Further Mathematical Methods (Linear Algebra)

Solutions For Problem Sheet 2

In this Problem Sheet, we looked at some questions involving linear transformations. As we did not cover this material in the lectures, it is vital that you read the hand-out 'Lecture 3: Linear Transformations' before you attempted these exercises. We also applied the two tests for linear independence which we devised in the lectures and examined how solutions to sets of simultaneous equations can be analysed in terms of the null space of a matrix.

1. We are asked to prove Theorems 3.3 and 3.7 from the hand-out, namely:

> **Theorem 3.3** Let V and W be vector spaces. If $T : V \to W$ is a linear transformation, then
>
> **a**. $T(\mathbf{0}) = \mathbf{0}$.
>
> **b**. $T(-\mathbf{v}) = -T(\mathbf{v})$.
>
> **c**. $T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v})$.
>
> for all vectors $\mathbf{u}, \mathbf{v} \in V$.
>
> **Proof:** Let V and W be vector spaces. If $T : V \to W$ is a linear transformation, then Definition 3.1 tells us that for any vectors $\mathbf{u}, \mathbf{v} \in V$ and any scalars α_1, α_2,
>
> $$T(\alpha_1\mathbf{u} + \alpha_2\mathbf{v}) = \alpha_1 T(\mathbf{u}) + \alpha_2 T(\mathbf{v}),$$
>
> where $T(\mathbf{u}), T(\mathbf{v}) \in W$. The results that we have to prove follow directly from this definition, i.e. for any vectors $\mathbf{u}, \mathbf{v} \in V$:
>
> **a**. Setting $\alpha_1 = \alpha_2 = 0$ in the definition we get
>
> $$\alpha_1\mathbf{u} + \alpha_2\mathbf{v} = \mathbf{0} + \mathbf{0} = \mathbf{0} \text{ and } \alpha_1 T(\mathbf{u}) + \alpha_2 T(\mathbf{v}) = \mathbf{0} + \mathbf{0} = \mathbf{0}.$$
>
> Thus, $T(\mathbf{0}) = \mathbf{0}$.
>
> **b**. Setting $\alpha_1 = 0$ and $\alpha_2 = -1$ in the definition we get
>
> $$\alpha_1\mathbf{u} + \alpha_2\mathbf{v} = \mathbf{0} - \mathbf{v} = -\mathbf{v} \text{ and } \alpha_1 T(\mathbf{u}) + \alpha_2 T(\mathbf{v}) = \mathbf{0} - T(\mathbf{v}) = -T(\mathbf{v}).$$
>
> Thus, $T(-\mathbf{v}) = -T(\mathbf{v})$.
>
> **c**. Setting $\alpha_1 = 1$ and $\alpha_2 = -1$ in the definition we get
>
> $$\alpha_1\mathbf{u} + \alpha_2\mathbf{v} = \mathbf{v} - \mathbf{u} \text{ and } \alpha_1 T(\mathbf{u}) + \alpha_2 T(\mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v}).$$
>
> Thus, $T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v})$.
>
> as required.

and,

> **Theorem 3.7** Let V and W be vector spaces. If $T : V \to W$ is a linear transformation, then the null space of T is a subspace of V and the range of T is a subspace of W.
>
> **Proof:** We are given that V and W are vector spaces and that $T : V \to W$ is a linear transformation. We prove the theorem by taking each case in turn:

- To show that the null space of T, denoted by $N(T)$, is a subspace of V we note that

$$N(T) = \{\mathbf{v} \in V \,|\, T(\mathbf{v}) = \mathbf{0}\},$$

by Definition 3.6. Now, clearly, if $\mathbf{x} \in N(T)$, then $\mathbf{x} \in V$ and so $N(T)$ is a sub*set* of V. To show that it is also a sub*space* of V, by Theorem 1.4, we have to show that it is closed under vector addition and scalar multiplication. So, considering any two vectors $\mathbf{x}, \mathbf{y} \in N(T)$, i.e. $T(\mathbf{x}) = \mathbf{0}$ and $T(\mathbf{y}) = \mathbf{0}$, we have:
  - Closure under vector addition since:

  $$T(\mathbf{x}+\mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y}) = \mathbf{0} + \mathbf{0} = \mathbf{0},$$

  and so $\mathbf{x}+\mathbf{y} \in N(T)$.
  - Closure under scalar multiplication since for any scalar α:

  $$T(\alpha\mathbf{x}) = \alpha T(\mathbf{x}) = \alpha\mathbf{0} = \mathbf{0},$$

  and so $\alpha\mathbf{x} \in N(T)$.

  Thus, $N(T)$ is a subspace of V, as required.
- To show that the range of T, denoted by $R(T)$, is a subspace of W we note that

$$R(T) = \{\mathbf{w} \in W \,|\, T(\mathbf{v}) = \mathbf{w} \text{ for some } \mathbf{v} \in V\},$$

by Definition 3.6. Now, clearly, if $\mathbf{x} \in R(T)$, then $\mathbf{x} \in W$ and so $R(T)$ is a sub*set* of W. To show that it is also a sub*space* of W, by Theorem 1.4, we have to show that it is closed under vector addition and scalar multiplication. So, considering any two vectors $\mathbf{x}, \mathbf{y} \in R(T)$, i.e. $T(\mathbf{x}') = \mathbf{x}$ and $T(\mathbf{y}') = \mathbf{y}$ for some $\mathbf{x}', \mathbf{y}' \in V$, we have:
  - Closure under vector addition since:

  $$T(\mathbf{x}'+\mathbf{y}') = T(\mathbf{x}') + T(\mathbf{y}') = \mathbf{x} + \mathbf{y},$$

  and so $\mathbf{x}+\mathbf{y} \in R(T)$ as there is a vector $\mathbf{x}'+\mathbf{y}' \in V$ such that $T(\mathbf{x}'+\mathbf{y}') = \mathbf{x}+\mathbf{y}$.[1]
  - Closure under scalar multiplication since for any scalar α:

  $$T(\alpha\mathbf{x}') = \alpha T(\mathbf{x}') = \alpha\mathbf{x},$$

  and so $\alpha\mathbf{x} \in R(T)$ as there is a vector $\alpha\mathbf{x}' \in V$ such that $T(\alpha\mathbf{x}') = \alpha\mathbf{x}$.[2]

  Thus, $R(T)$ is a subspace of W, as required.

