Chapter 1. Basic concepts

Lecture notes for MA1111

P. Karageorgis

pete@maths.tcd.ie

Matrices and vectors

- A matrix is a rectangular array of numbers arranged in rows and columns. A typical example is

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

- An $m \times n$ matrix is one that has m rows and n columns. Rows are always mentioned first by convention. The (i, j)th entry appears in the ith row and jth column. In this case, $A_{12} = 3$ and $A_{23} = 5$.
- A vector is a matrix that has only one column such as

$$\boldsymbol{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

- This vector has entries $x_1 = 2$ and $x_2 = 1$. Vectors can be used to describe direction, so they are sometimes depicted as arrows.

Addition and scalar multiplication

- To add two matrices or vectors of the same size, one adds their corresponding entries. As a typical example, one has

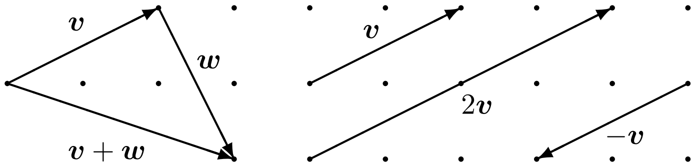
$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 0 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 1+0 & 2+2 \\ 3+1 & 4+3 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 4 & 7 \end{bmatrix}.$$

- To multiply a matrix/vector by a scalar, one multiplies each entry of the matrix/vector by that scalar. For instance,

$$A = \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix} \quad \Longrightarrow \quad 2A = \begin{bmatrix} 4 & 0 \\ 2 & 6 \end{bmatrix}.$$

- Addition and scalar multiplication of matrices satisfy the usual rules of arithmetic (commutative law, associative law and so on).
- One usually denotes by 0 the zero matrix whose entries are all zero.

Geometric interpretation of vectors



- The vectors depicted above are the vectors

$$v = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \qquad w = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \qquad v + w = \begin{bmatrix} 3 \\ -1 \end{bmatrix}, \qquad 2v = \begin{bmatrix} 4 \\ 2 \end{bmatrix}.$$

- Note that v, w and $v + w$ form the three sides of a triangle, while the scalar multiples of v are all vectors which are parallel to v.

Length and dot product

Definition 1.1 – Length and dot product

The length of a vector in $\mathbb{R}^n$ is defined by

$$||\boldsymbol{x}|| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}.$$

The dot product of two vectors in $\mathbb{R}^n$ is defined by

$$\boldsymbol{x} \cdot \boldsymbol{y} = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n.$$

- One has $\boldsymbol{x} \cdot \boldsymbol{x} = ||\boldsymbol{x}||^2$ and also $\boldsymbol{x} \cdot (\boldsymbol{y} + \boldsymbol{z}) = \boldsymbol{x} \cdot \boldsymbol{y} + \boldsymbol{x} \cdot \boldsymbol{z}$.
- Letting θ be the angle between $\boldsymbol{x}$ and $\boldsymbol{y}$, we have

$$\boldsymbol{x} \cdot \boldsymbol{y} = ||\boldsymbol{x}|| \cdot ||\boldsymbol{y}|| \cdot \cos\theta.$$

- Thus, two vectors $\boldsymbol{x}, \boldsymbol{y}$ are perpendicular if and only if $\boldsymbol{x} \cdot \boldsymbol{y} = 0$.

Cross product

Definition 1.2 – Cross product

The cross product of two vectors in $\mathbb{R}^3$ is the vector

$$\boldsymbol{x} \times \boldsymbol{y} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \times \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{bmatrix}.$$

- The cross product $\boldsymbol{x} \times \boldsymbol{y}$ is perpendicular to both $\boldsymbol{x}$ and $\boldsymbol{y}$.
- Letting θ be the angle between $\boldsymbol{x}$ and $\boldsymbol{y}$, we have

$$||\boldsymbol{x} \times \boldsymbol{y}|| = ||\boldsymbol{x}|| \cdot ||\boldsymbol{y}|| \cdot \sin\theta.$$

- Right hand rule: the cross product $\boldsymbol{x} \times \boldsymbol{y}$ points in the direction of your right thumb as you curl your fingers from $\boldsymbol{x}$ to $\boldsymbol{y}$.

Theorem 1.3 – Equation of a line

Given a point $A(x_0, y_0, z_0)$ and a vector $\boldsymbol{n}$ in $\mathbb{R}^3$, the equations

$$\left\{ \begin{array}{c} x = x_0 + n_1 t \\ y = y_0 + n_2 t \\ z = z_0 + n_3 t \end{array} \right\}$$

describe the line which passes through A in the direction of $\boldsymbol{n}$.

