Complex eigenvectors of a general 2×2 matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad A - \lambda I = \begin{bmatrix} a-\lambda & b \\ c & d-\lambda \end{bmatrix}$$

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

$$(a - \lambda)(d - \lambda) - bc = 0$$

$$\downarrow$$

$$\lambda^2 - (a + d)\lambda + (ad - bc) = 0$$

$$\downarrow$$

$$\lambda = \frac{(a + d) \pm \sqrt{(a + d)^2 - 4(ad - bc)}}{2}$$

If $\left((a + d)^2 - 4(ad - bc)\right) < 0$,

$\lambda_1 = \alpha + i\beta$ and $\lambda_2 = \alpha - i\beta$

$$\downarrow$$

$$\lambda_1 = (\lambda_2)^c$$

(we use ^{c} for complex conjugate)

- $\alpha = \dfrac{a + d}{2}$

- $\beta = \dfrac{\sqrt{4(ad - bc) - (a + d)^2}}{2}$

for $\lambda_1 = \alpha + i\beta$, we have $\left[\begin{array}{cc|c} a - \lambda_1 & b & 0 \\ c & d - \lambda_1 & 0 \end{array}\right]$

λ_1 was chosen from $\det(A - \lambda I) = 0$

$\downarrow$

the system has at least one free variable

$$\left[\begin{array}{cc|c} a-\lambda_1 & b & 0 \\ c & d-\lambda_1 & 0 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} a-\lambda_1 & b & 0 \\ 0 & (a-\lambda_1)(d-\lambda_1)-bc & 0 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} a-\lambda_1 & b & 0 \\ 0 & 0 & 0 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & \dfrac{b}{a-\lambda_1} & 0 \\ 0 & 0 & 0 \end{array}\right]$$

① choosing the free variable $x_2 = 1$ gives $x_1 = \dfrac{-b}{a-\lambda_1}$

$\downarrow$

$$\vec{x}_1 \text{ (for } \lambda_1) = \begin{bmatrix} \dfrac{-b}{a-\lambda_1} \\ 1 \end{bmatrix} = \begin{bmatrix} \dfrac{-b}{(a-\alpha)-i\beta} \\ 1 \end{bmatrix}$$

$A\vec{x}_1 = \lambda_1 \vec{x}_1$

take complex conjugates of both sides:

$(A\vec{x}_1)^c = (\lambda_1 \vec{x}_1)^c$

$\downarrow$

$A^c \vec{x}_1{}^c = \lambda_1{}^c \vec{x}_1{}^c$

$\downarrow$

$A \vec{x}_1{}^c = \lambda_1{}^c \vec{x}_1{}^c$

$\downarrow$

$A \vec{x}_1{}^c = \lambda_2 \vec{x}_1{}^c$

$\downarrow$

$\vec{x}_2 = \vec{x}_1{}^c$ is one of the eigenvectors associated with λ_2

$\vec{x}_1$ and $\vec{x}_2$ are vectors in $\mathbb{C}^2$ and can be uniquely decomposed as

- $\vec{x}_1 = \mathrm{Re}(\vec{x}_1) + i\,\mathrm{Im}(\vec{x}_1)$
- $\vec{x}_2 = \mathrm{Re}(\vec{x}_2) + i\,\mathrm{Im}(\vec{x}_2)$

The three images show $\mathrm{Re}(\vec{x}_1)$ and $\mathrm{Im}(\vec{x}_1)$

for $A = \begin{bmatrix} 0.5 & -1.2 \\ 1 & 1.5 \end{bmatrix}$ $\lambda \approx 1 \pm 0.975\,i$

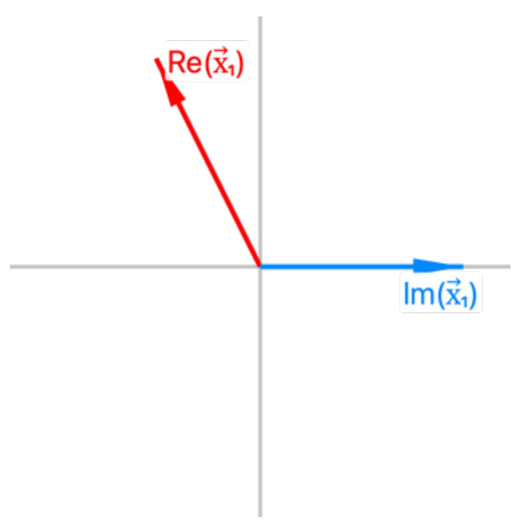
with different free variable choices:

Ⓐ pure real
Ⓑ pure imaginary
Ⓒ complex

Ⓐ

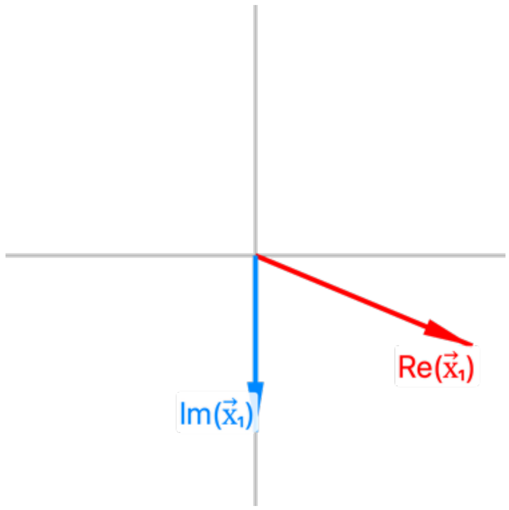
free variable $x_2 = 1$

$$\vec{x} = \begin{bmatrix} -0.337 \\ 0.674 \end{bmatrix} \pm i \begin{bmatrix} 0.657 \\ 0 \end{bmatrix}$$



free variable $x_1 = 1$

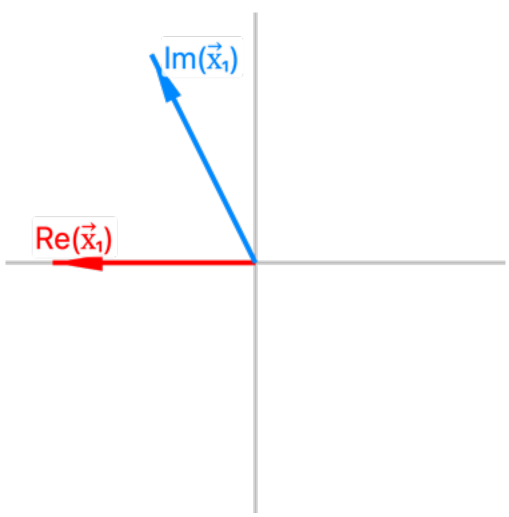
$$\vec{x} = \begin{bmatrix} 0.739 \\ -0.308 \end{bmatrix} \pm i \begin{bmatrix} 0 \\ -0.6 \end{bmatrix}$$



Ⓑ

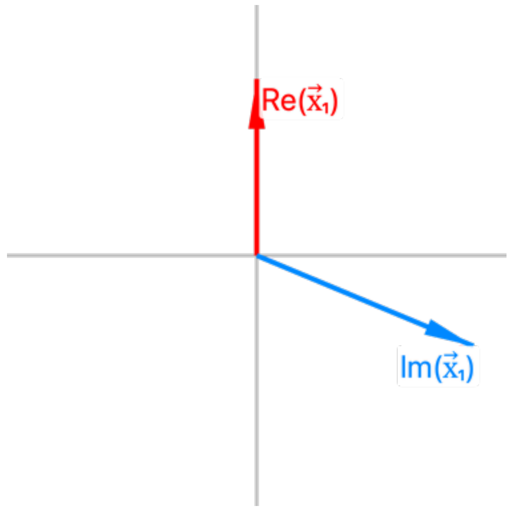
free variable $x_2 = i$

$$\vec{x} = \begin{bmatrix} -0.657 \\ 0 \end{bmatrix} \pm i \begin{bmatrix} -0.337 \\ 0.674 \end{bmatrix}$$



free variable $x_1 = i$

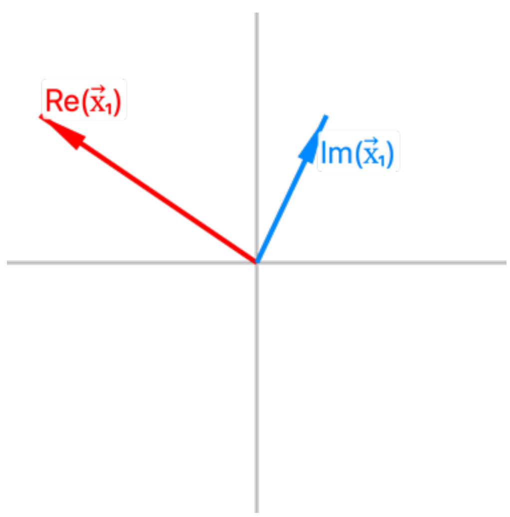
$$\vec{x} = \begin{bmatrix} 0 \\ 0.6 \end{bmatrix} \pm i \begin{bmatrix} 0.739 \\ -0.308 \end{bmatrix}$$



Ⓒ

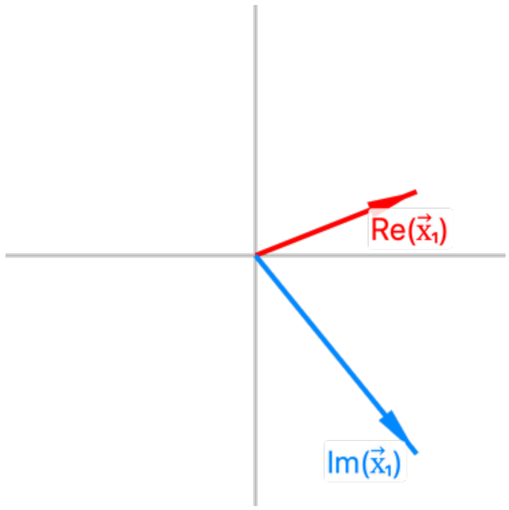
free variable $x_2 = 1 + i$

$$\vec{x} = \begin{bmatrix} -0.703 \\ 0.477 \end{bmatrix} \pm i \begin{bmatrix} 0.226 \\ 0.477 \end{bmatrix}$$



free variable $x_1 = 1 + i$

$$\vec{x} = \begin{bmatrix} 0.522 \\ 0.207 \end{bmatrix} \pm i \begin{bmatrix} 0.522 \\ -0.642 \end{bmatrix}$$



