Optimization and Computational Linear Algebra for Data Science
Outline

1. Vector spaces
    1. General definitions
    2. Linear dependency
    3. Proof of Theorem 2.1
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. Transpose of a matrix, symmetric matrices
4. Norm and dot product
    1. Norm
    2. Dot product
    3. Orthogonality
    4. Orthogonal projection and distance to a subspace
5. Matrices and orthogonality
    1. Gram-Schmidt orthogonalisation 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. Singular value decomposition
    1. The Spectral Theorem
    2. Singular value decomposition
    3. Interpretation and applications 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. Optimality conditions
    1. Local and global minimizers
    2. Constrained optimization
    3. The Lagrangian and the dual problem
    4. Kuhn Tucker Theorem
11. Gradient descent
    1. Gradient descent
    2. Gradient descent for smooth convex functions
    3. Gradient descent for smooth strongly convex functions