On Solving Generalized Nash Equilibrium Problems via Optimization*

Barbara Panicucci† Massimo Pappalardo† Mauro Passacantando†

Abstract. This paper deals with the generalized Nash equilibrium problem (GNEP), i.e. a noncooperative game in which the strategy set of each player, as well as his payoff function, depends on the strategies of all players. We consider an equivalent optimization reformulation of GNEP using a regularized Nikaido-Isoda function so that solutions of GNEP coincide with global minima of the optimization problem. We then propose a derivative-free descent type method with inexact line search to solve the equivalent optimization problem and we prove that our algorithm is globally convergent. The convergence analysis is not based on conditions guaranteeing that every stationary point of the optimization problem is a solution of GNEP. Finally, we present the performance of our algorithm on some examples.

Keywords. Generalized Nash equilibrium problem, Nikaido-Isoda function, descent method.

AMS Subject classification. 90C30 - 91A10 - 91B50.

1 Introduction

In this paper we consider the generalized Nash equilibrium problem (GNEP, for short). The GNEP problem is a noncooperative game in which, in contrast to the standard Nash equilibrium problem, the strategy set of each player depends on the strategies of all the other players as well as on his own strategy. Recently, GNEP has emerged as an effective and powerful tool for modelling a wide class of problems arising in many fields (electricity, telecommunications, transportation and others). For a detailed overview of GNEP we refer the reader to [5] and references therein.

Let us first recall the definition of the GNEP. There are N players, each labelled by an integer $i = 1, \dots, N$. Player i has a real valued payoff function $\theta_i(x)$ that depends on all players strategies $x = (x_1, \dots, x_N)$, where each component $x_i \in \mathbb{R}^{n_i}$ represents the strategy of the i-th player. The vector $x \in \mathbb{R}^n$, where $n = \sum_{i=1}^{N} n_i$, is also denoted by $x = (x_i, x_{-i})$, where x_{-i} denotes the strategy vector of all the players different from player i. Throughout this paper we assume that for all $i = 1, \dots, N$ the function θ_i is continuously differentiable on $\mathbb{R}^n$ and that $\theta_i(\cdot, x_{-i})$ is convex for all x_{-i}. We further assume

*This work has been supported by the National Research Program PRIN 20079PLLN7 "Nonlinear Optimization, Variational Inequalities, and Equilibrium Problems". This paper has been published in Optimization Letters, vol. 3 (2009), pp. 419-435.

†Department of Applied Mathematics - University of Pisa, via Buonarroti 1/c, 56127 Pisa - Italy (panicucc@mail.dm.unipi.it), (pappalardo@dma.unipi.it), (passacantando@dma.unipi.it).

that the strategy x_i of the i-th player belongs to a set $X_i\,(x_{-i})$ depending on the rival players' strategies x_{-i}.

Given an arbitrary tuple of rival players' strategies x_{-i}, the aim of player i is to choose a strategy x_i that solves the following optimization problem:

$$\min_{x_i \in X_i\,(x_{-i})} \quad \theta_i(x_i, x_{-i}).$$

A vector of strategies $x^* \in \prod_{i=1}^{N} X_i(x^*_{-i})$ is a solution of the GNEP if, for all $i = 1, \dots, N$, one has

$$\theta_i(x^*_i, x^*_{-i}) \le \theta_i(x_i, x^*_{-i}), \qquad \forall\ x_i \in X_i(x^*_{-i}).$$

In other words, x^* is a solution of the GNEP if no player can decrease his objective function by unilaterally changing his strategy. A solution of the GNEP is also termed as generalized Nash equilibrium.

In this paper we focus on a special class of GNEPs referred to as jointly convex GNEPs. More precisely we assume that there is a closed and convex set $X \subseteq \mathbb{R}^n$, which represents the joint constraints of all the players, such that

$$X_i(x_{-i}) = \{x_i \in \mathbb{R}^{n_i} :\ (x_i, x_{-i}) \in X\},$$

for all $i = 1, \dots, N$. This condition results to be verified in several applications (see e.g. [6]).

A known approach for solving the jointly convex GNEP is to consider an equivalent optimization reformulation. The basis for this reformulation is the so called Nikaido-Isoda function (NI function, for short) $\Psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, introduced for the first time in [9]:

$$\Psi(x, y) = \sum_{i=1}^{N} \left[\theta_i(x_i, x_{-i}) - \theta_i(y_i, x_{-i})\right].$$

In particular, the NI function provides an important subset of all the solutions of the jointly convex GNEP. A vector $x^* \in X$ is called a normalized Nash equilibrium (NNE, for short) of the jointly convex GNEP if

$$\max_{y \in X} \Psi(x^*, y) = 0. \tag{1}$$

It is worth noting that the Ky Fan's inequality [1] guarantees the existence of a NNE when X is compact. It is not difficult to check that a NNE is always a solution of the jointly convex GNEP, whereas the converse is not true in general.

By means of the NI function, the problem of finding NNE can be reformulated as the optimization problem (see [5])

$$\min_{x \in X} \psi(x),$$

where

$$\psi(x) = \sup_{y \in X} \Psi(x, y). \tag{2}$$

However, for any given x, the supremum in (2) may not be attained, or it may be attained at more than one point and, consequently the function ψ would be nonsmooth. In [7] the authors provided a smooth optimization reformulation of the problem of finding NNE, based on a regularization of the NI function. Let us consider the function

$$\psi_\alpha(x) = \max_{y \in X} \left[\Psi(x, y) - \frac{\alpha}{2}\, \|y - x\|^2\right], \tag{3}$$

where α is a positive parameter. Since under the given assumptions $\Psi(x, \cdot)$ is concave, the maximum in (3) is attained, for each $x \in X$, in an unique point denoted $y^\alpha(x)$. In [7] the following properties have been proved:

- ψ_α is nonnegative and continuously differentiable on X,
- x^* is a NNE if and only if $x^* \in X$ and $\psi_\alpha(x^*) = 0$.

