Algoritmo de ponto proximal para otimização em $\mathbb{R}^n$

Ilton Ferreira de Menezes [1], Glaydston de Carvalho Bento [2]

Instituto de Matemática e Estatística, Universidade Federal de Goiás, Campus II- Caixa Postal 131, CEP 74001-970 - Goiânia, GO, Brasil.

E-mail: iltomenezesufg@hotmail.com[1]; glaydston@mat.ufg.br[2]

Palavras chaves: ponto proximal, funções convexas, subdiferencial.

1 Introdução

Considere o seguinte problema de otimização

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

onde $f : \mathbb{R}^n \to \mathbb{R}$ (função objetivo) é não necessariamente diferenciável. O método de ponto proximal, introduzido por Martinet [1] e Rockafellar [2] na década de 70, é um método iterativo utilizado para resolver o problema (1). O método de ponto proximal gera uma sequência $\{x^k\} \subset \mathbb{R}^n$ como segue: dado $x_0 \in \mathbb{R}^n$,

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

onde $\{\lambda_k\}$ é uma sequência de números positivos. Sendo $f_k := f(x) + \lambda_k ||x - x^k||^2$ uma regularização de f, espera-se que f_k forneça algum tipo de aproximação do minimizador (caso exista) da função objetivo f.

Neste trabalho, seguindo [3], mostramos que se f é convexa e o conjunto dos minimizadores é não vazio, então a sequência gerada por (2) está bem definida e converge um minimizador da função objetivo.

2 Objetivos

- Objetivos gerais: Estudo de elementos de análise convexa e análise não diferenciável;
- Objetivos específicos: Estudo do método de ponto proximal clássico para otimização convexa.

[1] Orientando PIBIC CNPq

[2] Orientador

Revisado pelo orientador.

3 Metódologia

Estudo de livros e artigos e discussões semanais com o orientador.

4 Resultados e Discussões

4.1 Resultados preliminares

Nesta seção apresentamos algumas definições e resultados básicos de análise convexa que podem ser encontrados, por exemplo, em [4, 5, 6]. Daqui em diante, f sempre designará uma função definida em $\mathbb{R}^n$ e assumindo valores em $\mathbb{R}$, salvo menção explícita em contrário.

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

Definição 2. *Seja $D \subset \mathbb{R}^n$ um conjunto convexo. Uma função $f : D \to \mathbb{R}$ é dita ser convexa em D quando, para quaisquer $x, y \in D$ e $\alpha \in [0,1]$, tem-se*

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

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

Proposição 1. *Seja f é uma função convexa e $h : \mathbb{R}^n \to \mathbb{R}$ é uma função estritamente convexa. Então, $f + h : \mathbb{R}^n \to \mathbb{R}$ é uma função estritamente convexa.*

Demonstração. Segue imediatamente da última definição. □

Proposição 2. *Seja f uma função estritamente convexa. Se f possui um minimizador, ele é único.*

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

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

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

Teorema 3. *(Desigualdade de Jensen) Seja f uma função convexa. Então, para todo $m \in \mathbb{N}$, $x_i \in \mathbb{R}^n$ e $\lambda_i \geq 0$, $i = 1, ..., m$, tais que $\sum_{i=1}^{m} \lambda_i = 1$, tem-se*

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

Demonstração. A prova é feita por indução a partir da Definição 2. □

Definição 3. *Seja $x \in \mathbb{R}^n$ e $L > 0$. Uma função f é dita ser localmente Lipschitz em x com constante L, se existe uma $\epsilon > 0$ tal que*

$$|f(x) - f(y)| \leq L\|x - y\|, \qquad x, y \in \mathbb{R}^n.$$

Teorema 4. *Seja f uma função convexa. Então, f é localmente Lipschitz.*

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

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

com vértices $u^1, \dots, u^m$. É conhecido que $m = 2^n$ e $S = \operatorname{conv}(\{u^1, \dots, u^m\})$ (envoltória convexa de $\{u^1, \dots, u^m\}$), ver [5, página 121]. Então, para cada $y \in S$, existem $\lambda_i \geq 0$, $i = 1, \dots, m$, com $\sum_{i=1}^{m} \lambda_i = 1$, tal que $y = \sum_{i=1}^{m} \lambda_i u_i$. Denote por M o máximo de f em $\{u^1, \dots, u^m\}$. Visto que f é convexa e $\sum_{i=1}^{m} \lambda_i = 1$, do Teorema 4 segue que

