

ÁLGEBRA LINEAL, PARTE 3: TRANSFORMACIONES LINEALES¹

We continue following the relevant sections from our main references:

- ch. 5-7, 12 of Mota, Rumbos *Álgebra lineal en $\mathbb{R}^n$* ,
- ch. 8,5 of Vargas, Wachtel *Álgebra lineal I*.

Here we will return to studying ‘matrix algebra’ more completely, considering some basic structure of linear maps between vector spaces. In brief we will expand upon the mantra:

matrices represent linear functions.

Contents

1	Linear transformations	2
1.1	image (range), rank	4
1.2	kernel (null space), nullity	4
1.3	some examples	5
2	rank-nullity theorem	9

¹Connor Jackman, `connor.jackman@itam.mx`, 9.2026

1 Linear transformations

We recall again the origins of our subject. We are interested in some system of linear equations, say m equations in n variables

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m. \end{cases}$$

That is to say, with our more compact vector/matrix notations, we are considering a certain function (a *linear function*)

$$(*) \quad \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad \mathbf{x} \mapsto A\mathbf{x}$$

and, given $\mathbf{b} \in \mathbb{R}^m$, we seek the solution space $A^{-1}\{\mathbf{b}\} = \{\mathbf{x} : A\mathbf{x} = \mathbf{b}\} \subset \mathbb{R}^n$. Or, alternately, with our more geometric vector space language, we are asking about when is $\mathbf{b}$ in the subspace generated by the columns of A (the span of the columns of A). Here we develop more formally this ‘functional’ manner of thinking. First, note that $(*)$ has the useful ‘linearity’ property:

$$\bullet \quad A(c_1\mathbf{x}_1 + c_2\mathbf{x}_2) = c_1A\mathbf{x}_1 + c_2A\mathbf{x}_2$$

for any scalars c_j and vectors $\mathbf{x}_j$. From this property follows some general structure about solution spaces such as:

$$\bullet \quad \text{if } A\mathbf{x}_p = \mathbf{b} \text{ and } A\mathbf{x}_o = \mathbf{0} \text{ then also } A(\mathbf{x}_p + \mathbf{x}_o) = \mathbf{b} \text{ (and in fact any other solution is obtained in this way: } A^{-1}(\mathbf{b}) = \{\mathbf{x}_p + \mathbf{x}_o : A\mathbf{x}_o = \mathbf{0}\}).$$

This last linearity property above is characteristic enough to be a definition:

Definition 1. For vector spaces V, W , a function $L : V \rightarrow W$ is called linear (or a linear transformation/map) when:

$$L(a\mathbf{u} + b\mathbf{v}) = aL(\mathbf{u}) + bL(\mathbf{v})$$

for any scalars $a, b \in \mathbb{R}$ and vectors $\mathbf{u}, \mathbf{v} \in V$.

Remark. Almost anything can be reduced to a question about appropriate functions, so, studying functions is important. Most functions are complicated. Linear functions are not complicated. So we should understand them well (and one can get quite far by approximating complicated functions by linear ones, as you spend some time learning to do in calculus).

To justify the last definition:

Proposition 1. Let V, W be some vector spaces, and $\mathbf{v}_1, \dots, \mathbf{v}_n$ be some basis of V . Then any linear function $f : V \rightarrow W$ corresponds to a unique list of vectors $\mathbf{w}_1, \dots, \mathbf{w}_n \in W$ such that $f(\mathbf{v}_j) = \mathbf{w}_j$.

Proof. Since $\mathbf{v}_*$ is a basis for V , any vector in $\mathbf{v} \in V$ has a unique expansion (exercise: prove it!)

$$\mathbf{v} = x_1 \mathbf{v}_1 + \dots + x_n \mathbf{v}_n$$

in this basis. Since f is linear then

$$f(\mathbf{v}) = x_1 f(\mathbf{v}_1) + \dots + x_n f(\mathbf{v}_n)$$

so that the linear map is completely determined by the values $f(\mathbf{v}_j)$. $\square$

Remark. We should compare the list of vectors $\mathbf{w}_1 = f(\mathbf{v}_1), \dots, \mathbf{w}_n = f(\mathbf{v}_n)$ to the columns of the matrix A . Namely for $\mathbf{a}_1, \dots, \mathbf{a}_n \in W$ some columns of a matrix A then we can consider this matrix as a linear map

$$A : \mathbb{R}^n \rightarrow V, \quad \mathbf{x} \mapsto x_1 \mathbf{a}_1 + \dots + x_n \mathbf{a}_n$$

where $\mathbf{a}_j = A(\mathbf{e}_j)$ for $\mathbf{e}_j$ the standard basis of $\mathbb{R}^n$.

