MÉTODO TIPO-NEWTON PROXIMAL GLOBALIZADO PARA O PROBLEMA DE EQUILÍBRIO

Pedro Jorge Sousa dos Santos
Universidade Federal do Piauí
Campus Ministro Reis Velloso, Parnaíba - PI
pedrojorge@ufpi.edu.br

Paulo Sérgio Marques dos Santos
Universidade Federal do Piauí
Campus Ministro Reis Velloso, Parnaíba - PI
psergio@ufpi.edu.br

Susana Scheimberg
Universidade Federal do Rio de Janeiro
PESC/COPPE, Rio de Janeiro - RJ
susana@cos.ufrj.br

RESUMO

Neste trabalho apresentamos um algoritmo globalmente convergente para resolver o Problema de Equilíbrio o qual unifica problemas conhecidos, tais como: problemas de otimização, de desigualdade variacional e de equilíbrio de Nash. Este método é uma globalização do algoritmo do tipo-Newton não-suave localmente convergente apresentado em [Santos et al., 2016]. A principal ferramenta é o uso de uma função *gap* como função de mérito para a globalização. Alguns resultados numéricos são exibidos para ilustrar a performance do algoritmo.

PALAVRAS CHAVE. Problema de Equilíbrio, Convergência global, Método de Newton.

Tópico (Programação Matemática)

ABSTRACT

In this work we present a globally convergent algorithm for solving the Equilibrium Problem that unifies known problems such as: optimization problems, variational inequalities and Nash games. This method is a globalization of the locally convergent nonsmooth Newton-type algorithm presented in [Santos et al., 2016]. The main tool is the use of a gap function as a merit function for globalization. Some numerical results are displayed to illustrate the performance of the algorithm.

KEYWORDS. Equilibrium Problem, Global convergence, Newton's Method.

Topic (Mathematical Programming)

1. Introdução

Seja $K \subset \mathbb{R}^n$ convexo, fechado e não vazio e $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ uma bifunção de equilíbrio, i.e., $f(x,x) = 0$ para todo $x \in \mathbb{R}^n$. Consideraremos o seguinte problema de equilíbrio, denotado por EP(K, f):

$$\text{Encontrar} \quad x^* \in K \quad \text{tal que} \quad f(x^*, y) \geq 0 \quad \forall y \in K.$$

Este problema pode ser visto como uma formulação que unifica vários problemas importantes. De fato, ele inclui como caso particular, para uma apropriada escolha da bifunção f, problemas de otimização, problemas de equilíbrio de Nash, problemas do ponto fixo, desigualdades variacionais dentre outros; veja [Blum e Oettli, 1994].

Métodos do tipo-Newton para os problemas acima possuem propriedades de convergência locais atrativas, como por exemplo, taxa de convergência superlinear ou quadrática. No entanto, estes métodos possuem convergência apenas local no sentido de que é requerido pontos iniciais próximos a uma solução para se obter convergência do algoritmo. Portanto, para algoritmos do tipo-Newton serem práticos, é necessário que se faça uma 'globalização' do método no sentido de que o referido algoritmo aceite como ponto inicial um ponto viável arbitrário. Neste trabalho apresentaremos uma globalização do algoritmo do tipo-Newton dado em [Santos et al., 2016] para resolver EP(K, f) baseado na globalização apresentada em [Dreves et al., 2013] para equilíbrio de Nash.

Adotaremos a seguinte notação: O gradiente $\nabla\varphi(x)$ de uma função diferenciável $\varphi : \mathbb{R}^n \to \mathbb{R}$ será sempre um vetor coluna. Além disso, escreveremos $J\,\Phi(x)$ para o Jacobiano da função diferenciável $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ e $\nabla\Phi(x)$ para o transposto do Jacobiano. Para uma bifunção $\varphi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ a derivada parcial com respeito à primeira e à segunda variável será denotada por $\nabla_x\varphi(x,y) \in \mathbb{R}^n$ e $\nabla_y\varphi(x,y) \in \mathbb{R}^m$, respectivamente. A derivada parcial de segunda ordem com respeito primeiramente à segunda variável e depois com respeito à primeira variável será denotada por $\nabla^2_{yx}\varphi(x,y) \in \mathbb{R}^{m\times n}$, isto é, $(\nabla^2_{yx}\varphi)_{ij} = \frac{\partial^2\varphi}{\partial x_j \partial y_i}$.

Ao longo do trabalho assumiremos a seguinte hipótese padrão para métodos do tipo-Newton:

Hipótese 1.

(a) O conjunto $K \subset \mathbb{R}^n$ possui a seguinte representação

$$K = \{x \in \mathbb{R}^n \,;\, g(x) \leq 0\} \tag{1}$$

onde $g : \mathbb{R}^n \to \mathbb{R}^m$ é uma função duas vezes continuamente diferenciável e convexa.

