UNSW Mathematics Society Presents

MATH1231/1241 Workshop: Algebra

Presented by Isaiah Iliffe and Jay Liang

Overview I

Overview II

Probability
Conditional Probability
Random Variables
Cumulative Distribution Function
Probability Distribution
Expected Value and Variance
Binomial Distribution
Probability Density Function
Normal Distribution
[X] Exponential Distribution

1. Vector Spaces

MATH1241 S2 2018 Q1(i)

Prove that
$$S = \{\mathbf{x} \in \mathbb{R}^3 : x_1^2 = x_2 x_3\}$$
is **not** a subspace of $\mathbb{R}^3$.

Let $\mathbf{u} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

Now $0^2 = 1 \times 0$ so $\mathbf{u} \in S$, and $0^2 = 0 \times 1$ so $\mathbf{v} \in S$.

But $\mathbf{u} + \mathbf{v} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$, but $0^2 = 1 \times 1$ and so $\mathbf{u} + \mathbf{v} \notin S$.

Hence S is not closed under addition, and so S is not a subspace of $\mathbb{R}^3$.

Isaiah's question

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.

Isaiah makes the following argument:

> If S is closed under scalar multiplication, then for any $\mathbf{v} \in S$, multiplying by the zero scalar gives that $0\mathbf{v} = \mathbf{0} \in S$.
> So condition 1 of the subspace theorem is redundant.

Is it correct?

MATH1231 S2 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 $\mathbb{R}^3$.

1. Yes, $S = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{0}\}$

MATH1231 S2 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 $\mathbb{R}^3$.

2. No, $\mathbf{v_1} = \mathbf{v_2} = \mathbf{v_3} = \mathbf{v_4} = \mathbf{0}$

MATH1231 S2 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 $\mathbb{R}^3$.

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 $\mathbb{R}^3$.

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.

For S_1, we wish to see whether the only solution to

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

is $\lambda = 0$.
If $\lambda\mathbf{u} - \lambda\mathbf{v} = \mathbf{0}$, then since $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly independent, each coefficient must be 0 and so $\lambda = 0$ as required.

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.

For S_2, we wish to see whether the only solution to

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

is $\lambda_1 = \lambda_2 = 0$.
If $\lambda_1 \mathbf{u} + (\lambda_2 - \lambda_1)\mathbf{v} - \lambda_2 \mathbf{w} = \mathbf{0}$, then since $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly independent, each coefficient must be 0 and so $\lambda_1 = \lambda_2 = 0$ as required.

MATH1241 T2 2019 Q1(f)(i)

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

$$S_1 = \{\mathbf{u} - \mathbf{v}\},$$
$$S_2 = \{\mathbf{u} - \mathbf{v}, \mathbf{v} - \mathbf{w}\},$$
$$S_3 = \{\mathbf{u} - \mathbf{v}, \mathbf{v} - \mathbf{w}, \mathbf{w} - \mathbf{u}\}.$$

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

By inspection,

$$1(\mathbf{u} - \mathbf{v}) + 1(\mathbf{v} - \mathbf{w}) + 1(\mathbf{w} - \mathbf{u}) = \mathbf{0},$$

so S_3 is not linearly independent.

Suppose that V is a real vector space of dimension 95 containing two subspaces U of dimension 47 and W of dimension 56 . Let S be the intersection of U and W.
Find the greatest possible dimension and least possible dimension of S.
[**X**] Prove.

Greatest possible dimension: $\dim(S) = \dim(U \cap W) \leq \dim(U) = 47$.
Example: $V = \mathbb{R}^{95}$, $U = \text{span}\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{47}\}$,
$W = \text{span}\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{56}\}$, so $S = \text{span}\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{47}\}$
Proof of $\leq$: since $U \cap W \subseteq U$, we can extend the basis of $U \cap W$ to a basis of U. (Notes p42 Theorem 5)
So max dimension of S is 47.

