Course Outline - Mathematics for Computer Science

March 9, 2024

Mathematics for Computer Science

0.1 MODULE 01: Linear Algebra

0.1.1 Chapter 01: Systems of Linear equations: two variables

- Machine learning motivation
- Systems of sentences
- Systems of equations
- Systems of equations as lines
- A geometric notion of singularity
- Singular vs nonsingular matrices
- Linear dependence and independence
- The determinant

0.1.2 Chapter 02: Systems of Linear Equations: three variables

- Systems of equations (3×3)
- Singular vs non-singular (3×3)
- Systems of equations as planes (3×3)
- Linear dependence and independence (3×3)
- The determinant (3×3)

0.1.3 Chapter 03: Solving systems of Linear Equations: Elimination

- Machine learning motivation
- Solving non-singular systems of linear equations
- Solving singular systems of linear equations
- Solving systems of equations with more variables
- Matrix row-reduction
- Row operations that preserve singularity
- Gaussian elimination

0.1.4 Chapter 04: Solving systems of Linear Equations: Row Echelon Form and Rank

- The rank of a matrix
- The rank of a matrix in general
- Row echelon form
- Row echelon form in general
- Reduced row echelon form

0.1.5 Chapter 05: Vectors

- Norm of a vector
- Sum and difference of vectors
- Distance between vectors
- Multiplying a vector by a scalar
- The dot product
- Geometric Dot Product
- Multiplying a matrix by a vector
- **Assignment**: Vector Operations: Scalar Multiplication, Sum and Dot Product of Vectors

0.1.6 Chapter 06: Linear transformations

- Matrices as linear transformations
- Linear transformations as matrices
- Matrix multiplication
- The identity matrix
- Matrix inverse
- Which matrices have an inverse?
- Neural networks and matrices

0.1.7 Chapter 07: Determinants In-depth

- Machine Learning Motivation
- Singularity and rank of linear transformation
- Determinant as an area
- Determinant of a product
- Determinants of inverses

0.1.8 Chapter 08: Eigenvalues and Eigenvectors

- Bases in Linear Algebra
- Span in Linear Algebra
- Interactive visualization: Linear Span
- Eigenbases
- Eigenvalues and eigenvectors
- Principal Component Analysis (PCA)

[ ]:

[ ]:

[ ]:

[ ]:

[ ]:

[ ]:

0.2 MODULE 02: Probability & Statistics

0.2.1 Chapter 01: Introduction to probability

- Concept of probability: repeated random trials
- Conditional probability and independence
- Discriminative learning and conditional probability
- Bayes theorem

0.2.2 Chapter 02: Random variables

- Random variables
- Cumulative distribution function
- Discrete random variables: Bernoulli distribution
- Discrete random variables: Binomial distribution
- Probability mass function
- Continuous random variables: Uniform distribution
- Continuous random variables: Gaussian distribution
- Continuous random variables: Chi squared distribution
- Probability distribution function

0.2.3 Chapter 03: Describing distributions

- Measures of central tendency: mean, median, mode
- Expected values
- Quantiles and box-plots
- Measures of dispersion: variance, standard deviation

0.2.4 Chapter 04: Random vectors

- Joint distributions
- Marginal and conditional distributions
- Independence
- Measures of relatedness: covariance
- Multivariate normal distribution

0.2.5 Chapter 05: Sampling and point estimates

- Population vs. sample
- Describing samples: sample proportion and sample mean
- Distribution of sample mean and proportion: Central Limit Theorem
- Point estimates
- Biased vs Unbiased estimates

0.2.6 Chapter 06: Maximum likelihood estimation

- ML motivation example: Linear Discriminant Analysis
- Likelihood
- Intuition behind maximum likelihood estimation
- MLE: How to get the maximum using calculus

0.2.7 Chapter 07: Bayesian statistics

- ML motivation example: Naive Bayes
- Frequentist vs. Bayesian statistics
- A priori/ a posteriori distributions
- Bayesian estimators: posterior mean, posterior median, MAP

0.2.8 Chapter 08: Confidence intervals

- Margin of error
- Interval estimation
- Confidence Interval for mean of population
- CI for parameters in linear regression
- Prediction Interval

