Introduction to Linear Algebra

Bruno Benedetti
University of Miami

Fall 2024

Abstract

Some notes for the first lectures of my Linear Algebra course.

Contents

0 Preliminaries: Rings and fields

You are probably all familiar with the infinite set of *natural numbers* (aka "nonnegative integers")

$$\mathbb{N} = \{0, 1, 2, 3, \ldots, n, n+1, \ldots\}$$

with the set of *integers*

$$\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3, \ldots, n, -n, n+1, -n-1, \ldots\}$$

and with the set of *rational numbers* or *fractions*

$$\mathbb{Q} = \left\{ \frac{a}{b} \quad : \quad a \in \mathbb{Z}, \quad b \in \mathbb{N} \setminus \{0\} \right\},$$

where we consider two fractions $\frac{a}{b}$, $\frac{a'}{b'}$ identical if $ab' = a'b$. For example, $\frac{1}{2}$, $\frac{-1}{-2}$, and $\frac{3}{6}$ are the same. Addition and multiplication are defined by

$$\frac{a}{b} + \frac{c}{d} \stackrel{\text{def}}{=} \frac{ad + cb}{bd} \text{ and } \frac{a}{b}\,\frac{c}{d} \stackrel{\text{def}}{=} \frac{ac}{bd}.$$

You might also know larger sets $\mathbb{R}, \mathbb{C}$ from Calculus. Here is a more general framework:

0.1 Rings

Definition 1 (Ring)**.** A **ring with 1**, throughout called simply a *ring*, consists of a set R endowed with two internal operations $+$ and $\cdot$ that satisfy the following axioms:

(R1) The operation $+$ is *associative*. That is, for all x, y, z in R, $x + (y + z) = (x + y) + z$.

(R2) The operation $+$ is *commutative*. That is, for all x, y, z in R, $x + y = y + x$.

(R3) The operation $+$ has a unique *neutral element*. That is, there exists an element z in R such that for all x in R, $x + z = x$. From now on we denote such element by "0".

(R4) Every element has a unique *additive inverse*. That is, for all x in R there exists exactly one element y in A such that $x + y = 0$. From now on we denote such element by "$-x$".

(R5) The operation $\cdot$ is *associative*. That is, for all x, y, z in R, $x \cdot (y \cdot z) = (x \cdot y) \cdot z$.

(R6) The operation $\cdot$ has a unique *neutral element*, different than 0. That is, there exists a unique element $z \neq 0$ in R such that for all x in R, $xz = x$. From now on we denote such neutral element by "1".

(R7) The operation $\cdot$ *distributes* $+$: for all x, y, z in R, $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$.

Example 2. $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ are C-rings with 1. $\mathbb{N}$ is not a ring: additive inverses are not included.

Example 3. The set $\{\text{True}, \text{False}\}$ is a C-ring with the logical operations of XOR ("exclusive or": either-or, bot not both) and AND.

Notation. We write $a - b$ as a shortening of $a + (-b)$. Moreover, we usually write xy instead of $x \cdot y$. Note also that by associativity, it is not ambiguous to write $abcd$ instead of $a(b(cd))$ or of $(ab)(cd)$. In fact, no matter how you insert brackets, the result is always the same.

Remark 4. Some textbooks rephrase axiom (R3) as "The operation + has a neutral element". Uniqueness is anyway necessary: Were there two neutral elements z_1 and z_2, we would have $z_1 + z_2 = z_1$ (because z_2 is neutral) yet also $z_1 + z_2 = z_2$ (being z_1 neutral), so $z_1 = z_2$. Similarly for (R6): If we only knew that multiplicative inverses exist, then their uniqueness would follow automatically, since for neutral elements z_1, z_2 we would have simultaneously that $z_1 z_2 = z_1$ and $z_1 z_2 = z_2$. Finally, some textbooks rephrase axiom (R4) as "Every element has an *additive inverse*". Also in this case, uniqueness is implicit: Were there elements y_1, y_2 in A such that $x + y_1 = 0 = x + y_2$, then we would have $y_1 = y_1 + 0 = y_1 + (x + y_2) = (y_1 + x) + y_2 = 0 + y_2 = y_2$.

Note that in the definition of ring we are not demanding the commutativity of the $\cdot$ operation. We will see later that matrices form a ring where $ab \neq ba$ in general. We are also not demanding the existence of 'multiplicative inverses'. $\mathbb{Z}$ is a ring where multiplicative inverses are never included, except for 1 and -1.

Definition 5. Any ring (with 1) R is called *commutative* if it satisfies the extra axiom

(R8) The operation $\cdot$ is *commutative*. That is, for all x, y, z in R, $x \cdot y = y \cdot x$.

Definition 6. Let R be a ring (with 1). An element x in R is called *invertible in R* if there exists an element y in R such that $xy = yx = 1$.

For example, 1 is always invertible, because $1 \cdot 1 = 1$. Note that being invertible depends on the ring chosen: 3 is invertible in $\mathbb{Q}$, but it is not invertible in $\mathbb{Z}$.

Proposition 7. *Let R be a C-ring. For all a in R, $a \cdot 0 = 0 = 0 \cdot a$.*

Proof. Being 0 neutral element, $0 = 0 + 0$, and $a \cdot 0 + 0 = a \cdot 0$. So

$$a \cdot 0 + 0 = a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0,$$

where in the last step we used distributivity. So adding $-a \cdot 0$ to both sides, we get $0 = a \cdot 0$. Analogously one shows $0 = 0 \cdot a$. □

Corollary 8. *The element 0 is never invertible.*

Proof. By axiom (R6), the element we called 0 and the element we called 1 are different. Since $0 \cdot a$ is always 0, it can never be 1. □

Proposition 9. *Let x be an element of a ring R. Suppose x is "right invertible" (i.e. there exists an element y_1 in A such that $xy_1 = 1$) and "left invertible" (i.e. there exists an element y_2 in R such that $y_2 x = 1$). Then x is invertible, and $y_1 = y_2$.*

Proof. One has $y_2 = y_2(xy_1) = (y_2 x)y_1 = y_1$. □

Notation. In consequence of the previous proposition, any invertible element x has exactly one inverse, which we shall denote by x^{-1}.

Proposition 10. *Let x, y be elements of a ring R. If x, y are invertible, so is their product, as*

$$(xy)^{-1} = y^{-1}x^{-1}.$$

Proof. Clearly $(xy)(y^{-1}x^{-1}) = x(yy^{-1})x^{-1} = xx^{-1} = 1$ (make sure you get all steps!), and similarly $(y^{-1}x^{-1})(xy) = 1$. □

Corollary 11. *Let x be an element of a ring A. If x is invertible, then so is x^n, as*

$$(x^n)^{-1} = \left(x^{-1}\right)^n.$$

Proof. Exercise: Check that $\left(x^{-1}\right)^n \cdot x^n = 1$ and that $x^n \cdot \left(x^{-1}\right)^n = 1$. □

0.2 Fields

Definition 12 (Fields)**.** A **field** is a set $\mathbb{F}$ endowed with two internal operations $+$ and $\cdot$ that satisfy the following nine axioms:

(F1) The operation $+$ is *associative*. That is, for all x, y, z in $\mathbb{F}$, $x + (y + z) = (x + y) + z$.

(F2) The operation $+$ is *commutative*. That is, for all x, y, z in $\mathbb{F}$, $x + y = y + x$.

(F3) The operation $+$ has a unique *neutral element*. That is, there exists an element z in $\mathbb{F}$ such that for all x in $\mathbb{F}$, $x + z = x$. From now on we denote such element by "0".

(F4) Every element has a unique *additive inverse*. That is, for all x in $\mathbb{F}$ there exists exactly one element y in $\mathbb{F}$ such that $x + y = 0$. From now on we denote such element by "$-x$".

(F5) The operation $\cdot$ is *associative*. That is, for all x, y, z in $\mathbb{F}$, $x \cdot (y \cdot z) = (x \cdot y) \cdot z$.

(F6) The operation $\cdot$ is *commutative*. That is, for all x, y, z in $\mathbb{F}$, $x \cdot y = y \cdot x$.

(F7) The operation $\cdot$ has a unique *neutral element* different than 0. That is, there exists a unique element $z \neq 0$ in $\mathbb{F}$ such that for all x in $\mathbb{F}$, $xz = x$. From now on we denote such neutral element by "1".

(F8) Every element except 0 has a unique *multiplicative inverse*. That is, for all $x \neq 0$ in $\mathbb{F}$ there exists exactly one element y in $\mathbb{F}$ such that $xy = 1$. From now on we denote such element by "x^{-1}".

(F9) The operation $\cdot$ *distributes* $+$: for all x, y, z in $\mathbb{F}$, $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$.

Example 13. Fields are precisely commutative rings where every nonzero element is invertible.

1 Matrices

1.1 From linear systems to matrices

Definition 14. Let R be a ring. A *linear equation (over R)* is an expression of the type

$$a_1x_1 + \ldots + a_nx_n = b,$$

where $a_1, \ldots, a_n, b$ are elements of R, called *coefficients*. The x_i's are called *unknowns* or *variables*. A *solution* for the equation above is an element $(p_1, \ldots, p_n) \in R^n$ for which $a_1p_1 + \ldots + a_np_n = b$. Sometimes we write it $x_1 = p_1$, …, $x_n = p_n$.

Example 15. The equation $0x + 0y + 0z = 2$ has no solution.
The equation $2x = 3$ has exactly one solution, $x = 1.5$.
The equation $0x + 0y = 0$ has infinitely many solutions. In fact, any value of x and y will do.
The equation $x - 3y = 0$ has infinitely many solutions, but not any pair of values would work. If we set $y = t$, then x must equal $3t$. Thus the solution set is $\{(3t, t) : t \in \mathbb{R}\}$.

Definition 16. A *linear system* is a finite set of m linear equations in the same set of n variables, equations which we want to be hold true simultaneously. In other words, a linear system is an expression of the form

$$\begin{cases} a_{1,1}x_1 + \ldots + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + \ldots + a_{2,n}x_n = b_2 \\ \vdots \\ a_{m,1}x_1 + \ldots + a_{m,n}x_n = b_m \end{cases}$$

where the $a_{i,j}$'s and the b_i's are elements of R, called *coefficients*. We use the convention to start numbering top-to-bottom and left-to-right, because that is how we write in the Western world.

Notation When $n \leq 3$, for the variables we prefer to use the letters x, y, z.

Example 17. The system

$$\begin{cases} x - y = 1 \\ 2x + y = 6 \end{cases}$$

has the unique solution $(x, y) = (\frac{7}{3}, \frac{4}{3})$. Note that each of the two equations above had infinitely many solutions. There is a geometric application of this "solving problem": Namely, when $R = \mathbb{R}$, in the Cartesian plane, the equations $x - y = 1$ and $2x + y = 6$ represent two lines. Solving the system is therefore the way to find the intersection point of the two lines. This explains the name *linear*.

Example 18. The system

$$\begin{cases} x - y = 1 \\ 2x - 2y = 6 \end{cases}$$

has no solutions. If you fancy the geometric interpretation above, the equations $x - y = 1$ and $2x - 2y = 6$ over $\mathbb{R}$ represent two parallel lines.

Example 19. The system

$$\begin{cases} x + y - z = 1 \\ 2x - y + z = 6 \\ x - 2y + 2z = 5 \end{cases}$$

has infinitely many solutions, though not every triple will do: In fact, x is forced to be equal to -7. However, any triple $(-7, 8 + s, s)$, with $s \in R$, will do. Note also that the third equation is "superfluous", in the sense that it is obtainable by subtracting the first equation from the second one. Is there a geometric interpretation? Sure: $x + y - z = 1$ over $R = \mathbb{R}$ is the equation of a plane in $\mathbb{R}^3$, whereas $2x - y + z = 6$ is a different plane that intersects the previous one in a line ℓ. At this point a random third plane would probably intersect this line ℓ in a point; the third plane we chose, however, happens to contain ℓ entirely. So the final intersection is the line ℓ itself.

Example 20. The system

$$\begin{cases} x + y - z = 1 \\ 2x + 2y - 2z = 2 \end{cases}$$

has infinitely many solutions. (Note that the second equation is superfluous: It is merely the first equation multiplied through by a constant.) If we set $z = s$ and $y = t$, the only request we get from the linear system is $x = s - t + 1$. Thus the solution set is $\{(s - t + 1, t, s) : s, t \in R\}$. There are now "two degrees of freedom", in the sense that we are now free to pick s and t independently. The geometric interpretation is that we are intersecting two identical planes in three-dimensional space; thus the solution set is two-dimensional.

There are many more applications beside the geometric one that we stressed above. The main common aspect of all these applications is this: Nature is complex, but linear systems are easy to solve. So if we can approximate a complex problem in terms of a linear one, we are able to quickly provide approximate answers. This is basically the idea behind Calculus: How do you measure the length of a curve? It is difficult. But if you approximate it with small segments, and that's what derivatives are about, then you can compute their lengths and add them up... Since this course is called "linear algebra" and not "geometry of linear subspaces",

we will mostly present things algebraically, without explaining every time how to apply these finds to geometric intersection problem.

In these first lectures we shall deal with **three main questions**:
(A) How do we explain a system of linear equations to the computer?
(B) What techniques can we apply to solve it? (What moves are legitimate?)
(C) What is our goal? What is a "checkmate" situation, i.e. a system we certainly know how to solve?

Here are the answers:

(A). Matrices

We encode a linear system simply by means of a spreadsheet, where columns correspond to variables, and rows to equations. The only thing to pay attention to, is: please place the coefficients in the correct variable column.

Example 21.

$$\begin{cases} x+y-z=1 \\ 2x-y+z=6 \\ x-2z=5 \end{cases} \quad \rightsquigarrow \begin{pmatrix} 1 & 1 & -1 & 1 \\ 2 & -1 & 1 & 6 \\ 1 & 0 & -2 & 5 \end{pmatrix}$$

Note that the last column corresponds to the "constant term". So a system of m linear equations in n variables is translated into a spreadsheet (usually called "*augmented matrix*") of m rows, and $n+1$ columns.

Definition 22. An $m \times n$ *matrix* A is an array of numbers in a ring R [1], called *entries*, arranged in rows (numbered from top to bottom) and columns (numbered from left to right). The entry in row i and column j is denoted by $a_{i,j}$. Thus a generic $m \times n$ matrix looks like

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & & \ddots & \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix}$$

Definition 23. The elements $a_{i,i}$ are called *diagonal elements*. The first diagonal element is at the top left of the matrix; the last one, though, need not be at the bottom right. (It depends on whether $m < n$, $m = n$, or $m > n$.) An $m \times n$ matrix is called *square* if $m = n$. In a square matrix, the last diagonal element is at the bottom right corner.

Definition 24 (Identity matrix)**.** Let m be any positive integer. The identity matrix I_m is the $m \times m$ matrix that has 1s on the diagonal, and 0s elsewhere. This is always well-defined because any ring has elements called 0 and 1.

Note: The plural of "matrix" is "matrices".

Definition 25. Given two $m \times n$ matrices A, B with entries $a_{i,j}$, $b_{i,j}$ in the same ring R, we define $A + B$ to be the matrix C with entries $c_{i,j}$ defined by

$$c_{i,j} = a_{i,j} + b_{i,j}.$$

[1] For some application, e.g. the Gauss–Jordan algorithm, it will be necessary to require R to be a "field", like $\mathbb{R}$ or $\mathbb{Q}$. In this course, we will typically consider matrices with entries in $\mathbb{Q}$ or $\mathbb{R}$.

Definition 26. Given an $\ell \times m$ matrix A and an $m \times n$ matrix B, both with entries in some ring R, we define $A \times B$ (or simply "AB") to be the $\ell \times n$ matrix C with entries $c_{i,j}$ defined by

$$c_{i,j} = \sum_{k=1}^{m} a_{i,k} b_{k,j}.$$

Theorem 27. *Let m be any positive integer. Given any ring R, the set $M_{m,m}(R)$ of square $(m \times m)$ matrices forms a ring with respect to the two operations above. The zero element is the matrix with all entries 0_R, whereas the 1 is the identity I_m.*

Definition 28. Given an $m \times n$ matrix A with entries $a_{i,j}$ in a ring R, and an element $r \in R$, we define rA to be the $m \times n$ matrix with entries defined by

$$(rA)_{i,j} = r \cdot (a)_{i,j}.$$

In other words, rA is the matrix obtained by multiplying every single entry of A by r.

Notation. Let

$$\begin{cases} a_{1,1}x_1 + \ldots + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + \ldots + a_{2,n}x_n = b_2 \\ \vdots \\ a_{m,1}x_1 + \ldots + a_{m,n}x_n = b_m \end{cases}$$

be a linear system. If we adopt the convention that $\mathbf{x}$ is the $n \times 1$ matrix

$$\mathbf{x} \stackrel{\text{def}}{=} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

and $\mathbf{b}$ is the $m \times 1$ matrix

$$\mathbf{b} \stackrel{\text{def}}{=} \begin{pmatrix} b_1 \\ x_2 \\ \vdots \\ b_m \end{pmatrix},$$

then the linear system can be rewritten in a much more compact way as

$$A\mathbf{x} = \mathbf{b}.$$

(B). Legal moves

The following three moves obviously modify a linear system but do not change the solution set:

1. Interchange two equations.
2. Multiplying all terms of an equation by an invertible $q \in R$.
3. Replacing an equation $(*)$ by the sum of $(*)$ and q times another equation $(**)$ from the same linear system. (Or in other words, adding to one equation a constant times another equation.)

The third move may confuse you, but the crucial point is that the "other equation" (**) is brought along unchanged. This allows you to reverse the move: Just add to equation (*) $-q$ times equation (**). For this third step, q may be chosen to be zero, although this would result in no change to the system whatsoever.

Example 29. Obviously the linear systems below are equivalent

$$\begin{cases} x+y-z=1 \\ 2x-y+z=6 \\ x-2z=5 \end{cases} \quad \text{and} \quad \begin{cases} 7x+7y-7z=7 \\ 2x-y+z=6 \\ x-2z=5 \end{cases}$$

Example 30. Obviously the systems below are equivalent

$$\begin{cases} x+y-z=1 \\ 2x-y+z=6 \\ x-2z=5 \end{cases} \quad \text{and} \quad \begin{cases} 2x-y+z=6 \\ x+y-z=1 \\ x-2z=5 \end{cases}$$

Example 31. The systems below are equivalent: The move was, adding to equation (ii) -2 times equation (i).

$$\begin{cases} x+y-z=1 \\ 2x-y+z=6 \\ x-2z=5 \end{cases} \rightsquigarrow \begin{cases} x+y-z=1 \\ -3y+3z=4 \\ x-2z=5 \end{cases}$$

Note that equation (ii) was modified, but equation (i) was copy-pasted. If we now add to equation (ii) $+2$ times equation (i), we go back to the original system:

$$\begin{cases} x+y-z=1 \\ -3y+3z=4 \\ x-2z=5 \end{cases} \rightsquigarrow \begin{cases} x+y-z=1 \\ 2x-y+z=6 \\ x-2z=5 \end{cases}$$

These moves can obviously be explained to the computer, which understand spreasheets very well, and generalized to arbitrary rings, as follows. Suppose that we are allowed to perform these three moves on a $m \times n$ matrix with entries in a ring R.

1. Swap two rows.
2. Multiplying a row through by an invertible $q \in R$.
3. Add q times one row to another, for some $q \in R$ (invertible or not).

These moves, however many, do not affect the solution set of the corresponding linear system. The three moves above are called *elementary row operations.*

(C). The goal

When is a linear system immediately solvable? We want a "halting criterion" for the computer: The idea is to program it with a "do moves on your spreadsheet until you reach an ideal situation in which you can stop." What is this ideal situation?

Definition 32 (RRE form)**.** A matrix $m \times n$ with entries in a ring R is *in RRE form* if it satisfies the following four conditions:

(RRE1) If a row does not consist entirely of zeroes, the first nonzero number is a 1. (We call this a *leading 1.*)

(RRE2) If there are any rows that consist entirely of zeros, then they are grouped together at the bottom of the matrix.
(RRE3) In any two successive rows that do not consist entirely of zeros, the leading 1 in the lower row occurs farther to the right than the leading 1 in the higher row.
(RRE4) Each column that contains a leading 1 has zeros everywhere else in that column.

Remark 33. RRE stands for "Reduced Row Echelon". Anton's textbook uses the expression "*in Row Echelon form*" (without "reduced") to denote a matrix that satisfies (RRE1), (RRE2) and (RRE3), though not necessarily (RRE4). Some other textbooks use the same expression "in Row Echelon form" to describe a matrix that satisfies just (RRE2) and (RRE3). To avoid confusion, we will try not to use this expression altogether.

Non-Example 34. The following matrices are *not* in RRE form.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 8 & 1 \\ 0 & 1 & 0 & 6 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Example 35. The following three matrices are in RRE form.

$$\begin{pmatrix} 1 & 0 & 8 & 1 \\ 0 & 1 & 0 & 6 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Theorem 36. *A square matrix in RRE form is either a matrix with a zero row, or the identity.*

Proof. There is one leading 1 in each row by (RRE1), and the column of any leading 1 contains no other leading 1 by (RRE4). So the m leading ones are to be placed in distinct columns. But there are m columns, so by the Pigeonhole Principle, there is exactly one 1 in each column. Thus every row contains one 1, and no further element. So there is exactly one 1 in every row and column. Condition (RRE3) forces the ones to be on the diagonal. □

Why is this idea? Because **if an augmented matrix is already in RRE form, the system is already solved**!

Example 37. The three matrices in RRE form above, with entries in $\mathbb{R}$, correspond to the linear systems

$$\begin{cases} x + 8z = 1 \\ y = 6 \\ 0 = 0 \end{cases} \qquad \begin{cases} x = 0 \\ y = 0 \\ z = 0 \\ 0 = 1 \end{cases} \qquad \begin{cases} x = 1 \\ y = 2 \\ 0 = 0 \\ 0 = 0 \end{cases}$$

The first one has infinitely many solutions $\{(1 - 8z, 6, z) \ : \ z \in R\}$. The second system has no solution, due to the impossibility of the last equation. The third system has exactly one solution, $\{(1, 2)\}$.

In the next lecture, we will see a German strategy (called "Gauss-Jordan") to always get to the goal with the moves available. Or more formally, to "reduce any augmented matrix into RRE form, using (only) elementary row operations". The strategy only works if the entries are in a field.

