Chapter 22

Supplemental Analysis Results

In this chapter, we collect a set of supplementary results that can be used to streamline proofs and to broaden the range of applications of Part II. Many existence, uniqueness, and approximation statements can be reduced to structural principles: completeness (**Banach fixed point**), duality (**bounded linear functionals**), extension (**Hahn-Banach**), genericity (**Baire category**), compactness in function spaces (**Arzelà-Ascoli**), approximation by simple families (**Stone-Weierstrass**), and orthogonality in inner product spaces (**Hilbert space methods**). We end with two classical local theorems in multivariable analysis (**inverse and implicit function theorems**) and an application to **initial value problems**.

22.1 The Banach Fixed Point Theorem

Contractions And Complete Metric Spaces

Let (X, d) be a metric space. A map $T : X \to X$ is called a **contraction** if there exists a constant $q \in (0, 1)$ such that

$$d\big(T(\mathbf{x}), T(\mathbf{y})\big) \le q\, d(\mathbf{x}, \mathbf{y}) \quad \text{for all } \mathbf{x}, \mathbf{y} \in X.$$

The constant q is called a **contraction factor** (or contraction constant).

Examples

1. Let $(\mathbb{R}^n, \|\cdot\|)$ be equipped with the metric $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$ (any norm). Fix $q \in (0, 1)$ and define $T(\mathbf{x}) = q\mathbf{x}$. Then

$$d(T(\mathbf{x}), T(\mathbf{y})) = \|q\mathbf{x} - q\mathbf{y}\| = q\|\mathbf{x} - \mathbf{y}\| = q\, d(\mathbf{x}, \mathbf{y}),$$

so T is a contraction with contraction factor q.

2. On $(\mathbb{R}, |\cdot|)$, fix $a \in \mathbb{R}$ with $|a| < 1$ and $b \in \mathbb{R}$, and define

$$T(x) = ax + b.$$

Then

$$|T(x) - T(y)| = |a|\,|x - y| \le q\,|x - y| \quad \text{with} \quad q = |a| \in (0, 1),$$

so T is a contraction. □

3. Let (X, d) be any metric space and fix $\mathbf{x}_0 \in X$. Define $T(\mathbf{x}) = \mathbf{x}_0$ for all $x \in X$. Then

$$d(T(\mathbf{x}), T(\mathbf{y})) = d(\mathbf{x}_0, \mathbf{x}_0) = 0 \le q\, d(\mathbf{x}, \mathbf{y})$$

for every $q \in (0, 1)$, so every constant map is a contraction.

4. On $([0, 1], |\cdot|)$ define

$$T(x) = \frac{x^2}{3}.$$

For $x, y \in [0, 1]$,

$$|T(x) - T(y)| = \frac{1}{3}|x^2 - y^2| = \frac{1}{3}|x - y|\,|x + y| \le \frac{2}{3}|x - y|.$$

Thus T is a contraction with factor $q = \frac{2}{3}$.

5. Let $X = C([0, 1])$ with $\|f\|_\infty = \sup_{t \in [0,1]} |f(t)|$ and let $d(f, g) = \|f - g\|_\infty$. Fix $\lambda \in \mathbb{R}$ with $|\lambda| < 1$ and define

$$(Tf)(t) = \lambda \int_0^t f(s)\, ds.$$

Then for all $f, g \in X$,

$$\begin{aligned}\|Tf - Tg\|_\infty &= \sup_{t \in [0,1]} \left| \lambda \int_0^t (f - g)(s)\, ds \right| \le |\lambda| \sup_{t \in [0,1]} \int_0^t |(f - g)(s)|\, ds \\ &\le |\lambda| \sup_{t \in [0,1]} \int_0^t \|f - g\|_\infty\, ds \le |\lambda|\, \|f - g\|_\infty.\end{aligned}$$

So T is a contraction with factor $q = |\lambda|$.

6. On $(0, \infty)$ define $d(x, y) = |\ln x - \ln y|$. Fix $a \in (0, 1)$ and set $T(x) = x^a$. Then

$$d(T(x), T(y)) = |\ln(x^a) - \ln(y^a)| = a|\ln x - \ln y| = a\, d(x, y),$$

so T is a contraction with factor $q = a$.

Remark: if X has the discrete metric $d(\mathbf{x}, \mathbf{y}) = 1$ for $\mathbf{x} \neq \mathbf{y}$, then the only maps $T : X \to X$ that are contractions are the constant maps (otherwise one would have

$$1 = d(T(\mathbf{x}), T(\mathbf{y})) \leq q d(\mathbf{x}, \mathbf{y}) \leq q < 1$$

for some $\mathbf{x} \neq \mathbf{y}$).

Every contraction is Lipschitz (see exercises), hence uniformly continuous (see Chapter 3).

Existence, Uniqueness, and Iterative Convergence

Theorem (Banach Fixed Point Theorem)
Let (X, d) be a complete metric space and let $T : X \to X$ be a contraction with factor $q \in (0, 1)$. Then T has a unique fixed point $\mathbf{x}^* \in X$, and for every $\mathbf{x}_0 \in X$, the iterates $\mathbf{x}_{n+1} = T(\mathbf{x}_n)$ converge to $\mathbf{x}^*$.

Proof: this is the content of Theorem 114, applied to (X, d) and T. ■

Quantitative Error Bounds

Corollary 303 (A Priori Error Bound)
With the hypotheses of the Banach fixed point theorem, the iterates satisfy

$$d(\mathbf{x}_n, \mathbf{x}^*) \leq \frac{q^n}{1-q} d(\mathbf{x}_1, \mathbf{x}_0) \quad \text{for all } n \geq 0.$$

Proof:

■

Corollary 304 (A Posteriori Error Bound)
With the hypotheses of the Banach fixed point theorem, the iterates satisfy

$$d(\mathbf{x}_n, \mathbf{x}^*) \leq \frac{1}{1-q} d(\mathbf{x}_{n+1}, \mathbf{x}_n) \quad \text{for all } n \geq 0.$$

Proof:

■

Example: Consider the equation $x = \cos x$ on $\mathbb{R}$. Let $T(x) = \cos x$ and consider the interval $X = [0,1]$. Since $\cos([0,1]) \subseteq [\cos(1), 1] \subseteq [0,1]$, we have $T(X) \subseteq X$. Moreover, for all $x, y \in X$, the mean value theorem gives

$$|T(x) - T(y)| = |\cos x - \cos y| = |\sin(\xi)|\,|x - y| \leq \sin(1)\,|x - y|$$

for some ξ between x and y. Hence T is a contraction on X with factor $q = \sin(1) \in (0,1)$. Since $(X, |\cdot|)$ is complete, the Banach fixed point theorem implies that there is a unique $x^* \in X$ such that $x^* = \cos(x^*)$, and that the iteration $x_{n+1} = \cos(x_n)$ converges to x^* for every $x_0 \in X$. The a posteriori bound (Corollary 304) can be used to certify an accuracy target once $|x_{n+1} - x_n|$ is small. □

22.2 Normed Spaces, Linear Functionals, and Duality

A **normed vector space** is a vector space X over $\mathbb{K}$ equipped with a map $\|\cdot\| : X \to [0, \infty)$ satisfying

$$\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}, \qquad \|\alpha \mathbf{x}\| = |\alpha|\,\|\mathbf{x}\|, \qquad \|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|.$$

The norm induces a metric $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$.

Let X, Y be normed vector spaces over $\mathbb{K}$ and let $T : X \to Y$ be linear. We say that T is **bounded** if there exists $M \geq 0$ such that

$$\|T(\mathbf{x})\| \leq M\|\mathbf{x}\| \qquad \text{for all } \mathbf{x} \in X.$$

The smallest such M is the **operator norm** of T, denoted by

$$\|T\|_{\text{op}} = \sup_{\|\mathbf{x}\| \leq 1} \|T(\mathbf{x})\| = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|T(\mathbf{x})\|}{\|\mathbf{x}\|}.$$

By Theorem 140, for linear maps between normed spaces, boundedness is equivalent to continuity (and to continuity at 0).

We denote by $L_c(X, Y)$ the vector space of all **continuous linear** maps from X to Y.

Linear Functionals, The Dual Space, and Completeness

A **linear functional** on a normed space X is a linear map $\varphi : X \to \mathbb{K}$. The space of all bounded linear functionals is the **dual space** X^*, equipped with the norm

$$\|\varphi\| = \sup_{\|\mathbf{x}\|\leq 1} |\varphi(\mathbf{x})|.$$

Equivalently, $X^* = L_c(X, \mathbb{K})$.

Proposition 305
For every normed space X, the dual space $(X^*, \|\cdot\|)$ is a Banach space.

Proof: since $X^* = L_c(X, \mathbb{K})$ and $(\mathbb{K}, |\cdot|)$ is a Banach space, Theorem 141 implies that $L_c(X, \mathbb{K})$ is Banach. Hence $(X^*, \|\cdot\|)$ is complete. ■

Example.

1. If X is finite-dimensional, then every linear functional $\varphi : X \to \mathbb{K}$ is bounded (hence continuous). This follows from Corollary 143.
2. If K is compact and $X = C(K)$ with $\|f\|_\infty = \sup_{\mathbf{t}\in K} |f(\mathbf{t})|$, then for each $\mathbf{x} \in K$, the evaluation functional $\delta_{\mathbf{x}} : C(K) \to \mathbb{K}$ defined by $\delta_{\mathbf{x}}(f) = f(\mathbf{x})$ satisfies
$$|\delta_{\mathbf{x}}(f)| \leq \|f\|_\infty \quad \text{for all } f \in C(K),$$
and therefore $\|\delta_{\mathbf{x}}\| \leq 1$. Moreover, $\|\delta_{\mathbf{x}}\| = 1$ since $\delta_{\mathbf{x}}(\mathbf{1}) = 1$ and $\|\mathbf{1}\|_\infty = 1$, for the constant function $\mathbf{1}(\mathbf{x}) \equiv 1$. □

Finite-Dimensional Norm Equivalence

The equivalence of all norms in finite-dimensional spaces, as well as its standard consequences, are established in Proposition 142 and Corollaries 143-146. In particular, in finite dimension all norms induce the same notion of convergence and continuity, every linear map is continuous, every finite-dimensional normed space is complete, and compactness is characterized by closedness and boundedness.

