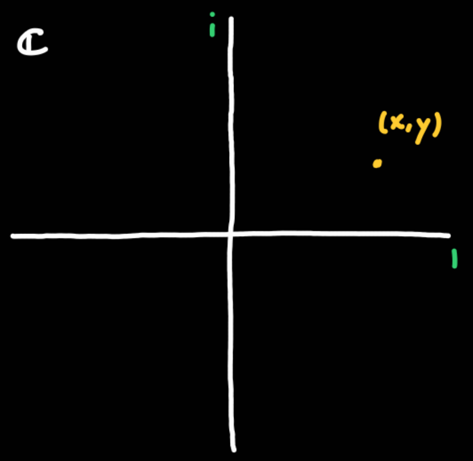
5. Complex numbers

Complex numbers are scalars that extend real numbers. We can think of them geometrically and algebraically. Geometrically, they are points in the two dimensional plane:



where the horizontal axis denotes the real part of the number, and the vertical axis denotes the imaginary part of the number.

The real axis is labeled by 1, the same one as in the real numbers, and the imaginary axis is labeled by i, an imaginary number satisfying $i^2=-1$. Algebraically, the point (x,y) of the complex plane is denoted $z=x+iy$, where x is said to be the real part of z and y is said to be the imaginary part of z. Given two complex numbers $z=x+iy$ and $w=a+ib$, their addition is:

$$z+w=(x+a)+i(y+b)$$

and their multiplication is:

$$zw=(xa-yb)+i(xb+ya).$$

The mnemonic to remember this multiplication is to distribute the parentheses:

$$zw = (x+iy)(a+ib) = xa + ixb + iya + i^2yb = xa + ixb + iya - yb = (xa-yb) + i(xb+ya)$$

using $i^2=-1$. The <u>modulus</u> of $z = x+iy$ is the length of the vector (x,y), and is denoted $|z|$. Namely:

$$|z| = \sqrt{x^2+y^2}.$$

The <u>conjugate</u> of $z = x+iy$ is the reflection of (x,y) across the real axis, denoted $\bar{z}$.

Namely:

$$\bar{z} = x - iy.$$

Applying these definitions we obtain the following properties of complex numbers:

(i) $|zw| = |z||w|$ for all z and w in $\mathbb{C}$.

(ii) $\overline{zw} = \bar{z}\,\bar{w}$ for all z and w in $\mathbb{C}$.

(iii) $zw = wz$ for all z and w in $\mathbb{C}$.

(iv) $z\bar{z} = |z|^2$ for all z in $\mathbb{C}$.

(v) $\frac{z}{w} = \frac{z\bar{w}}{|w|^2}$ for all z and w in $\mathbb{C}$.

(vi) Let $z = x+iy$, then $x = \frac{1}{2}(z+\bar{z})$ and $y = \frac{1}{2}(z-\bar{z})$.

In particular, if z is a non-zero complex number, we can divide by z.

<u>Example:</u> Let $z = 1+2i$ and $w = 3-4i$. Compute:

(i) $z+w$, $z-w$.

(ii) zw, $\bar{z}w$, $z\bar{w}$, $\bar{z}\bar{w}$, $\overline{zw}$.

(iii) $|zw|$, $|z|$, $|w|$, $|z||w|$.

(iv) $\frac{z}{w}$, $\frac{w}{z}$.

We have:

(i) $z+w = 1+2i+3-4i = 4-2i$, $z-w = 1+2i-(3-4i) = -2+6i$.

(ii) $zw = (1+2i)(3-4i) = 3-4i+6i+8 = 11+2i$,

$\bar{z}w = (1-2i)(3-4i) = 3-4i-6i-8 = -5-10i$,

$z\bar{w} = (1+2i)(3+4i) = 3+4i+6i-8 = -5+10i$,

$\bar{z}\bar{w} = (1-2i)(3+4i) = 3+4i-6i+8 = 11-2i$,

$\overline{zw} = \overline{11+2i} = 11-2i$.

