Next topics: Solving systems of linear equations

1. Gaussian elimination (today)
2. Gaussian elimination with partial pivoting (Week 9)
3. The method of LU-decomposition (Week 10)
4. Iterative techniques: Jacobi method (Week 11)
5. Revision (Week 12)
6. **Continuous Assessment Test**
   (Week 13 - Monday, 10th December, 7.00-9.00pm)

Solving systems of linear equations

Example: A mathematician has three children called X, Y and Z. Next year, Y will be half the age of X, while X will be three times as old as Z. In 5 years time, X will be twice as old as the sum of the current ages of the other two. How old are the children?

Review of matrices

Let m, $n \in \mathbb{N}$. A rectangle of numbers written like

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

where each $a_{ij} \in \mathbb{R}$ is called a **matrix** with m rows and n columns or an $m \times n$ matrix.

The above matrix may be denoted by $[a_{ij}]_{\substack{1 \le i \le m \\ 1 \le j \le n}}$ (or just by $[a_{ij}]$).

The number a_{ij} is called the (i, j) entry. It lies in the i^{th} row and j^{th} column.

The entries like a_{11}, a_{22}, a_{33} and so on are called the **diagonal entries**. These entries form the **main diagonal**. Notice that the main diagonal finishes in the bottom right corner only if the matrix has the same number of rows as columns, i.e. we have a "square matrix".

A matrix composed of only one column (that is an $m \times 1$ matrix) is called a column matrix or **vector**.

Two matrices are equal if and only if they have the same size and the corresponding entries are all equal.

For each m and n in $\mathbb{N}$, we use $\mathbb{M}_{m,n}$ for the set of all $m \times n$ matrices.

The $m \times n$ matrix whose entries are all zero is called the $m \times n$ **zero matrix** and is denoted $0_{m,n}$ or simply $\mathbf{0}$.

Matrix addition

If $A = [a_{ij}]$ and $B = [b_{ij}]$ are two matrices in $\mathbb{M}_{m,n}$ then we define $A + B \in \mathbb{M}_{m,n}$ by

$$A + B = [a_{ij} + b_{ij}] = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{bmatrix}$$

For example, $\begin{bmatrix} 1 & 2 & 1 \\ 4 & -3 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -3 & 2 \\ 5 & 4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 3 \\ 9 & 1 & 1 \end{bmatrix}$.

Note: We can only add two matrices if they are of the same size.

Scalar multiplication

If $A = [a_{ij}] \in \mathbb{M}_{m,n}$ and $k \in \mathbb{R}$ then we define the product of A with the scalar number k to be $kA \in \mathbb{M}_{m,n}$ given by

$$kA = [ka_{ij}] = \begin{bmatrix} ka_{11} & ka_{12} & \dots & ka_{1n} \\ ka_{21} & ka_{22} & \dots & ka_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ka_{m1} & ka_{m2} & \dots & ka_{mn} \end{bmatrix}$$

This operation is called **scalar multiplication**. Note that $kA = [ka_{ij}] = [a_{ij}k]$ so the notation Ak will also be used, and then $kA = Ak$.

For example $5 \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 5 & 10 \\ 20 & 15 \end{bmatrix}$ and $0 \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

Inner product

The next question is how to multiply two matrices. We start with a special case.

Let $n \in \mathbb{N}$ and let A be a $1 \times n$ matrix (a row matrix) and let B be an $n \times 1$ column matrix. That is $A = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix}$ while $B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$. We define the **(inner) product** of A and B to be

$$AB = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n.$$

Note that the inner product is of a row and a column of the same length (the length of a row or column matrix is the number of entries in it) and that the answer is a real number (not a matrix).

For example

$$\begin{bmatrix} 2 & -1 & 3 \end{bmatrix} \begin{bmatrix} -2 \\ 5 \\ 3 \end{bmatrix} = 2(-2) + (-1)5 + 3(3) = -4 - 5 + 9 = 0.$$

Matrix multiplication

Let m, n and $p \in \mathbb{N}$. Suppose $A \in M_{m,n}$ and $B \in M_{n.p}$. Then we define the product of A and B to be the $m \times p$ matrix formed as follows:

The (i,j) entry of AB is the inner product of the i^{th} row of A with the j^{th} column of B.

