A Quick Review of Linear Algebra

Most of this handout consists of review of Math 2250 (Elementary Linear Algebra), which is listed as a pre-requisite for this course. *You should already have studied vector spaces, matrices and linear transformations.* If any of these basic notions are unfamiliar or too abbreviated, please refer back to your linear algebra textbook (or any undergraduate textbook introduction to the subject).

A *vector space* has two kinds of objects: **vectors** and **scalars**. We denote by V the set of vectors, and by F the field of scalars. (Typically F is $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{Q}$, although F can be any *field.* Later we shall define precisely what is meant by a field; for now it may help to imagine simply that $F = \mathbb{R}$.) It also has two operations: **vector addition**, which is an operation of the form

$$\text{vector} + \text{vector} = \text{vector};$$

and **scalar multiplication**, which is an operation of the form

$$\text{scalar} \times \text{vector} = \text{vector}.$$

These operations are required to satisfy the following properties, which we take as axioms: for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and all $a, b \in F$, we have

(V1) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$;
(V2) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$;
(V3) there exists $-\mathbf{v} \in V$ such that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$;
(V4) $1\mathbf{v} = \mathbf{v}$;
(V5) $(a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}$;
(V6) $a(b\mathbf{v}) = (ab)\mathbf{v}$;
(V7) $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$.

Any set of vectors and scalars, and choice of operations of vector addition and scalar multiplication, satisfying (v1)–(V7), is called a **vector space over the field** F. In general a vector space is not required to have any kind of vector multiplication (i.e. vector$\times$ vector = vector); any vector space with such an operation, satisfying additional properties of the form $(a\mathbf{u})\mathbf{v} = a(\mathbf{uv})$ and $(\mathbf{uv})\mathbf{w} = \mathbf{u}(\mathbf{vw})$, is called an **algebra**). There is always an operation of vector subtraction, but this is defined in terms of vector addition by $\mathbf{u} - \mathbf{v} = \mathbf{u} + (-\mathbf{v})$ so it is not necessary to list this separately among our requirements.

A remark concerning notation: So far we have distinguished vectors from scalars by writing vectors in bold face font as $\mathbf{u}$, $\mathbf{v}$, etc. This convention is followed by many popular

introductory textbooks do. Other textbooks may denote vector quantities instead as $\vec{u}$, $\vec{v}$, etc. or as $\underline{u}$, $\underline{v}$, etc. (The underline is simply the standard convention used by authors as an instruction to the publisher to use a bold face font.) More advanced textbooks simply denote vectors as u, v, etc., trusting that the reader will understand which quantities are scalars and which are vectors, based on the context; or perhaps using the fact that scalars and vectors are often denoted using opposite ends of the alphabet. Still other books use Roman letters $u, v, \ldots$ for vectors and Greek letters $\alpha, \beta, \ldots$ for scalars; and in other textbooks this convention is reversed!

We list some examples of vector spaces. In each case we mention the dimension of the vector space, although we have not yet formally defined dimension.

Example 1: Euclidean Spaces

Let $V = \mathbb{R}^3$, the set of all ordered triples of real numbers of the form (a, b, c) with $a, b, c \in \mathbb{R}$. In this case the scalar field is $F = \mathbb{R}$. Vector addition and scalar multiplication are defined componentwise:

$$(a, b, c) + (a', b', c') = (a+a', b+b', c+c'), \quad t(a, b, c) = (ta, tb, tc).$$

The zero vector is

$$\mathbf{0} = (0, 0, 0)$$

and additive inverses are defined componentwise as

$$-(a, b, c) = (-a, -b, -c).$$

This gives a 3-dimensional vector space known as **Euclidean 3-space** in which vectors are interpreted as 'arrows', each with its own magnitude and direction. Vector addition follows the usual 'parallelogram law'; scalar multiplication by $t \in \mathbb{R}$ consists in preserving the direction, and multiplying the magnitude of each vector by t (assuming $t > 0$), or reversing the direction and multiplying the magnitude of each vector by $|t|$ (assuming $t < 0$).

