Matrix Methods for Computational Modeling and Data Analytics

Virginia Tech · Spring 2019

Mark Embree
embree@vt.edu

$\mathbf{Ax}=\mathbf{b}$

version of 14 February 2019

Chapter 4
Fundamentals of Subspaces

When does a linear system of equations, call it $\mathbf{Ax} = \mathbf{b}$, have a solution $\mathbf{x}$? When such a solution exists, is it unique? While we pose these abstract questions, we keep in mind the equations we derived for circuits and trusses in the last two chapters. In particular, as computational scientists, the existence and uniqueness of solutions can help us learn about our model, as when solutions to $\mathbf{A}^T\mathbf{KAx} = \mathbf{0}$ revealed the instability in the tipsy table in Figure 3.2. To develop a deep understanding of the $\mathbf{Ax} = \mathbf{b}$ problem, we need the language and tools of *subspaces*.

These notes draw heavily in spirit, details, and examples from the texts of Gilbert Strang[1] and Steve Cox[2].

[1] Gilbert Strang. *Introduction to Applied Mathematics*. Wellesley-Cambridge Press, Wellesley, MA, 1986

[2] Steven J. Cox. *Matrix Analysis* in Situ. Rice University, 2013

4.1 Subspaces

A fundamental skill of matrix theory is the ability to strategically partition n dimensional space into *subspaces* that simplify (or even trivialize) the problem we want to solve. Many such partitions are possible; the motivating application dictates which is best for a given situation. In the next few lectures we shall learn the first of several such techniques, which enables the solution of general linear systems of the form

$$\mathbf{Ax} = \mathbf{b}.$$

This decomposition of $\mathbb{R}^n$ will also make easy work of the least squares problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{b} - \mathbf{Ax}\|,$$

an essential tool for analyzing large data sets.

In data science, this least squares problem is an example of *regression*.

To describe the partitioning we have in mind, we need a few basic definitions and accompanying facts.

These facts are proved in basic linear algebra courses; we shall not dwell on all the proofs.

Definition 6 *A nonempty set of vectors* $\mathcal{S} \subseteq \mathbb{R}^n$ *is a* subspace *provided the following two conditions both hold:*

- *if* $\mathbf{v}, \mathbf{w} \in \mathcal{S}$, *then* $\mathbf{v} + \mathbf{w} \in \mathcal{S}$;
- *if* $\mathbf{v} \in \mathcal{S}$ *and* $\alpha \in \mathbb{R}$, *then* $\alpha\mathbf{v} \in \mathcal{S}$.

Mathematicians summarize these requirements by saying *a subspace is closed under vector addition and scalar multiplication.*

4.2 Subspaces: examples and counterexamples

Simple examples of subspaces include $\mathcal{S} = \{\mathbf{0}\}$ (the set containing only the zero vector) and $\mathcal{S} = \mathbb{R}^n$ (the set containing all vectors in $\mathbb{R}^n$). The most interesting subspaces occur between these extremes.

4.3 The column space and span

We want to know when $\mathbf{Ax} = \mathbf{b}$ has a solution. A special subspace provides a nice way to think about this problem. The *column space* of a matrix $\mathbf{A}$ is the set of all vectors that can be written in the form "$\mathbf{A}$ times a vector."

Definition 7 *The* column space *(or* range*) of* $\mathbf{A} \in \mathbb{R}^{m \times n}$ *is the set*

$$\mathcal{R}(\mathbf{A}) = \{\mathbf{Ax} : \mathbf{x} \in \mathbb{R}^n\}.$$

Thus $\mathcal{R}(\mathbf{A})$ contains all vectors $\mathbf{b}$ that can be written as $\mathbf{Ax}$ for some choice of $\mathbf{x}$. In other words, *if* $\mathbf{b}$ *is in the column space of* $\mathbf{A}$, *then by definition there must exist some vector* $\mathbf{x}$ *that solves* $\mathbf{Ax} = \mathbf{b}$. On the other hand, if $\mathbf{b}$ is not in $\mathcal{R}(\mathbf{A})$, then no vector $\mathbf{x}$ will give $\mathbf{Ax} = \mathbf{b}$.

Since $\mathbf{A} \in \mathbb{R}^{m \times n}$, the vector $\mathbf{Ax}$ is in $\mathbb{R}^m$, and so the vectors that make up $\mathcal{R}(\mathbf{A})$ all contain m entries: $\mathcal{R}(\mathbf{A})$ is a *subset* of $\mathbb{R}^m$. Is it a *subspace*? Yes. The proof of this fact illustrates the basic technique we use to prove that a set of vectors is a subspace.

Theorem 3 *The column space of any* $\mathbf{A} \in \mathbb{R}^{m \times n}$ *is a subspace of* $\mathbb{R}^m$.

Proof. To show that $\mathcal{R}(\mathbf{A})$ is a subspace, we must show that the two defining properties in Definition 6 hold. First, we must show that if $\mathbf{v}, \mathbf{w} \in \mathcal{R}(\mathbf{A})$, then $\mathbf{v} + \mathbf{w} \in \mathcal{R}(\mathbf{A})$.

Suppose $\mathbf{v}, \mathbf{w} \in \mathcal{R}(\mathbf{A})$. For this to be true, there must exist some vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ such that $\mathbf{Ax} = \mathbf{v}$ and $\mathbf{Ay} = \mathbf{w}$. Now apply $\mathbf{A}$ to $\mathbf{x} + \mathbf{y}$ to get

$$\mathbf{A}(\mathbf{x} + \mathbf{y}) = \mathbf{Ax} + \mathbf{Ay} = \mathbf{v} + \mathbf{w}.$$

We can write $\mathbf{Ax} = \mathbf{v}$ and $\mathbf{Ay} = \mathbf{w}$ for some $\mathbf{x}$ and $\mathbf{y}$ because of the definition of the column space: every vector in $\mathcal{R}(\mathbf{A})$ can be written as "$\mathbf{A}$ times a vector." We give these vectors the names $\mathbf{x}$ and $\mathbf{y}$ so we can easily work with them; we could have called them anything we liked.

We have showed that $\mathbf{v} + \mathbf{w}$ can be written in the form of "$\mathbf{A}$ times a vector," and so conclude that $\mathbf{v} + \mathbf{w} \in \mathcal{R}(\mathbf{A})$.

Now we must show that if $\mathbf{v} \in \mathcal{R}(\mathbf{A})$, then $\alpha\mathbf{v}$ is also in $\mathcal{R}(\mathbf{A})$ for any scalar $\alpha \in \mathbb{R}$. As before, if $\mathbf{v} \in \mathcal{R}(\mathbf{A})$, then there exists $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{Ax} = \mathbf{v}$. Now for any $\alpha \in \mathbb{R}$, apply $\mathbf{A}$ to $\alpha\mathbf{x}$ to get

$$\mathbf{A}(\alpha\mathbf{x}) = \alpha(\mathbf{Ax}) = \alpha\mathbf{v},$$

so $\alpha \mathbf{v}$ also has the form of "$\mathbf{A}$ times a vector," hence it is in $\mathcal{R}(\mathbf{A})$.

We have proved that $\mathcal{R}(\mathbf{A})$ is closed under vector addition and scalar multiplication, so it is a subspace. ■

For example, consider

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix},$$

so that

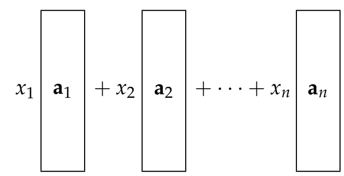
$$\mathbf{Ax} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}.$$

So we see that $\mathcal{R}(\mathbf{A})$ is the set of all vectors $\mathbf{Ax}$ that can be arrived at as a weighted sum of the columns of $\mathbf{A}$; here the values x_1 and x_2 describe how much of each column to take. The formal term for this "weighted sum" is *linear combination*.

Definition 8 *Let $x_1, \dots, x_n$ be any scalars. Then the vector*

$$x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + \cdots + x_n \mathbf{a}_n$$

is a linear combination *of the vectors $\mathbf{a}_1, \dots, \mathbf{a}_n$.*



The set of all linear combinations of a set of vectors is its *span*.

Definition 9 *The* span *of the vectors $\mathbf{a}_1, \dots, \mathbf{a}_n$ is the set of all linear combinations of these vectors:*

$$\operatorname{span}\{\mathbf{a}_1, \dots, \mathbf{a}_n\} = \{x_1 \mathbf{a}_1 + \cdots + x_n \mathbf{a}_n : x_1, \dots, x_n \in \mathbb{R}\}.$$

From these definitions we see that:

The column space is the *span* of the columns of a matrix

and

The span of $\mathbf{a}_1, \dots, \mathbf{a}_n$ is the column space of $\mathbf{A} = [\mathbf{a}_1 \ \cdots \mathbf{a}_n]$:

$$\operatorname{span}\{\mathbf{a}_1, \dots, \mathbf{a}_n\} = \mathcal{R}([\mathbf{a}_1 \ \cdots \ \mathbf{a}_n]).$$

Consider the two vectors

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \mathbf{a}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

The span of the single vector $\mathbf{a}_1$ consists of all vectors of the form

$$x_1 \mathbf{a}_1 = x_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_1 \end{bmatrix}.$$

