Eigenvectors and diagonalization

- eigenvectors
- diagonalization

Eigenvectors and eigenvalues

$\lambda \in \mathbb{C}$ is called an *eigenvalue* of $A \in \mathbb{C}^{n\times n}$ if

$A = \begin{bmatrix} i & 0 \\ 0 & 2i \end{bmatrix} \to \begin{cases} \lambda_1 = i \\ \lambda_2 = 2 \end{cases}$

$\lambda I - A = \begin{bmatrix} \lambda & 1 \\ -1 & \lambda \end{bmatrix}$

$\det(\lambda I - A) = \lambda^2 + 1 = 0$

$\lambda^2 = -1$

$$\mathcal{X}(\lambda) = \mathbf{det}(\lambda I - A) = 0$$

$\lambda = \pm i$

equivalent to: Characteristic polynomial

- there exists nonzero $v \in \mathbb{C}^n$ s.t. $(\lambda I - A)v = 0$, *i.e.*,

$\lambda = i$

$\lambda = -i$

$$Av = \lambda v$$

any such v is called an *eigenvector* of A (associated with eigenvalue λ)

- there exists nonzero $w \in \mathbb{C}^n$ s.t. $w^\mathsf{T}(\lambda I - A) = 0$, *i.e.*,

$$w^\mathsf{T} A = \lambda w^\mathsf{T}$$

any such w is called a *left eigenvector* of A

$A^T w = \lambda w$

if a scalar λ and a vector v exist, s.t.

$Av = \lambda v$

then λ is an eigenvalue of A and v is the corresponding eigenvector

Complex eigenvalues and eigenvectors

$Av = \lambda v$
$A(\alpha v) = \lambda(\alpha v)$

- if v is an eigenvector of A with eigenvalue λ, then so is αv, for any $\alpha \in \mathbb{C}$, $\alpha \neq 0$
- even when A is real, eigenvalue λ and eigenvector v can be complex

A, λ real

- when A and λ are real, we can always find a real eigenvector v associated with λ: if $Av = \lambda v$, with $A \in \mathbb{R}^{n \times n}$, $\lambda \in \mathbb{R}$, and $v \in \mathbb{C}^n$, then

$v = v_r + i v_i$

$$A\Re v = \lambda \Re v, \qquad A\Im v = \lambda \Im v$$

so $\Re v$ and $\Im v$ are real eigenvectors, if they are nonzero
(and at least one is)

- *conjugate symmetry*: if A is real and $v \in \mathbb{C}^n$ is an eigenvector associated with $\lambda \in \mathbb{C}$, then $\overline{v}$ is an eigenvector associated with $\overline{\lambda}$:

taking conjugate of $Av = \lambda v$ we get $\overline{Av} = \overline{\lambda v}$, so

$\overline{A}\,\overline{v} = \overline{\lambda}\,\overline{v}$

$$A\overline{v} = \overline{\lambda}\overline{v}$$

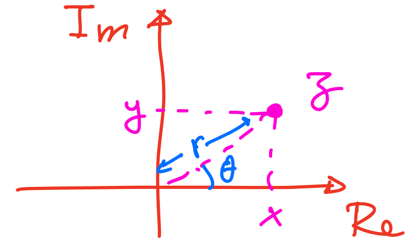
$z \in \mathbb{R}$
$\Updownarrow$
$z = \overline{z}$

$z = x + iy$
$\overline{z} = x - iy$
$z\overline{z} = x^2 + y^2$

we'll assume A is real from now on ...

$$z = x + iy$$

real imaginary



$$Re(z) = x$$

$$Im(z) = y$$

$$i^2 = -1$$

$$\bar{z} = x - iy$$

$$z = re^{i\theta}$$

$$= r(\cos\theta + i\sin\theta)$$

$$z\bar{z} = (x+iy)(x-iy) = x^2 - (iy)^2$$

$$= x^2 - \underbrace{(i)^2}_{-1} y^2$$

$$= x^2 + y^2 = |z|$$

$$z = re^{i\theta}$$

$$r = \sqrt{x^2+y^2}$$

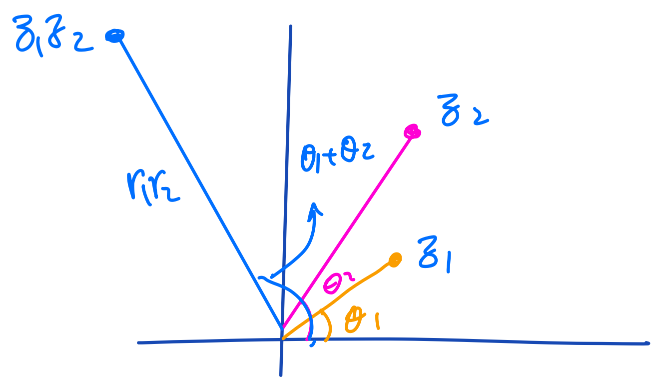
$$\theta = \tan^{-1} y/x$$

$$z_1 = x + iy$$

$$z_2 = a + ib$$

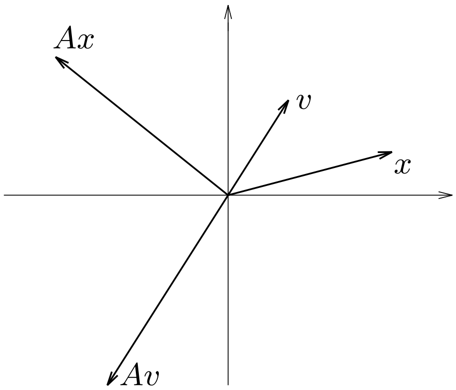
$$z_1 z_2 = (x+iy)(a+ib) = (xa - yb) + i(xb + ay)$$

$$r_1 e^{i\theta_1} \cdot r_2 e^{i\theta_2} = (r_1 r_2) e^{i(\theta_1+\theta_2)}$$



Scaling interpretation

$y = Ax$



(assume $\lambda \in \mathbb{R}$ for now; we'll consider $\lambda \in \mathbb{C}$ later)

if v is an eigenvector, effect of A on v is very simple: scaling by λ

- $\lambda \in \mathbb{R}$, $\lambda > 0$: v and Av point in same direction
- $\lambda \in \mathbb{R}$, $\lambda < 0$: v and Av point in opposite directions
- $\lambda \in \mathbb{R}$, $|\lambda| < 1$: Av smaller than v
- $\lambda \in \mathbb{R}$, $|\lambda| > 1$: Av larger than v

(we'll see later how this relates to stability of continuous- and discrete-time systems...)

Diagonalization

suppose $v_1, \ldots, v_n$ is a **_linearly independent_** set of eigenvectors of $A \in \mathbb{R}^{n\times n}$:

$$Av_i = \lambda_i v_i, \quad i = 1, \ldots, n$$

express as

$$A \underbrace{\begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}}_{T} = \underbrace{\begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}}_{T} \overbrace{\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}}^{\Lambda}$$

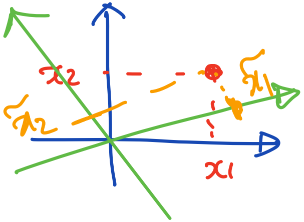
define $T = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$ and $\Lambda = \mathbf{diag}(\lambda_1, \ldots, \lambda_n)$, so

$$\boxed{AT = T\Lambda}$$

holds for ANY matrix A

$$T^{-1}AT = \Lambda \qquad A = T\Lambda T^{-1}$$

Diagonalization



$$\boxed{T^{-1}AT = \Lambda}$$

$$\begin{bmatrix}\lambda_1 & & \emptyset \\ & \ddots & \\ \emptyset & & \lambda_n\end{bmatrix}$$

- T invertible means $v_1, \ldots, v_n$ linearly independent
- similarity transformation by T *diagonalizes* A
- existence of invertible T such that

$$T^{-1}AT = \Lambda = \mathbf{diag}(\lambda_1, \ldots, \lambda_n)$$

  is equivalent to existence of a linearly independent set of n eigenvectors
- we say A is *diagonalizable*
- if A is not diagonalizable, it is sometimes called *defective*

Not all matrices are diagonalizable

example: $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$

- characteristic polynomial is $\mathcal{X}(s) = s^2$, so $\lambda = 0$ is only eigenvalue
- eigenvectors satisfy $Av = 0v = 0$, *i.e.*

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0$$

- so all eigenvectors have form $v = \begin{bmatrix} v_1 \\ 0 \end{bmatrix}$ where $v_1 \neq 0$
- thus, A cannot have two independent eigenvectors

