Math. Commun., Vol. **14**, No. 2, pp. 399-412 (2009)

Are adaptive Mann iterations really adaptive?

KAMIL S. KAZIMIERSKI[1,*]

[1] *Department of Mathematics and Computer Science, University of Bremen, 28 334 Bremen, Germany*

Received December 11, 2008; accepted October 17, 2009

Abstract. We show that the Adaptive Mann Iterations deserve to be named in such way. Namely, we will show that they adapt to the properties of the operator, even if the information on these properties are not known to the Mann Iteration a-priori. We will also show on an example that the Adaptive Mann Iterations perform better numerically than the usual Mann Iterations.
AMS subject classifications: Primary 47H06, 47H10; Secondary 54H25

Key words: adaptive Mann iterations, strongly accretive mappings, strongly pseudocontractive mappings, p-smooth Banach spaces, Banach spaces smooth of power type

1. Introduction

Let X be a real Banach space. For $p > 1$ the set-valued mapping $J_p : X \to 2^{X^*}$ given by

$$J_p(x) = \{x^* \in X^* \ : \ \langle x^*, x\rangle = \|x^*\|\|x\|, \|x^*\| = \|x\|^{p-1}\},$$

where X^* denotes the dual space of X and $\langle \cdot, \cdot\rangle$ denotes the duality pairing, is called the duality mapping of X. By j_p we denote a single-valued selection of J_p. Notice that as a consequence of the Hahn-Banach Theorem, $J_p(x)$ is nonempty for every x in X.

A Banach space X is said to be p-smooth, if it admits a weak polarization law, i.e. if there exists a positive constant G_p such that for all x, y in X and all $j_p(x) \in J_p(x)$

$$\|x - y\|^p \leq \|x\|^p - p\langle j_p(x), y\rangle + G_p\|y\|^p.$$

Notice that for $1 < p < \infty$ the sequence spaces ℓ_p, Lebesgue spaces L_p and Sobolev spaces W_p^m are $\min\{2, p\}$-smooth [9, 15]. Due to the polarization identity every Hilbert space is 2-smooth.

A map $T : X \to X$ is called strongly pseudocontractive, if for some $p > 1$ there exists a constant $k > 0$, such that for each x and y in X there is a $j_p(x-y) \in J_p(x-y)$ satisfying

$$\langle Tx - Ty, j_p(x - y)\rangle \leq (1 - k)\|x - y\|^p.$$

A map T is called strongly accretive if $(I - T)$ is strongly pseudocontractive.

*Corresponding author. *Email address:* `kamilk@math.uni-bremen.de` (K. S. Kazimierski)

Strongly accretive operators are of great importance in physics, since it is well known that many significant problems in physics can be modeled as time-dependent nonlinear equations of the form

$$\frac{du}{dt} + S(t)u = 0$$

, where S is a strongly accretive operator. The main focus of interest lies on the equilibrium points of such a system. Therefore many solvers of the equation

$$Su = 0$$

have been introduced in recent years. Notice that u solves the above equation, if (and only if) it is a fixed point of the mapping $T := I - S$. It is well-known (cf. e.g. [5, 3, 4, 7, 13, 14, 12]) that Mann iterations [11] are well suited methods for finding such a fixed point of the strongly pseudocontractive mapping T. We recall that the sequence (x_n) is a Mann iteration, if for some nonnegative sequence (α_n) the recursion

$$x_{n+1} = (1 - \alpha_n)x_n + \alpha_n T x_n \quad \text{with } x_0 \in C \tag{1}$$

is satisfied.

In [10] the author introduced a problem-adapted selection strategy for the update coefficients (α_n), if T is a strongly pseudocontractive mapping defined on some p-smooth space X.

In this paper we will introduce a new and improved version of the strategy mentioned before. We will also show that the strategy of this paper is adaptive in the sense that it adapts any continuity of the operator. This result is the main difference and improvement to the Mann iteration proposed by Chidume in [5]. Convergence rates results for Mann iterations are very rare. To the author's best knowledge this is the first time that convergence rates for accretive operators subject to their continuity were proven and hence it is the first time that the adaptivity of the Mann iterations was shown.

2. Construction of the adaptive Mann iterations

In this section we construct adaptive Mann iterations for strongly pseudocontractive and strongly accretive operators mapping X to X. The construction carried out in this section is an improved version of the construction presented in [10].

