The Expected-Projection Method: Its Behavior and Applications to Linear Operator Equations and Convex Optimization

Dan Butnariu
Department of Mathematics and Computer Science
University of Haifa, 31905 Haifa, Israel

November 8, 1995

Abstract

It was shown by Butnariu and Flåm [5] that, under some conditions, sequences generated by the expected projection method (EPM) in Hilbert spaces approximate almost common points of measurable families of closed convex subsets provided that such points exists. In this work we study the behavior of the EPM in the more general situation when the involved sets may or may not have almost common points and we give necessary and sufficient conditions for weak and strong convergence. Also, we show how the EPM can be applied to finding solutions of linear operator equations and to solving convex optimization problems.

Keywords: Convex set, Metric projection, Bochner integral, Asymptotic center of a sequence, Stochastic convex feasibility problem, Optimization problem.

Classification: 52A40, 90C30, 45B05

1 Introduction

1.1 This work is aimed at giving an answer to the following problem regarding the convergence of sequences generated in a Hilbert space via the so called *expected projection method*:

(CP): Let (Ω, A, μ) be a complete probability space and let Q be a measurable point-to-set mapping from Ω to the separable Hilbert space H such that, for each $\omega \in \Omega$, the set Q_ω is closed, convex and nonempty. Suppose that Q has an integrable[1] selector, that is, there exists an integrable function $\phi : \Omega \to H$ such that $\phi(\omega) \in Q_\omega$, for almost all $\omega \in \Omega$. In this case, the operator $\mathbf{P} : H \to H$, called the *expected projection operator* (with respect to Q), and given by

$$\mathbf{P}(x) = \int_\Omega P_\omega(x) d\mu(\omega), \tag{1}$$

where $P_\omega(x)$ denotes the (metric) projection of $x \in H$ onto the set Q_ω, is well-defined (cf. [5, Proposition 2.3]). An *expected-projection method (EPM) generated sequence* $\left\{x^k\right\}_{k\in\mathbf{N}}$ (with respect to Q) is defined recursively such that *the initial point* $x^0 \in H$ and

$$x^{k+1} = \mathbf{P}(x^k), \quad k \in \mathbf{N}. \tag{2}$$

The question is whether, or in what conditions, any EPM generated sequence converges (weakly or strongly) in H and, if affirmative, what can be said about the location of its limit.

1.2 The convergence problem (CP) appears in applied mathematics in various forms. Recall that a point $x^* \in H$ is called an *almost common point of the sets* Q_ω, $\omega \in \Omega$, if[2]

$$\mu\left(\{\omega \in \Omega; x^* \in Q_\omega\}\right) = 1. \tag{3}$$

The so called *stochastic convex feasibility problem (SCFP)* studied in [5] is that of finding an almost common point of the sets Q_ω, $\omega \in \Omega$. The SCFP is termed *consistent* if the sets Q_ω, $\omega \in \Omega$, have an almost common point; otherwise, the SCFP is called *inconsistent*. It was shown in [5] that, when the given SCFP is consistent and its data satisfy some additional conditions, any EPM generated sequence converges strongly and its limit is an almost common point of the sets Q_ω, $\omega \in \Omega$. Moreover, by combining [5, Lemma 4.6] and [3, Theorem 1] one can easily deduce that, *if the SCFP is consistent*, then any EPM generated sequence converges weakly to an almost common point of the sets Q_ω, $\omega \in \Omega$. These show that, when consistency of the given SCFP can be a priori guaranteed, the EPM generated sequences provide (weakly or strong) approximations of almost common points of the sets Q_ω, $\omega \in \Omega$.

It is known that the system of linear equations or/and inequalities (as those encountered in computed tomography and in signal processing - see [6] and the references therein), the Fredholm type integral equations (see [11]), some best approximation problems (see [5]), can be treated and solved as particular SCFPs. Deciding if such a SCFP is consistent may be difficult because of the amount and the randomness of the data. Even if such a decision is possible at the theoretical

[1]In what follows integrability and integrals (with repect to μ) are in the sense of Bochner (see [12]).

[2]Note that, since Q is measurable, the set $\{\omega \in \Omega; x^* \in Q_\omega\}$ is measurable too.

level (as, for instance, in the case of the Fredholm and Voltera type integral equations studied in [11]), the computer related errors (rounding) may determine that practical computation of the EPM generated sequence supposed to approximate a solution is done not for the original consistent SCFP but for a slightly different SCFP which is, in fact, inconsistent. These naturally lead to the following questions: (i) Do EPM generated sequences converge (weakly or strongly) when the given SCFP is inconsistent? and, if so, (ii) What can be said about the limit of such a sequence? Answering these questions means solving the convergence problem (CP).

