Mathematics for Machine Learning

— Introduction

Joseph Chuang-Chieh Lin

Department of Computer Science & Information Engineering,
Tamkang University

Fall 2023

Credits for the resource

- The slides are based on the textbook:
  - *Marc Peter Deisenroth, A. Aldo Faisal, and Cheng Soon Ong: Mathematics for Machine Learning. Cambridge University Press. 2020.*
  - *Howard Anton, Chris Rorres, Anton Kaul: Elementary Linear Algebra. Wiley. 2019.*
- We could partially refer to the monograph:
  Francesco Orabona: A Modern Introduction to Online Learning. https://arxiv.org/abs/1912.13213

Grading Policy

- Attendance (10%)
- Assignments & Quizzes (30%)
- Midterm Exam (30%)
  - 7 Nov. 2023.
- Final Exam (30%)
  - 2 Jan. 2024.

Outline

1 Introduction

Four pillars of ML

- Regression
  - Linear Algebra
  - Vector Algebra
- Dimensionality Reduction
  - Analytic Geometry
  - Probability & Distribution
- Density Estimation
  - Analytic Geometry
  - Probability & Distribution
- Classification
  - Matrix Decomposition
  - Optimization

Four pillars of ML

- Regression
  - Linear Algebra
  - Vector Algebra
- Dimensionality Reduction
  - Analytic Geometry
  - Probability & Distribution
- Density Estimation
  - Analytic Geometry
  - Probability & Distribution
- Classification
  - Matrix Decomposition
  - Optimization

Four pillars of ML

- Regression
  - Linear Algebra
  - Vector Algebra
- Dimensionality Reduction
  - Analytic Geometry
  - Probability & Distribution
- Density Estimation
  - Analytic Geometry
  - Probability & Distribution (quantification of uncertainty)
- Classification
  - Matrix Decomposition
  - Optimization (continuous optimization)

Part I.

Mathematics as the Foundation

Why linear algebra/vector algebra?

- The study of vectors and matrices.

Why linear algebra/vector algebra?

- The study of vectors and matrices.
- Represent numerical data as vectors.

Why linear algebra/vector algebra?

- The study of vectors and matrices.
- Represent numerical data as vectors.
- Represent a table of data as a matrix.

Why linear algebra/vector algebra?

- The study of vectors and matrices.
- Represent numerical data as vectors.
- Represent a table of data as a matrix.
- Formalize the *similarity* between vectors:
  - Analytic geometry (distance, norm, inner product, projection, ...

Why linear algebra/vector algebra?

- The study of vectors and matrices.
- Represent numerical data as vectors.
- Represent a table of data as a matrix.
- Formalize the *similarity* between vectors:
  - Analytic geometry (distance, norm, inner product, projection, ...
- Intuitive interpretation of the data and better efficiency for learning: *matrix decomposition*.

Part II:

Introductory Machine Learning

Topics

- Data, model & parameter estimation.
- Continuous Optimization.
- Linear regression.
  - Map the input $\mathbf{x} \in \mathbb{R}^d$ to corresponding observed function values $y \in \mathbb{R}$.
- Density estimation.
  - Find a probability distribution that describes the data.
- Principal Component Analysis
  - Matrix decomposition.
- Classification.

Terminologies

- i.e. $\Longrightarrow$ that is,
- e.g. $\Longrightarrow$ such as
- $\because \Longrightarrow$ because
- $\therefore \Longrightarrow$ therefore
- et al. $\Longrightarrow$ and others
- $\forall \Longrightarrow$ for any
- $\exists \Longrightarrow$ there exists
- a.k.a. $\Longrightarrow$ also known as
- w.r.t. $\Longrightarrow$ with respect to

Warm-up Exercise

Exercise

- Consider $\mathbf{x} = [x_1 \;\; x_2 \;\; x_3]^\top \in \mathbb{R}^3$ and $\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$.
- Compute $\mathbf{x}^\top \boldsymbol{A}\mathbf{x}$.
- Compute $\text{tr}(\boldsymbol{A}\mathbf{x}\mathbf{x}^\top)$.

Discussions