

ADMM for Multi-Block Problems

Hailiang ZHAO @ ZJU-CS

<http://hliangzhao.me>

October 24, 2022

Outline

I Direction Extension of ADMM

II Gaussian Back Substitution ADMM-GBS

III Prediction-Correction ADMM-PC

IV Linearized ADMM with Parallel Splitting LADMM-PS LADMM-SPU

References

Outline

I Direction Extension of ADMM

II Gaussian Back Substitution ADMM-GBS

III Prediction-Correction ADMM-PC

IV Linearized ADMM with Parallel Splitting LADMM-PS LADMM-SPU

References

Multi-Block Problem

We try to extend ADMM to solve the following problem with multi-blocks:

$$\min_{\mathbf{x}_i} \sum_{i=1}^m f_i(\mathbf{x}_i), \quad s.t. \quad \sum_{i=1}^m \mathbf{A}_i \mathbf{x}_i = \mathbf{b}. \quad (1)$$

We define

$$f(\mathbf{x}) = \sum_{i=1}^m f_i(\mathbf{x}_i) \quad (2)$$

$$\mathbf{A} = [\mathbf{A}_1, \dots, \mathbf{A}_m] \quad (3)$$

$$\mathbf{x} = (\mathbf{x}_1^T, \dots, \mathbf{x}_m^T)^T \quad (4)$$

for simplification.

Direct Extension of ADMM

The straightforward extension of the vanilla ADMM:

$$\tilde{\mathbf{x}}_i^{k+1} = \underset{\mathbf{x}_i}{\operatorname{argmin}} L_\beta(\tilde{\mathbf{x}}_1^{k+1}, \dots, \tilde{\mathbf{x}}_{i-1}^{k+1}, \mathbf{x}_i, \mathbf{x}_{i+1}^k, \dots, \mathbf{x}_m^k, \boldsymbol{\lambda}^k), \forall i, \quad (5)$$

$$\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + \beta \left(\sum_{i=1}^m \mathbf{A}_i \tilde{\mathbf{x}}_i^{k+1} - \mathbf{b} \right), \quad (6)$$

with $\mathbf{x}_i^k = \tilde{\mathbf{x}}_i^k$ (*no correction here*), where

$$\begin{aligned} L_\beta(\mathbf{x}_1, \dots, \mathbf{x}_m, \boldsymbol{\lambda}) &= \sum_i f_i(\mathbf{x}_i) + \langle \boldsymbol{\lambda}, \sum_i \mathbf{A}_i \mathbf{x}_i - \mathbf{b} \rangle \\ &\quad + \frac{\beta}{2} \left\| \sum_i \mathbf{A}_i \mathbf{x}_i - \mathbf{b} \right\|^2. \end{aligned} \quad (7)$$

However, it is proved that the above method might not convergent. Thus, we should make several modifications on the original ADMM for convergence guarantees.

Outline

I Direction Extension of ADMM

II Gaussian Back Substitution
ADMM-GBS

III Prediction-Correction
ADMM-PC

IV Linearized ADMM with Parallel Splitting
LADMM-PS
LADMM-SPU

References

ADMM-GBS

Now we introduce the Gaussian Back Substitution scheme. It first predicts $\tilde{\mathbf{x}}_i^{k+1}$ for all $i \in [m]$, and then corrects \mathbf{x}_i^{k+1} from $\tilde{\mathbf{x}}_i^{k+1}$. Denote

$$\mathbf{M} = \begin{bmatrix} \mathbf{A}_1^T \mathbf{A}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{A}_2^T \mathbf{A}_1 & \mathbf{A}_2^T \mathbf{A}_2 & \ddots & \mathbf{0} \\ \vdots & \vdots & \ddots & \mathbf{0} \\ \mathbf{A}_m^T \mathbf{A}_1 & \mathbf{A}_m^T \mathbf{A}_2 & \cdots & \mathbf{A}_m^T \mathbf{A}_m \end{bmatrix} \quad (8)$$

and

$$\mathbf{H} = \text{diag}(\mathbf{A}_1^T \mathbf{A}_1, \dots, \mathbf{A}_m^T \mathbf{A}_m), \quad (9)$$

\mathbf{M} and \mathbf{H} will be used in the following proofs frequently.

