CONVEX OPTIMIZATION:

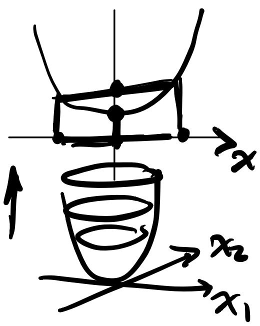
$$\min_{x \in \mathbb{R}^n} \; f(x) \quad \text{s.t.} \quad x \in \mathcal{X}$$

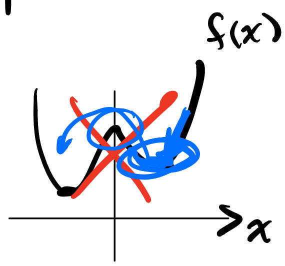
$f: \mathbb{R}^n \to \mathbb{R}$

f: convex/concave function

$\mathcal{X}$: convex set

Convex function
"bowl shaped"





$x, y \in \mathcal{X}$

$$f((1-\alpha)x+\alpha y) \le (1-\alpha)f(x)+\alpha f(y)$$

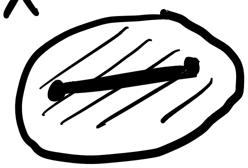
$0 \le \alpha \le 1$

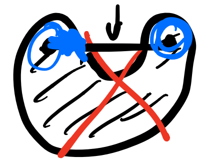
for a linear f.

$$f(\underline{(1-\alpha)x+\alpha y}) = \underline{(1-\alpha)f(x)+\alpha f(y)}$$

convex set

$\mathcal{X}$



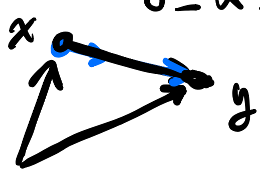


$\mathcal{X}$ convex if.

$x, y \in \mathcal{X}$

$\Rightarrow \boxed{(1-\alpha)x+\alpha y} \in \mathcal{X}$

$0 \le \alpha \le 1$ $\qquad \alpha \in \mathbb{R}$



$(1-\alpha)x+\alpha y$

$x + \underline{\alpha(y-x)}$

$f(x) = r^T x$ linear function

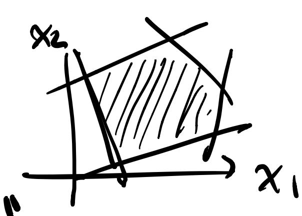
"linear program"

$f(x) = x^T Q x + r^T x$ Quadratic function

"quadratic program"

Polytopes

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



$$\boxed{\begin{aligned} Ax &= b \\ Cx &\le d \end{aligned}}$$

Applications

- stock portfolio optimization
- regression models
- robust control & estimation
- trajectory planning. rocket landings Behcet
- shortest path 1
  planning
  network routing
- Markov decision processes (MDPs) 2

"convex optimization" = "easily solvable"

Convex relaxations:

- combinatorial opt.
- classification
  image recognition
  low rank matrices/sparsity

Matrix Cookbook

Two books:

Convex Optimization ←
Stephen Boyd & Vanderberghe
(Readable)

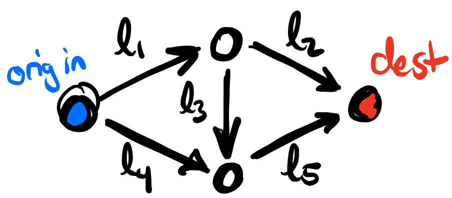
Convex Analysis ←
Rockafellar
(BIBLE)

Software:

Matlab: CVX   Python cvxpy

Preview: Network flow & MDPs

① Shortest Path



$G = (\mathcal{S}, \mathcal{E})$ $\searrow E \in \mathbb{R}^{|S| \times |E|}$ $\quad S \in \mathbb{R}^{|S|}$

$m = 1$ pop. mass

x_e: flow on edge e

①

$$\min_{x \in \mathbb{R}^{|E|}} l^T x$$

s.t. $Ex = Sm$, $x \geq 0$ $\Longleftrightarrow$

⊛

↓ dual

②

$$\max_{\lambda, v, \mu} v^T S m$$

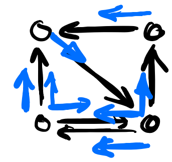
s.t. $\lambda \mathbb{1}^T = v^T E + \mu^T$, $\mu^T \geq 0$

v: value function
distance destination

μ: inefficiency on a particular edge

λ: total travel time

Markov Decision Process



S: states
A: actions (determine how transitions)
P: transition kernel
policy: when to choose different actions

