Linear Algebra 1: Lecture 8

Irena Penev

Winter 2022/2023

1 Linear transformations

We have already studied linear transformations from $\mathbb{F}^m$ to $\mathbb{F}^n$, for a field $\mathbb{F}$. The concept of a linear transformation can easily be generalized to arbitrary vector spaces, as follows.

Given vector spaces U and V over a field $\mathbb{F}$, we say that a function $f : U \to V$ is *linear* provided it satisfies the following two conditions:

1. for all $\mathbf{u}_1, \mathbf{u}_2 \in U$, we have $f(\mathbf{u}_1 + \mathbf{u}_2) = f(\mathbf{u}_1) + f(\mathbf{u}_2)$;
2. for all $\mathbf{u} \in U$ and $\alpha \in \mathbb{F}$, we have that $f(\alpha\mathbf{u}) = \alpha f(\mathbf{u})$.

Linear functions are also called *linear transformations*.

We note that in the definition of a linear function, the vector spaces (the domain and the codomain of the function) must be over the same field $\mathbb{F}$.

Example 1.1. *Let $\mathbb{P}_{\mathbb{R}}$ be the set of all polynomials with coefficients in $\mathbb{R}$. (Then $\mathbb{P}_{\mathbb{R}}$ is a vector space over $\mathbb{R}$.) Show that the function $D : \mathbb{P}_{\mathbb{R}} \to \mathbb{P}_{\mathbb{R}}$ given by*

$$D(\sum_{k=0}^{n} a_k x^k) \quad = \quad \sum_{k=1}^{n} k a_k x^{k-1}$$

for all integers $n \geq 0$ and real numbers $a_0, \dots, a_k$, is linear.

> ***Remark:*** *If you have studied calculus, you will notice that D is simply the derivative of polynomials.*

Proof. We need to check that D satisfies the two conditions from the definition of a linear function.

1. Fix $p(x), q(x) \in \mathbb{P}_{\mathbb{R}}$. Then there exists an integer $n \geq 0$ and real numbers $a_0, \dots, a_n, b_0, \dots, b_n$ such that

$$p(x) \quad = \quad \sum_{k=0}^{n} a_k x^k \qquad \text{and} \qquad q(x) \quad = \quad \sum_{k=0}^{n} b_k x^k.$$

(Note: We can take n to be any non-negative integer such that $n \geq \max\Big\{\deg(p(x)), \deg(p(x))\Big\}$. Coefficients a_k and b_k may possibly be zero.)

We now compute:

$$\begin{aligned}
D\Big(p(x)+q(x)\Big) &= D\Big((\sum_{k=0}^{n} a_k x^k) + (\sum_{k=0}^{n} b_k x^k)\Big) \\
&= D\Big(\sum_{k=0}^{n} (a_k+b_k) x^k\Big) \\
&= \sum_{k=1}^{n} k(a_k+b_k) x^{k-1} \\
&= (\sum_{k=1}^{n} k a_k x^{k-1}) + (\sum_{k=1}^{n} k b_k x^{k-1}) \\
&= D(\sum_{k=0}^{n} a_k x^k) + D(\sum_{k=0}^{n} a_k x^k) \\
&= D(p(x)) + D(q(x)).
\end{aligned}$$

2. Fix $p(x) \in \mathbb{P}_{\mathbb{F}}$ and $\alpha \in \mathbb{F}$. Then there exists an integer $n \geq 0$ and real numbers $a_0, \ldots, a_n$ such that

$$p(x) = \sum_{k=0}^{n} a_k x^k.$$

We now compute

$$\begin{aligned}
D\Big(\alpha p(x)\Big) &= D\Big(\alpha(\sum_{k=0}^{n} a_k x^k)\Big) \\
&= D\Big(\sum_{k=0}^{n} (\alpha a_k) x^k\Big) \\
&= \sum_{k=0}^{n} k(\alpha a_k) x^{k-1} \\
&= \alpha \sum_{k=0}^{n} k a_k x^{k-1} \\
&= \alpha D\Big(\sum_{k=0}^{n} a_k x^k\Big) \\
&= \alpha D\Big(p(x)\Big).
\end{aligned}$$

From 1. and 2., we conclude that D is linear. $\square$

Here is another example, for those of you who have studied calculus (if you haven't studied calculus yet, you may ignore Example 1.2 below).

