CS/ECE/ME 532 Sec. 004
Matrix Methods in Machine Learning

[Unit 3] Low Rank Representations and Singular Value Decomposition – Overview

Learning Objectives

At the end of this module, students will be able to:

- Find a low rank approximation to a matrix using given bases
- Use taste profiles to predict ratings/preferences
- Apply K-means to solve clustering problems
- Explain the dimensions of the SVD components for tall/narrow matrices and short/wide matrices
- Apply the economy or skinny SVD
- Apply the sum of outer products form of the SVD
- Write the rank-p approximation of a matrix using the SVD
- Use the right/left singular vectors to obtain orthonormal bases for the rows/columns of a matrix
- Use singular vectors to construct projection matrices
- Construct the pseudo inverse in terms of the SVD
- Define rank, operator norm, and ill-conditioning in terms of the singular values

Significance of Unit

In Unit 2, we looked at supervised learning. Unit 3 will focus on unsupervised learning because what we are trying to do with the low rank representation is to identify patterns that are present in the data. If we can describe high-dimensional data with a small number of patterns, we can make predictions, fill in missing entries, remove noise, or other non-ideal effects, and so on. So that's what a low rank representation is; it is describing data using a small number of patterns. We're going to look at two main forms: clustering and singular value decomposition. With a clustering problem, we take data and try to group vectors into vectors that have similar patterns, so cluster 1 would be a group of measurements that are similar and cluster 2 would be a group of measurements that are similar to each other, and so on. And that can be really useful because it can identify different conditions in the data. For example, I'm working on a project right now that is measuring the connectivity in the brain between different electrodes and that gives us a matrix of connectivities. We do this under different conditions: when the person is awake, sleeping, listening, reading, undergoing anesthesia, and we are trying to identify if whether different connectivity patterns or different networks, if you will, correspond to different states of the brain. So, this idea of clustering has many, many applications and is widely used. The other application that we will consider is the singular value decomposition, which may be the most powerful tool in linear algebra because it takes a matrix and gives us the patterns associated with the rows and associated with the columns. It gives us the patterns and tells us which ones are more important and how important they are relative to other patterns. If I have a matrix of data and I can describe it with a small number of patterns, singular value decomposition will tell me what those patterns are and how many I need to include. This allows me to do lots of things, like predict missing data in movie ratings and predict whether or not someone will like a movie. If you know there's a small number of patterns, it's easy to fill in those missing entries. Singular value decomposition will help us to solve systems of linear equations, which we talked about Unit 2. The applications of singular value decomposition will come up in Unit 4. In Unit 3, we will spend time understanding what singular value decomposition is and what its properties are, as well as continue with the theme of developing geometric insight.

Key Topics

1. Low-rank decomposition of matrices
2. Applications of low-rank matrix decompositions
   - 2.1. Taste profiles, patterns in data, and predicting ratings/preferences
   - 2.2. Unsupervised clustering: K-means
   - 2.3. Solving systems of linear equations using bases
     - 2.3.1. Obtaining unique solutions to nonunique problems using subspaces
     - 2.3.2. Classifiers using orthonormal bases
3. Gram-Schmidt for orthonormal bases. (did this in Unit 2)
4. Singular Value Decomposition: finding the "best" subspace
   - 4.1. Different forms - tall/narrow matrices, short/wide matrices, economy/skinny SVD, rank p form, sum of outer products form
   - 4.2. Orthonormal property of singular vectors
   - 4.3. Interpretation of singular vectors as bases for subspaces associated with the matrix
   - 4.4. Projection matrices and singular vectors
   - 4.5. Properties of singular values: rank, operator norm, ill-conditioning
   - 4.6. Define pseudo inverse

Learning Activities

- Instructional Units 3.1-3.3
- Activity 10
- Instructional Units 3.4, 3.5
- Activity 11
- Assignment 5
- Unit 3 Overview Quiz

Recommended Reading

- LE 9.0 Clustering in Non-negative Matrix Factorization
- LE 9.1 The K-means algorithm
- LE 6.1 The Singular Value Decomposition
- LE 6.2 Fundamental Subspaces
- LE 6.3 Matrix Approximation
- LE 6.5 Solving Least Squares Problems
- LE 6.7 Rank-Deficient and Underdetermined Systems