, isto é, $0 \in \partial f(\bar{x})$ e assim conclui-se que $\bar{x} \in \text{crit} f$;

□

3.2.2 Convergência Total

Em seguida será feita a análise de convergência, mostrando que uma sequência limitada gerada pelo MPPG, converge a algum ponto de acumulação os quais são pontos críticos generalizados da função objetivo f.

Teorema 3.5 *Assumindo que a sequência $\{x^k\}_{k\in\mathbb{N}}$ gerada pelo (MPPG), é limitada, $\overline{x}$ é um ponto de acumulação de $\{x^k\}_{k\in\mathbb{N}}$ e vale $(\mathcal{H}_1)$, $(\mathcal{H}_2)$ e $(\mathcal{H}_3)$. Seja $U\subset\mathbb{R}^n$ uma vizinhança de $\bar{x}$, $\eta\in]0,+\infty]$ e $\varphi:[0,\eta[\to\mathbb{R}_+$ uma função contínua e côncava como em (2-1) e (2-2). Se $\delta>0$ e $r\in]0,1[$ são constantes fixas, tais que $B\left(\bar{x},\frac{\delta}{\alpha}\right)\subset U$, então existe $k_0\in\mathbb{N}$ tal que:*

$$f(\bar{x})<f(x^k)<f(\bar{x})+\eta,\quad k\geq k_0, \tag{3-11}$$

$$\begin{aligned} q(\bar{x},x^{k_0})+\frac{1}{1-r}\sqrt{2\lambda_+\left[f(x^{k_0})-f(\bar{x})\right]}+\frac{r}{1-r}\sqrt{2\lambda_+\left[f(x^{k_0-1})-f(\bar{x})\right]}+\\ +\frac{\lambda_-}{r(1-r)2L\lambda_+}\left[\varphi(f(x^{k_0}))-\varphi(f(\bar{x}))\right]<\delta, \end{aligned} \tag{3-12}$$

$$x^{k_0+j}\in B_q(\bar{x},\delta),\quad j=1,2,\dots. \tag{3-13}$$

Além disso,

$$\sum_{k=0}^{+\infty}q(x^k,x^{k+1})<+\infty. \tag{3-14}$$

Em particular, a sequência $\{x^k\}$ converge para $\bar{x}$ o qual é um ponto crítico de f.

Demonstração. Pode-se considerar que x^k, $\bar{x}\in\mathrm{dom}(f)$, para todo $k\in\mathbb{N}$, pois caso $x^0\notin\mathrm{dom}(f)$, pela definição da sequência $\{x^k\}$ o próximo ponto x^1 estará em $\mathrm{dom}(f)$. Pelo item *ii*) da Proposição 3.1 $\{f(x^k)\}$ é bem definida e converge para o $\inf\limits_{k>0}f(x^k)=f(\bar{x})$. Trocando f por $f-\inf_{k>0}f(x^k)$, podemos assumir $\lim\limits_{k\to+\infty}f(x^k)=0$. Logo, temos sem perda de generalidade que, $f(\bar{x})=0$. Visto que a sequência $\{f(x^k)\}$ é positiva e decresce estritamente para 0, obtemos:

$$0<f(x^k),\quad k\in\mathbb{N}. \tag{3-15}$$

Em particular, existe um índice $N\in\mathbb{N}$ tal que:

$$0<f(x^k)<\eta,\quad k\geq N. \tag{3-16}$$

Agora defina a sequência $\{b_k\}$, dada por:

$$b_k=q(\bar{x},x^k)+\frac{1}{1-r}\sqrt{2\lambda_+f(x^k)}+\frac{r}{1-r}\sqrt{2\lambda_+f(x^{k-1})}+\frac{\lambda-}{r(1-r)2L\lambda+}\varphi(f(x^k)).$$

