Extra exercises

1. Solve the mass on the spring problem discussed in the tutorial on slides 32–35, 40 such that the end velocity is maximal, i.e. $\phi(x) = -x_2$.

2. Consider the control problem:

$$dx = udt + d\xi$$
$$C = \left\langle \frac{1}{2}x(T)^2 + \int_{t_0}^{T} dt \frac{1}{2}u(t)^2 \right\rangle$$

   with initial condition $x(t_0)$ and $\left\langle d\xi^2 \right\rangle = \nu dt$.

   (a) Solve the control problem in the deterministic case $\nu = 0$ using the PMP formalism.

   (b) Solve the control problem in the stochastic case using the Bellman equation.

   (c) Solve the control problem in the stochastic case using the path integral control methods and the Fokker Planck equation.

3. Consider the controlled random walk in one dimension

$$dx = udt + d\xi$$

   with initial position x at $t = 0$ and noise variance $\left\langle d\xi^2 \right\rangle = \nu dt$. When $u = 0$, the solution for the probability density at time t satisfies the Fokker-Planck equation

$$\partial_t \rho(y, t|x, 0) = \frac{1}{2}\nu \partial_y^2 \rho(y, t|x, 0)$$

   (a) Give an expression for the solution $\rho(y, t|x, o)$. Show that the solution satisfies the Fokker Planck equation.

   (b) Assume that there are two targets at $t = T$ at locations $x = \pm 1$. To make sure we arrive at the targets we define an end-cost function ϕ that has delta peaks at the targets:

$$\phi(x(T)) = \begin{cases} 0 & x(T) = \pm 1 \\ \infty & \text{otherwise} \end{cases}$$

   Compute the optimal cost-to-go for any x, t using the path integral formalism. Use $R = 1$ for the scaling of the control cost.

   (c) Derive that the optimal control satisfies

$$u^*(t, x) = \frac{\tanh(x/(\nu(T - t)) - x}{T - t}$$

4. Write a Matlab program for the control problem in exercise 3.

   (a) By varying ν, T, study numerically how the optimal control depends on these parameters.

   (b) Explain in words the delayed choice mechanism.

5. Consider the mountain car problem. A car is at the bottom of a valley and can accelerate forward or backwards. The problem is to find a control strategy that gets the car out of the valley. $L(x)$ is the shape of the valley. A definition of L can be found just below the exercise. Gravitational force is $F_g = -g \sin\alpha$, with $\tan\alpha = L' = \frac{dL}{dx}$ the slope. Since $-\pi/2 < \alpha < \pi/2$ we have $\cos\alpha = \sqrt{1 - \sin^2\alpha}$ and

$$F_g(x) = -g \frac{L'(x)}{\sqrt{1 + L'^2(x)}}$$

The second order system is described by position x and velocity v. The dynamics is

$$\begin{aligned} dx &= v dt \\ dv &= F_g(x) dt + u dt + d\xi \end{aligned}$$

The cost is

$$C(u) = \left\langle \phi(x(T)) + \int_0^T dt \frac{R}{2} u^2 \right\rangle$$

with $\phi(x) = 0$ for $x_{\min} < x < x_{\max}$ and $\phi(x) = A$ otherwise. Take $A = -1$ and $-x_{\min} = x_{\max} = 2$, $R = \nu = T = g = 1$. Feel free to change the values.

(a) Take $x(0) = 0.5$ and $v(0) = 0$. Simulate the uncontrolled dynamics and vary the parameters such that 1) the problem is too easy and all trajectories reach the top of the hill and 2) the problem is not too hard that no trajectories reach the top of the hill.

(b) With the parameter values found above, compute the optimal cost to go $J(x, v, t = 0)$ for $x = -2 : 0.1 : 2$ and $v = -2 : 0.1 : 2$ using MCMC. This is done by running n times the uncontrolled dynamics for each x, v pair:

$$J(x, v, t = 0) = -\lambda \log\left(\frac{1}{n} \sum_{\mu=1}^{n} \exp(-\phi(x(T))/\lambda)\right)$$

Plot $J(x, v, t = 0)$. Interpret the result.

(c) Design an optimal controller using the formula $u(x, t) = \frac{1}{dt} \frac{\mathbb{E} dW_t e^{-S}}{\mathbb{E} e^{-S}}$. x is the current state (position and velocity), t is the current time. dt is the time step. In theory, this should be infinitely small, but in practice better results are obtained with a quite large dt (for instance 0.1). The expressions are estimated with n trajectories, all starting at x, t as

$$\begin{aligned} \mathbb{E} e^{-S} &= \frac{1}{n} \sum_{\mu=1}^{n} \exp\left(-dt \sum_{i=N_1}^{N} V(x_i^\mu)\right) \\ \mathbb{E} W_t e^{-S} &= \frac{1}{n} \sum_{\mu=1}^{n} dW_t^\mu \exp\left(-dt \sum_{i=N_1}^{N} V(x_i^\mu)\right) \end{aligned}$$

with $t = N_1 dt$ and $T = N dt$. Note, that you can estimate both numerator and denominator from the same batch of samples.

```
function shape = L(x)
    shape = -1 - 1/2*(tanh(2*x + 2) - tanh(2*x - 2));
```

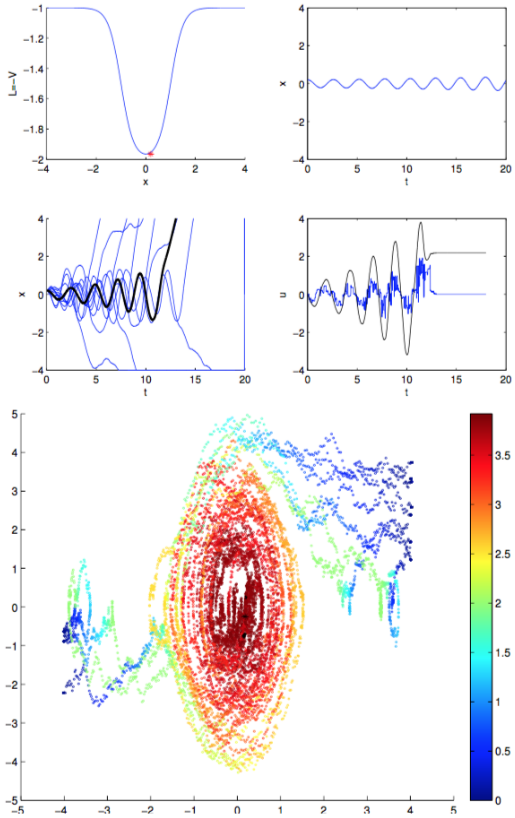

Figure 66: Mountain car problem. Top Left: Immediate cost $V(x) = -L(x)$. Initial position $x = 0.5, v = 0$. Top Right: Uncontrolled trajectories over time in the absence of noise do not reach the top of the hill. Bottom left: Uncontrolled trajectories over time with noise reach the top of the hill after much time (blue). Black is controlled deterministic trajectory. Bottom right: control over time u. For comparison also the velocity of the controlled trajectory is shown. Force u is synchronized with velocity. Bottom bottom: cost to go landscape at $t = 0$ versus position (horizontal) and velocity (vertical). Cost parameters $a = 0, b = 1, c = 0$. Number of sample paths $n = 10$. $R = 1, Ttotal = 20, T = 2, \nu = 1, G = 10, dt = 0.01$.