Lecture XXXIV: §4.6 Complex numbers & complex eigenvalues

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

§1 **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 $\varphi(\vec{v},\vec{w}) = \vec{v} \odot \vec{v}$ (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: $\underset{v \odot \vec{v}}{\underset{\parallel}{\varphi(\vec{v},\vec{w})}} = \underset{\vec{w} \odot \vec{v}}{\underset{\parallel}{\varphi(\vec{w},\vec{v})}}$

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

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

(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 $\vec{v} \odot \vec{w} = e$

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

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

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

A: $(a \cdot b)$ corresponds to $\begin{bmatrix} ab \\ 0 \end{bmatrix}$ 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

$$e \underset{\substack{\downarrow \\ \text{think } e' \text{ as neutral}}}{=} e \odot e' \underset{\substack{\downarrow \\ \text{think of } e \text{ as neutral}}}{=} e' \quad]$$

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? $\begin{bmatrix} a \\ b \end{bmatrix} \odot \begin{bmatrix} c \\ d \end{bmatrix} = \left(\begin{bmatrix} a \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ b \end{bmatrix} \right) \odot \left(\begin{bmatrix} c \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ d \end{bmatrix} \right)$

$$\overset{\text{Distribute}}{=} \begin{bmatrix} a \\ 0 \end{bmatrix} \odot \begin{bmatrix} c \\ 0 \end{bmatrix} + \begin{bmatrix} a \\ 0 \end{bmatrix} \odot \begin{bmatrix} 0 \\ d \end{bmatrix} + \begin{bmatrix} 0 \\ b \end{bmatrix} \odot \begin{bmatrix} c \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ b \end{bmatrix} \odot \begin{bmatrix} 0 \\ d \end{bmatrix}$$

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

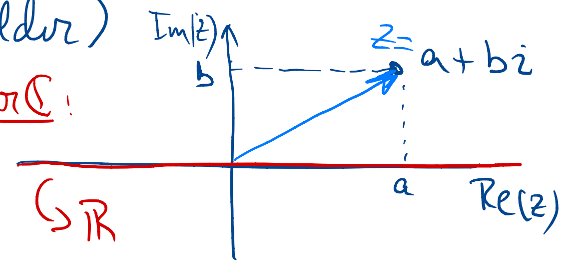
$$\underbrace{\begin{bmatrix} a\\0\end{bmatrix}\odot\begin{bmatrix} c\\0\end{bmatrix}}_{=\begin{bmatrix} ac\\0\end{bmatrix}\ (\text{from } \mathbb{R})} + \underbrace{\begin{bmatrix} a\\0\end{bmatrix}\odot\begin{bmatrix} 0\\d\end{bmatrix}}_{\text{NEED TO DEFINE}} + \underbrace{\begin{bmatrix} 0\\b\end{bmatrix}\odot\begin{bmatrix} c\\0\end{bmatrix}}_{=\begin{bmatrix} c\\0\end{bmatrix}\odot\begin{bmatrix} 0\\b\end{bmatrix}\ \text{NEED TO DEFINE THIS}} + \underbrace{\begin{bmatrix} 0\\b\end{bmatrix}\odot\begin{bmatrix} 0\\d\end{bmatrix}}_{\text{NEED TO DEFINE THIS}}$$

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 $\begin{bmatrix} a\\b\end{bmatrix}\odot\begin{bmatrix} a\\b\end{bmatrix}+\begin{bmatrix} 1\\0\end{bmatrix}=\begin{bmatrix} 0\\0\end{bmatrix}$.

§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 $z = a + i b$ (i is a placeholder)

Drawing for $\mathbb{C}$:



- Call $\mathbb{C}$ = set of complex numbers
- Names: a = "Real Part of z" = $\mathrm{Re}(z)$

  b = "Imaginary Part of z" = $\mathrm{Im}(z)$
- $a+ib = c+id$ means $a=c$ & $b=d$ (Re & Im parts agree)

Furthermore:
- $i = 0 + i\,1$ satisfies $\boxed{i^2=-1}$ (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:

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

By design, it satisfies three properties we wanted (1) Commutative (2) distributive (3) $1 = 1 + i0$ is neutral element

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

Example: $z = 1+i$, $w = 2+i3$ $\rightsquigarrow$ $zw = (1\cdot 2 - 1\cdot 3) + i(1\cdot 3 + 1\cdot 2)$
$= (2-3) + i(3+2) = -1+5i$

Q: How do we remember this definition?

A: Use distributive / commutative properties & $i^2 = -1$.

In the example:

$$\underline{(1+i)}(2+i3) = 1\cdot(2+i3) + i(2+3i)$$
$$= 2+i3 + i2 + 3i^2$$
$$= 2 + i(3+2) + 3(-1)$$
$$= (2-3) + i(3+2) = -1+i5$$

In general:

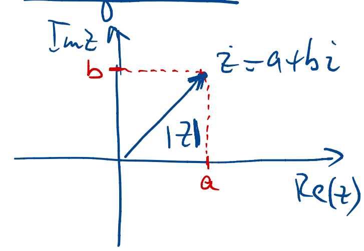
$$(a+ib)(c+id) = a(c+ib) + ib(c+id)$$
$$= ac + i(ad) + i(bc + ibd)$$
$$= ac + i(ad) + i(bc) + \underbrace{\boxed{i^2}}_{=-1}bd \underset{\substack{\uparrow \\ \text{regroup} \\ \text{real \& imaginary part}}}{=} (ac-bd) + i(ad+bc)$$

(This is why we defined multiplication as we did!)

