Journal of Computational Mathematics
Vol.27, No.4, 2009, 459–473.

http://www.global-sci.org/jcm
doi:10.4208/jcm.2009.27.4.013

A COMPARISON OF DIFFERENT CONTRACTION METHODS FOR MONOTONE VARIATIONAL INEQUALITIES*

Bingsheng He Xiang Wang and Junfeng Yang
Department of Mathematics, Nanjing University, Nanjing 210093, China
Email: hebma@nju.edu.cn

Abstract

It is interesting to compare the efficiency of two methods when their computational loads in each iteration are equal. In this paper, two classes of contraction methods for monotone variational inequalities are studied in a unified framework. The methods of both classes can be viewed as prediction-correction methods, which generate the same test vector in the prediction step and adopt the same step-size rule in the correction step. The only difference is that they use different search directions. The computational loads of each iteration of the different classes are equal. Our analysis explains theoretically why one class of the contraction methods usually outperforms the other class. It is demonstrated that many known methods belong to these two classes of methods. Finally, the presented numerical results demonstrate the validity of our analysis.

Mathematics subject classification: 65K10, 90C25, 90C30.
Key words: Monotone variational inequalities, Prediction-correction, Contraction methods.

1. Introduction

Let Ω be a nonempty closed convex subset of $\mathbb{R}^n$ and F be a continuous mapping from $\mathbb{R}^n$ into itself. A variational inequality problem, denoted by $\text{VI}(\Omega, F)$, is to determine a vector $u^* \in \Omega$ such that

$$(u - u^*)^T F(u^*) \geq 0, \qquad \forall\, u \in \Omega. \tag{1.1}$$

$\text{VI}(\Omega, F)$ problem includes nonlinear complementarity problem (when $\Omega = \mathbb{R}^n_+$) and system of nonlinear equations (when $\Omega = \mathbb{R}^n$) as its special cases and thus it has many applications [3,5]. The mapping F is said to be uniformly strong monotone (resp. monotone) on Ω if

$$(u - v)^T (F(u) - F(v)) \geq \mu \|u - v\|^2, \qquad \forall\, u, v \in \Omega,$$

where $\mu > 0$ (resp. $\mu = 0$) is a constant, F is Lipschitz continuous on Ω in the sense that there is a constant $L > 0$ such that

$$\|F(u) - F(v)\| \leq L\|u - v\|, \qquad \forall\, u, v \in \Omega.$$

Throughout this paper we assume that the operator F is monotone and Lipschitz continuous on Ω, and the solution set of $\text{VI}(\Omega, F)$, denoted by Ω^*, is nonempty.

In the literature, there are different types of methods for monotone $\text{VI}(\Omega, F)$ such as projection-contraction methods, continuous methods and cutting plane methods. Among these methods, the projection-contraction type of methods have attracted much attention for their simplicity. Let $P_\Omega(v)$ denote the projection of v onto Ω and u^k be the given current iterate. The

* Received July 2, 2007 / Revised version received April 6, 2008 / Accepted August 6, 2008 /

simplest projection method is the Goldstein-Levitin-Polyak approach [4, 11] which iteratively updates u^{k+1} according to the formula

$$u^{k+1} = P_{\Omega}[u^k - \beta_k F(u^k)]. \tag{1.2}$$

This method produces a convergent sequence for uniformly strong monotone $\mathrm{VI}(\Omega, F)$ when $0 < \beta_L \le \beta_k \le \beta_U < 2\mu/L^2$. The basic projection method (1.2) is called an explicit method because all the terms in its right hand side are known. There are also implicit approaches (whose right hand side includes the unknown vector) such as the Douglas-Rachford operator splitting method [2, 12] which determines u^{k+1} by the recursion form

$$u^{k+1} = P_{\Omega}[u^k - \beta_k F(u^k)] + (F(u^k) - F(u^{k+1})) \tag{1.3}$$

and the proximal point algorithm [13] which generates u^{k+1} by

$$u^{k+1} = P_{\Omega}[u^k - \beta_k F(u^{k+1})]. \tag{1.4}$$

These implicit methods produce convergent sequences for monotone $\mathrm{VI}(\Omega, F)$ when $0 < \beta_L \le \beta_k \le \beta_U < +\infty$. The sequence $\{u^k\}$ generated by (1.4) satisfies

$$\|u^{k+1} - u^*\|^2 \le \|u^k - u^*\|^2 - \|u^k - u^{k+1}\|^2, \quad \forall u^* \in \Omega^*.$$

The above inequality means that the new iterate u^{k+1} is closer to the solution set than the current point u^k. According to [1], the proximal point algorithm belongs to the class of Fejér contraction methods under Euclidean norm, or simply, contraction methods.

The main disadvantage of the implicit methods is that a subproblem should be solved in each iteration. Setting the u^{k+1} in (1.3) and (1.4) by u^k, we get the form (1.2), and the explicit method is convergent only for uniformly strong monotone (or co-coercive) $\mathrm{VI}(\Omega, F)$ when the parameter β_k is rigorously chosen. Instead of directly taking the left hand side of (1.2) as the new iterate, we set

$$\tilde{u}^k = P_{\Omega}[u^k - \beta_k F(u^k)] \tag{1.5}$$

as a predictor, the new iterate u^{k+1} (or called as corrector) will be generated by moving u^k in directions designed based on u^k and $\tilde{u}^k$. Such methods can be viewed as prediction-correction methods [9].

There are a number of contraction methods in the literature which belong to the prediction-correction methods. The purpose of this paper is to analyze the efficiency of the different methods whose computational loads in each iteration are equal. The paper is organized as follows. In section 2, we summarize preliminaries and define some basic concepts which will be used in this paper. Section 3 presents two criterions of the framework of the projection-contraction methods. In section 4, we analyze these two classes of methods theoretically and show that the iterates generated by the second class methods usually get more progress than those in the first class. Then, in section 5 we give linear and nonlinear applications with numerical experiments. As predicted by the analysis, the numerical results show the superiority of a class of methods clearly. Finally we give some conclusion remarks in section 6.

