Convex Relaxation

Cartoon

$X \in \mathbb{R}^{n\times n}$ $\quad X = X^T \succeq 0 \quad \Big| \; \cancel{\operatorname{rank}(X) = 1}$ ←

$X = \boxed{xx^T}$

$\min \quad f(x) + \boxed{\|X\|_2}$ ★

s.t. $\quad xx^T \succeq 0$

$$\min \; \operatorname{rank}(X) \quad \text{s.t. } A(X) = b \quad \longrightarrow \quad \min \; \|X\|_* = \sum_{i=1}^{m} \sigma_i(X) \quad \text{s.t. } A(X) = b$$

$X = xx^T = U \begin{bmatrix} \sigma_1 & & \\ & 0 & \\ & & \ddots \; 0 \end{bmatrix} V^T$ $\qquad X \Rightarrow X = U \begin{bmatrix} \sigma_1 & & \\ & 0 & \\ & & \ddots \; 0 \end{bmatrix} V^T$

Generalizations of Network Flow & MDP PROBLEMS

shortest path
MDPs
$\}\rightarrow$ linear programs

Shortest path:

$$\min_x \; c^T x$$
s.t. $Ex = S, \; x \geq 0$

LP $\rightarrow$ convex prog.

c_e: cost of traveling on edge e.

$c_e(x_e)$: congestion cost

potential function

$$\boxed{c^T x \rightarrow f(x) = \sum_e \int_0^{x_e} c_e(u)\, du \Rightarrow \frac{\partial f}{\partial x} = c(x)^T}$$

Optimality cond:

$$\underbrace{c^T}_{\frac{\partial f}{\partial x}} + v^T E - \mu^T = 0 \longrightarrow c(x)^T + v^T E - \mu^T = 0$$

x: now a population vector

solution: $\rightarrow$ equilibrium of a routing game

a whole population solving a shortest path problem

$$\min f(x)$$
s.t. $Ex = S\underbrace{m}_{\text{total population mass}}$ $\quad x \geq 0$

Literature: Wardrop Equilibrium
Patrickson (sp?) $\rightarrow$ Traffic Assignment Problem.

$$C(x) = \underset{\substack{\downarrow \\ \text{diagonal } Q \succ 0}}{Qx} + C \longrightarrow \quad c_e(x_e) = \underline{Q_{ee}}\, \underline{x_e} + c_e$$

$$f(x) = \sum_e \int_0^{x_e} c_e(u)\,du = \underline{\tfrac{1}{2} x^T Q x} + \underline{c^T x}$$

MDP case:

$$\max_y \; r^T y$$
$$\text{s.t. } Ay = Py,\ \mathbb{1}^T y = 1\ y \geq 0$$

$r_a \longrightarrow r_a(y_a)$ potential function

replace objective

$$f(y) = \sum_a \int_0^{y_a} r_a(u)\,du \qquad \frac{\partial f}{\partial y} = r(y)^T$$

MDP congestion game
- population of players all solving an MDP
(competition among uber drivers)

some literature: mean field game
stochastic game
Dan's PhD thesis

Algorithms:

- Interior Point Methods
- Simplex Method (LP)

Gradient Descent: $\leftarrow$

$\min\limits_{x} f(x)$ – find x s.t. $\frac{\partial f}{\partial x} = 0$

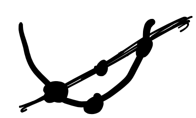
$\nabla f = \frac{\partial f}{\partial x}^T$

descent direction

$x^+ = x - \gamma \nabla f(x)$

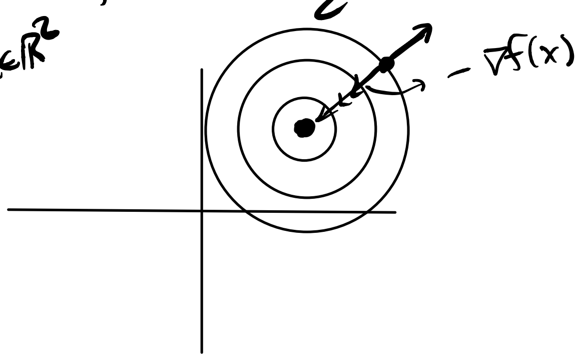
stepsize

1. compute descent direction
2. choose stepsize γ
   - fixed stepsize



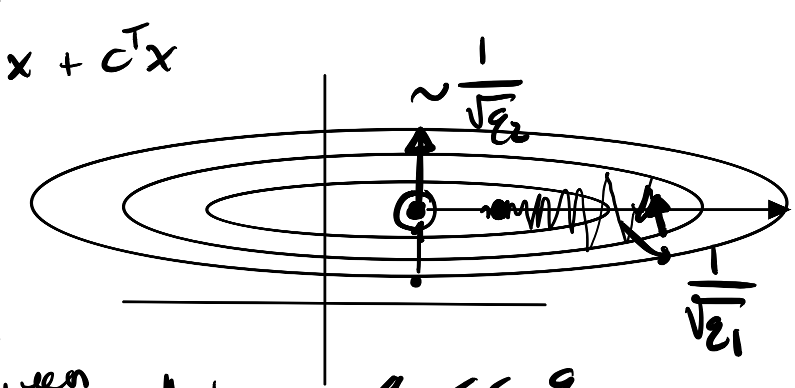
Ex. $f(x) = \frac{1}{2}x^Tx + c^Tx \leftarrow$

$x \in \mathbb{R}^2$



Ex. $f(x) = \frac{1}{2} x^T Q x + c^T x$

$Q = \begin{bmatrix} q_1 & 0 \\ 0 & q_2 \end{bmatrix}$



condition # of Q: ratio between largest & smallest eval

$q_1 << q_2$

"Q is poorly conditioned" $\Longrightarrow$ Gradient descent bad.

<u>Newton's</u> <u>Method</u>: 2nd order derivative information

$$x^+ = x - \gamma H^{-1} \nabla f(x)$$

H: Hessian $\quad H = \frac{\partial^2 f}{\partial x^2} \in \mathbb{R}^{n \times n}$ → symmetric

$$\left(\frac{\partial^2 f}{\partial x^2}\right)_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j} \qquad \underline{\frac{\partial^2 f}{\partial x_i \partial x_j}} = \underline{\frac{\partial^2 f}{\partial x_j \partial x_i}}$$

$g: \mathbb{R}^n \to \mathbb{R}^m$

$g(x) \in \mathbb{R}^m \quad x \in \mathbb{R}^n$

$$J = \frac{\partial g}{\partial x} \in \mathbb{R}^{m \times n}$$

if $g = \nabla f$

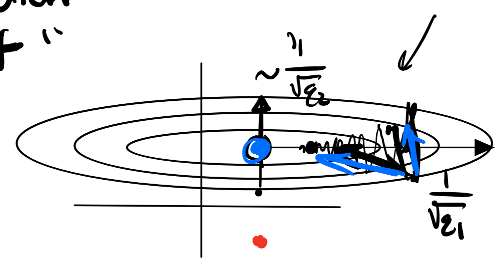
$$J = \frac{\partial g}{\partial x} = \frac{\partial^2 f}{\partial x^2} = H$$

J doesn't have to be symmetric

H: "how fast the gradient is changing"   "curvature of f"

$$x^+ = x - \gamma H^{-1} \nabla f(x)$$

"try not to go in directions with high curvature"   "do gradient descent"



if $f(x) = \frac{1}{2} x^T \underline{Q} x + c^T x$

$$\nabla f = Qx + c$$

$$H = Q$$

$$x^+ = x - \gamma H^{-1} \nabla f = x - \gamma Q^{-1}(Qx + c)$$

$$= x - \gamma(\underline{x} + \underline{Q^{-1}c})$$

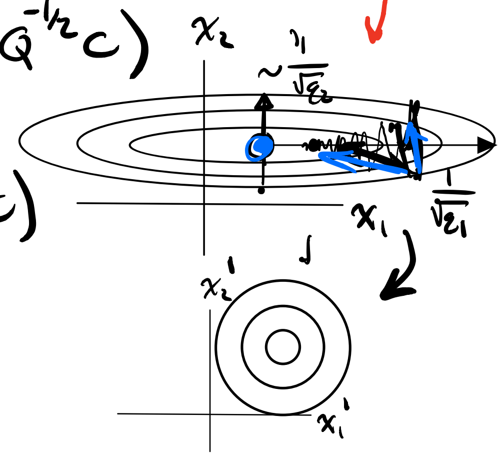
Interpret as a coord. transform $\boxed{x' = Q^{1/2} x} \Rightarrow x = Q^{-1/2} x'$

