ANALYTIXLABS

Introduction to Mathematical Foundation for Data Science

Module 1
Introduction to Sets and Functions

Course Instructor: Varun Shekhar

Content

- Lesson 1
  - Introduction to sets
  - Subsets
  - Set Builder Notation and Cardinality

- Lesson 2
  - Set Operations

- Lesson 3
  - Lists
  - Functions
  - Composition

Remember: The course is intended to make you get comfortable with concept of Mathematics required for data science. They are **NOT** in itself sufficient to make you **experts** in these concepts of mathematics

Introduction to Sets

- Sets and functions are foundational to the study of mathematics and is everywhere in quantitative disciplines, including statistics and data science.
- A **set** is an unordered collection of well defined objects. The objects in a set are called *elements*.
- We all use sets in our daily life as well. Some examples are below
  - Organization of utensils in the kitchen
  - Shopping malls as there will be separate and well defined portions for different kind of things
  - Playlists in your mobile are organized according to genres
  - The whole of universe can be considered as a set while galaxies can be considered as elements of it
  - Distribution of grocery items in the super market. Fruits, vegetables can be considered as different sets
- If a set S contains a finite number of elements $s1, s2, . . . , s_n$, we can write :

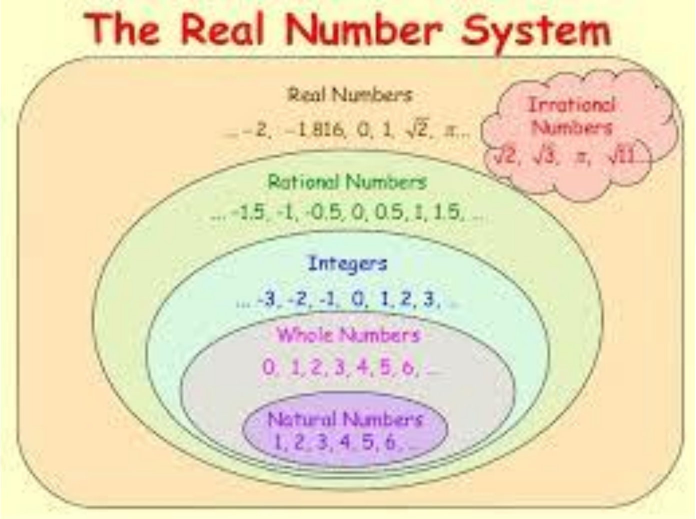
$$S = \{s1, s2, ..., s_n\}$$

Introduction to Sets

- If an element s belongs to set S, we write $s \in S$
- If an element s does not belong to set S, we write $s \notin S$
- If two sets have the same elements, then they are considered equal. For example, {1, 1, 2} = {1, 2}. For this reason, we typically list the elements of a set without duplication.
- The set containing no elements is called the **empty set** and is denoted $\varnothing$ or {}.
- Some sets with standard and specially typeset names include
  - R, the set of real numbers ; R = {-2, -1.816, 0, $\sqrt{2}$..}
  - Q, the set of rational numbers ; Q = {-1.5, 0, 1.1..}
  - Z, the set of integers; Z = {-3, -2, 0, 1, 3..}
  - N, the set of natural numbers; N = {1, 2, 3, 4..}

Subsets

- Suppose S and T are sets. If every element of T is also an element of S, then we say T is a subset of S, denoted $T \subset S$.
- Sets T and S are equal if $T \subset S$ and $S \subset T$
- Example:
  - We have $N \subset Z \subset Q \subset R$, since every natural number (N) is an integer (Z), every integer is rational (Q), and every rational number is real (R).



Set builder notation

- If S is a set and P is a property which each element of S either satisfies or does not satisfy, then

$$\{s \in S : \text{s satisfies P}\}$$

denotes the set of all elements in S which have the property P. This is called *set builder notation*. The colon (:) is read as ‘such that.’

