Finding Fixed Points by Averaging with Well-Behaved Maps

by

M. Epelman,
T.L. Magnanti and
G. Perakis

OR 322-97 October 1997

Finding Fixed Points by Averaging with Well-Behaved Maps

Marina Epelman,* Thomas L. Magnanti† and Georgia Perakis‡

October 29, 1997

Abstract

We introduce and examine a general framework for parametrically combining a well-behaved map g with a given map T into a composite map F_λ in order to find a fixed point of T. As special cases, our framework includes outside averaging $F_\lambda = \lambda g + (1-\lambda)T$ and inside averaging $F_\lambda = T(\lambda g + (1-\lambda)I)$. We establish conditions under which the trajectory of the sequence of fixed points x_λ to the map F_λ converge to a fixed point of T as λ approaches zero. We also establish convergence conditions for the iterative scheme $x_{k+1} = F_{\lambda_{k+1}}(x_k, T(x_k))$ which approximates the parametric fixed point trajectory x_λ.

Key Words:

Fixed Point Problems, Averaging Schemes, Nonexpansive Maps, Strongly-f-Monotone Maps.

*Operations Research Center, MIT, Cambridge, MA 02139.

†Sloan School of Management and Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA 02139.

‡Operations Research Center, MIT, Cambridge, MA 02139.

1 Introduction

Algorithms for solving problems in many applied settings establish a mapping T whose iterative application leads to a fixed point solution that solves the original problem. The convergence of this iterative procedure to a fixed point solution often requires strong assumptions on the algorithmic map T that restrict the algorithm's domain of applicability. In many instances, various forms of averaging will extend the range of applicability of the algorithmic (fixed point) map. To remedy this behavior, in this paper, we introduce a general averaging framework for solving the fixed point problem

$$\text{Find } \; x^* \in K \; \text{ satisfying } \; x^* = T(x^*) \tag{1}$$

defined over a given ground set $K \subseteq R^n$ when $T : K \rightarrow K$. We will let $FP(T)$ denote the set of fixed point solutions of (1) which we will assume to be nonempty.

The key idea that we explore is to average the fixed point (algorithmic) map T with maps g that are "well" behaved in order to counteract the "bad" properties of the map T. In particular, we will introduce a general averaging framework which includes as special cases averaging with two types of well-behaved maps: identity and contractive maps. By allowing us to combine these two types of maps in various ways, the framework will establish convergence for several forms of averaging. Our framework will include "outside averaging" as in $\lambda g(x) + (1 - \lambda)T(x)$. Halpern [8] and Browder [3] introduced this type of averaging and Bauschke [1] and Wittmann [14] studied it further in the special case when $g(x) =$ constant. Dunn [5] introduced the special case of $g(x) = x$, which Magnanti and Perakis further studied for variational inequality problems (see [10] and [11]). Our framework will also include "inside averaging" as in $T(\lambda g(x) + (1 - \lambda)x)$, with $0 \leq \lambda \leq 1$. It will permit averaging with the identity map (line search procedures), or outside and inside averaging with contractive maps as well as with the proximal point map. When applied to fixed point and variational inequality problems, this framework gives rise to certain known methods as well as several new ones.

Preliminaries:

Throughout our analysis, we will use a few elementary results concerning the relationship between a given map $G : K \to K$ and the map $I - G$.

Proposition 1 : *([4])*

I) A map G is contractive with contraction constant $a \in (0,1)$ if and only if $I - G$ is a strongly monotone map with constant $\frac{1-a}{2}$, that is,

$$(x - G(x) - y + G(y))^t(x - y) \geq \frac{1-a}{2}\|x - y\|^2 + \frac{1}{2}\|x - G(x) - y + G(y)\|^2, \forall x, y \in K.$$

II) A map G is nonexpansive if and only if $I - G$ is a strongly-f-monotone map with constant $\frac{1}{2}$, that is,

$$(x - G(x) - y + G(y))^t(x - y) \geq \frac{1}{2}\|x - G(x) - y + G(y)\|^2, \forall x, y \in K.$$

We will also use the following result.

Proposition 2 : *Suppose $K \subseteq R^n$ is a convex compact set and $G : K \to K$ is a nonexpansive map. Suppose further that $S \subseteq K$ is a closed convex set. Then some point $s^* \in S$ satisfies the condition*

$$(G(s^*) - s^*)^t(s - s^*) \leq 0 \quad \textit{for all} \quad s \in S.$$

Furthermore, if G is a contractive map, the point s^* is uniquely defined.

Proof: Suppose first that the mapping G is nonexpansive. Consider the mapping $\bar{G} : S \to S$: $\bar{G}(s) = \mathrm{Pr}_S[G(s)]$, defined by the Euclidean projection $\mathrm{Pr}_S[\cdot]$ onto the set S. This mapping satisfies the following inequality

$$\|\bar{G}(s_1) - \bar{G}(s_2)\| = \|\mathrm{Pr}_S[G(s_1)] - \mathrm{Pr}_S[G(s_2)]\| \leq \|G(s_1) - G(s_2)\| \leq \|s_1 - s_2\|.$$

The first and second inequalities follow from the nonexpansiveness of the projection operator and $G(\cdot)$, respectively. We conclude that the mapping $\bar{G}$ is nonexpansive and since its domain S is a compact convex set, this set contains a fixed point s^* of $\bar{G}$, (see, for example, [2]).

When the map G is contractive, a similar argument shows that the map $\bar{G}$ is also contractive, and therefore has a unique fixed point $s^* \in S$. Q.E.D.

Observe that this result says that the point s^* is a solution to a variational inequality problem defined on S with the mapping $s - G(s)$. Equivalently, s^* is a fixed point of the projection $Pr_S[G(s)]$ of $G(s)$ on S.

2 Averaging Trajectories.

2.1 Generalized Averaging Map.

As a first step to our analysis, we approximate the fixed point problem with the parameterized fixed point problem

$$x_\lambda = F_\lambda(x_\lambda, T(x_\lambda)). \tag{2}$$

What type of parameterized functions F_λ can we consider? Can we develop a characterization that will permit us to consider as special cases outside and inside averaging as is $F_\lambda(x_\lambda, T(x_\lambda)) = \lambda g(x_\lambda) + (1-\lambda)T(x_\lambda)$ and $F_\lambda(x_\lambda, T(x_\lambda)) = T(\lambda g(x_\lambda) + (1-\lambda)x_\lambda)$? Toward this purpose, for a set of parameters $\lambda \in \Lambda \subseteq R$, we consider a function $F_\lambda(x, T(x)) : \Lambda \times K \times K \to K$ that satisfies the following conditions:

A1 For all $\lambda \in \Lambda = [0,1]$, the parameterized problem $x_\lambda = F_\lambda(x_\lambda, T(x_\lambda))$ has a solution.

A2 $T(x)$ is a continuous map of x. The function $F_\lambda(x, T(x))$ is a continuous function of λ and x.

A3 $FP(F_0) = FP(T)$.

We refer to any map $F_\lambda(\cdot,\cdot)$ satisfying conditions A1–A3 as a *generalized averaging map*. Assume in addition that

A4 K is a closed, convex and bounded feasible region.

Remark:

Observe that conditions A4, A2 imply condition A1.

For situations satisfying these four assumptions, we can characterize the limit points of the sequence $\{x_\lambda\}$ induced by the parametrized fixed point problems (2).

Theorem 1 : *If conditions A1–A4 are valid and if $\lambda \to 0$, then every limit point of the sequence $\{x_\lambda\}$ is a fixed point solution of the map T.*

Proof:

Condition A1 states that the iterates induced by (2) exist. Condition A4 implies that the sequence $\{x_\lambda\}$ has limit points. Finally, conditions A2 and A3 imply that as λ approaches zero, the limit points of $\{x_\lambda\}$ are fixed point solutions of the original fixed point problem (1). Q.E.D.

The previous analysis has only characterized the limit points of the sequence of solutions of the parametric family of fixed point subproblems as solutions of the original problem. In order to further understand the convergence of the entire sequence as well as the nature of its limit points, we need to impose some additional conditions on the maps T and F_λ.

A5 The map T is nonexpansive.

A6 The map F_λ is of the form $F_\lambda = F(G_\lambda)$, with $F : K \to K$, $G_\lambda : K \times K \to K$ continuous functions of x, and, $\lambda \in \Lambda = [0, 1]$. In addition,

- F is a nonexpansive map (as a function of x) and $FP(F) \supseteq FP(T)$

- G_0 is a nonexpansive map (as a function of x) and $FP(G_0) \supseteq FP(T)$
- $\lim_{\lambda \to 0} \frac{\|G_\lambda(x_\lambda, T(x_\lambda)) - x_\lambda\|^2}{\lambda} = 0$

A7 $G_\lambda(x, T(x))^t(x - x^*) \leq \lambda G_1(x, T(x))^t(x - x^*) + (1 - \lambda)G_0(x, T(x))^t(x - x^*)$,

for all $\lambda \in [0, 1]$, all $x \in K$ and for all $x^* \in FP(T)$.

A8a $G_1(x, T(x))$ is a nonexpansive map as a function of x, or

A8b $G_1(x, T(x))$ is a contractive map as a function of x with contraction constant $a \in (0, 1)$.