MATH1241 Q5 2020

$\dim(V) = 95, \dim(U) = 47, \dim(W) = 56$

Least possible dimension of $\dim(U \cap W)$:
Intuition: to minimise intersection, "give" dimensions to U and V, minimising shared dimensions. Thus they have to share $47 + 56 - 95 = 8$ dimensions, so answer is 8.
Example: $V = \mathbb{R}^{95}$, $U = \text{span}\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{47}\}$,
$W = \text{span}\{\mathbf{e}_{48}, \ldots, \mathbf{e}_{95}, \mathbf{e}_1, \ldots, \mathbf{e}_8\}$, so $S = \text{span}\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_8\}$

MATH1241 Q5 2020

Suppose that V is a real vector space of dimension 95 containing two subspaces U of dimension 47 and W of dimension 56 . Let S be the intersection of U and W.
Fill in the blanks in the following proof that S is a subspace of V:
Proof:

1. S is a subset of the known vector space V.
2. For all $\mathbf{x}$ in _ and $\mathbf{y}$ in _, $\mathbf{x}$ and $\mathbf{y}$ __ which is a vector space so $\mathbf{x}+\mathbf{y}$ is in _. Similarly, $\mathbf{x}$ and $\mathbf{y}$ __ which is a vector space so $\mathbf{x}+\mathbf{y}$ is in _. Therefore $\mathbf{x}+\mathbf{y}$ is in __ and hence in _. This shows that S ____.
3. For all λ in $\mathbb{R}$ and $\mathbf{x}$ in _, $\mathbf{x}$ _____ therefore $\lambda\mathbf{x}$ is in ___ and hence in _. This shows that S ____.
4. Using the ____ and points 1, 2 and 3, we can conclude that S is a subspace of V.

MATH1241 Q5 2020

Prove that $\dim(U \cap W) \geq \dim(U) + \dim(W) - \dim(V)$

Proof: let B be a basis for $U \cap W$. Choose A disjoint from B such that $A \cup B$ is a basis for U, by extending the basis. Similarly choose C disjoint from B such that $C \cup B$ is a basis for W.
Now $|A| = \dim(U) - \dim(U \cap W)$, $|B| = \dim(U \cap W)$ and $|C| = \dim(W) - \dim(U \cap W)$. So $|A \cup B \cup C| = \dim(U) + \dim(W) - \dim(U \cap W)$.
So if we prove $A \cup B \cup C$ is linearly independent, then it has smaller size than dimension of V, so $\dim(U) + \dim(W) - \dim(U \cap W) \leq \dim(V)$, as required.

Construction: $B = \{\mathbf{b}_j\}_j$ is a basis for $U \cap W$. $A = \{\mathbf{a}_i\}_i$ is disjoint from B such that $A \cup B$ is a basis for U. $C = \{\mathbf{c}_k\}_k$ is disjoint from B such that $C \cup B$ is a basis for W.
RTP: $A \cup B \cup C$ is linearly independent.
Suppose $\sum_i \alpha_i \mathbf{a}_i + \sum_j \beta_j \mathbf{b}_j + \sum_k \gamma_k \mathbf{c}_k = \mathbf{0}$.
Then $\mathbf{u} = \sum_i \alpha_i \mathbf{a}_i \in W$ by closure, but $\mathbf{u} \in U$ by construction, so $\mathbf{u} \in U \cap W = \text{span}(B)$.
But $\mathbf{u} \in \text{span}(A)$ by construction, and since $A \cup B$ is linearly independent, $\text{span}(A) \cap \text{span}(B) = \{\mathbf{0}\}$. This means $\mathbf{u} = \mathbf{0}$, so all α_i are 0.
So $\sum_j \beta_j \mathbf{b}_j + \sum_k \gamma_k \mathbf{c}_k = \mathbf{0}$, which means all coefficients are 0 since $B \cup C$ is linearly independent.
Since all coefficients are forced to be 0, $A \cup B \cup C$ is linearly independent.
Remark: we have essentially proven Grassmann's identity (Google it).

2. 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).

MATH1251 NOVEMBER 2011 Q1 (b)

The mapping $T : V \to V$ is defined for $f \in V$ by

$$(Tf)(\theta) = f'(\theta) - f(\frac{\pi}{2} - \theta), \theta \in \mathbb{R}.$$

Tip: For the above problem, recall the fact that addition and scalar multiplication of functions is defined as:

$$(f + g)(x) = f(x) + g(x)$$

And,

$$(\lambda f)(x) = \lambda f(x).$$

Linearity of Differentiation

Recall that the linearity of differentiation implies that the derivative of a linear combination of functions is equal to linear combination of its derivatives.
i.e. $(af + g)'(x) = af'(x) + g'(x)$

MATH1251 NOVEMBER 2011 Q1 (b)

