Systems of Eqns: $A \in \mathbb{R}^{m \times n}$

$m > n$

A is tall

(lin ind cols)

$$y = Ax$$

Applications

data analysis

if redundant parameters

→ lin dep cols

prob no soln

(unless $y \in R(A)$

Find x to minimize

$$\min_x |y - Ax|^2 \Leftarrow$$

$$\Rightarrow \frac{\partial}{\partial x}|y-Ax|^2 = \frac{\partial}{\partial x}(y-Ax)^T(y-Ax)$$

$$= \frac{\partial}{\partial v} v^T v \Big|_{y-Ax} \frac{\partial(y-Ax)}{\partial x}$$

$$= -2(y-Ax)^T(+A) = 0$$

$$= y^TA - x^TA^TA = 0$$

$$x^T = y^TA(A^TA)^{-1}$$

$$\Rightarrow \boxed{x = (A^TA)^{-1}A^Ty}$$

A square invertible

$m = n$

$$y = Ax$$

$$x = A^{-1}y$$

unique soln

$m < n$

A is fat

→ (lin ind rows)

$$y = Ax$$

Applications

Constraints

if redundant constraints

→ lin dep rows

many solns

subspace of solns

if $\bar{x}$ satisfies

$y = A\bar{x}$ & $z \in N(A)$

$y = A(\bar{x} + z) = A\bar{x}$

set of solns

$\bar{x} + N(A)$

Find x

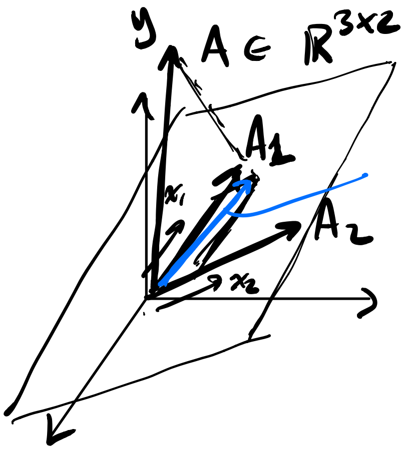
$$\min |x|^2 \Leftarrow$$

s.t. $y = Ax$ }—

$y \geq Ax$

Lagrangian:

$$\mathcal{L} = x^Tx - \lambda^T(y - Ax)$$



$A = \begin{bmatrix} A_1 & A_2 \end{bmatrix}$

$A \underbrace{(A^TA)^{-1}A^Ty}_{x}$

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial x} = 2x^T + \lambda^T A = 0 \\ \frac{\partial \mathcal{L}}{\partial \lambda} = -(y - Ax) = 0 \end{cases}$$

$$2x^T + \lambda^T A = 0$$

$$\left[2[-\ x^T -] + [\lambda^T][\ A\ ] = 0 \right][A^T]$$

$$2x^TA^T + \lambda^T \underbrace{AA^T}_{\text{invertible}} = 0$$

$y = Ax$

$$2y^T + \lambda^T AA^T = 0$$

$$\lambda^T = -2y^T(AA^T)^{-1}$$

Summary:

LS: $x = (\underline{A^TA})^{-1}\underline{A^T}y$

(tall A) $[A^T][\ ]$

MIN NORM: $x = \underline{A^T}(\underline{AA^T})^{-1}y$

(fat A) $[A][A^T]$

$2[A]_3 [A^T]$

$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 3\left[\ \right]^2 \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 3[A^T]^2 \quad 2[\boxtimes]^{-1} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$

$$2x^T = -\lambda^T A$$

$$2x^T = 2y^T(AA^T)^{-1}A$$

$$\Rightarrow \boxed{x^* = A^T(\underline{AA^T})^{-1}y}$$

Can show

$|x^*| \le |x^* + z|$ if $z \in \underline{N(A)}$

$|x^*|^2 \le |x^* + z|^2$

$(x^* + z)^T(x^* + z)$

$x^{*T}x^* + \cancel{z^Tx^*} + \cancel{x^{*T}z} + z^Tz$

$z^Tx^* = z^TA^T(AA^T)^{-1}y$

$(Az)^T = [0 \cdots 0]$

$|x^* + z|^2 = x^{*T}x^* + z^Tz$

$|x^*|^2 \le |x^*|^2 + \underbrace{|z|^2}_{c}$

Rank & Fund of Lin Alg.

Review: row rank (A) = # of lin ind rows = $\dim R(A^T)$

col rank (A) = # of lin ind cols = $\dim(R(A))$

$A = [A_1 \cdots A_n]$ $R(A)$ = lin comb of cols.

