Equivalent Differentiable Optimization Problems and Descent Methods for Asymmetric Variational Inequality Problems

Masao FUKUSHIMA

Department of Applied Mathematics and Physics

Faculty of Engineering

Kyoto University

Kyoto 606 Japan

Abstract   Whether or not the general asymmetric variational inequality problem can be formulated as a differentiable optimization problem has been an open question. This paper gives an affirmative answer to this question. We provide a new optimization problem formulation of the variational inequality problem and show that its objective function is continuously differentiable whenever the mapping involved in the latter problem is continuously differentiable. We also show that under appropriate assumptions on the latter mapping, any stationary point of the optimization problem is a global optimal solution, and hence solves the variational inequality problem. We discuss descent methods for solving the equivalent optimization problem and comment on systems of nonlinear equations and nonlinear complementarity problems.

1   Introduction

In this paper we consider the problem of finding a point $x^* \in R^n$ such that

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

where S is a nonempty closed convex subset of R^n, F is a mapping from R^n into itself, and $\langle \cdot, \cdot \rangle$ denotes the inner product in R^n. This problem is called the variational inequality problem and has been used to study various equilibrium models in economics, operations research, transportation and regional science [2,5,7,13].

In the special case where $S = R^n_+$, the nonnegative orthant in R^n, problem (1.1) can be rewritten as the nonlinear complementarity problem

$$x^* \geq 0, \ F(x^*) \geq 0 \ \text{ and } \langle x^*, F(x^*) \rangle \ = 0. \tag{1.2}$$

When $S = R^n$, problem (1.1) simply reduces to solving the system of nonlinear equations

$$F(x) = 0. \tag{1.3}$$

Therefore it is quite natural that many algorithms developed to solve the variational inequality problems are generalizations of the classical methods for systems of equations, such as Newton's method, SOR methods, projection methods and nonlinear Jacobi (diagonalization) methods [3,15].

Yet another approach for solving the variational inequality problem (1.1) is to transform it into an equivalent minimization problem. An obvious advantage of this approach lies in the fact that the latter problem may be solved by descent algorithms which possess a global convergence property. A typical situation where problem (1.1) can be reformulated as an optimization problem is that $F(x)$ is the gradient of a differentiable function $\theta : R^n \to R$, in which case problem (1.1) is equivalent to the problem

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

As is well known [14, Theorem 4.1.6], when the mapping F is differentiable, a necessary and sufficient condition for F to satisfy the above condition is that the Jacobian matrix $\nabla F(x)$ is symmetric for all x. Unfortunately we cannot expect this symmetry condition to hold in many practical equilibrium models.

For the general asymmetric variational inequality problem (1.1), a different optimization approach has been proposed [1,11,12], on the basis of the so called gap function $g : R^n \to R$ defined by

$$g(x) = \max \left\{ \langle F(x), x - y \rangle \,|\, y \in S \right\}. \tag{1.4}$$

2

This function was introduced by Auslender [1] for variational inequality problems and later by Hearn [8] for convex programming problems. It is easy to see that $g(x) \geq 0$ for all $x \in S$ and $g(x^*) = 0$ whenever x^* satisfies (1.1). Hence the variational inequality problem (1.1) is equivalent to the problem

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

By the definition (1.4), the gap function g is in general nondifferentiable. Auslender [1] shows that the function g is differentiable provided the set S is strongly convex. Unfortunately this assumption is not satisfied in many important applications such as traffic or market equilibrium problems where the set S is polyhedral convex.

Thus whether or not the general asymmetric variational inequality problem can be formulated as a differentiable optimization problem has been an open question [7]. This paper gives an affirmative answer to it. We shall provide a new optimization problem formulation of problem (1.1), in which the objective function is continuously differentiable whenever so is the mapping F. Since the set S is only assumed to be closed and convex, a wide variety of practical applications may be treated within this framework.

This paper is organized as follows: In Section 2 we review some preliminary results concerning the projection operators and monotone mappings. In Section 3 we formulate an optimization problem equivalent to the variational inequality problem (1.1). We also show that, under appropriate assumptions, the objective function of the optimization problem is differentiable and any stationary point is a global optimal solution of the problem. In Section 4 we discuss descent algorithms for solving the equivalent optimization problem. Finally in Section 5 we comment on systems of nonlinear equations and nonlinear complementarity problems.

