Diagonalization review

Suppose A is $n \times n$. When there is a basis for $\mathbb{R}^n$ of eigenvectors of A, $\vec{v}_1, \dots, \vec{v}_n$ with corresponding eigenvalues $\lambda_1, \dots, \lambda_n$, then A is similar to a diagonal matrix:

$$A = PDP^{-1}, \quad P = [\vec{v}_1 \cdots \vec{v}_n] \text{ and } D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

ex

$$A = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 1 & 4 & -3 & 0 \\ -1 & -2 & 0 & -3 \end{bmatrix} \qquad \lambda = 5, 5, -3, -3 \text{ (with multiplicities)}$$

$\lambda = 5$ eigenspace:

$$\text{Nul}(A - 5I_4) = \text{Nul}\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 4 & -8 & 0 \\ -1 & -2 & 0 & -8 \end{bmatrix} = \text{Nul}\begin{bmatrix} 1 & 4 & -8 & 0 \\ 0 & 2 & -8 & -8 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$= \text{Nul}\begin{bmatrix} 1 & 0 & 8 & 16 \\ 0 & 1 & -4 & -4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \text{Span}\left\{ \begin{bmatrix} -8 \\ 4 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -16 \\ 4 \\ 0 \\ 1 \end{bmatrix} \right\}$$

$\lambda = -3$ eigenspace:

$$\text{Nul}(A + 3I_4) = \text{Nul}\begin{bmatrix} 8 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 \\ 1 & 4 & 0 & 0 \\ -1 & -2 & 0 & 0 \end{bmatrix} = \text{Nul}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$= \text{Span}\left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

So $A = PDP^{-1}$ with $P = \begin{bmatrix} -8 & -16 & 0 & 0 \\ 4 & 4 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$ $D = \begin{bmatrix} 5 & & & \\ & 5 & & \\ & & -3 & \\ & & & -3 \end{bmatrix}$

(Note: to check you can calculate AP and PD, avoiding an inverse.)

Note also: diagonalization is not unique: the eigenvectors in P may come in any order!

Why diagonalize? A^n and $e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n$, which are used for studying dynamical systems, discrete and continuous.

Matrix of a transformation

Let V and W be finite dimensional vector spaces, $\dim V = n$, $\dim W = m$, $T: V \to W$ a linear transformation, $\mathcal{B}$ a basis of V and $\mathcal{C}$ a basis of W.

Let us organize all of this!

$$\begin{tikzcd} \mathbb{R}^n \arrow[d, "\mathcal{B}"'] & \mathbb{R}^m \arrow[d, "\mathcal{C}"] \\ V \arrow[r, "T"] & W \end{tikzcd}$$

Question: is there a transformation A from coordinates in $\mathbb{R}^n$ to coordinates in $\mathbb{R}^m$ so that $T(\mathcal{B}\vec{c}) = \mathcal{C}A\vec{c}$?

Yes: $A = \mathcal{C}^{-1} T \mathcal{B}$ is $\mathbb{R}^n \to \mathbb{R}^m$. This also gives that $T = \mathcal{C} A \mathcal{B}^{-1}$.

To calculate A, we may go column-by-column:

$$A\vec{e}_i = \mathcal{C}^{-1} T \mathcal{B} \vec{e}_i$$

ex $T: \mathbb{P}^2 \to \mathbb{P}^2 \qquad p(x) \mapsto p'(x)$

$\mathcal{B} = (1 \;\; x \;\; x^2)$

$\mathcal{C} = (1-x \;\; 1+x \;\; x^2)$

all for sake of exercise.

$$\mathcal{C}^{-1} T(\mathcal{B}\vec{e}_1) = \mathcal{C}^{-1} T(1) = \mathcal{C}^{-1}(0) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\mathcal{C}^{-1} T(\mathcal{B}\vec{e}_2) = \mathcal{C}^{-1} T(x) = \mathcal{C}^{-1}(1) = \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \end{bmatrix}$$

$$\mathcal{C}^{-1} T(\mathcal{B}\vec{e}_3) = \mathcal{C}^{-1} T(x^2) = \mathcal{C}^{-1}(2x) = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

So $A = \begin{bmatrix} 0 & 1/2 & -1 \\ 0 & 1/2 & 1 \\ 0 & 0 & 0 \end{bmatrix}$.

test: $T(1+x+x^2) = \mathcal{C} A \mathcal{B}^{-1}(1+x+x^2) = \mathcal{C} A \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \mathcal{C} \begin{bmatrix} -1/2 \\ 3/2 \\ 0 \end{bmatrix}$

$= \frac{-1}{2}(1-x) + \frac{3}{2}(1+x) = 1 + 2x.$

Full diagram:

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "A"] \arrow[d, "\mathcal{B}"'] & \mathbb{R}^m \arrow[d, "\mathcal{C}"] \\ V \arrow[r, "T"] & W \end{tikzcd}$$

Coordinate spaces

vector spaces.

