
Problem Set 1

1. Suppose that an investor wants to maximize its expected wealth $E[w_T]$ at time stage T . There are two investment options available: a fixed-interest savings account with interest rate of 3% in each time stage, or a stock whose price p_t fluctuates according to $P(p_{t+1} = (1 + r_u)p_t) = P(p_{t+1} = (1 - r_d)p_t) = 0.5$, where $0 \leq r_d \leq 1$ and $r_u \geq 0$. The investor must decide at each time stage which fraction of its current wealth to invest in each option.
 - (a) Suppose that there are no transaction costs involved in buying or selling stock. What is the optimal policy? Suppose that, instead of maximizing $E[w_T]$, the investor wants to maximize $E[\log w_T]$. What is the optimal policy? Do you expect the investor to be more or less conservative in this case?
 - (b) Suppose that buying or selling stock involves a transaction cost of 0.5% of the transaction value. Formulate this problem as an MDP, and write its Bellman's equation.
 - (c) Solve the problem numerically with $T = 20$, $r_d = 0.9$ and $r_u = 1.2$, for both the case of maximizing $E[w_T]$ and maximizing $E[\log w_T]$. Solve it again with $r_d = 0.7$ and $r_u = 1.4$. Analyze your results.
2. Let M_1, M_2, \dots, M_n be matrices with M_i having r_{i-1} rows and r_i columns for $i = 1, 2, \dots, n$ and some positive integers r_0, r_1, \dots, r_n . The problem is to choose the order for multiplying the matrices that minimizes the number of multiplications needed to compute the product $M_1 M_2 \dots M_n$. Assume that matrices are multiplied in the usual way.
 - (a) Formulate the problem as an MDP such that using backwards induction for finding the optimal order in which to multiply the matrices requires $O(n^3)$ operations.
 - (b) Find the optimal order in which to multiply the matrices when $n = 4$ and $(r_0, r_1, r_2, r_3, r_4) = (10, 30, 70, 2, 100)$.
 - (c) Using the same numerical values of part (b), solve the problem where the objective is instead to maximize the number of multiplications. What is the ratio of the maximum to minimum number of multiplications?
3. Show that Gauss-Seidel value iteration still converges to J^* if states are chosen in an arbitrary order, as long as each state is visited infinitely many times.
4. Let $c \in (0, 1]^{|S|}$ satisfy

$$\sum_{x \in S} c(x) = 1.$$

- (a) Show that, for every policy u ,

$$(I - \alpha P_u)^{-1} = \sum_{t=0}^{\infty} \alpha^t P_u^t.$$

- (b) Let $\mu_u = (1 - \alpha)c^T(I - \alpha P_u)^{-1}$. Show that μ_u is a probability distribution over S and $\mu_u(x) > 0$ for all x .

(c) For any $\nu \in (0, \infty)^{|\mathcal{S}|}$, define

$$\|J\|_{1,\nu} = \sum_{x \in \mathcal{S}} \nu(x) |J(x)|.$$

(d) Show that $\|\cdot\|_{1,c}$ is a norm.

(e) Let $J \leq J^*$ be arbitrary, and recall that u_J stands for a greedy policy with respect to J . Show that

$$\|J_{u_J} - J^*\|_{1,c} \leq \frac{1}{1-\alpha} \mu_{u_J}^T (TJ - J) \leq \frac{1}{1-\alpha} \|TJ - J\|_{1,\mu_{u_J}}.$$

(f) Show that, if we start value iteration with $J_0 \leq J^*$, we have $J_k \leq J^*$ for all k , and

$$\|J_{u_{J_k}} - J^*\|_{1,c} \leq \frac{\alpha^k}{1-\alpha} \|TJ_0 - J_0\|_{\infty}.$$

(g) In class, we have proved a bound on $\|J_{u_{J_k}} - J^*\|_{\infty}$. Compare the guarantees offered by the algorithm when the stopping criterion is $\|J_{u_{J_k}} - J^*\|_{1,c} \leq \epsilon$ versus $\|J_{u_{J_k}} - J^*\|_{\infty} \leq \epsilon$. Which one offers stronger guarantees? In which case does the algorithm stop first? Can you think of situations where it may make sense to use the first criterion?