MATH1231/1241 MathSoc Algebra Revision Session 2017 S2 Solutions

October 28, 2017

These solutions were written and typed up by Aaron Hassan. Please be ethical with this resource. The problems here were discussed in MathSoc's 2017 Semester 2 MATH1231/1241 Final Exam Revision Session, presented by Aaron Hassan and Samir Mustavi. Several problems were taken or adapted from UNSW past exam papers and homework sheets, and all copyright of the original questions belongs to UNSW's School of Mathematics and Statistics. These solutions are for the benefit of MathSoc members, so do not repost it on other forums or groups without asking for permission. If you appreciate this resource, please consider supporting us by coming to our events and buying our T-shirts! Also, best of luck for your finals! 😎

We cannot guarantee that our working is error-free, or that it would obtain full marks in the exam – please notify us of any errors or typos at unswmathsoc@gmail.com, or on our Facebook page. There are sometimes multiple methods of solving the same question. Remember that in the real final exam, you will be expected to explain your steps and working out.

---

1 Vector spaces

Problem 1

[MATH1231 2014 S2 Q1 iv)]

Let $S = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : z^2 = x^2 + y^2 \right\}$.

a) Prove that S is closed under scalar multiplication.

b) Prove that S is **not** a subspace of $\mathbb{R}^3$.

Solution 1

a) Let $\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in S$ and $\alpha \in \mathbb{R}$. Then $z^2 = x^2 + y^2$ and $\alpha\mathbf{x} = \begin{pmatrix} \alpha x \\ \alpha y \\ \alpha z \end{pmatrix} \in \mathbb{R}^3$. Now,

$$\begin{aligned}
(\alpha z)^2 &= \alpha^2 z^2 \\
&= \alpha^2 \left(x^2 + y^2\right) \\
&= \alpha^2 x^2 + \alpha^2 y^2 \\
&= (\alpha x)^2 + (\alpha y)^2 ,
\end{aligned}$$

which implies that $\alpha\mathbf{x} \in S$. Hence S is closed under scalar multiplication.

b) Let $\mathbf{v} = \begin{pmatrix}1\\0\\1\end{pmatrix} \in \mathbb{R}^3$ and $\mathbf{w} = \begin{pmatrix}0\\1\\1\end{pmatrix} \in \mathbb{R}^3$. Since $1^2 = 1^2 + 0^2$ and $1^2 = 0^2 + 1^2$, we have that $\mathbf{v}, \mathbf{w} \in S$. However, $\mathbf{v} + \mathbf{w} = \begin{pmatrix}1\\1\\2\end{pmatrix}$ and $2^2 \neq 1^2 + 1^2$. Hence $\mathbf{v} + \mathbf{w} \notin S$. Thus S is not closed under addition, so cannot be a subspace of $\mathbb{R}^3$.

Remark. As the above solution shows, to show that something is **not** a subspace, it suffices to come up with a single numerical example that demonstrates that one of the properties of a vector space does not hold.

Problem 2

[MATH1231 2016 S2 Q3 ii)]

Consider the set $S = \left\{ \begin{pmatrix}1\\1\\2\end{pmatrix}, \begin{pmatrix}2\\1\\3\end{pmatrix}, \begin{pmatrix}1\\3\\4\end{pmatrix} \right\}$ of vectors in $\mathbb{R}^3$.

a) Prove that S is linearly dependent.

b) Write the last vector in S as a linear combination of the other two.

Solution 2

a) Consider the equation

$$c_1 \begin{pmatrix}1\\1\\2\end{pmatrix} + c_2 \begin{pmatrix}2\\1\\3\end{pmatrix} + c_3 \begin{pmatrix}1\\3\\4\end{pmatrix} = \mathbf{0}, \qquad (*)$$

for $c_1, c_2, c_3 \in \mathbb{R}$. Equating components, we obtain a system of linear equations that is represented by the following augmented matrix:

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 1 & 1 & 3 & 0 \\ 2 & 3 & 4 & 0 \end{array}\right).$$

(That is, the matrix whose columns are the vectors in our given set.) We now row-reduce this. Performing $R_2 \rightsquigarrow R_2 - R_1$ and $R_3 \rightsquigarrow R_3 - 2R_1$, we have

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 1 & 1 & 3 & 0 \\ 2 & 3 & 4 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} \boxed{1} & 2 & 1 & 0 \\ 0 & \boxed{-1} & 2 & 0 \\ 0 & -1 & 2 & 0 \end{array}\right).$$

Performing $R_3 \rightsquigarrow R_3 - R_2$, we get that the row-echelon form is

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

Since there is a non-leading column (column 3) in the row-echelon form, there exists a non-zero solution for c_1, c_2, c_3 in equation $(*)$. Therefore, S is a linearly dependent set.

b) We use back-substitution from our row-echelon form above. Set $\boxed{c_3 = -1}$, then row 2 shows us that

$$-c_2 + 2c_3 = 0 \Rightarrow c_2 = 2c_3 \Rightarrow \boxed{c_2 = -2}.$$

Now from row 1, we get

$$c_1 + 2c_2 + c_3 = 0 \Rightarrow c_1 = -2c_2 - c_3 = -2 \times (-2) - (-1) = 4 + 1 \Rightarrow \boxed{c_1 = 5}.$$

Putting these values of c_k into equation $(*)$ from part a), we have

$$5 \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + (-2) \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} + (-1) \begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix} = \mathbf{0}.$$

Rearrange this to get the last vector as a linear combination of the other two:

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

Problem 3

[MATH1231 2014 S2 Q3 vi)]
Let $B = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ be a set of three non-zero vectors in $\mathbb{R}^3$.

a) State the definition for the set B to be a linearly independent set.

b) Prove that if B is an orthogonal set then B is linearly independent.

c) Hence explain why any orthogonal set of 3 non-zero vectors in $\mathbb{R}^3$ forms a basis for $\mathbb{R}^3$.

Solution 3

a) The set B is **linearly independent** if the only values of the scalars $\lambda_1, \lambda_2, \lambda_3$ for which

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \lambda_3 \mathbf{v}_3 = \mathbf{0}$$

are $\lambda_1 = \lambda_2 = \lambda_3 = 0$.
Remark. This question shows that it is important that you know the formal definitions of terms as given in the course pack! The definition for linear independence is given in Definition 1 on page 25 of the Algebra course pack notes (page 43 of the electronic PDF). Make sure to know the definitions of **all** key terms and statements of all key theorems. See pages xiii and xiv of the Algebra course pack notes for a list of definitions and terms that you are expected to know (pages 13-14 of the electronic PDF).

b) Assume B is an orthogonal set. Then by definition, $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ for all $i \neq j$, for $i, j \in \{1, 2, 3\}$. Now, suppose that

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \lambda_3 \mathbf{v}_3 = \mathbf{0}. \qquad (*)$$

Then dotting both sides with $\mathbf{v}_1$ gives

$$\begin{aligned}(\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \lambda_3\mathbf{v}_3) \cdot \mathbf{v}_1 &= \mathbf{0} \cdot \mathbf{v}_1 \\ \Rightarrow \lambda_1(\mathbf{v}_1 \cdot \mathbf{v}_1) + \lambda_2(\mathbf{v}_2 \cdot \mathbf{v}_1) + \lambda_3(\mathbf{v}_3 \cdot \mathbf{v}_1) &= 0.\end{aligned}$$

(expanding LHS and recalling that dotting with $\mathbf{0}$ always gives 0)

Now, since B is an orthogonal set, we have $\mathbf{v}_2 \cdot \mathbf{v}_1 = \mathbf{v}_3 \cdot \mathbf{v}_1 = 0$. Hence the left-hand side above is just $\lambda_1(\mathbf{v}_1 \cdot \mathbf{v}_1)$, so we have

$$\lambda_1(\mathbf{v}_1 \cdot \mathbf{v}_1) = 0.$$

But $\mathbf{v}_1 \cdot \mathbf{v}_1 = \|\mathbf{v}_1\|^2 \neq 0$, because we are told that B is a set of **non-zero** vectors, so the lengths of the vectors in B are non-zero. Therefore, $\lambda_1(\mathbf{v}_1 \cdot \mathbf{v}_1) = 0$ implies that $\lambda_1 = 0$.

