LEARNING OUTCOMES FOR THIS SECTION

1. *How to recognize when a set of* n *column vectors in* $\mathbb{R}^n$ *(or* $\mathbb{F}^n$*) forms a basis.*
   Think about the statement of Theorem 3.1.1 first, and try out a few of your own examples with $n = 2$ or $n = 3$, to get a sense of what it is saying. This is a good thing to do before studying the proof (which is not necessarily urgent).

2. *To recognize that elements of* $\mathbb{R}^n$ *(or* $\mathbb{F}^n$*) have different coordinates with respect to different bases*
   Think about this for some examples with $n = 2$.

3. *To use the change of basis matrix to write the coordinates of any vector in* $\mathbb{F}^n$ *with respect to a given basis*
   The instruction for how to do this is the "More important conclusion" above. Try it out for other vectors besides the v that was used here. Use a couple of such examples to satisfy yourself that it works, then go over the steps to think about *why* it works.

3.2 The Rank-Nullity Theorem

Let V and W be $\mathbb{F}$-vector spaces and let $\phi : V \to W$ be a linear transformation. Recall what this means:

- $\phi(u + v) = \phi(u) + \phi(v)$ for all $u, v \in V$, and $\phi(\lambda v) = \lambda\phi(v)$, for all $v \in V$ and $\lambda \in \mathbb{F}$.

Example 3.2.1. *If* A *is matrix in* $M_m \times n(\mathbb{F})$*, then left multiplication by* A *defines a linear transformation from* $\mathbb{F}^n$ *to* $\mathbb{F}^m$*. For example, the matrix* $A = \begin{bmatrix} 2 & 3 & 1 \\ 1 & -2 & 1 \end{bmatrix}$ *defines a linear transformation from* $\mathbb{R}^3$ *to* $\mathbb{R}^2$ *via*

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \to \begin{bmatrix} 2 & 3 & 1 \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 2a + 3b + c \\ a - 2b + c \end{bmatrix}.$$

For example, $\begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix} \to \begin{bmatrix} 3 \\ 7 \end{bmatrix}$ *under this transformation.*

Note that the images of the three standard basis vectors of $\mathbb{R}^3$ *under this transfomation are respectively the columns of* A*.*

Now suppose that $\dim V = n$ and $\dim W = m$. Once we choose bases $\mathcal{B}_V$ and $\mathcal{B}_W$ for V and W, *every* linear transformation from V to W looks like the one in Example 3.2.1 above. For example, the differential operator D, which sends every polynomial to its derivative, is a linear transformation from $\mathbb{R}[x]$ to $\mathbb{R}[x]$. But $\mathbb{R}[x]$ is an infinite-dimensional space, so we'll restrict our attention to the subspace P_3, which has dimension 4 and consists of all polynomials $a_1x^3 + a_2x^2 + a_3x + a_4$, of degree at most 3. The differential operator maps P_3 to P_2 (polynomials of degree at most 2).

Now write $\mathcal{B}_3 = \{x^3, x^2, x, 1\}$ and $\mathcal{B}_2 = \{x^2, x, 1\}$ - bases for P_3 and P_2 respectively. We take each of the four basis elements of $\mathcal{B}_3$ and look at its image in P_2 under D, considered as a vector in terms of its $\mathcal{B}_2$-coordinates. We have

$$x^3 \to 3x^2 \leftrightarrow \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}_{\mathcal{B}_2}, \ x^2 \to 2x \leftrightarrow \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}_{\mathcal{B}_2}, \ x \to 1 \leftrightarrow \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}_{\mathcal{B}_2}, \ 1 \to 0 \leftrightarrow \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}_{\mathcal{B}_2}.$$

The $\mathcal{B}_3$-coordinates of the element $p(x) = ax^3 + bx^2 + cx + d$ are given by the column $\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$, and the $\mathcal{B}_2$ coordinates of the derviative of p are given by

$$a\begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}_{\mathcal{B}_2} + b\begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}_{\mathcal{B}_2} + c\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}_{\mathcal{B}_2} + d\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}_{\mathcal{B}_2} = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}_{\mathcal{B}_3}.$$

The 3×4 matrix above is the matrix of D with respect to the bases $\mathcal{B}_3$ and $\mathcal{B}_2$. Its columns are the images under D of the elements of $\mathcal{B}_2$, written with respect to $\mathcal{B}_3$. To apply the operator to any polynomial $p(x)$, we can write it as a column vector (with respect to $\mathcal{B}_3$) and then multiply by the matrix. The result has the $\mathcal{B}_2$-coordinates of $p'(x)$.

Important Note: This matrix depends on the choice of bases! Suppose we keep the basis $\mathcal{B}_2$ of P_2, but take $\mathcal{C}_3 = \{x^3+x^2, x^2+x, x+1, 1\}$ as our basis of P_3. The matrix of the differential operator with respect to this choice has the $\mathcal{B}_2$-coordinates of the derivatives of elements of $\mathcal{C}_3$ as its columns, it is given by

$$\begin{bmatrix} 3 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}.$$

To use this matrix to determine the derivative of (for example) $f(x) = x^3 + 4x^2 - x - 2$, first write $f(x)$ with respect to $\mathcal{C}_3$: $1(x^3 + x^2) + 3(x^2 + x) - 4(x + 1) + 2(1)$. Then

$$\begin{bmatrix} 3 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ -4 \\ 2 \end{bmatrix}_{\mathcal{C}_3} = \begin{bmatrix} 3 \\ 8 \\ -1 \end{bmatrix}_{\mathcal{B}_2}.$$

The key points of this example are:

- that every linear transformation becomes a matrix product once bases have been chosen for the domain and target spaces, and
- that the matrix involved depends on the choice of bases,

and definitely not that this is a recommended method for differentiating polynomials, especially not the second version of it!

Now suppose that $\Phi : V \to W$ is a linear transformation of finite dimensional vector spaces. There are a couple of important subspaces, of U and V respectively, associated with Φ.

Definition 3.2.2. *The* kernel *of* Φ *, denoted* $\ker \Phi$*, is the set of elements of* V *whose image is the zero vector of* W*.*

$$\ker \Phi = \{v \in V : \Phi(v) = 0_W\} \subseteq V.$$

Definition 3.2.3. *The* image *of* Φ *, denoted* image Φ*, is the subset of* W *consisting of the images of all the elements of* V*.*

$$\text{image } \phi = \{\phi(v) : v \in V\} \subseteq W.$$

Since every linear transformation can be defined in terms of matrices, the concepts of kernel and image also have a matrix version

Example 3.2.4. *For the linear transformation from* $\mathbb{R}^3$ *to* $\mathbb{R}^2$ *defined as left multiplication by the matrix* $A = \begin{bmatrix} 2 & 3 & 1 \\ 1 & -2 & 1 \end{bmatrix}$ *(as in Example 3.2.1 above), the* kernel *consists of all vectors* $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ *for which*

$$\begin{bmatrix} 2 & 3 & 1 \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

In the matrix context, this is referred to as the (right) nullspace *of* A*.*

We can find it by row reduction; in this example it consists of all vectors of the form $t \begin{bmatrix} -5 \\ 1 \\ 7 \end{bmatrix}$ *where* $t \in \mathbb{R}$ *- a vector subspace of dimension 1 of* $\mathbb{R}^3$*.*

The image *of this linear transformation is the subspace of* $\mathbb{R}^2$ *consisting of all products of* A *with vectors in* $\mathbb{R}^3$ *- this is the linear span of the three columns of* A*. In the matrix context, it is called the* column space *of* A*. In this example, it is all of* $\mathbb{R}^2$*, since the first two columns of* A *(for example) span* $\mathbb{R}^2$*.*

We can note that in this example, the kernel (nullspace) and image (columnspace) have dimension 1 and 2, and 1+2=3, and 3 is the dimension of the domain $\mathbb{R}^3$*. This is not a coincidence, but a case of the* Rank-Nullity Theorem*.*

If $\phi : V \to W$ is a linear transformation, the kernel and image of Φ are *subspaces* of V and W respectively. To see this:

- for $\ker \phi$: Suppose that $u, v \in \ker \phi$. Then

$$\phi(u + v) = \phi(u) + \phi(v) = 0_W + 0_W = 0_w,$$

  so $\ker \phi$ is closed under addition in V.
  If $u \in \ker \phi$ and $\lambda \in \mathbb{F}$, then

$$\phi(\lambda u) = \lambda\phi(u) = \lambda 0_W = 0_W,$$

  so $\lambda u \in \ker \phi$ and $\ker \phi$ is closed under multplication by scalars.

