ADMM

Instructor: Jin Zhang

Department of Mathematics
Southern University of Science and Technology

Fall 2023

Contents

Model

$$H_{\text{opt}} = \min \left\{ H(x,z) \equiv h_1(x) + h_2(z) \ : \ Ax + Bz = c \right\}, \qquad (3)$$

where

(1). $h_1 : \mathbb{R}^n \to \overline{\mathbb{R}}$ and $h_1 : \mathbb{R}^n \to \overline{\mathbb{R}}$ are proper closed and convex.

(2). $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{m\times p}$ and $c \in \mathbb{R}^m$.

The Lagrangian associated with problem (3):

$$L(x,z;y) = h_1(x) + h_2(z) + \langle y, Ax + Bz - c\rangle.$$

The dual function of problem (3):

$$\begin{aligned} q(y) &= \inf_{x\in\mathbb{R}^n, z\in\mathbb{R}^p} L(x,z;y) \\ &= \inf_{x\in\mathbb{R}^n, z\in\mathbb{R}^p} \{h_1(x) + h_2(z) + \langle y, Ax + Bz - c\rangle\} \\ &= -h_1^*\left(-A^T y\right) - h_2^*\left(-B^T y\right) - \langle c, y\rangle. \end{aligned}$$

The dual problem of problem (3):

$$q_{\text{opt}} = \sup_{y\in\mathbb{R}^m} \left\{-h_1^*\left(-A^T y\right) - h_2^*\left(-B^T y\right) - \langle c, y\rangle.\right\} \quad \text{(maximization form)}$$

$$q_{\text{opt}} = \inf_{y\in\mathbb{R}^m} \left\{h_1^*\left(-A^T y\right) + h_2^*\left(-B^T y\right) + \langle c, y\rangle.\right\} \quad \text{(minimization form)}$$

Consider the update step:

$$
\begin{aligned}
& y^{k+1} = \underset{y\in\mathbb{R}^m}{\arg\min}\left\{h_1^*\left(-A^T y\right) + h_2^*\left(-B^T y\right) + \langle c, y\rangle + \frac{1}{2\rho}\left\|y - y^k\right\|^2\right\} \\
\overset{*}{\Leftrightarrow} & 0 \in \partial\left(h_1^*\left(-A^T \cdot\right) + h_2^*\left(-B^T \cdot\right) + \langle c, \cdot\rangle + \frac{1}{2\rho}\left\|\cdot - y^k\right\|^2\right)\left(y^{k+1}\right) \\
\overset{**}{\Rightarrow} & 0 \in -A\partial h_1^*\left(-A^T y^{k+1}\right) - B\partial h_2^*\left(-B^T y^{k+1}\right) + c + \frac{1}{\rho}\left(y^{k+1} - y^k\right) \\
\overset{***}{\Leftrightarrow} & \begin{cases} y^{k+1} & = y^k + \rho\left(Ax^{k+1} + Bz^{k+1} - c\right) \\ x^{k+1} & \in \underset{x\in\mathbb{R}^n}{\arg\min}\left\{\langle A^T y^{k+1}, x\rangle + h_1(x)\right\} \\ z^{k+1} & \in \underset{z\in\mathbb{R}^p}{\arg\min}\left\{\langle B^T y^{k+1}, z\rangle + h_2(z)\right\} \end{cases}
\end{aligned}
$$

where

$*$: Fermat's optimality condition

$**$: subdifferential (sum and affine) calculus rules

$***$: conjugate subgradient theorem

Notice that

$$
\begin{aligned}
& x^{k+1} \in \underset{x \in \mathbb{R}^n}{\arg\min} \left\{ \langle A^T y^{k+1}, x \rangle + h_1(x) \right\} \\
\Leftrightarrow & x^{k+1} \in \underset{x \in \mathbb{R}^n}{\arg\min} \left\{ \langle A^T \left(y^k + \rho \left(A x^{k+1} + B z^{k+1} - c \right) \right), x \rangle + h_1(x) \right\} \\
\Leftrightarrow & 0 \in A^T \left(y^k + \rho \left(A x^{k+1} + B z^{k+1} - c \right) \right) + \partial h_1(x^{k+1}) \\
\Leftrightarrow & 0 \in \nabla \left(\frac{\rho}{2} \left\| A \cdot + B z^{k+1} - c + \frac{1}{\rho} y^k \right\|^2 \right) (x^{k+1}) + \partial h_1(x^{k+1})
\end{aligned}
$$

Similarly,

$$
0 \in \nabla \left(\frac{\rho}{2} \left\| A x^{k+1} + B \cdot - c + \frac{1}{\rho} y^k \right\|^2 \right) (z^{k+1}) + \partial h_2(z^{k+1})
$$

Define $f : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}$ and $g : \mathbb{R}^n \times \mathbb{R}^p \to \overline{\mathbb{R}}$ with the values

$$f(x, z) = \frac{\rho}{2} \left\| Ax^{k+1} + B \cdot -c + \frac{1}{\rho} y^k \right\|^2 \text{ for any } x \in \mathbb{R}^n, z \in \mathbb{R}^p$$

$$g(x, z) = \begin{cases} h_1(x) + h_2(z) & \text{for } x \in \mathrm{dom}(h_1)\ z \in \mathrm{dom}(h_2) \\ \infty & \text{otherwise .} \end{cases}$$

Then

$$\begin{aligned} 0 &\in \nabla_x f(x^{k+1}, z^{k+1}) + \partial h_1(x^{k+1}) \\ &= \nabla_x f(x^{k+1}, z^{k+1}) + \partial_x h(x^{k+1}, z^{k+1}). \end{aligned}$$

Similarly,

$$0 \in \nabla_z f(x^{k+1}, z^{k+1}) + \partial_z h(x^{k+1}, z^{k+1}).$$

Hence

$$\begin{aligned}\left(x^{k+1}, z^{k+1}\right) &\in \underset{x\in\mathbb{R}^n, z\in\mathbb{R}^p}{\arg\min}\ f(x,z) + g(x,z) \\ &= \underset{x\in\mathbb{R}^n, z\in\mathbb{R}^p}{\arg\min} \left\{ h_1(x) + h_2(z) + \frac{\rho}{2}\left\| Ax + Bz - c + \frac{1}{\rho} y^k \right\|^2 \right\}.\end{aligned}$$

The Augmented Lagrangian Method:

- Initialization: $y^0 \in \mathbb{R}^m$, $\rho > 0$.
- General step: for any $k = 0, 1, 2, \cdots$ execute the following steps:
  - primal update step:

$$\left(x^{k+1}, z^{k+1}\right) \in \arg\min_{x,z} \left\{ h_1(x) + h_2(z) + \frac{\rho}{2} \left\| Ax + Bz - c + \frac{1}{\rho} y^k \right\|^2 \right\}$$

  - dual update step: $y^{k+1} = y^k + \rho \left(Ax^{k+1} + Bz^{k+1} - c \right)$.

