Data Privacy: Homeworks #2 Solutions

1. Show that the minimum mean square estimate is $\hat{X}(Y) = \mathbb{E}[X|Y]$. You can start from

$$\mathbb{E}\left[(\hat{X}(Y) - X)^2\right] = \mathbb{E}\left[((\hat{X}(Y) - \mathbb{E}[X|Y]) + (\mathbb{E}[X|Y] - X))^2\right] \tag{1}$$

Solution:

$$\begin{aligned}
\mathbb{E}\left[(\hat{X}(Y) - X)^2\right] =& \mathbb{E}\left[((\hat{X}(Y) - \mathbb{E}[X|Y]) + (\mathbb{E}[X|Y] - X))^2\right] && (2)\\
=& \mathbb{E}\left[(\hat{X}(Y) - \mathbb{E}[X|Y])^2\right] + 2\mathbb{E}\left[(\hat{X}(Y) - \mathbb{E}[X|Y])(\mathbb{E}[X|Y] - X)\right] \\
&+ \mathbb{E}\left[(\mathbb{E}[X|Y] - X)^2\right] && (3)\\
=& \mathbb{E}\left[(\hat{X}(Y) - \mathbb{E}[X|Y])^2\right] + 2\mathbb{E}\left[\mathbb{E}\left[(\hat{X}(Y) - \mathbb{E}[X|Y])(\mathbb{E}[X|Y] - X) \,\middle|\, Y\right]\right] \\
&+ \mathbb{E}\left[(\mathbb{E}[X|Y] - X)^2\right] && (4)\\
=& \mathbb{E}\left[(\hat{X}(Y) - \mathbb{E}[X|Y])^2\right] + 2\mathbb{E}\left[(\hat{X}(Y) - \mathbb{E}[X|Y])(\mathbb{E}[X|Y] - \mathbb{E}[X|Y])\right] \\
&+ \mathbb{E}\left[(\mathbb{E}[X|Y] - X)^2\right] && (5)\\
=& \mathbb{E}\left[(\hat{X}(Y) - \mathbb{E}[X|Y])^2\right] + \mathbb{E}\left[(\mathbb{E}[X|Y] - X)^2\right] && (6)\\
\geq& \mathbb{E}\left[(\mathbb{E}[X|Y] - X)^2\right]. && (7)
\end{aligned}$$

Clearly the equality holds iff $\hat{X}(Y) = \mathbb{E}[X|Y]$.

2. Let X be a random variable and $Z \sim \mathcal{N}(0, \sigma^2)$ be an independent Gaussian noise. Let $Y = X + Z$ be noise corrupted version of X.

   (a) Show that

$$\mathbb{E}[X|Y = y] = y + \sigma^2 \frac{\partial}{\partial y} \log P_Y(y). \tag{8}$$

   Hint: You can start from

$$P_Y(y) = \int P_X(x) P_Z(y - x)\, dx \tag{9}$$

   and compare it with $\frac{\partial}{\partial y} P_Y(y)$.

   (b) **(Bonus)** What about $\mathrm{Var}[X|Y = y]$?

Solution:

(a)

$$\partial_y P_Y(y) = \int P_X(x) \partial_y P_Z(y-x)\, dx \tag{10}$$

$$\partial_y P_Y(y) = \int P_X(x) \partial_y \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-y)^2}{2\sigma^2}\right) dx \tag{11}$$

$$\partial_y P_Y(y) = \int P_X(x) \frac{1}{\sqrt{2\pi\sigma^2}} \left(\frac{x-y}{\sigma^2}\right) \exp\left(-\frac{(x-y)^2}{2\sigma^2}\right) dx \tag{12}$$

$$\partial_y P_Y(y) = -\frac{y}{\sigma^2} P_Y(y) + \int P_X(x) \frac{1}{\sqrt{2\pi\sigma^2}} \left(\frac{x}{\sigma^2}\right) \exp\left(-\frac{(x-y)^2}{2\sigma^2}\right) dx \tag{13}$$

$$\partial_y P_Y(y) = -\frac{y}{\sigma^2} P_Y(y) + \int \left(\frac{x}{\sigma^2}\right) P_X(x) P_{Y|X}(y|x)\, dx \tag{14}$$

$$\partial_y P_Y(y) = -\frac{y}{\sigma^2} P_Y(y) + \int \left(\frac{x}{\sigma^2}\right) P_Y(y) P_{X|Y}(x|y)\, dx \tag{15}$$

$$\partial_y P_Y(y) = -\frac{y}{\sigma^2} P_Y(y) + P_Y(y) \frac{1}{\sigma^2} \mathbb{E}\left[X|Y\right] \tag{16}$$