Similarly, by dotting $(*)$ with $\mathbf{v}_j$ for $j = 2, 3$, we can deduce that $\lambda_j = 0$ for $j = 2, 3$. Therefore, the **only** solution to $(*)$ is $\lambda_1 = \lambda_2 = \lambda_3 = 0$. Hence B is a linearly independent set, and the proof is complete.
Remark. As usual, the typical procedure to show a set is linearly independent is to write down the equation $(*)$, and then show that all the scalars must be 0.

c) We have shown above that any orthogonal set of three non-zero vectors in $\mathbb{R}^3$ is linearly independent. Therefore, any such set is a set of three linearly independent vectors in a space of dimension 3 (since $\mathbb{R}^3$ has dimension 3), and is thus a basis for $\mathbb{R}^3$. (Remember in general, any set of n independent vectors in a vector space of dimension n is a basis for that space.)

Problem 4

[MATH1241 2016 S2 Q3 iii)]
The field $\mathbb{F} = \mathrm{GF}(4)$ has elements $\{0, 1, \alpha, \beta\}$, with addition and multiplication defined by the following tables.

$+$	0	1	α	β
0	0	1	α	β
1	1	0	β	α
α	α	β	0	1
β	β	α	1	0

$\times$	0	1	α	β
0	0	0	0	0
1	0	1	α	β
α	0	α	β	1
β	0	β	1	α

For the vectors $\mathbf{b}_1 = \begin{pmatrix}1\\\alpha\\\beta\end{pmatrix}, \mathbf{b}_2 = \begin{pmatrix}\beta\\1\\1\end{pmatrix}, \mathbf{b}_3 = \begin{pmatrix}1\\0\\\alpha\end{pmatrix}, \mathbf{v} = \begin{pmatrix}\alpha\\0\\0\end{pmatrix}$,

a) show that $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ is a basis for $\mathbb{F}^3$

b) explain without calculation why $\{\mathbf{b}_1, \mathbf{b}_1 + \mathbf{b}_2, \mathbf{b}_2 + \mathbf{b}_3, \mathbf{b}_3\}$ is a spanning set but not a basis for $\mathbb{F}^3$;

c) find the coordinate vector of $\mathbf{v}$ with respect to the ordered basis $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ of $\mathbb{F}^3$.

Solution 4

a) As usual, we place the vectors $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ as the columns of a matrix and row-reduce. Anticipating part c) of the problem, we also augment the vector $\mathbf{v}$. We thus have the augmented matrix

$$\left(\begin{array}{ccc|c} 1 & \beta & 1 & \alpha \\ \alpha & 1 & 0 & 0 \\ \beta & 1 & \alpha & 0 \end{array}\right).$$

(We are working in a special field here (GF(4)), which has its own rules for addition and multiplication (defined by the tables above). We will sometimes need to use these tables when performing row-reduction calculations.) Performing $R_2 \rightsquigarrow R_2 - \alpha R_1$ and $R_3 \rightsquigarrow R_3 - \beta R_1$, we get

$$\left(\begin{array}{ccc|c} 1 & \beta & 1 & \alpha \\ \alpha & 1 & 0 & 0 \\ \beta & 1 & \alpha & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} \boxed{1} & \beta & 1 & \alpha \\ 0 & 1-\alpha\beta & -\alpha & -\alpha^2 \\ 0 & 1-\beta^2 & \alpha-\beta & -\alpha\beta \end{array}\right).$$

Note that from the given multiplication table for the field, we see that $\alpha\beta = \alpha \times \beta = 1$. Therefore, $1 - \alpha\beta = 0$. Also, we know that $-\alpha \neq 0$, because $\alpha \neq 0$ (recall that in a general field, $-\alpha = 0$ iff $\alpha = 0$). Furthermore, the table shows that $\beta^2 = \beta \times \beta = \alpha$, so $1 - \beta^2 = 1 - \alpha \neq 0$ since $\alpha \neq 1$. So the last matrix is equal to

$$\left(\begin{array}{ccc|c} \boxed{1} & \beta & 1 & \alpha \\ 0 & 0 & -\alpha & -\alpha^2 \\ 0 & 1-\alpha & \alpha-\beta & -\alpha\beta \end{array}\right).$$

Swapping the bottom two rows, we have

$$\left(\begin{array}{ccc|c} \boxed{1} & \beta & 1 & \alpha \\ 0 & 0 & -\alpha & -\alpha^2 \\ 0 & 1-\alpha & \alpha-\beta & -\alpha\beta \end{array}\right) \sim \left(\begin{array}{ccc|c} \boxed{1} & \beta & 1 & \alpha \\ 0 & \boxed{1-\alpha} & \alpha-\beta & -\alpha\beta \\ 0 & 0 & \boxed{-\alpha} & -\alpha^2 \end{array}\right).$$

This is in row-echelon form. Now recalling that we said that $1 - \alpha \neq 0$ and $-\alpha \neq 0$, this shows that the row-echelon form of the left-hand matrix above has every row and column leading. As we know, this implies that $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ is a basis for $\mathbb{F}^3$.

b) Let $S = \{\mathbf{b}_1, \mathbf{b}_1 + \mathbf{b}_2, \mathbf{b}_2 + \mathbf{b}_3, \mathbf{b}_3\}$. From part a), we know that $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ is a basis for $\mathbb{F}^3$ and hence is spanning. So every vector in $\mathbb{F}^3$ can be written as a linear combination of $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$. But $\mathbf{b}_2$ is a linear combination of vectors in S (since it is equal to $(\mathbf{b}_2 + \mathbf{b}_3) - \mathbf{b}_3$), and so every vector in $\mathbb{F}^3$ can be written as a linear combination of $\mathbf{b}_1, \mathbf{b}_1 + \mathbf{b}_2, \mathbf{b}_2 + \mathbf{b}_3, \mathbf{b}_3$. Hence S is a spanning set for $\mathbb{F}^3$. (In fact in general, if V is a finite-dimensional vector space and S is a subset of V, then S is a spanning set for V iff every vector in a basis for V belongs to $\operatorname{span}(S)$ (proof is left as an exercise for the reader).)

This set S cannot be a basis for $\mathbb{F}^3$ though, because every basis of $\mathbb{F}^3$ must have precisely three elements (because every basis for a given finite-dimensional vector space has the same number of elements, and this number here is 3 due to the result of part a)). Since S has four elements, it cannot be a basis. (Alternatively, as S has four elements and is a subset of a three-dimensional vector space, it cannot be linearly independent, so is not a basis.)

c) Let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ and $[\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix}$ be the coordinate vector of $\mathbf{v}$ with respect to the ordered basis $\mathcal{B}$ (so our goal is to find the λ_j for $j = 1, 2, 3$). Then by definition, we have $\lambda_j \in \mathbb{F}$ for $j = 1, 2, 3$, and

$$\lambda_1 \mathbf{b}_1 + \lambda_2 \mathbf{b}_2 + \lambda_3 \mathbf{b}_3 = \mathbf{v}.$$

Using the given coordinates of $\mathbf{b}_j$ and $\mathbf{v}$, the above equation is equivalent to the linear system defined by the first augmented matrix written in the solution to part a), which we row-reduced there:

$$\left(\begin{array}{ccc|c} 1 & \beta & 1 & \alpha \\ \alpha & 1 & 0 & 0 \\ \beta & 1 & \alpha & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} \boxed{1} & \beta & 1 & \alpha \\ 0 & \boxed{1-\alpha} & \alpha - \beta & -\alpha\beta \\ 0 & 0 & \boxed{-\alpha} & -\alpha^2 \end{array}\right).$$

(Note that column j of the left-hand matrix is the column for λ_j, for $j = 1, 2, 3$.) Now we back-substitute to find the λ_j. From row 3, we have

$$-\alpha\lambda_3 = -\alpha^2 \Rightarrow \boxed{\lambda_3 = \alpha} \quad \text{(since } \alpha \neq 0\text{)}.$$

Now from row 2, we have

$$\begin{aligned} &(1-\alpha)\lambda_2 + (\alpha - \beta)\lambda_3 = -\alpha\beta \\ \Rightarrow\ &(1-\alpha)\lambda_2 + (\alpha - \beta)\cdot\alpha = -1. && (\alpha\beta = 1 \text{ from multiplication table}) \end{aligned}$$

Observe that from the addition table, we see that for every $x \in \mathbb{F}$, we have $x + x = 0$ (that is, the main anti-diagonal of the table is all 0's). Thus $x = -x$ for all $x \in \mathbb{F}$ (in other words, subtraction is the same as addition in this field). So $\alpha - \beta = \alpha + \beta = 1$ (using the addition table to find $\alpha + \beta$). Also, $1 - \alpha = 1 + \alpha = \beta$ (from the addition table). So the equation for λ_2 becomes

$$\beta\lambda_2 + 1\cdot\alpha = -1 \Rightarrow \beta\lambda_2 + \alpha = -1 \Rightarrow \lambda_2 = \beta^{-1}(-1-\alpha).$$

Now, $-1 - \alpha \in \mathbb{F}$ (by closure properties of a field) and so $-1 - \alpha = -(-1-\alpha) = 1 + \alpha$ (recalling $x = -x$ for all $x \in \mathbb{F}$). But $1 + \alpha = \beta$, so $-1 - \alpha = \beta$. Hence we have

$$\lambda_2 = \beta^{-1}(-1-\alpha) = \beta^{-1} \times \beta \Leftrightarrow \boxed{\lambda_2 = 1}.$$

Finally, from row 1, we have

$$\begin{aligned} &\lambda_1 + \beta\lambda_2 + \lambda_3 = \alpha \\ \Rightarrow\ &\lambda_1 + \beta + \alpha = \alpha && (\text{since } \lambda_2 = 1, \lambda_3 = \alpha) \\ &\Rightarrow \lambda_1 = -\beta = \beta. \end{aligned}$$

Thus

$$[\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} \beta \\ 1 \\ \alpha \end{pmatrix}.$$

2 Linear transformations

Problem 1

[MATH1231 2012 S2 Q3 v)]
Suppose A is a 3×2 matrix and let A^T denote its transpose.

a) Prove that the column space of the matrix AA^T is a subset of the column space of A and deduce that $\operatorname{rank}\left(AA^T\right) \leq \operatorname{rank}(A)$.

b) Deduce that $\operatorname{nullity}\left(AA^T\right) \geq 1$, and explain why the matrix AA^T can never be the identity.

Solution 1

a) Let the underlying field be $\mathbb{F}$ (typically $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$; also this $\mathbb{F}$ has nothing to do with the one in the previous problem, it simply represents a generic field). Note that AA^T is well-defined (A and A^T are compatible for any A in fact). It is a 3×3 matrix (since A has size 3×2 and A^T has size 2×3). Denote by $\mathrm{CS}(M)$ the column space of any matrix M. Recall the following general fact:

Fact. *A vector* $\mathbf{u}$ *is in* $\mathrm{CS}(M)$ *if and only if* $\mathbf{u} = M\mathbf{v}$ *for some vector* $\mathbf{v}$.

In other words, a vector is in $\mathrm{CS}(M)$ iff it can be written as "M times some vector". This is equivalent to the fact that $\mathrm{CS}(M) = \operatorname{im}(M)$ for any matrix M, which is a fact stated on page 99 of the Algebra course pack notes (page 117 of the PDF).

We need to show that $\mathrm{CS}\left(AA^T\right) \subseteq \mathrm{CS}(A)$ (recall that the symbol $\subseteq$ means "is a subset of"). To show this, let $\mathbf{y} \in \mathrm{CS}\left(AA^T\right)$. Then using the blue fact, we have

$$\mathbf{y} = \left(AA^T\right)\mathbf{x}$$

for some vector $\mathbf{x} \in \mathbb{F}^3$. Since matrix multiplication is associative, we have

$$\mathbf{y} = A\left(A^T\mathbf{x}\right) = A\mathbf{w},$$

where $\mathbf{w} := A^T\mathbf{x} \in \mathbb{F}^2$. This shows that $\mathbf{y} \in \mathrm{CS}(A)$ (using the blue fact). This proves that $\mathrm{CS}\left(AA^T\right) \subseteq \mathrm{CS}(A)$.

Now, recall that if W and V are vector spaces and $W \subseteq V$, then $\dim W \leq \dim V$. Using this, we have (recalling that column spaces are vector spaces)

$$\mathrm{CS}\left(AA^T\right) \subseteq \mathrm{CS}(A) \Rightarrow \dim\left(\mathrm{CS}\left(AA^T\right)\right) \leq \dim\left(\mathrm{CS}(A)\right),$$

and so $\operatorname{rank}\left(AA^T\right) \leq \operatorname{rank}(A)$, since rank is by definition the dimension of the column space.

Remark. We can similarly show that for *any* matrices A and B such that AB is defined, we have that $\mathrm{CS}(AB) \subseteq \mathrm{CS}(A)$, and so $\operatorname{rank}(AB) \leq \operatorname{rank}(A)$. In fact, we also always have $\operatorname{rank}(AB) \leq \operatorname{rank}(B)$. Proof of this is left as an exercise for the interested reader.

b) We know that $\operatorname{nullity}\left(AA^T\right) + \operatorname{rank}\left(AA^T\right) = 3$, by the rank-nullity theorem (since AA^T is 3×3, so has three columns). Hence

$$\operatorname{nullity}\left(AA^T\right) = 3 - \operatorname{rank}\left(AA^T\right)$$

$$\begin{aligned} &\geq 3 - \operatorname{rank}(A) \qquad\qquad \left(\text{since } \operatorname{rank}\left(AA^T\right) \leq \operatorname{rank}(A) \text{ from part a)}\right) \\ &\geq 1, \end{aligned}$$

since $\operatorname{rank}(A) \leq 2$, since A is a 3×2 matrix. (Recall that in general, an $m \times n$ matrix has rank less than or equal to both m and n.)

