University of Warwick
Department of Computer Science

CS131

Mathematics for Computer Scientists II

Cem Yilmaz

March 8, 2022

Contents

1 Number System

1.1 Converting to base n

We can utilise the division algorithm to achieve this. That is, for some base n to convert from base 10 we divide by n to get remainders.

Example 1.1. Division of binary

$$19 \div 2 = 9R1 \tag{1}$$
$$9 \div 2 = 4R1 \tag{2}$$
$$4 \div 2 = 2R0 \tag{3}$$
$$2 \div 2 = 1R0 \tag{4}$$
$$1 \div 2 = 0R1 \tag{5}$$

1.2 The division algorithm

Theorem 1.1. *The division algorithm*

Given any integers $a, b \in \mathbb{Z}$ and $b \neq 0$, there are unique integers $q, r \in \mathbb{Z}$ such that $a = qb + r$ and $0 \leq r < |b|$.

1.3 The Euclidean algorithm

The euclidean algorithm utilises the division algorithm to find $gcd(m, n) = b$ where $m, n, b \in \mathbb{Z}$.

Definition 1.1. Greatest Common Divisor

The greatest common divisors of two numbers m, n where $m, n \in \mathbb{Z}$ is the greatest number ζ such that $\zeta \mid m$ and $\zeta \mid n$. It is denoted as $gcd(m, n)$.

Then, through division, observe that $n = mb + r$ In particular, the key observation would be $gcd(r, m) = gcd(n, m) = b$. Repeat this process until one of the numbers reaches 0.

1.4 Modular Arithmetic

Modular arithmetic ensures that the two numbers have the same remainder.

Example 1.2. Modular Arithmetic

The numbers 19 and 21 are congruent to modulo 2. That is, they both have remainder 1.

$$19 \equiv 21 \bmod 2 \tag{6}$$

The notation $a \bmod n$ can also be used as a notation to denote the remainder of the integer a. Furthermore, the modular arithmetic can be subtracted, added and multiplied as usual. In particular,

Example 1.3. Properties of Modular Arithmetic

Consider the following examples

$$x \equiv 3 \bmod n \tag{7}$$
$$y \equiv 5 \bmod n \tag{8}$$
$$x + y \equiv 8 \bmod n \tag{9}$$
$$x \times y \equiv 15 \bmod n \tag{10}$$

Because of these properties, indeed

Corollary. *Power of modular arithmetic*

Then, from multiplication property, the following holds true. If for some integers x, y the property $x \equiv y \bmod n$ holds, then

$$x^k \equiv y^k \bmod n \tag{11}$$

also holds.

1.4.1 Use of Modular Arithmetic

Let N represent the number of bits used in a system. Then, there are 2^N bits of string length N can be used to represent the numbers in the integer range $[-2^{N-1}, 2^{N-1} - 1]$. In modular arithmetic, the integer x can be then represented as $x \bmod 2^N$. This is two's complement. For example,

Example 1.4. Two's Complement

Table 1: Two's Complement

x	$x \bmod 2^3$	String	x	$x \bmod 2^3$	String
0	0	000	-4	4	100
1	1	001	-3	5	101
2	2	010	-2	6	110
3	3	011	-1	7	111

1.5 Real Numbers

Real numbers consist of every possible numbers that are not complex. There are an infinite amount of real numbers in the interval $[0, 1]$. See CS 130 for this.

1.6 Rational Numbers

A rational number has the form $\frac{m}{n}$ where $m, n \in \mathbb{Z}$ and $n \neq 0$. We can always choose m and n s.t. $n \geq 1$ and $gd(m, n) = 1$.

1.7 Irrational Numbers

An algebraic number is a real number such as $\sqrt{2}$ and $-\sqrt{2}$. It is a solution of a polynomial equation with rational coefficients

Definition 1.2. Transcendental numbers

Transcendental numbers are real numbers which cannot be solutions of polynomial equations with rational coefficients. Examples include π and e.

1.8 Complex Numbers

Definition 1.3. Complex Number

The complex number i is defined as

$$i = \sqrt{-1} \tag{12}$$

And helps us expand our artillery in mathematics for further algebra. Complex numbers are usually denoted in the form

$$z = a + bi$$

where $a, b \in \mathbb{R}$. Furthermore, in the notation above, we follow the general notation

$$Im(z) = a$$
$$Re(z) = b$$

The set of complex numbers is denoted as $\mathbb{C}$

1.8.1 Argrand Diagram

The Argrand diagram is a plane of complex numbers where y axis is the $Im(z)$ and x axis is the $Re(z)$. In particular,

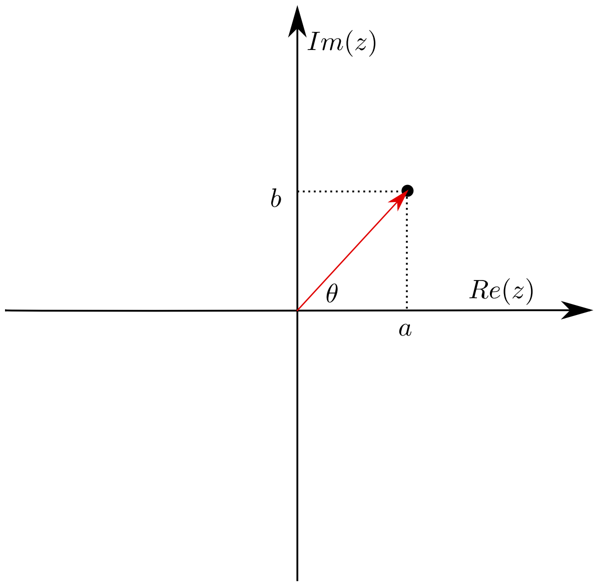

Figure 1: Argrand diagram of $a + bi$

Definition 1.4. Comlex Conjugate

The complex conjugate of $z = a + bi$ is denoted as $\overline{z}$ and is defined as

$$\overline{z} = a - bi \tag{13}$$

Another notation for this is z^*.

Definition 1.5. Modulus of Complex Numbers

The modulus of complex nmber $|z|$ is defined as

$$|z| = \sqrt{a^2 + b^2} \tag{14}$$

This shows the length of the red line in figure 1.

1.8.2 The Triangle Inequaliy

The triangle inequality is given as

$$|z + w| \leq |z| + |w|$$

A corollary of this is

$$||z||w|| \geq |z - w|$$

That is, if we consider

$$\begin{aligned} &|z| \leq |w| + |z - w| \\ \implies &|z| - |w| \leq |z - w| \end{aligned}$$

We similarly obtain for w

$$\begin{aligned} &|w| \leq |z| + |w - z| \\ \implies &|w| - |z| \leq |w - z| \end{aligned}$$

However

$$|w - z| = |z - w|$$

and therefore

$$||z| - |w|| \leq |z - w|$$

1.8.3 Polar representation

The polar representation of the number $a + bi$ can be written as $r(\cos\theta + i\sin\theta)$, where r is the modulus. In particular, $a = r\cos\theta$ and $b = \sin\theta$.

Theorem 1.2. *De Moivre's Theorem*

De Moivre'sT Theorem states that from the consequence of what is below

$$cis\phi \times cis\theta = cis\phi + \theta \tag{15}$$

We obtain the theorem

$$(cis\theta)^n = cisn\theta \tag{16}$$

Proof. The proof of this theorem is left as an exercise to the reader.
Hint: Begin with induction □

1.8.4 Fundamental Theorem of Algebra

Theorem 1.3. *Fundamental Theorem of Algebra*

Every polynomial of degree n with complex coefficients has exactly n not necessarily distinct solutions in $\mathbb{C}$.

