|  |
| --- |
| **Rating-based Recommender System** |
| **Problem**: Predict movie rating based on previous rated movie of that user    **Notation**: |
|  |
| **Content-based recommender system** |
| If x0 = 1  x = m(movies) \* n(content)  For each user j, learn theta || X \* theta = ratings of the m movies  For each theta (n \* #user) matrix  **Linear regression of content-based recommender system**    **Gradient Descent** |

**Collaborative Filtering**

|  |
| --- |
| **Goal** |
| Compute the feature of the movie given the rating and parameter theta |
|  |
| Fit (personal likes) to get (ratings) from (content)  **content(X) \* personal likes(theta) = ratings(y)**  Fit (content) to get (ratings) from (personal likes)  **personal likes(X) \* content(theta)=ratings(y)**  size(X) = #user \* (n+1) where x0 = 1 |
|  |
|  |
| Get which first? Personal likes? Content? |
| Ask user at the beginning of what they like, get an initial personal likes  Repeat{  Use initial personal likes to predict features  Use features to update personal likes  } |
|  |
| **Algorithm** |
| **Cost function**  **Goal**: minimizing x and theta simultaneously    **CFA (collaborative filtering algorithm)**  Initialize x and theta to small random values  Minimize J(theta, x), use gradient descent  For i from 1-#movie, j from 1-#user    **Regularization term no longer eliminate theta0 or x0** |

**Vectorization of collaborative filtering: Low rank matrix factorization**

|  |
| --- |
| **Organizing data**    **Predicted ratings**    X (#movie \* #content) \* theta(#content \* #user)  **Find related movies** |
|  |
| **Implementation detail: Mean normalization** |
| **Significance:**  Used to avoid all ratings of new users to be predicted as 0    When new user is added, collaborative filtering system will first predict all of his/her rating as 0,  **With mean normalization**   1. Compute average rating of each movie in mu, then subtract all past ratings from mu.      1. Pretend the modified Y as the data for actual ratings, to use to learn theta and x for the new user 2. Add back the mu     If j’s user is a new user (haven’t rated any movies), then it’s rating will be mu  Else, it also works on single movie that has no rating |
|  |
|  |