(iii) $|zw| = |11+2i| = \sqrt{11^2+2^2} = \sqrt{125} = 5\sqrt{5}$,

$|z| = |1+2i| = \sqrt{1^2+2^2} = \sqrt{5}$,

$$|w| = |3-4i| = \sqrt{3^2+4^2} = \sqrt{25} = 5,$$

$$|z||w| = \sqrt{5}\cdot 5 = 5\sqrt{5}.$$

(iv) $$\frac{z}{w} = \frac{1+2i}{3-4i} = \frac{1+2i}{3-4i}\cdot\frac{3+4i}{3+4i} = \frac{-5+10i}{25} = \frac{-1}{5}+\frac{2}{5}i,$$

$$\frac{w}{z} = \frac{3-4i}{1+2i} = \frac{3-4i}{1+2i}\cdot\frac{1-2i}{1-2i} = \frac{-5-10i}{5} = -1-2i.$$

The conceptual reason for using complex numbers is that every polynomial can be completely factored. More specifically, if $f(x) = a_0 + a_1x + \cdots + a_nx^n$ is a polynomial of degree n then it can be written as $f(x) = (x-z_1)\cdots(x-z_n)$ where $z_1,\dots,z_n$ are complex numbers.

<u>Example:</u> Factor the polynomial $f(x) = x^3 - \pi x^2 + x - \pi$. It has $x=\pi$ as a root since

$f(\pi) = \pi^3 - \pi\pi^2 + \pi - \pi = 0$, so $(x-\pi)$ divides $f(x)$. The long division:

$$\begin{array}{l|l} x^3 - \pi x^2 + x - \pi & x-\pi \\ \hline -(x^3-\pi x^2) & x^2+1 \\ \hline 0+0+x-\pi & \\ \quad -(x-\pi) & \\ \hline \quad 0+0 & \end{array}$$

gives $x^3 - \pi x^2 + x - \pi = (x-\pi)(x^2+1)$.

The roots of x^2+1 are: $$x = \frac{-0 \pm \sqrt{0-4\cdot1\cdot1}}{2\cdot1} = \pm\frac{\sqrt{-4}}{2} = \pm\sqrt{-1} = \pm i.$$

Indeed $f(i) = i^3 - \pi(i^2) + i - \pi = -i + \pi + i - \pi = 0$ and $f(-i) = (-i)^3 - \pi(-i)^2 + (-i) - \pi =$

$= (-1)\cdot(-i) - \pi\cdot 1\cdot(-1) - i - \pi = 0$, so $(x-i)$ and $(x+i)$ divide x^2+1. Multiplying

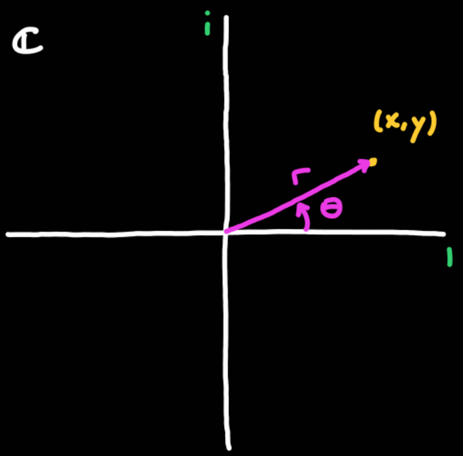
$(x-i)(x+i) = x^2 - i^2 = x^2+1$, so $f(x) = (x-i)(x+i)(x-\pi)$.

This is an example of a polynomial with real coefficients but only one real solution.

The polynomial $g(x) = x^3 - x^2 + x - 1$ has rational coefficients but only one rational solution. Moreover, both $f(x)$ and $g(x)$ have complex coefficients, and they have three complex solutions.

Since complex numbers are scalars, a matrix can have complex numbers in its entries.

Another geometric interpretation of the complex numbers comes from polar coordinates:



we can express (x,y) as (r,θ) where $r = \sqrt{x^2+y^2}$ and $\theta = \arctan\left(\frac{y}{x}\right)$, and we can express (r,θ) as (x,y) where $x = r\cos\theta$ and $y = r\sin\theta$.

This gives the following algebraic expression:

$$z = x+iy = r\cos\theta + ir\sin\theta = r(\cos\theta + i\sin\theta) = r\cdot e^{i\theta}$$

where in the last step we used <u>Euler's identity</u>:

$$e^{i\theta} = \cos\theta + i\sin\theta.$$

A rigorous proof of Euler's identity is beyond the scope of this course, but it can be achieved with any of the multiple ways of defining the exponential function for complex numbers. Formally, the function e^z with source $\mathbb{C}$ and target $\mathbb{C}$ satisfies:

(i) $\frac{d}{dz}(e^z) = e^z$ and $e^0 = 1$.

(ii) $e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}$, an absolutely convergent series.