Remark: augmented Lagrangian

The augmented Lagrangian associated with problem (3) and parameter $\rho > 0$ is defined to be

$$L_\rho(x, z; y) = h_1(x) + h_2(z) + \langle y, Ax + Bz - c\rangle + \frac{\rho}{2}\left\|Ax + Bz - c\right\|^2.$$

- $L_0 = L$ is the Lagrangian;
- L_ρ can be considered as a penalized version of the Lagrangian;
- The primal update step can be equivalently written as

$$\left(x^{k+1}, z^{k+1}\right) \in \arg\min_{x,z} L_\rho\left(x, z; y^k\right).$$

The augmented Lagrangian method is in general not an implementable method since the primal update step can be as hard to solve as the original problem.

One source of difficulty is the coupling term between the z and the x variables, which is of the form $\rho\left(x^T A^T B z\right)$.

The approach used in the ADMM to tackle this difficulty is to replace the exact minimization in the primal update step by one iteration of the alternating minimization method.

ADMM:

- Initialization: $x^0 \in \mathbb{R}^n, z^0 \in \mathbb{R}^p, y^0 \in \mathbb{R}^m,\ \rho > 0.$
- General step: for any $k = 0, 1, 2, \cdots$ execute the following steps:

$$(a)\ x^{k+1} \in \arg\min_x \left\{ h_1(x) + \frac{\rho}{2} \left\| Ax + Bz^k - c + \frac{1}{\rho} y^k \right\|^2 \right\}$$

$$(b)\ z^{k+1} \in \arg\min_z \left\{ h_2(x) + \frac{\rho}{2} \left\| Ax^{k+1} + Bz - c + \frac{1}{\rho} y^k \right\|^2 \right\}$$

$$(c)\ y^{k+1} = y^k + \rho \left(Ax^{k+1} + Bz^{k+1} - c \right).$$

2.1 Alternating Direction Proximal Method of Multipliers (AD-PMM)

Let $G \in \mathbb{R}^{n\times n}, Q \in \mathbb{R}^{p\times p}$ be two positive semidefinite matrices. Define $\|\cdot\|_G : \mathbb{R}^n \to \mathbb{R}_+$ and $\|\cdot\|_Q : \mathbb{R}^p \to \mathbb{R}_+$ with the values

$$\|x\|_G = \sqrt{x^T G x} \text{ for any } x \in \mathbb{R}^n;$$
$$\|z\|_Q = \sqrt{z^T Q z} \text{ for any } z \in \mathbb{R}^p.$$

AD-PMM:

- Initialization: $x^0 \in \mathbb{R}^n, z^0 \in \mathbb{R}^p, y^0 \in \mathbb{R}^m,\ \rho > 0$.
- General step: for any $k = 0, 1, 2, \cdots$ execute the following steps:

(a).

$$x^{k+1} \in \arg\min_x \left\{ h_1(x) + \frac{\rho}{2} \left\| Ax + Bz^k - c + \frac{1}{\rho} y^k \right\|^2 + \frac{\left\| x - x^k \right\|_G^2}{2} \right\};$$

(b).

$$z^{k+1} \in \arg\min_z \left\{ h_2(x) + \frac{\rho}{2} \left\| Ax^{k+1} + Bz - c + \frac{1}{\rho} y^k \right\|^2 + \frac{\left\| z - z^k \right\|_Q^2}{2} \right\}$$

(c). $y^{k+1} = y^k + \rho \left(Ax^{k+1} + Bz^{k+1} - c \right)$.

$\|A\|^2$

Choosing $G = \alpha I - A^T A$ with $\alpha \geq \rho\lambda_{\max}(A^T A)$, then the function that needs to be minimized in the step (a) of AD-PMM can be simplified as follows:

确定G的形式

$$
\begin{aligned}
& h_1(x) + \frac{\rho}{2}\left\|Ax + Bz^k - c + \frac{1}{\rho}y^k\right\|^2 + \frac{1}{2}\left\|x - x^k\right\|_G^2 \\
=& h_1(x) + \frac{\rho}{2}\left\|A(x - x^k) + Ax^k + Bz^k - c + \frac{1}{\rho}y^k\right\|^2 + \frac{1}{2}\left\|x - x^k\right\|_G^2 \\
=& h_1(x) + \frac{\rho}{2}\left\|A(x - x^k)\right\|^2 + \langle \rho Ax, Ax^k + Bz^k - c + \frac{1}{\rho}y^k \rangle \\
& + \frac{\alpha}{2}\left\|x - x^k\right\|^2 - \frac{\rho}{2}\left\|A(x - x^k)\right\|^2 + \text{constant} \\
=& h_1(x) + \rho\langle Ax, Ax^k + Bz^k - c + \frac{1}{\rho}y^k \rangle + \frac{\alpha}{2}\left\|x - x^k\right\|^2 + \text{constant.}
\end{aligned}
$$

Then the step (a) of AD-PMM:

$$x^{k+1} = \arg\min_{x\in\mathbb{R}^n} \left\{ h_1(x) + \rho\langle Ax, Ax^k + Bz^k - c + \frac{1}{\rho}y^k\rangle + \frac{\alpha}{2}\left\|x - x^k\right\|^2 \right\}$$
$$= \arg\min_{x\in\mathbb{R}^n} \left\{ \frac{h_1(x)}{\alpha} + \frac{1}{2}\left\|x - \left(x^k - \frac{\rho}{\alpha}A^T\left(Ax^k + Bz^k - c + \frac{1}{\rho}y^k\right)\right)\right\|^2 \right\}$$

That is,

$$x^{k+1} = \mathsf{prox}_{\frac{1}{\alpha}h_1}\left[x^k - \frac{\rho}{\alpha}A^T\left(Ax^k + Bz^k - c + \frac{1}{\rho}y^k\right)\right].$$

Similarly, choose $Q = \beta I - B^T B$ with $\beta \geq \rho\lambda_{\max}(B^T B)$, then the step (b) of AD-PMM:

$$z^{k+1} = \mathsf{prox}_{\frac{1}{\beta}h_2}\left[z^k - \frac{\rho}{\beta}B^T\left(Ax^{k+1} + Bz^k - c + \frac{1}{\rho}y^k\right)\right].$$

AD-LPMM:

- Initialization: $x^0 \in \mathbb{R}^n, z^0 \in \mathbb{R}^p, y^0 \in \mathbb{R}^m$, $\rho > 0$, $\alpha \geq \rho\lambda_{\max}(A^T A)$, $\beta \geq \rho\lambda_{\max}(B^T B)$.
- General step: for any $k = 0, 1, 2, \cdots$ execute the following steps:

(a). $x^{k+1} = \text{prox}_{\frac{1}{\alpha}h_1}\left[x^k - \frac{\rho}{\alpha}A^T\left(Ax^k + Bz^k - c + \frac{1}{\rho}y^k\right)\right]$;

(b). $z^{k+1} = \text{prox}_{\frac{1}{\beta}h_2}\left[z^k - \frac{\rho}{\beta}B^T\left(Ax^{k+1} + Bz^k - c + \frac{1}{\rho}y^k\right)\right]$;

(c). $y^{k+1} = y^k + \rho\left(Ax^{k+1} + Bz^{k+1} - c\right)$.

