Chapter 1: Introduction and Mathematical Foundations

Dr. Youssef SALMAN
Lebanese University - Faculty of Sciences - Section V

Applied Mathematics & Statistics for Computational Sciences

1. Overview

Mathematics serves as the language of modern computational sciences, especially in fields such as Artificial Intelligence (AI) and Machine Learning (ML). This chapter lays the groundwork for understanding how fundamental mathematical concepts support and shape AI/ML models. We begin by reviewing algebraic principles and then move to more structured objects such as vector spaces.

2. Overview on Artificial Intelligence and Machine Learning

Artificial Intelligence (AI) is a branch of computer science focused on building systems capable of performing tasks that typically require human intelligence. These tasks include learning, reasoning, problem-solving, perception, and natural language understanding. AI systems aim to simulate or augment human decision-making in environments that involve uncertainty, complexity, or vast amounts of data (Check Weiss (1999) for more info).

Machine Learning (ML) is a subfield of AI that enables systems to automatically learn patterns and improve from experience without being explicitly programmed. Instead of relying on hard-coded rules, ML algorithms infer relationships and make predictions from data through mathematical modeling and statistical analysis.

In essence, while AI is the broader discipline concerned with intelligent behavior, ML focuses specifically on data-driven methods that allow computers to learn and adapt (Check Bishop and Nasrabadi (2006) for more info).

3. The Role of Mathematics in AI and ML

Mathematics plays a central role in AI and ML for modeling data, training algorithms, evaluating performance, and making predictions. For instance, linear algebra is the foundation of neural networks and dimensionality reduction techniques; calculus is crucial for optimization methods like gradient descent; and probability theory underpins models that make decisions under uncertainty.

As an example, consider a facial recognition system:

- Images are represented as high-dimensional vectors (linear algebra).
- The system must learn to minimize error in recognition (calculus).
- It needs to quantify how confident it is in a prediction (probability and statistics).

4. Refresher: Algebraic Foundations

4.1. Functions and Equations

A function is a rule that assigns to each input exactly one output. In machine learning, models are functions mapping features to predictions.

Definition (Function): A function f from a set X to a set Y is written as $f : X \to Y$ such that for each $x \in X$, there is a unique $y \in Y$ where $y = f(x)$.

Example: A linear function $f(x) = 3x + 2$ maps real numbers to real numbers.

Application: In linear regression, we model the relationship between an input variable x (e.g., years of experience) and an output y (e.g., salary) with a linear function $y = ax + b$.

Quadratic and exponential functions also appear frequently:

- Quadratic: $f(x) = ax^2 + bx + c$, used in optimization and control.
- Exponential: $f(x) = ae^{bx}$, models population growth or learning rate decay.

4.2. Solving Equations

Solving equations is fundamental in finding model parameters.

- **Linear:** $3x + 2 = 0 \Rightarrow x = -\frac{2}{3}$
- **Quadratic:** $ax^2 + bx + c = 0$ has solutions $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$

5. Vectors and Vector Operations

In computational sciences, data points are often represented as vectors.

Definition (Vector): A vector in $\mathbb{R}^n$ is an ordered list of n real numbers, typically written as $\mathbf{v} = [v_1, v_2, \ldots, v_n]$.

Operations:

- Addition: $\mathbf{a} + \mathbf{b} = [a_1 + b_1, \ldots, a_n + b_n]$
- Scalar multiplication: $c \cdot \mathbf{v} = [cv_1, cv_2, \ldots, cv_n]$
- Dot product: $\mathbf{a} \cdot \mathbf{b} = \sum a_i b_i$

Real-World Example: In a movie recommendation system, each user and movie can be represented as vectors of preferences or features. The dot product between a user vector and a movie vector indicates the predicted rating.

6. Vector Spaces

6.1. Definition and Axioms

A vector space is a mathematical structure formed by a collection of vectors that can be added together and scaled by numbers.

Definition (Vector Space): A vector space V over a field $\mathbb{F}$ (typically $\mathbb{R}$) is a set equipped with:

- Vector addition: $\mathbf{u} + \mathbf{v} \in V$
- Scalar multiplication: $a \cdot \mathbf{v} \in V$ for all $a \in \mathbb{F}$, $\mathbf{v} \in V$

These operations must satisfy axioms such as associativity, distributivity, and existence of additive identity and inverses.

Example: The set $\mathbb{R}^3$ (3D vectors) is a vector space over $\mathbb{R}$.

6.2. Linear Independence, Basis, and Dimension

- Linear Independence: A set of vectors is linearly independent if no vector in the set is a linear combination of the others.
- Basis: A basis of a vector space is a linearly independent set of vectors that spans the space.
- Dimension: The number of vectors in a basis.

Example: In $\mathbb{R}^2$, the vectors $(1, 0)$ and $(0, 1)$ form a basis. The dimension is 2.
Application: In data science, PCA (Principal Component Analysis) finds a basis of a lower-dimensional subspace capturing most data variance.

7. Conclusion

This chapter has provided a structured review of the foundational mathematical concepts that underpin much of AI and ML. From basic algebraic equations to the structure of vector spaces, understanding these concepts is crucial for implementing and interpreting machine learning algorithms. Next, we transition to the Python environment, where these concepts will be explored through practical, hands-on coding exercises.

References

Bishop, C. M. and N. M. Nasrabadi (2006). *Pattern recognition and machine learning*, Volume 4. Springer.

Weiss, G. (1999). *Multiagent systems: a modern approach to distributed artificial intelligence*. MIT press.