Theoretical Background

Problem Statement

Consider an arbitrary system with nonlinear discrete-time dynamics $x(t+\Delta t) = f(x(t), u(t))$. Let $x \in \mathbb{R}^{n_x \times (N+1)}$ and $u \in \mathbb{R}^{n_u \times N}$ be the state and input trajectory of the system over a finite time-horizon N and denote the cost of this trajectory as $J(x, u)$. Then compose the following finite-horizon nonlinear optimal control problem:

$$\begin{aligned} \min_{x,u} \quad & J(x,u) \\ \text{subj. to } & x_0 = x(t) \\ & x_{i+1} = f(x_i, u_i) \quad \forall i \in [0, N-1] \end{aligned}$$

Probabilistic Inference-Based Predictive Control

Introduce an *optimality* parameter $\mathcal{O}$ depending on the system trajectory (x, u), which we define as

$$\mathcal{O} := \begin{cases} 1 & \text{if } (x,u) \text{ are optimal} \\ 0 & \text{else} \end{cases}$$

Of course, $\mathcal{O}(x, u)$ cannot easily be computed in advance. Instead, we take a probabilistic approach and assume $p(\mathcal{O}=1|x,u)$ follows a Boltzmann distribution with respect to the trajectory cost $J(x, u)$. By defining the expected cost for a sequence of control inputs as $\bar{J}(u) = \mathbb{E}_{p(x|u)}[J(x,u)]$, we obtain the probability distribution that a given sequence of inputs is optimal as

$$p(\mathcal{O}=1|u) = \eta^{-1} \exp(-\lambda^{-1} \bar{J}(u))$$

where λ denotes the *temperature* and η is the normalization constant.
Then, using Bayes' theorem:

$$\begin{aligned} p(u|\mathcal{O}=1) &\propto p(\mathcal{O}=1|u) p(u) \\ &\propto Z^{-1} \exp(-\lambda^{-1} \bar{J}(u)) p(u) \end{aligned}$$

where Z is the new normalization constant and $p(u)$ is the prior distribution of u.
Since $\int_{-\infty}^{\infty} p(u|\mathcal{O}=1) \mathrm{d}u = 1$, we can compute

$$Z = \int_{-\infty}^{\infty} \exp(-\lambda^{-1} \bar{J}(u)) p(u) \mathrm{d}u = \mathbb{E}_{p(u)}[\exp(-\lambda^{-1} \bar{J}(u))]$$

Directly sampling u from $p(u|\mathcal{O})$ is generally challenging as it can be a highly complex distribution. Therefore, we approximate it with a variational distribution $q(u;\theta)$ whose parameters θ are found by minimizing the Kullback-Leibler divergence:

$$\theta = \arg\min_{\theta} \mathrm{D}_{\mathcal{KL}}(p(u|\mathcal{O}=1) || q(u|\theta))$$

Model Predictive Path Integral Control (MPPI)

Choose a Gaussian distribution with fixed covariance as the variational distribution $q(u|\mu) = \mathcal{N}(u; \mu, \Sigma)$ with mean $\mu \in \mathbb{R}^N$ and covariance matrix $\Sigma \in \mathbb{R}^{N \times N}$.

$$\begin{aligned}
\mu &= \arg\min_{\mu} \mathrm{D}_{\mathcal{KL}} \left(p(u|\mathcal{O}=1) \| q(u;\mu) \right) \\
&= \arg\min_{\mu} \mathbb{E}_{p(u|\mathcal{O}=1)} \left[\log p(u|\mathcal{O}=1) - \log \mathcal{N}(u;\mu,\Sigma) \right] \\
&= \arg\max_{\mu} \mathbb{E}_{p(u|\mathcal{O}=1)} \left[\log \mathcal{N}(u;\mu,\Sigma) \right]
\end{aligned}$$

