A NOTE ON GLOBALLY CONVERGENT NEWTON METHOD FOR STRONGLY MONOTONE VARIATIONAL INEQUALITIES

Kouichi Taji
Nagoya University

Abstract   Newton's method for solving variational inequalities is known to be locally quadratically convergent. By incorporating a line search strategy for the regularized gap function, Taji et al. (Mathematical Programming, 1993) have proposed a modification of a Newton's method which is globally convergent and whose rate of convergence is quadratic. But the quadratic convergence has been shown only under the assumptions that the constraint set is polyhedral convex and the strict complementarity condition holds at the solution. In this paper, we show that the quadratic rate of convergence is also achieved without both the polyhedral convex assumption and the strict complementarity condition. Moreover, the line search procedure is simplified.

Keywords: Nonlinear programming, variational inequality, Newton's method, regularized gap function

1. Introduction

The variational inequality problem, denoted by VIP, is to find a vector $x^* \in S$ such that

$$\langle F(x^*), x - x^* \rangle \geq 0 \text{ for all } x \in S, \tag{1.1}$$

where $F : \Re^n \to \Re^n$ is a mapping and S is a nonempty closed convex subset of $\Re^n$. The symbol $\langle \cdot, \cdot \rangle$ denotes the inner product in $\Re^n$. The variational inequality problem is theoretically and practically useful, and has been used to study and formulate various equilibrium problems arising in economics and engineerings [2].

Newton's method, applied to VIPs, generates a sequence $\{x^k\}$, where $x^{k+1} \in S$ is a solution to the linearized VIP

$$\left\langle F(x^k) + \nabla F(x^k)^T(x^{k+1} - x^k), x - x^{k+1} \right\rangle \geq 0 \text{ for all } x \in S. \tag{1.2}$$

It is well known that, under suitable assumptions, Newton's method converges to a solution quadratically, provided that an initial point x^0 is chosen to be sufficiently close to a solution [2, 4]. Also, by incorporating line search to a merit function, Newton's method has shown to be modified to have globally convergent property [3, 7]. In particular, the globally convergent Newton method (GNEW), proposed by Taji et al.[7], makes use of the regularized gap function proposed by Fukushima [1] as a merit function and allows inexact line search of Armijo type. While, under the strongly monotone assumption, the method has been shown to be globally convergent, the quadratic rate of convergence has been shown under limited assumptions that the constraint set S is polyhedral convex and that the strict complementarity holds at a solution. These two assumptions are not necessary for quadratic convergence of the naive Newton's method (1.2).

In this paper, we show that the quadratic convergence of GNEW is established without both the polyhedral convex constraint and the strict complementarity. Furthermore, though the method originally uses two-step line search procedure to obtain quadratic convergence, we also show that a slightly simplified line search procedure is sufficient for both global and quadratic convergence.

We summarize several notations frequently used in this paper. The symbol $\langle \cdot, \cdot \rangle$ denotes the *inner product* in $\Re^n$, and $\|\cdot\|$ denotes the *Euclidean norm* in $\Re^n$ defined by $\|x\| = \sqrt{\langle x, x \rangle}$ or the induced matrix norm. For an $n \times n$ symmetric positive definite matrix G, we define the *G-norm* by $\|x\|_G = \sqrt{\langle x, Gx \rangle}$. $\text{Proj}_{S,G}(x)$ represents the *projection under the G-norm onto S*, which is the unique solution to the problem

$$\text{minimize} \quad \|y - x\|_G \quad \text{subject to } y \in S.$$

2. Preliminaries

In this section, we first introduce the regularized gap function and summarize its basic properties. Then we present a globally convergent Newton method for solving the variational inequality problem (1.1) and its convergent properties. We also introduce the strict complementarity condition.

For an arbitrarily chosen positive definite symmetric $n \times n$ matrix G, the regularized gap function, introduced by Fukushima [1], $f : \Re^n \to \Re$ for a variational inequality problem (1.1) is defined as

$$f(x) = \max \left\{ -\langle F(x), y - x \rangle - \frac{1}{2} \langle y - x, G(y - x) \rangle \; \middle| \; y \in S \right\} \tag{2.1}$$

$$= -\langle F(x), H(x) - x \rangle - \frac{1}{2} \langle H(x) - x, G(H(x) - x) \rangle, \tag{2.2}$$

where the mapping $H : \Re^n \to \Re^n$ is defined by

$$H(x) = \text{Proj}_{S,G}(x - G^{-1}F(x)). \tag{2.3}$$

We note that, by the positive definiteness of G, the maximum in (2.1) is uniquely attained at $y = H(x)$. It is known that the mapping H is nonexpansive, and x is a solution to (1.1) if and only if $H(x) = x$.