(iii) $e^z = \lim_{n\to\infty} \left(1 + \frac{z}{n}\right)^n$.

It also transforms sums into products, like the real exponential function:

$e^{z+w} = e^z e^w$ for all z and w in $\mathbb{C}$.

<u>Example:</u> Write $\cos\theta$ and $\sin\theta$ in terms of exponential functions.

We know that if $z = x+iy$ then $x = \frac{z+\bar{z}}{2}$ and $y = \frac{z-\bar{z}}{2i}$. Setting $z = e^{i\theta}$, knowing $e^{i\theta} = \cos\theta + i\sin\theta$ and $\overline{e^{i\theta}} = \cos\theta - i\sin\theta = \cos(-\theta) + i\sin(-\theta) = e^{-i\theta}$ gives:

$$\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2} \quad \text{and} \quad \sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}.$$

<u>Example:</u> Convert the following numbers from $x+iy$ to $re^{i\theta}$.

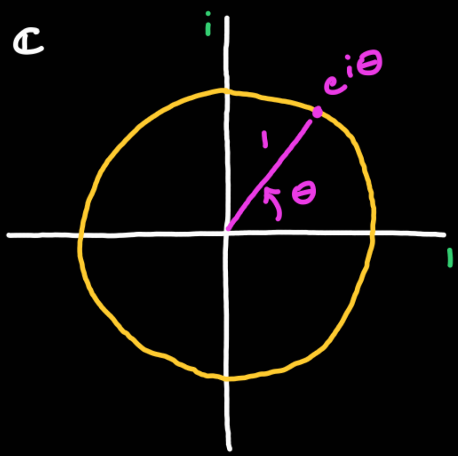
(i) $z = 1+\sqrt{3}i = \left(\sqrt{1^2+3}\right)e^{i\arctan\left(\frac{\sqrt{3}}{1}\right)} = 2e^{i\frac{\pi}{3}}$.

(ii) $z = 3 = 3 \cdot e^{i0} = 3$.

(iii) $z = -4i = -4e^{i\frac{\pi}{2}} = 4 \cdot (-1) \cdot e^{i\frac{\pi}{2}} = 4e^{i\pi}e^{i\frac{\pi}{2}} = 4e^{i\frac{3\pi}{2}}$.

Example: Draw all complex numbers of the form $e^{i\theta}$.

Noticing $|e^{i\theta}| = \sqrt{e^{i\theta}\,\overline{e^{i\theta}}} = \sqrt{e^{i\theta}e^{-i\theta}} = \sqrt{e^{i\theta - i\theta}} = \sqrt{e^0} = \sqrt{1} = 1$, we have that all these numbers are in the unit circle of the complex plane.



To obtain each point in the unit circle exactly once, we can ask that $0 \leq \theta < 2\pi$.

Each of these geometric interpretations of the complex numbers has an advantage. The Cartesian form is easier when computing sums, while the polar form is easier when computing products:

$$re^{i\theta}se^{i\phi} = (rs)e^{i(\theta+\phi)}.$$

Example: Compute the following.

(i) $\sqrt{-1}$. This is solving $x^2 = -1$.

(ii) $\sqrt{1}$. This is solving $x^2 = 1$.

(iii) $\sqrt[3]{1}$. This is solving $x^3 = 1$.

(iv) $\sqrt[5]{2}$. This is solving $x^5 = 2$.

(v) $\sqrt[5]{2e^{i\frac{\pi}{3}}}$. This is solving $x^5 = 2e^{i\frac{\pi}{3}}$.

Proceeding one by one, we have:

(i) $-1 = e^{-i\pi} = e^{i(-\pi + 2\pi n)}$ for all integers n, since $e^{2\pi n} = 1$.

Then $\sqrt{-1} = (-1)^{\frac{1}{2}} = \left(e^{i(2\pi n - \pi)}\right)^{\frac{1}{2}} = e^{i\left(\frac{2\pi n}{2} - \frac{\pi}{2}\right)} = e^{i\pi(n-\frac{1}{2})}$. The angle

$\pi(n-\frac{1}{2})$ must be between 0 and 2π, and n must be an integer,

so both $n=1$ and $n=2$ are valid. These give the angles $\theta = \frac{\pi}{2}$ and

$\theta = \frac{3\pi}{2}$, which correspond to:

$$e^{i\frac{\pi}{2}} = i \quad \text{and} \quad e^{i\frac{3\pi}{2}} = e^{i(\pi + \frac{\pi}{2})} = e^{i\pi} e^{i\frac{\pi}{2}} = -i.$$

Indeed $i^2 = -1$ and $(-i)^2 = -1$, so $\sqrt{-1}$ has two values.

