Problemset 4

Daniel Bestard, Michael Cameron, Hans-Peter Höllwirth, Akhil Lohia

March 16, 2017

1 Linear-Quadratic Problem with Forecasts

First of all let's set up the problem in order to make the proof. The dynamics of the problem is linear function of the form:

$$x_{k+1} = A_k x_k + B_k u_k + w_k$$

and the cost function is a quadratic function of the form:

$$\mathop{\mathbb{E}}_{w_k}\left\{ g_N(x_N) + \sum_{k=0}^{N-1} g_k(x_k, u_k, w_k) \right\}$$

where

$$g_N(x_N) = x_N' Q_N x_N$$

$$g_k(x_k, u_k, w_k) = x_k' Q_k x_k + u_k' R_k u_k$$

The matrices A_k, B_k, Q_k and R_k are given and the last two are positive semidefinite symmetric and positive definite symmetric, respectively.

The DP-algorithm that solves the minimization problem is:

$$J_N(x_N) = x_N' Q_N x_N$$

$$J_k(x_k) = \underset{u_k}{min} \mathop{\mathbb{E}}_{w_k|y_k}\left\{ x_k' Q_k x_k + u_k' R_k u_k + J_{k+1}(A_k x_k + B_k u_k + w_k) \right\}$$

By induction we get that:

$$
\begin{aligned}
J_{N-1}(x_{N-1}) = \min_{u_{N-1}} \mathop{\mathbb{E}}_{w_{N-1}|y_{N-1}} \Big\{ & x'_{N-1}Q_{N-1}x_{N-1} + u'_{N-1}R_{N-1}u_{N-1} + \\
& + \big(A_{N-1}x_{N-1} + B_{N-1}u_{N-1} + w_{N-1}\big)' Q_N \big(A_{N-1}x_{N-1} + B_{N-1}u_{N-1} + w_{N-1}\big) \Big\} \\
& = x'_{N-1}Q_{N-1}x_{N-1} + x'_{N-1}A'_{N-1}Q_N A_{N-1}x_{N-1} + \\
& + \min_{u_{N-1}} \Big\{ u'_{N-1}R_{N-1}u_{N-1} + u'_{N-1}B'_{N-1}Q_N B_{N-1}u_{N-1} + 2x'_{N-1}A'_{N-1}Q_N B_{N-1}u_{N-1} \Big\} + \\
& + \mathop{\mathbb{E}}_{w_{N-1}|y_{N-1}} \Big\{ w'_{N-1}Q_N w_{N-1} + 2x'_{N-1}A'_{N-1}Q_N w_{N-1} \Big\} + \\
& + \min_{u_{N-1}} \mathop{\mathbb{E}}_{w_{N-1}|y_{N-1}} \Big\{ 2u'_{N-1}B'_{N-1}Q_N w_{N-1} \Big\}
\end{aligned}
$$

By differentiating the previous expression and setting it to 0, we obtain the following result:

$$\big(R_{N-1} + B'_{N-1}Q_N B_{N-1}\big)u^*_{N-1} = -B'_{N-1}Q_N A_{N-1}x_{N-1} - B'_{N-1}Q_N \mathbb{E}[w_{N-1}|y_{N-1}]$$

Given the definitions provided previously we can note that the matrix $R_{N-1}+B'_{N-1}Q_N B_{N-1}$ is positive definite, which means that we can invert it and obtain the following optimal value:

$$
\begin{aligned}
u^*_{N-1} &= -\big(R_{N-1} + B'_{N-1}Q_N B_{N-1}\big)^{-1}\big(B'_{N-1}Q_N A_{N-1}x_{N-1} + B'_{N-1}Q_N \mathbb{E}[w_{N-1}|y_{N-1}]\big) \\
&= -\big(R_{N-1} + B'_{N-1}Q_N B_{N-1}\big)^{-1} B'_{N-1}Q_N \big(A_{N-1}x_{N-1} + \mathbb{E}[w_{N-1}|y_{N-1}]\big)
\end{aligned}
$$

Note that the previous expression is already of the form of the expression to be proved. If we substitute back the optimal value u^*_{N-1} in $J_{N-1}(x_{N-1})$ and continue doing this we get the general expression provided in the exercise, which is:

$$\mu^*_k(x_k, y_k) = -\big(R_k + B'_{N-1}K_{k+1}B_k\big)^{-1} B'_k K_{k+1}\big(A_k x_k + \mathbb{E}[w_k|y_k]\big) + \alpha_k$$

where K_{k+1} is a function of A_k, B_k, Q_k and R_k.

