Week 1

Vector

- For this course remember vector means rows or columns of numbers

Matrix

Types of Matrices

Row Matrix

$$\begin{pmatrix} a & b & c \end{pmatrix}$$

Column Matrix
Vector Matrix

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

Zero Matrix
Null Matrix

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

Diagonal Matrix

$$\begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}$$

Scalar Matrix

$$\begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix}$$

Unit Matrix

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

Upper Triangular Matrix

$$\begin{pmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{pmatrix}$$

Lower Triangular Matrix

$$\begin{pmatrix} a & 0 & 0 \\ b & c & 0 \\ d & e & f \end{pmatrix}$$

Properties of matrix addition and multiplication

1. $(A + B) + C = A + (B + C)$ (Associativity of addition)
2. $(AB)C = A(BC)$ (Associativity of multiplication)
3. $A + B = B + A$ (Commutativity of addition)
4. $AB \neq BA$ (In General)

Week 2 P1

$L2.1$ Determinants (Part3)

Important properties and Identities

Property 1 : Determinant of a product is product of the determinants

1. det(AB) = det(A)*det(B)
2. det(AB) = det(BA) here A and B are square matrices of size n
3. det($A^T A$) $= det(A)^2$ here ($det(A^t)$ = $det(A)$)

Property 2: Switching two rows or columns changes the sign of the determinant

$$det(A) = -det(\tilde{A})$$

Property 3: Adding multiple of a row to another row leaves the determinant unchanged (Same for columns)

Property 4: Scalar multiplication of a row by a constant t multiplied the determinant by t. (same for columns)

L2.2 Cramer's Rule

Crammer's Rule 3x3

$$\begin{aligned} 2x + 3y - 5z &= 1 \\ x + y - z &= 2 \\ 2y + z &= 8 \end{aligned} \qquad x = \frac{D_x}{D} \quad y = \frac{D_y}{D} \quad z = \frac{D_z}{D}$$

$$D = \begin{vmatrix} 2 & 3 & -5 \\ 1 & 1 & -1 \\ 0 & 2 & 1 \end{vmatrix} = -7 \qquad D_y = \begin{vmatrix} 2 & 1 & -5 \\ 1 & 2 & -1 \\ 0 & 8 & 1 \end{vmatrix} = -21$$

$$D_x = \begin{vmatrix} 1 & 3 & -5 \\ 2 & 1 & -1 \\ 8 & 2 & 1 \end{vmatrix} = -7 \qquad D_z = \begin{vmatrix} 2 & 3 & 1 \\ 1 & 1 & 2 \\ 0 & 2 & 8 \end{vmatrix} = 14$$

Week 2 P2

L2.3 Solution to a system of linear equations with an invertible coefficient matrix

- If $Ab = BA = I_{N \times N}$ and is denoted by A^{-1}, B is called inverse of A
- Inverse of a matrix exist iff $det(A) \neq 0$

Adjugate of a square matrix

- $adj(A) = C^{T}\ where\ C_{ij} = (-1)^{i+j}M_{ij}$ cofactor

- $A^{-1} = \frac{1}{det(A)}adj(A)$ A is nxn square matrix and det(A) is not 0

The solution of a system of linear equations with an invertible coefficient matrix

$x = A^{-1}b$ here A is invertible i.e det(A) is not 0

Solutions of a homogeneous system of linear equations

$Ax = 0$

- Unique solution which is 0 if $det(A) \neq 0$
- Infinite solution if $det(A) = 0$

Properties of Adjugate of a square matrix

- $adj(AB) = adj(B) \times adj(A)$
- $adj(A + B) = adj(A) + adj(B)$
- $adj(A^{T}) = adj(A)^{T}$
- $adj(A^{-1}) = adj(A)^{-1}$

The echelon form

A matrix is a **row echelon form** if

- The first non-zero element in each row, called the leading entry is 1

Week 2 P3

- Each leading entry is in a column to the right of the leading entry in the previous row.
- Rows with all zero elements, if any are below rows having a non-zero element
- For a non-zero row, the leading entry in the row is the only non-zero entry in its column.

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Row echelon form

$$\begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Reduced row echelon form

Row Reduction

1. Interchange two rows e.g $R1 \leftrightarrow R2$
2. Scalar multiplication of a row by a constant e.g $R1/3$
3. Adding multiple of a row to another row e.g $R1 - 3R2$

Recall from determinants

$A \sim\sim\sim B$

1. $Ri \leftrightarrow Rj \Rightarrow det(A) = - det(B)$
2. $Ri * c \Rightarrow det(A) = c * det(B)$
3. $Ri + cRj \Rightarrow det(A) = det(B)$

L2.6 Gaussian Elimination method

1. Augmented matrix is denoted by $[A \mid B]$ *we denote the like this*

System of Equations

Associated Augmented Matrix

$$\begin{Bmatrix} x_1 + x_2 + x_3 = 3 \\ 2x_1 + 3x_2 + x_3 = 5 \\ x_1 - x_2 - 2x_3 = -5 \end{Bmatrix} \Longleftrightarrow \left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 2 & 3 & 1 & 5 \\ 1 & -1 & -2 & -5 \end{array}\right]$$

Week 3 P1

L3.1 Introduction to vector spaces

Properties of addition and scalar multiplication

Let v, w and $v`$ be vectors in $\mathfrak{R}$ and $a, b \in \mathfrak{R}$

i. $v + w = w + v$
ii. $(v + w) + v` = v + (w + v`)$
iii. The $0^{\rightarrow}$ vector satisfies that $v + 0 = 0 + v = v$
iv. The vector $-v$ satisfies that $v + (-v) = 0$
v. $1v = v$
vi. $(ab)v = a(bv)$
vii. $a(v + w) = av + aw$
viii. $(a + b)v = av + bv$

A **vector space** is a set with two operations (called addition and scalar multiplication
With the above properties (i) to (viii).

Definition of a vector space

A vector space V over $\mathfrak{R}$ is a set along with two functions

$+ : V \times V \rightarrow V$ and $\bullet : \mathfrak{R} \times V \rightarrow V$

(i.e for each pair of elements $v_1 and\ v_2\ in\ V$, there is a unique element $v_1 + v_2\ in\ V$ and for each $c \in \mathfrak{R}\ and\ v \in V$ there is a unique element $c.v$ in V)

It is standard to suppress the . and only write cv instead of c.v
The function + and • are required to satisfy the above mentioned rules.(i-viii)

L3.2 Some properties of vector spaces

Cancellation law of Vector addition

⇒If $v_1, v_2, v_3 \in V\ such\ that\ v_1 + v_3 = v_2 + v_3\ then\ v_1 = v_2$
⇒ The vector 0 described in (iii) is unique
⇒ The vector v` described in (iv) is unique and it is standard to refer to it as -v
In any Vector space V the following statements are true

- $0v = 0$ for each $v \in V$ and $c0 = 0$ for each $c \in \mathfrak{R}$

Week 3 P2

L3.3 Linear dependence

Linear Combination

Let V be a vector space and $v_1, v_2, v_3, \ldots v_n \in V$. The ***linear combination*** of $v_1, v_2, v_3, \ldots v_n$ with coefficients $a_1, a_2, a_3, \ldots a_n \in \mathbb{R}$ is the vector $\sum_{i=1}^{n} a_i v_i \in V$

A vector $v \in V$ is a ***linear combination*** of $v_1, v_2, v_3, \ldots v_n$ if there exist some $a_1, a_2, a_3, \ldots a_n \in \mathbb{R}$

So that v = $\sum_{i=1}^{n} a_i v_i$

Definition of Linear Dependence

A set of vectors $v_1, v_2, v_3, \ldots v_n$ from a vector space V is said to be ***linearly dependent***, if there exist scalars $a_1, a_2, a_3, \ldots a_n$ *not all zero* such that

$$a_1 v_1 + a_2 v_2 + a_3 v_3 + \ldots + a_n v_n = 0$$

Equivalently , the 0 vector is a linear combination of $v_1, v_2, v_3, \ldots v_n$ with non-zero coefficients.

