Programming Assignment 2 Numerical Linear Algebra 1 Spring 2024

The solutions are due on Canvas by 11:59 PM on Monday, February 19, 2024

General Task

This programming assignment implements and empirically evaluates the transformation of a full-rank matrix $A \in \mathbb{R}^{n \times k}$ with $n \geq k$ to an upper triangular form, i.e.,

$$T_L^{-1} A T_R^{-1} \rightarrow \begin{pmatrix} U \\ 0 \end{pmatrix}$$

where $U \in \mathbb{R}^{k \times k}$ and nonsingular. Note A may be a rectangular matrix. Factoring such "tall" matrices is often used as a computational primitive in factoring square and nonsingular matrices, e.g., to solve $Ax = b$. You must produce code to use Gauss transforms and row/column permutations to achieve this transformation

$$P_r A P_c = LU$$

where $P_r \in \mathbb{R}^{n \times n}$ is a permutation matrix that interchanges rows, $P_c \in \mathbb{R}^{k \times k}$ is a permutation matrix that interchanges columns, $L \in \mathbb{R}^{n \times k}$ is a unit lower trapezoidal matrix, and $U \in \mathbb{R}^{k \times k}$ is a nonsingular upper triangular matrix.

This assignment empirically evaluates the predictions you make about the structure of the factors and the general performance of the method across a large set of randomly generated problems.

The Codes

1. Implement a code that computes the LU factorization of a matrix A without pivoting, using partial pivoting or using complete pivoting. The code that is capable of performing these three tasks, based on a user selection. Your code should also detect situations where the factorization may not proceed and exit gracefully. Clearly, this is the case when the candidate pivot set contains no acceptable value. In exact arithmetic this means they are all 0. Your code should also allow the detection of a set of candidate pivots that are all "too small" and warn the user. Your factorization routine, which must be a separate routine from the driver/tester routine, should accept the matrix A stored in a simple 2-dimensional array-like data structure and other relevant parameters such as n and a flag indicating what form of the factorization should be attempted, e.g., no pivoting, partial pivoting, or complete pivoting. The routine should return the matrices L and U stored within the array that contained A on input, i.e., you should implement the in-place algorithm strategy described in the class notes and lectures.

You should also keep a copy of A in an additional data structure for correctness checking. The 1 values on the diagonal of L, the 0 values in the upper triangular portion of L and the 0 values in the lower triangular portion of U should not be stored at any point in the code. The routine should also return the permutation matrices P_r and/or P_c appropriate to the pivoting strategy used. These permutation matrices should not be stored as full matrices within any array. The matrices, or equivalently the set of elementary permutations that are their factors, can be represented by at most n integers each.

2. Your code should be based on subroutines similar to the ones discussed below. You are free to change or augment the organization if you think it necessary but make your design clear in your solutions.

3. You will also need various test routines designed to evaluate and validate the correctness of the code and accomplish the tasks described below. You may code in any compiled and typed language you wish although C, C++, Julia, and Fortran are preferred. In all cases, however, you may not use standard libraries such as LAPACK or built-in matrix routines for pieces of your routines implementing the computations described above.

Suggested Subroutines for the factorization:

1. Initialization: This routine generates the matrix to be factored and places it in the array to be used during the factorization as well as another array used to preserve the matrix for use when assessing the accuracy of the factorization. It shoud also determine the position of the first desired pivot element (r_1, c_1) based on the pivoting scheme used.

2. Permutation: This routine takes the desired pivot element's indices, defines the appropriate elementary permutations P_{r_i} and P_{c_i}, stores their parameters appropriately, and applies them to the current active portion of the matrix.

3. FormGauss: This routine examines the appropriate part of the permuted active part of the matrix and determines the parameters determining M_i. You may find it convenient to return the parameters in a work vector and then place the values in the appropriate positions of the array in which A is being transformed and the elements of L and U are created.