Thus AA^T can never be the identity, because the identity has nullity 0^*, whereas AA^T has nullity at least 1.

Remark. We can similarly show that for *any* matrix A with strictly more rows than columns, AA^T can never equal the identity. More than this, we can show that for *any* matrix A with strictly more rows than columns, there is no matrix B for which AB equals the identity. Proof of this is left as an exercise for the interested reader.

Problem 2

[MATH1231 2013 S2 Q3 ii)]
Let

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 5 & 3 \\ 1 & 3 & 4 & 9 \end{pmatrix}.$$

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

b) Hence state the value of $\operatorname{nullity}(A)$. Give a reason.

Solution 2

a) To find $\ker(A)$ (and hence a basis for it), we solve $A\mathbf{x} = \mathbf{0}$, where $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$ and $\mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$.

As usual, this entails row-reducing the matrix A. We will only augment the zero vector at the very end, because this vector does not change when using elementary row operations, so there is no need to write it in every line. We have

$$\begin{aligned} A &= \begin{pmatrix} \boxed{1} & 2 & 3 & 4 \\ 2 & 3 & 5 & 3 \\ 1 & 3 & 4 & 9 \end{pmatrix} \\ &\xrightarrow[R_3 \rightsquigarrow R_3 - R_1]{R_2 \rightsquigarrow R_2 - 2R_1} \begin{pmatrix} \boxed{1} & 2 & 3 & 4 \\ 0 & \boxed{-1} & -1 & -5 \\ 0 & 1 & 1 & 5 \end{pmatrix} \\ &\xrightarrow{R_3 \rightsquigarrow R_3 - R_2} \left(\begin{array}{cccc|c} \boxed{1} & 2 & 3 & 4 & 0 \\ 0 & \boxed{-1} & -1 & -5 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right). \end{aligned}$$

* Because the kernel of the identity I clearly **only** contains the zero vector, so has dimension 0. Note that no non-zero vector can be in $\ker(I)$, because if $\mathbf{v} \neq \mathbf{0}$, then $I\mathbf{v} = \mathbf{v} \neq \mathbf{0}$.

Thus in solving $A\mathbf{x} = \mathbf{0}$, the last two columns in the left-hand part of the augmented part being non-leading implies that we can set $\boxed{x_3 = s}$ and $\boxed{x_4 = t}$ (free parameters). Then from row 2, we have

$$-x_2 - x_3 - 5x_4 = 0 \Rightarrow x_2 = -x_3 - 5x_4 \Rightarrow \boxed{x_2 = -s - 5t}.$$

Now, from row 1, we have

$$x_1 = -2x_2 - 3x_3 - 4x_4 = -2(-s - 5t) - 3s - 4t \Rightarrow \boxed{x_1 = -s + 6t}.$$

Thus $\mathbf{x}$ is in $\ker(A)$ iff

$$\begin{aligned}\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} &= \begin{pmatrix} -s + 6t \\ -s - 5t \\ s \\ t \end{pmatrix} \qquad \text{(for some scalars } s, t\text{)} \\ &= s \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 6 \\ -5 \\ 0 \\ 1 \end{pmatrix}.\end{aligned}$$

Thus as we can see, a basis for $\ker(A)$ is $\left\{ \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 6 \\ -5 \\ 0 \\ 1 \end{pmatrix} \right\}$.

Remark. We can (and should!) quickly **check** that each of the vectors we found in our kernel result in the zero vector when multiplied on the left by A. That is, quickly verify that

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 5 & 3 \\ 1 & 3 & 4 & 9 \end{pmatrix} \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 5 & 3 \\ 1 & 3 & 4 & 9 \end{pmatrix} \begin{pmatrix} 6 \\ -5 \\ 0 \\ 1 \end{pmatrix}.$$

(With some practice, you should be able to verify this mentally.) If one (or more) of the vectors in the basis for $\ker(A)$ you find do **not** give $\mathbf{0}$ when multiplied by A, then you will know you have made mistake somewhere.

b) From part a), we see that there are two vectors in a basis for $\ker(A)$. Hence the dimension of $\ker(A)$ is 2, i.e. $\text{nullity}(A) = 2$ (since $\text{nullity}(A)$ is by definition the dimension of $\ker(A)$).

Problem 3

[MATH1241 2014 S2 Q3 iii)]
Let $\mathcal{R}[\mathbb{R}]$ denote the vector space of all real-valued functions defined on $\mathbb{R}$. Let S be the subspace of $\mathcal{R}[\mathbb{R}]$ that is spanned by the **ordered** basis $\mathcal{B} = \{\cos x, \sin x\}$. Define the linear map $T : S \to S$ by

$$T(f) = f - 2f', \quad \text{where } f' \equiv \frac{\mathrm{d}f}{\mathrm{d}x}.$$

a) Calculate the matrix C that represents T with respect to the basis $\mathcal{B}$.

b) State the rank of the matrix C found in part a).

c) From part b), what can be deduced about the solutions $y \in S$ of the differential equation

$$y - 2y' = g,$$

where g is a given function in S?

d) Using parts a) and c), or otherwise, find all solutions $y \in S$ of the differential equation

$$y - 2y' = \cos x.$$

Solution 3

a) Note that since $S = \text{span}\,(\mathcal{B})$ has dimension 2 (since the basis $\mathcal{B}$ has precisely two elements), the matrix of T with respect to $\mathcal{B}$ must be a 2×2 matrix. From the "General Matrix Representation Theorem" (Theorem 1 of page 110 (page 128 of the electronic PDF) of Algebra course pack notes), we know that **the first column of C is $[T(\cos x)]_{\mathcal{B}}$** and **the second column of C is $[T(\sin x)]_{\mathcal{B}}$**.

Hence as usual, to get the columns of T, we evaluate T at each of the basis vectors (in order as they appear in $\mathcal{B}$) and then write down the results as their coordinate vectors (with respect to $\mathcal{B}$). Evaluating T at the first basis function,

$$\begin{aligned} T(\cos x) &= \cos x - 2\frac{\mathrm{d}}{\mathrm{d}x}(\cos x) \\ &= \cos x + 2\sin x = 1\cos x + 2\sin x, \end{aligned}$$

which implies that $[T(\cos x)]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$. Hence **the first column of C is $[T(\cos x)]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$**.

Evaluating T at the second basis function,

$$\begin{aligned} T(\sin x) &= \sin x - 2\frac{\mathrm{d}}{\mathrm{d}x}(\sin x) \\ &= \sin x - 2\cos x \\ &= -2\cos x + 1\sin x, \end{aligned}$$

so $[T(\sin x)]_{\mathcal{B}} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$. Thus **the second column of C is $[T(\sin x)]_{\mathcal{B}} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$**. Remembering that C is 2×2, we thus have found both columns of C, so we can write down C as

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

b) Clearly, C has two non-zero columns and the columns are not constant multiples of one another. Therefore, C has two linearly independent columns and so $\text{rank}\,(C) = 2$.

c) The differential equation $y - 2y' = g$ can be written as $T(y) = g$ (by definition of T).

