

Lecture XXXIV: §4.6 Complex numbers & complex eigenvalues

TODAY: want to discuss complex numbers & its use in the EV Problem

3. Motivation: Why complex numbers?

There are 2 motivations to define complex numbers

① Algebraic: We want to find a new set of numbers, enlarging \mathbb{R} so that every polynomial in one variable has a root (Example: $x^2 + 1$ has no roots in \mathbb{R})

② Geometric: Want to define some sort of multiplication in \mathbb{R}^2

This would be an operation $\mathbb{R}^2 \times \mathbb{R}^2 \xrightarrow{\varphi} \mathbb{R}^2$
 $(v, w) \longmapsto \varphi(v, w) = \text{a vector in } \mathbb{R}^2$

Write $\varphi(\vec{v}, \vec{w}) = \vec{v} \odot \vec{w}$ (not usual dot product because that gives back a number, not a vector!)

that satisfy the same properties that multiplication in \mathbb{R} has:

(1) Commutative: $\varphi(\vec{v}, \vec{w}) = \varphi(\vec{w}, \vec{v})$
 $v \overset{\parallel}{\odot} \vec{w} \quad \vec{w} \overset{\parallel}{\odot} v$

(2) Distributive: $\varphi(\vec{v}, \vec{w} + \vec{u}) = \varphi(\vec{v}, \vec{w}) + \varphi(\vec{v}, \vec{u})$
 $\vec{v} \overset{\parallel}{\odot} (\vec{w} + \vec{u}) \quad \vec{v} \overset{\parallel}{\odot} \vec{w} + \vec{v} \overset{\parallel}{\odot} \vec{u}$

(3) Must have a neutral element: a vector \vec{e} in \mathbb{R}^2 with
 $\vec{v} \odot \vec{e} = \vec{e} \odot \vec{v} = \vec{v}$ for all \vec{v} in \mathbb{R}^2

(4) Last thing we need: We want to be able to "invert" vectors $\vec{v} \neq \vec{0}$.

So given any $\vec{v} \neq \vec{0}$, we want to find \vec{w} so that $\vec{v} \odot \vec{w} = \vec{e}$

We want to think of \mathbb{R}^2 as a set of "numbers" with operations
 $+$ (coordinatewise) & \odot extending those in \mathbb{R} . For this to

To work, we need to view \mathbb{R} inside \mathbb{R}^2 . We do so by identifying a in \mathbb{R} with $\begin{bmatrix} a \\ 0 \end{bmatrix}$ in \mathbb{R}^2 . (Formally: we have an injective map $\mathbb{R} \hookrightarrow \mathbb{R}^2$ $a \mapsto \begin{bmatrix} a \\ 0 \end{bmatrix}$)

• This identification is compatible with $+$ on both sides because

$$\begin{bmatrix} a+b \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ 0 \end{bmatrix} + \begin{bmatrix} b \\ 0 \end{bmatrix} \iff a+b \text{ in } \mathbb{R}$$

• Q: What about \cdot in \mathbb{R} vs \odot in \mathbb{R}^2 ?

A: $(a \cdot b)$ corresponds to $\begin{bmatrix} a \cdot b \\ 0 \end{bmatrix}$ so this should be $\begin{bmatrix} a \\ 0 \end{bmatrix} \odot \begin{bmatrix} b \\ 0 \end{bmatrix}$

Claim: This gives a hint to what the neutral element should be.

Indeed, since 1 is neutral element for multiplication in \mathbb{R} , we are forced to have $e = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ (Because neutral element will be unique!)

[Indeed, if we had 2 neutral elements, called e & e' then

$$e = e \odot e' = e' \quad \left[\begin{array}{l} \downarrow \text{think } e' \text{ as neutral} \\ \downarrow \text{think of } e \text{ as neutral} \end{array} \right]$$

In summary: $\vec{v} \odot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \vec{v}$ must be true for all \vec{v} .

• Also: since all $a \neq 0$ in \mathbb{R} are invertible, this would guarantee that all $\begin{bmatrix} a \\ 0 \end{bmatrix} \neq \vec{0}$ are invertible, with $\begin{bmatrix} a \\ 0 \end{bmatrix}^{-1} = \begin{bmatrix} 1/a \\ 0 \end{bmatrix}$

This explains why we wanted property (4) to begin with.

