GD for Lin Reg

$y_i = ax_i + b \quad i = 1, \ldots, n$

$E(a,b) = \sum_i (y_i - ax_i - b)^2$

$a \leftarrow a - \eta \, \partial E / \partial a$

$$\frac{\partial E}{\partial a} = \frac{\partial}{\partial a} \sum_i (y_i - ax_i - b)^2$$

$$= -2 \sum_i (y_i - ax_i - b) x_i$$

$a \leftarrow a + 2\eta \sum_i (y_i - ax_i - b) x_i$

$b \leftarrow b - \eta \, \partial E / \partial b$

$$\frac{\partial E}{\partial b} = \frac{\partial}{\partial b} \sum_i (y_i - ax_i - b)^2$$

$$= -2 \sum_i (y_i - ax_i - b)$$

$b \leftarrow b + 2\eta \sum_i (y_i - ax_i - b)$

Matrix Factorization Model:

$U = n_u \times K$
$V = n_m \times K \quad V^T = K \times n_m$

$$\hat{Y} = U V^T$$

$$\begin{bmatrix} \cdots \hat{y}_{ij} \cdots \end{bmatrix} = \begin{bmatrix} - u_i - \end{bmatrix} \begin{bmatrix} | \\ v_j \\ | \end{bmatrix}$$

$$\hat{y}_{ij} = u_i \cdot v_j = (UV^T)_{ij} = \sum_s u_{is} v_{js}$$

Mask Matrix:

$$R_{ij} = \begin{cases} 0 & \text{if } y_{ij} \text{ was blank} \\ 1 & \text{if not} \end{cases}$$

$$N = \sum_{i,j} R_{ij}$$

Loss: $$E(U,V) = \frac{1}{N} \sum_{(i,j): R_{ij}=1} (y_{ij} - u_i \cdot v_j)^2$$

Grad Descent:

$$w \leftarrow w - \eta \, \partial E / \partial w$$

Will need to update the elements of u_i & v_j:

$$u_i = (u_{i1}, u_{i2}, \dots, u_{ik}, \dots, u_{iK})$$

$$v_j = (v_{j1}, v_{j2}, \dots, v_{jk}, \dots, v_{jK})$$

i & l are user indizes

Updating eq:

$$\frac{\partial E}{\partial u_{lk}} = \frac{1}{N} \sum_{(i,j): R_{ij}=1} \frac{\partial}{\partial u_{lk}} (y_{ij} - \hat{y}_{ij})^2$$

$$= \frac{1}{N} \cdot 2 \sum_{(\dots)} (y_{ij} - \hat{y}_{ij}) \left(\frac{\partial \hat{y}_{ij}}{\partial u_{lk}} \right)$$

Claim:

$$\frac{\partial \hat{y}_{ij}}{\partial u_{lk}} = \begin{cases} 0 & \text{if } l \neq i \\ v_{jk} & \text{if } l = i \end{cases}$$

EX: $K=2$

$$\hat{y}_{34} = u_{31}v_{41} + u_{32}v_{42}$$

If $l=5$, $k=1$, then

$$\frac{\partial \hat{y}_{34}}{\partial u_{51}} = \frac{\partial}{\partial u_{51}}(u_{31}v_{41} + u_{32}v_{42}) = 0$$

If $l=3$, then:

$$\frac{\partial \hat{y}_{34}}{\partial u_{31}} = \frac{\partial}{\partial u_{31}}(u_{31}v_{41} + u_{32}v_{42}) = v_{41}$$

Updating eq:

$$\frac{\partial E}{\partial u_{lk}} = \frac{1}{N} \sum_{(i,j):R_{ij}=1} \frac{\partial}{\partial u_{lk}} (y_{ij} - \hat{y}_{ij})^2$$

$$= \frac{-1}{N} \cdot 2 \sum_{(\ldots)} (y_{ij} - \hat{y}_{ij}) \left(\frac{\partial \hat{y}_{ij}}{\partial u_{lk}} \right)$$

$$= \frac{-2}{N} \sum_{j:R_{lj}=1} (y_{ij} - \hat{y}_{ij})(v_{jk})$$

$$u_{lk} \leftarrow u_{lk} - \eta \, \partial E / \partial u_{lk}$$

$$u_{lk} \leftarrow u_{lk} - \eta \left(\frac{-2}{N} \sum_{j:R_{lj}=1} (y_{lj} - \hat{y}_{lj}) v_{jk} \right)$$

$$u_{lk} \leftarrow u_{lk} + \frac{2\eta}{N} \sum_{(j:R_{lj}=1)} (y_{lj} - \hat{y}_{lj}) v_{jk}$$

A similar derivation for v_{lk} gives

$$v_{lk} \leftarrow v_{lk} + \frac{2\eta}{N} \sum_{(i:R_{il}=1)} (y_{lj} - \hat{y}_{lj}) u_{ik}$$

Vectorizing the GD eq.:

element-wise multiplication

Define $\Delta = (Y - UV^T) \otimes R$

$\Delta_{ij} = (y_{ij} - u_i \cdot v_j) \cdot r_{ij}$

Define the matrices:

$\frac{\partial E}{\partial U} :=$ the matrix w/ elements $\partial E / \partial u_{ik}$

$= \left[\frac{\partial E}{\partial u_{ik}} \right]_{i,k}$

$\frac{\partial E}{\partial V} = \left[\frac{\partial E}{\partial v_{jk}} \right]_{j,k}$

Claim:

$\frac{\partial E}{\partial U} = -\frac{2}{N} \Delta V$ &

$\frac{\partial E}{\partial V} = -\frac{2}{N} \Delta^T U$.

Why? Let's go element-by-element.

$$\left(\frac{\partial E}{\partial u}\right)_{ik} = \frac{\partial E}{\partial u_{ik}} = -\frac{2}{N}\sum_j \Delta_{ij} v_{jk}$$

$$= -\frac{2}{N}\sum_j (y_{ij} - u_i v_j) \cdot r_{ij} v_{jk}$$

$$= -\frac{2}{N}\sum_{j:r_{ij}=1} (y_{ij} - u_i v_j) v_{jk}$$