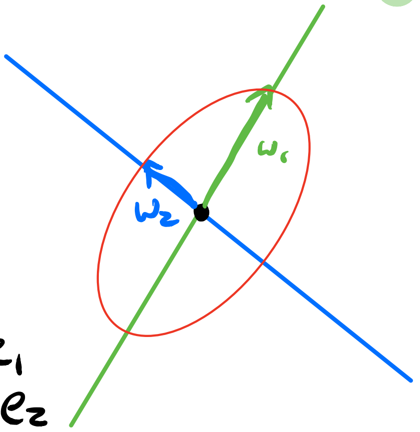
Geometry of the SVD

We've drawn pictures of triple products before:

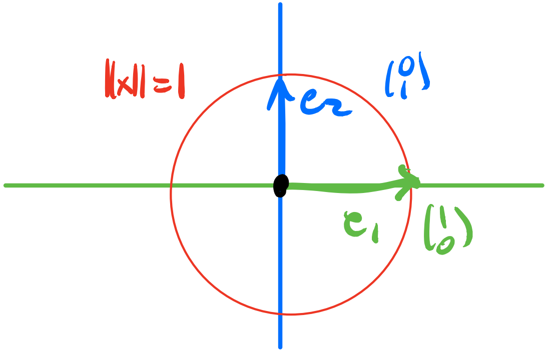
Diagonalization

$$A = \frac{1}{10}\begin{pmatrix}11 & 6\\ 9 & 14\end{pmatrix} = CDC^{-1} \qquad C = \begin{array}{c}\begin{array}{cc} w_1 & \ \ w_2\end{array}\\ \begin{pmatrix}2 & -1\\ 3 & 1\end{pmatrix}\end{array} \qquad D = \begin{pmatrix}2 & 0\\ 0 & 1/2\end{pmatrix}$$



