A New Merit Function and A Successive Quadratic Programming Algorithm for Variational Inequality Problems

Kouichi Taji
Graduate School of Information Science, Nara Institute of Science and Technology
8916-5 Takayama, Ikoma, Nara 630-01, JAPAN

Masao Fukushima
Graduate School of Information Science, Nara Institute of Science and Technology
8916-5 Takayama, Ikoma, Nara 630-01, JAPAN

Abstract

Recently, various merit functions for variational inequality problems have been proposed and their properties have been studied. Unfortunately these functions may not be easy to evaluate unless the constraints of the problem have a relatively simple structure. In this paper, we propose a new merit function for variational inequality problems with general convex constraints. At each point, the proposed function is defined as an optimal value of a quadratic programming problem whose constraints consist of a linear approximation of the given nonlinear constraints. We show that the set of constrained minima of the proposed merit function coincides with the set of solutions to the original variational inequality problem. We also show that this function is directionally differentiable in all directions and, under suitable assumptions, any stationary point of the function over the constraint set actually solves the original variational inequality problem. Finally, we propose a descent method for solving the variational inequality problem and prove its convergence. The method is closely related to a successive quadratic programming method for solving nonlinear programming problems.

1 Introduction

Let S be a nonempty closed convex subset of R^n and let F be a continuous mapping from R^n into R^n. The variational inequality problem (VI) is to find a vector $x^* \in S$ such that

$$\langle F(x^*), x - x^* \rangle \geq 0 \quad \text{for all} \quad x \in S, \tag{1}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product in R^n. This problem has been used to formulate and study various equilibrium models arising in economics, operations research, transportation and regional science [13].

To solve the variational inequality problem (1), many iterative methods, such as Newton method, projection methods, the linearized Jacobi method and the successive overrelaxation methods, have been proposed and their convergence properties have been examined [7].

On the other hand, various merit functions for VI have been proposed and their properties have been studied [1, 2, 5, 10, 14]. Among them, the gap function

$$g(x) = \max_y \{ \langle F(x), x - y \rangle \mid y \in S \}, \tag{2}$$

first introduced by Auslender [2], has the property that its minimum on S coincides with a solution to VI, and hence problem (1) can be reformulated as the optimization problem

$$\text{minimize } g(x) \quad \text{subject to } x \in S. \tag{3}$$

Based on this formulation, Marcotte [11] proposed a descent algorithm for monotone variational inequalities, and Marcotte and Dussault [12] presented a globally convergent modification of Newton method.

Fukushima [5] has proposed another optimization formulation of VI, using the regularized gap function

$$\hat{f}(x) = \max_y \left\{ -\langle F(x), y - x \rangle - \frac{1}{2} \langle y - x, G(y - x) \rangle \;\middle|\; y \in S \right\}, \tag{4}$$

where G is an $n \times n$ symmetric positive definite matrix. The function is shown to be differentiable whenever so is F, while the gap function g defined by (2) is generally nondifferentiable. Fukushima [5] has proposed a descent method for solving VI using the merit function $\hat{f}$. The merit function $\hat{f}$ has also been used in a globally convergent modification of Newton method by Taji, Fukushima and Ibaraki [15]. Independently, Auchmuty [1] has proposed a class of merit functions which includes the gap function g and Fukushima's merit function. Larsson and Patriksson [10] have developed and generalized Auchmuty's class of merit functions. Wu, Florian and Marcotte [16] have proposed a general descent framework for VI by using a class of gap functions.

In this paper, we propose a new merit function, which is defined by (4) with the set S replaced by its polyhedral outer approximation. This function has the advantage over the function $\hat{f}$, particularly when S is a general convex set specified

by nonlinear convex inequalities so that $\hat{f}$ is difficult to evaluate. We show that the proposed merit function has the property that its minimum on S is a solution to (1). We also show that this function is directionally differentiable in all directions and, under suitable assumptions, any stationary point of the function over the constraint set actually solves the original problem VI. Finally, we propose a descent method for solving VI and establish its convergence. We note that the method is closely related to a successive quadratic programming method for solving nonlinear programming problems.

2 A New Merit Function

In this section, we introduce a new merit function for VI which is a relaxed version of the merit function introduced by Fukushima [5]. In the remainder of this paper, we suppose that the set S consists of points $x \in R^n$ satisfying

$$c_i(x) \leq 0, \quad i = 1, 2, \ldots, m, \tag{5}$$

where $c_i : R^n \to R$ are twice continuously differentiable convex functions, and the Slater's constraint qualification holds, i.e., there exists an $\hat{x} \in R^n$ such that

$$c_i(\hat{x}) < 0 \quad \text{for all} \quad i = 1, 2, \ldots, m. \tag{6}$$

We also assume that the mapping F is continuously differentiable. Under these assumptions, it is known that x^* is a solution to (1) if and only if there exist Lagrange multipliers $\lambda_i^*, i = 1, \ldots, m$, such that the following conditions are satisfied [7, Proposition 2.2] :

$$\begin{aligned} &F(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla c_i(x^*) = 0, \\ &c_i(x^*) \leq 0,\ \lambda_i^* \geq 0,\ \lambda_i^* c_i(x^*) = 0,\ i = 1, 2, \ldots, m. \end{aligned} \tag{7}$$

For each $x \in R^n$, we define the set $T(x)$ as

$$T(x) = \{y \in R^n \mid c_i(x) + \langle \nabla c_i(x), y - x \rangle \leq 0, \quad i = 1, 2, \ldots, m\}. \tag{8}$$

Note that, for all $x \in R^n$, the set $T(x)$ is a polyhedral convex set which always contains the set S.

Using the set $T(x)$, we define the function f by

$$f(x) = \max \left\{ -\langle F(x), y - x \rangle - \frac{1}{2} \langle y - x, G(y - x) \rangle \;\middle|\; y \in T(x) \right\}, \tag{9}$$

where G is an $n \times n$ symmetric positive definite matrix. Note that the positive definiteness of G and the convexity of $T(x)$ guarantee that the maximum in (9) is always attained uniquely. Thus we can rewrite (9) as

$$f(x) = -\langle F(x), H(x) - x \rangle - \frac{1}{2} \langle H(x) - x, G(H(x) - x) \rangle, \tag{10}$$

where $H(x)$ is the unique solution y to the quadratic programming problem

$$\begin{array}{lll} \mathrm{QP}(x): & \text{minimize}_y & \frac{1}{2}\langle y-x, G(y-x)\rangle + \langle F(x), y-x\rangle \\ & \text{subject to} & y \in T(x). \end{array} \tag{11}$$

Note that by the optimality condition for (11), $H(x)$ satisfies the variational inequality

$$\langle F(x) + G(H(x)-x), y - H(x)\rangle \geq 0 \quad \text{for all } y \in T(x). \tag{12}$$

The next Lemma characterizes a solution of problem (1) as a fixed point of the mapping H.

Lemma 2.1 *x is a solution to (1) if and only if $H(x) = x$.*

Proof. Let x^* solve (1). It is known [3, page 143] that if the Slater's constraint qualification (6) holds, then x^* also satisfies the inequality

$$\langle F(x^*), y - x^*\rangle \geq 0 \quad \text{for all} \quad y \in T(x^*). \tag{13}$$

Since $H(x^*)$ solves $\mathrm{QP}(x^*)$, $H(x^*)$ satisfies

$$\langle F(x^*) + G(H(x^*) - x^*), y - H(x^*)\rangle \geq 0 \quad \text{for all } y \in T(x^*). \tag{14}$$

Since $H(x^*) \in T(x^*)$, it follows from (13) that

$$\langle F(x^*), H(x^*) - x^*\rangle \geq 0,$$

while since $x^* \in S \subset T(x^*)$, (14) implies

$$\langle F(x^*) + G(H(x^*) - x^*), x^* - H(x^*)\rangle \geq 0.$$

Adding the above two inequalities, we have

$$\langle G(H(x^*) - x^*), H(x^*) - x^*\rangle \leq 0.$$

Hence, we must have $x^* = H(x^*)$ because of the positive definiteness of G.

On the other hand, suppose $x^* = H(x^*)$. Then we have from (12)

$$\langle F(x^*), y - x^*\rangle \geq 0 \quad \text{for all} \quad y \in T(x^*),$$

which implies

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

because $S \subset T(x^*)$. Therefore, to prove that x^* solves (1), we only need to show that $x^* \in S$. Since $H(x^*) \in T(x^*)$ and $x^* = H(x^*)$, it follows from the definition (8) of $T(x)$ that $c_i(x^*) \leq 0$ for all $i = 1, 2, \ldots, m$. This completes the proof. □

Now let us consider the optimization problem

$$\text{minimize} \quad f(x) \quad \text{subject to} \quad x \in S, \tag{15}$$

where the function f is defined by (9). The next theorem establishes the equivalence between the variational inequality problem (1) and the optimization problem (15).

