Eigenvalues and eigenvectors

Diagonalisation of a matrix

- Let A be an $n \times n$ matrix, it defines a linear transformation $f: R^n \to R^n$ sending $x \in R^n$ to $Ax \in R^n$.
- Let $\mathcal{B} = \{b_1, \cdots, b_n\}$ be a basis of R^n, let $B = [b_1 \cdots b_n]$, then B is an invertible matrix. Recall that under the basis $\mathcal{B}$, the linear transformation f is given by the matrix $A' = B^{-1}AB$. The matrix A and A' are said to be conjugate or similar.
- Question: Is it possible to find a basis of R^n such that A' is in simple form? Say diagonal?

- Reformulation of the question:

Notice that

$$B^{-1} A B = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \Leftrightarrow A B = B \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

$$\Leftrightarrow A [\vec{b}_1 \cdots \vec{b}_n] = [\lambda_1 \vec{b}_1 \cdots \lambda_n \vec{b}_n]$$

$$\Leftrightarrow A \vec{b}_i = \lambda_i \vec{b}_i, \quad i = 1, \cdots, n.$$

- Definition: Eigenvector (特征向量), eigenvalue (特征值)

An **eigenvector** of an $n \times n$ matrix A is a nonzero vector $\mathbf{x}$ such that $A\mathbf{x} = \lambda\mathbf{x}$ for some scalar λ. A scalar λ is called an **eigenvalue** of A if there is a nontrivial solution $\mathbf{x}$ of $A\mathbf{x} = \lambda\mathbf{x}$; such an $\mathbf{x}$ is called an *eigenvector corresponding to* λ.[1]

- In other words, they are solutions of the equation

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

- Its set of solutions $\mathrm{Nul}(A - \lambda I)$ for an eigenvalue λ will be called the eigenspace (特征子空间) of eigenvalue λ, denoted V_λ.
- The matrix A is said to be diagonalizable (可对角化) if A is similar to a diagonal matrix, i.e. there exists an invertible matrix B such that $A = BDB^{-1}$ for a diagonal matrix.

- The above discussion implies the following theorem:

The Diagonalization Theorem

An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.

In fact, $A = PDP^{-1}$, with D a diagonal matrix, if and only if the columns of P are n linearly independent eigenvectors of A. In this case, the diagonal entries of D are eigenvalues of A that correspond, respectively, to the eigenvectors in P.

- The question is reduced to calculating the eigenvalues and eigenvectors, I.e. to find $\lambda \in R$ such that the equation $(A - \lambda I)x = 0$ has non-trivial solutions.

- Recall that the equation $(A - \lambda I)x = 0$ has non-trivial solutions if and only if $A - \lambda I$ is not invertible, I.e. $det(A - \lambda I) = 0$. So to find the eigenvalues, we need to solve the last equation.
- Definition: The equation $det(A - \lambda I) = 0$ is called the characteristic equation (特征方程) of A. By the complete expansion of determinant, $det(A - \lambda I)$ is a polynomial of degree n, we call it the characteristic polynomial of A.
- The above discussion implies the theorem:

A scalar λ is an eigenvalue of an $n \times n$ matrix A if and only if λ satisfies the characteristic equation

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

- As the determinant is invariant under conjugation, it is clear that

If $n \times n$ matrices A and B are similar, then they have the same characteristic polynomial and hence the same eigenvalues (with the same multiplicities).

- Attention

WARNINGS:

1. The matrices

$$\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \text{ and } \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

are not similar even though they have the same eigenvalues.

2. Similarity is not the same as row equivalence. (If A is row equivalent to B, then $B = EA$ for some invertible matrix E.) Row operations on a matrix usually change its eigenvalues.

- Example: Find the characteristic polynomial and the eigenvalues of the matrix

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

Solution The characteristic polynomial equals

$$\det(A-\lambda I) = \begin{vmatrix} -1-\lambda & 4 & -2 \\ -3 & 4-\lambda & 0 \\ -3 & 1 & 3-\lambda \end{vmatrix}$$

$$= (-1-\lambda)\begin{vmatrix} 4-\lambda & 0 \\ 1 & 3-\lambda \end{vmatrix} - 4\begin{vmatrix} -3 & 0 \\ -3 & 3-\lambda \end{vmatrix} + (-2)\cdot\begin{vmatrix} -3 & 4-\lambda \\ -3 & 1 \end{vmatrix}$$

$$= (-1-\lambda)(4-\lambda)(3-\lambda) - 4(-3)(3-\lambda) - 2(-3+3(4-\lambda))$$

$$= -(\lambda+1)(4-\lambda)(3-\lambda) + \underbrace{36-12\lambda-18+6\lambda}_{18-6\lambda\,=\,6(3-\lambda)}$$

$$= (3-\lambda)(\lambda^2-3\lambda-4+6)$$

$$= (3-\lambda)(\lambda^2-3\lambda+2)$$

$$= (3-\lambda)(\lambda-1)(\lambda-2)$$

Hence $\det(A-\lambda I)$ has roots $\lambda = 1, 2, 3$, they are the eigenvalues of A.

