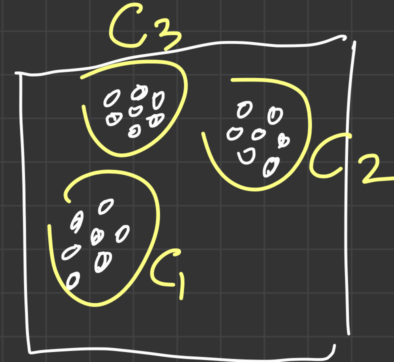
MATH 462

Today RL

5.11.2021

Side-Note Unsupervised Learning.

clustering.



Given $S_m = \{x_1, \dots, x_m\}$

$x_i \in \mathbb{R}^d$

Given k = predefn # clusters

Goal assign $y \in \mathcal{Y}_k$ to each x.

Given $d(x_i, x_j)$ pairwise distances.

Intuition $C_k = \{x_i \mid y(x_i) = k\}$

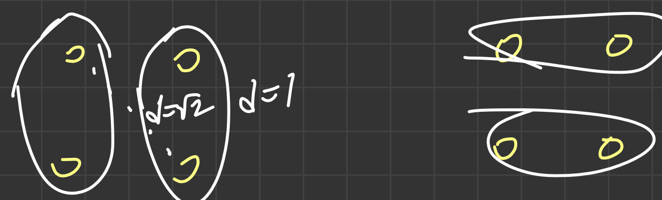
want $d(x_i, x_j)$ small $y(x_i) = y(x_j)$

$d(x_i, x_j)$ large if $y(x_i) \neq y(x_j)$

Theory Clustering has a loss. But nonconvex
so multiple local minima possible

practice simple algorithms
work well with random
initialization.
(sometimes run more than once)

Thm Find Global min of cluster loss
ill posed



Any reasonable cluster loss
should be decreasing in $d(x,x')$ $x,x' \in C$
increasing in $d(x,x')$ $x \in C$, $x' \in D$, $C \neq D$

Given assignment of $C_1, \dots, C_K$

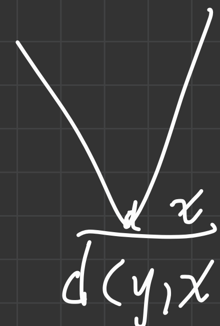
$$L_S(C) = \operatorname*{mean}_{x',x \in C} \left(d(x,x') \right)$$

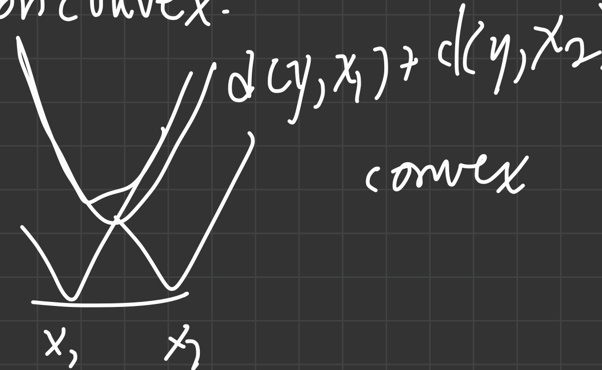
$$L_D(C,D) = \operatorname*{mean}_{x \in C,\ x' \in D} d(x,x')$$

Cluster Loss

$$\operatorname*{mean}_{C_i} L_S(C_i) - \operatorname*{mean}_{C \neq D} L_D(C,D)$$

Math why nonconvex.





convex

-dist concave.

Loss convex-convex not convex.

Math : Centroidal Voronoi Tesselation

Clustering good math.

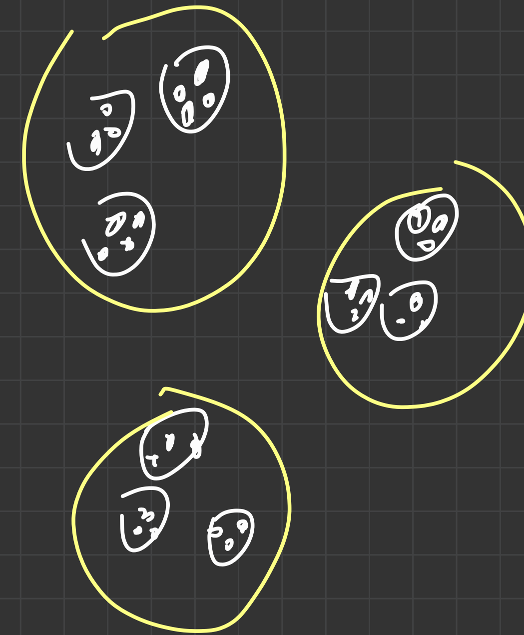
Semi Supervised Learning.

$$S_m = \{(x_1, y_1), \dots, (x_m, y_m)\}$$

$$m = 1000$$

$$S_n^u = \{x_1, \dots, x_n\}$$

$$n = 10^6$$

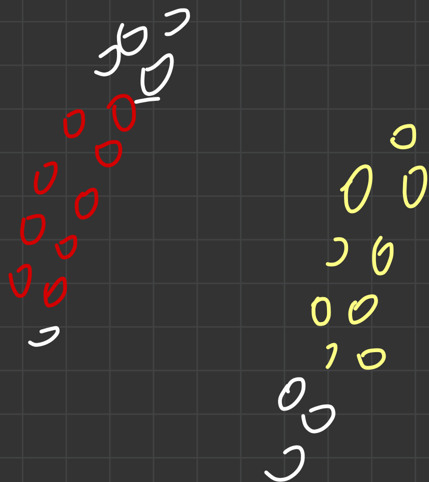


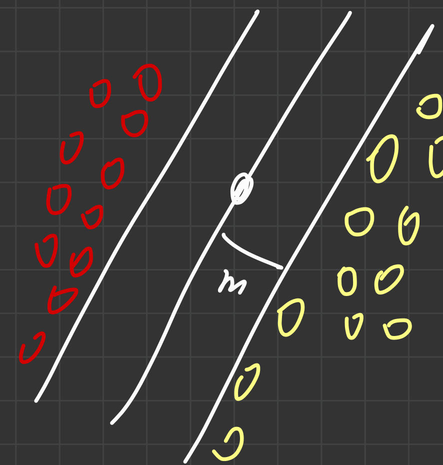
How to effectively use S_n^u ?

Theory / practise?

SVM margin loss

linear classification





SS SVM

unlabeled.

① classify with lables.

② predict each unl. x_i

③ go back and find best margin classifi using presumed labels

SimCLR

semisupervi feat. rep learning
using data augmentation.

RL

practise

✓ · chess
✓ · Go 2017
✓ · arcade ATARI.

× RL drive car

Deep RL vs. Small RL

Theory

✓ small RL

not capturing important aspects of problem

RL by example.

Deterministic
Stochastic

Reduce 2 player game to 1

Reduce ∞-horizon to finite

Back optimal & stoch control theory.

RL games

Game Theory Reduction. opt.

POV player 1.
RL max reward for my action given environ.
GAME: " " give player 2, min. my reward

$$\left[\max_{a, p1} \min_{a, p2} \text{Reward to P1} \right]$$

zero sum game.

Reduct of Game to RL:

model for P2 :

ex $P2$ play randomly $p = \frac{1}{3}, \frac{1}{3}, \frac{1}{3}$ for R, P, S

EX $P2$: R, P, S with $\frac{3}{6}$, $\frac{2}{6}$, $\frac{1}{6}$

winning states for P1 : $\frac{3}{6}$ $\frac{2}{6}$ $\frac{1}{6}$

a : R P S with prob. p_r p_p p_s

x = P2 R P S

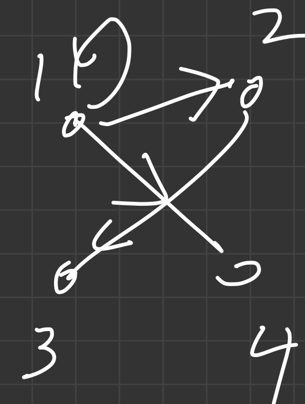
$\text{prob}(R, R) = p^1_R \, p^2_R$ etc

P1 \ P2	R	P	S	
R	0	-1	+1	$p^1_R \, p^2_R$ ...
P	+1	0	-1	⋮
S	-1	+1	0	⋮

EXP Reward

RPS   Stochastic
expectations

Control   Bellman



Abstract   Graph of states

$S = \{1, 2, 3, 4\}$

$a(1) = \{1, 2, 4\}$

$a(2) = \{3\}$

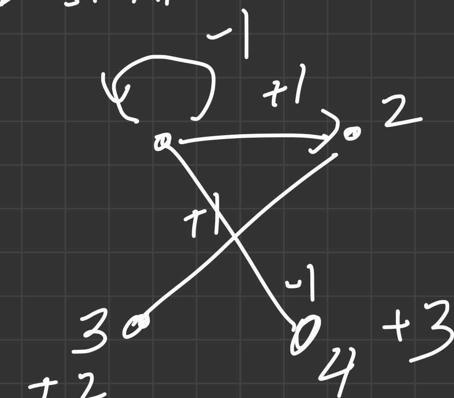
$a(3)$ = terminal game over

$a(4)$ = " "

Reward action

Reward state

search



Convert games / RL problem
which may not end to "finit" ones

Terminal time: end after T_f actions.