(b) As bifunções $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ e $\nabla_y f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ são continuamente diferenciáveis e, para cada x fixo, a função $f(x,\cdot) : \mathbb{R}^n \to \mathbb{R}$ é convexa.

Nosso trabalho está organizado como segue: Na Seção 2 relembramos noções básicas do método do tipo-Newton local dado em [Santos et al., 2016] e apresentamos alguns resultados teóricos necessários. Na Seção 3 definimos e analisamos o nosso algoritmo. Alguns resultados numéricos são reportados na Seção 4. A última seção contém nossa conclusão.

2. Preliminares

Nesta seção apresentaremos definições e resultados que serão utilizados ao longo do texto. Iniciamos com a definição de função continuamente diferenciável por partes.

Definição 1. *[Facchinei e Pang, 2003a, Definição 4.5.1] Uma aplicação $\Psi : \mathbb{R}^n \to \mathbb{R}^n$ é continuamente diferenciável por partes (ou PC^1) numa vizinhança de um ponto $x \in \mathbb{R}^n$, se Ψ é contínua e existe uma vizinhança $V(x)$ de x e um número finito de aplicações continuamente diferenciáveis definidas em $V(x)$, $\{\Psi_1, \Psi_2, \ldots, \Psi_k\}$, para algum inteiro k, tal que $\Psi(y) \in \{\Psi_1(y), \ldots, \Psi_k(y)\}$ para todo $y \in V(x)$. Cada aplicação Ψ_i é chamada uma C^1-parte de Ψ em x.*

Considere a bifunção regularizada $\Phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ dada por

$$\Phi(x, y) := f(x, y) + h(x, y) \tag{2}$$

onde $h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ satisfaz a seguinte hipótese:

Hipótese 2.

(a) As bifunções $h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ e $\nabla_y h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ são continuamente diferenciáveis;

(b) $h(x, \cdot) : \mathbb{R}^n \to \mathbb{R}$ é fortemente convexa para todo $x \in \mathbb{R}^n$;

(c) $\nabla_y h(x, x) = 0$ para todo $x \in \mathbb{R}^n$.

Proposição 1. *Suponha válidas as hipóteses 1 e 2. Então, o problema de otimização*

$$\min_y \Phi(x, y) \;\; \textit{sujeito a} \;\; y \in K \tag{3}$$

possui uma única solução para cada $x \in \mathbb{R}^n$, a qual denotaremos por $y(x)$.

Demonstração. Veja, por exemplo, [Izmailov e Solodov, 2009, Corolário 3.4.2]. □

A próxima proposição fornece algumas propriedades importantes da aplicação y. Para mais detalhes veja [Santos et al., 2016]. Observamos que não é necessário assumir condições de qualificação sobre o conjunto K para ganharmos o resultado abaixo.

Proposição 2. *Suponha válidas as hipóteses 1 e 2. Então,*

(a) a aplicação $y : \mathbb{R}^n \to \mathbb{R}^n$ é contínua;

(b) um ponto $x \in \mathbb{R}^n$ resolve EP(K, f) se, e somente se, $y(x) = x$.

É importante mencionar que embora a aplicação y seja contínua ela nem sempre é diferenciável, como poder ser observado no seguinte exemplo:

Exemplo 1. *Considere EP(K, f) com $n = 1$, a bifunção f dada por $f(x, y) = x(y - x)$ e o conjunto $K = [0, 1]$. Escolhendo $h(x, y) = (y - x)^2$ temos que a aplicação y é dada por*

$$y(x) = \begin{cases} 0 & , \textit{se } x \leq 0 \\ \frac{1}{2}x & , \textit{se } 0 < x < 2 \\ 1 & , \textit{se } x \geq 2 \end{cases}$$

que é contínua mas não diferenciável em toda a reta.

Com o objetivo de desenvolver um método local do tipo-Newton para EP(K, f) [Santos et al., 2016] consideram a aplicação $F : \mathbb{R}^n \to \mathbb{R}^n$ dada por

$$F(x) := y(x) - x$$

onde, de acordo com a Proposição 2, os zeros de F caracterizam as soluções do problema de equilíbrio. Embora a aplicação y não seja diferenciável é possível mostrar, sob hipóteses usuais, que ela é continuamente diferenciável por partes. De fato, denotando $I_m = \{1, \ldots, m\}$, seja $I_0(x) := \{i \in I_m \,;\, g_i(y(x)) = 0\}$ o conjunto de índices ativos em $y(x)$ e considere a condição de qualificação das restrições posto constante abaixo:

Hipótese 3. *Fixado $x \in \mathbb{R}^n$, a condição de qualificação posto constante (CRCQ) vale em $y(x)$, isto é, existe uma vizinhança $N(y(x))$ de $y(x)$ tal que para cada $J \subseteq I_0(x)$, o conjunto*

$$\{\nabla g_i(y) \,;\, i \in J\}$$

tem o mesmo posto (que depende de J) para todo $y \in N(y(x))$.

