2 Systems of Linear Equations

2.1 Introduction

This section is all about methods for solving systems of linear equations. A system of linear equations is a set of equations in the same variables. For example:

$$\begin{aligned} -x+y+2z &= 2 \\ 3x-y+z &= 6 \\ -x+3y+4z &= 4 \end{aligned}$$

This system has three equations and three unknowns, but in general this could be different. For example:

$$\begin{aligned} w-x+y+2z &= 2 \\ w+3x-y+z &= 6 \end{aligned}$$

In general a system of m linear equations in n unknowns will have the form:

$$\begin{aligned} a_{11}x_1+a_{12}x_2+a_{13}x_3+\ldots+a_{1n}x_n &= b_1 \\ a_{21}x_1+a_{22}x_2+a_{23}x_3+\ldots+a_{2n}x_n &= b_2 \\ \vdots &= \vdots \\ a_{m1}x_1+a_{m2}x_2+a_{m3}x_3+\ldots+a_{mn}x_n &= b_m \end{aligned}$$

Definition 2.1.1 *Given a system of m linear equations in n unknowns we can write this in matrix form as follows:*

$$\mathsf{AX}=\mathsf{B}$$

where $\mathsf{X}=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ *and* $\mathsf{B}=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ x_m \end{pmatrix}$ *are column matrices, and*

$\mathsf{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$ *is an $m \times n$ matrix.*

We can also use an ***Augmented Matrix*** *to represent the system of linear equations:*

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right)$$

Example 2.1.2.

$$\begin{aligned} w - x + y + 2z &= 2 \\ w + 3x - y + z &= 6 \end{aligned}$$

Could be written as

The Augmented matrix would be:

Remark 2.1.3 *You should have seen some matrix multiplication already (e.g. see Problem sheet 0 on blackboard). Notice that matrix multiplication is defined precisely so that the above equation works out.*

2.2 Matrix Algebra

We will very briefly go over Matrix algebra. You should make sure you go over the exercises on Problem sheet 0. For the moment we will mostly assume that the matrices take their values in $\mathbb{R}$ (at the end of this section we will see that we could have chosen to take values from any *Field F*).

If we want to add two matrices, they must have the **same size and shape (the same order)**. Then we can simply **add corresponding elements**. Formally:

Definition 2.2.1. Given $m \times n$ matrices, $\mathsf{A} = [a_{ij}]_{m\times n}$ and if $\mathsf{B} = [b_{ij}]_{m\times n}$, then the **(matrix) sum of** A **and** B is the $m \times n$ matrix $\mathsf{C} = [c_{ij}]_{m\times n}$ where $\boxed{}$

We can also multiply by a scalar product (any element of the field - here $\mathbb{R}$):

Definition 2.2.2. Let $\mathsf{A} = [a_{ij}]$ be any matrix, and let $\lambda \in \mathbb{R}$. Then the **scalar multiple of** A by λ, denoted by $\lambda\mathsf{A}$, is obtained by multiplying every element of A by λ. Thus if $\mathsf{A} = [a_{ij}]_{m\times n}$ then

$\boxed{\lambda\mathsf{A} = }$

See the handout sheet for properties of matrix addition and scalar multiplication.

We can also multiply two matrices together.

Definition 2.2.3. Let $\mathsf{A} = (a_{ij})_{p \times q}$ and $\mathsf{B} = (b_{ij})_{q \times r}$. Then the **matrix product of** A **and** B, denoted by AB, is the matrix C, where

$$\mathsf{C} = (c_{ij})_{p \times r}, \qquad \text{where } \boxed{c_{ij} = \qquad\qquad}$$

Hopefully you will have done lots of examples of this already. Let's have look at some properties of matrix multiplication.

Theorem 2.2.4. Matrix multiplication is associative. That is
Let $\mathsf{A}, \mathsf{B}, \mathsf{C}$ be matrices, and $\alpha \in \mathbb{R}$, then $(\mathsf{AB})\mathsf{C} = \mathsf{A}(\mathsf{BC})$.