Como $q(\bar{x},.)$, $\sqrt{.}$ e φ são funções contínuas, $\varphi(0)=0$ e $\{f(x^k)\}$ converge para 0, segue que 0 é um ponto de acumulação da sequência $\{b_k\}$. Então existe $k_0 = k_{j_0} > N$ tal que:

$$q(\bar{x},x^{k_0}) + \frac{1}{1-r}\sqrt{2\lambda_+ f(x^{k_0}))} + \frac{r}{1-r}\sqrt{2\lambda_+ f(x^{k_0-1})} + \frac{\lambda_-}{r(1-r)2L\lambda_+}\left[\varphi(f(x^{k_0}))\right] < \delta,$$

o que prova (3-12). Em particular, como $k_0 > N$, vale (3-11). Note que (3-12) implica que $x^{k_0} \in B_q(\bar{x},\delta)$, visto que $b_{k_0} < \delta$ e que os termos que determinam b_{k_0} são positivos. Pela hipótese $(\mathcal{H}_2)$, temos:

$$\alpha\|\bar{x}-x^{k_0}\| < q(\bar{x},x^{k_0}) \Rightarrow x^{k_0} \in B\left(\bar{x},\frac{\delta}{\alpha}\right) \subset U.$$

Assim, $x^{k_0} \in U \cap [0 < f < \eta]$. Como $\bar{x}$ é um ponto onde f satisfaz a desigualdade de Kurdyka-Lojasiewicz, temos:

$$\varphi'(f(x^{k_0}))\text{dist}(0,\partial f(x^{k_0})) \geq 1.$$

Logo, $0 \notin \partial f(x^{k_0})$. Pelo Lema 3-10 e pela Proposição 3-7 combinadas com a definição de $\text{dist}(0,\partial f(x^{k_0}))$, obtêm-se:

$$\text{dist}(0,\partial f(x^{k_0})) \leq \|w^{k_0}\| \leq \frac{1}{\lambda_-}q(x^{k_0-1},x^{k_0})\|v^{k_0}\| \leq \frac{L}{\lambda_-}q(x^{k_o-1},x^{k_0}),$$

pois, $w^{k_0} \in \partial f(x^{k_0})$ e $v^{k_0} \in \partial q(x^{k_0-1},\cdot)(x^{k_0})$.
Novamente pela desigualdade de Kurdyka-Lojasiewicz em $\bar{x}$, segue que,

$$\begin{aligned} 1 &\leq \varphi'(f(x^{k_0}))\text{dist}(0,\partial f(x^{k_0})), \\ &\leq \varphi'(f(x^{k_0}))\frac{L}{\lambda_-}q(x^{k_0-1},x^{k_0}). \end{aligned}$$

Daí,

$$\varphi'(f(x^{k_0})) \geq \frac{\lambda_-}{Lq(x^{k_0-1},x^{k_0})}. \tag{3-17}$$

Por outro lado, a concavidade da função φ implica que,

$$-\varphi(z) \geq -\varphi(t) - \varphi'(t)(z-t), \quad z,t \in \mathbb{R}.$$

Tomando $z = f(x^{k_0+1})$ e $t = f(x^{k_0})$, temos:

$$\varphi(f(x^{k_0})) - \varphi(f(x^{k_0+1})) \geq \varphi'(f(x^{k_0}))(f(x^{k_0}) - f(x^{k_0+1})).$$

Combinando essa última desigualdade com,

$$f(x^{k_0}) - f(x^{k_0+1}) \geq \frac{1}{2\lambda_+} q^2(x^{k_0}, x^{k_0+1})$$

e $\varphi'(s) > 0$, $s \in (0, \eta)$, tem-se:

$$\varphi(f(x^{k_0})) - \varphi(f(x^{k_0+1})) \geq \varphi'(f(x^{k_0})) \frac{1}{2\lambda_+} q^2(x^{k_0}, x^{k_0+1}). \tag{3-18}$$

Agora, usando as desigualdades (3-17) e (3-18), chegamos à seguinte expressão:

$$\begin{aligned} \varphi(f(x^{k_0})) - \varphi(f(x^{k_0+1})) &\geq \varphi'(f(x^{k_0})) \frac{1}{2\lambda_+} q^2(x^{k_0}, x^{k_0+1}), \\ &\geq \frac{\lambda_-}{Lq(x^{k_0-1}, x^{k_0})} \frac{1}{2\lambda_+} q^2(x^{k_0}, x^{k_0+1}). \end{aligned}$$

Logo,

$$\frac{q^2(x^{k_0}, x^{k_0+1})}{q(x^{k_0-1}, x^{k_0})} \leq \frac{2L\lambda_+}{\lambda_-} \left[\varphi(f(x^{k_0})) - \varphi(f(x^{k_0+1})) \right]. \tag{3-19}$$