EXAMPLES.

- The standard example is that any $m \times n$ matrix A corresponds to a linear transformation $\mathbb{R}^n \rightarrow \mathbb{R}^m, \mathbf{x} \mapsto A\mathbf{x}$ (matrix multiplication).
- Consider $V = C^1(\mathbb{R}, \mathbb{R}) = \{f : \mathbb{R} \rightarrow \mathbb{R} \mid f \text{ is differentiable with continuous derivative}\}$ and $W = C^0(\mathbb{R}, \mathbb{R}) = \{f : \mathbb{R} \rightarrow \mathbb{R} \mid f \text{ is continuous}\}$ with the pointwise addition and scaling operations we defined before. Then differentiation:

$$D : V \rightarrow W, \quad f \mapsto f',$$

is a linear transformations, as is integration $I : W \rightarrow V, (If)(x) = \int_0^x f(u)du$. Q: Does D have an inverse? Does I have an inverse?

Comment: D is defined on the larger vector space $\tilde{V} = \{f : \mathbb{R} \rightarrow \mathbb{R} \mid f \text{ is differentiable}\}$ with image contained in $\tilde{W} = \mathcal{F}(\mathbb{R}, \mathbb{R})$ the vector space of arbitrary functions $\mathbb{R} \rightarrow \mathbb{R}$. The image $D(\tilde{V}) \subset \tilde{W}$ is partially described by Darboux's theorem of analysis.

In general for some function between some sets $f : X \rightarrow Y$ we are interested in the typical questions ('solving' the system $f(x) = y$):

- for each $y \in Y$ what are its *pre-images*, or 'solution spaces', under f ? i.e. how to describe the subsets:

$$f^{-1}\{y\} = \{x \in X : f(x) = y\} \subset X.$$

- for which $y \in Y$ do solutions exist? ie when is $f^{-1}(y) \neq \emptyset$? i.e. how to describe the *image* of X under f :

$$\text{im}(f) = \{f(x) : x \in X\} \subset Y.$$

In general we do not know how to solve such questions. Restricting our attention to linear functions, as we will see, we will be able to give answers.

1.1 image (range), rank

Regarding when there exist solutions to a given linear system (consistency of the system) we have:

Definition 2. The image (or range) of a linear map $L : V \rightarrow W$ is the subset:

$$\text{im}(L) := \{L(\mathbf{v}) : \mathbf{v} \in V\} \subset W.$$

The rank of L is the dimension of its image: $\text{rank}(L) := \dim(\text{im}(L))$.

Remark. Implicit in the second part of this definition is that the image of a linear map is a subspace of W . Exercise: verify!

Remark. We have also called the rank of a matrix the number of pivots in its row echelon form. This is not an accident as we will see below: the number of pivots in the FER of a matrix always coincides with the dimension of the image of its corresponding linear map.

Remark. The image of the linear map $\mathbb{R}^n \rightarrow \mathbb{R}^m$ corresponding to an $m \times n$ matrix A (with m rows, and n columns) is equivalently the subspace of $\mathbb{R}^m$ generated (or spanned) by the columns of A . This is also called the column space of the matrix A . Likewise the subspace of $\mathbb{R}^n$ generated by (or spanned) by the rows of A is called the row space of the matrix A .

We recall that we are interested in the image of a linear map because it allows us to decide consistency of the corresponding linear systems (existence of solutions):

$$\text{There exist solutions to } L(\mathbf{v}) = \mathbf{w} \iff \mathbf{w} \in \text{im}(L).$$

1.2 kernel (null space), nullity

Regarding how unique solutions to a given linear system might be we have:

Definition 3. The kernel (or null space) of a linear map $L : V \rightarrow W$ is the subset:

$$\ker(L) := \{\mathbf{v} : L(\mathbf{v}) = \mathbf{0}\} \subset V.$$

The nullity of L is the dimension of its kernel: $\text{null}(L) := \dim(\ker(L))$.

Remark. Implicit in the second part of this definition is that the kernel of a linear map is a subspace of V . Exercise: verify!