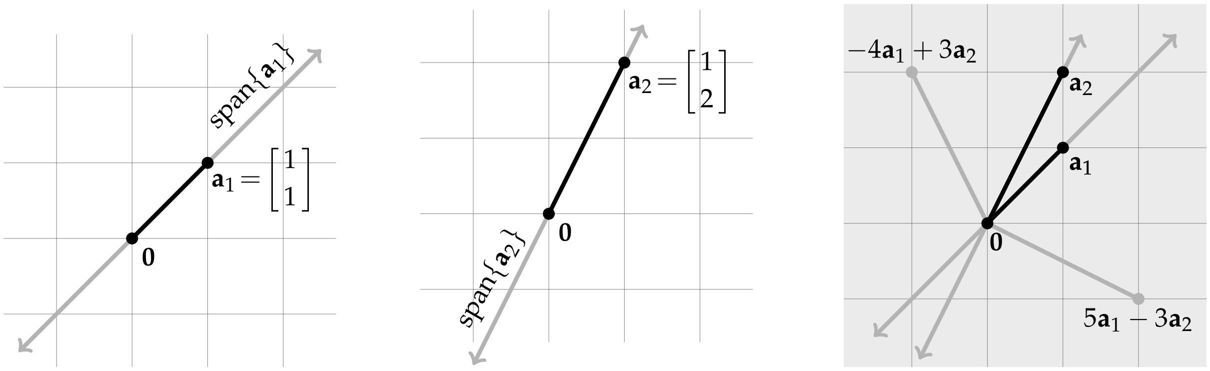


Figure 4.1: The subspaces span$\{\mathbf{a}_1\}$ (left), span$\{\mathbf{a}_2\}$ (middle), and span$\{\mathbf{a}_1, \mathbf{a}_2\}$ (right), shown in gray. In the first two plots, the subspaces are one-dimensional, consisting of all vectors that fall on the *lines* specified by $\mathbf{a}_1$ and $\mathbf{a}_2$. The last plot shows a two-dimensional subspace: it contains all points in the *plane* that can be written as a linear combination of $\mathbf{a}_1$ and $\mathbf{a}_2$.

We write

$$\text{span}\{\mathbf{a}_1\} = \left\{ \begin{bmatrix} x_1 \\ x_1 \end{bmatrix} : x_1 \in \mathbb{R} \right\}.$$

Similarly, the span of the vector $\mathbf{a}_2$ consists of all vectors

$$x_2 \mathbf{a}_2 = x_2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} x_2 \\ 2x_2 \end{bmatrix},$$

thus giving

$$\text{span}\{\mathbf{a}_2\} = \left\{ \begin{bmatrix} x_2 \\ 2x_2 \end{bmatrix} : x_2 \in \mathbb{R} \right\}.$$

To visualize these spaces, plot them in $\mathbb{R}^2$. As seen in Figure 4.1, as we consider all possible choices of x_1 and x_2, these subspaces trace out a *line* in two-dimensional space.

We plot these subspaces within two-dimensional space because the vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ each have two components: $\mathbf{a}_1, \mathbf{a}_2 \in \mathbb{R}^2$.

Now, consider the span of two vectors,

$$\text{span}\{\mathbf{a}_1, \mathbf{a}_2\} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2.$$

This subspace consists of all vectors of the form

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 = x_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_1 + 2x_2 \end{bmatrix},$$

so we write

$$\text{span}\{\mathbf{a}_1, \mathbf{a}_2\} = \left\{ \begin{bmatrix} x_1 + x_2 \\ x_1 + 2x_2 \end{bmatrix} : x_1, x_2 \in \mathbb{R} \right\}.$$

What vectors does this span contain? Suppose we want to know if some vector

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

is in span$\{\mathbf{a}_1, \mathbf{a}_2\}$. Since the span of the columns is just the column space of the matrix of columns,

$$\text{span}\{\mathbf{a}_1, \mathbf{a}_2\} = \mathcal{R}([\mathbf{a}_1, \mathbf{a}_2]),$$

we can determine if $\mathbf{b} \in \text{span}\{\mathbf{a}_1, \mathbf{a}_2\}$ by seeking x_1 and x_2 such that

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}. \tag{4.1}$$

Notice that we can find x_1 and x_2 provided the equation (4.1) has a solution for the specified values of b_1 and b_2. For our example we always find a solution, because the matrix has an inverse:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

For example, when

$$\mathbf{b} = \begin{bmatrix} -1 \\ 2 \end{bmatrix},$$

we compute

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

and so

$$\begin{bmatrix} -1 \\ 2 \end{bmatrix} = -4 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 3 \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

This vector is shown as a linear combination of $\mathbf{a}_1$ and $\mathbf{a}_2$ in the right plot of Figure 4.1. Similarly, this figure also shows

$$\begin{bmatrix} 2 \\ -1 \end{bmatrix} = 5 \begin{bmatrix} 1 \\ 1 \end{bmatrix} - 3 \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

In fact, because the matrix in (4.1) is invertible, for any vector $\mathbf{b} \in \mathbb{R}^2$, we can write $\mathbf{b} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2$ for some choice of x_1 and x_2. Hence, we say

Just extract x_1 and x_2 from $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

$$\text{span}\{\mathbf{a}_1, \mathbf{a}_2\} = \left\{ \begin{bmatrix} x_1 + x_2 \\ x_1 + 2x_2 \end{bmatrix} : x_1, x_2 \in \mathbb{R} \right\} = \mathbb{R}^2.$$

We saw the span of each of the single vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ was a *line*. The span of the two vectors together is a *plane*.

STUDENT EXPERIMENTS

4.13. Is the span of two vectors always a plane? Suppose we introduce the new vector

$$\mathbf{a}_3 = \begin{bmatrix} -2 \\ -2 \end{bmatrix}.$$

Sketch out imitations of the plots in Figure 4.1 for $\text{span}\{\mathbf{a}_3\}$ as well as $\text{span}\{\mathbf{a}_1, \mathbf{a}_3\}$ and $\text{span}\{\mathbf{a}_2, \mathbf{a}_3\}$. Are the subspaces $\text{span}\{\mathbf{a}_1, \mathbf{a}_3\}$ and $\text{span}\{\mathbf{a}_2, \mathbf{a}_3\}$ both planes?

4.4 *Null Space*

Let us return to the $\mathbf{Ax} = \mathbf{b}$ problem. In the last section we obtained a preliminary (true but incomplete) answer to the question: When does a solution $\mathbf{x}$ exist?

> The linear system $\mathbf{Ax} = \mathbf{b}$ has a solution if and only if $\mathbf{b} \in \mathcal{R}(\mathbf{A})$.

We turn now to the next critical question: Suppose for the moment that $\mathbf{Ax} = \mathbf{b}$ has a solution. When is that solution unique?

Suppose $\mathbf{y}$ and $\mathbf{z}$ both solve the equation, i.e.,

$$\mathbf{Ay} = \mathbf{b} \quad \text{and} \quad \mathbf{Az} = \mathbf{b}.$$

Then $\mathbf{Ay} = \mathbf{b} = \mathbf{Az}$, so

$$\mathbf{A}(\mathbf{y} - \mathbf{z}) = \mathbf{0}. \tag{4.2}$$

We see the special role the equation $\mathbf{Ax} = \mathbf{0}$ plays. If this equation only admits the trivial solution $\mathbf{x} = \mathbf{0}$, then the only possibility for $\mathbf{y}$ and $\mathbf{z}$ that satisfy equation (4.2) is

$$\mathbf{y} = \mathbf{z},$$

so any two solutions of $\mathbf{Ax} = \mathbf{b}$ (like our $\mathbf{y}$ and $\mathbf{z}$) must be identical: the solution of $\mathbf{Ax} = \mathbf{b}$ is *unique*.

Recall that the instability in the tipsy table was revealed by the infinitely many solutions to $\mathbf{A}^T\mathbf{KAx} = \mathbf{0}$ for the trivial load. As happens with many applications, the null space has a natural physical interpretation.

Definition 10 *The* null space *(or* kernel*) of a matrix* $\mathbf{A} \in \mathbb{R}^{m \times n}$ *is*

$$\mathcal{N}(\mathbf{A}) = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{Ax} = \mathbf{0}\}.$$

In words, the null space of $\mathbf{A}$ is the set of all vectors $\mathbf{x}$ that solve the equation $\mathbf{Ax} = \mathbf{0}$. Since $\mathbf{A0} = \mathbf{0}$, we always have $\mathbf{0} \in \mathcal{N}(\mathbf{A})$. The null space is most interesting when it contains nonzero vectors. For example, for the tipsy table in the last chapter,

$$\mathcal{N}(\mathbf{A}^T\mathbf{KA}) = \left\{ \begin{bmatrix} 0 \\ \gamma \\ 0 \\ \gamma \end{bmatrix} : \gamma \in \mathbb{C} \right\},$$

and $\mathbf{A}^T\mathbf{KAx} = \mathbf{f}$ had infinitely many solutions (if $f_2 + f_4 = 0$) or no solutions at all.

> A solution to the linear system $\mathbf{Ax} = \mathbf{b}$ is unique if and only if $\mathcal{N}(\mathbf{A}) = \{\mathbf{0}\}$.

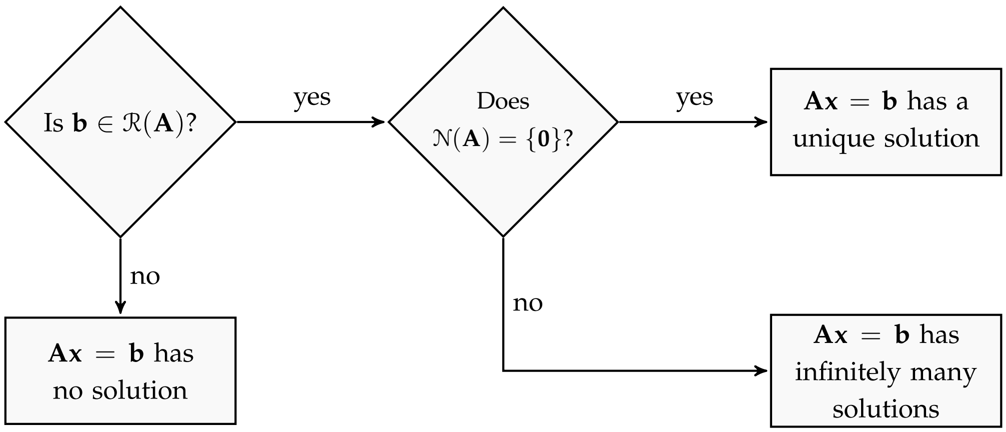


Figure 4.2: Decision process for analyzing if $\mathbf{Ax} = \mathbf{b}$ has a solution $\mathbf{x}$ for a given choice of $\mathbf{A}$ and $\mathbf{b}$.

