Диплом

Колтаков Михаил

Научный руководитель: Дахова Елизавета

3 февраля 2024 г.

Введём модель:

$$\begin{array}{ccc} X_{01}^1 & \dots & X_{01}^\tau \\ \dots & \dots & \dots \\ X_{0n}^1 & \dots & X_{0n}^\tau \end{array}$$

Где вдоль вертикальной оси меняются элементы батча, а вдоль горизонтальной — кадры. При этом для одного элемента батча можно ввести следующие латентные величины:

$$\begin{array}{ccccccc} X_0^1 & \to & X_1^1 & \to & \dots & \to & X_T^1 \\ \downarrow & & \downarrow & & & & \downarrow \\ X_0^2 & \to & X_1^2 & \to & \dots & \to & X_T^2 \\ \downarrow & & \downarrow & \ddots & & & \downarrow \\ \vdots & & \vdots & & \ddots & & \vdots \\ \downarrow & & \downarrow & & & \ddots & \downarrow \\ X_0^\tau & \to & X_1^\tau & \to & \dots & \to & X_T^\tau \end{array}$$

Объединим столбики в одну величину и дальше величины, обозначающие полное видео будем обозначать с волной: $\widetilde{X}$. Тогда у нас получается марковская цепь $\widetilde{X}_0 \to \widetilde{X}_1 \to \ldots \to \widetilde{X}_T$, для которой можно ввести следующее вариационное семейство:

$$q\left(\widetilde{X}_1, \ldots, \widetilde{X}_T \,|\, \widetilde{X}_0\right) = \prod_{t=1}^{T} q_t\left(\widetilde{X}_t \,|\, \widetilde{X}_{t-1}\right)$$

В данной работе будем считать плотность между двумя соседними шагами зашумления следующей:

$$\widetilde{X}_t \,|\, \widetilde{X}_{t-1} \overset{Q}{\sim} \mathcal{N}\left(\sqrt{1-\beta_t}\widetilde{X}_{t-1};\ \beta_t C\right), \text{ где } C = \begin{pmatrix} a & b & c & \dots \\ b & a & b & \dots \\ c & b & a & \dots \\ \vdots & & & \end{pmatrix} \text{ и полож. полуопред.}$$

Из этого по индукции получается следующее:

Обозн. $\widetilde{\alpha}_t = \prod_{s=1}^{t} \alpha_s, \quad \alpha_s = 1 - \beta_s$

База. $\widetilde{X}_1 \,|\, \widetilde{X}_0 \overset{Q}{\sim} \mathcal{N}\left(\sqrt{1-\beta_t}\widetilde{X}_0;\ \beta_t C\right) = \mathcal{N}\left(\sqrt{\widetilde{\alpha}_1}\widetilde{X}_0;\ (1-\widetilde{\alpha}_1)\, C\right)$ по утверждению выше.

Шаг. $\widetilde{X}_t \,|\, \widetilde{X}_{t-1} \overset{Q}{\sim} \mathcal{N}\left(\sqrt{1-\beta_t}\widetilde{X}_{t-1};\ \beta_t C\right) = \sqrt{1-\beta_t}\widetilde{X}_{t-1} + \sqrt{\beta_t}\varepsilon_t =$
$= \sqrt{1-\beta_t}\sqrt{\widetilde{\alpha}_{t-1}}\widetilde{X}_0 + \sqrt{1-\beta_t}\sqrt{1-\widetilde{\alpha}_{t-1}}\varepsilon'_{t-1} + \sqrt{\beta_t}\varepsilon_t = \sqrt{\widetilde{\alpha}_t}\widetilde{X}_0 + \sqrt{1-\widetilde{\alpha}_{t-1}\,(1-\beta_t)}\varepsilon'_t =$
$= \mathcal{N}\left(\sqrt{\alpha}_t\widetilde{X}_0; (1-\widetilde{\alpha}_t)\, C\right)$. Предпоследний переход выполнен используя свойства сложения дисперсий суммы нормальных распределений. Все ε и ε' имеют распределение $\mathcal{N}\left(\vec{0};\, C\right)$.

