Overview:

- Lagrangians
- KKT CONDITIONS
- DUALITY
- COMPUTE DUAL PROBLEMS
- CVXPY / CVX → SOFTWARE
- GEOMETRY OF LP
- NETWORK FLOW PROBLEMS

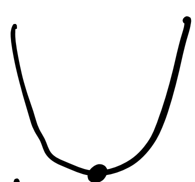
LOCAL OPTIMALITY CONDITIONS

$\min_x f(x)$

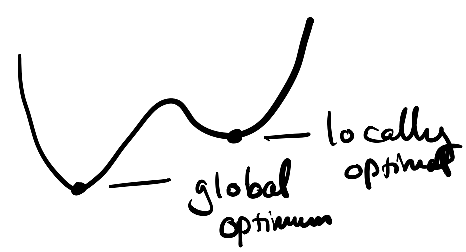
↓

convex

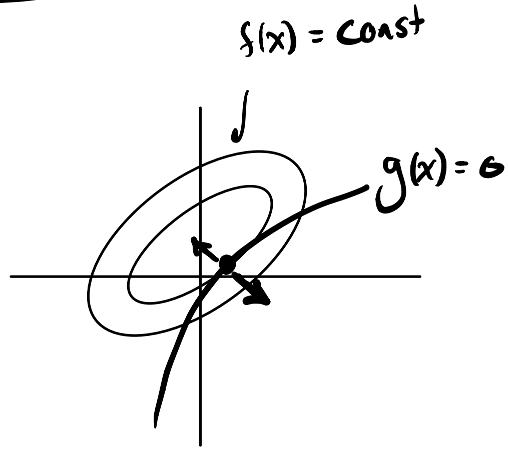
all local optima are global optima

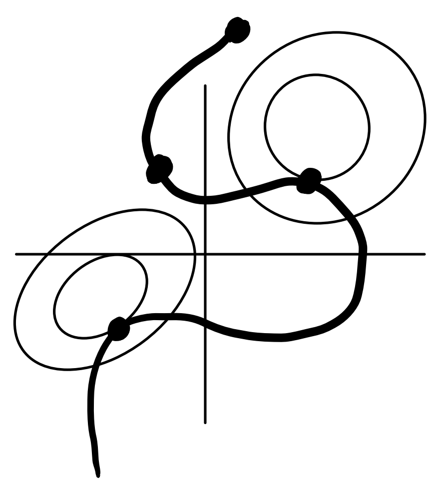


$\min_x f(x)$



KKT CONDITIONS ← LOCAL OPTIMALITY CONDITIONS





$x \in \mathbb{R}^n \quad y \in \mathbb{R}^n \qquad \alpha x + (1-\alpha) y$: Convex Combination of x & y

$0 \le \alpha \le 1$

Convexity: "bowl shaped functions"

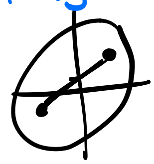
for Functions: $f: \mathbb{R}^n \to \mathbb{R}$

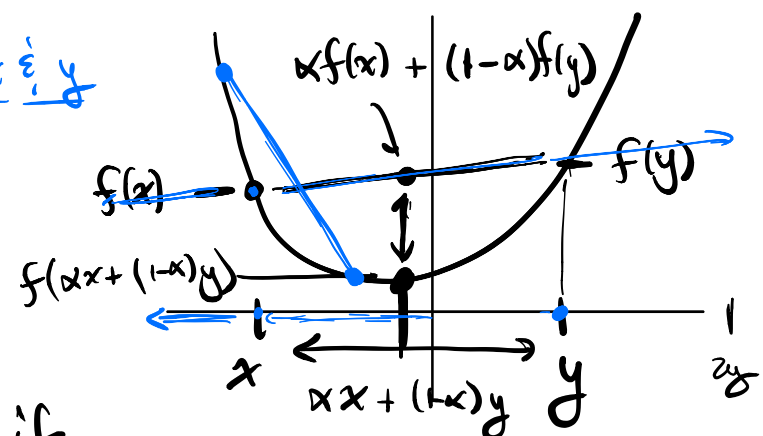
f is convex if

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

$0 \le \alpha \le 1$

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





f is concave if

$$f(\alpha x + (1-\alpha)y) \ge \alpha f(x) + (1-\alpha) f(y) \quad \forall x, y$$

Note: convex & concave = linear

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

f is strictly convex

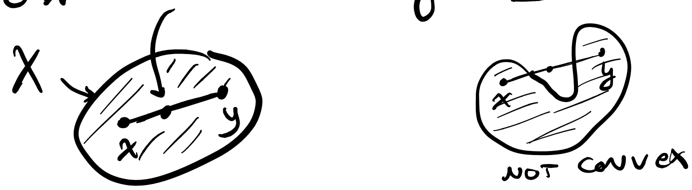
$$f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha) f(y)$$

Convexity

for sets $\mathcal{X} \subseteq \mathbb{R}^n$

$\mathcal{X}$ is convex if

$x, y \in \mathcal{X}$ $\quad \alpha x + (1-\alpha) y \in \mathcal{X}$



for a convex function f:

$\mathcal{X} = \{ x \mid f(x) \leq \text{const} \} \Longleftarrow$ convex set.

Lagrangian

$$\max_{x} f(x) \quad \leftarrow$$

$$\text{s.t.} \quad \underbrace{g(x) = 0}_{v} \quad \underbrace{h(x) \geq 0}_{\mu}$$

$$\mathcal{L}(x, v, \mu) = f(x) + v^T g(x) + \mu^T h(x)$$

Competition between x & v, μ

$$\max_{x} \min_{\underbrace{v, \mu \geq 0}} f(x) + v^T g(x) + \mu^T h(x)$$

force us to follow constraints

$$\max_{x} \min_{v,\mu \geq 0} f(x) + v^T \underline{g(x)} + \mu^T h(x) \Leftarrow$$

$g(x) \neq 0$ $\quad v$ can push $\mathcal{L}$ to $-\infty$

we have to choose x s.t. $\underline{g(x) = 0}$

$h(x) < 0$ $\quad \mu \geq 0$ can push $\underline{\mathcal{L}}$ to $\underline{-\infty}$

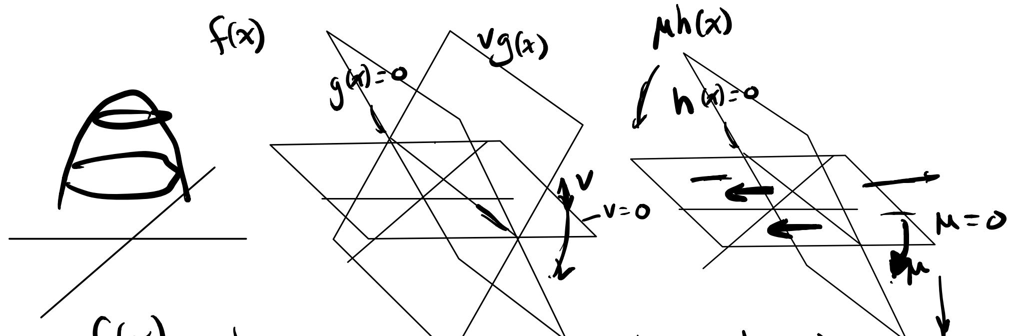
we have to choose x s.t. $\underline{h(x) \geq 0}$

Intuitively...

$$\max_{x} \; f(x) + I_v(g(x)) + I_\mu(h(x))$$

$$I_v = \begin{cases} -\infty & \text{if } g(x) \neq 0 \\ 0 & \text{if } g(x) = 0 \end{cases} \qquad I_\mu = \begin{cases} -\infty & \text{if } h(x) < 0 \\ 0 & \text{if } h(x) \geq 0 \end{cases}$$

Geometrically...