2. Preliminaries

Let G be an $n \times n$ positive definite matrix. The projection under G-norm is denoted by $P_{\Omega,G}(\cdot)$, *i.e.*,

$$P_{\Omega,G}(v) = \mathrm{argmin}\{\|v - u\|_G \mid u \in \Omega\}.$$

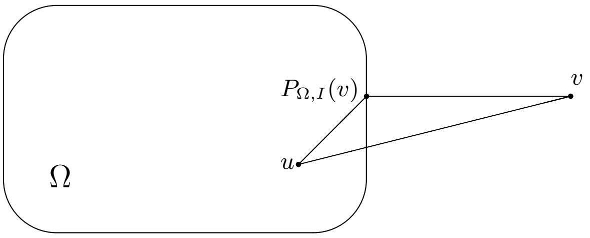

Fig. 1. Geometric interpretation of Inequality (2.3) with respect to Euclidean-norm

Especially, when $G = I$, $P_{\Omega,G}(v)$ is the projection to Ω with respect to the Euclidean-norm. From the above definition, it follows that

$$(v - P_{\Omega,G}(v))^T G(u - P_{\Omega,G}(v)) \le 0, \quad \forall\, v \in \mathbb{R}^n, \forall\, u \in \Omega. \tag{2.1}$$

Consequently, we have

$$\|P_{\Omega,G}(u) - P_{\Omega,G}(v)\|_G \le \|u - v\|_G, \qquad \forall\, u,\ v \in \mathbb{R}^n, \tag{2.2}$$

$$\|u - P_{\Omega,G}(v)\|_G^2 \le \|v - u\|_G^2 - \|v - P_{\Omega,G}(v)\|_G^2, \quad \forall\, v \in \mathbb{R}^n, \forall\, u \in \Omega. \tag{2.3}$$

Notice that variational inequality problem (1.1) is equivalent to finding $u^* \in \Omega$ such that

$$(Gu - Gu^*)^T G^{-1} F(u^*) \ge 0,\ \forall\, u \in \Omega,$$

where G is a positive definite matrix. Thus VI(Ω, F) is equivalent to the following projection equation

$$u = P_{\Omega,G}[u - G^{-1}F(u)]. \tag{2.4}$$

Therefore, solving VI(Ω, F) is equivalent to finding a zero point of the residue function

$$e(u, \beta) := u - P_{\Omega,G}[u - G^{-1}F(u)]. \tag{2.5}$$

To analyze the efficiency of the different methods, we give the following definitions.

Definition 2.1 (Test Vector) *For a given $u \in \Omega$ (or $\in \mathbb{R}^n$), $\tilde{u} \in \Omega$ is said to be a test vector of u if $\tilde{u}$ is generated from u by some well-defined rule and*

$$u = \tilde{u} \qquad \textit{iff} \qquad u \in \Omega^*. \tag{2.6}$$

For given u, there are many different ways to get $\tilde{u}$ which satisfies Definition 2.1. For example, $\tilde{u} = P_\Omega[u - F(u)]$ can be viewed as a test vector of u. In proximal point algorithm [13], for given u, the subproblem produces a $\tilde{u}$ which satisfies

$$\tilde{u} \in \Omega, \quad (u' - \tilde{u})^T[F(\tilde{u}) + (\tilde{u} - u)] \ge 0, \quad \forall u' \in \Omega. \tag{2.7}$$

It is easy to check that $\tilde{u}$ generated by (2.7) also satisfies (2.6) and thus is a test vector of u.

Definition 2.2 (Error Measure Function) *For the given current point u, let $\tilde{u} \in \Omega$ be a test vector of u. A continuous function $\varphi(u, \tilde{u}) : \mathbb{R}^{2n} \to \mathbb{R}$ is said to be an error measure function of VI(Ω, F) if there is a constant $c_0 > 0$, such that*

$$\varphi(u, \tilde{u}) \ge c_0 \|u - \tilde{u}\|^2. \tag{2.8}$$

Remark 2.1. From the definition of the test vector, it is natural to see that $\varphi(u,\tilde{u})=0$ implies $u\in\Omega^*$. In addition, the error measure function $\varphi(u,\tilde{u})$ usually has the following property

$$\varphi(u,\tilde{u})=0 \qquad \textit{iff} \qquad u\in\Omega^*.$$

Therefore, the error measure function $\varphi(u,\tilde{u})$ can be viewed as a measure since it measures the distance between u and Ω^*, just as its name implies.

3. Two Classes of Contraction Methods

The different contraction methods considered in this paper can be divided into two classes. Both of the methods use the same test vector $\tilde{u}$ as predictor and the difference is that they use different search directions to make correction. In order to derive our methods, we give two criterions for the search directions.

Criterion 3.1. *For the given u and its test point $\tilde{u}\in\Omega$, there exist an error measure function $\varphi(u,\tilde{u})$, a direction $d_1(u,\tilde{u})$ and a constant $\tau>0$, satisfying*

$$(u-u^*)^T d_1(u,\tilde{u})\ge\varphi(u,\tilde{u}),\quad \forall\, u^*\in\Omega^*, \tag{3.1}$$

and

$$\frac{\varphi(u,\tilde{u})}{\|d_1(u,\tilde{u})\|^2}\ge\tau. \tag{3.2}$$

Criterion 3.2. *For $u,\tilde{u}\in\Omega,\ \varphi(u,\tilde{u})$ and $d_1(u,\tilde{u})$ defined in Criterion 3.1, there is a direction $d_2(u,\tilde{u})$ which satisfies*

$$\tilde{u}=P_\Omega\{\tilde{u}-[d_2(u,\tilde{u})-d_1(u,\tilde{u})]\} \tag{3.3}$$

and

$$(\tilde{u}-u^*)^T d_2(u,\tilde{u})\ge\varphi(u,\tilde{u})-(u-\tilde{u})^T d_1(u,\tilde{u}),\quad \forall\, u\in\Omega,\ u^*\in\Omega^*. \tag{3.4}$$

Remark 3.1. Consider the equivalence between variational inequality problems and the projection equation (2.4), Condition 3.3 can also be written down as following if general G-norm is under consideration:

$$\tilde{u}=P_{\Omega,G}\{\tilde{u}-G^{-1}[d_2(u,\tilde{u})-d_1(u,\tilde{u})]\}.$$

Remark 3.2. In the algorithms, Conditions (3.1), (3.3) and (3.4) guarantee convergence while Condition (3.2) guarantees to avoid the slow convergence rate. It ought to be mentioned that condition (3.2) is very important for the design of algorithms. Although algorithms without (3.2) can also converge, the convergence rate is much slower than algorithms with this condition.

Lemma 3.1. *Criterion* 3.2 *implies Eq.* (3.1) *in Criterion* 3.1.

Proof. First, it follows from (3.3) that

$$(u'-\tilde{u})^T\{d_2(u,\tilde{u})-d_1(u,\tilde{u})\}\ge 0,\quad \forall\, u'\in\Omega \tag{3.5}$$

and thus (because $u^*\in\Omega$)

$$(\tilde{u}-u^*)^T\{d_1(u,\tilde{u})-d_2(u,\tilde{u})\}\ge 0. \tag{3.6}$$

Adding (3.4) and (3.6) we obtain

$$(\tilde{u}-u^*)^T d_1(u,\tilde{u}) \ge \varphi(u,\tilde{u}) - (u-\tilde{u})^T d_1(u,\tilde{u})$$

and thus

$$(u-u^*)^T d_1(u,\tilde{u}) \ge \varphi(u,\tilde{u}).$$

The lemma is then proved. □

Based on Criterion 3.1, we now give the framework of the first class of methods.

Algorithm 3.1 (General Form)

Let $d_1(u,\tilde{u})$ *satisfy Criterion* 3.1*. G is a positive definite matrix. The new iterate* $(u_1(\alpha))$ *is generated by:*

$$u_1(\alpha) = P_{\Omega,G}[u - \alpha G^{-1} d_1(u,\tilde{u})], \tag{3.7}$$

where

$$\alpha = \gamma\alpha^*, \qquad \alpha^* = \frac{\varphi(u,\tilde{u})}{\|G^{-1}d_1(u,\tilde{u})\|_G^2}, \qquad \gamma \in (0,2). \tag{3.8}$$

By a simple manipulation, it can be proven that

$$\|u(\alpha)-u^*\|_G^2 \le \|u-u^*\|_G^2 - \gamma(2-\gamma)\alpha^*\varphi(u,\tilde{u}), \quad \forall\, u^* \in \Omega^*. \tag{3.9}$$

Therefore, the new iterate $u_1(\alpha)$ is closer to the solution set Ω^* than u and the method belongs to the contraction methods [1] under G-norm. By considering the definitions of the test vector $\tilde{u}$, the error measure function $\varphi(u,\tilde{u})$ and (3.2), the convergence of Algorithm 3.1 follows from (3.9) and the results in [10].

If Criterions 3.1 and 3.2 are both satisfied, we can use the second class of methods.

Algorithm 3.2 (General Form)

Let $d_1(u,\tilde{u})$ *and* $d_2(u,\tilde{u})$ *satisfy both Criterion* 3.1 *and Criterion* 3.2*. The new iterate* $(u_2(\alpha))$ *is generated by*

$$u_2(\alpha) = P_{\Omega,G}[u - \alpha G^{-1} d_2(u,\tilde{u})], \tag{3.10}$$

where α *is just defined in* (3.8)*.*

Note that from (3.2) and (3.8) we have

$$\alpha^* \ge \tau/\|G^{-1}\| > 0, \tag{3.11}$$

which means the step size is bounded below. Thus both algorithms ensure to avoid the extreme slow convergence, see Remark 3.2.

4. The Main Results

To explain theoretically why the second class of methods usually outperform the first class, we define two profit functions. For any solution point $u^* \in \Omega^*$, let

$$\theta_1(\alpha) := \|u-u^*\|_G^2 - \|u_1(\alpha)-u^*\|_G^2 \tag{4.1}$$

and

$$\theta_2(\alpha) := \|u - u^*\|_G^2 - \|u_2(\alpha) - u^*\|_G^2 \tag{4.2}$$

be profit functions in the two classes of algorithms respectively. By setting

$$u(\alpha) = u - \alpha G^{-1} d_1(u, \tilde{u}), \tag{4.3}$$

we will prove two suitably introduced amounts

$$\theta_1(\alpha) \geq q_1(\alpha) = q(\alpha) + \|u(\alpha) - u_1(\alpha)\|_G^2, \tag{4.4}$$

$$\theta_2(\alpha) \geq q_2(\alpha) = q(\alpha) + \|u(\alpha) - u_2(\alpha)\|_G^2, \tag{4.5}$$

where

$$q(\alpha) = 2\alpha\varphi(u) - \alpha^2 \|G^{-1} d_1(u, \tilde{u})\|_G^2. \tag{4.6}$$

Finally, we show that

$$q_2(\alpha) \geq q_1(\alpha) + \|u_2(\alpha) - u_1(\alpha)\|^2. \tag{4.7}$$

This inequality together with (4.4) and (4.5) indicate the possible superiority of the second class of methods to the first class.

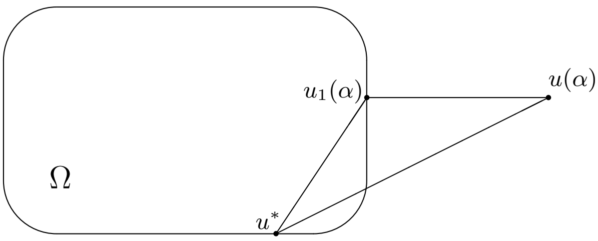

Fig. 2. Geometric interpretation of Inequality (4.11) under Euclidean-norm.

