If S has <u>all different</u> eigenvalues, we know we can diagonalize now: $S = Q\Lambda Q^T$.

But we can diagonalize even if eigenvalues are repeated. (I won't prove it.)

<u>Example</u> $S = \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$

<u>Eigenvalues</u>: $\begin{vmatrix} -1-\lambda & 1 & 1 \\ 1 & -1-\lambda & 1 \\ 1 & 1 & -1-\lambda \end{vmatrix} = (-1-\lambda)\begin{vmatrix} -1-\lambda & 1 \\ 1 & -1-\lambda \end{vmatrix} - 1\begin{vmatrix} 1 & 1 \\ 1 & -1-\lambda \end{vmatrix} + 1\begin{vmatrix} 1 & -1-\lambda \\ 1 & 1 \end{vmatrix}$

$= (-1-\lambda)(\lambda^2+2\lambda+\cancel{1}-\cancel{1}) - (-1-\lambda-1) + (1+1+\lambda)$

$= -\lambda^2-\lambda^3-2\lambda-2\lambda^2+\lambda+2+\lambda+2 = -\lambda^3-3\lambda^2+4 = 0$

By inspection: $\lambda = 1$ is a root. So can factor out $1-\lambda$:

$(1-\lambda)(\lambda^2+a\lambda+b) = -\lambda^3 + \underbrace{(1-a)}_{-3}\lambda^2 + \underbrace{(a-b)}_{0}\lambda + \underbrace{b}_{4}$

So $a=b=4$: $(1-\lambda)\underbrace{(\lambda^2+4\lambda+4)}_{(\lambda+2)^2} = 0 \longrightarrow \lambda = 1, -2, -2$.

<u>Eigenvectors for $\lambda = 1$</u>:

Solve $(S-I)\vec{x} = \vec{0}$

$\underbrace{\begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix}}_{\text{Entries in each row add to } 0}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

one eigenvector is $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

<u>Eigenvectors for $\lambda = -2$</u>:

Solve $(S+2I)\vec{x} = \vec{0}$

$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

So $x_1+x_2+x_3 = 0 \rightsquigarrow$

$\vec{x} = \begin{bmatrix} -x_2-x_3 \\ x_2 \\ x_3 \end{bmatrix} = x_2\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

So here is a basis of eigenvectors:

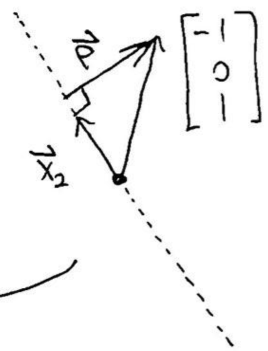
$$\underbrace{\begin{bmatrix}1\\1\\1\end{bmatrix}}_{\lambda=1},\ \underbrace{\begin{bmatrix}-1\\1\\0\end{bmatrix},\ \begin{bmatrix}-1\\0\\1\end{bmatrix}}_{\lambda=-2}$$

But it's not orthonormal, or even orthogonal.

rescale to unit vector

$$\vec{x}_1=\frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix},\quad \vec{x}_2=\frac{1}{\sqrt{2}}\begin{bmatrix}-1\\1\\0\end{bmatrix}$$

Use Gram-Schmidt process:



Vector $\vec{e}$ is orthogonal to $\vec{x}_2$, <u>and</u> it's still an eigenvector.

$$\vec{e}=\begin{bmatrix}-1\\0\\1\end{bmatrix}-P\begin{bmatrix}-1\\0\\1\end{bmatrix}=\begin{bmatrix}-1\\0\\1\end{bmatrix}-\frac{\vec{x}_2\vec{x}_2^{\,T}}{\underbrace{\vec{x}_2^{\,T}\vec{x}_2}_{=1}}\begin{bmatrix}-1\\0\\1\end{bmatrix}$$

projection onto span($\vec{x}_2$)