Since $\partial_y \log P_Y(y) = \frac{\partial_y P_Y(y)}{P_Y(y)}$, we have

$$\mathbb{E}\left[X|Y\right] = y + \sigma^2 \partial_y \log P_Y(y). \tag{17}$$

(b)

$$\partial_y^2 \log P_Y(y) = \frac{\partial_y^2 P_Y(y) P_Y(y) - (\partial_y P_Y(y))^2}{P_Y(y)^2} \tag{18}$$

$$= \frac{\partial_y^2 P_Y(y)}{P_Y(y)} - \frac{(\partial_y P_Y(y))^2}{P_Y(y)^2} \tag{19}$$

The second term is simply $(\partial_y \log P_Y(y))^2$. Then,

$$\partial_y^2 P_Y(y) = -\partial_y \frac{y}{\sigma^2} P_Y(y) + \partial_y P_Y(y) \frac{1}{\sigma^2} \mathbb{E}\left[X|Y\right] \tag{20}$$

$$= -\frac{1}{\sigma^2} P_Y(y) - \frac{y}{\sigma^2} \partial_y P_Y(y) + \partial_y \int P_X(x) \frac{1}{\sqrt{2\pi\sigma^2}} \left(\frac{x}{\sigma^2}\right) \exp\left(-\frac{(x-y)^2}{2\sigma^2}\right) dx \tag{21}$$

$$= -\frac{1}{\sigma^2} P_Y(y) - \frac{y}{\sigma^2} \partial_y P_Y(y) + \int P_X(x) \frac{1}{\sqrt{2\pi\sigma^2}} \left(\frac{x}{\sigma^2}\right)^2 \exp\left(-\frac{(x-y)^2}{2\sigma^2}\right) dx \tag{22}$$

$$= -\frac{1}{\sigma^2} P_Y(y) - \frac{y}{\sigma^2} \partial_y P_Y(y) + \int \left(\frac{x(x-y)}{\sigma^4}\right) P_{X|Y}(x|y) P_Y(y)\, dx \tag{23}$$

$$= -\frac{1}{\sigma^2} P_Y(y) - \frac{y}{\sigma^2} \left(-\frac{y}{\sigma^2} P_Y(y) + P_Y(y) \frac{1}{\sigma^2} \mathbb{E}\left[X|Y\right]\right)$$
$$+ \frac{1}{\sigma^2} \mathbb{E}\left[X^2|Y\right] P_Y(y) - \frac{y}{\sigma^4} \mathbb{E}\left[X|Y\right] P_Y(y) \tag{24}$$

Thus,

$$\frac{\partial_y^2 P_Y(y)}{P_Y(y)} = \frac{1}{\sigma^4}\left(-\sigma^2 + y^2 - 2y\mathbb{E}\left[X|Y\right] + \mathbb{E}\left[X^2|Y\right]\right) \tag{25}$$

On the other hand,

$$\frac{(\partial_y P_Y(y))^2}{P_Y(y)^2} =(\partial_y \log P_Y(y))^2 \tag{26}$$

$$=\frac{1}{\sigma^4}(\mathbb{E}\left[X|Y\right]^2 + y^2 - 2y\mathbb{E}\left[X|Y\right]). \tag{27}$$

Finally, we have

$$\sigma^4\partial_y^2 \log P_Y(y) =\mathbb{E}\left[X^2|Y\right] - \mathbb{E}\left[X|Y\right]^2 - \sigma^2 \tag{28}$$

$$=\mathrm{Var}\left[X|Y\right] - \sigma^2 \tag{29}$$

which implies

$$\mathrm{Var}\left[X|Y\right] = \sigma^2 + \sigma^4\partial_y^2 \log P_Y(y). \tag{30}$$

3. Let q be the forwarding distribution of DDPM. Show that

$$q(X^{(n-1)}|X^{(n)}, X^{(0)}) = \mathcal{N}\left(\mu_n(X^{(n)}|X^{(0)}), \frac{1-\bar{\alpha}_{n-1}}{1-\bar{\alpha}_n}\beta_n\right) \tag{31}$$

where

$$\mu_n(X^{(n)}|X^{(0)}) = \frac{1}{\sqrt{1-\beta_n}}\left(X^{(n)} + \beta_n \frac{\partial}{\partial_{X^{(n)}}} \log q(X^{(n)}|X^{(0)})\right) \tag{32}$$

Hint: Given $X^{(0)}$, all distributions are Gaussian. More precisely,

$$\begin{bmatrix} X^{(n-1)} \\ X^{(n)} \end{bmatrix}\Bigg| X^{0)} \tag{33}$$

is multivariate Gaussian. You can use the fact that if

$$\begin{bmatrix} U \\ V \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu_U \\ \mu_V \end{bmatrix}, \begin{bmatrix} \sigma_{UU} & \sigma_{UV} \\ \sigma_{VU} & \sigma_{VV} \end{bmatrix}\right) \tag{34}$$

then

$$U|V \sim \mathcal{N}\left(\mu_U + (\sigma_{UV}/\sigma_{VV})(V-\mu_V), \sigma_{UU} - \sigma_{UV}^2/\sigma_{VV}\right). \tag{35}$$

Also note that we have an analytic formula for

$$\frac{\partial}{\partial_{X^{(n)}}} \log q(X^{(n)}|X^{(0)}) \tag{36}$$

Solution:
Since $X^{(n-1)}|X^{(n)}, X^{(0)}$ is Gaussian, it is enough to compute the mean and variance. Clearly,

$$\begin{bmatrix} X^{(n-1)} \\ X^{(n)} \end{bmatrix} \Bigg| X^{0)} \sim \mathcal{N}\left(\begin{bmatrix} \sqrt{\bar{\alpha}_{n-1}}X^{(0)} \\ \sqrt{\bar{\alpha}_n}X^{(0)} \end{bmatrix}, \begin{bmatrix} 1-\bar{\alpha}_{n-1} & \sqrt{1-\beta_n}(1-\bar{\alpha}_{n-1}) \\ \sqrt{1-\beta_n}(1-\bar{\alpha}_{n-1}) & 1-\bar{\alpha}_n \end{bmatrix}\right) \tag{37}$$

Then, the conditional mean is

$$\begin{aligned} &\sqrt{\bar{\alpha}_{n-1}}X^{(0)} + \frac{\sqrt{1-\beta_n}(1-\bar{\alpha}_{n-1})}{1-\bar{\alpha}_n}(X^{(n)} - \sqrt{\bar{\alpha}_n}X^{(0)}) \\ =&\left(\sqrt{\bar{\alpha}_{n-1}} - \frac{\sqrt{1-\beta_n}(1-\bar{\alpha}_{n-1})}{1-\bar{\alpha}_n}\sqrt{\bar{\alpha}_n}\right)X^{(0)} + \frac{\sqrt{1-\beta_n}(1-\bar{\alpha}_{n-1})}{1-\bar{\alpha}_n}X^{(n)}. \end{aligned} \tag{38}$$

Note that

$$\begin{aligned} &\sqrt{\bar{\alpha}_{n-1}} - \frac{\sqrt{1-\beta_n}(1-\bar{\alpha}_{n-1})}{1-\bar{\alpha}_n}\sqrt{\bar{\alpha}_n} \\ =&\frac{\sqrt{\bar{\alpha}_{n-1}}}{1-\bar{\alpha}_n}(1-\bar{\alpha}_n-(1-\beta_n)(1-\bar{\alpha}_{n-1})) &(39)\\ =&\frac{\sqrt{\bar{\alpha}_{n-1}}}{1-\bar{\alpha}_n}(1-\bar{\alpha}_{n-1}(1-\beta_n)-(1-\beta_n)(1-\bar{\alpha}_{n-1})) &(40)\\ =&\frac{\sqrt{\bar{\alpha}_{n-1}}}{1-\bar{\alpha}_n}\beta_n. &(41) \end{aligned}$$

Thus, the conditional mean is

$$\begin{aligned} &\sqrt{\bar{\alpha}_{n-1}}X^{(0)} + \frac{\sqrt{1-\beta_n}(1-\bar{\alpha}_{n-1})}{1-\bar{\alpha}_n}(X^{(n)} - \sqrt{\bar{\alpha}_n}X^{(0)}) \\ =&\frac{\sqrt{\bar{\alpha}_{n-1}}}{1-\bar{\alpha}_n}\beta_n X^{(0)} + \frac{\sqrt{1-\beta_n}(1-\bar{\alpha}_{n-1})}{1-\bar{\alpha}_n}X^{(n)}. \end{aligned} \tag{42}$$

Recall $\bar{\alpha}_n = \bar{\alpha}_{n-1}(1-\beta_n)$. For comparison,

$$\begin{aligned} \mu_n(X^{(n)}|X^{(0)}) =&\frac{1}{\sqrt{1-\beta_n}}\left(X^{(n)} + \beta_n \partial_{X^{(n)}} \log q(X^{(0)}|X^{(0)})\right) &(43)\\ =&\frac{1}{\sqrt{1-\beta_n}}\left(X^{(n)} + \beta_n\left(-\frac{X^{(n)} - \sqrt{\bar{\alpha}_n}X^{(0)}}{1-\bar{\alpha}_n}\right)\right) &(44)\\ =&\frac{\sqrt{\bar{\alpha}_{n-1}}}{1-\bar{\alpha}_n}\beta_n X^{(0)} + \frac{1}{\sqrt{1-\beta_n}(1-\bar{\alpha}_n)}\left(1-\bar{\alpha}_n-\beta_n\right)X^{(n)} &(45)\\ =&\frac{\sqrt{\bar{\alpha}_{n-1}}}{1-\bar{\alpha}_n}\beta_n X^{(0)} + \frac{1-\bar{\alpha}_{n-1}}{(1-\bar{\alpha}_n)}\sqrt{1-\beta_n}X^{(n)}. &(46) \end{aligned}$$

On the other hand, the conditional variance is

$$\begin{aligned}
&1-\bar{\alpha}_{n-1}-\frac{(1-\beta_n)(1-\bar{\alpha}_{n-1})^2}{1-\bar{\alpha}_n} \\
&=\frac{1-\bar{\alpha}_{n-1}}{1-\bar{\alpha}_n}\left(1-\bar{\alpha}_n-(1-\beta_n)(1-\bar{\alpha}_{n-1})\right) && (47)\\
&=\frac{1-\bar{\alpha}_{n-1}}{1-\bar{\alpha}_n}\left(1-\bar{\alpha}_{n-1}(1-\beta_n)-(1-\beta_n)(1-\bar{\alpha}_{n-1})\right) && (48)\\
&=\frac{1-\bar{\alpha}_{n-1}}{1-\bar{\alpha}_n}\beta_n && (49)
\end{aligned}$$

4. Show that

$$D_{\mathrm{KL}}\left(\mathcal{N}(\mu_0,\sigma_0^2 I)\|\mathcal{N}(\mu_1,\sigma_1^2 I)\right)=\frac{1}{2\sigma_1^2}\|\mu_1-\mu_0\|^2+\frac{d}{2}\left(\frac{\sigma_0^2}{\sigma_1^2}-1\right)+d\log\frac{\sigma_1}{\sigma_0}. \tag{50}$$

where d is dimension of Gaussian distributions and $\mu_0,\mu_1\in\mathbb{R}^d,\sigma_1,\sigma_2>0$.

Solution:

$$\begin{aligned}
&D_{KL}(\mathcal{N}(\mu_0,\sigma_0^2 I)\|\mathcal{N}(\mu_1,\sigma_1^2 I)) \\
&=\mathbb{E}_{\mu_0,\sigma_0}\left[\log\frac{\frac{1}{\sqrt{(2\pi\sigma_0)^d}}\exp\left(-\frac{1}{2\sigma_0^2}\|X-\mu_0\|^2\right)}{\frac{1}{\sqrt{(2\pi\sigma_1)^d}}\exp\left(-\frac{1}{2\sigma_1^2}\|X-\mu_1\|^2\right)}\right] && (51)\\
&=d\log\frac{\sigma_1}{\sigma_0}+\mathbb{E}_{\mu_0,\sigma_0}\left[-\frac{1}{2\sigma_0^2}\|X-\mu_0\|^2+\frac{1}{2\sigma_1^2}\|X-\mu_1\|^2\right] && (52)
\end{aligned}$$