2. Consider the transformation $T : \mathbb{R}^4 \to \mathbb{R}^2$ given by

$$T\left([w, x, y, z]^t\right) = \begin{bmatrix} w+x+y \\ x+y+z \end{bmatrix}.$$

We are asked to show that this transformation is linear, and so we need to show that it satisfies the definition of a linear transformation[3] which means that we need to show that

$$T(\alpha\mathbf{x} + \alpha'\mathbf{x}') = \alpha T(\mathbf{x}) + \alpha' T(\mathbf{x}'),$$

for all vectors $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^4$ and all scalars $\alpha, \alpha' \in \mathbb{R}$. To do this, we take two general vectors $\mathbf{x} = [w, x, y, z]^t$ and $\mathbf{x}' = [w', x', y', z']^t$ in $\mathbb{R}^4$ and any two real numbers α and α' and note that

$$T(\alpha\mathbf{x} + \alpha'\mathbf{x}') = T\left(\alpha\begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} + \alpha'\begin{bmatrix} w' \\ x' \\ y' \\ z' \end{bmatrix}\right) = T\left(\begin{bmatrix} \alpha w + \alpha' w' \\ \alpha x + \alpha' x' \\ \alpha y + \alpha' y' \\ \alpha z + \alpha' z' \end{bmatrix}\right),$$

[1]Notice that $\mathbf{x}' + \mathbf{y}' \in V$ since $\mathbf{x}', \mathbf{y}' \in V$ and V is a vector space (i.e. it is closed under vector addition itself).
[2]Notice that $\alpha\mathbf{x}' \in V$ since $\mathbf{x}' \in V$ and V is a vector space (i.e. it is closed under scalar multiplication itself).
[3]That is, Definition 3.1.

which on applying the transformation yields

$$T(\alpha\mathbf{x}+\alpha'\mathbf{x}') = \begin{bmatrix}(\alpha w+\alpha' w')+(\alpha x+\alpha' x')+(\alpha y+\alpha' y')\\(\alpha x+\alpha' x')+(\alpha y+\alpha' y')+(\alpha z+\alpha' z')\end{bmatrix} = \alpha\begin{bmatrix}w+x+y\\x+y+z\end{bmatrix}+\alpha'\begin{bmatrix}w'+x'+y'\\x'+y'+z'\end{bmatrix}.$$

However, we know that

$$T\left(\begin{bmatrix}w\\x\\y\\z\end{bmatrix}\right) = \begin{bmatrix}w+x+y\\x+y+z\end{bmatrix} \text{ and } T\left(\begin{bmatrix}w'\\x'\\y'\\z'\end{bmatrix}\right) = \begin{bmatrix}w'+x'+y'\\x'+y'+z'\end{bmatrix},$$

and so, $T(\alpha\mathbf{x}+\alpha'\mathbf{x}') = \alpha T(\mathbf{x})+\alpha' T(\mathbf{x}')$, as required.

We are then asked to find the range and null-space of the transformation T, and to do this we note that:

- The range of T, i.e. the subspace of $\mathbb{R}^2$ given by the set

  $$R(T) = \{\mathbf{w}\in\mathbb{R}^2 \,|\, T(\mathbf{v}) = \mathbf{w} \text{ for some } \mathbf{v}\in\mathbb{R}^4\},$$

  is just the set of all vectors of the form

  $$T(\mathbf{v}) = \begin{bmatrix}w+x+y\\x+y+z\end{bmatrix},$$

  where $\mathbf{v} = [w,x,y,z]^t$ is any vector in $\mathbb{R}^4$. As such, a general vector in $R(T)$ can be written as

  $$T(\mathbf{v}) = w\begin{bmatrix}1\\0\end{bmatrix}+x\begin{bmatrix}1\\1\end{bmatrix}+y\begin{bmatrix}1\\1\end{bmatrix}+z\begin{bmatrix}0\\1\end{bmatrix},$$

  for any $w,x,y,z\in\mathbb{R}$. Thus, noting the linear dependence, we can see that

  $$R(T) = \text{Lin}\left\{\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}\right\} = \mathbb{R}^2,$$

  and this space is two-dimensional.

- The null space of T, i.e. the subspace of $\mathbb{R}^4$ given by the set

  $$N(T) = \{\mathbf{v}\in\mathbb{R}^4 \,|\, T(\mathbf{v}) = \mathbf{0}\},$$

  is just the set of all vectors of the form $[w,x,y,z]^t$ whose components satisfy the vector equation

  $$\begin{bmatrix}w+x+y\\x+y+z\end{bmatrix} = \mathbf{0}.$$

  So, equating components, we are looking for vectors in $\mathbb{R}^4$ whose components satisfy the simultaneous equations:

  $$\begin{aligned}w+x+y&=0\\x+y+z&=0\end{aligned}$$

  Now, as we have two equations in four variables, we take x and y (say) to be the free parameters r and s respectively. Thus, $w=-r-s$ and $z=-r-s$, which means that the null space contains vectors of the form

  $$\begin{bmatrix}w\\x\\y\\z\end{bmatrix} = r\begin{bmatrix}-1\\1\\0\\-1\end{bmatrix}+s\begin{bmatrix}-1\\0\\1\\-1\end{bmatrix},$$

and so, as these vectors are linearly independent, we have

$$N(T) = \text{Lin}\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ -1 \end{bmatrix} \right\},$$

and this subspace of $\mathbb{R}^4$ is two-dimensional.

We are also asked to verify the rank-nullity theorem which states that

$$\rho(T) + \eta(T) = n,$$

where $\rho(T)$ is the dimension of the range of T (i.e. 2), $\eta(T)$ is the dimension of the null space of T (i.e. 2) and n is the dimension of the domain of the transformation (i.e. 4). Comparing these numbers, it is easy to verify that the rank-nullity theorem holds in this case (as $2 + 2 = 4$).