①

$$\max_{y \in \mathbb{R}^{|A|}} r^T y$$

policy $\rightarrow y$ ✓ trans. kernel

s.t. $E_S y = P y$, $\mathbb{1}^T y = 1$ $y \geq 0$

stochastic local flow const.

global mass const.

dual ↓

② Bellman eqn ↓

$$\min_{\lambda, v, \mu} \lambda$$

s.t. $\lambda \mathbb{1}^T = r^T + v^T (E_S - P) + \mu^T$

$\mu \geq 0$

λ: total reward

q value ← v: value function

qlearning μ: inefficiency

Linear Algebra Review:

$x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times n}$ (in general)

Inner Products (Dot products)

$x, y \in \mathbb{R}^n$ $\quad y \cdot x = \langle y, x \rangle = y^T x = \sum_i y_i x_i$

Norm:

"length" 2-norm $\quad |x|_2 = \left(\sum_i |x_i|^2\right)^{1/2} = \sqrt{x^T x}$

Note: $|x| = |x|_2$

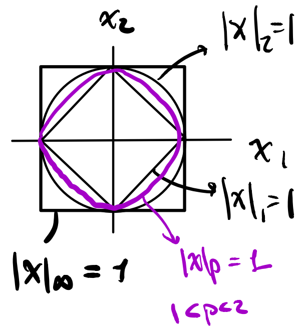
p-norm $\quad |x|_p = \left(\sum_i |x_i|^p\right)^{1/p}$

$1 \le p < \infty$

$p = 1 \quad |x|_1 = \left(\sum_i |x_i|\right)$

$\vdots$

$p = \infty \quad |x|_\infty = \max_i |x_i|$

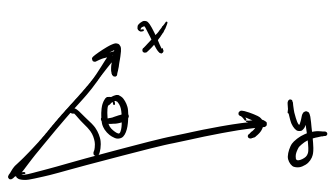


Unit Vector:

vector length 1 $\quad \dfrac{x}{|x|}$

Geometric defn of inner product:

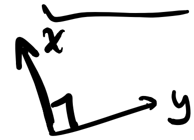
$$y^T x = |x||y|\cos\theta$$



$\theta = 0 \quad y^T x = |x||y|$

$\theta = \pi \quad y^T x = -|x||y|$

$\theta = \pi/2 \quad y^T x = 0$



$y^T x = 0$
orthogonal

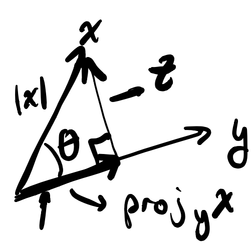
Projections

$$\text{proj}_y x = |x|\cos\theta \frac{y}{|y|} \frac{|y|}{|y|}$$

$$= \frac{y}{|y|^2} y^T x$$

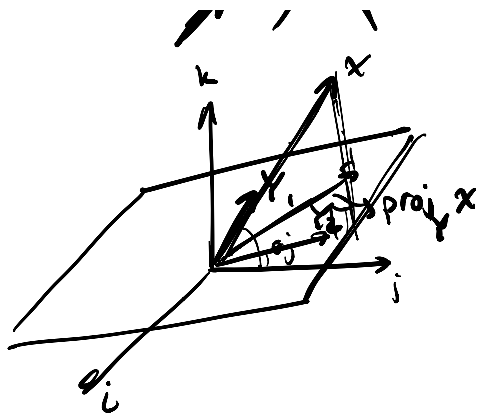
$$= \left[\frac{1}{|y|^2} y y^T\right] x$$

projection matrix



$$z = x - \text{proj}_y x$$

$$= \left[I - \frac{1}{|y|^2} y y^T\right] x$$



$$Y = [Y_1 \; Y_2]$$

$$\text{proj}_Y x = Y(Y^TY)^{-1}Y^Tx$$

Normal:

$$\text{proj}_Y(\text{proj}_Y x) = \text{proj}_Y x$$

$$Y(Y^TY)^{-1}Y^T\left(Y(Y^TY)^{-1}Y^T\right)$$

$$Y \; I \; (Y^TY)^{-1}Y^T$$

Block Matrix Multiplication

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} & \cdots & b_{1p} \\ b_{n1} & \cdots & b_{np} \end{bmatrix}$$

$$AB = \begin{bmatrix} a_{11}b_{11} + \cdots + a_{1n}b_{n1} & \cdots & a_{11}b_{1p} + \cdots + a_{1n}b_{np} \\ a_{m1}b_{11} + \cdots + a_{mn}b_{n1} & \cdots & \end{bmatrix}$$

