Slide 23-Eigenvalues

MAT2040 Linear Algebra

SSE, CUHK(SZ)

The motivation of the study for eigenvalue and eigenvector

Let $A \in \mathbb{R}^{n\times n}$, then $L(\mathbf{x}) = A\mathbf{x}$ is a linear operator from $\mathbb{R}^n$ to $\mathbb{R}^n$. Under this linear transformation, almost all the vectors will change their directions after the linear transformation.

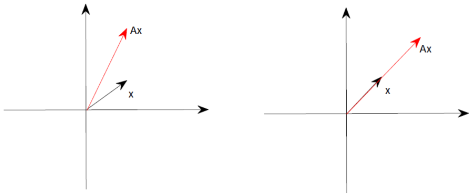

Figure: $\mathbf{x}$ is not an eigenvector. $\mathbf{x}$ is an eigenvector.

The motivation of the study for eigenvalue and eigenvector

If $\mathbf{x}$ and $A\mathbf{x}$ have the same/opposite direction, then there exists some constant λ such that $A\mathbf{x} = \lambda\mathbf{x}$, λ is called the eigenvalue of A, $\mathbf{x}$ is called the eigenvector of A.

Eigenvalues and eigenvectors are important concepts in linear algebra, and have many applications.

Definition 23.1 (**Eigenvalue and eigenvectors**) Let A be a square matrix with size $n \times n$ ($A \in \mathbb{R}^{n\times n}$ or $A \in \mathbb{C}^{n\times n}$), if there exists a scalar λ ($\lambda \in \mathbb{R}$ or $\lambda \in \mathbb{C}$) and nonzero vector $\mathbf{x}$ such that $A\mathbf{x} = \lambda\mathbf{x}$, then λ is called the **eigenvalue** (or **characteristic value**) and $\mathbf{x}$ is called the **eigenvector** (or **characteristic vector**) w.r.t λ.

Remark.

1. If $\mathbf{u}$ is an eigenvector w.r.t. eigenvalue λ, then so is $\alpha\mathbf{u}$ for any $\alpha \neq 0$. $A(\alpha\mathbf{u}) = \alpha A\mathbf{u} = \alpha\lambda\mathbf{u} = \lambda(\alpha\mathbf{u})$

2. If λ is the eigenvalue of A, then λ^s is the eigenvalue of A^s. Since $A^s\mathbf{x} = A^{s-1}A\mathbf{x} = \lambda A^{s-1}\mathbf{x} = \cdots = \lambda^s\mathbf{x}$

Example 23.2 Let

$$A = \begin{bmatrix} 0 & -3 \\ 1 & -4 \end{bmatrix}, \mathbf{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Then

$$A\mathbf{u} = \begin{bmatrix} 0 & -3 \\ 1 & -4 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 \\ -3 \end{bmatrix} = (-3)\mathbf{u}$$

Thus, -3 is the eigenvalue of A and the corresponding eigenvector is

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Theorem 23.3 Let Let A be a square matrix with size $n \times n$ ($A \in \mathbb{R}^{n\times n}$ or $A \in \mathbb{C}^{n\times n}$) and λ ($\lambda \in \mathbb{R}$ or $\lambda \in \mathbb{C}$), then the following statements are equivalent:

(a) λ is an eigenvalue of A.

(b) $(A - \lambda I)\mathbf{x} = \mathbf{0}$ has nontrivial solutions.

(c) $\text{Null}(A - \lambda I) \neq \{\mathbf{0}\}$, where $\text{Null}(A - \lambda I)$ is a subspace of $\mathbb{R}^n$ when $\lambda \in \mathbb{R}$ and $\text{Null}(A - \lambda I)$ is a subspace of $\mathbb{C}^n$ when $\lambda \in \mathbb{C}$. $\text{Null}(A - \lambda I)$ is called the **eigenspace** corresponding to λ.

(d) $A - \lambda I$ is singular.

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

Proof. Using the definition of Null space, determinant, matrix singular. It is easy to show that: λ is an eigenvalue of $A \Leftrightarrow (A - \lambda I)\mathbf{x} = \mathbf{0}$ has a nontrivial solution $\Leftrightarrow \text{Null}(A - \lambda I) \neq \{\mathbf{0}\} \Leftrightarrow A - \lambda I$ is singular $\Leftrightarrow$ $\det(A - \lambda I) = 0$.

