UNSW Mathematics Society presents

MATH1231/1241 Revision Seminar

(Higher) Mathematics 1B

Algebra

T2, 2020

Presented by Joanna Lin and Karen Zhang

Table of Contents I

Table of Contents II

Vector Spaces

Vector Spaces

Vector Space Definition

In the following definition, we assume that u, v and w are elements of a particular vector space. λ and μ are scalars in the field over which the vector space is defined.
A vector space V over a field $\mathbb{F}$ is a non-empty set in which addition of vectors and scalar multiplication are defined in such a way that the following axioms are satisfied:

1. **Closure under Addition**.
2. **Associative Law of Addition**: $(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$.
3. **Commutative Law of Addition**: $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$.
4. **Existence of a Zero**: There exists a $\mathbf{0} \in V$ such that $\mathbf{u}+\mathbf{0}=\mathbf{u}$.
5. **Existence of a Negative**: For each $\mathbf{u} \in V$, there exists a $\mathbf{v} \in V$ where $\mathbf{u}+\mathbf{v}=\mathbf{0}$.

Vector Space Definition (continued)

6. **Closure under Scalar Multiplication**.
7. **Associative Law of Scalar Multiplication**: $(\lambda\mu)\mathbf{u} = \lambda(\mu\mathbf{u})$.
8. **Multiplication by One**: For the scalar $1 \in \mathbb{F}$, $1\mathbf{u} = \mathbf{u}$.
9. **Scalar Distributive Law**: $(\lambda + \mu)\mathbf{u} = \lambda\mathbf{u} + \mu\mathbf{v}$.
10. **Vector Distributive Law**: $\lambda(\mathbf{u} + \mathbf{v}) = \lambda\mathbf{u} + \lambda\mathbf{v}$.

Examples and non-examples of vector spaces

Examples of vector spaces include: $\mathbb{R}^n$, $\mathbb{C}^n$, set of all polynomials and set $\mathbb{M}_{n,m}$ of real matrices.

Non-examples of vector spaces include: $\mathbb{R}^n$ defined over $\mathbb{C}$.

Subspaces

Here, we assume that S and V share the same field of scalars, as well as rules for addition and multiplication by scalars.

Subspace Definition

If S is a subset of a vector space V and is itself a vector space, then S is a subspace of V.

Subspace Theorem

A subset S of a vector space V is a subspace if and only if S

1. contains the zero vector,
2. is closed under addition and
3. is closed under scalar multiplication.

Using the Subspace Theorem

MATH1231 NOVEMBER 2011 Q1(ii)

Suppose A is a fixed matrix in $\mathbb{M}_{m,n}$. Apply the subspace theorem to show that

$$S = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\}$$

is a subspace of $\mathbb{R}^n$.

MATH1231 NOVEMBER 2011 Q1(ii)

$$S = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\}.$$

S is a subset of $\mathbb{R}^n$.
Zero vector. As $A\mathbf{0} = \mathbf{0}$, $\mathbf{0} \in S$.
Closure under addition. Let $\mathbf{x}, \mathbf{x}' \in S$. Then $A\mathbf{x} = \mathbf{0}$ and $A\mathbf{x}' = \mathbf{0}$. Now, $A(\mathbf{x} + \mathbf{x}') = A\mathbf{x} + A\mathbf{x}' = \mathbf{0} + \mathbf{0} = \mathbf{0}$ and so $\mathbf{x} + \mathbf{x}' \in S$. Hence, S is closed under addition.
Closure under scalar multiplication. Let $\mathbf{x} \in S$ and $\lambda \in \mathbb{R}$. Then $A\mathbf{x} = \mathbf{0}$. Now, $A(\lambda\mathbf{x}) = \lambda(A\mathbf{x}) = \lambda(\mathbf{0}) = \mathbf{0}$ and so $\lambda\mathbf{x} \in S$. Hence, S is also closed under scalar multiplication.
Hence, S is a subspace of $\mathbb{R}^n$ by the Subspace Theorem.

MATH1231 NOVEMBER 2018 Q1(i)

Prove that

$$S = \{\mathbf{x} \in \mathbb{R}^3 : x_1^2 = x_2 x_3\}$$

is **not** a subspace of $\mathbb{R}^3$.

MATH1231 NOVEMBER 2018 Q1(i)

$$S = \{\mathbf{x} \in \mathbb{R}^3 : x_1^2 = x_2 x_3\}$$

We see that the zero vector satisfies the equation restricting S. The equation is homogeneous, so it should be closed under scalar multiplication. Instead, we'll test a value for addition.

Let $\mathbf{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ and $\mathbf{x}' = \begin{pmatrix} 2 \\ 4 \\ 1 \end{pmatrix}$. Now $1^2 = 1 = 1 \times 1$ and $2^2 = 4 = 4 \times 1$ and so $\mathbf{x}, \mathbf{x}' \in S$. Now $\mathbf{x} + \mathbf{x}' = \begin{pmatrix} 1+2 \\ 1+4 \\ 1+1 \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \\ 2 \end{pmatrix}$, but $3^2 = 9 \neq 5 \times 2$ and so $\mathbf{x} + \mathbf{x}' \notin S$. Hence, S is not closed under addition, and so S is not a subspace of $\mathbb{R}^3$.

Spans

Note that when we talk about **linear combinations** and the span of a set of vectors, it is implicit that the set of vectors is **finite**.

Span is a Subspace

If S is a finite and non-empty set of vectors in a vector space V, then $\text{span}(S)$ is a subspace of V.
Furthermore, $\text{span}(S)$ is the smallest subspace containing S. This means $\text{span}(S)$ is a subspace of every subspace which contains S.

S is referred to as a **spanning set** of V if $\text{span}(S) = V$.

Testing Whether a Set Spans a Vector Space

To check whether a set spans a particular vector space.

1. Write out a matrix whose columns are the vectors in the set.
2. Perform row reduction until the matrix is in row-echelon form.
3. Check whether there are any **zero rows** in the matrix.

If there are zero rows, then the vectors do not span the vector space. If there are no zero rows, then the vectors span the vector space.

MATH1231 FEBRUARY 2012 Q1(ii)

Consider the set $S = \{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}, \mathbf{v_4}\}$ of vectors in $\mathbb{R}^4$ given by

$$S = \left\{ \begin{pmatrix} 1 \\ -2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -4 \\ 2 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 5 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ -4 \\ 8 \\ 7 \end{pmatrix} \right\}$$

and let $\mathbf{b} = \begin{pmatrix} -2 \\ 1 \\ -10 \\ -7 \end{pmatrix}$

1. Does $\mathbf{b}$ belong to span(S)? Give reasons.
2. Does S span $\mathbb{R}^4$? Give reasons.

MATH1231 FEBRUARY 2012 Q1(ii)

1. Consider the matrix

$$(A|\mathbf{b}) = \left(\begin{array}{cccc|c} 1 & 2 & 1 & 3 & -2 \\ -2 & -4 & 0 & -4 & 1 \\ 1 & 2 & 5 & 8 & -10 \\ 2 & 4 & 0 & 7 & 7 \end{array}\right) \longrightarrow \left(\begin{array}{cccc|c} 1 & 2 & 1 & 3 & -2 \\ 0 & 0 & 2 & 2 & -3 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

   The system represented by $(A|\mathbf{b})$ has at least one solution. Thus $\mathbf{b}$ is in span(S).

2. Since the row echelon form of A has a zero row, there will be some vectors in $\mathbb{R}^4$ for which there is an inconsistent row of zeroes on the left and a non-zero value on the right. Hence, S does not span $\mathbb{R}^4$.

$$\left(\begin{array}{cccc|c} 1 & 2 & 1 & 3 & b_1 \\ 0 & 0 & 2 & 2 & b_2 + 2b_1 \\ 0 & 0 & 0 & 1 & b_3 - 5b_1 - 2b_2 \\ 0 & 0 & 0 & 0 & b_4 + 7b_2 - 3b_3 + 15b_1 \end{array}\right)$$

Linear Independence

Linear Independence and Dependence

The set $S = \{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$, a subset of a vector space, is a **linearly independent** set if the only values of the scalars $\lambda_1, \ldots, \lambda_n$ for which

$$\lambda_1\mathbf{v_1} + \cdots + \lambda_n\mathbf{v_n} = \mathbf{0} \text{ are } \lambda_1 = \lambda_2 = \cdots = \lambda_n = 0.$$

Otherwise, the set is **linearly dependent**. That is, there exists $\lambda_1, \ldots, \lambda_2$ not all zero satisfying the above equation.

When proving things involving linear independence, start with these definitions of linear independence and dependence.

Intuitive Idea

Previously, we defined linear independence and dependence, but this theorem perhaps provides a more intuitive idea of ‘dependency’.

Linear Independence

A set of vectors S is linearly independent if and only if no vector in S can be written as a linear combination of the other vectors in S.

Linear Dependence

A set of vectors S is linearly dependent if and only if at least one vector in S is in the span of the other vectors in S.

Solving Linear Independence Problems

MATH1241 T2 2019 Q1(f)(i)

Let $B = \{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ be a basis for $\mathbb{R}^3$. Let

$$\begin{aligned} S_1 &= \{\mathbf{u} - \mathbf{v}\}, \\ S_2 &= \{\mathbf{u} - \mathbf{v}, \mathbf{v} - \mathbf{w}\}, \\ S_3 &= \{\mathbf{u} - \mathbf{v}, \mathbf{v} - \mathbf{w}, \mathbf{w} - \mathbf{u}\}. \end{aligned}$$

For each of S_1, S_2 and S_3, determine, with reasons, whether they are linearly independent.

MATH1241 T2 2019 Q1(f)(i)

$B = \{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly independent.

$$S_1 = \{\mathbf{u} - \mathbf{v}\}.$$

For S_1, consider the equation

$$\lambda_1(\mathbf{u} - \mathbf{v}) = \mathbf{0}.$$

We want to ensure that $\mathbf{u} - \mathbf{v}$ is not the zero vector.
Since $\mathbf{u}$ and $\mathbf{v}$ are linearly independent, one cannot be written as a linear combination of the other. Hence, $\mathbf{u} \neq \mathbf{v}$ and so $\mathbf{u} - \mathbf{v} \neq \mathbf{0}$.
The only choice of λ_1 is 0 and so S_1 is linearly independent.

Continued...

$B = \{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly independent.

$$S_2 = \{\mathbf{u} - \mathbf{v}, \mathbf{v} - \mathbf{w}\}.$$

For S_2, consider the equation

$$\lambda_1(\mathbf{u} - \mathbf{v}) + \lambda_2(\mathbf{v} - \mathbf{w}) = \mathbf{0}$$

which rearranges to

$$\lambda_1\mathbf{u} + (\lambda_2 - \lambda_1)\mathbf{v} + (-\lambda_2)\mathbf{w} = \mathbf{0}.$$

Since B is linearly independent, $\lambda_1 = \lambda_2 - \lambda_1 = -\lambda_2 = 0$. Hence $\lambda_1 = \lambda_2 = 0$ only, and so S_2 is linearly independent.

Continued...

$B = \{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly independent.

$$S_3 = \{\mathbf{u} - \mathbf{v}, \mathbf{v} - \mathbf{w}, \mathbf{w} - \mathbf{u}\}.$$

For S_3, consider

$$\lambda_1(\mathbf{u} - \mathbf{v}) + \lambda_2(\mathbf{v} - \mathbf{w}) + \lambda_3(\mathbf{w} - \mathbf{u}) = \mathbf{0}$$

which rearranges to

$$(\lambda_1 - \lambda_3)\mathbf{u} + (\lambda_2 - \lambda_1)\mathbf{v} + (\lambda_3 - \lambda_2)\mathbf{w} = \mathbf{0}.$$

Since B is linearly independent $\lambda_1 - \lambda_2 = \lambda_2 - \lambda_2 = \lambda_3 - \lambda_2 = 0$ and so $\lambda_1 = \lambda_2 = \lambda_3$. Letting $\lambda_1 = \lambda_2 = \lambda_3 = 1$, we observe that $(\mathbf{u} - \mathbf{v}) + (\mathbf{v} - \mathbf{w}) + (\mathbf{w} - \mathbf{u}) = \mathbf{0}$ and so $\lambda_1 = \lambda_2 = \lambda_3 = 0$ is not the only solution. Hence, S_3 is linearly dependent.

Testing Linear Independence

To prove that a set of vectors is linearly independent or dependent

1. Write out a matrix whose columns are the vectors in the set.
2. Perform row reduction until the matrix is in row-echelon form.
3. Check whether there are any **non-leading columns**.

If there are non-leading columns, then the vectors are linearly dependent. If there are no non-leading columns, then the vectors are linearly independent.
Alternatively, if you have a square matrix, take the determinant of the matrix. If the determinant is 0, then the column vectors are linearly dependent. Otherwise, it's linearly independent.

Spanning Sets and Linearly Independent Sets

Zero Rows vs Non-Leading Columns

With span problems, you are checking for zero rows and **with linear independence problems, you are checking for non-leading columns**.

For example, if a matrix is in row-echelon form, and there are zero rows and no non-leading columns, the set of vectors used to construct the original is still linearly independent. Consider:

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 5 \\ 1 & 2 & 4 \\ 1 & 2 & 3 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

The set of vectors forming the columns on the left are linearly independent (no non-leading columns), but they do not span $\mathbb{R}^4$.

Proving Linear Dependence/Independence

MATH1231 NOVEMBER 2016 Q3(ii)(Modified)

Consider the set $S = \{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_2}\}$ of vectors in $\mathbb{R}^3$, where

$$\mathbf{v_1} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \mathbf{v_2} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}, \mathbf{v_3} = \begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix}.$$

1. Prove that S is linearly dependent.
2. Write the last vector in S as a linear combination of the other two.
3. What is the relationship between $\text{span}(\mathbf{v_1}, \mathbf{v_2})$ and $\text{span}(\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3})$?

MATH1231 NOVEMBER 2016 Q3(ii)

1. Consider the matrix

$$A = \begin{pmatrix} \mathbf{v_1} & \mathbf{v_2} & \mathbf{v_3} \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 3 \\ 2 & 3 & 4 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

   Since there is a non-leading column in the above matrix, the homogeneous system $A\mathbf{x} = \mathbf{0}$ has solutions other than the zero vector. Hence, the set S is linearly dependent.

2. Suppose $\mathbf{v_3} = \lambda_1 \mathbf{v_1} + \lambda_2 \mathbf{v_2}$. Doing back substitution, $-\lambda_2 = 2 \implies \lambda_2 = -2$ and $\lambda_1 + 2\lambda_2 = 1 \iff \lambda_1 = 5$. Hence,

$$\begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix} = 5 \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + (-2) \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix}.$$

3. They are equal.

Bases and Dimension

Basis Definition

A set of vectors is a basis for vector space V if it is

1. an independent set and
2. a spanning set for V.

For a vector space V with a finite basis, every basis for V contains the same number of vectors. We can now define dimension.

Dimension Definition

The dimension of vector space V, denoted by $\dim(V)$, is the number of vectors in any basis for V.

Dimension of Subspaces

Note that the dimension of a subspace is often not equal to the dimension of the vector space that it is a subset of.

Dimension

Size of Independent Sets, Spanning Sets and Dimension

1. The number of vectors in any spanning set for vector space V is greater than or equal to the dimension of V.
2. The number of vectors in any linearly independent set in V is less than or equal to the dimension of V.

How to Confirm that a Set is a Basis

You can confirm that a set containing vectors in V is a basis of a vector space V if you know that it satisfies 2 of the following conditions:

1. The number of vectors in the set is equal to $\dim(V)$.
2. The set is linearly independent.
3. The set is a spanning set for V.

Constructing Bases

MATH1231 NOVEMBER 2011 Q1(i)

Consider a set of vectors $S = \{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}, \mathbf{v_4}\} \subset \mathbb{R}^3$.

1. Can S be a spanning set for $\mathbb{R}^3$? Give reason.
2. Will all such sets S be spanning sets? Give reason.
3. Suppose S consists of the vectors

$$\mathbf{v_1} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \mathbf{v_2} = \begin{pmatrix} 3 \\ 3 \\ 8 \end{pmatrix} \mathbf{v_3} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \mathbf{v_4} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

Determine a subset that forms a basis for $\mathbb{R}^3$.

MATH1231 NOVEMBER 2011 Q1(i)

1. Yes. The number of vectors in the set S exceeds 3, which is the number of vectors in a basis for $\mathbb{R}^3$. Some such sets S will contain a subset of 3 linearly independent vectors spanning $\mathbb{R}^3$.
2. No. S may not contain a subset of 3 linearly independent vectors, which is required to span $\mathbb{R}^3$.
3. Consider the matrix

$$A = \begin{pmatrix} 1 & 3 & 1 & -1 \\ 2 & 3 & -1 & 1 \\ 3 & 8 & 2 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 3 & 1 & -1 \\ 0 & -3 & -3 & 3 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

Hence $\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_4}\}$ is a basis for $\mathbb{R}^3$. *Note that we pick vectors out of the original set, not from the matrix in row echelon form.*

Bonus: Extending a set to a basis. Consider the set $W = \{\mathbf{v_1}, \mathbf{v_2}\}$ where $\mathbf{v_1} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ $\mathbf{v_2} = \begin{pmatrix} 3 \\ 3 \\ 8 \end{pmatrix}$ as in the previous question. Suppose we wanted to extend this to a basis. We augment matrix $B = \begin{pmatrix} \mathbf{v_1} & \mathbf{v_2} \end{pmatrix}$ with the standard basis vectors of $\mathbb{R}^3$.

$$\begin{pmatrix} 1 & 3 & 1 & 0 & 0 \\ 2 & 3 & 0 & 1 & 0 \\ 3 & 8 & 0 & 0 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 3 & 1 & 0 & 0 \\ 0 & -3 & -2 & 1 & 0 \\ 0 & 0 & -7 & -1 & 3 \end{pmatrix}.$$

Hence, a basis for $\mathbb{R}^3$ is $\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{e_1}\}$. Other times you will need to augment the matrix is when you have a set with more terms than a basis, but is still not a spanning set, and you are asked to find a basis with as many vectors from that set as possible.

Linear Transformations

Formal Definition of Linear Transformations

Linear Transformation/Map

Let V and W be two **vector spaces** over the same field $\mathbb{F}$. A function $T : V \to W$ is a linear transformation or linear map if it **preserves** addition and scalar multiplication.
Addition condition: $T(\mathbf{v} + \mathbf{v}') = T(\mathbf{v}) + T(\mathbf{v}')$.
Scalar multiplication condition: $T(\lambda \mathbf{v}) = \lambda T(\mathbf{v})$.

Key Things to Note in Definition

1. The domain and codomain of the linear transformation must be vector spaces. It is a good idea to address at the start of your proofs for linear transformations.
2. We don't refer to a transformation as being 'closed' under addition and scalar multiplication unlike vector spaces. Instead, say that it 'preserves' addition and scalar multiplication.

Proofs with Linear Transformations

MATH1241 NOVEMBER 2014 Q3(i)

Prove that the function $T : \mathbb{P}(\mathbb{R}) \to \mathbb{R}^2$ defined by

$$T(p) = \begin{pmatrix} p(0) \\ p(1) \end{pmatrix}, \quad \text{for all polynomials } p \in \mathbb{P}(\mathbb{R}),$$

is linear transformation.

MATH1241 NOVEMBER 2014 Q3(i)

The domain and codomain of T are vector spaces.

Addition condition. Let $p, q \in \mathbb{P}(\mathbb{R})$. Then

$$T(p+q) = \begin{pmatrix} (p+q)(0) \\ (p+q)(1) \end{pmatrix} = \begin{pmatrix} p(0)+q(0) \\ p(1)+q(1) \end{pmatrix}$$

$$= \begin{pmatrix} p(0) \\ p(1) \end{pmatrix} + \begin{pmatrix} q(0) \\ q(1) \end{pmatrix} = T(p) + T(q)$$

and so T preserves addition.

Scalar multiplication condition. Let $p \in \mathbb{P}(\mathbb{R})$ and $\lambda \in \mathbb{R}$. Then

$$T(\lambda p) = \begin{pmatrix} (\lambda p)(0) \\ (\lambda p)(1) \end{pmatrix} = \begin{pmatrix} \lambda \times p(0) \\ \lambda \times p(1) \end{pmatrix} = \lambda \begin{pmatrix} p(0) \\ p(1) \end{pmatrix} = \lambda T(p)$$

and so T also preserves scalar multiplication.

Hence, T is a linear map.

Alternative Proof

Combining the Addition and Scalar Multiplication Condition

A function $T : V \to W$ is a linear map if and only if

$$T(\lambda_1 \mathbf{v_1} + \lambda_2 \mathbf{v_2}) = \lambda_1 T(\mathbf{v_1}) + \lambda_2 T(\mathbf{v_2})$$

for all $\lambda_1, \lambda_2 \in \mathbb{F}$ and $\mathbf{v_1}, \mathbf{v_2} \in V$.

We can generalise the above theorem to:

$$T(\lambda_1 \mathbf{v_1} + \cdots + \lambda_n \mathbf{v_n}) = \lambda_1 T(\mathbf{v_1}) + \cdots + \lambda_n T(\mathbf{v_n})$$

which can be used to calculate function values. **Note** that the function value for any vector in the domain can be calculated if we know the function values for a basis.
Also **note** that, if $\{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ is a basis for the domain, then $\{T(\mathbf{v_1}), \ldots, T(\mathbf{v_n})\}$ spans the range (image).

MATH1231 NOVEMBER 2017 Q1(iii)(b)

Let $\mathbf{v_1} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \mathbf{v_2} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ and $\mathbf{v_3} = \begin{pmatrix} -2 \\ -1 \\ 1 \end{pmatrix}$.

We are given that $\mathbf{v_3} = -3\mathbf{v_1} + \mathbf{v_2}$.

Does there exist a linear map $T : \mathbb{R}^3 \to \mathbb{R}^3$ such that

$$T(\mathbf{v_1}) = \begin{pmatrix} 1 \\ 6 \\ 3 \end{pmatrix}, T(\mathbf{v_2}) = \begin{pmatrix} -2 \\ 16 \\ 2 \end{pmatrix} \text{ and } T(\mathbf{v_3}) = \begin{pmatrix} -6 \\ -3 \\ -8 \end{pmatrix}?$$

MATH1231 NOVEMBER 2017 Q1(iii)(b)

Assuming that such a linear map T exists,

$$T(\mathbf{v_3}) = T(-3\mathbf{v_1} + \mathbf{v_2}) = -3T(\mathbf{v_1}) + T(\mathbf{v_2}).$$

Now,

$$\text{RHS} = -3\begin{pmatrix} 1 \\ 6 \\ 3 \end{pmatrix} + \begin{pmatrix} -2 \\ 16 \\ 2 \end{pmatrix} = \begin{pmatrix} -5 \\ -2 \\ -7 \end{pmatrix}.$$

However

$$\text{LHS} = \begin{pmatrix} -6 \\ -3 \\ -8 \end{pmatrix} \neq \text{RHS}.$$

Hence, no such linear map T exists.

MATH1231 NOVEMBER 2011 Q2(iii)

Show that the map $T : \mathbb{R}^2 \to \mathbb{R}^3$, defined by

$$T(\mathbf{x}) = \begin{pmatrix} x_1 + x_2 \\ x_1 x_2 \\ -x_2 \end{pmatrix}, \quad \text{for } \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

is not linear.

Tip: If the transformed vector has a non-linear component (for the example above, this would be the 2nd component), then the map is non-linear. This isn't a proof but it's a good test to verify whether the map could be linear for yourself.

Use Counterexample to Prove a Transformation isn't Linear

When proving a map is not linear, you must provide a specific case where the map does not preserve addition or scalar multiplication.

MATH1231 NOVEMBER 2011 Q2(iii)

Consider the vector $\mathbf{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Then $T(\mathbf{v}) = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$ and so

$(-1)T(\mathbf{v}) = \begin{pmatrix} -2 \\ -1 \\ 1 \end{pmatrix}$. However,

$$T((-1)\mathbf{v}) = T\begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 + -1 \\ (-1)(-1) \\ -(-1) \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \neq (-1)T(\mathbf{v})$$

Hence, T does does not preserve scalar multiplication, and so T is not a linear map.

Matrices and Linear Maps

Matrix Multiplication is a Linear Transformation

For each $m \times n$ matrix A, the function $T_A : \mathbb{R}^n \to \mathbb{R}^m$, defined by

$$T_A(\mathbf{x}) = A\mathbf{x} \quad \text{for} \quad x \in \mathbb{R}^n,$$

is a linear map.

Matrix Representation Theorem

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map. Suppose vectors $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ are the standard basis vectors for $\mathbb{R}^n$. Then the jth column of the matrix A which has the property

$$T(\mathbf{x}) = A\mathbf{x}$$

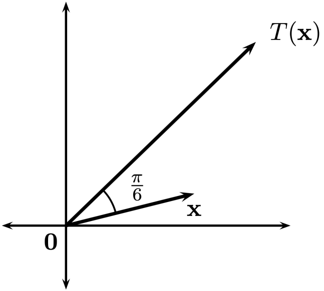
is given by $T(\mathbf{e}_j)$.

Geometric Example + Matrix Representation

MATH1231 NOVEMBER 2013 Q2(iv)

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear map which rotates a vector $\frac{\pi}{6}$ anticlockwise about the origin and doubles its length.

1. Show that $T(\mathbf{e_1}) = \begin{pmatrix} \sqrt{3} \\ 1 \end{pmatrix}$ where $\mathbf{e_1} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.
2. Find the matrix A such that $T(\mathbf{x}) = A\mathbf{x}$.



1. For the given transformation,
$$T(\mathbf{e_1}) = T\begin{pmatrix} 1\cos(0) \\ 1\sin(0) \end{pmatrix} = \begin{pmatrix} 2\cos(0+\frac{\pi}{6}) \\ 2\sin(0+\frac{\pi}{6}) \end{pmatrix} = \begin{pmatrix} \sqrt{3} \\ 1 \end{pmatrix}.$$
2. Similarly,
$$T(\mathbf{e_2}) = T\begin{pmatrix} 1\cos(\frac{\pi}{2}) \\ 1\sin(\frac{\pi}{2}) \end{pmatrix} = \begin{pmatrix} 2\cos(\frac{\pi}{2}+\frac{\pi}{6}) \\ 2\sin(\frac{\pi}{2}+\frac{\pi}{6}) \end{pmatrix} = \begin{pmatrix} -1 \\ \sqrt{3} \end{pmatrix}.$$

Hence, $A = \begin{pmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{pmatrix}$, by Matrix Representation Theorem.

For basic transformations, you can simply write the matrix form by 'factoring out' the vector, without using matrix representation theorem. For example,

$$T(\mathbf{x}) = \begin{pmatrix} x_2 \\ 2x_3 - x_1 \\ x_1 + x_2 + 5x_4 \\ 6x_2 - 3x_4 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 2 & 0 \\ 1 & 1 & 0 & 5 \\ 0 & 6 & 0 & -3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

Kernel

Kernel/Null Space Definition

The kernel of a linear transformation $T : V \to W$ is the set of all values in the domain V which map to the zero vector in the codomain. That is, the kernel is the subset of V defined by

$$\ker(T) = \{\mathbf{v} \in V : T(\mathbf{v}) = \mathbf{0}\}.$$

Similarly, for an $m \times n$ matrix A, the kernel is the subset of $\mathbb{R}^n$ defined by

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

The kernel of T is 'the set of all zeroes of T', similar to functions in calculus. The kernel of A is the set of all solutions of the **homogeneous equation** $A\mathbf{x} = \mathbf{0}$.

Geometric Interpretation of a Kernel

MATH1231 NOVEMBER 2012 Q2(vi) (Modified)

Suppose $\mathbf{b}$ is a non-zero vector in $\mathbb{R}^3$ and consider the projection map $T : \mathbb{R}^3 \to \mathbb{R}^3$, defined by $T(\mathbf{x}) = \dfrac{\mathbf{x} \cdot \mathbf{b}}{|\mathbf{b}|^2}\mathbf{b}$. You may assume that T is linear. Describe geometrically the kernel of T.

Note: Try out the proof that T is linear yourself.

The kernel consists of all vectors which, when projected onto $\mathbf{b}$, *results in a zero vector. That is, the vectors in the kernel must be perpendicular to the vector* $\mathbf{b}$ *because that is when no component of the vector is in the direction of* $\mathbf{b}$.

The kernel is the plane through the origin whose normal is $\mathbf{b}$.

Nullity

Before we define nullity, we must establish that:

A Kernel is a Vector Space

If $T : V \to W$ is a linear map, then $\ker(T)$ is a subspace of domain V.

Nullity

- The nullity of a linear map T is the dimension of $\ker(T)$.
- The nullity of a matrix A is the dimension of $\ker(A)$.

That is,

$$\text{nullity}(T) = \dim(\ker(T)) \quad \text{and} \quad \text{nullity}(A) = \dim(\ker(A)).$$

Image

Image

Let $T : V \to W$ be a linear map. Then the image of T is the set of all function values of T, that is, it is the subset of the codomain W defined by

$$\mathrm{im}(T) = \{\mathbf{w} \in W : \mathbf{w} = T(\mathbf{v}) \text{ for some } \mathbf{v} \in V\}.$$

Similarly, the image an $m \times n$ matrix A is

$$\mathrm{im}(A) = \{\mathbf{b} \in \mathbb{R}^m : \mathbf{b} = A\mathbf{x} \text{ for some } \mathbf{x} \in \mathbb{R}^n\}.$$

The image of a linear map corresponds to the range of a function as seen in calculus. Note that, for a matrix A, $\mathrm{im}(A) \equiv \mathrm{col}(A)$

Rank

Like the kernel...

An Image is a Vector Space

Let $T : V \to W$ be a linear map between vector spaces V and W. Then $\text{im}(T)$ is a subspace of the codomain W of T.

Now we can define the rank of a linear map and rank of a matrix...

Rank

- The rank of a linear map T is the dimension of $\text{im}(T)$.
- The rank of a matrix A is the dimension of $\text{im}(A)$.

That is,

$$\text{rank}(T) = \dim(\text{im}(T)) \quad \text{and} \quad \text{rank}(A) = \dim(\text{im}(A)).$$

Finding Rank, Nullity and Bases

Finding the Rank and Nullity

- $\text{nullity}(A) =$ number of non-leading columns in a row-echelon form U for A.
- $\text{rank}(A) =$ number of leading columns in row-echelon form U for A.

Finding a Basis for a Kernel vs Basis for an Image

- To find the basis of a kernel, we solve the homogeneous equation by doing back substitution on the matrix in row-echelon form against the zero vector.
- A basis for an image can be formed by extracting all column vectors from the original matrix corresponding to leading columns of the matrix in row-echelon form.

Rank-Nullity Theorem

Rank-Nullity Theorem

- Suppose V and W are finite dimensional vector spaces, and $T : V \to W$ is linear. Then

$$\text{rank}(T) + \text{nullity}(T) = \dim(V).$$

- For any matrix A

$$\text{rank}(A) + \text{nullity}(A) = \text{ number of columns of } A.$$

Since $\text{rank}(A) =$ number of leading columns and $\text{nullity}(A) =$ number of non-leading columns in row echelon form of A, we expect the result for matrices.

MATH1231 NOVEMBER 2011 Q1(iii)

A linear map $T : \mathbb{R}^m \to \mathbb{R}^n$ has rank k. State the nullity of T.

MATH1231 NOVEMBER 2011 Q2(iv)

Consider the matrix

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

The mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by

$$T(\mathbf{x}) = A\mathbf{x} - 2\mathbf{x}, \quad \text{for } x \in \mathbb{R}^3$$

1. Find the matrix B for T, such that $T(\mathbf{x}) = B\mathbf{x}$.
2. Explain why $\begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}$ belongs to the kernel of T.
3. Write down the rank and nullity of T.

MATH1231 NOVEMBER 2011 Q2(iv)

1. Note that $T(\mathbf{x}) = (A - 2I)\mathbf{x}$. Hence,

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

2. We test whether the vector maps to zero vector:

$$T\begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}\begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1(-3)+2(0)+3(1) \\ 1(-3)+2(0)+3(1) \\ 1(-3)+2(0)+3(1) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

By definition, $\begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}$ belongs to the kernel.

3. The rank and nullity of T is equal to that of B:

$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Hence, $\text{rank}(T) = 1$, $\text{nullity}(T) = 2$.

Bonus: Finding Basis for Image and Kernel

Kernel. Parameterising the variables in the non-leading columns, $x_1 + 2\lambda + 3\mu = 0 \iff x_1 = -2\lambda - 3\mu$. A vector in the kernel is given by:

$$\mathbf{x} = \begin{pmatrix} -2\lambda - 3\mu \\ \lambda \\ \mu \end{pmatrix} = \lambda \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}.$$

Hence, a basis for the kernel is $\left\{ \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} \right\}$.

Image. A basis is $\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$.

Eigenvalues and Eigenvectors

Definition

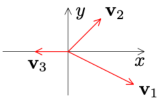
Let A be an $n \times n$ matrix, the scalar $\lambda \in \mathbb{F}$ and **v** a non-zero vector in $\mathbb{F}^n$. If

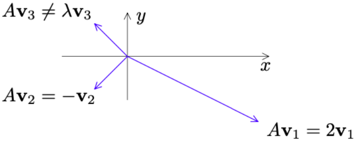
$$A\boldsymbol{v} = \lambda \boldsymbol{v},$$

then λ is the **eigenvalue** of A and **v** is the corresponding **eigenvector** of A.

By multiplying the matrix A to the vector **v**, we obtain a scalar of the same vector.

- Eigenvectors of A are vectors which point in the same/opposite direction when multiplied by A
- Eigenvalues are the ratios of lengths between A**v** and **v**





Computing eigenvalues and eigenvectors

Method

- λ is an eigenvalue of A iff $\det(A - \lambda I) = 0$
- **v** is an eigenvector of A that corresponds to λ iff $(A - \lambda I)\boldsymbol{v} = \mathbf{0}$, where $\boldsymbol{v} \neq \mathbf{0}$

Extra

- $\det(A - \lambda I) = 0$ is also known as the **characteristic polynomial** of degree n in λ.
- The **eigenspace** of A corresponding to λ is the set $E_\lambda = \{\boldsymbol{v} \in \mathbb{F}^n \mid A\boldsymbol{v} = \lambda\boldsymbol{v}\}$, and is a subspace of $\mathbb{F}^n$. (Proof: Subspace Theorem)

Eigenvalues and eigenvectors

Theorem 1: Existence of eigenvalues

An n x n matrix has exactly n eigenvalues if

- real and complex eigenvalues are counted
- eigenvalues are counted according to their multiplicity (i.e. $(\lambda - 1)^2 \longrightarrow \lambda = 1, 1)$

Sketch Proof: Consider the nth degree characteristic polynomial with exactly n roots.

Theorem 2: Independence of eigenvectors

If a matrix A has different eigenvalues for each corresponding eigenvector $\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_k}$, then the eigenvectors are linearly independent.

Diagonalisation

Definition

Given A is an n x n matrix with n **linearly independent** eigenvectors, there exists an invertible matrix M and a diagonal matrix D such that

$$A = MDM^{-1}$$

where the diagonal entries of D are the eigenvalues of A and the columns of M are the corresponding eigenvectors of A.

Theorem 3: Diagonalisation of matrices

A is diagonalisable iff it has n linearly independent eigenvectors.

For M to be invertible, the eigenvectors must be linearly independent

Example

1231 2018 Semester 2 Final Q2(v)

Consider the matrix $A = \begin{pmatrix} -1 & 3 \\ 2 & 0 \end{pmatrix}$.

1. Find the eigenvalues and eigenvectors of A.
2. Find a diagonal matrix D and matrix M such that $D = M^{-1}AM$.

Solutions

(1) For the matrix A, we have

$$\det(A-\lambda I) = \det\begin{pmatrix} -1-\lambda & 3 \\ 2 & 0-\lambda \end{pmatrix} = -\lambda(-1-\lambda)-2\times 3 = \lambda^2+\lambda-6.$$

By solving the characteristic equation $\lambda^2+\lambda-6=0$, we have $\lambda=-3$ and $\lambda=2$. To find the corresponding eigenvectors, we substitute the values for λ into $(A-\lambda I)\mathbf{v}=\mathbf{0}$. For $\lambda=-3$, we obtain $(A+3I)\mathbf{v}=\mathbf{0}$ which can be written as a system of linear equations and row reduced as follows:

$$\left(\begin{array}{cc|c} 2 & 3 & 0 \\ 2 & 3 & 0 \end{array}\right) \Longleftrightarrow \left(\begin{array}{cc|c} 2 & 3 & 0 \\ 0 & 0 & 0 \end{array}\right) \text{ so } v_2 = t,\ v_1 = -\frac{3}{2}t.$$

Therefore, the eigenvector for $\lambda=-3$ is

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} -\frac{3}{2}t \\ t \end{pmatrix} = t\begin{pmatrix} -3 \\ 2 \end{pmatrix}.$$

Solutions (contd.)

Similarly, for $\lambda = 2$ we have $(A - 2I)\mathbf{v} = \mathbf{0}$ and

$$\left(\begin{array}{cc|c} -3 & 3 & 0 \\ 2 & -2 & 0 \end{array}\right) \Longleftrightarrow \left(\begin{array}{cc|c} -3 & 3 & 0 \\ 0 & 0 & 0 \end{array}\right) \text{ so } \mathbf{v} = \begin{pmatrix} t \\ t \end{pmatrix} = t\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

(2) The diagonal matrix has the eigenvalues of A along the diagonal and zeroes otherwise;

$$D = \begin{pmatrix} -3 & 0 \\ 0 & 2 \end{pmatrix}.$$

M is a matrix whose columns are the eigenvectors of A;

$$M = \begin{pmatrix} -3 & 1 \\ 2 & 1 \end{pmatrix}.$$

Powers of Matrices

Evaluating powers of matrices

$$\begin{aligned} A^k &= \overbrace{MDM^{-1} \times MDM^{-1} \times MDM^{-1} \times \ldots \times MDM^{-1}}^{k \text{ times}} \\ &= MD^kM^{-1} \end{aligned}$$

Problems involving the powers of matrices can be simplified once we know the matrices D and M, since it's easy to calculate M^{-1} and D^k.

- $D^k = \begin{pmatrix} (\lambda_1)^k & 0 \\ 0 & (\lambda_2)^k \end{pmatrix}$ in the 2×2 case.

Diagonalisation

Corollary to Theorem 2: Matrices without repeated eigenvalues are diagonalisable

If A is an n x n matrix with n different eigenvalues, then A is diagonalisable.

Choose an eigenvector for each n eigenvalues $\implies$ eigenvectors must be independent (Theorem 2) $\implies$ A must be diagonalisable (Theorem 3).

Converse of Theorem 2 and Corollary

Note that the converse is false.

- Theorem 2: Linearly independent eigenvectors may have the same eigenvalue.
- Corollary: A diagonalisable matrix A may have repeated eigenvalues (does not have n different eigenvalues).

Example

1241 2017 Semester 2 Final Q3(ii)

Let $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}$.

1. Given that the eigenvalues of A are $1, 2, 3$, explain why A is diagonalisable.
2. Find an eigenvector for A for the eigenvalue $\lambda = 3$.
3. Let $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ and $f(x) = (x-1)(x-2)(x-3) = x^3 - 6x^2 + 11x - 6$. Show that $f(D) = D^3 - 6D^2 + 11D - 6I$ is the zero matrix.
4. Hence, prove that $f(A) = \mathbf{0}$.
5. Compute A^{-1} as a linear combination of A^2, A, I.

Solutions

(1) Since all three eigenvalues of A are different, A has independent eigenvectors (*Independence of Eigenvectors*). We also know that A is diagonalisable if and only if A has three independent eigenvectors. Hence, A is diagonalisable (*Diagonalisation of Matrices*).

(2) For $\lambda = 3$, we have $(A - 3\lambda)\mathbf{v} = \mathbf{0}$ which can be written as:

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

$$\mathbf{v} = \begin{pmatrix} \frac{1}{2}t \\ t \\ t \end{pmatrix} = t\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}.$$

Solutions (contd.)

Hence, the eigenvector for A that corresponds to $\lambda = 3$ is $\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$.

(3) Since $f(D) = D^3 - 6D^2 + 11D - 6I$, substituting $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ into $f(D)$ gives

$$f(D) = \begin{pmatrix} 1^3 & 0 & 0 \\ 0 & 2^3 & 0 \\ 0 & 0 & 3^3 \end{pmatrix} - 6 \begin{pmatrix} 1^2 & 0 & 0 \\ 0 & 2^2 & 0 \\ 0 & 0 & 3^2 \end{pmatrix} + 11 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} - 6 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Solutions (contd.)

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 27 \end{pmatrix} - \begin{pmatrix} 6 & 0 & 0 \\ 0 & 24 & 0 \\ 0 & 0 & 54 \end{pmatrix} + \begin{pmatrix} 11 & 0 & 0 \\ 0 & 22 & 0 \\ 0 & 0 & 33 \end{pmatrix} - \begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix}$$

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

Solutions (contd.)

(4) We know from part (1) that A is diagonalisable, so $A = MDM^{-1}$ where M is the invertible matrix with eigenvectors of A as its columns and D is the diagonal matrix with eigenvalues of A as entries. Therefore,

$$\begin{aligned}
f(A) &= f(MDM^{-1}) \\
&= (MDM^{-1})^3 - 6(MDM^{-1})^2 + 11(MDM^{-1}) - 6I \\
&= M(D^3)M^{-1} - M(6D^2)M^{-1} + M(11D)M^{-1} - M(6I)M^{-1} \\
&= M(D^3 - 6D^2 + 11D - 6I)M^{-1} \\
&= Mf(D)M^{-1} \\
&= \mathbf{0}
\end{aligned}$$

since $f(D) = \mathbf{0}$.

Solutions (contd.)

(5) From part (4), $f(A) = A^3 - 6A^2 + 11A - 6I = \mathbf{0}$. Rearranging this equation gives

$$
\begin{aligned}
I &= -\frac{1}{6}A^3 + A^2 - \frac{11}{6}A \\
A^{-1} &= -\frac{1}{6}A^2 + A - \frac{11}{6}I.
\end{aligned}
$$

Solving Systems of Differential Equations

Proposition

Given an n x n matrix A, $\mathbf{y}(t) = e^{\lambda t}\mathbf{v}$ is a non-zero solution of $\frac{d\mathbf{y}}{dt} = A\mathbf{y}$ iff A has eigenvalue λ and eigenvector $\mathbf{v}$.
The general solution to $\frac{d\mathbf{y}}{dt} = A\boldsymbol{y}$ will have the form

$$\mathbf{y}(t) = \alpha_1 e^{\lambda_1 t}\mathbf{v_1} + \alpha_2 e^{\lambda_2 t}\mathbf{v_2} + \cdots + \alpha_n e^{\lambda_n t}\mathbf{v_n},$$

with constants $\alpha_1, \alpha_2, \ldots, \alpha_n$, n eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and corresponding eigenvectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ of A.

