Week 2

Note Title 4/4/2021

LP in standard form

$$\min_{x \in \mathbb{R}^n} p^T x \quad \text{s.t.} \quad \left.\begin{array}{l} Ax \geq b \\ x \geq 0 \end{array}\right\} \underset{}{m+n \text{ linear inequality constraints}}$$

objective function

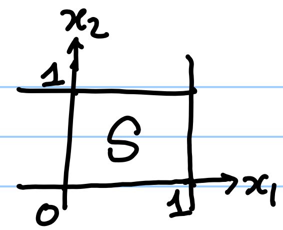
$S := \{x \in \mathbb{R}^n : Ax \geq b, x \geq 0\}$ is called the feasible region / constraint set of the LP.

Suggestion : think of the feasible region and the objective function as two separate entities

Example :

$$S = \{ x \in \mathbb{R}^2 : 0 \le x_1 \le 1,\ 0 \le x_2 \le 1 \}$$

$$= \{ x \in \mathbb{R}^2 : \underbrace{\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \ge \underbrace{\begin{bmatrix} -1 \\ -1 \end{bmatrix}}_{b},\ x \ge 0 \}$$



$0 \le x_1 \le 1$

$0 \le x_2 \le 1$

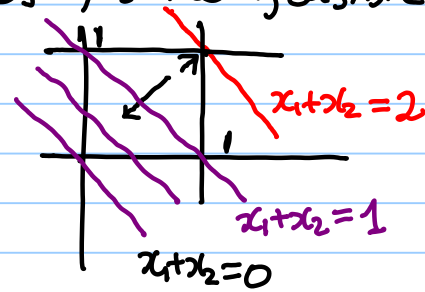
many different objectives , same feasible region :

$-x_1 - x_2$

- $p = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

minimizer $= \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

minimum value $= 0$



- $p = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$

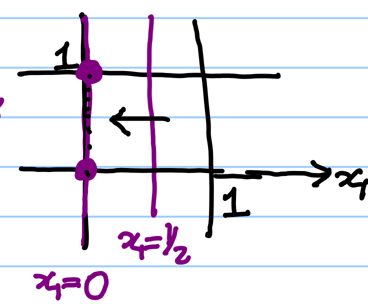
minimizer $= \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

minimum value $= -2$

- $p = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

any $(0, x_2)$, $0 \le x_2 \le 1$ is a minimizer, in particular, the vertices $(0,0)$, $(0,1)$ of S.

minimum value $= 0$



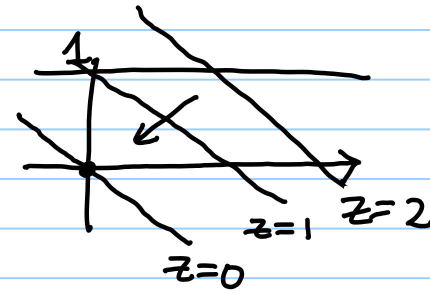
Another such easy example :

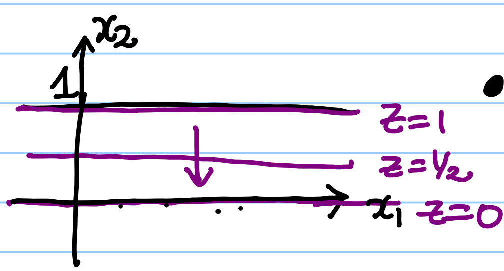
$$S = \{x \in \mathbb{R}^2 : x_2 \leq 1, \; x \geq 0\}$$

$$= \{x \in \mathbb{R}^2 : \underbrace{[0,-1]}_{A} x \geq \underbrace{[-1]}_{b}, \; x \geq 0\}$$

- if $p = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $z = x_1 + x_2$

minimizer $= \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, minimum value $= 0$





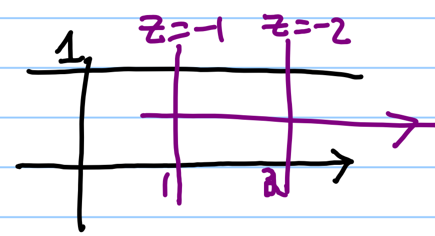
- if $p = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $z = x_2$,

Every point of the form $\begin{bmatrix} x_1 \\ 0 \end{bmatrix}$ $x_1 \geq 0$ is a minimizer

minimum value $= 0$

- if $p = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$, $z = -x_1$

no minimizer ! minimum value $= -\infty$



we say: this LP is <u>unbounded</u>

Fundamental theorem of LP :

If a LP is bounded (i.e. the minimum value is finite), then its minimizer is always attained at a vertex of the feasible region. (Rigorous proof omitted, hope it is evident enough to you.)