1.3 The main result proven in this paper, Theorem 2.2, shows that a necessary and sufficient condition for an EPM generated sequence to be weakly convergent is that the expected projection operator $\mathbf{P}$ has fixed a point. This implies that the answer to the question (i) asked above is affirmative because it may happen that $\mathbf{P}$ has fixed points even if the sets Q_ω, $\omega \in \Omega$, have no almost common points. Theorem 2.2 also answers the question (ii): The weak limit of an EPM generated sequence, whenever it exists, is a fixed point of $\mathbf{P}$. If, in addition, the point-to-set mapping Q has a square integrable selector, then the weak limit of any EPM generated sequence minimizes the average of the squared distances to the sets Q_ω. In the particular case when the sets Q_ω, $\omega \in \Omega$, have almost common points, then all fixed points of $\mathbf{P}$ are almost common points of the sets Q_ω, $\omega \in \Omega$ and conversely.

Theorem 2.2 allows to enlarge the area of applicability of the expected projection method. In Section 3 we show how this result can be used in order to solve linear operator equations $Tx = b$ in the space $\mathcal{L}^2([a,b])$. Corollary 3.2 shows that, under quite mild conditions, finding a solution of such an equation is equivalent to solving a SCFP. Also, it describes the behavior of the EPM generated sequences related to this SCFP in the case when the given linear operator equation has solutions as well as in the inconsistent case. It can be easily seen that Fredholm and Voltera type integral equations in $\mathcal{L}^2([a,b])$ fall in the class of linear operator equations whose solutions can be approximated via the EPM. Another category of problems which can be reduced to SCFPs and solved via the EPM is that of minimizing smooth convex functionals on compact subsets of $\mathbf{R}^n$ (see Section 3.3). Corollary 3.3 implies that large classes of such optimization problems can be solved numerically by an iterative method whose basic requirement is the availability of reliable techniques of determining integrals over the set of feasible solutions.

2 The convergence theorem

2.1 In this section we give our main result describing the behavior of the expected projection method generated sequences. To this end, some preliminary notions and results are needed. Recall that a *square-integrable selector* of the point-to-set mapping Q is a measurable function $\phi : \Omega \to H$ such that $\phi(\omega) \in Q_\omega$, for almost all $\omega \in \Omega$, and such that the functional $\omega \to \|\phi(\omega)\|^2$ is integrable. According to [1, Corollary 8.2.13], for each $x \in H$, the function $\omega \to d(x, Q_\omega) : \Omega \to \mathbf{R}_+$ is measurable and, therefore, so is the function $\omega \to d^2(x, Q_\omega)$. By consequence, for each $x \in H$, the function $f : H \to [0, +\infty]$ given by

$$f(x) = \int_\Omega \|P_\omega(x) - x\|^2 \, d\mu(\omega), \tag{4}$$

is well-defined. Of course, the set $Dom(f) = \{x \in H; f(x) < \infty\}$ my be empty. The next result shows that this can not happen when the point-to-set mapping Q has a square-integrable selector.

Lemma. *The following statements are equivalent:*
(i) $Dom(f) = H$;
(ii) $Dom(f) \neq \emptyset$;
(iii) For each point $x \in H$, the function $\omega \to P_\omega(x)$ is square-integrable;
(iv) The point-to-set mapping Q has a square-integrable selector;
(v) For some point $x^0 \in H$, the function $\omega \to P_\omega(x^0)$ is square-integrable.

Proof. The implications $(i) \Rightarrow (ii)$, $(iii) \Rightarrow (iv)$ and $(iii) \Rightarrow (v)$ are obvious. If $x^* \in Dom(f)$ and $x \in H$, then

$$f(x) \leq \int_\Omega [\|P_\omega(x) - P_\omega(x^*)\| + \|P_\omega(x^*) - x\|]^2 \, d\mu(\omega)$$

$$\leq \int_\Omega [\|P_\omega(x) - P_\omega(x^*)\| + \|P_\omega(x^*) - x^*\| + \|x^* - x\|]^2 \, d\mu(\omega)$$

$$\leq \int_\Omega [2 \cdot \|x^* - x\| + \|P_\omega(x^*) - x^*\|]^2 \, d\mu(\omega)$$

$$\leq 4 \cdot \|x^* - x\|^2 + f(x^*) + 4 \cdot \|x^* - x\| \cdot \int_\Omega \|P_\omega(x^*) - x^*\| \, d\mu(\omega) < \infty,$$

