ARTICLE IN PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Approximation Theory ■■■ (■■■■) ■■■–■■■

Journal of Approximation Theory

www.elsevier.com/locate/jat

A strongly convergent reflection method for finding the projection onto the intersection of two closed convex sets in a Hilbert space

Heinz H. Bauschke[a], Patrick L. Combettes[b,*], D. Russell Luke[c]

[a]Mathematics, Irving K. Barber School, UBC Okanagan, Kelowna, BC, Canada V1V 1V7
[b]Laboratoire Jacques-Louis Lions—UMR 7598, Université Pierre et Marie Curie—Paris 6, 75005 Paris, France
[c]Department of Mathematical Sciences, University of Delaware, Newark, Delaware 19716-2553, USA

Received 18 February 2005; accepted 11 January 2006

Communicated by Joseph Ward

Abstract

A new iterative method for finding the projection onto the intersection of two closed convex sets in a Hilbert space is presented. It is a Haugazeau-like modification of a recently proposed averaged alternating reflections method which produces a strongly convergent sequence.

Keywords: Best approximation problem; Convex set; Projection; Strong convergence

1. Introduction

Throughout this paper,

$$X \text{ is a real Hilbert space with inner product } \langle \cdot \mid \cdot \rangle \text{ and induced norm } \|\cdot\|, \tag{1}$$

and

$$A \text{ and } B \text{ are two closed convex sets in } X \text{ such that } C = A \cap B \neq \varnothing. \tag{2}$$

Given a point $x \in X$, the problem under consideration is the *best approximation problem*

$$\text{find } c \in C \text{ such that } \|x - c\| = \inf \|x - C\|. \tag{3}$$

* Corresponding author. Fax: +33 1 4427 7200.
E-mail addresses: heinz.bauschke@ubc.ca (H.H. Bauschke), plc@math.jussieu.fr (P.L. Combettes), rluke@math.udel.edu (D.R. Luke).

0021-9045/$ - see front matter
doi:10.1016/j.jat.2006.01.003

This problem, which was already studied by von Neumann in the 1930s in this general Hilbert space setting, is of fundamental importance in applied mathematics (see [5] for historical references, recent applications, algorithms, and further references).

The aim of this note is to present a new strongly convergent method—termed *Haugazeau-like Averaged Alternating Reflections* (HAAR)—for finding the solution of (3) iteratively. This algorithm is a modification of the *Averaged Alternating Reflections* (AAR) scheme, which we recently introduced in [4]. To describe AAR, we require some notation from convex analysis. Given any nonempty closed convex set S in X, denote the *projector* (best approximation operator) onto S by P_S. Furthermore, let I be the identity operator on X and let $R_S = 2P_S - I$ be the *reflector* with respect to S. We recall that the normal cone to S at $x \in S$ is defined by $N_S(x) = \big\{x^* \in X \mid (\forall s \in S)\ \ \langle x^* \mid s - x\rangle \leqslant 0\big\}$. Both AAR and HAAR rely upon the operator

$$T = \tfrac{1}{2} R_A R_B + \tfrac{1}{2} I, \tag{4}$$

and their analyses require the nonempty closed convex cone

$$K = N_{B-A}(0). \tag{5}$$

We are now ready to describe AAR and its asymptotic behavior (see also [4] for background).

Fact 1.1 (*AAR*). *Suppose that $x \in X$. Then the sequence of averaged alternating reflections* (*AAR*) $(T^n x)_{n\in\mathbb{N}}$ *converges weakly to a point in*

$$\operatorname{Fix} T = \big\{z \in X \mid Tz = z\big\} = C + K. \tag{6}$$

Moreover, the sequence $(P_B T^n x)_{n\in\mathbb{N}}$ is bounded and each of its weak cluster points lies in C.

Proof. Identity (6) was proved in [4, Corollary 3.9]. The statements regarding weak convergence and weak cluster points follow from [8, Theorem 1] applied to the normal cone operators N_A and N_B. (See also [3, Fact 5.9] and [4, Theorem 3.13(ii)].) □

Fact 1.1 implies that the weak cluster points of the sequence $(P_B T^n x)_{n\in\mathbb{N}}$ solve the *convex feasibility problem*

$$\text{find } c \in C. \tag{7}$$

Although such points solve (7), they may nonetheless be neither strong cluster points nor the solution of the best approximation problem (3) (see [4, Section 1] for a counterexample). These shortcomings of AAR motivated us to look for variants of AAR with better convergence properties. In Section 2, we investigate the relative geometry of the sets A and B, culminating in the formula $P_B P_{C+K} = P_C$ (see Corollary 2.9). This identity, Fact 1.1, and a consequence of the weak-to-strong convergence principle [2] lead in Section 3 to the precise formulation of HAAR. A crucial ingredient of HAAR is Haugazeau's [7] explicit projector onto the intersection of two halfspaces. Our main result (Theorem 3.3) guarantees strong convergence to the nearest point in C, i.e., to the solution of (3).

2. Relative geometry of two sets

We shall utilize the following notions from fixed point theory; see, e.g., [6].

Definition 2.1. Suppose that $R: X \to X$. Then:

(i) R is *firmly nonexpansive*, if

$$(\forall x \in X)(\forall y \in X) \quad \|Rx - Ry\|^2 + \|(I - R)x - (I - R)y\|^2 \leqslant \|x - y\|^2; \tag{8}$$

(ii) R is *nonexpansive*, if

$$(\forall x \in X)(\forall y \in X) \quad \|Rx - Ry\| \leqslant \|x - y\|. \tag{9}$$

It is well known, for example, that the projector onto a nonempty closed convex set is firmly nonexpansive.

Fact 2.2. *Suppose that $R: X \to X$. Then R is firmly nonexpansive if and only if $2R - I$ is nonexpansive.*

Proof. See [6, Theorem 12.1]. $\square$

Fact 2.3. *Suppose that S is a nonempty closed convex set in X and that $x \in X$. Then there exists a unique point $P_S x \in S$ such that $\|x - P_S x\| = \inf \|x - S\|$. The point $P_S x$ is characterized by*

$$P_S x \in S \quad \text{and} \quad (\forall s \in S) \quad \langle s - P_S x \mid x - P_S x\rangle \leqslant 0. \tag{10}$$

The induced operator $P_S: X \to S: x \mapsto P_S x$ is called the projector *onto S; it is firmly nonexpansive and consequently, the reflector $R_S = 2P_S - I$ is nonexpansive.*

The following property will be utilized repeatedly.

Fact 2.4. *Suppose that S is a nonempty closed convex set in X and that $z \in X$. Then for every $x \in X$, we have $P_{z+S}x = z + P_S(x - z)$.*

Proof. Use (10). $\square$

We record two additional auxiliary results.

Fact 2.5. *Suppose that U and V are two nonempty closed convex sets in X. Suppose furthermore that $u \in U$ and that $v \in V$. Then $N_{U+V}(u + v) = N_U(u) \cap N_V(v)$.*

Proof. See, e.g., [1, Section 4.6]. $\square$

Proposition 2.6. *Suppose that U and V are two nonempty closed convex sets in X such that $U \perp V$. Then $U + V$ is closed and $P_{U+V} = P_U + P_V$.*

Proof. Suppose that $(u_n)_{n\in\mathbb{N}}$ and $(v_n)_{n\in\mathbb{N}}$ are sequences in U and V, respectively, such that $(u_n + v_n)_{n\in\mathbb{N}}$ converges. For every $\{m, n\} \subset \mathbb{N}$, we have $\|(u_n + v_n) - (u_m + v_m)\|^2 = \|u_n - u_m\|^2 + \|v_n - v_m\|^2$. Hence $(u_n)_{n\in\mathbb{N}}$ and $(v_n)_{n\in\mathbb{N}}$ are both Cauchy sequences, since $(u_n + v_n)_{n\in\mathbb{N}}$ is. Thus $(u_n)_{n\in\mathbb{N}}$ and $(v_n)_{n\in\mathbb{N}}$ are both convergent, which implies that $\lim_{n\in\mathbb{N}} u_n + v_n \in U + V$.

Now let $x \in X$, $u \in U$, and $v \in V$. Since $\{u - P_U x, -P_U x\} \perp \{v - P_V x, -P_V x\}$, Fact 2.3 implies that

$$\begin{aligned} &\langle u + v - P_U x - P_V x \mid x - P_U x - P_V x\rangle \\ &\quad = \langle u - P_U x \mid x - P_U x\rangle + \langle u - P_U x \mid -P_V x\rangle \\ &\qquad + \langle v - P_V x \mid x - P_V x\rangle + \langle v - P_V x \mid -P_U x\rangle \\ &\quad = \langle u - P_U x \mid x - P_U x\rangle + \langle v - P_V x \mid x - P_V x\rangle \\ &\quad \leqslant 0. \end{aligned} \tag{11}$$

Using Fact 2.3 again, it follows that $P_{U+V} x = P_U x + P_V x$. □

Proposition 2.7. *Suppose that $c \in C$. Then $K = N_B(c) \cap \big(-N_A(c)\big) \subset (C - C)^\perp$.*

Proof. Using (5) and Fact 2.5, we deduce that

$$K = N_{B-A}(0) = N_{B+(-A)}\big(c + (-c)\big) = N_B(c) \cap N_{-A}(-c) = N_B(c) \cap \big(-N_A(c)\big). \tag{12}$$

Let $x \in K$. By (12), $\sup \langle x \mid B - c\rangle \leqslant 0$ and $\sup \langle -x \mid A - c\rangle \leqslant 0$. Since $C = A \cap B$, it follows that $\sup \langle x \mid C - c\rangle \leqslant 0$ and that $\sup \langle -x \mid C - c\rangle \leqslant 0$. Therefore, $x \in (C-c)^\perp = (C-C)^\perp$. □

Theorem 2.8. *Suppose that $x \in X$ and that $c \in C$. Then $P_{C+K} x = P_C x + P_K(x - c)$.*

Proof. Set $L = C - C$. Then $C - c \subset L$ and, by Proposition 2.7, $K \subset L^\perp$. Corollary 2.4 and Proposition 2.6 yield

$$\begin{aligned} P_{C+K} x &= P_{c+((C-c)+K)} x \\ &= c + P_{(C-c)+K}(x - c) \\ &= c + P_{C-c}(x - c) + P_K(x - c) \\ &= P_C x + P_K(x - c), \end{aligned} \tag{13}$$

which completes the proof. □

Corollary 2.9. *Suppose that $x \in X$. Then $P_B P_{C+K} x = P_C x$.*

Proof. Since $P_C x \in C$, Theorem 2.8 implies that $P_{C+K} x = P_C x + P_K(x - P_C x)$. Hence, using Proposition 2.7, we deduce that

$$P_{C+K} x - P_C x = P_K(x - P_C x) \in K \subset N_B(P_C x). \tag{14}$$

As $P_C x \in B$, this shows that $P_B P_{C+K} x = P_C x$. □

3. Main result

Definition 3.1. Suppose that $(x, y, z) \in X^3$ satisfies

$$\big\{w \in X \mid \langle w - y \mid x - y\rangle \leqslant 0\big\} \cap \big\{w \in X \mid \langle w - z \mid y - z\rangle \leqslant 0\big\} \neq \varnothing. \tag{15}$$

Set

$$\pi = \langle x - y \mid y - z\rangle, \quad \mu = \|x - y\|^2, \quad \nu = \|y - z\|^2, \quad \rho = \mu\nu - \pi^2, \tag{16}$$

and

$$Q(x, y, z) = \begin{cases} z, & \text{if } \rho = 0 \text{ and } \pi \geqslant 0; \\ x + (1 + \pi/\nu)(z - y), & \text{if } \rho > 0 \text{ and } \pi\nu \geqslant \rho; \\ y + (\nu/\rho)\big(\pi(x - y) + \mu(z - y)\big), & \text{if } \rho > 0 \text{ and } \pi\nu < \rho. \end{cases} \tag{17}$$

In [7], Haugazeau introduced the operator Q as an explicit description of the projector onto the intersection of the two halfspaces defined in (15). He proved in [7, Théorème 3 –1] that the sequence $(y_n)_{n\in\mathbb{N}}$ defined by $y_0 = x$ and

$$(\forall n \in \mathbb{N}) \quad y_{n+1} = Q\big(x, Q(x, y_n, P_B y_n), P_A Q(x, y_n, P_B y_n)\big) \tag{18}$$

converges strongly to $P_C x$. The next result is a particular application of the weak-to-strong convergence principle of [2], which will be used to reach the same conclusion for the proposed HAAR method.

Fact 3.2. *Suppose that $R\colon X \to X$ is nonexpansive and that* $\operatorname{Fix} R \neq \varnothing$. *Suppose furthermore that $x \in X$ and that $(\lambda_n)_{n\in\mathbb{N}}$ is a sequence in $\left]0, \frac{1}{2}\right]$ such that* $\inf_{n\in\mathbb{N}} \lambda_n > 0$. *Set $y_0 = x$ and define $(y_n)_{n\in\mathbb{N}}$ by*

$$(\forall n \in \mathbb{N}) \quad y_{n+1} = Q\big(x, y_n, (1 - \lambda_n) y_n + \lambda_n R y_n\big). \tag{19}$$

Then $(y_n)_{n\in\mathbb{N}}$ converges strongly to $P_{\operatorname{Fix} R} x$.

Proof. This follows from [2, Corollary 6.6(ii)]. □

We are now in a position to introduce HAAR and to establish its convergence properties.

Theorem 3.3 (*HAAR*). *Suppose that $x \in X$ and that $(\mu_n)_{n\in\mathbb{N}}$ is a sequence in* $]0, 1]$ *such that* $\inf_{n\in\mathbb{N}} \mu_n > 0$. *Define the sequence $(y_n)_{n\in\mathbb{N}}$ generated by Haugazeau-like averaged alternating reflections by $y_0 = x$ and*

$$(\forall n \in \mathbb{N}) \quad y_{n+1} = Q\big(x, y_n, (1 - \mu_n) y_n + \mu_n T y_n\big). \tag{20}$$

Then $(y_n)_{n\in\mathbb{N}}$ converges strongly to $P_{C+K} x$. Moreover, $(P_B y_n)_{n\in\mathbb{N}}$ converges strongly to $P_C x$.

Proof. Since the reflectors R_A and R_B are both nonexpansive (see Fact 2.3), so is their composition $R = R_A R_B$. Consequently, Fact 2.2 implies that T is firmly nonexpansive. Moreover, by Fact 1.1, $\operatorname{Fix} R = \operatorname{Fix}\big(\frac{1}{2} R + \frac{1}{2} I\big) = \operatorname{Fix} T = C + K$. The statement about strong convergence of $(y_n)_{n\in\mathbb{N}}$ follows from Fact 3.2 (with $\lambda_n = \mu_n/2$). Since $y_n \to P_{C+K} x$ and P_B is continuous, we further deduce that $(P_B y_n)_{n\in\mathbb{N}}$ converges strongly to $P_B P_{C+K} x$, which is equal to $P_C x$ by Corollary 2.9. □

Remark 3.4. Several comments on Theorem 3.3 are in order.

(i) While a detailed numerical study of HAAR lies outside the scope of this paper, we nonetheless briefly discuss a numerical example demonstrating the potential of HAAR. As in [4, Section 1] for AAR, we consider the case when $X = \mathbb{R}^2$, $A = \big\{(\xi_1, \xi_2) \in X \mid \xi_2 \leqslant 0\big\}$, and $B = \big\{(\xi_1, \xi_2) \in X \mid \xi_1 \leqslant \xi_2\big\}$. Let $x = (8, 4)$ so that $P_C x = (0, 0)$. Let $(y_n)_{n\in\mathbb{N}}$ be a sequence constructed as in Theorem 3.3 with $\mu_n \equiv 1$. Then $y_0 = x = (8, 4)$, $y_1 = (6, -2)$, and $y_n =$

$(0,0)$, for every $n \in \{2,3,\ldots\}$. Therefore, $P_B y_0 = (6,6)$, $P_B y_1 = (2,2)$, and $P_B y_n = (0,0)$, for every $n \in \{2,3,\ldots\}$. Thus HAAR converges to the solution $P_C x = (0,0)$ in just two steps. On the other hand, Dykstra's algorithm, which is a popular best approximation method (see, e.g., [5, Chapter 9]), requires infinitely many steps in this setting.

(ii) It is important to monitor the sequence $(P_B y_n)_{n\in\mathbb{N}}$ rather than $(y_n)_{n\in\mathbb{N}}$ in order to approximate $P_C x$. Indeed, let $A = B = \{0\}$ and $x \in X \setminus \{0\}$. Then $K = X$ and thus $(y_n)_{n\in\mathbb{N}}$ converges to $P_{C+K} x = P_X x = x$ but not to $P_C x = \{0\}$.

(iii) Theorem 3.3 can be utilized to handle best approximation problems with more than two sets. Suppose that $C_1, \ldots, C_J$ are finitely many closed convex sets in X such that

$$C = C_1 \cap \cdots \cap C_J \neq \varnothing. \tag{21}$$

As in our corresponding discussion for AAR in [4, Section 4], we employ Pierra's product space technique [9]. Let us take $(\omega_j)_{1\leqslant j\leqslant J}$ in $]0,1]$ such that $\sum_{j=1}^J \omega_j = 1$, and let us denote by $\mathbf{X}$ the Hilbert space X^J with the inner product $\big((x_j)_{1\leqslant j\leqslant J}, (y_j)_{1\leqslant j\leqslant J}\big) \mapsto \sum_{j=1}^J \omega_j \langle x_j, y_j\rangle$. Set

$$\mathbf{A} = \big\{(x,\ldots,x) \in \mathbf{X}: x \in X\big\} \quad \text{and} \quad \mathbf{B} = C_1 \times \cdots \times C_J, \tag{22}$$

and observe that the set $C = \bigcap_{j=1}^J C_j$ in X corresponds to the set $\mathbf{C} = \mathbf{A} \cap \mathbf{B}$ in $\mathbf{X}$. The projections of $\mathbf{x} = (x_j)_{1\leqslant j\leqslant J} \in \mathbf{X}$ onto $\mathbf{A}$ and $\mathbf{B}$ are given by

$$P_{\mathbf{A}}\mathbf{x} = \big(\textstyle\sum_{j=1}^J \omega_j x_j, \ldots, \sum_{j=1}^J \omega_j x_j\big) \quad \text{and} \quad P_{\mathbf{B}}\mathbf{x} = (P_{C_1} x_1, \ldots, P_{C_J} x_J), \tag{23}$$

respectively. Thus we have explicit formulae for $R_{\mathbf{A}} = 2P_{\mathbf{A}} - \mathbf{I}$ and $R_{\mathbf{B}} = 2P_{\mathbf{B}} - \mathbf{I}$, where $\mathbf{I}$ denotes the identity operator on $\mathbf{X}$. Let

$$\mathbf{T} = \tfrac{1}{2}(R_{\mathbf{A}} R_{\mathbf{B}} + \mathbf{I}), \tag{24}$$

let $x \in X$, and set $\mathbf{y}_0 = (x, x, \ldots, x) \in \mathbf{X}$. Define the sequence $(\mathbf{y}_n)_{n\in\mathbb{N}}$ recursively by

$$\mathbf{y}_{n+1} = \mathbf{Q}(\mathbf{y}_0, \mathbf{y}_n, \mathbf{T}\mathbf{y}_n), \tag{25}$$

where $\mathbf{Q}$ is defined on $\mathbf{X}^3$ in a manner analogous to Q on X^3 in Definition 3.1. Then Theorem 3.3 (with $\mu_n \equiv 1$) implies that $(P_{\mathbf{B}}\mathbf{y}_n)_{n\in\mathbb{N}}$ converges strongly to $P_{\mathbf{C}}\mathbf{y}_0 = (P_C x, \ldots, P_C x)$. Consequently, $(P_{\mathbf{A}} P_{\mathbf{B}} \mathbf{y}_n)_{n\in\mathbb{N}}$ converges strongly to $P_{\mathbf{C}}\mathbf{y}_0$ as well. Since this last sequence lies in $\mathbf{A}$, we identify it with some sequence $(a_n)_{n\in\mathbb{N}}$ in X via $(P_{\mathbf{A}} P_{\mathbf{B}}\mathbf{y}_n)_{n\in\mathbb{N}} = (a_n, \ldots, a_n)_{n\in\mathbb{N}}$. Altogether, the sequence $(a_n)_{n\in\mathbb{N}}$ converges strongly to $P_C x$.

Acknowledgment

H. H. Bauschke's work was supported in part by the Natural Sciences and Engineering Research Council of Canada.

References

[1] J.-P. Aubin, Optima and Equilibria, second ed., Springer, Berlin, 1998.

[2] H.H. Bauschke, P.L. Combettes, A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert spaces, Math. Oper. Res. 26 (2001) 248–264.

[3] H.H. Bauschke, P.L. Combettes, D.R. Luke, Phase retrieval, error reduction algorithm, and Fienup variants: a view from convex optimization, J. Opt. Soc. Amer. A 19 (2002) 1334–1345.

[4] H.H. Bauschke, P.L. Combettes, D.R. Luke, Finding best approximation pairs relative to two closed convex sets in Hilbert spaces, J. Approx. Theory 127 (2004) 178–192.
[5] F. Deutsch, Best Approximation in Inner Product Spaces, Springer, New York, 2001.
[6] K. Goebel, W.A. Kirk, Topics in Metric Fixed Point Theory, Cambridge University Press, Cambridge, 1990.
[7] Y. Haugazeau, Sur les Inéquations Variationnelles et la Minimisation de Fonctionnelles Convexes, Thèse, Université de Paris, France, 1968.
[8] P.L. Lions, B. Mercier, Splitting algorithms for the sum of two nonlinear operators, SIAM J. Numer. Anal. 16 (1979) 964–979.
[9] G. Pierra, Éclatement de contraintes en parallèle pour la minimisation d'une forme quadratique, Lecture Notes in Computer Science, vol. 41, Springer, New York, 1976, pp. 200–218.