Considere o conjunto dos multiplicadores de Lagrange associados ao problema (3),

$$M(x) = \{\lambda \in \mathbb{R}^m\,;\ (y(x), \lambda) \text{ satisfaz o sistema KKT associado a } (3)\},$$

e a seguinte família de subconjuntos do conjunto de índices ativos $I_0(x)$ dada por

$$B(x) = \{\mathcal{J} \subseteq I_0(x)\,;\ (\nabla g_i(y(x)))_{i\in J} \text{ é linearmente independente e}$$

$$supp(\lambda) \subseteq \mathcal{J} \text{ para algum } \lambda \in M(x)\}$$

onde, dado $\lambda \in \mathbb{R}^m$, $supp(\lambda) := \{i \in I_m\,;\ \lambda_i > 0\}$. Sob a Hipótese 3, $M(x)$ e $B(x)$ são sempre não vazios. Para mais detalhes veja [Janin, 1984; von Heusinger et al., 2012].

Proposição 3. *[Santos et al., 2016] Seja $x \in \mathbb{R}^n$ dado e assuma as Hipóteses 1 e 2. Suponha válida a Hipótese 3 em $y(x)$. Então a função $F(x) = y(x) - x$ é continuamente diferenciável por partes numa vizinhança de x, e o Jacobiano generalizado computável de F em x é dado por*

$$\partial_C F(x) = \{J\,y^{\mathcal{J}}(x) - I\,;\ \mathcal{J} \in B(x)\}, \tag{4}$$

onde

$$J\,y^{\mathcal{J}}(x) = C^{-1}A - C^{-1}D(D^T C^{-1} D)^{-1} D^T C^{-1} A, \tag{5}$$

com

$$\begin{aligned} A &= A(x) = -\nabla^2_{yx} f(x, y(x)) - \nabla^2_{yx} h(x, y(x)), \\ C &= C^{\mathcal{J}}(x) = \nabla^2_{yy} f(x, y(x)) + \nabla^2_{yy} h(x, y(x)) + \sum_{i\in\mathcal{J}} \lambda_i^{\mathcal{J}}(x) \nabla^2 g_i(y(x)), \\ D &= D^{\mathcal{J}}(x) = \nabla g_{\mathcal{J}}(y(x)), \end{aligned} \tag{6}$$

onde $\nabla g_{\mathcal{J}}(\cdot)$ é a notação para a matriz com vetores colunas $\{\nabla g_i(\cdot)\}_{i\in\mathcal{J}}$ e $I \in \mathbb{R}^{n\times n}$ denota a matriz identidade.

O método do tipo-Newton local utiliza então o Jacobiano generalizado computável de F para resolver a equação $F(x) = 0$. Para mais detalhes sobre o Jacobiano generalizado computável, veja [Sun et al., 1997; von Heusinger et al., 2012].

Considere agora a função

$$\varphi(x) := -\Phi(x, y(x)) = -\min_{y\in K} \Phi(x, y). \tag{7}$$

A função φ acima terá um papel crucial no processo de globalização do algoritmo. É interessante mencionar que embora a aplicação y nem sempre seja diferenciável, o mesmo não ocorre com a função φ. Para a prova da próxima proposição veja por exemplo [Mastroeni, 2003] ou [Facchinei e Pang, 2003b, Teorema 10.2.1].

Proposição 4. *Suponha válidas as hipóteses 1 e 2. Então, a função φ é continuamente diferenciável e seu gradiente é dado por*

$$\nabla\varphi(x) = -\nabla_x \Phi(x, y(x)) = -\nabla_x f(x, y(x)) - \nabla_x h(x, y(x)).$$

Considere agora duas bifunções distintas satisfazendo a Hipótese 2, digamos h_A e h_B, e considere as respectivas aplicações Φ_A e Φ_B, y_A e y_B, F_A e F_B, φ_A e φ_B definidas de modo análogo como o exposto acima utilizando-se h_A e h_B em (2). Considere também a função

$$\varphi_{AB}(x) := \varphi_A(x) - \varphi_B(x). \tag{8}$$

Temos o seguinte lema.

