Introduction

Department of Computer Engineering

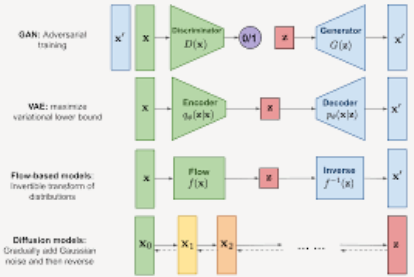
Sharif University of Technology

Hamid R. Rabiee rabiee@sharif.edu
Maryam Ramezani maryam.ramezani@sharif.edu

Course Description

How is our course?

Machine Learning Mathematics



Machine Learning Mathematics

01

Linear Algebra

Provides the foundation for manipulating data in high-dimensional spaces, essential for vector operations in machine learning.

02

Statistics & Probability

Offers tools to model uncertainty and make inferences about data, forming the backbone of many machine learning algorithms.

03

Optimization

Focuses on finding the best parameters for a model by minimizing or maximizing an objective function.

04

Information Theory

Analyzes data compression and transmission, helping quantify uncertainty and information gain in machine learning models.

05

Signal Processing

Involves analyzing and transforming data signals, important for handling time-series data and feature extraction in machine learning.

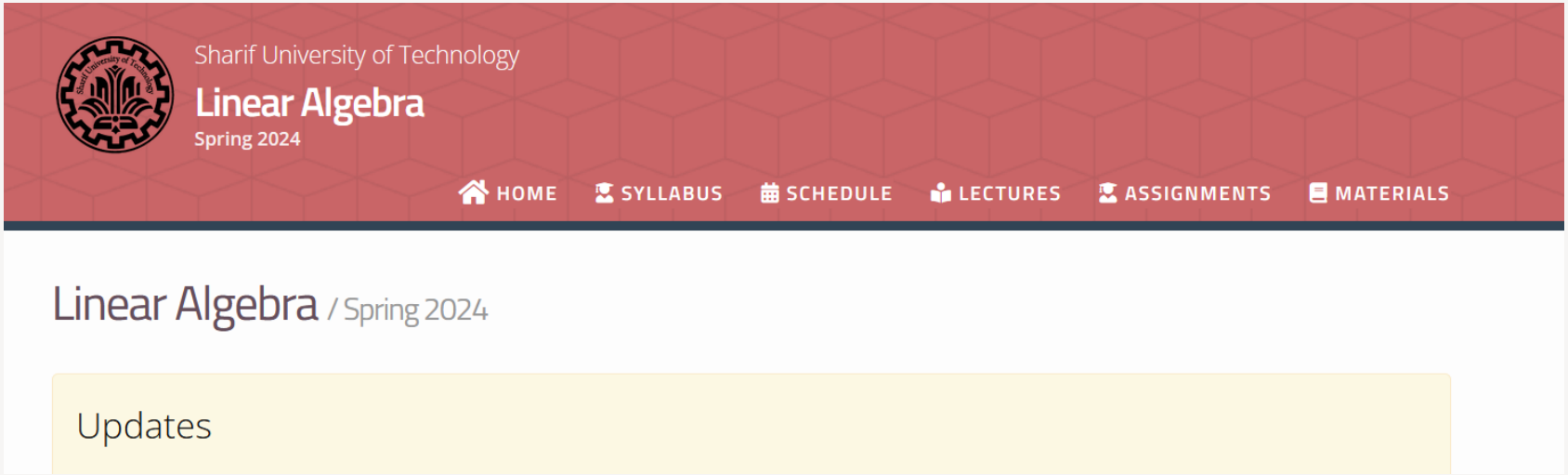
06

Graph Theory

Studies networks of nodes and edges, crucial for understanding relationships in data and complex models like Graph Neural Networks.

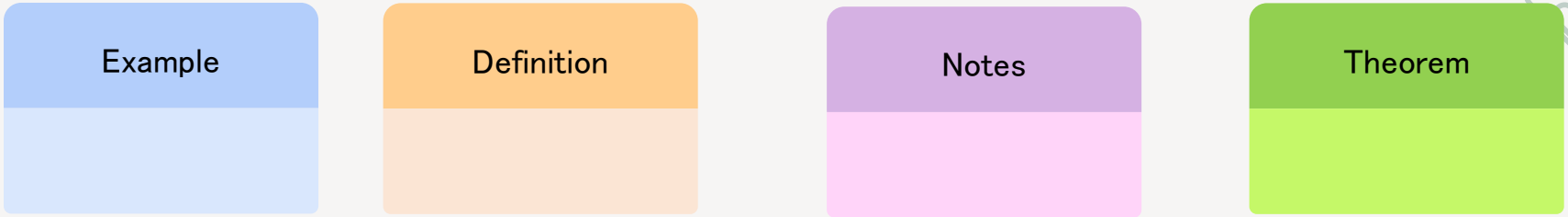
Course Material

https://sut-ce-courses.github.io/linearalgebra (jabrekh.ir)