Since C is a 2×2 matrix with rank 2, it is an invertible matrix. Therefore the matrix of T from part a) is an invertible matrix, which implies that the linear map $T : S \to S$ is an invertible linear map. Thus T is both one-to-one and onto, so the equation $T(y) = g$ has a unique solution for any $g \in S$. In other words, the differential equation in question **has a unique solution** for $y \in S$ (i.e. there exists one and only one solution $y \in S$, for any given $g \in S$).

d) The given differential equation can be written as the equation $T(y) = \cos x$. Let the coordinate vector of y with respect to $\mathcal{B}$ be $[y]_{\mathcal{B}} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$. Clearly the coordinate vector of $\cos x$ with respect to $\mathcal{B}$ is $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Also, by definition of the matrix of T, we have $[T(y)]_{\mathcal{B}} = C[y]_{\mathcal{B}}$. So writing $T(y) = \cos x$ as a matrix equation (using the matrix C), we have

$$T(y) = \cos x \Leftrightarrow C[y]_{\mathcal{B}} = [\cos x]_{\mathcal{B}}$$

(since two elements of a vector space are equal iff their coordinates with respect to a basis are equal)

$$\Leftrightarrow \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \qquad \text{(using our answer for } C \text{ in part a))}$$

This is a pair of simultaneous equations for y_1 and y_2. From the second row's equation, we get

$$2y_1 + y_2 = 0 \Rightarrow y_2 = -2y_1.$$

The first row's equation thus gives us

$$\begin{aligned} y_1 - 2y_2 &= 1 \\ \Rightarrow y_1 - 2(-2y_1) &= 1 \\ \Rightarrow 5y_1 &= 1 \\ \Rightarrow y_1 &= \frac{1}{5}. \end{aligned}$$

Now since $y_2 = -2y_1$, we have $y_2 = -\frac{2}{5}$. Thus $[y]_{\mathcal{B}} = \begin{pmatrix} 1/5 \\ -2/5 \end{pmatrix}$, that is, $\boxed{y = \frac{1}{5}\cos x - \frac{2}{5}\sin x}$. This is the (one and only) solution to the differential equation for $y \in S$.

Remark. You should check your answer by quickly verifying that $\left(\frac{1}{5}\cos x - \frac{2}{5}\sin x\right) - 2\left(\frac{1}{5}\cos x - \frac{2}{5}\sin x\right)'$ indeed equals $\cos x$ (if the answer you got for y did not satisfy $y - 2y' = \cos x$ when you tried to verify it, you would know you had made a mistake somewhere).

3 Eigenvalues and eigenvectors

Problem 1

[MATH1231 2013 S2 Q3 iii)]
Let

$$C = \begin{pmatrix} 1 & 2 \\ -2 & 5 \end{pmatrix}.$$

a) Find the eigenvalue(s) of C and for each eigenvalue, find the corresponding eigenvectors.

b) Is C diagonalisable? Give a reason for your answer.

Solution 1

a) We have

$$C - \lambda I = \begin{pmatrix} 1-\lambda & 2 \\ -2 & 5-\lambda \end{pmatrix}.$$

Hence the characteristic polynomial of C is

$$\begin{aligned} p_C(\lambda) := \det(C - \lambda I) &= (1-\lambda)(5-\lambda) - 2 \times (-2) \\ &= 5 - \lambda - 5\lambda + \lambda^2 + 4 \\ &= \lambda^2 - 6\lambda + 9 \\ &= (\lambda - 3)^2. \end{aligned}$$

The eigenvalues are the roots of $p_C(\lambda)$, which is just 3 (a repeated eigenvalue). The corresponding eigenspace is

$$\begin{aligned} E_3 &= \ker(C - 3I) \\ &= \ker \begin{pmatrix} 1-3 & 2 \\ -2 & 5-3 \end{pmatrix} \\ &= \ker \begin{pmatrix} -2 & 2 \\ -2 & 2 \end{pmatrix} \\ &= \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}. \end{aligned}$$

In summary, the only eigenvalue of C is 3, with corresponding eigenvector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ (or any non-zero scalar multiple of this).

b) **No.** For C to be diagonalisable, it must have two linearly independent eigenvectors (since it is a 2×2 matrix). But from part a), we see that C only has one independent eigenvector. Therefore, C is not diagonalisable. (Remember in general, an $n \times n$ matrix is diagonalisable iff it has n linearly independent eigenvectors.)

Problem 2

[MATH1231 2014 S2 Q2 iv)]

Consider the set S consisting of the vectors $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$ from $\mathbb{R}^3$ and let $\mathbf{u} = \begin{pmatrix} 1 \\ -1 \\ 12 \end{pmatrix}$.

a) Find scalars λ and μ such that $\mathbf{u} = \lambda \mathbf{v}_1 + \mu \mathbf{v}_2$.

b) A linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ has $\mathbf{v}_1, \mathbf{v}_2$ as eigenvectors with eigenvalues 2 and -1, respectively.

α) Find $T(\mathbf{u})$ as a linear combination of $\mathbf{v}_1, \mathbf{v}_2$.

β) Denote $T(T(\mathbf{u}))$ by $T^2(\mathbf{u})$, $T(T(T(\mathbf{u})))$ by $T^3(\mathbf{u})$, and so on. Express $T^n(\mathbf{u})$ as a linear combination of $\mathbf{v}_1, \mathbf{v}_2$, where n is a positive integer.

Solution 2

a) The equation $\mathbf{u} = \lambda\mathbf{v}_1 + \mu\mathbf{v}_2$ can be written as in augmented matrix form as

$$\left(\begin{array}{cc|c} 1 & 2 & 1 \\ 1 & 3 & -1 \\ 2 & -1 & 12 \end{array}\right),$$

where the first column is for λ and the second is for μ. Performing the row operations $R_2 \rightsquigarrow R_2 - R_1$ and $R_3 \rightsquigarrow R_3 - 2R_1$, we see

$$\left(\begin{array}{cc|c} 1 & 2 & 1 \\ 1 & 3 & -1 \\ 2 & -1 & 12 \end{array}\right) \sim \left(\begin{array}{cc|c} \boxed{1} & 2 & 1 \\ 0 & \boxed{1} & -2 \\ 0 & -5 & 10 \end{array}\right).$$

Performing $R_3 \rightsquigarrow R_3 + 5R_2$, the row-echelon form is

$$\left(\begin{array}{cc|c} \boxed{1} & 2 & 1 \\ 0 & \boxed{1} & -2 \\ 0 & 0 & 0 \end{array}\right).$$

Now from row 2, we have $\boxed{\mu = -2}$. So from row 1, we have

$$\begin{aligned} \lambda + 2\mu &= 1 \\ \Rightarrow \lambda - 4 &= 1 && (\text{as } \mu = -2) \\ \Rightarrow \boxed{\lambda = 5}&. \end{aligned}$$

So $\lambda = 5$ and $\mu = -2$.

b) Let $\lambda_1 = 2$ and $\lambda_2 = -1$, so from the given information on the eigenvalues and eigenvectors of T, $T(\mathbf{v}_j) = \lambda_j\mathbf{v}_j$, for $j = 1, 2$.