1.2 The Gauss–Jordan algorithm

A matrix with all entries equal to zero is already in reduced row echelon form. For all other matrices with entries in a **field** $\mathbb{F}$, we can run the following pair of algorithms:

Downward phase

- Find a leftmost nonzero element p. (That is, among all nonzero elements, pick one furthest to the left). With a row swap, bring it to the top row. Note that the columns left of p, if any exist, are all zeroes, because of the way p was chosen.
- Multiply the first row by p^{-1}, which exists because $\mathbb{F}$ is a field and $p \neq 0$. Now row 1 has a leading 1, in some column i.
- We want to zero out all entries in the i-th column, below the leading 1. To this end, for any j such that $a_{i,j} \neq 0$, add to row j the multiple of row 1 by the constant $-a_{ij}$. Note that this has no effect on the zeroes in the columns left of i (if $i > 1$). Also, there is no effect on other rows: Only row j is affected. For this reason, for fixed i, you may zero out the nonzero $a_{i,j}$'s in any order you like.
- Now column i has only one 1 in the top row, whereas all columns $< i$ are full of zeroes. Let A' be the matrix obtained from A by deleting the first row and the first i columns. Replace A with A' and start over.

At the end of this first algorithm, we have a matrix that satisfies (RRE1), (RRE2), (RRE3), but not necessarily (RRE4). Namely, *above* any leading 1 there might be nonzero entries. To achieve (RRE4), we run a second "correction algorithm":

Upward phase

- begin with the last nonzero row, and check if above its leading 1 the entries are all zero. If not, adds suitable multiple of the last nonzero row to the rows above to introduce zeros above the leading 1 of the last row.
- Now move up to the second-last nonzero row, and check if above its leading 1 the entries are all zero. If not, add suitable multiples of this row to the rows above to clear up the column of the leading 1.
- Now move up one row, and so on.

Example 38.

$$A = \begin{pmatrix} 0 & 0 & 2 & 1 & 0 \\ 0 & 3 & 6 & 0 & 9 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 0 & 3 & 6 & 0 & 9 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 0 & 1 & 2 & 0 & 3 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \end{pmatrix}$$

Since $\begin{pmatrix} 2 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$

we arrived from A to

$$A' = \begin{pmatrix} 0 & 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

This concludes the 'downward phase'. Now for the upward phase,

$$A' = \begin{pmatrix} 0 & 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 0 & 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

and we are done.

Remark 39. Inside the algorithm there are moments in which a choice is necessary; for example, at the beginning, if there are two different leftmost nonzero entries, either one can be chosen to become the first leading 1. But let R, R' be two RRE matrices obtained from the same matrix A, via different choices while performing the algorithm. If we view R, R' and A as augmented matrices, the associated linear systems must have the same solutions (because our moves do not change the solution set). Using this, one can prove that $R = R'$. In other words, the RRE form of a matrix is unique.

1.3 Elementary matrices and inverses

Definition 40. Let R be any ring. For fixed $m \in \mathbb{N}$, there are three types of so-called "elementary $m \times m$ matrices":

- $E_{i,j}$ is the matrix obtained from I_m by swapping rows i, j;
- $E_{i,j}(q)$ is the matrix obtained from I_m by replacing row i with "row i plus z times row j";
- $E_i(q)$ is the matrix obtained from I_m by multiplying row i by some **invertible** q in R.

For example, for $m = 2$ and $R = \mathbb{Q}$, we have:

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad E_{1,2} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad E_{1,2}(4) = \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix}, \quad \text{and } E_2(7) = \begin{pmatrix} 1 & 0 \\ 0 & 7 \end{pmatrix}.$$

These elementary matrices help us perform the analogous operations on a generic matrix A. In fact,

- $E_{i,j}A$ is the matrix obtained from A by swapping rows i, j;
- $E_{i,j}(q)A$ is obtained from A by replacing row i with "row i plus q times row j";
- $E_i(q)A$ is obtained from A by multiplying row i by some invertible q.

For example,

$$E_{1,2}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} c & d \\ a & b \end{pmatrix}; \; E_{1,2}(4)\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+4c & b+4d \\ c & d \end{pmatrix}; \; E_2(7)\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ 7c & 7d \end{pmatrix}.$$

Remark 41. An analogous result holds for columns:

- $AE_{i,j}$ is obtained from A by swapping columns i, j;
- $AE_{i,j}(q)$ is obtained from A by replacing column i with "column i plus q times column j";
- $AE_i(q)$ is obtained from A by multiplying column i by some invertible q.

For example,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} E_{1,2} = \begin{pmatrix} b & a \\ d & c \end{pmatrix}; \; \begin{pmatrix} a & b \\ c & d \end{pmatrix} E_{1,2}(4) = \begin{pmatrix} a+4b & b \\ c+4d & d \end{pmatrix}; \; \begin{pmatrix} a & b \\ c & d \end{pmatrix} E_2(7) = \begin{pmatrix} a & 7b \\ c & 7d \end{pmatrix}.$$

Proposition 42. *All elementary matrices are invertible.*

Proof. The inverse of $E_{i,j}$ is $E_{i,j}$ itself (swapping back!); the inverse of $E_{i,j}(q)$ is $E_{i,j}(-q)$ (re-subtracting back what we had added); and the inverse of $E_i(q)$ is $E_i(q^{-1})$ (scaling back). Note that for this third result, it was crucial to pick a q that is invertible. □

The algorithm we have seen in the previous section proves automatically the following:

Theorem 43 (Gauss)**.** *Given any nonzero $m \times n$ matrix A with entries in a field $\mathbb{F}$ (for example, $\mathbb{F} = \mathbb{R}$), there exists an $m \times n$ matrix M in RRE form and some elementary matrices $E_0, E_1, E_2, \ldots, E_\ell$ (all of size $m \times m$) such that*

$$E_\ell E_{\ell-1} \cdots E_1 E_0 A \; = \; M.$$

Proof. The effect of *any* legal move we perform during the Gauss–Jordan reduction is the same as left-multiplying the matrix by some suitable elementary matrix. □

This has a crucial consequence in terms of inverses. Note that **only for square matrices it makes sense to ask if they are invertible**, because square matrices form a ring, whereas $m \times n$ matrices with $m \neq n$ do not.

Theorem 44. *Given an $m \times m$ matrix A with entries in a field $\mathbb{F}$, the following are equivalent:*
(a) *A is invertible;*
(b) *$A\boldsymbol{x} = \boldsymbol{0}$ has only one solution;*
(c) *the Gauss–Jordan algorithm reduces A to I_m;*
(d) *A is a product of elementary matrices (of size $m \times m$).*

Proof. (a) ⇒ (b). Multiplying both sides of the equation $A\mathbf{x} = \mathbf{0}$ to the left by A^{-1}, we get $\mathbf{x} = \mathbf{0}$.

(b) ⇒ (c). Since $\mathbf{x} = \mathbf{0}$ is a solution, it must be the only one. Let A' (resp. I') be the $m \times (m+1)$ matrix obtained from A (resp. I_m) by appending to it an extra column of zeroes on the right. Running the Gauss–Jordan algorithm must then reduce the matrix A' to I'. The same steps, exactly in the same order, also reduce A to I_m.

(c) ⇒ (d). By Theorem 43, there are elementary matrices $E_0, E_1, E_2, \ldots, E_\ell$ (all of size $m \times m$) such that

$$E_\ell E_{\ell-1} \cdots E_1 E_0 A \;=\; I_m.$$

Now left-multiply both sides of the equation above by $(E_0)^{-1}(E_1)^{-1} \cdots (E_{\ell-1})^{-1}(E_\ell)^{-1}$. (Which exists, because elementary matrices are invertible.) We get

$$A \;=\; (E_0)^{-1}(E_1)^{-1} \cdots (E_{\ell-1})^{-1}(E_\ell)^{-1}.$$

(d) ⇒ (a). Any product of invertible matrices is an invertible matrix. □

Remark 45. Theorems 44 does not extend to matrices with entries in an arbitrary ring. Consider for example the positive integer 7. We can view 7 as a 1×1 matrix over the ring $\mathbb{Z}$. It is not true that 7 is invertible, but it is true that $7x = 0$ has only one solution. It is not true that the Gauss–Jordan algorithm reduces 7 to 1 (because if we try to get a leading 1, we should multiply by the inverse of 7, which does not exist.) And since we only defined elementary matrices of the type $E_i(q)$ *for q invertible*, 7 is not a product of elementary matrices.

Corollary 46. *Given an $m \times m$ matrix A with entries in a field $\mathbb{F}$, the following are equivalent:*
(a) *A is not invertible;*
(b) *$A\boldsymbol{x} = \boldsymbol{0}$ has some "nontrivial" solution $\boldsymbol{x} \neq \boldsymbol{0}$;*
(c) *the Gauss–Jordan algorithm reduces A to a matrix whose bottom row is made of zeroes;*
(d) *A cannot be written as product of elementary matrices.*

Proof. Straightforward from Theorems 44 and 36. □

Corollary 47. *Let A, B be $m \times m$ matrices with entries in a field. Then*

$$AB \text{ is invertible} \quad \Longleftrightarrow \text{ both } A, B \text{ are invertible.}$$

Proof. The "⇐" direction is clear: as we saw already, the inverse of AB is $B^{-1}A^{-1}$, since

$$(AB)B^{-1}A^{-1} = I_m = B^{-1}A^{-1}(AB).$$

For the other direction: exploiting Theorem 44, we will prove the invertibility of B by showing that $B\mathbf{x} = \mathbf{0}$ implies $\mathbf{x} = \mathbf{0}$. Indeed, $B\mathbf{x} = \mathbf{0}$ implies $ABtextbfx = A\mathbf{0} = \mathbf{0}$, which by Theorem 44 has only $\mathbf{x} = \mathbf{0}$ as solution. Thus B^{-1} exists. But then we can write A as a product of invertible matrices:

$$A = (AB)(B^{-1}).$$

Thus also A is invertible. □

Corollary 48. *Let A, B be $m \times m$ matrices with entries in a field. If $AB = I_m$, then also $BA = I_m$.*

Proof. If A is invertible, the result is easy: from $AB = I_m$, left-multiplying by A^{-1}, we get $B = A^{-1}$. Now suppose by contradiction that A is not invertible. Then its RRE form M of A has a bottom row of zeroes. By Gauss' theorem 43,

$$E_\ell E_{\ell-1} \cdots E_1 E_0 A \;=\; M.$$

So if we multiply to the right by B the equation above, since $AB = I_m$ we get

$$E_\ell E_{\ell-1} \cdots E_1 E_0 \;=\; MB.$$

So MB is invertible, because it equals a product of elementary matrices, which are invertible. But at the same time, we know that the last row of M is zero, so by the way we defined row-by-column multiplication, the last row of MB must also consist entirely of zeroes. This means that MB is not invertible. A contradiction. □

Remark 49. Theorem 44 gives an algorithmic, explicit way to compute the inverse of a matrix A, when it exists. The idea is to reduce A to a matrix in RRE form via the Gauss–Jordan algorithm. If the resulting RRE matrix has a row of zeroes at the bottom, then A was not invertible. If instead the resulting RRE matrix is the identity, then keeping track of the performed operations we can write

$$E_\ell E_{\ell-1} \cdots E_1 E_0 A \;=\; I_m,$$

whence right-multiplication by A^{-1} yields

$$E_\ell E_{\ell-1} \cdots E_1 E_0 \;=\; A^{-1}.$$

Remark 50. With a harder proof, Corollary 48 is still true for matrices in a commutative rings [2]. However, there are non-commutative rings R over which $AB = I_2$ but $BA \neq I_2$ [3].

1.4 Diagonal, triangular, symmetric matrices

Definition 51. A matrix A with entries in a ring R is called

- *lower triangular*, if $a_{i,j} = 0$ for all $i < j$;
- *upper triangular*, if $a_{i,j} = 0$ for all $i > j$; *diagonal*, if it is both upper and lower triangular.

Clearly, diagonal matrices are symmetric.

Example 52. $E_{i,j}(r)$ is upper triangular if $i < j$ and lower triangular if $i > j$. $E_i(r)$, with r invertible, is diagonal. $E_{i,j}$ is neither upper nor lower triangular.

[2] Reutenauer–Straubing, *Inversion of matrices over a commutative semiring*, J. Algebra 88 (1984), 350–360.
[3] J. C. Shepherdson, *Inverses and zero divisors in matrix rings*, Proc. London Math. Soc. 1, (1951).

Proposition 53. *If A and B are $m \times m$ lower triangular matrices, so are their sum and their product. The same is true for lower triangular (and thus, for diagonal).*

Proof. Suppose $a_{i,j} = b_{i,j} = 0$ for all $i < j$. Then obviously $a_{i,j} + b_{i,j} = 0$ for all $i < j$. Now given $i < j$, consider a new term k in $\{1, \ldots, m\}$. If $k < j$, then $b_{k,j}$ is zero. On the other hand, if $k \geq j$, then $k > i$, and so $a_{i,k}$ is zero. Either way,

$$(AB)_{i,j} = \sum k = 1^m a_{i,k} b_{k,j} = 0.$$

This shows that lower triangular matrices are closed with respect to sum and matrix multiplication. A completely analogous proof shows the same for upper triangular matrices. □

Definition 54. The *transpose* of an $m \times n$ matrix A is the $n \times m$ matrix $A^\top$ such that

$$(A^\top)_{i,j} = (A)_{j,i}.$$

In other words, $A^\top$ is obtained from A by reflection along the main diagonal. It is clear from the definition that $(A^\top)^\top = A$. Note that A is upper triangular if and only if $A^\top$ is lower triangular.

Remark 55. The "dot product" $\mathbf{v} \bullet \mathbf{w}$ from calculus, defined as

$$(v_1, \ldots, v_n) \bullet (w_1, \ldots, w_n) \stackrel{\text{def}}{=} v_1 w_1 + \ldots + v_n w_n,$$

can be seen as a matrix multiplication: It is simply $\mathbf{v}\mathbf{w}^\top$.

Definition 56. The (necessarily square!) matrices A for which $A = A^\top$ are called *symmetric matrices.*

Example 57. The identity and the elementary matrices $E_i(q)$ are symmetric for any (invertible) q. Instead $E_1 2(1)$ and E_{12} are not symmetric.

Proposition 58. *If A, B are two $m \times n$ matrices (same shape!), then*

$$(A + B)^\top = A^\top + B^\top.$$

In particular, the sum of two symmetric matrices is symmetric.

Proof. The first part is left as exercise to you. For the second one: if $A = A^\top$ and $B = B^\top$, then by the first part

$$(A + B)^\top = A^\top + B^\top = A + B.$$ □

Proposition 59. *If A is an $m \times n$ matrix with entries in a ring R, and r is an element of R, then*

$$(rA)^\top = r(A^\top).$$

In particular, the multiples of a symmetric matrix are symmetric.

Proof. Left to you. □

Proposition 60. *Let R be a* commutative *ring. Let A be an $\ell \times m$ matrix with entries in R. Let B be an $m \times n$ matrix with entries in R. Then*

$$(AB)^\top = B^\top A^\top.$$

Proof. Note first that if A is $\ell \times m$ and B is $m \times n$, then $A^\top$ is $m \times \ell$ and $B^\top$ is $n \times m$, so the expression $B^\top A^\top$ makes perfect sense (whereas $A^\top B^\top$, if $\ell \neq n$, does not). Let us check if the elements in row i and column j of the two matrices $(AB)^\top$ and $B^\top A^\top$ are the same. Indeed,

$$\left((AB)^\top\right)_{i,j} \stackrel{\text{def}}{=} (AB)_{j,i} = \sum_{k=1}^{m} a_{j,k} b_{k,i}, \qquad \text{and}$$

$$\left(B^\top A^\top\right)_{i,j} = \sum_{k=1}^{m} \left(B^\top\right)_{i,k} \left(A^\top\right)_{k,j} = \sum_{k=1}^{m} b_{k,i} a_{j,k}.$$

The commutativity of R enables us to conclude, because $a_{j,k} b_{k,i} = b_{k,i} a_{j,k}$ inside R. □

Proposition 61. *If A is an invertible square matrix with entries in a commutative ring, then $(A^\top)$ also is, and*

$$\left(A^\top\right)^{-1} = \left(A^{-1}\right)^\top.$$

In particular, if a symmetric matrix is invertible, and the entries are in a commutative ring, then the inverse is also symmetric.

Proof. If A is $m \times m$ invertible with entries in a commutative ring, by the previous proposition

$$A^\top \left(A^{-1}\right)^\top = \left(A^{-1} A\right)^\top = (I_m)^\top = I_m.$$

As for the second part: the zero matrix is clearly symmetric, but not invertible. That said, if A is symmetric *and* invertible, then by the first part

$$A^{-1} = \left(A^\top\right)^{-1} = \left(A^{-1}\right)^\top$$ □

Remark 62. The previous Propositions are not true for matrices over non-commutative rings. If r, s are elements of R such that $rs \neq sr$, we can view them as 1×1 matrices, and $(rs)^\top = rs \neq sr = s^\top r^\top$. It is more complicated to see this, but there are also non-commutative rings R such that some 2×2 matrix A with entries in R is invertible and symmetric, but has an inverse that is not symmetric.

Remark 63. The product of two symmetric matrices is typically **not** symmetric, even if the entries of A are from a field. You can see for yourself how the 'obvious' proof goes wrong:

$$(AB)^\top = B^\top A^\top = BA \quad (\text{not } AB!).$$

For a counterexample, look at the following matrices with entries in $\mathbb{Q}$:

$$\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 1 \\ 7 & 2 \end{pmatrix}.$$

However, when $AB = BA$, and the entries of A are from a commutative ring, then indeed $(AB)^\top = B^\top A^\top = BA = AB$, so the product of *commuting* symmetric matrices with entries in a commutative ring is indeed symmetric. A special case is the one below:

Proposition 64. *Let R be a commutative ring. If a matrix A with entries in R is symmetric, so are its power A^n, with $n \in \mathbb{N}$.*

Proof. AA^2 is the product of commuting symmetric matrices. Same for $A^3 = A \cdot A^2$, and so on. If we know that A^n is symmetric, then

$$(A^{n+1})^\top = (A \cdot A^n)^\top = (A^n)^\top \cdot A^\top = A^n \cdot A = A^{n+1}. \qquad \square$$

Remark 65. Let R be a non-commutative ring. Let r, s be elements such that $rs \neq sr$. Then

$$\begin{pmatrix} r & s \\ s & 0 \end{pmatrix}^2 = \begin{pmatrix} r & s \\ s & 0 \end{pmatrix} \begin{pmatrix} r & s \\ s & 0 \end{pmatrix} = \begin{pmatrix} r^2 + s^2 & sr \\ rs & 0 \end{pmatrix}$$

gives an example of a symmetric matrix A with A^2 not symmetric.

Remark 66. Remember that $AB \neq BA$ in general, so formulas like $(A+B)^2 = A^2 + 2AB + B^2$ do not work for matrices. In fact, all we can say is

$$(A + B)^2 = (A + B)(A + B) = A^2 + AB + BA + B^2.$$

The right-hand side equals $A^2 + 2AB + B^2$ precisely when $AB = BA$.

Proposition 67. *Let R be a commutative ring. For any $m \times n$ matrix A with entries in R, both matrices $AA^\top$, $A^\top A$ are symmetric.*

Proof. Left to you. Note that by Remark 65, the "R commutative" assumption is necessary. $\square$

Proposition 68. *For any square matrix A, the matrix $A + A^\top$ is symmetric.*

Proof. Left to you. $\square$

2 Trace and determinant of square matrices

In this section, we always work with *square* matrices. We start with an important definition:

Definition 69 (Similarity)**.** Let A, B be two $m \times m$ matrices with entries in a ring R. We say that B *is similar to* A if there exists an invertible matrix P such that

$$B = P^{-1}AP.$$

For example, any matrix is similar to itself (by taking $P = I_m$). Note that if B is similar to A, then automatically A is similar to B: In fact, if $B = P^{-1}AP$, then $PB = AP$ and $PBP^{-1} = A$; so setting $Q = P^{-1}$, we can write

$$Q^{-1}BQ = A.$$

For this reason, we usually say that "A and B are similar matrices" if either is similar to the other. Finally, note that if A is similar to B, and B is similar to C, then A is automatically similar to C. The proof is left to you.

2.1 The trace of a square matrix

Definition 70. The *trace* of a square $m \times m$ matrix A with entries in some ring R is the sum of its diagonal elements, i.e.

$$\operatorname{trace}(A) \stackrel{\text{def}}{=} \sum_{i=1}^{m} a_{i,i}.$$

Proposition 71. *For any square matrix* A, $\operatorname{trace}(A) = \operatorname{trace}(A^\top)$.

Proof. The diagonal elements of A and $A^\top$ are the same. □

Proposition 72. *For all* $m \times m$ *matrices* A, B *with entries in some ring* R*, and for all elements* $r \in R$*, one has*

$$\operatorname{trace}(A+B) = \operatorname{trace}(A) + \operatorname{trace}(B) \qquad \textit{and} \qquad \operatorname{trace}(rA) = r\operatorname{trace}(A).$$

Proof. Exercise. □

Proposition 73. *For all* $m \times m$ *matrices* A, B *with entries in a* commutative *ring*

$$\operatorname{trace}(AB) = \operatorname{trace}(BA)$$

even if this might be different from $\operatorname{trace}(A)\operatorname{trace}(B) = \operatorname{trace}(B)\operatorname{trace}(A)$.

Proof. One has

$$\operatorname{trace}(AB) = \sum_{i=1}^{m} (AB)_{i,i} = \sum_{i=1}^{m}\sum_{k=1}^{m} a_{i,k}b_{k,i} = \sum_{i=1}^{m}\sum_{k=1}^{m} b_{k,i}a_{i,k} = \sum_{i=1}^{m} (BA)_{i,i} = \operatorname{trace}(BA).$$

However, in the following example

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad BA = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

one immediately sees that

$$\operatorname{trace}(AB) = \operatorname{trace}(BA) = 1 \neq 0 = 0 \cdot 0 = \operatorname{trace}(A) \cdot \operatorname{trace}(B) = \operatorname{trace}(B) \cdot \operatorname{trace}(A). \qquad \square$$

Remark 74. The previous proof also shows that

$$\operatorname{trace}(AB) = \operatorname{trace}(BA)$$

even if A and B are rectangular, i.e. if A is $m \times n$ and B is $n \times m$ with $n \neq m$. In this case $\operatorname{trace}(A)$ and $\operatorname{trace}(B)$ are not even defined.

