/70

45

Denoising score matching

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

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

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

Noise-prediction

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

$\Longleftrightarrow$

Score-matching

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

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

* x conditioning disappears when taking expectation wrt x to give marginal score

*

Denoising score matching

* x conditioning disappears when taking expectation wrt x to give marginal score

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

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

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

Noise-prediction

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

$\Longleftrightarrow$

Score-matching *

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

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]