

Primal ADMM

$$L^P(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}) = \mathbf{c}^T \mathbf{x}_1 - \mathbf{y}^T (A\mathbf{x}_1 - \mathbf{b}) - \mathbf{s}^T (\mathbf{x}_1 - \mathbf{x}_2) + \frac{\beta}{2} \left(\|A\mathbf{x}_1 - \mathbf{b}\|^2 + \|\mathbf{x}_1 - \mathbf{x}_2\|^2 \right)$$

Primal ADMM Update

$$\nabla_{\mathbf{x}_1} L^P(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}) = \mathbf{c} - A^T \mathbf{y} - \mathbf{s} + \beta (A^T (A\mathbf{x}_1 - \mathbf{b}) + (\mathbf{x}_1 - \mathbf{x}_2))$$

Setting the gradient to 0, we obtain the update step for \mathbf{x}_1 :

$$\mathbf{x}_1 = (A^T A + I)^{-1} \left(\frac{1}{\beta} A^T \mathbf{y} + \frac{1}{\beta} \mathbf{s} - \frac{1}{\beta} \mathbf{c} + A^T \mathbf{b} + \mathbf{x}_2 \right)$$

For \mathbf{x}_2 :

$$\nabla_{\mathbf{x}_2} L^P = \mathbf{s} + \beta (\mathbf{x}_2 - \mathbf{x}_1)$$

Setting the gradient to 0, we obtain the update step for \mathbf{x}_2 :

$$\mathbf{x}_2 = \max \left\{ \mathbf{x}_1 - \frac{\mathbf{s}}{\beta}, 0 \right\}$$

where the max is computed component-wise to ensure that $\mathbf{x}_2 \geq 0$.

Primal with Block Splitting

For simplicity, we first consider splitting the problem into 2 blocks of equal size as follows:

$$\mathbf{x}_1 = \begin{bmatrix} \mathbf{x}_{1,1} \\ \mathbf{x}_{1,2} \end{bmatrix}$$

$$A = \begin{bmatrix} A_1 & A_2 \end{bmatrix}$$

Then, the Lagrangian can be expressed as:

$$\begin{aligned} L^P(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}) &= \mathbf{c}_1^T \mathbf{x}_{1,1} + \mathbf{c}_2^T \mathbf{x}_{1,2} - \mathbf{y}^T (A_1 \mathbf{x}_{1,1} + A_2 \mathbf{x}_{1,2} - \mathbf{b}) - \mathbf{s}_1^T (\mathbf{x}_{1,1} - \mathbf{x}_{2,1}) - \mathbf{s}_2^T (\mathbf{x}_{1,2} - \mathbf{x}_{2,2}) \\ &\quad + \frac{\beta}{2} \left(\|A_1 \mathbf{x}_{1,1} + A_2 \mathbf{x}_{1,2} - \mathbf{b}\|^2 + \|\mathbf{x}_{1,1} - \mathbf{x}_{2,1}\|^2 + \|\mathbf{x}_{1,2} - \mathbf{x}_{2,2}\|^2 \right) \end{aligned}$$

Taking the gradient with respect to the 1st block of \mathbf{x}_1 :

$$\nabla_{\mathbf{x}_{1,1}} L^P(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}) = \mathbf{c}_1 - A_1^T \mathbf{y} - \mathbf{s}_1 + \beta (A_1^T (A_1 \mathbf{x}_{1,1} + A_2 \mathbf{x}_{1,2} - \mathbf{b}) + (\mathbf{x}_{1,1} - \mathbf{x}_{2,1}))$$

Setting the gradient to 0, we obtain the update for $\mathbf{x}_{1,1}$:

$$\mathbf{x}_{1,1}^{k+1} = (A_1^T A_1 + I)^{-1} \left(\frac{1}{\beta} A_1^T \mathbf{y} + \frac{1}{\beta} \mathbf{s}_1 - \frac{1}{\beta} \mathbf{c}_1 + A_1^T \mathbf{b} + \mathbf{x}_{2,1}^k - A_1^T A_2 \mathbf{x}_{1,2}^k \right)$$

By symmetry, the update step for the 2nd block of \mathbf{x}_1 is:

$$\mathbf{x}_{1,2}^{k+1} = (A_2^T A_2 + I)^{-1} \left(\frac{1}{\beta} A_2^T \mathbf{y} + \frac{1}{\beta} \mathbf{s}_2 - \frac{1}{\beta} \mathbf{c}_2 + A_2^T \mathbf{b} + \mathbf{x}_{2,2}^k - A_2^T A_1 \mathbf{x}_{1,1}^{k+1} \right)$$

Note that here we use $\mathbf{x}_{1,1}^{k+1}$ to update $\mathbf{x}_{1,2}$. If we use a randomized update order, then we may end up updating $\mathbf{x}_{1,2}$ first instead.