We remark that similar results can be obtained substituting, in the definition of ψ_α, the Euclidean norm with a norm induced by a symmetric positive definite matrix.

Hence any NNE is a global minimizer of the smooth constrained optimization problem

$$\min_{x \in X} \ \psi_\alpha(x), \tag{4}$$

with zero optimal value. A classical method for finding NNE, based on the minimization of the function ψ, is the so called relaxation method, introduced in [12]. A modification of the relaxation method, based on the minimization of the function ψ_α, has been presented in [8]. The convergence analysis of the method in [8] relies in a main assumption which guarantees that every stationary point of problem (4) is a NNE.

The aim of this paper is to propose a new descent type algorithm for finding NNE, by solving the optimization problem (4). We will show that our algorithm is globally convergent to a NNE under an appropriate assumption on the payoff functions, which is not stronger than the one considered in [8] and does not guarantee stationary points of (4) to be NNE. The organization of the paper is as follows: in Section 2 we state the main assumption underlying our algorithm, we present some classes of GNEPs verifying it, and we discuss its relationship with the one used in [8]. In Section 3 we formally state our algorithm and we prove that it is globally convergent to a NNE. Finally, in Section 4 we present some numerical results.

2 Main assumption

The following assumption will be the key property to construct the descent direction and to guarantee the convergence of our algorithm:

for any given $\alpha > 0$ and $x \in X$,

$$\begin{aligned} &\sum_{i=1}^N \langle \nabla\theta_i(x_i, x_{-i}) - \nabla\theta_i(y_i^\alpha(x), x_{-i}), x - y^\alpha(x)\rangle \geq \\ &\qquad \geq \sum_{i=1}^N \left[\theta_i(x_i, x_{-i}) - \theta_i(y_i^\alpha(x), x_{-i})) - \langle \nabla_{x_i}\theta_i(y_i^\alpha(x), x_{-i}), x_i - y_i^\alpha(x)\rangle\right]. \end{aligned} \tag{5}$$

The rest of this section is devoted to a discussion of assumption (5). In the following we present some classes of GNEPs verifying (5) and then we provide a sufficient condition for assumption (5) to be satisfied.

Example 2.1. Let us consider that, for all $i = 1, \dots, N$ the payoff functions θ_i are separable, that is

$$\theta_i(x) = f_i(x_i) + g_i(x_{-i}),$$

where $f_i : \mathbb{R}^{n_i} \to \mathbb{R}$ is convex and $g_i : \mathbb{R}^{N-n_i} \to \mathbb{R}$. A simple calculation shows that, for any $y \in \mathbb{R}^n$, we have

$$\sum_{i=1}^N \langle \nabla\theta_i(x_i, x_{-i}) - \nabla\theta_i(y_i, x_{-i}), x - y\rangle = \sum_{i=1}^N \left[\nabla f_i(x_i) - \nabla f_i(y_i)\right] (x_i - y_i),$$

and

$$\sum_{i=1}^{N} [\theta_i(x_i, x_{-i}) - \theta_i(y_i, x_{-i})) - \langle \nabla_{x_i}\theta_i(y_i, x_{-i}), x_i - y_i\rangle] = \sum_{i=1}^{N} [f_i(x_i) - f_i(y_i) - \nabla f_i(y_i)(x_i - y_i)] \, .$$

Moreover, from the convexity of f_i it follows that

$$\nabla f_i(x_i)\,(x_i - y_i) \geq f_i(x_i) - f_i(y_i),$$

hence (5) holds.

Example 2.2. Let us consider the case where the payoff functions are quadratic, i.e. for all $i = 1, \ldots, N$ one has

$$\theta_i(x) = \frac{1}{2}\langle x_i, A_{ii}\, x_i\rangle + \sum_{\substack{j=1 \\ j\neq i}}^{N} \langle x_i, A_{ij}\, x_j\rangle,$$

where the matrices $A_{ij} \in \mathbb{R}^{n_i \times n_j}$ and A_{ii} are symmetric positive semidefinite. Then we have:

$$\sum_{i=1}^{N} \langle \nabla\theta_i(x_i, x_{-i}) - \nabla\theta_i(y_i, x_{-i}), x - y\rangle = \langle (x-y), \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} (x-y)\rangle,$$

and

$$\sum_{i=1}^{N} [\theta_i(x_i, x_{-i}) - \theta_i(y_i, x_{-i})) - \langle \nabla_{x_i}\theta_i(y_i, x_{-i}), x_i - y_i\rangle] = \frac{1}{2}\sum_{i=1}^{N} \langle x_i - y_i, A_{ii}\,(x_i - y_i)\rangle.$$

Therefore, if the matrix

$$\begin{pmatrix} \frac{1}{2}A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & \frac{1}{2}A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & \frac{1}{2}A_{nn} \end{pmatrix}$$

is positive semidefinite, (5) is satisfied.

