**ORIE 6300 Mathematical Programming I** October 27, 2016

# Problem Set 7

*Due Date: November 3, 2016*

1. Let $\gamma$ and $\tau$ be positive real numbers **that satisfy** $\gamma\tau < \frac{1}{\|A\|^2}$. Consider the Chambolle-Pock operator

$$T_{\mathrm{CP}} : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m \times \mathbb{R}^n$$
$$T_{\mathrm{CP}} \begin{bmatrix} y \\ x \end{bmatrix} := \begin{bmatrix} y - \gamma\,(Ax - b) \\ \max\{x + \tau\left(A^T(y - 2\gamma(Ax - b)) - c\right), 0\} \end{bmatrix},$$

In this exercise, we're going to prove that $T_{\mathrm{CP}}$ is firmly-nonexpansive in a Mahalanobis norm $\|x\|_Q$, i.e.,

$$\left(\forall z_1 \in \mathbb{R}^{m+n}\right), \left(\forall z_2 \in \mathbb{R}^{m+n}\right)$$
$$\|T_{\mathrm{CP}} z_1 - T_{\mathrm{CP}} z_2\|_Q^2 \le \|z_1 - z_2\|_Q^2 - \|(z_1 - T_{\mathrm{CP}} z_1) - (z_2 - T_{\mathrm{CP}} z_2)\|_Q^2, \qquad (1)$$

where

$$Q = \begin{bmatrix} \frac{1}{\gamma} I & -A \\ -A^T & \frac{1}{\tau} I \end{bmatrix}.$$

Define the *set-valued* mapping $M : \mathbb{R}^{m+n} \to 2^{\mathbb{R}^{m+n}}$: for all $z = (y, x) \in \mathbb{R}^{m+n}$,

$$Mz := \{-b\} \times (c + N_{\mathbb{R}^m_{\ge 0}}(x)) + \begin{bmatrix} 0 & A \\ -A^T & 0 \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix}.$$

(a) Let $z = (y, x) \in \mathbb{R}^{m+n}$. Show that

$$Q\,(z - T_{\mathrm{CP}} z) \in M T_{\mathrm{CP}} z$$

(**Hint:** use the projection inclusion formula $x - P_C(x) \in N_C(P_C(x))$).

(b) Let $z_1 = (y_1, x_1) \in \mathbb{R}^{m+n}$ and $z_2 = (y_2, x_2) \in \mathbb{R}^{m+n}$. Show that

$$(\forall u_1 \in M z_1), (\forall u_2 \in M z_2) \qquad \langle z_1 - z_2, u_1 - u_2 \rangle \ge 0$$

(this condition states that $M$ is a *monotone operator*). Using Part 1a, conclude that

$$\langle (z_1 - T_{\mathrm{CP}} z_1) - (z_2 - T_{\mathrm{CP}} z_2), T_{\mathrm{CP}} z_1 - T_{\mathrm{CP}} z_2 \rangle_Q \ge 0.$$

where for all $z, z' \in \mathbb{R}^{m+n}$, we have $\langle z, z' \rangle_Q = \langle Qz, z' \rangle$.

(c) Prove (1).

2. This exercise shows that solving a system of linear inequalities is essentially as hard as solving an LP.

   Let $P(A, b) = \{x \mid Ax = b, x \geq 0\}$. Suppose that $x^*$ is a minimizer of $\min_{x \in P(A,b)} c^T x$. Let $x_0 \in \mathbb{R}^n$ and for all $\gamma > 0$, define

$$x_\gamma = P_{P(A,b)}(x_0 - \gamma c).$$

   Prove that

$$\langle c, x_\gamma \rangle \leq \langle c, x^* \rangle + \frac{1}{2\gamma} \|x_0 - x^*\|^2.$$

   For which $\gamma > 0$ is $x_\gamma$ an $\varepsilon$-accuracy solution of the LP? (Recall that $x$ is an $\varepsilon$-accuracy solution if it is feasible and $\langle c, x \rangle < \langle c, x^* \rangle + \varepsilon$.)

3. In this exercise, we learn how to parallelize the Douglas-Rachford Splitting (DRS) algorithm and the Method of Alternating Projections (MAP) through the *product-space trick.*

   Consider $l$ closed convex sets $C_1, \ldots, C_l \subseteq \mathbb{R}^r$. Assume that $C_1 \cap \ldots \cap C_l \neq \emptyset$. Define $C = C_1 \times \cdots \times C_l$. Define the *diagonal* vector subspace $V \subseteq \mathbb{R}^{rl}$:

$$V := \{(x_1, \ldots, x_l) \in \mathbb{R}^{rl} \mid (\forall i)\ x_i \in \mathbb{R}^r, x_1 = x_2 = \cdots = x_l\}.$$

   (a) Given $z \in \mathbb{R}^{rl}$, compute $P_V z$ and determine $\text{Fix}(P_V)$.

   (b) Given $z \in \mathbb{R}^{rl}$, compute $P_C z$ and determine $\text{Fix}(P_C)$.

   (c) Determine $\text{Fix}(P_V P_C)$ and $\text{Fix}\left(\frac{1}{2}(2P_V - I) \circ (2P_C - I) + \frac{1}{2}I\right)$

   (d) Consider the primal-dual pair of linear programs

$$\min\{c^T x \mid Ax = b, x \geq 0\} \qquad \text{and} \qquad \max\{b^T y \mid A^T y \leq c\},$$

   and assume that there exists a primal-dual optimal solution, e.g., $(x^*, y^*) \in \mathbb{R}^{n+m}$. Define

$$D := \begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ c^T & -b^T & 0 \end{bmatrix} \qquad \text{and} \qquad d := \begin{bmatrix} b \\ c \\ 0 \end{bmatrix}.$$

   Note that $Dz = d$ has at least one solution because the LPs are solvable. Let $l = n+m+2$ and define

$$C_l := \left\{ \begin{bmatrix} x \\ y \\ s \end{bmatrix} \mid x \geq 0, s \geq 0 \right\}.$$

   Provide $l - 1$ sets $C_1, \ldots, C_{l-1} \subseteq \mathbb{R}^{m+2n}$ such that (1) $\{z \mid Dz = d\} = C_1 \cap \ldots \cap C_{l-1}$ and (2) for each $i = 1, \ldots, l-1$, the set $C_i$ is defined purely in terms of the $i$th rows of $D$ and $d$.

   As before, define $V \subseteq \mathbb{R}^{l(m+2n)}$ and $C := C_1 \times \ldots \times C_l$. Given $z \in \mathbb{R}^{l(m+2n)}$ compute $P_V P_C(z)$. What is the biggest computational drawback of this approach? Are there other ways to split $\{z \mid Dz = d\}$ into fewer sets? (There is no single correct answer.)