Since $\mathbb{E}_{\mu_0,\sigma_0}[X-\mu_0]=0$, we have The second term is simply

$$\begin{aligned}
&\mathbb{E}_{\mu_0,\sigma_0}\left[-\frac{1}{2\sigma_0^2}\|X-\mu_0\|^2+\frac{1}{2\sigma_1^2}\|X-\mu_1\|^2\right] \\
&\mathbb{E}_{\mu_0,\sigma_0}\left[-\frac{1}{2\sigma_0^2}\|X-\mu_0\|^2+\frac{1}{2\sigma_1^2}\|X-\mu_0\|^2+\frac{1}{2\sigma_1^2}\|\mu_1-\mu_0\|^2\right] && (53)\\
=&\mathbb{E}_{\mu_0,\sigma_0}\left[\left(\frac{1}{2\sigma_1^2}-\frac{1}{2\sigma_0^2}\right)\|X-\mu_0\|^2\right]+\frac{1}{2\sigma_1^2}\|\mu_1-\mu_0\|^2 && (54)\\
=&\left(\frac{1}{2\sigma_1^2}-\frac{1}{2\sigma_0^2}\right)d\sigma_0^2+\frac{1}{2\sigma_1^2}\|\mu_1-\mu_0\|^2 && (55)\\
=&\frac{1}{2\sigma_1^2}\|\mu_1-\mu_0\|^2+\frac{d}{2}\left(\frac{\sigma_0^2}{\sigma_1^2}-1\right). && (56)
\end{aligned}$$

5. Let $\sigma_t\geq 0$ be a smooth non-decreasing function for $0\leq t\leq T$. Define

$$\rho(t)=\sqrt{\frac{d}{dt}\sigma_t^2} \tag{57}$$

for $0 \le t \le T$. Consider the SDE

$$dX_t = \rho(t) dW_t \tag{58}$$

with initial condition $X_0 \sim p_0$ where W_t is standard Brownian motion. Show that $X_t|X_0 \sim \mathcal{N}(X_0, \sigma_t^2)$ by verifying that

$$p_t(x) = \int p_{t|0}(x|y)p_0(y)\,dy = \int \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{(x-y)^2}{2\sigma_t^2}\right) p_0(y)\,dy \tag{59}$$

satisfies the Fokker-Planck equation.

Solution:
The corresponding Fokker-Planck equation is given by

$$\partial_t p_t = \frac{\rho(t)^2}{2} \partial_x^2 p_t. \tag{60}$$

The left hand side is

$$\partial_t p_t = \int \partial_t \left(\frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{(x-y)^2}{2\sigma_t^2}\right) \right) p_0(y)\,dy \tag{61}$$

$$\begin{aligned}
&= \int \partial_t \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{(x-y)^2}{2\sigma_t^2}\right) p_0(y)\,dy \\
&\quad + \int \frac{1}{\sqrt{2\pi\sigma_t^2}} \partial_t \exp\left(-\frac{(x-y)^2}{2\sigma_t^2}\right) p_0(y)\,dy
\end{aligned} \tag{62}$$

$$\begin{aligned}
&= \int -\frac{\partial_t \sigma_t^2}{2\sqrt{2\pi}(\sigma_t^2)^{3/2}} \exp\left(-\frac{(x-y)^2}{2\sigma_t^2}\right) p_0(y)\,dy \\
&\quad + \int \frac{1}{\sqrt{2\pi\sigma_t^2}} (\partial_t \sigma_t^2) \left(\frac{(x-y)^2}{2\sigma_t^4}\right) \exp\left(-\frac{(x-y)^2}{2\sigma_t^2}\right) p_0(y)\,dy
\end{aligned} \tag{63}$$

$$\begin{aligned}
&= \frac{\rho(t)^2}{2} \int -\frac{1}{\sigma_t^2} \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{(x-y)^2}{2\sigma_t^2}\right) p_0(y)\,dy \\
&\quad + \frac{\rho(t)^2}{2} \int \frac{1}{\sqrt{2\pi\sigma_t^2}} \left(\frac{(x-y)^2}{\sigma_t^4}\right) \exp\left(-\frac{(x-y)^2}{2\sigma_t^2}\right) p_0(y)\,dy
\end{aligned} \tag{64}$$

