Submission for Deep Learning Exercise

Team: shallow_learning_group
Students: Batuhan Karaca

January 16, 2024

Pen Paper Task: Diffusion Models

If we have

$$x_t = \sqrt{\alpha_t} x_{t-1} + \sqrt{1-\alpha_t}\epsilon_{t-1} \tag{1}$$

then

$$x_t = (\sqrt{\prod_{\tau=1}^{t} \alpha_\tau}) x_0 + \sum_{\tau=0}^{t-2} (\sqrt{\prod_{i=\tau+2}^{t} \alpha_i - \prod_{i=\tau+1}^{t} \alpha_i})\epsilon_\tau + (\sqrt{1-\alpha_t})\epsilon_{t-1} \tag{2}$$

When $t = 1$, it is easy to see equations (1) and (2) are the same. For an arbitrary t

$$\begin{aligned}
x_{t-1} &= (\sqrt{\prod_{\tau=1}^{t-1} \alpha_\tau}) x_0 + \sum_{\tau=0}^{t-3} (\sqrt{\prod_{i=\tau+2}^{t-1} \alpha_i - \prod_{i=\tau+1}^{t-1} \alpha_i})\epsilon_\tau + (\sqrt{1-\alpha_{t-1}})\epsilon_{t-2} \\
(\sqrt{\alpha_t}) x_{t-1} &= (\sqrt{\prod_{\tau=1}^{t} \alpha_\tau}) x_0 + \sum_{\tau=0}^{t-3} (\sqrt{\prod_{i=\tau+2}^{t} \alpha_i - \prod_{i=\tau+1}^{t} \alpha_i})\epsilon_\tau + (\sqrt{\alpha_t - \alpha_t\alpha_{t-1}})\epsilon_{t-2} \\
&= (\sqrt{\prod_{\tau=1}^{t} \alpha_\tau}) x_0 + \sum_{\tau=0}^{t-2} (\sqrt{\prod_{i=\tau+2}^{t} \alpha_i - \prod_{i=\tau+1}^{t} \alpha_i})\epsilon_\tau \\
(\sqrt{\alpha_t}) x_{t-1} + (\sqrt{1-\alpha_t})\epsilon_{t-1} &= (\sqrt{\prod_{\tau=1}^{t} \alpha_\tau}) x_0 + \sum_{\tau=0}^{t-2} (\sqrt{\prod_{i=\tau+2}^{t} \alpha_i - \prod_{i=\tau+1}^{t} \alpha_i})\epsilon_\tau + (\sqrt{1-\alpha_t})\epsilon_{t-1} = x_t
\end{aligned}$$

Therefore, by induction we have proven that the equation (2) is true.

We should have an equation of the form

$$x_t = \mu + \sigma \cdot \epsilon \tag{3}$$

We found that $\mu = (\sqrt{\prod_{\tau=1}^{t} \alpha_\tau}) x_0 = (\sqrt{\tilde{a}}) x_0$. By the property of *sum of normally distributed random variables* if

$$X_i \sim N(\mu_i, \sigma_i^2) \tag{4}$$

and

$$Y = \sum_{i=1}^{n} c_i X_i \tag{5}$$

then

$$Y \sim N(\sum_{i=1}^{n} c_i \mu_i, \sum_{i=1}^{n} c_i^2 \sigma_i^2) \tag{6}$$

We know $\epsilon \sim N(0, I)$. Substituting $X_i = \epsilon_i$,

$$\begin{aligned}\sum_{\tau=0}^{t-2}\Big(\prod_{i=\tau+2}^{t}\alpha_i - \prod_{i=\tau+1}^{t}\alpha_i\Big) + (1-\alpha_t) &= \prod_{i=2}^{t}\alpha_i - \prod_{i=1}^{t}\alpha_i + \prod_{i=3}^{t}\alpha_i - \prod_{i=2}^{t}\alpha_i + \ldots + \prod_{i=t-1}^{t}\alpha_i - \prod_{i=t-2}^{t}\alpha_i + \prod_{i=t}^{t}\alpha_i - \prod_{i=t-1}^{t}\alpha_i + (1-\alpha_i)\\ &= \prod_{i=2}^{t}\alpha_i - \prod_{i=1}^{t}\alpha_i + \prod_{i=3}^{t}\alpha_i - \prod_{i=2}^{t}\alpha_i + \ldots + \prod_{i=t-1}^{t}\alpha_i - \prod_{i=t-2}^{t}\alpha_i + \alpha_t - \prod_{i=t-1}^{t}\alpha_i + 1 - \alpha_i\end{aligned}$$

Note that similar elements cancel each other and we find $\sigma^2 = 1 - \prod_{i=1}^{t}\alpha_i = 1 - \tilde{a}$.