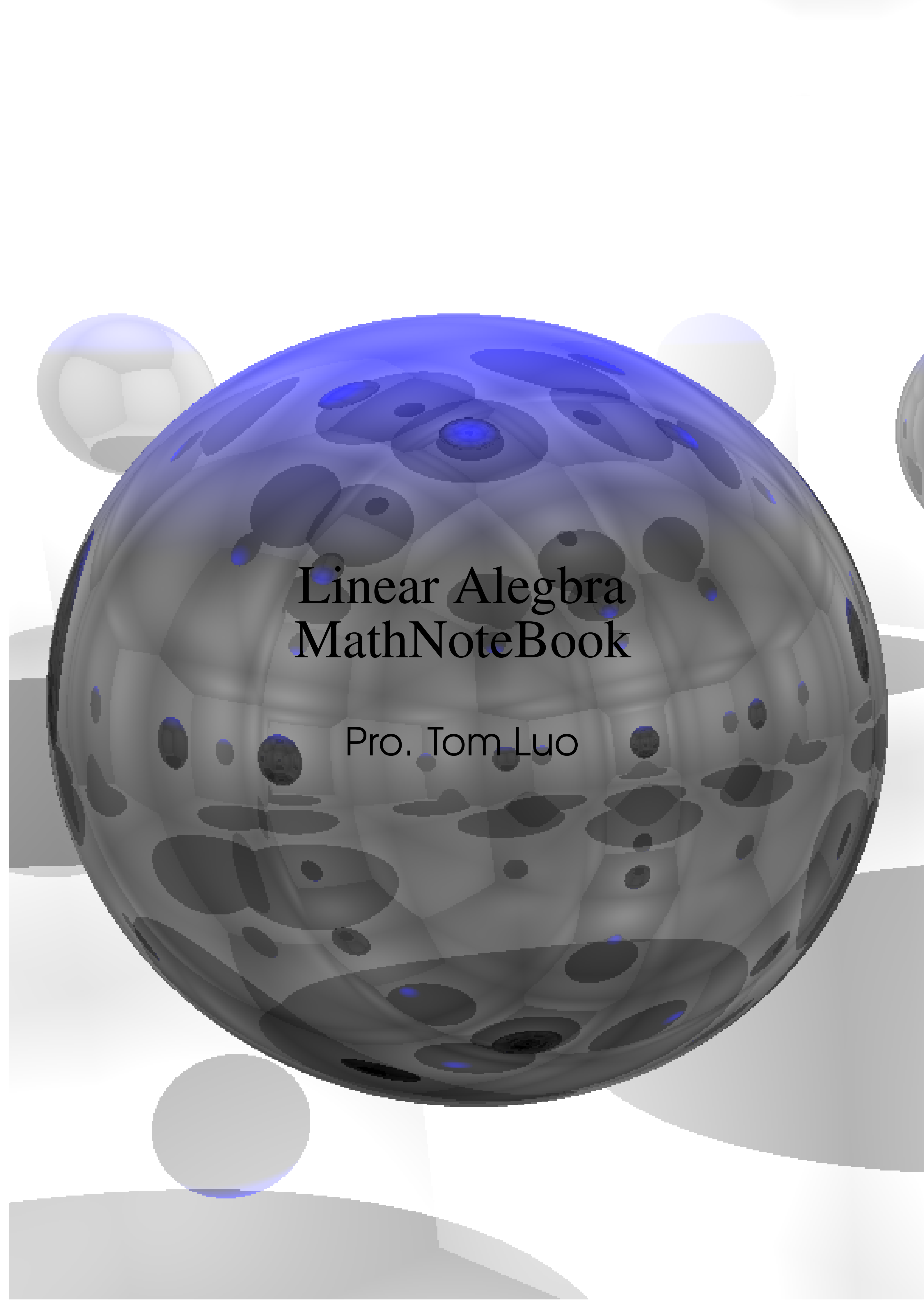

Linear Alegbra MathNoteBook

Pro. Tom Luo

16 — Week7

16.1 Friday

16.1.1 *Review*

- **Diagonalization:** If a $n \times n$ matrix is diagonalizable, it's equivalent to say it has n ind. eigenvectors. So its eigenvectors form a basis for $\mathbb{R}^n$. $(*)$
- If *eigenvalues are distinct*, then $(*)$ holds.