4. ApplyGauss: This routine applies M_i^{-1} to the active part of the matrix **and determines the next pivot element in the updated active part based on the pivoting strategy used.** The indices (r_{i+1}, c_{i+1}) should be returned to the calling routine for use in the next step of elimination. Note that you should be able to determine the indices in the same pass through the active part of the matrix that performs the rank-1 update defined by M_i^{-1}, i.e., you should not update the matrix on one pass through and then make a second pass over the candidate pivot set to determine the $i+1$ pivot element. Clearly, this step should be surpressed when applying M_{n-1}^{-1}.

Routines to support the evaluation of the factorization:

1. Implement a routine that accepts as input the 1-dimensional arrays specifying P_r and P_c and applies them a matrix stored in a 2-dimensional arrays and returns the result in a separate 2-dimensional array, i.e., $M_1 \leftarrow P_r M_2 P_c$.

2. Implement a matrix multiplicaiton routine that accepts as input the 2-dimensional array containing the information specifying the L and U matrices and returns in a separate 2-dimensional array the product $M = LU$ **or** the product $M = |L||U|$ based on user selection indicated by an input flag. Note that these matrix multiplications must use the information specifying L and U within the data structure. It **must not** expand them into separate arrays that include the 1 and 0 values that are not stored in the input array.

3. The test routines should include the computation of $\|W\|$ where W is a matrix. The norms used should be finite in computation, i.e., $\|W\|_1$, $\|W\|_\infty$, or $\|W\|_F$. You may use library routines to compute $\|W\|_2$ if you wish but you are definitely not expected to generate the code for the 2-norm.

Library Codes

You may use libaries and external routines in your test routines to generate solutions for comparisons, to generate historgrams, graphs and any other useful summary display mechanisms. Make sure when using library routines as part of your empirical analysis, you carefully check, e.g., the P_r, P_c, L and U generated by your routines and those generated by the library, e.g., MATLAB. These factors are not unique given that pivoting choices are not unique in general.

Metrics

There are several important metrics to use when assessing the code's correctness. These metrics should be computed in double precision espeicially if you have run your routines in single precision. Some suggestions follow:

1. When comparing matrices use more than one matrix norm, e.g., the finitely computable ones, $\|M\|_1$, $\|M\|_\infty$ and $\|M\|_F$.

2. Check the factorization accuracy

$$\frac{\|P_r A P_c - LU\|}{\|A\|}$$

where $\|A\| \geq 1$, i.e., relative error for large A.

3. If you have access to a standard library you may also use the results of its LU factorization algorithms. However, as noted above care must be taken since details of pivoting strategies may yield differences in the factors and permutations.

4. You should compute the growth factor

$$\gamma_\epsilon = \frac{\| \ |L_\epsilon||U_\epsilon| \ \|}{\|A\|}$$

where $L_\epsilon U_\epsilon = P_r A P_c$ is the **computed** factorization of $P_r A P_c$ using the selected pivoting strategy. This will be important for the next assignment that will assess the numerical stability of solving systems but it is also useful for checking the correctness of structured problems such as the one in the study questions with the large growth of elements.

Generation of Test Problems

A key consideration in this assignment is the generation of test problems. They must have full rank, i.e., linearly independent columns.

Some suggestions follow:

1. The matrices should be generated using double precision storage and computation. You can then coerce the type to single when storing the matrix before calling your single precision code.

2. Generate L and U so they are nonsingular unit lower trapezoidal and upper triangular matrices respectively. Evaluate their product to define A. By their structure these matrices have linearly independent columns and therefore their product has full-rank.

   This is useful for both small and large values of n and k. For small values you can also constrain L and U to have integer values so A will have integer values. Take care with the magnitude of the elements of L and U. Nicely conditioned problems tend to be specified by off-diagonal elements in L smaller than 1 in magnitude and diagonal elements in U that are not too small compared to off-diagonal elements in U.

3. Remember even if A is generated from L and U when you run your routine with partial or complete pivoting nontrivial permutation matrices P_r and/or P_c may result and you may return $\tilde{L}$ and $\tilde{U}$ such that $P_r A P_c = \tilde{L}\tilde{U}$.