α) We have $\mathbf{u} = 5\mathbf{v}_1 + 2\mathbf{v}_2$. Hence

$$\begin{aligned} T(\mathbf{u}) &= T(5\mathbf{v}_1 + 2\mathbf{v}_2) \\ &= 5T(\mathbf{v}_1) + 2T(\mathbf{v}_2) && (\text{since } T \text{ is linear}) \\ &= 5\lambda_1\mathbf{v}_1 + 2\lambda_2\mathbf{v}_2 \\ &= 10\mathbf{v}_1 - 2\mathbf{v}_2. && (\text{since } \lambda_1 = 2 \text{ and } \lambda_2 = -1) \end{aligned}$$

β) It is easy to show by induction the following two general facts about linear maps:

- for any linear map T, T^n is a linear map for any positive integer n (follows from induction and the fact that a composition of linear maps is linear);
- if T is a linear map with $T(\mathbf{v}) = \lambda\mathbf{v}$ for some $\mathbf{v}$ in the domain of T and scalar λ, then $T^n(\mathbf{v}) = \lambda^n\mathbf{v}$ for all positive integers n.

Using these facts, we have for any $n \in \mathbb{Z}^+$

$$\begin{aligned} T^n(\mathbf{u}) &= T^n(5\mathbf{v}_1 + 2\mathbf{v}_2) \\ &= 5T^n(\mathbf{v}_1) + 2T^n(\mathbf{v}_2) \\ \Rightarrow T^n(\mathbf{u}) &= (5 \times \lambda_1^n)\mathbf{v}_1 + (2 \times \lambda_2^n)\mathbf{v}_2, && \text{(using the "}T^n(\mathbf{v}) = \lambda^n\mathbf{v}\text{" fact)} \end{aligned}$$

which is the answer. (Observe that putting $n = 1$ gives us our answer for part α), as it should.)

Problem 3

[MATH1241 2015 S2 Q3 iv)]
A linear transformation $P : V \to V$ is said to be **idempotent** if $P(P(\mathbf{v})) = P(\mathbf{v})$ for all $\mathbf{v} \in V$ (in other words, $P^2 = P$).

a) Show that the only possible eigenvalues for an idempotent linear transformation are 0 and 1.

b) Show that if P is idempotent and P is neither the zero nor the identity transformation on V, then both 0 and 1 are eigenvalues.

Solution 3

For brevity, we will denote $P\mathbf{v} := P(\mathbf{v})$ and $P^2\mathbf{v} := P(P(\mathbf{v}))$.

a) Suppose λ is an eigenvalue of P, then $P\mathbf{v} = \lambda\mathbf{v}$ for some non-zero $\mathbf{v} \in V$. Applying P to both sides of this equation, we have

$$\begin{aligned} P(P\mathbf{v}) &= P\mathbf{v} \\ \Rightarrow P(\lambda\mathbf{v}) &= \lambda\mathbf{v} && \text{(since } P\mathbf{v} = \lambda\mathbf{v}) \\ \Rightarrow \lambda P\mathbf{v} &= \lambda\mathbf{v} && \text{(linearity of } P) \\ \Rightarrow \lambda(\lambda\mathbf{v}) &= \lambda\mathbf{v} && \text{(since } P\mathbf{v} = \lambda\mathbf{v}) \\ \Rightarrow \lambda^2\mathbf{v} &= \lambda\mathbf{v} \\ \Rightarrow \left(\lambda^2 - \lambda\right)\mathbf{v} &= \mathbf{0}. \end{aligned}$$

Since $\mathbf{v} \neq \mathbf{0}$, this implies* that

$$\begin{aligned} \lambda^2 - \lambda &= 0 \\ \Rightarrow \lambda(\lambda - 1) &= 0 \\ \Rightarrow \lambda &= 0 \text{ or } 1. \end{aligned}$$

Hence the only possible eigenvalues for an idempotent linear transformation are 0 and 1.

b) Assume P is idempotent and P is neither the zero nor the identity transformation on V. Since P is not the zero map, there exists a $\mathbf{v} \in V$ with $P\mathbf{v} \neq \mathbf{0}$. (Such $\mathbf{v}$ must be non-zero, since P maps zero to zero.) So we can write $P\mathbf{v} = \mathbf{w}$, where $\mathbf{w} \in V$ and $\mathbf{w} \neq \mathbf{0}$. Applying P to both sides of this equation, we have

$$P^2\mathbf{v} = P\mathbf{w}$$

* Recall that in a vector space, if $\alpha\mathbf{v} = \mathbf{0}$ and $\mathbf{v} \neq \mathbf{0}$, then $\alpha = 0$.

$$\begin{aligned}
&\Rightarrow P\mathbf{v} = P\mathbf{w} && \text{(remember } P^2 = P\text{, i.e. } P \text{ is idempotent)}\\
&\Rightarrow \mathbf{w} = P\mathbf{w}. && \text{(since } P\mathbf{v} = \mathbf{w}\text{)}
\end{aligned}$$

Thus $P\mathbf{w} = \mathbf{w} = 1\mathbf{w}$. Recalling that we said that $\mathbf{w} \in V$ and $\mathbf{w} \neq \mathbf{0}$, we have shown that there exists $\mathbf{w} \in V$ with $\mathbf{w} \neq \mathbf{0}$ with $P\mathbf{w} = 1\mathbf{w}$. This shows that 1 is an eigenvalue of P.

Now we show that 0 is an eigenvalue of P. By assumption that P is not the identity map in V, there exists $\mathbf{u} \in V$ with $P\mathbf{u} \neq \mathbf{u}$. So we can write $P\mathbf{u} = \mathbf{u} + \boldsymbol{\epsilon}$, where $\boldsymbol{\epsilon} \in V$ and $\boldsymbol{\epsilon} \neq \mathbf{0}$. Applying P to both sides, we have

$$\begin{aligned}
&P^2\mathbf{u} = P(\mathbf{u} + \boldsymbol{\epsilon}) &&\\
&\Rightarrow P\mathbf{u} = P\mathbf{u} + P\boldsymbol{\epsilon} && \text{(linearity and idempotency of } P\text{)}\\
&\Rightarrow P\boldsymbol{\epsilon} = \mathbf{0} = 0\boldsymbol{\epsilon}. &&
\end{aligned}$$

Thus we have shown that there exists $\boldsymbol{\epsilon} \in V$ with $\boldsymbol{\epsilon} \neq \mathbf{0}$ and $P\boldsymbol{\epsilon} = 0\boldsymbol{\epsilon}$. This shows that 0 is an eigenvalue of P, and completes the proof.

4 Introduction to probability and statistics

Problem 1

[MATH1231 2012 S2 Q3 iv)/MATH1241 2012 S2 Q3 i)]
A six-sided die, with faces numbered 1 to 6, is suspected of being unfair, so that the number 6 will occur more frequently than should happen by chance. During 300 test rolls of the die, the number 6 occurred 68 times.

a) Write down an expression for a tail probability that measures the chance of rolling a 6 at least 68 times.

b) Use the normal approximation to the binomial to estimate this probability.

c) Is this evidence that the die is unfair?

