Linear Algebra for Data Science

Module sequence M569A-D

M569 Module Sequence: The four one credit modules that make up Linear Algebra for Data Science graduate student sequence include:

- Math 569A: Matrices and Vector Spaces (a Linear Algebra primer)
- Math 569B: Geometric Techniques for Data Reduction
- Math 569C: Matrix Factorizations and Transformations
- Math 569D: Theoretical Foundations

The non-math major linear algebra primer M569-A is currently being offered on-line only. The courses MATH 569B-D are currently being offered experimentally as M580-A3 (M569B), M580-A4 (M569C) and M580-A5 (M569D).

MATH 569A: Matrices and Vector Spaces

This 15 week one credit module is available on-line only and serves as a bridge course to MATH 569B-D. The intended audience is non-math graduate students who desire to develop their understanding of the mathematics behind the algorithms in the Data Scientist's toolkit. It is the first course in a series of four one-credit modules and requires no previous knowledge of linear algebra. Students who have taken undergraduate linear algebra can use this course as a refresher, or start with the second module MATH 569B. Prerequisites: MATH 160.

1. Matrix operations
2. Linear systems and row operations
3. Underdetermined systems, RREF, rank/nullity
4. Solving linear systems: a geometric view
5. The matrix inverse and the LU decomposition
6. Vector spaces and subspaces
7. Sums of subspaces and the direct sum
8. Linear Combinations, span, column space.
9. Linear dependence, independence, the null space
10. The basis and change of basis
11. Basis for the row space
12. Basis for the column space
13. Orthogonality and matrix direct sum decompositions
14. Projections
15. Model fitting (Epilogue: Intro to the EVD and SVD)

M569B: Geometric Techniques for Data Reduction (MATH 580A3)

This 5 week one credit graduate course is offered during weeks 1-5. The intended audience is Math and non-math graduate students who desire to develop their understanding of the mathematics behind the tools in Data Science. It is the second course in a series of four one credit modules. The course requires previous knowledge of Linear Algebra commensurate with MATH 569A.

1. The Projection Matrix and its application to data sets.
2. Data modeling with $Ax = b$, a geometric perspective.
3. Hyperplanes, dot products and classification.
4. Determinants.
5. Eigenvalues and eigenvectors.
6. The characteristic polynomial.
7. Change of basis, similarity.
8. Diagonalization
9. Principal component analysis
10. Eigenbases and data reduction
11. Properties of PCA
12. Introduction to the Singular Value Decomposition
13. Image analysis and the SVD
14. Bases for the fundamental subspaces
15. Sample study

Math 569C: Matrix Factorizations and Transformations (MATH 580A4)

This 5 week one credit course is offered during Fall 2020 weeks 6-10. The intended audience is Math and non-math graduate students who desire to develop their understanding of the mathematics behind the tools in Data Science. It is the third course in a series of four one credit modules. The course requires previous knowledge of Linear Algebra consistent with the content of Math 569 A and B.

1. Graphs and Matrices
2. Multidimensional Scaling I: distance matrices and the algorithm
3. Multidimensional Scaling II: embedding of unit distance graphs and the circle
4. Fundamental Theorems of MDS
5. The Discrete Fourier transform
6. Angles between subspaces
7. Canonical correlation analysis
8. The GSVD and simultaneously diagonalization
9. Signal fraction analysis
10. Wavelets I: projections onto scaling and wavelet spaces
11. Wavelets II: recursive analysis and synthesis

12. Matching Pursuit
13. Sparse Dictionary Methods (KSVD)
14. Laplacian Eigenmaps and manifold learning
15. Subspace averaging

Math 569D: Theoretical Foundations

This five week one credit course (currently numbered M580A5) is offered during Fall 2020 weeks 11-15. The intended audience is Math and non-math graduate students who desire to develop their understanding of the mathematics behind the tools in Data Science. It is the fourth course in a series of four one credit modules. The course requires previous knowledge of Linear Algebra consistent with the content of Math 569 A, B and C.

1. Linear transformations, the rank-nullity theorem
2. Injection theorems
3. Surjection theorems
4. Isomorphism theorem
5. Matrix representation of a linear transformation
6. Inner product spaces
7. Best approximations
8. The adjoint
9. SVD revisited
10. The pseudo-inverse
11. The spectral theorem
12. PCA entropy criterion, circulant matrices
13. Mulitdimensional scaling theorems
14. Proof of the GSVD
15. The Courant-Fisher theorem