- These are called parametric equations. Here, the parameter t is arbitrary and each value of t gives rise to a point on the line.
- Note that the coefficients of t are the coordinates of $\boldsymbol{n}$, while the constant terms x_0, y_0, z_0 are the coordinates of the point.
- If a line passes through A and B, then its direction is $\overrightarrow{AB}$.

Theorem 1.4 – Equation of a plane

Given a point $A(x_0, y_0, z_0)$ and a vector $\boldsymbol{n}$ in $\mathbb{R}^3$, the equation

$$n_1(x - x_0) + n_2(y - y_0) + n_3(z - z_0) = 0$$

describes the plane which passes through A with normal vector $\boldsymbol{n}$.

- The vector $\boldsymbol{n}$ is perpendicular to every vector that lies on the plane.
- Note that the coefficients of x, y, z are the coordinates of $\boldsymbol{n}$.
- If a plane passes through A, B and C, then its normal vector is

$$\boldsymbol{n} = \overrightarrow{AB} \times \overrightarrow{AC}.$$

Linear equations

Definition 1.5 – Linear equation

An equation involving the unknowns $x_1, x_2, \ldots, x_n$ is linear, if it has the form $c_1x_1 + c_2x_2 + \ldots + c_nx_n = d$ for some constants $c_1, c_2, \ldots, c_n, d$.

- There is a standard procedure for solving any number of linear equations. It is called row reduction or Gauss-Jordan elimination.
- To perform row reduction, we form the associated augmented matrix

$$\left\{\begin{aligned} x - 2y &= 2 \\ x + 3y &= 7 \end{aligned}\right\} \quad \longleftrightarrow \quad \begin{bmatrix} 1 & -2 & 2 \\ 1 & 3 & 7 \end{bmatrix}$$

and then we try to simplify the matrix using row operations.

Definition 1.6 – Elementary row operations

There are three kinds of elementary row operations.

1. We can multiply/divide a row by a nonzero number.
2. We can add a multiple of one row to another row.
3. We can interchange two rows.

- One may use the first operation to ensure that the leftmost nonzero entry in each row is equal to 1. If it is equal to some other number, then we may simply divide by that number to make it equal to 1.
- Once the leftmost nonzero entry in a row is equal to 1, one may use the second operation to make all entries above/below it equal to zero.
- The third operation allows us to rearrange the rows, if needed.

Reduced row echelon form

Definition 1.7 – Reduced row echelon form

A matrix in reduced row echelon form has the following properties.

1. In every nonzero row, the leftmost nonzero entry is equal to 1, it is called a pivot, and every entry above/below each pivot is zero.
2. Pivots further down appear further to the right.
3. Zero rows, if any, appear at the bottom.

- Two typical examples of reduced row echelon forms are

$$\begin{bmatrix} \mathbf{1} & 0 & 3 & 0 & 1 \\ 0 & \mathbf{1} & 2 & 0 & 4 \\ 0 & 0 & 0 & \mathbf{1} & 6 \end{bmatrix}, \qquad \begin{bmatrix} 0 & \mathbf{1} & 1 & 0 & 2 & 1 \\ 0 & 0 & 0 & \mathbf{1} & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Here, each of the pivots is indicated in boldface font.

Case 1 – A system with no solutions

- We use row reduction to solve the system

$$\left\{\begin{aligned} x + 2y + z &= 4 \\ 2x + y + 2z &= 5 \\ 3x + 2y + 3z &= 4 \end{aligned}\right\}.$$

- In this case, row reduction gives

$$\begin{bmatrix} 1 & 2 & 1 & 4 \\ 2 & 1 & 2 & 5 \\ 3 & 2 & 3 & 4 \end{bmatrix} \longrightarrow \begin{bmatrix} \mathbf{1} & 2 & 1 & 4 \\ 0 & -3 & 0 & -3 \\ 0 & -4 & 0 & -8 \end{bmatrix} \longrightarrow \begin{bmatrix} \mathbf{1} & 2 & 1 & 4 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 2 \end{bmatrix}$$

$$\longrightarrow \begin{bmatrix} \mathbf{1} & 0 & 1 & 2 \\ 0 & \mathbf{1} & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

- Since the last row corresponds to the equation $0x + 0y + 0z = 1$, the given system has no solutions. This system actually describes three planes which have no point in common.