In particular, taking square roots of complex numbers is <u>not</u> a function.

(ii) $1 = e^{i0} = e^{i2\pi n}$ for all integers n,

$$\sqrt{1} = (1)^{\frac{1}{2}} = \left(e^{i2\pi n}\right)^{\frac{1}{2}} = e^{i\pi n},$$

and $0 \leq \pi n < 2\pi$ when $n=0$ and $n=1$. In those cases we obtain $e^{i0}=1$ and $e^{i\pi}=-1$.

(iii) $\sqrt[3]{1} = \left(e^{i2\pi n}\right)^{\frac{1}{3}} = e^{i\frac{2}{3}\pi n}$, so $n=0,1,2$ are valid, giving

$e^{i0}=1,\quad e^{i\frac{2\pi}{3}},\quad e^{i\frac{4\pi}{3}}$.

(iv) $\sqrt[5]{2} = \left(\underbrace{2}_{\text{as a real number}}e^{i2\pi n}\right)^{\frac{1}{5}} = \underbrace{2^{\frac{1}{5}}}\cdot e^{i\frac{2\pi}{5}n}$, so $n=0,1,2,3,4$ are the integers for which $0 \leq \frac{2\pi}{5}n < 2\pi$. These give $2^{\frac{1}{5}}, 2^{\frac{1}{5}}e^{i\frac{2\pi}{5}}$,

$2^{\frac{1}{5}}e^{i\frac{4\pi}{5}},\quad 2^{\frac{1}{5}}e^{i\frac{6\pi}{5}},\quad 2^{\frac{1}{5}}e^{i\frac{8\pi}{5}}$.

(v) $\sqrt[5]{\underbrace{2}_{\text{as a real number}}e^{i\frac{\pi}{3}}} = 2^{\frac{1}{5}}\cdot\left(e^{i\frac{\pi}{3}}\right)^{\frac{1}{5}} = 2^{\frac{1}{5}}\left(e^{i\left(\frac{\pi}{3}+2\pi n\right)}\right)^{\frac{1}{5}} = 2^{\frac{1}{5}}e^{i\pi\frac{(1+6n)}{15}}$

so $n=0,1,2,3,4$ are the integers for which $0 \leq \frac{\pi}{15}(1+6n) < 2\pi$,

giving $2^{\frac{1}{5}},\ 2^{\frac{1}{5}}e^{i\frac{7\pi}{15}},\ 2^{\frac{1}{5}}e^{i\frac{13\pi}{15}},\ 2^{\frac{1}{5}}e^{i\frac{19\pi}{15}},\ 2^{\frac{1}{5}}e^{i\frac{25\pi}{15}}$.

Combining the properties of the exponential and Euler's identity, we can remember many trigonometric identities. For example:

$$\cos(\alpha+\beta) + i\sin(\alpha+\beta) = e^{i(\alpha+\beta)} = e^{i\alpha}e^{i\beta} = (\cos\alpha + i\sin\alpha)(\cos\beta + i\sin\beta) =$$

$$= \left(\cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)\right) + i\left(\cos(\alpha)\sin(\beta) + \sin(\alpha)\cos(\beta)\right)$$

gives:

$$\cos(\alpha+\beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta), \text{ and}$$

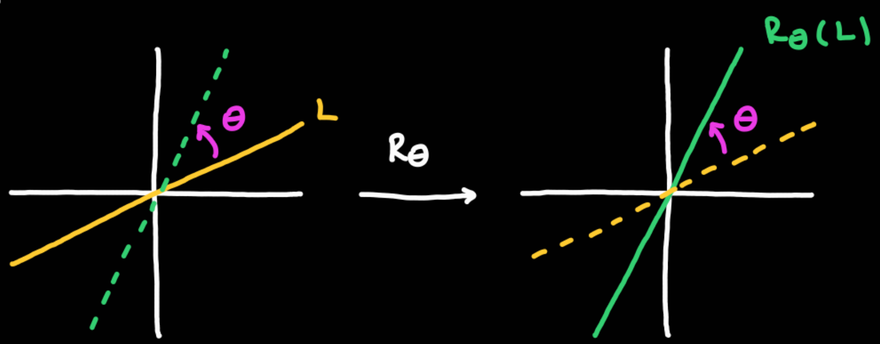
$$\sin(\alpha+\beta) = \cos(\alpha)\sin(\beta) + \sin(\alpha)\cos(\beta).$$