Final remark: The distributive property (2) reduces our search to 2 cases

We need to define $\begin{bmatrix} a \\ 0 \end{bmatrix} \odot \begin{bmatrix} 0 \\ b \end{bmatrix}$ & $\begin{bmatrix} 0 \\ c \end{bmatrix} \odot \begin{bmatrix} 0 \\ b \end{bmatrix}$

Why? $\begin{bmatrix} a \\ b \end{bmatrix} \odot \begin{bmatrix} c \\ d \end{bmatrix} = (\begin{bmatrix} a \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ b \end{bmatrix}) \odot (\begin{bmatrix} c \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ d \end{bmatrix})$
 $\stackrel{\text{distribute}}{=} \begin{bmatrix} a \\ 0 \end{bmatrix} \odot \begin{bmatrix} c \\ 0 \end{bmatrix} + \begin{bmatrix} a \\ 0 \end{bmatrix} \odot \begin{bmatrix} 0 \\ d \end{bmatrix} + \begin{bmatrix} 0 \\ b \end{bmatrix} \odot \begin{bmatrix} c \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ b \end{bmatrix} \odot \begin{bmatrix} 0 \\ d \end{bmatrix}$

Using the Distributive & Commutative Properties, we can rewrite this as

$$\begin{bmatrix} a \\ 0 \end{bmatrix} \odot \begin{bmatrix} c \\ 0 \end{bmatrix} + \underbrace{\begin{bmatrix} a \\ 0 \end{bmatrix} \odot \begin{bmatrix} 0 \\ d \end{bmatrix}}_{\text{NEED TO DEFINE}} + \underbrace{\begin{bmatrix} 0 \\ b \end{bmatrix} \odot \begin{bmatrix} c \\ 0 \end{bmatrix}}_{\text{NEED TO DEFINE THIS}} + \underbrace{\begin{bmatrix} 0 \\ b \end{bmatrix} \odot \begin{bmatrix} 0 \\ d \end{bmatrix}}_{\text{NEED TO DEFINE THIS}}$$

$\begin{bmatrix} a \\ 0 \end{bmatrix} \odot \begin{bmatrix} c \\ 0 \end{bmatrix}$ (from \mathbb{R})

Remember: We need to be able to solve $x^2 + 1 = 0$ in \mathbb{R}^2 . This means we must have $\begin{bmatrix} a \\ b \end{bmatrix}$ with $\begin{bmatrix} a \\ b \end{bmatrix} \odot \begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

§2. Definition of Complex Numbers:

Def: A complex number z is given by a vector of real numbers $\begin{bmatrix} a \\ b \end{bmatrix}$ written as $z = a + ib$ (i is a placeholder)

• Call \mathbb{C} = set of complex numbers

• Names: a = "Real Part of z " = $\text{Re}(z)$

b = "Imaginary Part of z " = $\text{Im}(z)$

• $a + ib = c + id$ means $a = c$ & $b = d$ (Re & Im parts agree)

Furthermore: $i = 0 + i1$ satisfies $i^2 = -1$ (so it's a soln of $x^2 + 1 = 0$!)

• a in \mathbb{R} is a complex number via $a = a + i \cdot 0$. (as the picture shows!)

Q: How do we add & multiply complex numbers?

① Addition: is done componentwise (that is, as we do in \mathbb{R}^2)

$$(a + ib) + (c + id) = (a + c) + i(b + d)$$

② Multiplication: is defined as follows:

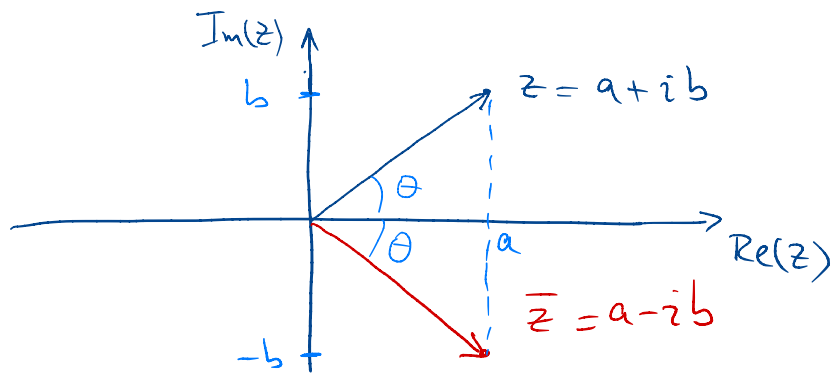
$$\begin{matrix} z = a + ib \\ w = c + id \end{matrix} \implies \boxed{zw = (ac - bd) + i(ad + bc)}$$

By design, it satisfies three properties we wanted

(this uses that $i^2 = -1$)

- (1) commutative
- (2) distributive
- (3) $1 = 1 + i \cdot 0$ is neutral element.

