Data-driven and Learning-based Control

3rd hands-on session

Erica Salvato

UNIVERSITÀ DEGLI STUDI DI TRIESTE

Table of Contents

What we know so far?

LQR requires linear systems or, at least, systems whose behavior is well approximated by that of the equilibrium points' surroundings. Another approach consists of repeating the

following step until convergence

1. Use linearization of the nonlinear system around a trajectory that is considered *nominal*
2. Compute a locally optimal feedback control law
3. Use the locally optimal feedback control law to obtain a new nominal trajectory

What we know so far?

LQR requires linear systems or, at least, systems whose behavior is well approximated by that of the equilibrium points' surroundings. Another approach consists of repeating the following step until convergence

1. Use linearization of the nonlinear system around a trajectory that is considered *nominal*
2. Compute a locally optimal feedback control law
3. Use the locally optimal feedback control law to obtain a new nominal trajectory

iLQR

What we know so far?

In practice would like to solve the problem of finding $\pi^*\left(x^{(k)}\right) = -Kx^{(k)}$ solution of:

$$\arg\min_{\pi} \sum_{k=0}^{\infty} {x^{(k)}}^{\top} Q x^{(k)} + {u^{(k)}}^{\top} R u^{(k)}$$

s.t.:

$$\mathbf{x}^{(\mathbf{k+1})} = \mathbf{f}\left(\mathbf{x}^{(\mathbf{k})}, \mathbf{u}^{(\mathbf{k})}\right)$$

where $f : \mathcal{X} \times \mathcal{U} \rightarrow \mathcal{X}$ is the non-linear state transition function

What we know so far?

The iLQR procedure is as follows:

1. Record a state-input trajectory of H steps from the non-linear system $\left\{\bar{x}^{(k)}, \bar{u}^{(k)}\right\}_{k=0}^{H-1}$

What we know so far?

1 A brief recap

The iLQR procedure is as follows:

1. Record a state-input trajectory of H steps from the non-linear system $\left\{\bar{x}^{(k)}, \bar{u}^{(k)}\right\}_{k=0}^{H-1}$

2. Linearize the non-linear system around the trajectory

$$\delta x^{(k+1)} = A(k)\delta x^{(k)} + B(k)\delta u^{(k)}$$

where $A(k) = \left[\frac{\partial f}{\partial x}\right]_{x=\bar{x}^{(k)}, u=\bar{u}^{(k)}}$, $B(k) = \left[\frac{\partial f}{\partial u}\right]_{x=\bar{x}^{(k)}, u=\bar{u}^{(k)}}$, while $\delta x^{(k)} = x^{(k)} - \bar{x}^{(k)}$ and $\delta u^{(k)} = u^{(k)} - \bar{u}^{(k)}$ are the variations of $x^{(k)}$ and $u^{(k)}$ from their nominal values.

The cost term then becomes

$$\left(x^{(k)} - \bar{x}^{(k)}\right)^\top Q \left(x^{(k)} - \bar{x}^{(k)}\right) + \left(u^{(k)} - \bar{u}^{(k)}\right)^\top R \left(u^{(k)} - \bar{u}^{(k)}\right)$$

3. Compute the quadratic approximation

$$\hat{x}^{(k)^\top} \hat{Q}(k) \hat{x}^{(k)} + \hat{u}^{(k)^\top} R \hat{u}^{(k)}$$

where $\hat{x}^{(k)} = \begin{bmatrix} x^{(k)} - \bar{x}^{(k)} \\ 1 \end{bmatrix}$ and $\hat{u}^{(k)} = u^{(k)} - \bar{u}^{(k)}$ is the homogeneous transformation of the time-variant linearization of the non-linear system

$$\hat{x}^{(k+1)} = \begin{bmatrix} A(k) & 0 \\ 0 & 1 \end{bmatrix} \hat{x}^{(k)} + \begin{bmatrix} B(k) \\ 0 \end{bmatrix} u^{(k)} = \hat{A}(k)\hat{x}^{(k)} + \hat{B}(k)u^{(k)}$$

while $\hat{Q}(k) \begin{bmatrix} Q & -q^{(k)} \\ -q^{(k)^\top} & d^{(k)} \end{bmatrix}$ where $\tilde{x}^{(k)^\top} Q \tilde{x}^{(k)} = d^{(k)}$ and $2\tilde{x}^{(k)^\top} Q = 2q^{(k)^\top}$.

4. Solve an LQR problem:
   Find $\pi^*\left(\hat{x}^{(k)}\right) = -K_k\hat{x}^{(k)}$ solution of:

$$\arg\min_{\pi} \sum_{k=0}^{\infty} \hat{x}^{(k)^\top} \hat{Q(k)} \hat{x}^{(k)} + \hat{u}^{(k)^\top} R\hat{u}^{(k)}$$

   s.t.:

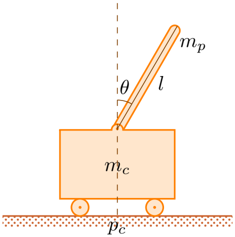
$$\hat{x}^{(k+1)} = \hat{A}(k)\hat{x}^{(k)} + \hat{B}(k)\hat{u}^{(k)}$$

   The solution is a controller K_k which returns $u^{(k)} = -K_k\left(x^{(k)} - \bar{x}^{(k)}\right) + \bar{u}^{(k)}$

5. Apply the controller on the non-linear system and store a new state-input trajectory of H steps from the non-linear system $\left\{\bar{x}^{(k)}, \bar{u}^{(k)}\right\}_{k=0}^{H-1}$ and repeat form step 2.