Example 1.2. *Let Diff*($\mathbb{R}$) *be the set of all differentiable functions from* $\mathbb{R}$ *to* $\mathbb{R}$*, and let Func*($\mathbb{R}$) *be the set of all functions from* $\mathbb{R}$*. (Then Diff*($\mathbb{R}$) *and Func*($\mathbb{R}$) *are vector spaces over* $\mathbb{R}$*.) Show that the function* $D : \mathit{Diff}(\mathbb{R}) \to \mathit{Func}(\mathbb{R})$ *given by* $F(f) = f'$ *for all* $f \in \mathit{Diff}(\mathbb{R})$ *is linear. (As usual,* f' *denotes the derivative of* f*.)*

Proof. 1. Fix $f, g \in \text{Diff}(\mathbb{R})$. Then by the properties of the derivative, we have that $D(f+g) = (f+g)' = f' + g' = D(f) + D(g)$.

2. Fix $f \in \text{Diff}(\mathbb{R})$ and $\alpha \in \mathbb{R}$. Then by the properties of the derivative, we have that $D(\alpha f) = (\alpha f)' = \alpha f' = \alpha D(f)$.

From 1. and 2., we conclude that D is linear. □

Proposition 1.3. *Let* U*,* V*, and* W *be vector spaces over a field* $\mathbb{F}$*. Then all the following hold:*

(a) for all linear functions $f, g : U \to V$*, we have that the function* $f + g$ *is linear;*[1]

(b) for all linear functions $f : U \to V$ *and scalars* $\alpha \in \mathbb{F}$*, the function* $\alpha f : U \to V$ *is linear;*[2]

(c) for all linear functions $f : U \to V$ *and* $g : V \to W$*, the function* $g \circ f$ *is liner.*[3]

$$\begin{tikzcd} U \arrow[r, "f"] \arrow[rr, bend left, "g \circ f"] & V \arrow[r, "g"] & W \end{tikzcd}$$

Proof. We prove (c). The proofs of (a) and (b) are left as an exercise. Fix linear transformations $f : U \to V$ and $g : V \to W$. We must show that $g \circ f$ is linear.

1. Fix $\mathbf{u}_1, \mathbf{u}_2 \in U$. Then

$$\begin{aligned}
(g \circ f)(\mathbf{u}_1 + \mathbf{u}_2) &= g(f(\mathbf{u}_1 + \mathbf{u}_2)) \\
&= g(f(\mathbf{u}_1) + f(\mathbf{u}_2)) && \text{because } f \text{ is linear} \\
&= g(f(\mathbf{u}_1)) + g(f(\mathbf{u}_2)) && \text{because } g \text{ is linear} \\
&= (g \circ f)(\mathbf{u}_1) + (g \circ f)(\mathbf{u}_2).
\end{aligned}$$

---

[1] As usual, the function $f + g : U \to V$ is defined by $(f+g)(\mathbf{u}) = f(\mathbf{u}) + g(\mathbf{u})$ for all $\mathbf{u} \in U$.

[2] As usual, the function $\alpha f : U \to V$ is defined by $(\alpha f)(\mathbf{u}) = \alpha(f(\mathbf{u}))$ for all $\mathbf{u} \in U$.

[3] As usual, the function $g \circ f : U \to W$ is defined by $(g \circ f)(\mathbf{u}) = g(f(\mathbf{u}))$ for all $\mathbf{u} \in U$.