$$\begin{aligned}
\mathcal{L}(\mu,\Sigma) &:= \mathbb{E}_{p(u|\mathcal{O}=1)} \left[\log \mathcal{N}(u;\mu,\Sigma) \right] \\
&= \mathbb{E}_{p(u|\mathcal{O}=1)} \left[-\frac{1}{2} \log((2\pi)^{n_u N} |\Sigma|^{n_u}) - \frac{1}{2} (u-\mu)^\top \Sigma^{-1} (u-\mu) \right] \\
&= -\frac{1}{2} \left(n_u N \log 2\pi - n_u \log |\Sigma^{-1}| + \mathbb{E}_{p(u|\mathcal{O}=1)} \left[(u-\mu)^\top \Sigma^{-1} (u-\mu) \right] \right)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \mathcal{L}(\mu,\Sigma)}{\partial \mu} &= -\frac{1}{2} \mathbb{E}_{p(u|\mathcal{O}=1)} [-2\Sigma^{-1}(u-\mu)] \\
&= \Sigma^{-1} \left(\mathbb{E}_{p(u|\mathcal{O}=1)}[u] - \mu \right) = 0
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \mu &= \mathbb{E}_{p(u|\mathcal{O}=1)}[u] \\
&= Z^{-1} \int_{-\infty}^{\infty} \exp(-\lambda^{-1} \bar{J}(u)) p(u) u \mathrm{d}u \\
&= \frac{\mathbb{E}_{p(u)}[\exp(-\lambda^{-1}\bar{J}(u))u]}{\mathbb{E}_{p(u)}[\exp(-\lambda^{-1}\bar{J}(u))]}
\end{aligned}$$

$\mathbb{E}_{p(u)}[\,\cdot\,]$ can be numerically approximated using Monte-Carlo methods and the prior $p(u)$ is typically taken as the previous optimal solution.

MPPI Algorithm

1. Sample K control input sequences $u^k \sim p(u) = \mathcal{N}(u;\mu,\Sigma)$
2. Simulate the system to obtain trajectories (x^k, u^k) and compute the associated trajectory costs $\bar{J}^k = J(x^k,u^k) - \min_k J(x^k,u^k)$
3. Find optimal μ through Monte-Carlo: $\mu \leftarrow \sum_{k=1}^{K} \mathrm{softmax}(-\lambda^{-1}\bar{J}^k) u^k$

Updating the Variational Covariance

Updating not only the mean μ but also the covariance Σ seems desirable as we would like to tighten the variance in the sampled control input sequences once we are close to the optimum. Similarly to above with μ, we can compute a derivative with respect to Σ as

$$\begin{aligned}
\frac{\partial \mathcal{L}(\mu,\Sigma)}{\partial \Sigma^{-1}} &= -\frac{1}{2} \frac{\partial}{\partial \Sigma^{-1}} \left(-\log |\Sigma^{-1}| + \mathbb{E}_{p(u|\mathcal{O}=1)} \left[(u-\mu)^\top \Sigma^{-1} (u-\mu) \right] \right) \\
&= \frac{1}{2} \left(\Sigma - \mathbb{E}_{p(u|\mathcal{O}=1)} \left[(u-\mu)(u-\mu)^\top \right] \right) = 0
\end{aligned}$$

$$
\begin{aligned}
\Rightarrow \Sigma &= \mathrm{Var}_{p(u|\mathcal{O}=1)}[u] \\
&= Z^{-1} \int_{-\infty}^{\infty} \exp(-\lambda^{-1} \bar{J}(u)) p(u)(u-\mu)(u-\mu)^\top \mathrm{d}u \\
&= \frac{\mathbb{E}_{p(u)}[\exp(-\lambda^{-1} \bar{J}(u))(u-\mu)(u-\mu)^\top]}{\mathbb{E}_{p(u)}[\exp(-\lambda^{-1} \bar{J}(u))]}
\end{aligned}
$$

Heuristically however, contrary to updating μ, the Monte-Carlo method does not work great for updating Σ in the same manner in practice. For the reason of limiting computational expense, we aim to keep K fairly small, which often leads to an underestimation of the optimal Σ. This can be solved to some extent by increasing the temperature λ, but finding a good balance between fast convergence of the mean and sufficiently large covariance for exploration seems difficult.