And indeed, we can use this theorem for quadratics. In particular, recall that for the quadratic

$$az^2 + bz + c = 0$$

We have discriminant Δ. That is,

$$\Delta = b^2 - 4ac$$

And if $\Delta \geq 0$, we have real solution(s). If we have $\Delta < 0$, then we have complex solutions.

2 Axioms

2.1 Algebraic Axioms

Axiom 1. Commutativity

It follows that

$$x + y = y + x \wedge x \times y = y \times x \tag{17}$$

Axiom 2. Associativity

It follows that

$$x + (y + z) = (x + y) + z \wedge x \times (y \times z) = (x \times y) \times z \tag{18}$$

Axiom 3. Distrubitivity of $\times$ over +

It follows that

$$x \times (y + z) = x \times y + x \times z \tag{19}$$

Axiom 4. Additive Identity

$\exists x . y + x = y$

$$\text{In particular, } x = 0 \tag{20}$$

Axiom 5. Multiplicative Identity

$\exists x . yx = y$

$$\text{In particular, } x = 1 \tag{21}$$

Axiom 6. Distinction

Multiplicative and additive identities are distinct. That is,

$$1 \neq 0 \tag{22}$$

So far, all the above axioms hold for $\mathbb{N}$. However, once we add the following axiom:

Axiom 7. Additive Inverse

$\exists - x . x + (-x) = 0$

So far, all above axioms hold for $\mathbb{Z}$. However, once we add the following axiom:

Axiom 8. Multiplicative Inverse

If $x \neq 0$, then $\exists x^{-1}.x \times x^{-1} = 1$

2.2 Ordering Axioms

Axiom 9. Transitivity of ordering

$$x < y \wedge y < z \implies x < z \tag{23}$$

Axiom 10. The trichotomy law

Exactly one of the following is true:

$$x < y \vee y < x \vee x = y \tag{24}$$

Axiom 11. Preservation of ordering under addition

If $x < y$, then

$$x + z < y + z \tag{25}$$

Axiom 12. Preservation of ordering under multiplication

If $0 < z$ and $x < y$ then

$$x \times z < y \times z \tag{26}$$

So far, all the above axioms hold for $\mathbb{Q}$. However, once we add the following axiom:

Axiom 13. Completeness

Every non-empty subset that is bounded above has a least upper bound.

2.3 Ordering

Definition 2.1. Upper bound

A real number u is an upper bound of S if $u \geq x \ \forall x \in S$

Definition 2.2. Lower bound

A real number u is a lower bound of S if $l \geq x \ \forall x \in S$

Definition 2.3. Supremum

A real number U is supremum of S if U is an upper bound of S and $U \leq u$ for every upper bound u of S. That is, it is the first upper bound.

Definition 2.4. Infimum

A real number L is the infimum of S if L is a lower bound of S and $L \geq l$ for every lower bound l of S. That is, it is the first lower bound.

2.4 Archimedes Property of Real

Theorem 2.1. *Archimedes Property of Reals*

Given any $\varepsilon \in \mathbb{R}^+$, $\exists n \in \mathbb{N}. n\varepsilon > 1$

Proof. Assume $n\varepsilon \leq 1$. Then,

$$\begin{aligned}
& \forall n \text{ that } \{n\varepsilon | n \in \mathbb{N}\} \text{ has an upper bound} && (27)\\
& \text{By completeness it has a least upper bound } l && (28)\\
\implies & \forall n, n\varepsilon \leq l && (29)\\
\implies & (n+1)\varepsilon \leq l && (30)\\
\iff & (n+1)\varepsilon - \varepsilon \leq l - \varepsilon && (31)\\
\iff & n\varepsilon \leq l - \varepsilon && (32)\\
\implies & l - \varepsilon \text{ is also an upper bound} && (33)
\end{aligned}$$

However, this is a contradiction since we already assumed that l is the least upper bound when clearly $l - \varepsilon < l$ □

3 Vectors

3.1 Addition

In 2 dimensional space, if $\underline{a} = (a_1, a_2)$ and $\underline{b} = (b_1, b_2)$ where $a, b \in \mathbb{R}^2$ and $\lambda \in \mathbb{R}$, then

$$\begin{aligned}
\underline{a} + \underline{b} &= (a_1 + b_1, a_2 + b_2)\\
\lambda \underline{a} &= (\lambda a_1, \lambda a_2)
\end{aligned}$$

Similar principle follow in n dimensional spaces.

3.2 Geometric Interpretation

A vector $\underline{p} = (p_1, p_2) \in \mathbb{R}^2$ with notation $\overline{OP}$ is seen as travelling from origin $O = (0, 0)$ to (p_1, p_2).

Definition 3.1. Unit Vector

A vector is called a unit vector if and only if its length is 1.

The distance between two vectors $\underline{a}$ and $\underline{b}$ is given by $|\underline{a} - \underline{b}|$.

Example 3.1. Example 1

Find the unit vector in $\mathbb{R}^2$ which has the same direction as $(2, -1)$.
We begin with finding the magnitude:

$$\sqrt{2^2 + 1^2} = \sqrt{5} \tag{34}$$

Then, we just have to scale it such that

$$(\frac{2}{\sqrt{5}, -\frac{1}{\sqrt{5}}}$$

3.3 Scalar Product

Definition 3.2. Scalar Product

The scalar product, also known as dot product, of vectors $\underline{a} = (a_1, a_2)$ and $\underline{b} = (b_1, b_2)$ in $\mathbb{R}^2$ is defined as follows:

$$a \cdot b = a_1 b_1 + a_2 b_2 = |\underline{a}||\underline{b}| \cos\theta$$

Where θ is the angle between $\underline{a}$ and $\underline{b}$. Note that if $\theta = \frac{\pi}{2}$, then the result is 0.

Proof. Write

$$\cos\theta = \cos(\beta - \alpha) = \cos\beta\cos\alpha + \sin\beta\sin\alpha$$

However, we can rewrite

$$\sin\alpha = \frac{a_2}{|\underline{a}|} \qquad \sin\beta = \frac{b_2}{|\underline{b}|}$$

$$\cos\alpha = \frac{a_1}{|\underline{a}|} \qquad \cos\beta = \frac{b_1}{|\underline{b}|}$$

Intuitively, in a graph:

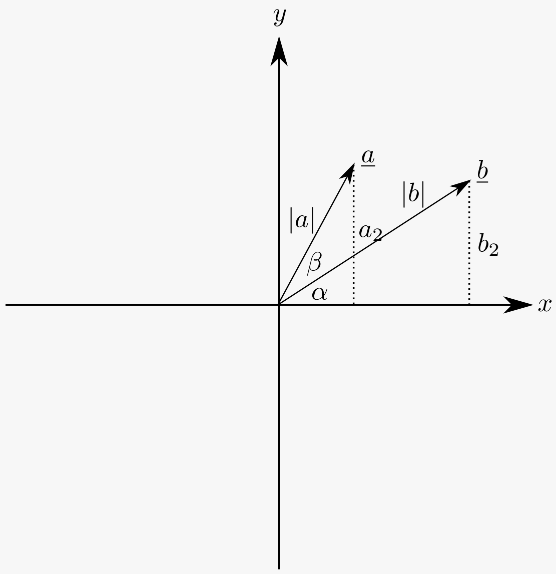

Figure 2: Dot Product Graph

□

Indeed, such rules can be generalised for n dimensions, if one continues to follow the patterns.

