Example III : sub-VP

Here are two equivalent interpretations. (Follow the derivation of appendix D of 2022 Lipman et al.)

Conditional VF for Fokker-Planck probability paths

The SDE for the sub-VP path is :

$$\mathrm{d}\mathbf{y} = -\frac{1}{2}T'(t)\mathbf{y} + \sqrt{T'(t)(1-e^{-2T(t)})}\mathrm{d}\mathbf{w}$$

where $T(t) = \int_0^t \beta(t)\mathrm{d}t, t \in [0,1]$. The SDE coefficients are therefore:

$$f_s(\mathbf{y}) = -\frac{T'(s)}{2}\mathbf{y}, \quad g_s = \sqrt{T'(s)(1-e^{-2T(t)})}$$

$$p_t(\mathbf{y}|\mathbf{y}_0) = \mathcal{N}(\mathbf{y}|e^{-\frac{1}{2}T(t)}\mathbf{y}_0, (1-e^{-T(t)})^2\mathbf{I})$$

Using Eq.40 from 2022 Lipman et al.: (the vector field satisfies the continuity equation with the probability path p_t, and therefore generates p_t)

$$w_t = f_t - \frac{g_t^2}{2}\nabla \log p_t$$

We have

$$\begin{aligned}
w_t(\mathbf{y}|\mathbf{y}_0) &= -\frac{T'(t)}{2}\mathbf{y} + \frac{1}{2}T'(t)(1-e^{-2T(t)}) \cdot \frac{(\mathbf{y}-e^{-\frac{1}{2}T(t)}\mathbf{y}_0)}{(1-e^{-T(t)})^2} \\
&= -\frac{T'(t)}{2}\mathbf{y} + \frac{1}{2}T'(t)(1+e^{-T(t)}) \cdot \frac{(\mathbf{y}-e^{-\frac{1}{2}T(t)}\mathbf{y}_0)}{1-e^{-T(t)}} \\
&= \frac{1}{2}T'(t)\left((1+e^{-T(t)}) \cdot \frac{(\mathbf{y}-e^{-\frac{1}{2}T(t)}\mathbf{y}_0)}{1-e^{-T(t)}} - \mathbf{y}\right) \\
&= \frac{1}{2}T'(t)\left(\frac{2e^{-T(t)}\mathbf{y} - \mathbf{y}_0(e^{-\frac{3}{2}T(t)} + e^{-\frac{1}{2}T(t)})}{1-e^{-T(t)}}\right)
\end{aligned}$$

And according to Lemma 1 from 2022 Lipman et al., we reverse the time and then get the conditional VF for the reverse probability path:

$$\tilde{w}_t(\mathbf{y}|\mathbf{y}_0) = -\frac{1}{2}T'(1-t)\left(\frac{2e^{-T(1-t)}\mathbf{y} - \mathbf{y}_0(e^{-\frac{3}{2}T(1-t)} + e^{-\frac{1}{2}T(1-t)})}{1-e^{-T(1-t)}}\right)$$

> **Lemma 1.** Consider a flow defined by a vector field $u_t(x)$ generating probability density path $p_t(x)$. Then, the vector field $\tilde{u}_t(x) = -u_{1-t}(x)$ generates the path $\tilde{p}_t(x) = p_{1-t}(x)$ when initiated from $\tilde{p}_0(x) = p_1(x)$.

Conditional VF for Flow Matching

$$p_{0t}(\mathbf{x}(t) \mid \mathbf{x}(0)) = \mathcal{N}(\mathbf{x}(t); \mathbf{x}(0)e^{-\frac{1}{2}\int_0^t \beta(s)\,\mathrm{d}s}, [1-e^{-\int_0^t \beta(s)\,\mathrm{d}s}]^2\mathbf{I})$$

$$p_t(\mathbf{x} \mid \mathbf{x}_1) = \mathcal{N}(\mathbf{x}|\alpha_{1-t}\mathbf{x}_1, (1-\alpha_{1-t}^2)^2 I), \text{ where } \alpha_t = e^{-\frac{1}{2}T(t)}, T(t) = \int_0^t \beta(s)\mathrm{d}s$$

We have: $\sigma_t(\mathbf{x}_1) = 1-\alpha_{1-t}^2$, and $\mu_t(\mathbf{x}_1) = \alpha_{1-t}\mathbf{x}_1$.

According to theorem 3 from [2022 Lipman et al.]:

> **Theorem 3.** Let $p_t(x \mid x_1)$ be a Gaussian probability path as in equation 10, and ψ_t its corresponding flow map as in equation 11. Then, the unique vector field that defines ψ_t has the form:
>
> $$u_t(x \mid x_1) = \frac{\sigma_t'(x_1)}{\sigma_t(x_1)}(x - \mu_t(x_1)) + \mu_t'(x_1). \tag{15}$$
>
> Consequently, $u_t(x \mid x_1)$ generates the Gaussian path $p_t(x \mid x_1)$.

We can derive:

$$u_t(\mathbf{x} \mid \mathbf{x}_1) = \frac{2\alpha_{1-t}\alpha_{1-t}'}{1-\alpha_{1-t}^2}(\mathbf{x} - \alpha_{1-t}\mathbf{x}_1) - \alpha_{1-t}\alpha_{1-t}'\mathbf{x}_1$$

$$u_t(\mathbf{x} \mid \mathbf{x}_1) = -\frac{1}{2}T'(1-t)\left(\frac{2e^{-T(1-t)}\mathbf{x} - \mathbf{x}_1(e^{-\frac{3}{2}T(1-t)} + e^{-\frac{1}{2}T(1-t)})}{1-e^{-T(1-t)}}\right)$$

This coincides with the derivation from the perspective of probability flow ODE.

Summary

Here we give a brief summary regarding the vector field(VF) of VE,VP,sub-VP,OT:

VE

$$u_t(\mathbf{x} \mid \mathbf{x}_1) = -\frac{\sigma_{1-t}'}{\sigma_{1-t}}(\mathbf{x} - \mathbf{x}_1).$$

VP

$$u_t(\mathbf{x} \mid \mathbf{x}_1) = \frac{\alpha_{1-t}'}{1-\alpha_{1-t}^2}(\alpha_{1-t}\mathbf{x} - \mathbf{x}_1) = -\frac{T'(1-t)}{2}\left[\frac{e^{-T(1-t)}\mathbf{x} - e^{-\frac{1}{2}T(1-t)}\mathbf{x}_1}{1-e^{-T(1-t)}}\right].$$

sub-VP

$$u_t(\mathbf{x} \mid \mathbf{x}_1) = -\frac{1}{2}T'(1-t)\left(\frac{2e^{-T(1-t)}\mathbf{x} - \mathbf{x}_1(e^{-\frac{3}{2}T(1-t)} + e^{-\frac{1}{2}T(1-t)})}{1-e^{-T(1-t)}}\right).$$

OT

$$u_t(x \mid x_1) = \frac{x_1 - (1-\sigma_{\min})x}{1-(1-\sigma_{\min})t}.$$

Optimal Transport (OT) has a very elegant representation. In the next section, we'll show through examples that OT also performs better than other methods at learning simple distributions.