Chapter 2

Approximation problem

In this chapter we study the best approximation problem in Hilbert and Banach spaces. In the focus are the questions of existence and uniqueness. The metric projection operators describe the best approximations. We will see in the next chapters that properties of the metric projections correspond to geometric properties of Banach spaces described by the duality mapping.

Metric projection operators are widely used in different areas of mathematics such as functional and numerical analysis, theory of optimization and approximation and for problems of optimal control; see Chapter 5.

The list of contributions to the theory of approximation is really very huge. Textbooks are for instance Achieser [1], Butzer and Berens [7], Cheney [8], Deutsch [9], de Vore and Lorentz [10], Holmes [12], Mhaskar and Pai [18], [20], Singer [22]. To original articles we refer at those places where they are used or where they have deep relations to the results presented.

2.1 Best approximation problem

Let (X, d) be a metric space and let C be a nonempty subset of X. For every $x \in X$, the distance between the point x and the set C is denoted by $\mathrm{dist}(x, C)$ and is defined by the following minimum equation

$$\mathrm{dist}(x, C) = \inf_{v \in C} d(x, v)\,. \tag{2.1}$$

We call the problem (2.1) the **best approximation problem** associated to C and $w \in C$ with $d(x, w) = \mathrm{dist}(x, C)$ a **best approximation** of x in C. We refer to the map $X \ni x \longmapsto \mathrm{dist}(x, C) \in \mathbb{R}$ as the **distance function** for C. As a first result we have

Lemma 2.1. *Let (X, d) be a metric space and let C be a nonempty subset of X. Then the distance function dist$(\cdot, C)$ is Lipschitz continuous with Lipschitz constant $L \le 1$.*

Proof:
Let $x, y \in X, w \in C$. Then $\mathrm{dist}(x, C) \le d(x, w) \le d(x, y) + d(y, w)$ and hence $\mathrm{dist}(x, C) - d(x, y) \le d(y, w)$. We conclude $\mathrm{dist}(x, C) - d(x, y) \le \mathrm{dist}(y, C)$. By symmetry, $\mathrm{dist}(y, C) - d(x, y) \le \mathrm{dist}(x, C)$. Thus, $|\mathrm{dist}(x, C) - \mathrm{dist}(y, C)| \le d(x, y)$. ∎

Remark 2.2. *Let* (X, d) *be a metric space and let* C *be a nonempty subset of* X. *If* C *is additional bounded we may consider the concept of farthest points: For* z *in* C *we call* z *a farthest point of* x *in* C *if*

$$d(x, z) = \mathit{far}(x, C) := \sup_{u \in C} d(x, u) .$$

This induces the mapping $F_C : C \rightrightarrows C$, *defined by*

$$F_C : C \ni x \longmapsto \{z \in C : d(x, z) = \mathit{far}(x, C)\} .$$

It is easily prooved that far $: C \longrightarrow C$ *is nonexpansive.* □

The basic questions in the best approximation problem concerns the **existence, uniquness** and quality of the **dependence of the best approximation** on the data (X, d, C, x) of the problem. We shall discuss these questions in this chapter. In the next chapter we present various applications of the best approximation problem.

The **metric projection operator** P_C is defined as follows:

$$P_C(x) := \{z \in C : d(x, z) = \mathrm{dist}(x, C)\} ,\ x \in X .$$

This operator P_C is a set valued mapping from X into X with range in C. We write: $P_C : X \rightrightarrows X$. Obviously, $P_C(x) = \{x\}$ for all $x \in C$. The domain $\mathrm{dom}(P_C)$ is the set $\{x \in X | P_C(x) \neq \emptyset\}$.

Clearly, $P_C(x)$ is a closed subset of C if C is closed. If $P_C(x) \neq \emptyset$ for every $x \in X$, then C is called **proximal**. Obviously, a necessary condition for proximality is the closedness of C. If $P_C(x)$ is a singleton for every $x \in X$, then C is said to be a **Chebyshev set**. Clearly, in this situation P_C is a **projection**: $P_C \circ P_C = P_C$.

A natural extension of the best approximation problem is to find a **best approximating pair** relative to two sets C, D in a metric space, i.e.:

Given subsets C, D in the metric space (X, d) find $(u, w) \in C \times D$ with

$$\mathrm{dist}(C, D) := \inf_{u' \in C, w' \in D} d(u', w') = d(u, w) . \tag{2.2}$$

Best approximating pairs may not exist in general. If D reduces to a singleton $\{x\}$ then the problem (2.2) reduces to the problem (2.1). On the other hand, when the problem (2.2) is consistent, i.e. $C \cap D \neq \emptyset$, then problem (2.2) reduces to the well known **feasibilty problem** for two sets and its solution set is $\{(x, x) : x \in C \cap D\}$. The feasibility problem captures a wide range of problems in applied mathematics and engineering. In Chapter 5 we shall study methods to solve the feasibility problem, namely the **alternate projection methods** using the metric projections P_C, P_D.

In the middle of the 19-th century Chebyshev proved that in the Banach space $C[0, 1]$ of continuous functions on $[0, 1]$ (endowed by the supremum norm) the subspace of polynomials of degree $\leq n$ and the subset $R_{m,n}$ of all rational functions

$$\frac{a_0 + a_1 t + \cdots + a_n t^n}{b_0 + b_1 t + \cdots + b_m t^m}$$

for fixed m, n are Chebyshev sets. In finite-dimensional spaces each nonempty closed subset is proximal due to the continuity of the norm and the Heine-Borel theorem. Moreover, in finite-dimensional euclidean spaces, Chebyshev sets are completely described by the Theorem of Motzkin (see [19] and [6, 13, 16, 19, 23]): a nonempty closed subset C in the euclidean space $\mathbb{R}^n$ is a Chebyshev set if and only if C is convex. Therefore, convexity should be an important tool to study projection operators.