Corollary 75. *Similar matrices have same trace.*

Proof. Suppose $B = P^{-1}AP$, for some invertible matrix P. Set $C = P^{-1}A$. By the previous proposition, $\operatorname{trace}(CP) = \operatorname{trace}(PC)$. But CP is B, and PC is A. □

Remark 76. Given three $m \times m$ matrices A, B, C, it is typically **not** true that

$$\operatorname{trace}(ABC) = \operatorname{trace}(ACB).$$

For example, take

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Then

$$ABC = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \quad \text{while} \quad ACB = \begin{pmatrix} c & d \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} d & 0 \\ 0 & 0 \end{pmatrix}.$$

What is true is that the trace (for matrices with entries in a commutative ring) is invariant under "circular shifts", i.e.

$$\operatorname{trace}(ABCD) = \operatorname{trace}(BCDA) = \operatorname{trace}(CDAB) = \operatorname{trace}(DABC).$$

This is very easy to prove: for example to show $\operatorname{trace}(ABCD) = \operatorname{trace}(BCDA)$, just call $B' \stackrel{\text{def}}{=} BCD$, and use the fact that we know already, namely, that

$$\operatorname{trace}(AB') = \operatorname{trace}(B'A).$$

Proposition 77. *Let A be a matrix with entries in the ring ($\mathbb{Z}$, $\mathbb{Q}$ or) $\mathbb{R}$. Then*

$$\operatorname{trace}(AA^\top) \geq 0,$$

with equality if and only if A is the zero matrix.

Proof. Let A be an $m \times n$ matrix. The (i,i) entry of $AA^\top$ is the dot product of the i-th row of A with itself. So it equals $\sum_{j=1}^{n}(a_{i,j})^2$. Thus all summands of the trace of $AA^\top$ are sums of squares, and in particular non-negative. Also, if some $a_{i,j}$ is not zero, then $(a_{i,j})^2$ is positive, so the trace of $AA^\top$ is also positive. So if the trace of $AA^\top$ is zero, all the $a_{i,j}$'s must be zero. □

2.2 The determinant, defined inductively, and Cauchy–Binet for fields

The determinant is a function defined on square matrices that has the four following properties:

- The determinant of the identity matrix is 1.
- The exchange of two rows multiplies the determinant by -1.
- Multiplying a row by an element r multiplies the determinant by this r.
- Adding a multiple of one row to another row does not change the determinant.

The difficult part is to give a definition. The official definition involves permutations, a topic you will see either in Abstract Algebra or in Discrete Mathematics. A second possible definition is to use the four properties above, because it turns out that they determine the function. In this course you will see instead a third, somewhat "recursive" definition, in which you'll be told how to compute the determinant of a 1×1 matrix; and then, assuming you know how to compute determinants of $(m-1) \times (m-1)$ matrices, you are going to be told how to compute the determinant of an $m \times m$ matrix. There is a problem with the way the book presents it, namely, circularity: It omits the proof of Theorem 2.1.1, which is needed for its definition of "determinant", but then it uses Theorem 2.1.1 to prove other three statements (Theorem 2.2.1, 2.2.2 and 2.2.3) which are somewhat necessary to prove Theorem 2.1.1 using this "recursive" definition of determinant. So I fixed it here for you.

Definition 78 (Determinant)**.** Let A be a square matrix with entries in a ring R.

- If A is a 1×1 matrix consisting of a single element a, we set $\det A \stackrel{\text{def}}{=} a$.
- If A is an $m \times m$ matrix with $m \geq 2$, we set

$$\det A \stackrel{\text{def}}{=} \sum_{i=1}^{m} (-1)^{i+1} \cdot a_{i,1} \cdot M_{i,1}^{A},$$

  where $M_{i,1}^{A}$, usually called "*the minor of* $(i,1)$ *in* A", is the determinant of the $(m-1) \times (m-1)$ matrix obtained from A by removing the i-th row and the 1st column.

Notation. The element $(-1)^{i+1} \cdot M_{i,1}^{A}$ is called "*the cofactor of* $(i,1)$ *in* A"; which explains why this way of computing the determinant is usually called a "cofactor expansion along the first column", and typically denoted by $C_{i,j}^{A}$. In these notes we use the notation $A^{-i,-j}$ for the matrix obtained from a matrix A by deleting the i-th row and the j-th column. So for example $\det A_{-i,1} = M_{i,1}^{A}$ and $(-1)^{i+1} \det A_{-i,1} = C_{i,j}^{A}$.

Example 79. $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = (-1)^2 \cdot a \cdot \det(d) + (-1)^3 \cdot c \cdot \det(b) = ad - cb.$

Remark 80. While $\operatorname{trace}(A+B) = \operatorname{trace}(A) + \operatorname{trace}(B)$, in general $\det(A+B) \neq \det A + \det B$. For example,

$$\det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 \neq 0 + 0 = \det \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \det \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Also, while $\operatorname{trace}(rA) = r \operatorname{trace}(A)$, in general $\det rA \neq r \det A$: for example,

$$\det \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = 4 \neq 2 \det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Example 81. $\det \begin{pmatrix} 1 & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{pmatrix} = (-1)^2 \cdot 1 \cdot \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - cb.$ In particular, $\det I_3 = 1$.

Proposition 82. *The determinant of an upper triangular matrix is the product of its diagonal elements. In particular,* $\det I_m = 1$.

Proof. The first column is zero below $a_{1,1}$. Hence,

$$\det A = a_{1,1} \cdot M_{1,1}^{A},$$

and $M_{1,1}^{A}$ is upper triangular of size $(m-1) \times (m-1)$. Iterating the argument, we conclude. $\square$

For Lower Triangular matrices, see Proposition 89 below.

Example 83. $\det \begin{pmatrix} 0 & a & b \\ 1 & 0 & 0 \\ 0 & c & d \end{pmatrix} = (-1)^3 \cdot 1 \cdot \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = -(ad - cb).$

Example 84. $\det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2a & 2b \\ 0 & c & d \end{pmatrix} = (-1)^2 \cdot 1 \cdot \det \begin{pmatrix} 2a & 2b \\ c & d \end{pmatrix} = 2(ad - cb).$

Example 85. $\det \begin{pmatrix} 1 & 0 & 0 \\ 0 & a+c & b+d \\ 0 & c & d \end{pmatrix} = (a+c)d - c(b+d) = ad - cb.$

Example 86. For a general 3×3 matrix,

$$\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} = a_{1,1}a_{2,2}a_{3,3} - a_{1,1}a_{3,2}a_{2,3} - a_{2,1}a_{1,2}a_{3,3} + a_{2,1}a_{3,2}a_{1,3} + a_{3,1}a_{1,2}a_{2,3} - a_{3,1}a_{2,2}a_{1,3}.$$

In general, it can be proven inductively that the determinant of an $m \times m$ matrix consists of $m! = m(m-1)(m-2)\cdots 1$ summands, and each summand is the product of m entries, no two of which are in the same row/column.

Here however we are *not* interested in figuring out the formula for a 5×5 matrix, say. Instead, the idea is to reduce the matrix into RRE form, and control how the determinant changes along the process. Recall that there are two possibilities for the RRE form of a square matrix: Either it is the identity, which has determinant 1, or it is a matrix with a row of zeroes, in which case the determinant is zero because of the next Lemma.

Lemma 87. *Let A, B, C be three $m \times m$ matrices that are identical except for the k-th row. Suppose the k-th row of C is the sum of the k-th rows of A and B. Then $\det A + \det B = \det C$.*

Proof. When $m = 1$ the claim is obvious. Now suppose the statement has already been proven for $(m-1) \times (m-1)$ matrices, and let's show it for $m \times m$ matrices. If we erase from A, B, C row k and column 1, we get three identical matrices. With the language of "minors", we have $M^A_{k,1} = M^B_{k,1} = M^C_{k,1}$. If instead we erase from A, B, C some row i different than k, and column 1, we still get three matrices A', B', C' that are identical, all of size $(m-1) \times (m-1)$, except for some row of C' that is the sum of the corresponding rows of A' and B'. So because we believe the theorem for matrices of size $(m-1) \times (m-1)$, we have that $\det A' + \det B' = \det C'$. But then, splitting the three cases $i = k$ and $i \neq k$, we get

$$\begin{aligned}
\det C &\overset{\text{def}}{=} \textstyle\sum_{i=1}^m (-1)^{i+1} c_{i,1} M^C_{i,1} \\
&= (-1)^{k+1} c_{k,1} M^C_{k,1} + \textstyle\sum_{i \neq k} (-1)^{i+1} c_{i,1} M^C_{i,1} \\
&= (-1)^{k+1} (a_{k,1} + b_{k,1}) M^C_{k,1} + \textstyle\sum_{i \neq k} (-1)^{i+1} c_{i,1} (M^A_{i,1} + M^B_{i,1}) \\
&= (-1)^{k+1} a_{k,1} M^C_{k,1} + \textstyle\sum_{i \neq k} (-1)^{i+1} c_{i,1} M^A_{i,1} + (-1)^{k+1} b_{k,1} M^C_{k,1} + \textstyle\sum_{i \neq k} (-1)^{i+1} b_{i,1} M^B_{i,1} \\
&= (-1)^{k+1} a_{k,1} M^A_{k,1} + \textstyle\sum_{i \neq k} (-1)^{i+1} a_{i,1} M^A_{i,1} + (-1)^{k+1} b_{k,1} M^B_{k,1} + \textstyle\sum_{i \neq k} (-1)^{i+1} b_{i,1} M^B_{i,1} \\
&= \textstyle\sum_{i=1}^m (-1)^{i+1} a_{i,1} M^A_{i,1} + \sum_{i=1}^m (-1)^{i+1} b_{i,1} M^B_{i,1} \\
&= \det A + \det B. \quad \square
\end{aligned}$$

Lemma 88. *The determinant of a matrix with a zero row is 0.*

Proof. If the k-th row of A consists of all zeroes, then we can apply the previous lemma with $B = C = A$: we get

$$\det A + \det A = \det A,$$

which by cancelation implies $\det A = 0$. □

Proposition 89. *The determinant of a lower triangular matrix is the product of its diagonal elements.*

Proof. Let A be a lower triangular matrix. We know $a_{1,j} = 0$ for all $j > 1$. Thus, the matrix obtained from A by erasing row i and column 1, if $i > 1$ has a first row of zeroes, and thus has determinant zero by Lemma 88. In other words $M^A_{i,1} = 0$ for all $i \geq 2$. But then

$$\det A = a_{1,1} \cdot M^A_{1,1} + \sum_{i \geq 2} (-1)^{i+1} a_{i,1} M^A_{i,1} = a_{1,1} \cdot M^A_{1,1},$$

and $M^A_{1,1}$ is a lower triangular of size $(m-1) \times (m-1)$. Iterating the argument, we conclude. □

Lemma 90. *If B is obtained from A by swapping 2 consecutive rows,* $\det B = -\det A$.

Proof. When $m = 2$, $\det \begin{pmatrix} c & d \\ a & b \end{pmatrix} = cb - ad = -(ad - cb) = -\det \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Now suppose the statement has already been proven for $(m-1) \times (m-1)$ matrices, and let's show it for an $m \times m$ matrix A. Let $k-1, k$ be the two swapped rows. By induction, when $i \notin \{k-1, k\}$ we believe $M^B_{i,1} = -M^A_{i,1}$; note also that when $i \notin \{k-1, k\}$, $b_{i,1} = a_{i,1}$. As for the remaining two cases, $i \in \{k-1, k\}$, we have that $M^B_{k-1,1} = M^A_{k,1}$, $b_{k-1,1} = a_{k,1}$, and symmetrically $M^B_{k,1} = M^A_{k-1,1}$ and $b_{k,1} = a_{k-1,1}$. So summing up,

$$\begin{aligned}
\det B &\overset{\text{def}}{=} \textstyle\sum_{i=1}^{m} (-1)^{i+1} b_{i,1} M^B_{i,1} \\
&= (-1)^k b_{k-1,1} M^B_{k-1,1} + (-1)^{k+1} b_{k,1} M^B_{k,1} + \textstyle\sum_{i \notin \{k-1,k\}} (-1)^{i+1} b_{i,1} M^B_{i,1} \\
&= (-1)^k a_{k,1} M^A_{k,1} + (-1)^{k+1} a_{k-1,1} M^A_{k-1,1} + \textstyle\sum_{i \notin \{k-1,k\}} (-1)^{i+1} a_{i,1} (-M^A_{i,1}) \\
&= -\left((-1)^{k+1} a_{k,1} M^A_{k,1} + (-1)^k a_{k-1,1} M^A_{k-1,1} + \textstyle\sum_{i \notin \{k-1,k\}} (-1)^{i+1} a_{i,1} M^A_{i,1} \right) \\
&= -\det A. \quad \square
\end{aligned}$$

Lemma 91. *If C is obtained from A by swapping any two rows,* $\det C = -\det A$.

Proof. To swap rows p and q, with $p < q$, we can perform $2(q-p) - 1$ swaps of adjacent rows, namely: with $(p, p+1), (p+1, p+2), \ldots, (q-1, q)$ (a total of $q - p$ steps) we bring row p to the q-th row, and row q to the $(q-1)$-st row; so with the subsequent $q - p - 1$ steps $(q-2, q-1), (q-3, q-2), \ldots (p, p+1)$ we bring back what was the original row q from the $(q-1)$-st to the p-th place. Now, the number $2(q-p) - 1$ is always **odd**, so we are flipping the determinant to its additive inverse an odd number of times. □

Lemma 92. *If A has two identical rows,* $\det A = 0$.

Proof. With the same proof technique of the previous Lemma, after performing a finite number of row swaps (which do not change the determinant, apart from possibly flipping it to its additive inverse), we may assume that the two identical rows are row 1 and row 2. Now for 2×2 matrices the statement is true: $\det \begin{pmatrix} a & b \\ a & b \end{pmatrix} = ab - ab = 0$. Now suppose the statement has already been proven for $(m-1) \times (m-1)$ matrices, and let's show it for an $m \times m$ matrix A. If A' is the matrix obtained from A by removing column 1 and some row $i \geq 3$, then A' has also two

equal rows and size $(m-1)\times(m-1)$, so by induction we believe that its determinant is zero. That is, $M^A_{i,1}=0$ for $i\geq 3$. But then

$$\begin{aligned}\det A &= a_{1,1}M^A_{1,1}-a_{2,1}M^A_{2,1}+\textstyle\sum_{i\geq 3}(-1)^{i+1}a_{i,1}\cdot M^A_{i,1}\\ &= a_{1,1}M^A_{1,1}-a_{2,1}M^A_{2,1}+0\\ &= 0,\end{aligned}$$

because $a_{1,1}=a_{2,1}$ and $M^A_{1,1}=M^A_{2,1}$. □

Lemma 93. *Let A, C be matrices with entries in a commutative ring R. If C is obtained from A by multiplying all elements of some row by the same element r in R, then $\det C = r\det A$.*

Proof. The claim is true for 1×1 matrices, obviously. Note that for the first time the commutativity of the ring is necessary, because already for $m=2$, if $rs\neq sr$,

$$\det\begin{pmatrix} r & r\\ -s & 0\end{pmatrix} = sr \neq rs = r\det\begin{pmatrix} 1 & 1\\ -s & 0\end{pmatrix}.$$

Now suppose the statement has already been proven for $(m-1)\times(m-1)$ matrices, and let's show it for $m\times m$ matrices, with $m\geq 2$. Let k be the row that gets multiplied. Then in particular $c_{k,1}=r\cdot a_{k,1}$. Since all other rows of the matrices C and A are identical, $M^C_{k,1}=M^A_{k,1}$. On the other hand, let i be any element different than k. Then $c_{i,1}=a_{i,1}$. Let A' and C', respectively, be the $(m-1)\times(m-1)$ matrices obtained from A and C, respectively, by deleting the first column and the i-th row. Then C' is identical to A' except for one row, which is the corresponding row of A' multiplied by r. But since we believe the claim for $(m-1)\times(m-1)$ matrices, $\det C' = r\cdot\det A'$. So

$$\begin{aligned}\det C &\overset{\text{def}}{=} \textstyle\sum_{i=1}^m(-1)^{i+1}c_{i,1}M^C_{i,1}\\ &= (-1)^{k+1}c_{k,1}M^C_{k,1}+\textstyle\sum_{i\neq k}^m(-1)^{i+1}c_{i,1}M^C_{i,1}\\ &= (-1)^{k+1}r\cdot a_{k,1}M^A_{k,1}+\textstyle\sum_{i\neq k}^m(-1)^{i+1}a_{i,1}\cdot r\cdot M^A_{i,1}\\ &= (-1)^{k+1}r\cdot a_{k,1}M^A_{k,1}+\textstyle\sum_{i\neq k}^m(-1)^{i+1}r\cdot a_{i,1}\cdot M^A_{i,1}\\ &= r\cdot\left((-1)^{k+1}a_{k,1}M^A_{k,1}+\textstyle\sum_{i\neq k}^m(-1)^{i+1}a_{i,1}\cdot M^A_{i,1}\right)\\ &= r\cdot\det A.\end{aligned}$$ □

Lemma 94. *Let A, C be matrices with entries in a commutative ring R. If C is obtained from A by replacing row k with "row k plus r times row h", for some $h\neq k$ and some r in R, then $\det C = \det A$.*

Proof. Let B be the matrix obtained from X by replacing row k with r times row h. Let B' be the matrix obtained from A by replacing row k with row h. From Lemma 87, we have

$$\det A + \det B = \det C.$$

So we have to show that $\det B = 0$. But B is obtained from B' by multiplying by r a single row of B'; thus $\det B = r\cdot\det B'$ by Lemma 93; and B' has two equal rows (namely, rows h and k), so $\det B' = 0$ by Lemma 88. □

Proposition 95. *Let R be any ring. The three elementary matrices have determinant*

$$\det E_{i,j} = -1, \quad \det E_i(r) = r, \quad \det E_{i,j}(r) = 1.$$

Proof. Apply Lemmas 91, 93, and 94 to $A = I_m$, which has determinant 1. □

Lemma 96. *Let E be an elementary $m \times m$ matrix. Let B be any $m \times m$ matrix. Then*

$$\det(EB) = \det E \cdot \det B.$$

Proof. We simply check in all three cases. If $E = E_{ij}$, then EB is obtained from B by swapping rows i and j, so indeed $\det E = -1$ by the previous proposition, and $\det(EB) = -\det B$ by Lemma 91. If $E = E_i(r)$ with r invertible, then $\det E = r$ by the previous proposition, and $\det(EB) = r \det B$ by 93. Finally, if if $E = E_i(z)$ with $z \neq 0$, then $\det E = 1$ and $\det(EB) = \det B$ by Lemma 94. □

Theorem 97 (Cauchy–Binet for fields)**.** *For any two $m \times m$ matrices A, B with entries in a field $\mathbb{F}$,*

$$\det(AB) = \det A \cdot \det B = \det(BA).$$

Moreover, A is invertible if and only if $\det A \neq 0$.

Proof. By Gauss–Jordan theorem, we can write

$$E_k\, E_{k-1} \cdots E_k\, A = M$$

for some elementary matrices E_h and some RRE square matrix M. By Theorem 36 , there are two cases:

(a) if $E_k\, E_{k-1} \cdots E_2\, E_1\, A = I_n$, then $A = E_1^{-1}\, E_2^{-1} \cdots E_{k-1}^{-1}\, E_k^{-1}$, so A is a product of elementary matrices. In this case by repeated application of Lemma 96 we get

$$\begin{aligned}
\det(AB) &= \det(E_1^{-1}E_2^{-1}\cdots E_k^{-1}B) = \det(E_1^{-1})\det(E_2^{-1}\cdots E_k^{-1}B) = \dots \\
&= \det(E_1^{-1})\det(E_2^{-1})\cdots\det(E_k^{-1})\det B \\
&= \det(E_1^{-1}E_2^{-1})\det(E_3^{-1})\cdots\det(E_k^{-1})\det B = \dots \\
&= \det(E_1^{-1}E_2^{-1}\cdots E_k^{-1})\det(B) \qquad = \det A \det B.
\end{aligned}$$

(b) if $E_k\, E_{k-1} \cdots E_2\, E_1\, A$ has a row of zeroes, then so does $(E_k\, E_{k-1} \cdots E_2\, E_1\, A)B$, because of how the row-by-column product is defined. So by Lemma 88 $\det((E_k\, E_{k-1} \cdots E_2\, E_1\, AB) = 0$. But again by repeated application of Lemma 96 we get

$$0 = \det(E_k\, E_{k-1} \cdots E_2\, E_1\, AB) = \det E_1 \det E_2 \cdots \det E_k \det(AB).$$

Since on the right hand side the first k factors of the are nonzero, it follows that $\det(AB) = 0$. Analogously,

$$0 = \det(E_k\, E_{k-1} \cdots E_2\, E_1\, A) = \det E_1 \det E_2 \cdots \det E_k \det A,$$

which implies $\det A = 0$.

So the formula $\det(AB) = \det A \cdot \det B$ is proven. Since $\det A$ and $\det B$ live in a commutative ring R, it follows that

$$\det(BA) = \det B \cdot \det A = \det A \cdot \det B.$$

As for the second claim: We just saw that when A is not invertible, $\det A = 0$. Conversely, when A is invertible, since $\det A \det A^{-1} = \det(AA^{-1}) = 1$ it must be $\det A \neq 0$. □

Remark 98. The proof of Theorem 97 gives an algorithmic, **explicit way to compute the determinant of a square matrix A with entries over a field**. The idea is to reduce A to a matrix in RRE form via the Gauss–Jordan algorithm. If the resulting RRE matrix has a row of

zeroes at the bottom, then $\det A = 0$. If instead the resulting RRE matrix is the identity, then keeping track of the performed operations we can write

$$E_\ell E_{\ell-1} \cdots E_1 E_0 A \;=\; I_m,$$

or in other words

$$A = (E_\ell E_{\ell-1} \cdots E_1 E_0)^{-1} = (E_0)^{-1}(E_1)^{-1} \cdots (E_\ell)^{-1},$$

whence by Theorem 97 we get

$$\det A = (\det E_0)^{-1}(\det E_1)^{-1} \cdots (\det E_\ell)^{-1}.$$

So what do we do to compute the determinant, if the ring R is not a field, e.g. if $R = \mathbb{Z}$? For $R = \mathbb{Z}$ there is a nice strategy available: just view the entries as elements of $\mathbb{Q}$, and compute the determinant over $\mathbb{Q}$ via matrix reduction. In the end, any fractions you might have generated along the way will simplify, and you will get an integer, because if all entries of a matrix are integers, so is $\det A$. This "trick" of finding a larger field your ring is contained into, is however not always possible: it only works for *domains*, i.e. rings where $rs = 0$ implies that either $r = 0$ or $s = 0$. (Since fields are domains, and subrings of domains are themselves domains, it won't possible for a non-domain R to be contained in any field.)

Remark 99. Theorem 97 does extend also to matrices with entries over any commutative ring, but one has to be careful in how to generalize the second part: For example, in the world of 1×1 matrices with entries in $\mathbb{Z}$, 2 is not invertible, but its determinant is 2, not 0. The correct generalization of the second part is that "A is invertible if and only if $\det A$ is invertible", cf. Theorem 106.