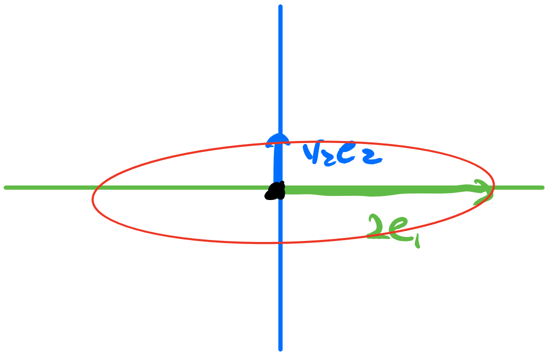
C^{-1}

$C^{-1}w_1 = e_1$
$C^{-1}w_2 = e_2$



D

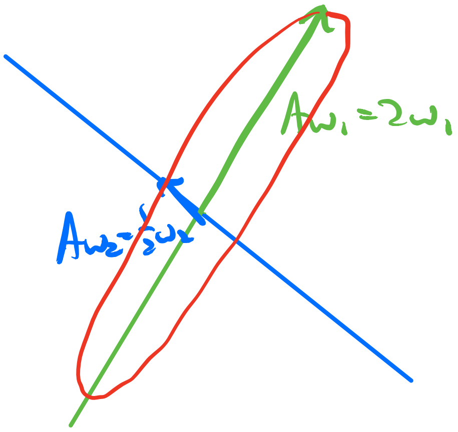
$De_1 = 2e_1$
$De_2 = \frac{1}{2}e_2$



C

$Ce_1 = w_1$
$Ce_2 = w_2$

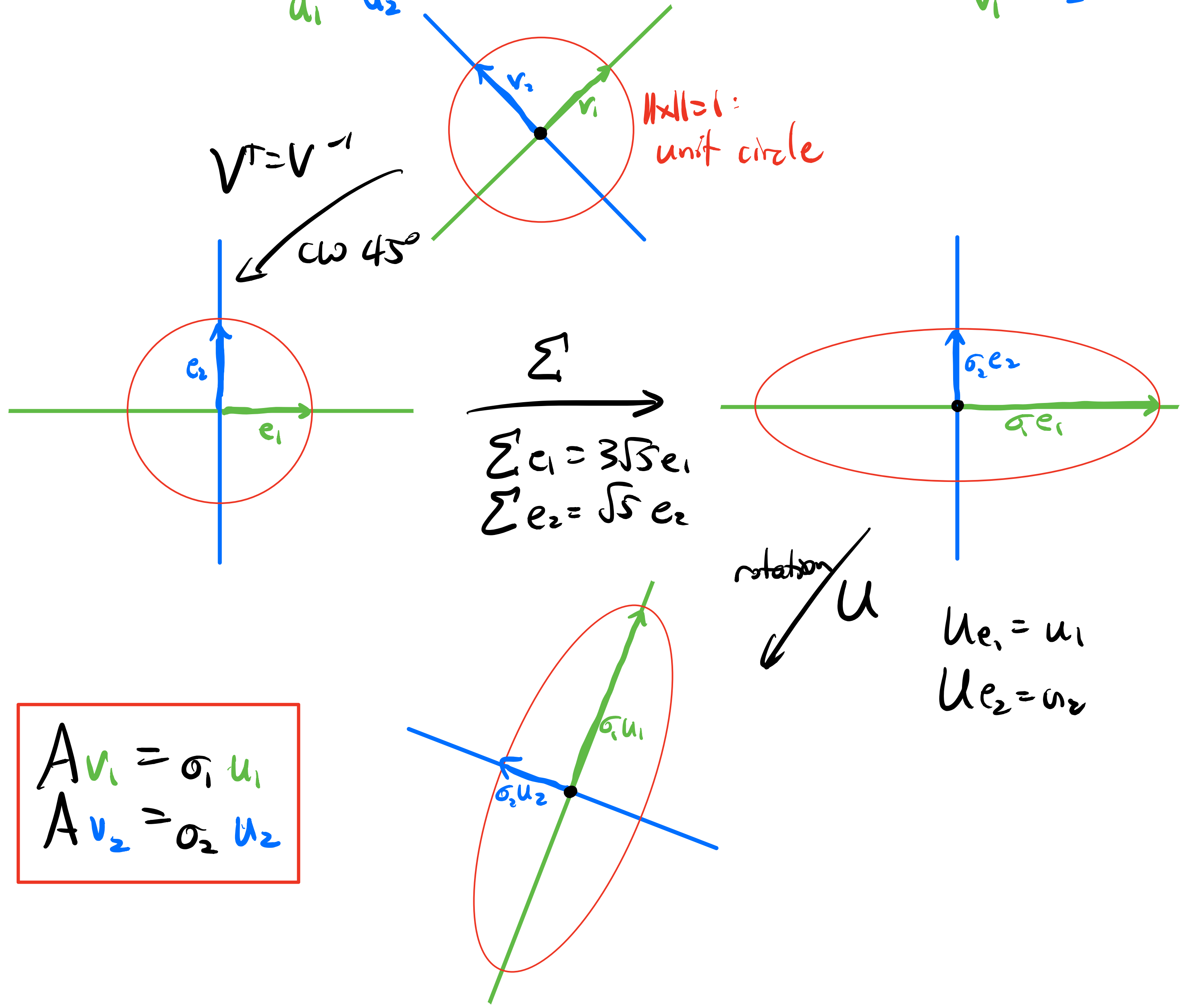
$$\boxed{\begin{aligned} Aw_1 &= 2w_1\\ Aw_2 &= \tfrac{1}{2}w_2\end{aligned}}$$



SVD: $A = \begin{pmatrix} 3 & 0 \\ 4 & 5 \end{pmatrix} = U \Sigma V^T$

ccw 45°

$$U = \frac{1}{\sqrt{10}} \begin{array}{c} \begin{pmatrix} 1 & -3 \\ 3 & 1 \end{pmatrix} \\ u_1 \quad u_2 \end{array} \qquad \Sigma = \begin{pmatrix} 3\sqrt{5} & 0 \\ 0 & \sqrt{5} \end{pmatrix} \qquad V = \frac{1}{\sqrt{2}} \begin{array}{c} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \\ v_1 \quad v_2 \end{array}$$



