Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 10, No. 1, pp. 199–213, January 2014

LAGRANGIAN-PPA BASED CONTRACTION METHODS FOR LINEARLY CONSTRAINED CONVEX OPTIMIZATION

Yanfei You*, Xiaoling Fu† and Bingsheng He‡

Abstract: For solving linearly constrained convex optimization problem $\min\{\theta(x) \mid Ax = b\ (or\ Ax \geq b),\ x \in \mathcal{X}\}$, this paper presents a class of contraction methods under the assumption that the minimization problem $\min\{\theta(x) + \frac{r}{2}\|x - a\|^2 \mid x \in \mathcal{X}\}$ has closed-form solution. The global convergence and the convergence rate of the proposed algorithms in an ergodic sense are proved. Finally, numerical results are provided to verify that the new methods are effective for some practical problems.

Key words: *convex optimization, proximal point algorithm, contraction methods, convergence rate*

Mathematics Subject Classification: *90C25, 90C30*

1 Introduction

Let $A \in \Re^{m\times n}$, $b \in \Re^m$, $\mathcal{X} \subset \Re^n$ be a closed convex set and $\theta(x) : \Re^n \to \Re$ be a proper convex function. We consider the following linearly constrained convex optimization

$$\min\{\theta(x) \mid Ax = b\,(or\ Ax \geq b),\ x \in \mathcal{X}\}. \tag{1.1}$$

This problem is popular in numerical optimization and has a wide rang of applications in different fields. In fact, plenty of convex optimization models for practical applications can be expressed as the problem (1.1). For example, the nearest approximation of correlation matrix and low-rank matrix completion [2] have the mathematical form of (1.1). In some practical applications, for any $r > 0$, the minimization problem

$$\min\{\theta(x) + \frac{r}{2}\|x - a\|^2 \mid x \in \mathcal{X}\} \tag{1.2}$$

either admits closed-form solution or can be efficiently solved up to a high precision.

Note that the Lagrangian function of the problem (1.1) is

$$L(x, \lambda) = \theta(x) - \lambda^T(Ax - b), \tag{1.3}$$

which is defined on $\mathcal{X} \times \Lambda$, where

$$\Lambda = \begin{cases} \Re^m, & \text{for the equality constraints} \quad Ax = b, \\ \Re^m_+, & \text{for the inequality constraints} \quad Ax \geq b. \end{cases}$$

*This author was supported by the NSFC grant 91130007
†This author was supported by the NSFC grant 70901018.
‡This author was supported by the NSFC Grant 91130007 and the MOEC fund 20110091110004.

ISSN 1348-9151

Let (x^*, λ^*) be a saddle point of the Lagrangian function, then we have

$$L_{\lambda\in\Lambda}(x^*, \lambda) \le L(x^*, \lambda^*) \le L_{x\in\mathcal{X}}(x, \lambda^*).$$

Finding a saddle point of the Lagrangian function is equivalent to getting a pair of (x^*, λ^*) such that

$$\begin{cases} x^* \in \mathcal{X}, & \theta(x) - \theta(x^*) + (x - x^*)^T(-A^T\lambda^*) \ge 0, \quad \forall\, x \in \mathcal{X}, \\ \lambda^* \in \Lambda, & (\lambda - \lambda^*)^T(Ax^* - b) \ge 0, \quad \forall\, \lambda \in \Lambda. \end{cases} \tag{1.4}$$

By denoting

$$u = \begin{pmatrix} x \\ \lambda \end{pmatrix}, \quad F(u) = \begin{pmatrix} -A^T\lambda \\ Ax - b \end{pmatrix} \tag{1.5}$$

and

$$\Omega = \mathcal{X} \times \Lambda, \tag{1.6}$$

the compact form of (1.4) can be written as the following mixed variational inequality:

$$\text{MVI}(\Omega, F, \theta) \qquad u^* \in \Omega, \quad \theta(x) - \theta(x^*) + (u - u^*)^T F(u^*) \ge 0, \quad \forall\, u \in \Omega. \tag{1.7}$$

Note that $\theta(x)$ is a convex function and the operator F in (1.5) is affine and monotone since

$$(u - v)^T(F(u) - F(v)) = 0, \quad \forall\, u,\, v \in \Re^{n+m}.$$

We denote the solution set of the problem (1.7) (and thus the set of the saddle points of the Lagrangian function) by Ω^*. For any $u^* = (x^*, \lambda^*) \in \Omega^*$, x^* is an optimal solution of (1.1). In this paper, we treat the linearly constrained convex optimization (1.1) in the frame of MVI(Ω, F, θ).

The rest of this paper is organized as follows. In Section 2, we review the basic idea of PPA and some related algorithms, and then show the motivation. Section 3 introduces the predictor via Lagrangian-PPA. In Section 4, we construct the Lagrangian-PPA based contraction methods and prove their convergence. Section 5 proves the convergence rate of the proposed method (4.14) in an ergodic sense. Finally, in Section 6, we report the numerical results for some practical applications.

2 Preliminaries of PPA and the Motivation

Proximal point algorithm (short as PPA), originally proposed by Martinet [15] and Rockafellar [16], is an attractive approach for MVI(Ω, F, θ) (1.7). As our proposed methods are based on PPA, we briefly review the idea of PPA. To solve MVI(Ω, F, θ), the k-th iteration of PPA begins with a given $u^k \in \Omega$ and $r_k > 0$, and obtains a $u^{k+1} \in \Omega$, such that

$$\theta(x) - \theta(x^{k+1}) + (u - u^{k+1})^T\big\{F(u^{k+1}) + r_k(u^{k+1} - u^k)\big\} \ge 0, \quad \forall u \in \Omega. \tag{2.1}$$

The generated sequence $\{u^k\}$ satisfies

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

and is Fejér monotone with respect to the solution set Ω^*.

Solving the PPA subproblem (2.1) is almost as difficult as the MVI (1.7), therefore the classical PPA only has abstract meaning and is not often used in practical computation. To overcome the drawback of the classical PPA, a customized PPA was proposed in [12] and further analyzed in [10]. The k-th iteration of the customized PPA begins with a given $u^k = (x^k, \lambda^k)$, and generates a $\tilde{u}^k$ via the following scheme:

$$\tilde{u}^k \in \Omega, \quad \theta(x) - \theta(\tilde{x}^k) + (u - \tilde{u}^k)^T \{F(\tilde{u}^k) + G(\tilde{u}^k - u^k)\} \geq 0, \quad \forall u \in \Omega, \tag{2.2}$$

where

$$G = \begin{pmatrix} rI_n & -A^T \\ -A & sI_m \end{pmatrix} \tag{2.3}$$

is a positive definite matrix. $\tilde{u}^k$ is a solution of (1.7) if and only if $u^k = \tilde{u}^k$. Using the notations of F (see (1.5)) and G, the subproblem (2.2) is to find $(\tilde{x}^k, \tilde{\lambda}^k) \in \Omega$, such that

$$\theta(x) - \theta(\tilde{x}^k) + \begin{pmatrix} x - \tilde{x}^k \\ \lambda - \tilde{\lambda}^k \end{pmatrix}^T \left\{ \begin{pmatrix} -A^T\tilde{\lambda}^k \\ A\tilde{x}^k - b \end{pmatrix} + \begin{pmatrix} r(\tilde{x}^k - x^k) - A^T(\tilde{\lambda}^k - \lambda^k) \\ -A(\tilde{x}^k - x^k) + s(\tilde{\lambda}^k - \lambda^k) \end{pmatrix} \right\} \geq 0,$$
$$\forall (x, \lambda) \in \Omega.$$

