A Trust Region Algorithm for Nash Equilibrium Problems*

Ya-xiang YUAN†
State Key Laboratory of Scientific/Engineering Computing,
Institute of Computational Mathematics and Scientific/Engineering Computing,
The Academy of Mathematics and Systems Sciences,
Chinese Academy of Sciences, P.O.Box 2719, Beijing, 100080, P.R.China,
E-mail: yyx@lsec.cc.ac.cn

Abstract

The Nash equilibrium problem is fundamental in economics and it is also a very special optimization problem. In this paper, we consider the application of trust region methods to Nash equilibrium problems. We propose a Jacobi-type trust region method for their solutions. The method includes different trust regions for each player, and the trial step is computed and accepted (or rejected) based on each individual utility function. An overall merit function is used and a non-standard technique is suggested to update the trust region bounds. Under certain conditions, we prove the global convergence and local superlinear convergence of the method.

Keywords: Nash equilibrium, trust region, convergence

AMS(MOS) Subject Classification: 65K05, 90C99

1 Introduction

The Nash Equilibrium Problem[18] is about N players ($v = 1, ..., N$), where each player v controls variable $x_v \in \Re^{n_v}$ and wants to minimize his own utility function u_v. Let $n = \sum_{i=1}^{N} n_v$ be the total number of variables and x be the vector of all variables

$$x = \begin{pmatrix} x_1, \\ \vdots \\ x_N \end{pmatrix} \in \Re^n . \tag{1.1}$$

The utility function u_v for the v-th player depends on all the variables x. Since the v-th player controls variables x_v, it is convenient to partition the variables x into two subsets: x_v and $\mathbf{x}^{-v}$, where $\mathbf{x}^{-v}$ are all the variables in x except those in x_v. Thus, $(x_v, \mathbf{x}^{-v})$ is a decomposition of x. Using such notations, the utility function $u_v(x)$ for the v-th player can be written as

$$u_v(x_v, \mathbf{x}^{-v}) . \tag{1.2}$$

*partly supported by Chinese NSF grant 10231060 and the Knowledge Innovation Program of CAS
†currently visiting D.A.M.T.P, University of Cambridge, England.

The aim of the v-th player is to minimize $u_v(x)$ by controlling the variables in x_v, namely

$$\min_{x_v \in X_v} u_v(x_v, \mathbf{x}^{-v}), \tag{1.3}$$

where $X_v \subset \Re^{n_v}$ is the set of possible strategies of player v. Normally X_v is compact convex set. Often, X_v is given by

$$X_v = \{x_v \mid \quad c_v(x_v) = 0,\ h_v(x_v) \leq 0,\ x_v \in \Re^{n_v}\}, \tag{1.4}$$

where $c_v \in \Re^{l_v}$ and $h_v \in \Re^{m_v}$ are continuous vector functions defined in $\Re^{n_v}$ with l_v and m_v being two positive integers. A Nash equilibrium point is a point x^* such that $(x^*)_v$ is a solution of

$$\min_{x_v \in X_v} u_v(x_v, (\mathbf{x}^*)^{-v}) \tag{1.5}$$

for all $v = 1, 2, ..., N$.

The Nash equilibrium problem (NEP)is fundamental in economics, and it has been studied extensively with many extensions, for example see [4, 8, 12, 13, 14, 15, 17, 20]. In this paper we propose a trust region algorithm for solving NEP. Trust region algorithms have been used successfully for nonlinear optimization, following the pioneer work of Powell [21, 22] and Fletcher [10, 11]. The general frame of trust region algorithms is at each iteration the original nonlinear (difficult) problem is replaced by an approximate model which is easy to solve, and then the approximate model is solved with the "trust region" constraint (requiring the solution of the approximate model in a region that is trusted). The solution of the model problem is called the trust region step, and it is either accepted or rejected, depending on whether the trust region step gives an acceptable improvement to the original problem. The trust region itself is also updated from iteration to iteration based on the goodness of the trial step. Trust region algorithms for nonlinear optimization have attracted many researches, for example see [1, 2, 3, 5, 6, 23, 24]. For detailed discussions on trust region algorithms, readers are referred to the nice monograph of Conn, Gould and Toint [7].

The NEP can be formulated as a variational inequality (VI) problem or a complementarity problem (CP), which, by differentiable merit functions, can be cased as either constrained or unconstrained optimization problems to which trust region methods can be applied. For example, see Facchinei and Pang [9] and Majig and Fukushima [16]. But, such indirect approaches need to minimize objective functions that depend on the gradients of the utility functions. Hence the application of standard trust region algorithms for optimization would require the Hessian matrices of all the utility functions. Another possibility is to reformulate the NEP into an optimization problem by using the Nikaido-Isoda function, and then to apply a standard trust region method for the corresponding optimization problem. In this approach, the objective function is

$$V(x) = \sum_{v=1}^{N} \left[u_v(x_v, \mathbf{x}^{-v}) - \min_{y_v \in X_v} u_v(y_v, \mathbf{x}^{-v}) \right] . \tag{1.6}$$

The definition of $V(x)$ depends on $\min_{y_v \in X_v} u_v(y_v, \mathbf{x}^{-v})$ and the application of a standard trust region for $\min V(x)$ will be very complicated because the evaluation of the objective function itself need to solve N minimization problems $\min_{y_v \in X_v} u_v(y_v, \mathbf{x}^{-v})$, not to mention that the gradient $\nabla V(x)$ is very difficult to obtain.