4. Randomly generated matrices tend to be nonsingular and reasonably conditioned. You can make sure any matrix is nonsingular by adding to the diagonal elements until the matrix is diagonally dominant by rows, columns or both.[1] This guarantees success of

[1] A matrix is strictly diagonally dominant by rows (columns) if the magnitude of each diagonal element is strictly larger than the sum of the magnitudes of all off-diagonal elements in the same row (column). A matrix may of course be diagonally dominant by rows and columns simultaneously, e.g., the identity.

the factorization if run without pivoting. As noted above, if you allow pivoting the routine may so do even for a diagonally dominant matrix depending on the form of dominance and the pivoting strategy used. It is also useful to note that after generating such a matrix, A you can apply random permutations $\tilde{P}$ and $\hat{P}$ to generate a new test matrix $\tilde{A} = \tilde{P}A\hat{P}$ that will not be diagonally dominant but will still be nonsingular. Of course, $\tilde{P}$ and $\hat{P}$ are not necessarily the permuations that will be generated by applying partial or complete pivoting to $\tilde{A}$.

5. You can generate a symmetric positive definite $A \in \mathbb{R}^{n\times n}$ from a lower triangular $\tilde{L}$ with positive diagonal elements (not necessarily 1) via $A = \tilde{L}\tilde{L}^T$. A is nonsingular by definition and factorization will succeed without pivoting. Note that your code will still produce an LU factorization since your factorization routine is not designed to exploit symmetry. However, there is a relationship between L, U and $\tilde{L}$. This is probed in the first set of structured factorization tasks.

6. Matrices with known structures that influence the structure or magnitude pattern of their factorizations are also useful. This is especially true if the patterns scale in a known way with n. Recall the example in the notes/homework problem that has large elements in U. Consider what should happen when no pivoting is used, partial pivoting and complete pivoting for various n values. Also, nonzero patterns such as banded matrices for A should produce specific nonzero patterns in L and U. Structure is the subject of the first set of empirical tasks discussed below.

7. Make sure that the matrices you use to check pivoting **actually require some pivoting when factored**.

8. You should run a range of problem sizes for each algorithm and problem type you evaluate.

9. Do not simply report the the factors or accuracy of the factorization for a small number of small systems. Think about how you would report the results of testing each of the routines with many matrices including those of sizes too large to display for any useful effect.

Empirical Tasks Set 1 : Structure

Consider the following structured matrices, predict the structure of their factors and factorization, and verify them empirically. Your solutions must include an explanation of your observations and justification for the conclusion that your predictions are correct.

1. Consider a matrix $A \in \mathbb{R}^{n\times n}$ that is is diagonal with positive elements, i.e., $\alpha_{ij} = e_i^T A e_j = 0$ for all $i \neq j$ and $\alpha_{ii} > 0$ for $1 \leq i, j \leq n$. For example, let $\alpha_{ii} = i$ or $\alpha_{ii} = n - i + 1$

$$n = 5 \rightarrow A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} 5 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

   Consider factoring with no pivoting, partial row pivoting and complete pivoting. What can be said about the L, U, P_r and P_c factors as a function of your choice of structure in the positive diagonal elements? What is the growth factor for your chosen problems and the pivoting choices?

2. Consider a matrix $A \in \mathbb{R}^{n\times n}$ that is is antidiagonal with positive elements, i.e., $\alpha_{1,n} > 0, \alpha_{2,n-1} > 0, \ldots, \alpha_{n-1,2} > 0, \alpha_{n,1} > 0$. For example,

$$n = 5 \rightarrow A = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 5 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} 0 & 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

   Consider factoring with no pivoting, partial row pivoting and complete pivoting. What can be said about the L, U, P_r and P_c factors as a function of your choice of structure in the positive antidiagonal elements? What is the growth factor for your chosen problems and pivoting choices?

