凸优化和单调变分不等式的收缩算法

第十六讲: 三块可分离凸优化问题的平行分裂增广 Lagrange 乘子法

Parallel splitting augmented Lagrangian methods for convex optimization with three separable blocks

南京大学数学系　　何炳生
hebma@nju.edu.cn

The context of this lecture is based on the publication [6]

1 Convex Optimization with 3 separable operators

We consider the linearly constrained convex optimization with 3 separable operators:

$$\min\{\theta_1(x)+\theta_2(y)+\theta_3(z)\,|\,Ax+By+Cz=b,\ x\in\mathcal{X},\ y\in\mathcal{Y},\ z\in\mathcal{Z}\} \quad (1.1)$$

where $\theta_1(x):\Re^{n_1}\to\Re,\ \theta_2(y):\Re^{n_2}\to\Re,\ \theta_3(z):\Re^{n_3}\to\Re$ are convex; $A\in\Re^{m\times n_1}$, $B\in\Re^{m\times n_2}$, $C\in\Re^{m\times n_3}$, $b\in\Re^m$; $\mathcal{X}\subset\Re^{n_1}$, $\mathcal{Y}\subset\Re^{n_2}$ and $\mathcal{Z}\subset\Re^{n_3}$ are given convex set. Let $n=n_1+n_2+n_3$.

This optimization problem is equivalent to find $(x^*,y^*,z^*,\lambda^*)\in\Omega$, such that

$$\left\{\begin{array}{l}\theta_1(x)-\theta_1(x^*)+(x-x^*)^T(-A^T\lambda^*)\ge 0,\\ \theta_2(y)-\theta_2(y^*)+(y-y^*)^T(-B^T\lambda^*)\ge 0,\\ \theta_3(z)-\theta_3(z^*)+(z-z^*)^T(-C^T\lambda^*)\ge 0,\\ (\lambda-\lambda^*)^T(Ax^*+By^*+Cz^*-b)\ge 0,\end{array}\right. \quad \forall\,(x,y,z,\lambda)\in\Omega, \qquad (1.2)$$

where

$$\Omega=\mathcal{X}\times\mathcal{Y}\times\mathcal{Z}\times\Re^m.$$

By denoting

$$u = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad w = \begin{pmatrix} x \\ y \\ z \\ \lambda \end{pmatrix} \quad F(w) = \begin{pmatrix} -A^T\lambda \\ -B^T\lambda \\ -C^T\lambda \\ Ax + By + Cz - b \end{pmatrix},$$

and

$$\theta(u) = \theta_1(x) + \theta_2(y) + \theta_3(z),$$

the optimal condition (1.2) can be written as

$$w^* \in \Omega, \quad \theta(u) - \theta(u^*) + (w - w^*)^T F(w^*) \geq 0, \quad \forall\, w \in \Omega.$$

For convenience, in this paper, we use the notations

$$v = \begin{pmatrix} y \\ z \\ \lambda \end{pmatrix}, \qquad \text{and} \qquad \mathcal{V}^* = \{(y^*, z^*, \lambda^*) \,|\, (x^*, y^*, z^*, \lambda^*) \in \Omega^*\}.$$

2 Full Parallel Splitting ALM

Full parallel splitting augmented Lagrangian method is a prediction-correction method. First it generates a predictor:

1. From given $(x^k, y^k, z^k, \lambda^k)$, obtain $\tilde{x}^k$, $\tilde{y}^k$ and $\tilde{z}^k$, by the following parallel manner:

$$\tilde{x}^k = \text{Argmin}\left\{ \begin{array}{c} \theta_1(x) - (\lambda^k)^T(Ax + By^k + Cz^k - b) \\ +\frac{\beta}{2}\|Ax + By^k + Cz^k - b\|^2 \end{array} \middle| x \in \mathcal{X} \right\} \tag{2.1a}$$

$$\tilde{y}^k = \text{Argmin}\left\{ \begin{array}{c} \theta_2(y) - (\lambda^k)^T(Ax^k + By + Cz^k - b) \\ +\frac{\beta}{2}\|Ax^k + By + Cz^k - b\|^2 \end{array} \middle| y \in \mathcal{Y} \right\} \tag{2.1b}$$

$$\tilde{z}^k = \text{Argmin}\left\{ \begin{array}{c} \theta_3(z) - (\lambda^k)^T(Ax^k + By^k + Cz - b) \\ +\frac{\beta}{2}\|Ax^k + By^k + Cz - b\|^2 \end{array} \middle| z \in \mathcal{Z} \right\} \tag{2.1c}$$

2. Update $\tilde{\lambda}^k$ by

$$\tilde{\lambda}^k = \lambda^k - \beta(A\tilde{x}^k + B\tilde{y}^k + C\tilde{z}^k - b). \tag{2.1d}$$

Note that the solution $(\tilde{x}^k, \tilde{y}^k, \tilde{z}^k) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ of (2.1a)-(2.1c) satisfies

$$\begin{cases} \theta_1(x) - \theta_1(\tilde{x}^k) + (x - \tilde{x}^k)^T \big\{ -A^T \lambda^k + \beta A^T \big(A\tilde{x}^k + By^k + Cz^k - b \big) \big\} \ge 0 \\ \theta_2(y) - \theta_2(\tilde{y}^k) + (y - \tilde{y}^k)^T \big\{ -B^T \lambda^k + \beta B^T \big(Ax^k + B\tilde{y}^k + Cz^k - b \big) \big\} \ge 0 \\ \theta_3(z) - \theta_3(\tilde{z}^k) + (z - \tilde{z}^k)^T \big\{ -C^T \lambda^k + \beta C^T \big(Ax^k + By^k + C\tilde{z}^k - b \big) \big\} \ge 0 \end{cases} \tag{2.2}$$

for all $(x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. Using

$$\tilde{\lambda}^k = \lambda^k - \beta(A\tilde{x}^k + B\tilde{y}^k + C\tilde{z}^k - b)$$

and by a manipulation, (2.2) can be rewritten as $(\tilde{x}^k, \tilde{y}^k, \tilde{z}^k) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$,

$$\theta(u) - \theta(\tilde{u}^k) + \begin{pmatrix} x - \tilde{x}^k \\ y - \tilde{y}^k \\ z - \tilde{z}^k \end{pmatrix}^T \begin{pmatrix} -A^T \tilde{\lambda}^k + \beta A^T [B(y^k - \tilde{y}^k) + C(z^k - \tilde{z}^k)] \\ -B^T \tilde{\lambda}^k + \beta B^T [A(x^k - \tilde{x}^k) + C(z^k - \tilde{z}^k)] \\ -C^T \tilde{\lambda}^k + \beta C^T [A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)] \end{pmatrix} \ge 0, \tag{2.3}$$

