Guest Lecture — Change of Basis & Matrix Diagonalization

(Jonathan) Tyler Reichanadter

0 Introduction ~ Who am I?

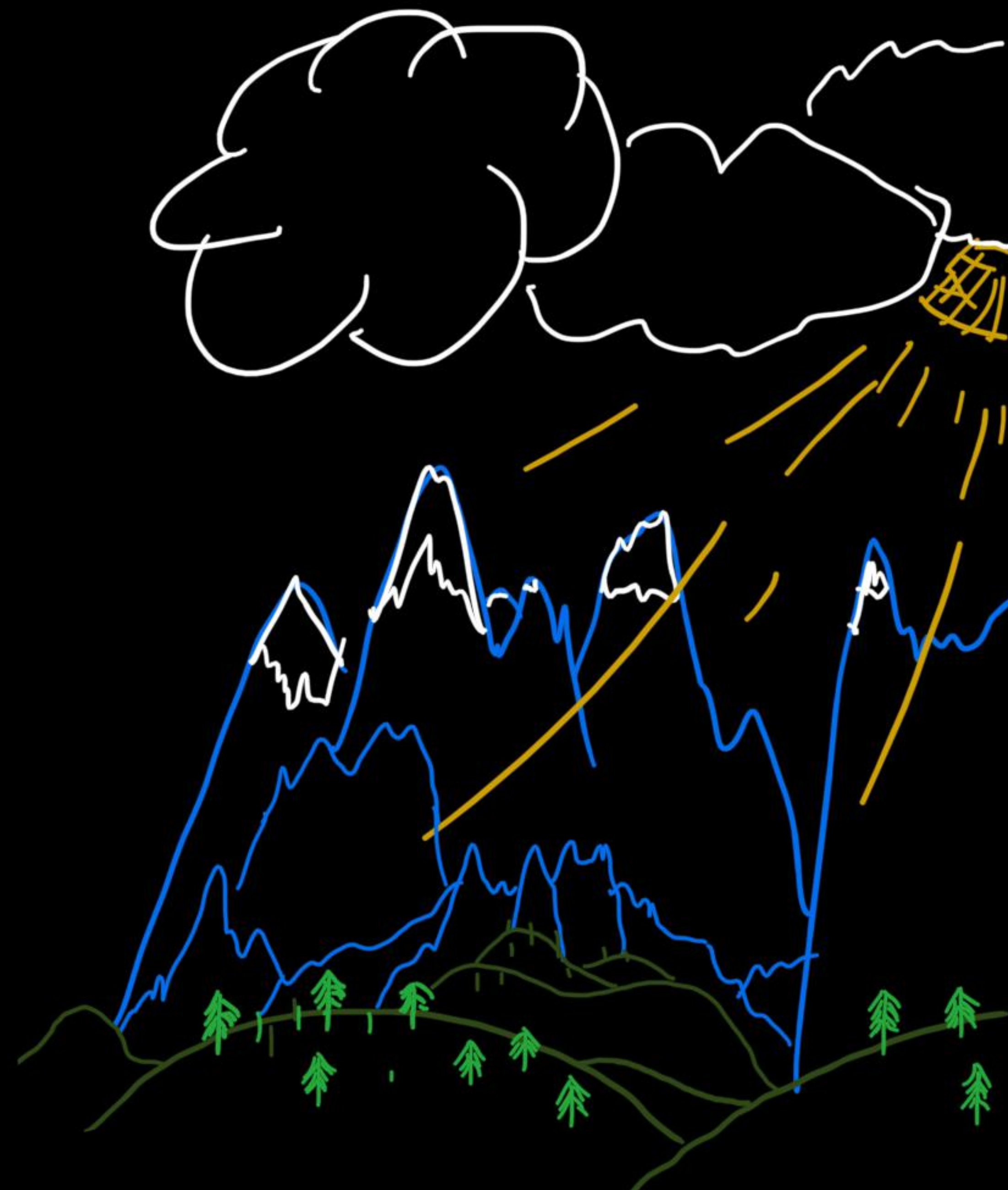
Howdy!

I'm Tyler

* I'm a GSI, but behind the curtains!
* 4th year PhD, research in "computational quantum materials"
* Enjoy weight-lifting, soccer, improv, through-hiking, moshi, etc.
* Raised in Seattle, undergrad for math and physics in Boulder, Colorado.

Let's zoom out for one moment...

... and take in the forest for the trees.



1 Recap

"What have we done in EECS 16A so far?"

(1) set of eqns

↓

(2) matrix-vector

↓

(3) gaussian elimination

↓

(4) span, linear independence

↓

(5) matrices as transformations

↓

(6) matrix inversion

↓

(7) column space, null space

↓

(8) eigenvalue & eigenspace

↓

(9) change of basis

Today

↓

[Module 1: DONE]

woot

yay!

Let's cover our eigenstuff once more.
Just to check our bases.

2 Eigenvalues & Eigenvectors

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

Any non-zero vector
Why not $\vec{v} = \vec{0}$?
$\vec{v} = \vec{0}$ becomes trivial case ($A\vec{0} = \vec{0}$ regardless of A)

Must be square! $A^{n \times n}$
But why?
$\vec{v}$ & $\lambda\vec{v} \in \mathbb{R}^n$
Otherwise an equation like this could never be true

Any number at all!
What does $\lambda = 0$ imply?

$A\vec{v} = \vec{0}$ A has a nontrivial nullspace!

$$v_1 \begin{bmatrix} \uparrow \\ \vec{a}_1 \\ \downarrow \end{bmatrix} \; v_2 \begin{bmatrix} \uparrow \\ \vec{a}_2 \\ \downarrow \end{bmatrix} \; \cdots \; v_n \begin{bmatrix} \uparrow \\ \vec{a}_n \\ \downarrow \end{bmatrix} \quad \vec{0}$$

* Equivalent statements:
  (1) Columns of A are linearly dependent (lin. ind.)
  (2) A is not invertible.
  (3) $\det(A) = 0$.

Refresher!

Suppose we have a vector space as defined

$$\text{span}\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k\}$$

(note each $\vec{a}_j \in \mathbb{R}^n$)

→ It Spans $\mathbb{R}^n$, if at least $k \geq n$ AND at least 'n' $\vec{a}_j$'s are lin. ind.

→ Is a Basis for $\mathbb{R}^n$, if $k = n$ AND ALL $\vec{a}_j$'s are lin. ind.

Eigenvalue/vector Machinery:

Getting eigenvalues

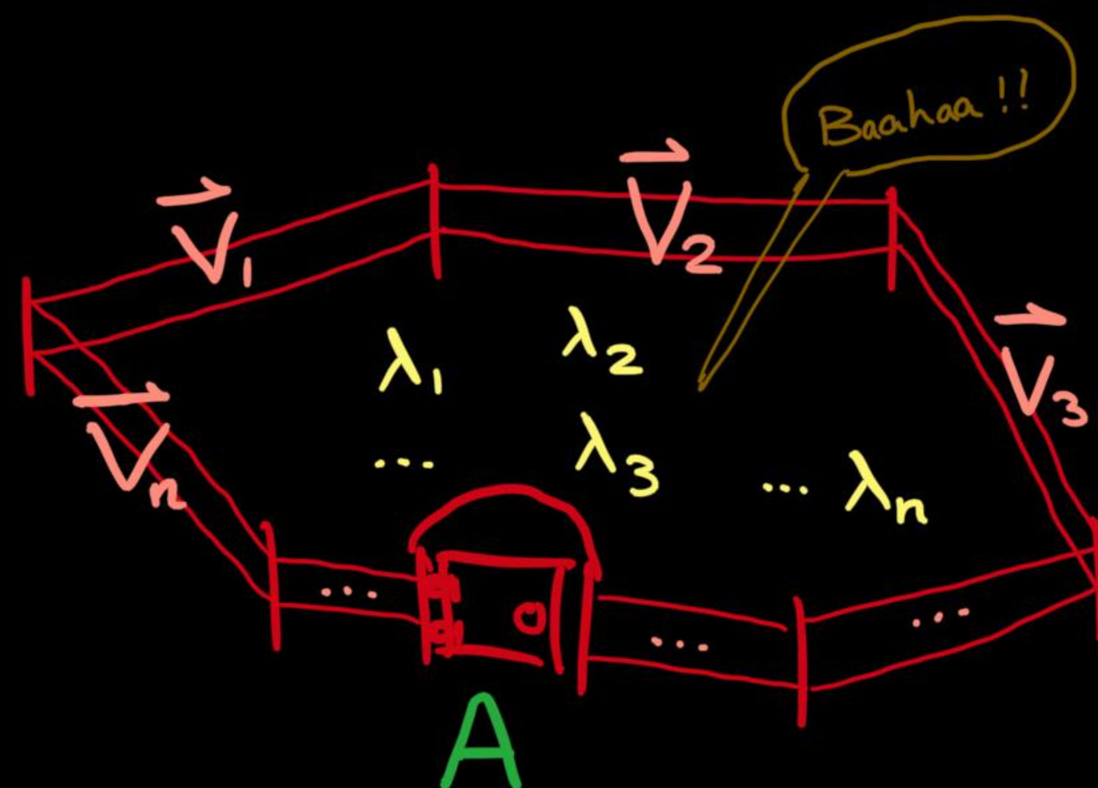
Getting Eigenvectors

1. Form $B_\lambda = A - \lambda I$
2. Find each λ yielding a non-empty nullspace for B_λ (equivalently, solve $\det(B_\lambda) = 0$).
3. For each λ, get the vector-space $\mathrm{null}(B_\lambda) = \mathrm{span}\{\vec{v}_1, \dots, \vec{v}_k\}$

* Note 1. If you are given λ's already, you can skip to step 3.

* Note 2. If you are given $\vec{v}$'s already, you get λ's literally through matrix multiplication $A\vec{v}_j$.

"In a broader sense, the eigenvalues form the heart of a matrix A, and the eigenvectors frame these eigenvalues"



Example: $A = \begin{bmatrix} 7 & 5 \\ 1 & 3 \end{bmatrix}$ $\lambda I = \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$

Step 1: $B_\lambda = A - \lambda I = \begin{bmatrix} 7-\lambda & 5 \\ 1 & 3-\lambda \end{bmatrix}$

If $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

Step 2: $\det(B_\lambda) = ad - bc$

$$= (7-\lambda)(3-\lambda) - 5$$
$$= \lambda^2 - 10\lambda + 21 - 5$$
$$= \lambda^2 - \underset{\beta}{10}\lambda + \underset{\gamma}{16} \equiv 0 \rightarrow \lambda = \frac{-\beta}{2} \pm \frac{1}{2}\sqrt{\beta^2 - 4\gamma}$$

You can always use quadratic formula for $A^{2\times 2}$ case!

$$= 5 \pm \frac{1}{2}\sqrt{100 - 64}$$
$$= 5 \pm \frac{1}{2}\sqrt{36}$$
$$= 5 \pm 3 \rightarrow \lambda = 2, 8$$

~ OR you can 'see' this factors ~

$$= (\lambda - 2)(\lambda - 8) \equiv 0 \rightarrow \lambda = 2, 8$$

Step 3: (for $\lambda_1 = 2$)

$A\vec{v}^{(1)} = \lambda_1 \vec{v}^{(1)} \rightsquigarrow B_{\lambda_1}\vec{v}^{(1)} = \vec{0}$ where $\vec{v}^{(1)} = \begin{bmatrix} v_1^{(1)} \\ v_2^{(1)} \end{bmatrix}$

$$\left[\begin{array}{cc|c} 5 & 5 & 0 \\ 1 & 1 & 0 \end{array}\right] \xrightarrow{\left\{ \begin{array}{l} R_1 \to \frac{1}{5}R_1 \\ R_2 \to R_2 - R_1 \end{array} \right\}} \left[\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right] \rightarrow v_1^{(1)} + \alpha = 0$$

Free parameter!
$v_2^{(1)} = \alpha$
(any $\alpha \in \mathbb{R}$)

$$\vec{v}^{(1)} = \begin{bmatrix} -\alpha \\ \alpha \end{bmatrix} = \alpha \begin{bmatrix} -1 \\ +1 \end{bmatrix}$$

$$\vec{v}^{(1)} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Finding $\vec{v}^{(2)}$ for B_{λ_2} is left as an exercise :)

~ Note! There are "infinite" choices for $\vec{v}^{(1)}$, but it must live in the null-space of B_{λ_1}, which happens to be a 1-dimensional vector space:

$$\vec{v}^{(1)} \in \text{span}\left\{ \begin{bmatrix} -1 \\ +1 \end{bmatrix} \right\}$$

3 Change of Basis

When we started, we just thought of $\vec{r}$ as an ordered list of numbers.
Now we can see it as "coordinates" as well!
Each term tells us how far to move in some "direction", which is also a vector!

$$\vec{r} = r_1 \vec{e}_1 + r_2 \vec{e}_2 + \cdots + r_n \vec{e}_n \iff \underbrace{\mathbb{I}\,\vec{r} = \vec{r}}_{\text{Equivalently}}$$

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

"Elementary column vectors"

What if we choose a new set of "directions" instead of $\vec{e}_j$'s?

Original basis

$$\vec{r} = r_1 \vec{e}_1 + r_2 \vec{e}_2$$

New basis

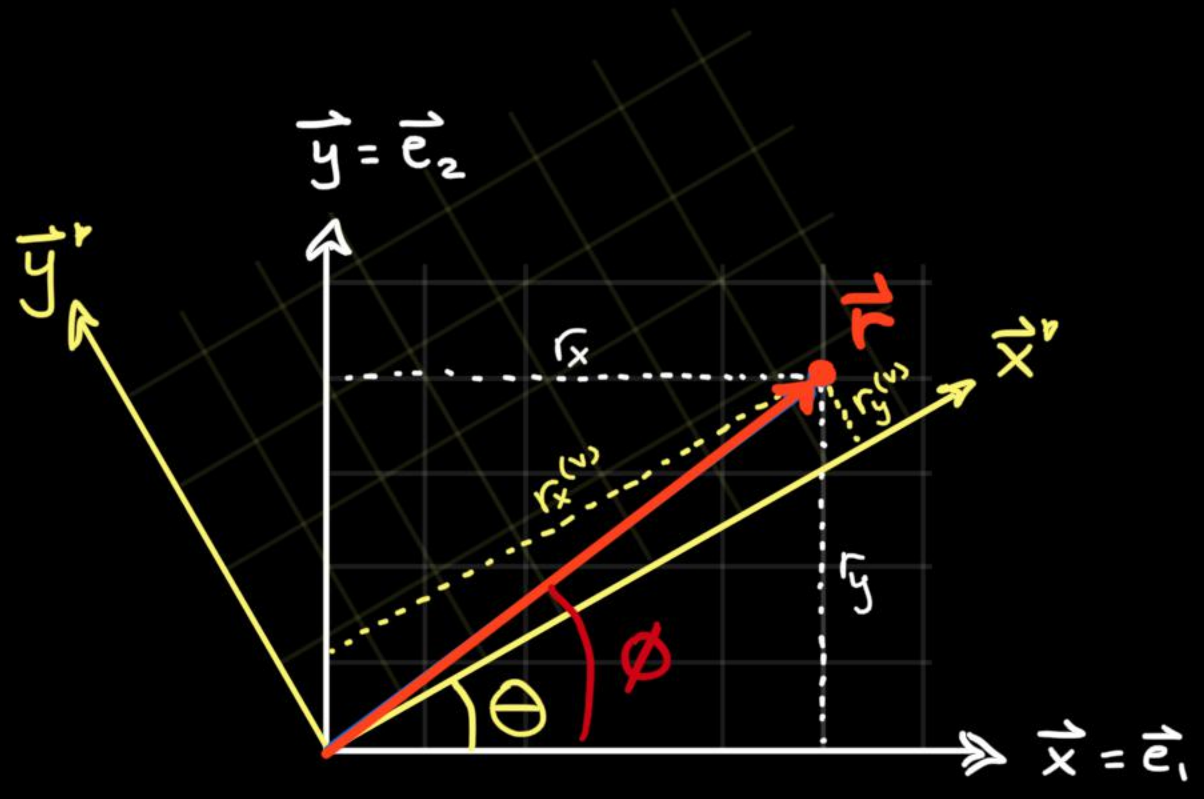
$$\equiv r_1^{(v)} \vec{x}' + r_2^{(v)} \vec{y}' \xrightarrow[\text{But check this out!}]{} = \underbrace{\begin{bmatrix} \uparrow & \uparrow \\ \vec{x}' & \vec{y}' \\ \downarrow & \downarrow \end{bmatrix}}_{V} \underbrace{\begin{bmatrix} r_1^{(v)} \\ r_2^{(v)} \end{bmatrix}}_{\vec{r}^{(v)}}$$

"Same" vector $\vec{r}$, but written in V basis.

But we want to get $\vec{r}^{(v)}$ <u>from</u> $\vec{r}$!
So we must invert:

$$\left\{ \vec{r}^{(v)} = V^{-1} \vec{r} \right\}$$

Example: Map $\vec{r}$ from $\{\vec{x}, \vec{y}\}$ basis to $\{\vec{x}', \vec{y}'\}$ basis!



$\cos(\theta) = C_\theta$
$\sin(\theta) = S_\theta$

$$\vec{x}' = \begin{bmatrix} C_\theta \\ S_\theta \end{bmatrix} \qquad \vec{y}' = \begin{bmatrix} -S_\theta \\ C_\theta \end{bmatrix}$$

$$V = \begin{bmatrix} C_\theta & -S_\theta \\ S_\theta & C_\theta \end{bmatrix} \equiv +\theta$$

CCW "Counter Clock-Wise"

