Complex numbers

$$x^2 + 1 = 0 \Rightarrow \quad x = \pm\sqrt{-1} = \pm i$$

Fundamental thm of algebra: Every real or complex polynomial of degree "n" has "n" roots (can be complex AND repeated)

Example: $x^4 - 1 = 0$ has 4 roots $\Rightarrow$

$$x = +1, -1, +i, -i$$

Euler's formula: $\cos(\theta) + i\sin(\theta) = e^{i\theta}$

Proof: (Taylor expansion)

$$\begin{aligned} e^{i\theta} = 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + ... &= 1 - \frac{\theta^2}{2} + ... + i(\theta - \frac{\theta^3}{6} + ...) \\ &= \cos(\theta) + i\sin(\theta) \end{aligned}$$

This means that $e^{i\theta}e^{-i\theta} = 1$

Proof:

$$\begin{aligned} (\cos(\theta) + i\sin(\theta))(\cos(\theta) - i\sin(\theta)) &= \cos^2\theta + \sin^2\theta + i\sin\theta\cos\theta - i\sin\theta\cos\theta \\ &= 1 \end{aligned}$$

Roots of unity: an nth root of unity $z^n = 1$

$$\exp\left(\frac{2k\pi i}{n}\right) = \cos\frac{2k\pi}{n} + i\sin\frac{2k\pi}{n}, \qquad k = 0, 1, \ldots, n-1$$

Euler's method

Differential equation governs the rate of change of a variable.

$$\frac{dx}{dt} = -x$$

This example is exponential decay.

If we know $x(t_0)$ we can compute x shortly after $(x(t_0 + \Delta t))$ with an approximation:

$$x(t_0 + \Delta t) = x(t_0) + \Delta t \frac{dx(t_0)}{dt}$$

Example:
Find $x(0.1)$ given $x(0) = 5$ using Euler's method:

$$\frac{dx(t)}{dt} = -x$$

Eigenvalues and eigenvectors

Example:

$$A = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$$

Find vectors that stay on their own span, e.g. $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

$$\begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Example: Consider a 3D rotation, the eigenvector of the rotation is the AXIS OF ROTATION with eigenvalue $\lambda = 1$

$$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ is an eigenvector with eigenvalue 1.

$A\vec{v} = \lambda\vec{v}$, λ: eigenvalue, $\vec{v}$: eigenvector

$$\begin{aligned} A\vec{v} &= \lambda I \vec{v} \\ A\vec{v} - \lambda I \vec{v} &= \vec{0} \\ (A - \lambda I)\vec{v} &= \vec{0} \end{aligned}$$

Trivial solution $\vec{v} = \vec{0}$. Only other way to get zero:

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

Example: Find eigenvalues of $A = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$

$$\begin{aligned} \det\left(\begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = \det\left(\begin{bmatrix} 3-\lambda & 1 \\ 0 & 2-\lambda \end{bmatrix}\right) &= 0 \\ \Rightarrow (3-\lambda)(2-\lambda) &= 0 \\ \Rightarrow \lambda &= 3, 2 \end{aligned}$$

Example: Find eigenvalues of $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

$$\det\left(\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = \det\left(\begin{bmatrix} -\lambda & -1 \\ 1 & -\lambda \end{bmatrix}\right) = 0$$
$$\Rightarrow \lambda^2 + 1 = 0$$
$$\Rightarrow \lambda = +i, -i$$

All vectors in the REAL plane are rotated $\Rightarrow$ no REAL vectors that stay on their own span.

Example: Find eigenvalues of $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

$$\det\left(\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = \det\left(\begin{bmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{bmatrix}\right) = 0$$
$$\Rightarrow (1-\lambda)(1-\lambda) = 0$$
$$\Rightarrow \lambda = 1$$

Only ONE eigenvalue/eigenvector. Find the eigenvector:

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = 1 \begin{bmatrix} a \\ b \end{bmatrix} \Rightarrow a + b = a \Rightarrow b = 0$$

The eigenvector is $\begin{bmatrix} a \\ 0 \end{bmatrix}$ where $a \in \mathbb{R}$.

Example: Find eigenvalues of $A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$.

$$\det\left(\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = \det\left(\begin{bmatrix} 2-\lambda & 0 \\ 0 & 2-\lambda \end{bmatrix}\right) = 0$$
$$\Rightarrow (2-\lambda)(2-\lambda) = 0$$
$$\Rightarrow \lambda = 2$$

Find the eigenvectors:

$$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = 2 \begin{bmatrix} a \\ b \end{bmatrix} \Rightarrow 2a = 2a; 2b = 2b \Rightarrow a \in \mathbb{R}, b \in \mathbb{R}$$

All vectors are eigenvectors of diagonal matrices!

