Available online at www.sciencedirect.com
SCIENCE DIRECT®

J. Math. Anal. Appl. 300 (2004) 362–374

Journal of
MATHEMATICAL
ANALYSIS AND
APPLICATIONS

www.elsevier.com/locate/jmaa

An approximate proximal-extragradient type method for monotone variational inequalities

Bing-sheng He [a,*,1], Zhen-hua Yang [a], Xiao-ming Yuan [b]

[a] *Department of Mathematics, Nanjing University, Nanjing, 210093, PR China*
[b] *Department of Mathematics, City University of Hong Kong, Hong Kong*

Received 8 June 2003

Available online 16 September 2004

Submitted by M.A. Noor

Abstract

Proximal point algorithms (PPA) are attractive methods for monotone variational inequalities. The approximate versions of PPA are more applicable in practice. A modified approximate proximal point algorithm (APPA) presented by Solodov and Svaiter [Math. Programming, Ser. B 88 (2000) 371–389] relaxes the inexactness criterion significantly. This paper presents an extended version of Solodov–Svaiter's APPA. Building the direction from current iterate to the new iterate obtained by Solodov–Svaiter's APPA, the proposed method improves the profit at each iteration by choosing the *optimal* step length along this direction. In addition, the inexactness restriction is relaxed further. Numerical example indicates the improvement of the proposed method.

Keywords: Monotone variational inequality; Inexact proximal point algorithm

* Corresponding author.
E-mail address: hebma@nju.edu.cn (B.-S. He).
[1] The author was supported by NSFC grant 10271054, MOEC grant 20020284027, and Jiangsu NSF grant BK2002075.

0022-247X/$ – see front matter
doi:10.1016/j.jmaa.2004.04.068

1. Introduction

Let C be a nonempty closed convex subset of R^n and F be a continuous monotone mapping from R^n into itself. The variational inequality problem is to determine a vector $x^* \in C$ such that

$$\mathrm{VI}(C, F) \quad (x - x^*)^T F(x^*) \geqslant 0, \quad \forall x \in C. \tag{1.1}$$

$\mathrm{VI}(C, F)$ includes nonlinear complementarity problems (when $C = R^n_+$) and systems of nonlinear equations (when $C = R^n$), and thus it has many important applications, e.g., see [6,8,9].

A classical method for solving variational inequality is the proximal point algorithm (abbreviated as PPA) [11,12]. Let $\lambda_{\min} > 0$ and $\{\lambda_k\} \subset [\lambda_{\min}, +\infty)$. For given $x^k \in C$ and λ_k, let x_*^{k+1} be the solution of following strongly monotone variational inequality:

$$\text{(PPA)} \quad \text{Find } x \in C \text{ such that } (x' - x)^T \bigl\{\bigl(x - x^k\bigr) + \lambda_k F(x)\bigr\} \geqslant 0, \quad \forall x' \in C. \tag{1.2}$$

The new iterate x^{k+1} of the exact version of PPA is taken by

$$x^{k+1} := x_*^{k+1}. \tag{1.3}$$

An equivalent recursion form of the exact PPA is

$$x^{k+1} = P_C\bigl[x^k - \lambda_k F\bigl(x^{k+1}\bigr)\bigr], \tag{1.4}$$

where P_C denotes the projection on C.

The ideal form (1.4) of PPA is often impractical since in many cases solving problem (1.2) exactly is either impossible or expensive. In 1976 Rockafellar [11,12] set up the fundamental convergence analysis for the approximate proximal point algorithm (abbreviated as APPA) to a general maximal monotone operator. The new iterate x^{k+1} of Rockafellar's APPAs is requested to satisfy condition

$$\bigl\|x^{k+1} - x_*^{k+1}\bigr\| \leqslant \nu_k, \qquad \sum_{k=0}^{\infty} \nu_k < +\infty \tag{1.5}$$

or

$$\bigl\|x^{k+1} - x_*^{k+1}\bigr\| \leqslant \nu_k \bigl\|x^k - x^{k+1}\bigr\|, \qquad \sum_{k=0}^{\infty} \nu_k < +\infty. \tag{1.6}$$

Since x_*^{k+1} is necessarily unknown, some upper bounds of $\|x^{k+1} - x_*^{k+1}\|$ are needful to be evaluated in order to implement such APPAs.

Extensive developments on APPA focus on different fields such as convex programming, mini-max problems, and variational inequality problems. To mention a few, see [1, 3–5,13]. The major challenges of APPA include accelerating convergence and designing inexactness restrictions that are easy to implement and tight for convergence.

