

北京邮电大学 2023—2024 学年第一学期

《神经网络与深度学习》课程实验作业（五）

实验内容：生成模型基础

注意事项：

- ① 本次实验包含两道题，共计 50 分；
- ② 所有实验结果需以实验报告 word 文档的形式进行提交，文件命名格式：实验五_姓名_学号.docx，文件中需要将作者设置为本人姓名；
- ③ 实验报告中可插入代码片段，完整代码无需放在实验报告中，和训练好的模型一起以压缩包的形式提交即可，压缩包命名格式：实验五代码_姓名_学号.zip；
- ④ GAN 有关实验附件，需以压缩包形式提交，压缩包命名格式：实验五支撑材料_姓名_学号.zip；
- ⑤ 作业提交截止时间：2023 年 1 月 15 日中午 12:00。

一、变分自编码器（VAE）（25 分）

ELBO 由两部分组成，即重建错误（RE）：

$$RE \stackrel{\text{def}}{=} \mathbb{E}_{\mathbf{x} \sim p_{data}(\mathbf{x})} [\mathbb{E}_{q_{\phi}(\mathbf{z}|\mathbf{x})} [\ln p_{\theta}(\mathbf{x}|\mathbf{z})]], \quad (1)$$

以及编码器与先验之间的正则项：

$$\Omega \stackrel{\text{def}}{=} \mathbb{E}_{\mathbf{x} \sim p_{data}(\mathbf{x})} [\mathbb{E}_{q_{\phi}(\mathbf{z}|\mathbf{x})} [\ln p_{\lambda}(\mathbf{z}) - \ln q_{\phi}(\mathbf{z}|\mathbf{x})]]. \quad (2)$$

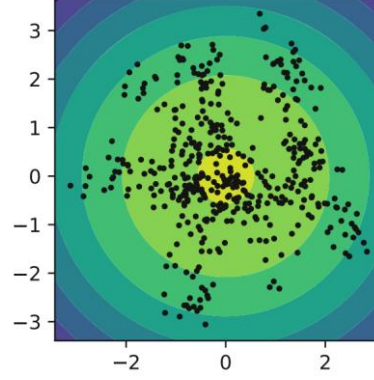
若对正则项进行进一步的变换，最终可以将正则项化简为：

$$\Omega = -\mathbb{CE}[q_{\phi}(\mathbf{z})||p_{\lambda}(\mathbf{z})] + \mathbb{H}[q_{\phi}(\mathbf{z}|\mathbf{x})], \quad (3)$$

其中第一项 $\mathbb{CE}[q_{\phi}(\mathbf{z})||p_{\lambda}(\mathbf{z})]$ 为混合后验和先验之间的交叉熵，第二项 $\mathbb{H}[q_{\phi}(\mathbf{z}|\mathbf{x})]$ 为 $q_{\phi}(\mathbf{z}|\mathbf{x})$ 与经验分布 $p_{data}(\mathbf{x})$ 的条件熵。

上式中，第一项实则蕴含了我们的目标，因为交叉熵会迫使聚合后验去匹配先验。据此，我们可以发现，你所选取的先验分布将会影响模型的性能。

请结合上述内容，基于 MNIST 手写数字数据集，实现不同先验分布下的 VAE。实验时，隐变量维度需设置为 2，并在不同先验下绘制聚合后验样本的可视化图像。实验预期结果如下所示，其中等高线得自标准高斯先验，黑点为采自聚合后验的样本。



- (1) 标准高斯: $p_{\lambda}(\mathbf{z}) = \mathcal{N}(\mathbf{z} | 0, \mathbf{I})$;
- (2) 混合高斯: $p_{\lambda}(\mathbf{z}) = \sum_{k=1}^K w_k \mathcal{N}(\mathbf{z} | \mu_k, \sigma_k^2)$, 其中 $\lambda = \{\{w_k\}, \{\mu_k\}, \{\sigma_k^2\}\}$ 是可训练的参数;
- (3) 其余先验 (自行确定)。

二、生成对抗网络 (GAN) (25 分)

在实验三中，我们在 Dog 数据集上对自编码器结构进行了实验。请基于该数据集，尝试自己搭建 GAN 架构，完成生成狗的生成任务。具体任务要求如下：

- (1) 当网络迭代次数为 10、30、50 时，随机生成 10 张图像，以展示生成图像质量随着迭代次数增加的提升；
- (2) FID 为评价生成图像质量的一种常用指标，值越小，则代表图像多样性越好，质量越好。FID 的计算公式为：

$$FID(x, g) = \|\mu_x - \mu_g\|^2 + Tr(\Sigma_x + \Sigma_g - 2\sqrt{\Sigma_x \Sigma_g}), \quad (4)$$

其中 Tr 表示矩阵对角线上元素的综合 (即矩阵的迹)， x 和 g 表示真实图片和生成

图片， μ 表示均值， σ 表示协方差矩阵。请根据该公式，编程实现 FID 函数；

- (3) 利用你训练好的模型，生成 300 张图片，并测试这些图片的 FID。