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

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

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

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

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

$$\alpha \leq f(z) \leq \beta, \qquad z \in B(u, 2\delta). \tag{3}$$

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

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

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

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

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

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

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

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

Assim, combinando última desigualdade com (4), concluimos que

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

Procedendo de modo análogo, é possível mostrar que

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

Portanto,

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

e o resultado segue. □

Corolário 5. *Seja f uma função convexa. Então, f é contínua.*

Demonstração. Segue do Teorema 4. □

Definição 4. *Sejam x e v um ponto e uma direção em $\mathbb{R}^n$, respectivamente. A derivada direcional de uma função f em x na direção v é definida por*

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

Proposição 6. *Seja $x \in \mathbb{R}^n$ e f uma função convexa. Então, a derivada direcional em cada direção $v \in I\!R^n$ existe e satisfaz*

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

Demonstração. Tome $v \in \mathbb{R}^n$ e definamos $\varphi : \mathbb{R} \to \mathbb{R}$ dada por

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

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

$$\varphi(t_2) - \varphi(t_1) = \frac{(t_2 - t_1)f(x) + t_1 f(x + t_2 v) - t_2 f(x + t_1 v)}{t_1 t_2}.$$

que, depois de algumas manipulações algébricas, fornece

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

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

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

de onde concluimos que φ decresce quando t tende a zero por valores positivos.

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

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

Usando que f é convexa, última igualdade nos fornece

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

de onde obtemos

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

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

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

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

Definição 5. *Seja S um conjunto não vazio em $\mathbb{R}^n$. A função $\sigma_S : \mathbb{R}^n \to \mathbb{R}$, definida por*

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

é chamada função suporte de S.

Proposição 7. *Sejam $A, B \subset \mathbb{R}^n$ conjuntos não vazios, convexos e compatos. Então,*

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

Definição 6. *Dado $x \in \mathbb{R}^n$, o subdiferencial de uma função f em x, denotado por $\partial f(x)$, é o conjunto dos vetores $s \in \mathbb{R}^n$ satisfazendo*

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

Teorema 8. *Seja f uma função convexa e $x \in \mathbb{R}^n$. Então,*

(i) $f'(x, v) = \sigma_{\partial f(x)}(v)$, *para todo* $v \in \mathbb{R}^n$*;*

(ii) $\partial f(x)$ *é um conjunto não vazio, convexo e compacto tal que* $\partial f(x) \subset B(0; L)$, *onde* L *é uma constante de Lipschitz de* f *em* x.

Demonstração. Ver, por exemplo, [6, pág. 14 e 15]. □

Teorema 9. *Sejam $f, g : \mathbb{R}^n \to \mathbb{R}$ funções convexas. Então, para cada $x \in \mathbb{R}^n$,*

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

Demonstração. Tome $v \in \mathbb{R}^n$. Visto que f e g são convexas, pelo item (ii) do Teorema 8 podemos concluir que existem $\tilde{s}_1 \in \partial f(x)$ e $\tilde{s}_2 \in \partial g(x)$ tal que

$$\sigma_{\partial f(x)}(v) = \langle \tilde{s}_1, v \rangle, \qquad \sigma_{\partial g(x)}(v) = \langle \tilde{s}_2, v \rangle. \tag{5}$$

Além disso, da Definição 5 segue que

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

Assim, de (5), obtemos que

$$\sigma_{\partial f(x) + \partial g(x)}(v) \geq \sigma_{\partial f(x)}(v) + \sigma_{\partial g(x)}(v). \tag{6}$$

Por outro lado,

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

o que implica que

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

Portanto, o resultado segue da última desigualdade combinada com a desigualdade (6). □

Teorema 10. *Sejam $f, g : \mathbb{R}^n \to \mathbb{R}$ funções convexas e $x \in \mathbb{R}^n$. Então,*

