MT251P – Lecture 1

Fiacre Ó Cairbre

Chapter 1 – Problem solving, Proof and Propositional Logic.

Section 1.1 – Problem Solving.

Mathematics has many features. The first feature we will discuss is problem solving.

1. Try to understand the problem first.

2. It may be helpful to draw a diagram.

3. You should never say that you cannot solve a problem after just trying for a few minutes.

4. Following on from 3 above, it's good to have more than one approach to a problem (if that one doesn't work).

5. The more problems you work on, the better you will (hopefully) become at problem solving.

6. If one approach doesn't work in a particular problem, you might still learn something from it.

Example 1.

Problem: – Consider a sports tournament (like tennis) where each game between two players produces one winner and one loser. The winner progresses to the next round and the loser drops out of the tournament. Eventually, there is a final between two players and the winner is called the champion. Suppose there are 91 players at the start of the tournament. How many games have been played in the tournament when the champion lifts the trophy?

Solution: –

What are we looking for? We are looking for a number (i.e. the number of games played). Now, we will show that there is something else (different from games) that has the same total number as number of games. Well, when the champion lifts the trophy at the end of the tournament, we know that the number of losers is exactly the same as the number of games played. The number of losers is 90 and so the number of games played is 90 and the problem is solved.

Remark 1.

Notice how the answer in example 1 above is independent of the structure of the tournament which might seem a little surprising. Also, example 1 can be generalised to n players (instead of 91). Try and convince yourself that the answer then is $n - 1$.

Example 2.

Problem: – Suppose there are twenty statements on a page and for $1 \leq k \leq 20$, the k^{th} statement reads

— There are exactly k false statements on this page. —

Which of the k statements are true and which are false?

Solution. – The first thing to notice is that at most one of the k statements can be true because we cannot have exactly t false statements and exactly w false statements where $t \neq w$. So, we have either exactly one true statement or no true statements.

Now, it's impossible to have no true statements because then all twenty statements are false which means that the 20^{th} statement is true which is a contradiction.

So, there is exactly one true statement which means exactly nineteen statements are false which means that the 19^{th} statement is true.

So, the answer is that the 19^{th} statement is true and all the other statements are false.

Example 3.

Problem: – A perfect number is defined as a positive integer that is the sum of all its positive divisors excluding itself. So, for example 6 is a perfect number because 1, 2, 3, 6 are the positive divisors of 6 and $6 = 1 + 2 + 3$. Find the next perfect number after 6. The classical Greeks (around 600 BC) were the first to consider perfect numbers and they called them perfect because they really liked them.

Solution: – We start checking numbers after 6. Notice that we can skip over prime numbers because the only positive divisors of a prime number p are 1 and p. After a little while we see that 28 is the next perfect number because its positive divisors are 1, 2, 4, 7, 14, 28 and $28 = 1 + 2 + 4 + 7 + 14$. I won't ask you to find the next perfect number after 28 and you will see why later.

Example 4 – First to 100 wins.

Problem: – Consider the following game between two players, A and B. Let $S = \{1, 2, 3, \ldots, 10\}$. A chooses any number from S and announces it. B chooses any number from S and adds it to A's number and announces the result. For, example, if A announced 6 and B chose 3, then B announces 9. A then chooses any number from S and adds it to the previous announcement and announces the result. For example, continuing on from above, if A chose 8, then A announces 17 ($=9 + 8$). Continue like this and the first player to announce 100 wins. Notice that during the game A and B can choose any number from S and it doesn't matter if the number was chosen before by A or B. Find a strategy that guarantees you win if you go first.

Solution: – Many people approach this problem by starting at the start of the game. However, it's more beneficial to start at the end of the game as follows. If you can announce the number 89, then you are guaranteed to win because you can announce 100 after your opponent's announcement. In a similar way, working back, we see that if you can announce the number 78, then you are guaranteed that you can announce 89 and so you are guaranteed to win. Continue working back like this and we see that the following strategy will guarantee a win:

You start by announcing 1 and then you announce 12 (no matter what your opponent announces), and then you announce 23, 34, 45, 56, 67, 78, 89, 100 to guarantee a win.

Remark 2.

There are still many unsolved problems in mathematics. Some of these problems have been around for thousands of years. Here is one example of an unsolved problem:

– Is there an odd perfect number? –

MT251P – Lecture 2

Fiacre Ó Cairbre

Section 1.2 – Proof.

The starting point for mathematics is a collection of accepted statements (also called axioms). One is not required to prove these axioms because they are accepted. How do we move on to obtain new mathematical statements? This is where deductive reasoning comes in. Deductive reasoning works in the following way:

We start with premises (which are accepted statements) and we then derive conclusions with certainty.

The last word certainty in the above makes deductive reasoning a very special type of reasoning.

Here is an example of deductive reasoning: Suppose my premises are that

All students in the class have brown hair. *(*)*

and

Brian is a student in the class. *(**)*

Then, I can conclude with certainty that

Brian has brown hair.

Mathematical proof.

The classical Greeks decided (around 600 BC) that in order to prove something in mathematics one needs to use deductive reasoning to make the conclusion. This is still the requirement for a mathematical proof today and is what distinguishes a mathematical proof from other forms of 'proof'.

Example 5.

Prove that the square of any odd integer is an odd integer.

Solution.

Here, the set of integers is the usual set $\{\ldots, -2, -1, 0, 1, 2, 3, \ldots\}$. Suppose n is an odd integer. Then we have that

$$n = 2k + 1, \quad \text{for some integer} \quad k$$

Thus

$$n^2 = (2k + 1)^2$$

$$= 4k^2 + 4k + 1$$

$$= 2(2k^2 + 2k) + 1$$

and so n^2 is odd and we are done.

Example 6.

Using a similar approach as in example 5, try to prove that the square of an even integer is an even integer.

Remark 3. The converse of a statement.

Suppose P is the statement that n is an odd integer and Q is the statement that n^2 is an odd integer. Then, in example 5 above we proved a statement of the form

$$P \Rightarrow Q$$

which means P implies Q (or in other words – if P is true, then Q is true). Suppose we want to prove the statement

$$Q \Rightarrow P$$

i.e. if n^2 is an odd integer, then n is an odd integer. This statement $Q \Rightarrow P$ is called the converse of $P \Rightarrow Q$. We will now prove $Q \Rightarrow P$ in the following example:

Example 7.

If n^2 is an odd integer, then prove that n is an odd integer.

Solution.

If n^2 is an odd integer, then n has to be odd because otherwise n is even and so (by example 6) n^2 is even which is impossible.

Remark 4 – Equivalent statements.

Using the P and Q from remark 3 we have now proved that

$$P \Rightarrow Q \qquad \text{and} \qquad Q \Rightarrow P \qquad (*)$$

The combination of the two statements in (*) above can be more concisely written as

$$P \iff Q$$

and so

$$n \;\text{ is odd } \iff n^2 \;\text{ is odd}$$

and we say

$$n \text{ is odd if and only if } n^2 \text{ is odd}$$

We say that the statements P and Q are equivalent when $P \iff Q$

Remark 5 – Counterexamples.

Counterexamples can be used to show that a certain statment is not valid. For example, -2 is a counterexample to the statement that $n^2 \geq 0$ implies $n \geq 0$.

Proof by contradiction.

Proof by contradiction works as follows:

Suppose we want to prove that the statement P is true. Well, we try to prove this by actually trying to prove that P is not false. How do we try to prove P is not false? Well, we suppose that P is false and try to show it leads to a contradiction and this will mean that the statement – P is false – cannot hold. Hence P must be true and we are done.

Example 8 – Prove that $\sqrt{2}$ is irrational.

Proof.

Now, we want to prove that $\sqrt{2}$ is irrational and so (using proof by contradiction) we will start by supposing $\sqrt{2}$ is not irrational (i.e. $\sqrt{2}$ is rational).

So, there are integers a, b with $b \neq 0$ such that

$$\sqrt{2} = \frac{a}{b}$$

Now, by dividing out any relevant common divisors of a and b, we can assume that the greatest common divisor of a and b is 1.

$$\sqrt{2} = \frac{a}{b}$$

$$\Rightarrow 2 = \frac{a^2}{b^2}$$

$$\Rightarrow a^2 = 2b^2$$

$$\Rightarrow a^2 \text{ is even}$$

$$\Rightarrow a \text{ is even}$$

$$\Rightarrow a = 2c, \quad \text{for some integer } c$$

$$\Rightarrow 4c^2 = 2b^2$$

$$\Rightarrow 2c^2 = b^2$$

$$\Rightarrow b^2 \text{ is even}$$

$$\Rightarrow b \text{ is even}$$

$$\Rightarrow 2 \text{ is a common divisor of } a \text{ and } b$$

which is a contradiction because the greatest common divisor of a and b is 1.

So, our initial assumption that $\sqrt{2}$ is rational is false. Hence $\sqrt{2}$ is irrational and we are done.

MT251P – Lecture 3

Fiacre Ó Cairbre

Section 1.3 – Propositional Logic.

Elementary and Compound Propositions.

A proposition is a statement that is either true or false, but not both.

Example 9.

Consider the following three statements:

P1: 3 is a prime number.

P2: $2 + 2 = 5$

P3: 81 is divisble by 3

The above three statements are all propositions. However the statement – Where are you going? – is not a propostion.

Remark 6.

A proposition takes exactly one of the values, true or false. Typically, we denote true by T or 1 and we denote false by F or 0. For example, P1 and P3 above take the value T or 1 and P2 above takes the value F or 0.

Our three propositions above are called elementary propositions because they cannot be broken down into simpler propositions. However, consider the proposition:

P4: $6 > 5$ and 110 is even.

Then, P4 is a compound proposition because it's constructed using the connective 'and' from the two simpler propositions, P5, P6 where:

P5: $6 > 5$

P6: 110 is even.

Remark 7.

We use the following three fundamental operations to construct compound propositions from elementary propositions

1. Negation (NOT).

The negation of a proposition P is denoted by $\neg P$ and is formed by negating the statement in P. So, $\neg P$ takes the value T $\iff$ P takes the value F. This will then mean that $\neg P$ takes the value F $\iff$ P takes the value T.

2. Conjunction (AND).

The conjunction of two propositions P and Q is denoted by $P \wedge Q$ and takes the value T $\iff$ P and Q both take the value T.

3. Disjunction (OR).

The disjunction of two propositions P and Q is denoted by $P \vee Q$ and takes the value T $\iff$ P or Q (or both) take the value T.

Truth Tables.

The expressions $\neg P$, $P \wedge Q$ and $P \vee Q$ are examples of formulas in propositional algebra. Formulas are built from variables (like P, Q etc.), logical operations (like $\neg$, $\wedge$, $\vee$ etc.) and the values T, F. Formulas define propositional functions (for example, $\neg P$ has one variable, $P \wedge Q$ has two variables etc). When we give a value to each variable in a formula, then the value of the formula is determined.

A truth table tells us what value the formula takes for all combinations of the variables. In a truth table, the T, F values in a column give us the value of the formula that appears on the top of the column.

Example 10.

The truth table for negation is

P	$\neg P$
T	F
F	T

Example 11.

The truth table for disjunction is

P	Q	$P \vee Q$
T	T	T
T	F	T
F	T	T
F	F	F

Example 12.

The truth table for conjunction is

P	Q	$P \wedge Q$
T	T	T
T	F	F
F	T	F
F	F	F

Example 13.

Construct the truth table for $\neg(P \wedge (\neg Q))$.

P	Q	$\neg Q$	$P \wedge (\neg Q)$	$\neg(P \wedge (\neg Q))$
T	T	F	F	T
T	F	T	T	F
F	T	F	F	T
F	F	T	F	T

Remark 8.

If formulas G, H have identical truth tables (meaning that we have identical columns under each of the two formulas at the right end of each table), then G and H are called equivalent. This means that G and H define the same propositional function and we write $G \sim H$ to denote that G is equivalent to H.

MT251P – Lecture 4

Fiacre Ó Cairbre

Implication and Conditional Statements.

Conditional statements of the form 'If P, then Q' are constructed using the implication operation $\Rightarrow$

Note that P implies Q (i.e. $P \Rightarrow Q$) is false if and only if P is true and Q is false.

The truth table for $P \Rightarrow Q$ is

P	Q	$P \Rightarrow Q$
T	T	T
T	F	F
F	T	T
F	F	T

Check that $P \Rightarrow Q$ is equivalent to $(\neg P) \vee Q$.

Remark 9.

The 'if and only if' statement is denoted by $P \iff Q$ and can now be defined as

$$P \iff Q = (P \Rightarrow Q) \wedge (Q \Rightarrow P)$$

Remark 10.

The operation $\neg$ has precedence over $\wedge$ which has precedence over $\vee$. For example

$$\neg P \wedge Q \text{ means } (\neg P) \wedge Q \text{ and doesn't mean } \neg(P \wedge Q)$$

Check that $(\neg P) \wedge Q$ is not equivalent to $\neg(P \wedge Q)$ by showing they don't have identical truth tables.

Tautology.

A formula is a tautology if it takes the value T for all possible values of its variables.

Example 14.

$P \vee \neg P$ is a tautology because its truth table is:

P	$\neg P$	$P \vee \neg P$
T	F	T
F	T	T

Example 15.

Suppose P, Q, R are three propositions. Then

$$(P \wedge Q) \wedge R \sim P \wedge (Q \wedge R) \qquad (*)$$

Proof:

P	Q	R	$P \wedge Q$	$(P \wedge Q) \wedge R$
T	T	T	T	T
T	T	F	T	F
T	F	T	F	F
T	F	F	F	F
F	T	T	F	F
F	T	F	F	F
F	F	T	F	F
F	F	F	F	F

P	Q	R	$Q \wedge R$	$P \wedge (Q \wedge R)$
T	T	T	T	T
T	T	F	F	F
T	F	T	F	F
T	F	F	F	F
F	T	T	T	F
F	T	F	F	F
F	F	T	F	F
F	F	F	F	F

Remark 11.

One can show that the following equivalences hold:

Idempotency: $P \vee P \sim P, \quad P \wedge P \sim P$

Associativity: $P \vee (Q \vee R) \sim (P \vee Q) \vee R, \quad P \wedge (Q \wedge R) \sim (P \wedge Q) \wedge R,$

Commutativity: $P \vee Q \sim Q \vee P, \quad P \wedge Q \sim Q \wedge P$

Distributivity: $P \wedge (Q \vee R) \sim (P \wedge Q) \vee (P \wedge R), \quad P \vee (Q \wedge R) \sim (P \vee Q) \wedge (P \vee R)$

Identity: Here we consider T to denote a formula which always has the value T and we consider F to denote a formula which always has the value F. Then, we have $P \vee F \sim P$, $P \vee T \sim T, \quad P \wedge F \sim F, \quad P \wedge T \sim P$

Complements: $P \vee \neg P \sim T, \quad P \wedge \neg P \sim F$

Involution: $\neg(\neg P) \sim P$

De Morgan's Laws: $\neg(P \vee Q) \sim \neg P \wedge \neg Q, \quad \neg(P \wedge Q) \sim \neg P \vee \neg Q$

Remark 12.

The equivalences in remark 11 can be used to prove other equivalences and to simplify propositional expressions by using the following two rules:

(i) If we replace a variable in two equivalent formulas by the same arbitrary formula in both, then we will end up with two equivalent formulas again.

(ii) A subformula within a formula can be replaced by an equivalent formula.

MT251P – Lecture 5

Fiacre Ó Cairbre

Example 16.

Use the equivalences in remark 11 to prove that

$$\neg(P \vee Q) \vee (\neg P \wedge Q) \sim \neg P$$

Proof:

$$\neg(P \vee Q) \vee (\neg P \wedge Q) \sim (\neg P \wedge \neg Q) \vee (\neg P \wedge Q), \qquad \text{by De Morgan's Law}$$

$$\sim \neg P \wedge (\neg Q \vee Q), \qquad \text{by using Distributivity}$$

$$\sim \neg P \wedge T, \qquad \text{by using the Complements}$$

$$\sim \neg P \qquad \text{by using the Identity}$$

Chapter 2 – Euclid's Elements.

Section 2.1 – Postulates and Common Notions.

Euclid lived in Alexandria and wrote the Elements around 300BC. His Elements essentially covered all known mathematics at the time. His approach was to accept very few axioms at the beginning and then, using deductive reasoning, he proved many results in mathematics.

A modern online version of the Elements has been made available by David Joyce and I will put a link to the website on moodle. It's important to note that if there are any differences between my notation, definitions etc. and the website above, then you should use my notation, definitions etc.

Postulates.

Note that we don't define what a point or line is but we will accept statements about points and lines. For example, we will accept that between any two different points, we can a draw a unique straight line (Postulate 1 below). We are free to do this in mathematics.

At the beginning of the Elements, Euclid states five postulates. These postulates are like axioms because he accepts them and doesn't prove them.

Postulate 1. – A unique straight line can be drawn between any two points.

Postulate 2. – A given finite straight line can be extended continuously in a straight line.

Postulate 3. – A point (its centre) and a distance (the length of the radius) define a circle.

Postulate 4. – All right angles are equal.

Postulate 5. – If a straight line falling on two straight lines makes the interior angles on the same side less than two right angles, then the two lines meet on that side on which the angles are less than two right angles.

Common Notions.

There are five so called common notions in the Elements. These are statements that Euclid accepted.

Common notion 1. – Things which equal the same thing also equal each other.

Common notion 2. – If equals are added to equals, then the sums are equal.

Common notion 3. – If equals are subtracted from equals, then the remainders are equal.

Common notion 4. – Things which coincide with one another equal one another.

Common notion 5. – The whole is greater than the part.

Some notation and conventions.

(i) We will denote points by capital letters like A, B, C.

(ii) The line between two points A, B is denoted by AB and its length is denoted by $|AB|$.

(iii) The angle formed by the two lines AB, AC at the point B is denoted by $\angle ABC$ (or $\angle CBA$).

(iv) Angles are measured in radians and the magnitude of an angle $\angle ABC$ is denoted by $|\angle ABC|$

(v) Denote the triangle with vertices at (non–collinear) points A, B, C by ΔABC.

(vi) We will denote Postulate 1 by P1, Postulate 2 by P2 etc. We will denote Common Notion 1 by CN1 etc.

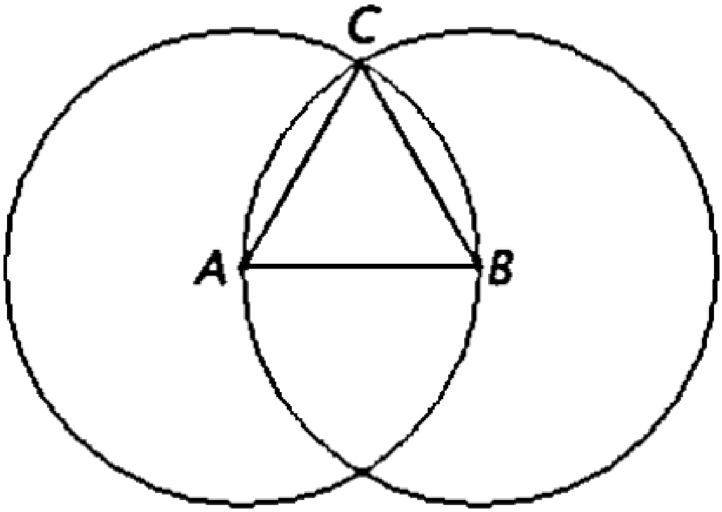
(vii) We will put E in front of proposition numbers to distinguish a result in the Elements from a result here. For example, we will label Proposition I.5 (i.e. Book I, Proposition 5) in the Elements as E.I.5 to distinguish it from Proposition 1.5 here, because Proposition 1.5 here will not be Proposition E.I.5.

Section 2.2. – Fundamental Theorems in Euclidean Geometry.

Proposition 1.1

To construct an equilateral triangle on a given line AB.

Proof.



Draw a circle O_1 with centre A and radius $|AB|$ (by P3) as in the figure above. Draw a circle O_2 with centre B and radius $|AB|$ (by P3). Denote the point of intersection of the two circles by C. Then $|AC| = |AB|$ and $|BC| = |AB|$. By CN1 we then have $|AC| = |BC| = |AB|$ and so the triangle ΔABC is an equilateral triangle and we are done.

Definition 1.

Two triangles ΔABC, ΔDEF are called congruent if

$$|AB| = |DE|, \quad |BC| = |EF|, \quad |AC| = DF|$$

and

$$|\angle ABC| = |\angle DEF|, \quad |\angle BCA| = |\angle EFD|, \quad |\angle CAB| = |\angle FDE|$$

Remark 1.

Propositions E.I.2 and E.I.3 show how we can construct a line at a point A equal to a given line BC and how we can cut from a longer line AB a section equal in length to a given shorter line CD. Proposition E.I.4 states the following first method for proving congruence of triangles.

Proposition 1.2 (SAS)

Suppose we have two triangles ΔABC, ΔDEF. If $|AB| = |DE|$, $|BC| = |EF|$ and $|\angle ABC = |\angle DEF|$, then ΔABC is congruent to ΔDEF.

Remark 2.

Euclid's proof of SAS uses 'superposition'. Nowadays, SAS is typically taken as an axiom and we will do that. We will use two other congruence results.

Fiacre Ó Cairbre

ASA.

Suppose we have two triangles ΔABC, ΔDEF. If $|\angle ABC| = |\angle DEF|$, $|BC| = |EF|$ and $|\angle BCA| = |\angle EFD|$, then ΔABC is congruent to ΔDEF.

SSS.

Suppose we have two triangles ΔABC, ΔDEF. If $|AB| = |DE|$, $|BC| = |EF|$ and $|AC| = |DF|$, then ΔABC is congruent to ΔDEF.

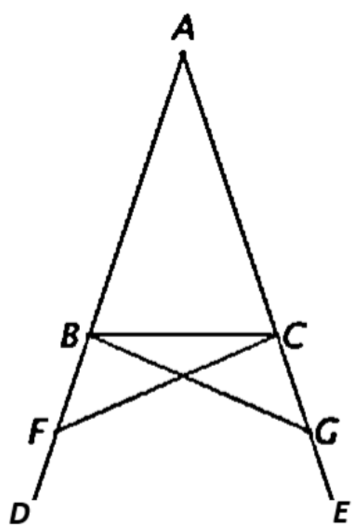
Proposition 1.3

Suppose we have an isosceles triangle ΔABC with $|AB| = |AC|$. Then,

$$|\angle ABC| = |\angle BCA|$$

.

Proof.



Choose F on AD and G on AE such that $|AG| = |AF|$ (by Proposition E.I.3). Join F to C and join B to G. Consider the triangles ΔABG, ΔACF. We have $|AC| = |AB|$, $|AF| = |AG|$, $|\angle FAC| = |\angle BAG|$ and so SAS implies that ΔABG is congruent to ΔACF. So, $|FC| = |BG|$, $|\angle AFC| = |\angle AGB|$.

Hence, $|\angle BFC| = |\angle CGB|$ and $|FC| = |BG|$. Also, since $|AF| = |AG|$ and $|AB| = |AC|$, then we get $|BF| = |CG|$ (by CN3). So, SAS implies that ΔBFC is congruent to ΔCGB.

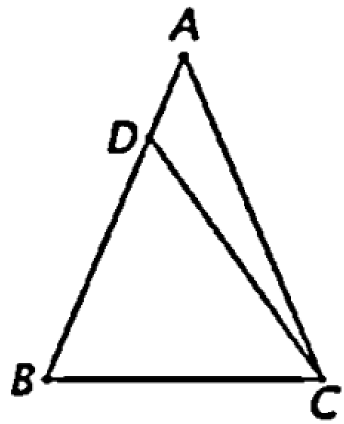
Then,

$$|\angle ABC| = |\angle ABG| - |\angle CBG| = |\angle ACF| - |\angle FCB| = |\angle BCA|$$

Proposition 1.4

Suppose ΔABC is a triangle with $|\angle ABC| = |\angle BCA|$. Then, $|AB| = |AC|$.

Proof.



We will prove this by contradiction. Suppose $|AB| \neq |AC|$. Then , we either have $|AB| > |AC|$ or $|AB| < |AC|$.

CASE 1. Suppose $|AB| > |AC|$.

Choose D on AB with $|BD| = |AC|$. Join C to D and use SAS to get that ΔABC is congruent to ΔDCB which contradicts CN5. Thus, $|AB| > |AC|$ is false.

CASE 2. Suppose $|AB| < |AC|$.

Use a similar approach as in CASE 1 to get thet $|AB| < |AC|$ is false.

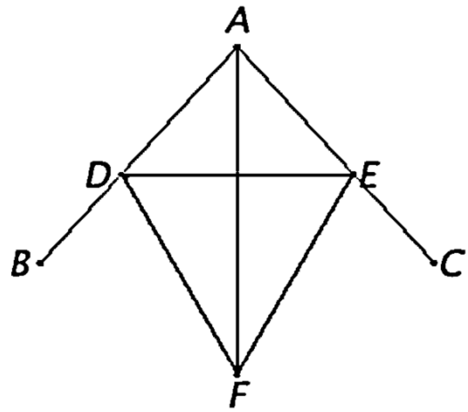
Overall then, we have $|AB| = |AC|$ and we are done.

Section 2.3 – Some elementary constructions.

Proposition 1.5

To construct the bisector of a given angle $\angle BAC$.

Proof.



Choose D on AB and E on AC with $|AE| = |AD|$. Join D to E and construct the equilateral triangle ΔDEF on DE. Join A to F. By SSS we get that ΔADF is congruent to ΔAEF and so $|\angle DAF| = |\angle EAF|$. Hence, AF bisects $\angle DAE$ which equals $\angle BAC$ and we are done.

Section 2.4 – Angles and Parallels.

Remark 3.

We will use the following results:

1. Suppose a line CD is drawn from a point C on the line AB between A and B. Then, $|\angle ACD| + |\angle DCB| = \pi$

2. Suppose C is a point not on a line AB. Then, there exists a unique line through C that is parallel to AB.

Remark 4.

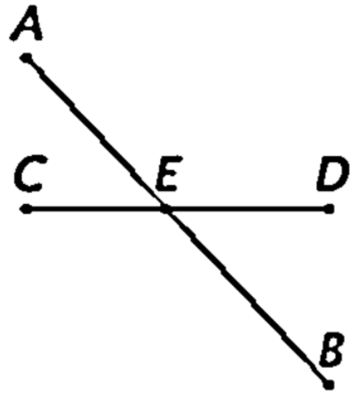
Result 2 in remark 3 above is equivalent to P5. It is a simpler version of P5.

Proposition 1.6

Suppose the line AB and CD intersect at the point E. Then $|\angle AEC| = |\angle BED|$.

The angles $\angle AEC$ and $\angle BED$ are called opposite angles.

Proof.



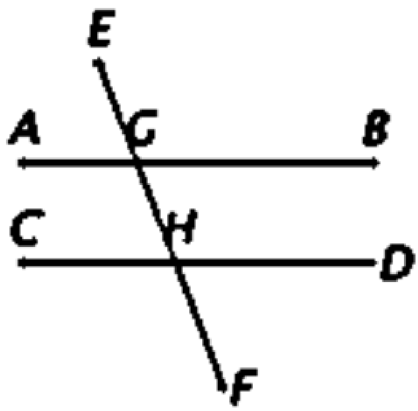
We have $|\angle AEC| + |\angle CEB| = \pi$ and $|\angle CEB| + |\angle BED| = \pi$. By CN1 we then get $|\angle AEC| + |\angle CEB| = |\angle CEB| + |\angle BED|$ Hence $|\angle AEC| = |\angle BED|$.

Fiacre Ó Cairbre

Proposition 1.7

Suppose AB and CD are parallel lines and EF is a third line that intersects AB at G and intersects CD at H. Then, $|\angle AGF| = |\angle DHE|$. The angles $\angle AGF$ and $\angle DHE$ are called alternate angles.

Proof.



We prove it by contradiction. Suppose $|\angle AGF| \neq |\angle DHE|$. Then, either $|\angle AGF| < |\angle DHE|$ or $|\angle AGF| > |\angle DHE|$.

CASE 1. Suppose $|\angle AGF| < |\angle DHE|$. Then

$$|\angle AGF| + |\angle EHC| < |\angle DHE| + |\angle EHC| = \pi$$

By P5, the lines AB and CD meet on the side of A, which is false. So, $|\angle AGF| < |\angle DHE|$ is false.

CASE 2. Suppose $|\angle AGF| > |\angle DHE|$. A similar approach as in CASE 1 will show that $|\angle AGF| > |\angle DHE|$ is false.

So, $|\angle AGF| = |\angle DHE|$.