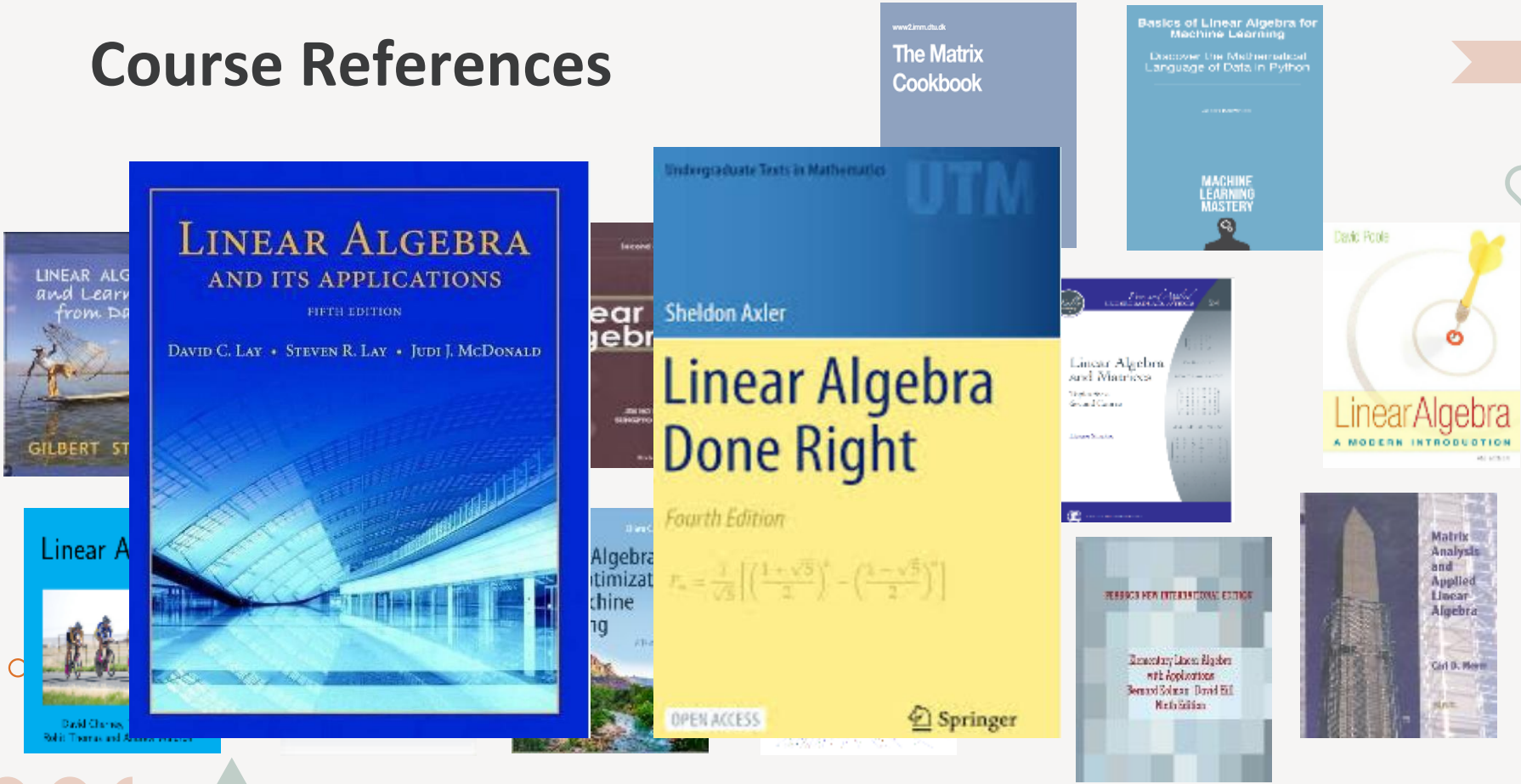
Course Method

- Slides



- Writing on board
- Geometric Interpretation and Intuition
- Notebooks

Course References



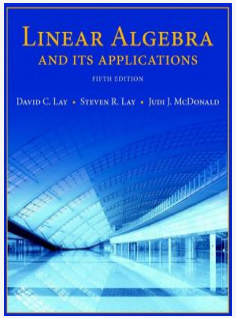
Reach Us!

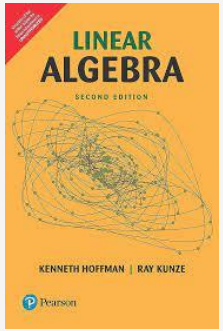
- Office Room: CE802-CE803
- Email:
  - Hamid R. Rabiee: (rabiee@sharif.edu)
  - Maryam Ramezani: (maryam.ramezani@sharif.edu)
- Course notes, homework and solutions, handouts, and other useful resources are available on the course page:
  - https://quera.org/course/18745/
  - Room: (Sunday & Tuesday: 13:30-15:00)
  - Ebnesina Alef 17
  - https://vc.sharif.edu/ch/rabiee
- Lead TA:
  - Mohammad Javad Ahmadpour: (mohamadahmadpour1383@gmail.com)
- Feedback
  - https://forms.gle/ox1gUHtbh9DCRgH29

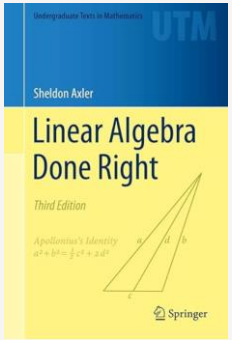
Resources

- Textbooks:
  - Sheldon Axler, Linear Algebra Done Right, fourth edition, 2023
  - Kenneth Hoffman and Ray A. Kunze.Linear Algebra. PHI Learning, 2004.
  - Gilbert Strang. Introduction to Linear Algebra. Wellesley–Cambridge Press, 2016.
  - David C. Lay, Steven R. Lay, and Judi J. McDonald.Linear Algebra and Its Applications. Pearson, 2016.

+Other textbooks and course materials.







Structure of the Course

Lectures

Goal: To introduce concepts in linear algebra, and motivate their use and importance.

Note: We try to cover useful materials in class, but we recommend you reading more!

Assignments

Purpose: To give you a chance to exercise your mind, and to solidify the concepts introduced to you in class.

Structure: Six theorical problems, and Three linear algebra practical problems.

Importance: Not important unless you want to learn the material and get a good grade.

Exams

Quiz: 5 with dropping lowest

Midterm: 1

Final: 1

Lecture Notes

- Lecture slide will be uploaded.
- Many times we will write on board, in real-time, during lecture to prove a theory or answer a question or add some additional explanations. It will be your responsibility to take notes.
- Slides links will be provided on site.

Syllabus

1403/07/01	S01	1	Elementary Row Operations, and Linear Equations
1403/07/03	S02		Row Reduction and Echelon Forms
1403/07/08	S03		Vector Space
1403/07/10	S04		Subspace

1403/07/15	S05	2	Independence (Linear and Affine)
1403/07/17	S06		Independence (Linear and Affine)
1403/07/22	S07		Bases, Dimension
1403/07/24	S08		Dimension and Rank
1403/07/29	S09		Dimension and Rank

Syllabus

1403/08/01	S10	3	Inner Product Space
1403/08/06	S11		Inner Product Space
1403/08/08	S12		Euclidean Norm, Inequalities and Angle
1403/08/13	S13		Orthogonality (Gram–Schmidt, etc.)

1403/08/15	S14	4	Linear Transformation
1403/08/20	S15		Linear Transformation
1403/08/22	S16		Change Basis
1403/08/27	S17		Inverse
1403/08/29	S18		Determinant
1403/09/04	S19		Determinant

Syllabus

1403/09/06	S20	5	Eigenvectors and Eigenvalues
1403/09/11	S21		Singular Values and Singular Vectors
1403/09/13	S22		Symmetric Matrices and Quadratic Forms
1403/09/18	S23		Diagonalization
1403/09/20	S24		Matrix Factorization
1403/09/22	Midterm Exam		
1403/09/25	S25		SVD

Syllabus

1403/09/27	S26	6	Norm Space
1403/10/02	S27		Derivatives
1403/10/04	S28		Derivatives
1403/10/09	S29		Least squares
1403/10/11	S30		Least squares
1403/10/30	Final Exam		

Assignments

- 6 Theory
- 3 Programming
    - Basic NumPy for array manipulation
    - Learning linear algebra and application of machine learning.
- You have a total 4 days of allowed late submission (for each theoretical and code parts). 0.5% of the assignment grade will be subtracted for each hour of delay.
- Homework released at the beginning of each part and your responsed should be uploaded on Quera.

Assignments

Assignment	Release	Dealine	Solution Release
T1	1403/07/01	1403/07/22	1403/07/23
P0	1403/07/03	1403/07/28	
T2	1403/07/15	1403/08/11	1403/08/12
T3	1403/08/01	1403/08/25	1403/08/26
T4	1403/08/15	1403/09/16	1403/09/17
P1	1403/09/04	1403/09/30	
T5	1403/09/06	1403/10/07	1403/10/09
T6	1403/09/27	1403/10/28	1403/10/30
P2	1403/10/02	1403/11/11	

Exams

Quizzes will be held as the following table on 13:00 for half an hour and will be graded in Grade Scope.

Exam	Time
Quiz 1	1403/07/24 13:00-13:30
Quiz 2	1403/08/13 13:00-13:30
Quiz 3	1403/08/27 13:00-13:30
Quiz 4	1403/09/18 13:00-13:30
Midterm	1403/09/22 09:00-12:00
Quiz 5	1403/10/11 13:00-13:30
Final	1403/10/30 09:00-12:00

TA Classes

TA Classes will be held on Wednesday from 18:00 till 20:00 in room https://vc.sharif.edu/rabiee-ta.

Date	Title
1403/07/18	Part1
1403/08/09	Part2
1403/08/23	Part3
1403/09/14	Part4
1403/09/21	Midterm
1403/10/05	Part5
1403/10/29	Final

Grading Policy

Quizzes

3 points (5 quizzes each 0.75 points; the lowest score will be dropped)

Written Assignments

5.6 points (each of the first four is 0.85 points, and each of the last two is 1.1 points)

Programming Assignments

2.4 points (preliminary exercises: 0.4 points, and the next two exercises are each 1 points)

Midterm

5 points (Lectures part 1,2,3,4)

Final

5 points (Lectures part 5,6)

Total: 21 Points

Introduction

Lets think about a question!

