Barcelona GSE - Stochastic Models and Optimization

Problem Set 4 - Solutions

Problem 1 (Linear-Quadratic Problem with Forecasts - [B05] Exercise 4.1)

Consider the linear-quadratic problem with perfect state information, where at the beginning of period k we have a forecast $y_k \in \{1, \ldots, n\}$ consisting of an accurate prediction that w_k will be selected according to a particular probability distribution $P_{k|y_k}$. The vectors w_k need not have zero mean under the distribution $P_{k|y_k}$. Show that the optimal control law is of the form

$$\mu_k^*(x_k, y_k) = -\big(B_k' K_{k+1} B_k + R_k\big)^{-1} B_k' K_{k+1}\big(A_k x_k + \mathbb{E}\{w_k | y_k\}\big) + \alpha_k,$$

where the matrices K_k are given by the discrete time Riccati equation and α_k are appropriate vectors.

Solution: As the hint suggests, the proof of the fact that

$$J_k(x_k, y_k) = x_k' K_k x_k + x_k' b_k(y_k) + c_k(y_k)$$

will follow by induction, deriving $\mu_k^*(x_k, y_k)$ along the way.

For $k = N$ the result is clearly true. Now, assume that $J_{k+1}(x_{k+1}, y_{k+1}) = x_{k+1}' K_{k+1} x_{k+1} + x_{k+1}' b_{k+1}(y_{k+1}) + c_{k+1}(y_{k+1})$. Then,

$$\begin{aligned}
J_k(x_k, y_k) &= \min_{u_k} \mathbb{E}\Big\{ x_k' Q_k x_k + u_k' R_k u_k + \sum_{i=1}^n p_i^{k+1}\big[x_{k+1}' K_{k+1} x_{k+1} + x_{k+1}' b_{k+1}(i) + c_{k+1}(i)\big] \;\Big|\; y_k \Big\} \\
&= x_k' Q_k x_k + \min_{u_k} \Big\{ u_k' R_k u_k + \mathbb{E}\Big\{ \big(A_k x_k + B_k u_k + w_k\big)' K_{k+1} \big(A_k x_k + B_k u_k + w_k\big) \\
&+ \big(A_k x_k + B_k u_k + w_k\big)' \sum_{i=1}^n p_i^{k+1} b_{k+1}(i) \;\Big|\; y_k \Big\} + \sum_{i=1}^n p_i^{k+1} c_{k+1}(i) \Big\} \\
&= x_k' Q_k x_k + x_k' A_k' K_{k+1} A_k x_k + 2 x_k' A_k' K_{k+1} \mathbb{E}\{w_k | y_k\} + \mathbb{E}\{w_k' K_{k+1} w_k | y_k\} + x_k' A_k' b_{k+1} \\
&+ b_{k+1}' \mathbb{E}\{w_k | y_k\} + \gamma_{k+1} + \min_{u_k} \Big\{ u_k' \big(R_k + B_k' K_{k+1} B_k\big) u_k + 2 u_k' B_k' K_{k+1} A_k x_k \\
&+ 2 u_k' B_k' K_{k+1} \mathbb{E}\{w_k | y_k\} + u_k' B_k' b_{k+1} \Big\},
\end{aligned}$$

where in the last step we have used the notation

$$\sum_{i=1}^n p_i^{k+1} b_{k+1}(i) = b_{k+1},$$

and

$$\sum_{i=1}^n p_i^{k+1} c_{k+1}(i) = \gamma_{k+1}.$$

We can determine the optimal control policy through the FOC:

$$2\big(R_k + B_k' K_{k+1} B_k\big) u_k^* + 2B_k' K_{k+1}\big(A_k x_k + \mathbb{E}\{w_k | y_k\}\big) + B_k' b_{k+1} = 0.$$

Since $R_k > 0$ and $K_{k+1} \geq 0$, we know that $R_k + B_k' K_{k+1} B_k > 0$. Consequently, the optimal control policy takes the form:

$$u_k^* = \mu_k^*(x_k, y_k) = -(R_k + B_k' K_{k+1} B_k\big)^{-1} B_k' K_{k+1}\big(A_k x_k + \mathbb{E}\{w_k | y_k\}\big) + \alpha_k,$$

with $\alpha_k = -\frac{1}{2}\big(R_k + B_k' K_{k+1} B_k\big)^{-1} B_k' b_{k+1}$, which is what we set out to prove.

Substituting this back to the expression for the optimal cost-to-go concludes the induction argument.