Example 2.3. Let us consider a generalization of the Example 2.2 as follows:

$$\theta_i(x) = \frac{1}{2}\langle x, Q^i\, x\rangle + \langle c^i, x\rangle + a^i = \frac{1}{2}\sum_{p,q=1}^{N} \langle x_p, Q^i_{pq}\, x_q\rangle + \sum_{j=1}^{N} c^i_j\, x_j + a^i$$

where $a^i \in \mathbb{R}$, $c^i \in \mathbb{R}^n$, the matrices $Q^i \in \mathbb{R}^{n\times n}$ are symmetric and $Q^i_{ii} \in \mathbb{R}^{n_i \times n_i}$ are positive semidefinite for all $i = 1, \ldots, N$. Then we have:

$$\sum_{i=1}^{N} \langle \nabla\theta_i(x_i, x_{-i}) - \nabla\theta_i(y_i, x_{-i}), x - y\rangle = \sum_{i=1}^{N}\sum_{j=1}^{N} \langle x_i - y_i, Q^i_{ij}\,(x_j - y_j)\rangle,$$

and

$$\sum_{i=1}^{N} [\theta_i(x_i, x_{-i}) - \theta_i(y_i, x_{-i})) - \langle \nabla_{x_i}\theta_i(y_i, x_{-i}), x_i - y_i\rangle] = \sum_{i=1}^{N} \langle x_i - y_i, Q^i_{ii}\,(x_i - y_i)\rangle.$$

It follows that if the matrix
$$\begin{pmatrix} \frac{1}{2}Q^1_{11} & Q^1_{12} & \cdots & Q^1_{1n} \\ Q^2_{21} & \frac{1}{2}Q^2_{22} & \cdots & Q^2_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ Q^n_{n1} & Q^n_{n2} & \cdots & \frac{1}{2}Q^n_{nn} \end{pmatrix}$$
is positive semidefinite, (5) holds.

We now show a general result that provides a sufficient condition for (5) to be satisfied.

Proposition 2.1. *If for each $y \in X$ the function $\Psi(\cdot, y)$ is convex on X, then (5) holds.*

Proof. Since Ψ is convex on X with respect to x, we have that

$$0 = \Psi(y^\alpha(x), y^\alpha(x)) \geq \Psi(x, y^\alpha(x)) + \langle \nabla_x \Psi(x, y^\alpha(x)), y^\alpha(x) - x \rangle \qquad \forall\, x \in X.$$

Moreover, the gradient of Ψ with respect to x is given by

$$\nabla_x \Psi(x, y^\alpha(x)) = \sum_{i=1}^{N} [\nabla \theta_i(x_i, x_{-i}) - \nabla \theta_i(y_i^\alpha(x), x_{-i})] + \begin{pmatrix} \nabla_{x_1} \theta_1(y_1^\alpha(x), x_{-1}) \\ \vdots \\ \nabla_{x_N} \theta_N(y_N^\alpha(x), x_{-N}) \end{pmatrix}.$$

Hence (5) holds. □

It is worth comparing the convexity of $\Psi(\cdot, y)$ and (5). The hypothesis of Proposition 2.1 means that
$$\Psi(z, y) \geq \Psi(x, y) + \langle \nabla_x \Psi(x, y), z - x \rangle, \tag{6}$$
for all $x, y, z \in X$. Whereas, (5) is equivalent to condition (6) for $z = y = y^\alpha(x)$. The following example shows that (5) is weaker than convexity of the function Ψ.

Example 2.4. Consider the GNEP with $N = 2$, $X = \{x \in \mathbb{R}^2 : \; x_1, x_2 \geq 1\}$ and payoff functions $\theta_1(x) = x_1\, x_2$ and $\theta_2(x) = x_2$. The NI function is

$$\Psi(x, y) = x_1\, x_2 - y_1\, x_2 + x_2 - y_2,$$

hence $\Psi(\cdot, y)$ is not convex on X.

Now let us consider $y^\alpha(x)$, the unique solution of

$$\max_{y \in X} [x_1\, x_2 - y_1\, x_2 + x_2 - y_2 - \frac{\alpha}{2}\,(y_1 - x_1)^2 - \frac{\alpha}{2}\,(y_2 - x_2)^2].$$

We obtain
$$y_1^\alpha(x) = \max\{1, x_1 - \frac{x_2}{\alpha}\}, \quad y_2^\alpha(x) = \max\{1, x_2 - \frac{1}{\alpha}\}.$$
Assumption (5) is equivalent to
$$(x_1 - y_1^\alpha)\,(x_2 - y_2^\alpha) \geq 0,$$
which is satisfied for all $x \in X$ and $\alpha > 0$.

The following examples allow us to point out the relationship between our assumption (5) and the one considered in [8], here reported for the sake of completeness.

Let $\alpha > 0$ be given. For any $x \in X$ such that $x \neq y^\alpha(x)$,

$$\sum_{i=1}^{N} \langle \nabla\theta_i(x_i, x_{-i}) - \nabla\theta_i(y_i^\alpha(x), x_{-i}), x - y^\alpha(x)\rangle > 0. \tag{7}$$

In [8] the authors showed that (7) suffices to guarantee that every stationary point of problem (4) is a global minimum and hence a NNE.

We now prove that assumption (5) is not stronger than (7) showing two examples in which (5) holds but (7) does not. Example 2.6 also shows that (5) does not guarantee that stationary points of (4) are NNE.

Example 2.5. Consider the GNEP with $N = 2$,

$$X = \left\{x \in \mathbb{R}^2 : \; x_1 \geq 1, \quad x_2 \geq 1, \quad x_1 + x_2 \leq 10\right\},$$

and payoff functions $\theta_1(x) = x_1\, x_2$ and $\theta_2(x) = -x_1\, x_2$. The point $x^* = (1, 9)$ is the unique NNE. We note that, for any $y \in \mathbb{R}^n$, we have

$$\sum_{i=1}^{N} \langle \nabla\theta_i(x_i, x_{-i}) - \nabla\theta_i(y_i, x_{-i}), x - y\rangle = (x_1 - y_1)\,(x_2 - y_2) + (x_1 - y_1)\,(y_2 - x_2) = 0,$$

and

$$\begin{aligned}&\sum_{i=1}^{N} \left[\theta_i(x_i, x_{-i}) - \theta_i(y_i, x_{-i})\right) - \langle \nabla_{x_i}\theta_i(y_i, x_{-i}), x_i - y_i\rangle]\\ &= x_1\, x_2 - y_1\, x_2 - x_2\,(x_1 - y_1) - x_1\, x_2 + x_1\, y_2 + x_1\,(x_2 - y_2)\\ &= 0,\end{aligned}$$

therefore condition (5) holds, but (7) does not hold for any given $\alpha > 0$.

