Quick Summary.

1. Given any square matrix $A = (a_{ij})$, we define its **determinant** variously denoted as $\det(A)$ or $|A|$ or $||A||$. The definition needs some auxiliary terms.
   - First for any (i, j) such that a_{ij} is defined, we define the **index of** $(\mathbf{i}, \mathbf{j})$ to be $\mathrm{ind}(i, j) = (-1)^{(i+j)}$.
   - Further we define the **minor of** $(\mathbf{i}, \mathbf{j})$ to be $\mathrm{minor}(A, i, j)$ which is the subdeterminant of A obtained by throwing away the i-th row and the j-th column.
   - We also define the **cofactor of** $(\mathbf{i}, \mathbf{j})$ to be $\mathrm{cofactor}(A, i, j) = \mathrm{ind}(i, j)\mathrm{minor}(A, i, j)$.
   - Then the determinant $|A|$ can be computed as $\sum a_{ij}\mathrm{cofactor}(A, i, j)$ where the sum is taken over all entries a_{ij} **coming from any chosen row or column.**

2. The value of the determinant gets multiplied by k if all entries in a single row or column are multiplied by k.

3. The value of the determinant gets multiplied by -1, if a single exchange of two rows or two columns is carried out. For more complicated permutations, we multiply by the sign of the permutation.

4. The value of the determinant is unchanged if we add a multiple of one row to another. Similar result holds for columns.

5. The value of a determinant is 0 if some row or column consists of zero entries only! (This follows from the definition.)

6. The value of a lower triangular determinant is equal to the product of its diagonal entries. Ditto for upper triangular. In general, this is how determinants are computed: reduce the determinant to upper or lower triangular form and then evaluate the product of the diagonal entries. If permutations are used along the way, then suitable sign is attached to the answer. Sometimes, an expansion along a suitable row/column is also used to reduce the work.

7. There is a more general expansion, the so-called Laplace expansion, which works with several rows (or columns) at once, instead of the single row (or column) as in the definition.

8. The adjoint of a matrix A is a matrix denoted by A^{adj} whose (i, j)-th entry is equal to $\mathrm{cofactor}(A, j, i)$. **Do notice the switch in the order!** The adjoint satisfies the identity

   $$AA^{\mathrm{adj}} = A^{\mathrm{adj}}A = |A|I.$$

   This lets us write the inverse of A as $A^{\mathrm{adj}}/|A|$. Of course it exists iff $|A| \neq 0$.

   We often use the notation $\mathrm{adj}(A)$ in place of A^{adj}.

9. For a general matrix M **its rank** $\mathrm{rank}(M)$ is defined to be the largest number r such that M has a nonzero subdeterminant of size r. Thus a square $n \times n$ matrix is invertible iff its rank is n. Rank of a matrix is obviously less than or equal to its rownum as well as colnum.

10. In general, **the equations $\mathbf{AX} = \mathbf{B}$ are solvable** iff $\text{rank}(A) = \text{rank}(A|B)$. Here $A|B$ stands for the augmented matrix. Obviously, if $\text{rank}(A) = \text{rownum}(A)$ then $AX = B$ is solvable for all B. The converse is true too!

11. **Cramer's Rule.** If we wish to solve a system $AX = B$ where A is a square $n \times n$ matrix we proceed thus.

    Define a convenient notation: $A(v, i)$ which is obtained from A by swapping the i-th column of A with the column v.

    We claim that when $AX = B$, then $\det(A)X_i = \det(A, B, i)$. The proof can be seen as follows. Consider $adj(A)AX = adj(A)B$. Since $adj(A)A = \det(A)I$ by comparing i-th row entries on both sides, we deduce that $\det(A)X_i = Row(adj(A), i)B$. If we recall the definition of the adjoint, we know that the i-th row of $adj(A)$ gives the sequence of cofactors of the i-th column of A. Thus the RHS is nothing but the $\det(A(B, i))$.

    - Assume $\det(A) \neq 0$. Then the above calculation gives a formula for the solution, namely $X_i = \frac{\det(A,B,i)}{\det(A)}$. This is the Cramer's rule.
    - If $\det(A) = 0$ but one of $\det(A(B, i))$ is not zero, then we get that the system is inconsistent.
    - If $\det(A) = 0$ and $\det(A(B, i)) = 0$ for all i, then it can be seen that one of the equations can be dropped (being dependent on the others) and one of the variables becomes free. Moving the free variables entries to the RHS (and treating them as constants) we reduce the problem to a smaller sized determinant!