Choices in computing $\vec{x}_1$:

① Choose which coordinate is the free variable (usually x_2)

② Choose the value of the free variable (usually 1)

③ Choose the scaling convention (usually normalized)

Eigenvectors of rotation-scaling matrix

$$A = s\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \qquad A - \lambda I = \begin{bmatrix} a - \lambda & -b \\ b & a - \lambda \end{bmatrix}$$

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

$$(a - \lambda)^2 + b^2 = 0$$

$$\downarrow$$

$$(a - \lambda) = \pm ib$$

$$\downarrow$$

$$\lambda = a \pm ib$$

Find $A\,\vec{x}_1 = \lambda_1\,\vec{x}_1$ corresponding to $\lambda_1 = a + ib$:

$$\begin{bmatrix} a - (a + ib) & -b \\ b & a - (a + ib) \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

λ_1 and λ_2 were computed so that the system has at least one free variable

$$\left[\begin{array}{cc|c} -ib & -b & 0 \\ b & -ib & 0 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} -ib & -b & 0 \\ 0 & 0 & 0 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & -i & 0 \\ 0 & 0 & 0 \end{array}\right]$$

- choose free variable $x_2 = 1$

$$\downarrow$$

$$x_1 = i\,x_2$$

$$\vec{x}_1 = \begin{bmatrix} i\,x_2 \\ x_2 \end{bmatrix} \rightarrow \vec{x}_1 = \begin{bmatrix} i \\ 1 \end{bmatrix}$$

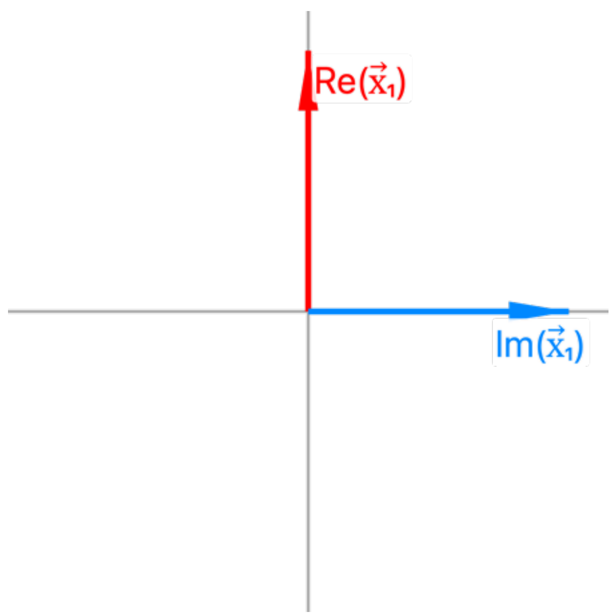
• normalize $\vec{x}_1$ using the complex length:

$$|\,\vec{x}_1\,|^2 = |\,i\,|^2 + |\,1\,|^2 = 2$$

$$\downarrow$$

$$|\,\vec{x}_1\,| = \sqrt{2}$$

$$\vec{x}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} i \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} + i \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$$



Ⓐ

A different real free-variable choice results in the same $\vec{x}$ after normalization

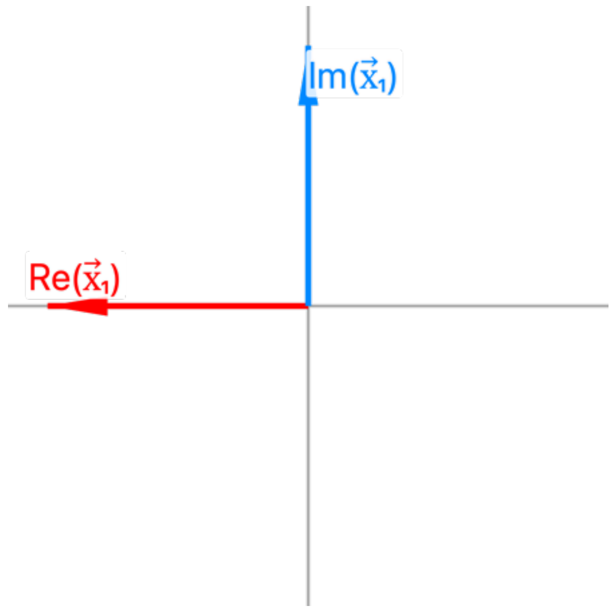
Ⓑ

A pure imaginary free-variable value produces the following $\vec{x}_1$:

For $\lambda_1 = a + ib$:

$x_1 = i\,x_2 = i^2 = -1$

$$\vec{x}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ i \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} + i \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$



Ⓒ

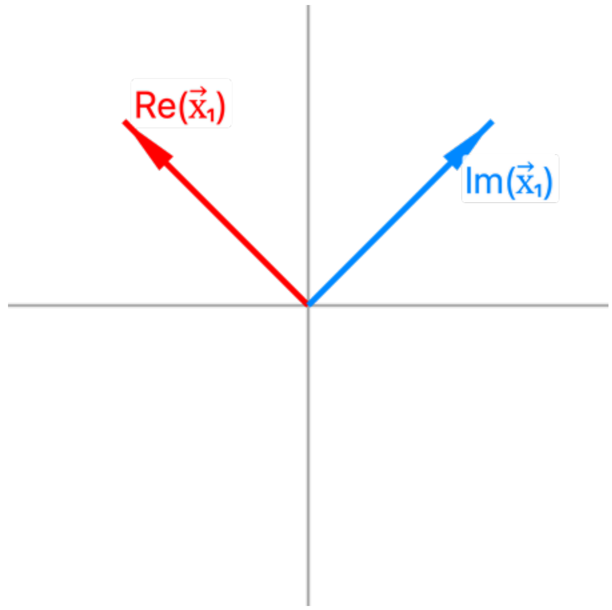
A general complex free-variable value gives another rotation of the pair

For example, $x_2 = 1 + i$:

For $\lambda_1 = a + i\,b$:

$x_1 = i\,x_2 = i\,(1 + i) = -1 + i$

$$\vec{x}_1 = \frac{1}{2} \begin{bmatrix} -1 + i \\ 1 + i \end{bmatrix} = \begin{bmatrix} \frac{-1 + i}{2} \\ \frac{1 + i}{2} \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix} + i \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$$



Eigenvectors $\vec{x}$ of a rotation-scaling matrix:
① $\mathrm{Re}(\vec{x})$ and $\mathrm{Im}(\vec{x})$ are orthogonal
② $|\,\mathrm{Re}(\vec{x})\,| = |\,\mathrm{Im}(\vec{x})\,|$
③ Changing the free variable rotates the pair and preserves the lengths and the angle

Why this stays true:

For $\lambda_1 = a + i\,b$ of a rotation-scaling matrix, row reduction gave

$$x_1 = i\,x_2$$

$$\downarrow$$

$$\vec{x} = x_2 \begin{bmatrix} i \\ 1 \end{bmatrix}$$

Suppose we choose $x_2 = p + i\,q$

$$\downarrow$$

$$\vec{x}' = x_2 \begin{bmatrix} i \\ 1 \end{bmatrix} = (p + i\,q)\begin{bmatrix} i \\ 1 \end{bmatrix} = p\begin{bmatrix} i \\ 1 \end{bmatrix} + i\,q\begin{bmatrix} i \\ 1 \end{bmatrix} = \begin{bmatrix} i\,p \\ p \end{bmatrix} + \begin{bmatrix} -q \\ i\,q \end{bmatrix} = \begin{bmatrix} -q \\ p \end{bmatrix} + i\begin{bmatrix} p \\ q \end{bmatrix}$$

$$\mathrm{Re}(\vec{x}) = \begin{bmatrix} -q \\ p \end{bmatrix} \qquad \mathrm{Im}(\vec{x}) = \begin{bmatrix} p \\ q \end{bmatrix}$$

$\downarrow$

for rotation-scaling matrix,

- $|\,\mathrm{Re}(\vec{x})\,| = |\,\mathrm{Im}(\vec{x})\,|$
- $\mathrm{Re}(\vec{x}) \cdot \mathrm{Im}(\vec{x}) = 0$

$\downarrow$

Changing the complex free variable gives another orthogonal equal-length pair $\mathrm{Re}(\vec{x})$, $\mathrm{Im}(\vec{x})$

From defective shear-scaling to rotation-scaling

Here, all computations use free variable $x_1 = 1$
(for the defective matrix, x_1 is the only possible free variable)

Ⓐ

Start with defective shear-scaling matrix $A = \begin{bmatrix} a & b \\ 0 & a \end{bmatrix}$

$$\det(A - \lambda I) = (a - \lambda)^2 = 0$$

$\downarrow$

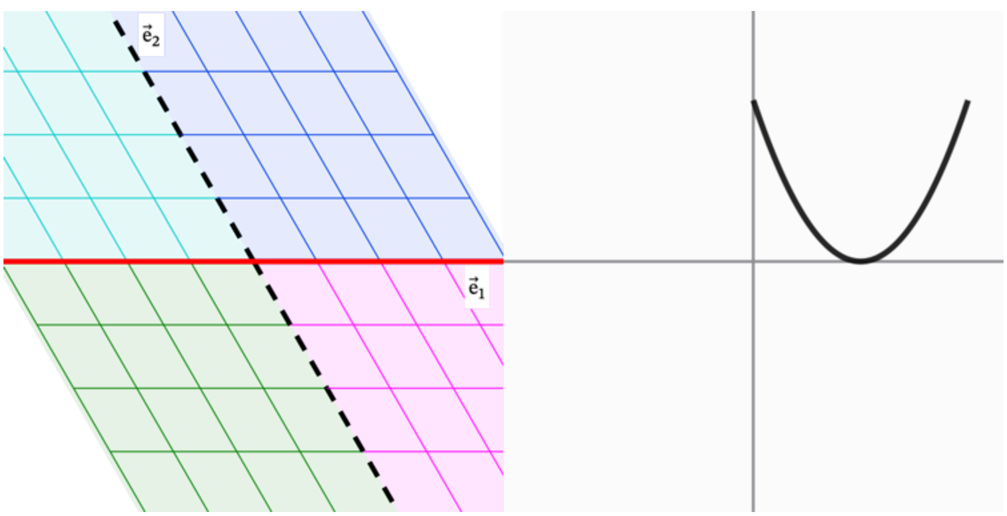
$$\lambda_1 = \lambda_2 = a = a_{11} = a_{22}$$

$$(A - \lambda I)\,\vec{x} = \begin{bmatrix} 0 & b \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$b \neq 0 \rightarrow x_2 = 0 \rightarrow \vec{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Only one eigenvector direction is available
(shown as the red line below)

$$\text{Example : } A = \begin{bmatrix} \approx 1.039 & -0.6 \\ 0 & \approx 1.039 \end{bmatrix}$$



Ⓑ

$$\text{Gradually adjust } a_{21} \text{ closer to -b: } A = \begin{bmatrix} a & b \\ c & a \end{bmatrix}, \quad b\,c < 0, \mid b \mid > \mid c \mid$$

$$\det(A - \lambda I) = (a - \lambda)^2 - b\,c = 0$$

Since $(b\,c) < 0$, denote $\omega = \sqrt{-b\,c}$

$$\downarrow$$

$$\lambda = a \pm i\,\omega$$

For $\lambda_1 = a + i\,\omega$:

$$(A - \lambda_1 I)\,\vec{x} = \begin{bmatrix} -i\,\omega & b \\ c & -i\,\omega \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\downarrow$$

$$x_2 = i\left(\frac{\omega}{b}\right)x_1$$

Choose free variable $x_1 = 1$

$$\vec{x}_1(\text{before normalization}) = \begin{bmatrix} 1 \\ i\,\dfrac{\omega}{b} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + i \begin{bmatrix} 0 \\ \dfrac{\omega}{b} \end{bmatrix}$$

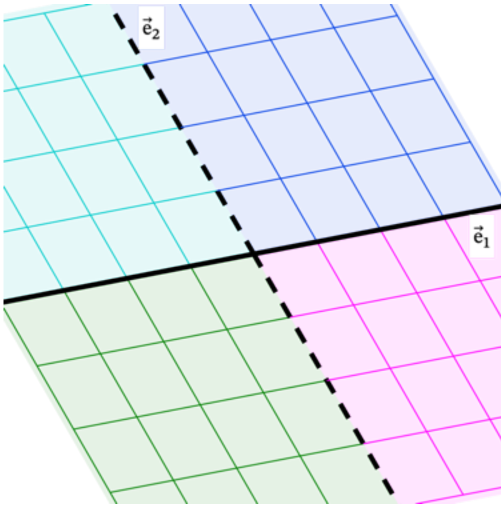
$$\downarrow$$

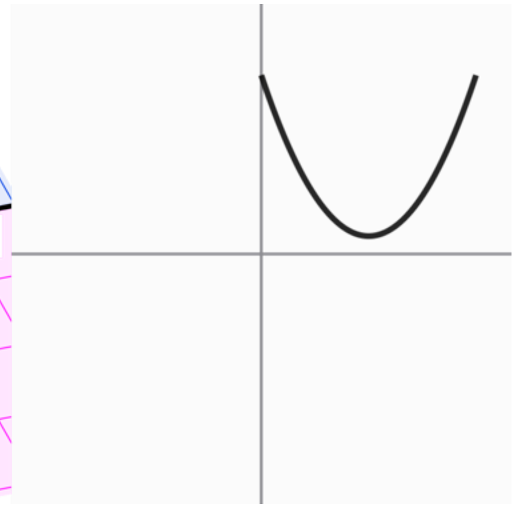
- $\mathrm{Re}(\vec{x}_1) \perp \mathrm{Im}(\vec{x}_1)$

(in these particular cases, because $a_{11} = a_{22}$)

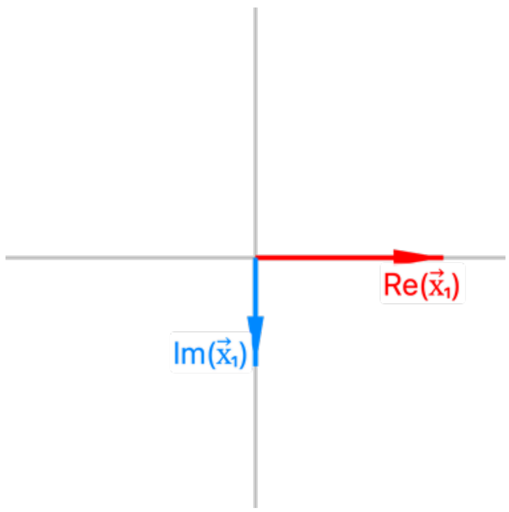
- $|\,\mathrm{Re}(\vec{x}_1)\,| > |\,\mathrm{Im}(\vec{x}_1)\,|$

Example ①: $A = \begin{bmatrix} \approx 1.039 & -0.6 \\ 0.2 & \approx 1.039 \end{bmatrix}$

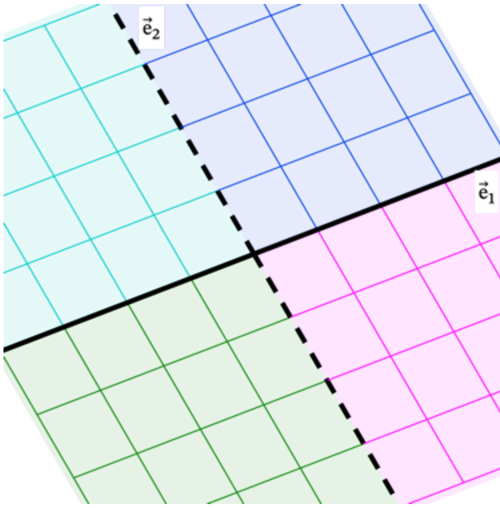


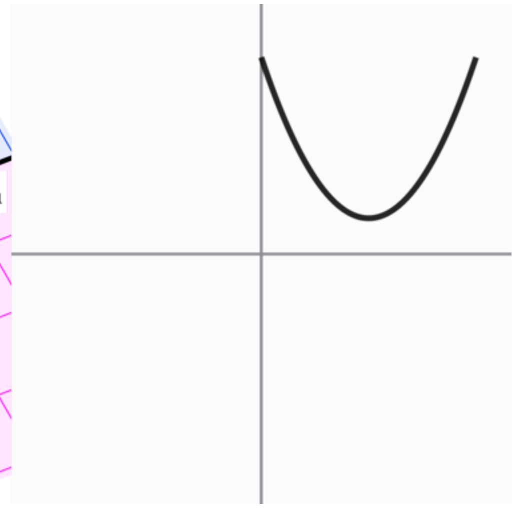


$$\vec{x} = \begin{bmatrix} 0.866 \\ 0 \end{bmatrix} \pm i \begin{bmatrix} 0 \\ -0.5 \end{bmatrix}$$

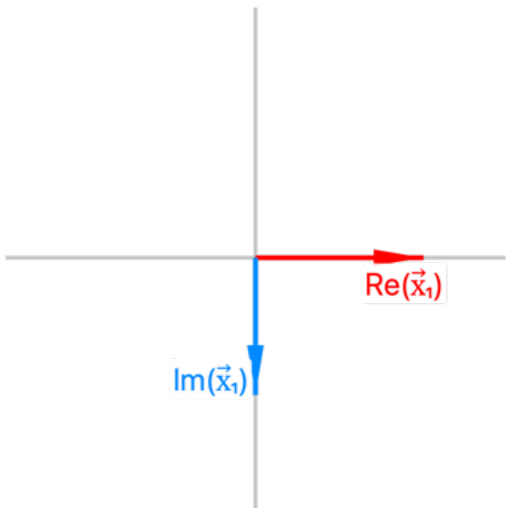


Example ②: $A = \begin{bmatrix} \approx 1.039 & -0.6 \\ 0.4 & \approx 1.039 \end{bmatrix}$





$$\vec{x} = \begin{bmatrix} 0.775 \\ 0 \end{bmatrix} \pm i \begin{bmatrix} 0 \\ -0.632 \end{bmatrix}$$

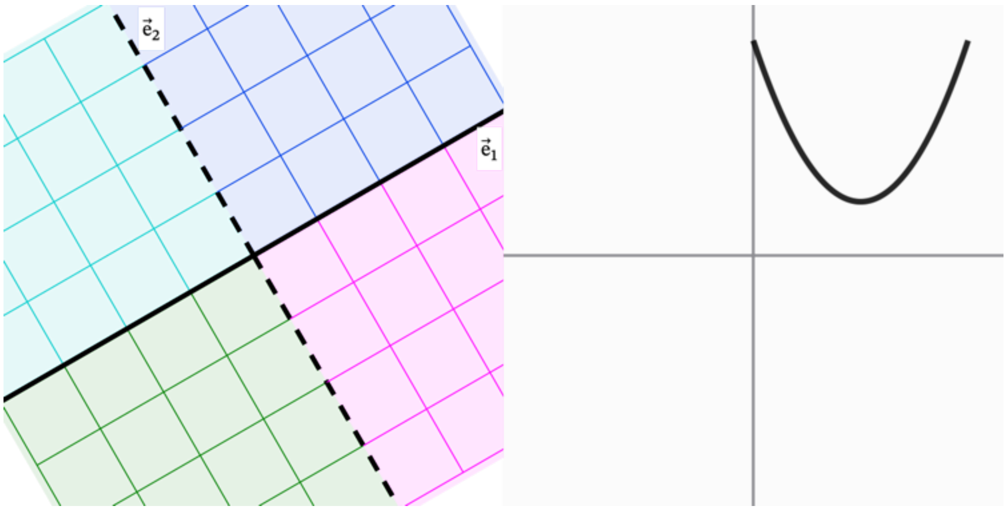


©

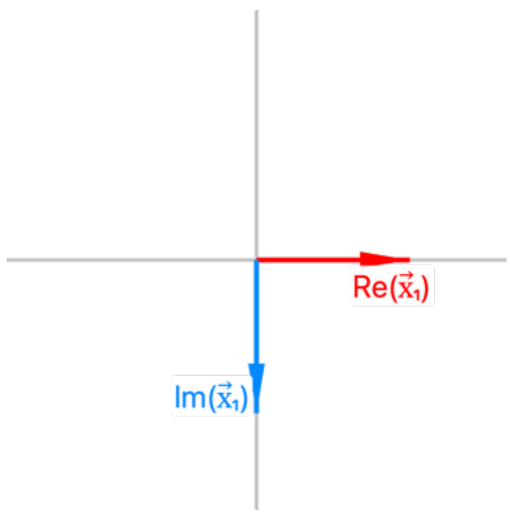
arrive at rotation-scaling matrix: $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ $b \neq 0$

- $\lambda = a \pm ib$
- $Re(\vec{x}_1) \perp Im(\vec{x}_1)$
- $| Re(\vec{x}_1) | = | Im(\vec{x}_1) |$

$$\text{Example : } A = \begin{bmatrix} \approx 1.039 & -0.6 \\ 0.6 & \approx 1.039 \end{bmatrix}$$



$$\vec{\mathbf{x}} = \begin{bmatrix} 0.707 \\ 0 \end{bmatrix} \pm i \begin{bmatrix} 0 \\ -0.707 \end{bmatrix}$$



Rotation-scaling factorization of 2×2 A with $\lambda \in \mathbb{C}$

A is a real-valued matrix with complex λ

- $A\vec{x} = \lambda\vec{x}$
- $\lambda = \alpha + i\beta$
- $\vec{x} = \mathrm{Re}(\vec{x}) + i\,\mathrm{Im}(\vec{x})$

$$A\left(\mathrm{Re}(\vec{x}) + i\,\mathrm{Im}(\vec{x})\right) = (\alpha + i\beta)\left(\mathrm{Re}(\vec{x}) + i\,\mathrm{Im}(\vec{x})\right)$$

$$\downarrow$$

$$A\,\mathrm{Re}(\vec{x}) + i\,A\,\mathrm{Im}(\vec{x}) = \alpha\,\mathrm{Re}(\vec{x}) + i^2\beta\,\mathrm{Im}(\vec{x}) + i\beta\,\mathrm{Re}(\vec{x}) + \alpha\, i\,\mathrm{Im}(\vec{x})$$

$$\downarrow$$

$$A\,\mathrm{Re}(\vec{x}) + i\,A\,\mathrm{Im}(\vec{x}) = \alpha\,\mathrm{Re}(\vec{x}) - \beta\,\mathrm{Im}(\vec{x}) + i\left(\beta\,\mathrm{Re}(\vec{x}) + \alpha\,\mathrm{Im}(\vec{x})\right)$$

Equate real and imaginary components separately:

- $A\,\mathrm{Re}(\vec{x}) = \alpha\,\mathrm{Re}(\vec{x}) - \beta\,\mathrm{Im}(\vec{x})$
- $A\,\mathrm{Im}(\vec{x}) = \beta\,\mathrm{Re}(\vec{x}) + \alpha\,\mathrm{Im}(\vec{x})$

$$\downarrow$$

$$A\begin{bmatrix} \mathrm{Re}(\vec{x}) \\ \mathrm{Im}(\vec{x}) \end{bmatrix} = \begin{bmatrix} A\,\mathrm{Re}(\vec{x}) \\ A\,\mathrm{Im}(\vec{x}) \end{bmatrix} = \left[\begin{array}{c|c} \alpha & -\beta \\ \hline \beta & \alpha \end{array}\right]\begin{bmatrix} \mathrm{Re}(\vec{x}) \\ \mathrm{Im}(\vec{x}) \end{bmatrix}$$

Block-transpose both sides:

$$A\left[\begin{array}{c|c} \mathrm{Re}(\vec{x}) & \mathrm{Im}(\vec{x}) \end{array}\right] = \left[\begin{array}{c|c} \mathrm{Re}(\vec{x}) & \mathrm{Im}(\vec{x}) \end{array}\right]\left[\begin{array}{c|c} \alpha & \beta \\ \hline -\beta & \alpha \end{array}\right]$$

Denote

$$X = \begin{bmatrix} \mathrm{Re}(\vec{x}) & \mathrm{Im}(\vec{x}) \end{bmatrix} \quad S = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$$

$$\downarrow$$

$$A\,X = X\,S$$

If X is invertible,

$$A = X\,S\,X^{-1}$$

Transformation proceeds as follows:

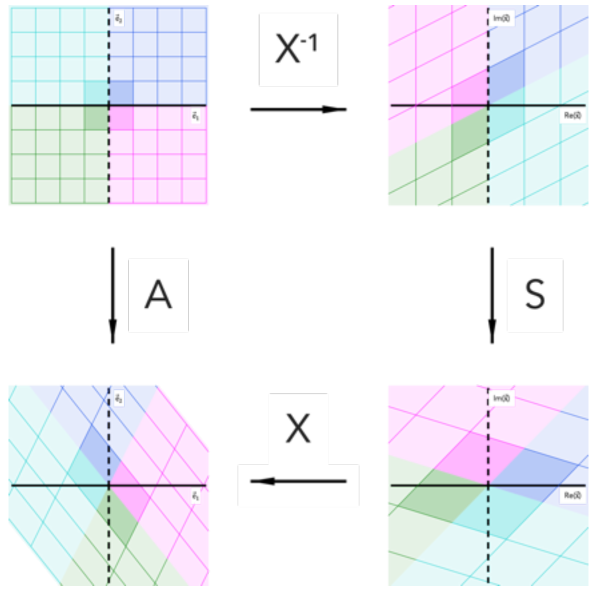
- X^{-1} changes coordinates to the $\begin{bmatrix} \mathrm{Re}(\vec{x}) \mid \mathrm{Im}(\vec{x}) \end{bmatrix}$ basis
- S rotates and scales in that basis
- X changes coordinates back to the standard basis

Note that $\det(S) = \alpha^2 + \beta^2 = |\lambda_1|^2 = |\lambda_2|^2$

Example with free variable $x_2 = 1$ and eigenvectors normalized:

$$A = \begin{bmatrix} 0.5 & -1.2 \\ 1 & 1.5 \end{bmatrix} =$$

$$\begin{bmatrix} \approx -0.337 & \approx 0.657 \\ \approx 0.674 & 0 \end{bmatrix} \begin{bmatrix} 1 & \approx 0.975 \\ \approx -0.975 & 1 \end{bmatrix} \begin{bmatrix} 0 & \approx 1.483 \\ \approx 1.522 & \approx 0.761 \end{bmatrix}$$



If matrix $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ is already in the rotation-scaling form,

choosing $\lambda = a - i\,b$ and free variable $x_2 = i$ gives $\vec{x} = \begin{bmatrix} 1 \\ i \end{bmatrix}$

$$\downarrow$$

$$\mathrm{Re}(\vec{x}) = \vec{e}_1 \text{ and } \mathrm{Im}(\vec{x}) = \vec{e}_2$$

$$\downarrow$$

$$X = I,\ S = A$$

Other choices give different invertible X matrices,
but the same final transformation A

Geometric meaning of $A\vec{x} = \lambda\vec{x}$ when $\lambda \in \mathbb{C}$

Suppose

- $A\vec{x} = \lambda\vec{x}$
- $\lambda \in \mathbb{C}$
- $\lambda = \alpha + i\beta$
- $\vec{x} = \mathrm{Re}(\vec{x}) + i\,\mathrm{Im}(\vec{x})$

If $\lambda(B) \in \mathbb{R}$, and $B\vec{v} = \lambda\vec{v}$, $\vec{v}$ is the direction that does not change

Here, we will try to find the equivalent for A and $\vec{x}$

from rotation-scaling theorem:

$$A\left[\begin{array}{c|c} \mathrm{Re}(\vec{x}) & \mathrm{Im}(\vec{x}) \end{array}\right] = \left[\begin{array}{c|c} \mathrm{Re}(\vec{x}) & \mathrm{Im}(\vec{x}) \end{array}\right]\left[\begin{array}{c|c} \alpha & \beta \\ \hline -\beta & \alpha \end{array}\right]$$

$$\downarrow$$

$$A\left[\begin{array}{c|c} \mathrm{Re}(\vec{x}) & \mathrm{Im}(\vec{x}) \end{array}\right] = \left[\begin{array}{c|c} \mathrm{Re}(\vec{x}) & \mathrm{Im}(\vec{x}) \end{array}\right] S \quad \text{where } S \text{ is a rotation-scaling matrix}$$

Unlike left-multiplication, right multiplication by S does not produce a distinctive visible pattern

The distinctive image appears when A itself is a rotation-scaling matrix where $\mathrm{Re}(\vec{x})$ and $\mathrm{Im}(\vec{x})$ are orthogonal and equal in length

$$A = \left[\begin{array}{c|c} a & -b \\ \hline b & a \end{array}\right]$$

$$\vec{x}(A) = \frac{1}{\sqrt{2}} \begin{bmatrix} i \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} + i \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$$

$$A\,\vec{x} = \frac{1}{\sqrt{2}} \begin{bmatrix} -b + ai \\ a + bi \end{bmatrix} = \begin{bmatrix} -\frac{b}{\sqrt{2}} \\ \frac{a}{\sqrt{2}} \end{bmatrix} + i \begin{bmatrix} \frac{a}{\sqrt{2}} \\ \frac{b}{\sqrt{2}} \end{bmatrix}$$

$\downarrow$

The new real and imaginary parts are still orthogonal and equal in length:

- $A\ \mathrm{Re}(\vec{x}) \perp A\ \mathrm{Im}(\vec{x})$
- $|A\ \mathrm{Re}(\vec{x})| = |A\ \mathrm{Im}(\vec{x})|$

So for a rotation-scaling matrix, the complex eigenvector produces a distinctive visible pair before and after applying A.

Rotation-scaling factorization for powers of A

$$\text{Rotation-scaling matrix } S = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} = \rho \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

- $\rho = \sqrt{\alpha^2 + \beta^2} = \sqrt{\det(S)}$
- $\theta = \operatorname{atan2}(\beta, \alpha)$

S rotates by θ and scales by ρ

↓

S^n rotates by nθ and scales by ρ^n

↓

$$S^n = \rho^n \left[\begin{array}{c|c} \cos(n\theta) & \sin(n\theta) \\ \hline -\sin(n\theta) & \cos(n\theta) \end{array}\right]$$

Suppose A is a real-valued matrix with complex λ:

$A = X S X^{-1}$

Ⓐ

Integer powers:

$A^2 = (X S X^{-1})(X S X^{-1})$

↓

$A^2 = X S^2 X^{-1}$

$A^n = X S^n X^{-1}$

can be proven in the same fashion for any integer n

Ⓑ

Rational powers:

Suppose

- $p = \dfrac{1}{q}$, where q is a positive integer
- we need to compute A^p

Use the principal angle $\theta = \operatorname{atan2}(\beta, \alpha)$

This chooses the principal fractional power of S:

$$S^{\frac{1}{q}} \text{ rotates by } \frac{\theta}{q} \text{ and scales by } \rho^{\frac{1}{q}}$$

Define

$$T = S^{p} = S^{\frac{1}{q}}$$

$$\downarrow$$

$$T^{q} = S$$

Define $B = X\,T\,X^{-1}$

$$B^{q} = (X\,T\,X^{-1})^{q} = X\,T^{q}\,X^{-1} = X\,S\,X^{-1} = A$$

$$\downarrow$$

$$B = A^{\frac{1}{q}}$$

$$\downarrow$$

$$A^{p} = B = X\,T\,X^{-1}$$

$$\downarrow$$

$$A^{p} = X\,S^{p}\,X^{-1}$$

Ⓒ

For a general rational power $t = s\,p$ where s is an integer:

$$A^{t} = \left(X\,S^{p}\,X^{-1}\right)^{s} = X\,S^{(s\,p)}\,X^{-1} = X\,S^{t}\,X^{-1}$$

Thus, for integer, fractional, and rational powers:

A^t is computed by

- moving to the $\begin{bmatrix} \mathrm{Re}(\vec{x}) & \mathrm{Im}(\vec{x}) \end{bmatrix}$ basis by X^{-1}
- applying S^t
- moving back by X

Rotation-scaling factorization for powers of A
(numerical example)
For a matrix with complex eigenvalues,
$A = X\,S\,X^{-1}$
so powers are computed by
$A^t = X\,S^t\,X^{-1}$

$$A = \begin{bmatrix} \approx 0.911 & \approx 0.212 \\ -0.142 & \approx 1.109 \end{bmatrix} =$$

$$\begin{bmatrix} \approx 0.443 & \approx -0.634 \\ \approx 0.634 & 0 \end{bmatrix} \begin{bmatrix} \approx 1.01 & \approx 0.142 \\ \approx -0.142 & \approx 1.01 \end{bmatrix} \begin{bmatrix} 0 & \approx 1.578 \\ \approx -1.578 & \approx 1.104 \end{bmatrix}$$

For the rotation-scaling matrix S,

$|\lambda| = \sqrt{\alpha^2 + \beta^2}$ and $\theta = \mathrm{atan2}(\beta, \alpha)$

So S^t rotates by $t\,\theta$ and scales by $|\lambda|^t$

Here $|\lambda| \approx 1.02 > 1$, so the powers rotate while expanding slowly

Parameter t runs from 0 to n:

- $t = 0$ gives the identity transformation
- integer t gives $A, A^2, \dots, A^n$
- non-integer t gives fractional powers between them

In the images below:

- green points are the starting points
- smaller dots: $A^t\ \vec{v}$ at fraction powers
- circled blue dots: $A^t\ \vec{v}$ at integer powers
- cyan lines: radius decreases after next step
- magenta lines: radius increases after next step

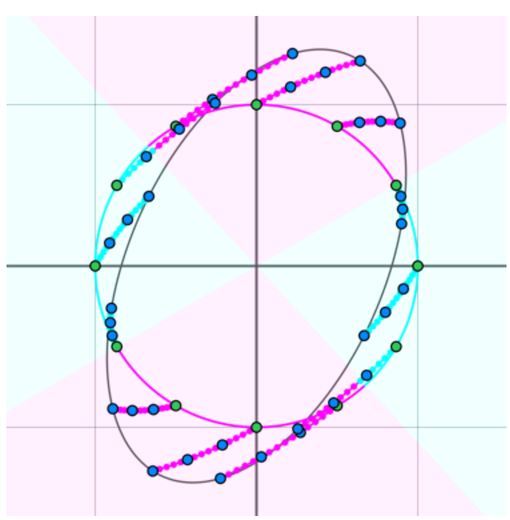
Each image starts with the unit circle in the coordinate system being used.
So these images are not consecutive steps applied to the same points

- For a generic matrix A, the unit circle usually becomes an ellipse
- For the rotation-scaling matrix S, the unit circle remains circular

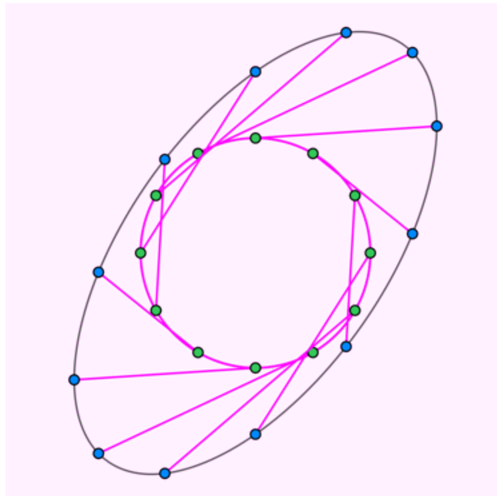
A^t on the standard unit circle

Start with points on the ordinary unit circle

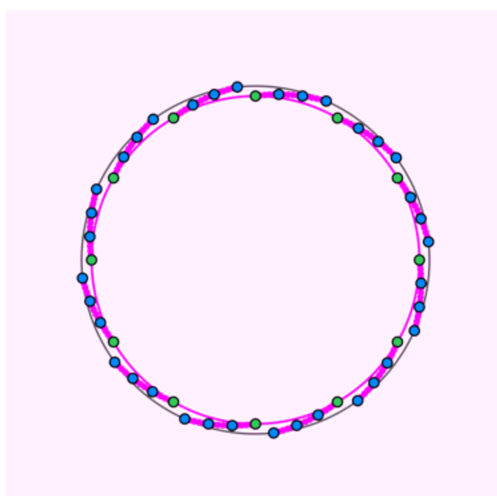
Apply A^t as t moves from 0 to n



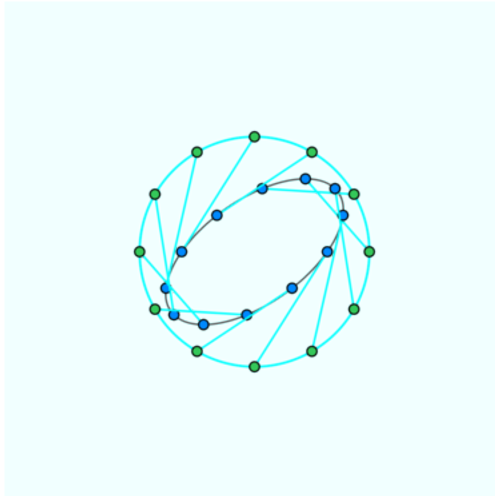
The same starting circle after X^{-1}
This moves the original starting points into the rotation-scaling basis
The circle is not expected to remain a circle



S^t on the unit circle in the rotation-scaling basis
This picture isolates what S does by itself
Starting again from the unit circle,
S^t rotates by $t\theta$ and scales by $|\lambda|^t$ as t moves from 0 to n



Move the S^t picture back by X
The change of basis turns the rotation-scaling picture back into the original coordinates



Following the same points through the factorization

The previous images restarted from the unit circle for each matrix
This image follows the same starting points through the factorization of the final power A^n