Example 2.6. Consider the GNEP with $N = 2$, $X = \left\{x \in \mathbb{R}^2 : \; x_1 \geq 1, \quad x_2 \geq 1\right\}$ and payoff functions $\theta_1(x) = \frac{1}{2}\, x_1^2$ and $\theta_2(x) = x_2$. The unique NNE is $x^* = (1, 1)$.
We see that, for any y, we have

$$\sum_{i=1}^{N} \langle \nabla\theta_i(x_i, x_{-i}) - \nabla\theta_i(y_i, x_{-i}), x - y\rangle = (x_1 - y_1)^2,$$

and

$$\sum_{i=1}^{N} \left[\theta_i(x_i, x_{-i}) - \theta_i(y_i, x_{-i})\right) - \langle \nabla_{x_i}\theta_i(y_i, x_{-i}), x_i - y_i\rangle] = \frac{1}{2}(x_1 - y_1)^2,$$

therefore (5) holds. In the following we show that assumption (7) does not hold.

Given an arbitrary $\alpha > 0$, the gap function is

$$\begin{aligned}\psi_\alpha(x) \quad &= \max_{y \in X} \left[\frac{1}{2}\, x_1^2 + x_2 - \frac{1}{2} y_1^2 - y_2 - \frac{\alpha}{2}(y_1 - x_1)^2 - \frac{\alpha}{2}(y_2 - x_2)^2\right]\\ &= \max_{y_1 \geq 1} \left[-\frac{(1+\alpha)}{2}\, y_1^2 + \alpha\, x_1\, y_1\right] + \max_{y_2 \geq 1} \left[-\frac{\alpha}{2}\, y_2^2 + (\alpha\, x_2 - 1)\, y_2\right] + \frac{(1-\alpha)}{2}\, x_1^2 - \frac{\alpha}{2}\, x_2^2 + x_2.\end{aligned}$$

It easy to show that the component functions of y^α are given by:

$$y_1^\alpha(x) = \begin{cases} 1 & \text{if } x_1 \leq 1 + \frac{1}{\alpha}, \\ \frac{\alpha}{1+\alpha}\, x_1 & \text{if } x_1 > 1 + \frac{1}{\alpha}, \end{cases}$$

and

$$y_2^\alpha(x) = \begin{cases} 1 & \text{if } x_2 \leq 1 + \frac{1}{\alpha}, \\ x_2 - \frac{1}{\alpha} & \text{if } x_2 > 1 + \frac{1}{\alpha}. \end{cases}$$

The gap function for this GNEP is

$$\psi_\alpha(x) = \begin{cases} \frac{x_1^2}{2\,(1+\alpha)} + \frac{1}{2\,\alpha} & \text{if } x_1, x_2 \geq 1 + \frac{1}{\alpha}, \\ (\frac{1-\alpha}{2})\, x_1^2 + \alpha\, x_1 - \frac{\alpha^2+\alpha-1}{2\alpha} & \text{if } 1 \leq x_1 \leq 1 + \frac{1}{\alpha} \leq x_2, \\ \left(\frac{1-\alpha}{2}\right)\, x_1^2 - \frac{\alpha}{2}\, x_2^2 + \alpha\, x_1 + (\alpha+1)\, x_2 - \frac{\alpha+3}{2} & \text{if } x_1, x_2 \in \left[1, 1 + \frac{1}{\alpha}\right], \\ -\frac{\alpha}{2}\, x_2^2 + (\alpha+1)\, x_2 - \frac{\alpha+2}{2} & \text{if } 1 \leq x_2 \leq 1 + \frac{1}{\alpha} \leq x_1. \end{cases}$$

Let us consider the case $1 \leq x_1 \leq 1 + \dfrac{1}{\alpha} \leq x_2$, we get

$$\nabla\psi_\alpha(x) = \begin{pmatrix} (1-\alpha)\, x_1 + \alpha \\ 0 \end{pmatrix}.$$

In particular, for $x_1 = 1$ we have $\nabla\psi_\alpha(x) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Therefore the points $(1, x_2)$ with $x_2 \geq 1 + \dfrac{1}{\alpha}$ are stationary points of problem (4), but they are not NNE since $\psi_\alpha(1, x_2) = \dfrac{1}{2\,\alpha} \neq 0$. Hence, in particular, (7) does not hold.

3 Solution method

In this section we present our algorithm and we prove that it is globally convergent to a NNE. In the following result we prove that (5) does not guarantee that the vector $y^\alpha(x) - x$ is a descent direction for ψ_α, however, it provides a useful condition that can be exploited to construct a descent type algorithm.