$$\Sigma e_1 = 3\sqrt{5} e_1$$
$$\Sigma e_2 = \sqrt{5} e_2$$

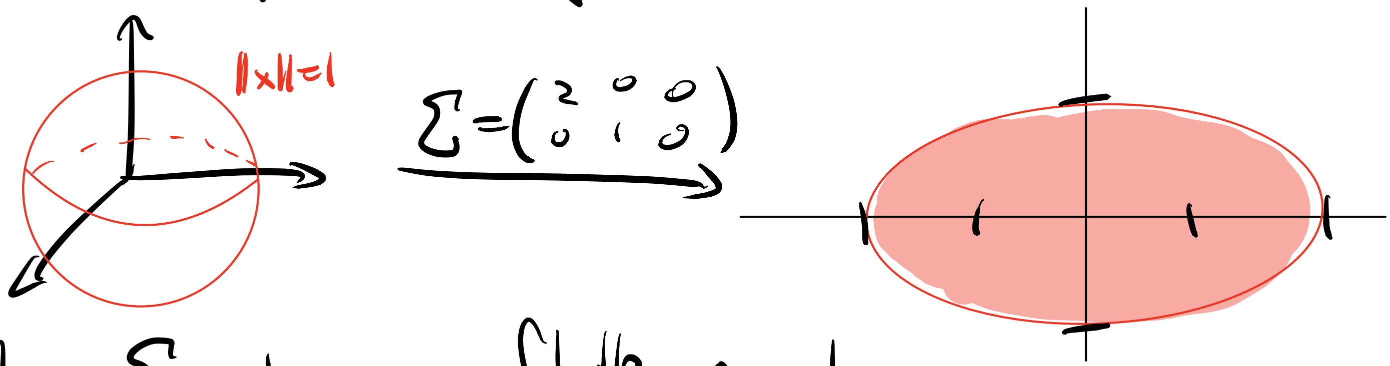
$$U e_1 = u_1$$
$$U e_2 = u_2$$

$$\boxed{\begin{aligned} A v_1 &= \sigma_1 u_1 \\ A v_2 &= \sigma_2 u_2 \end{aligned}}$$

$A =$ (first rotate/flip) (then stretch) (then rotate/flip)

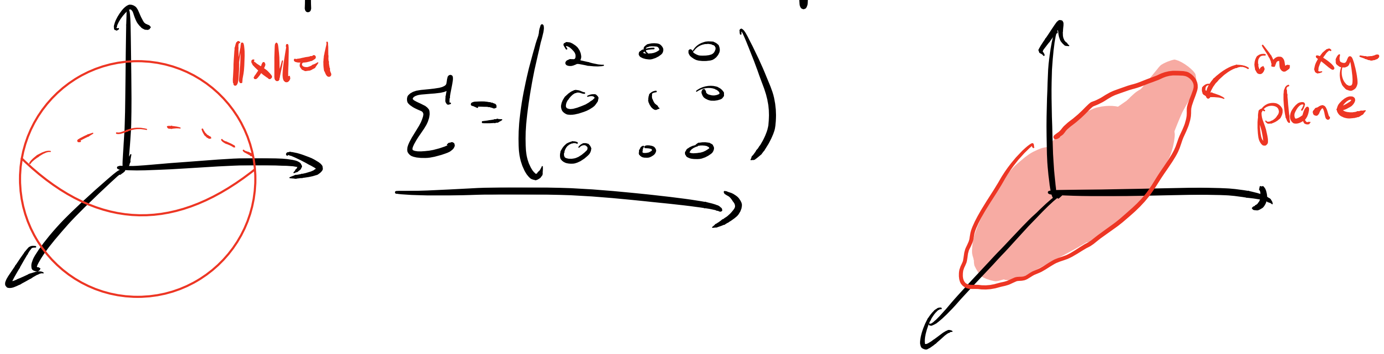
Notes / caveats:

- **Diagonalization:** start & end in $\{w_1, w_2\}$ basis
  SVD: start with $\{v_1, v_2\}$ & end with $\{u_1, u_2\}$ basis
- The V^T & U steps preserve lengths & angles (rotations / flips)
- The Σ step can change dimensions:



$$\xrightarrow{\Sigma = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}}$$

- The Σ step can flatten a sphere:



$$\xrightarrow{\Sigma = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}}$$

The Pseudo-Inverse

"Best possible" substitute for A^{-1} when A is not invertible.