The above inequality is decomposed as

$$\tilde{x}^k \in \mathcal{X}, \quad \theta(x) - \theta(\tilde{x}^k) + (x - \tilde{x}^k)^T \{-A^T(2\tilde{\lambda}^k - \lambda^k) + r(\tilde{x}^k - x^k)\} \geq 0, \quad \forall\, x \in \mathcal{X}, \tag{2.4a}$$

and

$$\tilde{\lambda}^k \in \Lambda, \quad (\lambda - \tilde{\lambda}^k)^T \{(Ax^k - b) + s(\tilde{\lambda}^k - \lambda^k)\} \geq 0, \quad \forall\, \lambda \in \Lambda. \tag{2.4b}$$

The solution of (2.4b) can be obtained directly by

$$\tilde{\lambda}^k = P_\Lambda[\lambda^k - \frac{1}{s}(Ax^k - b)],$$

where

$$P_\Lambda(\lambda) = \begin{cases} \lambda, & \text{if } \Lambda = \Re^m, \\ \max(0, \lambda), & \text{if } \Lambda = \Re^m_+. \end{cases}$$

After getting $\tilde{\lambda}^k$, we will obtain $\tilde{x}^k$ in (2.4a) via solving the following problem:

$$\min \big\{\theta(x) + \frac{r}{2}\big\|x - \big[x^k + \frac{1}{r}\, A^T(2\tilde{\lambda}^k - \lambda^k)\big]\big\|^2 \,\big|\, x \in \mathcal{X}\big\}. \tag{2.5}$$

Since $\tilde{\lambda}^k$ in (2.5) is known, this subproblem has the form of (1.2). In other words, under the assumption that the problem (1.2) has closed-form solution, the PPA subproblem (2.2) can be solved efficiently.

Setting $u = u^*$ in (2.2) gives

$$(\tilde{u}^k - u^*)^T G(u^k - \tilde{u}^k) \geq \theta(\tilde{x}^k) - \theta(x^*) + (\tilde{u}^k - u^*)^T F(\tilde{u}^k).$$

Using the structure of F yields $(\tilde{u}^k - u^*)^T F(\tilde{u}^k) = (\tilde{u}^k - u^*)^T F(u^*)$, and thus from the last inequality we get

$$(\tilde{u}^k - u^*)^T G(u^k - \tilde{u}^k) \geq \theta(\tilde{x}^k) - \theta(x^*) + (\tilde{u}^k - u^*)^T F(u^*). \tag{2.6}$$

Since $u^* \in \Omega^*$ and $\tilde{u}^k \in \Omega$, it follows from (1.7) that the right hand side of (2.6) is nonnegative and thus

$$(\tilde{u}^k - u^*)^T G(u^k - \tilde{u}^k) \geq 0. \tag{2.7}$$

Using $\gamma \in (0, 2)$ and

$$u^{k+1} = u^k - \gamma(u^k - \tilde{u}^k),$$

to obtain the new iterate u^{k+1}, it follows from (2.7) that the generated sequence $\{u^k\}$ satisfies

$$\|u^{k+1} - u^*\|_G^2 \leq \|u^k - u^*\|_G^2 - \gamma(2-\gamma)\|u^k - \tilde{u}^k\|_G^2, \quad \forall\, u^* \in \Omega^*.$$

Note that (2.2) can be regarded as the PPA in the context of G-norm (defined by $\|u\|_G = \sqrt{u^T G u}$). In order to ensure the convergence, the matrix G in (2.2) should be positive definite and thus the parameters r and s have to satisfy $rs > \|A^T A\|$. In practice, large parameters in the matrix G may lead to slow convergence. In this paper, we relax the restriction on the parameters r and s to

$$sr \geq \frac{1}{2}\|A^T A\|, \tag{2.8}$$

and propose a simple algorithm for solving (1.1).

3 Predictor via Lagrangian-PPA

For given $u^k = (x^k, \lambda^k)$, we obtain $\tilde{u}^k = (\tilde{x}^k, \tilde{\lambda}^k)$ by minimizing (resp. maximizing) the Lagrangian function with proximal terms. In details, we use the following primal-dual order:

$$\tilde{x}^k = \operatorname{Arg\,min}\{L(x, \lambda^k) + \frac{r}{2}\|x - x^k\|^2 \,|\, x \in \mathcal{X}\}, \tag{3.1a}$$

and

$$\tilde{\lambda}^k = \operatorname{Arg\,max}\{L(\tilde{x}^k, \lambda) - \frac{s}{2}\|\lambda - \lambda^k\|^2 \,|\, \lambda \in \Lambda\}. \tag{3.1b}$$

If one directly takes $\tilde{u}^k$ as the new iterate, it corresponds the classical Arrow-Hurwicz method [1]. Zhu and Chan used this classical Arrow-Hurwicz method to solve the Rudin Osher and Fatemi (ROF) image denoising problem [17]. See also [5] for a proof of convergence of the Arrow-Hurwicz method with very large rs. In the proposed method of this paper, we restrict rs to satisfy (2.8). We do not accept $\tilde{u}^k$ as an new iterate and call it as a predictor in the k-th iteration. Using the notation of the Lagrangian function, we conclude that the procedure to get $\tilde{u}^k = (\tilde{x}^k, \tilde{\lambda}^k)$ has the following equivalent form:

For given (x^k, λ^k) and $r, s > 0$,

$$\tilde{x}^k = \operatorname{Arg\,min}\big\{\theta(x) + \frac{r}{2}\big\|x - \big[x^k + \frac{1}{r}A^T\lambda^k\big]\big\|^2 \,\big|\, x \in \mathcal{X}\big\}, \tag{3.2a}$$

and

$$\tilde{\lambda}^k = P_\Lambda[\lambda^k - \frac{1}{s}(A\tilde{x}^k - b)]. \tag{3.2b}$$