- Theorem

The eigenvalues of a triangular matrix are the entries on its main diagonal.

Proof If A is triangular, then so is $A - \lambda I$, and we have

$$\det(A - \lambda I) = (a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda).$$

So the eigenvalues of A are $a_{11}, \cdots, a_{nn}$. $\boxtimes$

- Once we solve the characteristic equation $det(A - \lambda I) = 0$, we can proceed to determine the eigenspace V_λ, I.e. to solve the linear system $(A - \lambda I)x = 0$, to see whether A is diagonalizable.
- Proposition: Let $V_{\lambda_i}, i = 1, \cdots, r$ be eigenspaces with distinct eigenvalues $\lambda_1, \cdots, \lambda_r$. Let $\{b_{i,1}, \cdots, b_{i,d_i}\}$ be a basis for V_{λ_i} , then their union $\{b_{1,1}, \cdots, b_{1,d_1}, \cdots, b_{r,1}, \cdots, b_{r,d_r}\}$ remains linearly independent.
- Corollary

If $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are eigenvectors that correspond to distinct eigenvalues $\lambda_1, \ldots, \lambda_r$ of an $n \times n$ matrix A, then the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is linearly independent.

- Proof:

Suppose that there exists $x_{1,1}, \cdots, x_{1,d_1}, \cdots, x_{r,1}, \cdots, x_{r,d_r} \in \mathbb{R}$, such that

$$(*) \quad x_{1,1} b_{1,1} + \cdots + x_{1,d_1} b_{1,d_1} + \cdots + x_{r,1} b_{r,1} + \cdots + x_{r,d_r} b_{r,d_r} = 0$$

Applying $A - \lambda_1 I$ to $(*)$, get

$$(**) \quad \underbrace{(A-\lambda_1 I)(x_{1,1} b_{1,1} + \cdots + x_{1,d_1} b_{1,d_1})}_{0} + \underbrace{(A-\lambda_1 I)(x_{2,1} b_{2,1} + \cdots + x_{2,d_2} b_{2,d_2})}_{(\lambda_2-\lambda_1)(x_{2,1} b_{2,1} + \cdots + x_{2,d_2} b_{2,d_2})}$$

$$+ \cdots + \underbrace{(A-\lambda_1 I)(x_{r,1} b_{r,1} + \cdots + x_{r,d_r} b_{r,d_r})}_{(\lambda_r-\lambda_1)(x_{r,1} b_{r,1} + \cdots + x_{r,d_r} b_{r,d_r})} = 0$$

Similarly, applying $(A-\lambda_{r-1} I) \circ \cdots \circ (A-\lambda_2 I)$ to $(**)$, get

$$(\lambda_r - \lambda_{r-1}) \cdots (\lambda - \lambda_2)(\lambda - \lambda_1)(x_{r,1} b_{r,1} + \cdots + x_{r,d_r} b_{r,d_r}) = 0$$

hence $x_{r,1} = \cdots = x_{r,d_r} = 0$ as $b_{r,1}, \cdots, b_{r,d_r}$ form a basis of V_{λ_r}.

The equation $(*)$ then reduce to

$$x_{1,1} b_{1,1} + \cdots + x_{1,d_1} b_{1,d_1} + \cdots + x_{r-1,1} b_{r-1,1} + \cdots + x_{r-1,d_{r-1}} b_{r-1,d_{r-1}} = 0$$

Applying $(A - \lambda_{r-2} I) \circ \cdots \circ (A-\lambda_2 I) \circ (A - \lambda_1 I)$ to it, we can show $x_{r-1,1} = \cdots = x_{r-1,d_{r-1}} = 0$ as above.

Repeating the argument, we can show that all $x_{i,j} = 0$, hence the vectors are linearly independent. $\boxtimes$

- The equation $det(A - \lambda I) = 0$ may have roots λ which appears with multiplicity $m_\lambda > 1$. Let d_λ be the dimension of the eigenspace V_λ.
- Theorem: Suppose that all the roots of $det(A - \lambda I) = 0$ are real numbers, then $m_\lambda \geq d_\lambda$ for any root λ. The matrix A is diagonalizable if and only if $m_\lambda = d_\lambda$ for all the roots of the equation $det(A - \lambda I) = 0$.
- Corollary: If the roots of $det(A - \lambda I) = 0$ appear with multiplicity 1, then A must be diagonalizable.
- The corollary follows immediately from the theorem, as $1 = m_\lambda \geq d_\lambda \geq 1$ implies that $m_\lambda = d_\lambda = 1$.