Theorem 4.1. *For any $u^* \in \Omega^*$ and $\alpha \geq 0$, we have*

$$\theta_1(\alpha) \geq q_1(\alpha), \quad \forall\ \alpha \geq 0, \tag{4.8}$$

where

$$q_1(\alpha) = q(\alpha) + \|u(\alpha) - u_1(\alpha)\|_G^2 \tag{4.9}$$

and $q(\alpha)$ is defined by (4.6).

Proof. Note that (see the notations (4.3) and (3.7))

$$u_1(\alpha) = P_{\Omega,G}[u(\alpha)]. \tag{4.10}$$

Since $u_1(\alpha) \in \Omega$ and $u^* \in \Omega$, it follows from (2.3) that

$$\|u_1(\alpha) - u^*\|_G^2 \leq \|u(\alpha) - u^*\|_G^2 - \|u(\alpha) - u_1(\alpha)\|_G^2. \tag{4.11}$$

Substituting (4.11) in (4.1), we have

$$\theta_1(\alpha) \geq \|u - u^*\|_G^2 - \|u(\alpha) - u^*\|_G^2 + \|u(\alpha) - u_1(\alpha)\|_G^2. \tag{4.12}$$

From (4.3), (3.1) and (4.6) we have

$$\begin{aligned}
&\|u-u^*\|_G^2 - \|u(\alpha)-u^*\|_G^2 \\
&= \|u-u^*\|_G^2 - \|u-u^*-\alpha G^{-1}d_1(u,\tilde{u})\|_G^2 \\
&= 2\alpha(u-u^*)^T d_1(u,\tilde{u}) - \alpha^2\|G^{-1}d_1(u,\tilde{u})\|_G^2 \\
&\geq 2\alpha\varphi(u,\tilde{u}) - \alpha^2\|G^{-1}d_1(u,\tilde{u})\|_G^2 = q(\alpha). \qquad (4.13)
\end{aligned}$$

Since $q_1(\alpha) = q(\alpha) + \|u(\alpha) - u_1(\alpha)\|_G^2$, it follows from (4.12) and (4.13) that

$$\theta_1(\alpha) \geq q_1(\alpha), \quad \forall\ \alpha \geq 0$$

and the theorem is proved. □

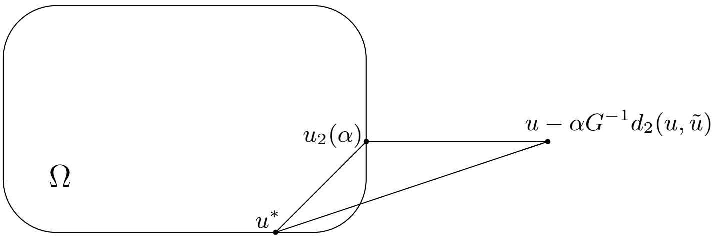

Fig. 3. Geometric interpretation of Inequality (4.15) under Euclidean-norm.

Theorem 4.2. *For any $u^* \in \Omega^*$ and $\alpha \geq 0$, we have*

$$\theta_2(\alpha) \geq q_2(\alpha), \qquad \forall\ \alpha \geq 0$$

where

$$q_2(\alpha) = q(\alpha) + \|u(\alpha) - u_2(\alpha)\|_G^2 \qquad (4.14)$$

and $q(\alpha)$ is defined by (4.6).

Proof. Since $u_2(\alpha) = P_{\Omega,G}[u - \alpha G^{-1}d_2(u,\tilde{u})]$ and $u^* \in \Omega$, it follows from (2.3) that

$$\|u_2(\alpha)-u^*\|_G^2 \leq \|u-\alpha G^{-1}d_2(u,\tilde{u})-u^*\|_G^2 - \|u-\alpha G^{-1}d_2(u,\tilde{u})-u_2(\alpha)\|_G^2. \qquad (4.15)$$

Consequently, using the definition of $\theta_2(\alpha)$ (see (4.2)), we get

$$\begin{aligned}
\theta_2(\alpha) &\geq \|u-u^*\|_G^2 - \|u-\alpha G^{-1}d_2(u,\tilde{u})-u^*\|_G^2 + \|u-\alpha G^{-1}d_2(u,\tilde{u})-u_2(\alpha)\|_G^2 \\
&= \|u-u_2(\alpha)\|_G^2 + 2\alpha(u-u^*)^T d_2(u,\tilde{u}) + 2\alpha(u_2(\alpha)-u)^T d_2(u,\tilde{u}) \\
&= \|u-u_2(\alpha)\|_G^2 + 2\alpha(\tilde{u}-u^*)^T d_2(u,\tilde{u}) + 2\alpha(u_2(\alpha)-\tilde{u})^T d_2(u,\tilde{u}). \qquad (4.16)
\end{aligned}$$

Since $u_2(\alpha) \in \Omega$, it follows from (3.5) that

$$(u_2(\alpha)-\tilde{u})^T d_2(u,\tilde{u}) \geq (u_2(\alpha)-\tilde{u})^T d_1(u,\tilde{u}). \qquad (4.17)$$

Substituting (3.4) and (4.17) in the right hand side of (4.16), we get

$$\begin{aligned}
\theta_2(\alpha) &\geq \|u-u_2(\alpha)\|_G^2 + 2\alpha\varphi(u) - 2\alpha(u-\tilde{u})^T d_1(u,\tilde{u}) + 2\alpha(u_2(\alpha)-\tilde{u})^T d_1(u,\tilde{u}) \\
&= \|u-u_2(\alpha)\|_G^2 + 2\alpha\varphi(u) + 2\alpha(u_2(\alpha)-u)^T d_1(u,\tilde{u}) \\
&= \|u-u_2(\alpha)-\alpha G^{-1}d_1(u,\tilde{u})\|_G^2 + 2\alpha\varphi(u) - \alpha^2\|G^{-1}d_1(u,\tilde{u})\|_G^2 \\
&= \|u(\alpha)-u_2(\alpha)\|_G^2 + q(\alpha). \qquad (4.18)
\end{aligned}$$

