ii. For all $j \neq \ell$ and all $c, d \in \mathbb{K}$, show that the row add operations $\vec{r}_k + d\,\vec{r}_\ell \mapsto \vec{r}_k$ and $\vec{r}_j + c\,\vec{r}_k \mapsto \vec{r}_j$ fail to commute

iii. For all $i \neq \ell$, $j \neq k$ and all $c, d \in \mathbb{K}$, show that the row add operations $\vec{r}_i + c\,\vec{r}_j \mapsto \vec{r}_i$ and $\vec{r}_k + d\,\vec{r}_\ell \mapsto \vec{r}_k$ fail to commute.

4.1.7 Problem. Find all solutions to the linear system with augmented matrix

$$\begin{bmatrix} 4 & 1 & 4 & 8 & b_1 \\ 0 & 3 & 1 & 8 & b_2 \\ 6 & 7 & 8 & 7 & b_3 \\ 9 & 2 & 9 & 8 & b_4 \end{bmatrix}$$

where the constant term vector $\vec{b}$ is either $\begin{bmatrix} 17 \\ 12 \\ 28 \\ 28 \end{bmatrix}$ or $\begin{bmatrix} 17.01 \\ 12.07 \\ 28.07 \\ 28.06 \end{bmatrix}$.

4.2 Reduced Row Echelon Form

HOW SHOULD WE FIND THE REDUCED ROW ECHELON FORM OF A MATRIX? The next algorithm establishes that every matrix is equivalent to a matrix in reduced row echelon form.

A description of this algorithm already appears in a Chinese mathematical text from 179 CE.

4.2.0 Algorithm (Row reduction)**.**
input: an $(m \times n)$-matrix $\mathbf{A} = [a_{j,k}]$.
output: an $(m \times n)$-matrix $\mathbf{B} = [b_{j,k}]$ in reduced row echelon form that is equivalent to $\mathbf{A}$.

Set $\mathbf{B} := \mathbf{A}$.	*initialize output matrix*
Set $r := 0$.	*initialize count of leading entries*
For k from 1 to n do	*loop over the columns*
If there exists j with $r < j \leqslant m$ and $b_{j,k} \neq 0$, then	*decide if there is another leading entry*
Set $r := r + 1$.	*increment count of leading entries*
Choose j with $r \leqslant j \leqslant m$ such that $b_{j,k} \neq 0$.	*select entry that will become leading one*
Multiply the j-th row in $\mathbf{B}$ by $b_{j,k}^{-1}$.	*make selected leading entry equal to* 1
Swap j-th and r-th rows in $\mathbf{B}$.	*position leading one in r-th row*
For ℓ from 1 to m with $\ell \neq r$ do	*looping over the other rows*
Multiply the r-th row by $-b_{\ell,k}$ and add it to the ℓ-th row in $\mathbf{B}$.	*create zero entries in k-th column*
Return the matrix $\mathbf{B}$.	

Correctness. The conditional "if" statement ensures that the first r rows are nonzero and the last $m - r$ rows are zero. The index k increases from 1 to n, so the chosen entry $b_{j,k}$ is the first nonzero entry in the j-th row. Since the index r counts the leading entries and $j > r$, the leading entry in the r-th row occurs to the right of the leading entry in the $(r-1)$-st row. The final loop, over ℓ, ensures that each leading one is the only nonzero entry its column. □

Although Algorithm 4.2.0 gives a uniform procedure for finding the reduced row echelon form a matrix, it can be more computational efficient to use a different sequence of elementary row operations to calculate the reduced row echelon form.

4.2.1 Corollary. *Every matrix is equivalent to a matrix in reduced row echelon form. Moreover, the reduced row echelon form can be obtained by a sequence of elementary operations.*

By identifying a linear system with its augmented matrix, we extend the notion of equivalent linear systems, namely those with the same solution set, to all matrices.

Proof. Apply the row reduction algorithm to a given matrix. □

4.2.2 Definition. The ***rank*** of a matrix $\mathbf{A}$ is the number r of nonzero rows in the reduced row echelon form. It is denoted $\operatorname{rank}(\mathbf{A}) = r$. When $\mathbf{A}$ is the coefficient matrix of a linear system, the r variables corresponding to the columns with leading ones are the ***leading variables*** and the complementary $n - r$ variables are the ***free variables***.

4.2.3 Proposition (Basic bounds on rank)**.** *The rank of an $(m \times n)$-matrix is at most the minimum of m and n.*

Proof. Let $\mathbf{A}$ be an $(m \times n)$-matrix. By definition, the rank of $\mathbf{A}$ is the number of nonzero rows in the reduced row echelon form, so we have $\operatorname{rank}(\mathbf{A}) = r \leqslant m$. On the other hand, if the leading entry in the j-th row lies in the k_j-th column, then the definition of reduced row echelon implies that $1 \leqslant k_1 < k_2 < \cdots < k_r \leqslant n$. We conclude that $r \leqslant n$ because there are at most n integers between 1 and n. □

4.2.4 Problem. Solve the linear system

$$\left\{\begin{array}{rrrrrl} x_1 &+ 2x_2 &- x_3 &+ 2x_4 &+ x_5 &= 2 \\ -x_1 &- 2x_2 &+ x_3 &+ 2x_4 &+ 3x_5 &= 6 \\ 2x_1 &+ 4x_2 &- 3x_3 &+ 2x_4 & &= 3 \\ -3x_1 &- 6x_2 &+ 2x_3 & &+ 3x_5 &= 9 \end{array}\right\}.$$