16.1.2 *Fibonacci Numbers*

We show a famous example, where eigenvalues tell how to find the formula for Fibonacci Numbers.
Every new Fibonacci number come from two previous ones:

$$\textbf{Fibonacci Number: } 0,1,1,2,3,5,8,13,\dots$$
$$\textbf{Fibonacci Equation: } \boldsymbol{F}_{k+2} = \boldsymbol{F}_{k+1} + \boldsymbol{F}_k, \quad \boldsymbol{F}_0 = 0, \boldsymbol{F}_1 = 1.$$

How to compute $\boldsymbol{F}_{100}$ without computing $\boldsymbol{F}_2$ to $\boldsymbol{F}_{99}$?
The key is to begin with a matrix equation $\boldsymbol{u}_{k+1} = \boldsymbol{A}\boldsymbol{u}_k$. We put two Fibonacci number into a vector $\boldsymbol{u}_k$, then you will see the matrix $\boldsymbol{A}$:

$$\text{Let } \boldsymbol{u}_k = \begin{bmatrix} \boldsymbol{F}_{k+1} \\ \boldsymbol{F}_k \end{bmatrix}. \text{ The rule } \begin{cases} \boldsymbol{F}_{k+2} = \boldsymbol{F}_{k+1} + \boldsymbol{F}_k \\ \boldsymbol{F}_{k+1} = \boldsymbol{F}_{k+1} \end{cases} \text{ is } \boldsymbol{u}_{k+1} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \boldsymbol{u}_k. \qquad \boldsymbol{u}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Every step we mutliply $\boldsymbol{u}_0$ by $\boldsymbol{A}$. After 100 steps we obtain $\boldsymbol{u}_{100} = \boldsymbol{A}^{100}\boldsymbol{u}_0$:

$$\boldsymbol{u}_{100} = \begin{bmatrix} \boldsymbol{F}_{101} \\ \boldsymbol{F}_{100} \end{bmatrix} = \boldsymbol{A}^{100}\boldsymbol{u}_0 = \boldsymbol{A}^{100} \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

But how to compute $\boldsymbol{A}^{100}$? If possible, you can diagonalize $\boldsymbol{A}$.
It's eay to show that for matrix $\boldsymbol{A} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$, we can decompose it into $\boldsymbol{A} = \boldsymbol{S}\boldsymbol{D}\boldsymbol{S}^{-1}$.

where $\boldsymbol{D} = \mathrm{diag}(\lambda_1, \lambda_2)$, $\boldsymbol{S} = \begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix}$.

And $\begin{bmatrix} \lambda_1 \\ 1 \end{bmatrix}$ is the eigenvector corresponding to λ_1, $\begin{bmatrix} \lambda_2 \\ 1 \end{bmatrix}$ is the eigenvector corresponding to λ_2.

You can verify $\lambda_1 = \frac{1+\sqrt{5}}{2}, \lambda_2 = \frac{1-\sqrt{5}}{2}$.

Thus we obtain $\boldsymbol{A}^{100} = \boldsymbol{S}\boldsymbol{D}^{100}\boldsymbol{S}^{-1}$. Hence we can compute $\boldsymbol{u}_{100}$:

$$
\begin{aligned}
\boldsymbol{u}_{100} = \boldsymbol{A}^{100}\boldsymbol{u}_0 = \boldsymbol{S}\boldsymbol{D}^{100}\boldsymbol{S}^{-1}\boldsymbol{u}_0 &= \boldsymbol{S}\begin{pmatrix} \lambda_1^{100} & \\ & \lambda_2^{100} \end{pmatrix}\boldsymbol{S}^{-1}\boldsymbol{u}_0 \\
&= \begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix}\begin{pmatrix} \lambda_1^{100} & \\ & \lambda_2^{100} \end{pmatrix}\begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix}^{-1}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \boldsymbol{F}_{101} \\ \boldsymbol{F}_{100} \end{bmatrix}
\end{aligned}
$$

After messy computation, we obtain

$$
\boldsymbol{F}_{100} = \frac{1}{\sqrt{5}}[\lambda_1^{100} - \lambda_2^{100}] = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^{100} - \left(\frac{1-\sqrt{5}}{2}\right)^{100}\right]
$$

Another way to compute $\boldsymbol{F}_{100}$:

We wet $\boldsymbol{S} = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 \end{bmatrix}$, where $\boldsymbol{x}_1 = \begin{bmatrix} \lambda_1 \\ 1 \end{bmatrix}, \boldsymbol{x}_2 = \begin{bmatrix} \lambda_2 \\ 1 \end{bmatrix}$. $\boldsymbol{x}_1, \boldsymbol{x}_2$ are eigenvectors of $\boldsymbol{A}$.

We let $\boldsymbol{u}_k = \begin{bmatrix} \boldsymbol{F}_{k+1} \\ \boldsymbol{F}_k \end{bmatrix}$.