Theorem 3.1. *If (5) holds, then for each $x \in X$ the vector $y^\alpha(x) - x$ satisfies the following condition:*

$$\langle \nabla\psi_\alpha(x), y^\alpha(x) - x \rangle \leq -\psi_\alpha(x) + \frac{\alpha}{2}\, \|y^\alpha(x) - x\|^2. \tag{8}$$

Proof. The function ψ_α is continuously differentiable on X and its gradient is given by (see [2])

$$\nabla\psi_\alpha(x) = \sum_{i=1}^{N} \left[\nabla\theta_i(x_i, x_{-i}) - \nabla\theta_i(y_i^\alpha(x), x_{-i})\right] + \begin{pmatrix} \nabla_{x_1}\theta_1(y_1^\alpha(x), x_{-1}) \\ \vdots \\ \nabla_{x_N}\theta_N(y_N^\alpha(x), x_{-N}) \end{pmatrix} + \alpha\, \left[y^\alpha(x) - x\right].$$

Therefore

$$\begin{aligned} \langle \nabla\psi_\alpha(x), y^\alpha(x) - x \rangle \quad &= \sum_{i=1}^{N} \langle \nabla\theta_i(x_i, x_{-i}) - \nabla\theta_i(y_i^\alpha(x), x_{-i}), y^\alpha(x) - x \rangle \\ &\quad + \sum_{i=1}^{N} \langle \nabla_{x_i}\theta_i(y_i^\alpha(x), x_{-i}), y_i^\alpha(x) - x_i \rangle + \alpha\, \|y^\alpha(x) - x\|^2 \end{aligned} \tag{9}$$

In view of (5), we obtain

$$\begin{aligned}\langle \nabla \psi_\alpha(x), y^\alpha(x) - x\rangle &\le \sum_{i=1}^{N} [\theta_i(y_i^\alpha(x), x_{-i})) - \theta_i(x_i, x_{-i})] + \alpha \, \|y^\alpha(x) - x\|^2 \\ &= -\Psi(x, y^\alpha) + \alpha \, \|y^\alpha(x) - x\|^2 \\ &= -\psi_\alpha(x) + \frac{\alpha}{2} \, \|y^\alpha(x) - x\|^2.\end{aligned}$$

□

Taking into account (8), the idea behind the algorithm to be proposed is to perform a line search along $y^\alpha(x) - x$, if it is a descent direction for ψ_α, or to reduce the parameter α otherwise. A similar idea was proposed in [11, 13] for solving monotone variational inequalities using the gap or D-gap function approach. We remark that we do not need to compute any derivative of the payoff functions and that we use an Armijo type stepsize rule, for which only evaluations of the function ψ_α are necessary. Now we formally state our algorithm as follows.

Algorithm

0. (Initial step)

   Let $\eta, \gamma \in (0, 1)$, and $\beta \in (0, \eta)$. Let $\{\alpha_k\}$ be a sequence strictly decreasing to 0.

   Choose any $x^0 \in X$ and set $k = 0$.

1. (Stopping criterion)

   If $\psi_{\alpha_k}(x^k) = 0$ **then** STOP, **else** set $k = k + 1$.

2. (Minimization of ψ_{α_k})

   2a. (Initialization) Set $\ell = 0$ and $z^0 = x^{k-1}$.

   2b. Compute $y^{\alpha_k}(z^\ell) = \arg\max_{y \in X} \left[\Psi(z^\ell, y) - \frac{\alpha_k}{2} \, \|y - z^\ell\|^2\right]$.

   If $-\psi_{\alpha_k}(z^\ell) + \frac{\alpha_k}{2} \, \left\|y^{\alpha_k}(z^\ell) - z^\ell\right\|^2 < -\eta \, \psi_{\alpha_k}(z^\ell)$

   then (line search)

   set $d^\ell = y^{\alpha_k}(z^\ell) - z^\ell$

   compute the smallest nonnegative integer m such that:

   $$\psi_{\alpha_k}(z^\ell + \gamma^m \, d^\ell) - \psi_{\alpha_k}(z^\ell) \le -\beta \, \gamma^m \, \psi_{\alpha_k}(z^\ell)$$

   set $t_\ell = \gamma^m$,

   else (update of x^k) set $x^k = z^\ell$ and return to step 1.

   2c. (Update of z^ℓ) Set $z^{\ell+1} = z^\ell + t_\ell \, d^\ell$, $\ell = \ell + 1$, and return to step 2b.

Theorem 3.2. *If (5) holds and the set X is bounded, then the algorithm proposed either stops at a NNE after a finite number of iterations, or generates a sequence $\{x^k\}$ such that any of its cluster points is a NNE, or generates a sequence $\{z^\ell\}$, for some fixed α_k, such that any of its cluster points is a NNE.*

Proof. First, we show that the algorithm is well defined, i.e. that the line search procedure is always finite. To this end assume, by contradiction, that there are $\ell, k \geq 0$ such that the inequality

$$\psi_{\alpha_k}(z^\ell + \gamma^m d^\ell) - \psi_{\alpha_k}(z^\ell) > -\beta\,\gamma^m\,\psi_{\alpha_k}(z^\ell),$$

holds for all $m \in \mathbb{N}$. Then we have:

$$\psi'_{\alpha_k}(z^\ell; d^\ell) = \lim_{m \to +\infty} \frac{\psi_{\alpha_k}(z^\ell + \gamma^m d^\ell) - \psi_{\alpha_k}(z^\ell)}{\gamma^m} \geq -\beta\,\psi_{\alpha_k}(z^\ell).$$

Combining (8) and step 2b, we get:

$$\psi'_{\alpha_k}(z^\ell; d^\ell) \leq -\psi_{\alpha_k}(z^\ell) + \frac{\alpha_k}{2}\,\|d^\ell\|^2 < -\,\eta\,\psi_{\alpha_k}(z^\ell),$$

therefore

$$(\eta - \beta)\,\psi_{\alpha_k}(z^\ell) < 0,$$

which is impossible because $\eta > \beta$ and $\psi_{\alpha_k}(z^\ell) \geq 0$. So the line search procedure is always finite.

There are three possible cases.

Case 1. The algorithm stops at x^k after a finite number of iterations. From the stopping criterion it follows that x^k is a NNE.

