TEMA Tend. Mat. Apl. Comput., **7**, No. 1 (2006), 53-61.

Funções Invexas Diferenciáveis e o Teorema de Karush-Kuhn-Tucker[1]

J. CERVELATI[2], M.A. ROJAS-MEDAR[3], Departamento de Matemática Aplicada, IMECC, UNICAMP, 13081-970 Campinas, SP, Brasil.

Resumo. Em 1980 surgiu o conceito de função invexa, esta classe de funções é maior do que a classe de funções convexas. Após esta descoberta, vários estudos foram feitos no intuito de utilizar esta nova classe de funções para garantir otimalidade para problemas de Programação Matemática. O objetivo deste trabalho é mostrar que as Condições de Karush-Kuhn-Tucker garantem otimalidade global se todas as funções do problema forem, ao invés de convexas, invexas.

1. Introdução

Em 1981, Hanson [6] introduziu uma nova classe de funções. Ele considerou funções diferenciáveis para as quais exista uma função $\eta : \Omega \times \Omega \to \mathbb{R}^n$ tal que:

$$f(x) - f(u) \geq \nabla f(u)\eta(x,u), \forall x, u \in \Omega. \tag{1.1}$$

É óbvio que funções convexas diferenciáveis satisfazem (1.1) com $\eta(x,u) = x-u$. No mesmo trabalho, Hanson mostrou que esta classe de funções é "maior" do que a classe de funções convexas diferenciáveis.

Foi observado que, se $F : \mathbb{R}^n \to \mathbb{R}$ é uma função convexa diferenciável e $\phi : \mathbb{R}^n \to \mathbb{R}^n$ é uma aplicação diferenciável com derivada inversível, então tomando $f = F \circ \phi$ e $x, u \in \mathbb{R}^n$, temos que:

$$\begin{aligned} f(x) - f(u) &= F(\phi(x)) - F(\phi(u)) \\ &\geq F'(\phi(u)).[\phi(x) - \phi(u)] \\ &= f'(u).[\phi'(u)]^{-1}.[\phi(x) - \phi(u)] \\ &= f'(u).\eta(x,u), \end{aligned}$$

onde $\eta(x,u) = \phi'(u)^{-1}.[\phi(x) - \phi(u)]$. Das considerações anteriores podemos notar o que resta da convexidade de F após o seu domínio ter sido distorcido por ϕ. Essa propriedade é justamente a propriedade (1.1). Baseado nestas observações, Craven [4] chamou tal propriedade de *invexity* e denominou estas funções de *invex functions* (uma alusão à invariante convexo). Em português, denominamos *invexidade* e *funções invexas*, respectivamente.

[1] Trabalho vencedor do 1º lugar do Prêmio "Beatriz Neves".
[2] ju@ime.unicamp.br; Bolsista de Iniciação Científica pela FAPESP Proc.02/12669-8.
[3] marko@ime.unicamp.br; Parcialmente financiado pelo CNPq, Proj.301354103-0.

A partir desta descoberta, vários estudos foram feitos no intuito de mostrar que no problema:

$$\begin{cases} Minimizar & f_0(x) \\ Sujeito \quad a & f_i(x) \quad \leq 0, i = 1, ..., p, \end{cases} \tag{1.2}$$

onde $f_0(x), f_i(x) : \Omega \to \mathbb{R}, i \in I$, são diferenciáveis no aberto $\Omega \subset \mathbb{R}^n$, ao invés das condições usuais de convexidade, $f_0, f_i, i \in I$, satisfazem (1.1), para uma mesma função $\eta(x, u)$, então as Condições de Karush-Kuhn-Tucker garantem otimalidade global.

2. Funções Invexas Diferenciáveis

Suponha Ω um aberto não vazio do $\mathbb{R}^n$ e $f : \Omega \to \mathbb{R}$ diferenciável.

Definição 2.1. *Dizemos que f é invexa em $u \in \Omega$ se, para todo $x \in \Omega$, existe uma função $\eta : \Omega \times \Omega \to \mathbb{R}^n$ tal que*

$$f(x) - f(u) \geq \nabla f(u)\eta(x, u).$$

Dizemos simplesmente que f é invexa em Ω, se f é invexa em cada ponto de Ω.