Theorem 4 *The null space of any $\mathbf{A} \in \mathbb{R}^{m \times n}$ is a subspace of $\mathbb{R}^n$.*

Proof. Recall that we must demonstrate two properties: (1) if $\mathbf{x}, \mathbf{y} \in \mathcal{N}(\mathbf{A})$, then $\mathbf{x} + \mathbf{y} \in \mathcal{N}(\mathbf{A})$; (2) if $\mathbf{x} \in \mathcal{N}(\mathbf{A})$ and $\alpha \in \mathbb{R}$, then $\alpha\mathbf{x} \in \mathcal{N}(\mathbf{A})$.

Both properties are easy to verify. If $\mathbf{x}, \mathbf{y} \in \mathcal{N}(\mathbf{A})$, then $\mathbf{Ax} = \mathbf{0}$ and $\mathbf{Ay} = \mathbf{0}$. Hence

$$\mathbf{A}(\mathbf{x} + \mathbf{y}) = \mathbf{Ax} + \mathbf{Ay} = \mathbf{0} + \mathbf{0} = \mathbf{0},$$

so $\mathbf{x} + \mathbf{y} \in \mathcal{N}(\mathbf{A})$.

Now if $\mathbf{x} \in \mathcal{N}(\mathbf{A})$ and $\alpha \in \mathbb{C}$, then

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

so $\alpha\mathbf{x} \in \mathcal{N}(\mathbf{A})$.

Since $\mathcal{N}(\mathbf{A})$ is closed under vector addition and scalar multiplication, it is a subspace. ■

To analyze existence and uniqueness of solutions to $\mathbf{Ax} = \mathbf{b}$, follow the flowchart in Figure 4.2. To use this procedure in practice, though, we need some way of computing $\mathcal{R}(\mathbf{A})$ and $\mathcal{N}(\mathbf{A})$. That is next on our agenda.

4.5 Linear independence, basis, dimension

Unless the subspace is trivial, it will contain infinitely many vectors. We seek an economical way of describing all these vectors, avoiding any redundancy.

Definition 11 *A set of vectors $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\} \subseteq \mathbb{R}^m$ is* linearly independent *provided no one of the vectors can be written as a linear combination of* *the others. Equivalently, the only choice of scalars* $x_1, \ldots, x_n$ *for which*

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots x_n\mathbf{a}_n = \mathbf{0}$$

is $x_1 = x_2 = \cdots = x_n = 0$.

Just as the column space has a close connection to the span of a set of vectors, the null space informs our understanding of linear independence. From the second part of the definition, we see that:

> The vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ are linearly independent if and only if the null space of $\mathbf{A} = [\mathbf{a}_1 \ \cdots \mathbf{a}_n]$ contains only the zero vector:
>
> $$\mathcal{N}([\mathbf{a}_1 \ \cdots \mathbf{a}_n]) = \{\mathbf{0}\}.$$

Definition 12 *A set of vectors* $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ *is a* basis *for a subspace* $\mathcal{S}$ *provided:*

A basis is the way we describe subspaces in MATLAB.

- span$\{\mathbf{a}_1, \ldots, \mathbf{a}_n\} = \mathcal{S}$;
- *the vectors* $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ *are linearly independent.*

Notice that a subspace can have many different bases. For example, for $\mathcal{S} = \mathbb{R}^2$ (the space of all vectors of length two),

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \mathbf{a}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

is a basis; so too is

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \mathbf{a}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix},$$

as also is

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \mathbf{a}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

In a basic linear algebra course, one shows that all bases for a given subspace $\mathcal{S}$ must contain the same number of vectors.

Figure 4.3 shows how a vector $\mathbf{b}$ can be written in all three of these bases. In each case,

$$\mathbf{b} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2,$$

where the coefficients x_1 and x_2 can be determined by solving the linear system of equations

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{b}.$$

So long as $\mathbf{b} \in \mathcal{R}([\mathbf{a}_1 \ \mathbf{a}_2]) = \text{span}\{\mathbf{a}_1, \mathbf{a}_2\}$, we can find the coefficients x_1 and x_2, as shown in Figure 4.3.

STUDENT EXPERIMENTS

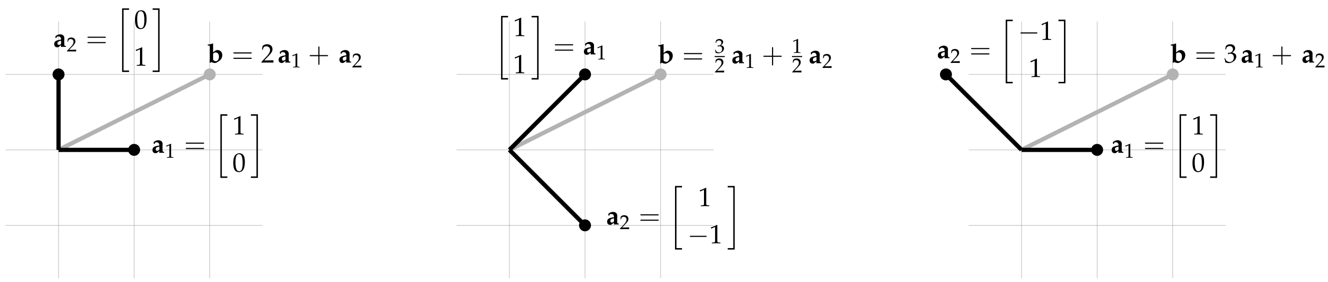


Figure 4.3: The vector $\mathbf{b} = [2, 1]^T$ is written as the linear combination of three different sets of basis vectors.

4.14. When will the coefficients x_1 and x_2 be unique?

Definition 13 *The number of elements in a basis for a subspace is called the* dimension *of the subspace. We denote the dimension of the subspace* $\mathcal{S}$ *by* $\dim(\mathcal{S})$.

For example, if $\mathbf{a}_1, \ldots, \mathbf{a}_n$ are linearly independent and

$$\mathcal{S} = \operatorname{span}\{\mathbf{a}_1, \ldots, \mathbf{a}_n\},$$

then $\dim(\mathcal{S}) = n$.

4.6 *Row reduction gives an echelon form*

In Chapter 2 we used row reduction to compute the solution to a linear system of equations (from the circuit model). Here we describe how the technique can be used to compute the column space $\mathcal{R}(\mathbf{A})$ and the null space $\mathcal{N}(\mathbf{A})$.

Definition 14 *A matrix is in* echelon form *under these conditions.*

- *The first nonzero entry in each row is called a* pivot.
- *All entries below a pivot must be zero.*
- *A pivot in row j must occur in column $k \geq j$.*
  (So pivots occur on or above the main diagonal of the matrix.)
- *All zero rows occur at the bottom of the matrix.*

A row containing a pivot is called a pivot row; *a column containing a pivot is called a* pivot column.

These requirements are similar, but less proscriptive, than the "reduced row echelon form" (RREF) often encountered in linear algebra classes. The RREF requires pivots to be equal to 1, and allows no other nonzero entries in any pivot column. Our use of the echelon form does not need these extra conditions. In any case, MATLAB's `rref` command will compute a perfectly good echelon form.

Any matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ can be transformed into echelon form using elementary row operations. For example, consider the matrix

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

Eliminate the $(3,1)$ entry with by subtracting the first row, which we encode, as in Chapter 2, with the help of an elementary matrix:

$$\mathbf{L}_1\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 0 & -1 & -1 \end{bmatrix}.$$

Now subtract the first row from the second to zero out the $(2,1)$ entry:

$$\mathbf{L}_2(\mathbf{L}_1\mathbf{A}) = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 0 & -1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & -1 & -1 \end{bmatrix}.$$

The $(1,1)$ entry of the matrix on the right is a pivot; below it we only have zeros. The $(2,2)$ entry will also be a pivot, but we must remove the $(3,2)$ entry beneath it. To do so, add the second row to the last,

$$\mathbf{L}_3(\mathbf{L}_2\mathbf{L}_1\mathbf{A}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & -1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

zeroing out the entire last row. This last matrix is in echelon form. We call it

$$\mathbf{A}_{\text{red}} = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

While that zero row is disaster for Gaussian elimination, it will not interfere with our determination of $\mathcal{R}(\mathbf{A})$ and $\mathcal{N}(\mathbf{A})$.

We have emphasized that the elementary row operations that transform $\mathbf{A}$ to $\mathbf{A}_{\text{red}}$ can be encoded as matrix-matrix multiplication:

$$(\mathbf{L}_3\mathbf{L}_2\mathbf{L}_1)\mathbf{A} = \mathbf{A}_{\text{red}}.$$

As with Gaussian elimination, these elementary row operations are all *reversible*, meaning that these elementary matrices $\mathbf{L}_j$ are invertible. Thus

$$\begin{aligned} \mathbf{A} &= (\mathbf{L}_1^{-1}\mathbf{L}_2^{-1}\mathbf{L}_3^{-1})\mathbf{A}_{\text{red}} \\ &= \mathbf{L}\mathbf{A}_{\text{red}}, \end{aligned}$$

where $\mathbf{L} = \mathbf{L}_1^{-1}\mathbf{L}_2^{-1}\mathbf{L}_3^{-1}$. While we have just worked this out for a 3×3 example, you can see that the same principles will hold in general:

$$\mathbf{A} = \mathbf{L}\mathbf{A}_{\text{red}}, \tag{4.3}$$

with $\mathbf{L}$ invertible.