Solution. Applying the row reduction algorithm to the associated augmented matrix, we obtain

$$\begin{bmatrix} \underline{1} & 2 & -1 & 2 & 1 & 2 \\ -1 & -2 & 1 & 2 & 3 & 6 \\ 2 & 4 & -3 & 2 & 0 & 3 \\ -3 & -6 & 2 & 0 & 3 & 9 \end{bmatrix} \xrightarrow[\sim]{\vec{r}_2 + \vec{r}_1 \mapsto \vec{r}_2} \begin{bmatrix} \underline{1} & 2 & -1 & 2 & 1 & 2 \\ 0 & 0 & 0 & 4 & 4 & 8 \\ 2 & 4 & -3 & 2 & 0 & 3 \\ -3 & -6 & 2 & 0 & 3 & 9 \end{bmatrix} \xrightarrow[\sim]{\vec{r}_3 - 2\vec{r}_1 \mapsto \vec{r}_3} \begin{bmatrix} \underline{1} & 2 & -1 & 2 & 1 & 2 \\ 0 & 0 & 0 & 4 & 4 & 8 \\ 0 & 0 & -1 & -2 & -2 & -1 \\ -3 & -6 & 2 & 0 & 3 & 9 \end{bmatrix}$$

$$\xrightarrow[\sim]{\vec{r}_4 + 3\vec{r}_1 \mapsto \vec{r}_4} \begin{bmatrix} 1 & 2 & -1 & 2 & 1 & 2 \\ 0 & 0 & 0 & 4 & 4 & 8 \\ 0 & 0 & \underline{-1} & -2 & -2 & -1 \\ 0 & 0 & -1 & 6 & 6 & 15 \end{bmatrix} \xrightarrow[\sim]{-\vec{r}_3 \mapsto \vec{r}_3} \begin{bmatrix} 1 & 2 & -1 & 2 & 1 & 2 \\ 0 & 0 & 0 & 4 & 4 & 8 \\ 0 & 0 & \underline{1} & 2 & 2 & 1 \\ 0 & 0 & -1 & 6 & 6 & 15 \end{bmatrix} \xrightarrow[\sim]{\substack{\vec{r}_3 \mapsto \vec{r}_2 \\ \vec{r}_2 \mapsto \vec{r}_3}} \begin{bmatrix} 1 & 2 & -1 & 2 & 1 & 2 \\ 0 & 0 & \underline{1} & 2 & 2 & 1 \\ 0 & 0 & 0 & 4 & 4 & 8 \\ 0 & 0 & -1 & 6 & 6 & 15 \end{bmatrix}$$

$$\xrightarrow[\sim]{\vec{r}_1 + \vec{r}_2 \mapsto \vec{r}_1} \begin{bmatrix} 1 & 2 & 0 & 4 & 3 & 3 \\ 0 & 0 & \underline{1} & 2 & 2 & 1 \\ 0 & 0 & 0 & 4 & 4 & 8 \\ 0 & 0 & -1 & 6 & 6 & 15 \end{bmatrix} \xrightarrow[\sim]{\vec{r}_4 + \vec{r}_2 \mapsto \vec{r}_4} \begin{bmatrix} 1 & 2 & 0 & 4 & 3 & 3 \\ 0 & 0 & 1 & 2 & 2 & 1 \\ 0 & 0 & 0 & \underline{4} & 4 & 8 \\ 0 & 0 & 0 & 8 & 8 & 16 \end{bmatrix} \xrightarrow[\sim]{\frac{1}{4}\vec{r}_3 \mapsto \vec{r}_3} \begin{bmatrix} 1 & 2 & 0 & 4 & 3 & 3 \\ 0 & 0 & 1 & 2 & 2 & 1 \\ 0 & 0 & 0 & \underline{1} & 1 & 2 \\ 0 & 0 & 0 & 8 & 8 & 16 \end{bmatrix}$$

$$\xrightarrow[\sim]{\vec{r}_1 - 4\vec{r}_3 \mapsto \vec{r}_1} \begin{bmatrix} 1 & 2 & 0 & 0 & -1 & -5 \\ 0 & 0 & 1 & 2 & 2 & 1 \\ 0 & 0 & 0 & \underline{1} & 1 & 2 \\ 0 & 0 & 0 & 8 & 8 & 16 \end{bmatrix} \xrightarrow[\sim]{\vec{r}_2 - 2\vec{r}_3 \mapsto \vec{r}_2} \begin{bmatrix} 1 & 2 & 0 & 0 & -1 & -5 \\ 0 & 0 & 1 & 0 & 0 & -3 \\ 0 & 0 & 0 & \underline{1} & 1 & 2 \\ 0 & 0 & 0 & 8 & 8 & 16 \end{bmatrix} \xrightarrow[\sim]{\vec{r}_4 - 8\vec{r}_3 \mapsto \vec{r}_4} \begin{bmatrix} 1 & 2 & 0 & 0 & -1 & -5 \\ 0 & 0 & 1 & 0 & 0 & -3 \\ 0 & 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The reduced row echelon form corresponds to the linear system

$$\left\{\begin{array}{l} x_1 + 2x_2 \qquad - x_5 = -5 \\ \qquad\quad x_3 \qquad\quad = -3 \\ \qquad\qquad x_4 + x_5 = 2 \end{array}\right\} = \left\{\begin{array}{l} x_1 = -5 - 2x_2 + x_5 \\ x_3 = -3 \\ x_4 = 2 - x_5 \end{array}\right\},$$

so the solution set is $\left\{ \begin{bmatrix} -5-2\,t_2+t_5 \\ t_2 \\ -3 \\ 2-t_5 \\ t_5 \end{bmatrix} \;\middle|\; t_2, t_5 \in \mathbb{K} \right\}$. □