where $\partial_t \sigma_t^2 = \rho(t)^2$.
On the other hand

$$\partial_x^2 p_t = \int \partial_x \left(\frac{1}{\sqrt{2\pi\sigma_t^2}} \left(-\frac{x-y}{\sigma_t^2}\right) \exp\left(-\frac{(x-y)^2}{2\sigma_t^2}\right) \right) p_0(y)\,dy \tag{65}$$

$$\begin{aligned}
&= \int \left(\frac{1}{\sqrt{2\pi\sigma_t^2}} \left(-\frac{1}{\sigma_t^2}\right) \exp\left(-\frac{(x-y)^2}{2\sigma_t^2}\right) \right) p_0(y)\,dy \\
&\quad + \int \left(\frac{1}{\sqrt{2\pi\sigma_t^2}} \left(\frac{(x-y)^2}{\sigma_t^4}\right) \exp\left(-\frac{(x-y)^2}{2\sigma_t^2}\right) \right) p_0(y)\,dy.
\end{aligned} \tag{66}$$

This concludes the proof.

6. Consider the ODE

$$dX_t = \left(f(X_t, t) - \frac{g^2(t)}{2} \partial_{X_t} \log p_t(X_t) \right) dt \tag{67}$$

with terminal condition $X_T \sim p_T$. Show that $\{p_t\}_{t=0}^T$ satisfies the Fokker–Planck equation

$$\partial_t p_t = -\partial_x (f p_t) + \frac{g^2}{2} \partial_x^2 p_t. \tag{68}$$

Hint: See the proof of Anderson's theorem and its derivation of the Fokker-Planck equation with SDE.

Solution:
Let reverse SDE by

$$d\bar{X}_t = (f(\bar{X}_t, t) - \frac{g^2(t)}{2} \nabla_x \log p_t(\bar{X}_t))dt \tag{69}$$

with $\bar{X}_T \sim p_T$. Alternatively, define $\{Y_t\}_{t=0}^T$ where $Y_t = \bar{X}_{T-t}$via

$$dY_t = -(f(Y_t, T-t) - \frac{g^2(T-t)}{2} \nabla_x p_{T-t}(Y_t))dt \tag{70}$$

with $Y_0 \sim p_T$.

Let $\{q_t\}_{t=0}^T$ be marginal densities of $\{Y_t\}_{t=0}^T$. Then, $\{q_t\}_{t=0}^T$ satisfies the FP equation

$$\partial_t q_t(y) = \partial_y \left((f(y, T-t) - \frac{g^2(T-t)}{2} \partial_y \log p_{T-t}(y)) q_t(y) \right). \tag{71}$$

Let $\{\bar{p}_t\}_{t=0}^T$ be marginal densities of $\{\bar{X}_t\}_{t=0}^T$. Since $\bar{p}_{T-t} = q_t$, then $\{\bar{p}_t\}_{t=0}^T$ satisfies the FP equation

$$\partial_t \bar{p}_t(x) = -\partial_x \left((f(x, t) - \frac{g^2(t)}{2} \partial_x \log p_t(x)) \bar{p}_t(x) \right). \tag{72}$$

The final step of the proof is proving that p_t solves the above reverse FP equation.

$$\begin{aligned} & -\partial_x \left((f(x,t) - \frac{g^2(t)}{2} \partial_x \log p_t(x)) p_t(x) \right) && (73) \\ =& -\partial_x (f p_t) + \frac{g^2(t)}{2} \partial_x^2 (p_t(x)) && (74) \\ =& \partial_x p_t && (75) \end{aligned}$$

where the last equation is from forward FP equation. Finally, p_t also satisfies equation 72, and therefore $p_t = \bar{p}_t$.

7. **Star shaped diffusion.** Consider the DDIM forward process.

$$q(X^{(N)}|X^{(0)}) = \mathcal{N}(\sqrt{\bar{\alpha}_N}X^{(0)}, (1-\bar{\alpha}_N)I) \tag{76}$$

$$q(X^{(n)}|X^{(n+1)}, X^{(0)}) = \mathcal{N}(\sqrt{\bar{\alpha}_n}X^{(0)} + \frac{\sqrt{1-\bar{\alpha}_n-\sigma_{n+1}^2}}{\sqrt{1-\bar{\alpha}_{n+1}}}(X^{(n+1)} - \sqrt{\bar{\alpha}_{n+1}}X^{(0)}), \sigma_{n+1}^2 I) \tag{77}$$

for $\sigma_n \geq 0$. Show that equation 77 matches the marginal distribution of DDPM.