The last equality of (4.18) follows from the definitions of $u(\alpha)$ and $q(\alpha)$. According to (4.14), the right hand side of (4.18) is $q_2(\alpha)$ and the assertion of the theorem is proved. $\square$

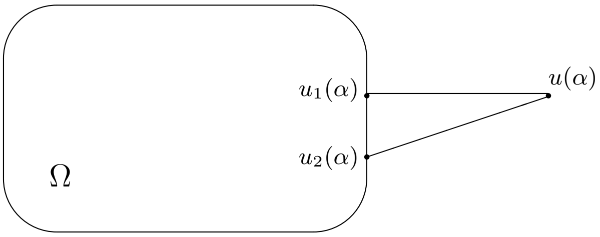

Fig. 4. Geometric interpretation of Inequality (4.21) under Euclidean-norm.

The assertion of this theorem follows directly from (4.20) and (4.21). $\square$

Theorem 4.3. *Let $q_1(\alpha)$ and $q_2(\alpha)$ be defined by* (4.9) *and* (4.14)*, respectively. Then we have*

$$q_2(\alpha) - q_1(\alpha) \geq \|u_2(\alpha) - u_1(\alpha)\|_G^2, \qquad \forall\ \alpha \geq 0. \tag{4.19}$$

Proof. It follows from (4.9) and (4.14) that

$$q_2(\alpha) - q_1(\alpha) = \|u(\alpha) - u_2(\alpha)\|_G^2 - \|u(\alpha) - u_1(\alpha)\|_G^2. \tag{4.20}$$

Note that $u_1(\alpha) = P_{\Omega,G}[u(\alpha)]$ (see the notations (4.3) and (3.1)) and $u_2(\alpha) \in \Omega$. By using (2.3), we obtain

$$\|u_2(\alpha) - u_1(\alpha)\|_G^2 \leq \|u(\alpha) - u_2(\alpha)\|_G^2 - \|u(\alpha) - u_1(\alpha)\|_G^2. \tag{4.21}$$

The assertion of this theorem follows directly from (4.20) and (4.21). $\square$

5. Applications to Some Existing Methods

5.1. Methods for monotone linear variational inequalities

We consider the monotone linear variational inequality LVI(Ω, M, q):

$$u^* \in \Omega, \qquad (u' - u^*)^T(Mu^* + q) \geq 0, \qquad \forall\, u' \in \Omega.$$

The method based on Criterion 3.1. It was proved that (see He [7] and Solodov and Tseng [14])

$$(u - u^*)^T(M^T + I)e(u) \geq \|e(u)\|^2, \quad \forall\ u^* \in \Omega^*, \tag{5.1}$$

where

$$e(u) = u - P_\Omega[u - (Mu + q)].$$

For the current point $u \in \Omega \setminus \Omega^*$, by letting

$$\tilde{u} = P_\Omega[u - (Mu + q)] \neq u, \tag{5.2}$$

$$\varphi(u, \tilde{u}) = \|u - \tilde{u}\|^2, \tag{5.3}$$

$$d_1(u, \tilde{u}) = (M^T + I)(u - \tilde{u}), \tag{5.4}$$

it is clear that $\tilde{u}$ is a test vector of u and $\varphi(u,\tilde{u})$ is an error measure function of LVI(Ω, M, q). Note that $d_1(u,\tilde{u}) \neq 0$, by observing that

$$(u-\tilde{u})^T d_1(u,\tilde{u}) = (u-\tilde{u})M^T(u-\tilde{u}) + \|u-\tilde{u}\|^2 > 0.$$

In this way Inequality (5.1) can be written as

$$(u-u^*)^T d_1(u,\tilde{u}) \geq \varphi(u,\tilde{u}), \quad \forall\, u^* \in \Omega^*$$

and thus the first condition (3.1) in Criterion 3.1 is satisfied. In this case, since

$$\frac{\varphi(u,\tilde{u})}{\|d_1(u,\tilde{u})\|^2} = \frac{\|u-\tilde{u}\|^2}{\|(M^T+I)(u-\tilde{u})\|^2} \geq \frac{1}{\|M^T+I\|^2},$$

the second condition (3.2) in Criterion 3.1 holds.

Based on Inequality (5.1), some Fejér monotone methods were established by He [7] (in Euclidean norm) and Solodov and Tseng [14] (in general G-norm). Of course, if the method under the Euclidean-norm is clear, the extension to the general G-norm is trivial. From the above analysis we get the algorithm based on Criterion 3.1 for solving linear variational inequality:

Algorithm 5.1 (LVI Form)

Given initial point u^0, $\epsilon > 0$, a positive definite matrix G and $\gamma \in (0,2)$. Repeat the following process until $\|u-\tilde{u}\|_2 < \epsilon$:

$$\begin{cases} \tilde{u} := P_\Omega[u-(Mu+q)] \\ \alpha := \dfrac{\|u-\tilde{u}\|_2^2}{\|G^{-1}(M^T+I)(u-\tilde{u})\|_G^2} \\ u := P_{\Omega,G}\{u-\gamma\alpha G^{-1}[(M^T+I)(u-\tilde{u})]\} \end{cases}$$

