Bài tập về nhà số 8

Nền tảng toán học của các mô hình tạo sinh – PIMA

Chủ đề: Mô hình khuếch tán

Người giải: Võ Hoàng Nhật Khang

1 Quá trình khuếch tán - Chuỗi Markov

1.1 Chứng minh $q(x_t|x_0)$ và cách lấy mẫu

Cho chuỗi Markov $x_0, x_1, \dots, x_T$ với

$$q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{\alpha_t}x_{t-1}, (1-\alpha_t)\mathbb{I}),$$

ta cần chứng minh

$$q(x_t|x_0) = \mathcal{N}(x_t; \sqrt{\bar{\alpha}_t}x_0, (1-\bar{\alpha}_t)\mathbb{I}),$$

với $\bar{\alpha}_t = \prod_{s=1}^t \alpha_s$.
Chứng minh:
Từ

$$x_t = \sqrt{\alpha_t}x_{t-1} + \sqrt{1-\alpha_t}\epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \mathbb{I}), \tag{1}$$

thay đệ quy

$$x_{t-1} = \sqrt{\alpha_{t-1}}x_{t-2} + \sqrt{1-\alpha_{t-1}}\epsilon_{t-1}, \quad \dots, \quad x_1 = \sqrt{\alpha_1}x_0 + \sqrt{1-\alpha_1}\epsilon_1. \tag{2}$$

Sau t bước

$$x_t = \sqrt{\bar{\alpha}_t}x_0 + \sum_{s=1}^t \sqrt{\bar{\alpha}_{s+1,t}(1-\alpha_s)}\epsilon_s, \tag{3}$$

với $\bar{\alpha}_{s+1,t} = \prod_{k=s+1}^t \alpha_k$ (bằng 1 nếu $s = t$).
Kỳ vọng

$$\mathbb{E}[x_t|x_0] = \sqrt{\bar{\alpha}_t}x_0. \tag{4}$$

Phương sai

$$\text{Var}(x_t|x_0) = \sum_{s=1}^t \bar{\alpha}_{s+1,t}(1-\alpha_s)\mathbb{I}. \tag{5}$$

Tính

$$\sum_{s=1}^{t} \bar{\alpha}_{s+1,t}(1-\alpha_s) = \sum_{s=1}^{t} \left(\prod_{k=s+1}^{t} \alpha_k \right) (1-\alpha_s) = 1 - \prod_{s=1}^{t} \alpha_s = 1 - \bar{\alpha}_t. \tag{6}$$

Vậy

$$q(x_t|x_0) = \mathcal{N}(x_t; \sqrt{\bar{\alpha}_t} x_0, (1-\bar{\alpha}_t)\mathbb{I}). \tag{7}$$

Cách lấy mẫu: Từ $q(x_t|x_0)$, lấy mẫu

$$x_t = \sqrt{\bar{\alpha}_t} x_0 + \sqrt{1-\bar{\alpha}_t}\epsilon, \quad \epsilon \sim \mathcal{N}(0, \mathbb{I}). \tag{8}$$

■

1.2 Phân phối đồng thời

Cho chuỗi Markov ngược $x_T, x_{T-1}, \ldots, x_0$, phân phối đồng thời là

$$p_\theta(x_0, x_1, \ldots, x_T) = p_\theta(x_T) \prod_{t=1}^{T} p_\theta(x_{t-1}|x_t). \tag{9}$$

Chứng minh:
Theo tính chất Markov

$$p_\theta(x_T, x_{T-1}, \ldots, x_0) = p_\theta(x_T) p_\theta(x_{T-1}|x_T) \cdots p_\theta(x_0|x_1). \tag{10}$$

Viết lại

$$p_\theta(x_0, x_1, \ldots, x_T) = p_\theta(x_T) \prod_{t=1}^{T} p_\theta(x_{t-1}|x_t) \textbf{ (Q.E.D)} \tag{11}$$

■

2 Suy diễn biến phân - Liên hợp phân phối chuẩn

Cho các phân phối

$$p(x) = \mathcal{N}(\mu_x, \sigma_x^2), \quad p(y \mid x) = \mathcal{N}(\mu_y x, \sigma_y^2), \tag{12}$$

với phân phối hậu nghiệm

$$p(x \mid y) = \mathcal{N}(\mu, \sigma^2), \tag{13}$$

$$\sigma^2 = \left(\frac{1}{\sigma_x^2} + \frac{\mu_y^2}{\sigma_y^2} \right)^{-1}, \quad \mu = \sigma^2 \left(\frac{\mu_x}{\sigma_x^2} + \frac{\mu_y y}{\sigma_y^2} \right). \tag{14}$$

