Eigenvalues and Eigenvectors

CS6015/LARP

Ack: Linear Algebra and Its Applications , Gilbert Strang

The Solution of $Ax = \lambda x$

- $Ax = \lambda x$ is a nonlinear equation; λ multiplies x. If we could discover λ, then the equation for x would be **linear**.
- We could write λIx in place of λx, and bring this term over to the left side:

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

The vector x is in the nullspace of $A - \lambda I$.
The number λ is chosen so that $A - \lambda I$ has a nullspace.

- We want a **nonzero** eigenvector x. The vector $x = 0$ always satisfies $Ax = \lambda x$, but it is useless.
- To be of any use, the nullspace of $A - \lambda I$ must contain vectors other than zero.
- In short, $\boldsymbol{A - \lambda I}$ **must be singular**.

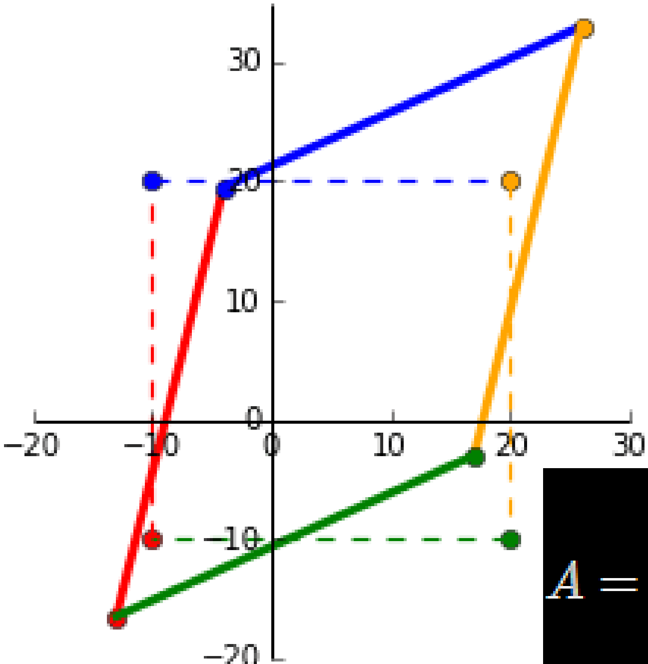
The Solution of $Ax = \lambda x$

5A The number λ is an eigenvalue of A if and only if $A - \lambda I$ is singular:

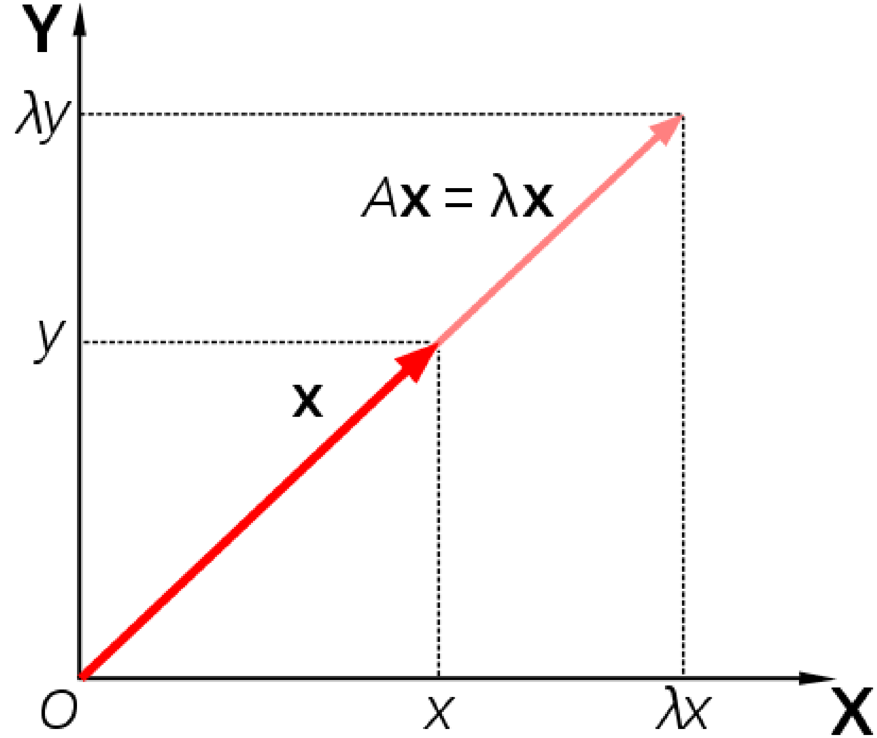
$$\det(A - \lambda I) = 0. \tag{10}$$

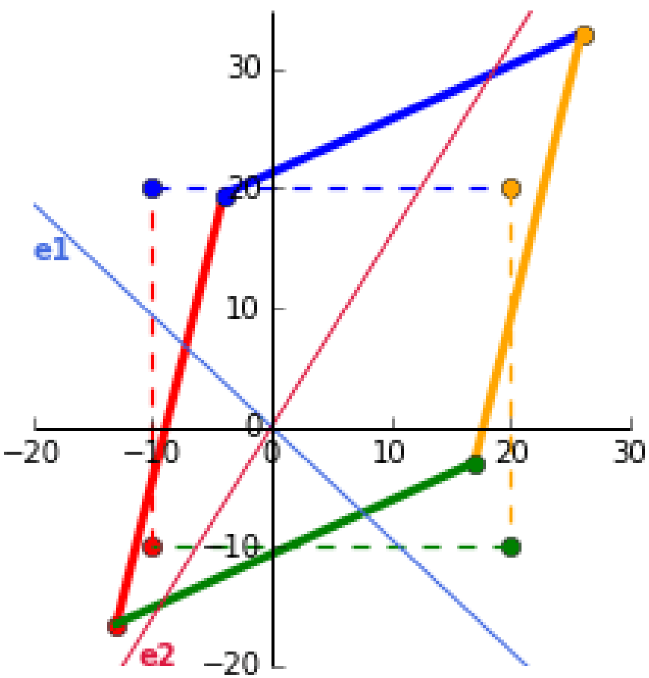
This is the characteristic equation. Each λ is associated with eigenvectors x:

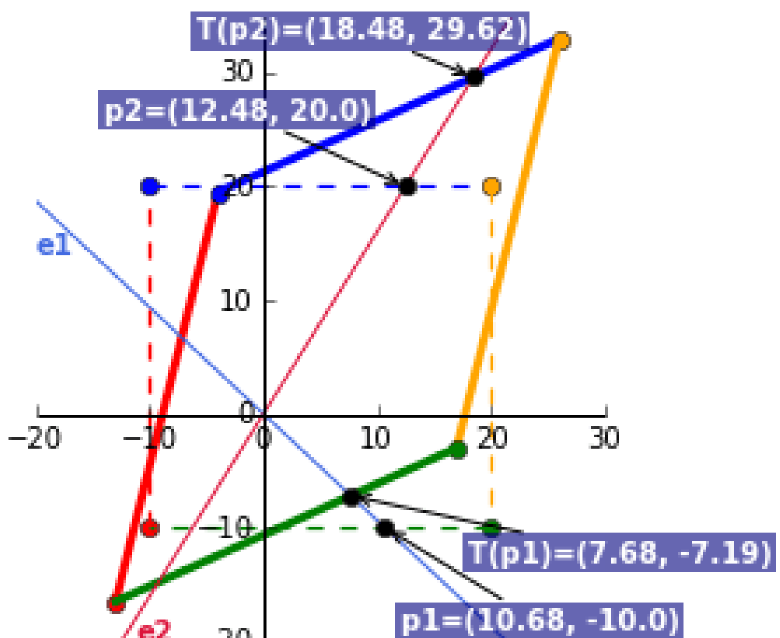
$$(A - \lambda I)x = 0 \qquad \text{or} \qquad Ax = \lambda x. \tag{11}$$



$$A = \begin{bmatrix} 1 & 0.3 \\ 0.45 & 1.2 \end{bmatrix}$$







The Solution of $Ax = \lambda x$

- Example:

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

we shift A by λI to make it singular:

Subtract λI $\quad A - \lambda I = \begin{bmatrix} 4-\lambda & -5 \\ 2 & -3-\lambda \end{bmatrix}$

Determinant $\quad |A - \lambda I| = (4-\lambda)(-3-\lambda) + 10 \quad$ or $\quad \lambda^2 - \lambda - 2$

- This is the **characteristic polynomial.**
- Its **roots**, where the **determinant is zero**, are the **eigenvalues**.

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

The Solution of $Ax = \lambda x$

Eigenvalues $$\lambda = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{1 \pm \sqrt{9}}{2} = -1 \text{ and } 2.$$

- There are two eigen values, because a quadratic has two roots.
- The values $\lambda = -1$ and $\lambda = 2$ lead to a solution of $Ax = \lambda x$ or $(A - \lambda I)x = 0$.

$$\lambda_1 = -1: \qquad (A - \lambda_1 I)x = \begin{bmatrix} 5 & -5 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

The solution (the first eigenvector) is any nonzero multiple of x_1:

Eigenvector for λ_1 $$x_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

The Solution of $Ax = \lambda x$

The solution (the first eigenvector) is any nonzero multiple of x_1:

$$\textbf{Eigenvector for } \lambda_1 \qquad x_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

The computation for λ_2 is done separately:

$$\lambda_2 = 2: \qquad (A - \lambda_2 I)x = \begin{bmatrix} 2 & -5 \\ 2 & -5 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

The second eigenvector is any nonzero multiple of x_2:

$$\textbf{Eigenvector for } \lambda_2 \qquad x_2 = \begin{bmatrix} 5 \\ 2 \end{bmatrix}.$$

For example, suppose

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

racteristic equation is

In both matrices, the columns are multiples of each other, so either column can be used; Eigenvectors ??

Thus, (1, -2) can be taken as an eigenvector associated with the eigenvalue -2; and (3, -1) as an eigenvector associated with the eigenvalue 3, as can be verified by multiplying them by *A*. (read **Cayley–Hamilton** theorem).

The Solution of $\boldsymbol{Ax = \lambda x}$

- The steps in solving $Ax = \lambda x$:

1. **Compute the determinant of $\boldsymbol{A - \lambda I}$.** With λ subtracted along the diagonal, this determinant is a polynomial of degree n. It starts with $(-\lambda)^n$.
2. **Find the roots of this polynomial.** The n roots are the eigenvalues of A.
3. **For each eigenvalue solve the equation $\boldsymbol{(A - \lambda I)x = 0}$.** Since the determinant is zero, there are solutions other than $x = 0$. Those are the eigenvectors.

The Solution of $\boldsymbol{Ax = \lambda x}$ (Recap)

- The key equation was $Ax = \lambda x$.
- Most vectors x will not satisfy such an equation.
- They **change direction** when multiplied by A, so that Ax is not a multiple of x.
- ***This means that only certain special numbers are eigenvalues, and only certain special vectors x are eigenvectors***.

