Optimization and Computational Linear Algebra for Data Science

Outline

1. Vector spaces
1. General definitions
2. Linear dependency
3. Basis, dimension

2. Linear transformations
1. Linear transformations
2. Matrix representation
3. Kernel and image

3. Rank
1. Definition of the rank
2. Properties of the rank
3. Invertible matrices
4. Transpose of a matrix, symmetric matrices

4. Norm and inner product
1. Norm
2. Inner product
3. Orthogonality
4. Orthogonal projection and distance to a subspace

5. Matrices and orthogonality
1. Gram-Schmidt orthogonalization method
2. Orthogonal matrices

6. Eigenvalues, eigenvectors and Markov chains
1. Eigenvalues and eigenvectors
2. Diagonalizable matrices
3. Application to Markov chains
4. Example: Google's PageRank algorithm

7. The spectral theorem and PCA
1. The Spectral Theorem
2. Application: Principal Component Analysis (PCA)
3. Singular value decomposition
4. Interpretations of the SVD

8. Graphs and Linear Algebra
1. Graphs
2. Graph Laplacian
3. Spectral clustering with the graph Laplacian
4. Spectral clustering as a relaxation
5. Spectral clustering beyond graphs

9. Convex functions
1. Convex sets
2. Convex functions

10. Linear regression
1. Least squares
2. Penalized least squares: Ridge regression and Lasso
3. Norms for matrices
4. Low-rank matrix estimation and matrix completion

11. Optimality conditions
1. Local and global minimizers
2. Constrained optimization
3. The Lagrangian and the dual problem

4. Kuhn Tucker Theorem

12. GRADIENT DESCENT

1. Gradient descent
2. Newton's method