

My solutions to  
Deep Learning: Foundations and Concepts

Dario Miro Konopatzki

## 20 Diffusion Models

### 20.1

Mean

$$\begin{aligned}\mathbb{E}[Z_t] &= \mathbb{E}\left[\sqrt{1-\beta_t}Z_{t-1} + \sqrt{\beta_t}\mathcal{E}_t\right] \\ &= \sqrt{1-\beta_t}\mathbb{E}[Z_{t-1}] + \sqrt{\beta_t}\mathbb{E}[\mathcal{E}_t] \quad \text{linearity of } \mathbb{E} \\ &= \sqrt{1-\beta_t}\mathbb{E}[Z_{t-1}] \quad \mathcal{E}_t \sim \mathcal{N}(0, 1) \text{ so i.p. } \mathbb{E}[\mathcal{E}_t] = 0\end{aligned}$$

$$\begin{aligned}\|\mathbb{E}[Z_t]\| &= \left\|\sqrt{1-\beta_t}\mathbb{E}[Z_{t-1}]\right\| \\ &= \left|\sqrt{1-\beta_t}\right| \|\mathbb{E}[Z_{t-1}]\| \\ &< \|\mathbb{E}[Z_{t-1}]\| \quad \left|\sqrt{1-\beta_t}\right| < 1 \text{ since } 0 < \beta_t < 1\end{aligned}$$

Auxiliary Calculations

$$\begin{aligned}\mathbb{E}[\mathcal{E}_t\mathcal{E}_t^\top] &= \mathbb{E}[\mathcal{E}_t\mathcal{E}_t^\top] - \mathbb{V}[\mathcal{E}_t] + \mathbb{V}[\mathcal{E}_t] \\ &= \mathbb{E}[\mathcal{E}_t\mathcal{E}_t^\top] - (\mathbb{E}[\mathcal{E}_t\mathcal{E}_t^\top] - \mathbb{E}[\mathcal{E}_t]\mathbb{E}[\mathcal{E}_t]^\top) + \mathbb{V}[\mathcal{E}_t] \\ &= \mathbb{E}[\mathcal{E}_t\mathcal{E}_t^\top] - \mathbb{E}[\mathcal{E}_t\mathcal{E}_t^\top] + \mathbb{E}[\mathcal{E}_t]\mathbb{E}[\mathcal{E}_t]^\top + \mathbb{V}[\mathcal{E}_t] \\ &= \mathbb{E}[\mathcal{E}_t]\mathbb{E}[\mathcal{E}_t]^\top + \mathbb{V}[\mathcal{E}_t] \\ &= \mathbb{I} \quad \mathbb{E}[\mathcal{E}_t] = 0, \mathbb{V}[\mathcal{E}_t] = \mathbb{I} \text{ since by assumption } \mathcal{E}_t \sim \mathcal{N}(0, \mathbb{I})\end{aligned}$$

$$\begin{aligned}\mathbb{E}[Z_t Z_t^\top] &= \mathbb{E}\left[\left(\sqrt{1-\beta_t}Z_{t-1} + \sqrt{\beta_t}\mathcal{E}_t\right)\left(\sqrt{1-\beta_t}Z_{t-1} + \sqrt{\beta_t}\mathcal{E}_t\right)^\top\right] \\ &= \mathbb{E}\left[(1-\beta_t)Z_{t-1}Z_{t-1}^\top + \sqrt{1-\beta_t}\sqrt{\beta_t}Z_{t-1}\mathcal{E}_t^\top\right]\end{aligned}$$

$$\begin{aligned}
& + \sqrt{\beta_t} \sqrt{1 - \beta_t} \mathcal{E}_t Z_{t-1}^\top + \beta_t \mathcal{E}_t \mathcal{E}_t^\top \Big] \\
& = (1 - \beta_t) \mathbb{E} [Z_{t-1} Z_{t-1}^\top] + \sqrt{1 - \beta_t} \sqrt{\beta_t} \mathbb{E} [Z_{t-1} \mathcal{E}_t^\top] \\
& \quad + \sqrt{\beta_t} \sqrt{1 - \beta_t} \mathbb{E} [\mathcal{E}_t Z_{t-1}^\top] + \beta_t \mathbb{E} [\mathcal{E}_t \mathcal{E}_t^\top] \quad \mathbb{E} \text{ linear} \\
& = (1 - \beta_t) \mathbb{E} [Z_{t-1} Z_{t-1}^\top] + \sqrt{1 - \beta_t} \sqrt{\beta_t} \mathbb{E} [Z_{t-1}] \mathbb{E} [\mathcal{E}_t^\top] \\
& \quad + \sqrt{\beta_t} \sqrt{1 - \beta_t} \mathbb{E} [\mathcal{E}_t] \mathbb{E} [Z_{t-1}^\top] + \beta_t \mathbb{E} [\mathcal{E}_t \mathcal{E}_t^\top] \quad Z_{t-1} \perp \mathcal{E}_t \\
& \stackrel{(\star)}{=} (1 - \beta_t) \mathbb{E} [Z_{t-1} Z_{t-1}^\top] + \beta_t \mathbb{E} [\mathcal{E}_t \mathcal{E}_t^\top] \\
& = (1 - \beta_t) \mathbb{E} [Z_{t-1} Z_{t-1}^\top] + \beta_t \mathbb{I}
\end{aligned}$$

( $\star$ )  $\mathcal{E}_t \sim \mathcal{N}(0, \mathbb{I})$ , i.p.  $\mathbb{E}[\mathcal{E}_t] = 0$

### Covariance

$$\begin{aligned}
\|\text{cov}(Z_t) - \mathbb{I}\| & = \|\mathbb{E} [Z_t Z_t^\top] - \mathbb{E}[Z_t] \mathbb{E}[Z_t]^\top - \mathbb{I}\| \\
& = \|(1 - \beta_t) \mathbb{E} [Z_{t-1} Z_{t-1}^\top] + \beta_t \mathbb{I} - (1 - \beta_t) \mathbb{E}[Z_{t-1}] \mathbb{E}[Z_{t-1}]^\top - \mathbb{I}\| \\
& = \|(1 - \beta_t) (\mathbb{E} [Z_{t-1} Z_{t-1}^\top] - \mathbb{E}[Z_{t-1}] \mathbb{E}[Z_{t-1}]^\top - \mathbb{I})\| \\
& = \|(1 - \beta_t) (\text{cov}(Z_{t-1}) - \mathbb{I})\| \\
& = |1 - \beta_t| \|\text{cov}(Z_{t-1}) - \mathbb{I}\| \\
& < \|\text{cov}(Z_{t-1}) - \mathbb{I}\| \quad |1 - \beta_t| < 1 \text{ since } 0 < \beta_t < 1
\end{aligned}$$