3.4 Linear Combination

Definition 3.3. Linear Combination

If $\underline{u}, \underline{v} \in \mathbb{R}^2$ and $\alpha, \beta \in \mathbb{R}$, then a vector of the form is

$$\alpha\underline{u} + \beta\underline{v} \tag{35}$$

is called a linear combination of $\underline{u}$ and $\underline{v}$

Example 3.2. Example 1: Linear Combination

In $\mathbb{R}^2$ the vector $(5,3)$ can be written in the form

$$(5,3) = 5(1,0) + 3(0,1)$$

Indeed, many combinations can exist.

The linear combinations can be extended to multiple dimensions. In fact, given non parallel vectors $\underline{u}, \underline{v} \in \mathbb{R}^2$, the vector $\alpha\underline{u} + \beta\underline{v}$ represents the diagonal of a parallelogram with sides $\alpha\underline{u}$ and $\beta\underline{v}$.

3.5 Span

Definition 3.4. Span

If $U = \{\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_m\}$ is a finite set of vectors in $\mathbb{R}^n$, then the span of U is the set of all linear combinations of $\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_m$

$$\text{Span } U = \{\alpha_1\underline{u}_1 + \alpha_2\underline{u}_2 + \ldots + \alpha_m\underline{u}_m | \alpha_1, \alpha_2, \ldots, \alpha_m \in \mathbb{R}\} \tag{36}$$

3.6 Subspaces

Definition 3.5. Subspaces

A subspace of $\mathbb{R}^n$ is a nonemtpy set S of $\mathbb{R}^n$ with the properties

$$\underline{u}, \underline{v} \in S \implies \underline{u} + \underline{v} \in S \tag{37}$$

$$\underline{u} \in S, \lambda \in \mathbb{R} \implies \lambda\underline{u} \in S \tag{38}$$

That is, there is closure under addition and closer under scalar multiplication.

If S is a subspace and $\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_m \in S$, then any linear combination of $\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_m$ also belongs to S. This follows from under addition and scalar multiplication by induction on m.

3.7 Linear Independence

Definition 3.6. Linear Dependence

Linear independence is when some set S with pairs of numbers have the property that none of the pair of numbers can be expressed as a linear combination of other two. In particular, a set $\{\underline{u}_1, \underline{u}_3, \ldots, \underline{u}_3\}$ of vectors, where

$$\underline{u}_i = (u_{i1}, u_{i2}, \ldots, u_{in}) \tag{39}$$

$\mathbb{R}^n$ is linearly dependent if there are numbers $a_1, a_2, \ldots, a_n \in \mathbb{R}$, not all zero, such that $a_1\underline{u}_1 + a_2\underline{u}_2 + \ldots + a_m\underline{u}_m = \underline{0}$. That is, each ordered pair for the sum of u_i becomes $(0, 0, \ldots, 0) = O$

Definition 3.7. Linear Independence

Linear dependence is when only the solution $\forall a_i = 0$ where $i \in \{1, 2, \ldots, m\}$ is the only solution to $a_1\underline{u}_1 + a_2\underline{u}_2 + \ldots + a_m\underline{u}_m = \underline{0}$.

Theorem 3.1. *Linear Independence*

A set $S = \{\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_m\}$ of nonzero vectors is linearly independent iff some $\underline{u}_r$ is a linear combination of its predecessors $\underline{u}_1, \ldots, \underline{u}_m$

3.8 Basis

Definition 3.8. Basis

Given subspace S of $\mathbb{R}^n$ (recall a non empty subset closed under addition and scalar multiplication), a set of vectors is called a basis of S if it is linearly independent set which spans S.

Example 3.3. Basis of R3

The set

$$S = \{(1,0,0),(0,1,0),(0,0,1)\} \tag{40}$$

is a basis of $\mathbb{R}^3$.

Definition 3.9. Standard Basis

In $\mathbb{R}^n$, the standard basis is the set $\{\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_n\}$ where $\underline{e}_r$ is the vector with r th component 1 and all other components 0. In particular, the example above is a standard basis for $\mathbb{R}^3$.

Theorem 3.2. *At most m vectors*

If the set $\{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_m\}$ spans S, a subspace of $\mathbb{R}^n$, then any linearly independent subset of S contains at most m vectors.

Proof. Suppose $\{\underline{w}_1, \underline{w}_2, \ldots, \underline{w}_p\}$ is a linearly independent subset of S. We wish to show that $p \leq m$.
We can express $\underline{w}_1 \in S$ as a linear combination $\{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_m\}$ since the v span S.
Thus the set $\{\underline{w}_1, \underline{v}_1, \ldots, \underline{v}_m\}$ is linearly dependent. Therefore, by the linear independent theorem, some vector $\underline{v}_i$ is a linear combination of its predecessors $\underline{w}_1, \underline{v}_1, \ldots, \underline{v}_{i-1}$
It follows that $\{\underline{w}_1, \underline{v}_1, \ldots, \underline{v}_{i-1}, \underline{v}_{i+1}, \ldots, \underline{v}_m\}$. I.e, we replace $\underline{v}_i$ with $\underline{w}_1$. We continue to iterate the same steps and thus we can get at most m amount of $\underline{w}_i$. $\square$

3.9 Dimension

Definition 3.10. Dimension

The dimension of a subspace $\mathbb{R}^n$ is the number of vectors in a bases for the subspace. In other words, it is the cardinality of the basis after being reduced.

Example 3.4. Example 1

Consider the set

$$S = \{(x,y,z)|x + 2y - z = 0\} \tag{41}$$

That is a subspace of $\mathbb{R}^3$. Find the basis and dimension of S.
To solve this problem, we can write

$$\begin{aligned} S &= \{(x,y,+2y)|x,y \in \mathbb{R}\} && (42)\\ &= \{x(1,0,1) + y(0,1,2)|x,y \in \mathbb{R}\} && (43)\\ &= span\{(1,0,1),(0,1,2)\} && (44) \end{aligned}$$

Thus, the dimension is 2.

4 Matrices

Definition 4.1. Matrix

A matrix can be seen as a table of information and is denoted as

$$A = [a_{ij}]_{m\times n} \tag{45}$$

shows a matrix of the form

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & \dots & \dots & a_{mn} \end{bmatrix} \tag{46}$$

Definition 4.2. Zero Matrix

The zero matrix, denoted as $O_{m\times n}$ is a $m \times n$ matrix whose all elements are 0.

4.1 Addition

The addition of matrices is defined in an easy manner. For two matrices $A = [a_{ij}]_{m\times n}$ and $B = [b_{ij}]_{m\times n}$, The addition $A + B = [a_{ij} + b_{ij}]_{m\times n}$

4.2 Multiplication

The multiplication of matrices way more complexly defined. For two matrices $A = [ij]_{m\times p}$ and $B = [b_{ij}]_{p\times n}$, the multiplication $A \times B = C$ is defined as

$$c_{ij} = \sum_{r=1}^{p} a_{ir} b_{rj} \tag{47}$$

In other words, the ij th element of AB is the scalar product of the i th row vector of A with the j th column of B.

4.3 Matrix Inverse

Definition 4.3. Inverse

Matrix B is the inverse of matrix A if A and B are square matrices of the same order and if

$$AB = I = BA \tag{48}$$

However, note that not all square matrices have an inverse. We will see how to check whether it exists and compute it using determinants.

