Q2.1 - Computing gradients

$$e_{ij}^2 = \left(r_{ij} - \sum_k p_{ik} q_{kj} \right)^2$$

Therefore,

$$e_{ij}^2 = \left(r_{ij} - \hat{r}_{ij} \right)^2$$

The gradient of pik and qkj is needed to minimise the mean squared error.The partial derivative with respect to pik and qkj gives the needed change.

- equation 1 (with respect to pik)

$$\Rightarrow \frac{\partial}{\partial p_{ik}} e_{ij}^2 \quad = -2 \left(r_{ij} - \sum_k p_{ik} q_{kj} \right) (q_{kj}) = -2 \left(r_{ij} - \hat{r}_{ij} \right) (q_{kj}) = -2 e_{ij} q_{kj}$$

- equation 2 (with respect to pik)

$$\Rightarrow \frac{\partial}{\partial q_{ik}} e_{ij}^2 \quad = -2 \left(r_{ij} - \sum_k p_{ik} q_{kj} \right) (p_{ik}) = -2 \left(r_{ij} - \hat{r}_{ij} \right) (p_{ik}) = -2 e_{ij} p_{ik}$$

Update Values:

$$\begin{aligned} \Rightarrow p'_{ik} &= p_{ik} + \alpha \frac{\partial}{\partial p_{ik}} e_{ij}^2 = p_{ik} + 2\alpha e_{ij} q_{kj} \\ \Rightarrow q'_{kj} &= q_{kj} + \alpha \frac{\partial}{\partial q_{kj}} e_{ij}^2 = q_{kj} + 2\alpha e_{ij} p_{ik} \end{aligned}$$

Q2.2 - Adding biases

$$\hat{r}ij = bu_i + bi_j + \sum_k pikq_{kj}$$

From 2.1, we know

$$e_{ij}^2 = (r_{ij} - \hat{r}_{ij})^2$$

Substitution w.r.t equation 1 and 2 :

$$e_{ij}^2 = \left(r_{ij} - \left(bu_i + bi_j + \sum_k p_{ik} q_{kj} \right) \right)^2$$

$$\Rightarrow \frac{\partial}{\partial bu_i} e_{ij}^2 = -2 \left(r_{ij} - \left(bu_i + bi_j + \sum_k p_{ik} q_{kj} \right) \right) = -2 \left(r_{ij} - \hat{r}ij \right) = -2eij$$

$$\Rightarrow \frac{\partial}{\partial bi_j} e_{ij}^2 = -2 \left(r_{ij} - \left(bu_i + bi_j + \sum_k p_{ik} q_{kj} \right) \right) = -2 \left(r_{ij} - \hat{r}ij \right) = -2eij$$

Update Values:

$$\Rightarrow b'u_i \quad = bu_i + \alpha \frac{\partial}{\partial bu_i} e_{ij}^2 = bu_i + 2\alpha e_{ij}$$

$$\Rightarrow b'i_j \quad = bi_j + \alpha \frac{\partial}{\partial bi_j} e_{ij}^2 = bi_j + 2\alpha e_{ij}$$