Metric projection, nonexpansive mappings and proximal methods

J. Baumeister[1]

October 18, 2016

[1]Dies sind Aufzeichnungen, die kritisch zu lesen sind, da sie noch nicht endgültig korrigiert sind, und daher auch nicht zitierfähig sind (**Not for quotation without permission of the author**). Hinweise auf Fehler und Verbesserungsvorschläge an *baumeister@math.uni-frankfurt.de*

Preface

Let (X, d) be a metric space (see below for the terminology) and let $F : X \longrightarrow X$ be a mapping. We say that F is Lipschitz continuous with Lipschitz constant $c \geq 0$ when

$$d(F(x), F(y)) \leq c d(x, y) \text{ for all } x, y \in X. \tag{1}$$

If (1) is satisfied with $c < 1$, then F is called a **contraction**. If If (1) is satisfied with $c = 1$, then F is called **nonexpansive**. If we have

$$d(F(x), F(y)) \geq c d(x, y) \text{ for all } x, y \in X. \tag{2}$$

with $c > 1$, then F is called **expanding**.

A point $z \in X$ with $F(z) = z$ is called a **fixed point** of F. We denote by $\mathrm{Fix}_X(F)$ the set of fixed points of F in X.

This monograph is mainly devoted to the study of nonexpansive mappings. In the focus is the question under which assumptions F has a fixed point and under which conditions the **orbit** $(F^n(x))_{n\in\mathbb{N}}$ defined by the successive approximation

$$x^{n+1} := F(x^n) = F^n(x^0)\,,\ n \in \mathbb{N}_0, x^0 := x\,, \tag{3}$$

converges to a fixed point. As we know, in the case of a contraction this question has a simple answer: the metric space has to be complete. In contrast to the situation of contractions, the theory concerning the existence of fixed points and the convergence of successive approximation in the case of nonexpansive mappings is rather extensive. We discuss the questions in the framework of Hilbert and Banach spaces only. For the most part of the considerations we present areas of current and active research.

In the main focus are realizations of the successive approximation when F is a composition of nonexpansive mappings $F_1, \dots, F_m : F = F_m \circ \cdots \circ F_1$. Then the iteration (1.4) may be adapted to this situation in the following way:

$$x^{n+1} := F_i(x^n)\,,\ n \in \mathbb{N}_0, x^0 := x\,,\ \text{ where } i = n \mod m + 1\,. \tag{4}$$

We call this method the **method of cyclic approximation (MCA)**. If $m = 2$ the method is called the **method of alternate approximation (MAA)**. In the analysis of these methods a subclass of nonexpansive mappings plays an outstanding role, namely the class of firmly nonexpansive mappings. These are nonexpansive mappings which have the property that their finite composition is a firmly nonexpansive mapping too.

Most of our considertions are devoted to metric projection operators. These are operators which assign to each point of a Banach space the best approximation in a subset of the space. Such mappings have nice properties in Hilbert spaces but are their study in Banach spaces is somewhat delicate.

The first chapter is devoted to the fixed point approach. We sketch the importance of the famous fixed point theorems of Banach, Brouwer and Schauder and give some hints to the algorithmic implementation of the successive approximation in the context of these fixed point theorems.

In the second chapter we discuss metric projections which describe the best approximation of points by elements of a convex set. Metric projections in Hilbert spaces are well studied examples of firmly nonexpansive mappings. These considerations are related to early contributions to the theory of approximation theory and successive projections. In the focus are the questions concerning existence, uniqueness, characterization of best approximations and the continuity of the metric projection.

The third chapter is devoted to the presentation of problems which may formulated nd analyzed as best approximation problems. We study the feasibility problem, matrix problems, variational inequalities, the inversion of radiographs and the recovery of signals and some other inverse problems.

The results of the first chapter are used in the fourth chapter to analyze the metric projection methods. These are methods which compute a metric projection of a „complicated set“ by an iteration of metric projections of simpler sets. We consider the **method of alternate projection (MAP)** and the **method of cyclic projection (MCP)**. In the focus is the convergence of the iteration process. Several specific cases and various variants of the iteration methods are presented.

The last chapter is devoted to the proximal point algorithm. This is an algorithmic approach for the computation of the the minimizers of a convex function. Here, the resolvents of maximal monotone operators – the subfifferential of a convex function is a such a maximal monotone operator – which are nonexpansive mappings play an important role. To present this a methods convex analysis plays an important role.