Solution: Let $p, q \in V$ and $\lambda_1, \lambda_2 \in \mathbb{R}$.
Then, $(Tp)(\theta) = p'(\theta) - p(\frac{\pi}{2} - \theta)$ and $(Tq)(\theta) = q'(\theta) - q(\frac{\pi}{2} - \theta)$.
Consider,

$$\begin{aligned}
T(\lambda_1 p + \lambda_2 q)(\theta) &= (\lambda_1 p + \lambda_2 q)'(\theta) - (\lambda_1 p + \lambda_2 q)(\frac{\pi}{2} - \theta) \\
&= \lambda_1 p'(\theta) + \lambda_2 q'(\theta) - \lambda_1 p(\frac{\pi}{2} - \theta) - \lambda_2 q(\frac{\pi}{2} - \theta) \\
&= (\lambda_1 p'(\theta) - \lambda_1 p(\frac{\pi}{2} - \theta)) + (\lambda_2 q'(\theta) - \lambda_2 q(\frac{\pi}{2} - \theta)) \\
&= \lambda_1 (p'(\theta) - p(\frac{\pi}{2} - \theta)) + \lambda_2 (q'(\theta) - q(\frac{\pi}{2} - \theta)) \\
&= \lambda_1 (Tp)(\theta) + \lambda_2 (Tq)(\theta).
\end{aligned}$$

Example

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

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

is a linear map.

Matrix Representation Theorem

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map. Suppose vectors $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ 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 $T(\mathbf{e}_j)$.

Example

MATH1251 NOVEMBER 2010 Q2 (iii)

Let $\mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$ be a fixed vector in $\mathbb{R}^3$. The function $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by

$$T(\mathbf{x}) = \mathbf{c} \times \mathbf{x}$$

for $\mathbf{x} \in \mathbb{R}^3$.
Where $\mathbf{c} \times \mathbf{x}$ is the cross product, is a linear map. Find a matrix C which represents T with respect to the standard basis for $\mathbb{R}^3$.

Solution: Consider the standard basis of $\mathbb{R}^3$:

$$\mathbf{e_1} = \begin{pmatrix}1\\0\\0\end{pmatrix}, \mathbf{e_2} = \begin{pmatrix}0\\1\\0\end{pmatrix}, \mathbf{e_3} = \begin{pmatrix}0\\0\\1\end{pmatrix}$$

And so,

$$\begin{aligned}
T(\mathbf{e_1}) &= \begin{pmatrix}c_1\\c_2\\c_3\end{pmatrix} \times \begin{pmatrix}1\\0\\0\end{pmatrix} \\
&= \det\begin{pmatrix}e_1 & e_2 & e_3\\ c_1 & c_2 & c_3\\ 1 & 0 & 0\end{pmatrix} \\
&= \mathbf{e_1}(0-0) - \mathbf{e_2}(0-c_3) + \mathbf{e_3}(0-c_2) \\
&= \begin{pmatrix}0 & c_3 & -c_2\end{pmatrix}^T
\end{aligned}$$

Solution continued:
Proceeding with the same method, we arrive at:

$$T(\mathbf{e_1}) = \begin{pmatrix} 0 \\ c_3 \\ -c_2 \end{pmatrix}, T(\mathbf{e_2}) = \begin{pmatrix} -c_3 \\ 0 \\ c_1 \end{pmatrix}, T(\mathbf{e_3}) = \begin{pmatrix} c_2 \\ -c_1 \\ 0 \end{pmatrix}.$$

And so, by applying the Matrix Representation Theorem,

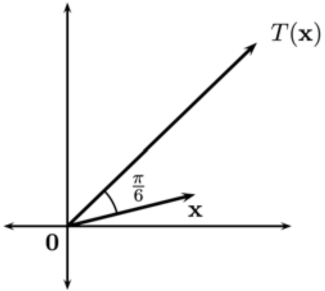
$$C = \begin{pmatrix} 0 & -c_3 & c_2 \\ c_3 & 0 & -c_1 \\ -c_2 & c_1 & 0 \end{pmatrix}$$

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.

Solution:
The kernel of T would consist of any vector where $\mathbf{x} \cdot \mathbf{b} = 0$. That is, all vectors in the kernel would be perpendicular to $\mathbf{b}$.

Furthermore, the kernel must contain $\mathbf{0}$, thus it is the plane through the origin which is normal to $\mathbf{b}$.

