MATH 2418 Linear Algebra. Week 5

Dr. James W. Miller

September 20 and 22, 2022

Summary of this Week's Goals

This week we will cover the LU decompositions of a square matrix presented in Sections 2.6 (Elimination = Factorization: $A = LU$) and 2.7 (Transposes and Permutations). By the end of the week, you should know how to invert the elementary matrices that perform row operations in the elimination process and express a matrix as a product of factors which are lower triangular and upper triangular ($A = LU$), diagonal ($A = LDU$), or a permutation matrix ($PA = LDU$). You will learn some definitions and facts related to the transpose of a matrix and some definitions and properties of symmetric matrices.

Announcements

- Your first midterm exam will be on Thursday, October 4, 8:30-9:45pm. It will cover sections 1.1 through 2.6.

2.5 Inverse Matrices (One More Example)

Gauss-Jordan Elimination

- Let's apply the Gauss-Jordan Elimination Method to find the inverse of the following matrix:

$$A = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 2 & 3 \\ 1 & 1 & 4 \end{bmatrix}$$

  - First, we write the augmented matrix:

$$\begin{bmatrix} A & I \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ -1 & 2 & 3 & 0 & 1 & 0 \\ 1 & 1 & 4 & 0 & 0 & 1 \end{bmatrix}$$

  - Next, pivot on the upper-left coefficent and produce zeros in the first column below the first pivot. This is equivalent to multiplying the augmented matrix on the left by a matrix we might call E_1. It is similar to an identity matrix but contains multipliers in the first column.

$$E_1 \begin{bmatrix} A & I \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 2 & 4 & 1 & 1 & 0 \\ 0 & 1 & 3 & -1 & 0 & 1 \end{bmatrix}, E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

  - Next, pivot on the second diagonal element and produce zeros in the second column below the 2. This is equivalent to multiplying the system on the left by a matrix we may call E_2.

$$E_2 E_1 \begin{bmatrix} A & I \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 2 & 4 & 1 & 1 & 0 \\ 0 & 0 & 1 & -3/2 & -1/2 & 1 \end{bmatrix}, E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1/2 & 1 \end{bmatrix}$$

– Next, pivot on the third diagonal element and produce zeros in the second column above the 1. This is equivalent to multiplying the system on the left by a matrix we may call E_3.

$$E_3E_2E_1\begin{bmatrix}A & I\end{bmatrix} = \begin{bmatrix}1 & 0 & 0 & 5/2 & 1/2 & -1\\ 0 & 2 & 0 & 7 & 3 & -4\\ 0 & 0 & 1 & -3/2 & -1/2 & 1\end{bmatrix}, E_3 = \begin{bmatrix}1 & 0 & -1\\ 0 & 1 & -4\\ 0 & 0 & 1\end{bmatrix}$$

– Our last step will be to multiply each row by the reciprocal of the pivot on each row to produce an identity matrix on the left side of the augmented system. This is equivalent to multiplying on the left by a diagonal matrix we may call D.

$$DE_3E_2E_1\begin{bmatrix}A & I\end{bmatrix} = \begin{bmatrix}1 & 0 & 0 & 5/2 & 1/2 & -1\\ 0 & 1 & 0 & 7/2 & 3/2 & -2\\ 0 & 0 & 1 & -3/2 & -1/2 & 1\end{bmatrix}, D = \begin{bmatrix}1 & 0 & 0\\ 0 & 1/2 & 0\\ 0 & 0 & 1\end{bmatrix}$$

– We have expressed the system in the following form:

$$DE_3E_2E_1\begin{bmatrix}A & I\end{bmatrix} = \begin{bmatrix}I & A^{-1}\end{bmatrix}, A^{-1} = \begin{bmatrix}5/2 & 1/2 & -1\\ 7/2 & 3/2 & -2\\ -3/2 & -1/2 & 1\end{bmatrix}$$

– We may also check that

$$A^{-1} = DE_3E_2E_1$$

That is, we have effectively multiplied $\begin{bmatrix}A & I\end{bmatrix}$ by A^{-1} to obtain $\begin{bmatrix}I & A^{-1}\end{bmatrix}$.

