DE GRUYTER DOI 10.1515/jiip-2014-0035 — J. Inverse Ill-Posed Probl. 2014; aop

Research Article

Fábio Margotti and Andreas Rieder

An inexact Newton regularization in Banach spaces based on the nonstationary iterated Tikhonov method

Abstract: A version of the nonstationary iterated Tikhonov method was recently introduced to regularize *linear* inverse problems in Banach spaces [7]. In the present work we employ this method as inner iteration of the inexact Newton regularization method REGINN [14] which stably solves *nonlinear* ill-posed problems. Further, we propose and analyze a Kaczmarz version of the new scheme which allows fast solution of problems which can be split into smaller subproblems. As special cases we prove strong convergence of Kaczmarz variants of the Levenberg–Marquardt and the iterated Tikhonov methods in Banach spaces.

Keywords: Inexact Newton regularization in Banach spaces, nonstationary iterated Tikhonov method, Levenberg–Marquardt scheme, Kaczmarz iteration

MSC 2010: 65J20, 65J15

Fábio Margotti, Andreas Rieder: Department of Mathematics, Karlsruhe Institute of Technology (KIT), 76128 Karlsruhe, Germany, e-mail: margotti@partner.kit.edu, andreas.rieder@kit.edu

1 Introduction

We consider nonlinear ill-posed problems

$$F(x) = y \tag{1.1}$$

in the abstract framework of Banach spaces, that is, $F\colon D(F) \subset X \to Y$ operates between *Banach spaces* X and Y where $D(F)$ denotes the domain of definition of F. Recently, this setting has been attracting and still attracts a lot of research since several real-life applications are naturally modeled with the help of Banach spaces, see, e.g., the first chapter in [16].

The starting point for our investigation is the Newton-type algorithm REGINN (REGularization based on INexact Newton iteration) [14] which improves the current iterate x_n via

$$x_{n+1} = x_n + s_n$$

by a correction step s_n obtained from approximately solving a local linearization of (1.1):

$$A_n s = b_n \tag{1.2}$$

where $A_n := F'(x_n)$ is the Fréchet derivative of F in x_n and $b_n := y - F(x_n)$ is the corresponding nonlinear residual. REGINN typically applies an iterative solver to (1.2), called *inner iteration*, to find a step satisfying

$$\|A_n s_n - b_n\| < \mu \|b_n\| \tag{1.3}$$

with a pre-defined constant $0 < \mu < 1$.

Now, assume that (1.1) splits into $d \in \mathbb{N}$ 'smaller' subproblems, that is, the Banach space Y factorizes into Banach spaces $Y_0, \dots, Y_{d-1}$: $Y = Y_0 \times Y_1 \times \cdots \times Y_{d-1}$. Accordingly, $F = (F_0, F_1, \dots, F_{d-1})^\top$, $F_j\colon D(F) \subset X \to Y_j$, and $y = (y_0, y_1, \dots, y_{d-1})^\top$. Our task can be recast as: find $x \in D(F)$ such that

$$F_j(x) = y_j, \quad j = 0, \dots, d-1. \tag{1.4}$$

The Kaczmarz variant of REGINN determines s_n from (1.3) where, however,

$$A_n := F'_{n \bmod d}(x_n) \quad \text{and} \quad b_n := y_{n \bmod d} - F_{n \bmod d}(x_n).$$

Brought to you by | De Gruyter / TCS
Authenticated
Download Date | 12/9/14 4:09 PM

Thus, the subsystems are processed cyclically breaking the large-scale system (1.1) into handy pieces. This kind of cycling strategy was initiated by Kaczmarz [9] and first analyzed in the context of nonlinear ill-posed problems by Kowar and Scherzer [10] and Haltmeier, Leitão and Scherzer [4]. For further result see, for example, [2, 8, 12].

We emphasize that systems like (1.4) arise quite naturally in applications where the data is measured by d individual experiments or observations. For instance, in electrical impedance tomography one wants to find the conductivity of an object by applying, say, d current patterns at the boundary and measuring the resulting voltages at the boundary as well.

In this work we consider the iterated Tikhonov regularization as suggested by Jin and Stals [7] as inner iteration of REGINN. The resulting scheme is called K-REGINN-IT which is short for Kaczmarz version of the REGINN-Iterated-Tikhonov method. As a byproduct we thus generalize the Levenberg–Marquardt regularization (Hanke [5]) to Banach spaces.

On the following pages we present a complete convergence analysis of K-REGINN-IT under the usual assumptions. As our setting is rather abstract and as our arguments are sometimes very technical we try to guide the reader gently through the exposition. Therefore, we collect needed properties and concepts of Banach spaces in Section 2. This material is taken from [3, 15–17]. In Section 3 we define K-REGINN-IT, prove its well-definedness and termination. Next we validate strong convergence in the noise-free situation (Section 4) and finally show the regularization property in Section 5.

To keep this exposition lean we restrained from presenting numerical examples. In a forthcoming paper we plan to compare numerically different types of K-REGINN methods. The impatient reader is referred to [13] where we solve the inverse problem of electrical impedance tomography by K-REGINN with an inner iteration of Landweber type.

2 Basic facts about the geometry of Banach spaces

If the context is clear, we always use a generic constant $C > 0$ even if it takes different values at different instances. Sometimes we write $a(x) \lesssim b(x)$ if and only if there exists a positive constant C independent of x such that $a(x) \leq C\, b(x)$ for all x.

In the following we formulate the assumptions on the Banach space X which we will need later to define and to analyze our method properly.

To cover the lack of an inner product in a general Banach space, we introduce the *duality mapping*. For an arbitrary (but fixed) number $p > 1$ this is the set-valued function $J_p \colon X \to 2^{X^*}$ defined by[1]

$$J_p(x) := \{x^* \colon X \to \mathbb{R} : x^* \text{ real-linear and continuous, } \langle x^*, x\rangle = \|x^*\|\,\|x\| \text{ and } \|x^*\| = \|x\|^{p-1}\}$$

where the symbol $\langle\cdot,\cdot\rangle$ denotes the inner product in Hilbert spaces as well as the duality pairing in a general Banach space. For all $x \in X$, $J_p(x) \neq \emptyset$ (see [3]) and the relation

$$J_r(x) = \|x\|^{r-t} J_t(x) \tag{2.1}$$

holds for each $r, t > 1$. The duality mapping J_2 is called the *normalized* duality mapping and as a consequence of the Riesz representation theorem, it is the identity operator in any real Hilbert space. A *selection* $j_p \colon X \to X^*$ of the duality mapping J_p is a single-valued not necessarily continuous function satisfying $j_p(x) \in J_p(x)$ for all $x \in X$. We will often use the estimates

$$\langle j_p(x), y\rangle \leq \|x\|^{p-1}\|y\| \quad \text{and} \quad \langle j_p(x), x\rangle = \|x\|^p$$

which are immediate consequences from the definition of J_p. Observe that the inner product also shares these properties for $p = 2$.

1 The duality mapping is often defined in a more general way associated with a so called *gauge function*. We prefer to use here this particular definition, which is actually the duality mapping associated with the gauge function $t \mapsto t^{p-1}$.

Brought to you by | De Gruyter / TCS
Authenticated
Download Date | 12/9/14 4:09 PM

We suppose now that the Banach space X has the following geometrical properties where we use the notation $a \vee b := \max\{a, b\}$.

Assumption 2.1. (a) X is reflexive.

(b) For each $1 < p < \infty$, the duality mapping $J_p \colon X \to X^*$ is single-valued, continuous and invertible with a continuous inverse satisfying

$$J_p^{-1} = J_{p^*}^* : X^* \to X^{**} \cong X, \tag{2.2}$$

where p and p^* are conjugate numbers, i.e., $\frac{1}{p} + \frac{1}{p^*} = 1$.

(c) There exists a number $1 < s < \infty$ such that, for all $1 < p < \infty$, the conjugate numbers of s and p satisfy

$$\|J_{p^*}^*(x^*) - J_{p^*}^*(y^*)\| \le C_{p^*,s^*}(\|x^*\| \vee \|y^*\|)^{p^*-s^*}\|x^* - y^*\|^{s^*-1},$$

for all $x^*, y^* \in X^*$, where $C_{p^*,s^*} > 0$ is a constant which depends only on p and s.

Remark 2.2. We would like to comment on Assumption 2.1: If X is s-convex, it is also uniformly convex and Assumption 2.1 (a) follows, see [3, Chapter II, Theorem 2.9].

If X is additionally uniformly smooth, the norms in X and X^* are both Fréchet differentiable [3, Chapter I, Theorem 3.12, and Chapter II, Theorem 2.13], which implies that each duality mapping is single-valued, continuous and invertible with continuous inverse [3, Chapter II, Corollary 3.15]. Furthermore, (2.2) holds due to [3, Chapter II, Corollary 3.5].

As a consequence of results from [17], Assumption 2.1 (c) is equivalent to X^* being an s^*-smooth Banach space, which, in turn, is equivalent to the s-convexity of X (see [16, Theorems 2.42 and 2.52]).

In particular, Assumption 2.1 holds whenever X is s-convex and uniformly smooth.