Firstly, We want to find linear combination of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ to get $\boldsymbol{u}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$:

$$
\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{\lambda_1 - \lambda_2}\left(\begin{bmatrix} \lambda_1 \\ 1 \end{bmatrix} - \begin{bmatrix} \lambda_2 \\ 1 \end{bmatrix}\right) \qquad \text{or} \qquad \boldsymbol{u}_0 = \frac{\boldsymbol{x}_1 - \boldsymbol{x}_2}{\lambda_1 - \lambda_2}
$$

Then we multiply $\boldsymbol{u}_0$ by $\boldsymbol{A}^{100}$ to get $\boldsymbol{u}_{100}$:

$$
\begin{aligned}
\boldsymbol{u}_{100} = \boldsymbol{A}^{100}\boldsymbol{u}_0 &= \frac{\boldsymbol{A}^{100}\boldsymbol{x}_1 - \boldsymbol{A}^{100}\boldsymbol{x}_2}{\lambda_1 - \lambda_2} \\
&= \frac{\boldsymbol{A}^{99}(\boldsymbol{A}\boldsymbol{x}_1) - \boldsymbol{A}^{99}(\boldsymbol{A}\boldsymbol{x}_2)}{\lambda_1 - \lambda_2} = \frac{\lambda_1\boldsymbol{A}^{99}\boldsymbol{x}_1 - \lambda_2\boldsymbol{A}^{99}\boldsymbol{x}_2}{\lambda_1 - \lambda_2} = \frac{\lambda_1^2\boldsymbol{A}^{98}\boldsymbol{x}_1 - \lambda_2^2\boldsymbol{A}^{98}\boldsymbol{x}_2}{\lambda_1 - \lambda_2} = \dots \\
&= \frac{\lambda_1^{100}\boldsymbol{x}_1 - \lambda_2^{100}\boldsymbol{x}_2}{\lambda_1 - \lambda_2}
\end{aligned}
$$

Since $\lambda_1 - \lambda_2 = \sqrt{5}$, finally we obtain the same result.

16.1.3 Imaginary Eigenvalues

The eigenvalues might not be real numbers sometimes.

■ **Example 16.1**

Consider the rotation matrix given by $\boldsymbol{K} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. It rotates our vector by 90°:

$$
\boldsymbol{K}\begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}.
$$

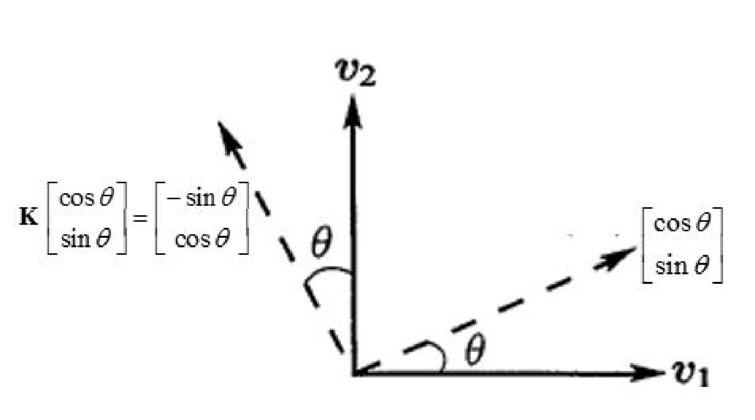

Figure 16.1: Rotate a vector by 90°.

This rotation matrix exists eigenvector and eigenvalue, which means $\exists \boldsymbol{v} \neq \mathbf{0}$ and λ s.t.

$$\boldsymbol{K}\boldsymbol{v} = \lambda \boldsymbol{v}.$$

However, this equation means this rotaion matrix doesn't change the direction of $\boldsymbol{v}$. But in geometric meaning it rotates vector $\boldsymbol{v}$ by 90°. Why? This phenomenon will not happen unless we go to imaginary eigenvectors. Let's compute eigenvalues and eigenvectors for $\boldsymbol{K}$ first:

$$P_{\boldsymbol{K}}(\lambda) = \begin{vmatrix} \lambda & 1 \\ -1 & \lambda \end{vmatrix} = \lambda^2 + 1 \implies \lambda_1 = i, \quad \lambda_2 = -i.$$

$$(\lambda_1 \boldsymbol{I} - \boldsymbol{K})\boldsymbol{x} = \begin{pmatrix} i & 1 \\ -1 & i \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \mathbf{0} \implies \boldsymbol{x} = \alpha \begin{pmatrix} 1 \\ -i \end{pmatrix}.$$

$$(\lambda_2 \boldsymbol{I} - \boldsymbol{K})\boldsymbol{x} = \begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \mathbf{0} \implies \boldsymbol{x} = \beta \begin{pmatrix} 1 \\ i \end{pmatrix}.$$

Moverover, we can do similar transformation for $\boldsymbol{K}$:

$$\boldsymbol{D} = \boldsymbol{S}^{-1}\boldsymbol{K}\boldsymbol{S} = \begin{pmatrix} i & \\ & -i \end{pmatrix} \qquad \text{where } \boldsymbol{S} = \begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix}.$$

■

R For motion in vector space, eigenvalues are "speed" and eigenvectors are "directions" under basis $\boldsymbol{S} = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \dots & \boldsymbol{x}_n \end{bmatrix}$.

$$\boldsymbol{v} = c_1\boldsymbol{x}_1 + \dots + c_n\boldsymbol{x}_n \xrightarrow{\text{postmultiply } \boldsymbol{A}} \boldsymbol{A}\boldsymbol{v} = c_1\lambda_1\boldsymbol{x}_1 + \dots + c_n\lambda_n\boldsymbol{x}_n.$$

$$\begin{pmatrix} c_1 & \dots & c_n \end{pmatrix} \xrightarrow{\text{rightmultiply } \boldsymbol{D} = \operatorname{diag}(\lambda_1, \dots, \lambda_n)} \begin{pmatrix} c_1\lambda_1 & \dots & c_n\lambda_n \end{pmatrix}.$$

16.1.4 Complex Numbers

Even when the matrix is real, its eigenvalues of this matrix may be complex numbers. Example: A 2 by 2 rotation matrix has no real eigenvectors. It rotates a vector by 90°. But it has complex eigenvalues i and $-i$.