2.6 Elimination = Factorization: $A = LU$

Factoring a Square Non-singular Matrix

- Let's consider the previous example again, but use elimination to triangularize the matrix rather than calculation its inverse. We will perform row operations only on the matrix A, not an augmented matrix.

$$A = \begin{bmatrix}1 & 0 & 1\\ -1 & 2 & 3\\ 1 & 1 & 4\end{bmatrix}$$

– First, pivot on the upper-left coefficent and produce zeros in the first column below the first pivot. This is equivalent to multiplying the augmented matrix on the left by a matrix we might call E_1. It is similar to an identity matrix but contains multipliers in the first column.

$$E_1A = \begin{bmatrix}1 & 0 & 1\\ 0 & 2 & 4\\ 0 & 1 & 3\end{bmatrix}, E_1 = \begin{bmatrix}1 & 0 & 0\\ 1 & 1 & 0\\ -1 & 0 & 1\end{bmatrix}$$

– Next, pivot on the second diagonal element and produce zeros in the second column below the 2. This is equivalent to multiplying the system on the left by a matrix we may call E_2.

$$E_2E_1A = \begin{bmatrix}1 & 0 & 1\\ 0 & 2 & 4\\ 0 & 0 & 1\end{bmatrix}, E_2 = \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & -1/2 & 1\end{bmatrix}$$

– The right hand side of this system is upper-triangular, so we can stop with the elimination steps now and write:

$$E_2 E_1 A = U,$$

where

$$U = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 4 \\ 0 & 0 & 1 \end{bmatrix}$$

– Next, we observe that the matrices E_1 and E_2 are both invertible. Their inverses are shown below, and have a simple intuitive form. Essentially, the matrices add a multiple of the pivot row to another row. This operation is inverted by subtracting that same multiple of the pivot row from the other row.

$$E_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

$$E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1/2 & 1 \end{bmatrix}$$

– Another remarkable observation that we can make at this point is that the product of the inverses also has a very simple form. Essentially, the product of the inverses simply collects all of the non-zero elements of each inverse into a single matrix. The product is lower-triangular, so we can call it L.

$$L = E_1^{-1} E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 1/2 & 1 \end{bmatrix}$$

Note: The product $E_2 E_1$ is not so simple because the matrices are multiplied in an order that does not produce a similar result. The simplicity of the product of the inverses comes about because of the order in which they are multiplied and the structure of the zeros in columns below the diagonal elements.

– From the relation $E_2 E_1 A = U$, we derive the relation $A = E_1^{-1} E_2^{-1} U = LU$.

$$A = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 2 & 3 \\ 1 & 1 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 1/2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 4 \\ 0 & 0 & 1 \end{bmatrix} = LU$$

– The matrix L has only ones on its main diagonal. The matrix U does not. We can make L and U more consistent in their appearance by rewriting U as the product of a diagonal matrix and an upper-triangular matrix. When we factor the diagonal elements of U into the matrix D, we must also divide each row of U by the diagonal element we factored from each row. We call this form an LDU decomposition of the matrix A.

$$A = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 2 & 3 \\ 1 & 1 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 1/2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} = LDU$$

Observations About the LU and LDU Factorizations of a Square Matrix

- The LU and LDU factorizations of a square matrix A are possible when elimination produces n non-zero pivots (meaning the columns of A are linearly independent, the matrix A is invertible, etc.) and the elimination process requires no row exchanges. Row exchanges will break the LU factorization. We'll handle this case in the next section using permutation matrices.
- After factoring A in the form LU, the matrix equation $A\mathbf{x} = \mathbf{b}$ becomes $LU\mathbf{x} = \mathbf{b}$. If we then let $\mathbf{c} = U\mathbf{x}$, we can break $A\mathbf{x} = \mathbf{b}$ into two triangular systems, $L\mathbf{c} = \mathbf{b}$ and $U\mathbf{x} = \mathbf{c}$. We solve the system $A\mathbf{x} = \mathbf{b}$ via the following steps:
  - Factor A as the product of L and U.
  - Solve $L\mathbf{c} = \mathbf{b}$ for the vector $\mathbf{c}$ using forward substitution.
  - Solve $U\mathbf{x} = \mathbf{c}$ for the vector $\mathbf{x}$ using back substitution.
- If you need solve $A\mathbf{x} = \mathbf{b}$ repetitively for a lot of different values of $\mathbf{b}$, the LDU factorization can help you do this very efficiently.
  - Factor A as the product of L, D and U. The important parts of L, D and U (the parts which aren't certain to be ones and zeros) can be stored in the same array of computer memory where A was once stored. Once A is factored, it no longer needs to be stored in its original form. The system becomes $LDU\mathbf{x} = \mathbf{b}$ and we can solve it via the algorithm below, where now $\mathbf{d} = DU\mathbf{x}$ and $\mathbf{c} = U\mathbf{x}$.
  - Solve $L\mathbf{d} = \mathbf{b}$ for the vector $\mathbf{c}$ using forward substitution.
  - Solve $\mathbf{d} = D\mathbf{c}$ by simply dividing each element of $\mathbf{d}$ by the corresponding diagonal element of D to obtain $\mathbf{c}$.
  - Solve $U\mathbf{x} = \mathbf{c}$ for the vector $\mathbf{x}$ using back substitution.