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,

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

MATH1251 NOVEMBER 2013 Q2 (i)

Let

$$B = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & -4 & 6 & -2 \\ -1 & 2 & -3 & 1 \end{pmatrix}.$$

A row-echelon form U of the augmented matrix $(B|I)$ is given by

$$U = \left(\begin{array}{cccc|ccc} 1 & 2 & 3 & 4 & 1 & 0 & 0 \\ 0 & 4 & 0 & 5 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 2 \end{array}\right)$$

a) Find a basis for the kernel of B.

b) Find a basis for $\text{im}(B)$.

c) State the rank of B and the nullity of B.

d) Without solving, does the equation $B\mathbf{x} = \begin{pmatrix} 1 & 0 & 1 \end{pmatrix}^T$ have a solution? Give a reason.

MATH1251 NOVEMBER 2013 Q2 (i)

Solution to a) and c): In U, it is observed that columns 3 and 4 are not leading columns and thus set $x_3 = \alpha$ and $x_4 = \beta$ as free parameters.

Suppose that the right side of the augmented matrix had just entries of 0, then:

$$\begin{aligned} 4x_2 + 5\beta &= 0 \\ x_1 + 2x_2 + 3\alpha + 4\beta &= 0 \end{aligned}$$

Solving this pair of simultaneous equations gives, $x_1 = -3\alpha - \frac{3}{2}\beta$, $x_2 = -\frac{5}{4}\beta$, $x_3 = \alpha$ and $x_4 = \beta$.

Thus, any vector in the kernel can be expressed as:

$$\alpha \begin{pmatrix} -3 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} -3/2 \\ -5/4 \\ 0 \\ 1 \end{pmatrix}$$

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,

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

MATH1251 NOVEMBER 2013 Q2 (i)

Let

$$B = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & -4 & 6 & -2 \\ -1 & 2 & -3 & 1 \end{pmatrix}.$$

A row-echelon form U of the augmented matrix $(B|I)$ is given by

$$U = \left(\begin{array}{cccc|ccc} 1 & 2 & 3 & 4 & 1 & 0 & 0 \\ 0 & 4 & 0 & 5 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 2 \end{array}\right)$$

a) Find a basis for the kernel of B.

b) Find a basis for $\mathrm{im}(B)$.

c) State the rank of B and the nullity of B.

d) Without solving, does the equation $B\mathbf{x} = \begin{pmatrix} 1 & 0 & 1 \end{pmatrix}^T$ have a solution? Give a reason.

Solution to b) and c): In U, it is observed that columns 1 and 2 are leading columns and so the first and second vector of B form the basis for the image of B (by Linear Independence).

So, a basis for $\text{im}(B)$ is the set containing vectors: $\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ -4 \\ 2 \end{pmatrix}.$

Solution to d): No.
Try this yourself. It is revision of determining the number of solutions when solving augmented matrices.

(Hint: Notice that $(1, 0, 1)^T = \mathbf{e_1} + \mathbf{e_3}$. Then using the reduced augmented matrix $(B|I)$, all that is left to do, is to check the final row of this augmented matrix.)

To summarise...

Finding the Rank and Nullity

- nullity(A) = number of non-leading columns in a row-echelon form U for A.
- 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.

Example

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. $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. 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 $\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$.

MATH1231 T3 2020 Q14

Let V be the set of ordered pairs (a, b) where $a, b \in \mathbb{R}$.
Define special addition and multiplication by a scalar in $\mathbb{R}$ by

$$(a, b) \oplus (c, d) = (-3 + a + c,\ 5 + b + d) \quad \text{for all } (a, b), (c, d) \in V,$$

$$\lambda \star (a, b) = (a\lambda - 3\lambda + 3, b\lambda + 5\lambda - 5)) \quad \text{for all } \lambda \in \mathbb{R} \text{ and } (a, b) \in V.$$

(b) V is a vector space. What are the additive and multiplicative identities?
(c) Find a real number k such that the mapping $T : V \to \mathbb{R}$ given by $T((a, b)) = a + b + k$ is linear.

Additive identity (x, y) must satisfy $(a, b) \oplus (x, y) = (a, b)$ for all a, b.
That is, $(-3 + a + x, 5 + b + y) = (a, b)$, and so
$-3 + a + x = a \implies x = 3$ and $5 + b + y = b \implies y = -5$.
So additive identity is $(3, -5)$.

MATH1231 T3 2020 Q14

Let V be the set of ordered pairs (a, b) where $a, b \in \mathbb{R}$.
Define special addition and multiplication by a scalar in $\mathbb{R}$ by

$$(a, b) \oplus (c, d) = (-3 + a + c,\ 5 + b + d) \quad \text{for all } (a, b), (c, d) \in V,$$

$$\lambda \star (a, b) = \ (a\lambda - 3\lambda + 3, b\lambda + 5\lambda - 5)) \quad \text{for all } \lambda \in \mathbb{R} \text{ and } (a, b) \in V.$$

(b) V is a vector space. What are the additive and multiplicative identities?
(c) Find a real number k such that the mapping $T : V \to \mathbb{R}$ given by $T((a, b)) = a + b + k$ is linear.