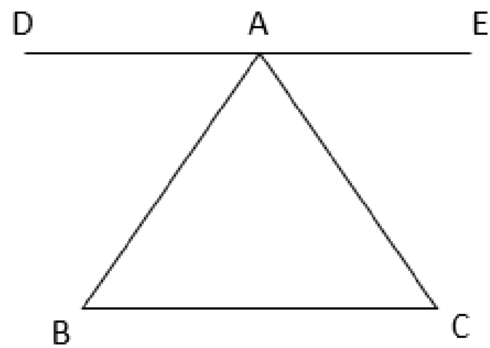
Proposition 1.8

Suppose ΔABC is a triangle. Then,

$$|\angle ABC| + |\angle BCA| + |\angle CAB| = \pi$$

.

Proof.



Draw a line DE through A that is parallel to BC. Then,

$$|\angle DAB| + |\angle CAB| + |\angle EAC| = \pi$$

.

Now,

$$|\angle DAB| = |\angle ABC| \quad \text{and} \quad |\angle EAC| = |\angle BCA|$$

.

So,

$$|\angle ABC| + |\angle CAB| + |\angle BCA| = \pi$$

.

Section 2.5 – Areas.

Note the following result:

Suppose ΔABC is a triangle and let AE denote the perpendicular from A onto BC (or an extension of BC if necessary). Then, the area of ΔABC is $\frac{1}{2}|BC||AE|$.

Proposition 1.9

Consider a triangle ΔABC. Choose D on AB and choose E on AC such that DE is parallel to BC. Then,

$$\frac{|AD|}{|DB|} = \frac{|AE|}{|EC|}$$

Definition 2.

Two triangles ΔABC and ΔDEF are called similar if

$$|\angle ABC| = |\angle DEF|, \quad |\angle BCA| = |\angle EFD|, \quad |\angle CAB| = |\angle FDE|$$

Proposition 1.10

Suppose AB and CD are lines and EF is a third line that intersects AB at G and intersects CD at H. If $|\angle EGB| = |\angle DHE|$, then AB and CD are parallel.

Proposition 1.11

Suppose ΔABC and ΔDEF are similar triangles. Then

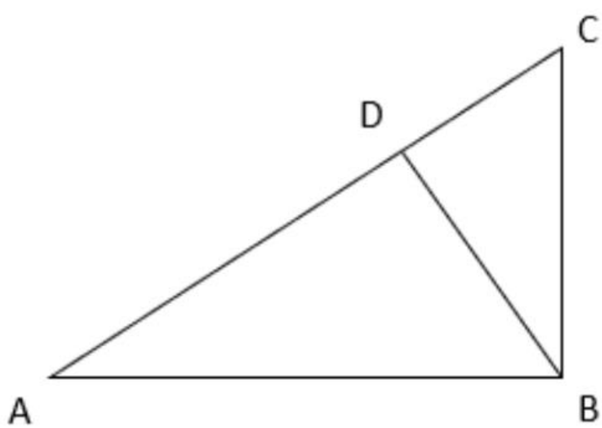
$$\frac{|DE|}{|AB|} = \frac{|EF|}{|BC|} = \frac{|DF|}{|AC|}$$

Proposition 1.12 – Pythagoras' Theorem.

Suppose ΔABC is a right angled triangle with $\angle ABC$ a right angle. Then,

$$|AB|^2 + |BC|^2 = |AC|^2$$

Proof.



Draw a perpendicular from B onto AC at D, The triangles ΔADB and ΔABC are similar and so Proposition 1.11 gives

$$\frac{|AB|}{|AD|} = \frac{|AC|}{|AB|}$$

Hence, $|AB|^2 = |AD||AC|$.

ΔBDC and ΔABC are similar and so

$$\frac{|DC|}{|BC|} = \frac{|BC|}{|AC|}$$

Thus, $|BC|^2 = |DC||AC|$. So,

$$|AB|^2 + |BC|^2 = |AC|(|AD| + |DC|)$$

$$= |AC|^2$$

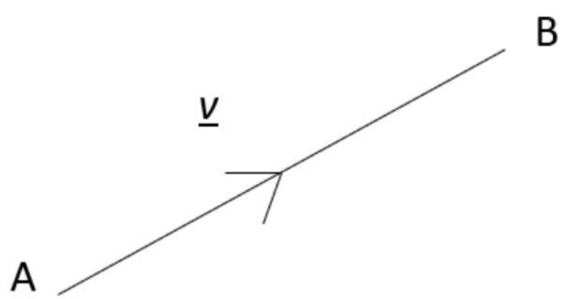
Chapter 3 – Vectors.

Section 3.1 – Introduction.

Remark 1.

Vectors can be used to describe things that require both magnitude and direction. For example, you could say that you are sailing at 30 mph in a south west direction. Here 30 is the magnitude and south west is the direction. If a thing only requires a magnitude, then it can be described by a scalar (or constant). For example, length is a scalar because it only requires a magnitude.

You can think of a vector as being a path (in a particular direction) between two points. Suppose A is your starting point (called the initial point) and B is your finishing point (called the terminal point).



The vector $\underline{v}$ starting at A and finishing at B is also denoted by

$$\underline{v} = \vec{v} = \vec{AB}$$

The magnitude of $\underline{v}$ is the distance from A to B.

MT251P – Lecture 8

Fiacre Ó Cairbre

Remark 2.

Once we know the intial point A and the terminal point B of a vector $\underline{v}$, then we can think of $\underline{v}$ in terms of the coordinates of A and B.

For example, denote the usual two–dimensional xy–plane by

$$\mathbb{R}^2 = \{(x,y) : x, y \in \mathbb{R}\}$$

and suppose

$$A = (x_1, y_1), \ B = (x_2, y_2) \in \mathbb{R}^2$$

Then the vector $\underline{v} = \vec{AB}$ can be written in many ways including:

$$\underline{v} = (x_2 - x_1, y_2 - y_1)$$

$$\underline{v} = v_1 i + v_2 j, \quad \text{where} \quad v_1 = x_2 - x_1, \ v_2 = y_2 - y_1$$

$\underline{v}$ is called the displacement vector.

Example 1.

Suppose $\underline{v} = \vec{AB}$, where $A = (2,1), \ B = (4,-6)$. Then,

$$\underline{v} = (4-2)i + (-6-1)j = 2i - 7j$$

Remark 3.

$0i + 0j$ is called the zero vector.

Theorem 1.

Suppose $\underline{w} = w_1 i + w_2 j$, $\underline{v} = v_1 i + v_2 j$ and $t \in \mathbb{R}$. Then

(a) $\underline{w} + \underline{v} = (w_1 + v_1)i + (w_2 + v_2)j$.

(b) $t\underline{w} = tw_1 i + tw_2 j$.

(c) The magnitude (or length) of $\underline{v}$ is denoted by $||\underline{v}||$ and satisfes

$$||\underline{v}|| = \sqrt{{v_1}^2 + {v_2}^2}$$

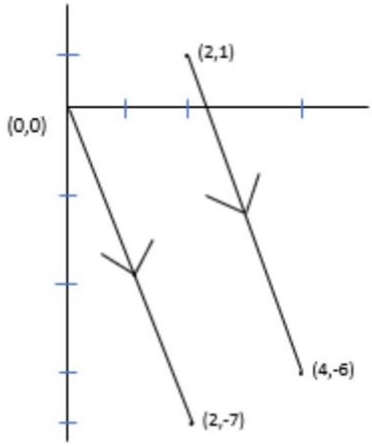
(d) The dot (or scalar) product of $\underline{w}$ and $\underline{v}$ is denoted by $\underline{w}.\underline{v}$ and is defined by

$$\underline{w}.\underline{v} = w_1 v_1 + w_2 v_2$$

(e) Suppose neither $\underline{w}$ nor $\underline{v}$ is the zero vector. Then, $\underline{w}.\underline{v} = ||\underline{w}||||\underline{v}|| \cos\theta$, where θ is the angle between $\underline{w}$ and $\underline{v}$ and $0 \leq \theta \leq \pi$.

Example 2.

Suppose $E = (0,0)$, $F = (2,-7)$ and $\underline{w} = \vec{EF}$. Then, $\underline{w} = \underline{v}$ from example 1 because $\underline{w} = 2i - 7j$. This is an example of the fact that $\vec{AB}$ (from example 1) may equal $\vec{EF}$ even though $A \neq E$ and $B \neq F$. The reason that $\vec{AB}$ and $\vec{EF}$ are actually the same vector is because $\vec{AB}$ and $\vec{EF}$ both have the same magnitude and direction.
The following picture shows $\vec{AB}$ and $\vec{EF}$



Example 3.

Suppose $\underline{w} = 2i - 3j$, $\underline{v} = 4i + 2j$. Then

$$\underline{w} + \underline{v} = 6i - j$$

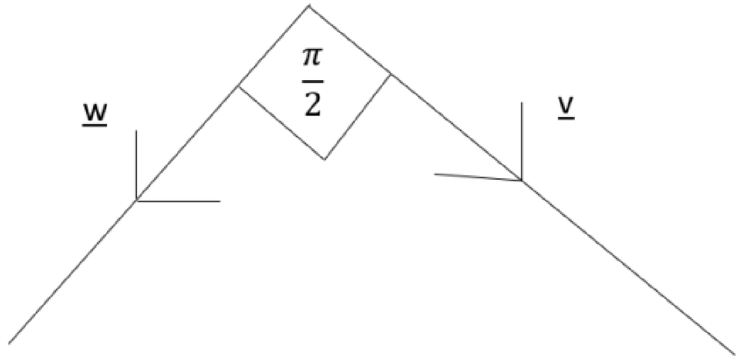
$$2\underline{v} = 8i + 4j$$

$$||\underline{w}|| = \sqrt{4+9} = \sqrt{13}$$

$$\underline{w}.\underline{v} = 2(4) + (-3)2 = 2$$

Definition 1.

The non–zero vectors $\underline{v}, \underline{w}$ are said to be perpendicular (or orthogonal) if the angle between $\underline{v}$ and $\underline{w}$ is $\frac{\pi}{2}$.



We also define the zero vector to be perpendicular to any vector.

Remark 4.

Two vectors $\underline{v}$, $\underline{w}$ in $\mathbb{R}^2$ are perpendicular $\iff$ $\underline{v}.\underline{w} = 0$

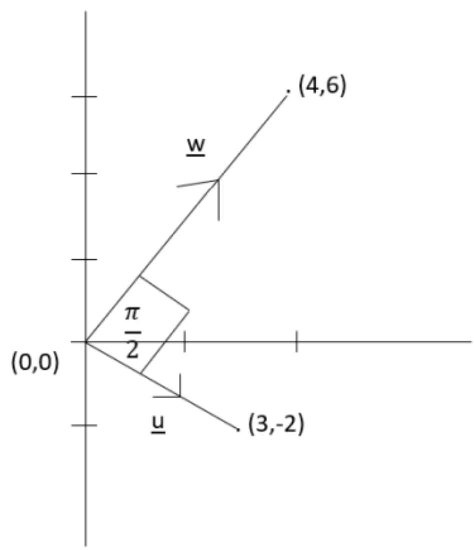
Example 4.

Supppose $\underline{u} = 3i - 2j$ and $\underline{w} = 4i + 6j$. Then, find the angle between $\underline{u}$ and $\underline{w}$.

Solution.

Note that $\underline{u}.\underline{w} = 0$ and so $\underline{u}$ and $\underline{w}$ are perpendicular. Hence, the angle between $\underline{u}$ and $\underline{w}$ is $\frac{\pi}{2}$.

Here is a picture of this example:



Remark 5.

Suppose $A = (0,0)$, $B = (x_1, x_2)$. Then the vector $\underline{v} = \vec{AB}$ is called a position vector. So, a position vector is a vector that has the origin $(0,0)$ as initial point.

MT251P – Lecture 9

Fiacre Ó Cairbre

Three–dimensional vectors.

Remark 6.

Suppose we are in three–dimensional space given by

$$\mathbb{R}^3 = \{(x, y, z) : x, y, z \in \mathbb{R}\}$$

and suppose

$$A = (x_1, y_1, z_1),\ B = (x_2, y_2, z_2) \in \mathbb{R}^3$$

The vector $\underline{u} = \vec{AB}$ can be written as

$$\underline{u} = u_1 i + u_2 j + u_3 k, \quad \text{where} \quad u_1 = x_2 - x_1,\ u_2 = y_2 - y_1,\ u_3 = z_2 - z_1$$

$\underline{u}$ is called the displacement vector.

$0i + 0j + 0k$ is called the zero vector.

Theorem 2.

Suppose $\underline{v} = v_1 i + v_2 j + v_3 k$, $\underline{w} = w_1 i + w_2 j + w_3 k$ and $t \in \mathbb{R}$.

(a) $\underline{v} + \underline{w} = (v_1 + w_1)i + (v_2 + w_2)j + (v_3 + w_3)k$.

(b) $t\underline{v} = tv_1 i + tv_2 j + tv_3 k$.

(c) The magnitude (or length) of $\underline{v}$ is denoted by $||\underline{v}||$ and satisfies

$$||\underline{v}|| = \sqrt{{v_1}^2 + {v_2}^2 + {v_3}^2}$$

(d) The dot (or scalar) product of $\underline{v}$ and $\underline{w}$ is denoted by $\underline{v}.\underline{w}$ and is defined by

$$\underline{v}.\underline{w} = v_1 w_1 + v_2 w_2 + v_3 w_3$$

(e) Suppose neither $\underline{v}$ nor $\underline{w}$ is the zero vector. Then, $\underline{v}.\underline{v} = ||\underline{v}||||\underline{v}|| \cos\theta$, where θ is the angle between $\underline{v}$ and $\underline{w}$ and $0 \leq \theta \leq \pi$.

Definition 2.

The non–zero vectors $\underline{v}, \underline{w}$ are said to be perpendicular if the angle between $\underline{v}$ and $\underline{w}$ is $\frac{\pi}{2}$.

We also define the zero vector to be perpendicular to any vector.

Remark 7.

Two vectors $\underline{v}$, $\underline{w}$ in $\mathbb{R}^3$ are perpendicular $\iff$ $\underline{v}.\underline{w} = 0$

Definition 3.

Suppose $\underline{u} = u_1 i + u_2 j + u_3 k$ and $\underline{w} = w_1 i + w_2 j + w_3 k$. The cross product of $\underline{u}$ and $\underline{w}$ is denoted by $\underline{u} \times \underline{w}$ and is defined by

$$\underline{u} \times \underline{w} = (u_2 w_3 - u_3 w_2) i + (u_3 w_1 - u_1 w_3) j + (u_1 w_2 - u_2 w_1) k$$

Example 5.