Provaremos a seguinte afirmação:

$$q(x^k, x^{k+1}) \leq rq(x^{k-1}, x^k) + \frac{M}{r} \left[\varphi(f(x^k)) - \varphi(f(x^{k+1})) \right], \tag{3-20}$$

para $k = k_0$, e onde $M = \dfrac{2L\lambda_+}{\lambda_-}$. De fato, temos duas possibilidades a consider:

a) $q(x^{k_0}, x^{k_0+1}) \geq rq(x^{k_0-1}, x^{k_0})$,

b) $q(x^{k_0}, x^{k_0+1}) < rq(x^{k_0-1}, x^{k_0})$.

Primeiro suponha que ocorra *a)*. Então, $q(x^{k_0-1}, x^{k_0}) \leq \dfrac{q(x^{k_0}, x^{k_0+1})}{r}$, multiplicando e dividindo o lado direito dessa última desigualdade por $q(x^{k_0}, x^{k_0+1})$ obtemos,

$$q(x^{k_0-1}, x^{k_0}) \leq \frac{q(x^{k_0}, x^{k_0+1})}{r} = \frac{q^2(x^{k_0}, x^{k_0+1})}{rq(x^{k_0}, x^{k_0+1})} \Longrightarrow q(x^{k_0}, x^{k_0+1}) \leq \frac{q^2(x^{k_0}, x^{k_0+1})}{rq(x^{k_0-1}, x^{k_0})}.$$

Consequentemente teremos:

$$q(x^{k_0}, x^{k_0+1}) \leq rq(x^{k_0-1}, x^{k_0}) + \frac{M}{r} \left[\varphi(f(x^{k_0})) - \varphi(f(x^{k_0+1})) \right].$$

Visto que $rq(x^{k_0-1},x^{k_0}) > 0$, a afirmação fica provada. Agora se ocorre $b)$, $q(x^{k_0},x^{k_0+1}) < rq(x^{k_0-1},x^{k_0})$, somando $\frac{M}{r}\left[\varphi(f(x^{k_0})) - \varphi(f(x^{k_0+1}))\right] \geq 0$. A desigualdade em $b)$ ainda continua sendo verificada. Assim, prova-se (3-20). Em seguida (3-13) será provado por indução sobre j. Já foi provado o caso em que $j = 0$ segue diretamente da prova de (3-12). Vejamos agora para $j = 1$, como $f(x^k) > 0$, $0 < \lambda_- \leq \lambda_k \leq \lambda_+, \quad \forall k \in \mathbb{N}$ e tem-se $r \in (0,1)$ podemos fazer a seguinte análise. Temos:

$$\frac{1}{2\lambda_+}q^2(x^{k_0},x^{k_0+1}) \leq f(x^{k_0}) - f(x^{k_0+1}),$$

Assim,

$$q(x^{k_0},x^{k_0+1}) \leq \sqrt{2\lambda_+[f(x^{k_0}) - f(x^{k_0+1})]},$$

tirado o termo $f(x^{k_0+1})$, pois quando o retiramos, ainda temos a desigualdade. Segue:

$$q(x^{k_0},x^{k_0+1}) \leq \sqrt{2\lambda_+ f(x^{k_0})} \leq \frac{1}{1-r}\sqrt{2\lambda_+ f(x^{k_0})}. \tag{3-21}$$

Além disso, combinando a desigualdade triangular com a desigualdade (3-21), obtemos por (3-12) que:

$$q(\bar{x},x^{k_0+1}) \leq q(\bar{x},x^{k_0}) + q(x^{k_0},x^{k_0+1}) \leq q(\bar{x},x^{k_0}) + \frac{1}{1-r}\sqrt{2\lambda_+ f(x^{k_0})} < \delta,$$

com isso mostramos (3-13) para $j = 1$. Agora tome $j \geq 1$ e suponha que (3-13) é verdade para todo $k = k_0, k_0+1, k_0+2, \ldots, k_0+j-1$, neste caso também temos (3-20) vale para $k = k_0 + j$, com $j = 0,1,\ldots,j-1$ isto é,

$$q(x^{k_0+j-1},x^{k_0+j}) \leq rq(x^{k_0+j-1-1},x^{k_0+j-1}) + \frac{M}{r}\left[\varphi(f(x^{k_0+j-1})) - \varphi(f(x^{k_0+j}))\right],$$