To the author's knowledge, no trust region methods have been proposed to solve equilibrium problems directly. We believe that it is an interesting question to ask whether trust region algorithms, a class of widely applied algorithms for nonlinear optimization, can be made suitable for these problems. This is the motivation of the current paper, and we will give a Jocobi-type trust region algorithm for solving Nash equilibrium problems and study the convergence properties of the method.

The paper is organized as follows. In the next section, we give a Jacobi type trust region algorithm for Nash equilibrium problems. Our trust region method uses different trust regions for each player, and the trial step for the variables controlled by each player is computed and accepted (or rejected) based on the corresponding individual utility function. Then, an overall merit function based on all the utility functions is defined for the updating of the trust regions. A non-standard technique is suggested to update the trust region bounds. In Section 3, we prove the global convergence of the method, and local convergence result is established in Section 4. In Section 5, possible generalizations of our method are discussed briefly.

2 A Jacobi-type Trust Region Algorithm for NEP

The trust region approach imposes constraints on the step-lengths of the changes to the variables. At the k-th iteration, assume that the current variables of the player v is $(x^{(k)})_v$. Let the bound on the norm of the changes to the variables controlled by the v-th player at the k-th iteration be denoted by $\Delta_{v,k}$. Thus the trust region subproblem for the v-th player at the k-th iteration is

$$\min \quad \phi_{v,\,k}(d_v) \tag{2.1}$$

$$\text{s. t.} \quad \|d_v\|_2 \leq \Delta_{v,\,k} \tag{2.2}$$

$$(x^{(k)})_v + d_v \in X_v \ , \tag{2.3}$$

where $\phi_{v,\,k}(d_v)$ is an approximation to the utility function $u_v((x^{(k)})_v + d_v,\, (\mathbf{x}^{(\mathbf{k})})^{-v})$, $d_v \in \Re^{n_v}$.

Define

$$g_{v,\,k} = \frac{\partial u_v(x)}{\partial x_v}|_{x=x^{(k)}}, \tag{2.4}$$

A natural choice for $\phi_{v,\,k}(d_v)$ is the second order model

$$\phi_{v,\,k}(d_v) = u_v(x^k) + d_v^T g_{v,\,k} + \frac{1}{2} d_v^T B_{v,\,k} d_v, \quad d_v \in \Re^{n_v}, \tag{2.5}$$

where $B_{v,\,k} \in \Re^{n_v \times n_v}$ is a symmetric matrix which approximates the Hessian matrix $\frac{\partial^2 u_v(x)}{\partial x_v^2}|_{x=x^{(k)}}$. Let $d_{v,\,k}$ be the solution of problem (2.1). We denote the predicted reduction of the v-th utility function by

$$Pred_{v,\,k} = \phi_{v,\,k}(0) - \phi_{v,\,k}(d_{v,\,k}), \tag{2.6}$$

and the actual reduction by

$$Ared_{v,\,k} = u_v((x^{(k)})_v,\, (\mathbf{x}^{(\mathbf{k})})^{-v}) - u_v((x^{(k)})_v + d_{v,\,k},\, (\mathbf{x}^{(\mathbf{k})})^{-v}). \tag{2.7}$$

The ratio between these two reductions plays an important role in a trust region algorithm. Letting

$$r_{v,\,k} = \frac{Ared_{v,\,k}}{Pred_{v,\,k}}, \tag{2.8}$$

we define

$$(x^{(k+1)})_v = \begin{cases} (x^{(k)})_v + d_{v,k}, & \text{if } r_{v,k} > 0, \\ (x^{(k)})_v & \text{otherwise.} \end{cases} \tag{2.9}$$

The above definition of the new iteration point x_{k+1} is based on the reduction of the utility function of each player, but, when an overall step is taken, we can have $u_v(x^{(k+1)}) > u_v(x^{(k)})$ even when $r_{v,k} > 0$, because of the contributions to $x^{(k+1)}$ from the other players. Hence we need to use criteria other than $r_{v,k}$ to update the trust region bounds $\Delta_{v,k}$. One natural merit function is $V(x)$ defined in (1.6), but the calculation of $V(x)$ requires the solutions of N minimization problems. Thus, we have to find an easy to compute merit function.

Define Π_{X_v} to be the projection mapping from $\Re^{n_v}$ onto X_v, namely

$$\Pi_{X_v}(y) = \text{argmin}_{x\in X_v}||x - y||_2, \quad \forall\, y \in \Re^{n_v}. \tag{2.10}$$