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

Demonstração. Tome $v \in \mathbb{R}^n$. Visto que f e g são convexas, pelo Teorema 8, segue que $\partial f(x)$ e $\partial g(x)$ são conjuntos não vazios, convexos e compactos tais que

$$\sigma_{\partial f(x)} v = f'(x, v), \qquad \sigma_{\partial g(x)} v = g(x, v). \tag{7}$$

Em particular, $\partial f(x) + \partial g(x)$ é um conjuno não vazio, convexo e compacto e, pelo Teorema 9,

$$\sigma_{\partial f(x) + \partial g(x)} v = \sigma_{\partial f(x)} v + \sigma_{\partial g(x)} v. \tag{8}$$

Por outro lado, visto que $(f+g)'(x, v) = f'(x, v) + g'(x, v)$, usando o item (i) do Teorema 8 junto com (7), obtemos

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

Combinando última igualdade com (8) junto com a arbitrariedade de v, o resultado segue da Proposição 7. $\square$

Definição 7. *Seja f uma função contínua. Diz-se que f e uma 1-coerciva se*

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

Teorema 11. *Seja f uma funcão contínua. Se f é 1-coerciva, então f possui pelo menos um mínimo global.*

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

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

Mas isto nos diz que

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

Visto que f é convexa, do Corolário (5), temos que f é contínua e consequentemente, $L_f(Mr)$ é fechado. Portanto, da última inclusão, $L_f(Mr)$ é um conjunto compacto. Assim, pelo Teorema de Weierstrass, f possui pelo menos um mínimo quando restrita a $L_f(Mr)$. Seja $y \in \mathbb{R}^n$ um minimizador de f restrita a $L_f(Mr)$. Afirmamos que y é um minimizador irrestrito de f. De fato, visto que y é um minimizador de f quando restrita ao conjunto $L_f(Mr)$, temos que $f(y) \leq f(x)$, para todo $x \in L_f(Mr)$. Por outro lado, para todo $x \notin L_f(Mr)$,

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

e o resultado segue. □

Proposição 12. *Seja f uma função convexa e $x \in \mathbb{R}^n$. Uma condição necessaria e suficiente para que x seja um minimizador de f é que $0 \in \partial f(x)$.*

Demonstração. Segue imediatamente da definição de subdiferencial. □

4.2 O método de ponto proximal

Nesta seção apresentamos uma análise de convergência do método de ponto proximal, ver por exemplo, [7, 3].

Proposição 13. *Seja f uma função convexa. A sequência $\{x^k\}$ gerada pelo método de ponto proximal* (2) *está bem definida.*

Demonstração. Pelo item (ii) do Teorema 8, para todo $k \in \mathbb{N}$, existe $s^k \in \mathbb{R}^n$ tal que

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

Assim, visto que $f_k(x) = f(x) + \lambda_k ||x^k - x||^2$, temos que

$$f_k(x) \geq f(x^k) + \langle s^k, x - x^k \rangle + \lambda_k ||x^k - x||^2.$$

Dividindo ambos os membros da desigualdade acima por $||x - x^k||$, obtemos

$$\frac{f_k(x)}{||x - x^k||} \geq \frac{f(x^k)}{||x - x^k||} + \left\langle s^k, \frac{x - x^k}{||x - x^k||} \right\rangle + \lambda_k ||x^k - x||,$$

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

$$\frac{f_k(x)}{||x - x^k||} \geq \frac{f(x^k)}{||x - x^k||} - \|s^k\| + \lambda_k ||x^k - x||.$$

Visto que $f(x^k)$ e $\|s^k\|$ são constantes e $\lambda_k > 0$, fazendo $\|x - x^k\|$ ir a $+\infty$ em ambos os membros da última desigualdade, segue que f_k é 1-coerciva. Assim, como $\lambda_k ||x^k - x||^2$ é estritamente convexa, o resultado segue do Teorema 11 combinado com as Proposições 1 e 2. □

Teorema 14. *Seja $f : \mathbb{R}^n \to \mathbb{R}$ uma função convexa. A sequência $\{x^k\} \subset \mathbb{R}^n$ gerada por (2) é caracterizada pela relação*

$$2\lambda(x^k - x^{k+1}) \in \partial f(x^{k+1}).$$

Demonstração. Sendo x^{k+1} dada por (2), o resultado segue da Proposição 12 combinada com o Teorema 10. $\square$

