Matrices and Gaussian Elimination (Part-2)

P. Sam Johnson

National Institute of Technology Karnataka (NITK)
Surathkal, Mangalore, India

Introduction

We have discussed that triangular system is easily solvable.

We shall discuss some procedures to convert a linear system into "two triangular system", called triangular factorization.

We shall also discuss some special matrices and applications.

Inner Product

Multiplication of a row matrix $\begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix}$ and a column matrix $\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ (of matching lengths) prodcues a single number :

$$\begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} := x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

This single quantity is called **the inner product of two vectors**, denoted by $\langle x, y \rangle$.

Matrix-Vector Product

Multiplying a matrix by a vector : **a row at a time.**

Each row of the matrix combines with the vector to give a component of the product. **There are n inner products when there are n rows.**

$$Ax := \begin{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\ \begin{pmatrix} a_{21} & a_{22} & \dots & a_{2n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\ \vdots \\ \begin{pmatrix} a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{pmatrix}.$$

By rows : $Ax = \begin{pmatrix} 1 & 1 & 6 \\ 3 & 0 & 3 \\ 1 & 1 & 4 \end{pmatrix} \begin{pmatrix} 2 \\ 5 \\ 0 \end{pmatrix} = \begin{pmatrix} 1.2 + 1.5 + 6.0 \\ 3.2 + 0.5 + 3.0 \\ 1.2 + 1.5 + 4.0 \end{pmatrix} = \begin{pmatrix} 7 \\ 6 \\ 7 \end{pmatrix}.$

Matrix-Vector Product

The **second way** is equally important. In fact it is more important.

Multiplying a matrix by a vector : **a column at a time.**

The product Ax is found at all once, and it is combination of the three columns of A.

By columns : $$Ax = 2\begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix} + 5\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + 0\begin{pmatrix} 6 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 7 \\ 6 \\ 7 \end{pmatrix}.$$

The answer is twice column 1 plus 5 times column 2.

It corresponds to "the column picture" of the linear system $Ax = b$.

We shall discuss the concept of "columun picture" later.

Matrix Notation for Individual Entries in a Matrix A

- The entry in the ith row and jth column is always denoted by a_{ij}.
- The first subscript gives the row number and the second subscript indicates the column.
- If A is an $m \times n$ matrix, then the index i goes 1 to m - there are m rows ; and the index j goes from 1 to n - there are n columns.
- Altogether, the matrix has mn entries, forming a rectangular array, and a_{mn} is the lower right corner.

Matrix Notation for Individual Entries in a Matrix A

- $\sum_{j=1}^{n} a_{ij}x_j$ is the ith component of Ax, formed the inner product of ith row of A with x. This sum takes us along the ith row of A, forming its inner product with x.
- The length of the rows (the number of columns in A) must match the length of x.
- An $m \times n$ matrix multiplies an n-dimensional vector (and produces an m-dimensional vector).
- Summations are simpler to work with than writing everything out in full, but they are not as good as matrix notation itself.
- Why is matrix notation preferred ?
  We want to get on with the connection between matrix multiplication and Guassian elimination.

Matrix-Matrix Product

Four different ways to look at matrix multiplication :

1. **Inner Product (Entry-wise) :** Each entry AB is the product of a row and a column :

$$(AB)_{ij} = \text{ row } i \text{ of } A \text{ times column } j \text{ of } B.$$

2. **Row Picture (Row-wise) :** Each row of AB is the product of a row and a matirx:

$$\text{row } i \text{ of } AB = \text{ row } i \text{ of } A \text{ times } B.$$

That is, if $\mathrm{r}_1, \mathrm{r}_2, \ldots, \mathrm{r}_m$ are the rows of A, then $AB = \begin{pmatrix} \mathrm{r}_1 B \\ \mathrm{r}_2 B \\ \vdots \\ \mathrm{r}_m B \end{pmatrix}$.

3. **Column Picture (column-wise) :** Each column of AB is the product of a matrix and a column :

$$\text{column } j \text{ of } AB = A \text{ times column } j \text{ of } B.$$

That is, if $\mathrm{c}_1, \mathrm{c}_2, \ldots, \mathrm{c}_n$ are the columns of B, then

$$AB = \begin{pmatrix} A\mathrm{c}_1 & A\mathrm{c}_2 & \cdots & A\mathrm{c}_n \end{pmatrix}.$$

Matrix-Matrix Product

4. To discuss the fourth method of matrix-matrix multiplication, we define **outer product :**

   Multiplication of a column matrix $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}$ and a row matrix $y = \begin{pmatrix} y_1 & y_2 & \dots & y_n \end{pmatrix}$ (may of different lengths) prodcues a matrix which is called **the outer product of two vectors**, denoted by xy^T :

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \begin{pmatrix} y_1 & y_2 & \dots & y_n \end{pmatrix} := \begin{pmatrix} x_1y_1 & x_1y_2 & \dots & x_1y_n \\ x_2y_1 & x_2y_2 & \dots & x_2y_n \\ \vdots & \vdots & \vdots & \vdots \\ x_my_1 & x_my_2 & \dots & x_my_n \end{pmatrix}.$$

   If $\mathrm{c}_1, \mathrm{c}_2, \dots, \mathrm{c}_p$ are the columns of $A_{m\times p}$ and $\mathrm{r}_1, \mathrm{r}_2, \dots, \mathrm{r}_p$ are the rows of $B_{p\times n}$, then

$$AB = \sum_{i=1}^{p} c_i r_i \quad \text{(sum of } p \text{ matrices).}$$

Matrix-Matrix Product

1. The ij-entry of AB is the inner product of the ith row of A and the jth column of B.
2. Rows of AB : $AB = \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ 2a+3c & 2b+3d \end{pmatrix}$.
   The first row of AB is $1[a\ b] + 0[c\ d] = [a\ b]$. The second row of AB is $2[a\ b] + 3[c\ d] = [2a+3c,\ 2b+3d]$. Each row of AB is a combination of the rows of B.
3. Columns of AB : $\begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ 2a+3c & 2b+3d \end{pmatrix}$.
   By columns, B consists of two columns side by side, and A multiplies each of them separately. Therefore each column of AB is the combination of the columns of A.
   $\begin{pmatrix} a \\ 2a+3c \end{pmatrix} = a \begin{pmatrix} 1 \\ 2 \end{pmatrix} + c \begin{pmatrix} 0 \\ 3 \end{pmatrix}$.
   $\begin{pmatrix} b \\ 2b+3d \end{pmatrix} = b \begin{pmatrix} 1 \\ 2 \end{pmatrix} + d \begin{pmatrix} 0 \\ 3 \end{pmatrix}$.
   The first columns of AB is "a" times column 1 plus "c" times column 2.
4. The matrix AB is a sum of two matrices :

$$
\begin{aligned}
AB &= \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \begin{pmatrix} a & b \end{pmatrix} + \begin{pmatrix} 0 \\ 3 \end{pmatrix} \begin{pmatrix} c & d \end{pmatrix} \\
&= \begin{pmatrix} a & b \\ 2a+3c & 2b+3d \end{pmatrix}.
\end{aligned}
$$

Exercises

Exercise 1.

