

1 Дано

1.

$$\nabla_{x_t} \log p(x_t|y) = \nabla_{x_t} \log p(x_t) + \nabla_{x_t} \log p(y|x_t) \quad (1)$$

2.

$$x_t = \sqrt{\bar{\alpha}_t} x_0 + \sqrt{1 - \bar{\alpha}_t} \epsilon, \quad \epsilon \sim \mathcal{N}(0, I) \quad (2)$$

3.

$$\nabla_{x_t} \log p(x_t) = -\frac{\epsilon_\theta(x_t, t)}{\sqrt{1 - \bar{\alpha}_t}} \quad (3)$$

4.

$$\boxed{\nabla_{x_t} \log p(y|x_t) = (1 - \sigma(f(x_t, t, y))) \cdot \nabla_{x_t} f(x_t, t, y)} \quad (4)$$

5.

$$\mu_\theta(x_t) = \frac{1}{\sqrt{\alpha_t}} \left(x_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} \epsilon_\theta(x_t, t) \right) \quad (5)$$

Нужно получить в общем виде

$$\tilde{\epsilon}(x_t, t, y) = \epsilon_\theta(x_t, t) - \lambda \cdot \sigma_t \cdot \nabla_{x_t} \log p(y|x_t) \quad (6)$$

2 Решение

Можно представить условный score:

$$\nabla_{x_t} \log p(x_t|y) = -\frac{\tilde{\epsilon}(x_t, t, y)}{\sqrt{1 - \bar{\alpha}_t}} \quad (7)$$

Подставляем (7) и (3) в (1):

$$-\frac{\tilde{\epsilon}(x_t, t, y)}{\sqrt{1 - \bar{\alpha}_t}} = -\frac{\epsilon_\theta(x_t, t)}{\sqrt{1 - \bar{\alpha}_t}} + \nabla_{x_t} \log p(y|x_t) \quad (8)$$

Умножаем на $-\sqrt{1 - \bar{\alpha}_t}$:

$$\boxed{\tilde{\epsilon}(x_t, t, y) = \epsilon_\theta(x_t, t) - \sqrt{1 - \bar{\alpha}_t} \cdot \nabla_{x_t} \log p(y|x_t)} \quad (9)$$

Подставляем (4) в (9):

$$\boxed{\tilde{\epsilon}(x_t, t, y) = \epsilon_\theta(x_t, t) - \sqrt{1 - \bar{\alpha}_t} \cdot (1 - p(y|x_t)) \cdot \nabla_{x_t} f(x_t, t, y)} \quad (10)$$

С коэффициентом guidance λ :

$$\boxed{\tilde{\epsilon}(x_t, t, y) = \epsilon_\theta(x_t, t) - \lambda \sqrt{1 - \bar{\alpha}_t} \cdot (1 - p(y|x_t)) \cdot \nabla_{x_t} f(x_t, t, y)} \quad (11)$$

3 Среднее

В DDPM обратный шаг (без условия):

$$\mu_\theta(x_t) = \frac{1}{\sqrt{\alpha_t}} \left(x_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} \epsilon_\theta(x_t, t) \right) \quad (12)$$

С условием заменяем ϵ_θ на $\tilde{\epsilon}$:

$$\tilde{\mu}(x_t, y) = \frac{1}{\sqrt{\alpha_t}} \left(x_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} \tilde{\epsilon}(x_t, t, y) \right) \quad (13)$$

Подставляем $\tilde{\epsilon}$ из (11):

$$\tilde{\mu}(x_t, y) = \frac{1}{\sqrt{\alpha_t}} \left(x_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} [\epsilon_\theta - \lambda \sqrt{1 - \bar{\alpha}_t} (1 - p(y|x_t)) \nabla_{x_t} f] \right) \quad (14)$$

Раскрываем:

$$\tilde{\mu}(x_t, y) = \frac{1}{\sqrt{\alpha_t}} \left(x_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} \epsilon_\theta \right) + \frac{\lambda(1 - \alpha_t)}{\sqrt{\alpha_t}} (1 - p(y|x_t)) \nabla_{x_t} f \quad (15)$$

Итоговая формула для условного среднего:

$$\tilde{\mu}(x_t, y) = \mu_\theta(x_t) + \frac{\lambda(1 - \alpha_t)}{\sqrt{\alpha_t}} (1 - p(y|x_t)) \nabla_{x_t} f(x_t, t, y)$$

(16)

4 Сэмплирование

Общая формула:

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(x_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} \tilde{\epsilon}(x_t, t, y) \right) + \sigma_t \xi, \quad \xi \sim \mathcal{N}(0, I) \quad (17)$$

Подставляем $\tilde{\epsilon}(x_t, t, y)$ из (11):

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(x_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} [\epsilon_\theta - \lambda \sqrt{1 - \bar{\alpha}_t} (1 - p(y|x_t)) \nabla_{x_t} f] \right) + \sigma_t \xi \quad (18)$$

Раскрываем скобки:

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(x_t - \underbrace{\frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} \epsilon_\theta}_{\mu_\theta(x_t)} + \lambda(1 - \alpha_t)(1 - p(y|x_t)) \nabla_{x_t} f \right) + \sigma_t \xi \quad (19)$$

Итог:

$$x_{t-1} = \underbrace{\frac{1}{\sqrt{\alpha_t}} \left(x_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} \epsilon_\theta(x_t, t) \right)}_{\mu_\theta(x_t)} + \frac{\lambda(1 - \alpha_t)}{\sqrt{\alpha_t}} (1 - p(y|x_t)) \nabla_{x_t} f + \sigma_t \xi$$

(20)