2. Fix $\mathbf{u} \in U$ and $\alpha \in \mathbb{F}$. Then

$$\begin{aligned} (g \circ f)(\alpha\mathbf{u}) &= g(f(\alpha\mathbf{u})) \\ &= g(\alpha f(\mathbf{u})) && \text{because } f \text{ is linear} \\ &= \alpha g(f(\mathbf{u})) && \text{because } g \text{ is linear} \\ &= \alpha(g \circ f)(\mathbf{u}). \end{aligned}$$

From 1. and 2., we conclude that $g \circ f$ is linear. ☐

Given vector spaces U and V over a field $\mathbb{F}$, the set of all linear transformations from U to V is denoted by $\mathrm{Hom}(U, V)$.[4] By parts (a) and (b) of Proposition 1.3, $\mathrm{Hom}(U, V)$ is a vector space over $\mathbb{F}$. The vector addition and scalar multiplication operations in $\mathrm{Hom}(U, V)$ are the addition and scalar multiplication of functions, as in parts (a) and (b) of Proposition 1.3. The zero vector in $\mathrm{Hom}(U, V)$ is the zero function, i.e. the function $f_0 : U \to V$ given by $f_0(\mathbf{u}) = \mathbf{0}_V$ for all $\mathbf{u} \in U$.

We remark that our next proposition (Proposition 1.4) was already proven in the context of linear functions from $\mathbb{F}^m$ to $\mathbb{F}^n$, where $\mathbb{F}$ is a field (see Proposition 1.4 from Lecture Notes 4). Here, we prove the proposition more generally, with an essentially identical proof.

Proposition 1.4. *Let U and V be vector spaces over a field $\mathbb{F}$, and let $f : U \to V$ be a linear transformation. Then $f(\mathbf{0}) = \mathbf{0}$.*[5]

Proof. We observe that

$$f(\mathbf{0}) \;=\; f(0 \cdot \mathbf{0}) \;\overset{(*)}{=}\; 0f(\mathbf{0}) \;\overset{(**)}{=}\; \mathbf{0},$$

where (*) follows from the fact that f is linear, and (**) follows from Proposition 2.3(a) from Lecture Notes 6.[6] ☐

For a function $f : A \to B$ and a set $A' \subseteq A$, we define

$$f[A'] \;:=\; \{f(a) \mid a \in A'\};$$

the *image* of f, denoted $\mathrm{Im}(f)$, is defined by $\mathrm{Im}(f) := f[A]$, i.e.

$$\mathrm{Im}(f) \;:=\; \{f(a) \mid a \in A\}.$$

---

[4] Linear transformations are sometimes called “homomorphisms,” which is where the notation $\mathrm{Hom}(U, V)$ comes from.

[5] Technically, this means $f(\mathbf{0}_U) = \mathbf{0}_V$, where $\mathbf{0}_U$ is the zero vector in U, and $\mathbf{0}_V$ is the zero vector in V.

[6] Proposition 2.3(a) from Lecture Notes 6 states that if V is a vector space over a field $\mathbb{F}$, then for all vectors $\mathbf{v} \in V$, we have that $0\mathbf{v} = \mathbf{0}$.

In some texts, the word "range" is used instead of "image."

Given a linear transformation $f : U \to V$, where U and V are vector spaces over a field $\mathbb{F}$, the *kernel* of f is defined to be the set

$$\mathrm{Ker}(f) := \{\mathbf{u} \in U \mid f(\mathbf{u}) = \mathbf{0}\}.$$

Example 1.5. *Let $\mathbb{P}_{\mathbb{R}}$ be the set of all polynomials with coefficients in $\mathbb{R}$. (Then $\mathbb{P}_{\mathbb{R}}$ is a vector space over $\mathbb{R}$.) Consider the function $D : \mathbb{P}_{\mathbb{R}} \to \mathbb{P}_{\mathbb{R}}$ given by*

$$D(\sum_{k=0}^{n} a_k x^k) \quad = \quad \sum_{k=1}^{n} k a_k x^{k-1}$$

for all integers $n \geq 0$ and real numbers $a_0, \dots, a_k$. In Example 1.1, it was shown that D is linear. Clearly, Ker(D) is the set of all constant polynomials, and Im(D) is the set of all polynomials (i.e. Im(D) $= \mathbb{P}_{\mathbb{R}}$).