① Start with identity grid

② Move into the rotation-scaling coordinates:

$$v \rightarrow X^{-1} v$$

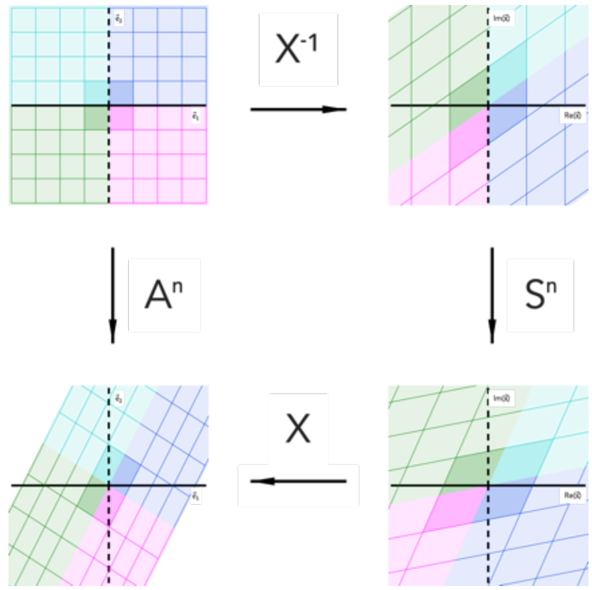
③ Apply the powers of S up to n:

$$X^{-1} v \rightarrow S^n X^{-1} v$$

④ Move the result back to the original coordinates:

$$S^n X^{-1} v \rightarrow X S^n X^{-1} v = A^n v$$

Here the panels are consecutive steps:
each panel uses the output from the previous one



Powers of A, more numerical examples

Ⓐ

We construct A as $A = X\,S\,X^{-1}$

- An arbitrary X is kept constant
- 6 instances of S are chosen as follows:

S is defined by λ and rotation angle θ and $\lambda(S) = \lambda(A)$

ⓐ

Since $|\lambda|$ controls expansion or contraction, we chose examples with

- $|\lambda| < 1$
- $|\lambda| = 1$
- $|\lambda| > 1$

For images of S^t:

- $|\lambda| < 1$: every orbit moves toward the origin
- $|\lambda| = 1$: every orbit stays on a circle
- $|\lambda| > 1$: every orbit moves away from the origin

For images of A:

- For any λ, radius may change within a given orbit
- Area scaling is consistent:

$\det(A) = \det(S) = |\lambda|^2$

ⓑ

θ choices: $\frac{\pi}{10}, \frac{\pi}{5}$

ⓒ

Chosen $X = \begin{bmatrix} 1 & 0.75 \\ 0 & 1 \end{bmatrix}$ $X^{-1} = \begin{bmatrix} 1 & -0.75 \\ 0 & 1 \end{bmatrix}$

Unlike S, X does not conform to any particular matrix type

Ⓑ

Orbit images:

ⓐ

The colored unit circle classifies starting directions:

- cyan: radius decreases after the next step
- magenta: radius increases after the next step
- gray: radius is unchanged

The circle changes color where the step preserves radius:

$$| A \vec{v} | = | \vec{v} |$$

$$\downarrow$$

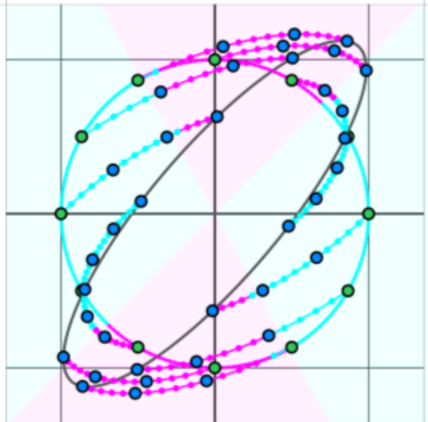
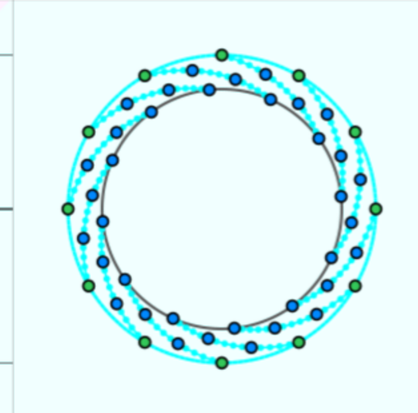
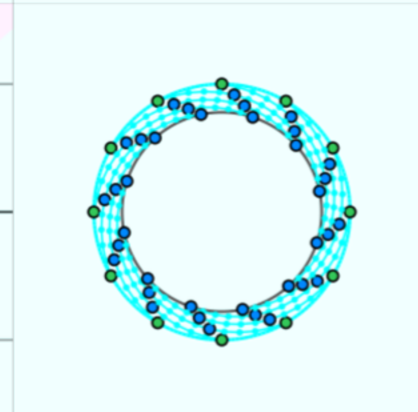
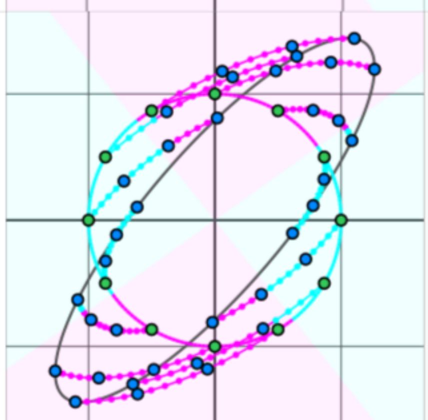
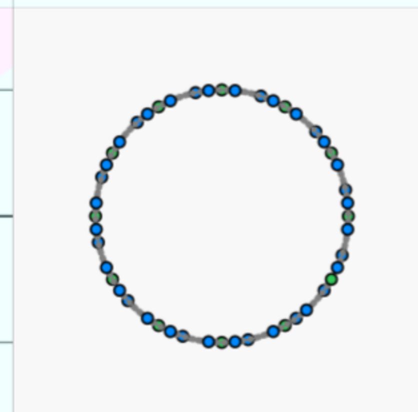
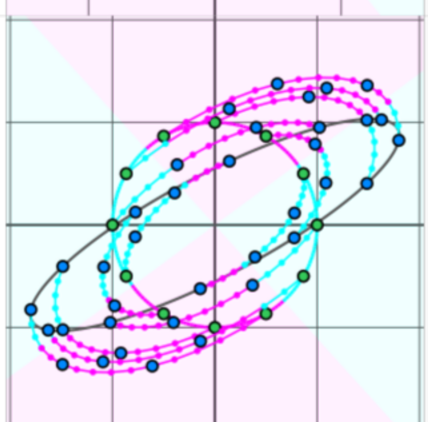
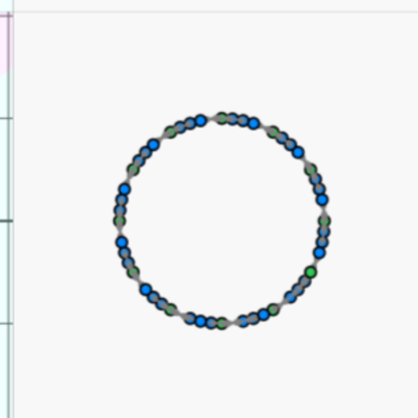
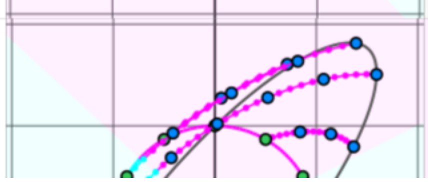
$$\vec{v}^T (A^T A - I) \vec{v} = 0$$

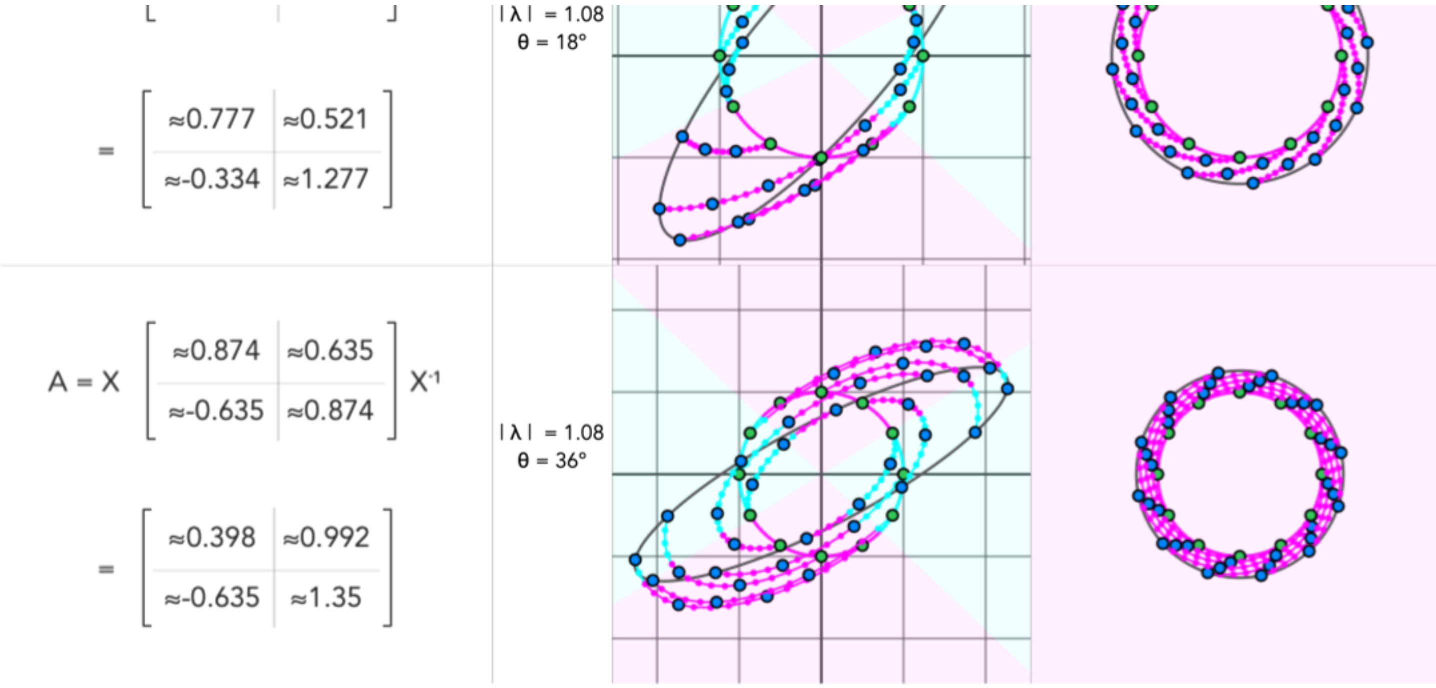
Generically, these transition directions form two lines through the origin, dividing the plane into two magenta sectors and two cyan sectors.

