5 Eigenvalues and Eigenvectors（特征值与特征向量）

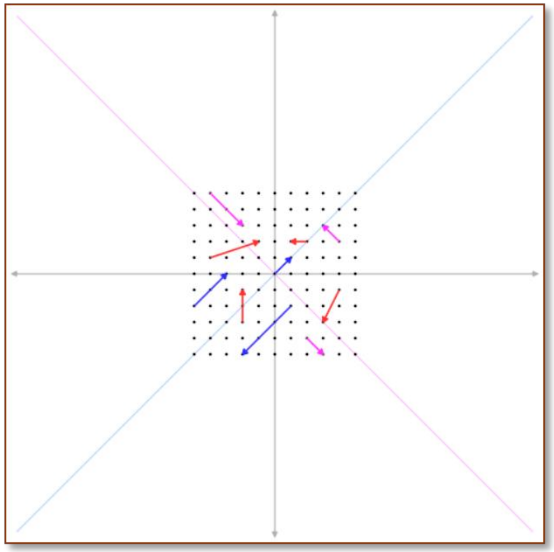
5.5

COMPLEX MATRICES

Operations in the Complex Case

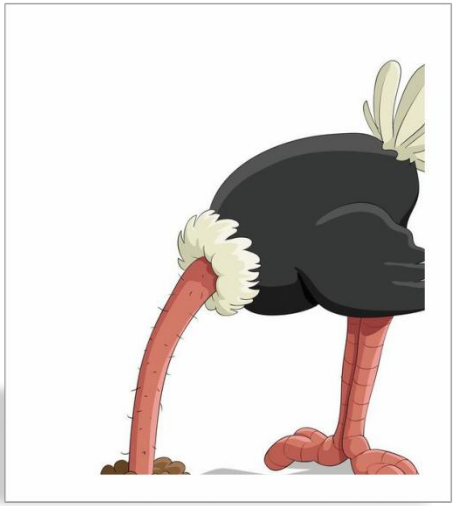
Hermitian Matrices

Unitary Matrices



Since $|\boldsymbol{A} - \lambda \boldsymbol{I}|$ is a polynomial of degree n, the equation always has exactly n roots, counting multiplicities, *provided that possibly complex roots are included.*

A real matrix has real coefficients in $|\boldsymbol{A} - \lambda \boldsymbol{I}|$, but the eigenvalues (as in rotations) may be *complex*.



We cannot avoid complex numbers and vectors any more.

The key is to let $\boldsymbol{A}$ act on the space $\mathbf{C}^n$.

The new definitions coincide with the old when the vectors and matrices are *real*.

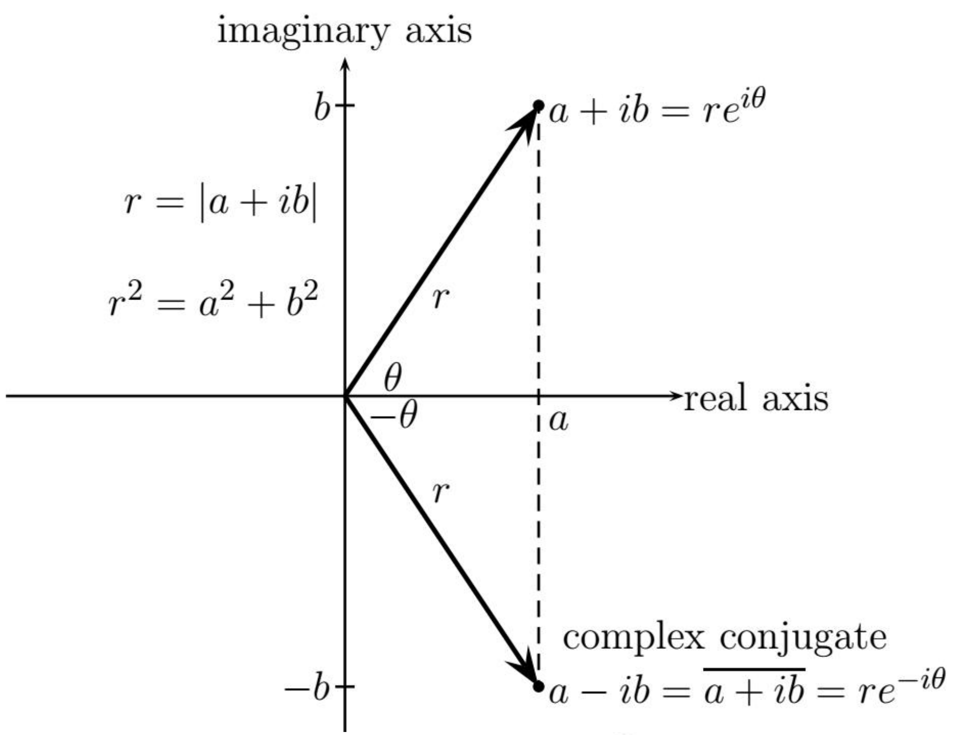
Main results:

1. ***Every real symmetric matrix (and Hermitian matrix) has real eigenvalues.***
2. ***Its eigenvectors can be chosen to be orthonormal.***

I. Some Definitions in the Complex Case

(1) Take a complex number $z = a + ib$, where $i = \sqrt{-1}$.