It is known [1] that $f(x) \geq 0$ for all $x \in S$, and that $f(x) = 0$ and $x \in S$ if and only if x is a solution to (1.1). Hence the problem (1.1) is equivalent to find a solution x^* to the optimization problem:

$$\text{minimize} \quad f(x) \quad \text{subject to} \quad x \in S$$

with $f(x^*) = 0$. It is also known [1] that the regularized gap function f is continuously differentiable whenever so is the mapping F, and its gradient is given by

$$\nabla f(x) = F(x) - [\nabla F(x) - G](H(x) - x). \tag{2.4}$$

Moreover, when the mapping F is strongly monotone with modulus μ on S, that is,

$$\langle F(y) - F(x), y - x \rangle \geq \mu \|y - x\|^2 \text{ for all } x, y \in S,$$

then f satisfies the inequality [7, Proposition 3.4]

$$f(x) \geq \left(\mu - \frac{1}{2}\|G\| \right) \|x - x^*\|^2 \text{ for all } x \in S. \tag{2.5}$$

The globally convergent Newton method for variational inequalities (1.1) proposed by [7] is the following.

GNEW

Step 0 Choose $x^0 \in S$, $0 < \beta < 1$, $0 < \gamma < 1$, $0 < \sigma < 1$, and a symmetric positive definite matrix G. Let $k := 0$.

Step 1 Find the unique solution $\bar{x}^k \in S$ that solves

$$\left\langle F(x^k) + \nabla F(x^k)^T(\bar{x}^k - x^k), x - \bar{x}^k \right\rangle \geq 0 \text{ for all } x \in S.$$

Let $d^k := \bar{x}^k - x^k$.

Step 2a If $f(x^k + d^k) \leq \gamma f(x^k)$, then set $\alpha_k := 1$ and go to Step 3.

Step 2b Otherwise set $\alpha_k := \beta^{l_k}$ where l_k is the smallest nonnegative integer l such that

$$f(x^k) - f(x^k + \beta^l d^k) \geq -\sigma\beta^l \left\langle \nabla f(x^k), d^k \right\rangle.$$

Step 3 Set $x^{k+1} := x^k + \alpha_k d^k$. Let $k := k + 1$. Return to Step 1.

The algorithm GNEW is shown [7] to be globally convergent under the assumptions that the mapping F is continuously differentiable and strongly monotone and that the matrix G is chosen sufficiently small to satisfy $\| G \| < 2\mu$. The rate of convergence of GNEW is also shown [7] to be quadratic, but it is proved with the two-step line search Step 2a and only under the assumptions that the set S is polyhedral and that the following strict complementarity condition originally introduced by [3] holds.

Definition 2.1 Suppose that S is polyhedral and that problem (1.1) has a unique solution x^*. Let S^* denote the minimal face of S containing x^*. Then we say that the *strict complementarity* holds at x^* if $x \in S$ and $\langle F(x^*), x - x^* \rangle = 0$ imply $x \in S^*$.

3. Main Results

In this section, we consider the behavior of the regularized gap function around the solution. For this purpose, we assume that the mapping F is continuously differentiable and that the transposed Jacobian ∇F is positive definite at a solution x^*. Hence, there is a neighborhood X of the solution x^* such that the mapping F is strongly monotone with modulus μ on X and Lipschitz continuous with modulus L on X, that is,

$$\|F(y) - F(x)\| \leq L\|y - x\| \text{ for all } x, y \in X.$$

Furthermore, we can assume without loss of generality that X is sufficiently small and Newton's method (1.2) is convergent quadratically on X, hence there exists a $\zeta > 0$ such that

$$\|N(x) - x^*\| \leq \zeta\|x - x^*\|^2, \quad N(x) \in X \tag{3.1}$$

holds for all $x \in X$, where $N(x) \in S$ denotes the solution of the linearized variational inequality:

$$\left\langle F(x) + \nabla F(x)^T(N(x) - x), y - N(x) \right\rangle \geq 0 \text{ for all } y \in S. \tag{3.2}$$

We note that $N(x) \in S \cap X$ for all $x \in X$ and X is assumed to be small enough to satisfy $X \subset \{x \in \Re^n \mid \zeta\|x - x^*\| < 1\}$.