Remark. We will see below that the number of free variables in the FER of a matrix always coincides with the dimension of the kernel of its corresponding linear map.

Recall our interest in the kernel of a linear map stems from its measuring uniqueness of solutions to the corresponding linear systems. For instance:

$$\text{A solution to } L(\mathbf{v}) = \mathbf{w} \text{ is unique} \iff \ker L = \{\mathbf{0}\}.$$

More precisely:

Proposition 2. *Let $L : V \rightarrow W$ be a linear map. Given $\mathbf{w}_o \in \text{im}(L)$ then there is some $\mathbf{v}_p \in V$ so that $L(\mathbf{v}_p) = \mathbf{w}_o$ (a particular solution). The full solution space to $L(\mathbf{v}) = \mathbf{w}_o$ is:*

$$L^{-1}\{\mathbf{w}_o\} := \{\mathbf{v} \in V : L(\mathbf{v}) = \mathbf{w}_o\} = \{\mathbf{v}_p + \mathbf{u} : \mathbf{u} \in \ker L\} =: \mathbf{v}_p + \ker L.$$

Proof. Given our particular solution $L(\mathbf{v}_p) = \mathbf{w}_o$ then $\mathbf{v} \in V$ is another solution iff $\mathbf{v} \in L^{-1}\{\mathbf{w}_o\}$ iff $L(\mathbf{v}) = \mathbf{w}_o = L(\mathbf{v}_p)$ iff $L(\mathbf{v} - \mathbf{v}_p) = L(\mathbf{v}) - L(\mathbf{v}_p) = \mathbf{w}_o - \mathbf{w}_o = \mathbf{0}$ iff $\mathbf{v} - \mathbf{v}_p = \mathbf{u} \in \ker L$ iff $\mathbf{v} = \mathbf{v}_p + \mathbf{u}$ some $\mathbf{u} \in \ker L$ iff $\mathbf{v} \in \mathbf{v}_p + \ker L$. $\square$

So, the dimension of a solution space to a given linear system (when consistent) has dimension that of the kernel (the nullity of the linear map).

1.3 some examples

- For $m < n$, the projection:

$$\mathbb{R}^n \rightarrow \mathbb{R}^m, (x_1, \dots, x_n) \mapsto (x_1, \dots, x_m)$$

is a surjective (onto) linear map: the image is all of $\mathbb{R}^m$. Its kernel is the $n - m$ dimensional subspace $x_j = 0$ for $j \in \{m + 1, \dots, n\}$ of $\mathbb{R}^n$.

- For $m > n$, the inclusion:

$$\mathbb{R}^n \rightarrow \mathbb{R}^m, (x_1, \dots, x_n) \mapsto (x_1, \dots, x_n, 0, \dots, 0)$$

is an injective (one-to-one) linear map: its kernel is only $\mathbf{0} \in \mathbb{R}^n$. Its image is the n -dimensional subspace $x_j = 0$ for $j \in \{n + 1, \dots, m\}$ of $\mathbb{R}^m$.

- To find the image of the linear map

$$L : \mathbb{R}^3 \rightarrow \mathbb{R}^4, L \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 4y + z \\ 2x + 3y \\ 3x + 2y + z \\ 4x + y \end{pmatrix}.$$

By definition, the image of L is the set of $\mathbf{b} \in \mathbb{R}^4$ such that the system:

$$\begin{pmatrix} 1 & 4 & 1 \\ 2 & 3 & 0 \\ 3 & 2 & 1 \\ 4 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix}$$

has a solution (is consistent). This we settle by reducing to FER form:

$$\left(\begin{array}{ccc|c} 1 & 4 & 1 & b_1 \\ 2 & 3 & 0 & b_2 \\ 3 & 2 & 1 & b_3 \\ 4 & 1 & 0 & b_4 \end{array} \right) \sim \left(\begin{array}{ccc|c} 1 & 0 & 0 & \frac{-3b_1+2b_2+3b_3}{10} \\ 0 & 1 & 0 & \frac{b_1+b_2-b_3}{5} \\ 0 & 0 & 1 & \frac{b_1-2b_2+b_3}{2} \\ 0 & 0 & 0 & b_1 - b_2 - b_3 + b_4 \end{array} \right)$$

which is consistent iff $b_1 - b_2 - b_3 + b_4 = 0$. Or, in summary the image of this linear map is the 3-dimensional subspace:

$$\text{im}(L) = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in \mathbb{R}^4 \mid x_1 - x_2 - x_3 + x_4 = 0 \right\} \subset \mathbb{R}^4.$$