Theorem 2.2 *Let the function $f\colon R^n \to R$ be defined by* (9). *Then $f(x) \geq 0$ for all $x \in S$. Moreover, $x \in S$ and $f(x) = 0$ if and only if x solves the variational inequality problem* (1). *Hence x solves* (1) *if and only if it solves the optimization problem* (15).

Proof. Recall that $x \in T(x)$ whenever $x \in S$. So it follows from (9) that

$$\begin{aligned} f(x) &\geq -\langle F(x), x - x\rangle - \frac{1}{2}\langle x - x, G(x - x)\rangle \\ &= 0, \end{aligned}$$

for any $x \in S$. We shall show that $x \in S$ and $f(x) = 0$ if and only if $H(x) = x$. This along with Lemma 2.1 proves the rest of the theorem. First suppose that $x \in S$ and $f(x) = 0$. Since $x \in T(x)$, (12) implies

$$\langle F(x) + G(H(x) - x), x - H(x)\rangle \geq 0,$$

namely

$$-\langle F(x), H(x) - x\rangle \geq \langle H(x) - x, G(H(x) - x)\rangle. \tag{16}$$

Hence, it follows from (10) and (16) that

$$f(x) \geq \frac{1}{2}\langle H(x) - x, G(H(x) - x)\rangle.$$

But since G is positive definite and $f(x) = 0$, we have $H(x) = x$. Next suppose that $H(x) = x$. Then, it follows from (10) that $f(x) = 0$. Moreover, by the feasibility of $H(x)$ to problem (11), we have

$$c_i(x) + \langle \nabla c_i(x), H(x) - x\rangle \leq 0, \ \ i = 1, \ldots, m.$$

Thus $H(x) = x$ implies $x \in S$. This completes the proof. □

Next we consider the continuity and differentiability of f.

Definition 2.3 [9] Let Φ be a point-to-set-mapping from X into 2^X, where X is a subset of R^n.

1. Φ is *open* at a point $\bar{x}$ if, for any sequence $\{x^k\}$ such that $x^k \to \bar{x}$, $\bar{y} \in \Phi(\bar{x})$ implies that there exists a sequence $\{y^k\}$ such that $y^k \in \Phi(x^k)$ and $y^k \to \bar{y}$.

2. Φ is *closed* at a point $\bar{x}$ if, for any sequence $\{x^k\}$ such that $x^k \to \bar{x}$, $y^k \in \Phi(x^k)$ and $y^k \to \bar{y}$ imply $\bar{y} \in \Phi(\bar{x})$.

3. Φ is *continuous* at a point $\bar{x}$ if it is both open and closed at $\bar{x}$. Φ is continuous on X if it is continuous at any point in X.

Proposition 2.4 *Let c_i, $i = 1, 2, \ldots, m$, be continuously differentiable and satisfy the Slater's constraint qualification (6). Then the point-to-set mapping T defined by (8) is continuous on R^n.*

Proof. Since $c_i(x) + \langle \nabla c_i(x), y - x \rangle$ is continuous with respect to (x, y) for each i, it follows from [9, Theorem 10] that T is closed at each x. By assumption, there is an $\hat{x}$ such that $c_i(\hat{x}) < 0$ for all $i = 1, 2, \ldots, m$. Moreover from the convexity of c_i, we have $c_i(x) + \langle \nabla c_i(x), \hat{x} - x \rangle \leq c_i(\hat{x}) < 0$. Hence, it follows from [9, Theorem 12] that T is open at each x. Therefore, T is continuous on R^n. □

Lemma 2.5 *Let the mapping F be continuous. If the Slater's constraint qualification (6) holds, then the mapping H is bounded on any bounded set.*