- If a set is linearly dependent, then so is every superset of it.

L3.4 Linear Independence Part-1

Definition of linear independence

A set of vectors $v_1, v_2, v_3, \ldots v_n$ from a vector space V is said to be ***linearly independent*** if $v_1, v_2, v_3, \ldots v_n$ are not ***linearly dependent .***

Equivalently : A set of vectors $v_1, v_2, v_3, \ldots v_n$ from a vector space V is said to be ***linearly independent*** if the only linear combination of $v_1, v_2, v_3, \ldots v_n$ which equals 0 is the linear combination with all coefficients 0.

Week 3 P3

The **0** vector

Let $v_1, v_2, v_3, \ldots v_n$ be a set of vectors containing the **0** vector. Suppose $v_i = 0$. Then we can choose $a_i = 1$ and $a_j = 0$ for $j \neq i$

Then the linear combination of $a_1 v_1 + a_2 v_2 + a_3 v_3 + \ldots + a_n v_n$ is 0 but not all coefficients are 0.

Hence, a set of vectors $v_1, v_2, v_3, \ldots v_n$ containing the **0** vectors is always a linearly dependent set.

- **Two non-zero vectors are *linearly independent* precisely when they are *not multiples of each other.***
- **If three vectors are *linearly independent* then *none of these vectors is a linear combination of the other two.***
- **To check $v_1, v_2, v_3, \ldots v_n \in \mathbb{R}^m$ are linearly independent we have to check that homogeneous system of linear equations $Vx = 0$ has only the trivial solution, where the j^{th} column of V is v_j.**
- **Any set of n vectors in $\mathbb{R}^2$ with $n \geq 3$ are linearly dependent.**
- **Any set of r vectors in $\mathbb{R}^n$ with $r \geq n$ are linearly dependent.**

- **If $det(A) \neq 0$ then vectors are linearly independent.**

Subspace

Def .

A subset W of a vector space V over $\mathbb{R}$ is called subspace of V if W is also a Vector space over $\mathbb{R}$ with the same operations , defined over V.

Suppose W is a given subset of V

Test for becoming a subspace

i. $0 \in W$

ii. $x \in W, y \in W \Rightarrow x + y \in W$ (closed under addition)

iii. $c \in \mathbb{R}, x \in W \Rightarrow cx \in W$ (closed under scalar multiplication)

Then we say W is subspace of V

Week 4 P1

L4.1 What is a basis for a vector space

Span of a set of Vectors

The span of a set S (of Vectors) is defined as the set of all finite linear combinations of elements (vectors) of S, and denoted by Span(S)

i.e $$Span(S) = \left\{ \sum_{i=1}^{n} a_i v_i \in V \mid a_1, a_2, \ldots a_n \in \mathbb{R} \right\}$$

Example:

Let $S = \{(1,0)\} \in \mathbb{R}^2$ Then

$Span(S) = \{a(1,0) \mid a \in \mathbb{R}\} = \{(a,0) \mid a \in \mathbb{R}\}$

Thus $Span(S)$ is the X-axis in $\mathbb{R}^2$

Spanning set for a vector space

Let V be a vector space. A set $S \subseteq V$ is a spanning set for V if $Span(S) = V$

Example:

If $S = \{(1,0), (0,1)\}$ then $Span(S) = \mathbb{R}^2$

If $S = \{(1,0), (0,1), (1,2)\}$ then $Span(S) = \mathbb{R}^2$

If $S = \{(1,1), (0,1)\}$ then $Span(S) = \mathbb{R}^2$

If $S = \{(1,0,0), (0,1,0), (0,0,1)\}$ then $Span(S) = \mathbb{R}^3$

What is the basis ?

A basis B of a vector space V is a linearly independent subset of V that spans V.
Example:

Let $e_i \in \mathbb{R}^n$ be the vector with i^{th} coordinate 1 and all other coordinates 0 e.g.

$e_1 = (1, 0, 0, \ldots, 0)$

The set $\epsilon = \{e_1, e_2, e_3, \ldots e_n\} \subseteq \mathbb{R}^n$ is a basis for $\mathbb{R}^n$

Week 4 P2

L4.2: Finding bases for vector spaces

Equivalent conditions for B to be a basis

The following conditions are equivalent to a subset $B \subseteq V$ being a basis:
i. B is linearly independent and Span(B) = V
ii. B is a maximal linearly independent set
iii. B is a minimal spanning set.

How do we find a basis ?

i) Start with the ϕ and keep appending vectors which are not in the span of the set thus far obtained , until we obtain a spanning set.

ii) Take a spanning set and keep deleting vectors which are linear combinations of the other vectors, until the remaining vectors satisfy that they are not a linear combination of the other remaining ones.

L4.3: What is the rank / dimension for a vector space.

The dimension (or rank) of a vector space is the **size (or cardinality) of a basis of the vector space.**

For this course: If B is a basis of V, then the rank is the number of elements in B.

For every vector space there exists a basis, and all bases of a vector space have the same number of elements (or cardinality); hence, the dimension (or rank) of a vector space (say V) is uniquely defined and denoted by $dim(V)$ (or $rank(V)$) respectively.

Dimension of $\mathbb{R}^n$

Recall the i-th standard basis vector is $\mathbb{R}^n$.
$e_i = (0, 0, 0..., 0, 1, 0... 0)$
i.e. the i-th coordinate is 1 and 0 elsewhere
Recall that the set $\{e_1, e_2, e_3,..., e_n\}$ is a basis of $\mathbb{R}^n$ called the standard basis.

Hence the dimension of $\mathbb{R}^n$ is n.

Week 4 P3

L4.3:

Rank of a matrix

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

⇒ The **column space** of a A is the subspace of $\mathbb{R}^m$ spanned by the column vector of A.

⇒ The **row space** of a A is the subspace of $\mathbb{R}^n$ spanned by the row vectors of A.

⇒ The dimension of the column space of A is defined as the **column rank** of A.

⇒ The dimension of row space of A is defined as the **row rank** of A.

Fact: **Column rank = row rank** and this number is called the rank of A.

L4.4: Rank and dimension using Gaussian elimination

Finding dimension and basis with a given spanning set.

e.g. Let us consider the vector space W spanned by the set S = {(1,0,1),(-2,-3,1),(3,3,0)}.

We will use the following steps to find the dimension and a basis for W and carry out the steps for our example.

Reduce to a matrix in row echelon form

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

⇒ The number of non-zero rows is the dimension of the vector

⇒ The vectors corresponding to the non-zero rows form the basis of the vector space W.

In the example , the final matrix is

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

Hence dimension of the vector space spanned by {(1,0,1),(-2,-3,1),(3,3,0)} is 2

And a basis is given by (1,0,1),(0,1,-1)

Week 4 P4

L4.4:

An alternative to the row-based method.

Column method

Example

Let us consider the vector space W spanned by the set S = {(1,0,1),(-2,-3,1),(3,3,0)}.
We will use the fact (see notes or slides) to find a basis for W which is a subset of S.
Form the matrix with the vectors in S as the columns.

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

Row reduce this matrix

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

This matrix is in row echelon form and columns with pivot entries (leading 1s) are the first and second columns.
Therefore (1,0,1),(-2,-3,1) which are the first and second vectors in S respectively, form a basis for W.

Week 5 P1

L5.1: The null space of a matrix: finding nullity and a basis Part 1

Solution space of a homogeneous system of linear equations.

Let A be an $m \times n$ matrix

The subspace $W = \{x \in \mathbb{R}^n \mid Ax = 0\}$ of $\mathbb{R}^n$ is called the solution space of the homogeneous system of linear equation $Ax = 0$ or the **null space** of A.