because the function $\omega \to \|P_\omega(x^*) - x^*\|$ is integrable as being square integrable. This proves $(ii) \Rightarrow (i)$. The implication $(iv) \Rightarrow (iii)$ follows from the fact that, for each square integrable selector ϕ of the mapping Q and for any $x \in H$, we have

$$\begin{aligned} f(x) &\leq \textstyle\int_\Omega \|\phi(\omega) - x\|^2 \, d\mu(\omega) \\ &\leq \|x\|^2 + \textstyle\int_\Omega \|\phi(\omega)\|^2 \, d\mu(\omega) + 2 \cdot \|x\| \cdot \textstyle\int_\Omega \|\phi(\omega)\| \, d\mu(\omega) < \infty, \end{aligned} \tag{5}$$

and

$$\begin{aligned} \textstyle\int_\Omega \|P_\omega(x)\|^2 \, d\mu(\omega) &\leq \textstyle\int_\Omega [\|P_\omega(x) - x\| + \|x\|]^2 \, d\mu(\omega) \\ &= f(x) + \|x\|^2 + 2 \cdot \|x\| \cdot \textstyle\int_\Omega \|P_\omega(x) - x\| \, d\mu(\omega) \\ &\leq f(x) + \|x\|^2 + 2 \cdot \|x\| \cdot \textstyle\int_\Omega \|\phi(\omega) - x\| \, d\mu(\omega), \end{aligned} \tag{6}$$

where
$$\int_\Omega \|\phi(\omega)-x\|\,d\mu(\omega) \le \int_\Omega \|\phi(\omega)\|\,d\mu(\omega) + \|x\| < \infty.$$
Observe that from (5) also follows that $(iv)\Rightarrow(i)$. Now, if (i) holds, then the function $\omega \to \|P_\omega(x)-x\|$ is square integrable and, by applying the first inequality in (6), we obtain (iii). The implication $(v)\Rightarrow(iii)$ results from the inequality
$$\|P_\omega(x)\| \le \left\|P_\omega(x)-P_\omega(x^0)\right\| + \left\|P_\omega(x^0)\right\| \le \left\|x-x^0\right\| + \left\|P_\omega(x^0)\right\|,$$
which is satisfied for each pair $(\omega, x) \in \Omega \times H$. □

2.2 Our answer to (CP) is stated as follows:

Theorem.

(A) *The following statements are equivalent:*

$(A1)$ *The set $Fix(\mathbf{P})$ of the fixed points of $\mathbf{P}$ is nonempty;*

$(A2)$ *There exists a point $x^0 \in H$ such that the EPM generated sequence $\left\{x^k\right\}_{k\in\mathbf{N}}$ with the initial point x^0 is bounded;*

$(A3)$ *There exists a point $x^0 \in H$ such that the EPM generated sequence $\left\{x^k\right\}_{k\in\mathbf{N}}$ with the initial point x^0 has a weak accumulation point and*
$$\lim_{k\to\infty}(x^{k+1}-x^k) = 0 \tag{7}$$

$(A4)$ *There exists a point $x^0 \in H$ such that EPM generated sequence $\left\{x^k\right\}_{k\in\mathbf{N}}$ with the initial point x^0 converges weakly;*

$(A5)$ *For each point $x^0 \in H$, the EPM generated sequence $\left\{x^k\right\}_{k\in\mathbf{N}}$ with the initial point x^0 converges weakly.*

(B) *The weak limit of an EPM generated sequence, whenever it exists, belongs to $Fix(\mathbf{P})$.*

(C) *If either one of the conditions $(i)-(v)$ of Lemma 2.1 is satisfied, then the function $f : H \to [0,\infty]$ given by (4) is everywhere finite, convex, continuously differentiable and $Fix(\mathbf{P}) = Arg\min(f)$. In this case, whenever an EPM generated sequence converges weakly its limit is a (global) minimizer of f.*

2.3 The proof of Theorem 2.2 is given below. Observe that Theorem 2.2 and the considerations contained in its proof imply that each of the following statements holds:

(j) If f is proper and $Arg\min(f) \neq \emptyset$, then f is everywhere finite and the weak limit of any EPM generated sequence exists and belongs to $Arg\min(f) = Fix(\mathbf{P})$.

(jj) If the set $\prod_\mu Q_\omega$ of the almost common points of Q_ω, $\omega \in \Omega$, is nonempty, then f is finite (any point in $\prod_\mu Q_\omega$ defines a constant selector of Q), $\inf_{x\in H} f(x) = 0$ and, therefore, $Fix(\mathbf{P}) = \prod_\mu Q_\omega$ and any EPM generated sequence $\left\{x^k\right\}_{k\in\mathbf{N}}$ converges weakly to a point in $\prod_\mu Q_\omega$.

2.4 We start the proof of Theorem 2.2 by observing that the implications $(A1)\Leftrightarrow(A2)$ result from [9, Theorem 5.2] because, according to [5, Theorem 3.2], the expected projection operator $\mathbf{P}$ is nonexpansive. Note that, according to Lemma 2.1, under the assumptions of (C) the function f defined by (4) is everywhere finite. Therefore, Proposition 2.5 in [5] applies and it shows that f is convex, continuously differentiable and, for each $x \in H$,

$$\nabla f(x) = 2 \cdot (x - \mathbf{P}(x)). \tag{8}$$

These imply that

$$Arg\min(f) = Fix(\mathbf{P}). \tag{9}$$

Hence, (C) follows from (A).

2.5 In order to prove (A) we use the following

Lemma. *If $Fix(\mathbf{P}) \neq \emptyset$ and if $\left\{x^k\right\}_{k\in\mathbf{N}}$ is an EPM generated sequence, then, for any $z \in Fix(\mathbf{P})$ and for each $k \in \mathbf{N}$,*

$$\left\|x^k - x^{k+1}\right\|^2 + \left\|x^{k+1} - z\right\|^2 \leq \left\|x^k - z\right\|^2. \tag{10}$$

Proof. Note that

$$< x^k - z, x^{k+1} - x^k >=< x^k - \mathbf{P}(z), x^{k+1} - x^k >$$

$$=< (x^k - x^{k+1}) + (x^{k+1} - \mathbf{P}(z)), x^{k+1} - x^k >$$

$$=< x^{k+1} - \mathbf{P}(z), x^{k+1} - x^k > - \left\|x^{k+1} - x^k\right\|^2$$

$$= \left[\left\|\mathbf{P}(z) - \mathbf{P}(x^k)\right\|^2 - < \mathbf{P}(x^k) - \mathbf{P}(z), x^k - z >\right] - \left\|\mathbf{P}(x^k) - x^k\right\|^2,$$

where the quantity between the square brackets is nonpositive because of Proposition 2.4 in [5]. This shows that

$$< x^k - z, x^{k+1} - x^k >\leq - \left\|x^{k+1} - x^k\right\|^2.$$

Hence,

$$\left\|x^{k+1} - z\right\|^2 = \left\|x^k - z\right\|^2 + \left\|x^{k+1} - x^k\right\|^2 + 2\cdot < x^k - z, x^{k+1} - x^k >$$

$$\leq \left\|x^k - z\right\|^2 + \left\|x^{k+1} - x^k\right\|^2 - 2 \cdot \left\|x^{k+1} - x^k\right\|^2$$

and the proof is complete. $\square$

2.6 The implication $(A1)\Rightarrow(A3)$ is a direct consequence of the following

Lemma. *If $Fix(\mathbf{P}) \neq \emptyset$, then any EPM generated sequence $\left\{x^k\right\}_{k\in\mathbf{N}}$ is bounded and satisfies* (7)*. Moreover, the series $\sum_{k=0}^{\infty}\left\|x^{k+1}-x^k\right\|^2$ converges.*

Proof. Suppose that $z \in Fix(\mathbf{P})$. Summing up the inequalities in (10) for $k = 0, 1, ..., n$ we obtain

$$\sum_{k=0}^{n}\left\|x^{k+1}-x^k\right\|^2 \leq \left\|x^0 - z\right\|^2.$$

This implies that the series $\sum_{k=0}^{\infty}\left\|x^{k+1}-x^k\right\|^2$ converges and, therefore, the equality (7) holds. Applying again (10), we deduce that the sequence $\left\{\left\|x^k - z\right\|\right\}_{k\in\mathbf{N}}$ is nonincreasing, hence convergent and, thus, bounded. Consequently, the sequence $\left\{x^k\right\}_{k\in\mathbf{N}}$ is bounded too. □

2.7 Recall (see [9, Section 4]) that the *asymptotic center* of the bounded sequence $\left\{y^k\right\}_{k\in\mathbf{N}}$ in H is the unique minimizer of the functional $h(\cdot;\left\{y^k\right\}_{k\in\mathbf{N}}) : H \to [0,\infty)$ defined by

$$h(x;\left\{y^k\right\}_{k\in\mathbf{N}}) = \overline{\lim}_{k\to\infty}\left\|y^k - x\right\|. \tag{11}$$

Also, if $\left\{y^k\right\}_{k\in\mathbf{N}}$ converges weakly, then its weak limit and its asymptotic center coincide. We use this fact in order to prove the next result from which the implication $(A3)\Rightarrow(A1)$ clearly follows .