$=1$ ($\vec{x}_2$ is a unit vector)

$$=\begin{bmatrix}-1\\0\\1\end{bmatrix}-\frac{1}{\sqrt{2}}\begin{bmatrix}-1\\1\\0\end{bmatrix}\underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix}-1&1&0\end{bmatrix}\begin{bmatrix}-1\\0\\1\end{bmatrix}}_{1}=\begin{bmatrix}-1\\0\\1\end{bmatrix}-\frac{1}{2}\begin{bmatrix}-1\\1\\0\end{bmatrix}=\begin{bmatrix}-1/2\\-1/2\\1\end{bmatrix}$$

Finally, $\vec{x}_3=\dfrac{1}{\|\vec{e}\|}\vec{e}=\dfrac{1}{\sqrt{(-\frac{1}{2})^2+(-\frac{1}{2})^2+1^2}}\begin{bmatrix}-1/2\\-1/2\\1\end{bmatrix}=\sqrt{\dfrac{2}{3}}\begin{bmatrix}-1/2\\-1/2\\1\end{bmatrix}=\dfrac{1}{\sqrt{6}}\begin{bmatrix}-1\\-1\\2\end{bmatrix}$

Now we can diagonalize S:

$$\begin{bmatrix}-1&1&1\\1&-1&1\\1&1&-1\end{bmatrix}=\underbrace{\begin{array}{c}\begin{bmatrix}1/\sqrt{3}&-1/\sqrt{2}&-1/\sqrt{6}\\1/\sqrt{3}&1/\sqrt{2}&-1/\sqrt{6}\\1/\sqrt{3}&0&2/\sqrt{6}\end{bmatrix}\\ \vec{x}_1\qquad\ \ \vec{x}_2\qquad\ \ \vec{x}_3\end{array}}_{Q}\underbrace{\begin{bmatrix}1&0&0\\0&-2&0\\0&0&-2\end{bmatrix}}_{\Lambda}\underbrace{\begin{bmatrix}1/\sqrt{3}&1/\sqrt{3}&1/\sqrt{3}\\-1/\sqrt{2}&1/\sqrt{2}&0\\-1/\sqrt{6}&-1/\sqrt{6}&2/\sqrt{6}\end{bmatrix}}_{Q^T=Q^{-1}}$$

Some <u>special</u> symmetric matrices: Suppose A is an $m \times n$ matrix (we don't need $m=n$).

Then $\underset{n\times n}{A^TA}$ and $\underset{m \times m}{AA^T}$ are both symmetric (not usually the same)

What's special about A^TA? For one thing, its eigenvalues are not just <u>real</u> numbers. They are also <u>positive</u> (or 0).

Why? Suppose $A^TA\vec{x} = \lambda\vec{x}$ with $\vec{x} \neq \vec{0}$.

Then
$$\underset{(A\vec{x})^T A x}{\vec{x}^T A^T A \vec{x}} = \underset{\lambda(\vec{x}^T\vec{x})}{\lambda(\vec{x}^T\vec{x})} \rightarrow \lambda = \frac{(A\vec{x})\cdot(A\vec{x})}{\vec{x}\cdot\vec{x}} = \frac{\|A\vec{x}\|^2}{\|\vec{x}\|^2} \geq 0$$

Let's arrange the eigenvalues in decreasing order:

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r > \lambda_{r+1} = 0, \ldots, \lambda_n = 0$$

$$\vec{v}_1, \vec{v}_2, \ldots \vec{v}_r, \underbrace{\vec{v}_{r+1}, \ldots, \vec{v}_n}_{\text{Basis for } N(A^TA)} \leftarrow \text{Orthonormal basis of eigenvectors}$$

Since each $\lambda_i \geq 0$, we can take their square roots: $\sigma_i = +\sqrt{\lambda_i}$

called the <u>singular values</u> of A

Also a basis for $N(A)$, because $N(A) = N(A^TA)$.

Why? If $\vec{x}$ in $N(A)$, the $A\vec{x} = \vec{0}$, so

$A^TA\vec{x} = A^T\vec{0} = \vec{0} \rightarrow \vec{x}$ is <u>also</u> in $N(A^TA)$.