Consider the vectors $\underline{u} = 3i - 2j + k$ and $\underline{w} = i + 3j + 3k$. Find

(a) $\underline{u} + \underline{w}$

(b) $||\underline{w}||$

(c) $\underline{u}.\underline{w}$

(d) $\underline{u} \times \underline{w}$

Solution.

(a) $\underline{u} + \underline{w} = 4i + j + 4k$

(b) $||\underline{w}|| = \sqrt{1+9+9} = \sqrt{19}$

(c) $\underline{u}.\underline{w} = 0$

(d) $\underline{u} \times \underline{w} = ((-2)(3) - (1)(3))i + ((1)(1) - (3)(3))j + ((3)(3) - (-2)(1))k = -9i - 8j + 11k$

Remark 8.

If $\underline{u}$ and $\underline{w}$ are vectors in $\mathbb{R}^3$, then $\underline{u} \times \underline{w}$ gives a vector which is perpendicular to both $\underline{u}$ and $\underline{w}$. In example 5, $\underline{m} = \underline{u} \times \underline{w}$ is perpendicular to both $\underline{u}$ and $\underline{w}$ because $\underline{m}.\underline{u} = 0$ and $\underline{m}.\underline{w} = 0$.

Example 6.

Consider the vectors $\underline{u} = i + 2j + k$ and $\underline{w} = 3i + j + 2k$. Find the angle θ between $\underline{u}$ and $\underline{w}$.

Solution.

$\underline{u}.\underline{w} = 7,\ ||\underline{u}|| = \sqrt{6}, ||\underline{w}|| = \sqrt{14}$ and so

$$7 = \sqrt{6}\sqrt{14}\cos\theta \quad \Rightarrow \quad \cos\theta = \frac{7}{\sqrt{84}} \quad \Rightarrow \quad \theta = \arccos(\frac{7}{\sqrt{84}})$$

Remark 9.

Suppose $A = (0,0,0),\ B = (x_1, x_2, x_3)$. The vector $\underline{v} = \vec{AB}$ is called a position vector.

Section 3.2 – Vectors in $\mathbb{R}^n$.

Definition 4.

For $n \geq 1$ we define

$$\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) : x_i \in \mathbb{R} \text{ for } 1 \leq i \leq n\}$$

The vector with initial point at the origin, $(0, 0, \ldots, 0)$, and terminal point $(x_1, x_2, \ldots, x_n)$ is called a position vector and is denoted by

$$\underline{x} = (x_1, x_2, \ldots, x_n)$$

Theorem 3.

Suppose $(x_1, x_2, \ldots, x_n)$ and $(y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$ and $k \in \mathbb{R}$.

(a) $\underline{x} + \underline{y} = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$

(b) $k\underline{x} = (kx_1, kx_2, \ldots, kx_n)$

(c) $||\underline{x}|| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$

(d) $\underline{x}.\underline{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n$

(e) $\underline{x}.\underline{y} = ||\underline{x}||||\underline{y}|| \cos\theta$, where θ is the angle between $\underline{x}$ and $\underline{y}$ and $0 \leq \theta \leq \pi$ and neither $\underline{x}$ nor $\underline{y}$ is the zero vector $(0, 0, \ldots, 0)$

MT251P – Lecture 10

Fiacre Ó Cairbre

Example 7.

Suppose $\underline{x} = (3, -2, 4, 6, 0)$ and $\underline{y} = (0, 1, 0, -2, 5)$ in $\mathbb{R}^5$.

(a) $\underline{x} + \underline{y} = (3, -1, 4, 4, 5)$

(b) $||\underline{x}|| = \sqrt{9 + 4 + 16 + 36} = \sqrt{65}$

(c) $\underline{x}.\underline{y} = -14$

Theorem 4.

Suppose $\underline{x} = (x_1, x_2, \ldots, x_n)$, $\underline{y} = (y_1, y_2, \ldots, y_n)$, $\underline{u} = (u_1, u_2, \ldots, u_n)$ and $\underline{w} = (w_1, w_2, \ldots, w_n) \in \mathbb{R}^n$. Then

(a) $\underline{y} - \underline{x}$ is the vector from $(x_1, x_2, \ldots, x_n)$ to $(y_1, y_2, \ldots, y_n)$. Also, the vector from $(x_1, x_2, \ldots, x_n)$ to $(y_1, y_2, \ldots, y_n)$ is parallel to the vector from $(u_1, u_2, \ldots, u_n)$ to $(w_1, w_2, \ldots, w_n)$ $\Longleftrightarrow$ $\underline{y} - \underline{x} = t(\underline{w} - \underline{u})$, for some $t \in \mathbb{R}$.

(b) The line L passing through $(x_1, x_2, \ldots, x_n)$ and $(y_1, y_2, \ldots, y_n)$ is given by

$$L = \{(x_1, x_2, \ldots, x_n) + t(y_1 - x_1, y_2 - x_2, \ldots, y_n - x_n) : t \in \mathbb{R}\}$$

L is also the line passing through $(x_1, x_2, \ldots, x_n)$ parallel to the vector $\underline{y} - \underline{x}$ and is called the parametric equation of L.

Furthermore the line M passing through $(x_1, x_2, \ldots, x_n)$ and parallel to the vector $(u_1, u_2, \ldots, u_n)$ is given by

$$M = \{(x_1, x_2, \ldots, x_n) + t(u_1, u_2, \ldots, u_n) : t \in \mathbb{R}\}$$

Example 8.

(a) Find the parametric equation of the line L passing through $(1, 2, -1)$ and $(2, 0, 2)$.

(b) Is $(3, -2, 5)$ on L? Is $(3, 0, 1)$ on L?

Solution.

(a) $L = \{(1, 2, -1) + t(1, -2, 3) : t \in \mathbb{R}\}$

(b) $(3, -2, 5)$ is on L because it corresponds to $t = 2$ in (a). $(3, 0, 1)$ is not on L because there is no $t \in \mathbb{R}$ satisfying

$$1 + t = 3, \quad 2 - 2t = 0, \quad -1 + 3t = 1 \qquad (*)$$

Definition 5.

The non–zero vectors $\underline{x}, \underline{y} \in \mathbb{R}^n$ are said to be perpendicular if the angle between them is $\frac{\pi}{2}$.

We also define the zero vector to be perpendicular to any vector.

Theorem 5.

Suppose $\underline{x}, \underline{y} \in \mathbb{R}^n$. Then

(a) $\underline{x}.\underline{y} = 0 \iff \underline{x}$ and $\underline{y}$ are perpendicular

Theorem 6 – Cauchy–Schwartz Inequality.

$|\underline{x}.\underline{y}| \leq ||\underline{x}|| ||\underline{y}||$, for all $\underline{x}, \underline{y} \in \mathbb{R}^n$ $\qquad (i)$

Also, equality holds in (i) above $\iff \underline{x} = t\underline{y}$, for some $t \in \mathbb{R}$ $\qquad (ii)$

Proof.

First we note that

$$(\underline{x} + s\underline{y}).(\underline{x} + s\underline{y}) \geq 0, \text{ for all } s \in \mathbb{R}$$

and so

$$\underline{x}.\underline{x} + s\underline{y}.\underline{x} + s\underline{x}.\underline{y} + s^2\underline{y}.\underline{y} \geq 0, \text{ for all } s \in \mathbb{R}$$

$$\Rightarrow ||\underline{x}||^2 + 2s\underline{x}.\underline{y} + s^2||\underline{y}||^2 \geq 0, \text{ for all } s \in \mathbb{R} \qquad (*)$$

This means that the left hand side of $(*)$ (as a quadratic equation in s) has at most one real root. So, by the quadratic formula we get

$$4(\underline{x}.\underline{y})^2 - 4||\underline{y}||^2||\underline{x}||^2 \leq 0$$

$$\Rightarrow (\underline{x}.\underline{y})^2 \leq ||\underline{x}||^2||\underline{y}||^2$$

$$\Rightarrow |\underline{x}.\underline{y}| \leq ||\underline{x}|| ||\underline{y}|| \qquad (**)$$

and so we have proved (i).

Now equality holds in $(**)$

$$\iff \text{ the quadratic in } s \text{ in } (*) \text{ has exactly one real root}$$

$$\iff (\underline{x} + q\underline{y}).(\underline{x} + q\underline{y}) = 0, \text{ for exactly one } q \in \mathbb{R}$$

$$\iff \underline{x} = -q\underline{y}$$

and so we have proved (ii) and we are done.

Remark 10.

The Cauchy–Schwartz inequality implies that for all non–zero $\underline{x}, \underline{y} \in \mathbb{R}^n$, we have

$$-1 \leq \frac{\underline{x}.\underline{y}}{||\underline{x}||||\underline{y}||} \leq 1$$

Hence, there is a unique θ such that

$$\cos\theta = \frac{\underline{x}.\underline{y}}{||\underline{x}||||\underline{y}||} \quad \text{and} \quad 0 \leq \theta \leq \pi \qquad (a)$$

MT251P – Lecture 11

Fiacre Ó Cairbre

Remark 10. continued.

This means that we can define the angle θ between the two non–zero vectors $\underline{x}, \underline{y}$ as the unique θ satisfying

$$\cos\theta = \frac{\underline{x}.\underline{y}}{||\underline{x}||||\underline{y}||}$$

and this agrees with theorem 3(e).

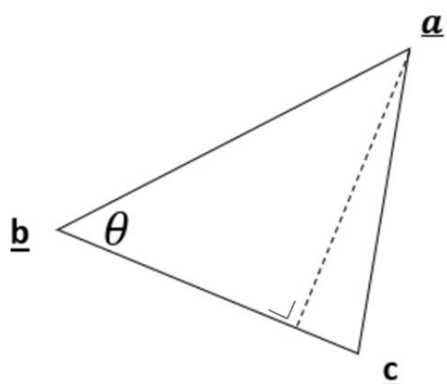
Example 9.

If θ is the angle between the vectors $\underline{x} = (2,0,1,0)$ and $\underline{y} = (1,3,0,0)$ in $\mathbb{R}^4$, then

$$\cos\theta = \frac{\underline{x}.\underline{y}}{||\underline{x}||||\underline{y}||} = \frac{2}{\sqrt{5}\sqrt{10}} = \frac{2}{\sqrt{50}}$$

Area of a triangle.

Three non–collinear vectors $\underline{a}$, $\underline{b}$, $\underline{c}$ in $\mathbb{R}^n$ define a triangle.



We will now derive a formula for the area A of the triangle in terms of the three vectors $\underline{a}$, $\underline{b}$, $\underline{c}$. Well, A is half the base by the perpendicular height.

Now, after translating $\underline{b}$ to the origin, the base can be described by the position vector $\underline{c} - \underline{b}$. The perpendicular height is $||\underline{a} - \underline{b}|||\sin\theta|$ where θ is the angle between $\underline{a} - \underline{b}$ and $\underline{c} - \underline{b}$.

So,

$$A = \frac{1}{2}||\underline{c} - \underline{b}||||\underline{a} - \underline{b}|||\sin\theta|$$

$$\Rightarrow A^2 = \frac{1}{4}||\underline{c}-\underline{b}||^2||\underline{a}-\underline{b}||^2\sin^2\theta$$

$$= \frac{1}{4}(||\underline{c}-\underline{b}||^2||\underline{a}-\underline{b}||^2 - ||\underline{c}-\underline{b}||^2||\underline{a}-\underline{b}||^2\cos^2\theta)$$

$$= \frac{1}{4}(||\underline{c}-\underline{b}||^2||\underline{a}-\underline{b}||^2 - ((\underline{c}-\underline{b}).(\underline{a}-\underline{b}))^2)$$

$$\Rightarrow A = \frac{1}{2}\sqrt{||\underline{c}-\underline{b}||^2||\underline{a}-\underline{b}||^2 - ((\underline{c}-\underline{b}).(\underline{a}-\underline{b}))^2}$$

Example 10.

Find the area A of the tiangle formed by the three vectors $\underline{a} = (1,1,0)$, $\underline{b} = (2,3,1)$, $\underline{c} = (0,2,2)$.

Solution.

$$||\underline{c}-\underline{b}||^2 = 6, \quad ||\underline{a}-\underline{b}||^2 = 6, \quad (\underline{c}-\underline{b}).(\underline{a}-\underline{b}) = 3$$

and so

$$A = \frac{1}{2}\sqrt{36-9} = \frac{1}{2}\sqrt{27}$$

Example 11.

Find the parametric equation of the line L which contains the point $(-1,4,3)$ and is parallel to the line K which has parametric equation

$$K = \{(9,0,-1) + t(-1,4,2) : t \in \mathbb{R}\}$$

Also, find another point on L different from $(-1,4,3)$.

Solution.

K is parallel to the vector $\underline{z} = (-1,4,2)$ and so L is parallel to $\underline{z}$. Hence, the parametric equation of L is

$$L = \{(-1,4,3) + t(-1,4,2) : t \in \mathbb{R}\} \qquad (*)$$

To find another point on L different from $(-1,4,3)$, we let t be any non–zero real number in $(*)$. For example, let $t = 1$ to get the point

$$(-1,4,3) + (-1,4,2) = (-2,8,5)$$

and so $(-2,8,5)$ is another point on L different from $(-1,4,3)$.

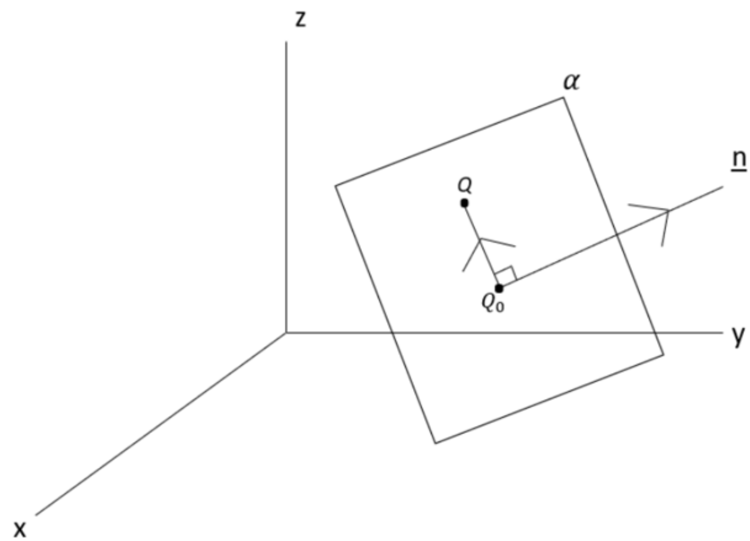
Section 3.3 – Planes in $\mathbb{R}^3$.

Definition 6.

Consider the plane α which contains the point $Q_0 = (x_0, y_0, z_0)$ and suppose the non–zero vector $\underline{n} = ai + bj + ck$ is perpendicular to α. Then, the equation of α is

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

Remark 11.



The motivation for the above definition is that if $Q = (x, y, z)$ is any point in α, then the vector $\vec{Q_0Q}$ is perpendicular to $\underline{n}$ and so the dot product

$$\underline{n}.\vec{Q_0Q} = 0$$

which means

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

Remark 12.

$\underline{n}$ is called a normal vector to α in definition 6 above.

Example 12.

Find the equation of the plane α which contains the point $(1, 2, -4)$ and has normal vector $\underline{n} = 3i - j + 2k$.

Solution.

The equation of α is

$$3(x-1)-(y-2)+2(z+4)=0$$

$$\Rightarrow 3x-y+2z+7=0$$

Example 13.

Find the equation of the plane α which contains the point $(1,-1,3)$ and is perpendicular to the line L with parametric equation

$$L=\{(2,1,-2)+t(-3,2,4): t\in\mathbb{R}\}$$

Solution.

L is parallel to $\underline{n}=-3i+2j+4k$ and so α is perpendicular to $-3i+2j+4k$. Hence, the equation of α is

$$-3(x-1)+2(y+1)+4(z-3)=0$$

$$\Rightarrow -3x+2y+4z=7$$

$$\Rightarrow 3x-2y-4z+7=0$$

Example 14.

Find the equation of the plane α containing the three points $(1,0,0)$, $(1,2,1)$, $(2,1,0)$.

MT251P – Lecture 12

Fiacre Ó Cairbre

Solution to example 14.

Let

$$A = (1, 0, 0), \quad B = (1, 2, 1), \quad C = (2, 1, 0)$$

Then, let

$$\underline{n} = \vec{AB} \times \vec{AC}$$

Then, $\underline{n}$ will be perpendicular to both $\vec{AB}$ and $\vec{AC}$. Also, $\underline{n}$ can be taken to be a normal vector to α. Now,

$$\underline{n} = (2j + k) \times (i + j)$$

$$= -i + j - 2k$$

and so the equation of α is

$$-(x - 1) + (y - 0) - 2(z - 0) = 0$$

$$\Rightarrow -x + y - 2z + 1 = 0$$

$$\Rightarrow x - y + 2z = 1$$

Example 15.

Find the parametric equation of the line L which is the intersection of the two planes

$$x + y + z = 1 \quad \text{and} \quad x + 2y + 3z = 2$$

Solution.

$$\underline{n} = i + j + k \quad \text{is a normal vector to the plane} \quad x + y + z = 1$$

$$\underline{s} = i + 2j + 3k \quad \text{is a normal vector to the plane} \quad x + 2y + 3z = 2$$

Now, a vector parallel to L is perpendicular to both $\underline{n}$ and $\underline{s}$. So, we can take $\underline{n} \times \underline{s}$ as a vector parallel to L.

Now,

$$\underline{n} \times \underline{s} = i - 2j + k$$

and so $i - 2j + k$ is a vector parallel to L. We will now show how to find a point on L. Well, one can show that any line in $\mathbb{R}^3$ must have at least one point with either $x = 0$, $y = 0$ or $z = 0$.

Try $z = 0$ to get

$$x + y = 1 \quad \text{and} \quad x + 2y = 2 \qquad\qquad (*)$$

$x = 0$, $y = 1$ will satisfy $(*)$. So, $(0, 1, 0)$ is a point on L.

So, overall, we have that L is a line containing the point $(0, 1, 0)$ and parallel to the vector $i - 2j + k$. Hence, the parametric equation of L is

$$L = \{(0, 1, 0) + t(1, -2, 1) : t \in \mathbb{R}\}$$

Theorem 7 – Shortest distance from a point to a plane.

Suppose (x_0, y_0, z_0) is a point not on the plane $ax + by + cz + d = 0$. Then, the shortest distance from (x_0, y_0, z_0) to this plane is

$$\left| \frac{ax_0 + by_0 + cz_0 + d}{\sqrt{a^2 + b^2 + c^2}} \right|$$

Example 16.

Find the shortest distance from $(2, 2, 3)$ to the plane $2x + 2y - 3z + 3 = 0$.

Solution.

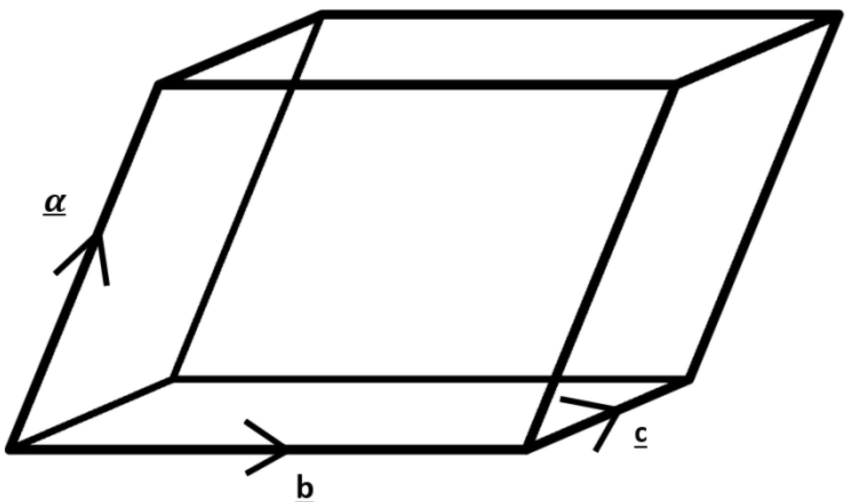
The required shortest distance is

$$\left| \frac{2(2) + 2(2) + 3(-3) + 3}{\sqrt{2^2 + 2^2 + (-3)^2}} \right|$$

$$= \frac{2}{\sqrt{17}}$$

Volume of a Parallelepiped in $\mathbb{R}^3$.

A parallelepiped is a three–dimensional generalisation of a parallelogram as in the picture. A parallelepiped is determined by the three vectors $\underline{a}, \underline{b}, \underline{c}$.



Theorem 8.

The volume of the above parallelepiped is

$$|\underline{a}.(\underline{b} \times \underline{c})|$$

$\underline{a}.(\underline{b} \times \underline{c})$ is called the triple product of the three vectors $\underline{a}, \underline{b}, \underline{c}$.

Example 17.

Find the volume of the parallelepiped determined by $\underline{a} = (1, 0, -1)$, $\underline{b} = (2, 1, 2)$, $\underline{c} = (3, 2, -1)$.

Solution.

The required volume V is given by

$$V = |\underline{a}.(\underline{b} \times \underline{c})|$$

Now, $\underline{b} \times \underline{c} = -5i + 8j + k$ and so $\underline{a}.(\underline{b} \times \underline{c}) = -6$. Hence, $V = 6$.

Chapter 4 – Matrices.

Section 4.1 – Systems of linear equations.

Definition 1.

A linear equation in the n variables $x_1, x_2, \ldots, x_n$ is given by

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

where $b,\ a_i$ are constants for $1 \leq i \leq n$.

Example 1.

$2x_1 + 3x_2 = 5$ is a linear equation. This is also the equation of a line in $\mathbb{R}^2$.

Example 2.

$2x_1 + 3x_2 - x_3 = 5$ is a linear equation. This is also the equation of a plane in $\mathbb{R}^3$.

Example 3.

(i) $3x - \frac{1}{2}y + 2 = 0$ is a linear equation.

(ii) $x^2 + 3y = 5$ is not a linear equation.

(iii) $\cos x + 4y - z = 2$ is not a linear equation.

Example 4.

Note that two lines in $\mathbb{R}_2$:

$$ax + by = c \quad \text{and} \quad rx + sy = t \quad \text{where} \quad a, b, c, r, s, t \quad \text{are constants}$$

intersect in either no points, exactly one point or or infinitely many points.

MT251P – Lecture 13

Fiacre Ó Cairbre

This theorem should have appeared after the Cauchy Schwartz inequality in Lecture 10.

Theorem – The Triangle Inequality.

$||\underline{x}+\underline{y}|| \leq ||\underline{x}|| + ||\underline{y}||$, for all $\underline{x},\ \underline{y} \in \mathbb{R}^n$.

Now, we return to chapter 4 and continue from the end of lecture 12.

Definition 2.

A system of k linear equations in the n variables $x_1, x_2, \ldots\ x_n$ is given by:

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2$$

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

$$a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n = b_k$$

where a_{ij}, b_i are constants for $1 \leq i \leq k,\ 1 \leq j \leq n$.

A solution of the above system of equations is a list of n numbers $s_1, s_2, \ldots, s_n$ such that $x_1 = s_1, x_2 = s_2, \ldots, x_n = s_n$ gives a solution to every equation in the above system. The set of all solutions is called the solution set.

Two systems of linear equations are called equivalent if they have the same solution set.

Definition 3.

A $k \times n$ matrix is a rectangular array of numbers with k rows (horizontal) and n columns (vertical).

Example 5.

$$\begin{pmatrix} 2 & 4 \\ -1 & 0 \\ 2 & 1 \end{pmatrix} \quad \text{is a} \quad 3 \times 2 \quad \text{matrix}$$

Remark 1.

A $k \times n$ matrix A can be written in the form:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & . & . \\ . & . & . & . \\ a_{k1} & a_{k2} & \dots & a_{kn} \end{pmatrix}$$

where a_{ij} is the entry in the i^{th} row and j^{th} column. We can also write $A = [a_{ij}]$.

Definition 4.

Two matrices $A = [a_{ij}]$ and $B = [bij]$ are defined to be equal if A and B have the same number of rows (say k) and the same number of columns (say n) and $a_{ij} = b_{ij}$, for $1 \leq i \leq k,\ 1 \leq j \leq n$.

Definition 5.

An elementary row operation on a matrix A consists of one of the following operations, where R_i denotes the i^{th} row of A:

(a) Interchange R_i and R_j.

(b) Replace R_i with cR_i, where $c \neq 0$.

(c) Replace R_i with $R_i + dR_j$, where $i \neq j$ and d is some constant.

Definition 6.

We say that two matrices are row equivalent if one is obtained from the other by a collection of elementary row operations.

A row in a matrix is called a zero row if all the entries in the row are 0. A row that is not a zero row is called a non–zero row.

Definition 7.

A $k \times n$ matrix B is said to be in reduced row echelon form (RREF) if satisfies the following conditions:

1. In a non–zero row, the first non–zero entry is a 1. We call this 1 a leading 1.

2. If there any zero rows, then they are grouped together at the bottom of the matrix.

3. In any two successive non–zero rows, the leading 1 in the lower row occurs further to the right than the leading 1 in the higher row.

4. Each column that contains a leading 1 (from some row) has zeros everywhere else.

Definition 8.

A $k \times n$ matrix B is said to be in row echelon form (REF) if it satisfies conditions 1,2 and 3 in definition 7 (but doesn't necessarily satisfy condition 4).

Example 6.

$$\begin{pmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 7 \\ 0 & 0 & 1 & -1 \end{pmatrix} \qquad \text{is in RREF}$$

Theorem 1.

Any matrix A is row equivalent to a unique matrix B which is in RREF.

Section 4.2 – How to find a solution set of a system of linear equations.

Remark 2.

We now look at finding the solution set of the following system linear equations S in definition 2:

First we define the augmented matrix A' of S given by

$$A' = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ . & . & . & . & . \\ . & . & . & . & . \\ a_{k1} & a_{k2} & \dots & a_{kn} & b_k \end{pmatrix}$$

Now, any elementary row operation on A' gives the augmented matrix of a system of linear equations that is equivalent to S.

We perform on A' a collection of elementary row operations until we obtain a matrix C which is in RREF. It will be convenient to read off the solution set of the system of linear equations V corresponding to C and this will be the same as the solution set of S. This strategy is called Gauss–Jordan elimination.

Example 7.

Use Gauss–Jordan elimination to find the solution set of the following system T of linear equations:

$$x_1 + 2x_2 + x_3 = 3$$

$$2x_1 + 5x_2 - x_3 = -4$$

$$3x_1 - 2x_2 - x_3 = 5$$

MT251P – Lecture 14

Fiacre Ó Cairbre

Example 7.

Use Gauss–Jordan elimination to find the solution set of the following system T of linear equations:

$$x_1 + 2x_2 + x_3 = 3$$

$$2x_1 + 5x_2 - x_3 = -4$$

$$3x_1 - 2x_2 - x_3 = 5$$

Solution.

The augmented matrix for T is

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

We perform elementary row operations on this augmented matrix. Replace R_2 with $R_2 - 2R_1$ and replace R_3 with $R_3 - 3R_1$ to get

$$\begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & -3 & -10 \\ 0 & -8 & -4 & -4 \end{pmatrix}$$

Replace R_3 with $R_3 + 8R_2$ to get

$$\begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & -3 & -10 \\ 0 & 0 & -28 & -84 \end{pmatrix}$$

Replace R_3 with $-\frac{1}{28}R_3$ to get

$$\begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & -3 & -10 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

Replace R_1 with $R_1 - R_3$ and replace R_2 with $R_2 + 3R_3$ to get

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

Replace R_1 with $R_1 - 2R_2$ to get

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

which is in RREF.

The system of linear equations V corresponding to C is:

$$x_1 = 2$$

$$x_2 = -1$$

$$x_3 = 3$$

The solution set of V is $x_1 = 2,\ x_2 = -1,\ x_3 = 3$. Hence, by remark 3, the solution set of the original system of linear equations T is also

$$x_1 = 2,\ x_2 = -1,\ x_3 = 3$$

Example 8.

Find the solution set of the following system T of linear equations:

$$x_1 - 3x_2 + 5x_3 = 3$$

$$x_1 - 2x_2 + 7x_3 = 5$$

$$2x_1 - 6x_2 + 10x_3 = 5$$

Solution.

The augmented matrix for T is

$$\begin{pmatrix} 1 & -3 & 5 & 3 \\ 1 & -2 & 7 & 5 \\ 2 & -6 & 10 & 5 \end{pmatrix}$$

Replace R_2 with $R_2 - R_1$ and replace R_3 with $R_3 - 2R_1$ to get

$$E = \begin{pmatrix} 1 & -3 & 5 & 3 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Note that E is not in RREF but we can actually stop here without having to go to RREF because the solution set of the system of linear equations W corresponding to E can be read off conveniently as follows:

The system of linear equations W is:

$$x_1 - 3x_2 + 5x_3 = 3$$

$$x_2 + 2x_3 = 2$$

$$0 = -1$$

and so W has no solutions. So, T has no solutions.