2 Preliminaries

Let $\| \cdot \|_G$ denote the norm in R^n defined by $\|x\|_G = \langle x, Gx \rangle^{\frac{1}{2}}$, where G is a symmetric positive definite matrix. Given the norm $\| \cdot \|_G$, the projection of a point x onto the closed

convex set S, denoted by $\mathrm{Proj}_{S,G}(x)$ is defined as the (unique) solution of the problem

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

The projection operator $\mathrm{Proj}_{S,G}(\cdot)$ is nonexpansive, i.e.,

$$\|\mathrm{Proj}_{S,G}(x) - \mathrm{Proj}_{S,G}(x')\|_G \leq \|x - x'\|_G \qquad \text{for all } x, x' \in R^n. \tag{2.1}$$

Moreover, for each $x \in R^n$, $\mathrm{Proj}_{S,G}(x)$ satisfies

$$\langle x - \mathrm{Proj}_{S,G}(x), G(y - \mathrm{Proj}_{S,G}(x)) \rangle \leq 0 \qquad \text{for all } y \in S. \tag{2.2}$$

A proof of these results for the special case where $G = I$ may be found in [1, Chapter 5] [10, Chapter 1], and can be extended to the present case in a straightforward manner.

The mapping $F : R^n \to R^n$ is said to be monotone on S if

$$\langle F(x) - F(x'), x - x' \rangle \geq 0 \qquad \text{for all } x, x' \in S, \tag{2.3}$$

and strictly monotone on S if strict inequality holds in (2.3) whenever $x \neq x'$. It is known [14, Theorem 5.4.3] that, if F is continuously differentiable and the Jacobian matrix $\nabla F(x)$ is positive definite for all $x \in S$, i.e.,

$$\langle d, \nabla F(x)d \rangle > 0 \qquad \text{for all } x \in S, \, d \in R^n \, (d \neq 0), \tag{2.4}$$

then F is strictly monotone on S. Note that we do not require $\nabla F(x)$ to be symmetric in (2.4). The mapping F is said to be strongly (or uniformly) monotone with modulus $\mu > 0$ on S if

$$\langle F(x) - F(x'), x - x' \rangle \geq \mu \|x - x'\|^2 \qquad \text{for all } x, x' \in S, \tag{2.5}$$

where $\| \cdot \|$ denotes the Euclidean norm in R^n. It can be shown [14, Theorem 5.4.3] that, when F is continuously differentiable, a necessary and sufficient condition for (2.5) is

$$\langle d, \nabla F(x)d \rangle \geq \mu \|d\|^2 \qquad \text{for all } x \in S, \, d \in R^n. \tag{2.6}$$

Clearly any strongly monotone mapping is strictly monotone.

3 Equivalent optimization problem

In the rest of the paper we suppose that an $n \times n$ symmetric positive definite matrix G is given. Let $x \in R^n$ be temporarily fixed and consider the following problem:

$$\text{minimize} \quad \langle F(x), y - x \rangle + \frac{1}{2} \langle y - x, G(y - x) \rangle \quad \text{subject to } y \in S. \tag{3.1}$$

Since this problem is essentially equivalent to the problem

$$\text{minimize} \quad \|y - (x - G^{-1}F(x))\|_G^2 \quad \text{subject to } y \in S,$$

the optimal solution of (3.1) is precisely $\text{Proj}_{S,G}(x - G^{-1}F(x))$, the projection of the point $x - G^{-1}F(x)$ onto the set S with respect to the norm $\|\cdot\|_G$. Thus for each x, we can uniquely determine the optimal solution of (3.1). Hereafter we shall denote it by $H(x)$ for notational simplicity. Namely we define the mapping $H : R^n \to R^n$ as

$$\boxed{H(x) = \text{Proj}_{S,G}(x - G^{-1}F(x)).} \tag{3.2}$$

The mapping H thus defined yields a fixed point characterization of the solution of the variational inequality problem (1.1).