01

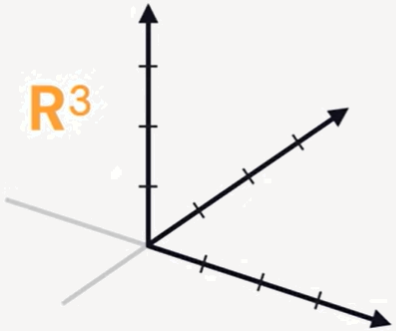
Vector

Basic Concept

Data Representations (Linear Algebra)

- How can we represent data (images, text, user preferences, etc.) in a way that computers can understand?
- Organize information into a vector!
  - A vector is a 1-dimensional array of numbers.
  - It has both a magnitude (length) and a direction
- The totality of a vectors with n entries is an n-dimensional vector space.

$$V = \nearrow = \begin{bmatrix} -3 \\ 0.7 \\ 2 \end{bmatrix}$$



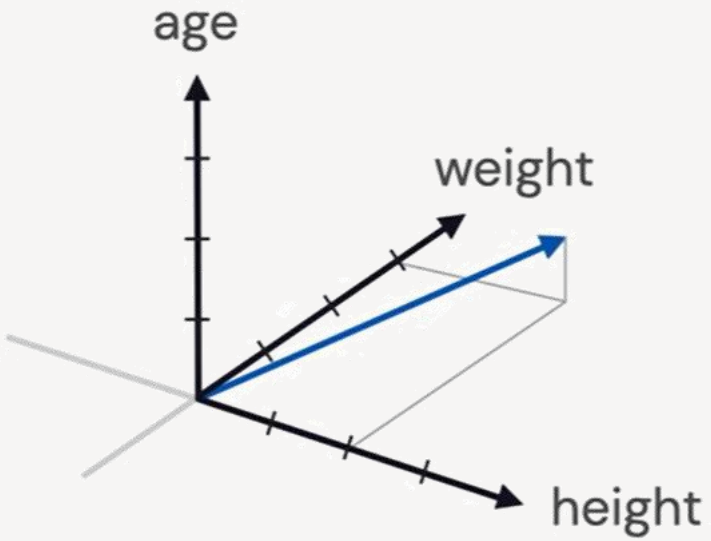
"3-dimensional space" consists of all vectors with 3 entries:

$$\begin{bmatrix} * \\ * \\ * \end{bmatrix}$$

Data Representations (Machine Learning)

- A **feature vector** is a vector whose entries represent the “features” of an object.
- The vector space containing them is called **feature space**.

$$P = \begin{bmatrix} 64 \\ 131 \\ 24 \end{bmatrix} \begin{matrix} \text{height} \\ \text{weight} \\ \text{age} \end{matrix}$$