Remark 3.

If in remark 2 we stop at REF instead of RREF, then the strategy is called Gaussian Elimination.

Example 9.

Use Gaussian elimination to find the solution set of the following system U of linear equations:

$$x_1 + 2x_2 + x_3 = 3$$

$$3x_1 - x_2 - 3x_3 = -1$$

$$2x_1 + 3x_2 + x_3 = 4$$

The augmented matrix for U is

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

Replace R_2 with $R_2 - 3R_1$ and replace R_3 with $R_3 - 2R_1$ to get

$$\begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & -7 & -6 & -10 \\ 0 & -1 & -1 & -2 \end{pmatrix}$$

Interchange R_3 and R_2 to get

$$\begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & -1 & -1 & -2 \\ 0 & -7 & -6 & -10 \end{pmatrix}$$

Replace R_2 with $-R_2$ to get

$$\begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & -7 & -6 & -10 \end{pmatrix}$$

Replace R_3 with $R_3 + 7R_2$ to get

$$C = \begin{pmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 4 \end{pmatrix}$$

which is in REF (but not RREF).

The system of linear equations V corresponding to C is:

$$x_1 + 2x_2 + x_3 = 3$$

$$x_2 + x_3 = 2$$

$$x_3 = 4$$

The solution set of V is $x_3 = 4$, $\;x_2 = 2 - x_3 = -2$, $\;x_1 = 3 - 2x_2 - x_3 = 3$. So, the solution set of U is also

$$x_1 = 3,\ x_2 = -2,\ x_3 = 4$$

Section 4.3 – Free variable case.

Remark 4.

Suppose, as in remark 2, the augmented matrix C is in RREF. If there are less non–zero rows in C than variables, then we are in what is called the free variable case. In this case. there will be infinitely many solutions and here is how we will read off the required solution set:

If a leading 1 in some row in C occurs in column k, then we say the variable x_k is a leading variable. Otherwise, we say x_k is a free variable. We can assign any real number value to the free variables and we can write the leading variables in terms of the free variables. So, the values of the leading variables will be determined by the values of the free variables.

Theorem 3.

Suppose S is a system of linear equations as in definition 2. Then, S has either no solution, a unique solution or infinitely many solutions.

Example 10.

Find the solution set of the following system W of linear equations:

$$x_1 + 6x_2 + 3x_4 = 2$$

$$12x_2 + 2x_1 + x_3 + 6x_4 = 0$$

$$3x_1 + x_3 + 18x_2 + 9x_4 = 2$$

Solution.

The augmented matrix for W is

$$\begin{pmatrix} 1 & 6 & 0 & 3 & 2 \\ 2 & 12 & 1 & 6 & 0 \\ 3 & 18 & 1 & 9 & 2 \end{pmatrix}$$

MT251P – Lecture 15

Continuation of example 10.

Replace R_2 with $R_2 - 2R_1$ and replace R_3 with $R_3 - 3R_1$ to get

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

Replace R_3 with $R_3 - R_2$ to get

$$C = \begin{pmatrix} 1 & 6 & 0 & 3 & 2 \\ 0 & 0 & 1 & 0 & -4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

which is in RREF.

By remark 4, we are in the free variable case. x_1 and x_3 are leading variables and x_2 and x_4 are free variables.

The system of linear equations V corresponding to C is:

$$x_1 + 6x_2 + 3x_4 = 2 \qquad (i)$$

$$x_3 = -4 \qquad (ii)$$

By remark 4, we say $x_2 = s$ and $x_4 = t$, where s can be any real number and t can be any real number. We write the leading variables, $x_1,\ x_3$ in terms of the free variables $x_2,\ x_4$ as follows:

$$x_1 = 2 - 6x_2 - 3x_4 = 2 - 6s - 3t \qquad (iii)$$

$$x_3 = -4 \qquad (ii)$$

So, the required solution set for W is

$$\{(2 - 6s - 3t, s, -4, t) : s, t \in \mathbb{R}\}$$

Section 4.4 – Matrix Operations.

Definition 9.

(i) If C is a $k \times n$ matrix, we say that $k \times n$ is the size of C.

(ii) Suppose A and B are two matrices. Then the sum, $A + B$ and the difference $A - B$ are only defined if A and B have the same size. In this case we add or subtract the matrices by adding or subtracting the corresponding entries.

(iii) If A is a matrix and t is a scalar (i.e. t is a real number), then tA is the matrix obtained by multiplying each entry of A by t.

Example 11.

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

$$\text{If } \; A = \begin{pmatrix} 2 & -3 & 4 \\ 9 & 0 & 2 \end{pmatrix} \quad \text{then } \; 3A = \begin{pmatrix} 6 & -9 & 12 \\ 27 & 0 & 6 \end{pmatrix}$$

Definition 10 – Matrix multiplication.

Suppose $A = [a_{ij}]$ is a $k \times n$ matrix and $B = [b_{ij}]$ is an $n \times p$ matrix. Then, the (matrix) product AB is a $k \times p$ matrix and is defined as $AB = [c_{ij}]$ where

$$c_{ij} = \sum_{t=1}^{n} a_{it} b_{tj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

Example 12.

$$A = \begin{pmatrix} 2 & 3 \\ -1 & 0 \\ 2 & 4 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & -2 \\ 3 & 2 \end{pmatrix}$$

$$AB = \begin{pmatrix} 2(1) + 3(3) & 2(-2) + 3(2) \\ -1(1) + 0(3) & -1(-2) + 0(2) \\ 2(1) + 4(3) & 2(-2) + 4(2) \end{pmatrix} = \begin{pmatrix} 11 & 2 \\ -1 & 2 \\ 14 & 4 \end{pmatrix}$$

Definition 11.

(i) A square matrix is a matrix with the same number of rows as columns.

(ii) If $A = [a_{ij}]$ is square matrix of size $n \times n$, then the entries $a_{11},\ a_{22},\ a_{33}, \ldots, a_{nn}$ are called the entries along the main diagonal of A.

(iii) The trace of a square matrix A is the sum of the entries along the main diagonal of A and is denoted by $tr(A)$, i.e.

$$tr(A) = \sum_{i=1}^{n} a_{ii}, \quad \text{if} \quad A = [a_{ij}] \quad \text{is of size} \;\; n \times n$$

Example 13.

$$A = \begin{pmatrix} 2 & -3 & 0 \\ 4 & 1 & 2 \\ 9 & -3 & -1 \end{pmatrix}$$

$tr(A) = 2$

Definition 12.

If A is a $k \times n$ matrix, then the transpose of A is denoted by A^T and is the $n \times k$ matrix we get by swapping the rows and columns of A, i.e. the i^{th} row of A^T is the i^{th} column of A and the j^{th} column of A^T is the j^{th} row of A.

Example 14.

If $A = \begin{pmatrix} -1 & 0 & 3 \\ 4 & 2 & 1 \end{pmatrix}$ then $A^T = \begin{pmatrix} -1 & 4 \\ 0 & 2 \\ 3 & 1 \end{pmatrix}$

Remark 5.

If A is a $k \times n$ matrix and B is an $n \times p$ matrix, then

(i) $(A^T)^T = A$

(ii) $(AB)^T = B^T A^T$.

Remark 6 – Properties of Matrix operations.

Suppose we have matrices $A,\ B,\ C$ and scalars $r,\ t$ and suppose that everything below is defined. Then,

1. $A + B = B + A$
2. $A + (B + C) = (A + B) + C$
3. $A(BC) = (AB)C$
4. $A(B + C) = AB + AC$
5. $(A + B)C = AC + BC$
6. $r(A + B) = rA + rB$
7. $(r + t)A = rA + tA$
8. $(rt)A = r(tA)$
9. $r(AB) = (rA)B = A(rB)$

MT251P – Lecture 16

Definition 13.

The identity $n \times n$ matrix is denoted by I_n and is the $n \times n$ matrix with all the entries along the main diagonal equal to one and all other entries equal to zero.

Example 15.

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Remark 7.

If A is any $k \times n$ matrix, then $I_k A = A$ and $AI_n = A$.

Definition 14.

(i) A symmetric matrix B is a square matrix such that $B^T = B$.

(ii) A diagonal matrix B is a square matrix such that all the entries not on the main diagonal are zero.

(iii) A lower triangular matrix is a square matrix such that all the entries above the main diagonal are zero. An upper triangular matrix is a square matrix such that all the entries below the main diagonal are zero.

Example 16.

$$\begin{pmatrix} 1 & 4 & 7 \\ 4 & 2 & -1 \\ 7 & -1 & 5 \end{pmatrix} \quad \text{is symmetric}$$

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix} \quad \text{is a diagonal matrix}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 3 & 2 & 0 \\ -1 & 3 & 4 \end{pmatrix} \quad \text{is lower triangular}$$

Chapter 5 – The Inverse and Determinant of a Matrix.

Remark 1.

Let $M_{k \times n}(\mathbb{R})$ denote the set of $k \times n$ matrices with real entries. So for example

$$\begin{pmatrix} 0 & -1 & 4 \\ 2 & 3 & 1 \end{pmatrix} \in M_{2 \times 3}(\mathbb{R})$$

Definition 1.

Suppose $A \in M_{n\times n}(\mathbb{R})$. Then, we say that A is invertible if there is a matrix $B \in M_{n\times n}(\mathbb{R})$ such that

$$AB = I_n = BA$$

In this case B is called an inverse of A.

Theorem 1.

Suppose $A \in M_{n\times n}(\mathbb{R})$ is invertible. Then,

(i) If B and C are both inverses of A, then $B = C$, i.e. the inverse of A is unique. Hence, we call B the inverse of A and we denote it by A^{-1}.

(ii) A^{-1} is invertible and $(A^{-1})^{-1} = A$

(iii) Suppose $W \in M_{n\times n}(\mathbb{R})$ is invertible. Then, AW is invertible and $(AW)^{-1} = W^{-1}A^{-1}$.

Remark 2.

Here is a proof of Theorem 1(i):

First note that $AB = AC = I_n$. So, we have

$$AB = AC \Rightarrow B(AB) = B(AC)$$

$$\Rightarrow (BA)B = (BA)C$$

$$\Rightarrow I_nB = I_nC$$

$$\Rightarrow B = C$$

Theorem 2.

Suppose $A \in M_{2\times 2}(\mathbb{R})$ with

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Then, A is invertible $\iff ad - bc \neq 0$

Also, if $ad - bc \neq 0$, then

$$A^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Example 1.

Suppose

$$A = \begin{pmatrix} 2 & -1 \\ 0 & 4 \end{pmatrix}$$

Then, A is invertible and

$$A^{-1} = \frac{1}{8}\begin{pmatrix} 4 & 1 \\ 0 & 2 \end{pmatrix}$$

Definition 2.

We know from Theorem 1 in chapter 4 that any matrix A is row equivalent to a unique matrix B which is in RREF. We call this matrix B the RREF of A.

Theorem 3.

An $n \times n$ matrix A is invertible $\iff$ the RREF of $A = I_n$.

Also, in this case where A is invertible, the collection of elementary row operations that take you from A to I_n will also take you from I_n to A^{-1}.

Example 2.

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

Does A^{-1} exist? If it does, then find A^{-1}.

Solution.

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

Interchange R_2 with R_1 (in both matrices A and I_3) to get

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

Replace R_1 with $\frac{1}{2}R_1$ (in both matrices) to get

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

MT251P – Lecture 17

Continuation of example 2.

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

Replace R_2 with $R_2 - 3R_1$ (in both matrices) to get

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 5 \\ 0 & 1 & 1 \end{pmatrix} \qquad \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ 1 & -\frac{3}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Interchange R_3 with R_2 (in both matrices) to get

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 5 \end{pmatrix} \qquad \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ 1 & -\frac{3}{2} & 0 \end{pmatrix}$$

Replace R_1 with $R_1 + \frac{1}{5}R_3$ and replace R_2 with $R_2 - \frac{1}{5}R_3$ (in both matrices) to get

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{pmatrix} \qquad \begin{pmatrix} 0.2 & 0.2 & 0 \\ -0.2 & 0.3 & 1 \\ 1 & -\frac{3}{2} & 0 \end{pmatrix}$$

Replace R_3 with $\frac{1}{5}R_3$ (in both matrices) to get

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 0.2 & 0.2 & 0 \\ -0.2 & 0.3 & 1 \\ 0.2 & -0.3 & 0 \end{pmatrix}$$

So, the RREF of $A = I_3$ and hence A is invertible. So, A^{-1} exists and

$$A^{-1} = \begin{pmatrix} 0.2 & 0.2 & 0 \\ -0.2 & 0.3 & 1 \\ 0.2 & -0.3 & 0 \end{pmatrix}$$

Theorem 4.

Suppose $A,\ B \in M_{n\times n}(\mathbb{R})$. Then

$$AB = I_n \iff BA = I_n$$

Remark 3.

A consequence of theorem 4 is that if you want to check that $C = A^{-1}$, then you only need to check that $AC = I_n$ and you don't need to check that $CA = I_n$.

Section 5.2 – The Determinant of a Matrix.

Definition 3.

Suppose A is an $n \times n$ matrix. The determinant of A is denoted by $\det(A)$ and is defined as follows:

If $A = (a_{11})$ is a 1×1 matrix, then $\det(A) = a_{11}$.

If $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ is a 2×2 matrix then $\det(A) = a_{11}a_{22} - a_{12}a_{21}$

Now, assume that the determinant has been defined for all $(n-1) \times (n-1)$ matrices and let $A = [a_{ij}]$ be an $n \times n$ matrix with a_{ij} in the i^{th} row and j^{th} column of A.

How do we find $\det(A)$?

Let $D(i,j)$ be the $(n-1) \times (n-1)$ matrix obtained by deleting row i and column j of A.

Let $M_{i,j} = \det D(i,j)$

Let $C_{i,j} = (-1)^{i+j} M_{i,j}$.

$C_{i,j}$ is called the ij cofactor of A.

Then we define

$$\det(A) = \sum_{j=1}^{n} a_{1j} C_{1j}$$

Example 3.

Suppose

$$A = \begin{pmatrix} 4 & 1 & 9 \\ 3 & 1 & -2 \\ -1 & 1 & 3 \end{pmatrix}$$

Find $\det(A)$.

