DETERMINANTS

1. Solving linear equations

The simplest type of equations are *linear*. The equation

$$ax = b \tag{1}$$

is a linear equation, in the sense that the function $f(x) = ax$ is linear[1] and it is equated to a value b, in the sense that we ask: for which values of the variable x does the function attains the value b. The answer is: for a unique x, namely b/a **if** $a \neq 0$. But if $a = 0$, then there is no such x if $b \neq 0$, while if $b = 0$ then there are infinitely many x satisfying this equation (all x in fact).

We can have more equations with more unknowns. Here is another example: solve the *linear system of equations*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned} \tag{2}$$

in the sense that we need to find all the values of ordered paris of unknowns (x_1, x_2) for which the *linear transformation* which takes (x_1, x_2) to the ordered pair $(a_{11}x_1 + a_{12}x_2, a_{21}x_1 + a_{22}x_2)$ has values equal to (b_1, b_2). A little work and ingenuity leads us to the formulas

$$x_1 = \frac{-a_{12}b_2 + a_{22}b_1}{a_{11}a_{22} - a_{12}a_{21}}, \quad x_2 = \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{12}a_{21}} \tag{3}$$

which means that if $a_{11}a_{22} - a_{12}a_{21} \neq 0$ then the system (2) has a unique solution. Otherwise, we can easily find that, just as in the previous example, we can have either no solutions (say, if all $a_{ij} = 0$ and one of the $b_i \neq 0$), or infinitely many (say, if all $a_{ij},\ b_j = 0$); it is not immidiately clear that the existence of a unique solution is ruled out in this case.

A general linear system of m linear equations with n unknowns, has the form

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ &a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ &\vdots \\ &a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{aligned} \tag{4}$$

which we need to solve for the unknowns $(x_1, x_2, \ldots, x_n)$. It will turn out that solutions of a general (4) behave much like in the previous simple examples: they can be given by algebraic operations, like (3), and the following situations can occur:

1. the system is *compatible*, i.e. it has solutions, in which case either

1a. the system is *determinate*, i.e. the solution is unique, or

1b. the system is *indeterminate*, i.e. there are more than one solutions (there will be infinitely many) or,

2. the system is *incompatible*, i.e. it has no solutions.

In the case the system is compatible we also want to describe its solutions.

Finally, what kind of numbers can $a_{ij},\ b_j\ x_j$ be, and can they be more general mathematical objects? Just like in the example (2), (3), general linear systems and their solutions can be written using the four basic

[1] From now on, ony functions of the tyoe $f(x) = ax$ will be called *linear*, while functions of the type $f(x) = ax + n$ will be called *linear-affine*.

arithmetic operations, satisfying the usual properties of commutativity, associativity, and distributivity: a set endowed with two operations which have the usual properties of addition and multiplication of real numbers is called a *field*. For example, the set of real numbers, $\mathbb{R}$ is, of course, a field, and so is the set of complex numbers $\mathbb{C}$, and, perhaps a less familiar, but very useful, is the residue classes modulo a prime number, $\mathbb{Z}_p$.

2. Determinants

We noticed the curious phenomenon that the formula (3) for the general solution of two equations with two unknowns has the same expression at the denominator of both x_1 and x_2, hence if this special number is not zero, then the system is determinate. With some work, one can find a similar phenomenon for a system of 3 equations with 3 unknowns: the system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \end{aligned} \tag{5}$$

has the solution

$$x_1 = \frac{a_{12}a_{23}b_3 - a_{12}a_{33}b_2 - a_{13}a_{22}b_3 + a_{13}a_{32}b_2 + a_{22}a_{33}b_1 - a_{23}a_{32}b_1}{a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}}, \tag{6}$$

$$x_2 = -\frac{a_{11}a_{23}b_3 - a_{11}a_{33}b_2 - a_{13}a_{21}b_3 + a_{13}a_{31}b_2 + a_{21}a_{33}b_1 - a_{23}a_{31}b_1}{a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}},$$

$$x_3 = \frac{a_{11}a_{22}b_3 - a_{11}a_{32}b_2 - a_{12}a_{21}b_3 + a_{12}a_{31}b_2 + a_{21}a_{32}b_1 - a_{22}a_{31}b_1}{a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}}$$

(see the Maple example to help with solving and writing up these formulas.)

In (6) the denominators are the same, again. Note that in (3) the denominator is made up od products of two of the a_{ij} having a plus sign or a minus sign in front. In (6) the denominators are products of three of the a_{ij} with some sign. You may also note that the products look like $a_{1i}a_{2j}a_{3k}$ where the numbers i, j, k are some permutation of the numbers $1, 2, 3$. And that all such possibilities are present. This is called the *determinant* of the matrix formed by the a_{ij}'s. You may note that the numerators also look like determinants...they are! We will now define the determinant of a general square matrix.

2.1. The definition of the determinant.

A *permutation* of the set of numbers $S_n = \{1, 2, \ldots, n\}$ is a function $\sigma : S_n \to S_n$ which is one-to-one[2] and onto[3]. Therefore the numbers $\sigma(1), \sigma(2), \ldots, \sigma(n)$ are just the numbers $1, 2, \ldots, n$ written in a different order.[4]

You may recall that the number of all possible permutations of $\{1, 2, \ldots, n\}$ is $n! := 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$. (Why?)

A simple notation for a permutation σ is $(\sigma(1), \sigma(2), \ldots, \sigma(n))$.

Suppose we have a permutation $\sigma : S_n \to S_n$. If for some $i < j$ it happens that $\sigma(i) > \sigma(j)$ then this is called an *inversion* of σ. Consider for example, the permutation $(4, 1, 3, 2)$. This means that $\sigma(1) = 4$, $\sigma(2) = 1$, $\sigma(3) = 3$, $\sigma(4) = 2$. Since $4 = \sigma(1) > 1 = \sigma(2)$, this is an inversion, and so are $4 = \sigma(1) > 3 = \sigma(3)$, and $4 = \sigma(1) > 2 = \sigma(4)$, and $3 = \sigma(3) > 2 = \sigma(4)$.

Denote by $N(\sigma)$ the number of inversions of the permutation σ.

---

[2] A one-to-one function (an injection) is a function σ which takes each value no more than once, that is, if $\sigma(x_1) = \sigma(x_2)$ then necessarily $x_1 = x_2$.

[3] A function $\sigma : A \to B$ is called onto B (it is surjective) if it takes every value in B, that is, for all $y \in B$ there is (at least one) $x \in A$ so that $\sigma(x) = y$.

[4] A function $\sigma : A \to B$ which is one-to-one and onto (a bijection) takes every value in B exactly once.

The *signature* of a permutation σ is $\operatorname{sgn}(\sigma) = (-1)^{N(\sigma)}$. That is, $\operatorname{sgn}(\sigma) = 1$ if σ has an even number of inversions and $\operatorname{sgn}(\sigma) = -1$ if it has an odd number.

Example. Let us consider all the permutations of $S_2 = \{1, 2\}$. There are $2! = 2$ possible permutations: σ_0, the permutation $1, 2$ (the identity, which is always a permutation) with signature 1 and σ_1, the permutation $2, 1$, with signature -1. Note that the denominator in (3) has the pleasant expression $a_{11}a_{22} - a_{12}a_{21} = \operatorname{sgn}(\sigma_0)a_{1\sigma_0(1)}a_{2\sigma_0(2)} + \operatorname{sgn}(\sigma_1)a_{1\sigma_1(1)}a_{2\sigma_1(2)}$.

Exercise. Find all the permutations of $S_3 = \{1, 2, 3\}$ and their signature. Find a formula for the denominator in (6).

Definition. The determinant of an $n \times n$ (square!) matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \tag{7}$$

is the number

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \sum_{\text{all permuations } \sigma} \operatorname{sgn}(\sigma)\, a_{1\sigma(1)}a_{2\sigma(2)} \dots a_{n\sigma(n)} \tag{8}$$

Remark 1. *(i) Each product $a_{1\sigma(1)}a_{2\sigma(2)} \dots a_{n\sigma(n)}$ contains exactly one a_{ij} from each row and from each column of the matrix A.*

(ii) With this geometric view, and inversion happens when the position of $a_{i\sigma(i)}$ in the matrix A is higher that that of $a_{j\sigma(j)}$ (this means $i < j$) and $a_{i\sigma(i)}$ is situated to the right of $a_{j\sigma(j)}$ (which means that $\sigma(i) > \sigma(j)$).

Exercise. Show that the denominators in (3), (6) equal $\det[a_{ij}]$. Show that the numerators are also determinants.

2.2. **Properties of determinants.**

2.2.1. *Determinant of the transpose matrix.* Recall that the *transpose*, A^T, of a matrix A is the matrix having columns equal to the rows of A, in the same order:

$$\text{if } A = [a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \quad \text{then} \quad A^T = [a_{ji}] = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & & & \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{bmatrix}$$

Of course, $(A^T)^T = A$.

By the definition (**??**) we have $\det A^T = \sum_\sigma (-1)^{N(\sigma)} a_{\sigma(1)1}a_{\sigma(2)2} \dots a_{\sigma(n)n}$. Each term in the sum can be reordered: $a_{\sigma(1)1}a_{\sigma(2)2} \dots a_{\sigma(n)n} = a_{1\tau(1)}a_{2\tau(2)} \dots a_{n\tau(n)}$ where τ is the permutation so that $\tau(j) = i$ whenever $\sigma(i) = j$ (it is the inverse function). Noting that $N(\sigma) = N(\tau)$ it follows that

$$\det A^T = \det A$$

2.2.2. *Antisymmetry in rows and columns:* If we exchange two columns (or two rows), the determinant changes sign:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{j1} & \dots a_{k1} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{j2} & \dots a_{k2} & \dots & a_{2n} \\ \vdots & & & \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{jn} & \dots a_{kn} & \dots & a_{nn} \end{vmatrix} = - \begin{vmatrix} a_{11} & a_{12} & \dots & a_{k1} & \dots a_{j1} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{k2} & \dots a_{j2} & \dots & a_{2n} \\ \vdots & & & \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{kn} & \dots a_{jn} & \dots & a_{nn} \end{vmatrix}$$

In particular

Corollary 2. *If two columns (or two rows) are equal, then the determinant is zero.*

2.2.3. *Linearity in rows (and in columns).* It is immediately seen from the definition (**??**) that

(9) $$\begin{vmatrix} a_{11} & a_{12} & \dots & \alpha b_{j1} + \beta c_{j1} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & \alpha b_{j2} + \beta c_{j2} & \dots & a_{2n} \\ \vdots & & & \vdots & & \\ a_{n1} & a_{n2} & \dots & \alpha b_{jn} + \beta c_{jn} & \dots & a_{nn} \end{vmatrix} =$$

$$\alpha \begin{vmatrix} a_{11} & a_{12} & \dots & b_{j1} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & b_{j2} & \dots & a_{2n} \\ \vdots & & & \vdots & & \\ a_{n1} & a_{n2} & \dots & b_{jn} & \dots & a_{nn} \end{vmatrix} + \beta \begin{vmatrix} a_{11} & a_{12} & \dots & c_{j1} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & c_{j2} & \dots & a_{2n} \\ \vdots & & & \vdots & & \\ a_{n1} & a_{n2} & \dots & c_{jn} & \dots & a_{nn} \end{vmatrix}$$

In particular, *a common factor can be extracted from a column (or row)*:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & \alpha b_{j1} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & \alpha b_{j2} & \dots & a_{2n} \\ \vdots & & & \vdots & & \\ a_{n1} & a_{n2} & \dots & \alpha b_{jn} & \dots & a_{nn} \end{vmatrix} = \alpha \begin{vmatrix} a_{11} & a_{12} & \dots & b_{j1} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & b_{j2} & \dots & a_{2n} \\ \vdots & & & \vdots & & \\ a_{n1} & a_{n2} & \dots & b_{jn} & \dots & a_{nn} \end{vmatrix}$$

And if $\alpha = 0$:

Corollary 3. *If a column (or a row) consists only of zeros, then the determinant is zero.*

2.2.4. *Adding to a column a multiple of another column.*

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{j1} + \lambda a_{k1} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{j2} + \lambda a_{k2} & \dots & a_{2n} \\ \vdots & & & \vdots & & \\ a_{n1} & a_{n2} & \dots & a_{jn} + \lambda a_{kn} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{j1} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{j2} & \dots & a_{2n} \\ \vdots & & & \vdots & & \\ a_{n1} & a_{n2} & \dots & a_{jn} & \dots & a_{nn} \end{vmatrix}$$

2.3. **Expansion of a determinant with respect to the a row (or column). Minors.** Fix some row, say the ith one. In the sum (8) collect all the terms containing a_{ij}; we obtain $a_{ij}A_{ij}$ where A_{ij} is a sum or difference of $n-1$ elements $a_{\ell k}$ where $\ell \neq i$ and $j \neq k$. Thus A_{ij} is the same if we replace the row i and column j with anything, say with 0's. A_{ij} is called *the cofactor of* a_{ij} *in the determinant* $\det A$.

