APM462: Homework 4

Due date: Tue, July 12 in class.

Suggested problems (not to be turned in):

(A) Assume that Q is a symmetric $n \times n$ matrix, with eigenvalues $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ and with an *orthonormal basis* of eigenvectors $w_1, \ldots, w_n$.

Since $w_1, \ldots, w_n$ is a basis, any vector $v \in \mathbb{R}^n$ can be written in the form

$$v = a_1 w_1 + \cdots + a_n w_n. \tag{1}$$

(In fact, $a_i = w_i^T v$ for every i — this follows by multiplying equation (1) by w_i^T on the left and using the fact that the vectors $w_1, \ldots, w_n$ are orthonormal.)

(a) Show that if $v = a_1 w_1 + \cdots + a_n w_n$ and at least one a_i is nonzero, then

$$\frac{v^T Q v}{v^T v} = \theta_1 \lambda_1 + \ldots + \theta_n \lambda_n, \qquad \text{where } \theta_i = \frac{a_i^2}{a_1^2 + \cdots + a_n^2}.$$

(b) Using part a) (if you like), prove that

$$\lambda_n = \text{largest eigenvalue of } Q = \max_{v \neq 0} \frac{v^T Q v}{v^T v}. \tag{2}$$

Hints: it may be convenient to break this into two parts: first, that $\frac{v^T Q v}{v^T v} \leq \lambda_n$ for every nonzero vector v, and second, that there is some choice of a nonzero vector v such that $\frac{v^T Q v}{v^T v} = \lambda_n$.

Remark. By almost the same argument, one can also show that

$$\lambda_1 = \text{smallest eigenvalue of } Q = \min_{v \neq 0} \frac{v^T Q v}{v^T v}. \tag{3}$$

(B) Consider the iterative process

$$x_{k+1} = \frac{1}{2}\left(x_k + \frac{a}{x_k}\right),$$

where $a > 0$. Assume the process converges: $\lim_{k \to \infty} x_k = x_\infty$.

(a) What is x_∞?

(b) Show that if $x_0 > \sqrt{a}$ then

$$\sqrt{a} \leq x_{k+1} \leq x_k \text{ for all } k \in \mathbb{N}.$$

(c) Show that if $0 < x_0 < \sqrt{a}$ then

$$0 < x_k \leq x_{k+1} \leq \sqrt{a} \text{ for all } k \in \mathbb{N}.$$

(d) What is the order of convergence?

Remark: This is a first year calculus question about sequences using induction.

The following 3 questions are to be handed in:

(1) Let $f(x,y) = x^2 + y^2 + xy - 3x$.

(a) Find an unconstrained local minimum point $\mathbf{x}_*$ for f.

solution. $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}\cdot Q\mathbf{x} - b\cdot\mathbf{x}$ where $Q = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, $b = (3,0)^T$, and $\mathbf{x} = (x,y)^T$. We know $\mathbf{x}_*$ saisfies $Q\mathbf{x}_* = b$, so $\mathbf{x}_* = Q^{-1}b = (2,-1)^T$. Note that the eigenvalues of Q are 3 and 1.

(b) Why is $\mathbf{x}_*$ actually a global minimum point?

solution. Q is positive definite, so f is strictly convex, so the local minimum is a global minimum.

(c) Using the method of steepest descent, what is the smallest k that will guarantee $E(\mathbf{x_k}) \leq 10^{-3}\, E(\mathbf{x_0})$. Here $E(\mathbf{x}) = f(\mathbf{x}) - f(\mathbf{x}^*)$. **Remark:** This is a staright forward "plug in" question.

solution. $E(\mathbf{x}) = f(\mathbf{x}) - f(\mathbf{x}^*) = (\frac{1}{2}\mathbf{x}\cdot Q\mathbf{x} - b\cdot\mathbf{x}) - (\frac{1}{2}\mathbf{x}_*\cdot Q\mathbf{x}_* - b\cdot\mathbf{x}_*) = \frac{1}{2}(\mathbf{x}-\mathbf{x}_*)\cdot Q(\mathbf{x}-\mathbf{x}_*) = q(\mathbf{x})$, where the function $q(\mathbf{x})$ was defined in leture. So we have $E(\mathbf{x_k}) \leq r^k E(\mathbf{x_0})$, where $r = \left(\frac{\Lambda-\lambda}{\Lambda+\lambda}\right)^2 = \left(\frac{3-1}{3+1}\right)^2 = \frac{1}{4}$. We want smallest integer k such that $4^{-k} \leq 10^{-3}$. Equivalently, $-k\log 4 \leq -3\log 10$, or $k \geq 3\frac{\log 10}{\log 4} = 4.98\ldots$. So $k = 5$.