- Conjugate (共轭) $\bar{z} = a - ib$.
- Absolute value $r = |z| = \sqrt{a^2 + b^2}$.
- Polar form: $a + ib = r(\cos\theta + i\sin\theta) = r\mathbf{e}^{i\theta}$.



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

Complex addition

$$(a + ib) + (c + id)$$
$$= (a + c) + i(b + d).$$

Multiplication

$$(a + ib)(c + id)$$
$$= (ac - bd) + i(bc + ad).$$

Three important properties:

1. The conjugate of a product equals the product of the conjugates:

$$\overline{(a+ib)(c+id)} = (ac-bd)-i(bc+ad)$$
$$= \overline{(a+ib)}\,\overline{(c+id)}.$$

2. The conjugate of a sum equals the sum of the conjugates:

$$\overline{(a+c)+i(b+d)} = (a+c)-i(b+d)$$
$$= \overline{(a+ib)}+\overline{(c+id)}.$$

3. Multiplying any $a+ib$ by its conjugate $a-ib$ produces a real number a^2+b^2:

$$(a+ib)(a-ib) = a^2+b^2 = r^2.$$

This distance r is the ***absolute value*** $|a+ib| = \sqrt{a^2+b^2}$.

For example,

$x = 3 + 4i$ times its conjugate $\bar{x} = 3 - 4i$ is the absolute value squared:

$$x\bar{x} = (3 + 4i)(3 - 4i) = 25 = |x|^2,$$

so $r = |x| = 5$.

To divide by $3 + 4i$, multiply numerator and denominator by its conjugate $3 - 4i$:

$$\frac{2+i}{3+4i} = \frac{2+i}{3+4i}\,\frac{3-4i}{3-4i} = \frac{10-5i}{25}.$$

In *polar coordinates*(极坐标), multiplication and division are easy:

$re^{i\theta}$ times $Re^{i\alpha}$ has absolute value rR and angle $\theta + \alpha$.

$re^{i\theta}$ divided by $Re^{i\alpha}$ has absolute value r/R and angle $\theta - \alpha$.

(2) Pick a complex vector $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbf{C}^n$ (*the complex vector space containing all vectors with n complex components*), where $x_j = a_j + \boldsymbol{i} b_j$.

- Vector addition: $\boldsymbol{x} + \boldsymbol{y} = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)^{\mathrm{T}}$.
- Scalar multiplication: $c\boldsymbol{x}$, $c \in \mathbf{C}$.
- The vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ are linearly *dependent* if some *nontrivial* combination gives $c_1\boldsymbol{v}_1 + \ldots + c_k\boldsymbol{v}_k = \mathbf{0}$; the c_j may now be complex.
- $\mathbf{C}^n$ is a complex vector space of dimension n. (The unit coordinate vectors are still in $\mathbf{C}^n$; they are still independent; and they still form a basis.)

For $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{C}^n$,

- The length squared $\|\boldsymbol{x}\|^2 = |x_1|^2 + |x_2|^2 + \cdots + |x_n|^2$.
- The conjugate: $\bar{\boldsymbol{x}} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)^{\mathrm{T}}$.
- Inner product: $\bar{\boldsymbol{x}}^{\mathrm{T}} \boldsymbol{y} = \bar{x}_1 y_1 + \cdots + \bar{x}_n y_n$.

  In particular, $\bar{\boldsymbol{x}}^{\mathrm{T}} \boldsymbol{x} = \bar{x}_1 x_1 + \cdots + \bar{x}_n x_n = \|\boldsymbol{x}\|^2$.

Attention: Length is computed differently.

The inner product, the definitions of symmetric and orthogonal matrices, all need to be modified for complex numbers.

For example,

$\boldsymbol{x} = \begin{bmatrix} 1 \\ i \end{bmatrix}$, then $\|\boldsymbol{x}\|^2 = 2$.

$\boldsymbol{y} = \begin{bmatrix} 2+i \\ 2-4i \end{bmatrix}$, then $\|\boldsymbol{y}\|^2 = 25$.

Also $\bar{\boldsymbol{y}}^{\mathrm{T}} \boldsymbol{y} = \overline{(2+i)}(2+i) + \overline{(2-4i)}(2-4i) = 5 + 20 = 25$.

(3) Let $\boldsymbol{A} = [a_{ij}]_{m\times n}$ be a complex matrix.

– The conjugate: $\overline{\boldsymbol{A}} = [\bar{a}_{ij}]_{m\times n}$.

– The conjugate transpose (共轭转置) : $\overline{\boldsymbol{A}}^{\mathrm{T}} = [\bar{a}_{ji}]_{n\times m}$,

called '***A* Hermitian**' (*A*的厄米特矩阵), denoted by $\boldsymbol{A}^H$.

(Instead of a bar for the conjugate and a T for the transpose, a superscript H combines both operations)

For example, $\begin{bmatrix} 2+i & 3i \\ 4-i & 5 \\ 0 & 0 \end{bmatrix}^H = \begin{bmatrix} 2-i & 4+i & 0 \\ -3i & 5 & 0 \end{bmatrix}$.

