Artificial Intelligence and Machine Learning (AIML)

2023–24

- **Last lecture**: introduced the concepts of AI and ML with some illustrated examples
- **This lecture**: essential mathematical background

Essential mathematical background

- To study AI and ML, you need **minimal background knowledge** of mathematics, just enough to understand the concepts and algorithms
- You may have some of this background already; this section is included for those without the minimal background and/or if you have forgotten the background mathematics
- Other, more specialized mathematical concepts, notation and terminology will be introduced later in the lecture material where needed
- This lecture will necessarily have to be a quick overview: you can find good introductions in the background mathematics reading list

Arithmetic expressions and rules of arithmetic

- **Expressions** and **evaluation rules** give precise meaning to arrangements of symbols and calculations with those symbols; the **order of operations** (addition, subtraction, multiplication, division, expanding brackets) is critical:

$50 + 12 - (7 \times 4^2)$ (power)
$= 62 - (7 \times 4 \times 4)$ (multiply)
$= 62 - (7 \times 16)$ (multiply again)
$= 62 - 112$ (multiply and remove brackets)
$= -50$ (subtraction)

AIML

Algebraic expressions and the rules of algebra

- Substituting a **letter** or **symbol** in place of a **number**, allows us to form **abstract expressions** which refer to quantities without knowing the actual value of the quantity
- The abstract **rules of algebra** are:

$a+b=b+a$	$a\times b=b\times a$	**(commutativity)**
$a+0=0+a=a$	$a\times 1=1\times a=a$	**(units)**
$a+(b+c)=(a+b)+c$	$a\times(b\times c)=(a\times b)\times c$	**(associativity)**
$a\times(b+c)=a\times b+a\times c$	$(a+b)\times c=a\times c+b\times c$	**(distributivity)**

- We apply these rules in sequence to **assumptions** given as expressions, to **derive** logical conclusions (**theorems**) which hold for any values
- We usually omit the multiplication sign "×" where the context is clear, e.g. $3yz=3\times y\times z$.

AIML

Mathematical terminology

- A **variable** is a symbol representing a quantity or collection of quantities; examples: X, Y, Z, a, b, c etc.; usually variables are assumed to have specific values in a particular context
- **Data** refers to measured quantities in the real world
- A **mathematical model** consists of **variables** and **adjustable parameters** to represent a real-world problem; AI and ML is essentially about the application of mathematical models to answering questions about the world through logical reasoning and computation

Mathematical terminology

- There are various kinds of **numbers: natural numbers** (whole numbers such as 3, 16) **integers** (whole numbers which can also be negative e.g. -5), **fractional** (**rational**) numbers (ratios of integers e.g. 3/7 or -15/8), and **real** (**continuous**) numbers which can take on any value on the **real line**
- In AI and ML, data are often either symbols or sequences of symbols (**discrete data**), or they are **points** in the **cartesian coordinate system (continuous data)**, given in terms of their coordinate values e.g. (1,3) or (0.5,-0.8) for their two-dimensional x-y coordinates; this can be extended to multiple dimensions e.g. (0.5,-0.8,3.6) for *x-y-z* (three-dimensional)

Special mathematical notation: special sets

- AI and ML makes extensive use of special mathematical notation; we try to minimize this usage in this module but you must be able to read the textbooks

Symbol	**Meaning**
$\mathbb{R}$, ***R***	Set of all **real (continuous)** numbers
$\mathbb{N}$, ***N***	Set of all **natural** numbers not including zero
$\mathbb{Z}$, ***Z***	Set of all **integer** numbers
$\mathbb{Q}$, ***Q***	Set of all **fractional** (**rational**) numbers
$\mathbb{C}$, ***C***	Set of all **complex** numbers

AIML

Special mathematical notation: logical statements

Symbol	Meaning
$\neg$	logical **"not"** statement
$\wedge$	logical **"and"** statement, e.g. $(x=3) \wedge (y=2)$, "x is 3 and y is 2"
$\vee$	logical **"or"** statement, e.g. $(x=3) \vee (x=-3)$ "x is 3 or −3"
$\in$	**is an element of** e.g. $-3 \in \mathbb{Z}$
$\Rightarrow$	logical **"if … then"** statement, e.g. $P \Rightarrow Q$, "P implies Q"
$\Leftrightarrow$	logical **"if and only if"** statement, "iff" e.g. $P \Leftrightarrow Q$, "P if and only if Q"
$\exists$	**"there exists"** quantifier
$\forall$	**"for all"** quantifier