Throughout this paper we assume that the operator F is monotone and Lipschitz continuous on C and that the solution set of $\mathrm{VI}(\Omega, F)$, denoted by C^*, is nonempty. We use x^* to denote any point in C^*.

2. Preliminaries and motivation

We use the concept of projection under the Euclidean norm, which is denoted by $P_C(\cdot)$, i.e.,

$$P_C(z) = \operatorname{argmin}\{\|z - x\| \mid x \in C\}.$$

It follows from this definition that

$$\{z - P_C(z)\}^T\{y - P_C(z)\} \leqslant 0, \quad \forall z \in R^n,\ \forall y \in C. \tag{2.1}$$

Consequently, we have

$$\|P_C(y) - P_C(z)\| \leqslant \|y - z\|, \quad \forall y, z \in R^n \tag{2.2}$$

and

$$\|P_C(y) - x\|^2 \leqslant \|y - x\|^2 - \|y - P_C(y)\|^2, \quad \forall y \in R^n,\ \forall x \in C. \tag{2.3}$$

Lemma 2.1 [2, p. 267]. *Let $\lambda > 0$, then x^* solves* VI(C, F) *if and only if*

$$x^* = P_C[x^* - \lambda F(x^*)].$$

Denote

$$e(x, \lambda) := x - P_C[x - \lambda F(x)]. \tag{2.4}$$

Then solving VI(C, F) is equivalent to finding a zero point of $e(x, \lambda)$. For given x, it is well known [7] that $\|e(x, \lambda)\|$ is a non-decreasing function of λ.

The following lemma can be viewed as a corollary of [14, Proposition 3.4]. It gives us an upper bound of $\|x - x_*^{k+1}\|$ for any $x \in C$.

Lemma 2.2. *Let F be monotone on C and x_*^{k+1} be the unique solution of the strongly monotone* VI *sub-problem* (1.2). *Then we have*

$$\Delta(x) \geqslant \|x - x_*^{k+1}\|^2, \quad \forall x \in C, \tag{2.5}$$

where

$$\begin{aligned} \Delta(x) := {} & 2\{x - P_C[x^k - \lambda_k F(x)]\}^T\{(x - x^k) + \lambda_k F(x)\} \\ & - \|x - P_C[x^k - \lambda_k F(x)]\|^2. \end{aligned} \tag{2.6}$$

Proof. For any fixed $x \in C$, we define

$$h(x, z) := (\|(x - x^k) + \lambda_k F(x)\|^2 - \|x^k - \lambda_k F(x) - z\|^2)$$

and it follows that

$$h(x, P_C[x^k - \lambda_k F(x)]) = \max\{h(x, z) \mid z \in C\}.$$

Since $x_*^{k+1} \in C$, we have

$$h(x, P_C[x^k - \lambda_k F(x)]) \geqslant h(x, x_*^{k+1}),$$

and this yields

$$2\{x - P_C[x^k - \lambda_k F(x)]\}^T \{(x - x^k) + \lambda_k F(x)\} - \|x - P_C[x^k - \lambda_k F(x)]\|^2$$
$$\geqslant 2(x - x_*^{k+1})^T \{(x - x^k) + \lambda_k F(x)\} - \|x - x_*^{k+1}\|^2. \tag{2.7}$$

On the other hand, since x_*^{k+1} is the solution of (1.2) and $x \in C$, we have

$$(x - x_*^{k+1})^T \{(x_*^{k+1} - x^k) + \lambda_k F(x_*^{k+1})\} \geqslant 0$$

and consequently

$$\begin{aligned}
&(x - x_*^{k+1})^T \{(x - x^k) + \lambda_k F(x)\} \\
&\quad \geqslant (x - x_*^{k+1})^T \{[(x - x^k) + \lambda_k F(x)] - [(x_*^{k+1} - x^k) + \lambda_k F(x_*^{k+1})]\} \\
&\quad = (x - x_*^{k+1})^T \{(x - x_*^{k+1}) + \lambda_k (F(x) - F(x_*^{k+1}))\} \\
&\quad \geqslant \|x - x_*^{k+1}\|^2,
\end{aligned} \tag{2.8}$$

where the last inequality comes from the monotonicity of F. Then the assertion follows from (2.7) and (2.8) directly. □