There are some rules to the inverse:

- If A has an inverse, then it is unique.
- If A is the inverse of B then B is the inverse of A.
- The inverse is denoted by A^{-1}.

4.4 Matrix Transposition

Definition 4.4. Matrix Transposition

The matrix transposition of some matrix $A = [a_{ij}]_{m\times n}$ denoted as A^T is the matrix $A^T = [a_{ji}]_{n\times m}$

4.5 Identity Matrices

Definition 4.5. Identity Matrices

The identity matrix of diagonal matrices is denoted as I and are square matrices which take the form such as

$$I = [1] \tag{49}$$

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{50}$$

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{51}$$

$$\ldots \tag{52}$$

These have the property that $A \times A^{-1} = I$. In particular, $BI = B$.

4.6 Determinant

Definition 4.6. Determinant

The determinant, denoted as for some matrix $A = [a_{ij}]_{m \times n}$

$$det(A) \tag{53}$$

$$|A| \tag{54}$$

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} \tag{55}$$

Determinants for this module only extended to 2×2 matrices. For 2×2 matrix, it is defined to be

$$ad - bc \tag{56}$$

Corollary. *A matrix is invertible if its determinan is non-zero*

We know that

$$det(A)det(A^{-1}) = det(AA^{-1}) = det(I) = 1$$

Meaning that $det(A) \neq 0$. Indeed, with this, it is easy to verify that a 2×2 matrix A is invertible if and only if its determinant is nonzero.

Let $\alpha = [a_{ij}]_{3\times 3}$ be a matrix. Then the determinant

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

is defined by

$$a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

And expanding these 2×2 matrices obtains us

$$|A| = a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})$$

Indeed, we can do some algebraic manipulation to give it a slightly different definitions:

$$|A| = -a_{21}(a_{12}a_{33} - a_{13}a_{32}) + a_{22}(a_{11}a_{33} - a_{13}a_{31}) - a_{23}(a_{11}a_{32} - a_{12}a_{31})$$
$$|A| = a_{31}(a_{12}a_{23} - a_{22}a_{13}) - a_{32}(a_{11}a_{23} - a_{21}a_{13}) + a_{33}(a_{11}a_{22} - a_{21}a_{12})$$

And with this, notice that the coefficients alternate in these different forms, i.e.

$$\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$$

As such, we can see a pattern:

Definition 4.7. Determinants of n matrices

If $A = [a_{ij}]$ is an $n \times n$ matrix, then the ij th minor M_{ij} of A is defined to be the determinant of the $(n-1) \times (n-1)$ matrix obtained from A by deleting the i th row and j th column. In other words, the ij th cofactor A_{ij} of A is defined by

$$A_{ij} = (-1)^{i+j} M_{ij} \tag{57}$$

For which, then, we can create

$$|A| = a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in} \qquad \text{expansion by } i \text{ th row} \tag{58}$$
$$|A| = a_{1j}A_{1j} + a_{2j}A_{2j} + \ldots + a_{nj}A_{nj} \qquad \text{expansion by the } j \text{ th column} \tag{59}$$

This obtains us the following interesting properties:

Corollary. *Elementary Row Operations*

If B is the matrix obtained from A by

1. Multiplying a row of A by a number λ, then $|B| = \lambda|A|$
2. Interchanging two rows of A, then $|B| = -|A|$
3. Adding a multiple of one row of A to another, then $|B| = |A|$

Proof. Proofs of 1 and 2 are left as an exercise to the reader.
We use the fact that any determinant with two equal rows is 0. This easily seen by number 2, where if we were to swap these rows, we would obtain $|M| = -|M|$ hence $|M| = 0$ Now, we consider adding a multiple to obtain

$$(a_{i1} + \lambda a_{k1})A_{i1} + (a_{i2} + \lambda a_{k2})A_{i2} + \ldots + (a_{in} + \lambda a_{kn})A_{kn} \tag{60}$$
$$\implies |A| + \lambda(a_{k1}A_{i1} + a_{k2}A_{i2} + \ldots + a_{kn}A_{i}n) \tag{61}$$

But the right side is equal to 0 from property 2 as one of the A_{ij} contains a_k already, therefore this is just equal to $|A|$. □

4.7 Linear Equations

You can express linear equations in matrices. For example, if we consider the equations

$$a_{11}x_1 + a_{12}x_2 = b_1$$
$$a_{21}x_1 + a_{22}x_2 = b_2$$

The equivalent matrix form is

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

And to solve such linear equations using row operations we

1. Swap (interchange) two equations
2. Multiply both sides of an equation by a nonzero number
3. Add a multiple of one equation to another equation

Example 4.1. Example

Consider the equation

$$x_1 - x_2 + x_3 = 1 \tag{62}$$
$$x_1 + x_2 + 2x_3 = 0 \tag{63}$$
$$2x_1 - x_2 + 3x_3 = 2 \tag{64}$$

Then, we can create an augment matrix such that

$$\left[\begin{array}{ccc|c} 1 & -1 & 1 & 1 \\ 1 & 1 & 2 & 0 \\ 2 & -1 & 3 & 2 \end{array}\right] \tag{65}$$

4.8 Row Equivalence

Definition 4.8. Row Equivalence

Matrices A and B are row equivalent ($A \sim B$) if A can be transformed to B using a finite (possibly 0) number of elementary row operations.

4.9 Row Echelon form

Definition 4.9. Row Echelon Form

A matrix is in row echelon form if the first nonzero entry in each row is further to the right than the first nonzero entry in the previous row.

4.10 Elementary Matrices

Elementary row operations can be performed by multiplying a matrix on the left by a suitable "elementary matrix". These are matrices that can be used to, for example, interchange rows.

Example 4.2. Elementary Matrix

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} d & e & f \\ a & b & c \\ g & h & i \end{bmatrix} \tag{66}$$

In this case, the left matrix is an elementary matrix.

More generally we define $n \times n$ elementary matrices as follows:

- E_{ij} is obtained from identity matrix I by swapping rows i and j.
- $E_i(\lambda)$ for $\lambda \neq 0$ is obtained from I by multiplying the entries in the i th row by λ.
- $E_{ij}(\mu)$ is obtained from I by adding μ times row j and to row i.

It is easy to verify that every elementary matrix is invertible and that its inverse is in fact another elementary matrix. In fact,

$$E_{ij}^{-1} = E_{ij} \text{ from } E_{ij}E_{ij} = I$$
$$E_i(\lambda) = E_i\left(\frac{1}{\lambda}\right) \text{ since } E_i(\lambda)\, E\left(\frac{1}{\lambda}\right) = I$$
$$E(\mu)E(-\mu) \text{ since } E(-\mu)(\mu) = I$$

Theorem 4.1. *Sequences of elementary rows*

If a sequence of elementary row operation transforms a square matrix A into I, then A is invertible and the same sequence transforms I into A^{-1}.

