Lecture 11
Eigenvectors and diagonalization

• eigenvectors

• dynamic interpretation: invariant sets

• complex eigenvectors & invariant planes

• left eigenvectors

• diagonalization

• modal form

• discrete-time stability
Eigenvectors and eigenvalues

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

$$\chi(\lambda) = \det(\lambda I - A) = 0$$

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 $A$
• 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 \( \overline{v} \) is an eigenvector associated with \( \overline{\lambda} \):
  taking conjugate of \( Av = \lambda v \) we get \( \overline{A} \overline{v} = \overline{\lambda} \overline{v} \), so
  \[
  A \overline{v} = \overline{\lambda} \overline{v}
  \]

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 $\lambda$

(what is $\lambda$ here?)
• $\lambda \in \mathbb{R}, \lambda > 0$: $v$ 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. . . )
Dynamic interpretation

suppose $Av = \lambda v, v \neq 0$

if $\dot{x} = Ax$ and $x(0) = v$, then $x(t) = e^{\lambda t}v$

several ways to see this, e.g.,

$$x(t) = e^{tA}v = \left( I + tA + \frac{(tA)^2}{2!} + \cdots \right) v$$

$$= \quad v + \lambda tv + \frac{\lambda^2 t^2}{2!} v + \cdots$$

$$= \quad e^{\lambda t}v$$

(since $(tA)^k v = (\lambda t)^k v$)
• for $\lambda \in \mathbb{C}$, solution is complex (we'll interpret later); for now, assume $\lambda \in \mathbb{R}$

• if initial state is an eigenvector $v$, resulting motion is very simple — always on the line spanned by $v$

• solution $x(t) = e^{\lambda t}v$ is called *mode* of system $\dot{x} = Ax$ (associated with eigenvalue $\lambda$)

• for $\lambda \in \mathbb{R}$, $\lambda < 0$, mode contracts or shrinks as $t \uparrow$

• for $\lambda \in \mathbb{R}$, $\lambda > 0$, mode expands or grows as $t \uparrow$
Invariant sets

A set $S \subseteq \mathbb{R}^n$ is invariant under $\dot{x} = Ax$ if whenever $x(t) \in S$, then $x(\tau) \in S$ for all $\tau \geq t$

i.e.: once trajectory enters $S$, it stays in $S$

**vector field interpretation:** trajectories only cut into $S$, never out
suppose $Av = \lambda v$, $v \neq 0$, $\lambda \in \mathbb{R}$

- line $\{ tv \mid t \in \mathbb{R} \}$ is invariant
  
  (in fact, ray $\{ tv \mid t > 0 \}$ is invariant)

- if $\lambda < 0$, line segment $\{ tv \mid 0 \leq t \leq a \}$ is invariant
Complex eigenvectors

suppose \( Av = \lambda v, \ v \neq 0, \ \lambda \) is complex

for \( a \in \mathbb{C}, \) (complex) trajectory \( a e^{\lambda t}v \) satisfies \( \dot{x} = Ax \)

hence so does (real) trajectory

\[
x(t) = \Re (a e^{\lambda t}v) = e^{\sigma t} \begin{bmatrix} v_{re} & v_{im} \end{bmatrix} \begin{bmatrix} \cos \omega t & \sin \omega t \\ -\sin \omega t & \cos \omega t \end{bmatrix} \begin{bmatrix} \alpha \\ -\beta \end{bmatrix}
\]

where

\[
v = v_{re} + jv_{im}, \quad \lambda = \sigma + j\omega, \quad a = \alpha + j\beta
\]

• trajectory stays in invariant plane \( \text{span}\{v_{re}, v_{im}\} \)

• \( \sigma \) gives logarithmic growth/decay factor

• \( \omega \) gives angular velocity of rotation in plane
Dynamic interpretation: left eigenvectors

suppose $w^T A = \lambda w^T$, $w \neq 0$

then

$$\frac{d}{dt}(w^T x) = w^T \dot{x} = w^T Ax = \lambda (w^T x)$$

i.e., $w^T x$ satisfies the DE $d(w^T x)/dt = \lambda (w^T x)$

hence $w^T x(t) = e^{\lambda t} w^T x(0)$

• even if trajectory $x$ is complicated, $w^T x$ is simple
• if, e.g., $\lambda \in \mathbb{R}$, $\lambda < 0$, halfspace $\{ z \mid w^T z \leq a \}$ is invariant (for $a \geq 0$)
• for $\lambda = \sigma + j\omega \in \mathbb{C}$, $(\Re w)^T x$ and $(\Im w)^T x$ both have form

$$e^{\sigma t} (\alpha \cos(\omega t) + \beta \sin(\omega t))$$
Summary

• *right eigenvectors* are initial conditions from which resulting motion is simple (*i.e.*, remains on line or in plane)

• *left eigenvectors* give linear functions of state that are simple, for any initial condition
Example 1: \[
\dot{x} = \begin{bmatrix}
-1 & -10 & -10 \\
1 & 0 & 0 \\
0 & 1 & 0
\end{bmatrix} x
\]

Block diagram:

\[ \mathcal{X}(s) = s^3 + s^2 + 10s + 10 = (s + 1)(s^2 + 10) \]

Eigenvalues are \(-1, \pm j\sqrt{10}\)
trajectory with $x(0) = (0, -1, 1)$:
left eigenvector associated with eigenvalue $-1$ is

$$g = \begin{bmatrix} 0.1 \\ 0 \\ 1 \end{bmatrix}$$

let's check $g^T x(t)$ when $x(0) = (0, -1, 1)$ (as above):
eigenvector associated with eigenvalue $j\sqrt{10}$ is

$$v = \begin{bmatrix} -0.554 + j0.771 \\ 0.244 + j0.175 \\ 0.055 - j0.077 \end{bmatrix}$$

so an invariant plane is spanned by

$$v_{re} = \begin{bmatrix} -0.554 \\ 0.244 \\ 0.055 \end{bmatrix}, \quad v_{im} = \begin{bmatrix} 0.771 \\ 0.175 \\ -0.077 \end{bmatrix}$$
for example, with $x(0) = v_{re}$ we have

\begin{align*}
\begin{array}{cccc}
\v1 & \v2 & \v3 \\
\hline
-1 & 0 & 0.1 \\
0.5 & -0.5 & 0.1 \\
0 & 0 & 0.1 \\
\end{array}
\end{align*}
Example 2: Markov chain

probability distribution satisfies \( p(t + 1) = Pp(t) \)

\[ p_i(t) = \text{Prob}( z(t) = i ) \] so \( \sum_{i=1}^{n} p_i(t) = 1 \)

\[ P_{ij} = \text{Prob}( z(t + 1) = i \mid z(t) = j ), \] so \( \sum_{i=1}^{n} P_{ij} = 1 \)
(such matrices are called stochastic)

rewrite as:

\[ [1 \ 1 \ \cdots \ 1]P = [1 \ 1 \ \cdots \ 1] \]

i.e., \( [1 \ 1 \ \cdots \ 1] \) is a left eigenvector of \( P \) with e.v. 1

hence \( \det(I - P) = 0 \), so there is a right eigenvector \( v \neq 0 \) with \( P v = v \)

it can be shown that \( v \) can be chosen so that \( v_i \geq 0 \), hence we can normalize \( v \) so that \( \sum_{i=1}^{n} v_i = 1 \)

**interpretation:** \( v \) is an *equilibrium distribution*; i.e., if \( p(0) = v \) then \( p(t) = v \) for all \( t \geq 0 \)

(if \( v \) is unique it is called the *steady-state distribution* of the Markov chain)
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

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

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

\[
AT = T\Lambda
\]

and finally

\[
T^{-1}AT = \Lambda
\]
• $T$ invertible since $v_1, \ldots, v_n$ linearly independent
• similarity transformation by $T$ diagonalizes $A$

conversely if there is a $T = [v_1 \cdots v_n]$ s.t.

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

then $AT = T\Lambda$, i.e.,

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

so $v_1, \ldots, v_n$ is a linearly independent set of eigenvectors of $A$

we say $A$ is diagonalizable if

• there exists $T$ s.t. $T^{-1}AT = \Lambda$ is diagonal
• $A$ has a set of linearly independent eigenvectors

(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

**fact:** if $A$ has distinct eigenvalues, *i.e.*, $\lambda_i \neq \lambda_j$ for $i \neq j$, then $A$ is diagonalizable

(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)
Modal form

suppose \( A \) is diagonalizable by \( T \)

define new coordinates by \( x = T\tilde{x} \), so

\[
T\dot{x} = AT\tilde{x} \iff \dot{x} = T^{-1}AT\tilde{x} \iff \dot{\tilde{x}} = \Lambda\tilde{x}
\]
in new coordinate system, system is diagonal (decoupled):

\[ \tilde{x}_1 \]
\[ \frac{1}{s} \]
\[ \lambda_1 \]

\[ \vdots \]

\[ \vdots \]

\[ \tilde{x}_n \]
\[ \frac{1}{s} \]
\[ \lambda_n \]

trajectories consist of \( n \) independent modes, \( i.e. \)

\[ \tilde{x}_i(t) = e^{\lambda_i t} \tilde{x}_i(0) \]

hence the name \textit{modal form}
Real modal form

when eigenvalues (hence $T$) are complex, system can be put in real modal form:

$$S^{-1}AS = \text{diag}\left(\Lambda_r, \begin{bmatrix} \sigma_{r+1} & \omega_{r+1} \\ -\omega_{r+1} & \sigma_{r+1} \end{bmatrix}, \ldots, \begin{bmatrix} \sigma_n & \omega_n \\ -\omega_n & \sigma_n \end{bmatrix}\right)$$

where $\Lambda_r = \text{diag}(\lambda_1, \ldots, \lambda_r)$ are the real eigenvalues, and

$$\lambda_i = \sigma_i + j\omega_i, \quad i = r + 1, \ldots, n$$

are the complex eigenvalues
block diagram of ‘complex mode’:

\[
\begin{array}{c}
\sigma \\
\downarrow \\
1/s \\
\downarrow \\
\omega \\
\downarrow \\
1/s \\
\downarrow \\
\sigma
\end{array}
\]
diagonalization simplifies many matrix expressions

e.g., resolvent:
\[(sI - A)^{-1} = (sTT^{-1} - T\Lambda T^{-1})^{-1}\]
\[= (T(sI - \Lambda)T^{-1})^{-1}\]
\[= T(sI - \Lambda)^{-1}T^{-1}\]
\[= T \text{diag}\left(\frac{1}{s - \lambda_1}, \ldots, \frac{1}{s - \lambda_n}\right) T^{-1}\]

powers (i.e., discrete-time solution):
\[A^k = (T\Lambda T^{-1})^k\]
\[= (T\Lambda T^{-1}) \cdots (T\Lambda T^{-1})\]
\[= T\Lambda^k T^{-1}\]
\[= T \text{diag}(\lambda_1^k, \ldots, \lambda_n^k) T^{-1}\]

(for \(k < 0\) only if \(A\) invertible, i.e., all \(\lambda_i \neq 0\))
exponential (i.e., continuous-time solution):

\[
e^A = I + A + A^2/2! + \cdots \\
= I + T\Lambda T^{-1} + (T\Lambda T^{-1})^2/2! + \cdots \\
= T(I + \Lambda + \Lambda^2/2! + \cdots)T^{-1} \\
= Te^{\Lambda}T^{-1} \\
= T \text{ diag}(e^{\lambda_1}, \ldots, e^{\lambda_n})T^{-1}
\]
Analytic function of a matrix

for any analytic function $f : \mathbb{R} \to \mathbb{R}$, 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 $A \in \mathbb{R}^{n \times n}$ (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

\[
\begin{align*}
  f(A) &= \beta_0 I + \beta_1 A + \beta_2 A^2 + \beta_3 A^3 + \cdots \\
  &= \beta_0 TT^{-1} + \beta_1 T \Lambda T^{-1} + \beta_2 (T \Lambda T^{-1})^2 + \cdots \\
  &= T \left( \beta_0 I + \beta_1 \Lambda + \beta_2 \Lambda^2 + \cdots \right) T^{-1} \\
  &= T \, \text{diag}(f(\lambda_1), \ldots, f(\lambda_n)) T^{-1}
\end{align*}
\]
Solution via diagonalization

assume $A$ is diagonalizable

consider LDS $\dot{x} = Ax$, with $T^{-1}AT = \Lambda$

then

$$x(t) = e^{tA}x(0) = Te^{\Lambda t}T^{-1}x(0) = \sum_{i=1}^{n} e^{\lambda_i t} (w_i^T x(0)) v_i$$

thus: any trajectory can be expressed as linear combination of modes
interpretation:

• (left eigenvectors) decompose initial state $x(0)$ into modal components $w_i^T x(0)$

• $e^{\lambda_i t}$ term propagates $i$th mode forward $t$ seconds

• reconstruct state as linear combination of (right) eigenvectors
application: for what $x(0)$ do we have $x(t) \to 0$ as $t \to \infty$?

divide eigenvalues into those with negative real parts

$$\Re \lambda_1 < 0, \ldots, \Re \lambda_s < 0,$$

and the others,

$$\Re \lambda_{s+1} \geq 0, \ldots, \Re \lambda_n \geq 0$$

from

$$x(t) = \sum_{i=1}^{n} e^{\lambda_i t} (w_i^T x(0)) v_i$$

condition for $x(t) \to 0$ is:

$$x(0) \in \text{span}\{v_1, \ldots, v_s\},$$

or equivalently,

$$w_i^T x(0) = 0, \quad i = s + 1, \ldots, n$$

(can you prove this?)
Stability of discrete-time systems

suppose $A$ diagonalizable

consider discrete-time LDS $x(t + 1) = Ax(t)$

if $A = TΛT^{-1}$, then $A^k = TΛ^kT^{-1}$

then

$$x(t) = A^t x(0) = \sum_{i=1}^{n} \lambda_i^t (w_i^T x(0)) v_i \to 0 \quad \text{as} \quad t \to \infty$$

for all $x(0)$ if and only if

$$|λ_i| < 1, \quad i = 1, \ldots, n.$$

we will see later that this is true even when $A$ is not diagonalizable, so we have

**fact:** $x(t + 1) = Ax(t)$ is stable if and only if all eigenvalues of $A$ have magnitude less than one