Improved relaxed CQ methods for solving the split feasibility problem

LI Min
(School of Economics and Management, Southeast University, Nanjing 210096)

Abstract: This paper presents some improved relaxed CQ methods to solve the split feasibility problem. These new methods, which are based on the modified relaxed CQ algorithm, generate the new iterate by searching the optimal step size along the descent direction. Global convergence of these new methods is proved under mild assumptions. Preliminary numerical results verify the computational preferences of the new methods.

Key words: CQ algorithm; split feasibility problem; step length

0 Introduction

Let C and Q be nonempty closed convex sets in R^N and R^M, respectively, and A an $M \times N$ real matrix. The problem,

$$\text{to find } x \in C \text{ with } Ax \in Q, \qquad \text{if such } x \text{ exists,} \tag{1}$$

was called the split feasibility problem (SFP) by Censor and Elfving [1]. The SFP (1) is equivalent to the following variational inequality (see Section 3 in [2])

$$x^* \in C, \quad \langle F(x^*), x - x^* \rangle \geq 0, \qquad \forall x \in C \tag{2}$$

where

$$F(x) = A^T (I - P_Q) Ax, \tag{3}$$

I and P_Q denote the identity operator and the orthogonal projections onto Q, respectively. In this paper, we always assume that the solution set of (1), denoted by C^*, is always nonempty.

To solve the SFP (1), Byrne [3] proposed the CQ algorithm, which generates the new iterate as follows

$$x^{k+1} = P_C[x^k - \gamma F(x^k)], \tag{4}$$

where $\gamma \in (0, 2/L)$, L denotes the largest eigenvalue of the matrix $A^T A$. However, sometimes the projections onto C and Q are difficult to calculate. If this case appears, the efficiency of the CQ algorithm, will be seriously affected. In [4], Yang presented a relaxed CQ algorithm for solving the SFP, where at k-th iteration, the projections onto C and Q were replaced with the halfspaces C_k and Q_k, respectively.

Note that the step length of the CQ algorithm and the relaxed version relies on the largest eigenvalue of the matrix $A^T A$. In [2], Qu and Xiu proposed a modified relaxed CQ algorithm

$$\tilde{x}^k = P_{C_k}[x^k - \alpha_k F_k(x^k)], \tag{5}$$

where

$$F_k(x^k) = A^T (I - P_{Q_k}) Ax^k, \quad \alpha_k \left\| F(x^k) - F(\tilde{x}^k) \right\| \leq \mu \left\| x^k - \tilde{x}^k \right\|, \quad 0 < \mu < 1, \tag{6}$$

and the new iterate x^{k+1} is updated by

Foundations: Grant from Southeast University (No. KJ2009370), Specialized Research Fund for the Doctoral Program of Higher Education Grant (No. 200802861031)
Brief author introduction:LI Min (1980-), Female, Associate Professor, Research interests: Operations Research; Optimization. E-mail: limin@seu.edu.cn

$$x^{k+1} = P_{C_k}[x^k - \alpha_k F_k(\tilde{x}^k)]. \tag{7}$$

This modified algorithm adopted a self-adaptive strategy in (6), which was in the manner of Armijo's rule, to determine the step length. Thus, the estimation of the largest eigenvalue of the matrix $A^T A$ is avoided.

This paper is to develop some improved relaxed methods for solving the SFP (1) based on the modified relaxed CQ algorithm in [2]. In particular, let x^k be the current iterate of SFP (1) and $x_{\mathrm{I}}^k = \tilde{x}^k$ be generated by (5), then add an optimal step length β_k to $-\alpha_k F_k(x_{\mathrm{I}}^k)$ in (7) to produce x_{II}^k. We may prove that $-(x^k - x_{\mathrm{II}}^k)$ is a descent direction of $\|x - x^*\|^2/2$ at x^k, where $x^* \in C^*$. Hence, two iterative methods are motivated to be presented. The first method sets $x^{k+1} = x_{\mathrm{II}}^k$. The second method produces the new iterate x^{k+1} by

$$x^{k+1} = P_{C_k}[x^k - \rho_k(x^k - x_{\mathrm{II}}^k)]$$

where ρ_k is the optimal step length along the direction $-(x^k - x_{\mathrm{II}}^k)$. Global convergence of the new methods is proved under the same mild assumptions as in [2].

The rest of this paper is organized as follows. In Section 1, we summarize some preliminaries. In Section 2, some improved relaxed CQ methods are presented, followed by some remarks. Then some contractive properties of the new methods are first analyzed. In particular, the strategy of determining the optimal step length of the new methods is investigated. Then, in Section 3, the global convergence of the new methods is proved. In Section 4, we apply the new methods to solve some numerical problems, and compare it with the algorithm in [2]. The numerical results are therefore reported. Finally, some conclusions are made in Section 5.

1 Preliminaries

First, we summarize some basic properties related to variational inequalities. Let Ω denote the given nonempty closed convex set in R^n and $P_\Omega(x)$ the projection of x onto Ω, that is,

$$P_\Omega(x) = \operatorname{Arg\,min}\{\|x - y\| \, y \in \Omega\}$$

From the above definition, it follows that

$$\langle P_\Omega(x) - x, z - P_\Omega(x)\rangle \ge 0, \quad \forall x \in R^n, \ \forall z \in \Omega. \tag{8}$$

Consequently, we have

$$\langle (I - P_\Omega)x - (I - P_\Omega)y, x - y\rangle \ge \|(I - P_\Omega)x - (I - P_\Omega)y\|^2, \quad \forall x, y \in R^n \tag{9}$$

and

$$\|P_\Omega(x) - z\|^2 \le \|x - z\|^2 - \|P_\Omega(x) - x\|^2, \quad \forall x \in R^n, \ \forall z \in \Omega. \tag{10}$$

Let F be a mapping from R^n into R^n. For any $x \in R^n$ and $\alpha > 0$, define

$$x(\alpha) = P_\Omega[x - \alpha F(x)], \qquad e(x,\alpha) = x - x(\alpha). \tag{11}$$

Note that $e(x,\alpha)$ is a continuous function of x because the projection mapping is non-expansive. The next lemma states a useful property of $\|e(x,\alpha)\|$.