The last condition (e) provides a method to calculate the eigenvalues.

Remark Let A be an $n \times n$ matrix, if λ is an eigenvalue of A, then λ is also the eigenvalue of A^T since $\det(A^T - \lambda I) = \det((A - \lambda I)^T) = \det(A - \lambda I)$. Thus A and A^T have the same eigenvalues.

Definition 23.4 (Characteristic Polynomial) Let A is a $n \times n$ matrix ($A \in \mathbb{R}^{n\times n}$ or $A \in \mathbb{C}^{n\times n}$) and λ is a variable, then

$$p_A(\lambda) = \det(A - \lambda I) = \begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix}$$

is called the characteristic polynomial of A, and

$$p_A(\lambda) = 0$$

is called the characteristic equation of A.

Example 23.5 Find the eigenvalues and corresponding eigenvectors for the following matrices

(1)

$$A = \begin{bmatrix} 3 & 2 \\ 3 & -2 \end{bmatrix}$$

The characteristic equation is

$$p_A(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 3-\lambda & 2 \\ 3 & -2-\lambda \end{vmatrix} = \lambda^2 - \lambda - 12 = 0.$$

$p_A(\lambda)$ is a polynomial with degree 2.

Thus, $\lambda = 4$ or $\lambda = -3$.

When $\lambda = 4$,

$$\mathrm{Null}(A - 4I) = \mathrm{Null}\left(\begin{bmatrix} -1 & 2 \\ 3 & -6 \end{bmatrix}\right) = \mathrm{Span}\left(\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right)$$

is the eigenspace corresponding to $\lambda = 4$. When $\lambda = -3$,

$$\mathrm{Null}(A + 3I) = \mathrm{Null}\left(\begin{bmatrix} 6 & 2 \\ 3 & 1 \end{bmatrix}\right) = \mathrm{Span}\left(\begin{bmatrix} 1 \\ -3 \end{bmatrix}\right)$$

is the eigenspace corresponding to $\lambda = -3$.

Thus $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ is the eigenvector w.r.t $\lambda = 4$ and $\begin{bmatrix} 1 \\ -3 \end{bmatrix}$ is the eigenvector w.r.t. $\lambda = -3$.

(2)

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

Then characteristic equation is

$$p_A(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 2-\lambda & -3 & 1 \\ 1 & -2-\lambda & 1 \\ 1 & -3 & 2-\lambda \end{vmatrix} = -\lambda(\lambda-1)^2 = 0$$

$p_A(\lambda)$ is a polynomial with degree 3.

Then $\lambda_1 = 0, \lambda_2 = \lambda_3 = 1$

$$A - 0I = \begin{bmatrix} 2 & -3 & 1 \\ 1 & -2 & 1 \\ 1 & -3 & 2 \end{bmatrix} \xrightarrow{\text{Row operations}} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

The eigenspace w.r.t. 0 is $\mathrm{Span}\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right)$.

$$A - I = \begin{bmatrix} 1 & -3 & 1 \\ 1 & -3 & 1 \\ 1 & -3 & 1 \end{bmatrix} \xrightarrow{\text{Row operations}} \begin{bmatrix} 1 & -3 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The eigenspace w.r.t. 1 is $\mathrm{Span}\left(\begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right)$.

(3)

$$A = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$$

The characteristic equation is

$$p_A(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 2 \\ -2 & 1-\lambda \end{vmatrix} = (\lambda - 1)^2 + 4 = 0.$$

Thus $\lambda = 1 \pm 2i$. When $\lambda = 1 + 2i$, then

$$A - (1+2i)I = \begin{bmatrix} -2i & 2 \\ -2 & -2i \end{bmatrix} \xrightarrow{\text{Row operations}} \begin{bmatrix} 1 & i \\ 0 & 0 \end{bmatrix}$$

The eigenvector w.r.t. $1 + 2i$ is $\begin{bmatrix} -i \\ 1 \end{bmatrix}$.

When $\lambda = 1 - 2i$,

$$A - (1-2i)I = \begin{bmatrix} 2i & 2 \\ -2 & 2i \end{bmatrix} \xrightarrow{\text{Row operations}} \begin{bmatrix} 1 & -i \\ 0 & 0 \end{bmatrix}$$

The eigenvector w.r.t. $1 - 2i$ is $\begin{bmatrix} i \\ 1 \end{bmatrix}$.