Further, we are asked to find the matrix for T with respect to the standard bases in $\mathbb{R}^4$ and $\mathbb{R}^2$. To do this, we note that the standard basis of $\mathbb{R}^4$ is given by the set of vectors

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\},$$

and so calculating the image of each of these vectors under the transformation T we find that

$$T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ T\left(\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \ T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \text{ and } T\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

where the vectors on the left-hand-sides of these expressions are automatically of the required form as we are dealing with the standard basis of $\mathbb{R}^2$ too. Thus, in this case,

$$\mathsf{A}_T = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}.$$

And, we are also asked to find the the matrix for T with respect to the bases

$$S = \left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 2 \\ 1 \end{bmatrix} \right\} \quad \text{and} \quad S' = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\},$$

of $\mathbb{R}^4$ and $\mathbb{R}^2$ respectively. So, as before, we start by calculating the image of each of these vectors under the transformation T, i.e.

$$T\left(\begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 3 \\ 3 \end{bmatrix}, \ T\left(\begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \ T\left(\begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \text{ and } T\left(\begin{bmatrix} 0 \\ 0 \\ 2 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 3 \end{bmatrix},$$

where the vectors on the left-hand-sides of these expressions are written in terms of the standard basis of $\mathbb{R}^2$. However, to find A_T in this case, we need them to be written in terms of the basis $S' = \{\mathbf{w}_1, \mathbf{w}_2\}$ where $\mathbf{w}_1 = [1, 1]^t$ and $\mathbf{w}_2 = [1, 2]^t$, i.e. we need to find $[T(\mathbf{v})]_{S'}$ for these vectors. To do this, we denote the vectors in S by $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ and $\mathbf{v}_4$, and note that the first expression gives

$$T(\mathbf{v}_1) = \begin{bmatrix} 3 \\ 3 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 0\begin{bmatrix} 1 \\ 2 \end{bmatrix} = 3\mathbf{w}_1 + 0\mathbf{w}_2 \implies [T(\mathbf{v}_1)]_{S'} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}_{S'},$$

the second expression gives

$$T(\mathbf{v}_2) = \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 0 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 0\mathbf{w}_1 + 1\mathbf{w}_2 \implies [T(\mathbf{v}_2)]_{S'} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}_{S'},$$

the third expression gives

$$T(\mathbf{v}_3) = \begin{bmatrix} 2 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 0 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 2\mathbf{w}_1 + 0\mathbf{w}_2 \implies [T(\mathbf{v}_3)]_{S'} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}_{S'},$$

and the fourth expression gives

$$T(\mathbf{v}_4) = \begin{bmatrix} 2 \\ 3 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 1\mathbf{w}_1 + 1\mathbf{w}_2 \implies [T(\mathbf{v}_4)]_{S'} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}_{S'}.$$

Consequently,

$$\mathsf{A}_T = \begin{bmatrix} 3 & 0 & 2 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix},$$

is the required matrix for the bases S and S' given above.

Lastly, we are asked to verify that the sets S and S' are bases of $\mathbb{R}^4$ and $\mathbb{R}^2$ respectively. To do this, we note that (by Theorem 2.17) these sets will be bases if they are linearly independent. So, for S, we use the determinant test for linear independence, i.e. we note that

$$\underbrace{\begin{vmatrix} 1 & 0 & 1 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 \end{vmatrix}}_{\text{Use Row 1.}} = \underbrace{\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 1 & 1 & 1 \end{vmatrix}}_{\text{Use Row 1.}} + \underbrace{\begin{vmatrix} 2 & 1 & 0 \\ 0 & 0 & 2 \\ 1 & 1 & 1 \end{vmatrix}}_{\text{Use Row 2.}} = (1-2) - 2(2-1) = -3,$$

which is non-zero. Thus, S is a linearly independent set and hence a basis. Whilst, for S', we obviously have a linearly independent set as neither vector is a scalar multiple of the other, and so S' is a basis too.

3. We are asked to use the Wronskian to show that the set of functions $\{e^{\alpha x}, e^{\beta x}, e^{\gamma x}\}$ [defined for all values of $x \in \mathbb{R}$] is linearly independent if α, β and γ are all different. To do this, we note that the Wronskian of a set of three functions, say $\{f_1, f_2, f_3\}$, is given by

$$W(x) = \begin{vmatrix} f_1 & f_2 & f_3 \\ f_1' & f_2' & f_3' \\ f_1'' & f_2'' & f_3'' \end{vmatrix},$$

and we know that these functions are linearly independent if there is some x in their domain such that $W(x) \neq 0$. So, in the case under consideration, the Wronskian is given by

$$W(x) = \begin{vmatrix} e^{\alpha x} & e^{\beta x} & e^{\gamma x} \\ \alpha e^{\alpha x} & \beta e^{\beta x} & \gamma e^{\gamma x} \\ \alpha^2 e^{\alpha x} & \beta^2 e^{\beta x} & \gamma^2 e^{\gamma x} \end{vmatrix} = e^{(\alpha+\beta+\gamma)x} \begin{vmatrix} 1 & 1 & 1 \\ \alpha & \beta & \gamma \\ \alpha^2 & \beta^2 & \gamma^2 \end{vmatrix},$$

where we have taken the common factor out of each column in the last step. Now, to show that this set of functions is linearly independent if α, β and γ are all different, it is convenient to simplify the determinant by using some column operations, i.e.

$$\begin{vmatrix} 1 & 1 & 1 \\ \alpha & \beta & \gamma \\ \alpha^2 & \beta^2 & \gamma^2 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ \alpha & \beta-\alpha & \gamma-\alpha \\ \alpha^2 & \beta^2-\alpha^2 & \gamma^2-\alpha^2 \end{vmatrix} \qquad :C_2 \to C_2 - C_1 \text{ and } C_3 \to C_3 - C_1.$$

$$= (\beta-\alpha)(\gamma-\alpha) \begin{vmatrix} 1 & 0 & 0 \\ \alpha & 1 & 1 \\ \alpha^2 & \beta+\alpha & \gamma+\alpha \end{vmatrix} \qquad :\text{Common factors from } C_2 \text{ and } C_3.$$

$$\therefore \begin{vmatrix} 1 & 1 & 1 \\ \alpha & \beta & \gamma \\ \alpha^2 & \beta^2 & \gamma^2 \end{vmatrix} = (\beta-\alpha)(\gamma-\alpha) \begin{vmatrix} 1 & 0 & 0 \\ \alpha & 1 & 0 \\ \alpha^2 & \beta+\alpha & \gamma-\beta \end{vmatrix} \qquad :C_3 \to C_3 - C_2.$$

So, evaluating the determinant we find that the Wronskian is given by:

$$W(x) = (\beta-\alpha)(\gamma-\alpha)(\gamma-\beta)e^{(\alpha+\beta+\gamma)x},$$

which is non-zero if α, β and γ are all different.[4] Thus, the set of functions $\{e^{\alpha x}, e^{\beta x}, e^{\gamma x}\}$ [defined for all values of $x \in \mathbb{R}$] is linearly independent in this case.

4. To show that the Wronskian for the functions $f(x)$ and $g(x)$ defined [for all values of $x \in \mathbb{R}$] by

$$f(x) = x^2 \quad \text{and} \quad g(x) = \begin{cases} x^2 & \text{for } x \geq 0 \\ 0 & \text{for } x \leq 0 \end{cases}$$

is identically zero, we note that the Wronskian of these functions is given by

$$W(x) = \begin{vmatrix} f & g \\ f' & g' \end{vmatrix},$$

and so,

- If $x \geq 0$, then
$$W(x) = \begin{vmatrix} x^2 & x^2 \\ 2x & 2x \end{vmatrix} = 0,$$
as the two columns are equal.
- If $x \leq 0$, then
$$W(x) = \begin{vmatrix} x^2 & 0 \\ 2x & 0 \end{vmatrix} = 0,$$
as the there is a column of zeroes.

But, we also note that this pair of functions is linearly independent as there is no [single] scalar $\alpha \in \mathbb{R}$ such that $g(x) = \alpha f(x)$ for *all* x in the domain.

Aside: This result is useful because, from the lectures, we have the theorem:

> Let the functions $f_1, f_2, \ldots, f_n$ be defined and $n-1$ times differentiable for all values of x in the interval $(a,b) \subseteq \mathbb{R}$. If $W(x) \neq 0$ for some $x \in (a,b)$, then the set of functions $\{f_1, f_2, \ldots, f_n\}$ is linearly independent.

However, this question gives us a counterexample to [the contrapositive of] the converse of this theorem, i.e.

[4] Note that the $e^{(\alpha+\beta+\gamma)x}$ factor in this expression is never zero. Indeed, we only need to show that the Wronskian is non-zero for *some* value of x in the domain of these functions. So, as the domain is specified to be all $x \in \mathbb{R}$, we could have set $x = 0$. Then, working with $W(0) = (\beta-\alpha)(\gamma-\alpha)(\gamma-\beta)$, the result follows.

Let the functions $f_1, f_2, \ldots, f_n$ be defined and $n-1$ times differentiable for all values of x in the interval $(a,b) \subseteq \mathbb{R}$. If $W(x) = 0$ for all $x \in (a,b)$, then the set of functions $\{f_1, f_2, \ldots, f_n\}$ is linearly dependent.

since we have two functions where the Wronskian is equal to zero for all $x \in \mathbb{R}$ and yet the functions are linearly independent.

5. Writing the set of simultaneous equations given in the question in the form $\mathsf{A}\mathbf{x} = \mathbf{b}$ we find that

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 0 & 1 & -1 \\ 0 & 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 6 \end{bmatrix},$$

To solve this we write this equation as an 'augmented matrix' and perform row operations on it until we reach the Row Reduced Echelon form, i.e.

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 4 \\ 2 & 0 & 1 & -1 & 2 \\ 0 & 2 & 1 & 3 & 6 \end{array}\right] \xrightarrow{R_2 \to R_2 - 2R_1} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 4 \\ 0 & -2 & -1 & -3 & -6 \\ 0 & 2 & 1 & 3 & 6 \end{array}\right] \xrightarrow[R_2 \to -R_2]{R_3 \to R_3 + R_2}$$

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 4 \\ 0 & 2 & 1 & 3 & 6 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow{R_1 \to 2R_1 + R_2} \left[\begin{array}{cccc|c} 2 & 0 & 1 & -1 & 2 \\ 0 & 2 & 1 & 3 & 6 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Then, performing the row operations $R_1 \to \frac{1}{2}R_1$ and $R_2 \to \frac{1}{2}R_2$, we find that the Row Reduced Echelon form is

$$\left[\begin{array}{cccc|c} 1 & 0 & \frac{1}{2} & -\frac{1}{2} & 1 \\ 0 & 1 & \frac{1}{2} & \frac{3}{2} & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right],$$

and using this, we know from MA100 that the general solution to our set of simultaneous equations is also the general solution of the set of simultaneous equations given by the Row Reduced Echelon form of the 'augmented matrix', i.e.

$$\begin{bmatrix} 1 & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & \frac{1}{2} & \frac{3}{2} \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix},$$

and this can be written as

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 - \frac{1}{2}\lambda_1 + \frac{1}{2}\lambda_2 \\ 3 - \frac{1}{2}\lambda_1 - \frac{3}{2}\lambda_2 \\ \lambda_1 \\ \lambda_2 \end{bmatrix},$$

where we have set $x_3 = \lambda_1$ and $x_4 = \lambda_2$. Thus, writing this vector out in full, we can see that the general solution is given by

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 0 \\ 0 \end{bmatrix} - \frac{\lambda_1}{2} \begin{bmatrix} 1 \\ 1 \\ -2 \\ 0 \end{bmatrix} + \frac{\lambda_2}{2} \begin{bmatrix} 1 \\ -3 \\ 0 \\ 2 \end{bmatrix},$$

which is in the form $\mathbf{x} = \mathbf{v} + \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2$ required by the question if we set $\alpha_1 = -\frac{1}{2}\lambda_1$, $\alpha_2 = \frac{1}{2}\lambda_2$,

$$\mathbf{v} = \begin{bmatrix} 1 \\ 3 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ -2 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{u}_2 = \begin{bmatrix} 1 \\ -3 \\ 0 \\ 2 \end{bmatrix}.$$

Moreover, we can see that the set of vectors $\{\mathbf{u}_1, \mathbf{u}_2\}$ is a basis for the null space since this space is the set of all solutions of the matrix equation $\mathsf{A}\mathbf{x} = \mathbf{0}$ and using the analysis above, we know that this is the same as the set of all solutions of the matrix equation

$$\begin{bmatrix} 1 & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & \frac{1}{2} & \frac{3}{2} \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

that is the set of all vectors of the form

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2}\lambda_1 + \frac{1}{2}\lambda_2 \\ -\frac{1}{2}\lambda_1 - \frac{3}{2}\lambda_2 \\ \lambda_1 \\ \lambda_2 \end{bmatrix},$$

where we have set $x_3 = \lambda_1$ and $x_4 = \lambda_2$. Thus, writing this vector out in full, we can see that the null space is given by all linear combinations of the form

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = -\frac{\lambda_1}{2} \begin{bmatrix} 1 \\ 1 \\ -2 \\ 0 \end{bmatrix} + \frac{\lambda_2}{2} \begin{bmatrix} 1 \\ -3 \\ 0 \\ 2 \end{bmatrix},$$

i.e. the set of vectors $\{\mathbf{u}_1, \mathbf{u}_2\}$ spans the null space of A. Also, since neither of these vectors is a scalar multiple of the other, they are linearly independent and so, by Definition 2.10, this set of vectors is a basis of the null space of A as required.

The set of all vectors given by the general solution of our set of simultaneous equations is an affine set, and following the lectures, the dimension of this affine set is two since it is obtained by shifting the null space of A (which is two-dimensional) through the fixed vector $\mathbf{v}$. In turn, the Cartesian equation of this affine set is given by the intersection of the two [three-dimensional] hyperplanes whose Cartesian equations are

$$2x_1 + x_3 - x_4 = 2 \quad \text{and} \quad 2x_2 + x_3 + 3x_4 = 6,$$

which can be obtained from the first and second components of the general solution by using the fact that $\lambda_1 = x_3$ and $\lambda_2 = x_4$.

Other Problems.

The Other Problems on this sheet were intended to help you revise some properties of matrices which you should have encountered in MA100. Everyone *should* have tried these as we will be using these techniques throughout the course.

You should recall that the *row* (or *column*) *space* of a matrix is defined to be the space spanned by the row (or column) vectors of a matrix. We shall denote the row and column spaces of a matrix A by $RS(\mathsf{A})$ and $CS(\mathsf{A})$ respectively.

6. We are asked to use row operations to find the rank of the matrix

$$\mathsf{A} = \begin{bmatrix} 1 & -1 & 2 & -1 \\ 2 & 1 & -2 & -2 \\ -1 & 2 & -4 & 1 \\ 3 & 0 & 0 & -3 \end{bmatrix}.$$

This should be revision from MA100, and so performing the appropriate row operations we find that:

$$\begin{bmatrix} 1 & -1 & 2 & -1 \\ 2 & 1 & -2 & -2 \\ -1 & 2 & -4 & 1 \\ 3 & 0 & 0 & -3 \end{bmatrix} \xrightarrow[\substack{R_3 \to R_3 + R_1 \\ R_4 \to R_4 - 3R_1}]{R_2 \to R_2 - 2R_1} \begin{bmatrix} 1 & -1 & 2 & -1 \\ 0 & 3 & -6 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 3 & -6 & 0 \end{bmatrix} \xrightarrow{R_2 \leftrightarrow R_3}$$
$$\begin{bmatrix} 1 & -1 & 2 & -1 \\ 0 & 1 & -2 & 0 \\ 0 & 3 & -6 & 0 \\ 0 & 3 & -6 & 0 \end{bmatrix} \xrightarrow[\substack{R_3 \to R_3 - 3R_2 \\ R_4 \to R_4 - 3R_2}]{R_1 \to R_1 + R_2} \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

is the Row Reduced Echelon form of the matrix A. So, since this has two leading ones, the rank of A is two.

Once we have the Row Reduced Echelon form of A, it is easy to write down a smallest spanning set for $RS(\mathsf{A})$ and $CS(\mathsf{A})$. To do this, we note that the *row space* of the matrix A, denoted by $RS(\mathsf{A})$, is the space spanned by the row vectors of this matrix. So, a smallest spanning set for $RS(\mathsf{A})$ would be

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -2 \\ 0 \end{bmatrix} \right\},$$

as the row operations have removed any linear dependence from the row vectors that we started with.[5] Similarly, we note that the *column space* of the matrix A, denoted by $CS(\mathsf{A})$, is the space spanned by the column vectors of this matrix. So, as the leading ones in the Row Reduced Echelon form of A indicate the linear independence of the *corresponding* column vectors in the original form of the matrix, we can see that the first two column vectors of the matrix — as given in the question — are linearly independent. Thus, a smallest spanning set for $CS(\mathsf{A})$ would be

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ -1 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 2 \\ 0 \end{bmatrix} \right\}.$$

However, the set of vectors which can be found from the first two columns of the Row Reduced Echelon form of A, namely

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\},$$

is **NOT** a smallest spanning set for $CS(\mathsf{A})$ since although *row* operations preserve information about the *dimension* of the column space, they destroy information about the specific space that is spanned by the column vectors. (To see that this is so, notice that these two sets clearly span *different* spaces, although these spaces do have the *same* dimension.)[6]

7. We are given the linear transformation $T : \mathbb{R}^4 \to \mathbb{R}^3$ where

$$T\left([w, x, y, z]^t\right) = \begin{bmatrix} w + x \\ x + y \\ z + w \end{bmatrix},$$

and we are asked to find the matrix which represents this transformation with respect to the standard basis of $\mathbb{R}^4$. To do this, we calculate the image of each of the vectors in the standard basis of $\mathbb{R}^4$

[5] Also, notice that the dimension of the row space is two, as you should expect. (See also Question 9.)

[6] Also, notice that the dimension of the column space is two, as you should expect. (See also Questions 8 and 9.)

under the transformation T (see Question 2) and we find that

$$T\left(\begin{bmatrix}1\\0\\0\\0\end{bmatrix}\right)=\begin{bmatrix}1\\0\\1\end{bmatrix},\ T\left(\begin{bmatrix}0\\1\\0\\0\end{bmatrix}\right)=\begin{bmatrix}1\\1\\0\end{bmatrix},\ T\left(\begin{bmatrix}0\\0\\1\\0\end{bmatrix}\right)=\begin{bmatrix}0\\1\\0\end{bmatrix},\text{ and } T\left(\begin{bmatrix}0\\0\\0\\1\end{bmatrix}\right)=\begin{bmatrix}0\\0\\1\end{bmatrix},$$

where the vectors on the left-hand-sides of these expressions are automatically of the required form because, as no other basis was specified, we use the standard basis of $\mathbb{R}^3$ too. Thus, in this case,

$$\mathsf{A}_T=\begin{bmatrix}1&1&0&0\\0&1&1&0\\1&0&0&1\end{bmatrix},$$

is the required matrix.

We are also asked to find the range and null-space of A_T and then use these to verify that the rank-nullity theorem holds for matrices. To do this we note that:

- The range of A_T, i.e. the subspace of $\mathbb{R}^3$ given by the set

$$R(\mathsf{A}_T)=\{\mathbf{w}\in\mathbb{R}^3\,|\,\mathsf{A}_T\mathbf{v}=\mathbf{w}\text{ for some }\mathbf{v}\in\mathbb{R}^4\},$$

  is just the set of all vectors of the form

$$\mathbf{w}=\begin{bmatrix}1&1&0&0\\0&1&1&0\\1&0&0&1\end{bmatrix}\begin{bmatrix}w\\x\\y\\z\end{bmatrix},$$

  where $\mathbf{v}=[w,x,y,z]^t$ is any vector in $\mathbb{R}^4$. Multiplying this matrix product out we (unsurprisingly) get

$$\mathbf{w}=\begin{bmatrix}w+x\\x+y\\w+z\end{bmatrix},$$

  and as such, a general vector in $R(\mathsf{A}_T)$ can be written as

$$\mathbf{w}=w\begin{bmatrix}1\\0\\1\end{bmatrix}+x\begin{bmatrix}1\\1\\0\end{bmatrix}+y\begin{bmatrix}0\\1\\0\end{bmatrix}+z\begin{bmatrix}0\\0\\1\end{bmatrix},$$

  for any $w,x,y,z\in\mathbb{R}$. Thus, clearly,

$$R(\mathsf{A}_T)=\text{Lin}\left\{\begin{bmatrix}1\\0\\1\end{bmatrix},\begin{bmatrix}1\\1\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\end{bmatrix}\right\},$$

  (notice that this is just the column space of A_T) and so removing any linearly dependent vectors, we can see that

$$R(\mathsf{A}_T)=\text{Lin}\left\{\begin{bmatrix}1\\0\\1\end{bmatrix},\begin{bmatrix}0\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\end{bmatrix}\right\}=\mathbb{R}^3,$$

  and this space is three-dimensional.

- The null space of A_T, i.e. the subspace of $\mathbb{R}^4$ given by the set

$$N(\mathsf{A}_T)=\{\mathbf{v}\in\mathbb{R}^4\,|\,\mathsf{A}_T\mathbf{v}=\mathbf{0}\},$$

is just the set of all vectors of the form $[w, x, y, z]^t$ whose components satisfy the vector equation

$$\begin{bmatrix} w+x \\ x+y \\ w+z \end{bmatrix} = \mathbf{0}.$$

So, equating components, we are looking for vectors in $\mathbb{R}^4$ whose components satisfy the simultaneous equations:

$$\begin{aligned} w+x &= 0 \\ x+y &= 0 \\ w+z &= 0 \end{aligned}$$

Now, as we have three equations in four variables, we take w (say) to be the free parameter r. Thus, $x = -r$, $y = -x = r$ and $z = -r$, which means that the null space contains vectors of the form

$$\begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = r \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix},$$

and so we have

$$N(\mathsf{A}_T) = \text{Lin}\left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} \right\},$$

and this subspace of $\mathbb{R}^4$ is one-dimensional.

Further, it is easy to verify that the rank-nullity theorem holds for matrices, i.e. that

$$\rho(\mathsf{A}_T) + \eta(\mathsf{A}_T) = n,$$

where $\rho(\mathsf{A}_T)$ is the dimension of the range of A_T (i.e. 3), $\eta(\mathsf{A}_T)$ is the dimension of the null space of A_T (i.e. 1) and n is the dimension of the domain of the transformation (i.e. 4), since $3 + 1 = 4$.

The set of all vectors $\mathbf{b}$ for which the set of simultaneous equations given by $\mathsf{A}_T\mathbf{x} = \mathbf{b}$ is consistent is just the set of all vectors $\mathbf{b}$ for which there is an $\mathbf{x} \in \mathbb{R}^4$ that satisfies the matrix equation $\mathsf{A}_T\mathbf{x} = \mathbf{b}$. That is, it is the set of all vectors $\mathbf{b} \in \mathbb{R}^3$ such that $\mathsf{A}_T\mathbf{x} = \mathbf{b}$ for some $\mathbf{x} \in \mathbb{R}^4$, which is of course, the range of A_T. Thus, for $\mathsf{A}_T\mathbf{x} = \mathbf{b}$ to be consistent, we require that

$$\mathbf{b} \in R(\mathsf{A}_T) = \text{Lin}\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} = \mathbb{R}^3,$$

and so the required set of vectors is $\mathbb{R}^3$.[7]

8. We are asked to prove that for any real $m \times n$ matrix A, $R(\mathsf{A}) = CS(\mathsf{A})$. To do this we take the matrix in question to be

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

and note that the range of A, denoted by $R(\mathsf{A})$, is defined as

$$R(\mathsf{A}) = \{\mathbf{y} \in \mathbb{R}^m \mid \mathsf{A}\mathbf{x} = \mathbf{y} \text{ for some } \mathbf{x} \in \mathbb{R}^n\},$$

[7] The same conclusion can be reached using the result proved in Question 8.

in this case. Thus, for *any* vector $\mathbf{y} \in R(\mathsf{A})$, there exists a vector $\mathbf{x} = [x_1, x_2, \ldots, x_n]^t \in \mathbb{R}^n$ such that $\mathbf{y} = \mathsf{A}\mathbf{x}$, i.e.

$$\begin{aligned}
R(\mathsf{A}) &= \left\{ \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^m \;\middle|\; \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n \right\} \\
&= \left\{ \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix} \in \mathbb{R}^m \;\middle|\; x_1, x_2, \ldots, x_n \in \mathbb{R} \right\} \\
&= \left\{ x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \in \mathbb{R}^m \;\middle|\; x_1, x_2, \ldots, x_n \in \mathbb{R} \right\} \\
\therefore \quad R(\mathsf{A}) &= \text{Lin}\left\{ \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \ldots, \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \right\}
\end{aligned}$$

and so, $R(\mathsf{A})$ is just the space spanned by the column vectors of A. But, this is the column space of A, and so we can see that

$$CS(\mathsf{A}) = \text{Lin}\left\{ \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \ldots, \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \right\} = R(\mathsf{A}),$$

as required.

9. We are asked to establish that, for any real $m \times n$ matrix, the dimension of the column space, the dimension of the row space, and the rank of the matrix are all equal. Firstly, we note that, from Question 8, $R(\mathsf{A}) = CS(\mathsf{A})$ and so

$$\dim(CS(\mathsf{A})) = \dim(R(\mathsf{A})) = \rho(\mathsf{A}),$$

(see Problem Sheet 1, Question 12). Secondly, to *establish* that the dimension of the row space is equal to the dimension of the column space, we make the observation[8] that

> When we find the Row Reduced Echelon form of the matrix A, we locate the leading ones. Once we have found them, we note that:
>
> - The rows which contain them yield linearly independent vectors and we use these row vectors as a basis for the row space. (Thus, the dimension of the row space is equal to the number of leading ones.)
> - The columns which contain them also yield linearly independent vectors and we use the *corresponding* column vectors from the *original* matrix as a basis for the column space. (Thus, the dimension of the column space is equal to the number of leading ones.)
>
> (This is one of the major advantages of having the Row Reduced Echelon form of a matrix.)

[8] The *proof* that this is the case is long and boring (even for me!), and so we appeal to this observation as we only have to 'establish' that the result holds. Incidentally, the proof proceeds in much the same way except it requires you to prove some theorems about how row operations affect matrices — you should have seen some of these in MA100.

and consequently,

$$\dim(CS(\mathsf{A})) = \dim(RS(\mathsf{A})).$$

Thus, we can see that

$$\rho(\mathsf{A}) = \dim(CS(\mathsf{A})) = \dim(RS(\mathsf{A})),$$

as required.

Harder Problems.

The Harder Problems on this sheet were intended to give you some further practice in proving results about vector spaces and linear transformations.

10. We are asked to prove the following two theorems. Firstly,

> **Theorem:** If $S \subseteq V$ is a [finite] linearly independent set of vectors that is not already a basis for [a finite dimensional vector space] V, then the set S can be augmented to form a new set S' which is a basis for V by adding appropriate vectors to S.
>
> **Proof:** Let $S \subseteq V$ be a finite linearly independent set of vectors that is not already a basis for a finite dimensional vector space V where, say, $\dim(V) = n$. Clearly, by Definition 2.10, this means that S fails to span V and so we perform the following construction (note that $j \geq 1$):
>
> - **Step** 0**:** Take a vector $\mathbf{v}_0 \in V$ such that $\mathbf{v}_0 \notin \text{Lin}(S)$ and construct the set $S_0 = \{\mathbf{v}_0\} \cup S$ which will also be linearly independent by Theorem 2.9. If S_0 spans V, then S_0 is a basis for V and we have finished.
> - **Step** j**:** Take a vector $\mathbf{v}_j \in V$ such that $\mathbf{v}_j \notin \text{Lin}(S_{j-1})$ and construct the set $S_j = \{\mathbf{v}_j\} \cup S_{j-1}$ which will also be linearly independent by Theorem 2.9. If S_j spans V, then S_j is a basis for V and we have finished.
>
> (Notice that the construction above must terminate since, by [the contrapositive of] Theorem 2.14, we can have no more than n linearly independent vectors in an n-dimensional space.)[9] Thus, S can be augmented to form a new set which is a basis for V by adding the appropriate vectors, as required.

and secondly,

> **Theorem:** If $S \subseteq V$ is a [finite] set of vectors which spans [a finite dimensional vector space] V but is not already a basis for V, then the set S can be reduced to form a new set S' which is a basis for V by removing appropriate vectors from S.
>
> **Proof:** Let $S \subseteq V$ be a finite set of vectors which spans, but is not already a basis for, a finite dimensional vector space V where, say, $\dim(V) = n$. Clearly, by Definition 2.10, this means that S fails to be linearly independent and so, some of the vectors in S must be expressible as linear combinations of the other vectors in S. We now perform the following construction (note that $j \geq 1$):
>
> - **Step** 0**:** Take any vector $\mathbf{v} \in S$ which is expressible as a linear combination of the other vectors in S and remove it from S to form the the set S_0[10] which will also span V.[11] If S_0 is linearly independent, then S_0 is a basis for V and we have finished.

[9] Indeed, if S contains m vectors, then we will need to perform $n - m$ steps in this construction — i.e. it will terminate after **step** $n - m - 1$.

