Convex Feasibility Problems

Laureate Prof. Jonathan Borwein with Matthew Tam
http://carma.newcastle.edu.au/DRmethods/paseky.html

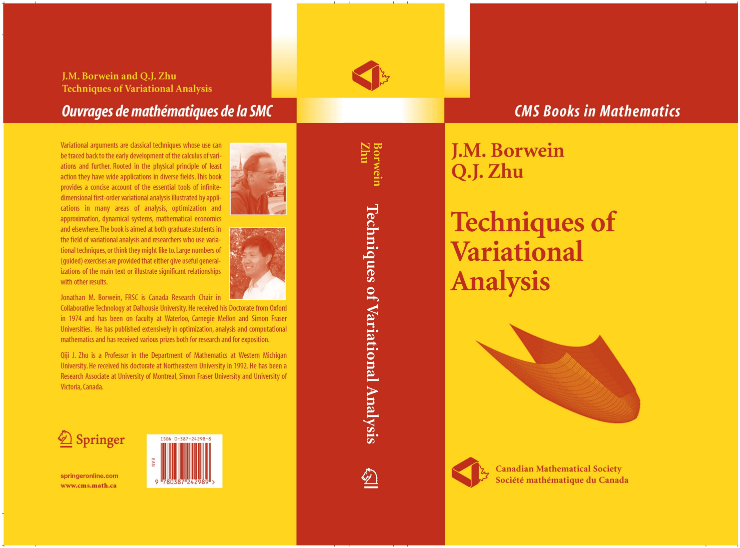
THE UNIVERSITY OF NEWCASTLE AUSTRALIA

CARMA

Spring School on Variational Analysis VI
Paseky nad Jizerou, April 19–25, 2015

Last Revised: May 6, 2016

Techniques of Variational Analysis



This lecture is based on Chapter 4.5: Convex Feasibility Problems

Abstract

Let X be a Hilbert space and let $C_n, n = 1, \ldots, N$ be convex closed subsets of X. The convex feasibility problem is to find some point

$$x \in \bigcap_{n=1}^{N} C_n,$$

when this intersection is non-empty.

In this talk we discuss projection algorithms for finding such a feasibility point. These algorithms have wide ranging applications including:

- solutions to convex inequalities,
- minimization of convex nonsmooth functions,
- medical imaging,
- computerized tomography, and
- electron microscopy

Abstract

Let X be a Hilbert space and let $C_n, n = 1, \ldots, N$ be convex closed subsets of X. The convex feasibility problem is to find some point

$$x \in \bigcap_{n=1}^{N} C_n,$$

when this intersection is non-empty.

In this talk we discuss projection algorithms for finding such a feasibility point. These algorithms have wide ranging applications including:

- solutions to convex inequalities,
- minimization of convex nonsmooth functions,
- medical imaging,
- computerized tomography, and
- electron microscopy

Projections

We start by defining projection to a closed convex set and its basic properties. This is based on the following theorem.

Theorem 4.5.1 (Existence and Uniqueness of Nearest Point)

Let X be a Hilbert space and let C be a closed convex subset of X. Then for any $x \in X$, there exists a unique element $\bar{x} \in C$ such that

$$\|x - \bar{x}\| = d(C; x).$$

Proof. If $x \in C$ then $\bar{x} = x$ satisfies the conclusion. Suppose that $x \notin C$. Then there exists a sequence $x_i \in C$ such that $d(C; x) = \lim_{i\to\infty} \|x - x_i\|$. Clearly, x_i is bounded and therefore has a subsequence weakly converging to some $\bar{x} \in X$.

Projections

We start by defining projection to a closed convex set and its basic properties. This is based on the following theorem.

Theorem 4.5.1 (Existence and Uniqueness of Nearest Point)

Let X be a Hilbert space and let C be a closed convex subset of X. Then for any $x \in X$, there exists a unique element $\bar{x} \in C$ such that

$$\|x - \bar{x}\| = d(C; x).$$

Proof. If $x \in C$ then $\bar{x} = x$ satisfies the conclusion. Suppose that $x \notin C$. Then there exists a sequence $x_i \in C$ such that $d(C; x) = \lim_{i \to \infty} \|x - x_i\|$. Clearly, x_i is bounded and therefore has a subsequence weakly converging to some $\bar{x} \in X$.

Since a closed convex set is weakly closed (Mazur's Theorem), we have $\bar{x} \in C$ and $d(C; x) = \|x - \bar{x}\|$. We show such $\bar{x}$ is unique. Suppose that $z \in C$ also has the property that $d(C; x) = \|x - z\|$. Then for any $t \in [0, 1]$ we have $t\bar{x} + (1 - t)z \in C$. It follows that

$$\begin{aligned} d(C; x) &\leq \|x - (t\bar{x} + (1 - t)z)\| = \|t(x - \bar{x}) + (1 - t)(x - z)\| \\ &\leq t\|x - \bar{x}\| + (1 - t)\|x - z\| = d(C; x). \end{aligned}$$

That is to say

$$t \to \|x - z - t(\bar{x} - z)\|^2 = \|x - z\|^2 - 2t\langle x - z, \bar{x} - z\rangle + t^2\|\bar{x} - z\|^2$$

is a constant mapping, which implies $\bar{x} = z$. ●

The nearest point can be characterized by the normal cone as follows.