Solution:
For $n = N$, it is clear that the distribution of $q(X^{(N)}|X^{(0)})$ matches the distribution.

$$X^{(N)} \overset{(d)}{=} \sqrt{\bar{\alpha}_N}X^{(0)} + Z^{(N)} \tag{78}$$

where $Z^{(N)} \sim \mathcal{N}(0, (1-\bar{\alpha}_N^2))$. Suppose DDIM forward process matches with DDPM at $n+1$, i.e.,

$$X^{(n+1)} \overset{(d)}{=} \sqrt{\bar{\alpha}_{n+1}}X^{(0)} + Z^{(n+1)} \tag{79}$$

where $Z^{(n+1)} \sim \mathcal{N}(0, (1-\bar{\alpha}_{n+1}^2))$. Then, we will show that DDIM forward process also matches with DDPM at n. From equation 77,

$$X^{(n)} \overset{(d)}{=} \sqrt{\bar{\alpha}_n}X^{(0)} + \frac{\sqrt{1-\bar{\alpha}_n-\sigma_{n+1}^2}}{\sqrt{1-\bar{\alpha}_{n+1}}}(X^{(n+1)} - \sqrt{\bar{\alpha}_{n+1}}X^{(0)}) + W^{(n+1)} \tag{80}$$

where $W^{(n+1)} \sim \mathcal{N}(0, \sigma_{n+1}^2 I)$. Then, from assumption where DDPM and DDIM matches at $n+1$, we have

$$X^{(n)} \overset{(d)}{=} \sqrt{\bar{\alpha}_n}X^{(0)} + \frac{\sqrt{1-\bar{\alpha}_n-\sigma_{n+1}^2}}{\sqrt{1-\bar{\alpha}_{n+1}}}Z^{(n+1)} + W^{(n+1)}. \tag{81}$$

The sum of two independent noise term has zero mean and covariance of

$$\frac{1-\bar{\alpha}_n-\sigma_{n+1}^2}{1-\bar{\alpha}_{n+1}}(1-\bar{\alpha}_{n+1})I + \sigma_{n+1}^2 I \tag{82}$$

$$= (1-\bar{\alpha}_n)I. \tag{83}$$

Thus,

$$X^{(n)} \sim \mathcal{N}(\sqrt{\bar{\alpha}_n}X^{(0)}, (1-\bar{\alpha}_n)I). \tag{84}$$

8. Let q be the forward process of star-diffusion model, which is given by

$$q(X^{(1)}, \ldots, X^{(N-1)}, X^{(N)}|X^{(0)}) = \prod_{n=1}^{N} q(X^{(n)}|X^{(0)})$$

$$q(X^{(n)}|X^{(0)}) \sim \mathcal{N}(\sqrt{\bar{\alpha}_n}X^{(0)}, (1-\bar{\alpha}_n)I).$$

Since it is non-Markovian process, the reverse process is given by

$$q(X^{(0)}, \ldots, X^{(N)}) = q(X^{(N)}) \prod_{n=1}^{N} q(X^{(n-1)}|X^{(n)}, \ldots, X^{(N)}),$$

which we approximate by

$$p_\theta(X^{(0)}, \ldots, X^{(N)}) = p_\theta(X^{(N)}) \prod_{n=1}^{N} p_\theta(X^{(n-1)}|X^{(n)}, \ldots, X^{(N)}).$$

Consider the variational lower bound

$$\mathcal{L}(\theta) \stackrel{def}{=} \mathbb{E}_q \left[\log \frac{p_\theta(X^{(0)}, \ldots, X^{(N)})}{q(X^{(1)}, \ldots, X^{(N)}|X^{(0)})} \right].$$

Show that

$$\begin{aligned}
\mathcal{L}(\theta) =& \mathbb{E}_q \left[\log p_\theta(X^{(0)}|X^{(1)}, \ldots, X^{(N)}) \right] \\
&- \mathbb{E}_q[\sum_{n=2}^{N} D(q(X^{(n-1)}|X^{(0)}) \| p_\theta(X^{(n-1)}|X^{(n)}, \ldots, X^{(N)}))] \\
&- \mathbb{E}_q[D(q(X^{(N)}|X^{(0)}) \| p_\theta(X^{(N)}))].
\end{aligned}$$