(2) Assume that Q is a symmetric $n\times n$ matrix, $c \in \mathbb{R}^n$ is a nonzero (column) vector, and μ is a positive number.

Consider the symmetric matrix $R = Q + \mu cc^T$.

Let $\lambda_i(Q)$ denote the ith eigenvalue of Q, and similarly and $\lambda_i(R)$ the ith eigenvalue of R, where they are arranged so that $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$, for both Q and R.

(a) Using formula (2) from exercise (A) above, prove that

$$\lambda_n(R) \geq \mu|c|^2 + \lambda_1(Q).$$

solution. By formula (2)

$$\lambda_n(R) = \max_{v\neq 0}\frac{v^TRv}{v^Tv} \geq \frac{c^TRc}{c^Tc} = \frac{c^TQc}{c^Tc} + \mu\frac{c^Tcc^Tc}{c^Tc}$$

By formula (3), $\frac{c^TQc}{c^Tc} \geq \lambda_1(Q)$, and since $c^Tc = |c|^2$, we deduce from the above that

$$\lambda_n(R) \geq \lambda_1(Q) + \mu\frac{(|c|^2)^2}{|c|^2} = \lambda_1(Q) + \mu|c|^2.$$

(b) Using formula (3) above, prove that if $n \geq 2$, then

$$\lambda_1(R) \leq \lambda_n(Q).$$

solution. If $n \geq 2$, then there must be a nonzero vector $w \in \mathbb{R}^n$ such that $w^T c = 0$. For this vector, $w^T Rw = w^T Qw$. Thus by formulas (3) and (2) (in that order),

$$\lambda_1(R) \leq \frac{w^T Rw}{w^T w} = \frac{w^T Qw}{w^T w} \leq \lambda_n(Q).$$

(c) Conclude that if Q is positive semidefinite, then the condition number of R satisfies

$$\text{condition number of } R = \frac{\lambda_n(R)}{\lambda_1(R)} \geq \frac{\mu|c|^2}{\lambda_n(Q)}.$$

Thus, the condition number is very large if μ is large compared to $\lambda_n(Q)$.

solution. If Q is positive semidefinite, then $\lambda_1(Q) \geq 0$, and part (a) implies that $\lambda_n(R) \geq \mu|c|^2$. So it immediately follows that

$$\text{condition number of } R = \frac{\lambda_n(R)}{\lambda_1(R)} \geq \frac{\mu|c|^2}{\lambda_n(Q)}.$$

(3) Suppose that you want to minimize $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T Q\mathbf{x} - \mathbf{b}^T\mathbf{x}$, where Q is *diagonal*, as well as being positive definite and symmetric. Let $\lambda_1, \cdots, \lambda_n$ be the eigenvalues of Q. Let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be the standard basis vectors for $\mathbb{R}^n$.

(a) Let $\mathbf{d}_i = \mathbf{e}_{i+1}$ for $i = 0, \ldots, n-1$, and show that $\{\mathbf{d}_0, \ldots, \mathbf{d}_{n-1}\}$ form a Q-orthogonal set.

solution. Write $\lambda_1, \ldots \lambda_n$ for the diagonal entries of Q. Then if $i \neq j$,

$$\mathbf{d}_i^T Q\mathbf{d_j} = \mathbf{e_{i+1}}^T Q\mathbf{e_{j+1}} = \mathbf{e_{i+1}}^T \lambda_{j+1}\mathbf{e_{j+1}} = \lambda_{j+1}\mathbf{e_{i+1}}^T \mathbf{e_{j+1}} = 0.$$

(b) Suppose that you try to minimize f using the Conjugate Directions method, with the Q-orthogonal set $\mathbf{d}_0, \ldots, \mathbf{d}_{n-1}$ found above, starting from a point $\mathbf{x}_0 = (a_1, \ldots, a_n)$.
Find $\mathbf{x}_k$ for every $k = 0, \ldots, n-1$.
Hint: $\mathbf{x_k} = (\frac{b_1}{\lambda_1}, \ldots, \frac{b_k}{\lambda_k}, a_{k+1}, \ldots, a_n)$.

solution. Let's write $\mathbf{b} = (b_1, \ldots, b_n)^T$, and as above, $\lambda_1, \ldots \lambda_n$ for the diagonal entries of Q.
We know that the minimizer is

$$\mathbf{x}^* = Q^{-1}\mathbf{b} = (\frac{b_1}{\lambda_1}, \ldots, \frac{b_n}{\lambda_n})$$

We will show that

$$\boxed{\mathbf{x_k} = (\frac{b_1}{\lambda_1}, \ldots, \frac{b_k}{\lambda_k}, a_{k+1}, \ldots, a_n).} \tag{4}$$

Thus, each step of the method replaces one component of the starting vector $\mathbf{x}_0$ with one component of the minimizer $\mathbf{x}^*$. We will prove (4) by induction. It is clear when $k = 0$. Now assume that it holds for $0, \ldots, k$. Then we have

$$\mathbf{g}_k = \nabla f(\mathbf{x}_k) = Q\mathbf{x}_k - b = (0, \ldots, 0, \lambda_{k+1}a_{k+1} - b_{k+1}, \ldots, \lambda_n a_n - b_n)$$

So

$$\mathbf{x}_{k+1} = \mathbf{x_k} - \frac{\mathbf{g}_k^T\mathbf{d}_k}{\mathbf{d}_k^T Q\mathbf{d}_k}\mathbf{d}_k = \mathbf{x_k} - \frac{\mathbf{g}_k^T\mathbf{e}_{k+1}}{\mathbf{e}_{k+1}^T Q\mathbf{e}_{k+1}}\mathbf{e}_{k+1} = \mathbf{x_k} - \frac{\lambda_{k+1}a_{k+1} - b_{k+1}}{\lambda_{k+1}}\mathbf{e}_{k+1}$$

and it is easy to see that the right-hand side is exactly the right-hand side of (4) for $k+1$.

(c) Show directly that for every $k \geq 1$, $\mathbf{x}_k$ minimizes f in the set

$$\mathbf{x}_0 + \mathfrak{B}_k,$$

where $\mathfrak{B}_k = \text{ span}\{\mathbf{d_0}, \ldots, \mathbf{d_{k-1}}\}$.

Hint: One way to do this is to write the restriction of f to $\mathbf{x}_0 + \mathfrak{B}_k$ as a function $\phi : \mathbb{R}^k \to \mathbb{R}$, where

$$\phi(y_1, \ldots, y_k) = f(\mathbf{x_0} + y_1\mathbf{d}_0 + \ldots + y_k\mathbf{d}_{k-1}),$$

and find the minimum of ϕ in $\mathbb{R}^k$, which is an unrestricted minimization problem.

Solution: Since $\mathbf{d}_0$ etc are just the standard basis vectors,

$$\phi(y_1, \ldots, y_k) = f(a_1 + y_1, \ldots, a_k + y_k, a_{k+1}, \ldots, a_n)$$

$$= \sum_{i=1}^{k}\left[\frac{1}{2}\lambda_i(a_i + y_i)^2 - b_i(a_i + y_i)\right] + \sum_{i=k+1}^{n}\left[\frac{1}{2}\lambda_i a_i^2 - b_i a_i\right].$$

When we minimize this (as a function of $(y_1, \ldots, y_k)$), we find that $y_i = b_i/\lambda_i - a_i$ for $i = 1, \ldots, k$. When we look back at f, this corresponds to the point

$$(\mathbf{x_0} + y_1\mathbf{d}_0 + \ldots + y_k\mathbf{d}_{k-1}) = (\frac{b_1}{\lambda_1}, \ldots, \frac{b_k}{\lambda_k}, a_{k+1}, \ldots, a_n),$$

So the minimum point is the one we found in part (b) via the Conjugate Gradient method.