Lemma 3.1. *For given $u^k = (x^k, \lambda^k)$, let $\tilde{u}^k$ be generated by* (3.2)*. Then we have*

$$\tilde{u}^k \in \Omega, \quad \big(\theta(x) - \theta(\tilde{x}^k)\big) + (u - \tilde{u}^k)^T F(\tilde{u}^k) \geq (u - \tilde{u}^k)^T Q(u^k - \tilde{u}^k), \quad \forall\, u \in \Omega, \tag{3.3}$$

where

$$Q = \begin{pmatrix} rI_n & A^T \\ 0 & sI_m \end{pmatrix}. \tag{3.4}$$

Proof. According to the first order optimal condition, $\tilde{x}^k$ generated by (3.2a) satisfies

$$\tilde{x}^k \in \mathcal{X}, \;\; \theta(x) - \theta(\tilde{x}^k) + (x - \tilde{x}^k)^T\{r(\tilde{x}^k - x^k) - A^T\lambda^k\} \ge 0, \;\; \forall x \in \mathcal{X}. \tag{3.5}$$

The updating form (3.2b) can be rewritten as $\tilde{\lambda}^k = P_\Lambda\{\tilde{\lambda}^k - [(\tilde{\lambda}^k - \lambda^k) + \frac{1}{s}(A\tilde{x}^k - b)]\}$, thus we have

$$\tilde{\lambda}^k \in \Lambda, \;\; (\lambda - \tilde{\lambda}^k)^T\{(A\tilde{x}^k - b) + s(\tilde{\lambda}^k - \lambda^k)\} \ge 0, \;\; \forall \lambda \in \Lambda. \tag{3.6}$$

Combining (3.5) and (3.6) together, we get

$$\begin{aligned}(\tilde{x}^k, \tilde{\lambda}^k) \in \Omega, \quad \theta(x) - \theta(\tilde{x}^k) + \begin{pmatrix} x - \tilde{x}^k \\ \lambda - \tilde{\lambda}^k \end{pmatrix}^T \Bigg\{ \begin{pmatrix} -A^T\tilde{\lambda}^k \\ A\tilde{x}^k - b \end{pmatrix} \\ + \begin{pmatrix} r(\tilde{x}^k - x^k) \\ 0 \end{pmatrix} + \begin{pmatrix} A^T(\tilde{\lambda}^k - \lambda^k) \\ s(\tilde{\lambda}^k - \lambda^k) \end{pmatrix} \Bigg\} \ge 0, \quad \forall (x, \lambda) \in \Omega.\end{aligned}$$

The compact form is

$$\tilde{u}^k \in \Omega, \quad \theta(x) - \theta(\tilde{x}^k) + (u - \tilde{u}^k)^T\{F(\tilde{u}^k) + Q(\tilde{u}^k - u^k)\} \ge 0, \quad \forall u \in \Omega,$$

where F and Q are defined in (1.5) and (3.4), respectively. The lemma is proved. □

Remark 3.2. We can change the order of obtaining the predictor $\tilde{u}^k$ and get some similar results. Under the same condition (2.8) for given $u^k = (x^k, \lambda^k)$, we obtain $\tilde{u}^k = (\tilde{x}^k, \tilde{\lambda}^k)$ in the following dual-primal order:

$$\tilde{\lambda}^k = \operatorname{Arg}\max\{L(x^k, \lambda) - \frac{s}{2}\|\lambda - \lambda^k\|^2 \,|\, \lambda \in \Lambda\}, \tag{3.7a}$$

and

$$\tilde{x}^k = \operatorname{Arg}\min\{L(x, \tilde{\lambda}^k) + \frac{r}{2}\|x - x^k\|^2 \,|\, x \in \mathcal{X}\}. \tag{3.7b}$$

By using the similar analysis as in Lemma 3.1, the predictor $\tilde{u}^k = (\tilde{x}^k, \tilde{\lambda}^k)$ generated by (3.7) satisfies

$$\begin{aligned}(\tilde{x}^k, \tilde{\lambda}^k) \in \Omega, \quad \theta(x) - \theta(\tilde{x}^k) + \begin{pmatrix} x - \tilde{x}^k \\ \lambda - \tilde{\lambda}^k \end{pmatrix}^T \Bigg\{ \begin{pmatrix} -A^T\tilde{\lambda}^k \\ A\tilde{x}^k - b \end{pmatrix} \\ + \begin{pmatrix} r(\tilde{x}^k - x^k) \\ -A(\tilde{x}^k - x^k) \end{pmatrix} + \begin{pmatrix} 0 \\ s(\tilde{\lambda}^k - \lambda^k) \end{pmatrix} \Bigg\} \ge 0, \quad \forall (x, \lambda) \in \Omega.\end{aligned}$$

Its compact form is

$$\tilde{u}^k \in \Omega, \; \theta(x) - \theta(\tilde{x}^k) + (u - \tilde{u}^k)^T\{F(\tilde{u}^k) + Q(\tilde{u}^k - u^k)\} \ge 0, \; \forall u \in \Omega, \tag{3.8}$$

where

$$Q = \begin{pmatrix} rI_n & 0 \\ -A & sI_m \end{pmatrix}. \tag{3.9}$$

If Q is a symmetric positive definite matrix, (3.3) coincides with (2.2) and thus it is a proximal point algorithm. Lemma 3.1 is the the base for constructing the contraction method and the convergence rate proofs in Section 4 and Section 5, respectively.

4 Lagrangian-PPA Based Contraction Method

In this section, we construct a contraction method based on the predictor via Lagrangian-PPA. All the contraction methods [7, 8, 9, 11] generate a sequence $\{u^k\}$ which is Fejér monotone with respect to the solution set Ω^*.

Lemma 4.1. *For given $u^k = (x^k, \lambda^k)$, let $\tilde{u}^k$ be generated by* (3.2)*. Then we have*

$$(u^k - u^*)^T Q(u^k - \tilde{u}^k) \geq (u^k - \tilde{u}^k)^T Q(u^k - \tilde{u}^k), \quad \forall\, u^* \in \Omega^*. \tag{4.1}$$

Proof. Lemma 3.1 is the base for the proof. Substituting $u = u^*$ in (3.3) implies

$$(\tilde{u}^k - u^*)^T Q(u^k - \tilde{u}^k) \geq \theta(\tilde{x}^k) - \theta(x^*) + (\tilde{u}^k - u^*)^T F(\tilde{u}^k).$$

By using the structure of F, we have $(\tilde{u}^k - u^*)^T F(\tilde{u}^k) = (\tilde{u}^k - u^*)^T F(u^*)$, and thus

$$(\tilde{u}^k - u^*)^T Q(u^k - \tilde{u}^k) \geq \theta(\tilde{x}^k) - \theta(x^*) + (\tilde{u}^k - u^*)^T F(u^*). \tag{4.2}$$

Since $\tilde{u}^k \in \Omega$ and u^* is a solution of $\mathrm{MVI}(\Omega, F)$, according to the definition of the mixed variational inequality (see (1.7)), the right hand side of (4.2) is nonnegative. Consequently, from (4.2) we get