Theorem 1.6. *Let U and V be vector spaces over a field $\mathbb{F}$, and let $f : U \to V$ be a linear transformation. Then the following hold:*

(a) Ker(f) is a subspace of U;

(b) for all subspaces W of U, $f[W]$ is a subspace of V;

(c) Im(f) is a subspace of V.

Proof. Since $\mathrm{Im}(f) = f[U]$, and U is a subspace of itself, we see that (b) implies (c). Thus, it is enough to prove (a) and (b). For both of these, we will apply Theorem 2.7 from Lecture Notes 6.

We first prove (a). Obviously, $\mathrm{Ker}(f) \subseteq U$. We now check that $\mathrm{Ker}(f)$ satisfies the three conditions from Theorem 2.7 from Lecture Notes 6.

1. By Proposition 1.4, we have that $f(\mathbf{0}) = \mathbf{0}$, and so $\mathbf{0} \in \mathrm{Ker}(f)$.[7]

2. Fix $\mathbf{u}_1, \mathbf{u}_2 \in \mathrm{Ker}(f)$. Then

$$\begin{aligned} f(\mathbf{u}_1 + \mathbf{u}_2) &= f(\mathbf{u}_1) + f(\mathbf{u}_2) && \text{because } f \text{ is linear} \\ &= \mathbf{0} + \mathbf{0} && \text{because } \mathbf{u}_1, \mathbf{u}_2 \in \mathrm{Ker}(f) \\ &= \mathbf{0}, \end{aligned}$$

and so $\mathbf{u}_1 + \mathbf{u}_2 \in \mathrm{Ker}(f)$.

3. Fix $\mathbf{u} \in \mathrm{Ker}(f)$ and $\alpha \in \mathbb{F}$. Then

$$\begin{aligned} f(\alpha\mathbf{u}) &= \alpha f(\mathbf{u}) && \text{because } f \text{ is linear} \\ &= \alpha\mathbf{0} && \text{because } \mathbf{u} \in \mathrm{Ker}(f) \\ &= \mathbf{0} && \text{by Proposition 2.3(b)} \\ & && \text{from Lecture Notes 6,} \end{aligned}$$

---

[7] Technically, we have that $f(\mathbf{0}_U) = \mathbf{0}_V$, and so $\mathbf{0}_U \in \mathrm{Ker}(f)$.

and so $\alpha\mathbf{u} \in \text{Ker}(f)$.

It now follows from Theorem 2.7 from Lecture Notes 6 that $\text{Ker}(f)$ is a subspace of U. So, (a) holds.

We now prove (b). Fix any subspace W of U. Clearly, $f[W] \subseteq V$. We now check that $f[W]$ satisfies the three conditions from Theorem 2.7 from Lecture Notes 6.

1. Since W is a subspace of U, we know that $\mathbf{0} \in W$, and by Proposition 1.4, we have that $f(\mathbf{0}) = \mathbf{0}$; so, $\mathbf{0} \in f[W]$.[8]

2. Fix $\mathbf{v}_1, \mathbf{v}_2 \in f[W]$; we must show that $\mathbf{v}_1 + \mathbf{v}_2 \in f[W]$. Then there exist $\mathbf{w}_1, \mathbf{w}_2 \in W$ such that $\mathbf{v}_1 = f(\mathbf{w}_1)$ and $\mathbf{v}_2 = f(\mathbf{W}_2)$. We now have that

$$\begin{array}{rcll} \mathbf{v}_1 + \mathbf{v}_2 & = & f(\mathbf{w}_1) + f(\mathbf{w}_2) & \text{because } \mathbf{v}_1 = f(\mathbf{w}_1) \\ & & & \text{and } \mathbf{v}_2 = f(\mathbf{w}_2) \\ & & & \\ & = & f(\mathbf{w}_1 + \mathbf{w}_2) & \text{because } f \text{ is linear.} \end{array}$$

Since $\mathbf{w}_1, \mathbf{w}_2 \in W$, and W is a subspace of U, we have that $\mathbf{w}_1 + \mathbf{w}_2 \in W$; consequently, $f(\mathbf{w}_1 + \mathbf{w}_2) \in f[W]$. So, by what we just showed, $\mathbf{v}_1 + \mathbf{v}_2 \in f[W]$.

