이동로봇 Tutorial 2차시

Model Predictive Path Integral

기계공학과
2025310186 박이새

2025. 11. 20

01

MPPI Algorithm

MPPI Algorithm

Pseudo Code

Algorithm 1: Sampling-Based MPC.

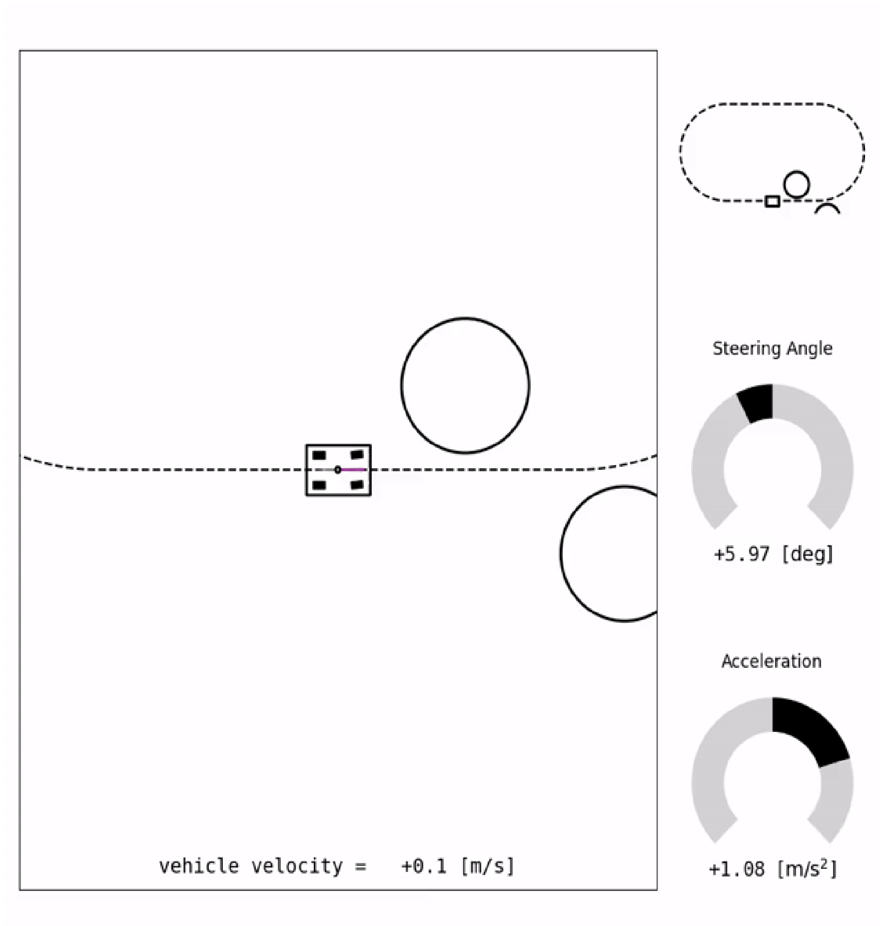
Given: $\mathbf{F}, g$: Transition Model;
K: Number of samples;
T: Number of time-steps;
$(\mathbf{u}_0, \mathbf{u}_1, \dots \mathbf{u}_{T-1})$: Initial control sequence;
$\Sigma, \phi, q, \gamma, \alpha$: Cost functions/parameters;
SGF: Savitsky–Galoy convolutional filter;
while *task not completed* **do**
  $\mathbf{x}_0 \leftarrow$ GetStateEstimate();
  for $k \leftarrow 0$ **to** $K-1$ **do**
    $\mathbf{x} \leftarrow \mathbf{x}_0$;
    Sample $\mathcal{E}^k = \left(\epsilon_0^k \dots \epsilon_{T-1}^k\right)$, $\epsilon_t^k \in \mathcal{N}(0, \Sigma)$;
    for $t \leftarrow 1$ **to** T **do**
      if $k < (1-\alpha)K$ **then**
        $\mathbf{v}_{t-1} = \mathbf{u}_{t-1} + \epsilon_{t-1}^k$;
      else
        $\mathbf{v}_{t-1} = \epsilon_{t-1}^k$;
      $\mathbf{x} \leftarrow \mathbf{F}(\mathbf{x}, g(\mathbf{v}_{t-1}))$;
      $S_k \mathrel{+}= \mathbf{q}(\mathbf{x}) + \gamma \mathbf{u}_{t-1}^{\mathrm{T}} \Sigma^{-1}(\mathbf{u}_t - \mathbf{v}_t)$;
    $S_k \mathrel{+}= \phi(\mathbf{x})$;
  ComputeWeights$(S_0, S_1, \dots S_k)$;
  for $t \leftarrow 0$ **to** $T-1$ **do**
    $U \leftarrow U + \text{SGF} * \left(\sum_{k=1}^{K} w_k \mathcal{E}^k\right)$;
  SendToActuators$(\mathbf{u}_0)$;
  for $t \leftarrow 1$ **to** $T-1$ **do**
    $\mathbf{u}_{t-1} \leftarrow \mathbf{u}_t$;
  $\mathbf{u}_{T-1} \leftarrow$ Intialize$(\mathbf{u}_{T-1})$;