Define the vector function

$$F(x) = \begin{pmatrix} \Pi_{X_1}\left(x_1 - \frac{\partial u_1(x)}{\partial x_1}\right) - x_1 \\ \Pi_{X_2}\left(x_2 - \frac{\partial u_2(x)}{\partial x_2}\right) - x_2 \\ \vdots \\ \Pi_{X_N}\left(x_N - \frac{\partial u_N(x)}{\partial x_N}\right) - x_N \end{pmatrix} \in \Re^n \;, \tag{2.11}$$

and the merit function

$$\psi(x) = \|F(x)\|_G^2 = \sum_{v=1}^{N} \left\| \Pi_{X_v}\left(x - \frac{\partial u_v(x)}{\partial x_v}\right) - x \right\|_{G_v}^2 \;, \tag{2.12}$$

where $G = Diag(G_1, G_2, ..., G_N)$ with $G_v \in \Re^{n_v \times n_v}$ $(v = 1, ..., N)$ being fixed symmetric positive definite matrices, and the norm $\|g\|_G$ denotes $\sqrt{g^T G g}$. Because the step definition (2.9) does not ensure a reduction of the merit function, it seems reasonable for us to reduce the trust region bounds if no sufficient reduction has been achieved for the best merit function up to the current iteration. Let

$$\eta_k = \min_{1\le i\le k} \psi(x_i)\,, \tag{2.13}$$

and

$$Pred_k = \sum_{v=1}^{N} Pred_{v,k}\,. \tag{2.14}$$

We define the ratio

$$\rho_k = \frac{\eta_k - \psi(x_{k+1})}{Pred_k}\,, \tag{2.15}$$

and we we always reduce all the trust region bounds $\Delta_{v,k}$ unless

$$\rho_k \ge \beta_1\,, \tag{2.16}$$

where $\beta_1 \in (0,1)$ is a constant, We let the trust region bounds have the form

$$\Delta_{v,k} = \frac{1}{\tau_v + t_{v,k}}\|\hat{g}_{v,k}\|_2 \quad (v = 1, 2, ..., N), \tag{2.17}$$

for some positive constant $\tau_v > 0$, where $t_{v,k} > 0$ is updated from iteration to iteration and

$$\hat{g}_{v,k} = (x^{(k)})_v - \Pi_{X_v}((x^{(k)})_v - g_{v,k}) \,. \tag{2.18}$$

It is easy to see that $\hat{g}_{v,k} = g_{v,k}$ if both $(x^{(k)})_v$ and $(x^{(k)})_v - g_{v,k}$ are in X_v. We call $\hat{g}_{v,k}$ the projected partial derivative. We adjust the ratio between the trust region bound $\Delta_{v,k}$ and the normal of projected partial derivative $\|\hat{g}_{v,k}\|_2$ instead of the trust region bound itself. Further, the ratio between $\Delta_{v,k}$ and $\|\hat{g}_{v,k}\|_2$ is not be increased unless both (2.16) holds and $r_{v,k}$ is larger than certain small positive number. Specifically, for each $v \in \{1, 2, ..., N\}$, we let

$$t_{v,k+1} = t_{v,k} + \delta_v \quad (v = 1, 2, ..., N), \tag{2.19}$$

if (2.16) fails, where each δ_v is a positive constant. Alternatively, if inequality (2.16) holds, we let

$$t_{v,k+1} = \begin{cases} \max[t_{v,k} - \delta_v,\, 0], & \text{if } r_{v,k} \geq \beta_2; \\ t_{v,k}, & \text{if } r_{v,k} \in (0, \beta_2]; \\ t_{v,k} + \delta_v, & \text{if } r_{v,k} \leq 0, \end{cases} \tag{2.20}$$

where $\beta_2 \in (0, 1)$ is a constant. On the other hand, in a classical trust region algorithm for nonlinear unconstrained optimization, if a reduction in the trust region bound is needed, the trust region bound will be reduced by a fraction (say, a half). The idea for the classic approach is to force a bound on the sum of the trial steps that do not give a sufficient reduction in the objective function in order to ensure the iterates converging to a stationary point. The update formulae (2.19) and (2.20) imply that, even the trust region bound is reduced at every iteration, we have

$$\sum_{k=1}^{\infty} \Delta_{v,k} = \infty \,, \tag{2.21}$$

provided that $\|\hat{g}_{v,k}\|_2$ is bounded away from zero.

Algorithm 2.1. *(A Jocobi-Type Trust Region Algorithm For NEP)*

Step 1 *Given an initial feasible point* $x^{(1)} \in X_1 \times X_2 \cdots \times X_N$.
Choose positive constants $\tau_v,\ \delta_v,\ t_{v,1}$ $(v = 1, 2, ..., N)$*;*
Choose $\beta_1 \in (0, 1)$, $\beta_2 \in (0, 1)$. *Set* $k := 1$.

Step 2 *If* $\sum_{v=1}^{N} \|\hat{g}_{v,k}\|_2 = 0$ *then stop.*
Solve (2.1) obtaining $d_{v,k}$ *for all* $v = 1, ..., N$.

Step 3 *Compute* $r_{v,k}$ *by (2.8), and define the next iterate point* $x^{(k+1)}$ *by (2.9).*

Step 4 *If (2.16) holds define* $t_{v,k+1}$ *by (2.20) otherwise by (2.19).*

Step 5 *Generate* $B_{v,k+1}$. *Set* $k := k + 1$. *Go to Step 2.*

The algorithm terminates if an iterate $x^{(k)}$ satisfies

$$\hat{g}_{v,k} = 0, \quad \forall v = 1, 2, ..., N. \tag{2.22}$$

We define such a point as a stationary point.

Definition 2.1. *We call x^* is a stationary point of the Nash equilibrium problem if x^* satisfies that*

$$F(x^*) = 0, \tag{2.23}$$

where $F(x)$ is defined by (2.11).

The solution of the Nash equilibrium problem is also a stationary point. If, for each v, the utility function $u_v(x_v, \mathbf{x}^{-v})$ as a function x_v is convex, a stationary point is also the solution of NEP.

