Announcements

- Assignment 6 due November 21
- Assignment 7 (last assignment) due November 30
- Final Exam:

  Where: To be determined

  When: Dec 19, 10:30-12:30

  Closed book exam. You may bring 1 sheet of handwritten notes and a calculator

Project Presentations

MICHIGAN ENGINEERING
UNIVERSITY of MICHIGAN ■ COLLEGE of ENGINEERING

Proposals Due today for in-class assignment credit

If you have not proposed a project yet then you have 24 hours to email me and I will assign you to an existing team of 2.

- Presentations will be in class Dec 5 and 7. All presentations are allotted 10 minutes.
- Send me your slides by Sept 1 if you want any feedback
- Drop by office hours if you want to discuss your project

For Next Week

Tuesday:

Guest speaker, Dr. Zheng Zhang, speaking about partially observable Markov decision processes (POMDPs)

Thursday:

We will discuss the following article:

- Alagoz, O., Maillart, L., Schaefer, A., Roberts, M., 2004, The Optimal Timing of Living-Donor Liver Transplant, Operations Research, 50(1), 1420-1430

There will be a short quiz on the article at the start of next class followed by discussion of the article

Linear Programming Method

MDPs can be formulated as linear programs:

- Decision variables: value function for each state ($v(s)$)
- Constraints: optimality equations

Advantages of this approach are:

1. Insights that can be gained by viewing MDPs through the lens of linear programming
2. Advances in large-scale linear programming offer advantages to solving MDPs efficiently

Note: Background Reading directory on canvas has resources for learning about linear programming (Lptutorial.pdf, AMPL-book.pdf)

Linear Programming Method

The following linear program (LP) formulation determines an optimal policy for a maximization problem

$$\text{Min } \sum_{j \in S} \alpha(j) v(j)$$

$$s.t.$$

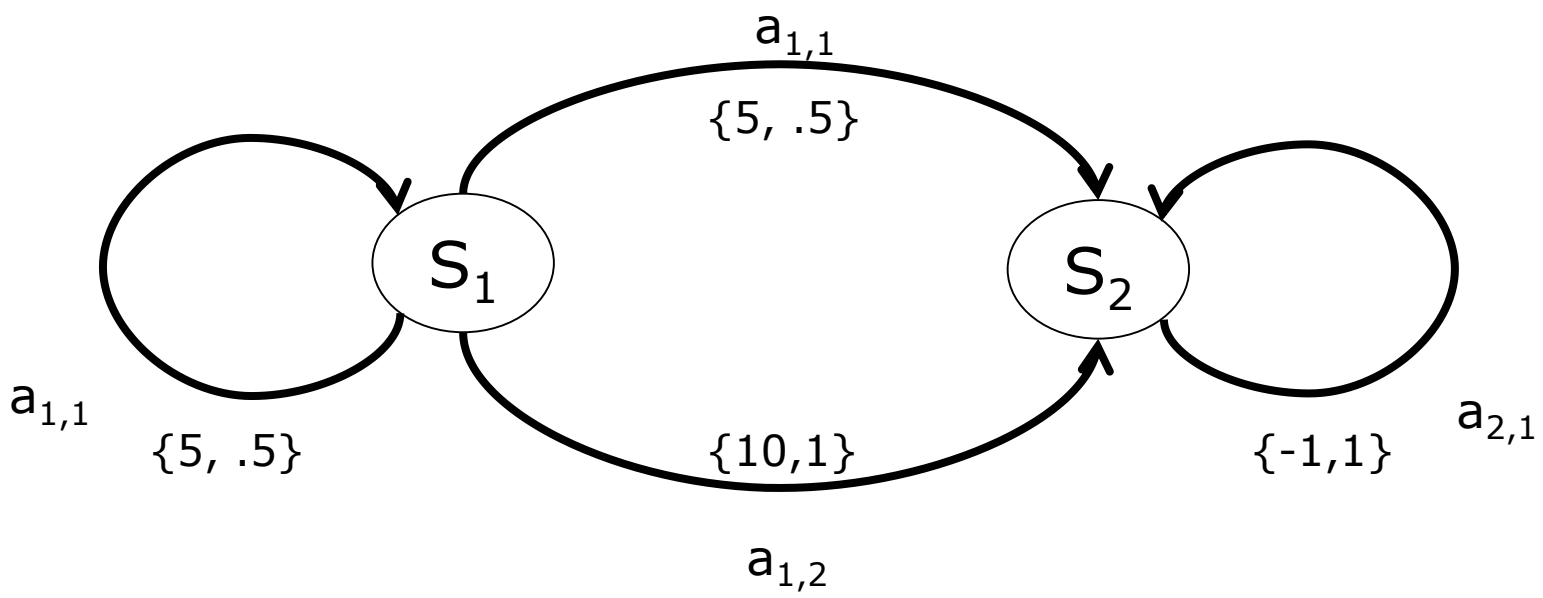
$$v(s) - \sum_{j \in S} \lambda p(j|s,a) v(j) \geq r(s,a), \ \forall a \in A(s), s \in S$$

$$v(s) \ \ urs, \forall s$$

where $\alpha(j), j \in S$ are positive scalars, often chosen to represent the initial state probability distribution ($\sum_{j \in S} \alpha(j) = 1$)

Example Revisited: 2 State MDP

In state S_1 actions $a_{1,1}$ and $a_{1,2}$ are available; in state S_2 only $a_{2,1}$ is available. Rewards and transition probabilities are defined below as {r,p}



Exercise: Formulate an LP and find an optimal policy. Assume λ = 0.95.

Linear Programming Method

The **primal problem** has |S| columns and |S| × |A| rows.

The **dual linear program** is:

$$\text{Max} \sum_{j \in S} \sum_{a \in A} r(s,a) x(s,a)$$

$$\text{s.t.}$$

$$\sum_{a \in A} x(j,a) - \sum_{s \in S} \sum_{a \in A} \lambda p(j|s,a) x(s,a) = \alpha(j), \forall j \in S$$

$$x(s,a) \geq 0, \forall s, a$$

and has |S| × |A| columns and |S| rows

Properties of the LP Formulation

Theorem (≈6.9.1, 6.9.3) Each dual feasible solution corresponds to a randomized stationary policy defined by

$$P\{d_x(s) = a\} = \frac{x(s,a)}{\sum_{a' \in A} x(s,a')}$$

And a basic feasible solution to the dual LP corresponds to a deterministic stationary policy.

Proof: See Puterman.

Properties of the LP Formulation

MICHIGAN ENGINEERING
UNIVERSITY of MICHIGAN ■ COLLEGE of ENGINEERING

Theorem (≈6.9.4): The following properties hold for the dual LP formulation of an infinite horizon MDP:

a) An optimal solution exists

b) The optimal dual solution, x^*, corresponds to an optimal deterministic policy

Proof: Complete in class

Properties of the LP Formulation

MICHIGAN ENGINEERING
UNIVERSITY of MICHIGAN ■ COLLEGE of ENGINEERING

Proposition (6.9.5): For any positive vector, α , the dual LP has the same optimal basis. Hence the optimal decision rule, d_{x^*} does not depend on α .

Proof: Completed in class

Constrained MDPs

The dual solution, $x(s,a)$, can be interpreted as the total discounted probability that the system occupies state s and chooses action a across all epochs

Intuitive Argument via Duality:

$$\sum_{s\in S} \alpha(s) v^*(s) = \sum_{s\in S} \sum_{a\in A} r(s,a)\, x^*(s,a)$$

Using this interpretation allows for constraints to be added:

$$\sum_{s\in S} \sum_{a\in A} c(s,a) x(s,a) \leq C$$