Verify the following :

1. *Matrix multiplication is associative:* $(AB)C = A(BC)$.
2. *Matrix operations are distributive:* $A(B + C) = AB + AC$ *and* $(B + C)D = BD + CD$.
3. *Matrix multiplication is not commutative: Usually* $FE \neq EF$.

Identity Matrix is the matrix which is $n \times n$ square matrix where the diagonal elements are ones and the other elements are all zeros. It is represented as $\mathbb{I}_n$ or just by $\mathbb{I}$, where n represents the size of the square matrix.

That is, identity matrix is a square matrix in which all the elements of the principal diagonal are ones and all other elements are zeros. The effect of multiplying a given matrix by an identity matrix is to leave the given matrix unchanged.

1. **Elementary matrix** E_{ij} is a square matrix and is obtained by replacing ij^{th}-element of the identity matrix $\mathbb{I}$ by $-\ell_{ij}$.
2. Pre-multiplying E_{ij} with A does the subtraction of $\ell_{ij}(Row_j)$ from Row_i according to Gaussian Elimination.
3. For example, when we pre-multiply $E_{31} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\ell_{31} & 0 & 1 \end{pmatrix}$ with $A_{3\times 3}$,

   the first row of A gets multiplied by ℓ_{31} and then it gets substracted from the third row. That is,

$$\text{new}(Row_3) \to -\ell_{31}(Row_1) + (Row_3).$$

Permutation Matrices

1. **Permutation matrices** are derived from the identity matrix $\mathbb{I}$. If we exchange the ith and jth rows of $\mathbb{I}$, we get a permutation matrix, denoted by P_{ij}.
2. Pre-multiplying P_{ij} with A exchanges the ith and jth rows of A.
3. For example, when we pre-multiply the permutation matrix $P_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ with $A_{3\times 3}$, 2nd and 3rd rows of A are interchanged.
4. Product of two permutation matrices is again a permutation matrix.

Exercises

Exercises 2.

1. *What about inverse of a permutation matrix P?*
2. *Is AE_{ij} possible? When possible, what is happening in A?*
3. *What should be done if you want exchange two columns?*
4. *Is AP_{ij} possible? When possible, what is happening in A?*

Triangular Factors and Row Exchanges

Given a system $Ax = \begin{pmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 5 \\ -2 \\ 9 \end{pmatrix} = b.$

First step : subtract 2 times the first equation from the **second**. The **elementary matirx** $E = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ should be pre-multiplied in $Ax = b$, we get

$$EAx = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 5 \\ -2 \\ 9 \end{pmatrix}.$$

Our original matrix subtracts 2 times the first component from the second, leaving the first and third components unchanged. After this step the new and simple system (equivalent to the old) is just $E(Ax) = Eb$.

Triangular Factors and Row Exchanges

In the above example, there are three elimination steps :

(a) subtract 2 times the first equation from the **second**

(b) subtract -1 times the first equation from the **third**

(c) subtract -1 times the second equation from the **third**.

The result is an equivalent but simpler system, with a new coefficient matrix U (**upper triangular matrix**) :

$$Ux = \begin{pmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 5 \\ -12 \\ 2 \end{pmatrix} = c.$$

Triangular Factors and Row Exchanges

The elementary matrices for steps (i), (ii) and (iii) repectively are

$$E = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad F = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \quad G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

The result of all three steps $GFEA = U$, where $GFE = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 1 & 1 \end{pmatrix}$ is a **lower triangular matrix**. We could multiply GFE together to find the single matrix that takes A to U (and also takes b to c). The product GFE is the true order of elimination. It is the matrix that takes the original A to the upper triangular U.

This is good, but the most important question is exactly the opposite. How would we get from U back to A? How can we undo the steps of Gaussian elimination.

Triangular Factors and Row Exchanges

A single step, say step (a), is not hard to undo. Instead of subtracting, we add twice the first row to the second. (Not twice the second row to the first!) The result of doing both the subtraction and the addition is to bring back the identity matrix.

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

If the **elementary matrix** E has the number $-\ell$ in the (i,j) position, then its inverse has $+\ell$ in that position.

That matrix is denoted by E^{-1}. Thus E^{-1} times E is the identity matrix.

Triangular Factors and Row Exchanges

The final problem is to undo the whose process at once, and the matrix $E^{-1}F^{-1}G^{-1}$ takes U back to A. Inverses come in the opposite order.

Thus $LU = A$, where $L = E^{-1}F^{-1}G^{-1}$.

Now we recognize the matrix L that takes U back to A. It is called L, because it is lower triangular. And it has a special property that can be seen only multiplying the three inverse matrices in the right order :

$$E^{-1}F^{-1}G^{-1} = \begin{pmatrix} 1 & & \\ \mathbf{2} & 1 & \\ & & 1 \end{pmatrix} \begin{pmatrix} 1 & & \\ & 1 & \\ \mathbf{-1} & & 1 \end{pmatrix} \begin{pmatrix} 1 & & \\ & 1 & \\ & \mathbf{-1} & 1 \end{pmatrix} = \begin{pmatrix} 1 & & \\ \mathbf{2} & 1 & \\ \mathbf{-1} & \mathbf{-1} & 1 \end{pmatrix} = L.$$

The special thing is that the entries below the diagonal are the multipliers $\ell = 2, -1$ and -1.

Triangular Factors and Row Exchanges

If no row exchanges are required, the orignial matrix A can be written as a product $A = LU$.

The matrix L is lower triangular, with 1's on the diagonal and the multipliers ℓ_{ij} (taken from elimination) below the diagonal.

U is the upper triangular matrix which appears after forward elimination and before back-substitution; its diagonal entries are **the pivots**.

