凸优化和单调变分不等式的收缩算法

第三讲：单调变分不等式投影收缩算法中的两对孪生方法

Two pairs of twin projection and contraction methods for monotone variational inequalities

南京大学数学系　　何炳生

hebma@nju.edu.cn

The context of this lecture is based on the publication [5, 7, 10]

We study the solution methods for nonlinear VI. For any $\beta > 0$, the variational inequalities

$$\textbf{VI}(\Omega, F) \quad u^* \in \Omega, \quad (u - u^*)^T F(u^*) \geq 0, \quad \forall u \in \Omega \tag{0.1}$$

and

$$\textbf{VI}(\Omega, F) \quad u^* \in \Omega, \quad (u - u^*)^T \beta F(u^*) \geq 0, \quad \forall u \in \Omega \tag{0.2}$$

have the same solution set. For given u, let $e(u, \beta) := u - P_\Omega[u - \beta F(u)]$, $\|e(u, \beta)\|$ is the error function, which measures how much u failed to be a solution point.

Recent years, the projection contraction algorithm for linear variational inequality [3, 4] has been successfully applied to robot motion planning and real-time control [2, 11]. Some scholars in Geotechnical Mechanics have solved the problems that worry them for a long time [12, 13] by the PC Algorithms [6, 7] successfully .

In the last lecture, we have introduced the PC Algorithms for monotone LVI and NVI whose search directions are based on FI1+FI2 and FI1+FI2+FI3, respectively. Actually, accompanied each PC Algorithm in the last lecture, there is a twin algorithm, whose search directions are based on FI1 and FI1+FI3, respectively. It is reasonable, among the twin algorithms, the method utilizes fewer fundamental inequalities, is more efficient.

It is interesting that the same step length is used when the new iteration points are updated by the different correction procedures. From the perspective of mathematics itself, there is such a clever inner connection, which also gives us the enjoyment of the beauty of mathematics.

1 Mathematical Backgrounds

Basic property of the projection mapping We need to list a few important properties of the projection operator. Readers interested in proving these results can refer to Lecture 2 of this series of lectures.

Lemma 1.1 *Let $\Omega \subset \Re^n$ be a closed convex set, then for any $v \in \Re^n$, we have*

$$(u - P_\Omega(v))^T (P_\Omega(v) - v) \geq 0, \quad \forall u \in \Omega. \tag{1.1}$$

The equivalent projection equation By using the properties of the projection, solving the variational inequality (0.1) is equivalent to finding a zero point of $e(u, \beta)$ which is defined by

$$e(u, \beta) := u - P_\Omega[u - \beta F(u)].$$

The projection contraction algorithm for monotone variational inequality (0.2) is a prediction-correction method. The prediction is provided by projection and the contraction is realized by correction. In the k-th iteration of the PC Algorithms, for given u^k and $\beta_k > 0$, the predictor $\tilde{u}^k$ is given by

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

Thus, $u^k \in \Omega^*$ (or $e(u, \beta) = 0$) if and only if $u^k = \tilde{u}^k$.

Set $v = u^k - \beta_k F(u^k)$ in (1.1), because $\tilde{u}^k = P_\Omega(v)$, it follows from (1.1) that

$$\tilde{u}^k \in \Omega, \quad (u - \tilde{u}^k)^T \{\tilde{u}^k - [u^k - \beta_k F(u^k)]\} \geq 0, \quad \forall u \in \Omega. \tag{1.3}$$

Consequently, we get

$$\tilde{u}^k \in \Omega, \quad (u - \tilde{u}^k)^T \beta_k F(u^k) \geq (u - \tilde{u}^k)^T (u^k - \tilde{u}^k), \quad \forall\, u \in \Omega. \tag{1.4}$$

Three fundamental inequalities Let u^* be a solution of VI (Ω, F). Since $\tilde{u}^k \in \Omega$, according to the definition of the variational inequality (0.1), we have

(FI1) $$(\tilde{u}^k - u^*)^T \beta_k F(u^*) \geq 0. \tag{1.5}$$

Since $u^* \in \Omega$, setting the any $u \in \Omega$ in (1.3) by u^*, it follows that

(FI2) $$(\tilde{u}^k - u^*)^T \big([u^k - \tilde{u}^k] - \beta_k F(u^k)\big) \geq 0. \tag{1.6}$$

The variational inequality considered in this series is monotone, thus

$$\text{(FI3)} \qquad (\tilde{u}^k - u^*)^T \big(\beta_k F(\tilde{u}^k) - \beta_k F(u^*)\big) \geq 0. \tag{1.7}$$

We call (1.5), (1.6) and (1.7) three fundamental inequalities. Although these inequalities are simple, they are very important. The search directions of the PC Algorithms are derived from these inequalities.

2 A pair of twin PC Algorithms for LVI

When the operator $F(u)$ in (0.1) is affine, $F(u) = Mu + q$, such variational inequality is linear (abbreviated LVI):

$$u^* \in \Omega, \quad (u - u^*)^T (Mu^* + q) \geq 0, \quad \forall\, u \in \Omega.$$

We say the LVI is monotone if $M^T + M$ is positive semidefinite, it is not necessarily that M is symmetric.

2.1 The ascent directions provided by the predictor

For linear variational inequality, we use a fixed $\beta > 0$.

- **The ascent direction provided by FI1** Because $F(u) = Mu + q$, the fundamental inequality (1.5) can be written as

$$\{(u^k - u^*) - (u^k - \tilde{u}^k)\}^T \beta\{(Mu^k + q) - M(u^k - u^*)\} \geq 0,$$

Since $(u^k - u^*)^T M(u^k - u^*) \geq 0$, it follows from the above inequality

$$(u^k - u^*)^T \underline{\beta[M^T(u^k - \tilde{u}^k) + (Mu^k + q)]} \geq (u^k - \tilde{u}^k)^T \beta(Mu^k + q), \tag{2.1}$$

If $u^k \in \Omega$, using $F(u) = Mu + q$, from (1.3) we get

$$(u^k - \tilde{u}^k)^T \beta_k (Mu^k + q) \geq \|u^k - \tilde{u}^k\|^2. \tag{2.2}$$

Thus, if $u^k \in \Omega$, $\beta_k \big(M^T(u^k - \tilde{u}^k) + (Mu^k + q)\big)$ is a ascent direction of $\frac{1}{2}\|u - u^*\|^2$ at u^k. **(2.1)-(2.2) is true only for $u^k \in \Omega$.**

- **The ascent direction provided by FI1+FI2** Adding (1.5) and (1.6), and using $F(u) = Mu + q$, we get

$$\{(u^k - u^*) - (u^k - \tilde{u}^k)\}^T \{(u^k - \tilde{u}^k) - \beta M(u^k - u^*)\} \ge 0.$$

Since $(u^k - u^*)^T M(u^k - u^*) \ge 0$, it follows from the above inequality

$$(u^k - u^*)^T \underline{(I + \beta M^T)(u^k - \tilde{u}^k)} \ge \|u^k - \tilde{u}^k\|^2. \qquad (2.3)$$

Thus, $(I + \beta M^T)(u^k - \tilde{u}^k)$ is a ascent direction of the unknown distance function of $\frac{1}{2}\|u - u^*\|^2$ at the point u^k. **(2.3) is true for any $u^k \in \Re^n$.**

A pair of the twin directions For $F(u) = Mu + q$, (1.4) can be written as

$$\tilde{u}^k \in \Omega, \quad (u - \tilde{u}^k)^T \beta(Mu^k + q) \ge (u - \tilde{u}^k)^T (u^k - \tilde{u}^k), \quad \forall\, u \in \Omega.$$

Adding $(u - \tilde{u}^k)^T \beta M^T (u^k - \tilde{u}^k)$ to the both sides of the last inequality, we get

$$\begin{aligned} \tilde{u}^k \in \Omega, \quad & (u - \tilde{u}^k)^T \underline{\beta[M^T(u^k - \tilde{u}^k) + (Mu^k + q)]} \\ & \ge (u - \tilde{u}^k)^T \underline{(I + \beta M^T)(u^k - \tilde{u}^k)}, \quad \forall u \in \Omega. \qquad (2.4) \end{aligned}$$

We call (2.4) the two directions

$$\beta[M^T(u^k - \tilde{u}^k) + (Mu^k + q)] \quad \text{and} \quad (I + \beta M^T)(u^k - \tilde{u}^k) \tag{2.5}$$

which lay on the left and right sides of (2.4), respectively, a pair of twin ascent directions for LVI. They are also the direction in (2.1) and (2.3), which derived from (FI1) and (FI1+FI2), respectively. For the notation convenience, we denote

$$g(u^k, \tilde{u}^k) = M^T(u^k - \tilde{u}^k) + (Mu^k + q). \tag{2.6}$$

Using this expression, (2.4) can be written as

$$\begin{aligned} \tilde{u}^k \in \Omega, \qquad & (u - \tilde{u}^k)^T \underline{\beta g(u^k, \tilde{u}^k)} \\ & \geq (u - \tilde{u}^k)^T \underline{(I + \beta M^T)(u^k - \tilde{u}^k)}, \quad \forall u \in \Omega. \end{aligned} \tag{2.7}$$

Notice that for the direction $(I + \beta M^T)(u^k - \tilde{u}^k)$, there is a constant $K > 0$, such that $\|(I + \beta M^T)(u^k - \tilde{u}^k)\| \leq K\|u^k - \tilde{u}^k\|$. However, for the related direction $g(u^k, \tilde{u}^k)$, we do not have the similar statement !