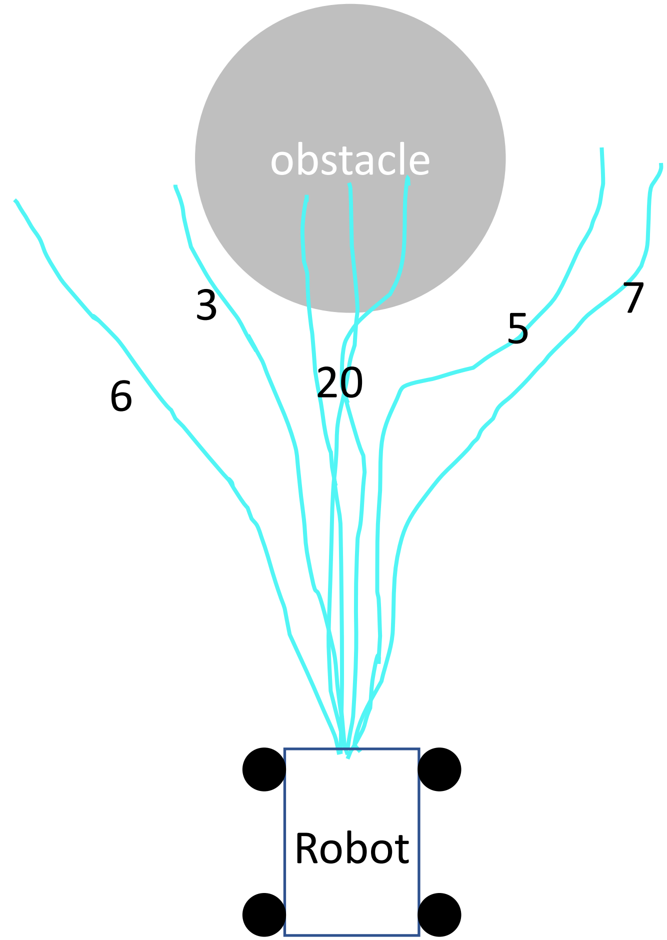
Algorithm 2: Information-Theoretic Weight Computation.

Given: $S_1, S_2, \dots S_K$: Trajectory costs;
λ: Inverse Temperature;
$\rho \leftarrow \min_k [S_k]$;
$\tilde{\eta} \leftarrow \sum_{k=1}^{K} \exp\left(-\frac{1}{\lambda}(S_k - \rho)\right)$;
for $k \leftarrow 1$ **to** K **do**
  $w_k \leftarrow \frac{1}{\tilde{\eta}} \exp\left(-\frac{1}{\lambda}(S_k - \rho)\right)$;
return $\{w_1, w_2, \dots w_K\}$

MPPI Algorithm

Example





Ref: https://github.com/MizuhoAOKI/python_simple_mppi

MPPI Algorithm

Pros and Cons

Pros

- Flexible: handles nonlinear dynamics and nonconvex cost functions
- Parallelizable: efficient on GPU/parallel hardware, suitable for real-time applications
- Derivative-free: does not require gradients of dynamics or cost functions
- Exploratory: inherent sampling encourages exploration in uncertain environments

Cons

- Sample-sensitive: performance depends heavily on the quality of sampled trajectories
- Computationally expensive: requires a large number of rollouts for good performance
- Hyperparameter-dependent: sensitive to noise variance, horizon length, temperature, etc.