Q: What else?

Viewing $z = a+ib$ as a vector in $\mathbb{R}^2$, we have a magnitude

$$\boxed{|z| = \sqrt{a^2+b^2} = \sqrt{Re(z)^2 + Im(z)^2}}$$



- We call it the modulus of the complex number z

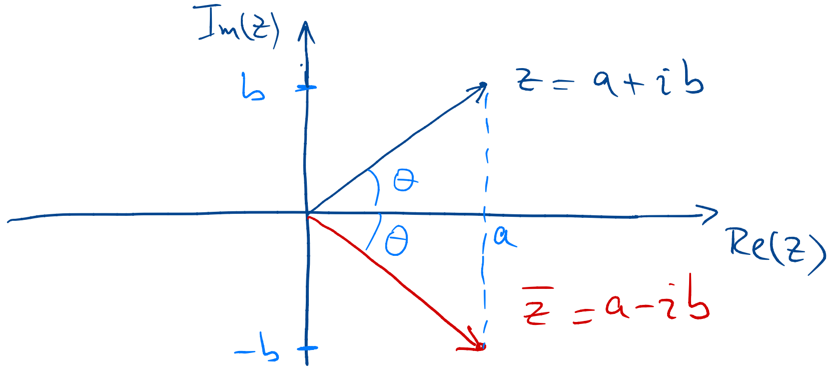
KEY PROPERTY: $|zw| = |z|\,|w|$ (Proof: write down both sides for $z = a+ib$, $w = c+id$)

§3. Next operation: Complex conjugation

Definition: Given $z = a+ib$, its complex conjugate is $\overline{z} = a - ib$

In particular: $z = \overline{z}$ if & only if $b = -b$ (so $b=0$ & z is in $\mathbb{R}$)

Visualization?



mirror image about Real axis

Properties of complex conjugation:

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

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

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

Why are they valid? Just follow the definitions!

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

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

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

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

Consequence: If $z \neq 0$, we have either $\mathrm{Re}(z) \neq 0$ or $\mathrm{Im}(z) \neq 0$.
In both cases, we get $|z|^2 = a^2+b^2 > 0$.
Since $z\cdot\bar{z} = |z|^2$, we divide by $|z|^2$ on both sides & regroup to get $z\left(\frac{\bar{z}}{|z|^2}\right) = 1$

Conclusion: $\boxed{z^{-1} = \frac{\bar{z}}{|z|^2} \quad \text{whenever } z \neq 0}$

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

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

② $\dfrac{2+i}{1+i} = \dfrac{2+i}{1+i}\,\dfrac{1-i}{1-i} = \dfrac{(2+i)(1-i)}{1^2+1^2} = \dfrac{(2+1)+i(-2+1)}{2} = \dfrac{3}{2} - \dfrac{i}{2}$.

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

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

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

$\left.\begin{array}{l} \overline{z+w} = 3-i \\ \bar{z}+\bar{w} = 1+i3+2-i4 = 3-i \end{array}\right\}$ we expected this from Property (1) on page 5

$\left\{\begin{array}{l} \bar{z}\cdot\bar{w} = (1+i3)(2-i4) = (2+12)+i(-4+3\cdot 2) = 14+2i \\ \overline{zw} = \overline{14-i2} = 14+i2 \end{array}\right.$

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

What does $\sqrt{}$ 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 $z^2 = 16 i^2$
$= -16$ ✓

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

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

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

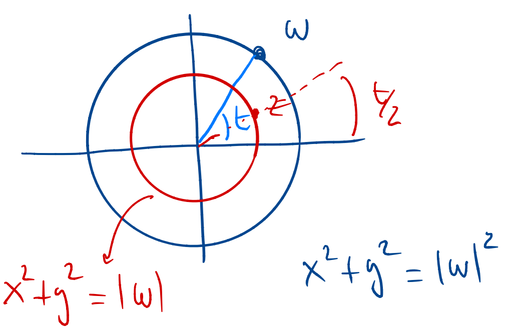
Since the unit circle has parameterization $x = \ \ t$ for $0 \le t \le 2\pi$
$y = \ \ t$

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

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

$$z^2 = |w| \left(\cos \frac{t}{2} + i \sin \frac{t}{2} \right)^2 \underset{\substack{\text{formulas for } \sin(2\alpha) = 2\sin(\alpha)\cos(\alpha) \\ \cos(2\alpha) = \cos^2(\alpha) - \sin^2(\alpha) \\ \text{for } \alpha = \frac{t}{2}}}{=} |w| (\cos t + i \sin t) = w \ ✓$$

Visually:



Q: What about 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 $\mathbb{R}[x]$, write

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n \quad a_0, \ldots, a_n \text{ in } \mathbb{R}$$

Then $\quad f(z) = a_0 + a_1 z + \cdots + a_n z^n = 0$

Take complex conjugate to get:

$$0 = \overline{0} = \overline{f(z)} = \overline{a_0 + a_1 z + \cdots + a_n z^n} = \overline{a_0} + \overline{a_1}\,\overline{z} + \cdots + \overline{a_n}(\overline{z})^n$$

$$\underset{\substack{\downarrow \\ a_i\text{'s are real numbers}}}{=} a_0 + a_1 \overline{z} + \cdots + a_n (\overline{z})^n$$

So $f(\overline{z}) = 0$, meaning $\overline{z}$ is also a root of $f(x)$, just as z.

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

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

$\rightarrow$ 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 $\mathbb{C}^2$.