Example 1. Everything is clear when A is a ***diagonal matrix***:

$$A = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \quad \text{has} \quad \lambda_1 = 3 \quad \text{with} \quad x_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \lambda_2 = 2 \quad \text{with} \quad x_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

On each eigenvector A acts like a multiple of the identity: $Ax_1 = 3x_1$ and $Ax_2 = 2x_2$. Other vectors like $x = (1,5)$ are mixtures $x_1 + 5x_2$ of the two eigenvectors, and when A multiplies x_1 and x_2 it produces the eigenvalues $\lambda_1 = 3$ and $\lambda_2 = 2$:

$$A \quad \text{times} \quad x_1 + 5x_2 \quad \text{is} \quad 3x_1 + 10x_2 = \begin{bmatrix} 3 \\ 10 \end{bmatrix}.$$

This is Ax for a typical vector x—not an eigenvector. But the action of A is determined by its eigenvectors and eigenvalues.

Example 2. The eigenvalues of a projection matrix are 1 or 0.

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \quad \text{has} \quad \lambda_1 = 1 \quad \text{with} \quad x_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \lambda_2 = 0 \quad \text{with} \quad x_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

- We have $\lambda = 1$ when x projects to itself, and $\lambda = 0$ when x projects to the zero vector.
- The column space of P is filled with eigenvectors, and so is the nullspace.
- If those spaces have dimension r and $n - r$, then $\lambda = 1$ is repeated r times and $\lambda = 0$ is repeated $n - r$ times (always n λ's):

Four eigenvalues allowing repeats

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{has} \quad \lambda = 1, 1, 0, 0.$$

- A zero eigenvalue signifies that the matrix is singular.

Example 3. The eigenvalues are on the main diagonal when A is triangular.

$$\det(A-\lambda I)=\begin{vmatrix} 1-\lambda & 4 & 5 \\ 0 & \frac{3}{4}-\lambda & 6 \\ 0 & 0 & \frac{1}{2}-\lambda \end{vmatrix}=(1-\lambda)(\tfrac{3}{4}-\lambda)(\tfrac{1}{2}-\lambda)$$

- The determinant is just the product of the diagonal entries.
- It is zero if $\lambda = 1, \lambda = \frac{3}{4}, or\ \lambda = \frac{1}{2}$
- The eigenvalues were already sitting along the main diagonal.

5B The ***sum*** of the n eigenvalues equals the sum of the n diagonal entries:

$$\textbf{Trace of} \quad A = \lambda_1 + \cdots + \lambda_n = a_{11} + \cdots + a_{nn}. \tag{15}$$

Furthermore, the ***product*** of the n eigenvalues equals the ***determinant*** of A.

For a 2 by 2 matrix, the trace and determinant tell us everything:

$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ has trace $a+d$, and determinant $ad-bc$

$$\det(A-\lambda I) = \det\begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} = \lambda^2$$

The eigenvalues are $\lambda =$.

Diagonalization of a Matrix

- The *eigenvectors* ***diagonalize*** *a matrix*

5C Suppose the n by n matrix A has n linearly independent eigenvectors. If these eigenvectors are the columns of a matrix S, then $S^{-1}AS$ is a diagonal matrix Λ. The eigenvalues of A are on the diagonal of Λ:

$$\textbf{Diagonalization} \qquad S^{-1}AS = \Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}. \qquad (1)$$

- We call S the "**eigenvector matrix**" and Λ the "**eigenvalue matrix**".

Diagonalization of a Matrix

5C Suppose the n by n matrix A has n linearly independent eigenvectors. If these eigenvectors are the columns of a matrix S, then $S^{-1}AS$ is a diagonal matrix Λ. The eigenvalues of A are on the diagonal of Λ:

$$\textbf{Diagonalization} \qquad S^{-1}AS = \Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}. \qquad (1)$$

Proof. Put the eigenvectors x_i in the columns of S, and compute AS by columns:

$$AS = A \begin{bmatrix} | & | & & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \lambda_1 x_1 & \lambda_2 x_2 & \cdots & \lambda_n x_n \\ | & | & & | \end{bmatrix}.$$

Then the trick is to split this last matrix into a quite different product $S\Lambda$:

$$\begin{bmatrix} \lambda_1 x_1 & \lambda_2 x_2 & \cdots & \lambda_n x_n \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}.$$

Diagonalization of a Matrix

Proof. Put the eigenvectors x_i in the columns of S, and compute AS by columns:

$$AS = A\begin{bmatrix} | & | & & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \lambda_1 x_1 & \lambda_2 x_2 & \cdots & \lambda_n x_n \\ | & | & & | \end{bmatrix}.$$

Then the trick is to split this last matrix into a quite different product $S\Lambda$:

$$\begin{bmatrix} \lambda_1 x_1 & \lambda_2 x_2 & \cdots & \lambda_n x_n \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}.$$

- It is **crucial to keep these matrices in the right order**.
- If Λ came before S (instead of after), then λ_1 would multiply the entries in the first row. Therefore,

$$AS = S\Lambda, \quad \text{or} \quad S^{-1}AS = \Lambda, \quad \text{or} \quad A = S\Lambda S^{-1}$$

- S is invertible, because its columns (the eigenvectors) were assumed to be independent.

Diagonalization of a Matrix (REMARKS)

Remark 1. If the matrix A has no repeated eigenvalues—the numbers $\lambda_1, \dots, \lambda_n$ are distinct—then its n eigenvectors are automatically independent. So, ***any matrix with distinct eigenvalues can be diagonalized.***

Remark 2. The diagonalizing matrix S is not unique. We can multiply the columns of S by any nonzero constants, and produce a new diagonalizing S.

Remark 3. Other matrices S will not produce a diagonal Λ.