Def: If Σ is an $m\times n$ diagonal matrix w/nonzero diagonal entries $\sigma_1, \dots, \sigma_r$, its pseudoinverse is $\Sigma^+ = n\times m$ diagonal matrix w/nonzero diagonal entries $\sigma_1^{-1}, \dots, \sigma_r^{-1}$.

$$\Sigma = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \rightsquigarrow \Sigma^+ = \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

NB: If Σ is invertible ($\Rightarrow$ square) then $\Sigma^+ = \Sigma^{-1}$:

$$\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}^+ = \begin{pmatrix} 1/3 & 0 \\ 0 & 1/2 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}^{-1}$$

Def: Let A be an $m\times n$ matrix with SVD

$$A = \sigma_1 u_1 v_1^T + \dots + \sigma_r u_r v_r^T \qquad A = U\Sigma V^T$$

The pseudoinverse of A is the $n\times m$ matrix

$$A^+ = \frac{1}{\sigma_1} v_1 u_1^T + \dots + \frac{1}{\sigma_r} v_r u_r^T \qquad A^+ = V\Sigma^+ U^T$$

NB: If A is invertible,

$$A^+A = (V\Sigma^+ \overset{I}{\cancel{U^T)(U}}\Sigma V^T) = V\overset{I}{\cancel{\Sigma^+\Sigma}}V^T = VV^T = I_n$$

so $A^+ = A^{-1}$.

NB: In general, for $i \le r$ we have $A^+u_i = \frac{1}{\sigma_i} v_i$

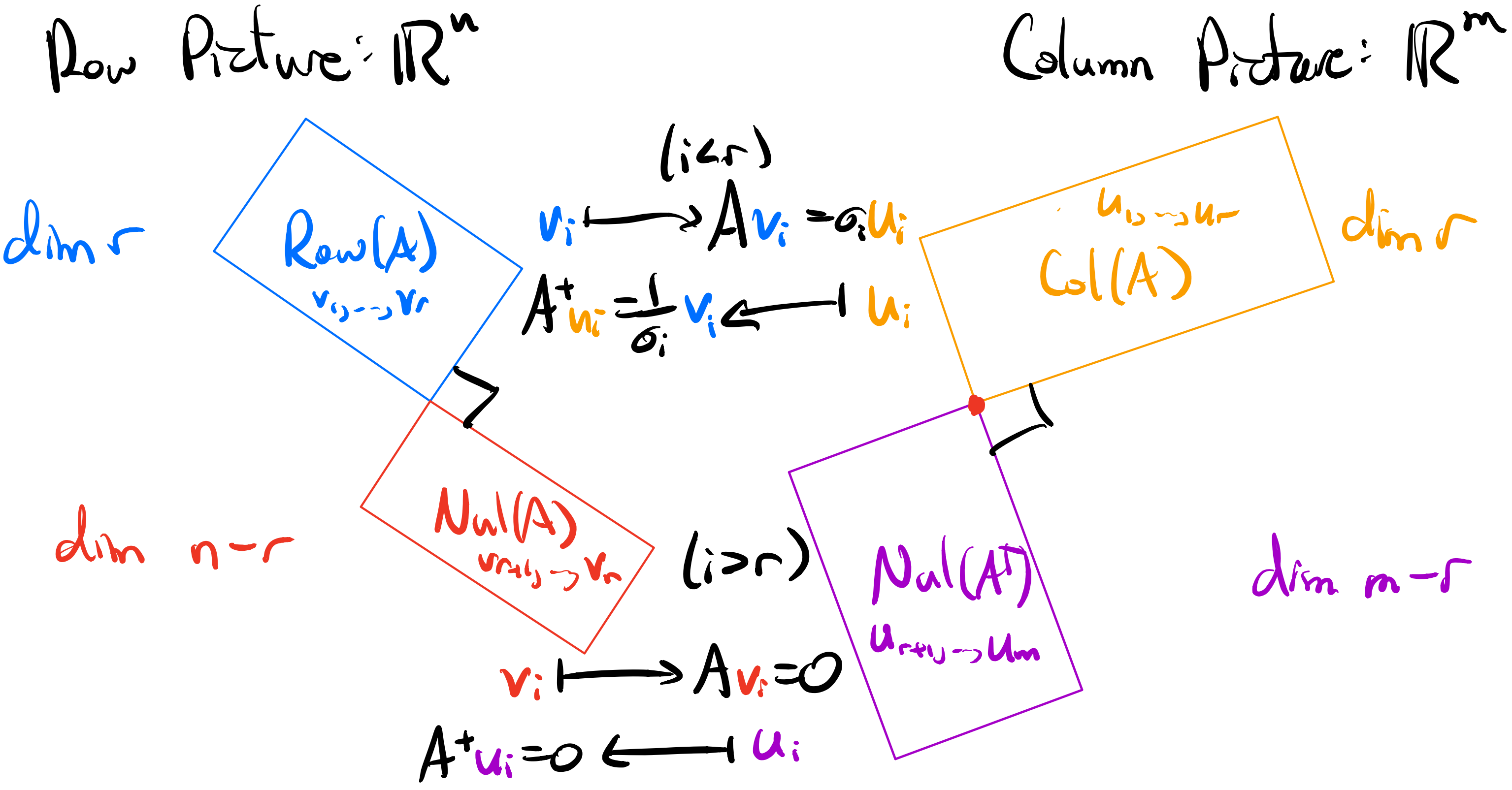
so $A^+Av_i = A^+(\sigma_i u_i) = \sigma_i A^+ u_i = \sigma_i \frac{1}{\sigma_i} v_i = v_i$