Proposition 3.1 *For each $x \in R^n$, let $H(x)$ be the unique optimal solution of problem (3.1), or equivalently, let $H : R^n \to R^n$ be the mapping defined by (3.2). Then x solves the variational inequality problem (1.1) if and only if x is a fixed point of the mapping H, i.e., $x = H(x)$.*

Proof See for example [2]. □

Now we are in a position to state the optimization problem equivalent to the variational inequality problem (1.1). Let us define the function $f : R^n \to R$ by

$$f(x) = -\langle F(x), H(x) - x \rangle - \frac{1}{2} \langle H(x) - x, G(H(x) - x) \rangle, \tag{3.3}$$

i.e., $f(x)$ is the negative of the optimal value of problem (3.1), and consider the problem

$$\text{minimize} \quad f(x) \quad \text{subject to } x \in S. \tag{3.4}$$

The next theorem establishes the equivalence of the variational inequality problem (1.1) and the optimization problem (3.4).

Theorem 3.1 *Let the function $f : R^n \to R$ be defined by (3.3). Then $f(x) \geq 0$ for all $x \in S$, and $f(x) = 0$ if and only if x solves the variational inequality problem (1.1). Hence x solves the optimization problem (3.4) if and only if it solves the variational inequality problem (1.1).*

Proof The function f can be written as

$$f(x) = \frac{1}{2}\{\|G^{-1}F(x)\|_G^2 - \|H(x) - (x - G^{-1}F(x))\|_G^2\}. \tag{3.5}$$

Since $\|G^{-1}F(x)\|_G$ equals the G-distance between x and $x - G^{-1}F(x)$, and $\|H(x) - (x - G^{-1}F(x))\|_G$ is the G-distance between $x - G^{-1}F(x)$ and its projection $H(x)$ onto S, we have $f(x) \geq 0$ for all $x \in S$. Moreover the definition of $H(x)$ implies that those distances are equal, i.e. $f(x) = 0$, if and only if $x = H(x)$. This along with Proposition 3.1 proves the first part of the theorem. The last part of the theorem is a direct consequence of the first part. $\square$

Note that Theorem 3.1 relies upon no specific assumptions on F and f. In order for the equivalent optimization problem (3.4) to be practically useful, however, the objective function f has to possess some desirable properties. Fortunately we can show that, for any closed convex set S, the function f is continuously differentiable whenever so is the mapping F. This is in contrast with the gap function g defined by (1.4), which in general fails to inherit the smoothness property from F.

Theorem 3.2 *If the mapping $F : R^n \to R^n$ is continuous, then the function $f : R^n \to R$ defined by (3.3) is also continuous. Furthermore, if F is continuously differentiable, then f is also continuously differentiable and its gradient is given by*

$$\nabla f(x) = F(x) - [\nabla F(x) - G](H(x) - x). \tag{3.6}$$

Proof Suppose that F is continuous. Then it follows from (2.1) and (3.2) that the mapping H is continuous. Hence, by (3.3), the function f is also continuous. To prove the second half of the theorem, let us define the function $h : R^n \times S \to R$ by

$$h(x,y) = \langle F(x), y - x \rangle + \frac{1}{2}\langle y - x, G(y - x) \rangle. \tag{3.7}$$

Clearly, when F is continuously differentiable, so is the function h. Since, by definition,

$$f(x) = -\min\{h(x,y)|y \in S\} \tag{3.8}$$

and since the minimum on the right-hand side of (3.8) is uniquely attained at $y = H(x)$, it follows from [1, Chapter 4, Theorem 1.7] that f is differentiable and $\nabla f(x)$ is given by

$$\nabla f(x) = -\nabla_x h(x, H(x)). \tag{3.9}$$

The formula (3.6) now follows from (3.7) and (3.9). □

Now let us have a closer look at the relationship between the variational inequality problem (1.1) and the optimization problem (3.4). Theorem 3.1 says that solving (1.1) amounts to finding a global optimal solution of problem (3.4). However, since the function f is in general nonconvex, problem (3.4) may have local minima or stationary points which do not minimize $f(x)$ over $x \in S$ globally. This could be a serious drawback of the present approach, because most of the existing optimization algorithms are only able to find local optimal solutions when applied to nonconvex problems. Fortunately, this difficulty can be completely avoided when the mapping F is continuously differentiable and $\nabla F(x)$ is positive definite for all $x \in S$.