Case 2. The algorithm generates an infinite sequence $\{x^k\}$. From condition at step 2b we have

$$\psi_{\alpha_k}(x^k) \leq \frac{\alpha_k}{2\,(1-\eta)}\,\left\|x^k - y^{\alpha_k}(x^k)\right\|^2 \qquad \forall\, k \in \mathbb{N}.$$

Since x^k and $y^{\alpha_k}(x^k)$ belong to X which is a bounded set, the norm $\|x^k - y^{\alpha_k}(x^k)\|$ is bounded above. Because $\lim_{k\to\infty} \alpha_k = 0$, we have

$$\lim_{k \to \infty} \psi_{\alpha_k}(x^k) = 0. \tag{10}$$

Let x^* be any cluster point of $\{x^k\}$ and x^{k_s} a subsequence converging to x^*. From the definition of ψ_{α_k} it follows that for each $y \in X$ we have

$$\psi_{\alpha_{k_s}}(x^{k_s}) \geq \Psi(x^{k_s}, y) - \frac{\alpha_{k_s}}{2}\,\|x^{k_s} - y\|^2 \qquad \forall\, s \in \mathbb{N}.$$

Moreover Ψ is continuous, $\lim_{k\to\infty} \alpha_k = 0$, and (10) holds, thus passing to the limit we obtain

$$0 \geq \Psi(x^*, y).$$

Since y is arbitrary, we have proved that $\max_{y \in X} \Psi(x^*, y) = 0$, that is x^* is a NNE.

Case 3. The algorithm generates an infinite sequence $\{z^\ell\}$ for a fixed $\alpha_k = \alpha$. Let us consider two possible subcases: either $\limsup_{\ell\to\infty} t_\ell > 0$, or $\limsup_{\ell\to\infty} t_\ell = 0$.

Subcase 3a. If $\limsup_{\ell\to\infty} t_\ell > 0$, then there exists $t^* > 0$ and a subsequence $\{t_{\ell_s}\}$ such that $t_{\ell_s} \geq t^* > 0$ for all $s \in \mathbb{N}$. Since the sequence $\{z^\ell\}$ is infinite, we have:

$$\psi_\alpha(z^{\ell_s}) - \psi_\alpha(z^{\ell_s+1}) \geq \beta\,t_{\ell_s}\,\psi_\alpha(z^{\ell_s}) \geq \beta\,t^*\,\psi_\alpha(z^{\ell_s}) \geq 0. \tag{11}$$

The sequence $\{\psi_\alpha(z^\ell)\}$ is monotone decreasing and bounded below, hence

$$\lim_{\ell \to \infty} [\psi_\alpha(z^\ell) - \psi_\alpha(z^{\ell+1})] = 0,$$

and in particular

$$\lim_{s\to\infty} [\psi_\alpha(z^{\ell_s}) - \psi_\alpha(z^{\ell_s+1})] = 0. \tag{12}$$

Using (11) and (12), we obtain $\lim_{s\to\infty} \psi_\alpha(z^{\ell_s}) = 0$ and thus $\lim_{\ell\to\infty} \psi_\alpha(z^\ell) = 0$. If z^* is any cluster point of $\{z^\ell\}$, then from the continuity of ψ we have $\lim_{\ell\to\infty} \psi_\alpha(z^\ell) = \psi_\alpha(z^*)$, hence $\psi_\alpha(z^*) = 0$, i.e. z^* is a NNE.

Subcase 3b. If $\limsup_{\ell\to\infty} t_\ell = 0$, then $\lim_{\ell\to\infty} t_\ell = 0$. From the step length rule it follows that for all $\ell \in \mathbb{N}$,

$$\psi_\alpha(z^\ell + \gamma^{-1} t_\ell d^\ell) - \psi_\alpha(z^\ell) > -\beta \gamma^{-1} t_\ell \psi_\alpha(z^\ell). \tag{13}$$

By the mean value theorem we have:

$$\psi_\alpha(z^\ell + \gamma^{-1} t_\ell d^\ell) - \psi_\alpha(z^\ell) = \langle \nabla\psi_\alpha(z^\ell + \delta_\ell \gamma^{-1} t_\ell d^\ell), \gamma^{-1} t_\ell d^\ell \rangle, \tag{14}$$

for some $\delta_\ell \in (0, 1)$. In view of (13) and (14) we obtain:

$$\langle \nabla\psi_\alpha(z^\ell + \delta_\ell \gamma^{-1} t_\ell d^\ell), d^\ell \rangle > -\beta \psi_\alpha(z^\ell) \qquad \forall \ \ell \in \mathbb{N}. \tag{15}$$

Let z^* be any cluster point of $\{z^\ell\}$. Since $\lim_{k\to\infty} t_\ell = 0$, passing to the limit in (15) and taking a subsequence if necessary, we get:

$$\langle \nabla\psi_\alpha(z^*), d^* \rangle \geq -\beta \psi_\alpha(z^*), \tag{16}$$

where $d^* = y^\alpha(x^*) - x^*$.