2.3 Cofactor expansions, adjoint matrices, and Cauchy–Binet for rings

To finish this chapter, it remains to reconcile the definition of determinant given here with the more general, but ambiguous one, given in Anton's book. We do this for all matrices with entries over a commutative ring.

Theorem 100. *For any square matrix A with entries in a commutative ring, one has*

$$\det A = \det A^\top.$$

First proof. The statement is obviously true for 1×1 matrices, and easily shown for 2×2 matrices, for which we see already the importance of the commutativity assumption:

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - cb = ad - bc = \det \begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

For $m \geq 3$, we will show the equivalent statement that the determinant of A (defined as a cofactor expansion along the first column) equals also the cofactor expansion along the first row. In other words, we will prove that

$$\sum_{i=1}^{m} (-1)^{i+1} \cdot a_{i,1} \cdot M^A_{i,1} = \sum_{j=1}^{m} (-1)^{1+j} \cdot a_{1,j} \cdot M^A_{1,j},$$

where $M^A_{1,j}$ is the determinant of the $(m-1)\times(m-1)$ matrix obtained from A by removing the 1-st row and the j-th column. In fact, the $i=1, j=1$ option gives the same summand on both sides, so the equation above can be simplified: we have to show that

$$\sum_{i=2}^{m}(-1)^{i+1}\cdot a_{i,1}\cdot M^A_{i,1}=\sum_{j=2}^{m}(-1)^{1+j}\cdot a_{1,j}\cdot M^A_{1,j}. \tag{1}$$

The idea to complete the proof is now to fix $i,j\geq 2$, and *compare the terms containing $a_{i,1}a_{1,j}$ on both sides of Equation 1* . I took this proof from ximera.osu.edu/oerlinalg.

Fix $i\geq 2$ and focus only of the term $a_{i,1}$. In the left hand side, it appears only in the term

$$(-1)^{i+1}a_{i,1}M^A_{i,1}=(-1)^{i+1}a_{i,1}\det A_{-i,-1}=(-1)^{i+1}a_{i,1}\det(A_{-i,-1})^{\top},$$

where the last step is justified by the fact that $A_{-i,-1}$ has size $(m-1)\times(m-1)$. In other words, we are allowed to evaluate $\det A_{-i,-1}$ by cofactor expansion on the first row. We just have to be careful about the subscripts: Since the first column of A was removed, the j-th column of A contains the $(j-1)$-st column of $A_{-i,-1}$. So on the left hand side of Equation 1, $a_{i,1}$ appears in an expression equal to

$$(-1)^{i+1}a_{i,1}\left[(-1)^{1+1}a_{1,2}\det(A_{-i,-1})_{-1,-1}+\ldots+x(-1)^{1+(j-1)}a_{i,1}\det(A_{-i,-1})_{-1,-(j-1)}+\ldots\right]$$

and more importantly, for fixed $j\geq 2$, $a_{i,1}a_{1,j}$ appears only in the term

$$(-1)^{i+1}a_{i,1}\left[a_{1,j}(-1)^{1+(j-1)}\det(A_{-i,-1})_{-1,-(j-1)}\right],$$

which can be rewritten as

$$(-1)^{i+j+1}a_{i,1}a_{1,j}\det(A_{-i,-1})_{-1,-(j-1)}. \tag{2}$$

This complete our analysis of the left hand side of Equation 1. Now let us check how $a_{i,1}a_{1,j}$ appears in the right hand side of Equation 1. First of all, $a_{1,j}$ appears only in the quantity

$$(-1)^{1+j}a_{1,j}M^A_{1,j}=(-1)^{1+j}a_{1,j}\det A_{-1,-j}.$$

If we can expand $\det A_{-1,-j}$ the usual way (along the first column), we get that the quantity above equals

$$(-1)^{1+j}a_{1,j}\left[(-1)^{1+1}a_{2,1}\det(A_{-1,-j})_{-1,-1}+\ldots+(-1)^{(i-1)+1}a_{i,1}\det(A_{-1,-j})_{-(i-1),-1}+\ldots\right].$$

More importantly, for fixed $i\geq 2$, $a_{i,1}a_{1,j}$ appears only in the term

$$(-1)^{1+j}a_{1,j}\left[a_{i,1}(-1)^{(i-1)+1}\det(A_{-1,-j})_{-(i-1),-1}\right],$$

which simplifies to

$$(-1)^{i+j+1}a_{i,1}a_{1,j}\det(A_{-1,-j})_{-(i-1),-1}. \tag{3}$$

If we compare equations (2) and (3), we see that we have obtained the same quantities: In fact, let B the matrix obtained by deleting the first and the ith row of A, and the first and the jth column of A. According to the order of the deletions, if we first delete row i and column 1, then in the next step we have to delete row 1 and column $j-1$ (!), so $B=A_{-i,-1})_{-1,-(j-1)}$. If instead we first delete row 1 and column j, then in a second round we should delete column 1 and the row now called $i-1$, so we see that $B=(A_{-1,-j})_{-(i-1),-1}$. □

Second proof, valid only for entries in a field. For matrices A with entries *in a field*, a much simpler proof that $\det A = \det A^\top$ is available, using the Gauss-Jordan reduction and Cauchy–Binet's theorem. First you verify that $\det E = \det E^\top$ for every elementary matrix E, which is super-easy because

$$E_{i,j}^\top = E_{i,j}, \quad E_i(r)^\top = E_i(r), \quad E_{i,j}(r)^\top = E_{j,i}(r).$$

Then we distinguish two cases: if A is not invertible, then neither is $A^\top$ (because $A^\top B = I_m$ would imply $B^\top A = I_m$), so both $\det A$ and $\det A^\top$ equal zero by Theorem 97. If instead A is invertible, then via Gauss-Jordan we can write

$$E_\ell E_{\ell-1} \cdots E_1 E_0 A \;=\; I_m,$$

whence by Cauchy-Binet's Theorem 97we get $\det A = (\det E_\ell \det E_{\ell-1} \cdots \det E_1 \det E_0)^{-1}$; but also, if we transpose both sides we get

$$A^\top E_0^\top E_1^\top \cdots E_{\ell-1}^\top E_\ell^\top = I_m,$$

whence by Cauchy-Binet's Theorem 97 we get $\det A^\top = (\det E_\ell^\top \det E_{\ell-1}^\top \cdots \det E_1^\top \det E_0^\top)^{-1}$. Since $\det E = \det E^\top$ for all elementary matrices, we conclude that also in this case $\det A^\top = \det A$. □

Corollary 101. *For matrices A with entries in any commutative ring R:*

(i) *If A has a zero column,* $\det A = 0$.
(ii) *If A has two identical columns,* $\det A = 0$.
(iii) *If C is obtained from A by multiplying all elements of some column by the same element r in R, then* $\det C = r \det A$.
(iv) *If C is obtained from A by replacing column k with "column k plus r times column h", for some $h \neq k$ and some r in R, then* $\det C = \det A$.

Proof. (i) From Theorem 100 and Lemma 88.
(ii) From Theorem 100 and Lemma 92.
(iii) From Theorem 100 and Lemma 93.
(iv) From Theorem 100 and Lemma 94.

□

Theorem 102 (Cofactor expansion along any column of row)**.** *For any square matrix A, for any fixed $j \in \{1, \ldots, m\}$, one has*

$$\det A = \sum_{i=1}^{m} (-1)^{i+j} \cdot a_{i,j} \cdot M_{i,j}^A,$$

where $M_{i,j}^A$, usually called "the (i,j)-minor of A"*, is the determinant of the $(m-1) \times (m-1)$ matrix obtained from A by removing the i-th row and the j-th column. The formula above is called "cofactor expansion along the j-th column of A". And also, for fixed $i \in \{1, \ldots, m\}$,*

$$\det A = \sum_{j=1}^{m} (-1)^{i+j} \cdot a_{i,j} \cdot M_{i,j}^A,$$

a formula called "cofactor expansion along the i-th row of A".

Proof. By performing $j-1$ swaps, move the column j to position 1. Let B be the matrix obtained. By Lemma 91, we know that $\det A^\top = (-1)^{j-1} \det B^\top$; so by Theorem 100

$$\det A = \det A^\top = (-1)^{j-1} \det B^\top = (-1)^{j-1} \det B.$$

But by definition,

$$\det B = (-1)^{i+1} b_{i,1} M^B_{i,1} = (-1)^{i+1} a_{i,j} M^A_{i,j}.$$

Putting together the previous two equations, we conclude:

$$\det A = (-1)^{j-1} \sum_{i=1}^{m} (-1)^{i+1} a_{i,j} M^A_{i,j} = \sum_{i=1}^{m} (-1)^{i+j-2} a_{i,j} M^A_{i,j} = \sum_{i=1}^{m} (-1)^{i+j} a_{i,j} M^A_{i,j}.$$

The second claim follows immediately from the first claim and the fact that $\det A = \det A^\top$. □

Inverse of a matrix via the adjoint

The best method to find the inverse of a matrix is certainly via the Gauss–Jordan reduction. However, since we made the effort to look into determinants and cofactor expansions, there is a formula of some interest, particularly if the ring R of the entries is not a field (and so the Gauss–Jordan reduction is unavailable).

Definition 103. Let A be an $m \times m$ matrix with entries in a commutative ring. The *cofactor matrix* C^A and its transpose, the *adjoint matrix* Adj^A, are defined as follows:

$$\mathrm{C}(A)_{i,j} \stackrel{\text{def}}{=} C^A_{i,j} = (-1)^{i+j} M^A_{i,j} = (-1)^{i+j} \det A_{-i,-j}$$

$$\operatorname{Adj}^A_{i,j} \stackrel{\text{def}}{=} C^A_{j,i} = (-1)^{i+j} M^A_{j,i} = (-1)^{i+j} \det A_{-j,-i}$$

Theorem 104. *For any $m \times m$ matrix A with entries in a commutative ring R,*

$$A \operatorname{Adj}^A = \det A \cdot I_m.$$

In particular, if $\det A$ is invertible in R, A has an inverse, namely, $(\det A)^{-1} \cdot \operatorname{Adj}^A$.

Proof. Let us call P the $m \times m$ matrix obtained from the row-by-column product $A \operatorname{Adj}^A$. By definition,

$$P_{i,j} = \sum_{k=1}^{m} a_{i,k} (\operatorname{Adj}^A)_{k,j} = \sum_{k=1}^{m} a_{i,k} C_{j,k} = \sum_{k=1}^{m} (-1)^{j+k} a_{i,k} M^A_{j,k}.$$

Now if $i = j$, we have

$$P_{i,i} = \sum_{k=1}^{m} (-1)^{i+k} a_{i,k} M^A_{i,k} = \det A$$

because of Theorem 102: It's the cofactor expansion along the i-th row of A. It remains to see that $P(i,j) = 0$ if $i \neq j$. That is completely not proven in Anton's book, but the point is this: consider the matrix B obtained from A by copy-pasting the i-th row into the j-th row of A. If we compute the determinant of B along the j-th row, by Theorem 102 we get

$$\det B = \sum_{k=1}^{m} (-1)^{j+k} b_{j,k} M^B_{j,k}.$$

But B and A are identical except for the j-th row, so $M^B_{j,k} = M^A_{j,k}$; and by the way B was constructed, $b_{j,k} = b_{i,k} = a_{i,k}$. So the expression above becomes

$$\det B = \sum_{k=1}^{m} (-1)^{j+k} a_{i,k} M^A_{j,k}.$$

Now the right hand side is precisely the quantity we want to be zero. But since B has two equal rows, by Lemma 92 we have $\det B = 0$. □

Theorem 105 (Cauchy–Binet for entries in $\mathbb{Z}$)**.** *For any two $m \times m$ matrices A, B with entries in the ring $\mathbb{Z}$,*

$$\det(AB) = \det A \cdot \det B = \det(BA).$$

Moreover, A is invertible if and only if $\det A$ *is* ± 1.

Proof. Since the entries of A are in $\mathbb{Z}$, they are also in $\mathbb{Q}$. So by viewing A, B as elements of $M_{m,m}(\mathbb{Q})$, Theorem 97 tells us that inside $\mathbb{Q}$

$$\det(AB) = \det A \cdot \det B = \det(BA).$$

A priori, as we said, this is an equality in $\mathbb{Q}$. But both sides of the equality above can be computed from the entries of A via sums and products; so this is actually an equality of elements in $\mathbb{Z}$. This shows the first part. Now if A is invertible within the world of matrices with entries in $\mathbb{Z}$, let B be its inverse: both $\det A$ and $\det B$ are integers, and by Cauchy–Binet

$$\det A \det B = \det(AB) = \det I_m = 1.$$

But the only integer divisors of 1 are ± 1. So either $\det A$ and $\det B$ are both 1, or they are both -1. As for the converse, this is given by Theorem 104: if $\det A = \pm 1$, then Theorem 104 constructs an inverse matrix all of whose elements are in $\mathbb{Z}$, because all entries of Adj^A are integers, and we only need to divide them by either 1 or -1. □

This gives us the idea to extend Cauchy–Binet to arbitrary commutative rings.

Theorem 106 (Cauchy–Binet for fields)**.** *For any two $m \times m$ matrices A, B with entries in a ring R,*

$$\det(AB) = \det A \cdot \det B = \det(BA).$$

Moreover, A is invertible if and only if $\det A$ *is invertible in R.*

Sketch of proof. For the first part, the idea is to use some abstract algebra, plus the proof of Cuachy–Binet for fields that we already have. One calls *domain* a commutative ring R in which if a product is zero, then one of the factors must be zero as well. For example, $\mathbb{Z}$ is a domain; and every field is a domain. It can be seen that if R is a domain, so is the ring of polynomials with coefficients in R and one variable x. We denote such ring by $R[x]$. Iterating the argument, if a ring R is a domain, so is its ring of polynomials with coefficients in R and n variables. We denote such ring by $R[x_1, \dots, x_n]$. For example, $\mathbb{Z}[x_1, \dots, x_n]$ is a domain. Now, given a domain D, there is a canonical way to construct the smallest field containing D; the idea is to follow the construction of $\mathbb{Q}$ from $\mathbb{Z}$. That is, we consider the set of fractions

$$\mathbb{F}_D \stackrel{\text{def}}{=} \left\{ \frac{a}{b} : a \in D, b \in D, b \neq 0 \right\}$$

with the convention that we consider two fractions $\frac{a}{b}, \frac{a'}{b'}$ equal if $ab' = a'b$. This set is called *field of fractions of D*, and one can see that this is a field: Exactly like for the fractions you

know, the inverse of an element $\frac{a}{b}$, with $a \neq 0$, is simply $\frac{b}{a}$. Now, the premise is over: We want to show Cauchy–Binet for $m \times m$ matrices with entries in a commutative ring R. The idea is to show it as an identity of polynomials with coefficients in $\mathbb{Z}$. That is, we consider the entries $a_{1,1}, \ldots, a_{m,m}$ as m^2 variables and to prove Cauchy–Binet as an identity for polynomials in $\mathbb{Z}[a_{1,1}, \ldots, a_{m,m}]$. But if we set $D = \mathbb{Z}[a_{1,1}, \ldots, a_{m,m}]$, then D is a domain for the reasons mentioned above (it is a ring of polynomials with $n = m^2$ variables over the domain $\mathbb{Z}$), so we can form its field of fractions $\mathbb{F}_D$. Inside $\mathbb{F}_D$, the identity by Cauchy–Binet will be true, because we have proved it for matrices with entry in any field. But then since this both sides of this identity are in D, the identity will hold for D as well. This establishes $\det(AB) = \det A \cdot \det B$. As for the second part: if A is invertible, then $\det A \det A^{-1} = \det(AA^{-1} = \det I_m = 1$, so $\det A$ is invertible. The converse implication is given by Theorem 104. □

Lemma 107. *Let A be a matrix with entries in a commutative ring. If A is upper triangular, and $i \leq j+1$, then $A_{-i,-j}$ is upper triangular. Moreover, when $i < j$, $\det A_{-i,-j} = 0$.*

Proof. We only have to prove the first statement; the second one follows by transposing. Set $M \stackrel{\text{def}}{=} A_{-i,-j}$. Then

$$m_{h,k} = \begin{cases} a_{h,k} & \text{if } h < i \text{ and } k < j; \\ a_{h,k+1} & \text{if } h < i \text{ and } k \geq j; \\ a_{h+1,k} & \text{if } h \geq i \text{ and } k < j; \\ a_{h+1,k+1} & \text{if } h \geq i \text{ and } k \geq j. \end{cases}$$

Now suppose that $i \leq j+1$ and $h > k$. This rules out the case "$h < i$ and $k \geq j$", because we would have

$$h < i \leq j+1 \leq k+1,$$

which (since they are integers) is the same as $h \leq k$, contradicting $h > k$. But if $h > k$, all three statements

$$h > k, \qquad h+1 > k, \qquad h+1 > k+1$$

are true; so either way, looking at the formula above, $m_{h,k}$ is copy-pasted from an entry of A with row-index larger that the column-index. But in the upper triangular matrix A, any element below the diagonal is zero. So $m_{h,k} = 0$. This proves that $M \stackrel{\text{def}}{=} A_{-i,-j}$ is upper triangular. Now suppose that $i < j$. The $(i+1)$-st row of A begins with at least $i+1$ zeroes; since the i-th row of A is deleted, and $j > i$, all of these $i+1$ zeroes will be in the i-th row of M, which therefore has a 0 on the main diagonal. But the determinant of an upper triangular matrix is the product of its diagonal elements. If one of them is zero, the product is zero. □

Note that the assumptions $i \leq j+1$ and $i < j$ are really necessary (and sharp bounds): The 3×3 matrix $E_{1,2}(1)$ is upper triangular, but if we remove row 3 and column 1, what we get is not upper triangular. And if from $E_{1,2}(1)$ we remove row 1 and column 1, what we get is I_2, which has determinant 1, not 0. As an exercise, you may show the "transpose statement": if B is lower triangular, and $i \geq j-1$, then $B_{-i,-j}$ is lower triangular; and if in addition $i > j$, then $\det B_{-i,-j} = 0$.

Theorem 108. *Let A be a square matrix with entries in a commutative ring.*
(1) *If A is upper triangular, so is the adjunct matrix* Adj^A.
(2) *If A is upper triangular and invertible, so is its inverse A^{-1}.*
(3) *If A is lower triangular, so is the adjunct matrix* Adj^A.
(4) *If A is lower triangular and invertible, so is its inverse A^{-1}.*
(5) *If A is diagonal, so is the adjunct matrix* Adj^A.
(6) *If A is diagonal and invertible, so is its inverse A^{-1}.*

Proof. All these statements follow from statement number (i). In fact (ii) follows from (i) via the formula $A^{-1} = (\det A)^{-1} \cdot \mathrm{Adj}^A$. Moreover, (iii) and (iv) are easily obtainable from (i) and (ii, respectively, by transposing. And then (v) follows from putting together (i) and (iii), while (vi) follows from putting together (ii) and (iv). So, let focus on statement (i).

Let A be upper triangular (i.e. $a_{i,j} = 0$ for all $i > j$). Recall that its cofactor matrix is

$$C(A)_{i,j} \stackrel{\text{def}}{=} C^A_{i,j} = (-1)^{i+j} M^A_{i,j} = (-1)^{i+j} \det A_{-i,-j}.$$

But when $i < j$, $A_{-i,-j}$ is upper triangular with determinant zero, because of Lemma 107. So when $i < j$, we have $C(A)_{i,j} = 0$, which means that the cofactor matrix is lower triangular. So its transpose, the adjunct matrix, is upper triangular. □

2.4 Consequences on linear systems with coefficients in an infinite field

We introduce the determinant of square matrices. We defined it by cofactor expansion along the first column, but it turns out that can be computed by cofactor expansion along *any* row or column of your convenience[4]. In addition to the four properties we claimed at the beginning, Cauchy–Binet's theorem revealed a very important fact:

- $\det(AB) = \det A \cdot \det B$;
- A is invertible if and only if $\det A$ is invertible in R.

When R is a field, the second statement becomes, "A is invertible if and only if $\det A \neq 0$."

Now let us consider again the situation of a linear system $A\mathbf{x} = \mathbf{b}$, with $A, \mathbf{b}$ matrices with entries in an **infinite** field, like $\mathbb{Q}$ or $\mathbb{R}$ or $\mathbb{C}$. (There are also finite fields: A field with two elements is given in Example 3).

Proposition 109. *Let A an $m \times m$ matrix and $\mathbf{b}$ an $m \times 1$ matrix, both with entries in an* infinite *field $\mathbb{F}$.*

- *If $\det A \neq 0$, then $A\mathbf{x} = \mathbf{b}$ has a unique solution, namely, $\mathbf{x} = A^{-1}\mathbf{b}$.*
- *If $\det A = 0$ and $\mathbf{b} = \mathbf{0}$, then $A\mathbf{x} = \mathbf{b}$ has infinitely many solutions.*
- *If $\det A = 0$ and $\mathbf{b} \neq \mathbf{0}$, then $A\mathbf{x} = \mathbf{b}$ may have either infinitely many solutions, or none.*

In particular, a linear system over an infinite field can only have 0 or 1 or $+\infty$ solutions.

Proof. Part (i) is clear from Theorem 97, because if $\det A \neq 0$, then A is invertible.
Part (ii) follows from Theorem 44: Since $\det A = 0$, the $A\mathbf{x} = \mathbf{b}$ has more than one solution. So it has a solution $\mathbf{x} \neq 0$. But then for any element $f \in \mathbb{F}$, $f \cdot \mathbf{x}$ is also a solution, because

$$A(f\mathbf{x}) = f(A\mathbf{x}) = f\mathbf{0} = \mathbf{0}.$$

So there at least as many solutions as elements of $\mathbb{F}$: Using that $\mathbb{F}$ is infinite, we conclude.
As for Part (iii): Some systems, like the one below, have zero solutions

$$\begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$ □

We claim though that if $A\mathbf{x} = \mathbf{b} \neq \mathbf{0}$, with $\det A = 0$, has one solution, then automatically it must have infinitely many others. In fact, let $\mathbf{y}$ be any of the infinitely many solution of $A\mathbf{x} = \mathbf{0}$. Then $\mathbf{y} + \mathbf{z}$ is another solution of $A\mathbf{x} = \mathbf{b}$, because

$$A(\mathbf{y} + \mathbf{z}) = A\mathbf{y} + A\mathbf{z} = \mathbf{0} + \mathbf{b} = \mathbf{b}.$$

Since $\mathbf{y} + \mathbf{z} \neq \mathbf{y}' + \mathbf{z}$ for $\mathbf{y} \neq \mathbf{y}'$, we conclude.

[4] This makes some previous result look simple, for example, that the determinant of a matrix equals the determinant of its transpose, or that swapping consecutive rows brings a minus sign to the determinant. Technically this is not a new proof though, because we used Lemma 91 and Theorem 100 for the proof of Theorem 102.