- Example: Suppose S denotes a set of all real numbers between 0 and 1. This can be written in a set builder notation as

$$\{s \in R : 0 < s < 1\}$$

Cardinality

- Given a set S, the cardinality of S, denoted |S|, denotes the number of elements in S.
- Let S = {1, 2, 3, 4}. Then Cardinality of S = 4 since number of elements in S is 4

Countably finite/Infinite sets

- A set is countably infinite if its elements can be counted in such a way that, even though the counting will take forever, you will get to any particular element in a finite amount of time.
- The set of integers is countably infinite, since they can be arranged sequentially: {0,1,-1,2,-2,3,-3,…} However, as suggested by the above arrangement, we can count off all the integers. Counting off every integer will take forever. But, if you specify any integer, say −10 ,234, 872, we will get to this integer in the counting process in a finite amount of time.
- A countably finite set is one which has finite elements in sequence; S = {1, 2, 3, 4}

Set operations

- Following are the set operations that we will cover:
  - Complement
  - Union
  - Intersection
  - Disjoint
  - Partition
  - Cartesian Product

Complement

- If A and S are sets and $A \subset S$, then the complement of A with respect to S, denoted $S \setminus A$ or A^c, is the set of all elements in S that are not in A. That is

$$A^c = \{s \in S : s \notin A\}$$

- Suppose $S = \{1, 2, 3, 4, 5\}$ and $A = \{4, 2\}$, then $A^c = \{1, 3, 5\}$

- Let's Suppose $A \subset S$ and $|S| = 20$ and $|A| = 13$, then $|S \setminus A| = 20 - 13 = 7$

Union

- Union is the operation of combining elements of different sets into one without duplicates
- The **union** of two sets S and T, denoted $S \cup T$, is the set containing all the elements of S and all the elements of T and no other elements. In other words, $s \in S \cup T$ if and only if either $s \in S$ or $s \in T$
- If S = {1, 2, 3, 4} and T = {2, 3, 6, 7, 8}, then $S \cup T$ = {1, 2, 3, 4, 6, 7, 8}

Intersection

- Intersection is the operation of finding elements which are present in all the sets
- The **Intersection** of two sets S and T, denoted $S \cap T$, is the set consisting of elements that are in both *S* and *T*. In other words, $s \in S \cap T$ if and only if either $s \in S$ or $s \in T$
- If S = {1, 2, 3, 4} and T = {2, 3, 6, 7, 8}, then $S \cap T$ = {2, 3}
- The Union and Intersection operations can be applied to any number of sets

Disjoint

- If the participating sets do not have any elements in common, they are called disjoint sets

- In other words, if S and T are two sets such that $S \cap T = \emptyset$, then S and T are disjoint sets

- $S = \{1, 2, 3\}$ and $T = \{5, 6, 7\}$ then $S \cap T = \{\}$, hence they are disjoint

- Problem: Find three sets A, B, and C which have $A \cap B \cap C = \{\}$, but for which all of the intersections $A \cap B$, $B \cap C$, and $A \cap C$ are nonempty.
  - **Solution :** $A = \{1, 2, 3\}$; $B = \{1, 4, 5\}$; $C = \{4, 2, 8\}$, Then
    - $A \cap B = \{1\}$
    - $B \cap C = \{4\}$
    - $A \cap C = \{2\}$
    - $A \cap B \cap C = \{\}$

Partition

- Suppose you and your partner wants to do a shopping of *set* of 5 items to your house. A good idea to do a faster shopping is to *partition* the items to be purchased such that each of your purchase sets do not have anything in common
- A **partition** of a set S is a collection of non-empty sets $S_1, S_2, \ldots, S_n$ such that

$$S = \bigcup_{i=1}^{n} S_i$$

And $S_1, S_2, \ldots, S_n$ are *disjoint*

- Find partitions of S = {1, 2, 3, 4, 5}
  - $S_1 = \{1\}$
  - $S_2 = \{2\}$
  - $S_3 = \{3\}$
  - $S_4 = \{4\}$
  - $S_5 = \{5\}$

Can there be a sixth partition??