We first show the next lemma.

Lemma 3.1 *Suppose that the mapping F is (not necessarily differentiable) Lipschitz continuous on some neighborhood X of a solution x^* to (1.1). Then there exists an $\bar{L} > 0$ such that the following inequality holds for all $x \in X$:*

$$f(x) \leq \langle F(x^*), x - x^* \rangle + \bar{L}\|x - x^*\|^2. \tag{3.3}$$

Proof. From the definition of the regularized gap function (2.2), we have

$$
\begin{aligned}
& f(x) - \langle F(x^*), x - x^* \rangle \\
={} & -\langle F(x), H(x) - x \rangle - \frac{1}{2}\langle H(x) - x, G(H(x) - x) \rangle - \langle F(x^*), x - x^* \rangle \\
\leq{} & \langle F(x^*) - F(x), x^* - x \rangle + \langle F(x), x^* - H(x) \rangle - \frac{1}{2}\langle H(x) - x, G(H(x) - x) \rangle \\
& + \langle F(x^*), H(x) - x^* \rangle \\
\leq{} & \langle F(x^*) - F(x), x^* - x \rangle + \langle F(x) - F(x^*), x^* - H(x) \rangle, \qquad (3.4)
\end{aligned}
$$

where the first inequality follows from the fact that x^* is a solution to (1.1) and $H(x) \in S$.

It follows from $H(x^*) = x^*$ and the nonexpansiveness of projection mapping that

$$
\begin{aligned}
\|x^* - H(x)\|_G &= \|H(x^*) - H(x)\|_G \\
&\leq \left\| \left(x^* - G^{-1}F(x^*)\right) - \left(x - G^{-1}F(x)\right) \right\|_G \\
&\leq \|x^* - x\|_G + \left\| G^{-1}\left(F(x^*) - F(x)\right) \right\|_G \\
&\leq \left(\|G\|^{1/2} + L\|G^{-1}\|^{1/2} \right) \|x^* - x\|, \qquad (3.5)
\end{aligned}
$$

where the last inequality follows from the Lipschitz continuity of F. Thus we have from (3.4) and (3.5) that

$$
\begin{aligned}
& f(x) - \langle F(x^*), x - x^* \rangle \\
\leq{} & \|F(x^*) - F(x)\|\|x^* - x\| + \|F(x) - F(x^*)\|_{G^{-1}}\|x^* - H(x)\|_G \\
\leq{} & L\|x^* - x\|^2 + L\|G^{-1}\|^{1/2}\left(\|G\|^{1/2} + L\|G^{-1}\|^{1/2} \right)\|x^* - x\|^2 \\
={} & \left(L + L\|G^{-1}\|^{1/2}\|G\|^{1/2} + L^2\|G^{-1}\| \right)\|x^* - x\|^2,
\end{aligned}
$$

which shows the lemma. $\qquad \square$

Remark 3.1 The same result has also obtained in the proof of [7, Theorem 5.1]. However, the theorem assumes that ∇F is Lipschitz continuous in the neighborhood of a solution x^*, which is stronger than that assumed in this lemma. We also remark that this lemma does not need the differentiability of the mapping F. $\qquad \square$

The next theorem is the main result of this section. The theorem shows that the Newton step $N(x)$ bounds the regularized gap function in cubic order.