ADMM-GBS

Lemma I.1

Suppose that $i \in [m]$ [$f_i(\mathbf{x}_i)$ is convex]. Then with (5) and (6), we have

$$\begin{aligned} f(\tilde{\mathbf{x}}^{k+1}) - f(\mathbf{x}^*) &+ \langle \boldsymbol{\lambda}^*, \sum_{i=1}^m \mathbf{A}_i \tilde{\mathbf{x}}_i^{k+1} - \mathbf{b} \rangle \\ &\leq \frac{1}{2\beta} (\|\boldsymbol{\lambda}^k - \boldsymbol{\lambda}^*\|^2 - \|\boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*\|^2) - \frac{\beta}{2} \|\tilde{\mathbf{x}}^{k+1} - \mathbf{x}^k\|_{\mathbf{H}}^2 \\ &- \beta(\mathbf{x}^k - \mathbf{x}^*)^T \mathbf{M}(\tilde{\mathbf{x}}^{k+1} - \mathbf{x}^k). \end{aligned} \quad (10)$$

We want to make the RHS of the inequality in a form of $\phi^k - \phi^{k+1}$ such that a recursion can be established.

ADMM-GBS

To this end, we need to define \mathbf{x}^{k+1} and find some $\mathbf{G} \succeq \mathbf{0}$ so

$$\begin{aligned} & -\frac{\beta}{2} \|\tilde{\mathbf{x}}^{k+1} - \mathbf{x}^k\|_{\mathbf{H}}^2 - \beta(\mathbf{x}^k - \mathbf{x}^*)^T \mathbf{M}(\tilde{\mathbf{x}}^{k+1} - \mathbf{x}^k) \\ & = \frac{\beta}{2} (\|\mathbf{x}^k - \mathbf{x}^*\|_{\mathbf{G}}^2 - \|\mathbf{x}^{k+1} - \mathbf{x}^*\|_{\mathbf{G}}^2) \end{aligned} \quad (11)$$

If we let

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \mathbf{D}(\tilde{\mathbf{x}}^{k+1} - \mathbf{x}^k), \quad (12)$$

then the RHS of (11) becomes

$$\begin{aligned} & -(\tilde{\mathbf{x}}^{k+1} - \mathbf{x}^k)^T \mathbf{D}^T \mathbf{G} \mathbf{D}(\tilde{\mathbf{x}}^{k+1} - \mathbf{x}^k) \\ & - 2(\mathbf{x}^k - \mathbf{x}^*)^T \mathbf{G} \mathbf{D}(\mathbf{x}^{k+1} - \mathbf{x}^k). \end{aligned} \quad (13)$$

It means \mathbf{D} and \mathbf{G} should satisfy that

$$\mathbf{D}^T \mathbf{G} \mathbf{D} = \mathbf{H} \text{ and } \mathbf{G} \mathbf{D} = \mathbf{M}. \quad (14)$$

ADMM-GBS

Assume that \mathbf{A}_i 's are all of full column rank. Then \mathbf{M} and \mathbf{H} are invertible. Let $\mathbf{D} = \mathbf{M}^{-T}\mathbf{H}$ and $\mathbf{G} = \mathbf{M}\mathbf{H}^{-1}\mathbf{M}^T$, then (12) becomes

$$\mathbf{x}^{k+1} = \mathbf{x}^k + \mathbf{M}^{-T}\mathbf{H}(\tilde{\mathbf{x}}^{k+1} - \mathbf{x}^k), \quad (15)$$

which can be computed by the famous Gaussian back substitution efficiently due to the lower-block-triangular structure of \mathbf{M} .