Teorema 15. *Seja f uma função convexa e $x \in \mathbb{R}^n$. Se a sequência $\{x^k\}$ é gerada por (2), então vale a desigualdade*

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

Demonstração. Note que

$$\|x - x^{k+1}\|^2 = \|x - x^k + x^k - x^{k+1}\|^2 = \|x - x^k\|^2 - \|x^{k+1} - x^k\|^2 + 2\langle x^k - x^{k+1}, x - x^{k+1}\rangle \quad (9)$$

Agora, visto que f é convexa, pelo Teorema 14, temos

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

Assim, pela definição de $\partial f(x^{k+1})$,

$$f(x) \geq f(x^{k+1}) + \langle 2\lambda_k(x^k - x^{k+1}), x - x^{k+1}\rangle = f(x^{k+1}) + 2\lambda_k\langle x^k - x^{k+1}, x - x^{k+1}\rangle,$$

e, consequentemente,

$$\frac{1}{\lambda_k}(f(x) - f(x^{k+1})) \geq 2\langle x^k - x^{k+1}, x - x^{k+1}\rangle.$$

Portanto, o resultado segue da última desigualdade combinada com a desigualdade (9). $\square$

Corolário 16. *Seja $\{x^k\}$ a sequência gerada por (2). Então, vale a desigualdade*

$$||x - x^{k+1}||^2 \leq ||x - x^k||^2 + \frac{1}{\lambda_k}(f(x) - f(x^{k+1})), \quad (10)$$

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

Demonstração. Segue do Teorema 15. $\square$

Definição 8. *Seja $\emptyset \neq U \subset \mathbb{R}^n$. Uma sequência $\{x^k\} \subset \mathbb{R}^n$ é dita ser Fejér convergente a U com respeito a norma euclidiana se*

$$\|x^{k+1} - u\| \leq \|x^k - u\|, \qquad u \in U. \quad (11)$$

Corolário 17. *Se* $\{x^k\}$ *é Fejér convergente ao conjunto* $\emptyset \neq U^* \subset I\!R^n$*, então* $\{x^k\}$ *é uma sequência limitada. Se, além disso, um ponto de acumulação* $\bar{x}$ *da sequência* $\{x^k\}$ *pertence a* U*, então* $\bar{x} = \lim_{k\to+\infty} x^k$*.*

Demonstração. Dado $u \in U$, a desigualdade (11) implica que $||u - x^k||^2 \leq ||u - x_0||^2$, para todo k. Deste modo a sequência $\{x^k\}$ é limitada. Agora sejam $\bar{x} \in U$ um ponto de acumulação de $\{x^k\}$ e $\{x_{k_j}\}$ uma subsequência de $\{x^k\}$ tal que $\lim_{j\to+\infty}(x_{k_j})=\bar{x}$. Como $\bar{x} \in U$, segue de (11) que a sequência de números reais positivos $\{||x^k - \bar{x}||\}$ é monótona não-decrescente e possui uma subsequência, a saber $\{||x_{k_j} - \bar{x}||\}$, convergindo para zero. Então a sequência converge para zero, isto é, $\lim_{k\to+\infty} ||x^k - \bar{x}||=0$, o que significa $\lim_{k\to+\infty} x^k=\bar{x}$. $\square$

Teorema 18. *Sejam* f *uma função convexa e* $\{x^k\}$ *a sequência gerada por* (2)*. Se a sequência* $\{\lambda_k\}$ *é tal que* $\displaystyle\sum_{k=0}^{\infty} \frac{1}{\lambda^k} = +\infty$*, então* $\lim_{k\to+\infty} f(x^k) = f_*$*, onde* $f_*=\inf_{x\in\mathbb{R}^n} f(x)$*. Se, além disso, o conjunto* U^* *dos minimizadores de* f *é não vazio, então* $\lim_{x\to\infty}(x^k) = x_*$*, com* $x_* \in U^*$*.*