Proof. Suppose that H is not bounded on some bounded set B. Then there is a sequence $\{x^k\}$ in B such that $\| H(x^k) \| \to \infty$. Since $\{F(x^k)\}$ is bounded on B by the continuity of F and since G is positive definite, it follows from (10) that $f(x^k) \to -\infty$. On the other hand, since $\hat{x} \in T(x^k)$ for all k, we have

$$\begin{aligned} f(x^k) &= \max \left\{ -\langle F(x^k), y - x^k \rangle - \frac{1}{2} \langle y - x^k, G(y - x^k) \rangle \;\middle|\; y \in T(x^k) \right\} \\ &\geq -\langle F(x^k), \hat{x} - x^k \rangle - \frac{1}{2} \langle \hat{x} - x^k, G(\hat{x} - x^k) \rangle, \end{aligned}$$

which is bounded below by the boundedness of $\{x^k\}$ and $\{F(x^k)\}$. This is a contradiction. □

Definition 2.6 Let ϕ be a function from R^n into $[-\infty, +\infty]$, and let x be a point where ϕ is finite. We say that ϕ is *directionally differentiable* at x in the direction d if the limit

$$\lim_{t \to 0, t > 0} \frac{\phi(x + td) - \phi(x)}{t} \tag{17}$$

exists, and we call the limit the *directional derivative* and denote it by $\phi'(x; d)$.

Theorem 2.7 *Suppose that the mapping F is continuous and c_i, $i = 1, 2, \ldots, m$, are continuously differentiable. Suppose also that the Slater's constraint qualification (6) holds. Then the function f defined by (10) is continuous on R^n. Moreover, if F is continuously differentiable and c_i, $i = 1, \ldots, m$, are twice continuously differentiable, then f is directionally differentiable in any direction $d \in R^n$ and its directional derivative $f'(x; d)$ is given by*

$$f'(x; d) = \min_{\lambda \in \Lambda(x)} \left\langle F(x) - [\nabla F(x) - G](H(x) - x) - \sum_{i=1}^{m} \lambda_i \nabla^2 c_i(x)(H(x) - x),\ d \right\rangle, \tag{18}$$

where $\Lambda(x)$ is defined by

$$\begin{aligned} \Lambda(x) \;=\; \{ \lambda \in R^m \mid\; & F(x) + G(H(x) - x) + \sum_{i=1}^{m} \lambda_i \nabla c_i(x) = 0, \lambda_i \geq 0, \\ & \lambda_i [c_i(x) + \langle \nabla c_i(x), H(x) - x \rangle] = 0, i = 1, \ldots, m \}. \end{aligned} \tag{19}$$

Proof. To prove the first half, it is sufficient to show that the mapping $H(x)$ is continuous. Under the given assumptions, the point-to-set mapping T is continuous by Proposition 2.4 and the objective function of (11) is continuous with respect to (x, y). Therefore, by [9, Theorem 8], the mapping H is closed. However, since H is everywhere defined and single-valued, that H must be continuous for any x. The second half follows from [8, Theorem 2]. □

Remark 2.8 If the set $\Lambda(x)$ is a singleton, f is differentiable at x and the gradient is given by

$$\nabla f(x) = F(x) - [\nabla F(x) - G](H(x) - x) - \sum_{i=1}^{m} \lambda_i \nabla^2 c_i(x)(H(x) - x).$$

A sufficient condition [4, Theorem 6] for $\Lambda(x)$ to be a singleton is that the vectors $\nabla c_i(x), i \in I(x)$, are linearly independent, where $I(x) = \{i \mid c_i(x) + \langle \nabla(x), H(x) - x\rangle = 0\}$, and the strict complementarity condition is satisfied i.e., $\lambda_i = 0$ implies $c_i(x) + \langle \nabla c_i(x), H(x) - x\rangle < 0$.

Remark 2.9 The merit function $\hat{f}$ defined by (4) also has the property that the set of minima of $\hat{f}$ on S coincides with the set of solutions to problem (1). Moreover, $\hat{f}$ is continuously differentiable when so is the mapping F. It is easy to see that $f(x) \geq \hat{f}(x)$ because $T(x)$ contains S for all x. In particular, when c_i are all linear, (4) coincides with (9). Thus, the function f is regarded as a relaxed version of $\hat{f}$.

To obtain a solution to (1) by solving problem (15), we need to find a global minimizer of f on S. It is therefore desirable to know conditions under which a stationary point of (15) is a global optimal solution, because most optimization algorithms only find a stationary point of the problem. The next theorem answers the question.