Proof. Suppose that the row operations applied to A correspond to elementary matrices $E_1, E_2, \dots, E_n$ where E_1 is applied first, E_2 next and so on. Then,

$$E_n E_{n-1} \dots E_2 E_1 A = I$$

Let $E = E_n E_{n-1} \dots E_2 E_1$, then $EA = I$. However, $AE = 1$ as since E is a product of invertible matrices, it can shown to have an inverse E^{-1}.
Now,

$$AE = IAE = (E^{-1}E)AE = E^{-1}(EA)E = E^{-1}IE = E^{-1}E = I$$

Therefore A is invertible and

$$A^{-1} = E = E_n E_{n-1} \dots E_2 E_1 = E_n E_{n-1} \dots E_2 E_1 I.$$

□

4.11 Adjoint Of A Matrix

Definition 4.10. Adjoint of a matrix

The adjoint of a matrix $A = [a_{ij}]_{n\times n}$ is given by cofactors A_{ij} where

$$\begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & & & \vdots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{bmatrix}^T \tag{67}$$

4.12 Finding The Inverse

Example 4.3. Finding inverse

Find the inverse of

$$\begin{bmatrix} 2 & -1 & 4 \\ 4 & 0 & 2 \\ 3 & -2 & 7 \end{bmatrix} \tag{68}$$

We create the formation

$$\left[\begin{array}{ccc|ccc} 2 & -1 & 4 & 1 & 0 & 0 \\ 4 & 0 & 2 & 0 & 1 & 0 \\ 3 & -2 & 7 & 0 & 0 & 1 \end{array}\right] \tag{69}$$

Then, we do algebra. In particular, we try to get the left hand side in identity matrix form. By doing so, we also change the identity matrix on the right, and as such, obtain the inverse. In particular, after the operations

$$\begin{aligned} \frac{1}{2}\text{row1} \implies \text{row2} - 4\text{row1} \implies \text{row3} - 3\text{row1} \implies \text{row}\frac{2}{2} \\ \implies \text{row3} + \frac{1}{2}\text{row2} \implies -2\text{row3} \implies \text{row1} - 2\text{row3} \implies \text{row2} + 3\text{row3} \\ \implies \text{row1} + \frac{1}{2}\text{row2} \end{aligned}$$

Obtains us

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & \frac{1}{2} & 1 \\ 0 & 1 & 0 & 11 & -1 & -6 \\ 0 & 0 & 1 & 4 & -\frac{1}{2} & -2 \end{array}\right] \tag{70}$$

Another way to find the inverse is to use the definitions of our determinants and adjoints. That is,

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

Example 4.4. Finding inverse using adjoint

Find the inverse of the matrix A and hence solve the system of equations

$$A = \begin{bmatrix} 2 & 1 & 4 \\ 1 & 0 & 2 \\ 2 & 3 & 1 \end{bmatrix} \tag{71}$$

$$2x + y + 4z = 2 \tag{72}$$

$$x + 2z = 3 \tag{73}$$

$$2x + 3y + z = -6 \tag{74}$$

The determinant is given by, where we choose the middle row for ease as there exists a 0,

$$-\begin{vmatrix} 1 & 4 \\ 3 & 1 \end{vmatrix} + 0 - 2\begin{vmatrix} 2 & 1 \\ 2 & 3 \end{vmatrix} \tag{75}$$

$$= -(1 - 12) - 2(6 - 2) \tag{76}$$

$$=3 \tag{77}$$

Finally computing the adjoint matrix

$$Adj(A) = \begin{bmatrix} \begin{vmatrix} 0 & 2 \\ 3 & 1 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 0 \\ 2 & 3 \end{vmatrix} \\ -\begin{vmatrix} 1 & 4 \\ 3 & 1 \end{vmatrix} & \begin{vmatrix} 2 & 4 \\ 2 & 1 \end{vmatrix} & -\begin{vmatrix} 2 & 1 \\ 2 & 3 \end{vmatrix} \\ \begin{vmatrix} 1 & 4 \\ 0 & 2 \end{vmatrix} & -\begin{vmatrix} 2 & 4 \\ 1 & 2 \end{vmatrix} & \begin{vmatrix} 2 & 1 \\ 1 & 0 \end{vmatrix} \end{bmatrix}^T \tag{78}$$

$$= \begin{bmatrix} -6 & 3 & 3 \\ 11 & -6 & -4 \\ 2 & 0 & -1 \end{bmatrix}^T \tag{79}$$

$$= \begin{bmatrix} -6 & 11 & 2 \\ 3 & -6 & 0 \\ 3 & -4 & -1 \end{bmatrix} \tag{80}$$

Then, we can write the system of equations as

$$A \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ -6 \end{bmatrix} \tag{81}$$

Multiplying both sides by A^{-1}

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = A^{-1} \begin{bmatrix} 2 \\ 3 \\ -6 \end{bmatrix} \tag{82}$$

$$=\frac{1}{3} \begin{bmatrix} -6 & 11 & 2 \\ 3 & -6 & 0 \\ 3 & -4 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ -6 \end{bmatrix} \tag{83}$$

$$= \begin{bmatrix} 3 \\ -4 \\ 0 \end{bmatrix} \tag{84}$$

4.13 Linear Independence of Matrices

Following the definition of linear independence from 3.7, we can further extend it in the use of matrices. For some matrix $A = [a_{ij}]_{m\times n}$ where

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \tag{85}$$

The column vectors of A is defined as the ordered pairs where

$$(a_{11}, a_{21}, \dots, a_{m1}), (a_{12}, a_{22}, \dots, a_{m2}, \dots, (a_{1n}, a_{2n}, \dots, a_{mn}) \tag{86}$$

Similarly, the column vectors of A is defined as

$$(a_{11}, a_{12}, \dots, a_{1n}), (a_{21}, a_{22}, \dots, a_{2n}), \dots, (a_{m1}, a_{m2}, \dots, a_{mn}) \tag{87}$$

Then, the following theorem holds:

Theorem 4.2. *Linear Independence using Determinant*

A set of n vectors in $\mathbb{R}^n$ is linearly independent and as a corollary a basis if and only if it is the set of column vectors of a matrix with non-zero determinant.

Proof. Consider the set of vectors $\{\underline{u}_1, \underline{u}_2, \dots, \underline{u}_n\}$ and let $\underline{u}_j = (u_{1j}, u_{2j}, \dots, u_{nj})$ for $1 \leq j \leq n$. Now consider the equation what defines dependence/independence:

$$\alpha_1\underline{u}_1 + \alpha_2\underline{u}_2 + \dots + \alpha_n\underline{u}_n = \underline{0} \text{ i.e. the origin } O \text{ vector} \tag{88}$$

If we expand each ordered pair vector we obtain (noting the range of j)

$$\alpha_1 u_{11} + \alpha_2 u_{12} + \dots + \alpha_j u_{1j} + \dots + \alpha_n u_{1n} = 0 \tag{89}$$
$$\alpha_1 u_{21} + \alpha_2 u_{22} + \dots + \alpha_j u_{2j} + \dots + \alpha_n u_{2n} = 0 \tag{90}$$
$$\vdots \tag{91}$$
$$\alpha_1 u_{n1} + \alpha_2 u_{n2} + \dots + \alpha_j u_{nj} + \dots + \alpha_n u_{nn} = 0 \tag{92}$$

Then, using the property of matrix multiplication we can write this as

$$\begin{bmatrix} u_{11} & u_{12} & \dots & u_{1j} & \dots & u_{1n} \\ u_{21} & u_{22} & \dots & u_{2j} & \dots & u_{2n} \\ \vdots & & & & & \vdots \\ u_{n1} & u_{n2} & \dots & u_{nj} & \dots & u_{nn} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{93}$$

Let us define the vector matrix above to be $U = [u_{ij}]_{n\times n}$. Consider two cases:

- $|U| \neq 0$ - the matrix is invertible and its inverse exists. If we multiply both sides by U^{-1} we obtain that all coefficients of α are 0. As such, it is linearly independent.
- $|U| = 0$, then, the inverse does not exist. Furthermore, using the property of this specific matrix that $|U| = |U^T|$, we know that $|U^T| = 0$ and as such it cannot be reduced to I by elementary row operations and can instead be reducible to a matrix with a row of zeroes. We can apply elementary column operations to produce a column of zeroes, and as such, we can obtain $\underline{0}$, meaning these vectors are linearly dependent.

