Intro to Model-Predictive Control and Model-Predictive Path Integral Control

Ian Abraham

May 28, 2020

Outline

Stochastic Optimal Control (Model Predictive Path Integral Control)

Talking points

- free-energy formulation of optimal control problem
- how to construct stochastic control problem
- derive optimal sample-based actions using MPPI

Concepts you should know!

- how to take expectations (and what they are)
- importance sampling – switching sample distributions in expectations
- Jensen's inequality

What in Path Integral Control?

- sample-based approach to solving stochastic MPC

What in Path Integral Control?

- sample-based approach to solving stochastic MPC
- does not require continuity/differentiability/convex assumptions

What in Path Integral Control?

- sample-based approach to solving stochastic MPC
- does not require continuity/differentiability/convex assumptions
- optimal actions generated through sampling from distribution of possible actions

What in Path Integral Control?

- sample-based approach to solving stochastic MPC
- does not require continuity/differentiability/convex assumptions
- optimal actions generated through sampling from distribution of possible actions
- requires lots of computations

Notation and Preliminaries: Dynamics

$s_{t+1} = f(s_t, v_t)$ assume general form of discrete stochastic dynamics

$v_t \sim \mathcal{N}(a_t, \Sigma)$ model uncertainty in how control signal enters dynamics

$\Sigma \in \mathbb{R}^{m \times m}$ is noise variance

Notation and Preliminaries: Distributions

$\mathbb{P}$: probability of action sequence in uncontrolled system (zero mean $a_t = 0$)

$\mathbb{Q}$: probability of action sequence in *open-loop* controlled system (mean of a_t)

Notation and Preliminaries: Distributions

$\mathbb{P}, \mathbb{Q}$: distributions representing actions associated with the uncontrolled system and open loop controlled system

$\mathbb{P}, \mathbb{Q}$ defined through density function

$$p(v) \propto \Pi_{t=0}^{T-1} \exp\left(-\frac{1}{2} v_t^\top \Sigma^{-1} v_t\right)$$
$$q(v) \propto \Pi_{t=0}^{T-1} \exp\left(-\frac{1}{2} \|v_t - a_t\|_{\Sigma^{-1}}^2\right)$$

for sequence of random actions $v = [v_0, \ldots, v_{T-1}]$

Notation and Preliminaries: Objective

assume general cost function $\ell(s)$ —remove dependence on a_t for now,

$$\min \mathbb{E}\left[\mathcal{J}(v) = \sum_{t=0}^{T-1} \ell(s_t) + m(s_T)\right]$$
$$\text{subject to } s_{t+1} = f(s_t, v_t)$$

since actions are stochastic, objective now taken at expectation
expectation over *all possible actions* v over some candidate distribution.

Free Energy Formulation

the tuple $(f, \mathcal{J}, \lambda)$ represents the free energy of control system
models how stochastic systems can be maintained in non-equilibrium states

Free Energy Formulation

the tuple $(f, \mathcal{J}, \lambda)$ represents the free energy of control system
models how stochastic systems can be maintained in non-equilibrium states

the free energy of stochastic control problem is defined as

$$\mathcal{F}(v) = -\lambda \log \left(\mathbb{E}_{\mathbb{P}} \left[\exp \left(-\frac{1}{\lambda} \mathcal{J}(v) \right) \right] \right)$$

where $\lambda \in \mathbb{R}^{+}$ is known as the temperature of the system (behaves like a scaling parameter)

Free Energy Formulation: Importance Sampling + Jensen's Inequality

what can the free energy tell us about solving the optimal control problem?

$$\mathcal{F}(v) = -\lambda \log \left(\mathbb{E}_{\mathbb{P}} \left[\exp \left(-\frac{1}{\lambda} \mathcal{J}(v) \right) \right] \right)$$

can we relate the free energy to the stochastic control problem?

Free Energy Formulation: Importance Sampling + Jensen's Inequality

what can the free energy tell us about solving the optimal control problem?

$$\begin{aligned}\mathcal{F}(v) &= -\lambda \log\left(\mathbb{E}_{\mathbb{P}}\left[\exp\left(-\frac{1}{\lambda}\mathcal{J}(v)\right)\right]\right) \\ &= -\lambda \log\left(\mathbb{E}_{\mathbb{Q}}\left[\frac{p(v)}{q(v)}\exp\left(-\frac{1}{\lambda}\mathcal{J}(v)\right)\right]\right)\end{aligned}$$

- change expectation to $\mathbb{Q}$ using importance sampling

Free Energy Formulation: Importance Sampling + Jensen's Inequality

what can the free energy tell us about solving the optimal control problem?