3 General vector spaces

Definition 110 (Vector spaces)**.** (cf. Anton, Ch. 4.1) Let $\mathbb{F}$ be a field. An $\mathbb{F}$-*vector space* is a set V together with an operation $(v, w) \mapsto v + w$ from $V \times V$ to V (called "*sum*") and an operation $(\lambda, v) \mapsto \lambda v$ from $\mathbb{F} \times V$ to V (called "*scalar multiplication*") that satisfy the following axioms:

(VS1) $+$ is associative. That is, for all u, v, w in V, $(u + v) + w = u + (v + w)$.

(VS2) $+$ is commutative. That is, for all v, w in V, $v + w = w + v$.

(VS3) $+$ has a unique neutral element. That is, $\exists! z$ in V such that for all v in V, $v + z = v$. We shall call this z "0".

(VS4) Every element has a unique additive inverse. That is, for all v in V $\exists! y$ in V such that $v + y = 0$. We shall call this y "$-v$".

(VS5) For all v in V and for all λ, μ in $\mathbb{F}$, $\lambda(\mu v) = (\lambda\mu)v$.

(VS6) For all v in V and for all λ, μ in $\mathbb{F}$, $(\lambda + \mu)v = \lambda v + \mu v$.

(VS7) For all v, w in V and for all λ in $\mathbb{F}$, $\lambda(v + w) = \lambda v + \lambda w$.

(VS8) For all v in V, $1v = v$.

The elements of $\mathbb{F}$ are usually called *scalars*, whereas the elements of V are usually referred to as *vectors*.

Remark 111. As with the "ring" axioms, in (VS3) and/or in (VS4) one might as well delete the word 'unique': The resulting definitions would be equivalent.

Non-Example 112. With respect to the usual sum and product of fractions, $\mathbb{Z}$ is not a $\mathbb{Q}$-vector space. The reason is hidden not in the eight axioms, but in the first sentence: The operation $(\lambda, v) \mapsto \lambda v$ should go from $\mathbb{Q} \times \mathbb{Z}$ to $\mathbb{Z}$, but $\frac{1}{2} \cdot 3$ is not in $\mathbb{Z}$.

Example 113. Any field $\mathbb{F}$ has a natural structure of vector space over any subfield $\mathbb{G} \subseteq \mathbb{F}$. In fact, there is an obvious scalar multiplication

$$\begin{array}{rl} \mathbb{G} \times \mathbb{F} & \longrightarrow \mathbb{F} \\ (\lambda, \alpha) & \mapsto \lambda\alpha. \end{array}$$

In particular (by choosing $\mathbb{G} = \mathbb{F}$), any field $\mathbb{F}$ is itself an $\mathbb{F}$-vector space.

Example 114. With the usual operations of sum and product, $\mathbb{R}$ is an $\mathbb{R}-$vector space and also a $\mathbb{Q}$-vector space. However, $\mathbb{Q}$ is a vector space over $\mathbb{Q}$, but not over $\mathbb{R}$. The reason is that the product of one element in $\mathbb{Q}$ and one in $\mathbb{R}$ is necessarily in $\mathbb{R}$, but not necessarily in $\mathbb{Q}$.

Remark 115. We saw that there exists a field with two elements. Since a field must contain a 0 and a 1, and (according to the convention of these notes) they must be different, the smallest field has two elements. That field is automatically a vector space over itself. However, it is not the smallest vector space! The vector space axioms do not require the existence of a special element different than 0 in V, so actually the set $\{0\}$ is a vector space over any field. Not every set is a vector space, though. For example, one can prove via field theory if a vector space has finite size, that size must be a prime power So, a six-element set cannot be given the structure of $\mathbb{F}$-vector space, no matter what you pick for $\mathbb{F}$, and no matter the operations.

Example 116. $\mathbb{R}^n$ is an $\mathbb{R}$-vector space. More generally, the Cartesian product $\mathbb{F}^n$ is an $\mathbb{F}$-vector space with the usual coordinate-wise sum and multiplication of (all entries of) a vector by a scalar: $\lambda \cdot (x_1, \ldots, x_n) \stackrel{\text{def}}{=} (\lambda x_1, \ldots, \lambda x_n)$.

Example 117. Let $\mathbb{F}$ be any field. Let m, n be positive integers. The set of $m \times n$ matrices with entries in $\mathbb{F}$ is an $\mathbb{F}$-vector space, with the usual entry-wise sum and multiplication of (all entries of) a matrix by a scalar.

Example 118. Let $\mathbb{F}$ be a field. Let d be a natural number. Define

$$P(d, \mathbb{F}) \stackrel{\text{def}}{=} \{a_0 + a_1 x + \ldots + a_d x^d \ : \ a_i \in \mathbb{F}\}.$$

This is the set of all *polynomials of degree* $\leq d$ *with coefficients in* $\mathbb{F}$, plus the polynomial 0 (which by some conventions does not have a degree, by some other conventions has degree -1). This $P(d, \mathbb{F})$ is an $\mathbb{F}$-vector space: Any two polynomials of degree $\leq d$ sum up to a polynomial of degree $\leq d$, or possibly the zero polynomial. And if we multiply a polynomial of degree k by a constant, we get either the zero polynomial if the constant is zero, or a polynomial of degree k if the constant is nonzero.

Example 119. Let $I \subseteq \mathbb{R}$ be an arbitrary subset of the real line (typically $I = \mathbb{N}$, or I an interval). The set of functions $f : I \longrightarrow \mathbb{R}$ is an $\mathbb{R}$-vector space with the operations

$$(f + g)(x) \stackrel{\text{def}}{=} f(x) + g(x) \qquad \text{and} \qquad (\lambda f)(x) \stackrel{\text{def}}{=} \lambda \cdot f(x).$$

Non-Example 120. (cf. Anton, Ch. 4.1, Ex. 7) On $\mathbb{R}^2$, with the usual sum, consider the following "exotic scalar multiplication" $\circ : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$:

$$\lambda \circ (x, y) \stackrel{\text{def}}{=} (\lambda x, 0).$$

This satisfies all axioms from (VS1) to (VS7), but not (VS8). So , $\mathbb{R}^2$ is not a vector space with respect to componentwise sum and $\circ$.

Non-Example 121. On $\mathbb{R}^2$, with the usual sum, consider the "exotic scalar multiplication" $\bullet : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$,

$$\lambda \circ (x, y) \stackrel{\text{def}}{=} \lambda^2 x, \lambda^2 y).$$

This satisfies all axioms, except (VS6).

Lemma 122 (Cancellation)**.** *Let $\mathbb{F}$ be a field. Let V be an $\mathbb{F}$-vector space. For all $u, v, w \in V$, if $u + v = u + w$ then $v = w$.*

Proof. Add $-u$ to both sides of the equation $u + v = u + w$. □

Proposition 123. *(cf. Anton, Theorem 4.1.1) Let $\mathbb{F}$ be a field. Let V be an $\mathbb{F}$-vector space. For all $v \in V$ and for all $\lambda \in \mathbb{F}$, one has:*
(1) $0v = 0$.
(2) $\lambda 0 = 0$.
(3) $(-1)v = -v$.
(4) *if* $\lambda v = 0$, *then either* $\lambda = 0$ *or* $v = 0$.

Proof. (1) Follows from $0v + 0v = (0 + 0)v = 0v$, via the Cancellation lemma.
(2) Follows from $\lambda 0 = \lambda(0 + 0) = \lambda 0 + \lambda 0$, via the Cancellation lemma.
(3) Follows from $(-1)v + v = (-1)v + 1v = (-1 + 1)v = 0v = 0$ by the item above.
(4) If $\lambda \neq 0$, since $\mathbb{F}$ is a field, λ is invertible. So we can multiply by λ^{-1} the equation $\lambda v = 0$ obtaining $1v = \lambda^{-1} 0$. Or in other words, $v = 0$. □

3.1 Subspaces, spans, linearly independent sets and bases

Definition 124. A *subspace* of an $\mathbb{F}$-vector space V is a nonempty subset $W \subseteq V$ that is a vector space with respect to the same operations (restricted to W).

Proposition 125. *(cf. Anton, Theorem 4.2.1) Let V be an $\mathbb{F}$-vector space. Let W be a nonempty subset of V. Then W is a subspace of $V \iff W$ satisfies*

(SS1) for each w_1, w_2 in W, for each λ in $\mathbb{F}$, the element $w_1 + \lambda w_2$ is in W.

Proof. The '$\Rightarrow$' direction is clear. For the opposite: the only interesting axioms to verify are (VS3) and (VS4), as all others are trivially satisfied. Since W is nonempty, it has a vector w in it. By choosing $\lambda = 0$ and $w_1 = w_2 = w$, we see that 0 is in W, so W satisfies (VS3). But then by choosing $w_1 = 0$ and $\lambda = -1$, we see that for any w_2 in W, the element $-1 \cdot w_2$ is also in W. But by Proposition 123, item (3), this means that for any w_2 in W, the element $-w_2$ is also in W. So W satisfies (VS4) as well. □

Proposition 126. *(cf. Anton, Theorem 4.2.4) The solution set of a homogeneous linear system $A\mathbf{x} = \mathbf{0}$ of m equations in n real unknowns, is a subspace of $\mathbb{R}^n$.*

Proof. Let us use Proposition 125. Let $\mathbf{x}$ and $\mathbf{y}$ be two solutions. This means $A\mathbf{x} = \mathbf{0}$ and $A\mathbf{y} = \mathbf{0}$. But then, for all λ in $\mathbb{R}$, $\mathbf{x} + \lambda\mathbf{y}$ is also a solution, because

$$A(\mathbf{x} + \lambda\mathbf{y}) = A\mathbf{x} + \lambda A\mathbf{y} = \mathbf{0} + \lambda \cdot \mathbf{0} = \mathbf{0}.$$ □

Note that the word "homogeneous" is crucial. If $A\mathbf{x} = \mathbf{b}$ and $A\mathbf{y} = \mathbf{b}$, then $A(\mathbf{x} + \lambda\mathbf{y}) = \mathbf{b} + \lambda \cdot \mathbf{b}$, which is typically not equal to $\mathbf{b}$ when $\mathbf{b} \neq \mathbf{0}$.

Definition 127. Let $\mathbb{F}$ be a field. Let V be an $\mathbb{F}$-vector space. Let $\emptyset \subsetneq X \subseteq V$ be an arbitrary, nonempty set. A *linear combination of elements of X over $\mathbb{F}$* is an expression of the type

$$a_1 x_1 + \ldots + a_n x_n,$$

for some $n \in \mathbb{N}$, for some $a_1, \ldots, a_n$ in $\mathbb{F}$, and for some $x_1, \ldots, x_n$ in X. A *trivial* linear combination is one where all the a_i's are zero.

Theorem 128. *(cf. Anton, Theorem 4.2.3) Let $\mathbb{F}$ be a field. Let V be an $\mathbb{F}$-vector space. Let $\emptyset \subsetneq X \subseteq V$. The set of all linear combinations of elements of X is a subspace of V, called* $\mathrm{span}(X)$*. It is the smallest subspace of V that contains X.*

Proof. Let us define $\mathrm{span}(X) \stackrel{\text{def}}{=} \{a_1 x_1 + \ldots + a_n x_n \,:\, n \in \mathbb{N}, a_i \in \mathbb{F}, x_i \in X\}$. Let us prove that this is a subspace using Proposition 125. Let λ be in $\mathbb{F}$. Let w_1, w_2 be in $\mathrm{span}(X)$. This means that $w_1 = a_1 x_1 + \ldots + a_\ell x_\ell$ and $w_2 = b_1 y_1 + \ldots + b_m y_m$, for some a_i, b_j in $\mathbb{F}$ and some x_i, y_j in X. If we set $n \stackrel{\text{def}}{=} \ell + m$, and for all $j \in \{1, \ldots, m\}$ we also set $x_{\ell+j} \stackrel{\text{def}}{=} y_j$ and $a_{\ell+j} \stackrel{\text{def}}{=} \lambda b_j$, we get

$$\begin{aligned} w_1 + \lambda w_2 &= a_1 x_1 + \ldots + a_\ell x_\ell + \lambda b_1 y_1 + \ldots + \lambda b_m y_m \\ &= a_1 x_1 + \ldots + a_\ell x_\ell + a_{\ell+1} x_{\ell+1} + \ldots + a_{\ell+m} x_{\ell+m} \\ &= a_1 x_1 + \qquad \ldots \qquad + a_n x_n, \end{aligned}$$

which means that $w_1 + \lambda w_2$ is also in $\mathrm{span}(X)$. So $\mathrm{span}(X)$ is a subspace. Now, clearly $X \subseteq \mathrm{span}(X)$, as any x in X can be written as $1 \cdot x$. Also, any subspace that contains X, must contain all expressions of the type $a_1 x_1 + \ldots + a_n x_n$, if the a_i's are chosen in $\mathbb{F}$ and the x_i's are chosen in X; which tells us that the smallest subspace containing X is $\mathrm{span}(X)$. □

Definition 129. Let $\mathbb{F}$ be a field. Let V be an $\mathbb{F}$-vector space. Let $X \subseteq V$ be an arbitrary set of vectors (possibly infinite, possibly empty). We say that X is:

- *linearly independent (LI)*, if for any $n \in \mathbb{N}$, for any vectors $x_1, \ldots, x_n$ in X and for any scalars $a_1, \ldots, a_n$ in $\mathbb{F}$, the following holds:

$$\text{if } a_1x_1 + \ldots + a_nx_n = 0, \quad \text{then } a_1 = a_2 = \ldots = a_n = 0.$$

  In other words, X is LI if and only if the only linear combination of finitely many elements of X that equals zero, is the trivial one.
- *a set of generators for V*, if for every v in V one can find finitely many elements $a_1, \ldots, a_n$ of $\mathbb{F}$ and elements $x_1, \ldots, x_n$ of $X \subseteq V$ such that

$$v = a_1x_1 + \ldots a_nx_n.$$

  In other words, if every element of V can be written as a linear combination of finitely many elements of X.
- *a basis for V*, if it is both LI and a set of generators.

Remark 130. "X is a set of generator for W" is synonymous with "$\text{span}(X) \supseteq W$", not with "$\text{span}(X) = W$". We usually say that "$X$ spans W" whenever "$\text{span}(X) = W$". Beware of the expression "X generates W", because some authors use it in the sense of "$\text{span}(X) = W$" and some others in the sense of "$\text{span}(X) \supseteq W$". We shall avoid the expression altogether.

Example 131. (cf. Anton, Theorem 4.3.2)

- The empty set is LI, but it is not a set of generator of any subspace.
- A one-element set $X = \{x\}$ is LI if and only if $v \neq 0$.
- A two-element set is LI if and only if neither vector is a scalar multiple of the other.
- Any set (finite or not) $X \subseteq V$ that contains the zero vector, is not LI. In fact, $\lambda 0 = 0$ for all values of $\lambda \in \mathbb{F}$, and not just for $\lambda = 0$.

Proposition 132. *Let V be an $\mathbb{F}$-vector space. Let $X \subseteq Y \subseteq V$.*
(1) *If Y is LI, so is X.*
(2) *If X is a set of generators for V, so is Y.*

Proof. Both items follow immediately from the observation that if $X \subseteq Y \subseteq V$, then any linear combination of elements of X is also a linear combination of elements of Y. □

Note: Anton's book defines the three notions (LI, set of generators, basis) only when X is finite. Here is a proof that the definitions of Anton's book agree with ours, when we restrict ourselves to the finite case:

Proposition 133. *Let $X = \{x_1, \ldots, x_m\}$ be a finite set in an $\mathbb{F}$-vector space V. The following two statements are equivalent to one another:*
(i) *X is LI;*
(ii) *the only scalars $a_1, \ldots, a_m$ in $\mathbb{F}$ for which $a_1x_1 + \ldots + a_mx_m = 0$ are $a_1 = a_2 = \ldots = a_m = 0$.*
Moreover, the following two statements are equivalent to one another:
(a) *X is a set of generators for V.*
(b) *for each v in V, there are scalars $a_1, \ldots, a_m$ in $\mathbb{F}$ for which $a_1x_1 + \ldots + a_mx_m = v$.*
Finally the following two statements are equivalent to one another:
(1) *X is a basis for V;*
(2) *for each v in V there is a unique choice of m scalars $a_1, \ldots, a_m$ in $\mathbb{F}$ such that $a_1x_1 + \ldots + a_mx_m = v$.*

Proof. '(i) ⇒ (ii)' is obvious: the LI definition works for all n, so in particular with $n = m$.
'(ii) ⇒ (i)': Suppose $n \leq m$, and $a_1x_1 + \ldots + a_nx_n = 0$. Then $a_1x_1 + \ldots + a_nx_n + 0x_{n+1} + \ldots + 0x_m = 0$. But this is a linear combination of $x_1, \ldots, x_m$, so by the assumption $a_1 = \ldots = a_n = 0$.
'(a) ⇒ (b)': Up to relabeling, suppose v is a linear combination of $x_1, \ldots, x_n$, with $n < m$. Then we can also write $v = a_1x_1 + \ldots + a_nx_n + 0x_{n+1} + \ldots + 0x_m$, so v is also a linear combination of $x_1, \ldots, x_m$.
'(b) ⇒ (a)': this is obvious, take $n = m$. '(1) ⇒ (2)': Suppose $a_1x_1 + \ldots + a_mx_m = v = b_1x_1 + \ldots + b_mx_m$, with the a_i's and the b_i's in $\mathbb{F}$. Then

$$(a_1 - b_1)x_1 + \ldots + (a_m - b_m)x_m = 0.$$

But since $\{x_1, \ldots, x_m\}$ is LI, we must have $a_1 = b_1, \ldots, a_m = b_m$.
'(2) ⇒ (1)': clearly $\{x_1, \ldots, x_m\}$ is a set of generators. By contradiction, suppose $c_1x_1 + \ldots + c_mx_m = 0$ for some scalars $c_1, \ldots, c_m$ that are not all zero. Then

$$(a_1 + c_1)x_1 + \ldots + (a_m + c_m)x_m = (a_1x_1 + \ldots + a_mx_m) + (c_1x_1 + \ldots + c_mx_m) = v + 0 = v$$

is another way to write v as linear combination of $\{x_1, \ldots, x_m\}$. For these last two implications, see also Anton, Theorem 4.4.1. □

Example 134. Inside $V = \mathbb{R}^3$, an $\mathbb{R}$-vector space, let $X = \{\mathbf{e}_1 = (1, 0, 0), \mathbf{e}_2 = (0, 1, 0)\}$. This X is LI because none of the two vectors is a scalar multiple of the other. X is not a set of generators for $\mathbb{R}^3$, because any linear combination of $\mathbf{e}_1$ and $\mathbf{e}_2$ will have third coordinate zero. However, consider $W \stackrel{\text{def}}{=} \mathbb{R}^2 \times \{0\}$. This is a subspace of $\mathbb{R}^3$ (check!) that contains X. Since any element of the form $(a, b, 0)$ can be written as $(a, b, 0) = a(1, 0, 0) + b(0, 1, 0) = a\mathbf{e}_1 + b\mathbf{e}_2$, our X is a set of generators for W.

3.2 Different bases, same cardinality

Lemma 135 ('Plus-minus lemma')**.** *(cf. Anton, Theorem 4.5.3) Let V be an $\mathbb{F}$-vector space. Let X be a subset of V (finite or not).*
(1) *if X is LI, then for any $v \in V$ not in X,*

$$v \notin \operatorname{span}(X) \iff X \cup \{v\} \text{ is LI}.$$

(2) *if X is not LI, then there is an x in X such that*

$$\operatorname{span} X = \operatorname{span}(X - \{x\}),$$

where $X - \{x\}$ is the set obtained removing x from X.

Proof. (1), ⇐: Were v in $\operatorname{span}(X)$, we could find $b_1, \ldots, b_n \in \mathbb{F}$ and $x_1, \ldots, x_n$ in X such that

$$b_1x_1 + \ldots + b_nx_n = v.$$

But then the expression

$$b_1x_1 + \ldots + b_nx_n - v = 0$$

would contradict the assumption that $X \cup \{v\}$ is linearly independent.
(1), ⇒: Suppose that for some $a_1, \ldots, a_n, a_{n+1} \in \mathbb{F}$, and for some $x_1, \ldots, x_n$ from X,

$$a_1x_1 + \ldots + a_nx_n + a_{n+1}v = 0.$$

If $a_{n+1} = 0$ then the linear independence of X implies that all a_i's are zero, and we are done. If instead $a_{n+1} \neq 0$, dividing the expression above by $-a_{n+1}$ and setting $b_i \stackrel{\text{def}}{=} a_i(a_{n+1})^{-1}$ we obtain

$$b_1 x_1 + \ldots + b_n x_n = v,$$

which contradicts the assumption $v \notin \text{span}(X)$.

(2): By the assumption, we can find scalars $a_1, \ldots a_n$ not all zero and vectors $x_1, \ldots, x_n$ inside X such that

$$a_1 x_1 + \ldots + a_n x_n = 0.$$

For any i such that a_i is not zero, we can take $a_i x_i$ to the right hand side and divide by $-a_i$; the result is an equation that expresses x_i as a linear combination of the other vectors. □

Lemma 136 (Steinitz Exchange Lemma)**.** *Let $\mathbb{F}$ be a field. Let V be a $\mathbb{F}$-vector space. Suppose that $L = \{\ell_1, \ldots, \ell_m\}$ is a LI set in V, and $G = \{g_1, \ldots, g_n\}$ is a set of generators for V. Then for all $k \in \{0, \ldots, m\}$, one has $k \leq n$, and up to relabeling the elements of G, one has*

$$V = \text{span}(\ell_1, \ldots, \ell_k, g_{k+1}, \ldots, g_n).$$

In particular (choosing $k = m$), one has $m \leq n$.