Multiplicative identity μ must satisfy $\mu \star (a, b) = (a, b)$ for all a, b.
That is, $(a\mu - 3\mu + 3, b\mu + 5\mu - 5) = (a, b)$, which is satisfied iff $\mu = 1$.
So multiplicative identity is 1.

MATH1231 T3 2020 Q14

Let V be the set of ordered pairs (a, b) where $a, b \in \mathbb{R}$.
Define special addition and multiplication by a scalar in $\mathbb{R}$ by

$$(a, b) \oplus (c, d) = (-3 + a + c,\ 5 + b + d) \quad \text{for all } (a, b), (c, d) \in V,$$

$$\lambda \star (a, b) = \ (a\lambda - 3\lambda + 3, b\lambda + 5\lambda - 5)) \quad \text{for all } \lambda \in \mathbb{R} \text{ and } (a, b) \in V.$$

(b) V is a vector space. What are the additive and multiplicative identities?
(c) Find a real number k such that the mapping $T : V \to \mathbb{R}$ given by $T((a, b)) = a + b + k$ is linear. Explain.

Since $(3, -5)$ is the additive identity and by linearity,
$T((3, -5)) = T((3, -5) + (3, -5)) = T((3, -5)) + T((3, -5))$, and so $T((3, -5)) = 0$.
Thus $3 - 5 + k = 0$, so $k = 2$.

$$(a, b) \oplus (c, d) = (-3 + a + c,\ 5 + b + d) \quad \text{for all } (a, b), (c, d) \in V,$$

$$\lambda \star (a, b) = \ (a\lambda - 3\lambda + 3, b\lambda + 5\lambda - 5)) \quad \text{for all } \lambda \in \mathbb{R} \text{ and } (a, b) \in V.$$

We now check that $T((a, b)) = a + b + 2$ is linear.

$$\begin{aligned}
&T(\lambda \star (a, b) \oplus \mu \star (c, d)) \\
&= T((a\lambda - 3\lambda + 3, b\lambda + 5\lambda - 5) \oplus (c\mu - 3\mu + 3, d\mu + 5\mu - 5)) \\
&= T((a\lambda - 3\lambda + 3 + c\mu - 3\mu + 3 - 3, 5 + b\lambda + 5\lambda - 5 + d\mu + 5\mu - 5) \\
&= a\lambda - 3\lambda + 3 + c\mu - 3\mu + 3 - 3 + 5 + b\lambda + 5\lambda - 5 + d\mu + 5\mu - 5 + 2 \\
&= \lambda(a + b + 2) + \mu(c + d + 2) \\
&= \lambda T((a, b)) + \mu T((c, d))
\end{aligned}$$

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 S2 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.

MATH1241 S2 2018 Q1(iv)

Consider the matrix $M = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$.

(a) Find a basis for $\ker(M)$.

(b) Find a basis for $\text{im}(M^T)$.

(c) Give a geometric description of $\ker(M)$ and $\text{im}(M)$ as subspaces of $\mathbb{R}^2$.

MATH1241 T2 2020 Q4

Suppose that $T : \mathbb{R}^n \to \mathbb{R}^3$ is a linear transformation, such that for some matrix A, for each $\mathbf{v} \in \mathbb{R}^n$, $T(\mathbf{v}) = A\mathbf{v}$.

It is given that the image of T is the span of the vector $\begin{pmatrix} 7 \\ 3 \\ 1 \end{pmatrix}$ and that the nullity of T is 5.

(a) Find the value of n.

(b) Find a possible matrix A.

(a) By rank-nullity theorem, $n = \text{rank}(T) + \text{nullity}(T) = 1 + 5 = 6$.

(b)

$$\begin{pmatrix} 7 & 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

MATH1241 2020 T2 Q3

(a) Consider the linear mapping $T : \mathbb{P}_{92} \to \mathbb{P}_{92}$ where $T(p) = p'$. State $\text{rank}(T)$ and $\text{nullity}(T)$.
(b) Consider a linear mapping $R : W \to \mathbb{P}_{53}$, where W is a subspace of $\mathbb{P}_{53}$. Suppose that $\dim(W) = 53$ and $\text{nullity}(R) = 0$. Extend $\{x^3, x^4, \ldots, x^{51}, x^{52}\}$ to a basis for W.