A <u>vertex</u> of $S=\{x\in\mathbb{R}^n : Ax\geqslant b, x\geqslant 0\}$

is the solution of any choice of n out of the $m+n$ linear equations

$$\left.\begin{cases} Ax=b \\ x=0 \end{cases}\right\} \begin{matrix} m \\ + \\ n \end{matrix}$$

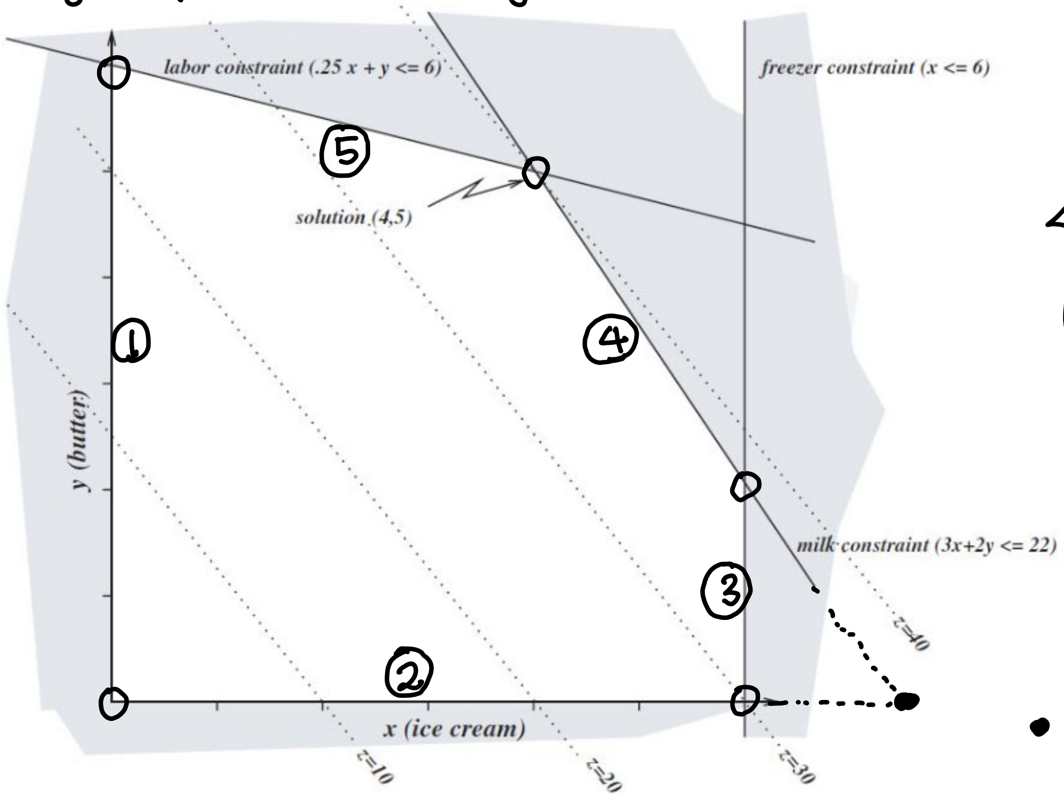
so that (i) the $n\times n$ linear equations are linearly independent

$\Updownarrow$

unique solution

(ii) the (unique) solution satisfies the remaining m inequality constraints.

Feasible region from the dairy problem:



Equations defining the boundary of S

$$
\left\{
\begin{array}{lcll}
x_1 & = & 0 & \text{— (1)} \\
x_2 & = & 0 & \text{— (2)} \\
-x_1 & = & -6 & \text{— (3)} \\
-3x_1 - 2x_2 & = & -22 & \text{— (4)} \\
-\frac{1}{4}x_1 - x_2 & = & -6 & \text{— (5)}
\end{array}
\right.
$$