Exercises

4.2.5 Problem. Determine which of the following statements are true. If a statement is false, then provide a counterexample.

i. The row reduction algorithm finds the reduced row echelon form of any matrix.
ii. The rank of a matrix equal the number of nonzero rows in the matrix.
iii. The rank of a matrix can be determined by applying the Row Reduction Algorithm.
iv. The rank of the zero matrix is zero.
v. The rank of a matrix is at most the number of rows in the matrix.
vi. The rank of a matrix is at most the number of columns in the matrix.

4.2.6 Problem. If the reduced row echelon form of the augmented matrix for some linear system is

$$\begin{bmatrix} 1 & -2 & 0 & 1 & 4 & 5 \\ 0 & 0 & 1 & 3 & 0 & -7 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

then find the solution set.

4.2.7 Problem. Solve the following system of linear equations:

$$\left\{ \begin{array}{rrrrrrl} & 2x_2 & -\,4x_3 & -\,5x_4 & +\;\;2x_5 & +\,5x_6 & = 0 \\ x_1 & -\;\,x_2 & +\;\;x_3 & +\,3x_4 & +\;\;\;x_5 & -\;\;x_6 & = 0 \\ 6x_1 & & -\,6x_3 & +\,5x_4 & +\,16x_5 & +\,7x_6 & = 0 \end{array} \right\}.$$

4.2.8 Problem. For what values of $k, \ell \in \mathbb{R}$, does the linear system associated to the augmented matrix

$$\begin{bmatrix} 1 & -1 & -4 & -2 \\ 1 & 0 & k-4 & \ell \\ 2 & k & 2(k+4) & 2k \end{bmatrix}$$

have

(i) exactly one solution,
(ii) no solutions, or
(iii) infinitely many solutions?

In case **(iii)**, describe the solution set.

4.2.9 Problem. Determine if the linear systems with the following augmented matrices are equivalent:

$$\mathbf{A} := \begin{bmatrix} 6 & 1 & -16 & -2 & 2 \\ -27 & 5 & 91 & -8 & -98 \\ 21 & -4 & -71 & 6 & 76 \end{bmatrix}, \quad \mathbf{B} := \begin{bmatrix} 9 & -2 & -31 & 2 & 32 \\ -3 & 20 & 49 & -22 & -152 \\ 3 & 39 & 69 & -43 & -279 \end{bmatrix}.$$

5
Spanning Sets

The solution set to a linear system inherits features from the ambient coordinate space. To understand this structure, we define the span of a set of vectors, we introduce the fundamental circuits, and we develop the basic arithmetic of matrices.

5.0 Vector Equations

How can we efficiently describe an infinite solution set? We merge addition and scalar multiplication into one concept.

5.0.0 Definition. Let m and n be positive integers. For any vectors $\vec{\boldsymbol{a}}_1, \vec{\boldsymbol{a}}_2, \dots, \vec{\boldsymbol{a}}_n$ in $\mathbb{K}^m$, a ***linear combination*** is any vector of the form $\vec{\boldsymbol{b}} := c_1\,\vec{\boldsymbol{a}}_1 + c_2\,\vec{\boldsymbol{a}}_2 + \cdots + c_n\,\vec{\boldsymbol{a}}_n$ where the coefficients $c_1, c_2, \dots, c_n$ are scalars in $\mathbb{K}$.

The ***vector equation***
$$x_1\,\vec{\boldsymbol{a}}_1 + x_2\,\vec{\boldsymbol{a}}_2 + \cdots + x_n\,\vec{\boldsymbol{a}}_n = \vec{\boldsymbol{b}}$$
is an alternative notation for the linear system whose augmented matrix is $\begin{bmatrix} \vec{\boldsymbol{a}}_1 & \vec{\boldsymbol{a}}_2 & \cdots & \vec{\boldsymbol{a}}_n & \vec{\boldsymbol{b}} \end{bmatrix}$.

5.0.1 Problem. Determine if the vector $\vec{\boldsymbol{b}} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$ a linear combination of the vectors $\vec{\boldsymbol{a}}_1 := \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix}$ and $\vec{\boldsymbol{a}}_2 := \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}$.