Lema 1. *A desigualdade*

$$h_B(x, y_B(x)) - h_A(x, y_B(x)) \leq \varphi_{AB}(x) \leq h_B(x, y_A(x)) - h_A(x, y_A(x))$$

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

Demonstração. De fato, segue da definição (7) que

$$\varphi_B(x) = -\Phi_B(x, y_B(x)) = -\min_{y \in K} \Phi_B(x, y) \geq -\Phi_B(x, y_A(x)).$$

Logo,

$$\varphi_{AB}(x) = \varphi_A(x) - \varphi_B(x) \leq -\Phi_A(x, y_A(x)) + \Phi_B(x, y_A(x)) = h_B(x, y_A(x)) - h_A(x, y_A(x)).$$

A desigualdade restante prova-se de modo análogo. □

Corolário 1. *Seja A e B matrizes quadradas de ordem n tal que A e $B - A$ são simétricas positivas definidas. Considere $h_A(x, y) := ||y - x||_A^2$ (e h_B definido de modo análogo). Neste caso, temos*

$$||y_B(x) - x||_{B-A}^2 \leq \varphi_{AB}(x) \leq ||y_A(x) - x||_{B-A}^2,$$

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

Proposição 5. *Suponha válido a Hipótese 1. Além disso, suponha A, B e h_A e h_B satisfazendo as hipóteses do Corolário 1. Então,*

(a) $\varphi_{AB}(x) \geq 0$ *para todo* $x \in \mathbb{R}^n$;

(b) x^* *resolve EP(K, f) se, e somente se,* $\varphi_{AB}(x^*) = 0$;

(c) φ_{AB} *é continuamente diferenciável e seu gradiente é dado por*

$$\nabla \varphi_{AB}(x) = \nabla_x \Phi_B(x, y_B(x)) - \nabla_x \Phi_A(x, y_A(x)).$$

Demonstração. (a) Segue diretamente da desigualdade à esquerda no Corolário 1.
(b) Segue da segunda parte da Proposição 2 e do Corolário 1.
(c) Segue diretamente da Proposição 4 e de (8). □

3. O Algoritmo

Utilizando os resultados e notações da seção anterior, estamos agora prontos para estabelecer o nosso algoritmo.

Algoritmo 1. *(Método de Newton globalizado para EP)*

(P.0) *Escolha* $x^0 \in \mathbb{R}^n$, $s > 1$, $\rho > 0$, $\tau \in (0, 1)$, $\sigma \in (0, 1)$ *e faça* $k := 0$.

(P.1) *Se* $||F_B(x^k)|| = ||y_B(x^k) - x^k|| = 0$, *pare;*

(P.2) *Calcule* $H_k \in \partial_C F_B(x^k)$.

(a) Encontre uma solução $d^k \in \mathbb{R}^n$ *do sistema linear*

$$H_k d^k = -F_B(x^k), \tag{9}$$

caso exista. Se não existe, ponha $d^k := -\nabla \varphi_{AB}(x^k)$ *e vá para o Passo (P.3), caso contrário, vá para (b).*

(b) Se

$$\varphi_{AB}(x^k + d^k) \leq \tau \varphi_{AB}(x^k), \tag{10}$$

ponha $x^{k+1} := x^k + d^k$, $k = k+1$ *e volte ao passo (P.1). Caso contrário, vá para (c).*

(c) Se d^k *não satisfaz a condição*

$$\langle \nabla \varphi_{AB}(x^k), d^k \rangle \leq -\rho ||d^k||^s, \tag{11}$$

ponha $d^k := -\nabla \varphi_{AB}(x^k)$. *Vá para o Passo (P.3).*

(P.3) *Calcule* $t_k := \max\{2^{-l};\, l = 0, 1, 2, \ldots\}$ *tal que*

$$\varphi_{AB}(x^k + t_k d^k) \leq \varphi_{AB}(x^k) + \sigma t_k \langle \nabla \varphi_{AB}(x^k), d^k \rangle. \tag{12}$$

Ponha $x^{k+1} = x^k + t_k d^k$, $k = k+1$ *e volte ao passo (P.1).*

Veremos na demonstração do Teorema 1 que o Algoritmo 1 encontra uma solução de EP(K, f) ou um ponto estacionário da função φ_{AB}. Para garantir que o ponto estacionário seja solução, devemos acrescentar alguma hipótese sobre os dados do problema e sobre as matrizes A e B. Por exemplo, obtemos esse resultado se considerarmos a seguinte hipótese:

Hipótese 4. *Para todo* $x \in \mathbb{R}^n$ *tal que* $y_A(x) \neq y_B(x)$*, a desigualdade*

$$\langle (\nabla_x f(x, y_B(x)) + \nabla_y f(x, y_B(x))) - (\nabla_x f(x, y_A(x)) + \nabla_y f(x, y_A(x))), y_B(x) - y_A(x) \rangle > 0$$

vale.