All of the preceding works just as well for $V = \mathbb{R}^n$ and $F = \mathbb{R}$, an n-dimensional vector space known as **Euclidean n-space**. The most familiar cases are $n = 2$ or 3, where vectors represent physical quantities such as velocity, force, acceleration, displacement, etc.

Example 2: Field extensions $E \supseteq F$

The set of complex numbers $V = \mathbb{C}$ is a 2-dimensional vector space over $F = \mathbb{R}$. More generally if F is any field and E is a field containing F as a subfield, then E is a vector space over F. For example every field is a (one-dimensional) vector space over itself. Also

$\mathbb{R}$ is a vector space over $\mathbb{Q}$ (of infinite dimension); also $\mathbb{C}$ is a vector space over $\mathbb{Q}$ (again of infinite dimension).

Example 3: F^n

Generalizing the first example, let F be any field of scalars (such as $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{Q}$) and let $V = F^n$ consisting of all n-tuples over F. This is an n-dimensional vector space over F. For example $\mathbb{C}^n$ is an n-dimensional vector space over $\mathbb{C}$; it may also be seen as a $2n$-dimensional vector space over $\mathbb{R}$.

Example 4: Polynomials

Let F be any field of scalars, and let $V = F[X]$, the set of all polynomials in X with coefficients in F. Then V is a vector space with the usual addition of polynomials, and multiplication by scalars. This is an infinite-dimensional vector space over F. Examples include $\mathbb{R}[X]$, $\mathbb{C}[X]$, $\mathbb{Q}[X]$, etc. The fact that polynomials have an extra operation of polynomial multiplication, means that $F[X]$ is more than a vector space: it is in fact an algebra.

Example 5: Functions

Let F be any field of scalars, such as $\mathbb{R}$. Let A be any set, and let V be the set of all real-valued functions defined on A. (For example A could be $\mathbb{R}$, or an interval like $[0,1] = \{x \in \mathbb{R} : 0 \leq x \leq 1\}$. Or A could be a subset of $\mathbb{R}^2$ or $\mathbb{R}^3$.) Addition of functions is defined pointwise as

$$(f+g)(x) = f(x) + g(x)$$

for all $f, g : A \to F$; $x \in A$ and scalar multiplication is also defined pointwise:

$$(af)(x) = a(f(x))$$

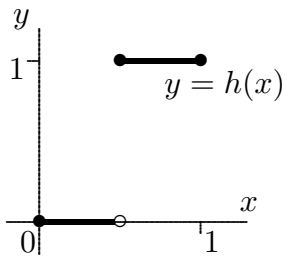
for every scalar $a \in F$ and function $f : A \to F$. Then V is a vector space over F. If A is an infinite set, then V is infinite-dimensional; in general the dimension of V is simply $|A|$, the number of elements in the domain A. As in the previous example, the fact that functions can be multiplied pointwise using the rule $(fg)(x) = f(x)g(x)$ means that V is more than just a vector space; it is an algebra.

Example 6: Continuous Functions

Modify the previous example by considering not *all* functions $A \to F$, but certain nice functions such as continuous functions. For example let V be the set of all continuous functions of the form $f : [0,1] \to \mathbb{R}$. The continuity requirement means that the graph

of f can be drawn in a single curve *without lifting one's pencil.* For example the function shown here:

$$h(x) = \begin{cases} 0, & \text{if } 0 \le x < 0.5; \\ 1, & \text{if } 0.5 \le x \le 1 \end{cases}$$



is *not* continuous, i.e. $h \notin V$. If f and g are continuous, and $c \in F$ is any scalar, then the functions $f + g$ and cf are also continuous. This means that continuous functions form a vector space (a subset of the space of all functions considered in the previous example, hence a *subspace.*) The vector space of all continuous functions $[0, 1] \to \mathbb{R}$ is infinite-dimensional; and because the product of any two continuous functions is continuous, we again have an algebra.