On the other hand, if $A^TA\vec{x} = \vec{0}$, then

$$\vec{x}^TA^TA\vec{x} = \vec{x}^T\vec{0} = 0$$
$$\| \qquad\qquad \|$$
$$(A\vec{x})^TA\vec{x} = \|A\vec{x}\|^2 \rightarrow A\vec{x} = \vec{0}.$$

So $\vec{x}$ is in $N(A)$ as well.

Note that $\dim N(A^TA) = \dim N(A) = n - r$, where $\underbrace{r = \text{rank}(A)}$

↗ same as rank of A^TA.

So the non-zero singular values $\sigma_1, \sigma_2, \ldots, \sigma_r$ go along with rank(A)-many orthonormal eigenvectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r$: $A^TA\vec{v}_i = \sigma_i^2 \vec{v}_i$

For $i = 1, 2, \ldots, r$, let's define $\vec{u}_i = \frac{1}{\sigma_i} A\vec{v}_i$ (vectors in $\mathbb{R}^m$ since A is $m \times n$)

What's special about $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r$?

① They are an orthonormal set in $\mathbb{R}^m$:

$$u_i^T \vec{u}_j = \frac{1}{\sigma_i \sigma_j}(Av_i)^T(A\vec{v}_j) = \frac{1}{\sigma_i\sigma_j} \vec{v}_i^T \underbrace{A^TA\vec{v}_j}_{\sigma_j^2 \vec{v}_j}$$

$$= \frac{\sigma_j}{\sigma_i} v_i^T \vec{v}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Because $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is orthonormal.

② They form a orthonormal basis of the column space $C(A)$:

- They are linearly independent because they are orthonormal.
- They are in the column space, because $C(A)$ = set of all $A\vec{x}$ for $\vec{x}$ in $\mathbb{R}^n$.
- They are enough for a basis since $\dim C(A) = \text{rank } r$.

③ They are eigenvectors for AA^T!

same eigenvalue λ_i

$$AA^T\vec{u}_i = \frac{1}{\sigma_i} A\underbrace{A^TA\vec{v}_i}_{\sigma_i^2\vec{v}_i} = \sigma_i A\vec{v}_i = \sigma_i^2\left(\frac{A\vec{v}_i}{\sigma_i}\right) = \sigma_i^2 \vec{u}_i$$

Now remember one of the big theorems:

$$C(A)^{\perp} = N(A^T) \text{ in } \mathbb{R}^m$$

We can get an <u>orthonormal</u> <u>basis</u> of $\mathbb{R}^m$ by <u>combining</u>

$$\underbrace{\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_r\}}_{\substack{\text{orthonormal basis} \\ \text{of } C(A)\text{, also} \\ \text{eigenvectors for } AA^T}} \text{ with } \underbrace{\{\vec{u}_{r+1}, \dots, \vec{u}_m\}}_{\substack{\text{orthonormal basis of } N(A^T)\text{, same as} \\ N(\cancel{A^TA}\ AA^T)\text{, so also eigenvectors for } AA^T \\ \text{(with eigenvalue 0).}}}$$

<u>Conclusion</u>: For <u>any</u> $m \times n$ matrix A, we've shown that we can find <u>orthonormal</u> bases of <u>both</u> $\mathbb{R}^m$ and $\mathbb{R}^n$ that are "good for A":

$\mathbb{R}^m$: $\{\underbrace{\vec{u}_1, \dots, \vec{u}_r}_{AA^T \vec{u}_i = \sigma_i^2 \vec{u}_i}, \underbrace{\vec{u}_{r+1}, \dots, \vec{u}_m}_{\text{basis of } N(A^T)}\}$ orthonormal basis of eigenvectors for AA^T

$\mathbb{R}^n$: $\{\underbrace{\vec{v}_1, \dots, \vec{v}_r}_{A^TA \vec{v}_i = \sigma_i^2 \vec{v}_i}, \underbrace{\vec{v}_{r+1}, \dots, \vec{v}_n}_{\text{basis of } N(A)}\}$ orthonormal basis of eigenvectors for A^TA