We call this method, i.e., (5), (6), and (15), *ADMM-GBS*. Note that $\tilde{\mathbf{x}}_i$ and \mathbf{x}_i are updated sequentially.

Ergodic Convergence Rate of ADMM-GBS

Theorem I.1

Suppose that $i \in [m]$ [$f_i(\mathbf{x}_i)$ is convex]. Then for ADMM-GBS we have

$$|f(\mathbf{x}^{\hat{K}+1}) - f(\mathbf{x}^*)| \leq \frac{C}{2(K+1)} + \frac{2\sqrt{C}\|\boldsymbol{\lambda}^*\|}{\sqrt{\beta}(K+1)} \quad (16)$$

$$\|\mathbf{A}\hat{\mathbf{x}}^{K+1} - \mathbf{b}\| \leq \frac{2\sqrt{C}}{\sqrt{\beta}(K+1)}, \quad (17)$$

where

$$\hat{\mathbf{x}}^{K+1} = \frac{1}{K+1} \sum_{k=0}^K \tilde{\mathbf{x}}^{k+1} \quad (18)$$

$$C = \frac{1}{\beta} \|\boldsymbol{\lambda}^0 - \boldsymbol{\lambda}^*\|^2 + \beta \|\mathbf{x}^0 - \mathbf{x}^*\|_{\mathbf{G}}^2. \quad (19)$$

Outline

I Direction Extension of ADMM

II Gaussian Back Substitution
ADMM-GBS

III Prediction-Correction
ADMM-PC

IV Linearized ADMM with Parallel Splitting
LADMM-PS
LADMM-SPU

References

ADMM-PC

ADMM-GBS is not simple enough since it needs to solve m small-scale linear systems. In the following, we give an improved strategy, which is based on the intuition that, *when solving (5), we actually only need $\mathbf{A}_i \mathbf{x}_i^k$ rather than \mathbf{x}_i^k* . Thus, we don't have to compute \mathbf{x}_i^k explicitly.

Denote $\mathbf{P} = \text{diag}(\mathbf{A}_1, \dots, \mathbf{A}_m)$ and

$$\mathbf{L} = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{I} & \mathbf{I} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{I} & \mathbf{I} & \cdots & \mathbf{I} \end{bmatrix} \quad (20)$$

Then $\mathbf{M} = \mathbf{P}^T \mathbf{L} \mathbf{P}$, $\mathbf{H} = \mathbf{P}^T \mathbf{P}$.

ADMM-PC

Similar to (11), we need to define \mathbf{x}^{k+1} and find $\mathbf{G}' \succeq \mathbf{0}$ such that:

$$\begin{aligned} & -\frac{\beta}{2} \|\tilde{\mathbf{x}}^{k+1} - \mathbf{x}^k\|_{\mathbf{H}}^2 - \beta(\mathbf{x}^k - \mathbf{x}^*)^T \mathbf{M}(\tilde{\mathbf{x}}^{k+1} - \mathbf{x}^k) \\ & = \frac{\beta}{2} (\|\mathbf{P}\mathbf{x}^k - \mathbf{P}\mathbf{x}^*\|_{\mathbf{G}'}^2 - \|\mathbf{P}\mathbf{x}^{k+1} - \mathbf{P}\mathbf{x}^*\|_{\mathbf{G}'}^2). \end{aligned} \quad (21)$$

If we let

$$\mathbf{P}\mathbf{x}^{k+1} = \mathbf{P}\mathbf{x}^k + \mathbf{D}'(\mathbf{P}\tilde{\mathbf{x}}^{k+1} - \mathbf{P}\mathbf{x}^k), \quad (22)$$

then the RHS of (21) becomes

$$\begin{aligned} & -(\mathbf{P}\tilde{\mathbf{x}}^{k+1} - \mathbf{P}\mathbf{x}^k)^T (\mathbf{D}')^T \mathbf{G}' \mathbf{D}' (\mathbf{P}\tilde{\mathbf{x}}^{k+1} - \mathbf{P}\mathbf{x}^k) \\ & - 2(\mathbf{P}\mathbf{x}^k - \mathbf{P}\mathbf{x}^*)^T \mathbf{G}' \mathbf{D}' (\mathbf{P}\tilde{\mathbf{x}}^{k+1} - \mathbf{P}\mathbf{x}^k). \end{aligned} \quad (23)$$