for all $(x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. Combining the last inequality with (2.1d), we have

$$\theta(u)-\theta(\tilde{u}^k)+\begin{pmatrix} x-\tilde{x}^k \\ y-\tilde{y}^k \\ z-\tilde{z}^k \\ \lambda-\tilde{\lambda}^k \end{pmatrix}^T \left\{ \begin{pmatrix} -A^T\tilde{\lambda}^k \\ -B^T\tilde{\lambda}^k \\ -C^T\tilde{\lambda}^k \\ A\tilde{x}^k+B\tilde{y}^k+C\tilde{z}^k-b \end{pmatrix} \right.$$

$$+\beta \begin{pmatrix} A^T \\ B^T \\ C^T \\ 0 \end{pmatrix} \big(A(x^k-\tilde{x}^k)+B(y^k-\tilde{y}^k)+C(z^k-\tilde{z}^k)\big)$$

$$\left. +\begin{pmatrix} \beta A^TA & 0 & 0 & 0 \\ 0 & \beta B^TB & 0 & 0 \\ 0 & 0 & \beta C^TC & 0 \\ 0 & 0 & 0 & \frac{1}{\beta}I_m \end{pmatrix} \begin{pmatrix} \tilde{x}^k-x^k \\ \tilde{y}^k-y^k \\ \tilde{z}^k-z^k \\ \tilde{\lambda}^k-\lambda^k \end{pmatrix} \right\} \ge 0,\ \forall w\in\Omega. \quad (2.4)$$

Based on the above analysis, we have the following lemma.

Lemma 2.1 *Let* $\tilde{w}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{z}^k, \tilde{\lambda}^k) \in \Omega$ *be generated by* (2.1) *from the given* $w^k = (x^k, y^k, z^k, \lambda^k)$. *Then, we have*

$$(\tilde{w}^k - w^*)^T H(w^k - \tilde{w}^k) \geq (\tilde{w}^k - w^*)^T \big(F(\tilde{w}^k) + \eta(u^k, \tilde{u}^k)\big), \qquad (2.5)$$

where

$$\eta(u^k, \tilde{u}^k) = \beta \begin{pmatrix} A^T \\ B^T \\ C^T \\ 0 \end{pmatrix} \big[A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k) + C(z^k - \tilde{z}^k)\big] \qquad (2.6)$$

and

$$H = \begin{pmatrix} \beta A^T A & 0 & 0 & 0 \\ 0 & \beta B^T B & 0 & 0 \\ 0 & 0 & \beta C^T C & 0 \\ 0 & 0 & 0 & \frac{1}{\beta} I_m \end{pmatrix}. \qquad (2.7)$$

Proof. Setting $(x, y, z, \lambda) = (x^*, y^*, z^*, \lambda^*)$ in (2.4), the assertion follows directly. $\Box$

Since F is monotone and $\tilde{w}^k \in \Omega$, it follows that

$$(\tilde{w}^k - w^*)^T F(\tilde{w}^k) \ge (\tilde{w}^k - w^*)^T F(w^*) \ge 0. \qquad (2.8)$$

In addition, by using $Ax^* + By^* + Cz^* = b$ and

$$\beta(A\tilde{x}^k + B\tilde{y}^k + C\tilde{z}^k - b) = (\lambda^k - \tilde{\lambda}^k),$$

we have

$$\begin{aligned} &(\tilde{w}^k - w^*)^T \eta(u^k, \tilde{u}^k) \\ &\quad = \quad (A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k) + C(z^k - \tilde{z}^k))^T \beta(A\tilde{x}^k + B\tilde{y}^k + C\tilde{z}^k - b) \\ &\quad = \quad (\lambda^k - \tilde{\lambda}^k)^T [A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k) + C(z^k - \tilde{z}^k)]. \qquad (2.9) \end{aligned}$$

Lemma 2.2 *Let* $\tilde{w}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{z}^k, \tilde{\lambda}^k) \in \Omega$ *be generated by* (2.1) *from the* *given* $w^k = (x^k, y^k, z^k, \lambda^k)$*. Then, we have*

$$(w^k - w^*)^T H(w^k - \tilde{w}^k) \geq \varphi(w^k, \tilde{w}^k), \quad \forall\, w^* \in \Omega^*, \tag{2.10}$$

where

$$\begin{aligned}\varphi(w^k, \tilde{w}^k) &= \|w^k - \tilde{w}^k\|_H^2 \\ &\quad +(\lambda^k - \tilde{\lambda}^k)^T \big(A(x^k - \tilde{x}^k)+B(y^k - \tilde{y}^k)+C(z^k - \tilde{z}^k)\big).\end{aligned} \tag{2.11}$$

In addition, we have

$$\varphi(w^k, \tilde{w}^k) \geq \frac{2-\sqrt{3}}{2}\|w^k - \tilde{w}^k\|_H^2. \tag{2.12}$$

Proof. First, using (2.5), (2.8) and (2.9) we obtain that

$$\begin{aligned}&(\tilde{w}^k - w^*)^T H(w^k - \tilde{w}^k) \\ &\quad \geq \quad (\lambda^k - \tilde{\lambda}^k)^T \big(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k) + C(z^k - \tilde{z}^k)\big).\end{aligned}$$

The first assertion of this lemma follows from the last inequality and the definition of $\varphi(w^k, \tilde{w}^k)$ directly.

Now, we turn to the second assertion (2.12). Notice that

$$\varphi(w^k,\tilde{w}^k) = \|w^k-\tilde{w}^k\|_H^2 + (\lambda^k-\tilde{\lambda}^k)^T\big(A(x^k-\tilde{x}^k)+B(y^k-\tilde{y}^k)+C(z^k-\tilde{z}^k)\big)$$

$$= \begin{pmatrix} x^k-\tilde{x}^k \\ y^k-\tilde{y}^k \\ z^k-\tilde{z}^k \\ \lambda^k-\tilde{\lambda}^k \end{pmatrix}^T \begin{pmatrix} \beta A^TA & 0 & 0 & \frac{1}{2}A^T \\ 0 & \beta B^TB & 0 & \frac{1}{2}B^T \\ 0 & 0 & \beta C^TC & \frac{1}{2}C^T \\ \frac{1}{2}A & \frac{1}{2}B & \frac{1}{2}C & \frac{1}{\beta}I_m \end{pmatrix} \begin{pmatrix} x^k-\tilde{x}^k \\ y^k-\tilde{y}^k \\ z^k-\tilde{z}^k \\ \lambda^k-\tilde{\lambda}^k \end{pmatrix}$$