Definição 2.2. *Dizemos que $u \in \Omega$ é um ponto estacionário de f, se $\nabla f(u) = 0$. Um ponto $u \in \Omega$ é um ponto de mínimo global de f (em relação a Ω), se $f(u) \leq f(x)$ para todo $x \in \Omega$.*

Teorema 2.1. *f é invexa em Ω se e somente se todo ponto estacionário de f é um mínimo global em Ω.*

Demonstração: Se f é invexa, é fácil verificar que $\nabla f(u) = 0 \Rightarrow f(x) \geq f(u)$ para todo $x \in \Omega$. Reciprocamente, suponha que todo ponto estacionário de f seja mínimo global em Ω, então

$$\eta(x, u) = \begin{cases} \frac{f(x)-f(u)}{\nabla^2 f(u)} \nabla f(u), & \text{se} \quad \nabla f(u) \neq 0 \\ 0, & \text{se} \quad \nabla f(u) = 0. \end{cases}$$

□

Corolário 2.1.1. *Se f não tem pontos estacionários então f é invexa.*

Exemplo 1: Consideremos a função $f : \mathbb{R} \to \mathbb{R}$ dada por $f(x) = 1 - \exp(-x^2)$. a função é claramente não convexa. No entanto f é invexa com $\eta(x, u)$ dada por:

$$\eta(x, u) = \begin{cases} \frac{-\exp(-x^2)+\exp(-u^2)}{2u\exp(-u^2)}, & \text{se} \quad u \neq 0 \\ 0, & \text{se} \quad u = 0. \end{cases}$$

Neste exemplo, podemos mostrar que a função é invexa sem necessariamente exibir a função $\eta(x, u)$. Basta notas que a derivada de f é dada por $\nabla f(x) = 2x\exp(-x^2)$. Logo, seu único ponto estacionário é $x = 0$, e é fácil verificar que tal ponto é mínimo global de f sobre $\mathbb{R}$. Então, pelo Teorema 2.1., f é invexa sobre $\mathbb{R}$

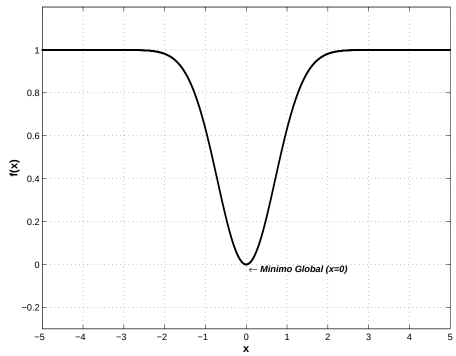

Figura 1: Gráfico da função do Exemplo 1

Teorema 2.2. *Se para cada $i = 1, 2, ...p, \lambda_i \geq 0$ e f_i são funções invexas para uma mesma função $\eta(x, u)$, então $\sum_{i=1}^{p} \lambda_i f_i$ é uma função invexa.*

Demonstração: Pela definição de invexidade $f_i(x) - f_i(u) \geq \eta(x, u)\nabla f_i(u)$, $i = 1, 2, ..., p$. Como $\lambda_i \geq 0$ para todo i, temos $\lambda_i(f_i(x) - f_i(u)) \geq \lambda_i(\eta(x, u)\nabla f_i(u))$, $i = 1, 2, ..., p$. Somando em i obtemos

$$\sum_{i=1}^{p} \lambda_i(f_i(x) - f_i(u)) \geq \sum_{i=1}^{p} \lambda_i(\eta(x, u)\nabla f_i(u))$$

e portanto

$$\sum_{i=1}^{p} \lambda_i f_i(x) - \lambda_i f_i(u) \geq \nabla \left[\sum_{i=1}^{p} \lambda_i f_i(u)\right] \eta(x, u).$$

Logo, a função $\sum_{i=1}^{p} \lambda_i f_i(u)$ é invexa. □

3. Invexidade X Convexidade

Inicialmente introduziremos alguns conceitos e resultados de convexidade generalizada para funções diferenciáveis.

Suponha o aberto Ω convexo.

Definição 3.1. *f é convexa em $u \in \Omega$ se para todo $x \in \Omega$*

$$f(x) - f(u) \geq \nabla f(u)(x - u).$$

Se f for convexa em cada ponto de Ω, então f é dita convexa.

Definição 3.2. *f é quasi-convexa em $u \in \Omega$ se para todo $x \in \Omega$*

$$f(x) \leq f(u) \Rightarrow \nabla f(u)(x - u) \leq 0.$$

A função f é quasi-convexa em Ω se o for em cada ponto de Ω.