02

Information-Theoretic Derivation of MPPI

Information-Theoretic Derivation of MPPI

Stochastic Control Setup

- Discrete-time dynamics $\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{v}_t)$

- Noise control input $\mathbf{v}_t \sim \mathcal{N}(\mathbf{u}_t, \boldsymbol{\Sigma})$

- Control sequence and noise sequence

$$\mathbf{U} = (\mathbf{u}_0, \ldots, \mathbf{u}_{T-1}), \quad \mathbf{V} = (\mathbf{v}_0, \ldots, \mathbf{v}_{T-1})$$

- Distribution over control noises

$$q(\mathbf{V} \mid \mathbf{U}, \boldsymbol{\Sigma}) = Z^{-T} \exp\Big(-\tfrac{1}{2} \textstyle\sum_{t=0}^{T-1} (\mathbf{v}_t - \mathbf{u}_t)^\top \boldsymbol{\Sigma}^{-1} (\mathbf{v}_t - \mathbf{u}_t) \Big)$$

- Cost

$$J(\mathbf{U}) = \mathbb{E}_{Q_{\mathbf{U},\boldsymbol{\Sigma}}}\Big[\phi(\mathbf{x}_T) + \textstyle\sum_{t=0}^{T-1} \ell(\mathbf{x}_t, \mathbf{u}_t) \Big]$$

Information-Theoretic Derivation of MPPI

Trajectory Cost and Free Energy

- Define state-only trajectory cost

$$S(\mathbf{V}) = \phi(\mathbf{x}_T) + \sum_{t=0}^{T-1} c(\mathbf{x}_t)$$

- Running cost decomposition

$$\ell(\mathbf{x}_t, \mathbf{u}_t) = c(\mathbf{x}_t) + \text{control cost term}$$

- Let P be a base distribution over $\mathbf{V}$ with density $p(\mathbf{V})$

- Define free energy

$$\mathcal{F}_\lambda(S, p, \mathbf{x}_0) = -\lambda \log \mathbb{E}_P\Big[\exp\big(-\tfrac{1}{\lambda} S(\mathbf{V})\big)\Big]$$

Information–Theoretic Derivation of MPPI

Free Energy Lower bound via KL

- Start from the free energy $\mathcal{F}_\lambda = -\lambda \log \mathbb{E}_P\left[e^{-\frac{1}{\lambda}S(\mathbf{V})}\right]$

- Change measure from P to the controlled distribution $Q_{\mathbf{U},\boldsymbol{\Sigma}}$

$$\mathcal{F}_\lambda = -\lambda \log \mathbb{E}_{Q_{\mathbf{U},\boldsymbol{\Sigma}}}\left[e^{-\frac{1}{\lambda}S(\mathbf{V})}\frac{p(\mathbf{V})}{q(\mathbf{V}|\mathbf{U},\boldsymbol{\Sigma})}\right]$$

- Apply Jensen’s inequality

$$\mathcal{F}_\lambda \leq -\lambda \mathbb{E}_{Q_{\mathbf{U},\boldsymbol{\Sigma}}}\left[\log\left(e^{-\frac{1}{\lambda}S(\mathbf{V})}\frac{p(\mathbf{V})}{q(\mathbf{V}|\mathbf{U},\boldsymbol{\Sigma})}\right)\right]$$

- Simplify

$$\mathcal{F}_\lambda \leq \mathbb{E}_{Q_{\mathbf{U},\boldsymbol{\Sigma}}}[S(\mathbf{V})] + \lambda D_{\mathrm{KL}}(Q_{\mathbf{U},\boldsymbol{\Sigma}} \parallel P)$$

$$D_{\mathrm{KL}}(Q_{\mathbf{U},\boldsymbol{\Sigma}} \parallel P) = \mathbb{E}_{Q_{\mathbf{U},\boldsymbol{\Sigma}}}\left[\log\frac{q(\mathbf{V}|\mathbf{U},\boldsymbol{\Sigma})}{p(\mathbf{V})}\right]$$

Information–Theoretic Derivation of MPPI

Free Energy Lower bound via KL

- Choose the base distribution as

$$p(\mathbf{V}) = q(\mathbf{V} \mid \tilde{\mathbf{U}}, \boldsymbol{\Sigma})$$

  where $\tilde{\mathbf{U}}$ is a nominal control sequence.
- Then the KL term

$$D_{\mathrm{KL}}\left(Q_{\mathbf{U},\boldsymbol{\Sigma}} \parallel Q_{\tilde{\mathbf{U}},\boldsymbol{\Sigma}}\right) = \frac{1}{2}\sum_{t=0}^{T-1}(\mathbf{u}_t - \tilde{\mathbf{u}}_t)^\top \boldsymbol{\Sigma}^{-1}(\mathbf{u}_t - \tilde{\mathbf{u}}_t) + c$$

