The University of Danang
University of Science and Technology

MATHEMATICS FOR COMPUTER SCIENCE

Chap 2. . Linear Algebra

Faculty of Information Technology
PHAM Cong Thang, PhD

References

1. M. P. Deisenroth, A. A. Faisal, and C. S. Ong, Mathematics for Machine Learning, Cambridge University Press; 1 edition (April 23, 2020), 398 pages.
2. E. Lehman, F. T. Leighton, A. R. Meyer, 2017, Mathematics for Computer Science, Eric Lehman Google Inc, 998 pages
3. A. Laaksonen, Competitive Programmer’s Handbook, 2018, 286 pages.
4. W. H. Press, S. A. Teukolsky, W.T. Vetterling, B. P. Flannery Numerical Recipes: The Art of Scientific Computing, Third Edition, Cambridge University Press, 1262 pages.
5. Other online/offline learning resources

Linear Algebra

- **Introduction**
- Eigenvectors and finding Eigenvectors
- Matrix Decompositions

Introduction

- When formalizing intuitive concepts, a common approach is to construct a set of objects (symbols) and a set of rules to manipulate these objects. This is known as an *algebra*.
- Linear algebra is the study of vectors and certain algebra rules to manipulate vectors.
  - The vectors many of us know from school are called "**geometric vectors**", which are usually denoted by a small arrow above the letter, e.g., $\vec{a}$, $\vec{b}$.
  - A bold letter is used to represent them, e.g., $\boldsymbol{a}$ and $\boldsymbol{b}$.
  - In general, vectors are special objects that can be added together and multiplied by scalars to produce another object of the same kind.

Introduction

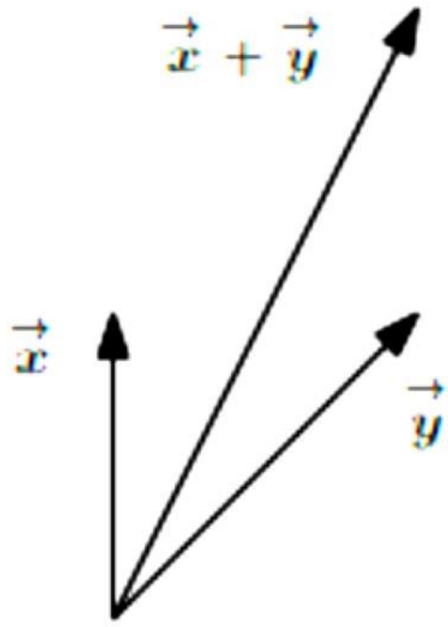
- From an abstract mathematical viewpoint, any object that satisfies these two properties can be considered a vector. Here are some examples of such vector objects:
  - **Geometric vectors**.
  - **Polynomials**
  - **Audio signals are vectors**
  - **Elements of $\mathbb{R}^n$ (tuples of n real numbers) are vectors**

Introduction

- Linear algebra focuses on the similarities between these vector concepts.
  - We can add them together and multiply them by scalars.
  - We will largely focus on vectors in $\mathbb{R}^n$ since most algorithms in linear algebra are formulated in $\mathbb{R}^n$.
  - We often consider data to be represented as vectors in $\mathbb{R}^n$.
  - We will focus on finite-dimensional vector spaces, in which case there is a 1:1 correspondence between any kind of vector and $\mathbb{R}^n$. When it is convenient, we will use intuitions about geometric vectors and consider array-based algorithms.
  - **Linear algebra plays an important role in machine learning and general Mathematics. The concept of a vector space and its properties underlie much of machine learning**