- To find the kernel of the linear map

$$L : \mathbb{R}^3 \rightarrow \mathbb{R}^4, \quad L \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 4y + z \\ 2x + 3y + z \\ 3x + 2y + z \\ 4x + y + z \end{pmatrix}.$$

By definition, the kernel (or null space) of L is the set of $\mathbf{x} \in \mathbb{R}^3$ solving the homogeneous system:

$$\begin{pmatrix} 1 & 4 & 1 \\ 2 & 3 & 1 \\ 3 & 2 & 1 \\ 4 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{0}$$

This we settle by reducing to FER form:

$$\begin{pmatrix} 1 & 4 & 1 \\ 2 & 3 & 1 \\ 3 & 2 & 1 \\ 4 & 1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1/5 \\ 0 & 1 & 1/5 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

so that z is a free variable and the full solution space is

$$x = -\frac{1}{5}z, \quad y = -\frac{1}{5}z$$

or, alternately the kernel of this map is the 1-dimensional subspace (line)

$$\ker L = \left\{ t \begin{pmatrix} -1 \\ -1 \\ 5 \end{pmatrix} \mid t \in \mathbb{R} \right\} = \text{span} \left\{ \begin{pmatrix} -1 \\ -1 \\ 5 \end{pmatrix} \right\} \subset \mathbb{R}^3$$

equivalently by $\ker L = \{(x, y, z)^T \in \mathbb{R}^3 \mid 5x + z = 0, \text{ and } 5y + z = 0\}$.

Remark. *The single free variable here determines all the solutions to this homogeneous system, so that the solution space is one-dimensional. In general: the number of free variables will be the same as the dimension of the solution space to said homogeneous system (the nullity, or dimension of kernel).*

- Lets suppose we want to find the image of the linear map from the previous item. We *could* proceed as in the above example to find:

$$\left(\begin{array}{ccc|c} 1 & 4 & 1 & x_1 \\ 2 & 3 & 1 & x_2 \\ 3 & 2 & 1 & x_3 \\ 4 & 1 & 1 & x_4 \end{array} \right) \sim \left(\begin{array}{ccc|c} 1 & 0 & \frac{1}{5} & \frac{-3x_1+4x_2}{5} \\ 0 & 1 & \frac{1}{5} & \frac{2x_1-x_2}{5} \\ 0 & 0 & 0 & x_1 - 2x_2 + x_3 \\ 0 & 0 & 0 & 2x_1 - 3x_2 + x_4 \end{array} \right)$$

so that the image is the 2-plane inside $\mathbb{R}^4$ defined by the pair of equations $x_1 - 2x_2 + x_3 = 0$ and $2x_1 - 3x_2 + x_4 = 0$ (each equation takes one dimension away from the image).

We may complain that this approach has some drawbacks: keeping track of the ‘image variables’ x_j in order to obtain the image described as the intersection of various level sets (the ‘consistency conditions’) is some work.

We can avoid keeping track of these ‘image variables’ x_j during the Gauss-Jordan elimination if we consider describing the image as a span of some generating vectors.

Since we know already that the image is the span of the three columns $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ our task is to determine dependencies among these three vectors in order to shorten this list of generating vectors (to a basis: minimal length). Now, finding dependencies between these three vectors is simpler since it amounts to considerations about a homogeneous system. Namely we consider only:

$$\left(\begin{array}{ccc} 1 & 4 & 1 \\ 2 & 3 & 1 \\ 3 & 2 & 1 \\ 4 & 1 & 1 \end{array} \right) \sim \left(\begin{array}{ccc} 1 & 0 & \frac{1}{5} \\ 0 & 1 & \frac{1}{5} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right)$$

so that the complete set of solutions to $x\mathbf{a}_1 + y\mathbf{a}_2 + z\mathbf{a}_3 = \mathbf{0}$ is $x = -\frac{1}{5}z, y = -\frac{1}{5}z$ or:

$$-\frac{1}{5}z\mathbf{a}_1 - \frac{1}{5}z\mathbf{a}_2 + z\mathbf{a}_3 = \mathbf{0}$$

where z is a free variable. Setting $z = 1$ we can eliminate $\mathbf{a}_3 = \frac{\mathbf{a}_1 + \mathbf{a}_2}{5}$ and thus the image is spanned by just the first two columns:

$$\text{im}(L) = \text{span}\left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 4 \\ 3 \\ 2 \\ 1 \end{pmatrix} \right\}$$

Remark. The 3rd column here corresponding to a free variable does not contribute to the span of the columns (being a linear combination of the 1st and 2nd pivot columns). Generally: to obtain a (minimal) system of vectors spanning the image (column space) it is sufficient to take those columns corresponding to those containing pivots in an FER reduction.

- To find the image and kernel of the linear map

$$\mathbb{R}^4 \rightarrow \mathbb{R}^3, \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} \mapsto \begin{pmatrix} x + 3y + z - w \\ 2x + 2y + z \\ 3x + y + z + w \end{pmatrix}.$$

Any of the methods described in the previous items will suffice to settle the question, but here we would like to present some alternatives. Let us consider the Gauss-Jordan reduction to FER form:

$$A = \begin{pmatrix} 1 & 3 & 1 & -1 \\ 2 & 2 & 1 & 0 \\ 3 & 1 & 1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1/4 & 1/2 \\ 0 & 1 & 1/4 & -1/2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Then, per the previous methods, we deduce directly:

- the pivots are in the 1st two columns so that $\text{im}(A) = \text{span}\left\{\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}\right\}$.
- In reducing to FER form the solution space of the homogeneous system (that is to say, the kernel) has not changed so that the kernel is the subspace: $x_1 + \frac{1}{4}x_3 + \frac{1}{2}x_4 = 0, x_2 + \frac{1}{4}x_3 - \frac{1}{2}x_4 = 0$. Alternately we may consider this subspace as those vectors $(x_1, x_2, x_3, x_4)^T \in \mathbb{R}^4$ perpendicular to the vectors $(1, 0, \frac{1}{4}, \frac{1}{2})^T$ and $(0, 1, \frac{1}{4}, -\frac{1}{2})^T$, or:

$$\ker A = \left(\text{span}\left\{\begin{pmatrix} 1 \\ 0 \\ \frac{1}{4} \\ \frac{1}{2} \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \frac{1}{4} \\ -\frac{1}{2} \end{pmatrix}\right\}\right)^\perp = \text{span}\left\{\begin{pmatrix} -\frac{1}{4} \\ -\frac{1}{4} \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 0 \\ 1 \end{pmatrix}\right\}$$

From this FER, our description of the kernel is quite complete (we have it described as the span of some generating vectors, and also as level sets). For the image however, our description is then not as complete, namely we only have it described with some generating vectors but not yet as level sets. For this, we may consider the transpose matrix:

$$A^T = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 1 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

which represents a certain linear map $\mathbb{R}^3 \rightarrow \mathbb{R}^4$. We observe the following:

- In reducing A^T to FER form we are replacing rows of A^T (columns of A) by their multiples and linear combinations of other rows. In particular, the span of the rows of A^T does not change upon reducing it to FER form, or in summary: *the span of the columns of A is the*

same as the span of the rows of A^T , is the same as the span of the rows of the FER reduction of A^T . Namely we have as well:

$$\text{im}(A) = \text{span}\left\{\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}\right\}.$$

So, for example, the dimension of the column space of A (the rank of A) is the same as the dimension of the row space of A (the rank of A^T).

- An $\mathbf{x} \in \mathbb{R}^3$ has $\mathbf{x} \in \ker A^T \iff A^T \mathbf{x} = \mathbf{0} \iff \mathbf{y} \cdot A^T \mathbf{x} = \mathbf{0}$ for any $\mathbf{y} \in \mathbb{R}^4$. We recall the usual dot product is (*the* characterizing property of transpose)

$$\mathbf{y} \cdot A^T \mathbf{x} = \mathbf{y}^T A^T \mathbf{x} = (A\mathbf{y})^T \mathbf{x} = (A\mathbf{y}) \cdot \mathbf{x}.$$

In conclusion: $\mathbf{x} \in \ker A^T \iff (A\mathbf{y}) \cdot \mathbf{x} = \mathbf{0}$ for all $\mathbf{y} \in \mathbb{R}^4 \iff \mathbf{x} \in (\text{im}(A))^\perp$, or:

$$\ker A^T = \text{im}(A)^\perp.$$

So, for our example since we have now more simply that $\ker A^T$ is by $x_1 = x_3, x_2 = -2x_3$ or is spanned by $(1, -2, 1)^T$, so that:

$$\text{im}(A) = \{\mathbf{x} \in \mathbb{R}^3 : x_1 - 2x_2 + x_3 = 0\}.$$

- Since $(A^T)^T = A$, we have also:

$$\ker A = \text{im}(A^T)^\perp$$

(however, we already obtain $\ker A$ efficiently from the FER of A).

2 rank-nullity theorem

From the last examples we have observed:

- the dimension of the null space (nullity) of a linear map is the same as the number of free variables in a reduction to FE form of some matrix representation.
- the dimension of the image (rank) of a linear map is the same as the number of pivots in a reduction to FE form of some matrix representation.

An $m \times n$ matrix, with m rows and n columns, corresponds to a linear map with domain $\mathbb{R}^n$. Since each column corresponds to either a pivot or free variable:

$$n = \dim(\text{domain}(L)) = \text{rank}(L) + \text{null}(L)$$

for a linear map L . Since counting dimensions is one of the most basic measures we have to start understanding a situation, these sorts of dimension counting formulas are foundational:

Theorem 1 (rank-nullity). *For $L : V \rightarrow W$ a linear map with domain a finite dimensional vector space then*

$$\dim V = \dim(\operatorname{im}(L)) + \dim(\ker L)$$

where $\operatorname{rank}(L) = \dim(\operatorname{im}(L))$ is the rank of L and $\operatorname{null}(L) = \dim(\ker L)$ is the nullity of L .

Proof. The kernel of L is a subspace of V . Let $\mathbf{u}_1, \dots, \mathbf{u}_k$ be some basis for $\ker(L) \subset V$. Extend this to a basis:

$$\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v}_1, \dots, \mathbf{v}_l$$

of V (the ‘extension Lemma’ from our intersection formula note). We claim that

$$(*) \quad L(\mathbf{v}_1), \dots, L(\mathbf{v}_l)$$

is a basis for $\operatorname{im}(L)$. Indeed:

1. $(*)$ are independent: if $\mathbf{0} = c_1 L(\mathbf{v}_1) + \dots + c_l L(\mathbf{v}_l)$ then (since L is linear) also $L(c_1 \mathbf{v}_1 + \dots + c_l \mathbf{v}_l) = \mathbf{0}$ so that $c_1 \mathbf{v}_1 + \dots + c_l \mathbf{v}_l \in \ker(L)$, and thus $c_1 \mathbf{v}_1 + \dots + c_l \mathbf{v}_l = b_1 \mathbf{u}_1 + \dots + b_k \mathbf{u}_k$ for some b_j ’s. But since $\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v}_1, \dots, \mathbf{v}_l$ are a basis (independent) this is possible only when all the c_j ’s and b_j ’s are zero.
2. $(*)$ span $\operatorname{im}(L)$: if $\mathbf{w} \in \operatorname{im}(L)$ then $\mathbf{w} = L(\mathbf{v})$ for some $\mathbf{v} \in V$. Since $\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v}_1, \dots, \mathbf{v}_l$ are a basis for V we have $\mathbf{v} = b_1 \mathbf{u}_1 + \dots + b_k \mathbf{u}_k + c_1 \mathbf{v}_1 + \dots + c_l \mathbf{v}_l$ for some coefficients b_j, c_j . Thus:

$$\mathbf{w} = L(b_1 \mathbf{u}_1 + \dots + b_k \mathbf{u}_k + c_1 \mathbf{v}_1 + \dots + c_l \mathbf{v}_l) = c_1 L(\mathbf{v}_1) + \dots + c_l L(\mathbf{v}_l)$$

since L is linear and $\mathbf{u}_j \in \ker(L)$ have $L(\mathbf{u}_j) = \mathbf{0}$.

So $(*)$ is a basis for $\operatorname{im}(L)$ and we have $\dim V = k + l = \dim(\ker L) + \dim(\operatorname{im}(L))$. $\square$

The basic situation underlying the rank-nullity theorem is captured with the standard projection:

$$\mathbb{R}^n \rightarrow \mathbb{R}^m, (x_1, \dots, x_n) \mapsto (x_1, \dots, x_m), \quad (n \geq m)$$

by considering the surjective (onto) linear map $L : V \rightarrow \operatorname{im}(L)$, where the domain V is partitioned (fibered) into parallel translates of $\ker L$, each such parallel translate a pre-image (solution space) over a certain image point:

