**Método do Gradiente Projetado Para Funções Convexas Generalizadas**

Milton Gabriel Garcia dos SANTOS [1].

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

E-mail: matematica_gabriel@yahoo.com.br

**Palavras chaves:** Gradiente Projetado, Quase convexidade, Pseudo Convexidade, Busca de Armijo.

# 1 Introdução

Em 2003, IUSEM (veja [1]]) estudou algumas propriedades de convergência do Método do Gradiente Projetado com a Busca de Armijo sobre as direções viáveis para problemas de otimização convexa, isto é, o problema inicial é:

$$\min_{s.a.\ x\ \in\ C} f(x), \tag{1}$$

onde $C \subset \mathbb{R}^n$ é não vazio, fechado e convexo e $f : \mathbb{R}^n \to \mathbb{R}$ é uma função continuamente diferenciável e convexa. O Método do Gradiente Projetado conceitual e apresentado como:

**Método 1** (Método do Gradiente Projetado com regra de Armijo )**.**

Inicialização. *Tome* $0 < \widehat{\beta} < \overline{\beta}$ *and* $x_0 \in C$. *Set* $k = 0$.
Critério de parada. *Se* $x^k$ *é estacionário então PARE. Caso contrário.*
Passo iterativo. *Calcule*

$$z^k = P_C(x^k - \beta_k \nabla f(x^k)), \qquad x^{k+1} = x^k + \gamma_k(z^k - x^k),$$

*onde* $\{\beta_k\} \subset [\widehat{\beta}, \overline{\beta}]$ *tal que* $\gamma_k = 2^{-j(k)}$ *sendo que*

$$j(k) = \{ \min j \in \mathbb{N} : f(x^{k+1}) \leq f(x^k) + \sigma 2^- j \nabla f(x^k)^t z^k - x^k\}, \ e \ \sigma \in (0,1).$$

*e VOLTE para* Critério de parada*.*

Foi provado para o caso convexo que:

[1] Bolsista CAPES

(i) Dada $\{x^k\}$ uma seqüência gerada pelo algoritmo acima, temos que se ela é infinita e, se o problema inicial possui ao menos uma solução, então os pontos de acumulação da seqüência gerada, quando existem, são estacionários.

(ii) Sob a hipótese de $f$ ser convexa e supondo que o problema inicial possui solução a seqüência $x^k$ infinita gerada pelo algoritmo acima converge a alguma solução do problema.

Tendo em mente os resultados expostos acima e baseado então, no artigo [2], nosso trabalho é estudar casos um pouco mais gerais, isto é, no caso que as funções objetivo são quase convexas ou pseudo convexas. Para este caso o Algoritmo continua sendo o mesmo, e a busca segue sendo a Busca de Armijo sob direções viáveis.

# 2 Resultados e Discussão

## 2.1 Definições Preliminares

**Definição 1.** *Dizemos que um ponto $\bar{x} \in \mathbb{R}^n$ é um ponto estacionário ou crítico para problema inicial, quando*

$$\langle \nabla f(\bar{x})\,,\, x - \bar{x} \rangle \geq 0, \qquad \forall\, x \in C.$$

**Definição 2.** ***(Funções Quase-convexas)*** *Uma função real $f$ definida sobre um conjunto $C \subset \mathbb{R}^n$ é dita quase-convexa sobre $\bar{x} \in C$ (com respeito a $C$) se para cada $x$ tal que $f(x) \leq f(\bar{x})$, a função $f$ não assume um valor maior que $f(\bar{x})$ sobre cada ponto da intersecção do segmento fechado $[\bar{x}, x]$, ou equivalentemente:*

$$\left.\begin{array}{l} x \in C \\ f(x) \leq f(\bar{x}) \\ 0 \leq \lambda \leq 1 \\ (1-\lambda)\bar{x} + \lambda x \in C \end{array}\right\} \Longrightarrow \quad f[(1-\lambda)\bar{x} + \lambda x] \leq f(\bar{x}).$$

*E dizemos que $f$ é quase-convexa em $C$ se é quase-convexa em cada ponto $x \in C$.*