Definition 16.1 — Complex Numbers. A cpmplex number $\boldsymbol{x} \in \mathbb{C}$ could be written as $\boldsymbol{x} = a + bi$, where $i^2 = -1$.
Its **complex conjugate** is defined as $\bar{\boldsymbol{x}} = a - bi$.
Its **modulus** is defined as $|\boldsymbol{x}| = \sqrt{a^2 + b^2} = \boldsymbol{x}\bar{\boldsymbol{x}}$.

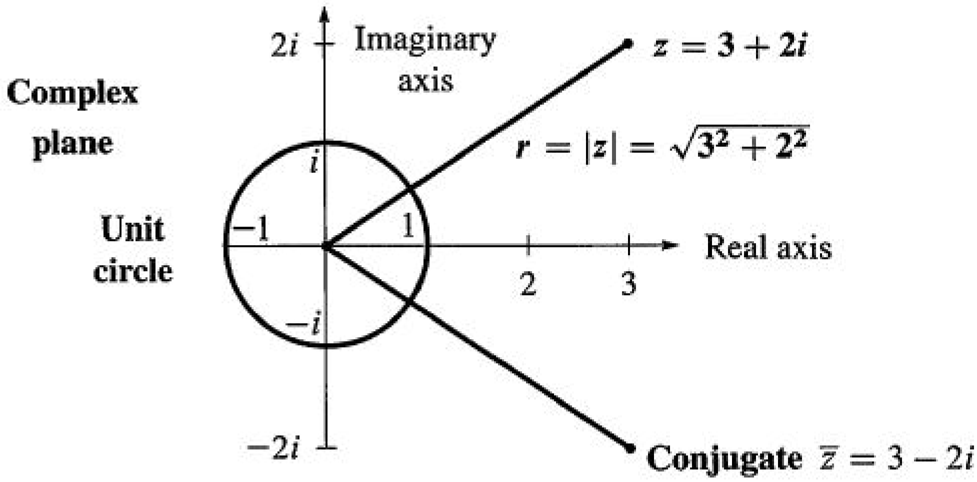

Figure 16.2: The number $z = a + bi$ corrsponds to the vector $\begin{bmatrix} a \\ b \end{bmatrix}$.

16.1.5 Complex Vectors

Definition 16.2 — Length (norm) for complex. For $z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} \in \mathbb{C}^n$, its **length (norm)** is defined as

$$\|z\| = \sqrt{|z_1|^2 + |z_2|^2 + \cdots + |z_n|^2} = \sqrt{\langle z, z \rangle} = \sqrt{z_1\bar{z}_1 + z_2\bar{z}_2 + \cdots + z_n\bar{z}_n}.$$

Before we introduce the definition of inner product for complex, let's introduce the *Hermition* of a vector in $\mathbb{C}^n$:

Definition 16.3 — Hermition. The hermition of a vector in $\mathbb{C}^n$ is its *conjugate transpose*.

$$\boldsymbol{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \qquad \boldsymbol{z}^{\mathrm{H}} = \bar{\boldsymbol{z}}^{\mathrm{T}} = \begin{bmatrix} \bar{z}_1 & \ldots & \bar{z}_n \end{bmatrix}.$$

Definition 16.4 — Inner product. The inner product of real or complex vectors $\boldsymbol{z}$ and $\boldsymbol{w}$ is $\boldsymbol{w}^{\mathrm{H}}\boldsymbol{z}$, which is defined as

$$\langle \boldsymbol{z},\boldsymbol{w}\rangle = \boldsymbol{w}^{\mathrm{H}}\boldsymbol{z} = \begin{bmatrix}\bar{w}_1 & \dots & \bar{w}_n\end{bmatrix}\begin{bmatrix}z_1\\ \vdots \\ z_n\end{bmatrix} = \bar{w}_1 z_1 + \dots + \bar{w}_n z_n.$$

Note that with complex vectors, $\boldsymbol{w}^{\mathrm{H}}\boldsymbol{z}$ is different from $\boldsymbol{z}^{\mathrm{H}}\boldsymbol{w}$. **The order of the vectors is now important!** In fact, $\boldsymbol{z}^{\mathrm{H}}\boldsymbol{w} = \bar{z}_1 w_1 + \dots + \bar{z}_n w_n$ is the complex conjugate of $\boldsymbol{w}^{\mathrm{H}}\boldsymbol{z}$.