Solution. We are asked to decide if there exists scalars $x_1, x_2 \in \mathbb{K}$ such that $x_1\,\vec{\boldsymbol{a}}_1 + x_2\,\vec{\boldsymbol{a}}_2 = \vec{\boldsymbol{b}}$. In other words, we must solve

$$\begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 \\ -2x_1 + 5x_2 \\ -5x_1 + 5x_2 \end{bmatrix} \quad \text{or} \quad \left\{ \begin{aligned} x_1 + 2x_2 &= 7 \\ -2x_1 + 5x_2 &= 4 \\ -5x_1 + 6x_2 &= 3 \end{aligned} \right\}.$$

The row reduction algorithm [4.2.0] gives

$$\begin{bmatrix} \underline{1} & 2 & 7 \\ -2 & 5 & 4 \\ -5 & 6 & -3 \end{bmatrix} \xrightarrow[\sim]{\substack{\vec{r}_2 + 2\vec{r}_1 \mapsto \vec{r}_2 \\ \vec{r}_3 + 5\vec{r}_1 \mapsto \vec{r}_3}} \begin{bmatrix} \underline{1} & 2 & 7 \\ 0 & 9 & 18 \\ 0 & 16 & 32 \end{bmatrix} \xrightarrow[\sim]{\substack{9^{-1}\vec{r}_2 \mapsto \vec{r}_2 \\ 16^{-1}\vec{r}_3 \mapsto \vec{r}_3}} \begin{bmatrix} 1 & 2 & 7 \\ 0 & \underline{1} & 2 \\ 0 & 1 & 2 \end{bmatrix} \xrightarrow[\sim]{\substack{\vec{r}_1 - 2\vec{r}_2 \mapsto \vec{r}_1 \\ \vec{r}_3 - \vec{r}_2 \mapsto \vec{r}_3}} \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

Hence, we have $\vec{\boldsymbol{b}} = 3\,\vec{\boldsymbol{a}}_1 + 2\,\vec{\boldsymbol{a}}_2$. □

Linear combinations create the most important subsets of the ambient coordinate space.

5.0.2 Definition. The ***span*** of the vectors $\vec{\boldsymbol{a}}_1, \vec{\boldsymbol{a}}_2, \dots, \vec{\boldsymbol{a}}_n \in \mathbb{K}^m$ is the set of all their linear combinations. We write

$$\operatorname{Span}(\vec{\boldsymbol{a}}_1, \vec{\boldsymbol{a}}_2, \dots, \vec{\boldsymbol{a}}_n) := \{ c_1\,\vec{\boldsymbol{a}}_1 + c_2\,\vec{\boldsymbol{a}}_2 + \cdots + c_n\,\vec{\boldsymbol{a}}_n \in \mathbb{K}^m \mid c_1, c_2, \dots, c_n \in \mathbb{K} \}.$$

The membership $\vec{b} \in \operatorname{Span}(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n)$ is equivalent to the linear system with the augmented matrix $\begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n & \vec{b} \end{bmatrix}$ being consistent or to the vector equation $x_1\,\vec{a}_1 + x_2\,\vec{a}_2 + \cdots + x_n\,\vec{a}_n = \vec{b}$ having at least one solution.

The span of one or two vectors is simply a line or plane respectively. When $\vec{0} \neq \vec{a} \in \mathbb{K}^m$, the set $\operatorname{Span}(\vec{a})$ consists of all scalar multiples of $\vec{a}$. Geometrically, this is the line in $\mathbb{K}^m$ passing through the origin and the point with position vector $\vec{a}$ relative to the origin. Similarly, if $\vec{a}_1$ and $\vec{a}_2$ are nonzero nonparallel vectors in $\mathbb{K}^3$, then $\operatorname{Span}(\vec{a}_1, \vec{a}_2)$ is the plane in $\mathbb{K}^3$ that contains the origin and the points with position vectors $\vec{a}_1, \vec{a}_2$ relative to the origin.

Since we have
$$x_1\,\vec{e}_1 + x_2\,\vec{e}_2 + \cdots + x_m\,\vec{e}_m = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$
the standard basis vectors $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_m$ span $\mathbb{K}^m$ or equivalently we have $\operatorname{Span}(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_m) = \mathbb{K}^m$.

5.0.3 Problem. Show that $\vec{a}_1 := \left[\begin{smallmatrix} 2 \\ 3 \end{smallmatrix}\right]$ and $\vec{a}_2 := \left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right]$ span $\mathbb{K}^2$.

Solution. We need to show that an arbitrary vector $\vec{b} = \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in \mathbb{K}^2$ is a linear combination of $\vec{a}_1$ and $\vec{a}_2$. Since we have

$$\begin{bmatrix} 2 & 1 & x \\ 3 & 2 & y \end{bmatrix} \xrightarrow[\sim]{\vec{r}_2 - \vec{r}_1 \mapsto \vec{r}_2} \begin{bmatrix} 2 & 1 & x \\ 1 & 1 & y-x \end{bmatrix} \xrightarrow[\sim]{\vec{r}_1 - 2\,\vec{r}_2 \mapsto \vec{r}_1} \begin{bmatrix} 0 & -1 & 3x-2y \\ 1 & 1 & y-x \end{bmatrix} \xrightarrow[\sim]{\substack{\vec{r}_1 \mapsto \vec{r}_2 \\ \vec{r}_2 \mapsto \vec{r}_1}} \begin{bmatrix} 1 & 1 & y-x \\ 0 & -1 & 3x-2y \end{bmatrix}$$
$$\xrightarrow[\sim]{\vec{r}_1 + \vec{r}_2 \mapsto \vec{r}_1} \begin{bmatrix} 1 & 0 & 2x-y \\ 0 & -1 & 3x-2y \end{bmatrix} \xrightarrow[\sim]{-\vec{r}_2 \mapsto \vec{r}_2} \begin{bmatrix} 1 & 0 & 2x-y \\ 0 & 1 & 2y-3x \end{bmatrix},$$

it follows that $(2x-y)\,\vec{a}_1 + (-3x+2y)\,\vec{a}_2 = \vec{b}$ and we deduce that $\operatorname{Span}(\vec{a}_1, \vec{a}_2) = \mathbb{K}^2$. □

Linear combinations also lead to a new matrix operation.