Lemma. *If x^* is a weak accumulation point of an EPM generated sequence $\left\{x^k\right\}_{k\in\mathbf{N}}$ which satisfies* (7)*, then $x^* \in Fix(\mathbf{P})$.*

Proof. Let $\left\{x^{k_p}\right\}_{p\in\mathbf{N}}$ be a subsequence of $\left\{x^k\right\}_{k\in\mathbf{N}}$ which converges weakly to x^*. Then, for each $p \in \mathbf{N}$, we have

$$\left\|x^{k_p} - \mathbf{P}(x^*)\right\| \leq \left\|x^{k_p} - x^{k_p+1}\right\| + \left\|x^{k_p+1} - \mathbf{P}(x^*)\right\|$$

$$= \left\|x^{k_p} - x^{k_p+1}\right\| + \left\|\mathbf{P}(x^{k_p}) - \mathbf{P}(x^*)\right\|$$

$$\leq \left\|x^{k_p} - x^{k_p+1}\right\| + \left\|x^{k_p} - x^*\right\|,$$

because $\mathbf{P}$ is nonexpansive. This implies

$$h(\ \mathbf{P}(x^*);\left\{x^{k_p}\right\}_{p\in\mathbf{N}}) = \overline{\lim}_{k\to\infty}\left\|x^{k_p} - \mathbf{P}(x^*)\right\|$$

$$\leq \overline{\lim}_{k\to\infty}\left(\left\|x^{k_p} - x^{k_p+1}\right\| + \left\|x^{k_p} - x^*\right\|\right)$$

$$\leq \overline{\lim}_{k\to\infty}\left\|x^{k_p} - x^*\right\| = h(x^*;\left\{x^{k_p}\right\}_{p\in\mathbf{N}}),$$

where h is the function defined by (11). Note that x^* is the asymptotic center of $\left\{x^{k_p}\right\}_{p\in\mathbf{N}}$ and, therefore, it is the unique minimizer of $h(\cdot;\left\{x^{k_p}\right\}_{p\in\mathbf{N}})$ over H. Since $\mathbf{P}(x^*)$ also minimizes the functional $h(\cdot;\left\{x^{k_p}\right\}_{p\in\mathbf{N}})$, it results that $x^* = \mathbf{P}(x^*)$. □

2.8 Recall that weakly convergent sequences are bounded. Therefore, the following implications hold $(A5)\Rightarrow(A4)\Rightarrow(A2)$. Now we prove that $(A1)\Rightarrow(A5)$. To this end, let $\left\{x^k\right\}_{k\in\mathbf{N}}$ be an EPM generated sequence. Then, according to $(A1)$ and Lemma 2.6, this sequence is bounded (and, consequently, has a weak accumulation point) and it satisfies (7). According to Lemma 2.7 all accumulation points of $\left\{x^k\right\}_{k\in\mathbf{N}}$ belong to $Fix(\mathbf{P})$. Suppose, by contradiction, that $\left\{x^k\right\}_{k\in\mathbf{N}}$ does not converge weakly. Then, $\left\{x^k\right\}_{k\in\mathbf{N}}$ has two different weak accumulation points x' and x'' and, necessarily, $x', x'' \in Fix(\mathbf{P})$. Applying Lemma 2.5 we deduce that the sequences $\left\{\left\|x^k - x'\right\|\right\}_{k\in\mathbf{N}}$ and $\left\{\left\|x^k - x''\right\|\right\}_{k\in\mathbf{N}}$ are nonincreasing and bounded. Therefore, the following limit exists and is finite

$$a = \lim_{k\to\infty}\left[\left\|x^k - x'\right\|^2 - \left\|x^k - x''\right\|^2\right]$$

$$= \|x'\|^2 - \|x''\|^2 + 2\cdot\lim_{k\to\infty} < x^k, x'' - x' > .$$

Let $\left\{x^{k_p}\right\}_{p\in\mathbf{N}}$ and $\left\{x^{h_q}\right\}_{q\in\mathbf{N}}$ be subsequences of $\left\{x^k\right\}_{k\in\mathbf{N}}$ converging weakly to x' and x'', respectively. Then,

$$a = \|x'\|^2 - \|x''\|^2 + 2\cdot\lim_{p\to\infty} < x^{k_p}, x'' - x' >$$

$$= \|x'\|^2 - \|x''\|^2 + 2\cdot < x', x'' - x' >$$

$$= -\|x' - x''\|.$$

Also, we have

$$a = \|x'\|^2 - \|x''\|^2 + 2\cdot\lim_{q\to\infty} < x^{h_q}, x'' - x' >$$

$$= \|x'\|^2 - \|x''\|^2 + 2\cdot < x'', x'' - x' >$$

$$= \|x' - x''\|.$$

These show that $\|x' - x''\| = a = -\|x' - x''\|$, i.e., $x' = x''$, a contradiction. □