Example: Solving $A\mathbf{x} = \mathbf{b}$ Using the LU Factorization

- Solve the system $A\mathbf{x} = \mathbf{b}$ using the LU factorization, where

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 0 & -1 \\ 1 & 1 & 0 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

- Factor A as the product of L and U.
  - First, pivot on the upper-left coefficent and produce zeros in the first column below the first pivot.

$$E_1 A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & -6 & -4 \\ 0 & -1 & -1 \end{bmatrix}, E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

  - Next, pivot on the second diagonal element and produce zeros in the second column below the -6.

$$E_2 E_1 A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & -6 & -4 \\ 0 & 0 & -1/3 \end{bmatrix}, E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1/6 & 1 \end{bmatrix}$$

  - We can now write

$$L = E_1^{-1} E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 1 & 1/6 & 1 \end{bmatrix}$$

and

$$U = \begin{bmatrix} 1 & 2 & 1 \\ 0 & -6 & -4 \\ 0 & 0 & -1/3 \end{bmatrix}$$

- Solve $L\mathbf{c} = \mathbf{b}$ for the vector $\mathbf{c}$ using forward substitution.

$$\begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 1 & 1/6 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

  - The first row gives $c_1 = 1$.
  - The second row gives $c_2 = -1 - 3c_1 = -1 - 3(1) = -4$.
  - The third row gives $c_3 = 1 - c_1 - \frac{1}{6}c_2 = 1 - 1 - \frac{1}{6}(-4) = \frac{2}{3}$.

- Solve $U\mathbf{x} = \mathbf{c}$ for the vector $\mathbf{x}$ using back substitution.

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & -6 & -4 \\ 0 & 0 & -1/3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -4 \\ 2/3 \end{bmatrix}$$

  - The third row gives $x_3 = (-3)\frac{2}{3} = -2$.
  - The second row gives $x_2 = -\frac{1}{6}\left[-4 + 4x_3\right] = -\frac{1}{6}\left[-4 + 4(-2)\right] = -\frac{1}{6}\left[-12\right] = 2$.
  - The first row gives $x_1 = 1 - 2x_2 - x_3 = 1 - 2(2) - (-2) = -1$.

- The solution is $\mathbf{x} = (x_1, x_2, x_3) = (-1, 2, -2)$.

2.7 Transposes and Permutations

The Transpose of a Matrix

- An $m \times n$ matrix A can be transposed to produce an $n \times m$ matrix which is denoted A^T (called "A-transpose").
- The rows of A are the columns of A^T and the columns of A are the rows of A^T.
- The ij-th element of A becomes the ji-th element of A^T.

$$\left(A^T\right)_{ij} = A_{ji}$$

- Example:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 4 \end{bmatrix}, A^T = \begin{bmatrix} 1 & -1 \\ 2 & 0 \\ 3 & 4 \end{bmatrix}$$

Properties of the Transpose Operator

- Sum: $(A+B)^T = A^T + B^T$
- Scalar Product: $(cA)^T = cA^T$
- Double-Transpose: $\left(A^T\right)^T = A$
- Matrix-Vector Product: $(A\mathbf{x})^T = \mathbf{x}^T A^T$

$$A\mathbf{x} = \begin{bmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$$

$$(A\mathbf{x})^T = x_1\mathbf{a}_1^T + x_2\mathbf{a}_2^T + \cdots + x_n\mathbf{a}_n^T = \begin{bmatrix}x_1 & x_2 & \cdots & x_n\end{bmatrix}\begin{bmatrix}\mathbf{a}_1^T\\\mathbf{a}_2^T\\\vdots\\\mathbf{a}_n^T\end{bmatrix} = \mathbf{x}^T A^T$$

