Eigenvalue Problems

Chia-Ping Chen

Department of Computer Science and Engineering
National Sun Yat-sen University

Linear Algebra

Outline

- Eigenvalues and eigenvectors
- Eigen-decomposition and diagonalization
- Difference equation
- Differential equation
- Complex matrices
- Similar matrices
- Change of basis

NOTATION

- $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$: eigenvalue equation of matrix $\boldsymbol{A}$
- $|\boldsymbol{A} - \lambda\boldsymbol{I}| = 0$: characteristic equation
- λ_i: eigenvalue
- $\{\lambda_1, \ldots, \lambda_k\}$: spectrum (the set of eigenvalues)
- $\boldsymbol{s}_i$: eigenvector corresponding to λ_i
- $\mathbb{E}_{\lambda_i}$: eigenspace corresponding to eigenvalue λ_i
- c^*: complex conjugate of c
- $\boldsymbol{A}^H$: Hermitian of $\boldsymbol{A}$
- $\boldsymbol{A} \sim \boldsymbol{B}$: similar matrices

Eigenvalue Equations

Definition (eigenvalue problem)

Let $\boldsymbol{A}$ be a square matrix.

- The eigenvalue equation or eigenvalue problem of $\boldsymbol{A}$ is

$$\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x}$$

- $\boldsymbol{x}$ is unknown vector, λ is unknown scalar
- Solutions of λ and $\boldsymbol{x} \neq \mathbf{0}$ are called eigenvalues and eigenvectors

Definition (characteristic equation/polynomial)

Let $\boldsymbol{A}$ be a square matrix.

- The eigenvalue equation $\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x}$ can be written as

$$(\boldsymbol{A} - \lambda \boldsymbol{I})\,\boldsymbol{x} = \boldsymbol{0}$$

- The characteristic equation of $\boldsymbol{A}$ is

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = 0$$

- The characteristic polynomial of $\boldsymbol{A}$ is

$$f(\lambda) = |\boldsymbol{A} - \lambda \boldsymbol{I}|$$

Lemma (eigenvalue and characteristic equation)

Let $\boldsymbol{A}$ be a square matrix. A scalar λ_i is an eigenvalue of $\boldsymbol{A}$ if and only if $|\boldsymbol{A} - \lambda_i \boldsymbol{I}| = 0$.

$$
\begin{aligned}
|\boldsymbol{A} - \lambda_i \boldsymbol{I}| = 0 &\Leftrightarrow (\boldsymbol{A} - \lambda_i \boldsymbol{I}) \text{ is singular} \\
&\Leftrightarrow \exists\, \boldsymbol{x} \neq \boldsymbol{0},\ (\boldsymbol{A} - \lambda_i \boldsymbol{I})\, \boldsymbol{x} = \boldsymbol{0} \\
&\Leftrightarrow \exists\, \boldsymbol{x} \neq \boldsymbol{0},\ \boldsymbol{A}\boldsymbol{x} = \lambda_i \boldsymbol{x} \\
&\Leftrightarrow \lambda_i \text{ is an eigenvalue of } \boldsymbol{A}
\end{aligned}
$$

Definition (eigenspace)

Let $\boldsymbol{A}$ be a square matrix and λ_i be an eigenvalue of $\boldsymbol{A}$. The eigenspace of $\boldsymbol{A}$ corresponding to λ_i is the set of vectors

$$\mathbb{E}_{\lambda_i} = \{\boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} = \lambda_i \boldsymbol{x}\}$$

- $\mathbb{E}_{\lambda_i}$ is the nullspace of $(\boldsymbol{A} - \lambda_i \boldsymbol{I})$

$$\begin{aligned}\mathbb{E}_{\lambda_i} &= \{\boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} = \lambda_i \boldsymbol{x}\} = \{\boldsymbol{x} \mid (\boldsymbol{A} - \lambda_i \boldsymbol{I})\,\boldsymbol{x} = \boldsymbol{0}\} \\ &= \mathcal{N}(\boldsymbol{A} - \lambda_i \boldsymbol{I})\end{aligned}$$

- If $\boldsymbol{A}$ is of order $n \times n$, $\mathbb{E}_{\lambda_i}$ is of dimension $n - \textbf{rank}(\boldsymbol{A} - \lambda_i \boldsymbol{I})$

$$\textbf{dim}(\mathbb{E}_{\lambda_i}) = \textbf{dim}(\mathcal{N}(\boldsymbol{A} - \lambda_i \boldsymbol{I})) = n - \textbf{rank}(\boldsymbol{A} - \lambda_i \boldsymbol{I})$$

Definition (spectrum)

Let $\boldsymbol{A}$ be a square matrix. The spectrum of $\boldsymbol{A}$ is the set of the eigenvalues of $\boldsymbol{A}$.

- The spectrum of $\boldsymbol{A}$ is the set of solutions to $|\boldsymbol{A}-\lambda\boldsymbol{I}|=0$
- Let $\boldsymbol{A}$ be a square matrix of order $n\times n$. Then $|\boldsymbol{A}-\lambda\boldsymbol{I}|$ is a polynomial of order n, and $|\boldsymbol{A}-\lambda\boldsymbol{I}|=0$ is an equation of order n.
- By the fundamental theorem of algebra, we have

$$\text{spectrum}(\boldsymbol{A})=\{\lambda_1,\ldots,\lambda_k\}$$

and

$$k\leq n$$

Lemma (eigenvectors of distinct eigenvalues)

Let $\boldsymbol{A}$ be a square matrix. Let $\boldsymbol{s}_1$ (resp. $\boldsymbol{s}_2$) be an eigenvector of $\boldsymbol{A}$ with eigenvalue λ_1 (resp. λ_2), where $\lambda_1 \neq \lambda_2$. Then $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ are linearly independent.

Suppose $c_1\boldsymbol{s}_1 + c_2\boldsymbol{s}_2 = \mathbf{0}$. Then

$$\lambda_1(c_1\boldsymbol{s}_1 + c_2\boldsymbol{s}_2) = \mathbf{0}$$

and

$$\boldsymbol{A}(c_1\boldsymbol{s}_1 + c_2\boldsymbol{s}_2) = c_1\lambda_1\boldsymbol{s}_1 + c_2\lambda_2\boldsymbol{s}_2 = \mathbf{0}$$

By subtraction, we have $c_2(\lambda_2 - \lambda_1)\boldsymbol{s}_2 = \mathbf{0}$. It follows that $c_2 = 0$ and $c_1 = 0$. Hence $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ are linearly independent.

Example (eigenvalue problem)

$$\boldsymbol{A} = \begin{bmatrix} 4 & -5 \\ 2 & -3 \end{bmatrix}$$

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = 0 \ \Rightarrow \ \lambda^2 - \lambda - 2 = 0 \ \Rightarrow \ \lambda_1 = 2, \, \lambda_2 = -1$$

For $\lambda_1 = 2$, the eigenspace is

$$\begin{bmatrix} 2 & -5 \\ 2 & -5 \end{bmatrix} \boldsymbol{s}_1 = \boldsymbol{0} \ \Rightarrow \ \mathbb{E}_{\lambda_1} = \left\{ \boldsymbol{s}_1 \,\middle|\, \boldsymbol{s}_1 = c \begin{bmatrix} \frac{5}{2} \\ 1 \end{bmatrix} \right\}$$

For $\lambda_2 = -1$, the eigenspace is

$$\begin{bmatrix} 5 & -5 \\ 2 & -2 \end{bmatrix} \boldsymbol{s}_2 = \boldsymbol{0} \ \Rightarrow \ \mathbb{E}_{\lambda_2} = \left\{ \boldsymbol{s}_2 \,\middle|\, \boldsymbol{s}_2 = c \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

Example (eigenvalue problem: projection matrix)

$$\boldsymbol{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

$$|\boldsymbol{P} - \lambda \boldsymbol{I}| = 0 \;\Rightarrow\; \lambda^2 - \lambda = 0 \;\Rightarrow\; \lambda_1 = 1,\, \lambda_2 = 0$$

$$(\boldsymbol{P} - \lambda_1 \boldsymbol{I})\boldsymbol{s}_1 = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \boldsymbol{s}_1 = \boldsymbol{0} \Rightarrow \mathbb{E}_{\lambda_1} = \left\{ \boldsymbol{s}_1 \,\middle|\, \boldsymbol{s}_1 = c \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

$$(\boldsymbol{P} - \lambda_2 \boldsymbol{I})\boldsymbol{s}_2 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \boldsymbol{s}_2 = \boldsymbol{0} \;\Rightarrow\; \mathbb{E}_{\lambda_2} = \left\{ \boldsymbol{s}_2 \,\middle|\, \boldsymbol{s}_2 = c \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$$

Example (complex eigenvalues)

$$\boldsymbol{K} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

$$|\boldsymbol{K} - \lambda \boldsymbol{I}| = 0 \;\Rightarrow\; \lambda^2 + 1 = 0 \;\Rightarrow\; \lambda_1 = i,\, \lambda_2 = -i$$

$$(\boldsymbol{K} - \lambda_1 \boldsymbol{I})\boldsymbol{s}_1 = \begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix} \boldsymbol{s}_1 = \boldsymbol{0} \Rightarrow \mathbb{E}_{\lambda_1} = \left\{ \boldsymbol{s}_1 \,\middle|\, \boldsymbol{s}_1 = c \begin{bmatrix} i \\ 1 \end{bmatrix} \right\}$$

$$(\boldsymbol{K} - \lambda_2 \boldsymbol{I})\boldsymbol{s}_2 = \begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix} \boldsymbol{s}_2 = \boldsymbol{0} \Rightarrow \mathbb{E}_{\lambda_2} = \left\{ \boldsymbol{s}_2 \,\middle|\, \boldsymbol{s}_2 = c \begin{bmatrix} -i \\ 1 \end{bmatrix} \right\}$$

Lemma (bound on the number of eigenvalues)

Let $\boldsymbol{A}$ be a square matrix of order $n \times n$. $\boldsymbol{A}$ has at most n distinct eigenvalues.

- A polynomial of order n cannot have more than n roots

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = \begin{vmatrix} a_{11} - \lambda & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} - \lambda \end{vmatrix}$$

- A space of dimension n cannot accommodate more than n linearly independent eigenvectors

Theorem (sum of eigenvalues = trace)

Let $\boldsymbol{A}$ be a square matrix.

- *The trace of $\boldsymbol{A}$ is the sum of diagonal elements*
- *Sum of the eigenvalues of $\boldsymbol{A}$ equals the trace of $\boldsymbol{A}$*

Factorize the polynomial $|\boldsymbol{A} - \lambda \boldsymbol{I}|$ by its roots

$$\begin{vmatrix} a_{11} - \lambda & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} - \lambda \end{vmatrix} = c(\lambda - \lambda_1) \dots (\lambda - \lambda_n)$$

For the λ^n and λ^{n-1} terms, the equality requires $c = (-1)^n$ and

$$(-1)^{n-1}(a_{11}+\cdots+a_{nn})\lambda^{n-1} = (-1)^n((-\lambda_1)+\cdots+(-\lambda_n))\lambda^{n-1}$$

$$\lambda_1 + \cdots + \lambda_n = a_{11} + \cdots + a_{nn}$$

Theorem (product of eigenvalues = determinant)

Let $\boldsymbol{A}$ be a square matrix. Product of the eigenvalues of $\boldsymbol{A}$ is equal to the determinant of $\boldsymbol{A}$.

The characteristic polynomial of $\boldsymbol{A}$ can be factorized by its roots

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = (-1)^n \prod_{i=1}^{n} (\lambda - \lambda_i)$$

Setting λ of both sides to 0, we get

$$|\boldsymbol{A}| = (-1)^n \prod_{i=1}^{n} (-\lambda_i) = \prod_{i=1}^{n} \lambda_i$$

The equality holds for repeated eigenvalues.

TIPS FOR FINDING EIGENVALUES

Let $\boldsymbol{A}$ be a square matrix. Eigenvalues of $\boldsymbol{A}$ may be found without solving the characteristic equation of $\boldsymbol{A}$.

- If $\boldsymbol{A}$ is singular, 0 is an eigenvalue
- If $\boldsymbol{A}$ has a constant row sum (or column sum), that constant is an eigenvalue
- If $\boldsymbol{A}$ is a triangular matrix, the diagonal elements of $\boldsymbol{A}$ are eigenvalues

Eigen-decomposition and Diagonalization

Definition (algebraic/geometric multiplicity)

Let $\boldsymbol{A}$ be a square matrix with spectrum $\{\lambda_1, \ldots, \lambda_k\}$.

- The characteristic polynomial of $\boldsymbol{A}$ can be expressed as

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = c \prod_{i=1}^{k} (\lambda - \lambda_i)^{\gamma_i}$$

- γ_i is called the algebraic multiplicity of λ_i
- The dimension of eigenspace $\mathbb{E}_{\lambda_i}$ is called the geometric multiplicity of λ_i, denoted by g_i

Suppose $\boldsymbol{A}$ is of order $n \times n$.

- The sum of algebraic multiplicities $\sum_i \gamma_i$ is exactly n
- The sum of geometric multiplicities $\sum_i g_i$ is at most n

Definition (defective matrix)

Let $\boldsymbol{A}$ be a square matrix of order $n \times n$ with spectrum $\{\lambda_1, \ldots, \lambda_k\}$. $\boldsymbol{A}$ is **defective** if

$$g_1 + \cdots + g_k < n$$

Consider

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \ \boldsymbol{B} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$\boldsymbol{A}$ is defective

$$n = 2, \lambda_1 = 0, \sum_i g_i = g_1 = \overbrace{n - \textbf{rank}(\boldsymbol{A} - \lambda_1 \boldsymbol{I})}^{\textbf{dim}\,\mathcal{N}(\boldsymbol{A} - \lambda_1 \boldsymbol{I})} = 2-1 = 1 < n$$

$\boldsymbol{B}$ is non-defective

$$n = 2, \lambda_1 = 0, \sum_i g_i = g_1 = n - \textbf{rank}(\boldsymbol{B} - \lambda_1 \boldsymbol{I}) = 2-0 = 2 = n$$

Definition (eigenbasis)

An **eigenbasis** of $\boldsymbol{A}$ is a basis containing eigenvectors of $\boldsymbol{A}$.

For a non-defective matrix $\boldsymbol{A}$ of order $n \times n$, we can construct an eigenbasis as follows.

- Let $\{\lambda_1, \ldots, \lambda_k\}$ be the spectrum of $\boldsymbol{A}$
- Let $\mathcal{B}_1, \ldots, \mathcal{B}_k$ be bases of eigenspaces $\mathbb{E}_{\lambda_1}, \ldots, \mathbb{E}_{\lambda_k}$
- Let $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$
- $\mathcal{B}$ is an eigenbasis: it is linearly independent and contains $\sum_{i=1}^{k} g_i = n$ eigenvectors of $\boldsymbol{A}$

Definition (eigenvector and eigenvalue matrix)

Let $\boldsymbol{A}$ be a non-defective matrix of order $n \times n$.

- From eigenbasis $\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n\}$ of $\boldsymbol{A}$, we can construct eigenvector matrix

$$\boldsymbol{S} = \begin{bmatrix} \boldsymbol{s}_1 & \ldots & \boldsymbol{s}_n \end{bmatrix}$$

- Let λ_i be the eigenvalue corresponding to $\boldsymbol{s}_i$. We can construct eigenvalue matrix

$$\boldsymbol{\Lambda} = \textbf{diag}(\lambda_1, \ldots, \lambda_n)$$

Theorem (eigen-decomposition)

Let $\boldsymbol{A}$ be a non-defective matrix. $\boldsymbol{A}$ can be decomposed by

$$\boldsymbol{A} = \boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^{-1}$$

This is eigenvalue decomposition or simply eigen-decomposition.

$$\begin{aligned}\boldsymbol{A}\begin{bmatrix}\boldsymbol{s}_1 & \dots & \boldsymbol{s}_n\end{bmatrix} &= \begin{bmatrix}\boldsymbol{A}\boldsymbol{s}_1 & \dots & \boldsymbol{A}\boldsymbol{s}_n\end{bmatrix} = \begin{bmatrix}\lambda_1\boldsymbol{s}_1 & \dots & \lambda_n\boldsymbol{s}_n\end{bmatrix} \\ &= \begin{bmatrix}\boldsymbol{s}_1 & \dots & \boldsymbol{s}_n\end{bmatrix}\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}\end{aligned}$$

It follows from $\boldsymbol{A}\boldsymbol{S} = \boldsymbol{S}\boldsymbol{\Lambda}$ that $\boldsymbol{A} = \boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^{-1}$.

Corollary (diagonalization of a matrix)

A non-defective square matrix can be diagonalized by its eigenvector matrix.

It follows from eigen-decomposition $\boldsymbol{A} = \boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^{-1}$ that

$$\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \boldsymbol{\Lambda} = \textbf{diag}(\lambda_1, \ldots, \lambda_n)$$

This is the **diagonalization** of $\boldsymbol{A}$.

Example (diagonalization of matrix)

$$\boldsymbol{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \quad \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\boldsymbol{K} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$$

Difference Equations (with an Eigenvalue Approach)

Definition (Fibonacci recurrence and numbers)

- Fibonacci recurrence

$$F_{k+1} = F_k + F_{k-1}$$

- Initial Fibonacci numbers

$$F_0 = 0, \quad F_1 = 1$$

- Fibonacci sequence

$$0 \quad 1 \quad 1 \quad 2 \quad 3 \quad 5 \quad 8 \quad 13 \quad \ldots$$

Definition (Fibonacci vectors and matrix)

- Fibonacci vectors

$$\boldsymbol{u}_k = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}$$

- Fibonacci matrix

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

k-step recurrence

- Fibonacci recurrence by matrix and vector

$$\begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_k \\ F_{k-1} \end{bmatrix}$$

That is

$$\boldsymbol{u}_k = \boldsymbol{A}\boldsymbol{u}_{k-1}$$

- k-step recurrence

$$\boldsymbol{u}_k = \boldsymbol{A}\boldsymbol{u}_{k-1} = \boldsymbol{A}(\boldsymbol{A}\boldsymbol{u}_{k-2}) = \cdots = \boldsymbol{A}^{k-1}\boldsymbol{u}_1 = \boldsymbol{A}^k\boldsymbol{u}_0$$

That is

$$\begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^k \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Power of a non-defective matrix

Let $\boldsymbol{A}$ be a non-defective matrix with eigen-decomposition $\boldsymbol{A} = \boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^{-1}$. Then

$$\boldsymbol{A}^k = \boldsymbol{S}\boldsymbol{\Lambda}^k\boldsymbol{S}^{-1}$$

$$\begin{aligned}
\boldsymbol{A}^k &= (\boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^{-1})^k \\
&= (\boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^{-1})(\boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^{-1})\dots(\boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^{-1}) \\
&= \boldsymbol{S}\boldsymbol{\Lambda}(\boldsymbol{S}^{-1}\boldsymbol{S})\boldsymbol{\Lambda}(\boldsymbol{S}^{-1}\boldsymbol{S})\dots(\boldsymbol{S}^{-1}\boldsymbol{S})\boldsymbol{\Lambda}\boldsymbol{S}^{-1} \\
&= \boldsymbol{S}\boldsymbol{\Lambda}^k\boldsymbol{S}^{-1}
\end{aligned}$$