$$f(x) = f(x') = \frac{1}{2} x'^T x' + c^T Q^{-1/2} x'$$

$$x'^+ = x' - \gamma \nabla f(x') = x' - \gamma(x' + Q^{-1/2} c)$$

$$Q^{1/2} x^+ = Q^{1/2} x - \gamma(Q^{1/2} x + Q^{-1/2} c)$$

$$x^+ = x - \gamma(x + Q^{-1} c)$$



Newton's Method w Equality Constraints

$$\min_x f(x) \quad \text{s.t.} \quad g(x) = 0 \quad \longrightarrow \quad \mathcal{L} = f(x) + v^T g(x)$$

$$\frac{\partial \mathcal{L}}{\partial x} = 0 \rightarrow \text{stationarity}$$

$$\frac{\partial \mathcal{L}}{\partial v} = 0 \rightarrow \text{feasibility}$$

$$(g(x) = 0)$$

Gradient Descent or Newton's Method on $\mathcal{L}$ instead of f.

$$\frac{\partial \mathcal{L}}{\partial x, v} = \begin{bmatrix} \frac{\partial \mathcal{L}}{\partial x} & \frac{\partial \mathcal{L}}{\partial v} \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} + v^T \frac{\partial g}{\partial x}, & g(x)^T \end{bmatrix}$$

$f - v|$

$$\frac{\partial^2 \mathcal{L}}{\partial (x,v)^2} = \begin{bmatrix} \frac{\partial^2 \mathcal{L}}{\partial x^2} & \frac{\partial^2 \mathcal{L}}{\partial v \partial x} \\ \frac{\partial^2 \mathcal{L}}{\partial x \partial v} & \frac{\partial^2 \mathcal{L}}{\partial v^2} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} + \sum_i v_i \frac{\partial^2 g_i}{\partial x^2} & \frac{\partial g}{\partial x}^T \\ \frac{\partial g}{\partial x} & 0 \end{bmatrix}$$

$$= \begin{bmatrix} Q & A^T \\ A & 0 \end{bmatrix}$$

$$Q = \frac{\partial^2 f}{\partial x^2} + \sum_i v_i \frac{\partial^2 g_i}{\partial x^2}$$

$$A = \frac{\partial g}{\partial x}$$

$$\frac{\partial f}{\partial x} x = \partial f$$

$$\boxed{\left(\frac{\partial f}{\partial x}\right)^T = \nabla f}$$

if.

$$g(x) = Ax - b$$

$$\frac{\partial g}{\partial x} = A$$

$$\begin{bmatrix} x^+ \\ v^+ \end{bmatrix} = \begin{bmatrix} x \\ v \end{bmatrix} - \gamma \begin{bmatrix} Q & A^T \\ A & 0 \end{bmatrix}^{-1} \begin{bmatrix} \nabla f + A^T v \\ (Ax-b)^T \end{bmatrix}$$

computational hard part

analytically previous lecture

$\to (AQ^{-1}A^T)^{-1}$

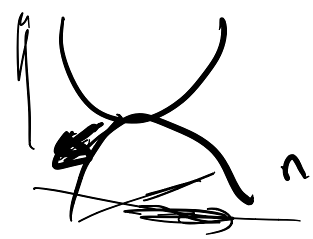
Ways to cheat in inverse computation:

Broyden Fletcher Goldfarb Shanno algorithm (BFGS)
(Hessian inverse approximation)

Davidon Fletcher Powell (DFP)

Local minimum:

- 1st order $\nabla f = 0$ → critical point "flat"
- 2nd order $\frac{\partial^2 f}{\partial x^2} > 0$ $\cup$



Interior Point Methods

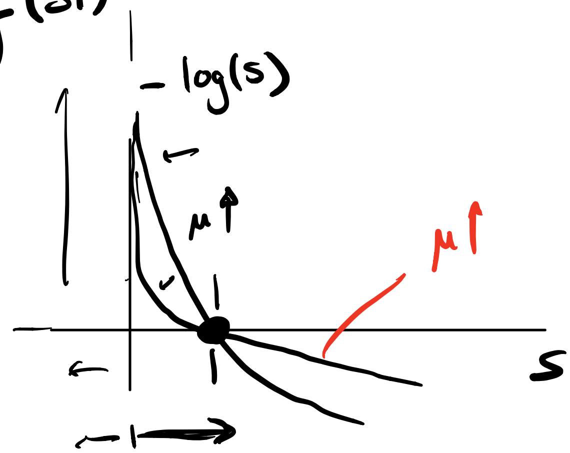
$$\min_x f(x)$$
$$\text{s.t.} \quad g(x) \geq 0$$

→ inequality constraint → equality constraint & barrier

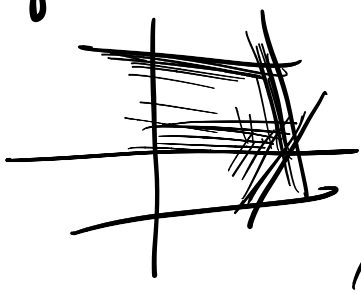
↓

$$\min_{x,s} \; \underline{f(x)} - \mu \sum_{i=1}^{m} \log(s_i)$$

$$\text{s.t.} \quad g(x) = s$$



higher dim:



$\mu \geq 0$

$$\mathcal{L} = f(x) - \mu \sum_{i=1}^{m} \log(s_i) + v^T (g(x) - s)$$

Newton's Method...

$$\frac{\partial \log}{\partial s_i} = \frac{1}{s_i}$$

$$\frac{\partial \mathcal{L}}{\partial x} = \frac{\partial f}{\partial x} + v^T \frac{\partial g}{\partial x}$$

$$\frac{\partial \mathcal{L}}{\partial s} = -\mu \left[\tfrac{1}{s_1} \cdots \tfrac{1}{s_m} \right] - v^T$$

$$= -\mu \mathbb{1}^T S^{-1} - v^T$$

$$S = \operatorname{diag}(s) = \begin{bmatrix} s_1 & & \\ & \ddots & \\ & & s_n \end{bmatrix}$$

$$\frac{\partial \mathcal{L}}{\partial v} = g(x) - s$$

$$\frac{\partial g}{\partial x} = A \qquad \frac{\partial^2 \mathcal{L}}{\partial x^2} = Q = \frac{\partial^2 f}{\partial x^2} + \sum_i v_i \frac{\partial^2 g}{\partial x}$$

$$\frac{\partial^2 \mathcal{L}}{(\partial x, s, v)^2} = \begin{bmatrix} Q & 0 & A^T \\ 0 & \mu S^{-2} & -I \\ A & -I & 0 \end{bmatrix}$$

$$\begin{pmatrix} \Delta x \\ \Delta s \\ \Delta v \end{pmatrix} = \begin{bmatrix} Q & 0 & A^T \\ 0 & \mu S^{-2} & -I \\ A & -I & 0 \end{bmatrix}^{-1} \begin{bmatrix} \nabla f + A^T v \\ -\mu S^{-1}\mathbb{1} - v \\ g(x) - s \end{bmatrix} \leftarrow$$

$$\begin{pmatrix} x^+ \\ s^+ \\ v^+ \end{pmatrix} = \begin{pmatrix} x \\ s \\ v \end{pmatrix} - \gamma \begin{pmatrix} \Delta x \\ \Delta s \\ \Delta v \end{pmatrix}$$

$$\frac{\partial \mathcal{L}}{\partial x} = \frac{\partial f}{\partial x} + v^T \frac{\partial g}{\partial x} = 0$$ stationarity constraint.

$$\left.\begin{aligned} \frac{\partial \mathcal{L}}{\partial s} &= -\mu\left[\tfrac{1}{s_1} \cdots \tfrac{1}{s_m}\right] - v^T = 0 \\ &= -\mu \mathbb{1}^T S^{-1} - v^T \end{aligned}\right\}$$

$$\frac{\partial \mathcal{L}}{\partial v} = g(x) - s = 0 \rightarrow$$ feasibility

$$-\mu \mathbb{1}^T S^{-1} - v^T = 0$$

$$-\mu \mathbb{1}^T = v^T S \Rightarrow \boxed{\mu} = v_i s_i = v_i g_i(x)$$

relaxing complementary slackness

Complementary slackness

$$v_i g_i(x) = \underline{0} \qquad v_i g_i(x) = \underline{\mu}$$

The original Lagrangian:

$$f(x) - \underline{v}^T g(x) \qquad v \geq 0 \quad \Rightarrow \left[\frac{\partial f}{\partial x} - v^T \frac{\partial g}{\partial x} = 0 \right]$$

with barrier functions

$$f(x) - \mu \sum_i \log(s_i) - v^T(g(x) - s)$$

$$\frac{\partial L}{\partial x} = \frac{\partial f}{\partial x} - v^T \frac{\partial g}{\partial x} = 0$$