В итоге индукция доказывает верность для любого t следующего выражения:

$$\widetilde{X}_t \,|\, \widetilde{X}_0 \sim \mathcal{N}\left(\sqrt{\widetilde{\alpha}_t}\widetilde{X}_0;\ (1-\widetilde{\alpha}_t)\,C\right)$$

А значит, можно провести такую репараметризацию, выделив среднее как постоянную компоненту и оставить случайность в отдельном слагаемом с нулевым средним:

$$\widetilde{X}_t \overset{dQ}{=\!=} \sqrt{\widetilde{\alpha}_t}\widetilde{X}_0 + \sqrt{1-\widetilde{\alpha}_t}\delta_t \Rightarrow \widetilde{X}_0 \overset{dQ}{=\!=} \frac{1}{\sqrt{\widetilde{\alpha}_t}}\left(\widetilde{X}_t - \sqrt{1-\widetilde{\alpha}_t}\delta_t\right) \quad \delta_t \sim \mathcal{N}\left(\vec{0};\ C\right)$$

Теперь перейдём к непрерывному случаю. В нём уравнение можно записать и преобразовать следующим образом:

$$\widetilde{X}_{t+dt} = \sqrt{1-\beta_t}\widetilde{X}_t + \sqrt{\beta_t}\varepsilon_t \quad \varepsilon_t \sim \mathcal{N}\left(\vec{0};\ C\right)$$

$$\widetilde{X}_{t+dt} = \left(1 - \frac{1}{2}\beta_t dt + \overline{\overline{o}}\,(dt)\right)\widetilde{X}_t + \sqrt{\beta_t dt}\varepsilon_t$$

$$d\widetilde{X}_t = -\frac{1}{2}\beta_t dt + \sqrt{\beta_t dt}\varepsilon_t$$

Так как матрица ковариаций ε_t не диагональна, оно не может быть приращением винеровского процесса, поэтому это стохастическое дифференциальное уравнение решать будет неудобно. Сведём его к винеровскому процессу. Так как матрица C положительно определена и диагональна, у неё существует спектральное разложение, такое, что матрица из собственных векторов ортогональна, то есть, $C = VSV^T$, где матрица S диагональна и все её значения неотрицательны в следствие положительной полуопределённости матрицы C. Значит, у матрицы S легко берётся корень, и можно ввести матрицу $V' = V\sqrt{S}$. Тогда

$$\sqrt{dt}\varepsilon_t = V'\varepsilon'_t \quad \varepsilon'_t \sim \mathcal{N}\left(\vec{0};\ dtI_\tau\right)$$

Значит, теперь можно переписать наше стох. дифференциальное уравнение через многомерный винеровский процесс:

$$d\widetilde{X}_t = -\frac{1}{2}\beta_t dt + \sqrt{\beta_t}V' d\vec{W}_t$$

Тогда можно воспользоваться формулой Фоккера–Планка, и получить последовательностью преобразований обратное СДУ и эквивалентное ОДУ.

$$\text{Прямое СДУ: } d\widetilde{X}_t = -\frac{1}{2}\beta_t dt + \sqrt{\beta_t}V' d\vec{W}_t$$

$$\text{Обратное СДУ: } d\widetilde{X}_t = \beta_t\left(-\frac{1}{2}\widetilde{X}_t + C\nabla_x \ln p_t\,(x)\Big|_{x=\widetilde{X}_t}\right)dt + \sqrt{\beta_t}d\vec{W}_t$$

$$\text{Эквивалентное ОДУ: } d\widetilde{X}_t = -\frac{1}{2}\beta_t\left(\widetilde{X}_t + C\nabla_x \ln p_t\,(x)\Big|_{x=\widetilde{X}_t}\right)dt$$