2.9 In order to prove (B), suppose that $\left\{x^k\right\}_{k\in\mathbf{N}}$ is an EPM generated sequence which converges weakly to x^*. Then, according to (A), we have that $Fix(\mathbf{P}) \neq \emptyset$ and (7) is satisfied. Now, application of Lemma 2.7 yelds that $x^* \in Fix(\mathbf{P})$ and the proof of the theorem is complete. □

3 Applications of the EPM

3.1 The EPM in the particular case when Ω is finite, $\mathcal{A} = 2^\Omega$ and $\mu : \mathcal{A} \to [0,1]$ is defined by

$$\mu(A) = \sum_{a \in A} \mu_a,$$

where, for $a \in \Omega$, μ_a are positive real numbers with $\sum_{a\in\Omega} \mu_a = 1$, reduces to the so called *simultaneous projection method* which originates in Cimmino's work [7]. Recent research concerning the behavior of the simultaneous projection method (see Censor and Zenios [6], Combettes [8], Iusem and De Pierro [10], Pierra [13], and the references therein) were mostly motivated by its applications to solving systems of convex inequalities appearing in computed tomography and signal processing. Kammerer and Nashed [11, Section 4] observed that an algorithm similar to that proposed by Cimmino for solving finite systems of linear equations (seen as problems of finding common points of hyperplanes in $\mathbf{R}^n$) can be applied to solve Fredholm and Voltera type integral equations (seen as SCFP of finding almost common points of infinite families of hyperplanes in a Hilbert space). The EPM can be viewed as an extension of the algorithm proposed in [11]. In what follows we show how the EPM can be used as an iterative algorithm for finding solutions of linear operator equations and of convex optimization problems. Other applications of the EPM are discussed in [5].

3.2 *Finding solutions to the linear operator equation $Tx = b$ in $L^2([a,b])$ via the EPM.* Let $\Omega = [a,b]$ be a real interval provided with the $\sigma-$algebra $\mathcal{A}$ of all its Lebesgue measurable subsets and with the probability measure $\mu = (b-a)^{-1} \cdot \lambda$, where λ stands for the Lebesgue measure on $\mathcal{A}$. Denote $H := \mathcal{L}^2([a,b])$, the Hilbert space of all square integrable functions on Ω with the inner product

$$< x, y >= \int_\Omega x(\omega) \cdot y(\omega) d\mu(\omega).$$

If $\xi : H \to \mathbf{R}$ is a linear continuous function, then we denote by ξ^* the unique element of H satisfying $\|\xi\|_* = \|\xi^*\|$ and $< \xi^*, x >= \xi(x)$, for any $x \in H$ (such a ξ^* exists by Riesz's theorem).

Let $T : H \to H$ be a linear operator and $b \in H$. We consider the (linear operator) equation

$$Tx = b. \tag{12}$$

This equation can be represented as a SCFP and solved by the EPM. Precisely, we have the following result:

Corollary. *If, for each $\omega \in \Omega$, the linear function $T_\omega : H \to \mathbf{R}$ given by $T_\omega x = (Tx)(\omega)$, is continuous and has $\|T_\omega\|_* \neq 0$, and if the function $\zeta : \Omega \to \mathbf{R}$ defined by*

$$\zeta(\omega) = \frac{b(\omega)}{\|T_\omega\|_*}, \tag{13}$$

is integrable, then

(*a*) *The point-to-set mapping* $Q : \Omega \to H$ *defined by*

$$Q_\omega := \{z \in H; T_\omega z = b(\omega)\}\,, \tag{14}$$

has closed, convex, nonempty values and is measurable;

(*b*) *For any pair* $(\omega,\, x) \in \Omega \times H$, *the metric projection of* x *onto the set* Q_ω *is given by*

$$P_\omega x = x + \frac{b(\omega) - T_\omega x}{\|T_\omega\|_*^2} \cdot T_\omega^*; \tag{15}$$

(*c*) *For each* $x \in H$, *the function* $\omega \to P_\omega x : \Omega \to H$ *is integrable;*

(*d*) *An EPM generated sequence* $\left\{x^k\right\}_{k\in\mathbf{N}}$ *with respect to* Q *converges weakly if and only if the linear operator equation*

$$\int_\Omega \frac{b(\omega) - T_\omega x}{\|T_\omega\|_*^2} \cdot T_\omega^* = 0, \tag{16}$$

has at least one solution in H*;*

(*e*) *If the equation* (16) *has solutions, then any EPM generated sequence with respect to* Q *converges weakly to a solution of* (16)*;*

