AMATH/MATH 516
FIRST HOMEWORK SET

This first problem set is mostly a warm-up exercise intended help you review some basic concepts from linear algebra and multi-variable calculus used in this course. Problems 1-3 are exercises that you should work out for your own benefit, but they will not be graded. Problems 4 and 5 will be graded and are due by class time Monday April 2. These problems set will be 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. I am very open with giving hints.

(1) Let Q be an $n \times n$ symmetric positive definite matrix. The following fact for symmetric matrices can be used to answer the questions in this problem.

Fact: *If M is a real symmetric $n \times n$ matrix, then there is a real orthogonal $n \times n$ matrix U ($U^TU = I$) and a real diagonal matrix $\Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$ such that $M = U\Lambda U^T$.*

(a) Show that the eigenvalues of Q^2 are the square of the eigenvalues of Q.

(b) If $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ are the eigen values of Q, show that

$$\lambda_n \|u\|_2^2 \leq u^TQu \leq \lambda_1 \|u\|_2^2 \ \forall \ u \in I\!R^n.$$

(c) If $0 < \underline{\lambda} < \bar{\lambda}$ are such that

$$\underline{\lambda}\|u\|_2^2 \leq u^TQu \leq \bar{\lambda}\|u\|_2^2 \ \forall \ u \in I\!R^n,$$

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: $\|Qu\|_2^2 = u^TQ^2u$.

(2) Consider the quadratic function $f : I\!R^n \mapsto I\!R$ given by

$$f(x) := \frac{1}{2}x^TQx - a^Tx + \alpha \ ,$$

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 $\text{Nul}(A^TA) = \text{Nul}(A)$ and $\text{Ran}(A^TA) = \text{Ran}(A^T)$.

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

(f) Provide a necessary and sufficient condition on the matrix A under which (lls) has a unique solution and then display this solution in terms of the data A and b.

(4) 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 $m \le n$, and $b \in I\!R^m$. Set

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

(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 ($P^2 = P$ and $P = P^T$) and $\mathrm{Ran}\,(P) = \mathrm{Nul}\,(A)$.
(c) Set $h(z) = f(x_0 + Pz)$ where x_0 is any point satisfying $Ax_0 = b$. Compute both $\nabla h(z)$ and $\nabla^2 h(z)$.
(d) Show that if $\bar{z}$ solves $\widehat{\mathcal{P}}:\ \min\{h(z)\ :\ z \in I\!R^n\}$, then $\bar{x} = x_0 + P\bar{z}$ solves $\mathcal{P}$. Conversely, show that if $\bar{x}$ solves $\mathcal{P}$, then there exists $\bar{z}$ solving $\widehat{\mathcal{P}}$ such that $\bar{x} = x_0 + P\bar{z}$.
(e) The set of first-order stationary points for the problem $\widehat{\mathcal{P}}$ is the set of points $\mathcal{S}_h = \{z \mid \nabla h(z) = 0\}$. We define the set of first-order stationary points for $\mathcal{P}$ to be $\mathcal{S}_f = \{x_0 + Pz \mid z \in \mathcal{S}_h\}$. Show that

$$\mathcal{S}_f = \{x \mid P\nabla f(x) = 0,\ Ax = b\} = \{x \mid Ax = b,\ \nabla f(x) \perp \mathrm{Nul}(A)\}\ .$$

(5) 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 $x^T Hx \ge 0$ 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 $\alpha \ge z^T Hz$.
(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

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

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 $1 \le k \le n$. 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: $\det(AB)=\det(A)\det(B)$, and the determinant of an upper or lower block triangular matrix is the product of the determinants of the diagonal blocks.