ECSE 506: Stochastic Control and Decision Theory

Aditya Mahajan

6 Jan 2022

Learning objectives

Develop ability to read and understand research papers in stochastic control.

Emphasis on understanding proofs. We will prove **every** result that we state in class.

Study examples from different application domains: communications, operations research, control systems, and power systems. Focus on being able to establish qualitative properties of optimal policies

Understand the role and limitations of models.

Course content

Stochastic optimization

Single decision made by single decision maker.

MDPs (Markov decision processes

Multiple decisions made by single DM with perfect information

POMDPs (Partially observable MDPs

Multiple decisions made by a single DM with imperfect information.

Decentralized control (also called Dec-POMDPs

Multiple decisions made by multiple DMs with imperfect info.

Background

Graduate probability

Conceptual understanding of random variables and conditional expectation

Real analysis

Basic understanding of limits and convergence, metric spaces, and completeness.

Optimization

Basic understanding of convexity and first and second order conditions for optimality.

Logistics

Assignments (20%)

- Weekly assignments; posted on the course website.
- Only one randomly selected question will be graded. Lowest assignment dropped.
- Solutions posted on myCourses and only accessible to registered students.
- If you are auditing and need access, send me a message.

Mid Term (40%)

- **If classes are online: Week of 28th March**
  Online exam. Available for 72hrs. Once you start, you'll have 2.5hrs to finish the exam.
- **If classes are in-person: 29th March**
  In person, 1.5 hr exam, during class time.

Term Project (20%)

- To be done either alone or in groups of two. Due end of term
- Critique one or two papers related to the course. Deliverables: project report and presentation.

Course Notes

Partial course notes available on the course website:
`https://adityam.github.io/stochastic-control`

While classes are online, the zoom recording will be available on myCourses.

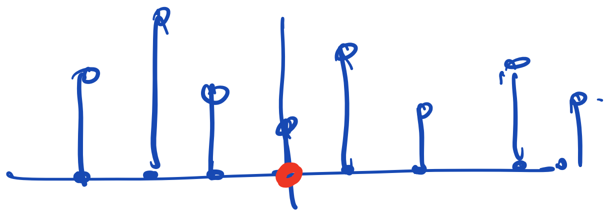
If/when in-person classes resume, video recordings may not be available. (Most rooms don't have video recording infrastructure).

Communication

Announcements and solutions posted on myCourses. Please check regularly.

All communication with the instruction should be via the discussion board on myCoures.

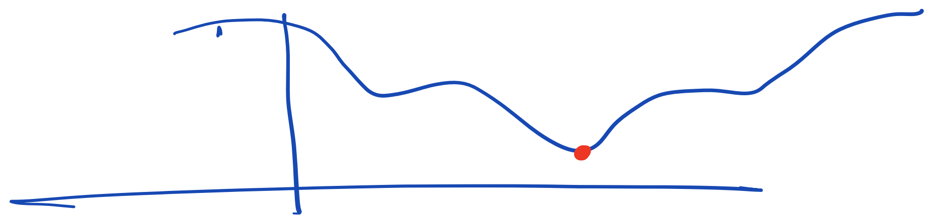
OPTIMIZATION



Brute force search

for $x \in \mathcal{X}$
  if $f(\text{argmin}) > f(x)$
    argmin $= x$
  end
end.

Given $f: \mathcal{X} \to \mathbb{R}$.



Find $x^* = \underset{x \in \mathcal{X}}{\arg\min} f(x)$

$\mathcal{X}$: Is cts. Calculus

Regularity cond. : convexity.

$\mathcal{X}$: Is discrete.

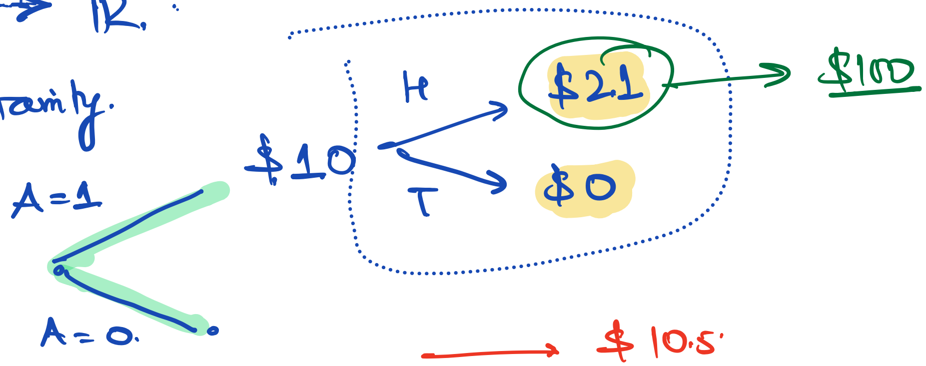
Decision prob vs optimization

INFORMATION / UNCERTAINTY.