      Example: Consider the system $AX = B$ where $A = \begin{pmatrix} 1 & 1 & 3 \\ 2 & 1 & 4 \\ 4 & 3 & 10 \end{pmatrix}$ and $B = \begin{pmatrix} -2 \\ 4 \\ 0 \end{pmatrix}$. We check that $\det(A) = 0$ and all $\det(A(B, i)) = 0$. In fact we see that the third equation is simply the sum of the second equation with twice the first. So, we drop it and see if some pair of variables (corresponding to a pair of columns) give a nonzero 2×2 determinant. We see that $\begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$ has a nonzero determinant, so the (remaining) third variable (say X_3) can be free. Thus we get a new pair of equations $PY = Q$ where $P = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$, $Y = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ and $Q = \begin{pmatrix} -2 - 3X_3 \\ 4 - 4X_3 \end{pmatrix}$. This can now be solved by the above rule, leaving X_3 free.

Theory behind Determinants.

After having learned how to compute a determinant of a square matrix, we now discuss the basic idea behind the notion of determinants. A determinant is a number associated with a sequence of n vectors in $\Re^n$. It is supposed to help us determine the various properties of the geometric object described by these vectors.

Thus, in $\Re^2$ consider two vectors which can be thought of as two arrows coming out from the origin. They define a parallelogram. If one of the vectors is simply a multiple of the other, then the parallelogram collapses to a line. If we take the two vectors $v_1 = \begin{pmatrix} a \\ b \end{pmatrix}$ and $v_2 = \begin{pmatrix} c \\ d \end{pmatrix}$, then we can form the determinant

$$\det(v_1, v_2) = \det(\begin{pmatrix} a & c \\ b & d \end{pmatrix}) = ad - bc = \Delta \text{ say.}$$

With some geometric calculations, it is easy to establish that $|\Delta|$ gives the area of the parallelogram and moreover, the sign of Δ even gives us a measure of the angle from v_1 to v_2.

Here is the calculation. Using the usual idea of polar coordinates, we can write

$$v_1 = \begin{pmatrix} a \\ b \end{pmatrix} = r_1 \begin{pmatrix} \cos(\theta_1) \\ \sin(\theta_1) \end{pmatrix} \text{ where } r_1 = \sqrt{a^2 + b^2}$$

and θ_1 is the angle made by v_1 with the x-axis measured counterclockwise.

Similarly, we write

$$v_2 = r_2 \begin{pmatrix} \cos(\theta_2) \\ \sin(\theta_2) \end{pmatrix} \text{ where } r_2 = \sqrt{c^2 + d^2}$$

and θ_2 is the angle made by v_2 with the x-axis measured counterclockwise.

Then we have

$$\Delta = r_1 r_2 \left(\cos(\theta_1)\sin(\theta_2) - \sin(\theta_1)\cos(\theta_2)\right) = r_1 r_2 \sin(\theta)$$

where $\theta = \theta_2 - \theta_1$ which is the angle measured from v_1 to v_2 counterclockwise.

Thus the determinant gives the signed area of the parallelogram constructed from v_1 towards v_2. It also lets us decide if the vectors are linearly dependent. We note that for dependent vectors, the angle θ is either 0 or π. In either case $\sin(\theta) = 0$. Thus $\Delta = \det(v_1, v_2) = 0$ iff v_1, v_2 are linearly dependent. [1]

We also note three natural properties enjoyed by these determinants:

1. If a vector is scaled by multiplying by a constant p then the determinant also gets multiplied by the same p.

2. If the vectors are swapped, the determinant gets multiplied by -1.

3. If a vector w is equal to $w_1 + w_2$, then

   $$\det(w, v) = \det(w_1, v) + \det(w_2, v).$$

   This is easily checked from the formula and it can also be geometrically verified for the areas of parallelograms.

[1] We have not paid careful attention to the zero vector in this analysis. Fortunately, when one of the two vectors is zero, then the vectors are clearly linearly dependent and also the $\det(v_1, v_2) = 0$. Thus, our conclusion is easily verified!

We can appeal to the geometric argument in three space $\Re^3$ as well. But in higher dimensions, we don't have a preconceived idea of "volumes". We therefore take a clue from the algebraic calculations and define the n-dimensional determinant as follows.

Axioms of determinants.

1. Given any ordered n-tuple $v_1, v_2, \cdots, v_n$ in $\Re^n$ the expression $\det(v_1, v_2, \cdots, v_n)$ shall be a well defined real number. Thus det is a function from n-tuples of vectors to $\Re$.

2. **Alternating property.** This function shall be an alternating function. This means, if any two vectors v_i, v_j are exchanged, then the value of the determinant shall be multiplied by -1.