3. Convergence Analysis of AD-PMM

$$\min\left\{H(x,z) = h_1 + h_2(z)\ :\ Ax + Bz = c\right\} \tag{3}$$

Assumption 4

(A). $h_1 : \mathbb{R}^n \to \overline{\mathbb{R}}$ and $h_2 : \mathbb{R}^p \to \overline{\mathbb{R}}$ are proper closed convex.

(B). $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{m\times p}, c \in \mathbb{R}^m, \rho > 0$.

(C). $G \in \mathbb{R}^{n\times n}, Q \in \mathbb{R}^{p\times p}$ are two positive semidefinite matrices.

(D). For any $a \in \mathbb{R}^n, b \in \mathbb{R}^p$ the optimal sets of the problem

$$\min_{x\in\mathbb{R}^n}\left\{h_1(x) + \frac{\rho}{2}\|Ax\|^2 + \frac{1}{2}\|x\|_G^2 + \langle a, x\rangle\right\}$$

- and

$$\min_{z\in\mathbb{R}^p}\left\{h_2(x)+\frac{\rho}{2}\left\|Bz\right\|^2+\frac{1}{2}\left\|z\right\|_Q^2+\langle b,z\rangle\right\}$$

are nonempty. (This guarantees that the AD-PMM method is actually a well-defined method.)

(E). There exists $\hat{x}\in \mathrm{ri}\,(\mathrm{dom}(h_1))$ and $\tilde{z}\in \mathrm{ri}\,(\mathrm{dom}(h_2))$ for which $A\hat{x}+B\tilde{z}=c$.

(F). Problem (3) has a nonempty optimal set, denoted by X^*, and the corresponding optimal value is H_{opt}.

Theorem (strong duality for problem (3) and it's dual

Suppose that Assumption 4 holds, and let $H_{\text{opt}}, q_{\text{opt}}$ be the optimal values of problem (3)(primal problem) and it's dual, respectively. Then, $H_{\text{opt}} = q_{\text{opt}}$ and the dual of problem (3) possesses an optimal solution.

Theorem ($O\left(\frac{1}{k}\right)$ rate of convergence of AD-PMM)

Suppose that Assumption 4 holds. Let $\{(x^k, z^k)\}_{k\geq 0}$ be the sequence generated by AD-PMM for solving problem (3). Let (x^*, z^*) be an optimal solution of problem (3) and y^* be an optimal solution of the dual of problem (3). Suppose that $\gamma > 0$ is any constant satisfying $\gamma \geq 2\|y^*\|$. Then for all $n \geq 0$,

$$H\left(x^{(n)}, z^{(n)}\right) - H_{\text{opt}} \leq \frac{\left\|x^* - x^0\right\|_G^2 + \left\|z^* - z^0\right\|_C^2 + \frac{1}{\rho}\left(\gamma + \left\|y^0\right\|\right)^2}{2(n+1)},$$

$$\left\|Ax^{(n)} + Bz^{(n)} - c\right\| \leq \frac{\left\|x^* - x^0\right\|_G^2 + \left\|z^* - z^0\right\|_C^2 + \frac{1}{\rho}\left(\gamma + \left\|y^0\right\|\right)^2}{\gamma(n+1)},$$

where $C = \gamma B^T B + Q$ and

$$x^{(n)} = \frac{1}{n+1}\sum_{k=0}^{n} x^{k+1}, z^{(n)} = \frac{1}{n+1}\sum_{k=0}^{n} z^{k+1}.$$

Proof: By steps (a) and steps (b) of AD-PMM,

$$-\rho A^T\left(Ax^{k+1}+Bz^k-c+\frac{1}{\rho}y^k\right)-G(x^{k+1}-x^k)\in\partial h_1(x^{k+1})$$

$$-\rho A^T\left(Ax^{k+1}+Bz^{k+1}-c+\frac{1}{\rho}y^k\right)-Q(z^{k+1}-z^k)\in\partial h_2(z^{k+1})$$

Use the notations:

$$\tilde{x}^k = x^{k+1}; \quad \tilde{z}^k = z^{k+1}, \quad \tilde{y}^k = y^k + \rho\left(Ax^{k+1} + Bz^k - c\right).$$

Then

$$h_1(x) - h_1(\tilde{x}^k) + \left\langle \rho A^T\left(A\tilde{x}^k + Bz^k - c + \frac{1}{\rho}y^k\right) + G(\tilde{x}^k - x^k), x - \tilde{x}^k \right\rangle \geq 0;$$
$$h_2(z) - h_2(\tilde{z}^k) + \left\langle \rho B^T\left(A\tilde{x}^k + B\tilde{z}^k - c + \frac{1}{\rho}y^k\right) + Q(\tilde{z}^k - z^k), z - \tilde{z}^k \right\rangle \geq 0.$$

Use the notations:

$$\tilde{y}^k = y^k + \rho\left(Ax^{k+1} + Bz^k - c\right).$$

Then

$$h_1(x) - h_1(\tilde{x}^k) + \left\langle A^T\tilde{y}^k + G(\tilde{x}^k - x^k), x - \tilde{x}^k \right\rangle \geq 0;$$
$$h_2(z) - h_2(\tilde{z}^k) + \left\langle B^T\tilde{y}^k + \left(\rho B^T B + Q\right)(\tilde{z}^k - z^k), z - \tilde{z}^k \right\rangle \geq 0.$$

Adding the above two inequalities and using the identity

$$y^{k+1} = y^k + \rho\left(A\tilde{x}^k + B\tilde{z}^k - c\right),$$

we can conclude that

$$H(x,z) - H(\tilde{x}^k, \tilde{z}^k) + \left\langle \begin{pmatrix} x - \tilde{x}^k \\ z - \tilde{z}^k \\ y - \tilde{y}^k \end{pmatrix}, \begin{pmatrix} A^T\tilde{y}^k \\ B^T\tilde{y}^k, \\ -A\tilde{x}^k - B\tilde{z}^k + c \end{pmatrix} - \begin{pmatrix} G(x^k - \tilde{x}^k) \\ C(z^k - \tilde{z}^k) \\ \frac{y^k - y^{k+1}}{\rho} \end{pmatrix} \right\rangle \geq 0$$

where $C = \rho B^T B + Q$.