- For $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, $\boldsymbol{x}^H = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$.
- Inner product $\overline{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{y}$ can also be written as $\boldsymbol{x}^H\boldsymbol{y}$.
- Orthogonal vectors have $\boldsymbol{x}^H\boldsymbol{y} = 0$.
- The squared length of $\boldsymbol{x}$ is $\boldsymbol{x}^H\boldsymbol{x}$.

Remark We note that $(\boldsymbol{AB})^H = \boldsymbol{B}^H\boldsymbol{A}^H$, and $(\boldsymbol{A}^H)^H = \boldsymbol{A}$.

II. Hermitian Matrices and Properties

Real cases: Symmetric matrices: $\boldsymbol{A} = \boldsymbol{A}^{\mathrm{T}}$.

With complex entries, this idea of symmetry has to be extended.

Generalization: ***matrices that equal their conjugate transpose***.

Definition 1 A matrix $\boldsymbol{A}$ is called a **Hermitian matrix** (A是厄米特矩阵) if $\boldsymbol{A}^H = \boldsymbol{A}$. (即满足：$\boldsymbol{A}$的共轭转置矩阵等于它本身)

For example,

$$\boldsymbol{A} = \begin{bmatrix} 2 & 3-3i \\ 3+3i & 5 \end{bmatrix} = \boldsymbol{A}^H,$$ so $\boldsymbol{A}$ is a Hermitian matrix.

The diagonal entries must be real; Each off-diagonal entry is matched with its mirror image across the main diagonal .

Remark *A real symmetric matrix is certainly Hermitian*. (For real matrices there is no difference between $\boldsymbol{A}^{\mathrm{T}}$ and $\boldsymbol{A}^H$.)

Property 1 *If* $\boldsymbol{A} = \boldsymbol{A}^H$, *then for all complex vectors* $\boldsymbol{x}$, *the number* $\boldsymbol{x}^H\boldsymbol{A}\boldsymbol{x}$ *is real.*

Proof. Notice that $\boldsymbol{x}^H\boldsymbol{A}\boldsymbol{x}$ is a number, and

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

That is to say, $\boldsymbol{x}^H\boldsymbol{A}\boldsymbol{x}$ is a number which is equal to its conjugate.

So $\boldsymbol{x}^H\boldsymbol{A}\boldsymbol{x}$ is a real number.

Property 2 *If* $\boldsymbol{A} = \boldsymbol{A}^H$, *then every eigenvalue is a real number.*

Proof. Let $\boldsymbol{A}$ be a Hermitian matrix, and assume $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$.

Then $\boxed{\boldsymbol{x}^H\boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}^H\lambda\boldsymbol{x} = \lambda\boldsymbol{x}^H\boldsymbol{x} = \lambda\|\boldsymbol{x}\|^2.}$

By Property 1, $\boldsymbol{x}^H\boldsymbol{A}\boldsymbol{x}$ is real.

And since $\boldsymbol{x} \neq \boldsymbol{0}$, $\|\boldsymbol{x}\|^2$ is real and positive,

thus $\lambda = \dfrac{\boldsymbol{x}^H\boldsymbol{A}\boldsymbol{x}}{\|\boldsymbol{x}\|^2}$, and so λ is a real number.

Property 2 *If* $\boldsymbol{A} = \boldsymbol{A}^H$, *then every eigenvalue is a real number.*

A Second Proof. (without using Property 1)

$$\bar{\lambda}\boldsymbol{x}^H\boldsymbol{x} = (\lambda\boldsymbol{x})^H\boldsymbol{x} = (\boldsymbol{A}\boldsymbol{x})^H\boldsymbol{x} = \boldsymbol{x}^H\boldsymbol{A}^H\boldsymbol{x} = \boldsymbol{x}^H\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}^H\boldsymbol{x}.$$

So $\bar{\lambda} = \lambda$, and λ is a real number.

For example,

$\boldsymbol{A} = \begin{bmatrix} 2 & 3-3i \\ 3+3i & 5 \end{bmatrix} = \boldsymbol{A}^H$, so $\boldsymbol{A}$ is a Hermitian matrix.

Then

$|\boldsymbol{A} - \lambda\boldsymbol{I}| = \begin{vmatrix} 2-\lambda & 3-3i \\ 3+3i & 5-\lambda \end{vmatrix}$

$= \lambda^2 - 7\lambda + 10 - |3-3i|^2$

$= \lambda^2 - 7\lambda - 8 = (\lambda - 8)(\lambda + 1).$

The eigenvalues of $\boldsymbol{A}$ are 8 and -1.

Property 3 *Let $\boldsymbol{A}$ be a Hermitian matrix (i.e., $\boldsymbol{A} = \boldsymbol{A}^H$), and λ_1, λ_2 be two different eigenvalues of $\boldsymbol{A}$. Then the eigenvectors corresponding to λ_1, λ_2 are orthogonal to each other.*

In particular, this is true for real symmetric matrices.