Example: Let $A = \begin{bmatrix} 2 & 0 & -1 \\ 3 & 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & -3 & 2 & 1 \\ 0 & 4 & 2 & 1 \\ 2 & 0 & 2 & 3 \end{bmatrix}$. Then $A \in \mathbb{M}_{2,3}$ and $B \in \mathbb{M}_{3,4}$ implies $AB \in \mathbb{M}_{2,4}$ and

$$AB = \begin{bmatrix} 2 & 0 & -1 \\ 3 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -3 & 2 & 1 \\ 0 & 4 & 2 & 1 \\ 2 & 0 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 0 & -6 & 2 & -1 \\ 5 & -5 & 10 & 7 \end{bmatrix}$$

In general $AB \neq BA$ for matrices

In order to multiply two matrices, A by B, the number of columns of A must equal the number of rows of B. In other words, if $A \in \mathbb{M}_{m,n}$ and $B \in \mathbb{M}_{p,q}$ then to form AB we must have that $n = p$. To form BA we must have $q = m$.

In particular we note that $AB \neq BA$ in this case. Even if we take two matrices A and B so that AB and BA are the same size, we **cannot** assume that $AB = BA$. For example, if $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$ then

$$AB = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}$$

while

$$BA = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 23 & 34 \\ 31 & 46 \end{bmatrix} \neq AB.$$

Elementary Row operations

Given a matrix, the following operations on the rows of that matrix are called **(elementary) row operations**.

1. Interchange rows, e.g.

$$\begin{bmatrix} 0 & 1 & 2 \\ 3 & 1 & -3 \\ 2 & 9 & 4 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 3 & 1 & -3 \\ 0 & 1 & 2 \\ 2 & 9 & 4 \end{bmatrix}$$

2. Multiply a row by a non-zero real number

$$\begin{bmatrix} 3 & 1 & -3 \\ 0 & 1 & 2 \\ 2 & 9 & 4 \end{bmatrix} \xrightarrow{R_1 \to \frac{1}{3} R_1} \begin{bmatrix} 1 & \frac{1}{3} & -1 \\ 0 & 1 & 2 \\ 2 & 9 & 4 \end{bmatrix}$$

3. Add a scalar multiple of one row to another

$$\begin{bmatrix} 1 & \frac{1}{3} & -1 \\ 0 & 1 & 2 \\ 2 & 9 & 4 \end{bmatrix} \xrightarrow{R_3 \to R_3 - 2R_1} \begin{bmatrix} 1 & \frac{1}{3} & -1 \\ 0 & 1 & 2 \\ 2-2(1) & 9-\frac{2}{3} & 4-2(-1) \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{3} & -1 \\ 0 & 1 & 2 \\ 0 & 8\frac{1}{3} & 6 \end{bmatrix}$$

Echelon form

A matrix $A = [a_{ij}]$ is said to be in **echelon form** if $a_{ij} = 0$ for $i > j$. For example, $\begin{bmatrix} 1 & -2 & -2 & 3 \\ 0 & -2 & 3 & 17 \\ 0 & 0 & 5 & 1 \end{bmatrix}$ is in echelon form. A square matrix in echelon form is called **upper triangular**.

Any matrix can be converted to echelon form by using elementary row operations. The steps involved are as follows. Let $A = [a_{ij}] \in \mathbb{M}_{m,n}$.

1. Look at the first column. We want all entries below the first one, a_{11}, to be zero. If necessary (that is, if $a_{11} = 0$), interchange two rows so that $a_{11} \neq 0$.
2. Add multiples of the first row to the other rows to ensure a_{21}, a_{31}, $\ldots, a_{m1}$ are all zero. For example, add $\frac{-a_{21}}{a_{11}} \times$ row 1 to row 2, etc.
3. Move to the second column. We want all the entries below a_{22}, that is a_{32}, a_{42}, etc., to be zero. If necessary, interchange two rows (**not** the first row) so that $a_{22} \neq 0$.
4. Add multiples of the second row to ensure $a_{32}, a_{42}, \ldots, a_{m2}$ are all zero.
5. Move along the columns in this way until the matrix is in echelon form.

Example:

Convert the matrix $\begin{bmatrix} 0 & 1 & 2 \\ 3 & 1 & -3 \\ 2 & 9 & 4 \end{bmatrix}$ to echelon form.

Systems of linear equations

A **linear equation** in one variable is an equation of the form

$$ax = b$$

where a and b are given real numbers, and x is a variable.

A **linear equation in n variables** is an equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

where $a_1, \ldots, a_n, b \in \mathbb{R}$. For example, $2x_1 + 3x_2 = 7$ is a linear equation in two variables; $2x - 3y + 4z = -5$ is a linear equation in three variables.

We can also consider a **system of m equations in n variables**. That is,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_2 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad\quad &\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m. \end{aligned}$$

The system of m equations above can also be written as **one** matrix equation:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

and we will usually write this as $AX = B$. When written in this form, a solution of the system will be a $n \times 1$ column matrix with n elements (or n-vector). A linear system may have no solutions, many solutions or just a unique solution.

The method of Gaussian Elimination for solving systems

Suppose the system is $AX = B$ where $A \in \mathbb{M}_{m,n}$ and $B \in \mathbb{M}_{m,1}$. Form the augmented matrix $[A \vdots B]$ and convert this matrix to echelon form to get a matrix $[A' \vdots B']$. The echelon system $A'X = B'$ has the same solutions as the original system $AX = B$ but is much easier to solve. This process of solving a linear system is called Gaussian elimination.

Example: Solve the system of linear equations,

$$\begin{array}{rcrcrcr} & & x_2 & + & 2x_3 & = & 3 \\ x_1 & + & 2x_2 & + & 3x_3 & = & -2 \\ -x_1 & - & 2x_2 & - & 2x_3 & = & 0 \end{array}$$

More examples

Solve, by Gaussian elimination, the following systems of linear equations.

$$
\text{(i)} \qquad
\begin{array}{rrrl}
x_1 & -2x_2 & -2x_3 & = 4 \\
3x_1 & -x_2 & -x_3 & = 7 \\
6x_1 & +x_2 & +x_3 & = 5.
\end{array}
$$

$$
\text{(ii)} \qquad
\begin{array}{rrrrl}
x_1 & +x_2 & +x_3 & -x_4 & = 0 \\
3x_1 & -x_2 & -x_3 & +2x_4 & = 0 \\
4x_1 & +5x_2 & +x_3 & -x_4 & = 0 \\
9x_1 & +6x_2 & +2x_3 & -x_4 & = 0
\end{array}
$$

Gaussian Elimination with partial pivoting

Example: Consider the following system

$$\begin{aligned}
\frac{2}{3}x_1 + \frac{2}{7}x_2 + \frac{1}{5}x_3 &= \frac{43}{15} \\
\frac{1}{3}x_1 + \frac{1}{7}x_2 - \frac{1}{2}x_3 &= \frac{5}{6} \\
\frac{1}{5}x_1 - \frac{3}{7}x_2 + \frac{2}{5}x_3 &= -\frac{12}{5}
\end{aligned} \tag{1}$$

Exercise: Using Gaussian elimination and back substitution show that the exact solution is $x_1 = 1$, $x_2 = 7$ and $x_3 = 1$.

Now write the system using four decimal digit rounding arithmetic

$$\begin{aligned} 0.6667x_1 + 0.2857x_2 + 0.2000x_3 &= 2.867 \\ 0.3333x_1 + 0.1429x_2 - 0.5000x_3 &= 0.8333 \\ 0.2000x_1 - 0.4286x_2 + 0.4000x_3 &= -2.400 \end{aligned} \tag{2}$$

Remarks:

1. To obtain the 4-digit floating point approximation of a number, we use the rounding method (e.g. $\frac{2}{3} = 0.66666\ldots$ gives 0.6667.)
2. Whenever performing a calculation (e.g. addition or multiplication) involving two or more numbers, the 4-digit rounded form is used for each number and then the result of the calculation is rounded as well. For example, $0.1234 \times 1.567 = 0.193368\ldots = 0.1934$. This is the source of **roundoff error accummulation**.
3. When performing row operations to get the echelon form, the entries below the pivot are implicitly set equal to **zero**, as opposed to being explicitly calculated, in order to avoid unnecessary calculations.

The first step of the Gaussian elimination yields

$$\begin{aligned} 0.6667x_1 + 0.2857x_2 \quad & +0.2000x_3 = 2.867 \\ +0.0001x_2 \quad & -0.6000x_3 = -0.5997 \\ -0.5143x_2 \quad & +0.3400x_3 = -3.260 \end{aligned} \tag{3}$$

where the row operations performed were

$$R_2 - \frac{0.3333}{0.6667} \times R_1 = R_2 - 0.4999 \times R_1$$
$$R_3 - \frac{0.2000}{0.6667} \times R_1 = R_3 - 0.3 \times R_1$$

and the above entries have been calculated as follows:

$$0.1429 - 0.2857 \times 0.4999 = 0.1429 - 0.1428 = 0.0001$$
$$0.8333 - 2.867 \times 0.4999 = 0.8333 - 1.433 = -0.5997$$
$$-0.4286 - 0.2857 \times 0.3 = -0.4286 - 0.08571 = -0.5143$$

and so on.

The next step of the Gaussian elimination gives

$$\begin{aligned} 0.6667x_1 + 0.2857x_2 \quad & +0.2000x_3 = 2.867 \\ +0.0001x_2 \quad & -0.6000x_3 = -0.5997 \\ & -3086x_3 = -3087 \end{aligned} \tag{4}$$

which gives the solution:

$x_3 = 1.000$ (exact)
$x_2 = 3.000$ (relative error=57%)
and $x_1 = 2.715$ (relative error=200%)!

Remark:

The large errors obtained in the previous example were obtained because of a cancellation (or loss of significance) error which occurred while working on the first column. This introduced a small pivot, 0.0001, in the second column which amplified the error. We cannot always avoid cancellation errors but at least we can try to avoid the use of very small pivots.

The partial pivoting strategy is the simplest scheme which eliminates the use of small pivots.

For each column j let

$$M_i = \max_{j \leq i} |A_{ij}|$$

and let i_0 be the smallest value of i $(i > j)$ for which this maximum is achieved. Then interchange rows j and i_0.

In other words, while working on column j, look for the entry with the largest absolute value below the pivot (A_{ij} with $i \geq j$). If this largest entry appears in more than one row, take the one closest to the pivot row, i_0, and swap it with the current pivot row, j.

Exercise: Apply the technique of Gaussian elimination with partial pivoting to system (2) and verify if the error has improved.

$$\begin{pmatrix} 0.6667 & 0.2857 & 0.2000 & 2.867 \\ 0.3333 & 0.1429 & -0.5000 & 0.8333 \\ 0.2000 & -0.4286 & 0.4000 & -2.400 \end{pmatrix} \longrightarrow \quad (\frac{R_2 - 0.4999 \times R_1}{R_3 - 0.3 \times R_1})$$

$$\begin{pmatrix} 0.6667 & 0.2857 & 0.2000 & 2.867 \\ 0 & 0.0001 & -0.6000 & -0.5997 \\ 0 & -0.5143 & 0.3400 & -3.260 \end{pmatrix} \longrightarrow (R_2 \longleftrightarrow R_3)$$

$$\begin{pmatrix} 0.6667 & 0.2857 & 0.2000 & 2.867 \\ 0 & -0.5143 & 0.3400 & -3.260 \\ 0 & 0.0001 & -0.6000 & -0.5997 \end{pmatrix} \longrightarrow (R_3 - 0.0001944 \times R_2)$$

$$\begin{pmatrix} 0.6667 & 0.2857 & 0.2000 & 2.867 \\ 0 & -0.5143 & 0.3400 & -3.260 \\ 0 & 0 & -0.5999 & -0.603 \end{pmatrix}$$

Partial pivoting gives the solution $x_3 = 1.001$, $x_2 = 7.000$ and $x_1 = 1.000$.