Note that the null space is a subspace of $\mathbb{R}^n$. The dimension of the null space is called **the nullity of A.**

Finding the nullity and a basis for the null space.

We have seen how to find the dimensions and a basis for the row space of A using row reduction.

We will use row reduction to also find the nullity and a basis for the null space of A.

Recall first how to find the solution space for system $Ax = b$ i.e Gaussian elimination.
⇒ Form the augmented matrix $[A|b]$
⇒ Apply the same row reduction operations on the augmented matrix that are used to row reduce A to obtain the augmented matrix $[R|C]$ where R is the matrix in reduced row echelon form obtained from A.
⇒ If the i-th column has the leading entry of some row, we call x_i a dependent variable.
⇒ If the i-th column does not have the leading entry of some row we call x_i an independent variable.

Finding the nullity and a basis for the null space.

$nullity(A) = number\ of\ independent\ variables.$

⇒ Assign arbitrary value t_i to the i-th independent variable.
⇒ Compute the value of each dependent variables in terms of $t_i s$ from the unique row it occurs in.
⇒ Every solution is obtained by letting $t_i s$ vary in $\mathbb{R}$

The vectors obtained by substituting $ti = 1$ and $tj = 0 \; \forall j \neq i$ as i varies constitutes a basis of the null space of A (i.e the solution space of $Ax = 0$)

L5.2: The null space of a matrix: finding nullity and a basis - Part 2

The rank-nullity theorem

Let A be $m \times n$ matrix.
Recall the row rank of A is the dimension of the row space of A and the column rank of A is the dimension of the column space of A. These are equal and are denoted by $rank(A)$.

$rank(A)$ is calculated as the number of non-zero rows of the matrix R in reduced row echelon form obtained by row reduction.

Note that for a matrix R in row echelon form the

Number of non-zero rows = number of dependent variables

For the corresponding homogeneous system $Rx = 0$

Hence, $rank(A) = number\ of\ nonzero\ rows\ of\ R$ = number of dependent variables of $Rx = 0$

$nullity(A)$ = number of independent variables of $Rx = 0$.

Therefore we have the rank-nullity theorem.
Theorem

For an $m \times n$ matrix A
$rank(A) + nullity(A) = n$ (dependent variable + independent variables = total variable)

How to check if a set of n vectors is a basis for $\mathbb{R}^n$

Short answer: Use determinants.

Suppose we are given n vectors of $\mathbb{R}^n$.
We write them as column of a matrix , thus obtaining an $n \times n$ (square) matrix
If the determinants of a matrix are 0, then the given set of vectors does not form a basis, otherwise it forms a basis.
Example:

The standard basis (1,0),(0,1) yields the Identity matrix/ with determinant 1.

The vector(1,-2),(5,-10) yields the matrix

1	5
-2	-10

With determinant 0. This is not a basis for $\mathbb{R}^2$.

Week 5 P2

L5.3: What is a linear mapping - Part 1

Grocery shop example

The prices of rice, dal and oil in shop A in the town of Malgudi are as follows:

	Rice (Per kg)	Dal (per kg)	Oil (per litre)
Shop A	45	125	150

The cost of 1 kg of rice , 2kg of dal and 1 litre of oil is
$1 \times 45 + 2 \times 125 + 1 \times 150 = 445.$
The cost of 2kg of rice,1 kg of dal and 2 litre of oil is
$2 \times 45 + 1 \times 125 + 2 \times 150 = 515.$
The cost of x_1 kg of rice, x_2 kg of dal and x_3 kg of oil is
$x_1 \times 45 + x_2 \times 125 + x_3 \times 150 = 45x_1 + 125x_2 + 150x_3$

Expressions and linear combinations

The term $45x_1 + 125x_2 + 150x_3$ is an expression.

We can equivalently think of it as a function c_A from $\mathbb{R}^3$ to $\mathbb{R}$

Since for every value of x_1,x_2,x_3 (with coefficients 45,125,150) , it is an example of a linear function.
Recall that linear combinations can be also expressed in terms of matrix multiplication e.g.

$c_A(x_1, x_2, x_3) = 45x_1 + 125x_2 + 150x_3 = \begin{bmatrix} 45 & 125 & 150 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$

L5.4: What is a linear mapping - Part 2

What is a linear mapping

A linear mapping f from $\mathbb{R}^n$ to $\mathbb{R}^m$ can be defined as follows :

$$f(x_1, x_2, ..., x_n) = (\sum_{j=1}^{n} a_{1j}x_j, \sum_{j=1}^{n} a_{2j}x_j, ..., \sum_{j=1}^{n} a_{mj}x_j)$$

Where the coefficients $a_{ij}s$ are real numbers (scalars). A linear mapping can be thought of as a collection of linear combinations.

We can write the expression on the RHS in matrix forms as Ax

$$\text{where } A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad \text{and} \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Linearity of linear mappings

It follows that a linear mapping satisfies linearity, i.e for any $c \in \mathbb{R}$ (Scalar)

$$f(x_1 + cy_1, x_2 + cy_2, ..., x_n + cy_n) = f(x_1, x_2, ..., x_n) + cf(y_1, y_2, ..., y_n)$$

L5.5: What is a linear transformation

Formal definition of Linear transformation

A function $f : V \to W$ between two vector spaces V and W is said to be a linear transformation if for any two vectors v_1 and v_2 in the vector space V and for any $c \in \mathbb{R}$ (scalar) the following conditions hold :

$\Rightarrow f(v_1 + v_2) = f(v_1) + f(v_2)$

$\Rightarrow f(cv_1) = cf(v_1)$

- Linear mappings are linear transformation

Examples

1. $f : \mathbb{R}^2 \to \mathbb{R}^2$ $\quad f(x, y) = (2x, y)$
2. $f : \mathbb{R}^2 \to \mathbb{R}^2$ $\quad f(x, y) = (2x, 0)$
3. $f : \mathbb{R}^3 \to \mathbb{R}^3$ $\quad f(x, y, z) = (x/2, 3y, 5z)$
4. $f : \mathbb{R}^3 \to \mathbb{R}^4$ $\quad f(x, y, z) = (4y - z, 3y + 11/19z, 5x - 2z, 23y)$

1-1 and onto functions

Recall that a function $f : V \to W$ is 1-1 (or injective) if $f(v_1) = f(v_2)$ implies $v_1 = v_2$

Recall that a function $f : V \to W$ is onto (or surjective) if for every $w \in W$ there exists $v \in V$
Such that $f(v) = w$
For a linear transformation, being 1-1 is equivalent to $f(v) = 0$ implies $v = 0$.

What is an Isomorphism

Recall that a function $f : V \to W$ is bijective (or a bijection) if it is 1-1 and onto.
Note that being a bijection is equivalent to : for any $w \in W$ there exists a unique $v \in V$ such that $f(v) = w$.
A linear transformation $f : V \to W$ between two vector spaces V and W is said to be an isomorphism if it is a bijection.
E.g.

$$f : \mathbb{R}^2 \to \mathbb{R}^2 \quad f(x, y) = (2x, y)$$

Bases determine linear transformations

Let V be a vector space with basis $\{v_1, v_2, \ldots, v_n\}$.

Let $f : V \to W$ be a linear transformation. Then the ordered vectors
$f(v_1), f(v_2), \ldots f(v_n)$ uniquely determine f.

Week 6 P1

L6.1: Linear transformations, ordered bases and matrices

Important property of finite dimensional vector spaces

Let V be a vector space with dimension n. Choose a basis $\{v_1, v_2, v_3..., v_n\}$.

Define $f : V \to \mathbb{R}^n$ by extending the function sending the basis vector v_i to the standard basis vector $e_i \in \mathbb{R}^n$ for each i.

