关于线性关系

$$x^{(t)}=\sqrt{\alpha_t}x^{(t-1)}+\sqrt{1-\alpha_t}z,\ z\sim\mathcal{N}(0,\mathbf{I})$$

$\sqrt{\alpha_t}$ 是对 $x^{(t-1)}$ 的一次线性变换，若把变换矩阵理解为 $\mathrm{H}=\sqrt{\alpha_t}\,\mathrm{I}$，因此这个迁移方程可以用如下线性逆变换描述：

$$y=\sqrt{\alpha_t}\,\mathrm{I}\,x+z,z\sim\mathcal{N}(0,(1-\alpha_t)\mathrm{I})$$

对于SBD使用的Score-Based方法是基于 $y=\mathrm{I}\,x+z$ 实现的，那么若要基于上述描述实现，只需要在涉及这个Score-Based方法前对目标 $x^{(t)}$ 进行一次关于 $\sqrt{\alpha_t}$ 的线性变换。

1. Diffusion过程推导

$\Rightarrow$ 后是修正了这个线性变换后的结果：

$$\begin{aligned}
x^{(t)}&=\sqrt{\alpha_t}x^{(t-1)}+\sqrt{1-\alpha_t}z_t\\
&=\sqrt{\alpha_t}(\sqrt{\alpha_{t-1}}x^{(t-2)}+\sqrt{1-\alpha_{t-1}}z_{t-1})+\sqrt{1-\alpha_t}z_t\\
&=\sqrt{\alpha_t\alpha_{t-1}}x^{(t-2)}+\sqrt{1-\alpha_t\alpha_{t-1}}\bar{z}_t\\
&=\cdots\\
&=\sqrt{\bar{\alpha}_t}x^{(0)}+\sqrt{1-\bar{\alpha}_t}z,\ z\sim\mathcal{N}(0,\mathbf{I})\\
\Rightarrow x_a^{(t)}&=\frac{x^{(t)}}{\sqrt{\alpha_t}}=x^{(t-1)}+\sqrt{\frac{1-\alpha_t}{\alpha_t}}z_t\\
&=\frac{x^{(t)}}{\sqrt{\bar{\alpha}_t}}=x^{(0)}+\sqrt{\frac{1-\bar{\alpha}_t}{\bar{\alpha}_t}}z,\ z\sim\mathcal{N}(0,\mathbf{I})
\end{aligned}$$

基于这个修正结果 $x_a^{(t)}$，使用 $x_a^{(t)}$ 和 $x^{(0)}$ 训练Score-Based计算梯度理论上会比原先的 $x^{(t)}$ 和 $x^{(0)}$ 更加准确。

2. Sampling过程推导

> 这个定义是第一版公式的，当时并没有细想这块内容。现在看这公式，发现它本身也暗示了这个变换关系，但是我当时实现时并没有做相关思考。

原先的采样公式如下所示：

$$\begin{aligned}
x^{(t)}&=\sqrt{\alpha_t}x^{(t-1)}+\sqrt{1-\alpha_t}z,\ z\sim\mathcal{N}(0,\mathbf{I})\\
\Rightarrow x^{(t-1)}&=\frac{x^{(t)}-\sqrt{1-\alpha_t}z}{\sqrt{\alpha_t}},\ z=\frac{x^{(t)}-\sqrt{\alpha_t}x^{(t-1)}}{\sqrt{1-\alpha_t}}
\end{aligned}$$

原先定义 Score-based 分布梯度：

$$\nabla_x log[q_\theta(x^{(t-1)}|x^{(t)})] \approx -z = -\frac{x^{(t)} - \sqrt{\alpha_t}x^{(t-1)}}{\sqrt{1-\alpha_t}}$$

> 新定义

引入这个变换修正:

$$x_a^{(t)} = \frac{x^{(t)}}{\sqrt{\alpha_t}} = x^{(t-1)} + \sqrt{\frac{1-\alpha_t}{\alpha_t}}z$$

$$\Rightarrow x^{(t-1)} = x_a^{(t)} - \sqrt{\frac{1-\alpha_t}{\alpha_t}}z,\ z = \sqrt{\frac{\alpha_t}{1-\alpha_t}}(x_a^{(t)} - x^{(t-1)})$$

重定义 Score-based 分布梯度:

$$\nabla_x log[q_\theta(x^{(t-1)}|x_a^{(t)})] \approx -z = \sqrt{\frac{\alpha_t}{1-\alpha_t}}(x^{(t-1)} - x_a^{(t)})$$

3. 综上

Diffusion部分并不用进行任何变换，只需要在所有Score-Based模块前对 $x^{(t)}$ 进行一次线性变换得到 $x_a^{(t)}$ 即可。代码修改量很小，下周会对这个改动进行验证。