Review

Last time: we did the outer product form SVD

A: $m \times n$ of rank r

$$A = \sigma_1 u_1 v_1^T + \cdots + \sigma_r u_r v_r^T$$

- $\sigma_1 \geq \cdots \geq \sigma_r > 0$ are the singular values
- $\{v_1, \ldots, v_r\}$ is an orthonormal set in $\mathbb{R}^n$
  - called the right singular vectors
  - forms a basis for Row(A)
  - orthonormal eigenvectors of A^TA:

    $$A^TAv_i = \sigma_i^2 v_i$$
- $\{u_1, \ldots, u_r\}$ is an orthonormal set in $\mathbb{R}^m$
  - called the left singular vectors
  - forms a basis for Col(A)
  - orthonormal eigenvectors of AA^T:

    $$AA^Tu_i = \sigma_i^2 u_i$$

The singular vectors are related by

$$Av_i = \sigma_i u_i \qquad A^Tu_i = \sigma_i v_i$$

SVD of A^T is

$$A^T = \sigma_1 v_1 u_1^T + \cdots + \sigma_r v_r u_r^T$$

NB: If A is a wide matrix $(m<n)$ then

$$A^TA : n\times n \qquad AA^T : m\times m \leftarrow \text{smaller}$$

So it's easier to compute eigenvalues & eigenvectors of AA^T!

$$\boxed{\text{If } A \text{ is wide, compute the SVD of } A^T.}$$

Eg: $A = \begin{pmatrix} -10 & 10 & -10 & 10 \\ 10 & 5 & 10 & 5 \end{pmatrix}$

$$A^TA = \begin{pmatrix} 200 & -50 & 200 & -50 \\ -50 & 125 & -50 & 125 \\ 200 & -50 & 200 & -50 \\ -50 & 125 & -50 & 125 \end{pmatrix}$$

yikes!

Let's compute the SVD of A^T instead.

$$AA^T = \begin{pmatrix} 400 & -100 \\ -100 & 250 \end{pmatrix} \qquad p(\lambda) = (\lambda - 450)(\lambda - 200)$$

$$\lambda_1 = 450 \Rightarrow \sigma_1 = \sqrt{450} = 15\sqrt{2} \quad u_1 = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ -1 \end{pmatrix} \quad v_1 = \frac{1}{\sigma_1} A^T u_1 = \frac{1}{\sqrt{10}}\begin{pmatrix} -2 \\ 1 \\ -2 \\ 1 \end{pmatrix}$$

$$\lambda_2 = 200 \Rightarrow \sigma_2 = \sqrt{200} = 10\sqrt{2} \quad u_2 = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad v_2 = \frac{1}{\sigma_2} A^T u_2 = \frac{1}{\sqrt{10}}\begin{pmatrix} 1 \\ 2 \\ 1 \\ 2 \end{pmatrix}$$

$$\Rightarrow A^T = 15\sqrt{2}\, v_1 u_1^T + 10\sqrt{2}\, v_2 u_2^T$$

$$\Rightarrow A = 15\sqrt{2}\, u_1 v_1^T + 10\sqrt{2}\, u_2 v_2^T$$

u_i are right-singular vectors of $A^T \rightsquigarrow$ left-singular vectors of A

SVD: Matrix Form

Let A be an $m \times n$ matrix of rank r.

Then $\boxed{A = U \Sigma V^T}$ where:

(square with orthonormal columns)

- $U = \begin{pmatrix} | & & | \\ u_1 & \cdots & u_m \\ | & & | \end{pmatrix}$ is an $m \times m$ orthogonal matrix
- $V = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix}$ is an $n \times n$ orthogonal matrix
- $\Sigma = \begin{pmatrix} \sigma_1 & & & 0 & & \\ & \ddots & & & & \\ & & \sigma_r & & & \\ 0 & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$ is an $m \times n$ diagonal matrix.

$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ are the singular values

Where did $u_{r+1}, \ldots, u_m$ and $v_{r+1}, \ldots, v_n$ come from??

They're orthonormal bases for the other two fundamental subspaces!

$Col(A)$: $\{u_1, \ldots, u_r\}$ $\qquad$ $Nul(A^T)$: $\{u_{r+1}, \ldots, u_m\}$

$Row(A)$: $\{v_1, \ldots, v_r\}$ $\qquad$ $Nul(A)$: $\{v_{r+1}, \ldots, v_n\}$

Procedure to Compute $A = U\Sigma V^T$:

(1) Compute the singular values and singular vectors

$$\{v_1, \dots, v_r\} \quad \{u_1, \dots, u_r\} \quad \sigma_1, \dots, \sigma_r$$

as before

(2) Find orthonormal bases

$$\{u_{r+1}, \dots, u_m\} \text{ for } \mathrm{Nul}(A^T)$$

$$\{v_{r+1}, \dots, v_n\} \text{ for } \mathrm{Nul}(A)$$

using Gram-Schmidt.

(3) $$U = \begin{pmatrix} | & & | & | & & | \\ u_1 & \cdots & u_r & u_{r+1} & \cdots & u_m \\ | & & | & | & & | \end{pmatrix} \quad V = \begin{pmatrix} | & & | & | & & | \\ v_1 & \cdots & v_r & v_{r+1} & \cdots & v_n \\ | & & | & | & & | \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} \sigma_1 & & & & 0 \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & 0 & \\ 0 & & & & \ddots \ 0 \end{pmatrix}$$ (same size as A)

Proof: Use the outer product version of matrix mult:

$$U\Sigma V^T = \begin{pmatrix} | & & | \\ u_1 & \cdots & u_m \\ | & & | \end{pmatrix} \begin{pmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & 0 \end{pmatrix} \begin{pmatrix} - & v_1 & - \\ & \vdots & \\ - & v_n & - \end{pmatrix}$$

$$= \begin{pmatrix} | & & | \\ u_1 & \cdots & u_m \\ | & & | \end{pmatrix} \begin{pmatrix} - & \sigma_1 v_1 & - \\ & \vdots & \\ - & \sigma_r v_r & - \\ - & 0 & - \\ & \vdots & \\ - & 0 & - \end{pmatrix}$$

$$= \sigma_1 u_1 v_1^T + \cdots + \sigma_r u_r v_r^T + 0 + \cdots + 0 \quad \checkmark$$

Eg: $A = \begin{pmatrix} -10 & 10 & -10 & 10 \\ 10 & 5 & 10 & 5 \end{pmatrix}$

(1) $A = 15\sqrt{2}\, u_1 v_1^T + 10\sqrt{2}\, u_2 v_2^T$ for

$$u_1 = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ -1 \end{pmatrix} \qquad v_1 = \frac{1}{\sqrt{10}}\begin{pmatrix} -2 \\ 1 \\ -2 \\ 1 \end{pmatrix}$$

$$u_2 = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ 2 \end{pmatrix} \qquad v_2 = \frac{1}{\sqrt{10}}\begin{pmatrix} 1 \\ 2 \\ 1 \\ 2 \end{pmatrix}$$

(2) $Nul(A^T) = \{0\}$ because $r = m$

$Nul(A)$: $\begin{pmatrix} -10 & 10 & -10 & 10 \\ 10 & 5 & 10 & 5 \end{pmatrix} \overset{\text{rref}}{\rightsquigarrow} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$

$$\overset{\text{PVF}}{\rightsquigarrow} \text{Span}\left\{ \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

[already orthogonal — usually need Gram-Schmidt] $\overset{\text{normalize}}{\rightsquigarrow}$ $v_3 = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ $v_4 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}$

(3) So $A = U\Sigma V^T$ for

$$U = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} \qquad \Sigma = \begin{pmatrix} 15\sqrt{2} & 0 & 0 & 0 \\ 0 & 10\sqrt{2} & 0 & 0 \end{pmatrix}$$

$$V = \begin{pmatrix} -2/\sqrt{10} & 1/\sqrt{10} & -1/\sqrt{2} & 0 \\ 1/\sqrt{10} & 2/\sqrt{10} & 0 & -1/\sqrt{2} \\ -2/\sqrt{10} & 1/\sqrt{10} & 1/\sqrt{2} & 0 \\ 1/\sqrt{10} & 2/\sqrt{10} & 0 & 1/\sqrt{2} \end{pmatrix}$$

NB: $A = U\Sigma V^T$ contains full orthogonal diagonalizations of A^TA and of AA^T:

$$\boxed{A^TA = V\begin{pmatrix} \sigma_1^2 & & & & \\ & \ddots & & & \\ & & \sigma_r^2 & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix}V^T \qquad AA^T = U\begin{pmatrix} \sigma_1^2 & & & & \\ & \ddots & & & \\ & & \sigma_r^2 & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix}U^T}$$

It also contains orthonormal bases for all four subspaces:

$$U = \begin{pmatrix} \overbrace{\begin{matrix} | & & | \\ u_1 & \cdots & u_r \\ | & & | \end{matrix}}^{\text{o.n. basis for } \mathrm{Col}(A)} & \overbrace{\begin{matrix} | & & | \\ u_{r+1} & \cdots & u_m \\ | & & | \end{matrix}}^{\text{o.n. basis for } \mathrm{Nul}(A^T)} \end{pmatrix}$$

$i \le r$ $\boxed{Av_i = \sigma_i u_i \uparrow \ \downarrow A^T u_i = \sigma_i v_i}$ $\boxed{Av_i = 0 \uparrow \ \downarrow A^T u_i = 0}$ $i > r$

$$V = \begin{pmatrix} \underbrace{\begin{matrix} | & & | \\ v_1 & \cdots & v_r \\ | & & | \end{matrix}}_{\text{o.n. basis for } \mathrm{Row}(A)} & \underbrace{\begin{matrix} | & & | \\ v_{r+1} & \cdots & v_n \\ | & & | \end{matrix}}_{\text{o.n. basis for } \mathrm{Nul}(A)} \end{pmatrix}$$

The Pseudo-Inverse

This is a matrix A^+ that is the "best possible" substitute for A^{-1} when A is not invertible.

- Works for non-square matrices:
  if A is $m\times n$ then A^+ is $n\times m$
- A^+b is the shortest least-squares solution of $Ax=b$.

First let's do diagonal matrices.

Def: If Σ is an $m\times n$ diagonal matrix with nonzero diagonal entries $\sigma_1,\ldots,\sigma_r$, its pseudo-inverse Σ^+ is the $n\times m$ diagonal matrix with nonzero diagonal entries $\sigma_1^{-1},\ldots,\sigma_r^{-1}$.

$$\underset{4\times 5}{\Sigma=\begin{pmatrix}3&0&0&0&0\\0&2&0&0&0\\0&0&1&0&0\\0&0&0&0&0\end{pmatrix}} \rightsquigarrow \underset{5\times 4}{\Sigma^+=\begin{pmatrix}1/3&0&0&0\\0&1/2&0&0\\0&0&1&0\\0&0&0&0\\0&0&0&0\end{pmatrix}}$$

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

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

Now let's do general matrices.

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

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

The pseudo-inverse of A is the $n \times m$ matrix

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

This has the same singular vectors (switch right & left) and reciprocal singular values.