Observation: when A is a matrix with size $n \times n$ ($A \in \mathbb{R}^{n \times n}$ or $A \in \mathbb{C}^{n \times n}$), then the characteristic polynomial of A is a polynomial with degree n.

Theorem 23. 9 (**Fundamental theorem in Algebra**) Every degree n polynomial with complex coefficients has exactly n complex roots. (Counting with multiplicity).

By using this theorem, there will be exactly n eigenvalues (counting with multiplicity) for any matrix A with size $n \times n$ ($A \in \mathbb{R}^{n \times n}$ or $A \in \mathbb{C}^{n \times n}$).

Theorem 23.10 (Product and Sum of Eigenvalues) Let $A = (a_{ij})_{n\times n}$ be a square matrix ($A \in \mathbb{R}^{n\times n}$ or $A \in \mathbb{C}^{n\times n}$), $\lambda_i(i = 1, 2, \cdots, n)$ are the eigenvalues, then

(1) $\det(A) = \prod_{i=1}^{n} \lambda_i$

(2) $\sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i$, where $\sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + \cdots + a_{nn}$ is called the trace of A. Denoted as $\mathrm{Trace}(A) = \sum_{i=1}^{n} a_{ii}$.

Proof. By definition, one has the characteristic polynomial is

$$p_A(\lambda) = \det(A - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix}$$

Now expand the above determinant along the first column, one has

$$p_A(\lambda) = \det(A - \lambda I) = (a_{11} - \lambda) \begin{vmatrix} a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix}$$

$$+(-1)^{2+1} a_{21} \begin{vmatrix} a_{12} & a_{13} & \cdots & a_{1n} \\ a_{32} & a_{33} - \lambda & \cdots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n2} & a_{n3} & \cdots & a_{nn} - \lambda \end{vmatrix}$$

$$+\sum_{i=3}^{n}(-1)^{i+1}a_{i1}\times$$

$$\begin{vmatrix} a_{12} & a_{13} & \cdots & a_{1,i-1} & a_{1,i} & a_{1,i+1} & \cdots & a_{1n} \\ a_{22}-\lambda & a_{23} & \cdots & a_{2,i-1} & a_{2,i} & a_{2,i+1} & \cdots & a_{2n} \\ a_{32} & a_{33}-\lambda & \cdots & a_{3,i-1} & a_{3,i} & a_{3,i+1} & \cdots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1,2} & a_{i-1,3} & \cdots & a_{i-1,i-1}-\lambda & a_{i-1,i} & a_{i-1,i+1} & \cdots & a_{i-1,n} \\ a_{i+1,2} & a_{i+1,3} & \cdots & a_{i+1,i-1} & a_{i+1,i} & a_{i+1,i+1}-\lambda & \cdots & a_{i+1,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n2} & a_{n3} & \cdots & a_{n,i-1} & a_{n,i} & a_{n,i+1} & \cdots & a_{nn}-\lambda \end{vmatrix}$$

$$=(a_{11}-\lambda)\det(M_{11})+\sum_{i=2}^{n}(-1)^{i+1}a_{i1}\det(M_{i1})$$

where $M_{i1}(i = 2, \cdots, n)$ does not contain $a_{11} - \lambda$ and $a_{ii} - \lambda$, thus, all the terms $(-1)^{i+1} a_{i1} \det(M_{i1})(i = 2, \cdots, n)$ only involves the product of $n-2$ diagonal elements from $A - \lambda I$. The highest-order term for $\sum_{i=2}^{n} (-1)^{i+1} a_{i1} \det(M_{i1})$ is λ^{n-2}.

Expanding $\det(M_{11})$ using the same manner, we can conclude that

$$(a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda)$$

is the only term that involves the product of more than $n-2$ diagonal elements from $A - \lambda I$. Hence the highest-order term (term of λ^n) and the second highest-order term (term of λ^{n-1}) are uniquely determined by the product $(a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda)$.

Comparing coefficients for the term λ^n $(a_{11}-\lambda)(a_{22}-\lambda)\cdots(a_{nn}-\lambda)$

gives the highest-order term, which is $(-1)^n\lambda^n$. Thus, the highest-order term (term of λ^n) in the characteristic polynomial $p_A(\lambda)$ is $(-1)^n\lambda^n$.