Definition 16.5 — Orthogonal. The two vectors of real or complex are *orthogonal* if their **inner product** is zero.

$$\boldsymbol{z} \perp \boldsymbol{w} \implies \langle \boldsymbol{z},\boldsymbol{w}\rangle = \boldsymbol{w}^{\mathrm{H}}\boldsymbol{z} = 0$$

Example 16.2
Given $\boldsymbol{z} = \begin{pmatrix}1\\ i\end{pmatrix}, \boldsymbol{w} = \begin{pmatrix}-i\\ 1\end{pmatrix}$.
Although we have $\boldsymbol{z}^{\mathrm{T}}\boldsymbol{w} = 0$, the two vectors are not perpendicular.
This is because $\langle \boldsymbol{z},\boldsymbol{w}\rangle = \boldsymbol{w}^{\mathrm{H}}\boldsymbol{z} = \begin{bmatrix}i & 1\end{bmatrix}\begin{bmatrix}1\\ i\end{bmatrix} = 2i \neq 0$.

Example 16.3 The inner product of $\boldsymbol{u} = \begin{bmatrix}1\\ i\end{bmatrix}$ with $\boldsymbol{v} = \begin{bmatrix}i\\ 1\end{bmatrix}$ is $\begin{bmatrix}-i & 1\end{bmatrix}\begin{bmatrix}1\\ i\end{bmatrix} = 0$.
Although those vectors $(1,i)$ and $(i,1)$ don't look perpendicular, actually they are! **A zero inner product still means vectors are orthogonal.**

Proposition 16.1 — Conjugate symmetry.
For two vectors $\boldsymbol{z}$ and $\boldsymbol{w} \in \mathbb{C}^n$, we have $\overline{\langle \boldsymbol{z},\boldsymbol{w}\rangle} = \langle \boldsymbol{w},\boldsymbol{z}\rangle$.

Verify:

$$\begin{aligned}\langle \boldsymbol{z},\boldsymbol{w}\rangle &= \boldsymbol{w}^{\mathrm{H}}\boldsymbol{z} = \bar{\boldsymbol{w}}^{\mathrm{T}}\boldsymbol{z} = \bar{w}_1 z_1 + \dots + \bar{w}_n z_n \\ \langle \boldsymbol{w},\boldsymbol{z}\rangle &= \boldsymbol{z}^{\mathrm{H}}\boldsymbol{w} = \bar{\boldsymbol{z}}^{\mathrm{T}}\boldsymbol{w} = \bar{z}_1 w_1 + \dots + \bar{z}_n w_n\end{aligned}$$

And since we have $\overline{\boldsymbol{w}\boldsymbol{v}} = \bar{\boldsymbol{w}}\bar{\boldsymbol{v}}$ and $\overline{\boldsymbol{w}+\boldsymbol{v}} = \bar{\boldsymbol{w}} + \bar{\boldsymbol{v}}$, it's easy to find that

$$\overline{\bar{w}_1 z_1 + \dots + \bar{w}_n z_n} = w_1 \bar{z}_1 + \dots + w_n \bar{z}_n. = \bar{z}_1 w_1 + \dots + \bar{z}_n w_n.$$

Hence $\overline{\langle \boldsymbol{z},\boldsymbol{w}\rangle} = \langle \boldsymbol{w},\boldsymbol{z}\rangle$. ∎

Proposition 16.2 — Sesquilinear.
For two vectors $\boldsymbol{z}$ and $\boldsymbol{w} \in \mathbb{C}^n$, we have

$$\langle \alpha\boldsymbol{z},\boldsymbol{w}\rangle = \alpha\langle \boldsymbol{z},\boldsymbol{w}\rangle \tag{16.1}$$

$$\langle \boldsymbol{z},\beta\boldsymbol{w}\rangle = \bar{\beta}\langle \boldsymbol{z},\boldsymbol{w}\rangle \tag{16.2}$$

for scalars α and β.

Verify:

$$\begin{aligned}\langle \alpha z, w\rangle &= w^{\mathrm{H}}(\alpha z)\\ &= \alpha(w^{\mathrm{H}}z)\\ &= \alpha\langle z, w\rangle.\end{aligned}$$

For equation (16.2), due to the conjugate symmetry, we derive

$$\langle z, \beta w\rangle = \overline{\langle \beta w, z\rangle}$$

Since $\langle \beta w, z\rangle = \beta\langle w, z\rangle = \beta\overline{\langle z, w\rangle}$, we obtain

$$\langle z, \beta w\rangle = \overline{\beta\overline{\langle z, w\rangle}} = \bar{\beta}\langle z, w\rangle.$$

■

Hermition of matrix

The hermition of a matrix $\boldsymbol{A}$ is given by

$$\boldsymbol{A}^{\mathrm{H}} := \bar{\boldsymbol{A}}^{\mathrm{T}}$$

And the rules for hermition usually comes frim transpose. For example, the hermition has the property