Problem 3 (**Asset Selling with Offer Estimation - [B05] Exercise 5.14**)

Consider a variation of the asset selling problem discussed in class: the offers $\{w_k\}$ are independent and identically distributed, taking values in a finite set, but their common distribution is unknown. Instead, it is known that this distribution is one out of two given distributions F_1 and F_2, and that the a priori probability that F_1 is the correct one is $q \in (0, 1)$.

Use Dynamic Programming to derive the optimal asset selling policy.

Solution: We use as sufficient statistic the pair (x_k, q_k), where x_k is the offer at hand at the beginning of period k and

$$q_k = \mathbb{P}\left(\text{distribution is } F_1 \mid w_0, \ldots, w_{k-1}\right).$$

The conditional probability q_k evolves according to

$$q_{k+1} = \frac{q_k F_1(w_k)}{q_k F_1(w_k) + (1-q_k) F_2(w_k)}, \qquad q_0 = q,$$

where $F_i(w_k)$ denotes probability of having an offer w_k under the distribution F_i.

Let $w^1, w^2, \ldots, w^n$ be the possible values w_k can take under either distribution. Then, we have the following DP algorithm:

$$J_N(x_N, q_N) = x_N$$

and

$$\begin{aligned} J_k(x_k, q_k) &= \max\left\{(1+r)^{N-k} x_k, \mathbb{E}[J_{k+1}(x_{k+1}, q_{k+1})]\right\} \\ &= \max\left\{(1+r)^{N-k} x_k, \sum_{i=1}^{N} \left(q_k F_1(w^i) + (1-q_k) F_2(w^i)\right) J_{k+1}\left(w^i, \frac{q_k F_1(w^i)}{q_k F_1(w^i) + (1-q_k) F_2(w^i)}\right)\right\}. \end{aligned}$$

As in the standard asset selling problem, we renormalize the cost-to-go function to account for the returns of the riskless asset:

$$V_k(x_k, q_k) = \frac{J_k(x_k, q_k)}{(1+r)^{N-k}}.$$

Then we have

$$V_k(x_k, q_k) = \max\left\{x_k, \alpha_k(q_k)\right\},$$

where

$$\alpha_k(q_k) = \frac{1}{1+r} \sum_{i=1}^{N} \left(q_k F_1(w^i) + (1-q_k) F_2(w^i)\right) J_{k+1}\left(w^i, \frac{q_k F_1(w^i)}{q_k F_1(w^i) + (1-q_k) F_2(w^i)}\right).$$

Consequently, the stopping sets have the form:

$$T_k = \{x \mid x \geq \alpha_k(q_k)\},$$

so the optimal asset selling policy is, again, of a threshold type.

Using the time-stationarity of the problem and the fact that $V_{N-1}(x, q) \geq V_N(x, q)$, for all x, q, it can be shown by induction that $V_k(x, q) \geq V_{k+1}(x, q)$, for all x, q, k, which implies that $\alpha_k(q) \geq \alpha_{k+1}(q)$, for all q, k.

Problem 4 (Inventory Control with Demand Estimation - [B05] Exercise 5.15)

Consider a variation of the inventory control problem discussed in class: the available inventory evolves according to the dynamical equation

$$x_{k+1} = x_k + u_k - w_k,$$

and the instantaneous cost is

$$cu_k + h\max(0, w_k - x_k - u_k) + p\max(0, x_k + u_k - w_k),$$

where c, h, and p are positive scalars, with $p > c$. There is no terminal cost.

The available inventory is perfectly observable at every period. The demand over different periods $\{w_k\}$ is independent and identically distributed, taking values in a finite set, whose underlying distribution is unknown. Instead, it is known that this distribution is one out of two given distributions F_1 and F_2, and that the a priori probability that F_1 is the correct one is $q \in (0, 1)$.

Use Dynamic Programming to characterize the optimal inventory management policy.

Solution: We use as sufficient statistic the pair (x_k, q_k), where x_k is the inventory at hand at the beginning of period k and

$$q_k = \mathbb{P}\left(\text{distribution is } F_1 \mid w_0, \ldots, w_{k-1}\right).$$

The conditional probability q_k evolves according to

$$q_{k+1} = \frac{q_k \mathbb{P}(w_k \mid F_1)}{q_k \mathbb{P}(w_k \mid F_1) + (1 - q_k)\mathbb{P}(w_k \mid F_2)}, \qquad q_0 = q,$$

where $\mathbb{P}(w_k \mid F_i)$ denotes probability of having demand w_k assuming the actual distribution is F_i.