Example 7: Solutions of Differential Equations

This example will be accessible to students who have taken a first course in differential equations. Let $F = \mathbb{R}$ and let V be the set of all differentiable functions $f : \mathbb{R} \to \mathbb{R}$ such that $f''(x) + f(x) = 0$ where $f''(x)$ denotes the second derivative of $f(x)$. Then V is a 2-dimensional vector space over $\mathbb{R}$; in fact V consists of all functions of the form $a \sin x + b \cos x$ where $a, b \in \mathbb{R}$. The dimension indicates the number of initial conditions needed to specify a particular solution of the differential equation; for example every particular solution $f \in V$ is uniquely specified by the two values $f(0)$ and $f'(0)$; or by the two values $f(0)$ and $f\left(\frac{\pi}{2}\right)$. This is analogous to every complex number being uniquely specified by its real part and its imaginary part; or every vector in $\mathbb{R}^2$ being specified by its two coordinates.

More generally if $a_0, a_1, \ldots, a_n \in \mathbb{R}$ with $a_n \neq 0$, then the solutions of

$$a_0 f(x) + a_1 f'(x) + a_2 f''(x) + \cdots + a_n f^{(n)}(x) = 0$$

form an n-dimensional vector space over $F = \mathbb{R}$. (Here $f^{(n)}$ denotes the n-th derivative of f.) This observation lies at the historical roots of the concepts of vector space and dimension. If $\mathbb{R}^n$ were the only vector space of interest then we would have no need to introduce axioms; all properties of Euclidean space would be studied merely in the context of that one example. It was the observation that many notions of geometric vectors apply equally well in other contexts (such as the set V of solutions of a linear differential equation) that provided the historical motivation for developing linear algebra as a subject. The beauty of the axiomatic approach to linear algebra is that it leads to

a uniform description of vector spaces, allowing intuition (often learned in the context of more geometric examples) to be readily applied to other settings (including the current example of solutions of differential equations).

Example 8: The Trivial Vector Space

Let F be any field, and let $V = \{\mathbf{0}\}$; so V consists of just one vector, the zero vector. We have $\mathbf{0} + \mathbf{0} = \mathbf{0}$ and $a\mathbf{0} = \mathbf{0}$ for all $a \in F$. What's more to say about this example? Its dimension is in fact zero.

Subspaces

Let V be a vector space over a field F, and let $U \subseteq V$ be any set of vectors in V. We say that U is a **subspace** of V if U is also a vector space, using the usual operations of vector addition and scalar multiplication for V. This condition is more subtle than it looks, because it implicitly requires that

$$(\text{vector in } U) + (\text{vector in } U) = (\text{vector in } U)$$

and

$$(\text{scalar in } F) \times (\text{vector in } U) = (\text{vector in } U).$$

It also implicitly requires that the zero vector of V lies in U, and that

$$-(\text{vector in } U) = (\text{vector in } U).$$

Note that in order to prove that U is a subspace of V, one does not need to verify the various commutative, associative and distributive laws; if these hold in all of V then they must hold in U. The main thing to check is that U has the zero vector (in particular U is not empty), and U is closed under vector addition and scalar multiplication. (Closure is implicitly assumed for every vector space; hence it is a requirement in particular for subspaces.) For example the subspaces of $\mathbb{R}^3$ (see Example 1 above) are

(i) the trivial subspace consisting of just the origin $\mathbf{0} = (0, 0, 0)$;
(ii) lines of $\mathbb{R}^3$ passing through the origin;
(iii) planes of $\mathbb{R}^3$ passing through the origin; and
(iv) $\mathbb{R}^3$ itself.

The subspaces listed in (i)–(iv) have dimension 0, 1, 2, 3 respectively.