2.2 Update the new iterate by the direction due to FI1+FI2

The correction uses the descent direction (the opposite of the ascent direction) of the distance function to make the new iteration point closer to the solution set. Based on the direction provided by FI1+FI2, the new iterate is updated by

$$u_{BD}^{k+1}(\alpha) = u^k - \alpha(I + \beta M^T)(u^k - \tilde{u}^k). \tag{2.8}$$

The lower index 'BD' means 'Bounded Direction'. For discussion how to determine the step length α, we denote the output of (2.8) by $u_{BD}^{k+1}(\alpha)$. Let us investigate the α-dependent reduction of the square of the distance

$$\vartheta_k^L(\alpha) := \|u^k - u^*\|^2 - \|u_{BD}^{k+1}(\alpha) - u^*\|^2. \tag{2.9}$$

According to the definition,

$$\begin{aligned} \vartheta_k^L(\alpha) &= \|u^k - u^*\|^2 - \|u^k - u^* - \alpha(I + \beta M^T)(u^k - \tilde{u}^k)\|^2 \\ &= 2\alpha(u^k - u^*)^T(I + \beta M^T)(u^k - \tilde{u}^k) \\ &\quad -\alpha^2\|(I + \beta M^T)(u^k - \tilde{u}^k)\|^2. \end{aligned} \tag{2.10}$$

For any given solution point u^*, (2.10) tell us that $\vartheta_k^L(\alpha)$ is a quadratic function of α. Since u^* is unknown, We can't directly find the maximum of $\vartheta_k^L(\alpha)$. With the help of (2.1), we have the following theorem.

Theorem 2.1 *Let $u_{BD}^{k+1}(\alpha)$ be updated by* (2.8). *Then for $\alpha > 0$, we have*

$$\vartheta_k^L(\alpha) \geq q_k^L(\alpha), \tag{2.11}$$

where $\vartheta_k^L(\alpha)$ defined by (2.9) *and*

$$q_k^L(\alpha) = 2\alpha\|u^k - \tilde{u}^k\|^2 - \alpha^2\|(I + \beta M^T)(u^k - \tilde{u}^k)\|^2. \tag{2.12}$$

Proof. The assertion derived directly from (2.10) by using (2.1). $\square$

Theorem 2.1 indicates that $q_k^L(\alpha)$ is a lower bound of $\vartheta_k^L(\alpha)$. The quadratic function $q_k^L(\alpha)$ reaches its maximum at

$$\alpha_k^* = \text{argmax}\{q_k^L(\alpha)\} = \frac{\|u^k - \tilde{u}^k\|^2}{\|(I + \beta M^T)(u^k - \tilde{u}^k)\|^2}. \tag{2.13}$$

From (2.13), it follows that

$$\alpha_k^* \geq \frac{1}{\|I + \beta M^T\|^2}. \tag{2.14}$$

In the practical computation, we take a relaxed factor $\gamma \in [1, 2)$ and updated the new iterate by

$$u_{BD}^{k+1} = u^k - \gamma \alpha_k^* (I + \beta M^T)(u^k - \tilde{u}^k), \tag{2.15}$$

Theorem 2.2 *Let* $u^{k+1} = u_{BD}^{k+1}$ *updated by* (2.15) *with* α_k^* *given by* (2.13)*, then*

$$\|u^{k+1} - u^*\|^2 \leq \|u^k - u^*\|^2 - \frac{\gamma(2-\gamma)}{\|I + \beta M^T\|^2}\|u^k - \tilde{u}^k\|^2. \tag{2.16}$$

Proof. According to (2.9) and (2.11), the u^{k+1} updated by (2.15) satisfies

$$\|u^{k+1} - u^*\|^2 \leq \|u^k - u^*\|^2 - q_k^L(\gamma \alpha_k^*). \tag{2.17}$$

According to the definitions of $q_k^L(\alpha)$ and α_k^* (see (2.12) and (2.13)), we get

$$q_k^L(\gamma\alpha_k^*) = \gamma(2-\gamma)\alpha_k^*\|u^k - \tilde{u}^k\|^2 \geq \frac{\gamma(2-\gamma)}{\|I+\beta M^T\|^2}\|u^k - \tilde{u}^k\|^2.$$

The last inequality follows from (2.14). The proof is complete. □

2.3 Update the new iterate by the direction due to FI1

The correction form (2.8) in §2.2 takes $(I+\beta M^T)(u^k - \tilde{u}^k)$ as the search direction. In this subsection, it is replaced by its related direction (see (2.5))

$$\beta[M^T(u^k - \tilde{u}^k) + (Mu^k + q)].$$

In §2.1, it is emphasized that (2.1)-(2.2) are true only for $u^k \in \Omega$. We use

$$u_{BLD}^{k+1}(\alpha) = P_\Omega\{u^k - \alpha\beta[M^T(u^k - \tilde{u}^k) + (Mu^k + q)]\}, \qquad (2.18)$$

to update the new iterate ensured in Ω. We denote the output of (2.18) by $u_{BLD}^{k+1}(\alpha)$. The lower index 'BLD' means 'Boundless Direction', because

$g(u^k, \tilde{u}^k) \not\to 0$ as $\mathrm{dist}(\tilde{u}^k, \Omega^*) \to 0$. For discussion how to determine the step length α, Let us investigate the α-dependent reduction of the square of the distance

$$\zeta_k^L(\alpha) = \|u^k - u^*\|^2 - \|u_{BLD}^{k+1}(\alpha) - u^*\|^2, \tag{2.19}$$

which is a function of α. We can not maximize $\zeta_k^L(\alpha)$ directly because it involves the unknown vector u^*. The following theorem indicates that for the same $\alpha > 0$, $\zeta_k^L(\alpha)$ is 'better than $\vartheta_k^L(\alpha)$ in (2.11).

Theorem 2.3 *Let* $u_{BLD}^{k+1}(\alpha)$ *be updated by* (2.18)*. Then for* $\zeta_k^L(\alpha)$ *defined in* (2.19) *with any* $\alpha > 0$*, we have*

$$\zeta_k^L(\alpha) \geq q_k^L(\alpha) + \|u_{BLD}^{k+1}(\alpha) - u_{BD}^{k+1}(\alpha)\|^2, \tag{2.20}$$

where $q_k^L(\alpha), u_{BD}^{k+1}(\alpha)$ *and* $u_{BLD}^{k+1}(\alpha)$ *are given by* (2.12), (2.8) *and* (2.18)*, respectively.*

Proof. By using the notation $g(u^k, \tilde{u}^k)$ in (2.6), the update form (2.18) can be written as $u_{BLD}^{k+1}(\alpha) = P_\Omega[u - \alpha\beta g(u^k, \tilde{u}^k)]$. Since $u^* \in \Omega$ and the projection operator is non-expansive, we have

$$\|u_{_{BLD}}^{k+1}(\alpha) - u^*\|^2 \ \le \ \|u^k - \alpha\beta g(u^k, \tilde{u}^k) - u^*\|^2.$$

However, according to the properties of projection and cosine theorem, we use more precise relations

$$\begin{aligned}\|u_{_{BLD}}^{k+1}(\alpha) - u^*\|^2 \quad &\le \quad \|u^k - \alpha\beta g(u^k, \tilde{u}^k) - u^*\|^2 \\ &\quad -\|u_{_{BLD}}^{k+1}(\alpha) - (u^k - \alpha\beta g(u^k, \tilde{u}^k))\|^2. \quad (2.21)\end{aligned}$$

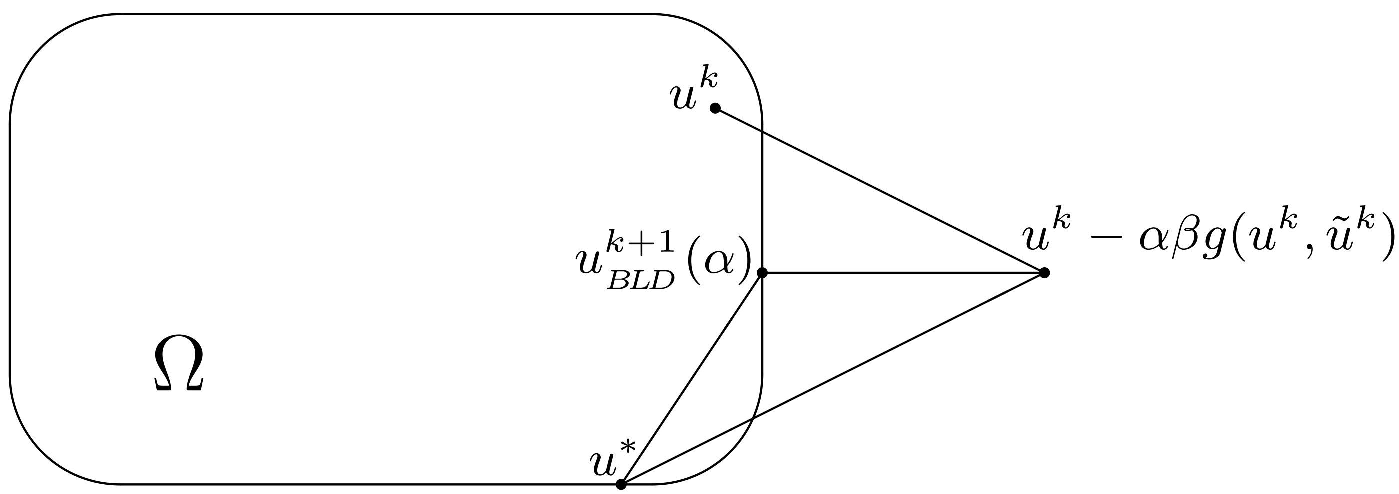

Fig.1 Geometric interpretation of the inequality (2.21)

Hence, by using $\zeta_k^L(\alpha)$ (see (2.19)) and (2.21), we have

$$\begin{aligned}
\zeta_k^L(\alpha) &\geq \|u^k - u^*\|^2 - \|(u^k - u^*) - \alpha\beta g(u^k, \tilde{u}^k)\|^2 \\
&\quad + \|(u_{BLD}^{k+1}(\alpha) - u^k) + \alpha\beta g(u^k, \tilde{u}^k)\|^2 \\
&= 2\alpha\beta(u^k - u^*)^T g(u^k, \tilde{u}^k) + 2\alpha\beta(u_{BLD}^{k+1}(\alpha) - u^k)^T g(u^k, \tilde{u}^k) \\
&\quad + \|u_{BLD}^{k+1}(\alpha) - u^k\|^2 \\
&= \|u_{BLD}^{k+1}(\alpha) - u^k\|^2 + 2\alpha(u_{BLD}^{k+1}(\alpha) - u^*)^T \beta g(u^k, \tilde{u}^k). \qquad (2.22)
\end{aligned}$$