Proof. By induction on k. The case $k = 0$ is clear. Suppose the claim holds for some $k < m$, and let us prove it for $k + 1$. Since

$$\ell_{k+1} \in V = \text{span}(\ell_1, \ldots, \ell_k, g_{k+1}, \ldots, g_n),$$

we can find elements $a_1, \ldots, a_n$ in $\mathbb{F}$ such that

$$u_{k+1} = \sum_{j=1}^{k} a_j \ell_j \; + \; \sum_{j=k+1}^{n} a_j g_j. \tag{4}$$

But since the ℓ_j's are linearly independent, at least one of $\{a_{k+1}, \ldots, a_n\}$ must be nonzero. This already implies $k + 1 \leq n$. Up to relabeling the elements $\ell_{k+1}, \ldots, \ell_n$, we will assume that $a_{k+1} \neq 0$. But then we can rewrite Equation 4 as

$$\ell_{k+1} - \sum_{j=1}^{k} a_j \ell_j \; - \; \sum_{j=k+2}^{n} a_j g_j = a_{k+1} g_{k+1},$$

and dividing by $a_{k+1} \neq 0$, we obtain that

$$g_{k+1} \text{ is in } \text{span}(\ell_1, \ldots, \ell_k, \ell_{k+1}, g_{k+2}, \ldots, g_n).$$

On the other hand, Equation 4 also tells us that

$$\ell_{k+1} \text{ is in } \text{span}(\ell_1, \ldots, \ell_k, g_{k+1}, g_{k+2}, \ldots, g_n).$$

So we conclude that

$$\text{span}(\ell_1, \ldots, \ell_k, \ell_{k+1}, g_{k+2}, \ldots, g_n) = \text{span}(\ell_1, \ldots, \ell_k, g_{k+1}, g_{k+2}, \ldots, g_n).$$

But our inductive assumption was that $V = \text{span}(\ell_1, \ldots, \ell_k, g_{k+1}, g_{k+2}, \ldots, g_n)$, and what we wanted to show is that $V = \text{span}(\ell_1, \ldots, \ell_k, \ell_{k+1}, g_{k+2}, \ldots, g_n)$. So we are done. □

Theorem 137 (Equal cardinality of bases)**.** *(cf. Anton's Theorems 4.5.1, 4.5.4, 4.5.5) Let V be a $\mathbb{F}$-vector space that has a finite set of generators.*
(1) *Any finite set of generators for V contains a basis of V.*
(2) *Any LI set in V is contained in a basis of V.*
(3) *Any two bases of V have the same number of elements, which we shall call "$\dim_{\mathbb{F}} V$".*
(4) *Any set with more than $\dim_{\mathbb{F}} V$ vectors of V is linearly dependent.*
(5) *Any set with less than $\dim_{\mathbb{F}} V$ vectors of V does not span V.*
(6) *Any LI set with exactly $\dim_{\mathbb{F}} V$ vectors is a basis.*
(7) *Any set of generators of V with exactly $\dim_{\mathbb{F}} V$ vectors is a basis.*

Proof. Let k be the size of a finite set of generators for V.
(1) Let X be a finite set of generators for V. If X is LI, it is a basis, and we are done. Otherwise, by Lemma 135 we can find a vector x_0 in X such that

$$V = \operatorname{span} X = \operatorname{span} X_1, \text{ where } X_1 \stackrel{\text{def}}{=} X - \{x_0\}.$$

If this X_1 is LI, we stop; if not, we find an x_1 in X such that

$$V = \operatorname{span} X = \operatorname{span} X_1 = \operatorname{span} X_2, \text{ where } X_2 \stackrel{\text{def}}{=} X_1 - \{x_1\}.$$

And so on. At each iteration, we discard one element from X. Since X is finite, at some point the algorithm terminates, and we have a basis.
(2) By Steinitz' Exchange Lemma V cannot contain LI sets with more than k elements. So, let X be a (necessarily finite) LI set in V. If $\operatorname{span}(X) \subsetneq V$, then X is the desired basis. If instead $\operatorname{span}(X) \subsetneq V$, choose an element v_1 in V outside $\operatorname{span}(X)$ and set

$$X_1 \stackrel{\text{def}}{=} X \cup \{v_1\}.$$

This X_1 is still LI by Lemma 135. If X_1 is a basis, we stop; otherwise, we repeat the argument above, choose a v_2 in V outside $\operatorname{span}(X_1)$ and set

$$X_2 \stackrel{\text{def}}{=} X_1 \cup \{v_2\} = X \cup \{v_1, v_2\}.$$

And so on. The algorithm cannot go on expanding the LI set forever, because we cannot have LI sets with more than k elements. So eventually it must stop, yielding a basis.
(3) First of all, we claim that because V has a finite set of k generators, then all LI sets in V must be finite. In fact, if V contained a LI sets L with infinitely many elements, by Proposition 132 any $(k+1)$-element subset of L would be still LI, in contradiction with Steinitz' Exchange Lemma. So the claim is proven; in particular, all bases of V must be finite. Now if $\{v_1, \dots v_m\}$ and $\{w_1, \dots, w_n\}$ are both bases, then in particular $\{v_1, \dots v_m\}$ is LI and $\{w_1, \dots, w_n\}$ is a set of generators, so by Steinitz' Exchange Lemma $m \leq n$; but also, $\{w_1, \dots, w_n\}$ is LI and $\{v_1, \dots v_m\}$ is a set of generators, so again by Steinitz' Exchange Lemma $n \leq m$. So $m = n$.
(4) Were the set LI, we could expand it to a basis with more than $\dim_{\mathbb{F}} V$ elements.
(5) Were the set spanning V, we could extract from it a basis with less than $\dim_{\mathbb{F}} V$ elements.
(6) Were the set not spanning V, we could expand to a basis with more than $\dim_{\mathbb{F}} V$ elements.
(7) Were the set not LI, we could extract from it a basis with less than $\dim_{\mathbb{F}} V$ elements. □

With minimal modifications, these facts are basically also true for vector spaces that do *not* have a finite set of generators. The proof is not algorithmic and requires a Lemma from set theory, which we state without proof:

Lemma 138. *Let $(E_j)_{j\in J}$ be finite sets, not necessarily disjoint, all contained in some set I. Let U be the union of the E_j. Then:*

- *If J is finite, then so is U.*
- *If J is infinite, then so is U. Moreover, U is in bijection with J.*

Theorem 139. *Let V be an arbitrary $\mathbb{F}$-vector space.*

- *Any set of generators for V contains a basis of V.*
- *Any LI set in V is contained in a basis of V.*
- *Between any two bases of V, there is a bijection.*

Sketch of proof.

- The family of generating sets for V is a nonempty family that can be partially ordered by inclusion. So it has at least one minimal element. Let B be any minimal element. This B must be LI, for otherwise a linear dependence would tell us (via Lemma 135) that there is some v in B such that $\operatorname{span} B = \operatorname{span}(B - \{v\})$; so $B - \{v\}$ would be a set of generators for V strictly smaller than B, a contradiction with the minimality of B. So B is an LI set of generators for V, or in other words, a basis of V.
- Let X be an LI set. Let C be the family of LI sets that contain X. Since C is a nonempty family that can be partially ordered by inclusion, it has one or more maximal elements. Let B be any maximal element. The span of B must be the whole V, because otherwise we could pick an element v in V outside $\operatorname{span}(B)$ and obtain via Lemma 135 that $B \cup \{v\}$ is still LI; a contradiction with the maximality of B. So B is a basis.
- We have already proven the claim in case V has a finite generating set. It only remains to discuss the case in which V has bases $(u_i)_{i\in I}$ and $(x_j)_{j\in J}$, with I, J both infinite. By definition of " $(u_i)_{i\in I}$ is a basis", every x_j can be written as

$$x_j = \sum_{i\in E_j} a_{ij} u_i, \text{ for some finite subsets } E_j \text{ of } I. \tag{5}$$

  Let $U \subseteq I$ be the union of all E_j. We claim that $U = I$. From the claim, the conclusion follows immediately by the previous Lemma, which tells us that U is in bijection with J. So let us prove the claim by contradiction.
  Let k be an element of I that does not belong to U. So, k does not belong to any of the E_j. Consider u_k, the corresponding element of the basis $(u_i)_{i\in I}$ for V. Since u_k is in V, and $(x_j)_{j\in J}$ form a basis, we can express u_k as linear combination of the x_j; from that expression, using Equation 5, we can in turn replace each x_j as a linear combination of only some of the u_i's, with u_k missing out, because it does not belong to any of the E_j. So in conclusion, we can write u_k as a linear combination of the other u_i's. This contradics the fact that the u_i's are a basis. □

We conclude with a very useful "basis criterion for $\mathbb{R}^n$", valid for any positive integer n.

Theorem 140. *(implied by Anton's Theorem 4.8.8) Let $\mathbb{F}$ be a field. Let n be any positive integer. Let $\mathbf{v_1}, \ldots, \mathbf{v_n}$ be vectors in $\mathbb{F}^n$, which means that each $\mathbf{v_i}$ can be viewed as a column of n scalars. Let A be the square matrix obtained by juxtaposing the columns $\mathbf{v_i}$, i.e.*

$$A \stackrel{\text{def}}{=} (\mathbf{v_1}|\mathbf{v_2}|\ldots|\mathbf{v_n}).$$

Then $\det A \neq 0$ if and only if $\mathbf{v_1}, \ldots, \mathbf{v_n}$ are a basis of $\mathbb{F}^n$.

Proof. First of all, notice that $\mathbb{F}^n$ has dimension n. In fact, if $\mathbf{e_i}$ is the vector with a 1 in position i and 0 in all other positions, then it is easy to see that

$$\{\mathbf{e_1}, \ldots, \mathbf{e_n}\}$$

form a basis of $\mathbb{F}^n$. So by Theorem 137, any LI set of n vectors in $\mathbb{R}^n$ is a basis.

Now let $\mathbf{x} = (x_1, \dots x_n)^\top$ be a solution of the system $A\mathbf{x} = \mathbf{0}$. Since the columns of A are $\mathbf{v_1}, \dots, \mathbf{v_n}$, the equation $A\mathbf{x} = \mathbf{0}$ can also be written as

$$x_1\mathbf{v_1} + x_2\mathbf{v_2} + \dots + x_n\mathbf{v_n} = 0.$$

So $\det A \neq 0$ if and only if the system $A\mathbf{x} = \mathbf{0}$ admits only the $\mathbf{x} = \mathbf{0}$ solution; if and only if the only linear combination of the columns of A yielding zero, is the trivial one; if and only if the set of the n vectors $\mathbf{v_1}, \dots, \mathbf{v_n}$ in the n-dimensional space $\mathbb{F}^n$ is LI; if and only if $\{\mathbf{v_1}, \dots, \mathbf{v_n}\}$ is a basis. □

Example 141. Do the vectors $(1, 2, 3)$, $(4, 5, 6)$ and $(7, 8, 9)$ span $\mathbb{R}^3$? Well,

$$\det \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix} = 1 \det \begin{pmatrix} 5 & 8 \\ 6 & 9 \end{pmatrix} - 2 \det \begin{pmatrix} 4 & 7 \\ 6 & 9 \end{pmatrix} + 3 \det \begin{pmatrix} 4 & 7 \\ 5 & 8 \end{pmatrix} = -3 + 12 - 9 = 0,$$

so they don't.

Remark 142. We chose to position the 3 vectors in $\mathbb{R}^3$ above as *columns*. Some other authors prefer to position them as *rows*. Nothing changes, of course, since $\det A = \det A^\top$ for all square matrices A.

This is a nice trick. Unfortunately, it only works for exactly n vectors in $\mathbb{R}^n$. But what if we have a set X of 4 vectors in $\mathbb{R}^3$, say? How do we know if X spans $\mathbb{R}^3$? Well, we could argue that they do span $\mathbb{R}^3$ if and only if we can extract a basis from X. In other words, if we form a 3×4 matrix whose columns are the vectors in X, the question boils down to: Does this matrix have a 3×3 submatrix with nonzero determinant?

Even more generally, suppose we have k vectors in $\mathbb{R}^n$, and suppose that they do not span $\mathbb{R}^n$. We may wonder: What is the dimension of the space they span, then? It could be any number between 0 and $n-1$. Is there a way to read off this number from the $k \times n$ matrix whose columns are the given k vectors? The answer is in the next section, where we will generalize and understand Theorem 140 deeper.

3.3 Grassmann's formula, linear maps, and the rank

It is easy to see that the intersection of two or more subspaces is a subspace:

Proposition 143. *(cf. Anton, Theorem 4.2.2) Let $(M_i)_{i \in I}$ be a family of subspaces of some $\mathbb{F}$-vector space V. Their intersection is also an $\mathbb{F}$-vector space.*

Proof. If a, b belong to the intersection, they belong to each M_i. Since each M_i is a subspace, for all $\lambda \in \mathbb{F}$, $a + \lambda b$ belongs to M_i. So $a + \lambda b$ belongs to the intersection of all M_i. By Proposition 125, we conclude. □

Obviously $M \cap N$ is the largest subspace contained in both M and N (the smallest being $\{0\}$). But what is the smallest subspace containing both M and N (the largest being all of V)?

Remark 144. The 'easy' answer, $M \cup N$, is wrong. In fact, the union is usually *not* a subspace. For example think of $V = \mathbb{R}^2$, $M = \text{span}(\mathbf{e_1})$, $M = \text{span}(\mathbf{e_2})$. Then $M \cup N$ is the two Cartesian axes together. However, this is not a vector space, because $\mathbf{e_1} + \mathbf{e_2}$ does not belong to it. So it's not closed with respect to addition.

Definition 145. Let M, N be subspaces of some $\mathbb{F}$-vector space V. Set

$$M + N \stackrel{\text{def}}{=} \{x + y \quad : \quad x \in M, y \in N\}.$$

If $M \cap N = \{0\}$, we sometimes write $M \oplus N$ instead of $M + N$.

Lemma 146. *$M + N$ is the smallest subspace of V that contains both M and N.*

Proof. Exercise. Hint: First show via Proposition 125 that $M + N$ is a subspace. Then writing $x = x + 0$ show that $M \subseteq M + N$, and similarly writing $y = 0 + y$ show that $N \subseteq M + N$. Finally, consider any subspace W that contains both M and N, and show that W must contain all sums of the form $x + y$, if x is in $M \subseteq W$ and y is in $N \subseteq W$. □

Lemma 147. *If X is a set of generators for M and N is a set of generators for N, then $X \cup Y$ is a set of generator for $M + N$.*

Proof. It is clear that $\text{span}(X \cup Y)$ contains both $\text{span}(X) = M$ and $\text{span}(Y) = N$. Now let W be a subspace of V that contains both M and N. Consider any expression of the form

$$a_1x_1 + \ldots + a_mx_m + b_1y_1 + \ldots + b_ny_n,$$

with $m, n \in \mathbb{N}$, $a_1, \ldots, a_m, b_1, \ldots, b_n$ in $\mathbb{F}$, $x_1, \ldots, x_m$ in X, and $y_1, \ldots, y_n$ in Y. Since $a_1x_1 + \ldots + a_mx_m$ is in $M \subseteq W$ and $b_1y_1 + \ldots + b_ny_n$ is in $N \subseteq W$, the expression above is in W. But then W contains $\text{span}(X \cup Y)$. □

Corollary 148 (Grassmann's formula)**.** *Let M, N be subspaces of some $\mathbb{F}$-vector space. If $\dim M$ and $\dim N$ are finite, one has*

$$\dim(M + N) = \dim M + \dim N - \dim(M \cap N).$$

In particular, when $M \cap N = \{0\}$, one has $\dim(M \oplus N) = \dim M + \dim N$.

Proof. Start with a basis $z_1, \ldots, z_n$ of $M \cap N$. By Theorem 139, this can be expanded to a basis $x_1, \ldots, x_\ell, z_1, \ldots, z_n$ of M. Or, also, it can be expanded to a basis $y_1, \ldots, y_m, z_1, \ldots, z_n$ of N. We claim that $B \stackrel{\text{def}}{=} \{x_1, \ldots, x_\ell, y_1, \ldots, y_m, z_1, \ldots, z_n\}$ is a basis for $M + N$. That B is a generating set is easy to see; the difficult part is to show that B is linearly independent. Suppose

$$a_1x_1 + \ldots + a_\ell x_\ell + b_1y_1 + \ldots + b_my_m + c_1z_1 + \ldots + c_nz_n = 0. \tag{6}$$

For convenience of notation, set

$$x \stackrel{\text{def}}{=} a_1x_1 + \ldots + a_\ell x_\ell, \quad y \stackrel{\text{def}}{=} b_1y_1 + \ldots + b_my_m, \quad z \stackrel{\text{def}}{=} c_1z_1 + \ldots + c_nz_n.$$

Equation 6 tells us that $x + y + z = 0$. But since x is in M and z is in M, so is $y = -(x + z)$. At the same time y is N by definition, so $y \in M \cap N$. Hence, we can write

$$y = d_1z_1 + \ldots + d_nz_n$$

for some scalars $d_1, \ldots, d_n$ in $\mathbb{F}$. Since $y = -(x + z)$, we obtain

$$x + z + d_1z_1 + \ldots + d_nz_n = 0.$$

Remembering what x and z stood for, this yields the combination

$$(c_1 + d_1)z_1 + \ldots + (c_n + d_n)z_n + a_1x_1 + \ldots + a_\ell x_\ell = 0.$$

But since $x_1, \ldots, x_\ell, z_1, \ldots, z_n$ are linearly independent, all coefficients above are zero. In particular, $a_1 = \ldots = a_\ell = 0$. Equation 6 becomes then a linear combination of $y_1, \ldots, y_m, z_1, \ldots, z_n$, which are also linearly independent. So $b_1, \ldots, b_m$ and $c_1, \ldots, c_n$ must be zero as well. □

Definition 149. (cf. Anton's Chapter 8.1) A *linear map* is a function $T : V \longrightarrow W$ between $\mathbb{F}$-vector spaces that satisfies

$$T(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 T(v_1) + \lambda_2 T(v_2).$$

Its *kernel* is $\ker T \stackrel{\text{def}}{=} \{v \text{ such that } T(v) = 0\}$. Its *image* is $\operatorname{im} T \stackrel{\text{def}}{=} \{T(v) \text{ such that } v \in V\}$. It is easy to see that $\ker T$ is a subspace of V, whereas $\operatorname{im} T$ is a subspace of W.

Example 150. Let P_n be the $\mathbb{R}$-vector space of polynomial functions

$$f(x) = a_0 + a_1 x + \ldots + a_n x^n$$

(of degree at most n). Then the derivative $\frac{d}{dx}$ is a linear map from P_n to P_{n-1}. It is a surjective map, so the image is the whole P_{n-1}; but it is not injective. The kernel of the derivative map consists of the whole P_0 (i.e. the constant polynomials).

Example 151. Let A be an $m \times n$ matrix with entries in $\mathbb{F}$. Then 'left-multiplication by A'

$$\begin{array}{rl} T: & \mathbb{F}^n \longrightarrow \mathbb{F}^m \\ & \mathbf{v} \longmapsto A\mathbf{v} \end{array}$$

is a linear map. Its kernel is

$$\ker T = \{\mathbf{x} \in \mathbb{F}^n \text{ such that } A\mathbf{x} = \mathbf{0}\}.$$

Its image is $\operatorname{im} T = \{A\mathbf{x} \ : \ \mathbf{x} \in \mathbb{F}^n\}$, but it is convenient to rewrite it, with a trick already used in Theorem 140, as follows: if $\mathbf{x} = (x_1, \ldots x_n)^\top$, and if $\mathbf{v_1}, \ldots, \mathbf{v_n}$ are the columns of A, then

$$\operatorname{im} T = \{x_1\mathbf{v_1} + \ldots + x_n\mathbf{v_n} \ : x_i \in \mathbb{F}\} = \operatorname{span}(\mathbf{v_1}, \ldots, \mathbf{v_n}).$$

The next result is sometimes called "rank–nullity theorem", because some textbooks like to refer to $\dim \ker T$ as the "nullity" of T and to $\dim \operatorname{im} T$ as the "rank" of T.

Theorem 152 (Rank-nullity theorem)**.** *(cf. Anton's Theorems 4.8.2 and 8.1.4) Let V be a finite-dimensional vector space. For any linear map $T : V \longrightarrow W$,*

$$\dim \ker T + \dim \operatorname{im} T = \dim V.$$

Proof. Set $n \stackrel{\text{def}}{=} \dim V$. Since $\ker T$ is a subspace of a finite-dimensional vector space, it has a finite basis $v_1, \ldots, v_k$. By Theorem 139, this basis can be expanded to a basis for V; since all bases of V must have the same number n of elements, this means that we can find a basis for V of the form $v_1, \ldots, v_k, x_1, \ldots x_{n-k}$. But then in W we have

$$\operatorname{im} T = \operatorname{span}(T(v_1), \ldots, T(v_k), T(x_1), \ldots T(x_{n-k})) = \operatorname{span}\left(T(x_1), \ldots T(x_{n-k})\right).$$

It remains to see that $T(x_1), \ldots T(x_{n-k})$ is a linearly independent set. Suppose there are scalars $\lambda_1, \ldots, \lambda_{n-k}$ in $\mathbb{F}$ such that

$$\lambda_1 T(x_1) + \ldots + \lambda_{n-k} T(x_{n-k}) \ = \ 0.$$

The left hand side, by definition of linear map, equals $T(\lambda_1 x_1 + \ldots + \lambda_{n-k} x_{n-k})$; so the formula above tells us that $\lambda_1 x_1 + \ldots + \lambda_{n-k} x_{n-k}$ is in $\ker T$, which is spanned by $v_1, \ldots, v_k$. Unless the λ_i's are all zero, this contradicts the fact that $v_1, \ldots, v_k, x_1, \ldots x_{n-k}$ are linearly independent. □

In case T is the left-multiplication by A discussed above, we obtain an important result.

Definition 153. Let A be an $m \times n$ matrix with entries in a field $\mathbb{F}$.
- The *rank* of A is the dimension of the subspace of $\mathbb{F}^m$ spanned by its n columns.
- The *nullity* of A is $\dim\{\mathbf{x} \in \mathbb{F}^n \text{ such that } A\mathbf{x} = \mathbf{0}\}$, a subspace of $\mathbb{F}^n$.