Use the identity for any positive semidefinite matrix P:

$$(a-b)^T P(c-d) = \frac{1}{2}\left(\|a-d\|_P^2 - \|a-c\|_P^2 + \|b-c\|_P^2 - \|b-d\|_P^2\right).$$

Then

$$\begin{aligned}
(x-\tilde{x}^k)^T G(x^k-\tilde{x}^k) &\geq \frac{1}{2}\left(\left\|x-\tilde{x}^k\right\|_G^2 - \left\|x-x^k\right\|_G^2 + \left\|\tilde{x}^k - x^k\right\|_G^2\right); \\
&\geq \frac{1}{2}\left\|x-\tilde{x}^k\right\|_G^2 - \frac{1}{2}\left\|x-x^k\right\|_G^2 \\
(z-\tilde{z}^k)^T Q(z^k-\tilde{z}^k) &\geq \frac{1}{2}\left(\left\|z-\tilde{z}^k\right\|_C^2 - \left\|z-z^k\right\|_C^2 + \left\|z^k-\tilde{z}^k\right\|_C^2\right); \\
\frac{1}{\rho}(y-\tilde{y}^k)^T(y^k-y^{k+1}) &\geq \frac{1}{2\rho}\left(\left\|y-y^{k+1}\right\|^2 - \left\|y-y^k\right\|^2\right) - \frac{1}{2}\left\|B(z^k-\tilde{z}^k)\right\|^2.
\end{aligned}$$

Denote

$$H = \begin{pmatrix} G & 0 & 0 \\ 0 & C & 0 \\ 0 & 0 & \frac{1}{\rho} I \end{pmatrix}, \; w = \begin{pmatrix} x \\ z \\ y \end{pmatrix}, \; w^k = \begin{pmatrix} x^k \\ z^k \\ y^k \end{pmatrix}, \; \tilde{w}^k = \begin{pmatrix} \tilde{x}^k \\ \tilde{z}^k \\ \tilde{y}^k \end{pmatrix}.$$

We obtain

$$\begin{aligned} \left\langle \begin{pmatrix} x - \tilde{x}^k \\ z - \tilde{z}^k \\ y - \tilde{y}^k \end{pmatrix}, \begin{pmatrix} G(x^k - \tilde{x}^k) \\ C(z^k - \tilde{z}^k) \\ \frac{y^k - y^{k+1}}{\rho} \end{pmatrix} \right\rangle \geq & \frac{1}{2} \left\| w - w^{k+1} \right\|_H^2 - \left\| w - w^k \right\|_H^2 \\ & + \frac{1}{2} \left\| z^k - \tilde{z}^k \right\|_C^2 - \frac{\rho}{2} \left\| B(z^k - \tilde{z}^k) \right\|^2 \\ \geq & \frac{1}{2} \left\| w - w^{k+1} \right\|_H^2 - \left\| w - w^k \right\|_H^2 . \end{aligned}$$

Then

$$H(x,z) - H(\tilde{x}^k, \tilde{z}^k) + \langle w - \tilde{w}^k, F\tilde{w}^k + \tilde{c} \rangle \geq \frac{1}{2} \left\| w - w^{k+1} \right\|_H^2 - \left\| w - w^k \right\|_H^2$$

where $F = \begin{pmatrix} 0 & 0 & A^T \\ 0 & 0 & B^T \\ -A & -B & 0 \end{pmatrix}$, $\tilde{c} = \begin{pmatrix} 0 \\ 0 \\ c \end{pmatrix}$.

F is skew symmetric ($F = -F^T$) implies that

$$\langle w - \tilde{w}^k, F(\tilde{w}^k - w) \rangle = \langle w - \tilde{w}^k, F^T(w - \tilde{w}^k) \rangle = \langle F(w - \tilde{w}^k), w - \tilde{w}^k \rangle,$$

and then $\langle w - \tilde{w}^k, F(\tilde{w}^k - w) \rangle = 0$,
and then
$\langle w - \tilde{w}^k, F\tilde{w}^k + \tilde{c} \rangle = \langle w - \tilde{w}^k, F(\tilde{w}^k - w) + Fw + \tilde{c} \rangle = \langle w - \tilde{w}^k, Fw + \tilde{c} \rangle$
and then

$$H(x,z) - H(\tilde{x}^k, \tilde{z}^k) + \langle w - \tilde{w}^k, Fw + \tilde{c} \rangle \geq \frac{1}{2} \left\| w - w^{k+1} \right\|_H^2 - \frac{1}{2} \left\| w - w^k \right\|_H^2,$$

Summing the above inequality over $k = 0, 1, \cdots, n$ yields

$$(n+1)H(x,z) - \sum_{k=0}^{n} H(\tilde{x}^k, \tilde{z}^k) + \left\langle (n+1)w - \sum_{k=0}^{n} \tilde{w}^k, Fw + \tilde{c} \right\rangle \geq -\frac{1}{2} \left\| w - w^0 \right\|_H^2$$

Then

$$H(x^{(n)}, z^{(n)}) - H(x,z) + \langle w^{(n)} - w, Fw + \tilde{c} \rangle \leq \frac{1}{2(n+1)} \left\| w - w^0 \right\|_H^2 .$$

Use again the skew-symmetry of F, then

$$H(x^{(n)}, z^{(n)}) - H(x,z) + \langle w^{(n)} - w, Fw^{(n)} + \tilde{c} \rangle \leq \frac{1}{2(n+1)} \left\| w - w^0 \right\|_H^2 .$$

Then

$$
\begin{aligned}
& H(x^{(n)}, z^{(n)}) - H(x, z) \\
& \quad + \langle x^{(n)} - x, A^T y^{(n)} \rangle + \langle z^{(n)} - z, B^T y^{(n)} \rangle + \langle y^{(n)} - y, -Ax^{(n)} - Bz^{(n)} + c \rangle \\
\leq & \frac{1}{2(n+1)} \left(\left\| x - x^0 \right\|_G^2 + \left\| z - z^0 \right\|_C^2 + \frac{1}{\rho} \left\| y - y^0 \right\|^2 \right)
\end{aligned}
$$