Remarks:

(i) Conditions A4 and A5 together imply that the set $FP(T)$ is nonempty. Condition A6 implies that $FP(T) \subseteq FP(F_0)$ since if $x^* \in FP(T)$ then $F_0(x^*) = F(G_0(x^*)) = F(x^*) = x^*$ from the set of inclusions in condition A6.

(ii) Usually, we let $F = T$ and $G_0 = I$ or $F = I$ and $G_0 = T$. In both cases, these mappings satisfy the set containment conditions in assumption A6.

By invoking these additional conditions, we can further characterize the limit points of the sequence induced by parametric fixed point problem (2).

Theorem 2 *:*

I) Suppose conditions A1–A8a are valid. Then as $\lambda \to 0$, *every limit point* $\bar{x}$ *of the sequence* $\{x_\lambda\}$ *is a fixed point solution of* $FP(T, K)$ *satisfying the condition*

$$\bar{x} \in FP(T): \quad (\bar{x} - G_1(\bar{x}, T(\bar{x})))^t(x - \bar{x}) \geq 0, \quad \text{for all} \quad x \in FP(T).$$

II) Suppose conditions A1–A8b are valid. Then as $\lambda \to 0$, *the sequence* $\{x_\lambda\}$ *converges to a fixed point solution* x^* *of* $FP(T, K)$ *satisfying the condition*

$$x^* \in FP(T): \quad (x^* - G_1(x^*, T(x^*)))^t(x - x^*) \geq 0, \quad \text{for all} \quad x \in FP(T).$$

Proof:

By condition A4, K is a compact convex set. The set of fixed points $FP(T)$ is known to be a closed convex set. Therefore, Proposition 2 implies that under conditions A8a or A8b, some point $x^* \in FP(T)$ satisfies the condition $(x^* - G_1(x^*, T(x^*)))^t(x - x^*) \geq 0$ for all $x \in FP(T)$.

I) Let us first assume that conditions A1–A8a hold. If $y_\lambda = G_\lambda(x_\lambda, T(x_\lambda))$, then, by the definition of x_λ, $x_\lambda = F(y_\lambda)$. Since $F(\cdot)$ is nonexpansive (assumption A6) and $FP(T) \subseteq FP(F)$, x^* is a fixed point of $F(\cdot)$ and

$$0 \leq ((y_\lambda - F(y_\lambda)) - (x^* - F(x^*)))^t (y_\lambda - x^*) = (y_\lambda - F(y_\lambda))^t (y_\lambda - x^*)$$

$$= (y_\lambda - x_\lambda)^t (y_\lambda - x^*) = (y_\lambda - x_\lambda)^t (y_\lambda - x_\lambda) + (y_\lambda - x_\lambda)^t (x_\lambda - x^*) \qquad (3)$$

$$= \|G_\lambda(x_\lambda, T(x_\lambda)) - x_\lambda\|^2 - (x_\lambda - G_0(x_\lambda, T(x_\lambda)))^t (x_\lambda - x^*) + (G_\lambda(x_\lambda, T(x_\lambda)) - G_0(x_\lambda, T(x_\lambda)))^t (x_\lambda - x^*).$$

Condition A7 implies that for all $\lambda \in [0, 1]$,

$$G_\lambda(x_\lambda, T(x_\lambda))^t (x_\lambda - x^*) \leq \lambda G_1(x_\lambda, T(x_\lambda))^t (x_\lambda - x^*) + (1 - \lambda) G_0(x_\lambda, T(x_\lambda))^t (x_\lambda - x^*),$$

or, upon rearranging terms,

$$(G_\lambda(x_\lambda, T(x_\lambda)) - G_0(x_\lambda, T(x_\lambda)))^t (x_\lambda - x^*) \leq \lambda (G_1(x_\lambda, T(x_\lambda)) - G_0(x_\lambda, T(x_\lambda)))^t (x_\lambda - x^*).$$

Substituting this inequality into (3), we conclude that for all $\lambda \in [0, 1]$,

$$0 \leq \|G_\lambda(x_\lambda, T(x_\lambda)) - x_\lambda\|^2 - (x_\lambda - G_0(x_\lambda, T(x_\lambda)))^t (x_\lambda - x^*) + \lambda (G_1(x_\lambda, T(x_\lambda)) - G_0(x_\lambda, T(x_\lambda)))^t (x_\lambda - x^*)$$

$$= \|G_\lambda(x_\lambda, T(x_\lambda)) - x_\lambda\|^2 - (1 - \lambda)(x_\lambda - G_0(x_\lambda, T(x_\lambda)))^t (x_\lambda - x^*) + \lambda (G_1(x_\lambda, T(x_\lambda)) - x_\lambda)^t (x_\lambda - x^*).$$

Since $x^* \in FP(T) \subseteq FP(G_0)$ and G_0 is a nonexpansive map (from condition A6),

$$(x_\lambda - G_0(x_\lambda, T(x_\lambda)))^t (x_\lambda - x^*) \geq 0.$$

Therefore, for all $\lambda \in (0, 1]$,

$$\frac{\|G_\lambda(x_\lambda, T(x_\lambda)) - x_\lambda\|^2}{\lambda} \geq (x_\lambda - G_1(x_\lambda, T(x_\lambda)))^t (x_\lambda - x^*).$$

Condition A6 implies that $\lim_{\lambda \to 0} (x_\lambda - G_1(x_\lambda, T(x_\lambda)))^t (x_\lambda - x^*) \leq 0$.

Therefore as $\lambda \to 0$, every limit point $\bar{x}$ (which is also a fixed point solution of the map T as we have shown in Theorem 1) satisfies the condition

$$(\bar{x} - G_1(\bar{x}, T(\bar{x})))^t(\bar{x} - x^*) \le 0. \tag{4}$$

Consequently, condition A8a and Proposition 1 imply that

$$(\bar{x} - G_1(\bar{x}, T(\bar{x})) - x^* + G_1(x^*, T(x^*)))^t(\bar{x} - x^*) \ge \frac{1}{2}\|\bar{x} - G_1(\bar{x}, T(\bar{x})) - x^* + G_1(x^*, T(x^*))\|^2.$$

Moreover, the inequality (4) and the definition of x^* imply that

$$(\bar{x} - G_1(\bar{x}, T(\bar{x})) - x^* + G_1(x^*, T(x^*)))^t(\bar{x} - x^*) \le 0.$$

Therefore, $\bar{x} - G_1(\bar{x}, T(\bar{x})) = x^* - G_1(x^*, T(x^*))$. Consequently, the definition of x^* and (4) imply that, for all $x \in FP(T)$,

$$(\bar{x} - G_1(\bar{x}, T(\bar{x})))^t(x - \bar{x}) = (x^* - G_1(x^*, T(x^*)))(x - x^*) + (\bar{x} - G_1(\bar{x}, T(\bar{x})))^t(x^* - \bar{x}) \ge 0,$$

establishing the first part of the theorem.

II) Let us now assume that conditions A1–A8b hold instead.

As in the previous argument, as $\lambda \to 0$, every limit point $\bar{x} \in FP(T)$ and satisfies the condition

$$(\bar{x} - G_1(\bar{x}, T(\bar{x})))^t(\bar{x} - x^*) \le 0. \tag{5}$$

Moreover, since G_1 is a contractive map with contraction constant a, condition A8b and Proposition 1 imply that

$$(\bar{x} - G_1(\bar{x}, T(\bar{x})) - x^* + G_1(x^*, T(x^*)))^t(\bar{x} - x^*) \ge \frac{1-a}{2}\|\bar{x} - x^*\|^2.$$

Therefore, inequality (5) and the definition of x^* imply that $\bar{x} = x^*$. Observe that the contractiveness of map G_1 implies that the point x^* is uniquely defined. Therefore, the entire sequence $\{x_\lambda\}$ converges to the point $x^* = Pr_F G_1(x^*, T(x^*)) \in FP(T)$. $\square$

This theorem states that any limit point $\bar{x}$ of the sequence $\{x_\lambda\}$ solves the variational inequality with the mapping $I - G_1$ over the general set $FP(T)$. In the special case when

$G_1(x) = g(x) = ax$ for some $0 < a < 1$, the condition becomes $(1-a)\bar{x}^t(x - \bar{x}) \geq 0$ for all $x \in FP(T)$ which is the optimality condition for the minimum norm problem defined on $FP(T)$. Therefore, the parametric fixed points converge to this minimum norm point.

To illustrate the type of limiting conditions we have developed, let us consider a numerical example.

Example:

Consider the fixed point problem $FP(T, [0,2] \times [-1,1])$ with the map $T(x) = x$, if $x_1 \geq 1$ and $T(x) = (2 - x_1, x_2)$, if $x_1 < 1$. The set of fixed point solutions is $FP(T) = \{(1, x_2) : x_2 \in [-1,1]\}$. Consider the contractive map $G_1(x) = g(x) = \frac{x}{2}$. Its fixed point solution is the point $x^* = (0,0)$. For this example, the averaging trajectory $x_\lambda^1 = \lambda g(x_\lambda^1) + (1-\lambda)T(x_\lambda^1)$ is $x_\lambda^1 = (\frac{4-4\lambda}{4-3\lambda}, 0)$. Similarly, the trajectory $x_\lambda^2 = T(\lambda g(x_\lambda^2) + (1-\lambda)x_\lambda^2)$ is $x_\lambda^2 = (\frac{4}{4-\lambda}, 0)$. As we have shown in Theorem 2, as λ approaches zero, the parametric fixed points in both cases converge to the fixed point solution with the minimum norm, that is, $x^* = Pr_F g(x^*, T(x^*)) = (1,0)$. Observe that although the trajectories x_λ^1 and x_λ^2 in this example are quite different, they both approach the same limit point.

