凸优化和单调变分不等式的收缩算法

第九讲: 基于线性变分不等式 PC 方法的求解复合凸优化问题的收缩算法

Contraction methods for composite convex optimization based on the PC Algorithms for LVIs

南方科技大学数学系　何炳生
hebma@nju.edu.cn

摘要. 上世纪九十年代, 我们发表了一些求解单调线性变分不等式(及凸二次优化)的投影收缩算法. 这些算法被他人用来求解一些应用问题有很好的表现. 在处理一类机器人控制问题上, 起到了其他算法不能替代的作用[3, 11, 12]. 本文介绍如何将这些方法推广, 用来求解复合凸优化问题.

1 Introduction

In the 1990s, we have published some projection and contraction algorithms for solving monotone linear variational inequalities [5, 6]. These algorithms can be applied to solve the constrained convex optimization problems

$$\min\{\frac{1}{2}x^T Hx + c^T x \,|\, x \in \mathcal{X}\} \tag{1.1}$$

and

$$\min\{\frac{1}{2}x^T Hx + c^T x \,|\, Ax = b (\text{or} \geq b), x \in \mathcal{X}\}, \tag{1.2}$$

where $H \in \Re^{n\times n}$ is a symmetric positive semi-definite matrix, $A \in \Re^{m\times n}$, $b \in \Re^m$, $c \in \Re^n$ and $\mathcal{X} \subset \Re^n$ is a closed convex set. The purpose of this article is to develop such algorithms to solve the following composite convex optimization problems:

$$\min\{\theta(x) + \frac{1}{2}x^T Hx + c^T x \,|\, x \in \mathcal{X}\} \tag{1.3}$$

and

$$\min\{\theta(x) + \frac{1}{2}x^T Hx + c^T x \,|\, Ax = b(\text{or} \ge b), x \in \mathcal{X}\}, \tag{1.4}$$

where $\theta(x) : \Re^n \to \Re$ is a convex function (not necessarily smooth), $H,\ A,\ b,\ c$ and $\mathcal{X}$ are the same as described in (1.1) and (1.2).

Throughout this article, we assume that the solution set (1.3) and (1.4) are nonempty. In addition, we assume that for any given constant $r > 0$ and vector $a \in \Re^n$, the subproblem

$$\min\{\theta(x) + \frac{r}{2}\|x - a\|^2 \mid x \in \mathcal{X}\} \tag{1.5}$$

has a closed-form solution or can be efficiently computed with a high precision.

The analysis of this note is based on the following lemma (proof is omitted here).

Lemma 1.1 *Let* $\mathcal{X} \subset \Re^n$ *be a closed convex set,* $\theta(x)$ *and* $f(x)$ *be convex functions and* $f(x)$ *is differentiable on an open set which includes* Ω*. Assume that the solution set of the minimization problem* $\min\{\theta(x) + f(x) \,|\, x \in \mathcal{X}\}$ *is nonempty. Then,*

$$x^* \in \arg\min\{\theta(x) + f(x) \,|\, x \in \mathcal{X}\} \tag{1.6a}$$

if and only if

$$x^* \in \mathcal{X}, \quad \theta(x) - \theta(x^*) + (x - x^*)^T \nabla f(x^*) \geq 0, \quad \forall\, x \in \mathcal{X}. \tag{1.6b}$$

2 The algorithms for solving (1.3) based on P-C Algorithm for the problem (1.1)

In (1.3), since the $\theta(x)$ is convex and H is semidefinite, by using Lemma 1.1, the optimal solution of (1.3), say x^*, satisfies

$$x^* \in \mathcal{X}, \quad \theta(x) - \theta(x^*) + (x - x^*)^T (Hx^* + c) \geq 0, \quad \forall\, x \in \mathcal{X}. \tag{2.1}$$

2.1 Key inequality for solving $\min\{\frac{1}{2}x^T H x + c^T x \mid x \in \mathcal{X}\}$

Set $\theta(x) = 0$ in (1.3), it is reduced to the problem (1.1) whose optimal solution x^* satisfies

$$x^* \in \mathcal{X}, \quad (x - x^*)^T (Hx^* + c) \geq 0, \quad \forall\, x \in \mathcal{X}. \qquad (2.2)$$

For solving (1.1) (or its equivalent (2.2)), we have proposed a class of projection and contraction algorithms [5, 6]. These algorithms are based on constructing the descent direction of the distance function $\frac{1}{2}\|x - x^*\|_G^2$, where G is some symmetric positive definite matrix.

For given $x^k \in \Re^n$ and $\beta > 0$, let

$$\tilde{x}^k = P_{\mathcal{X}}[x^k - \beta(Hx^k + c)]. \qquad (2.3)$$

x^k is the optimal solution of (1.1) (or its equivalent (2.2)) if and only if $\tilde{x}^k = x^k$.

The projector $\tilde{x}^k$ is the solution of the minimization problem,

$$\tilde{x}^k = \arg\min\{\frac{1}{2}\|x - [x^k - \beta(Hx^k + c)]\|^2 \,|\, x \in \mathcal{X}\}.$$

According to Lemma 1.1, we have

$$\tilde{x}^k \in \mathcal{X}, \quad (x - \tilde{x}^k)^T\{\tilde{x}^k - [x^k - \beta(Hx^k + c)]\} \geq 0, \quad \forall x \in \mathcal{X}.$$

Set the any vector $x \in \mathcal{X}$ in the above inequality by a solution point x^*, it follows that

$$(\tilde{x}^k - x^*)^T\{(x^k - \tilde{x}^k) - \beta(Hx^k + c)\} \geq 0. \tag{2.4}$$

On the other hand, since $\tilde{x}^k \in \mathcal{X}$, it follows from (2.2) that

$$(\tilde{x}^k - x^*)^T\beta(Hx^* + c) \geq 0. \tag{2.5}$$

Adding (2.4) and (2.5), we get

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

The above inequality can be rewritten as

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

Finally, by using the semi-positiveness of H, we get

$$(x^k - x^*)^T (I + \beta H)(x^k - \tilde{x}^k) \geq \|x^k - \tilde{x}^k\|^2. \tag{2.6}$$

The above inequality is the main basis for building the projection contraction algorithms for solving (1.1) (and its equivalent variational inequality (2.2)). We hope to establish the same inequality for the problem (1.3).