Lemma 1: ([2] Lemma 2.2) Let F be a mapping from R^n into R^n. For any $x \in R^n$ and

$\alpha > 0$, we have

$$\min\{1,\alpha\}\|e(x,1)\| \le \|e(x,\alpha)\| \le \max\{1,\alpha\}\|e(x,1)\| \tag{12}$$

In this paper, we assume that the projections P_C and P_Q are not easily calculated. Carefully speaking, the convex sets C and Q satisfy the following assumptions:

(H1) The set C is given by

$$C = \left\{ x \in R^N \,\middle|\, c(x) \le 0 \right\},$$

where $c: R^N \to R$ is a convex (not necessarily differentiable) function and C is nonempty. The set Q is given by

$$Q = \left\{ y \in R^M \,\middle|\, q(y) \le 0 \right\},$$

where $q: R^M \to R$ is a convex (not necessarily differentiable) function and Q is nonempty.

(H2) For any $x \in R^N$, at least one subgradient $\xi \in \partial c(x)$ can be calculated, where $\partial c(x)$ is a generalized gradient of $c(x)$ at x and is defined as follows:

$$\partial c(x) = \left\{ \xi \in R^N \,\middle|\, c(z) \ge c(x) + \langle \xi, z - x \rangle, \quad \text{for all} \quad z \in R^N \right\}.$$

For any $y \in R^M$, at least one subgradient $\eta \in \partial q(y)$ can be calculated, where

$$\partial q(y) = \left\{ \eta \in R^M \,\middle|\, q(u) \ge q(y) + \langle \eta, u - y \rangle, \quad \text{for all} \quad u \in R^M \right\}$$

The following lemma provides an important boundedness property of the subdifferential, see, e.g., [5].

Lemma 2: Suppose $h: R^n \to R$ is a convex function, then it is subdifferentiable everywhere and its subdifferentials are uniformly bounded on any bounded subset of R^n.

Denote

$$C_k = \left\{ x \in R^N \,\middle|\, c(x^k) + \langle \xi^k, x - x^k \rangle \le 0 \right\},$$

where ξ^k is an element in $\partial c(x^k)$, and

$$Q_k = \left\{ y \in R^M \,\middle|\, q(Ax^k) + \langle \eta^k, y - Ax^k \rangle \le 0 \right\},$$

where η^k is an element in $\partial q(Ax^k)$.

Remark 1: By the definition of subgradient, it is clear that the halfspaces C_k and Q_k contain C and Q, respectively. From the expressions of C_k and Q_k, the orthogonal projections onto C_k and Q_k may be directly calculated and then we have the following proposition (see [6,7]).

Proposition 1: For any $z \in R^N$.

$$P_{C_k}(z) = \begin{cases} z - \dfrac{c(x^k) + \langle \xi^k, z - x^k \rangle}{\|\xi^k\|^2}\xi^k, & \text{if } c(x^k) + \langle \xi^k, z - x^k \rangle > 0; \\ z, & \text{otherwise,} \end{cases}$$

and

$$P_{Q_k}(Az)=\begin{cases} Az-\dfrac{q(Ax^k)+\langle \eta^k, Az-Ax^k\rangle}{\|\eta^k\|^2}\eta^k, & \text{if } q(Ax^k)+\langle \eta^k, Az-Ax^k\rangle>0; \\ Az, & \text{otherwise,}\end{cases}$$

For every k, using Q_k we define the function $F_k : R^N \to R^N$ by

$$F_k(x)=A^T(I-P_{Q_k})Ax.$$

Although the function F_k depends on k, it has nice properties as shown in the following lemma.

Lemma 3: ([2], Lemma 4.2) For all $k=0,1,2,\cdots$, F_k is Lipschitz continuous on R^N with constant L and co-coercive on R^N with modulus $1/L$, where L is the largest eigenvalue of the matrix A^TA.

2 Improved relaxed CQ methods

In this section, we will propose two improved relaxed CQ methods and show how to determine the optimal step length. The detailed procedures of the new methods are presented as below:

Algorithm 1. Initialization: Choose $\mu\in(0,1)$, $\varepsilon>0$, $x^0\in R^N$ and $k=0$.

Step 1. Prediction: Choose an $\alpha_k>0$, such that

$$x_{\mathrm{I}}^k=P_{C_k}[x^k-\alpha_k F_k(x^k)], \tag{13}$$

and

$$\alpha_k\left\|F_k(x^k)-F_k(x_{\mathrm{I}}^k)\right\|\le\mu\left\|x^k-x_{\mathrm{I}}^k\right\|. \tag{14}$$

Step 2. Stopping Criterion: Compute

$$e_k(x^k,\alpha_k)=x^k-x_{\mathrm{I}}^k$$

If $\|e_k(x^k,\alpha_k)\|\le\varepsilon$, terminate the iteration with the approximate solution x^k. Otherwise, go to Step 3.

Step 3. Correction: The new iterate x^{k+1} is updated by

$$x^{k+1}=x_{\mathrm{II}}^k=P_{C_k}[x^k-\beta_k\alpha_k F_k(x_{\mathrm{I}}^k)] \tag{15}$$

where

$$\beta_k=\delta_k\beta_k^*,\qquad \beta_k^*=\frac{\langle x^k-x_{\mathrm{I}}^k, d_k(x^k,x_{\mathrm{I}}^k,\alpha_k)\rangle}{\|d_k(x^k,x_{\mathrm{I}}^k,\alpha_k)\|^2},\qquad \delta_k\in[\delta_L,\delta_U], \tag{16}$$

and

$$d_k(x^k,x_{\mathrm{I}}^k,\alpha_k)=x^k-x_{\mathrm{I}}^k-\alpha_k[F_k(x^k)-F_k(x_{\mathrm{I}}^k)]. \tag{17}$$

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

Algorithm 2. Initialization: Choose $\mu\in(0,1)$, $\varepsilon>0$, $x^0\in R^N$ and $k=0$.

