4. Eigenvalues and Eigenvectors

- linear dependence and linear span
- definition of eigenvalues
- important properties
- similarity transform
- diagonalization

Linear Independence

Definition: a set of vectors $\{v_1, v_2, \ldots, v_n\}$ is **linearly independent** if

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0 \Longrightarrow \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$$

equivalent conditions:

- coefficients of $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$ are uniquely determined, i.e.,

  $$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n$$

  implies $\alpha_k = \beta_k$ for $k = 1, 2, \ldots, n$

- no vector v_i can be expressed as a linear combination of the other vectors

examples:

- $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$ are independent
- $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ are independent
- $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 2 \\ 0 \end{bmatrix}$ are not independent
- $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix}$ are not independent

Linear span

Definition: the linear span of a set of vectors

$$\{v_1, v_2, \ldots, v_n\}$$

is the set of all linear combinations of $v_1, \ldots, v_n$

$$\text{span}\{v_1, v_2, \ldots, v_n\} = \{a_1 v_1 + a_2 v_2 + \cdots + a_n v_n \mid a_1, \ldots, a_n \in \mathbf{R}\}$$

example:

$\text{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$ is the hyperplane on $x_1 x_2$ plane

Definition

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

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

equivalent to:

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

$$Ax = \lambda x$$

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

Computing eigenvalues

- $\mathcal{X}(\lambda) = \det(\lambda I - A)$ is called the **characteristic polynomial** of A
- $\mathcal{X}(\lambda) = 0$ is called the **characteristic equation** of A

the characteristic equation provides a way to compute the eigenvalues of A

$$A = \begin{bmatrix} 5 & 3 \\ -6 & -4 \end{bmatrix}$$

$$\mathcal{X}(\lambda) = \begin{vmatrix} \lambda - 5 & -3 \\ 6 & \lambda + 4 \end{vmatrix} = \lambda^2 - \lambda - 2 = 0$$

solving the characteristic equation gives

$$\lambda = 2, -1$$

Computing eigenvectors

for each eigenvalue of A, we can find an associated eigenvector from

$$(\lambda I - A)x = 0$$

where x is a **nonzero** vector

for A in page 4-6, let's find an eigenvector corresponding to $\lambda = 2$

$$(\lambda I - A)x = \begin{bmatrix} -3 & -3 \\ 6 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0 \quad \Longrightarrow \quad x_1 + x_2 = 0$$

the equation has many solutions, so we can form the set of solutions by

$$\left\{ x \in \mathbf{R}^2 \;\middle|\; x = \begin{bmatrix} x_1 \\ -x_1 \end{bmatrix} \right\} = \operatorname{span} \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$$

this set is called the **eigenspace** of A corresponding to $\lambda = 2$

Eigenspace

eigenspace of A corresponding to λ is defined as the nullspace of $\lambda I - A$

$$\mathcal{N}(\lambda I - A)$$

equivalent definition: solution space of the homogeneous system

$$(\lambda I - A)x = 0$$

- an eigenspace is a vector space (by definition)
- 0 is in every eigenspace but it is not an eigenvector
- the *nonzero* vectors in an eigenspace are the eigenvectors of A

from page 4-7, any nonzero vector lies in the eigenspace is an eigenvector of A, $e.g.$, $x = \begin{bmatrix} -1 & 1 \end{bmatrix}^T$

same way to find an eigenvector associated with $\lambda = -1$

$$(\lambda I - A)x = \begin{bmatrix} -6 & -3 \\ 6 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0 \quad \Longrightarrow \quad 2x_1 + x_2 = 0$$

so the eigenspace corresponding to $\lambda = -1$ is

$$\left\{ x \;\middle|\; x = \begin{bmatrix} x_1 \\ -2x_1 \end{bmatrix} \right\} = \operatorname{span} \left\{ \begin{bmatrix} 1 \\ -2 \end{bmatrix} \right\}$$

and $x = \begin{bmatrix} 1 & -2 \end{bmatrix}^T$ is an eigenvector of A associated with $\lambda = -1$

Properties

- if A is $n \times n$ then $\mathcal{X}(\lambda)$ is a polynomial of order n
- if A is $n \times n$ then there are n eigenvalues of A
- even when A is real, eigenvalues and eigenvectors can be complex, *e.g.*,

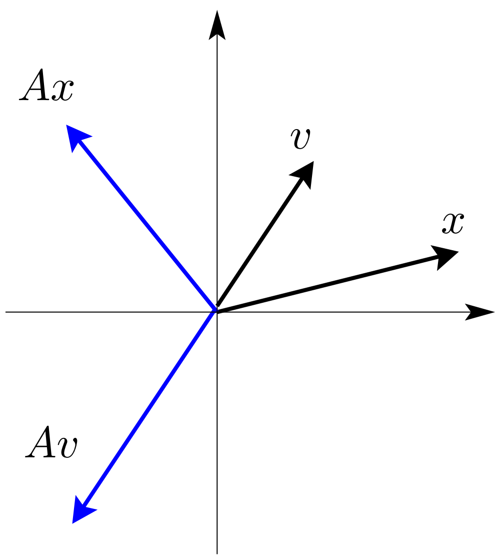
$$A = \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}, \quad A = \begin{bmatrix} -2 & 0 & 1 \\ -6 & -2 & 0 \\ 19 & 5 & -4 \end{bmatrix}$$

- if A and λ are real, we can choose the associated eigenvector to be real
- if A is real then eigenvalues must occur in complex conjugate pairs
- if x is an eigenvector of A, so is αx for any $\alpha \in \mathbf{C}$, $\alpha \neq 0$
- an eigenvector of A associated with λ lies in $\mathcal{N}(\lambda I - A)$

Scaling interpretation

assume λ is real

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



$\lambda > 0$	v and Av point in same direction
$\lambda < 0$	v and Av point in opposite directions
$\|\lambda\| < 1$	Av smaller than v
$\|\lambda\| > 1$	Av larger than v

Important facts

denote $\lambda(A)$ an eigenvalue of A

- $\lambda(\alpha A) = \alpha\lambda(A)$ for any $\alpha \in \mathbf{C}$
- $\mathbf{tr}(A)$ is the sum of eigenvalues of A
- $\det(A)$ is the product of eigenvalues of A
- A and A^T share the same eigenvalues ✎
- $\lambda(\overline{A^T}) = \overline{\lambda(A)}$ ✎
- $\lambda(A^m) = (\lambda(A))^m$ for any integer m
- A is invertible if and only if $\lambda = 0$ is not an eigenvalue of A ✎

Matrix powers

the mth power of a matrix A for a nonnegative integer m is defined as

$$A^m = \prod_{k=1}^{m} A$$

(the multiplication of m copies of A)

and A^0 is defined as the identity matrix, $i.e.$, $A^0 = I$

✌ **Facts:** if λ is an eigenvalue of A with an eigenvector v then

- λ^m is an eigenvalue of A^m
- v is an eigenvector of A^m associated with λ^m

Invertibility and eigenvalues

A is not invertible if and only if there exists a nonzero x such that

$$Ax = 0, \quad \text{or} \quad Ax = 0 \cdot x$$

which implies 0 is an eigenvalue of A

another way to see this is that

$$A \text{ is not invertible} \quad \Longleftrightarrow \quad \det(A) = 0 \quad \Longleftrightarrow \quad \det(0 \cdot I - A) = 0$$

which means 0 is a root of the characteristic equation of A

conclusion ✎ the following statements are equivalent

- A is invertible
- $\mathcal{N}(A) = \{0\}$
- $\lambda = 0$ is not an eigenvalue of A

Eigenvalues of special matrices

diagonal matrix:

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}$$

eigenvalues of D are the diagonal elements, *i.e.*, $\lambda = d_1, d_2, \ldots, d_n$

triangular matrix:

upper triangular | lower triangular

$$U = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix} \quad L = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

eigenvalues of L and U are the diagonal elements, *i.e.*, $\lambda = a_{11}, \ldots, a_{nn}$

Similarity transform

two $n \times n$ matrices A and B are said to be **similar** if

$$B = T^{-1}AT$$

for some invertible matrix T

T is called a **similarity transform**

✌ **invariant** properties under similarity transform:

- $\det(B) = \det(A)$
- $\mathbf{tr}(B) = \mathbf{tr}(A)$
- A and B have the same eigenvalues

$$\det(\lambda I - B) = \det(\lambda T^{-1}T - T^{-1}AT) = \det(\lambda I - A)$$

Diagonalization

an $n \times n$ matrix A is **diagonalizable** if there exists T such that

$$T^{-1}AT = D$$

is *diagonal*

- similarity transform by T *diagonalizes* A
- A and D are similar, so the entries of D must be the eigenvalues of A

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

- computing A^k is simple because $A^k = (TDT^{-1})^k = TD^kT^{-1}$

how to find a matrix T that diagonalizes A ?

suppose $\{v_1, \dots, v_n\}$ is a *linearly independent* set of eigenvectors of A

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

we can express this equation in the matrix form as

$$A \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

define $T = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}$ and $D = \mathbf{diag}(\lambda_1, \dots, \lambda_n)$, so

$$AT = TD$$

since T is invertible ($v_1, \dots, v_n$ are independent), finally we have

$$T^{-1}AT = D$$

conversely, if there exists $T = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$ that diagonalizes A

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

then $AT = TD$, or

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

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

conclusion: A is diagonalizable *if and only if*

n eigenvectors of A are linearly independent

(eigenvectors form a basis for $\mathbf{C}^n$)

- a diagonalizable matrix is called a **simple** matrix
- if A is not diagonalizable, sometimes it is called *defective*

Example

find T that diagonalizes

$$A = \begin{bmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{bmatrix}$$

the characteristic equation is

$$\det(\lambda I - A) = \lambda^3 - 11\lambda^2 + 39\lambda - 45 = 0$$

the eigenvalues of A are $\lambda = 5, 3, 3$

an eigenvector associated with $\lambda_1 = 5$ can be found by

$$(5 \cdot I - A)x = \begin{bmatrix} 1 & 0 & -1 \\ -2 & 2 & -2 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0 \quad \Longrightarrow \quad \begin{array}{c} x_1 - x_3 = 0 \\ x_2 - 2x_3 = 0 \\ x_3 \text{ is a free variable} \end{array}$$

an eigenvector is $v_1 = \begin{bmatrix} 1 & 2 & 1 \end{bmatrix}^T$

next, find an eigenvector associated with $\lambda_2 = 3$

$$(3 \cdot I - A)x = \begin{bmatrix} -1 & 0 & -1 \\ -2 & 0 & -2 \\ -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0 \quad \Longrightarrow \quad \begin{array}{c} x_1 + x_3 = 0 \\ x_2, x_3 \text{ are free variables} \end{array}$$

the eigenspace can be written by

$$\left\{ x \;\middle|\; x = x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\} = \text{span} \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

hence we can find two *independent* eigenvectors

$$v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad v_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

corresponding to the repeated eigenvalue $\lambda_2 = 3$

easy to show that v_1, v_2, v_3 are linearly independent

we form a matrix T whose columns are v_1, v_2, v_n

$$T = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

then v_1, v_2, v_3 are linearly independent if and only if T is invertible

by a simple calculation, $\det(T) = 2 \neq 0$, so T is invertible

hence, we can use this T to diagonalize A and it is easy to verify that

$$T^{-1}AT = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Not all matrices are diagonalizable

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

characteristic polynomial is $\det(\lambda I - A) = s^2$, so 0 is the only eigenvalue

eigenvector satisfies $Ax = 0 \cdot x$, *i.e.*,

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0 \quad \Longrightarrow \quad \begin{array}{c} x_2 = 0 \\ x_1 \text{ is a free variable} \end{array}$$

so all eigenvectors has form $x = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}$ where $x_1 \neq 0$

thus A cannot have *two independent* eigenvectors

Distinct eigenvalues

Theorem: if A has distinct eigenvalues, $i.e.$,

$$\lambda_i \neq \lambda_j, \quad i \neq j$$

then a set of corresponding eigenvectors are *linearly independent*

which further implies that A is diagonalizable

the converse is *false* – A can have repeated eigenvalues but still be diagonalizable

example: all eigenvalues of I are 1 (repeated eigenvalues) but I is diagonal

Proof by contradiction: assume the eigenvectors are dependent

(simple case) let $Ax_k = \lambda_k x_k$, $k = 1, 2$

suppose there exists $\alpha_1, \alpha_2 \neq 0$

$$\alpha_1 x_1 + \alpha_2 x_2 = 0 \tag{1}$$

multiplying (1) by A: $\quad \alpha_1 \lambda_1 x_1 + \alpha_2 \lambda_2 x_2 = 0$

multiplying (1) by λ_1: $\quad \alpha_1 \lambda_1 x_1 + \alpha_2 \lambda_1 x_2 = 0$

subtracting the above from the previous equation

$$\alpha_2(\lambda_2 - \lambda_1) x_2 = 0$$

since $\lambda_1 \neq \lambda_2$, we must have $\alpha_2 = 0$ and consequently $\alpha_1 = 0$

the proof for a general case is left as an exercise

Eigenvalues of symmetric matrices

A is an $n \times n$ (real) symmetric matrix, $i.e.$, $A = A^T$

x^* denotes $\bar{x}^T$ (complex conjugate transpose)

Facts ✌

- $y^* A y$ is real for all $y \in \mathbf{C}^n$
- all eigenvalues of A are real
- eigenvectors with distinct eigenvalues are orthogonal, i.e.,

$$\lambda_j \neq \lambda_k \quad \Longrightarrow \quad x_j^T x_k = 0$$

- there exists an *orthogonal* matrix U $(U^T U = U U^T = I)$ such that

$$A = U D U^T$$

(symmetric matrices are *always* diagonalizable)

MATLAB commands

`[V,D] = eig(A)` produces a diagonal matrix `D` of eigenvalues and a full matrix `V` whose columns are the corresponding eigenvectors

```
>> A = [5 3;-6 -4];
>> [V,D] = eig(A)
V =
    0.7071   -0.4472
   -0.7071    0.8944
D =
     2     0
     0    -1
```

$\lambda_1 = 2$ and $\lambda_2 = -1$ and the corresponding eigenvectors are

$$v_1 = \begin{bmatrix} 0.7071 & 0.7071 \end{bmatrix}^T, \quad v_2 = \begin{bmatrix} -0.4472 & 0.8944 \end{bmatrix}^T$$

note that the eigenvector is normalized so that it has unit norm

power of a matrix: use ^ to compute a power of A

```
>> A^3
ans =
    17     9
   -18   -10

>> eig(A^3)
ans =
     8
    -1

>> V*D^3*inv(V)
ans =
    17     9
   -18   -10
```

agree with the fact that the eigenvalue of A^3 is λ^3 and $A^3 = TD^3T^{-1}$

References

Chapter 5 in

H. Anton, *Elementary Linear Algebra*, 10th edition, Wiley, 2010

Lecture note on

Linear algebra, EE263, S. Boyd, Stanford university