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

$\mathbb{R}^n$	$\mathbb{C}^n$
$\langle x, y\rangle = x^{\mathrm{T}}y$	$\langle z, w\rangle = w^{\mathrm{H}}z$
$x^{\mathrm{T}}y = y^{\mathrm{T}}x$	$z^{\mathrm{H}}w = \overline{w^{\mathrm{H}}z}$
$\|x\|^2 = x^{\mathrm{T}}x$	$\|z\|^2 = z^{\mathrm{H}}z$
$x \perp y \Longleftrightarrow x^{\mathrm{T}}y = 0$	$z \perp w \Longleftrightarrow w^{\mathrm{H}}z = 0$

R *What aspects of eigenvalues/eigenvectors are not nice?*

- Some matrix are *non-diagonalizable*. (or equivalently, eigenvectors don't form a basis.)
- Eigenvalues can be *complex*.

We may ask, what matrix has all real eigenvalues? Let's focus on *real* matrix first. For real symmetric matrix, its eigenvalues are all real!

Proposition 16.3 For a *real symmetric* matrix $\boldsymbol{A}$,

- All eigenvalues are real numbers.
- Its eigenvectors corresponding to distinct eigenvalues are orthogonal.
- $\boldsymbol{A}$ is diagonalizable. More general, all eigenvectors of $\boldsymbol{A}$ are orthogonal!

Before the proof, let's introduce a useful formula: $\langle \boldsymbol{A}x, y\rangle = \langle x, \boldsymbol{A}^{\mathrm{H}}y\rangle$.

Verify: $\langle \boldsymbol{A}x, y\rangle = y^{\mathrm{H}}\boldsymbol{A}x = (\boldsymbol{A}^{\mathrm{H}}y)^{\mathrm{H}}x = \langle x, \boldsymbol{A}^{\mathrm{H}}y\rangle$

Proof.
- For the first part, suppose x is any eigenvectors of $\boldsymbol{A}$ corresponding to eigenvalue λ. Then we obtain

$$\langle \boldsymbol{A}x, x\rangle = \langle x, \boldsymbol{A}^{\mathrm{H}}x\rangle$$

  - For the LHS, $\langle \boldsymbol{A}\boldsymbol{x},\boldsymbol{x}\rangle = \langle \lambda \boldsymbol{x},\boldsymbol{x}\rangle = \lambda \langle \boldsymbol{x},\boldsymbol{x}\rangle$.
  - For the RHS, Since $\boldsymbol{A}$ is a real symmetric matrix, we have $\boldsymbol{A}^{\mathrm{H}} = \bar{\boldsymbol{A}}^{\mathrm{T}} = \boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}$.
    Hence $\langle \boldsymbol{x},\boldsymbol{A}^{\mathrm{H}}\boldsymbol{x}\rangle = \langle \boldsymbol{x},\boldsymbol{A}\boldsymbol{x}\rangle$. Moreover, $\langle \boldsymbol{x},\boldsymbol{A}\boldsymbol{x}\rangle = \langle \boldsymbol{x},\lambda\boldsymbol{x}\rangle = \bar{\lambda}\langle \boldsymbol{x},\boldsymbol{x}\rangle$.
    So we have $\langle \boldsymbol{x},\boldsymbol{A}^{\mathrm{H}}\boldsymbol{x}\rangle = \bar{\lambda}\langle \boldsymbol{x},\boldsymbol{x}\rangle$.

  Finally we have $\lambda\langle \boldsymbol{x},\boldsymbol{x}\rangle = \langle \boldsymbol{A}\boldsymbol{x},\boldsymbol{x}\rangle = \langle \boldsymbol{x},\boldsymbol{A}^{\mathrm{H}}\boldsymbol{x}\rangle = \bar{\lambda}\langle \boldsymbol{x},\boldsymbol{x}\rangle$.
  Since $\boldsymbol{x} \neq \boldsymbol{0}$, $\langle \boldsymbol{x},\boldsymbol{x}\rangle \neq 0$. Hence $\lambda = \bar{\lambda}$. So λ is real.
- For the second part, suppose $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are two eigenvectors corresponding to two **distinct** eigenvalues λ_1 and λ_2 respectively. Our goal is to show $\boldsymbol{x}_1 \perp \boldsymbol{x}_2$. We find that

$$\langle \boldsymbol{A}\boldsymbol{x}_1,\boldsymbol{x}_2\rangle = \langle \boldsymbol{x}_1,\boldsymbol{A}^{\mathrm{H}}\boldsymbol{x}_2\rangle$$