□

4.14 Linear Transformations

Definition 4.11. Linear Transformation

A function $T : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation if $\forall \underline{u}, \underline{v} \in \mathbb{R}^m$ and all $\lambda \in \mathbb{R}$ we obtain

$$T(\underline{u} + \underline{v}) = T(\underline{u} + T(\underline{v}) \wedge T(\lambda \underline{u}) = \lambda T(\underline{u}) \tag{94}$$

Corollary. *Zero vector*

If $T : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation then $T(\underline{0}) = \underline{0}$.

Corollary. *Every matrix defines a linear transformation*

By definition. In particular, let $M = [m_{ij}]_{m\times n}$, then the function $T : \mathbb{R}^m \to \mathbb{R}^n$ defined by $T(\underline{x}) = M\underline{x}$ for every $\underline{x} \in \mathbb{R}^m$ is a linear transformation. In this particular example, we regard $\mathbb{R}^m$ and $\mathbb{R}^n$ as column vectors so that $M\underline{x}$ is the product of an $n \times m$ matrix and an $m \times 1$ matrix which yields us an $n \times 1$ matrix. The linearity instantly follows from the definition of matrix multiplication:

$$M(\underline{u} + \underline{v}) = M\underline{u} + M\underline{v} \tag{95}$$

$$M(\lambda \underline{u}) = \lambda M(\underline{u}) \tag{96}$$

Definition 4.12. Matrix of linear transformation

Let $T : \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation, we define the matrix of a linear transformation and show that it relates coordinates in $\mathbb{R}^m$ to coordinates in $\mathbb{R}^n$.
Let $V = \{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_m\}$ be a basis for $\mathbb{R}^m$ and $W = \{\underline{w}_1, \underline{w}_2, \ldots, \underline{w}_n\}$ be a basis for $\mathbb{R}^n$. Then each of the vectors $T(\underline{v}_j)$ belongs to $\mathbb{R}^n$ and so is a linear combination of $\underline{w}_1, \underline{w}_2, \ldots, \underline{w}_n$. Hence for each j there are n numbers $\alpha_{1j}, \alpha_{2j}, \ldots, \alpha_{nj} \in \mathbb{R}$ with

$$T(\underline{v}_j) = \alpha_{1j}\underline{w}_1 + \alpha_{2j}\underline{w}_2 + \ldots + \alpha_{nj}\underline{w}_n \tag{97}$$

The matrix of T with respect to bases V and W is defined to be the $n \times m$ whose j th column contains the coefficients in the expansion of $T(\underline{v}_j)$, i.e. the matrix

$$\begin{bmatrix} \alpha_{11} & \ldots & \alpha_{1j} & \ldots & \alpha_{1m} \\ \alpha_{21} & \ldots & \alpha_{2j} & \ldots & \alpha_{2m} \\ \vdots & & & & \vdots \\ \alpha_{n1} & \ldots & \alpha_{nj} & \ldots & \alpha_{nm} \end{bmatrix}$$

In the special case where $m = n$ and $V = W$ we talk about the matrix of T with respect at basis V. Now, if $\underline{x} \in \mathbb{R}^m$ has coordinates $[x_1, x_2, \ldots, x_m]$ with respect to V, then

$$\begin{aligned}
\underline{x} =& x_1\underline{v}_1 + x_2\underline{v}_2 + \ldots + x_m\underline{v}_m \\
T(\underline{x}) =& x_1T(\underline{v}_1) + x_2T(v_2) + \ldots + x_mT(\underline{v}_m) \\
=& x_1(\alpha_{11}\underline{w}_1 + \alpha_{21}\underline{w}_2 + \ldots + \alpha_{n1}\underline{w}_n) + x_2(a_{12}\underline{w}_1 + \alpha_{22}\underline{w}_2 + \ldots + \alpha_{n2}\underline{w}_n) \\
& + \ldots + x_m(\alpha_{1m}\underline{w}_1 + \alpha_{2m}\underline{w}_2 + \ldots + \alpha_{nm}\underline{w}_n) \\
=& (x_1\alpha_{11} + x_2\alpha_{12} + \ldots + x_m\alpha_{1m})\underline{w}_1 + (x_1\alpha_{21} + x_2\alpha_{22} + \ldots + x_m\alpha_{2m})\underline{w}_2 \\
& + \ldots + (x_1\alpha_{n1} + x_2\alpha_{n2} + \ldots + x_m\alpha_{nm})\underline{w}_n
\end{aligned}$$

Having now expressed $T(\underline{x})$ as a linear combination of the vectors in W, so the coordinates $[y_1, y_2, \ldots, y_n]$ of

$T(\underline{x})$ with respect to W are, by definition, given by:

$$\begin{aligned}
\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} &= \begin{bmatrix} x_1\alpha_{11} + x_2\alpha_{12} + \ldots + x_m\alpha_{1m} \\ x_1\alpha_{21} + x_2\alpha_{22} + \ldots + x_m\alpha_{2m} \\ \vdots \\ x_1\alpha_{n1} + x_2\alpha_{n2} + \ldots + x_m\alpha_{nm} \end{bmatrix} \\
&= \begin{bmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1m} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2m} \\ \vdots & & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \\
&= M \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}
\end{aligned}$$

Where M is the vector of T which respect to V and W. This gives us the corollary:

Corollary. *M matrix of bases*

Let $T : \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation and let M be the matrix of T with respect to bases V in $\mathbb{R}^m$ and W in $\mathbb{R}^n$. Then the columns of M contain the coordinates of the images of the basis vectors in V with respect to the basis W. If $\underline{x} \in \mathbb{R}^m$ has coordinates $[x_1, x_2, \ldots, x_m]$ with respect to V, then the coordinates $[y_1, y_2, \ldots, y_n]$ of $T(\underline{x}) \in \mathbb{R}^n$ with respect to W are

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = M \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \tag{98}$$

Example 4.5. Example 1

Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be the linear transformation given by

$$T(x, y) = (y, x + y, x) \tag{99}$$

Find the matrix of T with respect ot the basis $V = \{(1, 1), (1, -2)\}$ of $\mathbb{R}^2$ and the basis $W = \{(1, 2, 0), (2, 1, 0), (0, 0, 1)\}$ of $\mathbb{R}^3$. If a vector $\underline{u}$ has coordinates $[2, 3]$ with respect to V, then what are the coordinates of $T(\underline{u})$ with respect to W?
We have

$$\begin{aligned} T(1, 1) &= (1, 2, 1) = (1, 2, 0) + 0(2, 1, 0) + (0, 0, 1) \\ T(1, -1) &= (-1, 0, 1) = \alpha(1, 2, 0) + \beta(2, 1, 0) + \gamma(0, 0, 1) \end{aligned}$$

Thus we obtain

$$\begin{aligned} \alpha &= \frac{1}{3} \\ \beta &= -\frac{2}{3} \\ \gamma &= 1 \end{aligned}$$

We now have expressed the images of the vectors in V as a linear combination of vectors in W. So the matrix of T with respect to V and W is

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

And if $\underline{u}$ has coordinates $[2, 3]$ with respect to V then

$$\begin{bmatrix} 1 & \frac{1}{3} \\ 0 & -\frac{2}{3} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \\ 5 \end{bmatrix}$$

