Online Self-Assessment for M.Sc. Mathematics Program

Youming Zhao

The math questions in this document are from `https://www.math.uni-potsdam.de/fileadmin/user_upload/images/Dateien/Self-assessment.pdf`. I have provided my solutions and proofs. I have already completed the first three sections and am currently finishing the last three sections. The latest version of this document is available at here.

1 Basic mathematical concepts and rigorous proofs

References: [5]

1. Prove that, for all integers $n \geq 1$, the term $n^3 - n$ is divisible by 3.

   Proof. Due to the factorization $n^3 - n = n(n-1)(n+1)$, we can see that $n^3 - n$ is the product of three consecutive non-negative integers. Since one of three consecutive integers must be divisible by 3, then the term $n^3 - n$ is divisible by 3. □

2. Let L_n be defined by $L_1 = 1$, $L_2 = 3$, and $L_n = L_{n-1} + L_{n-2}$ for $n \geq 3$. Use induction to show that $L_n \geq \frac{1}{2}\left(\frac{3}{2}\right)^n$ for all $n \geq 1$.

   Proof. Since $L_1 = 1 \geq \frac{3}{4} = \frac{1}{2} \times \frac{3}{2}$ and $L_2 = 3 \geq \frac{9}{8} = \frac{1}{2} \times (\frac{3}{2})^2$ hold, then the claim holds for $n = 1$ and $n = 2$. Suppose that $L_n \geq \frac{1}{2}\left(\frac{3}{2}\right)^n$ holds for all $n \leq k$, $k \geq 2$. Now we show that it holds for $n = k+1$.

$$\begin{aligned}
L_{k+1} &= L_k + L_{k-1} \\
&\geq \frac{1}{2}\left(\frac{3}{2}\right)^k + \frac{1}{2}\left(\frac{3}{2}\right)^{k-1} \\
&= \frac{1}{2}\left(\frac{3}{2}\right)^{k+1}\left(\frac{2}{3} + (\frac{2}{3})^2\right) \\
&= \frac{1}{2}\left(\frac{3}{2}\right)^{k+1}\frac{10}{9} \\
&> \frac{1}{2}\left(\frac{3}{2}\right)^{k+1}
\end{aligned}$$

   where the second line follows from the induction hypothesis. Thus, the claim also holds for $n = k+1$. This completes the proof. □

3. Prove that the set of prime numbers is not a finite set.

Proof. We prove the claim by contradiction. Suppose that the set of prime numbers is a finite set, denoted P. Let $|P| = n$ where $|P|$ denotes the cardinality of P. We denote these prime numbers by $p_1, p_2, \ldots, p_n$ which are listed in ascending order. Let $q = p_1 p_2 \cdots p_n + 1$. If q is prime, then we find another prime number, which contradicts the assumption $|P| = n$. If q is composite, then by the Fundamental Theorem of Arithmetic q can be uniquely represented as $p_1^{k_1} p_2^{k_2} \cdots p_n^{k_n}$ with non-negative integers $k_1, k_2, \cdots, k_n$. Since $q > 1$, then at least one of $k_1, k_2, \cdots, k_n$ is a positive, say, $k_i \geq 1$. Then we have $q = p_1 p_2 \cdots p_n + 1 = p_1^{k_1} p_2^{k_2} \cdots p_n^{k_n}$. Since k_i is a positive integer, we get

$$p_i^{k_i}(p_1^{k_1} p_2^{k_2} \cdots p_{i-1}^{k_{i-1}} p_{i+1}^{k_{i+1}} \cdots p_n^{k_n} - p_1 p_2 \cdots p_{i-1} p_{i+1} \cdots p_n) = 1$$

where the result in the parentheses is a nonzero integer. Due to the factorization, k_j's $(i \neq j)$ cannot be 1 simultaneously. This implies $p_i^{k_i} | 1$, which is impossible thanks to $p_i \geq 2$ and $k_i \geq 1$. Hence, there must exist other prime numbers between p_n and q. Both cases contradicts the supposition. Hence, the set of prime numbers is not a finite set. This completes the proof. □

4. Prove that the set of rational numbers $\mathbb{Q}$ is countably infinite. Then prove that the set $\mathbb{Q}^d$ of d-tuples of rationals is countably infinite for any $d \in \mathbb{N}$.

Proof. Let $A_i = \{\frac{1}{i}, \frac{2}{i}, \cdots\}$ where i is a positive integer. Therefore, A_i is countably infinite. According to the result that the union of countably many countable sets is still a countable set, the set of all positive rationals is countably infinite due to $\mathbb{Q}^+ = \bigcup_{i=1}^{\infty} A_i$. By the bijection $\phi(i) = -i$, we can conclude that the set of all the negative rationals $\mathbb{Q}^-$ is also countably infinite. Hence, $\mathbb{Q}$ is also countably infinite because of $\mathbb{Q} = \mathbb{Q}^+ \bigcup \mathbb{Q}^- \bigcup \{0\}$.

Now we prove the second claim by induction. When $d = 1$, $\mathbb{Q}$ is obviously countably infinite as shown above. Assume that $\mathbb{Q}^d$ is countably infinite, then $\mathbb{Q}^{d+1}$ can be represented as $\mathbb{Q}^d \times \mathbb{Q}$. Let $\mathbb{Q} = \{x_1, x_2, \cdots, x_k, \cdots\}$ and $A_k = \mathbb{Q}^d \times \{x_k\}$, then $\mathbb{Q}^{d+1} = \bigcup_{k=1}^{\infty} A_k$. Since $A_k \sim \mathbb{Q}^d$, then A_k is countably infinite by induction hypothesis. Furthermore, $\mathbb{Q}^{d+1}$ is countably infinite. Hence, the set of d-tuples of rationals is countably infinite. This completes the proof. □

2 Analysis

References: [8], [4]

1. Given a sequence $(a_n)_{n\in\mathbb{N}}$ of real numbers, give the definition of its infimum, supremum, limit inferior, limit superior, and limit. Give an example of a sequence that has no limit.

Solution. We define $\inf(a_n)_{n\in\mathbb{N}}$ to be the infimum of the set $\{a_n : n \in \mathbb{N}\}$, and $\sup(a_n)_{n\in\mathbb{N}}$ to be the supremum of the set $\{a_n : n \in \mathbb{N}\}$. We define a new sequence $(a_N^+)_{N\in\mathbb{N}}$ by the formula

$$a_N^- := \inf(a_n)_{n=N}^{\infty}.$$

Then we define the *limit inferior* of the sequence $(a_n)_{n\in\mathbb{N}}$, denoted $\liminf_{n\to\infty} a_n$, by the formula

$$\liminf_{n\to\infty} a_n := \sup(a_N^-)_{N\in\mathbb{N}}.$$

Similarly, we can define

$$a_N^+ := \sup(a_n)_{n=N}^{\infty}.$$

and define the *limit superior* of the sequence $(a_n)_{n\in\mathbb{N}}$, denoted $\limsup_{n\to\infty} a_n$, by the formula

$$\limsup_{n\to\infty} a_n := \inf(a_N^+)_{N\in\mathbb{N}}.$$

The *limit* L of the sequence $(a_n)_{n\in\mathbb{N}}$ is defined to satisfy that for all $\epsilon > 0$, there exists an $N > 0$ such that for all $n > N$, $|a_n - L| < \epsilon$. The sequence $(a_n)_{n\in\mathbb{N}}$ is said to converge to L if and only if $\limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n = L$. The sequence $\{-1, 1, -1, 1, \cdots\}$ has no limit since it has two subsequences which converge to two different reals, i.e. -1 and 1.

2. Given a sequence $(a_n)_{n\in\mathbb{N}}$ of real numbers, give the definition of the quantity $\sum_{i\in\mathbb{N}} a_i$. Describe sufficient conditions on the sequence such that the quantity $\sum_{i\in\mathbb{N}} a_i$ exists and is finite.

   Solution. The quantity $\sum_{i\in\mathbb{N}} a_i$ represents

$$\sum_{i=1}^{\infty} a_i = a_1 + a_2 + a_3 + \cdots$$

   which is called an infinite series, or just a series. We denote the partial sum by $s_n = \sum_{i=1}^{n} a_i$. If $\{s_n\}$ converges to s, i.e. $\lim_{n\to\infty} s_n = s$, we say that the series converges, and write

$$\sum_{i=1}^{\infty} a_i = s.$$

   By the Cauchy criterion, if for every $\epsilon > 0$ there is an integer N such that for all $m \geq n \geq N$

$$\left|\sum_{i=n}^{m} a_i\right| < \epsilon,$$

   then $\sum_{i=1}^{\infty} a_i$ converges. Further, $\sum_{i\in\mathbb{N}} a_i$ exists and is finite.

3. Give the definition for a function $f : \mathbb{R} \to \mathbb{R}$ to have the following properties at a point $x \in \mathbb{R}$: (i) left-continuity; (ii) right-continuity; (iii) continuity. Give the definition of a uniformly continuous function.

   Solution.

   (a) left-continuity: $\lim_{x\to x_{0-}} f(x) = f(x_0)$;
   (b) right-continuity: $\lim_{x\to x_{0+}} f(x) = f(x_0)$;
   (c) continuity: $\lim_{x\to x_{0-}} f(x) = \lim_{x\to x_{0+}} f(x) = f(x_0)$;
   (d) uniform continuity: $\forall\epsilon > 0$, $\exists\delta > 0$, $\forall x, x_0 \in \mathbb{R}$, if $|x - x_0| < \delta$, then $|f(x) - f(x_0)| < \epsilon$.

4. Give the definition of the derivative of a function $f : \mathbb{R} \to \mathbb{R}$ at a point $x \in \mathbb{R}$ and give the geometric interpretation of the derivative. Generalise this to the case where $f : \mathbb{R}^d \to \mathbb{R}$ for $d \in \mathbb{N}, d \geq 2$.

   Solution. The derivative of a function $f : \mathbb{R} \to \mathbb{R}$ at a point $x \in \mathbb{R}$ is defined by the following formula:

$$f'(x) = \lim_{\Delta x\to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

The geometric interpretation of the derivative is the slope of the tangent line to the graph of $f(x)$ at the point $(x, f(x))$. Generalising this to the case where $f : \mathbb{R}^d \to \mathbb{R}$ for $d \in \mathbb{N}, d \geq 2$, we have the derivative of f, $L \in \mathbb{R}^d$, which satisfies

$$\lim_{\|\Delta x\| \to 0} \frac{\|f(x + \Delta x) - (f(x) + \langle L, \Delta x \rangle)\|}{\|\Delta x\|} = 0$$

where $\| \cdot \|$ denotes l_2 norm, e.g. $\|\Delta x\| = \sqrt{(\Delta x_1)^2 + (\Delta x_2)^2 + \cdots + (\Delta x_d)^2}$.

5. Define the Riemann integral of a function $f : \mathbb{R} \to \mathbb{R}$ in terms of a limit.

   Solution. Suppose a function $f(x)$ is bounded on $[a, b]$, for any partition

   $$P : a = x_0 < x_1 < x_2 < \cdots < x_n = b,$$

   and for any $\xi_i \in [x_{i-1}, x_i]$, let $\Delta x_i = x_i - x_{i-1}$ and $\lambda = \max_{1 \leq i \leq n}(\Delta x_i)$, then when $\lambda \to 0$,

   $$\lim_{\lambda \to 0} \sum_{i=1}^{n} f(\xi_i) \Delta_i$$

   exists and the limit is independent with the partition P and the choice of x_i. Then we say that f is Riemann integrable on $[a, b]$ and the limit is its Riemann integral on $[a, b]$.

6. State the following theorems (including their hypotheses): (i) the intermediate value theorem; (ii) the mean value theorem; (iii) the fundamental theorem of calculus.

   Solution.

   (a) the intermediate value theorem: Let $a < b$, and let $f : [a, b] \to \mathbb{R}$ be a continuous function on $[a, b]$. Let y be a real number between $f(a)$ and $f(b)$, i.e., either $f(a) \leq y \leq f(b)$ or $f(a) \geq y \geq f(b)$. Then there exists at least one $c \in [a, b]$ such that $f(c) = y$.

   (b) the mean value theorem: Let $a < b$, and let $f : [a, b] \to \mathbb{R}$ be a function which is continuous $[a, b]$ and differentiable on (a, b). Then there exists an $x \in (a, b)$ such that $f'(x) = \frac{f(b) - f(a)}{b - a}$.

   (c) the fundamental theorem of calculus: Let $a < b$, and let $f : [a, b] \to \mathbb{R}$ be a Riemann integrable function. If $F : [a, b] \to \mathbb{R}$ is an antiderivative of f, then $\int_a^b f(x) \mathrm{d}x = F(b) - F(a)$.

7. Let $d \in \mathbb{N}, d > 1$, and $f : \mathbb{R}^d \to \mathbb{R}$. Let $k \in \{1, \ldots, d\}$ and define the k-th partial derivative $D_k f$ of f. Define the gradient of a scalar-valued, continuously differentiable function on $\mathbb{R}^d$, and the divergence of a continuously differentiable function $f : \mathbb{R}^d \to \mathbb{R}^d$.

   Solution. The k-the partial derivative, denoted $\frac{\partial f(x_1, \cdots, x_d)}{\partial x_k}$ is defined by the limit

   $$\lim_{\Delta \to 0} \frac{f(x_1, \cdots, x_k + \Delta x, \cdots, x_d) - f(x_1, \cdots, x_k, \cdots, x_d)}{\Delta x}.$$

   The gradient of a scalar-valued, continuously differentiable function on $\mathbb{R}^d$ is defined by the following column vector

   $$\nabla f = (\frac{\partial f}{\partial x_1}; \cdots ; \frac{\partial f}{\partial x_d}).$$

The divergence of a continuously differentiable function $f : \mathbb{R}^d \to \mathbb{R}^d$ is defined by

$$\mathrm{div} f = \frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2} + \cdots + \frac{\partial f_d}{\partial x_d}.$$

8. State a theorem about the existence and uniqueness of solutions to ordinary differential equations. Describe the main points of the proof of this theorem.

   Solution. Picard's Existence and Uniqueness Theorem: Let $D \subset \mathbb{R} \times \mathbb{R}^n$ be a closed rectangle with $(t_0, y_0) \in \mathrm{int} D$, the interior of D. Let $f : \mathbb{R} \to \mathbb{R}^n$ be a function that is continuous in t and Lipschitz continuous in y (with a Lipschitz constant independent from t). Then, there exists some constant $\epsilon > 0$ such that the Initial Value Problem (IVP):

   $$y'(t) = f(t, y(t)), \quad y(t_0) = y_0.$$

   has a unique solution $y(t)$ on the interval $[t_0 - \epsilon, t_0 + \epsilon]$.