Then f is an isomorphism.

The matrix corresponding to a linear transformation with respect to ordered bases

Let $f : V \to W$ be a linear transformation.
Let $\beta = v_1, v_2, v_3..., v_n$ be an ordered basis of V and $\gamma = w_1, w_2, w_3..., w_n$ be an ordered of W.
Each $f(v_i)$ can be uniquely written as a linear combination of $w_j s$, where $i = 1, 2,..., n$ and $j = 1, 2,... m$.

$$f(v_1) = a_{11}w_1 + a_{21}w_2 +...+ a_{m1}w_m$$
$$f(v_2) = a_{12}w_1 + a_{22}w_2 +...+ a_{m2}w_m$$
:
:
$$f(v_n) = a_{1n}w_1 + a_{2n}w_2 +...+ a_{mn}w_m$$

The matrix corresponding to the linear transformation f with respect to the ordered bases β and γ is given by

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \vdots & \vdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

E.g Let $V = W = \mathbb{R}^2$, $\beta = \gamma = (1, 0), (1, 1)$ and $f(x, y) = (2x, y)$.

$$f(1, 0) = (2, 0) = 2(1, 0) + 0(1, 1)$$
$$f(1, 1) = (2, 1) = 1(1, 0) + 1(1, 1)$$

Hence the matrix corresponding to f w.r.t the ordered bases

$\{(1, 0), (1, 1)\}$ is

$$\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$$

Week 6 P2

Recovering the linear transformation

Let β $= v_1, v_2, v_3 ..., v_n$ and γ $= w_1, w_2, w_3 ..., w_n$ be ordered bases of V and W respectively.
Suppose A is an $m \times n$ matrix. What is the corresponding linear transformation?

Let $v \in V$. Express $v = \sum_{j=1}^{n} c_j v_j$. Define

$f(v) = \sum_{j=1}^{n} c_j \sum_{i=1}^{m} A_{ij} w_i$

Check that f is a linear transformation!
Letting $c_k = 1$ and $c_j = 0$ for all $j \neq k$, we get that $f(v_k) = A_{1k} w_1 + A_{2k} w_2 + ... + A_{mk} w_m$.
Hence the matrix corresponding to f is indeed A.

Fixed ordered bases : Linear transformations ↔ matrices

Let β and γ be ordered bases for vector spaces V and W respectively where $n = dim(V)$ and $m = dim(W)$.
There is a bijection :
{ linear transformation from V to W } ↔ { $m \times n$ matrix }.

L6.2: Image and kernel of linear transformations

Definitions of kernel and image

Let $f : V \to W$ be a linear transformation.
Define the kernel of f (denoted by $ker(f)$) as :
$ker(f) = \{v \in V \mid f(v) = 0\}$.

Define the image of f (denoted by $Im(f)$) as :
$Im(f) = \{w \in W \mid \exists\, v \in V\, for\, which\, f(v) = w\}$.

$Im(f)$ is another name for the “range of the function f” .

Week 6 P3

The kernel and injectivity of a linear transformation

Recall that a function $f : V \to W$ is 1-1 (or injective) if $f(v_1) = f(v_2)$ implies $v_1 = v_2$.

Recall that a linear transformation f being 1-1 (or injective) is equivalent to $f(v) = 0$ implies $v = 0$.

Rewriting the last part in terms of $ker(f)$, we see that a linear transformation is 1-1 (or injective) is equivalent to $ker(f) = 0$.

A linear transformation f is 1-1 iff $ker(f) = 0$.

The image and surjectivity of a linear transformation

Recall that a function $f : V \to W$ is onto (or surjective) if for each $w \in W,$ exists some $v \in V$ such that $f(v) = w$.

It follows from the definition that a function $f : V \to W$ being onto (or surjective) is equivalent to $Range(f) = W$.

Writing this out for linear transformations, we see that : a linear transformation $f : V \to W$ is onto iff $Im(f) = W$.

Kernels and null spaces

Let $f : V \to W$ be a linear transformation. Let β $= v_1, v_2, v_3..., v_n$ and γ $= w_1, w_2, w_3..., w_n$ be ordered bases of V and W respectively.

Let A be the matrix corresponding to f with respect to β and γ.

Recall that for $v = \sum_{j=1}^{n} c_j v_j \in V,\ f(v) = \sum_{j=1}^{n} c_j \sum_{i=1}^{m} A_{ij} w_i$.

Hence, $f(v) = 0 \Leftrightarrow \sum_{j=1}^{n} A_{ij} c_j = 0$ for all i.

Thus, $v = \sum_{j=1}^{n} c_j v_j \in ker(f) \Leftrightarrow$ $\cdot\ c = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$

Is in the null space of A.

Week 6 P4

Images and column spaces

Let $f : V \to W$ be a linear transformation. Let $\beta = v_1, v_2, v_3,..., v_n$ and $\gamma = w_1, w_2, w_3,..., w_n$ be ordered bases of V and W respectively.
Let A be the matrix corresponding to f with respect to β and γ.

Recall that for $v = \sum_{j=1}^{n} c_j v_j \in V$, $f(v) = \sum_{j=1}^{n} c_j \sum_{i=1}^{m} A_{ij} w_i$.

Let $w = \sum_{i=1}^{m} d_i w_i \in W$. Then $w \in Im(f)$ precisely when there exist scalars c_j; $j = 1, 2,..., n$ such that $\sum_{j=1}^{n} A_{ij} c_j = d_i$ for all i.

Equivalently $w = \sum_{i=1}^{m} d_i w_i \in Im(f)$ if there exists a column vector

$$c = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

Such that the column vector

$$d = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_m \end{bmatrix} = Ac.$$

Hence, $w = \sum_{i=1}^{m} d_i w_i \in Im(f) \Leftrightarrow$

$$d = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_m \end{bmatrix}$$

Is in the column space of A.

Bases for the kernel and image of a linear transformation

Let $f : V \to W$ be a linear transformation. Let $\beta = v_1, v_2, v_3,..., v_n$ and $\gamma = w_1, w_2, w_3,..., w_n$ be ordered bases of V and W respectively.
Let A be the matrix corresponding to f with respect to β and γ.

The relation between kernels and null spaces derived earlier actually yields an isomorphism between them.

In particular, the vectors $\begin{bmatrix} c_{11} \\ c_{12} \\ \vdots \\ c_{1n} \end{bmatrix}, \begin{bmatrix} c_{21} \\ c_{22} \\ \vdots \\ c_{2n} \end{bmatrix}, \ldots, \begin{bmatrix} c_{k1} \\ c_{k2} \\ \vdots \\ c_{kn} \end{bmatrix}$ form a basis for the null space of A precisely when $v_1', v_2', \ldots, v_k' \in ker(f)$, where $v_i' = \sum_{j=1}^{n} c_{ij} v_j$, form a basis for $ker(f)$.

Week 6 P5

Similarly, the relation between images and column spaces derived earlier yields an isomorphism between them.

In particular, the vectors $\begin{bmatrix} d_{11} \\ d_{12} \\ \vdots \\ d_{1m} \end{bmatrix}, \begin{bmatrix} d_{21} \\ d_{22} \\ \vdots \\ d_{2m} \end{bmatrix}, \dots, \begin{bmatrix} d_{r1} \\ d_{r2} \\ \vdots \\ d_{rm} \end{bmatrix}$ form a basis for the column space of A precisely when $w'_1, w'_2, \dots, w'_r \in im(f)$, where $w'_i = \sum_{j=1}^{m} d_{ij} w_j$, form a basis for $im(f)$.

Note further that under this isomorphism, the columns of A, which form a spanning set of the column space of A, correspond to the images $f(v_i)$, which form a spanning set for $im(f)$.