Note $\boldsymbol{S}\boldsymbol{\Lambda}^k\boldsymbol{S}^{-1}$ is easier to compute than $\boldsymbol{A}^k$.

formula for Fibonacci vectors

- Fibonacci matrix $\boldsymbol{A}$ is non-defective
- Simplification of k-step recurrence

$$\begin{aligned}\boldsymbol{u}_k &= \boldsymbol{A}^k \boldsymbol{u}_0 = \boldsymbol{S}\boldsymbol{\Lambda}^k \boldsymbol{S}^{-1}\boldsymbol{u}_0 \\ &= \begin{bmatrix} \boldsymbol{s}_1 & \boldsymbol{s}_2 \end{bmatrix} \begin{bmatrix} \lambda_1^k & \\ & \lambda_2^k \end{bmatrix} \begin{bmatrix} \boldsymbol{s}_1 & \boldsymbol{s}_2 \end{bmatrix}^{-1} \boldsymbol{u}_0 \\ &= c_1\lambda_1^k \boldsymbol{s}_1 + c_2\lambda_2^k \boldsymbol{s}_2\end{aligned}$$

where

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{s}_1 & \boldsymbol{s}_2 \end{bmatrix}^{-1} \boldsymbol{u}_0$$

- We have $\boldsymbol{u}_k$ expressed by λ_i and $\boldsymbol{s}_i$

Eigenvalue problem of Fibonacci matrix

Recall $\boldsymbol{A} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$.

- Eigenvalues: solve $|\boldsymbol{A} - \lambda \boldsymbol{I}| = 0$

$$\lambda_1 = \frac{1+\sqrt{5}}{2}, \ \lambda_2 = \frac{1-\sqrt{5}}{2}$$

- Eigenvectors: solve $\boldsymbol{A}\boldsymbol{s}_i = \lambda_i \boldsymbol{s}_i$

$$(\boldsymbol{A} - \lambda_i \boldsymbol{I})\boldsymbol{s}_i = \begin{bmatrix} 1-\lambda_i & 1 \\ 1 & -\lambda_i \end{bmatrix} \boldsymbol{s}_i = \boldsymbol{0}$$

$$\Rightarrow \ \boldsymbol{s}_i = \begin{bmatrix} \lambda_i \\ 1 \end{bmatrix}, \ i = 1, 2$$

Exploiting the initial condition

With $\boldsymbol{u}_k = c_1\lambda_1^k\boldsymbol{s}_1 + c_2\lambda_2^k\boldsymbol{s}_2$, we still need to decide c_1 and c_2.

- Initial condition

$$\boldsymbol{u}_0 = c_1\lambda_1^0\boldsymbol{s}_1 + c_2\lambda_2^0\boldsymbol{s}_2 = c_1\boldsymbol{s}_1 + c_2\boldsymbol{s}_2$$

- Substitution of $\boldsymbol{s}_1, \boldsymbol{s}_2$ and $\boldsymbol{u}_0$

$$\begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

- Solve c_1 and c_2

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{\lambda_1 - \lambda_2} \\ -\frac{1}{\lambda_1 - \lambda_2} \end{bmatrix}$$

FORMULA FOR THE FIBONACCI NUMBERS

- kth Fibonacci vector

$$\begin{aligned}\boldsymbol{u}_k &= c_1\lambda_1^k\boldsymbol{s}_1 + c_2\lambda_2^k\boldsymbol{s}_2\\ &= \frac{1}{\lambda_1-\lambda_2}\lambda_1^k\begin{bmatrix}\lambda_1\\1\end{bmatrix} - \frac{1}{\lambda_1-\lambda_2}\lambda_2^k\begin{bmatrix}\lambda_2\\1\end{bmatrix}\\ &= \frac{1}{\lambda_1-\lambda_2}\begin{bmatrix}\lambda_1^{k+1}-\lambda_2^{k+1}\\ \lambda_1^k-\lambda_2^k\end{bmatrix} = \begin{bmatrix}F_{k+1}\\F_k\end{bmatrix}\end{aligned}$$

- kth Fibonacci number

$$\begin{aligned}F_k &= \frac{1}{\lambda_1-\lambda_2}\left(\lambda_1^k-\lambda_2^k\right)\\ &= \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^k - \left(\frac{1-\sqrt{5}}{2}\right)^k\right]\end{aligned}$$

Example (a Markov process)

Suppose $\frac{1}{10}$ of the population outside Asia move in and $\frac{2}{10}$ of the population inside Asia move out every year. What is the inside-Asia/outside-Asia population at the end of year k?

Let y_k (resp. z_k) be the population outside (resp. inside) Asia at the end of year k.

- Recurrence of population

$$\begin{aligned} y_{k+1} &= 0.9y_k + 0.2z_k \\ z_{k+1} &= 0.1y_k + 0.8z_k \end{aligned}$$

- In vector and matrix

$$\begin{bmatrix} y_{k+1} \\ z_{k+1} \end{bmatrix} = \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix} \begin{bmatrix} y_k \\ z_k \end{bmatrix}$$

- Population vectors and matrix

$$\boldsymbol{x}_k = \begin{bmatrix} y_k \\ z_k \end{bmatrix}, \quad \boldsymbol{A} = \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix}$$

- Year-to-year evolution of population

$$\boldsymbol{x}_{k+1} = \boldsymbol{A}\boldsymbol{x}_k$$

- Population vector at the end of year k

$$\boldsymbol{x}_k = \boldsymbol{A}\boldsymbol{x}_{k-1} = \boldsymbol{A}(\boldsymbol{A}\boldsymbol{x}_{k-2}) = \cdots = \boldsymbol{A}^k\boldsymbol{x}_0$$

- Eigen-decomposition of $\boldsymbol{A}$

$$\boldsymbol{A} = \boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^{-1} = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0.7 \end{bmatrix} \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

- Put the pieces together

$$\boldsymbol{x}_k = \boldsymbol{A}^k \boldsymbol{x}_0 = \boldsymbol{S}\boldsymbol{\Lambda}^k \boldsymbol{S}^{-1} \boldsymbol{x}_0$$

$$\begin{aligned}
\Rightarrow \quad \begin{bmatrix} y_k \\ z_k \end{bmatrix} &= \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0.7 \end{bmatrix}^k \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} y_0 \\ z_0 \end{bmatrix} \\
&= \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{3}(y_0 + z_0)(1)^k \\ -\frac{1}{3}(y_0 - 2z_0)(0.7)^k \end{bmatrix} \\
&= (y_0 + z_0)(1)^k \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix} + (y_0 - 2z_0)(0.7)^k \begin{bmatrix} \frac{1}{3} \\ -\frac{1}{3} \end{bmatrix}
\end{aligned}$$

Stationary Eigenvector

Consider population matrix $\boldsymbol{A}$.

- Non-negative
- Every column sums to 1
- Has eigenvalue 1
- Stationary eigenvector: let $\boldsymbol{\pi}$ be an eigenvector of $\boldsymbol{A}$ with eigenvalue 1

$$\boldsymbol{A}\boldsymbol{\pi} = 1 \cdot \boldsymbol{\pi} = \boldsymbol{\pi}$$

$$\boldsymbol{x}_0 = \boldsymbol{\pi} \Rightarrow \boldsymbol{x}_1 = \boldsymbol{A}\boldsymbol{x}_0 = \boldsymbol{\pi} \Rightarrow \cdots \Rightarrow \cdots \Rightarrow \boldsymbol{x}_n = \boldsymbol{\pi}$$

Linear Algebra and Differential Equations

Linear Differential Equation

Consider a linear differential equation

$$\frac{du}{dt} = au$$

where $u = u(t)$ is an unknown function of t and a is a constant.

- The equation is linear
- Let the initial condition be $u(0) = u_0$. The solution is

$$u(t) = u_0 e^{at} = e^{at} u_0$$

SYSTEM OF LINEAR DIFFERENTIAL EQUATIONS

Consider a system of linear differential equations

$$\begin{cases} \dfrac{dv}{dt} = a\,v + b\,w \\ \dfrac{dw}{dt} = c\,v + d\,w \end{cases}$$

Define $\boldsymbol{u} = \begin{bmatrix} v \\ w \end{bmatrix}$ and $\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. The system can be written as

$$\begin{aligned} \frac{d\boldsymbol{u}}{dt} &= \frac{d}{dt}\begin{bmatrix} v \\ w \end{bmatrix} = \begin{bmatrix} \frac{dv}{dt} \\ \frac{dw}{dt} \end{bmatrix} = \begin{bmatrix} a\,v + b\,w \\ c\,v + d\,w \end{bmatrix} \\ &= \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} v \\ w \end{bmatrix} \\ &= \boldsymbol{A}\boldsymbol{u} \end{aligned}$$

Decomposition with Eigenbasis

Let $\boldsymbol{A}$ be a non-defective $n \times n$ matrix. Consider

$$\frac{d\boldsymbol{u}}{dt} = \boldsymbol{A}\boldsymbol{u}$$

- $\boldsymbol{A}$ has an eigenbasis, say $\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n\}$
- A solution, say $\boldsymbol{u}$, can be expressed as

$$\boldsymbol{u}(t) = c_1(t)\boldsymbol{s}_1 + \cdots + c_n(t)\boldsymbol{s}_n$$

- $\boldsymbol{u}(t)$ varies with time and $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$ are time-invariant, so the coefficients $c_1(t), \ldots, c_n(t)$ must vary with time

MODE

Let $\boldsymbol{A}$ be a non-defective $n \times n$ matrix. Consider

$$\frac{d\boldsymbol{u}}{dt} = \boldsymbol{A}\boldsymbol{u}$$

- A solution that aligns with an eigenvector of $\boldsymbol{A}$ is a mode
- Let $\boldsymbol{u}(t) = c(t)\boldsymbol{s}$ be a mode, where $\boldsymbol{s}$ is an eigenvector of $\boldsymbol{A}$ with eigenvalue λ

$$\begin{aligned}\frac{d\boldsymbol{u}}{dt} = \boldsymbol{A}\boldsymbol{u} \;&\Rightarrow\; \frac{d\left(c(t)\boldsymbol{s}\right)}{dt} = \boldsymbol{A}\left(c(t)\boldsymbol{s}\right) = \lambda\left(c(t)\boldsymbol{s}\right)\\ &\Rightarrow\; \frac{dc(t)}{dt} = \lambda c(t)\\ &\Rightarrow\; c(t) = e^{\lambda t}c(0)\end{aligned}$$

- So a mode is proportional to $e^{\lambda t}\boldsymbol{s}$

Mixture of modes

A general solution is a linear combination of modes.