Remark 4. Not all matrices possess n linearly independent eigenvectors, so ***not all matrices are diagonalizable***.

Diagonalization of a Matrix (REMARKS)

The standard example of a "**defective matrix**" is

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

Its eigenvalues are $\lambda_1 = \lambda_2 = 0$, since it is triangular with zeros on the diagonal:

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

All eigenvectors of this A are multiples of the vector $(1,0)$:

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \qquad \text{or} \qquad x = \begin{bmatrix} c \\ 0 \end{bmatrix}$$

$\lambda = 0$ is a *double eigenvalue*—its *algebraic multiplicity* is 2. But the *geometric multiplicity* is 1—there is only one independent eigenvector. **We can't construct S.**

Diagonalization of a Matrix (REMARKS)

Diagonalizability of A depends on enough eigenvectors.
Invertibility of A depends on nonzero eigenvalues.

- *Diagonalization can fail only if there are repeated eigenvalues.*
- Even then, it does not always fail.
- $A = I$ has repeated eigenvalues $1, 1, \ldots, 1$ but it is already diagonal!

There is no shortage of eigenvectors in that case.

Diagonalization of a Matrix (REMARKS)

5D If eigenvectors $x_1, \ldots, x_k$ correspond to *different eigenvalues* $\lambda_1, \ldots, \lambda_k$, then those eigenvectors are linearly independent.

- Eigenvectors that come from distinct eigenvalues are automatically independent.
- A matrix with n distinct eigenvalues can be diagonalized. This is the typical case.

Examples of Diagonalization

Example 1. The projection $A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$ has eigenvalue matrix $\Lambda = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. The eigenvectors go into the columns of S:

$$S = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad \text{and} \quad AS = S\Lambda = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}.$$

That last equation can be verified at a glance. Therefore $S^{-1}AS = \Lambda$.

= ??

Example 2. The eigenvalues themselves are not so clear for a ***rotation***:

$$\textbf{90° rotation} \qquad K = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \text{has} \quad \det(K - \lambda I) = \lambda^2 + 1.$$

How can a vector be rotated and still have its direction unchanged?

- It can't—except for the zero vector, which is useless.
- The eigenvalues of K are imaginary numbers, $\lambda_1 = i$ and $\lambda_2 = -i$.
- In turning through 90°, they are multiplied by i or $-i$:

$$(K - \lambda_1 I)x_1 = \begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{and} \quad x_1 = \begin{bmatrix} 1 \\ -i \end{bmatrix}$$

$$(K - \lambda_2 I)x_2 = \begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{and} \quad x_2 = \begin{bmatrix} 1 \\ i \end{bmatrix}.$$

$$(K - \lambda_1 I)x_1 = \begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{and} \quad x_1 = \begin{bmatrix} 1 \\ -i \end{bmatrix}$$

$$(K - \lambda_2 I)x_2 = \begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{and} \quad x_2 = \begin{bmatrix} 1 \\ i \end{bmatrix}.$$

- The eigenvalues are distinct, even if imaginary, and the eigenvectors are independent. They go into the columns of S:

$$S = \begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix} \quad \text{and} \quad S^{-1}KS = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}.$$

- ***Complex numbers are needed even for real matrices.***
- If there are too few real eigenvalues, there are always n complex eigenvalues. (Complex includes real, when the imaginary part is zero.)

Powers and Products: A^k and AB

- *The eigenvalue of A^2 are exactly λ_1^2 , … , λ_n^2 , and every eigenvector of A is also an eigenvector of A^2*

$$A^2x = A\lambda x = \lambda Ax = \lambda^2 x.$$

- Thus λ^2 is an eigenvalue of A^2 , with the same eigenvector x.

- The same result comes from diagonalization, by squaring $S^{-1}AS = \Lambda$:

Eigenvalues of A^2 $\qquad (S^{-1}AS)(S^{-1}AS) = \Lambda^2 \quad \text{or} \quad S^{-1}A^2S = \Lambda^2.$

- The matrix A^2 is diagonalized by the same S, so the eigenvectors are unchanged. The eigenvalues are squared.
- This continues to hold for any power of A.

Powers and Products: A^k and AB

5E The eigenvalues of A^k are $\lambda_1^k, \ldots, \lambda_n^k$, and each eigenvector of A is still an eigenvector of A^k. When S diagonalizes A, it also diagonalizes A^k:

$$\Lambda^k = (S^{-1}AS)(S^{-1}AS)\cdots(S^{-1}AS) = S^{-1}A^kS. \tag{4}$$

Each S^{-1} cancels an S, except for the first S^{-1} and the last S.

- If A is invertible this rule also applies to its inverse (the power $k = -1$).
- The eigenvalues of A^{-1} are $\frac{1}{\lambda_i}$.

$$\textit{if} \quad Ax = \lambda x \quad \textit{then} \quad x = \lambda A^{-1}x \quad \textit{and} \quad \frac{1}{\lambda}x = A^{-1}x.$$

Powers and Products: A^k and AB

Example 3. If K is rotation through 90°, then K^2 is rotation through 180° (which means $-I$) and K^{-1} is rotation through $-90°$:

$$K = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \qquad K^2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \text{and} \quad K^{-1} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

The eigenvalues of K are i and $-i$; their squares are -1 and -1; their reciprocals are $1/i = -i$ and $1/(-i) = i$. Then K^4 is a complete rotation through 360°:

$$K^4 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and also} \quad \Lambda^4 = \begin{bmatrix} i^4 & 0 \\ 0 & (-i)^4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Powers and Products: A^k and AB

5F Diagonalizable matrices share the same eigenvector matrix S if and only if $AB = BA$.