Moreover, $\boxed{A\vec{v}_i = \sigma_i \vec{u}_i}$

We can use this equation to derive the <u>singular value decomposition</u> (SVD) of A:

Write: $V = \underbrace{\begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{bmatrix}}_{n \times n \text{ orthogonal matrix}}$, $U = \underbrace{\begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_m \end{bmatrix}}_{m \times m \text{ orthogonal matrix}}$, $\Sigma = \underbrace{\begin{bmatrix} \sigma_1 & & & & 0 \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & 0 & \\ 0 & & & & \ddots \; 0 \end{bmatrix}}_{m \times n \text{ "diagonal" matrix}}$

Now let's calculate:

$$AV = A\begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{bmatrix} = \begin{bmatrix} \underbrace{A\vec{v}_1 \; A\vec{v}_2 \; \cdots \; A\vec{v}_r}_{A\vec{v}_i = \sigma_i \vec{u}_i \text{ if } i \le r} & \underbrace{\vec{0} \; \cdots \; \vec{0}}_{A\vec{v}_i = \vec{0} \text{ if } i > r.} \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} \sigma_1\vec{u}_1 & \sigma_2\vec{u}_2 & \cdots & \sigma_r\vec{u}_r & \vec{0} & \cdots & \vec{0} \end{bmatrix}}_{}$$

Can write $\vec{0} = 0\vec{u}_{r+1}$, $\vec{0} = 0\vec{u}_{r+2}$ $\cdots$ $\vec{0} = 0\vec{u}_m$

m×n matrix; multiplying columns by σ_i = multiply on right with an m×n diagonal-like matrix

$$= \underbrace{\begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_m \end{bmatrix}}_{m\times m} \underbrace{\begin{bmatrix} \sigma_1 & & & & 0\text{'s} \\ & \sigma_2 & & & \\ & & \ddots & & \\ & & & \sigma_r & \\ 0\text{'s} & & & & 0 \cdots 0 \end{bmatrix}}_{m \times n} = U\Sigma$$

m×n (we will have columns of 0's on the right if $n>m$, and we will have rows of 0's at the bottom if $m>n$)

We calculated $AV = U\Sigma$. Also, V is orthogonal, $V^{-1} = V^T$.

So $\boxed{A = U\Sigma V^T}$ ← the singular value decomposition (SVD)

SVD shows that any m×n matrix A can be factored as:

(m×m orthogonal)(m×n diagonal-like)(n×n orthogonal)

↑ columns are the left singular vectors (orthonormal basis of eigenvectors for AA^T

↑ diagonal entries are the singular values (square roots of the positive real eigenvalues of AA^T and A^TA

↑ rows ~~columns~~ are the right singular vectors (orthonormal basis of eigenvectors for A^TA.

Example (Problem 7.2.4)

$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$ Let's find the SVD. First find eigenvalues and eigenvectors of $A^T A$.

$$A^T A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$ Eigenvalues: $\det(A^T A - \lambda I) =$

$$\begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = (2-\lambda)^2 - 1 = \lambda^2 - 4\lambda + 3 = (\lambda-3)(\lambda-1) = 0$$

So $\lambda_1 = 3 > \lambda_2 = 1 > 0 \rightsquigarrow \sigma_1 = \sqrt{3}, \sigma_2 = 1$ (square roots)

Eigenvectors for $\lambda = 3$: Solve $(A^T A - 3I)\vec{x} = \vec{0}$

$\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \rightsquigarrow x_1 = x_2 \rightsquigarrow \vec{x} = x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. The orthonormal basis vector $\vec{v}_1$ should be a unit vector: $\vec{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

Eigenvectors for $\lambda = 1$: Solve $(A^T A - I)\vec{x} = \vec{0}$.