- A general solution can be written as

$$\boldsymbol{u}(t) = c_1(t)\boldsymbol{s}_1 + \cdots + c_n(t)\boldsymbol{s}_n$$

- Substitute into the differential equation

$$\frac{d\boldsymbol{u}}{dt} = \boldsymbol{A}\boldsymbol{u} \;\Rightarrow\; \frac{d\left(\sum_i c_i(t)\boldsymbol{s}_i\right)}{dt} = \sum_i \lambda_i \left(c_i(t)\boldsymbol{s}_i\right)$$
$$\Rightarrow\; \sum_i \left(\frac{dc_i(t)}{dt} - \lambda_i c_i(t)\right)\boldsymbol{s}_i = \boldsymbol{0}$$

- Linear independence of $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$ requires

$$\frac{dc_i(t)}{dt} = \lambda_i c_i(t) \;\Rightarrow\; c_i(t) = e^{\lambda_i t} c_i(0)$$

Independence of modes

- We have

$$\begin{aligned} \boldsymbol{u}(t) &= \sum_{i=1}^{n} c_i(t)\boldsymbol{s}_i \\ &= c_1(0)e^{\lambda_1 t}\boldsymbol{s}_1 + \cdots + c_n(0)e^{\lambda_n t}\boldsymbol{s}_n \end{aligned}$$

- Each mode evolves with time exponentially and independently of the other modes.

- Eigenvector matrix of $\boldsymbol{A}$

$$\boldsymbol{S} = \begin{bmatrix} \boldsymbol{s}_1 & \dots & \boldsymbol{s}_n \end{bmatrix}$$

- Define

$$\boldsymbol{D}(t) = \textbf{diag}\left(e^{\lambda_1 t}, \dots, e^{\lambda_n t}\right), \; \boldsymbol{c}_0 = \begin{bmatrix} c_1(0) \\ \vdots \\ c_n(0) \end{bmatrix}$$

- We have

$$\begin{aligned} \boldsymbol{u}(t) &= c_1(0)e^{\lambda_1 t}\boldsymbol{s}_1 + \cdots + c_n(0)e^{\lambda_n t}\boldsymbol{s}_n \\ &= \boldsymbol{S}\boldsymbol{D}(t)\boldsymbol{c}_0 \end{aligned}$$

Theorem (solving linear differential equations)

Let $\boldsymbol{A}$ be a matrix with eigen-decomposition $\boldsymbol{A} = \boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^{-1}$. The solution to differential equation

$$\frac{d\boldsymbol{u}}{dt} = \boldsymbol{A}\boldsymbol{u}$$

with initial condition $\boldsymbol{u}(0) = \boldsymbol{u}_0$ is

$$\boldsymbol{u}(t) = \boldsymbol{S}\boldsymbol{D}(t)\boldsymbol{S}^{-1}\boldsymbol{u}_0$$

where $\boldsymbol{D}(t) = \textbf{diag}\left(e^{\lambda_1 t}, \ldots, e^{\lambda_n t}\right)$.

We have $\boldsymbol{u}(t) = \boldsymbol{S}\boldsymbol{D}(t)\boldsymbol{c}_0$. At $t = 0$

$$\boldsymbol{D}(0) = \boldsymbol{I} \;\Rightarrow\; \boldsymbol{u}_0 = \boldsymbol{S}\boldsymbol{c}_0 \;\Rightarrow\; \boldsymbol{c}_0 = \boldsymbol{S}^{-1}\boldsymbol{u}_0$$

Therefore

$$\boldsymbol{u}(t) = \boldsymbol{S}\boldsymbol{D}(t)\boldsymbol{S}^{-1}\boldsymbol{u}_0$$

Example (linear differential equation system)

Solve

$$\frac{d\boldsymbol{u}}{dt} = \boldsymbol{A}\boldsymbol{u}, \quad \boldsymbol{u} = \begin{bmatrix} v(t) \\ w(t) \end{bmatrix}, \quad \boldsymbol{A} = \begin{bmatrix} 4 & -5 \\ 2 & -3 \end{bmatrix}$$

with initial condition $v(0) = 8$ and $w(0) = 5$.

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = 0 \;\Rightarrow\; \lambda_1 = 2,\, \lambda_2 = -1 \;\Rightarrow\; \boldsymbol{s}_1 = \begin{bmatrix} \frac{5}{2} \\ 1 \end{bmatrix},\; \boldsymbol{s}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\boldsymbol{u} = \boldsymbol{S}\boldsymbol{D}\boldsymbol{S}^{-1}\boldsymbol{u}_0 = \begin{bmatrix} \frac{5}{2} & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} e^{2t} & 0 \\ 0 & e^{-t} \end{bmatrix} \begin{bmatrix} \frac{5}{2} & 1 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 8 \\ 5 \end{bmatrix}$$

$$= 2e^{2t} \begin{bmatrix} \frac{5}{2} \\ 1 \end{bmatrix} + 3e^{-t} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Definition (matrix as an exponent)

Let $\boldsymbol{A}$ be a square matrix. Define

$$e^{\boldsymbol{A}} \triangleq \boldsymbol{I} + \boldsymbol{A} + \frac{\boldsymbol{A}^2}{2!} + \frac{\boldsymbol{A}^3}{3!} + \dots$$

This is an extension of scalar exponential

$$e^a = 1 + a + \frac{a^2}{2!} + \frac{a^3}{3!} + \dots$$

A diagonal matrix as an exponent

Let $\boldsymbol{B} = \textbf{diag}(d_1, \ldots, d_n)$ be a diagonal matrix.

$$e^{\boldsymbol{B}} = \textbf{diag}\left(e^{d_1}, \ldots, e^{d_n}\right)$$

$$\begin{aligned}
e^{\boldsymbol{B}} &= \boldsymbol{I} + \boldsymbol{B} + \frac{\boldsymbol{B}^2}{2!} + \ldots \\
&= \textbf{diag}\,(1, \ldots, 1) + \textbf{diag}\,(d_1, \ldots, d_n) + \textbf{diag}\left(\frac{d_1^2}{2!}, \ldots, \frac{d_n^2}{2!}\right) + \ldots \\
&= \textbf{diag}\left(\left(1 + d_1 + \frac{d_1^2}{2!} + \ldots\right), \ldots, \left(1 + d_n + \frac{d_n^2}{2!} + \ldots\right)\right) \\
&= \textbf{diag}\left(e^{d_1}, \ldots, e^{d_n}\right)
\end{aligned}$$

A NON-DEFECTIVE MATRIX AS AN EXPONENT

Let $\boldsymbol{A}$ have eigen-decomposition $\boldsymbol{A} = \boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^{-1}$.

$$e^{\boldsymbol{A}} = \boldsymbol{S}e^{\boldsymbol{\Lambda}}\boldsymbol{S}^{-1}$$

$$\begin{aligned}
e^{\boldsymbol{A}} &= \boldsymbol{I} + \boldsymbol{A} + \frac{\boldsymbol{A}^2}{2!} + \dots \\
&= \boldsymbol{I} + \boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^{-1} + \frac{(\boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^{-1})^2}{2!} + \dots \\
&= \boldsymbol{S}\left(\boldsymbol{I} + \boldsymbol{\Lambda} + \frac{\boldsymbol{\Lambda}^2}{2!} + \dots\right)\boldsymbol{S}^{-1} \\
&= \boldsymbol{S}e^{\boldsymbol{\Lambda}}\boldsymbol{S}^{-1}
\end{aligned}$$

Theorem (matrix representation of solution)

Let $\boldsymbol{A}$ have eigen-decomposition $\boldsymbol{A} = \boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^{-1}$. The solution of a system of linear differential equations

$$\frac{d\boldsymbol{u}}{dt} = \boldsymbol{A}\boldsymbol{u}$$

is

$$\boldsymbol{u} = e^{\boldsymbol{A}t}\boldsymbol{u}_0$$

Proof.