2.2 Key inequality for solving $\min\{\theta(x) + \frac{1}{2}x^T H x + c^T x | x \in \mathcal{X}\}$

Our task is to solve (1.3). The purpose of this subsection is to construct the same key-inequality as (2.6). For given x^k and $\beta > 0$, we let

$$\tilde{x}^k = \arg\min\{\theta(x) + \frac{1}{2\beta}\|x - [x^k - \beta(Hx^k + c)]\|^2 \,|\, x \in \mathcal{X}\}. \tag{2.7}$$

This is an optimization problem as (1.5) and it is assumed to be solved without difficulty. According to Lemma 1.1, we have $\tilde{x}^k \in \mathcal{X}$ and

$$\theta(x) - \theta(\tilde{x}^k) + (x - \tilde{x}^k)^T \frac{1}{\beta}\big\{\tilde{x}^k - [x^k - \beta(Hx^k + c)]\big\} \geq 0, \ \forall x \in \mathcal{X}.$$

We rewrite it as $\tilde{x}^k \in \mathcal{X}$ and

$$\beta\theta(x) - \beta\theta(\tilde{x}^k) + (x - \tilde{x}^k)^T\{\beta(Hx^k + c) - (x^k - \tilde{x}^k)\} \geq 0, \ \forall x \in \mathcal{X}. \quad (2.8)$$

Since $\tilde{x}^k \in \mathcal{X}$, according to the optimal condition (2.1), we have

$$\beta\theta(\tilde{x}^k) - \beta\theta(x^*) + (\tilde{x}^k - x^*)^T \beta(Hx^* + c) \geq 0. \quad (2.9)$$

Setting the any $x \in \Omega$ in (2.8) by x^* and then adding it with (2.9), we obtain

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

Rewriting the above inequality in form

$$\{(x^k - x^*) - (x^k - \tilde{x}^k)\}^T\{(x^k - \tilde{x}^k) - \beta H(x^k - x^*)\} \geq 0$$

and using the semi-positivity of H, we get

$$(x^k - x^*)^T (I + \beta H)^T (x^k - \tilde{x}^k) \geq \|x^k - \tilde{x}^k\|^2. \qquad (2.10)$$

This is the same inequality as (2.6). Usually, we call the vector $\tilde{x}^k$ obtained by (2.7) the predictor in the k-th iteration of the proposed algorithm for solving the convex optimization problem (1.3).

2.3 Solving the optimization problem (1.3) by using the key inequality (2.10)

The inequality (2.10) can be written as

$$\langle (I + \beta H)(x^k - x^*), (x^k - \tilde{x}^k) \rangle \geq \|x^k - \tilde{x}^k\|^2.$$

Let $G = (I + \beta H)$, the above inequality tells us that $-(x^k - \tilde{x}^k)$ is the descent direction of the unknown distance function $\frac{1}{2}\|x - x^*\|_G^2$ at x^k.

We take

$$x^{k+1}(\alpha) = x^k - \alpha(x^k - \tilde{x}^k) \tag{2.11}$$

as the step length α dependent new iterate. In order to shorten the distance $\|x - x^*\|^2_{(I+\beta H)}$, we consider the following α-dependent benefit

$$\vartheta_k(\alpha) = \|x^k - x^*\|^2_{(I+\beta H)} - \|x^{k+1}(\alpha) - x^*\|^2_{(I+\beta H)}. \tag{2.12}$$

By using (2.10), we obtain

$$\begin{aligned} \vartheta_k(\alpha) &= \|x^k - x^*\|^2_{(I+\beta H)} - \|x^k - x^* - \alpha(x^k - \tilde{x}^k)\|^2_{(I+\beta H)} \\ &= 2\alpha(x^k - x^*)^T(I + \beta H)(x^k - \tilde{x}^k) \\ &\quad -\alpha^2\|x^k - \tilde{x}^k\|^2_{(I+\beta H)}. \end{aligned} \tag{2.13}$$

By using the (2.10), we have the following theorem.

Theorem 2.1 *For given x^k and any $\beta > 0$, let $\tilde{x}^k$ be a predictor generated by* (2.7) *and $x^{k+1}(\alpha)$ be updated by* (2.11)*. Then for any $\alpha > 0$, we have*

$$\vartheta_k(\alpha) \ge q_k(\alpha), \tag{2.14}$$

where $\vartheta_k(\alpha)$ *is defined by* (2.12) *and*

$$q_k(\alpha) = 2\alpha\|x^k - \tilde{x}^k\|^2 - \alpha^2\|x^k - \tilde{x}^k\|^2_{(I+\beta H)}. \tag{2.15}$$

Proof. The assertion of this theorem follows directly from (2.13) and (2.10). $\square$

Now, $q_k(\alpha)$ is a lower bound function of $\vartheta_k(\alpha)$. The quadratic function $q_k(\alpha)$ reaches its maximum at

$$\alpha_k^* = \frac{\|x^k - \tilde{x}^k\|^2}{(x^k - \tilde{x}^k)^T(I + \beta H)(x^k - \tilde{x}^k)}. \tag{2.16}$$

We use

$$x^{k+1} = x^k - \gamma\alpha_k^*(x^k - \tilde{x}^k), \quad \gamma \in (0, 2) \tag{2.17}$$

to determine the new iterate x^{k+1}. By using (2.15) and (2.16), we get

$$\begin{aligned} q(\gamma\alpha_k^*) &= 2\gamma\alpha_k^*\|x^k - \tilde{x}^k\|^2 - \gamma^2\alpha_k^*\big(\alpha_k^*\|x^k - \tilde{x}^k\|^2_{(I+\beta H)}\big) \\ &= \gamma(2 - \gamma)\alpha_k^*\|x^k - \tilde{x}^k\|^2. \end{aligned} \tag{2.18}$$

Thus, the profit of the k-th iteration

$$\begin{aligned}\vartheta(\gamma\alpha_k^*) &= \|x^k - x^*\|^2_{(I+\beta H)} - \|x^{k+1} - x^*\|^2_{(I+\beta H)} \\ &\geq q(\gamma\alpha_k^*) = \gamma(2-\gamma)\alpha_k^*\|x^k - \tilde{x}^k\|^2. \end{aligned} \tag{2.19}$$

Theoretically $\vartheta(\gamma\alpha_k^*) > 0$ when $\gamma \in (0,2)$, usually, we take $\gamma \in [1.2, 1.8]$.

