
VECTORS AND MATRICES

Preliminary Notes

Turbo Huang
November 2024
For Cambridge University

Contents

1	Complex numbers	3
1.1	Complex exponentiation	4
1.2	Roots of unity	5
1.3	Complex logarithms	6
1.4	Lines and circles	6
1.4.1	Complex equation of a line	6
1.4.2	Complex equation of a circle	6
2	Vectors	7
2.1	Definitions and properties	7
2.2	Scalar product	8
2.3	Cauchy-Schwarz Inequality	9
2.4	Vector product	10

1 Complex numbers

And so we begin! I suppose this is a better opportunity than any for me to share a profound foreword to the erudite, learned journey of mathematical enlightenment we are about to embark on with linear algebra - namely, to answer the question: why learn linear algebra at all? There are only three reasons. First, you're a massive sci-fi movie nerd and have accidentally stumbled upon the far-inferior version of the Matrix. Second, you're a massive basketball fan and received the worst surprise of your entire life when you searched "Jordan form" on YouTube in hopes of basketball enlightenment. Third, you're a massive German and have taken linear algebra for the sole purpose of pronouncing *eigenvalue* "the German way". Either or, I'm glad you're here with me; after all, if you're destined to become a machine learning dev earning seven figures and sunbathing in a luxury yacht, then - in the eternal words of wisdom of r/animememes - "don't say you love the anime if you haven't read the manga".

(Of course, besides all these completely unhinged things I'm talking about, I suppose there's also a few nuggets of mathematical beauty to be found in these curious morsels we call vectors and matrices here and there.)

Let's start with a return to form: complex numbers. In the realm of linear algebra specifically, complex numbers are important for two reasons. First, in the set of complex numbers \mathbb{C} , we can guarantee that a polynomial of degree n will have n roots by the Fundamental Theorem of Algebra; never again will we have to worry about equations like $\lambda^2 + 1 = 0$ making us more confused than tasting a burger from Pizza Hut and finding it delicious. This becomes particularly relevant when we have to deal with these polynomials, which arise when we find the eigenvalues of a particular matrix - more on that later.

Definition 1.1. (Complex number). We define the imaginary unit i as satisfying $i^2 = -1$; as such, we also define the set of complex numbers \mathbb{C} to encompass all numbers of the form

$$z = a + bi$$

where a and b are real. We write $a = \operatorname{Re}(z)$, $b = \operatorname{Im}(z)$, and the complex conjugate $\bar{z} = a - bi$ (a theorem in algebra will demonstrate that if z is a root of a polynomial, then \bar{z} is too).

But second of all - and much more thematically - while real numbers are *one-dimensional*, all lying upon the same infinitely long number line, complex numbers are two-dimensional; it is useful to think of $z = a + bi$ as a vector in the *complex plane*, $\begin{bmatrix} a \\ b \end{bmatrix}$. The representation of complex numbers as vectors is done on an *Argand plane*, analogous to the Cartesian plane but with the x-axis representing the real part of z and the y-axis representing the imaginary part.

What follows is a carousel of important results for complex numbers which are truly astounding in their mind-numbingness, not because of what they are but because we've seen them all before:

Definition 1.2. (Modulus and argument). Define the modulus of $z = a + bi$ as $|z| = \sqrt{a^2 + b^2}$; this is analogous to the length of its vector representation on the Argand plane. Define its argument as the angle its vector makes with the real axis: $\arg z = \tan^{-1}(\frac{b}{a})$. The modulus-argument pair (r, θ) can uniquely describe a complex number z , but each z has infinitely many arguments $\theta + 2k\pi$ (a full revolution, but not the French kind). We often take only the principal argument - $-\pi < \theta < \pi$.

Proposition 1.3. We have

$$z\bar{z} = a^2 + b^2 = |z|^2$$

and

$$z^{-1} = \frac{\bar{z}}{|z|^2}$$

Theorem 1.4. (Triangle inequality). For any two complex numbers z_1 and z_2 , we have

$$|z_1 + z_2| \leq |z_1| + |z_2|$$

which can be shown by the geometrical interpretation of the two complex numbers as vectors representing sides of a triangle.