On the other hand, since $\lambda_i(i=1,2,\cdots,n)$ are the roots of $p_A(\lambda)$, thus $p_A(\lambda)=(-1)^n(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)=(\lambda_1-\lambda)(\lambda_2-\lambda)\cdots(\lambda_n-\lambda)$.
Setting $\lambda=0$, then one has $p_A(0)=\det(A)=\prod_{i=1}^{n}\lambda_i$.

Comparing coefficients for the term λ^{n-1}

Look at the λ^{n-1} term of $(a_{11} - \lambda)(a_{22} - \lambda)\cdots(a_{nn} - \lambda)$ and $(\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_n - \lambda)$.

From the λ^{n-1} term of $(a_{11} - \lambda)(a_{22} - \lambda)\cdots(a_{nn} - \lambda)$, one can see that the coefficient of λ^{n-1} in the characteristic polynomial $p_A(\lambda)$ is $(-1)^{n-1}\sum_{i=1}^n a_{ii}$.

From $p_A(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_n - \lambda)$, one can see that the coefficient of λ^{n-1} is $(-1)^{n-1}\sum_{i=1}^n \lambda_i$. Comparing with $(-1)^{n-1}\sum_{i=1}^n a_{ii}$ and $(-1)^{n-1}\sum_{i=1}^n \lambda_i$, one has $\displaystyle\sum_{i=1}^n a_{ii} = \sum_{i=1}^n \lambda_i$.

Remark

1. A is nonsingular $\Leftrightarrow \det(A) \neq 0 \Leftrightarrow$ all eigenvalues $\lambda_i \neq 0$

2. A is nonsingular and λ is the eigenvalue of A, Then $\Leftrightarrow \lambda^{-1}$ is the eigenvalue of A^{-1}. $A\mathbf{x} = \lambda\mathbf{x} \Leftrightarrow A^{-1}\mathbf{x} = \frac{1}{\lambda}\mathbf{x}$.

Example 23.11 Let

$$A = \begin{bmatrix} 5 & -18 \\ 1 & -1 \end{bmatrix}$$

$\det(A) = 13$ and $\mathrm{Trace}(A) = 4$.

$$p_A(\lambda) = |A - \lambda I| = \begin{bmatrix} 5-\lambda & -18 \\ 1 & -1-\lambda \end{bmatrix} = \lambda^2 - 4\lambda + 13$$

$\lambda_1 = 2 - 3i, \lambda_2 = 2 + 3i.$

$\det(A) = \lambda_1\lambda_2 = 4 + 9 = 13,\ Trace(A) = \lambda_1 + \lambda_2 = 4.$

From above examples, one can see that even when A is a real matrix, the eigenvalues of A could be complex numbers. Thus, sometimes, we may need to deal with complex matrices.

Theorem 23.12 (Similar Matrices Have the Same Eigenvalues)
Let A, B are both $n \times n$ real matrices, if A B are similar, then two matrices have the same characteristic polynomial, and hence have the same eigenvalues.

Proof. Since A and B are similar, there exists a nonsingular matrix S, s.t. $B = S^{-1}AS$. The characteristic polynomial for B is

$$\begin{aligned}
p_B(\lambda) &= \det(B - \lambda I) \\
&= \det(S^{-1}AS - \lambda S^{-1}S) \\
&= \det(S^{-1}(A - \lambda I)S) \\
&= \det(S^{-1}) \det(A - \lambda I) \det(S) \\
&= \det(A - \lambda I) \\
&= p_A(\lambda)
\end{aligned}$$

where $\det(S^{-1}) = \frac{1}{\det(S)}$ is used. Thus, the characteristic polynomial of A and B are the same, they must have the same eigenvalues.

Example 23.13 Let

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & -2 \\ 6 & 6 \end{bmatrix}$$

It will be easy to check that

$$B = \begin{bmatrix} -1 & -2 \\ 6 & 6 \end{bmatrix} = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}^{-1} A \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}$$

Thus, A, B are similar and have the same eigenvalues.
In fact, $p_A(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 2-\lambda & 1 \\ 0 & 3-\lambda \end{vmatrix} = (2-\lambda)(3-\lambda)$.
Thus, both A and B have the eigenvalues 2, 3.