$A_{11} \in \mathbb{R}^{m_1 \times n_1}$ $\quad B_{11} \in \mathbb{R}^{n_1 \times p_1}$

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ A_{m1} & \cdots & A_{mn} \end{bmatrix} \quad B = \begin{bmatrix} B_{11} & B_{1p} \\ B_{n1} & B_{np} \end{bmatrix}$$

$$AB = \begin{bmatrix} A_{11}B_{11} + \cdots + A_{1n}B_{n1} \\ \vdots \end{bmatrix}$$

$$A = \begin{bmatrix} | & & | \\ A_1 & \cdots & A_n \\ | & & | \end{bmatrix} = \begin{bmatrix} - \bar{A}_1^T - \\ \vdots \\ - \bar{A}_m^T - \end{bmatrix} \quad B = [B_1 \cdots B_p] = \begin{bmatrix} - \bar{B}_1^T - \\ \vdots \\ - \bar{B}_n^T - \end{bmatrix}$$

$$AB = \begin{bmatrix} - \bar{A}_1^T - \\ \vdots \\ - \bar{A}_m^T - \end{bmatrix} [B_1 \cdots B_p] = \begin{bmatrix} \bar{A}_1^T B_1 & \cdots & \bar{A}_1^T B_p \\ \bar{A}_m^T B_1 & \cdots & \bar{A}_m^T B_p \end{bmatrix} \rightarrow \text{matrix of inner products}$$

$$AB = \begin{bmatrix} | & & | \\ A_1 & \cdots & A_n \\ | & & | \end{bmatrix} \begin{bmatrix} - & \bar{B}_1^T & - \\ & \vdots & \\ - & \bar{B}_n^T & - \end{bmatrix} = \underline{A_1 \bar{B}_1^T} \pm \cdots \pm A_n \bar{B}_n^T$$

$\Leftarrow$ dyadic expansion

$$\overset{q}{[y_1 \cdots y_n]} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \underline{y_1 x_1 + \cdots + y_n x_n}$$

$$m \left[\begin{bmatrix} A_1 \end{bmatrix} \overset{p}{\begin{bmatrix} - & \bar{B}_1^T & - \end{bmatrix}} \right]$$

$m \times p$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \begin{bmatrix} x_1 & x_2 \end{bmatrix} = \begin{bmatrix} y_1 x_1 & y_1 x_2 \\ y_2 x_1 & y_2 x_2 \end{bmatrix}$$

$$A \begin{bmatrix} | & & | \\ B_1 & \cdots & B_p \end{bmatrix} = \begin{bmatrix} AB_1 & AB_2 & \cdots & AB_p \end{bmatrix}$$

$$Ax = \begin{bmatrix} - & \bar{A}_1^T & - \\ & \vdots & \\ - & \bar{A}_m^T & - \end{bmatrix} x = \begin{bmatrix} \bar{A}_1^T x \\ \vdots \\ \bar{A}_m^T x \end{bmatrix} \leftarrow$$

inner product of x w/ ea. row of A.

$$\boxed{Ax = \begin{bmatrix} A_1 & \cdots & A_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \underset{q}{A_1} \underset{q}{x_1} + \cdots + \underset{q}{A_n} \underset{q}{x_n}}$$

$$\begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

Linear Combinations

$y \in \mathbb{R}^m$ lin comb of $[A_1 \cdots A_n]$

if $y = A_1 x_1 + \cdots + A_n x_n$ $\quad x \in \mathbb{R}^n$

if $\exists\, x$ s.t. $y = Ax$ for $A = [A_1 \cdots A_n]$

Span of $[A_1 \cdots A_n] = A$

$$\text{span of cols of } A = \{ y \in \mathbb{R}^m \mid y = Ax,\ x \in \mathbb{R}^n \}$$