$$\begin{aligned}
\mathcal{F}(v) &= -\lambda \log\left(\mathbb{E}_{\mathbb{P}}\left[\exp\left(-\frac{1}{\lambda}\mathcal{J}(v)\right)\right]\right) \\
&= -\lambda \log\left(\mathbb{E}_{\mathbb{Q}}\left[\frac{p(v)}{q(v)}\exp\left(-\frac{1}{\lambda}\mathcal{J}(v)\right)\right]\right) \\
&\leq -\lambda \mathbb{E}_{\mathbb{Q}}\left[\log\left(\frac{p(v)}{q(v)}\exp\left(-\frac{1}{\lambda}\mathcal{J}(v)\right)\right)\right]
\end{aligned}$$

- apply Jensen's inequality to simplify problem and present a very important bound

Free Energy Formulation: Simplifying

nice bound

$$\mathcal{F}(v) \leq -\lambda \mathbb{E}_{\mathbb{Q}} \left[\log \left(\frac{p(v)}{q(v)} \exp \left(-\frac{1}{\lambda} \mathcal{J}(v) \right) \right) \right]$$

but what does it all mean?

Free Energy Formulation: Simplifying

expanding out the log gives the free energy bound as

$$\begin{aligned}\mathcal{F}(v) &\leq -\lambda \mathbb{E}_{\mathbb{Q}}\left[\log\left(\frac{p(v)}{q(v)}\exp\left(-\frac{1}{\lambda}\mathcal{J}(v)\right)\right)\right] \\ &= \mathbb{E}_{\mathbb{Q}}\left[\mathcal{J}(v) - \lambda\log\left(\frac{p(v)}{q(v)}\right)\right]\end{aligned}$$

we get optimal control objective plus log term under $\mathbb{Q}$

Free Energy Formulation: Simplifying

expanding out the log gives the free energy bound as

$$\begin{aligned}
\mathcal{F}(v) &\leq -\lambda \mathbb{E}_{\mathbb{Q}}\left[\log\left(\frac{p(v)}{q(v)}\exp\left(-\frac{1}{\lambda}\mathcal{J}(v)\right)\right)\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[\mathcal{J}(v) - \lambda \log\left(\frac{p(v)}{q(v)}\right)\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[\mathcal{J}(v) + \frac{\lambda}{2}\sum_{t=0}^{T-1} a_t^\top \Sigma_t a_t\right]
\end{aligned}$$

free energy bound is just optimal control problem + regularization!

Working with distributions

free energy bounds the optimal control problem!
what can we do with that?

$$\mathcal{F}(v) \leq \mathbb{E}_{\mathbb{Q}}\left[\mathcal{J}(v) - \lambda \log\left(\frac{p(v)}{q(v)}\right)\right]$$

Working with distributions

free energy bounds the optimal control problem!
what can we do with that?

$$\mathcal{F}(v) \leq \mathbb{E}_{\mathbb{Q}}\left[\mathcal{J}(v) - \lambda \log\left(\frac{p(v)}{q(v)}\right)\right]$$

what if $\frac{p(v)}{q(v)} \propto \frac{1}{\exp(-\frac{1}{\lambda}\mathcal{J}(v))}$?

Working with distributions

free energy bounds the optimal control problem!
what can we do with that?

$$\mathcal{F}(v) \leq \mathbb{E}_{\mathbb{Q}}\left[\mathcal{J}(v) - \lambda \log\left(\frac{p(v)}{q(v)}\right)\right]$$

what if $\frac{p(v)}{q(v)} \propto \frac{1}{\exp(-\frac{1}{\lambda}\mathcal{J}(v))}$?

you end up with

$$\mathcal{F}(v) = \text{constant}$$

convince yourself

Working with distributions

what does this mean?

stoch. control problem at constant equilibrium (optima) when $\frac{p(v)}{q(v)} \propto \frac{1}{\exp(-\frac{1}{\lambda}\mathcal{J}(v))}$

we have a condition for optimality from free energy bound!
optimize free parameter a_t (mean) in $q(v)$

Optimal Distribution

using the condition for optimality, we can define the optimal control distribution $\mathbb{Q}^\star$ through its probability density function

$$q^\star(v) \propto \exp\left(-\frac{1}{\lambda}\mathcal{J}(v)\right) p(v)$$

we can't sample directly from $\mathbb{Q}^\star$, but we can evaluate whether action samples are likely optimal

KL-Control

we have $\mathbb{Q}$, but how can we get $\mathbb{Q}$ closer to $\mathbb{Q}^{\star}$?