para $j = 0,1,2,\ldots,j-1$. Assim, somando as j inequações obtidas. Teremos a seguinte expressão:

$$\begin{aligned}
\sum_{i=0}^{j-1} q(x^{k_0+i},x^{k_0+i+1}) &\leq r\sum_{i=0}^{j-1} q(x^{k_0+i-1},x^{k_0+i}) + \frac{M}{r}\left[\varphi(f(x^{k_0})) - \varphi(f(x^{k_0+j}))\right], \\
&\leq rq(x^{k_0-1},x^{k_0}) + r\sum_{i=1}^{j-1} q(x^{k_0+i-1},x^{k_0+i}) + \left[\varphi(f(x^{k_0})) - \varphi(f(x^{k_0+j}))\right]
\end{aligned}$$

Fazendo algumas manipulações algébricas obtemos:

$$(1-r)\sum_{i=0}^{j-1} q(x^{k_0+i}, x^{k_0+i+1}) \leq rq(x^{k_0-1}, x^{k_0}) + \left[\varphi(f(x^{k_0})) - \varphi(f(x^{k_0+j}))\right].$$

Isto é,

$$\sum_{i=0}^{j-1} q(x^{k_0+i}, x^{k_0+i+1}) \leq \frac{r}{1-r} q(x^{k_0-1}, x^{k_0}) + \frac{M}{r(1-r)}\left[\varphi(f(x^{k_0})) - \varphi(f(x^{k_0+j}))\right]. \quad (3\text{-}22)$$

E pela desigualdade triangular,

$$\begin{aligned} q(\bar{x}, x^{k_0+j}) \leq q(\bar{x}, x^{k_0}) + q(x^{k_0}, x^{k_0+1}) \quad &+ \quad q(x^{k_0+1}, x^{k_0+1+1}) + \ldots + \\ \ldots \quad &+ q(x^{k_0+i}, x^{k_0+i+1}), \end{aligned}$$

com $i = 0, 1, ..., j-1$, isto é:

$$q(\bar{x}, x^{k_0+j}) \leq q(\bar{x}, x^{k_0}) + q(x^{k_0}, x^{k_0+1}) + \sum_{i=1}^{j-1} q(x^{k_0+i}, x^{k_0+i+1}). \quad (3\text{-}23)$$

Assim combinando (3-23), (3-21) e $-\varphi(f(x^{k_0+j})) < 0$, chega-se à:

$$q(\bar{x}, x^{k_0+j}) \leq q(\bar{x}, x^{k_0}) + \frac{1}{1-r}\sqrt{2\lambda_+ f(x^{k_0})} + \frac{r}{1-r}\sqrt{2\lambda_+ f(x^{k_0})} + \frac{M}{r(1-r)}\left[\varphi(f(x^{k_0}))\right].$$

Como já mostrado em 3-12, $q(\bar{x}, x^{k_0+j}) < \delta$, implicando que $x^{k_0+j} \in B_q(\bar{x}, \delta)$, e assim conclui-se a prova por indução. Note que 3-14, segue imediatamente de 3-22, ou seja,

$$\sum_{k=0}^{+\infty} q(x^k, x^{k+1}) < +\infty.$$

Esse resultado junto com a hipótese $(\mathcal{H}_2)$ implica que $\{x^k\}$ é uma sequência de Cauchy. Logo convergente, e converge a $\bar{x}$, que é por sua vez um ponto de acumulação, e como visto na Proposição 3.4, $\bar{x}$ é um ponto crítico de f. □

3.2.3 Taxa de Convergência

Apresentaremos a seguir uma análise da taxa de convergência.

Teorema 3.6 *Assumindo que f satisfaz $(\mathcal{H}_1)$ - $(\mathcal{H}_4)$ e seja $\{x^k\}_{k\in\mathbb{N}}$ uma sequência gerada pelo MPPG limitada. Sendo $\varphi(s) = s^{1-\theta}$, $\theta \in [0,1)$, e $\bar{x}$ o ponto limite da sequência*

$\{x^k\}_{k\in\mathbb{N}}$. *Então:*

i) Se $\theta = 0$*, a sequência converge em um número finito de passos;*

ii) Se $\theta \in (0, \frac{1}{2}]$*, então existe* $C > 0$ *e* $Q \in [0,1)$ *tal que:*

$$q(\bar{x}, x^k) \leq CQ^k;$$

iii) Se $\theta \in (\frac{1}{2}, 1)$*, então existe* $C > 0$ *tal que:*

$$q(\bar{x}, x^k) \leq Ck^{-\frac{1-\theta}{2\theta-1}}.$$

Demonstração. Se $\theta = 0$, suponha por contradição que a sequência $\{x^k\}$ é infinitamente gerada. Tome $k_0 \in \mathbb{N}$ suficientemente grande, tal que, $x^k \in B_q(\bar{x}, \delta) \cap [0 < f < \eta]$, $\quad k \geq k_0$, (estamos supondo que $f(\bar{x}) = 0$). Como $\theta = 0$, $\varphi(s) = s$. Por (2-2), obtemos

$$1 \leq \operatorname{dist}(0, \partial f(x^k)), \quad k \geq k_0. \tag{3-24}$$

Por outro lado, a partir da Proposição 3.2 para cada $k \in \mathbb{N}$, existe $w^k \in \partial f(x^k)$ e $v^k \in \partial q(x^{k-1}, .)(x^k)$ tal que,

$$\|w^k\| \leq \frac{\|v^k\| q(x^{k-1}, x^k)}{\lambda_{k-1}} \leq \frac{M q(x^{k-1}, x^k)}{\lambda_-}, \tag{3-25}$$

onde a última desigualdade segue a partir do Lema 3.3 (com $v = v^k$, $\bar{z} = x^{k-1}$ e $\bar{x} = x^k$) e definição da sequência $\{\lambda_k\}$. Assim, visto que $\lim\limits_{k\to+\infty} q(x^{k-1}, x^k) = 0$ (isto segue a partir do item *iii*) da Proposição 3.1), obtemos que $\lim\limits_{k\to+\infty} w^k = 0$ o que é uma contradição com (3-24). Logo, o item (i) está provado. Deixe-nos definir a seguinte notação:

$$\Delta_k = \sum_{p=k}^{\infty} q(x^p, x^{p+1}).$$

Dessa maneira, $\Delta_{k-1} = \sum\limits_{p=k-1}^{\infty} q(x^p, x^{p+1})$ e, pela desigualdade triangular, segue que $q(\bar{x}, x^k) \leq \Delta_k$. Utilizaremos a última desigualdade para concluirmos as estimativas que aparecem em *ii*) e *iii*), ou seja, limitaremos Δ_k, e consequentemente $q(\bar{x}, x^k)$ estará limitado.

Observe que, $\Delta_{k-1} - \Delta_k = q(x^{k-1}, x^k)$. Pela expressão (3-20), temos

$$\sum_{p=k}^{N} q(x^p, x^{p+1}) \leq r \sum_{p=k}^{N} q(x^{p-1}, x^p) + \frac{M}{r} \sum_{p=k}^{N} [\varphi(f(x^p)) - \varphi(f(x^{p+1}))]. \tag{3-26}$$

Como,

$$\begin{aligned}\sum_{p=k}^{N} q(x^{p-1},x^{p}) &= q(x^{k-1},x^{k})+\sum_{p=k+1}^{N} q(x^{p-1},x^{p}),\\ &= q(x^{k-1},x^{k})+\sum_{p=k}^{N} q(x^{p},x^{p+1}).\end{aligned}$$

Substituindo $\sum_{p=k}^{N} q(x^{p-1},x^{p})$ em (3-26), obtemos,

$$\begin{aligned}\sum_{p=k}^{N} q(x^{p},x^{p+1}) &\leq \frac{r}{1-r}q(x^{k-1},x^{k})+\frac{M}{r(1-r)}\sum_{p=k}^{N}[\varphi(f(x^{p}))-\varphi(f(x^{p+1}))],\\ &\leq \frac{r}{1-r}q(x^{k-1},x^{k})+\frac{M}{r(1-r)}[\varphi(f(x^{k}))-\varphi(f(x^{N+1}))],\\ &\leq \frac{r}{1-r}q(x^{k-1},x^{k})+\frac{M}{r(1-r)}\varphi(f(x^{k})).\end{aligned}$$

Assim, $\Delta_k \leq \frac{r}{1-r}(\Delta_{k-1}-\Delta_k)+\frac{M}{r(1-r)}\varphi(f(x^k))$, sendo $\varphi(s)=s^{1-\theta}$, a equação fica como segue,

$$\Delta_k \leq \frac{r}{1-r}(\Delta_{k-1}-\Delta_k)+\frac{M}{r(1-r)}(f(x^k))^{1-\theta}. \tag{3-27}$$

Como já assumimos que $x^k \in B_q(\bar{x},\delta)\cup[0<f<\eta], \quad k\geq k_0$, e que $f(\bar{x})=0$, temos pela desigualdade de Kurdyka-Lojasiewicz,

$$\begin{aligned}1 &\leq \varphi'(f(x^k))\mathrm{dist}(0,\partial f(x^k)),\\ &\leq (1-\theta)f(x^k)^{-\theta}\|w^k\|.\end{aligned}$$

Para todo $w^k\in\partial f(x^k)$, consequentemente,

$$f(x^k)^{\theta}\leq(1-\theta)\|w^k\|,$$

ou ainda,

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

Voltando a equação (3-27), e substituindo a expressão acima e usando a limitação em (3-25) temos,

$$\begin{aligned}
\Delta_k &\leq \frac{r}{1-r}(\Delta_{k-1}-\Delta_k)+\frac{M(1-\theta)^{\frac{1-\theta}{\theta}}\|w^k\|^{\frac{1-\theta}{\theta}}}{r(1-r)}, \qquad (3\text{-}28)\\
&\leq \frac{r}{1-r}(\Delta_{k-1}-\Delta_k)+\frac{M(1-\theta)^{\frac{1-\theta}{\theta}}}{r(1-r)}\left[\frac{Lq(x^{k-1},x^k)}{\lambda_-}\right]^{\frac{1-\theta}{\theta}},\\
&\leq \frac{r}{1-r}(\Delta_{k-1}-\Delta_k)+\frac{M(1-\theta)^{\frac{1-\theta}{\theta}}}{r(1-r)}\left[\frac{L(\Delta_{k-1}-\Delta_k)}{\lambda_-}\right]^{\frac{1-\theta}{\theta}}.
\end{aligned}$$

Se supormos $\theta \in (\frac{1}{2},1)$, teremos, $\frac{1-\theta}{\theta}<1$. Como $\Delta_k \to 0$, quando $k\to\infty$, (basta olhar a expressão (3-28)). Daí temos que existem um inteiro positivo $k_1 \geq k_0$ e uma constante $\overline{C_1}$ tal que:

$$\Delta_k \leq \overline{C_1}(\Delta_{k-1}-\Delta_k)^{\frac{1-\theta}{\theta}},$$

ou simplesmente,

$$\Delta_k^{\frac{\theta}{1-\theta}} \leq C_1(\Delta_{k-1}-\Delta_k). \qquad (3\text{-}29)$$

Agora defina $h:(0,+\infty)\to\mathbb{R}$, por $h(s)=s^{\frac{-\theta}{1-\theta}}$ e seja $R\in(1,+\infty)$. Afirmamos que para cada $k\geq k_1$, existem $\mu>0$ e $\nu<0$ tal que:

$$\Delta_k^{\nu}-\Delta_{k-1}^{\nu}\geq\mu. \qquad (3\text{-}30)$$

Vamos considerar dois casos:

a) $h(\Delta_k)\leq Rh(\Delta_{k-1})$,
b) $h(\Delta_k)> Rh(\Delta_{k-1})$.

Suponha primeiro que ocorra a). Pela desigualdade em (3-29), temos:

$$1\leq\frac{C_1(\Delta_{k-1}-\Delta_k)}{\Delta_k^{\frac{\theta}{1-\theta}}}=C_1(\Delta_{k-1}-\Delta_k)h(\Delta_k).$$

Usando a) temos:

$$\begin{aligned}
1 &\leq C_1(\Delta_{k-1}-\Delta_k)h(\Delta_k),\\
&\leq RC_1(\Delta_{k-1}-\Delta_k))h(\Delta_{k-1}).
\end{aligned}$$

Agora, visto que $h(\Delta_{k-1}) = \min_{\Delta_k \leq s \leq \Delta_{k-1}} h(s)$, última desigualdade implica:

$$\begin{aligned} 1 &\leq RC_1 \int_{\Delta_k}^{\Delta_{k-1}} h(s)\mathrm{d}s, \\ &= RC_1 \int_{\Delta_k}^{\Delta_{k-1}} s^{\frac{-\theta}{1-\theta}}\mathrm{d}s, \\ &= RC_1 \left[\frac{s^{\frac{-\theta}{1-\theta}+1}}{\frac{-\theta}{1-\theta}+1} \right] \Big|_{\Delta_k}^{\Delta_{k-1}}, \\ &= RC_1 \left[\frac{1-\theta}{1-2\theta} (\Delta_{k-1}^{\frac{1-2\theta}{1-\theta}} - \Delta_k^{\frac{1-2\theta}{1-\theta}}) \right]. \end{aligned}$$

Logo, a desigualdade (3-30) é obtida para $\mu = \frac{2\theta-1}{(1-\theta)RC_1} > 0$ e $\nu = \frac{1-2\theta}{1-\theta} < 0$. Agora, assumindo $b)$, usando definição de h e $m = (\frac{1}{R})^{\frac{1-\theta}{\theta}} \in (0,1)$, adquirimos:

$$\Delta_k^{\frac{-\theta}{1-\theta}} > R\Delta_{k-1}^{\frac{-\theta}{1-\theta}} \Longrightarrow \Delta_k < m\Delta_{k-1},$$

e, levando em conta que $\nu < 0$, segue

$$\Delta_k^{\nu} > m^{\nu}\Delta_{k-1}^{\nu}.$$

Subtraindo Δ_{k-1}^{ν} em ambos os membros da última desigualdade, temos:

$$\begin{aligned} \Delta_k^{\nu} - \Delta_{k-1}^{\nu} &> m^{\nu}\Delta_{k-1}^{\nu} - \Delta_{k-1}^{\nu}, \\ &> (m^{\nu}-1)\Delta_{k-1}. \end{aligned}$$

Como $m^{\nu} - 1 > 0$ e $\Delta_\ell \longrightarrow 0^+$ quando $\ell \longrightarrow +\infty$, existe $\bar{\mu} > 0$ tal que, $(m^{\nu}-1)\Delta_{k-1}^{\nu} \geq \bar{\mu}, \quad k \geq k_1$ e além disso obtemos:

$$0 < \bar{\mu} \leq \Delta_k^{\nu} - \Delta_{k-1}^{\nu}. \tag{3-31}$$

Tomando $\hat{\mu} = \min\{\mu, \bar{\mu}\} > 0$, combinando (3-30) com (3-31) teremos, $0 < \hat{\mu} \leq \Delta_k^{\nu} - \Delta_{k-1}^{\nu}, \quad k \geq k_1$.

Somando essa última desigualdade com k variando de $N_1 \geq k_1$ até algum N maior que N_1 temos:

$$\sum_{k=N_1}^{N} \hat{\mu} \leq \sum_{k=N_1}^{N} (\Delta_k^{\nu} - \Delta_{k-1}^{\nu}),$$

que implica em

$$\hat{\mu}(N - N_1) \leq \Delta_N^{\nu} - \Delta_{N_1-1}^{\nu},$$

ou seja,

$$\Delta_N^{\nu} \geq \hat{\mu}N + \Delta_{N_1}^{\nu} - \hat{\mu}N_1. \tag{3-32}$$

Agora como, $\nu < 0$, podemos supor que $\Delta_{N_1}^{\nu} - \hat{\mu}N_1 \geq 0$. Logo, de (3-32),

$$\Delta_N^{\nu} \;\geq\; \hat{\mu}N,$$

e, consequentemente,

$$\Delta_N \leq CN^{\frac{1}{\nu}}, \quad C = \hat{\mu}^{\frac{1}{\nu}}. \tag{3-33}$$

Portanto, o item (iii) segue por considerando que $q(\bar{x}, x^k) \leq \Delta_k$.

Se $\theta \in (0, \frac{1}{2})$, segue de (3-28) que existe uma constante positiva C_2 tal que

$$\Delta_k \leq C_2(\Delta_{k-1} - \Delta_k),$$

para k suficientemente grande. Então:

$$\Delta_k \;\leq\; \frac{C_2}{1+C_2}\Delta_{k-1}.$$

Observe que por essa última expressão, obtemos:

$$\Delta_k \;\leq\; \left(\frac{C_2}{1+C_2}\right)^k \Delta_0 = Q^k \Delta_0.$$

Logo, o item ii) segue com $Q = \frac{C_2}{1+C_2} \in (0,1)$. □

Referências Bibliográficas

[1] ATTOUCH, H.; BOLTE, J. **On the convergence of the proximal algorithm for nonsmooth functions involving analytic features**. *Math. Program.*, 116(1-2, Ser. B):5–16, 2009.

[2] ATTOUCH, H.; BOLTE, J.; REDONT, P.; SOUBEYRAN, A. **Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the kurdyka-lojasiewicz inequality**. *Math. Oper. Res.*, 35(2):438–457, 2010.

[3] BENTO, G.; SOUBEYRAN, A. **A generalized inexact proximal point method for nonsmooth functions that satisfies kurdyka lojasiewicz inequality**. *Set-Valued and Variational Analysis*, p. 1–17, 2015.

[4] BIERSTONE, E.; MILMAN, P. D. **Semianalytic and subanalytic sets**. *Publications Mathématiques de l'IHÉS*, 67(1):5–42, 1988.

[5] BOCHNAK, J.; COSTE, M.; ROY, M.-F. **Real algebraic geometry**. Springer, 1998.

[6] BOLTE, J.; DANIILIDIS, A.; LEWIS, A. **The lojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems**. *SIAM Journal on Optimization*, 17(4):1205–1223, 2007.

[7] BOLTE, J.; DANIILIDIS, A.; LEWIS, A.; SHIOTA, M. **Clarke subgradients of stratifiable functions**. *SIAM J. on Optimization*, 18(2):556–572, May 2007.

[8] BOLTE, J.; DANIILIDIS, A.; LEY, O.; MAZET, L. **Characterizations of łojasiewicz inequalities: subgradient flows, talweg, convexity**. *Trans. Amer. Math. Soc.*, 362(6):3319–3363, 2010.

[9] BOLTE, J.; SABACH, S.; TEBOULLE, M. **Proximal alternating linearized minimization for nonconvex and nonsmooth problems**. *Math. Program.*, 146(1-2, Ser. A):459–494, 2014.

[10] COSTA, T. M. D. **Análise não suave e aplicações em otimização**. 2011.

[11] DI CONCILIO, A.; GERLA, G. **Quasi-metric spaces and point-free geometry**. *Math. Structures Comput. Sci.*, 16(1):115–137, 2006.

[12] GARCÍA-RAFFI, L. M. **Compactness and finite dimension in asymmetric normed linear spaces**. *Topology Appl.*, 153(5-6):844–853, 2005.

[13] IUSEM, A. **Proximal point methods in optimization**. 20 Colóquio Brasileiro de Matemática, 1995.

[14] LIMA, E. **Curso de análise, vol. 2**. Projeto Euclides. Instituto de Matemática Pura e Aplicada (IMPA), 2010.

[15] LOJASIEWICZ, S. **Une propriété topologique des sous-ensembles analytiques réels, colloques internationaux du cnrs, les équations aux dérivées partielles, 117, ed**. *B. Malgrange (Paris 1962), Publications du CNRS, Paris*, 1963.

[16] MARTINET, B. **Brève communication. régularisation d'inéquations variationnelles par approximations successives**. *ESAIM: Mathematical Modelling and Numerical Analysis - Modélisation Mathématique et Analyse Numérique*, 4(R3):154–158, 1970.

[17] MORDUKHOVICH, B. S.; SHAO, Y. H. **Nonsmooth sequential analysis in Asplund spaces**. *Trans. Amer. Math. Soc.*, 348(4):1235–1280, 1996.

[18] MORENO, F. G.; OLIVEIRA, P. R.; SOUBEYRAN, A. **A proximal algorithm with quasi distance. application to habit's formation**. *Optimization*, 61(12):1383–1403, 2012.

[19] ROCKAFELLAR, R. T. **Monotone operators and the proximal point algorithm**. *SIAM Journal on Control and Optimization*, 14(5):877–898, 1976.

[20] ROCKAFELLAR, R. T.; WETS, R. J.-B.; WETS, M. **Variational analysis**, volume 317. Springer, 1998.

[21] VINTER, R. B. **Optimal Control**. Springer, 2000.