MATH 387 ASSIGNMENT 3

DUE WEDNESDAY MARCH 14

Note: You are encouraged to do additional reading for this assignment, and strongly encouraged to type your solutions in LaTeX.

1. (a) Let $A \in \mathbb{R}^{n\times m}$ be a matrix with full column rank. Show that the reduced QR factorization
$$A = QR$$
exists and is unique, where $Q \in \mathbb{R}^{n\times m}$ has orthonormal columns and R is upper triangular with positive diagonal entries.
   (b) Recall that two matrices A and B are called *similar*, if there is an invertible matrix Θ such that $\Theta A \Theta^{-1} = B$. Show that similar matrices share the same collection of eigenvalues. In particular, if A is similar to B and B is diagonal, then we can simply read off the eigenvalues of A from the diagonal entries of B. If such B exists, then we say that A is *diagonalizable*. Show that even in exact arithmetic, there is no general procedure to construct Θ for any given diagonalizable A, such that $\Theta A \Theta^{-1}$ is diagonal, where by a "procedure" we mean a finite sequence of elementary operations, including taking n-th roots.
2. (a) Describe an algorithm for QR decomposition that is based on Givens rotations. Estimate the asymptotic complexity of the algorithm, and compare it to that of the Householder QR algorithm.
   (b) Adapt the Householder QR algorithm so that it can efficiently handle the case when $A \in \mathbb{R}^{n\times m}$ has lower bandwidth p and upper bandwidth q, i.e., when $a_{ij} = 0$ for $i - j > p$ or $j - i > q$.
   (c) A square matrix B is called *Hessenberg* if $b_{ij} = 0$ for $i - j > 1$, i.e., if all entries below the first sub-diagonal are zero. Come up with a procedure based on Householder reflections, that constructs an orthogonal matrix Q such that $QAQ^\top = B$, where A is a given square matrix, and B is a Hessenberg matrix. Show that in this setting, if A is symmetric, then we can make B tridiagonal. (In view of 1(b), this is about the best we can do for eigenvalue problems, without resorting to infinite processes.)
3. In this exercise, we will study the *Cholesky factorization* $A = R^\top R$, which is an adaptation of the LU factorization to symmetric and positive definite matrices. Recall that A is called *positive definite* if $x^\top Ax > 0$ for all nonzero x. Assume that $A \in \mathbb{R}^{n\times n}$ is symmetric and positive definite, and justify the following steps in detail.
   (a) All eigenvalues of A are positive.
   (b) All principal minors of A are positive, and therefore an LU factorization of A exists.

Date: Winter 2018.

(c) Let $A = LU$ be the LU factorization, and let D be the diagonal matrix consisting of the diagonal entries of U. Then $M = D^{-1}U$ satisfies $M = L^\top$, and hence $A = LDL^\top$.

(d) There exists a diagonal matrix Λ such that $\Lambda^2 = D$, and with $R = \Lambda L^\top$, we have the Cholesky factorization $A = R^\top R$, where R is upper triangular with positive diagonal entries.

(e) The entries of $R = [r_{ij}]$ satisfy the bound

$$r_{ij}^2 \leq a_{jj} \qquad (1 \leq i, j \leq n),$$

where a_{jj} are the diagonal entries of A. This indicates a strong stability property of the Cholesky factorization.

(f) The j-th column of the relation $A = R^\top R$ is

$$A_j = \sum_{k=1}^{j} r_{kj}(R^\top)_k = \sum_{k=1}^{j-1} r_{kj}(R^\top)_k + r_{jj}(R^\top)_j,$$

where $(R^\top)_k$ is the k-th column of $R^\top$, or the transposed k-th row of R. Let us rewrite it as

$$r_{jj}(R^\top)_j = A_j - \sum_{k=1}^{j-1} r_{kj}(R^\top)_k =: v. \qquad (*)$$

The vector $v \in \mathbb{R}^n$ depends only on the first $j-1$ rows of R, and hence the j-th row of R can be computed by

$$(R^\top)_j = \frac{1}{\sqrt{v_j}} v, \qquad (**)$$

where v_j is of course the j-th component of v. Taking the second equality of $(*)$ as a prescription to compute v, the relations $(*)$ and $(**)$, with $j = 1, \ldots, n$, define an algorithm to compute the Cholesky factor R.

(g) The purpose of the j-th step of the aforementioned algorithm is to compute the j-th row of R. Hence we only need to be compute the last $n - j + 1$ components of v in $(*)$. Taking this into account, we estimate the number multiplications in the Cholesky factorization algorithm as $\frac{1}{6}n^3 + O(n^2)$, which shows that it is twice as efficient as the Gaussian elimination.

4. In class, we have shown that if K is a square matrix with $\|K\| < 1$, then $I - K$ is invertible, and

$$I + K + K^2 + \ldots + K^m \to (I - K)^{-1} \qquad \text{as} \quad m \to \infty.$$

We can use this fact to design an iterative method to solve $Ax = b$. The starting point should be to somehow write A in terms of $I - K$, where K has small norm. We can write $A = I - (I - A)$ and set $K = I - A$, but we would need $\|I - A\| < 1$ to ensure convergence. As a simple way to introduce some flexibility, let us multiply $Ax = b$ by some number $\omega \in \mathbb{R} \setminus \{0\}$, to get

$$\omega A x = \omega b,$$

and then introduce $K = I - \omega A$, yielding

$$(I - K)x = \omega b \qquad \Longleftrightarrow \qquad Ax = b.$$

If $\|K\| = \|I - \omega A\| < 1$, then

$$x_m := (I + K + K^2 + \ldots + K^m)\omega b \to x.$$

The iterates x_m satisfy the recurrent relation

$$\begin{aligned} x_{m+1} &= \omega b + K(I + K + \ldots + K^m)\omega b = \omega b + Kx_m = \omega b + (I - \omega A)x_m \\ &= x_m + \omega(b - Ax_m), \end{aligned}$$

which is convenient for implementation.

(a) Assuming that $\|I - \omega A\| < 1$, derive an estimate on $\|x_m - x\|$ that goes to 0 geometrically as $m \to \infty$.

(b) Assuming that A is diagonalizable, and that all its eigenvalues are positive, estimate $\|I - \omega A\|$ in terms of λ_1, λ_n, and ω. Here λ_1 and λ_n are the smallest and the largest eigenvalues of A, respectively.

(c) In the estimate derived in (b), optimize the choice of the parameter ω.