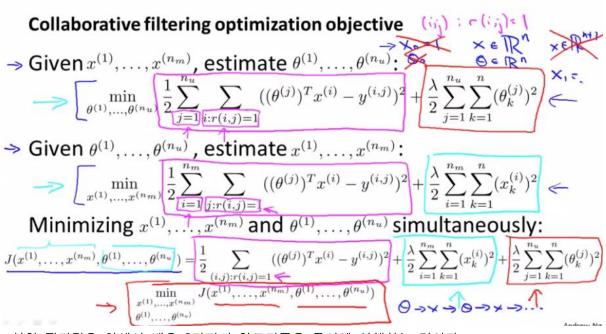
Coursera 16-4 Collaborative-filtering Algorithm

https://www.coursera.org/learn/machine-learning/lecture/f26nH/collaborative-filtering-algorithm



- 협업 필터링은 앞에서 배운 2가지의 알고리즘을 동시에 실행하는 것이다.
- 협업 필터링의 첫 번째 항은 평점을 매긴 모든 사용자에 대한 X(i), 세타에 대한 것 이다. 나머 지 두 개의 항은 세타와 X(i)의 정규화 항이다.
- 동시에 학습을 진행하고 있기 때문에 x0(절편값)이 필요하지 않다. x $\subseteq R^n$, heta $\subseteq R^n$

- Collaborative filtering algorithm \uparrow 1. Initialize $x^{(1)}, \dots, x^{(n_m)}, \theta^{(1)}, \dots, \theta^{(n_u)}$ to small random values.
- \rightarrow 2. Minimize $J(x^{(1)}, \dots, x^{(n_m)}, \theta^{(1)}, \dots, \theta^{(n_u)})$ using gradient descent (or an advanced optimization algorithm). E.g. for

$$x_{k}^{(i)} := x_{k}^{(i)} - \alpha \left(\sum_{j:r(i,j)=1} ((\theta^{(j)})^{T} x^{(i)} - y^{(i,j)}) \theta_{k}^{(j)} + \lambda x_{k}^{(i)} \right)$$

$$\theta_{k}^{(j)} := \theta_{\underline{k}}^{(j)} - \alpha \left(\sum_{i:r(i,j)=1} ((\theta^{(j)})^{T} x^{(i)} - y^{(i,j)}) x_{k}^{(i)} + \lambda \theta_{k}^{(j)} \right)$$

$$\frac{\partial}{\partial x_{k}^{(i)}}$$

$$\frac{\partial}{\partial x_{k}^{(i)}}$$

3. For a user with parameters θ and a movie with (learned) features \underline{x} , predict a star rating of $\underline{\theta^T x}$.

- small random values 초기값으로 학습을 진행하는 것이 neural network와
- x0가 없기 때문에 k=0인 특수케이스의 Gradient descent모델이 존재하지 않는다.

In the algorithm we described, we initialized $x^{(1)},\ldots,x^{(n_m)}$ and $\theta^{(1)},\ldots,\theta^{(n_u)}$ to small random values. Why is this?

- This step is optional. Initializing to all 0's would work just as well.
- Random initialization is always necessary when using gradient descent on any problem.
- ullet This ensures that $x^{(i)}
 eq heta^{(j)}$ for any i,j.
- ullet This serves as symmetry breaking (similar to the random initialization of a neural network's parameters) and ensures the algorithm learns features $x^{(1)},\dots,x^{(n_m)}$ that are different from each other.

단어공부

Sequentially: 순차적으로 Convention: 관습,관례,대회