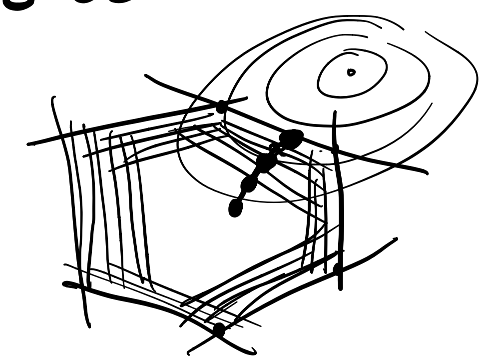
$$g_i(x) \geq 0$$

$$\frac{\partial L}{\partial s} = -\mu \mathbb{1}^T S^{-1} + v^T = 0$$

$$g_i(x) = s$$

$$\Rightarrow v^T = \mu \mathbb{1}^T S^{-1} \qquad v_i = \frac{\mu}{s_i}$$

More details about adjusting $\underline{\mu}$...



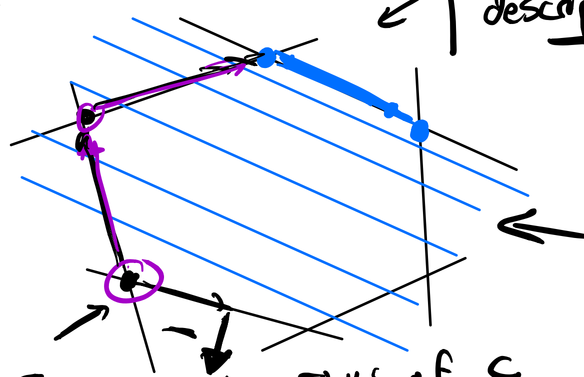
$\mu \to 0$

$x \to$ optimum

Simplex Method: row geometry

for solving linear programs.

$$\max_{x}\ r^T x \quad \text{s.t. } Ax = b,\ \cancel{Cx \geq d}$$



Dantzig, 1940's

like Gaussian Elimination w/ an objective & inequality constraints

General Form

$$\max_{z}\ r^T z$$

s.t. $Ez = f$, $Cz \geq d$

slack variables

$z = z^+ - z^-$ $\quad Cz + s = d,\ s \geq 0$

$z^+ > 0,\ z^- > 0$ $\quad \begin{bmatrix} C & -C & I \end{bmatrix}\begin{bmatrix} z^+ \\ z^- \\ s \end{bmatrix} = d$

$E(z^+ - z^-) = f$

$\begin{bmatrix} E & -E \end{bmatrix}\begin{bmatrix} z^+ \\ z^- \end{bmatrix} = f$

$z^+ > 0,\ z^- > 0$

$r^T z = \begin{bmatrix} r^T & -r^T \end{bmatrix}\begin{bmatrix} z^+ \\ z^- \end{bmatrix}$

Standard Form:

$$\max\ r^T x \quad \text{s.t. } Ax = b,\ x \geq 0$$

$$\max_x \begin{bmatrix} r^T & -r^T & 0 \end{bmatrix} x \qquad x = \begin{bmatrix} z^+ \\ z^- \\ s \end{bmatrix}$$

s.t. $\underbrace{\begin{bmatrix} E & -E & 0 \\ C & -C & I \end{bmatrix}}_{A} x = \underbrace{\begin{bmatrix} f \\ d \end{bmatrix}}_{b} \quad x \geq 0$

Geometry: (column geometry)

$$\max_x \; r^T x = F(x)$$

s.t. $[Ax = b, \; x \geq 0$

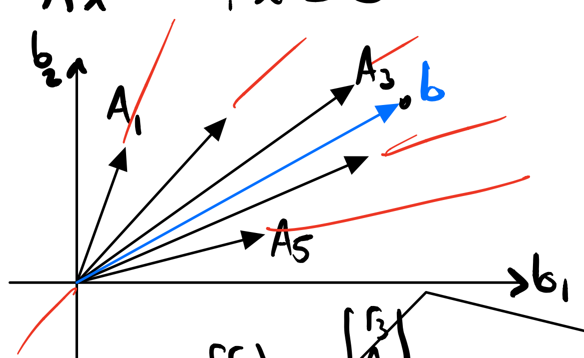
$$A = [A_1 \; A_2 \; A_3 \; A_4 \; A_5] \quad b = Ax = A_1 x_1 + \dots$$