22.3 The Hahn-Banach Theorem

A function $p : X \to \mathbb{R}$ on a real vector space X is **sub-linear** if for every $\mathbf{x}, \mathbf{y} \in X$, we have

$$p(\mathbf{x}+\mathbf{y}) \leq p(\mathbf{x}) + p(\mathbf{y}) \quad \text{and} \quad p(\alpha\mathbf{x}) = \alpha p(\mathbf{x}) \text{ for all } \alpha \geq 0.$$

Examples

1. If $(X, \|\cdot\|)$ is a normed vector space, then $p(\mathbf{x}) = \|\mathbf{x}\|$ is sub-linear thanks to the triangle inequality. More generally, any semi-norm $q : X \to [0, \infty)$ is sub-linear.

2. If $\varphi_1, \ldots, \varphi_m : X \to \mathbb{R}$ are linear, then

$$p(\mathbf{x}) = \max_{1 \leq i \leq m} \{\varphi_i(\mathbf{x})\}$$

is sub-linear. Likewise, if $\mathcal{F}$ is a non-empty family of linear functionals such that $p(\mathbf{x}) = \sup_{\varphi \in \mathcal{F}} \varphi(\mathbf{x}) \in \mathbb{R}$ for all $\mathbf{x} \in X$, then p is sub-linear.

3. On $X = \mathbb{R}$, the function

$$p(t) = t^+ = \max\{t, 0\}$$

is sub-linear.

4. Let $C \subseteq X$ be convex, with $\mathbf{0} \in C$ and so that for each $\mathbf{x} \in X$ there exists $\lambda > 0$ with $\mathbf{x} \in \lambda C = \{\lambda \mathbf{z} \mid \mathbf{z} \in C\}$. Define the **Minkowski functional** on C by

$$p_C(\mathbf{x}) = \inf\{\lambda > 0 \mid \mathbf{x} \in \lambda C\}.$$

Then $p_C : X \to [0, \infty)$ is sub-linear. □

The next result is a step on the way to the general Hahn-Banach Theorem.

Lemma 306 (One-Dimensional Extension)
Let X be a real vector space, let $p : X \to \mathbb{R}$ be sub-linear, let $Y < X$ be a subspace, and let $\varphi_0 : Y \to \mathbb{R}$ be linear with $\varphi_0(\mathbf{y}) \leq p(\mathbf{y})$ for all $\mathbf{y} \in Y$. If $\mathbf{x}_0 \in X \setminus Y$, then there exists a linear map $\varphi_1 : Y + \text{span}\{\mathbf{x}_0\} \to \mathbb{R}$ such that $\varphi_1|_Y = \varphi_0$ and $\varphi_1(\mathbf{x}) \leq p(\mathbf{x})$ for all $\mathbf{x} \in Y + \text{span}\{\mathbf{x}_0\}$.

Proof:

■

We can use Lemma 306 repeatedly to prove the finite-dimensional version of the Hahn-Banach Theorem.

Theorem 307 (Hahn-Banach, Finite-Dimensional Analytic Form)
Let X be a real vector space of finite dimension, let $p : X \to \mathbb{R}$ be sub-linear, and let $Y < X$ be a subspace. If $\varphi_0 : Y \to \mathbb{R}$ is linear and satisfies $\varphi_0(\mathbf{y}) \leq p(\mathbf{y})$ for all $\mathbf{y} \in Y$, then there exists a linear extension $\varphi : X \to \mathbb{R}$ such that $\varphi|_Y = \varphi_0$ and $\varphi(\mathbf{x}) \leq p(\mathbf{x})$ for all $\mathbf{x} \in X$.

Proof: if $Y = X$, there is nothing to prove. Otherwise, choose a basis $(\mathbf{e}_1, \ldots, \mathbf{e}_k)$ of Y and extend it to a basis $(\mathbf{e}_1, \ldots, \mathbf{e}_k, \mathbf{e}_{k+1}, \ldots, \mathbf{e}_n)$ of X. Define $Y_k = Y$ and for $j = k+1, \ldots, n$ set

$$Y_j = Y_{j-1} + \text{span}\{\mathbf{e}_j\}.$$

Starting from $\varphi_k = \varphi_0$ on Y_k, apply Lemma 306 successively to extend φ_{j-1} from Y_{j-1} to Y_j while preserving the domination $\varphi_j \leq p$. After finitely many steps we obtain a linear map $\varphi_n : X \to \mathbb{R}$ extending φ_0 and satisfying $\varphi_n \leq p$ on X. Take $\varphi = \varphi_n$. ■

The infinite-dimensional case is a harder to prove, however.

Theorem 308 (Hahn-Banach, Classical Form)
Let $(X, \|\cdot\|)$ be a real normed vector space, and let $Y < X$ be a linear subspace. Assume that X is separable. If $\varphi_0 : Y \to \mathbb{R}$ is a bounded linear functional, then there exists a bounded linear functional $\varphi : X \to \mathbb{R}$ such that $\varphi|_Y = \varphi_0$ and

$$\|\varphi\| = \|\varphi_0\|.$$

Proof: left as an exercise. ■

Applications

Theorem 308 is one of the main tools for showing that a normed space has "enough" continuous linear functionals. Here are some standard applications.

Corollary 309 (Normed-Space Extension)
Let X be a real normed space and $Y \subseteq X$ a subspace. If $\varphi_0 \in Y^*$, then there exists $\varphi \in X^*$ such that $\varphi|_Y = \varphi_0$ and $\|\varphi\| = \|\varphi_0\|$.

Proof:

■

Proposition 310 (Norming Functional On A Line)
Let X be a real normed space and $x \in X$, $x \neq 0$. Then there exists $\varphi \in X^*$ such that $\|\varphi\| = 1$ and $\varphi(x) = \|x\|$.

Proof: let $Y = \text{span}\{x\}$ and define $\varphi_0 : Y \to \mathbb{R}$ by $\varphi_0(\alpha x) = \alpha\|x\|$. Then

$$|\varphi_0(\alpha x)| = |\alpha|\,\|x\| = \|\alpha x\|,$$

so $\|\varphi_0\| = 1$ on Y. By Corollary 309, φ_0 extends to $\varphi \in X^*$ with $\|\varphi\| = 1$. In particular, $\varphi(x) = \varphi_0(x) = \|x\|$. ■

A subset C of a real vector space X is called **convex** if for all $x, y \in C$ and all $t \in [0, 1]$,

$$tx + (1 - t)y \in C.$$

Theorem 311 (Separation Of A Point From A Closed Convex Set)
Let X be a real normed space, let $C \subseteq X$ be non-empty, closed, and convex, and let $\mathbf{x}_0 \notin C$. Then there exist $\varphi \in X^*$ and $\alpha \in \mathbb{R}$ such that $\varphi(\mathbf{x}) \leq \alpha < \varphi(\mathbf{x}_0)$ for all $x \in C$.

Proof:

■

22.4 Baire Category Theorem

A subset A of a metric space X is **nowhere dense** if $\overline{A}$ has empty interior.

Examples

1. If $A \subseteq \mathbb{R}$ is finite, then $\overline{A} = A$ and $\text{int}(A) = \varnothing$ in the standard topology, so A is nowhere dense.
2. Since $\overline{\mathbb{Z}} = \mathbb{Z}$ and $\text{int}(\mathbb{Z}) = \varnothing$ in the standard topology, $\mathbb{Z}$ is nowhere dense.
3. Let $U \subseteq_O \mathbb{R}^n$. Then its boundary $\partial U \subseteq_C \mathbb{R}^n$ has empty interior, so ∂U is nowhere dense.
4. Let $U \subseteq_O X$ be non-empty. Then $U \subseteq \overline{U}$, so $\text{int}(\overline{U}) \supseteq U \neq \varnothing$ and so U is not nowhere dense.
5. In the standard topology, we have $\overline{\mathbb{Q}} = \mathbb{R}$, so $\text{int}(\overline{\mathbb{Q}}) = \text{int}(\mathbb{R}) = \mathbb{R} \neq \varnothing$. Thus $\mathbb{Q}$ is not nowhere dense.
6. In the standard topology $A \subseteq \mathbb{R}$ contains a non-degenerate interval (a, b), then $(a, b) \subseteq \text{int}(A) \subseteq \text{int}(\overline{A})$, so A is not nowhere dense. □

A set is of **first category** (or **meagre**) if it is a countable union of nowhere dense sets.

Examples

1. $\mathbb{R}$ is not meagre in itself in the standard topology, as a consequence of the Baire category theorem (Theorem 312).
2. Since $\mathbb{Q}$ is meagre in $\mathbb{R}$ (left as an exercise) and $\mathbb{R}$ is not meagre, its complement $\mathbb{R} \setminus \mathbb{Q}$ is not meagre.

3. If $A = \{a_1, a_2, \dots\}$, then
$$A = \bigcup_{n=1}^{\infty} \{a_n\},$$
and each singleton $\{a_n\}$ is nowhere dense. Hence A is meagre. In particular, $\mathbb{Q}$ and $\mathbb{Z}$ are meagre in $\mathbb{R}$.

4. Let $X = C([0,1])$ with $\|\cdot\|_\infty$, and let P_n be the subspace of polynomials of degree $\leq n$ (viewed as elements of $C([0,1])$). Each P_n is finite-dimensional, hence closed, and has empty interior in X, so it is nowhere dense. Therefore
$$\mathcal{P} := \bigcup_{n=0}^{\infty} P_n$$
is meagre in X. On the other hand, $\mathcal{P}$ is dense in X according to the Weierstrass approximation theorem (see Section 22.6). □

These concepts lead to the following result.

Theorem 312 (Baire Category Theorem)
If (X,d) is a complete metric space, then X is not meagre in itself. Equivalently, if $(U_n)_{n\in\mathbb{N}}$ is a sequence of dense open subsets of X, then $\bigcap_{n=1}^{\infty} U_n$ is dense in X.

Proof:

■

Let's take a look at a first important applications.

Theorem 313 (Uniform Boundedness Principle)
Let X be a Banach space and Y a normed space. Let $(T_n)_{n\in\mathbb{N}} \subseteq \mathcal{L}(X,Y)$, the space of bounded linear maps from X to Y. If for every $\mathbf{x} \in X$ the set $\{\|T_n(\mathbf{x})\| \mid n \in \mathbb{N}\}$ is bounded, then

$$\sup_{n\in\mathbb{N}} \{\|T_n\|_{\text{op}}\} < \infty.$$

Proof:

■

There are other classical applications of Baire's theorem,; we will discuss two more: the open mapping theorem and the closed graph theorem.

Theorem 314 (Open Mapping)
Let X, Y be Banach spaces and let $T : X \to Y$ be a surjective bounded linear map. Then T maps open sets to open sets.

Proof:

■

Theorem 315 (Closed Graph)
Let X, Y be Banach spaces and let $T : X \to Y$ be linear. If the graph

$$\mathcal{G}(T) = \{(\mathbf{x}, T(\mathbf{x})) \mid x \in X\} \subseteq X \times Y$$

is closed in $X \times Y$, then T is bounded.

Proof: consider $X \times Y$ with the norm $\|(\mathbf{x}, \mathbf{y})\| = \|\mathbf{x}\| + \|\mathbf{y}\|$, which makes it a Banach space. The graph $\mathcal{G}(T)$ is a closed subspace, hence Banach. Define $S : \mathcal{G}(T) \to X$ by $S(\mathbf{x}, T(\mathbf{x})) = \mathbf{x}$. Then S is linear, bounded, and bijective. Its inverse is $S^{-1}(\mathbf{x}) = (\mathbf{x}, T(\mathbf{x}))$. By Theorem 314, S^{-1} is bounded, so there exists C such that

$$\|\mathbf{x}\| + \|T(\mathbf{x})\| = \|S^{-1}(\mathbf{x})\| \leq C\|\mathbf{x}\|.$$

Thus $\|T(\mathbf{x})\| \leq (C - 1)\|\mathbf{x}\|$ for all $\mathbf{x}$, hence T is bounded. ■

22.5 Function Spaces And Compactness

Let K be a compact metric space. Denote by $C(K)$ the vector space of continuous functions $f : K \to \mathbb{R}$, equipped with the **sup norm**

$$\|f\|_\infty = \sup_{\mathbf{x}\in K} |f(\mathbf{x})|.$$

Proposition 316
$(C(K), \|\cdot\|_\infty)$ is a Banach space.

Proof:

■

A family $\mathcal{F} \subseteq C(K)$ is **pointwise bounded** if for each $\mathbf{x} \in K$, the set $\{f(\mathbf{x}) \mid f \in \mathcal{F}\}$ is bounded in $\mathbb{R}$. It is **equi-continuous** if for every $\varepsilon > 0$ there exists $\delta > 0$ such that for all $f \in \mathcal{F}$ and all $\mathbf{x}, \mathbf{y} \in K$,

$$d(\mathbf{x}, \mathbf{y}) < \delta \Longrightarrow |f(\mathbf{x}) - f(\mathbf{y})| < \varepsilon.$$

A subset A of a metric space X is **relatively compact** if $\overline{A} \subseteq_K X$.

In $\mathbb{R}^n$, the Bolzano-Weierstrass theorem can be read as a compactness criterion: boundedness implies relative compactness, so every bounded sequence has a convergent sub-sequence. The next result plays an analogous role in spaces of continuous functions.

Theorem 317 (Arzelà-Ascoli)
Let K be a compact metric space. A subset $\mathcal{F} \subseteq C(K)$ is relatively compact in $(C(K), \|\cdot\|_\infty)$ if and only if $\mathcal{F}$ is pointwise bounded and equi-continuous. Equivalently, every sequence in $\mathcal{F}$ has a uniformly convergent sub-sequence.

Proof:

■

Applications

1. A common use of Arzelà-Ascoli is to show that an operator $T : C(K) \to C(K)$ is compact by proving that T maps bounded subsets of $C(K)$ into families that are uniformly bounded and equi-continuous. Classical examples include many integral operators of the form
$$(Tf)(\mathbf{x}) = \int_K k(\mathbf{x}, \mathbf{y}) f(\mathbf{y})\, d\mathbf{y},$$
with continuous kernels k.

2. In Peano-type existence proofs, we construct a sequence of approximate solutions (for instance, Euler polygonal approximations) that is uniformly bounded and equi-continuous

on a compact time interval. Arzelà-Ascoli yields a uniformly convergent sub-sequence, and one then passes to the limit to obtain a solution.

3. When minimizing a functional over a class of continuous functions, we often show that a minimizing sequence is uniformly bounded and equi-continuous. Arzelà-Ascoli then provides a uniformly convergent sub-sequence, and additional hypotheses (such as lower semi-continuity) allow us to conclude that the limit is a minimizer.

4. If $\mathcal{F} \subseteq C^1([a,b])$ is uniformly bounded and the derivatives $\{f' \mid f \in \mathcal{F}\}$ are uniformly bounded, then $\mathcal{F}$ is equi-continuous. Arzelà-Ascoli implies that $\mathcal{F}$ is relatively compact in $C([a,b])$ (with the sup norm). This idea is a basic compactness tool in analysis and partial differential equations

22.6 Weierstrass Approximation Theorem

Dense subsets $\mathcal{A} \subseteq C(K)$ of ("simple") families of functions can be used to approximate much more complicated ones. For instance, the set of polynomials in one variable over the interval $[a,b]$ is dense in the set of continuous functions $C([a,b])$ with respect to $\|\cdot\|_\infty$.

Theorem 318 (Weierstrass Approximation Theorem)
Let $f:[a,b]\to\mathbb{R}$ be continuous. For every $\varepsilon>0$ there exists a polynomial p such that

$$\|f-p\|_\infty = \sup_{x\in[a,b]} |f(x)-p(x)| < \varepsilon.$$

We first prove the theorem on $[0,1]$, then transfer it to $[a,b]$ by an affine change of variables. Our proof will make use of the **Bernstein polynomials**. For $f \in C([0,1])$, the Bernstein polynomial indexed by n in x is

$$B_n(f)(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k(1-x)^{n-k}, \qquad x\in[0,1].$$

Note that $\deg B_n(f)(x) \leq n$ for all $f \in C([0,1])$.

Lemma 319
For every $x \in [0,1]$,

$$\sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k} = 1, \qquad \sum_{k=0}^{n}\left(\frac{k}{n}-x\right)^2\binom{n}{k}x^k(1-x)^{n-k} = \frac{x(1-x)}{n}.$$

Proof: left as an exercise. ■

Examples: when necessary, we let $X \sim \mathcal{B}(n,x)$, $x \in [0,1]$. Then $E[X] = nx$, $E[X^2] = n^2x^2 + nx(1-x)$, and $E[X^3] = n^3x^3 + 3n^2x^2(1-x) + nx(1-x)(1-2x)$

1. If $f(t) \equiv 1$, then

$$B_n(f)(x) = \sum_{k=0}^{n} 1 \cdot \binom{n}{k} x^k (1-x)^{n-k} = (x + (1-x))^n = 1.$$

2. If $f(t) = t$, then

$$B_n(f)(x) = \sum_{k=0}^{n} \frac{k}{n} \binom{n}{k} x^k (1-x)^{n-k} = \frac{1}{n} E\left[\frac{X}{n}\right] = \frac{1}{n} \cdot nx = x.$$

3. If $f(t) = t^2$, then

$$B_n(f)(x) = E\left[\left(\frac{X}{n}\right)^2\right] = \frac{1}{n^2} E[X^2] = x^2 + \frac{x(1-x)}{n}.$$

For instance,

$$B_4(t^2)(x) = x^2 + \frac{x(1-x)}{4} = \frac{3}{4}x^2 + \frac{1}{4}x.$$

4. If $f(t) = t^3$, then

$$B_n(f)(x) = E\left[\left(\frac{X}{n}\right)^3\right] = x^3 + \frac{3}{n}x^2(1-x) + \frac{1}{n^2}x(1-x)(1-2x).$$

For instance,

$$B_3(t^3)(x) = \frac{1}{9}x + \frac{2}{3}x^2 + \frac{2}{9}x^3.$$

5. If $f(t) = t(1-t) = t - t^2$, then

$$B_n(f)(x) = B_n(t)(x) - B_n(t^2)(x) = x - \left(x^2 + \frac{x(1-x)}{n}\right) = \left(1 - \frac{1}{n}\right) x(1-x).$$

6. If $f(t) = e^t$, then using the binomial generating function $E[s^X] = (1-x+xs)^n$, we have:

$$B_n(f)(x) = E\left[e^{X/n}\right] = E\left[\left(e^{1/n}\right)^X\right] = (1 - x + xe^{1/n})^n.$$

For instance,

$$B_4(e^t)(x) = (1 - x + xe^{1/4})^4.$$

7. If $f(t) = \sin(\pi t)$ and $n = 4$, then using the values

$$\sin(0) = 0, \quad \sin\left(\frac{\pi}{4}\right) = \frac{\sqrt{2}}{2}, \quad \sin\left(\frac{\pi}{2}\right) = 1, \quad \sin\left(\frac{3\pi}{4}\right) = \frac{\sqrt{2}}{2}, \quad \sin(\pi) = 0$$

yields

$$B_4(f)(x) = \sum_{k=0}^{4} \sin(k\pi/4)\binom{4}{k}x^k(1-x)^{4-k} = 4\cdot\frac{\sqrt{2}}{2}\,x(1-x)^3 + 6\,x^2(1-x)^2 + 4\cdot\frac{\sqrt{2}}{2}\,x^3(1-x).$$

Expanding, we get:

$$B_4(\sin \pi t)(x) = 2\sqrt{2}\,x + (6 - 6\sqrt{2})x^2 + (-12 + 8\sqrt{2})x^3 + (6 - 4\sqrt{2})x^4.$$

8. If $f(t) = \sqrt{t}$ and $n = 3$, then using the values

$$f(0) = 0, \qquad f\left(\tfrac{1}{3}\right) = \tfrac{\sqrt{3}}{3}, \qquad f\left(\tfrac{2}{3}\right) = \tfrac{\sqrt{6}}{3}, \qquad f(1) = 1$$

yields

$$B_3(f)(x) = \sum_{k=0}^{3} \sqrt{k/3}\binom{3}{k}x^k(1-x)^{3-k} = 3\cdot\tfrac{\sqrt{3}}{3}\,x(1-x)^2 + 3\cdot\tfrac{\sqrt{6}}{3}\,x^2(1-x) + x^3.$$

Expanding, we get

$$B_3(\sqrt{t})(x) = \sqrt{3}\,x + (-2\sqrt{3} + \sqrt{6})x^2 + (1 + \sqrt{3} - \sqrt{6})x^3.$$

9. If $f(t) = \log(1 + t)$ and $n = 3$, we have

$$f(0) = 0, \qquad f\left(\tfrac{1}{3}\right) = \log\left(\tfrac{4}{3}\right), \qquad f\left(\tfrac{2}{3}\right) = \log\left(\tfrac{5}{3}\right), \qquad f(1) = \log 2,$$

and so

$$B_3(f)(x) = 3\log\left(\tfrac{4}{3}\right)x(1-x)^2 + 3\log\left(\tfrac{5}{3}\right)x^2(1-x) + (\log 2)\,x^3.$$

10. If $f(t) = \left|t - \frac{1}{3}\right|$ and $n = 3$, we have

$$f(0) = \tfrac{1}{3}, \quad f\left(\tfrac{1}{3}\right) = 0, \quad f\left(t\frac{2}{3}\right) = \tfrac{1}{3}, \quad f(1) = \tfrac{2}{3},$$

and so

$$B_3(f)(x) = \tfrac{1}{3}(1-x)^3 + 3\cdot\tfrac{1}{3}\,x^2(1-x) + \tfrac{2}{3}x^3 = \tfrac{1}{3} - x + 2x^2 - \frac{2}{3}x^3. \qquad \square$$

We are now ready to provide the proof of Theorem 318.