Cartesian Product

- The *Cartesian product* of two sets A and B, denoted $A \times B$, is the set of all possible ordered pairs where the elements of A are first and the elements of B are second. In set-builder notation, $A \times B = \{(a, b) : a \in A \text{ and } b \in B\}$.
- ***Ordered Pair:*** An *ordered pair* is a pair of objects where one element is designated first and the other element is designated second, denoted (a, b).
- *Example:* Let $A = \{H, T\}$ and $B = \{1, 2, 3, 4, 5, 6\}$.
  - $A \times B = \{(H, 1), (H, 2), (H, 3), (H, 4), (H, 5), (H, 6), (T, 1), (T, 2), (T, 3), (T, 4), (T, 5), (T, 6)\}$
  - $B \times A = \{(1, H), (2, H), (3, H), (4, H), (5, H), (6, H), (1, T), (2, T), (3, T), (4, T), (5, T), (6, T)\}$
- *Solve:* A couple is planning their wedding. They have four nieces (Ankita, Priyanka, Deeksha, and Sangeetha) and three nephews (Rakshit, Abhinav, and Varun). How many different pairings are possible to have one boy and one girl as a ring bearer and flower girl?

Lists

- A **list** is an ordered collection of *finitely many* elements.

- A **list** is different from a set in that:

  - A list gives importance to the order in which elements appear but order does not matter in set. For example, list {1, 2, 3} ≠ list {2, 1, 3} but Set {1, 2, 3} is same as set {2, 1, 3}

  - A list takes into account redundancy of elements whereas a set does not. For example list {1, 1, 2} has three elements whereas set {1, 1, 2} has two elements

Functions

- If A and B are sets, then a **function** *f* **:** A→*B*, is an assignment to each element of A of some element of B.
- The set A is called the **domain** of *f* and B is called the **codomain** of *f* .
- The domain and codomain of a function should be considered part of the data of the function: to fully specify *f* , we must specify (i) the domain A, (ii) the codomain B, and (iii) the value of f (x) for each x ∈ A. Two
- functions f and g are considered equal if (i) they have the same domain and codomain and (ii) f (x) = g(x) for all x in the domain of f and g. For Example f(x) = |x| and g(x) = $\sqrt{x^2}$ are equal functions as they yield same values for all real values of x

Item	Count
Books	5
Binders	10
Envelopes	4

As seen in the table, each item in Item Set (Domain) has an associated value in the count set (Codomain). For Example, Book is associated with number 5

Range or Image of a function

- Let F(x) = 3x + 2. What goes inside the function (values of x) is the domain. If Domain x = {1, 2, 3}, then we get corresponding function outputs as {5, 8, 11) . This specific output that the function give for a given domain value is called "Range or Image" of a function
- How is Range different from Codomain?
  - The output of f(x) **can possibly** be {1, 2, 3, 4, 5, 6, 7, 8, 9, 10} . This possible range of values is called Codomain. But what explicitly comes out of the function, i.e, what we actually see as output solely depends what we give as an input. That subset of outputs is called Range or Image
  - So Range is a subset of Codomain
- Domain is an essential part of function since for the same function, a **different set of inputs** gives different set of outputs
- Remember: A general function cannot have **one to many** output. For example f(x) = 3 or 4. But **Many to One** is fine

Types of Functions

- Functions can of three types:
  - Injective
  - Surjective
  - Bijective

  (Note: *All explanations take **A** as input set **B** as output set)*

- **Injective function**:
  - Injective means we **cannot** have two or more A's that give the same B. For example $f(x) = x^2$. here f(-2) and f(2) both gives 4. Hence they **cannot** be injective.
  - $f(x) = x + 2$; $f(1) = 3$ and $f(-1) = 1$; so at any given point $f(x_1) \neq f(x_2)$. Such a function is called injective function
  - So a function f is **injective** if no two elements in the domain map to the same element in the codomain.
  - Injective function is a **One to One** function

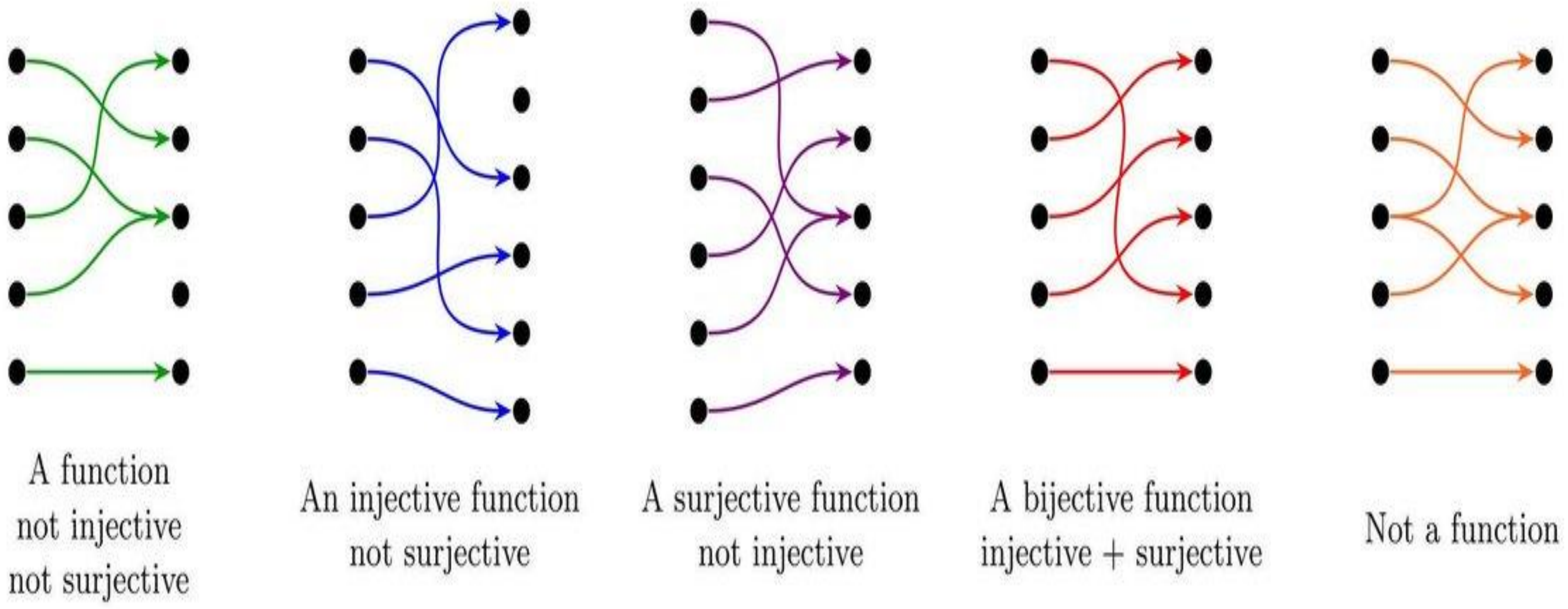
Types of Functions

- **Surjective function**:
  - Surjective means we can have one or more A's that give the same B. Simply put, Every B has some or the other A
  - Let's say A = {1, 2, 3, 5} and B = {6, 7} and *f* is such that f = {(1,6), (2,7), (3, 7), (4, 6)} then this is a surjective function. Because every element in B has one or the other mapping from A. No element in B is **left out**
  - A function f is **surjective** if the range of *f* is equal to the codomain of *f* ; in other words, if $b \in B$ implies that there exists $a \in A$ with $f(a) = b$
  - Surjective functions are also called **Onto** functions

Types of Functions

- **Bijective function**:
  - If a function is both injective and Surjective then it is a bijective function.
  - Let's take the same $f(x) = x^2$ example. But here we will say $f(x) = x^2$ is a function from the set of positive real numbers to positive real numbers
  - So in this example, our domain is a set of positive real number and range is also a set of positive real numbers
  - Now this is both injective and surjective; Injective because $f(2) = 4$ and $f(y) = 4$ implies $x = y = 2$. Surjective because no positive real numbers in codomain is left out, for every positive real number in B, there is one or the other positive real number in A

Types of Functions



Composition

- Function composition means applying one function to the result of another function
- If f : A → *B* and g : B → C, then the function $g \circ f$ which maps x ∈ A to g(f (x)) ∈ C is called the **composition** of g and f .
- Let's take an example f(x) = x + 2 ; g(x) = 2x + 3
  - G(f(x)) = 2(x+2) + 3 = 2x + 7 . i.e, you replace x of g(x) with f(x); in other words f(x) is an input to g(x)
- A function can compose itself. F(x) = 5x + 2
  - Then $f \circ f$ = f(f(x) = 5(5x+2) + 2 = 25 x + 10 + 2 = 25x +12

Keeping a Set of alarms at 4:30 AM, 5 AM, 5:30 AM

But finally Use a Sleep function till 12 PM !!

Thank You!!
Varun Shekhar
varunshekharrc92@gmail.com