Without Row Exchanges - An Example

Let us consider the linear system $\begin{cases} 2u + v + w = 5 \\ 4u - 6v = -2 \\ -2u + 7v + 2w = 9 \end{cases}$

- Step 1 elimination :
  $Row_2 \to Row_2 - 2Row_1, \quad Row_3 \to Row_3 - (-1)Row_1$
- Elementary matrices : $E_{21} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \quad E_{31} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$
- Step 2 elimination : $Row_3 \to Row_3 - (-1)Row_2$
- Elementary matrices : $E_{32} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$
- The product $E_{32}E_{31}E_{21}A$ becomes an upper diagonal matrix, say U.

Example - Gaussian Elimination

- $$\underbrace{\begin{pmatrix}1&0&0\\0&1&0\\0&1&1\end{pmatrix}}_{E_{32}}\underbrace{\begin{pmatrix}1&0&0\\0&1&0\\1&0&1\end{pmatrix}}_{E_{31}}\underbrace{\begin{pmatrix}1&0&0\\-2&1&0\\0&0&1\end{pmatrix}}_{E_{21}}\underbrace{\begin{pmatrix}2&1&1\\4&-6&0\\-2&7&2\end{pmatrix}}_{A}=\underbrace{\begin{pmatrix}2&1&1\\0&-8&-2\\0&0&1\end{pmatrix}}_{U}$$

- $$\underbrace{\begin{pmatrix}1&0&0\\2&1&0\\0&0&1\end{pmatrix}}_{F_{21}}\underbrace{\begin{pmatrix}1&0&0\\0&1&0\\-1&0&1\end{pmatrix}}_{F_{31}}\underbrace{\begin{pmatrix}1&0&0\\0&1&0\\0&-1&1\end{pmatrix}}_{F_{32}}\underbrace{\begin{pmatrix}2&1&1\\0&-8&-2\\0&0&1\end{pmatrix}}_{U}=\underbrace{\begin{pmatrix}2&1&1\\4&-6&0\\-2&7&2\end{pmatrix}}_{A}$$

We denote the inverse of E_{ij} by F_{ij}. That is, $F_{21}F_{31}F_{32} = (E_{32}E_{31}E_{21})^{-1}$ and $F_{21} = E_{21}^{-1}, F_{31} = E_{31}^{-1}, F_{32} = E_{32}^{-1}$. Thus

$$\underbrace{\begin{pmatrix}2&1&1\\4&-6&0\\-2&7&2\end{pmatrix}}_{A}=\underbrace{\begin{pmatrix}1&0&0\\2&1&0\\-1&-1&1\end{pmatrix}}_{L}\underbrace{\begin{pmatrix}2&1&1\\0&-8&-2\\0&0&1\end{pmatrix}}_{U}.$$

Triangular Factorization $A = LU$

One linear system = two triangular system. When L and U are known, A could be thrown away. We go from b to c by forward elimination (that uses L) and we go from c to x by back-substitution (that uses U). We can and should do without A, when its factors have been found. ($A = Lu$, $b = Ax = LUx$ implies $L^{-1}Ax = L^{-1}b = c$).

In matrix terms, elimination splits $Ax = b$ into two triangular systems: first $Lc = b$ and then $Ux = c$. This identical to $Ax = b$. Pre-multiply $Ux = c$ by L to give $LUx = Lc$, which is $Ax = b$.

Each triangular system can be solved in $n^2/2$ steps. The solution for any new right side b' can be found in only n^2 operations. That is far below the $n^3/3$ steps needed to factor A on the left hand side.

Triangular Factorization $A = LU$

The triangular factorization is often written $A = LDU$, where L and U have 1's on the diagonal and D is the diagonal matrix of pivots.

It is conventional, although completely confusing, to go on denoting this new upper triangular matrix by the same letter U. Whenever you see LDU, it is understood that U has 1's on the diagonal - in other words that each row was divided by the pivot. Then L and U are treated evenly. An example for $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ is

$$A = \begin{pmatrix} 1 & \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ & -2 \end{pmatrix} = \begin{pmatrix} 1 & \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & \\ & -2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ & 1 \end{pmatrix} = LDU.$$

That has the 1's on the diagonals of L and U, and the pivots 1 and -2 in D.

Unique Triangular Factorization

If $A = L_1 D_1 U_1$ and $A = L_2 D_2 U_2$, where the L's are lower triangular with unit diagonal, the U's are upper triangular with unit diagonal, and the D's are diagonal matrices with no zeros on the diagonal, then

$$L_1 = L_2, \; D_1 = D_2, \; U_1 = U_2.$$

The LDU factorization and the LU factorization are uniquely determined by A.

Example - Gaussian Elimination

- How to solve for x? We have not applied row operations on RHS b.

$$\boxed{Ax = b \implies LUx = b.}$$

To solve $Ax = b$, we solve two triangular systems in the order

$$\boxed{Ly = b \; ; \quad Ux = y.}$$

- If A remains same, but b changes (in any mathematical model), Gaussian Elimination provides both L and U. Only solution needs to be found for every changing vector b.
- We can also write as $A = LDU$, where D is the diagonal matrix with pivots on the diagonals, L and U are lower and upper triangular matrices with unit diagonal entries.

Row Exchanges and Permutation Matrices

We have to face a problem that the number we expect to use as a pivot might be zero. This could occur in the middle of a calculation, or it can happen at the very beginning (in case $a_{11} = 0$.) A simple example is $\begin{pmatrix} 0 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$. The difficulty is clear, no multiple of the first equation will remove the coefficient 3.

The remedy is equally clear.

Exchange the two equations, moving the entry 3 up into the pivot.

To express this in matrix terms, we need to find **the permutation matrix that produces the row exchange**. It is $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and multiply by P does exchange the rows.

Zero in the Pivot Location

The next difficult case is that **a zero in the pivot location** raises two possibilites: the trouble may be easy to fix, or it may be serious.

This is decided by looking below the zero. If there is a nonzero entry lower down in the same column, then a row exchange is carried out; the nonzero entry becomes the needed pivot, and estimation can get going again.

In the example, $A = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ d & e & f \end{pmatrix}$, everything depends on the number d.

If $d = 0$, the problem is incurable and matrix is **singular**. There is no hope for a unique solution.

Zero in the Pivot Location

If d is not zero, an exchange of rows 1 and 3, permutation matrix

$$P_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

will move d into the pivot, and stage 1 is complete. However the next pivot position also contains a zero.

