INSTRUCTIONS

Due in class on Tuesday, November 13, 2007. You can work in groups of 2-3 students and submit the entire assignment as a group. No late homeworks will be accepted without prior instructor approval. The purpose of this homework is to review interior point methods for linear, second order, and semidefinite programming.

A quick word on notation: Let $\langle u, v \rangle$ denote the dot product of $u$ and $v$. If $u$ and $v$ are vectors in $\mathbb{R}^n$, then $\langle u, v \rangle = u^Tv$. If $U$ and $V$ are matrices in $\mathcal{S}^n$ then the Frobenius inner product is defined as $\langle U, V \rangle = \text{trace}(UV)$. Let $F(x)$ denote a self-concordant barrier functional for the cone $K$ and let $D_F$ denote its domain. We will use $g(x)$ and $H(x)$ to denote the gradient and the Hessian of $F(x)$ at $x$. Note that $H(x)$ is symmetric positive definite over $D_F$. Moreover, we will use the following shorthand notation:

$$
g_x(y) = H(x)^{-1}g(y)$$

$$
H_x(y) = H(x)^{-1}H(y).$$

(1)

Suppose $u, v \in \mathbb{R}^n$, then the local inner product and norm at $x \in D_F$ are as follows:

$$
\langle u, v \rangle_x = \langle u, H(x)v \rangle$$

$$
||u||_x = \sqrt{\langle u, u \rangle_x}.
$$

(2)

For matrices $U, V \in \mathcal{S}^n$, the local Frobenius inner product and norm at $x \in D_F$ are as follows:

$$
\langle U, V \rangle_x = \langle U, H(x)V \rangle$$

$$
||U||_x = \sqrt{\langle U, U \rangle_x}.
$$

(3)

We define the spectral norm of $U \in \mathcal{S}^n$ as

$$
||U||^S = \max_{i=1, \ldots, n} |\lambda_i(U)|
$$

(4)

where $\lambda_i(U)$, $i = 1, \ldots, n$ denote the eigenvalues of $U$. Finally, the local spectral norm at $x \in D_F$ is defined as

$$
||U||^S_x = \max_{i=1, \ldots, n} |\lambda_i(H(x)^{\frac{1}{2}}UH(x)^{-\frac{1}{2}})|.
$$

(5)

1. A functional $F(x)$ is said to be self-concordant if for all $x \in D_F$ we have
(a) \( y \in D_F \) for all \( y \) satisfying \( ||y - x||_x < 1 \).

(b) If \( y \) satisfies \( ||y - x||_x < 1 \) and \( v \) is a nonzero vector, then

\[
1 - ||y - x||_x \leq \frac{||v||_y}{||v||_x} \leq \frac{1}{1 - ||y - x||_x}.
\] (6)

Show that \( F(x) = -\sum_{i=1}^{n} \log x_i \) (barrier functional for the linear cone) and \( F(X) = -\log \det X \) (barrier functional for the semidefinite cone) are self-concordant barrier functionals by showing that they satisfy the above definition. For the semidefinite cone, use the local Frobenius norm in (6).

2. Consider the following self concordant barrier functional

\[
F(x) = -\log(x_1^2 - x_2^2 - \ldots - x_n^2)
= -\log(x^T J x)
\] (7)

for the second order cone \( Q^n_+ = \{ x \in \mathbb{R}^n : x_1 \geq \sqrt{x_2^2 + \ldots + x_n^2} \} \) where

\[
J = \begin{pmatrix}
1 & -1 & \cdots & -1 \\
-1 & \ddots & \ddots & \vdots \\
\vdots & \ddots & -1 & -1 \\
-1 & \cdots & -1 & -1
\end{pmatrix}.
\]

Show the following:

(a) The gradient and the Hessian of \( F(x) \) are

\[
g(x) = -\frac{2}{x^T J x} J x
\]

and

\[
H(x) = \frac{4}{(x^T J x)^2} J x x^T J - \frac{2}{x^T J x} J
\]

respectively.

