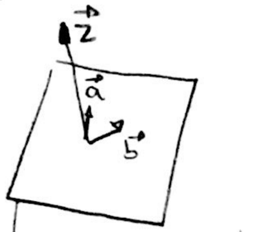
Eigenvalues and Eigenvectors

Given a square matrix $A \in \mathbb{R}^{n \times n}$, we say $\lambda \in \mathbb{C}$ is an eigenvalue of A and $x \in \mathbb{C}^n$ is the corresponding eigenvector if

→ set of complex #'s

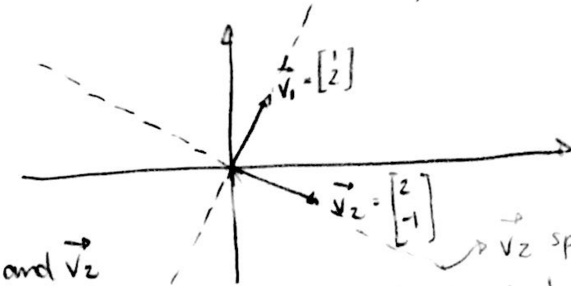
$$\underbrace{Ax = \lambda x}_{\text{scale}} \quad , x \neq 0$$

scale direction



→ plane is spanned by $\vec{a}$ and $\vec{b}$

Transformation takes the mirror image accross and the vectors $\vec{a}$ & $\vec{b}$ don't flip (just get scaled up, don't change directions → good basis)



→ $\vec{v}_1$ spans this line

→ $\vec{v}_2$ spans this line

$\vec{v}_1$ and $\vec{v}_2$ are examples of vectors that just get scaled up by a transformation $\mathbb{R}^2 \to \mathbb{R}^2$ that flips over vectors. $T(\vec{v}) = \lambda \vec{v}$

⇒ the lines that $\vec{v}_1$ and $\vec{v}_2$ span don't change → $\vec{v}_1$ & $\vec{v}_2$ are good basis.

→ transformation → eigenvector

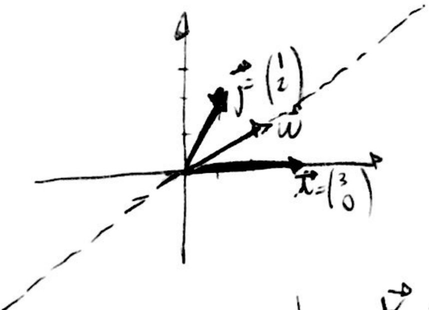
$$T(\vec{v}) = \lambda \vec{v}$$

eigenvalue associated with eigenvector

$$T(\vec{v}) = A\vec{v} \quad \to \quad A\vec{v} = \lambda \vec{v}$$

⇒ Matrices are linear transformations

For instance, consider the linear transformation $\mathbb{R}^2 \to \mathbb{R}^2$ (2 dimensions) that moves $\vec{i}$ to coordinates $\begin{bmatrix} 3 \\ 0 \end{bmatrix}$ and $\vec{j}$ to coordinates $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$: $A = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$



Now, image a vector $\vec{w}$ and its span on $\mathbb{R}^2$ (in blue)

If $\vec{w}$ is an eigenvector, then the transformation (ie. multiplication $A\vec{w}$) won't change $\vec{w}$ direction, will only scale $\vec{w}$ by some constant (eigenvalue) $2\vec{k} = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$

for example $\vec{k} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ ends up stretched by a factor (λ) of 2 : $A\vec{k} = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$

also $\vec{i} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$ ends up stretched by 3: $3\vec{i} = \begin{bmatrix} 9 \\ 0 \end{bmatrix}$ ⇒ $\vec{k}$ and $\vec{i}$ are eigenvectors of the transformation (A).

$A\vec{i} = \begin{bmatrix} 9 \\ 0 \end{bmatrix}$

→ linearity is going to imply that any other vector on the diagonal line spanned by $\vec{k}$ is gonna get "stretched" out by 2, and $\vec{i}$ by 3.

→ For this transformation, those are all the vectors with this special property of staying in their own span.

Eigenvalues and Eigenvectors

$A\vec{v} = \lambda\vec{v}$ → scalar × vector

matrix mult.

$A\vec{v} = (\lambda I)\vec{v}$ ⇒ both sides look like matrix vector mult. $\lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

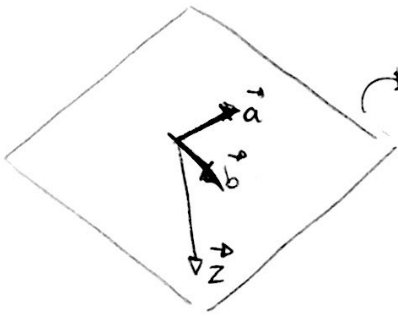
$A\vec{v} - \lambda I \vec{v} = 0$

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

e.g $\begin{bmatrix} 3-\lambda & 1 & 4 \\ 1 & 5-\lambda & 9 \\ 2 & 6 & 5-\lambda \end{bmatrix}$ → we want $\vec{v}$ s.t. the new matrix $* v = \vec{0}$ vector

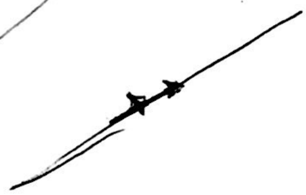
$\vec{v}$ should be non-zero.

The only way it's possible for the product of a matrix with a non-zero vector to be zero is if the transformation associated with that matrix squishes space into a lower dimension.



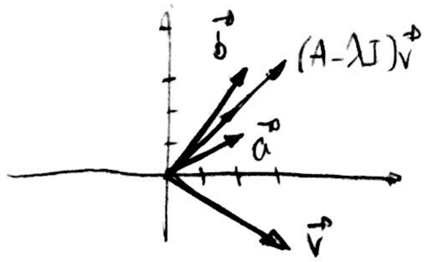
→ $\vec{z}$ cannot pop out of the plane.

the result is a line ($\vec{z}$ rotate the plane and it becomes a line, with $\vec{a}$ and $\vec{b}$ from $\vec{A}$ lying on the line)



→ This corresponds to a zero determinant for the matrix $\det(A - \lambda I) = 0$

Ex:



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

$A - \lambda I = \begin{bmatrix} 2-\lambda & 2 \\ 1 & 3-\lambda \end{bmatrix}$

$\det(A - \lambda I) = 0$ when $\lambda = 1$ //

→ when $\lambda = 1$ the matrix $(A - \lambda I)$ squishes space onto a line.

∴ there's a non zero vector $\vec{v}$ s.t. $(A - \lambda I)\vec{v}$ equals to the zero vector.

Returning to our 1st example

$A = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$ $\det\left(\begin{bmatrix} 3-\lambda & 1 \\ 0 & 2-\lambda \end{bmatrix}\right) = (3-\lambda)(2-\lambda) - (1 \cdot 0) = 0$ ⇒ $\lambda = 2$ or $\lambda = 3$

For $\lambda = 2$ $\begin{bmatrix} 3-2 & 1 \\ 0 & 2-2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ $\begin{cases} x + y = 0 \\ 0x + 0y = 0 \end{cases}$ $(x, y) = \text{span}\left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right)$

For $\lambda = 3$ $\begin{bmatrix} 3-3 & 1 \\ 0 & 2-3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ $(x, y) = \text{span}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right)$

Eigenvalues and Eigenvectors

Ex. A 2D transformation does not have to have eigenvectors.

Ex: $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ $\det(A - I\lambda) = \det\left(\begin{bmatrix} 0-\lambda & -1 \\ 1 & 0-\lambda \end{bmatrix}\right) = (-\lambda)(-\lambda) - (-1)(1)$

$= \lambda^2 + 1$

rotates every vector off its own span

(roots are imaginary $\lambda = i$ & $\lambda = -i$

$\Rightarrow$ complex # λ indicates no eigenvectors.

2 Methods to find eigen vectors

① solve for λ for $\det(A - \lambda I) = 0$

② for each λ, solve $A\vec{v} = \lambda\vec{v}$ or $(A - \lambda I)\vec{v} = \vec{0}$

1ST solve for $(A - \lambda I)\vec{v} = \vec{0}$

Ex $A = \begin{bmatrix} 2 & 0.8 \\ 0.8 & 0.6 \end{bmatrix}$ $\det\left(\begin{bmatrix} 2-\lambda & 0.8 \\ 0.8 & 0.6-\lambda \end{bmatrix}\right) = (2-\lambda)(0.6-\lambda) - 0.8^2$

$\{\lambda_1, \lambda_2\} = \{2.36, 0.23\}$

For $\lambda = 2.36$

$$\begin{pmatrix} 2-2.36 & 0.8 \\ 0.8 & 0.6-2.36 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \begin{cases} -0.36x + 0.8y = 0 \Rightarrow x = 2.2y \\ 0.8x - 1.76y = 0 \Rightarrow x = 2.2y \end{cases}$$