Corollary 154. *(cf. Anton's theorem 4.8.2) Given any $m \times n$ matrix A, the rank of A plus the nullity of A equals n.*

Proof. Apply the rank-nullity theorem to the map

$$\begin{aligned} T: \quad & \mathbb{F}^n \longrightarrow \mathbb{F}^m \\ & \mathbf{v} \longmapsto A\mathbf{v} \end{aligned}$$

and see Example 151 for how to interpret $\dim \operatorname{im} T$ as the dimension of the space spanned by the columns of A. □

Corollary 155. *The rank of any $m \times n$ matrix is at most $\min(m, n)$.*

Proof. By definition it is the dimension of a subspace of $\mathbb{R}^m$, so it is at most m. By the rank-nullity theorem, it is also at most n. □

Corollary 156. *(implied by Anton's Theorem 4.8.8) Let A be a square $m \times m$ matrix with entries in a field $\mathbb{F}$. A has rank m if and only if it is invertible.*

Proof. By the rank-nullity theorem, the rank of A is exactly m if and only if the nullity of A is 0; if and only if the subspace $\{\mathbf{x} : A\mathbf{x} = \mathbf{0}\}$ is equal to $\{0\}$; if and only if $A\mathbf{x} = \mathbf{0}$ has only $\mathbf{x} = \mathbf{0}$ as solution; if and only if A is invertible. □

It is possible to say something more about the rank, because of the following Lemmas:

Lemma 157. *If M is an RRE matrix with entries in a field $\mathbb{F}$, the following three quantities are equal:*
- *the rank of M, i.e. the dimension of the span of the columns of M;*
- *the rank of $M^\top$, i.e. the dimension of the span of the rows of M;*
- *the number of leading ones in M;*
- *the number of nonzero rows of M.*

Proof. Let M be an $m \times n$ matrix. Suppose M has r leading ones, which by definition of RRE matrix, are in the first r rows; so r is also the number of nonzero rows. Clearly $r \leq \min(m, n)$. The r columns of the leading ones have ones in different positions, so they form an LI set. This shows that the rank of the space spanned by the columns of M is at least r. Moreover, the bottom $m - r$ rows are made of zeroes, so any column vector in M ends with $m - r$ zeroes. So the span of columns of M is a subset (in fact, a subspace) of

$$\{\mathbf{x} \in \mathbb{F}^n : x_{m-r+1} = x_{m-r+2} = \ldots = x_m = 0\},$$

which has dimension r. So the rank of M is at most r. Hence, the rank of M is exactly r. Now let us look at the space generated by the rows of M. Since the last $m - r$ rows are zero, such space is generated already by the first r rows. It remains to see that the first r rows are linearly independent. This is clear because the leading ones are the only nonzero entry in each column. So the rows of M also span a space of dimension r. □

Lemma 158. *Let A be an $m \times n$ matrix with entries in a field $\mathbb{F}$. Let B be an $m \times n$ matrix obtained from A via an elementary row operation. Then:*

- *The span of the rows of A is the same as the span of the rows of B.*
- *The span of the columns of A may be different than the span of the columns of B, but the two subspaces have the same dimension.*

Proof. The first item is easy and left to you. For the second item: Any linear dependence between some columns of the matrix A can be expressed as we did above as a nonzero vector $\mathbf{x}$ with $A\mathbf{x} = 0$. Now let E be an elementary matrix. Let $B = EA$. Then $B\mathbf{x} = EA\mathbf{x} = \mathbf{0}$, which expresses a linear dependence among the same columns of B. So if certain columns of A are linearly dependent, the same columns in B are also linearly dependent. Conversely, if some columns of B are linearly dependent, this means that there exists a nonzero vector $\mathbf{y}$ such that $B\mathbf{y} = \mathbf{0}$. But then $A\mathbf{y} = E^{-1}B\mathbf{y} = E^{-1}\mathbf{0} = \mathbf{0}$. So the same columns in A are linearly dependent. In conclusion, if a set of columns is linearly independent (respectively linearly dependent) then it remains so under an elementary row operation. Hence, one elementary row operation does not change the size of the largest set of linearly independent columns. Finally, consider

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix} \text{ and } B = \begin{pmatrix} 1 & 3 \\ 0 & 0 \end{pmatrix}.$$

Then B is obtained from A by adding -2 times the first row to the second one. The span of the columns of A consists of all multiples of the vector $\binom{1}{2}$, whereas the span of the columns of B consists of all multiples of the vector $\binom{1}{0}$; so even if they both have dimension 1, they are different subspaces. □

Theorem 159. *Let A be an $m \times n$ matrix with entries in a field $\mathbb{F}$. The following quantities are equal:*

- *the rank of A, i.e. the dimension of the span of the columns of A;*
- *the rank of $A^\top$, i.e. the dimension of the span of the rows of A;*
- *the size of the largest square submatrix of A with nonzero determinant.*

Proof. Via elementary operations, let us reduce A to its RRE form M. By Lemma 158, this does not change the rank. By Lemma 157, the rank of M is the number r of leading ones. This shows that the rank of A is the number of leading ones in M. Moreover, by Lemma 158 the space spanned by all the rows does not change at all when passing from A to M. But the space spanned by the rows of M has also dimension equal to r.

So we proved that the first two quantities are equal. It remains to show that the number r of leading ones in the RRE form of A is also the size of the largest square submatrix of A with invertible determinant. From the space spanned by all n columns of A, extract a basis of r elements any way you want. Let B be the $m \times r$ submatrix of A formed by those r LI columns that span the same space spanned by all columns. Clearly B has also rank r. If $r = m$, we stop: B is square and has maximum rank, so by Corollary 156 it has nonzero determinant. If $r < m$, we apply to B the already proved equivalence of the first two quantities: The dimension of the span of the *rows* of B is equal to the dimension of the span of its columns, namely, r. So pick r rows of B that span such space and restrict to the submatrix C of B formed by them. This C is now an $r \times r$ matrix with rank r, so by Corollary 156 it has nonzero determinant. □

An important application of this is to linear systems of the form $A\mathbf{x} = \mathbf{0}$. Notice that if A' is obtained from A by appending to it a zero column on the right, then the reduction into RRE form of A and of A' involve multiplications with the same matrices. Now if A is an $m \times n$ matrix, then:

- the rank of A is the number of leading ones in the RRE form of A, and also in the RRE form of A'; they will correspond to the variables that are "determined" by the system, and must be equal to zero in a solution.

- the nullity of A is the dimension of $\{\mathbf{x} \in \mathbb{F}^n \text{ such that } A\mathbf{x} = \mathbf{0}\}$, i.e. the number of free parameters in a solution. These are the variables that can take an arbitrary value in a solution.

So a system $A\mathbf{x} = \mathbf{0}$ of m equations in n unknowns, if $r = rank(A)$, has a solution that contains exactly $n - r$ independent parameters ("degrees of freedom").

The same analysis is true for a system $A\mathbf{x} = \mathbf{b}$, except that this system might be impossible. So the correct statement is: a system $A\mathbf{x} = \mathbf{b}$ of m equations in n unknowns, if $r = rank(A)$, is either impossible, or it has a solution that contains exactly $n - r$ independent parameters ("degrees of freedom").

We conclude the section with a characterization of ranks of matrices.

Remark 160. Unlike the notion of "determinant", we defined the notion of "rank" only for matrices with entries in a field. Several inequivalent generalizations to commutative rings are possible: See e.g. https://mattbaker.blog/2022/12/24/linear-algebra-over-rings/

4 Orthogonality in $\mathbb{R}^n$ and in R^n, with R any ring

4.1 Orthogonal matrices and orthogonal vectors

Definition 161. Let A be a square, $n \times n$ matrix with entries in a commutative ring R. We say that A is *orthogonal* if

$$A^\top A = I_n,$$

or equivalently, if the inverse of A is the transpose of A. We say that A is a *rotation* if it is orthogonal, and $\det A = 1$.

Note that if A is orthogonal, so is $A^\top$. Also: From Cauchy–Binet's formula, if A is orthogonal, then

$$(\det A)^2 = \det A \cdot \det A = \det A^\top \cdot \det A = \det(A^\top A) = \det I_n = 1.$$

So it could be that $\det A = 1$, but if $R = \mathbb{Z}$, for example, it could also be that $\det A = -1$. Note that if R is $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$, then in R the equation $x^2 = 1$ has exactly two solutions ($x = 1$ and $x = -1$); but there are rings like $\mathbb{Z}_2$ in which the equation $x^2 = 1$ has only one solution, because $-1 = 1$; and there are also rings in which the equation $x^2 = 1$ has infinitely many solutions. For example, in the ring R formed by the 2×2 diagonal matrices with entries in $\mathbb{R}$, any element $A_t \stackrel{\text{def}}{=} \begin{pmatrix} 1 & t \\ 0 & -1 \end{pmatrix}$ has the property that $A_t A_t = I_2$.

Definition 162. The *dot product* of two vectors $\mathbf{x}$, $\mathbf{y}$ in R^n is the element of R defined by

$$\mathbf{x} \bullet \mathbf{y} \stackrel{\text{def}}{=} \sum_{i=1}^{n} x_i y_i.$$

Of course, the four quantities $\mathbf{x}\mathbf{y}^\top$, $\mathbf{x}^\top\mathbf{y}$, $\mathbf{y}^\top\mathbf{x}$ and $\mathbf{y}\mathbf{x}^\top$ are all equal to $\sum_{i=1}^n x_i y_i$, so any of these quantities can also be taken as definition of $\mathbf{x} \bullet \mathbf{y}$.

Lemma 163. *(cf. Anton's Theorems 3.2.2 and 3.2.3) Let R be a commutative ring. Let $\mathbf{u}, \mathbf{v}, \mathbf{w}$ be three vectors in R^n. Let $r \in R$. Then:*

(a) $\mathbf{u} \bullet \mathbf{v} = \mathbf{v} \bullet \mathbf{u}$;
(b) $\mathbf{u} \bullet (\mathbf{v} + \mathbf{w}) = \mathbf{u} \bullet \mathbf{v} + \mathbf{u} \bullet \mathbf{w}$;
(c) $r(\mathbf{u} \bullet \mathbf{v}) = (r\mathbf{u}) \bullet \mathbf{v}$;
(d) $\mathbf{0} \bullet \mathbf{v} = 0$.

Proof. Exercise. □

Lemma 164. *(cf. Anton's formula 26 on Ch.3) Let R be a commutative ring. Let $\mathbf{u}, \mathbf{v}$ be two vectors in R^n. For any $n \times n$ matrix A with entries in R,*

$$(A\mathbf{u}) \bullet \mathbf{v} = \mathbf{u} \bullet (A^\top \mathbf{v}).$$

Proof. $(A\mathbf{u}) \bullet \mathbf{v} = (A\mathbf{u})^\top \mathbf{v} = \mathbf{u}^\top A^\top \mathbf{v} = \mathbf{u} \bullet (A^\top \mathbf{v})$. □

Proposition 165. *Let R be a commutative ring. Let A be an $n \times n$ matrix with entries in R. A is symmetric if and only if for all $\mathbf{u}, \mathbf{v}$ in R^n one has*

$$(A\mathbf{u}) \bullet \mathbf{v} = \mathbf{u} \bullet (A\mathbf{v}).$$

Proof. Exercise. □

Proposition 166. *(similar to Anton's Theorem 7.1.3) Let R be any commutative ring. For an $n \times n$ matrix A with entries in R, the following are equivalent:*
(a) *A is orthogonal;*
(b) *$(A\mathbf{x}) \bullet (A\mathbf{y}) = \mathbf{x} \bullet \mathbf{y}$, for all $\mathbf{x}, \mathbf{y}$ in R^n.*
Furthermore, if R is either $\mathbb{Z}$ or a field, this third statement is equivalent to the other two:
(c) *$(A\mathbf{x}) \bullet (A\mathbf{x}) = \mathbf{x} \bullet \mathbf{x}$, for all $\mathbf{x}$ in R^n.*

Proof. '(a) $\Rightarrow$ (b)': By Lemma 164, $(A\mathbf{x}) \bullet (A\mathbf{y}) = \mathbf{x} \bullet (A^\top A\mathbf{y}) = \mathbf{x} \bullet \mathbf{y}$.
'(b) $\Rightarrow$ (a)': By Lemma 164, $\mathbf{x} \bullet A^\top A\mathbf{y} = A\mathbf{x} \bullet A\mathbf{y} = \mathbf{x} \bullet \mathbf{y}$. In other words, if we call B the matrix $A^\top A - I_n$, for all $\mathbf{x}, \mathbf{y}$ in R^n we have that

$$\mathbf{x}^\top B\mathbf{y} = \mathbf{x}^\top (A^\top A - I_n)\mathbf{y} = \mathbf{x}^\top A^\top A\mathbf{y} - \mathbf{x}^\top \mathbf{y} = \mathbf{x} \bullet A^\top A\mathbf{y} - \mathbf{x} \bullet \mathbf{y} = 0.$$

In particular, this is true for $\mathbf{x} = \mathbf{e_i}$ and $\mathbf{y} = \mathbf{e_j}$: but since

$$\mathbf{e_i}^\top B\mathbf{e_j} = b_{i,j},$$

we obtain that $b_{i,j} = 0$ for all i, j. So B is the zero matrix and $A^\top A = I_n$.
This shows that (a) and (b) are equivalent; clearly, (b) implies (c) by taking $\mathbf{x} = \mathbf{y}$. Now suppose that R is either $\mathbb{Z}$ or a field, and assume (c) holds. In particular, if we call B the matrix $A^\top A - I_n$, for all $\mathbf{x}$ in R^n we have that

$$\mathbf{x}^\top B\mathbf{x} = 0.$$

Apply this to $\mathbf{x} = \mathbf{e_i} + \mathbf{e_j}$. This tells us that

$$\begin{aligned} 0 &= (\mathbf{e_i} + \mathbf{e_j})^\top B(\mathbf{e_i} + \mathbf{e_j}) = \\ &= \mathbf{e_i}^\top B\mathbf{e_i} + \mathbf{e_i}^\top B\mathbf{e_j} + \mathbf{e_j}^\top B\mathbf{e_i} + \mathbf{e_i}^\top B\mathbf{e_i} \\ &= 0 + b_{i,j} + b_{j,i} + 0, \end{aligned}$$

which tells us that $b_{i,j} = -b_{j,i}$. But the matrix B is symmetric, because it is the sum of the symmetric matrix $A^\top A$ and of the diagonal matrix $-I$. So $b_{i,j} = b_{j,i}$. This tells us that $2b_{i,j} = 0$ for all i, j. By the assumptions on R, this implies that $b_{i,j} = 0$ for all i, j. □

Remark 167. If $R = \mathbb{Z}_2$, consider the matrix

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

Then A is not orthogonal. However, for all $\mathbf{x} = (x, y)^\top$ in $(\mathbb{Z}_2)^2$, one has

$$(A\mathbf{x}) \bullet (A\mathbf{x}) = \begin{pmatrix} x+y \\ 0 \end{pmatrix} \bullet \begin{pmatrix} x+y \\ 0 \end{pmatrix} = x^2 + 2xy + y^2 = x^2 + y^2 = \mathbf{x} \bullet \mathbf{x}.$$

So in the previous Proposition, when R is an arbitrary commutative ring, (c) might not imply (b) and (a).

Definition 168. Two vectors $\mathbf{x}$, $\mathbf{y}$ in R^n are *orthogonal* if $\mathbf{x} \bullet \mathbf{y} = 0$. In particular, the zero vector is orthogonal to any vector.

Theorem 169 (Pythagoras's theorem)**.** *If $\mathbf{x}$, $\mathbf{y}$ in R^n are* orthogonal*, then*

$$(\mathbf{x} + \mathbf{y}) \bullet (\mathbf{x} + \mathbf{y}) = \mathbf{x} \bullet \mathbf{x} + \mathbf{y} \bullet \mathbf{y}.$$

Proof. We have

$$(\mathbf{x}+\mathbf{y})\bullet(\mathbf{x}+\mathbf{y}) = \mathbf{x}\bullet\mathbf{x}+\mathbf{x}\bullet\mathbf{y}+\mathbf{y}\bullet\mathbf{x}+\mathbf{y}\bullet\mathbf{y},$$

but some of the summands are zero by assumption. □

Definition 170 (Orthogonal and orthonormal sets)**.** A set of vectors in R^n is called *orthogonal* if the dot product of any two distinct elements in it yields 0. A set of vectors in R^n is called *orthonormal* if it is orthogonal, and the dot product of any element with itself yields 1.

Theorem 171. *(cf. Anton's Theorem 7.1.1) Let A be an $n\times n$ matrix with entries in a commutative ring R.*
A is orthogonal $\iff$ its columns form an orthonormal set in R^n.

Proof. The (i,j)-element of $A^\top A$ is exactly obtained as the dot product of the i-th row of $A^\top$ (which is the same as the i-th column of A) with the j-th column of A. So, saying that the diagonal element (i,i) of $A^\top A$ is 1, means that the dot product of the i-th column of A with itself yields 1. Instead, saying that $A^\top A$ has zeroes off the diagonal, is the same as saying that for $i\neq j$, the dot product of the i-th column of A with the j-th yields 0. □

Theorem 172. *Let $\mathbb{F}$ be a field. Any orthonormal set in $\mathbb{F}^n$ is LI.*

Proof. Suppose

$$a_1\mathbf{x_1}+a_2\mathbf{x_2}+\ldots+a_m\mathbf{x_m}=\mathbf{0}$$

for some a_i's in R and some $\mathbf{x_1},\ldots,\mathbf{x_m}$ in our orthonormal set. If we take the dot product of the expression above with a single $\mathbf{x_i}$, we obtain

$$a_1(\mathbf{x_1}\bullet\mathbf{x_i})+a_2(\mathbf{x_2}\bullet\mathbf{x_i})+\ldots+a_i(\mathbf{x_i}\bullet\mathbf{x_i})+\ldots+a_m(\mathbf{x_m}\bullet\mathbf{x_i})=0,$$

which by the orthonormality assumption simplifies to $a_i(1)=0$. □

Remark 173. An orthogonal set need not be LI, because it may contain the zero vector. Or more generally, it may contain vectors whose scalar product with themselves is zero.

Is it possible to transform an orthogonal set "without the zero vector in it" into an orthonormal set? Well, the idea would be to "normalize" the vectors, i.e. divide each vector $\mathbf{x}$ by the quantity $\mathbf{x}\bullet\mathbf{x}$. However, can we make sure that $\mathbf{x}\bullet\mathbf{x}$ is a number different than zero? The scalar

$$\mathbf{x}\bullet\mathbf{x}=\sum_{i=1}^{n}(x_i)^2$$

is a sum of squares, but we are in an arbitrary field $\mathbb{F}$, so we cannot conclude that if a sum of squares is zero, then all numbers are zero. For example, if $\mathbb{F}=\mathbb{C}$, one has $1^2+i^2=0$. So the dot prodoct of the vector $(1,i)$ with itself, is zero.

For these reason, it is particularly interesting to focus on $\mathbb{R}^n$, as an $\mathbb{R}$-vector space.

4.2 The case of the Euclidean space $\mathbb{R}^n$

Inside $\mathbb{R}^n$ there is a dot product with the crucial property that

$$\mathbf{x}\bullet\mathbf{x}>0 \text{ if } \mathbf{x}\neq\mathbf{0} \tag{7}$$

(because it is a sum of squares). This property immediately triggers several others. For example:

Definition 174. The *norm* (or *length*) of a vector $\mathbf{x} = (x_1, \ldots, x_n)^\top$ in $\mathbb{R}^n$ is defined as

$$||\mathbf{x}|| \stackrel{\text{def}}{=} \sqrt{\mathbf{x} \bullet \mathbf{x}}.$$

Remark 175. Clearly, $||\mathbf{x}|| = 0$ if and only if $\mathbf{x} = \mathbf{0}$. Note also that $\|\lambda\mathbf{x}\| = |\lambda| \bullet \|\mathbf{x}\|$. Finally, note that for $n = 1$, $||x|| \stackrel{\text{def}}{=} \sqrt{x^2} = |x|$, so the norm coincides with the absolute value.

For $\mathbb{R}^n$ we can strengthen Theorem 172 as follows:

Theorem 176. *(cf. Anton's Thm 6.3.1) In $\mathbb{R}^n$, any orthogonal set of nonzero vectors is LI.*

Proof. Suppose

$$a_1\mathbf{x_1} + a_2\mathbf{x_2} + \ldots + a_m\mathbf{x_m} = \mathbf{0}$$

for some a_i's in $\mathbb{R}$ and some $\mathbf{x_1}, \ldots, \mathbf{x_m} \neq \mathbf{0}$ in our orthogonal set. If we take the dot product of the expression above with a single $\mathbf{x_i}$, we obtain

$$a_1(\mathbf{x_1} \bullet \mathbf{x_i}) + a_2(\mathbf{x_2} \bullet \mathbf{x_i}) + \ldots + a_i(\mathbf{x_i} \bullet \mathbf{x_i}) + \ldots + a_m(\mathbf{x_m} \bullet \mathbf{x_i}) = 0,$$

which by the orthogonality assumption simplifies to $a_i||\mathbf{x_i}||^2 = 0$. Since $\mathbf{x_i}$ is not zero, this implies $a_i = 0$. □

Here are two famous inequalities:

Theorem 177 (Cauchy–Schwarz inequality)**.** *For all vectors $\mathbf{x}, \mathbf{y}$ in $\mathbb{R}^n$ one has*

$$|\mathbf{x} \bullet \mathbf{y}| \;\leq\; ||\mathbf{x}||\,||\mathbf{y}||,$$

with equality if and only if $\{\mathbf{x}, \mathbf{y}\}$ is linearly dependent.

Proof. The inequality above involves nonnegative numbers. Now, to show that two real numbers a, b satisfy $0 \leq a \leq b$, it suffices to show $0 \leq a^2 \leq b^2$. So we are going to show that

$$(\mathbf{x} \bullet \mathbf{y})^2 \;\leq\; (\mathbf{x} \bullet \mathbf{x})(\mathbf{y} \bullet \mathbf{y}),$$

where the inequality is strict if and only if $\{\mathbf{x}, \mathbf{y}\}$ is LI. But a two-element set $\{\mathbf{x}, \mathbf{y}\}$ is LI if and only if $\mathbf{y} \neq \mathbf{0}$ and $\mathbf{x} \neq \lambda\mathbf{y}$ for any λ in $\mathbb{R}$. So we distinguish three cases:

- if $\mathbf{y} = \mathbf{0}$, then $(\mathbf{x} \bullet \mathbf{y})^2 = 0^2 = 0 = (\mathbf{x} \bullet \mathbf{x})(\mathbf{y} \bullet \mathbf{y})$;
- if $\mathbf{x} = \lambda\mathbf{y}$ for some λ, then $(\mathbf{x} \bullet \mathbf{y})^2 = (\lambda\mathbf{y} \bullet \mathbf{y})^2 = \lambda^2(\mathbf{y} \bullet \mathbf{y})^2 = (\lambda\mathbf{y}) \bullet (\lambda\mathbf{y})(\mathbf{y} \bullet \mathbf{y}) = (\mathbf{x} \bullet \mathbf{x})\,(\mathbf{y} \bullet \mathbf{y})$;
- if instead $\mathbf{y} \neq \mathbf{0}$ and $\mathbf{x} \neq \lambda\mathbf{y}$ for all λ, then by Equation 7 we have

$$0 < (\mathbf{x} - \lambda\mathbf{y}) \bullet (\mathbf{x} - \lambda\mathbf{y}) = (\mathbf{x} \bullet \mathbf{x}) - 2\lambda(\mathbf{x} \bullet \mathbf{y}) + \lambda^2(\mathbf{y} \bullet \mathbf{y}).$$

Since the inequality above holds true for all λ, in particular it holds for $\lambda = \frac{\mathbf{x} \bullet \mathbf{y}}{\mathbf{y} \bullet \mathbf{y}}$, which means

$$0 < (\mathbf{x} \bullet \mathbf{x}) - 2\frac{(\mathbf{x} \bullet \mathbf{y})^2}{\mathbf{y} \bullet \mathbf{y}} + \frac{(\mathbf{x} \bullet \mathbf{y})^2}{(\mathbf{y} \bullet \mathbf{y})^2}\mathbf{y} \bullet \mathbf{y} = \mathbf{x} \bullet \mathbf{x} - \frac{(\mathbf{x} \bullet \mathbf{y})^2}{(\mathbf{y} \bullet \mathbf{y})}.$$

Multiplying by $(\mathbf{y} \bullet \mathbf{y})$, which is positive by Equation 7, we get $(\mathbf{x} \bullet \mathbf{y})^2 < (\mathbf{x} \bullet \mathbf{x})\,(\mathbf{y} \bullet \mathbf{y})$. This proves the inequality. Note that the inequality holds with equality in the first two cases, but is strict in the third case. □

Theorem 178 (Triangle inequality)**.** *For any* $\mathbf{x}, \mathbf{y}$ *vectors of* $\mathbb{R}^n$*, one has*

$$||\mathbf{x}+\mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}||.$$

with equality if and only if one of the two vectors is a positive multiple of the other, or one of the two vectors is zero.

