Movie recommender system

Amarnath Mahadevuni

Objective:

Given a set of 1682 movies and 943 users partial ratings of those movies, predict how a user would rate a movie they have not watched yet.

- Define user preference vector : Theta
- Define movie feature vector : X
- Define cost function using existing ratings of movies by users.
- Calculate Theta and X to minimize error in rating

User Preference Vector

- Each user has specific characteristics they like in a movie.
- Represent these preferences as a 100 dimensional vector Theta.
- Eg. I like adventure and comedy, so my preference vector would have high values for adventure and comedy.

Movie feature Vector

- Each movie has certain characteristics of its own. Eg. Star wars is rich in adventure and drama.
- Represent these features as a 100 dimensional vector X

Collaborative-filtering

Both X and Theta are unknown.

So use iterative method - Collaborative filtering.

Initialize X and Theta randomly and define the cost function as :

J = Sum(Square(Theta * Transpose(X) - actual_rating)) over all user ratings.

Gradients for X and theta at each iteration:

dX =(Theta* Transpose(X)-actual_rating)*Theta over all ratings

dTheta =(Theta* Transpose(X)-actual_rating)*X over all ratings

Update X and Theta

$X = X - learning\_rate * dX$

$Theta = Theta - learning\_rate * dTheta$