The set of continuous functions $[0, 1] \to \mathbb{R}$ (see Example 6) is a subspace of the space of all functions $[0, 1] \to \mathbb{R}$ (Example 5).

The set V_n consisting of all polynomials $f(X) \in F[X]$ of degree $< n$ is an n-dimensional subspace of $F[X]$, and we have

$$\{0\} = V_0 \subset V_1 \subset V_2 \subset V_3 \subset \cdots \subset F[X].$$

Basis, Span and Dimension

Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be vectors in a vector space V over a field F. Any vector of the form

$$\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n \in V, \quad \text{where } a_1, a_2, \ldots, a_n \in F$$

is called a **linear combination** of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. Note that there is no restriction on the number n of terms, but it must be finite. Let $S \subseteq V$ be an arbitrary set of vectors. The **span** of S, denoted $\langle S \rangle$, is the set of all linear combinations of vectors $v_1, v_2, \ldots, v_n \in S$. It is easy to see that $\langle S \rangle$ is a subspace of V; and we say that this subspace is **spanned by** S; in other words, S **spans** the subspace $\langle S \rangle$. (Note: We first introduced '*span*' as a noun, then used it as a verb.) For example if $V_n \subset F[X]$ is the subspace consisting of all polynomials of degree less than n, then $V_n = \langle 1, X, X^2, \ldots, X^{n-1} \rangle$. Also if $S = \{1, X, X^2, 1+X, 1+X^2\}$ then $\langle S \rangle = V_3$; but in this case S contains 'redundant' polynomials which do not contribute anything to $\langle S \rangle$. In this case we say that S is *linearly dependent*. In general a subset $S \subseteq V$ is **linearly dependent** if one of its elements can be expressed as a linear combination of other elements in S. Such 'redundant' elements can be deleted without changing $\langle S \rangle$. In the case of $S = \{1, X, X^2, 1+X, 1+X^2\}$ we may delete $1+X$ and $1+X^2$, leaving $1, X, X^2$ as a set of polynomials spanning V_3. Alternatively we may delete X and X^2, leaving $1, 1+X, 1+X^2$ as a set of three polynomials spanning V_3. It is a fact that any set of vectors spanning V_3 has at least three members; and that if it has more than three members, it is linearly dependent.

If $S \subseteq V$ is not linearly dependent, we say it is **linearly independent**. For example the set $\{1, 1+X, 1+X^2\}$ is linearly independent. A subset $B \subset V$ is a **basis** if it is linearly independent, *and* it spans V. For example $\{1, X, X^2\}$ is a basis for V_3. So is $\{1, 1+X, 1+X^2\}$. There are many choices of basis for V_3, but every basis for V_3 has 3 members. This is what it means to say that V_3 has dimension 3. We state,without proof, some facts which are proved in any standard book on linear algebra.

Let V be any vector space. Then V has a basis. If S is any set which spans V (such as V itself) then S contains a basis (i.e. there exists a subset $B \subseteq V$ which is a basis for V. We may obtain B from S by repeatedly deleting any 'redundant' vectors. Conversely, suppose $S \subset V$ is any linearly independent subset. Then there exists a basis for V which contains S, i.e. there exists a basis B for V such that $B \supseteq S$. We may obtain B from S by repeatedly adding vectors which are not in the span of the previously chosen vectors.

The vector space V typically has many choices of basis, but any two bases for V ('*bases*' is the plural of '*basis*') have the same number of members. The number of vectors in any basis for V is the **dimension** of V.

