# Diffusion Tutorial

compiled by D.Gueorguiev, 6/17/2024

## Introductory Notes

The goal of generative modeling is: given iid samples from unknown distribution , construct a sampler for approximately the same distribution.

*Example*: given a training set of dog images from some underlying distribution , we want a method of producing new images of dogs from this distribution.

One way to solve this problem is to learn a transformation from some easy-to-sample distribution (such as Gaussian noise) to our target distribution . Diffusion models offer a general framework for learning such transformations. We want to reduce the problem of sampling from distribution to a sequence of easier sampling problems.

## Gaussian Diffusion

Let be a random variable in distributed according to the target distribution . Then construct a sequence of r.v.’s by successively adding independent Gaussian noise with some small scale :

, (1)

(1) represents forward process, which transforms the data distribution into a noise distribution. Thus (1) defines a joint distribution over , and we let denote the marginal distributions of each . Notice that at large step count , the distribution is nearly Gaussian, so we can approximate sample from by just sampling a Gaussian.

Now, suppose we can solve the following subproblem –

Given a sample marginally distributed as , produce a sample marginally distributed as .

We will call a method that does this a *reverse sampler*, since it tells us how to sample from assuming we can already sample from . If we had a reverse sampler, we could sample from our target by simply starting with a Gaussian sample from , and iteratively applying the reverse sampling procedure to get samples from , and finally .

The key insight of diffusion is, learning to reverse each step can be easier than learning to sample from target distribution in one step. There are many ways to construct reverse samplers, but for concreteness let us first see the standard diffusion sampler which we will call the *Denoising Diffusion Probabilistic Model* (DDPM) sampler.

The *ideal DDPM sampler* uses an obvious strategy: at time , given input (a sample from ), we output a sample from the conditional distribution

(2)

(2) represents a reverse sample. The problem is, it requires learning a generative model for the conditional distribution for every , which could be complicated. But if the per-step noise is sufficiently small, then it turns out this conditional distribution becomes simple:

Property of Diffusion Reverse Process

For small , and the Gaussian diffusion process defined in (1), the conditional distribution is itself close to Gaussian. That is, for all times and conditionings , there exists some mean parameter such that

(3)

For a given time and conditioning value , learning the mean of is sufficient to learn the full conditional distribution . This not obvious fact enables a drastic simplification – instead of having to learn an arbitrary distribution from scratch, we now know everything about this distribution except its mean, which we denote .

Learning the mean of is a much simpler problem than learning the conditional distribution itself; we can solve it by regression.

References

[1] [Introduction to Diffusion Models for Deep Learning, Ryan O'Connor, 2022 (online blog)](https://www.assemblyai.com/blog/diffusion-models-for-machine-learning-introduction/)

[2] [What are Diffusion Models? Lilian Weng, OpenAI, 2021 (online blog)](https://lilianweng.github.io/posts/2021-07-11-diffusion-models/)

[3] [Diffusion Models for Video Generation, Lilian Weng, OpenAI, 2024 (online blog)](https://lilianweng.github.io/posts/2024-04-12-diffusion-video/)

[4] [Step-By-Step Diffusion: An Elementary Tutorial, P. Nakkiran et al, 2024](https://github.com/dimitarpg13/deep_learning_for_image_processing/blob/main/literature/articles/generative_models/StepByStepDiffusionAnElementaryTutorial_Nakkiran_2024.pdf)

[5] [Tutorial on Diffusion Models for Imaging and Vision, Stanley Chan, 2024](https://github.com/dimitarpg13/deep_learning_for_image_processing/blob/main/literature/articles/generative_models/Tutorial_on_Diffusion_Models_for_Imaging_and_Vision_Chan_2024.pdf)

[6] [Understanding Diffusion Models: Unified Perspective, Calvin Luo, Google Brain, 2022](https://github.com/dimitarpg13/deep_learning_for_image_processing/blob/main/literature/articles/generative_models/Understanding_Diffusion_Models-A_Unified_Perspective_Luo_GoogleBrain_2022.pdf)

[7] [Lightweight Diffusion Models: A Survey, W. Song et al, 2024](https://github.com/dimitarpg13/deep_learning_for_image_processing/blob/main/literature/articles/generative_models/Lightweight_Diffusion_Models_A_Survey_Song_2024.pdf)

[8] [Deep Unsupervised Learning Using Nonequilibrium Thermodynamics, Jascha Sohl-Dickstein et al, Stanford U., 2015](https://github.com/dimitarpg13/deep_learning_for_image_processing/blob/main/literature/articles/generative_models/Deep_Unsupervised_Learning_using_Nonequilibrium_Thermodynamics_Sohl-Dickstein_2015.pdf)