Theorem 3.3 *Assume that the mapping $F : R^n \to R^n$ is continuously differentiable and its Jacobian matrix $\nabla F(x)$ is positive definite for all $x \in S$. If x is a stationary point of problem (3.4), i.e.,*

$$\langle \nabla f(x), y - x \rangle \geq 0 \quad \text{for all } y \in S, \tag{3.10}$$

then x is a global optimal solution of problem (3.4), and hence it solves the variational inequality problem (1.1).

Proof Suppose that x satisfies (3.10). Then substituting the formula (3.6) into (3.10) and putting $y = H(x)$ yield

$$\langle F(x) + G(H(x) - x), H(x) - x \rangle \geq \langle \nabla F(x)(H(x) - x), H(x) - x \rangle. \tag{3.11}$$

On the other hand, since $H(x) = \text{Proj}_{S,G}(x - G^{-1}F(x))$, it follows from (2.2) that

$$\langle x - G^{-1}F(x) - H(x), G(y - H(x)) \rangle \leq 0 \quad \text{for all } y \in S. \tag{3.12}$$

In particular, putting $y = x$ in (3.12), we obtain

$$\langle x - G^{-1}F(x) - H(x), G(x - H(x)) \rangle = \langle F(x) + G(H(x) - x), H(x) - x \rangle$$

$$\leq 0. \tag{3.13}$$

Thus by (3.11) and (3.13), we have

$$\langle \nabla F(x)(H(x) - x), H(x) - x \rangle \leq 0. \tag{3.14}$$

However, since $\nabla F(x)$ is positive definite, it follows from (2.4) that the inequality (3.14) holds only if $x = H(x)$. By Proposition 3.1, this implies that x solves the variational inequality problem (1.1) and, by Theorem 3.1, x is a global optimal solution of problem (3.4). $\quad\square$

4   Descent methods

Let the mapping $F : R^n \rightarrow R^n$ be continuously differentiable. Then by Theorem 3.2, the objective function f of problem (3.4) is also continuously differentiable and its gradient $\nabla f(x)$ can be calculated from (3.6). We may therefore apply any of the existing gradient-type algorithms to solve problem (3.4). In the following, however, we shall not further discuss the gradient-based approaches, but explore the possibility of utilizing the vector

$$\begin{aligned} d &= H(x) - x \\ &= \text{Proj}_{S,G}(x - G^{-1}F(x)) - x \end{aligned} \tag{4.1}$$

as a search direction at $x \in S$. Since (4.1) does not involve the Jacobian matrix $\nabla F(x)$, the vector d may be much easier to compute than the gradient $\nabla f(x)$, particularly when evaluation of $\nabla F(x)$ is expensive (see (3.6)). The next proposition shows that, under monotonicity assumptions on F, the vector d given by (4.1) is actually a descent direction of the function f at x.

Proposition 4.1 *Let the mapping $F : R^n \rightarrow R^n$ be continuously differentiable. If the Jacobian matrix $\nabla F(x)$ is positive definite on S, then for each $x \in S$ the vector $d = H(x) - x$ satisfies the descent condition*

$$\langle \nabla f(x), d \rangle < 0, \tag{4.2}$$

8

whenever $d \neq 0$. *Moreover if F is strongly monotone with modulus $\mu > 0$ on S, then for each* $x \in S$

$$\langle \nabla f(x), d \rangle \leq -\mu \|d\|^2. \tag{4.3}$$

Proof   From (3.6) and (4.1), we have

$$\langle \nabla f(x), d \rangle = \langle F(x) + G(H(x) - x), H(x) - x \rangle - \langle \nabla F(x)(H(x) - x), H(x) - x \rangle. \tag{4.4}$$

But since $H(x) = \mathrm{Proj}_{S,G}(x - G^{-1}F(x))$, it follows from (2.2) that

$$\langle x - G^{-1}F(x) - H(x), G(y - H(x)) \rangle \leq 0 \quad \text{for all } y \in S,$$

which in turn implies

$$\langle F(x) + G(H(x) - x), H(x) - x \rangle \leq 0. \tag{4.5}$$

Combining (4.4) and (4.5), we obtain

$$\langle \nabla f(x), d \rangle \leq -\langle \nabla F(x)(H(x) - x), H(x) - x \rangle = -\langle \nabla F(x)d, d \rangle. \tag{4.6}$$

Thus if $\nabla F(x)$ is positive definite, then (4.2) follows from (2.4) and (4.6), while if F is strongly monotone, then (4.6) together with (2.6) implies (4.3). $\square$

The following theorem establishes global convergence of the algorithm that uses the search directions $d = H(x) - x$ with the exact line search rule.

Theorem 4.1 *Let $\{x^k\}$ be a sequence generated by the iteration*

$$x^{k+1} = x^k + t_k d^k, \quad k = 0, 1, 2, \ldots,$$

where d^k are given by $d^k = H(x^k) - x^k$ and $t_k \in [0, 1]$ are determined from

$$f(x^k + t_k d^k) = \min\{f(x^k + td^k) | 0 \leq t \leq 1\}. \tag{4.7}$$

Suppose that S is compact and that the mapping $F : R^n \to R^n$ is continuously differentiable and $\nabla F(x)$ is positive definite for all $x \in S$. Then, for any starting point $x^0 \in S$, the generated sequence $\{x^k\}$ lies in S and converges to the unique solution of the variational inequality problem (1.1).

9

Proof First notice that the continuity of F together with the compactness of S implies that problem (1.1) has a solution [10, page 12]. Moreover since the given hypotheses on F imply that F is strictly monotone, the solution of (1.1) is unique [10, page 84].

Since the points x^k and $H(x^k) = x^k + d^k$ both belong to S and since $0 \leq t_k \leq 1$, it follows from the convexity of S that the sequence $\{x^k\}$ is contained in S. To prove global convergence of the algorithm we shall use Zangwill's Convergence Theorem A [16, page 91]. By Proposition 4.1, the direction d satisfies the descent condition (4.2) under the given hypotheses, and $d = H(x) - x$ is continuous with respect to x, whenever F is continuous. Moreover since the point-to-set mapping that assigns to each pair (x, d) the points determined by the exact line search (4.7) is closed in the sense of [16], the algorithmic map associated with the iterative procedure under consideration is also closed. Consequently, the above-mentioned convergence theorem guarantees that any accumulation point of the generated sequence $\{x^k\}$ solves problem (1.1). Since the solution of (1.1) is unique by the assumptions on F, the entire sequence is convergent to the solution of (1.1). □

Under the strong monotonicity assumption on the mapping F, we can establish a stronger convergence result which allows the use of inexact line search procedures. Specifically we consider here the following Armijo-type steplength rule: Let α and β be given constants such that $\alpha > 0$ and $0 < \beta < 1$. For the current iterate x^k and the search direction d^k, determine the steplength t_k as $t_k = \beta^{\ell_k}$ where ℓ_k is the smallest nonnegative integer ℓ such that

$$f(x^k + \beta^\ell d^k) \leq f(x^k) - \alpha\beta^\ell ||d^k||^2. \tag{4.8}$$

Theorem 4.2 *Let $\{x^k\}$ be a sequence generated by the iteration*

$$x^{k+1} = x^k + t_k d^k, \quad k = 0, 1, 2, ...,$$

where d^k are given by $d^k = H(x^k) - x^k$ and $t_k \in [0,1]$ are determined by the Armijo-type rule (4.8). Suppose that the set S is compact and the mapping $F : R^n \to R^n$ is differentiable and strongly monotone with modulus $\mu > 0$ on S. Suppose also that $F(x)$ and $\nabla F(x)$ are

Lipschitz continuous on S, i.e., there exist constants $L_0 > 0$ and $L_1 > 0$ such that

$$||F(x) - F(x')|| \leq L_0 ||x - x'|| \quad for \ all \ x, x' \in S, \tag{4.9}$$

and

$$||\nabla F(x) - \nabla F(x')|| \leq L_1 ||x - x'|| \quad for \ all \ x, x' \in S, \tag{4.10}$$

where the norm on the left-hand side of (4.10) is an appropriate matrix norm. Then the generated sequence $\{x^k\}$ lies in S and converges to the unique solution of the variational inequality problem (1.1) for any starting point $x^0 \in S$, if the positive constant α in (4.8) is chosen sufficiently small so that $\alpha < \mu$.