$A = \begin{bmatrix} -\bar{a}_1^T- \\ -\bar{a}_m^T- \end{bmatrix}$ $A^T = \begin{bmatrix} | & & | \\ \bar{a}_1 & \cdots & a_m \\ | & & | \end{bmatrix}$ $R(A^T)$ = lin combs of rows

---

row rank = col rank = rank rank(A) = rank(A^T)

$rk(A) = rk(A^T)$

rank(A) = $rk(A)$

Rank Nullity Thm: $A \in \mathbb{R}^{m \times n}$

$$\underbrace{rk(A)}_{k} + \underbrace{\dim N(A)}_{n-k} = n$$

Proof: by construction $A \in \mathbb{R}^{m \times n}$

$A \rightarrow$ take a set of lin ind cols of A that span $R(A)$

$\underline{A} = [\underbrace{B_1}_{} \underbrace{B_2}_{}] \Longrightarrow \underline{B_2} = \underline{B_1 V}$ for some matrix of coeffs V

cols are lin ind. & span $R(A)$

cols are lin dep on cols of B_1

$= B_1 [v_1 \cdots v_{n-k}]$

$B_1 \in \mathbb{R}^{m \times k}$ $B_2 \in \mathbb{R}^{m \times (n-k)}$ $V \in \mathbb{R}^{k \times (n-k)}$

Many different choices for B_1 $A = [\, | \; | \; | \; | \; | \; | \;\; | \;\; | \; | \,]$

$$A = [B_1 \;\; B_1 V] = B_1 [I \;\; V]$$

$$N = \begin{bmatrix} -V \\ I \end{bmatrix} \qquad AN = B_1 [I \;\; V] \begin{bmatrix} -V \\ I \end{bmatrix} = B_1 [-V + V] = 0$$

explicit construction of a **basis** for nullspace of A $\rightarrow$ cols of $N = \begin{bmatrix} -V \\ I \end{bmatrix}$

- **lin ind** $Nz = \begin{bmatrix} -V \\ I \end{bmatrix} z = \begin{bmatrix} -Vz \\ z \end{bmatrix} = 0 \Rightarrow z = 0$

- if $Ax = 0$ $\Rightarrow x \in \text{span}(N)$

$$x = \begin{bmatrix} -V \\ I \end{bmatrix} x_2$$

$\Downarrow$

cols of N are a basis for nullspace of A

$$Ax = [B_1 \;\; B_1 V] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$$= B_1 x_1 + B_1 V x_2 = 0$$

$$\Rightarrow B_1 x_1 = -B_1 V x_2$$

$$B_1^T (B_1 x_1 = -B_1 V x_2)$$

$$B_1^T B_1 x_1 = -B_1^T B_1 V x_2$$

since cols of B_1 are lin ind.

$B_1^T B_1 \rightarrow$ invertible

$$x_1 = -V x_2$$

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -V x_2 \\ I x_2 \end{bmatrix} = \begin{bmatrix} -V \\ I \end{bmatrix} x_2$$

Row perspective on Nullspace ...

$$A = \begin{bmatrix} - \bar{a}_1^T - \\ - \bar{a}_m^T - \end{bmatrix} \quad \text{if } x \in \mathcal{N}(A)$$
$$\bar{a}_i^T x = 0 \quad \forall i$$

$$Ax = \begin{bmatrix} \bar{a}_1^T x \\ \vdots \\ \bar{a}_m^T x \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \quad x \in \mathcal{N}(A) \text{ then}$$
$x \perp$ rows of A.

$\Rightarrow \underline{\mathcal{R}(A^T)} \perp \underline{\mathcal{N}(A)}$

Properties of Rank $\quad f(\underline{x}) = \underline{A}x \quad A \in \mathbb{R}^{m \times n}$

- $rk(A) \le \min(m, n)$

  $rk(A) = \min(m,n) \Longrightarrow$ full rank matrix

  full rank fat matrix has m lin ind rows
  full rank tall matrix has n lin ind cols

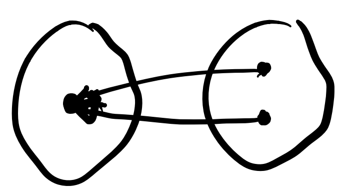
  not full rank $\rightarrow$ rank deficient

- $rk(A) = 0 \Rightarrow A = 0$
- A has rank n: "full col rank" has to be square or tall.

  f is one-to-one / injective $\rightarrow \mathcal{N}(A) = \{0\}$ nullspace is trivial

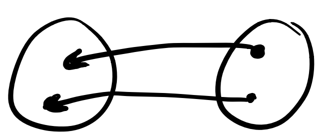
  every x has a unique $y = f(x)$
  $= Ax$