6. Eigen-analysis

6.1. Eigenvalues and eigenvectors

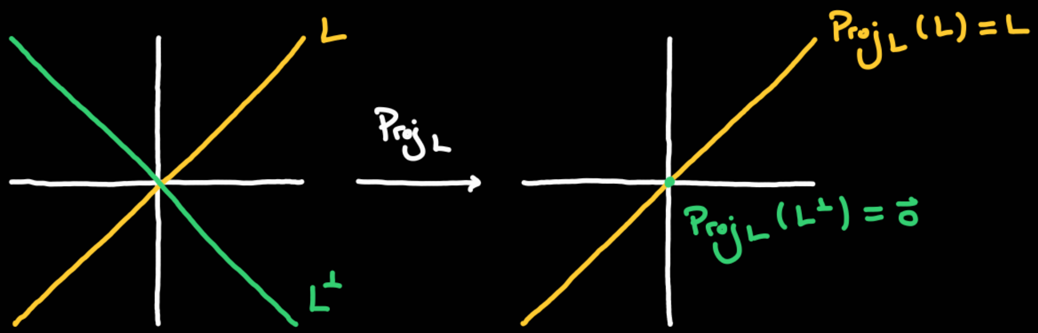
A square matrix A of size $n \times n$ can be seen as a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$. Geometrically, this may mangle a nice geometric figure F in $\mathbb{R}^n$ into some complicated figure $A(F)$. However, sometimes the figure does not change, namely $A(F) = F$. This is an intrinsic property of the matrix A, for a given A these figures may or may not exist. Since the easiest geometric figure is a line through the origin, let's see a few examples that preserve and do not preserve them.

Example: Rotation by an angle of θ counterclockwise does not preserve any line, except when $\theta = 0$.

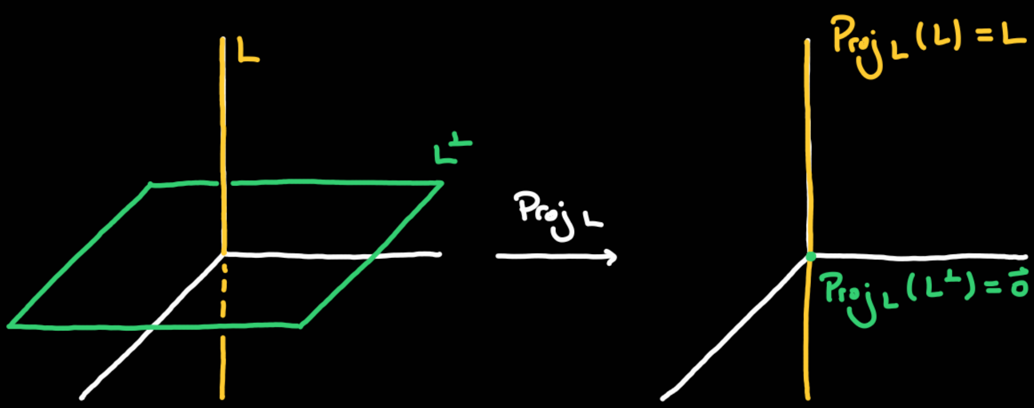


Example: Projection onto a line L preserves the line itself. Moreover, although all lines perpendicular to it are *not* preserved, their image remains perpendicular to L.

In two dimensions we have:

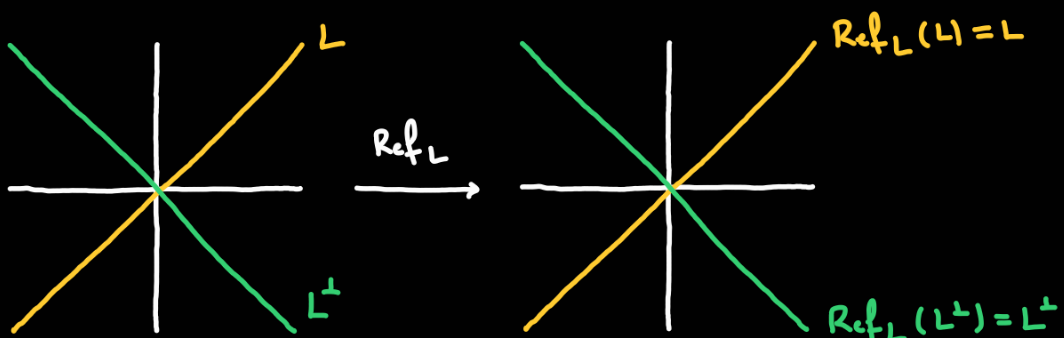


where $\vec{0}$ is in $L^\perp$. In three dimensions we have:

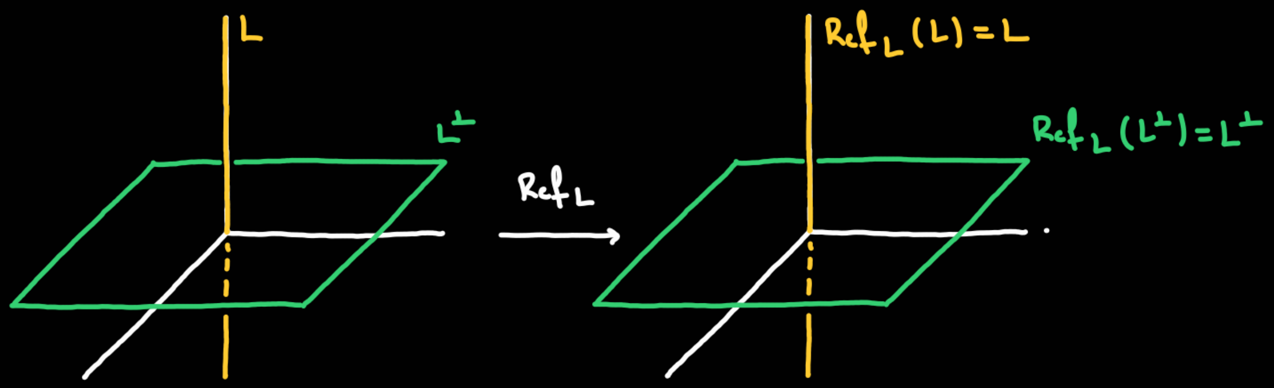


where $\vec{0}$ is in $L^\perp$.

Example: Reflection across a line L preserves both the line itself and the lines perpendicular to it. In two dimensions we have:



and in three dimensions we have:



The rigorous notions capturing this behavior are eigenvectors and their corresponding eigenvalues. An $n \times n$ matrix A has an <u>eigenvector</u> $\vec{v}$ when $\vec{v} \neq \vec{0}$ and $A\vec{v} = \lambda \vec{v}$ for some scalar λ, which is called the <u>eigenvalue</u> of $\vec{v}$. Although by definition an eigenvector $\vec{v}$ is <u>not</u> $\vec{0}$, it may have an eigenvalue $\lambda = 0$.

<u>Example:</u> The matrix $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ has two eigenvectors: $\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ with eigenvalue $\lambda_1 = 3$ and $\vec{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ with eigenvalue $\lambda_2 = -1$.

<u>Example:</u> The matrix $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ has two eigenvectors: $\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ with eigenvalue $\lambda_1 = 2$ and $\vec{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ with eigenvalue $\lambda_2 = 0$.

<u>Example:</u> The matrix $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ has only one eigenvector: $\vec{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ with eigenvalue $\lambda = 1$.

<u>Example:</u> The matrix $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ has <u>no</u> real eigenvectors and <u>no</u> real eigenvalues.

It has two complex eigenvectors: $\vec{v}_1 = \begin{bmatrix} i \\ 1 \end{bmatrix}$ with eigenvalue $\lambda_1 = i$ and

$\vec{v}_2 = \begin{bmatrix} -i \\ 1 \end{bmatrix}$ with eigenvalue $\lambda_2 = -i$.

Conceptually, the eigenvectors of a matrix capture "preferential directions" of the matrix. That is, directions that are invariant when multiplied by the matrix on the left, because the image of a vector in the direction of an eigenvector remains in the direction of that same eigenvector.

Example: Let $\vec{v}$ be an eigenvector of A with eigenvalue λ. Then $\vec{v}$ defines the line $\vec{x} = s\vec{v}$ for s scalars, and $A\vec{x} = As\vec{v} = sA\vec{v} = s\lambda\vec{v} = \lambda s\vec{v} = \lambda\vec{x}$ means that all vectors in that line are eigenvectors of eigenvalue λ.