B.2 Problem Set II

(due Wed. Oct. 7)[2]

Remark. As always, please ask for hints if you are stuck.

Remark. In this problem set, you might need the formula for differentiation under the integral sign, which I recall here:

$$\frac{d}{dy}\int_{\alpha(y)}^{\beta(y)} f(y,\xi)d\xi = \int_{\alpha(y)}^{\beta(y)} \frac{df(y,\xi)}{dy}d\xi + f(y,\beta(y))\frac{d\beta(y)}{dy} - f(y,\alpha(y))\frac{d\alpha(y)}{dy}.$$

Problem B.2.1 (reading)**.** Have a look at MATLAB's functions for LQR/LQG design, in the control systems toolbox (look in MATLAB's help and check the LQR/LQG section). The controllers and Kalman filters implemented are the steady-state ones mentioned in class ($k \to -\infty$ in the Riccati difference equation for LQR, $k \to +\infty$ for the Kalman filter). Their implementation requires solving algebraic Riccati equations (ARE), and note that we haven't discussed how to do that in practice. One way is to iterate the Riccati difference equation and wait for convergence, but that's not really an efficient way to solve the ARE. Check the help for the matrix equation solver available in MATLAB to solve this equation (`dare`). You will see that it mentions certain conditions for the solver to work, namely stabilizability and detectability of certain pairs of matrices. For an introduction to the steady-state versions of the LQR and LQG problems, please read pp.151-159 and pp.234-236 in Bertsekas. We definitely won't cover the numerical methods to solve the ARE in this course, so it's good to know that these solvers already exist when you need them. Most likely we won't have time to look at the theory behind the infinite-horizon LQR and LQG problems in more details, because that would mean discussing quite a bit of linear systems theory first for completeness and would take us off-topic.

Problem B.2.2. Do the exercises in the notes of chapter 4 and 5.

Problem B.2.3. Consider the inventory control problem of chapter 3, with no fixed cost ($K = 0$), and a terminal cost $c_N(x_N) = -c\,x_N$, where c is also the unit ordering cost as used in that chapter. This case arises when, at the end of the last period, we can obtain full reimbursement of the leftover units, and must incur the unit cost for each unit backlogged, on top of any shortage penalty incurred in the previous period. Assume that $c_H > 0$ and that the demand w has a continuous cumulative distribution function $F_w(y) := P(w \leq y)$ (you can assume that the probability distribution of w has a density). Let S be the solution of the following equation (the so-called *critical fractile solution*)

$$F(S) = \frac{c_B}{c_B + c_H}. \tag{B.1}$$

Show that the optimal policy is a base stock policy with the same base stock level $S_k = S$ at every period, where S is defined by (B.1) [hint: you will

[2] this version: Oct. 4 2009

probably have to show that $dJ_k^*(x)/dx = -c$ for $x < S$. Do not worry about taking derivatives, e.g. under the integral, and other mathematical technicalities, when you look for the solution; you can clean up your argument later if you want].

Problem B.2.4. Consider the LQR problem of chapter 4, and change in that model only the cost function (4.7) to

$$E\left\{\sum_{k=0}^{N-1}(x_k^T Q_k x_k + 2x^T S^T u + u_k^T R_k u_k) + x_N^T Q_N x_N\right\}. \tag{B.2}$$

That is, we add the cross terms $2x^T S^T u$ in the cost function. Derive the optimal control law. [hint: it's not hard using again the Schur complement, but instead you could try to find a change of control variable u to reduce this problem to the problem solved in the notes - for that, think about square completion...].

Problem B.2.5. Consider the LQR problem with *no disturbances* ($w_k = 0$). Use the discrete-time minimum principle discussed in chapter 2 to derive the solution to this problem. First, write explicitly the adjoint system for the co-state λ. Then, prove using backward induction that $\lambda_k = P_k x_k$, where P_k satisfies the Riccati difference equation. Finally, express the control u_k in terms of P_{k+1} and x_k but not λ_{k+1}.

Problem B.2.6. Consider a *scalar* linear system

$$x_{k+1} = a_k x_k + b_k u_k + w_k, \quad k = 0, 1, \ldots, N-1,$$

where $a_k, b_k \in \mathbb{R}$ and each w_k is a Gaussian random variable with zero mean and variance σ^2. We assume no constraints and independent disturbances. Show that the control law $\{\mu_0^*, \mu_1^*, \ldots, \mu_{N-1}^*\}$ that minimizes the cost function

$$E\Big\{\exp\Big[x_N^2 + \sum_{k=0}^{N-1}(x_k^2 + r u_k^2)\Big]\Big\}, \quad r > 0,$$

is linear in the state variable, assuming the optimal cost is finite for every x_0 (for a bonus, you can discuss when this cost is indeed finite). Does the certainty equivalence principle hold? Show by an example that the Gaussian assumption is essential for the result on the linearity of the optimal control law to hold (for analyses of multidimensional versions of this exercise, see [Jac73, Whi82, Whi90, Bas00]).

Hint: note the formula

$$\int_{-\infty}^{\infty} e^{-(ax^2+bx+c)}dx = \sqrt{\frac{\pi}{a}}e^{(b^2-4ac)/4a}, \text{ for } a > 0.$$

Problem B.2.7. Consider again the output tracking problem B.1.7, with the modified dynamics

$$x_{k+1} = Ax_k + Bu_k + w_k, \; k = 0, 1, \ldots, N-1,$$

where the disturbances w_k are i.i.d. Gaussian, zero mean, with covariance matrix ρI_3, for some $\rho > 0$. Note the following important point on the notation: we continue to assume *full state information*, the notation y_k just denotes the output we are interested in for tracking purposes (so you are allowed to design a controller $\mu_k(x_k)$, function of the state). We want to solve this problem using the LQR solution, which does not allow for the hard control constraint $\|u\|_\infty \leq U_{\max}$. Instead add a control cost in the objective to get

$$J = E\left[\sum_{k=1}^{N} \|y_k - \hat{y}_k\|^2 + u_k^T R u_k\right].$$

Ignoring the hard constraint $\|u\|_\infty \leq U_{\max}$ for the moment, derive the optimal controller for this objective. You can do this yourself (ask for hints if you are stuck), or find a function that is already implemented (maybe a function implementing this is available for MATLAB, I don't know - most likely not). Let me repeat that the controller you are looking for is of the form $x_k \mapsto \mu_k(x_k)$. Next, we want to use this controller heuristically to control the system subject to the constraint $\|u\|_\infty \leq U_{\max}$, for $\rho = 10^{-4}$, and $\rho = 10^{-2}$. The choice of the matrix R is left to you. To enforce $\|u\|_\infty \leq U_{\max} = 0.1$, saturate the control inputs obtained from your design to $U_{\max}$ when these inputs are outside of their allowed range. Note that you can experiment with the matrix R to obtain smaller inputs. Report the best performance $E[\sum_{k=1}^N \|y_k - \hat{y}_k\|^2]$ that you could obtain (not including the control cost), by approximating the expectation using averages over sufficiently many simulations so that you don't see variations in this average that are too large between two experiments. Discuss your simulation results precisely, in particular the impact of the choice of R. Plot some sample paths from your best designs.

Problem B.2.8. Consider a problem with imperfect state information where the system and observations are linear:

$$\begin{aligned} x_{k+1} &= A_k x_k + B_k u_k + w_k, \\ y_k &= C_k x_k + v_k. \end{aligned}$$

For simplicity, assume that all the quantities x_k, u_k, y_k, w_k, v_k are scalars. The initial state x_0 and the disturbances w_k and v_k are assumed Gaussian and mutually independent. x_0 has mean μ_{x_0} and variance $\sigma_{x_0}^2$. w_k, v_k are i.i.d with mean zero and known variances σ_w^2, σ_v^2.

1. Show that $E[x_0|\mathcal{I}_0], \ldots E[x_{N-1}|\mathcal{I}_{N-1}]$ constitute a sufficient statistic for this problem. [hint: start from the fact that the conditional distribution $P_{x_k|\mathcal{I}_k}$ is a sufficient statistic - can you say more about this distribution here ?]

2. Use the previous result to obtain an optimal policy of the single-stage problem involving the scalar system and observation

$$\begin{aligned} x_1 &= x_0 + u_0 \\ y_0 &= x_0 + v_0, \end{aligned}$$

and the cost function $E[|x_1|]$.

3. (bonus) Generalize part 2 for the case of the scalar system

$$\begin{aligned} x_{k+1} &= ax_k + u_k \\ y_k &= cx_k + v_k, \end{aligned}$$

and the cost function $E\left[\sum_{k=1}^{N} |x_k|\right]$. The scalars a and c are known.

Problem B.2.9. Consider a machine that can be in one of two states, good or bad. Suppose that the machine produces an item at the end of each period. The item produced is either good or bad depending on whether the machine is in a good or bad state at the beginning of the corresponding period. We suppose that once the machine is in a bad state it remains in that state until it is replaced. If the machine is in a good state at the beginning of a certain period, then with probability t it will be in the bad state at the end of the period. Once an item is produced, we may inspect the item at a cost I or not inspect. If an inspected item is found to be bad, the machine is replaced with a machine in good state at a cost R. The cost for producing a bad item is $C > 0$. Write a DP algorithm for obtaining an optimal inspection policy assuming a machine initially in good state and a horizon of N periods. Solve the problem for $t = 0.2$, $I = 1$, $R = 3$, $C = 2$, and $N = 8$. (to check your results: the optimal policy is to inspect at the end of the third period and not inspect in any other period).