- for image ϕ: Suppose that $w, z \in$ image ϕ. Then $w = \phi(u)$ and $z = \phi(v)$ for some u and v in V, and

$$w + z = \phi(u) + \phi(v) = \phi(u + v),$$

  so $w + z \in$ image ϕ and image ϕ is closed under addition in W.
  If $w \in \ker \phi$ and $\lambda \in \mathbb{F}$, then $w = \phi(u)$ for $u \in V$, and

$$\lambda w = \lambda\phi(u) = \phi(\lambda u),$$

  so $\lambda w \in$ image ϕ and image ϕ is closed under multplication by scalars.

Now we come to the Rank-Nullity Theorem, which relates the dimensions of the kernel, image and domain of a linear transformation. Equivalently, it relates the dimensions of the nullspace and column space of a matrix to the number of columns. The dimension of the image of a linear trasformation is called its *rank*, and the dimension of the kernel is called the *nullity*.

Theorem 3.2.5. *Rank-Nullity Theorem Let $\phi : V \to W$ be a linear transformation, where V and W are finite-dimensional vector spaces over a field $\mathbb{F}$. Then*

$$\dim(\ker \phi) + \operatorname{rank}\phi = \dim V.$$

Proof. Write n for $\dim V$ and k for $\dim(\ker \phi)$. Let $\{v_1, \dots v_k\}$ be a basis of $\ker \phi$. This may be extended to a basis $\mathcal{B} = \{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}$ of V. Since $\mathcal{B}$ spans V, every element of image phi has the form

$$\begin{aligned} \phi(a_1 v_1 + \cdots + a_k v_k + a_{k+1} v_{k+1} + \cdots + a_n v_n) &= \phi(a_{k+1} v_{k+1} + \cdots + a_n v_n) \\ &= a_{k+1}\phi(v_{k+1}) + \cdots + a_n \phi(v_n), \end{aligned}$$

for some scalars $a_{k+1}, \dots, a_n$ (and $a_1, \dots, a_k$). It follows that the set $\mathcal{B}' = \{\phi(v_k), \dots, \phi(v_n)\} \subseteq W$ is a spanning set of image ϕ. We now show that $\mathcal{B}'$ is a basis of image ϕ, by showing that it is linearly independent. Suppose not, and suppose that

$$c_{k+1}\phi(v_{k+1}) + \cdots + c_n\phi(v_n) = 0_W$$

for some scalars $c_{k+1}, \dots, c_n$. Then

$$\phi(c_{k+1}v_{k+1} + \cdots + c_n v_n) = 0_W \Longrightarrow c_{k+1}v_{k+1} + \cdots + c_n v_n \in \ker \phi.$$

But this means that $c_{k+1}v_{k+1} + \cdots + c_n v_n$ is a linear combination of $v_1, \dots, v_k$, contrary to the linear independence of $\{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}$.

We conclude that $\mathcal{B}'$ is a basis of image ϕ, which means that the image of ϕ has dimension $n - k$ (this is the rank of ϕ), and so

$$\dim(\ker \phi) + \operatorname{rank}\phi = k + n - k = n = \dim V,$$

as required. $\square$

Theorem 3.2.6. *Rank-Nullity Theorem, matrix version Let* A *be any* $m \times n$ *matrix, with entries in a field* $\mathbb{F}$. *Then* n *is the sum of the dimension of the right nullspace of* A *and the dimension of the column space of* A.

The dimension of the columns space of a matrix A is called the *column rank* of A.

LEARNING OUTCOMES FOR THIS SECTION

1. *To recall the defintion of a linear transformation*
as a function between vector spaces that respects the addition and scalar multiplication operations.

2. *To note that left multiplication by any* $m \times n$ *matrix is a linear transformation from* $\mathbb{F}^n$ *to* $\mathbb{F}^m$,
and that the columns of the matrix are the images of the standard basis vectors of $\mathbb{F}^n$

3. *That every linear transformation can be represented as left multiplication by a matrix*,
after choosing bases for the domain and target spaces. For relatively small and manageable examples, you should be able to write down the matrix that does this, and realize that it depends on the choice of basis (we will come back to this point).

4. *To recognize the terms kernel, image, nullspace, nullity, rank and column space*.

5. *To be able to state and interpret the Rank-Nullity Theorem, in its versions for matrices and for linear transformations*
The proof is important too, but understanding the statement is more important. One way to think of it informally is that if we apply a linear transformation to a space of dimension n, the image need not have the full dimension n, because some of the elements might be mapped to zero, and so not be "recoverable" in the image (these are the elements of the kernel). But the full dimension n has to be accounted for by the combination of the kernel or the image - their dimensions must add up to n.