ADMM-PC

Note that the LHS of (21) equals to

$$-\frac{\beta}{2}\|\mathbf{P}\tilde{\mathbf{x}}^{k+1} - \mathbf{P}\mathbf{x}^k\|^2 - \beta(\mathbf{P}\mathbf{x}^k - \mathbf{P}\mathbf{x}^*)^T \mathbf{L}(\mathbf{P}\tilde{\mathbf{x}}^{k+1} - \mathbf{P}\mathbf{x}^k). \quad (24)$$

Therefore, we should choose

$$\mathbf{D}' = \mathbf{L}^{-T} \text{ and } \mathbf{G}' = \mathbf{L}\mathbf{L}^T. \quad (25)$$

ADMM-PC

Now (22) can be explicitly re-written as

$$\begin{aligned}
 & \begin{pmatrix} \mathbf{A}_1 \mathbf{x}_1^{k+1} \\ \mathbf{A}_2 \mathbf{x}_2^{k+1} \\ \vdots \\ \mathbf{A}_{m-1} \mathbf{x}_{m-1}^{k+1} \\ \mathbf{A}_m \mathbf{x}_m^{k+1} \end{pmatrix} = \begin{pmatrix} \mathbf{A}_1 \mathbf{x}_1^k \\ \mathbf{A}_2 \mathbf{x}_2^k \\ \vdots \\ \mathbf{A}_{m-1} \mathbf{x}_{m-1}^k \\ \mathbf{A}_m \mathbf{x}_m^k \end{pmatrix} \\
 & + \begin{pmatrix} \mathbf{I} & -\mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & -\mathbf{I} & \cdots & \mathbf{0} & \mathbf{0} \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I} & -\mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{A}_1 \tilde{\mathbf{x}}_1^{k+1} - \mathbf{A}_1 \mathbf{x}_1^k \\ \mathbf{A}_2 \tilde{\mathbf{x}}_2^{k+1} - \mathbf{A}_2 \mathbf{x}_1^k \\ \vdots \\ \mathbf{A}_{m-1} \tilde{\mathbf{x}}_{m-1}^{k+1} - \mathbf{A}_{m-1} \mathbf{x}_{m-1}^k \\ \mathbf{A}_m \tilde{\mathbf{x}}_m^{k+1} - \mathbf{A}_m \mathbf{x}_m^k \end{pmatrix}.
 \end{aligned} \tag{26}$$

ADMM-PC

The RHS of (26) equals to

$$\begin{pmatrix} \mathbf{A}_1 \tilde{\mathbf{x}}_1^{k+1} + \mathbf{A}_2 \mathbf{x}_2^k - \mathbf{A}_2 \tilde{\mathbf{x}}_2^{k+1} \\ \mathbf{A}_2 \tilde{\mathbf{x}}_2^{k+1} + \mathbf{A}_3 \mathbf{x}_3^k - \mathbf{A}_3 \tilde{\mathbf{x}}_3^{k+1} \\ \vdots \\ \mathbf{A}_{m-1} \tilde{\mathbf{x}}_{m-1}^{k+1} + \mathbf{A}_m \mathbf{x}_m^k - \mathbf{A}_m \tilde{\mathbf{x}}_m^{k+1} \\ \mathbf{A}_m \tilde{\mathbf{x}}_m^{k+1} \end{pmatrix}. \quad (27)$$

The result shows that we can obtain $\mathbf{A}_i \mathbf{x}_i^{k+1}$ conveniently without solving linear systems.

However, \mathbf{x}_i^{k+1} may not exist given $\mathbf{A}_i \mathbf{x}_i^{k+1}$, making the next round iteration invalid. Thus we need to revise the iterations accordingly.

ADMM-PC

Introducing variables ξ_i^k , which play the role of $\mathbf{A}_i \mathbf{x}_i^k$, we iterate as follows:

$$\tilde{\mathbf{x}}_i^{k+1} = \underset{\mathbf{x}_i}{\operatorname{argmin}} \tilde{L}_i(\underbrace{\tilde{\mathbf{x}}_1^{k+1}, \dots, \tilde{\mathbf{x}}_{i-1}^{k+1}}_{\text{calculated in seq.}}, \mathbf{x}_i, \xi_{i+1}^k, \dots, \xi_m^k, \boldsymbol{\lambda}^k), \forall i, \quad (28)$$

$$\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + \beta \left(\sum_{i=1}^m \mathbf{A}_i \tilde{\mathbf{x}}_i^{k+1} - \mathbf{b} \right) \quad (29)$$

$$\xi^{k+1} = \xi^k + \mathbf{L}^{-T} (\mathbf{P} \tilde{\mathbf{x}}^{k+1} - \xi^k), \quad (30)$$

where $\tilde{L}_i(\mathbf{x}_1, \dots, \mathbf{x}_i, \xi_{i+1}, \dots, \xi_m, \boldsymbol{\lambda})$ is

$$\begin{aligned} & \sum_{j=1}^i f_j(\mathbf{x}_j) + \langle \boldsymbol{\lambda}, \sum_{j=1}^i \mathbf{A}_j \mathbf{x}_j + \sum_{j=i+1}^m \xi_j - \mathbf{b} \rangle \\ & + \frac{\beta}{2} \left\| \sum_{j=1}^i \mathbf{A}_j \mathbf{x}_j + \sum_{j=i+1}^m \xi_j - \mathbf{b} \right\|^2, \end{aligned} \quad (31)$$

ADMM-PC

We call (28), (29), and (30) ADMM with Prediction-Correction (*ADMM-PC*).

For ADMM-PC, we have a similar result:

Lemma I.2

Suppose that $i \in [m]$ [$f_i(\mathbf{x}_i)$ is convex]. Then for ADMM-PC, we have

$$\begin{aligned} & f(\tilde{\mathbf{x}}^{k+1}) - f(\mathbf{x}^*) + \langle \boldsymbol{\lambda}^*, \sum_{i=1}^m \mathbf{A}_i \tilde{\mathbf{x}}_i^{k+1} - \mathbf{b} \rangle \\ & \leq \frac{1}{2\beta} (\|\boldsymbol{\lambda}^k - \boldsymbol{\lambda}^*\|^2 - \|\boldsymbol{\lambda}^{k+1} - \boldsymbol{\lambda}^*\|^2) \\ & \quad + \frac{\beta}{2} \left(\|\boldsymbol{\xi}^k - \mathbf{P}\mathbf{x}^*\|_{\mathbf{L}\mathbf{L}^T}^2 - \|\boldsymbol{\xi}^{k+1} - \mathbf{P}\mathbf{x}^*\|_{\mathbf{L}\mathbf{L}^T}^2 \right). \end{aligned} \quad (32)$$

Ergodic Convergence Rate of ADMM-PC

ADMM-PC have the same convergence rate to ADMM-GBS:

Theorem I.2

Suppose that $i \in [m]$ [$f_i(\mathbf{x}_i)$ is convex]. Then for ADMM-PC (16) and (17) holds, but C changes to

$$\frac{1}{\beta} \|\boldsymbol{\lambda}^0 - \boldsymbol{\lambda}^*\|^2 + \beta \|\boldsymbol{\xi}^0 - \mathbf{P}\mathbf{x}^*\|_{\mathbf{L}\mathbf{L}^T}^2 \quad (33)$$

Outline

I Direction Extension of ADMM

II Gaussian Back Substitution
ADMM-GBS

III Prediction-Correction
ADMM-PC

IV Linearized ADMM with Parallel Splitting
LADMM-PS
LADMM-SPU

References

ADMM-GBS and ADMM-PC have the following drawbacks:

- ▶ It requires to maintain two groups of variables, the predicted and the corrected, thus increasing the memory cost (at least doubled)
- ▶ \mathbf{x}_i^k (ADMM-GBS) or ξ_i^k (ADMM-PC) could be sparse or low-rank, thus the memory consumption can be more

Nevertheless, we can use the linearized augmented Lagrangian method.