2 Computational Assignment on Linear-Quadratic Control

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

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

with $A = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}$, and $\{w_k\}$ being a sequence of IID, zero-mean, Gaussian random vectors with diagonal covariance matrix D. The cost structure is defined as

$$\begin{aligned} g_N(x_N) &= x_N^T C^T C x_N \\ g_k(x_k) &= x_k^T C^T C x_k + u_k^T R u_k \end{aligned}$$

with $C = \begin{bmatrix} 1 & 3 \end{bmatrix}$ and diagonal, positive-definite input-cost matrix R.

2.1 Controllability and observability conditions

Note that the matrix $[B, AB] = \begin{bmatrix} 4 & 2 & 7 & 11 \\ 1 & 3 & 10 & 10 \end{bmatrix}$ is full rank, thus ensuring controllability. Intuitively, *controllability* means that one is able to move the internal state x of a system from any initial state x_0 to any other final state x_k in a finite time period through a control sequence $\{u_k\}$ (ignoring any disturbances). In particular, *controllability* imposes the following structure on the system:

$$x_k = Bu_{k-1} + ABu_{k-2} + ... + A^{k-1}Bu_0 + A^k x_0$$

Furthermore, the observability condition is ensured because matrix $[C^T, A^T C^T] = \begin{bmatrix} 1 & 7 \\ 3 & 9 \end{bmatrix}$ is full rank. Intuitively, *observability* means that it is possible to determine the behavior of the entire system through the system's output $\{x_k\}$. In particular, given measurements of the form $z_k = Cx_k$, one can infer initial state x_0 through the relations

$$\begin{bmatrix} z_{k-1} \\ ... \\ z_1 \\ z_0 \end{bmatrix} = \begin{bmatrix} CA^{k-1} \\ ... \\ CA \\ C \end{bmatrix} x_0$$

2.2 Different covariance matrices

First, we compare the behavior of the system for two covariance matrices for the disturbances, $D_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $D_2 = \begin{bmatrix} 5 & 0 \\ 0 & 6 \end{bmatrix}$, with D_2 "much larger" than D_1, under optimal control (given by the discrete-time Riccati equation). $R = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$ and $x_0 = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$ are fixed.

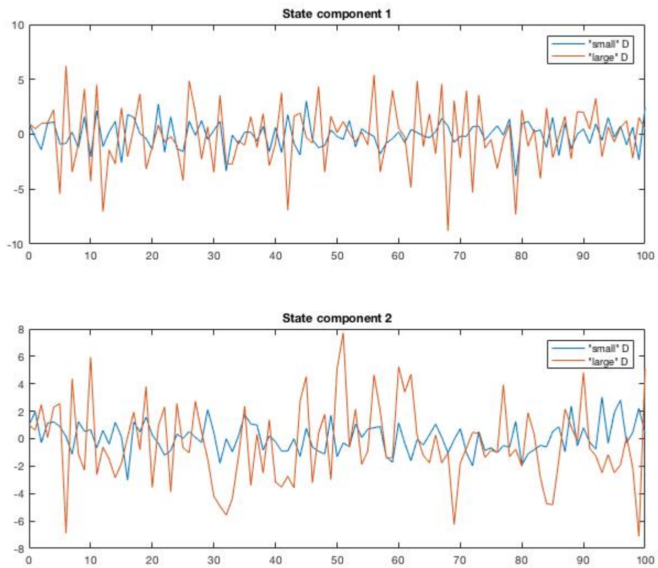

Figure 1: State transitions for covariance matrices D_1 (blue) and D_2 (orange)

We observe from *Figure 1* that in both cases, the state keeps fluctuating around mean 0. However, as expected the fluctuations are larger for the "larger" covariance matrix D_2.

2.3 Different initial states

Next, we compare the behavior of the system for two initial conditions, $x_0^{(1)} = \begin{bmatrix} 1 & 1 \end{bmatrix}$ and $x_0^{(2)} = \begin{bmatrix} 5 & 4 \end{bmatrix}$, with $x_0^{(1)}$ "much larger" than $x_0^{(2)}$, under optimal control (given by the discrete-time Riccati equation). $R = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$ and $D = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ are fixed.

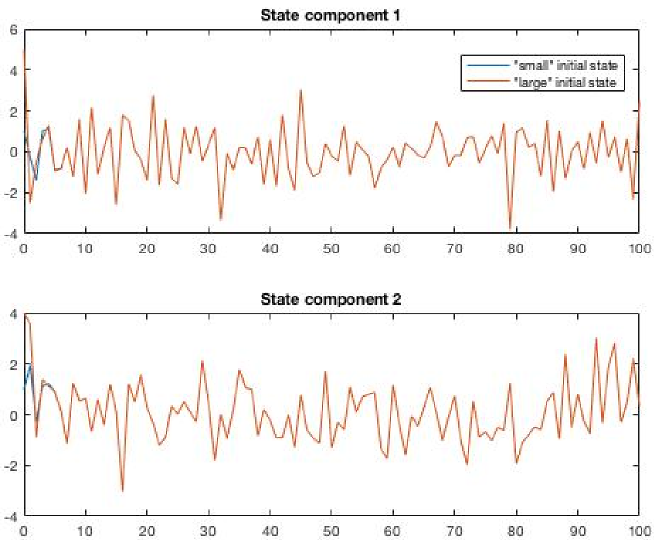

Figure 2: State transitions for initial states $x_0^{(1)}$ (blue) and $x_0^{(2)}$ (orange)

We observe from *Figure 2* that both initial conditions converge to the same state transitions within a few (less than 10) periods.

2.4 Different input-cost matrices

Next, we compare the behavior of the system for two input-cost matrices conditions, $R_1 = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$ and $R_2 = \begin{bmatrix} 18 & 0 \\ 0 & 22 \end{bmatrix}$, with R_2 "much larger" than R_1, under optimal control (given by the discrete-time Riccati equation). $D = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $x_0 = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$ are fixed.

We observe from *Figure 3* that in both cases, the state keeps fluctuating around mean 0. However, both systems always move into the same direction (up or down), the magnitude of the movement is significantly larger with input-cost matrix R_2.

2.5 Optimal control vs. steady-state control

Finally, we compare the behavior of the system under optimal control (given by the discrete-time Riccati equation) vs. steady-state control (given by the algebraic Riccati equation). $D = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $R = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$, and $x_0 = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$ are fixed.

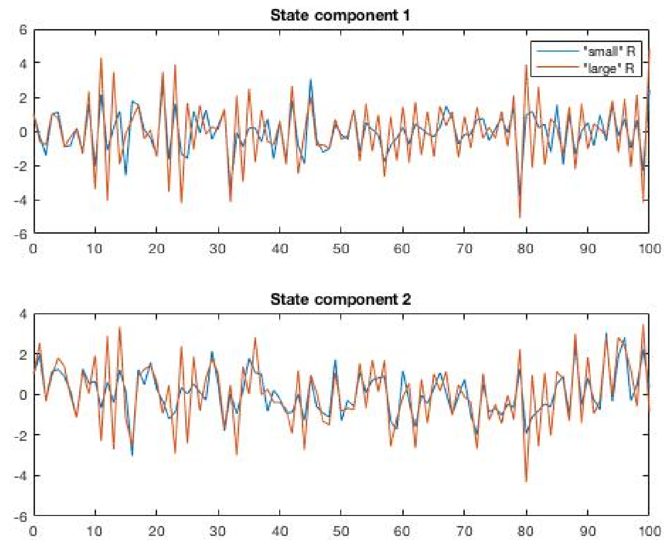

Figure 3: State transitions for input-cost matrices R_1 (blue) and R_2 (orange)

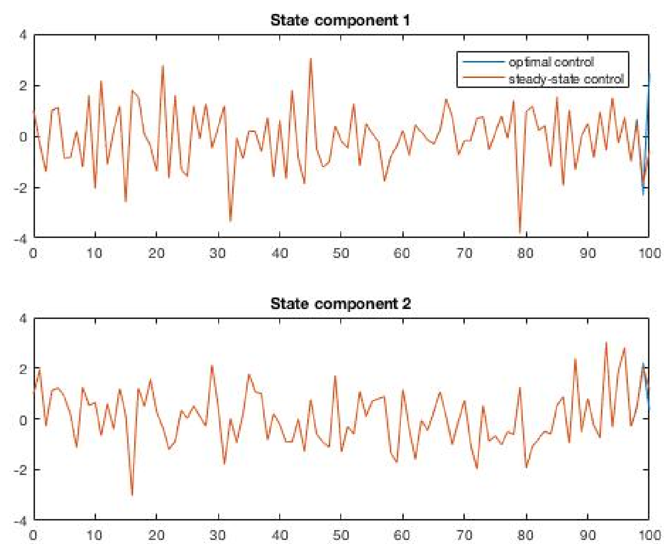

Figure 4: State transitions under optimal control (blue) and steady-state control (orange)

We observe from *Figure 4* that under both controls the system has identical state transitions, except for the last few (less than 5) periods. This result shows that the discrete-time Riccati equations converge to the algebraic solution within few iterations.

3 Asset Selling with Offer Estimation

4 Inventory Control with Demand Estimation

5 Robust Dynamic Programming

Consider a variation of the basic problem in which we do not have a probabilistic description of uncertainty. Instead, we want to find the closed-loop policy $\pi = \{\mu_0(.), .., \mu_{N-1}(.)\}$ with $\mu_k(x_k) \in U_k(x_k)$ that minimizes the maximum possible cost:

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

5.1 DP formulation

Using the principle of optimality, the DP like recursion for this variation of the basic problem looks like:

$$\begin{aligned} J_N(x_N) &= g_N(x_N) \\ J_k(x_k) &= \min_{u_k \in U_k(x_k)} \max_{w_k \in W_k(x_k, \mu_k(x_k))} [g_k(x_k, \mu_k(x_k), w_k) + J_{k+1}(f_k(x_k, \mu_k(x_k), w_k))] \end{aligned}$$

Note that, when we compare the DP formulation to the original basic problem, instead of minimizing the expected cost over w_k, here we minimize the maximum possible cost that can result from an action $u_k = \mu_k(x_k)$.

5.2 Reachability of a target tube

Now assume that at each stage k, the state x_k must belong to a given set X_k. A cost structure that fits the reachability problem within the general formulation looks as follows:

$$\begin{aligned} J_N(x_N) &= \begin{cases} 0 & \text{if } x_N \in X_N \\ 1 & \text{if } x_N \notin X_N \end{cases} \\ J_k(x_k) &= \min_{u_k \in U_k(x_k)} \max_{w_k \in W_k(x_k, \mu_k(x_k))} [J_{k+1}(f_k(x_k, \mu_k(x_k), w_k))] \end{aligned}$$

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, can be computed recursively as follows:

$$\bar{X}_N = X_N$$
$$\bar{X}_k = \{x_k : \exists \mu_k(x_k) \in U_k(x_k). \forall w_k \in W_k(x_k, \mu_k(x_k)). f_k(x_k, \mu_k(x_k), w_k) \in \bar{X}_{k+1}\}$$

For $x_k \in \bar{X}_k$, there must exist at least one action $u_k = \mu_k(x_k)$ in set $U_k(x_k)$ such that every possible outcome w_k in the outcome set $W_k(x_k, \mu_k(x_k))$ takes us to a state $x_{k+1} = f_k(x_k, \mu_k(x_k), w_k) \in \bar{X}_{k+1}$.