

Siddhant Mishra-Sharma (MIT/AI FI) Summer School

162

4

3

Demonstrating score matching

Need to compute the score $\nabla_{x_t} \log q(x_t)$

The *conditional score* $\nabla_{x_t} \log q(x_t | x)$ can be computed using the diffusion kernel

$$\nabla_{x_t} \log q(x_t | x) = - \frac{(x_t - x)}{\sigma_t^2} = - \frac{\epsilon}{\sigma_t}$$

Explain why x conditioning disappears

Noise-prediction

$$\frac{1}{2\sigma_q^2(t)} \frac{(1 - \alpha_t)^2}{(1 - \bar{\alpha}_t) \alpha_t} \left[\left\| \epsilon - \hat{\epsilon}_\theta(x_t, t) \right\|^2 \right]$$



Score-matching

$$\frac{1}{2\sigma_q^2(t)} \frac{(1 - \alpha_t)^2}{\alpha_t} \left[\left\| s_\theta(x_t, t) - \nabla \log p(x_t) \right\|^2 \right]$$

The noise- and score-prediction networks are equivalent up to a std-scaling

Denoising score matching

Need to compute the score $\nabla_{x_t} \log q(x_t)$

The *conditional* score $\nabla_{x_t} \log q(x_t | x)$ can be computed using the diffusion kernel

$$\nabla_{x_t} \log q(x_t | x) = -\frac{(x_t - x)}{\sigma_t^2} = -\frac{\epsilon}{\sigma_t}$$

Noise-prediction

$$\frac{1}{2\sigma_q^2(t)} \frac{(1 - \alpha_t)^2}{(1 - \bar{\alpha}_t) \alpha_t} \left[\left\| \epsilon - \hat{\epsilon}_\theta(x_t, t) \right\|^2 \right]$$



Score-matching

$$\frac{1}{2\sigma_q^2(t)} \frac{(1 - \alpha_t)^2}{\alpha_t} \left[\left\| s_\theta(x_t, t) - \nabla \log p(x_t) \right\|^2 \right]$$

The noise- and score-prediction networks are equivalent up to a std-scaling

The noise/score-prediction model and *latent diffusion*

[Rombach et al 2021]