One may equally well define a basis by saying: B is a basis of V iff every vector in V can be written as a *unique* linear combination of elements of B. Thus for example $\{1, X\}$ fails to be a basis for V_3 since there are polynomials $f(X) \in V_3$, such as X^2, that cannot be written as linear combinations of 1 and X. Also the set $S = \{1, X, X^2, 1+X, 1+X^2\}$ fails to be a basis since there are polynomials $f(X) \in V_3$ which expressible in more than one way as linear combinations of S; for example $f(X) = 2 - X$ can be written in at least two different ways in terms of S as

$$f(X) = 2(1) + (-1)X = 1(1) + (-3)X + 1(1+X).$$

The fact that $B = \{1, X, X^2\}$ is a basis for V_3 follows from the fact that every $f(X) \in V_3$ is expressible in the form $f(X) = a{\cdot}1 + bX + cX^2$ for unique values of $a, b, c \in F$. The fact that $B' = \{1, 1+X, 1+X^2\}$ is also a basis for V_3 means that every $f(X) \in V_3$ can also be written in the form $f(X) = \alpha{\cdot}1 + \beta(1+X) + \gamma(1+X^2)$ for unique values of $\alpha, \beta, \gamma \in F$; indeed

$$a + bX + cX^2 = \alpha + \beta(1+X) + \gamma(1+X^2)$$

has as its unique solution $\alpha = a-b-c,\ \beta = b,\ \gamma = c$.

Linear Transformations

Let V and W be vector spaces over the same field F. A function $T : V \to W$ is called **linear** if $T(a\mathbf{u} + b\mathbf{v}) = aT(\mathbf{u}) + bT(\mathbf{v})$ for all $a, b \in F$ and all $\mathbf{u}, \mathbf{v} \in V$. Moreover such functions are called **linear transformations**.

For example let V be the vector space over $F = \mathbb{R}$ consisting of all continuous functions $f : [0, 1] \to \mathbb{R}$. Let $W = \mathbb{R}^3$ and define $T : V \to W$ by $T(f) = (f(0), f(1), f(2))$. We have

$$\begin{aligned} T(af + bg) &= (af(0){+}bg(0),\ af(1){+}bg(1),\ af(2){+}bg(2)) \\ &= a(f(0), f(1), f(2)) + b(g(0), g(1), g(2)) \\ &= aT(f) + bT(g) \end{aligned}$$

so T is linear.

As another example consider the derivative operator $D : F[X] \to F[X]$ defined by: $D(f(X) = f'(X)$, the usual derivative of $f(X)$. The linearity of D is clear from

$$\begin{aligned} D(af(X) + bg(X)) &= \tfrac{d}{dX}\big(af(X) + bg(X)\big) \\ &= af'(X) + bg'(X) \\ &= aD(f(X)) + bD(g(X)). \end{aligned}$$

Matrices

An $m \times n$ **matrix** over a field F is a rectangular array of scalars in F of the form

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Note that $a_{ij} \in F$ denotes the entry in the i-th row ($i = 1, 2, \ldots, m$) and the j-th column ($j = 1, 2, \ldots, n$). Given also an $n \times p$ matrix over F of the form

$$B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix}$$

then the matrix product

$$AB = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mp} \end{bmatrix}$$

is the $m \times p$ matrix with entries defined by

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}\,.$$

(Note that this is simply the dot product of the i-th row of A with the j-th column of B.) The matrix product AB is only defined only if the number of columns of A equals the number of rows of B. For example

$$\begin{bmatrix} 2 & -1 & 0 \\ 3 & 1 & 4 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 6 & -3 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} 4 & 3 \\ 29 & -7 \end{bmatrix}; \qquad \begin{bmatrix} 5 & 0 \\ 6 & -3 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 0 \\ 3 & 1 & 4 \end{bmatrix} = \begin{bmatrix} 10 & -5 & 0 \\ 3 & -9 & -12 \\ 1 & -3 & -4 \end{bmatrix}.$$

Evidently AB is different from BA in general, as this example shows; thus matrix multiplication is not commutative. Matrices can be added only if they have the same size, for example

$$\begin{bmatrix} 0 & 3 & -2 \\ -6 & 1 & 11 \end{bmatrix} + \begin{bmatrix} 2 & -1 & 1 \\ 0 & 5 & 13 \end{bmatrix} = \begin{bmatrix} 2 & 2 & -1 \\ -6 & 6 & 24 \end{bmatrix}.$$

Matrix operations satisfy the laws

$$(AB)C = A(BC); \qquad A(B + C) = AB + AC; \qquad (A + B)C = AC + BC$$

whenever these expressions are defined.