Moreover, for all $\ell \in \mathbb{N}$, we have:

$$-\psi_\alpha(z^\ell) + \frac{\alpha}{2} \|z^\ell - y^\alpha(z^\ell)\|^2 < -\eta \psi_\alpha(z^\ell),$$

hence passing to the limit and taking a subsequence if necessary, by Theorem 3.1 it follows that

$$\langle \nabla\psi_\alpha(z^*), d^* \rangle \leq -\psi_\alpha(z^*) + \frac{\alpha}{2} \|d^*\|^2 \leq -\eta \psi_\alpha(z^*). \tag{17}$$

¿From (16) and (17) we get

$$(\eta - \beta) \psi_\alpha(z^*) \leq 0.$$

Since $\eta > \beta$ and $\psi_\alpha(z^*) \geq 0$, it follows that $\psi_\alpha(z^*) = 0$, i.e. z^* is a NNE. □

Remark 3.1. *Note that in the Algorithm the sequence $\{\alpha_k\}$ can be chosen adaptively, for example (see [11]) such as:*

$$\alpha_k = \begin{cases} \alpha_{k-1} & \textit{if } \psi_{\alpha_{k-1}}(x^{k-1}) \leq \nu_{k-1}, \\ \mu \alpha_{k-1} & \textit{otherwise}, \end{cases} \tag{18}$$

where $0 < \mu < 1$ and $\{\nu_k\}$ is a sequence decreasing to 0. Indeed, if the algorithm generates an infinite sequence $\{x^k\}$ with $\{\alpha_k\}$ chosen by (18), then either $\lim_{k\to\infty} \alpha_k = 0$, which can be treated as in the Case 2 of Theorem 3.2, or one has

$$\alpha_k = \bar{\alpha} \quad \textit{and} \quad \psi_{\bar{\alpha}}(x^k) \leq \nu_k \qquad \forall \ k > \bar{k},$$

hence $\lim_{k\to\infty} \psi_{\bar{\alpha}}(x^k) = 0$. Then, for each cluster point x^* of $\{x^k\}$, we have $\psi_{\bar{\alpha}}(x^*) = 0$, that is x^* is a NNE.

Remark 3.2. *The version of the algorithm discussed in this paper determines the stepsize using an Armijo-type rule. It can be proved that the same global convergence result holds with an exact line search as well.*

4 Numerical results

In the following we consider three examples to test the algorithm presented in the previous section. We implemented the algorithm in MATLAB and we used the Optimization Toolbox to compute y^{α_k}. We set the algorithm parameters as follows: $\eta = 0.5$, $\beta = 0.4$, $\gamma = 0.5$, and $\alpha_k = 5/5^k$. The algorithm stops whenever $\psi_{\alpha_k}(x^k) < 10^{-12}$.

For each example numerical results are summarized in two tables. We solved each example applying our algorithm starting from different points in the set X. In the first table the behaviour of the algorithm for different initial points is reported. For each initial point (first column) the number of optimization problems solved to compute y^{α_k} is given in column two, the number of outer iterations, i.e. the number of k updates, is given in column three, and the number of inner iterations, i.e. the number of performed line searches, is given in column four.

Second table is devoted to describe the algorithm steps starting from one initial point. For each outer iteration of the algorithm the value of k, α_k, and x^k are given in the first three columns. The value of $\psi_{\alpha_k}(x^k)$ is reported in column four. Then columns five to nine describe the inner iterations. In columns five, six, and seven the value of ℓ, z^ℓ, and $\psi_{\alpha_k}(z^\ell)$ are given respectively. Finally, column eight tells whether line search is performed and column nine gives the corresponding stepsize.

Example 4.1. Let us consider the two-player GNEP described in Example 2.5. The unique NNE is $(1, 9)$ and therefore the algorithm reaches it starting from any point. Tables 1 and 2 report numerical results on such example as described above.

starting point	# optimization problems	# outer iterations	# inner iterations
$(1,1)$	5	3	1
$(1,8)$	4	2	1
$(2,3)$	4	2	1
$(2,4)$	5	2	2
$(2,6)$	5	2	2
$(3,4)$	4	2	1
$(3,7)$	3	1	1
$(4,3)$	4	2	1
$(4,6)$	3	1	1
$(5,5)$	4	2	1
$(6,4)$	4	2	1
$(8,1)$	4	2	1
$(9,1)$	4	2	1

Table 1: Numerical results for Example 4.1 for different initial points.

Example 4.2. Let us consider a two-player GNEP with payoff functions

$$\theta_1(x_1, x_2) = \frac{1}{2}\, x_1^2 \qquad \text{and} \qquad \theta_2(x_1, x_2) = x_2,$$

k	α_k	x^k	$\psi_{\alpha_k}(x^k)$	ℓ	z^ℓ	$\psi_{\alpha_k}(z^\ell)$	line search	stepsize
0	5	$(2,4)$	2					
1	1			0	$(2,4)$	5.5	yes	1
				1	$(1,6)$	0.5	no	–
		$(1,6)$	0.5					
2	0.2			0	$(1,6)$	2.1	yes	1
				1	$(1,9)$	0	no	–
		$(1,9)$	0					

Table 2: Steps of the algorithm on Example 4.1

and the set $X = \{x \in \mathbb{R}^2 : \ x_1 \geq 1, \quad x_2 \geq 1, \quad x_1 + x_2 \leq 10\}$. The unique NNE is $(1,1)$ and therefore the algorithm reaches $(1,1)$ starting from any point. The corresponding numerical results are given in Tables 3 and 4.

starting point	# optimization problems	# outer iterations	# inner iterations
$(1,4)$	4	2	1
$(1,9)$	5	3	1
$(2,5)$	4	2	1
$(2,8)$	5	3	1
$(3,3)$	4	2	1
$(3,7)$	5	2	2
$(4,4)$	4	2	1
$(5,2)$	4	2	1
$(6,3)$	4	2	1
$(7,3)$	5	2	2
$(9,1)$	5	2	2

Table 3: Numerical results for Example 4.2 for different initial points.

k	α_k	x^k	$\psi_{\alpha_k}(x^k)$	ℓ	z^ℓ	$\psi_{\alpha_k}(z^\ell)$	line search	stepsize
0	5	$(7,3)$	4.1833					
1	1			0	$(7,3)$	12.75	no	–
		$(7,3)$	12.75					
2	0.2			0	$(7,3)$	22.0167	yes	1
				1	$(1,1.6667)$	0.1778	yes	1
				2	$(1,1)$	0	no	–
		$(1,1)$	0					