Definição 3.3. *f é pseudo-convexa em $u \in \Omega$ se para todo $x \in \Omega$*

$$\nabla f(u)(x-u) \geq 0 \Rightarrow f(x) \geq f(u).$$

Se f for pseudo-convexa em cada $u \in \Omega$, então f é pseudo-convexa em Ω.

Teorema 3.1. *Convexidade $\Rightarrow$ Pseudo-Convexidade $\Rightarrow$ Quasi-convexidade.*
As recíprocas não são verdadeiras.

Demonstração: Vide [7].

Teorema 3.2. *Pseudo-convexidade $\Rightarrow$ Invexidade.*
A recíproca é falsa. Além disso, existem funções invexas que não são quasi-convexas e funções quasi-convexas que não são invexas.

Demonstração: Seja f pseudo-convexa em $u \in \Omega$, então $\nabla f(u)(x-u) \geq 0 \Rightarrow f(x) - f(u) \geq 0$. Se f é invexa em $u \in \Omega$, pelo Teorema 2.1. temos que todo ponto estacionário é um mínimo global.

Tome $u \in \Omega$ um ponto estacionário, então $\nabla f(u) = 0$, e daí $f(x) - f(u) \geq 0$, para todo $u \in \Omega$, portanto u é um mínimo global. Assim Pseudo-convexidade implica Invexidade. $\square$

Podemos verificar que as recíprocas não são verdadeiras através dos exemplos abaixo:

Exemplo 2: Consideremos a função $f : \mathbb{R} \to \mathbb{R}$ dada por $f(x) = x^3$. Tome, $\nabla f(0)(-1-0) = 0 \geq 0$, mas $-1 = f(-1) < f(0) = 0$. Logo, f é quasi-convexa mas não é pseudo-convexa nem invexa, pois seu único ponto estacionário não é um mínimo global.

Exemplo 3: Consideremos a função $f : \mathbb{R} \to \mathbb{R}$ dada por $f(x) = x + x^3$. f não é convexa, no entanto é pseudo-convexa. De fato, se $\nabla f(u)(x-u) \geq 0$, então $x - u \geq 0$ desde que $\nabla f(u) = 1 + 3u^2$. Logo, $x \geq u$, o que implica $x^3 \geq u^3$. Podemos concluir então que $f(x) \geq f(u)$.

Exemplo 4: Consideremos a função $f : \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x, y) = x + x^3 - 10y^3 - y$. f não tem pontos estacionários, logo é invexa. Mas, tomando os pontos $(0,0)$ e $(2,1)$, obtemos $f(0,0) > f(2,1)$ mas $\nabla f(0,0)(2,1) > 0$. Assim sendo, f não é quasi-convexa e também não é pseudo-convexa.

Agora, apresentaremos alguns conceitos e resultados de invexidade generalizada para funções diferenciáveis. E, a seguir, será feito um paralelo entre os conceitos e resultados já apresentados de convexidade generalizada com os de invexidade generalizada.

Definição 3.4. *f é quasi-invexa em $u \in \Omega$ se existe $\eta : \Omega \times \Omega \to \mathbb{R}^n$ para todo $x \in \Omega$ tal que*

$$f(x) \leq f(u) \Rightarrow \nabla f(u)\eta(x, u) \leq 0.$$

A função f é quasi-invexa em Ω se o for em cada ponto de Ω.

Definição 3.5. *f é pseudo-invexa em $u \in \Omega$ se existe $\eta : \Omega \times \Omega \to \mathbb{R}^n$ para todo $x \in \Omega$ tal que*

$$\nabla f(u)\eta(x,u) \geq 0 \Rightarrow f(x) \geq f(u).$$

Se f for pseudo-invexa em cada $u \in \Omega$, então f é pseudo-invexa em Ω.

Teorema 3.3. *A classe das funções Pseudo-invexas coincide com a classe das funções invexas (ao contrário do que ocorre com as funções pseudo-convexas e convexas).*

Demonstração: Uma função f diferenciável é pseudo-invexa no aberto Ω se existir $\eta : \Omega \times \Omega \to \mathbb{R}^n$ tal que $f(x) - f(u) < 0 \Rightarrow \nabla f(u)\eta(x,u) < 0$.

