18.06

Zack Remscrim

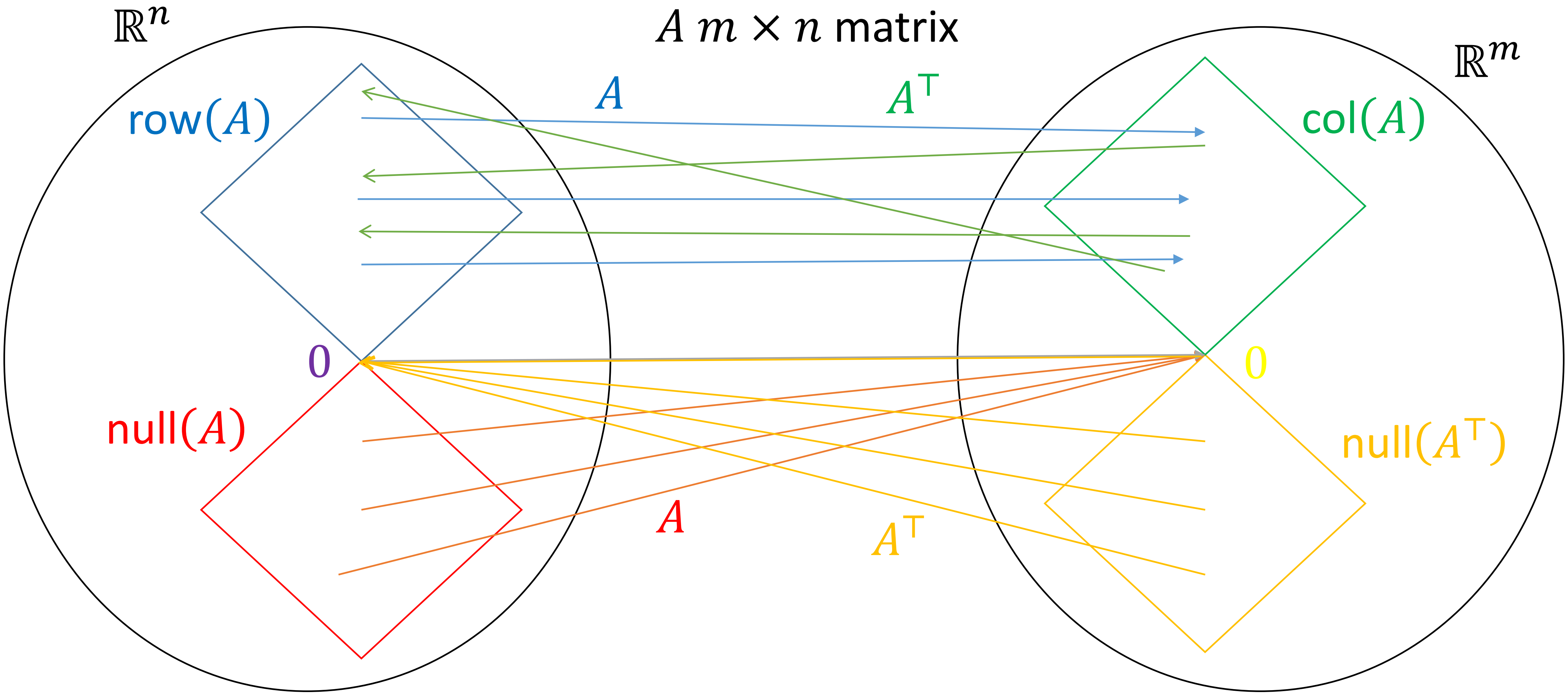
Four Fundamental Spaces

A $m \times n$ matrix

- $\text{col}(A)$
- $\text{row}(A)$
- $\text{null}(A)$
- $\text{null}(A^\top)$

Many questions about A can be answered by understanding these spaces

Four Fundamental Spaces



Rank Results

A $m \times n$ matrix, rank r

$A^\top$ also has rank r

A has (compact) SVD $U_1 \Sigma_r V_1^\top$

$A^\top$has (compact) SVD $(U_1 \Sigma_r V_1^\top)^\top$

$$= (V_1^\top)^\top \Sigma_r^\top U_1^\top$$

$$= V_1 \Sigma_r U_1^\top$$

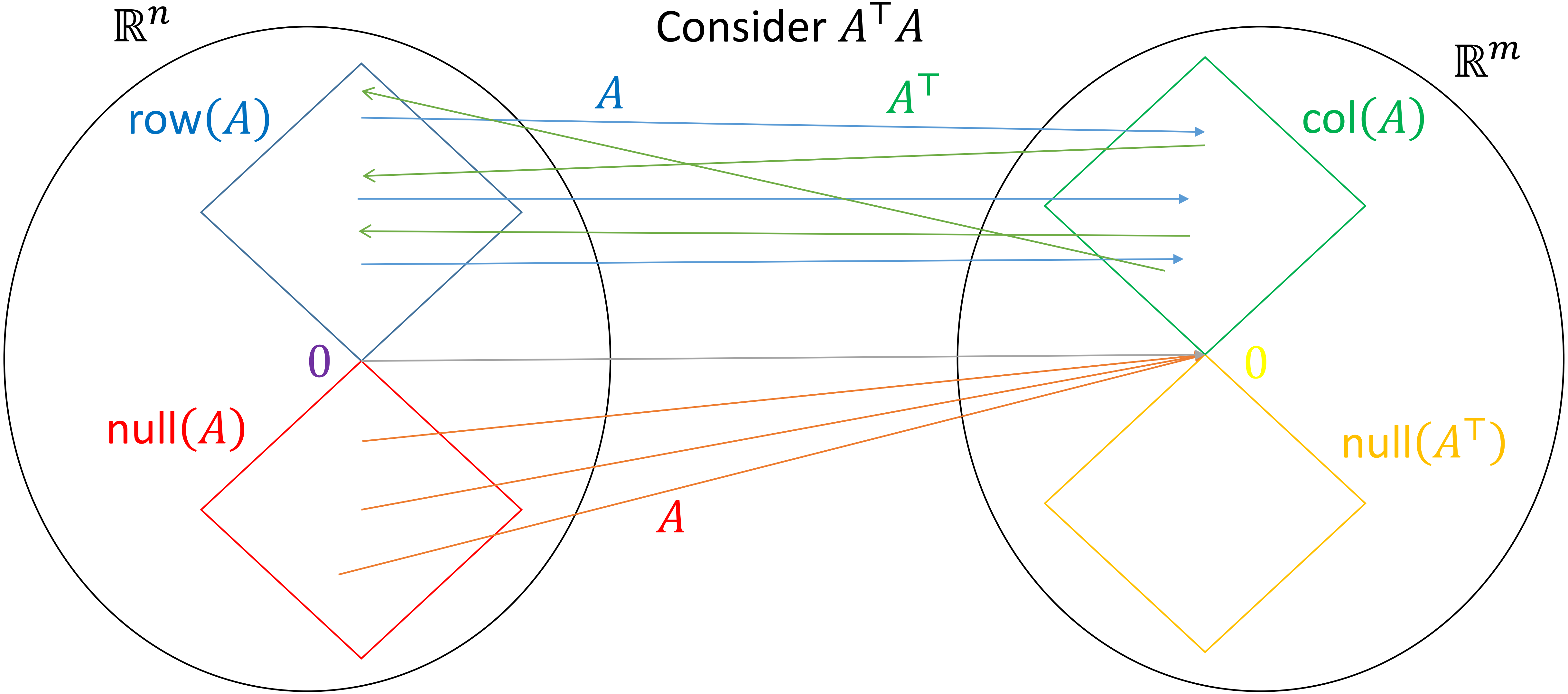
What about $A^\top A$?

