Some Mathematical Intuitions and Definitions in Calculus, Probability and Linear algebra

Pattern recognition

September 2024

1 Calculus

1.1 The concept of Limits ((ϵ, δ)-definition)

Suppose f:$\mathbb{R} \to \mathbb{R}$ is a function defined on the real line, and there are two real numbers p and L. One would say that the limit of f, as x approaches p, is L and written

$$\lim_{x \to p} f(x) = L$$

or alternatively, say $f(x)$ tends to L as x tends to p, and written:

$$f(x) \to L \text{ as } x \to p,$$

if for every number $\epsilon > 0$ there is a real $\delta > 0$ such that

$$\text{if } 0 < |x - a| < \delta \text{ then } |f(x) - L| < \epsilon$$

For example, we may say

$$\lim_{x \to 2} (4x + 1) = 9$$

because for every real $\epsilon > 0$ we can take $\delta = \epsilon/4$ such that for all real x, if $0 < |x - 2| < \delta$, then $|4x + 1 - 9| < \epsilon$.

1.2 The concept of Continuity (limits of functions)

A function f: $\mathbb{R} \to \mathbb{R}$ is continuous at a point c of its domain if the limit of $f(x)$, as x approaches c through the domain of f, exists and is equal to $f(c)$. In mathematical notation, this is written as

$$\lim_{x \to c} f(x) = f(c)$$

1.3 The concept of Continuity ((ϵ, δ)-definition)

Given a function f: $D \to \mathbb{R}$ defined on a subset D of the set $\mathbb{R}$ of real numbers. f is said to be continuous at a point x_0 when the following holds:

For any positive real number $\epsilon > 0$, however small, there exists some positive real number $\delta > 0$ such that for all x in the domain of f with $x_0 - \delta < x < x_0 + \delta$, the value of $f(x)$ satisfies

$$f(x_0) - \epsilon < f(x) < f(x_0) + \epsilon.$$

1.4 The concept of Convergent *sequence*

By a *sequence*, we mean a function f defined on the set J of all positive integers. If $f(n) = x_n$, for $n \in J$, it is customary to denote the sequence f by the symbol x_n, or sometimes by x_1, x_2, x_3, ... The values of f, that is, the elements x_n, are called the *terms* of the sequence. If A is a set and if $x_n \in A$ for all $n \in J$, the x_n is said to be a *sequence* in A, or a *sequence of elements of* A.

Note that the terms $x_1, x_2, x_3, ...$ of a sequence need not be distinct.

Next, a sequence a_n has the **limit** L and we write

$$\lim_{n\to\infty} a_n = L \ \textit{or}\ a_n \to L \text{ as } n \to \infty$$

if we can make terms a_n as close to L as we like by taking n sufficiently large. If $\lim_{n\to\infty} a_n$ exists, we say the sequence **converges**(or is **convergent**). Otherwise, we say the sequence **diverges** (or is **divergent**).

1.5 The concept of Convergent *series*

Given a sequence a_n, we use the notation

$$\sum_{n=p}^{q} a_n \ (p \leq q)$$

to denote the sum $a_p + a_{p+1} + \cdots + a_q$. With a_n we associate a sequence s_n, where

$$s_n = \sum_{k=1}^{n} a_k.$$

For s_n we also use the symbolic expression

$$a_1 + a_2 + a_3 + \ldots$$

or, more concisely,

$$\sum_{n=1}^{\infty} a_n$$

which we call an *infinite series*, or just a *series*. The numbers s_n are called the *partial sums* of the series.

If s_n converges to s, we say that the series *converges*, and write

$$\sum_{n=1}^{\infty} a_n = s.$$

The number s is called the sum of the series; but it should be clearly understood that *s is the limit of a sequence of sums*, and is not obtained simply by addition. If s_n diverges, the series is said to diverge.