Let X be a p-smooth Banach space, $T : X \to X$ strongly pseudocontractive and suppose T has a fixed point. If T has a fixed point, then by the strong pseudocontractivity this fixed point is unique (cf. e.g. [10]). Since X is p-smooth and T is strongly pseudocontractive, we have

$$\begin{aligned}\|x_{n+1}-x^*\|^p &\leq \|x_n-x^*\|^p-\alpha_n\cdot p\langle x_n-Tx_n,j_p(x_n-x^*)\rangle\\&\quad+\alpha_n^p\cdot G_p\|x_n-Tx_n\|^p\\&=\|x_n-x^*\|^p-\alpha_n\cdot p\langle x_n-x^*-(Tx_n-Tx^*),j_p(x_n-x^*)\rangle\\&\quad+\alpha_n^p\cdot G_p\|x_n-Tx_n\|^p\\&\leq\|x_n-x^*\|^p-\alpha_n\cdot p\|x_n-x^*\|^p+\alpha_n\cdot p(1-k)\|x_n-x^*\|^p\\&\quad+\alpha_n^p\cdot G_p\|x_n-Tx_n\|^p\ .\end{aligned}$$

We arrive at the central inequality of our construction

$$\|x_{n+1}-x^*\|^p\leq\|x_n-x^*\|^p-\alpha_n\cdot pk\|x_n-x^*\|^p+\alpha_n^p\cdot G_p\|x_n-Tx_n\|^p. \tag{2}$$

We notice that the right-hand side in (2) is smaller than the left-hand side as long as

$$0<\alpha_n<\alpha_n^+:=\left(\tfrac{pk}{G_p}\cdot\tfrac{\|x_n-x^*\|^p}{\|x_n-Tx_n\|^p}\right)^{\frac{1}{p-1}}\ . \tag{3}$$

By differentiation we can see that the right-hand side in (2) is minimal for

$$\alpha_n^*=\left(\tfrac{k}{G_p}\cdot\tfrac{\|x_n-x^*\|^p}{\|x_n-Tx_n\|^p}\right)^{\frac{1}{p-1}}\ . \tag{4}$$

By τ_n we denote the number defined by

$$\tau_p:=\left(\tfrac{1}{p}\right)^{\frac{1}{p-1}}\ . \tag{5}$$

Then we see that

$$\alpha_n^*=\tau_n\alpha_n^+.$$

We do not know the exact value of $\|x_n-x^*\|^p$. Assume that we know an upper bound for R_n on $\|x_n-x^*\|^p$, i.e.

$$\|x_n-x^*\|^p\leq R_n.$$

Then we replace $\|x_n-x^*\|^p$ in (3) by β_nR_n. We shall impose additional conditions on β_n later. For the time being assume that $\beta_n\in(0,1)$.

Notice that our assumption $\|x_n-x^*\|^p\leq R_n$ is not a boundedness condition for the operator. In fact, it is still possible that the range of the operator is unbounded. Notice further that we are allowed to overestimate $\|x_n-x^*\|^p$ by arbitrary large R_n. This is very important in real world applications. Usually some information on the magnitude of x^* is available (e.g. by the structure of the operator). Then it is easy to set $R_n=D(\|x_n\|+C)^p$, where C is the magnitude estimation and D is some big number.

We already noticed that $\alpha_n^*=\tau_p\alpha_n^+$. Our estimation of the exact distance $\|x_n-x^*\|^p$ of course also influences the optimal choice of τ_p. Therefore we replace it by some number $t_n\in(0,1)$. The introduction of this second parameter t_n is the main improvement to the construction proposed in [10]. We introduce now auxiliary variables

$$h_n := \left(\tfrac{(pk)^p}{G_p} \cdot \tfrac{R_n}{\|x_n - Tx_n\|^p}\right)^{\frac{1}{p-1}} \quad \text{and} \quad q := \tfrac{p}{p-1}. \tag{6}$$

Then all admissible α_n can be written as

$$\alpha_n = \tfrac{1}{pk} h_n \beta_n^{q-1} t_n$$

for some choice of β_n and t_n, where α_n is admissible if the related x_{n+1} admits $\|x_{n+1} - x^*\| \le \|x_n - x^*\|$.

For any fixed choice of $\alpha_n = \frac{1}{pk} h_n \beta_n^{q-1} t_n$ with $\beta_n \in (0,1)$ and $t_n \in (0,1)$ we have

$$\alpha_n^p G_p \|x_n - Tx_n\|^p = \tfrac{1}{(pk)^p} \cdot \tfrac{pk^{pq} R_n^q}{(G_p\|x_n - Tx_n\|^p)^{q-1}} \cdot \beta_n^{(q-1)p} t_n^p = h_n \beta_n^q t_n^p R_n.$$

By (2) we get for x_n with $\|x_n - x^*\|^p \ge \beta_n R_n$

$$\|x_{n+1} - x^*\|^p \le (1 - h_n t_n \beta_n^q + h_n \beta_n^q t_n^p) R_n\ . \tag{7}$$