Not injective



CODOM DOM

injective



CODOM DOM

if $z \in \mathcal{N}(A)$
$z \neq 0$

$Ax = y = Ax + z$

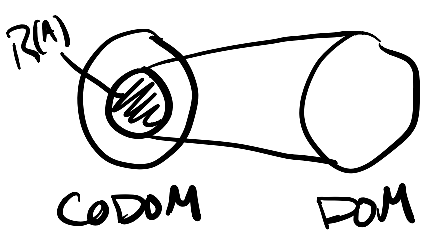
both x, $x+z$ give you y

- A has rank m: "full row rank" — has to be square or fat

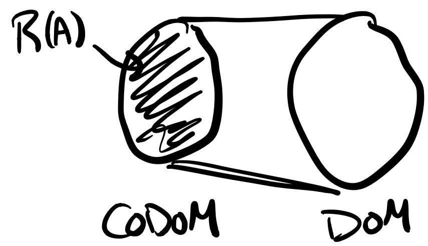
f is onto / surjective

every y has at least one x s.t. $y = f(x)$

NOT SURJECTIVE



SURJECTIVE



$R(A) = \mathbb{R}^m$

span of cols

---

full row rank	full col rank
$R(A) = \mathbb{R}^m$	$N(A) = 0$
span of cols	subspace $\perp$ to rows of A

- A is square $m = n$

A is invertible $\iff$ rank n

Note: rank 1 matrices / dyads — 1 col repeated w different scalings

$A = \underbrace{ab^T}_{\text{outer prod}}$ $\quad a = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix}$ $\quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$

or equiv 1 row repeated w diff. scalings

$$A = \begin{bmatrix} | & & | \\ a b_1 & \cdots & a b_n \\ | & & | \end{bmatrix} = \begin{bmatrix} a_1 b^T - \\ \vdots \\ a_m b^T - \end{bmatrix} \leftarrow$$

- $B \in \mathbb{R}^{n \times k}$

  $rk(AB) \leq \min(rk(A), rk(B))$

- $B \in \mathbb{R}^{n \times k}$, $rk(B) = n$

  $rk(AB) = rk(A)$

  unproven

- $C \in \mathbb{R}^{l \times m}$, $rk(C) = m$

  $rk(CA) = rk(A)$

- Sylvester's Rk inequality

  $rk(A) + rk(B) - n \leq rk(AB)$

- Frobenius

  $rk(AB) + rk(BC) \leq rk(B) + rk(ABC)$

- Subadditivity:

  $rk(A+B) \leq rk(A) + rk(B)$

  $\Rightarrow$ a rank-k matrix can be written as a sum of k rank-1 matrices and no fewer

$$A = \begin{bmatrix} a_1 & a_2 & a_3 = a_1 + a_2 \end{bmatrix} = \begin{bmatrix} a_1 \end{bmatrix} [1\ 0\ 1] + \begin{bmatrix} a_2 \end{bmatrix} [0\ 1\ 1]$$

$rk(A) = 2.$

- $\mathrm{rk}(A^TA) = \mathrm{rk}(AA^T) = \underline{\mathrm{rk}(A)} = \mathrm{rk}\underline{(A^T)}$

$\underline{A^TA}, \underline{AA^T}$: "Grammians"

$x \in N(A^TA)$ iff $x \in N(A)$

if $x \in N(A)$: $Ax = 0$ $\quad A^TAx = 0 = A^T0 = 0$

if $x \in N(A^TA)$ $\quad A^TAx = 0$ $\quad x^TA^TAx = 0$

$|\underline{Ax}|^2 = 0$ $\quad Ax = 0$

$x \in N(A)$

Defns: vector space $\mathcal{W}$

Orthogonal Complement

subspace $\mathcal{V}_1 \subseteq \mathcal{W}$

$$\mathcal{V}_1^\perp = \{ w \in \mathcal{W} \mid w^Tv = 0, v \in \mathcal{V}_1 \}$$

set of all vectors $\perp$ to all vectors in $\mathcal{V}_1$

Direct Sum:

$\mathcal{V}_1 \subseteq \mathcal{W}$ $\quad \mathcal{V}_2 \subseteq \mathcal{W}$

$$\underline{\mathcal{V}_1} \oplus \underline{\mathcal{V}_2} = \{ w \in \mathcal{W} \mid w = v_1 + v_2, v_1 \in \mathcal{V}_1 \ v_2 \in \mathcal{V}_2 \}$$

span of both spaces

$\mathcal{V} \subseteq \mathcal{W}$ $\quad \mathcal{W} = \mathcal{V} \oplus \mathcal{V}^\perp$

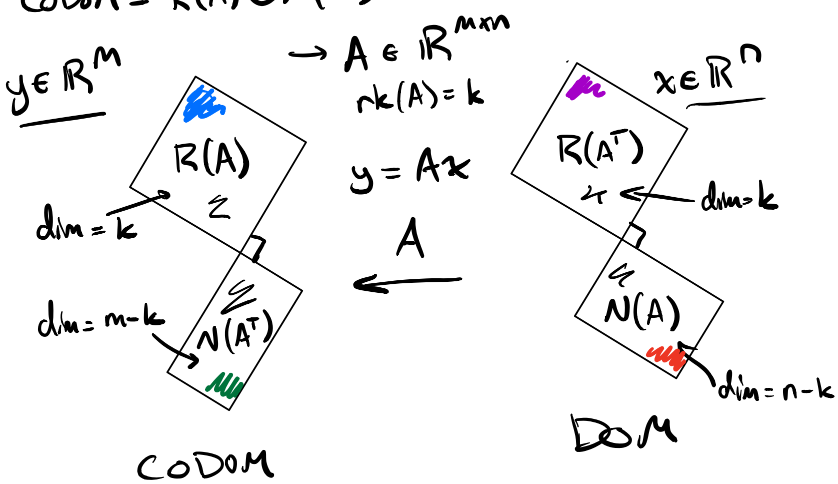
FUND. THM OF LIN ALG: $A \in \mathbb{R}^{m \times n}$

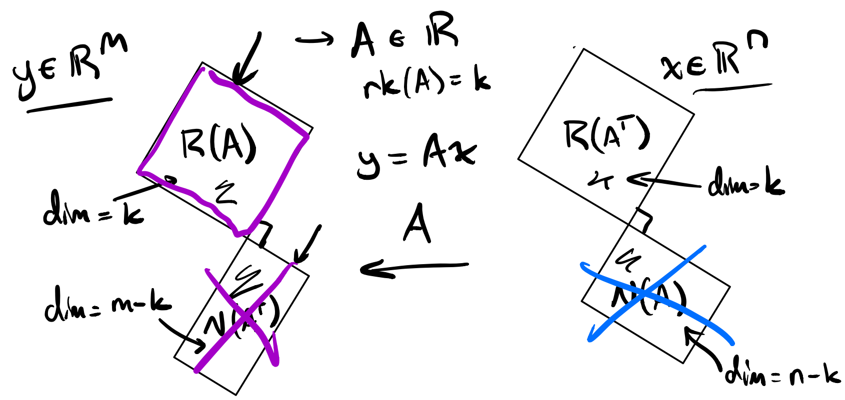
4 subspaces: $R(A)$, $R(A^T)$, $N(A)$, $N(A^T)$

$N(A)$: right nullspace

$N(A^T)$: left nullspace

CODOMAIN	$\xleftarrow{y = Ax}$	DOMAIN
$R(A), N(A^T)$		$R(A^T), N(A)$
$R(A)^\perp = N(A^T)$		$R(A^T)^\perp = N(A)$
$\text{CODOM} = R(A) \oplus N(A^T)$		$\text{DOM} = R(A^T) \oplus N(A)$





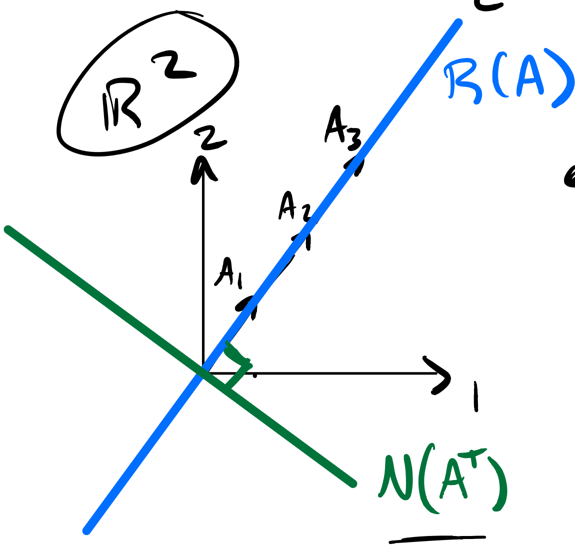
surjective

injective

Pictures

$A \in \mathbb{R}^{2\times 3}$ $A = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix}$

$$A = \begin{bmatrix} A_1 & A_2 & A_3 \end{bmatrix} = \begin{bmatrix} — \bar{a}_1^T — \\ — \bar{a}_2^T — \end{bmatrix}$$



$\xleftarrow{A}$