Lema 2. *Suponha válido a Hipótese 1 e sejam* A*,* B *e* h_A*,* h_B *satisfazendo as hipóteses do Corolário 1. Além disso, assuma válido a Hipótese 4. Então todo ponto estacionário da função* φ_{AB} *é solução do problema de equilíbrio EP*(K, f)*.*

Demonstração. Seja x um ponto estacionário da da função φ_{AB}. Da Proposição 5 temos que

$$\begin{aligned}
0 = \nabla \varphi_{AB}(x) &= \nabla_x \Phi_B(x, y_B(x)) - \nabla_x \Phi_A(x, y_A(x)) \\
&= \nabla_x f(x, y_B(x)) - \nabla_x f(x, y_A(x)) + \nabla_x h_B(x, y_B(x)) - \nabla_x h_A(x, y_A(x)) \\
&= \nabla_x f(x, y_B(x)) - \nabla_x f(x, y_A(x)) - \nabla_y h_B(x, y_B(x)) + \nabla_y h_A(x, y_A(x)) \\
&= \nabla_x f(x, y_B(x)) + \nabla_y f(x, y_B(x)) - \nabla_x f(x, y_A(x)) - \nabla_y f(x, y_A(x)) \\
&\quad -\nabla_y f(x, y_B(x)) + \nabla_y f(x, y_A(x)) - \nabla_y h_B(x, y_B(x)) + \nabla_y h_A(x, y_A(x)) \\
&= \nabla_x f(x, y_B(x)) + \nabla_y f(x, y_B(x)) - \nabla_x f(x, y_A(x)) - \nabla_y f(x, y_A(x)) \\
&\quad +\nabla_y \Phi_A(x, y_A(x)) - \nabla_y \Phi_B(x, y_B(x)).
\end{aligned} \tag{13}$$

Por outro lado, da definição de $y_A(x)$, temos que $y_A(x)$ satisfaz a condição de otimalidade

$$\langle \nabla_y \Phi_A(x, y_A(x)), z - y_A(x) \rangle \geq 0 \ \ \forall z \in K. \tag{14}$$

De modo análogo,

$$\langle \nabla_y \Phi_B(x, y_B(x)), z - y_B(x) \rangle \geq 0 \ \ \forall z \in K. \tag{15}$$

Tomando $z = y_B(x)$ em (14), $z = y_A(x)$ em (15) e somando as duas desigualdades obtemos

$$\langle \nabla_y \Phi_B(x, y_B(x)) - \nabla_y \Phi_A(x, y_A(x)), y_B(x) - y_A(x) \rangle \leq 0.$$

Tomando o produto interno em (13) com respeito a $y_B(x) - y_A(x)$ e considerando a equação acima, obtemos

$$\langle (\nabla_x f(x, y_B(x)) + \nabla_y f(x, y_B(x))) - (\nabla_x f(x, y_A(x)) + \nabla_y f(x, y_A(x))), y_B(x) - y_A(x) \rangle \leq 0.$$

Logo, por hipótese, $y_A(x) = y_B(x)$. Portanto, substituindo essa igualdade em (13), encontramos

$$\begin{aligned} 0 &= \nabla_y h_B(x, y_A(x)) - \nabla_y h_A(x, y_A(x)) \\ &= 2(B - A)(y_A(x) - x). \end{aligned}$$

O que implica que $y_A(x) = x$ uma vez que $B - A$ é uma matriz simétrica positiva definida. O resultado segue da Proposição 2. $\square$

Teorema 1. *Assuma a Hipótese 1 e suponha válida a Hipótese 3 para todo $x \in K$. Além disso, sejam A, B e h_A, h_B satisfazendo as hipóteses do Corolário 1 e assuma a Hipótese 4. Então, ou o algoritmo pára em uma solução do problema de equilíbrio, ou todo ponto de acumulação x^* da sequência gerada pelo algoritmo é uma solução de EP(K, f).*

Demonstração. A prova segue os mesmos passos de [Dreves et al., 2013, Teorema 3.2]. Nós a exibimos aqui por questão de completude.
Se o algoritmo pára no passo **(P.1)** então $\{x^k\}$ é uma solução de EP(K, f) por conta da Proposição 2(b). Caso contrário, considere uma subsequência $\{x^k\}$ convergindo para x^*.
Se, para um conjunto infinito de índices na sequência $\{x^k\}$, tivermos $d^k = -\nabla\varphi_{AB}(x^k)$ então, por argumentos padrão, x^* é um ponto estacionário de φ_{AB} e portanto, pelo Lema 2, x^* é uma solução de EP(K, f).
Suponha então que d^k é sempre obtido resolvendo-se a equação (9). Se (10) vale para uma quantidade infinita de índices ganhamos que $\varphi_{AB}(x^*) = 0$ uma vez que φ_{AB} é contínua, $\varphi_{AB}(x) \geq 0$ para todo $x \in \mathbb{R}^n$ e $\tau \in (0, 1)$. Logo, novamente pelo Lema 2, x^* é uma solução de EP(K, f).
Suponha então que, além de d^k ser obtido resolvendo-se a equação (9), a condição (11) é sempre satisfeita. De (11) temos que $||d^k||^{s-1} \leq -\frac{1}{\rho}\langle\nabla\varphi_{AB}(x^k), \frac{d^k}{||d^k||}\rangle$ o que implica que a sequência $\{d^k\}$ é limitada uma vez que $x^k \to x^*$, $\nabla\varphi_{AB}$ é contínua e $s > 1$. Sem perda de generalidade, suponha que $d^k \to d^*$.
Suponha $d^* \neq 0$. De (11) e (12), temos que $\varphi_{AB}(x^k + t_k d^k) - \varphi_{AB}(x^k) \leq \sigma t_k \langle\nabla\varphi_{AB}(x^k), d^k\rangle < 0$, e como $\varphi_{AB}(x^k + t_k d^k) - \varphi_{AB}(x^k) \to 0$, ganhamos que

