Homework assignment #7

Linearized ADMM. Consider the standard problem

$$\text{minimize} \quad f(x) + g(Ax) \tag{1}$$

where f and g are closed convex functions. In lecture 12 (page 12.31) we derived the *proximal method of multipliers* from the proximal point method applied to the primal-dual optimality conditions. Here we write the proximal method of multipliers as

$$\begin{aligned}(x_{k+1}, y_{k+1}) &= \operatorname*{argmin}_{x,y} \left(f(x) + g(y) + \frac{\tau}{2}\|Ax - y + u_k\|_2^2 + \frac{1}{2\sigma}\|x - x_k\|_2^2\right) \\ u_{k+1} &= u_k + Ax_{k+1} - y_{k+1}.\end{aligned}$$

The two parameters τ and σ correspond to $\tau = \sigma = t$ on page 12.31. Using different values can be interpreted as a simple "preconditioning" of the proximal point method (see page 12.29). The variable u_k corresponds to $u_k = z_k/t$ on page 12.31.

We note that the iteration is similar to the augmented Lagrangian method, with an extra term $\|x - x_k\|_2^2$ added to the augmented Lagrangian. Motivated by the interpretation of ADMM as a simplified augmented Lagrangian method, we can replace the joint minimization over x, y by an alternating minimization:

$$\begin{aligned}x_{k+1} &= \operatorname*{argmin}_{x} \left(f(x) + \frac{\tau}{2}\|Ax - y_k + u_k\|_2^2 + \frac{1}{2\sigma}\|x - x_k\|_2^2\right) \\ y_{k+1} &= \mathrm{prox}_{(1/\tau)g}(Ax_{k+1} + u_k) \\ u_{k+1} &= u_k + Ax_{k+1} - y_{k+1}.\end{aligned}$$

For general f and A, the optimization problem in the x-update may be expensive, because the second term in the cost function contributes a quadratic term $x^T A^T A x$. To avoid this, one can make a further simplification and linearize the second term around x_k:

$$\frac{1}{2}\|Ax - y_k + u_k\|_2^2 \approx \frac{1}{2}\|Ax_k - y_k + u_k\|_2^2 + (Ax_k - y_k + u_k)^T A(x - x_k).$$

If we omit the constant terms (in x), the simplified x-update is

$$\begin{aligned}x_{k+1} &= \operatorname*{argmin}_{x} \left(f(x) + \tau(Ax_k - y_k + u_k)^T Ax + \frac{1}{2\sigma}\|x - x_k\|_2^2\right) \\ &= \operatorname*{argmin}_{x} \left(f(x) + \frac{1}{2\sigma}\|x - x_k + \tau\sigma A^T(Ax_k - y_k + u_k)\|_2^2\right) \\ &= \mathrm{prox}_{\sigma f}(x_k - \tau\sigma A^T(Ax_k - y_k + u_k)).\end{aligned}$$

The resulting method is known as *linearized ADMM:*

$$\begin{aligned}
x_{k+1} &= \operatorname{prox}_{\sigma f}(x_k - \tau\sigma A^T(Ax_k - y_k + u_k)) \\
y_{k+1} &= \operatorname{prox}_{(1/\tau)g}(Ax_{k+1} + u_k) \\
u_{k+1} &= u_k + Ax_{k+1} - y_{k+1}.
\end{aligned}$$

Show that linearized ADMM is equivalent to PDHG applied to the dual of (1),

$$\text{maximize} \quad -g^*(z) - f^*(-A^T z).$$

PDHG for this problem is

$$\begin{aligned}
z_{k+1} &= \operatorname{prox}_{\tau g^*}(z_k + \tau A\tilde{x}_k) \\
\tilde{x}_{k+1} &= \operatorname{prox}_{\sigma f}(\tilde{x}_k - \sigma A^T(2z_{k+1} - z_k)).
\end{aligned}$$