LADMM-PS

Similar to (46) and (47) of *ADMM Slide: Part 2*, we use the following iterations:

$$\begin{aligned} \mathbf{x}^{k+1} = \operatorname{argmin}_{\mathbf{x}} & \left(f(\mathbf{x}) + \langle \boldsymbol{\lambda}^k, \mathbf{Ax} - \mathbf{b} \rangle \right. \\ & \left. + \frac{\beta}{2} \|\mathbf{Ax} - \mathbf{b}\|^2 + D_{\Psi}(\mathbf{x}, \mathbf{x}^k) \right), \end{aligned} \quad (34)$$

$$\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + \beta \left(\sum_{i=1}^m \mathbf{A}_i \mathbf{x}_i^{k+1} - \mathbf{b} \right), \quad (35)$$

where $\mathbf{L} = \operatorname{diag}(L_1 \mathbf{I}_{n_1}, \dots, L_m \mathbf{I}_{n_m})$, in which n_i is the dimension of \mathbf{x}_i and L_i is to be determined.

Note that we hope each \mathbf{x}_i^k can be calculated separately for all $i \in [m]$.

LADMM-PS

If we set

$$\Psi(\mathbf{x}) = \frac{\beta}{2} \|\mathbf{x}\|_{\mathbf{L}}^2 - \frac{\beta}{2} \|\mathbf{Ax}\|^2, \quad (36)$$

then with $\mathbf{Ax} - \mathbf{b} = (\mathbf{Ax}^k - \mathbf{b}) + (\mathbf{Ax} - \mathbf{Ax}^k)$, we have

$$\begin{aligned} & \frac{\beta}{2} \|\mathbf{Ax} - \mathbf{b}\|^2 + D_{\Psi}(\mathbf{x}, \mathbf{x}^k) \\ &= \frac{\beta}{2} \|\mathbf{Ax}^k - \mathbf{b}\|^2 + \beta \langle \mathbf{A}^T(\mathbf{Ax}^k - \mathbf{b}), \mathbf{x} - \mathbf{x}^k \rangle + \frac{\beta}{2} \|\mathbf{x} - \mathbf{x}^k\|_{\mathbf{L}}^2. \end{aligned}$$

If we choose $L_i \geq m \|\mathbf{A}_i\|^2$, then $\forall \mathbf{x}$,

$$\begin{aligned} \|\mathbf{x}\|_{\mathbf{L}}^2 &= \sum_i L_i \mathbf{x}_i^T \mathbf{x}_i \geq m \sum_i \|\mathbf{A}_i\|^2 \|\mathbf{x}_i\|^2 \geq m \sum_i \|\mathbf{A}_i \mathbf{x}_i\|^2 \\ &\geq \left\| \sum_i \mathbf{A}_i \mathbf{x}_i \right\|^2 \geq \|\mathbf{Ax}\|^2. \end{aligned} \quad (37)$$

LADMM-PS

In the above slide, we get $\|\mathbf{x}\|_L^2 \geq \|\mathbf{Ax}\|^2$. Now (34) becomes separable and the subproblem for each \mathbf{x}_i is

$$\begin{aligned}\mathbf{x}_i^{k+1} = \operatorname{argmin}_{\mathbf{x}_i} & \left(f_i(\mathbf{x}_i) + \langle \boldsymbol{\lambda}^k, \mathbf{A}_i \mathbf{x}_i \rangle \right. \\ & + \beta \langle \mathbf{A}_i^T \left(\sum_{j=1}^m \mathbf{A}_j \mathbf{x}_j^k - \mathbf{b} \right), \mathbf{x}_i - \mathbf{x}_i^k \rangle \\ & \left. + \frac{\beta L_i}{2} \|\mathbf{x}_i - \mathbf{x}_i^k\|^2 \right),\end{aligned}\tag{38}$$

which can be solved in parallel for $i \in [m]$.