In the next section we will show that indeed these types of averaging trajectories are special cases of the averaging trajectories we have considered in this section.

2.2 Examples of generalized averaging maps.

We next provide several examples of maps $F_\lambda(\cdot,\cdot)$ and show that they are special cases of the averaging framework we have introduced. Again, we assume that the ground set K is a convex compact set, the map $T : K \to K$ is nonexpansive and the map $g : K \to K$ is contractive. These special cases use various forms of inside and outside averaging. The fact that the fixed point map T is nonexpansive allows us to apply the proximal point map $(I + c(I - T))^{-1}$ for some constant c to the map $(I - T)$ in some of these examples. For a discussion of the proximal point map and its relationship to fixed point problems, see Eckstein and Bertsekas [6] and Rockafellar [13].

A Outside averaging with a contractive map

$F_\lambda(x, T(x)) = F_\lambda^{out}(x, T(x)) \stackrel{def}{=} \lambda g(x) + (1 - \lambda)T(x).$

In this case, $F = I$ and $G_\lambda = \lambda g + (1 - \lambda)T$.

Halpern [8] and Browder [3] introduced, and Bauschke [1] and Wittmann [14] further studied, a special case of this map (with the constant map $g(x) = c$ for some constant c).

B Inside averaging with a contractive map

$F_\lambda(x, T(x)) = F_\lambda^{in}(x, T(x)) \stackrel{def}{=} T(\lambda g(x) + (1 - \lambda)x).$

In this case, $F = T$ and $G_\lambda = \lambda g + (1 - \lambda)I$.

C Outside averaging with the proximal point map

$$F_\lambda(x, T(x)) = (1 - \lambda)J_{c(I-T)}(x) \stackrel{def}{=} (1 - \lambda)(I + c(I - T))^{-1}(x).$$

In this case, $F = I$ and $G_\lambda = (1 - \lambda)J_{c(I-T)}$.

D Inside averaging with the proximal point map

$F_\lambda(x, T(x)) = J_{c(I-T)}((1 - \lambda)x) \stackrel{def}{=} (I + c(I - T))^{-1}((1 - \lambda)x).$

In this case, $F = J_{c(I-T)}$ and $G_\lambda = (1 - \lambda)I$.

Note that method **C** (**D**) is analogous to outside (inside) averaging of the map $J_{c(I-T)}$ with $g \equiv 0$. Furthermore, like the methods **A** and **B**, we can apply averaging with a more general contractive map g (or a family of contractive maps g_λ as in method **F** below).

E Convex combinations of any F_λ maps that satisfy the conditions A1-A8a or A1-A8b

F Averaging with a family of contractive maps

In examples **A**–**E**, we can replace the map g with a family of maps $g_\lambda : K \to K$ if for some constants $\alpha \in [0, 1)$ and $\gamma \geq 0$:

$$||g_\lambda(x) - g_\lambda(y)|| \leq \alpha||x - y|| \ \forall x, y \in K, \ \forall \lambda \in [0, 1]$$

and

$$||g_{\lambda^1}(x) - g_{\lambda^2}(x)|| \leq \gamma|\lambda^1 - \lambda^2|, \ \forall x \in K, \ \forall \lambda^1, \lambda^2 \in [0,1].$$

Validity of the algorithm schemes **A**–**E**:

We next show that the maps **A**–**F** satisfy the conditions A1–A8.

- Conditions A4 and A5 follow by assumption.
- In all six cases, it is easy to see that the map $F_\lambda(x, T(x))$ is nonexpansive. Therefore, condition A4 implies condition A1. Furthermore, for $\lambda \in (0,1)$, $F_\lambda(x, T(x))$ is a contractive map.
- $||F_\lambda(x, T(x)) - F_{\bar{\lambda}}(x, T(x))|| \leq |\lambda - \bar{\lambda}|L(x)$ for some appropriately defined continuous function $L(x)$ of x for examples **A**–**F**.
  Therefore, F_λ is continuous as a function of λ.
- Conditions A2 and A3 follow from the corresponding definitions of $F_\lambda(x, T(x))$.
- We can easily check that in all six cases that the map $F_\lambda = F(G_\lambda)$ satisfies the properties of condition A6.
- $G_\lambda(x, T(x))^t(x - x^*) = \lambda G_1(x, T(x))^t(x - x^*) + (1 - \lambda)G_0(x, T(x))^t(x - x^*)$ in all of the examples **A**–**F**.
- In examples **A** and **B**, $G_1(x, T(x)) = g$, while in example **F**, $G_1(x, T(x)) = g_1$. In each case, G_1 is a contractive map. In examples **C** and **D**, $G_1(x, T(x)) = 0$, a nonexpansive map. In example **E**, the specific form of $G_1(x, T(x))$ will depend upon a particular convex combination being considered, but it is easy to see that in any situation it will be nonexpansive.

Comparing Some Averaging Trajectories

To complete this section, we compare the trajectories generated through some of the previous examples.

Consider the trajectories of the fixed points generated from the maps $F_\lambda^A(x) = (1-\lambda)T(x)$ and $F_\lambda^C(x) = (1-\lambda)J_{c(I-T)}(x)$, for $\lambda \in [0,1)$. The map $F_\lambda^A(x)$ is a special case of the map $F_\lambda^{out}(x)$ with $g(x) = 0$. Let

$$x_\lambda^A = (1-\lambda)T(x_\lambda^A) \quad \text{and} \quad x_\lambda^C = (1-\lambda)J_{c(I-T)}(x_\lambda^C) = (1-\lambda)((1+c)I - cT)^{-1}(x_\lambda^C).$$

We compare the two trajectories by noting that

$$x_\lambda^C = ((1+c)I - cT)\left(\frac{x_\lambda^C}{1-\lambda}\right) = \frac{(1+c)}{1-\lambda}x_\lambda^C - cT\left(\frac{x_\lambda^C}{1-\lambda}\right).$$

Therefore, $\frac{x_\lambda^C}{1-\lambda} = \frac{c}{c+\lambda}T\left(\frac{x_\lambda^C}{1-\lambda}\right)$ implying that

$$x_\lambda^C = (1-\lambda)x^A_{\frac{\lambda}{c+\lambda}}.$$

We might also consider the trajectories of fixed points generated from the maps $F_\lambda^B(x) = T((1-\lambda)x)$ and $F_\lambda^D(x) = ((c+1)I - cT)^{-1}((1-\lambda)x)$. The map $F_\lambda^B(x)$ is a special case of the map $F_\lambda^{in}(x)$ with $g(x) = 0$. As above, we can show that

$$x_\lambda^D = x^B_{\frac{\lambda}{c+\lambda}}.$$

Notice that in this case the trajectory of the fixed points generated by inside averaging with the proximal point mapping (that is, the sequence $\{x_\lambda^D\}$) is the same as the trajectory of the fixed points generated by inside averaging (that is, the sequence $\{x_\lambda^B\}$). Observe that whenever $c \geq 1$, the sequence $\{x_\lambda^D\}$ converges faster than the sequence $\{x_\lambda^B\}$.

3 Approximate Averaging Trajectories

The averaging framework described in the previous section provides an intuitive way of approximating fixed point solutions of a map T with the trajectory of fixed points solutions for a class of parameterized subproblems. Following this trajectory exactly as well as computing the fixed points of the subproblems typically will be very expensive computationally. Therefore, we would like to be able to approximate these trajectories.

We suggest two iterative schemes of approximating the generalized averaging trajectory. Both include as a special case outside averaging with a constant map as previously considered in the literature (see, for example, Halpern [8] and Bauschke [1]). The framework here, however, extends well beyond this very special case.

3.1 A General Averaging Framework

We consider the following approximation,

$$x_{k+1} = F_{\lambda_{k+1}}(x_k, T(x_k)), \tag{6}$$

which takes a single step towards solving the parameterized fixed point problem (2) and then alters the value of λ.

We consider a function $F_\lambda(x, T(x))$ defined on the set $\Lambda \times K \times K \to K$. Suppose, in addition, the function satisfies the following conditions:

B1 $FP(F_0) = FP(T)$.

B2 $F_\lambda(x, T(x)) = f_\lambda(x)$ is a continuous function of x and λ and $\|F_\lambda(x, T(x)) - F_\lambda(y, T(y))\| \leq \beta(\lambda)\|x - y\|$, with $\beta(\lambda) \geq 0$, $0 < \beta(0) \leq 1$ and $\beta'(0) < 0$ and $\beta(\lambda) \leq \beta(0) + \beta'(0)\lambda$.

B3 $\|F_\lambda(x, T(x)) - F_{\bar{\lambda}}(x, T(x))\| \leq |\lambda - \bar{\lambda}|L(x)$ for some continuous function $L : K \to R_+$.