On the other hand, if $\|x_n - x^*\|^p < \beta_n R_n$ we conclude again with (2) that

$$\|x_{n+1} - x^*\|^p \le (\beta_n + h_n \beta_n^q t_n^p) R_n. \tag{8}$$

The main idea of our construction is to choose β_n and t_n optimal, in the sense that the estimations in (7) and (8) are minimized. Thus such optimal β_n and t_n minimize

$$\max\{1 - h_n t_n \beta_n^q + h_n \beta_n^q t_n^p, \beta_n + h_n \beta_n^q t_n^p\}.$$

for fixed h_n. One can show (cf. Appendix), that the above maximum is minimal for β_n^{opt} and t_n^{opt}, where β_n^{opt} is the solution of the equation

$$\tau_p h_n \beta_n^{\frac{p+1}{p-1}} = 1 - \beta_n \tag{9}$$

and t_n^{opt} is defined via

$$t_n^{\mathrm{opt}} = \tau_p (\beta_n^{\mathrm{opt}})^{q-1},$$

where τ_p is the same as in (5). Then the optimal value of α_n, which we denote by α_n^{opt}, is given by

$$\begin{aligned}
\alpha_n^{\mathrm{opt}} &= \tfrac{1}{pk} h_n (\beta_n^{\mathrm{opt}})^{q-1} \tau_n (\beta_n^{\mathrm{opt}})^{q-1} \\
&= \tfrac{1}{pk} \tau_p h_n (\beta_n^{\mathrm{opt}})^{\frac{p+1}{p-1}} (\beta_n^{\mathrm{opt}})^{-1} \\
&= \tfrac{1}{pk} \tfrac{1-\beta_n^{\mathrm{opt}}}{\beta_n^{\mathrm{opt}}}\ .
\end{aligned}$$

One can check that

$$1 - h_n t_n^{\mathrm{opt}} (\beta_n^{\mathrm{opt}})^q + h_n (\beta_n^{\mathrm{opt}})^q (t_n^{\mathrm{opt}})^p = \beta_n^{\mathrm{opt}} + h_n (\beta_n^{\mathrm{opt}})^q (t_n^{\mathrm{opt}})^p\ .$$

By (7) or (8) we therefore get for α_n^{opt}

$$\begin{aligned}\|x_{n+1} - x^*\|^p &\leq (\beta_n^{\text{opt}} + h_n(\beta_n^{\text{opt}})^q (t_n^{\text{opt}})^p) R_n \\ &= \left(\beta_n^{\text{opt}} + \tfrac{1}{p}(1 - \beta_n^{\text{opt}})\beta_n^{\text{opt}}\right) R_n \\ &= \left(\left(1 + \tfrac{1}{p}\right)\beta_n^{\text{opt}} - \tfrac{1}{p}(\beta_n^{\text{opt}})^2\right) R_n \ .\end{aligned}$$

Thus the number R_{n+1} defined by

$$R_{n+1} := \left(\left(1 + \tfrac{1}{p}\right)\beta_n^{\text{opt}} - \tfrac{1}{p}(\beta_n^{\text{opt}})^2\right) R_n \tag{10}$$

fulfills

$$\|x_{n+1} - x^*\|^p \leq R_{n+1} < R_n,$$

where the right-hand side is true due to the fact that β_n^{opt} as defined by (9) is zero, only if x_n is already a fixed point of T. We of course assume that the iteration stops then.

Algorithm 1 (Adaptive Mann iteration)**.**

(S_0) *Choose an arbitrary $x_0 \in X$ with $\|x_0 - x^*\|^p \leq R_0$. Set $n = 0$.*

(S_1) *Stop if $Tx_n = x_n$, else compute the unique positive solution β_n of the equation*

$$p\left(\tfrac{k^p}{G_p} \cdot \tfrac{R_n}{\|x_n - Tx_n\|^p}\right)^{\frac{1}{p-1}} \beta_n^{\frac{p+1}{p-1}} = 1 - \beta_n. \tag{11}$$

(S_2) *Set*

$$\begin{aligned}\alpha_n &= \tfrac{1}{pk} \cdot \tfrac{1-\beta_n}{\beta_n} \\ R_{n+1} &= \left(\left(1 + \tfrac{1}{p}\right)\beta_n - \tfrac{1}{p}\beta_n^2\right) R_n.\end{aligned}$$

(S_3) *Set*

$$x_{n+1} = (1 - \alpha_n)x_n + \alpha_n T x_n.$$

(S_4) *Let $n \leftarrow (n+1)$ and go to step (S_1).*

The constants G_p and k are as introduced in Section 1.

The formulas in (S_1) and (S_2) follow directly from (3), (6), (9) and (10).

3. Convergence results

In this section we will state several theorems, which justify the name of adaptive Mann iterations. In fact, we will show that Algorithm 1 automatically adapts to the continuity of the underlying operator T.

In what follows let h_n and q be defined as in (6).

Theorem 1. *Let X be a p-smooth Banach space and $T : X \to X$ a strongly pseudocontractive map (with fixed point x^*), mapping bounded sets on bounded sets. Let R_0 be the initial guess on the distance $\|x_0 - x^*\|^p$. Then the sequence (x_n) defined by the Mann iteration of Algorithm 1 converges strongly to the fixed point of T (if T has a fixed point). The rate of convergence is given by*

$$\|x_n - x^*\| \le C n^{-\frac{p-1}{p}},$$

for some $C > 0$.