Proof

Example 2.2.5. Matrix multiplication is not commutative (i.e. $AB \neq BA$)

Proof: To show this we just need one counterexample. Lets try to make it as simple as possible.

Exercise 2.2.6. Let A, B be matrices with entries in $\mathbb{R}$. Show $\lambda AB = A(\lambda B)$.

Proof

2.3 Row Operations

Recall the definition of an Augmented Matrix from the first lecture. Here's an example to help.

Exercise 2.3.1. Find the Augmented matrix for the following system of linear equations

$$\begin{aligned} -x + y + 2z &= 2 \\ 3x - y + z &= 6 \\ -x + 3y + 4z &= 4 \end{aligned}$$

$$\begin{pmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \\ 4 \end{pmatrix}$$

$$\left(\begin{array}{ccc|c} -1 & 1 & 2 & 2 \\ 3 & -1 & 1 & 6 \\ -1 & 3 & 4 & 4 \end{array}\right)$$

From School you know how to solve systems of linear equations. There are 3 operations you can do:

- multiply an equation by a non-zero factor.
- Add a multiple of one equation to another
- Swap equations around.

In the augmented matrix format we can do these operations more efficiently.

Definition 2.3.2. Elementary row operations (e.r.o's) are performed on an augmented matrix. There are three allowable operations:

- **Multiply** a row by any (non-zero) number
- **Add to any row a multiple of another row**
- **Interchange** two rows

Note that the elementary row operations amount to the actions we could take on the original equations.

Remark 2.3.3 1. *Performing row operations preserves the solutions of a linear system.*

2. *Each row operation has an inverse row operation.*

Example 2.3.4.

$$\begin{aligned} 3x - 2y + z &= -6 \qquad (1) \\ 4x + 6y - 3z &= 5 \qquad (2) \\ -4x + 4y &= 12 \qquad (3) \end{aligned}$$

First multiply (3) by $\frac{1}{4}$:

$$-x + y = 3 \qquad (4)$$

Then add $3 \times (4)$ to (1) and $4 \times (4)$ to (2)

$$\begin{aligned} y + z &= 3 \qquad (5) \\ 10y - 3z &= 17 \qquad (6) \end{aligned}$$

Then take $10 \times (5)$ from (6)

$$-13z = -13 \qquad (7)$$

So $z = 1$. Plug this into (5):

$$y + 1 = 3$$

So $y = 2$. Plug this into (4):

$$-x + 2 = 3$$

So $x = -1$

Definition 2.3.5. Two systems of linear equations are **equivalent** if either:

-
-

Remark 2.3.6 *Equivalently, by Remark 2.3.3 two systems of linear equations are equivalent if and only if they have the same set of solutions.*

If a row consists of mainly 0s and 1s it becomes easier to read off the solutions to the equations. For example:

Example 2.3.7. If we are working in unknowns x, y, z:

$$\left(\begin{array}{ccc|c} -2 & 1 & 2 & 2 \\ 3 & -3 & 1 & 5 \end{array}\right)$$

Whereas

$$\left(\begin{array}{ccc|c} 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 5 \end{array}\right)$$

Definition 2.3.8. We say a matrix is in **echelon form (ef)** if must satisify the following:

- All of the zero rows are at the bottom.
- The first non-zero entry in each row is 1.
- The first non-zero entry in row i is strictly to the left of the first non-zero entry in row $i + 1$.

We say a matrix is in **row reduced echelon form (rref)** if it is in echelon form and:

- The first non-zero entry in row i appears in column j, then every other element in column j is zero.

Example 2.3.9.

2.4 Elementary matrices

Elementary row operations can be carried out using matrix multiplication.

Definition 2.4.1. Any matrix that can be obtained from an identity matrix by means of one elementary row operation is an **elementary matrix**.

There are three types of elementary matrix:

- The general form of the elementary matrix which multiplies a row by any (non-zero) number, α is of the form

  $$\mathsf{E_r}(\alpha) =$$

  where all elements on row r is multiplied by α.