The number a is now below it (the e above is useless) and if a is not zero, then another row exchange is called for

$$P_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

(exchange of rows 2 and 3).

Zero in the Pivot Location

There is a permutation matrix that will do both of the row exchanges at once, which is the product of the two separate permutaions

$$P_{23}P_{13} = P = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

(first exchange rows 1 and 3, then exchange rows 2 and 3).

If we had known what to do, we could have multiplied our matrix by P in the first place.

Then elimination order would have no difficulty with PA; it was only that the original order was unfortunate.

Example - Gaussian Elimination

- Triangular factorization: $\boxed{A = LU}$ $\boxed{L = F_{21}F_{31}F_{32}}$
- Diagonals of L are ones and diagonals of U are the pivots.
- In this example : $L = F_{21}F_{31}F_{32} = \begin{pmatrix} 1 & 0 & 0 \\ \ell_{21} & 1 & 0 \\ \ell_{31} & \ell_{32} & 1 \end{pmatrix}$
- All multiples used in elimination are the elements of L.

.

Summary

The theory of Gaussian elimination can be summarized as follows:
In the **nonsingular** case, there is a permutation matrix P that reorders the rows of A to avoid zeros in the pivot positions. In this case

(a) $Ax = b$ has a unique solution
(b) it is found by elimination with row exchanges
(c) with the rows reordered in advance, PA can be factored into LU.

In **singular** case, no reordering can produce a full set of pivots.

Caution about the lower triangular matrix L

Suppose elimination subtracts row 1 from row 2, creating $\ell_{21} = 1$. Then suppose it exchanges rows 2 and 3. If that exchange is done in advance, the multiplier will change to $\ell_{31} = 1$ in $PA = LU$.

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 3 \\ 2 & 5 & 8 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 3 & 6 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 \\ 0 & 3 & 6 \\ 0 & 0 & 2 \end{pmatrix} = U.$$

With the rows exchanged, we recover LU - but now $\ell_{31} = 1$ and $\ell_{21} = 2$.

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \text{ and } PA = LU.$$

With Row Exchanges

$$A = \begin{pmatrix} 3 & 4 & 7 \\ 6 & 8 & 3 \\ 1 & 2 & 1 \end{pmatrix} \to \begin{pmatrix} 3 & 4 & 7 \\ 0 & 0 & -11 \\ 0 & \frac{2}{3} & -\frac{4}{3} \end{pmatrix} \to \begin{pmatrix} 3 & 4 & 7 \\ 0 & \frac{2}{3} & -\frac{4}{3} \\ 0 & 0 & -11 \end{pmatrix}$$

- Step 1 : Multipliers are $\ell_{21} = 2$ and $\ell_{31} = \frac{1}{3}$.
- Step 2 : Row exchange needed. Pre-multiply by the permutation matrix

$$P_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

- We get $P_{23}A = LU$, where $L = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$.

What if more row exchanges are done on A at different stages of eliminatation?

- Consider $A = \begin{pmatrix} 0 & 3 & 2 \\ 0 & 0 & 4 \\ 2 & 4 & 9 \end{pmatrix}$
- $Row_1 \leftrightarrow Row_3$: $P_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$
- Elimination $\ell_{21} = \ell_{31} = 0$ (No elimination at step 1 after row exchange)
- $Row_2 \leftrightarrow Row_3$: $P_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$
- Elimination : $\ell_{32} = 0$ (No elimination at step 2 after row exchange)
- $U = \begin{pmatrix} 2 & 4 & 9 \\ 0 & 3 & 2 \\ 0 & 0 & 4 \end{pmatrix}$.
- We get $P_{23}P_{13}A = LU$ ($L = \mathbb{I}$ in this example).

Example - Gaussian Elimination

- For a nonsingular matrix A, there is a permutation matrix P that reorders the rows of A to avoid zeros in the pivot positions. Then $Ax = b$ has a unique solution. With the rows reordered in advance, PA can be factored into LU.
- In practice, we cannot reorder in advance. Still it is possible to obtain correct P, L, U matrices so that $PA = LU$.

Exercises 3.

1. *What multiple ℓ_{32} of row 2 of A will elimination subtract from row 3 of A? Use the factored form*

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 4 & 1 \end{bmatrix} \begin{bmatrix} 5 & 7 & 8 \\ 0 & 2 & 3 \\ 0 & 0 & 6 \end{bmatrix}.$$

   What will be the pivots? Will a row exchange be required?

2. (a) *Why does it take approximately $n^2/2$ multiplication-subtraction steps to solve each of $Lc = b$ and $Ux = c$?*
   (b) *How many steps does elimination use in solving 10 systems with the same 60 by 60 coefficient matrix A?*

3. *Apply elimination to produce the factors L and U for*

$$A = \begin{bmatrix} 2 & 1 \\ 8 & 7 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 4 & 4 \\ 1 & 4 & 8 \end{bmatrix}.$$

4. *Find a 4 by 4 permutation matrix that requires three row exchanges to reach the end of elimination (which is $U = I$).*

5. *How could you factor A into a product UL, upper triangular times lower triangular? Would they be the same factors as in $A = LU$?*

Exercises 4.

1. *Solve by elimination, exchanging rows when necessary:*

$$\begin{aligned} u + 4v + 2w &= -2 \\ -2u - 8v + 3w &= 32 \quad \text{and} \\ v + w &= 1 \end{aligned} \qquad\qquad \begin{aligned} v + w &= 0 \\ u + v &= 0 \\ u + v + w &= 1. \end{aligned}$$

   Which permutation matrices are required?

2. *(Move to 3 by 3) Forward elimination changes $Ax = b$ to a triangular $Ux = c$:*

$$\begin{aligned} x + y + z &= 5 \\ x + 2y + 3z &= 7 \\ x + 3y + 6z &= 11 \end{aligned} \qquad \begin{aligned} x + y + z &= 5 \\ y + 2z &= 2 \\ 2y + 5z &= 6 \end{aligned} \qquad \begin{aligned} x + y + z &= 5 \\ y + 2z &= 2 \\ z &= 2. \end{aligned}$$

   The equation $z = 2$ in $Ux = c$ comes from the original $x + 3y + 6z = 11$ in $Ax = b$ by subtracting $\ell_{31} =$ ________ times equation 1 and $\ell_{32} =$ ________ times the final equation 2. Reverse that to recover $\begin{bmatrix} 1 & 3 & 6 & 11 \end{bmatrix}$ in $\begin{bmatrix} A & b \end{bmatrix}$ from the final $\begin{bmatrix} 1 & 1 & 1 & 5 \end{bmatrix}$ and $\begin{bmatrix} 0 & 1 & 2 & 2 \end{bmatrix}$ and $\begin{bmatrix} 0 & 0 & 1 & 2 \end{bmatrix}$ in $\begin{bmatrix} U & c \end{bmatrix}$:

$$\text{Row 3 of } \begin{bmatrix} A & b \end{bmatrix} = (\ell_{31}\ \text{Row}\,1 + \ell_{32}\ \text{Row}\,2 + 1\ \text{Row}\,3) \text{ of } \begin{bmatrix} U & c \end{bmatrix}.$$

   In matrix notation this is multiplication by L. So $A = LU$ and $b = Lc$.