Chứng minh rằng

$$q(x_{t-1} \mid x_t, x_0) = \mathcal{N}(\mu_t, \sigma_t^2 I), \tag{15}$$

với

$$\sigma_t^2 = \frac{(1-\alpha_t)(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}, \quad \mu_t = \sigma_t^2 \left(\frac{\sqrt{\alpha_t}}{1-\alpha_t} x_t + \frac{\sqrt{\alpha_{t-1}}}{1-\bar{\alpha}_{t-1}} x_0 \right), \tag{16}$$

biết

$$q(x_t \mid x_{t-1}) = \mathcal{N}(\sqrt{\alpha_t} x_{t-1}, (1-\alpha_t)I), \quad q(x_t \mid x_0) = \mathcal{N}(\sqrt{\bar{\alpha}_t} x_0, (1-\bar{\alpha}_t)I), \tag{17}$$

$$\bar{\alpha}_t = \prod_{s=1}^{t} \alpha_s. \tag{18}$$

Chứng minh:
Ta chứng minh $q(x_{t-1} \mid x_t, x_0)$ là phân phối chuẩn với các tham số như yêu cầu.
Dựa trên tính chất Markov và $q(x_t \mid x_{t-1}, x_0) = q(x_t \mid x_{t-1})$, ta có

$$q(x_{t-1} \mid x_t, x_0) \propto q(x_t \mid x_{t-1}) q(x_{t-1} \mid x_0). \tag{19}$$

Ánh xạ

- Tiên nghiệm: $q(x_{t-1} \mid x_0) = \mathcal{N}(\sqrt{\bar{\alpha}_{t-1}} x_0, (1-\bar{\alpha}_{t-1})I)$:

$$\mu_x = \sqrt{\bar{\alpha}_{t-1}} x_0, \quad \sigma_x^2 = 1 - \bar{\alpha}_{t-1}. \tag{20}$$

- Likelihood: $q(x_t \mid x_{t-1}) = \mathcal{N}(\sqrt{\alpha_t} x_{t-1}, (1-\alpha_t)I)$:

$$\mu_y = \sqrt{\alpha_t}, \quad \sigma_y^2 = 1 - \alpha_t, \quad y = x_t. \tag{21}$$

Tính σ_t^2

$$\frac{1}{\sigma_t^2} = \frac{1}{1-\bar{\alpha}_{t-1}} + \frac{\alpha_t}{1-\alpha_t} = \frac{(1-\alpha_t) + \alpha_t(1-\bar{\alpha}_{t-1})}{(1-\alpha_t)(1-\bar{\alpha}_{t-1})}. \tag{22}$$

Rút gọn

$$(1-\alpha_t) + \alpha_t(1-\bar{\alpha}_{t-1}) = 1 - \alpha_t \bar{\alpha}_{t-1} = 1 - \bar{\alpha}_t. \tag{23}$$

$$\sigma_t^2 = \frac{(1-\alpha_t)(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}. \tag{24}$$

Tính μ_t

$$\mu_t = \sigma_t^2 \left(\frac{\sqrt{\bar{\alpha}_{t-1}} x_0}{1-\bar{\alpha}_{t-1}} + \frac{\sqrt{\alpha_t} x_t}{1-\alpha_t} \right) = \sigma_t^2 \left(\frac{\sqrt{\alpha_t}}{1-\alpha_t} x_t + \frac{\sqrt{\bar{\alpha}_{t-1}}}{1-\bar{\alpha}_{t-1}} x_0 \right). \tag{25}$$

Vậy

$$q(x_{t-1} \mid x_t, x_0) = \mathcal{N}(\mu_t, \sigma_t^2 I), \tag{26}$$

với

$$\sigma_t^2 = \frac{(1-\alpha_t)(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}, \quad \mu_t = \sigma_t^2 \left(\frac{\sqrt{\alpha_t}}{1-\alpha_t} x_t + \frac{\sqrt{\bar{\alpha}_{t-1}}}{1-\bar{\alpha}_{t-1}} x_0 \right) \; \textbf{(Q.E.D)} \tag{27}$$

■

3 Phân kỳ Kullback-Leibler giữa 2 phân phối chuẩn

Cho $p(x) = \mathcal{N}(\mu_p, \sigma_p)$, $q(x) = \mathcal{N}(\mu_q, \sigma_q)$, chứng minh

$$D_{\mathsf{KL}}(p\|q) = \frac{1}{2}\left(\log\frac{\sigma_q^2}{\sigma_p^2} + \frac{\sigma_p^2}{\sigma_q^2} - 1\right) + \frac{1}{2\sigma_q^2}\|\mu_p - \mu_q\|^2 \tag{28}$$

Chứng minh:
Sử dụng

$$D_{\mathsf{KL}}(p\|q) = \frac{1}{2}\left(\log\frac{|\Sigma_q|}{|\Sigma_p|} - d + \mathrm{tr}(\Sigma_q^{-1}\Sigma_p) + (\mu_q - \mu_p)^\top\Sigma_q^{-1}(\mu_q - \mu_p)\right). \tag{29}$$

Thay $d = 1$, $\Sigma_p = \sigma_p^2$, $\Sigma_q = \sigma_q^2$

$$D_{\mathsf{KL}}(p\|q) = \frac{1}{2}\left(\log\frac{\sigma_q^2}{\sigma_p^2} - 1 + \frac{\sigma_p^2}{\sigma_q^2} + \frac{(\mu_q - \mu_p)^2}{\sigma_q^2}\right). \tag{30}$$

Rút gọn

$$D_{\mathsf{KL}}(p\|q) = \frac{1}{2}\left(\log\frac{\sigma_q^2}{\sigma_p^2} + \frac{\sigma_p^2}{\sigma_q^2} - 1\right) + \frac{1}{2\sigma_q^2}\|\mu_p - \mu_q\|^2 \textbf{ (Q.E.D)} \tag{31}$$

■