KL-Control

we have $\mathbb{Q}$, but how can we get $\mathbb{Q}$ closer to $\mathbb{Q}^\star$?

use a distance measure between distributions (KL-divergence for example)

$$a^\star = \arg\min_a D_{\mathrm{KL}}\left(\mathbb{Q}^\star \| \mathbb{Q}\right)$$

where $a = [a_0, \ldots, a_{T-1}]$

KL-Control

KL-divergence is defined as

$$D_{\mathrm{KL}}(\mathbb{Q}^{\star}\|\mathbb{Q}) = \int_{\Omega} q^{\star}(v)\log\left(\frac{q^{\star}(v)}{q(v)}\right) dv$$

where $\Omega \subset \mathbb{R}^{m} \times \{0, \ldots, T-1\}$ is the sample space

KL-Control

let's optimize KL-divergence to get $a^\star$!

$$a^\star = \arg\min_a \left[\int_\Omega q^\star(v) \log\left(\frac{q^\star(v)}{q(v)}\right) dv \right]$$

KL-Control

using importance sampling, we can show that

$$
\begin{aligned}
a^\star &= \arg\min_a \left[\int_\Omega q^\star(v) \log\left(\frac{q^\star(v)}{q(v)}\right) dv \right] \\
&= \arg\min_a \int_\Omega q^\star(v) \log\left(\frac{q^\star(v)}{p(v)} \frac{p(v)}{q(v)}\right) dv \\
&= \arg\min_a \int_\Omega q^\star(v) \log\left(\frac{q^\star(v)}{p(v)}\right) dv - \int_\Omega q^\star(v) \log\left(\frac{q(v)}{p(v)}\right) dv
\end{aligned}
$$

KL-Control

using importance sampling, we can show that

$$\begin{aligned}
a^\star &= \arg\min_a \left[\int_\Omega q^\star(v) \log\left(\frac{q^\star(v)}{q(v)}\right) dv\right] \\
&= \arg\min_a \int_\Omega q^\star(v) \log\left(\frac{q^\star(v)}{p(v)} \frac{p(v)}{q(v)}\right) dv \\
&= \arg\min_a \underbrace{\int_\Omega q^\star(v) \log\left(\frac{q^\star(v)}{p(v)}\right)}_{\text{does not depend on a}} dv - \int_\Omega q^\star(v) \log\left(\frac{q(v)}{p(v)}\right) dv
\end{aligned}$$

KL-Control

convert to maximization

$$
\begin{aligned}
a^\star &= \arg\min_a \left[\int_\Omega q^\star(v) \log\left(\frac{q^\star(v)}{q(v)}\right) dv \right] \\
&= \arg\min_a \underbrace{\int_\Omega q^\star(v) \log\left(\frac{q^\star(v)}{p(v)}\right)}_{\text{does not depend on a}} dv - \int_\Omega q^\star(v) \log\left(\frac{q(v)}{p(v)}\right) dv \\
&= \arg\max_a \int_\Omega q^\star(v) \log\left(\frac{q(v)}{p(v)}\right) dv
\end{aligned}
$$

KL-Control

integrate out probabilities

$$
\begin{aligned}
a^\star &= \arg\max_a \int_\Omega q^\star(v) \log\left(\frac{q(v)}{p(v)}\right) dv \\
&= \arg\max_a \int_\Omega q^\star(v) \left(\sum_{t=0}^{T-1} -\frac{1}{2} a_t^\top \Sigma^{-1} a_t + a_t^\top \Sigma^{-1} v_t\right) dv \\
&= \arg\max_a \sum_{t=0}^{T-1} \left(-\frac{1}{2} a_t^\top \Sigma^{-1} a_t + a_t^\top \int_\Omega q^\star(v) \Sigma^{-1} v_t dv\right)
\end{aligned}
$$

KL-Control

integrate out probabilities

$$
\begin{aligned}
a^\star &= \arg\max_a \int_\Omega q^\star(v) \log\left(\frac{q(v)}{p(v)}\right) dv \\
&= \arg\max_a \int_\Omega q^\star(v) \left(\sum_{t=0}^{T-1} -\frac{1}{2} a_t^\top \Sigma^{-1} a_t + a_t^\top \Sigma^{-1} v_t\right) dv \\
&= \arg\max_a \sum_{t=0}^{T-1} \left(-\frac{1}{2} a_t^\top \Sigma^{-1} a_t + a_t^\top \int_\Omega q^\star(v) \Sigma^{-1} v_t dv\right)
\end{aligned}
$$

objective is quadratic w.r.t. a_t !