Exercise:

Consider the following system

$$
\begin{aligned}
3x_1 + x_2 + 4x_3 - x_4 &= 7\\
2x_1 - 2x_2 - x_3 + 2x_4 &= 1\\
5x_1 + 7x_2 + 14x_3 - 8x_4 &= 20\\
x_1 + 3x_2 + 2x_3 + 4x_4 &= -4
\end{aligned}
$$

1. Show, by direct substitution, that the exact solution is (1,-1,1,-1).
2. Solve the system using Gaussian elimination without pivoting and calculate the relative errors.
3. Solve the system using Gaussian elimination with partial pivoting and calculate the relative errors.

Scaled partial pivoting

Consider now the example

$$0.7x_1 + 1725x_2 = 1739$$
$$0.4352x_1 - 5.433x_2 = 3.271$$

The exact solution is $x_1 = 20$ and $x_2 = 1$. Partial pivoting would leave the equations unchanged since 0.7 is the largest entry in the first column. Gaussian elimination then gives

$$0.7x_1 + 1725x_2 = 1739$$
$$-1077x_2 = -1078$$

hence $x_2 = 1.001$ and $x_1 = 17.14$.

The large errors obtained in this case are due to the fact that the entries 0.7 and 0.4352 are not compared with the other coefficients in their corresponding rows. For instance, $0.7 > 0.4352$ but

$$\frac{0.7}{1725} < \frac{0.4325}{5.433}$$

If, instead, we choose the pivot which is largest in magnitude relative to the other coefficients in the equation, we need to swap the rows and use 0.4352 as pivot

$$0.4352x_1 - 5.433x_2 = 3.271$$
$$0.7x_1 + 1725x_2 = 1739$$

which gives $x_2 = 1.000$ and $x_2 = 20.000$.

Scaled partial pivoting

Construct a scale vector as follows

$$S_i = \max_{1 \le j \le n} |a_{ij}|, \qquad \text{for each} \quad 1 \le i \le n$$

(S_i is the largest absolute value in row i.)

During step i of the Gaussian elimination (column i), let

$$M_i = \max_{i \le j \le n} \left(\frac{|a_{ji}|}{S_j} \right)$$

and let j_0 be the smallest j for which this maximum occurs. If $j_0 > i$ then we interchange rows i and j_0.

Note: The scale vector S does not change during the Gaussian elimination process.

Exercise:

Consider the system

$$
\begin{aligned}
3x_1 + x_2 + 4x_3 - x_4 &= 7 \\
2x_1 - 2x_2 - x_3 + 2x_4 &= 1 \\
5x_1 + 7x_2 + 14x_3 - 8x_4 &= 20 \\
x_1 + 3x_2 + 2x_3 + 4x_4 &= -4
\end{aligned}
$$

Recall that the exact solution of this system is (1,-1,1,-1). Calculate the solution using Gaussian elimination with scaled partial pivoting and compare the errors with those obtained using (simple) partial pivoting.

The LU decomposition method

Consider the system of equations, written in matrix form

$$AX = B \tag{5}$$

where $A \in \mathbb{M}_{n,n}$ is a square matrix and X, $B \in \mathbb{M}_{n,1}$.

The LU decomposition (or factorisation) of the matrix A consists of writing the matrix A as the product of a **lower triangular** matrix L and an **upper triangular** matrix U, such that $A = LU$. (Recall that a matrix is called upper triangular if all entries below the main diagonal are zero and lower triangular if all entries above the main diagonal are zero.)

The procedure of solving the system of equations (5) using LU decomposition is the following.

Step 1: Determining the LU decomposition The matrices L and U can be determined by writing (in the case of a 3×3 matrix, for example)

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = L \cdot U = \begin{pmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{pmatrix} \cdot \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{pmatrix}$$

and solving this system for the entries l_{ij} and u_{ij}.

Step 2: Solving the LU decomposed system

After the LU factorisation, system (5) becomes

$$LUX = B$$

which can be solved as two successive problems:

$$(1) \quad LY = B; \qquad\qquad (2) \quad UX = Y.$$

Each of these two systems is easy to solve since the matrix of coefficients is either upper or lower triangular.