We call (34) and (38) linearized ADMM with parallel splitting (*LADMM-PS*). Similar to Bregman ADMM, LADMM-PS has the $\mathcal{O}(1/K)$ and the linear convergence rates under certain conditions.

LADMM-SPU

LADMM-PS works by replacing the serial update order in linearized ADMM by the parallel update order in the linearized augmented Lagrangian method.

Since for the two-block case, serial update order is faster than parallel update order, we may combine them for faster convergence when dealing with the multi-block case. That is, divide the m blocks into two partitions:

$$1, \dots, m' \quad \text{and} \quad m' + 1, \dots, m \quad (39)$$

and then update the two partitions in serial, while updating the blocks in the same partition in parallel.

The detailed iterates are:

$$\begin{aligned}
 \mathbf{x}_i^{k+1} = & \underset{\mathbf{x}_i}{\operatorname{argmin}} \left(f_i(\mathbf{x}_i) + \langle \boldsymbol{\lambda}^k, \mathbf{A}_i \mathbf{x}_i \rangle \right. \\
 & + \beta \langle \mathbf{A}_i^T \left(\sum_{t=1}^m \mathbf{A}_t \mathbf{x}_t^k - \mathbf{b} \right), \mathbf{x}_i - \mathbf{x}_i^k \rangle \\
 & \left. + \frac{m' \beta \|\mathbf{A}_i\|^2}{2} \|\mathbf{x}_i - \mathbf{x}_i^k\|^2 \right), \quad \forall i \in [1, m'], \quad (40)
 \end{aligned}$$

$$\begin{aligned}
 \mathbf{x}_j^{k+1} = & \underset{\mathbf{x}_j}{\operatorname{argmin}} \left(f_j(\mathbf{x}_j) + \langle \boldsymbol{\lambda}^k, \mathbf{A}_j \mathbf{x}_j \rangle \right. \\
 & + \beta \langle \mathbf{A}_j^T \left(\sum_{t=1}^m \mathbf{A}_t \mathbf{x}_t^{k+1} + \sum_{t=m'+1}^m \mathbf{A}_t \mathbf{x}_t^k - \mathbf{b} \right), \mathbf{x}_j - \mathbf{x}_j^k \rangle \\
 & \left. + \frac{(m - m') \beta \|\mathbf{A}_j\|^2}{2} \|\mathbf{x}_j - \mathbf{x}_j^k\|^2 \right), \quad \forall i \in [m' + 1, m]. \quad (41)
 \end{aligned}$$

LADMM-SPU

The detailed iterates are (cont'd):

$$\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k + \beta \left(\sum_{i=1}^m \mathbf{A}_i \mathbf{x}_i^{k+1} - \mathbf{b} \right). \quad (42)$$

We call (40), (41), and (42) ADMM with the Serial and the Parallel Update Orders (LADMM-SPU).

LADMM-SPU has exactly the same form of Bregman ADMM. Thus we can also have the sublinear and the linear convergence rates under different conditions, which is faster than making the m blocks parallel.

Outline

I Direction Extension of ADMM

II Gaussian Back Substitution
ADMM-GBS

III Prediction-Correction
ADMM-PC

IV Linearized ADMM with Parallel Splitting
LADMM-PS
LADMM-SPU

References

References

1. Lin, Zhouchen, Huan Li, and Cong Fang. *Alternating Direction Method of Multipliers for Machine Learning*. Springer Nature, 2022.
2. Boyd, Stephen, Stephen P. Boyd, and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.