Ma 661 Dynamic Programming and Reinforcement Learning
Darinka Dentcheva darinka.dentcheva@stevens.edu

Homework 1

due Wednesday, February 5, 2020

Problem 1. Consider the optimization problem

$$\begin{aligned} \max\ & 4z_1 + 5z_2 + 4z_3 + 8z_4 + 6z_5 \\ & 4z_1 + 6z_2 + 5z_3 + 9z_4 + 7z_5 \le 19 \\ & z_j \ge 0, \quad j = 1,\dots,5. \end{aligned} \tag{1}$$

Assume that all z_j may take only nonnegative integer values.

(a) Formulate a dynamic programming problem for solving problem (1). Describe the state space, the action space, the feasible action mapping, and the dynamic programming equation.
Decision times $t = 1,\dots 5$
State space $\mathscr{X} = \{0,1,\dots,19\}$
Control space $\mathscr{U} = \{0,1,2,\dots\}$
Feasible control mapping $U_t(x) = \{z \in \mathbb{N} : z \le \left\lfloor \frac{x}{a_i} \right\rfloor\}$, where $a = (4,6,5,9,7)$.
Reward function: $r_t(x,z) = c_t z$, where $c = (4,5,4,8,6)$.
Dynamics: $x_{t+1} = x_t - a_t z_t$.
Dynamic programming equations:

$$\begin{aligned} v_t(x) &= \max_{z\in U_t(x)} \{r_t(x,z) + v_{t+1}(x - a_t z)\}, \quad t = 1,\dots,5; \\ v_6(x) &= 0. \end{aligned}$$

(b) Use the dynamic programming equations to obtain the optimal solution.
The optimal solution is $z = (3,0,0,0,1)$.

Problem 2. Let $A_1,\dots,A_n$ be matrices with A_i having k_{i-1} rows and k_i columns for $i = 0,1,2,\dots,n$. and some positive integers $k_0,k_1,\dots,k_n$. The problem is to choose the order of multiplying the matrices that will minimize the number of multiplications needed to compute the product $A_1A_2\cdots A_n$.

(a) Formulate a dynamic programming problem to determine the optimal order in which to multiply the matrices. Clearly formulate the dynamic programming equation.

- Decision times $t = 1,2,\dots,n-1$;

- State space $\mathscr{X} = \{(k_0, k_1, \ldots, k_\ell) : k_i \in \mathbb{Z}, \ell = 2, \ldots, n\}$
  The state indicated the sizes of the matrix that have to be multiplied.
- Control space $\mathscr{U} = \{1, 2, \ldots, n-1\}$
  The control determines which matrices to multiply, for example, $u = 2$ means multiply matrix number two with the next matrix (the third one).
- Feasible control mapping $U_t(x) = \{1, 2, \ldots, |x| - 1\}$, where $|x|$ means the number of component the state has.
- Cost function: $c(x, u) = x_{u-1} * x_u * x_{u+1}$.
- Dynamics: if $x^t = (x_1^t, \ldots, x_\ell^t)$, then $x^{t+1} = (x_1^t, \ldots, x_{u-1}^t, x_{u+1}^t, \ldots, x_\ell^t)$
- We can use the forward method in this case. Dynamic programming equations:

$$v_{t+1}(x^{t+1}) = \min_{u \in U_t(x)} \{c(x^t, u) + v_{t+1}(x^{t+1})\}, \quad t = 1, \ldots, n;$$
$$v_1(x^1) = 0, \quad x^1 = (k_0, k_1, \ldots, k_n)$$

(b) Solve the problem numerically, when $n = 4$ and $(k_0, k_1, \ldots, k_4) = (10, 30, 70, 2, 100)$.
The best order is $(A_1(A_2A_3))A_4$.

Problem 3. A directed graph has node set $\mathscr{N} = \{1, \ldots, 10\}$ and the arc set given in the table on the next page. Determine the shortest path from node 1 to node 10 using dynamic programming.

The shortest path is $1 \to 4 \to 5 \to 10$.