CS294-40 Learning for Robotics and Control Lecture 12 - 10/07/2008

Separation Principle, Dynamics Modeling

Lecturer: Pieter Abbeel *Scribe: Pål From*

1 Announcements

- Milestone report: due on Sunday; $1-2$ pages with the results so far, $\frac{1}{2}-1$ page of future plans.

2 Separation Principle

Assume we have a linear system

$$x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1} \qquad \text{for } k = 1, 2, \ldots, H \tag{1}$$

with the quadratic cost

$$\mathbb{E}\left[x_H^\top P_H x_H + \sum_{k=0}^{H-1} (u_k^\top R_k u_k + x_k^\top P_k x_k)\right] \tag{2}$$

The input disturbances w_k are assumed to be independent, zero mean and have finite variance.

We need to find a rule for the control u_t given I_t. I_t contains the information available to the controller at time t, i.e. any (noisy) observations of the states at previous times $k = 1 \ldots t$ as well as the previous controls, u_k for $k = 1 \ldots t-1$.

We start by solving for the optimal policy for time $H-1$:

$$\arg\min_{u_{H-1}} \mathbb{E}\left[u_{H-1}^\top R_{H-1} u_{H-1} + x_H^\top P_H x_H \mid I_{H-1}\right] \tag{3}$$

We start by rewriting the second term, in particular, we will use the following:

$$\begin{aligned}
&\mathbb{E}\left[\left[x_H - \mathbb{E}\left[x_H \mid I_{H-1}\right]\right]^\top P_H \mathbb{E}\left[x_H - \mathbb{E}\left[x_H \mid I_{H-1}\right]\right] \mid I_{H-1}\right] \\
&= \mathbb{E}\left[x_H^\top P_H x_H \mid I_{H-1}\right] + \mathbb{E}\left[\mathbb{E}\left[x_H \mid I_{H-1}\right]^\top P_H \mathbb{E}\left[x_H \mid I_{H-1}\right] \mid I_{H-1}\right] - 2\mathbb{E}\left[x_H^\top P_H \mathbb{E}\left[x_H \mid I_{H-1}\right] \mid I_{H-1}\right] \\
&= \mathbb{E}\left[x_H^\top P_H x_H \mid I_{H-1}\right] + \mathbb{E}\left[x_H \mid I_{H-1}\right]^\top P_H \mathbb{E}\left[x_H \mid I_{H-1}\right] - 2\mathbb{E}\left[x_H \mid I_{H-1}\right]^\top P_H \mathbb{E}\left[x_H \mid I_{H-1}\right] \\
&= \mathbb{E}\left[x_H^\top P_H x_H \mid I_{H-1}\right] - \mathbb{E}\left[x_H \mid I_{H-1}\right]^\top P_H \mathbb{E}\left[x_H \mid I_{H-1}\right]
\end{aligned} \tag{4}$$

Using Eqn. (4) and linearity of expection, we can re-write Eqn. (3) as follows:

$$\arg\min_{u_{H-1}} \left(u_{H-1}^\top R_{H-1} u_{H-1} + \mathbb{E}\left[x_H \mid I_{H-1}\right]^\top P_H \mathbb{E}\left[x_H \mid I_{H-1}\right] + \mathbb{E}\left[\left[x_H - \mathbb{E}\left[x_H \mid I_{H-1}\right]\right]^\top P_H \left[x_H - \mathbb{E}\left[x_H \mid I_{H-1}\right]\right] \mid I_{H-1}\right]\right) \tag{5}$$

Interestingly, Eqn. (5) shows that *for a quadratic cost, the expected cost-to-go can be split into the cost-to-go for the expected state* $\mathbb{E}[x_H|I_{H-1}]$, *and an additional term* $\mathbb{E}\left[\left[x_H - \mathbb{E}\left[x_H \mid I_{H-1}\right]\right]^\top P_H \left[x_H - \mathbb{E}\left[x_H \mid I_{H-1}\right]\right] \mid I_{H-1}\right]$, *which accounts for the cost incurred by the uncertainty about the state.*

We will now use the second particularly interesting fact about the linear quadratic setting: $[x_H - \mathbb{E}[x_H \mid I_{H-1}]]$ is independent of $u_{0:H-1}$ so that we can exclude this term from the minimization. This property relies on the linearity of the system.

Intuitively, this property means that *the estimation error is not influenced by the control inputs we apply for a linear system.* We do this by showing that the linear terms in u are repeated in x_H and $\mathbb{E}[x_H|I_{H-1}]$ and thus cancel.

This can be seen by writing out the expressions for x_H and $\mathbb{E}[x_H|I_{H-1}]$:

$$\begin{aligned} x_H &= Ax_{H-1} + Bu_{H-1} + w_{H-1} \\ &= A^H x_0 + A^{H-1} w_0 + A^{H-2} w_1 + A^{H-3} w_3 + \cdots + w_{H-1} \\ &\quad + A^{H-1} B u_0 + A^{H-2} B u_1 + \cdots + B u_{H-1} \end{aligned} \tag{6}$$

and

$$\begin{aligned} \mathbb{E}[x_H|I_{H-1}] = A^H \mathbb{E}[x_0|I_{H-1}] + A^{H-1}\mathbb{E}[w_0|I_{H-1}] + \cdots + \mathbb{E}[w_{H-1}|I_{H-1}] \\ + A^{H-1}B\mathbb{E}[u_0|I_{H-1}] + \cdots + B \underbrace{\mathbb{E}[u_{H-1}|I_{H-1}]}_{u_{H-1}(I_{H-1})} \end{aligned} \tag{7}$$

By observing the two expressions we see that since the control enter the expressions as linear terms and we have a linear system, the difference will not be affected by our choice of $u_{0:H-1}$. We can thus exclude the term

$$\mathbb{E}\left[[x_H - \mathbb{E}[x_H \mid I_{H-1}]]^\intercal P_H [x_H - \mathbb{E}[x_H \mid I_{H-1}]] \mid I_{H-1}\right] \tag{8}$$

from (5) and obtain the following certainty equivalent expression:

$$\begin{aligned} &\arg\min_{u_{H-1}} \mathbb{E}\left[u_{H-1}^\intercal R_{H-1} u_{H-1} + x_H^\intercal P_H x_H \mid I_{H-1}\right] \\ &= \arg\min_{u_{H-1}} \left(u_{H-1}^\intercal R_{H-1} u_{H-1} + \mathbb{E}[x_H \mid I_{H-1}]^\intercal P_H [x_H \mid I_{H-1}]\right) \end{aligned} \tag{9}$$

We can now use that $x_H = Ax_{H-1} + Bu_{H-1} + w_{H-1}$ and get

$$\arg\min_{u_{H-1}} \left(u_{H-1}^\intercal R_{H-1} u_{H-1} + \mathbb{E}[Ax_{H-1} + Bu_{H-1} + w_{H-1}]^\intercal P_H \mathbb{E}[Ax_{H-1} + Bu_{H-1} + w_{H-1}]\right) \tag{10}$$

Up to the noise w_{H-1}, we now have the same setting as in Lecture 6 (which covered the linear quadratic regulator setting). Using a similar derivation, and the fact that w_{H-1} is assumed to be zero-mean and independent of the other variables, we obtain:

$$u_{H-1} = K_{H-1}\mathbb{E}[x_{H-1}|I_{H-1}]$$

for

$$K_{H-1} = -(R_{H-1} + B_{H-1}^T P_H B_{H-1})^{-1} B_{H-1}^T P_H A_{H-1}.$$

Now we plug this into our original objective, as we still have to solve for $u_0, \ldots, u_{H-2}$:

$$\arg\min_{u_0,\ldots,u_{H-2}} \mathbb{E}\left[\mathbb{E}[x_{H-1}|I_{H-1}]^\top P_{H-1}\mathbb{E}[x_{H-1}|I_{H-1}] + \sum_{t=0}^{H-2} u_t^\top R_t u_t + x_t^\top Q_t x_t\right]$$

for

$$P_{H-1} = Q_{H-1} + K_{H-1}^T R_{H-1} K_{H-1} + (A_{H-1} + B_{H-1}K_{H-1})^T Q_H (A_{H-1} + B_{H-1}K_{H-1}). \tag{11}$$

Now, we proceed by solving for u_{H-2} in a similar fashion. First observe, by using the same derivation as in Eqn. (4), that:

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{E}\left[x_{H-1} \mid I_{H-1}\right]^{\top} P_{H-1} \mathbb{E}\left[x_{H-1} \mid I_{H-1}\right] | I_{H-2}\right] \\
&= \mathbb{E}\left[x_{H-1} \mid I_{H-2}\right]^{\top} P_{H-1} \mathbb{E}\left[x_{H-1} \mid I_{H-2}\right] \\
&\quad + \mathbb{E}\left[\left[x_{H-1} - \mathbb{E}\left[x_{H-1} \mid I_{H-2}\right]\right]^{\top} P_{H-1} \left[x_{H-1} - \mathbb{E}\left[x_{H-1} \mid I_{H-2}\right]\right] \mid I_{H-2}\right]
\end{aligned} \tag{12}$$

We can show similar to earlier when solving for u_{H-1} that the last term does not contribute to the minimization, but will of course affect the total cost. We simply cannot affect the uncertainty of the state by our control inputs.

We repeat the same reasoning for every time step $t = H-1, H-2, \ldots, 0$.

Hence, for linear systems with quadratic cost, the following procedure results in optimal control:

- Estimate the states of the system with a Kalman filter, i.e. $\mathbb{E}\left[x_t \mid I_t\right]$
- LQR controller - controller assuming the outputs of the Kalman filter to be true, i.e. using $\mathbb{E}\left[x_t \mid I_t\right]$ in the controller as though it were the true state x_t.

