EN530.603 Applied Optimal Control
Midterm #2 Solutions
December 4, 2013

1. Consider the optimal control of the discrete-time nonlinear system

$$x_{k+1} = f(x_k) + G(x_k)u_k$$

where $x_k \in \mathbb{R}^n$ and $u_k \in \mathbb{R}^m$ with cost function

$$J = \frac{1}{2}x_N^T P_N x_N + \sum_{k=0}^{N-1}\left[\frac{1}{2}u_k^T R_k u_k + q(x_k)\right],$$

for given matrices P_N, R_k, and nonlinear functions q_k. The initial state x_0 is given.

a) Formulate Bellman's equation in terms of the value function V_k and derive the optimality conditions for this problem. Are there any requirements on the data (i.e. P_N,R_k, f, G, q) needed to ensure stability, i.e. that V_k decreases with k?

Solution. Bellman's equation takes the form

$$V_k(x_k) = \min_{u_k}\left\{\frac{1}{2}u_k^T R_k u_k + q(x_k) + V_{k+1}\left(f(x_k) + G(x_k)u_k\right)\right\}, \tag{1}$$

where for now we only know the value of V_k at the final boundary, i.e.

$$V_N = \frac{1}{2}x_N^T P_N x_N.$$

Performing the minimizing on the right-hand side of (1) results in the necessary conditions

$$u_k^T R_k + \partial_x V_{k+1}\left(f(x_k) + G(x_k)u_k\right) \cdot G(x_k) = 0, \tag{2}$$

or equivalently the optimal control is determined implicitly by the relationship

$$u_k^* = R_k^{-1} G(x_k)^T \nabla_x V_{k+1}\left(f(x_k) + G(x_k)u_k^*\right),$$

assuming R_k is symmetric and invertible. A sufficient condition for computing a minimizing u^* is that the Hessian of the right hand side in (1) is positive definite, i.e.

$$R_k + G(x_k^*)^T \nabla_x^2 V_{k+1}(x_{k+1}^*) G(x_k^*) > 0, \tag{3}$$

where the notation x_k^* indicates that the optimal control u^* has been used in evolving the discrete trajectory $x_{0:N}^*$. Next, we have that

$$V_k(x_k^*) = \frac{1}{2}u_k^* R_k u_k^* + q(x_k^*) + V_{k+1}(x_{k+1}^*),$$

and so the value function will decrease along solution trajectories $x^*_{0:N}$ i.e. $V_{k+1}(x^*_{k+1}) < V_k(x^*_k)$ only when

$$q(x^*_k) > -\frac{1}{2}(u^*_k)^T R_k u^*_k. \tag{4}$$

b) What are suitable methods for computing numerically the optimal sequence of controls $\{u^*_0, u^*_1, \ldots, u^*_{N-1}\}$? Sketch the process of setting up the problem numerically, based on your chosen method.

Solution. Since there are no constraints and the problem is already discrete, one can use either direct shooting (setup an NLP over $u_{0:N-1}$ only), direct collocation (setup an NLP over over $x_{1:N}, u_{0:N-1}$), or sweep methods such as stage-wise Newton or differential dynamic programming (DDP). For instance, in DDP we define the Q-function

$$Q_i(x_i, u_i) = \frac{1}{2}u_i^T R_i u_i + q(x_i) + V_{i+1}(f(x_i) + G(x_i)u_i), \tag{5}$$

expanded it to second-order

$$\Delta Q_i \approx \frac{1}{2}\begin{bmatrix} 1 \\ \delta x_i \\ \delta u_i \end{bmatrix}\begin{bmatrix} 0 & \nabla_x Q_i^T & \nabla_u Q_i^T \\ \nabla_x Q_i & \nabla_x^2 Q_i & \nabla_{xu} Q_i \\ \nabla_u Q_i & \nabla_{ux} Q_i & \nabla_u^2 Q_i \end{bmatrix}\begin{bmatrix} 1 \\ \delta x_i \\ \delta u_i \end{bmatrix}$$

and compute the optimal change in the control by

$$\delta u^*_i = K_i \cdot \delta x_i + \alpha_i k_i, \tag{6}$$

where

$$K_i = -\nabla_u^2 Q_i^{-1} \nabla_{ux} Q_i, \quad k_i = -\nabla_u^2 Q_i^{-1} \nabla_u Q_i.$$

which involves a backward pass to compute K_i and k_i and a forward pass to update the controls and states.

c) Denote the optimal trajectory computed in b) by $\{x^*_0, x^*_1, \ldots, x^*_N\}$. Imagine that we now want to derive a controller applicable also in the vicinity of x^* which you assume can be expressed according to

$$u(x_k) = u^*_k + K_k(x_k - x^*_k),$$

for any x_k that is close to x^*_k. Give the exact expression for the gain matrix K_k.