Theorem 4.5.2 (Normal Cone Characterization of Nearest Point)

Let X be a Hilbert space and let C be a closed convex subset of X. Then for any $x \in X$, $\bar{x} \in C$ is a nearest point to x if and only if

$$x - \bar{x} \in N(C; \bar{x}).$$

Proof. Noting that the convex function $f(y) = \|y - x\|^2/2$ attains a minimum at $\bar{x}$ over set C, this directly follows from the Pshenichnii–Rockafellar condition (Theorem 4.3.6):

$$0 \in \partial f(\bar{x}) + N(C; \bar{x}).$$

The nearest point can be characterized by the normal cone as follows.

Theorem 4.5.2 (Normal Cone Characterization of Nearest Point)

Let X be a Hilbert space and let C be a closed convex subset of X. Then for any $x \in X$, $\bar{x} \in C$ is a nearest point to x if and only if

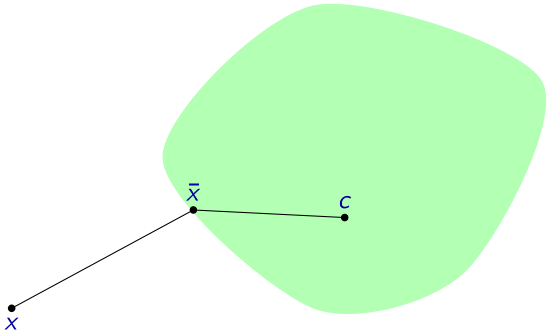
$$x - \bar{x} \in N(C; \bar{x}).$$

Proof. Noting that the convex function $f(y) = \|y - x\|^2/2$ attains a minimum at $\bar{x}$ over set C, this directly follows from the Pshenichnii–Rockafellar condition (Theorem 4.3.6):

$$0 \in \partial f(\overline{x}) + N(C; \overline{x}).$$

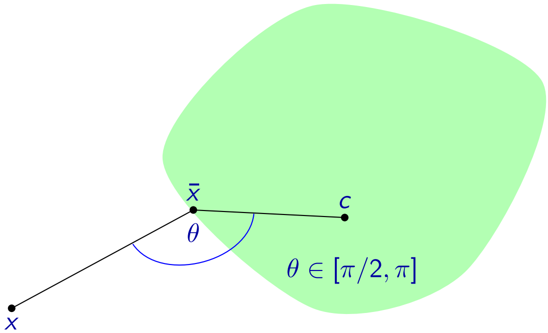
Projections

Geometrically, the normal cone characterization is:



$$x - \bar{x} \in N(C; \bar{x}) \iff \langle x - \bar{x}, c - \bar{x} \rangle \leq 0 \text{ for all } c \in C.$$

Geometrically, the normal cone characterization is:



$$x - \bar{x} \in N(C; \bar{x}) \iff \langle x - \bar{x}, c - \bar{x} \rangle \leq 0 \text{ for all } c \in C.$$

Projections

Definition 4.5.3 (Projection)

Let X be a Hilbert space and let C be a closed convex subset of X. For any $x \in X$ the unique nearest point $y \in C$ is called the projection of x on C and we define the projection mapping P_C by $P_C x = y$.

We summarize some useful properties of the projection mapping in the next proposition whose elementary proof is left as an exercise.

Proposition 4.5.4 (Properties of Projection)

Let X be a Hilbert space and let C be a closed convex subset of X. Then the projection mapping P_C has the following properties.

(i) for any $x \in C$, $P_C x = x$;

(ii) $P_C^2 = P_C$;

(iii) for any $x, y \in X$, $\|P_C y - P_C x\| \le \|y - x\|$.

Projections

Definition 4.5.3 (Projection)

Let X be a Hilbert space and let C be a closed convex subset of X. For any $x \in X$ the unique nearest point $y \in C$ is called the projection of x on C and we define the projection mapping P_C by $P_C x = y$.

We summarize some useful properties of the projection mapping in the next proposition whose elementary proof is left as an exercise.

Proposition 4.5.4 (Properties of Projection)

Let X be a Hilbert space and let C be a closed convex subset of X. Then the projection mapping P_C has the following properties.

(i) for any $x \in C$, $P_C x = x$;

(ii) $P_C^2 = P_C$;

(iii) for any $x, y \in X$, $\|P_C y - P_C x\| \leq \|y - x\|$.

Projections

Theorem 4.5.5 (Potential Function of Projection)

Let X be a Hilbert space and let C be a closed convex subset of X. Define

$$f(x) = \sup\left\{ \langle x, y \rangle - \frac{\|y\|^2}{2} \;\middle|\; y \in C \right\}.$$

Then f is convex, $P_C(x) = f'(x)$, and therefore P_C is a monotone operator.

Proof. It is easy to check that f is convex and

$$f(x) = \frac{1}{2}(\|x\|^2 - \|x - P_C(x)\|^2).$$

We need only show $P_C(x) = f'(x)$.

Theorem 4.5.5 (Potential Function of Projection)

Let X be a Hilbert space and let C be a closed convex subset of X. Define

$$f(x) = \sup\Big\{\langle x, y\rangle - \frac{\|y\|^2}{2} \;\Big|\; y \in C\Big\}.$$

Then f is convex, $P_C(x) = f'(x)$, and therefore P_C is a monotone operator.

Proof. It is easy to check that f is convex and

$$f(x) = \frac{1}{2}(\|x\|^2 - \|x - P_C(x)\|^2).$$

We need only show $P_C(x) = f'(x)$.

Fix $x \in X$. For any $y \in X$ we have

$$\|(x+y) - P_C(x+y)\| \leq \|(x+y) - P_C(x)\|,$$

so

$$\begin{aligned} \|(x+y) - P_C(x+y)\|^2 &\leq \|x+y\|^2 - 2\langle x+y, P_C(x)\rangle + \|P_C(x)\|^2 \\ &= \|x+y\|^2 + \|x - P_C(x)\|^2 - \|x\|^2 \\ &- 2\langle y, P_C(x)\rangle, \end{aligned}$$

hence $f(x+y) - f(x) - \langle P_C(x), y\rangle \geq 0$. On the other hand, since $\|x - P_C(x)\| \leq \|x - P_C(x+y)\|$ we get

$$\begin{aligned} f(x+y) - f(x) - \langle P_C(x), y\rangle &\leq \langle y, P_C(x+y) - P_C(x)\rangle \\ &\leq \|y\| \times \|P_C(x+y) - P_C(x)\| \\ &\leq \|y\|^2, \end{aligned}$$

which implies $P_C(x) = f'(x)$. ●

Projection Algorithms as Minimization Problems

We start with the simple case of the intersection of two convex sets. Let X be a Hilbert space and let C and D be two closed convex subsets of X. Suppose that $C \cap D \neq \emptyset$. Define a function

$$f(c,d) := \frac{1}{2}\|c-d\|^2 + \iota_C(c) + \iota_D(d).$$

We see that f attains a minimum at $(\bar{c},\bar{d})$ if and only if $\bar{c} = \bar{d} \in C \cap D$. Thus, the problem of finding a point in $C \cap D$ becomes one of minimizing function f.

We consider a natural descending process for f by alternately minimizing f with respect to its two variables. More concretely, start with any $x_0 \in D$. Let x_1 be the solution of minimizing

$$x \to f(x, x_0).$$

Projection Algorithms as Minimization Problems

We start with the simple case of the intersection of two convex sets. Let X be a Hilbert space and let C and D be two closed convex subsets of X. Suppose that $C \cap D \neq \emptyset$. Define a function

$$f(c, d) := \frac{1}{2}\|c - d\|^2 + \iota_C(c) + \iota_D(d).$$

We see that f attains a minimum at $(\bar{c}, \bar{d})$ if and only if $\bar{c} = \bar{d} \in C \cap D$. Thus, the problem of finding a point in $C \cap D$ becomes one of minimizing function f.

We consider a natural descending process for f by alternately minimizing f with respect to its two variables. More concretely, start with any $x_0 \in D$. Let x_1 be the solution of minimizing

$$x \to f(x, x_0).$$

Projection Algorithms as Minimization Problems

We start with the simple case of the intersection of two convex sets. Let X be a Hilbert space and let C and D be two closed convex subsets of X. Suppose that $C \cap D \neq \emptyset$. Define a function

$$f(c, d) := \frac{1}{2}\|c - d\|^2 + \iota_C(c) + \iota_D(d).$$

We see that f attains a minimum at $(\bar{c}, \bar{d})$ if and only if $\bar{c} = \bar{d} \in C \cap D$. Thus, the problem of finding a point in $C \cap D$ becomes one of minimizing function f.

We consider a natural descending process for f by alternately minimizing f with respect to its two variables. More concretely, start with any $x_0 \in D$. Let x_1 be the solution of minimizing

$$x \to f(x, x_0).$$

It follows from Theorem 4.5.2. that

$$x_0 - x_1 \in N(C; x_1).$$

That is to say $x_1 = P_C x_0$. We then let x_2 be the solution of minimizing

$$x \to f(x_1, x).$$

Similarly, $x_2 = P_D x_1$. In general, we define

$$x_{i+1} = \begin{cases} P_C x_i & i \text{ is even,} \\ P_D x_i & i \text{ is odd.} \end{cases} \tag{1}$$

This algorithm is a generalization of the classical von Neumann projection algorithm for finding points in the intersection of two subspaces.
We will show that in general x_i weakly converge to a point in $C \cap D$ and when $\text{int}(C \cap D) \neq \emptyset$ we have norm convergence.

Attracting Mappings and Fejér Sequences

We discuss two useful tools for proving the convergence.

Definition 4.5.6 (Nonexpansive Mapping)

Let X be a Hilbert space, let C be a closed convex nonempty subset of X and let $T\colon C \to X$. We say that T is nonexpansive provided that $\|Tx - Ty\| \leq \|x - y\|$.

Definition 4.5.7 (Attracting Mapping)

Let X be a Hilbert space, let C be a closed convex nonempty subset of X and let $T\colon C \to C$ be a nonexpansive mapping. Suppose that D is a closed nonempty subset of C. We say that T is attracting with respect to D if for every $x \in C \backslash D$ and $y \in D$,

$$\|Tx - y\| \leq \|x - y\|.$$

We say that T is k-attracting with respect to D if for every $x \in C \backslash D$ and $y \in D$,

$$k\|x - Tx\|^2 \leq \|x - y\|^2 - \|Tx - y\|^2.$$

Attracting Mappings and Fejér Sequences

We discuss two useful tools for proving the convergence.

Definition 4.5.6 (Nonexpansive Mapping)

Let X be a Hilbert space, let C be a closed convex nonempty subset of X and let $T\colon C \to X$. We say that T is nonexpansive provided that $\|Tx - Ty\| \leq \|x - y\|$.

Definition 4.5.7 (Attracting Mapping)

Let X be a Hilbert space, let C be a closed convex nonempty subset of X and let $T\colon C \to C$ be a nonexpansive mapping. Suppose that D is a closed nonempty subset of C. We say that T is attracting with respect to D if for every $x \in C \backslash D$ and $y \in D$,

$$\|Tx - y\| \leq \|x - y\|.$$

We say that T is k-attracting with respect to D if for every $x \in C \backslash D$ and $y \in D$,

$$k\|x - Tx\|^2 \leq \|x - y\|^2 - \|Tx - y\|^2.$$

Lemma 4.5.8 (Attractive Property of Projection)

Let X be a Hilbert space and let C be a convex closed subset of X. Then $P_C\colon X \to X$ is 1-attracting with respect to C.

Proof. Let $y \in C$. We have

$$\begin{aligned} \|x-y\|^2 - \|P_C x - y\|^2 &= \langle x - P_C x, x + P_C x - 2y\rangle \\ &= \langle x - P_C x, x - P_C x + 2(P_C x - y)\rangle \\ &= \|x - P_C x\|^2 + 2\langle x - P_C x, P_C x - y\rangle \\ &\geq \|x - P_C x\|^2. \end{aligned}$$

Note that if T is attracting (k-attracting) with respect to a set D, then it is attracting (k-attracting) with respect to any subset of D.

Lemma 4.5.8 (Attractive Property of Projection)

Let X be a Hilbert space and let C be a convex closed subset of X. Then $P_C : X \to X$ is 1-attracting with respect to C.

Proof. Let $y \in C$. We have

$$\begin{aligned} \|x-y\|^2 - \|P_C x - y\|^2 &= \langle x - P_C x, x + P_C x - 2y\rangle \\ &= \langle x - P_C x, x - P_C x + 2(P_C x - y)\rangle \\ &= \|x - P_C x\|^2 + 2\langle x - P_C x, P_C x - y\rangle \\ &\geq \|x - P_C x\|^2. \end{aligned}$$

●

Note that if T is attracting (k-attracting) with respect to a set D, then it is attracting (k-attracting) with respect to any subset of D.

Lemma 4.5.8 (Attractive Property of Projection)

Let X be a Hilbert space and let C be a convex closed subset of X. Then $P_C\colon X \to X$ is 1-attracting with respect to C.

Proof. Let $y \in C$. We have

$$\begin{aligned} \|x-y\|^2 - \|P_Cx-y\|^2 &= \langle x-P_Cx, x+P_Cx-2y\rangle \\ &= \langle x-P_Cx, x-P_Cx+2(P_Cx-y)\rangle \\ &= \|x-P_Cx\|^2 + 2\langle x-P_Cx, P_Cx-y\rangle \\ &\geq \|x-P_Cx\|^2. \end{aligned}$$

Note that if T is attracting (k-attracting) with respect to a set D, then it is attracting (k-attracting) with respect to any subset of D.

Attracting Mappings and Fejér Sequences

Definition 4.5.9 (Fejér Monotone Sequence)

Let X be a Hilbert space, let C be closed convex set and let (x_i) be a sequence in X. We say that (x_i) is Fejér monotone with respect to C if

$$\|x_{i+1} - c\| \le \|x_i - c\|, \quad \text{for all } c \in C \text{ and } i = 1, 2, \ldots$$

Next we summarize properties of Fejér monotone sequences.

Theorem 4.5.10 (Properties of Fejér Monotone Sequences)

Let X be a Hilbert space, let C be a closed convex set and let (x_i) be a Fejér monotone sequence with respect to C. Then

(i) (x_i) is bounded and $d(C; x_{i+1}) \le d(C; x_i)$.

(ii) (x_i) has at most one weak cluster point in C.

(iii) If the interior of C is nonempty then (x_i) converges in norm.

(iv) $(P_C x_i)$ converges in norm.

Definition 4.5.9 (Fejér Monotone Sequence)

Let X be a Hilbert space, let C be closed convex set and let (x_i) be a sequence in X. We say that (x_i) is Fejér monotone with respect to C if

$$\|x_{i+1} - c\| \leq \|x_i - c\|, \quad \text{for all } c \in C \text{ and } i = 1, 2, \ldots$$

Next we summarize properties of Fejér monotone sequences.

Theorem 4.5.10 (Properties of Fejér Monotone Sequences)

Let X be a Hilbert space, let C be a closed convex set and let (x_i) be a Fejér monotone sequence with respect to C. Then

(i) (x_i) is bounded and $d(C; x_{i+1}) \leq d(C; x_i)$.

(ii) (x_i) has at most one weak cluster point in C.

(iii) If the interior of C is nonempty then (x_i) converges in norm.

(iv) $(P_C x_i)$ converges in norm.

Proof. (i) is obvious.

Observe that, for any $c \in C$ the sequence $(\|x_i - c\|^2)$ converges and so does

$$(\|x_i\|^2 - 2\langle x_i, c\rangle). \tag{2}$$

Now suppose $c_1, c_2 \in C$ are two weak cluster points of (x_i). Letting c in (2) be c_1 and c_2, respectively, and taking limits of the difference, yields $\langle c_1, c_1 - c_2\rangle = \langle c_2, c_1 - c_2\rangle$ so that $c_1 = c_2$, which proves (ii).

To prove (iii) suppose that $B_r(c) \subset C$. For any $x_{i+1} \neq x_i$, simplifying

$$\|x_{i+1} - \left(c - h\frac{x_{i+1} - x_i}{\|x_{i+1} - x_i\|}\right)\|^2 \leq \|x_i - \left(c - h\frac{x_{i+1} - x_i}{\|x_{i+1} - x_i\|}\right)\|^2$$

we have

$$2h\|x_{i+1} - x_i\| \leq \|x_i - c\|^2 - \|x_{i+1} - c\|^2.$$

Proof. (i) is obvious.
Observe that, for any $c \in C$ the sequence $(\|x_i - c\|^2)$ converges and so does

$$(\|x_i\|^2 - 2\langle x_i, c\rangle). \qquad (2)$$

Now suppose $c_1, c_2 \in C$ are two weak cluster points of (x_i). Letting c in (2) be c_1 and c_2, respectively, and taking limits of the difference, yields $\langle c_1, c_1 - c_2\rangle = \langle c_2, c_1 - c_2\rangle$ so that $c_1 = c_2$, which proves (ii).
To prove (iii) suppose that $B_r(c) \subset C$. For any $x_{i+1} \neq x_i$, simplifying

$$\|x_{i+1} - \left(c - h\frac{x_{i+1} - x_i}{\|x_{i+1} - x_i\|}\right)\|^2 \leq \|x_i - \left(c - h\frac{x_{i+1} - x_i}{\|x_{i+1} - x_i\|}\right)\|^2$$

we have

$$2h\|x_{i+1} - x_i\| \leq \|x_i - c\|^2 - \|x_{i+1} - c\|^2.$$

Proof. (i) is obvious.
Observe that, for any $c \in C$ the sequence $(\|x_i - c\|^2)$ converges and so does

$$(\|x_i\|^2 - 2\langle x_i, c\rangle). \tag{2}$$

Now suppose $c_1, c_2 \in C$ are two weak cluster points of (x_i). Letting c in (2) be c_1 and c_2, respectively, and taking limits of the difference, yields $\langle c_1, c_1 - c_2\rangle = \langle c_2, c_1 - c_2\rangle$ so that $c_1 = c_2$, which proves (ii).
To prove (iii) suppose that $B_r(c) \subset C$. For any $x_{i+1} \neq x_i$, simplifying

$$\|x_{i+1} - \big(c - h\frac{x_{i+1} - x_i}{\|x_{i+1} - x_i\|}\big)\|^2 \leq \|x_i - \big(c - h\frac{x_{i+1} - x_i}{\|x_{i+1} - x_i\|}\big)\|^2$$

we have

$$2h\|x_{i+1} - x_i\| \leq \|x_i - c\|^2 - \|x_{i+1} - c\|^2.$$

For any $j > i$, adding the above inequality from i to $j-1$ yields

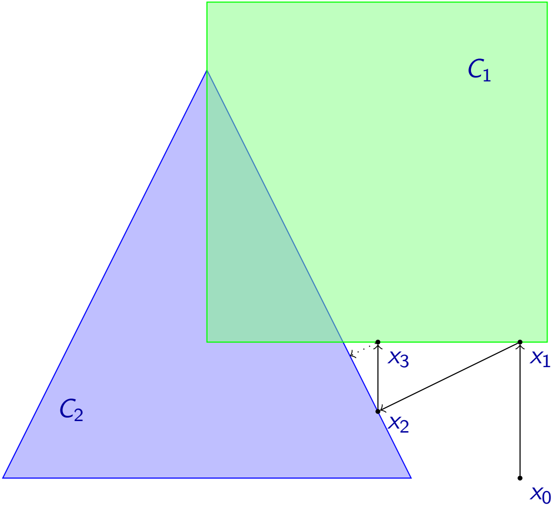
$$2h\|x_j - x_i\| \leq \|x_i - c\|^2 - \|x_j - c\|^2.$$

Since $(\|x_i - c\|^2)$ is a convergent sequence we conclude that (x_i) is a Cauchy sequence.
Finally, for natural numbers i, j with $j > i$, apply the parallelogram law $\|a - b\|^2 = 2\|a\|^2 + 2\|b\|^2 - \|a + b\|^2$ to $a := P_C x_j - x_j$ and $b := P_C x_i - x_j$ we obtain

$$\begin{aligned}
\|P_C x_j - P_C x_i\|^2 &= 2\|P_C x_j - x_j\|^2 + 2\|P_C x_i - x_j\|^2 \\
&\quad - 4\left\|\frac{P_C x_j + P_C x_i}{2} - x_j\right\|^2 \\
&\leq 2\|P_C x_j - x_j\|^2 + 2\|P_C x_i - x_j\|^2 \\
&\quad - 4\|P_C x_j - x_j\|^2 \\
&\leq 2\|P_C x_i - x_j\|^2 - 2\|P_C x_j - x_j\|^2 \\
&\leq 2\|P_C x_i - x_i\|^2 - 2\|P_C x_j - x_j\|^2.
\end{aligned}$$

We identify $(P_C x_i)$ as a Cauchy sequence, because $(\|x_i - P_C x_i\|)$ converges by (i).

The following example shows the first few terms of a sequence $\{x_n\}$ which is Fejér monotone with respect to $C = C_1 \cap C_2$.



Convergence of Projection Algorithms

Let X be a Hilbert space. We say a sequence (x_i) in X is *asymptotically regular* if

$$\lim_{i\to\infty} \|x_i - x_{i+1}\| = 0.$$

Lemma 4.5.11 (Asymptotical Regularity of Projection Algorithm)

Let X be a Hilbert space and let C and D be closed convex subsets of X. Suppose $C \cap D \neq \emptyset$. Then the sequence (x_i) defined by the projection algorithm

$$x_{i+1} = \begin{cases} P_C x_i & i \text{ is even,} \\ P_D x_i & i \text{ is odd.} \end{cases}$$

is asymptotically regular.

Let X be a Hilbert space. We say a sequence (x_i) in X is *asymptotically regular* if

$$\lim_{i\to\infty} \|x_i - x_{i+1}\| = 0.$$

Lemma 4.5.11 (Asymptotical Regularity of Projection Algorithm)

Let X be a Hilbert space and let C and D be closed convex subsets of X. Suppose $C \cap D \neq \emptyset$. Then the sequence (x_i) defined by the projection algorithm

$$x_{i+1} = \begin{cases} P_C x_i & i \text{ is even,} \\ P_D x_i & i \text{ is odd.} \end{cases}$$

is asymptotically regular.

Proof. By Lemma 4.5.8 both P_C and P_D are 1-attracting with respect to $C \cap D$. Let $y \in C \cap D$. Since x_{i+1} is either $P_C x_i$ or $P_D x_i$ it follows that

$$\|x_{i+1} - x_i\|^2 \leq \|x_i - y\|^2 - \|x_{i+1} - y\|^2.$$

Since $(\|x_i - y\|^2)$ is a monotone decreasing sequence, therefore the right-hand side of the inequality converges to 0 and the result follows. ●

Now, we are ready to prove the convergence of the projection algorithm.

Theorem 4.5.12 (Convergence for Two Sets)

Let X be a Hilbert space and let C and D be closed convex subsets of X. Suppose $C \cap D \neq \emptyset$ $(\mathrm{int}(C \cap D) \neq \emptyset)$. Then the projection algorithm

$$x_{i+1} = \begin{cases} P_C x_i & i \text{ is even,} \\ P_D x_i & i \text{ is odd.} \end{cases}$$

converges weakly (in norm) to a point in $C \cap D$.

Proof. Let $y \in C \cap D$. Then, for any $x \in X$, we have

$$\|P_C x - y\| = \|P_C x - P_C y\| \leq \|x - y\|, \quad \text{and}$$
$$\|P_D x - y\| = \|P_D x - P_D y\| \leq \|x - y\|.$$

Since x_{i+1} is either $P_C x_i$ or $P_D x_i$ we have that

$$\|x_{i+1} - y\| \leq \|x_i - y\|.$$

That is to say (x_i) is a Fejér monotone sequence with respect to $C \cap D$. By item (i) of Theorem 4.5.10 the sequence (x_i) is bounded, and therefore has a weakly convergent subsequence. We show that all weak cluster points of (x_i) belong to $C \cap D$. In fact, let (x_{i_k}) be a subsequence of (x_i) converging to x weakly.

Taking a subsequence again if necessary we may assume that (x_{i_k}) is a subset of either C or D. For the sake of argument let us assume that it is a subset of C and, thus, the weak limit x is also in C. On the other hand by the asymptotical regularity of (x_i) in Lemma 4.5.11 $(P_D x_{i_k}) = (x_{i_k+1})$ also weakly converges to x. Since $(P_D x_{i_k})$ is a subset of D we conclude that $x \in D$, and therefore $x \in C \cap D$. By item (ii) of Theorem 4.5.10 (x_i) has at most one weak cluster point in $C \cap D$, and we conclude that (x_i) weakly converges to a point in $C \cap D$. When $\operatorname{int}(C \cap D) \neq \emptyset$ it follows from item (iii) of Theorem 4.5.10 that (x_i) converges in norm. ●

Convergence of Projection Algorithms

Whether the alternating projection algorithm converged in norm without the assumption that

$$\mathrm{int}(C \cap D) \neq \emptyset,$$

or more generally of metric regularity, was a long-standing open problem.

Recently Hundal constructed an example showing that the answer is negative [5].

The proof of Hundal's example is self-contained and elementary. However, it is quite long and delicate, therefore we will be satisfied in stating the example.

Whether the alternating projection algorithm converged in norm without the assumption that

$$\operatorname{int}(C \cap D) \neq \emptyset,$$

or more generally of metric regularity, was a long-standing open problem.

Recently Hundal constructed an example showing that the answer is negative [5].

The proof of Hundal's example is self-contained and elementary. However, it is quite long and delicate, therefore we will be satisfied in stating the example.

Example 4.5.13 (Hundal)

Let $X = \ell_2$ and let $\{e_i \mid i = 1, 2, \dots\}$ be the standard basis of X. Define $v\colon [0, +\infty) \to X$ by

$$v(r) := \exp(-100r^3)e_1 + \cos\left((r - [r])\pi/2\right)e_{[r]+2} + \sin\left((r - [r])\pi/2\right)e_{[r]+3},$$

where $[r]$ signifies the integer part of r and further define

$$C = \{e_1\}^\perp \text{ and } D = \operatorname{conv}\{v(r) \mid r \geq 0\}.$$

Then the hyperplane C and cone D satisfies $C \cap D = \{0\}$.
However, Hundal's sequence of alternating projections x_i given by

$$x_{i+1} = P_D P_C x_i$$

starting from $x_0 = v(1)$ (necessarily) converges weakly to 0, but not in norm.

A related useful example is the moment problem.

Example 4.5.14 (Moment Problem)

Let X be a Hilbert lattice[1] with lattice cone $D = X^+$. Consider a linear continuous mapping A from X onto $\mathbb{R}^N$. The moment problem seeks the solution of $A(x) = y \in \mathbb{R}^N, x \in D$.

Define $C = A^{-1}(y)$. Then the moment problem is feasible iff

$$C \cap D \neq \emptyset.$$

A natural question is whether the projection algorithm converges in norm.

This problem is answered affirmatively in [1] for $N = 1$ yet remains open in general when $N > 1$.

[1] All Hilbert lattices are realized as $L_2(\Omega, \mu)$ in the natural ordering for some measure space.

Projection Algorithms for Multiple Sets

We now turn to the general problem of finding some points in

$$\bigcap_{n=1}^{N} C_n,$$

where C_n, $n = 1, \dots, N$ are closed convex sets in a Hilbert space X.

Let a_n, $n = 1, \dots, N$ be positive numbers. Denote

$$X^N := \{x = (x_1, x_2, \dots, x_N) \mid x_n \in X, n = 1, \dots, N\}$$

the product space of N copies of X with inner product

$$\langle x, y \rangle = \sum_{n=1}^{N} a_n \langle x_n, y_n \rangle.$$

Then X^N is a Hilbert space.

Define

$$C := C_1 \times C_2 \times \cdots \times C_N, \quad \text{and}$$
$$D := \{(x_1, \ldots, x_N) \in X^N : x_1 = x_2 = \cdots = x_N\}.$$

Then C and D are closed convex sets in X^N and

$$x \in \bigcap_{n=1}^{N} C_n \iff (x, x, \ldots, x) \in C \cap D.$$

Applying the projection algorithm (1) to the convex sets C and D defined above we have the following generalized projection algorithm for finding some points in

$$\bigcap_{n=1}^{N} C_n,$$

as we shall now explain.

Projection Algorithms for Multiple Sets

Define

$$C := C_1 \times C_2 \times \cdots \times C_N, \quad \text{and}$$
$$D := \{(x_1, \ldots, x_N) \in X^N : x_1 = x_2 = \cdots = x_N\}.$$

Then C and D are closed convex sets in X^N and

$$x \in \bigcap_{n=1}^{N} C_n \iff (x, x, \ldots, x) \in C \cap D.$$

Applying the projection algorithm (1) to the convex sets C and D defined above we have the following generalized projection algorithm for finding some points in

$$\bigcap_{n=1}^{N} C_n,$$

as we shall now explain.

Projection Algorithms for Multiple Sets

Denote $P_n = P_{C_n}$. The algorithm can be expressed by

$$x_{i+1} = \Big(\sum_{n=1}^{N} \lambda_n P_n \Big) x_i, \tag{3}$$

where $\lambda_n = a_n / \sum_{m=1}^{N} a_m$.

In other words, each new approximation is the convex combination of the projections of the previous step to all the sets $C_n, n = 1, \ldots, N$. It follows from the convergence theorem in the previous subsection that the algorithm (3) converges weakly to some point in $\bigcap_{n=1}^{N} C_n$ when this intersection is nonempty.

Projection Algorithms for Multiple Sets

Denote $P_n = P_{C_n}$. The algorithm can be expressed by

$$x_{i+1} = \Big(\sum_{n=1}^{N} \lambda_n P_n\Big) x_i, \tag{3}$$

where $\lambda_n = a_n / \sum_{m=1}^{N} a_m$.

In other words, each new approximation is the convex combination of the projections of the previous step to all the sets $C_n, n = 1, \ldots, N$. It follows from the convergence theorem in the previous subsection that the algorithm (3) converges weakly to some point in $\bigcap_{n=1}^{N} C_n$ when this intersection is nonempty.

Theorem 4.5.15 (Weak Convergence for N Sets)

Let X be a Hilbert space and let $C_n, n = 1, \dots, N$ be closed convex subsets of X. Suppose that $\bigcap_{n=1}^N C_n \neq \emptyset$ and $\lambda_n \geq 0$ satisfies $\sum_{n=1}^N \lambda_n = 1$. Then the projection algorithm

$$x_{i+1} = \left(\sum_{n=1}^N \lambda_n P_n \right) x_i,$$

converges weakly to a point in $\bigcap_{n=1}^N C_n$.

Proof. This follows directly from Theorem 4.5.12. ∎

Projection Algorithms for Multiple Sets

When the interior of $\bigcap_{n=1}^N C_n$ is nonempty we also have that the algorithm (3) converges in norm. However, since D does not have interior this conclusion cannot be derived from Theorem 4.5.12. Rather it has to be proved by directly showing that the approximation sequence is Fejér monotone w.r.t. $\bigcap_{n=1}^N C_n$.

Theorem 4.5.16 (Strong Convergence for N Sets)

Let X be a Hilbert space and let $C_n, n = 1, \ldots, N$ be closed convex subsets of X. Suppose that $\operatorname{int} \bigcap_{n=1}^N C_n \neq \emptyset$ and $\lambda_n \geq 0$ satisfies $\sum_{n=1}^N \lambda_n = 1$. Then the projection algorithm

$$x_{i+1} = \left(\sum_{n=1}^N \lambda_n P_n\right) x_i,$$

converges to a point in $\bigcap_{n=1}^N C_n$ in norm.

Projection Algorithms for Multiple Sets

When the interior of $\bigcap_{n=1}^N C_n$ is nonempty we also have that the algorithm (3) converges in norm. However, since D does not have interior this conclusion cannot be derived from Theorem 4.5.12. Rather it has to be proved by directly showing that the approximation sequence is Fejér monotone w.r.t. $\bigcap_{n=1}^N C_n$.

Theorem 4.5.16 (Strong Convergence for N Sets)

Let X be a Hilbert space and let $C_n, n = 1, \ldots, N$ be closed convex subsets of X. Suppose that $\operatorname{int} \bigcap_{n=1}^N C_n \neq \emptyset$ and $\lambda_n \geq 0$ satisfies $\sum_{n=1}^N \lambda_n = 1$. Then the projection algorithm

$$x_{i+1} = \Big(\sum_{n=1}^{N} \lambda_n P_n\Big) x_i,$$

converges to a point in $\bigcap_{n=1}^N C_n$ in norm.

Proof. Let $y \in \bigcap_{n=1}^{N} C_n$. Then

$$\begin{aligned} \|x_{i+1} - y\| &= \left\| \left(\sum_{n=1}^{N} \lambda_n P_n \right) x_i - y \right\| = \left\| \sum_{n=1}^{N} \lambda_n (P_n x_i - P_n y) \right\| \\ &\leq \sum_{n=1}^{N} \lambda_n \|P_n x_i - P_n y\| \leq \sum_{n=1}^{N} \lambda_n \|x_i - y\| = \|x_i - y\|. \end{aligned}$$

That is to say (x_i) is a Fejér monotone sequence with respect to $\bigcap_{n=1}^{N} C_n$. The norm convergence of (x_i) then follows directly from Theorems 4.5.10 and 4.5.15. ∎

Commentary and Open Questions

- We have proven convergence of the projection algorithm. It can be traced to von Neumann, Weiner and before, and has been studied extensively.
- We emphasize the relationship between the projection algorithm and variational methods in Hilbert spaces:
  - While projection operators can be defined outside of the setting of Hilbert space, they are not necessarily non-expansive.
  - In fact, non-expansivity of the projection operator characterizes Hilbert space in two more dimensions.
- The Hundal example clarifies many other related problems regarding convergence. Simplifications of the example have since been published.
  - What happens if we only allow "nice" cones?
- Bregman distance provides an alternative perspective into many generalizations of the projection algorithm.

Exercises

1. Let $T : \mathcal{H} \to \mathcal{H}$ be nonexpansive and let $\alpha \in [-, 1, 1]$. Show that $(I + \alpha T)$ is a *maximally monotone* continuous operator.
2. (Common projections) Prove formula for the projection onto each of the following sets:
    1. Half-space: $H := \{x \in \mathcal{H} : \langle a, x \rangle = b\}$, $0 \neq a \in \mathcal{H}$, $b \in \mathbb{R}$.
    2. Line: $L := x + \mathbb{R}y$ where $x, y \in \mathcal{H}$.
    3. Ball: $B := \{x \in \mathcal{H} : \|x\| \leq r\}$ where $r > 0$.
    4. Ellipse in $\mathbb{R}^2$: $E := \{(x, y) \in \mathbb{R}^2 : x^2/a^2 + y^2/b^2 = 1\}$.

       Hint: $P_E(u, v) = \left(\frac{a^2 u}{a^2 - t}, \frac{b^2 v}{b^2 - t}\right)$ where t solves

       $$\frac{a^2 u^2}{(a^2 - t)^2} + \frac{b^2 v^2}{(b^2 - t)^2} = 1.$$
3. (Non-existence of best approximations) Let $\{e_n\}_{n \in \mathbb{N}}$ be an orthonormal basis of an infinite dimensional Hilbert space. Define the set $A := \{e_1/n + e_n : n \in \mathbb{N}\}$. Show that A is norm closed and bounded but $d_A(0) = 1$ is not attained. Is A weakly closed?

References

- H.H. Bauschke & and J.M. Borwein. **On the convergence of von Neumann's alternating projection algorithm for two sets**. *Set-Valued Analysis* 1(2):185–212 (1993).
- J.M. Borwein & Q.J. Zhu. **Techniques of Variational Analysis**. Springer (2013).
- Y. Censor & A. Cegielski. **Projection methods: an annotated bibliography of books and reviews**. *Optimization*, accepted for publication. Preprint PDF.
- R. Escalante & M. Raydan, **Alternating projection methods**. SIAM Fundamentals of Algorithms (2011)
- H.S. Hundal. **An alternating projection that does not converge in norm**. *Nonlinear Analysis: Theory, Methods & Applications* 57(1):35–61 (2004).

Many resources (and definitions) available at:

`http://www.carma.newcastle.edu.au/jon/ToVA/`