1.1 Complex exponentiation

To extend exponentiation to complex numbers, we use the Taylor series definition of exponentiation:

Definition 1.5. (Exponential function). Define

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

which can be verified to satisfy the properties we expect from the exponential function, including $e^a e^b = e^{a+b}$. We assume that this sum converges for all complex numbers z .

Similarly, we would like to extend the trigonometric functions to the complex realm, where a geometric definition fails due to the budding, unhinged insanity that underlines the words "an angle of $39 + 46\pi i$ degrees":

Definition 1.6. (Complex sine and cosine). Define

$$\sin z = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!} = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \dots$$

and

$$\cos z = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!} = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \dots$$

From the two above results, we obtain a very important formula throughout all of math.

Theorem 1.7. (Euler's formula).

$$e^{iz} = \cos z + i \sin z$$

It almost feels like I ought to be wearing a suit and tie before I even dare to think about these symbols. We also note that any complex number can thus be written in terms of a complex exponential, as its modulus-argument form (r, θ) suggests it can be written as

$$z = r(\cos \theta + i \sin \theta) = re^{i\theta}$$

which allows us to state that multiplication between two complex numbers $z_1 = r_1 e^{i\theta_1}$ and $z_2 = r_2 e^{i\theta_2}$ requires the multiplication of their moduli and addition of their arguments.

1.2 Roots of unity

Definition 1.8. (Roots of unity). We refer to the complex roots of the equation $\omega^n = 1$ as the *n*th roots of unity; as this is a polynomial of deg *n*, we have *n* roots of unity, which can be completely described by

$$\omega = e^{\frac{2\pi ki}{n}}, \quad k = 0, 1, 2, 3, \dots, (n-1).$$

As a consequence of the above, we also have

$$\sum \omega = 1 + e^{\frac{2\pi i}{n}} + e^{\frac{4\pi i}{n}} + \dots + e^{\frac{2(n-1)\pi i}{n}} = 0$$

(You may have noticed that we've reached a critical mass of handwaving away statements without proof in this section. De Moivre is surely spinning in his grave. The reason why is because I can't be bothered to prove any of this stuff, so the proofs are left as an exercise to the reader.)

1.3 Complex logarithms

Definition 1.9. (Complex logarithms). Define the complex logarithm $\omega = \ln z$ as the number which satisfies $e^\omega = z$. If z is a complex number $z = re^{i\theta}$, then we have $e^\omega = re^{i\theta}$ and thus $\ln \frac{1}{r} = i\theta - \omega$ and $\omega = \ln re^{i\theta} = i\theta + \ln r$. (I'm just now realizing we could've got here with $\ln ab = \ln a + \ln b$.)

Definition 1.10. (Complex powers). Define z^α for complex z as $e^{\alpha \ln z}$ where we insist that the argument used for z is $-\pi < \theta < \pi$, the principal argument.

Definition 1.11. (De Moivre's Theorem).

$$\cos n\theta + i \sin n\theta = (\cos \theta + i \sin \theta)^n.$$

This can be proven by induction; it is functionally identical to stating that $e^{ni\theta} = (e^{i\theta})^n$, which is obvious over the reals but not so obvious over complex numbers.

1.4 Lines and circles

1.4.1 Complex equation of a line

The equation is not called "complex equation of a line" because it involves complex numbers, but because it is unnecessarily complex. Observe. What can we do if we want to find a line that goes through the point x_0 on the Argand diagram and is parallel to some complex number ω ? By what we know of vector equations for lines, we can write the line as $x = x_0 + \lambda\omega$ for some real scalar λ . If we rewrite this as $\frac{x-x_0}{\omega} = \lambda$ and take the conjugate of both sides, we obtain

$$\frac{\bar{x} - \bar{x}_0}{\bar{\omega}} = \bar{\lambda} = \lambda$$

as λ is real. Thus the conjugate of this expression is equal to itself:

$$\frac{\bar{x} - \bar{x}_0}{\bar{\omega}} = \frac{x - x_0}{\omega}$$

which gives us the equation of a line parallel to ω and passing through x_0 .