Special mathematical notation: set operations

Symbol	Meaning
$\varnothing$	the **empty** set, a set with no elements
$\cup$	**union** of two sets e.g. $\{5,6\} \cup \{-3,5\} = \{5,6,-3\}$
$\cap$	**intersection** of two sets e.g. $\{5,6\} \cap \{-3,5\} = \{5\}$
$\setminus$	**subtract** from a set e.g. $\{5,6,-3\} \setminus \{5,-3\} = \{6\}$
$\subset$	**subset** or is **contained** in a set, $\{5,-3\} \subset \mathbb{Z}$ is true, 5 and -3 are integer

AIML

Equational relations

- **Equations relate** two quantities or expressions; e.g. the equation $X = YZ - U$ means, "X is always equal in value to Y times Z minus U"
- Equations can also express **inequalities**, e.g. $X < YZ - U$ means, X is always **less** in value than $YZ - U$
- **Example**: assume $X = 1$ then $X \leq 1 \wedge X \geq 1$ means "X is both less than or equal to 1, and greater than or equal to 1", which is a **true** relation, but
  $X < 1 \wedge X > 1$ is a **false** relation

Standard mathematical relations

Symbol	Meaning
$=$	equal to
$<$	less than
$>$	greater than
$\leq$	less than or equal to
$\geq$	greater than or equal to
$\gg$	much greater than
$\ll$	much less than
$\approx$	approximately equal to
$\neq$	not equal to

AIML

Functions

- **Functions** map inputs to outputs; e.g. $f(x)=x+3$ states that if we put 4 into the function named f, we get 7 as the output; can think of this as a opaque "machine" which produces an output value given some input value
- Functions map values from one set into another, for instance the function $\sin(x)$ maps the set $\mathbb{R}$ (real numbers) onto the **subset** of the real numbers, $[-1,1]$
- We write the set of inputs and outputs of the function using the notation $f{:}\mathrm{A}\rightarrow\mathrm{B}$, which means "$f$ has input set A, and output set B"
- **Examples**: the trigonometric function $\sin{:}\mathbb{R}\rightarrow[-1,1]$, the function $g(x)=2x$ has type $g{:}\mathbb{R}\rightarrow\mathbb{R}$
- Functions can only map one single input value into a single value at the output; but they can map many input values onto the same single output value

AIML

Function composition, conditionals

- Functions can be **composed** by putting the output of one function into the input of another, so $g(f(x))$ means first put x into f, then put the result into g
- **Example**: $g(x)=10x$ and $f(x)=x^2$, then $f(g(x))=(10x)^2=100x^2$
- When we do not want to refer to the actual values, we write $g \circ f$
- **Conditional** functions are used to represent more complex maps:

$$f\,(y) = \begin{cases} y^2 & y \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

means that if y is not zero then $f(y) = y^2$ otherwise f returns 0.

AIML

Indexing and referencing

- For the convenience of organizing and computing with large amounts of data or variables, we use numerical **subscripts** (or sometimes **superscripts**, where the context is clear), for instance if we write:

  $X = \{X_1, X_2, X_3, \ldots, X_{N-1}, X_N\}$

  this tells us that the symbol (variable) X refers to an **indexed collection** (a set) of N values, which are **referenced** or **indexed** by an integer.

- We might have $X_{51} = 3$, $X_{52} = -8$ then we know that $X_i = 3$, $X_{i+1} = -8$ for the **index variable** $i = 51$.

AIML

Summation and products

- Given an indexed variable X, we can compute **sums** using **summation** notation, which is shorthand for adding up over a **range** of the index:

$$\sum_{n=3}^{15} X_n = X_3 + X_4 + \cdots + X_{14} + X_{15}$$

in this case, over the range 3,4,...,15. So, $\sum_{n=1}^{N} X_n$ would refer to adding up all values (terms) in a length-N indexed collection of values.

- We denote (repeated/iterated) **products** in a similar way:

$$\prod_{n=3}^{15} X_n = X_3 \times X_4 \times \cdots \times X_{14} \times X_{15}$$

References and further reading

- **Gill**, Chapter 1, Chapter 3