hw8 Optimization

Guillaume Jarry

November 2022

1 Problem 3.5

If we apply the operator

$$(x^{k+1}, u^{k+1}) = (M + \mathbb{F})^{-1}(M - \mathbb{H})(x^k, u^k)$$

Then we get the equality:

$$\begin{bmatrix} x^{k+1} \\ u^{k+1} \end{bmatrix} = \left(\begin{bmatrix} \frac{I}{\alpha} & 2A^T \\ 0 & \frac{I}{\beta} \end{bmatrix} + \begin{bmatrix} \partial f \\ \partial g^* \end{bmatrix} \right)^{-1} \begin{bmatrix} \frac{x^k}{\alpha} + A^T u^k - \nabla h(x^k) \\ Ax^k + \frac{u^k}{\beta} \end{bmatrix}$$

Threfore,

$$\frac{x^{k+1}}{\alpha} + 2A^T u^{k+1} + \partial f(x^k + 1) = \frac{1}{\alpha} x^k + A^T u^k - \nabla h(x^k)$$

$$\frac{u^{k+1}}{\beta} + \partial g^*(x^{k+1} = Ax^k + \frac{1}{\beta} u^k$$

We can then write the intermediate line:

$$u^{k+1} = \underset{u}{argmin}\{g^*(u) + \frac{1}{2\beta}\|u - (Ax^k + \beta u^k)\|^2\}$$

$$x^{k+1} = \underset{x}{argmin}\{f(x) + \frac{1}{2\alpha}\|x - (x^k - \alpha A^T(2u^{k+1} - u^k))\|^2\}$$

And thus we can recognize, because the argmin in the expression above will not be change if we multiply by a positive constant, we can express those expression as proximal operator:

$$u^{k+1} = Prox_{\beta g^*}(Ax^k + \beta u^k)$$

$$x^{k+1} = Prox_{\alpha f}(x^k - \alpha A^T(2u^{k+1} - u^k))$$

Then, if total duality holds,

2 Problem 3.6

Let us introduce the change of variable $z^k = x^k - \alpha u^k$ in PDHG. We get:

$$x^{k+1} = Prox_{\alpha f}(\tilde{z}^k)$$

This line tells us that if we manage to show that $(\tilde{z}^k)$ and (z^k) are equivalent, then (x^k) of PDHG and (x^{k+1}) of DRS are equivalent.

Because the proximal operator will not be changed by a multiplication by a positive constant, Moreau's identity for CCP functions gives us:

$$u^{k+1} = Prox_{\frac{1}{\alpha}g^*}(\frac{1}{\alpha}(\alpha u^k - x^k + 2x^{k+1})) = \frac{1}{\alpha}(2x^{k+1} - \tilde{z}^k) - \frac{1}{\alpha}Prox_g(2x^{k+1} - \tilde{z}^k)$$

And finally, we have:

$$\tilde{z}^{k+1} = x^{k+1} - \alpha u^{k+1} = x^{k+1} - 2x^{k+1} + \tilde{z}^k + Prox_{\alpha g}(2x^{k+1} - \tilde{z}^k)$$

And then by simplyfing the expression we get:

$$\tilde{z}^{k+1} = +Prox_{\alpha g}(2x^{k+1} + \tilde{z}^k) - x^{k+1}$$

And here we recongize exactly the expression of (z^k) of the DRS with $Prox_{\alpha f}$. Then we have established that the sequences (z^k) and $(\tilde{z}^k)$ are equivalent in DRS and PDHG, and therefore the sequences (x^k) of PDHG and $(x^{k+1/2})$ of DRS also are, whch proves that both algorithm are equivalent !

3 Problem 3.7

Let us pose $\mathbb{F} = \begin{bmatrix} 0 & A^T \\ -A & 0 \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} + \begin{bmatrix} \partial f \\ \partial g^* \end{bmatrix}$ and $s = \begin{bmatrix} N & -A^T \\ -A & M \end{bmatrix}$.

Then the problem becomes the fixed point iterator:

$$\begin{bmatrix} x^{k+1} \\ u^{k+1} \end{bmatrix} = (S+F)^{-1} S \begin{bmatrix} x^k \\ u^k \end{bmatrix}$$

.

Now, let us try to see under which conditions this operator is non-expensive.

4 Problem 3.8

$$(M+G)^{-1}(M-H)(x^k, u^k) = (M+\mathbb{G})^{-1} \begin{bmatrix} \frac{x^k}{\alpha} - \beta A^T A x^k - \nabla h(x^k) \\ \frac{u^k}{\beta} - b \end{bmatrix}$$

Which gives us the intermediate line:

$$\frac{x^{k+1}}{\alpha} + \partial f(x^{k+1}) - \beta A^T A x^{k+1} + A^T u^{k+1} = \frac{x^k}{\alpha} - \beta A^T A x^k - \nabla h(x^k)$$

$$\frac{u^{k+1}}{\beta} - Ax^{k+1} = \frac{1}{\beta} u^k - b$$

Then, by rearranging the second line, we can simplify the first line and make the term $A^T A x^{k=1}$ disappear.

$$u^{k+1} = u^k + \beta(Ax^{k+1} - b)$$

$$\frac{x^{k+1}}{\alpha} - \beta A^T A x^{k+1} + \partial f(x^k + 1) = \frac{x^k}{\alpha} - A^T u^k - A^T A \beta x^{k+1} + \beta A^T b - \beta A^T A x^k - \nabla h(x^k)$$

The second line can be summed up as, because f is CCP as (the gradient will be null on the optimal point):

$$x^{k+1} = \underset{x}{argmin}\{f(x) + \frac{1}{2\alpha} \|x - (x^k - \nabla h(x^k) - \alpha A^T (u^k + \beta(Ax^k - b))\|^2\}$$

Because the result of the proximal operator will not be impacted if the inner expression is multiplied by a positive constant, we get the algortihm that we were supposed to obtain:

$$x^{k+1} = Prox_{\alpha f}(x^k - \nabla h(x^k) - \alpha A^T (u^k + \beta(Ax^k - b))$$
$$u^{k+1} = u^k + \beta(Ax^{k+1} - b)$$

Then, let us look at convergence: