AMATH/MATH 516
FIRST HOMEWORK SET

Due by class time Thursday 04/05/07. The purpose of this problem set is to have you brush up and further develop your multi-variable calculus and linear algebra skills. The problem set will be very difficult for some and straightforward for others. If you are having any difficulty, please feel free to discuss the problems with me at any time. Don't delay in starting work on these problems!

1. Let Q be an $n \times n$ symmetric positive definite matrix.
   (a) Show that the eigenvalues of Q^2 are the square of the eigenvalues of Q.
   (b) If $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ are the eigen values of Q, show that
   $$\lambda_n \|u\|_2^2 \leq u^T Q u \leq \lambda_1 \|u\|_2^2 \ \forall\ u \in I\!\!R^n.$$
   (c) If $0 < \underline{\lambda} < \bar{\lambda}$ are such that
   $$\underline{\lambda}\|u\|_2^2 \leq u^T Q u \leq \bar{\lambda}\|u\|_2^2 \ \forall\ u \in I\!\!R^n,$$
   then all of the eigenvalues of Q must lie in the interval $[\underline{\lambda}, \bar{\lambda}]$.
   (d) Let $\underline{\lambda}$ and $\bar{\lambda}$ be as in Part (c) above. Show that
   $$\underline{\lambda}\|u\|_2 \leq \|Qu\|_2 \leq \bar{\lambda}\|u\|_2 \ \forall\ u \in I\!\!R^n \ .$$
   Hint: $\|Qu\|_2^2 = u^T Q^2 u$.
2. Consider the quadratic function $f : I\!\!R^n \mapsto I\!\!R$ given by
   $$f(x) := \frac{1}{2} x^T Q x - a^T x + \alpha \ ,$$
   where $Q \in I\!\!R^{n\times n}$, $a \in I\!\!R^n$, and $\alpha \in I\!\!R$.
   (a) Write expressions for both $\nabla f(x)$ and $\nabla^2 f(x)$. Since it is not assumed that f is symmetric, be careful in how you express $\nabla^2 f(x)$.
   (b) If it is further assumed that Q is symmetric, what is $\nabla^2 f$?
   (c) State first– and second–order necessary conditions for optimality in the problem $\min\{f(x) \,:\, x \in I\!\!R^n\}$.
   (d) State a sufficient condition on the matrix Q under which the problem $\min f$ has a unique global solution and then display this solution in terms of the data Q and a.
3. Consider the linear equation
   $$Ax = b,$$
   where $A \in I\!\!R^{m\times n}$ and $b \in I\!\!R^m$. When $n < m$ it is often the case that this equation is over-determined in the sense that no solution x exists. In such cases one often attempts to locate a 'best' solution in a least squares sense. That is one solves the *linear least squares problem*
   $$(\text{lls}) : \text{minimize } \frac{1}{2}\|Ax - b\|_2^2$$
   for x. Define $f : I\!\!R^n \mapsto I\!\!R$ by
   $$f(x) := \frac{1}{2}\|Ax - b\|_2^2.$$
   (a) Show that f can be written as a quadratic function, that is, it can be written in the same form as the function of the preceding exercise.
   (b) What are $\nabla f(x)$ and $\nabla^2 f(x)$?

(c) Show that $\nabla^2 f(x)$ is positive semi-definite.

(d) Show that a solution to (lls) must always exist.

(e) Provide a necessary and sufficient condition on the matrix A (**not on the matrix** $A^T A$) under which (lls) has a unique solution and then display this solution in terms of the data A and b.

4. A mapping $\langle \cdot, \cdot \rangle : I\!R^n \mapsto I\!R^n$ is said to be an inner product on $I\!R^n$ is for all $x, y, z \in I\!R^n$

$$\begin{array}{cll} \text{(i)} & \langle x, x \rangle \geq 0 & \text{Non-Negative} \\ \text{(ii)} & \langle x, x \rangle = 0 \Leftrightarrow x = 0 & \text{Positive} \\ \text{(iii)} & \langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle & \text{Additive} \\ \text{(iv)} & \langle \alpha x, y \rangle = \alpha \langle x, y \rangle \ \forall \, \alpha \in I\!R & \text{Homogeneous} \\ \text{(v)} & \langle x, y \rangle = \langle y, x \rangle & \text{Symmetric} \end{array}$$