Proof:

■

There is an analogous density statement for trigonometric polynomials in spaces of continuous periodic functions. In **Fourier analysis**, summability methods provide approximation schemes that are adapted to **periodic structure** and to norms such as L^2 (see Section 11.2).

Stone-Weierstrass Approximation

The next theorem generalizes Weierstrass approximation by replacing polynomials on an interval with an abstract algebra of functions on a compact space.

Theorem 320 (Stone-Weierstrass)
Let K be a compact Hausdorff space and let $\mathcal{A} \subseteq C(K, \mathbb{R})$ be a sub-algebra that contains the constant functions and separates points of K (that is, for $\mathbf{x} \neq \mathbf{y}$ there exists $g \in \mathcal{A}$ with $g(\mathbf{x}) \neq g(\mathbf{y})$). Then $\mathcal{A}$ is dense in $C(K, \mathbb{R})$ with respect to $\|\cdot\|_\infty$.

In the complex case, if $\mathcal{A} \subseteq C(K, \mathbb{C})$ contains the constants, separates points, and is closed under complex conjugation, then $\mathcal{A}$ is dense in $C(K, \mathbb{C})$.

Proof: left as an exercise. ■

Applications

1. Theorem 318 says that polynomials are dense in $C([a, b])$, so continuous functions on an interval can be uniformly approximated by polynomials. Stone-Weierstrass generalizes this. In practice, this lets us replace a complicated function $f \in C(K)$, K compact, by an **explicit approximant** built from a chosen "feature family" (coordinates, trigonometric functions, etc.).
2. On the torus $K = T$ (or on $[0, 2\pi]$ with endpoints identified), we can apply Stone-Weierstrass to the algebra generated by $\{1, \cos x, \sin x\}$ to show that trigonometric polynomials are dense in $C(K)$. Thus every continuous 2π-periodic function can be uniformly approximated by finite linear combinations of $\sin(nx)$ and $\cos(nx)$.
3. Let $K \subseteq \mathbb{R}^n$ be compact. The algebra of (restrictions to K of) real polynomials in n variables contains constants and separates points, hence is dense in $C(K)$ by Stone-Weierstrass. Therefore every continuous function on a compact set in $\mathbb{R}^n$ can be uniformly approximated by multivariate polynomials on K.

22.7 Hilbert Spaces and Orthogonal Expansions

An **inner product space** is a vector space H over $\mathbb{R}$ (or $\mathbb{C}$) equipped with an inner product $\langle \cdot, \cdot \rangle$ that is linear in the first argument, conjugate symmetric, and positive definite. The induced norm is $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$.

A **Hilbert space** is a complete inner product space. The basic notions are those of

- **orthogonality**: $\mathbf{x} \perp \mathbf{y}$ if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$, and of the
- **orthogonal complement**: for $M \subseteq H$, $M^\perp = \{\mathbf{x} \in H \mid \langle \mathbf{x}, \mathbf{z} \rangle = 0 \text{ for all } \mathbf{z} \in M\}$.

A sequence $(\mathbf{e}_n)$ in H is **orthonormal** if $\langle \mathbf{e}_n, \mathbf{e}_m \rangle = \delta_{n,m}$ (the **Kronecker** δ).

Examples

1. The Euclidean spaces $\mathbb{R}^n$ and $\mathbb{C}^n$, endowed with the standard inner products

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{k=1}^{n} x_k y_k \qquad \text{and} \qquad \langle \mathbf{x}, \mathbf{y} \rangle = \sum_{k=1}^{n} x_k \overline{y_k},$$

   are Hilbert spaces since, more generally, every finite-dimensional inner product space is complete, hence Hilbert (see exercises).

2. The space of square-summable sequences

$$\ell^2(\mathbb{C}) = \left\{ \mathbf{x} = (x_n)_{n\geq 1} \,\middle|\, \sum_{n=1}^{\infty} |x_n|^2 < \infty \right\}, \quad \text{with inner product } \langle \mathbf{x}, \mathbf{y} \rangle = \sum_{n=1}^{\infty} x_n \overline{y_n}$$

   is a Hilbert space (see exercises).

3. If $[a, b] \subseteq \mathbb{R}$, then the space of square-integrable functions

$$L^2([a, b]) = \left\{ f : [a, b] \to \mathbb{C} \,\middle|\, \int_a^b |f(t)|^2 \, dt < \infty \right\}$$

   (with functions identified up to equality almost everywhere) is a Hilbert space with inner product

$$\langle f, g \rangle = \int_a^b f(t)\overline{g(t)} \, dt.$$

   In particular, the space $L^2([-\pi, \pi])$ of 2π-periodic functions with the same inner product is a Hilbert space, and it is the natural setting for Fourier series and orthogonal expansions.

4. The space $M_{m\times n}(\mathbb{R})$ (or $M_{m\times n}(\mathbb{C})$) becomes a Hilbert space when endowed with the Frobenius inner product

$$\langle A, B \rangle = \operatorname{tr}(AB^{\mathsf{T}}) \quad \text{(real case)}, \qquad \langle A, B \rangle = \operatorname{tr}(AB^{*}) \quad \text{(complex case)}$$

   (see exercises).

5. If H is a Hilbert space and $M \subseteq_C H$, then M (with the restricted inner product) is again a Hilbert space.

6. The space $C([0, 1])$ with the sup norm is complete, but it is not induced by an inner product, so it is not a Hilbert space.

7. The spaces $L^p([a, b])$ for $p \neq 2$ are Banach spaces, but in general they are not Hilbert spaces.

8. In $\mathbb{R}^n$ with the standard inner product $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{k=1}^{n} x_k y_k$, we have

$$(1,0,0) \perp (0,1,0) \qquad \text{since} \qquad \langle (1,0,0), (0,1,0) \rangle = 0.$$

More generally, the standard basis vectors $\mathbf{e}_i$ and $\mathbf{e}_j$ satisfy $\mathbf{e}_i \perp \mathbf{e}_j$ for $i \neq j$.

9. Let $M = \text{span}\{(1,1)\} \subseteq \mathbb{R}^2$ with the usual inner product. Then

$$M^\perp = \{(x,y) \in \mathbb{R}^2 \mid x + y = 0\} = \text{span}\{(1,-1)\}.$$

10. Let $M = \text{span}\{(1,0,0),(0,1,0)\} \subseteq \mathbb{R}^3$ with the usual inner product. Then

$$M^\perp = \text{span}\{(0,0,1)\}.$$

11. In $\ell^2(\mathbb{C})$ with $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{n \geq 1} x_n \overline{y_n}$, the sequence (of sequences)

$$\mathbf{e}_1 = (1,0,0,\dots), \mathbf{e}_2 = (0,1,0,\dots), \dots, \mathbf{e}_n = (0,\dots,0,1,0,\dots), \dots$$

is orthonormal since $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{i,j}$.

12. In $L^2([-\pi,\pi])$ with $\langle f, g \rangle = \int_{-\pi}^{\pi} f(t)\overline{g(t)}\, dt$, the complex exponentials

$$e_n(t) = \frac{1}{\sqrt{2\pi}} e^{int}, \qquad n \in \mathbb{Z},$$

form an orthonormal sequence:

$$\langle e_n, e_m \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(n-m)t}\, dt = \delta_{n,m}.$$

13. Let $M = \text{span}\{\mathbf{1}\} \subseteq L^2([0,1])$, where $\mathbf{1}(t) = 1$. Then

$$M^\perp = \left\{ f \in L^2([0,1]) \;\middle|\; \int_0^1 f(t)\, dt = 0 \right\}.$$

Indeed, $f \in M^\perp$ if and only if $\langle f, \mathbf{1} \rangle = \int_0^1 f(t)\, dt = 0$.

14. In $M_{n\times n}(\mathbb{R})$ with $\langle A, B \rangle = \text{tr}(AB^\mathsf{T}) = \sum_{i,j} a_{ij} b_{ij}$, let M be the subspace of diagonal matrices. Then

$$M^\perp = \{A \in M_{n\times n}(\mathbb{R}) \mid a_{1,1} = \cdots = a_{n,n} = 0\},$$

the subspace of matrices with zero diagonal. □

Theorem 321 (Orthogonal Projection)
Let H be a Hilbert space and $M \subseteq H$ a closed subspace. For every $\mathbf{x} \in H$ there exists a unique $\mathbf{m}_\mathbf{x} \in M$ such that

$$\|\mathbf{x} - \mathbf{m}_\mathbf{x}\| = \inf_{\mathbf{y} \in M} \|\mathbf{x} - \mathbf{y}\|, \quad \textit{with } \mathbf{x} - \mathbf{m}_\mathbf{x} \in M^\perp.$$

Proof:

The importance of orthonormal sequences is indicated by the next result.

Theorem 322 (Bessel's Inequality)
If $(\mathbf{e}_n)$ is orthonormal in H, then for every $\mathbf{x} \in H$,

$$\sum_{n=1}^{\infty} |\langle \mathbf{x}, \mathbf{e}_n \rangle|^2 \leq \|\mathbf{x}\|^2.$$

Proof: fix N and let $M_N = \text{span}\{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$. The orthogonal projection of $\mathbf{x}$ onto M_N is

$$P_N \mathbf{x} = \sum_{n=1}^{N} \langle \mathbf{x}, \mathbf{e}_n \rangle \mathbf{e}_n.$$

By orthonormality,

$$\|P_N \mathbf{x}\|^2 = \left\langle \sum_{n=1}^{N} \langle \mathbf{x}, \mathbf{e}_n \rangle \mathbf{e}_n, \sum_{m=1}^{N} \langle \mathbf{x}, \mathbf{e}_m \rangle \mathbf{e}_m \right\rangle = \sum_{n=1}^{N} |\langle \mathbf{x}, \mathbf{e}_n \rangle|^2.$$

Since $\mathbf{x} - P_N \mathbf{x} \perp M_N$, the Pythagorean theorem gives

$$\|\mathbf{x}\|^2 = \|P_N \mathbf{x}\|^2 + \|\mathbf{x} - P_N \mathbf{x}\|^2 \geq \|P_N \mathbf{x}\|^2 = \sum_{n=1}^{N} |\langle \mathbf{x}, \mathbf{e}_n \rangle|^2.$$

Letting $N \to \infty$ yields Bessel's inequality. ■

If $(\mathbf{e}_n)$ is an **orthonormal basis** of H (that is, the closed linear span of $\{\mathbf{e}_n\}$ is H), then $P_N \mathbf{x} \to \mathbf{x}$ in norm, and Bessel's inequality becomes **Parseval's identity**

$$\|\mathbf{x}\|^2 = \sum_{n=1}^{\infty} |\langle \mathbf{x}, \mathbf{e}_n \rangle|^2, \qquad \mathbf{x} = \sum_{n=1}^{\infty} \langle \mathbf{x}, \mathbf{e}_n \rangle \mathbf{e}_n \quad (\text{convergence in } \|\cdot\|).$$

How the Hilbert Space Viewpoint Frames Fourier Series

In the Hilbert space $H = L^2([-\pi, \pi])$ with inner product

$$\langle f, g \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \overline{g(t)} \, dt,$$

Fourier coefficients are inner products against an orthonormal trigonometric system, and partial sums are orthogonal projections onto finite-dimensional subspaces. Quadratic mean convergence corresponds to convergence in the L^2 norm, and Parseval's identity is exactly the Hilbert space norm identity for an orthonormal basis (see Section 11.2).

22.8 The Inverse Function Theorem

The next result is a central cog of multivariate calculus.

Theorem 323 (Inverse Function Theorem)
Let $U \subseteq_O \mathbb{R}^n$ and let $F : U \to \mathbb{R}^n$ be continuously differentiable. If $\mathbf{a} \in U$ and $DF(\mathbf{a})$ is invertible, then there exist open neighborhoods V of $\mathbf{a}$ and W of $F(\mathbf{a})$ such that $F : V \to W$ is a bijection. Moreover, the inverse $F^{-1} : W \to V$ is continuously differentiable and

$$D(F^{-1})(F(\mathbf{x})) = \big(DF(\mathbf{x})\big)^{-1} \quad \text{for all } \mathbf{x} \in V.$$

Proof:

If $DF(\mathbf{a})$ is invertible, then $DF(\mathbf{x})$ remains invertible in a neighborhood of **a**, and F is locally open and locally one-to-one. This is the local ingredient behind change of variables formulas (see Section 21.7).

22.9 Implicit Function Theorem

Let $f, g : U \subseteq_O \mathbb{R}^n \to \mathbb{R}$ be defined on a punctured neighbourhood of **a**. We write

$$f(\mathbf{x}) = o\big(g(\mathbf{x})\big) \quad \text{as } \mathbf{x} \to \mathbf{a}$$

if $g(\mathbf{x}) \neq 0$ near **a** and

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{f(\mathbf{x})}{g(\mathbf{x})} = 0.$$

For $F : \mathbb{R}^n \to \mathbb{R}^m$, **differentiability at a** means that $\exists$ a linear map $DF(\mathbf{a})$ such that

$$F(\mathbf{a}+\mathbf{h}) = F(\mathbf{a}) + DF(\mathbf{a})\mathbf{h} + o(\|\mathbf{h}\|) \quad \text{as } \mathbf{h} \to \mathbf{0}.$$

Many local theorems depend on whether suitable blocks of $DF(\mathbf{a})$ are invertible.

Theorem 324 (Implicit Function Theorem)
*Let $U \subseteq \mathbb{R}^n \times \mathbb{R}^m$ be open and let $G : U \to \mathbb{R}^m$ be continuously differentiable. Suppose $G(\mathbf{a}, \mathbf{b}) = 0$ and that the partial derivative $D_{\mathbf{y}}G(\mathbf{a}, \mathbf{b}) : \mathbb{R}^m \to \mathbb{R}^m$ is invertible. Then there exist neighbourhoods V of **a** and W of **b** and a unique continuously differentiable map $\varphi : V \to W$ such that*

$$G(\mathbf{x}, \varphi(\mathbf{x})) \equiv 0 \quad \text{for all } \mathbf{x} \in V.$$

Moreover, the Jacobian matrix of φ is given by

$$D\varphi(\mathbf{x}) = -\big(D_{\mathbf{y}}G(\mathbf{x}, \varphi(\mathbf{x}))\big)^{-1} D_{\mathbf{x}}G(\mathbf{x}, \varphi(\mathbf{x})).$$

Proof:

■

The implicit function theorem is often used to provide local solutions to systems of equations. Let $G : \mathbb{R}^{n+m} \to \mathbb{R}^m$ be a C^1 map, and write points as $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^n \times \mathbb{R}^m$. If $G(\mathbf{a}, \mathbf{b}) = 0$ and $D_{\mathbf{y}}G(\mathbf{a}, \mathbf{b})$ is invertible, then $\exists$ open neighbourhoods $U_0 \subseteq \mathbb{R}^n$ of **a** and $V_0 \subseteq \mathbb{R}^m$ of **b**, and a unique C^1 map $\varphi : U_0 \to V_0$ such that for all $(\mathbf{x}, \mathbf{y}) \in U_0 \times V_0$,

$$G(\mathbf{x}, \mathbf{y}) \equiv 0 \quad \iff \quad \mathbf{y} = \varphi(\mathbf{x}).$$

Moreover, $\varphi(\mathbf{a}) = \mathbf{b}$ and the derivative of φ at **a** is

$$D\varphi(\mathbf{a}) = -(D_{\mathbf{y}}G(\mathbf{a}, \mathbf{b}))^{-1} D_{\mathbf{x}}G(\mathbf{a}, \mathbf{b}).$$

The implicit function theorem can also be applied to **initial value problems** (IVP; see Chapter 23).

Theorem 325 (Picard-Lindelöf)
Let $f : [t_0 - h_0, t_0 + h_0] \times \mathbb{R}^m \to \mathbb{R}^m$ be continuous and assume there exists $L \geq 0$ such that

$$\|f(t, \mathbf{u}) - f(t, \mathbf{v})\| \leq L\|\mathbf{u} - \mathbf{v}\|$$

for all $t \in [t_0 - h_0, t_0 + h_0]$ and all $\mathbf{u}, \mathbf{v}$ in a neighbourhood of $\mathbf{u}_0$. Then there exists $h \in (0, h_0]$ such that the initial value problem

$$\mathbf{u}'(t) = f(t, \mathbf{u}(t)), \qquad \mathbf{u}(t_0) = \mathbf{u}_0,$$

has a unique solution $\mathbf{u} \in C^1([t_0 - h, t_0 + h]; \mathbb{R}^m)$. Moreover, the Picard iterates

$$\mathbf{u}_{n+1}(t) = \mathbf{u}_0 + \int_{t_0}^{t} f(s, \mathbf{u}_n(s))\, ds$$

converge uniformly to $\mathbf{u}$ on $[t_0 - h, t_0 + h]$ with $n \to \infty$.

Proof:

■

Exercises

1. Prepare a 2-page summary of this chapter, with important definitions and results.
2. Let (X, d) be a complete metric space and let $T : X \to X$ be a contraction with factor $q \in (0, 1)$. Fix $\mathbf{x}_0 \in X$ and define $\mathbf{x}_{n+1} = T(\mathbf{x}_n)$.
   a) Prove that $d(\mathbf{x}_n, \mathbf{x}^*) \leq q^n d(\mathbf{x}_0, \mathbf{x}^*)$ for all $n \geq 0$, where $\mathbf{x}^*$ is the unique fixed point.
   b) Prove that $d(\mathbf{x}_{n+1}, \mathbf{x}_n) \leq q^n d(\mathbf{x}_1, \mathbf{x}_0)$ for all $n \geq 0$.

3. Does every contraction $T : X \to X$ on a metric space (X, d) have a fixed point?
4. Consider $T(\mathbf{x}) = \cos x$ on $I = [0, 1]$ with the usual metric.
   a) Show that $T(I) \subseteq I$ and that T is a contraction on I with factor $q = \sin(1)$.
   b) Using the a posteriori bound, find an explicit N (in terms of q) such that if $|x_{N+1} - x_N| < 10^{-6}$ then $|x_N - \mathbf{x}^*| < 10^{-4}$.
5. Apply Picard iteration to the IVP $u'(t) = u(t)$, $u(0) = 1$ on a symmetric interval $[-h, h]$.
   a) Write the Picard operator T on $C([-h, h])$.
   b) Starting from the constant function $\mathbf{u}_0(t) \equiv 1$, compute the first four iterates $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4$ explicitly.
   c) Guess the limit of the iteration and justify your guess.
6. Let X, Y be normed spaces and let $T : X \to Y$ be linear.
   a) Prove that T is bounded if and only if $\sup_{\|x\| \leq 1} \|T(\mathbf{x})\| < \infty$.
   b) Prove that if T is bounded, then $\|T(\mathbf{x}) - T(\mathbf{y})\| \leq \|T\|_{\text{op}} \|\mathbf{x} - \mathbf{y}\|$ for all $x, y \in X$.
7. Let $D = \text{diag}(d_1, \ldots, d_n)$ act on $\mathbb{R}^n$.
   a) Compute $\|D\|_{\text{op}}$ when $\mathbb{R}^n$ is equipped with $\|\cdot\|_\infty$.
   b) Compute $\|D\|_{\text{op}}$ when $\mathbb{R}^n$ is equipped with $\|\cdot\|_1$.
   c) Compute $\|D\|_{\text{op}}$ when $\mathbb{R}^n$ is equipped with $\|\cdot\|_2$.
8. Let $\mathbf{u} \in \mathbb{R}^m$ and $\mathbf{v} \in \mathbb{R}^n$, and define $T : \mathbb{R}^n \to \mathbb{R}^m$ by

$$T(\mathbf{x}) = (\mathbf{v} \cdot \mathbf{x})\, \mathbf{u}$$

   (where $\mathbf{v} \cdot \mathbf{x}$ is the usual Euclidean inner product). Is T a linear operator? If so, compute $\|T\|_{\text{op}}$ for the Euclidean norms on $\mathbb{R}^n$ and $\mathbb{R}^m$.
9. Let $A = \begin{pmatrix} 1 & -2 \\ 3 & 4 \end{pmatrix}$ act on $\mathbb{R}^2$. Compute $\|A\|_{\text{op}}$ when both domain and codomain are equipped with $\|\cdot\|_\infty$.
10. Let A be as in the previous exercise. Compute $\|A\|_{\text{op}}$ when both domain and codomain are equipped with $\|\cdot\|_1$.
11. Let A be as in the previous exercise. Compute $\|A\|_{\text{op}}$ when both domain and codomain are equipped with $\|\cdot\|_2$.
12. Define $S : \mathbb{R}^n \to \mathbb{R}^n$ by

$$S(x_1, \ldots, x_n) = (0, x_1, \ldots, x_{n-1}).$$

   Compute $\|S\|_{\text{op}}$ with respect to $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|_\infty$.

