AI assignment 3

Huda Feroz Ahmed (ha04081)

November 2019

Q1.1 – Computing gradients

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

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

To minimize the mean squared error, we need to know that the direction of maximum change for the values of p_{ik} and q_{kj} i.e. we need to know the gradient.
The required change can be found by taking the partial derivative w.r.t p_{ik} and q_{kj}.

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

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

To update values of p_{ik} and q_{kj}, we move in the found direction.
Updated values can be written as:

$$\text{updated = new - (learning rate} \times \text{direction)}$$

$$p'_{ik} = p_{ik} + \alpha \frac{\partial}{\partial p_{ik}} e_{ij}^2 = p_{ik} + 2\alpha e_{ij} q_{kj}$$

$$q'_{kj} = q_{kj} + \alpha \frac{\partial}{\partial q_{kj}} e_{ij}^2 = q_{kj} + 2\alpha e_{ij} p_{ik}$$

Q1.2 - Adding biases

$$\hat{r}_{ij} = bu_i + bi_j + \sum_k p_{ik} q_{kj} \tag{1}$$

We know that

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

Substituting equation 1 in equation 2:

$$e_{ij}^2 = (r_{ij} - (bu_i + bi_j + \sum_k p_{ik} q_{kj}))^2 \tag{3}$$

$$\frac{\partial}{\partial bu_i} e_{ij}^2 = -2(r_{ij} - (bu_i + bi_j + \sum_k p_{ik} q_{kj})) = -2(r_{ij} - \hat{r}_{ij}) = -2e_{ij} \tag{4}$$

$$\frac{\partial}{\partial bi_j} e_{ij}^2 = -2(r_{ij} - (bu_i + bi_j + \sum_k p_{ik} q_{kj})) = -2(r_{ij} - \hat{r}_{ij}) = -2e_{ij} \tag{5}$$

To update values of bu_i and bi_j, we move in the found direction.
Updated values can be written as:

updated = new - (learning rate $\times$ direction)

$$b'u_i = bu_i + \alpha \frac{\partial}{\partial bu_i} e_{ij}^2 = bu_i + 2\alpha e_{ij} \tag{6}$$

$$b'i_j = bi_j + \alpha \frac{\partial}{\partial bi_j} e_{ij}^2 = bi_j + 2\alpha e_{ij} \tag{7}$$