Block Splitting (for a general number of blocks)

We can easily extend the above result to a general number of blocks. Let U denote the set of blocks which have already been updated at the current iteration. To update block i of \mathbf{x}_1 , we compute:

$$\mathbf{x}_{1,i}^{k+1} = (A_i^T A_i + I)^{-1} \left(\frac{1}{\beta} A_i^T \mathbf{y} + \frac{1}{\beta} \mathbf{s}_i - \frac{1}{\beta} \mathbf{c}_i + A_i^T \mathbf{b} + \mathbf{x}_{2,i}^k - \sum_{j \neq i, j \in U} A_i^T A_j \mathbf{x}_{1,j}^{k+1} - \sum_{j \neq i, j \notin U} A_i^T A_j \mathbf{x}_{1,j}^k \right)$$

where A_i refers to the i^{th} block of columns of A .

Dual ADMM

$$L^d(\mathbf{y}, \mathbf{s}, \mathbf{x}) = -\mathbf{b}^T \mathbf{y} - \mathbf{x}^T (A^T \mathbf{y} + \mathbf{s} - \mathbf{c}) + \frac{\beta}{2} \|A^T \mathbf{y} + \mathbf{s} - \mathbf{c}\|^2$$

Dual ADMM Update

$$\nabla_{\mathbf{y}} L^d(\mathbf{y}, \mathbf{s}, \mathbf{x}) = -\mathbf{b} - A\mathbf{x} + \beta A (A^T \mathbf{y} + \mathbf{s} - \mathbf{c})$$

Setting the gradient to 0, we obtain the update step for \mathbf{y} :

$$\begin{aligned} A (A^T \mathbf{y} + \mathbf{s} - \mathbf{c}) &= \frac{1}{\beta} (A\mathbf{x} + \mathbf{b}) \\ AA^T \mathbf{y} &= \frac{1}{\beta} (A\mathbf{x} + \mathbf{b}) - A\mathbf{s} + A\mathbf{c} \\ \mathbf{y} &= (AA^T)^{-1} \left(\frac{1}{\beta} (A\mathbf{x} + \mathbf{b}) - A\mathbf{s} + A\mathbf{c} \right) \end{aligned}$$

For \mathbf{s} :

$$\nabla_{\mathbf{s}} L^d(\mathbf{y}, \mathbf{s}, \mathbf{x}) = -\mathbf{x} + \beta (A^T \mathbf{y} + \mathbf{s} - \mathbf{c})$$

Setting the gradient to 0, we obtain the update step for \mathbf{s} :

$$\mathbf{s} = \max \left\{ \frac{1}{\beta} \mathbf{x} - A^T \mathbf{y} + \mathbf{c}, 0 \right\}$$

Dual ADMM With Block Splitting

For simplicity, we first consider splitting the problem into 2 blocks of equal size as follows:

$$\begin{aligned} \mathbf{y} &= \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} \\ \mathbf{b} &= \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix} \\ A^T &= [A_1^T \quad A_2^T] \end{aligned}$$

Then the Lagrangian can be expressed as:

$$L^d(\mathbf{y}, \mathbf{s}, \mathbf{x}) = -\mathbf{b}_1^T \mathbf{y}_1 - \mathbf{b}_2^T \mathbf{y}_2 - \mathbf{x}^T (A_1^T \mathbf{y}_1 + A_2^T \mathbf{y}_2 + \mathbf{s} - \mathbf{c}) + \frac{\beta}{2} \|A_1^T \mathbf{y}_1 + A_2^T \mathbf{y}_2 + \mathbf{s} - \mathbf{c}\|^2$$

Differentiating with respect to \mathbf{y}_1 :

$$\nabla_{\mathbf{y}_1} L^d(\mathbf{y}, \mathbf{s}, \mathbf{x}) = -\mathbf{b}_1 - A_1 \mathbf{x} + \beta A_1 (A_1^T \mathbf{y}_1 + A_2^T \mathbf{y}_2 + \mathbf{s} - \mathbf{c})$$