Distinct eigenvalues

$A = I \longrightarrow \chi(\lambda) = \det(\lambda I - A)$

$= \det(\lambda I - I) = (\lambda - 1)^n = 0$

$\longrightarrow \boxed{\lambda = 1}$

any $x \in \mathbb{R}^n$ $\neq 0$

is an eigenvector

if A has distinct eigenvalues then A is diagonalizable

$T = I$

$T^{-1} A T = I^{-1} I I$

- distinct eigenvalues means $\lambda_i \neq \lambda_j$ for $i \neq j$
- the converse is false — A can have repeated eigenvalues but still be diagonalizable

Diagonalization and left eigenvectors

orthogonal matrix: $Q^{-1} = Q^T$

$AT = T\Lambda \rightarrow T^{-1}AT = \Lambda$

$A = T\Lambda T^{-1}$

$\boxed{T^{-1}A = \Lambda T^{-1}}$

rewrite $T^{-1}AT = \Lambda$ as $T^{-1}A = \Lambda T^{-1}$, or

$$\begin{bmatrix} w_1^T \\ \vdots \\ w_n^T \end{bmatrix} A = \Lambda \begin{bmatrix} w_1^T \\ \vdots \\ w_n^T \end{bmatrix}$$

$$\underbrace{\begin{bmatrix} w_1^T \\ \vdots \\ w_n^T \end{bmatrix}}_{T^{-1}} \underbrace{\begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}}_{T} = I$$

where $w_1^T, \ldots, w_n^T$ are the rows of T^{-1}

thus

$$w_i^T A = \lambda_i w_i^T$$

$$\begin{bmatrix} & \vdots & \\ \cdots & w_i^T v_j & \\ & \vdots & \end{bmatrix}$$

i.e., the rows of T^{-1} are (lin. indep.) left eigenvectors, normalized so that

$$w_i^T v_j = \delta_{ij}$$

$$w_i^T v_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

(*i.e.*, left & right eigenvectors chosen this way are *dual bases*)

if T is orthogonal $\rightarrow T^{-1} = T^T \rightarrow w_i = v_i$

Diagonalization

$$A = T\Lambda T^{-1}$$

$$A^2 = T\Lambda \cancel{T^{-1}}\cancel{T}\Lambda T^{-1}$$

$$= T\Lambda^2 T^{-1}$$

diagonalization simplifies many matrix expressions

powers ($i.e.$, discrete-time solution to $\boxed{x(k+1) = Ax(k)}$):

$$\begin{aligned} A^k &= \left(T\Lambda T^{-1}\right)^k \\ &= \left(T\Lambda T^{-1}\right)\cdots\left(T\Lambda T^{-1}\right) \\ &= T\Lambda^k T^{-1} \\ &= T\,\mathbf{diag}(\lambda_1^k, \ldots, \lambda_n^k)T^{-1} \end{aligned}$$

$$= T \begin{bmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{bmatrix} T^{-1}$$

(for $k < 0$ only if A invertible, $i.e.$, all $\lambda_i \neq 0$)

Analytic function of a matrix

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots$$

for any analytic function $f : \mathbb{R} \to \mathbb{R}$, *i.e.*, given by power series

$$x(t) = e^{At}$$

$$f(a) = \beta_0 + \beta_1 a + \beta_2 a^2 + \beta_3 a^3 + \cdots$$

we can define $f(A)$ for $A \in \mathbb{R}^{n \times n}$ (*i.e.*, overload f) as

$$f(A) = \beta_0 I + \beta_1 A + \beta_2 A^2 + \beta_3 A^3 + \cdots$$

substituting $A = T \Lambda T^{-1}$, we have

$$\beta_0 + \beta_1 \lambda_1 + \beta_2 \lambda_2^2 + \cdots = f(\lambda_1)$$

$$\begin{aligned} f(A) &= \beta_0 I + \beta_1 A + \beta_2 A^2 + \beta_3 A^3 + \cdots \\ &= \beta_0 T T^{-1} + \beta_1 T \Lambda T^{-1} + \beta_2 (T \Lambda T^{-1})^2 + \cdots \\ &= T \left(\beta_0 I + \beta_1 \Lambda + \beta_2 \Lambda^2 + \cdots \right) T^{-1} \\ &= T\, \textbf{diag}(f(\lambda_1), \ldots, f(\lambda_n)) T^{-1} \end{aligned}$$

$$\begin{bmatrix} \boxed{} & & \emptyset \\ & \boxed{f(\lambda_2)} & \\ \emptyset & & \ddots \quad \boxed{} \end{bmatrix}$$

$f(\lambda_1)$ $f(\lambda_n)$