Take any pair (x, y) satisfying $x = 2.2y$ eg $\vec{v}_1 = \begin{pmatrix} 2.2 \\ 1 \end{pmatrix}$

make them unit vector $\vec{v}_1 = \begin{pmatrix} 2.2 \\ 1 \end{pmatrix}$, $\underbrace{\|\vec{v}_1\|}_{\text{length}} = \sqrt{(2.2)^2 + 1^2} = 2.41$

normalize $\vec{v} = \begin{pmatrix} \frac{2.2}{\|v_1\|} \\ \frac{1}{\|v_1\|} \end{pmatrix} = \begin{pmatrix} 0.91 \\ 0.41 \end{pmatrix} \rightarrow$ unit vector (directional) $\vec{V} = \text{span}\left\{\begin{pmatrix} 0.91 \\ 0.41 \end{pmatrix}\right\}$

Repeat for $\lambda = 0.23$

2ND Eigenvalue decomposition $A = U \Lambda U^T$ → orthogonality implies $U^T U = I$

$U^{-1} A U = U^{-1} U \Lambda U^T U$

$U^{-1} A U = \Lambda$ for A symmetric only

Ex: $A = \begin{bmatrix} 2 & 0.8 \\ 0.8 & 0.6 \end{bmatrix}$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} \begin{bmatrix} 2 & 0.8 \\ 0.8 & 0.6 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0.8 \\ 0.8 & 0.6 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

$$\begin{cases} \begin{bmatrix} 2 & 0.8 \\ 0.8 & 0.6 \end{bmatrix} \begin{bmatrix} a \\ c \end{bmatrix} = \begin{bmatrix} a\lambda_1 \\ c\lambda_1 \end{bmatrix} & \Sigma \vec{v}_1 = \lambda_1 \vec{v}_1 \\ \begin{bmatrix} 2 & 0.8 \\ 0.8 & 0.6 \end{bmatrix} \begin{bmatrix} b \\ d \end{bmatrix} = \begin{bmatrix} b\lambda_2 \\ d\lambda_2 \end{bmatrix} & \Sigma \vec{v}_2 = \lambda_2 \vec{v}_2 \end{cases}$$

$\rightarrow$ solve for $(A - \lambda_1 I)\vec{v}_1 = 0$ and $(A - \lambda_2 I)\vec{v}_2 = 0$

Eigenvectors and Eigenvalues

Some properties

- Trace of A is the sum of its eigenvalues

$$trA = \sum_{i=1}^{n} \lambda_i$$

- The determinant of A is the product of its eigenvalues

$$\det A = |A| = \prod_{i=1}^{n} \lambda_i$$

- The rank of A is equal to the # of non-zero eigenvalues of A.
- If A is non-singular (square and full rank) then $1/\lambda_i$ is an eigenvalue of A^{-1} associated with eigenvector x_i, i.e., $A^{-1}x_i = \left(\frac{1}{\lambda_i}\right)x_i$
- The eigenvalues of a diagonal matrix $D = diag(d_1, \dots, d_n)$ are just $d_1, \dots, d_n$.
- We can write all eigenvector equations simultameously as

$$AX = X\Lambda$$ where columns of $X \in \mathbb{R}^{n\times n}$ are eigenvectors of A and Λ is the diagonal matrix ${}_{n\times n}$ with eigenvalues.

If the eigenvectors of A are L.I. $\rightarrow X$ is invertible

$\rightarrow X^T X = I \rightarrow A = X\Lambda X^{-1}$ (decomposition)

For symmetric matrix A:

$\rightarrow$ all eigenvalues of A are real

$\rightarrow$ the eigenvectors of A are orthonormal (and X is orthogonal (denoted as U))

$\rightarrow$ A can be decomposed into $A = U\Lambda U^T$ (since $U^{-1} = U^T$ for orthogonal)

$\rightarrow$ Hence $x^T A x = x^T(U\Lambda U^T)x = y^T \Lambda y = \sum_{i=1}^{n} \lambda_i y_i^2$

($y = U^T x$ can be represented in this form since U is full rank)

$\rightarrow \sum_{i=1}^{n} \lambda_i \boxed{y_i^2}$ always $+$, if

$$\begin{cases} \lambda_i > 0 & \rightarrow A \text{ is PD} \\ \lambda_i \geq 0 & \rightarrow A \text{ is PSD} \\ \lambda_i < 0 & \rightarrow A \text{ is ND} \\ \lambda_i \leq 0 & \rightarrow A \text{ is NSD} \end{cases}$$

Application

$$\max_{x \in \mathbb{R}^n} x^T A x \text{ s.t } \|x\|_2^2 = 1$$

(want to find vector (norm) which maximizes $x^T A x$, assuming ordered λ_i's $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ the opt x is x_1, the eigenvector for $\underset{\text{max value}}{\underline{\lambda_1}}$)

$$\min_{x \in \mathbb{R}^n} x^T A x \text{ s.t. } \|x\|_2^2 = 1$$

(opt solution is x_n and the minimal value is λ_n)