Table of Contents

2 A case study



$$\ddot{\theta} = \frac{g \sin(\theta) \; + \; \cos(\theta) \left[\frac{-F - m_p \, l \, \dot{\theta}^2 \sin(\theta)}{m_c + m_p} \right] - \frac{\mu_p \dot{\theta}}{m_p \, l}}{l \left[\frac{4}{3} - \frac{m_p \, cos^2(\theta)}{m_c + m_p} \right]}$$

$$\ddot{p}_c = \frac{F \; + \; m_p \, l \left[\dot{\theta}^2 \; \sin(\theta) \; - \; \ddot{\theta} \; \cos(\theta) \right]}{m_c \; + \; m_p}$$

Table of Contents

Hands-on outline

Given the cart and pole system previously defined, we would like to apply a first iteration of iLQR. Therefore, we have to perform the following steps:

1. Simulate the non-linear cart and pole system applying the provided control input sequence $\bar{u}^{(k)}$ with a sample time of 0.005 s. Record the resulting state trajectory $\bar{x}^{(k)}$ with an initial state of $\theta = 180^\circ$, $\dot{\theta} = 0$, $p_c = 0$, $\dot{p}_c = 0$. Plot the resulting trajectory.

Expectation: Create a code able to produce the nominal state trajectory from a nominal input trajectory

Hands-on outline

2. Linearize the non-linear cart and pole system around the nominal state-input trajectories $\{\bar{x}^{(k)}, \bar{u}^{(k)}\}$ previously obtained, and store the time-varying matrices of the linearized system $\{A(k), B(k)\}$

Expectation: Create a code able to produce the *time-varying* matrices of the linearized system around the nominal state-input trajectories $\{\bar{x}^{(k)}, \bar{u}^{(k)}\}$

Hands-on outline

2. Linearize the non-linear cart and pole system around the nominal state-input trajectories $\{\bar{x}^{(k)}, \bar{u}^{(k)}\}$ previously obtained, and store the time-varying matrices of the linearized system $\{A(k), B(k)\}$

Expectation: Create a code able to produce the *time-varying* matrices of the linearized system around the nominal state-input trajectories $\{\bar{x}^{(k)}, \bar{u}^{(k)}\}$

3. Compute the discrete-time homogeneous representation of the linear time-varying cart and pole system obtained at point 2 and store the matrices of the homogeneous representation

Expectation: Create a code able to convert the continuous-time time-varying linearization of the cart and pole system into a homogeneous discrete-time linear time-variant system.

4. Assume $Q = I$ and $R = 1$, the cost matrices for a swing-up cart and pole task, and compute $\hat{Q}(k) = \begin{bmatrix} Q & -q^{(k)} \\ -q^{(k)^\top} & d^{(k)} \end{bmatrix}$ where $\bar{x}^{(k)^\top} Q \bar{x}^{(k)} = d^{(k)}$ and $2\bar{x}^{(k)^\top} Q = 2q^{(k)^\top}$ for the entire nominal state trajectory and store them

Expectation: Create a code able to provide as output the quadratic approximation of the state cost component at each time instant.

Hands-on outline

4. Assume $Q = I$ and $R = 1$, the cost matrices for a swing-up cart and pole task, and compute $\hat{Q}(k) = \begin{bmatrix} Q & -q^{(k)} \\ -q^{(k)^\top} & d^{(k)} \end{bmatrix}$ where $\bar{x}^{(k)^\top} Q \bar{x}^{(k)} = d^{(k)}$ and $2\bar{x}^{(k)^\top} Q = 2q^{(k)^\top}$ for the entire nominal state trajectory and store them

Expectation: Create a code able to provide as output the quadratic approximation of the state cost component at each time instant.

5. Solve the finite-horizon LQR problem on the linearized version of the system and store the control gain matrices $\{K(k)\}$. Assume that $P_H = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$.

Expectation: Create a code able to solve a Riccati recursion starting from the terminal cost $\hat{x}^{(H)^\top} P_H \hat{x}^{(H)} = 0$. Recalling that the system is time-varying and also the cost.

Hands-on outline

6. Simulate the non-linear cart and pole system applying the computed control gain matrices $\{K(k)\}$ with a sample time of 0.005 sec. Record the resulting state and control input trajectory with an initial state of $\theta = 180^\circ$, $\dot{\theta} = 0$, $p_c = 0$, $\dot{p_c} = 0$. Plot the resulting trajectories

Expectation: Crate a code able to compute the state-input trajectories obtained by applying the control gains of the previous step on the non-linear system. (Recall you computed the gain based on the linearized version of the cart and pole)

6. Simulate the non-linear cart and pole system applying the computed control gain matrices $\{K(k)\}$ with a sample time of 0.005 sec. Record the resulting state and control input trajectory with an initial state of $\theta = 180^\circ$, $\dot{\theta} = 0$, $p_c = 0$, $\dot{p_c} = 0$. Plot the resulting trajectories

Expectation: Crate a code able to compute the state-input trajectories obtained by applying the control gains of the previous step on the non-linear system. (Recall you computed the gain based on the linearized version of the cart and pole)

If you repeat the procedure starting again from 2., and updating at each iteration the nominal trajectories with those obtained at 6., you can create an iLQR algorithm. You can stop the algorithm when $\theta \leq 10^\circ$ thus achieving the swing-up task.

Questions’ time!

UNIVERSITÀ DEGLI STUDI DI TRIESTE