   The proof relies on transforming the differential equation, and applying the Banach fixed-point theorem. By integrating both sides, any function satisfies the differential equation mush also satisfy the integral equation

   $$y(t) - y(t_0) = \int_{t_0}^{t} f(s, y(s)) \mathrm{d}s.$$

   A simple proof of existence of the solution is obtained by successive approximations. In this context, the method is known as Picard iteration.

9. Consider the following differential equation in $\mathbb{R}^2$:

   $$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix}' = \begin{pmatrix} 2 & -8 \\ -8 & -17 \end{pmatrix} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}.$$

   (a) State whether there exists a solution to the differential equation, and whether this solution is unique.

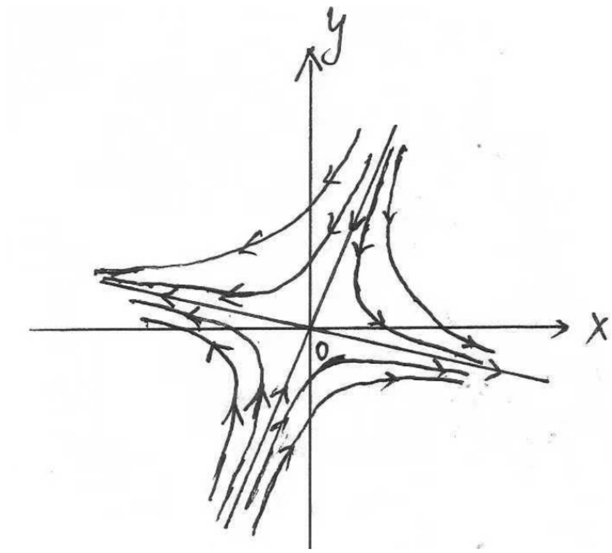
   Solution. Since the coefficient matrix is a constant matrix, then it is continuous for all $t \in \mathbb{R}$, then there exists a solution to the differential equation. The solution is not unique because the initial condition is not given.

   (b) Describe the long-time behaviour of solutions and their dependence on the initial condition $(x(0), y(0))^\top \in \mathbb{R}^2$. Justify your answers.

   Solution. By simple calculations, we get the eigenvalues, i.e. $\lambda_1 = \frac{-15+\sqrt{617}}{2} > 0$ and $\lambda_2 = \frac{-15-\sqrt{617}}{2} < 0$, with corresponding eigenvectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$. Thus, the general solution will be $c_1 \boldsymbol{v}_1 e^{\lambda_1 t} + c_2 \boldsymbol{v}_2 e^{\lambda_2 t}$ where c_1 and c_2 are determined by the initial condition $(x(0), y(0))$. In terms of the long-time behavior of solutions, it depends on the positive eigenvalue which results in the solution growing exponentially as $t \to \infty$. The only exception is that when the initial condition gives $c_1 = 0, c_2 \neq 0$, the solutions decay exponentially. The special case is the constant solution $(0, 0)^\top$ which is a stable state.

   (c) Draw the phase portrait of the differential equation.

   Solution. To draw the phase portrait, we need to get two eigenvectors $\boldsymbol{v_1} = (1, -0.36)^\top$ and $\boldsymbol{v_2} = (1, 2.74)^\top$. Then the phase portrait is shown as follows.



10. Consider the flow associated to the following vector field on $\mathbb{R}^2$:

$$F(x,y) = \begin{pmatrix} x(y-1) \\ -y^2 \end{pmatrix}.$$

Determine whether the flow has any critical points. If there are critical points, identify all of them.

Solution. By letting $x(y-1) = 0$ and $-y^2 = 0$, we get the critial point: (0,0).

11. State the Cauchy integral formula for holomorphic functions. Under what assumptions is this formula valid?

Solution. Suppose f is holomorphic in an open set that contains the closure of a disc D. If C denotes the boundary circle of this disc with the positive orientation, then

$$f(z) = \frac{1}{2\pi i}\int_C \frac{f(\xi)}{\xi - z}\mathrm{d}\xi, \quad \forall z \in D. \tag{1}$$

12. Use one or more theorems to justify why holomorphic functions are infinitely differentiable.

Solution. The result that holomorphic functions are infinitely differentiable follows from the following theorem, which has the common hypotheses as the Cauchy integral formula,

$$f^{(n)}(z) = \frac{n!}{2\pi i}\int_C \frac{f(\xi)}{(\xi - z)^{n+1}}\mathrm{d}\xi, \quad \forall z \in D, n = 1, 2, \cdots. \tag{2}$$

More specifically, for each $z_0 \in D$, we can apply the above theorem to a small enough circle centered at z_0 and contained in D and then the n-th order of derivative of f at z_0 is obtained.

3 Linear Algebra

References:

1. Let V, W be vector spaces and $\Phi : V \to W$ a linear map.

   (a) What does it mean to say that a subset $U \subseteq V$ is a subspace?

   Solution. Given U is a subspace of V, then $\boldsymbol{\alpha}+\boldsymbol{\beta} \in U$ and $k\boldsymbol{\alpha} \in U$ for all $\boldsymbol{\alpha}, \boldsymbol{\beta} \in U$ and any scalar k.

   (b) Define the kernel $\ker(\Phi)$ and show that it is a subspace of V.

   Proof. We define $\ker(\Phi)$ to be $\{\boldsymbol{\alpha} : \Phi(\boldsymbol{\alpha}) = \mathbf{0}\}$. Given $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \ker(\Phi)$, then $\Phi(\boldsymbol{\alpha}) = \mathbf{0}$ and $\Phi(\boldsymbol{\beta}) = \mathbf{0}$. Since Φ is a linear map, then $\Phi(\boldsymbol{\alpha}+\boldsymbol{\beta}) = \Phi(\boldsymbol{\alpha}) + \Phi(\boldsymbol{\beta}) = \mathbf{0}+\mathbf{0} = \mathbf{0}$ and $\Phi(k\boldsymbol{\alpha}) = k\Phi(\boldsymbol{\alpha}) = k \cdot \mathbf{0} = \mathbf{0}$. Therefore, $\boldsymbol{\alpha}+\boldsymbol{\beta} \in \ker(\Phi)$ and $k\boldsymbol{\alpha} \in \ker(\Phi)$. Thus, $\ker(\Phi)$ is a subspace of V. This completes the proof. □

   (c) Prove that $\Phi : V \to W$ is injective if and only if $\ker(\Phi) = \{\mathbf{0}\}$.

   Proof. Since Φ is a linear map, then $\Phi(\mathbf{0}) = \Phi(\boldsymbol{\alpha}-\boldsymbol{\alpha}) = \Phi(\boldsymbol{\alpha}) - \Phi(\boldsymbol{\alpha}) = \mathbf{0}$. Hence, $\mathbf{0} \in \ker\Phi$. With this, we first prove the necessity. Given Φ is injective, suppose that $\mathbf{0} \neq \boldsymbol{\beta} \in \ker\Phi$, then $\Phi(\boldsymbol{\beta}) = \mathbf{0} = \Phi(\mathbf{0})$. Since Φ is injective, then $\boldsymbol{\beta} = \mathbf{0}$, which contradicts the assumption. Thus, $\mathbf{0}$ is the unique element of $\ker\Phi$.

