6. LINEAR TRANSFORMATIONS

We have objects with some structure (vector spaces). This section is about structure-preserving maps between these objects. You will see a similar phenomenon in lots of other contexts too – whenever we have objects with some kind of structure, we can ask about structure-preserving maps between objects. (This can lead to further abstraction, which is explored in Category Theory, an interesting part of mathematics and a Part C course.)

6.1 What is a linear transformation?

Definition 144 *Let V, W be vector spaces over $\mathbb{F}$. We say that a map $T\colon V \to W$ is **linear** if*

(i) $T(v_1 + v_2) = T(v_1) + T(v_2)$ for all v_1, $v_2 \in V$ (preserves additive stucture);
and
(ii) $T(\lambda v) = \lambda T(v)$ for all $v \in V$ and $\lambda \in \mathbb{F}$ (preserves scalar multiplication).
*We call T a **linear transformation** or a **linear map**.*

Proposition 145 *Let V, W be vector spaces over $\mathbb{F}$, let $T\colon V \to W$.be a linear map. Then $T(0_V) = 0_W$.*

Proof. Note that $T(0_V) + T(0_V) = T(0_V + 0_V) = T(0_V)$, and hence $T(0_V) = 0_W$. ■

Proposition 146 *Let V, W be vector spaces over $\mathbb{F}$, let $T\colon V \to W$. The following are equivalent:*

(a) T is linear;

(b) $T(\alpha v_1 + \beta v_2) = \alpha T(v_1) + \beta T(v_2)$ for all v_1, $v_2 \in V$ and α, $\beta \in \mathbb{F}$;

(c) for any $n \geqslant 1$, if $v_1, \ldots, v_n \in V$ and $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ then

$$T(\alpha_1 v_1 + \cdots + \alpha_n v_n) = \alpha_1 T(v_1) + \cdots + \alpha_n T(v_n).$$

Proof. Exercise. ■

Example 147

- *Let V be a vector space. Then the **identity map** $id_V\colon V \to V$ given by $id_V(v) = v$ for all $v \in V$ is a linear map.*
- *Let V, W be vector spaces. The **zero map** $0\colon V \to W$ that sends every $v \in V$ to 0_W is a linear map.*

- *For* $m,\ n \geqslant 1$, *with* $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. *Then we define the left multiplication map*

$$L_A\colon \mathbb{R}^n_{\mathrm{col}} \to \mathbb{R}^m_{\mathrm{col}} \qquad \text{by} \qquad L_A(\mathbf{v}) = A\mathbf{v} \quad \text{for} \quad \mathbf{v} \in \mathbb{R}^n_{\mathrm{col}}.$$

This is a linear map. Similarly, we have a right multiplication map

$$R_A\colon \mathbb{R}^m \to \mathbb{R}^n \qquad \text{by} \qquad R_A(\mathbf{v}) = \mathbf{v}A \quad \text{for} \quad \mathbf{v} \in \mathbb{R}^m.$$

- *Take* $m,\ n,\ p \geqslant 1$ *with* $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. *The left multiplication map* $\mathcal{M}_{n\times p}(\mathbb{R}) \to \mathcal{M}_{m\times p}(\mathbb{R})$ *sending* X *to* AX *is a linear map.*

- *Let* V *be a vector space over* $\mathbb{F}$ *with subspaces* U, W *such that* $V = U \oplus W$. *For* $v \in V$ *there are unique* $u \in U$, $w \in W$ *such that* $v = u + w$. *Define* $P\colon V \to V$ *by* $P(v) = w$.

Proposition 148 *P is a linear map. P is called* ***projection of*** V ***onto*** W ***along*** U.

Proof. *Take* $v_1,\ v_2 \in V$ *and* $\alpha_1,\ \alpha_2 \in \mathbb{F}$. *Then there are* $u_1,\ u_2 \in U$, $w_1,\ w_2 \in W$ *such that* $v_1 = u_1 + w_1$ *and* $v_2 = u_2 + w_2$. *Now*

$$\begin{aligned}\alpha_1 v_1 + \alpha_2 v_2 &= \alpha_1(u_1 + w_1) + \alpha_2(u_2 + w_2)\\ &= (\alpha_1 u_1 + \alpha_2 u_2) + (\alpha_1 w_1 + \alpha_2 w_2)\end{aligned}$$

where $\alpha_1 u_1 + \alpha_2 u_2 \in U$ *and* $\alpha_1 w_1 + \alpha_2 w_2 \in W$. *So by uniqueness*

$$P(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 w_1 + \alpha_2 w_2 = \alpha_1 P(v_1) + \alpha_2 P(v_2).$$