Proof   Since (2.1), (3.2) and (4.9) ensure that the mapping H is Lipschitz continuous, it follows from (3.6) that the Lipschitz continuity (4.10) of $\nabla F(x)$ implies the same property of $\nabla f(x)$, i.e., there exists a constant $L > 0$ such that

$$||\nabla f(x) - \nabla f(x')|| \leq L ||x - x'|| \quad \text{for all } x, x' \in S. \tag{4.11}$$

It follows from (4.3) and (4.11) that, for $0 \leq t \leq 1$

$$
\begin{aligned}
f(x^k + td^k) - f(x^k) &= \int_0^t \langle \nabla f(x^k + sd^k), d^k \rangle ds \\
&= t \langle \nabla f(x^k), d^k \rangle + \int_0^t \langle \nabla f(x^k + sd^k) - \nabla f(x^k), d^k \rangle ds \\
&\leq -\mu t ||d^k||^2 + \int_0^t Ls ||d^k||^2 ds \\
&= (-\mu t + \frac{1}{2} L t^2) ||d^k||^2.
\end{aligned}
$$

Then it is not difficult to see that the inequality

$$f(x^k + td^k) \leq f(x^k) - \alpha t ||d^k||^2$$

holds for all t satisfying $0 \leq t \leq \min\{1, 2(\mu - \alpha)/L\}$, provided that $\alpha < \mu$. Consequently, from the manner in which the steplength t_k is determined, it follows that $t_k = \beta^{\ell_k}$ satisfies

$$t_k > \min\{\beta, 2\beta(\mu - \alpha)/L\} \tag{4.12}$$

for all k.

On the other hand, since $t_k \leq 1$, the convexity of S ensures that the generated sequence $\{x^k\}$ is contained in S. Therefore, by Theorem 3.1, the sequence $\{f(x^k)\}$ is bounded from below. Also (4.8) implies

$$f(x^{k+1}) \leq f(x^k) - \alpha t_k ||d^k||^2, \tag{4.13}$$

and hence $\{f(x^k)\}$ is monotonically decreasing. Thus it follows from (4.12) and (4.13) that

$$\lim_{k \to \infty} ||d^k|| = 0. \tag{4.14}$$

By the compactness of S, $\{x^k\}$ has at least one accumulation point. Moreover, since $d^k = H(x^k) - x^k$ and since the mapping H is continuous, (4.14) implies that any such accumulation point satisfies $H(x) = x$. Thus it follows from Proposition 3.1 that any accumulation point of $\{x^k\}$ solves the variational inequality problem (1.1). Since the strong monotonicity of F guarantees that the solution of (1.1) is unique, we can conclude that the entire sequence $\{x^k\}$ converges to the solution of (1.1). $\quad \Box$.

In Theorems 4.1 and 4.2, we have assumed the compactness of the set S in order to avoid such anomalies as nonexistence of a solution and divergence of a generated sequence. This assumption may appear so restrictive that the method cannot be applied to a wide range of problems including the nonlinear complementarity problem (1.2) and the system of nonlinear equations (1.3). From a practical viewpoint, however, this difficulty is not very serious. In fact, it may rather be helpful to introduce artificial bounds on the variables, which are supposed to contain a desired solution and prevent the method from generating meaningless points during the iterations.

The iterative methods considered above have a close relationship with the projection method which is currently one of the most popular methods for variational inequality problems [2,3,15]. In fact, using our notation, the projection method may be stated as

$$x^{k+1} = \text{Proj}_{S,G}(x^k - \rho G^{-1} F(x^k)), \tag{4.15}$$

where G is symmetric and positive definite and ρ is a positive parameter fixed throughout the iteration. It is known [2] that if the mapping is strongly monotone and the parameter ρ is chosen to be smaller than a certain threshold value which depends on the modulus of strong

monotonicity of F, then the iteration (4.15) converges to the unique solution of problem (1.1). Note however that the projection method does not resort to minimization of any merit function and hence it is not a descent method.