Theorem 3.1 *Suppose that the mapping F is continuously differentiable and ∇F is positive definite on the neighborhood X of the solution x^* of the variational inequality (1.1). Suppose also that Newton's method is quadratically convergent on X, and hence (3.1) holds for all $x \in X$. Then there exists an $L' > 0$ such that the regularized gap function (2.2) satisfies*

$$
f(N(x)) \leq L'\|x - x^*\|^3 \ \text{for all} \ x \in X. \qquad (3.6)
$$

Proof. First, it follows from the assumptions that there is a constant $M > 0$ such that $\|\nabla F(x)\| \leq M$ for all $x \in X$. From Lemma 3.1 and the fact that $N(x)$ is in X and is a solution to (3.2), we have

$$
\begin{aligned}
f(N(x)) &\leq \langle F(x^*), N(x) - x^* \rangle + \bar{L}\|N(x) - x^*\|^2 \\
&= \langle F(x^*) - F(x), N(x) - x^* \rangle + \langle F(x), N(x) - x^* \rangle + \bar{L}\|N(x) - x^*\|^2
\end{aligned}
$$

$$
\begin{aligned}
&\leq \quad \|F(x^*) - F(x)\|\|N(x) - x^*\| + \left\langle \nabla F(x)^T (N(x) - x), x^* - N(x) \right\rangle \\
&\quad + \bar{L}\|N(x) - x^*\|^2 \\
&\leq \quad L\|x^* - x\|\|N(x) - x^*\| + M\|N(x) - x\|\|x^* - N(x)\| + \bar{L}\|N(x) - x^*\|^2 \\
&\leq \quad \left((L + M)\|x^* - x\| + \left(M + \bar{L} \right)\|N(x) - x^*\| \right)\|N(x) - x^*\| \\
&\leq \quad \zeta \left((L + M) + \zeta \left(M + \bar{L} \right)\|x - x^*\| \right)\|x - x^*\|^3 \\
&\leq \quad \zeta \left((L + M) + \left(M + \bar{L} \right) \right)\|x - x^*\|^3.
\end{aligned}
$$

This shows the theorem with $L' = \zeta(L + 2M + \bar{L})$ (in the last inequality, $\zeta\|x - x^*\| < 1$ was used). □

From this theorem, we can show that the algorithm GNEW is locally quadratically convergent without the assumptions that the set S is polyhedral convex and that the strict complementarity holds (Definition 2.1). The next corollary guarantees that the step size α_k of Step 2a of the algorithm GNEW is equal to one for sufficiently large k.

Corollary 3.1 *Suppose that the assumptions of Theorem 3.1 hold. Suppose also that the matrix G is chosen such that $\|G\| < 2\mu$. Then for any positive parameter $\gamma < 1$, there is a neighborhood X' of the solution x^* such that $f(N(x)) \leq \gamma f(x)$ holds for all $x \in S \cap X'$.*
Proof. Since, from the assumptions, the inequalities (2.5) and (3.6) hold on $S \cap X$, then we have for all $x \in S \cap X$

$$
\begin{aligned}
f(N(x)) &\leq L'\|x - x^*\|^3 \\
&\leq \frac{2L'\|x - x^*\|}{2\mu - \|G\|} f(x).
\end{aligned}
$$

Therefore, the result holds for x sufficiently close to x^*, such as $\|x - x^*\| \leq \dfrac{2\mu - \|G\|}{2L'}\gamma$. □

The next corollary enables us to replace the two-step line search procedure, Steps 2a and 2b of the algorithm GNEW, with the following Armijo-type line search procedure:

Step 2' Set $\alpha_k := \beta^{l_k}$ where l_k is the smallest nonnegative integer l such that

$$
f(x^k) - f(x^k + \beta^l d^k) \geq \sigma \beta^l \left(\mu - \frac{1}{2}\|G\| \right)\|d^k\|^2. \tag{3.7}
$$

We note that the global convergence theorem can be shown in a way similar to the proof of [7, Theorem 4.1] for the simpler version of GNEW with Step 2'.

Remark 3.2 Practically, the cost of implementing Step 2' is equivalent to that of Steps 2a and 2b rather than using a parameter γ. However, there is a theoretical benefit in introducing Step 2' in that Step 2b is not shown to guarantee the acceptance of unit step size even if the iterate x^k is sufficiently close to a solution, while Step 2' must accept unit step size when the iterate is close to a region. □

Corollary 3.2 *Suppose that all the assumptions of Corollary 3.1 hold. Then there is a neighborhood X'' of x^* such that*

$$
f(x) - f(N(x)) \geq \sigma \left(\mu - \frac{1}{2}\|G\| \right)\|N(x) - x\|^2
$$

holds for all $x \in S \cap X''$, and hence $\alpha_k = 1$ holds for sufficiently large k.