Matrices can be used to represent linear transformations on vector spaces. For example if $D : V_4 \to V_3$ is the derivative operator then

$$D(a + bX + cX^2 + dX^3) = b + 2cX + 3dX^2$$

can be represented in matrix form as

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} b \\ 2c \\ 3d \end{bmatrix}.$$

Systems of Linear Equations

Matrices can also represent systems of linear equations; for example the system

$$\begin{aligned} 2x \ - y \ \ \ \ \ &= \ 5, \\ x + 3y - 2z &= -3, \\ x - 2y \ + z &= \ 6 \end{aligned}$$

can be represented as

$$A\mathbf{x} = \mathbf{b} \qquad \text{where } A = \begin{bmatrix} 2 & -1 & 0 \\ 1 & 3 & -2 \\ 1 & -2 & 1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 5 \\ -3 \\ 6 \end{bmatrix}.$$

To solve this system, note that the inverse of the matrix A is given by

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

i.e. the product $A^{-1}A = AA^{-1} = I$ is the identity matrix:

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

To solve $A\mathbf{x} = \mathbf{b}$, multiply both sides on the left by A^{-1} to obtain

$$\mathbf{x} = I\mathbf{x} = A^{-1}A\mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} -1 & 1 & 2 \\ -3 & 2 & 4 \\ -5 & 3 & 7 \end{bmatrix} \begin{bmatrix} 5 \\ -3 \\ 6 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 8 \end{bmatrix}.$$

Thus the unique solution of the linear system is given by x=4, y=3, z=8. (Remark: Inverting matrices is time-consuming! This is not the most efficient computational method for solving large linear systems. Rather, the importance of this approach lies in its theoretical simplicity.)

Determinants

Let A be a square matrix, i.e. an $n \times n$ matrix for some $n \geq 1$. The **main diagonal** of $A = \left(a_{ij}\right)$ consists of the entries lying on the diagonal from the upper-left to the lower-right corner, i.e. the entries $a_{11}\,, a_{22}\,, \ldots, a_{nn}\,.$ We say A is **diagonal** if all entries not lying on the main diagonal, are zero. We say A is **upper triangular** if all entries *below* the main diagonal are zero. Similarly, A is **lower triangular** if all entries *above* the main diagonal are zero. The **determinant** of A, denoted $\det A$, is a scalar value satisfying the following properties.

(D1) If A is diagonal, or upper or lower triangular, then $\det A$ equals the product of the entries on the main diagonal.

(D2) If A' is obtained from A by adding a multiple of one row to another, *or* by adding a multiple of one column to another, then $\det A' = \det A$.

(D3) If A' is obtained from A by interchanging two rows, *or* by interchanging two columns, then $\det A' = \det A$.

(D4) If A' is obtained from A by multiplying some row by a scalar c, *or* by multiplying some column by a scalar c, then $\det A' = c \det A$.

Together these rules suffice to efficiently compute any determinant. For example if

$$A = \begin{bmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{bmatrix}$$

then we compute

$$\begin{aligned}
\det A &= \det \begin{bmatrix} 1 & a & a^2 \\ 0 & b{-}a & b^2{-}a^2 \\ 1 & c & c^2 \end{bmatrix} && \text{using (D2)} \\
&= \det \begin{bmatrix} 1 & a & a^2 \\ 0 & b{-}a & b^2{-}a^2 \\ 0 & c{-}a & c^2{-}a^2 \end{bmatrix} && \text{using (D2)} \\
&= (b{-}a) \det \begin{bmatrix} 1 & a & a^2 \\ 0 & 1 & b{+}a \\ 0 & c{-}a & c^2{-}a^2 \end{bmatrix} && \text{using (D4)} \\
&= (b{-}a)(c{-}a) \det \begin{bmatrix} 1 & a & a^2 \\ 0 & 1 & b{+}a \\ 0 & 1 & c{+}a \end{bmatrix} && \text{using (D4)} \\
&= (b{-}a)(c{-}a) \det \begin{bmatrix} 1 & a & a^2 \\ 0 & 1 & b{+}a \\ 0 & 0 & c{-}b \end{bmatrix} && \text{using (D2)} \\
&= (b{-}a)(c{-}a)(c{-}b) && \text{using (D1).}
\end{aligned}$$