We can write the DP algorithm in terms of this sufficient statistic as follows:

$$J_N(x_N, q_N) = 0,$$

and

$$\begin{aligned} J_k(x_k, q_k) = \min_{u_k \geq 0} \Big\{ & cu_k \\ & + q_k \mathbb{E}\left[h\max(0, w_k - x_k - u_k) + p\max(0, x_k + u_k - w_k) + J_{k+1}\big(x_k + u_k - w_k, \phi(q_k, w_k)\big) \mid F_1\right] \\ & + (1 - q_k)\mathbb{E}\left[h\max(0, w_k - x_k - u_k) + p\max(0, x_k + u_k - w_k) + J_{k+1}\big(x_k + u_k - w_k, \phi(q_k, w_k)\big) \mid F_2\right] \Big\}, \end{aligned}$$

where

$$\phi(q_k, w_k) = \frac{q_k \mathbb{P}(w_k \mid F_1)}{q_k \mathbb{P}(w_k \mid F_1) + (1 - q_k)\mathbb{P}(w_k \mid F_2)}.$$

It can be shown inductively, as in the standard inventory management problem, that $J_k(x_k, q_k)$ is a convex function of x_k, for fixed q_k. Thus, for any given value of q_k, the minimization in the right-hand side of the DP recursion is exactly the same as before, with the probability distribution of w_k being the mixture of the distributions F_1 and F_2, with corresponding probabilities q_k and $(1 - q_k)$. It follows that for any given value of q_k, there is a threshold $S_k(q_k)$ such that it is optimal to order an amount $S_k(q_k) - x_k$, if $S_k(q_k) > x_k$, and to order nothing otherwise. In particular, $S_k(q_k)$ minimizes, over all $y \geq x_k$, the function

$$\begin{aligned} G_k(y, q_k) = & cy + q_k \mathbb{E}\left[h\max(0, w_k - y) + p\max(0, y - w_k) + J_{k+1}\big(y - w_k, \phi(q_k, w_k)\big) \mid F_1\right] \\ & + (1 - q_k)\mathbb{E}\left[h\max(0, w_k - y) + p\max(0, y - w_k) + J_{k+1}\big(y - w_k, \phi(q_k, w_k)\big) \mid F_2\right]. \end{aligned}$$

Problem 5 (**Robust Dynamic Programming**)

Consider the basic problem with the difference that instead of having a probabilistic description of the disturbance w_k, it is just known to belong to a given set $W_k(x_k, u_k)$. The goal is to find a closed-loop policy $\pi = \{\mu_0(\cdot), \ldots, \mu_{N-1}(\cdot)\}$, with $\mu_k(x_k) \in U_k(x_k)$, for all x_k and k, that minimizes the cost function

$$J_\pi(x_0) = \max_{w_k \in W_k(x_k, \mu_k(x_k))} \left\{ g_N(x_N) + \sum_{k=0}^{N-1} g_k\big(x_k, \mu_k(x_k), w_k\big) \right\}.$$

(a) Using the principle of optimality suggest a DP-like recursion that produces an optimal policy;

(b) A special case of the above problem is the "reachability of a target tube": at each stage k, the state x_k must belong to a given set X_k. Find a cost structure that fits the reachability problem within the general formulation, and propose a recursion to compute the set $\bar{X}_k$, i.e., the set that we must reach at stage k in order to be able to maintain the state of the system in the desired tube henceforth.

Solution: *Part (a).* Given the cost function, in this case the DP recursion could take the following form:

$$J_N(x_N) = g_N(x_N),$$

and

$$J_k(x_k) = \min_{u_k \in U_k(x_k)} \max_{w_k \in W_k(x_k, w_k)} \{g_k(x_k, u_k, w_k) + J_{k+1}\big(f(x_k, u_k, w_k)\big)\}.$$

Part (b). Having a cost function that penalizes being outside of the desired set would ensure that the subsequent states x_k belong to the set X_k. For example, the cost structure could be:

$$g_k(x_k) = \begin{cases} 0 & \text{if } x_k \in X_k \\ 1 & \text{if } x_k \notin X_k. \end{cases}$$

The objective is to reach at stage k the set $\bar{X}_k$ to be able to maintain the state of the system in the desired tube henceforth. If the state is in the required set, then the cost would be zero. So the set can be defined as:

$$\bar{X}_k = \{x_k \mid J_k(x_k) = 0\}.$$

The iterations can then start from set $\bar{X}_n = X_n$, and then for each iteration k the cost function should ensure that a control is chosen such that the next state stays within the required set. The set $\bar{X}_k$ can be defined as:

$$\bar{X}_k = \{x_k \in X_k \mid \exists u_k \in U_k(x_k) \text{ such that } f(x_k, u_k, w_k) \in \bar{X}_{k+1},\ \forall w_k \in W_k(x_k, u_k)\}.$$