Proof. Since from the assumptions, the inequalities (2.5), (3.1) and (3.6) hold on $S \cap X$, then we have for all $x \in S \cap X$

$$f(N(x)) - f(x) + \sigma \left(\mu - \frac{1}{2}\|G\| \right) \|N(x) - x\|^2$$

$$\leq L'\|x - x^*\|^3 - \left(\mu - \frac{1}{2}\|G\| \right) \|x - x^*\|^2 + \sigma \left(\mu - \frac{1}{2}\|G\| \right) (\|N(x) - x^*\| + \|x - x^*\|)^2$$

$$\leq L'\|x - x^*\|^3 - \left(\mu - \frac{1}{2}\|G\| \right) \|x - x^*\|^2 + \sigma \left(\mu - \frac{1}{2}\|G\| \right) \left(\zeta\|x - x^*\|^2 + \|x - x^*\| \right)^2$$

$$\leq \left\{ L'\|x - x^*\| - \left(\mu - \frac{1}{2}\|G\| \right) + \sigma \left(\mu - \frac{1}{2}\|G\| \right) (\zeta\|x - x^*\| + 1)^2 \right\} \|x - x^*\|^2.$$

Therefore, because $2\mu > \|G\|$ and $0 < \sigma < 1$, we have

$$L'\|x - x^*\| - \left(\mu - \frac{1}{2}\|G\| \right) + \sigma \left(\mu - \frac{1}{2}\|G\| \right) (\zeta\|x - x^*\| + 1)^2 \leq 0$$

for x sufficiently close to x^*, and hence, the result holds for x sufficiently close to x^*. $\qquad \square$

4. Concluding Remarks

In this paper, we have modified and completed the globally convergent Newton method (GNEW) for variational inequalities in two ways. One is to establish quadratic convergence without both the polyhedral assumption and the strict complementarity assumption. Another is to provide a simplified line search procedure.

There are some issues in implementing GNEW. Solving a linearized variational inequality (1.2) is not easy unless the constraint set S is simple such as polyhedral convex. To overcome this, we have already proposed [6], for solving a general convex constrained problem, another Newton method in which not only the mapping F but also convex constraints are linearized. The method is shown to be globally and superlinearly convergent, but the proof of superlinear convergence has not been satisfactory. We are preparing the next paper giving a rigorous proof of superlinear convergence of the Newton's method in [6] based on the idea of this paper.

Another issue is to know strongly monotone modulus μ in advance. In general, to estimate the strongly monotone modulus μ is not easy, even if the mapping F is known to be strongly monotone. The author hopes that the algorithm GNEW is incorporated into a proximal point algorithm [5], which is one of the promising method for solving monotone variational inequalities. In this case, subproblems are strongly monotone and its modulus is already known.

Acknowledgements

References

[1] M. Fukushima: Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems. *Mathematical Programming*, **53** (1992), 99–110.

[2] P.T. Harker and J.S. Pang: Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications. *Mathematical Programming*, **48** (1990), 161–220.

[3] P. Marcotte and J.P. Dussault: A note on a globally convergent Newton method for solving monotone variational inequalities. *Operations Research Letters*, **6** (1987), 35–42.

[4] J.S. Pang and D. Chan: Iterative methods for variational and complementarity problems. *Mathematical Programming*, **24** (1982), 284–313.

[5] R. T. Rockafellar: Monotone operators and the proximal point algorithm. *SIAM Journal on Control and Optimization*, **14** (1976), 877–898.

[6] K. Taji and M. Fukushima: A globally convergent Newton method for solving variational inequality problems with inequality constraints. In D.-Z. Du, L. Qi and R.S. Womersley (eds.): *Recent Advances in Nonsmooth Optimization* (World Scientific, Singapore, 1995), 405–417.

[7] K. Taji, M. Fukushima and T. Ibaraki: A globally convergent Newton method for solving strongly monotone variational inequalities. *Mathematical Programming*, **58** (1993), 369–383.

Kouichi Taji
Department of Mechanical Science and Engineering
Graduate School of Engineering
Nagoya University
Furo-cho, Chikusa, Nagoya 464-8603, Japan
E-mail: `taji@nuem.nagoya-u.ac.jp`