Deterministic optimization $\quad \min_{a \in A} c(a)$, $c : A \to \mathbb{R}$

Cost fn — Action space — Real #

$c : \underset{\text{Set of uncertainty}}{W} \times A \to \mathbb{R}$



Assuming : P_W is known

$10 · 1	~~$20~~ $21 · 0.5
	$0 · 0.5
✓	✓
$30 · 0.6	$21 0.5
$0 · 0.4	$0 0.5
✓	✓

$a_1 > a_2$

$a_2 > a_3$

$a_1 > a_3$

$\Longrightarrow$ Consistent

Utility maximizers

If a DM has consistent pref over a random outcomes.

$\Rightarrow$ $\exists$ a utility fn. s.t.

DM is a utility maximizer

Unique up to additive and multiplicative consts

$$\underset{a \in A}{\arg\max}\ \mathbb{E}\left[10\,u\left(c(w,a)\right) + 3\right]$$

$c : W \times A \to \mathbb{R}$

$$\underset{a \in A}{\arg\min}\ \mathbb{E}\left[c(w,a)\right]$$

$$\underset{a \in A}{\arg\max}\ \mathbb{E}\left[r(w,a)\right]$$

\$10 — \$21 (0.5), \$0 (0.5)

$A = 1$, $A = 0$

$\boxed{10.5 - 10 = 0.5\text{\$} \qquad = 0}$

\$100 — \$210 (0.5), \$0 (0.5)

$A_1 = 1$ \$5 10

$A = 0$ \$0

$\arg\max \{ \underline{0.5}, 0 \}$ $\arg\max \{ \underline{5}, 0 \}$

$A = 1$ $-\$0.5$ $\arg\max \{ -0.5, \underline{0} \}$

$A = 0$ 0

$$\arg\max \{ \theta_1, \theta_2, \ldots \theta_k \} \equiv \arg\max \{ \alpha\theta_1, \alpha\theta_2 \ldots \alpha\theta_k \}$$

if $\underline{\alpha > 0}$

$$\equiv \arg\max \{ \alpha\theta_1 + \underline{\beta}, \alpha\theta_2 + \underline{\beta}, \ldots \}$$