0.2.9 Chapter 09: Hypothesis testing

- ML Motivation: AB Testing
- Criminal trial
- Two types of errors
- Test for proportion and means
- Two sample inference for difference between groups
- ANOVA
- Power of a test

0.2.10 Chapter 10: Simple Linear Regression

- Variance and Standard Deviation
- Covariance and Covariance Matrix
- Correlation and Correlation Matrix
- Regression
- Regression Analysis
- Linear Regression
  - Fitting a Line using Gradiant Descent
  - Fitting a Line using Linear Least Squares (with one feature)

0.3 Bonus:

- Recap
  - Classical vs Emperical Probability
  - Marginal/Simple/Unconditional Probability (with examples)
  - Types of Events
  - Joint Probability & Conditional Probability
- Bayes' Theorem
  - Proof of Bayes' Theorem
  - Formulas of Bayes' Theorem
  - Examples of Bayes' Theorem
- Overview of Naïve Bayes' Classifier
- Naïve Bayes' Classifier for Datasets with Discrete Input Features
  - Example 1: Single Input FeatureExample

  - 2: Multiple Input Features
  - Example 3: Multiple Input Features
  - Example 4: Using Naïve Bayes' on Text Data
- Naïve Bayes' Classifier for Datasets with Continuous Input Features
  - Example 1: Single Input Feature
  - Example 2: Multiple Input Features

Assignments

[ ]:

[ ]:

[ ]:

[ ]:

[ ]:

0.4 MODULE 03: Calculus

0.4.1 Chapter 01: Introduction to Calculus, Functions and Limits

- What is Calculus?
- Functions and their implementation using Python
- What Are Limits?
- Precise Definition of Limit
- Limit Laws
- Infinities and Asymptotes
- Indeterminate Forms
- Limits in Python
- Limits with Plotting in Python

0.4.2 Chapter 02: Derivatives

- Example to motivate derivatives: Speedometer
- Derivative of common functions (c, x, x^2, 1/x)
- Meaning of e and the derivative of e^x
- Derivative of log x
- Existence of derivatives
- Properties of derivative
- Using Differentiation Rules to Differentiate a Function
- Using First Principle to Differentiate a Function
- Python Code to Differentiate a function using First Principle
- Plotting the function and its derivative
- How to Find Minima & Maxima using Derivatives?
- Partial Derivatives and Gradient

0.4.3 Chapter 03: Optimization with derivatives

- Intro to optimization
- Optimizing cost functions in ML: Squared loss
- Optimizing cost functions in ML: Log loss

0.4.4 Chapter 04: Gradients and optimization

- Intro to gradients
- Example to motivate gradients: Temperature
- Gradient notation
- Optimization using slope method: Linear regression

0.4.5 Chapter 05: Gradient Descent Algorithm

- Mountain Analogy for Gradient Descent Algorithm
- What is Gradient Descent and How to do Gradient Descent?
- Gradient Descent in Python
- Optimization using gradient descent: 1 variable
- Optimization using gradient descent: 2 variable
- Challenges of Gradient Descent Algorithm
  - Local Minima and Unfortunate Starting Point
  - Vanishing Gradient Problem
  - Exploding Gradient Problem

0.4.6 Chapter 06: Gradient descent for linear regression

- Recap of Linear Regression using OLS
- Example (Simple Linear Regression using Gradient Descent)
- Minimizing Error Function for Gradient Descent Algorithm
- Partial Derivatives of Error Function
- Calculate Regression Coefficients Using Gradient Descent Algorithm
- Fit the Line
- Carryout the Prediction

0.4.7 Chapter 07: Optimization in Neural Networks

- What is a Single Layer Perceptron?
- Perceptron with no activation and squared loss (linear regression)
- Perceptron with sigmoid activation and log loss (classification)
- Two-layer neural network with sigmoid activation and log loss
- Mathematics of Backpropagation ### Chapter 08: Beyond Gradient Descent: Newton's Method
- Root finding with Newton's method
- Adapting Newton's method for optimization
- Second derivatives and Hessians
- Multivariate Newton's method

[ ]:

[ ]:

[ ]:

[ ]:

[ ]:

[ ]:

[ ]:

[ ]: