Machine, Data and Learning (CS7.301)

Spring 2022, IIIT Hyderabad
11 Apr, Thursday (Lecture 18)

Taught by Prof. Vikram Pudi

Markov Decision Processes (contd.)

Linear Programming

Mathematical programming is used to find the best or optimal solution to a problem that requires limited resources. It involves conversion of a stated problem into a mathematical model, exploring the different solutions, and finding the optimal one.

Linear programming is a form of mathematical programming which constrains all functions involved to linear ones. More precisely, we need to maximise $Z = c_1x_1 + \cdots + c_nx_n$, subject to the constraints

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &\leq b_1 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &\leq b_m \end{aligned}$$

The decision variables x_i represent the levels of competing activities.

Dantzig's simplex algorithm is a popular algorithm to solve LP problems. However, LP formulations of MDPs are often slower than the value iteration algorithm.

We can formalise MDPs as LP problems by associating a value v_i with each state s_i. We want to maximise $\sum v_i$, subject to the constraint that

$$v_i \leq R(I, A) + \gamma \sum P(J \mid I, A)v_j$$

for all i.

However, a more popular formulation maximises $\sum_I \sum_A x_{ia}r_{ia}$, under the constraints

$$\sum_A x_{ja} - \sum_I \sum_A,$$

where x_{ia} is the number of times action a is taken in state i.
More simply, we maximise $r \cdot x$ under the constraint $Ax = \alpha$.