Exercises 5.

1. *What two elimination matrices* E_{21} *and* E_{32} *put* A *into upper triangular form* $E_{32}E_{21}A = U$*? Multiply by* E_{32}^{-1} *and* E_{21}^{-1} *to factor* A *into* $LU = E_{21}^{-1}E_{32}^{-1}U$*:*
$$A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 4 & 5 \\ 0 & 4 & 0 \end{bmatrix}.$$

2. *Solve* $Lc = b$ *to find* c*. Then solve* $Ux = c$ *to find* x*. What was* A*?*
$$L = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad U = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}.$$

3. *(Review) For which numbers* c *is* $A = LU$ *impossible—with three pivots?*
$$A = \begin{bmatrix} 1 & 2 & 0 \\ 3 & c & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

4. *Estimate the time difference for each new right-hand side* b *when* $n = 800$*. Create* $A = rand(800)$ *and* $b = rand(800, 1)$ *and* $B = rand(800, 9)$*. Compare the times from tic;* $A\backslash b$*; toc and tic;* $A\backslash B$*; toc (which solves for 9 right sides).*

Exercises

Exercises 6.

1. *If P_1 and P_2 are permutation matrices, so is P_1P_2. This still has the rows of I in some order. Give examples with $P_1P_2 \neq P_2P_1$ and $P_3P_4 = P_4P_3$.*
2. *Find a 3 by 3 permutation matrix with $P^3 = I$ (but not $P = I$). Find a 4 by 4 permutation $\hat{P}$ with $\hat{P}^4 \neq I$.*
3. *There are 12 "even" permutations of $(1, 2, 3, 4)$, with an even number of exchanges. Two of them are $(1, 2, 3, 4)$ with no exchanges and $(4, 3, 2, 1)$ with two exchanges. List the other ten. Instead of writing each 4 by 4 matrix, use the numbers $4, 3, 2, 1$ to give the position of the 1 in each row.*
4. *The matrix P that multiplies (x, y, z) to give (z, x, y) is also a rotation matrix. Find P and P^3. The rotation axis $a = (1, 1, 1)$ doesn't move, it equals Pa. What is the angle of rotation from $v = (2, 3, -5)$ to $Pv = (-5, 2, 3)$?*
5. *If P has 1s on the antidiagonal from $(1, n)$ to $(n, 1)$, describe PAP.*

Inverse

A **left inverse** of a matrix A is any matrix B such that $BA = \mathbb{I}$. A **right inverse** of A is any matrix C such that $AC = \mathbb{I}$.

A matrix B is said to be an **inverse** of A if it is both a left inverse and a right inverse of A. The matrix is **invertible** if there exists a matrix B such that $BA = \mathbb{I}$ and $AB = \mathbb{I}$. There is at most one such B, called **the inverse of** A and denoted by A^{-1}: $A^{-1}A = \mathbb{I}$ and $AA^{-1} = \mathbb{I}$.

- If there exists $x \neq 0$ such that $Ax = 0$, then A^{-1} does not exist.
- If inverse of a matrix exists, then it is unique.
- If A^{-1} exists, then $x = A^{-1}b$ is the unique solution of $Ax = b$.
- A product AB of invertible matrices has an inverse. It is found by multiplying the individual inverses in reverse order:

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Exercises 7.

1. *What is the inverse of a* 2×2 *matrix?*
2. *What is the inverse of*
   (a) *Identity matrix?*
   (b) *Diagonal matrix?*
   (c) *Lower and upper triangular matrices?*
   (d) *Elementary and permutation matrices?*
3. *If* $BA = \mathbb{I}$ *and* $AC = \mathbb{I}$*, then show that* $B = C$.

Gauss-Jordan Elimination is an algorithm that can be used to solve systems of linear equations and to find the inverse of any invertible matrix.

Gauss-Jordan Method

Consider the equation $AA^{-1} = \mathbb{I}$. If it is taken a column at a time, that equation determines the column of A^{-1}. The first column of the identity matrix $\mathbb{I}$ is the product of A and the first column of A^{-1}.

Consider a square matrix of order 3.

Let x_1, x_2, x_3 be the columns of A^{-1}. Then $Ax_1 = e_1, Ax_2 = e_2, Ax_3 = e_3$. Thus we have three systems of equations (or, in general n systems) and they all have the same coefficient matrix A.

The right sides are different, but it is possible to carry out elimination on **all systems simultaeously**. This is called the **Gauss-Jordan method**.

Gauss-Jordan Method

Instead of stopping at U and switching to back-substitution, it continues by subtracting multiples of a row, from the rows above. It produces zeros above the diagonal as well as below, and when it reaches the identity matrix we have found A^{-1}.

$$[A \;\; \mathbb{I}] \to [U \;\; L^{-1}] \to [\mathbb{I} \;\; A^{-1}]$$

The example keeps all three columns e_1, e_2, e_3, and operates on rows of length six : $[A \; e_1 \;\; e_2 \;\; e_3]$ becomes

$$\begin{pmatrix} 2 & 1 & 1 & 1 & 0 & 0 \\ 4 & -6 & 0 & 0 & 1 & 0 \\ -2 & 7 & 2 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 1 & 1 & 0 & 0 \\ 0 & -8 & -2 & -2 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{pmatrix} = [U \; L^{-1}].$$

Gauss-Jordan Method

The first half of elimination has gone from A to U, and now the second half will go from U to $\mathbb{I}$.