(*f*) *An element* $x^* \in H$ *is a solution of the equation* (12) *if and only if* x^* *is an almost common point of the sets* Q_ω, $\omega \in \Omega$*;*

(*g*) *If the equation* (12) *has solutions, then any EPM generated sequence with respect to* Q *converges weakly to a solution of it.*

Proof. The fact that each set Q_ω is convex, closed and nonempty is obvious. Note that the function $g : \Omega \times H \to \mathbf{R}$ defined by $g(\omega, x) = T_\omega x - b(\omega)$, has the properties that, for each $x \in H$, $g(\cdot, x)$ is measurable and, for each $\omega \in \Omega$, $g(\omega, \cdot)$ is continuous. Thus applying to g the Theorem 8.2.9 in [1] we deduce that the point-to-set mapping Q is measurable. Hence, (*a*) is proven. For an arbitrary $x \in H$, denote by z_ω the right hand side of the equation (15). Observe that

$$T_\omega z_\omega = b(\omega),$$

and that, for any $z \in H$,

$$< x - z_\omega, z_\omega - z >= 0.$$

These show that $P_\omega x = z_\omega$ is the metric projection of x onto Q_ω, i.e., (*b*) is proven. In order to prove (*c*) observe that, according to (*a*), (*b*) and to Corollary 8.2.13 in [1], for every $x \in H$, the function $\omega \to P_\omega x$ is measurable. Also, we have

$$\int_\Omega \|P_\omega(0)\|\, d\mu(\omega) = \int_\Omega |\zeta|\, d\mu(\omega),$$

where ζ is the function defined by (13). This implies that the function $\omega \to P_\omega(0)$ is integrable. If x is any element of H, then

$$\textstyle\int_\Omega \|P_\omega x\| \, d\mu(\omega) \leq \int_\Omega \|P_\omega x - P_\omega(0)\| \, d\mu(\omega) + \int_\Omega \|P_\omega(0)\| \, d\mu(\omega)$$

$$\leq \|x\| + \textstyle\int_\Omega \|P_\omega(0)\| \, d\mu(\omega),$$

proving that the function $\omega \to P_\omega x$ is integrable. Combining (a), (b) and (c) we deduce that the expected projection operator $\mathbf{P}$ with respect to the point-to-set mapping Q is well-defined and for any $x \in H$,

$$\mathbf{P}x = x + \int_\Omega \frac{b(\omega) - T_\omega x}{\|T_\omega\|_*^2} \cdot T_\omega^* d\mu(\omega). \tag{17}$$

Therefore, application of Theorem 2.2(A) yields (d) and (e). According to (14), the element $x^* \in H$ is a solution of (12) if and only if $x^* \in \bigsqcap_\mu Q_\omega$, i.e. (f) holds. By consequence, application of Remark 2.3(jj) implies (g). $\square$

3.3 *Finding solutions to convex programming problems with the EPM.* Consider the optimization problem (P) : Find $x \in X$ such that

$$f(x) = \inf\{f(y); y \in X\},$$

where $f : \mathbf{R}^n \to \mathbf{R}$ is convex and continuously differentiable and $X \subset \mathbf{R}^n$ is convex, compact and has nonempty interior. Clearly, the problem (P) has at least one optimal solution. Recall (see, for instance, [2, Section 3.5]) that (P) can be equivalently rewritten in variational form: Find $x \in X$ such that, for each $y \in X$,

$$< x - y, \nabla f(y) >\leq 0. \tag{18}$$

In this way, the optimization problem (P) can be transformed into an equivalent SCFP where $\Omega = X$, $\mathcal{A}$=the family of all Lebesgue measurable subsets of X, $\mu = (\lambda(X))^{-1} \cdot \lambda$ with λ being the Lebesgue measure on X, and $Q : \Omega \to \mathbf{R}^n$ is defined by

$$Q_y = \{x \in X; < x - y, \nabla f(y) >\leq 0\}. \tag{19}$$

Note that the point-to-set mapping Q is measurable because its graph is closed. Any common point of the sets Q_y, $y \in X$, satisfies (18) and, thus, is an optimal solution of (P). Moreover, we have that $\bigcap_{y \in X} Q_y = \bigsqcap_\mu Q_y$ because, if $x \in \bigsqcap_\mu Q_y$ and if, for some $y^* \in X$,

$$< x - y^*, \nabla f(y^*) >> 0,$$

then there exists an open neighborhood N of y^* such that, for all $y \in N$,

$$< x - y, \nabla f(y) >> 0,$$

and this is a contradiction because $\mu(N \cap X) > 0$. Hence, finding an almost common point of the sets Q_y, $y \in X$, amounts to finding an optimal solution of (P). Combining these facts, Theorem 2.2 and the well-known formula of the orthogonal projection on a half-space in $\mathbf{R}^n$ we get the next result which describes an iterative method of approximating solutions of (P).

