ECE 553, Spring 2008 **Problem Set #7 Solution**

Posted: May 2nd, 2008

Solutions:

1. The optimal controller is still the one given in the solution to the Problem 6 in Homework #5:
$$u^*(x,t) = -p(t)x - k(t), \quad t \geq 0.$$
The minimum expected cost is:
$$\begin{aligned}\min E\{J\} &= m(0) + \int_0^1 p(t)\ dt + p(0) + 2k(0) \\ &= m(0) + \frac{4e-2}{3e^2-1} + \ln\left(\frac{3e^2-1}{2e^2}\right),\end{aligned}$$
where
$$m(0) = -\int_0^1 (2k + k^2)\ dt.$$

2. For the LQG problem, we have shown in class that J can be equivalently written as
$$\begin{aligned} J &= E\left\{\int_{t_0}^{t_f} |u + R^{-1}B^T Px|_R^2\ dt\right\} + \int_{t_0}^{t_f} Tr[PDD^T]\ dt \\ &\equiv \sum_{i=0}^{n} E\left\{\int_{t_i}^{t_{i+1}} |u + R^{-1}B^T Px|_R^2\ dt\right\} + \int_{t_0}^{t_f} Tr[PDD^T]\ dt \end{aligned}$$
where $P(\cdot) \geq 0$ satisfies:
$$\dot{P} + A^T P + PA - PBR^{-1}B^T P + Q = 0; \quad P(t_f) = Q_f.$$
But now we don't have perfect state measurement $x(t)$ for all t. In the interval $[t_i, t_{i+1})$, decompose x into two parts, $x = y + z$, where
$$\begin{aligned}\dot{y} &= Ay + Bu, \quad y(t_i) = x(t_i), \\ \dot{z} &= Az + Dw, \quad z(t_i) = 0.\end{aligned}$$
Since u can only use information of $x(t_i)$ during the interval $[t_i, t_{i+1})$, conditioned on this value of $x(t_i)$, u and z are independent for all $t \in [t_i, t_{i+1})$, and hence J is equivalent to:
$$\sum_{i=0}^{n} E\left\{\int_{t_i}^{t_{i+1}} |u + R^{-1}B^T Py|_R^2\ dt\right\} + \sum_{i=0}^{n} E\left\{\int_{t_i}^{t_{i+1}} |R^{-1}B^T Pz|_R^2\ dt\right\} + \int_{t_0}^{t_f} Tr[PDD^T]\ dt.$$
From this, it readily follows that, in the interval $[t_i, t_{i+1})$, the unique optimal control is
$$\begin{aligned}u^*(t) &= -R^{-1}B^T Py(t), \\ \dot{y} &= (A - BR^{-1}B^T P)y, \quad y(t_i) = x(t_i).\end{aligned}$$
Let $\Phi(t,\tau)$ be the state transition matrix corresponds to $\tilde{A} \doteq A - BR^{-1}B^T P$. Then,
$$u^*(t) = -R^{-1}B^T P(t)\Phi(t, t_i)x(t_i), \quad t_i \leq t < t_{i+1}.$$

Note that this controller also obeys the separation principle between estimation and control since it is exactly the controller that solves the optimal control problem for the deterministic case with sampled state measurements.

The corresponding minimum cost of J is

$$J^* = \sum_{i=0}^{n} \int_{t_i}^{t_{i+1}} Tr[PBR^{-1}B^T PM^i(t)]\ dt + \int_{t_0}^{t_f} Tr[PDD^T]\ dt,$$

where M^i satisfies:

$$\dot{M}^i = AM^i + M^i A^T + DD^T, \quad M^i(t_i) = 0.$$

3. Since a is not time-varying, the state space model for this system is

$$\begin{aligned} \dot{x}_1 &= x_2, \quad x_1(0) = 1, \\ \dot{x}_2 &= x_3, \quad x_2(0) = 1, \\ \dot{x}_3 &= 0, \quad\ \ x_3(0) = a \sim N(0, \rho), \end{aligned}$$

with the measurement equation being

$$y = x_1 + v = [1, 0, 0] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + v.$$

Hence, in terms of notation adopted in class, we have

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad H = [1, 0, 0], \quad D = 0.$$