In the next section, we show that under certain conditions, the iterates generated by the above algorithm are not bounded away from stationary points.

3 Convergence Properties

One nice property of trust region methods for nonlinear optimization is that the predicted reduction of the trust region subproblem can be estimated by the residual of the optimality conditions. Here, we establish a similar result for the trust region subprolems for the NEP.

First we need the following simple result.

Lemma 3.1. *Assume that $(x^{(k)})_v \in X_v$, we have that*

$$g_{v,\,k}^T \hat{g}_{v,\,k} \geq \|\hat{g}_{v,\,k}\|_2^2. \tag{3.1}$$

Proof. For any $y \in X_v$, it follows from the definition of the projection Π_{X_v} that

$$(y - \Pi_{X_v}(y))^T (x - \Pi_{X_v}(y)) \leq 0, \quad \forall x \in X_v. \tag{3.2}$$

Let $y = (x^{(k)})_v - g_{v,\,k}$ and $x = (x^{(k)})_v$ in the above inequality, we obtain that

$$\left((x^{(k)})_v - g_{v,\,k} - \Pi_{X_v}((x^{(k)})_v - g_{v,\,k})\right)^T \left((x^{(k)})_v - \Pi_{X_v}((x^{(k)})_v - g_{v,\,k})\right) \leq 0, \tag{3.3}$$

which can be rewritten as

$$g_{v,\,k}^T \left((x^{(k)})_v - \Pi_{X_v}((x^{(k)})_v - g_{v,\,k})\right) \geq \left\|(x^{(k)})_v - \Pi_{X_v}((x^{(k)})_v - g_{v,\,k})\right\|_2^2. \tag{3.4}$$

This shows that the lemma is true. □

In order to analyze the convergence properties of our algorithm given in the previous section, we need the following fundamental result which is an extension of a result given by Powell [21].

Lemma 3.2. *Let $d_{v,\,k}$ be the solution of subproblem (2.1). Expression (2.4) has the property*

$$Pred_{v,\,k} \geq \frac{1}{2}\|\hat{g}_{v,\,k}\|_2 \min\left[\Delta_{v,\,k},\, \frac{\|\hat{g}_{v,\,k}\|_2}{\max[1,\, \|B_{v,\,k}\|_2]}\right]. \tag{3.5}$$

Proof. From the previous lemma and the definition of $d_{v,\,k}$, we have

$$\begin{aligned}
\phi_{v,\,k}(d_{v,\,k}) &= \min_{\|d_v\|_2 \leq \Delta_{v,\,k},\, (x^{(k)})_v + d_v \in X_v} \phi_{v,\,k}(d_v) \\
&\leq \min_{\alpha \in (0,1], \|\alpha \hat{g}_{v,\,k}\|_2 \leq \Delta_{v,\,k}} \phi_{v,\,k}(-\alpha \hat{g}_{v,\,k}) \\
&\leq \min_{\alpha \in (0,1], \|\alpha \hat{g}_{v,\,k}\|_2 \leq \Delta_{v,\,k}} \left[u_v((x^{(k)})_v) - \alpha\|\hat{g}_{v,\,k}\|_2^2 + \frac{1}{2}\alpha^2 \|B_{v,\,k}\|_2 \|\hat{g}_{v,\,k}\|_2^2\right].
\end{aligned} \tag{3.6}$$

If $\Delta_{v,k} \le \|\hat{g}_{v,k}\|_2$, it follows from Powell [21] and the above inequality that

$$Pred_{v,k} \ge \frac{1}{2}\|\hat{g}_{v,k}\|_2 \min\left[\Delta_{v,k},\ \frac{\|\hat{g}_{v,k}\|_2}{\|B_{v,k}\|_2}\right], \tag{3.7}$$

which implies (3.5). Now we assume that $\Delta_{v,k} > \|\hat{g}_{v,k}\|_2$. In this case, we have that

$$\begin{aligned} Pred_{v,k} &\ge \phi_{v,k}(0) - \phi_{v,k}\left(-\frac{1}{\max[1, \|B_{v,k}\|_2]}\hat{g}_{v,k}\right) \\ &= \frac{1}{\max[1, \|B_{v,k}\|_2]}\left(1 - \frac{\|B_{v,k}\|_2}{2\max[1, \|B_{v,k}\|_2]}\right)\|\hat{g}_{v,k}\|_2^2 \\ &\ge \frac{1}{2}\frac{\|\hat{g}_{v,k}\|_2^2}{\max[1,\ \|B_{v,k}\|_2]}. \end{aligned} \tag{3.8}$$

This completes our proof. □

Similar to trust region algorithms for unconstrained optimization problems, we do not need to solve the subproblem (2.1) exactly as long as we can compute an approximate solution $d_{v,k}$ that satisfies a weaker form of condition (3.5). Specifically, it is quite common for some practical trust region algorithms to solve the trust region subproblem (2.1) approximately so that $Pred_{v,k}$ satisfies the condition

$$Pred_{v,k} \ge \beta_3\|\hat{g}_{v,k}\|_2 \min\left[\Delta_{v,k},\ \frac{\|\hat{g}_{v,k}\|_2}{1 + \|B_{v,k}\|_2}\right], \tag{3.9}$$

where $\beta_3 \in (0,\ 0.5]$ is a constant. This inequality implies the following result:

Corollary 3.1. *Let $d_{v,k}$ be an inexact solution of subproblem (2.1) such that (3.9) holds. Then we have the bound*

$$Pred_{v,k} \ge \beta_3\|\hat{g}_{v,k}\|_2^2 \min\left[\frac{1}{\tau_v + t_{v,k}},\ \frac{1}{1 + \|B_{v,k}\|_2}\right]. \tag{3.10}$$

Lemma 3.3. *Let $d_{v,k}$ be an inexact solution of (2.1) such that (3.9) holds. If $B_{v,k}$ and $\frac{\partial^2 u_v(x)}{\partial x_v^2}$ are uniformly bounded and if*

$$\lim_{k\to\infty} t_{v,k} \to \infty\,, \tag{3.11}$$

then

$$\lim_{k\to\infty} r_{v,k} = 1\,. \tag{3.12}$$

Proof. Since $B_{v,k}$ are bounded uniformly, it follows from (2.17) and (3.10) that

$$Pred_{v,k} \ge \beta_3\|\hat{g}_{v,k}\|_2^2\frac{1}{\tau_v + t_{v,k}} = \beta_3\|\hat{g}_{v,k}\|_2\Delta_{v,k} = \beta_3(\tau_v + t_{v,k})\Delta_{v,k}^2\,, \tag{3.13}$$

for sufficiently large k. The uniformly boundedness of $B_{v,k}$ and $\frac{\partial^2 u_v(x)}{\partial x_v^2}$ imply that

$$|Ared_{v,k} - Pred_{v,k}| = O(\Delta_{v,k}^2)\,. \tag{3.14}$$

Therefore, from (3.11), (3.13) and (3.14) we deduce that (3.12) is true. □

We make the following assumptions:

Assumption 3.1. *Let $\{x^{(k)},\, k = 1, 2, ...\}$ be generated by our algorithm. We assume that*

1. *There exists a bounded convex set $\mathcal{S} \subset \Re^n$ such that $x^{(k)} \in \mathcal{S}$ for all k,*
2. *The functions $u_v(x)$ are continuously differentiable in $\mathcal{S}$ for all v. Furthermore, $\frac{\partial^2 u_v(x)}{\partial x_v^2}$ are uniformly bounded for all $x \in \mathcal{S}$.*
3. *$B_{v,\,k}$ are bounded uniformly for all v and all k.*

Now, we are going to show that under certain conditions the points $\{x_k\}$ generated by our method are not bounded away from stationary points.

Assumption 3.1 implies that there exists a positive number β_4 such that

$$\left\| \frac{\partial^2 u_v(x)}{\partial x_v^2} \right\|_2 \le \beta_4 \,, \quad \forall\, x \in \mathcal{S}, \quad v = 1, 2, ..., N, \tag{3.15}$$

and

$$1 + \|B_{v,\,k}\|_2 \le \beta_4 \,, \quad \forall\, k = 1, 2, ..., \quad v = 1, 2, ..., N. \tag{3.16}$$

Lemma 3.4. *Assume that all the conditions in Assumption 3.1 are satisfied, and that every $d_{v,\,k}$ satisfies (3.9). Then formulae (2.9) provides*

$$(x_{k+1})_v = (x_k)_v + d_{v,\,k}, \tag{3.17}$$

if

$$\Delta_{v,\,k} \le \frac{\beta_3}{2\beta_4} \|\hat{g}_{v,\,k}\|_2 \,, \tag{3.18}$$

where β_4 is the positive constant in (3.15) and (3.16).

Proof. Because $\Delta_{v,\,k} \le \beta_3 \|\hat{g}_{v,\,k}\|_2/(2\beta_4) \le \|\hat{g}_{v,\,k}\|_2/\beta_4$, it follows from (3.9) and (3.16) that

$$Pred_{v,\,k} \ge \beta_3 \|\hat{g}_{v,\,k}\|_2 \Delta_{v,\,k} \,. \tag{3.19}$$

On the other hand, the agrement between the first order terms of $Ared_{v,\,k}$ and $Pred_{v,\,k}$ provides

$$|Ared_{v,\,k} - Pred_{v,\,k}| \le \beta_4 \|d_{v,\,k}\|_2^2 \le \beta_4 \Delta_{v,\,k}^2. \tag{3.20}$$

Inequalities (3.18)-(3.20) imply

$$|1 - r_{v,\,k}| \le \frac{\beta_4}{\beta_3} \frac{\Delta_{v,\,k}}{\|\hat{g}_{v,\,k}\|_2} \le \frac{1}{2}, \tag{3.21}$$

so $r_{v,\,k}$ is at least 0.5 in formula (2.9). This completes our proof. □

The projected Cauchy step is the best point along the negative projected directional derivative direction within the trust region ball. Namely the Cauchy step $d_{v,\,k}^C = -\alpha_{v,\,k}^* \hat{g}_{v,\,k}$ is define by

$$\phi_{v,\,k}(d_{v,\,k}^C) = \min_{\alpha > 0, \|\alpha \hat{g}_{v,\,k}\|_2 \le \Delta_{v,\,k}} \phi_{v,\,k}(-\alpha \hat{g}_{v,\,k},). \tag{3.22}$$

We also assume that the computed trial step $d_{v,\,k}$ is not worse than the projected Cauchy Step, which is the condition:

Assumption 3.2. *Assume that every trial step $d_{v,k}$ generated by our method has the property*

$$\phi_{v,k}(d_{v,k}) \le \phi_{v,k}(d^C_{v,k}). \tag{3.23}$$

Now we can establish our main global convergence result as follows.