Convexity in a metric space needs some additional structure (geodesics). This the reason why we restrict ourselves to the case of metric projection operators in Hilbert spaces and Banach spaces. Then we may use the concept of convexity in an obvious manner; see the preliminaries in the preface. Additionally, differentiability of nonlinear mappings can be considered in order to study necessary and sufficient conditions for the best approximation.

The general theory of best approximation may be considered as the mathematical study that is motivated by the desire to find answers to the following basic questions:

Existence Which subsets are proximal?

Uniqueness Which subsets are Chebyshev sets?

Characterization How does one recognize when a given element is a best approximation?

Error of approximation How does one compute the error of approximations or at least sharp upper bounds for it?

Computation of best approximations Can one describe some useful algorithms for computing a best approximation?

Continuity of the best approximation process How does the metric projection vary as a function of the element to be approximated?

Applications What are problems in the applied sciences which are important motivations for developing the theory further?

In this monograph not all questions are answered in the same completeness and sharpness.

2.2 Proximal sets

The term *proximal set* was proposed by Killgrove and used first by Phelps. Let us start with some examples which show different cases concerning proximality.

Example 2.3.

(1) $\mathcal{X} := \mathbb{R}^2 := \{x = (u, v) : u, v \in \mathbb{R}\}$ *endowed with the* l_2*-norm* $\|(u, v)\|_2 := \sqrt{|u|^2 + |v|^2}$, $C := \{(u, v) : \|(u, v)\| \le 1\}$. *Then* C *is proximal and* $P_C(x) = \{(u, v)\|(u, v)\|_2^{-1}\}$ *for* $x \notin C$.

(2) $\mathcal{X} := \mathbb{R}^2 := \{x = (u, v) : u, v \in \mathbb{R}\}$ *endowed with the supremum* $\|(u, v)\|_\infty := \sup\{|u|, |v|\}$, $C := \{(0, v) : v \in \mathbb{R}\}$. *Then* C *is proximal and* $P_C(x) = \{(0, v) : |v| \leq 1\}$ *for* $x = (1, 0)$.

(3) $\mathcal{X} := \mathbb{R}^2 := \{x = (u, v) : u, v \in \mathbb{R}\}$ *endowed with the* l_1*-norm* $\|(u, v)\|_1 := |u| + |v|$ $C := \{(u, v) : v = \pm u\}$. *Then* C *is proximal and* $P_C(x) = \{(0, v) : |v| \leq 1\}$ *for* $x = (0, 1)$.

□

Definition 2.4. *Let* (X, d) *be a metric space and let* $C \subset X$. C *is called* **approximatively compact** *if for each* $x \in X$ *and for each sequence* $(u^n)_{n\in\mathbb{N}}$ *in* C *with* $\lim_n d(u^n, x) =$ *dist*(x, C) *there exists a subsequence* $(u^{n_k})_{k\in\mathbb{N}}$ *and* $u \in C$ *with* $\lim_k u^{n_k} = u$. □

Theorem 2.5. *Let* (X, d) *be a metric space and let* C *be a nonempty subset of* X. *Then we have:*

(a) *If* C *is approximately compact then* C *is proximal.*

(b) *If* C *is compact then* C *is proximal.*

Proof:
Ad (a) Let $x \in X$. Let $(u^n)_{n\in\mathbb{N}}$ be a minimal sequence. The point u in the definition 2.4 is in C and by the continuity of the metric $u \in P_C(x)$.
Ad (b) Obviously, a compact set is approximately compact. ■

Theorem 2.6. *Let* $\mathcal{X}$ *be a reflexive Banach space. Then every nonempty closed convex subset* C *of* $\mathcal{X}$ *is proximal.*

Proof:
Let $x \in \mathcal{X}$ and let $(u^n)_{n\in\mathbb{N}}$ be a minimal sequence in C; $\lim_n \|x - u^n\| = \mathrm{dist}(x, C), u^n \in C, n \in \mathbb{N}$. Then the sequence $(u^n)_{n\in\mathbb{N}}$ is bounded. Since $\mathcal{X}$ is reflexive this sequence contains a weakly convergent set. Let u be a weak cluster point of $(u^n)_{n\in\mathbb{N}}$. Since C is a closed convex set C is weakly closed; see [4]. Therefore $u \in C$. Due to the fact that the norm is sequential weakly lower semicontinuous we have $\|x - u\| = \mathrm{dist}(x, C)$. ■

Remark 2.7. *Theorem 2.6 can be reformulated in a stronger form, namely: A Banach space* $\mathcal{X}$ *is reflexive if and only if every nonempty closed convex subset* C *of* $\mathcal{X}$ *is proximal. See [17].* □

Example 2.8. *Consider the Banach space* l_1. *For any* $n \in \mathbb{N}$, *let* e^n *be the sequence in* l_1 *with* n*-th entry* 1 *and all other entries* 0. *Let* $b^n := (n+1)/ne^n, n \in \mathbb{N}$, *and let* $C := \overline{co}(b^1, b^2, \ldots, b^n, \ldots)$. *Then* C *is a nonempty convex closed subset of* l_1 *which is not proximal. Notice, that* l_1 *is a non-reflexive Banach space (with dual space* l_∞*).* □

Lemma 2.9. *Let* $\mathcal{X}$ *be a Banach space and let* $\lambda \in \mathcal{X}^*\backslash\{\theta\}$. *Then the following conditions are equivalent:*

(a) *ker*(λ) *is proximal.*

(b) λ *attains its norm, i.e.* $\langle \lambda, x\rangle = \|\lambda\|$ *for some* $x \in \overline{S}_1$.

