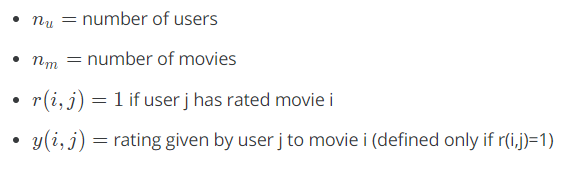
Recommender Systems

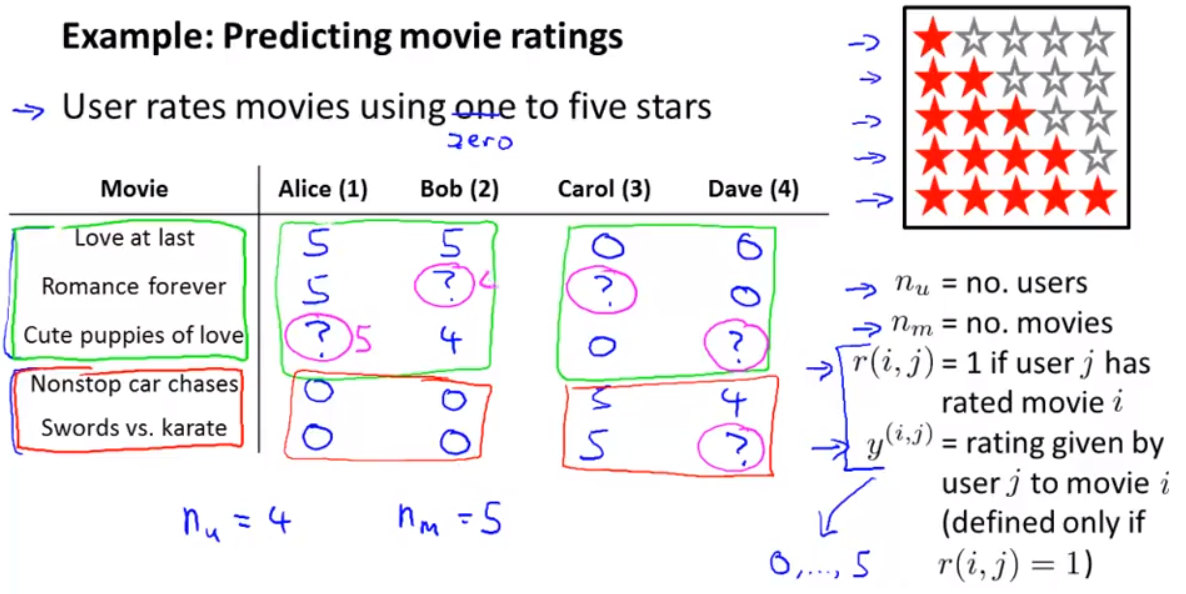
# Predicting Movie Ratings

## Problem Formulation

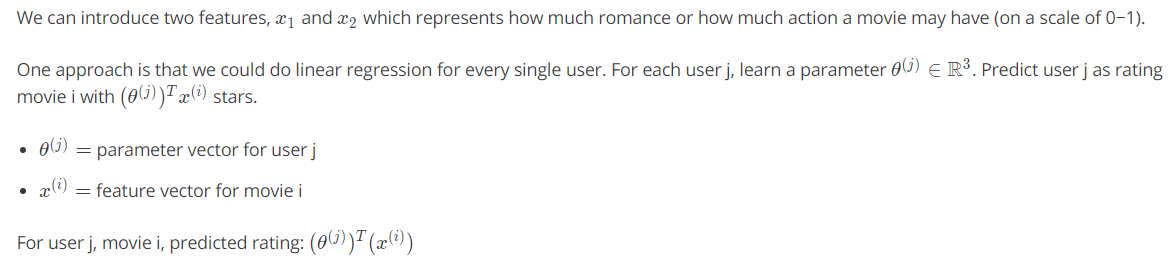
Recommendation is currently a very popular application of machine learning.

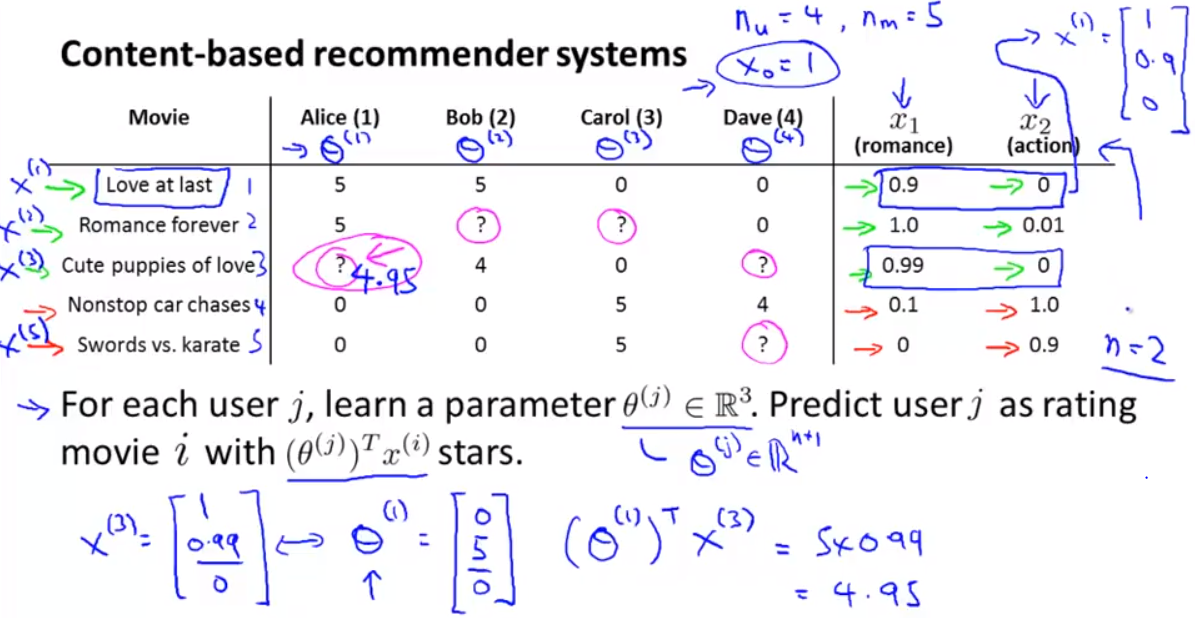
Say we are trying to recommend movies to customers. We can use the following definitions

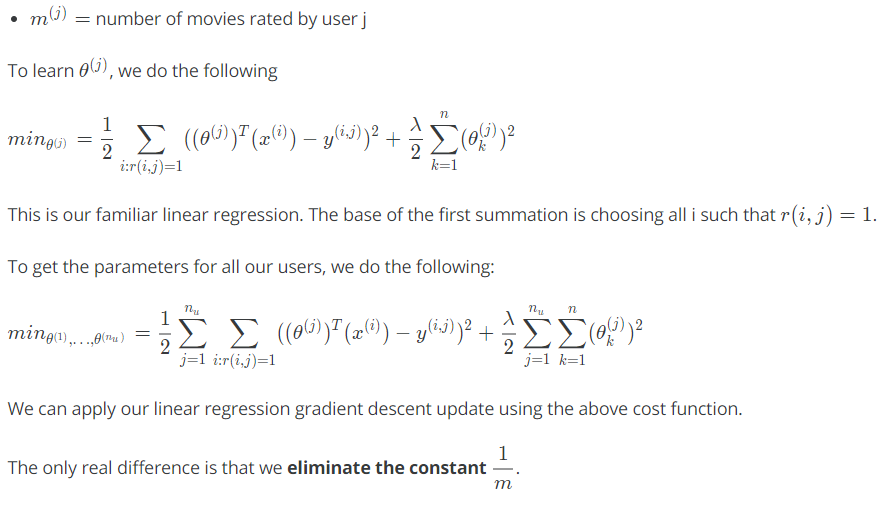


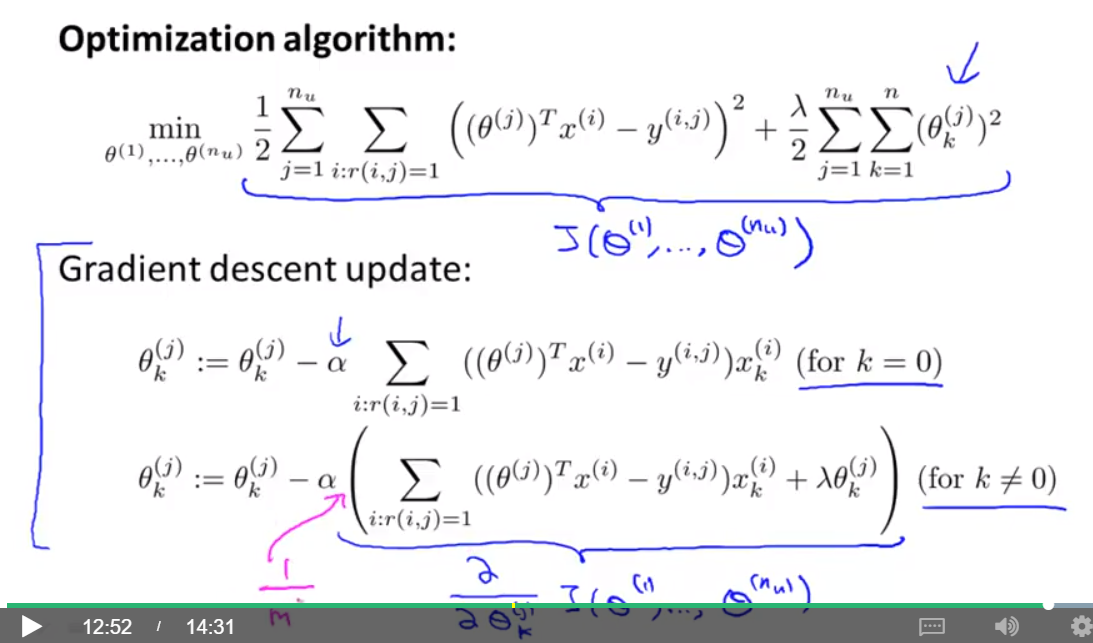


## Content Based Recommendations







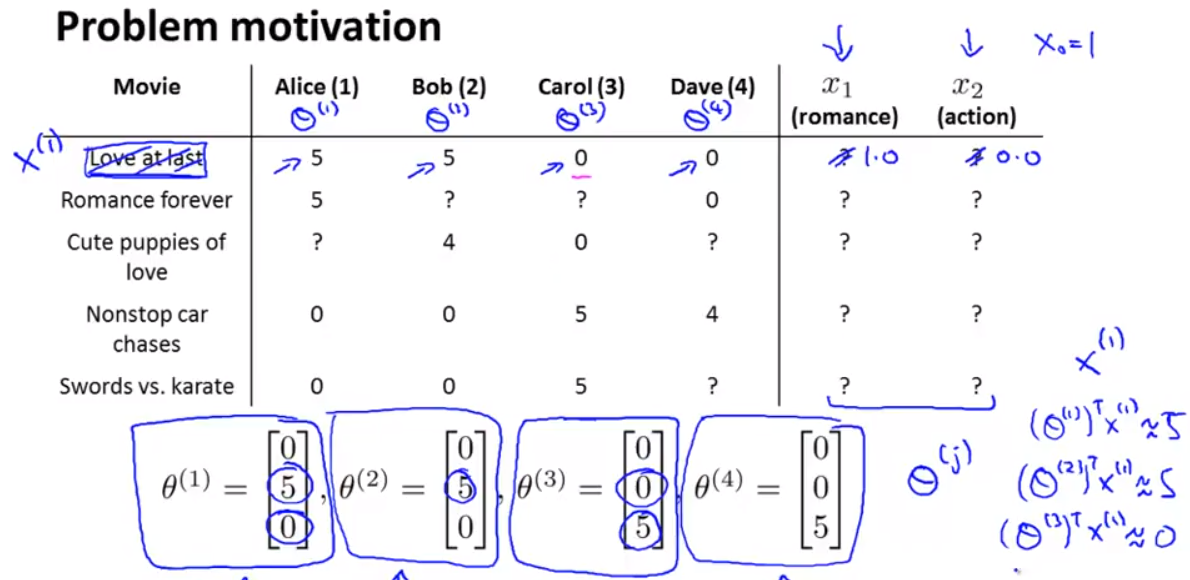


# Collaborative Filtering

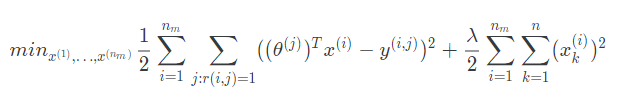
## Collaborative Filtering – Problem Formulation

