Detailed Discussions on Diffusion

1 Randomness in DDPM, SGM, and FM

1.1 What is the role of $\sigma_t z$ in the DDPM equation?

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(x_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}} \epsilon_\theta(x_t, t) \right) + \sigma_t z$$

Here, $\sigma_t z$ represents additional random noise, which plays a resampling role to ensure that the generated data retains a certain degree of randomness and is not entirely determined by the denoising model's prediction.

Intuitive explanation:

- Since the diffusion process introduces noise in the forward propagation, the reverse process must introduce some degree of randomness to ensure the plausibility of data recovery.
- Without this additional noise, the entire sampling process would become a deterministic denoising process, losing the stochastic nature of the diffusion model.

Mathematical explanation:

- $\sigma_t z$ is an extra noise term sampled from a Gaussian distribution, ensuring that some uncertainty remains in the sampling process: $z \sim \mathcal{N}(0, I)$.
- This term can be understood as preserving entropy, ensuring that information is not completely lost during the reverse diffusion process.

Impact:

- If $\sigma_t = 0$, the diffusion sampling process becomes a deterministic ODE sampling, similar to FM.
- In some improved diffusion methods (such as DDIM), σ_t can be adjusted to control the degree of determinism in sampling.

1.2 Why do DDPM and SGM require randomness?

DDPM and SGM employ stochastic diffusion models, where the path from a Gaussian distribution to the target distribution is inherently stochastic. This randomness arises due to the irreversible nature of the forward diffusion process:

- **DDPM forward diffusion**
  - Noise is incrementally added to the data x_0, eventually transforming x_T into pure Gaussian noise.
  - Since the forward process is stochastic (Markov process), the reverse process must incorporate randomness to match the real data distribution.
- **SGM (Score-Based Model) forward diffusion**
  - This is also a stochastic diffusion process, described by an SDE:

$$dx = f(x, t)dt + g(t)dW.$$

– The term $g(t)dW$ represents noise, ensuring that the data diffuses into a Gaussian distribution.

– During the reverse process, sampling follows the reverse SDE:

$$dx = [f(x,t) - g^2(t)\nabla_x \log q_t(x)]dt + g(t)d\bar{W}.$$

– Key points:

  * Since the forward diffusion process is stochastic, the reverse process must introduce additional randomness $g(t)d\bar{W}$ to ensure the generated data matches the true data distribution.
  * If $g(t)d\bar{W}$ is removed, SGM sampling reduces to an ODE (deterministic sampling).

Summary: DDPM and SGM require additional randomness because their forward diffusion process is inherently stochastic. Consequently, the reverse process must model this randomness.

1.3 Why does Flow Matching not require randomness?

The core idea of FM is to directly learn a continuous-time evolving vector field and generate data through deterministic ODE sampling.

SGM requires randomness, while FM does not, for the following reasons:

- **SGM (Score-Based Generative Model) uses SDE**
  - It requires score matching to learn the gradient of the data distribution.
  - Sampling uses SDE inversion, requiring randomness to explore different possibilities.
- **FM (Flow Matching) uses ODE**
  - It directly learns a vector field that maps noise to the real data distribution without relying on random noise.
  - The sampling is deterministic, following the same optimal trajectory each time.

In other words, FM finds the **"optimal path"** from a Gaussian distribution to the target data distribution, without requiring randomness to explore different paths.

1.4 Summary

Method	Does Forward Process Accumulate Noise?	Does Reverse Process Require Additional Randomness?	Reason
DDPM	Yes (noise added at each step)	Requires $\sigma_t z$	Reverse denoising needs to compensate for forward noise
SGM (SDE)	Yes (Wiener process diffusion)	Requires $g(t)d\bar{W}$	Reverse SDE needs to compensate for noise
FM	No noise accumulation	No additional randomness needed	Directly learns ODE, deterministic sampling

Table 1: Comparison of forward and reverse processes in different generative models

DDPM and SGM accumulate noise in their forward processes, so their reverse processes require additional randomness to correct the paths. FM does not accumulate noise in the forward process, so it does not require randomness in the reverse process and evolves deterministically via an ODE.

2 Why Can Flow Matching Still Generate Diverse Data?

Since FM sampling is deterministic, how can it still generate diverse data? The key lies in the diversity of the initial noise:

- FM still samples from Gaussian noise: Although FM sampling is deterministic, its input is a randomly sampled Gaussian noise $x_T \sim \mathcal{N}(0, I)$.
- Different initial noise samples x_T are mapped to different final samples x_0 through the ODE.
- The vector field $v_\theta(x, t)$ ensures that all initial points evolve into appropriate target data distributions.

This implies:

- The sampling process itself has no randomness (no $g(t)dW$), but the initial sample is random, ensuring diversity in the final output.
- Each x_T follows a fixed trajectory to x_0, but since x_T is random, the final generated data remains diverse.

Flow Matching decomposes the generative process into two parts:

- **Random initialization** (ensuring diversity) — Sampling the initial state from a Gaussian distribution.
- **Deterministic evolution** (ensuring data quality) — Using an ODE to learn the optimal transformation path.

This is similar to GANs: GANs also use a deterministic neural network (generator) to transform a random noise vector, but GAN training is unstable, whereas FM leverages ODEs for a smoother and more stable transformation process.