Note that condition B2 is valid if $\beta''(\lambda) \leq 0$, that is, $\beta(\lambda)$ is a concave function.
Consider stepsizes λ_k satisfying:

B4 $\lambda_k \to 0$.

B5 $\sum_k \lambda_k = +\infty$.

B6 $\sum_k |\lambda_k - \lambda_{k+1}| < +\infty$.

Wittmann [14] has used the same choice of stepsizes for the special case of outside averaging with a constant map. Finally, assume

B7 $L(x)$ is bounded over the feasible set K (for example, when the set K is bounded).

Remark:

We can view conditions B1-B3 as relaxed versions of conditions A1-A3. Observe that when $F_0(x, T(x)) = T$, the assumption $0 < \beta(0) \leq 1$ captures the nonexpansiveness of the map T as in condition A5. The conditions $\beta'(0) < 0$ and $\beta(\lambda) \leq \beta(0) + \beta'(0)\lambda$ imply that $\beta(\lambda) < 1$ for all $\lambda \in (0, 1]$. Consequently, F_λ is a contractive map for every fixed $\lambda \in (0, 1]$ and F_λ is a nonexpansive map for every fixed $\lambda \in [0, 1]$. In the case of inside and outside averaging, these assumptions imply the contractiveness of the map g, implying that for every fixed $\lambda \in [0, 1]$, $F_\lambda(x, T(x))$ is a continuous function of x.

Before analyzing the behavior of the general averaging map $F_\lambda(x, T(x))$, we examine several of its special cases. In particular, we show that several iterative schemes which are the algorithmic extensions of the examples **A**–**F** of the previous section satisfy properties B1–B7.

A Outside Averaging:

$$x_{k+1} = \lambda_{k+1} g(x_k) + (1 - \lambda_{k+1}) T(x_k)$$

Assume that

i) K is a bounded, closed, convex feasible region.

ii) g is a contractive map with a contraction constant $a \in (0, 1)$.

iii) T is a nonexpansive map.

iv) The stepsizes λ_{k+1} are chosen as in B4–B6.

Observe that in this case $F_\lambda(x, T(x)) = \lambda g(x) + (1-\lambda)T(x)$; consequently,

- $F_0(x, T(x)) = T(x)$, therefore, $FP(T) = FP(F_0)$ and B1 is satisfied.
- Conditions ii) and iii) imply that;
  $\|F_\lambda(x, T(x)) - F_\lambda(y, T(y))\| \leq \lambda a\|x-y\| + (1-\lambda)\|x-y\| = \beta(\lambda)\|x-y\|$ with $\beta(\lambda) = (1-\lambda(1-a))$ satisfies condition B2. In this case, $\beta(\lambda) = 1-\lambda(1-a) \geq 0$, $\beta(0) = 1$, $\beta'(0) = -(1-a) < 0$ and $\beta''(\lambda) = 0$.
- Condition iii) implies that
  $\|F_\lambda(x, T(x)) - F_{\bar{\lambda}}(x, T(x))\| \leq |\lambda - \bar{\lambda}|\|g(x)\| + |\lambda - \bar{\lambda}|\|T(x)\| = |\lambda - \bar{\lambda}|(\|g(x)\| + \|T(x)\|)$, which satisfies condition B3 with $L(x) = \|g(x)\| + \|T(x)\|$.
- B4, B5, B6 follow by definition.
- B7 follows since K is bounded.

B Inside Averaging:

$$x_{k+1} = T(\lambda_{k+1}g(x_k) + (1-\lambda_{k+1})x_k)$$

Assume that

i) K is a bounded, closed, convex feasible region.

ii) g is a contractive map with a contraction constant $a \in (0,1)$.

iii) T is a nonexpansive map.

iv) The stepsizes λ_{k+1} are chosen as in B4–B6.

Observe that in this case $F_\lambda(x, T(x)) = T(\lambda g(x) + (1-\lambda)x)$. Therefore,

- $F_0(x, T(x)) = T(x)$; consequently, $FP(T) = FP(F_0)$ and B1 is satisfied.
- Conditions ii) and iii) imply that
  $\|F_\lambda(x, T(x)) - F_\lambda(y, T(y))\| \leq \|\lambda g(x) + (1-\lambda)x - \lambda g(y) - (1-\lambda)y\| \leq \lambda a\|x - y\| + (1-\lambda)\|x-y\| = \beta(\lambda)\|x-y\|$ with $\beta(\lambda) = (1-\lambda(1-a))$ satisfies condition B2. In this case, $\beta(\lambda) = 1-\lambda(1-a) \geq 0$, $\beta(0) = 1$, $\beta'(0) = -(1-a) < 0$ and $\beta''(\lambda) = 0$.

- Condition iii) implies that $\|F_\lambda(x,T(x)) - F_{\bar\lambda}(x,T(x))\| \le \|\lambda g(x) + (1-\lambda)x - \bar\lambda g(\dot x) - (1-\bar\lambda)x\| \le |\lambda - \bar\lambda|\|g(x)\| + |\lambda-\bar\lambda|\|x\| = |\lambda-\bar\lambda|(\|g(x)\| + \|x\|)$, and so B3 is satisfied with $L(x) = \|g(x)\| + \|T(x)\|$.
- B4, B5, B6 follow by definition.
- B7 follows since K is bounded.

C Outside Averaging with the Proximal Point Mapping

$$x_{k+1} = (1-\lambda_{k+1})J_{c(I-T)}(x_k) \stackrel{def}{=} (1-\lambda_{k+1})(I + c(I-T))^{-1}(x_k).$$

Assume that

i) K is a bounded, closed, convex feasible region.

ii) T is a nonexpansive map.

iii) The stepsizes λ_{k+1} are chosen as in B4–B6.

- $F_0(x,T(x)) = (I+c(I-T))^{-1}(x)$; therefore, $FP(F_0) = FP(T)$ and B1 is satisfied.
- Since $J_{c(I-T)}$ is a nonexpansive map, $\|F_\lambda(x,T(x)) - F_\lambda(y,T(y))\| \le (1-\lambda)\|x-y\|$. Observe that $\beta(\lambda) = 1-\lambda$ and so $\beta(\lambda) \in [0,1]$, $\beta(0) = 1$, $\beta'(0) = -1 < 0$ and $\beta''(\lambda) \le 0$.
- $\|F_\lambda(x,T(x)) - F_{\bar\lambda}(x,T(x))\| \le |\lambda - \bar\lambda| L(x)$, with $L(x) = (I + c(I-T))^{-1}(x)$, which is bounded since K is a bounded set.
- B4–B6 follow by assumption.
- B7 follow since K is bounded.

D Inside Averaging with the Proximal Point Mapping

$$x_{k+1} = (I + c(I-T))^{-1}((1-\lambda_{k+1})x_k).$$

Assume that

i) K is a bounded, closed, convex feasible region.

ii) T is a nonexpansive map.

iii) The stepsizes λ_{k+1} are chosen as in B4–B6.

- $F_0(x, T(x)) = (I + c(I-T))^{-1}(x)$ and, therefore, $FP(F_0) = FP(T)$.
- Since $J_{c(I-T)}$ is a nonexpansive map, $\|F_\lambda(x,T(x)) - F_\lambda(y,T(y))\| \leq (1-\lambda)\|x-y\|$. Observe that $\beta(\lambda) = 1-\lambda$; therefore, $\beta(\lambda) \in [0,1]$, $\beta(0) = 1$, $\beta'(0) = -1 < 0$ and $\beta''(\lambda) \leq 0$.
- $\|F_\lambda(x,T(x)) - F_{\bar\lambda}(x,T(x))\| \leq |\lambda - \bar\lambda| L(x)$ with $L(x) = \|x\|$ bounded since K is a bounded set.
- B4–B6 follow by definition.
- Follows since K is bounded.

E Convex Combinations

Similar to the previous cases.

We next prove some convergence results for our general averaging framework for situations satisfying conditions B1-B7.

Lemma 1 *:*

Suppose $0 < \beta(0) \leq 1$, $\beta'(0) < 0$, $\beta(\lambda) \geq 0$, $\beta(\lambda) \leq \beta(0) + \beta'(0)\lambda$ *and* $\sum_k \lambda_k = +\infty$. *Then* $\prod_k \beta(\lambda_k) = 0$.