■

- *For* $A = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{R})$, *we define the* ***trace*** *of* A *to be*

$$\mathrm{trace}(A)\colon = a_{11} + a_{22} + \cdots + a_{nn},$$

(the sum of the entries on the main diagonal of A*). The map* $\mathrm{trace}\colon \mathcal{M}_{n\times n}(\mathbb{R}) \to \mathbb{R}$ *is a linear map.*

- *Let* $\mathbb{R}_n[x]$ *be the vector space of polynomials of degree at most* n. *Define* $D\colon \mathbb{R}_n[x] \to \mathbb{R}_n[x]$ *by* $p(x) \mapsto p'(x)$, *that is,*

$$D(a_n x^n + \cdots + a_1 x + a_0) = n a_n x^{n-1} + \cdots + a_1.$$

This is a linear map from $\mathbb{R}_n[x]$ *to* $\mathbb{R}_n[x]$. *We could also think of it as a linear map* $\mathbb{R}_n[x]$ *to* $\mathbb{R}_{n-1}[x]$.

- *Let* $C^1(\mathbb{R})$ *be the subspace of* $\mathbb{R}^{\mathbb{R}}$ *consisting of differentiable functions* $f\colon \mathbb{R} \to \mathbb{R}$. *The differential operator* $D\colon C^1(\mathbb{R}) \to \mathbb{R}^{\mathbb{R}}$ *sending* f *to* f' *is a linear map.*

- *Let* $C^\infty(\mathbb{R})$ *be the subspace of* $\mathbb{R}^{\mathbb{R}}$ *consisting of functions* $f\colon \mathbb{R} \to \mathbb{R}$ *that are infinitely differentiable. The differential operator* $D\colon C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ *sending* f *to* f' *is a linear map.*

- *Let* X *be a set, let* $V = \mathbb{R}^X$ *be the space of real valued functions on* X. *For* $a \in X$, *the* ***evaluation map*** $E_a\colon V \to \mathbb{R}$ *sending* f *to* $f(a)$ *is a linear map.*

6.2 Combining linear transformations

We can add linear transformations (pointwise), and we can multiply a linear transformation by a scalar (pointwise).

Proposition 149 *Let V, W be vector spaces over a field $\mathbb{F}$. For S, $T\colon V \to W$ and $\lambda \in \mathbb{F}$, we may define linear maps $S+T$ and λS by*

$$\begin{array}{rclllll} S+T\colon V & \to & W & \textit{by} & (S+T)(v) = S(v)+T(v) & \textit{for} & v \in V; \\ \lambda S\colon V & \to & W & \textit{by} & (\lambda S)(v) = \lambda S(v) & \textit{for} & v \in V. \end{array}$$

With these operations (and the zero map $0\colon V \to W$), the set of linear transformations $V \to W$ forms a vector space denoted $\mathrm{Hom}(V,W)$.

Proof. Firstly $S+T$ is a linear map as

$$\begin{aligned} (S+T)(\alpha_1 v_1 + \alpha_2 v_2) &= S(\alpha_1 v_1 + \alpha_2 v_2) + T(\alpha_1 v_1 + \alpha_2 v_2) \qquad \text{[by definition]} \\ &= \alpha_1 S(v_1) + \alpha_2 S(v_2) + \alpha_1 T(v_1) + \alpha_2 T(v_2) \qquad \text{[by linearity]} \\ &= \alpha_1 (S(v_1) + T(v_1)) + \alpha_2 (S(v_2) + T(v_2)) \qquad \text{[rearranging]} \\ &= \alpha_1 (S+T)(v_1) + \alpha_2 (S+T)(v_2) \qquad \text{[by definition]} \end{aligned}$$

showing linearity. That λS is linear is left as an exercise. Verifying the vector space axioms for $\mathrm{Hom}(V,W)$ involves showing:

(i) $S+T = T+S$: this follows from commutativity of $+$ in W.
(ii) $S+(T+U) = (S+T)+U$: this follows from associativity of $+$ in W.
(iii) $S+0 = S$: this follows from properties of 0_W.
(iv) S has an additive inverse $(-S)(v) \stackrel{\text{def}}{=} -(S(v))$.
(v)

$$\lambda(S+T) = (\lambda S) + (\lambda T); \quad (\lambda+\mu)S = (\lambda S) + (\mu S); \quad (\lambda\mu)S = \lambda(\mu S); \quad 1S = S.$$

These properties follow from the same properties for vectors in W as addition and scalar multiplication of linear maps is defined pointwise. ■

We can also compose linear transformations.