2 Linear algebra

2.1 The definition of a field

In mathematics, a **field** is a set on which *addition*, *subtraction*, *multiplication*, and *division* are defined and behave as the corresponding operations on rational and real numbers. The best known fields are *the field of rational numbers*, *the field of real numbers* and *the field of complex numbers*. Fields serve as foundational notions in several mathematical domains. This includes different branches of mathematical analysis, which are based on fields with additional structure.

Formally, a field is a set F together with two binary operations on F called *addition* and *multiplication*. A binary operation on F is a mapping $F \times F \to F$, that is, a *correspondence* that associates with each ordered pair of elements of F a uniquely determined element of F.

The result of the addition of a and b is called the *sum* of a and b, and is denoted $a + b$. Similarly, the result of the *multiplication* of a and b is called the *product* of a and b, and is denoted ab or $a \cdot b$. These operations are required to satisfy the following properties, referred to as ***field axioms*** (in these axioms, a, b, and c are arbitrary elements of the field F):

- **Associativity of addition and multiplication**: $a+(b+c) = (a+b)+c$, and $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.
- **Commutativity of addition and multiplication**: $a + b = b + a$, and $a \cdot b = b \cdot a$.
- **Additive and multiplicative identity**: there exist two distinct elements 0 and 1 in F such that $a + 0 = a$ and $a \cdot 1 = a$.
- **Additive inverses**: for every a in F, there exists an element in F, denoted -a, called the additive inverse of a, such that $a + (-a) = 0$.
- **Multiplicative inverses**: for every $a \neq 0$ in F, there exists an element in F, denoted by a^{-1} or $1/a$, called the multiplicative inverse of a, such that $a^{-1} = 1$.
- **Distributivity of multiplication over addition**: $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$.

An equivalent, and more succinct, definition is: a field has two commutative operations, called addition and multiplication; it is a group under addition with 0 as the additive identity; the nonzero elements are a group under multiplication with 1 as the multiplicative identity; and multiplication distributes over addition.

Even more succinctly: a field is a commutative ring where $0 \neq 1$ and all nonzero elements are invertible under multiplication.

2.2 The definition of a Vector Space

A vector space (or *linear space*) V over a field F consists of a set on which two operations (called *addition* and *scalar multiplication*, respectively) are defined so that for each pair of elements x, y in V there is a unique element $x + y$ in V, and for each element a in F and each element x in V there is a unique element ax in V, such that the following conditions hold.

- For all x, y in V, $x + y = y + x$ (commutativity of addition).
- For all x, y, z in V,$(x + y) + z = x + (y + z)$ (associativity of addition).
- There exists an element in V denoted by 0 such that $x + 0 = x$ for each x in V.
- For each element x in V there exists an element y in V such that $x+y = 0$.
- For each element x in V, $1x = x$.
- For each pair of elements a, b in F and each element x in V, $(ab)x = a(bx)$.
- For each element a in F and each pair of elements x, y in V, $a(x + y) = ax + ay$.
- For each pair of elements a, b in F and each element x in V, $(a + b)x = ax + bx$.

The elements $x + y$ and ax are called the sum of x and y and the product of a and x, respectively. The elements of the field F are called **scalars** and the elements of the vector space V are called **vectors**.

2.3 The definition of a Linear transformation

Let V and W be vector spaces (over F). We call a function $T : V \to W$ a linear transformation from V to W if, for all x, $y \in V$ and $c \in F$, we have

$$\text{(a) } T(x + y) = T(x) + T(y)$$

$$\text{(b) } T(cx) = cT(x)$$

2.4 The concept of the relationship of Linear transformation and matrix

Definition. A subset S of a vector space V is called *linearly dependent* if there exist a finite number of distinct vectors $u_1, u_2, \ldots, u_n$ in S and scalars $a_1, a_2, \ldots, a_n$, not all zero, such that $a_1u_1 + a_2u_2 + \cdots + a_nu_n = 0$. In this case we also say that the vectors of S are *linearly dependent.*

Definition. A subset S of a vector space that is not linearly dependent is called *linearly independent.* As before, we also say that the vectors of S are *linearly independent.*

Definition. A basis β for a vector space V is a *linearly independent* subset of V that generates V. If β is a basis for V, we also say that the vectors of β form a basis for V.