Fis a row, i. Collecting all the terms containing a_{i1}, then a_{i2} etc. formula (8) becomes

(10) $$\det A = a_{i1}A_{i1} + a_{i2}A_{i2} + \dots + a_{in}A_{in}$$

called the expansion of the determinant along the row i.

We can similarly expand along a column j:

(11) $$\det A = a_{1j}A_{1j} + a_{2j}A_{2j} + \dots + a_{nj}A_{nj}$$

We next show how to express the cofactors in terms of determinants of order $n-1$; thus we can evaluate determinants by reducing the order.

Denote by M_{ij} the determinant of the matrix obtained from A by deleting the row i and column j. M_{ij} is called a *minor*.

Theorem 4. *The cofactors are expressed in terms of minors as*

$$A_{ij} = (-1)^{i+j} M_{ij}$$

As a consequence

$$\det A = (-1)^{i+1} a_{i1} M_{i1} + (-1)^{i+2} a_{i2} M_{i2} + \ldots + (-1)^{i+n} a_{in} M_{in}$$

and

$$\det A = (-1)^{1+j} a_{1j} M_{1j} + (-1)^{2+j} a_{2j} M_{2j} + \ldots + (-1)^{n+j} a_{nj} M_{nj}$$

Indeed, clearly $A_{11} = M_{11}$. For other i, j in the matrix A we move the row i until it gets to position 1, being careful to preserve the order of the rows: we need $i-1$ switches with the nearest upper row. Similarly, we switch columns $j-1$ times. The minor in position $1, 1$ of the new matrix equals $(-1)^{i+j-2} M_{ij}$, hence our formula. □

Examples.

Determinants of order two:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

To calculate determinants of order three we can expand along any row or column. Say we choose to expand along the second row:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = -a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{22} \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} - a_{23} \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{33} \end{vmatrix}$$

The determinant of a diagonal matrix: assume all $a_{ij} = 0$ for $i \neq j$

$$\begin{vmatrix} a_{11} & 0 & \ldots & 0 \\ 0 & a_{22} & \ldots & 0 \\ \vdots & & & \\ 0 & 0 & \ldots & a_{nn} \end{vmatrix} = a_{11}a_{22}\ldots a_{nn}$$

More generally, the determinant of a (upper, or lower) triangular matrix:

$$\det \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ 0 & a_{22} & \ldots & a_{2n} \\ \vdots & & & \\ 0 & 0 & \ldots & a_{nn} \end{bmatrix} = a_{11}a_{22}\ldots a_{nn} = \det \begin{bmatrix} a_{11} & 0 & \ldots & 0 \\ a_{21} & a_{22} & \ldots & 0 \\ \vdots & & & \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}$$

2.4. **Testing linear dependence of vectors in $\mathbb{R}^n$.** Suppose we have k vectors in $\mathbb{R}^n$:

$$\mathbf{x}_i = (a_{i1},\, a_{i2}, \ldots, a_{in}), \;\; i = 1, 2 \ldots, k$$

and the question is to determine if they are linearly dependent or not. If one of them is $\mathbf{0}$, then they are dependent, so in the following we assume they are all not zero.

If $k \geqslant n+1$ then they are linearly dependent (why?). So assume $k \leqslant n$.

Assume they are dependent. Then one of them is a linear combination of the others, say $\mathbf{x}_k = \sum_{i=1}^{k-1} \lambda_i \mathbf{x}_i$. Let us form the matrix $A = [a_{ij}]$. Is may not be a square matrix, but it contains lots of square matrices: choose any k columns then we have a $k \times k$ matrix. All their determinants are zero by §2.2.4 and Corollary 3.

Conversely, if we find some $k \times k$ determinant not zero, then our k vectors are linearly independent.

2.5. **Cramer's rule.** We first note that

$$a_{1k}A_{1j} + a_{2k}A_{2j} + \ldots a_{nk}A_{nj} = 0 \quad \text{if } k \neq j \tag{12}$$

Indeed, by (11) this is the determinant of a matrix obtained from A by replacing its column j by a copy of its column k; its determinant is zero by Corollary 2.

