Infinite-Horizon Stochastic Optimal Control

Introduction

In this project, we will designed a trajectory tracking algorithm for a ground differential-drive robot. The objective follow the trajectory as close as possible without hitting any obstacles. We will try two different approach to solve this problem, Certainty Equivalent Control (CEC). CEC convert the stochastic problem to a deterministic one by fixing the noise variable at its exited value. Our implementation of Value Iteration will only update value function for N iterations instead of infinite numbers of iterations.

Problem Formulation

We can formulate the trajectory tracking problem as a discounted infinite-horizon stochastic optimal control problem. Define following variables:

$$
\begin{array}{ll}
\text{robot position:} & \mathbf{p} = [p_x, p_y]^\top \in \mathbb{R}^2 \\
\text{robot orientation:} & \theta \in [-\pi, \pi) \\
\text{reference position:} & \mathbf{r} = [r_x, r_y]^\top \in \mathbb{R}^2 \\
\text{reference orientation:} & \alpha \in [-\pi, \pi) \\
\text{robot position error:} & \tilde{\mathbf{p}} = \mathbf{p} - \mathbf{r} \\
\text{robot orientation error:} & \tilde{\theta} = \theta - \alpha \\
\text{linear velocity:} & v \in \mathbb{R} \\
\text{angular velocity:} & \omega \in \mathbb{R}
\end{array}
$$

Define error dynamics:

$$
\mathbf{e}_{t+1} = \begin{bmatrix} \tilde{\mathbf{p}}_{t+1} \\ \tilde{\theta}_{t+1} \end{bmatrix} = g(t, \mathbf{e}_t, \mathbf{u}_t, \mathbf{w}_t) = \underbrace{\begin{bmatrix} \tilde{\mathbf{p}}_t \\ \tilde{\theta}_t \end{bmatrix}}_{\mathbf{e}_t} + \underbrace{\begin{bmatrix} \tau \cos(\tilde{\theta}_t + \alpha_t) & 0 \\ \tau \sin(\tilde{\theta}_t + \alpha_t) & 0 \\ 0 & \tau \end{bmatrix}}_{\tilde{\mathbf{G}}(\mathbf{e}_t)} \underbrace{\begin{bmatrix} v_t \\ \omega_t \end{bmatrix}}_{\mathbf{u}_t} + \begin{bmatrix} \mathbf{r}_t - \mathbf{r}_{t+1} \\ \alpha_t - \alpha_{t+1} \end{bmatrix} + \mathbf{w}_t
$$

The objective is to find the policy that minimize the total costs:

$$
\begin{aligned}
V^*(\tau, \mathbf{e}) = \min_{\pi} \mathbb{E} & \left[\sum_{t=\tau}^{\infty} \gamma^t \left(\tilde{\mathbf{p}}_t^\top \mathbf{Q} \tilde{\mathbf{p}}_t + q(1-\cos(\tilde{\theta}_t))^2 + \mathbf{u}_t^\top \mathbf{R} \mathbf{u}_t \right) \mid \mathbf{e}_\tau = \mathbf{e} \right] \\
\text{s.t.} \quad & \mathbf{e}_{t+1} = g(t, \mathbf{e}_t, \mathbf{u}_t, \mathbf{w}_t), \quad \mathbf{w}_t \sim \mathcal{N}(\mathbf{0}, \operatorname{diag}(\boldsymbol{\sigma})^2), \quad t = \tau, \tau+1, \ldots \\
& \mathbf{u}_t = \pi(t, \mathbf{e}_t) \in \mathcal{U} \\
& \tilde{\mathbf{p}}_t + \mathbf{r}_t \in \mathcal{F}
\end{aligned}
$$

Technical Approach

Certainty Equivalent Control (CEC)

The Certainty Equivalent Control simplifies the problem by fix the noise $\mathbf{w}$ to its expected value, so in our case $\mathbf{w} = 0$, which means we would assume there is no noise. In addition, instead of looking at infinite time horizon $T = \infty$, we only look at finite-horizon for $T = k$ for some $k \in \mathbb{N}$. Therefore, the optimal value function V^* can be sampled as following:

$$
\begin{aligned}
V^*(\tau, \mathbf{e}) = \min_{\mathbf{u}_\tau, \ldots, \mathbf{u}_{\tau+T-1}} & \mathfrak{q}(\mathbf{e}_{\tau+T}) + \sum_{t=\tau}^{\tau+T-1} \gamma^t \left(\tilde{\mathbf{p}}_t^\top \mathbf{Q} \tilde{\mathbf{p}}_t + q(1-\cos(\tilde{\theta}_t))^2 + \mathbf{u}_t^\top \mathbf{R} \mathbf{u}_t \right) \\
\text{s.t.} \quad & \mathbf{e}_{t+1} = g(t, \mathbf{e}_t, \mathbf{u}_t, \mathbf{0}), \quad t = \tau, \ldots, \tau+T-1 \\
& \mathbf{u}_t \in \mathcal{U} \\
& \tilde{\mathbf{p}}_t + \mathbf{r}_t \in \mathcal{F}
\end{aligned}
$$

$$
\mathfrak{q}(\mathbf{e}_{\tau+T-1}) = \tilde{\mathbf{p}}_{\tau+T-1}^\top \mathbf{Q} \tilde{\mathbf{p}}_{\tau+T-1} + q(1-\cos(\tilde{\theta}_{\tau+T-1}))^2
$$

The above value function $V^*(\tau, e_\tau) = f(\mathbf{U}, \mathbf{E})$ where $\mathbf{U} = [\mathbf{u}_\tau^\top, \ldots, \mathbf{u}_{\tau+T-1}^\top]^\top$ and $\mathbf{E} = [\mathbf{e}_\tau^\top, \ldots, \mathbf{e}_{\tau+T}^\top]^\top$.

Since we require the car to avoid obstacles, we need the distance between the car to the center of the circle to be larger than the radius of the circle.

$$
(p_x - c_x)^2 + (p_y - c_y)^2 \geq r^2
$$

To rewrite this contain in terms of the optimization variables, we have $\mathbf{h} = ||\tilde{\mathbf{p}} + \mathbf{r} - \mathbf{c}||^2 \geq r^2$, where $\mathbf{c}$ is the center of the circle and r is the radius of the circle.

Now we can formulate the CEC problems a non-linear program of the form:

$$
\begin{aligned}
&\min_{\mathbf{U}} f(\mathbf{U}, \mathbf{E}) \\
\text{st. } &\mathbf{U}_{lb} \leq \mathbf{U} \leq \mathbf{U}_{ub} \\
&\mathbf{h}_{lb} \leq \mathbf{h(U, E)} \leq \mathbf{h}_{ub} \\
\text{where } &\mathbf{U}_{lb} = [0, -1]^\top,\ \mathbf{U}_{ub} = [1, 1]^\top,\ \mathbf{h}_{lb} = r^2,\ \mathbf{h}_{ub} = \infty
\end{aligned}
$$

Collision Avoidance

CEC is able to constantly avoid obstacles in the noise free case. It also has decent performance when noise is present.