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

and the assertion of this lemma follows directly. □

Now, we observe the right hand side of (4.1). Because

$$\begin{aligned}
&(u^k - \tilde{u}^k)^T Q(u^k - \tilde{u}^k)\\
&= r\|x^k - \tilde{x}^k\|^2 + s\|\lambda^k - \tilde{\lambda}^k\|^2 + (x^k - \tilde{x}^k)^T A^T(\lambda^k - \tilde{\lambda}^k)\\
&\geq r\|x^k - \tilde{x}^k\|^2 - \frac{1}{2}\|x^k - \tilde{x}^k\| \cdot \|A^T(\lambda^k - \tilde{\lambda}^k)\|\\
&\quad + s\|\lambda^k - \tilde{\lambda}^k\|^2 - \frac{1}{2}\|\lambda^k - \tilde{\lambda}^k\| \cdot \|A(x^k - \tilde{x}^k)\|\\
&\geq r\|x^k - \tilde{x}^k\|^2 - \frac{1}{4}\big\{r\|x^k - \tilde{x}^k\|^2 + \frac{1}{r}\|A^T(\lambda^k - \tilde{\lambda}^k)\|^2\big\}\\
&\quad + s\|\lambda^k - \tilde{\lambda}^k\|^2 - \frac{1}{4}\big\{s\|\lambda^k - \tilde{\lambda}^k\|^2 + \frac{1}{s}\|A(x^k - \tilde{x}^k)\|^2\big\}\\
&= \frac{3}{4}\big\{r\|x^k - \tilde{x}^k\|^2 + s\|\lambda^k - \tilde{\lambda}^k\|^2\big\} - \frac{1}{4}\big\{\frac{1}{s}\|A(x^k - \tilde{x}^k)\|^2 + \frac{1}{r}\|A^T(\lambda^k - \tilde{\lambda}^k)\|^2\big\},
\end{aligned}$$

using the condition (2.8), we obtain

$$(u^k - \tilde{u}^k)^T Q(u^k - \tilde{u}^k) \geq \frac{1}{4}\big\{r\|x^k - \tilde{x}^k\|^2 + s\|\lambda^k - \tilde{\lambda}^k\|^2\big\}. \tag{4.3}$$

Thus, the right hand side of (4.1) is positive whenever $\tilde{u}^k \neq u^k$. Let us define the matrices H and M by

$$H = \begin{pmatrix} rI_n & 0 \\ 0 & sI_m \end{pmatrix}, \qquad M = \begin{pmatrix} I_n & \frac{1}{r}A^T \\ 0 & I_m \end{pmatrix}, \tag{4.4}$$

respectively. Consequently, we have

$$Q = HM. \tag{4.5}$$

For given u^k and the predictor $\tilde{u}^k$ generated by (3.2), we let $u^{k+1}(\alpha)$ be an α-dependent new iterate defined by

$$u^{k+1}(\alpha) = u^k - \alpha M(u^k - \tilde{u}^k), \tag{4.6}$$

and consider how to choose the step size α.

Lemma 4.2. *For given* $u^k = (x^k, \lambda^k)$*, let* $\tilde{u}^k$ *be generated by* (3.2) *and* $u^{k+1}(\alpha)$ *be defined by* (4.6)*. Then for any* $\alpha \geq 0$*, we have*

$$\|u^k - u^*\|_H^2 - \|u^{k+1}(\alpha) - u^*\|_H^2 \geq q(\alpha), \quad \forall\, u^* \in \Omega^*, \tag{4.7}$$

where

$$q(\alpha) = 2\alpha(u^k - \tilde{u}^k)^T Q(u^k - \tilde{u}^k) - \alpha^2 \|M(u^k - \tilde{u}^k)\|_H^2. \tag{4.8}$$

Proof. For any but fixed $u^* \in \Omega^*$, we define

$$\vartheta(\alpha) = \|u^k - u^*\|_H^2 - \|u^{k+1}(\alpha) - u^*\|_H^2. \tag{4.9}$$

By using (4.6), we get

$$\begin{aligned} \vartheta(\alpha) &= \|u^k - u^*\|_H^2 - \|(u^k - u^*) - \alpha M(u^k - \tilde{u}^k)\|_H^2 \\ &= 2\alpha(u^k - u^*)^T HM(u^k - \tilde{u}^k) - \alpha^2 \|M(u^k - \tilde{u}^k)\|_H^2. \end{aligned}$$

To the term $(u^k - u^*)^T HM(u^k - \tilde{u}^k)$ in the right hand side of the last equation, using (4.1) and $Q = HM$ (see (4.5)) gives

$$\vartheta(\alpha) \geq q(\alpha),$$

immediately and the lemma is proved. □

Note that $q(\alpha)$ in (4.8) is a quadratic function of α and it reaches its maximum at

$$\alpha_k^* = \frac{(u^k - \tilde{u}^k)^T Q(u^k - \tilde{u}^k)}{\|M(u^k - \tilde{u}^k)\|_H^2}. \tag{4.10}$$

Lemma 4.3. *Under the condition* (2.8)*, we have*

$$\alpha_k^* > 1/4, \quad \forall\, k \geq 0. \tag{4.11}$$

Proof. By using (4.10), we prove the equivalent assertion of (4.11):

$$4(u^k - \tilde{u}^k)^T Q(u^k - \tilde{u}^k) - \|M(u^k - \tilde{u}^k)\|_H^2 > 0. \tag{4.12}$$

Note that

$$4(u^k - \tilde{u}^k)^T Q(u^k - \tilde{u}^k) - \|M(u^k - \tilde{u}^k)\|_H^2 = (u^k - \tilde{u}^k)^T (2Q + 2Q^T - M^T HM)(u^k - \tilde{u}^k).$$

Using $Q = HM$ (see (4.5)), we have

$$2(Q^T + Q) - M^T HM = 2(Q^T + Q) - M^T Q.$$

In the following we prove that the above matrix is positive definite. By a manipulation, we get

$$\begin{aligned} 2(Q^T + Q) - M^T Q &= \begin{pmatrix} 4rI_n & 2A^T \\ 2A & 4sI_m \end{pmatrix} - \begin{pmatrix} I_n & 0 \\ \frac{1}{r}A & I_m \end{pmatrix} \begin{pmatrix} rI_n & A^T \\ 0 & sI_m \end{pmatrix} \\ &= \begin{pmatrix} 4rI_n & 2A^T \\ 2A & 4sI_m \end{pmatrix} - \begin{pmatrix} rI_n & A^T \\ A & sI_m + \frac{1}{r}AA^T \end{pmatrix} \\ &= \begin{pmatrix} 2rI_n & A^T \\ A & sI_m \end{pmatrix} + \begin{pmatrix} rI_n & 0 \\ 0 & 2sI_m - \frac{1}{r}AA^T \end{pmatrix}. \end{aligned} \tag{4.13}$$

According to the condition (2.8), $rs > \frac{1}{2}\|A^T A\|$, both matrices in the right hand side of (4.13) are positive definite. Thus, (4.12) is true and the assertion is proved. □