- Therefore, the right–hand side of the bound becomes

$$\mathcal{F}_\lambda \leq \mathbb{E}_{Q_{\mathbf{U},\boldsymbol{\Sigma}}}\left[\phi(\mathbf{x}_T) + \sum_{t=0}^{T-1}\ell(\mathbf{x}_t, \mathbf{u}_t)\right] = J(\mathbf{U})$$

Information-Theoretic Derivation of MPPI

Optimal Distribution over Trajectories

- Define the optimal distribution

$$q^*(\mathbf{V}) = \frac{1}{\eta}\exp\Big(-\frac{1}{\lambda}S(\mathbf{V})\Big)p(\mathbf{V})$$

- Normalization constant

$$\eta = \int \exp\Big(-\frac{1}{\lambda}S(\mathbf{V})\Big)p(\mathbf{V})\,d\mathbf{V}$$

- This is a Boltzmann distribution over trajectories:
  lower-cost trajectories get higher probability

Information-Theoretic Derivation of MPPI

Control as KL Minimization

- Instead of minimizing $J(\mathbf{U})$ directly, we solve

$$\mathbf{U}^* = \arg\min_{\mathbf{U}} \, D_{\mathrm{KL}}\left(Q^* \parallel Q_{\mathbf{U},\mathbf{\Sigma}}\right)$$

- Expand the KL

$$D_{\mathrm{KL}}(Q^* \| Q_{\mathbf{U},\mathbf{\Sigma}}) = \mathbb{E}_{Q^*}\left[\log \frac{q^*(\mathbf{V})}{q(\mathbf{V}|\mathbf{U},\mathbf{\Sigma})}\right]$$

- Since q^* does not depend on

$$\mathbf{U}^* = \arg\max_{\mathbf{U}} \, \mathbb{E}_{Q^*}\left[\log q(\mathbf{V} \mid \mathbf{U}, \mathbf{\Sigma})\right]$$

Information-Theoretic Derivation of MPPI

Control as KL Minimization

- For Gaussian q , this is equivalent to

$$\mathbf{U}^* = \arg\min_{\mathbf{U}} \mathbb{E}_{Q^*}\left[\sum_{t=0}^{T-1} (\mathbf{v}_t - \mathbf{u}_t)^\top \mathbf{\Sigma}^{-1} (\mathbf{v}_t - \mathbf{u}_t) \right]$$

- Unconstrained minimizer:

$$\mathbf{u}_t^* = \mathbb{E}_{Q^*}[\mathbf{v}_t]$$

Information-Theoretic Derivation of MPPI

Importance Sampling

- We cannot sample directly from Q^*
  Use a proposal distribution $Q_{\hat{\mathbf{U}},\boldsymbol{\Sigma}}$

- We write this expectation as an integral

$$\mathbb{E}_{Q^*}[\mathbf{v}_t] = \int q^*(\mathbf{V})\,\mathbf{v}_t\,d\mathbf{V}$$

- Multiply and divide by the proposal density $q(\mathbf{V} \mid \hat{\mathbf{U}}, \boldsymbol{\Sigma})$

$$\mathbb{E}_{Q^*}[\mathbf{v}_t] = \int \frac{q^*(\mathbf{V})}{q(\mathbf{V}|\hat{\mathbf{U}},\boldsymbol{\Sigma})}\,q(\mathbf{V} \mid \hat{\mathbf{U}}, \boldsymbol{\Sigma})\,\mathbf{v}_t\,d\mathbf{V}$$

- Therefore,

$$\mathbb{E}_{Q^*}[\mathbf{v}_t] = \mathbb{E}_{Q_{\hat{\mathbf{U}},\boldsymbol{\Sigma}}}\left[\,w(\mathbf{V})\,\mathbf{v}_t\,\right],\quad w(\mathbf{V}) = \frac{q^*(\mathbf{V})}{q(\mathbf{V}|\hat{\mathbf{U}},\boldsymbol{\Sigma})}$$

Information-Theoretic Derivation of MPPI

Decomposing the Importance Weight

- Introduce the base distribution P with density $p(\mathbf{V})$
- We can factor $w(\mathbf{V})$ as

$$w(\mathbf{V}) = \left(\frac{q^*(\mathbf{V})}{p(\mathbf{V})}\right)\left(\frac{p(\mathbf{V})}{q(\mathbf{V}|\hat{\mathbf{U}},\mathbf{\Sigma})}\right)$$

- Using the optimal distribution

$$q^*(\mathbf{V}) = \frac{1}{\eta}\exp\Big(-\frac{1}{\lambda}S(\mathbf{V})\Big)p(\mathbf{V})$$

- We obtain

$$w(\mathbf{V}) = \frac{1}{\eta}\exp\Big(-\frac{1}{\lambda}S(\mathbf{V})\Big)\left(\frac{p(\mathbf{V})}{q(\mathbf{V}|\hat{\mathbf{U}},\mathbf{\Sigma})}\right)$$

Information-Theoretic Derivation of MPPI

Gaussian Ratio

- For the Gaussian base $p(\mathbf{V}) = q(\mathbf{V} \mid \tilde{\mathbf{U}}, \mathbf{\Sigma})$,
- We can write

$$\frac{p(\mathbf{V})}{q(\mathbf{V}|\hat{\mathbf{U}},\mathbf{\Sigma})} = \exp\left(-\frac{1}{2}\sum_{t=0}^{T-1} \left[D + 2(\hat{\mathbf{u}}_t - \tilde{\mathbf{u}}_t)^\top \mathbf{\Sigma}^{-1}\mathbf{v}_t\right]\right)$$

where

$$D = \hat{\mathbf{u}}_t^\top \mathbf{\Sigma}^{-1}\hat{\mathbf{u}}_t - \tilde{\mathbf{u}}_t^\top \mathbf{\Sigma}^{-1}\tilde{\mathbf{u}}_t$$

- The term D does not depend on $\mathbf{v}_t$, so it can be pulled out of the integral and canceled with the corresponding term in η.
- Thus we keep only the $\mathbf{v}_t$ – dependent part:

$$\frac{p(\mathbf{V})}{q(\mathbf{V}|\hat{\mathbf{U}},\mathbf{\Sigma})} \propto \exp\left(-\sum_{t=0}^{T-1} (\hat{\mathbf{u}}_t - \tilde{\mathbf{u}}_t)^\top \mathbf{\Sigma}^{-1}\mathbf{v}_t\right)$$

Information–Theoretic Derivation of MPPI

Final Weight and Control Law

- Substituting the Gaussian ratio into $w(\mathbf{V})$

$$w(\mathbf{V}) = \frac{1}{\eta} \exp\left(-\frac{1}{\lambda}\left[S(\mathbf{V}) + \lambda \sum_{t=0}^{T-1} (\hat{\mathbf{u}}_t - \tilde{\mathbf{u}}_t)^\top \mathbf{\Sigma}^{-1} \mathbf{v}_t \right] \right)$$

- The corresponding control update is

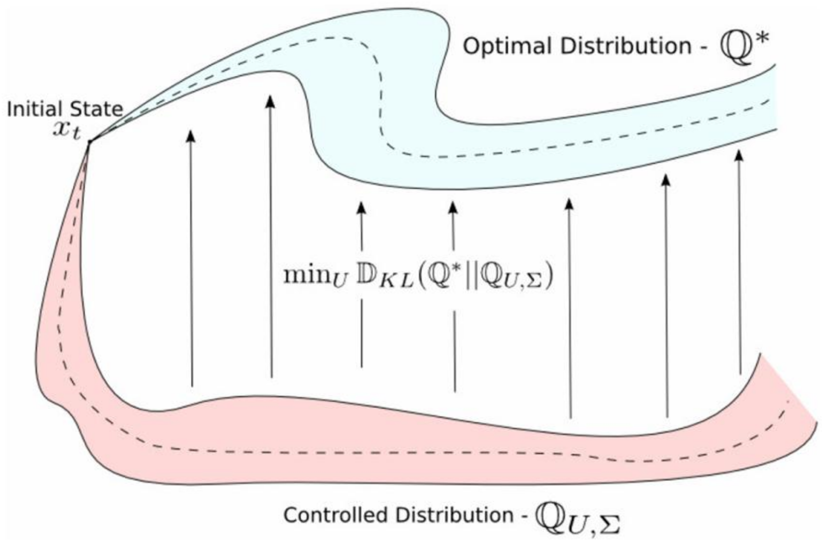
$$\mathbf{u}_t^* = \mathbb{E}_{Q_{\hat{\mathbf{U}},\mathbf{\Sigma}}}\left[w(\mathbf{V})\, \mathbf{v}_t \right]$$

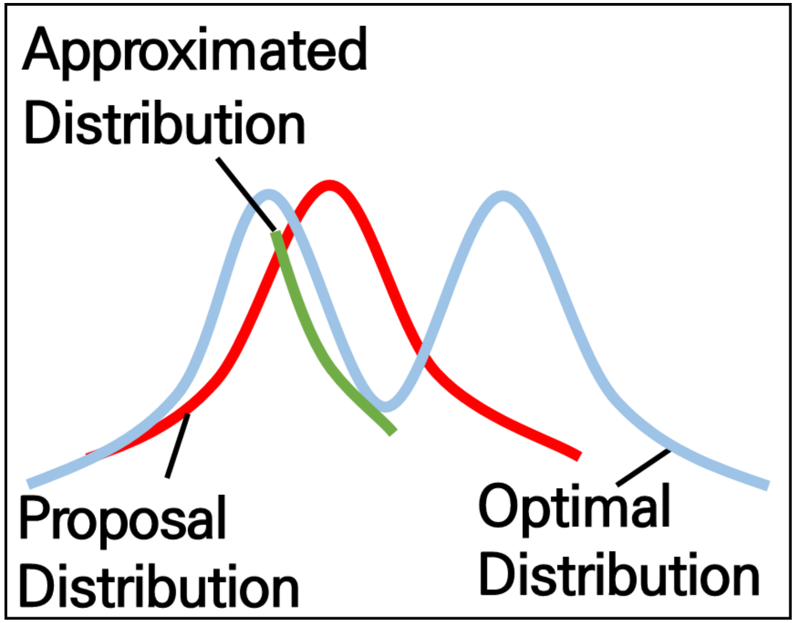
- In practice, this expectation is approximated by Monte Carlo

$$\mathbf{u}_t^* \approx \frac{\sum_{k=1}^{K} w^{(k)}\, \mathbf{v}_t^{(k)}}{\sum_{k=1}^{K} w^{(k)}} \qquad w^{(k)} \propto \exp\left(-\frac{1}{\lambda} \tilde{S}(\mathbf{V}^{(k)}) \right)$$

Information-Theoretic Derivation of MPPI

Example





03

Classical Path-Integral Derivation of MPPI

Classical Path-Integral Derivation of MPPI

Stochastic System and Optimal Control Cost

- Continuous-time stochastic dynamics

$$d\mathbf{x}_t = \big(\mathbf{f}(\mathbf{x}_t, t) + \mathbf{G}(\mathbf{x}_t, t)\,\mathbf{u}(\mathbf{x}_t, t)\big)\,dt + \mathbf{B}(\mathbf{x}_t, t)\,d\mathbf{w}_t$$

- State partition

$$\mathbf{x}_t = \begin{bmatrix} \mathbf{x}_t^{(a)} \\ \mathbf{x}_t^{(c)} \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} \mathbf{0} \\ \mathbf{G}_c \end{bmatrix}, \ \mathbf{B} = \begin{bmatrix} \mathbf{0} \\ \mathbf{B}_c \end{bmatrix}$$

- Cost functional

$$J(\mathbf{x}_t, t; \mathbf{u}) = \mathbb{E}_Q\left[\phi(\mathbf{x}_T) + \int_t^T \left(q(\mathbf{x}_s, s) + \tfrac{1}{2}\mathbf{u}_s^\top \mathbf{R}\mathbf{u}_s\right) ds\right]$$

- Value function

$$V(\mathbf{x}_t, t) = \min_{\mathbf{u}} J(\mathbf{x}_t, t; \mathbf{u})$$

Classical Path-Integral Derivation of MPPI

Stochastic HJB Equation

- Stochastic HJB

$$-\partial_t V = q + \mathbf{f}^\top V_{\mathbf{x}} - \frac{1}{2} V_{\mathbf{x}}^\top \mathbf{G}\mathbf{R}^{-1}\mathbf{G}^\top V_{\mathbf{x}} + \frac{1}{2}\mathrm{tr}\big(\mathbf{B}\mathbf{B}^\top V_{\mathbf{x}\mathbf{x}}\big)$$

- Boundary

$$V(\mathbf{x}_T, T) = \phi(\mathbf{x}_T)$$

- Optimal Control

$$\mathbf{u}^*(\mathbf{x}, t) = -\mathbf{R}^{-1}\mathbf{G}^\top V_{\mathbf{x}}(\mathbf{x}, t)$$

Classical Path-Integral Derivation of MPPI

Exponential Transform and Linearization of the HJB

- Define desirability transform

$$V(\mathbf{x},t) = -\lambda \log \Psi(\mathbf{x},t)$$

- Compute derivatives

$$V_{\mathbf{x}} = -\lambda \frac{\Psi_{\mathbf{x}}}{\Psi} \qquad V_{\mathbf{xx}} = -\lambda \frac{\Psi_{\mathbf{xx}}}{\Psi} + \lambda \frac{\Psi_{\mathbf{x}}\Psi_{\mathbf{x}}^\top}{\Psi^2}$$

- Substitute into the nonlinear HJB

$$\frac{\lambda}{\Psi}\partial_t\Psi = q - \lambda\frac{\mathbf{f}^\top\Psi_{\mathbf{x}}}{\Psi} + \frac{\lambda^2}{2\Psi^2}\Psi_{\mathbf{x}}^\top\mathbf{G}\mathbf{R}^{-1}\mathbf{G}^\top\Psi_{\mathbf{x}} - \frac{\lambda}{2\Psi}\mathrm{tr}(\mathbf{B}\mathbf{B}^\top\Psi_{\mathbf{xx}}) + \frac{\lambda}{2\Psi^2}\Psi_{\mathbf{x}}^\top\mathbf{B}\mathbf{B}^\top\Psi_{\mathbf{x}}$$

- Identify nonlinear gradient-quadratic terms

$$\frac{\lambda^2}{2\Psi^2}\Psi_{\mathbf{x}}^\top\mathbf{G}\mathbf{R}^{-1}\mathbf{G}^\top\Psi_{\mathbf{x}} \qquad \frac{\lambda}{2\Psi^2}\Psi_{\mathbf{x}}^\top\mathbf{B}\mathbf{B}^\top\Psi_{\mathbf{x}}$$

Classical Path-Integral Derivation of MPPI

Exponential Transform and Linearization of the HJB

- Noise-Control matching cancels these terms
  If

$$\mathbf{B}\mathbf{B}^\top = \lambda \mathbf{G}\mathbf{R}^{-1}\mathbf{G}^\top$$

then

$$\frac{\lambda^2}{2\Psi^2}\Psi_{\mathbf{x}}^\top \mathbf{G}\mathbf{R}^{-1}\mathbf{G}^\top \Psi_{\mathbf{x}} = \frac{\lambda}{2\Psi^2}\Psi_{\mathbf{x}}^\top \mathbf{B}\mathbf{B}^\top \Psi_{\mathbf{x}}$$

-> meaning the two nonlinear terms exactly cancel

- Final linear PDE

$$\partial_t \Psi = \Psi \, \frac{q}{\lambda} - \mathbf{f}^\top \Psi_{\mathbf{x}} - \frac{1}{2}\mathrm{tr}(\mathbf{B}\mathbf{B}^\top \Psi_{\mathbf{xx}})$$

Classical Path-Integral Derivation of MPPI

Feynman-Kac and Path Integral Representation

- Linear PDE solution via Feynman-Kac

$$\Psi(\mathbf{x}_{t_0}, t_0) = \mathbb{E}_P\left[e^{-\frac{1}{\lambda}\int q(\mathbf{x}_s, s)\, ds}\Psi(\mathbf{x}_T, T)\right]$$

- Terminal condition

$$\Psi(\mathbf{x}_T, T) = e^{-\phi(\mathbf{x}_T)/\lambda}$$

- Trajectory cost

$$S(\boldsymbol{\tau}) = \phi(\mathbf{x}_T) + \int q(\mathbf{x}_s, s)\, ds$$

- Path integral form

$$\Psi(\mathbf{x}_{t_0}, t_0) = \mathbb{E}_P[e^{-S(\boldsymbol{\tau})/\lambda}]$$

- Value function

$$V(\mathbf{x}, t) = -\lambda \log \mathbb{E}_P[e^{-S(\boldsymbol{\tau})/\lambda}]$$

Classical Path-Integral Derivation of MPPI

Continuous-Time Path Integral Optimal Control

- Path integral control law (continuous time)

$$\mathbf{u}^*(t_0)dt = \mathbf{g}(\mathbf{x}_{t_0}, t_0)\,\frac{\mathbb{E}_P\left[e^{-S(\boldsymbol{\tau})/\lambda}\mathbf{B}_c(\mathbf{x}_{t_0}, t_0)\,d\mathbf{w}\right]}{\mathbb{E}_P\left[e^{-S(\boldsymbol{\tau})/\lambda}\right]}$$

- Here, $\mathrm{g}(\mathrm{x}, t)$ is defined as

$$\mathbf{g}(\mathbf{x}, t) = \mathbf{R}(\mathbf{x}, t)^{-1}\mathbf{G}_c(\mathbf{x}, t)^\top\left(\mathbf{G}_c(\mathbf{x}, t)\mathbf{R}(\mathbf{x}, t)^{-1}\mathbf{G}_c(\mathbf{x}, t)^\top\right)^{-1}$$

- And B_c is the diffusion matrix acting on the controlled states

Classical Path-Integral Derivation of MPPI

Discrete-Time Path Integral Control

- Stochastic dynamics for uncontrolled system

$$d\mathbf{x}_t = \mathbf{f}(\mathbf{x}_t)\Delta t + \mathbf{B}(\mathbf{x}_t)\,\boldsymbol{\epsilon}_t\sqrt{\Delta t}$$

- Therefore

$$\mathbf{B}_c(\mathbf{x}_t)\,d\mathbf{w}_t = d\mathbf{x}_t^{(c)} - \mathbf{f}_c(\mathbf{x}_t)\,dt$$

- For controlled state component

$$d\mathbf{x}_t^{(c)} = \mathbf{f}_c(\mathbf{x}_t)\,dt + \mathbf{G}_c(\mathbf{x}_t)\mathbf{u}_t\,dt$$

- Discrete PI control law

$$\mathbf{u}^*(\mathbf{x}_{t_0}) = \mathbf{g}(\mathbf{x}_{t_0})\frac{\mathbb{E}_p[e^{-S/\lambda}(d\mathbf{x}_{t_0}^{(c)}/\Delta t - \mathbf{f}_c(\mathbf{x}_{t_0}))]}{\mathbb{E}_p[e^{-S/\lambda}]}$$

Classical Path-Integral Derivation of MPPI

Importance Sampling

- Goal: Replace sampling from inefficient distribution with a more effective distribution
- Start from discrete PI control

$$\mathbf{u}_t^* = \mathbf{g}(\mathbf{x}_t)\frac{\mathbb{E}_p\left[e^{-S(\tau)/\lambda}\left(\frac{d\mathbf{x}_t^{(c)}}{\Delta t} - \mathbf{f}_c\right)\right]}{\mathbb{E}_p[e^{-S(\tau)/\lambda}]}$$

- Write numerator and denominator as integrals over trajectories

$$\int e^{-S(\tau)/\lambda}\left(\frac{d\mathbf{x}_t^{(c)}}{\Delta t} - \mathbf{f}_c\right)p(\tau)\,d\tau, \quad \int e^{-S(\tau)/\lambda}p(\tau)\,d\tau$$

- Multiply numerator & denominator by $\frac{p(\tau)}{q(\tau)}$

$$\int e^{-S(\tau)/\lambda}\left(\frac{d\mathbf{x}_t^{(c)}}{\Delta t} - \mathbf{f}_c(\mathbf{x}_t)\right)\frac{p(\tau)}{q(\tau)}\,q(\tau)\,d\tau, \quad \int e^{-S(\tau)/\lambda}\frac{p(\tau)}{q(\tau)}\,q(\tau)\,d\tau$$

Classical Path-Integral Derivation of MPPI

Importance Sampling

- Convert integrals into expectations under q

$$\mathbf{u}_t^* = \mathbf{G}(\mathbf{x}_t)\frac{\mathbb{E}_q\left[e^{-S(\tau)/\lambda}\left(\frac{d\mathbf{x}_t^{(c)}}{\Delta t} - \mathbf{f}_c(\mathbf{x}_t)\right)\frac{p(\tau)}{q(\tau)}\right]}{\mathbb{E}_q\left[e^{-S(\tau)/\lambda}\frac{p(\tau)}{q(\tau)}\right]}$$

- Introduce likelihood ratio

$$w(\tau) = \frac{p(\tau)}{q(\tau)}$$

- Combine likelihood ratio with cost to form modified cost

$$\tilde{S} = S + \lambda \sum_t \log(p_t/q_t)$$

Classical Path–Integral Derivation of MPPI

MPPI Control Update

- Sample controls

$$\mathbf{u}_{t,k} = \mathbf{u}_t + \delta\mathbf{u}_{t,k}, \qquad \delta\mathbf{u}_{t,k} \sim \mathcal{N}(\mathbf{0}, \Sigma_u)$$

- For each rollout compute modified trajectory cost

$$\tilde{S}_k = \phi(\mathbf{x}_T) + \sum_{i=0}^{T-1} \tilde{q}_{i,k}$$

- Probability weights

$$w_k = \exp\left(-\frac{\tilde{S}_k}{\lambda}\right)$$

- Final MPPI control update

$$\mathbf{u}^*(t) = \mathbf{u}(t) + \frac{\sum_{k=1}^{K} w_k \, \delta\mathbf{u}_{t,k}}{\sum_{k=1}^{K} w_k}$$

Thank you!