Or, notice $T(c_1\vec{b_1} + \cdots + c_n\vec{b_n}) = c_1T(\vec{b_1}) + \cdots + c_nT(\vec{b_n})$
$$= \left(T(\vec{b_1}) \;\cdots\; T(\vec{b_n})\right)\vec{c}$$

Writing each $T(\vec{b_i})$ as $\mathcal{C}\vec{a_i}$,
$$= \mathcal{C}\left(\vec{a_1} \;\cdots\; \vec{a_n}\right)\vec{c}$$
Since $\vec{c} = \mathcal{B}^{-1}\vec{v}$, $\quad = \mathcal{C}A\mathcal{B}^{-1}\vec{v}$.

For linear operators $T: V \to V$, it tends to be useful to have the same basis on either end of T

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "A"] \arrow[d, "\mathcal{B}"'] & \mathbb{R}^n \arrow[d, "\mathcal{B}"] \\ V \arrow[r, "T"] & V \end{tikzcd}$$

since we might want to apply T (and thus A) repeatedly.

Here, $T = \mathcal{B}A\mathcal{B}^{-1}$.
"To compute T, find the coordinate vector relative to $\mathcal{B}$, mult. by A, then use the resulting coordinate to produce a lin. comb. of $\mathcal{B}$ to get a V vector."

We can diagonalize a linear operator by finding a basis $\mathcal{B}$ with respect to which the operator's matrix is diagonal.

ex an $n\times n$ matrix A is a linear operator $\vec{x} \mapsto A\vec{x}$, and $A = PDP^{-1}$ means P is basis of $\mathbb{R}^n$ with respect to which A has diagonal matrix D.

(it is hard finding reasonable non-matrices: f.d. v.space $\Rightarrow$ iso. to $\mathbb{R}^n$, after all.)

Complex numbers

Real numbers are somewhat anemic: not every polynomial has a root. This is an issue since then not every characteristic polynomial has a root. For instance, $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ has λ^2+1.

What if we just introduce roots to λ^2+1? The quadratic formula suggests $\pm\sqrt{-1}$ are the two roots: $(\sqrt{-1})^2+1=-1+1=0$. Let $i=\sqrt{-1}$ be the imaginary root (despite the name, this number is quite real to modern mathematicians).

"complex" as in tied together

Let us throw it in to $\mathbb{R}$, and let the complex numbers $\mathbb{C}$ be the smallest number system which contains i. In fact, every number in $\mathbb{C}$ is of the form $a+bi$ for $a,b\in\mathbb{R}$ due to the following rules:

(i) $(a+bi)(c+di) = ac+adi+bci+bdi^2$
$= (ac-bd)+(ad+bc)i$

(ii) $\dfrac{a+bi}{c-di}\cdot\dfrac{c+di}{c+di} = \dfrac{(ac-bd)+(ad+bc)i}{c^2+d^2}$

The complex conjugate of $a+bi$ is $\overline{a+bi}=a-bi$. It replaces all instances of i with $-i$. Properties:

(1) $(a+bi)(\overline{a-bi}) = a^2+b^2$
(2) $\overline{z+w}=\overline{z}+\overline{w}$ (easy to check)
(3) $\overline{zw}=\overline{z}\,\overline{w}$ (easy to check, just takes time)
(4) $\overline{z}^{-1}=\overline{z^{-1}}$

Thm (Fundamental theorem of algebra) Every non-constant polynomial has at least one complex root.

Consequence: by long division by $x-r$ whenever r is a root, every polynomial is a product of linear factors!

ex $x^3 - 2x^2 + 2x - 1$ has 1 as a root (by inspection)

$$\begin{array}{r l} & x^2 - x \quad\quad + 1 \\ x-1 \,\big)\! & \overline{x^3 - 2x^2 + 2x - 1} \\ & x^3 - x^2 \\ & \overline{\quad -x^2 + 2x} \\ & \quad -x^2 + x \\ & \overline{\quad\quad\quad x - 1} \end{array}$$

so $= (x-1)(x^2 - x + 1)$

$\frac{1 \pm \sqrt{1-4}}{2}$ other roots

$= \frac{1 \pm \sqrt{3}i}{2}$

$= (x-1)\left(x - \frac{1+\sqrt{3}i}{2}\right)\left(x - \frac{1-\sqrt{3}i}{2}\right)$

Another consequence: adding i is enough! All polys have roots!

ex $\mathbb{C}$ is a real vector space.

$\mathbb{C} = \operatorname{Span}\{1, i\}$. $\dim \mathbb{C} = 2$ since

$c1 + di = 0$ implies

$\sqrt{-1} = \frac{c}{d}$ but $-1 \neq \frac{c^2}{d^2}$. (positive)

$i : \mathbb{C} \to \mathbb{C}$ defined by $z \mapsto iz$ is a linear transformation.

