

Complex numbers

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

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)

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

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 $z^n = 1$

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

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.

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

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

$$A\vec{v} - \lambda I\vec{v} = \vec{0}$$

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

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

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

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

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

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

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

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

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

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

* 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 $\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$?

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

Plug this into the equation:

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

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

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

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

Find the eigenvalues:

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

Find the eigenvectors:

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

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

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

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

$$\text{RHS: } A(c_1 e^{\lambda_1 t} \vec{v}_1 + c_2 e^{\lambda_2 t} \vec{v}_2) = c_1 e^{\lambda_1 t} (A\vec{v}_1) + c_2 e^{\lambda_2 t} (A\vec{v}_2) = c_1 e^{\lambda_1 t} (\lambda_1 \vec{v}_1) + c_2 e^{\lambda_2 t} (\lambda_2 \vec{v}_2) \checkmark$$

Can also do it the looong way. Let

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

$$\text{LHS: } \begin{bmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{bmatrix} = \begin{bmatrix} c_1(-3)e^{\lambda t} + c_2(1)e^{\lambda t} \\ c_1(-3)(-1)e^{\lambda t} + c_2(1)e^{\lambda t} \end{bmatrix} = \begin{bmatrix} -3c_1 e^{\lambda t} + c_2 e^{\lambda t} \\ 3c_1 e^{\lambda t} + c_2 e^{\lambda t} \end{bmatrix}$$

$$\text{RHS: } \begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix} \left(c_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^{-3t} + c_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^t \right) = \begin{bmatrix} -3c_1 e^{-3t} + c_2 e^t \\ 3c_1 e^{-3t} + c_2 e^t \end{bmatrix} \checkmark$$

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 $x \rightarrow \infty, y \rightarrow \infty$)

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

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:

$$\begin{aligned} (A + bI)\vec{v} &= (\lambda + b)\vec{v} \\ A\vec{v} + bI\vec{v} &= \lambda\vec{v} + b\vec{v} \end{aligned}$$

$\Rightarrow \vec{v}$ 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 $\rightarrow \infty$?

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):

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

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 :

$$\begin{aligned} A^2 &= (V\Lambda V^{-1})(V\Lambda V^{-1}) = V\Lambda^2 V^{-1} \\ \Rightarrow A^n &= V\Lambda^n V^{-1} \end{aligned}$$

As $n \rightarrow \infty$, $\lambda > 1$ 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

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

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 $A = A^T$.

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?

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

Orthonormal matrices are orthogonal matrices with columns with unit norm - how do we make $V = [\vec{v}_1 \quad \vec{v}_2]$ 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 $\iff V^{-1} = V^T$.

Proof: (right-direction) We will show for a two column matrix, but applies to an N-D matrix:

$$V = [\vec{v}_1 \quad \vec{v}_2] \quad \text{and} \quad V^T = \begin{bmatrix} \vec{v}_1^T \\ \vec{v}_2^T \end{bmatrix}$$

$$\Rightarrow V^T V = \begin{bmatrix} \vec{v}_1^T \\ \vec{v}_2^T \end{bmatrix} [\vec{v}_1 \quad \vec{v}_2] = \begin{bmatrix} \vec{v}_1^T \vec{v}_1 & \vec{v}_1^T \vec{v}_2 \\ \vec{v}_2^T \vec{v}_1 & \vec{v}_2^T \vec{v}_2 \end{bmatrix}$$

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

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

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

Theorem: Symmetric matrices ($A = A^T$) have real eigenvalues and orthogonal eigenvectors.

Proof: (for orthogonal eigenvectors)

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

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

By definition,

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

How do we achieve equality for this expression? $V^{-1} = V^T$

Thus, V is an orthonormal matrix \Rightarrow eigenvectors are orthogonal.

Positive semi-definite matrices

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

Theorem: Positive semi-definite matrices are symmetric.

Proof:

$$\begin{aligned} S = A^T A &\Rightarrow S^T = (A^T A)^T = A^T (A^T)^T = A^T A \\ &\Rightarrow S = S^T \end{aligned}$$

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

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

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

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 .

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

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

Rewrite S as diagonalization:

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

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?

$\text{cov}(x, y) = \frac{1}{N_{\text{stim}}} \sum_i (x_i - \bar{x})(y_i - \bar{y})$ where i is for different stimuli

\Rightarrow covariance matrix $S = \frac{1}{N_{\text{stim}}} X X^\top = \begin{bmatrix} 34 & 32 \\ 32 & 32 \end{bmatrix}$.

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

$$\lambda = 65, 2; \quad \vec{v}_1 = \begin{bmatrix} 0.72 \\ 0.69 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} -0.69 \\ 0.72 \end{bmatrix}$$

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 :

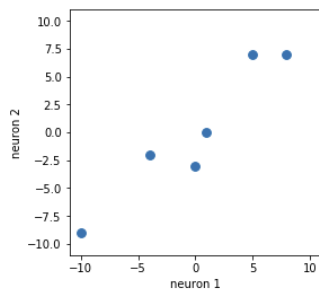
$$\text{proj}_{\vec{v}_1} \vec{x}_1 = \frac{\vec{x}_1^\top \vec{v}_1}{\|\vec{v}_1\|^2} \vec{v}_1 = (\vec{x}_1^\top \vec{v}_1) \vec{v}_1$$

```

In [1]: 1 import numpy as np
2 import matplotlib.pyplot as plt
3 %matplotlib inline
4
5 n1 = np.array([5,-4,8,-10,1,0])
6 n2 = np.array([7,-2,7,-9,0,-3])
7
8 A = np.concatenate((n1[np.newaxis,:], n2[np.newaxis,:]), axis=0)
9 print(A.shape)
10
11 # plot neuron activity
12 fig = plt.figure(figsize=(4,4))
13 ax = fig.add_subplot(111)
14 ax.scatter(n1,n2,s=60)
15 ax.set_xlabel('neuron 1')
16 ax.set_ylabel('neuron 2')
17 ax.set_xlim(-11,11)
18 ax.set_ylim(-11,11)
19 plt.show()

```

(2, 6)



```

In [2]: 1 print('covariance matrix')
2 covA = A @ A.T / A.shape[1]
3 print('[[ %2.0f, %2.0f ],\n [ %2.0f, %2.0f ]]'%(covA[0,0],covA[0,1],covA[1,0],covA[1,1]))
4
5 # find eigenvalues and eigenvectors of covariance matrix
6 lam, v = np.linalg.eig(A @ A.T / A.shape[1])
7
8 print('eigenvalues: %2.0f, %2.0f'%(lam[0],lam[1]))
9 print('eigenvectors: [%2.2f,%2.2f], [%2.2f,%2.2f]'%(v[0,0],v[1,0],v[0,1],v[1,1]))

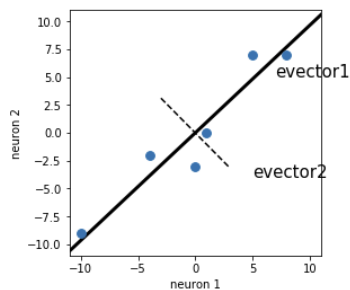
covariance matrix
[[ 34, 32 ],
 [ 32, 32 ]]
eigenvalues: 65, 2
eigenvectors: [0.72,0.69], [-0.69,0.72]

```

```

In [3]: 1 # plot EIGENVECTOR on top
2 fig = plt.figure(figsize=(4,4))
3 ax = fig.add_subplot(111)
4 ax.scatter(n1,n2,s=60)
5 ax.plot(np.array([-11,11]), np.array([-11,11])*v[1,0]/v[0,0],color='k', zorder=0, lw=3)
6 ax.text(7,5,'evector1',fontSize=15)
7 ax.plot(np.array([-3,3]), np.array([-3,3])*v[1,1]/v[0,1], '--',color='k', zorder=0)
8 ax.text(5,-4,'evector2',fontSize=15)
9 ax.set_xlabel('neuron 1')
10 ax.set_ylabel('neuron 2')
11 ax.set_xlim(-11,11)
12 ax.set_ylim(-11,11)
13 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)

$$X = [\vec{x}_1 \quad \dots \quad \vec{x}_n] \quad N \text{ neurons } \vec{x}_i \in \mathbb{R}^D$$

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.

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

if $\|\vec{u}_1\| \rightarrow \infty$, then variance will $\rightarrow \infty$. 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**:

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

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 maximizes 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 $\vec{u}_i \vec{u}_j^\top = \delta_{ij}$. 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:

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

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

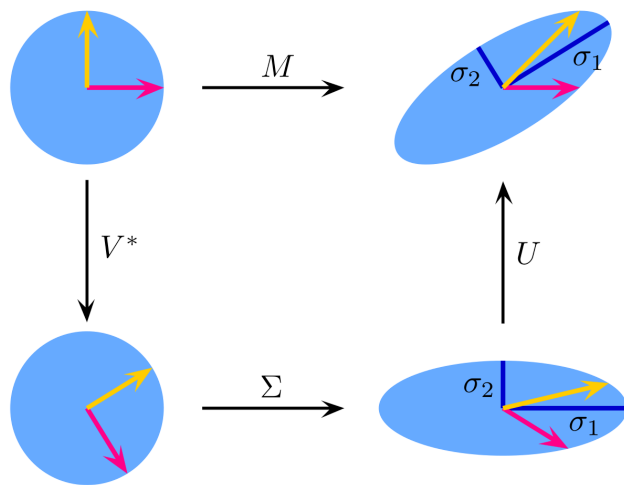
$$\begin{aligned} \frac{\partial J}{\partial z_{ni}} &= \frac{1}{N} \sum_{n=1}^N \frac{\partial}{\partial z_{ni}} \left(-2 \hat{\vec{x}}_n^\top \vec{x}_n + \hat{\vec{x}}_n^\top \hat{\vec{x}}_n \right) = \frac{1}{N} \sum_{n=1}^N \frac{\partial}{\partial z_{ni}} \left(-2 \left(\sum_{i=1}^M z_{ni} \vec{u}_i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

...

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 $U\Sigma V^T$ 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)



$$M = U \cdot \Sigma \cdot V^*$$

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

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

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

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

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

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