6.231 Dynamic Programming and Stochastic Control
Fall 2008

6.231 Dynamic Programming

Midterm Exam, Fall 2003

Prof. Dimitri Bertsekas

Problem 1

Consider the problem of inscribing an N-side polygon in a given circle, so that the polygon has maximal perimeter.

(a) Formulate the problem as a DP problem involving sequential placement of N points in the circle.

(b) Use DP to show that the optimal polygon is regular (all sides are equal).

Solution: Similar to Problem 1.27.

Problem 2

Consider an algorithm that may be viewed as a combined forward and backward label correcting method. The algorithm maintains an OPEN list of nodes, and two labels, d_i and e_i, for each node i. Initially, OPEN contains nodes 1 and N, $d_1 = 0$ and $d_i = \infty$ for $i \neq 1$, and $e_N = 0$ and $e_i = \infty$ for $i \neq N$. The algorithm also maintain a variable UPPER, initially set to a_{1N}. The typical iteration of the algorithm is as follows:

Step 1: Remove a node i from OPEN.

Step 2: For each $j > i$, if $d_i + a_{ij} < \min\{d_j, \text{UPPER}\}$, set $d_j = d_i + a_{ij}$, and place j in OPEN, if j is not already there. Furthermore, if $d_j + e_j < \text{UPPER}$, set $d_j + e_j = \text{UPPER}$.

Step 3: For each $j < i$, if $e_i + a_{ji} < \min\{e_j, \text{UPPER}\}$, set $e_j = e_i + a_{ji}$, and place j in OPEN, if j is not already there. Furthermore, if $d_j + e_j <$ UPPER, set $d_j + e_j =$ UPPER.

Step 4: If OPEN is empty stop; else go to Step 1.

Give an interpretation of the labels d_i and e_i, and of the sums $d_i + e_i$. Explain why the algorithm terminates with UPPER equal to the minimum cost (a reasonable justification is required here, not necessarily a complete rigorous proof).

Problem 3

A businessman operates during a time interval $[0, T]$ out of a van that he sets up in one of two locations. Being in location i at time t, where $i = 1, 2$, earns reward at a rate $g_i(t)$. Every switch between the two locations costs $c > 0$, but takes negligible time. Thus, for example, the reward for starting in location 1, switching to location 2 at time t_1, and switching back to 1 at time $t_2 > t_1$ earns total reward

$$\int_0^{t_1} g_1(t)\, dt + \int_{t_1}^{t_2} g_2(t)\, dt + \int_{t_2}^{T} g_1(t)\, dt - 2c.$$

We want to find a set of switching times that maximize the total reward. Assume that the function $g_1(t)-g_2(t)$ changes sign a finite number of times in the interval $[0, T]$. Formulate the problem as a finite horizon problem, and write the corresponding DP algorithm.

Solution: This is Problem 1.19.

Problem 4

Consider the asset selling problem of Section 4.4. The offers w_k are independent and identically distributed. However, the (common) distribution of the w_k is unknown. Instead it is known that this distribution is one out of two known distributions F_1 and F_2, and that the a priori probability that F_1 is the correct distribution is a given scalar q, with $0 < q < 1$.

(a) Formulate this as an imperfect state information problem, and identify the state, control, system disturbance, observation, and observation disturbance.

(b) Show that (x_k, q_k), where

$$q_k = P(d_k = 1 \mid w_0, \ldots, w_{k-1}),$$

is a suitable sufficient statistic, write a corresponding DP algorithm, and derive the form of the optimal selling policy.

Solution: Similar to midterm inventory control problem.

Problem 5

Consider a problem of finding a shortest path from a given origin node s to a given destination node t in a graph with nonnegative arc lengths. However, there is the constraint that the path should successively pass through at least one node from given node subsets $T_1, T_2, \ldots, T_N$ (i.e., for all k, pass through some node from the subset T_k after passing through at least one node of the subsets $T_1, \ldots, T_{k-1}$).

(a) Formulate this as a dynamic programming problem.

(b) Show that a solution can be obtained by solving a sequence of ordinary shortest path problems, each involving a single origin and multiple destinations.

Problem 6

Consider a problem of finding a shortest path from a given origin node s to a given destination node t in a graph with nonnegative arc lengths. Consider an algorithm that maintains two subsets of nodes, W and V, with the following properties:

(1) $s \in W$ and $t \in V$.

(2) If $i \in W$ and $j \notin W$, then the shortest distance from s to i is less than or equal to the shortest distance from s to j.

(3) If $i \in V$ and $j \notin V$, then the shortest distance from i to t is less than or equal to the shortest distance from j to t.

At each iteration the algorithm adds a new node to W and a new node to V (the Dijkstra algorithm can be used for this purpose), and terminates when W and V have a node in common. Let d_i^s be the shortest distance from s to i using paths all the nodes of which, with the possible exception of i, lie in W ($d_i^s = \infty$ if no such path exists), and let d_i^t be the shortest distance from i to t using paths all the nodes of which, with the possible exception of i, lie in V ($d_i^t = \infty$ if no such path exists).

(a) Show that upon termination, the shortest distance D_{st} from s to t is given by

$$D_{st} = \min_{i \in W} \{d_i^s + d_i^t\} = \min_{i \in W \cup V} \{d_i^s + d_i^t\} = \min_{i \in V} \{d_i^s + d_i^t\}.$$

(b) Show that the conclusion of part (a) holds if the algorithm is terminated once the condition

$$\min_{i \in W} \{d_i^s + d_i^t\} \le \max_{i \in W} d_i^s + \max_{i \in V} d_i^t$$

holds, even if the sets W and V have no node in common.

Problem 7

A famous but somewhat vain opera singer is scheduled to sing on N successive nights. If she is satisfied with her performance on a given night (which happens with probability p, independently of the previous history), she is happy to sing on the following night. If she is not satisfied, however, she sulks and is inclined to refuse to sing on the following night. In this case, the opera director may decide to send her an expensive gift, costing G dollars, which with probability q (and independently of the previous history) placates her into performing again on the following night. Given that each missed performance costs the opera house C dollars, the question is to find the optimal gift policy of the opera director.

(a) Formulate this as a DP problem, and characterize as best you can the optimal policy.

(b) Repeat part (a) for the case where the probability q is not constant, but rather is a decreasing function of the number of times he has sent her a gift so far.