The monograph has been prepared for a lecture on *Nonexpansive operators* at the departement of Computer Sciences and Mathematics. It benefits from a huge list of papers. It is not possible to present a complete bibliography concerning the different subjects. Instead, we prefer to mention just a few books and monographs of review character.

Sources for results and techiques in functional analysis are [1, 12, 13, 17]. References for convex analysis are [3, 5, 10, 15, 16]. The subject *metric projection* is studied in [4, 7, 9]. Monographs concerning the fixed point theory are [2, 6, 8, 11]. Proximal point algorithms are studied in [14].

Here are **notations and preliminaries** which are sufficient to start our considerations in the next chapter. **Troughout the monograph we implicetely assume that the vector spaces considered are real vector spaces and have dimensions at least one.**

A **metric space** is a nonempty set X endowed with a metric d; the pair (X, d) is then called a metric space. With respect to this metric convergence of sequences and their limits are well defined. The space (X, d) is called complete when each Cauchy sequence converges in X. The closure of a subset S of X is denoted by $\overline{\mathsf{S}}$ and we call a subset S of

X closed if $\overline{S} = S$ holds. The **diameter** of a set $S \subset X$ is given as

$$\mathrm{diam}(S) := \sup_{x,y \in S} d(x,y) .$$

A **normed space** $\mathcal{X}$ is a real vector space – the null vector is denoted by θ – endowed with a norm $\|\cdot\|$. Clearly, the pair $(\mathcal{X}, \|\cdot\|)$ becomes a metric space when we define the metric d as $d(u,v) := \|u-v\|$. If the resulting metric space is complete we call the space $\mathcal{X}$ a **Banach space**. In a normed space $\mathcal{X}$ we have with $x \in \mathcal{X}, r \geq 0$:

$$\begin{array}{rcl} B_r(x) & := & B_{1,\mathcal{X}}(x) := \{u \in \mathcal{X} : \|u-x\| < r\},\ B_r \ := \ B_r(\theta) \\ \overline{B}_r(x) & := & \overline{B}_{1,\mathcal{X}}(x) := \{u \in \mathcal{X} : \|u-x\| \leq r\},\ \overline{B}_r \ := \ \overline{B}_r(\theta) \end{array}$$

Let $\overline{S}_1 := \overline{S}_{1,\mathcal{X}} := \{x \in \mathcal{X} : \|x\| = 1\}$ denote the unit sphere in a Banach space $\mathcal{X}$.

Without further mention, all vector spaces will be assumed non-trivial, i.e. contain more than just the zero vector. A normed space $\mathcal{X}$ is **separable** iff there exists a countable subset S which is dense in $\mathcal{X}$. We know that $\mathcal{X}$ is separable if $\mathcal{X}^*$ is separable. The converse is true if $\mathcal{X}$ is reflexive.

A **pre-Hilbert space** $\mathcal{H}$ is a real vector space endowed with an inner product $\langle\cdot|\cdot\rangle$. This inner product generates a norm $\|\cdot\|$ in $\mathcal{H}$ via $\|x\| := \sqrt{\langle x|x\rangle}$. If the space $\mathcal{H}$ is complete with this associated norm we call the space $\mathcal{H}$ a **Hilbert space**. It is easy to prove that the so called **parallelogram identity** holds in a pre-Hilbert space $\mathcal{H}$:

$$\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2 ,\ u,v \in \mathcal{H} \tag{5}$$

Another tool in Hilbert spaces is the **Cauchy-Schwarz inequality**:

$$|\langle x|y\rangle| \leq \|x\|\|y\| \text{ for all } x,y \in \mathcal{H} \tag{6}$$

(5), (6) have many nice geometric and analytic consequences. Especially, the Cauchy-Schwarz inequality allows us to introduce the concept of orthogonality: $x,y \in \mathcal{H}$ are called **orthogonal** if $\langle x|y\rangle = 0$ holds. If S is a subset of $\mathcal{H}$ we call $S^\perp := \{x \in \mathcal{H} : \langle x|u\rangle = 0 \text{ for all } u \in S\}$ the orthogonal complement of S. If S is a subspace of $\mathcal{H}$ then $S^\perp$ is a closed subspace. If S is a closed subspace then we have $\mathcal{H} = S \oplus S^\perp$.