Let y^k be an approximate solution of (1.2) in the sense that

$$y^k \approx P_C[x^k - \lambda_k F(y^k)] \tag{2.9}$$

and define

$$\tilde{y}^k := P_C[x^k - \lambda_k F(y^k)]. \tag{2.10}$$

It follows from Lemma 2.2 that $\Delta(y^k)$ is an upper bound of $\|y^k - x_*^{k+1}\|^2$. In the practical computation of Rockafellar's APPAs [11], the new iterate is taken by

$$x^{k+1} := y^k,$$

and instead of (1.5) and (1.6), it is requested to satisfy

$$\Delta(y^k) \leqslant \nu_k^2, \qquad \sum_{k=0}^{\infty} \nu_k < +\infty \tag{2.11}$$

and

$$\Delta(y^k) \leqslant \nu_k^2 \|x^k - y^k\|^2, \qquad \sum_{k=0}^{\infty} \nu_k < +\infty, \tag{2.12}$$

respectively.

Recently, under a significantly relaxed inexactness restriction (in comparison with (2.12))

$$\Delta(y^k) \leqslant \nu \|x^k - y^k\|^2, \quad \nu < 1, \tag{2.13}$$

Solodov and Svaiter [13] proposed a new APPA in which the new iterate is given by

$$\text{(SS-method)} \quad x^{k+1} = P_C[x^k - \lambda_k F(y^k)]. \tag{2.14}$$

Indeed, this is a meaningful contribution in the area of APPAs. Note that the right-hand side of (2.14) is just $\tilde{y}^k$ (see (2.10)). Therefore, SS-method (2.14) can be written as

$$x^{k+1} = x^k - \alpha\left(x^k - \tilde{y}^k\right), \quad \alpha = 1, \tag{2.15}$$

which can be alternatively interpreted as follows: starting from x^k, SS-method moves along the direction $-(x^k - \tilde{y}^k)$ with the step size $\alpha = 1$. A natural question is whether there exists a better step size than 1 along the direction.

3. Main results

We extend the Solodov–Svaiter's formula (2.15) and present the following algorithm.

Algorithm. For given $x^0 \in C$ and $\lambda_{\min} > 0$, the sequence $\{x^k\}$ is generated by the iterative schemes:

Step 1. Find an approximate solution of (1.2), i.e., find y^k in the sense that

$$y^k \approx P_C\left[x^k - \lambda_k F\left(y^k\right)\right] \tag{3.1}$$

under the following inexactness restriction:

$$\Delta\left(y^k\right) \leqslant \nu\left(\left\|x^k - y^k\right\|^2 + \left\|x^k - \tilde{y}^k\right\|^2\right), \quad \nu < 1, \tag{3.2}$$

where

$$\tilde{y}^k = P_C\left[x^k - \lambda_k F\left(y^k\right)\right] \tag{3.3}$$

and $\Delta(\cdot)$ is defined by (2.6).

Step 2. Compute the new iterate

$$x^{k+1}(\alpha) = P_C\left[x^k - \alpha\left(x^k - \tilde{y}^k\right)\right], \tag{3.4}$$

where α is the step size. How to choose α will be specialized later.

An equivalent form of (3.4) is

$$x^{k+1}(\alpha) = P_C\left[x^k - \alpha\left(x^k - P_C\left[x^k - \lambda_k F\left(y^k\right)\right]\right)\right]. \tag{3.5}$$

Note that (3.3) is an extragradient step. In addition, by setting $y^k := x^{k+1}(\alpha)$ and $\alpha := 1$ in (3.5), it reduces to the classical PPA (1.4), thus the new method (3.1)–(3.4) is called an approximate proximal-extragradient type method.

The following theorem concerns how to choose the step size α in (3.4). Note that the technique developed in [7] and then extended for pseudo-monotone variational inequalities in [10] is useful in following analysis. For convenience, we denote

$$\zeta^k := y^k - \tilde{y}^k \tag{3.6}$$

and then we have

$$\Delta\left(y^k\right) = 2\left(\zeta^k\right)^T F\left(y^k\right) - \left\|\zeta^k\right\|^2. \tag{3.7}$$

Theorem 3.1. *Given $x^k \in C$ and $\lambda_k > 0$, let $y^k \in C$ be an approximate solution of* (1.2) *in the sense of* (3.1) *and the new iterate $x^{k+1}(\alpha)$ be given by* (3.4). *Then for any $\alpha > 0$ we have*

$$\Theta(\alpha) := \|x^k - x^*\|^2 - \|x^{k+1}(\alpha) - x^*\|^2 \geqslant \Psi(\alpha),$$

where

$$\Psi(\alpha) := 2\alpha\{\|x^k - \tilde{y}^k\|^2 - \lambda_k(\zeta^k)^T F(y^k)\} - \alpha^2\|x^k - \tilde{y}^k\|^2, \tag{3.8}$$