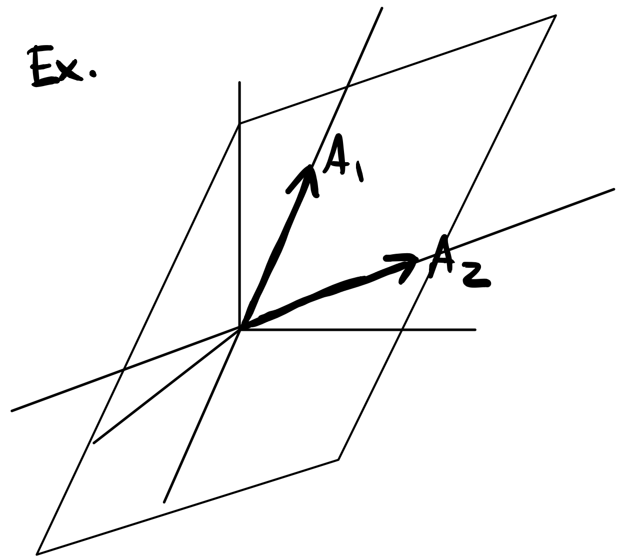
$$\text{range}(A) = \mathcal{R}(A) = \{ y \in \mathbb{R}^m \mid y = Ax,\ x \in \mathbb{R}^n \}$$

$\mathcal{R}(A_1) = ?$

$A = [A_1\ A_2]$

$\mathcal{R}(A) = ?$

Ex.



<u>Basis</u> for $\mathbb{R}^n$

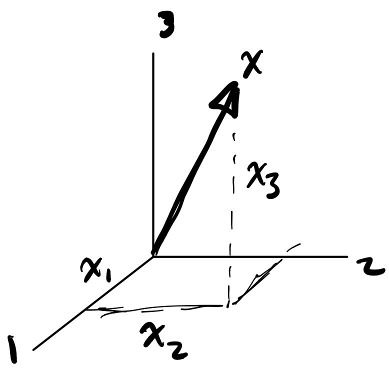
set of vectors that

- span the whole space
- lin independent

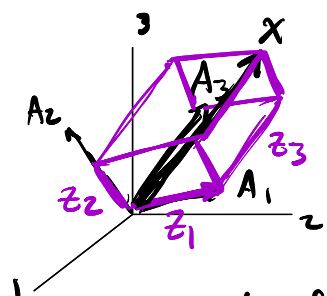
use a basis to represent other vectors

ex. standard basis

$$\begin{bmatrix} 1 & 0 & & 0 \\ 0 & 1 & & \vdots \\ \vdots & \vdots & & 0 \\ 0 & 0 & & 1 \end{bmatrix}$$



ex. basis $A = [A_1 A_2 A_3]$



what are coords of x w.r.t. the A basis

$$x = Az = A_1 z_1 + A_2 z_2 + A_3 z_3$$

$$z = \begin{bmatrix} 1 \\ 1/2 \\ 1 \end{bmatrix}$$

$$\underline{x = Az} \Rightarrow \underline{z = A^{-1}x}$$

$A = \begin{bmatrix} A_1 & A_2 & A_3 \end{bmatrix}$ $\quad |A_1| = 1 \quad |A_2| = 1 \quad |A_3| = 1$

$$z \stackrel{?}{=} A^T x = \begin{bmatrix} A_1^T \\ A_2^T \\ A_3^T \end{bmatrix} x = \begin{bmatrix} A_1^T x \\ A_2^T x \\ A_3^T x \end{bmatrix}$$

$A^T = A^{-1}$?

if $\underline{A^{-1} = A^T}$: A orthonormal transform $\quad$ rotation/reflection

all cols of A are orthogonal to ea. other and have length 1.