Proposition 150 *Let U, V, W be vector spaces over $\mathbb{F}$. Let $S\colon U \to V$ and $T\colon V \to W$ be linear. Then $T \circ S\colon U \to W$ is linear.*

Proof. Take u_1, $u_2 \in U$ and λ_1, $\lambda_2 \in \mathbb{F}$. Then

$$\begin{aligned} (T \circ S)(\lambda_1 u_1 + \lambda_2 u_2) &= T(S(\lambda_1 u_1 + \lambda_2 u_2)) \qquad \text{[definition of composition]} \\ &= T(\lambda_1 S(u_1) + \lambda_2 S(u_2)) \qquad [S \text{ is linear}] \\ &= \lambda_1 T(S(u_1)) + \lambda_2 T(S(u_2)) \qquad [T \text{ is linear}] \\ &= \lambda_1 (T \circ S)(u_1) + \lambda_2 (T \circ S)(u_2) \qquad \text{[definition of composition]} \end{aligned}$$

so $T \circ S$ is linear. ■

Remark 151 *We often write $T \circ S$ as TS. The notation $T \circ S$ removes any possible ambiguity about the order of the functions.*

Definition 152 *Let V, W be vector spaces and let $T\colon V \to W$ be linear. We say that T is **invertible** if there is a linear transformation $S\colon W \to V$ such that $ST = id_V$ and $TS = id_W$ (where id_V and id_W are the identity maps on V and W respectively). In this case, we call S the **inverse** of T, and write it as T^{-1}. An invertible linear map is called an **isomorphism.***

Remark 153 *T is a function, so if it is invertible then it has a unique inverse (you saw this in the Introduction to University Maths course), so there is no ambiguity in writing T^{-1}.*

Proposition 154 *Let V, W be vector spaces. Let $T\colon V \to W$ be linear. Then T is invertible if and only if T is bijective.*

Proof. ($\Rightarrow$) If T is invertible, then it is certainly bijective (see the Introduction to University Maths course).
($\Leftarrow$) Assume that T is bijective.

Then T has an inverse function $S\colon W \to V$, but it remains to show that S is linear. Take w_1, $w_2 \in W$ and λ_1, $\lambda_2 \in \mathbb{F}$. Let $v_1 = S(w_1)$, $v_2 = S(w_2)$. Then

$$T(v_1) = TS(w_1) = w_1 \qquad \text{and} \qquad T(v_2) = TS(w_2) = w_2.$$

Now

$$\begin{aligned} S(\lambda_1 w_1 + \lambda_2 w_2) &= S(\lambda_1 T(v_1) + \lambda_2 T(v_2)) \\ &= S(T(\lambda_1 v_1 + \lambda_2 v_2)) \qquad [\text{since } T \text{ is linear}] \\ &= \lambda_1 v_1 + \lambda_2 v \qquad [\text{as } S \text{ is inverse to } T] \\ &= \lambda_1 S(w_1) + \lambda_2 S(w_2). \end{aligned}$$

So S is linear. ∎

Proposition 155 *Let U, V, W be vector spaces. Let $S\colon U \to V$ and $T\colon V \to W$ be invertible linear transformations. Then $TS\colon U \to W$ is invertible, and $(TS)^{-1} = S^{-1}T^{-1}$.*

Proof. Exercise. ∎

Proposition 156 *(a) Let V, W be vector spaces with V finite-dimensional. If there is an invertible linear map $T\colon V \to W$ then $\dim V = \dim W$.*

(b) Let V, W be finite-dimensional vector spaces with $\dim V = \dim W$. Then there is an invertible linear map $T\colon V \to W$.

Consequently V and W are isomorphic if and only if $\dim V = \dim W$.

Proof. (a) Let $v_1, \ldots, v_n$ be a basis for V. It is left as an exercise to show that $Tv_1, \ldots, Tv_n$ is a basis for W. Then $n = \dim V = \dim W$.