$\tilde{y}^k$ and ζ^k are defined by (3.3) *and* (3.6), *respectively.*

Proof. Since $x^{k+1}(\alpha) = P_C[x^k - \alpha(x^k - \tilde{y}^k)]$ and $x^* \in C$, it follows from (2.2) that

$$\|x^k - x^* - \alpha(x^k - \tilde{y}^k)\| \geqslant \|x^{k+1}(\alpha) - x^*\|.$$

Then we have

$$\begin{aligned}\Theta(\alpha) &\geqslant \|x^k - x^*\|^2 - \|x^k - x^* - \alpha(x^k - \tilde{y}^k)\|^2 \\ &= 2\alpha(x^k - x^*)^T(x^k - \tilde{y}^k) - \alpha^2\|x^k - \tilde{y}^k\|^2 \\ &= 2\alpha\|x^k - \tilde{y}^k\|^2 + 2\alpha(\tilde{y}^k - x^*)^T(x^k - \tilde{y}^k) - \alpha^2\|x^k - \tilde{y}^k\|^2.\end{aligned} \tag{3.9}$$

Comparing the right-hand side of (3.9) and $\Psi(\alpha)$ in (3.8), it remains to prove

$$(\tilde{y}^k - x^*)^T(x^k - \tilde{y}^k) \geqslant (\tilde{y}^k - y^k)^T\lambda_k F(y^k). \tag{3.10}$$

Since $\tilde{y}^k = P_C[x^k - \lambda_k F(y^k)]$ and $x^* \in C$, it follows from (2.1) that

$$\{x^k - \lambda_k F(y^k) - \tilde{y}^k\}^T(\tilde{y}^k - x^*) \geqslant 0$$

and thus

$$(\tilde{y}^k - x^*)^T(x^k - \tilde{y}^k) \geqslant (\tilde{y}^k - x^*)^T\lambda_k F(y^k). \tag{3.11}$$

Note that $x^* \in C^*$ and F is monotone. We have

$$(y^k - x^*)^T F(y^k) \geqslant (y^k - x^*)^T F(x^*) \geqslant 0,$$

which implies

$$(\tilde{y}^k - x^*)^T\lambda_k F(y^k) \geqslant (\tilde{y}^k - y^k)^T\lambda_k F(y^k). \tag{3.12}$$

Therefore, inequality (3.10) follows from (3.11) and (3.12) directly and the proof is complete. □

The following relation is useful in following analysis and thus we list it as a proposition.

Proposition 3.1. *For y^k, $\tilde{y}^k$, $\Delta(y^k)$, and ζ^k defined in Sections* 2 *and* 3, *we have*

$$\|x^k - \tilde{y}^k\|^2 - \lambda_k(\zeta^k)^T F(y^k) = \frac{1}{2}\{(\|x^k - y^k\|^2 + \|x^k - \tilde{y}^k\|^2) - \Delta(y^k)\}. \tag{3.13}$$

Proof. Using $x^k - \tilde{y}^k = (x^k - y^k) + \zeta^k$ (see (3.6)), we rewrite

$$\|x^k - \tilde{y}^k\|^2 = \frac{1}{2}\big(\|(x^k - y^k) + \zeta^k\|^2 + \|x^k - \tilde{y}^k\|^2\big).$$

By some regrouping, we obtain

$$\begin{aligned}
&\|x^k - \tilde{y}^k\|^2 - \lambda_k(\zeta^k)^T F(y^k) \\
&\quad = \frac{1}{2}\big(\|(x^k - y^k) + \zeta^k\|^2 + \|x^k - \tilde{y}^k\|^2\big) - \lambda_k(\zeta^k)^T F(y^k) \\
&\quad = \frac{1}{2}\big(\|x^k - y^k\|^2 + \|\zeta^k\|^2 + \|x^k - \tilde{y}^k\|^2\big) - (\zeta^k)^T\big\{(y^k - x^k) + \lambda_k F(y^k)\big\} \\
&\quad = \frac{1}{2}\big(\|x^k - y^k\|^2 + \|x^k - \tilde{y}^k\|^2\big) - \Big((\zeta^k)^T\big((y^k - x^k) + \lambda_k F(y^k)\big) - \frac{1}{2}\|\zeta^k\|^2\Big) \\
&\quad \overset{(3.7)}{=} \frac{1}{2}\big(\|x^k - y^k\|^2 + \|x^k - \tilde{y}^k\|^2\big) - \frac{1}{2}\Delta(y^k). \qquad \square
\end{aligned}$$