4.7 *Computing the column space*

With the echelon form in hand, we are ready to compute $\mathcal{R}(\mathbf{A})$. The most efficient way to describe a subspace is with a basis, so we seek a set of basis vectors for the column space.

Start with the echelon form $\mathbf{A}_{\text{red}}$ of $\mathbf{A}$, such as

$$\mathbf{A}_{\text{red}} = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Suppose for a moment that we want the column space of $\mathbf{A}_{\text{red}}$ (rather than $\mathbf{A}$). The pivot structure makes $\mathcal{R}(\mathbf{A})$ very easy to characterize. Start with the first pivot column. It is a nonzero vector (since it contains the pivot, which must be nonzero). Moreover, it must be in $\mathcal{R}(\mathbf{A}_{\text{red}})$. In the example,

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

Thus we can use this pivot column as the first vector in a basis for $\mathcal{R}(\mathbf{A}_{\text{red}})$. Now look at the next pivot column, which also will be nonzero and in $\mathcal{R}(\mathbf{A}_{\text{red}})$, e.g.,

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

Now *here is the key point:* This second pivot column must be linearly independent of the first pivot column. Since the first pivot column has zero entries below the pivot, the pivot in the second column must occur in a row in which the first pivot column had a zero.

$$\begin{matrix} \text{first pivot} \longrightarrow \\ \text{zero beneath pivot} \longrightarrow \\ \end{matrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \begin{matrix} \\ \longleftarrow \text{second pivot} \\ \end{matrix}$$

The pivot structure ensures that the second pivot column is no multiple of the first: the two vectors are linearly independent.

The same argument applies to each remaining pivot column: regardless of the specific values in these vectors, the locations of the nonzero pivots ensures linear independence.

We have established the linear independence of the pivot columns. What about the non-pivot columns? By the structure of the pivots, *each of the non-pivot columns can be written as linear combinations of the pivot columns*, since these non-pivot columns only have nonzero entries in the pivot rows. Perhaps an example is helpful. For the matrix we have been analyzing,

$$\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix},$$

so the one non-pivot column is the sum of the two pivot columns, and hence is in their span.

Suppose $\mathbf{A}_{\text{red}}$ has r pivot columns, which we denote by

$$\widehat{\mathbf{a}}_1, \ldots, \widehat{\mathbf{a}}_r.$$

Since we have established that these pivot columns are linearly independent, and that any non-pivot column can be written as a linear combination of the pivot columns, we conclude that these vectors form a basis for $\mathcal{R}(\mathbf{A}_{\text{red}})$, and

$$\mathcal{R}(\mathbf{A}_{\text{red}}) = \text{span}\{\widehat{\mathbf{a}}_1, \ldots, \widehat{\mathbf{a}}_r\}.$$

This result is nice, *but how does it relate to our main interest*, $\mathcal{R}(\mathbf{A})$?

- Any vector in $\mathcal{R}(\mathbf{A})$ can be written as $\mathbf{Ax}$ for some $\mathbf{x} \in \mathbb{R}^n$.
- Since $\mathbf{A} = \mathbf{LA}_{\text{red}}$ as in (4.3), we have $\mathbf{Ax} = \mathbf{LA}_{\text{red}}\mathbf{x}$.
- Any vector $\mathbf{A}_{\text{red}}\mathbf{x}$ is in $\mathcal{R}(\mathbf{A}_{\text{red}})$, and so there exist constants $\gamma_1, \ldots, \gamma_r$ such that

  $$\mathbf{A}_{\text{red}}\mathbf{x} = \sum_{j=1}^{r} \gamma_j \widehat{\mathbf{a}}_j$$

  since $\mathcal{R}(\mathbf{A}_{\text{red}}) = \text{span}\{\widehat{\mathbf{a}}_1, \ldots, \widehat{\mathbf{a}}_r\}$.
- Thus any vector $\mathbf{Ax}$ in $\mathcal{R}(\mathbf{A})$ can be written as

  $$\mathbf{Ax} = \mathbf{LA}_{\text{red}}\mathbf{x} = \sum_{j=1}^{r} \gamma_j (\mathbf{L}\widehat{\mathbf{a}}_j).$$

  We conclude that

  $$\mathcal{R}(\mathbf{A}) = \text{span}\{\mathbf{L}\widehat{\mathbf{a}}_1, \ldots, \mathbf{L}\widehat{\mathbf{a}}_r\}.$$

- Thus the vectors $\mathbf{L}\widehat{\mathbf{a}}_1, \ldots, \mathbf{L}\widehat{\mathbf{a}}_r$ span $\mathcal{R}(\mathbf{A})$. Are they linearly independent, as required of a basis? Yes: For suppose there is some vector $\mathbf{y} \in \mathbb{R}^n$ for which

  $$[\, \mathbf{L}\widehat{\mathbf{a}}_1 \;\cdots\; \mathbf{L}\widehat{\mathbf{a}}_r \,]\, \mathbf{y} = \mathbf{0}.$$

  Then

  $$\mathbf{L}\, [\, \widehat{\mathbf{a}}_1 \;\cdots\; \widehat{\mathbf{a}}_r \,]\, \mathbf{y} = \mathbf{0},$$

  and, since $\mathbf{L}$ is invertible,

  $$[\, \widehat{\mathbf{a}}_1 \;\cdots\; \widehat{\mathbf{a}}_r \,]\, \mathbf{y} = \mathbf{L}^{-1}\mathbf{0} = \mathbf{0}.$$

  The only solution $\mathbf{y}$ to this last equation is $\mathbf{y} = \mathbf{0}$, since $\widehat{\mathbf{a}}_1, \ldots, \widehat{\mathbf{a}}_r$ are linearly independent. Thus $\mathbf{L}\widehat{\mathbf{a}}_1, \ldots, \mathbf{L}\widehat{\mathbf{a}}_r$ are linearly independent, and so form a basis for $\mathcal{R}(\mathbf{A})$.

- Finally, we seek to better understand the vectors $\mathbf{L}\widehat{\mathbf{a}}_1, \ldots, \mathbf{L}\widehat{\mathbf{a}}_r$.

  Let $k_1, \ldots, k_r$ be the indices of the pivot columns, and let $\mathbf{e}_\ell \in \mathbb{R}^n$ denote the vector that is zero in all entries, except for a 1 in the ℓth position. Then

  In the example, $k_1 = 1$ and $k_2 = 2$.

$$\widehat{\mathbf{a}}_\ell = \mathbf{A}_{\text{red}} \mathbf{e}_{k_\ell}$$

  plucks out the ℓth pivot column from $\mathbf{A}_{\text{red}}$. Thus

$$\mathbf{L}\widehat{\mathbf{a}}_\ell = \mathbf{L}\mathbf{A}_{\text{red}} \mathbf{e}_{k_\ell} = \mathbf{A}\mathbf{e}_{k_\ell} = \mathbf{a}_{k_\ell},$$

  where $\mathbf{a}_{k_\ell}$ is the k_ℓth column of $\mathbf{A}$.

  This is all summarized more clearly in words than symbols.

> To obtain a basis for $\mathcal{R}(\mathbf{A})$:
>
> - Compute an echelon form $\mathbf{A}_{\text{red}}$.
> - Identify the pivot columns in $\mathbf{A}_{\text{red}}$.
> - Take the columns of $\mathbf{A}$ corresponding to these pivot columns as a basis for $\mathcal{R}(\mathbf{A})$.

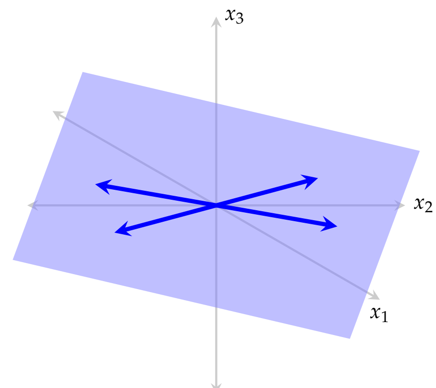


Figure 4.4: For this example, the column space $\mathcal{R}(\mathbf{A})$ is a two dimensional subspace (a plane). The blue arrows show the span of the individual basis vectors.

For the example above, the first two columns of $\mathbf{A}_{\text{red}}$ are the pivot columns, so the first two columns of $\mathbf{A}$ form a basis for $\mathcal{R}(\mathbf{A})$:

$$\mathcal{R}(\mathbf{A}) = \text{span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \right\},$$

shown in Figure 4.4. Notice: the first two columns of $\mathbf{A}_{\text{red}}$ *do not* form a basis for $\mathcal{R}(\mathbf{A})$. Beware of making this common mistake!

This is easy to see: the pivot columns of $\mathbf{A}_{\text{red}}$ both have zeros in their third entries, but the third row of $\mathbf{A}$ is nonzero, so $\mathcal{R}(\mathbf{A})$ contains vectors with nonzero third entries.

4.8 Computing the null space

Next we seek a procedure for constructing a basis for $\mathcal{N}(\mathbf{A})$. We describe the algorithm, give some examples, then explain why the procedure works.

First note that $\mathcal{N}(\mathbf{A}_{\text{red}}) = \mathcal{N}(\mathbf{A})$, so it suffices to find $\mathcal{N}(\mathbf{A}_{\text{red}})$ (which is simpler, since $\mathbf{A}_{\text{red}}$ is in echelon form).