Decomposing the cross term of the right hand side of (2.22) in form

$$\begin{aligned}
&(u_{BLD}^{k+1}(\alpha) - u^*)^T \beta g(u^k, \tilde{u}^k) \\
&= (u_{BLD}^{k+1}(\alpha) - \tilde{u}^k)^T \beta g(u^k, \tilde{u}^k) + (\tilde{u}^k - u^*)^T \beta g(u^k, \tilde{u}^k). \qquad (2.23)
\end{aligned}$$

To the first term of the right hand side of (2.23), since $u_{BLD}^{k+1}(\alpha) \in \Omega$, by using (2.7), we have

$$(u_{BLD}^{k+1}(\alpha) - \tilde{u}^k)^T \beta g(u^k, \tilde{u}^k) \geq (u_{BLD}^{k+1}(\alpha) - \tilde{u}^k)^T (I + \beta M^T)(u^k - \tilde{u}^k).$$

In other words,

$$(u_{BLD}^{k+1}(\alpha) - \tilde{u}^k)^T \beta g(u^k, \tilde{u}^k) \ge (u_{BLD}^{k+1}(\alpha) - u^k)^T (I + \beta M^T)(u^k - \tilde{u}^k)$$
$$+(u^k - \tilde{u}^k)^T (I + \beta M^T)(u^k - \tilde{u}^k). \quad (2.24)$$

To the second term of the right hand side of (2.23), $(\tilde{u}^k - u^*)^T \beta g(u^k, \tilde{u}^k)$, we split it in form

$$(\tilde{u}^k - u^*)^T \beta g(u^k, \tilde{u}^k) = (\tilde{u}^k - u^k)^T \beta g(u^k, \tilde{u}^k) + (u^k - u^*)^T \beta g(u^k, \tilde{u}^k)$$

By using (2.1), namely,

$$(u^k - u^*)^T \beta g(u^k, \tilde{u}^k) \ge (u^k - \tilde{u}^k)^T \beta (M u^k + q),$$

we get

$$\begin{aligned} &(\tilde{u}^k - u^*)^T \beta g(u^k, \tilde{u}^k) \\ &= \quad (\tilde{u}^k - u^k)^T \beta g(u^k, \tilde{u}^k) + (u^k - u^*)^T \beta g(u^k, \tilde{u}^k) \\ &\ge \quad (\tilde{u}^k - u^k)^T \beta g(u^k, \tilde{u}^k) + (u^k - \tilde{u}^k)^T \beta (M u^k + q). \end{aligned}$$

Consequently, by using the notation of $g(u^k, \tilde{u}^k)$, we get

$$\begin{aligned}(\tilde{u}^k - u^*)^T \beta g(u^k, \tilde{u}^k) \quad &\geq \quad (\tilde{u}^k - u^k)^T \{\beta g(u^k, \tilde{u}^k) - \beta(Mu^k + q)\} \\ &= \quad -\beta(u^k - \tilde{u}^k)^T M^T (u^k - \tilde{u}^k). \end{aligned} \tag{2.25}$$

Adding (2.24) and (2.25), it follows that

$$(u_{_{BLD}}^{k+1}(\alpha) - u^*)^T \beta g(u^k, \tilde{u}^k) \geq (u_{_{BLD}}^{k+1}(\alpha) - \tilde{u}^k)^T (I + \beta M^T)(u^k - \tilde{u}^k) + \|u^k - \tilde{u}^k\|^2.$$

Substituting the above inequality in (2.22) and using the notation of $q_k^{_L}(\alpha)$, we get

$$\begin{aligned}\zeta_k^{_L}(\alpha) \quad &\geq \quad \|u_{_{BLD}}^{k+1}(\alpha) - u^k\|^2 + 2\alpha(u_{_{BLD}}^{k+1}(\alpha) - u^k)^T (I + \beta M^T)(u^k - \tilde{u}^k) \\ &\qquad +2\alpha\|u^k - \tilde{u}^k\|^2 \\ &= \quad \|(u_{_{BLD}}^{k+1}(\alpha) - u^k) + \alpha(I + \beta M^T)(u^k - \tilde{u}^k)\|^2 \\ &\qquad -\alpha^2\|(I + \beta M^T)(u^k - \tilde{u}^k)\|^2 + 2\alpha\|u^k - \tilde{u}^k\|^2 \\ &= \quad \|u_{_{BLD}}^{k+1}(\alpha) - [u^k - \alpha(I + \beta M^T)(u^k - \tilde{u}^k)]\|^2 + q_k^{_L}(\alpha).\end{aligned}$$

Because $[u^k - \alpha(I + \beta M^T)(u^k - \tilde{u}^k)] = u_{BD}^{k+1}(\alpha)$ (see (2.8)), the last inequality is the same as (2.20). The proof is complete. $\square$

Theorem 2.3 tells us that $q_k^L(\alpha)$ is also a lower bound of $\zeta_k^L(\alpha)$. In the practical computation, whether the correction formula

$$\text{(PC-I)} \quad u_{BD}^{k+1} = u^k - \gamma\alpha_k^*(I + \beta M^T)(u^k - \tilde{u}^k) \tag{2.26}$$

or

$$\text{(PC-II)} \quad u_{BLD}^{k+1} = P_\Omega\big[u^k - \gamma\alpha_k^*\beta[M^T(u^k - \tilde{u}^k) + (Mu^k + q)]\big] \tag{2.27}$$

is used to update the new iterate u^{k+1}, both the step length α_k^* is given by (2.13), which is lower bounded from 0.

An effective iterative algorithm, the step size must be lower bounded from 0.

Using the formula (2.26) to update u^{k+1}, the advantage is that the correction does not need to do an extra projection. However, in many practical problems, the cost of the projection on Ω (e. g., $\Omega = \Re^n_+$ or a 'box') is not expensive. Thus, the correction formula (2.27) is often used. The reasons are explained in paper [10].

Based on the pair of twin directions offered by (2.4), the different algorithms (2.26) and (2.27), we get the following key inequalities for the contraction:

$$\vartheta_k^L(\alpha) = \|u^k - u^*\|^2 - \|u_{BD}^{k+1}(\alpha) - u^*\|^2 \;\geq\; q_k^L(\alpha),$$

$$\begin{aligned}\zeta_k^L(\alpha) &= \|u^k - u^*\|^2 - \|u_{BLD}^{k+1}(\alpha) - u^*\|^2 \\ &\geq\; q_k^L(\alpha) + \|u_{BLD}^{k+1}(\alpha) - u_{BD}^{k+1}(\alpha)\|^2,\end{aligned}$$

where $q_k^L(\alpha)$ is defined by (2.12) and reaches its maximum at α_k^* (see (2.13)).

The PC Algorithms for solving LVI, PC-I (2.26) and PC-II (2.27), are published in [3] and [4], respectively. Both the algorithms are successfully applied to robot motion planning and real-time control by Zhang and his students [2, 11].

In [2, 11], PC-I (2.26) is called **94LVI** because the method was published in 1994 and the title of the article is ‘A new method for a class of **L**inear **V**ariational **I**nequalities”; PC-II (2.27) is named as **E47** since it is described by **E**quations (**4**) to (**7**) in the original article [4]. The numerical experiments in [2, 11] verified that PC-II (2.27) is more efficient than PC-I (2.26).

3 A pair of twin PC Algorithms for NVI

For nonlinear variational inequalities (abbreviated NVI), we assume its operator F is Lipschitz continuous. In the projection for getting the predictor $\tilde{u}^k$ (1.2), the parameter β_k is chosen such that

$$\beta_k \|F(u^k) - F(\tilde{u}^k)\| \le \nu \|u^k - \tilde{u}^k\|, \quad \nu \in (0,1). \tag{3.1}$$

In the following, for analysis convenience, we omit the index k in β_k and assume that this β satisfies the condition (3.1).

3.1 The ascent directions provided by the predictor

- **The ascent direction provided by FI1+FI3** Adding (1.5) and (1.7), it follows that

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

  From the last inequality, we get

  $$(u^k - u^*)^T \beta F(\tilde{u}^k) \ge (u^k - \tilde{u}^k)^T \beta F(\tilde{u}^k). \tag{3.2}$$

When $u^k \in \Omega$, according to (1.4), we have

$$(u^k - \tilde{u}^k)^T \beta F(u^k) \geq \|u^k - \tilde{u}^k\|^2.$$

According to the last inequality and the assumption (3.1), it follows from the Cauchy-Schwarz inequality that

$$\begin{array}{ll} (u^k - \tilde{u}^k)^T \beta F(\tilde{u}^k) & \\ \quad = & (u^k - \tilde{u}^k)^T \beta F(u^k) - (u^k - \tilde{u}^k)^T \beta (F(u^k) - F(\tilde{u}^k)) \\ \quad \geq & (1 - \nu)\|u^k - \tilde{u}^k\|^2. \end{array} \tag{3.3}$$

The inequalities (3.2) and (3.3) tell us, for $u^k \in \Omega$, under the assumption (3.1), $\beta F(\tilde{u}^k)$ is an ascent direction of the unknown distance function $\frac{1}{2}\|u - u^*\|^2$ at the point u^k. **(3.2)-(3.3) is true only for $\boldsymbol{u^k \in \Omega}$.**

- **The ascent direction provided by FI1+FI2+ FI3**. Adding the fundamental inequalities (1.5), (1.6) and (1.7), we get

$$\{(u^k - u^*) - (u^k - \tilde{u}^k)\}^T d(u^k, \tilde{u}^k) \geq 0,$$

where

$$d(u^k, \tilde{u}^k) = (u^k - \tilde{u}^k) - \beta[F(u^k) - F(\tilde{u}^k)]. \quad (3.4)$$