So the coordinates of $T(\underline{u})$ with respect to W are $[3, -2, 5]$

4.15 Projection

Definition 4.13. Projection

Let $\underline{u}$ be a non-zero vector in $\mathbb{R}^2$. If $\underline{x} \in \mathbb{R}^2$ we define projection of $\underline{x}$ onto $\underline{u}$ to be the vector $P_{\underline{u}}(\underline{x})$ with the properties

$$P_{\underline{u}}(\underline{x}) \text{ is a multiple of } \underline{u} \tag{100}$$

$$\underline{x} - P_{\underline{u}}(\underline{x}) \text{ is perpendicular to } \underline{u} \tag{101}$$

Using these properties, we can show that for some $\alpha \in \mathbb{R}$

$$P_{\underline{u}} = \alpha \underline{u}$$

And using the other property

$$\begin{aligned} 0 &= (\underline{x} - P_{\underline{u}}(\underline{x})) \cdot \underline{u} \\ \implies 0 &= (\underline{x} - \alpha\underline{u}) \cdot \underline{u} \\ \implies 0 &= \underline{x} \cdot \underline{u} - \alpha|\underline{u}|^2 \\ \implies \alpha &= \frac{\underline{x} \cdot \underline{u}}{|\underline{u}|^2} \end{aligned}$$

Therefore we can define

With this, it is easy to check that it is a linear transformation. Left as an exercise for the reader.

4.16 Rotation through origin

Definition 4.14. Rotation through origin

Let $(x', y') = R_\theta(x, y)$, then we can write the rotation in matrix form as

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{102}$$

Proof. Consider $\theta \in [0, 2\pi)$, we define $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ to be the rotation of a point through an angle θ about the origin. Let us define points

$$\begin{aligned} (x, y) &= (r\cos\phi, r\sin\phi) \\ (x', y') &= (r\cos(\phi+\theta), r\sin(\phi+\theta)) \end{aligned}$$

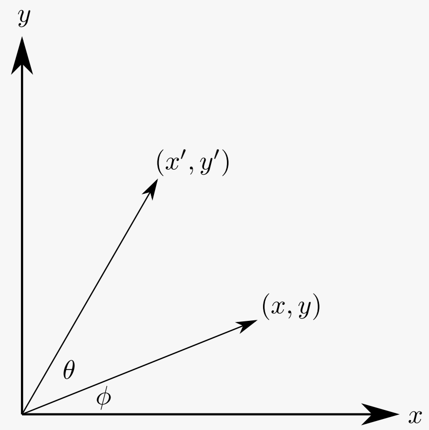

Figure 3: Rotation definitions

Then, through identities we have

$$\begin{aligned} (x', y') &= (r(\cos\phi\cos\theta - \sin\phi\sin\theta), r(\sin\phi\cos\theta + \sin\theta\cos\phi)) \\ (x', y') &= (x\cos\theta - y\sin\theta, y\cos\theta + x\sin\theta) \end{aligned}$$

For which we can collect like terms and write it in matrix form to obtain

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

□

4.17 Coordinates through basis

Definition 4.15. Coordinates

Let $V = \{\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n\}$ be a basis for $\mathbb{R}^n$. If $\underline{x} \in \mathbb{R}^n$, then $\underline{x}$ has a unique expansion as a linear combination

$$\underline{x} = \alpha_1 \underline{v}_1 + \alpha_2 \underline{v}_2 + \alpha_n \underline{v}_n \tag{103}$$

of these basis vectors. The coefficients $\alpha_1, \alpha_2, \dots, \alpha_n$ are called the coordinates of $\underline{x}$ with respect to the basis V.

Example 4.6. Example 1

Let $E = \{(1,0),(0,1)\}$ be the standard basis for $\mathbb{R}^2$ and let V be the basis $\{(1,-1),(2,3)\}$. Find the coordinates of $(1,2)$ with respect to E and V.

$$(1,2) = 1(1,0) + 2(0,1) \tag{104}$$

$$[1,2] \text{ for } E \tag{105}$$

$$\tag{106}$$

For V:

$$(1,2) = \alpha(1,-1) + \beta(2,3) \tag{107}$$

Solving these equations we obtain

$$\alpha = -\frac{1}{5} \tag{108}$$

$$\beta = \frac{3}{5} \tag{109}$$

Hence, the coordinates with respect to V are $\left[-\frac{1}{5}, \frac{3}{5}\right]$

4.18 Change of basis

If we have different basis in $\mathbb{R}^n$ then a given vector will have different coordinates with respect to each basis. The change in coordinates can be described by transition matrix

Definition 4.16. Transition Matrix

Let $V = \{\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n\}$ and $W = \{\underline{w}_1, \underline{w}_2, \dots, \underline{w}_n\}$ be two bases in $\mathbb{R}^n$. Suppose $\underline{x} \in \mathbb{R}^n$ has coordinates $[\alpha_1, \alpha_2, \dots, \alpha_n]$ with respect to V and coordinates $[\beta_1, \beta_2, \dots, \beta_n]$ with respect to W. We can find the list of two coordinates by using the identity transformation $I : \mathbb{R}^n \to \mathbb{R}^n$ which is defined by

$$I(\underline{x}) = \underline{x}, \quad \forall x \in \mathbb{R}^n \tag{110}$$

let M be the matrix of I with respect to the vases V and W. Then the coordinates $[\beta_1, \beta_2, \dots, \beta_n]$ of $I(\underline{x})$ i.e. of $\underline{x}$ with respect to W are given by

$$\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} = M \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} \tag{111}$$

4.19 Eigenvalues and Eigenvectors

Definition 4.17. Eigenvalues and Eigenvectors

Let A be a square matrix of order n. A number λ is called an eigenvalue of A if $A\underline{v} = \lambda\underline{v}$ for some non-zero column vector $\underline{v}$. When this is the case we call $\underline{v}$ an eigenvector of A corresponding to λ. These must be squares.

To find eigenvalues and eigenvectors we use the following property:

Definition 4.18. Characteristic Equation

A number λ is an eigenvalue of the matrix A if and only if

$$|A - \lambda I| = det(A - \lambda I) = 0 \tag{112}$$

This is also a polynomial of degree n in λ.

Proof.

$$\begin{aligned}
\lambda \text{ is an eigenvalue of } A &\iff A\underline{v} = \lambda\underline{v} \text{ for some non-zero } \underline{v} && (113)\\
&\iff (A - \lambda I)\underline{v} = \underline{0} \text{ for some } \underline{v} && (114)\\
&\iff |A - \lambda I| = 0 && (115)
\end{aligned}$$

□

Example 4.7. Example 1

Find the eigenvalues and eigenvectors of the matrix

$$A = \begin{bmatrix} -5 & 3 \\ 6 & -2 \end{bmatrix}$$

We have

$$\begin{aligned} A - \lambda I &= \begin{bmatrix} -5 & 3 \\ 6 & -2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -5-\lambda & 3 \\ 6 & -2-\lambda \end{bmatrix} \end{aligned}$$

We know that determinant must be 0, hence

$$\begin{aligned} (-5-\lambda)(-2-\lambda) - 18 &= 0 \\ \lambda^2 + 7\lambda - 8 &= 0 \\ (\lambda - 1)(\lambda + 8) &= 0 \\ \lambda &= 1, -8 \end{aligned}$$

Hence the eigenvalues for A are 1 and -8. To find the eigenvectors we consider each eigenvalue in turn:

$$\begin{aligned} \begin{bmatrix} -5 & 3 \\ 6 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} &= 1 \begin{bmatrix} x \\ y \end{bmatrix} \\ -6x + 3y &= 0 \\ 6x - 3y &= 0 \\ y &= 2x \end{aligned}$$

Therefore for any nonzero vector of the form $(x, 2x)$ is an eigenvector corresponding to value 1.

$$\begin{aligned} \begin{bmatrix} -5 & 3 \\ 6 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} &= -8 \begin{bmatrix} x \\ y \end{bmatrix} \\ 3x + 3y &= 0 \\ 6x + 6y &= 0 \\ y &= -x \end{aligned}$$

Hence any non-zero vector of the form $(x, -x)$ has is an eigenvector corresponding to the eigenvalue -8.

