MATH141-Linear Systems and Matrices

UNIVERSITY
OF WOLLONGONG
IN DUBAI

Dr. Assane Lo

University of Wolongong in Dubai

Introduction to Linear Systems

A **linear equation** in variables $x_1, x_2, ..., x_n$ has the form

$$a_1x_1 + a_2x_2 + a_3x_3 + \cdots + a_nx_n = d$$

where the numbers $a_1, ..., a_n \in \mathbb{R}$ are the equation's coefficients and $d \in \mathbb{R}$ is the constant.
An n-tuple $(s_1, s_2, ..., s_n) \in \mathbb{R}^n$ is a solution of, or satisfies, that equation if substituting the numbers $s_1, s_2, ..., s_n$ for the variables gives a true statement:

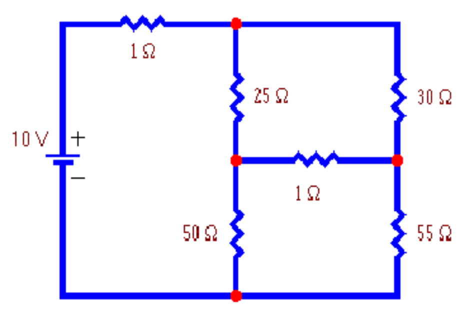
$$a_1s_1 + a_2s_2 + a_3s_3 + \cdots + a_ns_n = d$$

A **system of linear equations** or **linear system** is a collection of one or more linear equations involving the same variables $x_1, x_2, ..., x_n$.
It is of the form

$$\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n = d_1\\
&a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n = d_2\\
&\vdots\\
&a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n = d_m
\end{aligned}$$

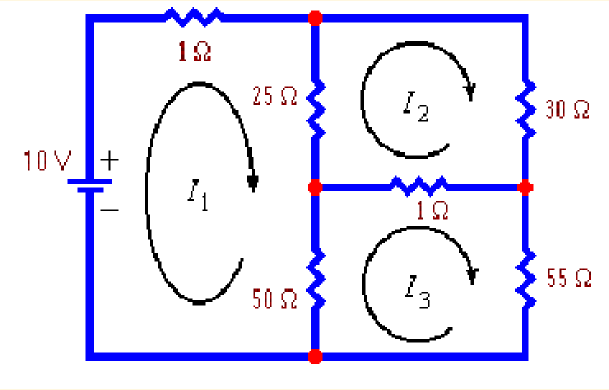
It has the solution $(s_1, s_2, ..., s_n) \in \mathbb{R}^n$ if that n-tuple is a solution of all of the equations in the system.

One of the most important applications of linear algebra to electronics is to analyze electronic circuits.



The goal is to calculate the current flowing in each branch of the circuit or to calculate the voltage at each node of the circuit.

To ind the current flowing in each branch of this circuit, we first observe that the number of loop currents required is 3.



After writing down Kirchoff's Voltage Law for each loop. The result is the following system of equations:

$$\begin{cases} 1i_1 + 25(i_1 - i_2) + 50(i_1 - i_3) = 10 \\ 25(i_2 - i_1) + 30i_2 + 1(i_2 - i_3) = 0 \\ 50(i_3 - i_1) + 1(i_3 - i_2) + 55i_3 = 0 \end{cases}$$

Example 1 The ordered pair $(-1, 5)$ is a solution of this system

$$\begin{aligned} 3x_1 + 2x_2 &= 7 \\ -x_1 + x_2 &= 6 \end{aligned}$$

In contrast, $(5, -1)$ is not a solution.

Finding the set of all solutions is solving the system.
No guesswork or good fortune is needed to solve a linear system.
There is an algorithm that always works.
The next example introduces that algorithm, called **Gauss' method** or **Gaussian elimination.**
It transforms the system, step by step, into one with a form that is easily solved.