Proof. For the proof we use a nice trick also used in [1] and [8]. Consider

$$\begin{aligned}\frac{1}{R_{n+1}^{q-1}} - \frac{1}{R_n^{q-1}} &= \frac{1 - \left[\left(1+\frac{1}{p}\right)\beta_n - \frac{1}{p}\beta_n^2\right]^{q-1}}{\left[\left(1+\frac{1}{p}\right)\beta_n - \frac{1}{p}\beta_n^2\right]^{q-1}} \cdot R_n^{1-q} \\ &\ge 1 - \left[\left(1+\frac{1}{p}\right)\beta_n - \frac{1}{p}\beta_n^2\right]^{q-1} \cdot R_n^{1-q} \ .\end{aligned}$$

Since $q > 1$, we have

$$\begin{aligned}\frac{1}{R_{n+1}^{q-1}} - \frac{1}{R_n^{q-1}} &\ge 1 - \left[\left(1+\frac{1}{p}\right)\beta_n - \frac{1}{p}\beta_n^2\right] \cdot R_n^{1-q} \\ &\ge (1 - \tfrac{1}{p}\beta_n)(1-\beta_n) \cdot R_n^{1-q}.\end{aligned}$$

Due to $0 < \beta_n < 1$ we get

$$\frac{1}{R_{n+1}^{q-1}} - \frac{1}{R_n^{q-1}} \ge \tfrac{1}{q}(1-\beta_n) \cdot R_n^{1-q}.$$

By the Taylor expansion of the right-hand side in (9) or (11) (as a function of β_n), a simple geometric argument or by employing the implicit function theorem, we get that

$$1 - \beta_n \ge \frac{\tau_p h_n}{1+\frac{p+1}{p-1}\tau_p h_n}.$$

Therefore

$$\begin{aligned}\frac{1}{R_{n+1}^{q-1}} - \frac{1}{R_n^{q-1}} &\ge \tfrac{1}{q}(1-\beta_n) \cdot R_n^{1-q} \\ &\ge \tfrac{1}{q} \cdot \frac{\tau_p h_n}{1+\frac{p+1}{p-1}\tau_p h_n} \cdot R_n^{1-q}.\end{aligned}$$

Notice that

$$\tfrac{x}{1+x} \ge \tfrac{1}{2}\min\{1, x\}.$$

Hence

$$\frac{1}{R_{n+1}^{q-1}} - \frac{1}{R_n^{q-1}} \ge \tfrac{1}{2q} \cdot \tfrac{p-1}{p+1} \min\{1, \tfrac{p+1}{p-1}\tau_p h_n\} R_n^{1-q}.$$

By construction we have $R_n^{1-q} \geq R_0^{1-q}$. Therefore the numbers R_n^{1-q} are uniformly bounded away from zero. Since T maps bounded sets on bounded sets, for M defined by

$$M := \{\|Tx - Tx^*\|^p : x \in X, \|x - x^*\|^p \leq R_0\},$$

we have that

$$M < \infty.$$

Hence

$$\|x_n - Tx_n\|^p \leq 2^{p-1}(\|x_n - x^*\|^p + \|Tx^* - Tx_n\|^p) \leq 2^{p-1}(R_0 + M) < \infty.$$

Therefore by definition of h_n the numbers $h_n R_n^{1-q}$ are also uniformly bounded away from zero. Thus

$$\frac{1}{R_{n+1}^{q-1}} - \frac{1}{R_n^{q-1}} \geq c$$

for some $c > 0$. We have

$$\frac{1}{R_{n+1}^{q-1}} \geq \frac{1}{R_{n+1}^{q-1}} - \frac{1}{R_0^{q-1}} = \sum_{k=0}^{n} \frac{1}{R_{k+1}^{q-1}} - \frac{1}{R_k^{q-1}} \geq (n+1)c.$$

Then

$$R_{n+1}^{q-1} \leq c^{-1}(n+1)^{-1}.$$

Finally, we arrive at

$$\|x_n - x^*\| \leq R_n^{1/p} \leq c^{-\frac{p-1}{p}} n^{-\frac{p-1}{p}},$$

since $-1/(q-1) \cdot 1/p = -(p-1)/p$. □

Remark 1. *Let us mention:*

1. *Notice that the proof of strong convergence (without any convergence rate) can be carried out in the same way as the proof of Theorem 3.2 in [10].*

2. *We also remark that the theorem is also true, if T is only locally strongly accretive.*

3. *The crucial point in the estimations above is that M (as defined above) is finite. It would therefore suffice to demand that the range of the R_0 norm ball around the fixed point is bounded, instead of the slightly more restrictive assumption that the range of all bounded sets is bounded.*

We can now prove that the Adaptive Mann iteration of Algorithm 1 automatically adapts the continuity of T.