The rank-nullity theorem for linear transformations

Let $T : V \to W$ be a linear transformation.
The rank of T (denoted $rank(T)$) is the dimension of $Im(T)$.
The nullity of T (denoted $nullity(T)$) is the dimension of $ker(T)$.
Reinterpreting the rank-nullity theorem for matrices, we obtain :
$rank(T) + nullity(T) = dim(V)$

Week 7 P1

L7.1: Equivalence and similarity of matrices

Equivalence of matrices

Let A and B be two matrices of order $m \times n$. We say A is **equivalent** to B if $B = QAP$ for some invertible $n \times n$ matrix P and for some invertible $m \times m$ matrix Q.

Other characteristics:

1) A can be transformed into B by a combination of elementary row and column operations.
2) $rank(A) = rank(B)$

Equivalence of matrices is an **equivalence relation** i.e.
⇒ A is equivalent to itself
⇒ A is equivalent to B implies B is equivalent to A.
⇒ A is equivalent to B and B to C implies A is equivalent to C.

Linear transformations and equivalence of matrices

Considere a linear transformation $T: V \to W$, two ordered bases $\beta_1$and β_2 for V, and two ordered bases γ_1 and γ_2 for W.
Let A be the matrix corresponding to T with respect to the bases β_1 and γ_1 and B be the matrix corresponding to T with respect to the bases β_2 and γ_2.
Then A is equivalent to B !

$$B = QAP$$

Similar matrices

An $n \times n$ matrix A is similar to an $n \times n$ matrix B if there exists an $n \times n$ invertible matrix P such that $B = P^{-1}AP$.
Note that similarity is an equivalence relation, i.e. :
⇒ A is similar to itself
⇒ A is similar to B implies B is similar to A.
⇒ A is similar to B and B to C implies A is similar to C.

Important properties of similar matrices

Suppose A and B are similar matrices. Then the following properties hold :
⇒ A and B are equivalent.
⇒ A and B have the same rank.

$\Rightarrow det(B) = det(P^{-1}AP) = det(P^{-1})det(A)det(P) = \frac{1}{det(P)}det(A)det(P) = det(A)$.

Week 7 P2

⇒ Several other invariants of A and B are the same such as the characteristic polynomial, minimal polynomial and eigenvalues (with multiplicity)

Linear transformations and similarity of matrices

Consider a linear transformation $T: V \rightarrow W$ and two ordered bases β and γ for V.
Let A be the matrix corresponding to T with respect to the basis β and B the matrix corresponding to T with respect to the basis γ.
Then A is similar to B !

$$B = P^{-1}AP$$

Why do we care about similarity ? Because under some basis, we hope that the corresponding matrix is a diagonal matrix which gives an easy geometric understanding of the linear transformation.

L7.2: Affine subspaces and affine mappings

Affine Subspaces

Let V be a vector space. An affine subspace of V is a subset L such that there exists $v \in V$ and a vector subspace $U \subseteq V$ such that
$L = v + U := \{v + u | u \in U\}$.
We say an affine subspace L is n-dimensional if the corresponding subspace U is n-dimensional.
The subspace U corresponding to an affine subspace is unique.
However the vector v is not unique and in fact can be any vector in L.
Affine subspaces are thus **translates** of a vector subspace of V.

Affine Subspaces in $\mathbb{R}^2$

⇒ Points
⇒ Lines
⇒ the entire plane $\mathbb{R}^2$
A subset which is not an affine subspace : the parabola $y = x^2 + 1$ or the curve $y^2 = x^3$.

Week 7 P3

Affine Subspaces in $\mathbb{R}^3$

⇒ Points
⇒ Lines
⇒ Planes
⇒ the entire space $\mathbb{R}^3$
Example: Two-dimensional affine subspaces in $\mathbb{R}^3$ can expressed as
$l = v + \lambda_1 v_1 + \lambda_2 v_2$ where $\lambda_1, \lambda_2 \in \mathbb{R}$ and v , v_1, v_2 are vectors in $\mathbb{R}^3$

The solution set to a system of linear equations

Let $Ax = b$ be a linear system of equations.
⇒ $b = 0$: In this case, it is a homogeneous system and as seen before , the solution set is subspace of $\mathbb{R}^n$, namely the null space $\eta(A)$ of A.
⇒ $b \notin$ column space of A : In this case, $Ax = b$ does not have a solution, so the solution set is the empty set.
⇒ $b \in$ column space of A : In this case, the solution set L is an affine subspace of $\mathbb{R}^n$.
Specifically it can be described as $L = v + \eta(A)$ where v is any solution of the equation $Ax = b$

Affine mappings of affine subspaces

Let L and $L`$ be affine subspaces of $V$ and $W$ respectively. Let $f: L \to L`$ be a function. Consider any vector $v \in L$ and the unique subspace $U \subseteq V$ such that $L = v + U$. Note that $f(v) \in L`$ and hence $L` = f(v) + U`$ where $U`$ is the unique subspace of W corresponding to $L`$.Then $f$ is an affine mapping from $L$ to $L`$ if the function $g: U \to U`$ defined by $g(u) = f(u + v) - f(v)$ is a linear transformation.
For a linear transformation $T: U \to U`$ and fixed vectors $v \in L$ and $v` \in L`$, an affine mapping $f$ can be obtained by defining $f(v + u) = v` + T(u)$, and in fact every affine mapping is obtained in this way.

An example and an important special case

Let $T(x, y, z) = (2x + 3y + 2, 4x - 5y + 3)$. Then this is an affine mapping from $\mathbb{R}^3$ to $\mathbb{R}^2$.
Let $T : V \to W$ be a linear transformation and $w \in W$, then the mapping

$T` : V \to W$

$T`(v) = w + T(v)$

Is an affine mapping from V to W.

Week 7 P4

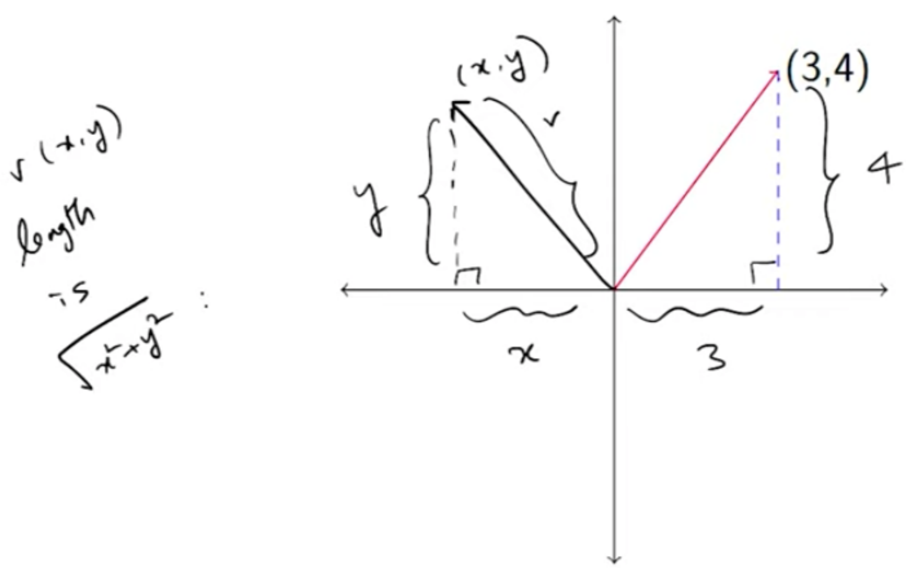
L7.3: Lengths and angles

The dot product of two vectors in $\mathbb{R}^2$

Consider the two vectors (3,4) and (2,7) in $\mathbb{R}^2$. The dot product of these two vectors gives us a scalar as follows:
$(3, 4).(2, 7) = 3 \times 2 + 4 \times 7 = 6 + 28 = 34$
For two general vectors (x_1, y_1) and (x_2, y_2) in $\mathbb{R}^2$, the dot product of these two vectors is the sclar computed as follows :
$(x_1, y_1).(x_2, y_2) = x_1x_2 + y_1y_2$.

The length of a vector in $\mathbb{R}^2$

Let us find the length of the vector $(3, 4)$ in $\mathbb{R}^2$.



Using Pythagoras' theorem, the length of the vector $(3, 4)$ is $\sqrt{3^2 + 4^2} = 5$ units.

Week 7 P5

The relation between length and dot product in $\mathbb{R}^2$

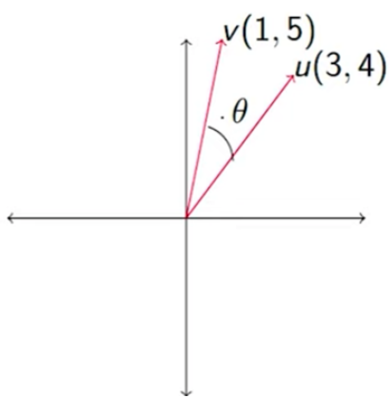
Observe that $(3, 4).(3, 4) = 3^2 + 4^2$, and hence the length of $(3, 4)$ is the square root of the dot product of the vector with itself.

Length of the vector $(3, 4) = \sqrt{(3, 4).(3, 4)} = \sqrt{3^2 + 4^2} = 5$

More generally, the length of the vector $(x, y) \in \mathbb{R}^2$ is $\sqrt{x^2 + y^2}$ = $\sqrt{(x, y).(x, y)}$

The angle between two vectors in $\mathbb{R}^2$

⇒ The angle between the vectors u and v and measures how far the direction is of v from u (or vice versa). e.g. θ is the angle between $u = (3, 4)$ and $v = (1, 5)$.



⇒It is measured in degrees (between 0 and 360) or radians (between 0 and 2Π).
⇒ The angle is often described by computing its trigonometric functions (e.g, sin, cos, tan).

The dot product and the angle between two vectors in $\mathbb{R}^2$

Let u and v be two vectors in $\mathbb{R}^2$. Then we can compute the angle θ between the vectors u and v using the dot products as :

$cos(\theta) = \frac{u.v}{\sqrt{(v.v)\times(u.u)}}$

The angle between two vectors in $\mathbb{R}^3$ and the dot product

The angle between the vectors u and v in $\mathbb{R}^3$ is the angle between them computed by passing a plane through them. (same logic as in $\mathbb{R}^2$)
It measures how far the direction is of v from u (or vice versa) on that plane.

Week 7 P6

L7.4: Inner products and norms on a vector space

Inner product on a vector space

An inner product on a vector space V is a function $\langle .,. \rangle : V \times V \rightarrow \mathbb{R}$ satisfying the following :
⇒ $\langle v, v \rangle > 0$ for all $v \in V \backslash \{0\}$; $\langle v, v \rangle = 0$ iff $v = 0$.
⇒ $\langle v_1 + v_2, v_3 \rangle = \langle v_1, v_3 \rangle + \langle v_2, v_3 \rangle$
⇒ $\langle v_1, v_2 \rangle = \langle v_2, v_1 \rangle$
⇒ $\langle cv_1, v_2 \rangle = c\langle v_2, v_1 \rangle = \langle v_1, cv_2 \rangle$ where $c \in \mathbb{R}$
A vector space V together with an inner product $\langle .,. \rangle$ is called an inner product space.

The dot product is an example of an inner product

Recall that the dot product of $u = (u_1, u_2,.., u_n)$ and $v = (v_1, v_2,..., v_n)$ be in $\mathbb{R}^n$ is

$$u.v = u_1 v_1 + u_2 v_2 +...+ u_n v_n.$$

This yield a function

$$\langle .,. \rangle : V \times V \rightarrow \mathbb{R} \,;\, \langle u, v \rangle = u.v$$

An example of an inner product on $\mathbb{R}^2$

The following is an example of an inner product on $\mathbb{R}^2$:

$$\langle .,. \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$$

$$\langle v, v \rangle = x_1 y_1 - (x_1 y_2 + x_2 y_1) + 2x_2 y_2$$

Where $u = (x_1, x_2)$ and $v = (y_1, y_2)$ be in $\mathbb{R}^2$

Norm on a vector space

A norm on a vector space V is a function

$$||.|| : V \rightarrow \mathbb{R}$$
$$x \rightarrow ||x||$$

Satisfying the following conditions:
⇒ $||x + y|| \leq ||x|| + ||y||$, for all $x, y \in V$
⇒ $||cx|| = |c|\, ||x||\, for\ all\ c \in \mathbb{R}$ and for all $x \in V$
⇒ $||x|| \geq 0$ for all $x \in V$; $||x|| = 0$ iff $x = 0$

Week 7 P7

Length as an example of a norm

Recall that the length of a vector $u = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$ is

$||u|| = \sqrt{(x_1^2 + x_2^2 + ... + x_n^2)}$

The length function $\mathbb{R}^n \to \mathbb{R}$ is a norm on $\mathbb{R}^n$.

An example of a norm on $\mathbb{R}^n$

The following is an example of a norm on $\mathbb{R}^n$:

Define $||u|| = |x_1| + |x_2| + .. |x_n|$ for $u = (x_1, x_2, .., x_n) \in \mathbb{R}^n$.

The inner product induces a norm

Let V be an inner product space with inner product $\langle .,. \rangle$.

Then the function $||.|| : V \to \mathbb{R}$ define by $||v|| = \sqrt{\langle v, v \rangle}$ is a norm of V.

Week 8 P1

L8.1: Orthogonality and linear independence

The geometric intuition of orthogonal vectors

If the angle θ between two vectors u and v in $\mathbb{R}^n$ is a right angle (i.e 90^o), then
$cos(\theta) = 0 = \frac{u.v}{||u||||v||}$ then $u.v = 0$.
e.g. $(1, 2, 3)$ and $(2, 2, -2)$ are orthogonal.

Orthogonal vectors

Two vectors u and v of an inner product space V are said to be orthogonal if $\langle u, v\rangle = 0$.
e.g. consider $\mathbb{R}^2$ with the inner product
$\langle u, v\rangle = x_1y_1 - (x_1y_2 + x_2y_1) + 2x_2y_2$ where $u = (x_1, x_2)$ and $v = (y_1, y_2)$.
Then the vectors $(1, 1)$ and $(1, 0)$ are orthogonal (w.r.t this inner product).

An orthogonal set of vectors

An orthogonal set of vectors of an inner product space V is a set of vectors whose elements are mutually orthogonal.
Explicitly, if $S = \{v_1, v_2,..., v_k\} \subseteq V$, then S is an orthogonal set of vectors if
$\langle v_i, v_j\rangle = 0$ for $i, j \in \{1, 2,..., k\}$ and $i \neq j$.
e.g. consider $\mathbb{R}^3$ with the usual inner product i.e. the dot product. Then the set
$S = \{(4, 3, -2), (-3, 2, -3), (-5, 18, 17)\}$ is an orthogonal set of vectors.

Orthogonality and linear independence

Let $\{v_1, v_2,..., v_k\}$ be an orthogonal set of vectors in the inner product space V.
Then $\{v_1, v_2,..., v_k\}$ is a linearly independent set of vectors.

What is an orthogonal basis

Let V be an inner product space. A basis consisting of mutually orthogonal vectors is called an orthogonal basis.
Since an orthogonal set of vectors is already linearly independent, an orthogonal set is a basis precisely when it is a maximal orthogonal set (i.e. there is no orthogonal set strictly containing this one).
If $dim(V) = n$, then

$$orthogonal\ basis \equiv orthogonal\ set\ of\ n\ vectors.$$

Week 8 P2

Example of orthogonal bases :

1. The standard basis
2. $\{(4, 3, -2), (-3, 2, -3), (-5, 18, 17)\} \subseteq \mathbb{R}^3$.
3. Consider $\mathbb{R}^2$ with the inner product

$\langle (x_1, x_2), (y_1, y_2)\rangle = x_1 y_1 - (x_1 y_2 + x_2 y_1) + 2x_2 y_2$.