ⓑ

Each curve is the orbit of one starting point:
$v, A v, A^2 v, \ldots, A^n v$ in the A^t column
$v, S v, S^2 v, \ldots, S^n v$ in the S^t column

- green points are the starting points on the unit circle
- small dots show the radius at fraction powers
- larger dots mark integer powers: $t = 1, 2, \ldots, n$

$A = X S X^{-1}$	$\vert\lambda\vert$ and θ	A^t	S^t
$A = X \begin{bmatrix} \approx 0.875 & \approx 0.284 \\ \approx -0.284 & \approx 0.875 \end{bmatrix} X^{-1}$ $= \begin{bmatrix} \approx 0.662 & \approx 0.444 \\ \approx -0.284 & \approx 1.088 \end{bmatrix}$	$\vert\lambda\vert = 0.92$ $\theta = 18°$		
$A = X \begin{bmatrix} \approx 0.744 & \approx 0.541 \\ \approx -0.541 & \approx 0.744 \end{bmatrix} X^{-1}$ $= \begin{bmatrix} \approx 0.339 & \approx 0.845 \\ \approx -0.541 & \approx 1.15 \end{bmatrix}$	$\vert\lambda\vert = 0.92$ $\theta = 36°$		
$A = X \begin{bmatrix} \approx 0.951 & \approx 0.309 \\ \approx -0.309 & \approx 0.951 \end{bmatrix} X^{-1}$ $= \begin{bmatrix} \approx 0.719 & \approx 0.483 \\ \approx -0.309 & \approx 1.183 \end{bmatrix}$	$\vert\lambda\vert = 1$ $\theta = 18°$		
$A = X \begin{bmatrix} \approx 0.809 & \approx 0.588 \\ \approx -0.588 & \approx 0.809 \end{bmatrix} X^{-1}$ $= \begin{bmatrix} \approx 0.368 & \approx 0.918 \\ \approx -0.588 & \approx 1.25 \end{bmatrix}$	$\vert\lambda\vert = 1$ $\theta = 36°$		
$A = X \begin{bmatrix} \approx 1.027 & \approx 0.334 \\ \approx -0.334 & \approx 1.027 \end{bmatrix} X^{-1}$			



Powers of A,
eigenvalues and eigenvectors of (A^T A)

$$A = \begin{bmatrix} \approx 0.911 & \approx 0.212 \\ -0.142 & \approx 1.109 \end{bmatrix} = X\,S\,X^{-1} =$$

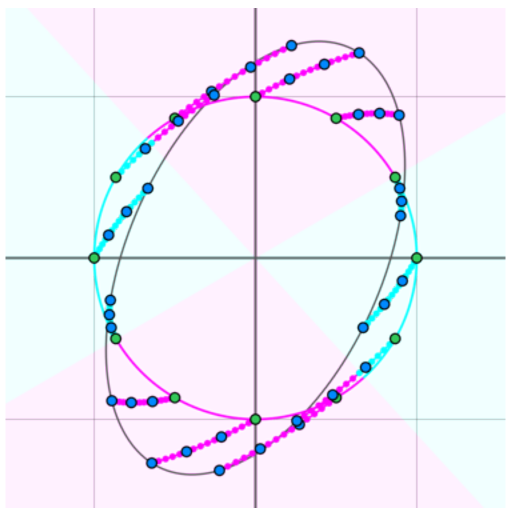
$$\begin{bmatrix} \approx 0.443 & \approx -0.634 \\ \approx 0.634 & 0 \end{bmatrix}\begin{bmatrix} \approx 1.01 & \approx 0.142 \\ \approx -0.142 & \approx 1.01 \end{bmatrix}\begin{bmatrix} 0 & \approx 1.578 \\ \approx -1.578 & \approx 1.104 \end{bmatrix}$$

Eigenvalues(A)

- $\lambda_1 \approx 1.01 + 0.142\,i$
- $\lambda_2 \approx 1.01 - 0.142\,i$
- $|\lambda_1| = |\lambda_2| \approx 1.02$
- $\dfrac{\text{ellipse area}}{\text{unit circle area}} = |\lambda\ \ |^t$ for any t

The image below shows successive powers of A with the ellipse outlining A^3

The colored sectors classify directions according to whether the next step increases or decreases the amplitude.



Color change occurs when

$$|A\vec{v}| = |\vec{v}|$$

$$\downarrow$$

$$(A\vec{v})^T (A\vec{v}) = \vec{v}^T \vec{v}$$

$$\downarrow$$

$$\vec{v}^T A^T A \vec{v} = \vec{v}^T I \vec{v}$$

$$\downarrow$$

$$\vec{v}^T (A^T A - I) \vec{v} = 0$$

To solve for $\vec{v}$, define $M = A^T A = \begin{bmatrix} \approx 0.85 & \approx 0.035 \\ \approx 0.035 & \approx 1.276 \end{bmatrix}$

with eigenvalues μ and normalized eigenvectors û

$$M = M^T$$
$$\downarrow$$

- μ_1 and $\mu_2 \in \mathbb{R}$
- μ_1 and $\mu_2 >= 0$
- $\hat{u}_1$ and $\hat{u}_2 \in \mathbb{R}^2$
- $\hat{u}_1 \cdot \hat{u}_2 = 0$

- $\hat{u}_i^T A^T A \hat{u}_i = (A \hat{u}_i)^T (A \hat{u}_i) = | A \hat{u}_i |^2$
- $(A^T A) \hat{u}_i = \mu_i \hat{u}_i$

$$\downarrow$$
$$\hat{u}_i^T \mu_i \hat{u}_i = \mu_i (\hat{u}_i^T \hat{u}_i) = \mu_i = | A \hat{u}_i |^2$$
$$\downarrow$$
$$| A \hat{u}_i | = \sqrt{\mu_i}$$

Define orthogonal matrix $U = [\ \hat{u}_1 \mid \hat{u}_2\ ]$

$$\downarrow$$
$$U^{-1} = U^T \text{ and } I = U\, I\, U^T$$
$$\downarrow$$

M can be orthogonally diagonalized as

$$M = A^T A = U D U^T = U \begin{bmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{bmatrix} U^T$$

$$\vec{v}^T (A^T A - I)\, \vec{v} = 0$$

$$\downarrow$$

$$\vec{v}^T (U D U^T - U I U^T)\, \vec{v} = 0$$

$$\downarrow$$

$$\vec{v}^T U (D - I) U^T \vec{v} = 0$$

$$\downarrow$$

$$(\vec{v}^T U) \begin{bmatrix} \mu_1 - 1 & 0 \\ 0 & \mu_2 - 1 \end{bmatrix} (U^T \vec{v}) = 0$$

$$\text{Denote } \vec{z} = (\vec{v}^T U) = \begin{bmatrix} \vec{z}_1 \\ \vec{z}_2 \end{bmatrix}$$

$$\downarrow$$

$$(\mu_1 - 1)\, \vec{z}_1^{\,2} + (\mu_2 - 1)\, \vec{z}_2^{\,2} = 0$$

$$\downarrow$$

- $\mu_1 > 1$ and $\mu_2 > 1$ → every direction expands
- $\mu_1 < 1$ and $\mu_2 < 1$ → every direction contracts
- $\mu_1 > 1 > \mu_2$ → some directions expand and some contract

In the mixed case, transition directions satisfy

$$\frac{z_2}{z_1} = \pm\sqrt{\frac{\mu_1 - 1}{1 - \mu_2}}$$

and form two lines through the origin in the [$\hat{u}_1$ | $\hat{u}_2$] coordinate system:

$$\vec{v} = U \vec{z}$$

$$\downarrow$$

U rotates these lines into color transitions in the original coordinates

The next image shows the same orbits, with the ellipse outlining A^1

⦿ red lines: eigenvector directions of $A^T A$

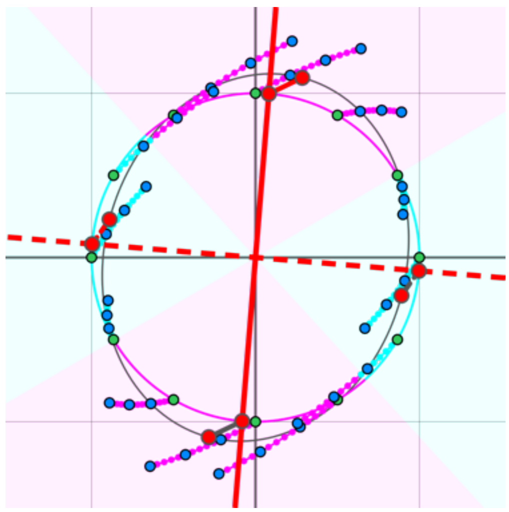
• $\hat{u}_1$ produces the largest radius $| A \hat{u}_1 |$

• $\hat{u}_2$ produces the smallest radius $| A \hat{u}_2 |$

⦿ short red segments: $\pm\hat{u}_i \rightarrow \pm A \hat{u}_i$

• $A \hat{u}_1$ lands at the largest-radius points of the A^1 ellipse

• $A \hat{u}_2$ lands at the smallest-radius points



Two more numerical examples:

①

$\mu_1 < 1$ and $\mu_2 < 1$

$$A = \begin{bmatrix} \approx 0.729 & \approx 0.169 \\ \approx -0.114 & \approx 0.888 \end{bmatrix} \quad A^T A = \begin{bmatrix} \approx 0.544 & \approx 0.022 \\ \approx 0.022 & \approx 0.816 \end{bmatrix}$$

eigenvalues(A)

• $\lambda_1 \approx 0.808 + 0.114\, i$

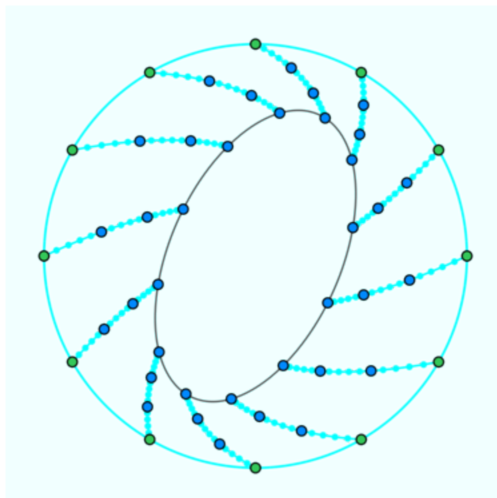
- $\lambda_2 \approx 0.808 - 0.114\ i$
- $|\lambda_1| = |\lambda_2| \approx 0.816$

eigenvalues($A^T A$)

- $\mu_1 \approx 0.818$
- $\mu_2 \approx 0.542$

↓

Every direction contracts



②

$\mu_1 > 1$ and $\mu_2 > 1$

$$A = \begin{bmatrix} \approx 1.093 & \approx 0.254 \\ \approx -0.17 & \approx 1.331 \end{bmatrix} \quad A^T A = \begin{bmatrix} \approx 1.223 & \approx 0.051 \\ \approx 0.051 & \approx 1.837 \end{bmatrix}$$

eigenvalues(A)

- $\lambda_1 \approx 1.212 + 0.17\ i$
- $\lambda_2 \approx 1.212 - 0.17\ i$
- $|\lambda_1| = |\lambda_2| \approx 1.224$

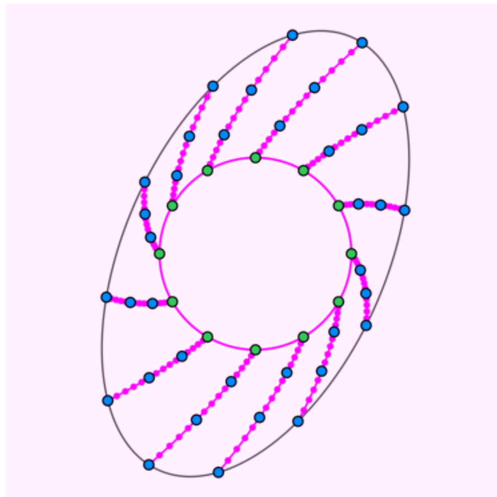
eigenvalues($A^T A$)

- $\mu_1 \approx 1.841$

• $\mu_2 \approx 1.219$

↓

Every direction expands



Rotation-scaling factorization
for matrix exponentials

Suppose A is a real-valued matrix with complex λ,
and we have the rotation-scaling factorization

$$A = X\,S\,X^{-1}$$

e^{tA} is defined by its Taylor series:

$$e^{tA} = I + tA + \left(\frac{t^2}{2!}\right)A^2 + \left(\frac{t^3}{3!}\right)A^3 + \ldots$$

Substitute $A^n = X S^n X^{-1}$:

$$e^{tA} = X I X^{-1} + t X S X^{-1} + \left(\frac{t^2}{2!}\right) X S^2 X^{-1} + \left(\frac{t^3}{3!}\right) X S^3 X^{-1} + \dots$$

Factor out X on the left and X^{-1} on the right:

$$e^{tA} = X \left(I + t S + \left(\frac{t^2}{2!}\right) S^2 + \left(\frac{t^3}{3!}\right) S^3 + \dots \right) X^{-1}$$

The expression in parentheses is the Taylor series for e^{ts}:

$$I + t S + \left(\frac{t^2}{2!}\right) S^2 + \left(\frac{t^3}{3!}\right) S^3 + \dots = e^{ts}$$

$$\downarrow$$

$$e^{tA} = X e^{ts} X^{-1}$$

So the exponential of A is computed by taking the exponential of the simpler rotation-scaling block S

Now compute e^{ts}

$$S = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$$

Split S into scaling and rotation-generator parts:

$$S = \alpha I + \beta J$$

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

The identity matrix I commutes with J

$$e^{ts} = e^{t(\alpha I + \beta J)} = e^{\alpha t I} e^{\beta t J}$$

Also,

$$e^{\alpha t I} = e^{\alpha t} I$$

So now we only need to compute $e^{\beta t J}$

Use the Taylor series:

$$e^{\beta t J} = I + \beta t J + \left(\frac{(\beta t)^2}{2!}\right) J^2 + \left(\frac{(\beta t)^3}{3!}\right) J^3 + \left(\frac{(\beta t)^4}{4!}\right) J^4 + \dots$$

The powers of J repeat:

- $J^0 = I$
- $J^1 = J$
- $J^2 = -I$
- $J^3 = -J$
- $J^4 = I$

Substitute these powers:

$$e^{\beta t J} = I + \beta t J - \left(\frac{(\beta t)^2}{2!}\right) I - \left(\frac{(\beta t)^3}{3!}\right) J$$

$$+ \left(\frac{(\beta t)^4}{4!}\right) I + \left(\frac{(\beta t)^5}{5!}\right) J - \ldots$$

Group the I terms and the J terms:

$$e^{\beta t J} = I\left(1 - \frac{(\beta t)^2}{2!} + \frac{(\beta t)^4}{4!} - \ldots\right)$$

$$+ J\left(\beta t - \frac{(\beta t)^3}{3!} + \frac{(\beta t)^5}{5!} - \ldots\right)$$

These are the Taylor series for cosine and sine:

$$e^{\beta t J} = I\cos(\beta t) + J\sin(\beta t)$$

$$e^{\beta t J} = \left[\begin{array}{c|c} \cos(\beta t) & \sin(\beta t) \\ \hline -\sin(\beta t) & \cos(\beta t) \end{array}\right]$$

$$\downarrow$$

$$e^{ts} = e^{\alpha t}\left[\begin{array}{c|c} \cos(\beta t) & \sin(\beta t) \\ \hline -\sin(\beta t) & \cos(\beta t) \end{array}\right]$$

$$e^{tA} = X\, e^{ts}\, X^{-1}$$

$$\downarrow$$

e^{tA} is computed by

- moving to the $\left[\ \mathrm{Re}(\vec{x})\ \ \mathrm{Im}(\vec{x})\ \right]$ basis by X^{-1}

- applying e^{ts}
- moving back by X

The second complex eigenvector

This page is only informational, since
rotation-scaling factorization uses only one eigenvector

As we previously showed: Suppose

$$A\,\vec{x}_1 = \lambda_1\,\vec{x}_1$$

$$\downarrow$$

$$\vec{x}_2 = \vec{x}_1{}^{c}$$

is one of the valid eigenvectors associated with λ_2,

which is the usual way to obtain $\vec{x}_2$

But if $\vec{x}_2$ is computed separately by solving

$$\det(A - \lambda_2\,I) = 0$$

the displayed result depends on the free-variable used

Assume x_2 is the free variable.
For λ_1, row reduction gives

$$x_1 = \left(\frac{-b}{a - \lambda_1}\right) x_2$$

Choosing $x_2 = s$ gives

$$\vec{x}_1 = \begin{bmatrix} \dfrac{-s\,b}{a - \lambda_1} \\ s \end{bmatrix} = \begin{bmatrix} \dfrac{-s\,b}{(a - \alpha) - i\beta} \\ s \end{bmatrix}$$

For λ_2, row reduction gives

$$x_1 = \left(- \frac{b}{a - \lambda_2} \right) x_2$$

To get $\vec{x}_2 = \vec{x}_1{}^c$ from row reduction,
choose the conjugate free-variable value $x_2 = s^c$

$$\vec{x}_2 = \begin{bmatrix} \dfrac{-s^c\,b}{a - \lambda_2} \\ s^c \end{bmatrix} = s^c \begin{bmatrix} \dfrac{-b}{(a - \alpha) + i\beta} \\ 1 \end{bmatrix} = \vec{x}_1{}^c$$

So row reduction gives $\vec{x}_2 = \vec{x}_1{}^c$ only if
the λ_2 system uses the conjugate free-variable value

In particular, this automatically happens when s is real

But if both systems use the same non-real free variable s,

the second vector is still a valid eigenvector for λ_2,
but it is not usually the complex conjugate of $\vec{x}_1$

with the same free variable s, $\vec{x}_2 = \begin{bmatrix} \dfrac{-s\,b}{a - \lambda_2} \\ s \end{bmatrix} = \begin{bmatrix} \dfrac{-s\,b}{(a - \alpha) + i\beta} \\ s \end{bmatrix}$

This $\vec{x}_2$ satisfies $A\,\vec{x}_2 = \lambda_2\,\vec{x}_2$,
but unless $s = s^c$, it is not the same displayed vector as $\vec{x}_1{}^c$.

Bottom line:

- The second eigenvector can always be obtained by conjugation
- Row reduction gives the conjugate vector only with conjugate-compatible free variables
- The rotation-scaling factorization only needs $\vec{x}_1 = \mathrm{Re}(\vec{x}_1) + i\,\mathrm{Im}(\vec{x}_1)$