- The general form of the elementary matrix which adds a multiple of a row by any non-zero number α to another is of the form

  $$\mathsf{E_{rs}}(\alpha) =$$

  where all elements of s are multiplied by α and added to row r.

- The general form of the elementary matrix which interchanges two rows is of the form

  $$\mathsf{E_{rs}} =$$

  where r and s are the rows to interchange.

Example 2.4.2. Find the string of elementary matrices that correspond to the following row operations:

$$\left(\begin{array}{ccc|c} 0 & 1 & 1 & 3 \\ 0 & 0 & -13 & -13 \\ -1 & 1 & 0 & 3 \end{array}\right) \xrightarrow{R_2 \mapsto -\frac{1}{13}R_2}$$
$$\left(\begin{array}{ccc|c} 0 & 1 & 1 & 3 \\ 0 & 0 & 1 & 1 \\ -1 & 1 & 0 & 3 \end{array}\right) \xrightarrow{R_1 \mapsto R_1 - R_2}$$
$$\left(\begin{array}{ccc|c} 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ -1 & 1 & 0 & 3 \end{array}\right) \xrightarrow{R_3 \mapsto -R_3 + R_1}$$
$$\left(\begin{array}{ccc|c} 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & -1 \end{array}\right) \xrightarrow{R_1 \mapsto R_2, R_2 \mapsto R_3, R_3 \mapsto R_1}$$
$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{array}\right)$$

Exercise 2.4.3. Find the string of elementary matrices corresponding to the following row operations.

$$\left(\begin{array}{ccc|c} 3 & -2 & 1 & -6 \\ 4 & 6 & -3 & 5 \\ -4 & 4 & 0 & 12 \end{array}\right) \xrightarrow{R_3 \mapsto \frac{1}{4}R_3}$$
$$\left(\begin{array}{ccc|c} 3 & -2 & 1 & -6 \\ 4 & 6 & -3 & 5 \\ -1 & 1 & 0 & 3 \end{array}\right) \xrightarrow[R_1 \mapsto R_1 + 3R_3]{R_2 \mapsto R_2 - 4R_3}$$
$$\left(\begin{array}{ccc|c} 0 & 1 & 1 & 3 \\ 0 & 10 & -3 & 17 \\ -1 & 1 & 0 & 3 \end{array}\right) \xrightarrow{R_2 \mapsto R_2 - 10R_1}$$

*****Mentimeter***************

Theorem 2.4.4. Let A be a $m \times n$ matrix and let E be an elementary $m \times m$ matrix. The matrix multiplication EA applies the same elementary row operation on A that was performed on the identity matrix to obtain E.

Proof: exercise.

2.5 More matrices

Definition 2.5.1. We say a matrix is square if it has the same number of rows as it does columns (i.e. its a member of $M_{n\times n}(F)$ for some field F).

Definition 2.5.2.

A square matrix $\mathsf{A} = a_{ij} \in M_{n\times n}(F)$ is said to be:

1. **upper triangular** if $a_{ij} = 0$ wherever $\boxed{\quad\quad}$. A has zeros for all its elements below the diagonal.
2. **lower triangular** if $a_{ij} = 0$ wherever $\boxed{\quad\quad}$. A has zeros for all its elements above the diagonal.
3. **diagonal** if $a_{ij} = 0$ wherever $\boxed{\quad\quad}$. That is to say A has zeros for all its elements except those on the main diagonal.

Example 2.5.3.

$$\underset{\text{Upper triangular}}{\begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 7 \\ 0 & 0 & 1 \end{pmatrix}} \quad \underset{\text{Lower triangular}}{\begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}} \quad \underset{\text{diagonal}}{\begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix}}$$

Definition 2.5.4. The $n \times n$ **identity matrix** is denoted by $\mathsf{I_n}$. An identity matrix has all of its diagonal entreis equal to 1 and all other entries equal to 0. It is called the identity matrix because it is the multiplicative identity matrix for $n \times n$ matrices, i.e.