Obviamente pseudo-invexidade implica invexidade (Teorema 3.2.). Por outro lado um ponto estacionário $u \in \Omega$ é também um mínimo global para uma função pseudo-invexa; de fato, se existir $x \in \Omega$ tal que $f(x) - f(u) < 0$, então temos $\nabla f(u)\eta(x,u) < 0$ e isso contradiz a hipótese de que $\nabla f(u) = 0$. Consequentemente, f é invexa. $\square$

Teorema 3.4. *Quasi-convexidade $\Rightarrow$ Quasi-invexidade.*

Invexidade $\Rightarrow$ Quasi-invexidade. As recíprocas não são verdadeiras.

Demonstração: f é quasi-convexa em $u \in \Omega$ se $f(x) - f(u) \leq 0 \Rightarrow \nabla f(u)(x-u) \leq 0$ para todo $x \in \Omega$. E é quasi-invexa em $u \in \Omega$ se existe $\eta : \Omega \times \Omega \to \mathbb{R}^n$ tal que $f(x) - f(u) \leq 0 \Rightarrow \nabla f(u)\eta(x,u) \leq 0$ para todo $x \in \Omega$. Claramente podemos concluir que se f é quasi-convexa, então ela é quasi-invexa quando $\eta(x,u) = x - u$.

Dizemos que f é invexa em $u \in \Omega$ se, para todo $x \in \Omega$, existe uma função $\eta : \Omega \times \Omega \to \mathbb{R}^n$ tal que $f(x) - f(u) \geq \nabla f(u)\eta(x-u) \leq 0$. Se f é quasi-invexa em $u \in \Omega$, temos que $f(x) - f(u) \leq 0 \Rightarrow \nabla f(u)\eta(x,u) \leq 0$, para todo $x \in \Omega$. Então $f(x) - f(u) \leq 0$, assim sendo $\eta(x,u)\nabla f(u) \leq 0$, portanto se f é invexa, então é quasi-invexa. $\square$

Exemplo 5: Consideremos a função $f : (0, \pi) \to \mathbb{R}$ dada por $f(x) = \sin^3(x)$. f é quasi-invexa com $\eta(x,u) = \cos(u)(\sin(x) - \sin(u))$. No entanto, f não é quasi-convexa, pois, para $x = \frac{3\pi}{4}$ e $u = \frac{\pi}{4}$, $f(x) = f(u)$, mas $\nabla f(u)(x-u) > 0$. E, o único ponto estacionário de f é $u = \frac{\pi}{2}$, o qual não é um ponto de mínimo global, logo f também não é invexa.

Em [5], Giorgi mostra que a classe de funções pseudo-convexas é a interseção das funções quasi-convexas com as invexas.

A partir deste resultado e de alguns resultados mostrados anteriormente é possível concluir que a maior classe de funções é a classe das funções quasi-invexas, dentro da qual temos as funções quasi-convexas e invexas. Como podemos visualizar no diagrama abaixo.

$$\begin{tikzcd}
\boxed{Convexas} \arrow[d, Rightarrow] & & \\
\boxed{Pseudo-Convexas} \arrow[r, Rightarrow] \arrow[d, Rightarrow] & \boxed{Invexas} \arrow[r, equal] \arrow[d, Rightarrow] & \boxed{Pseudo-invexas} \\
\boxed{Quasi-convexas} \arrow[r, Rightarrow] & \boxed{Quasi-invexas} &
\end{tikzcd}$$

4. Aplicações em Programação Matemática

Nesta seção mostraremos como os resultados de invexidade vistos nas seções anteriores podem ser aplicados em problemas de Programação Matemática.

Seja $\Omega \in \mathbb{R}^n$ não vazio e $f_0, f_i : \Omega \to \mathbb{R}, i = 1, ..., p$ funções diferenciáveis em Ω .

Considere o problema:

$$\begin{cases} Minimizar & f_0(x) \\ Sujeito \quad a & f_i(x) \quad \leq 0, \ i = 1, ..., p. \end{cases} \tag{4.1}$$

O conjunto dos pontos factíveis para (4.1), o qual suporemos não vazio, é dado por:

$$\mathbb{F} := \{x \in \Omega : f_i(x) \leq 0, i = 1, ..., p\}.$$

O objetivo aqui é encontrar candidatos ótimos à solução do problema (4.1).

Um resultado bastante conhecido e utilizado, é o Teorema de Karush-Kuhn-Tucker (KKT) que estabelece condições necessárias para (4.1) na forma de multiplicadores. Mas, usualmente impõe-se alguma regularidade sobre as funções de (4.1).