This is known as the *separation principle* for linear systems with quadratic costs: we don't have to explicitly account for uncertainty when deciding on our control inputs. We can be optimal by solving the estimation and the control problem separately. The estimator gives the optimal estimates of the states assuming no control and the controller is optimal assuming perfect state estimation.

Challenge problem: Can you find other systems for which the separation principle applies?

3 Modeling

We will consider an example dynamics model for a helicopter.

3.1 Helicopter model

We use the following state space to represent the state of the helicopter:

$$\text{state:} \qquad (n, e, d, \dot{n}, \dot{e}, \dot{d}, \underbrace{q_x, q_y, q_z, q_w}_{quaternion}, p, q, r) \tag{13}$$

where the quaternion represents a rotation θ about the axis $\vec{n} = [n_x, n_y, n_z]$, $\|\vec{n}\| = 1$ and can be written as

$$\begin{aligned}
q_x &= n_x \sin\frac{\theta}{2} \\
q_y &= n_y \sin\frac{\theta}{2} \\
q_z &= n_z \sin\frac{\theta}{2} \\
q_w &= \cos\frac{\theta}{2}
\end{aligned}$$

Note: two quaternions with opposite signs represent the same physical rotation, i.e.

$$q(\vec{n}, \theta + 2\pi) = -q(\vec{n}, \theta). \tag{14}$$

We have the following inputs

$$\text{input:} \qquad (\underbrace{\underbrace{u_{aileron}}_{\text{roll rate}}, \underbrace{u_{elevator}}_{\text{pitch rate}}}_{\text{cyclic control}}, \underbrace{u_{rudder}}_{\text{yaw rate}}, \underbrace{u_{collective}}_{\text{vertical trust}}) \tag{15}$$

Cyclic control means that the angle of the blade changes throughout the cycle.

3.2 Dynamic model

The dynamics model is given by

$$\begin{aligned} n_{t+1} &= n_t + \Delta t \cdot \dot{n}_t \\ e_{t+1} &= e_t + \Delta t \cdot \dot{e}_t \\ d_{t+1} &= d_t + \Delta t \cdot \dot{d}_t \end{aligned}$$

and for the quaternion

$$(q_x, q_y, q_z, q_w)_{t+1} = (q_x, q_y, q_z, q_w)_t * (\sin \frac{\theta}{2}\vec{n}, \cos \frac{\theta}{2}) \tag{16}$$

where $*$ is the quaternion product and

$$\vec{n} = \frac{(p, q, r)\Delta t}{\|(p, q, r)\Delta t\|_2}, \qquad \theta = \|(p, q, r)\Delta t\|_2. \tag{17}$$

Further, the moments are given by

$$\begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} = \begin{bmatrix} p \\ q \\ r \end{bmatrix} \times \left(I \cdot \begin{bmatrix} p \\ q \\ r \end{bmatrix} \right) + I \cdot \begin{bmatrix} \dot{p} \\ \dot{q} \\ \dot{r} \end{bmatrix} \tag{18}$$

The inertia matrix is given by

$$I = \begin{bmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{bmatrix}. \tag{19}$$

Often, approximating I by a constant times the identity matrix works well in practice. This allows (18) to be simplified: in particular, the first term on the right-hand side is then zero.

We have

$$\begin{aligned} p_{t+1} &= p_t + \Delta t \cdot \dot{p}_t \\ q_{t+1} &= q_t + \Delta t \cdot \dot{q}_t \\ r_{t+1} &= r_t + \Delta t \cdot \dot{r}_t \end{aligned}$$

The linear accelerations are found from

$$\begin{bmatrix} F_n \\ F_e \\ F_c \end{bmatrix} = m \begin{bmatrix} \ddot{n} \\ \ddot{e} \\ \ddot{d} \end{bmatrix} \tag{20}$$

This gives us the update for the velocities:

$$\begin{aligned} \dot{n}_{t+1} &= \dot{n}_t + \Delta t \cdot \ddot{n}_t \\ \dot{e}_{t+1} &= \dot{e}_t + \Delta t \cdot \ddot{e}_t \\ \dot{d}_{t+1} &= \dot{d}_t + \Delta t \cdot \ddot{d}_t \end{aligned}$$

So far we assumed $F_n, F_e, F_d, T_x, T_y, T_z$ were given.

It remains to study how to find the forces as a function of the inputs and the states, i.e., to find the following functions $f_\cdot$:

$$
\begin{aligned}
F_n &= f_{F_n}(s, u) & \qquad T_x &= f_{T_x}(s, u) \\
F_e &= f_{F_e}(s, u) & \qquad T_y &= f_{T_y}(s, u) \\
F_d &= f_{F_d}(s, u) & \qquad T_z &= f_{T_z}(s, u)
\end{aligned}
$$

There are two main approaches to this problem. The first is an in-debt study of fluid dynamics. Alternatively, we can fly the helicopter and find the function that fits the data the best. We will look into this second approach in the next lecture.