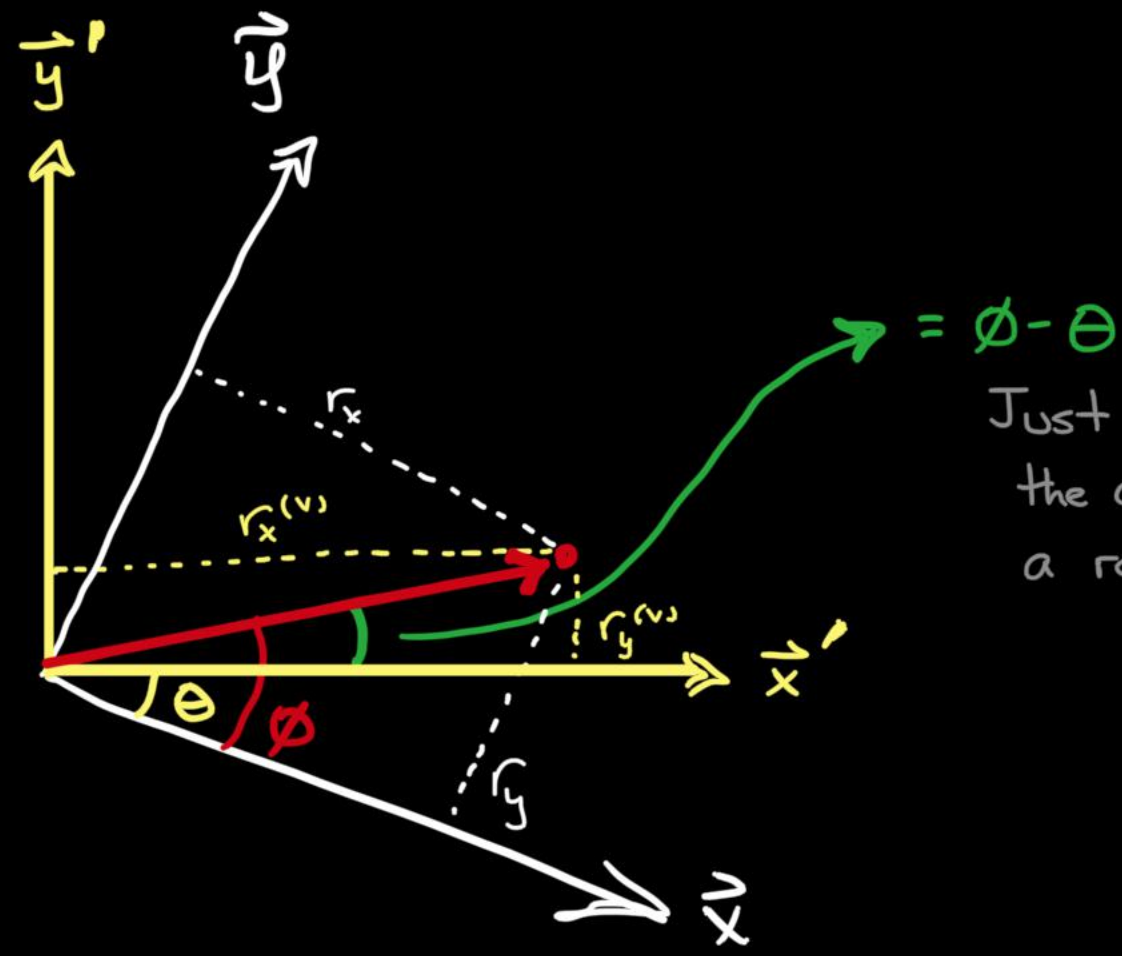
"Rotation by θ"

$$V^{-1} = \begin{bmatrix} C_\theta & S_\theta \\ -S_\theta & C_\theta \end{bmatrix}$$

Inverse is rotation by $-\theta$!

$$\vec{r}^{(v)} = \begin{bmatrix} C_\theta & S_\theta \\ -S_\theta & C_\theta \end{bmatrix} \underbrace{\begin{bmatrix} r_x \\ r_y \end{bmatrix}}_{\vec{r}}$$

Viewing $\vec{r}^{(v)}$ in V basis:



Just like from the original diagram, the change of basis here appears as a rotation by $-\theta$ CCW of $\vec{r}$!!

This was a cute example ♥, but can we generalize?

"What conditions are needed on

$$V = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} ?"$$

Hint 1: Recall that $\vec{r}^{(v)} = V^{-1}\vec{r}$.

Hint 2: What spaces do $\vec{r}$ and $\vec{r}^{(v)}$ live in?

$\Rightarrow$

1: All column vectors must be linearly independent.

2: $V^{n \times n}$ must be square since $\vec{r}, \vec{r}^{(v)} \in \mathbb{R}^n$.

Implies there are 'n' $\vec{v}_j$'s.

Thus...

The $\vec{v}_j$'s <u>must</u> form a <u>basis</u>!