Eg: $A = \begin{pmatrix} -10 & 10 & -10 & 10 \\ 10 & 5 & 10 & 5 \end{pmatrix} = 15\sqrt{2}\, u_1 v_1^T + 10\sqrt{2}\, u_2 v_2^T$

for $\quad u_1 = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ -1 \end{pmatrix} \qquad v_1 = \frac{1}{\sqrt{10}}\begin{pmatrix} -2 \\ 1 \\ -2 \\ 1 \end{pmatrix}$

$\quad u_2 = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ 2 \end{pmatrix} \qquad v_2 = \frac{1}{\sqrt{10}}\begin{pmatrix} 1 \\ 2 \\ 1 \\ 2 \end{pmatrix}$

$$\rightsquigarrow A^+ = \frac{1}{15\sqrt{2}} v_1 u_1^T + \frac{1}{10\sqrt{2}} v_2 u_2^T$$

$$= \frac{1}{15\sqrt{2}} \cdot \frac{1}{\sqrt{10}}\begin{pmatrix} -2 \\ 1 \\ -2 \\ 1 \end{pmatrix} \cdot \frac{1}{\sqrt{5}}\begin{pmatrix} 2 & -1 \end{pmatrix} + \frac{1}{10\sqrt{2}} \cdot \frac{1}{\sqrt{10}}\begin{pmatrix} 1 \\ 2 \\ 1 \\ 2 \end{pmatrix} \cdot \frac{1}{\sqrt{5}}\begin{pmatrix} 1 & 2 \end{pmatrix}$$

$$= \frac{1}{150}\begin{pmatrix} -4 & 2 \\ 2 & -1 \\ -4 & 2 \\ 2 & -1 \end{pmatrix} + \frac{1}{100}\begin{pmatrix} 1 & 2 \\ 2 & 4 \\ 1 & 2 \\ 2 & 4 \end{pmatrix} = \frac{1}{300}\begin{pmatrix} -5 & 10 \\ 10 & 10 \\ -5 & 10 \\ 10 & 10 \end{pmatrix}$$

NB:

$$A = \begin{array}{c} \begin{array}{cc} \quad\mathrm{Col}(A) \qquad & \qquad \mathrm{Nul}(A^T) \end{array} \\ \begin{pmatrix} | & & | & | & & | \\ u_1 & \cdots & u_r & u_{r+1} & \cdots & u_m \\ | & & | & | & & | \end{pmatrix} \end{array} \begin{pmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & 0 \end{pmatrix} \begin{array}{c} \begin{array}{cc} \quad\mathrm{Row}(A) \qquad & \qquad \mathrm{Nul}(A) \end{array} \\ \begin{pmatrix} | & & | & | & & | \\ v_1 & \cdots & v_r & v_{r+1} & \cdots & v_n \\ | & & | & | & & | \end{pmatrix}^T \end{array}$$

$$A^+ = \begin{array}{c} \begin{array}{cc} \quad\mathrm{Col}(A^+) \qquad & \qquad \mathrm{Nul}(A^{+T}) \end{array} \\ \begin{pmatrix} | & & | & | & & | \\ v_1 & \cdots & v_r & v_{r+1} & \cdots & v_n \\ | & & | & | & & | \end{pmatrix} \end{array} \begin{pmatrix} \sigma_1^{-1} & & & \\ & \ddots & & \\ & & \sigma_r^{-1} & \\ & & & 0 \end{pmatrix} \begin{array}{c} \begin{array}{cc} \quad\mathrm{Row}(A^+) \qquad & \qquad \mathrm{Nul}(A^+) \end{array} \\ \begin{pmatrix} | & & | & | & & | \\ u_1 & \cdots & u_r & u_{r+1} & \cdots & u_m \\ | & & | & | & & | \end{pmatrix}^T \end{array}$$

This is almost the SVD of A^+ (the singular values are just ordered backward: $\sigma_1^{-1} \leq \cdots \leq \sigma_r^{-1}$).