5.0.4 Definition. Let $\mathbf{A}$ be an $(m \times n)$-matrix whose columns are the vectors $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n \in \mathbb{K}^m$. For any vector $\vec{x} \in \mathbb{K}^n$, the ***product*** of the matrix $\mathbf{A}$ and the vector $\vec{x}$, denoted $\mathbf{A}\,\vec{x}$, is the linear combination of the columns of $\mathbf{A}$ with the corresponding entries of x as coefficients:

$$\mathbf{A}\,\vec{x} = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} := x_1\,\vec{a}_1 + x_2\,\vec{a}_2 + \cdots + x_n\,\vec{a}_n\,.$$

This definition requires that the number of columns in the matrix to equal the number of entries in the vector.

For instance, we have

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix} = 4 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 2 \\ -5 \end{bmatrix} + 7 \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \end{bmatrix} + \begin{bmatrix} 6 \\ -15 \end{bmatrix} + \begin{bmatrix} -7 \\ 21 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}.$$

An operation that is compatible with both addition and scalar multiplication is called a ***linear*** operation.

5.0.5 Proposition (Linearity of matrix multiplication)**.** *Let m and n be positive integers, let c and d be scalars in $\mathbb{K}$, and let $\vec{v}$ and $\vec{w}$ be vectors in $\mathbb{K}^n$. For any $(m \times n)$-matrix $\mathbf{A}$, we have $\mathbf{A}(c\,\vec{v} + d\,\vec{w}) = c(\mathbf{A}\vec{v}) + d(\mathbf{A}\vec{w})$.*

Proof. Let $\vec{\boldsymbol{a}}_1, \vec{\boldsymbol{a}}_2, \dots, \vec{\boldsymbol{a}}_n \in \mathbb{K}^m$ denote the columns of the matrix $\mathbf{A}$. The definition of matrix multiplication, the entrywise arithmetic of vectors, and the properties of a scalar multiplication give

$$\begin{aligned}\mathbf{A}(c\,\vec{\boldsymbol{u}} + d\,\vec{\boldsymbol{v}}) &= [\vec{\boldsymbol{a}}_1 \;\; \vec{\boldsymbol{a}}_2 \;\; \cdots \;\; \vec{\boldsymbol{a}}_n] \begin{bmatrix} cv_1 + dw_1 \\ cv_2 + dw_2 \\ \vdots \\ cv_n + dw_n \end{bmatrix} \\ &= (cv_1 + dw_1)\,\vec{\boldsymbol{a}}_1 + (cv_2 + dw_2)\,\vec{\boldsymbol{a}}_2 + \cdots + (cv_n + dw_n)\,\vec{\boldsymbol{a}}_n \\ &= c(v_1\,\vec{\boldsymbol{a}}_1 + v_2\,\vec{\boldsymbol{a}}_2 + \cdots + v_n\,\vec{\boldsymbol{a}}_n) + d(w_1\,\vec{\boldsymbol{a}}_1 + w_2\,\vec{\boldsymbol{a}}_2 + \cdots + w_n\,\vec{\boldsymbol{a}}_n) \\ &= c(\mathbf{A}\vec{\boldsymbol{v}}) + d(\mathbf{A}\vec{\boldsymbol{w}})\,. \qquad \square\end{aligned}$$

Exercises

5.0.6 Problem. Determine which of the following statements are true. If a statement is false, then provide a counterexample.

i. A linear combination of vectors is a scalar.

ii. A linear combination of a single vector is a scalar multiple of the vector.

iii. The zero vector is a linear combination of any subset of vectors.

iv. Every linear system can be expressed as a vector equation.

v. The span of a collection of vectors is a vector.

vi. The product of a matrix and a vector is a matrix.

vii. The product of a matrix and a vector is a vector.

viii. For the product of a matrix and a vector to be well-defined, the number of entries in the vector must equal the number of rows in the matrix.

ix. For an $(m \times n)$-matrix $\mathbf{A}$, a vector $\vec{\boldsymbol{v}} \in \mathbb{K}^n$, and a scalar $c \in \mathbb{K}$, we always have $\mathbf{A}(c\,\vec{\boldsymbol{v}}) = c(\mathbf{A}\vec{\boldsymbol{v}})$.

x. For an $(m \times n)$-matrix $\mathbf{A}$ and vectors $\vec{\boldsymbol{v}}, \vec{\boldsymbol{w}} \in \mathbb{K}^n$, we always have $\mathbf{A}(\vec{\boldsymbol{v}} + \vec{\boldsymbol{w}}) = (\mathbf{A}\vec{\boldsymbol{v}}) + (\mathbf{A}\vec{\boldsymbol{w}})$.

5.0.7 Problem. Solve the linear system $\left\{\begin{array}{rcrcrcr} & & y & + & 4z & = & -5 \\ x & + & 3y & + & 5z & = & -2 \\ 3x & + & 7y & + & 7z & = & 6 \end{array}\right\}$.