The error covariance equation is

$$\dot{\Sigma} = \Sigma A^T + A\Sigma - \Sigma H^T H \Sigma, \quad \Sigma(0) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \rho \end{bmatrix},$$

which can be solved explicitly to yield:

$$\Sigma(t) = \frac{20\rho}{\rho t^5 + 20} \begin{bmatrix} \frac{t^4}{4} & \frac{t^3}{2} & \frac{t^2}{2} \\ \frac{t^3}{2} & t^2 & t \\ \frac{t^2}{2} & t & 1 \end{bmatrix}, \quad t \geq 0.$$

The Kalman gain is

$$K(t) = \Sigma H^T = \frac{20\rho}{\rho t^5 + 20} \left[\frac{t^4}{4}, \frac{t^3}{2}, \frac{t^2}{2} \right],$$

and the filter for the system is

$$\begin{bmatrix} \dot{\hat{x}}_1 \\ \dot{\hat{x}}_2 \\ \dot{\hat{x}}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \end{bmatrix} + K(t)[y - \hat{x}_1], \quad \begin{bmatrix} \hat{x}_1(0) \\ \hat{x}_2(0) \\ \hat{x}_3(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

In the limit as $t \to \infty$, $\Sigma(t) = 0$ and $K(t) = 0$. This means that the random variable a is identified exactly as $t \to \infty$. Since $x_1(0)$ and $x_2(0)$ are both known, x_1 and x_2 also become exactly estimated as $t \to \infty$. Hence, as $t \to \infty$, the noisy measurements are of little extra information, and the Kalman gain also goes to zero. (This is possible largely due to $D = 0$.)

4. When the state equation and the expected cost function are, respectively,

$$\dot{x} = Ax + Bu + Dw + c, \quad x(t_0) = x_0,$$

$$J(u) = E\left\{x(t_f)^T Q_f x(t_f) + \int_0^{t_f} [x(t)^T Q(t)x(t) + u(t)^T R(t)u(t)]\, dt\right\},$$

with $R(\cdot) > 0$ and $Q(\cdot) \geq 0, Q_f \geq 0$, we have already shown in class that there exists a unique optimal control strategy given by

$$u^*(t) = -R^{-1}B^T[P(t)x(t) + k(t)],$$

where $P(\cdot)$ is the unique nonnegative definite solution of the RDE

$$\dot{P} + PA + A^T P - PBR^{-1}B^T P + Q = 0, \quad P(t_f) = Q_f,$$

and $k(\cdot)$ uniquely solves the ODE

$$\dot{k} + A^T k - PBR^{-1}B^T k + Pc = 0, \quad k(t_f) = 0.$$

(a) Specializing to the problem under consideration, $A = 0, B = 1, c = \frac{t^2}{2}, D = 1, Q_f = 1, Q = 0, R = 1$ yields

$$\begin{aligned} \dot{P} - P^2 &= 0, \qquad & P(1) &= 1, \\ \dot{k} - Pk + Pc &= 0, \qquad & k(1) &= 0. \end{aligned}$$

Solving this gives

$$P(t) = \frac{1}{2-t}, \quad k(t) = \frac{1-t^3}{6(2-t)}.$$

Therefore, the optimal control is

$$u^*(t) = -\frac{1}{2-t}\left[x(t) + \frac{1-t^3}{6}\right].$$

(b) To obtain the open loop optimal control, you should notice that this is very similar to that sampling case that we studied in Problem 2. But we here have a deterministic drift term $c = t^2/2$, and we need to employ the solution to the more general LQG model (as in the case (a)). Similarly, we decompose the state $x(t)$ into a deterministic part $y(t)$ and an uncontrolled stochastic part $z(t)$:

$$\begin{aligned} x &= y + z \\ \dot{y} &= Ay + Bu + c, \quad y(t_0) = x_0 \\ \dot{z} &= Az + Dw, \quad z(t_0) = 0. \end{aligned}$$

The open loop control u is independent of z and $E[z(t)] = 0$. Constructing the same $P(t)$ and $k(t)$ as in (a), J can be rewritten as:

$$\begin{aligned} J &= E\left\{\int_{t_0}^{t_f} |u + R^{-1}B^T(Py + k)|_R^2\, dt\right\} + 2\int_{t_0}^{t_f} k^T(t)c(t)\, dt \\ &+ \int_{t_0}^{t_f} \{Tr[PDD^T] - k^T(t)BR^{-1}B^T k(t)\}\, dt \\ &+ E\left\{\int_{t_0}^{t_f} z(t)^T PBR^{-1}B^T Pz(t)\, dt + x_0^T P(t_0)x_0 + 2x_0^T k(t_0)\right\}. \end{aligned}$$