Determinants are also denoted using vertical bars $|\ \ |$ as in the expression

$$\det A = \begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix} = (b-a)(c-a)(c-b).$$

A matrix A is invertible (i.e. it has an inverse matrix A^{-1} as described above) iff $\det A \neq 0$. Moreover determinants satisfy the rule

$$\det(AB) = (\det A)(\det B)$$

for any two $n \times n$ matrices A and B.

Fields

Finally, let us explain what a field is. A **field** is a set of numbers (called scalars) including special elements called zero and one (i.e. 0 and 1, which are required to be distinct) and two operations called addition and multiplication, such that for all $a, b, c \in F$ the following axioms hold:

(F1) $a + b = b + a$;
(F2) $(a + b) + c = a + (b + c)$;
(F3) $a + 0 = a$;

(F4) there exists an additive inverse $-a \in F$ such that $a + (-a) = 0$;
(F5) $ab = ba$;
(F6) $(ab)c = a(bc)$;
(F7) $1a = a$;
(F8) if $a \neq 0$, then there exists a multiplicative inverse $a^{-1} \in F$ such that $a^{-1}a = 1$;
(F9) $a(b + c) = ab + ac$.

The most important rule, which distinguishes fields (such as $\mathbb{R}$, $\mathbb{Q}$ or $\mathbb{C}$) from more general *rings* (such as $\mathbb{Z}$ or $\mathbb{R}[X]$) is (F8), which allows us to *divide* by any nonzero element $a \in F$: we define *division* by $b/a = a^{-1}b$. We also define *subtraction* in any field by $a - b = a + (-b)$.

The smallest field is the binary field $\mathbb{F}_2 = \{0, 1\}$ in which $1+1 = 0$ (this is the *integers mod 2.*) In fact for any prime p, the *integers mod p* forms a field $\mathbb{F}_p = \{0, 1, 2, \ldots, p-1\}$ with exactly p elements. If $a \in \mathbb{F}_p$ is nonzero, then to find its inverse $a^{-1} \in \mathbb{F}_p$ we note that $gcd(a, p) = 1$. By the Extended Euclidean Algorithm we can find $r, s \in \mathbb{Z}$ such that $ra + sp = 1$; then since $p = 0$ in $\mathbb{F}_p$ we obtain $ra = 1$ in $\mathbb{F}_p$ so $r = a^{-1}$. For example to find the inverse of $18 \in \mathbb{F}_{61}$ we first perform the Extended Euclidean Algorithm over $\mathbb{Z}$:

61	28	1
1	0	61
0	1	28
1	−2	5
−5	11	3
6	−13	2
−11	24	1

In $\mathbb{Z}$ we have $-11 \cdot 61 + 24 \cdot 28 = 1$, so in $\mathbb{F}_{61}$ we have $24 \cdot 28 = 1$, i.e. $28^{-1} = 24$. The field $\mathbb{F}_p$ is often denoted $\mathbb{Z}_p$ in undergraduate textbooks.

Fields are important because we can do linear algebra (in particular, solve linear equations) over any field of scalars. The usual algorithm for solving linear systems (Gaussian elimination, which consists of performing a sequence of elementary row operations on a matrix) requires division, and so can only proceed over a field. Fields are also important in our previous work since the Division Algorithm requires that the coefficients in our polynomials belong to a field.