Then

$$
\begin{aligned}
& H(x^{(n)}, z^{(n)}) - H(x, z) \\
& \quad + \langle y, Ax^{(n)} + Bz^{(n)} - c \rangle - \langle y^{(n)}, Ax + By - c \rangle \\
\leq & \frac{1}{2(n+1)} \left(\left\| x - x^0 \right\|_G^2 + \left\| z - z^0 \right\|_C^2 + \frac{1}{\rho} \left\| y - y^0 \right\|^2 \right)
\end{aligned}
$$

Let (x^*, z^*) be an optimal solution of problem (3). Then $H(x^*, z^*) = H_{\text{opt}}$ and $Ax^* + By^* = c$. Choose $x = x^*$ and $y = y^*$, we obtain

$$
\begin{aligned}
& H(x^{(n)}, z^{(n)}) - H_{\text{opt}} + \langle y, Ax^{(n)} + Bz^{(n)} - c\rangle \\
\leq & \frac{\left\|x^* - x^0\right\|_G^2 + \left\|z^* - z^0\right\|_C^2 + \frac{1}{\rho}\left\|y - y^0\right\|^2}{2(n+1)}.
\end{aligned}
$$

Since the above inequality holds for any $y \in \mathbb{R}^m$, we choose $y = \gamma \frac{Ax^{(n)} + Bz^{(n)} - c}{\left\|Ax^{(n)} + Bz^{(n)} - c\right\|}$ if $Ax^{(n)} + Bz^{(n)} - c \neq 0$ and $y = 0$ otherwise. Then

$$
\begin{aligned}
& H(x^{(n)}, z^{(n)}) - H_{\text{opt}} + \gamma \left\|Ax^{(n)} + Bz^{(n)} - c\right\| \\
\leq & \frac{\left\|x^* - x^0\right\|_G^2 + \left\|z^* - z^0\right\|_C^2 + \frac{1}{\rho}\left\|y - y^0\right\|^2}{2(n+1)} \\
\leq & \frac{\left\|x^* - x^0\right\|_G^2 + \left\|z^* - z^0\right\|_C^2 + \frac{1}{\rho}\left(\gamma + \left\|y^0\right\|\right)^2}{2(n+1)}
\end{aligned}
$$

Firstly, we have

$$H(x^{(n)}, z^{(n)}) - H_{\text{opt}} \leq \frac{\left\|x^* - x^0\right\|_G^2 + \left\|z^* - z^0\right\|_C^2 + \frac{1}{\rho}\left(\gamma + \left\|y^0\right\|\right)^2}{2(n+1)}$$

Second, recall that the value function associated with problem (3) is

$$v(t) = \min\left\{H(x,z) : Ax + Bz = c + t\right\},$$

v is convex and $y^* \in \partial v(0)$. Then

1. v is convex;
2. $-y^* \in \partial v(0)$;
3. $v(Ax^{(n)} + Bz^{(n)} - c) - v(0) \geq \langle -y^*, Ax^{(n)} + Bz^{(n)} - c\rangle$;
4. $\langle -y^*, Ax^{(n)} + Bz^{(n)} - c\rangle \geq -\frac{\gamma}{2}\left\|Ax^{(n)} + Bz^{(n)} - c\right\|$
5. $v(Ax^{(n)} + Bz^{(n)} - c) \leq H(x^{(n)}, z^{(n)})$ and $v(0) = H_{\text{opt}}$

Then $H(x^{(n)}, z^{(n)}) - H_{\text{opt}} \geq -\frac{\gamma}{2}\left\|Ax^{(n)} + Bz^{(n)} - c\right\|$.

Then

$$\left\| Ax^{(n)} + Bz^{(n)} - c \right\| \leq \frac{\left\| x^* - x^0 \right\|_G^2 + \left\| z^* - z^0 \right\|_C^2 + \frac{1}{\rho}\left(\gamma + \left\| y^0 \right\|\right)^2}{\gamma(n+1)}.$$

Model

$$\min_{x \in \mathbb{R}^n} f_1(x) + f_2(Ax), \tag{4}$$

where $f_1 : \mathbb{R}^n \to \overline{\mathbb{R}}$, $f_2 : \mathbb{R}^m \to \overline{\mathbb{R}}$ are proper closed convex and $A \in \mathbb{R}^{m \times n}$.

Problem (4) can rewritten as

$$\begin{aligned} \min_{x \in \mathbb{R}^n, z \in \mathbb{R}^m} \quad & h_1(x) + h_2(z) \\ \text{s.t.} \quad & Ax + Bz = c \end{aligned} \tag{4.1}$$

where $h_1 = f_1, h_2 = f_2, B = -I$ and $c = 0$.

The step (b) of ADMM for solving problem (4.1):

$$
\begin{aligned}
z^{k+1} &= \arg\min_z \left\{ h_2(x) + \frac{\rho}{2} \left\| Ax^{k+1} + Bz - c + \frac{1}{\rho} y^k \right\|^2 \right\} \\
&= \arg\min_z \left[f_2(z) + \frac{\rho}{2} \left\| Ax^{k+1} - z + \frac{1}{\rho} y^k \right\|^2 \right] \\
&= \mathsf{prox}_{\frac{1}{\rho} f_2} \left(Ax^{k+1} + \frac{1}{\rho} y^k \right).
\end{aligned}
$$

Algorithm 1 (ADMM for solving (4.1))

- Initialization: $x^0 \in \mathbb{R}^n, z^0, y^0 \in \mathbb{R}^m$, $\rho > 0$.
- General step: for any $k = 0, 1, 2, \cdots$ execute the following steps:

  (a). $x^{k+1} \in \arg\min_x \left\{ f_1(x) + \frac{\rho}{2} \left\| Ax - z^k + \frac{1}{\rho} y^k \right\|^2 \right\}$;

  (b). $z^{k+1} = \text{prox}_{\frac{1}{\rho} f_2} \left(Ax^{k+1} + \frac{1}{\rho} y^k \right)$;

  (c). $y^{k+1} = y^k + \rho \left(Ax^{k+1} - z^{k+1} \right)$.

Remark: Step (a) of Algorithm 1 might be difficult to compute since the minimization in step (a) is more involved than a prox computation due to the quadratic term $\frac{\rho}{2} x^T A^T Ax$.