3. Fix $\mathbf{v} \in f[W]$ and $\alpha \in \mathbb{F}$; we must show that $\alpha\mathbf{v} \in f[W]$. Since $\mathbf{v} \in f[W]$, we know that there exists some $\mathbf{w} \in W$ such that $\mathbf{v} = f(\mathbf{w})$. We now have that

$$\begin{array}{rcll} \alpha\mathbf{v} & = & \alpha f(\mathbf{w}) & \text{because } \mathbf{v} = f(\mathbf{w}) \\ & & & \\ & = & f(\alpha\mathbf{w}) & \text{because } f \text{ is linear.} \end{array}$$

Since $\mathbf{w} \in W$ and $\alpha \in \mathbb{F}$, and since W is a subspace of U, we have that $\alpha\mathbf{w} \in W$; consequently, $f(\alpha\mathbf{w}) \in f[W]$. So, by what we just showed, $\alpha\mathbf{v} \in f[W]$.

It now follows from Theorem 2.7 from Lecture Notes 6 that $f[W]$ is a subspace of V. So, (b) holds. □

2 The rank-nullity theorem for linear transformations

Proposition 2.1. *Let U and V be vector spaces over a field $\mathbb{F}$, and let $f : U \to V$ be a linear transformation. Let $\mathbf{u}_1, \dots, \mathbf{u}_k \in U$, and set $W := \mathit{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$. Then $f[W] = \mathit{Span}\Big(f(\mathbf{u}_1), \dots, f(\mathbf{u}_k)\Big)$. Moreover, $\mathit{dim}(f[W]) \leq \mathit{dim}(W) \leq k$.*

Proof. Let us show that $f[W] = \text{Span}(f(\mathbf{u}_1), \dots, f(\mathbf{u}_k))$. For this, we must prove the following:

---

[8] Technically, we have that $\mathbf{0}_U \in W$ and $f(\mathbf{0}_U) = \mathbf{0}_V$, and so $\mathbf{0}_V \in f[W]$.

(i) $f[W] \subseteq \text{Span}\Big(f(\mathbf{u}_1), \dots, f(\mathbf{u}_k)\Big)$;

(ii) $\text{Span}\Big(f(\mathbf{u}_1), \dots, f(\mathbf{u}_k)\Big) \subseteq f[W]$.

We first prove (i). Fix $\mathbf{v} \in f[W]$. Then there exists some $\mathbf{w} \in W$ such that $f(\mathbf{w}) = \mathbf{v}$. Since $W = \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$, we know that there exist scalars $\alpha_1, \dots, \alpha_k \in \mathbb{F}$ such that $\mathbf{w} = \alpha_1\mathbf{u}_1 + \dots + \alpha_k\mathbf{u}_k$. But now

$$\begin{aligned} \mathbf{v} &= f(\mathbf{w}) \\ &= f(\alpha_1\mathbf{u}_1 + \dots + \alpha_k\mathbf{u}_k) \\ &= \alpha_1 f(\mathbf{u}_1) + \dots + \alpha_k f(\mathbf{u}_k) \qquad \text{because } f \text{ is linear,} \end{aligned}$$

and it follows that $\mathbf{v} \in \text{Span}\Big(f(\mathbf{u}_1), \dots, f(\mathbf{u}_k)\Big)$. This proves (i).

We now prove (ii). Fix $\mathbf{v} \in \text{Span}\Big(f(\mathbf{u}_1), \dots, f(\mathbf{u}_k)\Big)$. Then there exist scalars $\alpha_1, \dots, \alpha_n \in \mathbb{F}$ such that $\mathbf{v} = \alpha_1 f(\mathbf{u}_1) + \dots + \alpha_k f(\mathbf{u}_k)$. Then

$$\begin{aligned} \mathbf{v} &= \alpha_1 f(\mathbf{u}_1) + \dots + \alpha_k f(\mathbf{u}_k) \\ &= f(\alpha_1\mathbf{u}_1 + \dots + \alpha_k\mathbf{u}_k) \qquad \text{because } f \text{ is linear.} \end{aligned}$$

Since $W = \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_k)$, we see that $\alpha_1\mathbf{u}_1 + \dots + \alpha_k\mathbf{u}_k \in W$; consequently, $\mathbf{v} = f(\alpha_1\mathbf{u}_1 + \dots + \alpha_k\mathbf{u}_k) \in f[W]$. This proves (ii).

We have now proven (i) and (ii), and it immediately follows that $f[W] = \text{Span}\Big(f(\mathbf{u}_1), \dots, f(\mathbf{u}_k)\Big)$.

It remains to show that $\dim(f[W]) \leq \dim(W) \leq k$. Since $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is a spanning set of W, Proposition 1.6 from Lecture Notes 7 guarantees that some subset of $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is a basis of W; after possibly permuting the elements of $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$, we may assume that, for some integer ℓ such that $0 \leq \ell \leq k$, $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell\}$ is a basis of W. Then $\dim(W) = \ell \leq k$. Next, since $\{\mathbf{u}_1, \dots, \mathbf{u}_\ell\}$ is a basis of W, we have that $W = \text{Span}(\mathbf{u}_1, \dots, \mathbf{u}_\ell)$. So, an argument completely analogous to the one above implies that $f[W] = \text{Span}\Big(f(\mathbf{u}_1), \dots, f(\mathbf{u}_\ell)\Big)$. But once again by Proposition 1.6 from Lecture Notes 7, some subset of $\{f(\mathbf{u}_1), \dots, f(\mathbf{u}_\ell)\}$ is a basis of $f[W]$, and we deduce that $\dim(f[W]) \leq \ell = \dim(W)$. This completes the argument. □

Given vector spaces U and V over a field $\mathbb{F}$ such that U is finite dimensional, and given a linear transformation $f : U \to V$, we define the *rank* of f, denoted by $\text{rank}(f)$, to be

$$\text{rank}(f) := \dim\Big(\text{Im}(f)\Big).$$

By Proposition 2.1, this is well-defined. We note that $\dim(\text{Ker}(f))$ is sometimes referred as the *nullity* of the linear transformation f. We now have the following theorem.

Rank–nullity theorem (linear transformation version). *Let U and V be vector spaces over a field $\mathbb{F}$, and assume that U is finite-dimensional. Then every linear transformation $f : U \to V$ satisfies*

$$\mathit{rank}(f) + \mathit{dim}\Big(\mathit{Ker}(f)\Big) \quad = \quad \mathit{dim}(U).$$

Proof. Since U is finite dimensional, so its subspace $\text{Ker}(U)$.[9] Let $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ be a basis of $\text{Ker}(f)$.[10] Then $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is a linearly independent set in U; since U is finite-dimensional, $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ can be extended to a basis $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_{k+\ell}\}$ (by Proposition 1.10 from Lecture Notes 7).

We now have that $\dim(\text{Ker}(f)) = k$ and $\dim(U) = k + \ell$.[11] So, to prove the theorem, it now suffices to show that $\dim\Big(\text{Im}(f)\Big) = \ell$. Clearly, it suffices to show that $\{f(\mathbf{u}_{k+1}), \dots, f(\mathbf{u}_{k+\ell})\}$ is a basis of $\text{Im}(f)$. Obviously, $f(\mathbf{u}_{k+1}), \dots, f(\mathbf{u}_{k+\ell}) \in \text{Im}(f)$, and so it suffices to show that $\{f(\mathbf{u}_{k+1}), \dots, f(\mathbf{u}_{k+\ell})\}$ is a linearly independent set, and that every vector in $\text{Im}(f)$ is a linear combination of $f(\mathbf{u}_{k+1}), \dots, f(\mathbf{u}_{k+\ell})$.