We can now solve systems with the same number of equations and unknowns:

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ &a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ &\vdots \\ &a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n \end{aligned} \tag{13}$$

if we assume $\det[a_{ij}] \neq 0$.

If $x_1, x_2, \ldots, x_n$ are numbers solving this system, we multiply the first equation by the cofactor A_{11}, the second one by A_{21} and so on, and we add these, obtaining

$$\begin{aligned} (a_{11}A_{11} + a_{21}A_{21} + \ldots + a_{n1}A_{n1})\, x_1 & \\ + (a_{12}A_{11} + a_{22}A_{21} + \ldots + a_{n1}A_{n2})\, x_2 & \\ + \ldots + (a_{1n}A_{11} + a_{2n}A_{21} + \ldots + a_{nn}A_{n1})\, x_n & \\ = A_{11}b_1 + A_{21}b_2 + \ldots + A_{n1}b_n & \end{aligned} \tag{14}$$

The coefficient of x_1 in (14) equals $\det A$ by (10), and the coefficients of $x_2, \ldots, x_n$ are zero by (12). The right hand side of (14) is the deteminant of the matrix B_1 obtained from A by replacing its first column with the numbers $b_1, b_2 \ldots, b_n$. We obtained

$$x_1 = \frac{\det B_1}{\det A}$$

We proceed similarly for all $j = 2, 3, \ldots, n$ by multiplying the first equation by A_{1j}, the second one by A_{2j} and adding them up:

$$x_j = \frac{\det B_j}{\det A}, \quad j = 1, 2, \ldots, n$$

where B_j is the matrix obtained from A by replacing its jth column with $b_1, b_2 \ldots, b_n$.

We should still show that these expressions are indeed solutions, which we leave as an exercise.

2.6. **Quasi-triangular matrices.** Suppose the matrix A has lots of zeroes, say:

$$A = \begin{bmatrix} a_{11} & \ldots & a_{1k} & 0 & \ldots & 0 \\ \vdots & & \vdots & & & \\ a_{k1} & \ldots & a_{kk} & 0 & \ldots & 0 \\ a_{k+1,1} & \ldots & a_{k+1,k} & a_{k+1,k+1} & \ldots & a_{k+1,n} \\ \vdots & & & & & \\ a_{n1} & \ldots & a_{nk} & a_{n,k+1} & \ldots & a_{nn} \end{bmatrix} = \begin{bmatrix} B & | & 0 \\ - & - & - \\ C & | & D \end{bmatrix}$$

where B is a $k \times k$ matrix and D is $(n-k) \times (n-k)$. Show the nice formula

$$\det A = \det B \, \det D$$

2.7. **Digression: the fundamental theorem of algebra.**

2.7.1. *Polynomials of degree two: roots and factorization.* Consider polynomials of degree two, with real coefficients: $p(x) = ax^2 + bx + c$. It is well known that $p(x)$ has real solutions $x_{1,2} = \dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ if $b^2 - 4ac \geq 0$ (where $x_1 = x_2$ when $b^2 - 4ac = 0$), and $p(x)$ has no real solutions if $b^2 - 4ac < 0$.

When the solutions are real, then the polynomial factors as

$$ax^2 + bx + c = a(x - x_1)(x - x_2)$$

In particular, if $x_1 = x_2$ then $p(x) = a(x - x_1)^2$ and x_1 is called a *double root*; x_1 is said to have *multiplicity two*. It is convenient to say that also in this case $p(x)$ has two roots.

If, on the other hand, if $p(x)$ has no real roots, then p cannot be factored within real numbers, and it is called *irreducible* (over the real numbers).

2.7.2. *Complex numbers and factorization of polynomials of degree two.* If $p(x) = ax^2 + bx + c$ is irreducible this means that $b^2 - 4ac < 0$ and we cannot take the square root of this quantity in order to calculate the two roots of $p(x)$. However, writing $b^2 - 4ac = (-1)\,(-b^2 + 4ac)$ and introducing the symbol i for $\sqrt{-1}$ we can write the zeroes of $p(x)$ as

$$x_{1,2} = \frac{-b \pm i\sqrt{-b^2 + 4ac}}{2a} = \frac{-b}{2a} \pm i\frac{\sqrt{-b^2 + 4ac}}{2a} \in \mathbb{R} + i\mathbb{R} = \mathbb{C}$$

Considering the two roots x_1, x_2 complex, we can still factor $ax^2 + bx + c = a(x - x_1)(x - x_2)$, only now the factors have complex coefficients. Within complex numbers every polynomial of degree two is reducible!