$A^{-1}A = I \qquad A^T A = I$

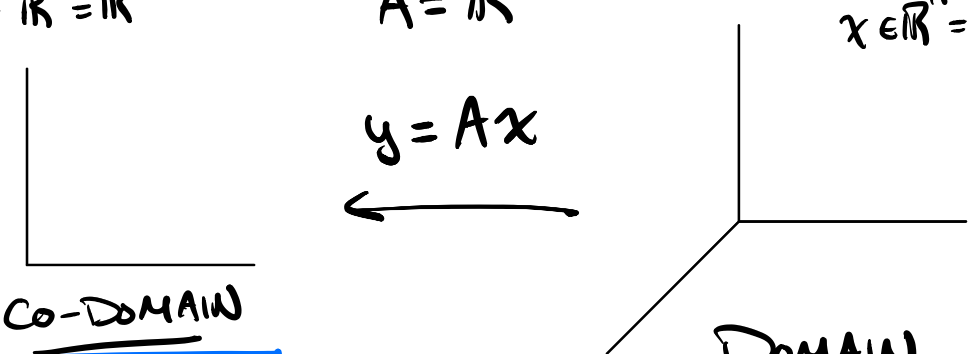
$$\begin{bmatrix} A_1^T \\ A_n^T \end{bmatrix} \begin{bmatrix} A_1 & \cdots & A_n \end{bmatrix} = \begin{bmatrix} A_1^T A_1 & & A_1^T A_n \\ & \ddots & \\ A_n^T A_1 & & A_n^T A_n \end{bmatrix} = I$$

$\underline{A_i^T A_j = 0} \quad i \neq j$

$x \in \mathbb{R}^n, y \in \mathbb{R}^m \qquad A \in \mathbb{R}^{m \times n}$

$y = Ax$

$y \in \mathbb{R}^m = \mathbb{R}^2 \qquad A = \mathbb{R}^{2\times 3} \qquad x \in \mathbb{R}^n = \mathbb{R}^3$



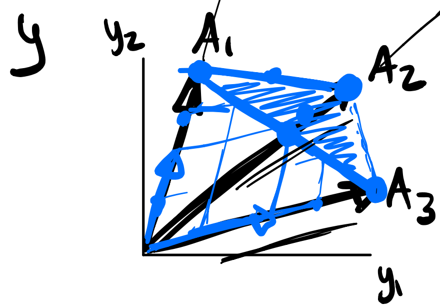
$y = Ax$

CO-DOMAIN $\qquad$ DOMAIN

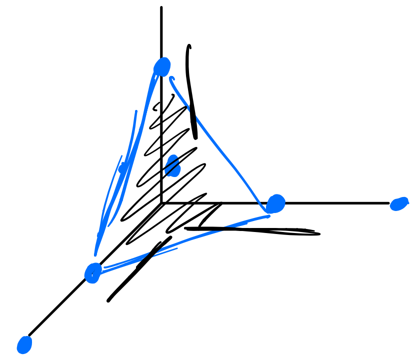
$\boxed{R(A): \text{subspace}}$

Column perspective:

$$A = [A_1 A_2 A_3]$$



$\xleftarrow{A}$



$x = \begin{bmatrix}1\\0\\0\end{bmatrix}$ $\quad x = \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix} \begin{bmatrix}1/2\\0\\1/2\end{bmatrix} \begin{bmatrix}1/2\\1/2\\0\end{bmatrix} \begin{bmatrix}1/3\\1/3\\1/3\end{bmatrix}$

$\begin{bmatrix}2\\0\\0\end{bmatrix} \begin{bmatrix}0\\2\\0\end{bmatrix}$ $\quad \frac{3}{4}A_1 + \frac{3}{4}A_3$

Convex hull or convex of cols of A $\quad$ elementwise

$$\Delta(A) = A\Delta = \{y \in \mathbb{R}^m \mid y = Ax, \sum_i x_i = 1, x \geq 0\}$$

simplex:

$$\Delta_n = \{x \in \mathbb{R}^n \mid \sum_i x_i = 1, x \geq 0\}$$

$\mathbb{1} = \begin{bmatrix}1\\\vdots\\1\end{bmatrix}$ $\quad \sum_i x_i = \mathbb{1}^T x$ $\quad [1 \cdots 1]\begin{bmatrix}x_1\\\vdots\\x_n\end{bmatrix}$

Row perspective

$A = [A_1 A_2 A_3]$

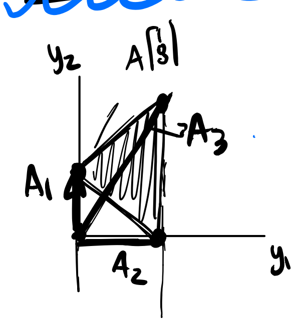
$A = \begin{bmatrix}- \bar{A}_1^T -\\ - \bar{A}_2^T -\end{bmatrix}$