In a vector space X convexity of a subset C is defined as follows: C is called **convex** if the following property holds:

$$[u,v] := \{z \in X : z = tu + (1-t)v, u,v \in C, t \in [0,1]\} \subset C \text{ for all } u,v \in X .$$

$[u,v]$ is called a **segment** joining u,v. The balls $B_r(x), \overline{B}_r(x)$ introduced above are convex sets. If a set $C \subset X$ is not convex we may consider the **convex hull**

$$\mathrm{co}(C) := \bigcap\{K \subset X | K \text{ convex}, C \subset K\}$$

The **closed convex hull** $\overline{\mathrm{co}}(C)$ of C is the closure of $\mathrm{co}(C)$. A point $x \in X$ is called an **extremal point** of a convex set $A \subset X$ if and only if $x \in [u,v]$ with $u,v \in A, u \neq v$, implies $x = u$ or $x = v$.

If $\mathcal{X}, \mathcal{Y}$ are normed spaces we may consider the space $\mathcal{B}(\mathcal{X},\mathcal{Y})$ of all linaer continuous mappings from $\mathcal{X} \longrightarrow \mathcal{Y}$. We set $\mathcal{B}(\mathcal{X}) := \mathcal{B}(\mathcal{X},\mathcal{Y})$. $\mathcal{B}(\mathcal{X},\mathcal{Y})$ is a normed spaced when it is endowed with the norm

$$\|T\| := \max\{\|T(x)\| : x \in \overline{X}_1\} .$$

If $\mathcal{X}, \mathcal{Y}$ are Banach spaces $\mathcal{B}(\mathcal{X}, \mathcal{Y})$ is a Banach space too. With I we denote the identity map between spaces. If $\mathsf{T} \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ we set

$$\begin{aligned} \ker(\mathsf{T}) &:= \{x \in \mathcal{X} : \mathsf{T}x = \theta\} \ \ (\textbf{nullspace/kernel of } \mathsf{T}) \\ \operatorname{ran}(\mathsf{T}) &:= \{y \in \mathcal{Y} : \mathsf{T}x = y \text{ for some } x \in \mathcal{X}\} \ \ (\textbf{range/image of } \mathsf{T}) \end{aligned}$$

The **dual space** $\mathcal{X}^*$ of a normed space $\mathcal{X}$ is defined as follows:

$$\mathcal{X}^* := \mathcal{B}(\mathcal{X}, \mathbb{R}) = \{\lambda : \mathcal{X} \longrightarrow \mathbb{R} : \lambda \text{ linear and continuous}\}.$$

The action of a functional $\lambda \in \mathcal{X}^*$ on the space $\mathcal{X}$ is called the **canonical pairing**. We use the following notation:

$$\lambda : \mathcal{X} \ni x \longmapsto \langle \lambda, x \rangle \in \mathbb{R}.$$

$\mathcal{X}^*$ is a normed space with respect to the norm

$$\| \cdot \|_* : \mathcal{X} \ni \lambda \longmapsto \sup\{\langle \lambda, x \rangle : x \in \overline{B}_1\} \in \mathbb{R}.$$

Actually, $\mathcal{X}^*$ is a Banach space. It is a result of the Hahn-Banach-Theorem that the norm in $\mathcal{X}$ can be regained by the dual space in the following sense:

$$\|x\| = \max_{\lambda \in \overline{B}_1} \langle \lambda, x \rangle, \ x \in \mathcal{X}.$$

The **bidual space** $\mathcal{X}^{**}$ of a normed space $\mathcal{X}$ is the dual space of $\mathcal{X}^*$. Clearly, each element x in $\mathcal{X}$ defines an element $\mu_x \in \mathcal{X}^{**}$ as follows:

$$\langle \mu_x, \lambda \rangle := \langle \lambda, x \rangle, \ \lambda \in \mathcal{X}^*.$$

If the mapping $\mathcal{X} \ni x \longmapsto \mu_x \in \mathcal{X}^{**}$ is surjective, then the space $\mathcal{X}$ is called **reflexive**. A necessary condition for reflexivity is the completeness of $\mathcal{X}$. A Banach space is reflexive if and only if its dual space is reflexive (see [17]). Another classification of reflexivity is the following one: A space $\mathcal{X}$ is reflexive if and only if the norm of each functional λ in $\mathcal{X}^*$ can be regained by $\|\lambda\| = \max_{x \in \mathcal{X}} \langle \lambda, x \rangle$ (Theorem of James).