(b) Let $n = \dim V = \dim W$. Let $v_1, \ldots, v_n$ be a basis for V and $w_1, \ldots, w_n$ be a basis for W. It is left as an exercise to show that

$$T\colon V \to W \qquad \text{given by} \qquad T\left(\sum \alpha_i v_i\right) = \sum \alpha_i w_i$$

is a well-defined, invertible linear map. ∎

Example 157 *Let $V = \mathbb{R}[x]$ denote the vector space of polynomials in x with real coefficients, and let W denote the vector space of real sequences $(a_n)_{n=0}^{\infty}$. Then V and W are both infinite-dimensional but are not isomorphic.*

Solution. The set $B = \{1, x, x^2, x^3, \ldots\}$ is a basis for V. That it is linearly independent shows that V is not finite-dimensional. The set of sequences $\{(\delta_{in})_{n=0}^{\infty} \mid i \geqslant 0\}$ is linearly independent and so W is also infinite-dimensional.

However the set $S = \{(t^n)_{n=0}^{\infty} \mid t \in \mathbb{R}\}$ is an uncountable linearly independent subset of W and hence W does not have a countable basis. We prove that S is linearly independent below. Suppose that

$$\alpha_1 (t_1^n) + \cdots + \alpha_k (t_k^n) = (0)$$

for real numbrs $\alpha_1, \ldots, \alpha_k, t_1, \ldots, t_k$ with the t_i distinct. Then for all $n \geqslant 0$ we have

$$\alpha_1 t_1^n + \cdots + \alpha_k t_k^n = 0.$$

These equations for $0 \leqslant n < k$ can be rewritten as the single matrix equation

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ t_1 & t_2 & t_3 & \cdots & t_k \\ t_1^2 & t_2^2 & t_3^2 & \cdots & t_k^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t_1^{k-1} & t_2^{k-1} & t_3^{k-1} & \cdots & t_k^k \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \vdots \\ \alpha_k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

As the t_i are distinct, then the above $k \times k$ matrix is invertible – this is proved in Linear Algebra II next term. Hence S is an uncountable linearly independent set. No such set exists in V and hence W is not isomorphic to V. ■

6.3 Rank and nullity

Definition 158 *Let V, W be vector spaces. Let $T : V \to W$ be linear. We define the **kernel** (or **null space**) of T to be*

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

*We define the **image** of T to be*

$$\operatorname{Im} T := \{T(v) \mid v \in V\}.$$

Here are some useful properties of kernels and images.

Proposition 159 *Let V, W be vector spaces over $\mathbb{F}$. Let $T \colon V \to W$ be linear. Then*

(a) $\ker T$ is a subspace of V and $\operatorname{Im} T$ is a subspace of W;

(b) T is injective if and only if $\ker T = \{0_V\}$.

(c) if A is a spanning set for V, then $T(A)$ is a spanning set for $\operatorname{Im} T$;

(d) if V is finite-dimensional, then $\ker T$ and $\operatorname{Im} T$ are finite-dimensional.

Proof. (a) Note that $\ker T \subseteq V$ and $\operatorname{Im} T \subseteq W$.

$\underline{\ker T}$ Note that $T(0_V) = 0_W$ so $0_V \in \ker T$.

Take v_1, $v_2 \in \ker T$ and $\lambda \in \mathbb{F}$, so $T(v_1) = T(v_2) = 0_W$.

Then $T(\lambda v_1 + v_2) = \lambda T(v_1) + T(v_2) = \lambda 0_W + 0_W = 0_W$, so $\lambda v_1 + v_2 \in \ker T$.

So, by the Subspace Test, $\ker T \leqslant V$.

$\underline{\operatorname{Im} T}$ We have $T(0_V) = 0_W$ so $0_W \in \operatorname{Im} T$.

Take w_1, $w_2 \in \operatorname{Im} T$ and $\lambda \in \mathbb{F}$. Then there are v_1, $v_2 \in V$ such that $T(v_1) = w_1$ and $T(v_2) = w_2$. Then $\lambda w_1 + w_2 = \lambda T(v_1) + T(v_2) = T(\lambda v_1 + v_2) \in \operatorname{Im} T$.

So, by the Subspace Test, $\operatorname{Im} T \leqslant W$.

(b) Say that T is 1–1. If $v \in \ker T$ then $Tv = 0_W = T(0_V)$. By injectivity we have $v = 0_V$ and so $\ker T = \{0_V\}$.

Conversely say that $\ker T = \{0_V\}$ and that $Tv_1 = Tv_2$. Then $T(v_1 - v_2) = Tv_1 - Tv_2 = 0_W$. Then $v_1 - v_2 = 0_V$ and so T is 1–1.

(c) Let A be a spanning set for V.

Take $w \in \operatorname{Im} T$. Then $w = T(v)$ for some $v \in V$.