Solution. There are several ways to proceed depending on how one chooses to approximate the nonlinear problem locally around the computed optimal trajectory $x^*_{0:N}$ and controls $u^*_{0:N}$. The most straightforward approach is to linearize the dynamics and minimize second order terms in the cost using LQR (first order terms ΔJ need not be considered since we are expanding around the optimal x^* where $\Delta J(x^*, u^*) = 0$ already). Another approach is to use the gain matrix K computed by DDP above. A third way is to use the gain matrix resulting from stage-wise Newton. Interestingly, all these will give similar but not exactly identical solutions. The LQR approach is to define

$$dx_k = x_k - x^*_k, \qquad du_k = u_k - u^*_k$$

and perform the optimization of the second-order cost terms given by

$$\Delta^2 J = \frac{1}{2} dx_N^T P_N dx_N + \frac{1}{2} \sum_{k=0}^{N-1} \left[du_k^T R_k du_k + dx_k^T Q_k dx_k) \right],$$

where $Q_k = \nabla_x^2 q(x_k^*)$, subject to the linearized dynamics

$$dx_{k+1} = A_k dx_k + B_k du_k,$$

where $A_k = \partial_x f(x_k^*)$ and $B_k = G_k(x_k^*)$. The optimal feedback control (derived in class) is

$$du_k^* = -(R_k + B_k^T P_{k+1} B_k)^{-1} B_k^T P_{k+1} A_k dx_k$$

where P_k is defined by

$$P_k = Q_k + A_k^T [P_{k+1} - P_{k+1} B_k (R_k + B_k^T P_{k+1} B_k)^{-1} B_k^T P_{k+1}] A_k.$$

and is computed backwards from the given boundary condition P_N. The required gain matrix K_k for our particular system is thus given by

$$K_k = -[R_k + G(x_k^*)^T P_{k+1} G(x_k^*)]^{-1} G(x_k^*)^T P_{i+1} \partial f(x_k^*).$$

2. Consider the first-order system with a scalar state x and dynamics

$$\dot{x}(t) = a x(t) + w(t),$$

for a given constant a, with disturbance density given by

$$E[w(t)] = 0, \qquad E[w(t)^2] = q_c \delta(t - \tau),$$

where q_c is the standard spectral density of $w(t)$.

a) Consider the discrete-time analog given

$$x_k = a_{k-1} x_{k-1} + w_{k-1},$$

for a given sampling rate Δt. Give the expressions for a_{k-1} and define the density of w_{k-1} (i.e. by computing the variance of w_{k-1} denoted by q_{k-1})

Solution. The term a_{k-1} corresponds to the state transition matrix and is defined by

$$a_{k-1} = e^{a\Delta t}.$$

The covariance of the discrete noise w_{k-1} is computed as

$$\begin{aligned} q_{k-1} &= \int_{t_{k-1}}^{t_k} e^{(t_k - \tau)a} \; q_c \; e^{(t_k - \tau)a} d\tau \\ &= q_c \int_{t_{k-1}}^{t_k} e^{2(t_k - \tau)a} d\tau \\ &= \frac{q_c}{2a} (e^{2a\Delta t} - 1). \end{aligned}$$

Note that if the exponential above is expanded to first order, it is also possible to use the approximation

$$q_{k-1} \approx \Delta t q_c,$$

where terms of order Δt^2 and higher have been discarded.

b) There is a sensor measuring x at each sampling interval Δt defined by the function

$$z = \cos(x) + v,$$

where v is a zero-mean noise term with variance $E[v^2] = r$. Assume that at some time t, there is an estimate of the system state with mean and variance given by

$$E[x(t)] = \hat{x}, \qquad E[(x(t) - \hat{x})^2] = p$$

A noisy measurement with value $\tilde{z}$ arrives at time $t + \Delta t$. What is your best estimate for the state and its variance at time $t + \Delta t$ in view of this measurement. Provide the full expressions for these quantities.

Solution. A first-order accurate solution is found by direct application of the extended Kalman filter (EKF). Let time t correspond to time index $k-1$ and let $(\hat{x}_{k|k-1}, p_{k|k-1})$ and $(\hat{x}_{k|k}, p_{k|k})$ denote the propagated means and covariances before and after measurement correction at time $t + \Delta t$, respectively. We have

$$\begin{aligned}
&\text{Prediction:} \\
&\quad \hat{x}_{k|k-1} = e^{a\Delta t}\hat{x} \\
&\quad p_{k|k-1} = pe^{2a\Delta t} + q_{k-1},
\end{aligned}$$

The measurement Jacobian is then given by

$$H = -\sin(\hat{x}_{k|k-1}).$$

and the prediceted values are substituted in the EKF correction step to obtain

$$\begin{aligned}
&\text{Correction:} \\
&\quad \hat{x}_{k|k} = e^{a\Delta t}\hat{x} + K[\tilde{z} - \cos(e^{a\Delta t}\hat{x})], \\
&\quad p_{k|k} = (1 - KH)(pe^{2a\Delta t} + q_{k-1}),
\end{aligned}$$

where the gain K is

$$K = \frac{(pe^{2a\Delta t} + q_{k-1})H}{(pe^{2a\Delta t} + q_{k-1})H^2 + r}.$$