Theorem 2. *Assume that additionally to the assumptions made in Theorem 1 the operator T is (locally) Hlder continuous, i.e. for every bounded set there exist some $0 < \gamma < 1$ and $\kappa > 0$, such that*

$$\|Tx - Ty\| \leq \kappa \|x - y\|^\gamma$$

for all x, y in that bounded set. Then the Mann iteration converges as

$$\|x_n - x^*\| \leq Cn^{-\frac{1}{1-\gamma}\cdot\frac{p-1}{p}},$$

for some $C > 0$.

Proof. Along the lines of the proof in Theorem 1 we get

$$\frac{1}{R_{n+1}^{(1-\gamma)(q-1)}} - \frac{1}{R_n^{(1-\gamma)(q-1)}} \geq \tfrac{1}{q} \cdot \frac{\tau_p h_n}{1+\frac{p+1}{p-1}\tau_p h_n} \cdot R_n^{(1-\gamma)(1-q)}.$$

We notice that since the identity operator is Lipschitz, it is also locally Hlder. Therefore the operator $I - T$ is also locally Hlder. We get then

$$\|x_n - Tx_n\|^p = \|(I-T)x_n - (I-T)x^*\|^p \leq C\|x_n - x^*\|^{\gamma p} \leq CR_n^{\gamma}$$

for some generic constant $C > 0$. Hence,

$$h_n \geq CR_n^{(1-\gamma)(q-1)}.$$

We arrive at

$$\frac{1}{R_{n+1}^{(1-\gamma)(q-1)}} - \frac{1}{R_n^{(1-\gamma)(q-1)}} \geq C.$$

Therefore as in the proof of Theorem 1 we conclude

$$R_{n+1}^{(1-\gamma)(q-1)} \leq Cn^{-1}.$$

Finally, we arrive at

$$\|x_n - x^*\| \leq R_n^{-1/p} \leq Cn^{-\frac{1}{1-\gamma}\cdot\frac{p-1}{p}}.$$

□

Remark 2. *Let us mention:*

1. *Notice that the Hlder exponent γ does not have to be given in Algorithm 1. Despite the fact that the algorithm does not have the information about the Hlder continuity, it is automatically able to achieve a faster convergence rate than for merely bounded operators. We think that this behavior of Algorithm 1 justifies the 'adaptivity' in the name.*
2. *We would also like to remark that the last theorem induces that Algorithm 1 may perform better than the algorithm proposed by Chidume in [5] for certain operators.*

Theorem 3. *Assume that additionally to the assumptions made in Theorem 1 the operator T is (locally) Lipschitz continuous, i.e. for every bounded set there exists some $\kappa > 0$, such that*

$$\|Tx - Ty\| \leq \kappa\|x - y\|$$

for all x, y in that bounded set. Then the Mann iteration converges (almost) linearly, i.e. there exists a constant $C > 0$, such that

$$\|x_n - x^*\| \leq C \cdot \exp(-n/C).$$

Proof. For this very special case the proof is rather simple. Consider

$$R_{n+1} = \left(\left(1+\tfrac{1}{p}\right)\beta_n - \tfrac{1}{p}\beta_n^2\right) R_n.$$

We show that β_n are bounded away from one. From (9) we see that $\beta_n \to 1$ only if $h_n \to 0$. With T also $I-T$ is (locally) Lipschitz. Consider therefore with some generic constant $C > 0$

$$h_n^{p-1} \geq C \tfrac{R_n}{\|x_n - Tx_n\|^p} = C \tfrac{R_n}{\|(I-T)x_n-(I-T)x^*\|^p} \geq C \tfrac{R_n}{R_n} \geq C.$$

Hence there exists some $\Gamma < 1$, such that

$$R_{n+1} \leq \Gamma R_n$$

for all $n \geq 1$. Therefore the sequence (R_n) converges geometrically. But (R_n) also majorizes the sequence $(\|x_n - x^*\|^p)$, which proves the claim. □

Remark 3. *Again we remark that Algorithm 1 does not need the information about the Lipschitz continuity to converge (almost) linearly.*

Notice that the results of Theorem 1 extend canonically to strongly accretive mappings. Therefore, as a consequence of Theorem 1 we can prove the following corollary:

Corollary 1. *Let X be a p-smooth Banach space, $f \in X$ and $S : X \to X$ strongly accretive. Suppose S maps bounded sets on bounded sets and suppose the equation*

$$Sx = f$$

has a solution. Then this solution is unique and the sequence (x_n) defined by the Mann iteration of Algorithm 1 with $T : X \to X$ and

$$T(x) = f + x - Sx$$

converges strongly to this unique solution x^* with the rate

$$\|x_n - x^*\| \leq C n^{-\frac{p-1}{p}}.$$

If in addition S is also γ-Hlder continuous, then the rate improves to

$$\|x_n - x^*\| \leq C n^{-\frac{1}{1-\gamma}\cdot\frac{p-1}{p}}.$$

If S is Lipschitz continuous, then the (x_n) converge linearly to the solution x^*.