Case 2 – A system with infinitely many solutions

- We use row reduction to solve the system

$$\left\{\begin{aligned} x + 2y + z &= 4 \\ 2x + y + 2z &= 5 \\ 3x + 2y + 3z &= 8 \end{aligned}\right\}.$$

- In this case, row reduction gives

$$\begin{bmatrix} 1 & 2 & 1 & 4 \\ 2 & 1 & 2 & 5 \\ 3 & 2 & 3 & 8 \end{bmatrix} \longrightarrow \begin{bmatrix} \mathbf{1} & 2 & 1 & 4 \\ 0 & -3 & 0 & -3 \\ 0 & -4 & 0 & -4 \end{bmatrix} \longrightarrow \begin{bmatrix} \mathbf{1} & 2 & 1 & 4 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

$$\longrightarrow \begin{bmatrix} \mathbf{1} & 0 & 1 & 2 \\ 0 & \mathbf{1} & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

- The corresponding equations are $x + z = 2$ and $y = 1$. Solving for the leftmost variable, we find that $x = 2 - z$, $y = 1$ and z is arbitrary.

Case 3 – A system with a unique solution

- We use row reduction to solve the system

$$\left\{\begin{aligned} x + 2y + z &= 4 \\ 2x + y + 2z &= 5 \\ 3x + 2y + 4z &= 4 \end{aligned}\right\}.$$

- In this case, row reduction gives

$$\begin{bmatrix} 1 & 2 & 1 & 4 \\ 2 & 1 & 2 & 5 \\ 3 & 2 & 4 & 4 \end{bmatrix} \longrightarrow \begin{bmatrix} \mathbf{1} & 2 & 1 & 4 \\ 0 & -3 & 0 & -3 \\ 0 & -4 & 1 & -8 \end{bmatrix} \longrightarrow \begin{bmatrix} \mathbf{1} & 2 & 1 & 4 \\ 0 & 1 & 0 & 1 \\ 0 & -4 & 1 & -8 \end{bmatrix}$$

$$\longrightarrow \begin{bmatrix} \mathbf{1} & 0 & 1 & 2 \\ 0 & \mathbf{1} & 0 & 1 \\ 0 & 0 & 1 & -4 \end{bmatrix} \longrightarrow \begin{bmatrix} \mathbf{1} & 0 & 0 & 6 \\ 0 & \mathbf{1} & 0 & 1 \\ 0 & 0 & \mathbf{1} & -4 \end{bmatrix}.$$

- Thus, there is a unique solution, namely $x = 6$, $y = 1$ and $z = -4$.

Number of solutions

Definition 1.8 – Free and leading variables

Consider the reduced row echelon form of the augmented matrix for a system of linear equations. The variables without pivots are called free and they are arbitrary. The variables with pivots are called leading and they can be expressed in terms of the free variables.

Theorem 1.9 – Linear systems

Consider a system of linear equations and let R be the reduced row echelon form of the associated augmented matrix.

1. If R has a pivot in the last column, the system has no solutions.
2. If R has no pivot in the last column and if there is at least one free variable, the system has infinitely many solutions.
3. If R has no pivot in the last column and if there are no free variables, the system has a unique solution.

Example: A system involving 5 unknowns

- We use row reduction to solve the system

$$\left\{ \begin{aligned} x_1 + 2x_2 + x_3 + 2x_4 + 3x_5 &= 1 \\ x_1 + x_2 \phantom{{}+x_3} + x_4 + 2x_5 &= 2 \\ 2x_1 + 3x_2 + x_3 + 2x_4 + 2x_5 &= 6 \end{aligned} \right\}.$$

- As one can easily check, the reduced row echelon form is

$$\begin{bmatrix} 1 & 2 & 1 & 2 & 3 & 1 \\ 1 & 1 & 0 & 1 & 2 & 2 \\ 2 & 3 & 1 & 2 & 2 & 6 \end{bmatrix} \longrightarrow \begin{bmatrix} \mathbf{1} & 0 & -1 & 0 & 1 & 3 \\ 0 & \mathbf{1} & 1 & 0 & -2 & 2 \\ 0 & 0 & 0 & \mathbf{1} & 3 & -3 \end{bmatrix}.$$

- The leading variables are x_1, x_2 and x_4. These can all be expressed in terms of the free variables x_3 and x_5. In fact, one has

$$\left\{ \begin{aligned} x_1 &= 3 + x_3 - x_5 \\ x_2 &= 2 - x_3 + 2x_5 \\ x_4 &= -3 - 3x_5 \end{aligned} \right\}.$$

Example: A system involving a parameter

- Suppose α is some fixed parameter and consider the system

$$\left\{ \begin{aligned} x + \alpha y &= 1 \\ x + y &= 2 \end{aligned} \right\}.$$

- In this case, the method of row reduction gives

$$\begin{bmatrix} 1 & \alpha & 1 \\ 1 & 1 & 2 \end{bmatrix} \longrightarrow \begin{bmatrix} \mathbf{1} & \alpha & 1 \\ 0 & 1-\alpha & 1 \end{bmatrix}$$

  and one needs to consider cases in order to proceed.