Then $\{(1, 1), (1, 0)\}$ is an orthogonal basis.

L8.2: What is an orthonormal basis?

An orthonormal set of vectors of an inner product space V is an orthogonal set of vectors such that the norm of each vector of the set is 1.

Explicitly , if $S = \{v_1, v_2, ..., v_k\} \subseteq V$, then S is an orthonormal set of vectors if

$\langle v_i, v_j\rangle = 0$ for $i, j \in \{1, 2, ..., k\}$ and $i \neq j$.

And $||v_i|| = 1 \ \forall i \in \{1, 2, .., k\}$

e.g. consider $\mathbb{R}^4$ with the usual inner product i.e. the dot product. Then the set

$\{(\frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, 0), (\frac{2}{\sqrt{42}}, \frac{1}{\sqrt{42}}, \frac{1}{\sqrt{42}}, \frac{6}{\sqrt{42}}), (\frac{2}{3}, 0, \frac{2}{3}, \frac{-1}{3})\}$ is an orthogonal set of vectors.

What is an orthonormal basis?

An orthonormal basis is an orthonormal set of vectors which forms a basis.

Equivalently : An orthonormal basis is an orthogonal basis where the norm of each vector is 1.

Equivalently : An orthonormal basis is a maximal orthonormal set.

Example : The standard basis w.r.t the usual inner product forms an orthonormal basis.

Obtaining orthonormal sets from orthogonal sets

Let V be an inner product space. If $\Gamma = \{v_1, v_2, ..., v_k\}$ is an orthogonal set of vectors, then we can obtain an orthonormal set of vectors β from Γ by

$\beta = \{\frac{v_1}{||v_1||}, \frac{v_2}{||v_2||}, ..., \frac{v_k}{||v_k||}\}$.

Example : Consider $\mathbb{R}^2$ with the usual inner product and the orthogonal basis

$\Gamma = \{(1, 3), (-3, 1)\}$

Then $\beta = \{\frac{1}{\sqrt{10}}(1, 3), \frac{1}{\sqrt{10}}(-3, 1)\}$ is an orthonormal basis of $\mathbb{R}^2$.

Week 8 P3

Why are orthonormal bases important?

Suppose $\Gamma = \{v_1, v_2,..., v_k\}$ is an orthonormal basis of an inner product space V and let $v \in V$.
Then v can be written as $v = c_1 v_1 + c_1 v_1 +... + c_n v_n$.
How do we find $c_1, c_2,..., c_n$? For any basis, this means writing a system of linear equations and solving it.
But since Γ is orthonormal , we can use the inner product and compute $c_i = \langle v , v_j \rangle$.

L8.3: Projections using inner products

Shortest distance in $\mathbb{R}^2$

A and B are points in the plane $\mathbb{R}^2$ and we want to find the nearest point from B on the line passing through A and the origin. Drop a perpendicular from B on to the line. Let a and b be the vectors corresponding to the points A and B respectively.

The projection of a vector to a subspace

Let V be an inner product space, $v \in V$ and $W \subseteq V$ be a subspace. Then the projection of v onto W is the vector in W, denoted by $projw(v)$, computed as follows :
Find an orthonormal basis $\{v_1, v_2,..., v_k\}$ for W.

Define $projw(v) = \sum_{i=1}^{n} \langle v , v_i \rangle v_i$.
Fact : The definition is independent of the chosen orthonormal basis (i.e. the expression on the RHS does not change even if you choose a different orthonormal basis).
The projection of v onto W is the vector in W closest to v. Note that "closest" is in terms of the distance based on the norm induced by the inner product.

Projection on a vector and orthogonal bases

Let V be an inner product space and $v, w \in V$. Define $proj_w(v) = proj_{<w>}(v)$.

Note that an orthonormal basis for $\langle w \rangle$ is $\frac{w}{||w||}$ and hence

$$proj_w(v) = \langle v, \frac{w}{||w||} \rangle \frac{w}{||w||} = \frac{\langle v,w \rangle}{||w||^2} w = \frac{\langle v,w \rangle}{\langle w,w \rangle} w.$$

Similarly, if $\{v_1, v_2,..., v_n\}$ is an orthogonal basis for a subspace W, then $\{\frac{v_1}{||v_1||}, \frac{v_2}{||v_2||},..., \frac{v_n}{||v_n||}\}$ is an orthonormal basis for W and hence

Week 8 P4

$$proj_w(v) = \sum_{i=1}^{n} \langle v, \frac{v_i}{||v_i||} \rangle \frac{v_i}{||v_i||} = \sum_{i=1}^{n} \frac{\langle v, v_i \rangle}{\langle v_i, v_i \rangle} = \sum_{i=1}^{n} proj_{v_i}(v).$$

Projection as a linear transformation

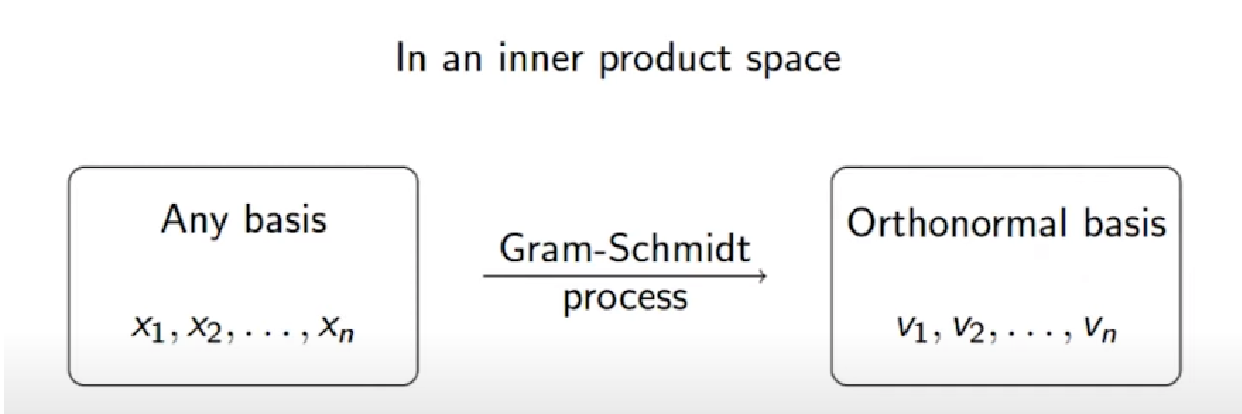
Let V be an inner product space and W be a subspace.
Then the projection of vectors in V to W is a linear transformation from V to W with image W.
Denote this linear transformation as P_W.

Some properties of the projection P_w

The linear transformation P_W has some interesting properties (some of which actually characterize it) :
i) $P_W(v) = v$, for all $v \in W$
ii) $Img(P_W) = W$
iii) $W^{\perp} = \{v | v \in V,\ such\ that\ \langle v, w \rangle = 0\ \forall\ w \in W\}$ is the null space of P_W
iv) $P_W^2 = P_W$
v) $||P_W(v)|| \leq ||v||$.

L8.4: The Gram-Schmidt process

An overview of the Gram-Schmidt process



Week 8 P5

Example and intuition

Consider the basis $\beta = \{(1, 2, 2), (-1, 0, 2), (0, 0, 1)\}$ for $\mathbb{R}^3$. Can we use this to obtain an orthonormal basis for $\mathbb{R}^3$?