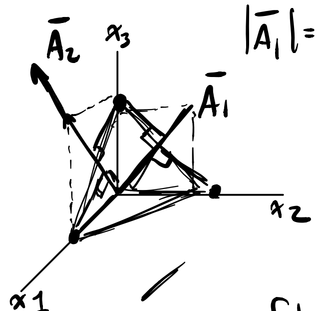
$y_1 = |\bar{A}_1| \frac{\bar{A}_1^T}{|\bar{A}_1|} x$

$y_2 = |\bar{A}_2| \frac{\bar{A}_2^T}{|\bar{A}_2|} x$

$|\bar{A}_2| = 2$

$|\bar{A}_1| = 1$



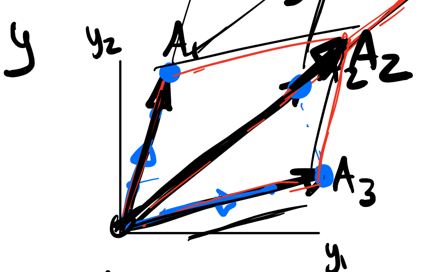


$x = \begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}$

Nullspace $\mathcal{N}(A) \subseteq$ DOMAIN

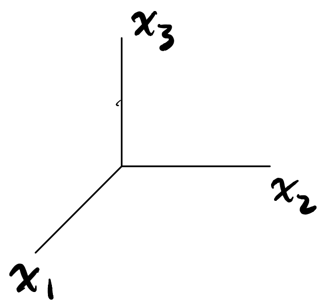
$$\mathcal{N}(A) = \{x \in \mathbb{R}^n \mid Ax = 0\}$$

Column perspective:



$A = [A_1 A_2 A_3]$

A



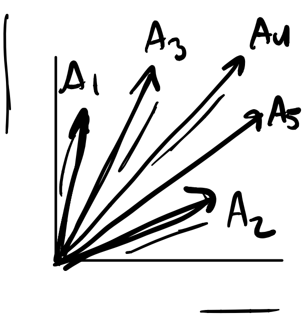
Redundant ways to reach a point in the codomain:

$$y = Ax \quad y = Ax' \quad x \neq x' \Big\}$$

$$x - x' \in \mathcal{N}(A)$$

$$Ax = y = Ax' \Rightarrow A(\underline{x - x'}) = 0$$

$$\left.\begin{aligned} y &= [A_1 A_2 A_3]\begin{bmatrix}1\\1\\0\end{bmatrix} \\ y &= [A_1 A_2 A_3]\begin{bmatrix}0\\2\\-1\end{bmatrix}\end{aligned}\right\} \to A\Bigg(\underbrace{\begin{bmatrix}1\\1\\0\end{bmatrix} - \begin{bmatrix}0\\2\\-1\end{bmatrix}}_{\begin{bmatrix}1\\-1\\1\end{bmatrix}}\Bigg) = 0$$



$A = [A_1 A_2 A_3 A_4 A_5]$

- use A_1, A_2 as basis vectors for $\mathbb{R}^2$
- compute coords of $A_3\ A_4, A_5$ w.r.t. A_1 & A_2

$[A_1 A_2]$ $\quad B_3, B_4, B_5 \in \mathbb{R}^2$

$A_3 = [A_1 A_2] B_3 \quad A_4 = [A_1 A_2] B_4$

$A_5 = [A_1 A_2] B_5$

$$A = [A_1 A_2] \begin{bmatrix} 1 & 0 & B_3 & B_4 & B_5 \\ 0 & 1 \end{bmatrix}$$

$$= \underbrace{[A_1 A_2] [I \;\; B]}\begin{bmatrix} -B \\ I \end{bmatrix}$$

$B = [B_3 \; B_4 \; B_5]$

$$-B + B = 0$$

the cols of

$\begin{bmatrix} -B \\ I \end{bmatrix} = \begin{bmatrix} -B_3 \; -B_4 \; -B_5 \\ I \end{bmatrix}$ span the nullspace of A.