(1), (2): $n=2$

(3), (4), (5): $m=3$

Note

- (1) + (3) gives a linearly dependent 2x2 system

  $\rightarrow$ no solution

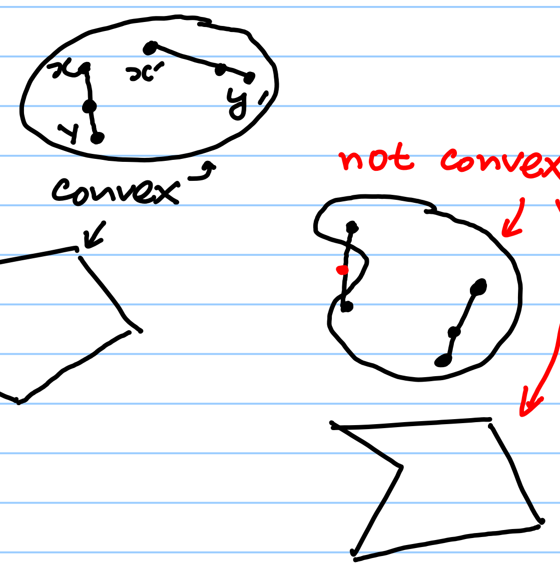
- (2) + (4) gives a linear independent 2x2 system, but the solution does not satisfy $-x_1 \geq -6$

Another fundamental fact:

$S = \{x \in \mathbb{R}^n : Ax \geq b,\ x \geq 0\}$ is always a convex subset of $\mathbb{R}^n$

[$C \subseteq \mathbb{R}^n$ is convex if $\forall x, y \in C$, $\forall t \in [0,1]$,

$$(1-t)x + ty \in S \ ]$$



<u>Proof:</u> Let $x, y \in S$, $0 \leq t \leq 1$.

$x \geq 0,\ y \geq 0 \Rightarrow (1-t)x + ty \geq 0$

Similarly, $Ax \geq b,\ Ay \geq b \Rightarrow \underbrace{(1-t)Ax + tAy}_{\substack{\parallel \\ A((1-t)x+ty)}} \geq b$

linearity of A!

so $(1-t)x + ty \in S$ $\blacksquare$

The fundamental theorem gives an obvious "naive algorithm" for solving LP:

Step 1: Solve for all vertices by taking all $\binom{m+n}{n}$ choices of n equations from the $m+n$ equations that define the feasible region.

Step 2: check which vertex gives the smallest z value.

Problem: does not handle unbounded LPs.

much bigger problem: Say if $n=100$, $m=50$ (not a big problem size in today's standard)

$$\binom{m+n}{n} = \binom{150}{50} \doteq 2 \times 10^{40}$$

Assume we can solve $\overset{1 \text{ million}}{\vee}$ 100×100 linear systems in 1 second, the naive algorithm will take

$$\binom{150}{50} / 60 / 60 / 24 / 365 / 10^6 \doteq 6.38 \times 10^{26} \text{ Years !}$$

Solving a LP is not the same as a linear system.

But we need techniques for the latter to solve LP.

The subject of linear algebra gives a systematic set of tools to answer the following basic questions:

Given a linear system $\overset{m\times n}{A}x = \overset{m\times 1}{b}$,

- does a solution exist?
- if so, what are the solutions? is the solution unique?
- algorithms for finding solutions?

Existence

- $\text{range}(A)$ (or $\text{Image}(A)$) $:= \{ Ax : x \in \mathbb{R}^n \} \subseteq \mathbb{R}^m$

  is a subspace of $\mathbb{R}^m$

  (Proof : if $y_1, y_2 \in \text{Image}(A)$, $\alpha_1, \alpha_2 \in \mathbb{R}$

  then $y_1 = Ax_1$, $y_2 = Ax_2$

  $$\alpha_1 y_1 + \alpha_2 y_2 = \alpha_1 A x_1 + \alpha_2 A x_2 \overset{\text{linearity}}{=} A(\alpha_1 x_1 + \alpha_2 x_2) \in \text{Image}(A) \;)$$

- $Ax = b$ has a solution $\iff$ $\underbrace{b \in \text{range}(A)}_{\text{relatively easy to determine because range(A) has a simple structure}}$

Working in tandem, these two facts are useful for checking existence of solutions.

In particular, if $\underbrace{\dim(\text{Image}(A))}_{= \text{rank}(A) \text{ (a.k.a. column rank of } A)} = m$, then $Ax=b$ has a solution $\forall b$.

