IE5571/8571 Assignment 1

You can discuss problems with other students but must write up your own solution.

Problem 1

A political campaign is entering the final stage. One of the candidates has enough funds left to purchase advertisements for a a total of five prime-time commercials on TV stations located in four different areas. Based on polling information, an estimate has been made of the number of additional votes that can be won in the different broadcasting regions (given in the table below in thousands of votes).

Commercials	Area 1	Area 2	Area 3	Area 4
0	0	0	0	0
1	4	6	5	3
2	7	8	9	7
3	9	10	11	12
4	12	11	10	14
5	15	12	9	16

Use dynamic programming to determine how to optimally allocate the five commercials over the four broadcasting regions to maximize the estimated number of votes won. You must specify the states and stages of the system, the controls, the dynamical system, and the cost function.

Problem 2

Farmer J currently has \$5000 and 10 tons of wheat. During month j the price of wheat is p_j (assumed known). During each month, J must decide how much wheat to buy or sell. There are three restrictions on each month's wheat transaction: 1) During any month, the amount of money spent on wheat cannot exceed the cash on hand at the beginning of the month; 2) during any month, J cannot sell more wheat than in possession at the beginning of the month; 3) because limited warehouse capacity, the ending inventory of wheat for each month cannot exceed 10 tons. Show how dynamic programming can be utilized to maximize the amount of cash farmer J has on hand at the end of 3 months.

Problem 3

Consider an inventory production system in which the state variables x_n is the stock level at the beginning of period n, for $n = 0, 1, \ldots, N$. The control $u_n \geq 0$ is the quantity ordered and immediately supplied at the beginning of period n. The demand during period n is given by Y_{n+1}. We assume that $\{Y_n\}_{n\geq 1}$ is a sequence of i.i.d non-negative random variables with probability density function f and cumulative distribution function F. The dynamics of the system are given by

$$x_{n+1} = x_n + u_n - Y_{n+1}, \quad n = 0, 1, \ldots, N-1.$$

We allow negative inventory by assuming excess demand is backlogged and filled when additional inventory becomes available. Our goal is to minimize the total expected operation cost

$$E\left[\sum_{n=0}^{N-1} c\left(x_n, u_n; Y_{n+1}\right)\right],$$

with the cost function

$$c\left(x_n, u_n; Y_{n+1}\right) = p \cdot u_n + h \cdot \left(x_n + u_n - Y_{n+1}\right)^+ + b \cdot \left(x_n + u_n - Y_{n+1}\right)^-,$$

where p is the unit production cost, h the unit holding cost, b the backlogging cost. Note that we assume $b > p$ and use the definitions $x^+ = \max\{x, 0\}$ and $x^- = -\min\{x, 0\}$.

(a) Write down a dynamic programming equation to study this system.

(b) Specify the value function and optimal policy when on stage $N-1$.

(c) **(8571 Only)** Can you say anything about the general structure of the optimal policy?

Problem 4

Suppose you initially have x_0 wealth. At each time point $n \in \{0, \dots, N-1\}$ you determine a proportion $u_n \in [0,1]$ of your total available wealth, x_n, to consume and put the rest in the bank. The bank offers an interest rate $r > 0$ per unit time. If we write $R = 1 + r$, then the dynamics of the wealth process are

$$x_{n+1} = Rx_n(1 - u_n), \quad n = 0, \dots, N_1.$$

For the wealth consumed at time n, you receive a utility $g(u_n x_n)$. For the wealth remaining at the terminal time N you receive a utility $g_N(x_N)$. The goal is to choose the controls $\{u_n\}$ to maximize your total utility, i.e., to solve the equation

$$v(x_0) = \max_{\{u_n\}} \left\{ g_N(x_n) + \sum_{n=0}^{N-1} g(u_n x_n) \right\}.$$

(a) Write down a dynamic programming equation for this system.

(b) Assume that $f(x) = x^\alpha$ and $g(x) = x^\alpha$ for $\alpha \in (0,1)$. Show that the optimal cost to go function can be written as

$$V_n(x) = c_n x^\alpha,$$

for constants c_n that are defined recursively. Also show that in this case the optimal proportion of wealth to consume is independent of current wealth.

Problem 5 (IE 8571 Only) Consider the controlled dynamical system, $x_{k+1} = f(x_k, u_k)$ with $k \in \{0, \dots, N-1\}$, state sequence $\{x_k\}_{k=0}^N$, and control sequence $\{u_k\}_{k=0}^{N-1}$. In addition assume that there are running costs g and terminal costs g_N. For initial condition x_0, define the solution to the optimal control problem as

$$J^*(x_0) = \min \left\{ g_N(x_N) + \sum_{k=0}^{N-1} g(x_k, u_k); u_k \in \mathcal{U}(x_k), x_{k+1} = f(x_k, u_k), k \in \{0, \dots, N-1\} \right\}.$$

In addition, for $k \in \{1, \dots, N-1\}$ define the $N-k$ tail problem as

$$J_k^*(x_k) = \min \left\{ g_N(x_N) + \sum_{m=k}^{N-1} g(x_m, u_m); u_m \in \mathcal{U}(x_m), x_{m+1} = f(x_m, u_m), m \in \{k, \dots, N-1\} \right\}.$$

Next define $V_N(x) = g_N(x)$, and for $k < N$ define

$$V_k(x) = \min_{u \in \mathcal{U}(x)} \left(g(x,u) + V_{k+1}\left(f(x,u)\right) \right).$$

Prove that for each $k \in \{0, 1, \dots, N\}$, $V_k = J_k^*$.

Problem 6 Exercise 2.4, 2.5 and 2.7 from Sutton and Barto

Problem 7 (IE 8571 Only) In the Besbes et al paper on Canvas, can you explain why the authors impose the restriction $V_T \leq T/K$? Do you think it would be possible to prove similar results with a weaker condition on V_T?