5 Concluding remarks

The variational inequality problems contain as special cases the nonlinear equations and nonlinear complementarity problems. In the case of nonlinear equations (1.3), the mapping H is simply written as

$$H(x) = x - G^{-1}F(x).$$

Hence problem (3.4) reduces to the problem of minimizing

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

which is to find the (weighted) least squares solution of the system of equations (1.3). Therefore the approach presented in this paper may be regarded as a very natural extension of the classical least squares methods for nonlinear equations [4,14]. Moreover, in the case of nonlinear equations, the iterative methods considered in the previous section take the form

$$x^{k+1} = x^k - t_k G^{-1}F(x^k). \tag{5.1}$$

Note that the iterative methods that use the same search directions as (5.1) but different steplength criteria have recently been studied in [6,9] for systems of nonlinear equations.

Since the evaluation of $f(x)$ for problem (1.1) involves minimization of a convex quadratic function over S, the practicality of the optimization approach based on problem (3.4) is dependent on the tractability of the set S. In the case of the nonlinear complementarity problem (1.2), the particular choice $G = I$ yields

$$H(x) = [x - F(x)]_+,$$

where, for any vector $z \in R^n$, $[z]_+$ denotes the vector with components $\max(0, z_i)$, $i = 1, 2, ..., n$. Hence the optimization problem (3.4) can be written as

$$\text{minimize } \frac{1}{2}\{\|F(x)\|^2 - \|x - F(x) - [x - F(x)]_+\|^2\} \quad \text{subject to } x \geq 0 \tag{5.2}$$

(cf. (3.5).) Since the objective function of problem (5.2) can readily be evaluated from $F(x)$, descent methods such as those studied in the previous section may effectively be applied to solve (5.2). Therefore we may expect that the present approach offers a new insight into the solution of nonlinear complementarity problems as well.

Acknowledgement   I am grateful to Professors T. Ibaraki and T. Hasegawa of Kyoto University for their helpful comments and support.

References

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

[2] S. Dafermos, "Traffic equilibrium and variational inequalities", *Transportation Science* 14 (1980) 42-54.

[3] S. Dafermos, "An iterative scheme for variational inequalities", *Mathematical Programming* 26 (1983) 40-47.

[4] J.E. Dennis, Jr. and R.B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations* (Prentice-Hall, Englewood Cliffs, N.J., 1983).

[5] M. Florian, "Mathematical programming applications in national, regional and urban planning", in: M. Iri and K. Tanabe, Eds., *Mathematical Programming: Recent Developments and Applications* (KTK Scientific Publishers, Tokyo, 1989) 57-81.

[6] J.H. Hammond and T.L. Magnanti, "Generalized descent methods for asymmetric systems of equations", *Mathematics of Operations Research* 12 (1987) 678-699.

[7] P.T. Harker and J.S. Pang, "Finite-dimensional variational inequality and nonlinear complementarity problems: A survey of theory, algorithms and applications", Working paper 87-12-06, Department of Decision Sciences, University of Pennsylvania, (Philadelphia, PA, 1987).

[8] D.W. Hearn, "The gap function of a convex program", *Operations Research Letters* 1 (1982) 67-71.

[9] T. Itoh, M. Fukushima and T. Ibaraki, "An iterative method for variational inequalities with application to traffic equilibrium problems", *Journal of the Operations Research Society of Japan* 31 (1988) 82-103.

[10] D. Kinderlehrer and G. Stampacchia, *An Introduction to Variational Inequalities and Their Applications* (Academic Press, New York, 1980).

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

[13] A. Nagurney, "Competitive equilibrium problems, variational inequalities and regional science", *Journal of Regional Science* 27 (1987) 503-517.

[14] J.M. Ortega and W.C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables* (Academic Press, New York, 1970).

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

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