Creating zeros above the pivots in the matrix, we reach A^{-1}: The matrix $[U\ L^{-1}]$ becomes

$$\begin{pmatrix} 2 & 1 & 0 & 2 & -1 & -1 \\ 0 & -8 & 0 & -4 & 3 & 2 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \frac{12}{16} & \frac{-5}{16} & \frac{-6}{16} \\ 0 & 1 & 0 & \frac{4}{8} & \frac{-3}{8} & \frac{-2}{8} \\ 0 & 0 & 1 & -1 & 1 & 1 \end{pmatrix} = [\mathbb{I}\ A^{-1}].$$

So $A^{-1} = \begin{pmatrix} \frac{12}{16} & \frac{-5}{16} & \frac{-6}{16} \\ \frac{4}{8} & \frac{-3}{8} & \frac{-2}{8} \\ -1 & 1 & 1 \end{pmatrix}$.

Gauss-Jordan Method

At the last step, we divided through by the pivots. The coefficient matrix in the left half became the identity. Since A went to $\mathbb{I}$, the same operations on the right half must have carried $\mathbb{I}$ into A^{-1}. Therefore we have computed the inverse. The number of operations required to find A^{-1} is given below :

The normal count for each new right-hand side is n^2, half in the forward direction and half in back-substitution. With n-right-hand sides $e_1, \ldots, e_n$ this makes n^3. After including the $n^3/3$ operations on A itself, the total seems to be $4n^3/3$.

This result is a little too high because of the zeros in the e_j. Forward elimination changes only the zeros below the 1. This part has only $n - j$ components, so the count for e_j is effectively changes to $(n-j)^2/2$. Summing over all j, the total for forward elimination is $n^3/6$. This is to be combined with the usual $n^3/3$ operations that are applies to A, and the $n(n^2/2)$ back-substitution steps that finally produce the columns x_j of A^{-1}.

The final operation count for computing A^{-1} is n^3 :

$$\textbf{Operation count} \quad \frac{n^3}{6} + \frac{n^3}{3} + n\left(\frac{n^2}{2}\right) = n^3.$$

Symmetric Matrices

The transpose of a lower triangular matrix is upper triangular. The transpose of A^T brings us back to A. If we add two matrices A and B and then transpose the result is the same as first transposing and then adding: $(A+B)^T = A^T + B^T$. Also, $(AB)^T = B^T A^T$ and $(A^{-1})^T = (A^T)^{-1}$.

A special class of matrices, probably the most important class of all.

A **symmetric matrix** is a matrix which equals its own transpose. That is, $A^T = A$.

The matric is necessarily square, and each entry on one side of the diagonal equals its "mirror image" on the other side $a_{ij} = a_{ji}$.

If A is symmetric, then A^{-1} is symmetric (if A^{-1} exits). Symmetric matrices appear in every subject whose laws are fair. "Each action has an equal and opposite reaction", and the entry which gives the action of i onto j is matched by the action of j onto i.

Symmetric Matrices

When elimination is applied to a symmetric matrix, $A^T = A$ is an advantage. The smaller matrices stay symmetric as elimination proceeds, and we can work with half the matrix. The lower right-hand corner remains symmetric :

$$\begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix} \to \begin{bmatrix} a & b & c \\ 0 & d - \frac{b^2}{a} & e - \frac{bc}{a} \\ 0 & e - \frac{bc}{a} & f - \frac{c^2}{a} \end{bmatrix}.$$

The work of elimination is reduced from $n^3/3$ to $n^3/6$. There is no need to store entires from both sides of the diagonal, or to store both L and U.

LDU **Factorization for Symmetric Matrices.** If A is symmetric, and if it can be factored into $A = LDU$ without row exchanges to destroy the symmetry, then the upper triangular U is the transpose of the lower triangular L. The factorization becomes $A = LDL^T$.

Relation Between Pivots and Matrix Inverse

Theorem 8.

Let A be an $n \times n$ matrix. Then A^{-1} exists if and only if A has n pivots.

Proof :

- If A has n pivots, then Gauss-Jordan elimination will be done successfully and we obtain $A \to \mathbb{I}$ by a sequence $E_1, E_2, \ldots, E_k$ of elementary martices, permutations or matrices which divide the row by pivot. Let $E_k E_{k-1} \cdots E_2 E_1 A = \mathbb{I}$ and $B = E_k E_{k-1} \cdots E_2 E_1$. Then $BA = \mathbb{I}$.
- Since A has n pivots, then system $A \begin{pmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{n1} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ has a solution. Similarly,

  $A \begin{pmatrix} c_{1j} \\ c_{2j} \\ \vdots \\ c_{nj} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$ has a solution. Let $C = (cij)$. Then $AC = \mathbb{I}$. Hence $B = C$. Thus A is invertible.

Proof (contd...)

- To prove the converse part, we assume that $B = A^{-1}$ exists. Then $AB = \mathbb{I}$.
- Suppose A has $< n$ pivots. This means elimination on A produces at least a zero row.
- Let M be the product of the sequence of steps (elimination) that led to the zero row. Then MA has a zero row.
- Also $AB = \mathbb{I}$. This implies $(MA)B = M$.
- Thus M has a zero row, but that is not possible. Thus A must has n pivots.

Transpose

The transpose of a matrix is found by interchanging its rows into columns or columns into rows. The transpose of the matrix is denoted by using the letter "T" in the superscript of the given matrix. The transpose of a matrix A is denoted by A^T.

- If A is lower triangular, then A^T is upper triangular.
- $(A+B)^T = A^T + B^T$.
- $(AB)^T = B^T A^T$.
- $(A^{-1})^T = (A^T)^{-1}$.

A **symmetric matrix** is a matrix that is equal to its transpose. That is, A is symmetric if $A = A^T$.

Exercises 9.

1. *If L is a lower triangular matrix and invertible, then prove that L^{-1} is also a lower triangular matrix.*
2. *Let P be a permutation matrix. Then show that $P^{-1} = P^T$.*
3. *Let L be a lower triangular matrix. Then L is invertible iff the diagonal elments of L are nonzero.*
4. *Suppose elimination fails because there is no pivot in column 3:*

$$\text{Missing pivot} \qquad A = \begin{bmatrix} 2 & 1 & 4 & 6 \\ 0 & 3 & 8 & 5 \\ 0 & 0 & 0 & 7 \\ 0 & 0 & 0 & 9 \end{bmatrix}.$$

Show that A cannot be invertible. The third row of A^{-1}, multiplying A, should give the third row $\begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}$ of $A^{-1}A = I$. Why is this impossible?

Exercises 10.