In the practical computation, we use

$$u^{k+1} = u^k - \alpha_k M(u^k - \tilde{u}^k), \tag{4.14a}$$

to update the new iterate, where

$$\alpha_k = \gamma\alpha_k^*, \qquad \gamma \in [1,2). \tag{4.14b}$$

Theorem 4.4. *For given $u^k = (x^k, \lambda^k)$, let $\tilde{u}^k$ be generated by (3.2) and u^{k+1} be updated by (4.14). Then we have*

$$\|u^{k+1} - u^*\|_H^2 \le \|u^k - u^*\|_H^2 - \frac{\gamma(2-\gamma)}{16}\|u^k - \tilde{u}^k\|_H^2. \tag{4.15}$$

Proof. Since we use (4.14) to update u^{k+1}, it follows from (4.7) that

$$\|u^k - u^*\|_H^2 - \|u^{k+1} - u^*\|_H^2 \ge q(\gamma\alpha_k^*).$$

By using (4.8) (4.10) and $\alpha^* \ge 1/4$, we obtain

$$q(\gamma\alpha_k^*) \ge \frac{1}{4}\gamma(2-\gamma)(u^k - \tilde{u}^k)^T Q(u^k - \tilde{u}^k). \tag{4.16}$$

In addition, it follows from (4.3) that

$$(u^k - \tilde{u}^k)^T Q(u^k - \tilde{u}^k) \ge \frac{1}{4}\|u^k - \tilde{u}^k\|_H^2.$$

Substituting it in (4.16), we obtain

$$q(\gamma\alpha_k^*) \ge \frac{1}{16}\gamma(2-\gamma)\|u^k - \tilde{u}^k\|_H^2,$$

and the theorem is proved. □

The inequality (4.14) indicates that the proposed algorithm is a contraction method. Because the predictor $\tilde{u}^k = (\tilde{x}^k, \tilde{\lambda}^k)$ is generated in the primal-dual order, we call the method (4.14) Primal-Dual Lagrangian-PPA based contraction method.

Theorem 4.5. *The sequence $\{u^k\}$ generated by the Primal-Dual Lagrangian-PPA based contraction method converges to some u^∞ which belongs to Ω^*.*

Proof. According to (4.15), it holds that $\{u^k\}$ is bounded and

$$\lim_{k\to\infty} \|u^k - \tilde{u}^k\| = 0. \tag{4.17}$$

So, $\{\tilde{u}^k\}$ is also bounded. Let u^∞ be a cluster point of $\{\tilde{u}^k\}$ and $\{\tilde{u}^{k_j}\}$ be a subsequence which converges to u^∞. It follows from (3.3) that

$$\tilde{u}^{k_j} \in \Omega, \quad \theta(x) - \theta(\tilde{x}^{k_j}) + (u - \tilde{u}^{k_j})^T F(\tilde{u}^{k_j}) \ge (u - u^{k_j})^T Q(u^{k_j} - \tilde{u}^{k_j}), \quad \forall\, u \in \Omega.$$

Since the matrix Q is nonsingular, it follows from the continuity of $\theta(x)$ and $F(u)$ that

$$u^\infty \in \Omega, \quad \theta(x) - \theta(x^\infty) + (u - u^\infty)^T F(u^\infty) \ge 0, \quad \forall\, u \in \Omega.$$

The above variational inequality indicates that u^∞ is a solution of MVI(Ω, F). By using (4.17) and $\lim_{j\to\infty} u^{k_j} = u^\infty$, the subsequence $\{u^{k_j}\}$ converges to u^∞. Due to (4.15), we have

$$\|u^{k+1} - u^\infty\|_H \le \|u^k - u^\infty\|_H,$$

and thus $\{u^k\}$ converges to u^∞. The proof is complete. □

Remark 4.6. For the predictor obtained by the dual-primal order (3.7), we can use the same update form (see (4.14))

$$u^{k+1} = u^k - \gamma\alpha_k^* M(u^k - \tilde{u}^k), \tag{4.18}$$

to update the new iterate where α_k^* is given by (4.10). Note that in this case, the matrix Q is defined in (3.9) and the matrix $M = H^{-1}Q$ (see (4.4)) is

$$M = \begin{pmatrix} I_n & 0 \\ -\frac{1}{s}A & I_m \end{pmatrix}. \tag{4.19}$$

We call the related method (4.18) Dual-Primal Lagrangian-PPA based contraction method. Similar to Theorem 4.5, for this method, we can obtain the following convergence theorem.

Theorem 4.7. *The sequence $\{u^k\}$ generated by the Dual-Primal Lagrangian-PPA based contraction method converges to some u^∞ which belongs to Ω^*.*

5 Convergence Rate in an Ergodic Sense

In this section, we show the convergence rate of the proposed Lagrangian-PPA based contraction method for the mixed variational inequality MVI(Ω, F, θ) (1.7). The concept of the complexity is the following theorem which is similar to Theorem 2.3.5 of [6] (see (2.3.2) in pp. 159). Only for the sake of completeness and clarity, we include it here as the following Theorem 5.1.

Theorem 5.1. *The solution set of VI(Ω, F, θ), denoted by Ω^*, is convex and can be characterized as*

$$\Omega^* = \bigcap_{u\in\Omega} \big\{\tilde{u} \in \Omega : \big(\theta(x) - \theta(\tilde{x})\big) + (u - \tilde{u})^T F(u) \geq 0\big\}. \tag{5.1}$$

Proof. . Indeed, if $\tilde{u} \in \Omega^*$, we have

$$\theta(x) - \theta(\tilde{x}) + (u - \tilde{u})^T F(\tilde{u}) \geq 0, \ \ \forall\, u \in \Omega.$$

By using the monotonicity of F on Ω, this implies

$$\theta(x) - \theta(\tilde{x}) + (u - \tilde{u})^T F(u) \geq 0, \ \ \forall\, u \in \Omega,$$

thus $\tilde{u}$ belongs to the right hand set in (5.1). Conversely, suppose $\tilde{u}$ belongs to the latter set. Let $u \in \Omega$ be arbitrary. The vector

$$\bar{u} = \alpha\tilde{u} + (1 - \alpha)u$$

belongs to Ω for all $\alpha \in (0, 1)$. Thus we have

$$\theta(\bar{x}) - \theta(\tilde{x}) + (\bar{u} - \tilde{u})^T F(\bar{u}) \geq 0. \tag{5.2}$$