Also has rank r

Can show using SVD, or Fundamental Spaces picture

Four Fundamental Spaces

Consider $A^\top A$



Orthogonal Spaces

Vector space V

We say vectors $u, w \in V$ are *orthogonal* when $u \cdot w = 0$

Given subspaces R, S of V, say R, S are *orthogonal* when

For every $r \in R$ and $s \in S$

$$r \cdot s = 0$$

For R subspace of V, $R^{\perp}$ is space of all $v \in V$ where

$r \cdot v = 0$ For every $r \in R$

Orthogonal Spaces

Example: In $\mathbb{R}^3$, xy-plane is orthogonal to z-axis

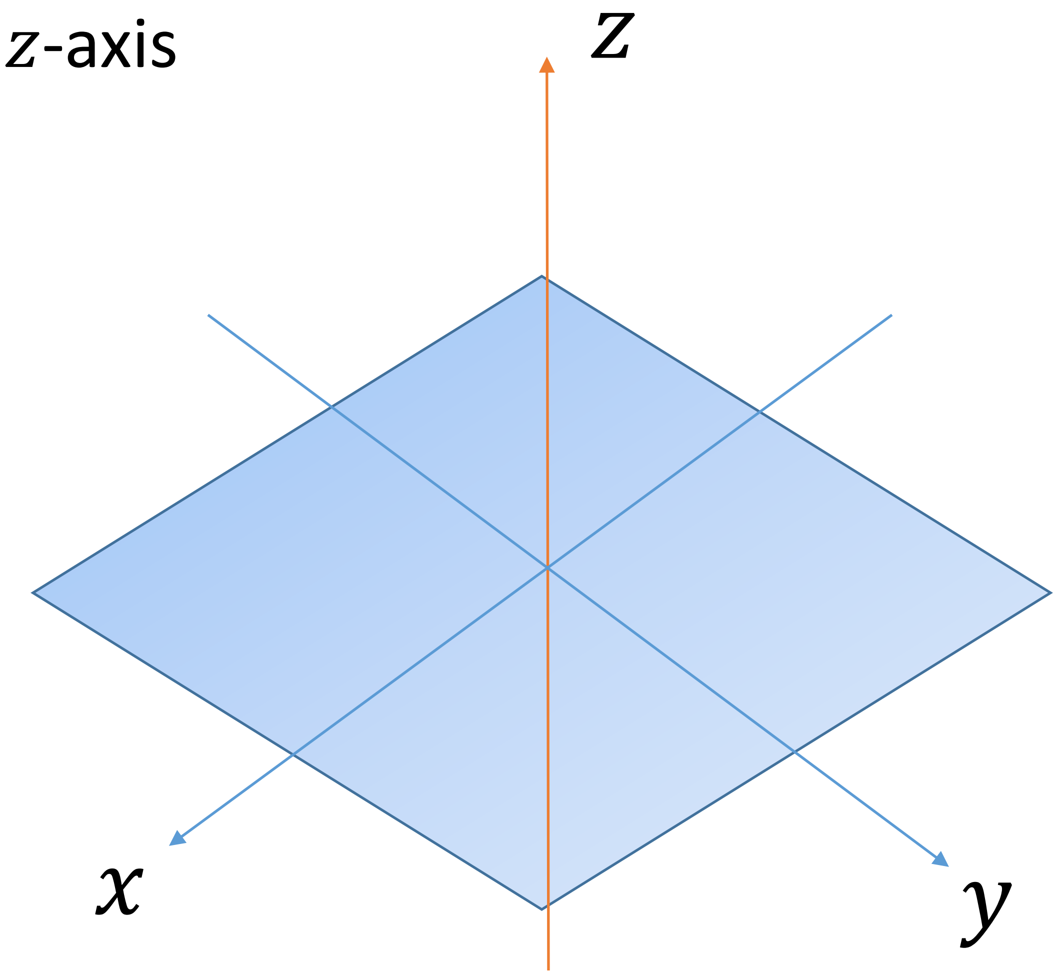
R: xy-plane

S: z-axis

$r = (a, b, 0) \in R$
$s = (0,0, c) \in S$

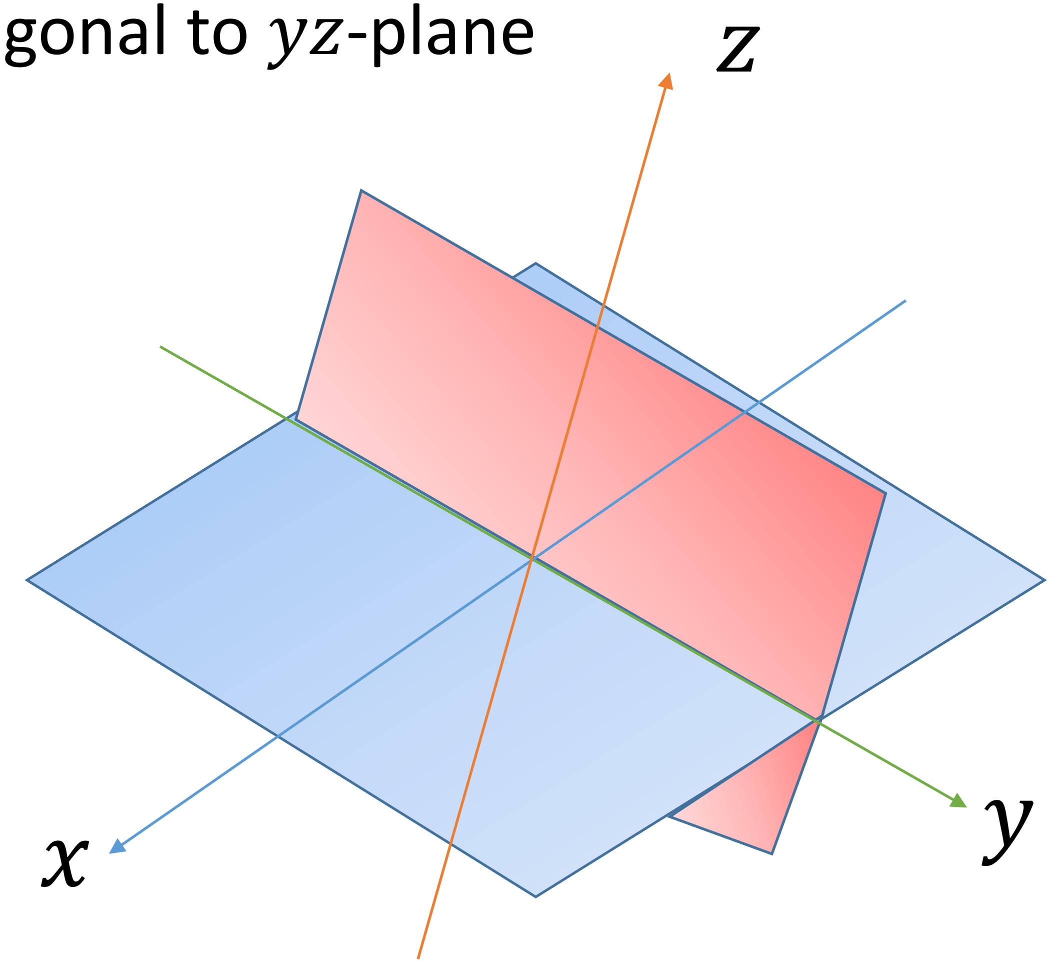
$r \cdot s = 0$

$R^\perp = S$



Orthogonal Spaces

Non-Example: In $\mathbb{R}^3$, xy-plane is *not* orthogonal to yz-plane



$r = (0,1,0) \in xy$-plane

$s = (0,1,0) \in yz$-plane

$r \cdot s = 1$

Four Fundamental Spaces

$\text{row}(A)^\perp = \text{null}(A)$

From SVD:

$\text{row}(A) = \text{col}(V_1)$

$\text{null}(A) = \text{col}(V_2)$

$\text{col}(A)^\perp = \text{null}(A^\top)$

From SVD:

$\text{col}(A) = \text{col}(U_1)$

$\text{null}(A^\top) = \text{col}(U_2)$

Matrices as Functions

A $m \times n$ matrix is a function

input: $x \in \mathbb{R}^n$

output: $y \in \mathbb{R}^m$

$y = Ax$

A is linear:

For any $x_1, x_2 \in \mathbb{R}^n$

and any $c_1, c_2 \in \mathbb{R}$

$$A(c_1x_1 + c_2x_2) = c_1A(x_1) + c_2A(x_2)$$

Linear Transformations

For V, W vector spaces (over $\mathbb{R}$)

Say a function T from V to W is linear if

For all $x_1, x_2 \in V$

and all $c_1, c_2 \in \mathbb{R}$

$$T(c_1 x_1 + c_2 x_2) = c_1 T(x_1) + c_2 T(x_2)$$

In some sense nothing new

Value in abstraction

Examples of Linear Transformations

$T: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$

$v = (x, y) \in \mathbb{R}^2$

$T(v) = (x, 0)$

Projection onto x-axis

$\frac{d}{dx}: V \longrightarrow W$

V space of polynomials in x of degree at most 4

W space of polynomials in x of degree at most 3

For $f(x) \in V$

$$\frac{d}{dx}(f(x)) = f'(x)$$

Non-Example of Linear Transformations

$T: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$

$$v = (x, y) \in \mathbb{R}^2$$

$$T(v) = (x^2 + y^2, x^2 - y^2)$$

$$T(cv) = ((cx)^2 + (cy)^2, (cx)^2 - (cy)^2)$$

$$= c^2(x^2 + y^2, x^2 - y^2)$$

$$= c^2 T(v)$$

Another Example of a Linear Transformation

$\nabla: V \longrightarrow W$

V space of differentiable functions $f: \mathbb{R}^n \longrightarrow \mathbb{R}$

W space of functions $h: \mathbb{R}^n \longrightarrow \mathbb{R}^n$

For $f(x) \in V$

$$\nabla\big(f(x)\big) = \left(\frac{\partial}{\partial x_1} f(x), \frac{\partial}{\partial x_2} f(x), \dots, \frac{\partial}{\partial x_n} f(x)\right)$$

$f(x + \Delta x) \approx f(x) + \nabla f(x)^{\top} \Delta \mathrm{x}$

Minimization Problems

Consider $Ax = b$

May or may not have a solution

If no solution, what is best approximate solution?