Example:

Solve the following system using the LU decomposition method.

$$
\begin{aligned}
x_1 + x_2 - x_3 &= 1 \\
x_1 + 2x_2 - 2x_3 &= 5 \\
-2x_1 + x_2 + x_3 &= 10
\end{aligned}
$$

The Inverse of a Matrix

Let $A \in \mathbb{M}_{n,n}$. If there exists $B \in M_{n,n}$ such that $AB = I = BA$ then we say A is **invertible** and B is an **inverse of** A. We write $B = A^{-1}$.

Note that if B is an inverse of A, then A is an inverse of B.

Procedure to invert a matrix:

We can use elementary row operations to invert a 2×2 invertible matrix. Start with $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and "augment" the identity matrix to form the 2×4 matrix

$$\left[\begin{array}{cc:cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right].$$

Now use elementary row operations to convert the left hand side to the identity matrix but performing the row operations across the entire row.

When the left hand matrix becomes equal to the identity, the right hand matrix will be the inverse of A.

Example:

$$\begin{bmatrix} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{bmatrix} \xrightarrow{R_2 \to R_2 - 3R_1} \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{bmatrix} \xrightarrow{R_1 \to R_1 + R_2} \begin{bmatrix} 1 & 0 & -2 & 1 \\ 0 & -2 & -3 & 1 \end{bmatrix}$$

$$\xrightarrow{R_2 \to -\frac{1}{2} R_2} \begin{bmatrix} 1 & 0 & -2 & 1 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{bmatrix}.$$

Hence the inverse of $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ is $\begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}$.

This procedure can be used to invert any $n \times n$ matrix, if it is invertible.

Example: Find the inverse of the matrix $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ -1 & -2 & -2 \end{bmatrix}$.

The answer is $A^{-1} = \begin{bmatrix} -2 & 2 & 1 \\ 1 & -2 & -2 \\ 0 & 1 & 1 \end{bmatrix}$. Check that $AA^{-1} = I = A^{-1}A$.

Note: A square matrix is invertible if and only if it does not contain a zero on the main diagonal after it has been reduced to echelon form by elementary row operations.

For example, the matrix $A = \begin{bmatrix} 1 & 2 & -3 \\ 3 & -1 & -5 \\ 2 & -3 & -2 \end{bmatrix}$ has no inverse!

Solving systems of linear equations

The inverse of a matrix gives a new method for solving a system of linear equations:

$$\text{If} \quad AX = B \qquad \text{then} \qquad X = A^{-1}B$$

provided A is invertible.

Exercise:
Solve the following system of equations by inverting the matrix of coefficients.

$$x + 2y + 3z = 1$$
$$x - y + 4z = 4$$
$$2x - 2y + z = 8$$

Properties of the LU decomposition method:

1: Not every matrix has an LU-decomposition. For example,

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

cannot be written as the product of a lower triangular and upper triangular matrices.

However, given any invertible matrix A, it is always possible to rearrange the rows of A so that the resulting matrix does have an LU-decomposition.

2. When a matrix has an LU-decomposition, that decomposition is not unique. Given the factorisation $A = LU$, if we define the new matrices

$$L' = LD \qquad \text{and} \qquad U' = D^{-1}U$$

where D is any invertible diagonal matrix, then L' is lower triangular, U' is upper triangular and $A = L'U'$ so we found another LU-factorisation.

In fact, once a matrix A admits an LU-factorisation, there are an infinite number of choices for the matrices L and U!

3. The LU-decomposition is unique up to scaling by a diagonal matrix. This means that, if A admits two LU-factorisations,

$$A = L_1 U_1 = L_2 U_2$$

then we must have

$$L_1 = L_2 D \qquad \text{and} \qquad U_2 = D U_1$$

for some diagonal matrix D.

4. If A admits a LU-factorisation then we can find matrices L and U such that L has 1's along the diagonal. (Similarly, we can find a decomposition in which U has 1's along the diagonal.)

5. The decomposition $A = LU$ is unique if we require that L (or U) has 1's along its diagonal!