$$t_k\langle\nabla\varphi_{AB}(x^k), d^k\rangle \to 0. \tag{16}$$

Suponha que $t_k \to 0$. Novamente de (12), temos

$$\frac{\varphi_{AB}(x^k + 2t_k d^k) - \varphi_{AB}(x^k)}{2t_k} > \sigma\langle\nabla\varphi_{AB}(x^k), d^k\rangle.$$

Passando o limite e levando em conta que φ_{AB} é C^1, temos $\langle\nabla\varphi_{AB}(x^*), d^*\rangle \geq \sigma\langle\nabla\varphi_{AB}(x^*), d^*\rangle$ o que nos dá $\langle\nabla\varphi_{AB}(x^*), d^*\rangle \geq 0$ ($\sigma \in (0, 1)$) contradizendo (11) uma vez que $d^* \neq 0$. Logo, existe $c > 0$ tal que $t_k > c$ para todo k. Contudo, (11) e (16) implicam que $d^k \to 0$ o que é uma contradição com $d^* \neq 0$. Portanto, $d^* = 0$.
Da Proposição 3, H_k é o Jacobiano de uma das C^1-partes de F_B o que implica que $\{H_k\}$ é limitada. Já que $||F_B(x^k)|| = ||H_k d^k|| \leq ||H_k|| \cdot ||d^k||$ e $d^k \to 0$ segue que $F_B(x^*) = 0$ e, pela Proposição 2(b), x^* é uma solução de EP(K, f). $\square$

Para analisar a convergência local do Algoritmo 1 assumiremos a não singularidade dos elementos do Jacobiano generalizado computável de F (que pode ser alcançada assumindo a Hipótese 5 abaixo; veja [Santos et al., 2016, Lema 2.2]).

Hipótese 5. *Para cada $\mathcal{J} \in B(x)$ e $\lambda \in M(x)$, temos*

$$d^T\left(M(x, y(x)) + \sum_{i\in\mathcal{J}} \lambda_i \nabla^2 g_i(y(x))\right) d \neq 0 \quad \forall\, d \in \mathcal{T}^{\mathcal{J}}(x),\ d \neq 0$$

onde $M(x, y) := \nabla^2_{yy}\Phi(x, y) + \nabla^2_{yx}\Phi(x, y)$ e $\mathcal{T}^{\mathcal{J}}(x) := \{d \in \mathbb{R}^n\,;\ \nabla g_i(y(x))^T d = 0 \forall i \in \mathcal{J}\}$.

O próximo lema é uma consequência direta de [Santos et al., 2016, Lemas 2.3 e 3.1]. Veja também [von Heusinger et al., 2012; Dreves et al., 2013].

Lema 3. *Seja x^* uma solução de EP(K, f). Assuma a Hipótese 1 e suponha válidas as Hipóteses 3 e 5 em x^*. Considere $\{x^k\}$ uma sequência arbitrária convergindo para x^*. Se d^k é uma solução de $H_k d^k = -F(x^k)$, então temos que*

$$||x^k + d^k - x^*|| = o(||x^k - x^*||).$$

Além disso, se o Jacobiano de todas as ∇g_i e de $\nabla_y \Phi$ são localmente Lipschitz contínuas, então

$$||x^k + d^k - x^*|| = O(||x^k - x^*||).$$

Teorema 2. *Assuma a Hipótese 1 e suponha válidas as Hipóteses do Corolário 1. Seja x^* uma solução de EP(K, f) e suponha que x^* é um ponto de acumulação da sequência $\{x^k\}$ gerada pelo Algoritmo 1. Se as Hipóteses 3 e 5 valem em x^*, então toda a sequência x^k converge para x^*. Além disso, existe k_0 tal que para $k \geq k_0$ o sistema (9) é sempre solúvel, a desigualdade (10) é satisfeita e $\{x^k\}$ converge superlinearmente para x^*. Além disso, se o Jacobiano de todas as ∇g_i e de $\nabla_y \Phi$ são localmente Lipschitz contínuas, a taxa de convergência é quadrática.*