Setting the gradient to 0, we obtain the update for \mathbf{y}_1 :

$$\begin{aligned} A_1 (A_1^T \mathbf{y}_1^{k+1} + A_2^T \mathbf{y}_2^k + \mathbf{s} - \mathbf{c}) &= \frac{1}{\beta} (A_1 \mathbf{x} + \mathbf{b}_1) \\ A_1 A_1^T \mathbf{y}_1^{k+1} &= \frac{1}{\beta} (A_1 \mathbf{x} + \mathbf{b}_1) - A_1 (A_2^T \mathbf{y}_2^k + \mathbf{s} - \mathbf{c}) \\ \mathbf{y}_1^{k+1} &= (A_1 A_1^T)^{-1} \left(\frac{1}{\beta} (A_1 \mathbf{x} + \mathbf{b}_1) - A_1 (A_2^T \mathbf{y}_2^k + \mathbf{s} - \mathbf{c}) \right) \end{aligned}$$

By symmetry, the update for \mathbf{y}_2 is:

$$\mathbf{y}_2^{k+1} = (A_2 A_2^T)^{-1} \left(\frac{1}{\beta} (A_2 \mathbf{x} + \mathbf{b}_2) - A_2 (A_1^T \mathbf{y}_1^{k+1} + \mathbf{s} - \mathbf{c}) \right)$$

assuming that we update \mathbf{y}_2 after \mathbf{y}_1 .

Dual Block Splitting (for a general number of blocks)

We can easily extend the above result to a general number of blocks. Let U denote the set of blocks which have already been updated at the current iteration. To update block i of \mathbf{y} :

$$\mathbf{y}_i^{k+1} = (A_i A_i^T)^{-1} \left(\frac{1}{\beta} (A_i \mathbf{x} + \mathbf{b}_i) - \sum_{j \neq i, j \in U} A_i A_j^T \mathbf{y}_j^{k+1} - \sum_{j \neq i, j \notin U} A_i A_j^T \mathbf{y}_j^k - A_i (\mathbf{s} - \mathbf{c}) \right)$$

where A_i refers to rows of A .

Interior-Point ADMM for Primal

$$L_\mu^p(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}) = \mathbf{c}^T \mathbf{x}_1 - \mu \sum_j \ln(x_{1,j}) - \mathbf{y}^T (A \mathbf{x}_1 - \mathbf{b}) - \mathbf{s}^T (\mathbf{x}_1 - \mathbf{x}_2) + \frac{\beta}{2} (\|A \mathbf{x}_1 - \mathbf{b}\|^2 + \|\mathbf{x}_1 - \mathbf{x}_2\|^2)$$

TODO: everything past here is not complete.

Update

$$\nabla_{\mathbf{x}_1} L_\mu^p(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}) = \mathbf{c} - \mu \left(\frac{1}{\mathbf{x}_1} \right) - A^T \mathbf{y} - \mathbf{s} + \beta (A^T (A \mathbf{x}_1 - \mathbf{b}) + (\mathbf{x}_1 - \mathbf{x}_2))$$

Setting the gradient to 0, we obtain the update step for x_1 :

$$\begin{aligned} 0 &= \mathbf{c} - \mu \left(\frac{1}{\mathbf{x}_1} \right) - A^T \mathbf{y} - \mathbf{s} + \beta (A^T (A \mathbf{x}_1 - \mathbf{b}) + (\mathbf{x}_1 - \mathbf{x}_2)) \\ &= \end{aligned}$$

Interior-Point ADMM for Dual

$$L_\mu^d(\mathbf{y}, \mathbf{s}, \mathbf{x}) = -\mathbf{b}^T \mathbf{y} - \mu \sum_j \ln(s_j) - \mathbf{x}^T (A^T \mathbf{y} + \mathbf{s} - \mathbf{c}) + \frac{\beta}{2} \|A^T \mathbf{y} + \mathbf{s} - \mathbf{c}\|^2$$

Update for \mathbf{y} :

$$\nabla_{\mathbf{y}} L_\mu^d(\mathbf{y}, \mathbf{s}, \mathbf{x}) = -\mathbf{b} - A \mathbf{x} + \beta A (A^T \mathbf{y} + \mathbf{s} - \mathbf{c})$$

Setting to 0, we get the update:

$$\mathbf{y} = A^{-T} \left(\frac{1}{\beta} A^{-1} (A\mathbf{x} + \mathbf{b}) - \mathbf{s} + \mathbf{c} \right)$$

Update for \mathbf{s} :

$$\nabla_{\mathbf{s}} L_{\mu}^d(\mathbf{y}, \mathbf{s}, \mathbf{x}) = -\mu \left(\frac{1}{\mathbf{s}} \right) - \mathbf{x} + \beta (A^T \mathbf{y} + \mathbf{s} - \mathbf{c})$$