Proof:

Since the log function is concave and $\beta(\lambda) \leq \beta(0) + \beta'(0)\lambda$,

$$\log \beta(\lambda) \leq \log \beta(0) + \frac{\beta'(0)}{\beta(0)}\lambda.$$

Therefore, since $\log \beta(0) \leq \log(1) = 0$,

$$\log \beta(\lambda) \leq \frac{\beta'(0)}{\beta(0)}\lambda$$

which implies that

$$e^{\left(\frac{\beta'(0)}{\beta(0)}\right)\lambda_k} \geq \beta(\lambda_k) \;\; \forall k.$$

Therefore,

$$\prod_k \beta(\lambda_k) \leq e^{\left(\frac{\beta'(0)}{\beta(0)}\right)\sum_k \lambda_k} = 0,$$

since $\sum_k \lambda_k = +\infty$ and $\left(\frac{\beta'(0)}{\beta(0)}\right) < 0$. Consequently, $\prod_k \beta(\lambda_k) = 0$. □

Theorem 3 *: If the set K is bounded, then conditions B1–B7 imply that every limit point of the sequence $\{x_k\}$ induced from the averaging scheme (6) is a fixed point of the map $T : K \to K$.*

Proof:

Since K is bounded, for some constant $L > 0$, $\|x_{k+1} - x_k\| \leq L$ and $L(x) \leq L$ for all $k \geq 0$ and all $x \in K$. Therefore, for any $k > 0$,

$$\|x_{k+1} - x_k\| = \|F_{\lambda_{k+1}}(x_k, T(x_k)) - F_{\lambda_k}(x_{k-1}, T(x_{k-1}))\|$$

$$\leq \|F_{\lambda_{k+1}}(x_k, T(x_k)) - F_{\lambda_{k+1}}(x_{k-1}, T(x_{k-1}))\| + \|F_{\lambda_{k+1}}(x_{k-1}, T(x_{k-1})) - F_{\lambda_k}(x_{k-1}, T(x_{k-1}))\|$$

$$\leq \beta(\lambda_{k+1})\|x_k - x_{k-1}\| + |\lambda_{k+1} - \lambda_k| L(x_k)$$

$$\leq \beta(\lambda_{k+1})\|x_k - x_{k-1}\| + |\lambda_{k+1} - \lambda_k| L.$$

Adding these telescopic inequalities and using the fact that $\beta(\lambda_k) \leq 1$ for all k, shows that

$$\|x_{k+1} - x_k\| \leq \|x_{m+1} - x_m\| \prod_{n=m}^{k+1} \beta(\lambda_{n+1}) + L \sum_{n=m}^{k+1} |\lambda_n - \lambda_{n+1}| \, \forall k \geq m > 0.$$

Consequently,

$$\overline{\lim_k} \|x_{k+1} - x_k\| \leq \|x_{m+1} - x_m\| \prod_{n=m}^{\infty} \beta(\lambda_{n+1}) + L \sum_{n=m}^{\infty} |\lambda_n - \lambda_{n+1}|$$

$$\leq L \prod_{n=m}^{\infty} \beta(\lambda_{n+1}) + L \sum_{n=m}^{\infty} |\lambda_n - \lambda_{n+1}| \longrightarrow_{m,k\to\infty} 0 \;\; (\text{since } \; \beta(\lambda_k) < 1 \; \forall k).$$

Therefore, $\|x_{k+1} - x_k\| \longrightarrow_{k\to\infty} 0$. Also,

$$\|x_k - F_0(x_k, T(x_k))\| \leq \|x_k - x_{k+1}\| + \|x_{k+1} - F_0(x_k, T(x_k))\|$$

$$= \|x_k - x_{k+1}\| + \|F_{\lambda_{k+1}}(x_k, T(x_k)) - F_0(x_k, T(x_k))\|$$

$$\leq \|x_k - x_{k+1}\| + \lambda_{k+1} L(x_k) \longrightarrow_{k\to\infty} 0$$

Therefore, every limit point of the sequence $\{x_k\}$ is a fixed point solution of the map F_0. Property B1 implies that it is also a fixed point solution of T. □

This theorem states that the limit points of the sequence induced by the iterative scheme (6) are, indeed, fixed points of T. Can we, as before, characterize these limiting fixed points? In order to do so, we consider a stronger version of assumptions B1–B7, obtained by replacing conditions B2 and B3 with the conditions

B2′ $F_\lambda = F(G_\lambda)$.

- The map F is nonexpansive (as a function of x) and $FP(F) \supseteq FP(T)$.
- The map G_0 is nonexpansive (as a function of x), $FP(G_0) \supseteq FP(T)$ and

$$\|G_\lambda(x, T(x)) - G_\lambda(y, T(y))\| \leq \beta(\lambda)\|x - y\|, \ \forall x, y \in K$$

with $\beta(\lambda) \geq 0$ and $0 < \beta(0) \leq 1$ and $\beta'(0) < 0$ and $\beta(\lambda) \leq \beta(0) + \beta'(0)\lambda$.

- For all $y \in K$ and all $\lambda \in [0, 1]$,

$$G_\lambda(x^*)^t(y - x^*) \leq \lambda G_1(x^*, T(x^*))^t(y - x^*) + (1 - \lambda)G_0(x^*)^t(y - x^*) \ \forall x^* \in FP(T).$$

B3′ $\|G_\lambda(x, T(x)) - G_{\bar{\lambda}}(x, T(x))\| \leq |\lambda - \bar{\lambda}| L(x)$ for some continuous function $L(x)$ of x.

In particular,

In case **A**, $F = I$ and $G_\lambda = \lambda g(x) + (1 - \lambda)T(x)$.

In case **B**, $F = T$ and $G_\lambda = \lambda g(x) + (1 - \lambda)x$.

In case **C**, $F = I$ and $G_\lambda = (1 - \lambda)(I + c(I - T))^{-1}$.

In case **D**, $F = (I + c(I - T))^{-1}$ and $G_\lambda = (1-\lambda)x$.

We can check that cases A–E satisfy the new properties B2′, B3′.

Remark: Properties B2′ and B3′ imply properties B2 and B3.

For simplicity, in the remainder of this paper we will use the abbreviated notation $G_{\lambda_{k+1}}(x, T(x)) = G_{\lambda_{k+1}}(x)$.

Before stating the main theorem of this section, we introduce the following lemma.

Lemma 2 *: Suppose that for given constants $\mu_j \in [0,1]$, $j \geq k_0 \geq 0$, the relation*

$$||y_{k+1} - x^*||^2 \leq \frac{3}{-\beta'(0)} \mu_{k+1}\epsilon + (1 - \mu_{k+1})||y_k - x^*||^2 \tag{7}$$

is valid for all $k \geq k_0$. Let $k \geq m \geq k_0$ and $1 - \alpha = \prod_{n=m}^{k}(1 - \mu_{n+1})$. Then $0 \leq \alpha \leq 1$ and

$$||y_{k+1} - x^*||^2 \leq \frac{3}{-\beta'(0)} \epsilon\alpha + (1 - \alpha)||y_m - x^*||^2. \tag{8}$$

Proof: We will establish this result using backward induction. Relation (7) implies that the inequality (8) is valid for $m = k$. Suppose that the inequality is also valid for some $m \leq k$. Applying (8) to the term $||y_m - x^*||^2$, we obtain

$$||y_{k+1} - x^*||^2 \leq \frac{3}{-\beta'(0)} \epsilon\alpha + (1 - \alpha)||y_m - x^*||^2$$

$$\leq \frac{3}{-\beta'(0)} \epsilon\alpha + (1 - \alpha)\left(\frac{3}{-\beta'(0)} \epsilon\mu_m + (1 - \mu_m)||y_{m-1} - x^*||^2\right)$$
$$= \frac{3}{-\beta'(0)} \epsilon(\alpha + (1 - \alpha)\mu_m) + (1 - \alpha)(1 - \mu_m)||y_{m-1} - x^*||^2.$$

Let $\alpha' = 1 - (1 - \alpha)(1 - \mu_m) = \alpha + (1 - \alpha)\mu_m$. Since α' is a convex combination of μ_m and 1, $0 \leq \alpha' \leq 1$. Moreover, since we can rewrite the previous expression as

$$||y_{k+1} - x^*||^2 \leq \frac{3}{-\beta'(0)} \epsilon\alpha' + (1 - \alpha')||y_{m-1} - x^*||^2,$$

we have completed the inductive argument and, therefore, established relation (8). Q.E.D.

Theorem 4 *: Conditions B1, B2′, B3′–B8 imply that the entire sequence $\{x_k\}$ induced by the iterative scheme (6) converges to a fixed point solution x^* satisfying the property $x^* = Pr_{FP(T)} G_1(x^*)$.*

Proof:

Condition B2′ implies that $\beta(1) \leq \beta(0) + 1 \cdot \beta'(0) < 1$, indicating that the map $G_1(x)$ is a contraction. Therefore, according to Proposition 2, a unique point $x^* \in FP(T)$ satisfies the property $x^* = Pr_{FP(T)} G_1(x^*)$ or, equivalently, $(x - x^*)^t (G_1(x^*) - x^*) \leq 0$ for all $x \in FP(T)$.