Example 2 Solve the system

$$\begin{aligned} &3x_3 = 9 \\ &x_1 + 5x_2 - 2x_3 = 2 \\ &\frac{1}{3}x_1 + 2x_2 = 3 \end{aligned}$$

Solution:

We repeatedly transform it until it is in a form that is easy to solve. First, swap row 1 with row 3 to get

$$\begin{aligned}
&\frac{1}{3}x_1 + 2x_2 = 3\\
&x_1 + 5x_2 - 2x_3 = 2\\
&3x_3 = 9
\end{aligned}$$

Next, multiply row 1 by 3 to get

$$
\begin{aligned}
&x_1 + 6x_2 = 9 \\
&x_1 + 5x_2 - 2x_3 = 2 \\
&3x_3 = 9
\end{aligned}
$$

Finally add -1 times row 1 to row 2 to get

$$
\begin{aligned}
&x_1 + 6x_2 = 9 \\
&-x_2 - 2x_3 = -7 \\
&3x_3 = 9
\end{aligned}
$$

The third step is the only nontrivial one. We've mentally multiplied both sides of the first row by -1,
mentally added that to the old second row, and written the result in as the new second row.
Now we can find the value of each variable.
The bottom equation shows that $x_3 = 3$. Substituting 3 for x_3 in the middle equation shows that $x_2 = 1$.
Substituting those two into the top equation gives that $x_1 = 3$
and so the system has a unique solution: the solution set is $\{(3, 1, 3)\}$.
This method of solving a linear equation is called Gauss Method and is based on the following Theorem

If a linear system is changed to another by one of these operations:

(1) an equation is swapped with another
(2) an equation has both sides multiplied by a nonzero constant
(3) an equation is replaced by the sum of itself and a multiple of another

then the two systems have the same set of solutions.

Remark

Each of those three operations has a restriction.
Multiplying a row by 0 is not allowed because obviously that can change the solution set of the system.
Similarly, adding a multiple of a row to itself is not allowed because adding -1 times the row to itself has the effect of multiplying the row by 0.
The three operations from the Theorem above are the elementary reduction operations, or row operations, or Gaussian operations.
They are swapping, multiplying by a scalar or rescaling, and pivoting.

When writing out the calculations, we will abbreviate 'row i' by 'ρ_i'.

Example 3 Solve the system

$$\begin{aligned} &x + y = 0 \\ &2x - y + 3z = 3 \\ &x - 2y - z = 3 \end{aligned}$$

The first transformation of the system involves using the first row to eliminate the x in the second row and the x in the third.
To get rid of the second row's 2x, we multiply the entire first row by -2, add that to the second row, and write the result in as the new second row.
To get rid of the third row's x, we multiply the first row by -1, add that to the third row, and write the result in as the new third row. That is

$$\xrightarrow[-\rho_1+\rho_3]{-2\rho_1+\rho_2} \begin{array}{rcrcrcl} x & + & y & & & = & 0 \\ & & -3y & + & 3z & = & 3 \\ & & -3y & - & z & = & 3 \end{array}$$

To finish we transform the second system into a third system, where the last equation involves only one unknown.
This transformation uses the second row to eliminate y from the third row.

$$\xrightarrow{-\rho_2+\rho_3} \begin{array}{rrrr} x + & y & & = 0 \\ & -3y + & 3z & = 3 \\ & & -4z & = 0 \end{array}$$

Now we are set up for the solution. The third row shows that $z = 0$. Substitute that back into the second row to get $y = -1$, and then substitute back into the first row to get $x = 1$.

In each row, the first variable with a nonzero coefficient is the row's **leading variable**.

A system is in **echelon form** if each leading variable is to the right of the leading variable in the row above it (except for the leading variable in the first row).

The only operation needed in the examples above is pivoting.

Here is a linear system that requires the operation of swapping equations.

After the first pivot

Example 4 Solve the system

$$
\begin{aligned}
&x - y = 0 \\
&2x - 2y + z + 2w = 4 \\
&y + w = 0 \\
&2z + w = 5
\end{aligned}
$$

Solution

$$\xrightarrow{-2\rho_1+\rho_2}\quad \begin{aligned} x - y \qquad\qquad &= 0 \\ z + 2w &= 4 \\ y \qquad + \quad w &= 0 \\ 2z + \quad w &= 5 \end{aligned}$$

The second equation has no leading y. To get one, we look lower down in the system for a row that has a leading y and swap it in to get

$$\xrightarrow{\rho_2\leftrightarrow\rho_3}\quad \begin{aligned} x - y \qquad\qquad &= 0 \\ y \qquad + \quad w &= 0 \\ z + 2w &= 4 \\ 2z + \quad w &= 5 \end{aligned}$$

(Had there been more than one row below the second with a leading y then we could have swapped in any one.) The rest of Gauss' method goes as before.

$$\xrightarrow{-2\rho_3+\rho_4}\begin{array}{rcrcrcr} x-y & & & & & = & 0 \\ y & + & & & w & = & 0 \\ & & z & + & 2w & = & 4 \\ & & & & -3w & = & -3 \end{array}$$

Back-substitution gives $w = 1$, $z = 2$, $y = -1$, and $x = -1$.