- Matrix Product: $(AB)^T = B^T A^T$

$$\begin{aligned}(AB)^T &= \left(A\begin{bmatrix}\mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n\end{bmatrix}\right)^T = \begin{bmatrix}A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_n\end{bmatrix}^T\\ &= \begin{bmatrix}(A\mathbf{b}_1)^T\\(A\mathbf{b}_2)^T\\\vdots\\(A\mathbf{b}_n)^T\end{bmatrix} = \begin{bmatrix}\mathbf{b}_1^T A^T\\\mathbf{b}_2^T A^T\\\vdots\\\mathbf{b}_n^T A^T\end{bmatrix} = \begin{bmatrix}\mathbf{b}_1^T\\\mathbf{b}_2^T\\\vdots\\\mathbf{b}_n^T\end{bmatrix} A^T = B^T A^T\end{aligned}$$

- Inverse: $\left(A^{-1}\right)^T = \left(A^T\right)^{-1}$, provided A^{-1} exists.

$$\begin{aligned}A^T\left(A^{-1}\right)^T &= \left(A^{-1}A\right)^T = I^T = I\\ \left(A^{-1}\right)^T A^T &= \left(AA^{-1}\right)^T = I^T = I\end{aligned}$$

  The above equations show that $\left(A^{-1}\right)^T$ meets the requirement to be an inverse of A^T. That is, $\left(A^{-1}\right)^T = \left(A^T\right)^{-1}$.

- Inner Product of Column Vectors: $\mathbf{x}^T\mathbf{y} = c$ (a scalar)

$$\begin{bmatrix}x_1 & x_2 & \cdots & x_n\end{bmatrix}\begin{bmatrix}y_1\\y_2\\\vdots\\y_n\end{bmatrix} = x_1y_1 + x_2y_2 + \cdots + x_ny_n = c$$

- Outer Product of Column Vectors: $\mathbf{x}\mathbf{y}^T = A$ (a matrix with linearly dependent columns)

$$\mathbf{x}\mathbf{y}^T = \mathbf{x}\begin{bmatrix}y_1 & y_2 & \cdots & y_n\end{bmatrix} = \begin{bmatrix}y_1\mathbf{x} & y_2\mathbf{x} & \cdots & y_n\mathbf{x}\end{bmatrix} = A$$

  All of the columns of the outer product are a scalar multiple of $\mathbf{x}$, so the columns are linearly dependent.

Symmetric Matrices

- A square matrix S is said to be symmetric if $S = S^T$. Each element of the matrix satisfies $s_{ij} = s_{ji}$
- Examples:

$$S = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix}, I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, D = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}, B = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}$$

- For any $m \times n$ matrix A, A^TA is a symmetric $n \times n$ matrix and AA^T is a symmetric $m \times m$ matrix.
- If a symmetric matrix S is factored into LDU with no row exchanges, then $U = L^T$, so $S = LDL^T$.

Permutation Matrices

- A matrix P containing the rows (or columns) of the identity matrix in any order is called a permutation matrix.
- There are $n!$ permutation matrices having the size $n \times n$.
- Permutation matrices are classified as even or odd depending on how many row swaps are required to produce them. There are an equal number of even and odd permutation matrices of any given size.
- The two 2×2 permutation matrices are:

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, P_{21} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

  Both of these are symmetric and are their own inverses. The matrix I is an even permutation (no row swaps) and the matrix P_{21} is odd (one row swap).

- The six 3×3 permutation matrices are:

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, P_{21} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, P_{31} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, P_{32} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix},$$
$$P_{32}P_{21} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, P_{21}P_{32} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

  - The matrices I, P_{21}, P_{31} and P_{32} are symmetric and are their own inverses.
  - The other two 3×3 permutation matrices may be expressed as the product of two row-swap permutation matrices. They are not symmetric, and they are inverses of each other.
  - The matrices I, $P_{32}P_{21}$, and $P_{21}P_{32}$ are even permutations since they involve zero or two row swaps. The matrices P_{21}, P_{31} and P_{32} are odd because they each involve one row swap.

- For any permutation matrix, $P^{-1} = P^T$.
- Multiplying the matrix A on the left by a permutation matrix rearranges the rows of A. Multiplication by P on the right rearranges the columns.

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} d & e & f \\ g & h & i \\ a & b & c \end{bmatrix}$$

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} c & a & b \\ f & d & e \\ i & g & h \end{bmatrix}$$

- When a matrix A requires row swaps for elimination to produce n non-zero pivots, the row swaps may be done first, before any elimination steps are performed. After swapping rows, it becomes possible to produce an LDU factorization of PA, the matrix resulting from the necessary row swaps. Thus, we may write $PA = LDU$.