Proof:
Set $U := \ker(\lambda)$.
$(a) \implies (b)$ Since $\lambda \neq \theta$ there exists $z \in X\backslash U$. Because U is proximal there exists $u \in U$ with $\|z-u\| = \mathrm{dist}(z, U)$. Let $x := (z-u)\|z-u\|^{-1}$. Clearly, $\|x\| = 1$ and

$$\mathrm{dist}(x, U) = \|z-u\|^{-1}\mathrm{dist}(z-u, U) = \|z-u\|^{-1}\mathrm{dist}(z, U) = 1\,.$$

Let $y \in \mathcal{X}$. Then $y = \langle\lambda, y\rangle\langle\lambda, x\rangle^{-1}x + v$ for some $v \in U$. Now, since U is a subspace with $\theta \in U$ we obtain

$$\begin{aligned}
\frac{|\langle\lambda, y\rangle|}{|\langle\lambda, x\rangle|} &= \frac{|\langle\lambda, y\rangle|}{|\langle\lambda, x\rangle|}\mathrm{dist}(x, U) = \frac{|\langle\lambda, y\rangle|}{|\langle\lambda, x\rangle|}\inf_{z\in U}\|x-z\| \\
&= \inf_{z\in U}\left\|\frac{\langle\lambda, y\rangle}{\langle\lambda, x\rangle}x + u - z\right\| = \inf_{z\in U}\|y-z\| \\
&= \mathrm{dist}(y, U) \le \|y\|
\end{aligned}$$

Therefore, $|\langle\lambda, y\rangle| \le \langle\lambda, x\rangle|\|y\|$ for all $y \in X$. Hence, $\|\lambda\| \le |\langle\lambda, x\rangle| \le \|\lambda\|$ since $\|x\| = 1$. Thus, $\|\lambda\| = |\langle\lambda, y\rangle|$.
$(b) \implies (a)$ Suppose λ attains its norm at $x \in \overline{S}_1$, i.e. $\langle\lambda, x\rangle = \|\lambda\|$ for some $x \in \overline{S}_1$. Let $u \in U$. Then

$$1 = \frac{\langle\lambda, x\rangle}{\|\lambda\|} = \frac{\langle\lambda, x-u\rangle}{\|\lambda\|} \le \frac{\|\lambda\|\|x-u\|}{\|\lambda\|} = \|x-u\|\,.$$

Therefore, $1 \le \mathrm{dist}(x, U) \le \|x-\theta\| = 1$. Thus $\theta \in P_U(x)$. Let $y \in X$. We can write $y = ax + u$ for some $a \in \mathbb{R}, u \in U$. It follows $P_U(y) = P_U(ax+u) = aP_U(x) + u \neq \emptyset$. Hence, U is proximal. ■

Lemma 2.10. *Let $\mathcal{H}$ be a Hilbert space and let $C, C_1, \dots, C_m$ be nonempty closed convex subsets of $\mathcal{H}$. The we have:*

(1) $P_C(x) = P_{C-y}(x-y) + y$ *for all* $x, y \in \mathcal{H}$.

(2) $P_{rC}(x) = rP_C(\frac{1}{r}x)$ *for all* $x \in \mathcal{H}, r \neq 0$.

(3) $(P_{C_1} \circ \cdots \circ P_{C_m})^n(x) = ((P_{C_1-y} \circ \cdots \circ P_{C_m-y})^n(x-y) + y$ *for all* $x, y \in \mathcal{H}, n \in \mathbb{N}$.

Proof:
Ad (1) Let $x, y \in \mathcal{H}$. Then for all $u \in C$

$$\|x - P_{C-y}(x-y) - y\| = \|x - y - (u-y)\| = \|x-u\|$$

which shows that $P_C(x) = P_{C-y}(x-y) + y$.
Ad (2) Let $x \in Hs, r > 0$. Then for all $u \in C$

$$\|x - rP_C(\frac{1}{r}x)\| = r\|\frac{1}{r}x - P_C(\frac{1}{r}x)\| \le r\|\frac{1}{r}x - u\| = \|x - ru\|$$

which shows that $P_{rC}(x) = rP_C(\frac{1}{r}x)$.
Ad (3) This is a consequence of (1). ■

2.3 Chebyshev sets

The term *Chebyshev set* was introduced by Stechkin in honor of the founder of best approximation theory. Let us start with some examples which show different cases concerning Chebyshev sets.

Example 2.11.

(1) $\mathcal{X} := \mathbb{R}^2 := \{x = (u,v) : u,v \in \mathbb{R}\}$ *endowed with the* l_2*-norm* $\|(u,v)\|_2 := \sqrt{|u|^2 + |v|^2}$, $C := \{(u,v) : \|(u,v)\| \le 1\}$. *Then* C *is a Chebyshev set since* $P_C(x) = \{(u,v)\|(u,v)\|_2^{-1}\}$ *for all* $x = (u,v) \ne \theta$.

(2) $\mathcal{X} := \mathbb{R}^2 := \{x = (u,v) : u,v \in \mathbb{R}\}$ *endowed with the supremum* $\|(u,v)\|_\infty := \sup\{|u|,|v|\}$, $C := \{(0,v) : v \in \mathbb{R}\}$. *Then* C *is proximal but not a Chebyshev set. In fact,* $P_C(x) = \{(0,v) : |v| \le 1\}$ *for* $x = (1,0)$.

(3) $\mathcal{X} := \mathbb{R}^2 := \{x = (u,v) : u,v \in \mathbb{R}\}$ *endowed with the* l_1*-norm* $\|(u,v)\|_1 := |u| + |v|$ $C := \{(u,v) : v = \pm u\}$. *Then* C *is proximal but not a Chebyshev set. In fact,* $P_C(x) = \{(0,v) : |v| \le 1\}$ *for* $x = (0,1)$.

(4) $\mathcal{X} := \mathbb{R}^2 := \{x = (u,v) : u,v \in \mathbb{R}\}$ *endowed with the norm* $\|(u,v)\| := |u-v| + \sqrt{u^2+v^2}$, $C := \{(u,0) : |u| \le 1\}$. *Then* C *is a Chebyshev set.*

□

Definition 2.12. *A normed space* $(\mathcal{X}, \|\cdot\|)$ *is called* **strictly convex** *or* **rotund**, *if for all* $x,y \in \mathcal{X}$ *with* $\|x\| = \|y\| = 1, x \ne y$, *we have* $\|x+y\| < 2$. □

Example 2.13. *The Banach space* $(C[0,1], \|\cdot\|_\infty)$ *is not strictly convex. To show this, choose* $x(t) := 1, y(t) := t$, $t \in [0,1]$. *Then we have* $x \ne y$ *and* $\|x\|_\infty = \|y\|_\infty = 1$. *But it holds* $\|x+y\|_\infty = 2$.

Lemma 2.14. *Let* $\mathcal{X}$ *be a Banach space. Then the following assertions are equivalent:*

(a) $\mathcal{X}$ *is strictly convex.*

(b) If $x,y \in \mathcal{X}$ *with* $\|x\| = \|y\| = 1, x \ne y$, *we have* $\|tx + (1-t)y\| < 1$.

(c) If $u,v \in \mathcal{X}$ *and* $\|u+v\| = \|u\| + \|v\|$ *then* $u = tv$ *for some* $t \ge 0$.

(d) Every point in the boundary $\overline{S}_1$ *of* $\overline{B}_1$ *is an extremal point of* $\overline{B}_1$.

Proof:
$(a) \implies (b)$ Let $x,y \in \mathcal{X}$ with $\|x\| = \|y\| = 1, x \ne y$. Assume that $\|t_*x + (1-t_*)y\| = 1$ for some $t_* \in (0,1)$. Obviously (by (a) $t_* \ne \frac{1}{2}$. Consider the case $t_* < \frac{1}{2}$. Using the point $2t_*x + (1-2t_*)x$ we obtain the contradiction $\|t_*x + (1-t_*)y\| < 1$. The case $t_* > \frac{1}{2}$ is similar.
$(b) \implies (c)$ Let $u,v \in \mathcal{X}$ with $\|u+v\| = \|u\| + \|v\|$. If $u = \theta$ or $v = \theta$ nothings has to be proved. Let $u,v \ne \theta$. Then we have

$$\Big\|\frac{\|u\|}{\|u\|+\|v\|}\frac{u}{\|u\|} + \frac{\|v\|}{\|u\|+\|v\|}\frac{v}{\|v\|}\Big\| = 1 .$$

Then by (b)

$$\frac{u}{\|u\|} = \frac{v}{\|v\|}$$

and we have $u = \|u\|\|v\|^{-1}v$.
$(c) \implies (a)$ Let $x, y \in \mathcal{X}$ with $\|x\| = \|y\| = 1$. If $\|x+y\| = 2$ then $\|x+y\| = \|x\|+\|y\|$. This implies $x = ty$ for some $t \geq 0$. Obviously $t = 1$ and we have $x = y$.
$(a) \iff (d)$ If x belongs to the segment $[u, v]$ we may assume without loss of generality that $x = \frac{1}{2}(u+v)$. Then the equivalence of (a) and (d) is obviously true. ■

Theorem 2.15. *Let $\mathcal{X}$ be a Banach space. Then the following conditions are equivalent:*

(a) $\mathcal{X}$ is reflexive and strictly convex.

(b) Every nonempty convex closed subset C of $\mathcal{X}$ is a Chebyshev set.

Proof:
$(a) \implies (b)$ We know from Theorem 2.6 already that every closed convex subset C is proximal. Let $a := \text{dist}(x, C)$ and let $z, z' \in P_C(x)$. Then $\|x - z\| = \|x - z'\| = a$. Since $u := \frac{1}{2}(z + z') \in C$ we have $\|x - u\| \geq a$. Using the triangle inequality we obtain $\|x - u\| \leq a$. This implies $\|x - u\| = a$ and since $\mathcal{X}$ is strictly convex we conclude $u = z = z'$.
$(b) \implies (a)$ We know from Remark 2.16 that $\mathcal{X}$ is a reflexive space. Applying (b) to the set $\overline{B}_1$ we obtain that $\mathcal{X}$ is strictly convex. ■

Remark 2.16. *Theorem 2.15 can be reformulated in a stronger form, namely: a reflexive Banach space $\mathcal{X}$ is strictly convex if and only if every nonempty closed convex subset C of $\mathcal{X}$ is a Chebyshev set; see [17].* □

In a Hilbert space every nonempty closed convex subset is a Chebyshev set since a Hilbert space is reflexive (see the preliminaries in the preface) and strictly convex due to the parallelogram identity. We shall give a proof of the fact that a nonempty closed convex subset is a Chebyshev set based on the parallelogram identity and the completeness only.

Theorem 2.17. *Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty convex closed subset of $\mathcal{H}$. Then C is Chebyshev set. Additionally, we have the property that each minimal sequence for the minimization of $C \ni y \longmapsto \|x - y\| \in \mathbb{R}$ converges to $P_C(x)$.*

Proof:
We prove $P_C(x)$ is not empty.
Let $(y^n)_{n\in\mathbb{N}}$ be sequence with $\lim_n \|x - y^n\| = a := \text{dist}(x, C)$. We want to show that $(y^n)_{n\in\mathbb{N}}$ is a Cauchy sequence. For m, n we have

$$\begin{aligned}
\|y^n - y^m\|^2 &= \|(x - y^m) - (x - y^n)\|^2 \\
&= -\|(x - y^m) + (x - y^n)\|^2 + 2\|x - y^m\|^2 + 2\|x - y^n\|^2 \\
&= -4\|\frac{1}{2}(y^m + y^n) - x\|^2 + 2\|x - y^m\|^2 + 2\|x - y^n\|^2 \\
&\leq -4a^2 + 2\|x - y^m\|^2 + 2\|x - y^n\|^2
\end{aligned}$$

This shows that $\lim_{m,n} \|y^n - y^m\| = 0$ and $(y^n)_{n\in\mathbb{N}}$ is a Cauchy sequence. Therefore there exists $y := \lim_n y^n$ and by the continuity of the norm we have $\|x - y\| = a$, i.e. $y \in P_C(x)$.