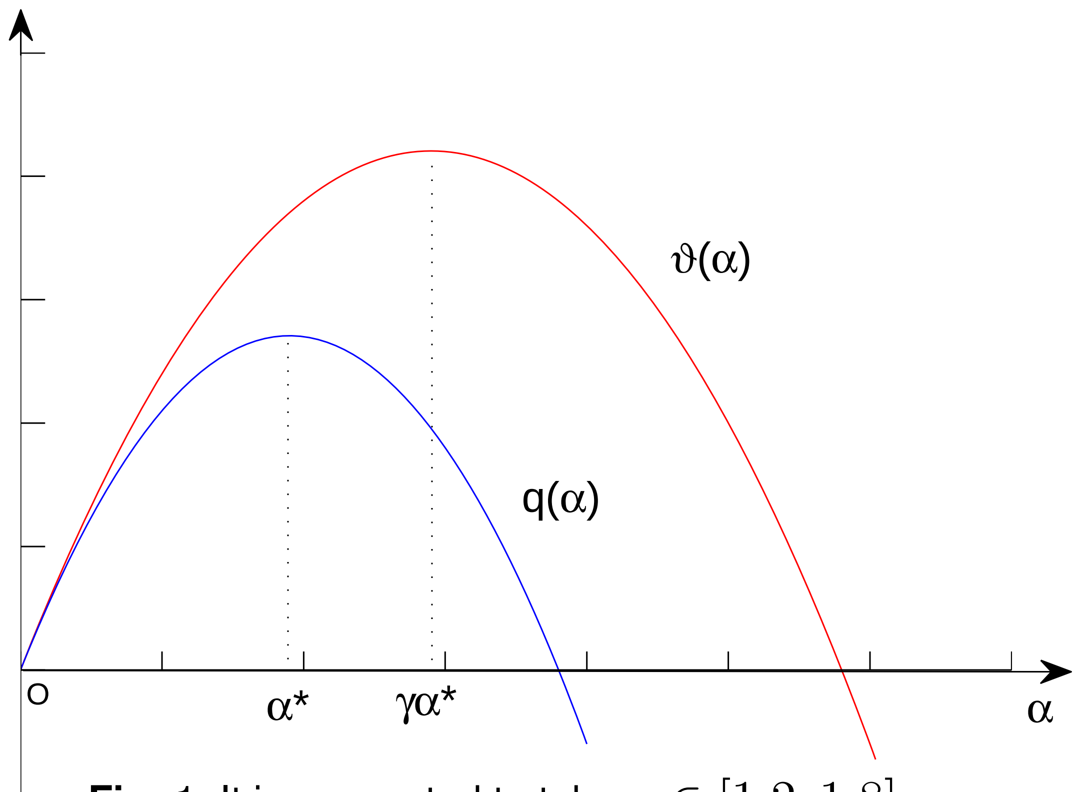

Fig. 1 It is suggested to take $\gamma \in [1.2, 1.8]$

For solving the composite convex optimization problem (1.3), we use (2.7) to generate the predictor $\tilde{x}^k$ and (2.17) to update the corrector x^{k+1}. Together with some strategy for adjusting the parameter β, we have the following algorithm.

Algorithm 2.1 Prediction-Correction method for solving the composite convex optimization problem (1.3):

Start with given x^0 and $\beta > 0$. For $k = 0, 1, \ldots$, do:

1. Prediction: $\tilde{x}^k = \arg\min\{\theta(x) + \frac{1}{2\beta}\|x - [x^k - \beta(Hx^k + c)]\|^2 \mid x \in \mathcal{X}\}$.
2. Correction: $x^{k+1} = x^k - \gamma\alpha_k^*(x^k - \tilde{x}^k)$,

   where $\alpha_k^* = \dfrac{\|x^k - \tilde{x}^k\|^2}{(x^k - \tilde{x}^k)^T(I + \beta H)(x^k - \tilde{x}^k)}$ and $\gamma \in (0, 2)$.
3. Adjust the parameter β if necessary

$$r = \frac{\beta\|x^k - \tilde{x}^k\|_H^2}{\|x^k - \tilde{x}^k\|^2}, \qquad \beta = \begin{cases} (\beta/r) * 0.9, & \text{if } r > 1 \quad \text{or} \quad r < 0.4; \\ \beta, & \text{otherwise.} \end{cases}$$

$k := k + 1$ and go to 1.

According to our numerical experiments, the parameter β should be selected in range

$$\frac{2}{5}\|x^k-\tilde{x}^k\| \le \beta\|x^k-\tilde{x}^k\|_H^2 \le \|x^k-\tilde{x}^k\|.$$

We can also adjust the parameters every five or ten iterations. In practical, after some iterations, the algorithm will automatically find a suitable fixed β.

Theorem 2.2 *For solving the composite convex optimization problem* (1.3)*, the sequence* $\{x^k\}$ *generated by Algorithm 2.1 with fixed* $\beta > 0$ *satisfies*

$$\|x^{k+1}-x^*\|_{(I+\beta H)}^2 \le \|x^k-x^*\|_{(I+\beta H)}^2 - \frac{\gamma(2-\gamma)}{\|I+\beta H\|}\|x^k-\tilde{x}^k\|^2. \quad (2.20)$$

Proof. From (2.16) we have $\alpha_k^* \ge \frac{1}{\|I+\beta H\|}$. Together with (2.19), it follows the assertion (2.20) directly. $\square$

Generally speaking, whether the selection of parameters is appropriate or not, it will affect the convergence speed. Without loss of the generality, Theorem 2.2 gives us the key inequality for the convergence of Algorithm 2.1.

3 The algorithms for solving (1.4) based on P-C Algorithm for the problem (1.2)

The Lagrangian function of the linearly constrained optimization problem (1.4) is

$$L(x,\lambda) = \theta(x) + \frac{1}{2}x^T Hx + c^T x - \lambda^T (Ax - b),$$

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

$$\Lambda = \begin{cases} \Re^m, & \text{if the lineary constraints in (1.4) is } Ax = b, \\ \Re^m_+, & \text{if the lineary constraints in (1.4) is } Ax \geq b. \end{cases}$$

Let $(x^*, \lambda^*) \in \mathcal{X} \times \Lambda$ be a saddle point of the Lagrangian function, we have

$$L_{\lambda \in \Re^m}(x^*, \lambda) \leq L(x^*, \lambda^*) \leq L_{x \in \mathcal{X}}(x, \lambda^*).$$

This tells us that $(x^*, \lambda^*) \in \mathcal{X} \times \Lambda$ and

$$\begin{cases} L(x, \lambda^*) - L(x^*, \lambda^*) \geq 0, & \forall\, x \in \mathcal{X}, \\ L(x^*, \lambda^*) - L(x^*, \lambda) \geq 0, & \forall\, \lambda \in \Lambda. \end{cases}$$

Thus, finding a saddle point of the Lagrange function is equivalent to finding $(x^*, \lambda^*) \in \mathcal{X} \times \Lambda$ such that

$$\begin{cases} \theta(x) - \theta(x^*) + (x - x^*)^T (Hx^* + c - A^T \lambda^*) \geq 0, & \forall\, x \in \mathcal{X}, \\ \qquad\qquad\qquad\qquad (\lambda - \lambda^*)^T (Ax^* - b) \geq 0, & \forall\, \lambda \in \Lambda. \end{cases} \tag{3.1}$$

The above variational inequality can be written in a compact form

$$u^* \in \Omega, \quad \theta(x) - \theta(x^*) + (u - u^*)^T (Mu^* + q) \geq 0, \quad \forall\, u \in \Omega, \tag{3.2a}$$

where

$$u = \begin{pmatrix} x \\ \lambda \end{pmatrix}, \quad M = \begin{pmatrix} H & -A^T \\ A & 0 \end{pmatrix}, \quad q = \begin{pmatrix} c \\ -b \end{pmatrix} \quad \text{and} \quad \Omega = \mathcal{X} \times \Lambda. \tag{3.2b}$$

Although the matrix M is not symmetric, however, for any u, we have $u^T M u = x^T H x \geq 0$. The variational inequality (3.2) is monotone.