Theorem 3.1. *Denote the Jacobi matrix of $F(x)$ by $J(x)$, if $Diag(\delta_1 G_1, \delta_2 G_2, ..., \delta_N G_n)J(x)$ is uniformly positive definite for all $x \in \mathcal{S}$, and if all the conditions in Assumptions 3.1 and 3.2 are satisfied, then the iterate points $\{x_k\}$ generated by our method are not bounded away from stationary points, namely either $\psi(x_k) = 0$ for some k or*

$$\lim_{k\to\infty} \eta_k = 0\,. \tag{3.24}$$

Proof We prove the theorem by contradictions. If the theorem is not true, there exists a positive constant β_5 such that

$$\psi(x_k) \ge \beta_5 \tag{3.25}$$

for all k. Thus, there is another positive constant β_6 such that

$$\sum_{v=1}^{N} \|\hat{g}_{v,k}\|_2^2 \ge \beta_6\,, \tag{3.26}$$

holds for all k.

First, we assume that there are only finitely many k such that $\rho_k \ge \beta_1$. In this case, there exists an integer $\bar{k}$ such that

$$t_{v,k} = t_{v,\bar{k}} + \delta_v(k-\bar{k}) \quad \forall\, k \ge \bar{k},\ v = 1, 2, ..., N. \tag{3.27}$$

The above relation gives that

$$\Delta_{v,k} = \frac{\|\hat{g}_{v,k}\|_2}{\tau_{v,\bar{k}} + \delta_v(k-\bar{k})} = \frac{\|\hat{g}_{v,k}\|_2}{\delta_v k} + O\left(\frac{\|\hat{g}_{v,k}\|_2}{k^2}\right) \to 0\,. \tag{3.28}$$

Thus, Lemma 3.3 implies that there exists $\hat{k} \ge \bar{k}$ such that (3.17) holds for all $k \ge \hat{k}$ and all $v = 1, 2, ..., N$. Moreover, for sufficiently large k, it follows from (3.27) that

$$d^C_{v,k} = -\frac{\hat{g}_{v,k}}{\tau_{v,\bar{k}} + \delta_v(k-\bar{k})} \tag{3.29}$$

and

$$\phi_{v,k}(d^C_{v,k}) = -\frac{\|\hat{g}_{v,k}\|_2^2}{\tau_{v,\bar{k}} + \delta_v(k-\bar{k})} + O(\Delta^2_{v,k}) = -\frac{\|\hat{g}_{v,k}\|_2^2}{\delta_v k} + O\left(\frac{\|\hat{g}_{v,k}\|_2^2}{k^2}\right). \tag{3.30}$$

Our assumption $\phi_{v,k}(d_{v,k}) \le \phi_{v,k}(d^C_{v,k})$ implies that

$$\hat{g}^T_{v,k} d_{v,k} + O(\Delta^2_{v,k}) \le -\frac{\|\hat{g}_{v,k}\|_2^2}{\delta_v k} + O\left(\frac{\|\hat{g}_{v,k}\|_2^2}{k^2}\right). \tag{3.31}$$

which gives

$$\hat{g}^T_{v,k} d_{v,k} \le -\frac{\|\hat{g}_{v,k}\|_2^2}{\delta_v k} + O\left(\frac{\|\hat{g}_{v,k}\|_2^2}{k^2}\right). \tag{3.32}$$

Due to the fact that $\|d_{v,k}\|_2 \le \Delta_{v,k}$, we have that

$$d_{v,k} = -\frac{\hat{g}_{v,k}}{\delta_v k} + O\left(\frac{\|\hat{g}_{v,k}\|_2}{k^2}\right). \tag{3.33}$$

This relation can be rewritten as

$$\begin{aligned} F(x_k) &= -kDiag(\delta_1 I, \delta_2 I, ..., \delta_N I)d_k + O\left(\frac{\|\hat{g}_{v,k}\|_2}{k}\right) \\ &= -kDiag(\delta_1 I, \delta_2 I, ..., \delta_N I)d_k + O(\|d_k\|_2). \end{aligned} \tag{3.34}$$

Equation (3.33) also implies that

$$Pred_{v,k} = -\frac{\|\hat{g}_{v,k}\|_2^2}{\delta_v k} + O(\|d_{v,k}\|_2^2), \tag{3.35}$$

and

$$k\|d_{v,k}\|_2^2 = \frac{\|\hat{g}_{v,k}\|_2^2}{\delta_v^2 k} + O\left(\frac{\|\hat{g}_{v,k}\|_2^2}{k^2}\right). \tag{3.36}$$

The above two relations yield that

$$\begin{aligned} k\|d_k\|_2^2 &\ge \frac{1}{\max_{1\le v\le N}\delta_v} Pred_k + O(\|d_k\|_2^2) \\ &\ge \frac{\sum_{v=1}^N \|\hat{g}_{v,k}\|_2^2}{k\max_{1\le v\le N}\delta_v^2} + O(\|d_k\|_2^2). \end{aligned} \tag{3.37}$$