Because $\theta(x)$ is convex, we have

$$\theta(\bar{x}) \leq \alpha\theta(\tilde{x}) + (1 - \alpha)\theta(x).$$

Substituting it in (5.2), we get

$$(\theta(x) - \theta(\tilde{x})) + (u - \tilde{u})^T F(\alpha\tilde{u} + (1 - \alpha)u) \geq 0,$$

for all $\alpha \in (0,1)$. Letting $\alpha \to 1$ yields

$$(\theta(x) - \theta(\tilde{x})) + (u - \tilde{u})^T F(\tilde{u}) \geq 0.$$

Thus $\tilde{u} \in \Omega^*$. Now, we turn to prove the convexity of Ω^*. For each fixed but arbitrary $u \in \Omega$, the set

$$\{\tilde{u} \in \Omega : \theta(\tilde{x}) + \tilde{u}^T F(u) \leq \theta(x) + u^T F(u)\},$$

is convex, therefore, its equivalent expression

$$\{\tilde{u} \in \Omega : \big(\theta(x) - \theta(\tilde{x})\big) + (u - \tilde{u})^T F(u) \geq 0\}$$

is convex. Since the intersection of any number of convex sets is convex, it follows that the solution set of $\mathrm{VI}(\Omega, F, \theta)$ is convex. □

According to Theorem 5.1, for given $\epsilon > 0$, we say $\bar{u}$ is an ϵ-approximate solution when

$$\sup_{D_\Omega(\bar{u})} \big\{\theta(\bar{x}) - \theta(x) + (\bar{u} - u)^T F(u)\big\} \leq \epsilon, \tag{5.3}$$

where[§]

$$D_\Omega(\bar{u}) = \big\{u \in \Omega \mid \|u - \bar{u}\| \leq 1\big\}.$$

In what follows, we show the convergence rate of the proposed Lagrangian-PPA contraction method for $\mathrm{MVI}(\Omega, F, \theta)$ (1.7) based on Lemma 3.1. Using the fact (see the notation of $F(u)$ in (1.5))

$$(u - \tilde{u}^k)^T F(u) = (u - \tilde{u}^k)^T F(\tilde{u}^k),$$

it follows from $\alpha_k > 0$ and (3.3) that

$$\tilde{u}^k \in \Omega, \quad \alpha_k\big(\theta(x) - \theta(\tilde{x}^k) + (u - \tilde{u}^k)^T F(u)\big) \geq \alpha_k (u - \tilde{u}^k)^T Q(u^k - \tilde{u}^k), \quad \forall\, u \in \Omega. \tag{5.4}$$

Lemma 5.2. *For given $u^k = (x^k, \lambda^k)$, let $\tilde{u}^k$ be generated by* (3.2) *and the new iterate u^{k+1} be updated by* (4.14). *Then we have*

$$\alpha_k\big(\theta(x) - \theta(\tilde{x}^k) + (u - \tilde{u}^k)^T F(u)\big) + \frac{1}{2}\|u - u^k\|_H^2 \geq \frac{1}{2}\|u - u^{k+1}\|_H^2,\ \forall w \in \Omega. \tag{5.5}$$

Proof. Using $Q = HM$ (see (4.5)) and the relation (4.14) to deal with the right hand side of (5.4), we get

$$\alpha_k (u - \tilde{u}^k)^T Q(u^k - \tilde{u}^k) = \alpha_k (u - \tilde{u}^k)^T HM(u^k - \tilde{u}^k) = (u - \tilde{u}^k)^T H(u^k - u^{k+1}). \tag{5.6}$$

Applying the identity

$$(a - b)^T H(c - d) = \frac{1}{2}\{\|a - d\|_H^2 - \|a - c\|_H^2\} + \frac{1}{2}\{\|c - b\|_H^2 - \|d - b\|_H^2\},$$

to the right hand side of (5.6) with

$$a = u, \quad b = \tilde{u}^k, \quad c = u^k, \quad \text{and} \quad d = u^{k+1},$$

[§]The reader may ask why we take $D_\Omega(\bar{u})$ in (5.3) instead of Ω. Consider the unconstrained optimization problem $\min_{x \in \Re^n} f(x)$, where $f(x) : \Re^n \to \Re$ is a convex differential function. The optimal condition is $\|\nabla f(x)\| = 0$. We say $\tilde{x}$ is an approximate solution if $\|\nabla f(\tilde{x})\| \leq \epsilon$. This condition is equivalent to $\sup_{\|x - \tilde{x}\| \leq 1} (x - \tilde{x})^T \nabla f(\tilde{x}) \leq \epsilon$. It is well known we can not expect $\sup_{x \in \Re^n} (x - \tilde{x})^T \nabla f(\tilde{x}) \leq \epsilon$, unless $\|\nabla f(\tilde{x})\| = 0$.

we thus obtain

$$\alpha_k(u-\tilde{u}^k)^TQ(v^k-\tilde{v}^k) = \frac{1}{2}\big(\|u-u^{k+1}\|_H^2-\|u-u^k\|_H^2\big)+\frac{1}{2}(\|u^k-\tilde{u}^k\|_H^2-\|u^{k+1}-\tilde{u}^k\|_H^2). \quad (5.7)$$

For the last term of (5.7), we have

$$\begin{aligned} &\|u^k-\tilde{u}^k\|_H^2-\|u^{k+1}-\tilde{u}^k\|_H^2 \\ &\quad = \|u^k-\tilde{u}^k\|_H^2-\|(u^k-\tilde{u}^k)-(u^k-u^{k+1})\|_H^2 \\ &\quad \overset{(4.14)}{=} \|u^k-\tilde{u}^k\|_H^2-\|(u^k-\tilde{u}^k)-\alpha_kM(u^k-\tilde{u}^k)\|_H^2 \\ &\quad = 2\alpha_k(u^k-\tilde{u}^k)HM(u^k-\tilde{u}^k)-\alpha_k^2\|M(u^k-\tilde{u}^k)\|_H^2 \\ &\quad \overset{(4.8)}{=} q(\alpha_k). \end{aligned} \quad (5.8)$$

Since $q(\alpha_k) > 0$, it follows from (5.7) and (5.8) that

$$\alpha_k(u-\tilde{u}^k)^TQ(v^k-\tilde{v}^k) = \frac{1}{2}\big(\|u-u^{k+1}\|_H^2-\|u-u^k\|_H^2\big).$$

Substituting it in (5.4), the assertion of this lemma is proved. □

Theorem 5.3. *For any integers $t>0$, we have $\tilde{u}_t \in \Omega$,*

$$\big(\theta(\tilde{x}_t)-\theta(x)\big)+(\tilde{u}_t-u)^TF(u) \le \frac{2}{\gamma(t+1)}\|u-u^0\|_H^2, \quad \forall u\in\Omega, \quad (5.9)$$

where

$$\tilde{u}_t = \frac{1}{\Upsilon_t}\sum_{k=0}^{t}\alpha_k^*\tilde{u}^k, \qquad \textit{and} \qquad \Upsilon_t = \sum_{k=0}^{t}\alpha_k^*. \quad (5.10)$$