So we see:

$$\mathrm{Col}(A) = \mathrm{Row}(A^+) = \mathrm{Row}(A^T)$$
$$\mathrm{Nul}(A^T) = \mathrm{Nul}(A^+) = \mathrm{Nul}(A^T)$$
$$\mathrm{Row}(A) = \mathrm{Col}(A^+) = \mathrm{Col}(A^T)$$
$$\mathrm{Nul}(A) = \mathrm{Nul}(A^{+T}) = \mathrm{Nul}(A^{TT})$$

NB: If A is invertible then $r=m=n$ and Σ is invertible, so $\Sigma^+ = \Sigma^{-1}$ and

$$AA^+ = (U\Sigma V^T)(V\Sigma^+ U^T)$$
$$= U\Sigma \overbrace{(V^TV)}^{=I_n} \Sigma^{-1} U^T = U \overbrace{\Sigma\Sigma^{-1}}^{=I_n} U^T = \overbrace{UU^T}^{=I_n} = I_n$$

$$\boxed{A \text{ is invertible} \iff A^{-1} = A^+}$$

So what are A^+A and AA^+ if A is not invertible?

Prop: $\boxed{\begin{array}{l} A^+A = \text{projection onto } \mathrm{Row}(A) \\ AA^+ = \text{projection onto } \mathrm{Col}(A) \end{array}}$

Proof: $AA^+ = (U\Sigma V^T)(V\Sigma^+U^T) = U\Sigma \underbrace{(V^TV)}^{=I_n}\Sigma^+U^T$

$= U\Sigma\Sigma^+U^T = U\begin{pmatrix} \sigma_1 & & & 0 \\ & \ddots & & \\ & & \sigma_r & \\ 0 & & & 0 \end{pmatrix}\begin{pmatrix} \sigma_1^{-1} & & & 0 \\ & \ddots & & \\ & & \sigma_r^{-1} & \\ 0 & & & 0 \end{pmatrix}U^T$

$= \begin{pmatrix} | & & | \\ u_1 & \cdots & u_m \\ | & & | \end{pmatrix}\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & 0 \end{pmatrix}\begin{pmatrix} - & u_1^T & - \\ & \vdots & \\ - & u_m^T & - \end{pmatrix} = \begin{pmatrix} | & & | \\ u_1 & \cdots & u_m \\ | & & | \end{pmatrix}\begin{pmatrix} - & u_1^T & - \\ & \vdots & \\ - & u_r^T & - \\ - & 0 & - \\ & \vdots & \\ - & 0 & - \end{pmatrix}$

$= u_1u_1^T + \cdots + u_ru_r^T$

This is the outer product formula for P_V, $V = \mathrm{Col}(A)$ because $\{u_1, \ldots, u_r\}$ is an orthonormal basis for $\mathrm{Col}(A)$.

A^+A: similar. ✓

Vector form: for $i \leq r$ we have

same singular vectors
reciprocal singular values

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

$$AA^+u_i = A\left(\frac{1}{\sigma_i}v_i\right) = \frac{1}{\sigma_i}Av_i = \frac{1}{\sigma_i}\cdot\sigma_iu_i = u_i$$

But for $i > r$ we have

$A^+Av_i = A^+\cdot 0 = 0 \quad (v_i \in \mathrm{Nul}(A))$

$AA^+u_i = A\cdot 0 = 0 \quad (u_i \in \mathrm{Nul}(A^T) = \mathrm{Nul}(A^+))$

NB
$v_i \in Row(A)$
$\Rightarrow P_{Row(A)} v_i = v_i$
$u_i \in Col(A)$
$\Rightarrow P_{Col(A)} u_i = u_i$