3.1 Key inequality for solving the minimization problem $\min\{\frac{1}{2}x^T Hx + c^T x \mid Ax = (\text{or } \geq)\, b, x \in \mathcal{X}\}$

Set $\theta(x) = 0$ (1.4), the problem is reduced to the linear constrained quadratic programming (1.2), the related variational inequality (3.2) is reduced to

$$u^* \in \Omega, \quad (u - u^*)^T (Mu^* + q) \geq 0, \quad \forall\, u \in \Omega, \tag{3.3a}$$

where

$$u = \begin{pmatrix} x \\ \lambda \end{pmatrix}, \quad M = \begin{pmatrix} H & -A^T \\ A & 0 \end{pmatrix}, \quad q = \begin{pmatrix} c \\ -b \end{pmatrix} \text{ and } \Omega = \mathcal{X} \times \Lambda. \tag{3.3b}$$

For solving (1.2) (or its equivalent (3.3)), we have proposed a projection and contraction method [5] whose search direction is established in the following way.

First, for given $u^k \in \Re^{m+n}$ and $\beta > 0$, let

$$\tilde{u}^k = P_\Omega[u^k - \beta(Mu^k + q)]. \tag{3.4}$$

Since

$$\tilde{u}^k = \arg\min\{\frac{1}{2}\|u - [u^k - \beta(Mu^k + q)]\|^2 | u \in \Omega\},$$

according to Lemma 1.1, it follows that

$$\tilde{u}^k \in \Omega, \quad (u - \tilde{u}^k)^T\{\tilde{u}^k - [u^k - \beta(Mu^k + q)]\} \ge 0, \quad \forall u \in \Omega.$$

Setting the any $u \in \Omega$ in the above inequality by u^*, we get

$$(\tilde{u}^k - u^*)^T\{(u^k - \tilde{u}^k) - \beta(Mu^k + q)\} \ge 0. \tag{3.5}$$

Because $\tilde{u}^k \in \Omega$ and $\beta > 0$, according to (3.3a), we have

$$(\tilde{u}^k - u^*)^T\beta(Mu^* + q) \ge 0. \tag{3.6}$$

Adding the above two inequalities, we get

$$(\tilde{u}^k - u^*)^T\{(u^k - \tilde{u}^k) - \beta M(u^k - u^*)\} \ge 0. \tag{3.7}$$

It can be written as

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

Consequently, using the semi-positivity of M ($v^TMv \ge 0$), it follows that

$$(u^k - u^*)^T(I + \beta M^T)(u^k - \tilde{u}^k) \ge \|u^k - \tilde{u}^k\|^2, \quad \forall u^* \in \Omega^*. \tag{3.8}$$

The above inequality is the main basis for building the projection and contraction algorithms for solving the linearly constrained convex quadratic optimization (1.2) (and its equivalent variational inequality (3.2)). We hope to establish the same inequality for the composite optimization problem (1.4).

3.2 Key inequality for solving the minimization problem $\min\{\theta(x) + \frac{1}{2}x^THx + c^Tx \mid Ax = (\text{or} \ge)\, b, x \in \mathcal{X}\}$

Our task is to solve the problem (1.4). This subsection will construct the same key-inequality as (3.8). For given $u^k = (x^k, \lambda^k)$ and $\beta > 0$, let

$$\tilde{x}^k = \arg\min\{\theta(x) + \frac{1}{2\beta}\|x - [x^k - \beta(Hx^k + c - A^T\lambda^k)]\|^2 \mid x \in \mathcal{X}\} \tag{3.9a}$$

and

$$\tilde{\lambda}^k = \arg\min\{\frac{1}{2}\|\lambda - [\lambda^k - \beta(Ax^k - b)] \mid \lambda \in \Lambda\}, \tag{3.9b}$$

Since (3.9a) is a convex optimization problem as (1.5) which has assumed to have closed form solution. Thus, there is no difficulty to obtain $\tilde{x}^k$ and $\tilde{\lambda}^k$ in parallel. According to Lemma 1.1, we have $\tilde{x}^k \in \mathcal{X}$ and

$$\theta(x) - \theta(\tilde{x}^k) + (x - \tilde{x}^k)^T \frac{1}{\beta}\big\{\tilde{x}^k - [x^k - \beta(Hx^k + c - A^T\lambda^k)]\big\} \geq 0, \ \forall x \in \mathcal{X}.$$

Taking a solution point x^* as the any point $x \in \mathcal{X}$, it follows that

$$\beta\theta(x^*) - \beta\theta(\tilde{x}^k) + (\tilde{x}^k - x^*)^T\{(x^k - \tilde{x}^k) - \beta(Hx^k + c - A^T\lambda^k)\} \geq 0. \tag{3.10}$$

On the other hand, since $\tilde{x}^k \in \mathcal{X}$, it follows from (3.1) that

$$\beta\theta(\tilde{x}^k) - \beta\theta(x^*) + (\tilde{x}^k - x^*)^T\beta\{Hx^* + c - A^T\lambda^*\} \geq 0, \ \forall\, x^* \in \mathcal{X}^*. \tag{3.11}$$

Adding (3.10) and (3.11), we obtain

$$(\tilde{x}^k - x^*)^T\{(x^k - \tilde{x}^k) - \beta[H(x^k - x^*) - A^T(\lambda^k - \lambda^*)]\} \ge 0. \quad (3.12)$$

For (3.9b), according to Lemma 1.1, we have

$$\tilde{\lambda}^k \in \Lambda, \quad (\lambda - \tilde{\lambda}^k)^T\big\{\tilde{\lambda}^k - [\lambda^k - \beta(Ax^k - b)]\big\} \ge 0, \quad \forall\, \lambda \in \Lambda.$$

Set the fixed any point $\lambda \in \Lambda$ in the last inequality by λ^*, it follows that

$$(\tilde{\lambda}^k - \lambda^*)^T\{(\lambda^k - \tilde{\lambda}^k) - \beta(Ax^k - b)\} \ge 0. \quad (3.13)$$

On the other hand, since $\tilde{\lambda}^k \in \Lambda$, according the second part of (3.1), we have

$$(\tilde{\lambda}^k - \lambda^*)^T\beta(Ax^* - b) \ge 0. \quad (3.14)$$

Adding (3.13) and (3.14), it follows that

$$(\tilde{\lambda}^k - \lambda^*)^T\{(\lambda^k - \tilde{\lambda}^k) - \beta A(x^k - x^*)\} \ge 0. \quad (3.15)$$

Write (3.12) and (3.15) together,

$$\begin{pmatrix} \tilde{x}^k - x^* \\ \tilde{\lambda}^k - \lambda^* \end{pmatrix}^T \left\{ \begin{pmatrix} x^k - \tilde{x}^k \\ \lambda^k - \tilde{\lambda}^k \end{pmatrix} - \beta \begin{bmatrix} H & -A^T \\ A & 0 \end{bmatrix} \begin{pmatrix} x^k - x^* \\ \lambda^k - \lambda^* \end{pmatrix} \right\} \geq 0.$$

Using the notation in (3.3), it can be written as

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

Clearly, it is same inequality as (3.7) in §3.1. It can be written as

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

Consequently, using the semi-positivity of M ($v^T M v \geq 0$), it follows that

$$(u^k - u^*)^T (I + \beta M^T)(u^k - \tilde{u}^k) \geq \|u^k - \tilde{u}^k\|^2, \quad \forall u^* \in \Omega^*. \quad (3.16)$$

This is the same key inequality as (3.8). For solving the linearly constrained composite convex optimization problem (1.4), we call the vector $\tilde{u}^k = (\tilde{x}^k, \tilde{\lambda}^k)$ obtained from (3.9) a predictor in the k-iteration.

3.3 Solving the optimization problem (1.4) by using the key inequality (3.16)

The task of the contraction method is to generate the new iterate which is more closed to the solution set. Thus, we use

$$u^{k+1}(\alpha) = u^k - \alpha(I + \beta M^T)(u^k - \tilde{u}^k) \tag{3.17}$$

to update the new iterate which is dependent on the step length α.

In the following we discuss how to chose the step length α. For this purpose, we consider the following α-dependent benefit function

$$\vartheta_k(\alpha) := \|u^k - u^*\|^2 - \|u^{k+1}(\alpha) - u^*\|^2. \tag{3.18}$$

According to the definition,

$$\begin{aligned} \vartheta_k(\alpha) &= \|u^k - u^*\|^2 - \|u^k - u^* - \alpha(I + \beta M^T)(u^k - \tilde{u}^k)\|^2 \\ &= 2\alpha(u^k - u^*)^T(I + \beta M^T)(u^k - \tilde{u}^k) \\ &\quad -\alpha^2\|(I + \beta M^T)(u^k - \tilde{u}^k)\|^2. \end{aligned} \tag{3.19}$$

Theorem 3.1 *For given* u^k *and any* $\beta > 0$*, let* $\tilde{u}^k$ *be a predictor generated by* (3.9) *and* $u^{k+1}(\alpha)$ *be updated by* (3.17)*. Then for any* $\alpha > 0$*, we have*

$$\vartheta_k(\alpha) \ge q_k(\alpha), \tag{3.20}$$

where $\vartheta_k(\alpha)$ *is defined by* (3.18) *and*

$$q_k(\alpha) = 2\alpha\|u^k - \tilde{u}^k\|^2 - \alpha^2\|(I + \beta M^T)(u^k - \tilde{u}^k)\|^2. \tag{3.21}$$

Proof. The assertion of this theorem follows from (3.19) and (3.16) directly. $\square$

Theorem 3.1 tells us $q_k(\alpha)$ is a lower bound function of the profit function $\vartheta_k(\alpha)$. The quadratic function $q_k(\alpha)$ (3.21) reaches its maximum at

$$\alpha_k^* = \text{argmax}\{q_k(\alpha)\} = \frac{\|u^k - \tilde{u}^k\|^2}{\|(I + \beta M^T)(u^k - \tilde{u}^k)\|^2}. \tag{3.22}$$

Notice that our intention is to maximize the quadratic profit function $\vartheta_k(\alpha)$ (see (3.19)), which includes the unknown solution u^*. As a remedy, we maximize its lower bound function $q_k(\alpha)$.

In the practical computation, we take a $\gamma \in [1, 2)$ and

$$u^{k+1} = u^k - \gamma\alpha_k^*(I + \beta M^T)(u^k - \tilde{u}^k), \tag{3.23}$$

to produce the new iterate. The reason to take $\gamma \in [1, 2)$ is illustrated in Fig. 1 of §2.3. Using (3.18) and (3.20), the new iterate u^{k+1} updated by (3.23) satisfies

$$\|u^{k+1} - u^*\|^2 \le \|u^k - u^*\|^2 - q_k(\gamma\alpha_k^*). \tag{3.24}$$

According to the definitions of $q_k(\alpha)$ and α_k^* (see (3.21) and (3.22)), we get

$$\begin{aligned} q_k(\gamma\alpha_k^*) &= 2\gamma\alpha_k^*\|u^k - \tilde{u}^k\|^2 - \gamma^2\alpha_k^*(\alpha_k^*\|(I + \beta M^T)(u^k - \tilde{u}^k)\|^2) \\ &= 2\gamma\alpha_k^*\|u^k - \tilde{u}^k\|^2 - \gamma^2\alpha_k^*\|u^k - \tilde{u}^k\|^2 \\ &= \gamma(2 - \gamma)\alpha_k^*\|u^k - \tilde{u}^k\|^2. \end{aligned}$$

Thus, from (3.24) we get

$$\|u^{k+1} - u^*\|^2 \le \|u^k - u^*\|^2 - \gamma(2 - \gamma)\alpha_k^*\|u^k - \tilde{u}^k\|^2. \tag{3.25}$$

Algorithm 3.1 Prediction-Correction method for solving the composite convex optimization problem (1.4):

Start with given $u^0 = (x^0, \lambda^0)$ and $\beta > 0$. For $k = 0, 1, \ldots$, do:

1. Prediction:

   $\tilde{x}^k = \arg\min\{\theta(x) + \frac{1}{2\beta}\|x - [x^k - \beta(Hx^k + c - A^T\lambda^k)]\|^2 \mid x \in \mathcal{X}\}$

   and $\tilde{\lambda}^k = \arg\min\{\frac{1}{2}\|\lambda - [\lambda^k - \beta(Ax^k - b)] \mid \lambda \in \Lambda\}$.

2. Correction: $u^{k+1} = u^k - \gamma\alpha_k^*(I + \beta M^T)(u^k - \tilde{u}^k)$,

   where $\alpha_k^* = \dfrac{\|u^k - \tilde{u}^k\|^2}{\|(I + \beta M^T)(u^k - \tilde{u}^k)\|^2}$. and $\gamma \in (0, 2)$.

