Barcelona GSE - Stochastic Models and Optimization

Problem Set 4

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.

Hint: Prove by induction that

$$J_k(x_k, y_k) = x_k' K_k x_k + x_k' b_k(y_k) + c_k(y_k), \qquad k = 0, 1, \ldots, N,$$

where $b_k(y_k)$ is an n-dimensional vector and $c_k(y_k)$ is a scalar.

Problem 2 (Computational Assignment on Linear-Quadratic Control)

Consider a horizon of $N = 100$ time periods and a discrete time, homogeneous, linear system, i.e.,

$$x_{k+1} = Ax_k + Bu_k + w_k,$$

where A and B are $2{\times}2$ matrices and $\{w_k\}$ is a sequence of IID, zero-mean, Gaussian random vectors. Choose A and B such that the matrix $[B, AB]$ is full rank; this is the so-called *controllability* condition. Let x_0 be the *nonzero* initial condition of the system, and D the *diagonal* covariance matrix of the disturbances.

Choose a $1{\times}2$ vector C, such that the matrix $[C', A'C']$ is full rank; this is the so-called *observability* condition. Also, let R be a *diagonal, positive definite* $2{\times}2$ matrix. The two define the cost structure:

$$g_k(x_k, u_k) = x_k' C' C x_k + u_k' R u_k,$$

and

$$g_N(x_N) = x_N' C' C x_N.$$

(i) What is the intuition behind the controllability and observability conditions? What structure do they impose on the system?

(ii) Fix R and x_0, and compare the behavior of the system for two covariance matrices for the disturbances, one "much larger" than the other, under optimal control (given by the discrete-time Riccati equation);

(iii) Fix R and D, and compare the behavior of the system for two initial conditions, one "much larger" than the other, under optimal control;

(iv) Fix x_0 and D, and compare the behavior of the system for two input-cost matrices, one "much larger" than the other, under optimal control;

(v) Fix R, x_0, and D, and compare the behavior of the system under optimal control vs. steady-state control (given by the algebraic Riccati equation).

Note: in evaluating the "behavior of the system," you are asked to plot the state of the system as a function of time along one sample path of the disturbances (i.e., *not* the average of many sample paths). Thus, to "compare the behavior of the system," you only need to plot the state of the system in the cases you are comparing, and wherever possible on the same sample path of the disturbances so that the comparison is meaningful.

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, 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.

Hint: Use as sufficient statistic the pair (x_k, q_k), where

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

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

Problem 5 (Robust Dynamic Programming)

Consider a variation of the basic problem discussed throughout the course: instead of having a probabilistic description of uncertainty, it is just known that w_k belongs 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.