Remark 4. *If in addition to strong pseudocontractivity in Theorem 1 (resp. strong accretivity in Corollary 1) we assume that T (resp. S) is Lipschitz continuous, then the existence of the fixed point of T (resp. solution of $Sx = f$) follows from [2, 6].*

We remark that our approach can also be extended to set-valued strong pseudocontractions and strongly accretive mappings without any difficulties.

4. Numerical example

In this section we consider a simple example to visualize the strength of the adaptive Mann iterations. We consider the rotation

$$\mathrm{Rot}_{a,\phi} = \begin{pmatrix} \cos(\phi) & a \cdot \sin(\phi) \\ -a \cdot \sin(\phi) & \cos(\phi) \end{pmatrix}, \quad a \geq 1,$$

as a mapping from X to X, where X is the space $\mathbb{R}^2$ equipped with the euclidean norm. Hence the space X is a Hilbert space. One can check that $\mathrm{Rot}_{a,\phi}$ is strongly accretive with

$$k = 1 - \cos(\phi).$$

On the other hand, $\mathrm{Rot}_{a,\phi}$ is not a contraction for $a \geq 1$, since

$$\| \mathrm{Rot}_{a,\phi} \, x\|^2 = (\cos(\phi)^2 + a^2 \sin(\phi)^2)\|x\|^2.$$

Further, it is clear, that the fixed point is $(0,0)$. In all our simulations the starting point was $x_0 = (1,0)^T$ and the value of a was 10.

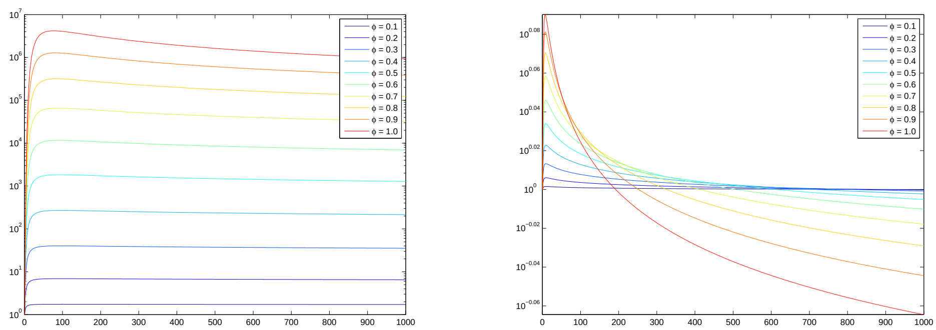
A classical assumption on the step-sizes α_n is given by

$$\sum \alpha_n^2 = 0 \qquad \text{and} \qquad \sum \alpha_n = \infty. \tag{12}$$

Hence a classical candidate for a step-size is

$$\alpha_n = \frac{1}{n}.$$

However, if condition (12) is fulfilled for some (α_n), then it is also fulfilled for all $c \cdot (\alpha_n)$ with $0 < c < \infty$.



Results for $c = 1$ *Results for* $c = 0.1$

Figure 1. *Convergence results of standard step-size rule* $\alpha_n = c/n$ *for the first* 1000 *iterations and variable values of* ϕ *from* 0.1 *to* 1.0 *in* 0.1 *steps. The y-axis denotes the squared distance of the current iterate to the fixed-point*

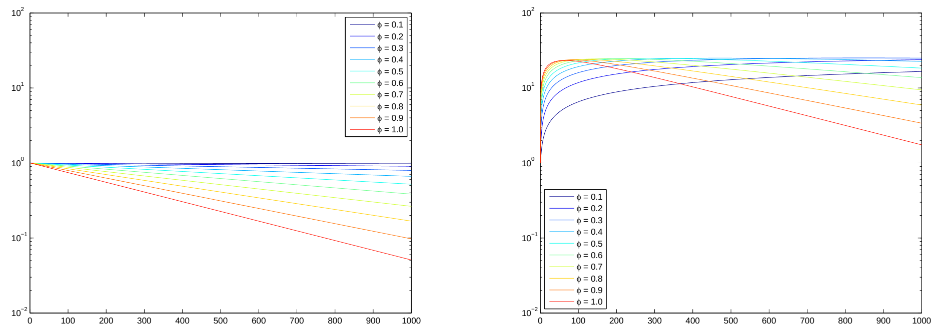
In the first step of our analysis we have to clarify how the choice of a particular c influences the behavior of the Mann iteration. A typical result is displayed in

Figure 1. We see that for the choice $c = 1$ the iterates diverge from the fixed point in the few first iterations. They begin to converge only in the late iterations. We further see that the convergence is lethargic in the sense that the lines representing the convergence are very flat. This behavior was to be expected.