Problem (4) can also be rewritten as

$$\begin{aligned}\min_{x, w \in \mathbb{R}^n, z \in \mathbb{R}^m} \; & f_1(w) + f_2(z) \\ \text{s.t.} \quad & \begin{pmatrix} A \\ I \end{pmatrix} x - \begin{pmatrix} z \\ w \end{pmatrix} = 0. \end{aligned} \tag{4.2}$$

Problem (4.2) fits model (3):

$$\min_{x \in \mathbb{R}^n, \tilde{z} \in \mathbb{R}^{m+n}} \left\{ h_1(x) + h_2(\tilde{z}) \,|\, \tilde{A}x + B\tilde{z} = c \right\}.$$

where $h_1 : \mathbb{R}^n \to \overline{\mathbb{R}}$ with the values $h_1(x) = 0$;
$h_2 : \mathbb{R}^{m+n} \to \overline{\mathbb{R}}$ with the values $h_2(\tilde{z}) = f_1(w) + f_2(z)$, where $\tilde{z} = \binom{z}{w}$.
$\tilde{A} = \binom{A}{I}$, $B = -I$, $c = 0$.

Notice that the dual vector $y \in \mathbb{R}^{m+n}$ is of the form $y = \binom{y_1}{y_2}$, where $y_1 \in \mathbb{R}^m$, $y_2 \in \mathbb{R}^n$.

The step (a) of ADMM for solving problem (4.2):

$$
\begin{aligned}
x^{k+1} &\in \arg\min_{x} \left\{ h_1(x) + \frac{\rho}{2} \left\| \tilde{A}x + B\tilde{z}^k - c + \frac{1}{\rho} y^k \right\|^2 \right\} \\
&= \arg\min_{x \in \mathbb{R}^n} \left[\left\| Ax - z^k + \frac{1}{\rho} y_1^k \right\|^2 + \left\| x - w^k + \frac{1}{\rho} y_2^k \right\|^2 \right] \\
&= \left(I + A^T A\right)^{-1} \left(A^T \left[z^k - \frac{1}{\rho} y_1^k \right] + w^k - \frac{1}{\rho} y_2^k \right).
\end{aligned}
$$

The step (b) of ADMM for solving problem (4.2):

$$\begin{aligned}
\begin{pmatrix} z^{k+1} \\ w^{k+1} \end{pmatrix} &= \tilde{z}^{k+1} \\
&\in \arg\min_{\tilde{z}} \left\{ h_2(x) + \frac{\rho}{2} \left\| \tilde{A} x^{k+1} + B\tilde{z} - c + \frac{1}{\rho} y^k \right\|^2 \right\} \\
&= \begin{pmatrix} \arg\min_{z} \left[f_2(z) + \frac{\rho}{2} \left\| A x^{k+1} - z + \frac{1}{\rho} y_1^k \right\|^2 \right] \\ \arg\min_{w} \left[f_1(w) + \frac{\rho}{2} \left\| x^{k+1} - z + \frac{1}{\rho} y_2^k \right\|^2 \right] \end{pmatrix} \\
&= \begin{pmatrix} \mathrm{prox}_{\frac{1}{\rho} f_2} \left(A x^{k+1} + \frac{1}{\rho} y_1^k \right) \\ \mathrm{prox}_{\frac{1}{\rho} f_1} \left(x^{k+1} + \frac{1}{\rho} y_2^k \right) \end{pmatrix}
\end{aligned}$$

The step (c) of ADMM for solving problem (4.2):

$$
\begin{aligned}
\begin{pmatrix} y_1^{k+1} \\ y_2^{k+1} \end{pmatrix} &= y^{k+1} \\
&= y^k + \rho \left(\tilde{A} x^{k+1} + B \tilde{z}^{k+1} - c \right) \\
&= \begin{pmatrix} y_1^k \\ y_2^k \end{pmatrix} + \rho \left(\begin{pmatrix} A x^{k+1} \\ x^{k+1} \end{pmatrix} - \begin{pmatrix} z^{k+1} \\ w^{k+1} \end{pmatrix} \right)
\end{aligned}
$$

Algorithm 2 (ADMM for solving (4.2))

- Initialization: $x^0, w^0, y_2^0 \in \mathbb{R}^n, z^0, y_1^0 \in \mathbb{R}^m$, $\rho > 0$.
- General step: for any $k = 0, 1, 2, \cdots$ execute the following steps:

(a). $x^{k+1} = (I + A^T A)^{-1} \left(A^T \left[z^k - \frac{1}{\rho} y_1^k \right] + w^k - \frac{1}{\rho} y_2^k \right);$

(b). $z^{k+1} = \text{prox}_{\frac{1}{\rho} f_2} \left(Ax^{k+1} + \frac{1}{\rho} y_1^k \right);$

$w^{k+1} = \text{prox}_{\frac{1}{\rho} f_1} \left(x^{k+1} + \frac{1}{\rho} y_2^k \right);$

(c). $y_1^{k+1} = y_1^k + \rho \left(Ax^{k+1} - z^{k+1} \right);$

$y_2^{k+1} = y_2^k + \rho \left(x^{k+1} - w^{k+1} \right).$

Remark: Algorithm 2 might still be too computationally demanding since it involves the evaluation of the inverse of $I + A^T A$, which might be a difficult task in large- scale problems.

Algorithm 3 (AD-LPMM for solving (4.1))

- Initialization: $x^0 \in \mathbb{R}^n, z^0, y^0 \in \mathbb{R}^m$, $\rho > 0$, $\alpha \geq \rho\lambda_{\max}(A^T A), \beta \geq \rho$.
- General step: for any $k = 0, 1, 2, \cdots$ execute the following steps:

(a). $x^{k+1} = \text{prox}_{\frac{1}{\alpha} f_1}\left[x^k + \frac{\rho}{\alpha} A^T \left(Ax^k - z^k + \frac{1}{\rho} y^k\right)\right];$

(b). $z^{k+1} = \text{prox}_{\frac{1}{\beta} f_2}\left[z^k + \frac{\rho}{\beta} \left(Ax^{k+1} - z^k + \frac{1}{\rho} y^k\right)\right];$

(c). $y^{k+1} = y^k + \rho\left(Ax^{k+1} - z^{k+1}\right).$

Remark: The above scheme has the advantage that it only requires simple linear algebra operations (no more than matrix/vector multiplications) and prox evaluations.

Example (l_1-regularized least squares).

Consider the problem

$$\min_{x \in \mathbb{R}^n} \left\{ \frac{1}{2} \|Ax - b\|_2^2 + \lambda \|x\|_1 \right\},$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ and $\lambda > 0$, which fits the composite model with $f_1(x) = \lambda \|x\|_1$ and $f_2(y) = \frac{1}{2} \|y - b\|_2^2$.