Consequently, we have

$$(u^k - u^*)^T d(u^k, \tilde{u}^k) \ge (u^k - \tilde{u}^k)^T d(u^k, \tilde{u}^k). \quad (3.5)$$

According to the notation $d(u^k, \tilde{u}^k)$ (3.4) and the assumption (3.1), it follows from the Cauchy-Schwarz inequality that

$$(u^k - \tilde{u}^k)^T d(u^k, \tilde{u}^k) \ge (1-\nu)\|u^k - \tilde{u}^k\|^2. \quad (3.6)$$

The inequalities (3.5) and (3.6) tell us, under the assumption (3.1), $d(u^k, \tilde{u}^k)$ is an ascent direction of the unknown distance function $\frac{1}{2}\|u - u^*\|^2$ at the point u^k. **(3.5)-(3.6) is true for any $u^k \in \Re^n$.**

A pair of the twin directions Notice that the inequality (1.4) is derived from the basic property of the projection. Adding the term

$$(u - \tilde{u}^k)^T\{-\beta[F(u^k) - F(\tilde{u}^k)]\},$$

to the both sides of (1.4), we get

$$\tilde{u}^k \in \Omega, \quad (u - \tilde{u}^k)^T \underline{\beta F(\tilde{u}^k)} \geq (u - \tilde{u}^k)^T \underline{d(u^k, \tilde{u}^k)}, \quad \forall u \in \Omega, \tag{3.7}$$

where $d(u^k, \tilde{u}^k)$ is given by (3.4). We call the directions

$$\beta F(\tilde{u}^k) \quad \text{and} \quad d(u^k, \tilde{u}^k) = (u^k - \tilde{u}^k) - \beta[F(u^k) - F(\tilde{u}^k)]$$

which lay on the left and right sides of (3.7), respectively, a pair of twin ascent directions for NVI. They are derived from FI1+FI3 and FI1+FI2+FI3, respectively.

3.2 Update the new iterate by the direction due to FI1+FI2+FI3

The correction uses the descent direction (the opposite of the ascent direction) of the distance function to make the new iteration point closer to the solution set. Based on the direction provided by FI1+FI2+FI3, the new iterate is updated by

$$u^{k+1}_{BD}(\alpha) = u^k - \alpha d(u^k, \tilde{u}^k). \tag{3.8}$$

where $d(u^k, \tilde{u}^k)$ is given by (3.4). The lower index 'BD' means 'Bounded

Direction' because $\|d(u^k, \tilde{u}^k)\| < 2\|u^k - \tilde{u}^k\|$. For discussion how to determine the step length α, we denote the output of (2.8) by $u_{BD}^{k+1}(\alpha)$. Let us investigate the α-dependent reduction of the square of the distance

$$\vartheta_k^N(\alpha) := \|u^k - u^*\|^2 - \|u_{BD}^{k+1}(\alpha) - u^*\|^2. \tag{3.9}$$

According to the definition,

$$\begin{aligned}\vartheta_k^N(\alpha) &= \|u^k - u^*\|^2 - \|u^k - u^* - \alpha d(u^k, \tilde{u}^k)\|^2 \\ &= 2\alpha(u^k - u^*)^T d(u^k, \tilde{u}^k) - \alpha^2 \|d(u^k, \tilde{u}^k)\|^2. \end{aligned} \tag{3.10}$$

For any given solution point u^*, (3.10) tell us that $\vartheta_k^N(\alpha)$ is a quadratic function of α. Since u^* is unknown, we can't directly find the maximum of $\vartheta_k^N(\alpha)$. With the help of (3.5), we have

Theorem 3.1 *Let* $u_{BD}^{k+1}(\alpha)$ *be updated by* (3.8) *with* $d(u^k, \tilde{u}^k)$ *given by* (3.4)*. Then for* $\alpha > 0$*, we have*

$$\vartheta_k^N(\alpha) \geq q_k^N(\alpha), \tag{3.11}$$

where

$$q_k^N(\alpha) = 2\alpha(u^k - \tilde{u}^k)^T d(u^k, \tilde{u}^k) - \alpha^2 \|d(u^k, \tilde{u}^k)\|^2. \qquad (3.12)$$

Proof. The assertion derived directly from (3.10) by using (3.5). □

Theorem 3.1 indicates that $q_k^N(\alpha)$ is a lower bound of $\vartheta_k^N(\alpha)$. The quadratic function $q_k^N(\alpha)$ reaches its maximum at

$$\alpha_k^* = \text{argmax}\{q_k^N(\alpha)\} = \frac{(u^k - \tilde{u}^k)^T d(u^k, \tilde{u}^k)}{\|d(u^k, \tilde{u}^k)\|^2}. \qquad (3.13)$$

In the practical computation, similarly as in §2, we take a relaxed factor $\gamma \in [1, 2)$ and updated the new iterate by

$$u_{BD}^{k+1} = u^k - \gamma\alpha_k^* d(u^k, \tilde{u}^k), \qquad (3.14)$$

where $d(u^k, \tilde{u}^k)$ is given by (3.5) and is α_k^* given by (3.13).

Theorem 3.2 *Let* $u^{k+1} = u_{BD}^{k+1}$ *be updated by* (3.14). *Then for any* $\gamma \in (0, 2)$, *we have*

$$\|u^{k+1}-u^*\|^2 \le \|u^k-u^*\|^2 - \frac{1}{2}\gamma(2-\gamma)(1-\nu)\|u^k-\tilde{u}^k\|^2. \quad (3.15)$$

Proof. According to (3.9) and (3.11), the u^{k+1} updated by (3.14) satisfies

$$\|u^{k+1}-u^*\|^2 \le \|u^k-u^*\|^2 - q_k^N(\gamma\alpha_k^*). \quad (3.16)$$

According to the definitions $q_k^N(\alpha)$ and α_k^* (see (3.12) and (3.13)), we get

$$\begin{aligned} q_k^N(\gamma\alpha_k^*) &= 2\gamma\alpha_k^*(u^k-\tilde{u}^k)^T d(u^k,\tilde{u}^k) - \gamma^2(\alpha_k^*)^2\|d(u^k,\tilde{u}^k)\|^2 \\ &= \gamma(2-\gamma)\alpha_k^*(u^k-\tilde{u}^k)^T d(u^k,\tilde{u}^k). \end{aligned}$$

In fact, by using (3.4) and (3.1), we have

$$\begin{aligned} & 2(u^k-\tilde{u}^k)^T d(u^k,\tilde{u}^k) - \|d(u^k,\tilde{u}^k)\|^2 \\ &= d(u^k,\tilde{u}^k)^T\{2(u^k-\tilde{u}^k) - d(u^k,\tilde{u}^k)\} \\ &= \|u^k-\tilde{u}^k\|^2 - \beta_k^2\|[F(u^k)-F(\tilde{u}^k)]\|^2 \; > \; 0. \end{aligned} \quad (3.17)$$

Thus, $\alpha_k^* > \frac{1}{2}$, and consequently, it follows that

$$q_k^N(\gamma\alpha_k^*) > \frac{1}{2}\gamma(2-\gamma)(u^k - \tilde{u}^k)^T d(u^k, \tilde{u}^k).$$

Substituting it in (3.16), we get

$$\|u^{k+1} - u^*\|^2 \leq \|u^k - u^*\|^2 - \frac{1}{2}\gamma(2-\gamma)(u^k - \tilde{u}^k)^T d(u^k, \tilde{u}^k).$$

Using (3.6) to the term $(u^k - \tilde{u}^k)^T d(u^k, \tilde{u}^k)$ in the right hand side of the last inequality, we get the assertion (3.15) and the theorem is proved. □

3.3 Update the new iterate by the direction due to FI1+FI3

The correction step (3.8) in §3.3 takes $d(u^k, \tilde{u}^k)$ as the search direction. In this section, it is replaced by $\beta F(\tilde{u}^k)$ and finished the correction with an additional projection. Namely, we let

$$u_{BLD}^{k+1}(\alpha) = P_\Omega[u^k - \alpha\beta F(\tilde{u}^k)], \tag{3.18}$$

to update the α-dependent new iterate ensured in Ω. We denote the output of (3.18) by $u_{_{BLD}}^{k+1}(\alpha)$. The lower index 'BLD' means 'Boundless Direction', because $F(\tilde{u}^k) \not\to 0$ as $\text{dist}(\tilde{u}^k, \Omega^*) \to 0$. For discussion how to determine the step length α, Let us investigate the α-dependent reduction of the square of the distance

$$\zeta_k^N(\alpha) = \|u^k - u^*\|^2 - \|u_{_{BLD}}^{k+1}(\alpha) - u^*\|^2 \qquad (3.19)$$

which is a function of α. We can not maximize $\zeta_k^N(\alpha)$ directly because it involves the unknown vector u^*. The following theorem indicates that for the same $\alpha > 0$, $\zeta_k^N(\alpha)$ is 'better than $\vartheta_k^N(\alpha)$ in (3.11)

Theorem 3.3 *Let $u_{_{BLD}}^{k+1}(\alpha)$ be updated by* (3.18)*. Then for $\zeta_k^N(\alpha)$ defined in* (3.19) *with any $\alpha > 0$, we have*

$$\zeta_k^N(\alpha) \geq q_k^N(\alpha) + \|u_{_{BLD}}^{k+1}(\alpha) - u_{_{BD}}^{k+1}(\alpha)\|^2, \qquad (3.20)$$

where $q_k^N(\alpha)$, $u_{_{BD}}^{k+1}(\alpha)$ and $u_{_{BLD}}^{k+1}(\alpha)$ are given by (3.12), (3.8) *and* (3.18)*, respectively.*

Proof. First, similarly as (2.21), because $u^{k+1}(\alpha) = P_{\Omega}[u - \alpha\beta F(\tilde{u}^k)]$ and $u^* \in \Omega$, we have (see Fig. 1)

$$\|u^{k+1}_{BLD}(\alpha) - u^*\|^2 \le \|u^k - \alpha\beta F(\tilde{u}^k) - u^*\|^2 - \|u^k - \alpha\beta F(\tilde{u}^k) - u^{k+1}_{BLD}(\alpha)\|^2. \tag{3.21}$$