Proof. If the same S diagonalizes both $A = S\Lambda_1 S^{-1}$ and $B = S\Lambda_2 S^{-1}$, we can multiply in either order:

$$AB = S\Lambda_1 S^{-1} S\Lambda_2 S^{-1} = S\Lambda_1\Lambda_2 S^{-1} \quad \text{and} \quad BA = S\Lambda_2 S^{-1} S\Lambda_1 S^{-1} = S\Lambda_2\Lambda_1 S^{-1}.$$

Since $\Lambda_1\Lambda_2 = \Lambda_2\Lambda_1$ (diagonal matrices always commute) we have $AB = BA$.

In the opposite direction, suppose $AB = BA$. Starting from $Ax = \lambda x$, we have

$$ABx = BAx = B\lambda x = \lambda Bx.$$

Thus x and Bx are both eigenvectors of A, sharing the same λ (or else $Bx = 0$).

Complex Matrices

- We now introduce the space $\mathbf{C}^n$ of vectors with n complex components.
- Addition and matrix multiplication follow the same rules as before.
- ***Length is computed differently***
- The old way, the vector in $\mathbf{C}^2$ with components $(1, i)$ would have zero length: $1^2 + i^2 = 0$ ***which is not good.***
- The correct length squared is $1^2 + |i|^2 = 2$
- The inner product, the transpose, the definitions of symmetric and orthogonal matrices, all need to be modified for complex numbers.

Complex Matrices

We particularly want to find out about *symmetric matrices* and *Hermitian matrices*: ***Where are their eigenvalues, and what is special about their eigenvectors***?

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

Complex Numbers and Their Conjugates

The real numbers a and the imaginary numbers ib are special cases of complex numbers; they lie on the axes

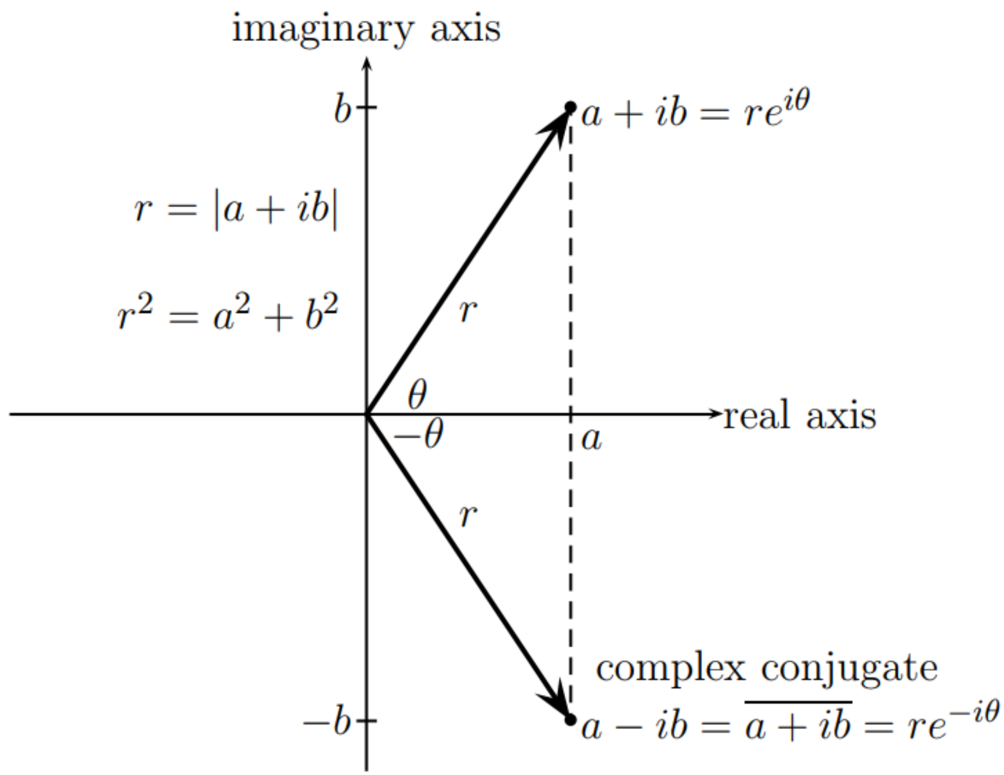


Fig: The complex plane, with $a + ib = re^{i\theta}$ and its conjugate $a - ib = re^{-i\theta}$

Complex Numbers and Their Conjugates

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

Multiplication $$\begin{aligned}(a+ib)(c+id) &= ac+ibc+iad+i^2bd \\ &= (ac-bd)+i(bc+ad)\end{aligned}$$

- The **complex conjugate** of $a + ib$ is the number $a - ib$. The ***sign*** of the ***imaginary part is reversed.***
- **It is the mirror image across the real axis**
- Any real number is its own conjugate, since $b = 0$.

Complex Numbers and Their Conjugates

The conjugate is denoted by a bar or a star: $(a+ib)^* = \overline{(a+ib)} = a - ib$.

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 $a^2 + b^2$:

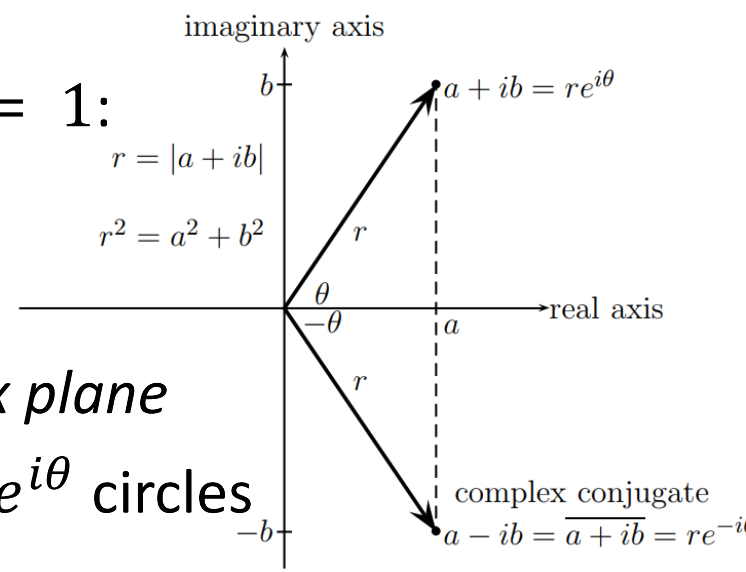
Absolute value $\quad (a+ib)(a-ib) = a^2+b^2 = r^2$

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

Complex Numbers and Their Conjugates

- Trigonometry connects the sides a and b to the hypotenuse r by $a = r\cos\theta$ and $b = r\,sin\theta$.
- Combining these two equations moves us into polar coordinates:

Polar form $\quad a + ib = r(\cos\theta + i\sin\theta) = re^{i\theta}$



Most important special case is when $r = 1$:

$$a + ib \text{ is } e^{i\theta} = \cos\theta + i\sin\theta$$

- It falls on the ***unit circle*** in the *complex plane*
- As θ varies from 0 to 2π, this number $e^{i\theta}$ circles around zero at the constant radial distance:

$$|e^{i\theta}| = \sqrt{\cos^2\theta + \sin^2\theta} = 1$$

Complex Numbers and Their Conjugates

Example 1. $x = 3 + 4i$ times its conjugate $\bar{x} = 3 - 4i$ is the absolute value squared:

$$x\bar{x} = (3+4i)(3-4i) = 25 = |x|^2 \quad \text{so} \quad r = |x| = 5.$$

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, 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 $\theta - \alpha$.

Lengths and Transposes in the Complex Case

The complex vector space $\mathbf{C}^n$ contains all vectors x with n complex components:

$$\textbf{Complex vector} \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{with components} \quad x_j = a_j + ib_j.$$

In the new definition of length, each x_j^2 is replaced by its modulus $|x_j|^2$:

$$\textbf{Length squared} \qquad \|x\|^2 = |x_1|^2 + \cdots + |x_n|^2$$

Hermitian Matrices

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

The matrix $\bar{A}^T = A^H = A^*$ is called a "***Hermitian***":

$$A^{\mathrm{H}} = \overline{A}^{\mathrm{T}} \quad \text{has entries} \quad (A^{\mathrm{H}})_{ij} = \overline{A_{ji}}.$$

Conjugate transpose

$$\begin{bmatrix} 2+i & 3i \\ 4-i & 5 \\ 0 & 0 \end{bmatrix}^{\mathrm{H}} = \begin{bmatrix} 2-i & 4+i & 0 \\ -3i & 5 & 0 \end{bmatrix}$$

- This symbol A^H gives official recognition to the fact that, with complex entries, it is seldom that we want only the transpose of A.
- It is the *conjugate transpose* A^H that becomes appropriate.
- A real symmetric matrix is certainly Hermitian. The eigenvalues are real

Hermitian Matrices

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

Every entry of A contributes to $x^{\mathrm{H}}Ax$. Try the 2 by 2 case with $x = (u, v)$:

$$x^{\mathrm{H}}Ax = \begin{bmatrix} \bar{u} & \bar{v} \end{bmatrix} \begin{bmatrix} 2 & 3-3i \\ 3+3i & 5 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

$$= 2\bar{u}u + 5\bar{v}v + (3-3i)\bar{u}v + (3+3i)u\bar{v}$$

$$= \textbf{real} + \textbf{real} + (\textbf{sum of complex conjugates}).$$

Property 2 If $A = A^{\mathrm{H}}$, every eigenvalue is real.

Proof. Suppose $Ax = \lambda x$. ***The trick is to multiply by*** x^{H}: $x^{\mathrm{H}}Ax = \lambda x^{\mathrm{H}}x$. The left-hand side is real by Property 1, and the right-hand side $x^{\mathrm{H}}x = \|x\|^2$ is real and positive, because $x \neq 0$. Therefore $\lambda = x^{\mathrm{H}}Ax / x^{\mathrm{H}}x$ must be real. Our example has $\lambda = 8$ and $\lambda = -1$:

$$|A - \lambda I| = \begin{vmatrix} 2-\lambda & 3-3i \\ 3+3i & 5-\lambda \end{vmatrix} = \lambda^2 - 7\lambda + 10 - |3-3i|^2$$)

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

Hermitian Matrices

Property 3 Two eigenvectors of a real symmetric matrix or a Hermitian matrix, if they come from different eigenvalues, are orthogonal to one another.

The proof starts with $Ax = \lambda_1 x$, $Ay = \lambda_2 y$, and $A = A^{\mathrm{H}}$:

$$(\lambda_1 x)^{\mathrm{H}} y = (Ax)^{\mathrm{H}} y = x^{\mathrm{H}} A y = x^{\mathrm{H}} (\lambda_2 y).$$

- The outside numbers are $\lambda_1 x^H y = \lambda_2 x^H y$, since the λ's are real.
- Now use the assumption $\lambda_1 \neq \lambda_2$, which forces the conclusion that $x^H y = 0$.
- In our example,

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

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

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

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

- These two eigenvectors are orthogonal:

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

5O A real symmetric matrix can be factored into $A = Q\Lambda Q^{\mathrm{T}}$. Its orthonormal eigenvectors are in the orthogonal matrix Q and its eigenvalues are in Λ.

- In geometry or mechanics, this is the ***principal axis theorem***. It gives the right choice of ***axes for an ellipse***.
- Those axes are perpendicular, and they ***point along the eigenvectors of the corresponding matrix***.
- In mathematics the formula $A = Q\Lambda Q^T$ is known as the ***spectral theorem***.

Hermitian Matrices

In mathematics the formula $A = Q\Lambda Q^T$ is known as the ***spectral theorem***.

$$A = Q\Lambda Q^{\mathrm{T}} = \begin{bmatrix} | & & | \\ x_1 & \cdots & x_n \\ | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} - & x_1^{\mathrm{T}} & - \\ & \vdots & \\ - & x_n^{\mathrm{T}} & - \end{bmatrix}$$

$$= \lambda_1 x_1 x_1^{\mathrm{T}} + \lambda_2 x_2 x_2^{\mathrm{T}} + \cdots + \lambda_n x_n x_n^{\mathrm{T}}.$$

- The spectral theorem $A = Q\Lambda Q^T$ has been proved only when the eigenvalues of A are distinct. Then there are certainly n independent eigenvectors, and A can be safely diagonalized.
- Nevertheless it is true that even with repeated eigenvalues, a symmetric matrix still has a complete set of orthonormal eigenvectors.
- The extreme case is the identity matrix, which has $\lambda = 1$ repeated n times—and no shortage of eigenvectors.

Unitary Matrices

A complex matrix with **orthonormal columns** is called a ***unitary matrix.***

Two analogies:

1. *A Hermitian (or symmetric) matrix can be compared to a real number.*
2. *A unitary (or orthogonal) matrix can be compared to a number on the unit circle.*

Unitary matrix $\quad U^{\mathrm{H}}U = I, \quad UU^{\mathrm{H}} = I, \quad \text{and} \quad U^{\mathrm{H}} = U^{-1}$

Unitary Matrices

Property 1′ $(Ux)^{\mathrm{H}}(Uy) = x^{\mathrm{H}}U^{\mathrm{H}}Uy = x^{\mathrm{H}}y$ and lengths are preserved by U:

Length unchanged $$\|Ux\|^2 = x^{\mathrm{H}}U^{\mathrm{H}}Ux = \|x\|^2. \qquad (11)$$

Property 2′ Every eigenvalue of U has absolute value $|\lambda| = 1$.

This follows directly from $Ux = \lambda x$, by comparing the lengths of the two sides: $\|Ux\| = \|x\|$ by Property 1′, and always $\|\lambda x\| = |\lambda|\|x\|$. Therefore $|\lambda| = 1$.

Property 3′ Eigenvectors corresponding to different eigenvalues are orthonormal.

- **Example:**

$$U = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix} \text{ has eigenvalues } e^{it} \text{ and } e^{-it}.$$

The orthogonal eigenvectors are $x = (1, -i)$ and $y = (1, i)$. (Remember to take conjugates in $x^{\mathrm{H}}y = 1 + i^2 = 0$.) After division by $\sqrt{2}$ they are orthonormal.

- **Skew-symmetric matrix:** $\boldsymbol{K^T = -K}$
- **Skew-Hermitian matrix:** $\boldsymbol{K^H = -K}$

If A is Hermitian then $K = iA$ is skew-Hermitian.

- The eigenvalues of K are purely imaginary instead of purely real; we multiply i. The eigenvectors are not changed.

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

Example:

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

Real versus Complex

$\mathbf{R}^n$ (n real components)	$\leftrightarrow$	$\mathbf{C}^n$ (n complex components)
length: $\|x\|^2 = x_1^2 + \cdots + x_n^2$	$\leftrightarrow$	length: $\|x\|^2 = \|x_1\|^2 + \cdots + \|x_n\|^2$
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 $\|Qx\| = \|x\|$	$\leftrightarrow$	$(Ux)^{\mathrm{H}}(Uy) = x^{\mathrm{H}}y$ and $\|Ux\| = \|x\|$

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

Virtually every step in this chapter has involved the combination $S^{-1}AS$. The eigenvectors of A went into the columns of S, and that made $S^{-1}AS$ a diagonal matrix (called Λ). When A was symmetric, we wrote Q instead of S, choosing the eigenvectors to be orthonormal. In the complex case, when A is Hermitian we write U—it is still the matrix of eigenvectors. Now we look at all combinations $M^{-1}AM$—*formed with any invertible M on the right and its inverse on the left*. The invertible eigenvector matrix S may fail to exist (the defective case), or we may not know it, or we may not want to use it.

Similarity Transformations

- The matrices A and $M^{-1}AM$ are "**similar**".
- Going from one to the other is a ***similarity transformation***.
- A whole family of matrices $M^{-1}AM$ is similar to A, and there are two questions:

1. What do these similar matrices $M^{-1}AM$ have in common?

2. With a special choice of M, what special form can be achieved by $M^{-1}AM$?

Similarity Transformations

- The matrices A and $M^{-1}AM$ are "**similar**".
- The family of matrices $M^{-1}AM$ includes A itself, by choosing $M = I$.
- ***Similar matrices share the same eigenvalues.***

5P Suppose that $B = M^{-1}AM$. Then A and B have the **same eigenvalues. Every eigenvector** x **of** A **corresponds to an eigenvector** $M^{-1}x$ **of** B.