Proof Let μ be a root of $\det(A - \lambda I) = 0$.

By definition, $d_\mu = \dim(\text{Nul}(\underbrace{A - \mu I}))$, Let $\vec{v}_1, \cdots \vec{v}_{d_\mu}$ be

V_μ

a basis of V_μ. Then A acts on them as

$$(*) \qquad A\vec{v}_i = \mu \vec{v}_i, \quad i = 1, \cdots, d_\mu.$$

By a previous theorem, the set $\{\vec{v}_1, \cdots, \vec{v}_{d_\mu}\}$ can be extended to a basis $\{\vec{v}_1, \cdots, \vec{v}_{d_\mu}, \vec{v}_{d_\mu + 1}, \cdots, \vec{v}_n\}$ of V.

By (*), with respect to this basis, A is now of the form

$$\begin{array}{cc} & \\ d_\mu & \left(\begin{array}{ccc|ccc} \mu & & & \times & \cdots & \times \\ & \ddots & & \vdots & & \vdots \\ & & \mu & \times & \cdots & \times \\ \hline & & & \times & \cdots & \times \\ & & & \vdots & & \vdots \\ & & & \times & \cdots & \times \end{array}\right) \\ & \qquad\qquad\qquad C \end{array}$$

Hence

$$\det(A - \lambda I) = \det(\mu I_{d_\mu} - \lambda I_{d_\mu}) \cdot \det(C - \lambda I_{n-d_\mu})$$
$$= (\mu - \lambda)^{d_\mu} \cdot \det(C - \lambda I_{n-d_\mu}).$$

$\Rightarrow (\mu - \lambda)^{d_\mu} \mid \det(A - \lambda I)$ as polynomials in λ $\Rightarrow m_\mu \geq d_\mu$.

For the second assertion, the necessity is clear. Indeed, if A is conjugate to $\begin{bmatrix} \lambda_1 & & & & \\ & \ddots \lambda_1 & & & \\ & & \ddots & & \\ & & & \lambda_r & \\ & & & & \ddots \lambda_r \end{bmatrix}$, with λ_i appearing n_i times,

then the eigenspace with eigenvalue λ_i is of dim n_i, and that

$$\det(A - \lambda I) = \begin{vmatrix} \lambda_1 - \lambda & & & & \\ & \ddots \lambda_1 - \lambda & & & \\ & & \ddots & & \\ & & & \lambda_r - \lambda & \\ & & & & \ddots \lambda_r - \lambda \end{vmatrix} = \prod_{i=1}^{r} (\lambda_i - \lambda)^{n_i}$$

Hence $m_{\lambda_i} = d_{\lambda_i} = n_i$, for $i = 1, \cdots, r$.

For the sufficiency, let $\lambda_1, \cdots, \lambda_r$ be the roots of $\det(A-\lambda I) = 0$, let V_{λ_i} be the corresponding eigenspace, let $\{v_{i,1}, \cdots, v_{i,d_{\lambda_i}}\}$ be a basis of V_{λ_i}. According to a previous proposition, the union $\mathcal{B} = \{v_{1,1}, \cdots, v_{1,d_{\lambda_1}}; \cdots; v_{r,1}, \cdots, v_{r,d_{\lambda_r}}\}$ remains linearly independent. On the other hand, $d_{\lambda_i} = m_{\lambda_i}$ and we have

$$\underbrace{m_{\lambda_1} + \cdots + m_{\lambda_r}}_{= d_{\lambda_1} + \cdots + d_{\lambda_r}} = \deg(\det(A - \lambda I)) = n.$$

So $\mathcal{B}$ forms a basis of $\mathbb{R}^n$. In other word, $\mathbb{R}^n$ admits a basis consisting of eigenvalues of A, hence A must be diagonalisable. $\boxed{\checkmark}$

- These results are summarized in the book as the theorem:

Let A be an $n \times n$ matrix whose distinct eigenvalues are $\lambda_1, \ldots, \lambda_p$.

a. For $1 \le k \le p$, the dimension of the eigenspace for λ_k is less than or equal to the multiplicity of the eigenvalue λ_k.

b. The matrix A is diagonalizable if and only if the sum of the dimensions of the eigenspaces equals n, and this happens if and only if (*i*) the characteristic polynomial factors completely into linear factors and (*ii*) the dimension of the eigenspace for each λ_k equals the multiplicity of λ_k.

c. If A is diagonalizable and $\mathcal{B}_k$ is a basis for the eigenspace corresponding to λ_k for each k, then the total collection of vectors in the sets $\mathcal{B}_1, \ldots, \mathcal{B}_p$ forms an eigenvector basis for $\mathbb{R}^n$.