To see this, note that $\mathbf{A}\mathbf{x} = \mathbf{0}$ means $\mathbf{L}\mathbf{A}_{\text{red}}\mathbf{x} = \mathbf{0}$, so $\mathbf{A}_{\text{red}}\mathbf{x} = \mathbf{L}^{-1}\mathbf{0} = \mathbf{0}$. Similarly, $\mathbf{A}_{\text{red}}\mathbf{x} = \mathbf{0}$ implies $\mathbf{L}\mathbf{A}_{\text{red}}\mathbf{x} = \mathbf{A}\mathbf{x} = \mathbf{0}$.

To obtain a basis for $\mathcal{N}(\mathbf{A}) = \mathcal{N}(\mathbf{A}_{\text{red}})$:

- Identify the non-pivot columns, which we will call *free columns*. If $\mathbf{A}_{\text{red}}$ has r pivots, then it has $n-r$ free columns.
- Suppose the free columns have indices $k_1, \ldots, k_{n-r}$. Then in the equation $\mathbf{A}_{\text{red}}\mathbf{x} = \mathbf{0}$, we regard the entries $x_{k_1}, \ldots, x_{k_{n-r}}$ as *free variables*.
- To obtain $n-r$ basis vectors for $\mathcal{N}(\mathbf{A}_{\text{red}})$, for each $\ell = 1, \ldots, n-r$, set the free variable $x_{k_\ell} = 1$, while setting all other free variables to zero. Then solve $\mathbf{A}_{\text{red}}\mathbf{x}_\ell = \mathbf{0}$ to get the basis vector $\mathbf{x}_\ell$.

In MATLAB `null(A)` returns a matrix whose columns give a basis for $\mathcal{N}(\mathbf{A})$. These basis vectors are scaled to be unit vectors. For a more cosmetically appealing basis, try `null(A,'r')`.

This procedure is best illustrated with examples. The 3×3 matrix we have been studying has $n - r = 3 - 2 = 1$ free variable, x_3. In this case, we will have only $n - r = 1$ basis vectors for $\mathcal{N}(\mathbf{A})$. To compute this vector, we set $x_3 = 1$ and solve

$$\mathbf{A}_{\text{red}}\mathbf{x} = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

which equates to the conditions

$$\begin{aligned} x_1 + x_2 + 2 &= 0 \\ x_2 + 1 &= 0 \\ 0 &= 0. \end{aligned}$$

The echelon form of $\mathbf{A}_{\text{red}}$ gives these equations triangular structure that allows us to find the unknown values x_1 and x_2 by back substitution, giving

$$x_1 = -1, \qquad x_2 = -1.$$

Thus, the single basis vector for $\mathcal{N}(\mathbf{A})$ is

$$\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}.$$

That last example was easy, because $\mathbf{A}_{\text{red}}$ only gave one free variable. To illustrate how to handle multiple free variables, we splice a new column into the 3×3 matrix above to give

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 & 2 \\ 1 & 2 & 2 & 3 \\ 1 & 2 & 0 & 1 \end{bmatrix}, \qquad \mathbf{A}_{\text{red}} = \begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Now columns 1 and 3 are pivot columns, while columns 2 and 4 are the free columns. Thus we have $n - r = 4 - 2 = 2$ free variables, x_2 and x_4. Construct the two basis vectors for $\mathcal{N}(\mathbf{A})$ one at a time:

- Set $x_2 = 1$ and $x_4 = 0$, and then solve $\mathbf{A}_{\text{red}}\mathbf{x} = \mathbf{0}$:

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ 1 \\ x_3 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

  whose solution gives

$$x_1 = -2, \qquad x_3 = 0$$

  and hence the vector

$$\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} \in \mathcal{N}(\mathbf{A}_{\text{red}}) = \mathcal{N}(\mathbf{A}).$$

- Set $x_2 = 0$ and $x_4 = 1$, and then solve $\mathbf{A}_{\text{red}}\mathbf{x} = \mathbf{0}$:

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ 0 \\ x_3 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

  whose solution gives

$$x_1 = -1, \qquad x_3 = -1$$

  and hence the vector

$$\begin{bmatrix} -1 \\ 0 \\ -1 \\ 1 \end{bmatrix} \in \mathcal{N}(\mathbf{A}_{\text{red}}) = \mathcal{N}(\mathbf{A}).$$

The 0-1 structure of the free variables ensures that the two vectors we have found in $\mathcal{N}(\mathbf{A})$ are linearly independent, and we have

$$\mathcal{N}(\mathbf{A}) = \text{span}\left\{ \begin{bmatrix} -2 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ -1 \\ 1 \end{bmatrix} \right\}.$$

Student experiments

4.15. Compute the column space and null space for

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 & 2 & 1 \\ 1 & 2 & 2 & 3 & 2 \\ 1 & 2 & 2 & 3 & 3 \\ 1 & 2 & 2 & 3 & 3 \end{bmatrix}.$$

Why does the procedure described above work? The explanation is simpler if we focus on a special case. Suppose the r pivots in $\mathbf{A}_{\text{red}}$ occur in columns $1, \ldots, r$, so that we have pivots in positions $(1,1)$ to (r,r) of $\mathbf{A}_{\text{red}}$. Thus we can partition $\mathbf{A}_{\text{red}}$ as

$$\mathbf{A}_{\text{red}} = \begin{bmatrix} \mathbf{T} & \mathbf{N} \\ \mathbf{0} & \mathbf{0} \end{bmatrix},$$

where $\mathbf{T}$ is an $r \times r$ upper triangular matrix with the pivots on the main diagonal, $\mathbf{N}$ is a $r \times (n-r)$ matrix, and the zero blocks have $m - r$ rows. To find a basis for $\mathcal{N}(\mathbf{A}_{\text{red}})$, we seek linearly independent vectors for which

$$\begin{bmatrix} \mathbf{T} & \mathbf{N} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{z} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix},$$

which breaks into the two equations

$$\mathbf{Ty} + \mathbf{Nz} = \mathbf{0}$$

$$\mathbf{0} = \mathbf{0}.$$

Only the first of these equations is useful; it implies

$$\mathbf{y} = -\mathbf{T}^{-1}\mathbf{Nz}.$$

We know $\mathbf{T}$ is invertible since it is a square upper triangular matrix with no zeros on the main diagonal.

So, for any value of $\mathbf{z} \in \mathbb{R}^{n-r}$, we can determine a unique corresponding value of $\mathbf{y} \in \mathbb{R}^r$. (Here the components of $\mathbf{z}$ are the free variables.) How many different ways can we choose $\mathbf{z} \in \mathbb{R}^{n-r}$ to give solutions

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

to $\mathbf{A}_{\text{red}}\mathbf{x} = \mathbf{0}$ that are linearly independent? Since $\mathbf{z}$ can be any vector of length $n - r$, there are thus $n - r$ linearly independent choices for $\mathbf{z}$. The simplest choice successively takes $\mathbf{z}$ to be the various columns of the $(n-r) \times (n-r)$ identity matrix. This makes it easy to write down all the choices for $\mathbf{z}$ at once: we seeks solutions to

$$\begin{bmatrix} \mathbf{T} & \mathbf{N} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{Y} \\ \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix},$$

which implies that $\mathbf{Y} = -\mathbf{T}^{-1}\mathbf{N}$. Hence the columns of

$$\mathbf{X} = \begin{bmatrix} -\mathbf{T}^{-1}\mathbf{N} \\ \mathbf{I} \end{bmatrix} \in \mathbb{R}^{n \times (n-r)}$$

form a basis for $\mathcal{N}(\mathbf{A}_{\text{red}}) = \mathcal{N}(\mathbf{A})$. The columns of $\mathbf{X}$ are precisely the basis vectors constructed by the procedure described at the start of this section.

In the 3×3 example above, we have

$$\mathbf{T} = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}, \quad \mathbf{N} = \begin{bmatrix} 3 \\ -1 \end{bmatrix},$$

giving the single basis vector

$$\mathbf{x} = \begin{bmatrix} -\mathbf{T}^{-1}\mathbf{N} \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}.$$

To arrive at this result, we assumed that the r pivots occurred in the first r columns of $\mathbf{A}$. If this is not the case, the argument above get a little more intricate, but the same principles apply.

STUDENT EXPERIMENTS

4.16. Suppose the pivots do not fall in the first r columns of $\mathbf{A}_{\text{red}}$. Explain how you can take construct a matrix $\mathbf{P} \in \mathbb{R}^{n\times n}$ whose rows are the same as the identity matrix, but in a scrambled order, so that

$$\mathbf{A}_{\text{red}}\mathbf{P} = \begin{bmatrix} \mathbf{T} & \mathbf{N} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

has the form described above, with $\mathbf{T} \in \mathbb{R}^{r\times r}$ containing the r pivots on the main diagonal. Now explain how to adjust the argument above to construct $\mathbf{X} \in \mathbb{R}^{n\times(n-r)}$ such that $\mathbf{A}_{\text{red}}\mathbf{X} = \mathbf{0}$.

4.9 *The Fundamental Theorem of Linear Algebra*

The chore of computing column spaces and null spaces has a payoff, for we shall see in this section how these spaces partition $\mathbb{R}^m$ and $\mathbb{R}^n$ in a way that will help us solve important application problems.

