

条件扩散模型

扩散后验采样

图片 x_0 \longleftrightarrow x_T 白噪音

目标: 采样 $q_0(x_0 | y)$

$$q_t(x_t | y) = \int q_t(x_t | x_0 | y) dx_0 \\ = \int q_0(x_0 | y) q_t(x_t | x_0) dx_0$$

$$x_0 \sim q_0(x_0 | y)$$

$$dx_t = f(t) x_t dt + g(t) dW_t$$

$$x_t \sim q_t(x_t | y)$$

计算 $\nabla_x \log q_{T-t}(x | y)$

$$= \nabla_x \log \frac{q_{T-t}(y | x) q_{T-t}(x)}{q(y)}$$

$$= \nabla_x \log q_{T-t}(y | x) + \nabla_x \log q_{T-t}(x)$$

$$= \nabla_x \log q_{T-t}(y | x) + \nabla_x \log q_{T-t}(x)$$

↑
score

$$q_t(y | x) = \int q_t(y | x_0 | x) dx_0$$

$$= \int q_t(y | x_0 | x) q_t(x_0 | x) dx_0$$

$$= \int q_0(y | x_0) \underbrace{q_{t_0}(x_0 | x)}_{\text{计算}} dx_0$$

↑
似然函数

Tweedie 公式

$$q_{0t}(x_t | x_0) = N(x_t | \lambda_t x_0, \delta_t^2 I)$$

$$q_{t0}(x_0 | x_t) = \frac{q_{0t}(x_t | x_0) q_0(x_0)}{q_t(x_t)}$$

$$q_t(x_t) = \int q_{0t}(x_t | x_0) q_0(x_0) dx_0$$

$$\nabla q_t(x_t) = \int q_{0t}(x_t | x_0) \frac{x_t - \lambda_t x_0}{\delta_t^2} q_0(x_0) dx_0$$

$$\frac{\nabla q_t(x_t)}{q_t(x_t)} = \int q_{0t}(x_0 | x_t) \frac{x_t - \lambda_t x_0}{\delta_t^2} dx_0$$

$$= \frac{x_t}{\delta_t^2} - \frac{\lambda_t}{\delta_t^2} E_{q_{t0}(x_0 | x_t)} x_0$$

$$E_{q_{t0}(x_0 | x_t)} x_0 = \frac{x_t}{\lambda_t} - \frac{\delta_t^2}{\lambda_t} \nabla \log q_t(x_t)$$

Twisted 扩散采样

$$\begin{aligned}
 q(x_{0:T} | y) &= \frac{q(x_{0:T}, y)}{q(y)} = \frac{q(y | x_{0:T}) q(x_{0:T})}{q(y)} \\
 &= \frac{q(y | x_0) q_T(x_T) \prod_{t=0}^{T-1} q_{t+1,t}(x_t | x_{t+1})}{q(y)}
 \end{aligned}$$

↑
似然函数

如果有 x_T, x_{T-1}, \dots, x_0 , 满足这个条件分布,

那么 $x_0 \sim q_0(x_0 | y)$

重要性采样

$x_T \sim q_T(x_T)$

扩散
模型

$x_t | x_{t+1} \sim q_{t+1,t}(x_t | x_{t+1}) \quad t = T-1, T-2, \dots, 2, 1$

$q_{t+1,t}(x_t | x_{t+1}) = N(x_t | m_{t+1|t}, b_{t+1|t}^2)$

比如 OU 过程

$x_t = Y_{T-t}$

$dy_t = [y_t + 2s_0(t, y_t)]dt + \sqrt{2} dw_t$

+

$dx_t = -[x_t + 2s_0(t, x_t)]dt + \sqrt{2} dw_t \quad (dt = -1)$

权重

$m_{t+1|t} = x_{t+1} + (x_{t+1} + 2s_0(t, x_{t+1})) \quad b_{t+1|t}^2 = 2$

最后权重 $w_{0:T} \propto q(y | x_0)$

联合分布满足 $q(x_{0:T} | y)$

序贯蒙特卡洛

$$q(x_{0:T} | y) = q(x_T | y) \prod_{t=0}^{T-1} q(x_t | x_{t+1}, y)$$

$$q(x_T | y) \propto q_T(y | x_T) q_T(x_T)$$

$$q(x_t | x_{t+1}, y) = \frac{q_t(y | x_t, \cancel{x_{t+1}}) q_{t+1}(x_t | x_{t+1})}{q_{t+1}(y | x_{t+1})}$$

$$q_{t+1|t}(x | x_{t+1}) \approx \mathcal{N}(x_t | m_{t+1|t}, \delta_{t+1|t}^2)$$