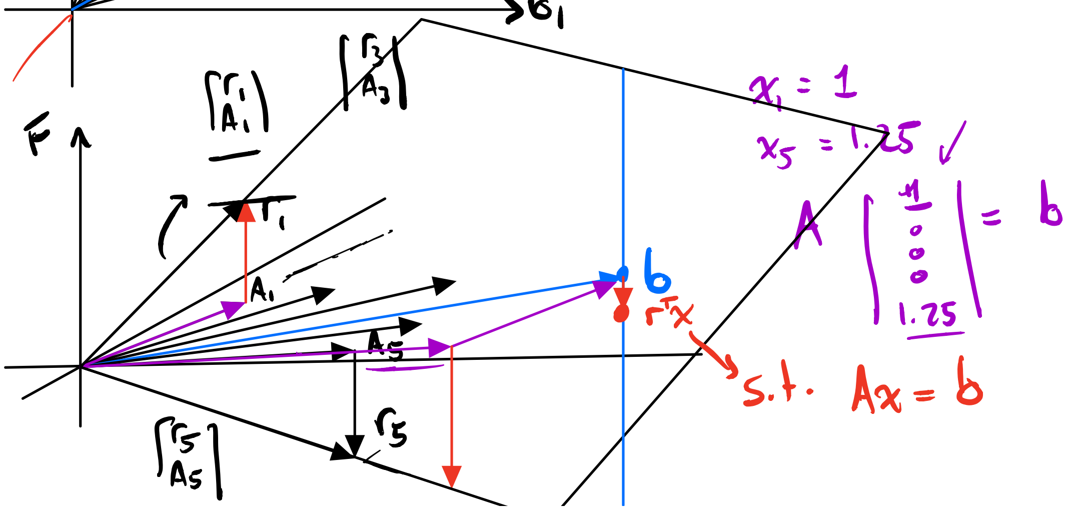
$Ax = b, \; x \geq 0$

$x \in \mathbb{R}^5$: coeffs of cols

$b \in \mathbb{R}^2$

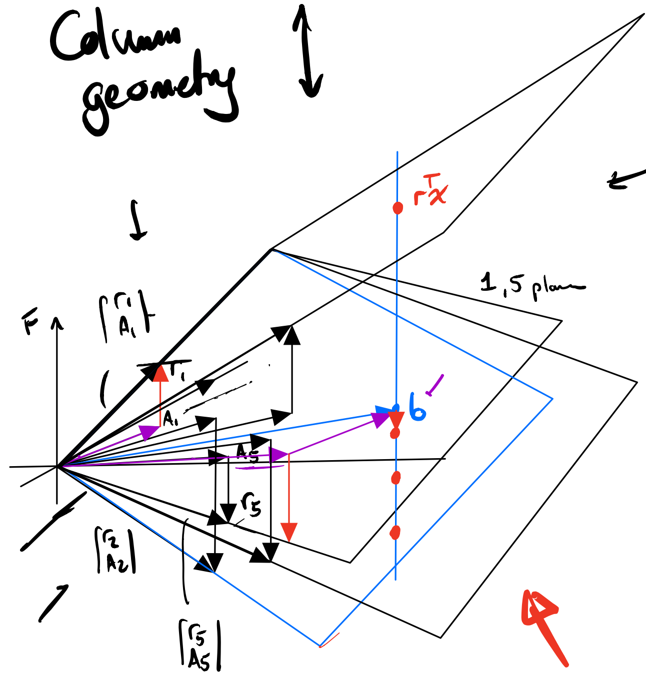
$A \in \mathbb{R}^{2\times 5}$



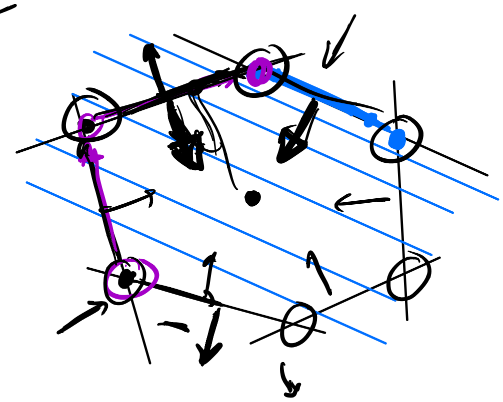


$$[r_1 \cdots r_5]\begin{bmatrix}1\\0\\0\\0\\1.25\end{bmatrix} = r_1(1) + r_5(1.25) =$$