5.0.8 Problem. Consider an arbitrary system of linear equations $\mathbf{A}\,\vec{\boldsymbol{x}} = \vec{\boldsymbol{b}}$ where the coefficient matrix $\mathbf{A}$ and constant term vector $\vec{\boldsymbol{b}}$ have all real entries.

(i) If $\mathbf{A}\,\vec{\boldsymbol{x}} = \vec{\boldsymbol{b}}$ has more than one solution, then prove that it has infinitely many solutions.

(ii) If there is a solution with entries in the complex numbers, then show that there is also a solution with entries in the real numbers.

5.0.9 Problem. Solve the linear system

$$\left\{\begin{array}{rcrcrcl} \mathrm{i}\,z_1 & + & (1+\mathrm{i})\,z_2 & & & = & \mathrm{i} \\ (1-\mathrm{i})\,z_1 & + & z_2 & - & \mathrm{i}\,z_3 & = & 1 \\ & & \mathrm{i}\,z_2 & + & z_3 & = & 1 \end{array}\right\}.$$

5.0.10 Problem. Does $\vec{v} := \begin{bmatrix} 1 - \mathrm{i} \\ 1 + \mathrm{i} \\ 1 - 2\mathrm{i} \end{bmatrix}$ belong to $\mathrm{Span}(\vec{w}_1, \vec{w}_2, \vec{w}_3)$ where

$$\vec{w}_1 := \begin{bmatrix} 0 \\ -\mathrm{i} \\ -1 \end{bmatrix}, \qquad \vec{w}_2 := \begin{bmatrix} -\mathrm{i} \\ 1+\mathrm{i} \\ -\mathrm{i} \end{bmatrix}, \qquad \vec{w}_3 := \begin{bmatrix} 2 \\ -2+2\mathrm{i} \\ 3 \end{bmatrix}.$$

5.1 Kernel of a Matrix

WHAT DOES MATRIX MULTIPLICATION MAP TO THE ZERO VECTOR? Understanding the collection of all vectors that become to the zero vector when multiplied by a matrix is surprisingly useful.

5.1.0 Definition. The ***kernel*** of an $(m \times n)$-matrix $\mathbf{A}$ is set of all vectors $\mathbb{K}^n$ such that their product with $\mathbf{A}$ equals the zero vector;

$$\mathrm{Ker}(\mathbf{A}) := \left\{ \vec{x} \in \mathbb{K}^n \;\middle|\; \mathbf{A}\,\vec{x} = \vec{\mathbf{0}} \right\}.$$

Some introductory textbooks refer to this set as the 'null space'.

The word "kernel" comes from a diminutive of "corn seed", but the mathematical usage follows the more general meaning of the nucleus of a structure or a core.

5.1.1 Lemma. *Any linear combination of vectors lying in the kernel of a matrix also belongs to the kernel.*

Proof. Let $\mathbf{A}$ be an $(m \times n)$-matrix. For any vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k \in \mathbb{K}^n$ lying in the kernel of the matrix $\mathbf{A}$ and any scalars $c_1, c_2, \dots, c_k \in \mathbb{K}$, the linearity of matrix multiplication [5.0.5] gives

$$\begin{aligned} \mathbf{A}(c_1\,\vec{v}_1 + c_2\,\vec{v}_2 + \cdots + c_k\,\vec{v}_k) &= c_1(\mathbf{A}\,\vec{v}_1) + c_2(\mathbf{A}\,\vec{v}_2) + \cdots + c_k(\mathbf{A}\,\vec{v}_k) \\ &= c_1\,\vec{\mathbf{0}} + c_2\,\vec{\mathbf{0}} + \cdots + c_k\,\vec{\mathbf{0}} = \vec{\mathbf{0}}. \end{aligned}$$

so we conclude that $c_1\,\vec{v}_1 + c_2\,\vec{v}_2 + \cdots + c_k\,\vec{v}_k \in \mathrm{Ker}(\mathbf{A})$. ☐

5.1.2 Problem. Solve the linear system $\mathbf{A}\,\vec{x} = \vec{\mathbf{0}}$ where

$$\mathbf{A} := \begin{bmatrix} 0 & \underline{1} & -2 & 0 & 1 & 5 & 0 \\ 0 & 0 & 0 & \underline{1} & 3 & -7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \underline{1} \end{bmatrix}.$$

Solution. Since the matrix $\mathbf{A}$ is in reduced row echelon form, we obtain the general solution by isolating the leading variables:

$$\left\{ \begin{aligned} x_2 - 2x_3 + x_5 + 5x_6 &= 0 \\ x_4 + 3x_5 - 7x_6 &= 0 \\ x_7 &= 0 \end{aligned} \right\} = \left\{ \begin{aligned} x_2 &= 2x_3 - x_5 - 5x_6 \\ x_4 &= -3x_5 + 7x_6 \\ x_7 &= 0 \end{aligned} \right\}.$$

It follows that $\mathrm{Ker}(\mathbf{A})$ is the set

$$\left\{ \begin{bmatrix} x_1 \\ 2x_3 - x_5 - 5x_6 \\ x_3 \\ -3x_5 + 7x_6 \\ x_5 \\ x_6 \\ 0 \end{bmatrix} \;\middle|\; x_1, x_3, x_5, x_6 \in \mathbb{K} \right\} = \left\{ x_1 \begin{bmatrix} \underline{1} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 2 \\ \underline{1} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 0 \\ -1 \\ 0 \\ -3 \\ \underline{1} \\ 0 \\ 0 \end{bmatrix} + x_6 \begin{bmatrix} 0 \\ -5 \\ 0 \\ 7 \\ 0 \\ \underline{1} \\ 0 \end{bmatrix} \;\middle|\; x_1, x_2, x_5, x_6 \in \mathbb{K} \right\},$$

or equivalently

$$\operatorname{Ker}(\mathbf{A}) = \operatorname{Span}\left(\begin{bmatrix} \underline{1} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ \underline{1} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 0 \\ -3 \\ \underline{1} \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -5 \\ 0 \\ 7 \\ 0 \\ \underline{1} \\ 0 \end{bmatrix} \right). \qquad \square$$

To go directly from the matrix in reduced row echelon form to the vectors that span the kernel, we need some notation and terminology.