   Now we show the sufficiency. Given $\Phi(\boldsymbol{\alpha}) = \Phi(\boldsymbol{\beta})$, then $\Phi(\boldsymbol{\alpha}) - \Phi(\boldsymbol{\beta}) = \Phi(\boldsymbol{\alpha}-\boldsymbol{\beta}) = \mathbf{0}$. Since $\ker(\Phi) = \{\mathbf{0}\}$, then $\boldsymbol{\alpha}-\boldsymbol{\beta} = \mathbf{0}$. Thus, $\boldsymbol{\alpha} = \boldsymbol{\beta}$, which implies that Φ is injective. This completes the proof. □

2. Let $\Phi : V \to W$ be a linear map between finite dimensional vector spaces.

   (a) Define the rank of Φ and the nullity of Φ.

   Solution. The rank of Φ is the dimension of its image, i.e. $\mathrm{rank}(\Phi) = \dim(\Phi(V)) = \mathrm{rank}(A)$, where A is the matrix that defines Φ. The nullity of Φ is the dimension of its kernel, i.e. $\mathrm{nullity}(\Phi) = \dim(\ker(\Phi))$.

   (b) Find the rank of Φ_A for the linear map $\Phi_A : \mathbb{R}^4 \to \mathbb{R}^3 : x \to Ax$ given by the matrix

   $$A = \begin{pmatrix} 1 & 1 & 2 & 3 \\ 1 & 0 & 2 & 2 \\ 3 & 2 & 6 & 8 \end{pmatrix}.$$

   Solution. The rank of Φ_A is equal to the rank of A, which is the dimension of the row space of A. It is easy to see that the first two rows are linearly independent and the third row is the linear combination of the first two rows. Hence, the dimension of the row space of A is 2. Thus, $\mathrm{rank}(\Phi_A) = 2$.

   (c) How many vectors must there be in a basis for the null space of A? Briefly justify your answer.

   Solution. There are 2 columns in a basis for the null space of A. Recall the Rank-Nullity Theorem,

   $$\mathrm{rank}(\Phi_A) + \mathrm{nullity}(\Phi_A) = \dim(V)$$

   Since $\mathrm{rank}(\Phi_A) = 2$ and $\dim(V) = 4$, we have $\dim(\ker(\Phi)) = \mathrm{nullity}(\Phi) = 4 - 2 = 2$. Therefore, the dimension of the null space of A is 2.

3. Let V be a finite dimensional vector space over a field F.

(a) Define the span $\{v_1, \ldots, v_n\}$ of a list of vectors $v_i \in V$.

Solution. The span $\{v_1, \ldots, v_n\}$ of a list of vectors $v_i \in V$ is all the linear combinations of them.

(b) What does it mean to say that the list $v_1, \ldots, v_n$ is (i) linearly independent, (ii) a basis of V.

Solution. We say that the list $v_1, \ldots, v_n$ is (i) linearly independent if the solution to $c_1v_1 + c_2v_2 + \cdots + c_nv_n = 0$ is only when all its coefficients c_i are 0. When the span $\{v_1, \ldots, v_n\}$ of a list of vectors $v_i \in V$ is exactly V and the list $\{v_1, \ldots, v_n\}$ is linearly independent, then we say that the list $\{v_1, \ldots, v_n\}$ is a basis of V.

4. Let Φ be a linear operator on a finite-dimensional complex inner product space V with adjoint Φ^*, and let U be a subspace of V.

(a) Define what it means to say that U is a Φ-invariant subspace, and define the orthogonal space $U^\perp$.

Solution. We say that U is a Φ-invariant subspace if $\Phi(U) \subseteq U$. The orthogonal space $U^\perp$ is defined by $\{\boldsymbol{v} \in V | (\boldsymbol{v}, U) = 0\}$. In other words, for all $\boldsymbol{u} \in U$, $(\boldsymbol{v}, \boldsymbol{u}) = 0$ holds.

(b) Show that if U is Φ-invariant, then $U^\perp$ is Φ^*-invariant.

Proof. For all $\boldsymbol{\alpha} \in U, \boldsymbol{\beta} \in U^\perp$, we have

$$(\boldsymbol{\alpha}, \Phi^*(\boldsymbol{\beta})) = (\Phi(\boldsymbol{\alpha}), \boldsymbol{\beta}) = 0, \tag{3}$$

which implies that $U^\perp$ is Φ^*-invariant. □

(c) State the Spectral Theorem for normal operators, using orthogonal projections.

Solution. Let V be an n-dimensional inner product space. and Φ be a normal operator on V. Suppose $\lambda_1, \cdots, \lambda_n$ are the distinct eigenvalues of Φ, then $V = \bigoplus_{i=1}^k V_{\lambda_i}$. Let $\boldsymbol{E}_i$ be the orthogonal projections from V to V_{λ_i}, Φ has the following decomposition:

$$\Phi = \lambda_1 \boldsymbol{E}_1 + \lambda_2 \boldsymbol{E}_2 + \cdots + \lambda_n \boldsymbol{E}_n. \tag{4}$$

5. In this question A, B, and C are sets, $f : A \to B$ and $g : B \to C$ are maps, and $h = g \circ f$ is the composed map, so $h : A \to C$.

(a) Suppose that h is surjective. Does it follow that f is surjective? Justify your answer.

Solution. Given h is surjective, f is not necessarily surjective. An example is $f : \mathbb{R} \to \mathbb{R}$ with $f = x^2$ and $g : \mathbb{R} \to \mathbb{R}_+$ with $g = x^2$. In this example, h is surjective but f is not surjective.

(b) Suppose that h is injective. Does it follow that g is injective? Justify your answer.

Solution. Given h is injective, it follows that g is injective. Since h is injective, then $h(a) \neq h(b)$ for all $a \neq b \in A$. Then $f(a) \neq f(b)$, otherwise $g(f(a)) = g(f(b))$ contradicts the assumption $h(a) \neq h(b)$. Thus, we have $g(f(a)) \neq g(f(b))$ given $f(a) \neq f(b)$, which implies that g is injective.

4 Measure theory

References: [2], [3], [1]

1. Define a sigma-algebra $\mathcal{F}$ on a nonempty set Ω. What does it mean for a sigma-algebra to be countably generated? What does it mean for $\mathcal{G}$ to be a sub-sigma algebra of $\mathcal{F}$?

   Solution. Given $\mathcal{F}$ is a collection of subsets of a nonempty set Ω, $\mathcal{F}$ is called a σ-algebra if it satisfies

   (a) $\emptyset \in \mathcal{F}$;

   (b) if $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$;

   (c) if A_n $(n = 1, 2, \cdots)$, then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

   A countably generated sigma-algebra is a σ-algebra $\mathcal{F}$ of subsets of a set Ω such that there exists a countable family $\mathcal{E} = \{E_n : n \in \mathbb{N}\}$ such that $\mathcal{F} = \sigma(\mathcal{E})$, where $\sigma(\mathcal{E})$ is the σ-algebra generated by $\mathcal{E}$. Given $\mathcal{G}$ is a sub-sigma algebra of $\mathcal{F}$, we conclude that $\mathcal{G}$ is also a σ-algebra and $\mathcal{G} \subset \mathcal{F}$.

2. Let $d \in \mathbb{N}$ be arbitrary. Define the Borel sigma-algebra on $\mathbb{R}^d$.

   Solution. The σ-algebra generated by the collection of all open sets in $\mathbb{R}^d$ is called Borel sigma-algebra on $\mathbb{R}^d$.

3. Define what a measurable space is and what a measure on a measurable space is.

   Solution. Consider a set X and a σ-algebra $\mathcal{F}$ on X. Then the tuple $(X, \mathcal{F})$ is called a measurable space. A set function μ is called a measure if the following conditions hold:

   - Non-negativity: For all $E \in \mathcal{F}$, $\mu(E) \geq 0$.
   - $\mu(\emptyset) = 0$.
   - Countable additivity (σ-addivity): For all countable collections $\{E_k\}_1^{\infty}$ of disjoint open sets in $\mathcal{F}$,
   $$\mu\left(\bigcup_{k=1}^{\infty} E_k\right) = \sum_{k=1}^{\infty} \mu(E_k).$$

4. Let $(S_1, \mathcal{S}_1)$ and $(S_2, \mathcal{S}_2)$ be two measurable spaces. What does it mean for a function $f : S_1 \to S_2$ to be $(\mathcal{S}_1, \mathcal{S}_2)$-measurable?

   Solution. That means for all $E \in \mathcal{S}_2$,
   $$f^{-1}(E) := \{x \in S_1 \mid f(x) \in E\} \in \mathcal{S}_1.$$

5. State the invariance properties of d-dimensional Lebesgue measure on $\mathbb{R}^d$.

   Solution.

   - The translation-invariance property of d-dimensional Lebesgue measure on $\mathbb{R}^d$ is that if E is a measurable set and $h \in \mathbb{R}^d$, then the set $E_h = E + h = \{x + h | x \in E\}$ is also measurable and their measures are equal, i.e., $m(E_h) = m(E)$.

- If $T : \mathbb{R}^n \to \mathbb{R}^n$ is linear and $E \in \mathcal{F}$, then $T(E) \in \mathcal{F}$ and $m(T(E)) = |\det(T)| \cdot m(T)$.

6. Define the Lebesgue integral on $\mathbb{R}$. For a subset S of $\mathbb{R}$, let $\mathbb{I}_S$ denote the indicator function of S, and let $\mathbb{Q}$ denote the set of rational numbers. Calculate the value of the Lebesgue integral $\int_{\mathbb{R}} \mathbb{I}_{\mathbb{Q}}(x)\, \mathrm{dLeb}(x)$, where 'Leb' denotes Lebesgue measure.

   Solution.

   - Simple functions: Let Ω be a measurable subset of $\mathbb{R}$, and let $f : \Omega \to \mathbb{R}$ be a measurable function. We say that f is a simple function if the image of $f(\Omega)$ is finite.
   - Lebegue integral of simple functions: Let Ω be a measurable subset of $\mathbb{R}$, and let $f : \Omega \to \mathbb{R}$ be a simple function which is non-negative. We then define the Lebesgue integral $\int_\Omega f$ of f on Ω by

   $$\int_\Omega f := \Sigma_{\lambda \in f(\Omega); \lambda > 0} \lambda m(\{x \in \Omega | f(x) = \lambda\}).$$

   - Majorization: Let $f : \Omega \to \mathbb{R}$ and $g : \Omega \to \mathbb{R}$. We say that f majorizes g, or g minorizes f, if and only if we have $f(x) \geq g(x)$ for all $x \in \Omega$.
   - Lebesgue integral for non-negative functions: Let Ω be a measurable subset of $\mathbb{R}$, and let $f : \Omega \to [0, \infty]$ be a measurable function and non-negative. We then define the Lebesgue integral $\int_\Omega f$ of f on Ω to be

   $$\int_\Omega f := \sup \left\{ \int_\Omega s : s \text{ is simple and non-negative, and minorizes } f \right\}.$$

   - Lebesgue integral: Let Ω be a measurable subset of $\mathbb{R}$, and let $f : \Omega \to [0, \infty]$ be a measurable function. Let

   $$f^+ := \max(f, 0); \quad f^- := \min(f, 0). \tag{5}$$

   Then both f^+ and f^- are non-negative. When $x \in \Omega$, we have

   $$f = f^+ - f^-, \quad |f| = f^+ + f^-. \tag{6}$$

   The integral of f is only defined if at least one of $\int_\Omega f^+ < \infty$ and $\int_\Omega f^- < \infty$ holds, in which case we define

   $$\int_\Omega f = \int_\Omega f^+ - \int_\Omega f^-. \tag{7}$$

   A measurable function is called Lebesgue integrable if $\int_\Omega f^+ < \infty$ and $\int_\Omega f^- < \infty$. Since $|f| = f^+ + f^-$, this is equivalent to $\int_\Omega |f| < \infty$, in which case we say that f is absolutely integrable.

   Based on the above definitions, we have

   $$\int_{\mathbb{R}} \mathbb{I}_{\mathbb{Q}}(x)\, \mathrm{dLeb}(x) = 1 \cdot m(Q) + 0 \cdot m(Q^c) = 1 \cdot 0 + 0 \cdot \infty = 0. \tag{8}$$

   where $0 \cdot \infty = 0$ by convention.

7. State the monotone convergence theorem, the dominated convergence theorem, and Fatou's lemma.

   Solution.

   - Lebesgue monotone convergence theorem: Let Ω be a measurable subset of $\mathbb{R}^n$, and let $(f)_{n=1}^{\infty}$ be a sequence of non-negative measurable functions from Ω to $\mathbb{R}$ which are increasing the sense that
   $$0 \leq f_1(x) \leq f_2(x) \leq f_3(x) \leq \dots \text{ for all } x \in \Omega. \tag{9}$$
   Then we have
   $$0 \leq \int_\Omega f_1 \leq \int_\Omega f_2 \leq \int_\Omega f_3 \leq \dots \tag{10}$$
   and
   $$\int_\Omega \sup_n f_n = \sup_n \int_\Omega f_n. \tag{11}$$
   - Lebesgue dominated convergence theorem: Let Ω be a measurable subset of $\mathbb{R}^n$, and let $(f)_{n=1}^{\infty}$ be a sequence of measurable functions from Ω to $\bar{\mathbb{R}}$ which converges pointwise. Suppose also that there is an absolutely integrable function $F : \Omega \to [0, \infty]$ such that $|f_n(x)| \leq F(x)$ for all $x \in \Omega$ and all $n = 1, 2, 3, \dots$. Then
   $$\int_\Omega \lim_{n\to\infty} f_n = \lim_{n\to\infty} \int_\Omega f_n. \tag{12}$$
   - Fatou's lemma: Let Ω be a measurable subset of $\mathbb{R}^n$, and let $(f)_{n=1}^{\infty}$ be a sequence of non-negative measurable functions from Ω to $[0, \infty]$. Then
   $$\int_\Omega \liminf_{n\to\infty} f_n \leq \liminf_{n\to\infty} \int_\Omega f_n. \tag{13}$$

8. Let μ, ν be two measures on a measurable space $(S, \mathcal{S})$. Define what it means for μ to be absolutely continuous with respect to ν and what it means for μ and ν to be mutually equivalent. Give an example of two measures μ and ν such that μ is absolutely continuous with respect to ν but not mutually equivalent to ν.

   Solution.

   - Absolute continuity: Let $\mathcal{N}_\mu = \{A \in \mathcal{S} | \mu(A) = 0\}$ and $\mathcal{N}_\nu = \{A \in \mathcal{S} | \nu(A) = 0\}$. If $\mathcal{N}_\nu \subseteq \mathcal{N}_\mu$, then we say that μ is absolutely continuous w.r.t. ν (denoted $\mu \ll \nu$). If $\mathcal{N}_\mu \subseteq \mathcal{N}_\nu$, then we say that ν is absolutely continuous w.r.t. μ (denoted $\nu \ll \mu$).
   - Mutual equivalence: if $\nu \ll \mu$ and $\mu \ll \nu$, then μ and ν are mutually equivalent.