3. Adjust the parameter β if necessary

$$r = \frac{\|(I + \beta M^T)(u^k - \tilde{u}^k)\|}{\|u^k - \tilde{u}^k\|}, \quad \beta = \begin{cases} \beta * (2.5/r) & \text{if } r > 3 \\ \beta * (2.5/r) & \text{if } r < 2 \\ \beta & \text{otherwise.} \end{cases}$$

$k := k + 1$ and go to 1.

For solving the problem (1.4), the k-th iteration of the Algorithm3.1 start from u^k, produces the predictor $\tilde{u}^k$ by (3.9), and updated the new iterate u^{k+1} by (3.23). For this algorithm, we have the following theorem.

Theorem 3.2 *Let* $\{u^k\}$ *be the sequence generated by the Algorithm3.1 for the problem* (1.4)*. Then we have*

$$\|u^{k+1} - u^*\|^2 \le \|u^k - u^*\|^2 - \frac{\gamma(2-\gamma)}{\|I + \beta M^T\|^2}\|u^k - \tilde{u}^k\|^2. \tag{3.26}$$

Proof. In fact, by using (3.22), we get

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

Thus, it follows from (3.25) that

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

Substituting it in (3.24), the assertion (3.26) follows directly. $\Box$

Remark Since the predictor $\tilde{u}^k$ generated by (3.9) satisfies (3.16), as in [6], we can take

$$u^{k+1} = u^k - \gamma(I + \beta M)^{-1}(u^k - \tilde{u}^k) \tag{3.27}$$

as the new iterate. For $G = (I + \beta M^T)(I + \beta M)$, by using (3.16), we get

$$\begin{aligned} \|u^{k+1} - u^*\|_G^2 &= \|(u^k - u^*) - \gamma(I + \beta M)^{-1}(u^k - \tilde{u}^k)\|_G^2 \\ &= \|u^k - u^*\|_G^2 - 2\gamma(u^k - u^*)^T(I + \beta M)^T(u^k - \tilde{u}^k) \\ &\quad + \gamma^2\|(I + \beta M)^{-1}(u^k - \tilde{u}^k)\|_G^2 \\ &\le \|u^k - u^*\|_G^2 - \gamma(2 - \gamma)\|u^k - \tilde{u}^k\|^2. \end{aligned}$$

Thus, for solving the optimization problem (1.4), if we use (3.9) to offer the predictor and (3.27) to update the new iterate, then the sequence $\{u^k\}$ satisfies

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

4 Convergence rate of the Algorithm 2.1

This section studies the convergence rate of the Algorithm 2.1 for the composite convex optimization problem (1.3) (or its equivalent variational inequality (2.1)) :

$$\theta(x) - \theta(x^*) + (x - x^*)^T (Hx^* + c) \geq 0, \quad \forall x \in \mathcal{X}. \tag{4.1}$$

For given $x^k \in \Re^n$, the predictor $\tilde{x}^k$ in Algorithm 2.1 is offered by (2.7). The new iterate of the k-th iteration of Algorithm 2.1 is updated by

$$x^{k+1} = x^k - \gamma \alpha_k^* (x^k - \tilde{x}^k), \tag{4.2}$$

where

$$\alpha_k^* = \frac{\|x^k - \tilde{x}^k\|^2}{\|x^k - \tilde{x}^k\|_G^2}, \qquad G = I + \beta H \qquad \text{and} \qquad \gamma \in (0, 2). \tag{4.3}$$

It was proved [6] that the sequence $\{x^k\}$ generated by the Algorithm 2.1 satisfies

$$\|x^{k+1} - x^*\|_G^2 \leq \|x^k - x^*\|_G^2 - \gamma(2 - \gamma)\alpha_k^* \|x^k - \tilde{x}^k\|^2. \tag{4.4}$$

Recall that $\mathcal{X}^*$ can be characterized as (see Theorem 2.1 in [8])

$$\mathcal{X}^* = \bigcap_{x \in \mathcal{X}} \big\{\tilde{x} \in \mathcal{X} : \theta(x) - \theta(\tilde{x}) + (x - \tilde{x})^T (Hx + c) \geq 0\big\}.$$

This implies that $\tilde{x} \in \mathcal{X}$ is an approximate solution of (4.1) with the accuracy ϵ if it satisfies

$$\tilde{x} \in \mathcal{X} \quad \text{and} \quad \inf_{x \in \mathcal{D}(\tilde{x})} \big\{\theta(x) - \theta(\tilde{x}) + (x - \tilde{x})^T (Hx + c)\big\} \geq -\epsilon,$$

where

$$\mathcal{D}(\tilde{x}) = \{x \in \mathcal{X} \mid \|x - \tilde{x}\| \leq 1\}.$$

In this section, we show that, for given $\epsilon > 0$, in $O(1/\epsilon)$ iterations the Algorithm 2.1 can find a $\tilde{x}$ such that

$$\tilde{x} \in \Omega \quad \text{and} \quad \sup_{x \in \mathcal{D}(\tilde{x})} \big\{\big(\theta(\tilde{x}) - \theta(x)\big) + (\tilde{x} - x)^T (Hx + c)\big\} \leq \epsilon, \quad (4.5)$$

where

$$\mathcal{D}(\tilde{x}) = \{x \in \Omega \mid \|x - \tilde{x}\| \leq 1\}.$$

In this sense, we will establish the algorithmic convergence complexity for the Algorithm 2.1.

4.1 Main theorem for complexity analysis

This subsection proves some main theorems for the complexity analysis. Now, we prove the key inequality for the complexity analysis of the Algorithm 2.1.

Theorem 4.1 *For given $x^k \in \Re^n$, let $\tilde{x}^k$ be generated by* (2.7)*. If he new iterate x^{k+1} is updated by* (4.2) *with any $\gamma \in (0,2)$, then we have*

$$\begin{aligned} &\gamma\alpha_k^*\beta\{\theta(x) - \theta(\tilde{x}^k) + (x-\tilde{x}^k)^T(Hx+c)\} \\ &\quad \geq \quad \frac{1}{2}\big(\|x-x^{k+1}\|_G^2 - \|x-x^k\|_G^2\big) + \frac{1}{2}q_k(\gamma), \ \ \forall\, x \in \Omega, \end{aligned} \tag{4.6}$$

where

$$q_k(\gamma) = \gamma(2-\gamma)\alpha_k^*\|x^k - \tilde{x}^k\|^2. \tag{4.7}$$

Proof. Since $\tilde{x}^k$ is the solution of (2.7), we have (see (2.8))

$$\beta\theta(x) - \beta\theta(\tilde{x}^k) + (x-\tilde{x}^k)^T\{\beta(Hx^k+c) - (x^k - \tilde{x}^k)\} \geq 0, \ \forall x \in \mathcal{X}.$$

Thus, we have

$$\beta\theta(x)-\beta\theta(\tilde{x}^k)+(x-\tilde{x}^k)^T\beta(Hx^k+c)\ge(x-\tilde{x}^k)^T(x^k-\tilde{x}^k),\ \forall x\in\mathcal{X}.$$

Adding the term $(x-\tilde{x}^k)^T\beta H(x^k-\tilde{x}^k)$ to the both sides of the above inequality and using $(I+\beta H)=G$, we obtain

$$\begin{aligned}&\beta\theta(x)-\beta\theta(\tilde{x}^k)+(x-\tilde{x}^k)^T\{\beta(Hx^k+c)+\beta H(x^k-\tilde{x}^k)\}\\ \ge\quad&(x-\tilde{x}^k)^TG(x^k-\tilde{x}^k),\quad\forall\, x\in\mathcal{X}.\end{aligned}$$

By using the identity

$$\beta(Hx^k+c)+\beta H(x^k-\tilde{x}^k)=\beta(Hx+c)+\beta H(\tilde{x}^k-x)+2\beta H(x^k-\tilde{x}^k)$$

Then, we rewrite the above inequality in our desirable form

$$\begin{aligned}&\beta\theta(x)-\beta\theta(\tilde{x}^k)+(x-\tilde{x}^k)^T\{\beta(Hx+c)+2\beta H(x^k-\tilde{x}^k)\}\\ \ge\quad&(x-\tilde{x}^k)^TG(x^k-\tilde{x}^k)+\beta\|x-\tilde{x}^k\|_H^2,\ \forall x\in\mathcal{X}.\end{aligned}$$

By using the Cauchy-Schwarz inequality, from the above inequality we obtain

$$\begin{aligned}&\beta\theta(x) - \beta\theta(\tilde{x}^k) + (x - \tilde{x}^k)^T\beta(Hx + c)\\ &\quad\geq \quad (x - \tilde{x}^k)^T G(x^k - \tilde{x}^k) + \beta\|x - \tilde{x}^k\|_H - 2(x - \tilde{x}^k)^T\beta H(x^k - \tilde{x}^k)\\ &\quad\geq \quad (x - \tilde{x}^k)^T G(x^k - \tilde{x}^k) - \beta\|x^k - \tilde{x}^k\|_H^2, \quad \forall\, x \in \mathcal{X},\end{aligned}$$

and thus

$$\begin{aligned}&\gamma\alpha_k^*\beta\{\theta(x) - \theta(\tilde{x}^k) + (x - \tilde{x}^k)^T(Hx + c)\}\\ &\quad\geq \quad (x - \tilde{x}^k)^T G\gamma\alpha_k^*(x^k - \tilde{x}^k) - \gamma\alpha_k^*\beta\|x^k - \tilde{x}^k\|_H^2, \quad \forall\, x \in \mathcal{X}.\end{aligned}$$

Because $\gamma\alpha_k^*(x^k - \tilde{x}^k) = (x^k - x^{k+1})$ (see (4.2)), thus we have

$$\begin{aligned}&\gamma\alpha_k^*\beta\{\theta(x) - \theta(\tilde{x}^k) + (x - \tilde{x}^k)^T(Hx + c)\}\\ &\quad\geq \quad (x - \tilde{x}^k)^T G(x^k - x^{k+1}) - \gamma\alpha_k^*\beta\|x^k - \tilde{x}^k\|_H^2, \ \forall\, x \in \mathcal{X}. \quad (4.8)\end{aligned}$$

To the crossed term in the right hand side of (4.8), $(x - \tilde{x}^k)^T G(x^k - x^{k+1})$, using an identity

$$(a-b)^T G(c-d) = \frac{1}{2}\left(\|a-d\|_G^2 - \|a-c\|_G^2\right) + \frac{1}{2}\left(\|c-b\|_G^2 - \|d-b\|_G^2\right),$$

we obtain

$$\begin{aligned}
&(x-\tilde{x}^k)^T G(x^k - x^{k+1}) \\
&= \frac{1}{2}\left(\|x-x^{k+1}\|_G^2 - \|x-x^k\|_G^2\right) + \frac{1}{2}\left(\|x^k-\tilde{x}^k\|_G^2 - \|x^{k+1}-\tilde{x}^k\|_G^2\right).
\end{aligned} \tag{4.9}$$

Using $x^{k+1} = x^k - \gamma\alpha_k^*(x^k - \tilde{x}^k)$ to the last part of the right hand side of (4.9), we get

$$\begin{aligned}
&\|x^k-\tilde{x}^k\|_G^2 - \|x^{k+1}-\tilde{x}^k\|_G^2 \\
&= \|x^k-\tilde{x}^k\|_G^2 - \|(x^k-\tilde{x}^k) - \gamma\alpha_k^*(x^k-\tilde{x}^k)\|_G^2 \\
&= 2\gamma\alpha_k^*\|x^k-\tilde{x}^k\|_G^2 - \gamma^2\alpha_k^*(\alpha_k^*\|x^k-\tilde{x}^k\|_G^2) \qquad (\text{see (4.3)}) \\
&= 2\gamma\alpha_k^*\|x^k-\tilde{x}^k\|^2 + 2\gamma\alpha_k^*\beta\|x^k-\tilde{x}^k\|_H^2 - \gamma^2\alpha_k^*\|x^k-\tilde{x}^k\|^2 \\
&= \gamma(2-\gamma)\alpha_k^*\|x^k-\tilde{x}^k\|^2 + 2\gamma\alpha_k^*\|x^k-\tilde{x}^k\|_H^2.
\end{aligned}$$

Substituting it in the right hand side of (4.9) and using the definition of $q_k(\gamma)$, we obtain

$$\begin{aligned}(x - \tilde{x}^k)^T G(x^k - x^{k+1}) \\ = \quad \frac{1}{2}\big(\|x - x^{k+1}\|_G^2 - \|x - x^k\|_G^2\big) + \frac{1}{2} q_k(\gamma) + \gamma\alpha_k^*\beta\|x^k - \tilde{x}^k\|_H^2.\end{aligned} \tag{4.10}$$

Adding (4.8) and (4.10), the theorem is proved. $\square$

By setting $x = x^*$ in (4.6), we get

$$\begin{aligned}\|x^k - x^*\|_G^2 - \|x^{k+1} - x^*\|_G^2 \\ \geq \quad 2\gamma\alpha_k^*\beta\{(\theta(\tilde{x}^k) - \theta(x^*) + (\tilde{x}^k - x^*)^T(Hx * +c)\} + q_k(\gamma).\end{aligned}$$

Because $\theta(\tilde{x}^k) - \theta(x^*) + (\tilde{x}^k - x^*)^T(Hx^* + c) \geq 0$, it follows from the last inequality and (4.7) that

$$\|x^{k+1} - x^*\|_G^2 \leq \|x^k - x^*\|_G^2 - \gamma(2-\gamma)\alpha_k^*\|x^k - \tilde{x}^k\|^2.$$

Thus, the contraction property (4.4) is a byproduct of Theorem 4.1.