All of the systems seen so far have the same number of equations as unknowns.
All of them have a unique solution.
We will now discuss the situations where we have more equations than variables.
Example 5. Solve the linear system

$$\begin{aligned} x + 3y &= 1 \\ 2x + y &= -3 \\ 2x + 2y &= -2 \end{aligned}$$

Solution

This system has 3 equations with 2 variables. We will still use Gauss Method

$$\xrightarrow[-2\rho_1+\rho_3]{-2\rho_1+\rho_2} \begin{aligned} x + 3y &= 1 \\ -5y &= -5 \\ -4y &= -4 \end{aligned}$$

This shows that one of the equations is redundant. Echelon form gives

$$\xrightarrow{-(4/5)\rho_2+\rho_3} \begin{aligned} x + 3y &= 1 \\ -5y &= -5 \\ 0 &= 0 \end{aligned}$$

Thus $y = 1$ and $x = -2$. The '$0 = 0$' is derived from the redundancy.

Example 6 Solve the linear system

$$\begin{aligned} x + 3y &= 1 \\ 2x + y &= -3 \\ 2x + 2y &= 0 \end{aligned}$$

Solution Again using Gauss Method, we have

$$\xrightarrow[-2\rho_1+\rho_3]{-2\rho_1+\rho_2} \begin{aligned} x + \quad 3y &= \ \ 1 \\ -5y &= -5 \\ -4y &= -2 \end{aligned}$$

Here the system is inconsistent: no pair of numbers satisfies all of the equations simultaneously. Echelon form makes this inconsistency obvious.

$$\xrightarrow{-(4/5)\rho_2+\rho_3} \begin{aligned} x + \quad 3y &= \ \ 1 \\ -5y &= -5 \\ 0 &= \ \ 2 \end{aligned}$$

This system does not have a solution. The solution set is empty.

A system of linear equations has

1. no solution, or
2. exactly one solution, or
3. infinitely many solutions.

Matrix Notation

An $n \times m$ **matrix** is a rectangular array of numbers with n rows and m columns. Each number in the matrix is called an **entry**.
Matrices are usually named by upper case roman letters, e.g. A. Each entry is denoted by the corresponding lower-case letter, e.g. a_{ij} is the number in row i and column j of the array.

For instance,

$$A = \begin{pmatrix} 1 & 2.2 & 5 \\ 3 & 4 & -7 \end{pmatrix}$$

has two rows and three columns, and so is a 2×3 matrix. (Read that "two by three"; the number of rows is always stated first.)

The entry in the second row and first column is $a_{21} = 3$. Note that the order of the subscripts matters: $a_{12} \neq a_{21}$ since $a_{12} = 2.2$.

The parentheses around the array are a typographic device so that when two matrices are side by side we can tell where one ends and the other starts. We may also use brackets instead of parentheses.

The collection of all $m \times n$ matrices will be denoted by $\mathcal{M}_{m\times n}$.

The essential information of a linear system can be recorded compactly in a matrix.
This will considerably simplify the notation in the Gaussian elimination process.
Example 7 Consider the linear system

$$\begin{aligned} x_1 + 2x_2 &= 4 \\ x_2 - x_3 &= 0 \\ x_1 + 2x_3 &= 4 \end{aligned}$$

Because the elementary operations in Gauss method only involve the coefficients, we may write the system in the matrix form

$$\left(\begin{array}{ccc|c} 1 & 2 & 0 & 4 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 2 & 4 \end{array}\right)$$

The vertical bar just reminds us of the difference between the coefficients on the systems's left hand side and the constants on the right.
When a bar is used to divide a matrix into parts, we call it an **augmented matrix**. The **coefficient matrix** is

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

with this matrix notation, Gauss' method goes as following:

$$\left(\begin{array}{ccc|c} 1 & 2 & 0 & 4 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 2 & 4 \end{array}\right) \xrightarrow{-\rho_1+\rho_3} \left(\begin{array}{ccc|c} 1 & 2 & 0 & 4 \\ 0 & 1 & -1 & 0 \\ 0 & -2 & 2 & 0 \end{array}\right) \xrightarrow{2\rho_2+\rho_3} \left(\begin{array}{ccc|c} 1 & 2 & 0 & 4 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