   Consider the measurable space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, where $\mathcal{B}(\mathbb{R})$ is the Borel σ-algebra on $\mathbb{R}$. Let ν be the Lebesgue measure on $\mathbb{R}$. Let $\mu(A) = \int_A f(x) d\nu(x)$, where $f(x)$ is the characteristic function of the interval $[0, 1]$, namely
   $$f(x) = \begin{cases} 1, & x \in [0,1] \\ 0, & x \notin [0,1]. \end{cases} \tag{14}$$
   Thus, we have $\mu(A) = \nu(A \cap [0, 1])$. To see that μ is absolutely continuous w.r.t. ν. If $\nu(A) = 0$, then $\mu(A) = \nu(A \cap [0, 1]) = 0$ due to $A \cap [0, 1] \subseteq A$. However, if $B = (2, 3)$, then $\mu(B) = \nu(B \cap [0, 1]) = \nu(\emptyset) = 0$, but $\nu(B) = 1$. Hence, μ and ν are not mutually equivalent.

5 Geometry / Topology

1. Give the definition of a topology on a nonempty set and define all the terms that you use in the definition.

   Solution. Given a set X, and $\mathcal{T}$ is a collection of subsets of X, then $\mathcal{T}$ is called a topology (space) if it satisfies

   (a) $X, \emptyset \in \mathcal{T}$;

   (b) if $A, B \in \mathcal{T}$, then $A \bigcap B \in \mathcal{T}$;

   (c) if $\mathcal{T}_1 \subset \mathcal{T}$, then $\bigcup_{A \in \mathcal{T}_1} \in \mathcal{T}$.

2. What does it mean for a topology to be Hausdorff?

   Solution. A topological space with the property that two distinct points can always be surrounded by disjoint open sets is called a *Hausdorff space*.

3. Define what is a homeomorphism of topological spaces.

   Solution. A function $f : X \to Y$ between two topological spaces is a homeomorphism if it has the following properties:

   - f is a bijection,
   - f is continuous,
   - the inverse function f^{-1} is continuous.

4. Give the definition of a connected set, a path connected set, and a simply connected set. Give an example of a set that is connected but not path connected.

   Solution. Let X be a topological space.

   - Disconnectedness: We say that X is disconnected if there exist open sets $U, V \subset X$ such that $U \cap V = \emptyset$, $U \subset X$ and $V \subset X$, and $U \cup V = X$.
   - Connectedness: X is connected if it is not disconnected.
   - Path-connectedness: X is path-connected if for all points $x, y \in X$ there exists a path from x to y, that is a continuous map $\gamma : [0, 1] \to X$ such that $\gamma(0) = x$ and $\gamma(1) = y$.
   - Simple connectedness: X is simply connected if it is path connected and any simple closed curve can be shrunk to a point continuously in the set.

   Let $S = \left\{ \left(x, \sin \frac{1}{x}\right) | x \in (0, 1] \right\}$ and $T = \{0\} \times [-1, 1]$. Since $\sin \frac{1}{x}$ is a continuous map under $(0, 1]$, then S is connected. It is easy to verify that the closure $\bar{S} = S \cup T$ is connected. Therefore, $\bar{S}$ is connected. However, $\bar{S}$ is not path-connected. To see this, suppose $\bar{S}$ is path-connected, then there exists a path $f : [0, 1] \to \bar{S}$ with $f(0) = (0, 0) \in \bar{S}$ and $f(1) = (1, \sin 1) \in \bar{S}$. Note that when f is restricted on $(0, 1]$, we have $f_{(0,1]}(x) = \sin \frac{1}{x}$. Since a path is a continuous map, then we should have $(0, 0) = \lim_{x \to 0^+} f(x)$. However, $\lim_{x \to 0^+} f(x) = (\lim_{x \to 0^+} x, \lim_{x \to 0^+} \sin \frac{1}{x})$ where $\lim_{x \to 0^+} \sin \frac{1}{x}$ does not exist. This is a contradiction. Hence, $\bar{S}$ is not path-connected.

5. Give the definition of a d-dimensional real manifold. Give an example of a 2-dimensional real manifold.

   Solution.

   - d-dimensional real manifold: A topological space $\mathcal{M}$ is called a d-dimensional real manifold if every point $x \in \mathcal{M}$ has a neighborhood homeomorphic to Euclidean space $\mathbb{R}^d$.
   - Example: A sphere S^2 is a 2-dimensional real manifold, where $S^2 = \{(x, y, z)\mathbb{R}^3 | x^2 + y^2 + z^2 = 1\}$.

6. What is the tangent space to a manifold $\mathcal{M}$ at a point $p \in \mathcal{M}$? What is the tangent space to the example of the 2-manifold you gave for the previous question?

   Solution.

   - Tangent vector: A tangent vector at p is the velocity $c'(0)$ of a smooth curve $c : \mathbb{R} \to \mathcal{M}$ with $c(0) = p$.
   - Tangent space: The tangent space $\mathrm{T}_p\mathcal{M}$ is the set of all tangent vectors at p.

   The tangent space $\mathrm{T}_p S^2$ is the set of all the vectors in $\mathbb{R}^3$ that are orthogonal to p, i.e. $\mathrm{T}_p S^2 = \{v \in \mathbb{R}^3\} | v \cdot p = 0\}$.

7. A non-empty intersection of two planes in $\mathbb{R}^3$ is a line. Prove that a non-empty intersection of two spheres in $\mathbb{R}^3$ is a circle (considering a point as a circle of radius 0).

   Solution. Given two spheres in $\mathbb{R}^3$ as follows,

   $$(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2 = R_1^2 \tag{15}$$
   $$(x - x_2)^2 + (y - y_2)^2 + (z - z_2)^2 = R_2^2 \tag{16}$$

   where (x_1, y_1, z_1) and (x_2, y_2, z_2) are two centers of the spheres, and R_1 and R_2 are their radii. Subtracting the first equation from the second equation yields

   $$2(x_1 - x_2)x + 2(y_1 - y_2)y + 2(z_1 - z_2)z = x_1^2 - x_2^2 + y_1^2 - y_2^2 + z_1^2 - z_2^2 + R_2^2 - R_1^2 \tag{17}$$

   which can be rewritten as

   $$Ax + By + Cz = D \tag{18}$$

   where $A = 2(x_1 - x_2)$, $B = 2(y_1 - y_2)$, $C = 2(z_1 - z_2)$, and $D = x_1^2 - x_2^2 + y_1^2 - y_2^2 + z_1^2 - z_2^2 + R_2^2 - R_1^2$. This is the standard form of a plane in $\mathbb{R}^3$. Therefore, the intersection of two spheres lies in this plane. Now we can consider the curve generated by cutting any one of the two spheres with the plane as the claimed intersection in of the question. Then this curve is definitely a circle since the points on the curve are equidistant to each center of the spheres.

8. Let $\gamma : \mathbb{R} \to \mathbb{R}^2$ be a regular unit-speed smooth curve such that all tangent lines to γ intersect in a point. Show that γ is a straight line.