---

in general.

if A has k lin ind cols.

lin ind cols $\qquad$ $n-k$ cols

break up A into $A = [M \mid M\overbrace{B}]$

lin dep cols

$A = M [I \mid B]$ $\qquad$ $n-k$ cols

$N(A) = \mathcal{R}\left(\begin{bmatrix} -B \\ I \end{bmatrix}\right)$

$B_4 = [A_1 A_2]^{-1} A_4$

$A_4 = [A_1 A_2] B_4$

$A = [A_1 A_2 A_3 A_4 A_5]$

$= [A_1 A_2] \underline{[A_1 A_2]^{-1} [A_1 A_2} \; A_3 A_4 A_5]$

$$\begin{bmatrix} 1 & 0 & \underbrace{[A_1 A_2]^{-1} A_3}_{B_3} & \underbrace{[A_1 A_2]^{-1} A_4}_{B_4} & [A_1 A_2]^{-1} A_5 \\ 0 & 1 \end{bmatrix}$$

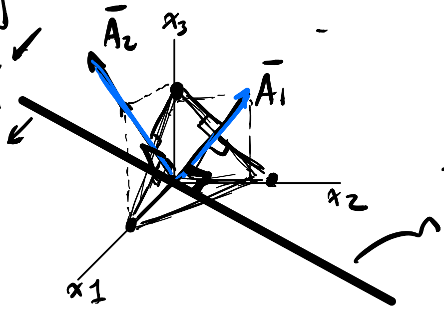
Row perspective

$A = [A_1\ A_2\ A_3]$

$A = \begin{bmatrix} -\ \bar{A}_1^T\ - \\ -\ \bar{A}_2^T\ - \end{bmatrix}$

$y_1 = |\bar{A}_1| \dfrac{\bar{A}_1^T x}{|\bar{A}_1|}$

$y_2 = |\bar{A}_2| \dfrac{\bar{A}_2^T x}{|\bar{A}_2|}$



$N(A)$

$Ax = 0$

$\left.\begin{aligned} \bar{A}_1^T x &= 0 \\ \bar{A}_2^T x &= 0 \end{aligned}\right\} \rightarrow$ $x \perp$ to all the rows

$R(A)$:
col perspective
more intuitive

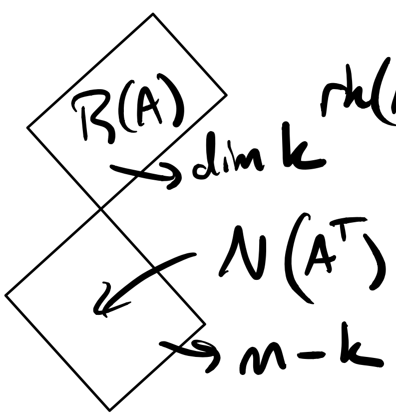
$N(A)$
row perspective
more intuitive

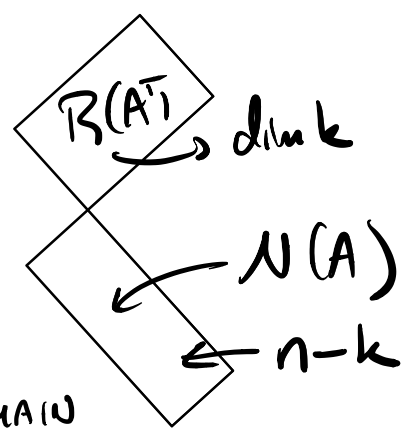
Fundamental Thm of LA:

$A \in \mathbb{R}^{m \times n}$

CODOMAIN $\longleftarrow$ DOMAIN

$rk(A) = k$





CODOMAIN
$R(A) \oplus^{\perp} N(A^T)$

DOMAIN
$R(A^T) \oplus^{\perp} N(A)$

Rank of a matrix:

col rank: # of lin ind cols.

row rank: # of lin ind rows

col rank = row rank = rank

$$\mathrm{rk}(A) = \mathrm{rk}(A^T)$$

Rank-Nullity theorem

$$\dim N(A) = n - \mathrm{rk}(A)$$

take all the lin ind cols of A.

coeffs of dep cols.

$$A = [M \;\; MB]$$

in the $N(A)$

$$= M[I \;\; B]\begin{bmatrix} -B \\ I \end{bmatrix}$$

$\begin{bmatrix} -B \\ I \end{bmatrix} \in \mathbb{R}^{n \times n-k}$ has $n-k$ lin ind cols.