Uniqueness

- $\text{null}(A)$ (or $\text{Ker}(A)$) $:= \{x \in \mathbb{R}^n : Ax = 0\}$

  is a subspace of $\mathbb{R}^n$

  (Proof: If $x_1, x_2 \in \text{null}(A)$, $\alpha_1, \alpha_2 \in \mathbb{R}$ then

  $$Ax_1 = 0, \quad Ax_2 = 0$$

  $$A(\alpha_1 x_1 + \alpha_2 x_2) \overset{\text{linearity}}{=} \alpha_1 A x_1 + \alpha_2 A x_2 = \alpha_1 \cdot 0 + \alpha_2 \cdot 0 = 0$$

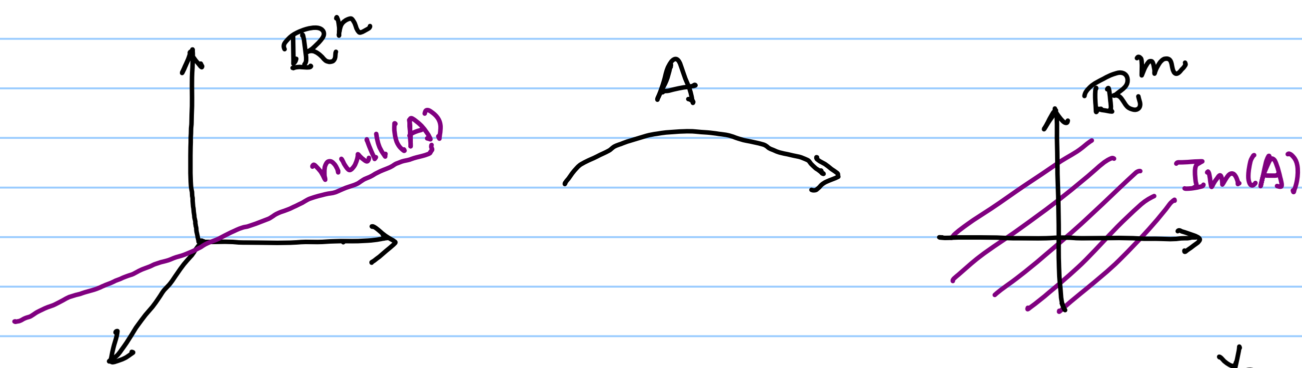
  So $\alpha_1 x_1 + \alpha_2 x_2 \in \text{null}(A)$ ∎)

- $Ax = b \iff A(x+z) = b$ for any $z \in \text{null}(A)$

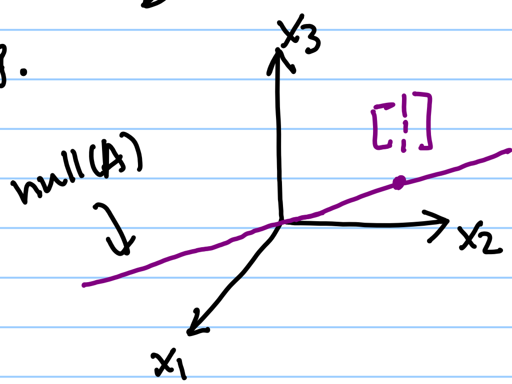
Working in tandem, these two facts are useful for determining the set of all solutions for a given linear system $Ax = b$

$$\text{solution set} = \{\text{a particular solution} + z : z \in \text{null}(A)\}$$

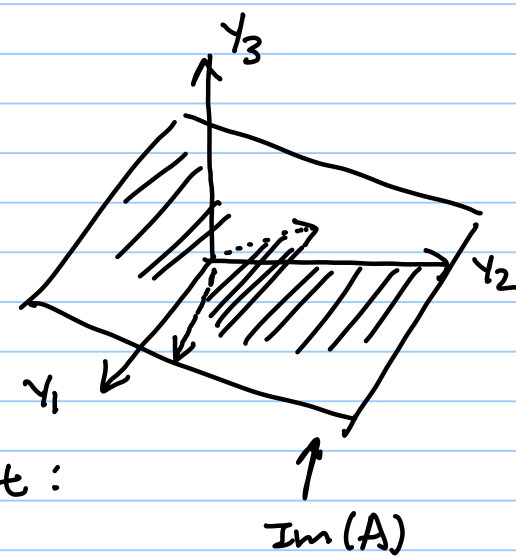
In particular, if $\dim(\text{null}(A)) \overset{\text{nullity}(A)}{=} 0$, the solution, if exists, is unique.