Proof. Summing the inequality (5.5) over $k=0,1,\dots,t$, and using $\alpha_k=\gamma\alpha_k^*$, we obtain

$$\gamma\Big(\big(\sum_{k=0}^{t}\alpha_k^*\big)\theta(x)-\sum_{k=0}^{t}\alpha_k^*\theta(\tilde{x}^k)\Big)+\gamma\Big(\big(\sum_{k=0}^{t}\alpha_k^*\big)u-\sum_{k=0}^{t}\alpha_k^*\tilde{u}^k\Big)^TF(u)+\frac{1}{2}\|u-u^0\|_H^2 \ge 0, \ \ \forall\, u\in\Omega.$$

By using (5.10), it follows that

$$\Big(\frac{1}{\Upsilon_t}\sum_{k=0}^{t}\alpha_k^*\theta(\tilde{x}^k)-\theta(x)\Big)+\Big(\frac{1}{\Upsilon_t}\sum_{k=0}^{t}\alpha_k^*\tilde{u}^k-u\Big)^TF(u) \le \frac{1}{2\gamma\Upsilon_t}\|u-u^0\|_H^2, \quad \forall u\in\Omega. \quad (5.11)$$

Since

$$\tilde{x}_t = \frac{1}{\Upsilon_t}\sum_{k=0}^{t}\alpha_k^*\tilde{x}^k \qquad \text{is a convex conbination of } \tilde{x}^0,\tilde{x}^1,\dots,\tilde{x}^k,$$

and $\theta(x)$ is convex, we have

$$\theta(\tilde{x}_t) \le \frac{1}{\Upsilon_t}\sum_{k=0}^{t}\alpha_k^*\theta(\tilde{x}^k).$$

Substituting it in (5.11) and using (5.10), we obtain

$$\big(\theta(\tilde{x}_t)-\theta(x)\big)+(\tilde{u}_t-u)^TF(u) \le \frac{1}{2\gamma\Upsilon_t}\|u-u^0\|_H^2, \quad \forall u\in\Omega. \quad (5.12)$$

Since $\alpha_k^* > 1/4$, $\forall\, k>0$, we have

$$\Upsilon_t \ge \frac{1}{4}(t+1).$$

Substituting it in (5.12), the assertion of this theorem follows directly. □

For any substantial set $\mathcal{D} \subset \Omega$, the proposed contraction methods reach

$$\big(\theta(\tilde{x}_t) - \theta(x)\big) + (\tilde{u}_t - u)^T F(u) \le \epsilon, \quad \forall\, w \in \mathcal{D}, \qquad \text{in at most} \qquad t = \left\lceil \frac{D^2}{2\gamma\epsilon} \right\rceil$$

iterations, where $\tilde{u}_t$ is defined in (5.10) and $D = \sup\{\|u - u^0\| \,|\, u \in \mathcal{D}\}$. The convergence rate is in an ergodic sense.

Remark 5.4. Similar to Theorem 5.3, we can obtain the convergence rate of Dual-Primal Lagrangian-PPA based contraction method.

6 Numerical Results

In this section, we report the results of experiments to demonstrate the efficiency of Lagrangian-PPA based contraction method. To that end, we compare it with PPA [12] on different applications. Our experiments focus on efficiency and speed, and evaluate the methods in terms of their number of iterations and computational times. All the codes were written by Matlab R2009b version and all the numerical experiments were performed on a Lenovo desktop computer with Intel (R) Core(TM) i5 CPU with 3.2GHz and 3.5GB RAM.

6.1 Nearest Correlation Matrix Problem

The approximate correlation matrix problem is prevalent in many fields and has led to much interest in computing the nearest correlation matrix to a given matrix $C \in \Re^{n\times n}$, that is, solving the problem

$$\min\{\frac{1}{2}\|X - C\|_F^2 \mid \operatorname{diag}(X) = e,\, X \in S_+^n\}, \tag{6.1}$$

where $e \in \Re^n$ is the vector whose entries are all 1s, S_+^n denotes the cone of positive definite symmetric matrices, $\operatorname{diag}(X)$ is the vector of diagonal elements of X, and $\|\cdot\|_F$ denotes the matrix Fröbenis norm $\|X\|_F = (\operatorname{trace}(X^TX))^{\frac{1}{2}}$.

Here, we apply the dual-primal Lagrangian-PPA(abbreviated as as L-PPA) contraction method for solving (6.1), and its subproblem is specified into:

$$\tilde{\lambda}^k = \lambda^k - \frac{1}{s}(\operatorname{diag}(X^k) - e) \tag{6.2a}$$

and

$$\min\{\|X - \tfrac{1}{1+r}(rX^k + \operatorname{diag}(\tilde{\lambda}^k) + C)\|_F^2 \,|\, X \in S_+^n\}. \tag{6.2b}$$

It is easy to see that (6.2b) admits a closed-form solution

$$\tilde{X}^k = P_{S_+^n}[\tfrac{1}{1+r}(rX^k + \operatorname{diag}(\tilde{\lambda}^k) + C)],$$

where $P_{S_+^n}$ denotes the projection operator onto S_+^n which can be completed by an eigenvalue decomposition. In our experiments, we apply mexsvd to conduct the eigenvalue decomposition. We constructed test data sets like those of [13] and the stopping criterion for solving (6.1) was:

$$\max\{\max_{ij} |X_{ij}^k - \tilde{X}_{ij}^k|,\, \max_j |\lambda_j^k - \tilde{\lambda}_j^k|\} \le 10^{-5}.$$

In order to illustrate the efficiency of our L-PPA, we performed numerical experiments with PPA [12], with n ranging from 500 to 3000. Note that $A = I$, we used initial stepsize $s = 0.5, r = 1.01/s$ for PPA, $s = 0.5 \times 0.8, r = 0.65/s$ for L-PPA. The comparisons between these algorithms for solving (6.1) are presented in Table 1.

Table 1. Performance of PPA and L-PPA

Size	PPA		L-PPA	
n	It.	CPU	It.	CPU
500	27	2.80	22	2.27
1000	31	21.24	25	17.01
1500	36	79.61	29	63.47
2000	40	199.31	33	159.86
3000	49	1091.91	37	873.74

Both PPA and L-PPA quickly converged in a few iterations. PPA took much more iterations to achieve accurate solution and the run times of L-PPA were about 80% of PPA. Both methods employed an easy implementable framework and required one function evaluations(involving SVD computation) per iteration which are costly. In facts, L-PPA was faster than PPA which supports that optimal update stepsize can improve performance. Of course, care should be taken: For instance, the cost of computing optimal stepsize α_k here is negligible, compared to the computation of SVD.