Let $\bar{x}$ be a limit point of the sequence $\{x_k\}$ induced by the iterates (6). Then since $FP(T) \subseteq FP(G_0)$, $\bar{x}$ is also a fixed point of G_0, implying that $(G_0(\bar{x}) - x^*)^t (G_1(x^*) - x^*) \leq 0$. Therefore, $\limsup_k (G_0(x_k) - x^*)^t ((G_1(x^*) - x^*) \leq 0$, implying that for any $\epsilon > 0$, $(G_0(x_k) - x^*)^t ((G_1(x^*) - x^*) \leq \epsilon$ for k sufficiently large. Define $L = \sup_{x \in K} L(x)$. Since $\lambda_k \to 0$ as $k \to \infty$, for k sufficiently large, $3L^2 \lambda_k \leq \epsilon$ for arbitrarily small $\epsilon > 0$. Let k_0 be large enough so that for all $k \geq k_0$ both of these conditions are valid. Then if we let $y_{k+1} = G_{\lambda_{k+1}}(x_k)$,

$$\|y_{k+1} - x^*\|^2 = \|G_{\lambda_{k+1}}(x_k) - G_0(x^*)\|^2$$

$$= \|G_{\lambda_{k+1}}(x_k) - G_{\lambda_{k+1}}(x^*)\|^2 + \|G_{\lambda_{k+1}}(x^*) - G_0(x^*)\|^2 + 2(G_{\lambda_{k+1}}(x_k) - G_{\lambda_{k+1}}(x^*))^t (G_{\lambda_{k+1}}(x^*) - G_0(x^*))$$

$$= \|G_{\lambda_{k+1}}(x_k) - G_{\lambda_{k+1}}(x^*)\|^2 + \|G_{\lambda_{k+1}}(x^*) - G_0(x^*)\|^2 + 2(G_{\lambda_{k+1}}(x_k) - G_0(x_k))^t (G_{\lambda_{k+1}}(x^*) - G_0(x^*))$$

$$+ 2(G_0(x_k) - G_0(x^*))^t (G_{\lambda_{k+1}}(x^*) - G_0(x^*)) - 2\|G_{\lambda_{k+1}}(x^*) - G_0(x^*)\|^2$$

$$\leq (\beta(\lambda_{k+1}))^2 \|x_k - x^*\|^2 + 3L^2 \lambda_{k+1}^2 + 2(G_0(x_k) - G_0(x^*))^t (G_{\lambda_{k+1}}(x^*) - G_0(x^*)).$$

To obtain the previous inequality, we have applied assumptions B2′ and B3′ twice, once after using the Cauchy-Schwartz inequality. Since $x_k = F(y_k)$ and x^* is a fixed point of F, which is a nonexpansive map,

$$\|y_{k+1} - x^*\|^2 \leq (\beta(\lambda_{k+1}))^2 \|y_k - x^*\|^2 + 3L^2 \lambda_{k+1}^2 + 2(G_0(x_k) - G_0(x^*))^t (G_{\lambda_{k+1}}(x^*) - G_0(x^*)).$$

Applying condition B2$'$ to the last term of this inequality and using the fact that $G_0(x^*) = x^*$ gives

$$\|y_{k+1} - x^*\|^2 \leq (\beta(\lambda_{k+1}))^2\|y_k - x^*\|^2 + 3L^2\lambda_{k+1}^2 + 2\lambda_{k+1}(G_0(x_k) - x^*)^t(G_1(x^*) - G_0(x^*))$$

$$\leq (\beta(\lambda_{k+1}))^2\|y_k - x^*\|^2 + 3L^2\lambda_{k+1}^2 + 2\lambda_{k+1}\epsilon \leq (\beta(\lambda_{k+1}))^2\|y_k - x^*\|^2 + 3\epsilon\lambda_{k+1}.$$

The properties of $\beta(\lambda)$ in B2$'$ imply that

$$(\beta(\lambda_{k+1}))^2 \leq \beta(\lambda_{k+1}) \leq \beta(0) + \beta'(0)\lambda_{k+1} \leq 1 + \beta'(0)\lambda_{k+1},$$

with $\beta'(0) < 0$. Therefore,

$$\|y_{k+1} - x^*\|^2 \leq (1 + \beta'(0)\lambda_{k+1})\|y_k - x^*\|^2 + 3\epsilon\lambda_{k+1}.$$

Letting $-\beta'(0)\lambda_k = \mu_k$ and noticing that $\mu_k \in [0,1]$ for all k sufficiently large, we can rewrite this relation as

$$\|y_{k+1} - x^*\|^2 \leq \frac{3}{-\beta'(0)}\mu_{k+1}\epsilon + (1 - \mu_{k+1})\|y_k - x^*\|^2.$$

Lemma 2 implies relation (8) and, therefore, that

$$\|y_{k+1} - x^*\|^2 \leq \frac{3}{-\beta'(0)}\epsilon + \prod_{n=\bar{k}}^{k}(1 - \mu_{n+1})\|y_{\bar{k}} - x^*\|^2. \tag{9}$$

In this expression, $\bar{k}$ is any value chosen so that all the previous facts contingent on the k being "large enough" are valid. Lemma 1 and equation (9) imply that

$$\overline{\lim_k}\|y_{k+1} - x^*\|^2 \leq \frac{3}{-\beta'(0)}\epsilon.$$

Therefore, since ϵ is arbitrary,

$$\|y_{k+1} - x^*\| \longrightarrow_{k\to\infty} 0.$$

Observe that $x_{k+1} = F(y_{k+1})$ and $FP(F) \supseteq FP(T)$ imply that the sequence $\{x_k\}$ converges to the fixed point solution x^* of T satisfying the property $x^* = Pr_{FP(T)}(G_1(x^*))$. $\square$

3.2 Path-following algorithms

In this subsection we will present another variant of the previous averaging framework by considering multiple applications of the averaging map $F_{\lambda_{k+1}}$. That is, we approximate the trajectories of the parameterized fixed point subproblems more accurately by applying the approximate averaging maps several times. In return, we need not place any restrictions on the sequence $\{\lambda_k\}$.

To provide some motivation, we consider the sequence x_λ induced by the parameterized fixed point subproblems (2). When conditions A1–A8b are valid, the sequence $\{x_\lambda\}$ converges to a fixed point solution x^* as $\lambda \to 0$ (see Theorem 2, part II). If, in addition, $F_\lambda(\cdot)$ is a contractive map for every $\lambda > 0$, then we can evaluate the fixed point x_λ as

$$x_\lambda = \lim_{k \to \infty} F_\lambda^k(X_0, T(X_0)) \tag{10}$$

for any starting point $X_0 \in K$.

Since x_λ converges to a fixed point solution of (1) as $\lambda \to 0$, we might consider an iterative scheme that "follows closely" the path of fixed points $\{x_\lambda\}$. That is, given a sequence $\lambda_k \to 0$, $\lambda_k \in (0,1)$ and a sequence of nonnegative integers $\{m_k\}$, for an arbitrary $X_0 \in K$, we can approximate the original fixed point problem through

$$X_{k+1} = F_{\lambda_{k+1}}^{m_{k+1}}(X_k, T(X_k)) \tag{11}$$

by ensuring that the iterates X_k stay "close" to the fixed point solutions of the corresponding subproblems $x_{\lambda_k} = F_{\lambda_k}(x_{\lambda_k}, T(x_{\lambda_k}))$.

The following result formalizes this iterative scheme.

Theorem 5 *: Let $\{\lambda_k\}$ and $\{m_k\}$ be scalars satisfying conditions A1-A8b and assume F_λ satisfies condition B2 with $\beta(\lambda_k)^{m_k} \to 0$ as $k \to \infty$. Then the sequence $\{X_k\}$ induced by relation (11) converges to the fixed point solution $x^* = \lim_{\lambda \to 0} x_\lambda$.*

Proof: Since the set K is bounded, for some constant L, $\|x-y\| \leq L$ for all $x, y \in K$. Furthermore, since x_{λ_k} is a fixed point solution of the map $F_{\lambda_k}(x_{\lambda_k}, T(x_{\lambda_k})$ and, therefore, of the map $F_{\lambda_k}^{m_k}(x_{\lambda_k}, T(x_{\lambda_k}),$

$$\|X_{k+1} - x_{\lambda_{k+1}}\| = \|F_{\lambda_{k+1}}^{m_{k+1}}(X_k, T(X_k)) - F_{\lambda_{k+1}}^{m_{k+1}}(x_{\lambda_{k+1}}, T(x_{\lambda_{k+1}}))\|$$

$$\leq \beta(\lambda_{k+1})^{m_{k+1}} \|X_k - x_{\lambda_{k+1}}\| \leq \beta(\lambda_{k+1})^{m_{k+1}} L \longrightarrow_{k \to \infty} 0.$$

Since $x_{\lambda_k} \longrightarrow_{k\to\infty} x^*$ by Theorem 2, this inequality implies that the sequence X_k converges to the fixed point solution $x^* = \lim_{\lambda \to 0} x_\lambda$. $\square$

4 Approximate Averaging Trajectories; an Alternate Approach

To this point, we have introduced a general averaging framework using contractive averaging maps F_λ for each fixed value of λ. We next study the general averaging framework when the map F_λ is nonexpansive. Averaging with the identity map (line search procedures) is a type of averaging that has received much attention in the literature (see for example [5], [10] [11]). For this type of averaging, the averaging map F_λ is nonexpansive rather than contractive. *Can our averaging framework incorporate this type of averaging?* In this section, we answer this question by illustrating how our averaging framework works. In the context of this relaxed assumption, that is, when the map F_λ is nonexpansive.

We consider a continuous function $F_\lambda(x, T(x))$ of x and λ, mapping $\Lambda \times K \times K$ into K. We impose several conditions on F_λ:

C1 $FP(F_0) = FP(T)$.

C2 $F_\lambda = F(G_\lambda)$

- F is a nonexpansive map and $FP(T) \subseteq FP(F)$.
- G_0 is a nonexpansive map and $FP(T) \subseteq FP(G_0)$.
- G_1 is a nonexpansive map and $FP(T) \subseteq FP(G_1)$.

