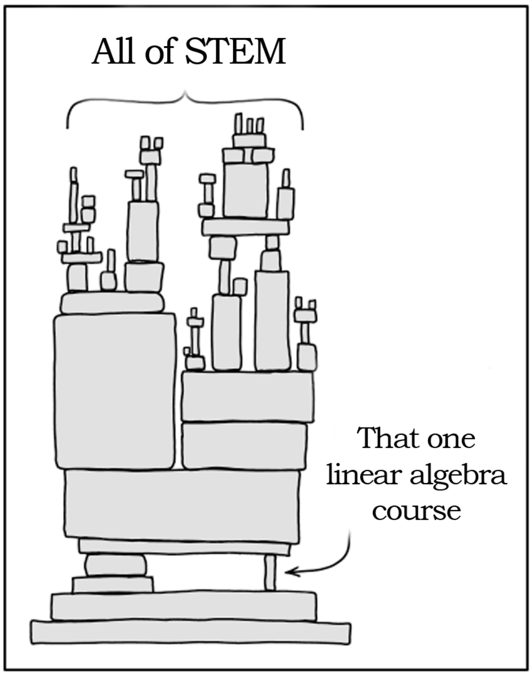
Applied Linear Algebra in Data Analysis



Source: https://twitter.com/Quasilocal/status/1664701458200121351

The above figure is not that far away from the truth. Linear algebra today is as important as calculus, if not more.

Linear algebra serves as the first step towards high-dimensional thinking, a prerequisite skill for modern data analysis. Paired with probability theory and optimisation, it comprises a potent trio of tools driving the current machine learning revolution.

This course aims to provide a solid foundation in applied linear algebra, basic probability theory, multivariate regression, and basic optimisation. But why these three topics?

- **Linear Algebra:** Think in high dimensions & visualise complex structures.
- **Probability Theory:** Navigate uncertainty with precision. Probability theory provides a principled approach to handling the unpredictable nature of real-world data.
- **Optimisation:** Seek the best solutions. Optimisation allows us to navigate the vast solution space and pinpoint the most efficient problem-solving paths.

Throughout this course, we will:

a) Introduce fundamental concepts in linear algebra, matrix methods, probability theory, multivariate regression, and basic optimisation.
b) Demonstrate the practical application of these concepts in real-world scenarios.
c) Provide hands-on experience for applying these concepts to solve real-world problems.

This course is a great foundational course for anyone interested in machine learning. It will equip you with the basic tools needed to understand concepts in machine learning and data science.

Course Content

(Modules in orange cover concepts, and Modules in blue are applications)

Vectors and Vector Spaces; Matrices; Solving Linear Equations; Orthogonality; Matrix Inverses; Signal processing: Fourier and Wavelet Transforms; Linear Least Squares and its Variants; Eigenvectors and Eigenvalues; Systems theory: Linear Dynamical Systems; Positive Definiteness & Semi-definiteness; Matrix Norms; Singular Value Decomposition; Dimensionality Reduction: Principal Component Analysis; Image compression; Basic Probability Theory; Basic Statistics; Linear Regression Models; Logistic Regression; Machine learning: Cross-validation; Linear Models for Regression and Classification; Optimization: A brief introduction; Linear programming; Neuromechanics: Linear Programming applied to neuromuscular control, and radiation planning.

Assignments

- This is a hands-on course heavy on pen-and-paper and programming assignments. Students are encouraged to use Python or Julia for their programming assignments.

- You have a cumulative 5-day late submission grace period for your assignments for the entire course. If you miss the 6 PM deadline for an assignment, a submission made before 6 PM the following day will be considered 1-day late submission.
- If you use up your 5-day grace period, all subsequent late assignments will receive 0 marks.

Resources

The topics covered in this course are quite mature, so it's no surprise that there are numerous wonderful resources. Here are some resources that cover the topics discussed in this course. There are several wonderful resources:

1. G Strang, Introduction to Linear Algebra. Wellesley, *MA: Wellesley-Cambridge Press*, 1993.
2. CD Meyer, Matrix Analysis and Applied Linear Algebra. *Siam*; 2000 Jun 1.
3. S Boyd and L Vandenberghe, Introduction to Applied Linear Algebra – Vectors, Matrices, and Least Squares. Online Book.
4. Online lectures: Linear Algebra by Prof. Gilbert Strang.
5. Selected online lectures: Matrix Methods in Data Analysis, Signal Processing, and Machine Learning, by Prof. Gilbert Strang.
6. Selected online lectures: Linear Dynamical Systems by Prof. Stephen Boyd.
7. Gelman, Andrew, Jennifer Hill, and Aki Vehtari. Regression and other stories. Cambridge University Press, 2021.

Grading

The final grade for the course will consider the performance on the assignments, quizzes, mid-term, and the final exam.

Assignments	15
Quizzes	10
Mid-term	20
Final	55
Total	100

An absolute grading scheme will be followed for the course:

S: >=90; **A**: >=80; **B**: >= 70; **C**: >= 60; **D**: >=50; **E**: >=40; **F**: <40

Class Schedule

Week	Day	Date	Details
1	1	4/8/25	Concepts in Vector Spaces
	2	5/8/25	
	3	7/8/25	
2	4	11/8/25	Matrices
	5	12/8/25	
	6	14/8/25	
3	7	18/8/25	Solving linear equations
	8	19/8/25	
	9	21/8/25	
4	10	25/8/25	Orthogonality
	11	26/8/25	
	12	28/8/25	**Midterm 1**
5	13	1/9/25	Matrix Inverses
	14	2/9/25	
	15	4/9/25	**Application**: Signal Processing – Fourier & Wavelet Transform
6	16	8/9/25	Eigenvectors and Eigenvalues
	17	9/9/25	
	18	11/9/25	**Application**: Linear Dynamical Systems
7	19	15/9/25	Review with TAs
	20	16/9/25	
	21	18/9/25	**Midterm 2**
8	22	22/9/25	Positive Definite Matrices
	23	23/9/25	Singular Value Decomposition
	24	25/9/25	**Application:** Dimensionality Reduction: Principal Component Analysis; Image compression
9	25	29/9/25	Linear Least Squares and its Variants
	26	30/9/25	
	27	2/10/25	**Application**: Signal Processing, CT reconstruction
10	28	6/10/25	Probability and Statistics
	29	7/10/25	
	30	9/10/25	
11	31	13/10/25	Probability and Statistics
	32	14/10/25	
	33	16/10/25	
12	34	20/10/25	Review with TAs
	35	21/10/25	
	36	23/10/25	**Midterm 3**
13	37	27/10/25	Linear Regression Models
	38	28/10/25	
	39	30/10/25	
14	40	3/11/25	Linear Regression Models
	41	4/11/25	
	42	6/11/25	
15	43	10/11/25	Linear Regression Models
	44	11/11/25	

	45	13/11/25	
16	46	17/11/25	Application: Machine Learning
	47	18/11/25	
	48	20/11/25	Course Conclusion
		24/11/25	Course Review
		25/11/25	
		27/11/25	Final Exam

Assignment Schedule

Assignment	Details	Due on
1	Vector spaces	25/8/25
2	Matrices	
3	Solution to Linear Equations	
4	Orthogonality & Matrix Inverses	15/9/25
5	Signal Processing	
6	Eigenvectors and Eigenvalues, Linear Dynamical Systems	
7	Positive Definite Matrices, SVD and Dimensionality Reduction	6/10/25
8	Linear Least Squares	
9	Probability & Statistics	20/10/25
10	Linear Regression Models	17/11/25
11	Machine Learning Applications	24/11/25