ROB-GY 6323 Reinforcement Learning And Optimal Control For Robotics — Implementation of a SQP for Non linear Optimal Control

Vineela Reddy Pippera Badguna [vp2504]

October 2024

1. Algorithm Outline

1. **Initialize:** Set initial states $\theta_0 = 0$, $\omega_0 = 0$, and guess values for θ_n, ω_n and control sequence $u_n \ for N = 0$ to 300 ; initialize guess Lagrange multipliers λ. and

2. **Iterate until convergence:**
   - Compute gradient and Hessian of cost, Jacobian $G(\bar{x})$, and constraint residual $g(\bar{x})$.
   - Solve the KKT system for primal (Δx) and dual ($\Delta\lambda$) steps.
   - Perform line search for step size α, then update $\bar{x}$ and λ.
   - Check convergence: $\|\nabla_x L\| <$ tol and $\|g(\bar{x})\| <$ tol.

2. Gradient of the Running Cost

The running cost is:

$$J(\theta_n, \omega_n, u_n) = \sum_{n=0}^{300} 10(\theta_n - \pi)^2 + 0.1\omega_n^2 + 0.1u_n^2$$

The gradient with respect to θ_n, ω_n, and u_n is:

$$\nabla J(\bar{x}) = \begin{bmatrix} 2 \cdot 10(\bar{\theta}_0 - \pi) \\ 2 \cdot 0.1 \cdot \bar{\omega}_0 \\ 2 \cdot 0.1 \cdot \bar{u}_0 \\ \vdots \\ 2 \cdot 10(\bar{\theta}_{300} - \pi) \\ 2 \cdot 0.1 \cdot \bar{\omega}_{300} \\ 2 \cdot 0.1 \cdot \bar{u}_{300} \end{bmatrix}$$

3. Hessian of the Running Cost

The Hessian is block diagonal, with each block:

$$H_n = \begin{bmatrix} 20 & 0 & 0 \\ 0 & 0.2 & 0 \\ 0 & 0 & 0.2 \end{bmatrix}$$

The full Hessian is:

$$H = \mathrm{diag}(H_0, H_1, \ldots, H_{300})$$

4. Linear Approximation of the Constraints

The pendulum dynamics are:

$$\begin{aligned} \theta_{n+1} &= \theta_n + \Delta t \cdot \omega_n \\ \omega_{n+1} &= \omega_n + \Delta t \cdot (u_n - g \sin \theta_n) \end{aligned}$$

The Jacobian at time step n is:

$$G_n = \begin{bmatrix} 1 & \Delta t & 0 \\ -\Delta t \cdot g \cos(\bar{\theta}_n) & 1 & \Delta t \end{bmatrix}$$

The residual vector $g(\bar{x})$ is:

$$g(\bar{x}) = \begin{bmatrix} (\theta_0 + \Delta t \cdot \omega_0) - \theta_1 \\ (\omega_0 + \Delta t \cdot (u_0 - g \sin \theta_0)) - \omega_1 \\ (\theta_1 + \Delta t \cdot \omega_1) - \theta_2 \\ (\omega_1 + \Delta t \cdot (u_1 - g \sin \theta_1)) - \omega_2 \\ \vdots \\ (\theta_{299} + \Delta t \cdot \omega_{299}) - \theta_{300} \\ (\omega_{299} + \Delta t \cdot (u_1 - g \sin \theta_{299})) - \omega_{300} \end{bmatrix}$$

The linearized constraint for all steps is:

$$G(\bar{x}) \Delta x = g(\bar{x})$$

where $G(\bar{x})$ is block diagonal, composed of G_n for all steps.

5. Inequality Constraints

For each control input u_n at each time step n, the inequality bounds are:

$$-4 \leq u_n \leq 4, \quad \text{for } n = 0, 1, 2, \ldots, N$$

where $N = 300$.

$$H = \begin{bmatrix} 0 & 0 & -1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & -1 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & -1 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$h\_vec = \begin{bmatrix} u_0 + 4 \\ u_0 - 4 \\ u_1 + 4 \\ u_1 - 4 \\ \vdots \\ u_{299} + 4 \\ u_{299} - 4 \\ u_{300} + 4 \\ u_{300} - 4 \end{bmatrix}$$

6. Answers to Questions from Notebook

Question 1, SQP solver:
320 iterations (`convergence criteria met, target reached`).

Question 2, SQP solver with inequality constraints:
531 iterations (`convergence criteria met, target & max_iter not reached`).

All the plots for each question are available in the notebook.

Comparing solution of question 2 with question 1.

In Question 1, the solution converges after 320 iterations, successfully guiding the pendulum to reach the target state. However, in Question 2, despite meeting the convergence criteria, the pendulum fails to reach the target. This limitation likely arises from the restricted control input range of $[-4, 4]$, which may not provide sufficient force to achieve the desired outcome. Adjusting the weight on the control input did not yield improvement, and the convergence in Question 2 is slower compared to Question 1.