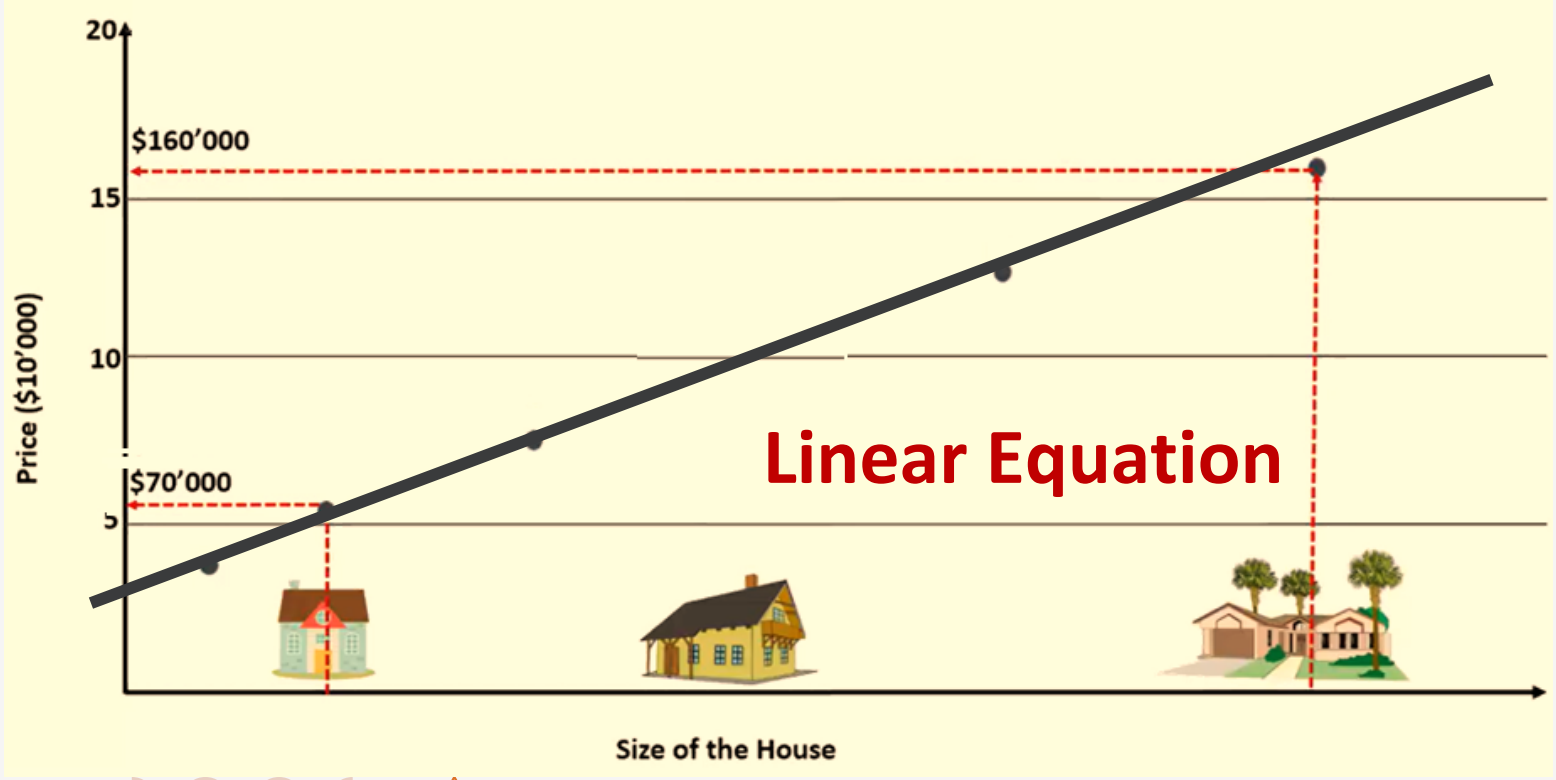
02

Equation with Matrix and Vector Format

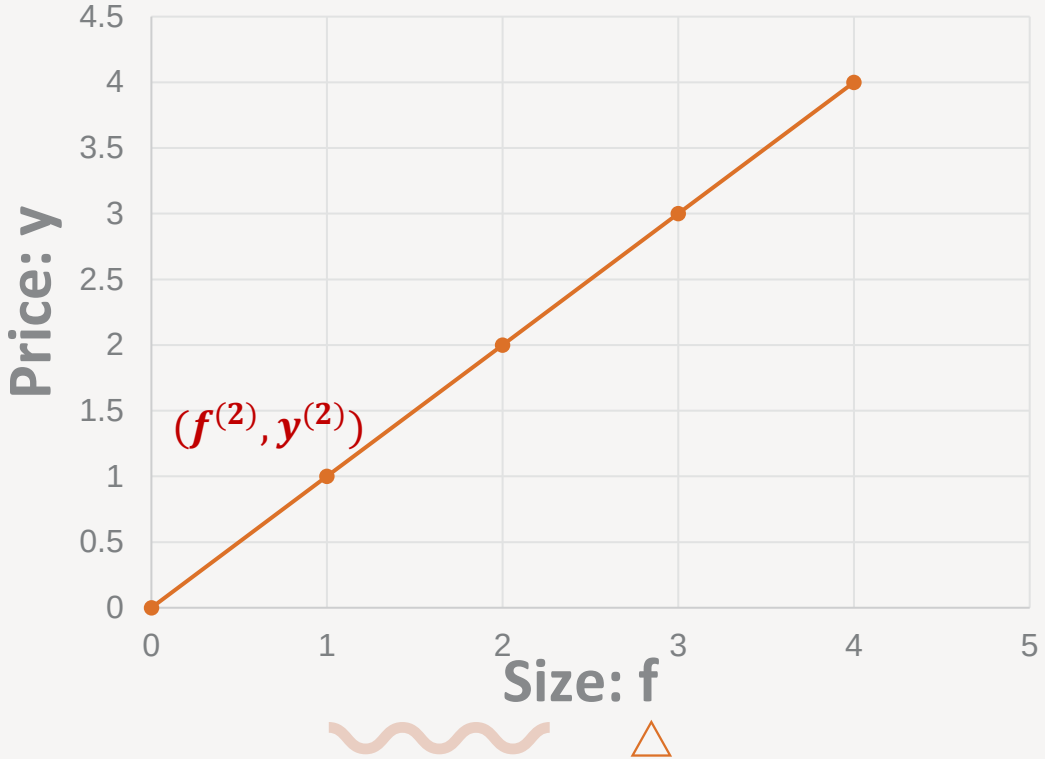
House Pricing Example

Price Problem

Price Problem



Linear Equation without offset



$$y = af$$

How to convert it to matrix-vector multiplication?

$$Ax = b$$

Linear Equation with offset

$$y = af + c$$

How to convert it to matrix-vector multiplication?