Let $v_1 = (1, 2, 2)$. We want a vector which is orthogonal to v_1, i.e. a vector in $\langle v_1 \rangle^{\perp}$, so we use the projection P_{v1} to v_1.

Define $v_2 = (-1, 0, 2) - P_{v1}((-1, 0, 2)) = (-\frac{4}{3}, -\frac{2}{3}, \frac{4}{3})$

We want a vector which is orthogonal to both v_1 and v_2, i.e. a vector in $Span(\{v_1, v_2\})^{\perp}$, so we use the projection $P_{Span(\{v_1, v_2\})}$ to $Span(\{v_1, v_2\})$.

Define $v_3 = (0, 0, 1) - P_{v_1}((0, 0, 1)) - P_{v_2}((0, 0, 1)) = (\frac{2}{9}, -\frac{2}{9}, \frac{1}{9})$

Thus $\{v_1, v_3, v_3\}$ is an orthogonal basis and dividing each vector by its norms yields an orthonormal basis $\{(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}), (-\frac{2}{3}, -\frac{1}{3}, \frac{2}{3}), (\frac{2}{3}, -\frac{2}{3}, \frac{1}{3})\}$.

The Gram-Schmidt process

Let V be an inner product space with a basis $\{x_1, x_2, .., x_n\}$.

Define the orthogonal basis $\{v_1, v_2, .., v_n\}$ and the corresponding orthonormal basis $\{w_1, w_2, .., w_n\}$ as follows :

$v_1 = x_1;\qquad w_1 = \frac{v_1}{||v_2||}$

$v_2 = x_2 - \langle x_2, w_1 \rangle\ ;\qquad w_2 = \frac{v_1}{||v_2||}$

: : :

: : :

$v_i = x_i - \langle x_i, w_1 \rangle w_1 - \langle x_i, w_2 \rangle w_2 - ... - \langle x_i, w_{i-1} \rangle w_{i-1};\qquad w_i = \frac{v_i}{||v_i||}$

: : :

$v_n = x_n - \langle x_n, w_1 \rangle w_1 - \langle x_n, w_2 \rangle w_2 - ... - \langle x_n, w_{n-1} \rangle w_{n-1};\qquad w_n = \frac{v_n}{||v_n||}$

Main Take-homes

Theorem: Any finite-dimensional vector space with an inner product has an orthonormal basis. Any basis can be changed to an orthonormal basis using the Gram-Schmidt process.

Week 8 P6

L8.5: Orthogonal transformations and rotations

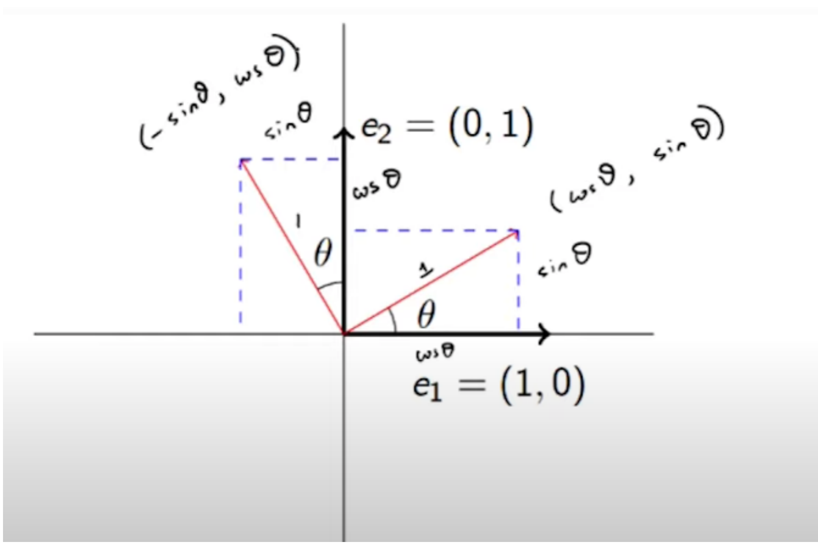
Let V be an inner product space and T be a linear transformation from V to V. T is said to be orthogonal transformation if

$$\langle T_V, T_W \rangle = \langle v, w \rangle \quad \forall\, v, w \in V.$$

When $V = \mathbb{R}^n$ with the usual inner product, a linear transformation $T: \mathbb{R}^n \to \mathbb{R}^n$ is orthogonal if it preserves angles and lengths.
Fact : It is enough to demand that the linear transformation preserves lengths. In that case, angles automatically get preserved (think of triangle congruences).

Finding the rotation matrix in $\mathbb{R}^2$

Consider the standard basis {(1,0),(0,1)} of $\mathbb{R}^2$. Rotate the plane by an angle θ. The vectors obtained after rotation tell us the matrix corresponding to this linear transformation.



Let T_θ be the corresponding linear transformation. Then $T_\theta(1, 0) = (cos(\theta), sin(\theta))$ and $T_\theta(0, -1) = (-sin(\theta), cos(\theta))$. Thus the matrix corresponding to this linear transformation is

$$R_\theta = \begin{pmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{pmatrix}.$$

Note $R_\theta^T = R_{-\theta}$ and $R_\theta^T R_\theta = R_\theta R_\theta^T = I$

Week 8 P7

Further note that since angles and lengths are preserved and the standard basis is orthonormal, the rotated vectors are also orthonormal and therefore yield and orthonormal basis of $\mathbb{R}^2$.

Rotations in $\mathbb{R}^3$

Consider the rotations about the axes in $\mathbb{R}^3$. Since these clearly preserve angles and distances and are linear transformations, they are orthogonal transformations.

Rotations about the axes can be described by considering its effect on the standard basis $\{e_1, e_2, e_3\}$.

When considering the rotation about the Z-axis, e_3 remains unchanged and the XY-plane gets rotated exactly as in the previous case of $\mathbb{R}^2$. Therefore its matrix is

RE VIDEOS

$$T_3(\theta) = \begin{pmatrix} cos(\theta) & -sin(\theta) & 0 \\ sin(\theta) & cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Similarly, the matrix corresponding to rotation about the X-axis is

$$T_1(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & cos(\theta) & -sin(\theta) \\ 0 & sin(\theta) & cos(\theta) \end{pmatrix}$$

and the matrix corresponding to rotation about the Y-axis is

$$T_2(\theta) = \begin{pmatrix} cos(\theta) & 0 & -sin(\theta) \\ 0 & 1 & 0 \\ sin(\theta) & 0 & cos(\theta) \end{pmatrix}.$$

Notice : $T_i(\theta)^T = T_i(-\theta)$ and $T_i(\theta)^T T_i(\theta) = T_i(\theta) T_i(\theta)^T = I$.

Week 8 P8

Another example of an orthogonal transformation

Let us define a linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$, where

$$T(x_1, x_2, x_3) = \frac{1}{3}(x_1 - 2x_2 + 2x_3, 2x_1 - x_2 - 2x_3, 2x_1 + 2x_2 + x_3).$$

Then evaluating T on the standard basis $\{e_1, e_2, e_3\}$ yields :

$$T(e_1) = v_1 = \left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right)$$
$$T(e_2) = v_2 = \left(-\frac{2}{3}, -\frac{1}{3}, \frac{2}{3}\right)$$
$$T(e_3) = v_3 = \left(\frac{2}{3}, -\frac{2}{3}, \frac{1}{3}\right) \quad .$$

Thus, the matrix corresponding to T is $A = \frac{1}{3}\begin{pmatrix} 1 & -2 & 2 \\ 2 & -1 & -2 \\ 2 & 2 & 1 \end{pmatrix}$.

Orthogonal matrices

As $\{v_1, v_2, v_3\}$ is an orthonormal set, the linear transformation T is an orthogonal transformation.

Observe that $AA^T = A^T A = I_3$.

A square matrix A is called an orthogonal matrix if $AA^T = A^T A = I_3$