E.g.



$$A = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix}$$



Yet another basic (but less obvious) fact:

$$\text{rank}(A) + \text{nullity}(A) = n$$

In particular, if $m=n$, then $\text{rank}(A) = n \iff \text{nullity}(A) = 0$

$\Updownarrow$ sol. exists $\forall b$ $\qquad$ $\Updownarrow$ solution unique, $\forall b$

A very basic viewpoint: when thinking of linear systems

$$Ax = b,$$

leave the vector b alone first, and think of A as a (linear) map.

Think: $\mathbb{R}^n \xrightarrow{A} \mathbb{R}^m$

Do not just treat A as a boring array of numbers!

Image(A) and null(A), being (linear) subspaces, can be described precisely by <u>basis elements</u>. You probably remember something called the

"row reduced echelon form" of a matrix as a tool to determine

Image(A), null(A) etc.

In this course, we devolop an alternate procedure, called Jordan exchange, that serves the same purpose, and is also convenient for describing the simplex method for LP.

Let $A \in \mathbb{R}^{m \times n}$, the following tableau reminds us to think of A as a linear map $x \mapsto y = Ax$

$$x \to \boxed{A} \to y$$

input/independent variables

output/dependent variables

$$\begin{array}{cc|ccccc|}
 & & x_1 & \cdots & x_s & \cdots & x_n \\
\hline
y_1 & = & A_{11} & \cdots & A_{1s} & \cdots & A_{1n} \\
\vdots & & \vdots & & \vdots & & \vdots \\
y_r & = & A_{r1} & \cdots & A_{rs} & \cdots & A_{rn} \\
\vdots & & \vdots & & \vdots & & \vdots \\
y_m & = & A_{m1} & \cdots & A_{ms} & \cdots & A_{mn} \\
\hline
\end{array}$$

Jordan exchange is about interchanging the roles of some of the independent variables and dependent variables.

Basic step :

If $A_{rs} \neq 0$

$$y_r = A_{r1} x_1 + \cdots + A_{rs} x_s + \cdots + A_{rn} x_n$$

$$\Rightarrow \quad x_s = \underbrace{\frac{1}{A_{rs}}}_{B_{rs}} y_r + \sum_{\substack{j=1 \\ j \neq s}}^{n} \underbrace{\left(\frac{-A_{rj}}{A_{rs}} \right)}_{B_{rj}} x_j$$

Next, substitute this expression to the rest of the equations in order to express y_i $(i \neq r)$ as a linear combination of $x_1, \cdots, x_{r-1}, x_{r+1}, \cdots, x_n$ and y_r :

$$
\begin{aligned}
y_i &= \sum_{\substack{j=1\\ j\neq s}}^{n} A_{ij}x_j + A_{is}\left(\frac{1}{A_{rs}}y_r + \sum_{\substack{j=1\\ j\neq s}}^{n} \frac{-A_{rj}}{A_{rs}}x_j\right) \\
&= \sum_{\substack{j=1\\ j\neq s}}^{n} B_{ij}x_j + B_{is}y_r,
\end{aligned}
$$

where

$$
B_{is} = \frac{A_{is}}{A_{rs}}, \qquad B_{ij} = \left(A_{ij} - \frac{A_{is}}{A_{rs}}A_{rj}\right) = \left(A_{ij} - B_{is}A_{rj}\right) \qquad \forall i \neq r,\ j \neq s.
$$

As such, we have the following tableau that represents the same set of linear relations, but with y_r as an independent variable, and x_s as a dependent variable.

