Data Science Math Skills

Katherine G. Pe

Contents

Data Science Math Skills

Course Information

Data Science Math Skills by **Duke University** is a online course you can take on this site Coursera.

Motivation for Learning & Re-learning

A lot of graduate school students struggle with Data Science courses only because of their lack of knowledge and/or understanding of Mathematics for Data Science. The course gives an overview of Mathematical concepts you will encounter while learning Data Science.

Supplemental Notes and Videos

Here's how I make my notes:

My notes include videos from Khan Academy and other websites. The content's the same, and often a bit better due to lack of errors. The text are usually from the Coursera video transcripts.

I indicate **why** it is important to learn such concepts through **Further Reading** notes.

Sets and What They're Good For

Set Basics and Vocabulary

- Set Theory
- Set Theory Operations

Further Reading

> A set is the fundamental discrete structure on which all other discrete structures are built.

Those who studied Discrete Mathematics or read a book about it will probably just re-learn a lot from this course on Set Basics.

- Applications of Set Theory in Computer Science - A list of the most obvious applications of Set Theory.
- Discrete Mathematics and Its Applications - I read most of the book as a supplemental material for a Discrete Math course. The book clearly states why a set is the foundational structure in Computer Science.

Venn Diagrams

- What are Venn diagrams?

Further Reading

- A Visual Explanation of SQL Joins

The Infinite World of Real Numbers

- What are Real Numbers?
- Multi-step Inequalities

The Jagged S Symbol

- Sigma/Summation Notation

  $\sum_{n=1}^{10} n^2$
- The Sigma has similarities to the factorial symbol, but it suggests that you add the values of i based on the stopping point n

Further Reading

- Graph-based Machine Learning

Descartes Was Really Smart

Plotting Points

- The x-axis is going to be the set of all points x-y in the Cartesian plane, x-y in R2, such that their y coordinate is zero.
- We divide the Cartesian plane into four separate regions, and these we call **quadrants**.
- Coordinate plane: quadrants

Distance Formula

- Distance formula

  $d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$

Point-Slope Formula for Lines

- Point Slope Form

  $y - b = m(x - a)$
- Calculating the Slope

  $(y2 - y1/x2 - x1)$

Slope-Intercept Formula

- Slope Intercept Equation
- Intro to Slope Intercept Form

  $y = mx + b$

Input-Output Machines

Functions: Mapping from Sets to Sets

- Functions on Sets - UCLA

  $f : A \to B$

Functions: Graphing in the Cartesian Plane

- Functions on a Coordinate Plane
- Vertical line test
- The **vertical line** test says: *any vertical line, intersects the graph of a function once.* If it intersects it more than once, we violate things here.

Increasing and Decreasing Functions

- A function is increasing whenever:

  $a < b$

  $f(a) < f(b)$

  $f(x) = 2^x$ – An example of an increasing function
- A function is decreasing whenever:

  $a > b$

  $f(a) > f(b)$

  $g(x) = 3^{-x}$ – An example of a decreasing function

Composition and Inverse

- How to find the inverse of a composite function
- Evaluating Composite functions
- Not every function has an inverse.
- If the graph of f fails the *horizontal line test*, the graph has no inverse.

This is about that derivative stuff

Tangent Lines - Slope of a Graph at a Point

- How to find the slope & equation of tangent line at a given point
- Finding the Equation Of A Tangent Line
- A **Tangent Line** is a line which locally touches a curve at one and only one point.
- Formulas:

  $f'(a)$

Tangent Lines - The Derivative Function

- The derivative & tangent line equations
- Formulas:

  $f'(x) = 2x$

  $f(x) = x^2$

Fast Growth, Slow Growth

Using Integer Exponents

- 0 as an exponent:

  $2^0 = 1$
- Negative exponents:

  $x^{-n} = 1/x^n$
- Scientific Notation

Simplification Rules for Algebra using Exponents

- Multiplication rule

  $x^n x^m = x^{(n+m)}$
- Power to a power

  $(x^n)m = x^{nm}$
- Product to a power

  $(xy)^n = x^n y^n$
- Fraction to a power

  $(x/y)^n = x^n/y^n$
- Division and negative powers

  $x^n/x^n = x^{n-m}$
- Fractional exponents

  $x^(a/b) =^b \sqrt{x^a}$

How Logarithms & Exponents are Related

- Intro to logarithms
- Exponential form:

  $b^x = N$
- Logarithmic form:

  $x = log_b = N$
- The logarithm of any base of 1 is 0, just as any number raised to 0 is 1.

General rules

- Product rule

  $log(xy) = log(x) + log(y)$

- Quotient rule

  $log\frac{x}{y} = log(x) - log(y)$

- Power and root rule

  $log(x^n) = nlog(x)$

Change of Base

- The base of **10**, **2** and the natural log *(e)* are common in Data Science.
- Formula for changing the base:

  $log_a(b) = \frac{log_x(b)}{log_x(a))}$

The Rate of Growth of Continuous Processes

- Exponential rate of growth can be a **discrete** exponential rate of growth or a **continuous** exponential rate of growth.
- Note: those who studied Finance will likely see the practicality of knowing, but not of the terms of themselves.
- Discrete rate of growth:

  $1.00(1 + r)t$

- Euler's contstant e
- What is e?

  2.71828

- Euler is pronounced *Oyler*. He's a mathematician most famously known of his solution to the graph theory problem: Seven Bridges of Königsberg
- Problem: A baby elephant weighing 200 kg grows at a continuously compounded rate of 5%/year. How much does it weigh in 3 years?

  $(200kg)e^{(0.05)(3)} = 232.4kg$