$$= \begin{pmatrix} \sqrt{\beta}A(x^k-\tilde{x}^k) \\ \sqrt{\beta}B(y^k-\tilde{y}^k) \\ \sqrt{\beta}C(z^k-\tilde{z}^k) \\ \sqrt{1/\beta}(\lambda^k-\tilde{\lambda}^k) \end{pmatrix}^T \begin{pmatrix} I_m & 0 & 0 & \frac{1}{2}I_m \\ 0 & I_m & 0 & \frac{1}{2}I_m \\ 0 & 0 & I_m & \frac{1}{2}I_m \\ \frac{1}{2}I_m & \frac{1}{2}I_m & \frac{1}{2}I_m & I_m \end{pmatrix} \begin{pmatrix} \sqrt{\beta}A(x^k-\tilde{x}^k) \\ \sqrt{\beta}B(y^k-\tilde{y}^k) \\ \sqrt{\beta}C(z^k-\tilde{z}^k) \\ \sqrt{1/\beta}(\lambda^k-\tilde{\lambda}^k) \end{pmatrix}. \tag{2.13}$$

Because the eigenvalues of the matrix

$$\begin{pmatrix} 1 & 0 & 0 & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix} \quad \text{are } 1,\ 1 \text{ and } 1 \pm \frac{\sqrt{3}}{2},$$

The smallest eigenvalue is $\frac{2-\sqrt{3}}{2}$. Therefore,

$$\begin{aligned} \varphi(w^k, \tilde{w}^k) \quad &\geq \quad \frac{2-\sqrt{3}}{2} \begin{pmatrix} \sqrt{\beta}A(x^k - \tilde{x}^k) \\ \sqrt{\beta}B(y^k - \tilde{y}^k) \\ \sqrt{\beta}C(z^k - \tilde{z}^k) \\ \sqrt{1/\beta}(\lambda^k - \tilde{\lambda}^k) \end{pmatrix}^T \begin{pmatrix} \sqrt{\beta}A(x^k - \tilde{x}^k) \\ \sqrt{\beta}B(y^k - \tilde{y}^k) \\ \sqrt{\beta}C(z^k - \tilde{z}^k) \\ \sqrt{1/\beta}(\lambda^k - \tilde{\lambda}^k) \end{pmatrix} \\ &= \quad \frac{2-\sqrt{3}}{2} \|w^k - \tilde{w}^k\|_H^2. \end{aligned}$$

The lemma is proved. □

According to Lemma 2.2, we have

$$(w^k - w^*)^T H(w^k - \tilde{w}^k) \geq \varphi(w^k, \tilde{w}^k) \geq \frac{2-\sqrt{3}}{2} \|w^k - \tilde{w}^k\|_H^2.$$

This give us the possibility to produce a new iterate which is more closed to the solution set by the following correction.

Correction: Based on the predictor by (2.1), update the new iterate w^{k+1} by:

$$w^{k+1} = w^k - \alpha_k (w^k - \tilde{w}^k), \qquad \alpha_k = \gamma \alpha_k^*, \qquad \gamma \in (0,2) \tag{2.14a}$$

where

$$\alpha_k^* = \frac{\varphi(w^k, \tilde{w}^k)}{\|w^k - \tilde{w}^k\|_H^2}. \tag{2.14b}$$

By using (2.10) and (2.14), we obtain

$$\begin{aligned}
\|w^{k+1} - w^*\|_H^2 &= \|(w^k - w^*) - \gamma\alpha_k^*(w^k - \tilde{w}^k)\|_H^2 \\
&= \|w^k - w^*\|_H^2 - 2\gamma\alpha_k^*(w^k - w^*)^T H (w^k - \tilde{w}^k) \\
&\quad + \gamma^2 (\alpha_k^*)^2 \|w^k - \tilde{w}^k\|_H^2 \\
&\le \|w^k - w^*\|_H^2 - 2\gamma\alpha_k^* \varphi(w^k, \tilde{w}^k) \\
&\quad + \gamma^2 (\alpha_k^*)^2 \|w^k - \tilde{w}^k\|_H^2 \\
&\le \|w^k - w^*\|_H^2 - \gamma(2-\gamma)\alpha_k^* \varphi(w^k, \tilde{w}^k).
\end{aligned} \tag{2.15}$$

In comparison with the computational load for obtaining $(\tilde{x}^k, \tilde{y}^k, \tilde{z}^k)$, the calculation cost for step-size α_k^* is slight.

Convergence Using (2.14) and (2.14b), we obtain

$$\|w^{k+1} - w^*\|_H^2 \le \|w^k - w^*\|_H^2 - \frac{\gamma(2-\gamma)(7-4\sqrt{3})}{4}\|w^k - \tilde{w}^k\|_H^2.$$

♣ B. S. He, Parallel splitting augmented Lagrangian methods for monotone structured variational inequalities, Computational Optimization and Applications, 42, 195-212, 2009.

3 Partially Parallel Splitting ALM

In the partially parallel splitting ALM, x is an intermediate variable. The iteration begins with $v^k = (y^k, z^k, \lambda^k)$ and produces new iterate v^{k+1}. It is still a prediction-correction method.

Prediction:

1. From given (y^k, z^k, λ^k), obtain $\tilde{x}^k, \tilde{y}^k$ and $\tilde{z}^k$, by the following partially parallel manner:

$$\tilde{x}^k = \text{Argmin}\left\{ \begin{array}{c} \theta_1(x) - (\lambda^k)^T(Ax + By^k + Cz^k - b) \\ +\frac{\beta}{2}\|Ax + By^k + Cz^k - b\|^2 \end{array} \middle| x \in \mathcal{X} \right\} \quad \text{(3.1a)}$$

$$\tilde{y}^k = \text{Argmin}\left\{ \begin{array}{c} \theta_2(y) - (\lambda^k)^T(A\tilde{x}^k + By + Cz^k - b) \\ +\frac{\beta}{2}\|A\tilde{x}^k + By + Cz^k - b\|^2 \end{array} \middle| y \in \mathcal{Y} \right\} \quad \text{(3.1b)}$$

$$\tilde{z}^k = \text{Argmin}\left\{ \begin{array}{c} \theta_3(z) - (\lambda^k)^T(A\tilde{x}^k + By^k + Cz - b) \\ +\frac{\beta}{2}\|A\tilde{x}^k + By^k + Cz - b\|^2 \end{array} \middle| z \in \mathcal{Z} \right\} \quad \text{(3.1c)}$$

2. Update $\tilde{\lambda}^k$ by

$$\tilde{\lambda}^k = \lambda^k - \beta(A\tilde{x}^k + B\tilde{y}^k + C\tilde{z}^k - b). \quad \text{(3.1d)}$$

Note that in (3.1b) and (3.1c), we use $\tilde{x}^k$ generated by (3.1a).