(a) Image is any polynomial in $\mathbb{P}_{91}$, so rank is 92. Kernel is any constant polynomial, so nullity is 1.
(b) Nullity 0, so W contains no constant functions. So add x, x^2, x^{53}.

Let $\mathbb{P}_3 = \mathbb{P}_3(\mathbb{R})$ be the space of real polynomials of degree at most 3. Lucas's favourite linear mapping $T : \mathbb{P}_3 \to \mathbb{R}_3$ is defined by

$$T(p) = \begin{pmatrix} p(3) \\ p(-1) \\ p(-1) \end{pmatrix}.$$

(b) By calculating explicitly $\ker T$, find the nullity and rank of T.
(c) Charlie's favourite mapping $A : \mathbb{P}_3 \to \mathbb{R}_3$ is given by

$$A(p) = \begin{pmatrix} p(-652) \\ p(347) \\ p(-15) \end{pmatrix}.$$

Lucas thinks it will be too hard to determine the nullity and rank of A, but Charlie thinks they can easily be found. Who do you agree with?

MATH1231 2020 Q11

$T : \mathbb{P}_3 \to \mathbb{R}_3$ is defined by

$$T(p) = \begin{pmatrix} p(3) \\ p(-1) \\ p(-1) \end{pmatrix}.$$

Kernel: let $p(x) = ax^3 + bx^2 + cx + d$. If $p(3) = p(-1) = 0$, then $27a + 9b + c + d = 0$ and $-a + b - c + d = 0$. This is satisfied exactly when $26a + 10b + 2d = 0 \implies d = -13a + 5b$, and $c = -a + b + d = -14a + 6b$. Two parameters, so nullity is 2.
By rank-nullity theorem, rank is 2.

MATH1231 2020 Q11

$A : \mathbb{P}_3 \to \mathbb{R}_3$ is given by

$$A(p) = \begin{pmatrix} p(-652) \\ p(347) \\ p(-15) \end{pmatrix}.$$

Kernel: $p(-652) = p(347) = p(-15) = 0$, so
$p(x) = a(x + 652)(x - 347)(x + 15)$, which has one parameter so nullity is 1.
Rank 3.

Use the following Maple output to answer the questions below.

```
> with(LinearAlgebra):
> A := <<-8, 2, -17, -1>|<1, -10, 7, 11>|<5, -9, 24, 14>
       |<-1, -31, 4, 30>>;
```

$$A := \begin{bmatrix} -8 & 1 & 5 & -1 \\ 2 & -10 & -9 & -31 \\ -17 & 7 & 24 & 4 \\ -1 & 11 & 14 & 30 \end{bmatrix}$$

```
> LinearSolve(A, ZeroVector(4));
```

$$\begin{bmatrix} 0 \\ -4\ t_4 \\ t_4 \\ t_4 \end{bmatrix}$$

a) Is the set

$$S = \left\{ \begin{pmatrix} -8 \\ 2 \\ -17 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -10 \\ 7 \\ 11 \end{pmatrix}, \begin{pmatrix} 5 \\ -9 \\ 24 \\ 14 \end{pmatrix}, \begin{pmatrix} -1 \\ -31 \\ 4 \\ 30 \end{pmatrix} \right\}$$

linearly independent? Give reasons.

b) Find a basis for ker(A).

3. Eigenvalues and Eigenvectors

Definition

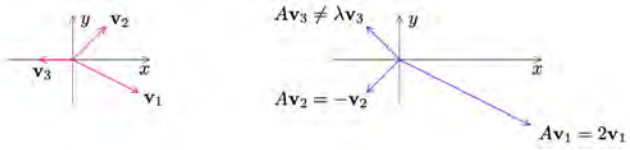
Let A be an n x n matrix, the scalar $\lambda \in \mathbb{F}$ and $\mathbf{v}$ a non-zero vector in $\mathbb{F}^{\mathrm{n}}$. If

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

then λ is the **eigenvalue** of A and $\mathbf{v}$ is the corresponding **eigenvector** of A.

By multiplying the matrix A to the vector $\mathbf{v}$, 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 $A\mathbf{v}$ and $\mathbf{v}$



Computing eigenvalues and eigenvectors

Method

- λ is an eigenvalue of A iff $\det(A - \lambda I) = 0$
- $\mathbf{v}$ 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 $E_\lambda = \{\boldsymbol{v} \in \mathbb{F}^{\mathrm{n}} \mid A\boldsymbol{v} = \lambda \boldsymbol{v}\}$, and is a subspace of $\mathbb{F}^{\mathrm{n}}$. (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 $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$, 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

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

MATH1231 2018 Semester 2 Final Q2(v)

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

a) Find the eigenvalues and eigenvectors of A.

b) Find a diagonal matrix D and matrix M such that $D = M^{-1}AM$.

MATH1231 2018 Semester 2 Final Q2(v)

Solution to a): For the matrix A, we have

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

Solving the characteristic equation, we obtain $\lambda = -3$ and $\lambda = 2$.

Now, to find the corresponding eigenvectors, we substitute the values of λ back into $(A - \lambda I)\mathbf{v} = \mathbf{0}$:

For $\lambda = -3$, we have $(A + 3I)\mathbf{v} = \mathbf{0}$

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

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

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

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

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

Solution to b): 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.

Example

MATH1241 2017 Semester 2 Final Q3(ii)

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

a) Given that the eigenvalues of A are $1, 2, 3$, explain why A is diagonalisable.

b) Find an eigenvector for A for the eigenvalue $\lambda = 3$.

c) 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.

d) Hence, prove that $f(A) = \mathbf{0}$.

e) Compute A^{-1} as a linear combination of A^2, A, I.

MATH1241 2017 Semester 2 Final Q3(ii)

Solution to a): 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*).

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

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

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

Solution to c): 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

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

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

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

MATH1241 2017 Semester 2 Final Q3(ii)

Solution to d): We know from part a) 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 $f(D) = \mathbf{0}$.

MATH1241 2017 Semester 2 Final Q3(ii)

Solution to e): From part d), $f(A) = A^3 - 6A^2 + 11A - 6I = \mathbf{0}$.
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 $\frac{d\boldsymbol{y}}{dt} = A\boldsymbol{y}$ 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 $\alpha_1, \alpha_2, \ldots, \alpha_n$, eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and corresponding eigenvectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ of A.

Example

MATH1251 NOVEMBER 2012 Q2 (iii)

a) Find the eigenvalues and the corresponding eigenvectors for

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

b) Hence find the solution of the differential equations

$$\frac{dy_1}{dt} = 2y_1 - 2y_2$$
$$\frac{dy_2}{dt} = 4y_1 - 4y_2$$

satisfying $y_1(0) = 2, y_2(0) = 3$.

c) Let $\mathbf{y}(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix}$. Find $\lim_{t\to\infty} \mathbf{y}(t)$.

Solution to a):

$$\begin{aligned}\det(A-\lambda I) &= \det\begin{pmatrix}2-\lambda & -2\\ 4 & -4-\lambda\end{pmatrix}\\ &= (2-\lambda)(-4-\lambda)+8\\ &= -8+2\lambda+\lambda^2+8\\ &= \lambda(2+\lambda)\end{aligned}$$

So, the eigenvalues are 0 and -2.

When $\lambda = 0$, then $ker\begin{pmatrix}2 & -2\\ 4 & -4\end{pmatrix} = ker\begin{pmatrix}2 & -2\\ 0 & 0\end{pmatrix}$

So the eigenvector for 0 is $\begin{pmatrix}1\\1\end{pmatrix}$.

Attempt $\lambda = -2$ yourselves. The answer is the eigenvector, $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

Solution to b): The general solution is: $\mathbf{y}(t) = \alpha_1 e^{\lambda_1 t}\mathbf{v_1} + \alpha_2 e^{\lambda_2 t}\mathbf{v_2}$, where λ_1, λ_2 are the eigenvalues and $\mathbf{v_1}, \mathbf{v_2}$ are the eigenvectors. So, substituting in the eigenvalues and eigenvectors previously calculated, we obtain:

$$\mathbf{y}(t) = \alpha_1 e^0 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \alpha_2 e^{-2t} \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Substituting in $y_1(0) = 2$, we get $2 = \alpha_1 + \alpha_2$

And subtituting in $y_2(0) = 3$, we get $3 = \alpha_1 + 2\alpha_2$

Solving simultaneously, yields $\alpha_1 = \alpha_2 = 1$

MATH1251 NOVEMBER 2012 Q2 (iii)

Solution to c): The general solution as found in part b) was

$$\mathbf{y}(t) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + e^{-2t} \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

So, as $t \to \infty$, then $e^{-2t} \to 0$, leaving us with just $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Higher order ODE example (Extra)

Higher Order ODEs

Solve the equation

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

Solutions (contd.)