We first show that $\{f(\mathbf{u}_{k+1}), \dots, f(\mathbf{u}_{k+\ell})\}$ is linearly independent. Fix $\alpha_{k+1}, \dots, \alpha_{k+\ell} \in \mathbb{F}$ such that

$$\alpha_{k+1} f(\mathbf{u}_{k+1}) + \dots + \alpha_{k+\ell} f(\mathbf{u}_{k+\ell}) \quad = \quad \mathbf{0}.$$

We must show that $\alpha_{k+1} = \dots = \alpha_{k+\ell}$. We have that

$$\begin{aligned} \mathbf{0} \quad &= \quad \alpha_{k+1} f(\mathbf{u}_{k+1}) + \dots + \alpha_{k+\ell} f(\mathbf{u}_{k+\ell}) \\ &\overset{(*)}{=} \quad f(\alpha_{k+1}\mathbf{u}_{k+1} + \dots + \alpha_{k+\ell}\mathbf{u}_{k+\ell}), \end{aligned}$$

where (*) follows from the linearity of f. Thus, $\alpha_{k+1}\mathbf{u}_{k+1} + \dots + \alpha_{k+\ell}\mathbf{u}_{k+\ell} \in \text{Ker}(f)$. Since $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is a basis of $\text{Ker}(f)$, it follows that there exist scalars $\beta_1, \dots, \beta_k \in \mathbb{F}$ such that

$$\alpha_{k+1}\mathbf{u}_{k+1} + \dots + \alpha_{k+\ell}\mathbf{u}_{k+\ell} \quad = \quad \beta_1\mathbf{u}_1 + \dots + \beta_k\mathbf{u}_k.$$

Consequently,

$$(-\beta_1)\mathbf{u}_1 + \dots + (-\beta_k)\mathbf{u}_k + \alpha_{k+1}\mathbf{u}_{k+1} + \dots + \alpha_{k+\ell}\mathbf{u}_{k+\ell} \quad = \quad \mathbf{0}.$$

---

[9] This follows from Theorem 1.12 of Lecture Notes 7.

[10] It is possible that $k = 0$, in which case, $\{\mathbf{u}_1, \dots, \mathbf{u}_k\} = \emptyset$ and $\text{Ker}(f) = \{\mathbf{0}\}$.

[11] This is because $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is a basis of $\text{Ker}(f)$, and $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_{k+\ell}\}$ is a basis of U.

Since $\{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_{k+1}\}$ is a basis of U, and in particular, a linearly independent set in U, we see that $-\beta_1 = \dots = -\beta_k = \alpha_{k+1} = \dots = \alpha_{k+\ell} = 0$. In particular, $\alpha_{k+1} = \dots = \alpha_{k+\ell} = 0$. This proves that $\{f(\mathbf{u}_{k+1}), \dots, f(\mathbf{u}_{k+\ell})\}$ is linearly independent.

It remains to show that every vector in $\text{Im}(f)$ is a linear combination of the vectors $f(\mathbf{u}_{k+1}), \dots, f(\mathbf{u}_{k+\ell})$. Fix $\mathbf{v} \in \text{Im}(f)$. By Proposition 2.1, $\mathbf{v}$ is a linear combination of the vectors $f(\mathbf{u}_1), \dots, f(\mathbf{u}_k), f(\mathbf{u}_{k+1}), \dots, f(\mathbf{u}_{k+\ell})$. So, there exist scalars $\alpha_1, \dots, \alpha_k, \alpha_{k+1}, \alpha_{k+\ell} \in \mathbb{F}$ such that

$$\mathbf{v} \quad = \quad \alpha_1 f(\mathbf{u}_1) + \dots + \alpha_k f(\mathbf{u}_k) + \alpha_{k+1} f(\mathbf{u}_{k+1}) + \dots + \alpha_{k+\ell} f(\mathbf{u}_{k+\ell}).$$

But since $\mathbf{u}_1, \dots, \mathbf{u}_k \in \text{Ker}(f)$, we know that $f(\mathbf{u}_1) = \dots = f(\mathbf{u}_k) = \mathbf{0}$, and we deduce that

$$\begin{aligned} \mathbf{v} &= \alpha_1 f(\mathbf{u}_1) + \dots + \alpha_k f(\mathbf{u}_k) + \alpha_{k+1} f(\mathbf{u}_{k+1}) + \dots + \alpha_{k+\ell} f(\mathbf{u}_{k+\ell}) \\ &= \alpha_1 \mathbf{0} + \dots + \alpha_k \mathbf{0} + \alpha_{k+1} f(\mathbf{u}_{k+1}) + \dots + \alpha_{k+\ell} f(\mathbf{u}_{k+\ell}) \\ &= \alpha_{k+1} f(\mathbf{u}_{k+1}) + \dots + \alpha_{k+\ell} f(\mathbf{u}_{k+\ell}). \end{aligned}$$