Hence, by using $\zeta_k^N(\alpha)$ (see (3.19)) and (3.21), we have

$$\begin{aligned}
\zeta_k^N(\alpha) \quad &\ge \quad \|u^k - u^*\|^2 - \|(u^k - u^*) - \alpha\beta F(\tilde{u}^k)\|^2 \\
&\qquad +\|(u^k - u^{k+1}_{BLD}(\alpha)) - \alpha\beta F(\tilde{u}^k)\|^2 \\
&= \quad 2\alpha(u^k - u^*)^T \beta F(\tilde{u}^k) + 2\alpha(u^{k+1}_{BLD}(\alpha) - u^k)^T \beta F(\tilde{u}^k) \\
&\qquad +\|u^k - u^{k+1}_{BLD}(\alpha)\|^2 \\
&= \quad \|u^k - u^{k+1}_{BLD}(\alpha)\|^2 + 2\alpha(u^{k+1}_{BLD}(\alpha) - u^*)^T \beta F(\tilde{u}^k).
\end{aligned} \tag{3.22}$$

Decomposing the cross term of the right hand side of (3.22) in form

$$(u^{k+1}_{BLD}(\alpha) - u^*)^T \beta F(\tilde{u}^k) = (u^{k+1}_{BLD}(\alpha) - \tilde{u}^k)^T \beta F(\tilde{u}^k) + (\tilde{u}^k - u^*)^T \beta F(\tilde{u}^k).$$

Since F is monotone and u^* is a solution of VI(Ω, F), we have

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

Substituting it in the right hand side of (3.22), it follows that

$$\zeta_k^N(\alpha) \geq \|u_{BLD}^{k+1}(\alpha) - u^k\|^2 + 2\alpha(u_{BLD}^{k+1}(\alpha) - \tilde{u}^k)^T \beta F(\tilde{u}^k). \qquad (3.23)$$

Since $u_{BLD}^{k+1}(\alpha) \in \Omega$, replacing the any $u \in \Omega$ in (3.7) by $u_{BLD}^{k+1}(\alpha)$, we get

$$(u_{BLD}^{k+1}(\alpha) - \tilde{u}^k)^T \beta F(\tilde{u}^k) \geq (u_{BLD}^{k+1}(\alpha) - \tilde{u}^k)^T d(u^k, \tilde{u}^k).$$

Substituting it in the right hand side of (3.23),

$$\begin{aligned} \zeta_k^N(\alpha) &\geq \|u_{BLD}^{k+1}(\alpha) - u^k\|^2 + 2\alpha(u_{BLD}^{k+1}(\alpha) - \tilde{u}^k)^T d(u^k, \tilde{u}^k) \\ &= \|u_{BLD}^{k+1}(\alpha) - u^k\|^2 + 2\alpha(u_{BLD}^{k+1}(\alpha) - u^k)^T d(u^k, \tilde{u}^k) \\ &\quad +2\alpha(u^k - \tilde{u}^k)^T d(u^k, \tilde{u}^k). \end{aligned} \qquad (3.24)$$

By using the notations $q_k^N(\alpha)$ (see (3.12)) and $u_{BD}^{k+1}(\alpha)$ (3.8), we get

$$\begin{aligned}\zeta_k^N(\alpha) &\geq \|(u_{BLD}^{k+1}(\alpha) - u^k) + \alpha d(u^k, \tilde{u}^k)\|^2 - \alpha^2 \|d(u^k, \tilde{u}^k)\|^2 \\ &\quad +2\alpha(u^k - \tilde{u}^k)^T d(u^k, \tilde{u}^k) \\ &= \|u_{BLD}^{k+1}(\alpha) - (u^k - \alpha d(u^k, \tilde{u}^k))\|^2 + q_k^N(\alpha) \\ &= \|u_{BLD}^{k+1}(\alpha) - u_{BD}^{k+1}(\alpha)\|^2 + q_k^N(\alpha).\end{aligned}$$

This is just the assertion (3.20) and the theorem is proved. □

Theorem 3.3 tells us that $q_k^N(\alpha)$ is also the lower bound of $\zeta_k^N(\alpha)$. In practical computation, with the same predictor $\tilde{u}^k$ given by (1.2) which satisfied (3.1), the corrector u^{k+1} is updated by

$$\text{(PC Method-I)} \qquad u_I^{k+1} = u_{BD}^{k+1}(\alpha) = u^k - \gamma \alpha_k^* d(u^k, \tilde{u}^k) \qquad (3.25)$$

or

$$\text{(PC Method-II)} \qquad u_{II}^{k+1} = u_{BLD}^{k+1}(\alpha) = P_\Omega[u^k - \gamma \alpha_k^* \beta_k F(\tilde{u}^k)], \qquad (3.26)$$

where α_k^* is given by (3.13), which is lower bounded from $1/2$.

Based on the same predictor, if we use the formula (3.25) to update u^{k+1}, the advantage is that the correction does not need to do an extra projection. However, in many practical problems, the cost of the projection on Ω (for example, $\Omega = \Re^n_+$ or a 'box') is not expensive. Thus, the correction formula (3.26) is often used.

By using theorem 3.2, no matter which of the twin methods is applied, the generated sequence $\{u^k\}$ satisfies

$$\|u^{k+1} - u^*\|^2 \le \|u^k - u^*\|^2 - \frac{1}{2}\gamma(2-\gamma)(1-\nu)\|u^k - \tilde{u}^k\|^2.$$

Based on this inequality, we can prove the following convergence theorem.

Theorem 3.4 *Assume that the operator F in VI(Ω, F) is Lipschitz continuous and its solution set Ω^* is nonempty. Then the sequence $\{u^k\}$ generated by* (3.25) *or* (3.26) *converges to some solution point of VI(Ω, F).*

The projection contraction algorithms introduced in this section have successfully applied to solve many geotechnical engineering problems [12, 13].

4 Applications and numerical experiments

We use examples of linear and nonlinear variational inequalities to illustrate the efficiency of the twin algorithms — PC Method-I and PC Method-II.

4.1 Applied the different PC Methods for LVI

For LVI, we use the ”sum of the shortest distance” mentioned in Lecture 1 as an example. This problem is equivalent to a min-max problem whose corresponding LVI with a skew symmetric matrix. For a detailed description of this kind of problem, see $\S$5 of Lecture 1.
The test examples taking from SIAM J. on Optimization.

- G. L. XUE AND Y. Y. YE, *An efficient algorithm for minimizing a sum of Euclidean norms with applications*, SIAM Optim. 7 (1997), 1017-1039.

Fig.1 depicters the structure of the network, where $b_{[i]}, i = 1, \ldots, 10$ are regular points whose coordinates are given. The connection between $x_{[j]}, j = 1, \ldots, 8$ and $b_{[i]}$ are also given. Fig. 2 gives the positions of $x_{[j]}, j = 1, \ldots, 8$ when the sum of the shortest distance is reached.

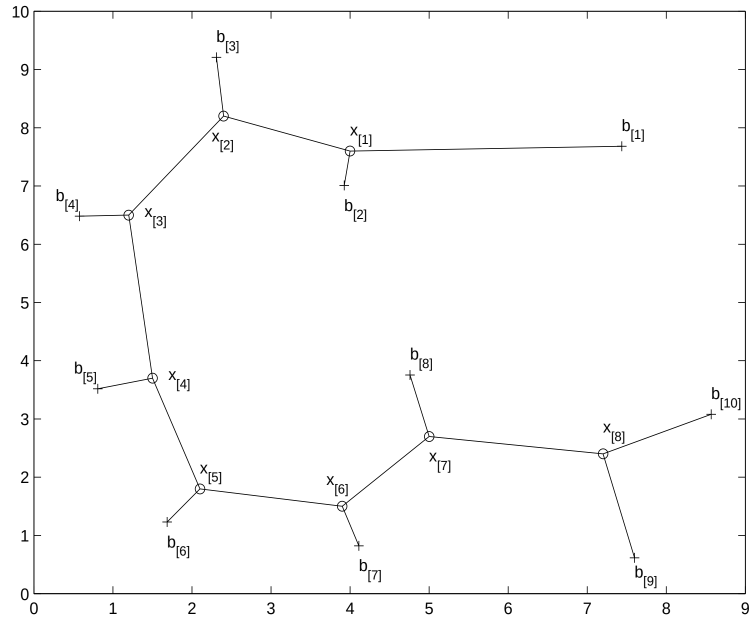

Fig. 1 Ordering of the topology

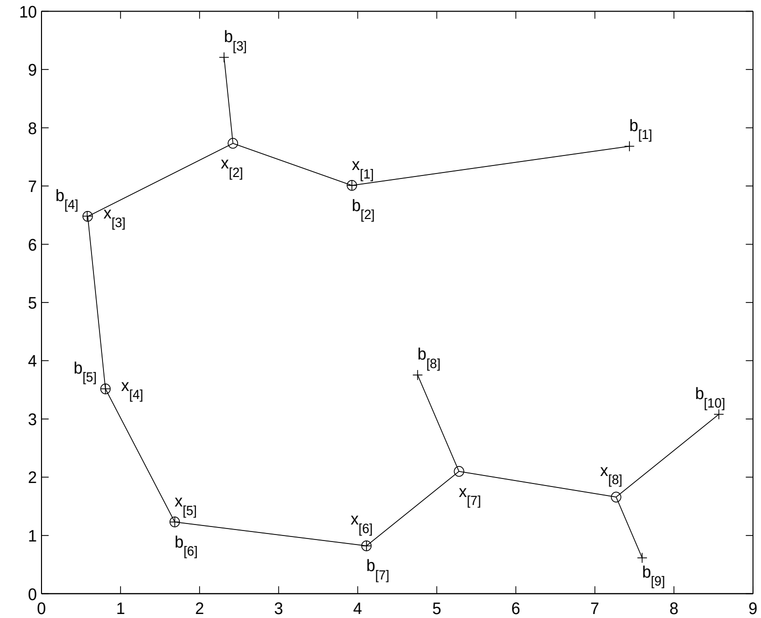

Fig. 2. Optimal solution in Euclidean-norm

The coordinates of the 10 regular points

	x-coordinate	y-coordinate		x-coordinate	y-coordinate
$b_{[1]}$	7.436490	7.683284	$b_{[6]}$	1.685912	1.231672
$b_{[2]}$	3.926097	7.008798	$b_{[7]}$	4.110855	0.821114
$b_{[3]}$	2.309469	9.208211	$b_{[8]}$	4.757506	3.753666
$b_{[4]}$	0.577367	6.480938	$b_{[9]}$	7.598152	0.615836
$b_{[5]}$	0.808314	3.519062	$b_{[10]}$	8.568129	3.079179

The update forms of using the contraction method I (3.25) and II (3.26) are

$$\text{(PC Method-I)} \qquad u^{k+1} = u^k - \gamma\alpha_k^*(I + M^T)(u^k - \tilde{u}^k),$$

and

$$\text{(PC Method-II)} \qquad u^{k+1} = P_\Omega\{u^k - \gamma\alpha_k^*[M^T(u^k - \tilde{u}^k) + (Mu^k + q)]\},$$

respectively. The numerical results are listed in the following table.

