7 Eigenvalues and Eigenvectors

7.1 Definition

• For a given n×n matrix A, consider the equation :

(1) $\boxed{\mathrm{Ax} = \lambda \mathrm{x}}$ "eigenvalue problem"

x : unknown vector
λ : unknown scalar

when (1) has a solution $\mathrm{x} \neq 0$,
· a value of λ for which a nonzero x exists is called an eigenvalue
· the corresponding nonzero x is called an eigenvector

• How to solve an eigenvalue problem?
$(\mathrm{A}\text{-}\lambda\mathrm{I})\mathrm{x} = 0$
For nonzero x, we need $\det(\mathrm{A}\text{-}\lambda\mathrm{I}) = 0$.
$\Rightarrow$Compute n roots (eigenvalues)
For each λ, solve $(\mathrm{A}\text{-}\lambda\mathrm{I})\mathrm{x} = 0$ for nonzero x.
(If x is an eigenvector, so is cx.)

note Elimination does not preserve λ.

ex) $U = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

ex) $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$

$$|A - \lambda I| = \begin{vmatrix} -\lambda & 1 \\ -1 & -\lambda \end{vmatrix} = \lambda^2 + 1 = 0 \qquad \Rightarrow \lambda_1 = i, \quad \lambda_2 = -i$$

$$\begin{bmatrix} -i & 1 \\ -1 & -1 \end{bmatrix} x_1 = 0 \qquad \Rightarrow x_1 = \begin{bmatrix} 1 \\ i \end{bmatrix}$$

$$\begin{bmatrix} i & 1 \\ -1 & 1 \end{bmatrix} x_2 = 0 \qquad \Rightarrow x_2 = \begin{bmatrix} 1 \\ -i \end{bmatrix}$$

7.2 Symmetric, Skew-symmetric, and Orthogonal Matrices

• For a real matrix A
· symmetric if $\mathrm{A}^T = \mathrm{A}$

$\cdot$ skew-symmetric if A^T = -A
$\cdot$ orthogonal if $A^T = A^{-1}$

- For a complex matrix A
  - $\cdot$ Hermitian if $\bar{A}^T = A$ $\leftarrow$ $\bar{a}_{jj} = a_{jj}$ (real diagonal)
  - $\cdot$ skew-Hermitian if $\bar{A}^T$ = -A $\leftarrow$ $\bar{a}_{jj} = -a_{jj}$ (pure imaginary or 0 on the diag.)
  - $\cdot$ unitary if $\bar{A}^T = A^{-1}$

Theorem [Eigenvalue]
(i) The eigenvalues of a Hermitian matrix (and of a symmetric matrix) are real.
(ii) The eigenvalues of a skew-Hermitian matrix (and of a shew-symmetric matrix) are pure imaginary or 0.
(iii) The eigenvalues of a unitary matrix (and of an orthogonal matrix) have absolute value 1.

Proof
(i) Let $\bar{A}^T = A$.

$$Ax = \lambda x \quad \Rightarrow \quad \bar{x}^T Ax = \bar{x}^T \lambda x = \lambda(\underbrace{|x_1|^2 + \cdots + |x_n|^2}_{real,nonzero}) \qquad (*)$$

$$(\bar{Ax}) = x^T \bar{A}\bar{x} = (x^T \bar{A}\bar{x})^T = \bar{x}^T \bar{A}^T = \bar{x}^T Ax \qquad (**)$$

$$\Rightarrow \bar{x}^T Ax \text{ is real.}$$

$$\therefore \lambda = \frac{real}{real} = real$$

(ii) Let $\bar{A}^T$ = -A.
Then (**) becomes $\overline{(\bar{x}^T Ax)}$ = -$\bar{x}^T Ax$ = pure imaginary or 0

$$\therefore \lambda = 0$$

(iii) Let $\bar{A}^T = A^{-1}$.
Then $\underbrace{Ax = \lambda x}_{\square} \quad \Rightarrow \quad \underbrace{(\bar{A}\bar{x})^T = (\bar{\lambda}\bar{x})^T = \bar{\lambda}\bar{x}^T}_{\blacksquare}$

multiply $\square$ and $\blacksquare$, then

$$(\bar{A}\bar{x})^T = \bar{\lambda}\bar{x}^T \lambda x = |\lambda|^2 \bar{x}^T x$$

$$|\lambda|^2 = 1.$$

7.3 Similarity and Diagonalization

Definition n×n matrices A and $\widehat{A}$ are *similar* if $\underline{\widehat{A} = P^{-1}AP}$ for some nonsingular matrix P.

Theorem If $\widehat{A}$ is similar to A, then $\widehat{A}$ has the same eigenvalues as A. And if x is an eigenvector of A, then $y = P^{-1}x$ is an eigenvector of $\widehat{A}$ corresponding to the same eigenvalues

Proof

$$\text{Let } Ax=\lambda x \ (x\neq 0)$$
$$P^{-1}Ax=P^{-1}\ \lambda x$$
$$P^{-1}APP^{-1}x=\lambda\ P^{-1}x$$
$$\widehat{A}\ P^{-1}x$$
$$\therefore P^{-1}x \text{ is an eigenvectors of } \widehat{A} \text{ corresponding to eigenvalues } \lambda$$

Theorem [Diagonalization] $\longmapsto$ (Strang pg. 288, Kreyzig pg.394)

Suppose the n × n matrix A has n linearly independent eigenvectors $x_1, \ldots, x_n$.

Let X = $[x_1 \cdots x_n]$: eigenvector matrix

Then $X^{-1}AX=\Lambda=\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}$: eigenvalue matrix(diagonalized!)

Proof

$$AX=A[x_1 \cdots x_n]=[Ax_1 \cdots Ax_n]=[\lambda_1 x_1 \cdots \lambda_n x_n]$$
$$=[x_1 \cdots x_n]\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix} = X\Lambda$$
$$\therefore X^{-1}AX = \Lambda \longleftarrow \text{We assumed } [x_1 \cdots x_n] \text{ are linearly independent.}$$

Note $\underbrace{\textit{Invertibility}}_{\text{whether an eigenvalue is 0 or not}}$ vs $\underbrace{\textit{Diagonalizability}}_{\text{whether there are n linearly independent eigvectors or not}}$

$\Rightarrow$ No connection!

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

$$\det(\text{A-}\lambda\text{I}) = \begin{vmatrix} 1-\lambda & 0 \\ 0 & 1-\lambda \end{vmatrix} = 0 \Rightarrow \lambda_{1,2} = 1 \text{ (repeated eigenvalues)}$$

$$\text{eigenvectors satisfy} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \text{x} = 0 \Rightarrow \text{x} = \text{c}\begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ (no second eigenvector)}$$

$\therefore$ A cannot be diagonalized.

Theorem Eigenvectors $x_1 \cdots x_j$ that correspond to *distinct* eigenvalues are *linearly independent*. Thus, an n $\times$ n matrix with n distinct eigenvalues is *diagonalizable*.

Proof

$$c_1x_1 + \cdots + c_{j-1}x_{j-1} + c_jx_j = \mathbf{0} \tag{1}$$

$$\lambda_1c_1x_1 + \cdots + \lambda_{j-1}c_{j-1}x_{j-1} + \lambda_jc_jx_j = \mathbf{0} \tag{2}$$

$$\lambda_j(1) \Rightarrow \lambda_jc_1x_1 + \cdots + \lambda_jc_{j-1}x_{j-1} + \lambda_jc_jx_j = \mathbf{0} \tag{3}$$

$$(2)-(3) \Rightarrow (\lambda_1-\lambda_j)c_1x_1 + \cdots + (\lambda_{j-1}-\lambda_j)c_{j-1}x_{j-1} = \mathbf{0} \tag{4}$$

$$\mathbf{A}(4) \Rightarrow (\lambda_1-\lambda_j)c_1x_1 + \cdots + (\lambda_{j-1}-\lambda_j)\lambda_{j-1}c_{j-1}x_{j-1} = \mathbf{0} \tag{5}$$

$$\lambda_{j-1}(4) \Rightarrow (\lambda_1-\lambda_j)\lambda_{j-1}c_1x_1 + \cdots + (\lambda_{j-1}-\lambda_j)\lambda_{j-1}c_{j-1}x_{j-1} = \mathbf{0} \tag{6}$$

$$(5)-(6) \Rightarrow (\lambda_1-\lambda_j)(\lambda_1-\lambda_{j-1})c_1x_1 + \cdots + (\lambda_{j-2}-\lambda_j)(\lambda_{j-2}-\lambda_{j-1})c_{j-1}x_{j-1} = \mathbf{0} \tag{7}$$

$\Longrightarrow$ continue this process. Eventually,

$$(\lambda_1-\lambda_j)(\lambda_1-\lambda_{j-1}\cdots(\lambda_1-\lambda_2)c_1x_1) = \mathbf{0}$$

$\Rightarrow c_1 = 0$

$\Rightarrow$ similarly, every $c_i = 0$

$\therefore x_1, \cdots, x_j$ are linearly independent.

when $\lambda_1, \cdots, \lambda_n$ are all different, we have linearly independent $x_1, \cdots, x_n$. Thus, $\mathbf{A}$ is diagonalizable in that case.