The general step of Algorithm 1:

(a). $x^{k+1} \in \arg\min_x \left\{ \lambda \|x\|_1 + \frac{\rho}{2} \left\| Ax - z^k + \frac{1}{\rho} y^k \right\|^2 \right\}$;

(b). $z^{k+1} = \frac{\rho A x^{k+1} + y^k + b}{1+\rho}$;

(c). $y^{k+1} = y^k + \rho \left(A x^{k+1} - z^{k+1} \right)$.

Remark: this algorithm is completely useless since it suggests to solve an l_1-regularized least squares problem by a sequence of l_1-regularized least squares problems.

The general step of Algorithm 2:

(a). $x^{k+1} = \left(I + A^T A\right)^{-1} \left(A^T \left[z^k - \frac{1}{\rho} y_1^k \right] + w^k - \frac{1}{\rho} y_2^k \right);$

(b). $z^{k+1} = \frac{\rho A x^{k+1} + y_1^k + b}{1+\rho};$

$w^{k+1} = \mathcal{T}_{\frac{\lambda}{\rho}} \left(x^{k+1} + \frac{1}{\rho} y_2^k \right);$

(c). $y_1^{k+1} = y_1^k + \rho \left(A x^{k+1} - z^{k+1} \right);$

$y_2^{k+1} = y_2^k + \rho \left(x^{k+1} - w^{k+1} \right).$

Remark: this algorithm will require to compute the inverse of $I + A^T A$ in a preprocess. These operations might be difficult to execute in large-scale problems

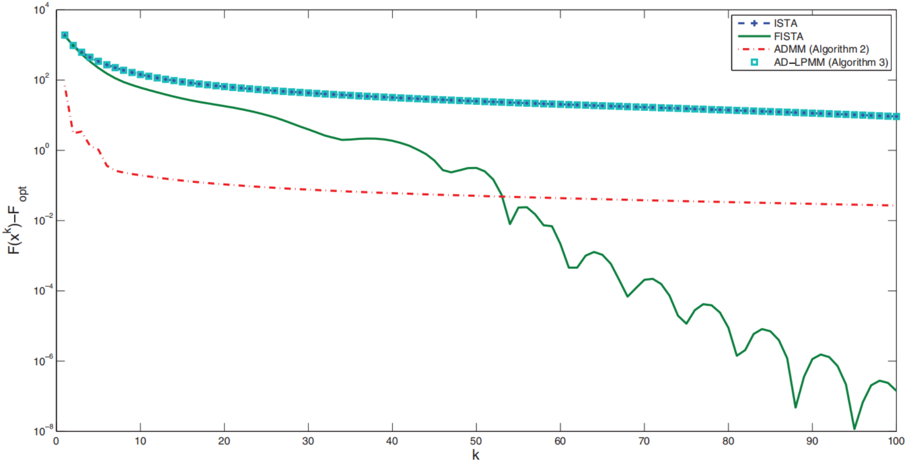
The general step of Algorithm 3 with $\alpha = \lambda_{\max}\left(A^T A\right)\rho$ and $\beta = \rho$:

(a). $x^{k+1} = \mathcal{T}_{\frac{\lambda}{\rho\lambda_{\max}(A^T A)}}\left[x^k - \frac{1}{\lambda_{\max}(A^T A)}\left(Ax^k - z^k + \frac{1}{\rho}y^k\right)\right]$;

(b). $z^{k+1} = \frac{\rho Ax^{k+1} + y^k + b}{1+\rho}$;

(c). $y^{k+1} = y^k + \rho\left(Ax^{k+1} - z^{k+1}\right)$.

Remark: The dominant computations in this algorithm are matrix/vector multiplications.



ISTA
FISTA
ADMM (Algorithm 2)
AD–LPMM (Algorithm 3)
$F(x^k)-F_{opt}$
k
10^4
10^2
10^0
10^{-2}
10^{-4}
10^{-6}
10^{-8}
0
10
20
30
40
50
60
70
80
90
100

Remark

1. ISTA and AD-LPMM (Algorithm 3) exhibit the same performance, while ADMM (Algorithm 2) seems to outperform both of them. This is actually not surprising since the computations carried out at each iteration of ADMM (solution of linear systems) are much heavier than the computations per iteration of AD-LPMM and ISTA (matrix/vector multiplications).
   In that respect, the comparison is in fact not fair and biased in favor of ADMM.
2. FISTA significantly outperforms ADMM starting from approximately 50 iterations despite the fact that it is a simpler algorithm that requires substantially less computational effort per iteration.
   One possible reason is that FISTA is a method with a provably $O(1/k^2)$ rate of convergence in function values, while ADMM is only guaranteed to converge at a rate of $O(1/k)$.

Example (robust regression)

Consider the problem

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_1 ,$$

where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$, which fits the composite model with $f_1(x) = 0$ and $f_2(y) = \|y - b\|_1$.

The general step of Algorithm 1:

(a). $x^{k+1} = \arg\min_x \left\| Ax - z^k + \frac{1}{\rho} y^k \right\|^2$;

(b). $z^{k+1} = \mathcal{T}_{\frac{1}{\rho}} \left(Ax^{k+1} + \frac{1}{\rho} y^k - b \right) + b$;

(c). $y^{k+1} = y^k + \rho \left(Ax^{k+1} - z^{k+1} \right)$.

The general step of Algorithm 2:

(a). $x^{k+1} = \left(I + A^T A\right)^{-1} \left(A^T \left[z^k - \frac{1}{\rho} y_1^k\right] + w^k - \frac{1}{\rho} y_2^k\right);$

(b). $z^{k+1} = \mathcal{T}_{\frac{1}{\rho}} \left(Ax^{k+1} + \frac{1}{\rho} y_1^k - b\right) + b;$

$w^{k+1} = x^{k+1} + \frac{1}{\rho} y_2^k;$

(c). $y_1^{k+1} = y_1^k + \rho \left(Ax^{k+1} - z^{k+1}\right);$

$y_2^{k+1} = y_2^k + \rho \left(x^{k+1} - w^{k+1}\right).$

The general step of Algorithm 2 with $y_2^0 = 0$:

(a). $x^{k+1} = \left(I + A^T A\right)^{-1} \left(A^T \left[z^k - \frac{1}{\rho} y_1^k\right] + x^k\right);$

(b). $z^{k+1} = \mathcal{T}_{\frac{1}{\rho}} \left(Ax^{k+1} + \frac{1}{\rho} y_1^k - b\right) + b;$

(c). $y_1^{k+1} = y_1^k + \rho \left(Ax^{k+1} - z^{k+1}\right).$

The general step of Algorithm 3 with $\alpha = \lambda_{\max}\left(A^T A\right)\rho$ and $\beta = \rho$:

(a). $x^{k+1} = x^k - \frac{1}{\lambda_{\max}(A^T A)}\left(Ax^k - z^k + \frac{1}{\rho}y^k\right)$;

(b). $z^{k+1} = \mathcal{T}_{\frac{1}{\rho}}\left(Ax^{k+1} + \frac{1}{\rho}y^k - b\right) + b$;

(c). $y^{k+1} = y_1^k + \rho\left(Ax^{k+1} - z^{k+1}\right)$.