Visualization?



mirror image
about Real axis

Properties of complex conjugation:

$$(1) \overline{z+w} = \overline{z} + \overline{w}$$

$$(2) \overline{z \cdot w} = \overline{z} \cdot \overline{w}$$

$$(3) z \cdot \overline{z} = |z|^2$$

Why are they valid? Just follow the definitions!

$$(1) \overline{z+w} = \overline{(a+ib) + (c+id)} = \overline{(a+c) + i(b+d)} = a+c - i(b+d) = (a-ib) + (c-id) = \overline{z} + \overline{w}$$

$$(2) \overline{z \cdot w} = \overline{(a+ib)(c+id)} = \overline{(ac-bd) + i(ad+bc)} = ac-bd - i(ad+bc)$$

$$\overline{z} \cdot \overline{w} = (a-ib) \cdot (c-id) = (ac-bd) + i(-ad-bc)$$

$$(3) z \cdot \overline{z} = (a+ib) \overline{(a+ib)} = (a+ib)(a-ib) = (a^2+b^2) + i(ab+ba) = a^2+b^2 = |z|^2$$

Consequence: If $z \neq 0$, we have either $\text{Re}(z) \neq 0$ or $\text{Im}(z) \neq 0$.

In both cases, we get $|z|^2 = a^2 + b^2 > 0$,

Since $z \cdot \overline{z} = |z|^2$, we divide by $|z|^2$ on both sides &

regroup to get $z \left(\frac{\overline{z}}{|z|^2} \right) = 1$

Conclusion: $z^{-1} = \frac{\overline{z}}{|z|^2}$ whenever $z \neq 0$

Application: Use this to write any quotient $\frac{a+ib}{c+id}$ as a complex number.

How? $\frac{a+ib}{c+id} = \frac{a+ib}{c+id} \frac{c-id}{c-id} = \frac{(a+ib)(c-id)}{c^2+d^2} = \frac{ac-bd}{c^2+d^2} + i \frac{bc-ad}{c^2+d^2}$

↳ multiply & divide by the complex conjugate of the denominator.

Examples: ① $z = 1+i \rightsquigarrow \bar{z} = 1-i$, $|z|^2 = 1^2 + 1^2 = 2 > 0$ 6

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

② $\frac{z+i}{1+i} = \frac{z+i}{1+i} \frac{1-i}{1-i} = \frac{(z+i)(1-i)}{1^2+1^2} = \frac{(z+1)+i(-z+1)}{2} = \frac{z}{2} - \frac{i}{2}$

③ $z = 1-i3 \rightsquigarrow \bar{z} = 1+i3$

$w = 2+i4 \rightsquigarrow \bar{w} = 2-i4$

$z+w = 3+i$, $z\bar{w} = (1-i3)(2+i4) = (2+4\cdot 3) + i(4-3\cdot 2) = 14 - i2$

$\overline{z+w} = 3-i$

$\bar{z} + \bar{w} = 1+i3 + 2-i4 = 3-i$

} we expected this from Property (1) on page 5

$\bar{z} \cdot \bar{w} = (1+i3)(2-i4) = (2+12) + i(-4+3\cdot 2) = 14 + 2i$

$\overline{z\bar{w}} = \overline{14 - i2} = 14 + i2$

→ These 2 agree, as expected from Property (2) on page 5

§4 Roots of polynomials in $\mathbb{R}[x]$ & $\mathbb{C}[x]$:

• Main reason for defining \mathbb{C} is the following Theorem:

• Fundamental Theorem of Algebra: Every polynomial of degree ≥ 1 in 1 variable with complex coefficients has a root in \mathbb{C} .

("Short-hand definition: \mathbb{C} is algebraically closed ")

Example: (Quadratic polynomials)

$P(x) = ax^2 + bx + c$ with a, b, c in \mathbb{C} $a \neq 0$

Roots: $\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ in \mathbb{C}

What does $\sqrt{\quad}$ mean for \mathbb{C} ?