We may want transform between two distinct bases $\{\vec{u}_1, \vec{u}_2\}$ and $\{\vec{v}_1, \vec{v}_2\}$:

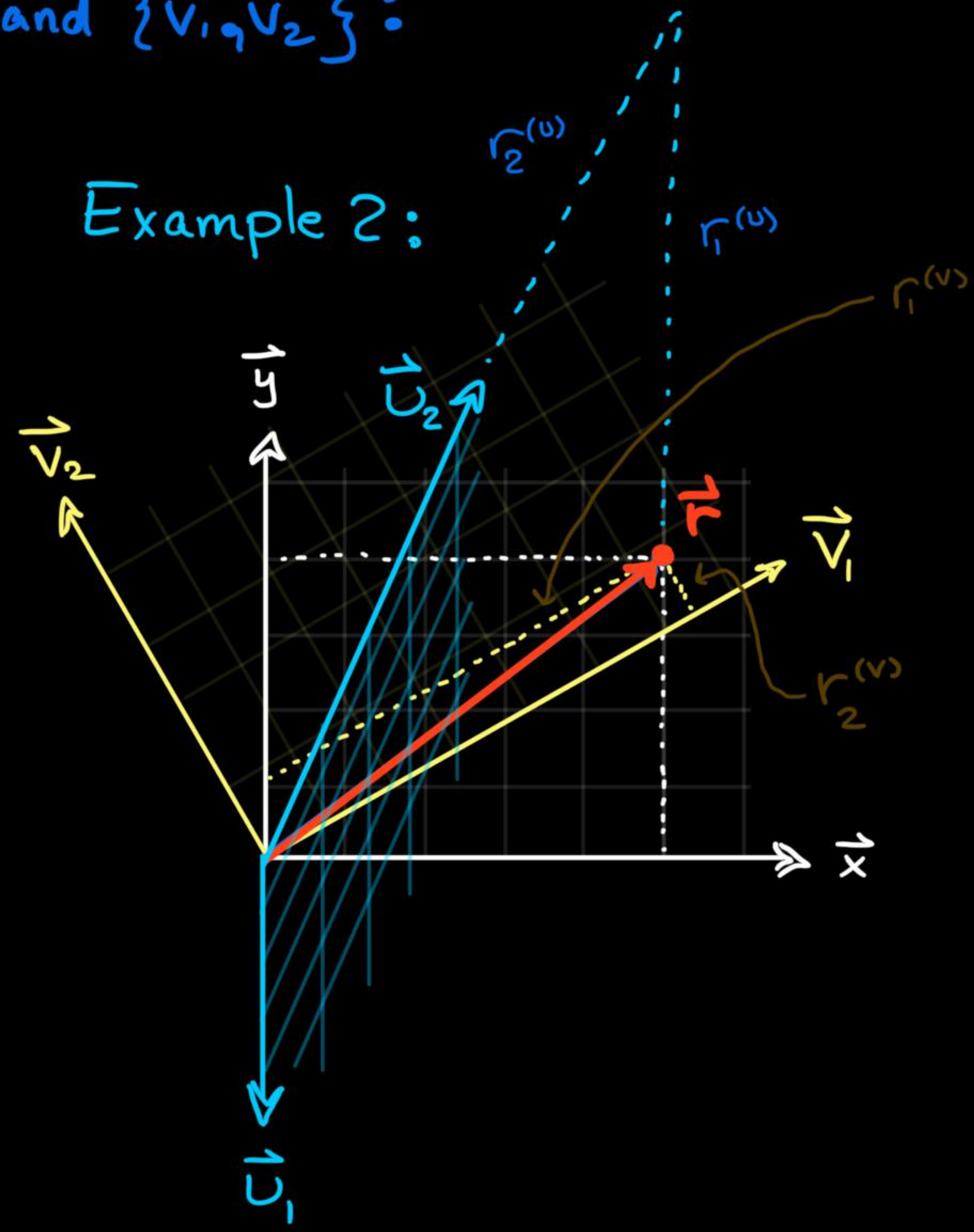
$$\vec{r}^{(u)} \longrightarrow \vec{r}^{(v)}$$

Example 2:

$$\vec{r} = r_1^{(v)} \vec{v}_1 + r_2^{(v)} \vec{v}_2$$
$$= r_1^{(u)} \vec{u}_1 + r_2^{(u)} \vec{u}_2$$

$$\underbrace{\vec{r} = V \vec{r}^{(v)} = U \vec{r}^{(u)}}$$

$$\boxed{\vec{r}^{(v)} = V^{-1} U \vec{r}^{(u)}}$$



Epiphany moment

"Can we use eigenvectors as our basis??"

(Terrible pikachu drawing)

Is that legal? Yes!

Well, terms & conditions apply...

Yes, always if all λ's are distinct.

Most times yes, even if some λ's repeat.

BUT WAIT!! What if the eigenvectors are linearly dep? ☹

Fear Not!

- **Claim:** Given the eigenvals & vecs $(\lambda_1, \lambda_2, \dots, \lambda_n)$ & $(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n)$ for matrix A, if $\lambda_1 \neq \lambda_2$ then $\vec{v}_1$ & $\vec{v}_2$ are linearly ind.

Proof: BY CONTRADICTION

"There exists" "such that"

↳ Suppose they are lin. dep., then $\exists\ \alpha, \beta \neq 0$ s.t. $\alpha\vec{v}_1 + \beta\vec{v}_2 = \vec{0}$.

↳ Now

$$A(\alpha\vec{v}_1 + \beta\vec{v}_2) = \alpha A\vec{v}_1 + \beta A\vec{v}_2$$
$$= \alpha\lambda_1\vec{v}_1 + \beta\lambda_2\vec{v}_2$$
$$= \lambda_1(\alpha\vec{v}_1 + \beta\vec{v}_2) + (\lambda_2 - \lambda_1)\beta\vec{v}_2$$

However, $A(\alpha\vec{v}_1 + \beta\vec{v}_2) = A\vec{0} = \vec{0}$.

So...

Can't be $\vec{0}$ by assumption

Can't be $\vec{0}$ by eigvec assumption.

$$\vec{0} = \cancel{\lambda(\vec{0})} + (\lambda_2 - \lambda_1)\beta\vec{v}_2$$

Non-zero by initial setup

Contradiction!

$\therefore\ \alpha\vec{v}_1 + \beta\vec{v}_2 \neq \vec{0}$

Thus $\vec{v}_1$ & $\vec{v}_2$ must be linearly independent! $\square$

Bonus Note:

So if all n eigenvalues λ_j's are distinct for $A \in \mathbb{R}^{n\times n}$, then span$\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ forms a basis for $\mathbb{R}^n$!!

- But... what if you have less than 'n' eigenvalues for A??

↳ Answer:
1. There are always 'n' eigenvalues, but some may be identical, eg $\lambda_1 = \lambda_2 = 3$.
2. If $\lambda_1 = \lambda_2 = 3$ for example, then $null(A - \lambda_1 I)$ will most often be 2D.

(It's possible to still be 1D, but this is somewhat rare. In this case A is said to be "non-diagonalizable".

4 Matrix Diagonalization

Completing the story: Eigenvectors + Basis Change

$$\vec{x} = x_1^{(v)}\vec{v}_1 + x_2^{(v)}\vec{v}_2 + \cdots + x_n\vec{v}_n \longrightarrow \vec{x}^{(v)} = \begin{bmatrix} x_1^{(v)} \\ x_2^{(v)} \\ \vdots \\ x_n^{(v)} \end{bmatrix}$$

where: $A\vec{v}_j = \lambda_j \vec{v}_j$

Step 1: Convert $\vec{x}$ to $\vec{x}^{(v)}$ in eigvec basis $\left(\vec{x}^{(v)} = V^{-1}\vec{x} \right)$

Step 2:
$$\begin{aligned} A\vec{x}^{(v)} &= x_1^{(v)} A\vec{v}_1 + x_2^{(v)} A\vec{v}_2 + \cdots + x_n^{(v)} A\vec{v}_n \\ &= x_1^{(v)}\lambda_1\vec{v}_1 + x_2^{(v)}\lambda_2\vec{v}_2 + \cdots + x_n^{(v)}\lambda_n\vec{v}_n \\ &= \Lambda\vec{x}^{(v)} \end{aligned}$$

$$\begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$

Step 3: Convert back from V basis: $A\vec{x} = V\underbrace{(A\vec{x})^{(v)}}_{\Lambda V^{-1}\vec{x}}$

Today's Punchline:

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

If non-diag, this becomes an upper triangular matrix

columns of eig-vecs

eigenvalues on diagonal

inv(V)

Diagonalizing Diagram

$$\begin{tikzcd} \vec{x} \arrow[r, bend left, "V^{-1}"] \arrow[d, "A"'] & \vec{x}^{(v)} \arrow[l, bend left, "V"] \arrow[d, "\Lambda"] \\ A\vec{x} \arrow[r, bend left, "V^{-1}"] & (A\vec{x})^{(v)} \arrow[l, bend left, "V"] \end{tikzcd}$$

Step 1: Change to eigen-basis.

Step 2: Apply 'A' with Λ.

Step 3: Change back.

Is this really useful? Yes !!

* Example 1: Compute A^2

$$A^2 = V\Lambda \underbrace{V^{-1}V}_{\text{Identity!}}\Lambda V^{-1} = V\underbrace{\Lambda\Lambda}_{\text{Diagonal Matrices}}V^{-1}$$

$$\begin{bmatrix}\lambda_1 & & 0\\ & \lambda_2 & \\ 0 & & \ddots & \lambda_n\end{bmatrix}\begin{bmatrix}\lambda_1 & & \\ & \lambda_2 & \\ & & \ddots & \lambda_n\end{bmatrix} = \begin{bmatrix}\lambda_1^2 & & \\ & \lambda_2^2 & \\ & & \ddots & \lambda_3^2\end{bmatrix}$$

* Example 2: Compute A^N

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

In summary, from the diagonal form we can easily compute matrix products!

$$A^N = V\begin{bmatrix}\lambda_1^N & & 0\\ & \lambda_2^N & \\ 0 & & \ddots & \lambda_n^N\end{bmatrix}V^{-1}$$

Fun fact: This works beyond integers, e.g. $N = 1/2$.
Even crazier... it works for any function $f(A)$!!

Ex: $$\log(A) = V\begin{bmatrix}\log(\lambda_1) & & 0\\ & \log(\lambda_2) & \\ 0 & & \ddots & \log(\lambda_n)\end{bmatrix}V^{-1}$$

* Example 3: Compute inverse A^{-1}

Use the same trick!

$$A^{-1} = V\underbrace{\begin{bmatrix}\frac{1}{\lambda_1} & & 0\\ & \frac{1}{\lambda_2} & \\ 0 & & \ddots & \frac{1}{\lambda_3}\end{bmatrix}}_{\Lambda^{-1}}V^{-1}$$

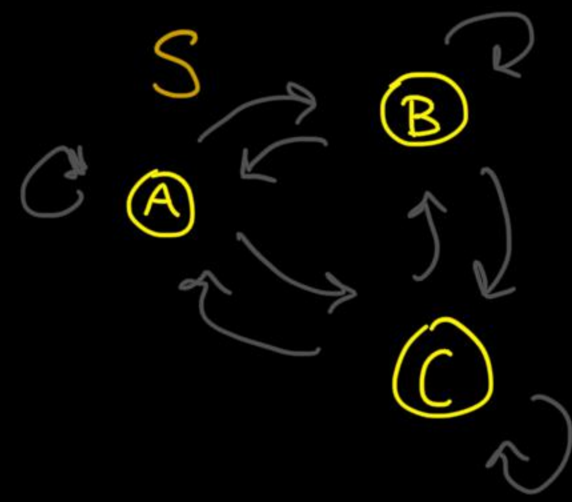
Food for thought: What happens if $\lambda_1 = 0$??

Verify
$$\begin{aligned} AA^{-1} &= V\Lambda V^{-1}V\Lambda^{-1}V^{-1}\\ &= V\underbrace{\Lambda\Lambda^{-1}}V^{-1}\\ &= VV^{-1}\\ &= I \quad ✓\end{aligned}$$

$$\begin{bmatrix}\lambda_1 & & 0\\ & \lambda_2 & \\ 0 & & \ddots & \lambda_n\end{bmatrix}\begin{bmatrix}1/\lambda_1 & & 0\\ & 1/\lambda_2 & \\ 0 & & \ddots & 1/\lambda_n\end{bmatrix} = \begin{bmatrix}1 & & \\ & 1 & \\ & & \ddots & \\ & & & 1\end{bmatrix} = I$$

* Example 4: Transition matrices!

Suppose we have some system where $S\vec{x}_{N-1} = \vec{x}_N$. Can we find out what happens in the long run? (steady state)



$$\vec{x}_N = \begin{bmatrix} x_N^{(A)} \\ x_N^{(B)} \\ x_N^{(C)} \end{bmatrix}$$

→ Start with $\vec{x}_0$, so $\vec{x}_1 = S\vec{x}_0$

$$\vec{x}_2 = S\vec{x}_1 = S(S\vec{x}_0) = S^2\vec{x}_0$$

thus $\vec{x}_{final} = \lim_{N\to\infty} S^N \vec{x}_0$

→ Diagonalize $S = V\Lambda V^{-1}$, so then $\lim_{N\to\infty} S^N = \lim_{N\to\infty} V\Lambda^N V^{-1}$

$$= V \begin{bmatrix} \lim_{N\to\infty}(\lambda_1)^N & & & \\ & \lim_{N\to\infty}(\lambda_2)^N & & \\ & & \ddots & \\ & & & \lim_{N\to\infty}(\lambda_n)^N \end{bmatrix} V^{-1}$$

Cases:

(a) $|\lambda_2| > 1$, blows up $\lambda_2^N \to \pm\infty$

(b) $|\lambda_2| < 1$, decays $\lambda_2^N \to 0$

(c) $\lambda_2 = +1$, constant $\lambda_2^N = 1$

this is the steady state ✓

(d) $\lambda_2 = -1$, alternating! $\lambda_2^N \to \pm 1$

weird infinite oscillations, but it cannot occur if S and $\vec{x}_0$ all have nonnegative values. ◡̈

Fin ☺