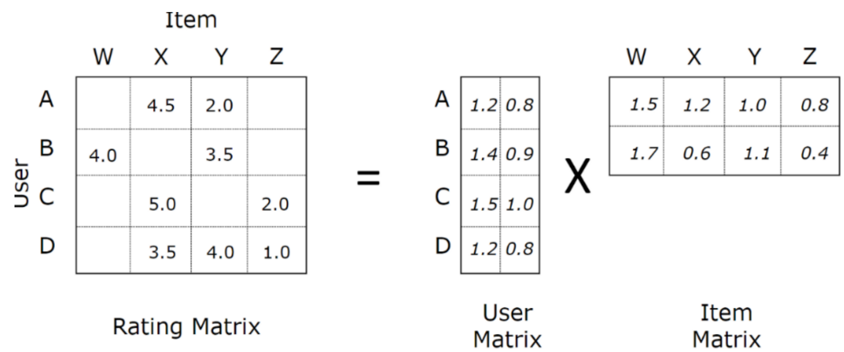

$$X_{N \times M} \approx U_{N \times K} \times V_{K \times M}^{T}$$

$$x_{n,m} = u_n^T v_m = \sum_{k=1}^{K} u_{n,k} v_{m,k}$$

$$L = \sum_{(m,n) \in \Omega} \left(x_{n,m} - u_n^T v_m \right)^2$$

$$L^{reg} = \sum_{(m,n) \in \Omega} \left(x_{n,m} - u_n^T v_m \right)^2 + \sum_{n=1}^{N} \lambda_u \|u_n\|_2^2 + \sum_{m=1}^{M} \lambda_v \|v_m\|_2^2$$

$$w = \left(X^T X + \lambda I_{d+1} \right)^{-1} X^T y$$

$$\underset{v_m}{arg\,min} \sum_{n \in \Omega_{c_m}} \left(x_{n,m} - u_n^T v_m \right)^2 + \lambda_v \|v_m\|_2^2$$

$v_m^{\text{¿}} = \left(U^T U + \lambda_v I_K \right)^{-1} U^T x_m$ Dimensions are given below,

$$[K \times 1] = \left([K \times N] \times [N \times K] + [K \times K] \right) \times [K \times N] \times [N \times 1]$$

But we cannot take all of $U^T U$ as there are missing entries.

$$[K \times 1] = \left(\left[K \times |\Omega_{cm}| \right] \times \left[|\Omega_{cm}| \times K \right] + [K \times K] \right) \times \left[K \times |\Omega_{cm}| \right] \times \left[|\Omega_{cm}| \times 1 \right]$$

This is equivalent to,

$$v_m^{\text{¿}} = \left(\sum_{n \in \Omega_{cm}} u_n u_n^T + \lambda_v I_K \right)^{-1} \sum_{n \in \Omega_{cm}} x_{n,m} u_n$$

$$[K \times 1] = \left(\sum_{n \in \Omega_{cm}} [K \times 1] \times [1 \times K] + [K \times K] \right) \sum_{n \in \Omega_{cm}} x_{n,m} [K \times 1]$$

Similarly,

$$\underset{u_n}{arg\,min} \sum_{m \in \Omega_{r_n}} \left(x_{n,m} - u_n^T v_m \right)^2 + \lambda_u \| u_n \|_2^2$$

$$u_n^{¿} = \left(V^T V + \lambda_u I_K \right)^{-1} V^T x_n$$

$$u_n^{¿} = \left(\sum_{m \in \Omega_{rn}} v_m v_m^T + \lambda_u I_K \right)^{-1} \sum_{m \in \Omega_{rn}} x_{n,m} v_m$$

Dyadic data, each

$$L_{reg} = \| y - Xw \| + \lambda \| w \|_2^2$$

$$\frac{\partial L_{reg}}{\partial w} = -2 \sum_{i=1}^{n} \left(y_i - x_i^T w \right) x_i + 2 \lambda w$$

2 will be absorbed into η

If no analytical solution
Gradient descent

1. Start with an initial value of $w = w^{(0)}$
2. Update w by moving along the gradient of the loss function L
   $$w^{(t)} = w^{(t-1)} - \eta \frac{\partial L}{\partial w}$$
3. Repeat until converge