Machine Learning HW8

Student ID: 111652015

Problem 1

Definition (Notes on 10/17)

$$L_{SSM}(\theta) = \ \mathbb{E}_{x\sim p(x)}\big[\|S(x;\theta)\|^2\big] \ + \ \mathbb{E}_{x\sim p(x)}\mathbb{E}_{v\sim p(v)}\Big[\,2\,v^\top\nabla_x\big(v^\top S(x;\theta)\big)\Big]\,.$$

Goal Show that

$$L_{SSM}(\theta) = \mathbb{E}_{x\sim p(x)}\mathbb{E}_{v\sim p(v)}\,\Big[\|v^\top S(x;\theta)\|^2 + 2\,v^\top\nabla_x\big(v^\top S(x;\theta)\big)\Big]\,.$$

Lemma Assume $\mathbb{E}_v[vv^\top] = I$. Then for any vector a,

$$\mathbb{E}_v\big[(v^\top a)^2\big] = a^\top\mathbb{E}_v[vv^\top]\,a = \|a\|^2.$$

Proof Apply the identity with $a = S(x;\theta)$:

$$\mathbb{E}_v\big[\|v^\top S(x;\theta)\|^2\big] = \|S(x;\theta)\|^2.$$

Taking $\mathbb{E}_x$ on both sides and substituting into the given definition yields

$$\begin{aligned} L_{SSM}(\theta) &= \mathbb{E}_x\mathbb{E}_v\big[\|v^\top S(x;\theta)\|^2\big] + \mathbb{E}_x\mathbb{E}_v\big[2\,v^\top\nabla_x(v^\top S(x;\theta))\big] \\ &= \mathbb{E}_x\mathbb{E}_v\Big[\|v^\top S(x;\theta)\|^2 + 2\,v^\top\nabla_x(v^\top S(x;\theta))\Big]\,. \end{aligned}$$

Proof of Lemma Note that

$$(v^\top a)^2 = (a^\top v)(v^\top a) = a^\top(vv^\top)a.$$

Taking expectations and using linearity,

$$\mathbb{E}\big[(v^\top a)^2\big] = \mathbb{E}\Big[a^\top(vv^\top)a\Big] = a^\top\,\mathbb{E}[vv^\top]\,a = a^\top I a = \|a\|^2.$$

Problem 2 (Briefly explain SDE)

Stochastic Differential Equation (SDE) A noisy dynamical system can be modeled by the SDE

$$\mathrm{d}x_t \ = \ f(x_t,t)\,\mathrm{d}t \ + \ G(x_t,t)\,\mathrm{d}W_t, \qquad x(0) = x_0. \tag{1}$$

Here $x_t \in R^d$ is the unknown stochastic process; W_t is a Brownian motion ; f is the drift that determines the average trend , and G is the diffusion matrix that injects noise into the system.

Itô integral form Equation (1) is equivalent to the Itô integral equation

$$x_t \;=\; x_0 \;+\; \int_0^t f(x_s, s)\,\mathrm{d}s \;+\; \int_0^t G(x_s, s)\,\mathrm{d}W_s. \tag{2}$$

Definition of the Itô integral The stochastic integral in (2) is defined as the mean-square limit

$$\int_0^t G(x_s, s)\,\mathrm{d}W_s \;=\; \lim_{n\to\infty} \sum_{k=0}^{n-1} G\big(x(t_k), t_k\big)\,\big[W(t_{k+1}) - W(t_k)\big],$$

for partitions $0 = t_0 < t_1 < \cdots < t_n = t$, where the increments satisfy

$$W(t_{k+1}) - W(t_k) \;\sim\; \mathcal{N}\big(0,\,(t_{k+1} - t_k)I\big),$$

and increments over disjoint intervals are independent.

Existence and Uniqueness (Common Conditions) f, G globally Lipschitz + linear growth $\Rightarrow$ unique strong solution.

Euler–Maruyama Method Let the step size be $\Delta t = T/N$ and define grid points $t_n = n\,\Delta t$ for $n = 0, 1, \ldots, N$. Set the initial value $X_0 = x_0$. The Euler–Maruyama iteration is

$$X_{n+1} \;=\; X_n \;+\; f\big(X_n, t_n\big)\,\Delta t \;+\; G\big(X_n, t_n\big)\,\Delta W_n, \qquad n = 0, 1, \ldots, N-1,$$

where the noise increments satisfy

$$\Delta W_n \;\sim\; \mathcal{N}\big(0,\,\Delta t\, I\big), \quad \text{independently for each } n.$$

Problem 3 Unanswered Questions

1.What is the difference between a strong solution and a weak solution of an SDE?

2.Does the SDE have an analytic (closed-form) solution?