Theorem 2.10 *Suppose that the mapping F is continuously differentiable, ∇F is positive definite, the convex functions c_i, $i = 1, \ldots, m$, are twice continuously differentiable and the Slater's constraint qualification (6) is satisfied. If $x \in S$ and*

$$f'(x; y - x) \geq 0 \quad \textit{for all } y \in S, \tag{20}$$

then x is a solution to (1).

Before proving the theorem, we show the next lemma. The lemma will also be used to derive a descent condition in the next section.

Lemma 2.11 *Let the mapping F be continuously differentiable and the convex functions c_i, $i = 1, \ldots, m$, be twice continuously differentiable. If $d = H(x) - x$, then we have*

$$f'(x; d) \leq -\langle d, \nabla F(x) d\rangle + \min_{\lambda \in \Lambda(x)} \left(\sum_{i \in I_+} \lambda_i c_i(x) \right), \tag{21}$$

where $I_+ = \{i | c_i(x) > 0\}$.

Proof. Since $d = H(x) - x$, it follows from the Karush-Kuhn-Tucker conditions for problem (11) that d together with a Lagrange multiplier vector λ satisfies

$$F(x) + Gd + \sum_{i=1}^{m} \lambda_i \nabla c_i(x) = 0, \tag{22a}$$

$$c_i(x) + \langle \nabla c_i(x), d \rangle \leq 0, \qquad \lambda_i \geq 0, \tag{22b}$$

$$\lambda_i [c_i(x) + \langle \nabla c_i(x), d \rangle] = 0, \quad i = 1, \ldots, m. \tag{22c}$$

From (18) and (22a) and (22c), we have

$$\begin{aligned} f'(x;d) &= \min_{\lambda \in \Lambda(x)} \left\langle F(x) - [\nabla F(x) - G]d - \sum_{i=1}^{m} \lambda_i \nabla^2 c_i(x) d,\ d \right\rangle \\ &= \min_{\lambda \in \Lambda(x)} \left\{ -\langle d, \nabla F(x) d \rangle - \left\langle d, \sum_{i=1}^{m} \lambda_i \nabla c_i(x) \right\rangle - \left\langle d, \sum_{i=1}^{m} \lambda_i \nabla^2 c_i(x) d \right\rangle \right\} \\ &= \min_{\lambda \in \Lambda(x)} \left\{ -\langle d, \nabla F(x) d \rangle + \sum_{i=1}^{m} \lambda_i c_i(x) - \sum_{i=1}^{m} \lambda_i \langle d, \nabla^2 c_i(x) d \rangle \right\}. \end{aligned} \tag{23}$$

Since the convexity of c_i ensures that $\nabla^2 c_i(x)$ is positive semi-definite, i.e., $\langle d, \nabla c_i^2(x) d \rangle \geq 0$, and since $\sum_{i \notin I_+} \lambda_i c_i(x) \leq 0$, we have from (23)

$$f'(x;d) \leq -\langle d, \nabla F(x) d \rangle + \min_{\lambda \in \Lambda(x)} \left(\sum_{i \in I_+} \lambda_i c_i(x) \right).$$

This completes the proof. □

Proof of Theorem 2.10. Suppose that $x \in S$ satisfies (20). First note that, under the Slater's constraint qualification, we have

$$T(x) \subset x + \mathrm{cl}\{\alpha(y - x) \mid y \in S, \alpha > 0\}, \tag{24}$$

where cl denotes closure of a set [3, page 143]. Thus, from (20), (24) and the positive homogeneity of $f'(x;\cdot)$, we can deduce that

$$f'(x; y - x) \geq 0 \ \text{ for all } y \in T(x). \tag{25}$$

Since $x \in S$ implies $I_+ = \phi$, it follows from Lemma 2.11 that

$$f'(x; H(x) - x) \leq -\langle H(x) - x, \nabla F(x)(H(x) - x) \rangle. \tag{26}$$

Since $H(x) \in T(x)$, we have from (25) and (26)

$$\langle H(x) - x, \nabla F(x)(H(x) - x) \rangle \leq 0.$$

But since $\nabla F(x)$ is positive definite, we have $x = H(x)$. Therefore, it follows from Lemma 2.1 that x is a solution to (1). □

3 Successive Quadratic Programming Algorithm

In this section, we present a successive quadratic programming algorithm for solving the variational inequality problem (1) and prove its convergence.