Now we begin to investigate how to choose the step size α in (3.4). Substituting (3.13) into (3.8), we have

$$\Psi(\alpha) = \alpha\big\{\big(\|x^k - y^k\|^2 + \|x^k - \tilde{y}^k\|^2\big) - \Delta(y^k)\big\} - \alpha^2\|x^k - \tilde{y}^k\|^2. \tag{3.14}$$

In fact, $\Psi(\alpha)$ is the tight lower bound of the improvement obtained at the kth iteration of the proposed method. Since $\Psi(\alpha)$ is a quadratic function of α, it reaches its maximum at

$$\alpha_k^* = \frac{(\|x^k - y^k\|^2 + \|x^k - \tilde{y}^k\|^2) - \Delta(y^k)}{2\|x^k - \tilde{y}^k\|^2} \tag{3.15}$$

with

$$\Psi(\alpha_k^*) = \frac{1}{2}\alpha_k^*\big(\big(\|x^k - y^k\|^2 + \|x^k - \tilde{y}^k\|^2\big) - \Delta(y^k)\big). \tag{3.16}$$

Under inexactness restriction (3.2), it follows from (3.15) and (3.16) that

$$\alpha_k^* \geqslant \frac{1-\nu}{2} \quad \text{and} \quad \Psi(\alpha_k^*) \geqslant \frac{(1-\nu)^2}{4}\big(\|x^k - y^k\|^2 + \|x^k - \tilde{y}^k\|^2\big). \tag{3.17}$$

To accelerate convergence, we propose a relaxation factor $\gamma_k \in [\gamma_L, \gamma_U] \subset (0, 2)$. Thus in practice the step size α in (3.4) at the kth iteration is

$$\alpha = \gamma_k \alpha_k^*. \tag{3.18}$$

Therefore, the new iterate x^{k+1} is generated by

$$x^{k+1} = P_C\big[x^k - \gamma_k\alpha_k^*(x^k - \tilde{y}^k)\big]. \tag{3.19}$$

By simple manipulations we obtain

$$\begin{aligned}
\Psi(\gamma_k\alpha_k^*) &\overset{(3.14)}{=} \gamma_k\alpha_k^*\big(\|x^k - y^k\|^2 + \|x^k - \tilde{y}^k\|^2 - \Delta(y^k)\big) - \big(\gamma_k^2\alpha_k^*\big)\big(\alpha_k^*\|x^k - \tilde{y}^k\|^2\big) \\
&\overset{(3.15)}{=} \left(\gamma_k\alpha_k^* - \frac{1}{2}\gamma_k^2\alpha_k^*\right)\big(\|x^k - y^k\|^2 + \|x^k - \tilde{y}^k\|^2 - \Delta(y^k)\big) \\
&\overset{(3.16)}{=} \gamma_k(2-\gamma_k)\Psi(\alpha_k^*).
\end{aligned} \tag{3.20}$$

It follows from Theorem 3.1 and (3.17) that

$$\|x^{k+1} - x^*\|^2 \leqslant \|x^k - x^*\|^2 - \frac{\gamma_L(2-\gamma_U)(1-\nu)^2}{4}\big(\|x^k - y^k\|^2 + \|x^k - \tilde{y}^k\|^2\big). \tag{3.21}$$

Theorem 3.2. *Given $x^k \in C$ and $\lambda_k > 0$, let $y^k \in C$ be an approximate solution of* (1.2) *in the sense of* (3.1) *and the new iterate x^{k+1} be generated by* (3.19). *If the inexactness criterion* (3.2) *holds, then $\{x^k\}$ converges to some $x^\infty \in C^*$.*

Proof. It follows from (3.21) that there is a constant $c_0 > 0$ such that

$$\|x^{k+1} - x^*\|^2 \leqslant \|x^k - x^*\|^2 - c_0\big(\|x^k - y^k\|^2 + \|x^k - \tilde{y}^k\|^2\big), \quad \forall x^* \in C^*. \tag{3.22}$$

This means that the sequence $\{x^k\}$ is bounded. In addition, we have

$$\sum_{k=1}^{\infty} c_0\big(\|x^k - y^k\|^2 + \|x^k - \tilde{y}^k\|^2\big) \leqslant \|x^0 - x^*\|^2.$$