- If it happens that $\alpha = 1$, then the second row gives rise to a contradiction, so the system has no solutions. If $\alpha \neq 1$, we get

$$\begin{bmatrix} \mathbf{1} & \alpha & 1 \\ 0 & 1 & 1/(1-\alpha) \end{bmatrix} \longrightarrow \begin{bmatrix} \mathbf{1} & 0 & 1-\alpha/(1-\alpha) \\ 0 & \mathbf{1} & 1/(1-\alpha) \end{bmatrix},$$

  so the system has a unique solution.

Matrix multiplication: examples

- The product of a row vector with a column vector is defined by

$$\begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 y_1 + x_2 y_2 + \ldots + x_n y_n \end{bmatrix}.$$

- To compute the product AB of two matrices, one multiplies each of the rows of A with each of the columns of B. For instance,

$$\left[\begin{array}{cc} 2 & 3 \\ \hline 1 & 2 \end{array}\right] \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2+6 \\ 1+4 \end{bmatrix} = \begin{bmatrix} 8 \\ 5 \end{bmatrix}$$

and one may similarly compute

$$\left[\begin{array}{cc} 3 & 1 \\ \hline 1 & 2 \end{array}\right] \left[\begin{array}{c|c|c} 1 & 2 & 1 \\ 1 & 3 & 2 \end{array}\right] = \left[\begin{array}{c|c|c} 4 & 9 & 5 \\ \hline 3 & 8 & 5 \end{array}\right].$$

Matrix multiplication: warnings

The commutative law $AB = BA$ is not generally true for matrices.

- As a typical example, we note that

$$AB = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$
$$BA = \begin{bmatrix} 2 & 3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}.$$

The binomial theorem is not generally true for matrices.

- In fact, one needs the commutative law to justify the computation

$$\begin{aligned}(A+B)^2 = (A+B)(A+B) &= A^2 + AB + BA + B^2 \\ &= A^2 + 2AB + B^2.\end{aligned}$$

It is not true that $AB = 0$ implies $A = 0$ or $B = 0$.

Matrix multiplication: properties

Definition 1.10 – Matrix multiplication

Suppose A is an $m \times n$ matrix and B is an $n \times p$ matrix. Then their product AB is the $m \times p$ matrix whose (i,j)th entry is given by

$$(AB)_{ij} = (i\text{th row of } A) \cdot (j\text{th column of } B) = \sum_{k=1}^{n} A_{ik}B_{kj}.$$

Theorem 1.11 – Properties of matrix multiplication

The following properties hold whenever the products are defined.

- Associative law: $(AB)C = A(BC)$.
- Left distributive law: $A(B+C) = AB + AC$.
- Right distributive law: $(A+B)C = AC + BC$.
- Identity matrix: If I_n is $n \times n$ with diagonal entries equal to 1 and all other entries equal to 0, then $AI_n = A$ and $I_nB = B$.

Block multiplication

- Matrix multiplication of 2×2 matrices gives

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x & y \\ z & w \end{bmatrix} = \begin{bmatrix} ax+bz & ay+bw \\ cx+dz & cy+dw \end{bmatrix}.$$

- We shall sometimes need the analogous formula

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} X & Y \\ Z & W \end{bmatrix} = \begin{bmatrix} AX+BZ & AY+BW \\ CX+DZ & CY+DW \end{bmatrix}$$

  in the case that A, B, C, D and X, Y, Z, W are matrices.

- This formula is especially useful when a large matrix can be decomposed into smaller, simpler blocks. For instance,

$$\begin{bmatrix} I_m & A \\ 0 & I_n \end{bmatrix} \begin{bmatrix} B & C \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} B & C \\ 0 & 0 \end{bmatrix}.$$

Linear systems revisited

Theorem 1.12 – Linear systems

Every system of linear equations can be expressed in the form

$$A\boldsymbol{x} = \boldsymbol{y},$$

where the matrix A contains the coefficients of all variables, $\boldsymbol{x}$ is the vector of unknowns and $\boldsymbol{y}$ contains the constant coefficients.

- As we already know, one can solve such a system using row reduction of the associated augmented matrix $[A\ \ \boldsymbol{y}]$.
- Matrix multiplication can be used to develop an alternative approach for solving systems of linear equations.
- The new approach is only valid for square $n \times n$ matrices A, but it has its own advantages. It will also help us understand row reduction.