Last time

$$\begin{bmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{bmatrix}$$

- Non-diagonalizable matrices
- Orthogonal diagonalizability
- Spectral theorem

Ex Let's orthogonally diagonalize

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

$$f_A(\lambda) = \det \begin{bmatrix} 1-\lambda & 1 & 1 \\ 1 & 1-\lambda & 1 \\ 1 & 1 & 1-\lambda \end{bmatrix}$$

$$= (1-\lambda)\det\begin{bmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{bmatrix} - \det\begin{bmatrix} 1 & 1 \\ 1 & 1-\lambda \end{bmatrix}$$

$$+ \det\begin{bmatrix} 1 & 1-\lambda \\ 1 & 1 \end{bmatrix}$$

$$= (1-\lambda)(\lambda^2 - 2\lambda) + \lambda + \lambda$$

$$= -\lambda^3 + \lambda^2 + 2\lambda^2 - 2\lambda + 2\lambda$$

$$= -\lambda^2(\lambda - 3) \qquad \lambda = 0, 3$$

$$E_0 = \ker\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \ker\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \operatorname{Span}\left(\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}\right)$$

$$E_3 = \ker\begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} = \ker\begin{bmatrix} 1 & 1 & -2 \\ 0 & 3 & -3 \\ 0 & -3 & 3 \end{bmatrix} = \ker\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

$= \text{Span}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

<u>Sanity check</u>: $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = 0.$

To get an ONEB, we apply Gram-Schmidt to E_0.

$$\vec{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \qquad \vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

$$\vec{u}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

$$\vec{v}_2^{\perp} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} - \left(\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \right) \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

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

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

$$\vec{u}_2 = \frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$$

ONB for E_0: $\left\{ \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \right\}$

ONB for E_3: $\left\{ \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$

ONEB: $\left\{ \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}.$

Orthogonal diagonalization

(0) If A is not symmetric, stop.

(1) Find a basis for each E_λ.

(2) Apply Gram-Schmidt to obtain an ONB for each E_λ.

(3) If $\vec{u}_1, \ldots, \vec{u}_n$ are the vectors from (2), then the matrix

$$S = \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_n \end{bmatrix}$$

is orthogonal, and $S^{-1}AS$ is diagonal.

Our last topics will be applications of the last several lectures. The first is based on the following innocent seeming observations.

<u>Observation 1</u> For <u>any</u> matrix A, A^TA is symmetric.

Indeed, $(A^TA)^T = A^T(A^T)^T = A^TA$.

Observation 2 Given orthogonal eigenvectors $\vec{v}_1$ and $\vec{v}_2$ of A^TA, $A\vec{v}_1$ and $A\vec{v}_2$ are still orthogonal.

Indeed,
$$
\begin{aligned}
A\vec{v}_1 \cdot A\vec{v}_2 &= (A\vec{v}_1)^T A\vec{v}_2 \\
&= \vec{v}_1^T A^T A \vec{v}_2 \\
&= \vec{v}_1 \cdot A^T A \vec{v}_2 \\
&= \vec{v}_1 \cdot \lambda_2 \vec{v}_2 \\
&= \lambda_2 (\vec{v}_1 \cdot \vec{v}_2) \\
&= 0.
\end{aligned}
$$

Ex $A = \begin{bmatrix} 6 & 2 \\ -7 & 6 \end{bmatrix}$, $A^TA = \begin{bmatrix} 6 & -7 \\ 2 & 6 \end{bmatrix}\begin{bmatrix} 6 & 2 \\ -7 & 6 \end{bmatrix} = \begin{bmatrix} 85 & -30 \\ -30 & 40 \end{bmatrix}$

$$f_{A^TA}(\lambda) = \lambda^2 - 125\lambda + 2500$$

$$= (\lambda - 100)(\lambda - 25) \qquad \lambda = 25, 100$$

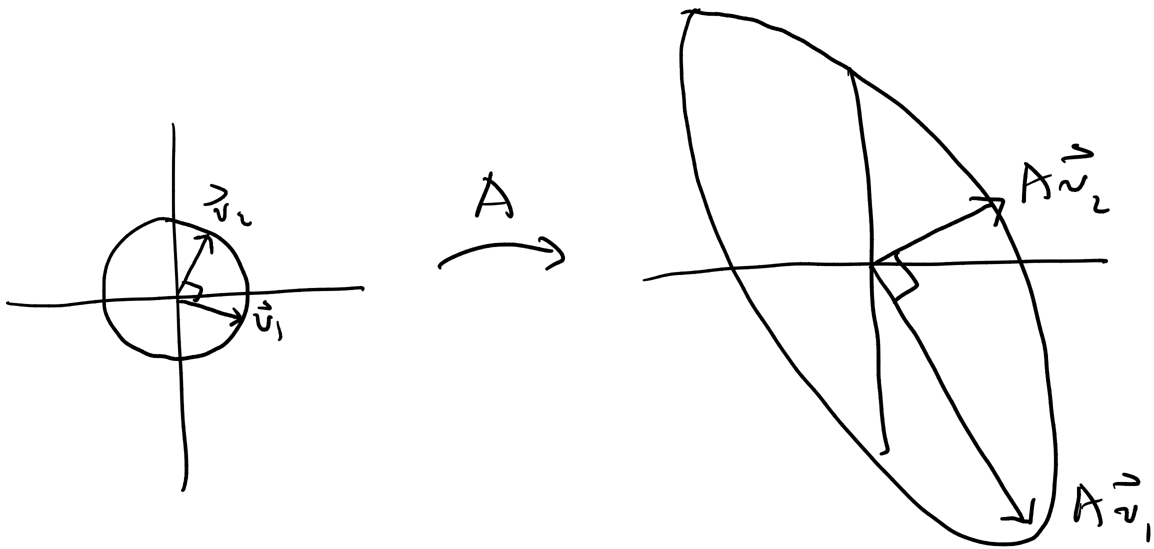
$$E_{100} = \ker\begin{bmatrix} -15 & -30 \\ -30 & -60 \end{bmatrix} = \ker\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} = \operatorname{Span}\begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

$$E_{25} = \ker\begin{bmatrix} 60 & -30 \\ -30 & 15 \end{bmatrix} = \ker\begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix} = \operatorname{Span}\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\vec{v}_1 = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ -1 \end{bmatrix} \qquad \vec{v}_2 = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$A\vec{v}_1 = \begin{bmatrix} 6 & 2 \\ -7 & 6 \end{bmatrix} \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} 10 \\ -20 \end{bmatrix}$$

$$A\vec{v}_2 = \begin{bmatrix} 6 & 2 \\ -7 & 6 \end{bmatrix} \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} 10 \\ 5 \end{bmatrix}$$



The effect of A is to transform the unit circle into an ellipse. The lengths of the axes of the ellipse are

$$|A\vec{v}_1| = \sqrt{\frac{100}{5} + \frac{400}{5}} = 10 = \sqrt{\lambda_1}$$

$$|A\vec{v}_2| = \sqrt{\frac{100}{5} + \frac{25}{5}} = 5 = \sqrt{\lambda_2}.$$

<u>Def</u> The <u>singular values</u> of the $m \times n$ matrix A are the square roots of the eigenvalues of A^TA. We denote them

$\sigma_1, \ldots, \sigma_n$ and list them in descending order:

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0.$$

For any $m \times n$ matrix A, there is a basis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ such that

(1) $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is orthonormal

(2) $\{A\vec{v}_1, \ldots, A\vec{v}_n\}$ is orthogonal

(3) $|A\vec{v}_i| = \sigma_i$.

Ex $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$, $A^TA = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$

$$f_{A^TA}(\lambda) = \det \begin{bmatrix} 1-\lambda & 1 & 0 \\ 1 & 2-\lambda & 1 \\ 0 & 1 & 1-\lambda \end{bmatrix}$$

$$= (1-\lambda)(\lambda^2 - 3\lambda + 1) - (1-\lambda)$$

$$= -\lambda^3 + 3\lambda^2 + \lambda^2 - \lambda - 3\lambda + 1 - 1 + \lambda$$

$$= -\lambda^3 + 4\lambda^2 - 3\lambda$$

$$= -\lambda(\lambda^2 - 4\lambda + 3)$$

$$= -\lambda(\lambda - 3)(\lambda - 1)$$

$\lambda = 0, 1, 3$

The example shows that, in general, the image of A has dimension equal to the number of nonzero singular values.

$$\boxed{\begin{aligned} \operatorname{rank}(A) &= \#\text{ of } \sigma_i \neq 0 \\ &= \#\text{ of } \lambda_i \neq 0 \\ &= \operatorname{rank}(A^TA) \end{aligned}}$$

As we've seen, $\{A\vec{v}_1, \ldots, A\vec{v}_n\}$ is orthogonal, and $|A\vec{v}_i| = \sigma_i$. If A has rank r, then $A\vec{v}_i \neq 0$ for $1 \leq i \leq r$, so $\{\vec{u}_1, \ldots, \vec{u}_r\}$ is orthonormal, where

$$\vec{u}_i = \frac{1}{\sigma_i} \vec{v}_i$$

Expanding this to an ONB for $\mathbb{R}^m$, we have the equations