Uses of eigenvalues/eigenvectors

- dynamical systems - governs timescales
- low-dimensional representations

Differential Equations

systems of equations governing dynamics

Example: Exponential decay
x = firing rate of a neuron, τ = timescale of neuron

$$\begin{aligned}
\frac{dx}{dt} &= -x/\tau \qquad \text{(firing rate decays to zero)} \\
\Rightarrow \int \frac{dx}{x} &= \int \frac{-dt}{\tau} \\
\Rightarrow \ln(x) &= -t/\tau + c \\
\Rightarrow x(t) &= e^{-t/\tau + c} = e^{-t/\tau} e^{c} \\
x(t) &= c e^{-t/\tau} \quad \text{where} \quad c = x(0)
\end{aligned}$$

* neuron's firing rate decays with timescale τ *

Example: Add another neuron as input:

$$\begin{aligned}
\frac{dx}{dt} &= -x + 2y \\
\frac{dy}{dt} &= -y + 2x
\end{aligned}$$

We can rewrite this as a matrix multiplication:

$$\begin{bmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{bmatrix} = A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Suppose the solution takes the form $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} e^{\lambda t}$, what are λ and $\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$?

$$\text{LHS:} \quad \begin{bmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{bmatrix} = \begin{bmatrix} \frac{d}{dt}\left(a_1 e^{\lambda t}\right) \\ \frac{d}{dt}\left(a_2 e^{\lambda t}\right) \end{bmatrix} = \begin{bmatrix} a_1(\lambda e^{\lambda t}) \\ a_2(\lambda e^{\lambda t}) \end{bmatrix} = \lambda \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} e^{\lambda t}$$

Plug this into the equation:

$$\begin{aligned}
\lambda \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} e^{\lambda t} &= \begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} e^{\lambda t} \\
\Rightarrow \lambda \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} &= \begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}
\end{aligned}$$

Let $\vec{v} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ and $A = \begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix}$

$$\Rightarrow \lambda \vec{v} = A \vec{v}$$

What are λ and $\vec{v} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$?

Find the eigenvalues:

$$\det\left(\begin{bmatrix}-1 & 2\\ 2 & -1\end{bmatrix} - \lambda\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\right) = \det\left(\begin{bmatrix}-1-\lambda & 2\\ 2 & -1-\lambda\end{bmatrix}\right) = 0$$
$$\Rightarrow (-1-\lambda)(-1-\lambda) - 4 = 0$$
$$\Rightarrow \lambda^2 + 2\lambda - 3 = 0$$
$$\Rightarrow (\lambda+3)(\lambda-1) = 0$$
$$\Rightarrow \lambda = -3, +1$$

Find the eigenvectors:

$$\begin{bmatrix}-1 & 2\\ 2 & -1\end{bmatrix}\begin{bmatrix}a_1\\ a_2\end{bmatrix} = -3\begin{bmatrix}a_1\\ a_2\end{bmatrix} \Rightarrow -a_1 + 2a_2 = -3a_1 \Rightarrow a_2 = -a_1 \Rightarrow \begin{bmatrix}a_1\\ a_2\end{bmatrix} = \begin{bmatrix}1\\ -1\end{bmatrix}$$
$$\begin{bmatrix}-1 & 2\\ 2 & -1\end{bmatrix}\begin{bmatrix}a_1\\ a_2\end{bmatrix} = +1\begin{bmatrix}a_1\\ a_2\end{bmatrix} \Rightarrow -a_1 + 2a_2 = a_1 \Rightarrow a_2 = a_1 \Rightarrow \begin{bmatrix}a_1\\ a_2\end{bmatrix} = \begin{bmatrix}1\\ 1\end{bmatrix}$$

On your own check that these are eigenvectors.

Two solutions of the differential equation:

$$\begin{bmatrix}x\\ y\end{bmatrix} = c_1 e^{-3t}\begin{bmatrix}1\\ -1\end{bmatrix}, \quad \begin{bmatrix}x\\ y\end{bmatrix} = c_2 e^{t}\begin{bmatrix}1\\ 1\end{bmatrix}$$

Any **linear combinations** are also solutions, let's check this. Let

$$\begin{bmatrix}x\\ y\end{bmatrix} = c_1 e^{\lambda_1 t}\vec{v}_1 + c_2 e^{\lambda_2 t}\vec{v}_2$$

$$\text{LHS:} \quad \begin{bmatrix}\frac{dx}{dt}\\ \frac{dy}{dt}\end{bmatrix} = c_1(\lambda_1 e^{\lambda_1 t})\vec{v}_1 + c_2(\lambda_2 e^{\lambda_2 t})\vec{v}_2 = c_1 e^{\lambda_1 t}(\lambda_1\vec{v}_1) + c_2 e^{\lambda_2 t}(\lambda_2\vec{v}_2)$$
$$\text{RHS:} \quad A(c_1 e^{\lambda_1 t}\vec{v}_1 + c_2 e^{\lambda_1 t}\vec{v}_2) = c_1 e^{\lambda_1 t}(A\vec{v}_1) + c_2 e^{\lambda_2 t}(A\vec{v}_2) = c_1 e^{\lambda_1 t}(\lambda_1\vec{v}_1) + c_2 e^{\lambda_2 t}(\lambda_2\vec{v}_2)\checkmark$$

Can also do it the looong way. Let

$$\begin{bmatrix}x\\ y\end{bmatrix} = c_1 e^{-3t}\begin{bmatrix}1\\ -1\end{bmatrix} + c_2 e^{t}\begin{bmatrix}1\\ 1\end{bmatrix}$$

$$\text{LHS:} \quad \begin{bmatrix}\frac{dx}{dt}\\ \frac{dy}{dt}\end{bmatrix} = \begin{bmatrix}c_1(-3)e^{\lambda t} + c_2(1)e^{\lambda t}\\ c_1(-3)(-1)e^{\lambda t} + c_2(1)e^{\lambda t}\end{bmatrix} = \begin{bmatrix}-3c_1 e^{\lambda t} + c_2 e^{\lambda t}\\ 3c_1 e^{\lambda t} + c_2 e^{\lambda t}\end{bmatrix}$$
$$\text{RHS:} \quad \begin{bmatrix}-1 & 2\\ 2 & -1\end{bmatrix}\left(c_1\begin{bmatrix}1\\ -1\end{bmatrix}e^{-3t} + c_2\begin{bmatrix}1\\ 1\end{bmatrix}e^{t}\right) = \begin{bmatrix}-3c_1 e^{-3t} + c_2 e^{t}\\ 3c_1 e^{-3t} + c_2 e^{t}\end{bmatrix}\checkmark$$

Theorem: If you start on an eigenvector, you STAY on an eigenvector.
"Proof": (using Euler's method)

$$\begin{bmatrix} x(t+\Delta t) \\ y(t+\Delta t) \end{bmatrix} \approx \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} + \Delta t \begin{bmatrix} \frac{dx(t)}{dt} \\ \frac{dy(t)}{dt} \end{bmatrix}$$

If $\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \vec{v}$ (eigenvector of A), then $\begin{bmatrix} \frac{dx(t)}{dt} \\ \frac{dy(t)}{dt} \end{bmatrix} = A\vec{v} = \lambda \vec{v} \Rightarrow$

$$\begin{bmatrix} x(t+\Delta t) \\ y(t+\Delta t) \end{bmatrix} \approx \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} + \Delta t \ \lambda \begin{bmatrix} x(t) y(t) \end{bmatrix} = (1 + \Delta t \lambda)\vec{v}$$

still on eigenvector

What does $\lambda > 0$ versus $\lambda < 0$ mean?
* note we can think about these neurons as "groups of neurons" *

How can we make the system stable? (don't want $x \to \infty, y \to \infty$)

$$\begin{bmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{bmatrix} = \begin{bmatrix} -3 & 2 \\ 2 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \left(\begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix} - 2I \right) \begin{bmatrix} x \\ y \end{bmatrix}$$

The neurons will decay faster, why would this make this stable?

