ORIE 6300 Mathematical Programming I November 14, 2014

Problem Set 11

Due Date: November 21, 2014

1. (5 points) Let $A \in \Re^{m \times n}$ have rank m, and let $P_A = I - A^T(AA^T)^{-1}A$. Show that $P_A = P_A^T = P_A^2$ and then that $u^T P_A u = \|P_A u\|^2$ for every $u \in \Re^n$, so that P_A is positive semidefinite.

2. (10 points) Recall that when we determined the affine-scaling direction, we looked at a transformation $x \to \hat{x} = \bar{X}^{-1}x$, where $\bar{X} = diag(x)$ for $x > 0$, intended to map x to e, the vector of all ones. Then the original LP in terms of the transformed variables becomes $\min(\bar{X}c)^T\hat{x}$ subject to $A\bar{X}\hat{x} = b, \hat{x} \geq 0$.

   (a) (2 points) Write the dual of the problem expressed in terms of transformed variables.

   (b) (2 points) If $(\hat{y}, \hat{s})$ are the transformed variables for the dual, then show that $x_j s_j = \hat{x}_j \hat{s}_j$, and thus that $x^T s = \hat{x}^T \hat{s}$ under the transformation.

   (c) (3 points) Recall the potential function $G_q(x, s) = q \ln(x^T s) - \sum_{j=1}^n \ln(x_j s_j)$ for $q = n + \sqrt{n}$. Give the gradient $g = \nabla_x G_q(x, s)$ with respect to x, and give its evaluation at the current point $(\hat{x}, \hat{s})$ in the transformed space.

   (d) (3 points) Give the projection of the gradient g onto the null space of $A\bar{X}$; this gives the affine-scaling direction for the gradient of the potential function.

   The direction of the last part is the direction used in some potential reduction algorithms.

3. (8 points) Suppose we are given points $x \in \mathcal{F}^\circ(P) = \{x \in \Re^n : Ax = b, x > 0\}$ and $(y, s) \in \mathcal{F}^\circ(D) = \{y \in \Re^m, s \in \Re^n : A^T y + s = c, s > 0\}$. Suppose that $x_j s_j \geq \theta\mu$, where $\mu = \frac{1}{n}x^T s$ and $\theta > 0$ is some parameter. Further suppose that we solve the following system for $(\Delta x, \Delta y, \Delta s)$:

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -XSe + \sigma\mu e, \end{bmatrix}$$

   where $\sigma \in [0, 1]$ is a parameter, $X = diag(x)$, and $S = diag(s)$. For any $\alpha \geq 0$, let $(x(\alpha), y(\alpha), s(\alpha)) = (x, y, s) + \alpha(\Delta x, \Delta y, \Delta s)$.

   Consider
$$\bar{\alpha} = \max\{\hat{\alpha} \in [0, 1] : X(\alpha)S(\alpha) \geq \theta\mu(\alpha)e \text{ for all } \alpha \in [0, \hat{\alpha}]\},$$
   where $X(\alpha) = diag(x(\alpha))$, $S(\alpha) = diag(s(\alpha))$, and $\mu(\alpha) = \frac{1}{n}x(\alpha)^T s(\alpha)$. Let $(x', y', s') = (x(\bar{\alpha}), y(\bar{\alpha}), s(\bar{\alpha}))$. Show that either:

   - x' is the optimal solution to $\min(c^T x : Ax = b, x \geq 0)$, and (y', s') to $\max(b^T y : A^T y + s = c, s \geq 0)$; or
   - $x' \in \mathcal{F}^\circ(P)$ and $(y', s') \in \mathcal{F}^\circ(D)$.

   Further show that only the second possibility can occur if $\sigma > 0$.