   Proof. Let $\boldsymbol{r}(\mathrm{t})$ be a space curve with regular parametrization. According to the given hypotheses, we have

   $$\boldsymbol{r}(t) - \boldsymbol{R}_0 = \lambda(t)\boldsymbol{r}'(t). \tag{19}$$

Taking derivatives w.r.t. t on both sides gives

$$\boldsymbol{r}'(t) = \lambda'(t)\boldsymbol{r}'(t) + \lambda(t)\boldsymbol{r}''(t) \tag{20}$$

which implies that $\boldsymbol{r}'(t) \times \boldsymbol{r}''(t) = 0$. Therefore, the curvature

$$\kappa(t) = \frac{\boldsymbol{r}'(t) \times \boldsymbol{r}''(t)}{|\boldsymbol{r}'(t)|^3} = 0, \tag{21}$$

showing that $\boldsymbol{r}(\mathrm{t})$ is a straight line. □

9. Let S be a compact connected surface which is not homeomorphic to a sphere. Show that there are points on S where the Gaussian curvature is positive, negative and zero.

10. Let $\boldsymbol{\alpha}(\mathrm{t})$ be a space curve with an arbitrary regular parametrization. Show that its curvature is given by

$$\kappa(t) = \frac{||\boldsymbol{\alpha}' \times \boldsymbol{\alpha}''||}{||\boldsymbol{\alpha}'||^3}.$$

Proof. We can reparameterize $\boldsymbol{\alpha}(t)$ by arclength s, obtaining $\boldsymbol{\beta}(s)$. Then we have $\boldsymbol{\alpha}(t) = \boldsymbol{\beta}(s(t))$. By the chain rule, we get

$$\boldsymbol{\alpha}' = \frac{\mathrm{d}s(t)}{\mathrm{d}t}\frac{\mathrm{d}\boldsymbol{\beta}(s(t))}{\mathrm{d}s} = v(t)\mathbf{T}(s) \tag{22}$$

where $v(t) = \frac{\mathrm{d}s(t)}{\mathrm{d}t}$ is the speed and $\mathbf{T}$ represents the unit tangent vector. Furthermore, we have

$$\mathbf{T}(s(t))' = \frac{\mathrm{d}s(t)}{\mathrm{d}t}\frac{\mathrm{d}\mathbf{T}(s(t))}{\mathrm{d}s} = v(t)k(s(t))\mathbf{N}(s(t)) \tag{23}$$

where $k(s(t))$ denotes the curvature and $\mathbf{N}(s(t))$ represents the principal normal vector. Finally, we get

$$\boldsymbol{\alpha}'' = v'(t)\mathbf{T}(s(t)) + v(t)^2 k(s(t))\mathbf{N}(s(t)). \tag{24}$$

Now we have

$$\boldsymbol{\alpha}' \times \boldsymbol{\alpha}'' = (v(t)\mathbf{T}(s(t))) \times \left(v(t)v'(t)\mathbf{T}(s(t)) + v(t)^2 k(s(t))\mathbf{N}(s(t))\right) \tag{25}$$

$$= v(t)^3 k(s(t))\mathbf{T}(s(t)) \times \mathbf{N}(s(t)) \tag{26}$$

Since both the curvature k and the speed v are nonnegative, then $v(t)^3 k(s(t)) = \|\boldsymbol{\alpha}' \times \boldsymbol{\alpha}''\|$ holds, which follows from $\|\mathbf{T}(s(t)) \times \mathbf{N}(s(t))\| = 1$. According to (22), $\|\boldsymbol{\alpha}'\| = v(t)\|\mathbf{T}(s)\| = v(t)$. Let $\kappa(t) = k(s(t))$, then we get the desired the result

$$\kappa(t) = \frac{\|\boldsymbol{\alpha}' \times \boldsymbol{\alpha}''\|}{v(t)^3} = \frac{\|\boldsymbol{\alpha}' \times \boldsymbol{\alpha}''\|}{\|\boldsymbol{\alpha}'\|^3}. \tag{27}$$

□

11. Define the catenoid as the surface S_f parametrized by

$$f(u, v) = (\cosh u \cos v, \cosh u \sin v, u)$$

and the helicoid as the surface S_g parametrized by

$$g(u, v) = (\sinh u \cos v, \sinh u \sin v, v).$$

(a) Show that, for any point $(u, v) \in \mathbb{R}^2$, the Gaussian curvature of S_f at $f(u, v)$ coincides with the Gaussian curvature of S_g at $g(u, v)$.

(b) Show that the mean curvature of both S_f and S_g is identically zero.

6 Probability theory

References: [10], [9]

1. Define a probability space and define every object involved in the definition.

   Solution. A probability space or a probability triple $(\Omega, \mathcal{F}, P)$ consists of three elements:

   - A sample space, Ω, which is the set of all possible outcomes.
   - An event space, which is a set of events, $\mathcal{F}$, an event being a set of outcomes in the sample space.
   - A probability function (measure), P, which assigns, to each event in the event space, a probability, a number being between 0 and 1 (inclusive).

2. State what sigma-additivity or countable additivity of a probability measure means.

   Solution. The probability measure $P : \mathcal{F} \to [0, 1]$ - a function on $\mathcal{F}$ such that:

   - P is countably additive (a.k.a. σ-additive): if $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$ is a countable collection of pairwise disjoint sets, then $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.
   - the measure of the entire space is equal to 1: $P(\Omega) = 1$.

3. Let E be a nonempty set. Give the definition of an E-valued random variable X on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then, define the cumulative distribution function (c.d.f.) of X and state all the properties that imply a function $F : \mathbb{R} \to \mathbb{R}$ is the cumulative distribution function of a $\mathbb{R}$-valued random variable.

   Solution. An E-valued random variable X is a measurable function $X : \Omega \to E$.

   The cumulative distribution function (c.d.f.) of X is defined as

   $$F_X(x) = \mathbb{P}(X \le x), \ \forall x \in \mathbb{R}.$$

   Every cumulative function F_X is non-decreasing and right-continuous. Furthermore,

   $$\lim_{x \to -\infty} F_X(x) = 0, \qquad \lim_{x \to +\infty} F_X(x) = 1.$$

4. Give the definition of mutual independence of a collection of measurable events, and of a collection of random variables. Let X and Y be two $\mathbb{R}$-valued random variables on a probability space. Define such that the covariance of X and Y is zero. Does this imply that X and Y are independent? If yes, then give a proof. If not, then give a counterexample.

   Solution.

   - Mutual independence of a collection of measurable events: A collection of measurable events $\{A_i\}_{i \in I}$ is said to be mutually independent if for any finite subset $\{A_{i_1}, A_{i_2}, \ldots, A_{i_n}\}$ of these events, the following holds:

     $$\mathbb{P}\left(\bigcap_{j=1}^{n} A_{i_j}\right) = \prod_{j=1}^{n} \mathbb{P}(A_{i_j}). \tag{28}$$

     This means that the probability of the intersection of any finite number of these events is equal to the product of their individual probabilities.

- Mutual independence of a collection of random variables: A collection of random variables $\{X_i\}_{i\in I}$ is said to be mutually independent if for any finite subset $\{X_{i_1}, X_{i_2}, \dots, X_{i_n}\}$ of these events, and for any collection of Borel sets $\{B_{i_1}, B_{i_2}, \dots, B_{i_n}\}$, the following holds:
$$\mathbb{P}\left(X_{i_1} \in B_{i_1}, X_{i_2} \in B_{i_2}, \dots, X_{i_n} \in B_{i_n}\right) = \prod_{j=1}^{n} \mathbb{P}(X_{i_j} \in B_{i_j}). \tag{29}$$
This means that the joint probability distribution of the random variables, evaluated at any finite number of Borel sets, is equal to the product of their marginal distributions evaluated at these sets.
- Given the covariance of X and Y is zero, it is not necessarily true that X and Y are independent. For example[7], consider three outcomes of (X, Y), $(-1, 1)$, $(0, -2)$, and $(1, 1)$, all with the same probability $1/3$. Obviously, they are not independent because X determines Y. However, their covariance is 0 as shown below.
$$\operatorname{cov}(X, Y) = E[XY] - E[X]E[Y] = E[XY] = \frac{1}{3} \times (-1) + \frac{1}{3} \times 0 + \frac{1}{3} \times 1 = 0. \tag{30}$$

5. If two random variables are equal in distribution, does this imply they are equal almost surely? If yes, then give a proof. If not, then give a counterexample.