Solution.

$$C_{11} = M_{11} = \det D(1,1) = \det \begin{pmatrix} 1 & -2 \\ 1 & 3 \end{pmatrix} = 5$$

$$C_{12} = -M_{12} = -\det D(1,2) = -\det \begin{pmatrix} 3 & -2 \\ -1 & 3 \end{pmatrix} = -7$$

$$C_{13} = M_{13} = \det D(1,3) = \det \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix} = 4$$

So,

$$\det(A) = \sum_{j=1}^{3} a_{1j}C_{1j} = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13}$$

$$= 4C_{11} + C_{12} + 9C_{13}$$

$$= 20 - 7 + 36 = 49$$

Theorem 5.

Suppose $A = [a_{ij}]$ is an $n \times n$ matrix with a_{ij} in the i^{th} row and j^{th} column of A. Then

(i) $\det A = \sum_{j=1}^{n} a_{ij}C_{ij}$, for any i satisfying $1 \leq i \leq n$ (this is called expansion along the i^{th} row).

Also, $\det A = \sum_{i=1}^{n} a_{ij}C_{ij}$, for any j satisfying $1 \leq j \leq n$ (this is called expansion along the j^{th} column).

(ii) $\det A = \det A^T$

(iii) If A is upper triangular (i.e. $a_{ij} = 0$ whenever $i > j$), then $\det A = a_{11}a_{22}a_{33} \ldots a_{nn}$.

If A is lower triangular (i.e. $a_{ij} = 0$ whenever $i < j$), then $\det A = a_{11}a_{22}a_{33} \ldots a_{nn}$.

(iv) If B is the matrix obtained from A by multiplying one row of A by a scalar s, then $\det B = s \det A$.

(v) If B is the matrix obtained from A by swapping two rows of A, then $\det B = -\det A$.

(vi) If two rows of A are identical, then $\det A = 0$.

(vii) If A has a zero row, then $\det A = 0$

(viii) If B is the matrix obtained from A by adding s times one row of A to another row of A, then $\det B = \det A$.

(ix) Properties (iv), (v), (vi), (vii) and (viii) remain valid if the word, row, is replaced everywhere by the word, column,

(x) $\det(AB) = \det A \det B$, where B is any $n \times n$ matrix. This is a very useful property of determinants.

MT251P – Lecture 18

Example 4.

Suppose $A = \begin{pmatrix} 2 & 9 & 3 \\ 0 & 1 & 4 \\ 0 & 4 & -\frac{1}{2} \end{pmatrix}$

Find $\det A$.

Solution.

Use (i) in Theorem 5 above and expand along the first column of A to get

$$\det(A) = a_{11}C_{11} + a_{21}C_{21} + a_{31}C_{31}$$

$$= a_{11}C_{11}$$

$$= 2\det\begin{pmatrix} 1 & 4 \\ 4 & -\frac{1}{2} \end{pmatrix}$$

$$= 2(-\frac{33}{2})$$

$$= -33$$

Theorem 6

Suppose A is an $n \times n$ matrix. Then, A is invertible $\iff$ $\det A \neq 0$.

Proof

We will prove $\Rightarrow$.

$$AA^{-1} = I_n$$

$$\Rightarrow \det(AA^{-1}) = \det I_n = 1$$

$$\Rightarrow \det A \det(A^{-1}) = 1$$

$$\Rightarrow \det A \neq 0$$

Example 5.

Suppose $A = \begin{pmatrix} 2 & 0 & 9 & 3 \\ 9 & 4 & 4 & 4 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 4 & -\frac{1}{2} \end{pmatrix}$

Find $\det A$ and use it to show A is invertible.

Solution.

Expand along the second column of A to get

$$\det A = 4 \det \begin{pmatrix} 2 & 9 & 3 \\ 0 & 1 & 4 \\ 0 & 4 & -\frac{1}{2} \end{pmatrix}$$

$$= 4(-33) \quad \text{by example 4}$$

$$= -132$$

So, $\det A \neq 0$ and hence A is invertible.

Example 6.

Suppose $B = \begin{pmatrix} 5 & 2 & 3 & 0 \\ 9 & -2 & 4 & 0 \\ 1 & 4 & 5 & -2 \\ 0 & 3 & 0 & 0 \end{pmatrix}$

Find $\det(B^2)$.

Solution.

Expand along the fourth row of B to get

$$\det B = 3 \det \begin{pmatrix} 5 & 3 & 0 \\ 9 & 4 & 0 \\ 1 & 5 & -2 \end{pmatrix}$$

$$= 3 \left(-2 \det \begin{pmatrix} 5 & 3 \\ 9 & 4 \end{pmatrix} \right)$$

$$= -6(-7) = 42$$

So, by (x) in Theorem 5, we get $\det(B^2) = (\det B)^2 = (42)^2 = 1764$.

Remark 4.

If A is an invertible $n \times n$ matrix, then the proof in Theorem 6 shows that

$$\det(A^{-1}) = \frac{1}{\det A}$$

Example 7.

Suppose $A = \begin{pmatrix} 3 & 0 & 0 \\ 4 & 2 & 0 \\ 5 & 5 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 2 & 0 & 8 \\ 2 & 0 & 2 \\ 2 & 3 & 2 \end{pmatrix}$

Find $\det(A^{-1}B)$.

Solution.

$$\det(A^{-1}B) = \det(A^{-1}) \det B$$

$$\frac{\det B}{\det A}$$

Now, A is lower triangular and so $\det A = 3(2)(1) = 6$. Also, we can find $\det B$ by expanding along the second column to get

$$\det B = -3 \det \begin{pmatrix} 2 & 8 \\ 2 & 2 \end{pmatrix}$$

$$= -3(4 - 16) = 36$$

So,

$$\det(A^{-1}B) = \frac{36}{6} = 6$$

Theorem 7.

Suppose A is an $n \times n$ matrix and $b_1, b_2, \ldots, b_n$ are real numbers. Then, the system of linear equations

$$A \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_n \end{pmatrix}$$

has a unique solution $\iff$ $\det A \neq 0$.

Example 8.

Find the values of t for which

$$A_t = \begin{pmatrix} t-1 & 3 & -1 \\ 0 & t-2 & 4 \\ 0 & 0 & t+2 \end{pmatrix}$$

is not invertible.

Solution.

$\det A_t = (t-1)(t-2)(t+2)$ because A_t is upper triangular.

So, $\det A_t = 0 \iff t = -2, 1, 2$. Hence, A_t is not invertible exactly when $t = -2, 1, 2$.

MT251P – Lecture 19

Chapter 6 – Subspaces, Dimension and Rank.

Section 6.1 – Linear Independence and Basis.

Definition 1.

The vectors $\underline{u_1}, \underline{u_2}, \ldots, \underline{u_k} \in \mathbb{R}^n$ are called linearly independent if the only $\alpha_i \in \mathbb{R}$ for $1 \leq i \leq k$ that satisfy

$$\alpha_1 \underline{u_1} + \alpha_2 \underline{u_2} + \cdots + \alpha_k \underline{u_k} = \underline{0}$$

are $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$

Definition 2.

If the vectors $\underline{u_1}, \underline{u_2}, \ldots, \underline{u_k} \in \mathbb{R}^n$ are not linearly independent, then we call them linearly dependent.

Example 1.

(a) $(1, 0, 0), (0, 2, 0), (0, 0, -1)$ are linearly independent in $\mathbb{R}^3$ since

$$\alpha_1(1, 0, 0) + \alpha_2(0, 2, 0) + \alpha_3(0, 0, -1) = (0, 0, 0)$$

$$\iff (\alpha_1, 2\alpha_2, -\alpha_3) = (0, 0, 0)$$

$$\iff \alpha_1 = \alpha_2 = \alpha_3 = 0$$

(b) $(2, 0, 0), (0, 2, 0), (2, 4, 0)$ are linearly dependent in $\mathbb{R}^3$ since

$$(2, 0, 0) + 2(0, 2, 0) + (-1)(2, 4, 0) = (0, 0, 0)$$

Example 2.

Are $(0, 2, 1), (-3, 1, 1), (-6, 6, 4)$ linearly independent in $\mathbb{R}^3$?

Solution.

$$\alpha_1(0, 2, 1) + \alpha_2(-3, 1, 1) + \alpha_3(-6, 6, 4) = (0, 0, 0)$$

$$\iff -3\alpha_2 - 6\alpha_3 = 0, \quad 2\alpha_1 + \alpha_2 + 6\alpha_3 = 0, \quad \alpha_1 + \alpha_2 + 4\alpha_3 = 0$$

This is a system of three linear equations in $\alpha_1, \alpha_2, \alpha_3$. One can show there are infinitely many solutions for $\alpha_1, \alpha_2, \alpha_3$ and so $(0, 2, 1), (-3, 1, 1), (-6, 6, 4)$ are linearly dependent.

Definition 3.

A non–empty subset S of $\mathbb{R}^n$ is called a subspace of $\mathbb{R}^n$ if the following two conditions are satisfied:

(i) $\underline{x} + \underline{y} \in S, \quad$ for all $\quad \underline{x}, \underline{y} \in S$

(ii) $\alpha\underline{x} \in S, \quad$ for all $\quad \alpha \in \mathbb{R}, \quad \underline{x} \in S$

Example 3.

Consider $S = \{(0,t) : t \text{ is a prime number}\}$ as a subset of $\mathbb{R}^2$. Is S a subspace of $\mathbb{R}^2$?

Solution.

No, because $(0,3), (0,5) \in S$ but $(0,3) + (0,5) = (0,8) \notin S$.

Example 4.

Consider $S = \{(t,0,0) : t \text{ is an even integer}\}$ as a subset of $\mathbb{R}^3$. Is S a subspace of $\mathbb{R}^3$?

Solution.

No, because $\frac{1}{2}(2,0,0) \notin S$.

Example 5.

Consider $S = \{(0,0,0,0)\}$ as a subset of $\mathbb{R}^4$. Is S a subspace of $\mathbb{R}^4$?

Solution.

Yes, because the two conditions in definition 3 are satisfied.

Definition 4.

The vectors $\underline{u_1}, \underline{u_2}, \ldots, \underline{u_k}$ span (or generate) a subspace S of $\mathbb{R}^n$ if every $\underline{x} \in S$ can be written in the form:

$$\underline{x} = \alpha_1\underline{u_1} + \alpha_2\underline{u_2} + \cdots + \alpha_k\underline{u_k}, \quad \text{for some} \quad \alpha_i \in \mathbb{R}, \quad 1 \le i \le k$$

In this case we call $\{\underline{u_1}, \underline{u_2}, \ldots, \underline{u_k}\}$ a spanning set of S.

Definition 5.

Suppose $\underline{u_1}, \underline{u_2}, \ldots, \underline{u_k}$ are vectors in $\mathbb{R}^n$. A linear combination of $\underline{u_1}, \underline{u_2}, \ldots, \underline{u_k}$ is an expression of the form:

$$\alpha_1\underline{u_1} + \alpha_2\underline{u_2} + \cdots + \alpha_k\underline{u_k}, \quad \text{where} \quad \alpha_i \in \mathbb{R}, \quad 1 \le i \le k$$

MT251P – Lecture 20

Example 6.

(i) $(4,-4,0)$ is a linear combination of $(2,0,0), (\frac{1}{2},-1,0)$ because

$$(4,-4,0) = (2,0,0) + 4(\frac{1}{2},-1,0)$$

(ii) $(4,-2,1)$ is not a linear combination of $(2,0,0), (\frac{1}{2},-1,0)$. because there are no $\alpha_1, \alpha_2 \in \mathbb{R}$ such that

$$(4,-2,1) = \alpha_1(2,0,0) + \alpha_2(\frac{1}{2},-1,0)$$

Example 7.

$S = \{(a,b,0) : a,b \in \mathbb{R}\}$ is subspace of $\mathbb{R}^3$.

Definition 6.

Suppose $\underline{u_1}, \underline{u_2}, \ldots, \underline{u_k}$ are vectors in $\mathbb{R}^n$. The set

$$S(\underline{u_1}, \underline{u_2}, \ldots, \underline{u_k}) = \{\alpha_1\underline{u_1} + \alpha_2\underline{u_2} + \cdots + \alpha_k\underline{u_k},\ \alpha_i \in \mathbb{R},\quad 1 \le i \le k\}$$

consisting of all linear combinations of $\underline{u_1}, \underline{u_2}, \ldots, \underline{u_k}$ is called the subspace generated by $\underline{u_1}, \underline{u_2}, \ldots, \underline{u_k}$.

Remark 1.

One can check that $S(\underline{u_1}, \underline{u_2}, \ldots, \underline{u_k})$ is a subspace of $\mathbb{R}^n$.

Definition 7.

Suppose S is a subspace of $\mathbb{R}^n$. The set $\{\underline{u_1}, \underline{u_2}, \ldots, \underline{u_k}\}$ is a basis for S if the following two conditions are satisfied:

(i) $\underline{u_1}, \underline{u_2}, \ldots, \underline{u_k}$ are linearly independent.

(ii) $\underline{u_1}, \underline{u_2}, \ldots, \underline{u_k}$ span S.

Example 8.

Suppose $\underline{e_i}$ is the vector in $\mathbb{R}^4$ with zero in every coordinate except the i^{th} coordinate where there is a one. Then $B = \{\underline{e_1}, \underline{e_2}, \underline{e_3}, \underline{e_4}\}$ is a basis for $\mathbb{R}^4$.

Proof.

We first prove that B is linearly independent. Suppose $\alpha_i \in \mathbb{R},\ 1 \le i \le 4$ and

$$\alpha_1\underline{e_1} + \alpha_2\underline{e_2} + \alpha_3\underline{e_3} + \alpha_4\underline{e_4} = (0,0,0,0)$$

Thus,

$$\alpha_1(1,0,0,0)+\alpha_2(0,1,0,0)+\alpha_3(0,0,1,0)+\alpha_4(0,0,0,1)=(0,0,0,0)$$

$$\iff (\alpha_1,\alpha_2,\alpha_3,\alpha_4)=(0,0,0,0)$$

$$\iff \alpha_1=\alpha_2=\alpha_3=\alpha_4=0$$

and so B is linearly independent.

We next prove that B spans $\mathbb{R}^4$. So, suppose $\underline{x}=(x_1,x_2,x_3,x_4)\in\mathbb{R}^4$. Then, we can write

$$\underline{x}=x_1\underline{e}_1+x_2\underline{e}_2+x_3\underline{e}_3+x_4\underline{e}_4$$

.

and so $\underline{x}$ is a linear combination of $\underline{e}_1,\underline{e}_2,\underline{e}_3,\underline{e}_4$ and hence B spans $\mathbb{R}^4$.

So, B is basis for $\mathbb{R}^4$.

Example 9.