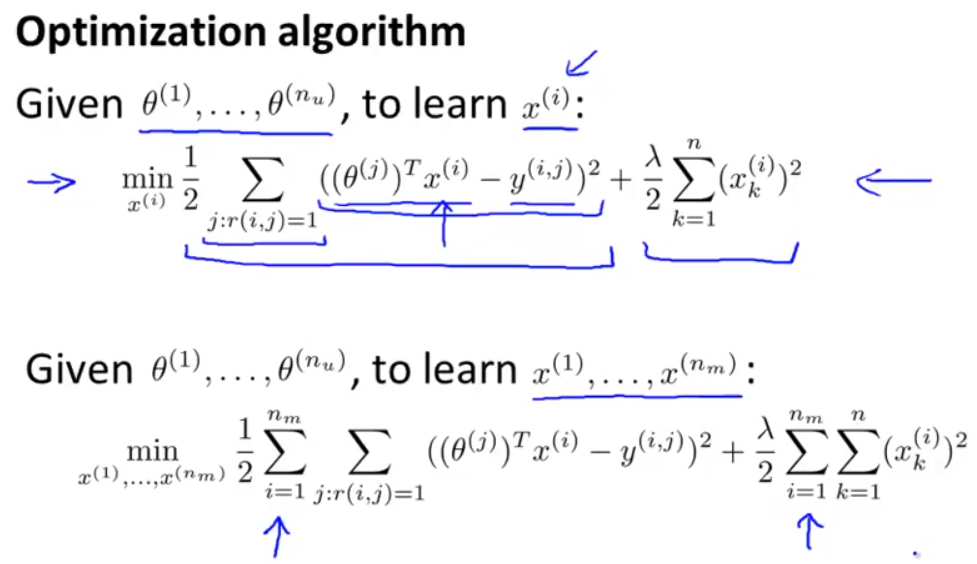
It can be very difficult to find features such as "amount of romance" or "amount of action" in a movie. To figure this out, we can use feature finders.

We can let the users tell us how much they like the different genres, providing their parameter vector immediately for us.



To infer the features from given parameters, we use the squared error function with regularization over all the users:

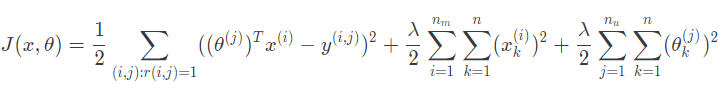


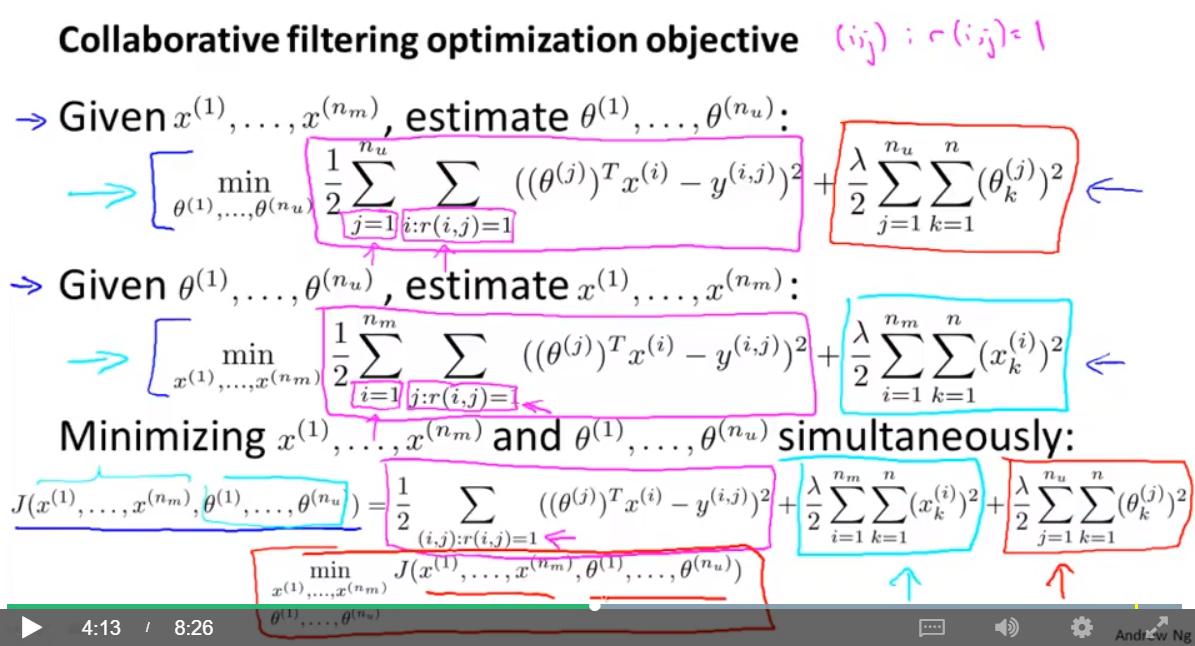


You can also **randomly guess** the values for theta to guess the features repeatedly. You will actually converge to a good set of features.