We further observe that for big values of ϕ the iterates diverge more than for small values. And in general they converge rather slow. This is remarkable, since the bigger ϕ, the more accretive the problem. Therefore we expect the iterates to converge faster with increasing the value of ϕ. We conclude that for $c = 1$ the iterates behave counterintuitive. However, the expected behavior that stronger accretiveness (bigger k) implies faster convergence if we choose $c = 0.1$. Although we do not visualize it here, we remark that if the value of c is chosen too small, the results are again getting counterintuitive. Altogether we see that the right choice of the constant c is crucial in applications. However, the standard convergence proofs do not deliver any information on the appropriate choice of c.

This is completely different for the adaptive Mann iterations as we can see in Figure 2. First we have chosen the initial estimate on the distance to 1, which is the true distance of the start point $x_0 = (1, 0)^T$ to the fixed point $(0, 0)$. We see that for adaptive Mann iterations the full information on accretivity is used immediately by the iteration. The more accretive the mapping, the faster the convergence of the Mann iterations.

On the other hand, the information about the exact distance to the fixed point is usually not available. However, usually some sensible estimate on the magnitude of the distance is available due to the constraints of the problem. What happens if we overestimate the value R_0 by several orders of magnitude? An example is again given in Figure 2. We have tested our results for $R_0 = 100$. Hence we overestimated the distance function by two orders of magnitude. But again the accretiveness is



Results for $R_0 = 1$ *Results for* $R_0 = 100$

Figure 2. *Convergence results of adaptive Mann iterations for the first* 1000 *iterations and variable values of* ϕ *from* 0.1 *to* 1.0 *in* 0.1 *steps. The* y*-axis denotes the squared error of the current iterate to the fixed-point*

transported by means of convergence speed. The more accretive the mapping, the faster the convergence. Altogether we see that the results for Mann iterations always agree with the intuition.

At last we notice that the adaptive Mann iterations converge faster than the standard version. The slope of the distances in Figure 1 (representing the standard Mann iteration) is getting flatter and flatter with increasing the number of iterations. This is not the case for the adaptive Mann iterations in Figure 1. This was to be expected after the results of Theorem 3.

We summarize: In our numerical simulations adaptive Mann iterations converged faster than standard Mann iterations. And the results of the adaptive Mann iterations with respect to the dependency on the accretivity were more intuitive in comparison with the standard Mann iterations. Altogether we dare say that for the class of operators discussed in this paper (i.e. strongly accretive) the adaptive Mann iterations have better convergence properties than the standard Mann iterations.

5. Appendix

We show that for fixed h_n the maximum

$$\max\{1 - h_n t_n \beta_n^q + h_n \beta_n^q t_n^p, \beta_n + h_n \beta_n^q t_n^p\} \tag{13}$$

is minimal for β_n^{opt} and t_n^{opt}, with β_n^{opt} the solution of the equation

$$\tau_p h_n \beta_n^{\frac{p+1}{p-1}} = 1 - \beta_n \tag{14}$$

and t_n^{opt} defined via

$$t_n^{\text{opt}} = \tau_p (\beta_n^{\text{opt}})^{q-1},$$

where τ_p is the same as in (5).

We consider the function

$$(\beta_n, t_n) \mapsto \max\{1 - h_n t_n \beta_n^q + h_n \beta_n^q t_n^p, \beta_n + h_n \beta_n^q t_n^p\}.$$

One can easily see that for small β_n and small t_n the maximum is dominated by the first term. For large β_n and large t_n the maximum is dominated by the second term. We use this observation in our analysis.

For fixed h_n and β_n the function

$$t_n \mapsto 1 - h_n t_n \beta_n^q + h_n \beta_n^q t_n^p \tag{15}$$

is minimal for

$$t_n = \tau_p\ .$$

Next we notice that the function

$$\beta_n \mapsto 1 - h_n \tau_n \beta_n^q + h_n \beta_n^q \tau_n^p$$

is monotonically decreasing in β_n. We also notice that for $t_n = \tau_p$ the second term of maximum (13), i.e. the function

$$\beta_n \mapsto \beta_n + h_n \beta_n^q \tau^p$$

is monotonically decreasing. Hence the first term of maximum (13) is bigger than the second as long as

$$1 - h_n\tau_n\beta_n^q + h_n\beta_n^q\tau_n^p \geq \beta_n + h_n\beta_n^q\tau^p$$

or equivalently

$$1 - \beta_n \geq h_n\tau_n\beta_n^q.$$

We know that for $\beta_n \in (0,1)$ this inequality is fulfilled in some subinterval

$$(0, \beta_n^-),$$

where β_n^- is the solution of

$$\tau_p h_n\beta_n^q = 1 - \beta_n. \tag{16}$$

Altogether we have proven, that for $(\beta_n, t_n) \in (0, \beta_n^-) \times (0,1)$ the maximum (13) is minimal for $(\beta_n, t_n) = (\beta_n^-, \tau_n)$.