Only the first term above involves the control u (the sepearation principle again), and this gives the optimal controller

$$u^*(t) = -R^{-1}B^T(Py^*(t) + k(t)),$$

where

$$\dot{y}^* = (A - BR^{-1}B^TP)y^* - BR^{-1}B^Tk + c; \quad y^*(t_0) = x_0$$

and $P(t)$ and $k(t)$ are as above. For the specific problem under consideration, we have

$$\dot{y}^* = -\frac{1}{2-t}y^* - \frac{1-t^3}{6(2-t)} + \frac{t^2}{2}, \quad y^*(0) = 1.$$

This yields

$$y^*(t) = \frac{7}{6} \cdot \frac{2-t}{12} + \frac{t^3-1}{6}.$$

Hence the control is

$$u^*(t) = -\left[\frac{1}{2-t}y^* + \frac{1-t^3}{6(2-t)}\right] = -\frac{7}{72}.$$

Note that interestingly the open loop control is independent of time.

(c) The difference between J_O and J_f is

$$J_O - J_f = E\left\{\int_{t_0}^{t_f} z(t)^T PBR^{-1}B^TPz(t)\,dt\right\} = Tr\int_{t_0}^{t_f} PBR^{-1}B^TPM\,dt,$$

where $M(t) = E[z(t)z(t)^T]$ satisfies the Lyapunov differential equation:

$$\dot{M} = AM + MA^T + DD^T, \quad M(t_0) = 0.$$

For the specific problem, $M(t) = t$ and it gives

$$J_O - J_f = \int_0^1 \frac{t}{(2-t)^2}\,dt = 1 - \ln 2.$$

5. The LQG discounted cost problem was discussed in class under the perfect state feedback. Without noisy state measurements, the controller part will be the same as in that case (except the state is now from the filter), and the filter part will clearly not make use of the discount factor. The parameters are

$$A = 0, \beta = 2, B = 1, D = 1.5, H = 1, G = 1, Q = R = 1.$$

Then $A_\beta = -1$, and the systems are obviously both controllable and observable. Hence the infinite horizon problem is well-defined, *i.e.* there exists an optimal feedback controller that is stationary:

$$u^*(t) = \bar{P}\hat{x}(t),$$

where

$$-\bar{P}^2 + 1 - 2\bar{P} = 0 \;\Rightarrow\; \bar{P} = -1 + \sqrt{2} = 0.414.$$
$$\dot{\hat{x}} = -\bar{P}\hat{x} + K(y - \hat{x}), \quad \hat{x}(0) = 1,$$

where K is the Kalman gain. In our case, $\dot{\Sigma} = -\Sigma^2 + \frac{9}{4}, \Sigma(0) = 0$. For $t \to \infty$, $\Sigma(t) \to \bar{\Sigma} = \frac{3}{2} = K$. Then the steady state optimal control is

$$u^*(t) = -0.414\hat{x}(t), \quad \dot{\hat{x}} = -0.414\hat{x} + 1.5[y - \hat{x}], \quad \hat{x}(0) = 1.$$

The corresponding value of J is

$$\begin{aligned} J^* &= Tr[\bar{P}\Sigma(0)] + Tr[\bar{P}\hat{x}_0^2] + \frac{1}{\beta}Tr[\bar{P}DD^T] + \int_0^\infty e^{-\beta t} Tr[\Sigma \bar{P} B R^{-1} B^T \bar{P}]\, dt \\ &\approx 0.414 \cdot 1 + \frac{1}{2} \cdot \frac{9}{4} \cdot 0.414 + \int_0^\infty e^{-2t}(0.414)^2 \Sigma(t)\, dt. \end{aligned}$$

Solving $\Sigma(t)$ yields $\Sigma(t) = \frac{3}{2} \cdot \frac{1-e^{-3t}}{1+e^{-3t}}$. Substitute it into the expression for J^* and obtain

$$J^* = 0.943131.$$

If we use instead the steady value $\Sigma = \bar{\Sigma}$ for all t, then the value is

$$J^* = 1.008,$$

which is slightly higher, as expected.

Please report to yima@uiuc.edu if you find any typos.