Corollary. *Suppose that the function $f : \mathbf{R}^n \to \mathbf{R}$ is a convex and continuously differentiable and that X is a convex compact subset of $\mathbf{R}^n$ such that $Int(X) \neq \emptyset$. Then, for any initial point $x^0 \in \mathbf{R}^n$, the sequence recursively defined by*[3]

$$x^{k+1} = x^k + \frac{1}{\lambda(X)} \cdot \int_X \min\left[0, \frac{< x^k - y, \nabla f(y) >}{\|\nabla f(y)\|^2}\right] \cdot \nabla f(y) d\lambda(y) \tag{20}$$

converges to a minimizer of the function f over X.

In general, computing the multiple integrals involved in (20) may be difficult even if we are able to determine explicit formulae for f and ∇f. However, in practical problems, X is usually given as the solution set of a system of inequalities

$$g_i(x) \leq 0, \quad 1 \leq i \leq m,$$

where, for each i, the function $g_i : \mathbf{R}^n \to \mathbf{R}$ is convex. In such a case, one can apply the algorithm described by in [4, Section 4] for computing the triangulation $\mathcal{T}_1$ as defined there for some small number $\varepsilon > 0$. This triangulation $\mathcal{T}_1$ is a collection of simplices of diameter less that the given ε, included in X, and whose union form a convex polytope with Hausdorff distance to X no greater than ε and having all its vertices in ∂X. Therefore, the Lebesgue sum

$$S_\varepsilon^k = \sum_{\tau \in \mathcal{T}_1} F_k(bar(\tau)) \cdot \lambda(\tau),$$

where $F_k(y)$ stands for the integrand in (20) and $bar(\tau)$ denotes the barycenter of the simplex τ, approximates the integral in (20) with an error which converges to zero as $\varepsilon \searrow 0$.

Acknowledgment

The author is grateful to Professor Alfredo Iusem for comments helping to improve an earlier version of this work.

[3] Here we make the usual convention that the integrand in the next equation is zero whenever $\nabla f(y) = 0$.

References

[1] J.-P. Aubin and H. Frankowska: Set-Valued Analysis, *Birkhäuser*, Basel, 1990.

[2] D.P. Bertsekas and J.N. Tsitsiklis: Parallel and Distributed Computation: Numerical Methods, *Prentice-Hall International*, 1993.

[3] L.M. Bregman: The method of successive projections for finding a common point of convex sets, *Dokl. Acad. Nauk SSSR*, 162, 3, 1965, 688-692.

[4] D. Butnariu and Y. Censor: A method for approximating the solution set of a system of convex inequalities by polytopes, *J. Computational and Applied Mathematics*, 36, 1991, 289-304.

[5] D. Butnariu and S.D. Flåm: Strong convergence of the expected projection method in Hilbert spaces, *J. Numer. Funct. Anal. Optim.*, to appear.

[6] Y. Censor and S. Zenios: Introduction to Methods of Parallel Optimization, *Instituto de Matematica Pura e Applicada*, Rio de Janiero, Brazil, 1993.

[7] G. Cimmino: Calcolo approsimato per le solutioni di sistemi di equationi lineari, *La Riserca Scientifica, Roma*, XVI, Anno IX, 2, 1938, 326-333.

[8] P.L. Combettes: Construction of a common fixed point of firmly nonexpansive mappings, *C.R. Acad. Sci. Paris*, t. 320, Série I, 1995, in publication.

[9] K. Goebel and S. Reich: Uniform convexity, hyperbolic geometry, and nonexpansive mappings, *Marcel Dekker*, New York, 1984.

[10] A. Iusem and A.R. De Pierro: Convergence results for an accelerated nonlinear Cimmino algorithm, *Numer. Math.*, 49, 1986, 367-378.

[11] W.J. Kammerer and M.Z. Nashed: Iterative methods for best approximate solutions of linear integral equations of the first and second kind, *J. Math. Anal. Appl.*, 40, 1972, 547-573.

[12] J. Mikusinski: The Bochner Integral, *Academic Press*, New York, 1978.

[13] G. Pierra: Decomposition trough formalization in a product space, *Math. Progr.*, 28, 1984, 96-115.

[14] R.R. Phelps: Convex Functions, Monotone Operators and Differentiability, *Springer Verlag*, Berlin, 1993.