First, we show that the vector

$$d = H(x) - x$$

is a descent direction of the penalty function $\theta_r : R^n \to R$ defined by

$$\theta_r(x) = f(x) + r \sum_{i=1}^{m} \max(0, c_i(x)), \tag{27}$$

where r is a sufficiently large positive parameter.

Theorem 3.1 *Let the mapping F be continuously differentiable and the convex functions c_i, $i = 1, \ldots, m$, be twice continuously differentiable. If $\nabla F(x)$ is positive definite and*

$$\| \lambda \|_\infty \leq r \quad \textit{for all } \lambda \in \Lambda(x),$$

then the vector $d = H(x) - x$ satisfies the descent condition

$$\theta_r'(x; d) < 0,$$

whenever $d \neq 0$.

Proof. Let $I_+ = \{i \,|\, c_i(x) > 0\}$ and $I_0 = \{i \,|\, c_i(x) = 0\}$. By Theorem 2.7 and [6, Lemma 3.1], θ_r is directionally differentiable and the directional derivative is given by

$$\theta_r'(x; d) = f'(x; d) + r \sum_{i \in I_+} \langle \nabla c_i(x), d \rangle + r \sum_{i \in I_0} \max(0, \langle \nabla c_i(x), d \rangle). \tag{28}$$

First note that $d = H(x) - x$ together with a Lagrange multiplier vector λ satisfies

$$F(x) + Gd + \sum_{i=1}^{m} \lambda_i \nabla c_i(x) = 0, \tag{29a}$$

$$c_i(x) + \langle \nabla c_i(x), d \rangle \leq 0, \qquad \lambda_i \geq 0, \tag{29b}$$

$$\lambda_i [c_i(x) + \langle \nabla c_i(x), d \rangle] = 0, \quad i = 1, \ldots, m. \tag{29c}$$

Then (29b) yields

$$\sum_{i \in I_0} \max(0, \langle \nabla c_i(x), d \rangle) = 0. \tag{30}$$

By Lemma 2.11, we have

$$f'(x; d) \leq -\langle d, \nabla F(x) d \rangle + \min_{\lambda \in \Lambda(x)} \left(\sum_{i \in I_+} \lambda_i c_i(x) \right). \tag{31}$$

Hence, it follows from (28), (29b), (30) and (31) that

$$\begin{aligned}\theta_r'(x;d) &\leq -\langle d, \nabla F(x)d\rangle + \min_{\lambda\in\Lambda(x)}\left(\sum_{i\in I_+}\lambda_i c_i(x)\right) + r\sum_{i\in I_+}\langle \nabla c_i(x), d\rangle \\ &\leq -\langle d, \nabla F(x)d\rangle + \min_{\lambda\in\Lambda(x)}\left(\sum_{i\in I_+}(\lambda_i - r)c_i(x)\right) \\ &< 0,\end{aligned}$$

because, by assumption, $\nabla F(x)$ is positive definite and $\| \lambda \|_\infty \leq r$ for all $\lambda \in \Lambda(x)$. The proof is complete. □

Algorithm

Step 0 Choose x^0, $r > 0$ and a symmetric positive definite matrix G. Let $k := 0$

Step 1 Find the unique solution d^k of the quadratic programming problem:

$$\begin{array}{ll}\text{minimize} & \frac{1}{2}\langle d, Gd\rangle + \langle F(x^k), d\rangle \\ \text{subject to} & c_i(x^k) + \langle \nabla c_i(x^k), d\rangle \leq 0, \quad i = 1, 2, \ldots, m.\end{array}$$

Step 2 Set $x^{k+1} := x^k + t_k d^k$, where

$$t_k := \arg\min_{0\leq t\leq 1}\theta_r(x^k + td^k).$$

Let $k := k + 1$. Go to step 1.

The next theorem establishes the global convergence of this algorithm.

Theorem 3.2 *Suppose that $\nabla F(x)$ is positive definite on R^n. Suppose also that the parameter r is chosen sufficiently large. If the sequence $\{x^k\}$ generated by the above algorithm is bounded, then $\{x^k\}$ converges to the unique solution to the variational inequality problem* (1).