C3 $G_\lambda(x^*)^t(y-x^*) \leq \lambda G_1(x^*)^t(y-x^*) + (1-\lambda)G_0(x^*)^t(y-x^*)$, $\forall y \in FP(T)$.

C4 Either F is a contractive map (with a contraction constant $a \in (0,1)$) and G_λ is a nonexpansive map: $\|G_\lambda(x) - G_\lambda(x^*)\|^2 \leq \|x - x^*\|^2$,

or F is a nonexpansive map (with a "contractive" constant $a = 1$), $FP(G_0) = FP(T)$ and some $\gamma(\lambda) \in R^+$ satisfies the condition,

$\|G_\lambda(x) - G_\lambda(x^*)\|^2 \leq \|x - x^*\|^2 - \gamma(\lambda)\|x - G_0(x)\|^2$.

C5 $\|G_\lambda(x) - G_{\bar{\lambda}}(x)\| \leq L(x)|\lambda - \bar{\lambda}|$,

and $L(x) \leq L \ \forall x \in K$.

The stepsizes λ_k satisfy the conditions

C6 The $\lim_k \lambda_k = 0$ and $\sum_k \gamma(\lambda_k) = +\infty$.

As in Section 2, we will consider the general iteration

$$x_{k+1} = F_{\lambda_{k+1}}(x_k). \tag{12}$$

Question:

When is condition C4 true?

The following lemma provides some sufficient conditions.

Lemma 3 : *The following conditions imply condition C4.*

a) G_0 *is firmly nonexpansive, that is,*

$\|G_0(x) - G_0(x^*)\|^2 \leq \|x - x^*\|^2 - \|x - G_0(x)\|^2$.

b) $\|G_\lambda(x) - G_\lambda(x^*)\|^2 \leq \gamma(\lambda)\|G_0(x) - G_0(x^*)\|^2$, *with* $\gamma(\lambda) \in [0,1]$.

Proof: Conditions b) and a) imply that

$$\|G_\lambda(x) - G_\lambda(x^*)\|^2 \leq \gamma(\lambda)\|G_0(x) - G_0(x^*)\|^2 \leq \gamma(\lambda)\|x - x^*\|^2 - \gamma(\lambda)\|x - G_0(x)\|^2 \leq$$

$$\|x - x^*\|^2 - \gamma(\lambda)\|x - G_0(x)\|^2.$$

□

Theorem 6 *: Conditions C1–C6 imply that the entire sequence $\{x_k\}$ induced by the averaging scheme (12) converges to a fixed point solution.*

Proof: Let x^* be a fixed point of the map T, and set $y_{k+1} = G_{\lambda_{k+1}}(x_k)$ and $x_{k+1} = F(y_{k+1})$. The nonexpansiveness of the map F, the definition of the sequence $\{y_k\}$, the fact that x^* is a fixed point of F, G_0, and G_1, and condition C3 imply that

$$\|x_{k+1} - x^*\|^2 \leq a\|y_{k+1} - x^*\|^2 = a\|G_{\lambda_{k+1}}(x_k) - G_0(x^*)\|^2$$

$$= a\|G_{\lambda_{k+1}}(x_k) - G_{\lambda_{k+1}}(x^*)\|^2 - a\|G_{\lambda_{k+1}}(x^*) - G_0(x^*)\|^2 + 2a(G_{\lambda_{k+1}}(x_k) - G_0(x^*))^t(G_{\lambda_{k+1}}(x^*) - G_0(x^*))$$

$$\leq a\|G_{\lambda_{k+1}}(x_k) - G_{\lambda_{k+1}}(x^*)\|^2 - a\|G_{\lambda_{k+1}}(x^*) - G_0(x^*)\|^2 + 2\lambda_{k+1}a(G_{\lambda_{k+1}}(x_k) - G_0(x^*))^t(G_1(x^*) - G_0(x^*))$$

$$\leq a\|G_{\lambda_{k+1}}(x_k) - G_{\lambda_{k+1}}(x^*)\|^2.$$

Condition C4 implies that either

(1) F is contractive (with $a \in (0,1)$) and G_λ is nonexpansive. Then for some $a \in (0,1)$,

$$\|x_{k+1} - x^*\|^2 \leq a\|x_k - x^*\|^2.$$

This inequality implies the convergence of the sequence $\{x_k\}$ to a fixed point solution x^* of T.

(2) F is nonexpansive (with $a = 1$) and condition C4 implies that if we consider a fixed point x^* of the map T, then

$$\|x_{k+1} - x^*\|^2 \leq \|x_k - x^*\|^2 - \gamma(\lambda_{k+1})\|x_k - G_0(x_k)\|^2. \tag{13}$$

Relation (13) implies that the sequence $\{\|x_k - x^*\|^2\}$ converges for all fixed points $x^* \in FP(T)$. We now consider two subcases.

(i) For some limit point $\bar{x}$ of the sequence $\{x_k\}$, $\|\bar{x} - G_0(\bar{x})\|^2 = 0$ implying that $\bar{x} \in FP(G_0) = FP(T)$. But since $\{\|x_k - x^*\|^2\}$ is convergent for all fixed points x^* of T, the entire sequence $\{x_k\}$ converges to a fixed point solution $\bar{x}$ of T.

(ii) For some sufficiently large K and some constant $B > 0$, $\|x_k - G_0(x_k)\|^2 \geq B$ for all $k \geq K$. Consequently,

$$\infty > \lim \|x_k - x^*\|^2 - \|x_K - x^*\|^2 \leq - \sum_{k=K}^{+\infty} \gamma(\lambda_{k+1}) B = -\infty$$

which is a contradiction. Therefore, $\|\bar{x} - G_0(\bar{x})\|^2 = 0$. Then, as before, the entire sequence $\{x_k\}$ converges to a fixed point solution of T. $\square$

We next examine the application of these conditions to the special cases **A**–**F** that we introduced in Section 1.

Special Cases:

A Outside Averaging

$$F_\lambda^{out}(x) = \lambda g(x) + (1 - \lambda) T(x)$$

- The map T is *firmly nonexpansive.*
- The map g is *nonexpansive* and $FP(g) \supseteq FP(T)$.
- $\lambda_k > 0$ with $\sum_k \lambda_k (1 - \lambda_k) = +\infty$.

Observe that in this method, g could be the identity map. In this case, the method reduces to averaging with the identity map (see for example [5], [10], [11]).

Note that $G_\lambda = \lambda g(x) + (1 - \lambda) T(x)$ and $F = I$. Condition C1 is valid since $F_0(x) = T(x)$.

The fact that $g(x^*) = T(x^*) = x^*$ implies that $G_\lambda(x^*) = x^*$. Therefore, condition C4 applies with $\gamma(\lambda) = \lambda(1 - \lambda)$ since

$$\|G_\lambda(x) - G_\lambda(x^*)\|^2 \leq \|\lambda g(x) + (1 - \lambda) T(x) - x^*\|^2 \leq$$

$$\lambda \|g(x) - x^*\|^2 + (1 - \lambda) \|x - x^*\|^2 - \lambda(1 - \lambda) \|x - T(x)\|^2 \leq$$

$$\|x - x^*\|^2 - \lambda(1 - \lambda) \|x - T(x)\|^2.$$

We can also check that conditions C2, C3, C5 and C6 hold.

In the special case of $g(x) = x$, this result becomes Dunn's averaging result [5].

B Inside Averaging

$$F_\lambda^{in}(x) = T(\lambda g(x) + (1-\lambda)x),$$

- The map T is *contractive.*
- The map g is *nonexpansive* and $FP(g) \supseteq FP(T)$.
- $\lambda_k > 0$ with $\sum_k \lambda_k(1-\lambda_k) = +\infty$.

Observe that $G_\lambda = \lambda g(x) + (1-\lambda)x$ and $F = T$. Condition C1 is valid since $F_0(x) = T(x)$. Condition C4 applies with $\gamma(\lambda) = \lambda(1-\lambda)$ since

$$\|G_\lambda(x) - G_\lambda(x^*)\| = \|\lambda g(x) + (1-\lambda)x - x^*\|^2 =$$

$$\|x - x^*\|^2 + \lambda^2\|x - g(x)\|^2 - 2\lambda(x - g(x))^t(x - x^*) \leq \|x - x^*\|^2 - \lambda(1-\lambda)\|x - g(x)\|^2.$$

The last inequality follows from the nonexpansiveness of the map g (see Proposition 1) as well as the fact that $x^* - g(x^*) = 0$ which follows from the assumption that $FP(g) \supseteq FP(T)$.

Conditions C2, C3, C5 and C6 also hold.

C Outside Averaging with the Proximal Point Mapping

$$F_\lambda(x) = \lambda g(x) + (1-\lambda)(I + c(I-T))^{-1}(x).$$

- The map T is *nonexpansive.*
- The map g is *nonexpansive* and $FP(g) \supseteq FP(T)$.
- $\lambda_k > 0$ with $\sum_k (1-\lambda_k)^2 = +\infty$.

In this case, $G_\lambda = \lambda g(x) + (1-\lambda)(I + c(I-T))^{-1}$ and $F = I$. Condition C1 is valid since $F_0(x) = (I + c(I-T))^{-1}(x)$.