We prove that $P_C(x)$ is a singleton.
Let $u, v \in P_C(x) : \|x-u\| = \|x-v\| = a$. By the convexity of C we have $\frac{1}{2}(u+v) \in C$ and therefore

$$\|x - \frac{1}{2}(u+v)\| \le \frac{1}{2}\|x-u\| + \frac{1}{2}\|x-v\| = a .$$

Hence

$$\|u-v\|^2 = \|(x-u)-(x-v)\|^2 = 2\|x-u\|^2+\|x-v\|^2-4\|x-\frac{1}{2}(u+v)\|^2 \le 2a^2+2a^2-4a^2 = 0$$

and we conclude $u = v$.
The additional assertion follows from the first part of our proof. ■

Remark 2.18. *There are some very surprising results concerning the chebyshevian property: First, there exist no finite-dimensional Chebyshev subspace in* $L_1[0,1]$. *Second, there exists a separable reflexive Banach space without any finite-dimensional Chebyshev subspace.*

A very important open problem is: Does there exist a non-convex Chebyschev subset in the Hilbert space. This question goes back to Efimov, Klee and Stechkin who have adduced plausible evidence to support the conjecture that there exist non-convex Chebyshev subsets; see [2]. On the other hand, they gave sufficient conditions for Chebyshev sets in Hilbert space beeing convex. Here is a list of such results concerning this question:

(1) If C *is a Chebyshev set in* $\mathbb{R}^2$ *then* C *is convex (Bunt [6], 1934).*

(2) If C *is a Chebyshev set in* $\mathbb{R}^n$ *then* C *is convex (Kritikos [16], 1938).*

(3) If C *is a boundedly compact*[1] *Chebyshev set, then* C *is convex (Efimov and Stechkin [11], 1959).*

(4) If C *is a weakly closed Chebyshev set in a Hilbert space then* C *is convex (Klee [13], 1961).*

(5) If a set C *is a Chebyshev set in a Hilbert space and each half space intersects* C *in a proximal set, then* C *is convex (Singer [21], 1967).*

(6) If C *is an approximately compact Chebyshev set in a Hilbert space, then* C *is convex (Klee [14], 1967).*

(7) If C *is a Chebyshev set in a Hilbert space with a continuous metric projection then* C *is convex (Asplund [2], 1969).*

□

2.4 Well posedness of the approximation problem

Definition 2.19. *Let* $\mathcal{X}$ *be a Banach space and let* C *be a nonempty subset of* $\mathcal{X}$. *Then the approximation problem* (2.1) *is called* **stable** *if every sequence* $(x^n)_{n\in\mathbb{N}}$ *in* C *with* $\lim_n \|x - x^n\| = dist(x, C)$ *(minimizing sequence) satisfies* $\lim_n dist(x^n, P_C(x)) = 0$. *If in addition the solution set* $P_C(x)$ *is a singleton set, the approximation problem* (2.1) *is called* **strongly solvable**. □

[1] C is called **boundedly compact** if for any $r > 0$ the set $\overline{B}_r \cap C$ is compact.

The property that the best approximation problem (2.1) is **strongly solvable** is saying that the best approximation problem is **well posed in the sense of Hadamard**: there exists a unique solution which depends continously on the point being approximated.

Corollary 2.20. *Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty closed convex subset of $\mathcal{H}$. Then the best approximation problem* (2.1) *is strongly solvable.*

Proof:
See Theorem 2.17. ■

For the discussion of the well posedness-property in Banach spaces we have to introduce some other facts concerning the geometric properties of Banach spaces.

Definition 2.21. *A Banach space $\mathcal{X}$ is an* **E-space** *if the following conditions hold:*

(1) $\mathcal{X}$ is reflexive.

(2) $\mathcal{X}$ is strictly convex.

(3) If $x \in \overline{S}_1$ is the weak limit of the sequence $(x^n)_{n\in\mathbb{N}}$ in $\overline{S}_1$ then x is the strong limit of this sequence.

□

Definition 2.22. *Let $\mathcal{X}$ be a Banach space. We say that $\mathcal{X}$ has a* **Kadek-Klee norm** *if for every sequence $(x^n)_{n\in\mathbb{N}}$ in $\mathcal{X}$ with $w - \lim_n x^n = x$ and $\lim_n \|x^n\| = \|x\|$ the sequence $(x^n)_{n\in\mathbb{N}}$ converges in norm to x.* □

A Banach space has a Kadek-Klee norm if the weak and strong convergence coincides on the unit sphere. It is easy to check that the property (3) in Definition 2.21 is equivalent to the fact that $\mathcal{X}$ has the Kadek-Klee property.

Clearly, since weak convergence is equivalent to the strong convergence in a finite-dimensional normed space, every finite-dimensional normed space has a Kadek-Klee norm. Each Hilbert space, each uniformly convex space and all l_p-, $L_p(\Omega)-$spaces, $1 < p < \infty$, have a Kadek-Klee norm.

Remark 2.23. *A more suggestive definition of an E-space is the following: A Banach space $\mathcal{X}$ is an E-space if it is strictly convex and every weakly closed subset is approximately compact; see Theorem 2.24 and [12, 15]. The equivalence of this definition with that given in Definition 2.21 can proved by using two deep results in functional analysis: firstly, a Banach space is reflexive iff its closed unit ball is weakly compact; secondly, a nonempty closed convex subset C of a Banach space $\mathcal{X}$ is weakly compact iff every $\lambda \in \mathcal{X}^*$ attains its maximum on C. The last result is a very useful result since it characterizes weak compactness without using the weak topology. We know from the Bishop-Phelps theorem (see [5]) that in a Banach space the set of linear functionals which attain their maximum on a nonempty bounded closed convex subset is a (norm) dense subset of $\mathcal{X}^*$.* □