Demonstração. Seja U^* o conjunto dos minimizadores da função convexa f. Se $x^k \in U^*$, para algum $k > 0$, é fácil ver que $x^k = x^{k+1} = x^{k+2} = x^{k+3} = \ldots$. Assim, $f(x^k) = f_*$ e não há mais nada a fazer. Agora, suponha que para todo $k \in \mathbb{N}$, $x^k \notin U^*$. Por (2), segue que a sequência $\{f(x^k)\}$ é estritamente decrescente. Afirmamos que $\lim_{k\to+\infty} f(x^k) = f_*$. De fato, suponhamos, por absurdo, que $\lim_{k\to+\infty} f(x) > f_*$. Então, existe $x \in I\!R^n$ e $\delta > 0$ tal que

$$f(x) < f(x^k) - \delta, \qquad k \in \mathbb{N}. \tag{12}$$

Substituindo a desigualdade (12) na desigualdade (10), obtemos

$$\|x - x^{k+1}\|^2 \leq \|x - x^k\|^2 - \frac{\delta}{\lambda_k}, \qquad k \in \mathbb{N}. \tag{13}$$

e, consequentemente,

$$\frac{\delta}{\lambda_k} \leq ||x - x^k||^2 - ||x - x^{k+1}||^2,$$

para todo $k > 0$. Somando termo a termo

$$\sum_{k=0}^{j} \frac{1}{\lambda_k} \leq \frac{1}{\delta}(||x - x_0||^2 - ||x - x^{j+1}||^2) \leq \frac{1}{\delta}||x - x_0||^2,$$

para todo j, o que contraria $\displaystyle\sum_{k=0}^{\infty} \frac{1}{\lambda_k} = +\infty$. Logo, $\lim_{k\to+\infty} f(x^k) = f_*$. Suponhamos agora que $U^* \neq \emptyset$ e tome $\bar{x} \in U^*$. Deste modo $f(\bar{x}) \leq f(x^k)$, para todo k. Substituindo x por $\bar{x}$

em (10), obtemos

$$||\bar{x} - x^{k+1}||^2 \leq ||\bar{x} - x^k||^2, \qquad k \in \mathbb{N},$$

assegurando que a sequência $\{x^k\}$ é Fejér convergente a U^*. Logo, pelo corolário 17, temos que $\{x^k\}$ é limitada. Seja $\{x^{k_j}\}$ uma subsequência convergente de $\{x^k\}$, tal que $\lim_{j\to+\infty} x^{k_j} = x_*$. Da primeira parte $\lim_{j\to+\infty} f(x^{k_j}) = f_*$. Assim, visto que f é convexa, em particular f contínua (Corolário 5) e, consequentemente, $f(x_*) = f_*$. Mas isto nos diz que $x_* \in U^*$. Portanto, novamente pelo corolário 17, concluimos que a sequência $\{x^k\}$ converge a x_*, isto é, $\lim_{k\to\infty}(x^k) = x_*$. $\square$

5 conclusão

Neste trabalho estudamos o método de ponto proximal e apresentamos uma análise de convergência da sequência gerada pelo mesmo a uma solução do problema de minimização convexa irrestrito, no caso que a função objetivo é não necessariamente diferenciável.

Referências

[1] Martinet B., Régularization, d'inéqations variationelles par approximations successives, Revue Française d'Informatique et de Recherche Opertionelle, (1970), 154-159.

[2] Rockafellar, R. T., Monotone operators and the proximal point method, SIAM, J. Control. Optim., 14 (1970), 154-159.

[3] Silva, S. R. P, Algoritmo de Ponto Proximal para Otimização em $\mathbb{R}^n$, UFG, Dissertação de mestrado, (2002).

[4] Tiel, J. V., Convex analysis: an introductory text, Royal Neterland Meteorological Institute, 1984.

[5] Izmailov, A., Solodov, M., Otimização Volume 1. Condições de otimalidade, elementos de análise conexa e de dualidade Rio de Janeiro: IMPA,2005.

[6] Mäkelä, M. M., Nonsmooth optimization: analysis and algorithms with applications to optimal control Marko M., Mäkelä and Pekka Neitaanmäki.

[7] Iusen, A.N., Proximal point methods in optimization, 20. Colóquio Brasileiro de Matemática,IMPA, (1995).