Analysis The solution $(\tilde{x}^k, \tilde{y}^k, \tilde{z}^k) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ of (3.1a)-(3.1c) satisfies

$$\begin{cases} \theta_1(x) - \theta_1(\tilde{x}^k) + (x - \tilde{x}^k)^T\{-A^T\lambda^k + \beta A^T\big(A\tilde{x}^k + By^k + Cz^k - b\big)\} \geq 0 \\ \theta_2(y) - \theta_2(\tilde{y}^k) + (y - \tilde{y}^k)^T\{-B^T\lambda^k + \beta B^T\big(A\tilde{x}^k + B\tilde{y}^k + Cz^k - b\big)\} \geq 0 \\ \theta_3(z) - \theta_3(\tilde{z}^k) + (z - \tilde{z}^k)^T\{-C^T\lambda^k + \beta C^T\big(A\tilde{x}^k + By^k + C\tilde{z}^k - b\big)\} \geq 0 \end{cases} \tag{3.2}$$

for all $(x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. Using $\tilde{\lambda}^k = \lambda^k - \beta(A\tilde{x}^k + B\tilde{y}^k + C\tilde{z}^k - b)$ and by a manipulation, (3.2) can be rewritten as $(\tilde{x}^k, \tilde{y}^k, \tilde{z}^k) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$,

$$\theta(u) - \theta(\tilde{u}^k) + \begin{pmatrix} x - \tilde{x}^k \\ y - \tilde{y}^k \\ z - \tilde{z}^k \end{pmatrix}^T \begin{pmatrix} -A^T\tilde{\lambda}^k + \beta A^T[B(y^k - \tilde{y}^k) + C(z^k - \tilde{z}^k)] \\ -B^T\tilde{\lambda}^k + \beta B^T[\quad 0 \quad + C(z^k - \tilde{z}^k)] \\ -C^T\tilde{\lambda}^k + \beta C^T[B(y^k - \tilde{y}^k) + \quad 0 \quad] \end{pmatrix} \geq 0,$$

for all $(x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. Combining the last inequality with (3.1d), we have

$$\theta(u)-\theta(\tilde{u}^k)+\begin{pmatrix} x-\tilde{x}^k \\ y-\tilde{y}^k \\ z-\tilde{z}^k \\ \lambda-\tilde{\lambda}^k \end{pmatrix}^T \left\{ \begin{pmatrix} -A^T\tilde{\lambda}^k \\ -B^T\tilde{\lambda}^k \\ -C^T\tilde{\lambda}^k \\ A\tilde{x}^k+B\tilde{y}^k+C\tilde{z}^k-b \end{pmatrix} \right.$$

$$+\beta\begin{pmatrix} A^T \\ B^T \\ C^T \\ 0 \end{pmatrix}\big(B(y^k-\tilde{y}^k)+C(z^k-\tilde{z}^k)\big)$$

$$\left. +\begin{pmatrix} 0 & 0 & 0 \\ \beta B^T B & 0 & 0 \\ 0 & \beta C^T C & 0 \\ 0 & 0 & \frac{1}{\beta}I_m \end{pmatrix}\begin{pmatrix} \tilde{y}^k-y^k \\ \tilde{z}^k-z^k \\ \tilde{\lambda}^k-\lambda \end{pmatrix}\right\} \geq 0,\ \forall w\in\Omega. \qquad (3.3)$$

Based on the above analysis, we have the following lemma.

Lemma 3.1 *Let* $\tilde{w}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{z}^k, \tilde{\lambda}^k) \in \Omega$ *be generated by* (3.1) *from the given* $v^k = (y^k, z^k, \lambda^k)$. *Then, we have*

$$(\tilde{v}^k - v^*)^T H(v^k - \tilde{v}^k) \geq (\tilde{w}^k - w^*)^T \big(F(\tilde{w}^k) + \eta(u^k, \tilde{u}^k)\big), \tag{3.4}$$

where

$$\eta(u^k, \tilde{u}^k) = \beta \begin{pmatrix} A^T \\ B^T \\ C^T \\ 0 \end{pmatrix} \big(B(y^k - \tilde{y}^k) + C(z^k - \tilde{z}^k)\big) \tag{3.5}$$

and

$$H = \begin{pmatrix} \beta B^T B & 0 & 0 \\ 0 & \beta C^T C & 0 \\ 0 & 0 & \frac{1}{\beta} I_m \end{pmatrix}. \tag{3.6}$$

Proof. Setting $(x, y, z, \lambda) = (x^*, y^*, z^*, \lambda^*)$ in (3.3), and using $v = (y, z, \lambda)$ the assertion follows directly. $\Box$

Since F is monotone and $\tilde{w}^k \in \Omega$, it follows that

$$(\tilde{w}^k - w^*)^T F(\tilde{w}^k) \geq (\tilde{w}^k - w^*)^T F(w^*) \geq 0. \tag{3.7}$$

In addition, by using $Ax^* + By^* + Cz^* = b$ and

$$\beta(A\tilde{x}^k + B\tilde{y}^k + C\tilde{z}^k - b) = (\lambda^k - \tilde{\lambda}^k),$$

we have

$$\begin{aligned} &\eta(u^k, \tilde{u}^k)^T(\tilde{w}^k - w^*) \\ = \quad & (B(y^k - \tilde{y}^k) + C(z^k - \tilde{z}^k))^T \beta(A\tilde{x}^k + B\tilde{y}^k + C\tilde{z}^k - b) \\ = \quad & (B(y^k - \tilde{y}^k) + C(z^k - \tilde{z}^k))^T(\lambda^k - \tilde{\lambda}^k). \end{aligned} \tag{3.8}$$

Lemma 3.2 *Let* $\tilde{w}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{z}^k, \tilde{\lambda}^k) \in \Omega$ *be generated by* (3.1) *from the given* $v^k = (y^k, z^k, \lambda^k)$. *Then, we have*

$$(v^k - v^*)^T H(v^k - \tilde{v}^k) \geq \varphi(v^k, \tilde{v}^k), \quad \forall\, w^* \in \Omega^*, \tag{3.9}$$

where

$$\varphi(v^k, \tilde{v}^k) = \|v^k - \tilde{v}^k\|_H^2 \\ +(\lambda^k - \tilde{\lambda}^k)^T \big(B(y^k - \tilde{y}^k) + C(z^k - \tilde{z}^k)\big). \qquad (3.10)$$

Proof. First, using (3.4), (3.7) and (3.8) we obtain that

$$(\tilde{v}^k - v^*)^T H(v^k - \tilde{v}^k) \\ \geq \quad (\lambda^k - \tilde{\lambda}^k)^T \big(B(y^k - \tilde{y}^k) + C(z^k - \tilde{z}^k)\big).$$

The assertion of this lemma follows from the last inequality and the definition of $\varphi(v^k, \tilde{v}^k)$ directly. □

Lemma 3.3 *Let* $\tilde{w}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{z}^k, \tilde{\lambda}^k) \in \Omega$ *be generated by* (3.1) *from the given* $v^k = (y^k, z^k, \lambda^k)$. *For* H *defined in* (3.6) *and* $\varphi(v^k, \tilde{v}^k)$ *defined in* (3.10), *we have*

$$\varphi(v^k, \tilde{v}^k) \geq \frac{2 - \sqrt{2}}{2} \|v^k - \tilde{v}^k\|_H^2. \qquad (3.11)$$

Proof. According to the definition of (3.10), we obtain

$$\varphi(v^k,\tilde v^k) = \|v^k-\tilde v^k\|_H^2 + (\lambda^k-\tilde\lambda^k)^T\big(B(y^k-\tilde y^k)+C(z^k-\tilde z^k)\big)$$

$$= \begin{pmatrix} y^k-\tilde y^k \\ z^k-\tilde z^k \\ \lambda^k-\tilde\lambda^k \end{pmatrix}^T \begin{pmatrix} \beta B^TB & 0 & \frac{1}{2}B^T \\ 0 & \beta C^TC & \frac{1}{2}C^T \\ \frac{1}{2}B & \frac{1}{2}C & \frac{1}{\beta}I_m \end{pmatrix} \begin{pmatrix} y^k-\tilde y^k \\ z^k-\tilde z^k \\ \lambda^k-\tilde\lambda^k \end{pmatrix}$$

$$= \begin{pmatrix} \sqrt{\beta}B(y^k-\tilde y^k) \\ \sqrt{\beta}C(z^k-\tilde z^k) \\ \sqrt{1/\beta}(\lambda^k-\tilde\lambda^k) \end{pmatrix}^T \begin{pmatrix} I_m & 0 & \frac{1}{2}I_m \\ 0 & I_m & \frac{1}{2}I_m \\ \frac{1}{2}I_m & \frac{1}{2}I_m & I_m \end{pmatrix} \begin{pmatrix} \sqrt{\beta}B(y^k-\tilde y^k) \\ \sqrt{\beta}C(z^k-\tilde z^k) \\ \sqrt{1/\beta}(\lambda^k-\tilde\lambda^k) \end{pmatrix}.$$

Because the matrix

$$\begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix}$$

is positive definite and its eigenvalues are 1 and $1 \pm \frac{\sqrt{2}}{2}$, thus, we obtain

$$\begin{aligned} \varphi(v^k, \tilde{v}^k) &\geq \frac{2-\sqrt{2}}{2} \begin{pmatrix} \sqrt{\beta}B(y^k - \tilde{y}^k) \\ \sqrt{\beta}C(z^k - \tilde{z}^k) \\ \sqrt{1/\beta}(\lambda^k - \tilde{\lambda}^k) \end{pmatrix}^T \begin{pmatrix} \sqrt{\beta}B(y^k - \tilde{y}^k) \\ \sqrt{\beta}C(z^k - \tilde{z}^k) \\ \sqrt{1/\beta}(\lambda^k - \tilde{\lambda}^k) \end{pmatrix} \\ &= \frac{2-\sqrt{2}}{2} \|v^k - \tilde{v}^k\|_H^2, \end{aligned}$$

and the assertion is proved. □

Correction Based on the predictor by (3.1), update the new iterate v^{k+1} by

$$v^{k+1} = v^k - \alpha_k(v^k - \tilde{v}^k), \qquad \alpha_k = \gamma\alpha_k^*, \qquad \gamma \in (0,2) \tag{3.12a}$$

where

$$\alpha_k^* = \frac{\varphi(v^k, \tilde{v}^k)}{\|v^k - \tilde{v}^k\|_H^2} \tag{3.12b}$$

By using (3.9) and (3.12), we obtain

$$\|v^{k+1} - v^*\|_H^2 = \|(v^k - v^*) - \gamma\alpha_k^*(v^k - \tilde{v}^k)\|_H^2$$
$$\begin{aligned} &= \|v^k - v^*\|_H^2 - 2\gamma\alpha_k^*(v^k - v^*)^T H(v^k - \tilde{v}^k) + \gamma^2(\alpha_k^*)^2\|v^k - \tilde{v}^k\|_H^2 \\ &\le \|v^k - v^*\|_H^2 - 2\gamma\alpha_k^*\varphi(v^k, \tilde{v}^k) + \gamma^2(\alpha_k^*)^2\|v^k - \tilde{v}^k\|_H^2 \\ &= \|v^k - v^*\|_H^2 - \gamma(2-\gamma)\alpha_k^*\varphi(v^k, \tilde{v}^k). \end{aligned} \quad (3.13)$$

In comparison with the computational load for obtaining $(\tilde{x}^k, \tilde{y}^k, \tilde{z}^k)$, the calculation cost for step-size α_k^* is slight.

Convergence Using (3.11) and (3.12b), it follows from (3.13) that

$$\|v^{k+1} - v^*\|_H^2 \le \|v^k - v^*\|_H^2 - \frac{\gamma(2-\gamma)(3-2\sqrt{2})}{2}\|v^k - \tilde{v}^k\|_H^2.$$

From (3.11) we know that $\alpha_k^* \ge \frac{2-\sqrt{2}}{2}$. In the correction step, we can also take the fixed step size $\alpha_k \equiv \alpha \in (0, 2-\sqrt{2})$, the method is still convergent.