For $A \in M_{n\times n}(\mathbb{R})$, $\mathsf{I_n}A = A\mathsf{I_n} = A$

Definition 2.5.5. If, for a square matrix B, if there exists another square matrix B^{-1} such that $\mathsf{B}\mathsf{B}^{-1} = \mathsf{I} = \mathsf{B}^{-1}\mathsf{B}$, then we say that B is invertible, and B^{-1} **is an inverse of** B.

It is important to realise that **the matrix** B **might not have an inverse:** B^{-1} **might not exist**.

Definition 2.5.6. A matrix without an inverse is called a **singular** matrix.

Example 2.5.7. Let $\mathsf{A} = \begin{pmatrix} 2 & 0 \\ 1 & -1 \end{pmatrix}$, verify that it has an inverse: $\mathsf{B} = \begin{pmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & -1 \end{pmatrix}$.

Theorem 2.5.8. The inverse of a given matrix is unique. If there exist square matrices A, B, C such that AB = I = CA, then B = C.

This theorem shows that if a matrix A is invertible, we can talk about the inverse of A, denoted by A^{-1}. In some circumstances, we can say that a matrix is invertible, and we can find an expression for its inverse, without knowing exactly what the matrix is.

Exercise 2.5.9. Suppose $A, B \in M_{n\times n}(\mathbb{R})$ are both invertible. Show that AB is invertible by finding its inverse.

Definition 2.5.10. If $\mathsf{A} = [a_{ij}]_{m\times n}$, then the **Transpose of** A is $\mathsf{A}^\mathsf{T} = [a_{ji}]_{n\times m}$.

Example 2.5.11. If

$$\mathsf{A} = \begin{pmatrix} 1 & 0 & 5 \\ 4 & 2 & 1 \end{pmatrix}, \qquad \text{then} \qquad \mathsf{A}^\mathsf{T} =$$

and we can see that the transpose of a 2×3 matrix must be a 3×2 matrix.

Exercise 2.5.12. Let $A \in M_{n\times m}(\mathbb{R})$, $B \in M_{m\times p}(\mathbb{R})$, $(AB)^T = B^T A^T$.

Proof:

First remark that $B^T A^T$ is defined and has order $p \times n$, not also AB has order $n \times p$ so $(AB)^T$ has order $p \times n$.

Let $\mathsf{A} = (a_{ij})$ and $\mathsf{B} = (b_{ij})$

- The ij^{th} entry of AB is $\sum_{k=1}^m a_{ik}b_{kj}$. This is the ji^{th} entry of $(AB)^T$
- The ji^{th} entry of $B^T A^T$ is $\sum_{k=1}^m (b^T)_{jk}(a^T)_{ki} = \sum_{k=1}^m (b)_{kj}(a)_{ik} = \sum_{k=1}^m a_{ik}b_{kj}$

Theorem 2.5.13. Given an invertible square matrix A, then A^T is also invertible, and $(\mathsf{A}^\mathsf{T})^{-1} = (\mathsf{A}^{-1})^\mathsf{T}$.

2.6 Inverses using row operations

We can use Elementary matrices to find inverses of matrices (it they exist).

Theorem 2.6.1. Every elementary matrix is invertible and the inverse is also an elementary matrix.

Proof

Theorem 2.6.2. If the square matrix A can be row-reduced to an identity matrix by a sequence of elementary row operations, then A is invertible and the inverse of A is found by applying the same sequence of elementary row operations to I.

Proof

Example 2.6.3. Let $\mathsf{A} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & 0 \\ 3 & 0 & 4 \end{pmatrix}$ find A^{-1}.

The method consists of writing the identity matrix I to the right of our given matrix, and then using the same elementary row operations on both matrices to turn the left-hand matrix into I. When this has been achieved, the right-hand matrix will have been transformed into the inverse matrix, A^{-1}.

First, we construct the augmented matrix $\mathsf{A}|\mathsf{I}$, by writing the identity matrix to the right of the matrix A,

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 & 1 & 0 \\ 3 & 0 & 4 & 0 & 0 & 1 \end{array}\right)$$

After our row operations, this matrix will be transformed into $\mathsf{I}|\mathsf{A}^{-1}$.