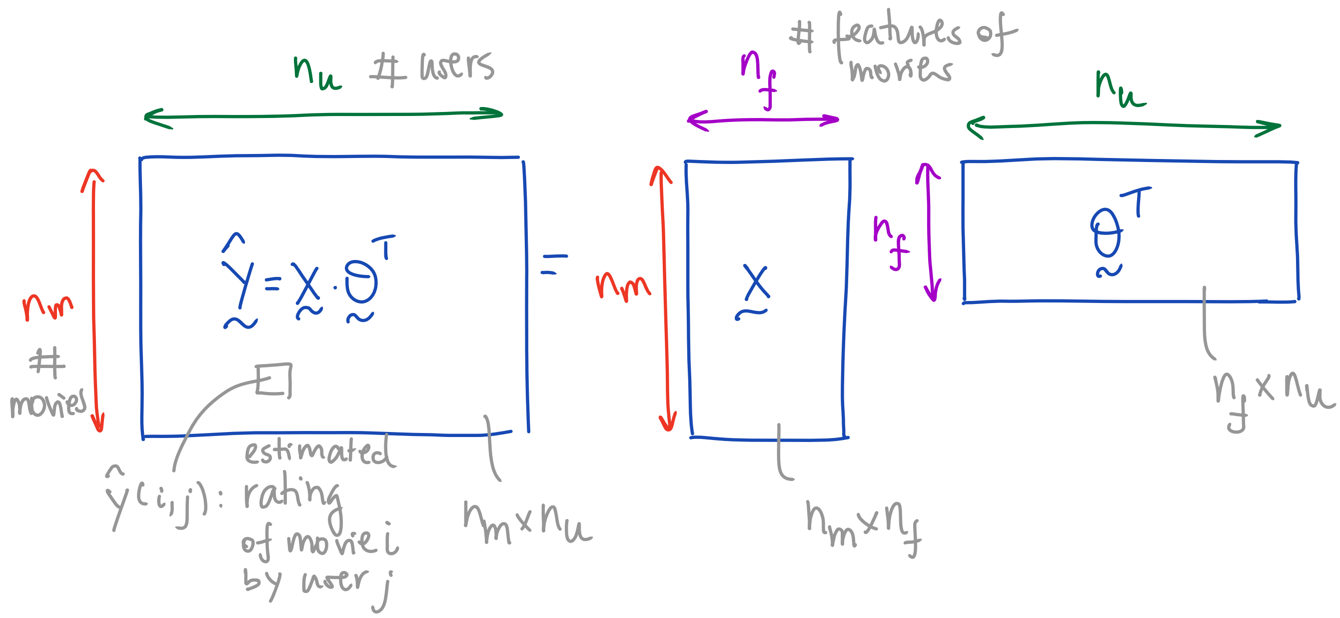
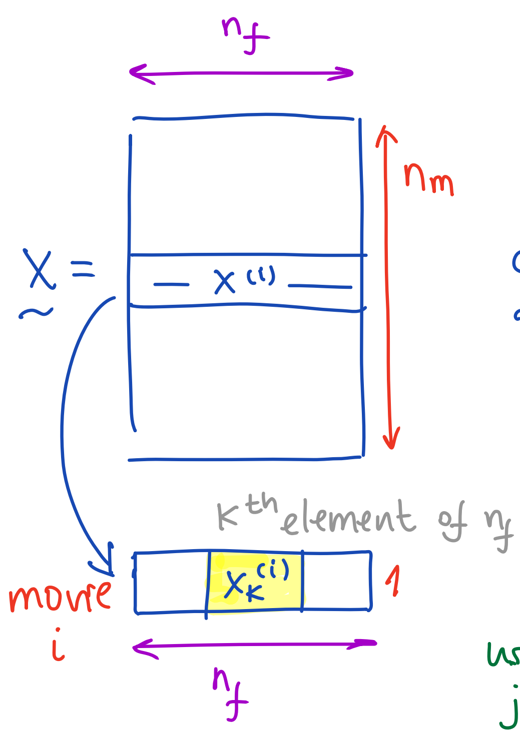
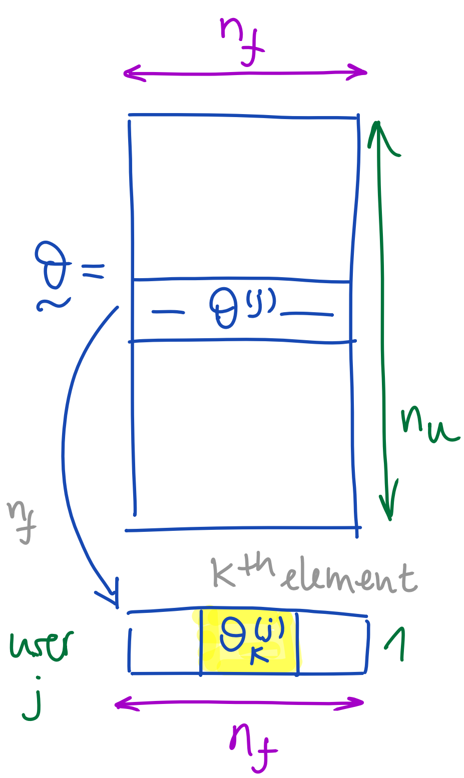
Recommender Systems – Low Rank Matrix Factorization for Collaborative Learning





