Marc Peter Deisenroth
A. Aldo Faisal
Cheng Soon Ong

Mathematics For Machine Learning

Notes Written By Jin Hyoung Joo

Contents

Chapter 1
Linear Algebra

1.1 Vectors

Definition 1.1 (Vectors) Vectors are special objects that can be added together and multiplied by scalars to produce another object of the same kind. Any object that satisfies these two properties can be considered a vector.

Definition 1.2 (Groups) A group $(G, \otimes)$ is a set G and an operation $\otimes : G \times G \to G$ where the following properties hold.

1. Closure of G under $\otimes$: $\forall x, y \in G, x \otimes y \in G$
2. Associativity: $\forall x, y, z \in G, (x \otimes y) \otimes z = x \otimes (y \otimes z)$
3. Neutral Element: $\exists e \in G, \forall x \in G, x \otimes e = x$ and $e \otimes x = x$
4. Inverse Element: $\forall x \in G, \exists y \in G, x \otimes y = e \in G$ and $y \otimes x = e$, e is the neutral element.

The group is an Abelian group if $\forall x, y \in G, x \otimes y = y \otimes x$ (commutative).

Definition 1.3 (Vector Spaces) A real-valued vector space $(V, +, \cdot)$ is a set V with two operations, $+ : V \times V \to V$ and $\cdot : \mathbb{R} \times V \to V$ where the following conditions hold.

1.
2.
3.
4.

1.2 Linear Systems

Definition 1.4 (System of Linear Equations) A system of linear equations include multiple linear equations that may lead to a solution, or no solutions, or infinitely many solutions.

$$a_{11}x_1 + \cdots + a_{1n}x_n = b_1$$

$$\vdots$$

$$a_{m1}x_1 + \cdots + a_{mn}x_n = b_m$$

A system of linear equations can be transformed into a matrix form.

$$Ax = b, A \in \mathbb{R}^{m \times n}, x \in \mathbb{R}^{n \times 1}, y \in \, > \times \nVdash$$

1.3 Matrices

Definition 1.5 (Matrix) A real-valued $m \times n$ matrix A is an mn-tuple of elements $a_{ij}, i = 1, \cdots, m, j = 1, \cdots, n$, ordered in a rectangular scheme consisting of m rows and n columns.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, a_{ij} \in \mathbb{R}$$

Operations performed on matrices include:

- Matrix Addition: $A + B$
- Matrix Multiplication: AB
- Hadamard Product (Element-wise multiplication): $(A \odot B)_{ij} = A_{ij}B_{ij}$

Properties of matrices include:

- Associativity:

$$\forall A \in \mathbb{R}^{m \times n}, \forall B \in \mathbb{R}^{n \times p}, \forall C \in \mathbb{R}^{p \times q}, (AB)C = A(BC)$$

- Distributivity:

$$\forall A, B \in \mathbb{R}^{m \times n}, \forall C, D \in \mathbb{R}^{n \times p}, (A + B)C = AC + BC$$
$$A(C + D) = AC + AD$$

- Multiplication with the identity matrix:

$$\forall A \in \mathbb{R}^{m \times n}, I_m A = A I_n = A$$

Definition 1.6 (Inverse Matrix) Given a square matrix $A \in \mathbb{R}^{n \times n}$, the inverse matrix of A is denoted as A^{-1} and has the property of $AB = BA = I_n$. Not all matrices have an inverse, and those that do are called regular/invertible/nonsingular matrices. Otherwise, they are called noninvertible/singular.

Definition 1.7 (Transpose Matrix) Given a square matrix $A \in \mathbb{R}^{n \times n}$, the transpose of A is denoted as $A^\intercal$, and is a matrix that has the rows and columns interchanged ($B_{ij} = A_{ji}$). A matrix is symmetric if $A = A^\intercal$.

Some important properties of inverses and transposes are as below.

$$AA^{-1} = A^{-1}A$$
$$(AB)^{-1} = B^{-1}A^{-1}$$
$$(A^\intercal)^\intercal = A$$
$$(AB)^\intercal = B^\intercal A^\intercal$$
$$(A + B)^\intercal = A^\intercal + B^\intercal$$