Quando $f : C \to \mathbb{R}$ é diferenciável e quase-convexa em $x$, então a definição acima é equivalente a condição:

$$f(y) \leq f(x) \Rightarrow \langle \nabla f(x)\,,\, y - x \rangle \leq 0, \qquad \forall\, x, y \in C.$$

**Definição 3.** *Seja $f$ uma função real definida sobre algum conjunto aberto do $\mathbb{R}^n$ contendo o conjunto $C$, $f$ é dita pseudo-convexa sobre $\bar{x} \in C$ (com respeito a $C$) se é diferenciável sobre $\bar{x}$ e dado $x \in C$ temos,*

$$\langle \nabla f(\bar{x}) \, , \, x - \bar{x} \rangle \geq 0 \Rightarrow f(x) \geq f(\bar{x}).$$

*Dizemos que $f$ é pseudo-convexa sobre $C$, se é pseudo-convexa sobre cada $x \in C$.*

**Teorema 1.** *Sejam $C \subset \mathbb{R}^n$ um conjunto convexo, e $f$ uma função real definida em algum conjunto aberto que contém $C$. Se $f$ é pseudo-convexa em $C$ então $f$ é quase-convexa em $C$.*

*Demonstração.* Veja página 143 de [3]. □

Seja $w \in \mathbb{R}^n$ e $C \subset \mathbb{R}^n$ não vazio, fechado e convexo. O operador $P_C : \mathbb{R}^n \rightarrow C$ definido por

$$P_C(w) := \operatorname{argmin} \|x - w\|, \; x \in C,$$

é chamado a *Projeção* de $w$ sobre o conjunto $C$.

## 2.2 Análise de Convergência

**Proposição 2.** *Sejam $\{x^k\}$ e $\{z^k\}$ a sequências geradas pelo método* (1)*.*

**i.** *A iteração $x^k \in C$ para todo k;*

**ii.** *Se $x^k$ não é ponto estaionário, então $\langle \nabla f(x^k), z^k - x^k \rangle < 0$;*

**iii.** *Para cada k, j(k) está bem definido.*

**Proposição 3.** *Se o problema inicial* (1) *possui solução. Então todo ponto de acumulação de $\{x^k\}$ é ponto estacionário, onde essa seqüência é gerada pelo método* (1)*.*

**Lema 4.** *Para todo $x \in C$ e todo k, nós temos:*

$$\|x^{k+1} - x\|^2 \; \leq \; \|x^k - x\|^2 \; + \; \eta(f(x^k) - f(x^{k+1})) \; + \; 2\gamma_k \beta_k \langle \nabla f(x^k), x - x^k \rangle,$$

*onde $\eta = 2\widetilde{\beta}/\delta$.*

**Corolário 5.** *Se f é quaseconvexa, $x \in C$ e $f(x) \leq f(x^k) \; \forall \, k$, então*

$$\eta f(x) \; \leq \; \|x^{k+1} - x\|^2 + \eta f(x^{k+1}) \; \leq \; \|x^k - x\|^2 + \eta f(x^k),$$

*para todo k. Em particular, $\{\|x^k - x\|^2 + \eta f(x^k)\}$ e $\{x^k\}$ são sequências convergentes.*

**Teorema 6.** *[Principal] Assumindo que f é quase convexa. Então temos:*

**i.** *A sequência* $\{x^k\}$ *converge para um ponto estacionário, ou*

**ii.** *O problema inicial não tem solução,*

$$\lim_{k\to\infty} \|x^k\| = \infty, \qquad \lim_{k\to\infty} f(x^k) = \inf\{f(x)|x \in C\}.$$

*Demonstração.* Existe então, dois casos para se analisado. Primeiro caso, quando a sequência possui ao menos um ponto de acumulação e o segundo quando a sequência não possui pontos de acumuulação.

1º Caso) Seja $\bar{x}$ um ponto de acumulação de $x^k$, então existe uma subsequência $x^{k_j}$ convergindo a $\bar{x}$ e daí temos que $f(x^{k_j}) \to f(\bar{x})$ e como $\{f(x^k)\}$ é convergente resulta que $f(x^k) \to f(\bar{x})$, o que nos dá $f(\bar{x}) \leq f(x^k)$ e pelo corolário (5) $x^k$ é convergente e pela proposição 3 $\bar{x}$ é um ponto estacionário.