Teorema 4.1 (Condições de Karush-Kuhn-Tucker). *Seja $x^* \in \mathbb{F}$ uma solução ótima (local ou global) de (4.1), e $f_0, f_i, i = 1, ..., p$ obedecem alguma condição de regularidade. Então, existem escalares $\lambda_i, i = 1, ..., p$, tais que:*

$$\nabla f_0(x^*) + \sum_{i=1}^{p} \lambda_i \nabla f_i(x^*) \quad = \quad 0, \tag{4.2}$$

$$\lambda_i \quad \geq \quad 0, \ i = 1, ..., p, \tag{4.3}$$

$$\lambda_i f_i \quad = \quad 0, \ i = 1, ..., p. \tag{4.4}$$

Se todas as funções do problema (4.1) forem convexas, então as Condições de Karush-Kuhn-Tucker também são suficientes.

É possível encontrar na literatura várias condições que garantem a condição de regularidade exigida no teorema acima (por exemplo [1],[7]), das quais, as mais conhecidas são a Condição de Slater e a Condição de Mangasarian-Fromovitz.

No entanto, vêm sendo realizados vários estudos com o propósito de ampliar a classe de funções para as quais as Condições de Karush-Kuhn-Tucker sejam suficientes.

Após o surgimento da noção de função invexa, foi possível concluir que tal resultado pode ser utilizado para o caso em que as funções são invexas para uma mesma função η, como em [9].

Teorema 4.2. *Suponha que x^* seja um ponto factível para* (4.1) *e que as funções f_0, f_i sejam invexas em x^* para uma mesma função η e que existam multiplicadores $\lambda_i \in \mathbb{R}$, $i = 1, ..., p$, tal que as condições de Karush-Kuhn-Tucker sejam satisfeitas em x^*. Então x^* é um ponto de mínimo global para* (4.1).

Demonstração: De acordo com Teorema 2.2. $f_0 + \sum_{i=1}^{p} \lambda_i f_i$ é invexa em x^*, das Condições de Karush-Kuhn-Tucker segue que x^* é ponto estacionário de $f_0 + \sum_{i=1}^{p} \lambda_i f_i$. Então, pelo Teorema 2.1., x^* é ponto de mínimo global para $f_0 + \sum_{i=1}^{p} \lambda_i f_i$. Deste modo,

$$f_0(x) + \sum_{i=1}^{p} \lambda_i f_i(x) \geq f_0(x^*) + \sum_{i=1}^{p} \lambda_i f_i(x^*),$$

para todo ponto factível de (4.1).

Mas $\lambda_i f_i(x^*) = 0$ para todo i e, portanto

$$f_0(x) + \sum_{i=1}^{p} \lambda_i f_i(x) \geq f_0(x^*),$$

de onde obtemos que $f_0(x) \geq f_0(x^*)$, pois $\lambda_i f_i(x) \leq 0$ para todo $x \in \mathbb{F}$. Logo podemos concluir que x^* é solução global para (4.1). $\square$

Observação: Foi introduzida em [8] por Martin uma nova definição de invexidade, a qual foi dada o nome de *weak invexity*, em português chamada de *invexidade fraca*, como em [2].

Com esta nova definição foi possível caracterizar os problemas de Programação Matemática invexos nos moldes do Teorema 2.1..

Exemplo 6:

$$\begin{array}{lr} \text{Minimizar} & x - \sin(y) \\ \text{Sujeito a} & \sin(x) - 4\sin(y) \leq 0, \\ & 2\sin(x) + 7\sin(y) + x - 6 \leq 0, \\ & 2x + 2y - 3 \leq 0, \\ & 4x^2 + 4y^2 - 9 \leq 0, \\ & -\sin(x) \leq 0, \\ & -\sin(y) \leq 0, \\ & (x, y) \in \mathbb{R}^2. \end{array}$$

Podemos facilmente verificar a natureza não convexa da função objetivo e das restrições. No entanto, todas elas satisfazem as condições de invexidade com $\eta(z, u)$ dada por:

$$\eta(z, u) = \left(\frac{\sin(u_1) - \sin(x)}{\cos(x)}, \frac{\sin(u_2) - \sin(y)}{\cos(y)} \right), \text{ onde } \ z = (x, y) \quad \text{e} \quad u = (u_1, u_2).$$