Deterministic optimization $\min_{a \in A} \underline{c(a)}$

Stochastic optimization $\min_{a \in A} \underbrace{\mathbb{E}[c(W,a)]}_{C(a)}$

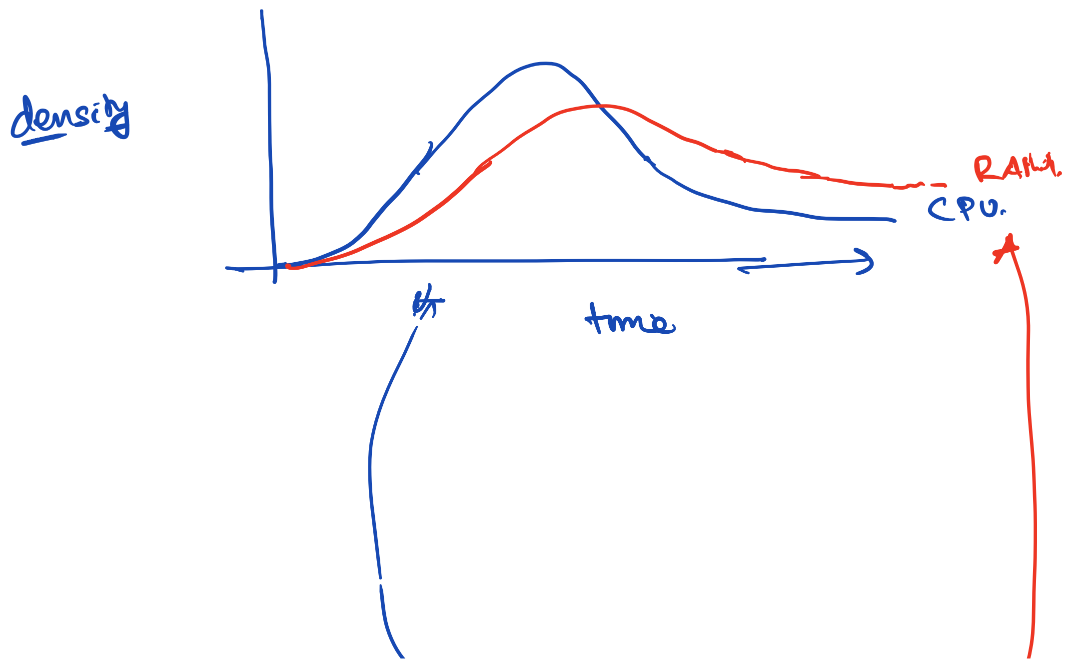
Example

CPU/RAM

A = 1

A = 2.

Max the time to failure

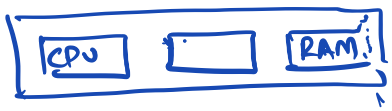


w^{CPU} $\qquad$ w^{RAM} ✓ $\qquad$ $W = \{w_1^{CPU}, w_2^{CPU}, w_1^{RAM}, w_2^{RAM}\}$

$w_1^{CPU} \quad w_2^{CPU} \qquad w_1^{RAM}, \ w_2^{RAM}$

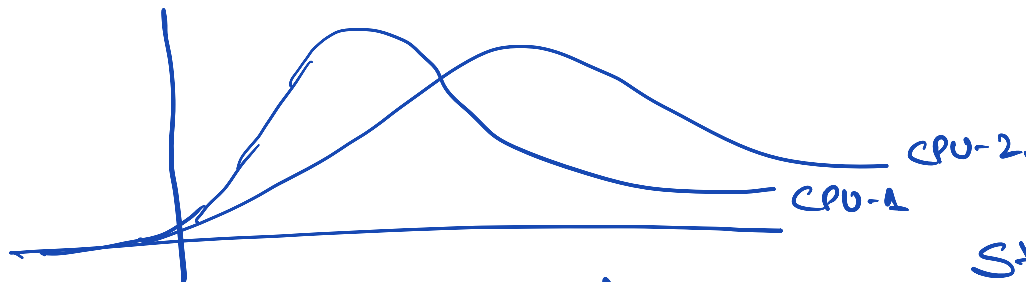
$$\mathbb{E}[r(W, \underline{A=1})] = \mathbb{E}[\min\{\max\{w_1^{CPU}, w_2^{CPU}\}, w_1^{RAM}\}]$$

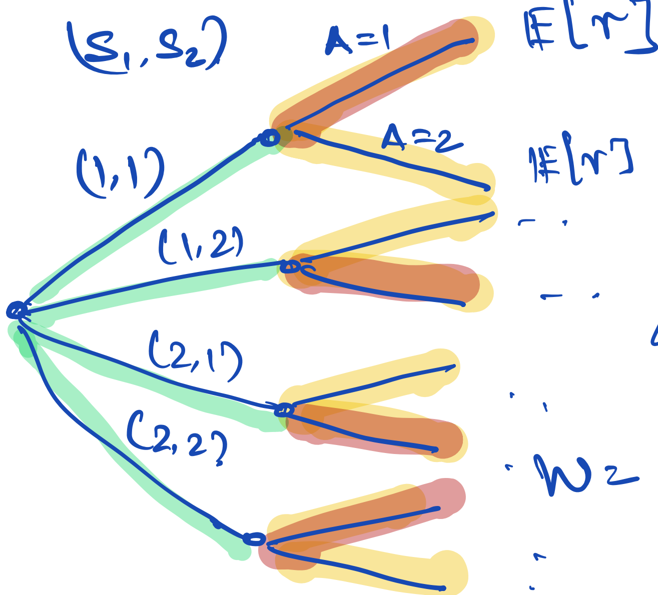
$$\mathbb{E}[r(W, \underline{A=2})] = \mathbb{E}[\min\{w_1^{CPU}, \max\{w_1^{RAM}, w_2^{RAM}\}\}]$$