Two vectors $x, y \in I\!R^n$ are said to be orthogonal in the inner product $\langle \cdot, \cdot \rangle$ if $\langle x, y \rangle = 0$

Unless otherwise specified, we use the notation $\langle x, y \rangle$ to designate the usual Euclidean inner product:

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i \ .$$

(a) Let $\langle x, y \rangle$ be the Euclidean inner product on $I\!R^n$. Given $A \in I\!R^{n \times n}$, show that $A = 0$ if and only if

$$\langle x, Ay \rangle = 0 \quad \forall \ x, y \in I\!R^n \ .$$

(b) Let $H \in I\!R^{n \times n}$ be symmetric and positive definite (i.e. $H = H^T$ and $x^T H x > 0 \ \forall \ x \in I\!R^n \setminus \{0\}$). Show that the bi-linear form given by

$$\langle x, y \rangle_H = x^T H y \quad \forall \ x, y \in I\!R^n$$

defines an inner product on $I\!R^n$.

(c) Every inner product defines a transformation on the space of linear operators called the *adjoint*. For the Euclidean inner product on $I\!R^n$, this is just the usual transpose. Given a linear transformation $M : I\!R^n \mapsto I\!R^n$, the adjoint is defined by the relation

$$\langle y, Mx \rangle = \langle M^* y, x \rangle \ , \quad \text{for all } x, y, \in I\!R^n.$$

The inner product given above, $\langle \cdot, \cdot \rangle_H$, also defines an adjoint mapping which we can denote by M^{T_H}. Show that

$$M^{T_H} = H^{-1} M^T H \ .$$

(d) The matrix $P \in I\!R^{n \times n}$ is said to a projection if $P^2 = P$. Clearly, if P is a projection, then so is $I - P$. The subspace $P I\!R^n = \mathrm{Ran\,(P)}$ is called the subspace that P projects onto. A projection is said to be orthogonal with respect to a given inner product $\langle \cdot, \cdot \rangle$ on $I\!R^n$ if and only if

$$\langle (I - P)x, Py \rangle = 0 \quad \forall \ x, y \in I\!R^n \ ,$$

that is, the subspaces $\mathrm{Ran\,(P)}$ and $\mathrm{Ran\,(I - P)}$ are orthogonal in the inner product $\langle \cdot, \cdot \rangle$. Show that the projection P is orthogonal with respect to the inner product $\langle \cdot, \cdot \rangle_H$ (defined above), where $H \in I\!R^{n \times n}$ is symmetric and positive definite, if and only if

$$P = H^{-1} P^T H \ .$$

5. Consider the minimization problem

$$\mathcal{P} : \quad \begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & Ax = b \ , \end{array}$$

where $f : I\!R^n \mapsto I\!R$ is assumed to be twice continuously differentiable, $A \in I\!R^{m \times n}$ has full rank with $m \leq n$, and $b \in I\!R^m$. Set

$$P := I - A^T (A A^T)^{-1} A \ .$$

(a) Show that P is well–defined, that is, show that the matrix AA^T is non–singular.

(b) Show that P is the *orthogonal projector* onto the nulspace of A. That is, show that P is an orthogonal projector and Ran (P) = Nul (A).

(c) Show that if $\bar{x} \in I\!R^n$ is a locally optimal solution to $\mathcal{P}$, then $P\nabla f(\bar{x}) = 0$.

(d) Set $h(z) = f(x_0 + Pz)$ where x_0 is any point satisfying $Ax_0 = b$. Let $\mathcal{S}_1$ be the set of first–order stationary points for the problems $\mathcal{P}$ and let $\mathcal{S}_2$ be the set of first–order stationary points for the problem $\min\{h(z) : z \in I\!R^n\}$. Show that $\mathcal{S}_1 = x_0 + P(\mathcal{S}_2)$. Recall that the first-order necessary conditions for $\mathcal{P}$ are $\nabla f(x) \in \mathrm{Nul}(A)^\perp$ where $\mathrm{Nul}(A)^\perp$ is the subspace orthogonal to the null-space of A. Display both the gradient and Hessian of h.