13. Let $P:\mathbb{R}^n\to\mathbb{R}^n$ be the coordinate projection

$$P(\mathbf{x}_1,\ldots,\mathbf{x}_n)=(\mathbf{x}_1,\ldots,\mathbf{x}_k,0,\ldots,0).$$

Compute $\|P\|_{\text{op}}$ with respect to $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|_\infty$.

14. Let $X=C([0,1])$ with $\|\cdot\|_\infty$, and define $T:X\to X$ by

$$(Tf)(x)=\int_0^x f(t)\,dt.$$

Compute $\|T\|_{\text{op}}$.

15. Let $X=C([0,1])$ with $\|\cdot\|_\infty$, and fix $\mathbf{x}_0\in[0,1]$. Define $\delta_{\mathbf{x}_0}:X\to\mathbb{R}$ by $\delta_{\mathbf{x}_0}(f)=f(\mathbf{x}_0)$. Compute $\|\delta_{\mathbf{x}_0}\|_{\text{op}}$.

16. Let $X=\mathbb{R}^n$ with $\|\cdot\|_\infty$, and define $\varphi:X\to\mathbb{R}$ by

$$\varphi(x_1,\ldots,x_n)=x_1+\cdots+x_n.$$

Compute $\|\varphi\|_{\text{op}}$.

17. Let $T:\mathbb{R}^2\to\mathbb{R}^2$ be defined by $T(\mathbf{x})=A\mathbf{x}$ with

$$A=\begin{pmatrix}2&-1\\0&1\end{pmatrix}.$$

Compute $\|T\|_{\text{op}}$ when $\mathbb{R}^2$ is equipped with the norm $\|x\|_\infty=\max\{|\mathbf{x}_1|,|\mathbf{x}_2|\}$.

18. Let X be a real normed space and let $\mathbf{x}\in X$ with $\mathbf{x}\neq\mathbf{0}$.

    a) Use Hahn-Banach to prove that there exists $\varphi\in X^*$ with $\|\varphi\|=1$ and $\varphi(\mathbf{x})=\|\mathbf{x}\|$.

    b) Is the functional φ unique for every $\mathbf{x}\neq\mathbf{0}$?

19. Show that the functions defined in the examples on 556 are all sub-linear.

20. Let X be a real normed space and $Y\subseteq X$ a subspace. Is it the case that every $\varphi_0\in Y^*$ admits an extension $\varphi\in X^*$ with $\|\varphi\|=\|\varphi_0\|$.

21. **A proof of the Hahn-Banach theorem.** Throughout, let $(X,\|\cdot\|)$ be a real normed vector space, assume that X is separable, let $Y<X$ be a subspace, and let $\varphi_0:Y\to\mathbb{R}$ be a bounded linear functional. Set

$$p(x)=\|\varphi_0\|\,\|\mathbf{x}\|\qquad(\mathbf{x}\in X).$$

    a) Show that $p:X\to\mathbb{R}$ is sub-linear and continuous.

    b) Show that $\varphi_0(\mathbf{y})\leq p(\mathbf{y})$ for all $\mathbf{y}\in Y$.

c) Let $(\mathbf{x}_n)_{n\in\mathbb{N}}$ be a countable dense subset of X. For each $n \in \mathbb{N}$, set

$$Z_n = Y + \operatorname{span}\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}.$$

Show that Z_n is a subspace of X, that $Z_n \subseteq Z_{n+1}$, and that

$$Z_\infty := \bigcup_{n=1}^{\infty} Z_n$$

is dense in X.

d) Using Lemma 306A (one-dimensional extension), prove by induction that there exists a sequence of linear maps

$$\varphi_n : Z_n \to \mathbb{R} \qquad (n \in \mathbb{N})$$

such that

$$\varphi_n|_Y = \varphi_0, \qquad \varphi_{n+1}|_{Z_n} = \varphi_n, \qquad \text{and} \qquad \varphi_n(\mathbf{z}) \le p(\mathbf{z}) \text{ for all } \mathbf{z} \in Z_n.$$

e) Define $\varphi_\infty : Z_\infty \to \mathbb{R}$ by $\varphi_\infty(\mathbf{z}) = \varphi_n(\mathbf{z})$ whenever $\mathbf{z} \in Z_n$. Show that φ_∞ is well-defined, linear, extends φ_0, and satisfies

$$\varphi_\infty(\mathbf{z}) \le p(\mathbf{z}) \qquad \text{for all } \mathbf{z} \in Z_\infty.$$

f) Show that for all $\mathbf{z} \in Z_\infty$,

$$|\varphi_\infty(\mathbf{z})| \le \|\varphi_0\| \, \|\mathbf{z}\|.$$

(Hint: use $\varphi_\infty(\mathbf{z}) \le p(\mathbf{z})$ and $-\varphi_\infty(\mathbf{z}) = \varphi_\infty(-\mathbf{z}) \le p(-\mathbf{z}) = p(\mathbf{z})$.)

g) Deduce that φ_∞ is Lipschitz on Z_∞:

$$|\varphi_\infty(\mathbf{z}) - \varphi_\infty(\mathbf{w})| \le \|\varphi_0\| \, \|\mathbf{z} - \mathbf{w}\| \qquad \text{for all } \mathbf{z}, \mathbf{w} \in Z_\infty.$$

h) For each $\mathbf{x} \in X$, choose a sequence $(\mathbf{z}_k) \subseteq Z_\infty$ such that $\mathbf{z}_k \to x$. Prove that $(\varphi_\infty(\mathbf{z}_k))$ is a Cauchy sequence in $\mathbb{R}$ and define

$$\varphi(\mathbf{x}) = \lim_{k\to\infty} \varphi_\infty(\mathbf{z}_k).$$

i) Show that $\varphi : X \to \mathbb{R}$ is well-defined (the limit does not depend on the approximating sequence), linear, and satisfies $\varphi|_Y = \varphi_0$.

j) Show that $\varphi(\mathbf{x}) \le p(\mathbf{x})$ for all $\mathbf{x} \in X$, and conclude that

$$|\varphi(\mathbf{x})| \le \|\varphi_0\| \, \|\mathbf{x}\| \qquad \text{for all } \mathbf{x} \in X.$$

In particular, $\|\varphi\| \le \|\varphi_0\|$.

k) Show that

$$\|\varphi_0\| \leq \|\varphi\|.$$

(Hint: the supremum defining $\|\varphi_0\|$ is taken over $\{\mathbf{y} \in Y \mid \|y\| \leq 1\}$, which is contained in $\{\mathbf{x} \in X \mid \|x\| \leq 1\}$, and $\varphi = \varphi_0$ on Y.)

l) Conclude that $\|\varphi\| = \|\varphi_0\|$, completing the proof of Theorem 308.

22. Let (X, d) be a complete metric space.

  a) Prove that $\mathbb{Q}$ is meagre in $\mathbb{R}$.

  b) Is countable intersection of dense open subsets of X open?

23. Fill in the details for the examples on pp. 22.4-22.4.

24. Is the uniform boundedness principle valid if X is a normed space that is not complete?

25. This series of exercises relate to the **Cantor set**.

  a) Let $C_0 = [0, 1]$. Given C_n, obtain C_{n+1} by removing from each closed interval component of C_n its open middle third. Define the **Cantor set** by

$$C := \bigcap_{n=0}^{\infty} C_n.$$

    i. Write C_1 and C_2 explicitly as finite unions of closed intervals.

    ii. Prove that each C_n is a finite union of 2^n closed intervals, each of length 3^{-n}.

    iii. Prove that $C_{n+1} \subseteq C_n$ for all n.

  b) i. Prove that each C_n is closed in $\mathbb{R}$.

    ii. Deduce that C is closed.

    iii. Deduce that C is compact.

  c) i. Prove that C contains no non-degenerate interval.

    ii. Deduce that int$(C) = \varnothing$.

  d) Prove that C is nowhere dense in $\mathbb{R}$.

  e) Let $U = [0, 1] \setminus C$.

    i. Prove that U is open in $[0, 1]$ (with the subspace topology).

    ii. Show that U is a countable union of pairwise disjoint open intervals.

    iii. Prove that U is dense in $[0, 1]$.

  f) i. Show every $x \in [0, 1]$ has a base-3 expansion $x = \sum_{k=1}^{\infty} a_k 3^{-k}$ with $a_k \in \{0, 1, 2\}$, and that some numbers have two such expansions.

    ii. Prove that $x \in C$ if and only if x has a base-3 expansion using only the digits 0 and 2.

    iii. Prove that if $x \in C$, then there exists a base-3 expansion of x with digits only in $\{0, 2\}$ (even if x also has another expansion).

g) Using the ternary characterization above, prove that C is uncountable. (Hint: map sequences $(\varepsilon_k)_{k\geq 1} \in \{0,1\}^{\mathbb{N}}$ to points of C by sending ε_k to the ternary digit $2\varepsilon_k$.)

h) i. Fix $x \in C$ and $r > 0$. Prove that $(x-r, x+r)$ contains a point of C distinct from x.

ii. Deduce that C has no isolated points.

iii. Conclude that C is **perfect**, meaning C is closed and every point of C is a limit point of C.

i) Prove that C is totally disconnected: the only connected subsets of C are singletons. (Hint: use the fact that points of C can be separated at some finite stage C_n by disjoint closed intervals.)

j) Let

$$C_L := C \cap \left[0, \frac{1}{3}\right], \qquad C_R := C \cap \left[\frac{2}{3}, 1\right].$$

i. Prove that $C = C_L \cup C_R$ and $C_L \cap C_R = \varnothing$.

ii. Prove that the map $x \mapsto 3x$ is a bijection from C_L onto C.

iii. Prove that the map $x \mapsto 3x - 2$ is a bijection from C_R onto C.

k) Assume the definition and basic properties of Lebesgue measure m on $\mathbb{R}$ are known.

i. Compute the total length removed at stage n.

ii. Show that $m(C) = 0$ by estimating $m(C_n)$ and passing to the limit.

l) Prove that C is meagre in $\mathbb{R}$. (Hint: use exercise d).)