- Corollary

An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

EXAMPLE 3 Diagonalize the following matrix, if possible.

$$A = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix}$$

That is, find an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$.

Solution 1) To begin with, we need to find the eigenvalues. The characteristic equation of A equals

$$\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 3 & 3 \\ -3 & -5-\lambda & -3 \\ 3 & 3 & 1-\lambda \end{vmatrix} = (1-\lambda)\begin{vmatrix} -5-\lambda & -3 \\ 3 & 1-\lambda \end{vmatrix} - 3\begin{vmatrix} -3 & -3 \\ 3 & 1-\lambda \end{vmatrix} + 3\begin{vmatrix} -3 & -5-\lambda \\ 3 & 3 \end{vmatrix}$$

$$= (1-\lambda)\left[(-5-\lambda)(1-\lambda) + 9\right] - 3\left[3(\lambda-1) + 9\right] + 3\left[-9 + 3(5+\lambda)\right]$$

$$= (1-\lambda)(\lambda^2 + 4\lambda + 4) - (9\lambda + 18) + 9\lambda + 18$$

$$= (1-\lambda)(\lambda+2)^2$$

So the eigenvalues are $\lambda = 1$, $\lambda = -2$.

2) We proceed to determine the eigenspaces, i.e. we need to solve the equation $(A - \lambda I)\vec{x} = 0$ for $\lambda = 1$ and $\lambda = -2$.

For $\lambda = 1$, we get the linear system

$$\begin{bmatrix} 0 & 3 & 3 \\ -3 & -6 & -3 \\ 3 & 3 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

performing row reduction, get

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

$$\Rightarrow \quad \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

x_3 is free variable

$$\Rightarrow \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -x_3 \\ x_3 \\ x_3 \end{pmatrix} = x_3 \cdot \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

↑ eigenvector of eigenvalue 1.

For the eigenvalue $\lambda = -2$, we get the linear system

$$\begin{pmatrix} 3 & 3 & 3 \\ -3 & -3 & -3 \\ 3 & 3 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0,$$

reduced to

$$x_1 + x_2 + x_3 = 0.$$

$$\Rightarrow \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -x_2 - x_3 \\ x_2 \\ x_3 \end{pmatrix} = x_2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

So the eigenspace V_{-2} is of dim 2, with basis $\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$.

3) We can conclude that $\mathbb{R}^3$ has a basis consisting of eigenvectors of A, hence A is diagonalisable.

Let $P = \begin{pmatrix} -1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$, then $AP = P \cdot \begin{pmatrix} 1 & & \\ & -2 & \\ & & -2 \end{pmatrix}$,

eigenvectors   eigenvalues

so $A = P \begin{pmatrix} 1 & & \\ & -2 & \\ & & -2 \end{pmatrix} P^{-1}$.

EXAMPLE 4 Diagonalize the following matrix, if possible.

$$A = \begin{bmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{bmatrix}$$

Solution 1) We begin by determining the eigenvalues of A.

The characteristic equation of A equals

$$\det(A - \lambda I) = \begin{vmatrix} 2-\lambda & 4 & 3 \\ -4 & -6-\lambda & -3 \\ 3 & 3 & 1-\lambda \end{vmatrix}$$

$$= (2-\lambda)\begin{vmatrix} -6-\lambda & -3 \\ 3 & 1-\lambda \end{vmatrix} - 4 \begin{vmatrix} -4 & -3 \\ 3 & 1-\lambda \end{vmatrix} + 3 \begin{vmatrix} -4 & -6-\lambda \\ 3 & 3 \end{vmatrix}$$

$$= (2-\lambda)\left[(-6-\lambda)(1-\lambda)+9\right] - 4(4\lambda-4+9) + 3(3\lambda+18-12)$$

$$= (2-\lambda)(\lambda^2+5\lambda+3) - 16\lambda - 20 + 9\lambda + 18$$

$$= -\lambda^3 - 5\lambda^2 - 3\lambda + 2\lambda^2 + 10\lambda + 6 - 7\lambda - 2$$

$$= -\lambda^3 - 3\lambda^2 + 4$$

$$= (1-\lambda)(\lambda+2)^2$$

2) We proceed to determine the eigenspace

For $\lambda = -2$, we need to solve the linear system

$$\begin{pmatrix} 4 & 4 & 3 \\ -4 & -4 & -3 \\ 3 & 3 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

reduced to $$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

$$\Rightarrow \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

x_2 is the free variable

$$\Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -x_2 \\ x_2 \\ 0 \end{pmatrix} = x_2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

The dimension of the eigenspace V_{-2} is of dim $1 <$ multiplicity m_{-2} of the eigenvalue, so A can not be diagonalized.