CPU < Manf 1, Manf 2 $\qquad s_1 \in \{1, 2\}$

RAM → M1, M2 $\qquad s_2 \in \{1, 2\}$





State $\quad s = (s_1, s_2)$

$r : S \times W \times A \to \mathbb{R}.$

$$\left\{ \begin{array}{ll} w_1^{CPU,1} & w_2^{CPU,1} \\ w_1^{CPU,2} & w_2^{CPU,2} \\ w_1^{RAM,1} & w_2^{RAM,1} \\ w_1^{RAM,2} & w_2^{RAM,2} \end{array} \right\}$$

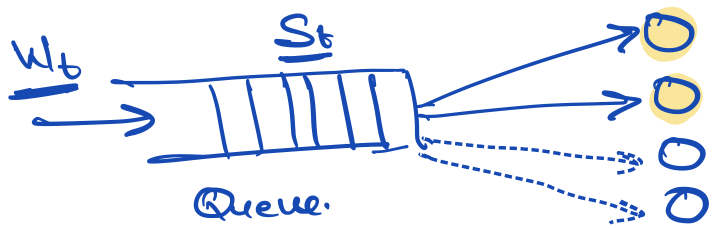
$$r((s_1, s_2), W, \underline{1}) = \mathbb{E}\left[\underline{\min\left\{ \max\left\{ w_1^{CPU, s_1}, w_2^{CPU, s_1} \right\}, w_1^{RAM, s_2} \right\}}\right]$$

__Policy__ $\quad \pi : S \to A$

$$\max_{\pi : S \to A} \mathbb{E}\left[r(\underline{S}, W, \underline{\pi(S)}) \right] \Rightarrow$$ Converted to a __sequence__ of para. opt. prob.

$\pi : S \to A$ → functional opt.

Web service



1 - CPU $\quad p$

2 - CPU

3 - CPU

4 CPU $\quad A_t \in \{1,2,3,4\}$

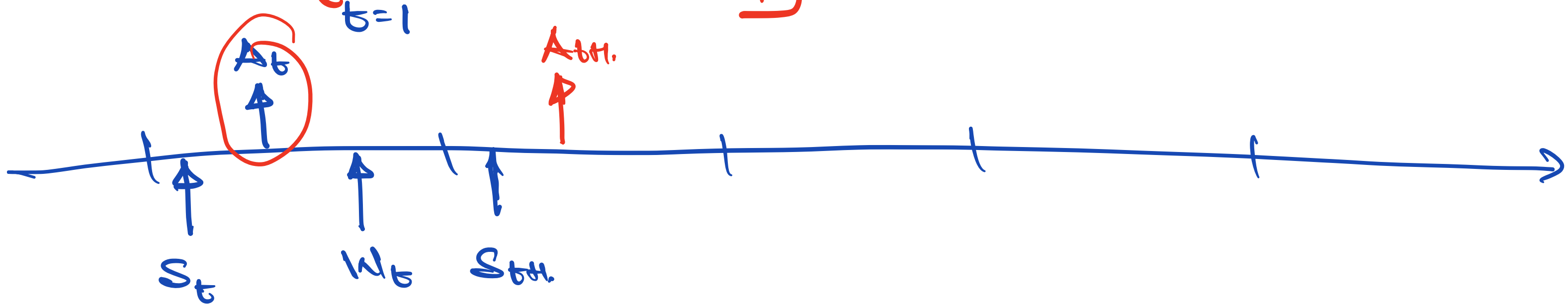
$$S_{t+1} = S_t - \min\{A_t, S_t\} + W_t$$

↳ Random new jobs

$$c(S_t, A_t) = h(S_t) + p A_t$$

↳ Holding cost

$$\mathbb{E}\left[\sum_{t=1}^{T} c(S_t, A_t)\right]$$



Info I_t available at time t.

$A_t = \pi_t(I_t)$ → Full info MDP

→ Partial | POMDP

$$\min_{\pi_1, \pi_2 \ldots \pi_T} \mathbb{E}\left[\sum_{t=1}^{T} c(S_t, A_t) \right]$$

Under some cond. $\Rightarrow$ seq. of para. optimization