Therefore, we have

$$\lim_{k\to\infty} \|x^k - y^k\| = 0 \quad \text{and} \quad \lim_{k\to\infty} \|x^k - \tilde{y}^k\| = 0,$$

and consequently $\{y^k\}$ is also bounded. Moreover, since $\zeta^k = (x^k - \tilde{y}^k) - (x^k - y^k)$, we have

$$\lim_{k\to\infty} \|\zeta^k\| = 0.$$

Since $\|e(y^k, \lambda)\|$ is a non-decreasing function of λ, it follows from $\lambda_k \geqslant \lambda_{\min}$ that

$$\begin{aligned}
\|e(y^k, \lambda_{\min})\| &\leqslant \|e(y^k, \lambda_k)\| = \|y^k - P_C[y^k - \lambda_k F(y^k)]\| \\
&\overset{(3.6)}{=} \|\zeta^k + \tilde{y}^k - P_C[y^k - \lambda_k F(y^k)]\| \\
&\overset{(3.3)}{\leqslant} \|\zeta^k\| + \|P_C[x^k - \lambda_k F(y^k)] - P_C[y^k - \lambda_k F(y^k)]\| \\
&\overset{(2.2)}{\leqslant} \|\zeta^k\| + \|x^k - y^k\|
\end{aligned}$$

and thus

$$\lim_{k\to\infty} e(y^k, \lambda_{\min}) = 0. \tag{3.23}$$

Let x^∞ be a cluster point of $\{y^k\}$ and the subsequence $\{y^{k_j}\}$ converges to x^∞. Since $e(x, \lambda)$ is a continuous function of x, it follows from (3.23) that

$$e(x^\infty, \lambda_{\min}) = \lim_{j\to\infty} e(y^{k_j}, \lambda_{\min}) = 0.$$

According to Lemma 2.1, x^∞ is a solution point of $\mathrm{VI}(C, F)$.

From $\lim_{k\to\infty}\|x^k - y^k\| = 0$ and $\lim_{j\to\infty} y^{k_j} = x^\infty$, we know $\lim_{j\to\infty} x^{k_j} = x^\infty$. Note that inequality (3.22) is true for all solution points of $\mathrm{VI}(C, F)$, hence we have

$$\|x^{k+1} - x^\infty\|^2 \leqslant \|x^k - x^\infty\|^2, \quad \forall k \geqslant 0; \tag{3.24}$$

and it follows that the sequence $\{x^k\}$ converges to x^∞. □

4. Relation to Solodov–Svaiter's method

The method proposed by Solodov and Svaiter in [13] is a specific implementation of the proposed method where $\alpha = 1$. In [13] the inexactness restriction for finding y^k is set as

$$\Delta(y^k) \leqslant \nu \|x^k - y^k\|^2, \quad \nu < 1, \tag{4.1}$$

and then the new iterate is generated by

$$x^{k+1} = P_C\big[x^k - \lambda_k F(y^k)\big]. \tag{4.2}$$

It is clear that condition (4.1) is more restrictive than condition (3.2). Under restriction (4.1), it follows from (3.15) that

$$\alpha_k^* = \frac{\|x^k - \tilde{y}^k\|^2 + \|x^k - v^k\|^2 - \Delta(y^k)}{2\|x^k - \tilde{y}^k\|^2} > 0.5.$$

Therefore, SS-method can be viewed as a special form of (3.19) by setting

$$\gamma_k := \frac{1}{\alpha_k^*} \in (0, 2). \tag{4.3}$$

From Eq. (3.14) we have

$$\Psi(1) = \|x^k - y^k\|^2 - \Delta(y^k)$$

and applying the inexactness restriction (4.1), we obtain

$$\Psi(1) \geqslant (1 - \nu)\|x^k - y^k\|^2. \tag{4.4}$$

It follows from Theorem 3.1 that the sequence $\{x^k\}$ generated by SS-method satisfies

$$\|x^{k+1} - x^*\|^2 \leqslant \|x^k - x^*\|^2 - (1 - \nu)\|x^k - y^k\|^2. \tag{4.5}$$

Hence, convergence of SS-method follows from (4.5) immediately.

5. Numerical example for network equilibrium problems

In this section, we compare the efficiency of the proposed method with SS-method. As an application we use the example in the traffic equilibrium problems [15].