Step 1. Prediction: Choose an $\alpha_k>0$, such that

$$x_{\mathrm{I}}^k=P_{C_k}[x^k-\alpha_k F_k(x^k)], \tag{18}$$

and

$$\alpha_k \left\| F_k(x^k) - F_k(x_{\mathrm{I}}^k) \right\| \le \mu \left\| x^k - x_{\mathrm{I}}^k \right\|. \tag{19}$$

Step 2. Stopping Criterion: Compute

$$e_k(x^k, \alpha_k) = x^k - x_{\mathrm{I}}^k .$$

If $\left\| e_k(x^k, \alpha_k) \right\| \le \varepsilon$, terminate the iteration with the approximate solution x^k. Otherwise, go to Step 3.

Step 3. Correction: The corrector x_{II}^k is updated by

$$x_{\mathrm{II}}^k = P_{C_k}[x^k - \beta_k \alpha_k F_k(x_{\mathrm{I}}^k)] \tag{20}$$

where

$$\beta_k = \delta_k \beta_k^*, \qquad \beta_k^* = \frac{\left\langle x^k - x_{\mathrm{I}}^k, d_k(x^k, x_{\mathrm{I}}^k, \alpha_k) \right\rangle}{\left\| d_k(x^k, x_{\mathrm{I}}^k, \alpha_k) \right\|^2}, \qquad \delta_k \in [\delta_L, \delta_U] \subseteq (0,2), \tag{21}$$

and

$$d_k(x^k, x_{\mathrm{I}}^k, \alpha_k) = x^k - x_{\mathrm{I}}^k - \alpha_k [F_k(x^k) - F_k(x_{\mathrm{I}}^k)]. \tag{22}$$

Step 4. Extension: The new iterate x^{k+1} is updated by

$$x^{k+1} = P_{C_k}[x^k - \rho_k (x^k - x_{\mathrm{II}}^k)] \tag{23}$$

where

$$\rho_k = \gamma_k \rho_k^*, \quad \rho_k^* = \frac{\left\| x^k - x_{\mathrm{II}}^k \right\|^2 + \beta_k \alpha_k \left\langle x_{\mathrm{II}}^k - x_{\mathrm{I}}^k, F_k(x_{\mathrm{I}}^k) \right\rangle}{\left\| x^k - x_{\mathrm{II}}^k \right\|^2}, \quad \gamma_k \in [\gamma_L, \gamma_U] \subseteq (0,2). \tag{24}$$

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

Remark 2: In the prediction step, if the selected α_k satisfies $0 < \alpha_k < \mu / L$ (L is the largest eigenvalue of the matrix $A^T A$), then from Lemma 3, we have

$$\alpha_k \left\| F_k(x^k) - F_k(x_{\mathrm{I}}^k) \right\| \le \alpha_k L \left\| x^k - x_{\mathrm{I}}^k \right\| \le \mu \left\| x^k - x_{\mathrm{I}}^k \right\| \tag{25}$$

and thus Condition (14) or (19) is satisfied. Without loss of generality, we can assume that $\inf\{\alpha_k\} = \alpha_{\min} > 0$. Since we do not know the value of $L > 0$ but it exists, in practice, a self-adaptive scheme is adopted to find such a suitable $\alpha_k > 0$. For given x^k and a trial $\alpha_k > 0$, along with the value of $F_k(x^k)$, we set the trial x_{I}^k as follows:

$$x_{\mathrm{I}}^k = P_{C_k}[x^k - \alpha_k F_k(x^k)].$$

Then calculate

$$r_k := \frac{\alpha_k \left\| F_k(x^k) - F_k(x_{\mathrm{I}}^k) \right\|}{\left\| x^k - x_{\mathrm{I}}^k \right\|},$$

if $r_k \le \mu$, the trial x_{I}^k is accepted as predictor; Otherwise, reduce α_k by $\alpha_k := 0.9 \mu \alpha_k \min(1, 1/r_k)$ to get a new smaller trial α_k and repeat this procedure. In the case that the predictor has been accepted, a good initial trial α_{k+1} for next iteration is prepared by the following strategy:

$$\alpha_{k+1}=\begin{cases}\dfrac{0.9\mu}{r_k}\alpha_k & \text{if } r_k\le\nu,\\ \alpha_k & \text{otherwise},\end{cases}\qquad (\text{usually } \nu\in[0.4,0.5]). \tag{26}$$

Condition (14) or (19) ensures that $\alpha_k\|F_k(x^k)-F_k(x_{\mathrm{I}}^k)\|$ is smaller than $\|x^k-x_{\mathrm{I}}^k\|$, however, too small $\alpha_k\|F_k(x^k)-F_k(x_{\mathrm{I}}^k)\|$ leads to slow convergence. The proposed adjusting strategy (26) is intended to avoid such a case as indicated in [8,9]. Actually, it is very important to balance the quantity of $\alpha_k\|F_k(x^k)-F_k(x_{\mathrm{I}}^k)\|$ and $\|x^k-x_{\mathrm{I}}^k\|$ in practical computation. Note that there are at least two times to utilize the value of function in the prediction step: one is $F_k(x^k)$, and the other is $F_k(x_{\mathrm{I}}^k)$ for testing whether the Condition (14) or (19) holds. When α_k is selected well enough, x_{I}^k will be accepted after only one trial and in this case, the prediction step exactly utilizing the value of concerned function twice in one iteration.

Remark 3: As x_{I}^k (and resulted $F_k(x_{\mathrm{I}}^k)$) is determined by x^k and α_k, the vector $d_k(x^k,x_{\mathrm{I}}^k,\alpha_k)=x^k-x_{\mathrm{I}}^k-\alpha_k[F_k(x^k)-F_k(x_{\mathrm{I}}^k)]$ in (22) is a function of x^k and α_k at all. In addition, the correction step does not require any new function evaluations.

Remark 4: In the extension step, we only use the function value $F_k(x_{\mathrm{I}}^k)$ which is obtained in the prediction step. Therefore, the extension step also does not require any new function evaluations.