$$Ax = b$$

House Features

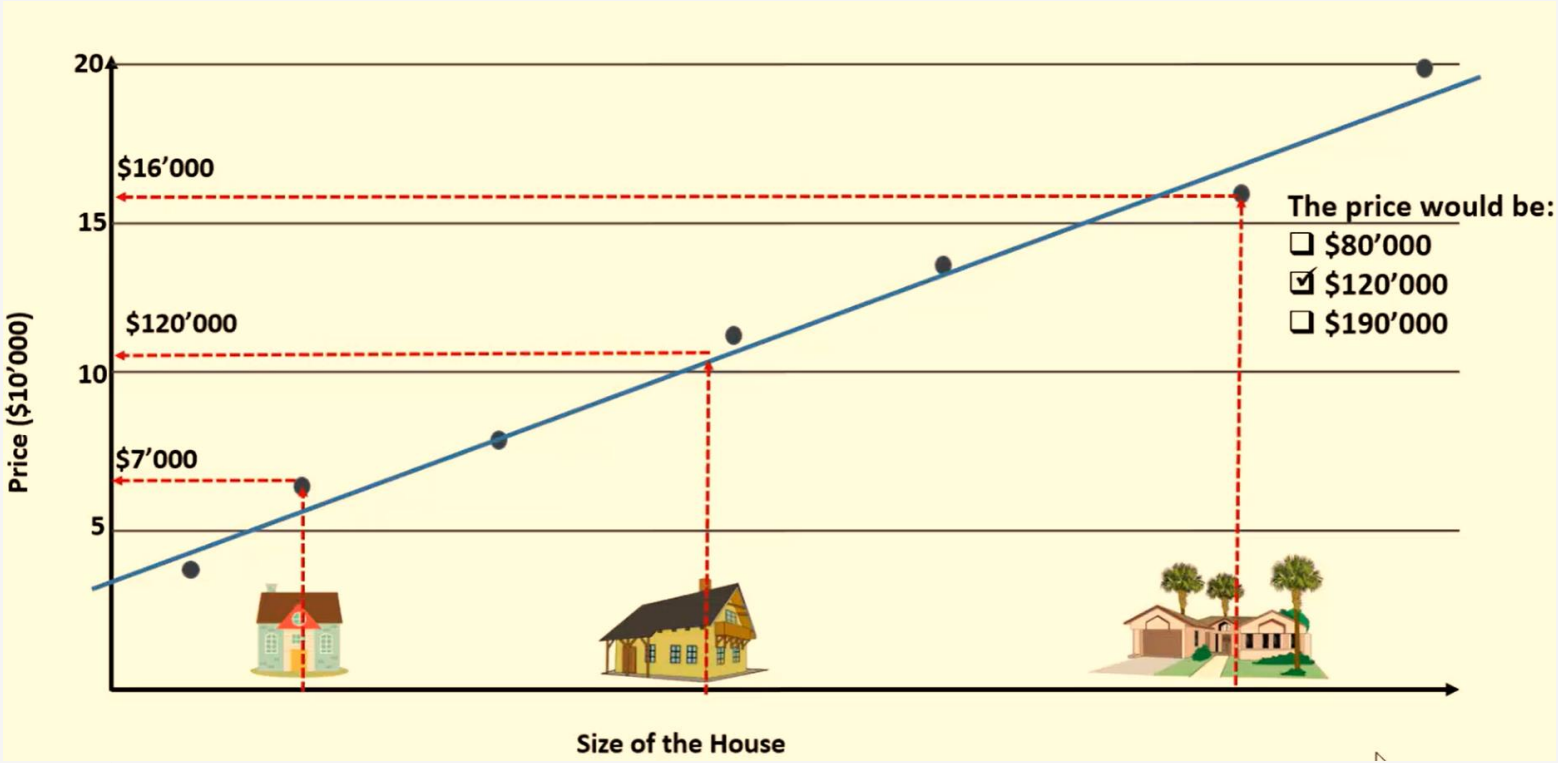
- #Room
- Size
- #Bedroom
- Age
- Address features: Street, Alley, …
- Size of part1, part2, part3, part4
- Floors
- #Bathrooms
- …

N number of training data with M featues:

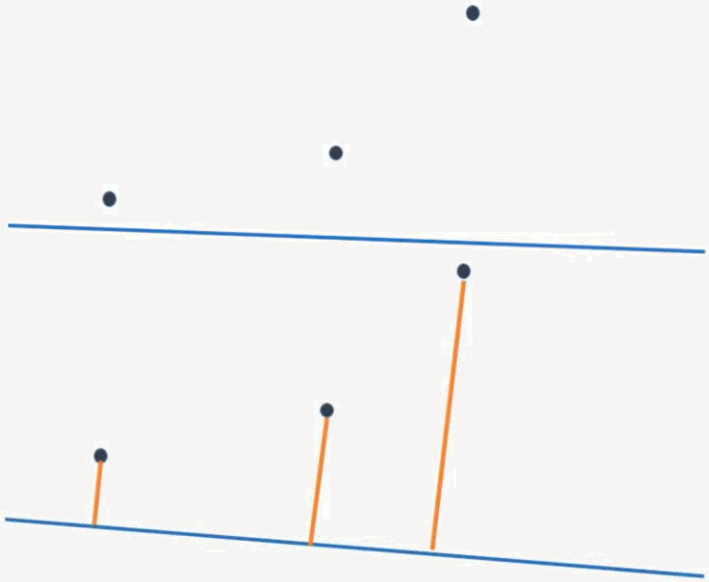
$$A_{N\times M}x_{M\times 1} = b_{N\times 1}$$