Solution:

$$\begin{aligned}
\mathcal{L}(\theta) =& \mathbb{E}_q \left[\log \frac{p_\theta(X^{(0)}, \ldots, X^{(N)})}{q(X^{(1)}, \ldots, X^{(N)}|X^{(0)})} \right] \\
=& \mathbb{E}_q \left[\log \frac{p_\theta(X^{(N)}) \prod_{n=1}^{N} p_\theta(X^{(n-1)}|X^{(n)}, \ldots, X^{(N)})}{\prod_{n=1}^{N} q(X^{(n)}|X^{(0)})} \right] \\
=& \mathbb{E}_q \left[\log p_\theta(X^{(0)}|X^{(1)}, \ldots, X^{(N)}) \right] \\
&- \sum_{n=2}^{N} \mathbb{E}_q[\log \frac{q(X^{(n-1)}|X^{(0)})}{p_\theta(X^{(n-1)}|X^{(n)}, \ldots, X^{(N)}))}] \\
&- \mathbb{E}_q[\log \frac{q(X^{(N)}|X^{(0)})}{p_\theta(X^{(N)}))}] \\
=& \mathbb{E}_q \left[\log p_\theta(X^{(0)}|X^{(1)}, \ldots, X^{(N)}) \right] \\
&- \mathbb{E}_q[\sum_{n=2}^{N} D(q(X^{(n-1)}|X^{(0)}) \| p_\theta(X^{(n-1)}|X^{(n)}, \ldots, X^{(N)}))] \\
&- \mathbb{E}_q[D(q(X^{(N)}|X^{(0)}) \| p_\theta(X^{(N)}))].
\end{aligned}$$

9. **Guidance.** In diffusion models, classifier-guided conditional generation uses

$$\nabla_x \log p_t(x) + \omega \nabla_x \log p_t(y \,|\, x)$$

with a classifier scale parameter $\omega \geq 1$. However, many papers in the literature add classifier guidance to an already conditional model via

$$\nabla_x \log p_t(x \mid y) + s\nabla_x \log p_t(y \mid x)$$

with $s \geq 0$. Show that the two are equivalent with $s = \omega - 1$.

Solution:
By Bayes's theorem, we have $p_t(x \mid y)p_t(y) = p_t(y \mid x)p_t(x)$. So we have

$$\begin{aligned}
\nabla_x \log p_t(x \mid y) + s\nabla_x \log p_t(y \mid x) &= \nabla_x \log\{p_t(x \mid y)p_t(y)\} + s\nabla_x \log p_t(y \mid x) \\
&= \nabla_x \log\{p_t(y \mid x)p_t(x)\} + s\nabla_x \log p_t(y \mid x) \\
&= \nabla_x \log p_t(x) + (s+1)\nabla_x \log p_t(y \mid x).
\end{aligned}$$

Therefore, with $s = \omega - 1$, we can see that the two formula are equivalent.

10. **Langevin** Let $p(x)$ is a probability density function that is smooth and strictly positive for all $x \in \mathbb{R}^d$. Let $\{p_t\}_{t\in[0,T]}$ be the marginal density functions of the Langevin SDE

$$dX_t = \frac{1}{2}\nabla_{X_t} \log p(X_t)dt + dW_t.$$

Show that if $p_0 = p$, then $p_t = p$ for all $t > 0$.

Solution:
The Fokker-Planck equation of Langevin SDE is

$$\partial_t p_t = -\frac{1}{2}\nabla_{X_t} \cdot (\nabla_{X_t} \log p(X_t) \cdot p_t) + \frac{1}{2}\Delta p_t.$$

If $p_t = p$, then

$$\begin{aligned}
\partial_t p_t &= -\frac{1}{2}\nabla_{X_t} \cdot (\nabla_{X_t} \log p(X_t) \cdot p(X_t)) + \frac{1}{2}\Delta p(X_t) \\
&= \frac{1}{2}\left(-\nabla_{X_t} \cdot \nabla_{X_t} p(X_t) + \Delta p(X_t)\right) \\
&= 0
\end{aligned}$$

We have $p_0 = p$ and we know that there is a unique solution. Therefore, the unique solution is $p_t = p$.