The solution is

$$\boldsymbol{u} = \boldsymbol{S}\boldsymbol{D}\boldsymbol{S}^{-1}\boldsymbol{u}_0 = \boldsymbol{S}e^{\boldsymbol{\Lambda}t}\boldsymbol{S}^{-1}\boldsymbol{u}_0$$

We have

$$\begin{aligned}
\boldsymbol{S}e^{\boldsymbol{\Lambda}t}\boldsymbol{S}^{-1} &= \boldsymbol{S}\left(\boldsymbol{I} + \boldsymbol{\Lambda}t + \frac{\boldsymbol{\Lambda}^2t^2}{2!} + \dots\right)\boldsymbol{S}^{-1} \\
&= \boldsymbol{I} + (\boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^{-1})t + \frac{(\boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^{-1})^2t^2}{2!} + \dots \\
&= \boldsymbol{I} + \boldsymbol{A}t + \frac{\boldsymbol{A}^2t^2}{2!} + \dots \\
&= e^{\boldsymbol{A}t}
\end{aligned}$$

Hence $\boldsymbol{u} = e^{\boldsymbol{A}t}\boldsymbol{u}_0$. □

HIGH-ORDER LINEAR DIFFERENTIAL EQUATION*

A high-order linear differential equation can be converted to a system of first-order linear differential equations.

For example, consider a third-order linear differential equation

$$\frac{d^3y}{dt^3} + b\frac{d^2y}{dt^2} + c\frac{dy}{dt} = 0$$

Define

$$v = \frac{dy}{dt},\; w = \frac{dv}{dt},\; \boldsymbol{u} = \begin{bmatrix} y \\ v \\ w \end{bmatrix},\; \boldsymbol{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -c & -b \end{bmatrix}$$

Then

$$\frac{d\boldsymbol{u}}{dt} = \frac{d}{dt}\begin{bmatrix} y \\ v \\ w \end{bmatrix} = \begin{bmatrix} \frac{dy}{dt} \\ \frac{dv}{dt} \\ \frac{dw}{dt} \end{bmatrix} = \begin{bmatrix} v \\ w \\ -bw - cv \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -c & -b \end{bmatrix}\begin{bmatrix} y \\ v \\ w \end{bmatrix} = \boldsymbol{A}\boldsymbol{u}$$

LINEAR PARTIAL DIFFERENTIAL EQUATION*

A linear partial differential equation can be converted to a system of first-order linear differential equations.

Consider heat equation

$$\frac{\partial u(t,x)}{\partial t} = \frac{\partial^2 u(t,x)}{\partial x^2}$$

Discretizing x to n points, we have

$$\frac{d\boldsymbol{u}}{dt} = \boldsymbol{A}\boldsymbol{u}, \quad \boldsymbol{u} = \begin{bmatrix} u_1(t) \\ \cdot \\ \cdot \\ u_n(t) \end{bmatrix}, \ \boldsymbol{A} = \begin{bmatrix} -2 & 1 & & \\ 1 & -2 & \cdot & \\ & \cdot & \cdot & 1 \\ & & 1 & -2 \end{bmatrix}$$

where $u_i(t) = u(t, x_i)$.

Complex Matrix

COMPLEX NUMBERS

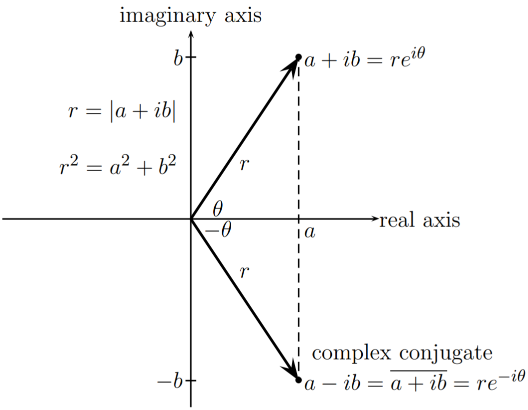

Figure 5.4: The complex plane, with $a + ib = re^{i\theta}$ and its conjugate $a - ib = re^{-i\theta}$.

Complex vector and complex matrix

- A vector with complex elements is a complex vector
- A matrix with complex elements is a complex matrix

Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be complex vectors of size n.

- **Inner product** of $\boldsymbol{x}$ and $\boldsymbol{y}$

$$(\boldsymbol{x}, \boldsymbol{y}) = x_1^* y_1 + \cdots + x_n^* y_n$$

- **Length** (or **norm**) of $\boldsymbol{x}$

$$\|\boldsymbol{x}\|^2 = (\boldsymbol{x}, \boldsymbol{x})$$

- **Orthogonality** of $\boldsymbol{x}$ and $\boldsymbol{y}$

$$(\boldsymbol{x}, \boldsymbol{y}) = 0 \quad \Leftrightarrow \quad \boldsymbol{x} \perp \boldsymbol{y}$$

Example (complex vectors)

Decide the inner product, lengths and orthogonality for

$$\boldsymbol{x} = \begin{bmatrix} 3-2i \\ 2+i \end{bmatrix}, \ \boldsymbol{y} = \begin{bmatrix} 5 \\ -1-i \end{bmatrix}$$

Definition (the Hermitian of a matrix)

Let $\boldsymbol{A}$ be a complex matrix. The **Hermitian** of $\boldsymbol{A}$ is

$$\boldsymbol{A}^H = (\boldsymbol{A}^*)^T = \left(\boldsymbol{A}^T\right)^*$$

- Relationship between elements

$$a_{ij}^H = a_{ji}^*$$

- Hermitian of Hermitian

$$\left(\boldsymbol{A}^H\right)^H = \boldsymbol{A}$$

- Hermitian of product

$$(\boldsymbol{AB})^H = \boldsymbol{B}^H \boldsymbol{A}^H$$

Lemma (inner product and Hermitian)

Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be complex vectors.

- *The inner product of $\boldsymbol{x}$ and $\boldsymbol{y}$ is*

$$(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}^H \boldsymbol{y}$$

- *Furthermore*

$$(\boldsymbol{x}, \boldsymbol{A}\boldsymbol{y}) = \boldsymbol{x}^H \boldsymbol{A}\boldsymbol{y} = (\boldsymbol{A}^H\boldsymbol{x})^H \boldsymbol{y} = (\boldsymbol{A}^H\boldsymbol{x}, \boldsymbol{y})$$

$$(\boldsymbol{A}\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{A}\boldsymbol{x})^H \boldsymbol{y} = \boldsymbol{x}^H \boldsymbol{A}^H\boldsymbol{y} = (\boldsymbol{x}, \boldsymbol{A}^H\boldsymbol{y})$$

Definition (Hermitian matrix)

Let $\boldsymbol{A}$ be a complex matrix. $\boldsymbol{A}$ is **Hermitian** if

$$\boldsymbol{A}^H = \boldsymbol{A}$$

That is

$$a_{ij}^H = a_{ji}^* = a_{ij}$$

For example

$$\begin{bmatrix} 2 & 3-3i \\ 3+3i & 5 \end{bmatrix}$$

Properties of Hermitian matrix

Let $\boldsymbol{A}$ be Hermitian.

- $\boldsymbol{x}^H \boldsymbol{A} \boldsymbol{x}$ is real for any $\boldsymbol{x}$

$$\begin{aligned} \boldsymbol{x}^H \boldsymbol{A} \boldsymbol{x} &= \boldsymbol{x}^H \boldsymbol{A}^H \boldsymbol{x} = \boldsymbol{x}^H \boldsymbol{A}^H \left(\boldsymbol{x}^H\right)^H = \left(\boldsymbol{x}^H \boldsymbol{A} \boldsymbol{x}\right)^H \\ &= \left(\boldsymbol{x}^H \boldsymbol{A} \boldsymbol{x}\right)^* \end{aligned}$$

- Eigenvalues of $\boldsymbol{A}$ are real

$$\boldsymbol{A}\boldsymbol{s}_i = \lambda_i \boldsymbol{s}_i \;\Rightarrow\; \boldsymbol{s}_i^H \boldsymbol{A} \boldsymbol{s}_i = \lambda_i \boldsymbol{s}_i^H \boldsymbol{s}_i \;\Rightarrow\; \lambda_i = \frac{\boldsymbol{s}_i^H \boldsymbol{A} \boldsymbol{s}_i}{\boldsymbol{s}_i^H \boldsymbol{s}_i} \in \mathbb{R}$$

- Eigenspaces of $\boldsymbol{A}$ are orthogonal

$$\begin{aligned} (\boldsymbol{A}\boldsymbol{s}_1, \boldsymbol{s}_2) = (\boldsymbol{s}_1, \boldsymbol{A}\boldsymbol{s}_2) &\Rightarrow (\lambda_1^* - \lambda_2)(\boldsymbol{s}_1, \boldsymbol{s}_2) = 0 \\ &\Rightarrow (\boldsymbol{s}_1, \boldsymbol{s}_2) = 0 \end{aligned}$$

Example (properties of Hermitian matrix)

$$\boldsymbol{A} = \begin{bmatrix} 2 & 3-3i \\ 3+3i & 5 \end{bmatrix}$$

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = 0 \;\Rightarrow\; \lambda_1 = 8,\ \lambda_2 = -1$$

$$(\boldsymbol{A} - 8\boldsymbol{I})\,\boldsymbol{s}_1 = \boldsymbol{0} \;\Rightarrow\; \boldsymbol{s}_1 = \begin{bmatrix} \frac{1-i}{2} \\ 1 \end{bmatrix}$$

$$(\boldsymbol{A} + \boldsymbol{I})\,\boldsymbol{s}_2 = \boldsymbol{0} \;\Rightarrow\; \boldsymbol{s}_2 = \begin{bmatrix} i-1 \\ 1 \end{bmatrix}$$

$$(\boldsymbol{s}_1, \boldsymbol{s}_2) = \left(\frac{1-i}{2}\right)^* \cdot (i-1) + 1 \cdot 1 = 0$$

Definition (unitary matrix)

Let $\boldsymbol{U}$ be a complex matrix. $\boldsymbol{U}$ is **unitary** if

$$\boldsymbol{U}^{-1} = \boldsymbol{U}^{H}$$

That is

$$\boldsymbol{U}^{H}\boldsymbol{U} = \boldsymbol{U}\boldsymbol{U}^{H} = \boldsymbol{I}$$

Properties of Unitary Matrix

Let $\boldsymbol{U}$ be unitary.