4 Extension to problems with 4 operators

We consider the following optimization problem with 4 separable operators:

$$\min\{\sum_{i=1}^{4}\theta_i(x_i)\,|\,\sum_{i=1}^{4}A_ix_i=b,\,x_i\in\mathcal{X}_i\} \tag{4.1}$$

where $\theta_i(x_i):\Re^{n_i}\rightarrow\Re$, is convex, $A_i\in\Re^{m\times n_i}$, $\mathcal{X}_i\subset\Re^{n_i}$ is given convex set, $i=1,\ldots,4$; $b\in\Re^m$. This optimization problem is equivalent to find $w^*=(x_1^*,x_2^*,x_3^*,x_4^*,\lambda^*)\in\Omega$, such that

$$\left\{\begin{array}{l}\theta_1(x_1)-\theta_1(x_1^*)+(x_1-x_1^*)^T(-A_1^T\lambda^*)\geq 0,\\ \theta_2(x_2)-\theta_2(x_2^*)+(x_2-x_2^*)^T(-A_2^T\lambda^*)\geq 0,\\ \theta_3(x_3)-\theta_3(x_3^*)+(x_3-x_3^*)^T(-A_4^T\lambda^*)\geq 0,\\ \theta_4(x_4)-\theta_4(x_4^*)+(x_4-x_4^*)^T(-A_4^T\lambda^*)\geq 0,\\ (\lambda-\lambda^*)^T(A_1x_1^*+A_2x_2^*+A_3x_3^*+A_4x_4^*-b)\geq 0,\end{array}\right. \quad \forall\, w\in\Omega, \tag{4.2}$$

where

$$\Omega=\mathcal{X}_1\times\mathcal{X}_2\times\mathcal{X}_3\times\mathcal{X}_4\times\Re^m.$$

By denoting

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}, \quad w = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ \lambda \end{pmatrix} \quad F(w) = \begin{pmatrix} -A_1^T\lambda \\ -A_2^T\lambda \\ -A_3^T\lambda \\ -A_4^T\lambda \\ \sum_{i=1}^4 A_i x_i - b, \end{pmatrix},$$

$$\text{and} \qquad \theta(x) = \theta_1(x_1) + \theta_2(x_2) + \theta_3(x_3) + \theta_4(x_4),$$

the optimal condition (4.2) can be written as

$$w^* \in \Omega, \quad \theta(x) - \theta(x^*) + (w - w^*)^T F(w^*) \geq 0, \quad \forall\, w \in \Omega.$$

Full parallel splitting augmented Lagrangian method

If we extend the full parallel splitting augmented Lagrangian method to solve (4.1), similarly as in Lemma 2.2 (see (2.10) and (2.11)), we have

$$(w^k - w^*)^T H(w^k - \tilde{w}^k) \geq \varphi(w^k, \tilde{w}^k), \quad \forall\, w^* \in \Omega^*, \tag{4.3}$$

where

$$\varphi(w^k, \tilde{w}^k) = \|w^k - \tilde{w}^k\|_H^2 + (\lambda^k - \tilde{\lambda}^k)^T \big(\sum_{i=1}^4 A_i(x_i^k - \tilde{x}_i^k)\big) \tag{4.4}$$

and

$$H = \textbf{diag}(\beta A_1^T A_1, \beta A_2^T A_2, \beta A_3^T A_3, \beta A_4^T A_4, \frac{1}{\beta} I_m).$$

Note that, similar as (2.13), we have

$$\varphi(w^k, \tilde{w}^k) = \begin{pmatrix} \sqrt{\beta}A_1(x_1^k - \tilde{x}_1^k) \\ \sqrt{\beta}A_2(x_2^k - \tilde{x}_2^k) \\ \sqrt{\beta}A_3(x_3^k - \tilde{x}_3^k) \\ \sqrt{\beta}A_4(x_4^k - \tilde{x}_4^k) \\ \sqrt{1/\beta}(\lambda^k - \tilde{\lambda}^k) \end{pmatrix}^T \begin{pmatrix} I & 0 & 0 & 0 & I/2 \\ 0 & I & 0 & 0 & I/2 \\ 0 & 0 & I & 0 & I/2 \\ 0 & 0 & 0 & I & I/2 \\ I/2 & I/2 & I/2 & I/2 & I \end{pmatrix} \begin{pmatrix} \sqrt{\beta}A_1(x_1^k - \tilde{x}_1^k) \\ \sqrt{\beta}A_2(x_2^k - \tilde{x}_2^k) \\ \sqrt{\beta}A_3(x_3^k - \tilde{x}_3^k) \\ \sqrt{\beta}A_4(x_4^k - \tilde{x}_4^k) \\ \sqrt{1/\beta}(\lambda^k - \tilde{\lambda}^k) \end{pmatrix}.$$

In other words, for the problem with 4 separable operators, if we use the full parallel splitting augmented Lagrangian method, we will met a matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 1 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix}.$$

The above matrix is positive semi-definite and its eigenvalues are $0, 1, 1, 1$ and 2. The vector $(1, 1, 1, 1, -2)^T$ is the eigenvector related to the eigenvalue 0. This means, even through $\|w^k - \tilde{w}^k\|_H \neq 0$, it can not guarantee that $\varphi(w^k, \tilde{w}^k) > 0$. Therefore, we suggest to use the partially splitting augmented Lagrangian method.

Partially parallel splitting augmented Lagrangian method

In the partially parallel splitting augmented Lagrangian method, x_1 is an intermediate variable during the iterations. After solving the x_1-subproblem, then solve the x_2, x_3, x_4-subproblems in the parallel manner !

Prediction Similarly as the prediction step (3.1) for problems with 3 operators

From given $(x_2^k, x_3^k, x_4^k, \lambda^k)$, obtain $\tilde{w}^k$ by the following partially parallel manner:

$$\tilde{x}_1^k = \text{Argmin}\left\{\begin{array}{c}\theta_1(x_1) - (\lambda^k)^T(A_1x_1 + A_2x_2^k + A_3x_3^k + A_4x_4^k - b) \\ + \frac{\beta}{2}\|A_1x_1 + A_2x_2^k + A_3x_3^k + A_4x_4^k - b\|^2\end{array}\middle| x_1 \in \mathcal{X}_1\right\} \quad (4.5a)$$

$$\tilde{x}_2^k = \text{Argmin}\left\{\begin{array}{c}\theta_2(x_2) - (\lambda^k)^T(A_1\tilde{x}_1^k + A_2x_2 + A_3x_3^k + A_4x_4^k - b) \\ + \frac{\beta}{2}\|A_1\tilde{x}_1^k + A_2x_2 + A_3x_3^k + A_4x_4^k - b\|^2\end{array}\middle| x_2 \in \mathcal{X}_2\right\} \quad (4.5b)$$

$$\tilde{x}_3^k = \text{Argmin}\left\{\begin{array}{c}\theta_3(x_3) - (\lambda^k)^T(A_1\tilde{x}_1^k + A_2x_2^k + A_3x_3 + A_4x_4^k - b) \\ + \frac{\beta}{2}\|A_1\tilde{x}_1^k + A_2x_2^k + A_3x_3 + A_4x_4^k - b\|^2\end{array}\middle| x_3 \in \mathcal{X}_3\right\} \quad (4.5c)$$

$$\tilde{x}_4^k = \text{Argmin}\left\{\begin{array}{c}\theta_4(x_4) - (\lambda^k)^T(A_1\tilde{x}_1^k + A_2x_2^k + A_3x_3^k + A_4x_4 - b) \\ + \frac{\beta}{2}\|A_1\tilde{x}_1^k + A_2x_2^k + A_3x_3^k + A_4x_4 - b\|^2\end{array}\middle| x_4 \in \mathcal{X}_4\right\} \quad (4.5d)$$

Update $\tilde{\lambda}^k$ by

$$\tilde{\lambda}^k = \lambda^k - \beta(A_1\tilde{x}_1^k + A_2\tilde{x}_2^k + A_3\tilde{x}_3^k + A_4\tilde{x}_4^k - b). \quad (4.5e)$$

Note that in (4.5b), (4.5c) and (4.5d), we use $\tilde{x}_1^k$ generated by (4.5a).

Similar as (3.3), for the $\tilde{w}^k$ generated by (4.5), we have

$$\theta(x)-\theta(\tilde{x}^k)+\begin{pmatrix} x_1-\tilde{x}_1^k \\ x_2-\tilde{x}_2^k \\ x_3-\tilde{x}_3^k \\ x_4-\tilde{x}_4^k \\ \lambda-\tilde{\lambda}^k \end{pmatrix}^T \left\{ \begin{pmatrix} -A_1^T\tilde{\lambda}^k \\ -A_2^T\tilde{\lambda}^k \\ -A_3^T\tilde{\lambda}^k \\ -A_4^T\tilde{\lambda}^k \\ A_1\tilde{x}_1^k+A_2\tilde{x}_2^k+A_3\tilde{x}_3^k+A_4\tilde{x}_4^k-b \end{pmatrix} \right.$$

$$+\beta\begin{pmatrix} A_1^T \\ A_2^T \\ A_3^T \\ A_4^T \\ 0 \end{pmatrix}\big[A_2(x_2^k-\tilde{x}_2^k)+A_3(x_3^k-\tilde{x}_3^k)+A_4(x_4^k-\tilde{x}_4^k)\big]$$

$$\left. +\begin{pmatrix} 0 & 0 & 0 & 0 \\ \beta A_2^TA_2 & 0 & 0 & 0 \\ 0 & \beta A_3^TA_3 & 0 & 0 \\ 0 & 0 & \beta A_4^TA_4 & 0 \\ 0 & 0 & 0 & \frac{1}{\beta}I_m \end{pmatrix}\begin{pmatrix} \tilde{x}_2^k-x_2^k \\ \tilde{x}_3^k-x_3^k \\ \tilde{x}_4^k-x_4^k \\ \tilde{\lambda}^k-\lambda \end{pmatrix}\right\} \geq 0, \ \forall w\in\Omega. \quad (4.6)$$

For convenience we use the notations

$$v = \begin{pmatrix} x_2 \\ x_3 \\ x_4 \\ \lambda \end{pmatrix} \quad \text{and} \quad \mathcal{V}^* = \{(x_2^*, x_3^*, x_4^*, \lambda^*) \,|\, (x_1^*, x_2^*, x_3^*, x_4^*, \lambda^*) \in \Omega^*\}.$$

Based on the above analysis, we have the following lemma.

Lemma 4.1 *Let* $\tilde{w}^k = (\tilde{x}_1^k, \tilde{x}_2^k, \tilde{x}_3^k, \tilde{x}_4^k, \tilde{\lambda}^k) \in \Omega$ *be generated by* (4.5) *from the given* $v^k = (x_2^k, x_3^k, x_4^k, \lambda^k)$*. Then, we have*

$$(\tilde{v}^k - v^*)^T H(v^k - \tilde{v}^k) \geq (\tilde{w}^k - w^*)^T \big(F(\tilde{w}^k) + \eta(x^k, \tilde{x}^k)\big), \tag{4.7}$$

where

$$H = \textbf{diag}(\beta A_2^T A_2, \beta A_3^T A_3, \beta A_4^T A_4, \frac{1}{\beta} I_m) \tag{4.8}$$

and

$$\eta(x^k,\tilde{x}^k)=\beta\begin{pmatrix}A_1^T\\A_2^T\\A_3^T\\A_4^T\\0\end{pmatrix}\left[A_2(x_2^k-\tilde{x}_2^k)+A_3(x_3^k-\tilde{x}_3^k)+A_4(x_4^k-\tilde{x}_4^k)\right] \quad (4.9)$$

Lemma 4.2 *Let* $\tilde{w}^k=(\tilde{x}_1^k,\tilde{x}_2^k,\tilde{x}_3^k,\tilde{x}_4^k,\tilde{\lambda}^k)\in\Omega$ *be generated by* (4.5) *from the given* $v^k=(x_2^k,x_3^k,x_4^k,\lambda^k)$. *Then, we have*

$$(v^k-v^*)^T H(v^k-\tilde{v}^k)\ge\varphi(v^k,\tilde{v}^k),\quad\forall\, v^*\in\mathcal{V}^*, \quad (4.10)$$

where

$$\begin{aligned}\varphi(v^k,\tilde{v}^k)&=\|v^k-\tilde{v}^k\|_H^2\\&\quad+(\lambda^k-\tilde{\lambda}^k)^T\left[A_2(x_2^k-\tilde{x}_2^k)+A_3(x_3^k-\tilde{x}_3^k)+A_4(x_4^k-\tilde{x}_4^k)\right].\end{aligned} \quad (4.11)$$

The proofs of Lemma 4.2 is similar as those in Lemma 3.2.

Lemma 4.3 *Let* $\tilde{w}^k = (\tilde{x}_1^k, \tilde{x}_2^k, \tilde{x}_3^k, \tilde{x}_4^k, \tilde{\lambda}^k) \in \Omega$ *be generated by* (4.5) *from the given* $v^k = (x_2^k, x_3^k, x_4^k, \lambda^k)$. *For* H *defined in* (4.8) *and* $\varphi(v^k, \tilde{v}^k)$ *defined in* (4.11), *we have*

$$\varphi(v^k, \tilde{v}^k) \geq \frac{2-\sqrt{3}}{2}\|v^k - \tilde{v}^k\|_H^2. \tag{4.12}$$

Proof. According to the definition of (4.11), we have

$$\begin{aligned}
\varphi(v^k, \tilde{v}^k) &= \|v^k - \tilde{v}^k\|_H^2 \\
&\quad +(\lambda^k - \tilde{\lambda}^k)^T\big[A_2(x_2^k - \tilde{x}_2^k) + A_3(x_3^k - \tilde{x}_3^k) + A_4(x_4^k - \tilde{x}_4^k)\big] \\
&= \begin{pmatrix} \sqrt{\beta}A_2(x_2^k - \tilde{x}_2^k) \\ \sqrt{\beta}A_3(x_3^k - \tilde{x}_3^k) \\ \sqrt{\beta}A_4(x_4^k - \tilde{x}_4^k) \\ \sqrt{1/\beta}(\lambda^k - \tilde{\lambda}^k) \end{pmatrix}^T \begin{pmatrix} I_m & 0 & 0 & \frac{1}{2}I_m \\ 0 & I_m & 0 & \frac{1}{2}I_m \\ 0 & 0 & I_m & \frac{1}{2}I_m \\ \frac{1}{2}I_m & \frac{1}{2}I_m & \frac{1}{2}I_m & I_m \end{pmatrix} \begin{pmatrix} \sqrt{\beta}A_2(x_2^k - \tilde{x}_2^k) \\ \sqrt{\beta}A_3(x_3^k - \tilde{x}_3^k) \\ \sqrt{\beta}A_4(x_4^k - \tilde{x}_4^k) \\ \sqrt{1/\beta}(\lambda^k - \tilde{\lambda}^k) \end{pmatrix}.
\end{aligned}$$

Note that the eigenvalues of the symmetric matrix

$$\begin{pmatrix} 1 & 0 & 0 & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix} \quad \text{are} \quad 1, \ 1 \text{ and } 1 \pm \frac{\sqrt{3}}{2},$$

and the smallest eigenvalue is $\frac{2-\sqrt{3}}{2}$. Therefore,

$$\begin{aligned} \varphi(v^k, \tilde{v}^k) \quad &\geq \quad \frac{2-\sqrt{3}}{2} \begin{pmatrix} \sqrt{\beta} A_2(x_2^k - \tilde{x}_2^k) \\ \sqrt{\beta} A_3(x_3^k - \tilde{x}_3^k) \\ \sqrt{\beta} A_4(x_4^k - \tilde{x}_4^k) \\ \sqrt{1/\beta}(\lambda^k - \tilde{\lambda}^k) \end{pmatrix}^T \begin{pmatrix} \sqrt{\beta} A_2(x_2^k - \tilde{x}_2^k) \\ \sqrt{\beta} A_3(x_3^k - \tilde{x}_3^k) \\ \sqrt{\beta} A_4(x_4^k - \tilde{x}_4^k) \\ \sqrt{1/\beta}(\lambda^k - \tilde{\lambda}^k) \end{pmatrix} \\ &= \quad \frac{2-\sqrt{3}}{2} \|v^k - \tilde{v}^k\|_H^2, \end{aligned}$$

and the assertion is proved. $\square$

Correction Based on the predictor by (4.5), update the new iterate v^{k+1} by

Update the new iterate v^{k+1} by

$$v^{k+1} = v^k - \alpha_k (v^k - \tilde{v}^k), \qquad \alpha_k = \gamma \alpha_k^*, \qquad \gamma \in (0, 2) \tag{4.13a}$$

where

$$\alpha_k^* = \frac{\varphi(v^k, \tilde{v}^k)}{\|v^k - \tilde{v}^k\|_H^2} \tag{4.13b}$$

By using (4.10) and (4.13), we obtain

$$\|v^{k+1} - v^*\|_H^2 \le \|v^k - v^*\|_H^2 - \gamma(2-\gamma)\alpha_k^* \varphi(v^k, \tilde{v}^k).$$

Convergence Using (4.12) and (4.13b), we obtain

$$\|v^{k+1} - v^*\|_H^2 \le \|v^k - v^*\|_H^2 - \frac{\gamma(2-\gamma)(7-4\sqrt{3})}{4} \|v^k - \tilde{v}^k\|_H^2. \tag{4.14}$$

This inequality (4.14) is essential for the convergence.

5 Conclusion Remarks

All the methods in this note are prediction-correction methods. The predictor offers us a descent direction of the unknown function $\frac{1}{2}\|w - w^*\|_H^2$ and the correction produces the new iterate which is more closed to the solution set.

The first ADMM-based prediction-correction method was published by Ye and Yuan [13]. For solving the problem (1.1), we suggest to use the method in Section 3, whose prediction and correction is (3.1) and (3.12), respectively. Even though it needs to compute the step size in the correction step in each iteration, the cost is usually small in comparison with the computational load for solving the subproblem in the prediction step.

Since 2012, we have some new publication about the multi-block problems. First, without correction, the direct extension of the ADMM for the problem (1.1) is not necessarily convergent [2]. For multi-block problems, we suggest to use the methods [7, 8, 9, 11].

References

[1] S. Boyd, N. Parikh, E. Chu, B. Peleato and J. Eckstein, Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers, Nov. 2010.

[2] C.H. Chen, B.S. He, Y.Y. Ye and X. M. Yuan, The direct extension of ADMM for multi-block convex minimization problems is not necessary convergent, Mathematical Programming, 155 (2016) 57-79.

[3] D. Gabay, “Applications of the Method of Multipliers to Variational Inequalities,” in *Augmented Lagrange Methods: Applications to the Solution of Boundary-valued Problems*, M. Fortin and R. Glowinski, eds., North Holland, Amsterdam, The Netherlands, 299-331, 1983.

[4] B.S. He, A class of projection and contraction methods for monotone variational inequalities, Applied Mathematics and Optimization, **35**, pp. 69-76, 1997.

[5] B.S. He, “Inexact implicit methods for monotone general variational inequalities”, Math. Program. Vol. 86, Series A, pp. 199-217, 1999

[6] B.S. He, Parallel splitting augmented Lagrangian methods for monotone structured variational inequalities, Computational Optimization and Applications, 42, 195-212, 2009.

[7] B.S. He, L.S. Hou, and X.M. Yuan, On Full Jacobian Decomposition of the Augmented Lagrangian Method for Separable Convex Programming, SIAM J. Optim., 25 (2015) 2274 − 2312.

[8] B.S. He, M. Tao and X.M. Yuan, Alternating Direction Method with Gaussian Back Substitution for Separable Convex Programming, SIAM J. Optim. 22(2012), 313-340.

[9] B.S. He, M. Tao and X.M. Yuan, A splitting method for separable convex programming, IMA J. Numerical Analysis, 31(2015), 394-426.

[10] B. S. He and M.-H. Xu, A general framework of contraction methods for monotone variational inequalities, Pacific J. Optimization, 4, 195-212, 2008.

[11] B.S. He and X.M. Yuan, Block-wise Alternating Direction Method of Multipliers for Multiple-block Convex Programming and Beyond, SMAI J. Computational Mathematics 1 (2015) 145-174.

[12] R.T. Rockafellar, Monotone operators and the proximal point algorithm, SIAM J. Control Optim., 14 (1976), pp. 877-898.

[13] C. H. Ye and X. M. Yuan, A Descent Method for Structured Monotone Variational Inequalities, Optimization Methods and Software, Vol. 22, pp. 329-338, 2007.