6.2 Matrix Completion Problem

In this subsection we report experiments which demonstrate the effective performance of the L-PPA approach on matrix completion problem [2, 3]. Considering the recovery of a low-rank matrix X from under-sampled data, the matrix completion problem can be cast as the following convex programming:

$$\min\{\|X\|_* \mid X_\Pi = M_\Pi\}, \tag{6.3}$$

where $M \in \Re^{l\times n}$ is a given matrix, $\|X\|_*$ denotes the nuclear norm which is defined as the sum of all singular values of X, and $(\cdot)_\Pi : \Re^{l\times n} \to \Re^{l\times n}$ is the sampling operator defined by $X_{ij} = M_{ij}$ if $(i, j) \in \Pi$ and 0 otherwise. Applying L-PPA to this problem, the predictor takes the form

$$\tilde\lambda^k_\Pi = \lambda^k_\Pi - \frac{1}{s}(X^k_\Pi - M_\Pi)$$

and

$$\tilde X^k \quad = \quad \text{Argmin}\Big\{\frac{1}{r}\|X\|_* + \frac{1}{2}\Big\|X - (X^k + \frac{1}{r}\tilde\lambda^k_\Pi)\Big\|_F^2\Big\}. \tag{6.4}$$

In fact, the closed-form solution of (6.4) is given by:

$$\tilde\sigma_i = \max(\sigma_i - \frac{1}{r}, 0),$$

and

$$\tilde X^k = U^k\tilde\Lambda^k(V^k)^T,$$

where $\tilde\Lambda^k = \text{diag}(\tilde\sigma_1, \tilde\sigma_2, \cdots, \tilde\sigma_n)$.

Its computational cost is dominated by singular value decomposition, hence, we combine a Lanczos method PROPACK [14] to accomplish the singular value decomposition. We conducted the test examples of (6.3) in the way suggested by [3] and compared our L-PPA against PPA. Since $\|A^TA\| \le 1$, the parameters were set as follows: $s = 160, r = 1.01/s$ for PPA, $s = 160 \times 0.8, r = 0.65/s$ for L-PPA. All initial points were set to 0. As [3], we use the stopping criterion

$$\text{relative error} = \frac{\|X^k_\Pi - M_\Pi\|_F}{\|M_\Pi\|_F} \le 10^{-3}$$

as the stopping criterion. In Table 2, we report the numerical performance of these methods in various scenarios. In the following, sr denotes the sampling ratio as [18] and fr = rank $*$ $(l + n - \text{rank})/(\text{sr} * l * n)$ is the freedom of set, all test matrices were square.

Table 2. Comparison of PPA and L-PPA

Problems				PPA			L-PPA		
n	rank	sr	fr	rel.err.	It.	CPU	rel.err	It.	CPU
200	15	0.43	0.33	1.86e-004	48	3.48	1.75e-004	45	3.07
500	10	0.16	0.25	1.02e-004	57	6.62	9.77e-005	42	4.97
500	20	0.24	0.33	1.16e-004	41	5.62	1.83e-004	33	5.84
1000	10	0.12	0.17	1.00e-004	70	35.99	1.08e-004	52	29.97
1000	50	0.50	0.20	1.43e-004	40	41.85	1.23e-004	30	29.99
2000	10	0.39	0.25	1.28e-004	142	300.82	1.26e-004	111	246.85

From Table 2, we observe that L-PPA outperformed PPA in both iteration numbers and CPU times. Our numerical results show that L-PPA was competitive and it provided more robust performance than PPA.

The above experimental results confirm the efficiency of L-PPA, in particular showing its nice convergent behavior. In addition, the comparison highlights the fact that L-PPA is faster than PPA. The key reason is the less rigorous step-size choice and the optimal 'α_k'. Overall, L-PPA provide a simple and elegant framework for solving various practical problems arising from linearly constrained convex optimization.

References

[1] K.J. Arrow, L. Hurwicz and H. Uzawa, Studies in linear and non- linear programming. in *Stanford Mathematical Studies in the Social Sciences, vol. II.*, H.B. Chenery, S.M. Johnson, S. Karlin, T. Marschak, R.M. Solow (eds.) , Stanford University Press, Stanford, 1958.

[2] J.F. Cai, E.J. Candès and Z.W. Shen, A singular value thresholding algorithm for matrix completion, *SIAM Journal on Optimization* 20 (2010) 1956–1982.

[3] E.J. Candès and B. Recht, Exact matrix completion via convex optimization, *Foudations of Computational Mathematics* 9 (2008) 717–772.

[4] A. Chambolle and T. Pock, A first-order primal-dual algorithms for convex problem with applications to imaging, *Journal of Mathematical Imaging and Vision* 40 (2011) 120–145.

[5] E. Esser, X.Q. Zhang and T.F. Chan, A general framework for a class of first order primal- dual algorithms for convex pptimization in imaging science, *SIAM Journal on Imaging Sciences* 3 (2010) 1015–1046.

[6] F. Facchinei and J.S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Vol. I. Springer Series in Operations Research. Springer Verlag, New York, 2003.

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

[8] B.S. He, Inexact implicit methods for monotone general variational inequalities, *Mathematical Programming* Series A 86 (1999) 199–217.

[9] B.S. He, X.L. Fu and Z.K. Jiang, Proximal point algorithm using a linear proximal term, *JOTA* 141 (2009) 209–239.

[10] B.S. He and Y. Shen, On the convergence rate of customized proximal point algorithm for convex optimization and saddle-point problem (in Chinese), *Scientia Sinica Mathematica* 42 (2012) 515–525.

[11] B.S. He and M.H. Xu, A general framework of contraction methods for monotone variational inequalities, *Pacific Journal of Optimization* 4 (2008) 195–212.

[12] B.S. He and X.M. Yuan, Convergence analysis of primal-dual algorithms for a saddle-point problem: From contraction perspective, SIAM *Journal on Imaging Science* 5 (2012) 119–149.

[13] B.S. He and X.M. Yuan, A contraction method with implementable proximal regularization for linearly constrained convex programming, *Optimization Online*, 2010.

[14] R.M. Larsen, PropackCsoftware for large and sparse svd calculation, http://sun.stanford.edu/ rmunk/PROPACK/, 2005.

[15] B. Martinet, Regularisation d'inéquations variationelles par approximations succesives, *Revue Franqaise dInformatique et de Recherche Operationelle* 4 (1970) 154–159.

[16] R.T. Rockafellar, Monotone operators and the proximal point algorithm, *SIAM Journal on Control and Optimization* 14 (1976) 877–898.

[17] M. Zhu and T. Chan, An efficient primal-dual hybird gradient algorithm for total variation image restoration, CAM Reports 08-34, UCLA, Center for Applied Math, 2008.

[18] S.Q. Ma, D. Goldfarb and L.F. Chen. Fixed Point and Bregman Iterative Methods for Matrix Rank Minimization, *Mathematical Programming* 128 (2011) 321–353.

Manuscript received 10 April 2013
revised 8 July 2013
accepted for publication 2 November 2013

Yanfei You
Department of Mathematics, Nanjing University, Nanjing, 210093, China
E-mail address: `yanfeiyou@gmail.com`

Xiaoling Fu
Institute of Systems Engineering, Southeast University, Nanjing, 210096, China
E-mail address: `fuxlnju@hotmail.com`

Bingsheng He
Department of Mathematics, Nanjing University, Nanjing, 210093, China
E-mail address: `hebma@nju.edu.cn`