Definition. Let V be a finite-dimensional vector space. An ordered basis for V is a basis for V endowed with a specific order; that is, an ordered basis for V is a finite sequence of *linearly independent* vectors in V that generates V.

Example. In F^3, $\beta = \{e_1, e_2, e_3\}$ can be considered an ordered basis. Also $\gamma = \{e_2, e_1, e_3\}$ is an ordered basis, but $\beta \neq \gamma$ as ordered bases.

Now that we have the concept of ordered basis, we can identify abstract vectors in an n-dimensional vector space with n-tuples. This identification is provided through the use of coordinate vectors, as introduced next.

Definition. Let $\beta = \{u_1, u_2, \ldots, u_n\}$ be an ordered basis for a finite dimensional vector space V. For $x \in V$, let $a_1, a_2, ..., a_n$ be the unique scalars such that

$$x = \sum_{i=1}^{n} a_i u_i$$

We define the coordinate vector of x relative to β, denoted $[x]_\beta$, by

$$[x]_\beta = \begin{pmatrix} a_1 \\ a_2 \\ \cdots \\ a_n \end{pmatrix}$$

Notice that $[u_i]_\beta = e_i$ in the preceding definition.

Example Let $V = P_2(R)$, and let $\beta = 1, x, x^2$ be the standard ordered basis for V. If $f(x) = 4 + 6x - 7x^2$, then

$$[f]_\beta = \begin{pmatrix} 4 \\ 6 \\ 7 \end{pmatrix}$$

.

Let us now proceed with the promised matrix representation of a linear transformation. Suppose that V and W are finite-dimensional vector spaces with ordered bases $\beta = \{v_1, v_2, \ldots, v_n\}$ and $\gamma = \{w_1, w_2, \ldots, w_m\}$, respectively.

Let $T : V \to W$ be linear. Then for each j, $1 \leq j \leq n$, there exist unique scalars $a_{ij} \in F$, $1 \leq i \leq m$, such that

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i, \text{ for } 1 \leq j \leq n.$$

.

Definition. Using the notation above, we call the $m \times n$ matrix A defined by $A_{ij} = a_{ij}$ the **matrix representation of T in the ordered bases β and** γ and write $A = [T]_\beta^\gamma$. If $V = W$ and $\beta = \gamma$, then we write$A = [T]_\beta$.

3 Probability

3.1 σ-algebra

Definition. Let X be some set, and let $P(X)$ represent its power set. Then a subset $\Sigma \subseteq P(X)$ is called a σ-algebra if and only if it satisfies the following three properties:

1. X is in Σ.
2. Σ is closed *under complementation*: If some set A is in Σ, then its complement $X/A \in \Sigma$.
3. Σ is closed *under countable unions:*: If some set $A_1, A_2, A_3, \dots$ are in Σ, then $\forall n \in \mathbb{N}$, $A = \cup_{n=1}^{\infty} A_n$.

Definition. Let $J = \{\sigma\text{-algebras on } \mathbb{R} \text{ that contain all open sets of } \mathbb{R}\}$. Let

$$\mathcal{B}(\mathbb{R}) := \cap_{\Sigma \in J} \Sigma$$

Then $\mathcal{B}(\mathbb{R})$ is called the *Borel σ-algebra* of $\mathbb{R}$. An element of $\mathcal{B}(\mathbb{R})$ is called a *Borel set* of $\mathbb{R}$.