  - For LHS, $\langle \boldsymbol{A}\boldsymbol{x}_1,\boldsymbol{x}_2\rangle = \langle \lambda_1\boldsymbol{x}_1,\boldsymbol{x}_2\rangle = \lambda_1\langle \boldsymbol{x}_1,\boldsymbol{x}_2\rangle$.
  - For RHS, $\langle \boldsymbol{x}_1,\boldsymbol{A}^{\mathrm{H}}\boldsymbol{x}_2\rangle = \langle \boldsymbol{x}_1,\boldsymbol{A}\boldsymbol{x}_2\rangle = \langle \boldsymbol{x}_1,\lambda_2\boldsymbol{x}_2\rangle = \bar{\lambda}_2\langle \boldsymbol{x}_1,\boldsymbol{x}_2\rangle$. Since we have shown that all eigenvalues are real for $\boldsymbol{A}$, we obtain $\bar{\lambda} = \lambda$.
    Hence $\langle \boldsymbol{x}_1,\boldsymbol{A}^{\mathrm{H}}\boldsymbol{x}_2\rangle = \lambda_2\langle \boldsymbol{x}_1,\boldsymbol{x}_2\rangle$.

  Hence $\lambda_1\langle \boldsymbol{x}_1,\boldsymbol{x}_2\rangle = \langle \boldsymbol{A}\boldsymbol{x}_1,\boldsymbol{x}_2\rangle = \langle \boldsymbol{x}_1,\boldsymbol{A}^{\mathrm{H}}\boldsymbol{x}_2\rangle = \lambda_2\langle \boldsymbol{x}_1,\boldsymbol{x}_2\rangle$.
  Since $\lambda_1 \neq \lambda_2$, we obtain $\langle \boldsymbol{x}_1,\boldsymbol{x}_2\rangle = 0$. That is to say, $\boldsymbol{x}_1 \perp \boldsymbol{x}_2$.
- The proof for the third part is not required.

■

16.1.6 Spectral Theorem

We state a theorem without proving it:

Theorem 16.1 — Spectral Theorem. Any real symmetric matrix $\boldsymbol{A}$ has the factorization

$$\boldsymbol{A} = \boldsymbol{Q}\Lambda\boldsymbol{Q}^{\mathrm{T}}, \tag{16.3}$$

where $\Lambda \in \mathbb{R}^{n\times n}$ is diagonal matrix, $\boldsymbol{Q} \in \mathbb{R}^{n\times n}$ is orthogonal matrix.

Proof. From proposition (16.3) we know that $\boldsymbol{A}$ is *diagonalizable*, which means there exists invertible matrix $\boldsymbol{Q}$ and diagonal matrix Λ such that

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

Then we know that all eigenvalues of $\boldsymbol{A}$ are real numbers, so Λ is a real matrix.
Since all eigenvectors $\boldsymbol{x}_1,\dots,\boldsymbol{x}_n$ are orthogonal, matrix $\boldsymbol{Q} = \begin{bmatrix}\boldsymbol{x}_1 & \dots & \boldsymbol{x}_n\end{bmatrix}$ is orthogonal matrix. ■

R $\boldsymbol{A} = \boldsymbol{Q}\Lambda\boldsymbol{Q}^{\mathrm{T}} = \boldsymbol{Q}\Lambda\boldsymbol{Q}^{-1}$. So $\boldsymbol{A}$ could be diagonalized by an orthogonal matrix.
If we let $\boldsymbol{Q} = \begin{bmatrix}q_1 & \dots & q_n\end{bmatrix}$, $\Lambda = \operatorname{diag}(\lambda_1,\dots,\lambda_n)$, then $\boldsymbol{A}$ could be written as:

$$\boldsymbol{A} = \boldsymbol{Q}\Lambda\boldsymbol{Q}^{\mathrm{T}} = \begin{bmatrix}q_1 & \dots & q_n\end{bmatrix}\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}\begin{bmatrix}q_1^{\mathrm{T}} \\ \vdots \\ q_n^{\mathrm{T}}\end{bmatrix}$$

Or equivalently,

$$\boldsymbol{A} = \lambda_1 q_1 q_1^{\mathrm{T}} + \lambda_2 q_2 q_2^{\mathrm{T}} + \dots + \lambda_n q_n q_n^{\mathrm{T}} \tag{16.4}$$

Note that for each term $q_i q_i^{\mathrm{T}}$ is the **projection matrix** to q_i. Hence this theorm says that a real symmetric matrix is a combination of projection matrices.

■ **Example 16.4**
If we write $\boldsymbol{A}$ as combination of projection matrix, we can have a deep understanding for $\boldsymbol{Ax}$:

$$\boldsymbol{A} = \sum_{j=1}^{n} \lambda_j q_j q_j^{\mathrm{T}} \implies \boldsymbol{Ax} = \sum_{j=1}^{n} \lambda_j q_j q_j^{\mathrm{T}} \boldsymbol{x} = \sum_{j=1}^{n} \lambda_j (q_j q_j^{\mathrm{T}} \boldsymbol{x}).$$

If we set $n = 2$, it's clear to find that

$$\boldsymbol{x} = c_1 q_1 + c_2 q_2 \implies \boldsymbol{Ax} = \lambda_1 c_1 q_1 + \lambda_2 c_2 q_2$$