Start from $Ax = \lambda x$ and substitute $A = MBM^{-1}$:

$$\textbf{Same eigenvaluc} \qquad MBM^{-1}x = \lambda x \quad \text{which is} \quad B(M^{-1}x) = \lambda(M^{-1}x). \qquad (1)$$

The eigenvalue of B is still λ. The eigenvector has changed from x to $M^{-1}x$.

Similarity Transformations

We can also check that $A - \lambda I$ and $B - \lambda I$ have the same determinant:

Product of matrices $\quad B - \lambda I = M^{-1}AM - \lambda I = M^{-1}(A - \lambda I)M$

Product rule $\quad \det(B - \lambda I) = \det M^{-1} \det(A - \lambda I) \det M = \det(A - \lambda I).$

- The polynomials $\det(A - \lambda I)$ and $\det(B - \lambda I)$ are equal.
- Their roots—the eigenvalues of A and B—are the same. ***Here are matrices B similar to A.***

Example 1. $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ has eigenvalues 1 and 0. Each B is $M^{-1}AM$:

$$\text{If } M = \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix}, \text{ then } B = \begin{bmatrix} 1 & b \\ 0 & 0 \end{bmatrix}: \quad \text{triangular with } \lambda = 0 \text{ and } 0.$$

$$\text{If } M = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \text{ then } B = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}: \quad \text{projection with } \lambda = 0 \text{ and } 0.$$

$$\text{If } M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \text{ then } B = \text{an arbitrary matrix with } \lambda = 0 \text{ and } 0.$$

Diagonalizing Symmetric and Hermitian Matrices

- The triangular form will show that any *symmetric* or *Hermitian* matrix—whether its *eigenvalues are distinct or not*—has a **complete set of orthonormal eigenvectors**.
- We need a unitary matrix such that $U^{-1}AU$ is diagonal.
- This triangular T must be diagonal, because it is also Hermitian when $A = A^H$:
- The diagonal matrix $U^{-1}AU$ represents a key theorem in linear algebra.

$$T = T^{\mathrm{H}} \qquad (U^{-1}AU)^{\mathrm{H}} = U^{\mathrm{H}}A^{\mathrm{H}}(U^{-1})^{\mathrm{H}} = U^{-1}AU$$

Diagonalizing Symmetric and Hermitian Matrices

The diagonal matrix $U^{-1}AU$ represents a key theorem in linear algebra.

$$T = T^{\mathrm{H}} \qquad (U^{-1}AU)^{\mathrm{H}} = U^{\mathrm{H}}A^{\mathrm{H}}(U^{-1})^{\mathrm{H}} = U^{-1}AU$$

5S **(Spectral Theorem)** Every real symmetric A can be diagonalized by an orthogonal matrix Q. Every Hermitian matrix can be diagonalized by a unitary U:

$$\textbf{(real)} \qquad Q^{-1}AQ = \Lambda \quad \text{or} \quad A = Q\Lambda Q^{\mathrm{T}}$$

$$\textbf{(complex)} \qquad U^{-1}AU = \Lambda \quad \text{or} \quad A = U\Lambda U^{\mathrm{H}}$$

The columns of Q (or U) contain orthonormal eigenvectors of A.

Diagonalizing Symmetric and Hermitian Matrices

***Remark* 1.** In the real symmetric case, the eigenvalues and eigenvectors are real at every step. That produces a *real* unitary U—an orthogonal matrix.

***Remark* 2.** A is the limit of symmetric matrices with *distinct* eigenvalues. As the limit approaches, the eigenvectors stay perpendicular. This can fail if $A \neq A^{\mathrm{T}}$:

$$A(\theta) = \begin{bmatrix} 0 & \cos\theta \\ 0 & \sin\theta \end{bmatrix} \quad \text{has eigenvectors} \quad \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}.$$

As $\theta \to 0$, the *only* eigenvector of the nondiagonalizable matrix $\left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$ is $\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$.

Normal Matrices

- The matrix N is normal if it commutes with N^H: $NN^H = N^H N$.
- **Normal matrices are exactly those that have a complete set of orthonormal eigenvectors.**

Read about Jordan form

Remark 2. A is the limit of symmetric matrices with *distinct* eigenvalues. As the limit approaches, the eigenvectors stay perpendicular. This can fail if $A \neq A^{\mathrm{T}}$:

$$A(\theta) = \begin{bmatrix} 0 & \cos\theta \\ 0 & \sin\theta \end{bmatrix} \quad \text{has eigenvectors} \quad \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}.$$

As $\theta \to 0$, the *only* eigenvector of the nondiagonalizable matrix $\left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$ is $\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$.

Example 3. The spectral theorem says that this $A = A^{\mathrm{T}}$ can be diagonalized:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{with repeated eigenvalues} \quad \lambda_1 = \lambda_2 = 1 \text{ and } \lambda_3 = -1.$$

$\lambda = 1$ has a plane of eigenvectors, and we pick an orthonormal pair x_1 and x_2:

$$x_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad x_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \text{and} \quad x_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \quad \text{for } \lambda_3 = -1.$$

These are the columns of Q. Splitting $A = Q\Lambda Q^{\mathrm{T}}$ into 3 columns times 3 rows gives

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \lambda_1 \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \lambda_3 \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Since $\lambda_1 = \lambda_2$, those first two projections $x_1 x_1^{\mathrm{T}}$ and $x_2 x_2^{\mathrm{T}}$ (each of rank 1) combine to give a projection P_1 of rank 2 (onto the plane of eigenvectors). Then A is

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \lambda_1 P_1 + \lambda_3 P_3 = (+1) \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} + (-1) \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (5)$$