## Collaborative Filtering Algorithm

To speed things up, we can simultaneously minimize our features and our parameters:

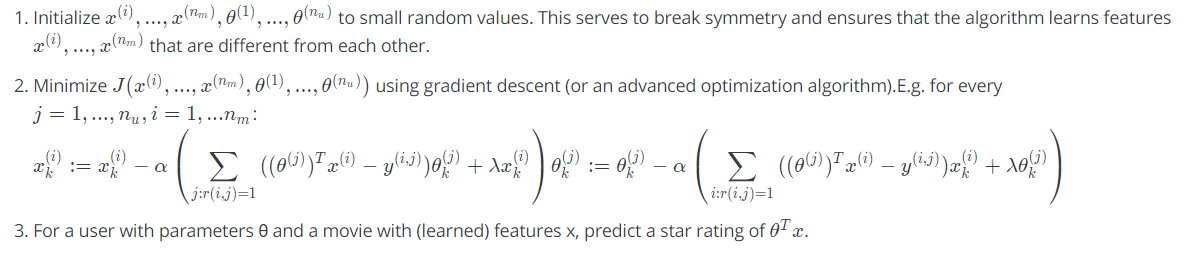




It looks very complicated, but we've only combined the cost function for theta and the cost function for x.

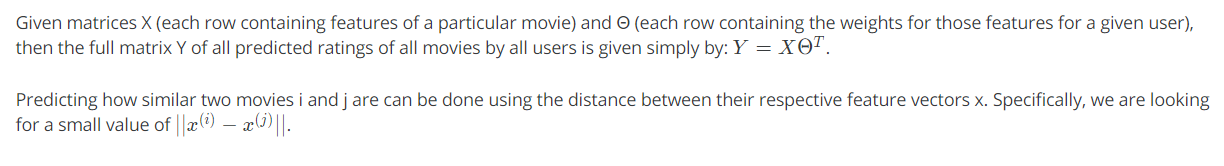
Because the algorithm can learn them itself, the bias units where x0=1 have been removed, therefore x∈ℝn and θ∈ℝn.

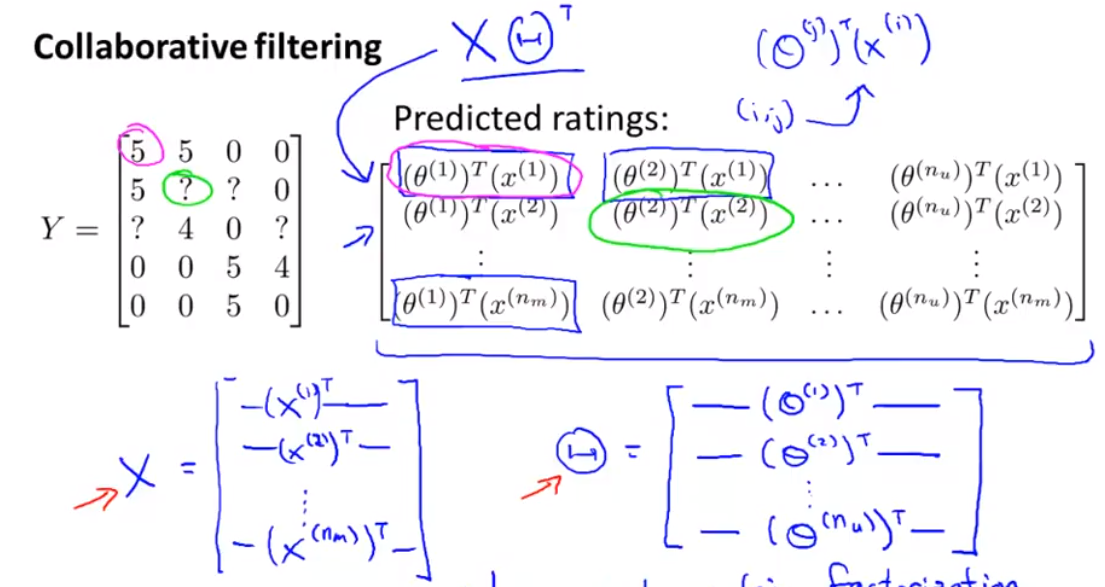
These are the steps in the algorithm:



# Low rank matrix factorisation

## Vectorization





## Implementation detail: Mean Normalisation

If the ranking system for movies is used from the previous lectures, then new users (who have watched no movies), will be assigned new movies incorrectly. Specifically, they will be assigned θ with all components equal to zero due to the minimization of the regularization term. That is, we assume that the new user will rank all movies 0, which does not seem intuitively correct.

We rectify this problem by normalizing the data relative to the mean. First, we use a matrix Y to store the data from previous ratings, where the ith row of Y is the ratings for the ith movie and the jth column corresponds to the ratings for the jth user.