Consider a network $[N, L]$ of nodes N and directed links L, which consists of a finite sequence of connecting links with a certain orientation. Let a, b, etc., denote the links, and let p, q, etc., denote the paths. We let ω denote an origin/destination (O/D) pair of nodes of the network and P_ω denotes the set of all paths connecting O/D pair ω. Let x_p represent the traffic flow on path p and d_ω denote the traffic demand between O/D pair ω, which must satisfy

$$d_\omega = \sum_{p \in P_\omega} x_p,$$

where $x_p \geqslant 0$, $\forall p$. Let f_a denote the link load on link a, which must satisfy the following conservation of flow equation:

$$f_a = \sum_{p \in P} \delta_{ap} x_p,$$

where $\delta_{ap} = 1$, if link a is contained in path p, and 0, otherwise. Let $t = \{t_a \mid a \in L\}$ be the row vector of link costs, with t_a denoting the user cost of traversing link a which is given by

$$t_a(f_a) = t_a^0 \left[1 + 0.15 \left(\frac{f_a}{C_a} \right)^4 \right], \tag{5.1}$$

where t_a^0 is the free-flow travel cost on link a and C_a is the designed capacity of link a. A user travelling on path p incurs a (path) travel cost θ_p satisfying

$$\theta_p = \sum_{a \in L} \delta_{ap} t_a.$$

Let A be the path-arc incidence matrix of the given problem. Since x is the path-flow, the arc-flow f is given by

$$f = A^T x.$$

For given link travel cost vector t_a, the path travel cost vector θ is given by

$$\theta = At.$$

Hence, the path travel cost vector θ is a mapping of the path-flow x and its mathematical form is

$$\theta(x) = At\bigl(A^T x\bigr).$$

Associated with every O/D pair ω, there is a travel disutility λ_ω, which is defined as

$$\lambda_\omega(d) = -m_\omega \log(d_\omega) + q_\omega. \tag{5.2}$$

Note that both the path costs and the travel disutilities are functions of the flow pattern x. The traffic network equilibrium problem is to seek the path flow pattern x^*, which induces a demand pattern $d^* = d(x^*)$, for every O/D pair ω and each path $p \in P_\omega$,

$$F_p(x) = \theta_p(x) - \lambda_\omega\bigl(d(x)\bigr).$$

The problem is a monotone variational inequality in the space of path-flow pattern x:

$$\text{Find} \quad x^* \geqslant 0 \text{ such that } \bigl(x - x^*\bigr)^T F(x^*) \geqslant 0, \quad \forall x \geqslant 0. \tag{5.3}$$

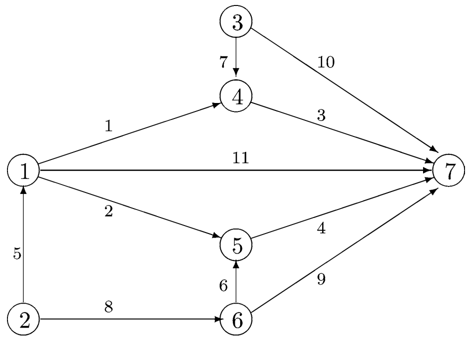

Fig. 1. The network used for the numerical test.

Table 1
The free-flow travel cost and the designed capacity of links in (5.1)

Link	Free-flow travel time t_a^0	Capacity C_a	Link	Free-flow travel time t_a^0	Capacity C_a
1	6	200	7	5	150
2	5	200	8	10	150
3	6	200	9	11	200
4	16	200	10	11	200
5	6	100	11	15	200
6	1	100	–	–	–

Table 2
The O/D pairs and the coefficient m and q in (5.2)

No. of the pair	O/D pair	m_ω	q_ω
1	$(1, 7)$	25	$25 \log 600$
2	$(2, 7)$	33	$33 \log 500$
3	$(3, 7)$	20	$20 \log 500$
4	$(6, 7)$	20	$20 \log 400$

For the test we take the example used in [15]. The network is depicted in Fig. 1. The free-flow travel cost and the designed capacity of links in (5.1) are given in Table 1, the O/D pairs and the coefficient m and q in the disutility function (5.2) are given in Table 2. For this example, there are together 12 paths for the 4 given O/D pairs as listed in Table 4.