Proof. To prove the global convergence, we shall use Zangwill's global convergence Theorem A [17, page 91]. Since $\{x^k\}$ is bounded, it follows from [6, Lemma 3.3] that there exists a positive number $\bar{r}$ such that $\| \lambda^k \|_\infty \leq \bar{r}$ for all k, where λ^k is an arbitrary optimal Lagrange multiplier vector of the quadratic program solved in Step 1 at iteration k. Assuming that $r \geq \bar{r}$, we have from Theorem 3.1 that d^k satisfies the descent condition $\theta_r'(x^k; d^k) < 0$, whenever x^k is not a solution to (1). Moreover, since the map $d = H(x) - x$ and the function θ_r are continuous with respect to x and since the exact line search strategy on a bounded interval is closed in the sense of [17], the overall algorithmic map is closed. Therefore, the above mentioned theorem [17] guarantees that any accumulation point of $\{x^k\}$ is a solution to (1). Since problem (1) has at most one solution by the positive definiteness of $\nabla F(x)$, we conclude that the entire sequence converges to the solution to (1). □

Remark 3.3 When F is a gradient mapping of some differentiable convex function φ, problem (1) corresponds to a necessary and sufficient optimality condition for the convex programming problem

$$\begin{array}{ll} \text{minimize} & \varphi(x) \\ \text{subject to} & c_i(x) \leq 0, \quad i = 1, 2, \ldots, m. \end{array} \tag{32}$$

We may therefore apply the algorithm to (32) with the identification $F = \nabla\varphi$. Then the subproblem solved in Step 1 is the same as that of a standard successive quadratic programming problem [6], but the merit function used in the line search is quite different.

References

[1] G. Auchmuty, Variational Principles for Variational Inequalities, *Numerical Functional Analysis and Optimization*, 10 (1989) 863-874.

[2] A. Auslender, *Optimisation: Méthodes Numériques*, (Masson, Paris, 1976).

[3] M.S. Bazaraa and C.M. Shetty, *Foundations of Optimization*, (Springer-Verlag, Berlin Heidelberg, 1976).

[4] A.V. Fiacco, G.P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, (SIAM, Philadelphia, 1990).

[5] M. Fukushima, Equivalent Differentiable Optimization Problems and Descent Methods for Asymmetric Variational Inequality Problems, *Mathematical Programming*, 53 (1992) 99-110.

[6] S.P. Han, A Globally Convergent Method for Nonlinear Programming, *Journal of Optimization Theory and Applications*, 22 (1977) 297-309.

[7] P.T. Harker, J.S. Pang, Finite-Dimensional Variational Inequality and Nonlinear Complementarity Problems: A Survey of Theory, Algorithms and Applications, *Mathematical Programming*, 48 (1990) 161-220.

[8] W. Hogan, Directional Derivatives for Extremal-Value Functions with Applications to the Completely Convex Case, *Operations Research*, 21 (1973) 187-209.

[9] W. Hogan, Point-to-Set Maps in Mathematical Programming, *SIAM Review*, 15 (1973) 591-603.

[10] T. Larsson, M. Patriksson, A Class of Gap Functions for Variational Inequalities, *Mathematical Programming*, 64 (1994) 53-79.

[11] P. Marcotte, A New Algorithm for Solving Variational Inequalities with Application to the Traffic Assignment Problem, *Mathematical Programming*, 33 (1985) 339-351.

[12] P. Marcotte, J.P. Dussault, A Note on a Globally Convergent Newton Method for Solving Variational Inequalities, *Operations Research Letters*, 6 (1987) 35-42.

[13] A. Nagurney, *Network Economics: A Variational Inequality Approach*, (Kluwer Academic Publishers, Massachusetts, 1993).

[14] J.S. Pang, Newton's Method for B-differentiable Equations, *Mathematics of Operations Research*, 15 (1990) 311-341.

[15] K. Taji, M. Fukushima, T. Ibaraki, A Globally Convergent Newton Method for Solving Strongly Monotone Variational Inequalities, *Mathematical Programming*, 58 (1993) 369-383.

[16] J.H. Wu, M. Florian, P. Marcotte, A General Descent Framework for the Monotone Variational Inequality Problem, *Mathematical Programming*, 61 (1993) 281-300.

[17] W.I. Zangwill, *Nonlinear Programming: A Unified Approach*, (Prentice-Hall, N.J., 1969).