1.4.2 Complex equation of a circle

A circle with center x_0 and radius r , in abstract terms, is simply a locus of points a distance r away from the center x_0 . (In less abstract terms, it can be referred to as "half a Venn diagram" or "the inferior donut".) This can be formulated as

$$|x - x_0| = r$$

or, squaring both sides,

$$\begin{aligned} |x - x_0|^2 &= r^2 \\ (x - x_0)(\bar{x} - \bar{x}_0) &= r^2 \end{aligned}$$

2 Vectors

2.1 Definitions and properties

Over the myriad centuries of our blessed lifetimes, we've known vectors as many things: they can be really tall numbers, really squiggly arrows, the only place where a mathematician knows the meaning of the word "bold", and, above all else, the orange jumpsuit guy from Despicable Me who embodies the biological fact that rhombus-shaped human bodies are peak evolution. But now comes a time where we study vectors not as actual *things* - arrays of numbers, directions, arrows - but as abstract math objects, just like what numbers are to us.

Definition 2.1. (Vector space). A *vector space* V over \mathbb{R} or \mathbb{C} is a collection of entities ("vectors") $\mathbf{v} \in V$, in which we define two operations: addition of two vectors and scalar multiplication with a vector.

These "vectors" (notice that I am stubbornly using quotation marks here, because technically, we aren't supposed to know what vectors are yet; we're defining them here) must satisfy the following properties under addition and scalar multiplication:

Vector addition

1. Commutativity: $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$
2. Associativity (order does not change result): $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$
3. Identity: there is a vector $\mathbf{0}$ such that $\mathbf{a} + \mathbf{0} = \mathbf{a}$.
4. Existence of inverses: for all vectors \mathbf{a} , there exists another vector in V which we denote $-\mathbf{a}$ such that $\mathbf{a} + (-\mathbf{a}) = \mathbf{0}$.

Scalar multiplication

Here λ, μ denote scalars and all bolded letters represent vectors.

1. Scalar distributive property: $\lambda(\mathbf{a} + \mathbf{b}) = \lambda\mathbf{a} + \lambda\mathbf{b}$.
2. Vector distributive property: $\mathbf{a}(\lambda + \mu) = \lambda\mathbf{a} + \mu\mathbf{a}$
3. Associativity: $\lambda(\mu\mathbf{a}) = \mu(\lambda\mathbf{a})$
4. Identity: $1\mathbf{a} = \mathbf{a}$.

It is implicitly stated that the sum of two "vectors" in V (suspicious name, that) will be another "vector" (also in V), and the scalar multiplication of a vector will give another vector; thus we can say that vector spaces are *closed under addition and scalar multiplication*; that is, these two operations map elements of V to V .

Notice also that vectors (as we understand them; the matrix kind) of any size will all satisfy these properties, be it that they have two or three or ten entries; in fact, vectors with a single entry - which are just scalars - also satisfy these properties. Thus, any \mathbb{R}^n - the

set of n -tuples of real numbers - are vector spaces; all lines through the origins are vector spaces (as in, all points on these lines), but lines not through the origin are not - they do not contain $(0,0)$, the zero vector.

The geometric meaning of vectors, of course, are more familiar to us: objects \mathbf{v} with a length denoted $|\mathbf{v}|$, and a direction which remains unchanged under scalar multiplication (either parallel or antiparallel).

Definition 2.2. (Unit vector). A unit vector, usually denoted with a little hat like this - $\hat{\mathbf{v}}$ - is a vector with length 1.

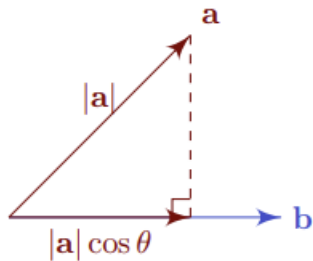
2.2 Scalar product

In *addition* (ba dum-tss) to addition and multiplication, we define a third operation in a vector space V - the scalar product - that takes two vectors \mathbf{a}, \mathbf{b} as inputs and returns a scalar.

Thus far, the one instance of a scalar product we are familiar with is the dot product; but the dot product is not the only scalar product that can be defined. We know that the dot product in $\mathbb{R}^{\mathbb{F}}$ and $\mathbb{R}^{\mathbb{F}}$ between vectors \mathbf{a} and \mathbf{b} is defined $\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos \theta$ where θ is the angle between the two, and satisfies:

1. $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$
2. $\mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|^2 \geq 0$
3. $\mathbf{a} \cdot \mathbf{a} = 0 \iff \mathbf{a} = \mathbf{0}$
4. If $\mathbf{a} \cdot \mathbf{b} = 0$, then \mathbf{a}, \mathbf{b} are perpendicular.

From the definition $\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos \theta$, we know that the dot product gives us information on the angle between two vectors; indeed, from the definition of cosine in a triangle, we know that this expression represents the component of \mathbf{a} parallel to \mathbf{b} , also known as the *projection*: However, this is only one of many possible scalar products, or *inner products*



$\langle \mathbf{x} | \mathbf{y} \rangle$, that can be defined. Similar to a dot product, they are defined by the following properties:

1. Commutativity: $\langle \mathbf{x} | \mathbf{y} \rangle = \langle \mathbf{y} | \mathbf{x} \rangle$
2. Distributive and associative properties: $\langle \mathbf{x} | (\lambda \mathbf{y} + \mu \mathbf{z}) \rangle = \lambda \langle \mathbf{x} | \mathbf{y} \rangle + \mu \langle \mathbf{x} | \mathbf{z} \rangle$
3. Identity: if $\langle \mathbf{x} | \mathbf{x} \rangle = 0$, then $\mathbf{x} = 0$; $\langle \mathbf{x} | \mathbf{x} \rangle \geq 0$.

Accordingly, we define the *norm* of a vector \mathbf{x} as follows:

Definition 2.3. (Norm). The *norm* of a vector \mathbf{x} is defined as

$$|\mathbf{x}| = \sqrt{\langle \mathbf{x} | \mathbf{x} \rangle} \quad (1)$$

The above definitions encompass far more than the dot product; for instance, if we define the vector space V as the set of all real integrable functions $f(x)$, then we can verify that the operation

$$\langle f(x) | g(x) \rangle = \int_0^1 f(x)g(x) dx \quad (2)$$

is a valid inner product. This generality will be useful in establishing the power of the following result.

2.3 Cauchy-Schwarz Inequality

Theorem 2.4. (Cauchy-Schwarz inequality). For all vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{\kappa}$, we have

$$|\langle \mathbf{x} | \mathbf{y} \rangle| \leq |\mathbf{x}| |\mathbf{y}|. \quad (3)$$

In other words, the norm of the inner product is greater than or equal to the product of the norms. This applies to any inner product on any real vector space, as the proof below will demonstrate.

Proof. Consider the norm $|\langle \mathbf{x} - \lambda \mathbf{y} | \mathbf{x} - \lambda \mathbf{y} \rangle|$, which is non-negative by definition. Thus its square

$$(\mathbf{x} - \lambda \mathbf{y}) \cdot (\mathbf{x} - \lambda \mathbf{y})$$

is also non-negative, leading to

$$|\mathbf{x}|^2 - 2\lambda \mathbf{x} \cdot \mathbf{y} + \lambda^2 |\mathbf{y}|^2 \geq 0$$

where \cdot denotes any inner product. Reorganizing this as a quadratic equation in λ and recognizing that it has at least one real solution, we have its discriminant

$$\Delta = b^2 - 4ac = (4(\mathbf{x} \cdot \mathbf{y})^2) - 4|\mathbf{x}|^2 |\mathbf{y}|^2 \geq 0$$

leading to the desired result. □

We derived this proof entirely based on the axiomatic properties of inner products instead of a specific inner product (e.g. the dot product), so this inequality is applicable across any inner product in any real vector space - like the integral-based one defined above.

2.4 Vector product

Similar to the scalar product, we can also define a binary operation known as the *vector product* over a vector space that takes two vectors and outputs another vector,