Showing in graph, we have

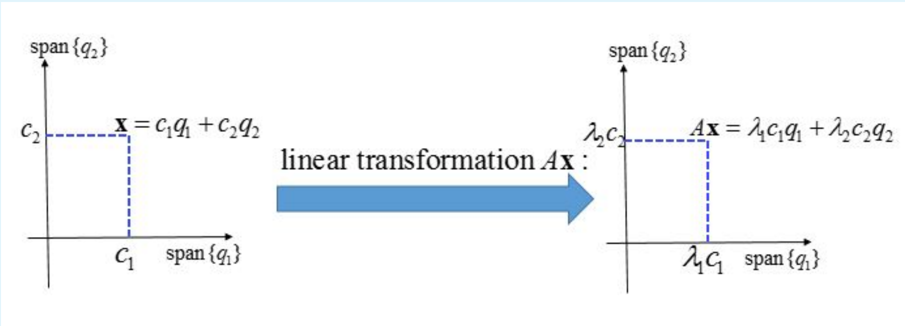

Figure 16.3: Linear transformation of $\boldsymbol{A}$.

The formula $\boldsymbol{A} = \sum_{j=1}^{n} \lambda_j q_j q_j^{\mathrm{T}}$ or $\boldsymbol{A} = \boldsymbol{Q}\Lambda\boldsymbol{Q}^{\mathrm{T}}$ are called **eigendecomposition** or **eigenvalue decomposition**.
And sometimes $\{\lambda_1, \dots, \lambda_n\}$ are called **spectum** of $\boldsymbol{A}$.
Also, we can extend our result from real symmetric matrix into complex:

16.1.7 Hermition matrix

Definition 16.6 — Hermition matrix. A matrix $\boldsymbol{M} \in \mathbb{C}^{n\times n}$ is said to be **Hermition** if $\boldsymbol{M} = \boldsymbol{M}^{\mathrm{H}}$.

Example: $\boldsymbol{M} = \begin{bmatrix} 3 & 2-i \\ 2+i & 4 \end{bmatrix}$ is hermition matrix since $\boldsymbol{M} = \boldsymbol{M}^{\mathrm{H}}$.
If $\boldsymbol{M}$ is a real matrix, then $\boldsymbol{M} = \boldsymbol{M}^{\mathrm{H}} \iff \boldsymbol{M} = \boldsymbol{M}^{\mathrm{T}}$. So if the real matrix is hermition matrix, that is to say it is real symmetric matrix.
Hermition matrix has many interesting properties:

Proposition 16.4 If $\boldsymbol{M} = \boldsymbol{M}^{\mathrm{H}}$, then $\boldsymbol{x}^{\mathrm{H}}\boldsymbol{Mx} \in \mathbb{R}$ for any complex vectors $\boldsymbol{x}$.

Proof. We set $\alpha := \boldsymbol{x}^{\mathrm{H}}\boldsymbol{Mx}$. Since α is a number (easy to check), we obtain $\alpha^{\mathrm{T}} = \alpha$.
Thus $\bar{\alpha} = \alpha^{\mathrm{H}} = (\boldsymbol{x}^{\mathrm{H}}\boldsymbol{Mx})^{\mathrm{H}} = \boldsymbol{x}^{\mathrm{H}}\boldsymbol{Mx} = \alpha$.
Hence α is real. ■

Proposition 16.5 If $\boldsymbol{M} = \boldsymbol{M}^{\mathrm{H}}$, then $\langle \boldsymbol{x}, \boldsymbol{My}\rangle = \langle \boldsymbol{Mx}, \boldsymbol{y}\rangle$.

Proof. By definition,

$$\langle \boldsymbol{x}, \boldsymbol{My}\rangle = (\boldsymbol{My})^{\mathrm{H}}\boldsymbol{x} = \boldsymbol{y}^{\mathrm{H}}\boldsymbol{M}^{\mathrm{H}}\boldsymbol{x} = \boldsymbol{y}^{\mathrm{H}}\boldsymbol{Mx} = \langle \boldsymbol{Mx}, \boldsymbol{y}\rangle.$$

■

And we have general orthogonal matrices for complex matrices:

Definition 16.7 — Unitary. A unitary matrix is a complex matrix that has **orthonormal columns**. In other words, $\boldsymbol{U}$ is unitary if $\boldsymbol{U}^{\mathrm{H}}\boldsymbol{U} = \boldsymbol{I}$.

And the spectral theorm can also apply for Hermition matrix:

Theorem 16.2 Any Hermition matrix $\boldsymbol{M}$ can be factorized into

$$\boldsymbol{M} = \boldsymbol{U}\Lambda\boldsymbol{U}^{\mathrm{H}}$$

where Λ is a real diagonal matrix, $\boldsymbol{U}$ is a complex unitary matrix.

R What good points does Hermition matrix has?
- It is diagonalizable.
- Its eigenvectors form orthogonal basis.
- Its eigenvalues are all real.