Find x such that $||Ax - b||^2$ as small as possible

Earlier, used QR decomposition

Now (matrix) calculus

Or using subspaces

Minimization: Matrix Calculus

Minimize $f(x) = \|Ax - b\|^2 = (Ax - b)^\top (Ax - b)$

Idea: Compute $\nabla f(x)$, equate to 0, and solve

$$
\begin{aligned}
df &= (Adx)^\top (Ax - b) + (Ax - b)^\top (Adx) \\
&= (Adx)^\top (Ax - b) + (Adx)^\top (Ax - b) \\
&= 2(Adx)^\top (Ax - b) \\
&= 2(dx)^\top A^\top (Ax - b) \\
&= \left(2A^\top (Ax - b)\right)^\top dx
\end{aligned}
$$

$$
\begin{aligned}
df &= \left(\nabla f(x)\right)^\top dx \\
\nabla f(x) &= 2A^\top (Ax - b)
\end{aligned}
$$

Minimization: Matrix Calculus

Minimize $f(x) = \|Ax - b\|^2 = (Ax - b)^\top(Ax - b)$

Idea: Compute $\nabla f(x)$, equate to 0, and solve

$$\nabla f(x) = 2A^\top(Ax - b)$$

Solve $A^\top(Ax - b) = 0$

$$A^\top Ax = A^\top b$$

If $A^\top A$ invertible,

$$x = (A^\top A)^{-1}A^\top b$$

What if $A^\top A$ not invertible?

Does $A^\top(Ax - b) = 0$ always have a solution?

The Equation $A^\top(Ax - b) = 0$

Does this equation always have a solution?

Yes!

Boring case: if $b \in \mathrm{col}(A)$

Any solution to $Ax = b$ works

Interesting case: if $b \notin \mathrm{col}(A)$

$\tilde{b} \in \mathrm{col}(A)$ projection of b onto $\mathrm{col}(A)$

$\tilde{b} - b$ orthogonal to $\mathrm{col}(A)$

$\tilde{b} - b \in \mathrm{null}(A^\top)$

$A^\top(\tilde{b} - b) = 0$

Any solution to $Ax = \tilde{b}$ works

Minimization: Matrix Calculus

Minimize $f(x) = \|Ax - b\|^2 = (Ax - b)^\top (Ax - b)$

Idea: Compute $\nabla f(x)$, equate to 0, and solve

$\nabla f(x) = 2A^\top (Ax - b)$

Solve $A^\top (Ax - b) = 0$

$\tilde{b}$ projection of b onto col(A)

Any solution to $Ax = \tilde{b}$ works

Can find $\tilde{b}$ and solution to $Ax = \tilde{b}$ with SVD

Minimization: Subspaces

Minimize $f(x) = \|Ax - b\|^2$

Idea: For any x, if $y = Ax$, then $y \in \text{col}(A)$

and if $y \in \text{col}(A)$, $y = Ax$ for some x

Minimum occurs when $y \in \text{col}(A)$ as close to b as possible

By definition, when $y = \tilde{b}$

So any solution to $Ax = \tilde{b}$ works

More Linear Transformations

V is space of $m \times n$ matrices

$T_K : V \longrightarrow \mathbb{R}$

K single $m \times n$ matrix

For $A \in V$,

$$T_K(\mathrm{A}) = \mathrm{sum}(K.* A)$$
$$= \sum_{i,j} K_{i,j} A_{i,j}$$

$\mathrm{tr} : V \longrightarrow \mathbb{R}$

$$\mathrm{tr}(\mathrm{A}) = \sum_i A_{i,i}$$

$T_K(\mathrm{A}) = \mathrm{tr}(\mathrm{K}^\top \mathrm{A})$

More Matrix Calculus

Previously: considered $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$

$$f(x + \Delta x) \approx f(x) + \mathrm{g}^\top \Delta \mathrm{x}$$

$$g = \nabla f(x)$$

$$df = g^\top dx$$

Now: consider $f\colon \mathbb{R}^{m \times n} \longrightarrow \mathbb{R}$

$$f(A + \Delta A) \approx f(A) + \mathrm{tr}(G^\top \Delta \mathrm{A})$$

$$dA = \mathrm{tr}(G^\top dA)$$