So, $\mathbf{v}$ is a linear combination of the vectors $f(\mathbf{u}_{k+1}), \dots, f(\mathbf{u}_{k+\ell})$. This completes the argument. □

In Lecture Notes 7, we gave a slightly informal proof of the rank-nullity theorem for matrices. We are now in a position to give a fully formal proof of that theorem. Indeed, as we shall see, the rank-nullity theorem for matrices readily follows from the rank-nullity theorem for linear transformations. But first, we need the following simple proposition.

Proposition 2.2. *Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times m}$. Let $f : \mathbb{F}^m \to \mathbb{F}^n$ be given by $f(\mathbf{u}) = A\mathbf{u}$ for all $\mathbf{u} \in \mathbb{F}^m$. Then f is linear. Moreover, all the following hold:*

(a) Col(A) = Im(f);

(b) Nul(A) = Ker(f);

(c) rank(A) = rank(f).

Proof. The fact that f is linear follows from Proposition 1.1 from Lecture Notes 4. It remains to prove (a), (b), and (c). Set $A = \begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix}$.

For (a), we observe that

$$\begin{aligned}
\mathrm{Col}(A) &= \mathrm{Span}(\mathbf{a}_1, \dots, \mathbf{a}_m) \\
&= \{x_1\mathbf{a}_1 + \dots + x_m\mathbf{a}_m \mid x_1, \dots, x_m \in \mathbb{F}\} \\
&= \{\begin{bmatrix} \mathbf{a}_1 & \dots & \mathbf{a}_m \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} | x_1, \dots, x_m \in \mathbb{F}\} \\
&= \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{F}^m\} \\
&= \{f(\mathbf{x}) \mid \mathbf{x} \in \mathbb{F}^m\} \\
&= \mathrm{Im}(f).
\end{aligned}$$

For (b), we observe that

$$\begin{aligned}
\mathrm{Nul}(A) &= \{\mathbf{x} \in \mathbb{F}^m \mid A\mathbf{x} = \mathbf{0}\} \\
&= \{\mathbf{x} \in \mathbb{F}^m \mid f(\mathbf{x}) = \mathbf{0}\} \\
&= \mathrm{Ker}(f).
\end{aligned}$$

Finally, for (c), we observe that

$$\begin{aligned}
\mathrm{rank}(f) &= \dim\Big(\mathrm{Im}(f)\Big) && \text{by definiton} \\
&= \dim\Big(\mathrm{Col}(A)\Big) && \text{by (a)} \\
&= \mathrm{rank}(A) && \text{by Theorem 3.2 from Lecture Notes 7.}
\end{aligned}$$

This completes the argument. □

Rank–nullity theorem (matrix version). *Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n \times m}$. Then*

$$rank(A) + dim\Big(Nul(A)\Big) = \underbrace{m}_{\substack{= \text{ number of} \\ \text{columns of } A}}.$$

Proof. Let $f : \mathbb{F}^m \to \mathbb{F}^n$ be given by $f(\mathbf{u}) = A\mathbf{u}$ for all $\mathbf{u} \in \mathbb{F}^m$. By

Proposition 2.2, f is linear. We now have the following:

$$\begin{aligned}
\operatorname{rank}(A) + \dim\Big(\operatorname{Nul}(A)\Big) &\overset{(*)}{=} \operatorname{rank}(f) + \dim\Big(\operatorname{Ker}(f)\Big) \\
&\overset{(**)}{=} \dim(\mathbb{F}^m) \\
&= m
\end{aligned}$$

where (*) follows from Proposition 2.2, and (**) follows from the rank-nullity nullity theorem for linear transformations. □