Proof. Since $\lambda \leq |\lambda|$ for all $\lambda \in \mathbb{R}$, in particular we have that for all $\mathbf{x}, \mathbf{y}$ vectors of $\mathbb{R}^n$

$$\mathbf{x} \bullet \mathbf{y} \leq |\mathbf{x} \bullet \mathbf{y}|.$$

But then applying the line above and the Cauchy–Schwarz inequality we obtain

$$\begin{aligned} ||\mathbf{x}+\mathbf{y}||^2 &= (\mathbf{x}+\mathbf{y}) \bullet (\mathbf{x}+\mathbf{y}) \\ &= (\mathbf{x} \bullet \mathbf{x}) + 2(\mathbf{x} \bullet \mathbf{y}) + (\mathbf{y} \bullet \mathbf{y}) \\ &\leq ||\mathbf{x}||^2 + 2|\mathbf{x} \bullet \mathbf{y}| + ||\mathbf{y}||^2 \\ &\leq ||\mathbf{x}||^2 + 2||\mathbf{x}||\,||\mathbf{y}|| + ||\mathbf{y}||^2 = \\ &= (||\mathbf{x}|| + ||\mathbf{y}||)^2. \end{aligned}$$

This proves the inequality. To have equality, all steps above must be equality. In particular, we also must have $\mathbf{x} \bullet \mathbf{y} = |\mathbf{x} \bullet \mathbf{y}|$, which tells us that $\mathbf{x} \bullet \mathbf{y} \geq 0$, and we must have equality in the Cauchy-Schwarz inequality, which tells us that $\{\mathbf{x}, \mathbf{y}\}$ is linearly dependent. So either $\mathbf{y} = \mathbf{0}$, or $\mathbf{x} = \mathbf{0}$, or $\mathbf{x} = \lambda \mathbf{y} \neq \mathbf{0}$ for some $\lambda \neq 0$. But $\mathbf{x} \bullet \mathbf{y} \geq 0$ can then be rewritten as $\lambda(\mathbf{y} \bullet \mathbf{y}) \geq 0$, and we know that $(\mathbf{y} \bullet \mathbf{y})$ is positive by Equation 7; so λ must be positive as well. □

Another formula of Euclidean geometry is the *law of cosines*. It says that in a triangle of edge lengths a, b, c, if α is the angle opposite to the edge of length a one has

$$a^2 = b^2 + c^2 - 2bc\cos\alpha.$$

(A particular case is Pythagora's theorem: when α is a right angle, $a^2 = b^2 + c^2$.) This formula is absorbed into our definition of "dot product", as follows:

Theorem 179. *For any* $\{\mathbf{x}, \mathbf{y}\}$ *set of LI vectors of* $\mathbb{R}^n$*, in* $\mathbb{R}$ *one has the identity*

$$\mathbf{x} \bullet \mathbf{y} = ||\mathbf{x}||\,||\mathbf{y}||\,\cos\alpha,$$

where α *is the angle formed by* $\mathbf{x}, \mathbf{y}$ *in the unique plane of* $\mathbb{R}^n$ *spanned by them.*

Proof. By the law of cosines applied to the triangle formed by the vectors $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{x}-\mathbf{y}$, which is opposite to α,

$$||\mathbf{x}-\mathbf{y}||^2 = ||\mathbf{x}||^2 + ||\mathbf{y}||^2 - 2||\mathbf{x}||||\mathbf{y}||\cos\alpha.$$

But then,

$$\begin{aligned} 2||\mathbf{x}||||\mathbf{y}||\cos\alpha &= ||\mathbf{x}||^2 + ||\mathbf{y}||^2 - ||\mathbf{x}-\mathbf{y}||^2 = \\ &= (\mathbf{x} \bullet \mathbf{x}) + (\mathbf{y} \bullet \mathbf{y}) - (\mathbf{x}-\mathbf{y}) \bullet (\mathbf{x}-\mathbf{y}) = \\ &= 2(\mathbf{x} \bullet \mathbf{y}). \qquad \square \end{aligned}$$

Corollary 180. *'Perpendicular' (= forming an angle* α *equal to* 90 *or* 270 *degrees) is synonymous with 'orthogonal'. In fact, given any non-zero vectors* $\{\mathbf{x}, \mathbf{y}\}$*, we have*

$$0 = \mathbf{x} \bullet \mathbf{y} = ||\mathbf{x}||\,||\mathbf{y}||\,\cos\alpha \quad \Longleftrightarrow \quad \cos\alpha = 0 \quad \Longleftrightarrow \quad \alpha = 90^o \text{ or } 270^o.$$

Finally, notice the following fact: Given a vector $(a, b, c)^\top$ in $\mathbb{R}^3$, what is the set of all vectors perpendicular to it? It can be written as

$$\{(x, y, z)^\top \ : \ (a, b, c)(x, y, z)^\top = 0\},$$

and so it is the plane of equation $ax + by + cz = 0$. This generalizes to all dimensions: The hyperplane orthogonal to a given nonzero vector $\mathbf{a}$ is the hyperplane of equation

$$\mathbf{a} \bullet \mathbf{x} = 0.$$

4.3 Projections and the Gram-Schmidt algorithm in $\mathbb{R}^n$

Definition 181 (Projection)**.** Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. We define the *projection in direction* $\mathbf{x}$ *of* $\mathbf{y}$ as

$$\mathrm{proj}_{\mathbf{x}}(\mathbf{y}) \stackrel{\text{def}}{=} \frac{\mathbf{x} \bullet \mathbf{y}}{\mathbf{x} \bullet \mathbf{x}} \mathbf{x} \text{ if } \mathbf{x} \neq \mathbf{0}, \qquad \text{and } \mathrm{proj}_{\mathbf{x}}(\mathbf{y}) \stackrel{\text{def}}{=} 0 \text{ if } \mathbf{x} = \mathbf{0}.$$

Remark 182. Note that $\mathrm{proj}_{\mathbf{x}}(\mathbf{y})$ is a vector proportional to $\mathbf{x}$. The wording "in direction $\mathbf{x}$" suggests that $\mathrm{proj}_{\mathbf{x}}(\mathbf{y})$ is invariant up to rescaling $\mathbf{x}$: indeed, for any $\lambda \neq 0$,

$$\mathrm{proj}_{\lambda\mathbf{x}}(\mathbf{y}) = \frac{(\lambda\mathbf{x}) \bullet \mathbf{y}}{(\lambda\mathbf{x}) \bullet (\lambda\mathbf{x})} \lambda\mathbf{x} = \frac{\lambda^2}{\lambda^2} \frac{\mathbf{x} \bullet \mathbf{y}}{\mathbf{x} \bullet \mathbf{x}} \mathbf{x} = \mathrm{proj}_{\mathbf{x}}(\mathbf{y}).$$

However, $\mathrm{proj}_{\mathbf{x}}(\mathbf{y})$ does change if we rescale $\mathbf{y}$: In fact, it is easy to see that

$$\mathrm{proj}_{\mathbf{x}}(\lambda\mathbf{y}) = \lambda \, \mathrm{proj}_{\mathbf{x}}(\mathbf{y}).$$

Lemma 183. *For all* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$*, the vector*

$$\mathbf{z} \stackrel{\text{def}}{=} \mathbf{y} - \mathrm{proj}_{\mathbf{x}}(\mathbf{y}) \;\; \textit{is orthogonal to } \mathbf{x}.$$

More generally, for all $\mathbf{x_1}, \ldots, \mathbf{x_m}, \mathbf{y}$ *in* $\mathbb{R}^n$*, if* $\{\mathbf{x_1}, \ldots, \mathbf{x_m}\}$ *is an orthogonal set, then so is* $\{\mathbf{x_1}, \ldots, \mathbf{x_m}, \mathbf{z}\}$*, where*

$$\mathbf{z} \stackrel{\text{def}}{=} \mathbf{y} - \mathrm{proj}_{\mathbf{x_1}}(\mathbf{y}) - \ldots - \mathrm{proj}_{\mathbf{x_m}}(y).$$

Proof. The first claim, which is also the case $m = 1$ of the second claim, is obvious if $\mathbf{x} = \mathbf{0}$ (any vector is orthogonal to $\mathbf{0}$!) and an easy computation if $\mathbf{x} \neq \mathbf{0}$:

$$\begin{aligned} \mathbf{x} \bullet (\mathbf{y} - \mathrm{proj}_{\mathbf{x}}(\mathbf{y})) &= (\mathbf{x} \bullet \mathbf{y}) - (\mathbf{x} \bullet \mathrm{proj}_{\mathbf{x}}(\mathbf{y})) \\ &\stackrel{\text{def}}{=} (\mathbf{x} \bullet \mathbf{y}) - (\mathbf{x} \bullet \tfrac{\mathbf{x} \bullet \mathbf{y}}{\mathbf{x} \bullet \mathbf{x}} \mathbf{x}) \\ &= (\mathbf{x} \bullet \mathbf{y}) - \tfrac{\mathbf{x} \bullet \mathbf{y}}{\mathbf{x} \bullet \mathbf{x}} (\mathbf{x} \bullet \mathbf{x}) = 0. \end{aligned}$$

For the second claim, we proceed by induction on m. Set $\mathbf{w} \stackrel{\text{def}}{=} \mathbf{x} - \mathrm{proj}_{\mathbf{x_1}}(\mathbf{y}) - \ldots - \mathrm{proj}_{\mathbf{x_{m-1}}}(\mathbf{y})$. By inductive assumption, $\mathbf{w}$ is orthogonal to all of $\mathbf{x_1}, \ldots, \mathbf{x_{m-1}}$. Thus for all $i \in \{1, \ldots, m-1\}$,

$$\begin{aligned} \mathbf{x_i} \bullet \big(\mathbf{y} - \mathrm{proj}_{\mathbf{x_1}}(\mathbf{y}) - \ldots - \mathrm{proj}_{\mathbf{x_m}}(\mathbf{y})\big) &= \mathbf{x_i} \bullet \big(\mathbf{w} - \mathrm{proj}_{\mathbf{x_m}}(\mathbf{y})\big) \\ &= (\mathbf{x_i} \bullet \mathbf{w}) - \big(\mathbf{x_i} \bullet \mathrm{proj}_{\mathbf{x_m}}(\mathbf{y})\big) \\ &= \quad 0 \quad - \quad 0, \end{aligned}$$

because $\mathrm{proj}_{\mathbf{x_m}}(\mathbf{y})$ is "proportional to $\mathbf{x_m}$" and therefore orthogonal to all $\mathbf{x_i}$'s with $i < m$ (here we used the assumption that the $\mathbf{x_i}$'s are an orthogonal set). It remains to see whether the last vector $\mathbf{x_m}$ is also orthogonal to $\mathbf{y} - \mathrm{proj}_{\mathbf{x_1}}(\mathbf{y}) - \ldots - \mathrm{proj}_{\mathbf{x_m}}(\mathbf{y})$. Indeed,

$$\begin{aligned} \mathbf{x_m} \bullet \big(\mathbf{y} - \mathrm{proj}_{\mathbf{x_1}}(\mathbf{y}) - \ldots - \mathrm{proj}_{\mathbf{x_m}}(y)\big) &= (\mathbf{x_m} \bullet \mathbf{y}) - (\mathbf{x_m} \bullet \mathrm{proj}_{\mathbf{x_1}}(\mathbf{y})) - \ldots - (\mathbf{x_m} \bullet \mathrm{proj}_{\mathbf{x_m}}(\mathbf{y})) \\ &= (\mathbf{x_m} \bullet \mathbf{y}) - 0 - \ldots - 0 - (\mathbf{x_m} \bullet \mathrm{proj}_{\mathbf{x_m}}(\mathbf{y})) \\ &= (\mathbf{x_m} \bullet \mathbf{y}) - \mathbf{x_m} \bullet \left(\tfrac{\mathbf{x_m} \bullet \mathbf{y}}{\mathbf{x_m} \bullet \mathbf{x_m}} \mathbf{x_m}\right) \\ &= (\mathbf{x_m} \bullet \mathbf{y}) - \tfrac{\mathbf{x_m} \bullet \mathbf{y}}{\mathbf{x_m} \bullet \mathbf{x_m}} (\mathbf{x_m} \bullet \mathbf{x_m}) = 0. \quad \square \end{aligned}$$

The next theorem is what authorized me to "draw" for you $\mathrm{proj}_{\mathbf{x}}(\mathbf{y})$ the way I did in class:

Theorem 184. *For all* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$*, there is a unique vector* $\mathbf{z}$ *orthogonal to* $\mathbf{x}$ *such that*

$$\mathbf{y} = \lambda\mathbf{x} + \mathbf{z},$$

for some $\lambda \in \mathbb{R}$*. In fact, if* $\mathbf{x} \neq \mathbf{0}$*, the scalar* λ *is also uniquely determined, i.e. there is a unique way to decompose* $\mathbf{y} = \lambda\mathbf{x} + \mathbf{z}$ *with* λ *in* $\mathbb{R}$ *and* $\mathbf{z}$ *orthogonal to* $\mathbf{x}$*.*

Proof. If $\mathbf{x} = \mathbf{0}$, the (forced!) choice $\mathbf{z} = \mathbf{y}$ works, because it is orthogonal to $\mathbf{0}$. If $\mathbf{x} \neq \mathbf{0}$, the choice $\lambda = \frac{\mathbf{x} \bullet \mathbf{y}}{\mathbf{x} \bullet \mathbf{x}}$ works, because then $\lambda \mathbf{x} \stackrel{\text{def}}{=} \text{proj}_{\mathbf{x}}(\mathbf{y})$ and $\mathbf{z} = \mathbf{y} - \text{proj}_{\mathbf{x}}(\mathbf{y})$ is indeed orthogonal to $\mathbf{x}$ by Lemma 183. So the "existence" is clear. As for uniqueness, suppose

$$\mu \mathbf{x} + \mathbf{w} = \mathbf{y} = \lambda \mathbf{x} + \mathbf{z},$$

with $\lambda, \mu \in \mathbb{R}$ and $\mathbf{w}, \mathbf{z}$ orthogonal to $\mathbf{x}$. Let us rewrite the equality above as

$$(\mu - \lambda)\mathbf{x} = \mathbf{z} - \mathbf{w}.$$

If we take the dot product of both sides of this equality by $\mathbf{x}$, we obtain

$$(\mu - \lambda)||\mathbf{x}||^2 = 0 - 0,$$

which implies $\mu - \lambda = 0$. So $\mu = \lambda$. But then

$$\mathbf{w} = \mathbf{y} - \mu \mathbf{x} = \mathbf{y} - \lambda \mathbf{x} = \mathbf{z}.$$ □

Theorem 185 (Gram-Schmidt)**.** *(cf. Anton's Theorem 6.3.5) Given a finite set of linearly independent vectors $\mathbf{v_1}, \ldots, \mathbf{v_m}$ in $\mathbb{R}^n$, there exist a set of orthonormal vectors $\mathbf{o_1}, \ldots, \mathbf{o_m}$ with the same span, and in fact, with the stronger property that $\text{span}(\mathbf{v_1}, \ldots, \mathbf{v_j}) = \text{span}(\mathbf{o_1}, \ldots, \mathbf{o_j})$ for all $j \leq m$. In particular, every nonzero subspace of $\mathbb{R}^n$ has an orthonormal basis.*
Moreover, any orthonormal set of vectors in any subspace $W \subseteq \mathbb{R}^n$ can be completed to an orthonormal basis of W.

Proof. Via the following recursive procedure, called *Gram–Schmidt algorithm*. Define

$$\begin{aligned}
\mathbf{o_1} &\stackrel{\text{def}}{=} \mathbf{v_1} \\
\mathbf{o_2} &\stackrel{\text{def}}{=} \mathbf{v_2} - \text{proj}_{\mathbf{o_1}} \mathbf{v_2} \\
&\vdots \\
\mathbf{o_j} &\stackrel{\text{def}}{=} \mathbf{v_j} - \text{proj}_{\mathbf{o_1}} \mathbf{v_j} - \text{proj}_{\mathbf{o_2}} \mathbf{v_j} - \ldots - \text{proj}_{\mathbf{o_{j-1}}} \mathbf{v_j}.
\end{aligned}$$

By Lemma 183, an elementary induction shows the $\mathbf{o_i}$'s are all orthogonal to one another. So by Theorem 176 the set of the $\mathbf{o_i}$'s is linearly independent. We claim that for all j,

$$\text{span}(\mathbf{v_1}, \ldots, \mathbf{v_j}) = \text{span}(\mathbf{o_1}, \ldots, \mathbf{o_j}).$$

In fact, the very definition of $\mathbf{o}_j$ immediately implies that

$$\mathbf{v_j} = \text{proj}_{\mathbf{o_1}} \mathbf{v_j} + \text{proj}_{\mathbf{o_2}} \mathbf{v_j} + \ldots + \text{proj}_{\mathbf{o_{j-1}}} \mathbf{v_j} + \mathbf{o_j},$$

which is a way to write $\mathbf{v_j}$ as a linear combination of $\{\mathbf{o_1}, \ldots, \mathbf{o_j}\}$. In other words, $\mathbf{v_j}$ is in the span of $\{\mathbf{o_1}, \ldots, \mathbf{o_j}\}$, and thus

$$\text{span}(\mathbf{v_1}, \ldots, \mathbf{v_j}) \subseteq \text{span}(\mathbf{o_1}, \ldots, \mathbf{o_j}).$$

But both lists above are linearly independent, so $\text{span}(\mathbf{v_1}, \ldots, \mathbf{v_j})$ is a j-dimensional subspace of $\text{span}(\mathbf{o_1}, \ldots, \mathbf{o_j})$, which has also dimension j. So the two spaces are equal. Therefore, we obtained a set of *nonzero orthogonal* vectors $\mathbf{o_1}, \ldots, \mathbf{o_m}$ with the same span of $\mathbf{v_1}, \ldots, \mathbf{v_m}$. Let us "normalize", i.e. let us replace each $\mathbf{o_i}$ with $\mathbf{o_i'} \stackrel{\text{def}}{=} \frac{\mathbf{o_i}}{||\mathbf{o_i}||}$. This yields the desired orthonormal set, because obviously $\text{span}(\mathbf{o_1}, \ldots, \mathbf{o_m}) = \text{span}(\mathbf{o_1'}, \ldots, \mathbf{o_m'})$ and

$$\mathbf{o_i'} \bullet \mathbf{o_i'} = \frac{\mathbf{o_i}}{||\mathbf{o_i}||} \bullet \frac{\mathbf{o_i}}{||\mathbf{o_i}||} = \frac{1}{||\mathbf{o_i}||^2} (\mathbf{o_i'} \bullet \mathbf{o_i'}) = 1.$$

This proves that every nonzero subspace of $\mathbb{R}^n$ has an orthonormal basis. But it also proves something stronger, because of the following observation: if $\mathbf{v_1}$ and $\mathbf{v_2}$ are orthogonal, then $\mathbf{o_2} \stackrel{\text{def}}{=} \mathbf{v_2}$. More generally, if $\mathbf{v_1}, \ldots, \mathbf{v_j}$ are orthogonal, then $\mathbf{o_i} = \mathbf{v_i}$ for all $i \leq j$. And in particular, if we start with a list of vectors $\mathbf{v_1}, \ldots, \mathbf{v_m}$ the first j of which are orthonormal, then $\mathbf{o'_i} = \mathbf{o_i} = \mathbf{v_i}$ for all $i \leq j$. So if we have an orthonormal set of j vectors $\mathbf{v_1}, \ldots, \mathbf{v_j}$ in any subspace $W \subseteq \mathbb{R}^n$, since these vectors are LI we can complete them to a basis $\mathbf{v_1}, \ldots, \mathbf{v_j}, \mathbf{v_{j+1}}, \ldots, \mathbf{v_m}$ of W; and then when we can apply Gram-Schmidt to this basis, obtaining an orthonormal set that spans W, or in other words, an orthonormal basis for W, such that the first j vectors of this basis are exactly $\mathbf{v_1}, \ldots, \mathbf{v_j}$. □

Example 186. (See Example 8, Chapter 6 in book) Consider the three vectors

$$\mathbf{u_1} = (1,1,1)^\top, \quad \mathbf{u_2} = (0,1,1)^\top, \quad \mathbf{u_3} = (0,0,1)^\top.$$

The Gram-Schmidt algorithm, illustrated in the book, gives an orthogonal basis for $\mathbb{R}^3$, namely,

$$\mathbf{o_1} = (1,1,1)^\top, \quad \mathbf{o_2} = (-\frac{2}{3}, \frac{1}{3}, \frac{1}{3})^\top, \quad \mathbf{o_3} = (0, -\frac{1}{2}, -\frac{1}{2})^\top.$$

By normalizing, we get an orthonormal basis:

$$\mathbf{o'_1} = (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})^\top, \quad \mathbf{o'_2} = (-\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}})^\top, \quad \mathbf{o'_3} = (0, -\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})^\top.$$

You may object: But the three vectors we started with, in the previous example, were an LI set!, so they spanned $\mathbb{R}^3$ – and without any computation necessary, we already know an orthonormal basis for $\mathbb{R}^3$, namely $\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}$. So what is the point of using the Gram–Schmidt algorithm? Well, in the list of orthogonal vectors we found, *the first one is* $\mathbf{u_1}$. Note that none of $\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}$ is $\mathbf{u_1}$. This also explains why the Gram-Schmidt algorithm yields different results if you reshuffle the original list of vectors: If you started with

$$(0,0,1)^\top, (0,1,1)^\top, (1,1,1)^\top$$

the algorithm would output an orthogonal basis whose first vector is $(0,0,1)^\top$, and whose second vector $\mathbf{w}$ has the property that

$$\text{span}((0,0,1)^\top, (0,1,1)^\top) = \text{span}((0,0,1)^\top, \mathbf{w}^\top).$$