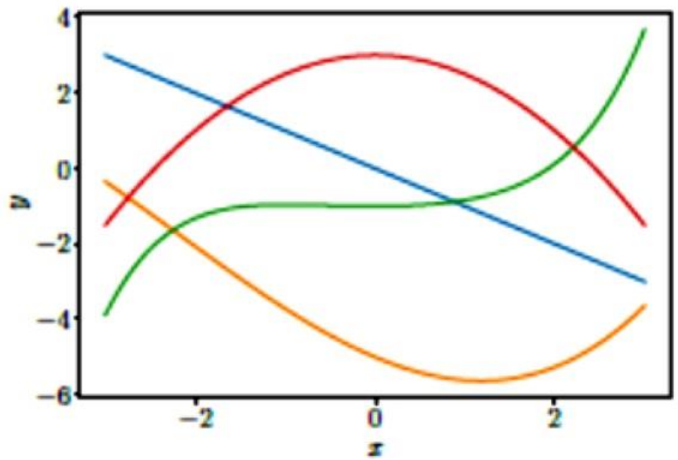
Introduction - vector objects

- **Geometric vectors**. This example of a vector may be familiar from high school mathematics and physics. Geometric vectors are directed segments, which can be drawn (at least in two dimensions).
  - Two geometric vectors $\vec{x}, \vec{y}$ can be added, such that $\vec{x} + \vec{y} = \vec{z}$ is another geometric vector.
  - Multiplication by a scalar $\lambda\vec{x}$ is also a geometric vector. In fact, it is the original vector scaled by λ
  - Interpreting vectors as geometric vectors enables us to use our intuitions about direction and magnitude to reason about mathematical operations.



Introduction - vector objects

- **Polynomials** are also vectors:
  - Two polynomials can be added together, which results in another polynomial
  - They can be multiplied by a scalar $\lambda \in \mathbb{R}$, and the result is a polynomial as well. Therefore, polynomials are (rather unusual) instances of vectors.
  - Note that polynomials are very different from geometric vectors. While geometric vectors are concrete "drawings", polynomials are abstract concepts.



Introduction - vector objects

- **Audio signals** are vectors.
  - Audio signals are represented as a series of numbers.
  - We can add audio signals together, and their sum is a new audio signal.
  - If we scale an audio signal, we also obtain an audio signal.

Introduction - vector objects

- **Elements of $\mathbb{R}^n$ (tuples of n real numbers)** are vectors
  - $\mathbb{R}^n$ is more abstract than polynomials, and it is the concept we focus on in this subject.
  - For instance, $a = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \in \mathbb{R}^3$ is an example of a triplet of numbers. Adding two vectors $a,\ b \in \mathbb{R}^n$ component-wise results in another vector: $a + b = c \in \mathbb{R}^n$
  - Multiplying $a \in \mathbb{R}^n\ by\ \lambda \in \mathbb{R}$ results in a scaled vector $\lambda a \in \mathbb{R}^n$
  - Considering vectors as elements of $\mathbb{R}^n$ has an additional benefit that it loosely corresponds to arrays of real numbers on a computer.
  - **Many programming languages support array operations, which allow for convenient implementation of algorithms that involve vector operations.**

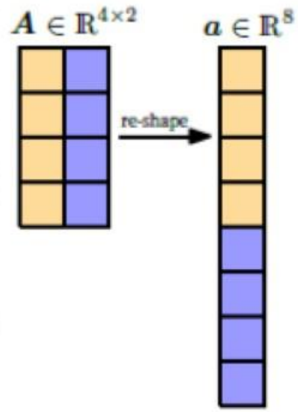
Introduction – Matrix

- Matrices
  - Matrices play a central role in linear algebra.
  - They can be used to compactly represent systems of linear equations, but they also represent linear functions (linear mappings)

Introduction – Matrix

- With $m; n \in N$ a real-valued $(m; n)$ *matrix* A is an $m \cdot n$-tuple of $elements\ a_{ij}$ $, i = 1, \ldots m, j = 1 \ldots n$, which is ordered according to a rectangular scheme consisting of m rows and n columns:

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



  - By convention $(1; n)$ −matrices are called *rows* and $(m; 1)$ −matrices are called column *columns*. These special matrices are also called *row/column vectors*.
  - Rmn is the set of all real-valued $(m; n)$ −matrices. $A \in R^{m\times n}$ can be equivalently represented as a $\in R^{m\times n}$ by stacking all n columns of the matrix into a long vector

Introduction – Matrix operations

- **Matrix Addition and Multiplication:**
  - The sum of two matrices $A \in R^{m\times n}, B \in R^{mn}$ is defined as the elementwise sum, i.e.,

$$\boldsymbol{A}+\boldsymbol{B} := \begin{bmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & & \vdots \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{bmatrix} \in \mathbb{R}^{m\times n}$$

  - For matrices$A \in R^{m\times n}, B \in \mathbb{R}^{n\times k}$, the elements of the product:

$C = AB \in \mathbb{R}^{m\times k}$ are computed as

$$c_{ij} = \sum_{l=1}^{n} a_{il} b_{lj}, \qquad i = 1, \ldots, m, \quad j = 1, \ldots, k$$

Introduction – Matrix operations

- **Inverse and Transpose:**
  - **Inverse:** Consider a square matrix $A \in \mathbb{R}^{n \times n}$. Let matrix $B \in \mathbb{R}^{n \times n}$ have the property that $AB = I_n = BA$. B is called the inverse of A and denoted by A^{-1}.
  - Unfortunately, not every matrix A possesses an inverse A^{-1}.
    - If this regular inverse does exist, A is called *regular/invertible/nonsingular*,
    - otherwise *singular/noninvertible*.

Introduction – Matrix operations

- **Inverse and Transpose:**
  - **Transpose:** For $A \in \mathbb{R}^{m\times n}$ the matrix $B \in \mathbb{R}^{m\times n}$ with $b_{ij} = a_{ji}$ is called the *transpose* of A. We write $B = A^T$
    - In general,A can be obtained by writing the columns of A as the rows of A
    - *Matrix $A \in \mathbb{R}^{n\times n}$ is symmetric if $A = A^T$*
    - *The **sum of symmetric matrices is always symmetric**. However, **their product is always defined, it is generally not symmetric***
- **Multiplication by a Scalar:** Matrices are multiplied by a scalar $\lambda \in \mathbb{R}$.
  - Let $A \in \mathbb{R}^{m\times n}$ and $\lambda \in \mathbb{R}$ then $\lambda A = K$, $K_{ij} = \lambda\, a_{ij}$. Practically, λ scales each element of A

Introduction – Matrix operations

- The determinant $\det(A)$ or $|A|$ of a square matrix A:
  - A number encoding certain properties of the matrix.
  - A matrix is invertible if and only if its determinant is nonzero.
  - The determinant of a product of square matrices equals the product of their determinants: $\det(AB) = \det(A) \cdot \det(B$
  - The determinant of a matrix tell us whether or not the matrix is invertible, can be used to find a formula for the inverse of a matrix, and give us a method for solving linear equations
  - If A has one row of zeros then $detA = 0$
  - If A is a triangular matrix then $detA$ is the product of the entries on the main diagonal

Introduction – Matrix operations

- The determinant $\det(A)$ or $|A|$ of a square matrix A:
  - Determinants are important concepts in linear algebra. A determinant is a mathematical object in the analysis and solution of systems of linear equations.
  - Determinants are only defined for square matrices $A \in \mathbb{R}^{n \times n}$, i.e., matrices with the same number of rows and columns.