Examples

First Order ODEs

Solve the system

$$\frac{dy_1}{dx} = 6y_1 + 2y_2$$
$$\frac{dy_2}{dx} = -2y_1 + y_2$$

subject to the initial conditions $y_1(0) = -5$ and $y_2(0) = 7$.

Higher Order ODEs

Solve the equation

$$\frac{d^2y}{dt^2} - 2\frac{dy}{dt} - 24y = 0.$$

Solutions

First Order ODEs. This system can be written as $\frac{d\mathbf{y}}{dt} = A\mathbf{y}$, where $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ and A is the coefficient matrix $\begin{pmatrix} 6 & 2 \\ -2 & 1 \end{pmatrix}$.

$$\det(A - \lambda I) = \lambda^2 - 7\lambda + 10 = (\lambda - 2)(\lambda - 5)$$

For $\lambda = 2, \mathbf{v} = \alpha \begin{pmatrix} 1 \\ -2 \end{pmatrix}$ and so $\mathbf{y} = e^{2t} \begin{pmatrix} 1 \\ -2 \end{pmatrix}$.

For $\lambda = 5, \mathbf{v} = \alpha \begin{pmatrix} 2 \\ -1 \end{pmatrix}$ and so $\mathbf{y} = e^{5t} \begin{pmatrix} 2 \\ -1 \end{pmatrix}$.

Hence, the general solution is
$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = Ce^{2t} \begin{pmatrix} 1 \\ -2 \end{pmatrix} + De^{5t} \begin{pmatrix} 2 \\ -1 \end{pmatrix}$, where C, D are constants.

Solutions (contd.)

Next, we need to find the values of the constant coefficients. By substituting the initial conditions into $y_1 = Ce^{2t} + 2De^{5t}$ and $y_2 = -2Ce^{2t} - De^{5t}$, we have

$$C + 2D = -5 \text{ and } -2C - D = 7$$

and solving simultaneously gives $C = -3$ and $D = -1$. Therefore, the solutions are

$$\begin{aligned} y_1 &= -3e^{2t} - 2e^{5t} \\ y_2 &= 6e^{2t} + e^{5t}. \end{aligned}$$

Solutions (contd.)

Higher Order ODEs. Let $y_1 = y$ and $y_2 = \frac{dy}{dt}$. So,

$$\frac{dy_1}{dt} = y_2$$
$$\frac{dy_2}{dt} = 24y_1 + 2y_2.$$

This system can be written as $\frac{d\mathbf{y}}{dt} = A\mathbf{y}$, where $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ and A is the coefficient matrix $\begin{pmatrix} 0 & 1 \\ 24 & 2 \end{pmatrix}$. Solving this coefficient matrix gives eigenvalues and eigenvectors

$$\lambda = 6, \mathbf{v} = \alpha \begin{pmatrix} 1 \\ 6 \end{pmatrix}$$

$$\lambda = -4, \mathbf{v} = \alpha \begin{pmatrix} 1 \\ -4 \end{pmatrix}.$$

Solutions (contd.)

Therefore, the general solution is

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = Ce^{6t}\begin{pmatrix} 1 \\ 6 \end{pmatrix} + De^{-4t}\begin{pmatrix} 1 \\ -4 \end{pmatrix},$$ where C, D are constants.

Note that we set the solution of the original equation to $y = y_1$, hence $y = Ce^{6t} + De^{-4}$.

Probability and Statistics

Probability

Definition

A probability on a sample space S is a function $P : S \to \mathbb{R}$ with the properties:

- $0 \leq P(A) \leq 1$ for all $A \subseteq S$
- $P(\varnothing) = 0$ and $P(S) = 1$
- $P(A \cup B) = P(A) + P(B)$ for all disjoint $A, B \subseteq S$

Some Fundamental Rules

- $P(A^c) = 1 - P(A)$ (Complement)
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- If $A \subseteq B$, then $P(A) \leq P(B)$
- If S is finite, then $\displaystyle\sum_{a \in S} P(\{a\}) = 1$

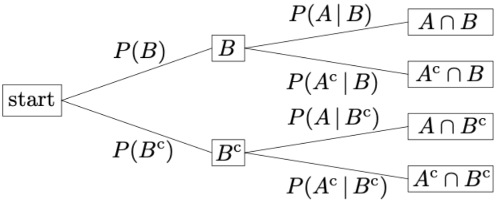
Conditional Probability

Definition

Let A and B be events in a sample space S such that $P(B) \neq 0$.

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Probability of A under the assumption that B has occurred.
Drawing tree diagrams help!



Or remember the following...

Conditional Probability

3 Important Rules

- Multiplication Rule:

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A).$$

If $B_1, B_2, \ldots B_n$ partition S, then

- Total Probability Rule:

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_n)P(B_n).$$

- Bayes' Rule:

$$P(B_k|A) = \frac{P(A|B_k)P(B_k)}{P(A)}.$$

Examples

1231 2019 T2 Final Q1(e)

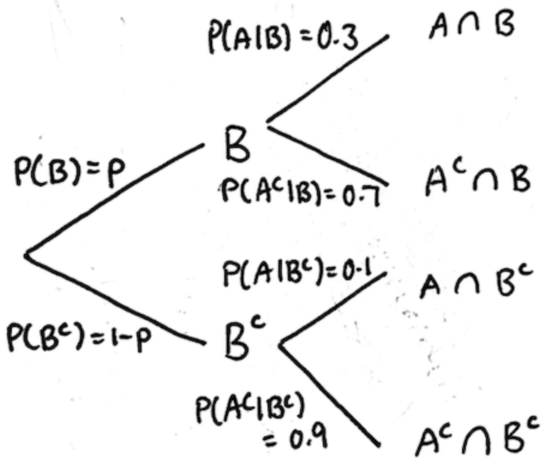
You are given the following information about the event A and B.

- $P(A) = 0.2$
- $P(B) = p$
- $P(A|B) = 0.3$
- $P(A^c|B^c) = 0.9$

1. Draw a tree diagram representing this information.
2. Find p.

Solutions

(1) Tree Diagram



Solutions (contd.)

(2) The total probability rule gives

$$\begin{aligned} P(A) &= P(A|B)P(B) + P(A|B^c)P(B^c) \\ &= 0.3p + 0.1(1-p) \\ &= 0.2p + 0.1. \end{aligned}$$

Since $P(A) = 0.2$,

$$\begin{aligned} 0.2 &= 0.2p + 0.1 \\ p &= \frac{0.1}{0.2} \\ &= 0.5. \end{aligned}$$

Examples

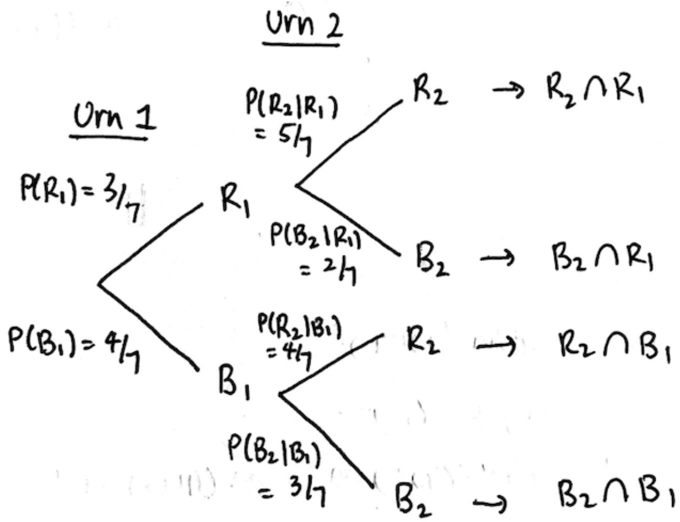
1241 2015 Semester 2 Final Q3(i)

Urn 1 contains 3 red balls and 4 blue balls. Urn 2 contains 4 red balls and 2 blue balls. A ball is drawn at random from Urn 1 and placed in Urn 2, and then a ball is drawn at random from the 7 balls now in Urn 2.

1. What is the probability that the ball drawn from Urn 2 is blue?
2. Given that the ball drawn from Urn 2 is red, what is the probability that the ball transferred was blue?
3. Let A be the event that the ball drawn from Urn 1 is blue and B be the event that the ball drawn from Urn 2 is blue. Are A and B statistically independent?

Solutions

Tree Diagram:



Solutions (contd.)

(1) By the total probability rule,

$$\begin{aligned} P(B_2) &= P(R_1)P(B_2|R_1) + P(B_1)P(B_2|B_1) \\ &= \frac{3}{7} \times \frac{2}{7} + \frac{4}{7} \times \frac{3}{7} \\ &= \frac{18}{49}. \end{aligned}$$

(2) By Bayes' Rule,

$$\begin{aligned} P(B_1|R_2) &= \frac{P(R_2 \cap B_1)}{P(R_2)} = \frac{P(B_1)P(R_2|B_1)}{P(B_1)P(R_2|B_1) + P(R_1)P(R_2|R_1)} \\ &= \frac{\frac{4}{7} \times \frac{4}{7}}{\frac{4}{7} \times \frac{4}{7} + \frac{3}{7} \times \frac{5}{7}} \\ &= \frac{16}{31}. \end{aligned}$$

Solutions (contd.)

(3) Intuitively, events A and B are not statistically independent because drawing a red or blue ball from Urn 1 (whether or not event A occurs) to place in Urn 2 will change the sample space of event B. Clearly,

$$P(A \cap B) = \frac{4}{7} \times \frac{3}{7} \neq \frac{4}{7} \times \frac{2}{6} = P(A)P(B).$$

Mutual Independence

Definition

Events A and B are mutually independent if

$$P(A \cap B) = P(A)P(B).$$

This concept can be extended to events $A_1, A_2, \ldots, A_k$ by taking any two or more of these events.

Similar to saying $P(A) = P(A|B)$ or $P(B) = P(B|A)$.

Random Variables

Definition

A random variable on a sample space S is a function $X : S \to \mathbb{R}$.

$$P(X = x) = P(\{s \in S \mid X(s) = x\}).$$

and similarly for $P(X \leq x), P(X \in A)$, etc.

E.g. Toss 2 dice, function X = sum the 2 dice values. Elements = $[2, 2], [1, 2], [5, 6]$ etc., and $S_X = \{2, 3, \ldots, 12\}$.

Discrete Random Variable

A random variable $X : S \to \mathbb{R}$ is discrete if its image $\{X(s) \mid s \in S\}$ is countable.

Cumulative Distribution Function

Definition

The cumulative distribution function of a random variable X is the function denoted by F_X, given as

$$F_X : \mathbb{R} \to \mathbb{R} \text{ where } F_X(x) = P(X \leq x).$$

Properties

- If $a \leq b$, then $F(a) \leq F(b)$ i.e. F is non-decreasing
- If $a \leq b$, then $P(a < X \leq b) = F(b) - F(a)$
- $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to \infty} F(x) = 1$

Probability Distribution

Definition

The probability distribution function of a discrete random variable X is a description of the probabilities of all events associated with X. Often denoted as

$$p_k = P(X = x_k),$$

where the range of $X = \{x_1, x_2, \dots\}$. A collection of real numbers p_k forms a probability distribution iff $p_k \geq 0 \ \forall \ k,$ and $\sum_{\text{all} k} p_k = 1.$

Note: The cumulative distribution function for a discrete random variable X is

$$F(x) = \sum_{k \leq x} p_k.$$

Expected Value

Definition

The mean/expected value of a discrete random variable X is

$$E(X) = \sum_{\text{all}\, k} x_k p_k.$$

Theorem: Expected value of a function of a random variable

Let X be a discrete random variable with values x_k and corresponding probabilities p_k. Let $g : \mathbb{R} \to \mathbb{R}$. The random variable $Y = g(X)$ has expected value

$$E(Y) = E(g(X)) = \sum_{\text{all}\, k} g(x_k) p_k.$$

Variance

Definition

The variable of a random discrete variable X with mean μ is

$$Var(X) = E((X-\mu)^2).$$

Note: Standard Deviation of X is given by $SD(X) = \sqrt{(Var(X))}$.
An alternative formula for variance (easier to calculate) is

$$Var(X) = E(X^2) - (E(X))^2.$$

Scaling and shifting of a random variable

Theorem

Given a and b are real constants,

- $E(aX + b) = aE(X) + b$
- $Var(aX + b) = a^2 Var(X)$
- $SD(aX + b) = |a| SD(X)$

Examples

1231 2019 T2 Final Q1(b)

Let X be a discrete random variable with the following probability distribution.

x	1	2	3	4	5
P(X=x)	0.1	0.3	0.1	p	q

Given that $E(X) = 3.3$,

1. Find the missing probabilities p and q.
2. Calculate $Var(X)$.

Solutions

(1) For probability distributions, we know that $\sum_{\text{all} k} p_k = 1$, hence

$$\begin{aligned} 0.1 + 0.3 + 0.1 + p + q &= 1 \\ p + q &= 0.5. \end{aligned}$$

Since we're given $E(X) = \sum_{\text{all} k} x_k p_k = 3.3$, we have

$$\begin{aligned} 1 \times 0.1 + 2 \times 0.3 + 3 \times 0.1 + 4 \times p + 5 \times q &= 3.3 \\ 4p + 5q &= 2.3. \end{aligned}$$

By solving simultaneous equations with $p + q = 0.5$ and $4p + 5q = 2.3$, we obtain $p = 0.2$ and $q = 0.3$.

Solutions (contd.)

(2) Substituting the table values into $Var(X) = E(X^2) - (E(X))^2$,

$$Var(X) = 1^2 \times 0.1 + 2^2 \times 0.3 + 3^2 \times 0.1 + 4^2 \times 0.2 + 5^2 \times 0.3 - 3.3^2 = 2.01.$$

Bernoulli Trials/Process

Definition

Bernoulli Trials are experiments with **two possible outcomes**, usually written as $p =$ **success** and $q = 1 - p =$ **failure**.
Bernoulli Processes consist of a sequence of identical Bernoulli Trials where the events S_k of a success on the kth trial are mutually independent.

Number of Successes in a Bernoulli Process

Let X be a random variable that gives the total number of successes in a Bernoulli Process consisting of n trials with probability $p =$ success. Then,

$$P(X = k) = \binom{n}{k} p^k q^{n-k} \text{ for } k = 0, 1, 2, \ldots, n.$$

Binomial Distribution

Definition

The binomial distribution with n trials and probability $p =$ success is the function

$$B(n, p, k) = \binom{n}{k} p^k (1-p)^{n-k} \text{ for } k = 0, 1, 2, \ldots, n.$$

Note that the binomial distribution forms a probability distribution.

Expected Value and Variance

If $X \sim B(n, p)$ then

- $E(X) = np$
- $Var(X) = npq$

Geometric Distribution

Definition

The geometric distribution with probability $p =$ success is the function

$$G(p, k) = (1-p)^{k-1}p \text{ for } k = 0, 1, 2, \ldots.$$

The first success occurs on the kth trial iff the first $k-1$ trials failed and the kth trial is a success. Note that the geometric distribution forms a probability distribution.

Expected Value and Variance

If $X \sim G(p)$ then

- $E(X) = \frac{1}{p}$
- $Var(X) = \frac{q}{p^2}$

Sign Tests

Sign Tests

Sign tests are used to decide whether a comparison between two sets of figures or a set of figures and a standard can be assumed as pure luck, or if there is evidence of a difference between the two sets or the set and standard.

algebra	7	7	8	3	9	10	8	10	9	7	7	2	9	8	9
calculus	4	7	7	9	10	9	9	7	5	6	4	1	8	8	7

Number of students who did better at algebra $= X \sim B\left(13, \frac{1}{2}\right)$.

$P(X \geq 10) = \sum_{k=10}^{13} \binom{13}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{13-k} \approx 0.0461$. Since $4.61\% < 5\%$, it's very unlikely that the difference was pure luck. Hence, the students are better at algebra than calculus.

Probability Density Function

Continuous Random Variable

A random variable is continuous iff its cumulative function $F_X(x)$ is continuous.

- Variable can take possible values that cover an interval of the real line

Definition

If X is a continuous variable with cumulative function $F(x)$, the probability density function is defined by

- $f(x) = F'(x)$ if F is differentiable at x
- $f(x) = \lim_{x \to a^-} F'(x)$ if F is not differentiable at $x = a$

Probability Density Function

Properties

Consider a continuous random variable X with cumulative distribution function $F(x)$ and probability density function $f(x)$. Then,

- $f(x) \geq 0,\ x \in \mathbb{R}$
- $F(x) = \displaystyle\int_{-\infty}^{x} f(t)dt$
- $\displaystyle\int_{-\infty}^{\infty} f(t)dt = 1$
- if $a \leq b$, then $P(a \leq X \leq b) = P(a < X \leq b) = \displaystyle\int_{a}^{b} f(x)dx$

Probability Density Function

Expected Value

Expected value of a continuous random variable $=$
$E(X) = \int_{-\infty}^{\infty} xf(x)dx$, where $f(x)$ is the probability density function.
Expected value of a **function** of a random variable $=$
$E(Y) = \int_{-\infty}^{\infty} g(x)f(x)dx$, where $Y = g(X)$, $g : \mathbb{R} \to \mathbb{R}$, and $f(x)$ is the probability density function.

Var, SD and scaling/shifting

Similar to discrete random variable.

Examples

1231 2017 Semester 2 Final Q2(iii)

The probability density function f of a continuous random variable X is given by

$$f(x) = \begin{cases} kx^2 & \text{for } 0 \le x \le 3 \\ 0 & \text{otherwise,} \end{cases}$$

where k is a constant.

1. Find the value of k.
2. Evaluate $E(X)$ and $Var(X)$.

1241 example: 2016 Semester 2 Q3(iv)

Solutions

(1) Since f is a probability density function, $\displaystyle\int_{-\infty}^{\infty} f(x)dx = 1$. So,

$$\int_{-\infty}^{\infty} f(x)dx = \int_{0}^{3} kx^2 dx = \left[\frac{kx^3}{3}\right]_0^3 = 9k = 1.$$

Therefore, $k = \frac{1}{9}$.

Solutions (contd.)

(2) Using the definition $E(X) = \int_{-\infty}^{\infty} xf(x)dx$, we have

$$E(X) = \int_0^3 \frac{1}{9}x^3 dx = \left[\frac{x^4}{36}\right]_0^3 = \frac{9}{4}.$$

Using the definition $Var(X) = E(X^2) - (E(X))^2$, first find $E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)dx$. So,

$$E(X^2) = \int_{-\infty}^{\infty} \frac{1}{9}x^4 dx = \left[\frac{x^5}{45}\right]_0^3 = \frac{27}{5}.$$

Hence, $Var(X) = E(X^2) - (E(X))^2 = \frac{27}{5} - (\frac{9}{4})^2 = \frac{27}{80}$.

Probability Density Function

Standardisation

If $E(X) = \mu$ and $Var(X) = \sigma^2$, then the random variable $Z = \frac{X-\mu}{\sigma}$ is the standardised random variable obtained from X.

- Variable has mean $E(Z) = 0$ and $Var(Z) = 1$

Normal Distribution

Definition

A continuous random variable with probability density function

$$\phi(x) : \mathbb{R} \to \mathbb{R} \text{ where } \phi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

has normal distribution with mean μ and variance σ^2.

- Written as $X \sim N(\mu, \sigma^2)$

Standard normal distribution has $\mu = 0$ and $\sigma = 1$ ($X \sim N(0, 1)$).

Normal Distribution

To calculate normal probabilities

Finding the normal probability involves integrating the probability density function, however this cannot be done in terms of elementary functions. So, we reduce to the standard normal distribution by using the change of variable $Z = \frac{X-\mu}{\sigma}$,

$$
\begin{aligned}
P(X \leq x) &= P\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) \\
&= P\left(Z \leq \frac{x-\mu}{\sigma}\right) \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\mu}{\sigma}} e^{-\frac{z^2}{2}}\,dz
\end{aligned}
$$

The value of this integral for various z can be found on the table.

Examples

1231 2018 Semester 2 Final Q1(iii)

Suppose the final mark for a student in MATH1131 is approximately normally distributed with mean 59 and standard deviation 7.44. Given that the pass mark is 50, what percentage of MATH1131 students are expected to pass?

Solutions

Let X be the mark for a student in MATH1131. We're given that $\mu = 59$ and $\sigma = 7.44$, so

$$\begin{aligned} P(X \geq 50) &= P\left(\frac{X-\mu}{\sigma} \geq \frac{50-59}{7.44}\right) \\ &= P(Z \geq -1.21) \\ &= 1 - P(Z \leq -1.21) \\ &= 1 - 0.1131 \\ &= 0.8869. \end{aligned}$$

Therefore, 88.69% of students are expected to pass.

Examples

1241 2019 T2 Final Q2(a)

The time (in seconds) of a group of athletes running the 400m are found to be approximately normally distributed with mean 48.7 and variance 1.9. Estimate the value of t such that the fastest 2% of these athletes achieve a time of t seconds or better.

Solutions

Let X be the time of the group of athletes running in 400m. In this question, we are given $\mu = 48.7$, $\sigma = \sqrt{1.9} = 1.378$ and $P(X \geq t) = 0.02$. So,

$$\begin{aligned} P(X \geq t) &= P\left(\frac{X-\mu}{\sigma} \geq \frac{t-48.7}{\sqrt{1.9}}\right) \\ &= P\left(Z \geq \frac{t-48.7}{\sqrt{1.9}}\right) \\ &= 1 - P\left(Z \leq \frac{t-48.7}{\sqrt{1.9}}\right) = 0.02. \end{aligned}$$

Hence, $P\left(Z \leq \dfrac{t-48.7}{\sqrt{1.9}}\right) = 0.98.$

Solutions (contd.)

