J. Math. Anal. Appl. 421 (2015) 747–753

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

The alternating algorithm in a uniformly convex and uniformly smooth Banach space

Allan Pinkus

Department of Mathematics, Technion, Haifa, Israel

ARTICLE INFO

Article history:
Received 2 April 2014
Available online 2 July 2014
Submitted by P. Nevai

Keywords:
Alternating algorithm
Uniformly convex
Uniformly smooth

ABSTRACT

Let X be a uniformly convex and uniformly smooth Banach space. Assume that the M_i, $i = 1, \dots, r$, are closed linear subspaces of X, P_{M_i} is the best approximation operator to the linear subspace M_i, and $M := M_1 + \cdots + M_r$. We prove that if M is closed, then the alternating algorithm given by repeated iterations of

$$(I - P_{M_r})(I - P_{M_{r-1}}) \cdots (I - P_{M_1})$$

applied to any $x \in X$ converges to $x - P_M x$, where P_M is the best approximation operator to the linear subspace M. This result, in the case $r = 2$, was proven in Deutsch [4].

1. Introduction

In any uniformly convex Banach space X each closed convex set C is Chebyshev, see e.g., Cheney [3, p. 22]. That is, to each $x \in X$ there exists a unique best approximation from C. Let $P_C x$ denote this best approximation. Thus $P_C x \in C$ and

$$\|x - P_C x\| \le \|x - y\|$$

for all $y \in C$. Assume that the M_i, $i = 1, \dots, r$, are closed linear subspaces of the uniformly convex and uniformly smooth Banach space X. The question we consider is how and when we might find a best approximation to $x \in X$ from

$$M := M_1 + \cdots + M_r$$

in the space X based on the knowledge of these P_{M_i}.

E-mail address: pinkus@technion.ac.il.

http://dx.doi.org/10.1016/j.jmaa.2014.06.076
0022-247X/© 2014 Elsevier Inc. All rights reserved.

Historically the first method of this form that was studied seems to have been the alternating algorithm. This algorithm goes under various names in different settings. Particular variants have been called, among other things, the von Neumann Alternating Algorithm, the Cyclic Coordinate Algorithm, the Cyclic Projection Algorithm, the Schwarz Domain Decomposition method, and the Diliberto–Straus Algorithm.

The basic idea is the following. We start with $x_1 := x \in X$. We then find a best approximation $m_1 = P_{M_1}x_1$ to x_1 from M_1, and set $x_2 := x_1 - m_1$. We then find a best approximation $m_2 = P_{M_2}x_2$ to x_2 from M_2, and set $x_3 := x_2 - m_2$, then find a best approximation $m_3 = P_3x_3$ to x_3 from M_3, etc., and after cycling through all the subspaces M_j, $j = 1, \dots, r$, we then start again, i.e., after finding a best approximation from M_r we then go to M_1.

More precisely, set

$$E := (I - P_{M_r})(I - P_{M_{r-1}}) \cdots (I - P_{M_1}).$$

Thus, for each $x \in X$,

$$Ex = x - m_1 - \cdots - m_r,$$

where m_i is a best approximation to $x - m_1 - \cdots - m_{i-1}$ from M_i, $i = 1, \dots, r$, i.e., $m_i = P_{M_i}(x - m_1 - \cdots - m_{i-1})$. Now consider

$$\lim_{s\to\infty} E^s x.$$

Note that $E^s x = x - y_s$ for some $y_s \in M$. The hope is that this scheme will converge, and converge to $x - P_M x$.

In the Hilbert space setting von Neumann already in 1933 (see von Neumann [7]) showed the desired convergence of the above-mentioned alternating algorithm in the case of two subspaces, without demanding closure of the sum. This was extended to any finite number of subspaces in Halperin [6]. See Deutsch [4,5] for a discussion of this method. There are, by now, numerous different proofs of this result, and many additional algorithms based on the knowledge of these $\{P_{M_i}\}_{i=1}^r$. As an example, in any Hilbert space iterations of the operator $E = I - \sum_{i=1}^r \mu_i P_{M_i}$, where $\mu_i \in (0, 2)$ and $\sum_{i=1}^r \mu_i < 2$, applied to x always converge to $x - P_M x$. See Bauschke, Borwein [1] for additional examples. The linearity and orthogonality properties of the P_{M_i} in Hilbert space are both relevant and expedient. However in the non-Hilbert space setting, where the P_{M_i} are not linear operators, there seem to be very few results, and many of these are rather specialized.

In this note we prove that if X is a uniformly convex and uniformly smooth Banach space, the M_i are closed linear subspaces, and M is also closed, then this classic alternating algorithm necessarily converges as desired. (Note that the closure of the M_i does not imply the closure of M.)

This result is not valid in every normed linear space, see e.g., Deutsch [4]. In any normed linear space X which is not smooth, one can construct two linear subspaces M_1 and M_2, and an $x \in X$ for which $P_{M_i}x = 0$, $i = 1, 2$, and yet the zero element is not a best approximation to x from $M_1 + M_2$.

2. Main result

In the proof of our main result we assume that both the M_i, $i = 1, \dots, r$, and M are closed. This closure finds its expression in the following fundamental result, see Browder [2] (also in Bauschke, Borwein [1]).

Lemma 2.1. *Let X be a Banach space. Assume*

$$M = M_1 + \cdots + M_r$$

where M, $M_1, \dots, M_r$ are closed linear subspaces of X. Then there exists a $\kappa > 0$ such that each $m \in M$ has a representation of the form

$$m = m_1 + \cdots + m_r$$

where $m_j \in M_j$, $j = 1, \dots, r$, and

$$\|m_1\| + \cdots + \|m_r\| \le \kappa \|m\|.$$

Proof. Introduce the norm

$$\big\|(m_1, \dots, m_r)\big\| = \|m_1\| + \cdots + \|m_r\|$$

on $M_1 \times \cdots \times M_r$. Let

$$L : M_1 \times \cdots \times M_r \to M_1 + \cdots + M_r = M$$

be the linear map given by

$$L(m_1, \dots, m_r) = m_1 + \cdots + m_r.$$

This map is continuous and onto M. Since M is complete, it follows from the Open Mapping Theorem that there exists a $\kappa > 0$ such that for each $m \in M$ there exists $(m_1, \dots, m_r) \in M_1 \times \cdots \times M_r$ such that $L(m_1, \dots, m_r) = m$ and $\|(m_1, \dots, m_r)\| \le \kappa \|m\|$, i.e.,

$$\|m_1\| + \cdots + \|m_r\| \le \kappa \|m\|. \qquad \square$$

The proof of the following result in the case $r = 2$ is due to Deutsch [4]. If the P_{M_i} are linear, which is rare indeed, then this result, for any finite r, without the demand of the closure of M, and for a smooth uniformly convex X, is in Reich [8].

Theorem 2.2. *Let X be a uniformly convex and uniformly smooth Banach space, and let M_i, $i = 1, \dots, r$, be closed linear subspaces. If $M := M_1 + \cdots + M_r$ is closed, then*

$$\lim_{n\to\infty} \big[(I - P_{M_r})(I - P_{M_{r-1}}) \cdots (I - P_{M_1})\big]^n x = x - P_M x$$

for each $x \in X$.

Proof. For notational ease, set $P_j = P_{M_j}$ and $P = P_M$. For every P_j as above, we always have

$$\big\|(I - P_j)x\big\| \le \|x\|. \tag{2.1}$$

Set

$$H^{sr+j} = (I - P_j) \cdots (I - P_1) E^s$$

for $j \in \{1, \dots, r\}$ and $s \in \mathbb{Z}_+$, the set of nonnegative integers. Note that $H^{sr+r} = E^{s+1}$. Now

$$H^k x = x - y_k =: x_k$$

where $y_k \in M$. Furthermore, from (2.1)

$$\|x_k\| \le \|x_{k-1}\| \tag{2.2}$$

for all k. As such

$$\lim_{k\to\infty} \|x_k\|$$

exists. If

$$\lim_{k\to\infty} \|x_k\| = 0,$$

then there is nothing to prove since this implies that $x \in M$, by our assumption of closure, and y_k converges to x. As such, we assume that

$$\lim_{k\to\infty} \|x_k\| = C > 0.$$

For each $k \in \mathbb{N}$, let ϕ_k be the continuous linear functional on X (as X is smooth, ϕ_k is uniquely defined) satisfying

(a) $\|\phi_k\| = \|x_k\|$,
(b) $\phi_k(x_k) = \|x_k\|^2$.

For $k = sr+j$, $j \in \{1, \dots, r\}$, we also have from the best approximation property and the definition of H^{sr+j} that

(c) $\phi_{sr+j}(m_j) = 0$ all $m_j \in M_j$, $j \in \{1, \dots, r\}$.

For the existence of these $\{\phi_k\}$, see e.g., Singer [9, p. 18]. Note that

$$x_{sr+j} - x_{sr+j-1} \in M_j,$$

and thus

$$\phi_{sr+j}(x_{sr+j-1}) = \phi_{sr+j}(x_{sr+j}) = \|x_{sr+j}\|^2.$$

We first claim that for every $j, k \in \{1, \dots, r\}$ we have

$$\lim_{s\to\infty} \|\phi_{sr+j} - \phi_{sr+k}\| = 0.$$

It obviously suffices to prove that

$$\lim_{s\to\infty} \|\phi_{sr+j} - \phi_{sr+j-1}\| = 0,$$

for $j \in \{2, \dots, r\}$. Now

$$\begin{aligned}
\frac{\|\phi_{sr+j} + \phi_{sr+j-1}\|}{2} &\ge \frac{(\phi_{sr+j} + \phi_{sr+j-1})(x_{sr+j-1})}{2\|x_{sr+j-1}\|} \\
&= \frac{\phi_{sr+j}(x_{sr+j-1}) + \phi_{sr+j-1}(x_{sr+j-1})}{2\|x_{sr+j-1}\|} \\
&= \frac{\|x_{sr+j}\|^2 + \|x_{sr+j-1}\|^2}{2\|x_{sr+j-1}\|}
\end{aligned}$$

$$\begin{aligned} &\geq \|x_{sr+j}\| \\ &= \|\phi_{sr+j}\|. \end{aligned}$$

Thus

$$\frac{\|\phi_{sr+j} + \phi_{sr+j-1}\|}{2} \geq \|\phi_{sr+j}\|. \tag{2.3}$$

Furthermore

$$\lim_{s\to\infty} \|\phi_{sr+j}\| = \lim_{s\to\infty} \|x_{sr+j}\| = C > 0.$$

As X is uniformly smooth, then X^* is uniformly convex. By this is meant that the *modulus of uniform convexity* defined by

$$\delta(\varepsilon) = \inf\left\{1 - \frac{\|f+g\|}{2} : f, g \in X^*,\ \|f\| = \|g\| = 1,\ \|f-g\| \geq \varepsilon\right\}$$

is strictly positive for $\varepsilon > 0$. ($\delta(\varepsilon)$ is, by definition, an increasing function of $\varepsilon \in (0, 2]$ that decreases to zero as ε decreases to zero.)

Set

$$f_s = \frac{\phi_{sr+j}}{\|\phi_{sr+j-1}\|}, \qquad g_s = \frac{\phi_{sr+j-1}}{\|\phi_{sr+j-1}\|}$$

and $\varepsilon = \|f_s - g_s\|$. Since $\|f_s\| \leq \|g_s\| = 1$, we have

$$\delta(\varepsilon) \leq 1 - \frac{\|f_s + g_s\|}{2}$$

which, when substituting for ε, f_s and g_s, gives us

$$\|\phi_{sr+j-1}\|\,\delta\left(\frac{\|\phi_{sr+j} - \phi_{sr+j-1}\|}{\|\phi_{sr+j-1}\|}\right) \leq \|\phi_{sr+j-1}\| - \frac{\|\phi_{sr+j} + \phi_{sr+j-1}\|}{2}.$$

From (2.3) we obtain

$$\|\phi_{sr+j-1}\|\,\delta\left(\frac{\|\phi_{sr+j} - \phi_{sr+j-1}\|}{\|\phi_{sr+j-1}\|}\right) \leq \|\phi_{sr+j-1}\| - \|\phi_{sr+j}\|.$$

From (2.2) and (a) the right-hand side tends to zero as s tends to ∞. As $\|\phi_{sr+j-1}\|$ is bounded away from zero we must have, by the definition of the modulus of convexity,

$$\lim_{s\to\infty} \|\phi_{sr+j} - \phi_{sr+j-1}\| = 0.$$

We therefore obtain

$$\lim_{s\to\infty} \|\phi_{sr+j} - \phi_{sr+k}\| = 0$$

for every $j, k \in \{1, \ldots, r\}$.

We recall that

$$H^{sr+j}x = x_{sr+j} = x - y_{sr+j}.$$

We can write

$$x - x_{sr+j} = y_{sr+j} = \widetilde{m}_1 + \cdots + \widetilde{m}_r$$

for some $\widetilde{m}_\ell \in M_\ell$, $\ell = 1, \ldots, r$ (these elements also depend upon s and j) where, by Lemma 2.1, for some fixed $\kappa > 0$,

$$\|\widetilde{m}_1\| + \cdots + \|\widetilde{m}_r\| \le \kappa \|y_{sr+j}\| \le \kappa\big[\|x\| + \|x_{sr+j}\|\big] \le 2\kappa\|x\|.$$

Let y be any element of M satisfying $\|x - y\| \le \|x\|$. Set

$$y = m_1 + \cdots + m_r$$

where $m_\ell \in M_\ell$, $\ell = 1, \ldots, r$, and

$$\|m_1\| + \cdots + \|m_r\| \le \kappa\|y\| \le 2\kappa\|x\|,$$

for $\kappa > 0$ as in Lemma 2.1. Now, for y as above,

$$\begin{aligned}\|x_{sr+j}\|^2 &= \phi_{sr+j}(x_{sr+j}) = \phi_{sr+j}(x - y_{sr+j}) = \phi_{sr+j}\left(x - \sum_{k=1}^r \widetilde{m}_k\right)\\ &= \phi_{sr+j}(x-y) + \phi_{sr+j}\left(y - \sum_{k=1}^r \widetilde{m}_k\right) = \phi_{sr+j}(x-y) + \phi_{sr+j}\left(\sum_{k=1}^r (m_k - \widetilde{m}_k)\right).\end{aligned}$$

Since $\phi_{sr+k}(m_k) = 0$ for all $m_k \in M_k$ we have

$$\begin{aligned}\|x_{sr+j}\|^2 &= \phi_{sr+j}(x-y) + \sum_{k=1}^r (\phi_{sr+j} - \phi_{sr+k})(m_k - \widetilde{m}_k)\\ &\le \|\phi_{sr+j}\|\,\|x-y\| + \sum_{k=1}^r \|\phi_{sr+j} - \phi_{sr+k}\|\,\|m_k - \widetilde{m}_k\|\\ &\le \|\phi_{sr+j}\|\,\|x-y\| + \sum_{k=1}^r \|\phi_{sr+j} - \phi_{sr+k}\|\,4\kappa\|x\|\\ &= \|x_{sr+j}\|\,\|x-y\| + 4\kappa\|x\| \sum_{k=1}^r \|\phi_{sr+j} - \phi_{sr+k}\|.\end{aligned}$$

Let $s \to \infty$ and recall that we have

$$\lim_{s\to\infty} \|\phi_{sr+j} - \phi_{sr+k}\| = 0$$

for $j, k \in \{1, \ldots, r\}$. Furthermore

$$\lim_{s\to\infty} \|x_{sr+j}\| = C > 0.$$

Thus, we obtain

$$\lim_{s\to\infty} \|x_{sr+j}\| \le \|x - y\|$$

for every $y \in M$ satisfying $\|x - y\| \le \|x\|$, and therefore

$$\lim_{s\to\infty} \|x - y_{sr+j}\| \le \min_{y\in M} \|x - y\| = \|x - Px\|.$$

It easily follows from the definition of uniform convexity that y_{sr+j} must converge to Px in X as $s \to \infty$. This is valid for each $j \in \{1, \dots, r\}$. $\square$

If the M_i are all finite-dimensional linear subspaces, then M is closed and Theorem 2.2 can be more easily proven.

As was mentioned, in the Hilbert space setting convergence is obtained without the necessity of the closure of the sum. It is still an open question as to whether this closure property is necessary here. For more on this, see Deutsch [5, p. 234]. In addition, if we are in a Hilbert space, and the sum of the spaces is closed, then there exist constants C and θ, where $C > 0$, $0 \le \theta < 1$, such that

$$\|P_M x - y_n\| \le C\theta^n$$

for all $n \in \mathbb{N}$. In the optimization literature this is called *linear convergence*. Nothing is known in this setting regarding convergence rates.

Acknowledgment

The author would like to thank the referee for her/his input.

References

[1] H.H. Bauschke, J.M. Borwein, On projection algorithms for solving convex feasibility problems, SIAM Rev. 38 (1996) 367–426.
[2] F.E. Browder, Convergence theorems for sequences of nonlinear operators in Banach spaces, Math. Z. 100 (1967) 201–225.
[3] E.W. Cheney, Introduction to Approximation Theory, McGraw–Hill, New York, 1966.
[4] F. Deutsch, The alternating method of von Neumann, in: W. Schempp, K. Zeller (Eds.), Multivariate Approximation Theory, in: Internat. Ser. Numer. Math., vol. 51, Birkhäuser, Basel, 1979, pp. 83–96.
[5] F. Deutsch, Best Approximation in Inner Product Spaces, Springer-Verlag, New York, 2001.
[6] I. Halperin, The product of projection operators, Acta Sci. Math. (Szeged) 23 (1962) 96–99.
[7] J. von Neumann, Functional Operators, vol. II. The Geometry of Orthogonal Spaces, Ann. of Math. Stud., vol. 22, Princeton University Press, Princeton, NJ, 1950, this is a reprint of mimeographed lecture notes first distributed in 1933.
[8] S. Reich, Nonlinear semigroups, accretive operators, and applications, in: V. Lakshmikantham (Ed.), Nonlinear Phenomena in Mathematical Sciences, Academic Press, New York, 1982, pp. 831–838.
[9] I. Singer, Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces, Grundlehren Math. Wiss., vol. 171, Springer-Verlag, Berlin, Heidelberg, 1970.