[see computer demo.]

$$
\begin{array}{cc|ccccccc|}
 & & x_1 & \cdots & x_{s-1} & y_r & x_{s+1} & \cdots & x_n \\
\hline
y_1 & = & B_{11} & \cdots & & B_{1s} & & \cdots & B_{1n} \\
\vdots & & \vdots & & & \vdots & & & \vdots \\
y_{r-1} & = & & & & & & & \\
x_s & = & B_{r1} & \cdots & & B_{rs} & & \cdots & B_{rn} \\
y_{r+1} & = & & & & & & & \\
\vdots & & \vdots & & & \vdots & & & \vdots \\
y_m & = & B_{m1} & \cdots & & B_{ms} & & \cdots & B_{mn} \\
\hline
\end{array}
$$

A simple geometric way to solve a system of two equations in two unknowns is to plot the corresponding lines and determine the point where they intersect. Of course, this technique fails when the lines are parallel to one another. A key idea in linear algebra is that of *linear dependence*, which is a generalization of the idea of parallel lines. Given a matrix $A \in \mathbf{R}^{m \times n}$, we may ask if any of its rows are redundant. In other words, is there a row $A_{k\cdot}$ that can be expressed as a linear combination of the other rows? That is,

$$A_{k\cdot} = \sum_{\substack{i=1 \\ i \neq k}}^{m} \lambda_i A_{i\cdot}. \tag{2.7}$$

If so, then the rows of A are said to be *linearly dependent*.

The idea of linear independence extends also to functions, including the linear functions y defined by $y(x) := Ax$ that we have been considering above. The functions $y_i(x)$, $i = 1, 2, \ldots, m$, defined by $y(x) := Ax$ are said to be linearly dependent if

$$z'y(x) = 0 \qquad \forall x \in \mathbf{R}^n \text{ for some nonzero } z \in \mathbf{R}^m$$

and linearly independent if

$$z'y(x) = 0 \qquad \forall x \in \mathbf{R}^n \implies z = 0. \tag{2.8}$$

The equivalence of the linear independence definitions for matrices and functions is clear when we note that

$$\begin{aligned} & z'Ax = 0 \qquad \forall x \in \mathbf{R}^n \text{ for some nonzero } z \in \mathbf{R}^m \\ \iff & z'A = 0 \qquad \text{for some nonzero } z \in \mathbf{R}^m. \end{aligned}$$

Thus the functions $y(x)$ are linearly independent if and only if the rows of the matrix A are linearly independent.

Theorem 2.2.3 (Steinitz). *For a given matrix $A \in \mathbf{R}^{m \times n}$, the linear functions y, defined by $y(x) := Ax$, are linearly independent if and only if for the corresponding tableau all the y_i's can be exchanged with some m independent x_j's.*

(In particular, $m \leq n$ if the rows are linearly independent.)

more generally:

Definition 4.1.2. The row rank of a matrix $A \in \mathbf{R}^{m \times n}$ is the maximum number of linearly independent rows that are present in A.

Theorem 4.1.3. *Given $A \in \mathbf{R}^{m \times n}$, form the tableau $y := Ax$. Using Jordan exchanges, pivot as many of the y's to the top of the tableau as possible. The rank of A is equal to the number of y's pivoted to the top.*

- pivot as many y's to the top as possible
- reorder the rows and columns

$$
\begin{array}{c|c|}
 & x_1 \cdots\cdots x_n \\
\hline
\begin{matrix} y_1 = \\ \vdots \\ y_m = \end{matrix} & A \\
\hline
\end{array}
\quad \xrightarrow{\qquad\qquad\qquad\qquad} \quad
\begin{array}{c|c|c|}
 & y_{I_1} & x_{J_2} \\
\hline
x_{J_1} = & B_{I_1 J_1} & B_{I_1 J_2} \\
\hline
y_{I_2} & B_{I_2 J_1} & 0 \\
\hline
\end{array}
$$

$$\{1, \cdots, n\} = J_1 \cup J_2$$

$$\{1, \cdots\cdots, m\} = I_1 \cup I_2$$

same size

Not found in the textbook, (the book never mentions 'null space' or 'nullity')

$$|I_1| = |J_1| = \text{rank}(A)$$

$$|J_2| = \text{nullity}(A)$$

$$\text{rank}(A) + \text{nullity}(A) = |J_1| + |J_2| = n$$

Theorem 4.1.4. *Let* $A \in \mathbf{R}^{m \times n}$*; then* rank$(A)$*—the number of linearly independent rows of* A*—is equal to the number of linearly independent columns of* A.