- $\|\boldsymbol{U}\boldsymbol{x}\| = \|\boldsymbol{x}\|$ for any $\boldsymbol{x}$

$$\|\boldsymbol{U}\boldsymbol{x}\|^2 = (\boldsymbol{U}\boldsymbol{x}, \boldsymbol{U}\boldsymbol{x}) = \left(\boldsymbol{x}, \boldsymbol{U}^H\boldsymbol{U}\boldsymbol{x}\right) = (\boldsymbol{x}, \boldsymbol{x}) = \|\boldsymbol{x}\|^2$$

- Eigenvalue of $\boldsymbol{U}$ has modulus 1

$$\boldsymbol{U}\boldsymbol{s}_i = \lambda_i\boldsymbol{s}_i \Rightarrow \|\boldsymbol{s}_i\| = \|\boldsymbol{U}\boldsymbol{s}_i\| = \|\lambda_i\boldsymbol{s}_i\| = |\lambda_i|\|\boldsymbol{s}_i\|$$
$$\Rightarrow |\lambda_i| = 1$$

- Eigenspaces of $\boldsymbol{U}$ are orthogonal

$$\begin{aligned} & (\boldsymbol{U}\boldsymbol{s}_1, \boldsymbol{U}\boldsymbol{s}_2) = (\boldsymbol{s}_1, \boldsymbol{U}^H\boldsymbol{U}\boldsymbol{s}_2) = (\boldsymbol{s}_1, \boldsymbol{s}_2) \\ & (\boldsymbol{U}\boldsymbol{s}_1, \boldsymbol{U}\boldsymbol{s}_2) = (\lambda_1\boldsymbol{s}_1, \lambda_2\boldsymbol{s}_2) = \lambda_1^*\lambda_2(\boldsymbol{s}_1, \boldsymbol{s}_2) \\ \Rightarrow \; & (1 - \lambda_1^*\lambda_2)(\boldsymbol{s}_1, \boldsymbol{s}_2) = 0 \\ \Rightarrow \; & (\boldsymbol{s}_1, \boldsymbol{s}_2) = 0 \end{aligned}$$

Example (unitary matrices)

Rotation matrix

$$\begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}$$

Permutation matrix

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Fourier matrix

$$\frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & \omega & \omega^2 & \omega^3 \\ 1 & \omega^2 & \omega^4 & \omega^6 \\ 1 & \omega^3 & \omega^6 & \omega^9 \end{bmatrix}, \ \omega = e^{i\frac{2\pi}{4}}$$

Definition (skew-Hermitian matrix)

Let $\boldsymbol{K}$ be a complex matrix. $\boldsymbol{K}$ is skew-Hermitian if

$$\boldsymbol{K}^H = -\boldsymbol{K}$$

Let $\boldsymbol{A}$ be Hermitian.

- $(i\boldsymbol{A})$ is skew-Hermitian since

$$(i\boldsymbol{A})^H = -i\boldsymbol{A}^H = -i\boldsymbol{A} = -(i\boldsymbol{A})$$

- For example

$$\boldsymbol{K} = i\boldsymbol{A} = i\begin{bmatrix} 2 & 3-3i \\ 3+3i & 5 \end{bmatrix} = \begin{bmatrix} 2i & 3+3i \\ -3+3i & 5i \end{bmatrix}$$

$$\boldsymbol{K}^H = \begin{bmatrix} 2i & 3+3i \\ -3+3i & 5i \end{bmatrix}^H = \begin{bmatrix} -2i & -3-3i \\ 3-3i & -5i \end{bmatrix} = -\boldsymbol{K}$$

FROM REAL ELEMENTS TO COMPLEX ELEMENTS

- Complex vectors are extension of real vectors
- Matrix Hermitian is the extension of matrix transpose
- Hermitian matrix is the extension of symmetric matrix
- Unitary matrix is the extension of orthogonal matrix
- Skew-Hermitian is the extension of anti-symmetric

Similar Matrices

Definition (similar matrices)

Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be matrices. $\boldsymbol{A}$ and $\boldsymbol{B}$ are **similar** if there exists an invertible matrix $\boldsymbol{M}$ such that

$$\boldsymbol{B} = \boldsymbol{M}^{-1}\boldsymbol{A}\boldsymbol{M}$$

- Notation

$$\boldsymbol{A} \sim \boldsymbol{B}$$

- An equivalence relation

$$\begin{aligned} (\boldsymbol{A} \sim \boldsymbol{B}) \;&\Rightarrow\; (\boldsymbol{B} \sim \boldsymbol{A}) \\ (\boldsymbol{A} \sim \boldsymbol{B}) \wedge (\boldsymbol{B} \sim \boldsymbol{C}) \;&\Rightarrow\; (\boldsymbol{A} \sim \boldsymbol{C}) \end{aligned}$$

Example (similar matrices)

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad M_1 = \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix}, \quad M_2 = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

- M_1 and M_2 are invertible

$$M_1^{-1} = \begin{bmatrix} 1 & -b \\ 0 & 1 \end{bmatrix}, \quad M_2^{-1} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

- Similarity

$$A \sim M_1^{-1} A M_1 = \begin{bmatrix} 1 & b \\ 0 & 0 \end{bmatrix} = B_1$$

$$A \sim M_2^{-1} A M_2 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = B_2$$

$$B_1 \sim B_2$$

Properties of similar matrices

Let A and B be similar matrices.

- They have the same eigenvalues
- Their eigenvectors are related

Lemma (eigenvalues of similar matrices)

If $\boldsymbol{A} \sim \boldsymbol{B}$, $\boldsymbol{A}$ and $\boldsymbol{B}$ have the same spectrum.

Proof.

Let $\boldsymbol{B} = \boldsymbol{M}^{-1}\boldsymbol{A}\boldsymbol{M}$ and λ be an eigenvalue of $\boldsymbol{B}$.

$$\begin{aligned} |\boldsymbol{B} - \lambda \boldsymbol{I}| = 0 \;&\Rightarrow\; |\boldsymbol{M}^{-1}\boldsymbol{A}\boldsymbol{M} - \lambda \boldsymbol{M}^{-1}\boldsymbol{M}| = 0 \\ &\Rightarrow\; |\boldsymbol{M}^{-1}(\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{M}| = 0 \\ &\Rightarrow\; |\boldsymbol{M}^{-1}|\,|(\boldsymbol{A} - \lambda \boldsymbol{I})|\,|\boldsymbol{M}| = 0 \\ &\Rightarrow\; |\boldsymbol{A} - \lambda \boldsymbol{I}| = 0 \end{aligned}$$

Hence λ is an eigenvalue of $\boldsymbol{A}$. □

Lemma (eigenvectors of similar matrices)

Suppose $\boldsymbol{A} \sim \boldsymbol{B}$ with $\boldsymbol{B} = \boldsymbol{M}^{-1}\boldsymbol{A}\boldsymbol{M}$. Let $\boldsymbol{s}$ be an eigenvector of $\boldsymbol{A}$ with eigenvalue λ.

- *$\boldsymbol{t} = \boldsymbol{M}^{-1}\boldsymbol{s}$ is an eigenvector of $\boldsymbol{B}$*
- *$\boldsymbol{t}$ also corresponds to eigenvalue λ*

Proof.

$$\begin{aligned} \boldsymbol{A}\boldsymbol{s} = \lambda\boldsymbol{s} &\Rightarrow \boldsymbol{M}^{-1}\boldsymbol{A}\boldsymbol{s} = \lambda\boldsymbol{M}^{-1}\boldsymbol{s} \\ &\Rightarrow \left(\boldsymbol{M}^{-1}\boldsymbol{A}\boldsymbol{M}\right)\left(\boldsymbol{M}^{-1}\boldsymbol{s}\right) = \lambda\left(\boldsymbol{M}^{-1}\boldsymbol{s}\right) \\ &\Rightarrow \boldsymbol{B}\boldsymbol{t} = \lambda\boldsymbol{t} \end{aligned}$$

Hence $\boldsymbol{t}$ is an eigenvector of $\boldsymbol{B}$ with eigenvalue λ. □

Example (eigenvectors of similar matrices)

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \boldsymbol{B}_1 = \begin{bmatrix} 1 & b \\ 0 & 0 \end{bmatrix}, \boldsymbol{M}_1 = \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix}$$

- $\boldsymbol{B}_1 \sim \boldsymbol{A}$ since $\boldsymbol{B}_1 = \boldsymbol{M}_1^{-1}\boldsymbol{A}\boldsymbol{M}_1$
- $\boldsymbol{s}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\boldsymbol{s}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ are eigenvectors of $\boldsymbol{A}$ with eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 0$
- Thus

$$\boldsymbol{t}_1 = \boldsymbol{M}_1^{-1}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \boldsymbol{t}_2 = \boldsymbol{M}_1^{-1}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -b \\ 1 \end{bmatrix}$$

are eigenvectors of $\boldsymbol{B}_1$, also with eigenvalues 1 and 0

Change of Basis

Definition (identity transformation)

The **identity transformation** $\boldsymbol{I} : \mathbb{V} \mapsto \mathbb{V}$ maps any vector to itself

$$\boldsymbol{I}(\boldsymbol{x}) = \boldsymbol{x}$$

- Linearity

$$\begin{aligned}\boldsymbol{I}(c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2) &= c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2 \\ &= c_1\boldsymbol{I}(\boldsymbol{x}_1) + c_2\boldsymbol{I}(\boldsymbol{x}_2)\end{aligned}$$

- Change of basis: one basis for domain $\mathbb{V}$ and one basis for range $\mathbb{V}$

Definition (change of basis)

Let $\mathbb{V}$ be a vector space. In a **change of basis**, the basis used to represent vectors in $\mathbb{V}$ is changed from one to another.

