Flow matching:

↳ u_t^θ (θ: parameters)

Neural network

Goal: Marginal Vector field. (sim. with Neural Net)

↳ $u_t^\theta \approx u_t^{target}$

• Flow matching Loss: (marginal vector field.)

↳ $L_{fm}(\theta) = \mathbb{E}\left[\| u_t^\theta(x) - u_t^{target}(x)\|^2\right]$

From any time 0-1, choose a Data point, ⊥ validate in intermediate distribution

- $t \sim$ Unif ← Uniform in $[0,1]$
- $z \sim p_{data}$ ← draw data point
- $x \sim p_t(\cdot|z)$ ← draw conditional prob. Path.

* Nevertheless, we can't compute it!

↓ Thus we use a conditional loss

• Conditional Flow matching Loss:

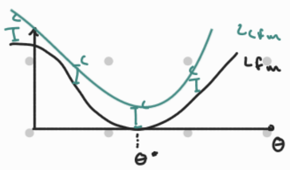
↳ $L_{cfm}(\theta) = \mathbb{E}\left[\| u_t^\theta(x) - u_t^{target}(x|z)\|^2\right]$ (cond. vector field.)

$t \sim$ Unif
$z \sim p_{data}$
$x \sim p_t(\cdot|z)$

↳ why would we need this if we want to generate points? we need this to compute

* $L_{fm}(\theta) = L_{cfm}(\theta) + C$

↳ for $C < 0$ independent of θ



Theorem 4.1

* ① For minimizer θ^* of L_{cfm}: $\boxed{u_t^{\theta^*} = u_t^{target}}$

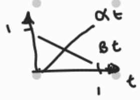
② $\nabla_\theta L_{cfm}(\theta) = \nabla_\theta L_{fm}(\theta)$

→ SGD the Same.

Stochastic Gradient Descent

→ L_{cfm} for Gaussian Cond. Path:

$p_t(\cdot|z) = N(\underbrace{\alpha_t z}_{\text{mean}}, \underbrace{\beta_t^{2T} I_d}_{\text{Variance}})$



we know α_t, β_t

→ $u_t^{target}(x|z) = \left(\dot{\alpha}_t - \frac{\dot{\beta}_t}{\beta_t}\alpha_t\right) z + \frac{\dot{\beta}_t}{\beta_t} x$

Diff.(?) ← $\underset{\text{uniform Gaussian}}{\overset{\text{noise}}{\varepsilon}} \sim N(0, I_d) \Rightarrow \alpha_t z + \beta_t \varepsilon \overset{def}{=} x \sim p_t(\cdot|z)$

Gaussian Distribution!

$\dot{\alpha}_t = \frac{d}{dt}\alpha_t$, $\dot{\beta}_t = \frac{d}{dt}\beta_t$

$L_{cfm}(\theta) = \mathbb{E}_{t\sim Unif,\, z\sim p_{data},\, x\sim N(\alpha_t z, \beta_t^2 I_d)}\left[\left\| u_t^\theta(x) - \left(\dot{\alpha}_t - \frac{\dot{\beta}_t}{\beta_t}\alpha_t\right) z - \frac{\dot{\beta}_t}{\beta_t} x\right\|^2\right]$

$x = \alpha_t z + \beta_t \varepsilon$

$= \mathbb{E}\left[\| u_t^\theta(\alpha_t z + \beta_t \varepsilon) - (\dot{\alpha}_t z + \dot{\beta}_t z\|^2\right]$

$t \sim$ Unif
$z \sim p_{data}$
$\varepsilon \sim N(0, I_d)$

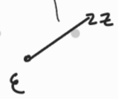
↳ we sample time Uniformly, sample data point from dataset, sample noise

* we Predict Velocities

ε • → • z

Conditional OT path: $\alpha_t = t$; $\beta_t = 1 - t$

Optimal transport $\quad \dot{\alpha}_t = 1$; $\dot{\beta}_t = -1$



$\overbrace{\|u_t^\theta(x) - (z - \varepsilon)\|^2}^{}$ → same.

$$L_{CFM}(\theta) = \mathbb{E}_{\substack{t \sim \text{Unif} \\ z \sim p_{data} \\ \varepsilon \sim N(0, I_d)}} \left[\| u_t^\theta (\underbrace{tz + (1-t)\varepsilon}_{\text{Straight line}}) - (\underbrace{z - \varepsilon}_{\text{Difference.}}) \|^2 \right]$$

Straight line → we grab a point on line between ε and Data

Difference. → Diff between ε and Data

Proof n.)

↳ $\|a - b\|^2 = \|a\|^2 - 2a^T b + \|b\|^2 \quad (a, b) \in \mathbb{R}^d$ (vectors)

- $L_{FM}(\theta) = \mathbb{E}_{t,z,x}\left[\|u_t^\theta(x) - u_t^{target}(x)\|^2\right] = \mathbb{E}_{t,z,x}\left[\|\cancel{u_t^\theta(x)}\|^2 - \overbrace{2u_t^\theta(x)^T u_t^{target}(x)}^{\text{marginal vector field}} + \|\cancel{u_t^{target}(x)}\|^2\right]$ → constants
- $L_{CFM}(\theta) = \mathbb{E}_{t,z,x}\left[\|u_t^\theta(x) - u_t^{target}(x|z)\|^2\right] = \mathbb{E}_{t,z,x}\left[\|\cancel{u_t^\theta}(x)\|^2 - \underbrace{2u_t^\theta(x) u_t^{target}(x|z)}_{\text{Cond. vector field}} + \|\cancel{u_t^{target}}(x|z)\|^2\right.$

same.

↳ Neural Network same.