4.2 Convergence rate of the Algorithm 2.1.

This section uses Theorem 4.1 to show the convergence rate of the Algorithm 2.1.

Theorem 4.2 *Let the sequence* $\{x^k\}$ *be generated by the Algorithm 2.1. Then for any integer* $t > 0$*, it holds that*

$$\theta(\tilde{x}_t) - \theta(x) + (\tilde{x}_t - x)^T(Hx + c) \le \frac{\|I + \beta H\|}{2\gamma\beta(t+1)}\|x - x^0\|_G^2, \ \forall x \in \mathcal{X}, \tag{4.11}$$

where

$$\tilde{x}_t = \frac{1}{\Upsilon_t}\sum_{k=0}^{t}\alpha_k^*\tilde{x}^k \qquad \text{and} \qquad \Upsilon_t = \sum_{k=0}^{t}\alpha_k^*. \tag{4.12}$$

Proof. For the convergence rate proof, we allow $\gamma \in (0, 2]$. In this case, we still have $q_k(\gamma) \ge 0$. From (4.6) we get

$$\begin{aligned} &\alpha_k^*(\theta(x) - \theta(\tilde{x}^k)) + \alpha_k^*(x - \tilde{x}^k)^T(Hx + c) \\ \ge \quad & \frac{1}{2\gamma\beta}\|x - x^{k+1}\|_G^2 - \frac{1}{2\gamma\beta}\|x - x^k\|_G^2, \ \ \forall\, x \in \mathcal{X}. \end{aligned}$$

Summing the above inequality over $k = 0, \ldots, t$, we obtain

$$\Big((\sum_{k=0}^{t} \alpha_k^*) \theta(x) - \sum_{k=0}^{t} \alpha_k^* \theta(\tilde{x}^k) \Big) + \Big((\sum_{k=0}^{t} \alpha_k^*) x - \sum_{k=0}^{t} \alpha_k^* \tilde{x}^k \Big)^T (Hx + c)$$
$$\geq \quad -\frac{1}{2\gamma\beta} \|x - x^0\|^2, \quad \forall x \in \mathcal{X}.$$

Using the notations of Υ_t and $\tilde{x}_t$ in the above inequality, we derive

$$\Big(\frac{1}{\Upsilon_t} (\sum_{k=0}^{t} \alpha_k^* \theta(\tilde{x}^k)) - \theta(x) \Big) + (\tilde{x}_t - x)^T (Hx + c) \leq \frac{\|x - x^0\|^2}{2\gamma\beta\Upsilon_t}, \quad \forall x \in \mathcal{X}. \tag{4.13}$$

Indeed, $\tilde{x}_t \in \mathcal{X}$ because it is a convex combination of $\tilde{x}^0, \tilde{x}^1, \ldots, \tilde{x}^t$. Since $\theta(x)$ is convex and

$$\tilde{x}_t = \frac{1}{\Upsilon_t} \sum_{k=0}^{t} \alpha_k^* \tilde{x}^k \quad \text{and} \quad \Upsilon_t = \sum_{k=0}^{t} \alpha_k^*.$$

we have $\theta(\tilde{x}_t) \le \frac{1}{\Upsilon_t}\big(\sum_{k=0}^{t} \alpha_k^* \theta(\tilde{x}^k)\big)$. Thus, it follows from (4.13) that

$$\theta(\tilde{x}_t) - \theta(x) + (\tilde{x}_t - x)^T (Hx + c) \le \frac{\|x - x^0\|^2}{2\gamma\beta\Upsilon_t}, \quad \forall x \in \mathcal{X}. \tag{4.14}$$

Because $\alpha_k^* \ge 1/\|I + \beta H\|$ for all $k > 0$ (see (4.3)), it follows from (4.12) that

$$\Upsilon_t \ge \frac{t+1}{\|I + \beta H\|}.$$

Substituting it in (4.14), the proof is complete. $\square$

Thus the Algorithm 2.1. has $O(1/t)$ convergence rate. For any substantial set $\mathcal{D} \subset \mathcal{X}$, it reaches

$$\theta(\tilde{x}_t) - \theta(x) + (\tilde{x}_t - x)^T (Hx + c) \le \epsilon, \quad \forall\, x \in \mathcal{D}(\tilde{x}_t),$$

in at most

$$t = \left\lceil \frac{\|I + \beta H\| d^2}{2\gamma\beta\epsilon} \right\rceil \text{ iterations,}$$

where $\tilde{x}_t$ is defined in (4.12) and $d = \sup\{\|x - x^0\| \,|\, x \in \mathcal{D}(\tilde{x}_t)\}$. This convergence rate is in the ergodic sense, the statement (4.11) suggests us to take a larger parameter $\gamma \in (0, 2]$ in the correction steps of the the Algorithm 2.1.

References

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

[2] C. H. Chen, X. L. Fu, B.S. He and X. M. Yuan, On the Iteration Complexity of Some Projection Methods for Monotone Linear Variational Inequalities, JOTA, **172**, 914-928, 2017.

[3] D. Chen and Y. Zhang, A Hybrid Multi-Objective Scheme Applied to Redundant Robot Manipulators, IEEE Transactions on Automation Science and Engineering, **14**, 1337–1350, 2017.

[4] F. Facchinei and J. S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity problems, Volume I*, Springer Series in Operations Research, Springer-Verlag, 2003.

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

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

[7] B.S. He, A class of projection and contraction methods for monotone variational inequalities, *Appl. Math. & Optimi.*, **35**, 69–76, 1997.

[8] B.S. He and X.M. Yuan, On the $O(1/n)$ Convergence Rate of the Douglas-Rachford Alternating Direction Method， SIAM J. Numer. Anal. **50**, 700-709, 2012.

[9] A. Nemirovski, *Prox-method with rate of convergence* $O(1/t)$ *for variational inequality with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems*, SIAM J. Optim. 15 (2005), pp. 229-251.

[10] P. Tseng, *On accelerated proximal gradient methods for convex-concave optimization*, manuscript, University of Washington, USA, 2008.

[11] Yunong Zhang, Senbo Fu, Zhijun Zhang, and Lin Xiao, On the LVI-Based Numerical Method (E47 Algorithm) for Solving Quadratic Programming Problems, IEEE International Conference on Automation and Logistics, 2011.

[12] Yunong Zhang, Long Jin, Robot Manipulator Redundancy Resolution, Hoboken, Wiley, 2017

[13] 何炳生, 论求解单调变分不等式的一些投影收缩算法, 计算数学, **18**, 54-60, 1996.