From the table of standard normal probabilities, $P(Z \leq 2.06) = 0.9803$, so

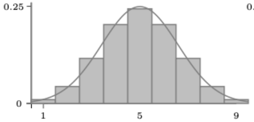
$$\begin{aligned} \frac{t - 48.7}{\sqrt{1.9}} &= 2.06 \\ t &= \sqrt{1.9} \times 2.06 + 48.7 \\ &= 51.540. \end{aligned}$$

Hence, t is estimated to be 51.540 seconds.

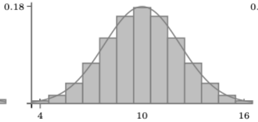
Normal Approximation to the Binomial Distribution

Using Normal Distributions

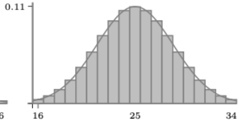
Normal distributions are used to approximate the binomial distribution $B(n, p)$ for large values of n. Generally, normal distributions are used to model experiments involving a large number of identical and independent trials with several possible outcomes.



$B(10, \frac{1}{2})$ and $N(5, \frac{5}{2})$



$B(20, \frac{1}{2})$ and $N(10, 5)$



$B(50, \frac{1}{2})$ and $N(25, \frac{25}{2})$

Normal Approximation to the Binomial Distribution

Continuity Correction

A continuity correction is an adjustment that is made when a continuous distribution (e.g. Normal) is used to approximate a discrete distribution (e.g. Binomial).

Binomial Distribution		Normal Approximation		Notes
P(x = c)	P(x = 10)	P(c - 0.5 < x < c + 0.5)	P(9.5 < x < 10.5)	Includes c
P(x > c)	P(x > 10)	P(x > c + 0.5)	P(x > 10.5)	Does not include c
P(x <= c)	P(x <= 10)	P(x < c + 0.5)	P(x < 10.5)	Includes c
P(x < c)	P(x < 10)	P(x < c - 0.5)	P(x < 9.5)	Does not include c
P(x >= c)	P(x >= 10)	P(x > c - 0.5)	P(x > 9.5)	Includes c
P(a < x < b)	P(9 < x < 11)	P(a - 0.5 < x < b + 0.5)	P(8.5 < x < 11.5)	

NOTE: For continuous random variables, $<$ and $\leq$ are the same since continuous random variables cannot equal any specific value.

Examples

1231 2015 Semester 2 Q2(iv)

Denver attempted an online examination in which he answered 50 multiple choice questions. In each question, there were 5 choices with only 1 correct answer. Denver chose all the answers randomly. Let X be the random variable counting the number of questions he guessed correctly.

1. Find $E(X)$ and $Var(X)$.
2. Use the normal approximation to the binomial distribution to find the probability that Denver correctly guessed 12 or more questions.

Solutions

(1) From the question, we can see that this is a binomial distribution with $n = 50$ trials and success probability $p = \frac{1}{5}$ and failure probability $q = 1 - p = \frac{4}{5}$. Hence, if $X \sim B(50, \frac{1}{5})$ then

$$E(X) = np = 50 \times \frac{1}{5} = 10 \text{ and } Var(X) = npq = 50 \times \frac{1}{5} \times \frac{4}{5} = 8.$$

Solutions (contd.)

(2) X is normally distributed with expected value $\mu = 10$ and standard deviation $\sigma = \sqrt{Var(X)} = \sqrt{8}$. Therefore, using the approximation $Y \sim N(10, 8)$

$$\begin{aligned}
P(X \geq 12) &\simeq P(Y \geq 11.5) \\
&= P\left(\frac{Y - 10}{\sqrt{8}} \geq \frac{11.5 - 10}{\sqrt{8}}\right) \\
&= P(Z \geq 0.53) \\
&= 1 - P(Z \leq 0.53) \\
&= 1 - 0.7019 \\
&= 0.2981.
\end{aligned}$$

[X] Exponential Distribution

Definition

A continuous random variable T with probability density function

$$f : \mathbb{R} \to \mathbb{R} \text{ where } f(t) = \begin{cases} \lambda e^{-\lambda t} & \text{if } t > 0 \\ 0 & \text{if } t \leq 0 \end{cases}$$

has exponential distribution with parameter λ.

- $T \sim Exp(\lambda)$

Expected value and Variance

- $E(T) = \frac{1}{\lambda}$
- $Var(T) = \frac{1}{\lambda^2}$

Examples

Exponential Distribution

Suppose that on a highway, cars pass at an average rate of 5 cars per minute. Assume that the duration of time between successive cars follows the exponential distribution.

1. On average, how many seconds elapse between two successive cars?
2. Find the probability that after a car passes by, the next car will pass within the next 30 seconds.
3. Find the probability that after a car passes by, the next car will not pass for at least another 20 seconds.

Solutions

The exponential distribution can be used to model waiting times where events may occur at any time, not specific to integer times.

(1) Since cars pass at an average rate of 5 cars/min, $\frac{60}{5} = 12$ seconds should elapse between each car on average.

(2) Let T be the time (seconds) between successive cars. From part (1), we know that the average or $E(\lambda) = \frac{1}{\lambda} = 12$. By rearranging, we have our parameter $\lambda = \frac{1}{12}$, and so $T \sim Exp(\frac{1}{12})$. The cumulative distribution function of T is found by integrating the probability density function,

$$P(T \leq t) = F(t) = \lambda \int_0^t e^{-\lambda x} dx = 1 - e^{-\lambda t}, t > 0.$$

So, $P(T \leq 30) = 1 - e^{-\frac{30}{12}} \approx 0.9179$.

Solutions (contd.)

(3) For the probability that the next car will not pass for at least another 20 seconds,

$$\begin{aligned} P(X \geq 20) &= 1 - P(X \leq 20) \\ &= 1 - (1 - e^{-\frac{20}{12}}) \\ &= e^{-\frac{20}{12}} \\ &\approx 0.1889. \end{aligned}$$

Supplementary

Field

Field Definition

In the following definition, we assume that x, y and z are elements of a particular field.
A field $\mathbb{F}$ is a non-empty set for which rules of addition and multiplication are defined and all of the following axioms are satisfied:

1. **Closure under Addition**.
2. **Associative Law of Addition**: $(x + y) + z = x + (y + z)$.
3. **Commutative Law of Addition**: $x + y = y + x$.
4. **Existence of a Zero**: There exists a $0 \in \mathbb{F}$ such that $x + 0 = 0 + x = x$.
5. **Existence of a Negative**: For each $x \in \mathbb{F}$, there exists a $y \in \mathbb{F}$ where $x + y = 0$.

Field Definition (continued)

6. **Closure under Multiplication**.
7. **Associative Law of Multiplication**: $(xy)z = x(yz)$.
8. **Commutative Law of Multiplication**: $xy = yx$.
9. **Existence of a One**: There exists a $1 \in \mathbb{F}$ such that $1x = x1 = x$.
10. **Existence of an Inverse**: For each $x \in \mathbb{F}$, there exists a $y \in \mathbb{F}$ where $xy = 1$.
11. **(Left) Distributive Law**: $x(y+z) = xy + xz$.
12. **(Right) Distributive Law**: $(x+y)z = xz + yz$.

Examples and non-examples of fields

Examples of fields are: $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ and $\mathbb{Z}_2$.
Non-examples of fields are: $\mathbb{N}$ and $\mathbb{Z}$.

Properties of Vector Spaces

Properties Derived from Axioms

1. **Uniqueness of zero.**
2. **Cancellation property**: $\mathbf{u} + \mathbf{w} = \mathbf{u} + \mathbf{v} \implies \mathbf{w} = \mathbf{v}$.
3. **Uniqueness of negatives.**
4. **Multiplication by zero scalar**: $0\mathbf{v} = \mathbf{0}$.
5. **Multiplication by zero vector**: $\lambda\mathbf{0} = \mathbf{0}$.
6. **Negative of zero**: $-\mathbf{0} = \mathbf{0}$.
7. **Multiplication by -1**: $(-1)\mathbf{v} = -\mathbf{v}$.
8. **Zero products**: $\lambda\mathbf{v} = \mathbf{0} \implies \lambda = 0$ and/or $\mathbf{v} = 0$.
9. **Cancellation property**: $\lambda\mathbf{v} = \mu\mathbf{v} \implies \lambda = \mu$ if $\mathbf{v} \neq \mathbf{0}$.

Linear Independence/Dependence Properties

Uniqueness of Linear Combinations

Suppose $\mathbf{v} \in \text{span}(S)$. The values of the scalars in the linear combination for a vector $\mathbf{v}$ are unique if and only if S is a linearly independent set.

Adding Vectors to Linearly Independent Sets

If S is a linearly independent set in a vector space V and $\mathbf{v} \in V$ but not in $\text{span}(S)$ then adding $\mathbf{v}$ to S results in a linearly independent set.

Adding Vectors to Sets

If a vector is added to a set, then the span of the new set is equal to the span of the original set if and only if the additional vector is in the span of the original set. *The new set is linearly dependent.*

Dropping Vectors from Linearly Dependent Sets

If S is a linearly dependent set of vectors, then it is possible to drop at least one vector from S to obtain a new set with the same span as S.

Dropping Vectors from Linearly Independent Sets

If S is a linearly independent set, then dropping any vector from S results in a smaller span than $\text{span}(S)$.

'At Least One'

In the theorems we've just discussed, 'at least one' does not mean that the statement is true for each and every vector in the set!

Constructing Bases: General Method

Extending a Set to a Basis

Suppose A is a matrix whose columns are from the linearly independent set $S = \{\mathbf{v_1}, \dots, \mathbf{v_n}\}$ which is a subset of $\mathbb{R}^m$, followed by the standard basis vectors of $\mathbb{R}^m$. To create a basis for $\mathbb{R}^m$, choose the vectors from S and the standard basis vectors which correspond to the leading columns in the row-echelon form of A.

Note: If a set is linearly dependent, then you can simply extract a linearly independent set of vectors from the set to apply the above theorem.

Reducing a Spanning Set to a Basis

Suppose A is a matrix whose columns are from the set $S = \{\mathbf{v_1}, \dots, \mathbf{v_n}\}$. To create a basis for span(S), delete the vectors from S from which correspond to the non-leading columns in the row-echelon form of A.

[X] Coordinate Vectors

Ordered Basis and Coordinate Vector

Let $B = \{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ be an ordered set that is a basis for a vector space. If

$$\mathbf{v} = x_1\mathbf{v_1} + \cdots + x_n\mathbf{v_n},$$

then the **coordinate vector of v with respect to the ordered basis** B is given by

$$[\mathbf{v}]_B = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Note that taking coordinate vectors is a linear transformation.

Proving that a Transformation isn't Linear

Linear Maps Preserves the Zero and All Negatives

Consider the function $T : V \to W$. If either

1. $T(\mathbf{0}) \neq \mathbf{0}$ or
2. $T(-\mathbf{v}) \neq -T(\mathbf{v})$ for some $\mathbf{v} \in V$

then T is not a linear map.

If you cannot make use of the above theorem to prove that a transformation is not linear (for, say $T : \mathbb{R} \to \mathbb{R}$ where $T(x) = \sin(x)$), then provide a counterexample to one of the conditions of a linear map as specified in the definition (i.e. show that the map either does not preserve addition or scalar multiplication for at least specific set of values)

Linear Maps and Geometric Meaning

Stretching/Compression, Rotation and Reflection

- Stretch or compress: $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$
- Rotate: $\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$
- Reflect: $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$, $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$

Vector Operations

- taking **dot product** with a fixed vector: $T(\mathbf{x}) = \mathbf{b} \cdot \mathbf{x}$.
- **projection** on a fixed vector: $T(\mathbf{x}) = \text{proj}_{\mathbf{b}}\, \mathbf{x} = \dfrac{\mathbf{x} \cdot \mathbf{b}}{|\mathbf{b}^2|}\mathbf{b}$.

Determining Whether Solutions Exist

Determine Whether a Solution Exists

The equation $A\mathbf{x} = \mathbf{b}$ has:

1. no solution if $\text{rank}(A) \neq \text{rank}(A|\mathbf{b})$, and
2. at least one solution if $\text{rank}(A) = \text{rank}(A|b)$. Further,
   - if $\text{nullity}(A) = 0$, the solution is unique, whereas,
   - if $\text{nullity}(A) = n > 0$, then the general solution is of the form

$$\mathbf{x} = \mathbf{x}_p + \lambda_1\mathbf{k}_1 + \cdots + \lambda_n\mathbf{k}_n \quad \text{for} \quad \lambda_1, \ldots, \lambda_n \in \mathbb{R},$$

where $\mathbf{x}_p$ is a particular solution of $A\mathbf{x} = \mathbf{b}$ and $\{\mathbf{k}_1, \ldots, \mathbf{k}_2\}$ is a basis of $\ker(A)$.

We already knew this before we defined rank and nullity using our knowledge of the relationship between the existence of solutions and the columns of a matrix.

Vector Spaces

MATH1231 NOVEMBER 2012 Q1(iii)

The vector space $\mathbb{P}_2$ consists of all polynomials, with real coefficients, of degree at most 2. Suppose W is the subset of $\mathbb{P}_2$ defined by

$$W = \{p \in \mathbb{P}_2 : p'(0)p''(0) = 0\},$$

where $'$ denotes differentiation.

1. Show that W is closed under scalar multiplication.
2. Determine, with reasons, whether or not W is a subspace of $\mathbb{P}_2$

MATH1231 NOVEMBER 2012 Q1(iii)

$$W = \{p \in \mathbb{P}_2 : p'(0)p''(0) = 0\}.$$

1. Let $p \in W$ and $\lambda \in \mathbb{R}$. Then, $p'(0)p''(0) = 0$. Now, $(\lambda p)'(0)(\lambda p)''(0) = \lambda^2 \times p'(0)p''(0) = \lambda^2 \times 0 = 0$ and so $\lambda p \in W$. That is, W is closed under scalar multiplication.
2. Let $p, q \in W$. Then, $p'(0)p''(0) = 0$ and $q'(0)q''(0) = 0$. Now,

$$\begin{aligned}
&(p+q)'(0)(p+q)''(0) \\
&= (p'+q')(0)(p''+q'')(0) \\
&= (p'(0)+q'(0))(p''(0)+q''(0)) \\
&= p'(0)p''(0) + p'(0)q''(0) + q'(0)p''(0) + q'(0)q''(0) \\
&= p'(0)q''(0) + q'(0)p''(0).
\end{aligned}$$

Since this does not seem to equal 0 all the time, we should find a counterexample.

We note that at least one of $p'(0)$ and $p''(0)$ is zero, and at least one of $q'(0)$ and $q''(0)$ is zero to satisfy the condition for p and q to be in W.

Suppose $p(x) = x^2$ and $q(x) = x$. Note that $p, q \in \mathbb{P}_2$ and $p'(x) = 2x$ and $p''(x) = 2 \implies p'(0) = 0$ and $p''(0) = 2 \implies p'(0)p''(0) = 0$ and $q'(x) = 1$ and $q''(x) = 0 \implies q'(0) = 1$ and $q''(0) = 0 \implies q'(0)q''(0) = 0$. Therefore, our chosen polynomials satisfies $p, q \in W$.

Now $(p+q)'(x)(p+q)''(x) = 0 \times 0 + 1 \times 2 = 2 \neq 0$. Hence $p + q \notin W$, and so W is not closed under addition. Hence, W is not a vector space.

MATH1241 NOVEMBER 2010 Q2(iv)

Let $\mathbb{P}_3(\mathbb{R})$ denote the vector space of polynomials of degree 3 or less with real coefficients and let S be given by

$$S = \{p \in \mathbb{P}_3(\mathbb{R}) : (x-3)p'(x) - 2p(x) = 0 \text{ for all } x \in \mathbb{R}\}.$$

Show that S is a subspace of $\mathbb{P}_3(\mathbb{R})$.

MATH1241 NOVEMBER 2010 Q2(iv)

$$S = \{p \in \mathbb{P}_3(\mathbb{R}) : (x-3)p'(x) - 2p(x) = 0 \text{ for all } x \in \mathbb{R}\}.$$

S is a subset of $\mathbb{P}_3(\mathbb{R})$.
Zero Vector. Let p be the zero polynomial, defined by $p(x) \equiv 0$. Clearly, $p'(x) \equiv 0$, and so $(x-3)p'(x) - 2p(x) = 0$. Hence $p \in S$, and so S contains the zero polynomial.
Closure under addition. Let $p, q \in S$. Then $(x-3)p'(x) - 2p(x) = 0$ and $(x-3)q'(x) - 2q(x) = 0$. Now,

$$\begin{aligned}
&(x-3)(p+q)'(x) - 2(p+q)(x)\\
&= (x-3)(p'+q')(x) - 2(p(x)+q(x))\\
&= (x-3)(p'(x)+q'(x)) - 2p(x) - 2q(x)\\
&= ((x-3)p'(x) - 2p(x)) + ((x-3)q'(x) - 2q(x))\\
&= 0 + 0 = 0
\end{aligned}$$

and so $p + q \in S$. Hence, S is closed under addition.
Closure under scalar multiplication. Let $p \in S$ and $\lambda \in \mathbb{R}$. Then $(x-3)p'(x) - 2p(x) = 0$. Now,

$$\begin{aligned}(x-3)(\lambda p)'(x) - 2(\lambda p)(x) &= (x-3)(\lambda \times p'(x)) - 2\lambda \times p(x) \\ &= \lambda[(x-3)p'(x) - 2p(x)] \\ &= \lambda \times 0 = 0\end{aligned}$$

and so $\lambda p \in S$. Hence S is closed under scalar multiplication.
Hence, S is a subspace of $\mathbb{P}_3(\mathbb{R})$ by the Subspace Theorem.

Linear Transformations

MATH1231 NOVEMBER 2014 3(vi)(b)

Let $B = \{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}$ be a set of three non-zero vectors in $\mathbb{R}^3$.

1. Prove that if B is an orthogonal set the B is a linearly independent set.
2. Hence, explain why any orthogonal set of 3 non-zero vectors in $\mathbb{R}^3$ forms a basis for $\mathbb{R}^3$.

MATH1231 NOVEMBER 2014 3(vi)(b)

1. If B is an orthogonal set, then $\mathbf{v_1} \cdot \mathbf{v_2} = 0, \mathbf{v_1} \cdot \mathbf{v_3} = 0$ and $\mathbf{v_2} \cdot \mathbf{v_3} = 0$. Consider the equation

$$\lambda_1 \mathbf{v_1} + \lambda_2 \mathbf{v_2} + \lambda_3 \mathbf{v_3} = \mathbf{0}. \tag{1}$$

   Now, taking the dot product of $\mathbf{v_1}$ and both sides (1), we obtain:

$$\begin{aligned} \mathbf{v_1} \cdot (\lambda_1 \mathbf{v_1} + \lambda_2 \mathbf{v_2} + \lambda_3 \mathbf{v_3}) &= \mathbf{v_1} \cdot \mathbf{0} \\ \lambda_1(\mathbf{v_1} \cdot \mathbf{v_1}) + \lambda_2(\mathbf{v_1} \cdot \mathbf{v_2}) + \lambda_3(\mathbf{v_1} \cdot \mathbf{v_3}) &= 0 \\ \lambda_1 |\mathbf{v_1}|^2 &= 0. \end{aligned}$$

   Since $\mathbf{v_1}$ is not the zero vector, $|\mathbf{v_1}|^2 \neq 0$. Hence $\lambda_1 = 0$. Repeating the process with $\mathbf{v_2}$ and $\mathbf{v_3}$, we find that the only solution of λ_1, λ_2 and λ_3 to (1) is $\lambda_1 = \lambda_2 = \lambda_3 = 0$.
   Hence, B is a linearly independent set.

2. Since $\dim(\mathbb{R}^3) = 3$, and there are 3 linearly independent vectors in an orthogonal set of 3 non-zero vectors, such a set must span $\mathbb{R}^3$, forming a basis.

MATH1231 NOVEMBER 2015 Q3(ii)

Let $S = \{p_1, p_2, p_3\}$ be a set of polynomials over $\mathbb{R}$, where

$$p_1(x) = 1 + x + 2x^2, \tag{1}$$
$$p_2(x) = 3 + 2x + x^2, \tag{2}$$
$$p_3(x) = 2 + 5x + x^2. \tag{3}$$

1. Show that S is linearly independent.
2. Explain why S forms a basis for $\mathbb{P}_2(\mathbb{R})$.

MATH1231 NOVEMBER 2015 Q3(ii)

$S = \{p_1, p_2, p_3\}$ where $p_1(x) = 1 + x + 2x^2$, $p_2(x) = 3 + 2x + x^2$ and $p_3(x) = 2 + 5x + x^2$.

1. Consider the matrix A whose columns consist of the coefficients of each polynomial.

$$A = \begin{pmatrix} 1 & 3 & 2 \\ 1 & 2 & 5 \\ 2 & 1 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 3 & 2 \\ 0 & -1 & 3 \\ 0 & 0 & -18 \end{pmatrix}$$

   Hence, the homogeneous equation $A\mathbf{x} = \mathbf{0}$ has a unique solution, and so S is a linearly independent set.

2. *Note that the dimension of* $\mathbb{P}_n$ *is* $n + 1$. Since there are no zero rows in the row-echelon form of A, there will be no inconsistent rows for any coefficients of polynomials in $\mathbb{P}_2(\mathbb{R})$ and so S is also spanning set for $\mathbb{P}_2(\mathbb{R})$. Hence S is a basis for $\mathbb{P}_2(\mathbb{R})$.