Now there are $v_1, \ldots, v_n \in A$ and $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ such that $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$. So

$$w = T(\alpha_1 v_1 + \cdots + \alpha_n v_n) = \alpha_1 T(v_1) + \cdots + \alpha_n T(v_n) \in \langle T(A) \rangle.$$

So $T(A)$ spans $\operatorname{Im} T$.

(d) Assume that V is finite-dimensional. Then $\ker T \leqslant V$ so $\ker T$ is finite-dimensional. Also, $\operatorname{Im} T$ is finite-dimensional by (iii). ■

Corollary 160 *Given a matrix A the image of L_A is $\operatorname{Col}(A)$, the column space of A and the image of R_A is $\operatorname{Row}(A)$, the row space of A.*

Proof. The canonical basis $\mathbf{e}_1, \ldots, \mathbf{e}_m$ spans $\mathbb{R}^m$ and hence the rows of A, $\mathbf{r}_1 = \mathbf{e}_1 A, \ldots, \mathbf{r}_m = \mathbf{e}_m A$ span $\operatorname{Im} R_A$. Hence $\operatorname{Im} R_A \subseteq \operatorname{Row}(A)$. Conversely if $\mathbf{v} \in \operatorname{Row}(A)$ then

$$\mathbf{v} = \alpha_1 \mathbf{r}_1 + \cdots + \alpha_m \mathbf{r}_m = (\alpha_1 \mathbf{e}_1 + \cdots + \alpha_m \mathbf{e}_m) A \in \operatorname{Im} R_A.$$

Likewise $\operatorname{Im} L_A = \operatorname{Col}(A)$. ■

Definition 161 *Let V, W be vector spaces with V finite-dimensional. Let $T : V \to W$ be linear. We define the* ***nullity*** *of T to be nullity$(T) := \dim(\ker T)$, and the* ***rank*** *of T to be rank$(T) := \dim(\operatorname{Im} T)$.*

Given the previous corollary, the rank of L_A equals the column rank of A and the rank of R_A equals the row rank of A.

The next theorem is very important!

Theorem 162 (Rank-Nullity Theorem) *Let V, W be vector spaces with V finite-dimensional. Let $T \colon V \to W$ be linear. Then*

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

Proof. Take a basis $v_1, \ldots, v_n$ for $\ker T$, where $n = \text{nullity}(T)$ and extend this to a basis $v_1, \ldots, v_n, v_1', \ldots, v_r'$ of V so that $\dim V = n + r$. We claim that $B = \{Tv_1', \ldots, Tv_r'\}$ is a basis for $\operatorname{Im} T$.

B spans $\operatorname{Im} T$:
By Proposition 159(c), $T(v_1), \ldots, T(v_n), T(v_1'), \ldots, T(v_r')$ span $\operatorname{Im} T$, being the images of a basis. But $v_1, \ldots, v_n \in \ker T$, meaning $T(v_1) = \cdots = T(v_n) = 0_W$, so in fact $Tv_1', \ldots, Tv_r'$ span $\operatorname{Im} T$.

B is linearly independent:
Take $\alpha_1, \ldots, \alpha_r \in \mathbb{F}$ such that

$$\alpha_1 T(v_1') + \cdots + \alpha_r T(v_r') = 0_W.$$

As T is linear, we can rewrite this as $T(\alpha_1 v_1' + \cdots + \alpha_r v_r') = 0_W$. So $\alpha_1 v_1' + \cdots + \alpha_r v_r' \in \ker T$. As $v_1, \ldots, v_n$ is a basis for $\ker T$, there are $\beta_1, \ldots, \beta_n \in \mathbb{F}$ such that

$$\alpha_1 v_1' + \cdots + \alpha_r v_r' = \beta_1 v_1 + \cdots + \beta_n v_n,$$

But $v_1, \ldots, v_n, v_1', \ldots, v_r'$ are linearly independent (being a basis for V), so

$$\beta_1 = \cdots = \beta_n = \alpha_1 = \cdots = \alpha_r = 0.$$

This shows $w_1, \ldots, w_r$ are linearly independent and the claim follows.

Now using the claim we have $\text{rank}T = r$, and so $\dim(V) = n + r = \text{nullity}(T) + \text{rank}(T)$. ∎

Here are a couple of useful results in their own right that also illustrate the usefulness of the Rank-Nullity Theorem.

Corollary 163 *Let V be a finite-dimensional vector space. Let $T \colon V \to V$ be linear. The following are equivalent:*
(a) T is invertible;
(b) rank$T = \dim V$;
(c) nullity$T = 0$.