We need to find z in \mathbb{C} with $z^2 = b^2 - 4ac$

We can always do this!

Example: say $b^2 - 4ac = -16$ then $z = 4i$ satisfies $z^2 = 16i^2 = -16$

In general, say we want to solve $z^2 = w$ for w in \mathbb{C} .

then $|z^2| = |z|^2 = |w|$ so $|z| = \sqrt{|w|}$
KEY PROP page 4

$\frac{w}{|w|}$ has modulus 1 so it lies in the circle $x^2 + y^2 = 1$
where $x = \operatorname{Re}(w)$, $y = \operatorname{Im}(w)$

Since the unit circle has parameterization $x = \sin t$ for $0 \leq t \leq 2\pi$
 $y = \cos t$

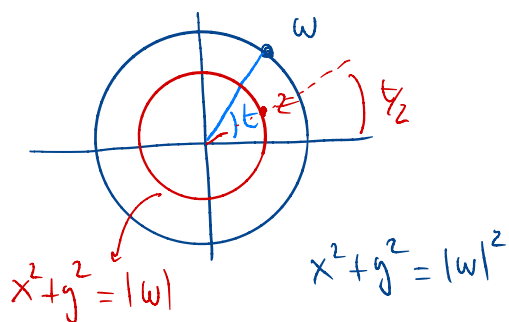
then we write $\frac{w}{|w|} = \sin t + i \cos t \Rightarrow w = |w| (\sin t + i \cos t)$

Proposal $z = \sqrt{|w|} \left(\sin \frac{t}{2} + i \cos \frac{t}{2} \right)$ satisfies

$$z^2 = |w| \left(\sin \frac{t}{2} + i \cos \frac{t}{2} \right)^2 = |w| (\sin t + i \cos t) = w \quad \checkmark$$

trig formulas \Rightarrow $\sin(2\alpha) = 2\sin(\alpha)\cos(\alpha)$
 $\cos(2\alpha) = \cos^2(\alpha) - \sin^2(\alpha)$
 $\Rightarrow \alpha = \frac{t}{2}$

Visually:



Q: What abouts polynomials in $\mathbb{R}[x]$?

A: Two types of roots:
(1) Real roots
(2) complex, non-real roots: they come in pairs (z & \bar{z}) (conjugate pairs)

Why? If $f(z)=0$ & f is in $\mathbb{R}[z]$, write

$$f(x) = a_0 + a_1x + \dots + a_nx^n \quad a_0, \dots, a_n \text{ in } \mathbb{R}$$

Then $f(z) = a_0 + a_1z + \dots + a_nz^n = 0$

Take complex conjugate to get:

$$0 = \overline{0} = \overline{f(z)} = \overline{a_0 + a_1z + \dots + a_nz^n} = \overline{a_0} + \overline{a_1}\overline{z} + \dots + \overline{a_n}(\overline{z})^n$$
$$= a_0 + a_1\overline{z} + \dots + a_n(\overline{z})^n$$

So $f(\overline{z}) = 0$, meaning \overline{z} is also a root of $f(x)$, just as z .
 a_i 's are real numbers

Example: $f(x) = x^3 - x^2 + x - 1 = (x-1)(x^2+1) = (x-1)(x-i)(x+i)$

- 1 real root $z=1$
- 2 complex conjugate roots: i & $-i = \overline{i}$.

Application: If $P_A(t)$ is the characteristic polynomial of an $n \times n$ matrix A with real entries, then the eigenvalues of A are the roots of $P_A(t)$. They are either real numbers or come in conjugate pairs.

Example: $A = \begin{bmatrix} 3 & 1 \\ -2 & 1 \end{bmatrix} \rightsquigarrow P_A(t) = \det \begin{pmatrix} 3-t & 1 \\ -2 & 1-t \end{pmatrix} = t^2 - 4t + 5$

Roots $\lambda = \frac{4 \pm \sqrt{-4}}{2} = \frac{4 \pm 2i}{2} = 2 \pm i$

→ 2 complex eigenvalues: $2+i$ & $2-i$

Next time: We'll discuss how to find eigenvectors for them. They will not be vectors in \mathbb{R}^2 , but rather in \mathbb{C}^2 .