Theorem 2.24. *Let $\mathcal{X}$ be an E-space and let C be a nonempty closed convex subset of $\mathcal{X}$. Then the best approximation problem* (2.1) *is strongly solvable.*

Proof:
Clearly, C is Chebyshev set by $(1),(2)$ in Definition 2.21. We have to show that the approximation problem (2.1) is stable.
Let $x \in \mathcal{X}$ and let $(x^n)_{n\in\mathbb{N}}$ be sequence in C with $\lim_n \|x - x^n\| = \text{dist}(x, C)$. Obviously, the sequence $(x^n)_{n\in\mathbb{N}}$ is bounded. Since $\mathcal{X}$ is reflexive there exists a weakly convergent subsequence $(x^{n_k})_{k\in\mathbb{N}}$. Let $z := w - \lim_k x^{n_k}$. Since C is closed and convex it is weakly closed too and we have $z \in C$. Then $x - z = w - \lim_k(x - x^{n_k})$ and by the weak lower semicontinuity of the norm

$$\|x - z\| \le \liminf_k \|x - x^{n_k}\| = \text{dist}(x, C)\,.$$

This implies $z = P_C(x)$ and we conclude that $(x^n)_{n\in\mathbb{N}}$ converges weakly to $z = P_C(x)$.
If $a := \text{dist}(x, C) = 0$ then $z = x$ and $\lim_n x^n = x$ since $\lim_n \|x - x^n\| = a = 0$. Now assume $a > 0$. Then we may assume that $\|x - x^n\| \ge a/2 > 0, n \in \mathbb{N}$. Since $\mathcal{X}$ is reflexive there exists a subsequence $(x^{n_k})_{k\in\mathbb{N}}$ with

$$\frac{x - z}{a} = w - \lim_k \frac{x - x^{n_k}}{\|x - x^{n_k}\|}\,.$$

We know $\|x - z\| = a$. Now we apply (3) in Definition 2.21 and obtain

$$\frac{x - z}{a} = \lim_k \frac{x - x^{n_k}}{\|x - x^{n_k}\|}$$

and finally, $x - z = \lim_k(x - x^{n_k})$. ■

In Section 4.1 we will see that the property to be an E-space can be described as the uniform convexity of the dual space. There, we can give examples of E-spaces. The crucial point is the verification of the condition (3) in the definition of E-spaces.

2.5 Lower bounds for the approximation error

Here we are interested in upper and lower bounds for the „error“ $\text{dist}(x, C)$ for approximating x by vectors in C. Of course, upper bounds are easily obtained from the inequality

$$\text{dist}(x, C) \le \|x - u\|\,,\ u \in C\,.$$

Therefore, we have to develop an approach to obtain lower bounds. The idea is to find a way to write down the infimum $\text{dist}(x, C) = \inf_{u\in C} \|x - u\|$ by a supremum over a set which can be handled. Usually, such a way uses duality arguments.

Let $\mathcal{X}$ be a Banach space, let C be a nonempty closed convex subset and let $x \in \mathcal{X}\backslash C$. Then we have

$$\inf_{u\in C} \|x - u\| = \inf_{y\in\mathcal{X}} (\|x - y\| + \delta_C(y)) = \inf_{y\in\mathcal{X}} (\|y\| + \delta_C(x - y))\,,$$

where δ_C is the **characteristic function** of convex analysis:

$$\delta_C(z) := \begin{cases} 0 & \text{if } z \in C \\ \infty & \text{if } z \notin C \end{cases}.$$

Now, we may use the duality theorem for convex programs (see Theorem 10.62):

$$\inf_{y\in\mathcal{X}}(\|y\|+\delta_C(x-y)) = \sup_{\lambda\in\mathcal{X}^*,\|\lambda\|\le 1}\ \sup_{u\in C}\langle\lambda, x-u\rangle \tag{2.3}$$

since we have

$$\nu^*(\lambda) = \delta_{\overline{B}_1}(\lambda)\,,\ \delta_C(x-\cdot)^*(\lambda) = \sup_{u\in C}\langle\lambda, x-u\rangle\,,\ \lambda\in\mathcal{X}^*\,,$$

where $\nu : \mathcal{X}\ni z \longmapsto \|z\|\in\mathbb{R}$ is the norm function.

Theorem 2.25. *Let $\mathcal{X}$ be a Banach space, let C be a nonempty closed convex subset of $\mathcal{X}$ and let $x\in\mathcal{X}\backslash C$. Then*

$$dist(x, C) = \max_{\lambda\in\mathcal{X}^*,\|\lambda\|\le 1}(\langle\lambda, x\rangle - \sup_{u\in C}\langle\lambda, u\rangle) \tag{2.4}$$

Proof:
We have to argue that the supremunm in (2.3) is actually a maximum. But this is consequence of the fact that the norm function is continuous and has range in $\mathbb{R}$. ■

The result above may be interpreted that the distance from a point x to a set C may be seen as the maximum of the distances from x to hyperplanes that seperate x and C; see Theorem 10.62.

Now, the formula in (2.4) provides us with an easy way of obtaining **lower bounds** for $\mathrm{dist}(x, C)$:

$$\text{If } \lambda\in\mathcal{X}^* \text{ with } \|\lambda\|\le 1 \text{ then } \langle\lambda, x\rangle - \sup_{u\in C}\langle\lambda, u\rangle \le \mathrm{dist}(x, C)\,. \tag{2.5}$$

2.6 Appendix: Convexity I

2.7 Conclusions and comments

[3]

2.8 Exercises

1.) Let $\mathcal{X}$ be a strictly convex Banach space and let $C\subset\mathcal{X}$ be a nonempty closed subset of $\mathcal{X}$. Suppose $y\in P_C(x)$. Then

$$P_C(tx+(1-t)y) = \{y\} \text{ for all } t\in(0,1)\,.$$

2.) Let C be a nonempty subset of the euclidean space $\mathbb{R}^n$. Show the equivalence of the following conditions:

(a) C is a Chebyshev set

(b) C is closed and convex

3.) Let $\mathbb{R}^2$ be equipped with the l_1-norm and let $C := \{x = (x_1, x_2) : x_2 = \pm x_1\}$. Compute $P_C(x)$ for $x = (0, 1)$ and show that C is no Chebyshev set.

4.) Let $\mathbb{R}^n$ equipped with the norm

$$\|(x_1, x_2)\|_{12} := |x_1 - x_2| + (x_1^2 + x_2^2)^{\frac{1}{2}} .$$

Then the unit ball $C := \{(x_1, x_2)\|_{12} \leq 1\}$ is a Chebyshev set.

5.) Consider the space $C[0, 1]$ of continuous functions on $[0, 1]$ equipped with the supremum norm $\|\cdot\|_\infty$. Find an equivalent norm $\|\cdot\|$ such that $(C[0, 1], \|\cdot\|)$ is strictly convex.

6.) Let $\mathcal{X}$ be a Banach space and let C be a nonempty subset of $\mathcal{X}$. We set

$$\hat{C} := \{x \in \mathcal{X} : \|x\| = \inf_{y \in C} \|x - y\|\} .$$

Show:

(a) $C \cap \hat{C} = \emptyset$ or $C \cap \hat{C} = \{\theta\}$.

(b) $\hat{C}$ is closed.

(c) C is a Chebyshev set if and only if $\mathcal{X} = C \oplus \hat{C}$.

7.) Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty closed convex subset of $\mathcal{H}$. We define:

$$\|P_C\| := \sup_{x \in \mathcal{H} \backslash \{\theta\}} \frac{\|P_C(x)\|}{\|x\|} .$$

Show:

(1) $\|P_C\| = \begin{cases} 0 & \text{if } C = \{\theta\} \\ 1 & \text{if } C \neq \{\theta\} \end{cases}$ if $\theta \in C$.

(2) $\|P_C\| = \infty$ if $\theta \notin C$.

8.) Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty closed convex subset of $\mathcal{H}$. We define:

$$\|P_C\| := \sup_{x \in \mathcal{H} \backslash \{\theta\}} \frac{\|P_C(x)\|}{\|x\|} .$$

Show:

(1) $\|P_A \circ P_B\| = \begin{cases} \sup_{x \in B \backslash \{\theta\}} \dfrac{\|P_A(x)\|}{\|x\|} & \text{if } B \neq \{\theta\} \\ 0 & \text{if } B = \{\theta\} \end{cases}$

(2) If A, B are closed convex subsets of $\mathcal{H}$ with $\theta \in A \cap B$ then $\|P_A \circ P_B\| \leq 1$.

9.) Let $\mathcal{X}$ be a normed space. Then the following conditions are equivalent:

(a) $\mathcal{X}$ is uniformly convex.

(b) If $(x^n)_{n \in \mathbb{N}}, (y^n)_{n \in \mathbb{N}}$ are sequences in $\mathcal{X}$ with $\|x^n\| = \|y^n\| = 1, n \in \mathbb{N}$, then we have:

$$\text{From } \lim_n \|\frac{1}{2}(x^n + y^n)\| = 1 \text{ we conclude } \lim_n (x^n - y^n) = \theta .$$

(c) If $(x^n)_{n\in\mathbb{N}}, (y^n)_{n\in\mathbb{N}}$ are sequences in $\mathcal{X}$ with $\limsup_n \|x^n\| \le 1, \limsup_n \|y^n\| \le 1, n \in \mathbb{N}$, we have:

$$\text{From } \lim_n(\|\frac{1}{2}(x^n + y^n)\|) = 1 \text{ we conclude } \lim_n(x^n - y^n) = \theta\,.$$

10.) Let $\mathcal{X}$ be a uniformly convex Banach space. If $(x^n)_{n\in\mathbb{N}}$ is a sequence in $\mathcal{X}$ with

$$\alpha := \lim_n \|x^n\| = \lim_{n,m} \|\frac{1}{2}(x^n + x^m)\|\,,$$

Then this sequence is convergent.

11.) Let $\mathcal{X}$ be a uniformly convex Banach space and let $(x^n)_{n\in\mathbb{N}}$ be a sequence in X. Then the following conditions are equivalent:

(a) $\lim_n x^n = x$.

(b) $w - \lim_n x^n = x$, $\lim_n \|x^n\| = \|x\|$.

12.) Let $\mathcal{H}$ be a Hilbert space, let A, B be a closed subsets of $\mathcal{H}$ and let $x \in \mathcal{H}$. Show: If $P_B(x) \in A$, then $P_A(x) = P_B(x)$.

13.) Let $\mathcal{X}$ be a Banach space and let $U \subset \mathcal{X}$ be closed subspace which is a Chebyshev space. Show: $P_U^{-1}(x) = x + P_U^{-1}(\theta)$ for all $x \in \mathcal{X}$.

14.) Let $\mathcal{X}$ be a Banach space and let $U \subset \mathcal{X}$ be closed subspace which is a Chebyshev space. Set $U^\theta := P_U^{-1}(\theta)$. Show $\mathcal{X} = U \oplus U^\theta$.

15.) Let $\mathcal{X}$ be a Banach space and let $H := H_{\lambda,\alpha} := \{x \in \mathcal{X} : \langle\lambda, x\rangle = \alpha\}$ be a hyperplane ($\lambda \in \mathcal{X}^*, \lambda \ne \theta, \alpha \in \mathbb{R}$). Show $\text{dist}(x, H) = |\langle\lambda, x\rangle - \alpha|\|\lambda\|^{-1}$ for all $x \in \mathcal{X}$.

16.) Let $\mathcal{X}$ be a Banach space and let $H := H_{\lambda,\alpha} := \{x \in \mathcal{X} : \langle\lambda, x\rangle = \alpha\}$ be a hyperplane ($\lambda \in \mathcal{X}^*, \lambda \ne \theta, \alpha \in \mathbb{R}$) which is a Chebyshev set. Show that P_H is linear.

17.) Define $f : \mathbb{R} \ni r \longmapsto \frac{1}{2}\text{dist}(r, 2\mathbb{Z}) \in \mathbb{R}$. Set $C := \text{graph}(f)$ and consider $\mathbb{R}^2$ endowed with the l_1-norm. Show that C is a nonconvex Chebyshev set.

18.) Let $\mathcal{H}$ be a Hilbert space and let C be a nonempty closed convex subset fof $\mathcal{H}$. Suppose that $U : \mathcal{H} \longrightarrow \mathcal{H}$ is a surjective linear isometry. Then $U(C)$ is a nonempty closed convex subset of $\mathcal{H}$ and $P_{U(C)} = U \circ P_C \circ U^*$.

19.) Consider $\mathbb{R}^n$ endowed with the l_2-norm. Let

$$U := \{x = (x_1, \dots, x_n) : \sum_{i=1}^n x_i = 0\}.$$

(a) Show that U is a linear subspace of $\mathcal{H}$ with $\dim U = n - 1$.

(b) Compute $P_U(e^i), i = 1, 2, \dots$. (Here, $e^i = (\delta_{ij})$.)

(c) Compute $P_U(x)$ for all $x \in \mathbb{R}^n$.

20.) Consider $\mathbb{R}^n$ endowed with the l_2-norm. Let

$$C := \{x = (x_1, \dots, x_n) : x_1 \le x_2 \le \dots \le x_n\}.$$

Show that C is a Chebyshev set and a convex cone. Compute $P_C(x)$ for all $x \in \mathbb{R}^n$.

21.) Let $\mathcal{X}$ be a Banach space and let U be a linear subspace. Denote by $U^\perp$ the set

$$\{\lambda \in \mathcal{X}^* : \langle \lambda, u \rangle = 0 \text{ for all } u \in U\}.$$

Show: For all $x \in \mathcal{X} \backslash U$ we have

$$\operatorname{dist}(x, U) = \max_{\lambda \in U^\perp \cap \overline{B}_1} |\langle \lambda, x \rangle| .$$

22.) Consider the Banach space c_0 of the real sequences converging to zero. c_0 is endowed with the supremum norm. Let

$$U := \{x = (x^n)_{n \in \mathbb{N}} : \sum_{n=1}^{\infty} 2^{-n} x^n = 0\}.$$

Show:

(a) U is a closed linear subspace of c_0.

(b) For every $x \in c_0 \backslash U$ there exists no $u \in U$ with $\operatorname{dist}(x, U)$.

23.) Let A, B, C be convex subsets of $\mathbb{R}^n$ Suppose that B is closed and C is bounded. Show that $A + C \subset B + C$ implies $A \subset B$.

Bibliography

[1] N.I. Achieser. *Theory of Approximation.* Dover Publications, New York, 1992.

[2] E. Asplund. Chebyshev sets in Hilbert space. *Trans. Amer. Math. Soc.*, 240:235–240, 1969.

[3] A. Assadi, H. Haghshenas, and T.D. Narang. A look at proximal and Chebyschev sets in Banach spaces. *Le Matematiche*, LXIX:71–87, 2014.

[4] J. Baumeister. Konvexe Analysis, 2014. Skriptum WiSe 2014/15, Goethe–Universität Frankfurt/Main.

[5] J.M. Borwein and A.S. Lewis. *Convex Analysis and Nonlinear Optimization. Theory and Examples.* Springer, New York, 2006.

[6] L. Bunt. Multivariate quadrature on adaptive sparse grids. *Computing*, 71:89–114, 2003.

[7] P.P. Butzer and H. Berens. *Semi-groups of operators and approximation.* Springer, New York, 1967.

[8] E.W. Cheney. *Introduction to approximation theory, 2nd ed.* AMS publsihing, Providence, 1982.

[9] F. Deutsch. *Best approximation in inner product spaces.* Springer, New York, 2001.

[10] R.A. DeVore and G.G. Lorentz. *Constructive Approximation.* Springer, Berlin, 1993.

[11] N.V. Efimov and S.B. Stechkin. Support properties of sets in Banach spaces and Chebyshev sets. *Dokl. Akad. Nauk SSSR*, 127:254–257, 1959.

[12] R. Holmes. *A Course on Optimization and Best Approximation.* Springer, 1971.

[13] V. Klee. Convexity of Chebyshev sets. *Math. Ann.*, 142:292–304, 1961.

[14] V. Klee. Remarks on nearest points in normed linear spaces. In *Proc. Colloquium on Convexity*, pages 168–176. Univ. Mat. Inst. Copenhagen, 1967.

[15] P. Kosmol and D. Müller-Wichards. *Optimization in function spaces:: with stability considerations in Orlicz spaces.* de Gruyter, Berlin, 2011.

[16] M. Kritikos. Sur quelques propriéetées des ensembles convexes. *Bull. Math. de la Soc. Roumaine des sciences*, 40:87–92, 1938.

[17] J. Li. The metric projection and its applications to solving variational inequalities in Banach spaces. *Fixed point Theory*, 5:285–298, 2004.

[18] H.N. Mhaskar and D.V. Pai. *Fundamentals of approximation theory.* CRC Press, Boca Raton, 2000.

[19] T.S. Motzkin. Sur quelques proprietes characteritiques des ensembles convexes. *Rend. Acad. dei Lincei (Roma)*, 21:562–567, 1935.

[20] M.J.D. Powell. *Approximation theory and methods.* Cambridge IUniversity Press, New York, 1981.

[21] I. Singer. Some open problems on best approximations, 1967. Seminaire Choquet.

[22] I. Singer. *The theory of best approximation and functional analysis*, volume 13 of *Series in applied mathematics.* SIAM, Philadelphia, 1974.

[23] I. Vlasov. Chebyshev sets in Banach spaces. *Soviet Math. Dokl.*, pages 1373–1374, 2.