April 3 2025 Conditional Flow Matching

Lipman+ ICML 2023

A breakthrough in training CNFs:

idea: try to directly regress vector field, using a trick.

Give up on max likelihood training

$$L_{FM} = \mathbb{E}_{\substack{t \sim U[0,1] \\ x \sim p_t(x)}} \| \underbrace{u_\theta(x,t)}_{NN} - u_t(x) \|^2$$

Obvious problem: this is totally intractable since u_t & p_t are not known! (that's the whole pt)

Idea/trick: what if we instead try to regress conditional paths

$$p_t(x|y) \nearrow p_0(x|y) = \mathcal{N}(0,1)$$

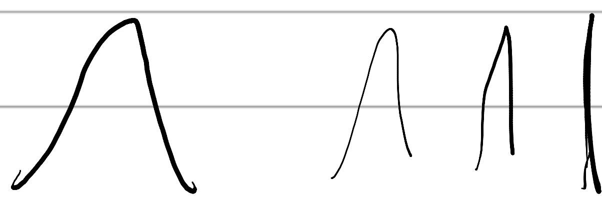
$$\searrow p_1(x|y) = \underbrace{\delta(x-y)}_{\text{building blocks of gen dist'n.}}$$

$$\boxed{p_t(x) = \int dx_1\, p_t(x|x_1)\, p_{data}(x_1)}$$

correct bdy conditions

$$\begin{cases} p_0(x) = p_{base} = \mathcal{N}(0,1) \\ p_1(x) = p_{data}(x) \end{cases}$$

Very obvious (but not only) choice for $p_t(x|y)$: $\mathcal{N}(x|\gamma_t y, \sigma_t^2 \mathbb{1})$



with $\begin{cases} (\gamma_0, \sigma_0) = (0,1) \\ (\gamma_1, \sigma_1) = (1,0) \end{cases}$

$t=0$ $\qquad$ $t=1$

"Gaussian probability path"

$X_t \sim \rho_t(x \mid x_1)$

$X_t = \gamma_t x_1 + \sigma_t \varepsilon \qquad \varepsilon \sim N(0,1)$

then

$u_t(x \mid x_1) = \frac{dX_t}{dt} = \dot{\gamma}_t x_1 + \dot{\sigma}_t \varepsilon$ conditional vector field

Claim: $u_t(x) = \int dx_1 \, \frac{u_t(x \mid x_1)\, \rho_t(x \mid x_1)\, \rho_{data}(x_1)}{\rho_t(x)}$

Pf: start from $\rho_t(x) = \int dx_1 \, \rho_t(x \mid x_1)\, \rho_{data}(x_1)$

apply continuity eqn to both sides

$$\frac{\partial \rho_t}{\partial t} = \int dx_1 \, \frac{\partial \rho_t(x \mid x_1)}{\partial t} \rho_{data}(x_1)$$

$$= \int dx_1 \, -\vec{\nabla}_x \cdot \left(u_t(x \mid x_1)\, \rho_t(x \mid x_1) \right) \rho_{data}(x_1)$$

$$= -\vec{\nabla}_x \left(\rho_t(x)\, u_t(x) \right)$$

So we identify $u_t(x)$ w/ formula above.

Brilliant insight: regressing <u>conditional vector field</u> is equiv to regressing unconditional one!

$$L_{CFM} = \mathbb{E}_{\substack{t \sim U(0,1) \\ x_t \sim p_t(x_t|x_1) \\ x_1 \sim p_{data}}} \| u_\theta(x_t, t) - u_t(x_t|x_1) \|^2$$

<u>Claim</u>: $u_{\theta^*}(x,t)$ matches that of $\min L_{FM}$

<u>Pf</u>: $\| u_\theta(x_t, t) - u_t(x_t|x_1) \|^2$

$$= \|u_\theta(x_t,t)\|^2 - 2\, u_\theta(x_t,t) \cdot u_t(x_t|x_1) + \underbrace{\|u_t(x_t|x_1)\|^2}_{\text{const indep of } \theta}$$

$$\mathop{\mathbb{E}}_{\substack{x_t \sim p_t(x_t|x_1) \\ x_1 \sim p_{data}}} \|u_\theta(x_t,t)\|^2 = \mathop{\mathbb{E}}_{x_t \sim p_t(x_t)} \|u_\theta(x_t,t)\|^2$$

since indep of x_1

$$\mathop{\mathbb{E}}_{x_t \sim p_t(x_t)} u_\theta(x_t,t) \cdot u_t(x_t) =$$

$$\int dx_t\, \cancel{p_t(x_t)}\, u_\theta(x_t,t) \cdot \int dx_1 \frac{u_t(x_t|x_1)\, p_t(x_t|x_1)\, p_{data}(x_1)}{\cancel{p_t(x_t)}}$$

$$= \mathop{\mathbb{E}}_{\substack{x_t \sim p_t(x_t|x_1) \\ x_1 \sim p_{data}}} u_\theta(x_t,t) \cdot u_t(x_t|x_1)$$

So two losses have same θ dep!

$$L_{CFM}(\theta) = L_{FM}(\theta) + \text{const}$$

must have same minim!

So CFM objective gets us correct uncond'l vector field!

The final objective is shockingly easy!

Still have choices for γ_t, σ_t

Simplest: $\gamma_t = t, \sigma_t = 1-t$

$$u_t(x_t | x_1) = x_1 - \varepsilon$$

$$\hookrightarrow \varepsilon = \frac{x_t - x_1 \gamma_t}{\sigma_t}$$

$$\boxed{\mathbb{E}_{\substack{t \sim U[0,1] \\ \varepsilon \sim N(0,1) \\ x_1 \sim p_{data}}} \| u_\theta(x_t, t) - (x_1 - \varepsilon) \|^2}$$

$$\hookrightarrow = \varepsilon \sigma_t + x_1 \gamma_t$$

Incredibly simple setup in the end!