Proof. Let $\boldsymbol{x}_1,\ \boldsymbol{x}_2$ be the eigenvectors of $\boldsymbol{A}$ corresponding to λ_1, λ_2, respectively. Then

$$\boldsymbol{A}\boldsymbol{x}_1 = \lambda_1 \boldsymbol{x}_1, \text{ and } \qquad \boldsymbol{A}\boldsymbol{x}_2 = \lambda_2 \boldsymbol{x}_2.$$

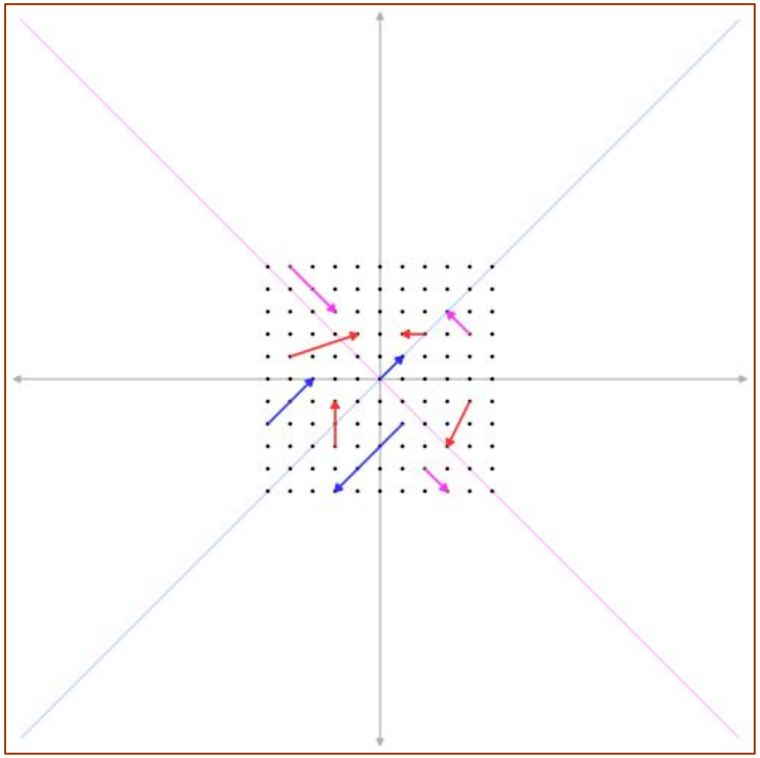
Hence

$$\lambda_1 \boldsymbol{x}_1^H \boldsymbol{x}_2 = (\lambda_1 \boldsymbol{x}_1)^H \boldsymbol{x}_2 = (\boldsymbol{A}\boldsymbol{x}_1)^H \boldsymbol{x}_2 = \boldsymbol{x}_1^H \boldsymbol{A}^H \boldsymbol{x}_2$$
$$= \boldsymbol{x}_1^H \boldsymbol{A} \boldsymbol{x}_2 = \boldsymbol{x}_1^H \lambda_2 \boldsymbol{x}_2 = \lambda_2 \boldsymbol{x}_1^H \boldsymbol{x}_2.$$

Since $\lambda_1 \neq \lambda_2$, we conclude that $\boldsymbol{x}_1^H \boldsymbol{x}_2 = 0$, and $\boldsymbol{x}_1,\ \boldsymbol{x}_2$ are orthogonal.

Recall Example 3 in § 5.1 :

Find the eigenvalues and eigenvectors of the matrix $\boldsymbol{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$.



http://en.wikipedia.org/wiki/File:Eigenvectors-extended.gif

The *Eigenvectors*

$k_1\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ $\quad$ $k_2\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$(k_1 \neq 0)$ $\quad$ $(k_2 \neq 0)$

corresponding respectively to

the *Eigenvalues*:

$\lambda_1 = 1$ $\quad$ $\lambda_2 = 3$

For example,

$$\boldsymbol{A}=\begin{bmatrix} 2 & 3-3i \\ 3+3i & 5 \end{bmatrix}=\boldsymbol{A}^H\text{, so } \boldsymbol{A} \text{ is a Hermitian matrix.}$$

The eigenvalues of $\boldsymbol{A}$ are 8 and -1.

$$(\boldsymbol{A}-8\boldsymbol{I})\boldsymbol{x}=\begin{bmatrix} -6 & 3-3i \\ 3+3i & -3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{x}=\begin{bmatrix} 1 \\ 1+i \end{bmatrix}.$$

$$(\boldsymbol{A}+\boldsymbol{I})\boldsymbol{y}=\begin{bmatrix} 3 & 3-3i \\ 3+3i & 6 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{y}=\begin{bmatrix} 1-i \\ -1 \end{bmatrix}.$$

The two eigenvectors are orthogonal :

$$\boldsymbol{x}^H\boldsymbol{y}=\begin{bmatrix} 1 & 1-i \end{bmatrix}\begin{bmatrix} 1-i \\ -1 \end{bmatrix}=0.$$

The next is one of the great theorems in linear algebra.

Theorem 1 (*Spectral Theorem*, part I) *A real symmetric matrix* (实对称矩阵) $\boldsymbol{A}$ *can be factored into*

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

Its orthonormal eigenvectors are in the orthogonal matrix $\boldsymbol{Q}$ *and its eigenvalues are in* $\boldsymbol{\Lambda}$.

Proof. (We only prove this for $\boldsymbol{A}$ with distinct eigenvalues.)

Let $\boldsymbol{Q}$ be the matrix with columns being n eigenvectors of $\boldsymbol{A}$ which are orthonormal. Then

$$\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q} = \boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n),$$

and so $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{-1} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{\mathrm{T}}$.

(*Even with repeated eigenvalues, a symmetric matrix still has a complete set of orthonormal eigenvectors.* – Next section)

Remark 1

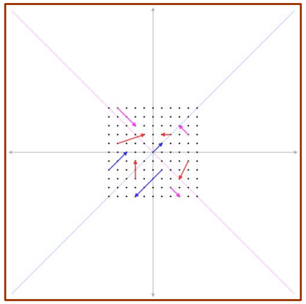
$$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^T = [\boldsymbol{x}_1 \quad \cdots \quad \boldsymbol{x}_n]\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}\begin{bmatrix} \boldsymbol{x}_1^{\mathrm{T}} \\ \vdots \\ \boldsymbol{x}_n^{\mathrm{T}} \end{bmatrix}$$

$$= \lambda_1 \boldsymbol{x}_1 \boldsymbol{x}_1^{\mathrm{T}} + \lambda_2 \boldsymbol{x}_2 \boldsymbol{x}_2^{\mathrm{T}} + \ldots + \lambda_n \boldsymbol{x}_n \boldsymbol{x}_n^{\mathrm{T}}.$$

So $\boldsymbol{A}$ becomes a combination of one-dimensional projections—which are the special matrices $\boldsymbol{x}_i \boldsymbol{x}_i^{\mathrm{T}}$ of rank 1, multiplied by λ_i.

Example 1 $\boldsymbol{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ has eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 3$.

The eigenvectors, with length scaled to 1, are



$$\boldsymbol{x}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix} \text{ and } \boldsymbol{x}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

$$\boldsymbol{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} + 3\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \lambda_1 \boldsymbol{x}_1 \boldsymbol{x}_1^{\mathrm{T}} + \lambda_2 \boldsymbol{x}_2 \boldsymbol{x}_2^{\mathrm{T}}.$$

— ~~combination of two one-dimensional projections.~~

Remark 2 If $\boldsymbol{A}$ is *real* and its eigenvalues *happen to be real*, then its eigenvectors are also real.

(solve $(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{x}=\boldsymbol{0}$ and compute by elimination.)

But they will not be orthogonal unless $\boldsymbol{A}$ is symmetric:

$\boldsymbol{A}=\ \boldsymbol{Q\Lambda Q}^{\mathrm{T}}$ leads to $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$.

Remark 3 If $\boldsymbol{A}$ is *real*, all complex eigenvalues come in conjugate pairs: $\boldsymbol{Ax}=\lambda\boldsymbol{x}$ and $\boldsymbol{A}\overline{\boldsymbol{x}}=\overline{\lambda}\overline{\boldsymbol{x}}$.

(This is true because $\boldsymbol{A}\overline{\boldsymbol{x}}=\overline{\boldsymbol{Ax}}=\overline{\lambda\boldsymbol{x}}=\overline{\lambda}\overline{\boldsymbol{x}}$)

Hence $\overline{\lambda}$ is also an eigenvalue of $\boldsymbol{A}$, with $\overline{\boldsymbol{x}}$ a corresponding eigenvector.

If $a+ib$ *is an eigenvalue of a real matrix, so is* $a-ib$.

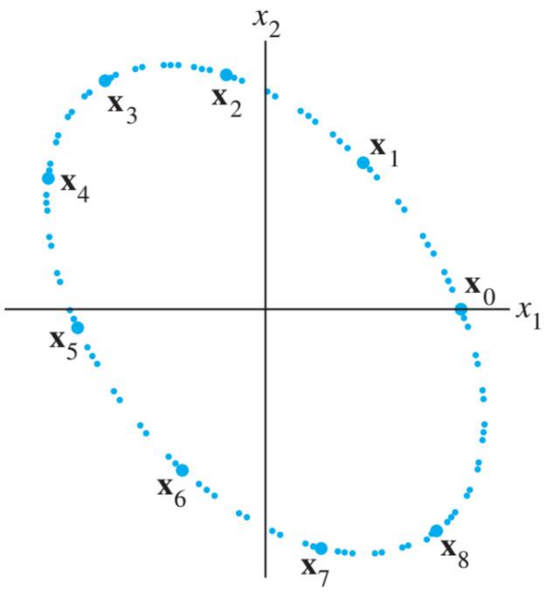
(Note: If $\boldsymbol{A}=\boldsymbol{A}^{\mathrm{T}}$ then $b=0$. 实对称矩阵的特征值都是实数)

For example, $A = \begin{bmatrix} 0.5 & -0.6 \\ 0.75 & 1.1 \end{bmatrix}$

The eigenvalues are $\lambda_{1,2} = 0.8 \pm 0.6i$.

The basis for the eigenspace corresponding to λ_1 and λ_2 are

$\boldsymbol{v}_1 = \begin{bmatrix} -2+4i \\ 5 \end{bmatrix}$ and $\boldsymbol{v}_2 = \begin{bmatrix} -2-4i \\ 5 \end{bmatrix}$.



Iterates of a point $\boldsymbol{x}_0$ under the action of a matrix with a complex eigenvalue

One way to see how multiplication by $\boldsymbol{A}$ affects points is to plot an arbitrary initial point – say, $\boldsymbol{x}_0 = (2,0)^T$ – and then to plot

$\boldsymbol{x}_1 = \boldsymbol{A}\boldsymbol{x}_0, \ \boldsymbol{x}_2 = \boldsymbol{A}\boldsymbol{x}_1, \ \boldsymbol{x}_3 = \boldsymbol{A}\boldsymbol{x}_2, \ldots$

Example 2 Let $\boldsymbol{A} = \begin{bmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{bmatrix}$.

Find an orthogonal matrix $\boldsymbol{Q}$ such that $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}$ is a diagonal matrix.

Solution

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = \begin{vmatrix} 2-\lambda & 2 & -2 \\ 2 & 5-\lambda & -4 \\ -2 & -4 & 5-\lambda \end{vmatrix} = \begin{vmatrix} 2-\lambda & 2 & -2 \\ 0 & 1-\lambda & 1-\lambda \\ -2 & -4 & 5-\lambda \end{vmatrix}$$

$$= \begin{vmatrix} 2-\lambda & 2 & -4 \\ 0 & 1-\lambda & 0 \\ -2 & -4 & 9-\lambda \end{vmatrix} = (1-\lambda)\begin{vmatrix} 2-\lambda & -4 \\ -2 & 9-\lambda \end{vmatrix}$$

$$= (1-\lambda)^2(10-\lambda).$$

So the eigenvalues of $\boldsymbol{A}$ are $\lambda_1 = 1$ (Algebraic multiplicity is 2) and $\lambda_2 = 10$ (Algebraic multiplicity is 1).

For the eigenvalue λ_1=1, by $(\boldsymbol{A}-\lambda_1\boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}$, i.e.,

$$\begin{bmatrix} 1 & 2 & -2 \\ 2 & 4 & -4 \\ -2 & -4 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The basis for the eigenspace of λ_1 : $\boldsymbol{x}_1 = (-2, 1, 0)^{\mathrm{T}}$, $\boldsymbol{x}_2 = (2, 0, 1)^{\mathrm{T}}$.

By Gram-Schmidt orthogonalization, let

$\boldsymbol{\beta}_1 = \boldsymbol{x}_1 = (-2, 1, 0)^{\mathrm{T}}$,

$$\boldsymbol{\beta}_2 = \boldsymbol{x}_2 - \frac{\boldsymbol{x}_2^{\mathrm{T}}\boldsymbol{\beta}_1}{\boldsymbol{\beta}_1^{\mathrm{T}}\boldsymbol{\beta}_1}\boldsymbol{\beta}_1 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} - \frac{-4}{5}\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{2}{5} \\ \frac{4}{5} \\ 1 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix},$$

and normalize $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$ into:

$$\gamma_1 = \frac{\boldsymbol{\beta}_1}{\|\boldsymbol{\beta}_1\|} = \frac{\sqrt{5}}{5}[-2, \ 1, \ 0]^{\mathrm{T}}, \qquad \gamma_2 = \frac{\boldsymbol{\beta}_2}{\|\boldsymbol{\beta}_2\|} = \frac{\sqrt{5}}{15}[2, \ 4, \ 5]^{\mathrm{T}}.$$

For the eigenvalue λ_2=10, by $(\boldsymbol{A}-\lambda_2\boldsymbol{I})\boldsymbol{x}=\boldsymbol{0}$, i.e., $\begin{bmatrix} -8 & 2 & -2 \\ 2 & -5 & -4 \\ -2 & -4 & -5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

We can get $\boldsymbol{x}_3$ =(1, 2, – 2)$^{\mathrm{T}}$ and the corresponding unit vector:

$$\gamma_3 = \tfrac{1}{3}\begin{bmatrix} 1, & 2, & -2 \end{bmatrix}^{\mathrm{T}}.$$

Take the orthogonal matrix

$$\boldsymbol{Q}=\left[\gamma_1,\gamma_2,\gamma_3\right]=\begin{bmatrix} -2\sqrt{5}/5 & 2\sqrt{5}/15 & 1/3 \\ \sqrt{5}/5 & 4\sqrt{5}/15 & 2/3 \\ 0 & \sqrt{5}/3 & -2/3 \end{bmatrix}$$

which will make

$$\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q} = \mathrm{diag}(\lambda_1,\ \lambda_2,\ \lambda_3) = \mathrm{diag}(1,\ 1,\ 10).$$

III. Unitary Matrices

A *real* orthogonal matrix— $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Q} = \boldsymbol{I}$.

For *complex* matrices, the transpose will be replaced by the conjugate transpose. The condition will become $\boldsymbol{U}^{H}\boldsymbol{U} = \boldsymbol{I}$.

The new letter $\boldsymbol{U}$ reflects the new name: ***A complex matrix with orthonormal columns is called a unitary matrix.***

Definition 2 A matrix $\boldsymbol{U}$ is called a **unitary matrix** (酉矩阵) if $\boldsymbol{U}^{H} = \boldsymbol{U}^{-1}$.

Equivalently, $\boldsymbol{U}^{H}\boldsymbol{U} = \boldsymbol{I}$, and $\boldsymbol{U}\boldsymbol{U}^{H} = \boldsymbol{I}$.

Unitary matrices have many nice properties.

Theorem 2 Let $\boldsymbol{U}$ be a unitary matrix. Then the following hold.

1. *Inner products and lengths are preserved by* $\boldsymbol{U}$*.*

Proof. $(\boldsymbol{Ux})^{\mathrm{H}}(\boldsymbol{Uy}) = \boldsymbol{x}^{\mathrm{H}}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{Uy} = \boldsymbol{x}^{\mathrm{H}}\boldsymbol{y}$, and $\|\boldsymbol{Ux}\|^2 = \|\boldsymbol{x}\|^2$.

2. *Every eigenvalue of* $\boldsymbol{U}$ *has absolute value* $|\lambda| = 1$*.*

Proof. This follows directly from $\boldsymbol{Ux} = \lambda\boldsymbol{x}$, by comparing the lengths of the two sides: $\|\boldsymbol{Ux}\| = \|\boldsymbol{x}\|$, and always $\|\lambda\boldsymbol{x}\| = |\lambda| \cdot \|\boldsymbol{x}\|$.

Therefore $|\lambda| = 1$.

3. *Eigenvectors of* $\boldsymbol{U}$ *corresponding to different eigenvalues are orthogonal (and can be scaled to orthonormal).*

Proof. Start with $\boldsymbol{Ux} = \lambda_1\boldsymbol{x}$ and $\boldsymbol{Uy} = \lambda_2\boldsymbol{y}$, and take inner products:

$$\boldsymbol{x}^{\mathrm{H}}\boldsymbol{y} = (\boldsymbol{Ux})^{\mathrm{H}}(\boldsymbol{Uy}) = (\lambda_1\boldsymbol{x})^{\mathrm{H}}(\lambda_2\boldsymbol{y}) = \overline{\lambda_1}\lambda_2\boldsymbol{x}^{\mathrm{H}}\boldsymbol{y}.$$

Comparing the left to the right, $\overline{\lambda_1}\lambda_2 = 1$ or $\boldsymbol{x}^{\mathrm{H}}\boldsymbol{y} = 0$. But Property 2 is $\overline{\lambda_1}\lambda_1 = 1$, so we cannot also have $\overline{\lambda_1}\lambda_2 = 1$. Thus $\boldsymbol{x}^{\mathrm{H}}\boldsymbol{y} = 0$ and the eigenvectors are orthogonal (*and can be scaled to unit length*).

Check the properties by working on the following matrices.

Example 3 $\boldsymbol{U} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ has eigenvalues $\mathbf{e}^{i\theta}$ and $\mathbf{e}^{-i\theta}$.

The orthonormal eigenvectors are $\boldsymbol{x}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -i \end{bmatrix}$ and $\boldsymbol{x}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ i \end{bmatrix}$.

Example 4 $\boldsymbol{P} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$ has eigenvalues $-1,\ i,\ -i,\ 1$.

The orthonormal eigenvectors are $\begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} \\ \frac{i}{2} \\ -\frac{1}{2} \\ -\frac{i}{2} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} \\ -\frac{i}{2} \\ -\frac{1}{2} \\ \frac{i}{2} \end{bmatrix}, \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}$.

Let $\boldsymbol{A}$ be a matrix of degree n. Let λ be an eigenvalue of $\boldsymbol{A}$.

The *eigenspace* V_λ is the subspace spanned by the eigenvectors of $\boldsymbol{A}$ corresponding to λ.

By Gram-Schmidt procedure, an eigenspace has an orthonormal basis.

Lemma 1 *A Hermitian matrix has a complete set of ~~orthonormal eigenvectors.~~* (more discussions in Section 5.6)

Remark Assume that $\boldsymbol{A}$ is Hermitian. From each eigenspace of $\boldsymbol{A}$, choose an orthonormal basis by Gram-Schmidt process.

Since any two vectors corresponding to different eigenvalues are orthogonal, the eigenvectors in these orthonormal bases are orthonormal, i.e., there are n eigenvectors of $\boldsymbol{A}$ which are orthonormal.

Let $\boldsymbol{A}$ be a Hermitian matrix of degree n, and let $\boldsymbol{v}_1,\ \boldsymbol{v}_2,\ .\ .\ .\ ,\ \boldsymbol{v}_n$ be a complete set of orthonormal eigenvectors, corresponding to eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$ respectively.

Let $\boldsymbol{U} = [\boldsymbol{v}_1,\ \boldsymbol{v}_2,\ .\ .\ .\ ,\ \boldsymbol{v}_n]$, then $\boldsymbol{U}$ is a unitary matrix, and

$$\begin{aligned} \boldsymbol{AU} &= \boldsymbol{A}\,[\boldsymbol{v}_1\ \boldsymbol{v}_2\ .\ .\ .\ \boldsymbol{v}_n] = [\boldsymbol{Av}_1\ \boldsymbol{Av}_2\ .\ .\ .\ \boldsymbol{Av}_n] \\ &= [\lambda_1\boldsymbol{v}_1\ \ \lambda_2\boldsymbol{v}_2\ \ \dots\ \lambda_n\boldsymbol{v}_n] = \boldsymbol{U}\mathrm{diag}(\lambda_1, \lambda_2, ..., \lambda_n). \end{aligned}$$

Thus $\boldsymbol{U}$ diagonalizes $\boldsymbol{A}$: $\boldsymbol{U}^{-1}\boldsymbol{AU} = \mathrm{diag}(\lambda_1, \lambda_2, ..., \lambda_n)$.

This gives the following important theorem.

Theorem 3 (*Spectral Theorem*)

(1) *Each real symmetric matrix* $\boldsymbol{A}$ *can be diagonalized by an orthogonal matrix* $\boldsymbol{Q}$.

(2) *Every Hermitian matrix* $\boldsymbol{A}$ *can be diagonalized by a unitary matrix* $\boldsymbol{U}$.

The columns of $\boldsymbol{Q}$ *(or* $\boldsymbol{U}$*) consist of orthonormal eigenvectors of* $\boldsymbol{A}$.

Skew-Hermitian matrices

- Skew-symmetric matrices satisfy $\boldsymbol{K}^{\mathrm{T}} =- \boldsymbol{K}$.
- Skew-Hermitian matrices (反厄米特矩阵) satisfy $\boldsymbol{K}^{\mathrm{H}} =- \boldsymbol{K}$.

Property If $\boldsymbol{A}$ is Hermitian then $\boldsymbol{K} = i\boldsymbol{A}$ is skew-Hermitian. (i.e., If $\boldsymbol{A} = \boldsymbol{A}^{\mathrm{H}}$, and $\boldsymbol{K} = i\boldsymbol{A}$, then $\boldsymbol{K}^{\mathrm{H}} =- \boldsymbol{K}$.)

Remark The eigenvalues of $\boldsymbol{K}$ are purely imaginary instead of purely real (反厄米特矩阵的特征值是纯虚数). For example,

$$\boldsymbol{A} = \begin{bmatrix} 2 & 3-3i \\ 3+3i & 5 \end{bmatrix} = \boldsymbol{A}^{\mathrm{H}}$$, so $\boldsymbol{A}$ is a Hermitian matrix.

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

The diagonal entries are multiples of i (allowing zero).

The eigenvalues are $8i$ and $-i$.

The eigenvectors are still orthogonal, and we still have $\boldsymbol{K} = \boldsymbol{U\Lambda U}^{\mathrm{H}}$ — with a unitary $\boldsymbol{U}$ instead of a real orthogonal $\boldsymbol{Q}$, and with $8i$ and $-i$ on the diagonal of $\boldsymbol{\Lambda}$.

Real versus Complex

$\mathbf{R}^n$ (n real components)	$\leftrightarrow$	$\mathbf{C}^n$ (n complex components)
length: $\lVert x\rVert^2 = x_1^2 + \cdots + x_n^2$	$\leftrightarrow$	length: $\lVert x\rVert^2 = \lvert x_1\rvert^2 + \cdots + \lvert x_n\rvert^2$
transpose: $A_{ij}^{\mathrm{T}} = A_{ji}$	$\leftrightarrow$	Hermitian transpose: $A_{ij}^{\mathrm{H}} = \overline{A_{ji}}$
$(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$	$\leftrightarrow$	$(AB)^{\mathrm{H}} = B^{\mathrm{H}}A^{\mathrm{H}}$
inner product: $x^{\mathrm{T}}y = x_1y_1 + \cdots + x_ny_n$	$\leftrightarrow$	inner product: $x^{\mathrm{H}}y = \bar{x}_1y_1 + \cdots + \bar{x}_ny_n$
$(Ax)^{\mathrm{T}}y = x^{\mathrm{T}}(A^{\mathrm{T}}y)$	$\leftrightarrow$	$(Ax)^{\mathrm{H}}y = x^{\mathrm{H}}(A^{\mathrm{H}}y)$
orthogonality: $x^{\mathrm{T}}y = 0$	$\leftrightarrow$	orthogonality: $x^{\mathrm{H}}y = 0$
symmetric matrices: $A^{\mathrm{T}} = A$	$\leftrightarrow$	Hermitian matrices: $A^{\mathrm{H}} = A$
$A = Q\Lambda Q^{-1} = Q\Lambda Q^{\mathrm{T}}$ (real Λ)	$\leftrightarrow$	$A = U\Lambda U^{-1} = U\Lambda U^{\mathrm{H}}$ (real Λ)
skew-symmetric $K^{\mathrm{T}} = -K$	$\leftrightarrow$	skew-Hermitian $K^{\mathrm{H}} = -K$
orthogonal $Q^{\mathrm{T}}Q = I$ or $Q^{\mathrm{T}} = Q^{-1}$	$\leftrightarrow$	unitary $U^{\mathrm{H}}U = I$ or $U^{\mathrm{H}} = U^{-1}$
$(Qx)^{\mathrm{T}}(Qy) = x^{\mathrm{T}}y$ and $\lVert Qx\rVert = \lVert x\rVert$	$\leftrightarrow$	$(Ux)^{\mathrm{H}}(Uy) = x^{\mathrm{H}}y$ and $\lVert Ux\rVert = \lVert x\rVert$

The columns, rows, and eigenvectors of Q and U are orthonormal, and every $\lvert\lambda\rvert = 1$

Key words:

Real Hermitian matrices are symmetric; real unitary matrices are orthogonal.

Spectral Theorem:

(1) *Each real symmetric matrix* ***A*** *can be diagonalized by an orthogonal matrix* ***Q***.

(2) *Every Hermitian matrix* ***A*** *can be diagonalized by a unitary matrix* ***U***.

Homework

See Blackboard