Suppose $\underline{e_i}$ is the vector in $\mathbb{R}^n$ with zero in every coordinate except the i^{th} coordinate where there is a one. Then $B=\{\underline{e_1},\underline{e_2},\ldots,\underline{e_n}\}$ is a basis for $\mathbb{R}^n$. The proof of this is similar to example 8. Also, this basis is called the standard basis for $\mathbb{R}^n$.

Theorem 1.

Suppose $\underline{y_1},\underline{y_2},\ldots,\underline{y_w}$ are vectors in $S(\underline{u_1},\underline{u_2},\ldots,\underline{u_k})$. If $w>k$, then $\underline{y_1},\underline{y_2},\ldots,\underline{y_w}$ are linearly dependent.

Remark 2.

If $\underline{u_1},\underline{u_2},\ldots,\underline{u_k}$ is a basis for a subspace T of $\mathbb{R}^n$, then $T=S(\underline{u_1},\underline{u_2},\ldots,\underline{u_k})$.

Theorem 2.

Suppose $\{\underline{u_1},\underline{u_2},\ldots,\underline{u_k}\}$ and $\{\underline{y_1},\underline{y_2},\ldots,\underline{y_r}\}$ are two bases for a subspace T of $\mathbb{R}^n$. Then $k=r$. The unique number of vectors in a basis for T is called the dimension of T and is denoted by $\dim T$.

Example 10.

$\mathbb{R}^2$ has dimension 2 because $\{(1,0),(0,1)\}$ is a basis for $\mathbb{R}^2$. Also, $\mathbb{R}^n$ has dimension n because $\{\underline{e_1},\underline{e_2},\ldots,\underline{e_n}\}$ is a basis for $\mathbb{R}^n$.

Theorem 3.

Suppose S is a subspace of $\mathbb{R}^n$ and suppose that C is a subset of S such that the number of elements in C is $\dim S$. Then C is linearly independent $\iff$ C spans S.

Remark 3.

Theorem 3 shows that if the number of elements in C is $\dim S$, then we only have to check one of the two conditions (i), (ii) in definition 7 in order to check if C is a basis for S.

Example 11

Prove that $C = \{(-1, 0, 3), (0, 2, -2), (0, 0, 5)\}$ is a basis for $\mathbb{R}^3$.

Proof.

First, note that $\mathbb{R}^3$ is a subspace of $\mathbb{R}^3$ and that C is a subset of $\mathbb{R}^3$ such that the number of elements in C is $\dim \mathbb{R}^3 = 3$. Then, remark 3 shows that if we can show C is linearly independent, then C will be a basis for $\mathbb{R}^3$.

So, we will prove that C is linearly independent and then we will be done. So, suppose

$$\alpha_1(-1, 0, 3) + \alpha_2(0, 2, -2) + \alpha_3(0, 0, 5) = (0, 0, 0) \qquad (*)$$

for some $\alpha_i \in \mathbb{R}$ for $1 \leq i \leq k$

Then,

$$(*) \iff (-\alpha_1, 2\alpha_2, 3\alpha_1 - 2\alpha_2 + 5\alpha_3) = (0, 0, 0)$$

$$\iff \alpha_1 = 0,\ 2\alpha_2 = 0,\ 3\alpha_1 - 2\alpha_2 + 5\alpha_3 = 0$$

$$\iff \alpha_1 = 0,\ \alpha_2 = 0,\ \alpha_3 = 0$$

Hence C is linearly independent and we are done.

Section 6.2 – Row Space, Column Space and Rank.

Remark 4.

(i) Suppose A is a $k \times n$ matrix. We can consider each row of A as a vector in $\mathbb{R}^n$ (called a row vector). Similarly, we can consider each column of A as a vector in $\mathbb{R}^k$ (called a column vector).

(ii) If W is a spanning set for L, then we say that L is spanned by W.

MT251P – Lecture 21

Fiacre Ó Cairbre

Definition 8.

(i) The row space of a $k \times n$ matrix A is the subspace $R(A)$ of $\mathbb{R}^n$ that is spanned by the k row vectors of A. The row rank of A is defined to be $\dim R(A)$.

(ii) The column space of a $k \times n$ matrix A is the subspace $C(A)$ of $\mathbb{R}^k$ that is spanned by the n column vectors of A. The column rank of A is defined to be $\dim C(A)$.

Remark 5.

A $k \times n$ matrix A can be considered as a function from $\mathbb{R}^n$ to $\mathbb{R}^k$ as follows:

Suppose

$$\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{pmatrix}$$

is a column vector in $\mathbb{R}^n$. Then,

$$\underline{y} = A\underline{x} \qquad (*)$$

will give a column vector in $\mathbb{R}^k$.

In $(*)$ above $A\underline{x}$ means the product of the $k \times n$ matrix A by the $n \times 1$ matrix $\underline{x}$.

Example 12.

Consider the 2×3 matrix

$$A = \begin{pmatrix} 2 & -1 & 3 \\ 4 & 1 & 2 \end{pmatrix}$$

and suppose

$$\underline{x} = \begin{pmatrix} 2 \\ -2 \\ 5 \end{pmatrix} \in \mathbb{R}^3$$

Then,

$$A\underline{x} = \begin{pmatrix} 2 & -1 & 3 \\ 4 & 1 & 2 \end{pmatrix} \begin{pmatrix} 2 \\ -2 \\ 5 \end{pmatrix} = \begin{pmatrix} 21 \\ 16 \end{pmatrix} \in \mathbb{R}^2$$

Definition 9.

Suppose A is a $k \times n$ matrix. Then, the image space of A is denoted by $Im(A)$ and is defined as:

$$Im(A) = \{A\underline{x} : \underline{x} \in \mathbb{R}^n\}$$

Remark 6.

In definition 9, suppose G_i is the i^{th} column of A considered as a column vector in $\mathbb{R}^k$ for $1 \leq i \leq n$ and suppose

$$\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

Then, $A\underline{x} = x_1 G_1 + x_2 G_2 + \cdots + x_n G_n$ and so $Im(A)$ is the set of all linear combinations of the n column vectors of A. Hence, $Im(A) = C(A)$.

Definition 10.

Suppose A is $k \times n$ matrix. Then, the rank of A is denoted by rank (A) and is defined to be $\dim Im(A)$ which is also $\dim C(A)$.

Definition 11.

Suppose A is $k \times n$ matrix. Then, the kernel of A is denoted by $\ker(A)$ and is defined as:

$$\ker(A) = \{\underline{x} \in \mathbb{R}^n : A\underline{x} = 0\}$$

Definition 12.

One can show that $\ker(A)$ is a subspace of $\mathbb{R}^n$. The nullity of A is then defined to be $\dim \ker(A)$.

Theorem 4 – Rank–Nullity Theorem.

Suppose A is $k \times n$ matrix. Then

$$\dim \ker A + \dim ImA = n$$

i.e. rank $A+$ nullity of $A = n$.

Theorem 5.

Suppose A is a $k \times n$ matrix. Then, the row rank of A is equal to the column rank of A. Also, the row rank of A is equal to the rank of A

Remark 7 – How do we find the rank of a matrix?

Suppose A is a $k \times n$ matrix. Here is the strategy for finding rank A: You first perform elementary row operations on A and stop when you have an REF matrix C. Then, rank A is the number of non–zero rows in C.

Example 13.

Find the rank of the matrix

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

Solution.

Interchange R_1 with R_2 to get

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

Replace R_2 with $R_2 - 2R_1$ and replace R_3 with $R_3 - 5R_1$ to get

$$\begin{pmatrix} 1 & -2 & -3 \\ 0 & 3 & 9 \\ 0 & 6 & 18 \end{pmatrix}$$

Replace R_3 with $R_3 - 2R_2$ to get

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

Replace R_2 with $\frac{1}{3}R_2$ to get

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

C is in REF and so the rank of A is the number of non–zero rows in C. Hence, rank $A = 2$.

Theorem 6.

Suppose A is an $n \times n$ matrix. Then, A is invertible $\iff$ rank $A = n$.

Example 14

Is $A = \begin{pmatrix} 2 & -1 & 3 \\ 1 & -2 & -3 \\ 5 & -4 & 3 \end{pmatrix}$ invertible?

Solution.

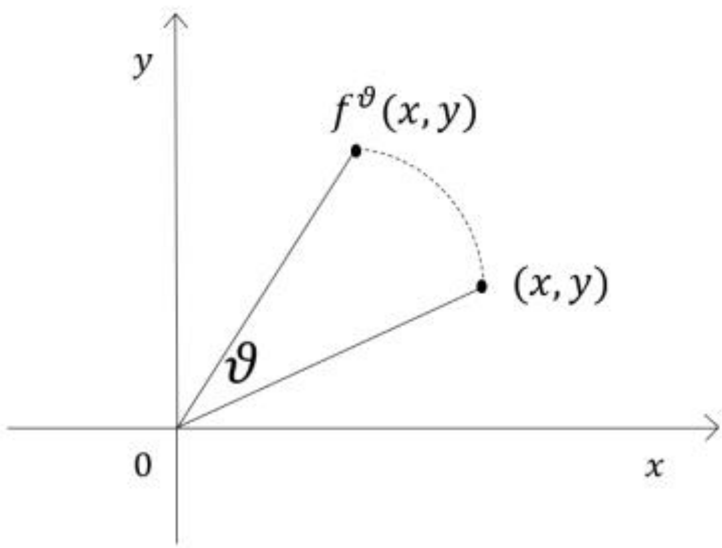
Note that from example 13 we have that rank $A = 2$ and so theorem 6 implies that A is not invertible.

Chapter 7 – Applications of Matrices to Geometry.

Section 7.1 – Rotations.

Remark 1.

Rotations appear in many important applications of mathematics. Consider the usual xy plane denoted by $\mathbb{R}^2 = \{(x, y) : x, y \in \mathbb{R}\}$ and consider the anti–clockwise rotation about the origin $(0, 0)$ through an angle $\theta \in [0, 2\pi)$. Denote this rotation by f^θ.

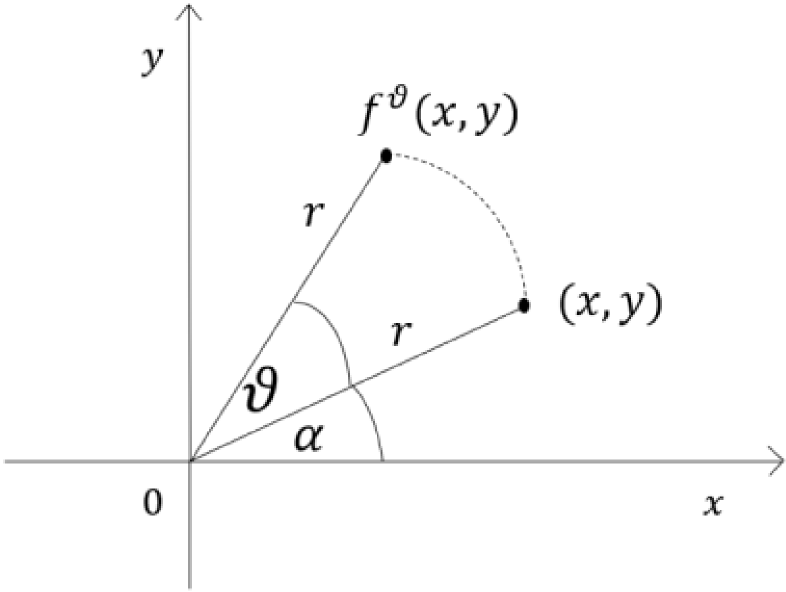


Theorem 1.

Using the notation in remark 1, we have that

$$f^\theta(x, y) = (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta), \quad \text{for all} \quad (x, y) \in \mathbb{R}^2$$

Proof.



Let α be the angle that the line joining (x, y) to the origin, makes with the positive x-axis. Let r be the distance from (x, y) to the origin. Then,

$$x = r\cos\alpha \qquad \text{and} \qquad y = r\sin\alpha$$

Now,

$$f^{\theta}(x, y) = (a, b), \qquad \text{where} \quad a = r\cos(\alpha + \theta),\ b = r\sin(\alpha + \theta)$$

So,

$$a = r(\cos\alpha\cos\theta - \sin\alpha\sin\theta)$$

$$= x\cos\theta - y\sin\theta$$

and

$$b = r(\sin\alpha\cos\theta + \sin\theta\cos\alpha)$$

$$= y\cos\theta + x\sin\theta$$

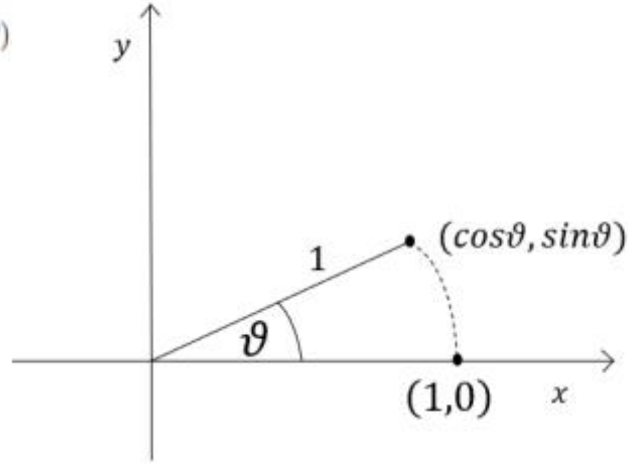
and so

$$f^{\theta}(x, y) = (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta)$$

and we are done.

Example 1.

$f^{\theta}(1,0) = (\cos\theta, \sin\theta)$



Remark 2.

Consider the following 2×2 matrix

$$A_{\theta} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Notice that

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix}$$

and so

$$A_{\theta} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix}$$

So, the rotation f^{θ} corresponds to multiplication on the left by the matrix A_{θ}. This shows how matrices arise naturally in rotations and it's an example of an application of matrices to geometry.

Remark 3.

Note that in theorem 1 we write the elements of $\mathbb{R}^2$ as row (or horizontal) vectors and in remark 2 we write the elements of $\mathbb{R}^2$ as column (or vertical) vectors.

Remark 4.

Many important problems in science, engineering, computer animation, special effects in movies, space navigation etc. involve rotations.

Example 2.

Suppose $\theta = \frac{\pi}{2}$. Then

$$A_\theta = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

So, the anti–clockwise rotation about the origin $(0,0)$ through an angle $\frac{\pi}{2}$ corresponds to multiplication on the left by the matrix

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

Example 3.

Suppose $\theta = \frac{\pi}{4}$. Then

$$A_\theta = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

So, the anti–clockwise rotation about the origin $(0,0)$ through an angle $\frac{\pi}{4}$ corresponds to multiplication on the left by the matrix

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$