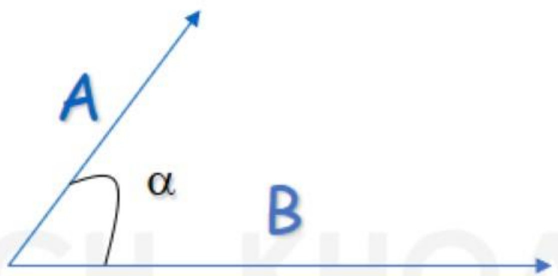
$$\det(\boldsymbol{A}) = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

Introduction – Matrix operations

- The determinant $\det(A)$ or $|A|$ of a square matrix A:
  - Adding a multiple of any row to another row, or a multiple of any column to another column, does not change the determinant.
  - Interchanging two rows or two columns affects the determinant by multiplying it by -1.
  - Multiplying one row (or column) by a nonzero number k affects the determinant by multiplying it by k.

Introduction – Vector operations

- Inner (dot) Product: $v \cdot w$
  - The inner product is a **SCALAR**
  - $v \cdot w = (x_1, x_2) \cdot (y_1, y_2) = x_1 y_1 + x_2 y_2$
  - $v \cdot w = \|v\| \, \|v\| \cos(\alpha)$
  - $v \cdot w = 0 \iff v \perp w$
  - vectors $v \cdot w$ are "**columns**", then dot product is $w^T v$
  - $v \cdot w = (x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$
  - $\lambda \cdot v = \lambda(x_1, x_2) = (\lambda x_1, \lambda x_2)$



Introduction – Vector operations

- Dot Product
  - The dot product as a matrix multiplication

$$A \cdot B = A^T B = [a \quad b \quad c] \begin{bmatrix} d \\ e \\ f \end{bmatrix} = ad + be + cf$$

  - The dot product of a vector with itself $\|A\|^2 = A^T A = aa + bb + cc$
  - The dot product is also related to the angle between the two vectors

$$A \cdot B = \|A\| \, \|B\| \cos(\alpha)$$

Introduction – Vector operations

- Norms: the norm of a vector $\|x\|$
  - Absolutely homogeneous: $\lambda\|x\| = \|\lambda\|\|x\|$
  - Triangle inequality: $\|x + y\| \leq \|x\| + \|y\|$
  - Positive definite: $\|x\| \geq 0$ and $\|x\| = 0 \Leftrightarrow x = 0$
  - The Manhattan norm on $\mathbb{R}^n$ is defined for $x \in \mathbb{R}^n$ (is also called l_1 norm)

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|,$$

where $\|\cdot\|$ is the absolute value, The Manhattan norm is also called l_1 norm

  - Euclidean Norm on $\mathbb{R}^n$ is defined for $x \in \mathbb{R}^n$ (is also called l_2 norm)

$$\|x\|_1 = \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{x^T x}$$

Introduction – Vector operations

- Norms:
  - l_p norm of x is defined is ($p \geq 1$):

$$\|x\|_p = \sqrt[p]{\sum_{i=1}^{n} |x_i|^p}$$

  - l_∞ norm (maximum norm or uniform norm) of x is defined

$$\|x\|_p = \max\{|x_1|, |x_2|, \dots, |x_n|\}$$

Introduction – Vector operations

- Vector Addition:
  - $A + B = (x_1, x_2) + (y_1 + y_2) = (x_1 + x_2, y_1 + y_2)$
- Scalar Product:
  - $\lambda A = \lambda(x_1, x_2) = (\lambda x_1, \lambda x_2)$

Introduction - Matrix Norms

- Common matrix norms for a matrix A

column-sum norm $\|A\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{n} |a_{ij}|$

Frobenius norm

row-sum norm $\|A\|_\infty = \max_{1 \le i \le n} \sum_{j=1}^{n} |a_{ij}|$

spectral norm (2 norm) $\|A\|_2 = (\mu_{\max})^{1/2}$

μ_{max} is the largest eigenvalue of $A^T A$.

Introduction – Vector operations

- Inner Products
  - Inner products allow for the introduction of intuitive geometrical concepts, such as the length of a vector and the angle or distance between two vectors.
  - A major purpose of inner products is to determine whether vectors are orthogonal to each other.
- *Cross product or vector product is an operation on two vectors in three-dimensional space*