Free Mathematics Courses for Data Science & Machine Learning

<= Previous post
Next post =>

Tags: Courses, Data Science, Machine Learning, Mathematics, MOOC

It's no secret that mathematics is the foundation of data science. Here are a selection of courses to help increase your maths skills to excel in data science, machine learning, and beyond.

By Matthew Mayo, KDnuggets.

comments

Are you interested in learning the foundations to a successful data science career? Or are you looking to brush up on your maths, or strengthen your understanding by extending that base?

This is a selection of maths courses, collections of courses, and specializations which are freely available online, and which can help achieve your data science mathematics goals. They have been separated into the broad topics of mathematical foundations, algebra, calculus, statistics & probability, and those especially relevant to data science & machine learning.

Take a look at the list and closer inspect those which may be of interest to you. I hope you find something useful.

Original image by Chuk Yong from Pixabay

Mathematical Foundations

These courses are intended to help lay the foundation for learning more advanced maths, as well as foster the development of mathematical thinking. Descriptions come directly from the respective course websites.

- **Introduction to Logic, Stanford** (course)
  This course is an introduction to Logic from a computational perspective. It shows how to encode information in the form of logical sentences; it shows how to reason with information in this form; and it provides an overview of logic technology and its applications - in mathematics, science, engineering, business, law, and so forth.

- **Introduction to Mathematical Thinking, Stanford** (course)
  Professional mathematicians think a certain way to solve real problems, problems that can arise from the everyday world, or from science, or from within mathematics itself. The key to success in school math is to learn to think inside-the-box. In contrast, a key feature of mathematical thinking is thinking outside-the-box – a valuable ability in today's world. This course helps to develop that crucial way of thinking.

- **High School Mathematics, MIT** (collection of courses)
  In this section we have provided a collection of mathematics courses and resources from across MIT. Some are materials that were used to teach MIT undergraduates, while others were designed specifically for high school students.

Algebra

These algebra courses run the gamut from introductory algebra to linear models and matrix algebra. Algebra is helpful in computation and data science generally, and encompasses some of the main concepts in powering some machine learning algorithms, including neural networks. Descriptions come directly from the respective course websites.

- **Algebra I, Khan Academy** (course)
  Course covers algebra foundations, solving equations & inequalities, working with units, linear equations & graphs, forms of linear equations, systems of equations, inequalities (systems & graphs), functions, sequences, absolute value & piecewise functions, exponents & radicals, exponential growth & decay, quadratics (multiplying & factoring), quadratic functions & equations, irrational numbers.

- **Algebra II, Khan Academy** (course)
  Course covers polynomial arithmetic, complex numbers, polynomial factorization, polynomial division, polynomial graphs, rational exponents & radicals, exponential models, logarithms, transformations of functions, equations, trigonometry, modeling, rational functions.

- **Linear Algebra, MIT** (course)
  This is a basic subject on matrix theory and linear algebra. Emphasis is given to topics that will be useful in other disciplines, including systems of equations, vector spaces, determinants, eigenvalues, similarity, and positive definite matrices.

- **Linear Algebra - Foundations to Frontiers, University of Texas at Austin** (course)
  Through short videos, exercises, visualizations, and programming assignments, you will study Vector and Matrix Operations, Linear Transformations, Solving Systems of Equations, Vector Spaces, Linear Least-Squares, and Eigenvalues and Eigenvectors. In addition, you will get a glimpse of cutting edge research on the development of linear algebra libraries, which are used throughout computational science.

- **Introduction to Linear Models and Matrix Algebra, Harvard** (course)
  In this introductory online course in data analysis, we will use matrix algebra to represent the linear models that commonly used to model differences between experimental units. We perform statistical inference on these differences. Throughout the course we will use the R programming language to perform matrix operations.

Calculus

These calculus courses cover topics from preparatory precalculus through to differentiation, integration, to multivariate calculus and differential equations. Calculus has broad uses, generally, and contains core concepts which power neural networks work. Descriptions come directly from the respective course websites.

- **Precalculus, Khan Academy** (course)
  Course covers complex numbers, polynomials, composite functions, trigonometry, vectors, matrices, series, conic sections, probability and combinatorics.

- **Calculus 1, Khan Academy** (course)
  Course covers limits and continuity, derivatives: definitions and basic rules, derivatives: chain rule and other advanced topics, applications of derivatives, analyzing functions, integrals, differential equations, applications of integrals.

- **Calculus 2, Khan Academy** (course)
  Course covers integrals review, integration techniques, differential equations, applications of integrals, parametric equations, polar coordinates, and vector-valued functions, series.

- **Multivariable calculus, Khan Academy** (course)
  Course covers thinking about multivariate functions, derivatives of multivariate functions, applications of multivariate derivatives, integrating multivariate functions, Green's, Stokes', and the divergence theorems.

- **Differential equations, Khan Academy** (course)
  Course covers first order differential equations, second order differential equations, Laplace transform.