It is possible for eigenvalues of a real matrix to be complex. This happens due to the quadratic.

4.20 Diagonalisation of a matrix

Theorem 4.3. *Diagonalisation*

Let A be an $n \times n$ matrix. If

$$V^{-1}AV = D = diag(\lambda_1, \lambda_2, \ldots, \lambda_n) \tag{116}$$

where V is the $n \times n$ matrix whose columns are $\left[\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n\right]$ then $\left[\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n\right]$ are the eigenvectors of A, and $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the corresponding eigenvalues.

Proof. If $V^{-1}AV = D = diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$ holds, then we know that because it is a set of column vectors of a matrix with non-zero determinant, it is linearly independent. That is, because V must be non-singular, its determinant is non-zero. Now let $\underline{e}_j$ be the $n \times 1$ column matrix whose j th element is 1 but has all other elements zero. Now consider the product $V\underline{e}_j$. From laws of matrix algebra, this must be $\underline{v}_j$, the j th column of V. Therefore we have the implications

$$\begin{aligned} V^{-1}AV &= D \\ V^{-1}AV\underline{e}_j &= D\underline{e}_j \\ V^{-1}A(V\underline{e}_j) &= \lambda_j\underline{e}_j \\ A\underline{v}_j = V\lambda_j\underline{e}_j = \lambda_j V\underline{e}_j &= \lambda_j\underline{v}_j \end{aligned}$$

But $A\underline{v}_j = \lambda_j\underline{v}_j$ defines $\underline{v}_j$ and λ_j to be the j th eigenvector and associated eigenvalue of A, respectively. This completes the proof. □

Example 4.8. Example 1

Find a diagonal matrix D an invertible matrix P such that $P^{-1}AP = D$ where A is the matrix

$$A = \begin{bmatrix} 4 & 2 & 2 \\ -1 & 1 & -1 \\ 0 & 0 & 2 \end{bmatrix} \tag{117}$$

Using the characteristic equation $|A - | = 0$ we obtain

$$\begin{bmatrix} 4-\lambda & 2 & 2 \\ -1 & 1-\lambda & -1 \\ 0 & 0 & 2-\lambda \end{bmatrix} = 0 \tag{118}$$

which in turn obtains us

$$(2-\lambda)(\lambda-2)(\lambda-3) = 0$$
$$\lambda = 2, 2, 3$$

Now (x, y, z) is an eigenvector corresponding to the eigenvalue 2 if and only if

$$\begin{bmatrix} 4 & 2 & 2 \\ -1 & 1 & -1 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 2 \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

which leads us to

$$x + y + z = 0$$

Therefore any vector of the form $(x, y, -x - y)$ where x and y are not both zero is an eigenvector corresponding to the repeated eigenvalue 2. Now for 3:

$$\begin{bmatrix} 4 & 2 & 2 \\ -1 & 1 & -1 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 3 \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

We obtain the eigenvector $(-2y, y, 0)$ where y is non-zero for the eigenvalue 3. Hence we obtain

$$P = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 1 \\ -1 & -1 & 0 \end{bmatrix}$$
$$D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

The P comes from $x(1, 0, -1)$, $y(0, 1, -1)$, $y(-2, 1, 0)$. The diagonals come from their respective eigenvalues.

5 Sequences

Definition 5.1. Sequence

A sequence (a_n) is an infinite list of numbers $(a_0, a_1, a_2, a_3, \ldots)$. The index of the first term will usually be 0, but sometimes we will start with a_1. As an example:

$$2^{-n} \text{ denotes the sequence } (1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots) \tag{119}$$

5.1 Limit of a Sequence

Definition 5.2. Convergence of a sequence

A sequence a_n of real numbers is convergent to a limit $l \in \mathbb{R}$ if for every $\varepsilon > 0$ there is an integer N (that depends on ε) with $|a_n - l| < \varepsilon$ for all $n > N$. When a_n converges to l we write

$$\lim_{n\to\infty} a_n = l \tag{120}$$

(121)

Example 5.1. Example 1

The sequence $\frac{1}{n}$ for $n > 0$ converges to 0. To see this, let $\varepsilon > 0$. Then there is an N with $\frac{1}{N} < \varepsilon$ and so for $n > N$ we obtain

$$\left|\frac{1}{n} - 0\right| = \frac{1}{n} < \frac{1}{N} < \varepsilon \tag{122}$$

$$\therefore \frac{1}{n} \text{ converges to } 0 \tag{123}$$

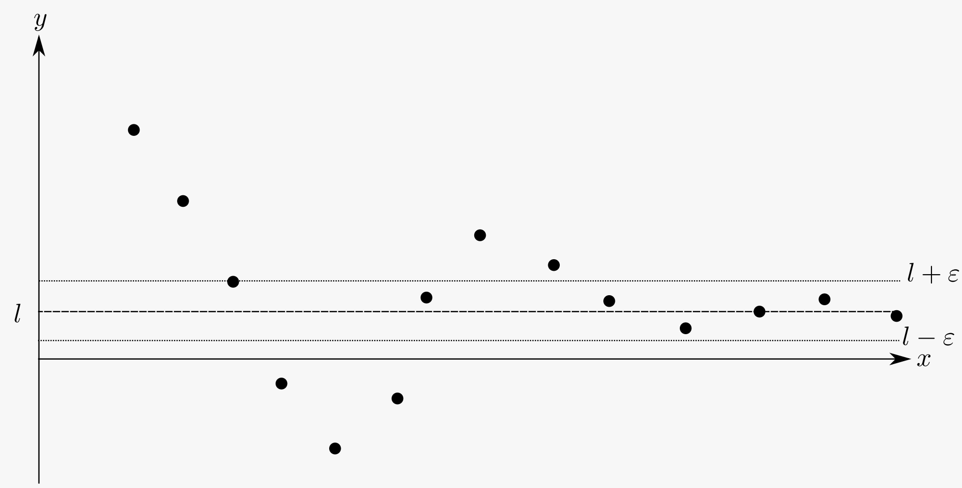

Figure 4: seqconvergece

5.2 Rules for Convergent Sequences

If a_n, b_n and c_n are convergent sequences with $a_n \to \alpha$, $b_n \to \beta$ and $c_n \to \gamma$, then

1. Sum - $a_n + b_n \to \alpha + \beta$
2. Scalar multiple - $\lambda a_n \to \lambda\alpha$
3. Product - $a_n b_n \to \alpha\beta$
4. Reciprocal - $\frac{1}{a_n} \to \frac{1}{\alpha}$
5. Quotient - $\frac{b_n}{a_n} \to \frac{\alpha}{\beta}$

6. Hybrid - $\frac{b_n c_n}{a_n} \to \frac{\beta\gamma}{\alpha}$