26. Let K be a compact metric space and let $\mathcal{F} \subseteq C(K)$.

a) If $\mathcal{F}$ is relatively compact in $(C(K), \|\cdot\|_\infty)$, is $\mathcal{F}$ pointwise bounded?

b) If $\mathcal{F}$ is pointwise bounded, is $\mathcal{F}$ relatively compact?

27. Consider the family $\mathcal{F} = \{f_n\}_{n\geq 1} \subseteq C([0,1])$ given by $f_n(x) = x^n$. Is $\mathcal{F}$ relatively compact in $(C([0,1]), \|\cdot\|_\infty)$?

28. If X is a normed space and $A \subseteq X$ is closed and bounded, is A necessarily compact?

29. Show that in a complete metric space X, any pre-compact subset A is also relatively compact, and *vice-versa*.

30. Show that every compact metric space K contains a countable dense subset.

31. In the proof of the Arzelà-Ascoli theorem, show that the last line of the proof holds: that relative compactness of $\mathcal{F}$ in $C(K)$ is equivalent to every sequence in $\mathcal{F}$ having a uniformly convergent sub-sequence.

32. For $f(t) = e^{-t}$, compute the Bernstein polynomials $B_n(f)(x)$ for $x \in [0,1]$. Does any pattern emerge? Is there any function f such that $B_n(f)(x) = f(x)$ for all $x \in [0,1]$?

33. Prove the identities of Lemma 319 (it might be helpful to lean on the knowledge of the binomial distribution from probability).

34. Are the Bernstein polynomials of degree n "linear" over the space $C([0,1])$? That is, for $n \geq 0$, $a, b \in \mathbb{R}$ and $f, g \in C([0,1])$, is it the case that

$$B_n(af + bg) = aB_n(f) + bB_n(g)?$$

35. Prove the Stone-Weierstrass theorem using the following steps.

    a) Let K be a compact Hausdorff space and let $C(K)$ denote $C(K, \mathbb{R})$ with the norm $\|f\|_\infty = \sup_{\mathbf{x}\in K} |f(\mathbf{x})|$. Let $\mathcal{A} \subseteq C(K)$ be a sub-algebra, and let $\overline{\mathcal{A}}$ be its closure in $\|\cdot\|_\infty$. Prove that $\overline{\mathcal{A}}$ is again a sub-algebra of $C(K)$ and that it contains the constant functions if $\mathcal{A}$ does.

    b) Show that if $f \in C(K)$ and p is a real polynomial, then $p \circ f \in C(K)$. Show moreover that if $f \in \mathcal{A}$, then $p \circ f \in \mathcal{A}$.

    c) Fix $M > 0$. Using the Weierstrass approximation theorem on $[-M, M]$, show that there exists a sequence of polynomials (p_m) such that

    $$\sup_{|t|\leq M} \big|p_m(t) - |t|\big| \longrightarrow 0.$$

    d) Let $f \in \overline{\mathcal{A}}$. Prove that $|f| \in \overline{\mathcal{A}}$. (Hint: first reduce to the case where $\|f\|_\infty \leq M$, apply part c) to f, and use part b) inside $\overline{\mathcal{A}}$).

    e) For $f, g \in C(K)$ show that

    $$\max(f,g) = \frac{f + g + |f-g|}{2}, \qquad \min(f,g) = \frac{f + g - |f-g|}{2}.$$

    Deduce from part d) that if $f, g \in \overline{\mathcal{A}}$, then $\max(f,g), \min(f,g) \in \overline{\mathcal{A}}$. Conclude that $\overline{\mathcal{A}}$ is a sub-lattice of $C(K)$.[1]

    f) Assume that $\mathcal{A}$ separates points of K.[2] Show that for any distinct $\mathbf{x}, \mathbf{y} \in K$ there exists $h \in \mathcal{A}$ such that $h(\mathbf{x}) = 0$ and $h(\mathbf{y}) = 1$. (Hint: start from $g \in \mathcal{A}$ with $g(\mathbf{x}) \neq g(\mathbf{y})$ and apply an affine change of variables).

    g) Assume that $\mathcal{A}$ contains the constant functions and separates points. Let $\mathbf{x} \neq \mathbf{y}$ and let $a, b \in \mathbb{R}$ be arbitrary. Show that there exists $u \in \mathcal{A}$ with $u(\mathbf{x}) = a$ and $u(\mathbf{y}) = b$. *Hint:* use part f) and linear combinations.

    h) Assume that $\mathcal{A}$ contains constants and separates points. Let $F \subseteq K$ be closed and let $\mathbf{x} \in K \setminus F$. Show that there exists $u \in \overline{\mathcal{A}}$ such that $u(\mathbf{x}) = 0$ and $u(\mathbf{y}) \geq 1$ for all $\mathbf{y} \in F$. (Hint: for each $\mathbf{y} \in F$, use part f) to find $u_\mathbf{y} \in \mathcal{A}$ with $u_\mathbf{y}(\mathbf{x}) = 0$ and $u_\mathbf{y}(\mathbf{y}) = 1$. Use continuity to get an open neighbourhood $V_\mathbf{y}$ of $\mathbf{y}$ on which $u_\mathbf{y} > 1/2$. Use compactness of F to extract finitely many $\mathbf{y}_1, \ldots, \mathbf{y}_N$ so that $F \subseteq \bigcup_i V_{\mathbf{y}_i}$, then combine them using max (part e) and rescale.)

[1] A **sublattice** of $C(K)$ is a subset $\mathcal{L} \subseteq C(K)$ such that for all $f, g \in \mathcal{L}$, the pointwise functions $\max(f,g)$ and $\min(f,g)$ also belong to $\mathcal{L}$.

[2] That is, for any $\mathbf{x} \neq \mathbf{y} \in K$, $\exists g \in \mathcal{A}$ such that $g(\mathbf{x}) \neq g(mmy)$.

i) Let $f \in C(K)$ and let $\varepsilon > 0$. Fix $\mathbf{x} \in K$. Show that there exists $g_{\mathbf{x}} \in \overline{\mathcal{A}}$ such that

$$g_{\mathbf{x}}(\mathbf{x}) = f(\mathbf{x}) \qquad \text{and} \qquad g_{\mathbf{x}} \leq f + \varepsilon \quad \text{on } K.$$

(Hint: for each $\mathbf{y} \in K$ use part g) to build $u_{\mathbf{x},\mathbf{y}} \in \mathcal{A}$ with $u_{\mathbf{x},\mathbf{y}}(\mathbf{x}) = f(\mathbf{x})$ and $u_{\mathbf{x},\mathbf{y}}(\mathbf{y}) = f(\mathbf{y})$. For each fixed $\mathbf{y}$, continuity gives an open neighbourhood $V_{\mathbf{y}}$ on which $u_{\mathbf{x},\mathbf{y}} \leq f + \varepsilon$. Use compactness to extract finitely many $V_{\mathbf{y}_i}$ covering K, and set $g_{\mathbf{x}} = \min_i \{u_{\mathbf{x},\mathbf{y}_i}\}$ (part e).)

j) With the assumptions of part i), show that there exist finitely many points $\mathbf{x}_1, \ldots, \mathbf{x}_m \in K$ such that

$$g := \max\{g_{\mathbf{x}_1}, \ldots, g_{\mathbf{x}_m}\}$$

satisfies

$$f \leq g \leq f + \varepsilon \quad \text{on } K.$$

(Hint: for each x, the inequality $g_{\mathbf{x}}(\mathbf{x}) = f(\mathbf{x})$ and continuity imply that $g_{\mathbf{x}} > f - \varepsilon$ on some neighbourhood of x. Use compactness of K to choose finitely many such neighbourhoods.)

k) Assume that $\mathcal{A} \subseteq C(K)$ is a sub-algebra that contains the constants and separates points. Use part j) to show that for every $f \in C(K)$ and every $\varepsilon > 0$ there exists $g \in \overline{\mathcal{A}}$ with $\|f - g\|_\infty < \varepsilon$. Conclude that $\overline{\mathcal{A}} = C(K)$, that is, $\mathcal{A}$ is dense in $C(K)$.

l) Let $\mathcal{A} \subseteq C(K, \mathbb{C})$ be a sub-algebra containing the constants and separating points. Assume in addition that $f \in \mathcal{A}$ implies $\overline{f} \in \mathcal{A}$. Show that the set $\text{Re}(\mathcal{A}) = \{\text{Re}(f) \mid f \in \mathcal{A}\}$ is a real sub-algebra of $C(K, \mathbb{R})$ containing constants and separating points, and deduce that $\mathcal{A}$ is dense in $C(K, \mathbb{C})$.

