作业4

1 用于推荐的潜在特征（100 分）

注意：请使用原生的 Python 库，如 numpy、pandas 等来解决这个问题。运行时间通常需要几分钟，但具体时间可能取决于您使用的系统。

这个问题的目标是实现随机梯度下降算法，用于构建一个潜在因子推荐系统。我们可以使用它向用户推荐电影。

假设我们有一个评分矩阵 R。该矩阵的元素 R_{iu} 对应于用户 u 对项目 i 的评分。矩阵 R 的大小为 $m \times n$，其中 m 是电影的数量，n 是用户的数量。

矩阵的大多数元素是未知的，因为每个用户只能对少数电影进行评分。

我们的目标是找到两个矩阵 P 和 Q，使得 $R \simeq QP^T$。矩阵 Q 的维度是 $m \times k$，矩阵 P 的维度是 $n \times k$。这里的 k 是算法的一个参数。

我们将误差定义为

$$E = \left(\sum_{(i,u) \in rating} (R_{iu} - q_i \cdot p_u^T)^2 \right) + \lambda \left[\sum_u \|p_u\|_2^2 + \sum_i \|q_i\|_2^2 \right]. \qquad (1)$$

在这里，$\sum_{(i,u) \in rating}$ 表示我们仅对用户对该项进行了评分的 (用户, 项目) 对进行求和，即矩阵 R 的 (i, u) 元素是已知的。q_i 表示矩阵 Q 的第 i 行（对应于一个项目），p_u 表示矩阵 P 的第 u 行（对应于一个用户 u）。q_i 和 p_u 都是大小为 k 的行向量。λ 是正则化参数。$\|\cdot\|_2$ 是 L_2 范数，$\|p_u\|_2^2$ 是 L_2 范数的平方，即 p_u 的元素平方和。

(a) [40 分]

用 ϵ_{iu} 表示误差 E 相对于 R_{iu} 的导数。ϵ_{iu} 的表达式是什么？在随机梯度下降算法中，q_i 和 p_u 的更新方程是什么？请给出你的推导，并在最终的 q_i 和 p_u 表达式中使用 ϵ_{iu}。

(b) [60 分]

实现该算法。从磁盘读取矩阵 R 的每个元素，并为每个元素更新 ϵ_{iu}、q_i 和 p_u。

强调一下，不允许将矩阵 R 存储在内存中。你必须逐个从磁盘读取每个元素 R_{iu}，并在每次迭代中应用更新方程（对每个元素都进行更新）。算法的每次迭代都会读取整个文件。

选择 $k = 20$、$\lambda = 0.1$ 和 迭代次数 $= 40$。从 $\eta = 0.1$ 开始，找到一个良好的学习率 η 的值（不允许修改 k 或 λ）。对于下面讨论的训练集 ratings.train.txt，在 40 次迭代后，误差 E 应该小于 65000；你应该观察到 q_i 和 p_u 均不再变化。

根据 η 的值，可能会遇到以下情况：

- 如果 η 太大，误差函数可能会收敛到一个较高的值，或者可能不会单调递减。它甚至可能发散，使向量 p 和 q 的分量等于 ∞。
- 如果 η 太小，误差函数没有足够的时间显著减小并达到收敛。因此，它可能会单调递减但不收敛，即在 40 次迭代后可能具有较高的值，因为它尚未收敛。

使用压缩包中的 data 中的数据集来解决这个问题。它包含以下文件：

- ratings.train.txt：这是矩阵 R。每个条目由用户 ID、电影 ID 和评分组成。

绘制目标函数 E（在式(1)中定义）在训练集上随迭代次数变化的值。您找到的 η 是多少？

你可以使用任何编程语言来实现这一部分，但建议使用 Python 以提高速度。Python 应该能够在几分钟内获得一个可运行的解决方案。

提示：如果您对算法的某些步骤不确定如何进行，以下提示可能会帮助您，如果您有其他方法，不一定需要遵循这些提示。

- P 和 Q 的初始化：我们希望对于所有用户 u 和物品 i，都有 q_i 和 p_u，使得 $q_i \cdot p_u^T \in [0, 5]$。实现这一目标的一个好方法是将 P 和 Q 的所有元素初始化为 $[0, \sqrt{5/k}]$ 范围内的随机值。
- 更新方程：在每次更新中，我们使用 p_u 来更新 q_i，并使用 q_i 来更新 p_u。使用旧值计算 q_i 和 p_u 的新值，然后更新向量 q_i 和 p_u。
- 你应该在完成一次完整的训练迭代之后，计算 E。在迭代过程中计算 E 是不正确的，因为 P 和 Q 仍然在更新中。

需要提交的内容:

(1) ϵ_{iu} 的方程。在随机梯度下降算法中的更新方程 [(a)]

(2) η 的值。绘制 E 与迭代次数的折线图。确保你的图表有 y 轴，以便我们能够读取 E 的值。只需要使用你选择的 η 绘制一张图即可 [(b)]

(3) 请将所有代码上传到 gitee[(b)]