$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \rightsquigarrow x_1 + x_2 = 0 \rightsquigarrow \vec{x} = x_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

$\vec{v}_2$ should be a unit vector: $\vec{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Note that $\vec{v}_1 \perp \vec{v}_2$, ~~should~~ so $\{\vec{v}_1, \vec{v}_2\}$ is indeed an orthonormal basis of $\mathbb{R}^2$. Get orthogonal matrix $V = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$

Left singular vectors $\vec{u}_1, \vec{u}_2, \vec{u}_3$

$$\vec{u}_1 = \frac{1}{\sigma_1} A\vec{v}_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

$$\vec{u}_2 = \frac{1}{\sigma_2} A\vec{v}_2 = \frac{1}{1} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

What about $\vec{u}_3$? It needs to a unit basis vector of $N(AA^T) = N(A^T)$

Solve $A^T\vec{x} = \vec{0}$: $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \xrightarrow{\text{Row 1} - \text{Row 2}} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix} \rightarrow \begin{matrix} x_1 = x_3 \\ x_2 = -x_3 \end{matrix}$

So $\vec{x} = \vec{x}_3 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$. $\vec{u}_3$ needs to be a <u>unit</u> vector: $\vec{u}_3 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$

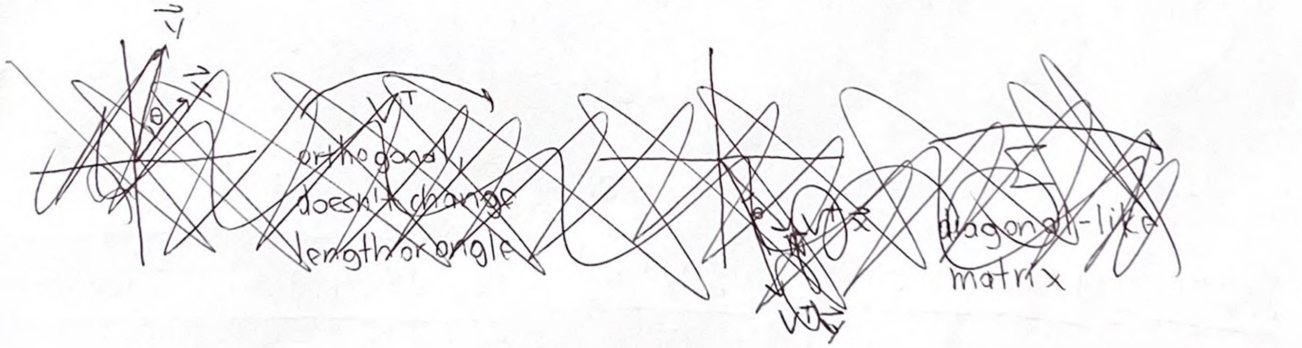
We can now write $U = \begin{bmatrix} 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \\ 2/\sqrt{6} & 0 & -1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \end{bmatrix}$ ← Note that the columns are $\perp$

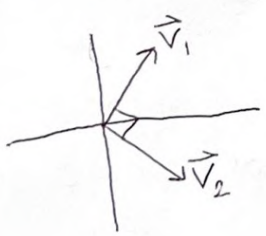
Finally, the SVD is $A = U\Sigma V^T = \begin{bmatrix} 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \\ 2/\sqrt{6} & 0 & -1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$

Check this is correct: $U\Sigma V^T = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 2/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} = A$ ✓

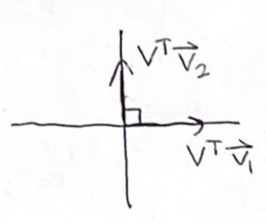
What are some things we can do with SVD?

<u>Geometry</u> $A = U\Sigma V^T$

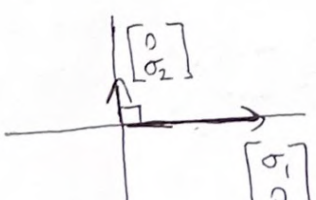




V^T, orthogonal, doesn't change length or angle



Σ, just scales the x and y-axis vectors



U, orthogonal, doesn't change length or angle