A change of basis is an identity transformation

Let $\mathcal{B}$ and $\mathcal{B}'$ be bases of $\mathbb{V}$. Consider the change of basis from $\mathcal{B}$ to $\mathcal{B}'$.

- It is identity transformation since vectors are not changed
- That means the change of basis is $\boldsymbol{I} : \mathbb{V} \mapsto \mathbb{V}$, where $\mathcal{B}$ is the domain basis and $\mathcal{B}'$ is the range basis

MATRIX FOR CHANGE OF BASIS

Consider the change of basis from $\mathcal{B} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ to $\mathcal{B}' = \{\boldsymbol{v}'_1, \ldots, \boldsymbol{v}'_n\}$ for space $\mathbb{V}$.

- It is $\boldsymbol{I} : \mathbb{V} \to \mathbb{V}$
- It has a matrix representation $[\boldsymbol{I}_{\mathcal{B}\mathcal{B}'}]$
- Suppose

$$\boldsymbol{v}_j = \sum_{i=1}^{n} m_{ij} \boldsymbol{v}'_i, \; j = 1, \ldots, n$$

Then

$$\boldsymbol{I}(\boldsymbol{v}_j) = \boldsymbol{v}_j = \sum_{i=1}^{n} m_{ij} \boldsymbol{v}'_i, \quad j = 1, \ldots, n$$

Hence

$$[\boldsymbol{I}_{\mathcal{B}\mathcal{B}'}] = \{m_{ij}\}$$

Representation of a vector with a basis

Let $\mathbb{V}$ be a vector space and $\mathcal{B} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ be a basis of $\mathbb{V}$.

- A vector $\boldsymbol{x} \in \mathbb{V}$ is a linear combination of the basis vectors

$$\boldsymbol{x} = \sum_{i=1}^{n} x_i \boldsymbol{v}_i$$

- Given $\mathcal{B}$, $\boldsymbol{x}$ can be represented by

$$[\boldsymbol{x}_{\mathcal{B}}] = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \Leftrightarrow \boldsymbol{x} = \sum_{i=1}^{n} x_i \boldsymbol{v}_i$$

Theorem (representation in two bases)

Let $\boldsymbol{x}$ be a vector. The representation of $\boldsymbol{x}$ in bases $\mathcal{B}$ and $\mathcal{B}'$ are related by

$$[\boldsymbol{x}_{\mathcal{B}'}] = [\boldsymbol{I}_{\mathcal{B}\mathcal{B}'}]\,[\boldsymbol{x}_{\mathcal{B}}]$$

Suppose $\boldsymbol{x} = \sum_{j=1}^{n} x_j \boldsymbol{v}_j$ and $\boldsymbol{v}_j = \sum_{i=1}^{n} m_{ij}\boldsymbol{v}'_i$. We have

$$\sum_{j=1}^{n} x_j \boldsymbol{v}_j = \sum_{j=1}^{n} x_j \sum_{i=1}^{n} m_{ij}\boldsymbol{v}'_i = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} m_{ij}x_j\right)\boldsymbol{v}'_i = \sum_{i=1}^{n} x'_i\boldsymbol{v}'_i$$

Hence

$$x'_i = \sum_{j=1}^{n} m_{ij}x_j,\ \ i = 1,\ldots,n$$

That is

$$[\boldsymbol{x}_{\mathcal{B}'}] = [\boldsymbol{I}_{\mathcal{B}\mathcal{B}'}]\,[\boldsymbol{x}_{\mathcal{B}}]$$

Theorem (inverse matrix and change of basis)

Let $\mathcal{B}$ and $\mathcal{B}'$ be bases of $\mathbb{V}$.

$$[\boldsymbol{I}_{\mathcal{B}'\mathcal{B}}] = [\boldsymbol{I}_{\mathcal{B}\mathcal{B}'}]^{-1}$$

Proof.

For any $\boldsymbol{x} \in \mathbb{V}$, it follows from $[\boldsymbol{x}_{\mathcal{B}}] = [\boldsymbol{I}_{\mathcal{B}'\mathcal{B}}][\boldsymbol{x}_{\mathcal{B}'}]$ and $[\boldsymbol{x}_{\mathcal{B}'}] = [\boldsymbol{I}_{\mathcal{B}\mathcal{B}'}][\boldsymbol{x}_{\mathcal{B}}]$ that

$$[\boldsymbol{x}_{\mathcal{B}}] = [\boldsymbol{I}_{\mathcal{B}'\mathcal{B}}][\boldsymbol{I}_{\mathcal{B}\mathcal{B}'}][\boldsymbol{x}_{\mathcal{B}}]$$

Hence

$$[\boldsymbol{I}_{\mathcal{B}'\mathcal{B}}]\,[\boldsymbol{I}_{\mathcal{B}\mathcal{B}'}] = \boldsymbol{I}$$

$$[\boldsymbol{I}_{\mathcal{B}'\mathcal{B}}] = [\boldsymbol{I}_{\mathcal{B}\mathcal{B}'}]^{-1}$$

□

Theorem (similarity and linear transformation)

Let $\boldsymbol{T} : \mathbb{V} \mapsto \mathbb{V}$ *be linear transformation and* $\mathcal{B} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ *and* $\mathcal{B}' = \{\boldsymbol{v}_1', \ldots, \boldsymbol{v}_n'\}$ *be bases of* $\mathbb{V}$.

- *The matrix representation of* $\boldsymbol{T}$ *using* $\mathcal{B}$ *and* $\mathcal{B}'$, *respectively* $[\boldsymbol{T}_{\mathcal{B}\mathcal{B}}]$ *and* $[\boldsymbol{T}_{\mathcal{B}'\mathcal{B}'}]$, *must be similar.*
- *That is*

$$[\boldsymbol{T}_{\mathcal{B}'\mathcal{B}'}] \sim [\boldsymbol{T}_{\mathcal{B}\mathcal{B}}]$$

Proof.

Suppose $\boldsymbol{T}$ maps $\boldsymbol{x}$ to $\boldsymbol{y}$. From identity transformation

$$[\boldsymbol{x}_{\mathcal{B}'}] = [\boldsymbol{I}_{\mathcal{B}\mathcal{B}'}]\,[\boldsymbol{x}_{\mathcal{B}}]\,,\;\; [\boldsymbol{y}_{\mathcal{B}'}] = [\boldsymbol{I}_{\mathcal{B}\mathcal{B}'}]\,[\boldsymbol{y}_{\mathcal{B}}]$$

From linear transformation $\boldsymbol{T}$

$$[\boldsymbol{y}_{\mathcal{B}'}] = [\boldsymbol{T}_{\mathcal{B}'\mathcal{B}'}]\,[\boldsymbol{x}_{\mathcal{B}'}]\,,\;\; [\boldsymbol{y}_{\mathcal{B}}] = [\boldsymbol{T}_{\mathcal{B}\mathcal{B}}]\,[\boldsymbol{x}_{\mathcal{B}}]$$

It follows that

$$\begin{aligned}
& [\boldsymbol{I}_{\mathcal{B}\mathcal{B}'}]\,[\boldsymbol{y}_{\mathcal{B}}] = [\boldsymbol{T}_{\mathcal{B}'\mathcal{B}'}]\,([\boldsymbol{I}_{\mathcal{B}\mathcal{B}'}]\,[\boldsymbol{x}_{\mathcal{B}}]) \\
\Rightarrow\;\; & [\boldsymbol{y}_{\mathcal{B}}] = [\boldsymbol{I}_{\mathcal{B}\mathcal{B}'}]^{-1}\,[\boldsymbol{T}_{\mathcal{B}'\mathcal{B}'}]\,[\boldsymbol{I}_{\mathcal{B}\mathcal{B}'}]\,[\boldsymbol{x}_{\mathcal{B}}] \\
\Rightarrow\;\; & [\boldsymbol{T}_{\mathcal{B}\mathcal{B}}] = [\boldsymbol{I}_{\mathcal{B}\mathcal{B}'}]^{-1}\,[\boldsymbol{T}_{\mathcal{B}'\mathcal{B}'}]\,[\boldsymbol{I}_{\mathcal{B}\mathcal{B}'}]
\end{aligned}$$

Hence $[\boldsymbol{T}_{\mathcal{B}\mathcal{B}}] \sim [\boldsymbol{T}_{\mathcal{B}'\mathcal{B}'}]$. □

Definition (eigenbasis of linear transform)

Let $\mathbb{V}$ be a space and $\boldsymbol{T} : \mathbb{V} \mapsto \mathbb{V}$ be a linear transformation.

