

Computational Data Mining

Fatemeh Mansoori

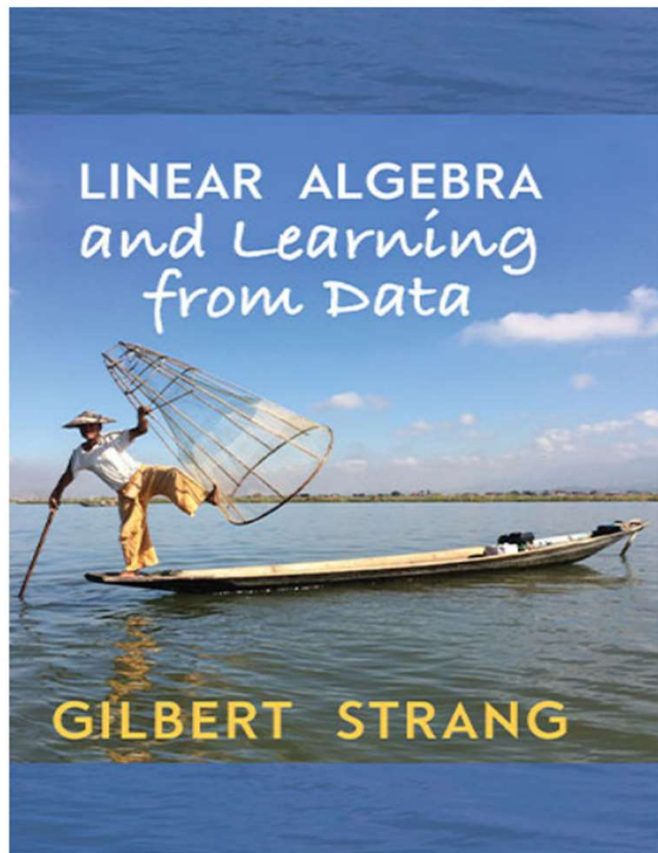
Course material

- <https://github.com/uisf-course/CDM> last year
- <https://github.com/uisf-course/CDM> Sp404 This year

Reference Book

- Main Reference :
 - Elden, L., Matrix Methods in Data Mining and Pattern Recognition, SIAM, 2007

References book



Textbook: [Linear Algebra and Learning from Data](#) by Gilbert Strang (2019).
(Additional readings will be posted for each lecture)

MIT Course :

- Matrix Methods In Data Analysis, Signal Processing, And Machine Learning

Course Topics

- Column space, basis, rank, rank-1 matrices, $A=CR$, and $AB=\sum(\text{col})(\text{row})$
 - Textbook 1.1–1.3
- Matrix multiplication by blocks and columns-times-rows
- The "famous four" matrix factorizations: [LU](#), [diagonalization \$X\Lambda X^{-1}\$ or \$Q\Lambda Q^T\$](#) , [QR](#), and the [SVD \$U\Sigma V\$](#)
- The [four fundamental subspaces](#) for an $m \times n$ matrix A of rank r , mapping "inputs" $x \in \mathbb{R}^n$ to "outputs" $Ax \in \mathbb{R}^m$
- Orthogonal bases and unitary matrices " Q "
- Eigenproblems, diagonalization

Course Topics

- (Symmetric/Hermitian) positive definite ("SPD") matrices
- [Low-rank approximation](#) and the Eckart–Young theorem.
- Diagonalizing the covariance matrix: [Principal components analysis \(PCA\)](#).
- [pseudo-inverse](#) $A^+ = V\Sigma^+U^T$
- Least-squares solve
- QR factorization : Gram–Schmidt algorithm
- Randomized linear algebra
- Matrix Calculus

Course Topics

- Matrix calculus
- Backpropagation for neural networks.
- Course Project :
 - Choose a topic related to the material in the course
 - Recommender System Using Matrix Factorization with Alternating Least Square Method
 - Non leaner matrix factorization for drug repositioning
 - Matrix Completion Methods for Recommendation Systems