Solution 1

a) Let X be the random variable "number of 6's rolled in 300 test rolls of the die". Then under the assumption the die is **fair**, the rolls are independent, each with probability $p = \frac{1}{6}$ of resulting in a 6. Thus $X \sim \text{Bin}\left(300, \frac{1}{6}\right)$ (written $B\left(300, \frac{1}{6}\right)$ in the course pack, i.e. Binomial with $n = 300, p = \frac{1}{6}$). Hence the tail probability that measures the chance of rolling a 6 at least 68 times is

$$\begin{aligned}
\mathbb{P}(X \geq 68) &= \sum_{k=68}^{300} \mathbb{P}(X = k)\\
&= \sum_{k=68}^{300} \binom{300}{k} \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{300-k},
\end{aligned}$$

using the formula for the PMF of a Binomial random variable (with $n = 300, p = \frac{1}{6}$).

b) Since $X \sim \text{Bin}(n, p)$ where $n = 300, p = \frac{1}{6}$, using the formula for the mean and variance of a Binomial random variable, we have

$$\mathbb{E}[X] = np = 300 \times \frac{1}{6} \Rightarrow \mathbb{E}[X] \equiv \mu_X = 50$$

and

$$\text{Var}(X) = np(1-p) = 50 \times \frac{5}{6} = \frac{125}{3} \Rightarrow \text{SD}(X) = \sqrt{\text{Var}(X)} \equiv \sigma_X = \sqrt{\frac{125}{3}}.$$

Since n is large, the Normal approximation to the Binomial implies that we can approximate the distribution of X by a random variable $Y \sim \mathcal{N}(\mu_X, \sigma_X^2)$. Recall that to use the Normal approximation to the Binomial, we must apply the **continuity correction** $\mathbb{P}(X \geq 68) \approx \mathbb{P}(Y \geq 67.5)$. Thus we have

$$\begin{aligned}
\mathbb{P}(X \geq 68) &\approx \mathbb{P}(Y \geq 67.5) \\
&= \mathbb{P}\left(Z \geq \frac{67.5 - \mu_X}{\sigma_X}\right) \qquad \left(\text{where } Z := \tfrac{Y - \mu_X}{\sigma_X} \sim \mathcal{N}(0,1) \text{ since } Y \sim \mathcal{N}(\mu_X, \sigma_X^2)\right) \\
&= \mathbb{P}\left(Z \geq \frac{67.5 - 50}{\sqrt{\frac{125}{3}}}\right) \\
&\approx \mathbb{P}(Z \geq 2.71) \quad \left(\text{since } \tfrac{67.5-50}{\sqrt{\frac{125}{3}}} = 2.71 \text{ to two decimal places, using a calculator}\right) \\
&= 1 - \mathbb{P}(Z \leq 2.71) \\
&= 1 - \Phi(2.71) \qquad (\text{where } \Phi(\cdot) \text{ is the standard normal CDF}) \\
&\approx 1 - 0.9966 \\
&\qquad (\text{since } \Phi(2.71) \approx 0.9966 \text{ from the standard normal probability table}) \\
&= 0.0034.
\end{aligned}$$

So the final answer is $\mathbb{P}(X \geq 68) \approx 0.0034$. That is, the probability is estimated as 0.34% using the Normal approximation.

c) Yes. The above shows that assuming the die is fair, the probability of getting a result "at least as extreme as obtained" (i.e. at least 68 rolls coming up as 6) is about 0.0034, which is very low (less than 5% (in fact, less than 1%)). Therefore, this **is** evidence that the die is unfair (has probability greater than $\frac{1}{6}$ of showing a 6).

Problem 2

[MATH1231 2013 S2 Q1 iii)]
A collection of discs consists of DVDs and Blu-ray discs. In the collection, 75% are DVDs and 25% are Blu-ray discs. Among the DVDs, 60% are movies. Among the Blu-ray discs, 90% are movies.

a) Find the probability that a disc chosen randomly from the collection is a movie.

b) Find the probability that a randomly chosen disc from the collection is a Blu-ray disc given that it is a movie.

Solution 2

Let D, B, M be the event that a disc chosen randomly from the collection is a DVD, Blu-ray disc, movie, respectively (note that $D = B^c$). Note that from the given data, we have

$$\mathbb{P}(D) = 0.75, \quad \mathbb{P}(B) = 0.25, \quad \mathbb{P}(M \mid D) = 0.6, \quad \mathbb{P}(M \mid B) = 0.9.$$

a) We are asked to find $\mathbb{P}(M)$. We have

$$\begin{aligned}\mathbb{P}(M) &= \mathbb{P}(M \mid D)\,\mathbb{P}(D) + \mathbb{P}(M \mid B)\,\mathbb{P}(B) && \text{(Total Probability Rule)}\\ &= 0.6 \times 0.75 + 0.9 \times 0.25 && \text{(given data)}\\ &= 0.45 + 0.225\\ &= 0.675.\end{aligned}$$

Hence the answer is 0.675.

b) We are asked to find $\mathbb{P}(B \mid M)$. We have

$$\begin{aligned}\mathbb{P}(B \mid M) &= \frac{\mathbb{P}(M \mid B)\,\mathbb{P}(B)}{\mathbb{P}(M)} && \text{(Bayes' Rule)}\\ &= \frac{0.9 \times 0.25}{0.675} && \text{(given data and using part a))}\\ &= \frac{1}{3}.\end{aligned}$$

Hence the answer is $\frac{1}{3}$.

Problem 3

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

$$f(x) = \begin{cases} 3(1-x)^2 & \text{for } 0 \le x \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

a) Sketch the graph of $y = f(x)$.

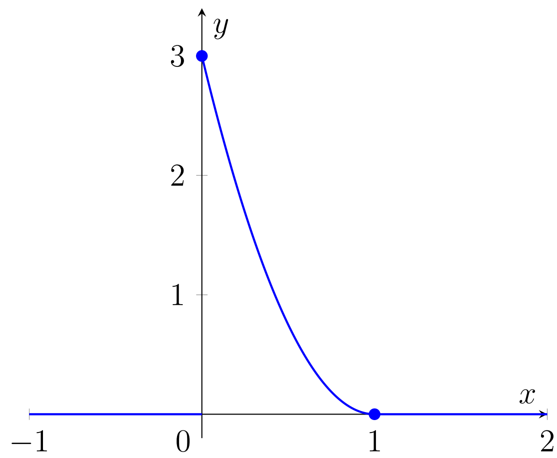
b) Find $\mathbb{E}[X]$ and $\mathrm{Var}(X)$.

c) Find $\mathbb{P}\left(\frac{1}{2} < \sin(\pi X) < \frac{1}{\sqrt{2}}\right)$.

d) The **median** of a distribution is defined to be the real number m such that $\mathbb{P}(X \le m) = \frac{1}{2}$. Find the median of the above distribution.