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

Now subtract $2I$ and find new eigenvalues

$$\begin{aligned} \det(A - 2I - \lambda_{\text{new}} I) = \det(A - (2 + \lambda_{\text{new}})I) &= 0 \\ \Rightarrow 2 + \lambda_{\text{new}} &= \lambda \\ \Rightarrow \lambda_{\text{new}} &= \lambda - 2 \end{aligned}$$

Theorem: Eigenvalues of $A + bI$ are $\lambda + b$ where λ are eigenvalues of A and eigenvectors are the same as the eigenvectors of A.
Proof:

$$(A + bI)\vec{v} = (\lambda + b)\vec{v}$$
$$A\vec{v} + bI\vec{v} = \lambda\vec{v} + b\vec{v}$$

$\Rightarrow \vec{v}$ is also an eigenvector of $A + bI$

New system $\begin{bmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{bmatrix} = \begin{bmatrix} -3 & 2 \\ 2 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$ has eigenvalues $\lambda = -5, -1$ and eigenvectors $\vec{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

How else can we prevent neurons from $\to \infty$?
add inhibitory neurons!

Example: Inhibitory neuron

$$\begin{bmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Which neuron in this system is the "excitatory neuron" and which is the "inhibitory neuron"?

Find eigenvalues (recall from video):

$$\begin{aligned} \text{det}\left(\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) = \text{det}\left(\begin{bmatrix} -\lambda & -1 \\ 1 & -\lambda \end{bmatrix} \right) &= 0 \\ \Rightarrow \lambda^2 + 1 &= 0 \\ \Rightarrow \lambda &= +i, -i \end{aligned}$$

Therefore, x and y are functions of $e^{it} = \cos(t) i \sin(t)$ $\Rightarrow$ OSCILLATIONS!

Make phase diagram and show oscillation.

See what happens when $\lambda = -1 \pm i$.

Diagonalizing a matrix

Let A be a matrix, with λ_1, λ_2 eigenvalues and $\vec{x}_1, \vec{x}_2$ eigenvectors. Let's multiply A with its eigenvectors:

$$A\begin{bmatrix}\vec{x}_1 & \vec{x}_2\end{bmatrix} = \begin{bmatrix}A\vec{x}_1 & A\vec{x}_2\end{bmatrix} = \begin{bmatrix}\lambda_1\vec{x}_1 & \lambda_2\vec{x}_2\end{bmatrix}.$$

We can rewrite the RHS side as a matrix multiplication:

$$\begin{bmatrix}\lambda_1\vec{x}_1 & \lambda_2\vec{x}_2\end{bmatrix} = \begin{bmatrix}\vec{x}_1 & \vec{x}_2\end{bmatrix}\begin{bmatrix}\lambda_1 & 0 \\ 0 & \lambda_2\end{bmatrix} = V\Lambda$$

where we term the eigenvector matrix V and the diagonal matrix with the eigenvalues Λ. Now let's rewrite the first expression and try to **diagonalize** A:

$$AV = \Lambda V$$

Multiply by V^{-1} on both sides.

$$V^{-1}AV = V^{-1}V\Lambda = \Lambda$$

V diagonalizes A.

Can decompose A into V e'vectors and Λ e'values:

$$A = V\Lambda V^{-1}$$

This also makes it easy to compute powers of A:

$$A^2 = (V\Lambda V^{-1})(V\Lambda V^{-1}) = V\Lambda^2 V^{-1}$$

$$\Rightarrow A^n = V\Lambda^n V^{-1}$$

As $n \to \infty$, $\lambda > 1$ will dominate and therefore the transformation will tend towards its corresponding eigenvector.

Similar matrices

Let B be similar to A: $B = M^{-1}AM$ and $B\vec{x} = \lambda\vec{x}$. What are the eigenvalues and eigenvectors of A?

$$B\vec{x} = M^{-1}AM\vec{x} = \lambda\vec{x}$$

Multiply both sides by M

$$M(M^{-1}AM)\vec{x} = M(\lambda\vec{x})$$
$$\Rightarrow A(M\vec{x}) = \lambda(M\vec{x})$$

Same eigenvalues, eigenvectors transformed by M.
* any matrix is similar to the diagonal matrix (through eigenvectors): $\Lambda = V^{-1}AV$ *

Symmetric matrices

Definition: A is **symmetric** if $A = A^\top$.

Example: Recall from last class, $A = \begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix}$ has $\lambda = -3, +1$ and eigenvectors $\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

These eigenvalues are **real** numbers. These eigenvectors are also **orthogonal**. How do we check orthogonality?

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix}^\top \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1 - 1 = 0 \ \checkmark$$

Orthonormal matrices are orthogonal matrices with columns with unit norm - how do we make $V = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 \end{bmatrix}$ have columns of unit norm?

$$V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

Prove for yourself that this works.

Theorem: Matrix V orthonormal $\iff V^{-1} = V^\top$.
Proof: (right-direction) We will show for a two column matrix, but applies to an N-D matrix:

$$V = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 \end{bmatrix} \quad \text{and} \quad V^\top = \begin{bmatrix} \vec{v}_1 \\ \vec{v}_2 \end{bmatrix}$$

$$\Rightarrow V^\top V = \begin{bmatrix} \vec{v}_1^\top \\ \vec{v}_2^\top \end{bmatrix} \begin{bmatrix} \vec{v}_1 & \vec{v}_2 \end{bmatrix} = \begin{bmatrix} \vec{v}_1^\top \vec{v}_1 & \vec{v}_1^\top \vec{v}_2 \\ \vec{v}_2^\top \vec{v}_1 & \vec{v}_2^\top \vec{v}_2 \end{bmatrix}$$

Since $\vec{v}_1$ and $\vec{v}_2$ are orthogonal, $\vec{v}_1^\top \vec{v}_2 = 0$. Since $\vec{v}_1$ and $\vec{v}_2$ are unit norm, $\vec{v}_1^\top \vec{v}_1 = 1$ and $\vec{v}_2^\top \vec{v}_2 = 1 \Rightarrow$

$$V^\top V = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \Rightarrow V^\top = V^{-1}$$

Check for $V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ that the inverse is the transpose.

> **Theorem:** Symmetric matrices ($A = A^\top$) have real eigenvalues and orthogonal eigenvectors.

Proof: (for orthogonal eigenvectors)
Rewrite A as $A = V^{-1} \Lambda V$

$$\Rightarrow A^\top = (V^{-1} \Lambda V)^\top = V^\top \Lambda (V^{-1})^\top$$

By definition,

$$A = A^\top \Rightarrow V^{-1} \Lambda V = V^\top \Lambda (V^{-1})^\top$$

How do we achieve equality for this expression? $V^{-1} = V^\top$
Thus, V is an orthonormal matrix $\Rightarrow$ eigenvectors are orthogonal.

Positive semi-definite matrices

Definition: S is **positive semi-definite** if $S = A^\top A$.

Theorem: Positive semi-definite matrices are symmetric.
Proof:

$$S = A^\top A \Rightarrow S^\top = (A^\top A)^\top = A^\top (A^\top)^\top = A^\top A$$
$$\Rightarrow S = S^\top$$

> **Theorem:** Positive semi-definite matrices have all eigenvalues $\lambda \geq 0$

Proof: Let λ, $\vec{v}$ be eigenvalues and eigenvectors of S

$$\begin{aligned}
A^\top A\vec{v} &= \lambda\vec{v} \\
\vec{v}^\top(A^\top A\vec{v}) &= \vec{v}^\top(\lambda\vec{v}) \\
\vec{v}^\top(A^\top A\vec{v}) &= \vec{v}^\top(\lambda\vec{v}) \\
(A\vec{v})^\top(A\vec{v}) &= \lambda\vec{v}^\top\vec{v} \\
||A\vec{v}||^2 &= \lambda||\vec{v}||^2 \quad \text{norm is always positive} \\
\Rightarrow \lambda &\geq 0 \quad \forall S = A^\top A
\end{aligned}$$

When will $\lambda = 0$?
* if S does not have independent columns (determinant = 0) *

Example: $S = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, find eigenvalues λ_1, λ_2 and eigenvectors $\vec{v}_1, \vec{v}_2$.

$$\begin{aligned}
\det\left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} - \lambda\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = \det\left(\begin{bmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{bmatrix}\right) &= 0 \\
\Rightarrow \lambda^2 - 2\lambda + 1 &= 0 \\
\Rightarrow \lambda(\lambda - 2) &= 0 \\
\Rightarrow \lambda &= 0, 2
\end{aligned}$$

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = 2\begin{bmatrix} a \\ b \end{bmatrix} \Rightarrow a + b = 2a \Rightarrow b = a \Rightarrow \vec{v}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = 0\begin{bmatrix} a \\ b \end{bmatrix} \Rightarrow a + b = 0 \Rightarrow b = -a \Rightarrow \vec{v}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Rewrite S as diagonalization:

$$S = V^\top \Lambda V = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^\top \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

There is a full column of zeros $\Rightarrow \vec{v}_2$ doesn't matter. Can then write S as

$$S = \frac{1}{2}\begin{bmatrix}1\\1\end{bmatrix}^\top (2)\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}1\\1\end{bmatrix}^\top \begin{bmatrix}1\\1\end{bmatrix}$$

What is the rank of S?

Example: 2 neurons' firing rates in response to 6 different stimuli (normalized to zero average firing rate):

$$X = \begin{bmatrix} 5 & 7 \\ -4 & 2 \\ 8 & 7 \\ -10 & -9 \\ 1 & 0 \\ 0 & -3 \end{bmatrix}$$

How are these neurons' firing rates covarying?
cov$(x,y) = \frac{1}{N_{\text{stim}}}\Sigma_i(x_i - \bar{x})(y_i - \bar{y})$ where i is for different stimuli
$\Rightarrow$ covariance matrix $S = \frac{1}{N_{\text{stim}}}XX^\top = \begin{bmatrix}34 & 31\\31 & 32\end{bmatrix}$.

What are the eigenvalues and eigenvectors? S is positive semi-definite so $\lambda \geq 0$.

$$\lambda = 11, 0.4; \quad \vec{v}_1 = \begin{bmatrix}0.72\\0.69\end{bmatrix}, \vec{v}_2 = \begin{bmatrix}-0.69\\0.72\end{bmatrix}$$

What do you notice when you plot these vectors?

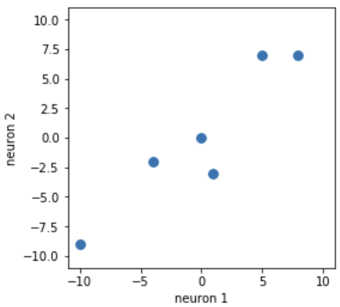
What is the projection of X onto $\vec{v}_1$ and $\vec{v}_2$?
Neuron 1 $\vec{x}_1$ onto $\vec{v}_1$:

$$\text{proj}_{\vec{v}_1}\vec{x}_1 = \frac{\vec{x}_1^\top \vec{v}_1}{||\vec{v}_1||^2}\vec{v}_1 = (\vec{x}_1^\top \vec{v}_1)\vec{v}_1$$

In [1]:

```
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline

n1 = np.array([5,-4,8,-10,1,0])
n2 = np.array([7,-2,7,-9,-3,0])

A = np.concatenate((n1[np.newaxis,:], n2[np.newaxis,:]), axis=0)
print(A.shape)

# plot neuron activity
fig = plt.figure(figsize=(4,4))
ax = fig.add_subplot(111)
ax.scatter(n1,n2,s=60)
ax.set_xlabel('neuron 1')
ax.set_ylabel('neuron 2')
ax.set_xlim(-11,11)
ax.set_ylim(-11,11)
plt.show()
```

```
(2, 6)
```



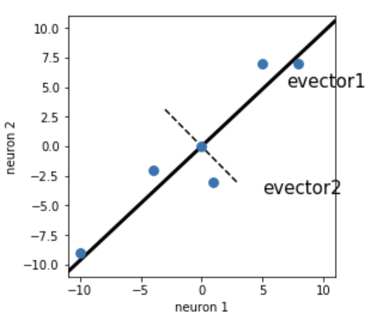
In [2]:

```
print('covariance matrix')
print(A @ A.T)

# find eigenvalues and eigenvectors of covariance matrix
lam, v = np.linalg.eig(A @ A.T)

print('eigenvalues: %2.0f, %2.0f'%(lam[0],lam[1]))
print('eigenvectors: [%2.2f,%2.2f], [%2.2f,%2.2f]'%(v[0,0],v[1,0],v[0,1],v[1,1]))
```

```
covariance matrix
[[206 186]
 [186 192]]
eigenvalues: 385, 13
eigenvectors: [0.72,0.69], [-0.69,0.72]
```

In [3]:

```
# plot EIGENVECTOR on top
fig = plt.figure(figsize=(4,4))
ax = fig.add_subplot(111)
ax.scatter(n1,n2,s=60)
ax.plot(np.array([-11,11]), np.array([-11,11])*v[1,0]/v[0,0],color='k', zorder=0, lw=3)
ax.text(7,5,'evector1',fontsize=15)
ax.plot(np.array([-3,3]), np.array([-3,3])*v[1,1]/v[0,1],'--',color='k', zorder=0)
ax.text(5,-4,'evector2',fontsize=15)
ax.set_xlabel('neuron 1')
ax.set_ylabel('neuron 2')
ax.set_xlim(-11,11)
ax.set_ylim(-11,11)
plt.show()
```



Principal components analysis

- most data is HIGH-dimensional, how do we visualize/understand it?
- PCA is a linear dimensionality reduction technique
- The first PC is a projection that captures the MOST variance in the data
- find a low-dimensional space that **preserves** as much variance in the original data as possible.
- can use low-D summary and it may be more interpretable
- can also use as a **pre-processing** step before doing classification or regression – a low-dimensional regression has FEWER parameters so it acts as a "regularization" step

PCA derivation: (maximum variance)

$$X = \begin{bmatrix} \vec{x}_1 & \dots & \vec{x}_n \end{bmatrix} \qquad N \text{ neurons } \vec{x}_i \in \mathbb{R}^D$$

If $\vec{x}_i$ are mean 0, then covariance $S = \frac{1}{N_{\text{stim}}} X X^\top$.
We want to find principal component $\vec{u}_1$ that maximizes variance of projection of data onto it.

$$\begin{aligned} \max_{\vec{u}_1} \ \text{var}_i(\vec{u}_1^\top \vec{x}_i) &= \frac{1}{N_{\text{stim}}} \Sigma_i (\vec{u}_1^\top \vec{x}_i)(\vec{u}_1^\top \vec{x}_i)^\top \\ &= \frac{1}{N_{\text{stim}}} \Sigma_i \vec{u}_1^\top (\vec{x}_i \vec{x}_i^\top) \vec{u}_1 \\ &= \vec{u}_1^\top S \vec{u}_1 \end{aligned}$$

if $||\vec{u}_1|| \to \infty$, then variance will $\to \infty$. We therefore need to **constrain** this optimization such that the norm of $\vec{u}_1 < \infty$. We choose $\vec{u}_1 = 1$. To do constrained optimization we use **Lagrange multipliers**:

$$\begin{aligned} \mathcal{L}(\vec{u}_1, \lambda) &= \vec{u}_1^\top S \vec{u}_1 - \lambda(\vec{u}_1^\top \vec{u}_1 - 1) \\ \frac{\partial}{\partial \vec{u}_1} \mathcal{L}(\vec{u}_1, \lambda) &= \frac{\partial}{\partial \vec{u}_1} \left(\vec{u}_1^\top S \vec{u}_1 - \lambda(\vec{u}_1^\top \vec{u}_1 - 1) \right) & \frac{\partial}{\partial \lambda} \mathcal{L}(\vec{u}_1, \lambda) &= \frac{\partial}{\partial \lambda} \left(\vec{u}_1^\top S \vec{u}_1 - \lambda(\vec{u}_1^\top \vec{u}_1 - 1) \right) \\ \frac{\partial}{\partial \vec{u}_1} \mathcal{L}(\vec{u}_1, \lambda) &= 2S\vec{u}_1 - 2\lambda \vec{u}_1 = 0 & \frac{\partial}{\partial \lambda} \mathcal{L}(\vec{u}_1, \lambda) &= \vec{u}_1^\top \vec{u}_1 - 1 = 0 \\ &\Rightarrow S\vec{u}_1 = \lambda \vec{u}_1 & &\Rightarrow ||\vec{u}_1||^2 = 1 \end{aligned}$$

Can you tell what $\vec{u}_1$ should be to satisfy this equation?

Let $\vec{u}_1$ be an eigenvector of the covariance matrix S, which eigenvector maximuizes the variance?

$$\begin{aligned} \max_{\vec{u}_1} \ \text{var}_i(\vec{u}_1^\top \vec{x}_i) &= \vec{u}_1^\top S \vec{u}_1 \text{ where } \vec{u}_1 \text{ is an eigenvector} \\ &= \vec{u}_1^\top (\lambda \vec{u}_1) \\ &= \lambda \vec{u}_1^\top \vec{u}_1 = \lambda ||\vec{u}_1||^2 = \lambda \end{aligned}$$

PCA: The eigenvector with the largest eigenvalue is the first principal component. The next principal components are the following eigenvectors.