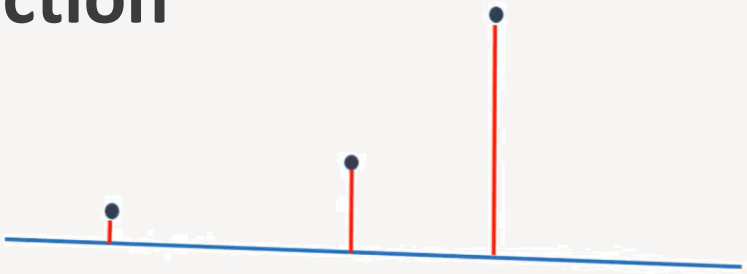
Linear Equation

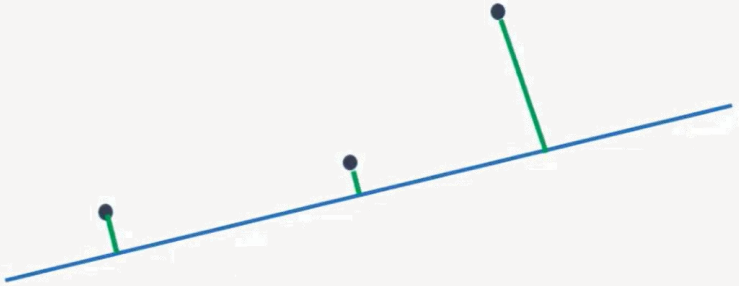
Least Squares Error Correction



Least Squares Error Correction

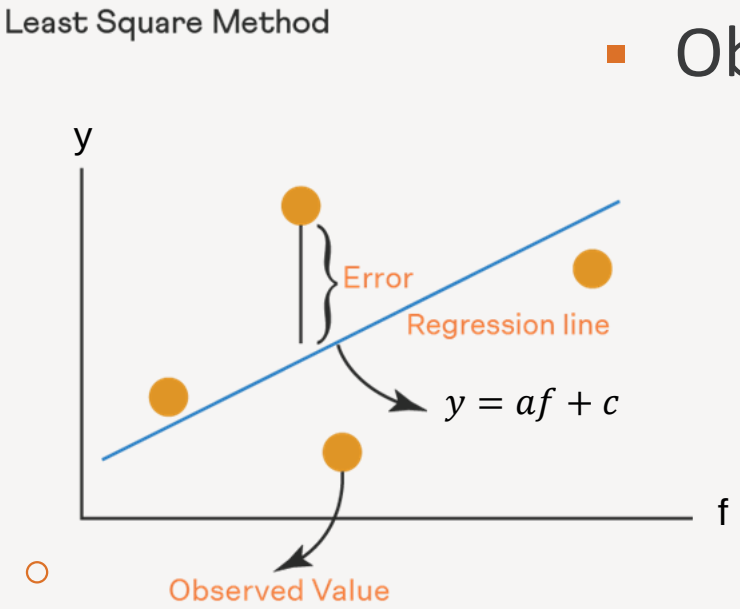






- Error 1:
- Error 2:
- Error 3:

Least Squares



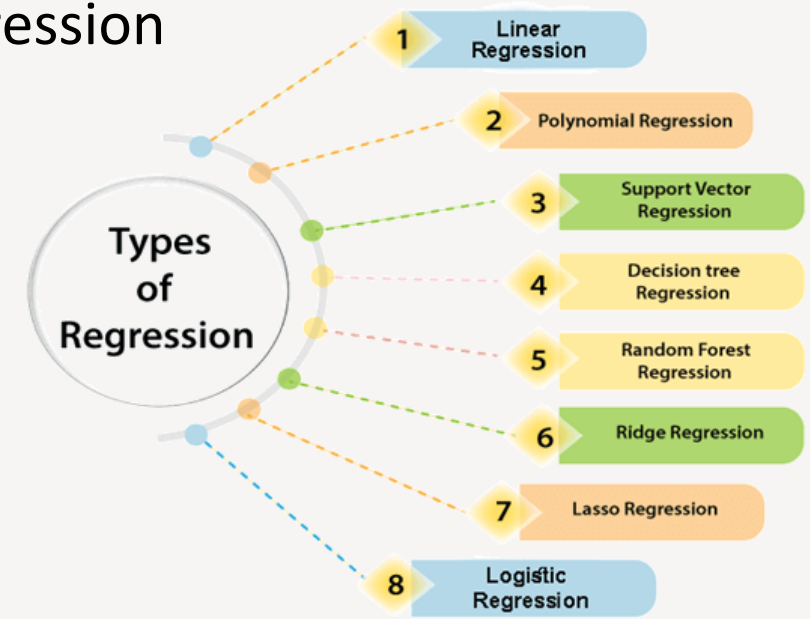
- Objective for “m” features:

$$\hat{y}=a_1 f_1 + a_2 f_2 + \cdots + a_m f_m + c$$

$$\min ||y - \hat{y}||$$

Linear Algebra and Machine Learning Application

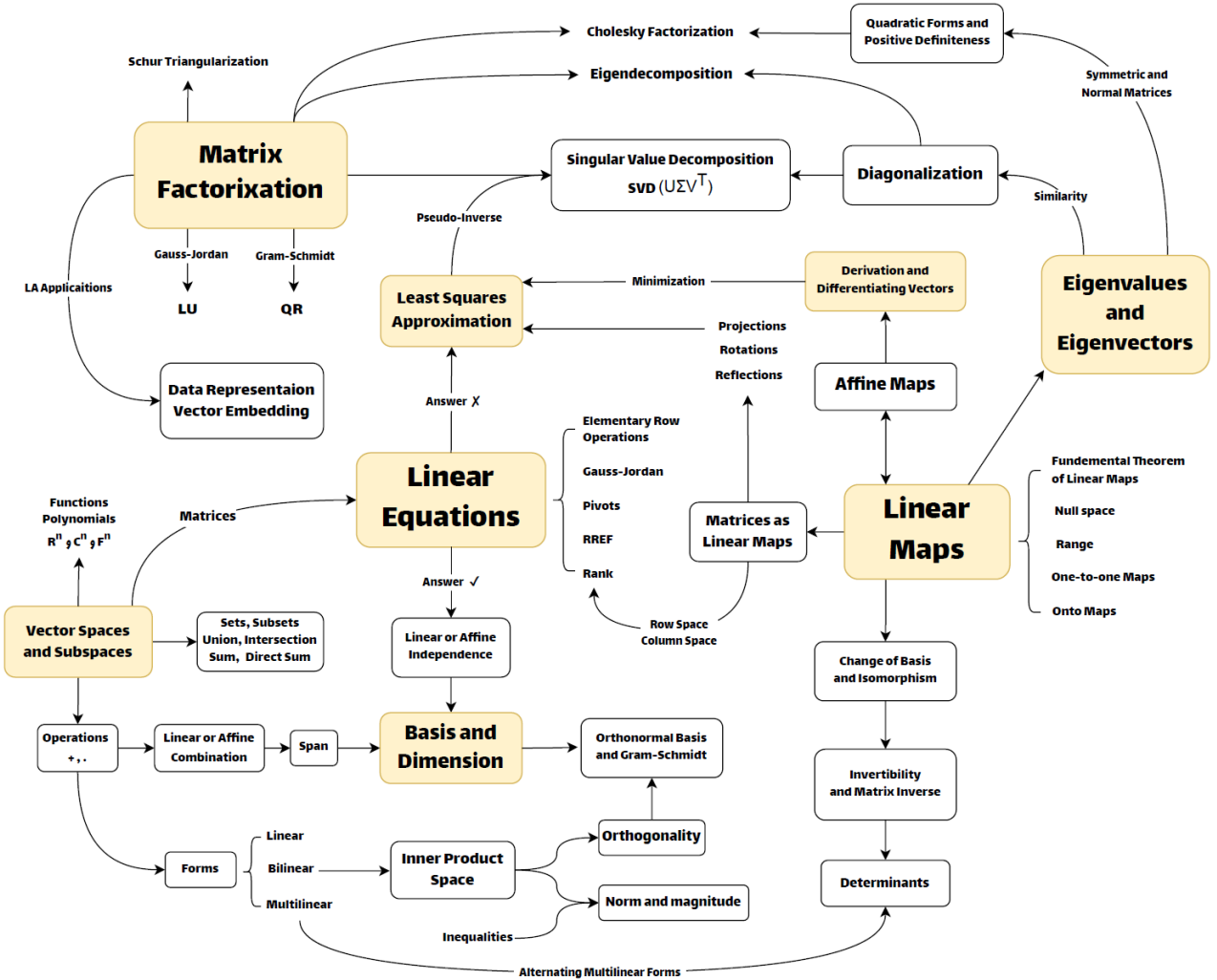
- $Ax = b \quad \rightarrow \quad x = A^{-1}b$ Inverse of Matrix/Pseudo Inverse
- Regression



02

Road Map

Step by Step


Cholesky Factorization
Quadratic Forms and Positive Definiteness
Schur Triangularization
Eigendecomposition
Symmetric and Normal Matrices
Matrix Factorixation
Singular Value Decomposition
SVD $(U\Sigma V^T)$
Diagonalization
Similarity
Pseudo-Inverse
Gauss-Jordan
Gram-Schmidt
LA Applicaitions
LU
QR
Least Squares Approximation
Minimization
Derivation and Differentiating Vectors
Eigenvalues and Eigenvectors
Projections
Rotations
Reflections
Data Representaion Vector Embedding
Affine Maps
Answer ✗
Elementary Row Operations
Gauss-Jordan
Pivots
RREF
Rank
Linear Equations
Matrices as Linear Maps
Linear Maps
Fundemental Theorem of Linear Maps
Null space
Range
One-to-one Maps
Onto Maps
Functions
Polynomials
R^n ,C^n ,F^n
Matrices
Answer ✓
Row Space
Column Space
Vector Spaces and Subspaces
Sets, Subsets
Union, Intersection
Sum, Direct Sum
Linear or Affine Independence
Change of Basis and Isomorphism
Operations
+ , .
Linear or Affine Combination
Span
Basis and Dimension
Orthonormal Basis and Gram-Schmidt
Invertibility and Matrix Inverse
Linear
Bilinear
Multilinear
Forms
Inner Product Space
Orthogonality
Norm and magnitude
Determinants
Inequalities
Alternating Multilinear Forms