In the final step of the elimination, the second row stands for $x_2 - x_3 = 0$ and the first row stands for $x_1 + 2x_2 = 4$ so the system becomes

$$\begin{aligned} x_1 + 2x_2 &= 4 \\ x_2 - x_3 &= 0 \\ 0 &= 0 \end{aligned}$$

We have $x_3 = x_2$ and $x_1 = 4 - 2x_2$

The solution (x_1, x_2, x_3) can be written as

$$(x_1, x_2, x_3) = (4 - 2x_2, x_2, x_2)$$

The equation has infinitely many solutions and the solution set is given by

$$\{(4 - 2x_2, x_2, x_2) \mid x_2 \in \mathbb{R}\}$$

A vector (or column vector) is a matrix with a single column. A matrix with a single row is a row vector. The entries of a vector are its components.

Vectors are an exception to the convention of representing matrices with capital roman letters. We use lower-case roman or greek letters.

Example

$$x = \begin{pmatrix} 1 \\ \text{-2} \\ 9 \end{pmatrix}$$

The linear equation

$$a_1x_1 + a_2x_2 + a_3x_3 + \cdots + a_nx_n = d$$

with unknowns $x_1, x_2, ..., x_n$ is satisfied by

$$s = \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix}$$

if $a_1s_1 + a_2s_2 + a_3s_3 + \cdots + a_ns_n = d$
A vector satisfies a linear system if it satisfies each equation in the system.

The style of description of solution sets that we will use involves adding the vectors, and also multiplying them by real numbers.
We need to define these operations.
Let

$$u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \text{ and } v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

The vector sum of u and v is

$$u + v = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} + \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{pmatrix}$$

In general, two matrices with the same number of rows and the same number of columns add in this way, entry-by-entry.
The scalar multiplication of the real number r and the vector v is

$$r \cdot v = \begin{pmatrix} rv_1 \\ \vdots \\ rv_n \end{pmatrix}$$

In general, any matrix is multiplied by a real number in this entry-by-entry way.

Example 8

$$\begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} + \begin{pmatrix} 3 \\ -1 \\ 4 \end{pmatrix} = \begin{pmatrix} 2+3 \\ 3-1 \\ 1+4 \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \\ 5 \end{pmatrix}$$

and

$$7 \cdot \begin{pmatrix} 1 \\ 4 \\ -1 \\ -3 \end{pmatrix} = \begin{pmatrix} 7 \\ 28 \\ -7 \\ -21 \end{pmatrix}$$

A column or row vector of all zeros is defined to be the zero vector and is denoted by O.

Matrix Operations

A rectangular array of numbers of the form

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

is called an $m \times n$ matrix, with m rows and n columns.

We count rows from the top and columns from the left. Hence

$$\begin{pmatrix} a_{i1} & \dots & a_{in} \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

represent respectively the i-th row and the j-th column of the matrix, and a_{ij} represents the entry in the matrix on the i-th row and j-th column.

Example: Consider the 3×4 matrix

$$\begin{pmatrix} 2 & 4 & 3 & -1 \\ 3 & 1 & 5 & 2 \\ -1 & 0 & 7 & 6 \end{pmatrix}$$

$$\begin{pmatrix} 3 & 1 & 5 & 2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 3 \\ 5 \\ 7 \end{pmatrix}$$

represent respectively the 2-nd row and the 3-rd column of the matrix, and 5 represents the entry in the matrix on the 2-nd row and 3-rd column.

We now consider the question of arithmetic involving matrices. First of all, let us study the problem of addition.
Suppose that the two matrices

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \dots & b_{mn} \end{pmatrix}$$

both have m rows and n columns. Then we write

$$A + B = \begin{pmatrix} a_{11} + b_{11} & \dots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \dots & a_{mn} + b_{mn} \end{pmatrix}$$

and call this the sum of the two matrices A and B.

Example 8 Suppose that

$$A = \begin{pmatrix} 2 & 4 & 3 & -1 \\ 3 & 1 & 5 & 2 \\ -1 & 0 & 7 & 6 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 2 & -2 & 7 \\ 0 & 2 & 4 & -1 \\ -2 & 1 & 3 & 3 \end{pmatrix}.$$

Then

$$A + B = \begin{pmatrix} 2+1 & 4+2 & 3-2 & -1+7 \\ 3+0 & 1+2 & 5+4 & 2-1 \\ -1-2 & 0+1 & 7+3 & 6+3 \end{pmatrix} = \begin{pmatrix} 3 & 6 & 1 & 6 \\ 3 & 3 & 9 & 1 \\ -3 & 1 & 10 & 9 \end{pmatrix}.$$

Remark. There is no definition for adding matrices having different sizes. For instance we cannot add the following two matrices

$$\begin{pmatrix} 2 & 4 & 3 & -1 \\ -1 & 0 & 7 & 6 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 2 & 4 & 3 \\ 3 & 1 & 5 \\ -1 & 0 & 7 \end{pmatrix}.$$

Suppose that A, B, C are m $\times$ n matrices. Suppose further that O_{mn} represents the m $\times$ n matrix with all entries zero. Then

(a) $A + B = B + A$;
(b) $A + (B + C) = (A + B) + C$;
(c) $A + O_{mn} = A$; and
(d) there is an m $\times$ n matrix A' such that $A + A' = O_{mn}$

Suppose that the matrix

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

has m rows and n columns, and that c $\in \mathbb{R}$.

Then

$$cA = \begin{pmatrix} ca_{11} & \dots & ca_{1n} \\ \vdots & & \vdots \\ ca_{m1} & \dots & ca_{mn} \end{pmatrix}$$

and call this the product of the matrix A by the scalar c.

Example 9 Suppose that

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

Then

$$2A = \begin{pmatrix} 4 & 8 & 6 & -2 \\ 6 & 2 & 10 & 4 \\ -2 & 0 & 14 & 12 \end{pmatrix}.$$

Suppose that A, B are m $\times$ n matrices, and that c, d $\in \mathbb{R}$. Suppose further that O_{mn} represents the m $\times$ n matrix with all entries zero. Then

(a) $c(A + B) = cA + cB$;
b) $(c + d)A = cA + dA$;
(c) $0A = O_{mn}$; and
(d) $c(dA) = (cd)A$.

Matrix Multiplication

The question of multiplication of two matrices is rather more complicated. To motivate this, let us consider the representation of a system of linear equations

$$\begin{aligned} a_{11}x_1 + \ldots + a_{1n}x_n &= b_1, \\ &\vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n &= b_m, \end{aligned}$$

in the form $Ax = b$, where

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

represent the coefficients

and

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

represents the variables. This can be written in full matrix notation by

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Can you work out the meaning of this representation?

Now let us define matrix multiplication more formally.
Suppose that

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_{11} & \dots & b_{1p} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{np} \end{pmatrix}$$

are respectively an m $\times$ n matrix and an n $\times$ p matrix.

Then the matrix product AB is given by the m $\times$ p matrix

$$AB = \begin{pmatrix} q_{11} & \cdots & q_{1p} \\ \vdots & & \vdots \\ q_{m1} & \cdots & q_{mp} \end{pmatrix},$$

where for every i = 1, . . . , m and j = 1, . . . , p, we have

$$q_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = a_{i1} b_{1j} + a_{i2} b_{2j} + \ldots + a_{in} b_{nj}$$

Remark.

Note first of all that the number of columns of the first matrix must be equal to the number of rows of the second matrix.

On the other hand, for a simple way to work out q_{ij}, the entry in the i-th row and j-th column of AB, we observe that the i-th row of A and the j-th column of B are respectively

$$\begin{pmatrix} a_{i1} & \dots & a_{in} \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{pmatrix}.$$

We now multiply the corresponding entries – from a_{i1} with b_{j1}, and so on, until a_{in} with b_{nj} – and then add these products to obtain q_{ij}.

Example 10 Consider the matrices

$$A = \begin{pmatrix} 2 & 4 & 3 & -1 \\ 3 & 1 & 5 & 2 \\ -1 & 0 & 7 & 6 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 4 \\ 2 & 3 \\ 0 & -2 \\ 3 & 1 \end{pmatrix}.$$

Note that A is a 3 $\times$ 4 matrix and B is a 4 $\times$ 2 matrix, so that the product AB is a 3 $\times$ 2 matrix. Let us calculate the product

$$AB = \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \\ q_{31} & q_{32} \end{pmatrix}.$$

Consider first of all q_{11}. To calculate this, we need the 1-st row of A and the 1-st column of B, so let us cover up all unnecessary information, so that

$$\begin{pmatrix} 2 & 4 & 3 & -1 \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{pmatrix} \begin{pmatrix} 1 & \times \\ 2 & \times \\ 0 & \times \\ 3 & \times \end{pmatrix} = \begin{pmatrix} q_{11} & \times \\ \times & \times \\ \times & \times \end{pmatrix}.$$

From the definition we have

$$q_{11} = 2 \cdot 1 + 4 \cdot 2 + 3 \cdot 0 + (-1) \cdot 3 = 2 + 8 + 0 - 3 = 7.$$

Consider next q_{12}. To calculate this, we need the 1-st row of A and the 2-nd column of B, so let us cover up all unnecessary information, so that

$$\begin{pmatrix} 2 & 4 & 3 & -1 \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{pmatrix} \begin{pmatrix} \times & 4 \\ \times & 3 \\ \times & -2 \\ \times & 1 \end{pmatrix} = \begin{pmatrix} \times & q_{12} \\ \times & \times \\ \times & \times \end{pmatrix}.$$

From the definition we have

$$q_{12} = 2 \cdot 4 + 4 \cdot 3 + 3 \cdot (-2) + (-1) \cdot 1 = 8 + 12 - 6 - 1 = 13.$$

We do the same for the remaining entries to get

$$AB = \begin{pmatrix} 2 & 4 & 3 & -1 \\ 3 & 1 & 5 & 2 \\ -1 & 0 & 7 & 6 \end{pmatrix} \begin{pmatrix} 1 & 4 \\ 2 & 3 \\ 0 & -2 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 7 & 13 \\ 11 & 7 \\ 17 & -12 \end{pmatrix}.$$

Suppose that A is an $m \times n$ matrix, B is an $n \times p$ matrix and C is an $p \times r$ matrix. Then $A(BC) = (AB)C$.

(a) Suppose that A is an $m \times n$ matrix and B and C are $n \times p$ matrices. Then $A(B + C) = AB + AC$.

(b) Suppose that A and B are $m \times n$ matrices and C is an $n \times p$ matrix. Then $(A + B)C = AC + BC$.

Suppose that A is an $m \times n$ matrix, B is an $n \times p$ matrix, and that $c \in \mathbb{R}$. Then $c(AB) = (cA)B = A(cB)$

Inversion of Matrices

For the remainder of this chapter, we shall deal with square matrices, those where the number of rows equals the number of columns.

The n $\times$ n matrix

$$I_n = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix},$$

where

$$a_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases}$$

is called the identity matrix of order n.

Note that

$$I_1 = (1) \qquad \text{and} \qquad I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The following result is relatively easy to check. It shows that the identity matrix I_n acts as the identity for multiplication of n $\times$ n matrices.

The Gauss-Jordan Method for Finding the inverse of a Matrix

In this section, we shall discuss a technique by which we can find the inverse of a square matrix, if the inverse exists.
Before we discuss this technique, let us recall the three elementary row operations that we discussed previously .
These are: (1) interchanging two rows; (2) adding a multiple of one row to another row; and (3) multiplying one row by a non-zero constant.

By an elementary n $\times$ n matrix, we mean an n $\times$ n matrix obtained from I_n by an elementary row operation.
Suppose that A is an n $\times$ n matrix, and suppose that B is obtained from A by an elementary row operation. Suppose further that E is an elementary matrix obtained from I_n by the same elementary row operation. Then

$$B = EA$$

Example 11 Consider the matrices

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \quad \text{and} \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Let us interchange rows 1 and 2 of A and do likewise for I_3. We obtain respectively

$$\begin{pmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Note that

$$\begin{pmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Let us interchange rows 2 and 3 of A and do likewise for I_3. We obtain respectively

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Note that

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Let us add 3 times row 1 to row 2 of A and do likewise for I_3. We obtain respectively

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 3a_{11}+a_{21} & 3a_{12}+a_{22} & 3a_{13}+a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Note that

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 3a_{11}+a_{21} & 3a_{12}+a_{22} & 3a_{13}+a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Let us add -2 times row 3 to row 1 of A and do likewise for $\mathbf{I}_3$. We obtain respectively

$$\begin{pmatrix} -2a_{31}+a_{11} & -2a_{32}+a_{12} & -2a_{33}+a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Note that

$$\begin{pmatrix} -2a_{31}+a_{11} & -2a_{32}+a_{12} & -2a_{33}+a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

In other words, we consider an array with the matrix A on the left and the matrix I_n on the right.
We now perform elementary row operations on the array and try to reduce the left hand half to the matrix I_n
If we succeed in doing so, then the right hand half of the array gives the inverse A^{-1}.

Example. Consider the matrix

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 3 & 0 & 3 \\ -2 & 3 & 0 \end{pmatrix}.$$

To find A^{-1} we consider the array

$$(A \mid I_3) = \begin{pmatrix} 1 & 1 & 2 & 1 & 0 & 0 \\ 3 & 0 & 3 & 0 & 1 & 0 \\ -2 & 3 & 0 & 0 & 0 & 1 \end{pmatrix}$$

We now perform elementary row operations on this array and try to reduce the left hand half to the
matrix I_3. Note that if we succeed, then the final array is clearly in reduced row echelon form.
We therefore follow the same procedure as reducing an array to reduced row echelon form.
Adding -3 times row 1 to row 2, we obtain

$$\begin{pmatrix} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & -3 & -3 & -3 & 1 & 0 \\ -2 & 3 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Adding 2 times row 1 to row 3, we obtain

$$\begin{pmatrix} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & -3 & -3 & -3 & 1 & 0 \\ 0 & 5 & 4 & 2 & 0 & 1 \end{pmatrix}.$$

Multiplying row 3 by 3, we obtain

$$\begin{pmatrix} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & -3 & -3 & -3 & 1 & 0 \\ 0 & 15 & 12 & 6 & 0 & 3 \end{pmatrix}.$$

Adding 5 times row 2 to row 3, we obtain

$$\begin{pmatrix} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & -3 & -3 & -3 & 1 & 0 \\ 0 & 0 & -3 & -9 & 5 & 3 \end{pmatrix}.$$

Multiplying row 1 by 3, we obtain

$$\begin{pmatrix} 3 & 3 & 6 & 3 & 0 & 0 \\ 0 & -3 & -3 & -3 & 1 & 0 \\ 0 & 0 & -3 & -9 & 5 & 3 \end{pmatrix}.$$

Adding 2 times row 3 to row 1, we obtain

$$\begin{pmatrix} 3 & 3 & 0 & -15 & 10 & 6 \\ 0 & -3 & -3 & -3 & 1 & 0 \\ 0 & 0 & -3 & -9 & 5 & 3 \end{pmatrix}.$$

Adding -1 times row 3 to row 2, we obtain

$$\begin{pmatrix} 3 & 3 & 0 & -15 & 10 & 6 \\ 0 & -3 & 0 & 6 & -4 & -3 \\ 0 & 0 & -3 & -9 & 5 & 3 \end{pmatrix}.$$

Adding 1 times row 2 to row 1, we obtain

$$\begin{pmatrix} 3 & 0 & 0 & -9 & 6 & 3 \\ 0 & -3 & 0 & 6 & -4 & -3 \\ 0 & 0 & -3 & -9 & 5 & 3 \end{pmatrix}.$$

Multiplying row 1 by 1/3, row 2 by -1/3 and row 3 by -1/3, we obtain

$$\begin{pmatrix} 1 & 0 & 0 & -3 & 2 & 1 \\ 0 & 1 & 0 & -2 & 4/3 & 1 \\ 0 & 0 & 1 & 3 & -5/3 & -1 \end{pmatrix}.$$

Note now that the array is in reduced row echelon form, and that the left hand half is the identity matrix I_3.
It follows that the right hand half of the array represents the inverse A^{-1}.
Hence

$$A^{-1} = \begin{pmatrix} -3 & 2 & 1 \\ -2 & 4/3 & 1 \\ 3 & -5/3 & -1 \end{pmatrix}.$$

The Transpose Matrix

The transpose of an $n \times m$ matrix $A = (a_{ij})_{\substack{i=1,\dots,n \\ j=1,\dots,m}}$ is the $m \times n$ matrix

$$A^T = (a_{ji})_{\substack{i=1,\dots,n \\ j=1,\dots,m}} .$$

A matrix M is symmetric if

$$M = M^T.$$

Example

$$\begin{pmatrix} 2 & 5 & 6 \\ 1 & 3 & 4 \end{pmatrix}^T = \begin{pmatrix} 2 & 1 \\ 5 & 3 \\ 6 & 4 \end{pmatrix}$$

(a) Only square matrices can be symmetric

(b) The transpose of a column vector is a row vector, and vice-versa.

(c) Taking the transpose of a matrix twice does nothing. i.e. $\left(A^T\right)^T = A$

Let M,N be matrices such that MN makes sense. Then

$$(MN)^T = N^T M^T.$$