

Course: Machine Learning

Assignment: Week 7 _ Written Assignment

Student: Han Yun Chen (112652010)

1. Explain the concept of score matching and describe how it is used in diffusion models

Basically, score matching is a way to learn the data distribution without directly computing the normalization constant of $p(x)$, which is usually very hard.

Instead of learning $p(x)$ itself, we learn its gradient:

$$S(x) = \nabla_x \log p(x)$$

This score tells us which direction the data is more likely to appear.

In practice, we train a neural network $S_\theta(x)$ to approximate this gradient. Since the real score $\nabla_x \log p(x)$ is unknown, we use losses like implicit score matching (ISM) or denoising score matching (DSM).

DSM adds Gaussian noise to the data ($x = x_0 + \sigma\epsilon$) and trains the model to predict how to denoise it:

$$L_{DSM}(\theta) = \mathbb{E} \left[\frac{1}{\sigma^2} \|\sigma S_\theta(x_0 + \sigma\epsilon) + \epsilon\|^2 \right].$$

In diffusion models, this score function is used to reverse the noising process:

starting from pure noise, the model gradually denoises the sample using the predicted scores, turning noise into realistic images.

So, score matching is basically the training objective behind diffusion models.

2. Unanswered Questions

- Why do we train on multiple noise levels instead of one fixed σ ?
- How is DSM related to denoising autoencoders?