(b) The barrier parameter

\[
\theta(F) = \sup_{x \in D_F} (\nabla F(x)^T \nabla^2 F(x)^{-1} \nabla F(x))
= \frac{1}{2}.
\]

3. Let \( x \in \mathbb{R}^n \) and let \( F(x) \) be a self concordant barrier functional for a cone \( \mathcal{K} \). Show that

\[
\min_v \frac{||v||_y^2}{||v||_x^2} = \frac{1}{||H_x(y)^{-1}||_x^2}
\]
where $||.||_x^S$ denotes the local spectral norm defined in (5). Hence, show that

$$||H_x(y)^{-1}||_x^S \leq \frac{1}{(1 - ||y - x||_x)^2}.$$ 

**Hint:** Use the definition of self-concordance from Problem 1.

4. Consider the conic program

$$\min \ c^T x$$
$$\text{s.t.} \quad Ax = b$$
$$x \in \mathcal{K}. \quad (8)$$

Let $F(x)$ be a self-concordant barrier functional for the cone $\mathcal{K}$ and let $g(x)$ and $H(x)$ denote its gradient and Hessian, respectively. While solving (8) with an interior point method, we generate the following family of problems

$$\min \ t c^T x + F(x)$$
$$\text{s.t.} \quad Ax = b \quad (9)$$

where $t > 0$ is a parameter.

(a) We derived the following optimality conditions

$$A^T y + s = c$$
$$Ax = b$$
$$ts + g(x) = 0 \quad (10)$$

for (9) in class. Note that the first two equations in (10) are linear while the third equation is nonlinear. Let $(\bar{x}, \bar{y}, \bar{s})$ be a solution satisfying $A\bar{x} = b$ and $A^T \bar{y} + \bar{s} = c$. We will apply one iteration of Newton’s method starting at $(\bar{x}, \bar{y}, \bar{s})$ to (10). Applying Newton’s method to (10) gives the following linear system of linear equations

$$A^T \Delta y + \Delta s = 0$$
$$A\Delta x = 0$$
$$t(\bar{s} + \Delta s) + g(\bar{x}) + H(\bar{x})\Delta x = 0 \quad (11)$$

which we will use to solve for $(\Delta x, \Delta y, \Delta s)$. Show that the solution to (11) is given by

$$\Delta y = (AH(\bar{x})^{-1}A^T)^{-1}(AH(\bar{x})^{-1}\bar{s} - t^{-1}A\bar{x})$$
$$\Delta s = -A^T \Delta y$$
$$\Delta x = \bar{x} - tH(\bar{x})^{-1}(\bar{s} + \Delta s). \quad (12)$$

Show that one can directly compute $\Delta x$ for (11) as

$$\Delta x = -(I - H(\bar{x})^{-1}A^T(AH(\bar{x})^{-1}A^T)^{-1}A)H(\bar{x})^{-1}(t\bar{s} + g(\bar{x})). \quad (13)$$
(b) What is the Newton system (11) and the solution to the Newton system (12) when \( \mathcal{K} \) is a positive semidefinite cone of size \( n \)?

5. Let \( Q \) be a positive definite matrix such that

\[
Q^{-1}SQ^{-1} = QXQ \tag{14}
\]

Given \( X, S > 0 \), the solution \( Q \) to (14) is given by

\[
Q = P^{\frac{1}{2}} \quad \text{where} \quad P = S^{\frac{1}{2}}(S^{\frac{1}{2}}XS^{\frac{1}{2}})^{-\frac{1}{2}}S^{\frac{1}{2}}. \tag{15}
\]

Note that the matrix equation (14) can also be written as \( Q^2XQ^2 = S \). So, we can solve (14) by first solving \( PXP = S \) for \( P \), and then computing \( Q = P^{\frac{1}{2}} \). The matrix equation \( PXP = S \) is also called a Riccati equation. You need to show the following:

(a) \( P \) is symmetric and positive definite so that \( Q = P^{\frac{1}{2}} \) is well defined.

(b) \( Q^{-1}SQ^{-1} = QXQ \) or equivalently \( Q^2XQ^2 = S \).

The Nesterov Todd direction in semidefinite programming employs \( Q \) as a scaling matrix in the computation of the search direction.