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

$$\begin{aligned}\frac{dy_1}{dt} &= y_2 \\ \frac{dy_2}{dt} &= 24y_1 + 2y_2.\end{aligned}$$

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 $y = Ce^{6t} + De^{-4}$.

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 $A \subseteq S$
- $P(\emptyset) = 0$ 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 $\displaystyle\sum_{a \in S} P(\{a\}) = 1$

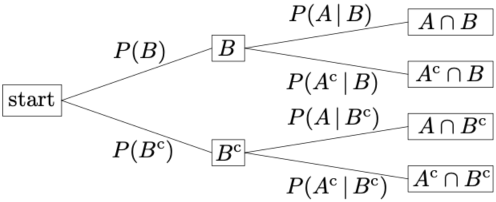
Conditional Probability

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)

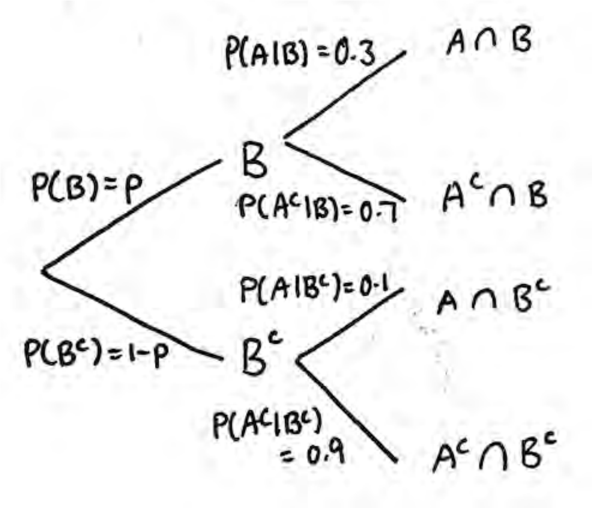
You are given the following information about the event A and B.

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

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 $P(A) = 0.2$,

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

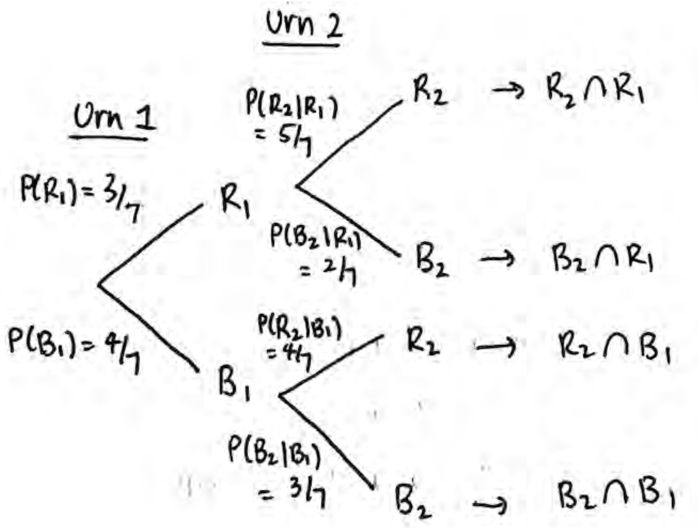
Examples

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?

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

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

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 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 $= X \sim B\left(13, \frac{1}{2}\right)$.

$P(X \geq 10) = \displaystyle\sum_{k=10}^{13} \binom{13}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{13-k} \approx 0.0461$. 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
- $f(x) = \lim\limits_{x \to a^-} F'(x)$ 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

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

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, $\int_{-\infty}^{\infty} f(x)dx = 1$. 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

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

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,

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

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.

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

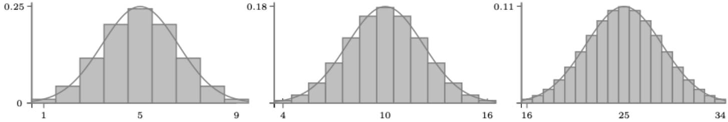
$$\begin{aligned}\frac{t - 48.7}{\sqrt{1.9}} &= 2.06 \\ t &= \sqrt{1.9} \times 2.06 + 48.7 \\ &= 51.540.\end{aligned}$$

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 $N(5, \frac{5}{2})$ $\quad$ $B(20, \frac{1}{2})$ and $N(10, 5)$ $\quad$ $B(50, \frac{1}{2})$ 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,

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