Demonstração. Novamente, a prova segue os mesmos passos de [Dreves et al., 2013, Teorema 3.6]. Nós a exibimos aqui por questão de completude.
De modo análogo ao de [Dreves et al., 2013, Teorema 3.6] é possível provar que, sob as hipóteses acima, a sequência $\{x^k\}$ converge para x^*. Além disso, prova-se que existe $c > 0$ tal que

$$\frac{c}{2}||x - x^*|| \leq ||F_B(x)|| \tag{17}$$

para todo x suficientemente próximo a x^*.
Quanto à segunda parte do teorema temos que o sistema (9) é sempre solúvel para k suficientemente grande devido a [Santos et al., 2016, Proposição 3.2]. Provemos agora que a desigualdade (10) é sempre satisfeita para k grande o suficiente. Seja $L > 0$ a constante de Lipschitz local de F_A em torno de x^* (a qual existe uma vez que F_A é PC^1, logo localmente Lipschitz contínua). Do Corolário 1 e do Lema 3 temos

$$\begin{array}{rcl} \sqrt{\varphi_{AB}(x^k + d^k)} & \overset{\text{Corolário 1}}{\leq} & ||F_A(x^k + d^k)||_{B-A} \\ & = & ||F_A(x^k + d^k) - F_A(x^*)||_{B-A} \\ & \leq & \sqrt{||B-A||} \cdot ||F_A(x^k + d^k) - F_A(x^*)|| \\ & \leq & \sqrt{||B-A||} \cdot L \cdot ||x^k + d^k - x^*|| \\ & \overset{\text{Lema 3}}{=} & o(||x^k - x^*||). \end{array}$$

Portanto, para k grande o suficiente, temos

$$\begin{array}{rcl} \sqrt{\varphi_{AB}(x^k + d^k)} & \leq & \sqrt{\tau}\eta\frac{c}{2}||x^k - x^*|| \\ & \overset{(17)}{\leq} & \sqrt{\tau}\eta||F_B(x^k)|| \\ & \leq & \sqrt{\tau}||F_B(x^k)||_{B-A} \\ & \overset{\text{Corolário 1}}{\leq} & \sqrt{\tau}\sqrt{\varphi_{AB}(x^k)} \end{array}$$

onde η é uma constante da equivalência entre as normas $||\cdot||$ e $||\cdot||_{B-A}$. Assim, (10) é satisfeito. Deste modo, para k suficientemente grande, o método globalizado coincide com o método local e portanto herda as propriedades de taxa de convergência do método. □

4. Resultados numéricos

Nesta seção avaliamos o desempenho de nosso algoritmo apresentando alguns resultados numéricos. Nós utilizamos o MATLAB® R2017a em um MacBook Intel Core i7 com 8 GB RAM para obter nossos resultados. As matrizes escolhidas para a definição de h_A e h_B foram $A = \frac{10^{-2}}{2} \cdot I$ e $B = \frac{1}{2} \cdot I$, onde I é a matriz identidade de ordem n. O critério de parada adotado foi $||y_B(x^k) - x^k|| < \varepsilon$. Em todos os exemplos faremos $\varepsilon = 10^{10}$ o que significa que nós requeremos uma precisão muito alta para a solução. Como parâmetros para o algoritmo utilizamos $s = 2,1$, $\rho = 10^{-8}$, $\tau = 0,5$ e $\sigma = 10^{-2}$.

Analisamos os mesmos problemas do método local em [Santos et al., 2016]. O primeiro exemplo consiste de um problema de equilíbrio de Nash generalizado que se encontra em [von Heusinger et al., 2012], os outros dois são problemas de equilíbrio quadráticos que se encontram em [Santos e Scheimberg, 2011] e [Nasri et al., 2016].

Para um resumo dos resultados veja a Tabela 1. Cada ponto inicial utilizado possui todas as coordenadas iguais.

Exemplo 2. *Considere o problema de equilíbrio de Nash com dois jogadores dados em [von Heusinger et al., 2012] e sua reformulação como um problema de equilíbrio EP(K, f).*
A bifunção f é dada por

$$f(x, y) = \sum_{v=1}^{2} [\theta_v(y^v, x^{-v}) - \theta_v(x^v, x^{-v})],$$

com

$$\theta_1(x_1, x_2) = \frac{1}{2}x_1^2 - x_1x_2 \ \ e \ \ \theta_2(x_1, x_2) = x_2^2 + x_1x_2,$$

e o conjunto comum de restrições é dado por $K = \{x \in \mathbb{R}^2 \,;\, x_1 \geq 1,\, x_2 \geq 0,\, x_1 + x_2 \geq 1\}$.