Note that the two roots of a quadratic polynomial with real coefficients are complex conjugate: if $a, b, c \in \mathbb{R}$ and $x_{1,2} \notin \mathbb{R}$ then $x_2 = \overline{x_1}$.

2.8. **The fundamental theorem of algebra.** It is absolutely remarkable that *any polynomial* can be completely factored using complex numbers:

Theorem 5. The fundamental theorem of algebra

Any polynomial $p(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0$ *with coefficients* $a_j \in \mathbb{C}$ *can be factored*

$$a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0 = a_n(x - x_1)(x - x_2)\ldots(x - x_n) \tag{15}$$

for a unique set of complex numbers $x_1, x_2, \ldots, x_n$ *(not necessarily distinct), called the roots of the polynomial* $p(x)$.

Remark. With probability one, the zeroes $x_1, \ldots, x_n$ of polynomials $p(x)$ are distinct. Indeed, if x_1 is a double root (or has higher multiplicity) then both relations $p(x_1) = 0$ and $p'(x_1) = 0$ must hold. This means that there is a relation between the coefficients $a_0, \ldots a_n$ of $p(x)$ (the multiplet $(a_0, \ldots a_n)$ belongs to an n dimensional surface in $\mathbb{C}^{n+1}$).

2.8.1. *Factorization within real numbers.* If we want to restrict ourselves only within real numbers then we can factor any polynomial into factors of degree one or two:

Theorem 6. Factorization within real numbers

Any polynomial of degree n *with real coefficients can be factored into factors of degree one or two with real coefficients.*

Theorem 6 is an easy consequence of the (deep) Theorem 5. Indeed, first, factor the polynomial in complex numbers (15). Then note that the zeroes $x_1, x_2, \ldots, x_n$ come in pairs of complex conjugate numbers, since if z satisfies $p(z) = 0$, then also its complex conjugate $\overline{z}$ satisfies $p(\overline{z}) = 0$. Then each pair of factors $(x - z)(x - \overline{z})$ must be replaced in (15) by its expanded value:

$$(x - z)(x - \overline{z}) = x^2 - (z + \overline{z})\,x + |z|^2$$

which is an irreducible polynomial of degree 2, with real coefficients. $\Box$

2.9. **The Vandermonde determinant.** Show that

$$V_n(a_1, a_2, \ldots, a_n) = \begin{vmatrix} 1 & 1 & \ldots & 1 \\ a_1 & a_2 & \ldots & a_n \\ a_1^2 & a_2^2 & \ldots & a_n^2 \\ \vdots & & & \\ a_1^{n-1} & a_2^{n-1} & \ldots & a_n^{n-1} \end{vmatrix} = \prod_{i<j} (a_j - a_i)$$

To see that, first replace a_n by an unknown x:

$$V_n(a_1, a_2, \ldots, x) = \begin{vmatrix} 1 & 1 & \ldots & 1 \\ a_1 & a_2 & \ldots & x \\ a_1^2 & a_2^2 & \ldots & x \\ \vdots & & & \\ a_1^{n-1} & a_2^{n-1} & \ldots & x^{n-1} \end{vmatrix}$$

Expanding $V_n(a_1, a_2, \ldots, x)$ along the last column we see that $V_n(a_1, a_2, \ldots, x)$ is a polynomial in x of degree (at most) $n-1$, with the leading coefficient equal to $V_{n-1}(a_1, a_2, \ldots, a_{n-1})$. For $x = a_1$ the determinant vanishes, and also for $x = a_2, \ \ldots, x = a_{n-1}$. Therefore $V_n(a_1, a_2, \ldots, x) = C_n(x - a_1)(x - a_2)\ldots(x - a_{n-1})$ where C_n is the coefficient of x^{n-1}. We noted that $C_n = V_{n-1}(a_1, a_2, \ldots, a_{n-1})$, therefore

$$V_n(a_1, a_2, \ldots, x) = C_n(x - a_1)(x - a_2)\ldots(x - a_{n-1})\, V_{n-1}(a_1, a_2, \ldots, a_{n-1})$$

Induction on n finishes the proof.