- A vector $\boldsymbol{v}_i$ such that $\boldsymbol{T}(\boldsymbol{v}_i) = \lambda_i \boldsymbol{v}_i$ is an eigenvector of $\boldsymbol{T}$ with eigenvalue λ_i
- A basis $\mathcal{B} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ of $\mathbb{V}$ is an eigenbasis

Properties

- Matrix representation for $\boldsymbol{T}$ using $\mathcal{B}$ is diagonal

$$[\boldsymbol{T}_{\mathcal{B}\mathcal{B}}] = \mathbf{diag}(\lambda_1, \ldots, \lambda_n) = \boldsymbol{\Lambda}$$

- Let $\mathcal{B}' = \{\boldsymbol{v}'_1, \ldots, \boldsymbol{v}'_n\}$ be a basis, and $\boldsymbol{v}_j = \sum_i s_{ij} \boldsymbol{v}'_i$. Then $[\boldsymbol{I}_{\mathcal{B}\mathcal{B}'}] = \{s_{ij}\} = \boldsymbol{S}$ and

$$[\boldsymbol{T}_{\mathcal{B}'\mathcal{B}'}] = [\boldsymbol{I}_{\mathcal{B}\mathcal{B}'}] [\boldsymbol{T}_{\mathcal{B}\mathcal{B}}] [\boldsymbol{I}_{\mathcal{B}\mathcal{B}'}]^{-1} = \boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^{-1}$$

Example (eigenbasis for projection)

Let $\boldsymbol{T}$ be the projection to the line L at angle θ to the horizontal axis. Find the matrix for $\boldsymbol{T}$ using standard basis $\mathcal{B}' = \{\boldsymbol{e}_x, \boldsymbol{e}_y\}$.

- Eigenvectors of $\boldsymbol{T}$ are $\boldsymbol{v}_1 = \cos\theta\,\boldsymbol{e}_x + \sin\theta\,\boldsymbol{e}_y$ with $\lambda_1 = 1$ and $\boldsymbol{v}_2 = -\sin\theta\,\boldsymbol{e}_x + \cos\theta\,\boldsymbol{e}_y$ with $\lambda_2 = 0$
- The matrix for $\boldsymbol{T}$ using eigenbasis $\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ is

$$[\boldsymbol{T}_{\mathcal{B}\mathcal{B}}] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \boldsymbol{\Lambda}$$

- The matrix for the change of basis from $\mathcal{B}$ to $\mathcal{B}'$ is

$$[\boldsymbol{I}_{\mathcal{B}\mathcal{B}'}] = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

- The matrix for $\boldsymbol{T}$ using standard basis is

$$[\boldsymbol{T}_{\mathcal{B}'\mathcal{B}'}] = [\boldsymbol{I}_{\mathcal{B}\mathcal{B}'}]\,[\boldsymbol{T}_{\mathcal{B}\mathcal{B}}]\,[\boldsymbol{I}_{\mathcal{B}\mathcal{B}'}]^{-1} = \begin{bmatrix} \cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{bmatrix}$$

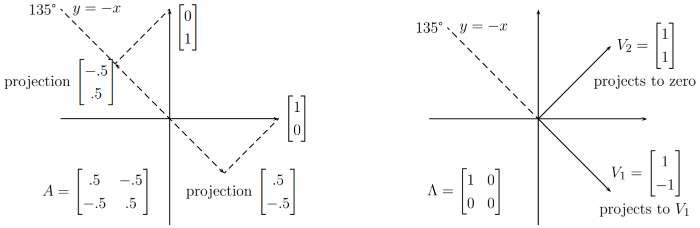

Figure 5.5: Change of basis to make the projection matrix diagonal.

Normal Matrix and Orthonormal Eigenbasis

Triangularizability and Diagonalizability

Let $\boldsymbol{A}$ be a square matrix.

- $\boldsymbol{A}$ is triangularizable if $\boldsymbol{A}$ is similar to a triangular matrix
- $\boldsymbol{A}$ is diagonalizable if $\boldsymbol{A}$ is similar to a diagonal matrix

Let $\boldsymbol{A}$ have eigen-decomposition $\boldsymbol{A} = \boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^{-1}$. Then

$$\boldsymbol{A} \sim \boldsymbol{\Lambda}$$

so $\boldsymbol{A}$ is diagonalizable (and triangularizable).

Lemma (Schur lemma)

Let $\boldsymbol{A}$ *be a square matrix. There exists a unitary matrix* $\boldsymbol{U}$ *such that*

$$\boldsymbol{U}^H\boldsymbol{A}\boldsymbol{U} = \boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}$$

is an upper-triangular matrix.

Note

$$\boldsymbol{A} \sim \boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}$$

and $\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}$ is triangular. Hence, Schur lemma guarantees every square matrix is triangularizable.

Definition (normal matrix)

Let $\boldsymbol{N}$ be a square matrix. $\boldsymbol{N}$ is **normal** if

$$\boldsymbol{N}\boldsymbol{N}^H = \boldsymbol{N}^H\boldsymbol{N}$$

- Unitary matrix is normal

$$\boldsymbol{U}\boldsymbol{U}^H = \boldsymbol{U}^H\boldsymbol{U} = \boldsymbol{I}$$

- Hermitian matrix is normal

$$\boldsymbol{H}\boldsymbol{H}^H = \boldsymbol{H}\boldsymbol{H} = \boldsymbol{H}^H\boldsymbol{H}$$

Theorem (normal matrix can be diagonalized)

Let $\boldsymbol{N}$ be a normal matrix. $\boldsymbol{N}$ is diagonalizable.

Proof

Let $\boldsymbol{U}$ be unitary and $\boldsymbol{\Gamma} = \boldsymbol{U}^H\boldsymbol{N}\boldsymbol{U}$ is upper-triangular. Note

$$\begin{aligned}
\boldsymbol{\Gamma}\boldsymbol{\Gamma}^H &= \boldsymbol{U}^H\boldsymbol{N}\boldsymbol{U}(\boldsymbol{U}^H\boldsymbol{N}\boldsymbol{U})^H \\
&= \boldsymbol{U}^H\boldsymbol{N}\boldsymbol{N}^H\boldsymbol{U} \\
&= \boldsymbol{U}^H\boldsymbol{N}^H\boldsymbol{N}\boldsymbol{U} \\
&= \boldsymbol{U}^H\boldsymbol{N}^H\boldsymbol{U}\boldsymbol{U}^H\boldsymbol{N}\boldsymbol{U} \\
&= \boldsymbol{\Gamma}^H\boldsymbol{\Gamma}
\end{aligned}$$

COMPLETING THE PROOF

For the first diagonal element of $\mathbf{\Gamma}\mathbf{\Gamma}^H$ and $\mathbf{\Gamma}^H\mathbf{\Gamma}$

$$\begin{aligned} (\mathbf{\Gamma}\mathbf{\Gamma}^H)_{11} &= (\mathbf{\Gamma}^H\mathbf{\Gamma})_{11} \\ \Rightarrow \sum_k \gamma_{1k}\gamma_{1k}^* &= \sum_k \gamma_{k1}^*\gamma_{k1} = |\gamma_{11}|^2 \\ \Rightarrow \gamma_{1k} &= 0, \quad k > 1 \end{aligned}$$

For the second diagonal element

$$\begin{aligned} (\mathbf{\Gamma}\mathbf{\Gamma}^H)_{22} &= (\mathbf{\Gamma}^H\mathbf{\Gamma})_{22} \\ \Rightarrow \sum_k \gamma_{2k}\gamma_{2k}^* &= \sum_k \gamma_{k2}^*\gamma_{k2} = |\gamma_{12}|^2 + |\gamma_{22}|^2 = |\gamma_{22}|^2 \\ \Rightarrow \gamma_{2k} &= 0, \quad k > 2 \end{aligned}$$

Row by row, we can show the elements of $\mathbf{\Gamma}$ to the right of the diagonal are 0. Hence $\mathbf{\Gamma}$ is diagonal and $\boldsymbol{N}$ is diagonalizable.

Normal matrix has orthonormal eigenbasis

Let $\boldsymbol{N}$ be a normal matrix. $\boldsymbol{N}$ has an orthonormal eigenbasis.

Proof.

Let $\boldsymbol{U}$ be unitary and diagonalize $\boldsymbol{N}$. It means $\boldsymbol{U}^H\boldsymbol{N}\boldsymbol{U} = \boldsymbol{\Gamma}$ where $\boldsymbol{\Gamma}$ is diagonal. Let $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ be the columns of $\boldsymbol{U}$.

- $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ are eigenvectors of $\boldsymbol{N}$

$$\boldsymbol{U}^H\boldsymbol{N}\boldsymbol{U} = \boldsymbol{\Gamma} \;\Rightarrow\; \boldsymbol{N}\boldsymbol{U} = \boldsymbol{U}\boldsymbol{\Gamma} \;\Rightarrow\; \boldsymbol{N}\boldsymbol{u}_i = \gamma_{ii}\boldsymbol{u}_i$$

- $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ is an orthonormal eigenbasis

$$\boldsymbol{U}^H\boldsymbol{U} = \boldsymbol{I} \;\Rightarrow\; \boldsymbol{u}_i^H\boldsymbol{u}_j = \delta_{ij}$$

□

Theorem (spectral theorem)

Let $\boldsymbol{A}$ be a real symmetric matrix. Then

$$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^T$$

where $\boldsymbol{\Lambda}$ is real and diagonal, and $\boldsymbol{Q}$ is real and orthogonal.

- $\boldsymbol{A}$ is normal, so it has orthonormal eigenbasis
- $\boldsymbol{A}$ is Hermitian, so its eigenvalues are real
- $\boldsymbol{\Lambda}$ is eigenvalue matrix and $\boldsymbol{Q}$ is eigenvector matrix
- As a sum

$$\begin{aligned}\boldsymbol{A} &= \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^T \\ &= \lambda_1 \boldsymbol{q}_1 \boldsymbol{q}_1^T + \cdots + \lambda_n \boldsymbol{q}_n \boldsymbol{q}_n^T\end{aligned}$$

Example (spectral decomposition)

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

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