$f(x) \quad + \quad v\,g(x) \quad + \quad \mu\,h(x)$

CLAIM:

$$\max_x \left(\min_v L(x,v) \right) \le \min_v \left(\max_x L(x,v) \right)$$

PROOF: very general argument

$$L(x,v) \le \max_x L(x,v)$$

$$\forall x \ \forall v \qquad \forall v$$

$$\min_v \Big(\underset{\forall x}{L(x,v)} \Big) \le \min_v \Big(\max_x L(x,v) \Big)$$

true for any x ...

pick x that maximizes the LHS...

$$\max_x \left(\min_v L(x,v) \right) \le \min_v \left(\max_x L(x,v) \right)$$

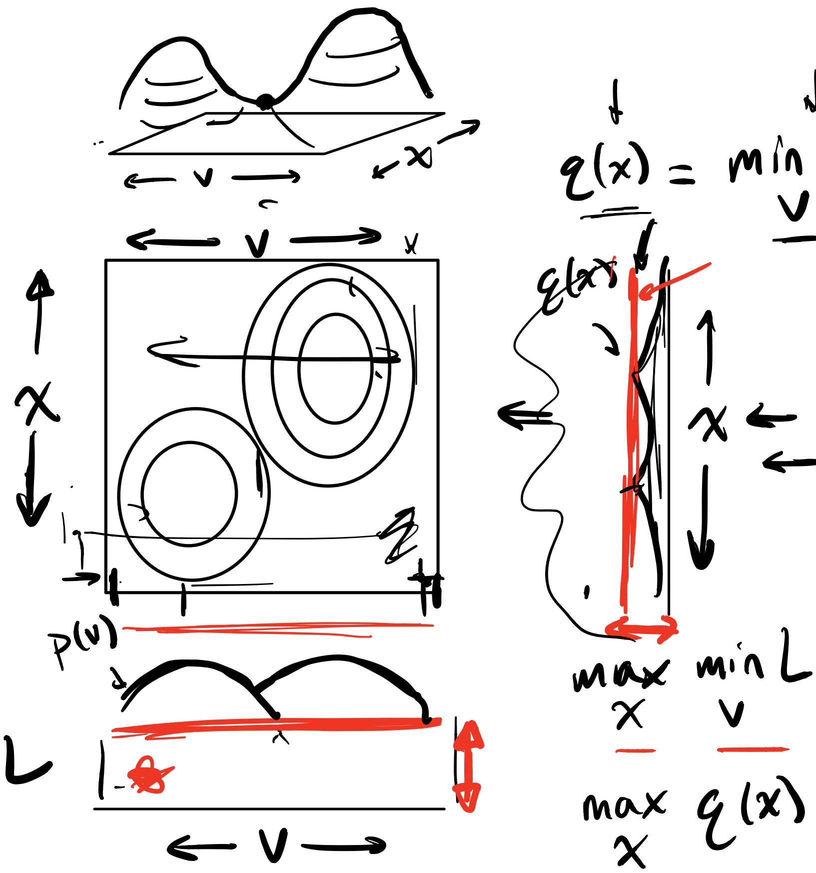
Ex.

$$\begin{array}{c|c|c} & \multicolumn{2}{c}{\longleftarrow v \longrightarrow} \\ \hline x \uparrow & L=1 & L=10 \\ \downarrow & L=20 & L=2 \end{array} \qquad \begin{array}{c} \left(\min_v L \right)(x) \\ \hline 1 \\ 2 \end{array} \qquad \max_x \left(\min_v L \right) = 2$$

$$\begin{array}{c|c} 20 & 10 \end{array} \longleftarrow \left(\max_x L \right)(v)$$

$$\min_v \left(\max_x L \right) = \boxed{10}$$

Ex:



$$q(x) = \min_{v} L(x,v)$$

$$\max_{x} L$$

$$\max_{x} \min_{v} L(x,v)$$

$$\max_{x} q(x)$$

$$p(v) = \max_{x} L(x,v)$$

$$\min_{v} \max_{x} L = \min_{v} p(v)$$

"the lowest hill is higher than the highest valley" ?

$$p^* = \underbrace{\max_x \left(\min_v L(x,v)\right)}_{\text{original optimization}} \le \underbrace{\min_v \left(\max_x L(x,v)\right)}_{} = d^*$$

primal opt. problem

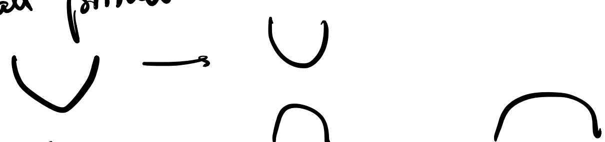
dual opt problem

Duality

$$\boxed{d^* - p^* \ge 0}$$ duality gap

Note: could have switched the role of primal and dual

$$p^* = \underbrace{\min_x \max_v L(x,v)}_{\text{call primal}} \ge \underbrace{\max_v \min_x L(x,v)}_{\text{dual}} = d^*$$



$$p^* - d^* \ge 0$$

Strong duality: $p^* = d^* \rightarrow$ duality gap $= 0$

for convex problems
+ technical conditions
(constraint qualifications) $\Rightarrow$ strongly dual

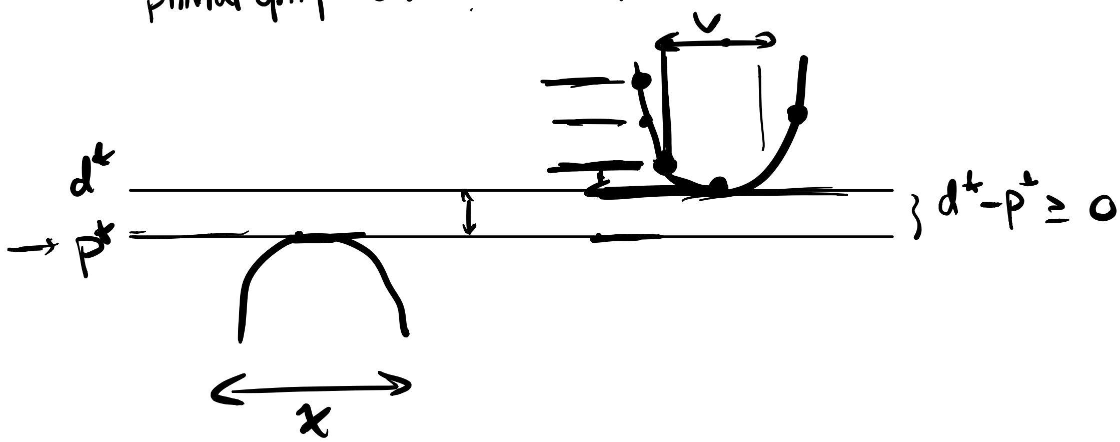
Dual problem can give us information even before we solve it...

$$p^* = \max_x \left(\min_v L(x,v)\right) \leq \min_v\left(\max_x L(x,v)\right) = \underline{d^*}$$

original optimization

primal opt. problem

dual opt problem



COMPUTING DUAL PROBLEMS:

LINEAR PROGRAM (LP)

PRIMAL $\max_{x,s} \; r^T x = f(x)$

s.t. $\underbrace{Ax = b}_{v}, \quad \underbrace{Cx = s + d}_{w}, \quad \underbrace{s \geq 0}_{\mu}$

Lagrangian:

$$\mathcal{L}(x,s,v,w,\mu) =$$

$$r^T x + v^T(Ax - b) + w^T(Cx - s - d) + \mu^T s$$

$$\max_{x,s\geq 0} \; \min_{v,w,\mu\geq 0} \mathcal{L} \Longrightarrow \min_{v,w,\mu\geq 0} \left(\max_{x,s\geq 0} \mathcal{L} \right)$$

$$(r^T + v^T A + w^T C)x + (-w^T + \mu^T)s \underbrace{- v^T b - w^T d}_{q(v,w)}$$