5.1.3 Definition. For all $1 \leqslant j \leqslant m$ and all $1 \leqslant k \leqslant n$, let $a_{j,k} \in \mathbb{K}$ denote the (j,k)-entry of an $(m \times n)$-matrix $\mathbf{A}$. Assume that the matrix $\mathbf{A}$ is in reduced row echelon form and set $r := \operatorname{rank}(\mathbf{A})$. For all $1 \leqslant j \leqslant r$, the leading entry in the j-th row lies in the k_j-th column and $1 \leqslant k_1 < k_2 < \cdots < k_r \leqslant n$. When the ℓ-th column of $\mathbf{A}$ does not contain a leading one, the associated ***fundamental circuit*** is the vector

$$\vec{\boldsymbol{e}}_\ell - a_{1,\ell}\, \vec{\boldsymbol{e}}_{k_1} - a_{2,\ell}\, \vec{\boldsymbol{e}}_{k_2} - \cdots - a_{r,\ell}\, \vec{\boldsymbol{e}}_{k_r} \in \mathbb{K}^n .$$

The number of nonzero entries is at most $r+1$. When the number of columns containing leading ones to the left of ℓ-th column is i, the fundamental circuit equals $\vec{\boldsymbol{e}}_\ell - a_{1,\ell}\, \vec{\boldsymbol{e}}_{k_1} - a_{2,\ell}\, \vec{\boldsymbol{e}}_{k_2} - \cdots - a_{i,\ell}\, \vec{\boldsymbol{e}}_{k_i}$ because $a_{i+1,\ell} = a_{i+2,\ell} = \cdots = a_{m,\ell} = 0$. Since each column not containing a leading one produces a fundamental circuit, the matrix $\mathbf{A}$ has $n - r$ distinct fundamental circuits.

We may visualize the reduced row echelon form as

$$\begin{array}{cc} & \begin{matrix} & k_1 & & k_i & & \ell & & k_{i+1} & \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ \\ i \\ i{+}1 \\ i{+}2 \\ \\ m \end{matrix} & \begin{bmatrix} \cdots & \underline{1} & \cdots & 0 & \cdots & a_{1,\ell} & \cdots & 0 & \cdots \\ \cdots & 0 & \cdots & 0 & \cdots & a_{2,\ell} & \cdots & 0 & \cdots \\ & \vdots & & \vdots & & \vdots & & \vdots & \\ \cdots & 0 & \cdots & \underline{1} & \cdots & a_{i,\ell} & \cdots & 0 & \cdots \\ \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots & \underline{1} & \cdots \\ \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots \\ & \vdots & & \vdots & & \vdots & & \vdots & \\ \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots \end{bmatrix} \end{array}$$

5.1.4 Proposition. *The kernel of a matrix is spanned by the fundamental circuits associated to its reduced row echelon form.*

Proof. Let $\mathbf{A}$ be an $(m \times n)$-matrix. Since elementary row operations produce equivalent linear systems, we may assume without loss of generality that $\mathbf{A}$ is in reduced row echelon form. To demonstrate the equality between the kernel of $\mathbf{A}$ and the span of the fundamental circuits, we prove containment in both directions.

For any two set $\mathscr{A}$ and $\mathscr{B}$, the following conditions are equivalent:
a. for all $a \in \mathscr{A}$, we have $a \in \mathscr{B}$,
b. $\mathscr{A}$ is a subset of $\mathscr{B}$,
c. $\mathscr{B}$ is a superset of $\mathscr{A}$,
d. $\mathscr{A} \cap \mathscr{B} = \mathscr{A}$, and
e. $\mathscr{A} \cup \mathscr{B} = \mathscr{B}$.
When $\mathcal{A}$ is a subset of $\mathcal{B}$, we write $\mathcal{A} \subseteq \mathcal{B}$. Similarly, when $\mathcal{B}$ is a superset of $\mathcal{A}$, we write $\mathcal{B} \supseteq \mathcal{A}$. By design, the 'subset' and 'superset' symbols remind us of '$\leqslant$' and '$\geqslant$' respectively.