For analysis, we consider the following general forms of correction step and extension step:

$$x_{\mathrm{II}}^k(\beta)=P_{C_k}[x^k-\beta\alpha_kF_k(x_{\mathrm{I}}^k)] \quad\text{and}\quad x^{k+1}(\rho)=P_{C_k}[x^k-\rho(x^k-x_{\mathrm{II}}^k)]. \tag{27}$$

Lemma 4: Given x^k, $x^*\in C^*$ and $\alpha_k>0$, let $x_{\mathrm{I}}^k\in C_k$ be the predictor and $x_{\mathrm{II}}^k(\beta)$ be given by the general form of the corrector. Then for any $\beta>0$ we have

$$\Theta_k(\beta)=\|x^k-x^*\|^2-\|x_{\mathrm{II}}^k(\beta)-x^*\|^2\ge\Phi_k(\beta)\ge Q_k(\beta), \tag{28}$$

where

$$\Phi_k(\beta)=\|x^k-x_{\mathrm{II}}^k(\beta)\|^2+2\beta\alpha_k\left\langle x_{\mathrm{II}}^k(\beta)-x_{\mathrm{I}}^k,F_k(x_{\mathrm{I}}^k)\right\rangle \tag{29}$$

and

$$Q_k(\beta)=2\beta\left\langle x^k-x_{\mathrm{I}}^k,d_k(x^k,x_{\mathrm{I}}^k,\alpha_k)\right\rangle-\beta^2\|d_k(x^k,x_{\mathrm{I}}^k,\alpha_k)\|^2. \tag{30}$$

Proof: Since $x^*\in C\subseteq C_k$ and $x_{\mathrm{II}}^k(\beta)=P_{C_k}[x^k-\beta\alpha_kF_k(x_{\mathrm{I}}^k)]$, it follows from (10) that

$$\|x_{\mathrm{II}}^k(\beta)-x^*\|^2\le\|x^k-\beta\alpha_kF_k(x_{\mathrm{I}}^k)-x^*\|^2-\|x^k-\beta\alpha_kF_k(x_{\mathrm{I}}^k)-x_{\mathrm{II}}^k(\beta)\|^2. \tag{31}$$

Consequently, using the definition of $\Theta_k(\beta)$, we get

$$\begin{aligned}\Theta_k(\beta)&\ge\|x^k-x^*\|^2+\|x^k-x_{\mathrm{II}}^k(\beta)-\beta\alpha_kF_k(x_{\mathrm{I}}^k)\|^2-\|x^k-x^*-\beta\alpha_kF_k(x_{\mathrm{I}}^k)\|^2\\&=\|x^k-x_{\mathrm{II}}^k(\beta)\|^2+2\beta\alpha_k\left\langle x_{\mathrm{II}}^k(\beta)-x^*,F_k(x_{\mathrm{I}}^k)\right\rangle.\end{aligned} \tag{32}$$

It follows from $Ax^*\in Q\subseteq Q_k$ that

$$F_k(x^*)=0.$$

Since $x_{\mathrm{I}}^k \in C_k$, using the monotonicity of F_k and the above equality, we have

$$\left\langle x_{\mathrm{I}}^k - x^*, F_k(x_{\mathrm{I}}^k)\right\rangle \ge \left\langle x_{\mathrm{I}}^k - x^*, F_k(x^*)\right\rangle \ge 0,$$

and consequently

$$\left\langle x_{\mathrm{II}}^k(\beta) - x^*, F_k(x_{\mathrm{I}}^k)\right\rangle \ge \left\langle x_{\mathrm{II}}^k(\beta) - x_{\mathrm{I}}^k, F_k(x_{\mathrm{I}}^k)\right\rangle. \tag{33}$$

Applying (33) to the last term in the right hand side of (32), we obtain

$$\Theta_k(\beta) \ge \left\| x^k - x_{\mathrm{II}}^k(\beta)\right\|^2 + 2\beta\alpha_k \left\langle x_{\mathrm{II}}^k(\beta) - x_{\mathrm{I}}^k, F_k(x_{\mathrm{I}}^k)\right\rangle. \tag{34}$$

The first assertion follows immediately. Since $x_{\mathrm{I}}^k = P_{C_k}[x^k - \alpha_k F_k(x^k)]$ and $x_{\mathrm{II}}^k(\beta) \in C_k$, it follows from (8) that for any $\beta > 0$,

$$0 \ge 2\beta \left\langle x_{\mathrm{II}}^k(\beta) - x_{\mathrm{I}}^k, [x^k - \alpha_k F_k(x^k)] - x_{\mathrm{I}}^k\right\rangle. \tag{35}$$

Adding (29) and (35) together and using the notation of $d_k(x^k, x_{\mathrm{I}}^k, \alpha_k)$ in (22), we obtain

$$\begin{aligned}\Phi_k(\beta) &\ge \left\| x^k - x_{\mathrm{II}}^k(\beta)\right\|^2 + 2\beta\left\langle x_{\mathrm{II}}^k(\beta) - x_{\mathrm{I}}^k, x^k - x_{\mathrm{I}}^k - \alpha_k[F_k(x^k) - F_k(x_{\mathrm{I}}^k)]\right\rangle \\ &= \left\| x^k - x_{\mathrm{II}}^k(\beta)\right\|^2 + 2\beta\left\langle x_{\mathrm{II}}^k(\beta) - x_{\mathrm{I}}^k, d_k(x^k, x_{\mathrm{I}}^k, \alpha_k)\right\rangle.\end{aligned} \tag{36}$$

Regrouping the first two terms of the right hand side of (36), we get

$$\begin{aligned}&\left\| x^k - x_{\mathrm{II}}^k(\beta)\right\|^2 + 2\beta\left\langle x_{\mathrm{II}}^k(\beta) - x_{\mathrm{I}}^k, d_k(x^k, x_{\mathrm{I}}^k, \alpha_k)\right\rangle \\ &= \left\| x^k - x_{\mathrm{II}}^k(\beta)\right\|^2 + 2\beta\left\langle (x_{\mathrm{II}}^k(\beta) - x^k) + (x^k - x_{\mathrm{I}}^k), d_k(x^k, x_{\mathrm{I}}^k, \alpha_k)\right\rangle \\ &= \left\| x^k - x_{\mathrm{II}}^k(\beta)\right\|^2 + 2\beta\left\langle x_{\mathrm{II}}^k(\beta) - x^k, d_k(x^k, x_{\mathrm{I}}^k, \alpha_k)\right\rangle + 2\beta\left\langle x^k - x_{\mathrm{I}}^k, d_k(x^k, x_{\mathrm{I}}^k, \alpha_k)\right\rangle \\ &= \left\| x^k - x_{\mathrm{II}}^k(\beta) - \beta d_k(x^k, x_{\mathrm{I}}^k, \alpha_k)\right\|^2 + 2\beta\left\langle x^k - x_{\mathrm{I}}^k, d_k(x^k, x_{\mathrm{I}}^k, \alpha_k)\right\rangle - \beta^2\left\| d_k(x^k, x_{\mathrm{I}}^k, \alpha_k)\right\|^2.\end{aligned}$$

Substituting this into (36), we obtain

$$\Phi_k(\beta) \ge 2\beta\left\langle x^k - x_{\mathrm{I}}^k, d_k(x^k, x_{\mathrm{I}}^k, \alpha_k)\right\rangle - \beta^2\left\| d_k(x^k, x_{\mathrm{I}}^k, \alpha_k)\right\|^2$$

and the second assertion is proved. □

Note that $Q_k(\beta)$ is a quadratic function of β and it reaches its maximum at

$$\beta_k^* = \frac{\left\langle x^k - x_{\mathrm{I}}^k, d_k(x^k, x_{\mathrm{I}}^k, \alpha_k)\right\rangle}{\left\| d_k(x^k, x_{\mathrm{I}}^k, \alpha_k)\right\|^2}, \tag{37}$$

with

$$Q_k(\beta_k^*) = \beta_k^* \left\langle x^k - x_{\mathrm{I}}^k, d_k(x^k, x_{\mathrm{I}}^k, \alpha_k)\right\rangle. \tag{38}$$

We set the step length β_k by $\beta_k = \delta_k \beta_k^*$, where $\delta_k \in [\delta_L, \delta_U] \subseteq (0,2)$ is a relaxation factor.

Lemma 5: The step length β_k in the prediction step satisfies:

$$Q_k(\beta_k) \ge \frac{\delta_L(2-\delta_U)(1-\mu)}{2}\left\| x^k - x_{\mathrm{I}}^k\right\|^2, \tag{39}$$

for all $k \ge 0$.

Proof: See (2.15), (5.5) and Theorem 2 in [9]. □

By simple manipulations we obtain

$$
\begin{aligned}
Q_k(\delta_k\beta_k^*) &\overset{(30)}{=} 2\delta_k\beta_k^*\left\langle x^k - x_{\mathrm{I}}^k, d_k(x^k, x_{\mathrm{I}}^k, \alpha_k)\right\rangle - (\delta_k^2\beta_k^*)(\beta_k^*\left\|d_k(x^k, x_{\mathrm{I}}^k, \alpha_k)\right\|^2) \\
&\overset{(37)}{=} (2\delta_k\beta_k^* - \delta_k^2\beta_k^*)\left\langle x^k - x_{\mathrm{I}}^k, d_k(x^k, x_{\mathrm{I}}^k, \alpha_k)\right\rangle \\
&\overset{(38)}{=} \delta_k(2-\delta_k)Q_k(\beta_k^*).
\end{aligned} \tag{40}
$$

Lemma 6: Given x^k and $x^* \in C^*$, let the corrector x_{II}^k be given by (20), then we have

$$
2\left\langle x^k - x^*, x^k - x_{\mathrm{II}}^k\right\rangle \ge \Phi_k(\beta_k) + \left\|x^k - x_{\mathrm{II}}^k\right\|^2. \tag{41}
$$

Proof: Note that

$$
x_{\mathrm{II}}^k - x^* = (x^k - x^*) - (x^k - x_{\mathrm{II}}^k).
$$

Substituting this into (28), we have

$$
2\left\langle x^k - x^*, x^k - x_{\mathrm{II}}^k\right\rangle - \left\|x^k - x_{\mathrm{II}}^k\right\|^2 \ge \Phi_k(\beta_k)
$$

and the proof is complete. □

Remark 5: Since $\Phi_k(\beta_k) \ge Q_k(\beta_k) \ge 0$, $-(x^k - x_{\mathrm{II}}^k)$ is a descent direction of $\left\|x - x^*\right\|^2/2$ at x^k, where x^* is any solution point.

Theorem 1: Given x^k and $x^* \in C^*$, let the corrector x_{II}^k be generated by (20), and the new iterate $x^{k+1}(\rho)$ be given by the general form (27). Then for any $\rho > 0$, we have

$$
\Lambda_k(\rho) = \left\|x^k - x^*\right\|^2 - \left\|x^{k+1}(\rho) - x^*\right\|^2 \ge \Psi_k(\rho), \tag{42}
$$

where

$$
\Psi_k(\rho) = \rho\{\Phi_k(\beta_k) + \left\|x^k - x_{\mathrm{II}}^k\right\|^2\} - \rho^2\left\|x^k - x_{\mathrm{II}}^k\right\|^2, \tag{43}
$$

β_k and $\Phi_k(\beta_k)$ are defined in (21) and (29), respectively.