$\max_{x,s\geq 0} \mathcal{L}$

using $\frac{\partial \mathcal{L}}{\partial x} = 0$ and $\frac{\partial \mathcal{L}}{\partial s} = 0$

$$\frac{\partial \mathcal{L}}{\partial x} = r^T + v^T A + w^T C = 0 \qquad \frac{\partial \mathcal{L}}{\partial s} = -w^T + \mu^T = 0$$

$$w^T = \mu^T$$

Dual PROBLEM (LP)

$$\min_{v,\mu} \; -v^T b - \mu^T d = g(v,\mu)$$

$$\underbrace{r^T + v^T A + \mu^T C = 0}_{x} \qquad \underbrace{\mu \geq 0}_{s}$$

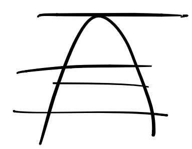
rewrite $\mathcal{L}$

$$\mathcal{L}(x, s, v, \mu) = -v^T b - \mu^T d + (r^T + v^T A + \mu^T C)x + \mu^T s$$

Quadratic Program $\qquad Q = Q^T \succ 0$

Primal

$$\boxed{\begin{aligned} &\min_x \; \tfrac{1}{2} x^T Q x + c^T x \\ &\text{s.t.} \quad \underbrace{Ax = b}_{v}, \quad \underbrace{Cx \geq d}_{\mu} \end{aligned}}$$



$$\mathcal{L}(x, v, \mu) = \tfrac{1}{2} x^T Q x + c^T x + v^T(Ax - b) + \mu^T(Cx - d)$$

$$\min_x \max_{v, \mu \geq 0} \mathcal{L}$$

$$\max_{v, \mu \geq 0} \left(\min_x \mathcal{L} \right) \qquad \min_x \mathcal{L} \Rightarrow \frac{\partial \mathcal{L}}{\partial x} = 0$$

$$\frac{\partial \mathcal{L}}{\partial x} = x^T Q + c^T + v^T A + \mu^T C = 0$$

$$\Rightarrow x^T = -(c^T + v^T A + \mu^T C) Q^{-1}$$

call $z^T = c^T + v^T A + \mu^T C$

$$\mathcal{L}(x,v,\mu) = \frac{1}{2}x^TQx + (c^T + v^TA + \mu^TC)x - v^Tb - \mu^Td$$

$\min_x \mathcal{L}$ ← plug in x as a function of v & μ

$$\frac{1}{2}\underbrace{z^TQ^{-1}}_{x^T}QQ^{-1}z - z^TQ^{-1}z - v^Tb - \mu^Td$$

$$\frac{1}{2}z^TQ^{-1}z - z^TQ^{-1}z - v^Tb - \mu^Td$$

$\min_x \mathcal{L} =$

DUAL PROBLEM:

$$\boxed{\begin{aligned} \max_{z,v,\mu}\quad & -\frac{1}{2}z^TQ^{-1}z - v^Tb - \mu^Td \\ \text{s.t.}\quad & z^T = c^T + v^TA + \mu^TC,\ \mu \geq 0 \end{aligned}}$$

$\underline{Q} \succ 0 \Rightarrow \underline{Q}^{-1} \succ 0$

<u>Solving</u> convex problems.

cvxpy : cvx for python

<u>CVX</u> : Matlab version

Simple LP:

OPTIMIZATION ON A SIMPLEX:

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

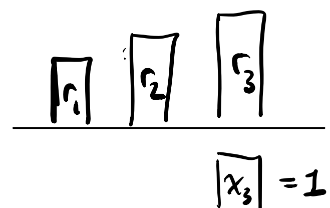
$$= \{ x \in \mathbb{R}^n \mid \mathbb{1}^T x = 1 ,\ x \geq 0 \}$$

PRIMAL

variables

mass or money distribution

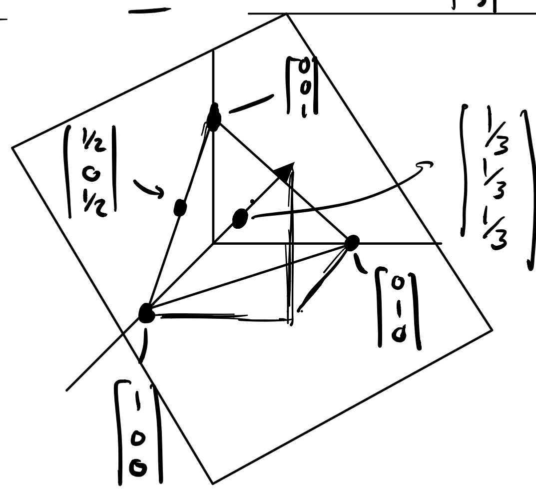
$$\boxed{\begin{aligned} \max_x \quad & r^T x = \sum_i r_i x_i \\ \text{s.t.} \quad & \mathbb{1}^T x = 1 ,\ x \geq 0 \end{aligned}}$$

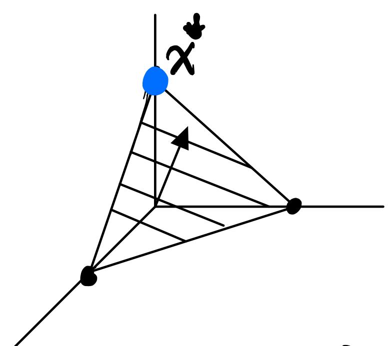


r: reward vector

$$r^T = [r_1 , r_2 \cdots r_n]$$

$$\left.\begin{aligned} \mathbb{1}^T x &= 1 \\ x &\geq 0 \end{aligned}\right]$$





$$f(x) = r^T x \qquad \frac{\partial f}{\partial x} = r^T$$

for linear functions: const gradient

$$\max_{x} \; r^T x$$

s.t. $\underbrace{\mathbb{1}^T x = 1}_{\lambda \in \mathbb{R}}$, $\underbrace{x \geq 0}_{\mu}$

$\mu \in \mathbb{R}^n_+$ $(\mu \geq 0)$

$$\mathcal{L}(x, \lambda, \mu) = r^T x - \lambda(\mathbb{1}^T x - 1) + \mu^T x$$

$$= (r^T - \lambda \mathbb{1}^T + \mu^T)\, x \; \boxed{+\lambda}$$

$$\frac{\partial \mathcal{L}}{\partial x} = r^T + \lambda \mathbb{1}^T + \mu^T = 0$$

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

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

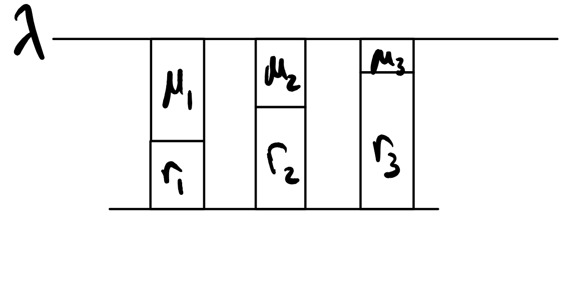
Dual Prob.

prices or rewards or inefficiencies

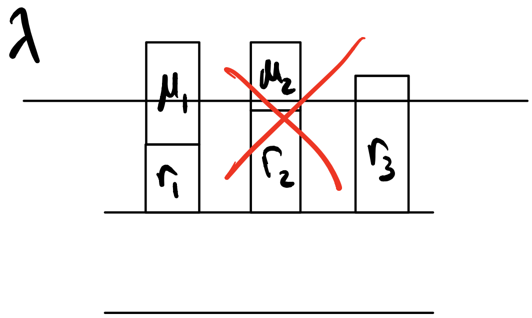
$$\boxed{\begin{array}{ll} \min\limits_{\lambda, \mu} & \lambda \\ \text{s.t.} & \lambda \mathbb{1}^T = r^T + \mu^T, \quad \mu \geq 0 \end{array}}$$