On the other hand, under the assumption on $J(x)$ in the theorem, there exists a positive constant β_7 such that

$$d^T Diag(\delta_1 G_1, \delta_2 G_2, ..., \delta_N G_n)J(x)d \ge \beta_7\|d\|_2^2 \tag{3.38}$$

holds for all $x \in \mathcal{S}$ and all $d \in \Re^n$. Thus, from (3.34) and (3.38) we can obtain that

$$\begin{aligned} \psi(x_{k+1}) &= \|F(x_{k+1})\|_G^2 = \|F(x_k + d_k)\|_G^2 \\ &= \|F(x_k) + J(x_k)d_k\|_G^2 + O(\|d_k\|_2^2) \\ &= \|F(x_k)\|_G^2 + 2F(x_k)^T GJ(x_k)d_k + O(\|d_k\|_2^2) \\ &= \psi(x_k) - 2kd_k^T Diag(\delta_1 G_1, \delta_2 G_2, ..., \delta_N G_n)J(x_k)d_k + O(\|d_k\|_2^2) \\ &\le \psi(x_k) - 2k\beta_7\|d_k\|_2^2 + O(\|d_k\|_2^2) \end{aligned} \tag{3.39}$$

for all large k. Inequalities (3.39) and (3.37) show that

$$\psi(x_{k+1}) \le \psi(x_k) - \frac{\beta_7 \sum_{v=1}^N \|\hat{g}_{v,k}\|_2^2}{k\max_{1\le v\le N}\delta_v^2}, \tag{3.40}$$

for all sufficiently large k, which is impossible because of (3.26) and the facts that $\beta_7 > 0$ and $\delta_v > 0$ for all v. The contraction shows that (3.24) is true if there are only finitely many k such that $\rho_k \ge \beta_1$.

To complete our proof, we now assume there are infinitely many k such that $\rho_k \ge \beta_1$. Define the set

$$\mathcal{I} = \{k \mid \quad \rho_k \ge \beta_1\}. \tag{3.41}$$

The definition implies that

$$\eta_k - \eta_{k+1} \geq \beta_1 Pred_k, \tag{3.42}$$

Thus

$$\sum_{k\in\mathcal{I}} Pred_k < +\infty, \tag{3.43}$$

which in turns shows that

$$\sum_{k\in\mathcal{I}} Pred_{v,\,k} < +\infty, \tag{3.44}$$

for all $v = 1, 2, ..., N$. Because $\Delta_{v,\,k} \leq \frac{1}{\tau_v}\|\hat{g}_{v,\,k}\|_2$, we can see that

$$\sum_{k\in\mathcal{I}} \Delta_{v,\,k}\|\hat{g}_{v,\,k}\|_2 < +\infty, \tag{3.45}$$

for all $v = 1, 2, ..., N$. Therefore

$$\sum_{k\in\mathcal{I}} \frac{\|\hat{g}_{v,\,k}\|_2^2}{\tau_v + t_{v,\,k}} < +\infty\,, \quad v = 1, 2, ..., N. \tag{3.46}$$

Thus,

$$\sum_{k\in\mathcal{I}}\sum_{v=1}^{N} \frac{\|\hat{g}_{v,\,k}\|_2^2}{\tau_v + t_{v,\,k}} < +\infty\,, \tag{3.47}$$

which gives that

$$\sum_{k\in\mathcal{I}} \frac{1}{\max_{1\leq v\leq N}[\tau_v + t_{v,\,k}]} \sum_{1\leq v\leq N} \|\hat{g}_{v,\,k}\|_2^2 < +\infty\,. \tag{3.48}$$

The above inequality and (3.26) imply that

$$\sum_{k\in\mathcal{I}} \frac{\beta_6}{\max_{1\leq v\leq N}[\tau_v + t_{v,1}] + k\max_{1\leq v\leq N}\delta_v} < +\infty\,. \tag{3.49}$$

Therefore, it follows that

$$\sum_{k\in\mathcal{I}} \frac{1}{k} < \infty\,. \tag{3.50}$$

Let $I_k = |\mathcal{I} \cap \{i\,|\,1 \leq i \leq k\}|$ the be number of indices that in $\mathcal{I}$ which is not greater than k. Due to relation (3.50), we have that

$$\lim_{k\to\infty} \frac{I_k}{k} = 0\,. \tag{3.51}$$

This shows that

$$\lim_{k\to\infty} \frac{t_{v,\,k}}{k} = \delta_v. \tag{3.52}$$

Thus, we have

$$\Delta_{v,\,k} = \frac{\|\hat{g}_{v,\,k}\|_2}{\delta_v k} + o\left(\frac{\|\hat{g}_{v,\,k}\|_2}{k}\right)\,. \tag{3.53}$$

Now, the above relation is very similar to (3.28). Hence, similar to the first part of our proof, we can also derive that (3.40) holds for all large k. This is a contradiction because $\psi(x_k)$ is bounded below. Thus, we see that the theorem is true. □

For the special case when $J(x)$ is positive definite, we can choose $G = I$ and $\delta_1 = \delta_2 = ... = \delta_N$. Then, we can easily see that the convergence result follows.

Corollary 3.2. *If $J(x)$ is uniformly positive definite for all $x \in \mathcal{S}$, and if all the conditions in Assumptions 3.1 and 3.2 are satisfied, we can let $G_v = I$ for all v and let $\delta_v = \delta$ for all v so that the iterate points $\{x_k\}$ generated by our method are not bounded away from stationary points.*

4 Discussion

We have given a Jocobi type trust region algorithm for Nash Equilibrium Problems. One natural extension is to consider a Gauss-Seidel type of trust region method. The obvious difficult is that the each player v will have their individual up to date point for defining the merit function to adjust the trust region, which would make the adjusting of trust regions more complicated.