$i \leq r$

$$\boxed{\begin{array}{l} A^+ A v_i = v_i \\ A A^+ u_i = u_i \end{array}}$$

$i > r$

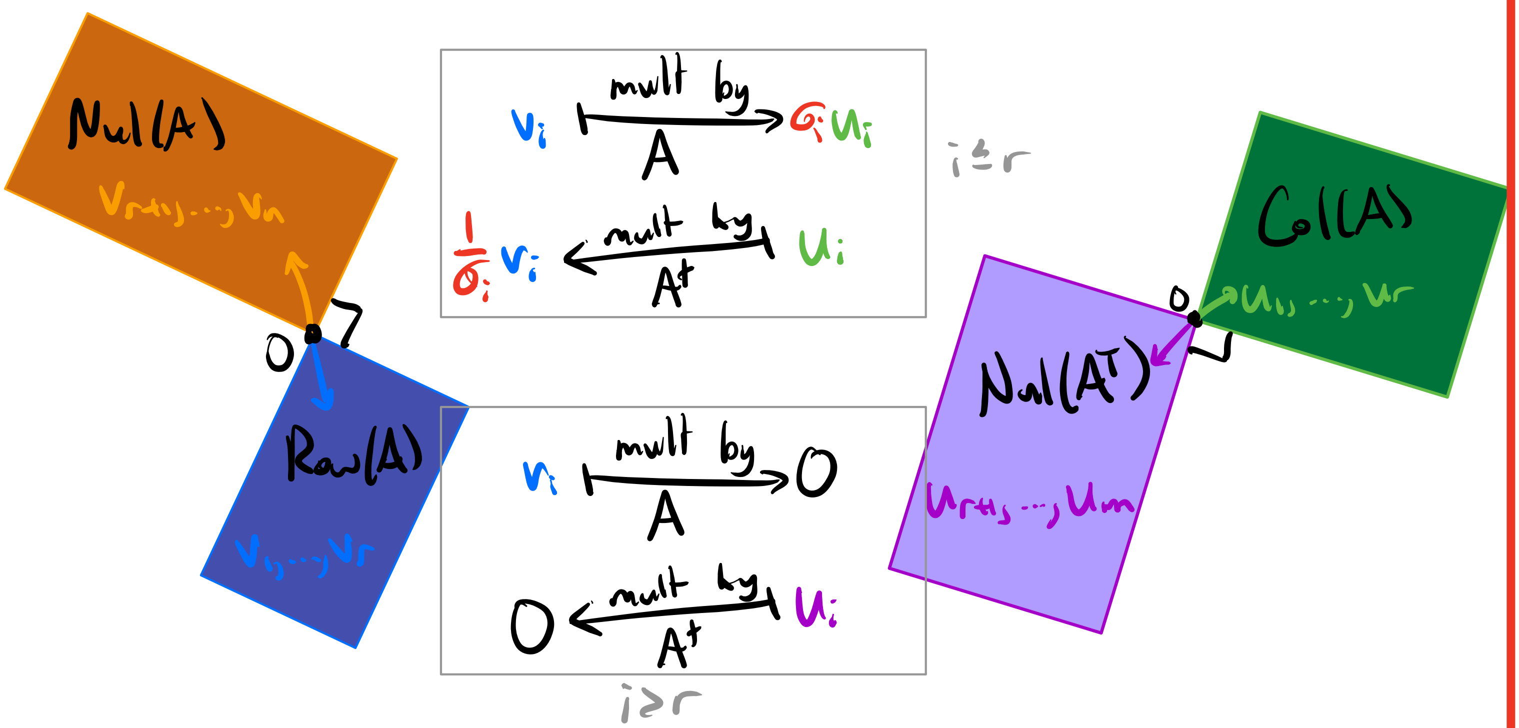
$$\boxed{\begin{array}{l} A^+ A v_i = 0 \\ A A^+ u_i = 0 \end{array}}$$

NB
$v_i \in Nul(A) = Row(A)^\perp$
$\Rightarrow P_{Row(A)} v_i = 0$
$u_i \in Nul(A^T) = Col(A)^\perp$
$\Rightarrow P_{Col(A)} u_i = 0$

The Big Picture Revisited

for an $m \times n$ matrix A of rank r

Row Picture: $\mathbb{R}^n$ Column Picture: $\mathbb{R}^m$



Recall: A projection matrix P_V is the identity matrix
$\iff V$ is all of $\mathbb{R}^n$