PCA derivation: (minimize residuals)
Introduce D orthonormal basis vectors $\vec{u}_i$ such that $\vec{u}_i\vec{u}_j = \delta_{ij}$. We can represent each neuron $\vec{x}_n$ as

$$\vec{x}_n = \sum_{i=1}^{D}(\vec{x}_n^\top \vec{u}_i)\vec{u}_i$$

where each neuron is a sum of $\vec{u}_i$ with weights of the projection of $\vec{x}_n$ onto the vectors $\vec{u}_i$. How do we choose $\vec{u}_i$ to minimize the error of the reconstruction of the original data with only M vectors?

$$\hat{\vec{x}}_n = \sum_{i=1}^{M} z_{ni}\vec{u}_i + \sum_{i=M+1}^{D} b_i\vec{u}_i$$

b_i are the same for all neurons (to make an M dimensional representation). Minimize reconstruction error:

$$J = \frac{1}{N}\sum_{n=1}^{N}||\vec{x}_n - \hat{\vec{x}}_n||^2 = \frac{1}{N}\sum_{n=1}^{N}\vec{x}_n^\top\vec{x}_n - 2\hat{\vec{x}}_n\vec{x}_n + \hat{\vec{x}}_n^\top\hat{\vec{x}}_n$$

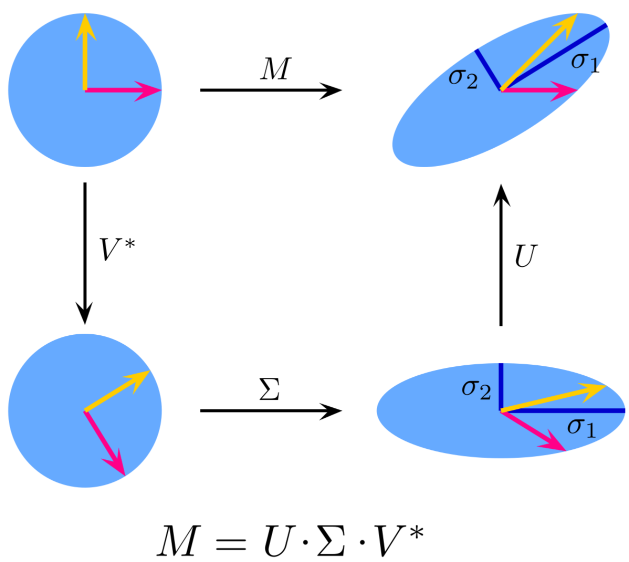
How do we minimize? Take derivative with respect to each variable.

$$\begin{aligned}\frac{\partial J}{\partial z_{ni}} &= \frac{1}{N}\sum_{n=1}^{N}\frac{\partial}{\partial z_{ni}}\left(-2\hat{\vec{x}}_n^\top\vec{x}_n + \hat{\vec{x}}_n^\top\hat{\vec{x}}_n\right) = \frac{1}{N}\sum_{n=1}^{N}\frac{\partial}{\partial z_{ni}}\left(-2\left(\sum_{i=1}^{M} z_{ni}\vec{u}_i\right)^\top\vec{x}_n + \left(\sum_{i=1}^{M} z_{ni}\vec{u}_i\right)^\top\left(\sum_{i=1}^{M} z_{ni}\vec{u}_i\right)\right)\\
&= \frac{1}{N}\sum_{n=1}^{N}\left(-2\left(\sum_{i=1}^{M}\frac{\partial}{\partial z_{ni}} z_{ni}\vec{u}_i\right)^\top\vec{x}_n + \frac{\partial}{\partial z_{ni}}\left(\sum_{i=1}^{M} z_{ni}\vec{u}_i\right)^\top\left(\sum_{i=1}^{M} z_{ni}\vec{u}_i\right)\right)\end{aligned}$$

...
continue this as an exercise (see PRML by Bishop for help)

Singular value decomposition (is basically PCA)

Definition: Singular value decomposition of a matrix M decomposes it into 3 matrices $U\Sigma V^\top$ where U and V are orthonormal and Σ is diagonal. If M has only real (not complex) entries, then U, V and Σ are also real. (pic from wikipedia)



What are U and V and how do they relate to PCA?

Suppose $X = U\Sigma V^\top$. Compute covariance:

$$\begin{aligned} XX^\top &= (U\Sigma V^\top)(U\Sigma V^\top)^\top \\ &= U\Sigma V^\top V\Sigma^\top U^\top \\ &= U\Sigma^2 U^\top \text{ (because } V \text{ is orthonormal)} \end{aligned}$$

U are the eigenvectors of $S = XX^\top$ which from above are the principal components.
Solve for V:

$$\begin{aligned} X &= U\Sigma V^\top \\ \Sigma^{-1}U^\top X &= V^\top \\ X^\top U\Sigma^{-1} &= V \end{aligned}$$

So $V = X^\top U\Sigma^{-1}$, data rotated by U and then inverse scaled by Σ.