3. Consider a matrix $A \in \mathbb{R}^{n\times n}$ that is is the sum of a diagonal matrix and an antidiagonal matrix with positive elements, i.e., and X nonzero pattern. Consider factoring with no pivoting, partial row pivoting and complete pivoting. What can be said about the L, U, P_r and P_c factors as a function of your choice of structure in the positive antidiagonal elements? If complete pivoting is there any structure that follows for a particular choice of pivot elements?

4. Consider a matrix $A \in \mathbb{R}^{n\times n}$ that is unit lower triangular and the elements in the strict lower part all have magnitude less than 1, i.e., $|\lambda_{ij}| < 1$ for $i > j$, $\lambda_{ij} = 0$ for $i < j$, and $\lambda_{ii} = 1$. Consider factoring with no pivoting, partial row pivoting and complete

pivoting. What can be said about the L, U, P_r and P_c factors? (Note that your code does not know that A is unit lower triangular and will eliminate the elements below the diagonal by applying Gauss transforms.)

5. Consider a lower triangular matrix A again but this time let the diagonal elements be positive and not 1 and the elements in the strictly lower triangular part be larger than 1, e.g.,

$$A = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 3 & 2 & 0 & 0 & 0 \\ 4 & 3 & 2 & 0 & 0 \\ 5 & 4 & 0 & 2 & 0 \\ 6 & 5 & 4 & 3 & 2 \end{pmatrix}.$$

   Consider factoring with no pivoting, partial row pivoting and complete pivoting. What can be said about the L, U, P_r and P_c factors?

6. Consider a matrix $A \in \mathbb{R}^{n \times n}$ that is tridiagonal with all elements on the main diagonal and the first super and subdiagonals nonzero, i.e., $\alpha_{ii} \neq 0$, $\alpha_{i+1,i} \neq 0$, and $\alpha_{i,i+1} \neq 0$. Of course, these must be such that the matrix is nonsingular.

   Suppose A is, additionally, strictly diagonally dominant by rows and columns. Consider factoring with no pivoting, partial row pivoting and complete pivoting. What can be said about the L, U, P_r and P_c factors?

7. Consider a matrix $A \in \mathbb{R}^{n \times n}$ with

   - $\alpha_{ij} = e_i^T A e_j = -1$ when $i > j$, i.e., all elements strictly below the diagonal are -1;
   - $\alpha_{ii} = e_i^T A e_i = 1$, i.e., all elements on the diagonal are 1;
   - $\alpha_{in} = e_i^T A e_n = 1$, i.e., all elements in the last column of the matrix are 1;
   - all other elements are 0

   e.g., for $n = 4$ we have

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

   Consider factoring with no pivoting, partial row pivoting and complete pivoting. What can be said about the L, U, P_r and P_c factors? What is the growth factor for the different pivoting choices?

8. Suppose $A \in \mathbb{R}^{n \times n}$ is a symmetric positive definite generated from a lower triangular $\tilde{L}$ with positive diagonal elements (not necessarily 1) via $A = \tilde{L}\tilde{L}^T$. A is nonsingular by definition and factorization will succeed without pivoting. Note that your code will still produce an LU factorization since your factorization routine is not designed to exploit symmetry. What is the relationship between L, U and $\tilde{L}$?

Empirical Tasks Set 2 : General Trends

This set requries the generation of a large set of problems grouped and empirically analyzed by problem size n and, if appropriate the class of matrix problem considered.

For each problem size and class of problem, generate many example problems and evaluate the various metrics discussed earlier You should present your results in a form appropriate to characterize these metrics over a large data set, i.e., too large to look at each problem individually. This can be done, for example, by graphs and histograms. The latter is particularly useful for detecting outliers in the performance such as large factorization error. These outliers can be discussed in more detail and explained if you wish.

You should also include empirical evidence of the $O(n^3)$ complexity giving the appropriate trend in time required to factor a matrix. It is recommended that this is done with a matrix that does not require pivoting such as a matrix that is strictly diagonally dominant by row and column.

Empirical Tasks Set 3 : Rectangular Matrices

All of the constructions above were square nonsingular matrices. You should include a few rectangular examples in your tests to demonstrate that your code works correctly.