Condition C4 applies with $\gamma(\lambda) = (1-\lambda)^2$. That is,

a) $G_0(x) = (I + c(I-T))^{-1}(x)$ is firmly nonexpansive (when T is nonexpansive).

b) $\|G_\lambda(x) - G_\lambda(x^*)\|^2 \leq [\|(1-\lambda)[(I + c(I-T))^{-1}(x) - x^*] + \lambda[g(x) - g(x^*)]\|]^2$ (after expanding and using the nonexpansiveness of map g as well as the fact that $FP(g) \supseteq FP(T)$)

$$\leq (1-\lambda)^2\|G_0(x) - G_0(x^*)\|^2 + \lambda^2\|x - x^*\|^2 + 2\lambda(1-\lambda)\|x - x^*\|^2.$$

Substituting a) in b) implies that

$$\|G_\lambda(x) - G_\lambda(x^*)\|^2 \leq \|x - x^*\|^2 - (1-\lambda)^2\|x - G_0(x)\|^2.$$

Conditions C2, C3, C5 and C6 also hold.

D Inside Averaging with the Proximal Point Mapping

$$F_\lambda(x) = (I + c(I-T))^{-1}(\lambda g(x) + (1-\lambda)x).$$

- The map T is *nonexpansive.*
- The map g is *nonexpansive* and $FP(g) \supseteq FP(T)$.
- $\lambda_k > 0$ with $\sum_k \lambda_k(1-\lambda_k) = +\infty$.

In this case, $G_\lambda = \lambda g + (1-\lambda)I$ and $F = (I + c(I-T))^{-1}$. Condition C1 is valid since $F_0(x) = (I + c(I-T))^{-1}(x)$.

Averaging as we did for method **C** shows that condition C4 applies with $\gamma(\lambda) = \lambda(1-\lambda)$. Conditions C2, C3, C5 and C6 also hold.

E Convex Combinations of Averaging Maps

Similar to the other cases.

4.1 Generalized Averaging with a Class of Contractive Maps

In this subsection we will establish an averaging scheme for solving fixed point problems with maps T of *bounded expansion.* We consider the special case of outside averaging, assuming that g is a contractive map

$$x_{k+1} = \lambda_{k+1}g(x_k) + (1-\lambda_{k+1})T(x_k)$$

and impose the following assumptions.

D1 The map T has fixed point solutions, and it is a map of *bounded expansion* around its solutions, that is,

$\|T(x_k) - x^*\|^2 \leq B\|x_k - x^*\|^2$, for some constant $B > 1$.

D2 The map g is contractive around its unique fixed point solution (with contractive constant $a \in (0,1)$).

D3 $FP(T) \supseteq FP(g)$.

D4 The step sizes $\lambda_{k+1} \in [c, 1]$, with $c > \frac{B-1}{B-a}$.

Theorem 7 *: Conditions D1–D4 imply that the outside averaging scheme induces a sequence that converges to the unique fixed point solution of g, which is also a fixed point solution of T.*

Proof: Conditions D1–D4 imply the following relationships:

$$\|x_{k+1} - x^*\|^2 = \|\lambda_{k+1}(g(x_k) - x_k) + (1 - \lambda_{k+1})(T(x_k) - x_k) + x_k - x^*\|^2$$

$$= \|\lambda_{k+1}(g(x_k) - x_k) + (1 - \lambda_{k+1})(T(x_k) - x_k)\|^2 + \|x_k - x^*\|^2$$

$$+2\lambda_{k+1}(g(x_k) - x_k)(x_k - x^*) + 2(1 - \lambda_{k+1})(T(x_k) - x_k)(x_k - x^*)$$

(expanding the first term implies that)

$$= \|x_k - x^*\|^2 + \lambda_{k+1}^2\|x_k - g(x_k)\|^2 + (1 - \lambda_{k+1})^2\|x_k - T(x_k)\|^2$$

$$+2\lambda_{k+1}(1-\lambda_{k+1})(x_k-T(x_k))^t(x_k-g(x_k))-2\lambda_{k+1}(x_k-g(x_k))(x_k-x^*)-2(1-\lambda_{k+1})(x_k-T(x_k))(x_k-x^*)$$

(Proposition 1 implies in turn that)

$$\leq \|x_k - x^*\|^2 + \lambda_{k+1}^2\|x_k - g(x_k)\|^2 + (1 - \lambda_{k+1})^2\|x_k - T(x_k)\|^2 + \lambda_{k+1}(1 - \lambda_{k+1})\|g(x_k) - x_k\|^2$$

$$+\lambda_{k+1}(1 - \lambda_{k+1})\|T(x_k) - x_k\|^2 - \lambda_{k+1}(1 - a)\|x_k - x^*\|^2 - \lambda_{k+1}\|x_k - g(x_k)\|^2$$

$$+(B-1)(1-\lambda_{k+1})\|x_k - x^*\|^2 - (1-\lambda_{k+1})\|x_k - T(x_k)\|^2$$
$$= \|x_k - x^*\|^2 - \|x_k - x^*\|^2[\lambda_{k+1}(B-a) - (B-1)].$$

Consequently,

$$\|x_{k+1} - x^*\|^2 \leq (1-C)\|x_k - x^*\|^2,$$

with $1 > C = c(B-a) - (B-1) > 0$. Therefore, the sequence $\{x_k\}$ contracts to a fixed point solution of T, namely, the one that is also the unique fixed point solution of g. □

Example:

Consider the expansive map $T(x_1, x_2) = 3(x_2, -x_1)$ which has a fixed point solution $x^* = (0,0)$. Observe that the constant $B = 3$ (condition D1), since

$$\|T(x) - T(x^*)\| = 3\|x\| = 3\|x - x^*\|.$$

Suppose we select $g(x) = \frac{1}{6}T(x) = \frac{1}{2}(x_2, -x_1)$ whose fixed point solution is also the point $x^* = (0,0)$ (condition D3). The map g is a contraction with constant $a = \frac{1}{2}$ (condition D2) since

$$\|g(x) - g(x^*)\| = \frac{1}{2}\|x\| = \frac{1}{2}\|x - x^*\|.$$

If we choose stepsizes $1 \geq \lambda_{k+1} \geq c > \frac{3-1}{3-\frac{1}{2}} = \frac{4}{5}$, then we can easily see that the sequence

$$x_{k+1} = \lambda_{k+1}\frac{1}{6}T(x_k) + (1-\lambda_{k+1})T(x_k) = T(x_k) - \frac{5}{6}\lambda_{k+1}T(x_k).$$

Of course, the difficulty in implementing this approach is the construction of a contractive map g whose unique fixed point solution is also a fixed point solution of T. Since, in general, this construction seems to require knowledge of the fixed points of the map T, this scheme appears at this point to be of only theoretical interest. The other methods that we have introduced don't suffer from this limitation. A natural direction for fututre research would be to test the methods computationally and to make necessary refinements to improve their convergence behavior.

Acknowledgment: This research has been supported by NSF grant DMI-9634736.

References

[1] H.H. Bauschke. The approximation of fixed points of compositions of nonexpansive mappings in Hilbert space. *Preprint.* Center for Experimental and Constructive Mathematics, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6, 1994.

[2] F.E. Browder. Fixed-point theorems for noncompact mappings in Hilbert space. *Proc. Nat. Acad. Sci. U.S.A.*, 53: 1272-1276, 1965.

[3] F.E. Browder. Convergence of approximants to fixed points of nonexpansive nonlinear mappings in Banach spaces. *Archive for Rational Mechanics and Analysis*, 24:82–90, 1967.

[4] F. E. Browder and W.V. Petryshn. Construction of fixed points of nonlinear mappings in a Hilbert space. *Journal of Analysis and Applications*, 20:197-228.

[5] J.C. Dunn. On the recursive averaging processes and Hilbert space extensions of the contraction mapping principle. *Journal of the Franklin Institute*, 295:117-133, 1973.

[6] J. Eckstein and D.P. Bertsekas. On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Mathematical Programming*, 55:293–318, 1992.

[7] A. Genel and J. Lindenstrauss. An example concerning fixed points. *Israel Journal of Mathematics*, 22:81–86, 1975.

[8] B. Halpern. Fixed points of nonexpanding maps. *Bulletin of the American Mathematical Society*, 73:957–961, 1967.

[9] S. Kaniel. Construction of a fixed point for contractions in Banach space. *Israel Journal of Mathematics*, 9:535–540, 1971.

[10] T. L. Magnanti and G. Perakis. Averaging schemes for variational inequalities and systems of equations. *Mathematics of Operations Research*, 22(3):568-587, 1997.

[11] T. L. Magnanti and G. Perakis. Solving variational inequalities and fixed point problems by averaging and optimizing potentials. *preprint*, 1997.

[12] S. Reich. Weak convergence theorems for nonexpansive mappings in Banach spaces. *Journal of Mathematical Analysis and Applications*, 67:274–276, 1979.

[13] R.T. Rockafellar. Monotone operators and the proximal point algorithm. *SIAM Journal on Control and Optimization*, 14(5):877–898, 1976.

[14] R. Wittmann. Approximation of fixed points of nonexpansive mappings. *Archiv der Mathematik*, 58:486–491, 1992.