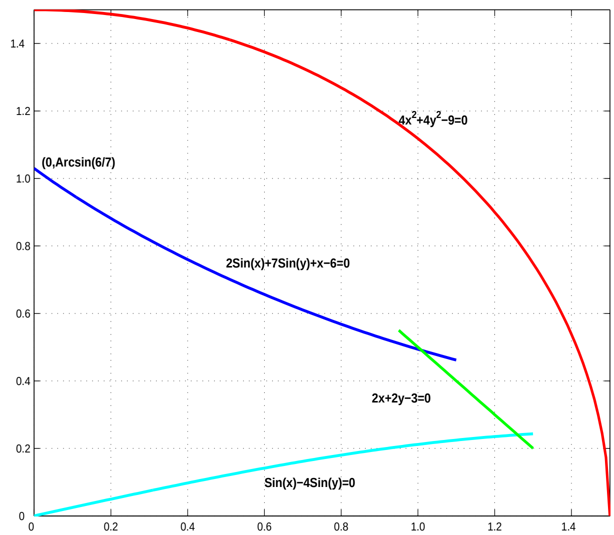

Figura 2: Gráfico da função do Exemplo 6

A função Lagrangeana $\mathcal{L}(z,\lambda) = f_0(z) + \sum_{i=1}^{p} \lambda_i f_i(x)$ é:

$$\begin{aligned} \mathcal{L}(z,\lambda) &= x - \sin(y) \\ &\quad + \lambda_1(\sin(x) - 4\sin(y)) \\ &\quad + \lambda_2(2\sin(x) + 7\sin(y) + x - 6) \\ &\quad + \lambda_3(2x + 2y - 3) \\ &\quad + \lambda_4(4x^2 + 4y^2 - 9) \\ &\quad + \lambda_5(-\sin(x)) \\ &\quad + \lambda_6(-\sin(y)), \end{aligned}$$

fica fácil verificar que as Condições de Karush-Kuhn-Tucker são satisfeitas para $\lambda = (0, \frac{1}{7}, 0, 0, \frac{10}{7}, 0)$ e $z_0 = \left(0, \arcsin\left(\frac{6}{7}\right)\right)$. Note que para utilizarmos o resultado do Teorema 4.2., não é necessário que conheçamos a função $\eta(z,u)$. É suficiente saber que ela existe.

Agradecimentos

Gostaríamos de agradecer a Sociedade Brasileira de Matemática Aplicada e Computacional por ter escolhido este trabalho para receber o Prêmio “Beatriz Neves”, ao Instituto de Matemática Estatística e Computação Científica - IMECC - UNICAMP pelo suporte e a FAPESP por ter financiado o projeto de Iniciação Científica.

Abstract. The concept of invex function was introduced by Hanson in 1980, this class of functions is bigger than the class of convex functions. After this, some studies have been made with the intention to use this new class of functions to guarantee optimality for problems of Mathematical Programming. The objective of this paper is to show that the Karush-Kuhn-Tucker Conditions guarantee global optimality if all the functions of the problem are, instead of convex, invex.

Referências

[1] R. Barbolla, E. Cerdá e P. Sanz, "Optimización. Cuestiones, Ejercicios y Aplicaciones a la Economia". Prentice Hall, Madrid, 2000.

[2] A.J. Brandão, M.A. Rojas-Medar e G.N. Silva, Uma introdução às funções invexas diferenciáveis com aplicações em otimização. *Boletim da Sociedade Paranaense de Matemática*, **19**, No. 1-2 (1999), 51-65.

[3] A. Ben-Israel e B. Mond, What´s Invexity? *J. Austral. Math. Soc. Ser. B*, **28** (1986), 1-9.

[4] B.D. Craven, Invex functions and constrained local minima. *Bull. Austral. Math. Soc.*, **24** (1981), 357-366.

[5] G. Giorgi, A note on the relashionships between convexity and invexity. *J. Austral. Math. Soc. Ser. B*, **32** (1990), 97-99.

[6] M.A. Hanson, On sufficiency of the Kuhn-Tucker conditions. *J. Math. Anal. Appl.*, **80** (1981), 545-550.

[7] O.L. Mangasarian, Nonlinear Programming. *Classics in Applied Mathematics*, SIAM, **10**, 1994.

[8] D.H. Martin, The essence of invexity. *J. Math. Anal. Appl.*, **47** (1985), 65-76.

[9] A.C. Moretti e M.A. Rojas-Medar, Condiciones suficientes de optimalidad em programación no lineal. *Cubo Matemática Educacional*, **3**, No. 2 (2001), 129-146.