The method based on Criterion 3.2. For monotone LVI(Ω, M, q) and the notations given by (5.2)-(5.4), we let

$$d_2(u,\tilde{u}) = M^T(u-\tilde{u}) + (Mu+q), \tag{5.5}$$

and will prove that the conditions in Criterion 3.2 are satisfied. It is easy to check that

$$\tilde{u} = P_\Omega[u-(Mu+q)] = P_\Omega[\tilde{u}-(d_2(u,\tilde{u})-d_1(u,\tilde{u}))]$$

and thus the first condition in Criterion 3.2 is satisfied. In [6], it was proved that

$$(u-u^*)^T[M^Te(u)+(Mu+q)] \geq e(u)^T(Mu+q), \qquad \forall\, u\in\Omega,\ u^*\in\Omega^*.$$

By using notations of $\tilde{u}$ and $d_2(u,\tilde{u})$, the above inequality can be rewritten as

$$(u-u^*)^T d_2(u,\tilde{u}) \geq (u-\tilde{u})^T(Mu+q), \qquad \forall\, u\in\Omega,\ u^*\in\Omega^*.$$

Consequently,

$$(\tilde{u}-u^*)^T d_2(u,\tilde{u}) \geq (u-\tilde{u})^T[(Mu+q)-d_2(u,\tilde{u})], \qquad \forall\, u\in\Omega,\ u^*\in\Omega^*. \tag{5.6}$$

By using

$$\begin{aligned}&(Mu+q)-d_2(u,\tilde{u})=(u-\tilde{u})-d_1(u,\tilde{u}),\\&\|u-\tilde{u}\|^2=\varphi(u,\tilde{u}),\end{aligned}$$

it follows from (5.6) that

$$\begin{aligned}(\tilde{u}-u^*)^T d_2(u,\tilde{u}) &\geq (u-\tilde{u})^T\{(u-\tilde{u})-d_1(u,\tilde{u})\}\\&=\varphi(u,\tilde{u})-(u-\tilde{u})^T d_1(u,\tilde{u})\end{aligned}$$

and thus the second condition in Criterion 3.2 is satisfied.

Now we give the algorithm based on Criterion 3.2 for solving linear variational inequality as follows (note that the step size is the same as in Algorithm 5.1):

Algorithm 5.2 (LVI Form)

Given initial point u^0, $\epsilon > 0$, *a positive definite matrix* G *and* $\gamma \in (0,2)$. *Repeat the following process until* $\|u-\tilde{u}\|_2 \leq \epsilon$

$$\begin{cases}\tilde{u} := P_\Omega[u-(Mu+q)]\\ \alpha := \dfrac{\|u-\tilde{u}\|_2^2}{\|G^{-1}(M^T+I)(u-\tilde{u})\|_G^2}\\ u := P_{\Omega,G}\{u-\gamma\alpha G^{-1}[M^T(u-\tilde{u})+(Mu+q)]\}\end{cases}$$

Numerical experiments. We implement Algorithms 5.1 and 5.2 to Example 1 in [15] for finding the shortest network in a given full Steiner topology. Based on

$$\|d\|_2=\max_{\xi\in B_2}\xi^T d, \qquad \text{where} \qquad B_2=\{\xi \mid \|\xi\|_2\leq 1\},$$

the l_2-norm problem was translated to a min-max problem and its equivalent form is a monotone linear variational inequality [9]. For l_1-norm and l_∞-norm distance problems, we translate the problems to a linear variational inequality by using

$$\begin{aligned}\|d\|_1&=\max_{\xi\in B_\infty}\xi^T d, \qquad \text{where} \qquad B_\infty=\{\xi \mid \|\xi\|_\infty\leq 1\}\\ \|d\|_\infty&=\max_{\xi\in B_1}\xi^T d, \qquad \text{where} \qquad B_1=\{\xi \mid \|\xi\|_1\leq 1\},\end{aligned}$$

respectively. The example is tested with $G=I$ and starting point $u^0=0$ under l_1, l_2 and l_∞ norms. The numerical results are given in Tables 1–3. From the numerical results we can see that with the same accuracy $\epsilon = 10^{-10}$ both Algorithm 5.1 and Algorithm 5.2 get the same total distance. However, as the theoretical analysis indicated, Algorithm 5.2 performs better. It only use about 54% iterations of Algorithm 5.1 and save nearly 50% CPU-time. All tests are run on a Lenovo Pentium 4 CPU 2.66GHz 256M PC.

Table 1. Shortest network under l_1 norm

Algorithm 5.1			Algorithm 5.2		
Iteration	$\|e(u)\|_\infty$	Total Distance	Iteration	$\|e(u)\|_\infty$	Total Distance
20	1.2e+000	29.0823040485	20	1.4e-001	28.8805008662
40	3.7e-002	28.6777786413	40	1.1e-004	28.6660178525
60	9.7e-004	28.6661448683	60	1.3e-007	28.6658582049
80	2.5e-005	28.6658649129	80	1.4e-010	28.6658580002
100	6.7e-007	28.6658581765	81	1.0e-010	28.6658580000
120	1.8e-008	28.6658580046			
140	4.8e-010	28.6658580001			
149	9.4e-011	28.6658580000			
CPU-time		0.031 Sec.	CPU-time		0.016 Sec.

Table 2. Shortest network under l_2 norm

Algorithm 5.1			Algorithm 5.2		
Iteration	$\|e(u)\|_\infty$	Total Distance	Iteration	$\|e(u)\|_\infty$	Total Distance
20	1.2e+000	25.7823694981	20	1.3e-001	25.4681520030
40	7.1e-002	25.3776304969	40	5.0e-004	25.3563526162
60	3.3e-003	25.3568683223	60	4.4e-006	25.3560698260
80	1.8e-004	25.3561050662	80	4.0e-008	25.3560677986
100	1.1e-005	25.3560698181	100	3.7e-010	25.3560677795
120	6.4e-007	25.3560678958	106	9.2e-011	25.3560677793
140	3.9e-008	25.3560677857			
160	2.4e-009	25.3560677797			
180	1.5e-010	25.3560677793			
183	9.5e-011	25.3560677793			
CPU-time		0.234 Sec.	CPU-time		0.125 Sec.