Table 1. Shortest network under l_2 norm.

PC Method-I			PC Method-II		
Iteration	$\|e(u)\|_\infty$	Total Distance	Iteration	$\|e(u)\|_\infty$	Total Distance
40	7.1e-002	25.3776304969	40	5.0e-004	25.3563526162
80	1.8e-004	25.3561050662	80	4.0e-008	25.3560677986
120	6.4e-007	25.3560678958	106	9.2e-011	25.3560677793
160	2.4e-009	25.3560677797			
183	9.5e-011	25.3560677793			
CPU-time		0.234 Sec.	CPU-time		0.125 Sec.

Here we take $\gamma = 1.8$. If $\gamma = 1$, 80% more iterations are needed for both methods.

PC Method-II for the problem in the Euclidean-norm

```
clear; % Steiner Minimum Tree * Read the coordinate of the regular points%(1)
 P1=[7.436490, 3.926097, 2.309469, 0.577367, 0.808314;                    %(2)
     7.683284, 7.008798, 9.208211, 6.480938, 3.519062];                   %(3)
 P2=[1.685912, 4.110855, 4.757506, 7.598152, 8.568129;                    %(4)
     1.231672, 0.821114, 3.753666, 0.615836, 3.079179];                   %(5)
 b=[P1,P2,zeros(2,7)]; x=zeros(2,8); z=zeros(2,17);  eps=1;   k=0; tic;  %(6)
while (eps > 10^(-10) & k<= 200)  k=k+1;   %% Beginning of an iteration  %(7)
 Ax=[x(:,1), x, x(:,8), x(:,1:7)-x(:,2:8)];   Axb=Ax-b;  %% Compute Ax-b %(8)
 ATz=z(:,2:9) + [z(:,1), -z(:,11:17)] + [z(:,11:17), z(:,10)];   % A^Tz  %(9)
 L2=0; for j=1:17  L2=L2 + norm(Axb(:,j),2); end;            % Length-2 %(10)
 if mod(k,20)==0 fprintf('k=%3d stopc=%9.1e L2=%13.10f\n',k,eps,L2);end;%(11)
 Pz=z+Axb; Dp=diag(1./max(1,sqrt(diag(Pz'*Pz)))); Pz=Pz*Dp; %P(z+(Ax-b))%(12)
 Ex = ATz;   Ez = z-Pz;  t=trace(Ex'*Ex)+ trace(Ez'*Ez);  eps=sqrt(t);  %(13)
 AEx= [Ex(:,1), Ex, Ex(:,8), Ex(:,1:7)-Ex(:,2:8)];      % Compute AEx   %(14)
 ATEz=Ez(:,2:9) + [Ez(:,1),-Ez(:,11:17)] + [Ez(:,11:17),Ez(:,10)]; %ATEz%(15)
  ta = trace(AEx'*AEx)+trace(ATEz'*ATEz);  alpha=t*1.8/(t+ta); %% Step L%(16)
  x =x-(ATz - ATEz)*alpha;                               %% New x and z %%(17)
  z =z-(AEx - Axb)*alpha; Dz=diag(1./max(1,sqrt(diag(z'*z)))); z=z*Dz; %%(18)
 end;                                            %% End of an iteration %%(19)
 toc; fprintf(' k=%3d    eps=%9.1e   Length-2=%13.10f \n', k,eps,L2);  %%(20)
```

把上面第(18) 行改成 `z =z-(AEx + Ez)*alpha;` 就是处理同一问题的(CM-D1)程序.

Fig. 3 and 4 depict the convergence tendencies of Contraction Method–2 for the minimum sum of the distance in the Euclidean-norm with different starting points.

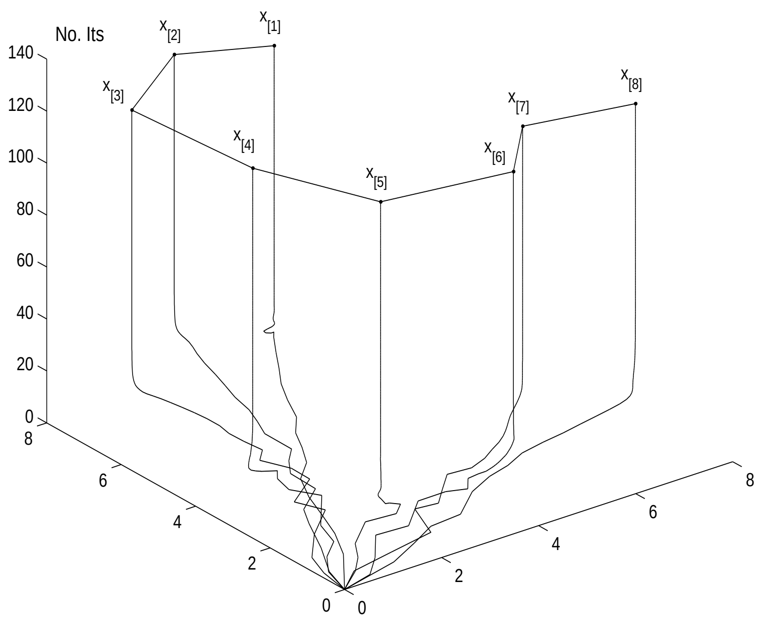

Fig. 3. Convergence tendency, $x^0 = 0$

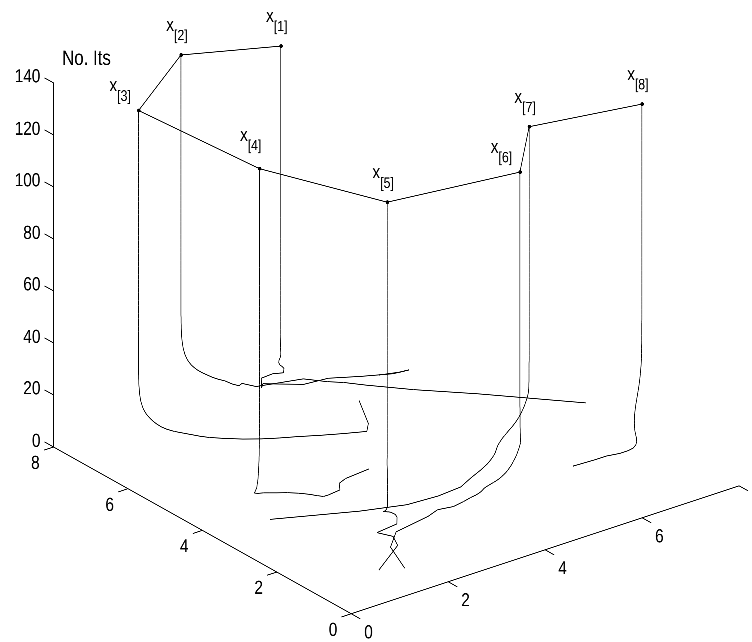

Fig. 4. Convergence tendency, x^0 random

对 l_1-模和 l_∞-模, 我们也用两种不同方法做了计算比较, 计算结果如下：

Table 2. Shortest network under l_1 norm.

PC Method-I			PC Method-II		
Iteration	$\|e(u)\|_\infty$	Total Distance	Iteration	$\|e(u)\|_\infty$	Total Distance
40	3.7e-002	28.6777786413	40	1.1e-004	28.6660178525
80	2.5e-005	28.6658649129	81	1.0e-010	28.6658580000
120	1.8e-008	28.6658580046			
149	9.4e-011	28.6658580000			
CPU-time		0.031 Sec.	CPU-time		0.016 Sec.

Table 3. Shortest network under l_∞ norm.

PC Method-I			PC Method-II		
Iteration	$\|e(u)\|_\infty$	Total Distance	Iteration	$\|e(u)\|_\infty$	Total Distance
40	9.0e-002	21.1322990353	40	2.1e-003	21.1145131146
80	4.4e-005	21.1129244226	80	4.1e-010	21.1129135002
120	2.4e-008	21.1129135060	84	7.4e-011	21.1129135000
150	9.2e-011	21.1129135000			
CPU-time		0.187 Sec.	CPU-time		0.094 Sec.

Fig. 5 and Fig. 6 depict the optimal solutions of the minimum sum of the distance in the l_1-norm and l_∞-norm, respectively.