Consequence:

- $A^+A = I_n \iff A$ has full column rank
  $(Row(A) = Nul(A)^\perp = \{0\}^\perp = \mathbb{R}^n)$
- $AA^+ = I_m \iff A$ has full row rank
  $(Col(A) = \mathbb{R}^m)$

(matrix B with $BA = I_n$)

NB: This shows that:

- A has full column rank $\iff A$ admits a left inverse
- A has full row rank $\iff A$ admits a right inverse

(matrix B with $AB = I_m$)

[See HW5 for the "$\Leftarrow$" implications.)

Eg: $A = \begin{pmatrix} -10 & 10 & -10 & 10 \\ 10 & 5 & 10 & 5 \end{pmatrix}$ $\quad A^+ = \frac{1}{300}\begin{pmatrix} -5 & 10 \\ 10 & 10 \\ -5 & 10 \\ 10 & 10 \end{pmatrix}$

$$A^+A = \frac{1}{300}\begin{pmatrix} -5 & 10 \\ 10 & 10 \\ -5 & 10 \\ 10 & 10 \end{pmatrix}\begin{pmatrix} -10 & 10 & -10 & 10 \\ 10 & 5 & 10 & 5 \end{pmatrix} = \begin{pmatrix} 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \end{pmatrix}$$

$$AA^+ = \frac{1}{300}\begin{pmatrix} -10 & 10 & -10 & 10 \\ 10 & 5 & 10 & 5 \end{pmatrix}\begin{pmatrix} -5 & 10 \\ 10 & 10 \\ -5 & 10 \\ 10 & 10 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$(Col(A) = \mathbb{R}^2 \Rightarrow$ projection is $I_2)$

Now we can compute exactly what A^+b is:

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

Proof: First note $A\hat{x} = AA^+b =$ projection of b onto $\mathrm{Col}(A)$

$\Rightarrow \hat{x} = A^+b$ solves $A\hat{x} = b_{\mathrm{Col}(A)}$

$\Rightarrow \hat{x}$ is a least-squares solution of $Ax = b$.

Note
$$\hat{x} = \frac{1}{\sigma_1} v_1 u_1^T b + \cdots + \frac{1}{\sigma_r} v_r u_r^T b$$
$$= \frac{1}{\sigma_1}(u_1 \cdot b) v_1 + \cdots + \frac{1}{\sigma_r}(u_r \cdot b) v_r$$
$$\in \mathrm{Span}\{v_1, \dots, v_r\} = \mathrm{Row}(A).$$

$\boxed{\hat{x} \in \mathrm{Row}(A)}$

Any other solution $\hat{x}'$ has the form $\hat{x}' = \hat{x} + y$ for $y \in \mathrm{Nul}(A)$.

(The least-squares solutions are the solutions of $A\hat{x} = b_{\mathrm{Col}(A)}$.)

Note $y \in \mathrm{Row}(A)^\perp \Rightarrow \hat{x} \cdot y = 0$.

$$\|x'\|^2 = \|\hat{x} + y\|^2 = (\hat{x} + y) \cdot (\hat{x} + y) = \hat{x} \cdot \hat{x} + \cancel{2\hat{x} \cdot y}^{\,0} + y \cdot y$$

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

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

Eg: $A = \begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix} = 2\cdot \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ 1\end{pmatrix}\cdot\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\end{pmatrix}$

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

The shortest least-squares solution of $Ax = b = \begin{pmatrix}3\\ 1\end{pmatrix}$ is $\hat{x} = A^+ b = \frac{1}{4}\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}\begin{pmatrix}3\\ 1\end{pmatrix} = \begin{pmatrix}1\\ 1\end{pmatrix}$.

All other least-squares solutions differ by $\mathrm{Nul}(A) = \mathrm{Span}\left\{\begin{pmatrix}-1\\ 1\end{pmatrix}\right\}$.