but $A^+Av_i = 0$ for $i > r$ b/c $v_i \in \mathrm{Nul}(A)$

Eg: $A = \begin{pmatrix}1&1\\1&1\end{pmatrix} \overset{SVD}{=} \frac{1}{\sqrt{2}}\begin{pmatrix}1&1\\1&-1\end{pmatrix}\cdot\begin{pmatrix}2&0\\0&0\end{pmatrix}\cdot\frac{1}{\sqrt{2}}\begin{pmatrix}1&1\\1&-1\end{pmatrix}^T$

$\Rightarrow A^+ = \frac{1}{\sqrt{2}}\begin{pmatrix}1&1\\1&-1\end{pmatrix}\cdot\begin{pmatrix}1/2&0\\0&0\end{pmatrix}\cdot\frac{1}{\sqrt{2}}\begin{pmatrix}1&1\\1&-1\end{pmatrix} = \frac{1}{4}\begin{pmatrix}1&1\\1&1\end{pmatrix}$

BIG PICTURE



What are A^+A / AA^+ if A is not invertible?

Prop: A^+A = matrix for orthogonal projection onto Row(A)

AA^+ = matrix for orthogonal projection onto Col(A)

→ $A^+Av_i = v_i$ for $i \le r$

$A^+Av_i = 0$ for $i > r$

If P = matrix for orthogonal projection onto Row(A)

$Pv_i = v_i$ for $i \le r$ b/c $v_i \in$ Row(A)

$Pv_i = 0$ for $i > r$ b/c $v_i \in \text{Nul}(A) = \text{Row}(A)^\perp = \text{Nul}(P)$

$\Rightarrow P = A^+A$

Prop: For any $b \in \mathbb{R}^m$, $\hat{x} = A^+ b$ is the shortest least-squares solution of $Ax = b$.

Proof: $\hat{x} = A^+ b$

$\rightsquigarrow A\hat{x} = AA^+ b =$ projection of b onto $Col(A) = \hat{b}$

$\Rightarrow A\hat{x} = \hat{b} \Rightarrow \hat{x}$ is a least-□ solution.

NB: $A^+ b = \frac{1}{\sigma_1}(u_1 \cdot b)\, v_1 + \cdots + \frac{1}{\sigma_r}(u_r \cdot b)\, v_r \in Row(A)$

Any other soln of $Ax = \hat{b}$ is $x = \hat{x} + y$ for $y \in Nul(A)$

$y \perp \hat{x} \in Row(A) = Nul(A)^\perp$

$\Rightarrow \|x\|^2 = \|\hat{x} + y\|^2 = \|\hat{x}\|^2 + \|y\|^2 \geq \|\hat{x}\|^2$

$\Rightarrow \hat{x}$ is shortest.

//

Eg: $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ $\quad b = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$

$\hat{x} = A^+ b = \frac{1}{4}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$Nul(A) = Span\left\{\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right\} \Rightarrow$

any other least-□ solution is

$x = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + a\begin{pmatrix} 1 \\ -1 \end{pmatrix}$