MATH 387 ASSIGNMENT 2

DUE TUESDAY MARCH 8

1. (Trefethen-Bau) Recall that Gaussian elimination yields a factorization $A = LU$, where L has unit diagonal but U in general does not. Describe the factorization that results if this process is varied in the following ways.
   (a) Elimination by columns from left to right, rather than by rows from top to bottom, so that A is made lower triangular.
   (b) Gaussian elimination applied after a preliminary scaling of the columns of A by a diagonal matrix D. What form does a system $Ax = b$ take under this rescaling? Is it the equations or the unknowns that are rescaled by D?
   (c) Gaussian elimination carried further, so that after A (assumed nonsingular) is brought to upper triangular form, additional operations ("backward elimination") are carried out so that this upper triangular matrix is made diagonal.
2. (Trefethen-Bau) Gaussian elimination $PA = LU$ can be used to compute the inverse A^{-1} of a nonsingular matrix $A \in \mathbb{R}^{n\times n}$, although it is rarely really necessary to do so.
   (a) Describe an algorithm for computing A^{-1} by solving n systems of equations, and show that the number of floating point multiplication/division operations taken by the algorithm is bounded by $Cn^3 + O(n^2)$ as $n \to \infty$. What is the best value for C?
   (b) Describe a variant of your algorithm, taking advantage of sparsity, that reduced the operation count to $cn^3 + O(n^2)$ with $c \sim C/2$.
   (c) Suppose one wishes to solve m systems of equations $Ax^{(k)} = b^{(k)}$, $k = 1, \ldots, m$, or equivalently, a block system $AX = B$ with $B \in \mathbb{R}^{n\times m}$. What is the asymptotic operation count (a function of n and m) for doing this (i) directly from the LU factorization, and (ii) with a preliminary computation of A^{-1}?
3. (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 $QAQ^T = B$, where A is a given square matrix, and B is a Hessenberg matrix.

Date: Winter 2016.

4. (Isaacson-Keller) A matrix $A = [a_{ik}] \in \mathbb{R}^{n\times n}$ is called *symmetric* if $a_{ik} = a_{ki}$ for all i, k, and is called *positive definite* if $x^T Ax \geq 0$ for all $x \in \mathbb{R}^n$, with $x^T Ax = 0$ only when $x = 0$. Suppose that $A \in \mathbb{R}^{n\times n}$ is symmetric and positive definite.
   (a) Show that $a_{ii} > 0$ for all i.
   (b) Show that $\max_i a_{ii} = \max_{i,k} |a_{ik}|$.
   (c) Let $A_k = [a_{ij}^{(k)}]$ be the matrix that enters in the k-th step of the Gaussian elimination process (with $A_1 = A$). Show that for each $k = 1, \ldots, n$, the submatrix $[a_{ij}^{(k)}]_{k\leq i,j\leq n}$ is symmetric and positive definite. Conclude that Gaussian elimination does not break down (hence in particular, that A is invertible).
   (d) Show that $a_{ii}^{(k)} \leq a_{ii}^{(k-1)}$ for $k \leq i \leq n$ and for all $k = 2, 3, \ldots, n$. Conclude that for Gaussian elimination in exact arithmetics, the growth factor is 1. Note that in exact arithmetics, the growth factor would be defined by
$$g(A) = \frac{\max_{i,j,k} |a_{ij}^{(k)}|}{\max_{i,j} |a_{ij}|}.$$
5. Assuming exact arithmetic, show that $g(A) \leq 2^{n-1}$ for any matrix $A \in \mathbb{R}^{n\times n}$, for Gaussian elimination with partial pivoting $PA = LU$.
6. (a) Let U be an upper triangular matrix with no zeroes on its diagonal. Let $\tilde{x} \in \mathbb{R}^n$ be the result of back-substitution applied to the system $Ux = b$ in floating point arithmetic (with the "machine epsilon" $\varepsilon > 0$). Show that there exists an upper triangular matrix $\tilde{U}$, such that $\tilde{U}\tilde{x} = b$ in exact arithmetics and that the entries of $\tilde{U} - U$ can be bounded in absolute value by an expression depending only on ε, n, and U. Argue that back-substitution is backward stable.
   (b) Recall that Gaussian elimination in floating point arithmetics produces matrices $\tilde{L}$ and $\tilde{U}$, where $\tilde{L}$ is lower triangular with unit diagonal and $\tilde{U}$ is upper triangular, satisfying
$$\|\tilde{L}\tilde{U} - A\|_\infty \leq \frac{3ng\varepsilon}{(1-\varepsilon)^2}\|A\|_\infty.$$
   Turn this into the following bound
$$\|\tilde{L}\tilde{U} - A\| \leq C_n g\varepsilon\|A\|, \qquad \text{for all small } \varepsilon,$$
   where $\|\cdot\|$ is the matrix norm induced by the Euclidean norm in $\mathbb{R}^n$. In particular, try get a near-optimal value for the constant C_n.
   (c) By combining the preceding two results, perform a backward error analysis of the Gaussian elimination process for solving the equation $Ax = b$. That is, complete the analysis we did in class by taking into account the round-off errors of the forward elimination (solution of $\tilde{L}y = b$) and back substitution (solution of $\tilde{U}x = y$).
7. 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

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

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.