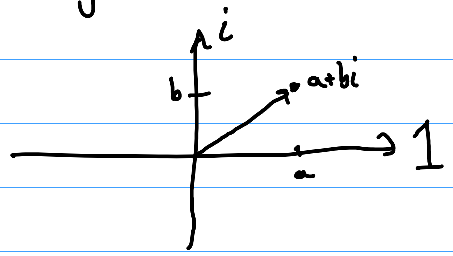
$i(a+bi) = -b + ai$, so in basis $(1 \ i)$,

matrix is $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

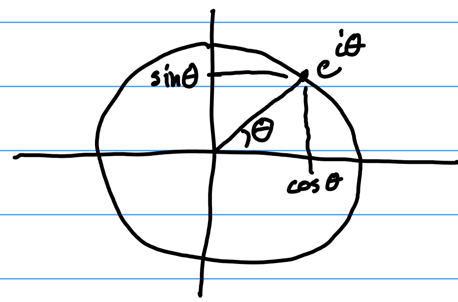
In fact, $\mathbb{C}$ is the same as having $a+bi$ be $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$.

$\det\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = a^2 + b^2$, which is zero only for the zero matrix.

Argand diagram:



Euler's formula $e^{i\theta} = \cos\theta + i\sin\theta$ is convenient notation justified by Taylor series. On the Argand diagram,



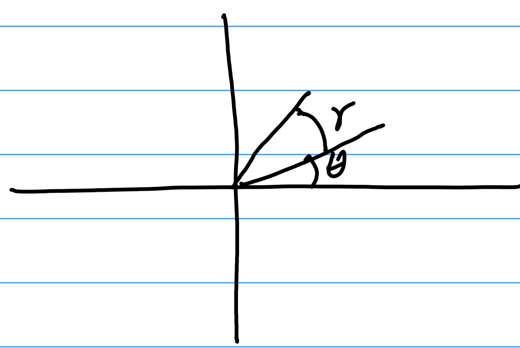
Every complex number can be written as $re^{i\theta}$, for $r \in \mathbb{R}$ non-negative the distance to 0.

for $z = a+bi$

$$z = \sqrt{a^2+b^2}\; e^{i\,\mathrm{atan}(b,a)} \quad \text{or} \quad = \underbrace{|z|}_{\text{magnitude}}\, e^{i\,\overbrace{\arg(z)}^{\text{argument}}}$$

atan(b,a) is like $\tan^{-1}(\frac{b}{a})$ but reaches every angle $[0, 2\pi)$

$(re^{i\theta})(se^{i\gamma}) = (rs)\,e^{i(\theta+\gamma)}$, so angles are added, magnitudes multiplied.

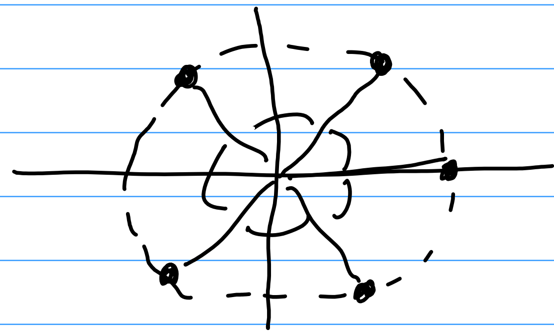


In other words: multiplying $\mathbb{C}$ by $re^{i\theta}$ scales the plane by r while rotating by θ CCW.

$(re^{i\theta})^n = r^n e^{ni\theta} = r^n(\cos(n\theta) + i\sin(n\theta))$

ex Solve $z^5 = 1$. For $z = re^{i\theta}$, $r^5 e^{5i\theta} = 1$, so $r = 1$ and 5θ is a multiple of 2π. $\theta = 0, \frac{2\pi}{5}, 2\cdot\frac{2\pi}{5}, 3\cdot\frac{2\pi}{5}, 4\cdot\frac{2\pi}{5}$.

Picture:



(vertices of regular pentagon)

Thm If z is a root of a polynomial with real coefficients, so is $\bar{z}$.

Pf If $a_n z^n + \cdots + a_1 z + a_0 = 0$

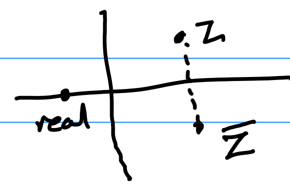
$$\overline{a_n z^n + \cdots + a_1 z + a_0} = \bar{0}$$

$$\overline{a_n}\, \bar{z}^n + \cdots + \overline{a_1}\, \bar{z} + \overline{a_0} = 0 \qquad a_n, \ldots, a_0 \text{ real, so}$$

$$a_n \bar{z}^n + \cdots + a_1 \bar{z} + a_0 = 0$$

Thus $\bar{z}$ is also a root.

For real roots, this says nothing, but for non-real roots, they come in conjugate pairs:



(This means every real poly factors into real quadratic and linear terms.)

ex $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ has $\lambda = i, -i$

$\lambda = i$: $\text{Nul}\begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix} = \text{Span}\left\{\begin{bmatrix} 1 \\ -i \end{bmatrix}\right\}$

$\lambda = -i$: $\text{Nul}\begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix} = \text{Span}\left\{\begin{bmatrix} 1 \\ i \end{bmatrix}\right\}$

$A = PDP^{-1} \qquad P = \begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix} \qquad D = \begin{bmatrix} i & \\ & -i \end{bmatrix}$

This leaves $\mathbb{R}^2$ to go to $\mathbb{C}^2$, a complex (not real) vector space! Same, but $\mathbb{C}$ scalars rather than $\mathbb{R}$!