Solutions for Homework 3 Foundations of Computational Math 1 Fall 2012

Problem 3.1

Suppose $A \in \mathbb{R}^{n\times n}$ is a nonsymmetric nonsingular diagonally dominant matrix with the following nonzero pattern (shown for $n = 6$)

$$\begin{pmatrix} * & * & * & * & * & * \\ * & * & 0 & 0 & 0 & 0 \\ * & 0 & * & 0 & 0 & 0 \\ * & 0 & 0 & * & 0 & 0 \\ * & 0 & 0 & 0 & * & 0 \\ * & 0 & 0 & 0 & 0 & * \end{pmatrix}$$

It is known that a diagonally dominant (row or column dominant) matrix has an LU factorization and that pivoting is not required for numerical reliability.

3.1.a. Describe an algorithm that solves $Ax = b$ in a stable manner as efficiently as possible.

3.1.b. Given that the number of operations in the algorithm is of the form $Cn^k + O(n^{k-1})$, where C is a constant independent of n and $k > 0$, what are C and k?

Solution:

Since the matrix is assumed to be diagonally dominant, no pivoting is required for stability and existence as was discussed in class and the text. Suppose we simply apply Gaussian elimination without pivoting. After one step of LU factorization the $n-1 \times n-1$ matrix to be factored on steps 2 to n is dense and $\frac{2}{3}n^3 + O(n^2)$ computations are needed to complete the factorization. That is, the linear combination of the first row with each row $2 \le i \le 6$ replaces all of the 0 elements with nonzero elements:

$$M_1^{-1}A = \begin{pmatrix} * & * & * & * & * & * \\ 0 & * & * & * & * & * \\ 0 & * & * & * & * & * \\ 0 & * & * & * & * & * \\ 0 & * & * & * & * & * \\ 0 & * & * & * & * & * \end{pmatrix}$$

These new nonzeros must be eliminated on subsequent steps.

Since the matrix is diagonally dominant any permuted PAP^T, i.e., using the same general permutation on the left and right, keeps diagonal elements on the diagonal but changes the ordering. Off-diagonal elements stay in their original rows and columns but occupy different positions. Therefore, PAP^T is also diagonally dominant and can be factored without pivoting.

Now consider step one of the LU factorization where we interchange rows and columns to choose the pivot element. We do not need to do so for stability or existence as long as we choose a pivot element from the diagonal elements of the matrix, i.e., $P_i = Q_i$.

Choose

$$P_1 = \begin{pmatrix} e_n & e_2 & \dots & e_{n-1} & e_1 \end{pmatrix} = Q_1$$

i.e., interchange rows n and 1 and columns n and 1 thereby choosing α_{nn} as the pivot for step one. For example, with $n = 6$ the matrix structure for the permuted matrix is

$$P_1 A P_1^T = \begin{pmatrix} * & 0 & 0 & 0 & 0 & * \\ 0 & * & 0 & 0 & 0 & * \\ 0 & 0 & * & 0 & 0 & * \\ 0 & 0 & 0 & * & 0 & * \\ 0 & 0 & 0 & 0 & * & * \\ * & * & * & * & * & * \end{pmatrix}$$

So M_1^{-1} only needs to combine row 1 with row 6 (in general with row n) to produce a first column of 0 below the diagonal, i.e.,

$$M_1^{-1} P_1 A P_1^T = \begin{pmatrix} * & 0 & 0 & 0 & 0 & * \\ 0 & * & 0 & 0 & 0 & * \\ 0 & 0 & * & 0 & 0 & * \\ 0 & 0 & 0 & * & 0 & * \\ 0 & 0 & 0 & 0 & * & * \\ 0 & * & * & * & * & * \end{pmatrix}$$

No further pivoting is needed and each step combines row i with row n to produce ultimately

$$M_5^{-1} M_4^{-1} M_3^{-1} M_2^{-1} M_1^{-1} P_1 A P_1^T = \begin{pmatrix} * & 0 & 0 & 0 & 0 & * \\ 0 & * & 0 & 0 & 0 & * \\ 0 & 0 & * & 0 & 0 & * \\ 0 & 0 & 0 & * & 0 & * \\ 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & * \end{pmatrix}$$

Each elimination step works only on a 2×2 dense matrix and eliminates one element in the last row – this requires 1 division, 1 multiplication and 1 addition. The total number of computations is therefore $3n + O(1)$ for the factorization to compute L and U. It requires $2n + O(1)$ to apply the M_i^{-1} to b, and $3n + O(1)$ for the backsolve.

Problem 3.2

It is known that if partial or complete pivoting is used to compute $PA = LU$ or $PAQ = LU$ of a nonsingular matrix then the elements of L are less than 1 in magnitude, i.e., $|\lambda_{ij}| \leq 1$. Now suppose $A \in \mathbb{R}^{n \times n}$ is a symmetric positive definite matrix, i.e., $A = A^T$ and $x \neq 0 \rightarrow$

$x^TAx > 0$. It is known that A has a factorization $A = LL^T$ where L is lower triangular with positive elements on the main diagonal (the Cholesky factorization). Does this imply that $|\lambda_{ij}| \leq 1$? If so prove it and if not give an $n \times n$ symmetric positive definite matrix with $n > 3$ that is a counterexample and justify that it is indeed a counterexample.
Solution:

A counterexample is easily constructed by specifying a lower triangular L with positive diagonal elements and off-diagonal elements greater than 1. For example,

$$\begin{pmatrix} 4 & 8 & 8 & 8 \\ 8 & 20 & 24 & 24 \\ 8 & 24 & 36 & 40 \\ 8 & 24 & 40 & 52 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 4 & 2 & 0 & 0 \\ 4 & 4 & 2 & 0 \\ 4 & 4 & 4 & 2 \end{pmatrix} \begin{pmatrix} 2 & 4 & 4 & 4 \\ 0 & 2 & 4 & 4 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

It is easily seen that A is is a symmetric positive definite matrix since for any $x \neq 0$ we have $x^TAx = x^T(LL^T)x = (x^TL)(L^Tx) = y^Ty = \|y\|_2^2 > 0$. Therefore, since $A = LL^T$ and A is symmetric postive definite we have that L is its Cholesky factor. Clearly, we have $|\lambda_{ij}| > 1$ for some i, j.

Problem 3.3

Suppose $PAQ = LU$ is computed via Gaussian elimination with complete pivoting. Show that there is no element in e_i^TU, i.e., row i of U, whose magnitude is larger than $|\mu_{ii}| = |e_i^TUe_i|$, i.e., the magnitude of the (i, i) diagonal element of U.
Solution:

This result follows easily from two basic facts of LU factorization.

We know that e_i^TU is produced in its final form on step i of the algorithm. So, at the beginning of step i of the algorithm we have

$$A^{(i-1)} = \begin{pmatrix} U_{11} & U_{12} \\ 0 & A_{i-1} \end{pmatrix}$$

where $U_{11} \in \mathbb{R}^{i-1 \times i-1}$ and $U_{12} \in \mathbb{R}^{i-1 \times n-i+1}$ contain the first $i-1$ rows of U. The permutations P_i and Q_i are used to move an element of A_{i-1} with maximal magnitude to the pivot position. The first row of the permuted A_{i-1} contains the values of $e_i^TUe_j$ for $j \geq i$ since the eliminations on step i leave that row unchanged. Since μ_{ii} is therefore the pivot element it must be larger in magnitude that any element of A_{i-1}. In particular, it must be larger in magnitude that any $e_i^TUe_j$ for $j > i$ since they are elements of A_{i-1}.

Problem 3.4

Suppose you are computing a factorization of the $A \in \mathbb{C}^{n \times n}$ with partial pivoting and at the beginning of step i of the algorithm you encounter the active part of the matrix with the form

$$TA = A^{(i-1)} = \begin{pmatrix} U_{11} & U_{12} \\ 0 & A_{i-1} \end{pmatrix}$$

where $U_{11} \in \mathbb{R}^{i-1 \times i-1}$ and nonsingular, and $U_{12} \in \mathbb{R}^{i-1 \times n-i+1}$ contain the first $i-1$ rows of U. Show that if the first column of A_{i-1} is all 0 then A must be a singular matrix.

Solution:

The matrix A_{i-1} is clearly singular since it has linearly dependent columns. Since $A^{(i-1)}$ is a block upper triangular matrix we have

$$\det\left(A^{(i-1)}\right) = \det\left(U_{11}\right)\det\left(A_{i-1}\right) = 0$$

Since we know that

$$A^{(i-1)} = TA$$

and T is the product of a series of Gauss transforms and elementary permutation matrices, T must be nonsingular and the only way for $A^{(i-1)}$ to be singular is for A to be singular.

A more constructive proof of the singularity of $A^{(i-1)}$ is to create a vector $x \neq 0$ such that $A^{(i-1)}x = 0$. This is easily done and the construction is demonstrated by considering the case with $n = 5$. Let $i = 3$ and

$$A^{(i-1)}x = \left(\begin{array}{cc|ccc} \mu_{11} & \mu_{12} & \mu_{13} & \mu_{14} & \mu_{15} \\ 0 & \mu_{22} & \mu_{23} & \mu_{24} & \mu_{25} \\ \hline 0 & 0 & 0 & \mu_{34} & \mu_{35} \\ 0 & 0 & 0 & \mu_{44} & \mu_{45} \\ 0 & 0 & 0 & \mu_{54} & \mu_{55} \end{array}\right) \left(\begin{array}{c} \xi_1 \\ \xi_2 \\ \hline \xi_3 \\ 0 \\ 0 \end{array}\right)$$

The matrix U_{11} is nonsingular since $\mu_{11} \neq 0$ and $\mu_{22} \neq 0$.

Given any value of $\xi_3 \neq 0$ there are unique values of ξ_1 and ξ_2 such that

$$\begin{pmatrix} \mu_{11} & \mu_{12} \\ 0 & \mu_{22} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} -\xi_3\mu_{13} \\ -\xi_3\mu_{23} \end{pmatrix}$$

Hence, we have

$$A^{(i-1)}x = \left(\begin{array}{cc|ccc} \mu_{11} & \mu_{12} & \mu_{13} & \mu_{14} & \mu_{15} \\ 0 & \mu_{22} & \mu_{23} & \mu_{24} & \mu_{25} \\ \hline 0 & 0 & 0 & \mu_{34} & \mu_{35} \\ 0 & 0 & 0 & \mu_{44} & \mu_{45} \\ 0 & 0 & 0 & \mu_{54} & \mu_{55} \end{array}\right) \left(\begin{array}{c} \xi_1 \\ \xi_2 \\ \hline \xi_3 \\ 0 \\ 0 \end{array}\right) = \left(\begin{array}{c} 0 \\ 0 \\ \hline 0 \\ 0 \\ 0 \end{array}\right)$$

with $x \neq 0$ and therefore $A^{(i-1)}$ is singular. The argument clearly generalizes to any n and i.

Problem 3.5

Let x and y be two vectors in $\mathbb{R}^n$.

3.5.a. Show that given x and y the value of $\|x - \alpha y\|_2$ is minimized when

$$\alpha_{min} = \frac{x^T y}{y^T y}$$

3.5.b. Show that $x = y\alpha_{min} + z$ where $y^T z = 0$, i.e., x is easily written as the sum of two orthogonal vectors with specifed minimization properties.

Solution: Minimizing $\|x - \alpha y\|_2$ is minimized when $\|x - \alpha y\|_2^2$ is minimized. We therefore have from the calculus

$$\begin{aligned} f(\alpha) &= \|x - \alpha y\|_2^2 \\ &= \alpha^2 (y^T y) - 2(x^T y)\alpha + x^T x \\ f'(\alpha) &= 2\alpha(y^T y) - 2(x^T y) \\ f'(\alpha_{min}) &= 0 \\ \alpha_{min} &= \frac{x^T y}{y^T y} \end{aligned}$$

Now define $z = x - \alpha_{min} y$ and it follows that

$$\begin{aligned} y^T z =& y^T (x - \alpha_{min} y) \\ =& y^T x - \alpha_{min} y^T y \\ =& y^T x - \frac{x^T y}{y^T y} y^T y \\ =& 0 \end{aligned}$$

as desired.

We can also look at applying the Householder reflector approach to this simple problem to verify that it creates the same solution. We have

$$\|x - y\alpha\|_2^2 = \|Hx - Hy\alpha\|_2^2$$
$$\|b + e_1 \gamma \alpha\|_2^2, \quad \gamma = \pm \|y\|_2$$

where H is the Householder reflector that transforms y to $-e_1\gamma$, i.e., a one-step triangularization of the "matrix" y.

We have

$$Hy = -\gamma e_1, \quad \gamma = \pm\|y\|_2$$

$$H = I + \beta w w^T, \quad w = y + \gamma e_1, \quad \beta = -\frac{2}{w^T w} = -\frac{1}{\gamma^2 + \gamma\eta_1}$$

$$\eta_1 = e_1^T y, \quad w^T w = 2(\gamma^2 + \gamma\eta_1), \quad w^T y = \gamma^2 + \gamma\eta_1, \quad Hx = x + \beta w w^T x$$

$$\begin{aligned} e_1^T H x &= \xi_1 + \beta e_1^T w w^T x \\ &= \xi_1 + \beta(\eta_1 + \gamma)(x^T y + \gamma\xi_1) \\ &= \xi_1 - \frac{(\eta_1 + \gamma)(x^T y + \gamma\xi_1)}{(\gamma^2 + \gamma\eta_1)} \\ &= \xi_1 - \frac{x^T y}{\gamma} + \xi_1 = -\frac{x^T y}{\gamma} \end{aligned}$$

Therefore, the 1×1 "triangular" system we solve to get α_{min} is

$$\gamma\alpha_{min} = \frac{x^T y}{\gamma} \to \alpha_{min} = \frac{x^T y}{\gamma^2} = \frac{x^T y}{y^T y}$$

as desired.