KL-Control solution

take derivative of each a_t, set to zero and solve for $a^\star$

$$a^\star = \arg\max_a \sum_{t=0}^{T-1} \left(-\frac{1}{2} a_t^\top \Sigma^{-1} a_t + a_t^\top \int_\Omega q^\star(v) \Sigma^{-1} v_t dv \right)$$

optimal control solution is then

$$a_t^\star = \int_\Omega q^\star(v) v_t dv$$

Calculating control

how do you compute this?

$$
\begin{aligned}
a_t^\star &= \int_\Omega q^\star(v) v_t dv \\
&= \mathbb{E}_{\mathbb{Q}^\star}\left[v_t\right]
\end{aligned}
$$

can't generate v samples from $q^\star(v)$

Calculating control

how do you compute this?

$$\begin{aligned} a_t^\star &= \int_\Omega q^\star(v) v_t dv \\ &= \mathbb{E}_{\mathbb{Q}^\star}\left[v_t\right] \end{aligned}$$

can't generate v samples from $q^\star(v)$, but we can from $\mathbb{Q}$

Calculating control

use important sampling to switch distribution

$$
\begin{aligned}
a_t^\star &= \int_\Omega q^\star(v)\frac{q(v)}{q(v)} v_t dv \\
&= \mathbb{E}_{\mathbb{Q}}\left[\frac{q^\star(v)}{q(v)} v_t\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[\omega(v) v_t\right]
\end{aligned}
$$

can samples from $\mathbb{Q}$

Sample-based formulation

expectations can be written out as a sum over samples
since samples are generated from $\mathbb{Q}$, we make a change of variables $a_t + \epsilon_t = v_t$

$$a_t^\star = \mathbb{E}_\mathbb{Q}\left[\omega(v) v_t\right]$$

we obtain recursive solution

$$a_t^{i+1} = a_t^i + \sum_{n=1}^{N} \omega(\epsilon^n)\epsilon_t^n \qquad \text{s.t. } \omega(\epsilon^n) = \frac{q^\star(a^i + \epsilon^n)}{q(a^i + \epsilon^n)}$$

where $\epsilon^n = [\epsilon_0^n, \ldots, \epsilon_{T-1}^n]$
$\epsilon_t^n \sim \mathcal{N}(0, \Sigma)$

MPPI Algorithm: Prediction step

initialize ℓ, f, T, N, a^0, Σ, λ

```
i = 0
s̄_0 ← getStateEstimate()
for n ∈ [1, ..., N] do
    s_0 ← s̄_0
    sample ε^n = [ε^n_0, ..., ε^n_{T−1}]
    for t ∈ [0, ..., T − 1] do
        s_{t+1} = f(s_t, a^i_t + ε^n_t)
        J(ε^n) += ℓ(s_t) + λ a^i_t^⊤ Σ^{−1} ε^k_t
    end for
    J(ε^n) += m(s_T)
    using J(ε^n), calculate importance ratio
    ω(ε^n) = q*(a^i + ε^n) / q(a^i + ε^n)
end for
```

$$s_{t+1} = f(s_t, a_t^i + \epsilon_t^n)$$

$$\mathcal{J}(\epsilon^n) += \ell(s_t) + \lambda {a_t^i}^\top \Sigma^{-1} \epsilon_t^k$$

$$\mathcal{J}(\epsilon^n) += m(s_T)$$

$$\omega(\epsilon^n) = \frac{q^\star(a^i + \epsilon^n)}{q(a^i + \epsilon^n)}$$

MPPI Algorithm: Update step

actions are recursively updated with the control samples

1: **for** $a_t^i \in a^i$ **do**
2: $\quad a_t^{i+1} \leftarrow a_t^i + \sum_{n=1}^{N} \omega(\epsilon^n)\epsilon_t^n$
3: **end for**

apply a_0^{i+1} to robot shift action section

1: **for** $t \in [1, \ldots, T-1]$ **do**
2: $\quad a_t^{i-1} \leftarrow a_t^{i+1}$
3: **end for**
4: init a_{T-1}^{i+1}
5: $i+=1$

Some implementation details

- beneficial to parallelize
- can add arbitrary dynamic constraints (like saturation) to f
- control variance Σ can lead to shaky behavior
- undersampling can also lead to poor performance

Overview

- free-energy formulation
- constructing stochastic control problem
- sample-based approach

Citations I

I. Abraham, A. Broad, A. Pinosky, B. Argall, and T. D. Murphey.
Hybrid control for learning motor skills.
IEEE Robotics and Automation Letters, 5(2):2864–2871, 2020.

I. Abraham, A. Handa, N. Ratliff, K. Lowrey, T. D. Murphey, and D. Fox.
Model-based generalization under parameter uncertainty using path integral control.
IEEE Robotics and Automation Letters, 5(2):2864–2871, 2020.

E. Theodorou, J. Buchli, and S. Schaal.
A generalized path integral control approach to reinforcement learning.
journal of machine learning research, 11(Nov):3137–3181, 2010.

G. Williams, A. Aldrich, and E. A. Theodorou.
Model predictive path integral control: From theory to parallel computation.
Journal of Guidance, Control, and Dynamics, 40(2):344–357, 2017.

Citations II

G. Williams, P. Drews, B. Goldfain, J. M. Rehg, and E. A. Theodorou.
Aggressive driving with model predictive path integral control.
In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1433–1440. IEEE, 2016.