Another possible extension of the current work is to applying trust region to the general equilibrium problem[19], namely finding $x^* \in C$ such that

$$f(x^*, y) \geq 0\,, \quad \forall\, y \in C\,, \tag{4.1}$$

where C is a nonempty closed convex subset of R^n and $f(x, y)$ is a continuous function from $C \times C$ to R with the following properties: 1) $f(x, .)$ is convex on C for all $x \in C$ and $f(x, x) = 0$ for all $x \in C$.

Acknowledgement

This work was carried out while I was visited University of Cambridge. I would like to thank Professor M.J.D. Powell and Professor A. Iserles for their kindly hosts. I am also grateful to Professor Masao Fukushiman and Professor J.S. Pang for their comments on an earlier version of the paper.

References

[1] R.H. Byrd, J.C. Gilbert and J. Nocedal, A trust region method based on interior point techniques for nonlinear programming, *Math. Program.* 89(2000) 149-185.

[2] R.H. Byrd, R.B. Schnabel and G.A. Shultz, A trust region algorithm for nonlinear constrained optimization, *SIAM J. Optimization* 24(1987) 1152-1170.

[3] M.R. Celis, J.E. Dennis and R.A. Tapia, A trust region strategy for nonlinear equality constrained optimization, in: P.T. Boggs, R.H. Byrd and R.B. Schnabel, eds., *Nuermical Optimization 1984*, (SIAM, Philadelphia, 1984) pp.71-82.

[4] X.J. Chen and M. Fukushima, A smoothing method for a mathematical program with P-Matrix linear complementarity constraints, *Computational Optimization and Applications* 27(2004) 223-246.

[5] A.R. Conn, N.I.M. Gould, D. Orban and Ph.L. Toint, A primal-dual trust-region algorithm for non-convex nonlinear programming, *Math. Program.* 87(2000) 215-249.

[6] A.R. Conn, N.I.M. Gould and Ph.L. Toint, Global convergence of a class of trust region algoirthms for optimization with simple bunds, *SIAM J. Optimization*, 25(1988) 433-460.

[7] A.R. Conn, N.I.M. Gould and Ph.L. Toint, *Turst Region Methods*, MPS/SIAM Series on Optimization, SIAM, Philadelphia, 2000.

[8] F. Facchinei and Ch. Kanzow, Generalized Nash equilibrium problems, *4OR* 5(2007) 173-210.

[9] F. Facchinei and J.S. Pang, *Finite-Dimensional Variational Inequality and Complementarity Problems, Vols. 1 and 2*, Springer, 2003.

[10] R. Fletcher, *Practical Methods of Optimization, Vol 1. Unconstraind Optimization*, (John Wiley, New York, 1980).

[11] R. Fletcher, A model algorithm for composite NDO problems, *Math. Program. Studies* 17(1982) 67-76.

[12] M. Fukushima, Z.Q. Luo and J.S. Pang, A global convergent sequential quadratic programmking algorithm for mathematical programs with linear complementarity constraints, *Computational Optimization and Applications* 10(1998) 5-34.

[13] M. Fukushima and J.S. Pang, Some feasibility issues in mathematical programs with equilibrium constraints, *SIAM J. Optimization* 8(1998) 673-681.

[14] S. Govindan and R. Wilson, A global method method to compute Nash equilibria, *J. Economic Theory* 110(2003) 65-86.

[15] X.W. Liu, G. Perakis and J. Sun, A robust SQP method for mathematical programs with linear complementarity constraints, *Computational Optimization and Application*, 34(2006) 5-33.

[16] M-A Majig and M. Fukushima, Restricted-step Josephy-Newton method for general variational inequalities with polyhedral constraints, Report, Kyoto University, 2008.

[17] R.B. Myerson, Nash equilibrium and the history of economic theory, *J. Enonomic Literature*, 37(1999) 1067-1082.

[18] J. Nash, Equilibrium points in n-person games, *Proceedings of the National Academy of Sciences of the United State of America* 36(1950) 48-49.

[19] T.T.V. Nguyen, J.J. Strodiot and V.H. Nguyen, A boundle method for solving equilibrium problems, *Math. Program. Ser. B* 116(2009) 529-552.

[20] J.S. Pang, G.Scutari, F. Facchinei and C.X. Wang, Distributed power allocation with rate constraints in Gaussian parallel interference channels, Research Report, 2007. Submitted to *IEEE Transsctions on Information Theory*.

[21] M.J.D. Powell, A new algorithm for unconstrained optimization, in; J.B. Rosen, O.L. Mangassarian and Ritter, eds., *Nonlinear Programming* (Academic Press, New York, 1970).

[22] M.J.D. Powell, On the global convergence of trust region algorithms for unconstrained minimization, *Math. Program.* 29(1984) 296-303.

[23] M.J.D. Powell and Y. Yuan, A trust region algorithm for equality constrained optimization, *Math. Program.* 49(1990) 189-211.

[24] Ph.L. Toint, Non-monotone trust region algorithms for nonlinear optimization subject to convex constraints, *Math. Program.* 77(1997) 69-94.