$\supseteq$: Applying Lemma 5.1.1, it suffices to show that every fundamental circuit belongs to $\operatorname{Ker}(\mathbf{A})$. Assuming that it does not contain a leading one, focus on the ℓ-th column. If the number of leading ones to left of the ℓ-th column is i, then we have $0 \leqslant i \leqslant r$ and the definition of reduced row echelon form implies that $a_{j,\ell} = 0$ for all $j > i$. The fundamental circuit associated to the ℓ-th column is the vector $\vec{\boldsymbol{v}} := \vec{\boldsymbol{e}}_\ell - a_{1,\ell}\, \vec{\boldsymbol{e}}_{k_1} - a_{2,\ell}\, \vec{\boldsymbol{e}}_{k_2} - \cdots - a_{i,\ell}\, \vec{\boldsymbol{e}}_{k_i}$, so the definition of matrix multiplication gives

$$\mathbf{A}\, \vec{\boldsymbol{v}} = 1 \begin{bmatrix} a_{1,\ell} \\ a_{2,\ell} \\ \vdots \\ a_{i,\ell} \\ 0 \\ \vdots \\ 0 \end{bmatrix} - a_{1,\ell} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} - a_{2,\ell} \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} - \cdots - a_{i,\ell} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{1,\ell} - a_{1,\ell} \\ a_{2,\ell} - a_{2,\ell} \\ \vdots \\ a_{i,\ell} - a_{i,\ell} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \vec{\mathbf{0}} .$$

$\subseteq$: It remains to show that any vector in the kernel of $\mathbf{A}$ is a linear combination of the fundamental circuits. To accomplish this, let $\ell_1, \ell_2, \dots, \ell_{n-r}$ denote the columns of $\mathbf{A}$ that do not contain a leading one and let $\vec{\boldsymbol{v}}_1, \vec{\boldsymbol{v}}_2, \dots, \vec{\boldsymbol{v}}_{n-r}$ be the associated fundamental circuits. Consider an arbitrary vector $\vec{\boldsymbol{w}}$ in $\operatorname{Ker}(\mathbf{A})$ and the linear combination $\vec{\boldsymbol{u}} := \vec{\boldsymbol{w}} - w_{\ell_1}\,\vec{\boldsymbol{v}}_1 - w_{\ell_2}\,\vec{\boldsymbol{v}}_2 - \cdots - w_{\ell_{n-r}}\,\vec{\boldsymbol{v}}_{n-r}$. Since the first inclusion already establishes that any fundamental circuit lies in the kernel, Lemma 5.1.1 implies that $\vec{\boldsymbol{u}} \in \operatorname{Ker}(\mathbf{A})$. By definition, we have $\vec{\boldsymbol{v}}_i = \vec{\boldsymbol{e}}_{\ell_i} - a_{1,\ell_i}\,\vec{\boldsymbol{e}}_{k_1} - a_{2,\ell_i}\,\vec{\boldsymbol{e}}_{k_2} - \cdots - a_{r,\ell_i}\,\vec{\boldsymbol{e}}_{k_r}$, so we obtain

$$\begin{aligned}
\vec{\boldsymbol{u}} &= \vec{\boldsymbol{w}} - w_{\ell_1}\vec{\boldsymbol{v}}_1 - w_{\ell_2}\vec{\boldsymbol{v}}_2 - \cdots - w_{\ell_{n-r}}\vec{\boldsymbol{v}}_{n-r} \\
&= (w_1\,\vec{\boldsymbol{e}}_1 + w_2\,\vec{\boldsymbol{e}}_2 + \cdots + w_m\,\vec{\boldsymbol{e}}_m) - \sum_{i=1}^{n-r} w_{\ell_i} \left(\vec{\boldsymbol{e}}_{\ell_i} - \sum_{j=1}^{r} a_{j,\ell_i}\,\vec{\boldsymbol{e}}_{k_j} \right) \\
&= \sum_{i=1}^{n-r} (w_{\ell_i} - w_{\ell_i})\vec{\boldsymbol{e}}_{\ell_i} + \sum_{j=1}^{r} \left(w_{k_j} + \sum_{i=1}^{n-r} (w_{\ell_i}\,a_{j,\ell_i}) \right) \vec{\boldsymbol{e}}_{k_j} \\
&= c_1\,\vec{\boldsymbol{e}}_{k_1} + c_2\,\vec{\boldsymbol{e}}_{k_2} + \cdots + c_r\,\vec{\boldsymbol{e}}_{k_r}
\end{aligned}$$

where $c_j := w_{k_j} + \sum_{i=1}^{n-r} w_{\ell_i}\,a_{j,\ell_i} \in \mathbb{K}$ for all $1 \leqslant j \leqslant r$. Thus, we have

$$\vec{\boldsymbol{0}} = \mathbf{A}\,\vec{\boldsymbol{u}} = c_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \cdots + c_r \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_r \\ 0 \\ \vdots \\ 0 \end{bmatrix} .$$

from which we deduce that $c_1 = c_2 = \cdots = c_r = 0$, $\vec{\boldsymbol{u}} = \vec{\boldsymbol{0}}$, and $\vec{\boldsymbol{w}} = w_{\ell_1}\vec{\boldsymbol{v}}_1 + w_{\ell_2}\vec{\boldsymbol{v}}_2 + \cdots + w_{\ell_{n-r}}\vec{\boldsymbol{v}}_{n-r}$. □

Exercises

5.1.5 Problem. Determine which of the following statements are true. If a statement is false, then provide a counterexample.

i. The zero vector always belongs to the kernel of a matrix.
ii. Every vector belongs to the kernel of the zero matrix.
iii. If a vector belongs to the kernel of a matrix, then all its scalar multiples also belong to the kernel.
iv. There is a circuit vector associated to every column of a matrix in reduced row echelon form.
v. A circuit vector also has at least two nonzero entries.