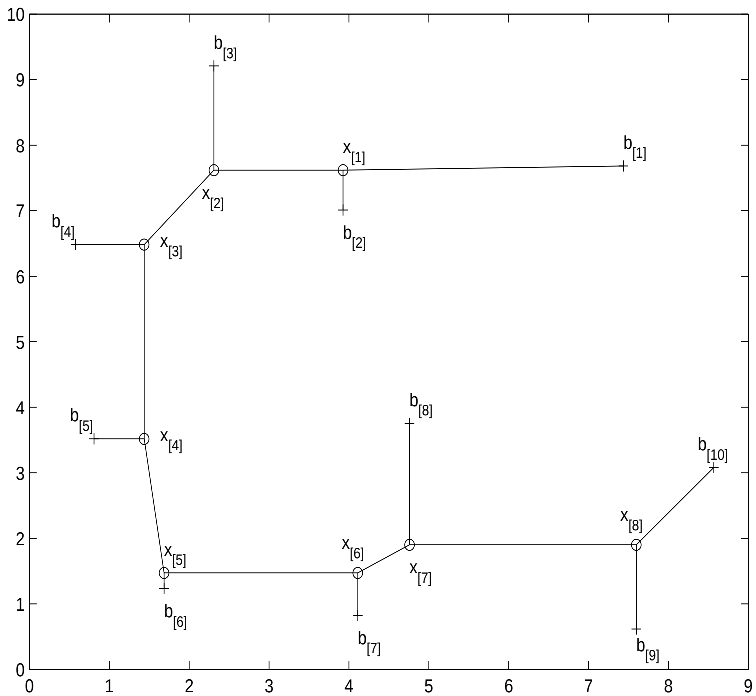

Fig. 5. Optimal solution, l_1-norm

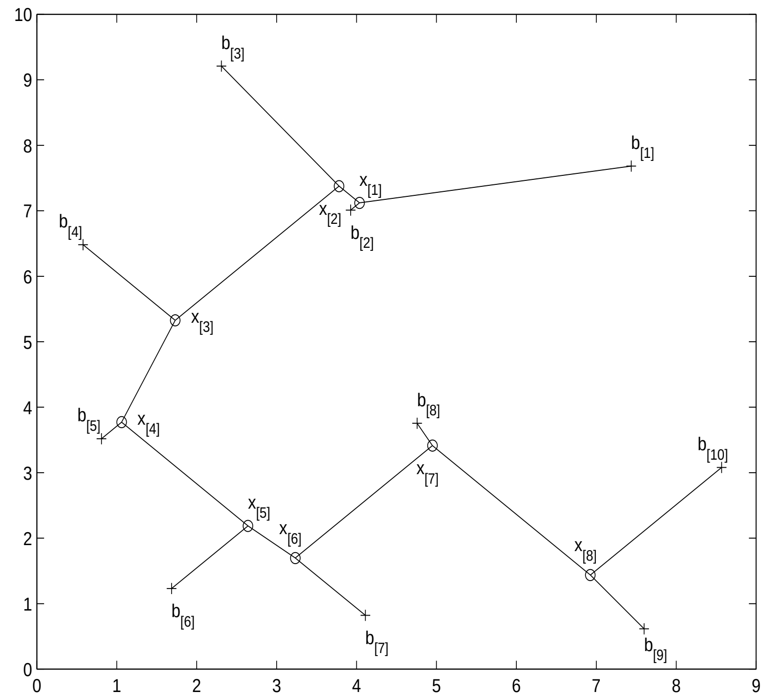

Fig. 6. Optimal solution, l_∞-norm

Fig. 7 and 8 depict the convergence tendencies of Contraction Method–2 with random starting points for the minimum sum of the distance in the in the l_1-norm and l_∞-norm, respectively.

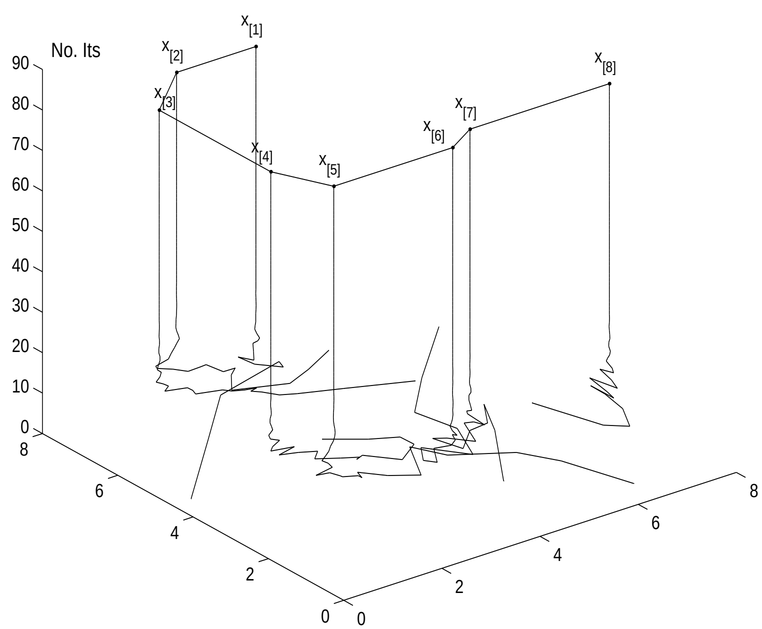

Fig. 7. Convergence tendency, l_1-norm

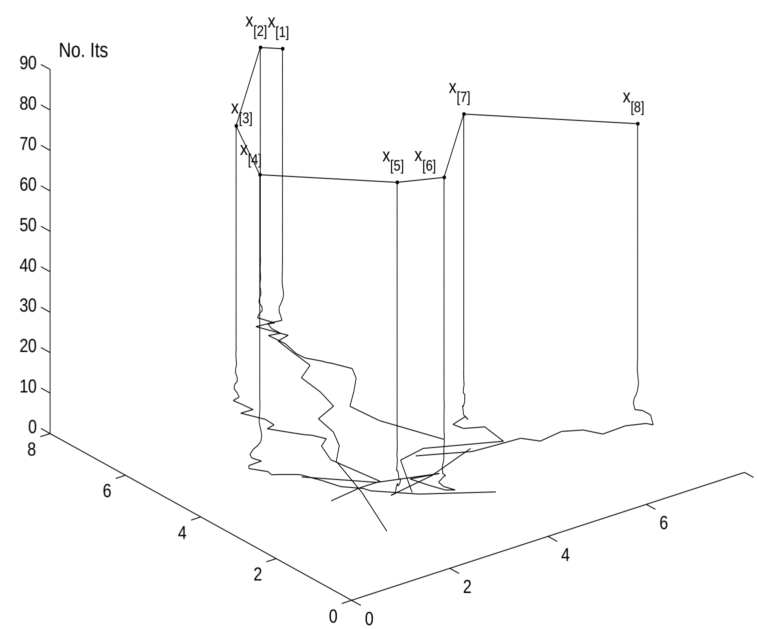

Fig. 8. Convergence tendency, l_∞-norm

4.2 Test examples of the NCP

For comparing the efficiency of PC Method-I and PC Method-II, we test the nonlinear complementarity problem (a class of VI(Ω, F) with $\Omega = \Re^n_+$)

$$u \geq 0, \qquad F(u) \geq 0, \qquad u^T F(u) = 0.$$

In the test examples, we take

$$F(u) = D(u) + Mu + q,$$

The linear part $Mu + q$ is generated by using Matlab, it produced by

```
A=(rand(n,n)-0.5)*10; B=(rand(n,n)-0.5)*10;  B=B-B'; M=A'*A+B;
q=(rand(n,1)-0.5)*1000;        or        q=(rand(n,1)-1.0)*500;
```

In the nonlinear part $D(u)$, each element is given by $D_j(u) = d_j * \arctan(u_j)$, where d_j is a random variable between $(0, 1)$.

We use the algorithms (3.25) and (3.26) in §3.3 to solve the test problems.

Notice that PC Method-I is just the PC Algorithm in Lecture 2 for NVI.

Set $\gamma\alpha_k^* \equiv 1$ in PC Method-II, it reduced to the Refined EG.

In all the tests, each element of the initial u^0 is a random variable in $(0, 10)$.

PC Method-II 的程序

PC Method-II:

Step 0. Set $\beta_0 = 1$, $\nu \in (0,1)$, $u^0 \in \Omega$ and $k = 0$.

Step 1. $\tilde{u}^k = P_\Omega[u^k - \beta_k F(u^k)]$,

$r_k := \beta_k \|F(u^k) - F(\tilde{u}^k)\| / \|u^k - \tilde{u}^k\|$,

while $r_k > \nu$, $\beta_k := \frac{2}{3}\beta_k * \min\{1, \frac{1}{r_k}\}$,

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

$r_k := \beta_k \|F(u^k) - F(\tilde{u}^k)\| / \|u^k - \tilde{u}^k\|$,

end(while)

$$d(u^k, \tilde{u}^k) = (u^k - \tilde{u}^k) - \beta_k[F(u^k) - F(\tilde{u}^k)],$$

$$\alpha_k = \frac{(u^k - \tilde{u}^k)^T d(u^k, \tilde{u}^k)}{\|d(u^k, \tilde{u}^k)\|^2},$$

$$u^{k+1} = P_\Omega[u^k - \gamma \alpha_k \beta_k F(\tilde{u}^k)],$$

If $r_k \le \mu$ then $\beta_k := \beta_k * 1.5$, **end(if)**

Step 2. $\beta_{k+1} = \beta_k$ and $k = k+1$, go to Step 1.

From PC Method-I to PC Method-II, 只是将

$u^{k+1} = u^k - \gamma \alpha_k d(u^k, \tilde{u}^k)$ 改成了 $u^{k+1} = P_\Omega[u^k - \gamma \alpha_k \beta_k F(\tilde{u}^k)]$.