Proof: Since

$$
\left\|x^k - x^* - \rho(x^k - x_{\mathrm{II}}^k)\right\| \ge \left\|x^{k+1}(\rho) - x^*\right\|, \tag{44}
$$

it follows that

$$
\begin{aligned}
\Lambda_k(\rho) &\ge \left\|x^k - x^*\right\|^2 - \left\|x^k - x^* - \rho(x^k - x_{\mathrm{II}}^k)\right\|^2 \\
&= 2\rho\left\langle x^k - x^*, x^k - x_{\mathrm{II}}^k\right\rangle - \rho^2\left\|x^k - x_{\mathrm{II}}^k\right\|^2.
\end{aligned} \tag{45}
$$

Inequality (42) follows from Lemma 6 and (43) directly and the proof is complete. □

Since $\Psi_k(\rho)$ is a quadratic function of ρ, it reaches its maximum at

$$
\rho_k^* = \frac{\Phi_k(\beta_k) + \left\|x^k - x_{\mathrm{II}}^k\right\|^2}{2\left\|x^k - x_{\mathrm{II}}^k\right\|^2} \overset{(29)}{=} \frac{\left\|x^k - x_{\mathrm{II}}^k\right\|^2 + \beta_k\alpha_k\left\langle x_{\mathrm{II}}^k - x_{\mathrm{I}}^k, F_k(x_{\mathrm{I}}^k)\right\rangle}{\left\|x^k - x_{\mathrm{II}}^k\right\|^2} \tag{46}
$$

with

$$
\Psi_k(\rho_k^*) = \frac{1}{2}\rho_k^*\left\{\Phi_k(\beta_k) + \left\|x^k - x_{\mathrm{II}}^k\right\|^2\right\} \ge \Psi_k(1). \tag{47}
$$

It follows from Lemma 4, Lemma 5 and (46) that

$$\rho_k^* \geq \frac{1}{2} \qquad \text{and} \qquad \Psi_k(\rho_k^*) \geq \frac{1}{4}\left\{\tau_0 \left\|x^k - x_{\mathrm{I}}^k\right\|^2 + \left\|x^k - x_{\mathrm{II}}^k\right\|^2\right\} \tag{48}$$

for some constant $\tau_0 > 0$. For fast convergence, we propose a relaxation factor $\gamma_k \in [\gamma_L, \gamma_U] \subseteq (0,2)$ and set the step length ρ_k by $\rho_k = \gamma_k \rho_k^*$. By simple manipulations we obtain

$$\begin{aligned} \Psi_k(\gamma_k\rho_k^*) &\overset{(43)}{=} \gamma_k\rho_k^*\left\{\Phi_k(\beta_k) + \left\|x^k - x_{\mathrm{II}}^k\right\|^2\right\} - (\gamma_k^2\rho_k^*)(\rho_k^*\left\|x^k - x_{\mathrm{II}}^k\right\|^2) \\ &\overset{(46)}{=} (\gamma_k\rho_k^* - \frac{1}{2}\gamma_k^2\rho_k^*)\left\{\Phi_k(\beta_k) + \left\|x^k - x_{\mathrm{II}}^k\right\|^2\right\} \\ &\overset{(47)}{=} \gamma_k(2-\gamma_k)\Psi_k(\rho_k^*). \end{aligned} \tag{49}$$

It follows from Theorem 1 that

$$\left\|x^{k+1} - x^*\right\|^2 \leq \left\|x^k - x^*\right\|^2 - \frac{\gamma_L(2-\gamma_U)}{4}\left\{\tau_0\left\|x^k - x_{\mathrm{I}}^k\right\|^2 + \left\|x^k - x_{\mathrm{II}}^k\right\|^2\right\}. \tag{50}$$

3 Convergence

It follows from (28) and (39) that for Algorithm 1, there exists a constant $\tau_1 > 0$, such that

$$\left\|x^{k+1} - x^*\right\|^2 \leq \left\|x^k - x^*\right\|^2 - \tau_1\left\|x^k - x_{\mathrm{I}}^k\right\|^2. \tag{51}$$

From (50), we have for Algorithm 2, there exists a constant $\tau_2 > 0$, such that

$$\left\|x^{k+1} - x^*\right\|^2 \leq \left\|x^k - x^*\right\|^2 - \tau_2\left\{\left\|x^k - x_{\mathrm{I}}^k\right\|^2 + \left\|x^k - x_{\mathrm{II}}^k\right\|^2\right\}. \tag{52}$$

The convergence result of the proposed methods in this paper is based on the following theorem.

Theorem 2: Let $\{x^k\}$ be a sequence generated by the proposed methods (Algorithms 1 and 2), α_k be a positive sequence and $\inf\{\alpha_k\} = \alpha_{\min} > 0$. If the solution set of the SFP is nonempty, then $\{x^k\}$ converges to a solution point of the SFP.

Proof: First, from (51) or (52) we get

$$\lim_{k\to\infty}\left\|x^k - x_{\mathrm{I}}^k\right\| = 0. \tag{53}$$

Again, it follows from (51) or (52) that the sequence $\{x^k\}$ is bounded. Let x^∞ be a cluster point of $\{x^k\}$ and the subsequence $\{x^{k_i}\}$ converges to x^∞. We are ready to show that x^∞ is a solution point of the SFP.

First, we show that $x^\infty \in C$. Since $x_{\mathrm{I}}^k \in C_{k_i}$, then by the definition of C_{k_i}, we have

$$c(x^{k_i}) + \left\langle \xi^{k_i}, x_{\mathrm{I}}^{k_i} - x^{k_i}\right\rangle \leq 0, \qquad \forall i = 1,2,\cdots.$$

Passing onto the limit in this inequality and taking into account (53) and Lemma 1, we obtain that

$$c(x^\infty) \leq 0.$$

Hence, we conclude $x^\infty \in C$.

Next, we need to show $Ax^\infty \in Q$. Note that

$$e_k(x,\alpha) = x - P_{C_k}[x - \alpha F_k(x)], \qquad k = 0,1,2,\cdots.$$

Then from Lemma 1, Remark 2 and (53), we have

$$\begin{aligned}\lim_{k_i\to\infty}\left\|e_{k_i}(x_{\mathrm{I}}^{k_i},1)\right\| &\le \lim_{k_i\to\infty}\frac{\left\|x^{k_i}-x_{\mathrm{I}}^{k_i}\right\|}{\min\{1,\alpha_{k_i}\}}\\ &\le \lim_{k_i\to\infty}\frac{\left\|x^{k_i}-x_{\mathrm{I}}^{k_i}\right\|}{\min\{1,\alpha_{\min}\}}\\ &=0.\end{aligned} \tag{54}$$

Using (8) and $x^*\in C_{k_i}$, we have for all $\forall i=1,2,\cdots$

$$\left\langle x^{k_i}-F_{k_i}(x^{k_i})-P_{C_{k_i}}[x^{k_i}-F_{k_i}(x^{k_i})],x^*-P_{C_{k_i}}[x^{k_i}-F_{k_i}(x^{k_i})]\right\rangle\le 0,$$

that is,

$$\left\langle e_{k_i}(x_{\mathrm{I}}^{k_i},1)-F_{k_i}(x^{k_i}),x^{k_i}-x^*-e_{k_i}(x_{\mathrm{I}}^{k_i},1)\right\rangle\ge 0. \tag{55}$$

It follows from (9) and $Ax^*\in Q_{k_i}$ that

$$\begin{aligned}&\left\langle F_{k_i}(x^{k_i}),x^{k_i}-x^*\right\rangle\\ &=\left\langle F_{k_i}(x^{k_i})-F_{k_i}(x^*),x^{k_i}-x^*\right\rangle\\ &=\left\langle A^T(I-P_{Q_{k_i}})Ax^{k_i}-A^T(I-P_{Q_{k_i}})Ax^*,x^{k_i}-x^*\right\rangle\\ &=\left\langle (I-P_{Q_{k_i}})Ax^{k_i}-(I-P_{Q_{k_i}})Ax^*,Ax^{k_i}-Ax^*\right\rangle\\ &\ge\left\|(I-P_{Q_{k_i}})Ax^{k_i}-(I-P_{Q_{k_i}})Ax^*\right\|^2\\ &=\left\|(I-P_{Q_{k_i}})Ax^{k_i}\right\|^2.\end{aligned}$$

From (55) and the above inequality we know for all $i=1,2,\cdots$,

$$\begin{aligned}&\left\langle x^{k_i}-x^*,e_{k_i}(x_{\mathrm{I}}^{k_i},1)\right\rangle\\ &\ge\left\|e_{k_i}(x_{\mathrm{I}}^{k_i},1)\right\|^2-\left\langle F_{k_i}(x^{k_i}),e_{k_i}(x_{\mathrm{I}}^{k_i},1)\right\rangle+\left\langle F_{k_i}(x^{k_i}),x^{k_i}-x^*\right\rangle\\ &\ge\left\|e_{k_i}(x_{\mathrm{I}}^{k_i},1)\right\|^2-\left\langle F_{k_i}(x^{k_i}),e_{k_i}(x_{\mathrm{I}}^{k_i},1)\right\rangle+\left\|(I-P_{Q_{k_i}})Ax^{k_i}\right\|^2.\end{aligned} \tag{56}$$

Since

$$\left\|F_{k_i}(x^{k_i})\right\|=\left\|F_{k_i}(x^{k_i})-F_{k_i}(x^*)\right\|\le L\left\|x^{k_i}-x^*\right\|,\quad \forall i=1,2,\cdots$$

and $\{x^{k_i}\}$ is bounded, the sequence $\left\{F_{k_i}(x^{k_i})\right\}$ is also bounded. Therefore, from (54) and (56) we get

$$\lim_{k_i\to\infty}\left\|(I-P_{Q_{k_i}})Ax^{k_i}\right\|=0,$$

that is,

$$\lim_{k_i\to\infty}P_{Q_{k_i}}(Ax^{k_i})-Ax^{k_i}=0. \tag{57}$$

Since $P_{Q_{k_i}}(Ax^{k_i})\in Q_{k_i}$, we have

$$q(Ax^{k_i})+\left\langle \eta^{k_i}, P_{Q_{k_i}}(Ax^{k_i})-Ax^{k_i}\right\rangle \le 0 .$$

Letting $k_i \to \infty$ from Lemma 1 and (57), we deduce that

$$q(Ax^{\infty}) \le 0 ,$$

that is, $Ax^{\infty} \in Q$. Therefore, x^{∞} is a solution of the SFP. Because the subsequence $\{x^{k_i}\}$ converges to x^{∞}, for an arbitrary scalar $\varepsilon > 0$, there exists a $k_l > 0$ such that

$$\left\| x^{k_l} - x^{\infty} \right\| \le \varepsilon .$$

On the other hand, since x^{∞} is a solution point, it follows from (51) or (52) that

$$\left\| x^{k} - x^{\infty} \right\| \le \left\| x^{k_l} - x^{\infty} \right\| \le \varepsilon \qquad \forall k \ge k_l ,$$

and thus the sequence $\{x^k\}$ converges to x^{∞}, which is a solution point of the SFP. □

4 Numerical results

In this section, we apply the proposed methods to solve the following split feasibility problems (Examples 1 and 2), which were tested in [10], to verify the effectiveness and computational superiority compared to the modified relaxed CQ algorithm in [2].

All the codes were written in Matlab and run on an HP Compaq 6910p notebook. For the CQ algorithm in [2], Algorithms 1 and 2, we take $\varepsilon = 10^{-10}$, $\alpha_0 = 1$, $\mu = 0.9$, $\nu = 0.4$, $\delta_k \equiv 1.8$, and $\gamma_k \equiv 1.8$. Since the test problems are from [10], we also list the original results by the halfspace-relaxation projection method in [10]. The numerical results for Examples 1 and 2 are reported in Tables 1-8.

Example 1 (*A convex feasibility problem*). Let $C = \left\{ x \in R^3 \middle| x_2^2 + x_3^2 - 4 \le 0 \right\}$, $Q = \left\{ x \in R^3 \middle| x_3 - 1 - x_1^2 \le 0 \right\}$. Find some point x in $C \cap Q$.

Tab. 1 Results for Example 1 using Qu and Xiu method in [10]

Starting points	Number of iterations	CPU(s)	Approximate solution
$(1,2,3,0,0,0)^T$	43	0.0500	$(0.3213, 0.2815, 0.1425)^T$
$(1,1,1,1,1,1)^T$	67	0.0910	$(0.8577, 0.8577, 1.3097)^T$
$(1,2,3,4,5,6)^T$	85	0.1210	$(1.1548, 0.8518, 1.8095)^T$

Tab. 2 Results for Example 1 using Qu and Xiu method in [2]

Starting points	Number of iterations	CPU(s)	Approximate solution
$(1,2,3)^T$	5	0.1250	$(1.0000, 1.1094, 1.6641)^T$
$(1,1,1)^T$	0	0.0320	$(1.0000, 1.0000, 1.0000)^T$
rand(3,1)*10	130	0.0780	$(0.8665, 0.6369, 1.7508)^T$

Tab. 3 Results for Example 1 using Algorithm 1

Starting points	Number of iterations	CPU(s)	Approximate solution
$(1,2,3)^T$	5	0.1870	$(1.0000, 1.1094, 1.6641)^T$
$(1,1,1)^T$	0	0.0310	$(1.0000, 1.0000, 1.0000)^T$
rand(3,1)*10	2	0.0940	$(1.0748, 0.6630, 1.6190)^T$

Tab. 4 Results for Example 1 using Algorithm 2

Starting points	Number of iterations	CPU(s)	Approximate solution
$(1,2,3)^T$	1	0.1560	$(1.0000, 0.7538, 1.1308)^T$
$(1,1,1)^T$	0	0.0310	$(1.0000, 1.0000, 1.0000)^T$
rand(3,1)*10	2	0.1100	$(0.6778, 0.4818, 1.3998)^T$

Example 2 (A split feasibility problem). Let

$$A=\begin{pmatrix} 2 & -1 & 3 \\ 4 & 2 & 5 \\ 2 & 0 & 2 \end{pmatrix},$$

$C=\left\{x\in R^3 \middle| x_1+x_2^2+2x_3\le 0\right\}$, $Q=\left\{x\in R^3 \middle| x_1^2+x_2-x_3\le 0\right\}$. Find some point $x\in C$ with $Ax\in Q$.

Tab. 5 Results for Example 2 using Qu and Xiu method in [10]

Starting points	Number of iterations	CPU(s)	Approximate solution
$(1,2,3,0,0,0)^T$	1890	2.7740	$(-0.1203, 0.0285, 0.0582)^T$
$(1,1,1,1,1,1)^T$	2978	4.2860	$(0.8603, -0.1658, -0.5073)^T$
$(1,2,3,4,5,6)^T$	3317	4.8570	$(3.6522, -0.1526, -2.3719)^T$

Tab. 6 Results for Example 2 using Qu and Xiu method in [2]

Starting points	Number of iterations	CPU(s)	Approximate solution
$(1,2,3)^T$	64	0.1570	$(-0.4019, 0.0674, 0.1967)^T$
$(1,1,1)^T$	81	0.0940	$(0.3568, 0.0343, -0.2652)^T$
rand(3,1)*10	105	0.0940	$(0.8747, 0.0795,-0.6876)^T$

Tab. 7 Results for Example 2 using Algorithm 1

Starting points	Number of iterations	CPU(s)	Approximate solution
$(1,2,3)^T$	4	0.1410	$(-0.4024, 0.0658, 0.1958)^T$
$(1,1,1)^T$	5	0.0940	$(0.3532, 0.0392, -0.2707)^T$
rand(3,1)*10	8	0.0940	$(0.8768, 0.0604, -0.6844)^T$

Tab. 8 Results for Example 2 using Algorithm 2

Starting points	Number of iterations	CPU(s)	Approximate solution
$(1,2,3)^T$	6	0.1720	$(-0.4305, 0.0774, 0.1048)^T$
$(1,1,1)^T$	1	0.1090	$(0.2000, -0.6000, -0.6000)^T$
rand(3,1)*10	7	0.1090	$(0.7984, -0.0384, -0.9042)^T$

These numerical data justify the computational superiority of the proposed methods over the modified relaxed CQ algorithm in [2] and the halfspace-relaxation projection method in [10].

5 Conclusion

For solving the split feasibility problem, this paper presents some improved relaxed CQ methods which are based on the modified relaxed CQ algorithm in [2]. The additional computational load resulted by the new methods is negligible, compared to the algorithm in [2]. The preliminary numerical tests show that the proposed methods are attractive in practice.

References

[1] CENSOR Y, ELFVING T. A multiprojection algorithm using Bregman projections in a product space[J]. Numerical Algorithms, 1994, 8: 221～239.
[2] QU B, XIU N H. A note on the CQ algorithm for the split feasibility problem[J]. Inverse Problems, 2005, 21: 1655～1665.
[3] BYRNE C L. Iterative oblique projection onto convex sets and the split feasibility problem[J]. Inverse Problems, 2002, 18: 441～453.
[4] YANG Q Z. The relaxed CQ algorithm solving the split feasibility problem[J]. Inverse Problems, 2004, 20: 1261～1266.
[5] ROCKAFELLAR R T. Convex Analysis[M]. Princeton, NJ: Princeton University Press, 1970.
[6] FUKUSHIMA M. A relaxed projection method for variational inequalities[J]. Mathematical Programming, 1986, 35: 58～70.
[7] POLYAK B T. Minimization of unsmooth functionals[J]. USSR Computational Mathematics and Mathematical Physics, 1969, 9: 14～29.
[8] HE B S, LIAO L-Z. Improvements of some projection methods for monotone nonlinear variational inequalities[J]. Journal of Optimization Theory and Applications, 2002, 112: 111～128.
[9] HE B S, YUAN X M, ZHANG J. J. Z. Comparison of two kinds of prediction-correction methods for monotone variational inequalities[J]. Computational Optimization and Applications, 2004, 27: 247～267.
[10] QU B, XIU N H. A new halfspace-relaxation projection method for the split feasibility problem[J]. Linear Algebra and its Applications, 2008, 428: 1218～1229.

求解分裂可行问题的改进的松弛 CQ 方法

李敏

（东南大学，经济管理学院，南京 210096）

摘要：本文提出求解分裂可行问题的改进的松弛 CQ 方法。这些新方法是基于修正的松弛 CQ 算法，沿着下降方向搜索最优步长产生新的迭代点。在适度的假设条件下，新方法是全局收敛的。初步的数值结果显示了新方法在计算上的优越性。

关键词：CQ 算法；分裂可行问题；步长

中图分类号：O224；O221.2