To set the stage, recall from Chapter 1 that the angle between two vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^\ell$ is defined via the inner product (dot product) to be

$$\angle(\mathbf{v}, \mathbf{w}) = \cos^{-1}\left(\frac{|\mathbf{v}^T\mathbf{w}|}{\|\mathbf{v}\|\,\|\mathbf{w}\|}\right).$$

Two vectors are *orthogonal* when they meet at a right angle, i.e.,

$$\mathbf{v}^T\mathbf{w} = 0.$$

We often indicate orthogonality with the $\perp$ notation:

Pronounce $\perp$ as "perp".

$$\mathbf{v} \perp \mathbf{w}.$$

We can also talk about orthogonality of entire subspaces. If *all* the vectors in the subspace $\mathcal{S}_1$ are orthogonal those in the subspace $\mathcal{S}_2$, we write $\mathcal{S}_1 \perp \mathcal{S}_2$. Keep these ideas in mind as we proceed.

Alongside the column space $\mathcal{R}(\mathbf{A})$, we also have the *row space*.

Definition 15 *The* row space *of a matrix* $\mathbf{A} \in \mathbb{R}^{m\times n}$ *is the set*

$$\mathcal{R}(\mathbf{A}^T) = \{\mathbf{A}^T\mathbf{y} : \mathbf{y} \in \mathbb{R}^m\}.$$

Notice that while $\mathcal{R}(\mathbf{A})$ contains vectors of length n, $\mathcal{R}(\mathbf{A}^T)$ contains vectors of length m:

$$\mathcal{R}(\mathbf{A}^T) \subseteq \mathbb{R}^n.$$

If you have already computed an echelon form $\mathbf{A}_{\text{red}}$ of $\mathbf{A}$, then it is easy to compute a basis for the row space: you simply take the transpose of the *pivot rows from* $\mathbf{A}_{\text{red}}$. We emphasize: you take rows from $\mathbf{A}_{\text{red}}$, not $\mathbf{A}$, as justified in the margin to the right.

Above in (4.3) we saw that $\mathbf{A} = \mathbf{L}\mathbf{A}_{\text{red}}$. If $\mathbf{c} \in \mathcal{R}(\mathbf{A}^T)$, then $\mathbf{A}^T\mathbf{y} = \mathbf{c}$ for some $\mathbf{y} \in \mathbb{R}^m$. But

$$\begin{aligned}\mathbf{c} = \mathbf{A}^T\mathbf{y} &= (\mathbf{L}\mathbf{A}_{\text{red}})^T\mathbf{y} \\ &= \mathbf{A}_{\text{red}}^T\mathbf{L}^T\mathbf{y} = \mathbf{A}_{\text{red}}^T(\mathbf{L}^T\mathbf{y}),\end{aligned}$$

so $\mathbf{c} \in \mathcal{R}(\mathbf{A}_{\text{red}}^T)$, hence $\mathcal{R}(\mathbf{A}^T) \subseteq \mathcal{R}(\mathbf{A}_{\text{red}}^T)$. With a similar argument one can show $\mathcal{R}(\mathbf{A}_{\text{red}}^T) \subseteq \mathcal{R}(\mathbf{A}^T)$, so $\mathcal{R}(\mathbf{A}^T) = \mathcal{R}(\mathbf{A}_{\text{red}}^T)$. Now the echelon structure of $\mathbf{A}_{\text{red}}$, with zeros in all entries before a pivot, and only one pivot per column, ensures that

To obtain a basis for $\mathcal{R}(\mathbf{A}^T)$:

- Compute an echelon form $\mathbf{A}_{\text{red}}$;
- Identify the pivot rows in $\mathbf{A}_{\text{red}}$;
- Take the transpose of the pivot rows of $\mathbf{A}_{\text{red}}$ as your basis for $\mathcal{R}(\mathbf{A}^T)$.

For our running 3×3 example,

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 1 & 0 & 1 \end{bmatrix}, \qquad \mathbf{A}_{\text{red}} = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

where the first two rows of $\mathbf{A}_{\text{red}}$ are the pivot rows, we thus obtain

$$\mathcal{R}(\mathbf{A}^T) = \text{span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}. \tag{4.4}$$

In the procedure described above and justified in the margin, we obtain the same number of basis vectors for the row space $\mathcal{R}(\mathbf{A}^T)$ as we had for the column space $\mathcal{R}(\mathbf{A})$: both are equal to the number of pivots, r. This important quantity is called the *rank* of $\mathbf{A}$.

Definition 16 *The number of pivots in an echelon form of* $\mathbf{A} \in \mathbb{R}^{m \times n}$ *is called the* rank *of* $\mathbf{A}$, *denoted* $\text{rank}(\mathbf{A})$.

We summarize as follows:

$$\dim(\mathcal{R}(\mathbf{A})) = \dim(\mathcal{R}(\mathbf{A}^T)) = \text{rank}(\mathbf{A}).$$

Now we are ready to observe a fascinating relationship between the row space and the null space. Both $\mathcal{R}(\mathbf{A}^T)$ and $\mathcal{N}(\mathbf{A})$ are subspaces of $\mathbb{R}^n$. Take a vector in each of these spaces, $\mathbf{x} \in \mathcal{N}(\mathbf{A})$ and $\mathbf{c} \in \mathcal{R}(\mathbf{A}^T)$. Since $\mathbf{c} \in \mathcal{R}(\mathbf{A}^T)$, there exists some $\mathbf{y}$ such that $\mathbf{c} = \mathbf{A}^T\mathbf{y}$. Thus

$$\mathbf{c}^T\mathbf{x} = (\mathbf{A}^T\mathbf{y})^T\mathbf{x} = \mathbf{y}^T(\mathbf{A}^T)^T\mathbf{x} = \mathbf{y}^T\mathbf{A}\mathbf{x} = \mathbf{y}^T(\mathbf{A}\mathbf{x}),$$

but since $\mathbf{x} \in \mathcal{N}(\mathbf{A})$, we must have $\mathbf{A}\mathbf{x} = \mathbf{0}$, so

$$\mathbf{c}^T\mathbf{x} = \mathbf{y}^T\mathbf{0} = 0,$$

the $\mathbf{c}$ and $\mathbf{x}$ are orthogonal. Since $\mathbf{c} = \mathbf{A}^T\mathbf{y} \in \mathcal{R}(\mathbf{A}^T)$ and $\mathbf{x} \in \mathcal{N}(\mathbf{A})$ represent arbitrary vectors in these subspaces, we must conclude that *every vector in* $\mathcal{R}(\mathbf{A}^T)$ *is orthogonal to every vector in* $\mathcal{N}(\mathbf{A})$, and vice versa. Thus, the entire subspaces are orthogonal to one another.

The row space $\mathcal{R}(\mathbf{A}^T)$ is orthogonal to the null space $\mathcal{N}(\mathbf{A})$:
$$\mathcal{R}(\mathbf{A}^T) \perp \mathcal{N}(\mathbf{A}).$$

Now assemble the three following facts.

- A basis for $\mathcal{R}(\mathbf{A}^T)$ contains r vectors, since $\dim(\mathcal{R}(\mathbf{A}^T)) = r$; call these basis vectors $\mathbf{r}_1, \ldots, \mathbf{r}_r$.
- A basis for $\mathcal{N}(\mathbf{A})$ contains $n - r$ vectors, since $\dim(\mathcal{N}(\mathbf{A})) = n - r$; call these basis vectors $\mathbf{n}_1, \ldots, \mathbf{n}_{n-r}$.
- The only vector in both $\mathcal{R}(\mathbf{A}^T)$ and $\mathcal{N}(\mathbf{A})$ is the zero vector, so the bases for $\mathcal{R}(\mathbf{A}^T)$ and $\mathcal{N}(\mathbf{A})$ have no vectors in common.

So, the r basis vectors for $\mathcal{R}(\mathbf{A}^T) \subseteq \mathbb{R}^n$ and the $n - r$ basis vectors for $\mathcal{N}(\mathbf{A}) \subseteq \mathbb{R}^n$ altogether give n linearly independent vectors in $\mathbb{R}^n$. *Since $\mathbb{R}^n$ is an n-dimensional space, these n vectors form a basis for $\mathbb{R}^n$:*

$$\mathbb{R}^n = \operatorname{span}\{\mathbf{r}_1, \ldots, \mathbf{r}_r, \mathbf{n}_1, \ldots, \mathbf{n}_{n-r}\}.$$

This fact has enormous consequences! For example, any vector $\mathbf{x} \in \mathbb{R}^n$ can be written *uniquely* as a linear combination of these basis vectors:
$$\mathbf{x} = \sum_{j=1}^{r} \gamma_j \mathbf{r}_j + \sum_{k=1}^{n-r} \gamma_{r+k} \mathbf{n}_k.$$
Write these sums individually as

$$\mathbf{x}_R = \sum_{j=1}^{r} \gamma_j \mathbf{r}_j, \qquad \mathbf{x}_N = \sum_{k=1}^{n-r} \gamma_{r+k} \mathbf{n}_k,$$

and notice that these pieces are orthogonal, since

$$\begin{aligned}\mathbf{x}_R^T \mathbf{x}_N &= \left(\sum_{j=1}^{r} \gamma_j \mathbf{r}_j \right)^T \left(\sum_{k=1}^{n-r} \gamma_{r+k} \mathbf{n}_k \right) \\ &= \sum_{j=1}^{r} \sum_{k=1}^{n-r} \gamma_j \gamma_{r+k} \mathbf{r}_j^T \mathbf{n}_k = \sum_{j=1}^{r} \sum_{k=1}^{n-r} \gamma_j \gamma_{r+k} \, 0 = 0,\end{aligned}$$

where we have used the fact that $\mathbf{r}_j^T \mathbf{n}_k = 0$ since $\mathcal{R}(\mathbf{A}^T) \perp \mathcal{N}(\mathbf{A})$.

Any vector $\mathbf{x} \in \mathbb{R}^n$ can be written uniquely as $\mathbf{x} = \mathbf{x}_R + \mathbf{x}_N$, with $\mathbf{x}_R \in \mathcal{R}(\mathbf{A}^T)$ and $\mathbf{x}_N \in \mathcal{N}(\mathbf{A})$.
The components $\mathbf{x}_R$ and $\mathbf{x}_N$ are orthogonal: $\mathbf{x}_R \perp \mathbf{x}_N$.

Mathematicians use a short-hand notation for the idea that any vector in $\mathbb{R}^n$ can be decomposed *uniquely* as the sum of a vector in $\mathcal{R}(\mathbf{A}^T)$ and a vector in $\mathcal{N}(\mathbf{A})$:

$$\mathbb{R}^n = \mathcal{R}(\mathbf{A}^T) \oplus \mathcal{N}(\mathbf{A}).$$

Were this decomposition not unique, we would only write $\mathbb{R}^n = \mathcal{R}(\mathbf{A}^T) + \mathcal{N}(\mathbf{A})$.

Our running 3×3 example has the row space

$$\mathcal{R}(\mathbf{A}^T) = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$$

and null space

$$\mathcal{N}(\mathbf{A}) = \operatorname{span}\left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \right\}.$$

Verify that the basis vectors for $\mathcal{R}(\mathbf{A}^T)$ are orthogonal to the basis vector for $\mathcal{N}(\mathbf{A})$. This orthogonality is evident in the plot of three dimensional space in Figure 4.5.

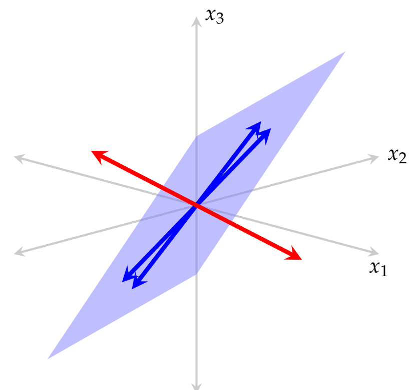


Figure 4.5: Illustration of the one-dimensional null space $\mathcal{N}(\mathbf{A})$ (red line) and the two-dimensional row space $\mathcal{R}(\mathbf{A}^T)$ (blue plane) for the 3×3 example. The blue arrows show the span of the basis vectors for $\mathcal{R}(\mathbf{A}^T)$.

The row and null spaces decompose $\mathbb{R}^n$. Of course, a similar decomposition teases $\mathbb{R}^m$ apart. To see this, simply apply the arguments above to $\mathbf{A}^T$ in place of $\mathbf{A}$. The row space of $\mathbf{A}^T$ is $\mathcal{R}(\mathbf{A})$, the column space of $\mathbf{A}$ whose computation we detailed above. We have a special name for the null space of $\mathbf{A}^T$.

Definition 17 *The* left null space *of a matrix* $\mathbf{A} \in \mathbb{R}^{m \times n}$ *is the set*

$$\mathcal{N}(\mathbf{A}^T) = \{\mathbf{y} \in \mathbb{R}^m : \mathbf{A}^T\mathbf{y} = \mathbf{0}\}.$$

The space $\mathcal{N}(\mathbf{A}^T)$ is called the *left* null space of $\mathbf{A}$ because if $\mathbf{y} \in \mathcal{N}(\mathbf{A}^T)$, then $\mathbf{y}^T\mathbf{A} = \mathbf{0}$.

Again, we have $\mathcal{N}(\mathbf{A}) \subseteq \mathbb{R}^n$, but

$$\mathcal{N}(\mathbf{A}^T) \subseteq \mathbb{R}^m.$$

As $\operatorname{rank}(\mathbf{A}) = \operatorname{rank}(\mathbf{A}^T) = r$, the left null space has dimension $m - r$:

$$\boxed{\dim(\mathcal{N}(\mathbf{A}^T)) = m - \operatorname{rank}(\mathbf{A}).}$$

There is no quick route to a basis for $\mathcal{N}(\mathbf{A}^T)$, beyond row-reducing $\mathbf{A}^T$ to echelon form $(\mathbf{A}^T)_{\text{red}}$ (which is not the same as $(\mathbf{A}_{\text{red}})^T$), and solving $(\mathbf{A}^T)_{\text{red}}\mathbf{y} = \mathbf{0}$ for $m - \operatorname{rank}(\mathbf{A})$ linearly independent vectors $\mathbf{y}$.

Applying the above arguments to $\mathbf{A}^T$, we have the following.

The column space $\mathcal{R}(\mathbf{A})$ is orthogonal to the left null space $\mathcal{N}(\mathbf{A}^T)$:

$$\mathcal{R}(\mathbf{A}) \perp \mathcal{N}(\mathbf{A}^T).$$

Similarly, the r basis vectors for $\mathcal{R}(\mathbf{A})$ and the $m-r$ basis vectors for $\mathcal{N}(\mathbf{A}^T)$ together decompose $\mathbb{R}^m$. We write this as

$$\mathbb{R}^m = \mathcal{R}(\mathbf{A}) \oplus \mathcal{N}(\mathbf{A}^T).$$

> Any vector $\mathbf{y} \in \mathbb{R}^m$ can be written uniquely as $\mathbf{y} = \mathbf{y}_R + \mathbf{y}_N$, with $\mathbf{y}_R \in \mathcal{R}(\mathbf{A})$ and $\mathbf{y}_N \in \mathcal{N}(\mathbf{A}^T)$.
>
> The components $\mathbf{y}_R$ and $\mathbf{y}_N$ are orthogonal: $\mathbf{y}_R \perp \mathbf{y}_N$.

Finally, we can collect these results in one major theorem. While it might look quite abstract, it provides a key that unlocks many problems from practical applications, like least-squares approximation. Hence Gil Strang calls this the *Fundamental Theorem of Linear Algebra*,[3] which we shall abbreviate from time to time as *FTLA*.

[3] Gilbert Strang. The Fundamental Theorem of Linear Algebra. *Amer. Math. Monthly*, 100:848–855, 1993

Theorem 5 *The column space $\mathcal{R}(\mathbf{A})$ is orthogonal to the left null space $\mathcal{N}(\mathbf{A}^T)$, and*

$$\mathcal{R}(\mathbf{A}) \oplus \mathcal{N}(\mathbf{A}^T) = \mathbb{R}^m. \tag{4.5}$$

The row space $\mathcal{R}(\mathbf{A}^T)$ is orthogonal to the null space $\mathcal{N}(\mathbf{A})$, and

$$\mathcal{R}(\mathbf{A}^T) \oplus \mathcal{N}(\mathbf{A}) = \mathbb{R}^n. \tag{4.6}$$

The plot in Figure 4.5 gives an impression of the orthogonality of these spaces, but the result is easier to visualize in two dimensions. Consider the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}, \qquad \mathbf{A}_{\text{red}} = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}.$$

Using the techniques described above, you can determine:

$$\mathcal{R}(\mathbf{A}) = \text{span}\left\{\begin{bmatrix} 1 \\ 3 \end{bmatrix}\right\}, \quad \mathcal{N}(\mathbf{A}^T) = \text{span}\left\{\begin{bmatrix} -3 \\ 1 \end{bmatrix}\right\},$$
$$\mathcal{R}(\mathbf{A}^T) = \text{span}\left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right\}, \quad \mathcal{N}(\mathbf{A}) = \text{span}\left\{\begin{bmatrix} -2 \\ 1 \end{bmatrix}\right\}.$$

Figure 4.6 shows the orthogonality of the fundamental subspaces.

4.10 Least squares, take one

We know now that $\mathbf{Ax} = \mathbf{b}$ has a solution only when $\mathbf{b} \in \mathcal{R}(\mathbf{A})$. When $\mathbf{b} \notin \mathcal{R}(\mathbf{A})$, we might relax our search for an exact solution, and instead be content with the best approximate solution we can find. In other words, we seek the vector $\mathbf{x} \in \mathbb{R}^n$ that minimizes the misfit

$$\|\mathbf{b} - \mathbf{Ax}\|,$$

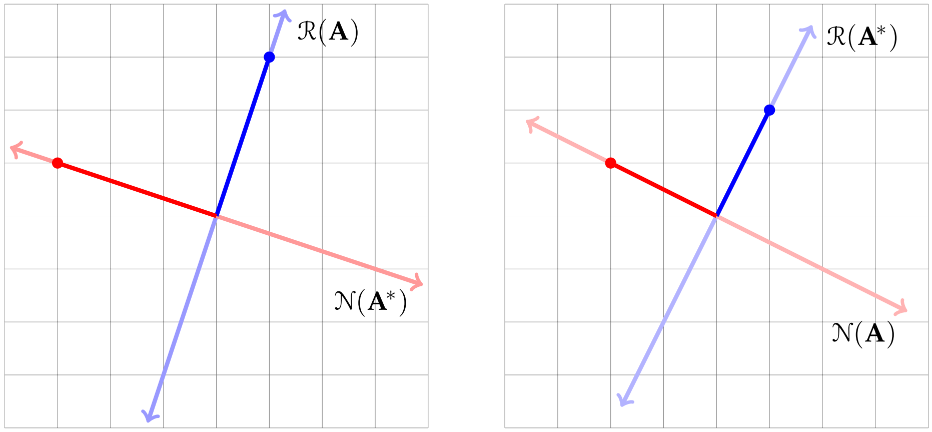


Figure 4.6: The four fundamental subspaces of the 2×2 matrix $\mathbf{A}$. Notice the orthogonality of the subspaces described in the FTLA. In each plot, notice that any vector in the plane can be expressed as the linear combination of the basis vectors for the fundamental subspaces (shown as the darker lines).

or, equivalently, the squared misfit

$$\|\mathbf{b} - \mathbf{A}\mathbf{x}\|^2 = \sum_{j=1}^{m} \Big| b_j - \sum_{k=1}^{n} a_{j,k} x_k \Big|^2;$$

This latter form explains why this approximation is called a *least squares problem*.