In mathematics, a *Borel set* is any set in a topological space that can be formed from **open sets** (or, equivalently, from closed sets) through the operations of *countable union*, *countable intersection*, and *relative complement.*

For a topological space X, the collection of all Borel sets on X forms a σ-algebra, known as the Borel algebra or Borel σ-algebra. The Borel algebra on X is the smallest σ-algebra containing all open sets (or, equivalently, all closed sets; note the second property of the definition).

Borel sets are important in measure theory, since any measure defined on the open sets of a space, or on the closed sets of a space, must also be defined on all Borel sets of that space. Any measure defined on the Borel sets is called a *Borel measure.*

Example. Examples of *Borel sets*:

1. $(a, b) \in \mathcal{B}(\mathbb{R})$;
2. $[a, b] \in \mathcal{B}(\mathbb{R})$;
3. $[a, b) \in \cap_{i=1}^{\infty}(a - \frac{1}{n}, b) \in \mathcal{B}(\mathbb{R})$;

3.2 The definition of a measurable function

Definition. Let X be a set, and let Σ is a σ-algebra over X. A set function μ from Σ to the extended real number line is called a *measure* if the following conditions hold:

- **Non-negativity**: For all $E \in \Sigma$, $\mu(E) \geq 0$.
- $\mu(\varnothing) = 0$.
- **Countable additivity** (or σ-additivity): For all countable collections $\{E_k\}_{k=1}^{\infty}$ of *pairwise disjoint sets* (i.e., $\forall k \in (N)$, $E_k \in \Sigma$ and $E_i \cap E_j = \emptyset$ for $i \neq j$) in Σ,
$$\mu(\cup_{k=1}^{\infty} E_k) = \sum_{k=1}^{\infty} \mu(E_k)$$

If at least one set E has finite measure, then the requirement $\mu(\varnothing) = 0$is met automatically due to countable additivity:

$$\mu(E_k) = \mu(E \cup \varnothing) = \mu(E) + \mu(\varnothing),$$

and therefore $\mu(\varnothing) = 0$.

The pair (X, Σ) is called a *measurable space*, and the members of Σ are called *measurable sets.*

The triple (X, Σ, μ) is called a *measure space.* A **probability measure** is a measure with total measure one, that is, $\mu(X) = 1$. A **probability space** is a measure space with a probability measure, denoted as (X, Σ, μ).

Definition. Given a measure space (X, Σ, μ). A function $f : X \to \mathbb{R}$ is said to be *measurable* if for all $B \in \mathcal{B}(\mathbb{R})$

$$f^{-1}(B) \in \Sigma$$

3.3 Some concepts in probability described in mathematics

Proposition. If $f : \mathbb{R} \to \mathbb{R}$ is a *nonnegative measurable function which satisfies* $\int_{\mathbb{R}} f(x)\,dx = 1$, then

$$\mathbb{P}(A) := \int_A f(x)\,dx$$

for $A \in \mathbb{R}$ is a *probability measure* on $\mathbb{R}$.

Remark. Proof by *Lebesgue monotone convergence theorem*, with the help of the Convergence phenomenon from the Countable additivity property of the measurable function f.

Definition. Given a *probability space* (X, Σ, μ) and a *random variable* X: $\Sigma \to \mathbb{R}$. For $B \in \mathcal{B}(R)$, we write

$$P(X \in B) := P(X^{-1}(B))$$

Example. On $\mathbb{R}$, for $A \in \mathbb{R}$, define

$$P(A) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{x^2}{2}}, dx$$

By theorem, P is a probability measure on $\mathbb{R}$. The function $X \to \mathbb{R} \to \mathbb{R}$ defined by

$$X(x) := x$$

is called the standard normal random variable. Note that

$$P(X \leq a) = P((-\infty, a]) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{x^2}{2}}, dx$$

Definition. Given a *probability space* (X, Σ, μ) and a *random variable space* X on Σ. The probability distribution function For $B \in \mathcal{B}(R)$, we write

$$F_X(x) := P(X \leq x) = P(X^{-1}((-\infty, x]))$$

for $x \in \mathbb{R}$.