

# HW7\_report

110652012 施品光

## Part1

1. Score matching is a method to train a model to estimate the gradient of the log-probability density, called the *score function*:

$$s_{\theta}(x) \approx \nabla_x \log p_{\text{data}}(x)$$

Instead of modeling  $p(x)$  directly, score matching minimizes the difference between the model's output and the true score, without computing the intractable normalizing constant.

2. In score-based (diffusion) generative models, this idea is applied to noisy data  $x_t$ .

The model  $s_{\theta}(x_t, t)$  learns to predict the score  $\nabla_{x_t} \log p_t(x_t)$  at each noise level  $t$ .

During sampling, the model uses these estimated scores to iteratively denoise random noise and recover realistic data by integrating a reverse-time stochastic differential equation (SDE).

## Part2

How does sliced score matching differ mathematically and conceptually from standard score matching?