Eigenvectors and diagonalization

- eigenvectors
- diagonalization
**Eigenvectors and eigenvalues**

λ ∈ ℂ is called an *eigenvalue* of \( A \in \mathbb{C}^{n \times n} \) if

\[
\chi(\lambda) = \text{det}(\lambda I - A) = 0
\]

equivalent to:

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

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

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

  any such \( w \) is called a *left eigenvector* of \( A \)
Complex eigenvalues and eigenvectors

- if $v$ is an eigenvector of $A$ with eigenvalue $\lambda$, then so is $\alpha v$, for any $\alpha \in \mathbb{C}$, $\alpha \neq 0$
- even when $A$ is real, eigenvalue $\lambda$ and eigenvector $v$ can be complex
- when $A$ and $\lambda$ are real, we can always find a real eigenvector $v$ associated with $\lambda$: if $Av = \lambda v$, with $A \in \mathbb{R}^{n \times n}$, $\lambda \in \mathbb{R}$, and $v \in \mathbb{C}^n$, then
  
  $$A\Re v = \lambda \Re v, \quad A\Im v = \lambda \Im v$$

  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 $\bar{v}$ is an eigenvector associated with $\bar{\lambda}$:
  
  taking conjugate of $Av = \lambda v$ we get $\overline{Av} = \overline{\lambda v}$, so
  
  $$A\bar{v} = \overline{\lambda \bar{v}}$$

we'll assume $A$ is real from now on . . .
Scaling interpretation

(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 $\lambda$

- $\lambda \in \mathbb{R}, \lambda > 0$: $v$ 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... )
suppose \( v_1, \ldots, v_n \) is a \textit{linearly independent} set of eigenvectors of \( A \in \mathbb{R}^{n \times n} \): 

\[
A v_i = \lambda_i v_i, \quad i = 1, \ldots, n
\]

express as 

\[
A \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}
\]

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

\[
AT = T\Lambda
\]
Diagonalization

- $T$ invertible means $v_1, \ldots, v_n$ linearly independent
- similarity transformation by $T$ diagonalizes $A$
- existence of invertible $T$ such that
  \[ T^{-1}AT = \Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n) \]
  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 \( \chi(s) = s^2 \), 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

if $A$ has distinct eigenvalues then $A$ is diagonalizable

- 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

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

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

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

\[w_i^T A = \lambda_i w_i^T\]
i.e., the rows of $T^{-1}$ are (lin. indep.) left eigenvectors, normalized so that

\[w_i^T v_j = \delta_{ij}\]
(i.e., left & right eigenvectors chosen this way are dual bases)
Diagonalization simplifies many matrix expressions

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

\[
A^k = (T \Lambda T^{-1})^k
= (T \Lambda T^{-1}) \cdots (T \Lambda T^{-1})
= T \Lambda^k T^{-1}
= T \text{diag}(\lambda_1^k, \ldots, \lambda_n^k) T^{-1}
\]

(for \(k < 0\) only if \(A\) invertible, \(i.e., \text{all } \lambda_i \neq 0\))
Analytic function of a matrix

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

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

we can define \( f(A) \) for \( A \in \mathbb{R}^{n \times n} \) (\textit{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

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

\[
= \beta_0 TT^{-1} + \beta_1 T \Lambda T^{-1} + \beta_2 (T \Lambda T^{-1})^2 + \cdots
\]

\[
= T \left( \beta_0 I + \beta_1 \Lambda + \beta_2 \Lambda^2 + \cdots \right) T^{-1}
\]

\[
= T \text{diag}(f(\lambda_1), \ldots, f(\lambda_n))T^{-1}
\]