1. *If the inverse of A^2 is B, show that the inverse of A is AB. (Thus A is invertible whenever A^2 is invertible.)*
2. (a) *Find the inverses of the permutation matrices*

$$P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad P = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

   (b) *Explain for permutations why P^{-1} is always the same as P^T. Show that the 1s are in the right places to give $PP^T = I$.*
3. *Use the Gauss-Jordan method to invert*

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Exercises 11.

1. *Compute the symmetric LDL^T factorization of*

$$A = \begin{bmatrix} 1 & 3 & 5 \\ 3 & 12 & 18 \\ 5 & 18 & 30 \end{bmatrix} \quad \textit{and} \quad A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}.$$

2. *Find the inverse of*

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

3. *If* $A = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ *and* $B = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$, *compute* A^TB, B^TA, AB^T, *and* BA^T.
4. *Suppose A is invertible and you exchange its first two rows to reach B. Is the new matrix B invertible? How would you find B^{-1} from A^{-1}?*
5. *(Remarkable) If A and B are square matrices, show that $I - BA$ is invertible if $I - AB$ is invertible. Start from $B(I - AB) = (I - BA)B$.*
6. *Prove that a matrix with a column of zeros cannot have an inverse.*

Exercises 12.

1. *Prove that A is invertible if $a \neq 0$ and $a \neq b$ (find the pivots and A^{-1}):*

$$A = \begin{bmatrix} a & b & b \\ a & a & b \\ a & a & a \end{bmatrix}.$$

2. *True or false (with a counterexample if false and a reason if true):*
   (a) *A 4 by 4 matrix with a row of zeros is not invertible.*
   (b) *A matrix with 1s down the main diagonal is invertible.*
   (c) *If A is invertible then A^{-1} is invertible.*
   (d) *If A^T is invertible then A is invertible.*
3. *Verify that $(AB)^T$ equals $B^T A^T$ but those are different from $A^T B^T$:*

$$A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \qquad AB = \begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix}.$$

In case $AB = BA$ (not generally true!), how do you prove that $B^T A^T = A^T B^T$?

Exercises 13.

1. *Factor these symmetric matrices into $A = LDL^T$. The matrix D is diagonal:*

$$A = \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 1 & b \\ b & c \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}.$$

2. *If $A = A^T$ needs a row exchange, then it also needs a column exchange to stay symmetric. In matrix language, PA loses the symmetry of A but ______ recovers the symmetry.*
3. *Here is a new factorization of A into triangular times symmetric:*

$$\textit{Start from } A = LDU. \textit{ Then } A \textit{ equals } L(U^T)^{-1} \textit{ times } U^T DU.$$

   Why is $L(U^T)^{-1}$ triangular? Its diagonal is all 1s. Why is $U^T DU$ symmetric?
4. *A square* northwest matrix *B is zero in the southeast corner, below the antidiagonal that connects $(1, n)$ to $(n, 1)$. Will B^T and B^2 be northwest matrices? Will B^{-1} be northwest or southeast? What is the shape of $BC =$* northwest times southeast*? You are allowed to combine permutations with the usual L and U (southwest and northeast).*

Special Matrices and Applications

We now discuss two points. The first is to explain one way in which large systems of linear equations can arise in practice. The other goal is to illustrate, by the same application, the special properties that coefficient matrices frequently have. It is unusual to meet large matrices that look as if there were constructed at random. Almost always there is a pattern, visible even at first sight - frequently a pattern of symmetry, and of very many zero entries. In the latter case, since a sparse matrix contains far fewer than n^2 pieces of information, the computations ought to work out much more simply than for a full matrix.

We shall look particularly at band matrices, whose nonzero entries are concentrated near the main diagonal, to see how this property is reflected in the elimination process. In fact we look at one special band matrix.

Special Matrices and Applications

Our example comes from changing a continuous problem into a discrete one. The continuous problem will have infinitely many unknowns (its asks for $u(x)$ at every x), and it cannot be solved exactly on a computer. Therefore it has to be approximated by a discrete problem - the more unknowns we keep, the better will be the accuracy and the greater the expense.

As a simple but still very typical continuous problem, our choice falls on the differential equation $-\frac{d^2u}{dx^2} = f(x), \quad 0 \leq x \leq 1$. This is a linear equation for the unknown function u, with inhomogeneous term f. There is some arbitrariness left in the problem, because any combination $C = Dx$ could be added to any solution.

The sum would constitute another solution, since the second derivative of $C + Dx$ contributes nothing. Therefore the uncertainty left by these two arbitrary constants C and D will be removed by adding a "boundary condition" at each end of the interval: $u(0) = 0, \quad u(1) = 0$.

Special Matrices and Applications

The result is a two-point boundary-value problem, describing not a transient but a steady-state phenomenon - the temperature distribution in a rod, for example, with ends fixed at 0° and with a heat source $f(x)$.

Remember that our goal is to produce that is discrete, or finite-dimensional - in other words, a problem in linear algebra.

For that reason we cannot accept more a finite amount of information about f, say its values at the equally spaced points $x = h, x = 2h, \ldots, x = nh$. And what we compute will be approximate values $u_1, \ldots, u_n$ for the true solution u at these same points. At the ends $x = 0$ and $x = 1 = (n+1)h$, we are already given the correct boundary values $u_0 = 0, u_{n+1} = 0$.

Special Matrices and Applications

The first question is, How de we replace the derivative d^2u/dx^2?

For du/dx there are several alternatives

$$\frac{du}{dx} \approx \frac{u(x+h)-u(x)}{h} \text{ or } \frac{u(x)-u(x-h)}{h} \text{ or } \frac{u(x+h)-u(x-h)}{2h}.$$

The last, because it is symmetric about x, is actually the most accurate. For the second derivative there is just one combination that uses only the values at x and $x \pm h$:

$$\frac{d^2u}{dx^2} \approx \frac{u(x+h)-2u(x)+u(x+h)}{h^2}.$$

It also has the merit of being symmetric about x.

At a typical meshpoint $x = jh$, the differential equations $-d^2u/dx^2 = f(x)$ is now replaced by this discrete analogue

$$\frac{d^2u}{dx^2} \approx \frac{u(x+h) - 2u(x) + u(x+h)}{h^2},$$

after multiplying through by h^2, $-uj + 1 + 2u_j - u_{j-1} = h^2 f(jh)$. There are n equations of exactly this form, one for every value $j = 1, \ldots, n$.