Solution 3

a) The graph is a parabolic arc between $x = 0$ and 1, and $f(x)$ identically zero outside this interval. The y-intercept is 3 because $f(0) = 3(1-0)^2 = 3$, and the vertex of the parabola is at $x = 1$, when $y = 0$. The graph of $y = f(x)$ is sketched below.



b) We have

$$\begin{aligned}
\mathbb{E}[X] &= \int_0^1 x f(x)\,\mathrm{d}x \\
&= \int_0^1 x \cdot 3(1-x)^2\,\mathrm{d}x \\
&= 3\int_0^1 x(1-2x+x^2)\,\mathrm{d}x \\
&= 3\int_0^1 \left(x-2x^2+x^3\right)\,\mathrm{d}x \\
&= 3\left[\frac{x^2}{2}-\frac{2x^3}{3}+\frac{x^4}{4}\right]_0^1 \\
&= 3\left(\frac{1}{2}-\frac{2}{3}+\frac{1}{4}\right) \\
&= 3\left(\frac{6-8+3}{12}\right) \\
&= \frac{1}{4}.
\end{aligned}$$

Also,

$$\begin{aligned}
\mathbb{E}\left[X^2\right] &= \int_0^1 x^2 f(x)\,\mathrm{d}x \\
&= \int_0^1 x^2 \cdot 3(1-x)^2\,\mathrm{d}x \\
&= 3\int_0^1 x^2\left(1-2x+x^2\right)\,\mathrm{d}x \\
&= 3\int_0^1 \left(x^2-2x^3+x^4\right)\,\mathrm{d}x \\
&= 3\left(\frac{1}{3}-\frac{2}{4}+\frac{1}{5}\right) \\
&= 3\times\frac{20-30+12}{60}
\end{aligned}$$

$$= \frac{1}{10}.$$

Hence

$$\begin{aligned}
\operatorname{Var}(X) &= \mathbb{E}\left[X^2\right] - (\mathbb{E}[X])^2 \\
&= \frac{1}{10} - \left(\frac{1}{4}\right)^2 \\
&= \frac{1}{10} - \frac{1}{16} \\
&= \frac{6}{160} \\
\Rightarrow \operatorname{Var}(X) &= \frac{3}{80}.
\end{aligned}$$

c) Note that X can only take values between 0 and 1 (since its PDF is identically zero outside the interval $(0,1)$). For $t \in [0,1]$, we have $\frac{1}{2} < \sin(\pi t) < \frac{1}{\sqrt{2}}$ iff $\frac{1}{6} < t < \frac{1}{4}$ or $\frac{3}{4} < t < \frac{5}{6}$ (draw the graph of $y = \sin(\pi t)$ for $0 \le t \le 1$ to see this). Therefore,

$$\begin{aligned}
\mathbb{P}\left(\frac{1}{2} < \sin(\pi X) < \frac{1}{\sqrt{2}}\right) &= \mathbb{P}\left(\frac{1}{6} < X < \frac{1}{4}\right) + \mathbb{P}\left(\frac{3}{4} < X < \frac{5}{6}\right) \\
&= \left[F\left(\frac{1}{4}\right) - F\left(\frac{1}{6}\right)\right] + \left[F\left(\frac{5}{6}\right) - F\left(\frac{3}{4}\right)\right],
\end{aligned}$$

where F is the CDF of X. Note that for $x \in [0,1]$, the CDF of X is

$$\begin{aligned}
F(x) &= \int_0^x f(t)\,\mathrm{d}t \\
&= \int_0^x 3(1-t)^2\,\mathrm{d}t \\
&= \left[-(1-t)^3\right]_0^x \\
&= -(1-x)^3 - \left(-(1-0)^3\right) \\
\Rightarrow F(x) &= 1 - (1-x)^3.
\end{aligned}$$

(You should check to make sure that this expression satisfies $F(0) = 0, F(1) = 1$, and is an increasing function. This is because $\mathbb{P}(X \le 0) = 0, \mathbb{P}(X \le 1) = 1$, and a CDF must be increasing.)

Thus $F(b) - F(a) = (1-a)^3 - (1-b)^3$ for all $a, b \in [0,1]$, so

$$\begin{aligned}
\mathbb{P}\left(\frac{1}{2} < \sin(\pi X) < \frac{1}{\sqrt{2}}\right) &= \left[F\left(\frac{1}{4}\right) - F\left(\frac{1}{6}\right)\right] + \left[F\left(\frac{5}{6}\right) - F\left(\frac{3}{4}\right)\right] \\
&= \left[\left(1-\frac{1}{6}\right)^3 - \left(1-\frac{1}{4}\right)^3\right] + \left[\left(1-\frac{3}{4}\right)^3 - \left(1-\frac{5}{6}\right)^3\right] \\
&= \frac{145}{864}.
\end{aligned}$$

d) We have $\mathbb{P}(X \le m) = \frac{1}{2}$, i.e $F(m) = \frac{1}{2}$. Using our formula for $F(x)$ from the previous part, we have

$$\begin{aligned} 1-(1-m)^3 &= \frac{1}{2} \\ \Rightarrow (1-m)^3 &= \frac{1}{2} \\ \Rightarrow 1-m &= \frac{1}{\sqrt[3]{2}} \\ \Rightarrow m &= 1-\frac{1}{\sqrt[3]{2}}. \end{aligned}$$

Thus the median is $1-\dfrac{1}{\sqrt[3]{2}}$.

Problem 4

(The log-normal distribution) Let X be a normal random variable with mean μ and standard deviation σ. Find the probability density function of the random variable $Y := e^X$. (Y is said to be a *log-normal* random variable with parameters μ and σ.)

Solution 4

Recall that the standard normal PDF is given by

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} = \Phi'(z).$$

Note that since $X \sim \mathcal{N}(\mu, \sigma^2)$, we have that $\frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$. We use a typical approach to computing a PDF: *first find the CDF, then differentiate it*. Let the CDF of Y be $F_Y(y)$. For $y > 0$, the CDF of Y is

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \le y) \\ &= \mathbb{P}\left(e^X \le y\right) && \text{(since } Y = e^X\text{)} \\ &= \mathbb{P}(X \le \ln y) && \text{(since } e^X \le Y \text{ is equivalent to } X \le \ln Y\text{)} \\ &= \mathbb{P}\left(Z \le \frac{\ln y - \mu}{\sigma}\right), \quad \text{where } Z := \frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1) && \text{(standardising)} \\ \Rightarrow F_Y(y) &= \Phi\left(\frac{\ln y - \mu}{\sigma}\right), \end{aligned}$$

where $\Phi(z) = \mathbb{P}(Z \le z)$ is the standard normal CDF. Therefore, the PDF of Y for $y > 0$ is

$$\begin{aligned} f_Y(y) &= \frac{\mathrm{d}}{\mathrm{d}y}\left(F_Y(y)\right) \\ &= \frac{\mathrm{d}}{\mathrm{d}y}\left(\Phi\left(\frac{\ln y - \mu}{\sigma}\right)\right) \end{aligned}$$

$$
\begin{aligned}
&= \Phi'\left(\frac{\ln y - \mu}{\sigma}\right) \cdot \frac{\mathrm{d}}{\mathrm{d}y}\left(\frac{\ln y - \mu}{\sigma}\right) && \text{(chain rule)} \\
&= \phi\left(\frac{\ln y - \mu}{\sigma}\right) \cdot \left(\frac{1}{\sigma y}\right) && \text{(recalling that } \Phi' = \phi) \\
&= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln y - \mu}{\sigma}\right)^2} \cdot \left(\frac{1}{\sigma y}\right) \\
&\Rightarrow \boxed{f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma y} e^{-\frac{1}{2}\left(\frac{\ln y - \mu}{\sigma}\right)^2}}.
\end{aligned}
$$

This is the PDF for $y > 0$. The PDF for $y \leq 0$ is identically 0, since the range of possible values for Y is the strictly positive reals, since $Y = e^X$ and exponentials are positive.