(row rank $=$ column rank)

$$\dim\big(\underbrace{\{u^T A : u \in \mathbb{R}^m\}}_{\text{a linear subspace of } \mathbb{R}^m}\big) \qquad \dim\big(\underbrace{\{Ax : x \in \mathbb{R}^n\}}_{\text{a linear subspace of } \mathbb{R}^n}\big)$$

$$\underset{\underset{\mathbb{R}^m}{\cap}}{u} \longrightarrow \boxed{A^T} \longrightarrow \underset{\underset{\mathbb{R}^n}{\cap}}{A^T u} \qquad\qquad \underset{\underset{\mathbb{R}^n}{\cap}}{x} \longrightarrow \boxed{A} \longrightarrow \underset{\underset{\mathbb{R}^m}{\cap}}{y = Ax}$$

Note : $(u^T A)^T = A^T u$

Solving a general linear system $\overset{m\times n}{A}x=\overset{m\times 1}{b}$ (sec 2.4)

Warmup exercise, connect the dots :

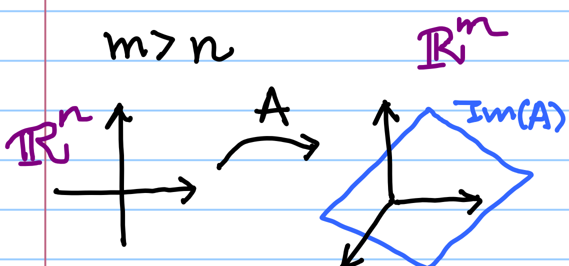
$m>n$ • • typically unique solution

$m=n$ • • typically no solution

$m<n$ • • typically infinitely solution

"typically" means for "almost all" choices of A and b

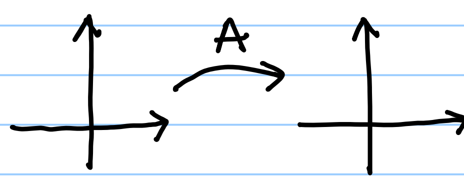
$m>n$



typically, rank(A) = n

A is typically injective but never surjective

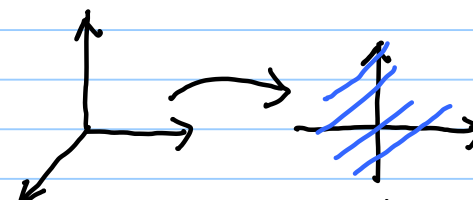
$m=n$



typically, rank(A) = n = m

typically both injective and surjective, i.e. invertible

$m<n$



typically, rank(A) = m

typically surjective but never injective

Here is an algorithm for solving $Ax=b$, it handles all cases, even the atypical ones:

1. Write the system in the following tableau form:

$$
\begin{array}{rcl}
 & & \begin{array}{cc} x & 1 \end{array} \\
y & = & \begin{array}{|c|c|} \hline A & -b \\ \hline \end{array}
\end{array}
$$

   Our aim is to seek x and y related by this tableau *such that* $y = 0$.

2. Pivot as many of the y_i's to the top of the tableau, say y_{I_1}, until no more can be pivoted, in which case we are blocked by a tableau as follows (with row and column reordering):

$$
\begin{array}{rc|cc|c|}
 & & y_{I_1} & x_{J_2} & 1 \\ \hline
x_{J_1} & = & B_{I_1J_1} & B_{I_1J_2} & d_{I_1} \\
y_{I_2} & = & B_{I_2J_1} & 0 & d_{I_2} \\ \hline
\end{array}
$$

   We now ask the question: Is it possible to find x and y related by this tableau such that $y = 0$?

3. <u>The system is solvable if and only if $d_{I_2} = 0$</u>, since *we require* $y_{I_1} = 0$ *and* $y_{I_2} = 0$. When $d_{I_2} = 0$, we obtain by writing out the relationships in the tableau explicitly that

$$
\left.
\begin{aligned}
y_{I_1} &= 0, \\
y_{I_2} &= B_{I_2J_1} y_{I_1} = 0, \\
x_{J_2} & \quad \text{is arbitrary}, \\
x_{J_1} &= B_{I_1J_2} x_{J_2} + d_{I_1}.
\end{aligned}
\right\} \rightarrow
\begin{bmatrix} x_{J_1} \\ x_{J_2} \end{bmatrix} = \begin{bmatrix} d_{I_1} \\ 0 \end{bmatrix} + \overbrace{\begin{bmatrix} B_{I_1J_2} \\ I \end{bmatrix}} x_{J_2}
$$

a particular solution

Columns of this matrix forms a basis of the null space of A

identity matrix, size $|J_2|$

Ex 2-4-1 : see class demo.