3. **Linearity.** The determinant shall be a linear function of each of its arguments $v_1, v_2, \cdots, v_n$. This means:

$$\det(w_1 + w_2, v_2, \cdots, v_n) = \det(w_1, v_2, \cdots, v_n) + \det(w_2, v_2, \cdots, v_n)$$

   and

$$\det(kv_1, v_2, \cdots, v_n) = k\det(v_1, v_2, \cdots, v_n) \text{ for any } k \in \Re.$$

   Similar conditions hold for each of the n arguments.[2]

4. If v_1, v_2, v_n are the unit vectors, i.e. the matrix with columns $v_1, v_2, \cdots v_n$ in order gives the identity matrix I_n, then $\det(v_1, v_2, \cdots, v_n) = 1$.

Outline of the argument.

It can be shown that subject to these conditions, there exists a unique determinant function. Moreover, it satisfies all the properties that we informally asserted.

Here is a sketch of the argument.

1. We prove the result by induction on the size n of the determinant.

   Our inductive statement is:

   Let A be the matrix formed by the vectors $v_1, v_2, \cdots, v_n$ as columns. We note that the entries of v_1 are $A(1,1), A(2,1), \cdots, A(n,1)$.

   For any $i = 1, 2, \cdots n$ consider vectors $w_2^i, w_3^i, \cdots, w_n^i$ in $\Re^{n-1}$ obtained by dropping the i-th entry from each of the $n-1$ vectors $v_2, v_3, \cdots, v_n$.

   Then

$$\det(v_1, v_2, \cdots, v_n) = \det(A) = A(1,1)cofactor(A,1,1) + A(2,1)cofactor(A,2,1) + \cdots + A(n,1)cofacto$$

   where

$$cofactor(A,i,1) = (-1)^{i+1}minor(A,i,1) = (-1)^{i+1}\det(w_2^i, w_3^i, \cdots w_n^i).$$

2. Thus the starting case shall be $n = 2$ and here we know everything already.

[2] A little thought may show that in view of the alternating condition, it would be enough to assume this condition just for the argument v_1, since we can deduce it for other arguments by swapping the vectors and then swapping them back.

3. Now assume the result for $n = 2$.

   For $n = 3$, our formula gives the usual expansion for the first column. The formula is easily seen to satisfy the linearity condition for the first argument v_1 and the linearity as well as the alternating property for the second and third components is easily seen from the inductive assumption.

   The only non trivial calculation is the proof that $\det(v_1, v_2, v_3) = -\det(v_2, v_1, v_3)$. This is checked by an easy but tedious calculation of expanding both determinants completely.

4. Now assume the result for all determinants of size less than or equal to $n-1$ and define $\det(v_1, v_2, \cdots, v_n)$ as in the statement, where the cofactors are defined and satisfy known properties by induction hypothesis.

   We note that the linearity of the formula is again easily seen to be true and as before, it is enough to prove that $\det(v_1, v_2, \cdots, v_n) = -\det(v_2, v_1, \cdots, v_n)$. We may either do the easy but tedious calculation or make the following shortcut. Note that any vector in $\Re^n$ can be written as a sum of at most n-vectors which have only one non zero entry. Using this, we can assume that our v_1 and v_2 each have only one non zero entry. In that case, the cofactor expansions by the first two vectors give an easy formula for both sides.

   Suppose v_1 has only one non zero entry p in ith position and v_2 has only one non zero entry in position j. We invite the reader to verify these statements:

   - If $i = j$, then $\det(v_1, v_2, \cdots, v_n) = \det(v_2, v_1, \cdots, v_n) = 0$. So the claim is true.
   - If $i < j$ then $\det(v_1, v_2, \cdots, v_n) = (-1)^{i+j-1}\Delta$ where Δ is the determinant formed by vectors $w_3, \cdots w_n$ which are obtained from $v_3, \cdots, v_n$ after droppin their i-th and j-th entries. Moreover $\det(v_2, v_1, \cdots, v_n) = (-1)^{j+i}\Delta$. Thus $\det(v_1, v_2, \cdots, v_n) = -\det(v_2, v_1, \cdots, v_n)$.
   - If $i > j$, then the calculation is similar, except $\det(v_1, v_2, \cdots, v_n) = (-1)^{i+j}\Delta$ and $\det(v_2, v_1, \cdots, v_n) = (-1)^{j+i-1}\Delta$.

5. Thus, the inductive step is complete and the result is proved.

Once the existence of the determinant function is proved, it is easy to prove uniqueness as well as all the properties asserted in the quick introduction.