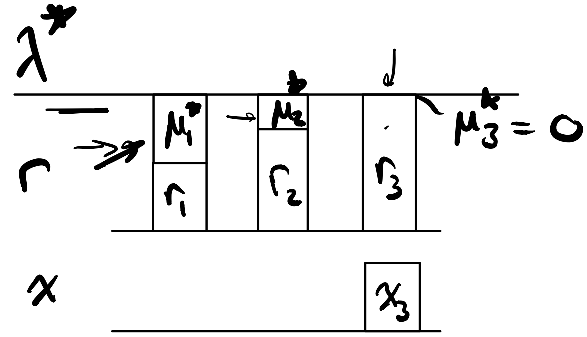
$$[\lambda \;\; \lambda \;\; \lambda] = [r_1 \;\; r_2 \;\; r_3] + [\mu_1 \;\; \mu_2 \;\; \mu_3]$$

$$\lambda = r_i + \mu_i$$



$$\lambda \mathbb{1}^T \geq r^T$$





x : dual variable for $\lambda \mathbb{1}^T = r^T + \mu^T$

Complementary slackness

$$x_i \mu_i = 0$$

$\mu_1 > 0 \Rightarrow x_1 = 0 \longrightarrow x_3 = 1$

$\mu_2 > 0 \Rightarrow x_2 = 0$

---

x_i : mass on option i
money on option i

$\mathbb{1}^T x = 1$ budget constraint

$x \geq 0$ spend my own money

r_i : value of good i

λ : upper bound on payoff

λ^* : optimal payoff

μ_i : inefficiency of option i ex. $\mu_1^* = r_3 - r_1$
or
regret for choosing i

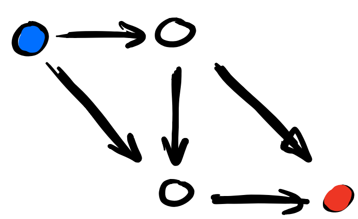
for best option, j, $\mu_j^* = 0$
best option totally efficient

$\mu_i^* x_i^* = 0$: complementary slackness
"at optimum, no inefficient options are chosen."

Network Flow

Applications:

- shortest path
- traffic flow
  - road traffic
  - cyber traffic
- (circuit theory)



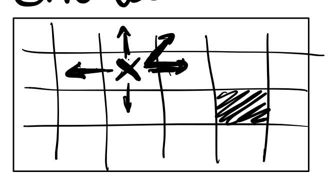
Markov Decision Processes

"stochastic network flow"
transitions between network locations not deterministic