   Solution. No. Let $X, Y : [0, 1] \to \mathbb{R}$ be defined by $x \mapsto x$ and $x \mapsto 1 - x$ respectively. Since $X = Y$ if only if $x = 1/2$, then $P(X \neq Y) = 1$. However, both X and Y have the uniform distribution on $[0, 1]$[6].

6. Let (X, Y) be a $\mathbb{R}^2$-valued random variable with joint p.d.f. $f_{X,Y}$. Define the marginal p.d.f. f_X and the conditional p.d.f. $f_{X|Y}$.

   Solution.

   - The marginal p.d.f. f_X:
   $$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x, y)dY, \quad \forall x \in \mathbb{R} \tag{31}$$
   - The conditional p.d.f. $f_{X|Y}$:
   $$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}. \tag{32}$$

7. Let X be a $\mathbb{R}^d$-valued random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $p \geq 1$. Define what it means for X to belong to $L^p(\mathbb{P})$.

   Solution. $X \in L^p(\mathbb{P})$ if
   $$\int_{\Omega} \|X(\omega)\|^p \, d\mathbb{P}(\omega) < \infty \tag{33}$$
   where $\|X(\omega)\|$ denotes the Euclidean norm of X in $\mathbb{R}^d$.

8. Give the definition of a probability density function (p.d.f.) of a $\mathbb{R}^d$-valued random variable. Write down the formula for the p.d.f. of a $\mathbb{R}^d$-valued normal random variable with mean $m \in \mathbb{R}^d$ and strictly positive definite covariance matrix $C \in \mathbb{R}^{d\times d}$. Make sure to specify all the necessary assumptions on m and C for the equation to be valid.

   Solution. A probability density function (p.d.f.) of a $\mathbb{R}^d$-valued random variable X is a function $f_X : \mathbb{R}^d \to [0, \infty)$ that satisfies the following properties:

   - Non-negativity: $f_X(x) \geq 0$ for all $x \in \mathbb{R}^d$.
   - Normalization: The integral of f_X over the entire space $\mathbb{R}^d$ equals 1:

   $$\int_{\mathbb{R}^d} f_X(x)\, dx = 1.$$

   - Probability measure: For any Borel set $A \subseteq \mathbb{R}^d$,

   $$\mathbb{P}(X \in A) = \int_A f_X(x)\, dx.$$

   Let X be a $\mathbb{R}^d$-valued normal random variable with mean vector $m \in \mathbb{R}^d$ and a strictly positive definite covariance matrix $C \in \mathbb{R}^{d\times d}$. The p.d.f. of X is given by:

   $$f_X(x) = \frac{1}{(2\pi)^{d/2}\sqrt{\det(C)}} \exp\left(-\frac{1}{2}(x-m)^T C^{-1}(x-m)\right),$$

   where $m \in \mathbb{R}^d$ is the mean vector, $C \in \mathbb{R}^{d\times d}$ is the strictly positive definite covariance matrix, meaning C is symmetric and all its eigenvalues are positive, ensuring that C^{-1} exists and is also positive definite, and $\det(C)$ denotes the determinant of the covariance matrix C.

References

[1] Wikipedia contributors. Measurable function. Wikipedia, 2024. `https://en.wikipedia.org/wiki/Measurable_function`.

[2] Wikipedia contributors. Measurable space. Wikipedia, 2024. `https://en.wikipedia.org/wiki/Measurable_space`.

[3] Wikipedia contributors. Measure (mathematics). Wikipedia, 2024. `https://en.wikipedia.org/wiki/Measure_(mathematics)`.

[4] Wikipedia contributors. Picard–lindelöf theorem — wikipedia, the free encyclopedia. `https://en.wikipedia.org/wiki/Picard%E2%80%93Lindel%C3%B6f_theorem`, 2024. [Online; accessed 14-July-2024].

[5] Keith Devlin. Mathematical thinking. Coursera, 2024. `https://www.coursera.org/learn/mathematical-thinking`.

[6] drhab. Equal distribution does not imply eqaul random variables. Mathematics Stack Exchange, 2020. `https://math.stackexchange.com/questions/3593226/prove-random-variables-are-not-almost-surely-equal-but-are-equal-in-distributio`.

[7] Prof. D. Joyce. Covariance and correlation. Clark University, 2014. https://math.clarku.edu/~djoyce/ma217/covar.pdf.

[8] Terence Tao. *Analysis I*. Hindustan Book Agency, New Delhi, 3rd edition, 2006.

[9] Wikipedia contributors. Cumulative distribution function. https://en.wikipedia.org/wiki/Cumulative_distribution_function, 2024. Accessed: 2024-07-13.

[10] Wikipedia contributors. Probability space. https://en.wikipedia.org/wiki/Probability_space, 2024. Accessed: 2024-07-13.

References

[1] Rick Durrett, *Elementary probability for applications*, Cambridge: Cambridge University Press, 2009 (English).

[2] Joseph E. Fields, *A Gentle Introduction to the Art of Mathematics*, https://aimath.org/textbooks/approved-textbooks/fields/.

[3] Charles M. Grinstead and J. Laurie Snell, *Introduction to Probability*, https://aimath.org/textbooks/approved-textbooks/grinstead-snell/.

[4] Richard Hammack, *Book of Proof*, https://aimath.org/textbooks/approved-textbooks/hammack/.

[5] Oliver Knill, *Linear Algebra and Vector Calculus I*, https://www.ams.org/open-math-notes/omn-view-listing?listingId=110802.

[6] Jörg Liesen and Volker Mehrmann, *Linear algebra*, Springer, 2015.

[7] Dirk Lorenz, *Measure Theory with a Geometric Touch*, https://www.ams.org/open-math-notes/omn-view-listing?listingId=110655.

[8] Walter Rudin, *Real and complex analysis*, 3rd ed., New York, NY: McGraw-Hill, 1987.

[9] Theodore Shifrin, *Differential Geometry: A First Course in Curves and Surfaces*, https://www.ams.org/open-math-notes/omn-view-listing?listingId=10755.

[10] Ted Sundstrom, *Mathematical Reasoning: Writing and Proof*, https://aimath.org/textbooks/approved-textbooks/sundstrom/.

[11] Irena Swanson, *Introduction to Analysis*, https://www.ams.org/open-math-notes/omn-view-listing?listingId=110809.

[12] Sergei Treil, *Linear Algebra Done Wrong*, https://aimath.org/textbooks/approved-textbooks/treil/.

[13] William F. Trench, *Introduction to Real Analysis*, https://aimath.org/textbooks/approved-textbooks/trench/.

[14] Elias Zakon, *Mathematical Analysis: Volume I*, https://www.trillia.com/zakon-analysisI.html.

[15] —, *Mathematical Analysis: Volume II*, https://www.trillia.com/zakon-analysisII.html.