Example (basis pursuit)

Consider the problem

$$\min \|x\|_1, \quad \text{s.t. } Ax = b,$$

where $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$, which fits the composite model with $f_1(x) = \|x\|_1$ and $f_2(y) = \delta_{\{b\}}(y)$.

The general step of Algorithm 1 with $z^0 = b$:

(a). $x^{k+1} \in \arg\min_x \left\{ \|x\|_1 + \frac{\rho}{2} \left\| Ax - b + \frac{1}{\rho} y^k \right\|^2 \right\}$;

(b). $y^{k+1} = y^k + \rho \left(Ax^{k+1} - b \right)$.

Remark: Algorithm 1 is actually not particularly implementable since its first update step does not seem to be simpler to solve than the original problem.

The general step of Algorithm 2 with $z^0 = b$:

(a). $x^{k+1} = \left(I + A^T A\right)^{-1} \left(A^T \left[b - \frac{1}{\rho} y_1^k\right] + w^k - \frac{1}{\rho} y_2^k\right)$;

(b). $w^{k+1} = \mathcal{T}_{\frac{1}{\rho}} \left(x^{k+1} + \frac{1}{\rho} y_2^k\right)$;

(c). $y_1^{k+1} = y_1^k + \rho \left(Ax^{k+1} - b\right)$,
$y_2^{k+1} = y_2^k + \rho \left(x^{k+1} - w^{k+1}\right)$.

The general step of Algorithm 3 with $\alpha = \lambda_{\max}\left(A^T A\right)\rho$, $\beta = \rho$ and $z^0 = b$:

(a). $x^{k+1} = \mathcal{T}_{\frac{\lambda}{\rho\lambda_{\max}(A^T A)}}\left[x^k - \frac{1}{\lambda_{\max}(A^T A)}\left(Ax^k - b + \frac{1}{\rho}y^k\right)\right]$;

(b). $y^{k+1} = y^k + \rho\left(Ax^{k+1} - b\right)$.

Example (minimizing $\sum_{i=1}^{p} g_i(A_i x)$)

Consider the problem

$$\min_{x\in\mathbb{R}^n} \sum_{i=1}^{p} g_i\left(A_i x\right),$$

where $g_1,\cdots,g_p$ are proper closed convex and $A_i \in \mathbb{R}^{m\times n}$ for all $i=1,\cdots,p$, which fits the composite model with $f_1(x) = \|x\|_1$, $f_2(y) = \sum_{i=1}^{p} g_i(y_i)$ and $A = \left(A_1^T,\cdots,A_p^T\right)^T$.

The general step of Algorithm 1:

(a). $x^{k+1} \in \arg\min\limits_{x} \sum_{i=1}^{p} \left\| A_i x - z_i^k + \frac{1}{\rho} y_i^k \right\|^2$;

(b). $z_i^{k+1} = \mathsf{prox}_{\frac{1}{\rho} g_i} \left(A_i x^{k+1} + \frac{1}{\rho} y_i^k \right)$;

(c). $y_i^{k+1} = y_i^k + \rho \left(A_i x^{k+1} - z_i^{k+1} \right)$.

The general step of Algorithm 1 with rank$(A) = n$:

(a). $x^{k+1} = \left(\sum_{i=1}^{p} A_i^T A_i \right)^{-1} \sum_{i=1}^{p} A_i^T \left(z_i^k - \frac{1}{\rho} y_i^k \right)$;

(b). $z_i^{k+1} = \mathsf{prox}_{\frac{1}{\rho} g_i} \left(A_i x^{k+1} + \frac{1}{\rho} y_i^k \right)$;

(c). $y_i^{k+1} = y_i^k + \rho \left(A_i x^{k+1} - z_i^{k+1} \right)$.

The general step of Algorithm 2:

(a). $x^{k+1} = \left(I + \sum_{i=1}^{p} A_i^T A_i\right)^{-1} \left(\sum_{i=1}^{p} A_i^T \left[z_i^k - \frac{1}{\rho} y_{1,i}^k\right] + w^k - \frac{1}{\rho} y_2^k\right);$

(b). $z_i^{k+1} = \text{prox}_{\frac{1}{\rho} g_i}\left(A_i x^{k+1} + \frac{1}{\rho} y_{1,i}^k\right);$

$w_i^{k+1} = x_i^{k+1} + \frac{1}{\rho} y_{2,i}^k;$

(c). $y_{1,i}^{k+1} = y_{1,i}^k + \rho\left(A_i x^{k+1} - z_i^{k+1}\right);$

$y_{2,i}^{k+1} = y_{2,i}^k + \rho\left(x^{k+1} - w_i^{k+1}\right).$

The general step of Algorithm 2 with $y_{2,i}^0 = 0$:

(a). $x^{k+1} = \left(I + \sum_{i=1}^{p} A_i^T A_i\right)^{-1} \left(\sum_{i=1}^{p} A_i^T \left[z_i^k - \frac{1}{\rho} y_{1,i}^k\right] + x^k\right);$

(b). $z_i^{k+1} = \text{prox}_{\frac{1}{\rho} g_i}\left(A_i x^{k+1} + \frac{1}{\rho} y_{1,i}^k\right);$

(c). $y_{1,i}^{k+1} = y_{1,i}^k + \rho\left(A_i x^{k+1} - z_i^{k+1}\right).$

Remark: Algorithm 2 is not simpler than Algorithm 1.

The general step of Algorithm 3 with $\alpha = \lambda_{\max}\left(\sum_{i=1}^{p} A_i^T A_i\right)\rho$ and $\beta = \rho$:

(a). $x^{k+1} = x^k - \frac{1}{\lambda_{\max}(\sum_{i=1}^{p} A_i^T A_i)} \sum_{i=1}^{p} A_i^T \left(A_i x^k - z_i^k + \frac{1}{\rho} y_i^k\right)$;

(b). $z_i^{k+1} = \mathrm{prox}_{\frac{1}{\rho} g_i}\left(A_i x^{k+1} + \frac{1}{\rho} y_i^k\right)$;

(c). $y_i^{k+1} = y_1^k + \rho\left(A_i x^{k+1} - z_i^{k+1}\right)$.