Matlab Code of Contraction Method–D2 for NCP

```
function PC_G(n,M,q,d,xstart,tol,pfq)                                    %(1)
fprintf('PC Method use Direction D1 with gamma a*  n=%4d \n',n);         %(2)
x=xstart;  Fx= d.*atan(x) + M*x + q;  stopc=norm(x-max(x-Fx,0),inf);     %(3)
beta=1;        k=0;         l=0;         tic;                            %(4)
while (stopc>tol && k<=2000)                                             %(5)
  if mod(k,pfq)==0 fprintf(' k=%4d    epsm=%9.3e  \n',k,stopc); end;     %(6)
    x0=x;   Fx0=Fx;    k=k+1;                                            %(7)
    x=max(x0-Fx0*beta,0);        Fx=d.*atan(x) + M*x + q;    l=l+1;      %(8)
       dx=x0-x;   df=(Fx0-Fx)*beta;                                      %(9)
       r=norm(df)/norm(dx);                                              %(10)
       while r>0.9     beta=0.7*beta*min(1,1/r);   l=l+1;                %(11)
         x=max(x0-Fx0*beta,0);    Fx=d.*atan(x) + M*x + q;               %(12)
         dx=x0-x;   df=(Fx0-Fx)*beta;     r=norm(df)/norm(dx);           %(13)
       end;                                                              %(14)
       dxf=dx-df;    r1=dx'*dxf;    r2=dxf'*dxf;     alpha=r1/r2;        %(15)
       x=max(x0- Fx*beta*alpha*1.9,0);                                   %(16)
       Fx= d.*atan(x) + M*x + q;     l=l+1;                              %(17)
       ex=x-max(x-Fx,0);     stopc=norm(ex,inf);                         %(18)
       if r <0.4  beta=beta*1.5;  end;                                   %(19)
end;  toc;   fprintf(' k=%4d   epsm=%9.3e     l=%4d  \n',k,stopc,l); %%%%
```

NCP 的计算结果 1　Easy Problems $q \in (-500, 500)$

	PC Method-I			PC Method-II		
$n =$	No. It	No. F	CPU	No. It	No. F	CPU
500	448	941	0.15	372	792	0.12
1000	475	995	1.37	410	852	1.17
1500	507	1064	3.17	416	887	2.64
2000	515	1080	5.53	418	892	4.55

NCP 的计算结果 2　Hard Problems $q \in (-500, 0)$

	PC Method-I			PC Method-II		
$n =$	No. It	No. F	CPU	No. It	No. F	CPU
500	908	1913	0.30	799	1704	0.27
1000	980	2068	2.87	857	1824	2.53
1500	941	1983	5.88	834	1771	5.25
2000	1112	2352	12.18	986	2105	10.87

PC Method-II converges faster than PC Method-I.

♣ 程序在附件的 Codes-03 中：运行 demo.m 输入 n 就可以, 其中也可以选择不同问题类型. PCd1.m 和 PCd2.m 分别是 PC Method-I 和 PC Method-II 的子程序.

5 更一般的孪生方法的统一框架

我们在 [8, 9] 中给出了求解变分不等式更一般的孪生收缩算法.

首先, 定义了预测点(或称检验点) . 对给定的 u^k, 根据一定法则生成的 $\tilde{u}^k \in \Omega$ 说成是一个预测点, 如果 $\boxed{u^k = \tilde{u}^k \quad \Leftrightarrow \quad u^k \in \Omega^*.}$

例如, 对给定的 u^k 和 $\beta > 0$, 由投影 $\boxed{\tilde{u}^k = P_\Omega[u^k - \beta F(u^k)]}$ 给出的 $\tilde{u}^k$ 是按确定的法则给出的, 它是一个预测点, 但这不是给出预测点的惟一方法.

统一框架. 对给定的 u^k, 设 $\tilde{u}^k \in \Omega$ 是 u^k 的一个预测点. 设有基于 $(u^k, \tilde{u}^k)$ 的一对孪生的方向 $d_1(u^k, \tilde{u}^k)$, $d_2(u^k, \tilde{u}^k)$ 和误差度量函数 $\varphi(u^k, \tilde{u}^k) \ge 0$, 它们满足以下条件:

1. 它们满足关系式

$$\tilde{u}^k \in \Omega, \quad (u - \tilde{u}^k)^T d_2(u^k, \tilde{u}^k) \ge (u - \tilde{u}^k)^T d_1(u^k, \tilde{u}^k), \quad \forall u \in \Omega. \tag{5.1a}$$

2. 存在常数 $K > 0$, 使得

$$\|d_1(u^k, \tilde{u}^k)\| \le K\|u^k - \tilde{u}^k\|. \tag{5.1b}$$

3. 对任意的 $u^* \in \Omega^*$, 有

$$(\tilde{u}^k - u^*)^T d_2(u^k, \tilde{u}^k) \geq \varphi(u^k, \tilde{u}^k) - (u^k - \tilde{u}^k)^T d_1(u^k, \tilde{u}^k), \qquad \text{(5.1c)}$$

4. $\varphi(u^k, \tilde{u}^k)$ 是 VI(Ω, F) 的误差度量函数, 即存在常数 $\delta > 0$, 使得

$$\varphi(u^k, \tilde{u}^k) \geq \delta \|u^k - \tilde{u}^k\|^2 \quad \& \quad \varphi(u^k, \tilde{u}^k) = 0 \Leftrightarrow u^k = \tilde{u}^k. \qquad \text{(5.1d)}$$

对误差度量函数 $\varphi(u^k, \tilde{u}^k)$ 而言, $d_1(u^k, \tilde{u}^k), d_2(u^k, \tilde{u}^k)$ 都是有利方向.

Lemma 5.1 如果统一框架中的条件 (5.1a) 和 (5.1c) 满足, 则有

$$(u^k - u^*)^T d_1(u^k, \tilde{u}^k) \geq \varphi(u^k, \tilde{u}^k), \quad \forall u^k \in \Re^n, u^* \in \Omega^*. \qquad \text{(5.2)}$$

证明. 因为 $u^* \in \Omega$, 以 u^* 代 (5.1a) 中的 u, 就有

$$(\tilde{u}^k - u^*)^T d_1(u^k, \tilde{u}^k) \geq (\tilde{u}^k - u^*)^T d_2(u^k, \tilde{u}^k).$$

再根据条件 (5.1c), 得到

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

从上式直接得到 (5.2), 引理得证. $\square$

Lemma 5.2 如果统一框架中的条件 (5.1a) 和 (5.1c) 满足, 则有

$$(u^k - u^*)^T d_2(u^k, \tilde{u}^k) \ge \varphi(u, \tilde{u}), \quad \forall u^k \in \Omega, u^* \in \Omega^*. \tag{5.3}$$

证明. 因为 $u^k \in \Omega$, 以 u^k 代 (5.1a) 中的 u, 我们有

$$(u^k - \tilde{u}^k)^T d_2(u^k, \tilde{u}^k) \ge (u^k - \tilde{u}^k)^T d_1(u^k, \tilde{u}^k). \tag{5.4}$$

将 (5.4) 和 (5.1c) 相加, 得到

$$(u^k - u^*)^T d_2(u^k, \tilde{u}^k) \ge \varphi(u^k, \tilde{u}^k).$$

引理得证. □

根据提供的 $d_1(u^k, \tilde{u}^k)$ 和 $d_2(u^k, \tilde{u}^k)$ 一对孪生方向, 我们可以构造一对算法

(Contraction Method-I) $\quad u^{k+1} = u^k - \gamma \alpha_k^* d_1(u^k, \tilde{u}^k),$

(Contraction Method-II) $\quad u^{k+1} = P_\Omega\{u^k - \gamma \alpha_k^* d_2(u^k, \tilde{u}^k)\},$

其中 $\alpha_k^* = \varphi(u^k, \tilde{u}^k)/\|d_1(u^k, \tilde{u}^k)\|^2$, $\gamma \in (0, 2)$, 由(5.1b)和(5.1d), 步长是有界的.

孪生方向, 相同步长, 是 PC 方法中最优美的篇章. 证明需技巧, 使用很方便!

References

[1] E. Blum and W. Oettli, Mathematische Optimierung. Grundlagen und Verfahren. Ökonometrie und Unternehmensforschung. Berlin-Heidelberg-New York: Springer-Verlag, 1975.

[2] D. S. Guo and Y.N. Zhang, Simulation and experimental verification of weighted velocity and acceleration minimization for robotic redundancy resolution, *IEEE Transactions on Automation Science and Engineering*, 2014, 11: 1203–1217.

[3] B.S. He, A new method for a class of linear variational inequalities, *Math. Progr.*, **66**, 137–144, 1994.

[4] B.S. He, Solving a class of linear projection equations, *Numerische Mathematik*, **68**, 71–80, 1994.

[5] B.S. He, A globally linearly convergent projection and contraction method for a class of linear complementarity problems. Schwerpunktprogramm der DFG Anwendungsbezogene Optimierung und Steuerung, No. 352, 1992

[6] B.S. He, A class of projection and contraction methods for monotone variational inequalities, *Applied Mathematics and optimization*, **35**, 69–76, 1997.

[7] B.S He and L.-Z Liao, Improvements of some projection methods for monotone nonlinear variational inequalities, JOTA, **112**, 111-128, 2002

[8] B.S. He, L.Z. Liao, and X. Wang, Proximal-like contraction methods for monotone variational inequalitiesin a unified framework I: Effective quadruplet and primary methods, Comput. Optim. Appl., **51**, 649-679, 2012.

[9] B.S. He, L.Z. Liao, and X. Wang, Proximal-like contraction methods for monotone variational inequalities in a unified framework II: General methods and numerical experiments, Comput. Optim. Appl. **51**, 681-708, 2012

[10] B.S. He, X.M. Yuan and J.J.Z. Zhang, *Comparison of two kinds of prediction-correction methods for monotone variational inequalities*, Computational Optimization and Applications, **27**, 247-267, 2004.

[11] L. Xiao and Y. N. Zhang, Acceleration-level repetitive motion planning and its experimental verification on six-link planar robot manipulator, *IEEE Transactions on Control System Technology*, 2013, 21: 906–914.

[12] H. Zheng, F. Liu and X.L. Du, Complementarity problem arising from static growth of multiple cracks and MLS-based numerical manifold method, Computer Methods in Applied Mechanics and Engineering, 295 (2015) 150-171.

[13] H. Zheng, P. Zhang and X.L. Du, Dual form of discontinuous deformation analysis, Computer Methods in Applied Mechanics and Engineering, 305 (2016) 196-216.