Proof. (a) $\Rightarrow$ (b):
Assume that T is invertible. Then T is bijective, so is surjective, so $\operatorname{Im} T = V$, meaning $\text{rank}T = \dim V$.
(b) $\Rightarrow$ (c):
Assume that $\text{rank}T = \dim V$. Then, by Rank-Nullity, $\text{nullity}T = 0$.
(c) $\Rightarrow$ (a):
Assume that $\text{nullity}T = 0$ so that $\ker T = \{0_V\}$. Then T is injective.
Also, by Rank-Nullity, $\text{rank}T = \dim V$ and $\operatorname{Im} T \leqslant V$, so $\operatorname{Im} T = V$, so T is surjective.
So T is bijective, so T is invertible ∎

The next result is important, and we'll use it again later in the course.

Corollary 164 *Let V be a finite-dimensional vector space. Let $T \colon V \to V$ be linear. Then any one-sided inverse of T is a two-sided inverse, and so is unique.*

Proof. Suppose that T has a right inverse $S\colon V \to V$, so $T \circ S = id_V$. Since id_V is surjective, T is surjective, so $\text{rank}T = \dim V$.

So, by the previous corollary T is invertible, say with two-sided inverse S'.

Then $S' = S' \circ id_V = S' \circ (T \circ S) = (S' \circ T) \circ S = id_V \circ S = S$.

Hence S is the (unique) two-sided inverse.

If instead we suppose that T has a left inverse $S\colon V \to V$, so $S \circ T = id_V$, then id_V is injective so that T is injective and hence $\text{nullity}T = 0$, and the argument is similar to the previous one. ■

Which, when reworded in terms of matrices, is the following result.

Corollary 165 *Let A, B be square matrices of the same size. If AB is invertible then A and B are invertible.*

Lemma 166 *Let V and W be vector spaces, with V finite-dimensional. Let $T\colon V \to W$ be linear and $U \leqslant V$. Then*

$$\dim U - \text{nullity}T \leqslant \dim T(U) \leqslant \dim U.$$

In particular, if T is injective then $\dim T(U) = \dim U$.

Proof. Let $S\colon U \to W$ be the restriction of T to U (that is, $S(u) = T(u)$ for all $u \in U$). Then S is linear, and $\ker S \leqslant \ker T$ so $\text{nullity}S \leqslant \text{nullity}T$. Also, $\operatorname{Im} S = T(U)$. By Rank-Nullity,

$$\begin{aligned}
\dim T(U) &= \dim \operatorname{Im} S = \dim U - \text{nullity}S \leqslant \dim U; \\
\dim T(U) &= \dim U - \text{nullity}S \geqslant \dim U - \text{nullity}T.
\end{aligned}$$

If T is injective, then $\text{nullity}T = 0$, so $\dim T(U) = \dim U$. ■

Remark 167 ***A take on the Rank-Nullity Theorem using matrices.***

Let A be an $m \times n$ matrix and consider its reduced form $\mathrm{RRE}(A)$. We know that there are as many leading 1s as there are non-zero rows and this is the row rank of A. We also know that the kernel of A can be parameterized by assigning parameters to every column which does not have a leading 1. Hence

$$\begin{aligned}
\textit{row rank of } A &= \textit{number of leading } 1s; \\
\textit{nullity of } A &= \textit{number of columns without leading } 1s.
\end{aligned}$$

Hence

$$\begin{aligned}
n &= \dim \mathbb{R}^n_{\text{col}} \\
&= (\textit{number of leading } 1s) + (\textit{number of columns without leading } 1s) \\
&= (\textit{row rank of } A) + (\textit{nullity of } A)\,.
\end{aligned}$$

Corollary 168 ***(Criteria for Invertibility)*** *Let A be an $n \times n$ matrix. The following statements are equivalent:*

(a) A is invertible.

(b) A has a left inverse.
(c) A has a right inverse.
(d) $\mathrm{Row}(A) = \mathbb{R}^n$.
(e) The columns of A are linearly independent.
(f) The rows of A are linearly independent.
(g) The only solution $\mathbf{x}$ *in* $\mathbb{R}^n_{\text{col}}$ *to the system* $A\mathbf{x} = \mathbf{0}$ *is* $\mathbf{x} = \mathbf{0}$.
(h) The row rank of A is n.
(i) $\mathrm{RRE}(A) = I_n$.

Proof. These are separately left as exercises. Some of the equivalencies have already been demonstrated. ■