used for time dependent discrete decision making problems

Examples

- Grid world

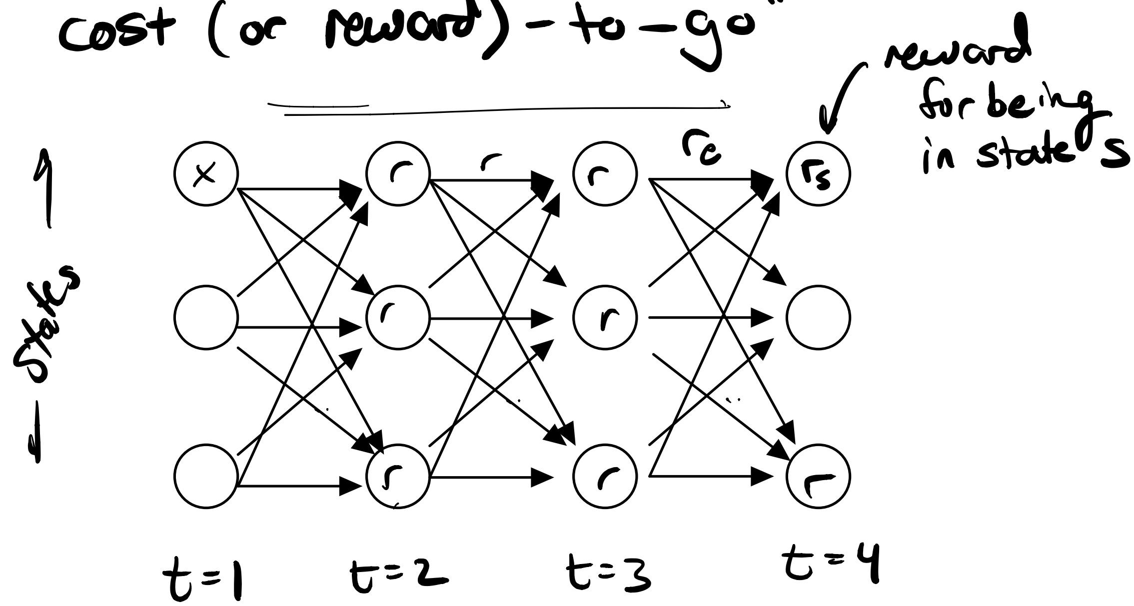


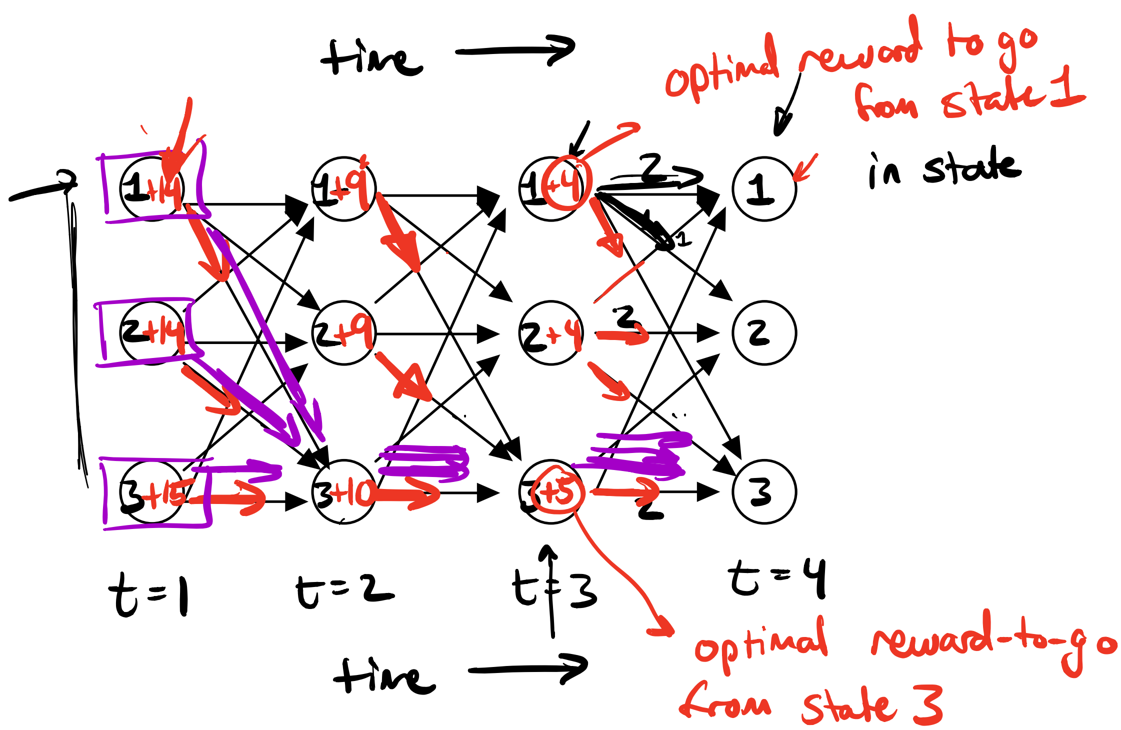
- Chess
- GO → AlphaGo
- Reinforcement Learning

Dynamic Programming:

" for time dependent optimization... Solve backwards from the end. "

"cost (or reward) - to - go"





Bellman Equation:

states s, actions at state s A_s

$a \in A_s$

for ea $a \in A_s \longrightarrow s'$

optimal reward-to-go from state s at time t: $V_s^*(t)$

WILL CLEAN UP NEXT WEEK

$$V_s^*(t) = \max_{a \in A_s} \left(r_a + V_{s'}^*(t+1) \right)$$

for a transitioning from s to s'

Convex duality in these mass flow problems...

Primal Problem $\longleftrightarrow$ dual problem

mass flowing forward in time (mass conservation equation)

cost/reward information flowing backward in time (Bellman eqn) state S : $|S|$

time T : $|T|$

actions A : $|A|$

$|S||T||A|$: dynamic prog.

$(|S||A|)^{|T|}$