Mathematics for Machine Learning

Samuel Ireson
samuel-a.ireson@db.com

Klejdi Sevdari
klejdi-a.sevdari@db.com

Finlay Pillar
finlay-pillar@db.com

Wing Lam Wong
wing-lam-a-wong@db.com

October 30, 2024

1. INTRODUCTION

We discuss motivations for studying the mathematics behind machine learning, give an outline of the content of the course, and mention the references which were used in putting it together.

Machine learning is a term which is commonly heard and spoken, and one less commonly understood. This is in part due to the rapid growth in experimental machine learning techniques, and the lagging in mathematical explanations of these techniques. Another prominent reason is that the mathematical understanding of machine learning can, in enterprise settings, be considered as 'less useful' or 'less relevant'. Some people may also perceive the inner workings of machine learning to be a 'black box', and a subject which they couldn't understand.

We will focus on tackling the final obstruction. The perception of the value of understanding finds itself in the hands of others, but the understanding in question is firmly in our own. There are only a few mathematical concepts needed to gain a working understanding of machine learning, and we will cover these as they are needed.

We hope to present a self-contained course, covering the key topics in detail, gradually working towards popular models and algorithms which are seen in real-world machine learning applications. We will place a heavy focus on examples, ones which are hopefully relevant and familiar to the reader, rooting the work we are completing in the context of its application.

For those interested in reading more than we have time to cover, we recommend the work of Schwartz et al.[3], Bishop[1], and Faisal et al.[2], all of which were to a lesser or greater extent consulted in the writing of this course.

1.1 WHAT IS LEARNING?

Sometimes, it is those concepts of which we have an intuitive idea which are hardest to rigorously define. What does it mean to learn? What does it mean to have learned? What does it mean to understand? These are all relevant questions in the theory of machine learning, especially the first, and we will first consider them in the context of human learning.

In the loosest sense, *learning* is the process of converting past experiences into applicable knowledge. For humans, the input to this learning can come in any number of formats. For example, we could consider the knowledge obtained from reading a book. What exactly is the process occurring in the mind of the reader which allows them to retain, and importantly, re-apply this information. We note that the process is decidedly not memorisation. As a child who has memorised their times tables cannot claim the ability of multiplication, the reader who cannot call upon knowledge in unseen circumstances cannot claim to fully understand what they have read. This gives a good hint as to an important facet of learning systems; *generalisation.*

For machines, the process is not dissimilar. The input for a learning algorithm is data, and the output (the expertise) are some tuned parameters which allow the model produce outputs from unseen data in the future. There is more nuance to this, but as we move towards mathematical rigour, it will be useful to keep this in mind.

1.2 STRUCTURE OF THE COURSE

As previously mentioned, there are only few mathematical concepts which are absolutely necessary for understanding machine learning. An overview of the structure of this course is as follows,

- Linear Algebra

- Multivariable calculus

- Probability

- Optimisation

- Linear regression

- Logistic regression

Those familiar with the above topics may note that this course is structured in a 'bottom-up' fashion – we build on the theory as we move forwards. This is somewhat in contrast to other machine learning courses which take a 'top-down' approach in which the theory behind the application is uncovered as it is encountered. Both of these strategies have their strengths, and courses which are mathematically inclined typically take the same approach as us. It is however necessary to have patience with this approach. If you have previously only encountered machine learning in a practical, hands-on setting, you may struggle to understand how the early chapters of this course fit into the wider picture, but eventually all will be clear.

2. Linear Algebra

3. Multivariable Calculus

3.1 Introduction

4. Probability

Test

5. OPTIMISATION THEORY

6. Linear Regression

7. Logistic Regression