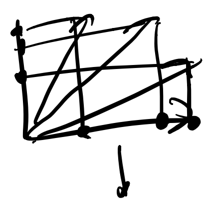
Column geometry

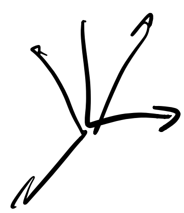


$$-\begin{bmatrix}-c_1^T-\\-c_6^T-\end{bmatrix}x = Cx \geq d$$



$$A \quad \begin{array}{c} \\ 6 \end{array}\begin{array}{c} 2 \\ \begin{bmatrix} \quad \end{bmatrix} \end{array}$$





$$Ax = b, \quad x \geq 0$$

$$A = \begin{array}{c} \\ 2 \end{array}\begin{array}{c} 3 \\ \begin{bmatrix}A_1 & A_2 & A_3\end{bmatrix} \end{array}$$

$$\max_x \quad r^T x$$
$$\text{s.t.} \quad Ax = b, \; x \geq 0$$

Tableau:

$$\begin{array}{lc} & \begin{array}{ccc} & \overbrace{\quad}^{x} & b \end{array} \\ \text{objective row} & \begin{bmatrix} 1 & -r^T & 0 \\ 0 & A & b \end{bmatrix} \\ \text{Constraint rows} \rightarrow & \end{array} \qquad \begin{bmatrix} F \\ b_1 \\ \vdots \\ b_m \end{bmatrix}$$

$$\left[\begin{array}{cccccc|c} 1 & -r_1 & -r_2 & -r_3 & -r_4 & -r_5 & 0 \\ 0 & A_1 & A_2 & A_3 & A_4 & A_5 & b_1 \\ 0 & & & & & & b_2 \end{array}\right] =$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & A_1 & A_2 \\ 0 & & \end{bmatrix} \left[\begin{array}{cccc} 1 & -r_1-r_2 & -r_3-r_4-r_5 & 0 \\ 0 & \begin{bmatrix}1 & 0 \\ 0 & 1\end{bmatrix} & \begin{bmatrix} \bar{A} \end{bmatrix} & \bar{b} \\ 0 & & & \end{array}\right]$$

$$[A_1 A_2]\,\bar{A} = [A_3 A_4 A_5]$$

$$[A_1 A_2]\,\bar{b} = b$$

$$\left[\begin{array}{cccccc|c} 1 & -r_1 & -r_2 & -r_3 & -r_4 & -r_5 & 0 \\ 0 & A_1 & A_2 & A_3 & A_4 & A_5 & b_1 \\ 0 & & & & & & b_2 \end{array}\right] =$$

$$\left[\begin{array}{cccc|c} 1 & -r_1 & -r_2 & -r_3-r_4-r_5 & 0 \\ 0 & 1 & 0 & \begin{bmatrix} \quad \end{bmatrix} & \bar{b}_1 \\ 0 & 0 & 1 & & \bar{b}_2 \end{array}\right] \longrightarrow$$

$x_1 \quad x_2$

solution

$$x = \begin{bmatrix} \bar{b}_1 \\ \bar{b}_2 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

cashing out

$$r_1\bar{b}_1 + r_2\bar{b}_2 = r_1x_1 + r_2x_2$$

$$\left[\begin{array}{cccccc|c} 1 & 0 & 0 & \boxed{+} & \boxed{+} & \boxed{+} & \boxed{} \\ 0 & 1 & 0 & \cdot & \cdot & \cdot & \bar{b}_1 \\ 0 & 0 & 1 & \cdot & \cdot & \cdot & \bar{b}_2 \end{array}\right]$$

basis cols — unused variables

reward for using x_1 & x_2 to solve $Ax = b$

height of the plane over b

how to change basis vectors

Note: r_j & b_i will keep changing w row operations

A_{ij}

$$\left[\begin{array}{cccccc|c} 1 & 0 & 0 & -r_3 & -r_4 & -r_5 & * \\ 0 & 1 & 0 & A_{13} & A_{14} & A_{15} & b_1 \\ 0 & 0 & 1 & A_{23} & A_{24} & A_{25} & b_2 \end{array}\right]$$

select 'a' pivot column with
a positive r_j

select a pivot row (the basis col
I want to swap
out)