Next we consider the region $(\beta_n^-, 1) \times (0,1)$. We notice that for fixed $\beta_n > \beta_n^-$ the function

$$t_n \mapsto \beta_n + h_n\beta_n^q\tau_n^p \tag{17}$$

is monotonically increasing. On the other hand, for small values of t_n the function in (17) is bigger than the function in (15). From the above considerations we know that for $t_n = \tau_n$ we have $1 - h_n t_n\beta_n^q + h_n\beta_n^q t_n^p < \beta_n + h_n\beta_n^q t_n^p$. And we know that in the interval $(0, \tau_p)$ the function in (15) is monotonically decreasing. Together, we have that for fixed $\beta_n > \beta_n^-$ maximum (13) is minimal, if

$$1 - h_n t_n\beta_n^q + h_n\beta_n^q t_n^p = \beta_n + h_n\beta_n^q t_n^p$$

or equivalently

$$1 - \beta_n = h_n t_n\beta_n^q. \tag{18}$$

We now consider the second term in maximum (13) along the line of (β_n, t_n) fulfilling (18). (Of course we could also consider the first term, since along line (18) both sides have the same value.) We have then

$$\beta_n + h_n\beta_n^q t^p = \beta_n + h_n\beta_n^q \left(\tfrac{1-\beta_n}{\beta_n^q} \cdot \tfrac{1}{h_n^p}\right).$$

The right hand-side in the last equation, as function of β_n, is minimal for β_n fulfilling

$$\tau_p h_n\beta_n^{\frac{p+1}{p-1}} = 1 - \beta_n. \tag{19}$$

We denote the solution of the above equation by β_n^+. Then the point (β_n^+, t_n^+) with t_n^+ defined by

$$t_n^+ = \tau_p(\beta_n^+)^{q-1}$$

is located on the line described by (18). We notice that

$$q = \tfrac{p}{p-1} < \tfrac{p+1}{p-1}.$$

Therefore comparing (16) and (19), we see that

$$\beta_n^- < \beta_n^+.$$

Hence, for $(\beta_n, t_n) \in (\beta_n^-, 1)\times(0,1)$ maximum (13) is minimal for $(\beta_n, t_n) = (\beta_n^+, t_n^+)$. Since the point (β_n^-, τ_n) is also located on the line fulfilling (18), we get that (β_n^+, t_n^+) is the minimizer of maximum (13) in the whole square $(0,1)\times(0,1)$. This shows the assertion.

References

[1] K. Bredies, D. A. Lorenz, *Iterated hard shrinkage for minimization problems with sparsity constraints*, SIAM Journal on Scientific Computing, to appear.

[2] F. Browder, *Nonlinear mappings of nonexpansive and accretive type in Banach spaces*, Bull. Amer. Math. Soc. **73**(1967), 875–882.

[3] C. E. Chidume, *An iterative process for nonlinear lipschitzian strongly accretive mappings in L_p spaces*, J. Math. Anal. Appl. **151**(1990), 453–461.

[4] C. E. Chidume, *Iterative solution of nonlinear equations with strongly accretive operators*, J. Math. Anal. Appl. **192**(1995), 502–518.

[5] C. Chidume, *Steepest descent method for locally accretive mappings*, J. Korean Math. Soc. **33**(1996), 1–12.

[6] K. Deimling, *Zeros of accretive operators*, Manuscripta Math. **13**(1974), 365–374.

[7] L. Deng, *An iterative process for nonlinear lipschitz and strongly accretive mappings in uniformly convex and uniformly smooth Banach spaces*, Nonlinear Analysis **24**(1995), 981–987.

[8] J. C. Dunn, *Rates of convergence for conditional gradient algorithms near singular and nonsingular extremals*, SIAM Journal on Control and Optimization **17**(1979), 187–211.

[9] O. Hanner, *On the uniform convexity of L_p and ℓ_p*, Ark. Mat. **3**(1956), 239–244.

[10] K. S. Kazimierski, *Adaptive Mann iterations for nonlinear accretive and pseudocontractive operator equations*, Mathematical Communications **13**(2008), 33–44.

[11] W. Mann, *Mean value in iteration*, Proc. Am. Math. Soc. **4**(1953), 506–510.

[12] M. A. Noor, *Some developments in general variational inequalities*, Appl. Math. Compu. **152**(2007), 199–277.

[13] M. O. Osilike, *Ishikawa and Mann iteration methods with errors for nonlinear equations of the accretive type*, J. Math. Anal. Appl. **213**(1997), 91–105.

[14] A. Rafiq, *Mann iterative scheme for nonlinear equations*, Mathematical Communications **12**(2007), 25–31.

[15] Z. -B. Xu, G. Roach, *Characteristic inequalities of uniformly convex and uniformly smooth Banach spaces*, J. Math. Anal. Appl. **157**(1991), 189–210.