(e) Show that if $P\nabla f(\bar{x}) = 0$ and f is convex, the $\bar{x}$ is a globally optimal solution to $\mathcal{P}$.

6. Let $H \in I\!R^{n\times n}_s$, $u \in I\!R^n$, and $\alpha \in I\!R$ where $I\!R^{n\times n}_s$ is the linear space of all real symmetric $n \times n$ matrices. Recall that H is said to be *positive definite* if $x^T Hx > 0$ for all $x \in I\!R^n$ with $x \neq 0$. Moreover, H is said to be *positive semi–definite* if $x^T Hx \geq 0$ for all $x \in I\!R^n$. We consider the block matrix

$$\hat{H} := \begin{bmatrix} H & u \\ u^T & \alpha \end{bmatrix}.$$

(a) Show that $\hat{H}$ is positive semi–definite if and only if H is positive semi–definite and there exists a vector $z \in I\!R^n$ such that $u = Hz$ and $\alpha \geq z^T Hz$.

(b) Show that $\hat{H}$ is positive definite if and only if H is positive definite and $\alpha > u^T H^{-1} u$.

(c) Let $A \in I\!R^{m\times n}$, $b \in I\!R^m$, $c \in I\!R^n$, and $\delta \in I\!R$. Use either Part (a) or Part (b) to show that $x \in I\!R^n$ is a solution to the quadratic inequality

$$(Ax + b)^T (Ax + b) \leq c^T x + \delta$$

if and only if the block matrix

$$\begin{bmatrix} I & (Ax + b) \\ (Ax + b)^T & (c^T x + \delta) \end{bmatrix}$$

is positive semi–definite.

(d) Suppose H is positive definite. Show that

$$\begin{bmatrix} H & u \\ 0 & (\alpha - u^T H^{-1} u) \end{bmatrix} = \begin{bmatrix} I & 0 \\ (-H^{-1}u)^T & 1 \end{bmatrix} \begin{bmatrix} H & u \\ u^T & \alpha \end{bmatrix}.$$

(e) Recall that the *kth principal minor* of a matrix $B \in I\!R^{n\times n}$ is the determinant of the upper left–hand corner k×k–submatrix of B for $1 \leq k \leq n$. Use an induction argument and Parts (b) and (d) above to show that H is positive definite if and only if every principal minor of H is positive.

Note: Your argument **must** use either Part (a) or Part (b) above.

Hint: det(AB)=det(A)det(B).

7. Let $A \in I\!R^{m\times n}$, $c \in I\!R^n$, $a \in I\!R^m$, $\delta > 0$, and $H \in I\!R^{n\times n}$ with H symmetric positive definite. Consider the problem

$$\begin{array}{lll} \mathcal{P} & \min_{x\in I\!R^n} c^T x & \\ & \text{subject to} & Ax = a \\ & & \|x_0 - x\|_H \leq \delta \end{array}$$

where $x_0 \in I\!R^n$ satisfies $Ax_0 = a$,

$$\|z\|_H = (z^T Hz)^{1/2} = [\langle z, z\rangle_H]^{1/2},$$

and the inner product $\langle\cdot,\cdot\rangle_H$ is defined in part (b) of problem 4 above.

(a) Suppose $H = LL^T$ for some non–singular matrix $L \in \mathbb{R}^{n \times n}$, e.g. $L = H^{1/2}$. If Q is the orthogonal projector onto the null–space of AL^{-T} in the usual (or Euclidean) inner product, show that the operator P given by

$$P = L^{-T} Q L^T$$

is the orthogonal projector onto the null–space of A with respect to the inner product $\langle \cdot, \cdot \rangle_H$.

(b) Show that

$$\bar{x} = x_0 - \delta \| PH^{-1} c \|_H^{-1} PH^{-1} c$$

solves $\mathcal{P}$ where P is as given in part (b) above.

Hint: It may be helpful to first reduce the problem to one of the form

$$\begin{array}{ll} \min \hat{c}^T w & \\ \text{subject to} & \hat{A} w = 0 \\ & \|w\|_2^2 \leq \delta^2 \ . \end{array}$$

It is also helpful to apply results relating least–squares to orthogonal projection.