Values of each of the n_f features for each movie i



Degree of affinity of each movie feature (n_f) given by each user j

Exercise sizes

$n_m = 1682$ movies

$n_u = 943$ users

$n_f = 100$ features for each movie

$n_f < n_m, n_u$, I understand

Learnt automatically.

- romance
- modern
- sci-fi
- action
- blockbuster
- ...

Goal: find $\underset{\sim}{X}$ & $\underset{\sim}{\theta}$ so that the difference between $\underset{\sim}{Y}$ and $\underset{\sim}{\hat{Y}} = \underset{\sim}{X} \cdot \underset{\sim}{\theta}^T$ is minimum, being $\underset{\sim}{X}$ & $\underset{\sim}{\theta}$ the ones defined above.

Cost and Gradient (unregularized)

Even though the cost $J(\underset{\sim}{X}, \underset{\sim}{\theta})$ is quite straightforward to compute vectorized, the gradient $\underset{\sim}{\nabla J}$ is a little bit tricky.

$$J(\underset{\sim}{X}, \underset{\sim}{\theta}) = J(x^{(1)}, \dots, x^{(n_m)}, \theta^{(1)}, \dots \theta^{(n_u)}) =$$

$$= \frac{1}{2} \sum_{(i,j):r(i,j)=1} \Big(\underbrace{(\theta^{(j)})^T x^{(i)}}_{\text{column vectors, not as defined above!}} - y^{(i,j)} \Big)^2$$

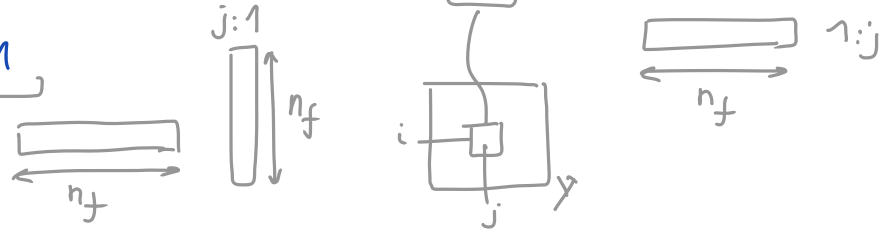
$\frac{\partial J}{\partial \theta}$

$$= \frac{1}{2} \sum_{(i,j):r(i,j)=1} \big(\underbrace{\underset{\sim}{X} \cdot \theta^T - Y}_{\text{matrices, as defined above}} \big)^2$$

$\frac{\partial J}{\partial x}$

$$\left(\frac{\partial J}{\partial x^{(i)}} \right)^T = \left[\frac{\partial J}{\partial x_1^{(i)}}, \dots, \frac{\partial J}{\partial x_k^{(i)}}, \dots, \frac{\partial J}{\partial x_{n_f}^{(i)}} \right] =$$

$$= \underbrace{\sum_{j:r(i,j)=1}}_{\text{Sum over all } j \text{ values, only take values when } r(i,j)=1} \Big(\overbrace{x^{(i)} \cdot (\theta^{(j)})^T}^{\text{These have always values}} - \overbrace{y^{(i,j)}}^{\text{This has values iff } r(i,j)=1} \Big) \cdot \theta^{(j)}$$



Sum over all j values, only take values when $r(i,j) = 1$

These have always values

This has values iff $r(i,j)=1$