36. Show that every finite-dimensional inner product space is complete.

37. This series of questions relates to the space $L^2([a, b])$ of square integrable functions over $[a, b]$ with inner product

$$\langle f, g \rangle = \int_a^b f(t)\overline{g(t)}\, dt.$$

Some knowledge of measure theory (see Chapter 21) is assumed.

a) $L^2([a, b])$ is the set of (equivalence classes of) measurable functions $f : [a, b] \to \mathbb{C}$ such that

$$\int_a^b |f(t)|^2\, dt < \infty.$$

Explain why, if $f = f'$ and $g = g'$ almost everywhere, then

$$\int_a^b f(t)\overline{g(t)}\, dt = \int_a^b f'(t)\overline{g'(t)}\, dt.$$

Conclude that $\langle f, g \rangle$ is well-defined on $L^2([a, b])$.

b) Verify directly from properties of the integral that for all scalars $\alpha, \beta \in \mathbb{C}$ and all $f_1, f_2, g \in L^2([a,b])$,

$$\langle \alpha f_1 + \beta f_2,\, g\rangle = \alpha\langle f_1, g\rangle + \beta\langle f_2, g\rangle, \qquad \langle f, g\rangle = \overline{\langle g, f\rangle}.$$

c) Show that for every $f \in L^2([a,b])$,

$$\langle f, f\rangle = \int_a^b |f(t)|^2\, dt \geq 0.$$

Then prove that

$$\langle f, f\rangle = 0 \quad \Longrightarrow \quad f = 0 \text{ almost everywhere}.$$

Conclude that $\langle \cdot, \cdot\rangle$ is an inner product on $L^2([a,b])$.

d) Prove the **Cauchy-Schwarz inequality**

$$|\langle f, g\rangle| \leq \|f\|_2\, \|g\|_2, \qquad \|f\|_2 := \sqrt{\langle f, f\rangle}.$$

Deduce that $\|\cdot\|_2$ is a norm on $L^2([a,b])$ and that

$$\|f - g\|_2^2 = \langle f - g, f - g\rangle$$

defines the metric associated to the inner product.

e) Let (f_n) be a Cauchy sequence in $(L^2([a,b]), \|\cdot\|_2)$. Choose a subsequence (f_{n_k}) such that

$$\|f_{n_{k+1}} - f_{n_k}\|_2 \leq 2^{-k} \quad \text{for all } k \geq 1.$$

Define

$$g_k := f_{n_{k+1}} - f_{n_k}, \qquad G_m(t) := \sum_{k=1}^{m} |g_k(t)|.$$

Show that $(G_m(t))_{m\geq 1}$ is increasing for each t, and set

$$G(t) := \sup_{m\geq 1} G_m(t) = \sum_{k=1}^{\infty} |g_k(t)|.$$

f) Use the Cauchy-Schwarz inequality to prove

$$\int_a^b |g_k(t)|\, dt \leq (b-a)^{1/2}\, \|g_k\|_2.$$

Deduce that

$$\sum_{k=1}^{\infty} \int_a^b |g_k(t)|\, dt < \infty,$$

and then use the monotone convergence theorem to obtain

$$\int_a^b G(t)\, dt = \sum_{k=1}^{\infty} \int_a^b |g_k(t)|\, dt < \infty.$$

Conclude that $G(t) < \infty$ for almost every t, hence $\sum_k g_k(t)$ converges absolutely for almost every t.

g) Define, for almost every t,

$$f(t) := f_{n_1}(t) + \sum_{k=1}^{\infty} g_k(t) = \lim_{m\to\infty} f_{n_m}(t).$$

Show that f is measurable. Then prove that

$$\|f_{n_m} - f\|_2 \to 0.$$

(A standard way is to use Fatou's lemma[3] on $|f_{n_m} - f|^2$ and the fact that the tails $\sum_{k\ge m} g_k$ are small in L^2 because of the 2^{-k} bound.)

h) Use the fact that (f_n) is Cauchy: given $\varepsilon > 0$, choose N such that $\|f_n - f_m\|_2 < \varepsilon$ for all $n, m \ge N$. Choose m large enough that $n_m \ge N$ and $\|f_{n_m} - f\|_2 < \varepsilon$. Then for $n \ge N$,

$$\|f_n - f\|_2 \le \|f_n - f_{n_m}\|_2 + \|f_{n_m} - f\|_2 < 2\varepsilon.$$

Conclude that $\|f_n - f\|_2 \to 0$.

i) **Conclude.** We have shown that every Cauchy sequence in $L^2([a, b])$ converges in $\|\cdot\|_2$ to an element of $L^2([a, b])$. Therefore $L^2([a, b])$ is complete with respect to the norm induced by the inner product, hence it is a Hilbert space.

38. Show that the space $M_{m\times n}(\mathbb{R})$ (or $M_{m\times n}(\mathbb{C})$) endowed with the Frobenius inner product

$$\langle A, B\rangle = \mathrm{tr}(AB^{\mathsf{T}}) \qquad \text{(real case)}, \qquad \langle A, B\rangle = \mathrm{tr}(AB^{*}) \qquad \text{(complex case)}$$

is a Hilbert space.

39. Work out the details of the examples on p. 573.

40. Let H be an inner product space.

a) Prove the **Cauchy-Schwarz inequality:** $|\langle \mathbf{x}, \mathbf{y}\rangle| \le \|\mathbf{x}\|\,\|\mathbf{y}\|$.

b) Deduce the **triangle inequality** $\|\mathbf{x} + \mathbf{y}\| \le \|\mathbf{x}\| + \|\mathbf{y}\|$ for the induced norm.

41. Let H be an inner product space. Prove the **parallelogram identity:** for all $\mathbf{x}, \mathbf{y} \in H$,

$$\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2.$$

42. Let H be an inner product space and $M \subseteq H$.

[3] **Fatou's Lemma:** if (f_n) is a sequence of measurable functions with $f_n \ge 0$, then

$$\int \liminf_{n\to\infty} f_n \le \liminf_{n\to\infty} \int f_n.$$

(Integrals may take the value $+\infty$.)

a) Prove that $M^\perp$ is a linear subspace of H.

b) Prove that $M \subseteq (M^\perp)^\perp$.

c) If M is a subspace, prove that $M^\perp$ is closed.

43. Let $H = \mathbb{R}^n$ with the standard inner product, and let $M = \text{span}\{\mathbf{u}\}$ where $\mathbf{u} \neq \mathbf{0}$.

    a) Show that the unique minimizer of $\|\mathbf{x} - \mathbf{m}\|$ over $\mathbf{m} \in M$ is

    $$P_M\mathbf{x} = \frac{\langle \mathbf{x}, \mathbf{u}\rangle}{\langle \mathbf{u}, \mathbf{u}\rangle}\mathbf{u}.$$

    b) Prove that $\mathbf{x} - P_M\mathbf{x} \in M^\perp$.

    c) Prove the **Pythagorean identity** $\|\mathbf{x}\|^2 = \|P_M\mathbf{x}\|^2 + \|\mathbf{x} - P_M\mathbf{x}\|^2$.

44. Let H be a Hilbert space and $M \subseteq H$ a closed subspace. Assume that $\mathbf{m}_1, \mathbf{m}_2 \in M$ both minimize $\|\mathbf{x} - \mathbf{m}\|$ over $\mathbf{m} \in M$. Prove that $\mathbf{m}_1 = \mathbf{m}_2$.

45. Let H be a Hilbert space and let $(\mathbf{x}_n)$ and $(\mathbf{y}_n)$ be sequences in H with $\mathbf{x}_n \to \mathbf{x}$ and $\mathbf{y}_n \to \mathbf{y}$. Prove that $\langle \mathbf{x}_n, \mathbf{y}_n\rangle \to \langle \mathbf{x}, \mathbf{y}\rangle$.

46. Let $(\mathbf{e}_n)$ be an orthonormal sequence in a Hilbert space H. Is it the case that $\langle \mathbf{x}, \mathbf{e}_n\rangle \to 0$ as $n \to \infty$ for every $\mathbf{x} \in H$?

47. Let $(\mathbf{e}_n)$ be an orthonormal sequence in a Hilbert space H. Assume that $\langle \mathbf{x}, \mathbf{e}_n\rangle = 0$ for all n implies $\mathbf{x} = \mathbf{0}$. Is the linear span of $\{\mathbf{e}_n\}$ dense in H?

48. In $H = L^2([0,1])$ with inner product $\langle f, g\rangle = \int_0^1 f(t)g(t)\,dt$, let $M = \text{span}\{1\}$. Compute the orthogonal projection of $f(t) = t$ onto M and compute the L^2-distance from t to M.

49. If $F : U \to \mathbb{R}^n$ is C^1 on an open set $U \subseteq \mathbb{R}^n$ with $\mathbf{a} \in U$ and if $DF(\mathbf{a})$ is invertible, is F globally one-to-one on U?

50. Let $F : \mathbb{R}^2 \to \mathbb{R}^2$ be given by

    $$F(\mathbf{x}, y) = (e^x \cos y,\ e^x \sin y).$$

    Compute $DF(0,0)$ and verify it is invertible. Use the inverse function theorem to explain why F has a C^1 local inverse near $(0,0)$, and identify this local inverse explicitly.

51. Define $F : \mathbb{R}^2 \to \mathbb{R}^2$ by

    $$F(x, y) = (e^x \cos y,\ e^x \sin y).$$

    a) Compute $DF(x,y)$ and $\det DF(x,y)$.

    b) Use the inverse function theorem to show that F is locally invertible at every point of $\mathbb{R}^2$.

    c) Identify an explicit local inverse near a point (x_0, y_0) (in terms of log and an appropriate branch of the argument).

52. Let $F : U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be C^1 and assume that det $DF(\mathbf{a}) \neq 0$ for some $\mathbf{a} \in U$.

   a) Show that there exist neighbourhoods V of $\mathbf{a}$ and W of $F(\mathbf{a})$ such that $F : V \to W$ is a C^1 bijection with a C^1 inverse.

   b) Prove that there exists $c > 0$ and a neighbourhood $V_0 \subseteq V$ of $\mathbf{a}$ such that

   $$\|F(\mathbf{x}) - F(\mathbf{z})\| \geq c\,\|\mathbf{x} - \mathbf{z}\| \quad \text{for all } \mathbf{x}, \mathbf{z} \in V_0.$$

   (Hint: use continuity of $(DF(\mathbf{x}))^{-1}$ near $\mathbf{a}$ and the mean value formula.)

53. Consider $G : \mathbb{R}^2 \to \mathbb{R}$ given by $G(x, y) = x^2 + y^2 - 1$.

   a) Verify that $D_{\mathbf{y}}G(0, 1) \neq 0$.

   b) Use the implicit function theorem to obtain a C^1 function φ defined near $x = 0$ such that $G(x, \varphi(x)) \equiv 0$ and $\varphi(0) = 1$.

   c) Compute $\varphi'(0)$ using the implicit differentiation formula.

54. Let $G : \mathbb{R}^{n+m} \to \mathbb{R}^m$ be C^1, and suppose $G(\mathbf{a}, \mathbf{b}) = \mathbf{0}$ with $(\mathbf{a}, \mathbf{b}) \in \mathbb{R}^n \times \mathbb{R}^m$. Assume that the $m \times m$ matrix $D_{\mathbf{y}}G(\mathbf{a}, \mathbf{b})$ (the derivative with respect to the last m variables) is invertible.

   a) State precisely what the inverse function theorem implies about solving $G(\mathbf{x}, \mathbf{y}) = 0$ for y as a function of x near $(\mathbf{a}, \mathbf{b})$.

   b) Show that the resulting function φ satisfies $\varphi(\mathbf{a}) = \mathbf{b}$ and compute $D\varphi(\mathbf{a})$ in terms of $D_xG(\mathbf{a}, \mathbf{b})$ and $D_{\mathbf{y}}G(\mathbf{a}, \mathbf{b})$.