Table 3. Shortest network under l_∞ norm

Algorithm 5.1			Algorithm 5.2		
Iteration	$\|e(u)\|_\infty$	Total Distance	Iteration	$\|e(u)\|_\infty$	Total Distance
20	1.5e+000	22.1882071270	20	6.8e-001	21.8459714979
40	9.0e-002	21.1322990353	40	2.1e-003	21.1145131146
60	2.0e-003	21.1133865209	60	9.0e-007	21.1129140170
80	4.4e-005	21.1129244226	80	4.1e-010	21.1129135002
100	1.0e-006	21.1129137566	84	7.4e-011	21.1129135000
120	2.4e-008	21.1129135060			
140	5.9e-010	21.1129135001			
150	9.2e-011	21.1129135000			
CPU-time		0.187 Sec.	CPU-time		0.094 Sec.

5.2. Methods for monotone nonlinear variational inequalities

We consider monotone nonlinear variational inequality

$$u^* \in \Omega, \qquad (u' - u^*)^T F(u^*) \geq 0, \qquad \forall\, u' \in \Omega.$$

The method based on Criterion 3.1. For $u \in \Omega \setminus \Omega^*$, let

$$\tilde{u} = P_\Omega[u - \beta F(u)] \neq u, \tag{5.7}$$

where $\beta > 0$ is chosen to satisfy

$$\beta\|F(u) - F(\tilde{u})\| \le \nu\|u - \tilde{u}\|, \quad \nu \in (0,1). \tag{5.8}$$

It was proved (see [8], Inequality (22) and [14]) that

$$(u - u^*)^T d_1(u,\tilde{u}) \ge (u - \tilde{u})^T d_1(u,\tilde{u}), \quad \forall\, u^* \in \Omega^*, \tag{5.9}$$

where

$$d_1(u,\tilde{u}) = (u - \tilde{u}) - \beta(F(u) - F(\tilde{u})). \tag{5.10}$$

Note that $d_1(u,\tilde{u}) \ne 0$, by observing that

$$(u - \tilde{u})^T d_1(u,\tilde{u}) \ge (1-\nu)\|u - \tilde{u}\|^2 > 0.$$

It is clear that $\tilde{u}$ generated by (5.7) is a test vector of u. Let

$$\varphi(u,\tilde{u}) = (u - \tilde{u})^T d_1(u,\tilde{u}). \tag{5.11}$$

Using Cauchy-Schwarz inequality, it follows from (5.8) and (5.10) that

$$\begin{aligned}\varphi(u,\tilde{u}) &= (u - \tilde{u})^T d_1(u,\tilde{u})\\ &= \|u - \tilde{u}\|^2 - (u - \tilde{u})^T \beta\big(F(u) - F(\tilde{u})\big)\\ &\ge (1-\nu)\|u - \tilde{u}\|^2,\end{aligned}$$

which implies that $\varphi(u,\tilde{u})$ is an error measure function of VI(Ω, F). Using the notation of $\varphi(u,\tilde{u})$ and $d_1(u,\tilde{u})$, (5.9) can be written as

$$(u - u^*)^T d_1(u,\tilde{u}) \ge \varphi(u,\tilde{u}), \quad \forall\, u^* \in \Omega^* \tag{5.12}$$

and thus Condition (3.1) holds. In addition, we have

$$\begin{aligned}2\varphi(u,\tilde{u}) &= 2(u - \tilde{u})^T d_1(u,\tilde{u})\\ &= 2\|u - \tilde{u}\|^2 - 2\beta(u - \tilde{u})^T\big(F(u) - F(\tilde{u})\big)\\ &\ge \|u - \tilde{u}\|^2 - 2\beta(u - \tilde{u})^T\big(F(u) - F(\tilde{u})\big) + \beta^2\|F(u) - F(\tilde{u})\|^2\\ &= \|d_1(u,\tilde{u})\|^2.\end{aligned}$$

Consequently,

$$\frac{\varphi(u,\tilde{u})}{\|d_1(u,\tilde{u})\|^2} \ge \frac{1}{2} \tag{5.13}$$

and (3.2) is satisfied. From the above analysis, we now give the algorithm based on Criterion 3.1 to solve nonlinear monotone variational inequality, see Algorithm 5.3.

Remark 5.1. Since the theme of this paper is to compare two types of projection-contraction methods and the differences is mainly derived by two different search directions, we omit the technique of adjusting β to satisfy (5.8) both in Algorithm 5.3 and the next one, namely Algorithm 5.4. Readers can get some references in [9].

Algorithm 5.3 (NVI Form)

Given initial point u^0, $\epsilon > 0$, $\nu \in (0,1)$, *a positive definite matrix* G *and* $\gamma \in (0,2)$. *Repeat the following iteration until* $\|u - \tilde{u}\|_2 < \epsilon$*:*

$$\begin{cases} \begin{cases} \tilde{u} := P_\Omega[u - \beta F(u)] \\ \quad \text{Ajust } \beta \text{ till (5.8) is satisfied} \end{cases} \\ e := u - \tilde{u} \\ d_1 := e - \beta(F(u) - F(\tilde{u})) \\ \alpha := \dfrac{e^T d_1}{\|G^{-1} d_1\|_G^2} \\ u := P_{\Omega,G}[u - \gamma\alpha G^{-1} d_1] \end{cases}$$

The method based on Criterion 3.2. For the monotone nonlinear VI(Ω, F), $\tilde{u}$, $\varphi(u,\tilde{u})$ and $d_1(u,\tilde{u})$ defined in this subsection, we let

$$d_2(u,\tilde{u}) = \beta F(\tilde{u}), \tag{5.14}$$

and will prove that the conditions in Criterion 3.2 are satisfied. It follows from (5.7) and (5.10) that

$$\tilde{u} = P_\Omega\{\tilde{u} - [d_2(u,\tilde{u}) - d_1(u,\tilde{u})]\} \tag{5.15}$$

and thus the first condition in Criterion 3.2 holds. Since F is monotone, we have

$$(\tilde{u} - u^*)^T \beta F(\tilde{u}) \geq (\tilde{u} - u^*)^T \beta F(u^*) \geq 0.$$

Using (5.11) and (5.14), it follows from above inequality that

$$(\tilde{u} - u^*)^T d_2(u,\tilde{u}) \geq \varphi(u,\tilde{u}) - (u - \tilde{u})^T d_1(u,\tilde{u}) \tag{5.16}$$

and the second condition in Criterion 3.2 is satisfied.

From the above analysis, we get the algorithm based on Criterion 3.2 to solve nonlinear monotone variational inequality (note that the step size is the same as in Algorithm 5.3):

Algorithm 5.4 (NVI Form)

Given initial point u^0, $\epsilon > 0$, $\nu \in (0,1)$, *a positive definite matrix* G *and* $\gamma \in (0,2)$. *Repeat the following process until* $\|u - \tilde{u}\|_2 < \epsilon$*:*

$$\begin{cases} \begin{cases} \tilde{u} := P_\Omega[u - \beta F(u)] \\ \quad \text{Ajust } \beta \text{ till (5.8) is satisfied} \end{cases} \\ e := u - \tilde{u} \\ d_1 := e - \beta(F(u) - F(\tilde{u})) \\ d_2 := \beta F(\tilde{u}) \\ \alpha := \dfrac{e^T d_1}{\|G^{-1} d_1\|_G^2} \\ u := P_{\Omega,G}[u - \gamma\alpha G^{-1} d_2] \end{cases}$$

Numerical experiments. In this experiment, we implement Algorithms 5.3 and 5.4 to the first example in [9]. We take $G = I$ and the origin as the initial point, $\gamma = 1.8$, $\nu = 0.9$, the initial $\beta = 0$ and the stop criterion $\epsilon = 10^{-7}$. The numerical results are shown in Table 4. From the results, we can see that the second class of methods also performs better than the first class. Algorithm 5.4 only uses 67% $\sim$ 87% iterations and save about 30% CPU-time. All tests are run on a Lenovo Pentium 4 CPU 2.66GHz 256M PC.

Table 4. Numerical results of nonlinear variational inequality

	Algorithm 5.3		Algorithm 5.4	
Problem Size	Iteration	CPU-time (Sec.)	Iteration	CPU-time (Sec.)
100	488	0.067	383	0.043
200	636	0.203	460	0.140
500	671	2.828	467	1.969
800	539	6.047	365	4.109
1000	587	8.891	510	7.688

6. Conclusions

In this paper, we compare the efficiency of two classes of contraction methods for solving monotone variational inequalities in a unified framework. Under this framework, convergence analysis of many existing methods becomes much easier, and it indicates the possible superiority of the second class of methods. Numerical results verify the validity of our theoretical analysis.

Acknowledgments. The research of the first author is supported by Jiangsu Province NSF BK2008255, The Cultivation Fund of the Key Scientific and Technical Innovation Project Ministry of Education of China 708044 and The Doctoral Fund of Ministry of Education of China 20060284001.

References

[1] E. Blum and W. Oettli, Mathematische Optimierung, Econometrics and Operations Research XX, Springer Verlag, 1975.

[2] J. Douglas, and H.H. Rachford, On the numerical solution of the heat conduction problem in 2 and 3 space variables, *Trans. Amer. Math. Soc.*, **82** (1956), 421-439.

[3] R. Glowinski, Numerical Methods for Nonlinear Variational Problems, Springer-Verlag: New York, Berlin, Heidelberg, Tokyo, 1984.

[4] A. A. Goldstein, Convex programming in Hilbert space, *Bull. Amer. Math. Soc.*, **70** (1964), 709-710.

[5] P. T. Harker and J.S. Pang, Finite-dimensional variational inequality and nonlinear complementarity problems: A survey of theory algorithms and applications, *Math. Program.*, **48** (1990), 161-220.

[6] B.S. He, A projection and contraction method for a classs of linear complementarity problems and its application in convex quadratic programming, *Appl. Math. Opt.*, **25** (1992), 247-262.

[7] B.S. He, A new method for a class of linear variational inequalities, *Math. Program.*, **66** (1994), 137-144.

[8] B.S. He, A class of projection and contraction methods for monotone variational inequalities, *Appl. Math. Opt.*, **35** (1997), 69-76.

[9] B.S. He, X.M. Yuan and J.J.Z. Zhang, Comparison of two kinds of prediction-correction methods for monotone variational inequalities, *Comput. Optim. Appl.*, **27** (2004), 247-267.

[10] B.S. He, M.J. Qian and Y.M. Wang, Study on approximate proximal point algorithms for monotone variational inequalities, Proceedings of the 2003' International Conference on Numerical Optimization and Numerical Linear Algebra, edited by Y.X. Yuan, 2003.

[11] E.S. Levitin and B.T. Polyak, Constrained minimization problems, *USSR Comput. Math. & Math. Phys.*, **6** (1966), 1-50.

[12] P.L. Lions and B. Mercier, Splitting algorithms for the sum of two nonlinear operators, *SIAM J. Numer. Anal.*, **16** (1979), 964-979.

[13] R.T. Rockafellar, Augmented lagrangians and applications of the proximal point algorithm in convex programming, *Math. Oper. Res.*, **1** (1976), 97-116.

[14] M.V. Solodov and P. Tseng, Modified projection-type methods for monotone variational inequalities, *SIAM J. Control Optim.*, **34** (1996), 1814-1830.

[15] G.L. Xue and Y.Y. Ye, An efficient algorithm for minimizing a sum of Euclidean norms with applications, *SIAM J. Optimiz.*, **7** (1997), 1017-1036.