比如 OU 过程

$$X_t = Y_{T-t}$$

$$dY_t = [Y_t + 2S_0(t, Y_t)]dt + \sqrt{2} dW_t$$

$$dX_t = -[X_t + 2S_0(t, X_t)]dt + \sqrt{2} dW_t \quad (dt = -1)$$

$$m_{t+1|t} = X_{t+1} + (X_{t+1} + 2S_0(t, X_{t+1})) \quad \delta_{t+1|t}^2 = 2$$

$$q_0(y | \hat{X}_0(t+1), x) \propto e^{-\frac{1}{2} \frac{\|x - m\|^2}{\delta_{t+1|t}^2}}$$

$$m = X_{t+1} + \nabla_x \log q_0(y | \hat{X}_0(t+1), X_{t+1}) \cdot \delta_{t+1|t}^2$$

$$m + m_{t+1|t} = X_{t+1} + 2(X_{t+1} + 2S_0(t, X_{t+1}) + \nabla_x \log q_0)$$

渐近精确

$$\begin{aligned} & q_T(x_T) \cdot q_0(y | \hat{x}_0(T, x_T)) \prod_{t=0}^{T-1} \tilde{q}_{t+1|t}(x_t | x_{t+1}, y) w_t \\ &= q_T(x_T) q_0(y | \hat{x}_0(T, x_T)) \prod_{t=0}^{T-1} \frac{q_0(y | \hat{x}_0(t, x_t)) q_{t+1|t}(x_t | x_{t+1})}{q_0(y | \hat{x}_0(t+1, x_{t+1}))} \\ &= q_T(x_T) \cdot \underbrace{q_0(y | \hat{x}_0(0, x_0))}_{q_0(y | x_0)} \prod_{t=0}^{T-1} q_{t+1|t}(x_t | x_{t+1}) \end{aligned}$$

SMC Diffusion Sampler

考虑方差爆炸 SDE

$$dX_t = \sigma dW_t$$

$$X_t = X_0 + \sqrt{\sigma^2 t} W$$

前向方程

$$\partial_t q_t(x) = \frac{\sigma^2}{2} \Delta q_t(x)$$

后向方程

$$\partial_t p_t(x) = -\frac{\sigma^2}{2} \Delta p_t(x) \quad p_t = q_{T-t}$$

粒子方程

$$dX_t = \sigma^2 \nabla_x \log p_t(x) dt + \sigma dW_t$$

$$= \sigma^2 \nabla_x \log q_{T-t}(x) dt + \sigma dW_t \quad (0 \rightarrow T)$$

$$dX_t = -\sigma^2 \nabla_x \log q_t(x) dt - \sigma dW_t \quad (T \rightarrow 0)$$

(dt = -1)

$$X_{t-1} = X_t + \sigma^2 \nabla_x \log q_t(X_t) + \sigma dW_t$$

$$\sim N(X_{t-1} | X_t + \sigma^2 \nabla_x \log q_t(X_t), \sigma^2)$$

Tweedie 公式

$$q_t(y|x) \approx q_0(y | \hat{X}_0(t, x))$$

$$E_{q_{t_0}(x_0|x)} x_0 = \hat{x}_0(t, x)$$

$$\begin{aligned} \nabla q_t(x) &= \nabla \int q(x_0|x) dx_0 \\ &= \int \nabla q_t(x|x_0) q(x_0) dx_0 \\ &= \int q_{t_0}(x|x_0) \frac{-(x-x_0)}{b^2 t} q(x_0) dx_0 \end{aligned}$$

$$\frac{\nabla q_t(x)}{q_t(x)} = \int q_{t_0}(x_0|x) \frac{x_0 - x}{b^2 t} dx_0$$

$$\hat{x}_0(t, x) = b^2 t \nabla \log q_t(x) + x$$

proposal

$$q_0(y | \hat{x}_0(t, x)) q_{t+t}(x | X_{t+t})$$

$$p_\eta(y - G(x - b^2 t S_0(t, x))) \mathcal{N}(x | b^2 S_0(t, X_{t+t}), b^2)$$

$$\approx e^{-\frac{1}{2} \frac{\|x - \tilde{m}\|^2}{\sigma^2}} \quad e^{-\frac{1}{2} \frac{\|x - \tilde{m}\|^2}{\sigma^2}}$$

$$\tilde{m} = b^2 S_0(t, X_{t+t})$$

$$m \approx X_{t+t} + b^2 \nabla_x \log p_\eta(X_{t+t})$$

$$m + \tilde{m} = X_{t+t} + b^2 (\nabla_x \log p_\eta(X_{t+t}) + S_0(t, X_{t+t}))$$

$$C = b^2$$