OR Infinite Horizon w. discounting.

Reward r

Becomes $(0.99)^t r$

want to reach terminal action.

payoff $+1$.

get there quickly: $(0.99)^{10} \cdot 1$

slowly $(0.99)^{100} 1$.

Defined surrogate loss ($\gamma = 0.99$ γ^t)
designed to encourage shorter solns.

Surrogates
— (Early RL & games ✓
Video games ✗
problems

Chess payoff +1 −1 ½ draw.

value to position [−1, +1]

easy position : win / lose / draw (assuming optimal play)
states

sum score not Reward.

$X \xrightarrow{\text{features}} f(X) \rightarrow$ approximated value fn $V(f(X))$

board

Bellman Eqn

look ahead many moves

$X_{T+k}(a_1(\cdot), a_2(\cdot))$

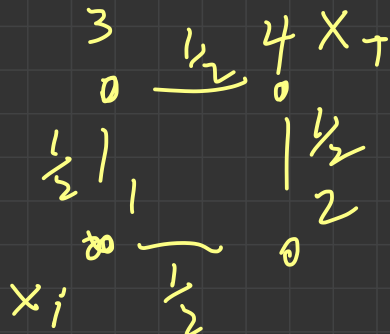
many actions.

choose Best for each player.

Combine the value at future states.

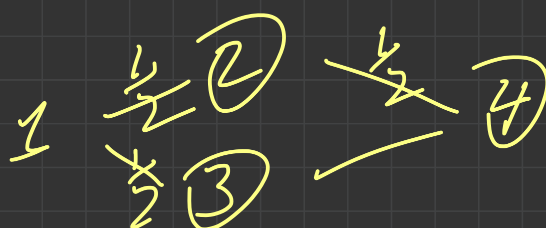
...rm start at 1
ends at 4

$R = 4$



Random

Reward



$p(1,4)$ in 2 steps $= P_{14} = \frac{1}{2}$

$p(1,1)$ " " $= P_{11} = \frac{1}{2}$

$$R(1) = P_{14} R_4 \left(\tfrac{1}{3}\right)^2 + P_{11} R(1) \left(\tfrac{1}{3}\right)^2$$

$$R(1) = \frac{1}{2} \cdot \frac{4}{9} + \frac{1}{2} R(1)/9$$

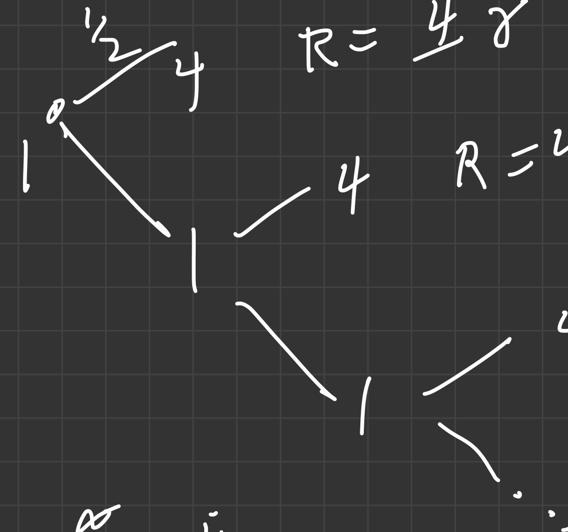
$$\left(\tfrac{1}{3}\right)^T R_4$$

$$R(1)\left(1 - \tfrac{1}{18}\right) = \frac{1}{2} \cdot \frac{4}{9}$$

$$R(1) = \tfrac{2}{9} \Big/ \tfrac{17}{18}$$

$P(1 \to 4)$ in 2 steps $= \frac{1}{2}$

$P(1 \to 1)$ in 2 step $= \frac{1}{2}$



$R = 4\gamma$

$R = 4\gamma^2$

$4\gamma^3$

$\gamma = \left(\frac{1}{3}\right)^2$

$$R(1) = 4\sum^{\infty} \gamma^i$$

$$= \frac{4}{1-\gamma}$$

Bryson Ex.

4 nbrs each state

control