Some applications suggest other characterizations of optimality, aside from least squares. For example, the problem of finding $\mathbf{x}$ to minimize

$$\sum_{j=1}^{m} \Big| b_j - \sum_{k=1}^{n} a_{j,k} x_k \Big|,$$

arises in signal processing, often associated with the field of *compressive sensing* or *sparse approximation*.

The FTLA provides the key to unlock these least squares problems. Start by decomposing

$$\mathbf{b} = \mathbf{b}_R + \mathbf{b}_N, \qquad \mathbf{b}_R \in \mathcal{R}(\mathbf{A}), \quad \mathbf{b}_N \in \mathcal{N}(\mathbf{A}^T).$$

Now for any $\mathbf{x} \in \mathbb{R}^n$,

$$\begin{aligned} \mathbf{b} - \mathbf{A}\mathbf{x} &= \mathbf{b}_R + \mathbf{b}_N - \mathbf{A}\mathbf{x} \\ &= (\mathbf{b}_R - \mathbf{A}\mathbf{x}) + \mathbf{b}_N, \end{aligned}$$

where we have grouped terms strategically: By its form we must have $\mathbf{A}\mathbf{x} \in \mathcal{R}(\mathbf{A})$, and since $\mathcal{R}(\mathbf{A})$ is a subspace, the sum of two $\mathcal{R}(\mathbf{A})$ vectors is also in $\mathcal{R}(\mathbf{A})$. The FTLA tells us that $\mathcal{R}(\mathbf{A}) \perp \mathcal{N}(\mathbf{A}^T)$, so we can apply the Pythagorean Theorem to obtain

$$\|\mathbf{b} - \mathbf{A}\mathbf{x}\|^2 = \|\mathbf{b}_R - \mathbf{A}\mathbf{x}\|^2 + \|\mathbf{b}_N\|^2 \tag{4.7}$$

for any $\mathbf{x}$. To solve the least squares problem, to get the best approximate solution to $\mathbf{A}\mathbf{x} = \mathbf{b}$, we need to pick $\mathbf{x}$ to minimize the expression (4.7). No choice of $\mathbf{x}$ can reach the $\|\mathbf{b}_N\|^2$ term; on the other hand, we can pick $\mathbf{x}$ to reach any vector $\mathbf{A}\mathbf{x}$ in $\mathcal{R}(\mathbf{A})$. In particular, we can pick $\mathbf{x}$ so that $\mathbf{A}\mathbf{x} = \mathbf{b}_R$, and with this choice

We seek $\mathbf{x}$ that solves $\mathbf{A}\mathbf{x} = \mathbf{b}_R$. As emphasized in Figure 4.2, this equation always has a solution since $\mathbf{b}_R \in \mathcal{R}(\mathbf{A})$.

$$\begin{aligned} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|^2 &= \|\mathbf{b}_R - \mathbf{A}\mathbf{x}\|^2 + \|\mathbf{b}_N\|^2 \\ &= \|\mathbf{b}_R - \mathbf{b}_R|^2 + \|\mathbf{b}_N\|^2 \\ &= \|\mathbf{b}_N\|^2, \end{aligned}$$

and this choice must be optimal, since $\|\mathbf{b}_R - \mathbf{A}\mathbf{x}\|^2 \geq 0$ for all $\mathbf{x}$.

> The approximation problem
>
> $$\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|$$
>
> is solved by any $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{A}\mathbf{x} = \mathbf{b}_R$.

STUDENT EXPERIMENTS

4.17. State conditions in which the solution $\mathbf{x}$ to the least squares problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|$$

is unique. In other words, when does the equation $\mathbf{A}\mathbf{x} = \mathbf{b}_R$ have a unique solution?

One last problem remains: How do we solve $\mathbf{A}\mathbf{x} = \mathbf{b}_R$, when $\mathbf{A}$ is a rectangular matrix? In particular, we would ideally solve the least squares problem without explicitly constructing $\mathbf{b}_R$. Notice that for an optimal choice of $\mathbf{x}$ that solves $\mathbf{A}\mathbf{x} = \mathbf{b}_R$, we have

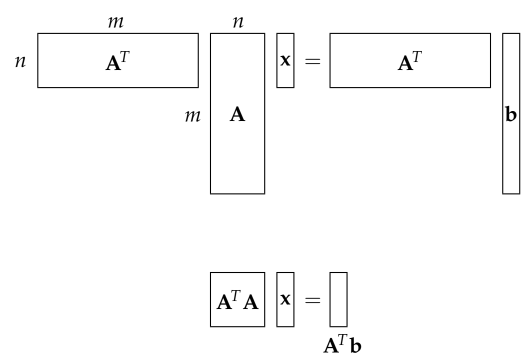
$$\mathbf{b} - \mathbf{A}\mathbf{x} = (\mathbf{b}_R - \mathbf{A}\mathbf{x}) + \mathbf{b}_N = 0 + \mathbf{b}_N.$$

Premultiply this equation by $\mathbf{A}^T$ and recall that $\mathbf{b}_N \in \mathcal{N}(\mathbf{A}^T)$ to obtain

$$\begin{aligned} \mathbf{A}^T(\mathbf{b} - \mathbf{A}\mathbf{x}) &= \mathbf{A}^T\mathbf{b}_N \\ &= \mathbf{0}, \end{aligned}$$

and so we rearrange to find

$$\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}. \tag{4.8}$$



The optimal $\mathbf{x}$ that solves the least squares problem also solves the linear system (4.8). This system is called the *normal equations*. Notice that $\mathbf{A}^T\mathbf{A} \in \mathbb{R}^{n\times n}$. In many applications, the columns of $\mathbf{A}$ correspond to variables we are trying to fit, while the rows of $\mathbf{A}$ represent observations (or the results of experiments) that provide the data for the fit. We can often collect many more observations than we have variables, so $m \gg n$. Yet the system $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}$ only requires that we solve a small $n \times n$ system to find the best fit $\mathbf{x}$.

When does $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}$ have a *unique* solution? It is easy to see that $\mathcal{N}(\mathbf{A}) \subset \mathcal{N}(\mathbf{A}^T\mathbf{A})$: for if $\mathbf{x} \in \mathcal{N}(\mathbf{A})$, then $\mathbf{A}\mathbf{x} = \mathbf{0}$, and so $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{0} = \mathbf{0}$; hence $\mathbf{x} \in \mathcal{N}(\mathbf{A}^T\mathbf{A})$. If we can also show that $\mathcal{N}(\mathbf{A}^T\mathbf{A}) \subset \mathcal{N}(\mathbf{A})$, then we can conclude that $\mathcal{N}(\mathbf{A}^T\mathbf{A}) = \mathcal{N}(\mathbf{A})$.

Suppose $\mathbf{x} \in \mathcal{N}(\mathbf{A}^T\mathbf{A})$, so that $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{0}$. Premultiply this equation by $\mathbf{x}^T$ to get

$$\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = 0.$$

But since $0 = \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = \|\mathbf{A}\mathbf{x}\|^2$, we must have $\mathbf{A}\mathbf{x} = \mathbf{0}$. Thus $\mathbf{x} \in \mathcal{N}(\mathbf{A})$, and hence $\mathcal{N}(\mathbf{A}^T\mathbf{A}) \subset \mathcal{N}(\mathbf{A})$.

Theorem 6 *The normal equations* $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}$ *have a unique solution provided* $\mathcal{N}(\mathbf{A}) = \{\mathbf{0}\}$.

4.11 The Fredholm Alternative

As a coda, we return to the problem of solving $\mathbf{A}\mathbf{x} = \mathbf{b}$, and recover a famous result that is often invoked but too seldom explained. There are various formulations. (Citation for this one – Wikipedia?).

Suppose $\mathbf{A}\mathbf{x} = \mathbf{b}$ has no solution. From above we know that this means $\mathbf{b} \notin \mathcal{R}(\mathbf{A})$, so that when, via the FTLA, we decompose $\mathbf{b}$ as

$$\mathbf{b} = \mathbf{b}_R + \mathbf{b}_N, \qquad \mathbf{b}_R \in \mathcal{R}(\mathbf{A}), \mathbf{b}_N \in \mathcal{N}(\mathbf{A}^T),$$

we must have $\mathbf{b}_N \neq \mathbf{0}$. Hence

$$\mathbf{b}_N^T\mathbf{b} = \mathbf{b}_N^T\mathbf{b}_R + \mathbf{b}_N^T\mathbf{b}_N = 0 + \|\mathbf{b}_N\|^2 \neq 0,$$

and $\mathbf{A}^T\mathbf{b}_N = \mathbf{0}$.

The Fredholm Alternative becomes more interesting when applied to integral equations, which can be interpreted as infinite-dimensional matrices. For details, see Ramm:

A. G. Ramm. A simple proof of the Fredholm alternative and a characterization of the Fredholm operators. *Amer. Math. Monthly*, 108:855–860, 2001

Proposition 1 (Fredholm Alternative) *For every matrix* $\mathbf{A} \in \mathbb{R}^{m\times n}$ *and vector* $\mathbf{b} \in \mathbb{R}^m$, *exactly one of the following is true:*

- *Either* $\mathbf{A}\mathbf{x} = \mathbf{b}$ *has a solution* $\mathbf{x} \in \mathbb{R}^n$;
- *Or there exists some nonzero* $\mathbf{y} \in \mathbb{R}^m$ *such that* $\mathbf{A}^T\mathbf{y} = \mathbf{0}$ *and* $\mathbf{y}^T\mathbf{b} \neq 0$.