Exemplo 3. *Considere os dois problemas de equilíbrio que se encontram em [Santos e Scheimberg, 2011] e [Nasri et al., 2016], onde $K = \{x \in \mathbb{R}^5 \,;\, \sum_{i=1}^5 x_i \geq -1,\, -5 \leq x_i \leq 5,\, i = 1, \ldots, 5\}$, e a bifunção f é da forma*

$$f(x, y) = \langle Px + Qy + q,\, y - x\rangle.$$

As matrizes P, Q e o vetor q são dados por

$$P_1 = \begin{bmatrix} 3,1 & 2 & 0 & 0 & 0 \\ 2 & 3,6 & 0 & 0 & 0 \\ 0 & 0 & 3,5 & 2 & 0 \\ 0 & 0 & 2 & 3,3 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}, \ P_2 = \begin{bmatrix} 3,1 & 2 & 0 & 0 & 0 \\ 2 & 3,6 & 0 & 0 & 0 \\ 0 & 0 & 3,5 & 2 & 0 \\ 0 & 0 & 2 & 3,3 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix},$$

$$Q = \begin{bmatrix} 1,6 & 1 & 0 & 0 & 0 \\ 1 & 1,6 & 0 & 0 & 0 \\ 0 & 0 & 1,5 & 1 & 0 \\ 0 & 0 & 1 & 1,5 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix} \ \ e \ \ q = \begin{bmatrix} 1 \\ -2 \\ -1 \\ 2 \\ -1 \end{bmatrix},$$

onde o i-ésimo problema considera $P = P_i$, $i = 1, 2$.

Tabela 1: Sumário dos resultados numéricos.

Exemplo	Ponto inicial	Número de iterações
Exemplo 2	10	11
Exemplo 2	100	11
Exemplo 2	1000	11
Exemplo 3 $(P = P_1)$	10	2
Exemplo 3 $(P = P_1)$	100	2
Exemplo 3 $(P = P_1)$	1000	2
Exemplo 3 $(P = P_2)$	10	2
Exemplo 3 $(P = P_2)$	100	2
Exemplo 3 $(P = P_2)$	1000	2

5. Observações finais

Neste trabalho apresentamos um método do tipo-Newton proximal globalizado para o problema de equilíbrio cujo conjunto viável é definido por desigualdades convexas. Nosso algoritmo foi baseado no método do tipo-Newton local de [Santos et al., 2016] e na globalização apresentada em [Dreves et al., 2013] para encontrar um equilíbrio normalizado do problema de equilíbrio de Nash generalizado. A principal ferramenta foi o estabelecimento de uma função do tipo *gap* como função de mérito que se mostrou útil para a globalização do método local apresentado em [Santos et al., 2016]. Os exemplos numéricos mostram que o método é capaz de encontrar uma solução do problema independente do ponto inicial escolhido.

Referências

Blum, E. e Oettli, W. (1994). From optimization and variational inequalities to equilibrium problems. *The Mathematics Student*, 63(1–4):123–145.

Dreves, A., von Heusinger, A., Kanzow, C., e Fukushima, M. (2013). A globalized newton method for the computation of normalized nash equilibria. *Journal of Global Optimization*, 56(2):327–340. ISSN 1573-2916. URL `http://dx.doi.org/10.1007/s10898-011-9824-9`.

Facchinei, F. e Pang, J. S. (2003a). *Finite-Dimensional Variational Inequalities and Complementary Problems*, volume 1. Springer-Verlag, New York.

Facchinei, F. e Pang, J. S. (2003b). *Finite-Dimensional Variational Inequalities and Complementary Problems*, volume 2 of *Springer series in operations research*. Springer.

Izmailov, A. e Solodov, M. (2009). *Otimização, volume 1: Condições de Otimalidade, Elementos de Análise Convexa e de Dualidade*. IMPA, 2 edition.

Janin, R. (1984). Directional derivative of the marginal function in nonlinear programming. *Mathematical Programming Study*, 21:110–126.

Mastroeni, G. (2003). Gap functions for equilibrium problems. *Journal of Global Optimization*, 27:411–426.

Nasri, M., Matioli, L. C., Ferreira, E. M. S., e Silveira, A. (2016). Implementation of augmented lagrangian methods for equilibrium problems. *Optimization Theory and Applications*, 168(3): 971–991.

Santos, P. J. S., Scheimberg, S., e Santos, P. S. M. (2016). Método tipo-newton proximal para o problema de equilíbrio. *Anais do XLVIII SBPO*, p. 2654–2665.

Santos, P. S. M. e Scheimberg, S. (2011). An inexact subgradient algorithm for equilibrium problems. *Computational & Applied Mathematics*, 30(1):91–107.

Sun, D., Fukushima, M., e Qi, L. (1997). *Complementary and Variational Problems: State of the Art*, chapter A computable generalized Hessian of the D-gap function and Newton-type methods for variational inequality problems, p. 452–472. SIAM, Philadelphia.

von Heusinger, A., Kanzow, C., e Fukushima, M. (2012). Newton's method for computing a normalized equilibrium in the generalized nash game through fixed point formulation. *Mathematical Programming*, 132:99–123.