Table 4: Steps of the algorithm on Example 4.2

Example 4.3. We consider a GNEP with $N = 5$, where each player i controls a single variable $x_i \in \mathbb{R}$, the payoff functions are given by:

$$\begin{aligned} \theta_1(x) &= \tfrac{1}{x_1} + x_2 \\ \theta_2(x) &= \tfrac{1}{x_2} + x_3 \\ \theta_3(x) &= x_3 + x_4 \\ \theta_4(x) &= x_4 + x_5 \\ \theta_5(x) &= x_5 + x_1 \end{aligned}$$

and the set $X = \left\{ x \in \mathbb{R}^5 : \; 10 \leq \sum_{i=1}^{5} x_i \leq 20, \quad x_i \geq 1 \quad \forall\, i = 1, \dots, 5 \right\}$. In this example the algorithm converges to $(8.5, 8.5, 1, 1, 1)$ starting from all the considered initial points. The corresponding numerical results are given in Tables 5 and 6.

starting point	# optimization problems	# outer iterations	# inner iterations
$(2, 2, 5, 3, 8)$	9	5	3
$(1, 2, 5, 10, 1)$	12	6	5
$(5, 5, 3, 2, 5)$	8	5	2
$(1, 7, 4, 2, 1)$	15	6	8
$(4, 1, 6, 4, 5)$	12	6	5
$(2, 2, 4, 6, 4)$	9	5	3
$(1, 5, 7, 1, 1)$	13	6	6
$(2, 2, 2, 5, 5)$	9	5	3
$(5, 4, 1, 5, 3)$	11	6	4
$(2, 1, 2, 2, 8)$	12	6	5
$(4, 4, 7, 2, 3)$	8	5	2

Table 5: Numerical results for Example 4.3 for different initial points.

References

[1] J.P. Aubin, and I. Ekeland, Applied nonlinear analysis, New York:John Wiley & Sons Inc., 1984.

[2] A. Auslender, Optimization. Mthodes numriques, Paris: Masson, 1976.

[3] E. Cavazzuti, M. Pappalardo, and M. Passacantando, Nash equilibria, variational inequalities, and dynamical systems, *Journal of Optimization Theory and Applications* (2002), vol. 114, pp. 491–506.

[4] F. Facchinei, A. Fisher, and V. Piccialli, On generalized Nash games and variational inequalities, *Operations Research Letters* (2007), vol. 35, pp. 159–164.

[5] F. Facchinei and C. Kanzow, Generalized Nash equilibrium problems, *4OR* (2007), vol. 5, pp. 173–210.

k	α_k	x^k	$\psi_{\alpha_k}(x^k)$	ℓ	z^ℓ	$\psi_{\alpha_k}(z^\ell)$	line search	stepsize
0	5	$(2,1,2,2,8)$	0.38014					
1	1			0	$(2,1,2,2,8)$	1.73477	no	-
		$(2,1,2,2,8)$	1.73477					
2	0.2			0	$(2,1,2,2,8)$	4.78178	no	-
		$(2,1,2,2,8)$	4.78178					
3	0.04			0	$(2,1,2,2,8)$	8.74326	yes	1
				1	$(3.7628,3.2978,1,1,1)$	0.08806	no	-
		$(3.7628,3.2978,1,1,1)$	0.08806					
4	0.008			0	$(3.7628,3.2978,1,1,1)$	0,19054	yes	1
				1	$(6.6366,6.3602,1,1,1)$	0.04093	no	-
		$(6.6366,6.3602,1,1,1)$	0.04093					
5	0.0016			0	$(6.6366,6.3602,1,1,1)$	0.06617	yes	1
				1	$(8.5087,8.4913,1,1,1)$	1.37e-07	no	-
		$(8.5087,8.4913,1,1,1)$	1.37e-07					
6	0.0003			0	$(8.5087,8.4913,1,1,1)$	2.23e-07	yes	1
				1	$(8.5001,8.4999,1,1,1)$	2.56e-11	yes	1
				2	$(8.5000,8.5000,1,1,1)$	2.81e-15	no	-
		$(8.5000,8.5000,1,1,1)$	2.81e-15					

Table 6: Steps of the algorithm on Example 4.3

[6] P. T. Harker, Generalized Nash games and quasi-variational inequalities, *European Journal of Operational Research* (1991), vol. 54, pp. 81–94.

[7] A. von Heusinger and C. Kanzow, Optimization reformulations of the generalized Nash equilibrium problem using Nikaido-Isoda-type functions, *Computational Optimization and Applications*, to appear.

[8] A. von Heusinger and C. Kanzow, Relaxation methods for generalized Nash equilibrium problem with inexact line search, *Preprint 282, Institute of Mathematics, University of Würzburg, Würzburg, February 2008.*

[9] H. Nikaido and K. Isoda, Note on non-cooperative convex games, *Pacific Journal of Mathematics* (1955), vol. 5, pp. 807–815.

[10] B. Panicucci, M. Pappalardo and M. Passacantando, A globally convergent descent method for nonsmooth variational inequalities, *Computational Optimization and Applications*, to appear.

[11] M. V. Solodov and P. Tseng, "Some methods based on the d-gap function for solving monotone variational inequalities", Computational Optimization and Applications, vol. 17, pp. 255–277, 2000.

[12] S. Uryasev and R.Y. Rubinstein, On relaxation algorithms in computation of noncooperative equilibria *IEEE Transaction on Automatic Control* (1994), vol. 39, pp. 1263–1267.

[13] D. L. Zhu and P. Marcotte, "Modified descent methods for solving the monotone variational inequality problem", Operations Research Letters, vol. 14, pp. 111–120, 1993.