Besides the norm-convergence we have in a Banach space $\mathcal{X}$ another type of convergence, namely the **weak convergence**. A sequence $(x^n)_{n \in \mathbb{N}}$ **converges weakly** to $x \in \mathcal{X}$ iff $\lim_n \langle \lambda, x^n \rangle = \langle \lambda, x \rangle$ for all $\lambda \in X^*$; we write $x = w - \lim_n x^n$. Due to the Hahn-Banach theorem the limit x is uniquely determined. It is known that the norm in $\mathcal{X}$ is weakly lower semicontinuous, i.e. $\|x\| \le \liminf_n \|x^n\|$ if $(x^n)_{n \in \mathbb{N}}$ converges weakly to x. An important result is that in a reflexive Banach space $\mathcal{X}$ the unit ball $\overline{B}_1$ is a **weakly compact** subset of $\mathcal{X}$ which means that every sequence in $\overline{B}_1$ has a weakly convergent subsequence.

As we know from the Riesz-representation theorem, the dual space $\mathcal{H}^*$ of a Hilbert space $\mathcal{H}$ may be isometrically identified with the space $\mathcal{H}$ by the so called **Riesz-isomorphism** $R_{\mathcal{H}}$:

$$R_{\mathcal{H}} : \mathcal{H} \ni x \longmapsto \lambda_x \in \mathcal{H}, \ \langle \lambda_x, u \rangle := \langle x | u \rangle, u \in \mathcal{H}.$$

A simple consequence of this fact is that every Hilbert space is reflexive.
In the background of this result is the fact that each closed subspace U in a Hilbert space $\mathcal{H}$ has an orthogonal complement $U^\perp$ So, we have $\mathcal{H} = U \oplus U^\perp$. Associated to this

decomposition of $\mathcal{H}$ are the projections P_{U} and $\mathsf{P}_{\mathsf{U}^\perp}$. We shall consider these projections in a vera detailed manner.

Throughout the monograph we assume that the spaces used (metric spaces, normed spaces, inner product spaces) are complete. This is an assumption which is not necessary for the most part of the results. But for the very important theorems one cannot go without this assumption. Therefore we state all our result under the completeness assumption.

Frankfurt, im Oktober 2016 Johann Baumeister

Bibliography

[1] J. Baumeister. Funktionalanalysis, 2013. Skriptum einer Vorlesung, Goethe–Universität Frankfurt/Main.

[2] J. Baumeister. Nichtlineare Funktionalanalysis, 2013. Skriptum WiSe 2013, Goethe–Universität Frankfurt/Main.

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

[4] H.H. Bauschke, E. Matouskova, and S. Reich. Projection and proximal point methods: convergence results and counterexamples. *Nonlinear analysis*, 56:715–738, 2004.

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

[6] A. Cegielski. *Iterative methods for fixed point problems in Hilbert spaces.* Springer, New York, 2007.

[7] Y. Censor and A. Cegielski. Projection methods: an annotated bibliography of books and reviews, 2014. arXiv:1406.6143v2.

[8] K. Deimling. *Nonlinear Functional analysis.* Springer, Berlin, 1985.

[9] K. Goebel and W.A. Kirk. *Topics in metric fixed point theory.* Cambridge University Press, Cambridge, 1990.

[10] R. Holmes. *Geometric functional analysis and its applications.* Springer, 1975.

[11] V.I.. Istratescu. *Fixed point theorems: an introduction.* Reidel, Dordrecht, 1988.

[12] W. Kaballo. *Grundkurs Funktionalanalysis.* Spektrum/Akademischer Verlag, Heidelberg, 2010.

[13] L.V. Kantorovich and G.P. Akilov. *Functional analysis in normed spaces.* Pergamon, New York, 1964.

[14] N. Parikh and S. Boyd. Proximal algorithms. *Foundations of Trends in Optimization*, 1:123–231, 2013.

[15] R.R. Phelps. *Convex functions, monotone operators and differentiability.* Springer, Berlin, 1993.

[16] R.T. Rockafellar. *Convex analysis.* Princeton University press, Princeton, 1970.

[17] D. Werner. *Funktionalanalysis.* Springer, 2002.