[10] Technically, this is the set $S_0 = S - \{\mathbf{v}\}$, where '$-$' denotes 'set difference'.

[11] See the second theorem in Question 4 of Problem Sheet 1.

- **Step j:** Take any vector $\mathbf{v} \in V$ which is expressible as a linear combination of the other vectors in S_{j-1} and remove it from S_{j-1} to form the the set S_j[12] which will also span V.[13] If S_j is linearly independent, then S_j is a basis for V and we have finished.

(Notice that the construction above must terminate since, by [the contrapositive of] the third theorem in Question 4 of Problem Sheet 1, we need at least n vectors to span an n-dimensional space.)[14] Thus, S can be reduced to form a new set which is a basis for V by removing the appropriate vectors, as required.

11. We are asked to prove the rank-nullity theorem (i.e. Theorem 3.9 in the hand-out) in the cases where $\eta(T) = 0$ and $\eta(T) = n$. We shall consider each of these cases in turn:

- **Case 1:** Assume that $\eta(T) = 0$. We further assume (as before) that $\dim(V) = n$ and as such, we can take the set of vectors $S' = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ to be a basis for V. Now, to prove the rank-nullity theorem in this case we need to find $\rho(T)$, and to do this, we want to show that the set of vectors $S'' = \{T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_n)\} \subseteq W$ is a basis for $R(T)$. Of course, to show this, we need to show that this set both spans $R(T)$ and is linearly independent (recall Definition 2.10):

  - **To show that S'' spans $R(T)$:** For any vector $\mathbf{v} \in V$ we can write

    $$\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n,$$

    and since $T : V \to W$ is a linear transformation, by Theorem 3.2, we can apply it to both sides of this expression to get

    $$T(\mathbf{v}) = \alpha_1 T(\mathbf{v}_1) + \alpha_2 T(\mathbf{v}_2) + \cdots + \alpha_n T(\mathbf{v}_n). \tag{1}$$

    Thus, every vector in $R(T)$ (i.e. every vector in W of the form $T(\mathbf{v})$ for some $\mathbf{v} \in V$) can be written as a linear combination of the vectors in S'' and so this set spans $R(T)$ (as required).

  - **To show that S'' is linearly independent:** As $\eta(T) = 0$, by Definition 2.16, the null space of T is $\{\mathbf{0}\}$, and as such, $T(\mathbf{v}) = \mathbf{0}$ iff $\mathbf{v} = \mathbf{0}$.[15] Also, as the set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis, we know that

    $$\mathbf{v} = \mathbf{0} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n \text{ iff } \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0,$$

    as these vectors are linearly independent. Consequently, using these two facts, we have

    $$T(\mathbf{v}) = \mathbf{0} \text{ iff } \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0,$$

    in Equation (1), and so the vectors in S'' are linearly independent (as required).

  Thus, S'' is a basis of $R(T)$, and as this set contains n vectors, $\rho(T) = n$. So, in this case, the rank-nullity theorem holds since

  $$\eta(T) + \rho(T) = 0 + n = n = \dim(V),$$

  as we assumed that V was an n-dimensional space in the proof.

[12] Technically, this is the set $S_j = S_{j-1} - \{\mathbf{v}\}$.

[13] See the second theorem in Question 4 of Problem Sheet 1.

[14] Indeed, if S contains m vectors, then we will need to perform $m - n$ steps in this construction — i.e. it will terminate after **step** $m - n - 1$.

[15] Since $\mathbf{0}$ is the only vector in the null space of T.

- **Case 2:** Assume that $\eta(T) = n$. We further assume (as before) that $\dim(V) = n$ and as such, we can take the set of vectors $S' = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ to be a basis for V. Now, as the null space of T is a subspace of V (see Theorem 3.7) and $\dim(V) = \eta(T) = n$, we can see that $V = N(T)$ (see Problem Sheet 1, Question 12). Thus, S' will be a basis for $N(T)$ too, and so we have

  $$T(\mathbf{v}_1) = T(\mathbf{v}_2) = \cdots = T(\mathbf{v}_n) = \mathbf{0}. \tag{2}$$

  Now, to prove the rank-nullity theorem in this case we need to find $\rho(T)$, and to do this, we need to find the range of T. But, $R(T)$ consists of the set of all vectors in W that are the image of some vector $\mathbf{v} \in V$, so taking such a general vector in V, say

  $$\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n,$$

  we can see that, by Theorem 3.2, $R(T)$ is the set of all vectors of the form

  $$T(\mathbf{v}) = \alpha_1 T(\mathbf{v}_1) + \alpha_2 T(\mathbf{v}_2) + \cdots + \alpha_n T(\mathbf{v}_n).$$

  However, by Equation (2), this entails that $T(\mathbf{v}) = \mathbf{0}$ for all such $\mathbf{v} \in V$ and so we have $R(T) = \{\mathbf{0}\}$. Hence, by Definition 2.16, $\rho(T) = 0$. So, in this case, the rank-nullity theorem holds since

  $$\eta(T) + \rho(T) = n + 0 = n = \dim(V),$$

  as we assumed that V was an n-dimensional space in the proof.

Finally, we note that the proof given in the handout does not apply in these cases because:

- In **Case 1**, we assume that $\eta(T) = 0$ and as such the null space of T has no basis,[16] whereas the proof in the handout assumes that we do have a basis for the null space of T.
- In **Case 2**, we find that $\eta(T) = n$ and as such a basis for the null space of T is already a basis for V, whereas the proof in the handout assumes that we can construct a basis for V by adding [linearly independent] vectors to a basis for the null space.[17]

12. We are asked to prove Theorems 3.12 and 3.13 from the hand-out. There is really no need for me to do this here as the proofs are essentially the arguments given at the beginning of Sections 3.3.3 and 3.3.4 respectively. If you are interested, you can try and 'reconstruct' the proofs yourself.

[16] To see why this is the case, see the discussion that immediately follows Definition 2.16.

[17] And, in turn, these new vectors allow us to go on and generate a basis for $R(T)$, although in **Case 2**, the range of T has no basis since $R(T) = \{\mathbf{0}\}$.