The first and last equations include the expressions u_0 and u_{n+1}, **where are not unknowns** - their values are the boundary conditions, and they are shifted to the right side of the equation and contribute to the non-homogeneous terms (in our example, they are zero).

It is easy to understand the previous equation as a steady-state equation, in which the flows $(u_i - u_{j+1})$ coming from the right and $(u_j - u_{j-1})$ coming from the left are balanced by the loss of $h^2 f(jh)$ at the center.

The structure of the n equations $-uj + 1 + 2u_j - u_{j-1} = h^2 f(jh)$ can be better visualized in matrix form $Au = b$. We shall choose $h = \frac{1}{6}$, or $n = 5$:

$$\begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \end{bmatrix} = h^2 \begin{bmatrix} f(h) \\ f(2h) \\ f(3h) \\ f(4h) \\ f(5h) \end{bmatrix}.$$

From now on, we will work with the above equation, and it is not essential to look back at the source of the problem.

The matrix A possesses many special properties, and three of those properties are fundamental :

1. **The matrix is trigiagonal.** All its nonzero entries lie on the main diagonal and the two adjacent diagonals. Outside this band there is nothing : $a_{ij} = 0$ if $|i-j| > 1$.
2. **The matrix is symmetric.** Each entry a_{ij} equals its mirror image a_{ji}, so that $A^T = A$. Therefore the upper triangular U will be the transpose of the lower triangular L, and the final factorization will be $A = LDL^T$. This symmetry of A reflects the symmetry of the original differential equation. If there had been an odd derivative like d^3u/dx^3 of du/dx. A would not have been symmetric.
3. **The matrix is positive definite.** It says that the pivots are positive but symmetry with positive pivots does have one immediate consequence: Row exchanges are unnecessary both in theory and in practice. The product of the pivots is the **determinant** of A.

Special Matrices and Applications

We can summarize the final result in several ways. The most revealing is to look at the LDU factorization of A :

$$A = \begin{bmatrix} 1 & & & & \\ \frac{-1}{2} & 1 & & & \\ & \frac{-2}{3} & 1 & & \\ & & \frac{-3}{4} & 1 & \\ & & & \frac{-4}{5} & 1 \end{bmatrix} \begin{bmatrix} \frac{2}{1} & & & & \\ & \frac{3}{2} & & & \\ & & \frac{4}{3} & & \\ & & & \frac{5}{4} & \\ & & & & \frac{6}{5} \end{bmatrix} \begin{bmatrix} 1 & \frac{-1}{2} & & & \\ & 1 & \frac{-2}{3} & & \\ & & 1 & \frac{-3}{4} & \\ & & & 1 & \frac{-4}{5} \\ & & & & 1 \end{bmatrix}.$$

The L and U factors of a tridiagonal matrix are bidiagonal. These factors have more or less the same structure of zeros as A itself. Note that L and U are transposes of one another, as was expected from the symmetry, and that the pivots d_i are all positive. The pivots are obviously converging to a limiting value of $+1$, as n gets large. Such matrices make a computer very happy.

These simplifications lead to a complete change in the usual operation count. At each elimination stage only two operations are needed, and there are n such stages. Therefore in place of $n^3/3$ operations we need only $2n$; the computation is quicker by orders of magnitude. And the same is true of back-substitution; instead of $n^2/2$ operations we again need only $2n$. Thus the number of operations for a tridiagonal system is proportional to n, not to a higher power of n. Tridiagonal systems $Ax = b$ can be solved almost instantaneously.

Suppose, more generally, that A is a band matrix; its entries are all zero except within the band $|i - j| < w$. The "half bandwidth" is $w = 1$ for a **diagonal matrix**, $w = 2$ for a **tridiagonal matrix**, and $w = n$ for a full matrix. The first stage of elimination requires $w(w - 1)$ operations, and after this stage we still have bankwidth w. Since there are about n stages, elimination on a bank matric must require about $w^2 n$ operations.

The operation count is proportional to n, and now we see that it is proportional also to the square of w. As w approaches n, the matrix becomes full, and the count again is roughly n^3. A more exact count depends on the fact that in the lower right corner the bandwidth is no longer w; there is no room for that many bands.

The precise number of divisions and multiplication-subtractions that produce L, D, and U (without assuming a symmetric A) is

$$P = \frac{1}{3}w(w-1)(3n-2w+1).$$

For a full matrix, which has $w = n$, we recover $P = \frac{1}{3}n(n-1)(n+1)$.

To summarize : A band matrix A has triangular factors L and U that lie within the same band, and both elimination and back-substitution are very fast. This is our last operation count, but we must emphasize the main point.

Roundoff Error

For a finite difference matrix like A, the inverse is a full matrix. Therefore, in solving $Ax = b$, we are actually much worse off knowing A^{-1} than knowing L and U. Multiplying A^{-1} by b takes n^2 steps, whereas $4n$ are sufficient to solve $Lc = b$ and then $Ux = c$ - the forward elimination and back-substitution that produce $x = U^{-1}c = U^{-1}L^{-1}b = A^{-1}b$.

In theory, the nonsingular case is completed. Row exchanges may be necessary to achieve a full set of pivots; then back-substitution solves $Ax = b$. In practice, other row exchanges may be equally necessary - or the computed solution can easily become worthless. Remember that for a system of moderate size, say 100 by 100, elimination involves a third of a million operations. With each operation we must expect a roundoff error. Normally, we keep a fixed number of significant digits (say three, for an extremely weak computer). Then adding two numbers of different sizes gives $0.345 + 0.00123 \to 0.346$ and the last digits in the smaller number are completely lost.

Roundoff Error

The question is, how do all these individual roundoff errors contribute to the final error in the solution?

This is not an easy problem. It was attacked by John von Neumann, who was the leading mathematician at the time when computers suddenly made a million operations possible.

In fact, the combination of Gauss and von Neumann gives the simple elimination algorithm a remarkabble distinguished history, although even von Neumann got a very complicated estimate of the roundoff error; it was Wilkinson who found the right way to answer the question, and his books are now classics.

References

1. G. Strang, Linear Algebra and Its Applications, Thomson Asia, 2003.
2. W. Cheney and D. Kincaid, Linear Algebra: Theory and Applications, Jones & Bartlett Student Edition, 2014.
3. S. Lang, Linear Algebra, 3rd Edition, Springer, 2004
4. S. Kumaresan, Linear Algebra: A Geometric Approach, PHI, 2008