MAT 167: Applied Linear Algebra
Lecture 26: Reviews/Study Guide for the Final Exam

Naoki Saito

Department of Mathematics
University of California, Davis

December 3 & 5, 2025

Outline

Outline

Basic Notions Covered by the Midterm Exam

- Matrix-vector multiplication
- Range, nullspace, basis, rank, dimension, linear independence
- Vector & matrix norms
- Condition number of a matrix
- Inner product
- Orthogonality and orthonormal basis (orthogonal matrix)
- Projectors (including orthogonal projectors)
- QR factorization (classical & modified Gram–Schmidt orthogonalization, Householder triangularization, and Givens rotations)
- Least Squares Problems (pseudoinverse, normal equations)

Outline

SVD

- Definition and meaning of SVD
- Reduced SVD vs Full SVD
- Relation of the singular values and matrix 2-norm and Frobenius norm
- Similarity and difference between SVD and Eigenvalue Decomposition
- How range and nullspace are represented by some singular vectors

The Best Rank k Approximation

- Why the first k terms of the SVD of A is the best rank k approximation of A?
- Condition number and SVD
- Rank and SVD

SVD and the Least Squares Problem

- Solution of the Least Squares Problem using the reduced SVD
- Pseudoinverse and SVD
- Pseudoinverse and Orthogonal Projectors

PCA and SVD

- The meaning of Principal Component Analysis (a.k.a. Karhunen-Loève Transform)
- Centered data matrix and Covariance matrix
- classical vs neoclassical data analysis settings
- PCA as an eigenvalue decomposition of a covariance matrix
- Rank of a covariance matrix
- PCA using SVD of a centered data matrix

NNMF and SVD

- The motivation and definition of NonNegative Matrix Factorization
- Computing the NNMF of a given nonnegative matrix using the Alternating Least Squares method
- Its relationship with SVD
- Application to term-document matrix analysis

Outline

Clustering and Classification

- Difference between clustering (unsupervised learning) and classification (supervised learning)
- The k-means algorithm for clustering
- The k-nearest neighbor method for classification

Handwritten Digit Classification

- Notion of training and test datasets
- A simple classification based on the distances from the mean (representative) digit images
- The k-nearest neighbor classification
- Classification using SVD basis vectors (left singular vectors)

Text Mining

- Preprocessing (stop words elimination; stemming)
- How to construct term-document matrices
- The Vector Space Model
- How to measure the similarity between a query and documents
- Performance modeling: Precision vs Recall
- Latent Semantic Indexing $\Longrightarrow$ the best rank k approximation of the term-document matrix
- Using k-means clustering for text mining
- Using NNMF for text mining

Web Document Searches

- Link structure of the Web: inlinks and outlinks of webpages
- Idea of an authority score and a hub score of a webpage
- Authority scores and hub scores often mutually reinforce each other
- Web graphs, adjacency matrices

HITS

- The HITS (Hyperlink Induced Topic Search) algorithm
- Neighborhood graphs based on query terms
- Power iteration on authority scores and hub scores
- Basics of power iteration (why it works, etc.)
- Strengths and weaknesses of HITS

PageRank

- PageRank computation is independent from query terms
- PageRank vector and power iteration
- Markov chains/random walks interpretation
- "Google" matrix; row stochastic matrices; dangling nodes
- irreducible and reducible Markov chains
- how to force irreducibility
- importance (PageRank value) vs relevancy (query topic)
- Importance of sparsity in matrix computations
- Strengths and weaknesses of PageRank