We test the problem with SS-method and the proposed method. Since the mapping $F(x)$ contains logarithmic function, we begin with starting point $x^0 = (1, 1, \dots, 1)$. We take $\lambda_k \equiv 50$ and use the improved extra-gradient method [7] to solve the sub-problem (1.2) approximately. Compared with SS-method, the inexactness restriction of the proposed method is further relaxed (see (4.1) and (3.2)). However, for comparison, we use the same inexactness restriction (2.13) with $\nu = 0.9$. The relaxation factor γ_k in (3.19) should lie in $[1, 2)$. In the test, we take $\gamma_k \equiv 1.5$. All codes are written in Matlab and run on an IBM T40 notebook computer. The computation stops as soon as $\|e(x^k, 1)\|_\infty \leqslant 10^{-6}$ (see (2.4)). We report the iteration numbers and the CPU time in Table 3. Since the subproblems are solved approximately by the iterative method in [7], the number of total inner

Table 3
Iteration numbers and CPU time of different methods

Method	No. of outer iterations	No. of total inner iterations	CPU-time (sec)
SS-method	20	182	0.22
New-method	12	111	0.13

Table 4
The optimal path flow

O/D pairs	Path no.	Link on the path	Optimal path-flow
	1	(1, 3)	165.3145
O/D pair (1, 7)	2	(2, 4)	0
	3	(11)	138.5735
	4	(5, 1, 3)	82.5281
	5	(5, 2, 4)	0
O/D pair (2, 7)	6	(5, 11)	55.7871
	7	(8, 6, 4)	0
	8	(8, 9)	87.0260
O/D pair (3, 7)	9	(7, 3)	19.7549
	10	(10)	229.9747
O/D pair (6, 7)	11	(9)	178.5600
	12	(6, 4)	0

Table 5
The optimal link flow

Link no.	Link flow	Link no.	Link flow	Link no.	Link flow	Link no.	Link flow
1	247.8426	4	0	7	19.7549	10	229.9747
2	0	5	138.3152	8	87.0260	11	194.3606
3	267.5974	6	0	9	265.5860	–	–

iterations is also reported. The optimal path flow and link flow are given in Tables 4 and 5, respectively.

By choosing the optimal step size α_k^* (see (3.15)), we extended SS-method in [13] to a new version. The computational load for choosing α_k^* is quite tiny. Comparison with SS-method shows that the new method is attractive in practice.

Acknowledgment

The authors thank the anonymous referee for his careful reading of earlier versions of this paper and constructive suggestions.

References

[1] A. Auslender, M. Teboulle, Entropic proximal decomposition methods for convex programs and variational inequalities, Math. Programming 91 (2001) 33–47.
[2] D.P. Bertsekas, J.N. Tsitsiklis, Parallel and Distributed Computation, Numerical Methods, Prentice Hall, Englewood Cliffs, NJ, 1989.
[3] X.J. Chen, M. Fukushima, Proximal quasi-Newton methods for nondifferentiable convex optimization, Math. Programming 85 (1998) 313–334.
[4] J. Eckstein, Nonlinear proximal point algorithms using Bregman functions, with applications to convex programming, Math. Oper. Res. 18 (1993) 202–226.
[5] J. Eckstein, Approximate iterations in Bregman-function-based proximal algorithms, Math. Programming 83 (1998) 113–123.
[6] M.C. Ferris, J.S. Pang, Engineering and economic applications of complementarity problems, SIAM Rev. 39 (1997) 669–713.
[7] B.S. He, L.-Z. Liao, Improvements of some projection methods for monotone nonlinear variational inequalities, J. Optim. Theory Appl. 112 (2002) 111–128.
[8] M.A. Noor, Some recent advances in variational inequalities, Part I. Basic concepts, New Zealand J. Math. 26 (1997) 53–80.
[9] M.A. Noor, Some recent advances in variational inequalities, Part II. Other concepts, New Zealand J. Math. 26 (1997) 229–255.
[10] M.A. Noor, Extragradient method for pseudomonotone variational inequalities, J. Optim. Theory Appl. 117 (2003) 475–488.
[11] R.T. Rockafellar, Augmented Lagrangians and applications of the proximal point algorithm in convex programming, Math. Oper. Res. 1 (1976) 97–116.
[12] R.T. Rockafellar, Monotone operators and the proximal point algorithm, SIAM J. Control Optim. 14 (1976) 877–898.
[13] M.V. Solodov, B.F. Svaiter, Error bounds for proximal point subproblems and associated inexact proximal point algorithms, Math. Programming, Ser. B 88 (2000) 371–389.
[14] K. Taji, M. Fukushima, T. Ibaraki, A globally convergent Newton method for solving strongly monotone variational inequalities, Math. Programming 58 (1993) 369–383.
[15] H. Yang, M.G.H. Bell, Traffic restraint, road pricing and network equilibrium, Transportation Res. B 31 (1997) 303–314.