2º Caso ) Suponhamos que a sequência $x^k$ não possui nenhum ponto de acumulação, ou seja, $\lim_{k\to\infty} \|x^k\| = \infty$ e suponha que o problema inicial possua um minimizador, isto é, existe $\widehat{x} \in C \; tal \; que \; f(\widehat{x}) \leq f(x^k)$ para todo k. Pelo corolário 4 $x^k$ converge, o que contradiz a hipóteses de $\lim_{k\to\infty} \|x^k\| = \infty$ e então o problema inicial não possui solução.
Além disso, afirmamos que $\lim_{k\to\infty} f(x^k) = \inf\{f(x) : x \in C\}$
Observamos que se $\lim_{k\to\infty} f(x^k) = -\infty \Rightarrow \lim_{k\to\infty} f(x^k) = \inf\{f(x) : x \in C\}$.
Agora se $\lim_{k\to\infty} f(x^k) = \bar{f} \geq \inf\{f(x) : x \in C\}$. Assumimos que $\bar{f} > \inf\{f(x) : x \in C\}$.
Pela continuidade de f existe um $\bar{x}$ tal que $f(\bar{x}) = \bar{f}$ e $f(\bar{x}) \leq f(x^k)$ para todo k, pois essa sequência é decrescente. Isto implica pelo corolário (5) que $x^k$ é convergente, contradizendo nossa hipótese inicial, ou seja, $\lim_{k\to\infty} f(x^k) = \inf\{f(x) : x \in C\}$.

□

**OBS:** Provamos então que, se a seqüência possui pontos de acumulação, ela converge para algum desses pontos e o mesmo é estacionário. Se a seqüência não possui pontos de acumulação, o conjunto solução é vazio.

Defina $T = \{x \in C : f(x) \leq f(x^k) \, \forall \, k\}$, podemos notar que se o conjunto solução do problema inicial não for vazio ou existir pontos de acumulação temos que $T \neq \emptyset$

**Corolário 7.** *Assumindo que f é pseudo convexa. então o conjunto solução do problema inicial é não vazio se, e somente se, existe ao menos um ponto de acumulação da sequência.*

*Demonstração.* Pelo teorema 5 da página 143 de [3], pseudo convexidade nos garante que f é estritamente quase convexa e como f é continua pelo teorema 3 da pagina 139 de [3] f é quase convexa.

$(\Rightarrow)$ Suponha que o conjunto solução é não vazio, isto é, existe um minimizador, $\bar{x}$ tal que $f(\bar{x}) \leq f(x^k) \; \forall \; k$ pelo corolário 5 $x^k$ converge e pelo teorema 6 converge a um ponto estacionário e esse ponto é de acumulação.

$(\Leftarrow)$ Suponha que exista pelo menos um ponto de acumulação $\bar{x}$ de $x^k$ o que implica que $f(\bar{x}) \leq f(x^k)$ para todo k.

Pelo corolário 4 e 5 e o teorema 6, $x^k$ converge a um ponto estacionário, isto é, $\langle \nabla f(\bar{x}), x - \bar{x} \rangle \geq 0$, mas pela pseudo convexidade temos que $f(x) \geq f(\bar{x})$ para todo $x \in C$ , logo o conjunto solução é não vazio, o que conclui a demonstração. □

# 3 Conclusões

Com este trabalho podemos ver alguns resultados importantes para casos mais gerais, mesmo enfraquecendo a hipóteses de convexidade da nossa função objetivo. Provando boas propriedades de convergência, como a existência de solução somente se a sequência possuir ponto de acumulação, e que se existir esses pontos a sequência gerada pelo Algoritmo converge para algum desses pontos de acumulação, e estes são estacionários.

# Referências


[1] A. N. IUSEM - *On the convergence properties of the projected gradient method for convex optimization*, Computational and Applied Mathematics, 22, 37-52 (2003).

[2] J.Y. Bello Cruz, L.R. Lucambio Pérez - *Convergence of a projected gradient method variant for quasiconvex objectives*, Nonlinear Analysis, 10, 2917-2922 (2010).

[3] Olvi L. Mangasarian - *Nonlinear Programming*, In Applied mathematics., (1994).