

(a) [40 分]

用  $\epsilon_{iu}$  表示误差  $E$  相对于  $R_{iu}$  的导数。 $\epsilon_{iu}$  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  $\epsilon_{iu}$ 。

$$(a) \cdot \frac{\partial E}{\partial R_{iu}} = 2(R_{iu} - q_i \cdot p_u^T) \cdot$$

$$\Rightarrow \epsilon_{iu} = 2(R_{iu} - q_i \cdot p_u^T)$$

$$\frac{\partial E}{\partial q_i} = -2 \sum_{(i,u) \in \text{Rating}} (R_{iu} - q_i \cdot p_u^T) \cdot p_u + 2\lambda q_i \quad i \text{ 固定.}$$

$$\frac{\partial E}{\partial p_u} = -2 \sum_{(i,u) \in \text{Rating}} (R_{iu} - q_i \cdot p_u^T) q_i + 2\lambda p_u \quad u \text{ 固定.}$$

对于 SGD，每次更新只使用样本  $(i, u)$  的信息

$$\Rightarrow \nabla(q_i) = -2(R_{iu} - q_i \cdot p_u^T) \cdot p_u + 2\lambda q_i$$

$$= -\epsilon_{iu} \cdot p_u + 2\lambda q_i$$

$$\nabla(p_u) = -2(R_{iu} - q_i \cdot p_u^T) \cdot q_i + 2\lambda p_u$$

$$= -\epsilon_{iu} \cdot q_i + 2\lambda p_u$$

$$\Rightarrow q_i \leftarrow q_i + \eta \cdot (\epsilon_{iu} \cdot p_u - 2\lambda q_i) \quad \text{其中 } \eta \text{ 代表学习率.}$$

$$p_u \leftarrow p_u + \eta \cdot (\epsilon_{iu} \cdot q_i - 2\lambda p_u)$$