- **Introduction to Calculus, University of Sydney** (course)
  The focus and themes of the Introduction to Calculus course address the most important foundations for applications of mathematics in science, engineering and commerce. The course emphasises the key ideas and historical motivation for calculus, while at the same time striking a balance between theory and application, leading to a mastery of key threshold concepts in foundational mathematics.

Statistics & Probability

Statistics and probability are the foundations of data science, more so than any other family of mathematical concepts. These courses will help prepare you to look at data through the statistical lens and with a critical probabilistic eye. Descriptions come directly from the respective course websites.

- **Statistics and probability, Khan Academy** (course)
  Course covers analyzing categorical data, displaying and comparing quantitative data, summarizing quantitative data, modeling data distributions, exploring bivariate numerical data, study design, probability, counting, permutations, and combinations, random variables, sampling distributions, confidence intervals, significance tests, two-sample inference for the difference between groups, inference for categorical data, advanced regression, analysis of variance

- **Fundamentals of Statistics, MIT** (course)
  Statistics is the science of turning data into insights and ultimately decisions. Behind recent advances in machine learning, data science and artificial intelligence are fundamental statistical principles. The purpose of this class is to develop and understand these core ideas on firm mathematical grounds starting from the construction of estimators and tests, as well as an analysis of their asymptotic performance

- **Data Science: Probability, Harvard** (course)
  We will introduce important concepts such as random variables, independence, Monte Carlo simulations, expected values, standard errors, and the Central Limit Theorem. These statistical concepts are fundamental to conducting statistical tests on data and understanding whether the data you are analyzing is likely occurring due to an experimental method or to chance.

- **Probability - The Science of Uncertainty and Data, MIT** (course)
  The course covers all of the basic probability concepts, including: multiple discrete or continuous random variables, expectations, and conditional distributions, laws of large numbers, the main tools of Bayesian inference methods, an introduction to random processes (Poisson processes and Markov chains)

- **Improving your statistical inferences, Eindhoven University of Technology** (course)
  First, we will discuss how to correctly interpret p-values, effect sizes, confidence intervals, Bayes Factors, and likelihood ratios, and how these statistics answer different questions you might be interested in. Then, you will learn how to design experiments where the false positive rate is controlled, and how to decide upon the sample size for your study, for example in order to achieve high statistical power. Subsequently, you will learn how to interpret evidence in the scientific literature given widespread publication bias, for example by learning about p-curve analysis. Finally, we will talk about how to do philosophy of science, theory construction, and cumulative science, including how to perform replication studies, why and how to pre-register your experiment, and how to share your results following Open Science principles.

- **Introduction to Probability and Data, Duke University** (course)
  This course introduces you to sampling and exploring data, as well as basic probability theory and Bayes' rule. You will examine various types of sampling methods, and discuss how such methods can impact the scope of inference. A variety of exploratory data analysis techniques will be covered, including numeric summary statistics and basic data visualization. You will be guided through installing and using R and RStudio (free statistical software), and will use this software for lab exercises and a final project. The concepts and techniques in this course will serve as building blocks for the inference and modeling courses in the Specialization.

- **Probability Theory and Mathematical Statistics, Penn State** (course)
  Courseware for a pair of related courses covers introduction to probability, discrete distributions, continuous distributions, bivariate distributions, ditributions of functions of random variables, estimation, hypothesis testing, nonparametric methods, bayesian methods, and more.

Mathematics for Data Science & Machine Learning

These are mathematics topics directly related to data science and machine learning. They may include material from courses above, and may also be more elementary than some of above as well. However, they can be useful for brushing up on material you may not have studied in a while, and which is especially pertinent to the practice of data science. Descriptions come directly from the respective course websites.

- **Data Science Math Skills, Duke University** (course)
  Data science courses contain math—no avoiding that! This course is designed to teach learners the basic math you will need in order to be successful in almost any data science math course and was created for learners who have basic math skills but may not have taken algebra or pre-calculus. Data Science Math Skills introduces the core math that data science is built upon, with no extra complexity, introducing unfamiliar ideas and math symbols one-at-a-time.

- **Essential Math for Machine Learning: Python Edition, Microsoft** (course)
  This course is not a full math curriculum; it's not designed to replace school or college math education. Instead, it focuses on the key mathematical concepts that you'll encounter in studies of machine learning. It is designed to fill the gaps for students who missed these key concepts as part of their formal education, or who need to refresh their memories after a long break from studying math.

- **Mathematics for Machine Learning, Imperial College London** (specialization)
  For a lot of higher level courses in Machine Learning and Data Science, you find you need to freshen up on the basics in mathematics - stuff you may have studied before in school or university, but which was taught in another context, or not very intuitively, such that you struggle to relate it to how it’s used in Computer Science. This specialization aims to bridge that gap, getting you up to speed in the underlying mathematics, building an intuitive understanding, and relating it to Machine Learning and Data Science.