As a substitute for the polarization identity

$$\frac{1}{2}\|x - y\|^2 = \frac{1}{2}\|x\|^2 - \langle x, y\rangle + \frac{1}{2}\|y\|^2,$$

which holds in real Hilbert spaces, we introduce the *Bregman distance* $\Delta_p \colon X \times X \to \mathbb{R}$,

$$\Delta_p(x, y) := \frac{1}{p}\|x\|^p - \langle J_p(y), x\rangle + \frac{1}{p^*}\|J_p(y)\|^{p^*}.$$

In any real Hilbert space, $\Delta_2(x, y) = \frac{1}{2}\|x - y\|^2$. A straightforward calculation shows the equality[2]

$$\Delta_p(x, y) = \frac{1}{p}\|x\|^p - \frac{1}{p}\|y\|^p - \langle J_p(y), x - y\rangle$$

and the *three-points identity*

$$\Delta_p(x, y) = \Delta_p(z, y) - \Delta_p(z, x) + \langle J_p(x) - J_p(y), x - z\rangle, \tag{2.3}$$

for all $x, y, z \in X$. Further,

$$\Delta_p(x, y) \ge \frac{1}{p}\|x\|^p + \frac{1}{p^*}\|y\|^p - \|y\|^{p-1}\|x\|.$$

Using now Young's inequality[3], we find that $\Delta_p(\cdot, \cdot) \ge 0$. Moreover, if $(x_n)_{n\in\mathbb{N}} \subset X$ is a sequence and $x \in X$ is a fixed vector, then $\Delta_p(x, x_n) \le \rho$ implies

$$\|x_n\|^{p-1}\Big(\frac{1}{p^*}\|x_n\| - \|x\|\Big) \le \rho.$$

Considering now the cases $\frac{1}{p^*}\|x_n\| - \|x\| \le \frac{1}{2p^*}\|x_n\|$ and $\frac{1}{p^*}\|x_n\| - \|x\| > \frac{1}{2p^*}\|x_n\|$, we conclude the implication

$$\Delta_p(x, x_n) \le \rho \implies \|x_n\| \le 2p^*(\|x\| \vee \rho^{1/p}). \tag{2.4}$$

2 This equivalent form is probably the most common definition of Bregman distance in Banach spaces with single-valued duality mappings.

3 For all $a, b \ge 0$, $ab \le \frac{a^p}{p} + \frac{b^{p^*}}{p^*}$.

Brought to you by | De Gruyter / TCS
Authenticated
Download Date | 12/9/14 4:09 PM

Therefore, $(x_n)_{n\in\mathbb{N}}$ is bounded. A similar result can be proven if $\Delta_p(x_n, x) \le \rho$. The continuity of the duality mapping (Assumption 2.1 (b)) is handed down to both arguments of the Bregman distance Δ_p. Of course $x = y$ implies $\Delta_p(x, y) = \Delta_p(y, x) = 0$. The reciprocal and other important results are true under further properties of X:

Assumption 2.3. (a) The functional $\|\cdot\|^p$ is strictly convex for any $1 < p < \infty$.
(b) Any sequence $(x_n)_{n\in\mathbb{N}} \subset X$ satisfying $x_n \rightharpoonup x$ and $\|x_n\| \to \|x\|$ as $n \to \infty$, converges strongly: $\|x_n - x\| \to 0$ as $n \to \infty$.
(c) If $1 < p \le s < \infty$, then

$$\Delta_p(x, y) \ge C\|x - y\|^s$$

for all $x, y \in B_R(0, \|\cdot\|) := \{z \in X : \|z\| \le R\}$ where $C > 0$ depends only on p, s and R.

Assumption 2.3 (a) implies the strict convexity of Δ_p in its first argument, which in turn ensures that $x = y$ whenever $\Delta_p(x, y) = 0$. In fact, if $x \ne y$, then for $\lambda \in (0, 1)$, we derive the contradiction

$$0 \le \Delta_p(\lambda x + (1 - \lambda)y, y) < \lambda\Delta_p(x, y) + (1 - \lambda)\Delta_p(y, y) = 0.$$

For our convergence analysis we rely on both Assumptions 2.1 and 2.3. We emphasize that the properties of X listed therein are not entirely independent of each other: some of them can be proven assuming the others. Note that both assumptions hold true in case X is s-convex and uniformly smooth (for Assumption 2.3 see [3, Chapter II, Propositions 1.6 and 2.8] and [16, Corollary 2.61 (a)]; for Assumption 2.1 see Remark 2.2). Important examples are the spaces $L^p(\Omega)$, $l^p(\mathbb{N})$, and $W^{n,p}(\Omega)$ with $1 < p < \infty$. They are uniformly smooth[4] and $\max\{p, 2\}$-convex. For the Lebesgue spaces the duality mapping $J_p\colon L^p(\Omega) \to L^{p^*}(\Omega)$ is given by

$$J_p(f) = |f|^{p-1}\,\mathrm{sign}(f). \tag{2.5}$$

3 The K-REGINN-IT method

We will need a bunch of standard assumptions about the structure of the nonlinearity F.

Assumption 3.1. (a) Equation (1.1) has a solution $x^+ \in X$ and for a given and fixed number $1 < p < \infty$, there exists a $\rho > 0$ such that

$$B_\rho(x^+, \Delta_p) := \{v \in X : \Delta_p(x^+, v) < \rho\} \subset D(F).$$

(b) Suppose that all the functions F_j, $j = 0, \dots, d-1$, are continuously Fréchet differentiable in $B_\rho(x^+, \Delta_p)$ and that their Fréchet derivatives $F_j'\colon B_\rho(x^+, \Delta_p) \to \mathcal{L}(X, Y_j)$ are uniformly bounded by a constant $M > 0$:

$$\|F_j'(v)\| \le M \quad \text{for all } v \in B_\rho(x^+, \Delta_p) \text{ and } j = 0, \dots, d-1.$$

(c) (Tangential Cone Condition (TCC)): Suppose that

$$\|F_j(v) - F_j(w) - F_j'(w)(v - w)\| \le \eta\|F_j(v) - F_j(w)\|$$

for all $v, w \in B_\rho(x^+, \Delta_p)$ and $j = 0, \dots, d-1$, where $0 \le \eta < 1$ is a constant.

We suppose to access only noisy versions $y_j^{\delta_j}$ of the exact but unknown data $y_j = F_j(x^+)$ satisfying

$$\|y_j - y_j^{\delta_j}\| \le \delta_j. \tag{3.1}$$

The positive *noise levels* δ_j, $j = 0, \dots, d-1$, are assumed to be known. Further, define the maximal noise level

$$\delta := \max\{\delta_j : j = 0, \dots, d-1\} > 0. \tag{3.2}$$

As the spaces Y_j are arbitrary, the duality mapping J_r does not need to be single-valued (for any $r > 1$). Then, $j_r\colon Y_j \to Y_j^*$ represents a selection of J_r.

4 They are actually $\min\{p, 2\}$-smooth, which is a stronger property.

Brought to you by | De Gruyter / TCS
Authenticated
Download Date | 12/9/14 4:09 PM

Now, we define K-REGINN-IT recursively: Let $x_n \in D(F)$ be given. The inner iteration to compute the Newton step s_n starts with setting $z_{n,0} := x_n$ and produces $z_{n,k+1}$ recursively as minimizer in X of the strict convex functional

$$T^\delta_{n,k}(z) := \frac{1}{r}\|b^\delta_n - A_n(z - x_n)\|^r + \alpha_n \Delta_p(z, z_{n,k}) \tag{3.3}$$

with $\alpha_n > 0$. Here,

$$A_n := F'_{[n]}(x_n) \quad \text{and} \quad b^\delta_n := y^{\delta_{[n]}}_{[n]} - F_{[n]}(x_n)$$

where $[n] := n \bmod d$ denotes the remainder of integer division. Note that the minimizer $z_{n,k+1}$ of $T^\delta_{n,k} : X \to \mathbb{R}$ exists and is unique due to the strict convexity of Δ_p. Set $s_{n,k} := z_{n,k} - x_n$ and $x_{n+1} := x_n + s_{n,k_n}$ where the final (inner) index k_n is determined as follows: choose $\tau > 1$, $\mu \in (0, 1)$ and $k_{\max} \in \mathbb{N} \cup \{\infty\}$. Define $k_n = 0$ in case of

$$\|b^\delta_n\| \le \tau\delta_{[n]}. \tag{3.4}$$

Otherwise set

$$k_{\text{REG}} := \min\{k \in \{1, \dots, k_{\max}\} : \|b^\delta_n - A_n s_{n,k}\| < \mu\|b^\delta_n\|\}, \tag{3.5}$$

using $\min \emptyset = \infty$. Finally,

$$k_n = \begin{cases} k_{\text{REG}}, & k_{\text{REG}} \le k_{\max}, \\ k_{\max}, & k_{\text{REG}} > k_{\max}. \end{cases}$$

Note that $x_{n+1} = z_{n,k_n}$ and $x_{n+1} = x_n$ if and only if (3.4) holds. The outer iteration stops as soon as the discrepancy principle (3.4) is satisfied d times in a row. Our approximate solution of (1.1) is then x_N where $N = N(\delta)$ is the smallest number[5] which satisfies

$$\|y^{\delta_j}_j - F_j(x_N)\| \le \tau\delta_j, \quad j = 0, \dots, d - 1. \tag{3.6}$$

See Algorithm 1 for an implementation in pseudocode.

As $z_{n,k+1}$ is the minimizer of $T^\delta_{n,k}$, we have $T^\delta_{n,k}(z_{n,k+1}) \le T^\delta_{n,k}(z_{n,k})$ yielding

$$\|b^\delta_n - A_n s_{n,k+1}\| \le \|b^\delta_n - A_n s_{n,k}\|, \quad k = 0, \dots, k_n - 1, \tag{3.7}$$

where equality holds only if $z_{n,k+1} = z_{n,k}$ as $z_{n,k+1} \ne z_{n,k}$ results in $\alpha_n \Delta_p(z_{n,k+1}, z_{n,k}) > 0$.

Using the *optimality condition* $0 \in \partial T^\delta_{n,k}(z_{n,k+1})$, we arrive at

$$\alpha_n(J_p(z_{n,k+1}) - J_p(z_{n,k})) \in A^*_n J_r(b^\delta_n - A_n s_{n,k+1}).$$

Hence, there exists some selection j_r such that

$$J_p(z_{n,k+1}) = J_p(z_{n,k}) + \frac{1}{\alpha_n} A^*_n j_r(b^\delta_n - A_n s_{n,k+1}).$$

Remark 3.2. By definition of $s_{n,k}$ the above equality can be rewritten as the implicit iteration

$$z_{n,k+1} = J^*_{p^*}\Big(J_p(z_{n,k}) + \frac{1}{\alpha_n} A^*_n j_r(b^\delta_n - A_n(z_{n,k+1} - x_n))\Big) \tag{3.8}$$

which can be solved for $z_{n,k+1}$ by a fixed point iteration. The convergence is guaranteed in case X is 2-convex, Y is 2-smooth and $\alpha_n \ge \alpha_{\min}$ with $\alpha_{\min} > 0$ being a constant large enough, see Appendix A. Alternatively, one may apply a gradient method like steepest descent (see, e.g., [1]) directly to the nonlinear functional (3.3) to find its minimizer. This approach has the advantage of requiring weaker geometrical properties of the spaces X and Y. The convergence speed is however strongly affected by convexity and smoothness properties of these spaces (see more details in [16, Section 5.3]).

5 The number N is chosen by a posteriori strategy, it thus depends actually on δ and y^δ: $N = N(\delta, y^\delta)$. But we stick to the simpler notation $N = N(\delta)$.

Brought to you by | De Gruyter / TCS
Authenticated
Download Date | 12/9/14 4:09 PM

Algorithm 1. K-REGINN-IT

Input: x_N; (y^{δ_j}, δ_j); F_j; F'_j, $j = 0, \dots, d-1$; μ; $k_{\max}$; τ;

Output: x_N with $\|y_j^{\delta_j} - F_j(x_N)\| \le \tau\delta_j$, $j = 0, \dots, d-1$;

```
ℓ := 0; x_0 := x_N; c := 0;
while c < d do
  for j = 0 : d − 1 do
    n := ℓd + j;
    b_n^δ := y_j^{δ_j} − F_j(x_n); A_n := F'_j(x_n);
    if ‖b_n^δ‖ ≤ τδ_j then
      x_{n+1} := x_n; c := c + 1;
    else
      k := 0; z_{n,0} := x_n;
      choose α_n > 0 properly;
      repeat
        z_{n,k+1} := arg min_{z∈X} ( (1/r)‖b_n^δ − A_n(z − x_n)‖^r + α_n Δ_p(z, z_{n,k}) )
        k := k + 1;
      until ‖b_n^δ − A_n(z_{n,k} − x_n)‖ < μ‖b_n^δ‖ or k = k_max
      x_{n+1} := z_{n,k}; c := 0;
    end if
  end for
  ℓ := ℓ + 1;
end while
x_N := x_{ℓd−c};
```

Remark 3.3. If $k_{\max} = 1$, then $k_n \in \{0, 1\}$ for all $n \in \mathbb{N}$ and x_{n+1} minimizes

$$T_{n,0}^{\delta}(x) = \frac{1}{r}\|y_{[n]}^{\delta_{[n]}} - F_{[n]}(x_n) - F'_{[n]}(x_n)(x - x_n)\|^r + \alpha_n \Delta_p(x, x_n).$$

In real Hilbert spaces this functional reads (with $p = r = 2$)

$$T_{n,0}^{\delta}(x) = \frac{1}{2}\|y_{[n]}^{\delta_{[n]}} - F_{[n]}(x_n) - F'_{[n]}(x_n)(x - x_n)\|^2 + \frac{\alpha_n}{2}\|x - x_n\|^2$$

revealing K-REGINN-IT as Kaczmarz version of the Levenberg–Marquardt method in Banach spaces. In case of a linear problem ($F_j = A_j$ are linear for all j) we have

$$T_{n,0}^{\delta}(x) = \frac{1}{r}\|y_{[n]}^{\delta_{[n]}} - A_{[n]}x\|^r + \alpha_n \Delta_p(x, x_n)$$

and the method is now a Kaczmarz version of the iterated Tikhonov method defined in [7] for the particular case $d = 1$.

In the next theorem we prove that K-REGINN-IT is well defined and terminates.

Theorem 3.4. *Let X and Y be Banach spaces with X satisfying Assumptions 2.1 and 2.3 with $1 < p \le s \le r$. Let Assumption 3.1 hold true and start with $x_0 \in B_\rho(x^+, \Delta_p)$. Choose $\overline{\alpha} > 0$ and define the constants*

$$C_0 := 3\frac{2^{r-1}}{2^{r-1} - 1} p^* M(\|x^+\| \vee \rho^{\frac{1}{p}}) > 0$$

and

$$C_1 := 2p^*\left(\|x^+\| \vee \left(\rho + \frac{C_0^r}{\overline{\alpha} 2^{r-1}}\right)^{\frac{1}{p}}\right) > 0.$$

Brought to you by | De Gruyter / TCS
Authenticated
Download Date | 12/9/14 4:09 PM

Define $\alpha_{\min} := \min\{\overline{\alpha}, \tilde{\alpha}\} > 0$, *where* $\tilde{\alpha} := C_2^{\frac{1}{1-s^*}} C_0^{\frac{(r-1)(s^*-1)-1}{s^*-1}} > 0$ *with*

$$0 < C_2 < \big(2^{\frac{1}{r^*}} C_{p^*,s^*} C_1^{(p^*-s^*)(p-1)} M^{s^*}\big)^{-1}$$

where C_{p^*,s^*} *is the constant from Assumption 2.1* (c). *Additionally, assume that the constant of the TCC in Assumption 3.1* (c) *satisfies* $0 \le \eta < C_3$ *where*

$$C_3 := \frac{1}{2^{r-1}} - C_4^r > 0$$

with

$$C_4 := C_{p^*,s^*} C_1^{(p^*-s^*)(p-1)} M^{s^*} C_2 > 0.$$

Further, let $\alpha_n \in [\alpha_{\min}, \alpha_{\max}]$ *for some* $\alpha_{\max} > \alpha_{\min}$, $\tau > (\eta+1)(C_3-\eta)^{-1}$ *and* $\mu \in ((\frac{\eta+1}{\tau} + \eta)C_3{}^{-1}, 1)$. *Then, there exists an* $N(\delta)$ *such that all iterates* $\{x_1, \dots, x_{N(\delta)}\}$ *of* K-REGINN-IT *are well defined and remain in* $B_\rho(x^+, \Delta_p)$. *Moreover, only the final iterate satisfies* (3.6) *and*

$$\Delta_p(x^+, x_n) \le \Delta_p(x^+, x_{n-1}) \tag{3.9}$$

for all $1 \le n \le N(\delta)$ *where equality holds if and only if* (3.4) *applies. Furthermore,*

$$\|x_n\| \le C \quad \text{for all } \delta > 0 \text{ and } n \le N(\delta), \tag{3.10}$$

with $C > 0$ *being independent of* n, $N(\delta)$, *and* δ.

Proof. First, observe that the bounds on C_2 imply $C_4^r < 2^{-(r-1)}$ which makes C_3 well defined. The lower bound on τ guarantees that we can select μ from an open interval.

We argue inductively: Suppose that $x_0, \dots, x_n$ are in $B_\rho(x^+, \Delta_p)$ and (3.9) holds. Further assume that x_n is not the final iterate, i.e., (3.6) is not satisfied for $N = n$. If x_n satisfies (3.4), then $x_{n+1} = x_n \in B_\rho(x^+, \Delta_p)$ and (3.9) becomes an equality. In case x_n violates (3.4), x_{n+1} is also well defined as we demonstrate in three steps:

(1) We show that

$$\|z_{n,l}\| \le C_1 \quad \text{for all } l \le k_n. \tag{3.11}$$

(2) We use (3.11) to validate that k_n is finite. Thus, x_{n+1} is well defined in X.

(3) We derive the monotonicity (3.9). Then, x_{n+1} is in $B_\rho(x^+, \Delta_p)$.

(1) Defining $e_n := x^+ - x_n$ and, by (2.3), we find that

$$\begin{aligned}
\Delta_p(x^+, z_{n,k+1}) - \Delta_p(x^+, z_{n,k}) &= -\Delta_p(z_{n,k+1}, z_{n,k}) + \langle J_p(z_{n,k+1}) - J_p(z_{n,k}), z_{n,k+1} - x^+\rangle \\
&\le \langle J_p(z_{n,k+1}) - J_p(z_{n,k}), z_{n,k+1} - x^+\rangle \\
&= \frac{1}{\alpha_n}\langle j_r(b_n^\delta - A_n s_{n,k+1}), A_n(s_{n,k+1} - e_n)\rangle \\
&= \frac{1}{\alpha_n}\langle j_r(b_n^\delta - A_n s_{n,k+1}), (b_n^\delta - A_n e_n) - (b_n^\delta - A_n s_{n,k+1})\rangle \\
&\le \frac{1}{\alpha_n}(\|b_n^\delta - A_n s_{n,k+1}\|^{r-1}\|b_n^\delta - A_n e_n\| - \|b_n^\delta - A_n s_{n,k+1}\|^r)
\end{aligned} \tag{3.12}$$

for $k = 0, \dots, k_n - 1$. Now, using Assumption 3.1 (c), we get for all $k \le k_n - 2$,

$$\begin{aligned}
\|b_n^\delta - A_n e_n\| &\le \|y_{[n]}^{\delta_{[n]}} - y_{[n]}\| + \|F_{[n]}(x^+) - F_{[n]}(x_n) - F'_{[n]}(x_n)(x^+ - x_n)\| \\
&\le \delta_{[n]} + \eta\|F_{[n]}(x^+) - F_{[n]}(x_n)\| \\
&\le \delta_{[n]} + \eta(\|y_{[n]}^{\delta_{[n]}} - y_{[n]}\| + \|b_n^\delta\|) \\
&\le \delta_{[n]}(\eta + 1) + \eta\|b_n^\delta\| \\
&< \Big(\frac{\eta+1}{\tau} + \eta\Big)\|b_n^\delta\|.
\end{aligned} \tag{3.13}$$

As

$$\Big(\frac{\eta+1}{\tau} + \eta\Big)\|b_n^\delta\| \le \mu\|b_n^\delta\| \le \|b_n^\delta - A_n s_{n,k+1}\|^r, \quad k \le k_n - 2,$$

Brought to you by | De Gruyter / TCS
Authenticated
Download Date | 12/9/14 4:09 PM

and in view of (3.13) we conclude that the right-hand side of (3.12) is negative. Then, for all $l \le k_n$,

$$\sum_{k=0}^{l-1}(\Delta_p(x^+, z_{n,k+1}) - \Delta_p(x^+, z_{n,k})) \le \frac{1}{\alpha_n}\|b_n^\delta - A_n s_{n,l}\|^{r-1}\Big(\Big(\frac{\eta+1}{\tau}+\eta\Big)\|b_n^\delta\| - \|b_n^\delta - A_n s_{n,l}\|\Big)$$
$$\le \frac{(\frac{\eta+1}{\tau}+\eta)}{\alpha_n}\|b_n^\delta - A_n s_{n,l}\|^{r-1}\|b_n^\delta\|.$$

From (3.7), $\alpha_n \ge \overline{\alpha}$, $\tau > \frac{\eta+1}{C_3-\eta}$ and $C_3 \le \frac{1}{2^{r-1}}$ we deduce that

$$\Delta_p(x^+, z_{n,l}) \le \Delta_p(x^+, z_{n,0}) + \frac{1}{\overline{\alpha}2^{r-1}}\|b_n^\delta\|^r. \tag{3.14}$$

From (3.13) and $(\frac{\eta+1}{\tau}+\eta) \le C_3 \le \frac{1}{2^{r-1}}$,

$$\|b_n^\delta\| - \|A_n e_n\| \le \|b_n^\delta - A_n e_n\| \le \Big(\frac{\eta+1}{\tau}+\eta\Big)\|b_n^\delta\| \le \frac{1}{2^{r-1}}\|b_n^\delta\|,$$

yielding

$$\|b_n^\delta\| \le \frac{2^{r-1}}{2^{r-1}-1}M\|e_n\|.$$

By $\|e_n\| \le \|x_n\| + \|x^+\|$, the induction hypotheses $x_n \in B_\rho(x^+, \Delta_p)$, and by (2.4), we see that

$$\|b_n^\delta\| \le C_0.$$

As $z_{n,0} = x_n \in B_\rho(x^+, \Delta_p)$, it follows from (3.14),

$$\Delta_p(x^+, z_{n,l}) \le \rho + \frac{C_0^r}{\overline{\alpha}2^{r-1}},$$

which in view of (2.4) implies (3.11).

(2) In the next step we show that $k_n < \infty$ by bounding the expression

$$\frac{\|b_n^\delta\|}{\alpha_n}\sum_{k=0}^{l-1}\|b_n^\delta - A_n s_{n,k}\|^{r-1}, \quad l \le k_n,$$

see (3.18) below. To this end we consider $\|b_n^\delta - A_n s_{n,k}\|^r$. As the function $\|\cdot\|^r$ is convex, we obtain

$$\|b_n^\delta - A_n s_{n,k}\|^r \le (\|b_n^\delta - A_n s_{n,k+1}\| + \|A_n(s_{n,k+1} - s_{n,k})\|)^r$$
$$\le (\|b_n^\delta - A_n s_{n,k+1}\| + M\|z_{n,k+1} - z_{n,k}\|)^r$$
$$\le 2^{r-1}(\|b_n^\delta - A_n s_{n,k+1}\|^r + M^r\|z_{n,k+1} - z_{n,k}\|^r)$$

which yields

$$-\|b_n^\delta - A_n s_{n,k+1}\|^r \le -\frac{1}{2^{r-1}}\|b_n^\delta - A_n s_{n,k}\|^r + M^r\|z_{n,k+1} - z_{n,k}\|^r. \tag{3.15}$$

To bound the rightmost term we use Assumption 2.1 (c) and note that $p^* - s^* \ge 0$ for $p \le s$,

$$\|z_{n,k+1} - z_{n,k}\| = \|J_{p^*}^*(J_p(z_{n,k+1})) - J_{p^*}^*(J_p(z_{n,k}))\|$$
$$\le C_{p^*,s^*}(\|J_p(z_{n,k+1})\| \vee \|J_p(z_{n,k})\|)^{p^*-s^*}\|J_p(z_{n,k+1}) - J_p(z_{n,k})\|^{s^*-1}$$
$$= C_{p^*,s^*}(\|z_{n,k+1}\| \vee \|z_{n,k}\|)^{(p^*-s^*)(p-1)}\Big\|\frac{1}{\alpha_n}A_n^* j_r(b_n^\delta - A_n s_{n,k+1})\Big\|^{s^*-1}$$
$$\le C_{p^*,s^*}C_1^{(p^*-s^*)(p-1)}\alpha_n^{1-s^*}M^{s^*-1}\|b_n^\delta - A_n s_{n,k+1}\|^{(r-1)(s^*-1)}$$

for all $k \le k_n - 1$. We proceed using (3.7), $\alpha_n \ge \tilde{\alpha}$, and $(r-1)(s^*-1) - 1 \ge 0$ for $r \ge s$:

$$M^r\|z_{n,k+1} - z_{n,k}\|^r \le (C_{p^*,s^*}C_1^{(p^*-s^*)(p-1)}M^{s^*})^r\frac{\|b_n^\delta\|^{r[(r-1)(s^*-1)-1]}}{\alpha_n^{(s^*-1)r}}\|b_n^\delta - A_n s_{n,k}\|^r$$
$$\le (C_{p^*,s^*}C_1^{(p^*-s^*)(p-1)}M^{s^*})^r\Bigg[\frac{C_0^{\frac{(r-1)(s^*-1)-1}{s^*-1}}}{\alpha_n}\Bigg]^{r(s^*-1)}\|b_n^\delta - A_n s_{n,k}\|^r$$
$$\le (C_{p^*,s^*}C_1^{(p^*-s^*)(p-1)}M^{s^*}C_2)^r\|b_n^\delta - A_n s_{n,k}\|^r$$
$$= C_4^r\|b_n^\delta - A_n s_{n,k}\|^r.$$

Brought to you by | De Gruyter / TCS
Authenticated
Download Date | 12/9/14 4:09 PM

From (3.15),

$$-\|b_n^\delta - A_n s_{n,k+1}\|^r \le -\Big(\frac{1}{2^{r-1}} - C_4^r\Big)\|b_n^\delta - A_n s_{n,k}\|^r = -C_3\|b_n^\delta - A_n s_{n,k}\|^r \tag{3.16}$$

for all $k \le k_n - 1$. Inserting (3.16) into (3.12) we, in view of (3.7), arrive at

$$\Delta_p(x^+, z_{n,k+1}) - \Delta_p(x^+, z_{n,k}) \le \frac{1}{\alpha_n}\|b_n^\delta - A_n s_{n,k}\|^{r-1}\Big(\Big(\frac{\eta+1}{\tau} + \eta\Big)\|b_n^\delta\| - C_3\|b_n^\delta - A_n s_{n,k}\|\Big). \tag{3.17}$$

Since $-C_3\|b_n^\delta - A_n s_{n,k}\| \le -\mu C_3\|b_n^\delta\|$ for all $k \le k_n - 1$, we have that

$$\Big(\frac{\eta+1}{\tau} + \eta\Big)\|b_n^\delta\| - C_3\|b_n^\delta - A_n s_{n,k}\| \le -C_5\|b_n^\delta\|$$

with $C_5 := \mu C_3 - (\frac{\eta+1}{\tau} + \eta) > 0$. Further,

$$\sum_{k=0}^{l-1}(\Delta_p(x^+, z_{n,k+1}) - \Delta_p(x^+, z_{n,k})) \le -C_5\frac{\|b_n^\delta\|}{\alpha_n}\sum_{k=0}^{l-1}\|b_n^\delta - A_n s_{n,k}\|^{r-1}$$

for all $l \le k_n$ resulting in

$$C_5\frac{\|b_n^\delta\|}{\alpha_n}\sum_{k=0}^{l-1}\|b_n^\delta - A_n s_{n,k}\|^{r-1} \le \Delta_p(x^+, x_n) - \Delta_p(x^+, z_{n,l}) < \infty. \tag{3.18}$$

As $\|b_n^\delta - A_n s_{n,k}\| \ge \mu\|b_n^\delta\|$ for all $k \le k_n - 1$,

$$C_5\,\mu^{r-1}\frac{\|b_n^\delta\|^r}{\alpha_n}\,l < \infty,$$

for all $l \le k_n$, which shows that $l < \infty$ and then $k_n < \infty$. Hence $x_{n+1} = z_{n,k_n}$ is well defined.

(3) Setting $l = k_n$ in (3.18) gives

$$C_5\frac{\|b_n^\delta\|}{\alpha_n}\sum_{k=0}^{k_n-1}\|b_n^\delta - A_n s_{n,k}\|^{r-1} \le \Delta_p(x^+, x_n) - \Delta_p(x^+, x_{n+1}) \tag{3.19}$$

which finally validates (3.9): $\Delta_p(x^+, x_{n+1}) < \Delta_p(x^+, x_n) \le \rho$.

To complete the proof of Theorem 3.4 it remains to demonstrate that Algorithm 1 terminates. Define the set $I := \{n \in \mathbb{N} : \|b_n^\delta\| > \tau\delta_{[n]}\}$ and suppose that I has infinitely many elements. Using again $\|b_n^\delta - A_n s_{n,k}\| \ge \mu\|b_n^\delta\|$ for all $k \le k_n - 1$ and $\alpha_n \le \alpha_{\max}$, it follows from (3.19) that

$$C_5\mu^{r-1}\frac{\|b_n^\delta\|^r}{\alpha_{\max}}k_n \le \Delta_p(x^+, x_n) - \Delta_p(x^+, x_{n+1})$$

for all $n \in I$. But for $n \notin I$, $k_n = 0$ and the above inequality trivially holds. Therefore,

$$\sum_{n=0}^{\infty}\|b_n^\delta\|^r k_n \lesssim \sum_{n=0}^{\infty}(\Delta_p(x^+, x_n) - \Delta_p(x^+, x_{n+1})) \le \Delta_p(x^+, x_0) < \infty.$$

Define now $\delta_{\min} := \min_{0\le j\le d-1}\delta_j > 0$ and observe that $k_n \ge 1$ for all $n \in I$. Hence,

$$\sum_{n\in I}(\tau\delta_{\min})^r \le \sum_{n\in I}\|b_n^\delta\|^r k_n = \sum_{n=0}^{\infty}\|b_n^\delta\|^r k_n < \infty$$

which can hold only if the number of elements in I is finite. Thus, $N(\delta)$ is the largest element in I plus 1. From inequality (3.9),

$$\Delta_p(x^+, x_n) \le \Delta_p(x^+, x_0) < \infty$$

for all $\delta > 0$ and $n \le N(\delta)$. It follows that

$$\|x_n\| \le C$$

for all $n \le N(\delta)$ and some $C > 0$ independent of n, N and δ. □

Brought to you by | De Gruyter / TCS
Authenticated
Download Date | 12/9/14 4:09 PM

Weak convergence of the K-REGINN-IT method is an immediate consequence of (3.10) and the reflexivity of X, see [11, Corollary 3.5].

Corollary 3.5. *Let all the assumptions of Theorem 3.4 hold true. If the operators F_j, $j = 0, \dots, d-1$, are weakly sequentially closed then for any sequence $(y_j^{(\delta_j)_i})_{i\in\mathbb{N}}$ with $\delta^{(i)} = \max\{(\delta_j)_i : j = 0, \dots, d-1\} \to 0$ as $i \to \infty$, the sequence $(x_{N(\delta^{(i)})})_{i\in\mathbb{N}}$ contains a subsequence that converges weakly to a solution of (1.1) in $B_\rho(x^+, \Delta_p)$. If x^+ is the unique solution of (1.1) in $B_\rho(x^+, \Delta_p)$, then $(x_{N(\delta)})_{\delta>0}$ converges weakly to x^+ as $\delta = \max\{\delta_j : j = 0, \dots, d-1\} \to 0$.*

Remark 3.6. The constants C_0, C_1 in Theorem 3.4 depend on the unknown solution x^+. But as $x_0 \in B_\rho(x^+, \Delta_p)$, we conclude that both constants are bounded in p, $\|x_0\|$, and ρ.

Remark 3.7. At a first glance the restriction $s \le r$ in the above theorem might affect the computation of a minimizer of $T_{n,k}^\delta$ (see (3.3)) via (3.8). This, however, is not the case due to (2.1). For instance, if $Y = L^{1.1}$ and $s \ge 2$, we can realize J_r on Y by (2.1) and (2.5).

Remark 3.8. The monotonicity estimate (3.9) actually holds in a more general setting:

$$\Delta_p(\vartheta_n, x_{n+1}) \le \Delta_p(\vartheta_n, x_n) \tag{3.20}$$

whenever ϑ_n is a solution of the $[n]$-th equation $y_{[n]} = F_{[n]}(\vartheta_n)$.

Remark 3.9. Let $k_{\max} = \infty$ and assume that (3.4) is violated by x_n. Following [5] we find

$$\begin{aligned}\|y_{[n]}^{\delta_{[n]}} - F_{[n]}(x_{n+1})\| &\le \|y_{[n]}^{\delta_{[n]}} - F_{[n]}(x_n) - F'_{[n]}(x_n)(x_{n+1} - x_n)\| + \|F_{[n]}(x_{n+1}) - F_{[n]}(x_n) - F'_{[n]}(x_n)(x_{n+1} - x_n)\| \\ &\le \mu\|y_{[n]}^{\delta_{[n]}} - F_{[n]}(x_n)\| + \eta\|F_{[n]}(x_{n+1}) - F_{[n]}(x_n)\|\end{aligned}$$

so that

$$\|y_{[n]}^{\delta_{[n]}} - F_{[n]}(x_{n+1})\| \le \Lambda\|y_{[n]}^{\delta_{[n]}} - F_{[n]}(x_n)\| \quad \text{where } \Lambda := \frac{\mu+\eta}{1-\eta}.$$

Now, if $0 \le \eta < C_3(1+2C_3)^{-1}$ and $\tau > (1+\eta)(C_3(1-2\eta)-\eta)^{-1}$, then $(\frac{1+\eta}{\tau}+\eta)C_3^{-1} < 1-2\eta$ and restricting μ to $((\frac{1+\eta}{\tau}+\eta)C_3^{-1}, 1-2\eta)$ yields $\Lambda < 1$.

4 Convergence in the noise-free setting

From now on, we need to differ clearly between the noisy ($\delta > 0$) and the noise-free ($\delta = 0$) situations. For this reason we exclusively mark quantities by a superscript δ when the data is corrupted by noise: x_n^δ, b_n^δ, A_n^δ etc. Thus, x_n, b_n, A_n etc. originate from exact data. Note that the starting guess is chosen independently of δ:

$$x_0^\delta = x_0.$$

Algorithm 1 is well defined in the noiseless situation when we set $\delta_j = 0$, $\tau = \infty$, and $\tau\delta_j = 0$. Then, the discrepancy principle (3.4) is replaced by $\|b_n\| = 0$, in which case $x_{n+1} = x_n$. Termination only occurs in the unlikely event that an iterate x_N satisfies $\|y_j - F_j(x_N)\| = 0$ for $j = 0, \dots, d-1$, i.e., x_N solves (1.4). In general, Algorithm 1 does not stop but produces a sequence which converges strongly to a solution of (1.1) as we will prove in this section, see Theorem 4.2 below.

Except for the termination statement, all results of Theorem 3.4 hold true with an even larger interval for the selection of the tolerances: $\mu \in (\frac{\eta}{C_3}, 1)$. Accordingly, the constant in (3.19) is replaced by $C_5 := \mu C_3 - \eta > 0$. Further, $N(\delta) = \infty$ in case we have no premature termination.

With the next lemma we prepare our convergence proof for the exact data case.

Lemma 4.1. *Assume all the hypotheses from Theorem 3.4 but with $\mu \in (\frac{\eta}{C_3}, 1)$. Then,*

$$\Delta_p(x_n, x_{n+1}) \lesssim \Delta_p(x^+, x_n) - \Delta_p(x^+, x_{n+1}) \tag{4.1}$$

for all $n \in \mathbb{N}$.

Brought to you by | De Gruyter / TCS
Authenticated
Download Date | 12/9/14 4:09 PM

Proof. From (2.3),

$$\Delta_p(x_n, x_{n+1}) \leq \Delta_p(x^+, x_{n+1}) - \Delta_p(x^+, x_n) + |\langle J_p(x_{n+1}) - J_p(x_n), x^+ - x_n\rangle|. \tag{4.2}$$

But from the definition of the scheme and from properties of j_r,

$$\begin{aligned}
|\langle J_p(x_{n+1}) - J_p(x_n), x^+ - x_n\rangle| &= |\langle J_p(z_{n,k_n}) - J_p(z_{n,0}), x^+ - x_n\rangle| \\
&= \Big|\sum_{k=0}^{k_n-1}\langle J_p(z_{n,k+1}) - J_p(z_{n,k}), x_n - x^+\rangle\Big| \\
&= \frac{1}{\alpha_n}\Big|\sum_{k=0}^{k_n-1}\langle j_r(b_n - A_n s_{n,k+1}), A_n(x_n - x^+)\rangle\Big| \\
&\leq \frac{1}{\alpha_n}\sum_{k=0}^{k_n-1}\|b_n - A_n s_{n,k}\|^{r-1}\|A_n(x_n - x^+)\| \\
&\leq (\eta + 1)\frac{\|b_n\|}{\alpha_n}\sum_{k=0}^{k_n-1}\|b_n - A_n s_{n,k}\|^{r-1} \\
&\leq \frac{\eta + 1}{C_5}(\Delta_p(x^+, x_n) - \Delta_p(x^+, x_{n+1})).
\end{aligned}$$

We have used the TCC (Assumption 3.1 (c)), (3.7) and (3.19). Inserting this result in (4.2), we arrive at (4.1) with constant $\frac{\eta+1}{C_5} - 1 > 0$. □

In the following convergence proof we adapt ideas from our work [13]. The line of argumentation is similar to [13] but we decided to present all details for the reader's convenience.

Theorem 4.2. *Let X and Y be Banach spaces with X satisfying Assumptions 2.1 and 2.3 with $1 < p \leq s \leq r$. Let Assumption 3.1 hold true and start with $x_0 \in B_\rho(x^+, \Delta_p)$. Choose the three constants $\alpha_{\min}$, $\alpha_{\max}$ and C_3 as in Theorem 3.4 and assume that the constant of the TCC in Assumption 3.1* (c) *satisfies $0 \leq \eta < C_3$. Additionally, let $k_{\max} < \infty$ in case $d > 1$. If $\alpha_n \in [\alpha_{\min}, \alpha_{\max}]$ and $\mu \in (\frac{\eta}{C_3}, 1)$, then* K-REGINN-IT *either stops after finitely many iterations with a solution of* (1.1) *or the sequence $(x_n)_{n\in\mathbb{N}} \subset B_\rho(x^+, \Delta_p)$ converges strongly in X to a solution of* (1.1)*. If x^+ is the unique solution in $B_\rho(x^+, \Delta_p)$, then $x_n \to x^+$ as $n \to \infty$.*

Proof. If Algorithm 1 stops after a finite number of iterations, then the current iterate is a solution of (1.1). Otherwise, $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence, as we will prove now.

Let $m, l \in \mathbb{N}$ with $m \leq l$. We consider first the case $d > 1$. Write $m = m_0 d + m_1$ and $l = l_0 d + l_1$ with $m_0, l_0 \in \mathbb{N}$ and $m_1, l_1 \in \{0, \dots, d-1\}$. Of course $m_0 \leq l_0$. Let $z_0 \in \{m_0, \dots, l_0\}$ which will be explicitly determined below, see (4.7). Define $z := z_0 d + z_1$, where $z_1 = l_1$ if $z_0 = l_0$ and $z_1 = d-1$ otherwise. This setting guarantees $m \leq z \leq l$. From Assumption 2.3 (c),

$$\|x_m - x_l\|^s \leq 2^s(\|x_m - x_z\|^s + \|x_z - x_l\|^s) \lesssim \Delta_p(x_z, x_m) + \Delta_p(x_z, x_l).$$

Identity (2.3) implies now that

$$\|x_m - x_l\|^s \lesssim \beta_{m,z} + \beta_{l,z} + f(z, m, l) \tag{4.3}$$

with $\beta_{m,z} := \Delta_p(x^+, x_m) - \Delta_p(x^+, x_z)$ and

$$f(z, m, l) := |\langle J_p(x_z) - J_p(x_m), x_z - x^+\rangle| + |\langle J_p(x_z) - J_p(x_l), x_z - x^+\rangle|.$$

By monotonicity (3.9), we conclude that $\Delta_p(x^+, x_n) \to \gamma \geq 0$ as $n \to \infty$. Thus, $\beta_{m,z}$ and $\beta_{l,z}$ converge to zero as $m \to \infty$ (which causes $z \to \infty$ and $l \to \infty$). Further,

$$f(z, m, l) \leq \sum_{n=m}^{l-1}|\langle J_p(x_{n+1}) - J_p(x_n), x_z - x^+\rangle|. \tag{4.4}$$

As in the proof of Lemma 4.1 we find

$$|\langle J_p(x_{n+1}) - J_p(x_n), x_z - x^+\rangle| \leq \frac{1}{\alpha_n}\sum_{k=0}^{k_n-1}\|b_n - A_n s_{n,k}\|^{r-1}\|A_n(x_z - x^+)\|. \tag{4.5}$$

Brought to you by | De Gruyter / TCS
Authenticated
Download Date | 12/9/14 4:09 PM

From Assumption 3.1 (c),

$$
\begin{aligned}
\|A_n(x_z - x^+)\| &\le \|A_n(x_n - x^+)\| + \|A_n(x_z - x_n)\| \\
&\le \|b_n\| + \|b_n - A_n(x^+ - x_n)\| + \|F_{[n]}(x_z) - F_{[n]}(x_n)\| + \|F_{[n]}(x_z) - F_{[n]}(x_n) - F'_{[n]}(x_n)(x_z - x_n)\| \\
&\le (\eta + 1)(\|b_n\| + \|F_{[n]}(x_z) - F_{[n]}(x_n)\|) \\
&\le (\eta + 1)(2\|b_n\| + \|y_{[n]} - F_{[n]}(x_z)\|). \qquad (4.6)
\end{aligned}
$$

Observe that in the last norm, the operator $F_{[n]}$ is applied in the "wrong" vector x_z. To estimate this norm, we use Assumption 3.1. Write $n = n_0 d + n_1$ with $n_0 \in \{m_0, \dots, l_0\}$ and $n_1 \in \{0, \dots, d-1\}$. Then,

$$
\begin{aligned}
\|y_{[n]} - F_{[n]}(x_z)\| &= \|y_{n_1} - F_{n_1}(x_{z_0 d + z_1})\| \\
&\le \|y_{n_1} - F_{n_1}(x_{z_0 d + n_1})\| + \sum_{j=0}^{d-1} \|F_{n_1}(x_{z_0 d + j + 1}) - F_{n_1}(x_{z_0 d + j})\| \\
&\le \|y_{n_1} - F_{n_1}(x_{z_0 d + n_1})\| + \frac{1}{1-\eta} \sum_{j=0}^{d-1} \|F'_{n_1}(x_{z_0 d + j})(x_{z_0 d + j + 1} - x_{z_0 d + j})\| \\
&\le \Big(1 + \frac{M}{1-\eta}\Big) \sum_{j=0}^{d-1} (\|y_j - F_j(x_{z_0 d + j})\| + \|x_{z_0 d + j + 1} - x_{z_0 d + j}\|).
\end{aligned}
$$

Now we fix $z_0 \in \{m_0, \dots, l_0\}$ such that

$$
\sum_{j=0}^{d-1} (\|y_j - F_j(x_{z_0 d + j})\| + \|x_{z_0 d + j + 1} - x_{z_0 d + j}\|) \le \sum_{j=0}^{d-1} (\|y_j - F_j(x_{\ell d + j})\| + \|x_{\ell d + j + 1} - x_{\ell d + j}\|), \quad \ell \in \{m_0, \dots, l_0\}. \qquad (4.7)
$$

Hence,

$$
\|y_{[n]} - F_{[n]}(x_z)\| \le \Big(1 + \frac{M}{1-\eta}\Big) \sum_{j=0}^{d-1} (\|y_j - F_j(x_{n_0 d + j})\| + \|x_{n_0 d + j + 1} - x_{n_0 d + j}\|). \qquad (4.8)
$$

Inserting (4.8) in (4.6), (4.6) in (4.5), and (4.5) in (4.4), we arrive at

$$
f(z, m, l) \lesssim \sum_{n=m}^{l-1} \frac{\|b_n\|}{\alpha_n} \sum_{k=0}^{k_n - 1} \|b_n - A_n s_{n,k}\|^{r-1} + g(z, m, l) + h(z, m, l) \qquad (4.9)
$$

where

$$
\begin{aligned}
g(z, m, l) &:= \sum_{n=m}^{l-1} \frac{1}{\alpha_n} \sum_{k=0}^{k_n - 1} \|b_n - A_n s_{n,k}\|^{r-1} \sum_{j=0}^{d-1} \|y_j - F_j(x_{n_0 d + j})\|, \\
h(z, m, l) &:= \sum_{n=m}^{l-1} \frac{1}{\alpha_n} \sum_{k=0}^{k_n - 1} \|b_n - A_n s_{n,k}\|^{r-1} \sum_{j=0}^{d-1} \|x_{n_0 d + j + 1} - x_{n_0 d + j}\|.
\end{aligned}
$$

The first term on the right-hand side of (4.9) can be estimated by (3.19). It remains to estimate g and h. As

$$
\sum_{k=0}^{k_n - 1} \|b_n - A_n s_{n,k}\|^{r-1} \le k_{\max} \|b_n\|^{r-1}
$$

and $\|b_n - A_n s_{n,k}\| \ge \mu \|b_n\|$ for all $k \le k_n - 1$, we have

$$
\begin{aligned}
g(z, m, l) &\le \frac{k_{\max}}{\alpha_{\min}} \sum_{n_0 = m_0}^{l_0} \Big(\sum_{n_1 = 0}^{d-1} \|y_{n_1} - F_{n_1}(x_{n_0 d + n_1})\|^{r-1} \sum_{j=0}^{d-1} \|y_j - F_j(x_{n_0 d + j})\| \Big) \\
&\lesssim \sum_{n_0 = m_0}^{l_0} \sum_{n_1 = 0}^{d-1} \|y_{n_1} - F_{n_1}(x_{n_0 d + n_1})\|^r \\
&= \sum_{n = m_0}^{l_0 d + d - 1} \|b_n\|^r \\
&\le \frac{\alpha_{\max}}{\mu^{r-1}} \sum_{n = m_0}^{l_0 d + d - 1} \frac{\|b_n\|}{\alpha_n} \sum_{k=0}^{k_n - 1} \|b_n - A_n s_{n,k}\|^{r-1}. \qquad (4.10)
\end{aligned}
$$

Brought to you by | De Gruyter / TCS
Authenticated
Download Date | 12/9/14 4:09 PM

Similarly, we estimate the last term in (4.9),

$$\begin{aligned} h(z,m,l) &\leq \frac{k_{\max}}{\alpha_{\min}} \sum_{n_0=m_0}^{l_0} \bigg(\sum_{n_1=0}^{d-1} \|y_{n_1} - F_{n_1}(x_{n_0 d+n_1})\|^{r-1} \sum_{j=0}^{d-1} \|x_{n_0 d+j+1} - x_{n_0 d+j}\| \bigg) \\ &\lesssim \sum_{n_0=m_0}^{l_0} \bigg(\sum_{n_1=0}^{d-1} \|y_{n_1} - F_{n_1}(x_{n_0 d+n_1})\|^{r} + \sum_{n_1=0}^{d-1} \|x_{n_0 d+n_1+1} - x_{n_0 d+n_1}\|^{r} \bigg) \\ &= \sum_{n=m_0}^{l_0 d+d-1} \|b_n\|^r + \sum_{n=m_0}^{l_0 d+d-1} \|x_{n+1} - x_n\|^r \\ &\leq \frac{\alpha_{\max}}{\mu^{r-1}} \sum_{n=m_0}^{l_0 d+d-1} \frac{\|b_n\|}{\alpha_n} \sum_{k=0}^{k_n-1} \|b_n - A_n s_{n,k}\|^{r-1} + \sum_{n=m_0}^{l_0 d+d-1} \|x_{n+1} - x_n\|^r. \end{aligned} \tag{4.11}$$

Relying on Assumption 2.3 (c) once again, we obtain

$$\|x_{n+1} - x_n\|^s \lesssim \Delta_p(x_n, x_{n+1}) \overset{(4.1)}{\lesssim} \Delta_p(x^+, x_n) - \Delta_p(x^+, x_{n+1}).$$

As $r \geq s$, we have for m, l large enough that

$$\sum_{n=m_0}^{l_0 d+d-1} \|x_{n+1} - x_n\|^r \leq \sum_{n=m_0}^{l_0 d+d-1} \|x_{n+1} - x_n\|^s \lesssim \Delta_p(x^+, x_{m_0}) - \Delta_p(x^+, x_{l_0 d+d}) = \beta_{m_0, l_0 d+d}.$$

Plugging this bound into (4.11), inserting then inequalities (4.11) and (4.10) in (4.9), (4.9) in (4.3), and using inequality (3.19), we end up with

$$\|x_m - x_l\|^s \lesssim \beta_{m,z} + \beta_{l,z} + \beta_{m_0, l_0 d+d}. \tag{4.12}$$

Now we consider the case $d = 1$ (where $k_{\max} = \infty$ is allowed). This situation is easier because we only need to change the definition x_z in (4.7) to the vector with the smallest residuum in the outer iteration, i.e., choose $z \in \{m, \dots, l\}$ such that $\|b_z\| \leq \|b_n\|$, for all $n \in \{m, \dots, l\}$. Then, from (4.6),

$$\|A_n(x_z - x^+)\| \leq (\eta + 1)(2\|b_n\| + \|b_z\|) \leq 3(\eta + 1)\|b_n\|$$

which, together with (4.5) and (4.4), leads to

$$f(z,m,l) \leq 3(\eta+1) \sum_{n=m}^{l-1} \frac{\|b_n\|}{\alpha_n} \sum_{k=0}^{k_n-1} \|b_n - A_n s_{n,k}\|^{r-1}.$$

Plugging now this result into (4.3) and using again (3.19), we arrive at (4.12) with m_0 and $l_0 d + d$ replaced by m and l, respectively.

In any case the right-hand side of inequality (4.12) converges to zero as $m \to \infty$ revealing $(x_n)_{n\in\mathbb{N}}$ to be a Cauchy sequence. As X is complete, it converges to some $x_\infty \in X$. Observe that $k_n \geq 1$ if $\|b_n\| \neq 0$ and as $\|b_n - A_n s_{n,k}\| \geq \mu\|b_n\|$ for all $k \leq k_n - 1$,

$$\frac{\mu^{r-1}}{\alpha_{\max}} \sum_{n=0}^{\infty} \|b_n\|^r \leq \sum_{n=0}^{\infty} \frac{\|b_n\|}{\alpha_n} k_n(\mu\|b_n\|)^{r-1} \leq \sum_{n=0}^{\infty} \frac{\|b_n\|}{\alpha_n} \sum_{k=0}^{k_n-1} \|b_n - A_n s_{n,k}\|^{r-1} \overset{(3.19)}{<} \infty.$$

Then, $\|y_{[n]} - F_{[n]}(x_n)\| = \|b_n\| \to 0$ as $n \to \infty$ and as the operators F_j are continuous for all $j = 0, \dots, d-1$, we have $y_j = F_j(x_\infty)$. If (1.1) has only one solution in $B_\rho(x^+, \Delta_p)$, then $x_\infty = x^+$. □

5 Regularization property

In this section we validate that K-REGINN-IT is a regularization scheme for solving (1.1) with noisy data y^δ. Indeed, we show that the family $(x^\delta_{N(\delta)})_{\delta>0}$ of outputs of Algorithm 1 relative to the inputs $(y^\delta)_{\delta>0}$ converges strongly to solutions of (1.1) with exact data y.

Brought to you by | De Gruyter / TCS
Authenticated
Download Date | 12/9/14 4:09 PM

To avoid possible wrong interpretations, we will not use the notation δ_j, $j = 0, \dots, d-1$, as in (3.1) any more. Instead, when we write δ_i, we mean a positive number in a sequence of numbers δ as defined in (3.2), i.e., $\delta_i := \max\{(\delta_j)_i : j = 0, \dots, d-1\} > 0$.

We follow ideas from [11] and [13]. In a first step we investigate the stability of the scheme, i.e., we study the behavior of the n-th iterate x_n^δ as δ approaches zero. The sets $\mathcal{X}_n$ defined below play an important role.

Definition 5.1. Let $\mathcal{X}_0 := \{x_0\}$ and define $\mathcal{X}_{n+1}$ from $\mathcal{X}_n$ by the following procedure: for each $\xi \in \mathcal{X}_n$, define $\sigma_{n,0}(\xi) := \xi$ and $\sigma_{n,k+1}(\xi)$ as the minimizer of

$$W_{n,k}(z) := \frac{1}{r}\|\widetilde{b}_n - F'_{[n]}(\xi)(z-\xi)\|^r + \alpha_n \Delta_p(z, \sigma_{n,k}(\xi)), \tag{5.1}$$

where $\widetilde{b}_n := y_{[n]} - F_{[n]}(\xi)$. Define

$$k_{\mathrm{REG}}(\xi) := \min\{k \in \{1, \dots, k_{\max}\} : \|\widetilde{b}_n - F'_{[n]}(\xi)(\sigma_{n,k}(\xi) - \xi)\| < \mu\|\widetilde{b}_n\|\} \tag{5.2}$$

and

$$k_n(\xi) := \begin{cases} 0, & \widetilde{b}_n = 0, \\ k_{\mathrm{REG}}(\xi), & k_{\mathrm{REG}}(\xi) \le k_{\max}, \\ k_{\max}, & k_{\mathrm{REG}}(\xi) > k_{\max}. \end{cases}$$

Then $\sigma_{n,k}(\xi) \in \mathcal{X}_{n+1}$ for $k = 1, \dots, k_n(\xi)$ in case $k_n(\xi) \ge 1$ and only for $k = 0$ in case $k_n(\xi) = 0$. We call $\xi \in \mathcal{X}_n$ the *predecessor* of the vectors $\sigma_{n,k}(\xi) \in \mathcal{X}_{n+1}$ and these ones *successors* of ξ

Of course $x_n \in \mathcal{X}_n$ and $\mathcal{X}_n$ is finite for all $n \in \mathbb{N}$. Moreover, from (3.9) we get

$$\Delta_p(x^+, \xi_{n+1}) \le \Delta_p(x^+, \xi_n)$$

whenever $\xi_{n+1} \in \mathcal{X}_{n+1}$ is a successor of $\xi_n \in \mathcal{X}_n$. We emphasize that the sets $\mathcal{X}_n$, $n \in \mathbb{N}_0$, are defined with respect to exact data y.

The proof of the next lemma basically adapts ideas of [7] and [11].

Lemma 5.2. *Let all the assumptions of Theorem 3.4 hold true. If $\delta_i \to 0$ as $i \to \infty$, then for $n \le N(\delta_i)$ with $\delta_i > 0$ sufficiently small, the sequence $(x_n^{\delta_i})_{i\in\mathbb{N}}$ splits into convergent subsequences, all of which converge to elements of $\mathcal{X}_n$.*

Proof. We prove the statement by induction. For $n = 0$, $x_0^{\delta_i} = x_0 \in \mathcal{X}_0$. Now, suppose that for some $n \in \mathbb{N}$ with $n + 1 \le N(\delta_i)$ for i large enough, $(x_n^{\delta_i})_{i\in\mathbb{N}}$ splits into convergent subsequences, all of which converge to elements of $\mathcal{X}_n$. To simplify the notation, let $(x_n^{\delta_i})_{i\in\mathbb{N}}$ itself be a subsequence which converges to an element of $\mathcal{X}_n$, say,

$$\lim_{i\to\infty} x_n^{\delta_i} = \xi \in \mathcal{X}_n. \tag{5.3}$$

We must prove that the sequence $(x_{n+1}^{\delta_i})_{i\in\mathbb{N}}$ splits into convergent subsequences, each one converging to a point of $\mathcal{X}_{n+1}$. Let us prove beforehand by induction over k that

$$z_{n,k}^{\delta_i} \to \sigma_{n,k}(\xi) \quad \text{as } i \to \infty \text{ for all } k \le k_n(\xi). \tag{5.4}$$

In the remainder of this proof we suppress the dependence of the $\sigma_{n,k}$ on ξ.

For $k = 0$, $z_{n,0}^{\delta_i} = x_n^{\delta_i} \to \xi = \sigma_{n,0}$ as $i \to \infty$. Suppose for some $k \le k_n(\xi) - 1$ that $z_{n,k}^{\delta_i} \to \sigma_{n,k}$ as $i \to \infty$. As the family $(z_{n,k+1}^{\delta_i})_{\delta_i>0}$ is uniformly bounded (see (3.11)) and X is reflexive (Assumption 2.1 (a)), there exists, by picking a subsequence if necessary, some $\bar{z} \in X$ such that $z_{n,k+1}^{\delta_i} \rightharpoonup \bar{z}$ as $i \to \infty$.

Next we show that $\bar{z} = \sigma_{n,k+1}$: For all $g \in Y_{[n]}^*$,

$$\langle g, F'_{[n]}(x_n^{\delta_i})s_{n,k+1}^{\delta_i}\rangle = \langle g, (F'_{[n]}(x_n^{\delta_i}) - F'_{[n]}(\xi))s_{n,k+1}^{\delta_i}\rangle + \langle g, F'_{[n]}(\xi)s_{n,k+1}^{\delta_i}\rangle.$$

But since $s_{n,k+1}^{\delta_i} = z_{n,k+1}^{\delta_i} - x_n^{\delta_i} \rightharpoonup \bar{z} - \xi =: \bar{s}$ as $i \to \infty$ and $F'_{[n]}(\xi)^* g \in X^*$, we have

$$\langle g, F'_{[n]}(\xi)s_{n,k+1}^{\delta_i}\rangle = \langle F'_{[n]}(\xi)^* g, s_{n,k+1}^{\delta_i}\rangle \to \langle F'_{[n]}(\xi)^* g, \bar{s}\rangle = \langle g, F'_{[n]}(\xi)\bar{s}\rangle.$$

Brought to you by | De Gruyter / TCS
Authenticated
Download Date | 12/9/14 4:09 PM

Now, as $F'_{[n]}$ is continuous and $x_n^{\delta_i} \to \xi$,

$$|\langle g, (F'_{[n]}(x_n^{\delta_i}) - F'_{[n]}(\xi)) s_{n,k+1}^{\delta_i} \rangle| \le \|g\|_{Y^*_{[n]}} \|F'_{[n]}(x_n^{\delta_i}) - F'_{[n]}(\xi)\|_{\mathcal{L}(X, Y_{[n]})} \|s_{n,k+1}^{\delta_i}\|_X \to 0 \quad \text{as } i \to \infty$$

because $\|s_{n,k+1}^{\delta_i}\| \le \|z_{n,k+1}^{\delta_i}\| + \|x_n^{\delta_i}\|$ is uniformly bounded. Then,

$$\langle g, F'_{[n]}(x_n^{\delta_i}) s_{n,k+1}^{\delta_i} \rangle \to \langle g, F'_{[n]}(\xi) \bar{s} \rangle \tag{5.5}$$

and as $g \in Y^*_{[n]}$ is arbitrary,

$$F'_{[n]}(x_n^{\delta_i}) s_{n,k+1}^{\delta_i} \rightharpoonup F'_{[n]}(\xi) \bar{s}.$$

From (3.1) we conclude that

$$b_n^{\delta_i} - F'_{[n]}(x_n^{\delta_i}) s_{n,k+1}^{\delta_i} \rightharpoonup \tilde{b}_n - F'_{[n]}(\xi) \bar{s}$$

and then

$$\|\tilde{b}_n - F'_{[n]}(\xi) \bar{s}\| \le \liminf \|b_n^{\delta_i} - F'_{[n]}(x_n^{\delta_i}) s_{n,k+1}^{\delta_i}\|. \tag{5.6}$$

Now, Assumption 2.1 (b) guarantees that J_p is continuous and similarly to (5.5) we get

$$\langle J_p(z_{n,k}^{\delta_i}), z_{n,k+1}^{\delta_i} \rangle = \langle J_p(z_{n,k}^{\delta_i}) - J_p(\sigma_{n,k}), z_{n,k+1}^{\delta_i} \rangle + \langle J_p(\sigma_{n,k}), z_{n,k+1}^{\delta_i} \rangle \to \langle J_p(\sigma_{n,k}), \bar{z} \rangle$$

which in turn implies

$$\begin{aligned} \Delta_p(\bar{z}, \sigma_{n,k}) &= \frac{1}{p} \|\bar{z}\|^p + \frac{1}{p^*} \|\sigma_{n,k}\|^p - \langle J_p(\sigma_{n,k}), \bar{z} \rangle \\ &\le \liminf \Big(\frac{1}{p} \|z_{n,k+1}^{\delta_i}\|^p + \frac{1}{p^*} \|z_{n,k}^{\delta_i}\|^p - \langle J_p(z_{n,k}^{\delta_i}), z_{n,k+1}^{\delta_i} \rangle \Big) \\ &= \liminf \Delta_p(z_{n,k+1}^{\delta_i}, z_{n,k}^{\delta_i}). \end{aligned} \tag{5.7}$$

From (5.3), (5.6), (5.7) and due to the minimality property of $z_{n,k+1}^{\delta_i}$,

$$W_{n,k}(\bar{z}) \le \liminf T_{n,k}^{\delta_i}(z_{n,k+1}^{\delta_i}) \le \liminf T_{n,k}^{\delta_i}(\sigma_{n,k+1}) = \lim_{i \to \infty} T_{n,k}^{\delta_i}(\sigma_{n,k+1}) = W_{n,k}(\sigma_{n,k+1}),$$

where $W_{n,k}$ and $T_{n,k}^{\delta_i}$ are defined in (5.1) and (3.3), respectively. Using minimality and uniqueness of $\sigma_{n,k+1}$, we finally conclude $\sigma_{n,k+1} = \bar{z}$. Then $z_{n,k+1}^{\delta_i} \rightharpoonup \sigma_{n,k+1}$ and $s_{n,k+1}^{\delta_i} \rightharpoonup \sigma_{n,k+1} - \xi$ which implies that $\bar{s} = \sigma_{n,k+1} - \xi$.

Below we will establish that

$$\Delta_p(z_{n,k+1}^{\delta_i}, z_{n,k}^{\delta_i}) \to \Delta_p(\sigma_{n,k+1}, \sigma_{n,k}) \quad \text{as } i \to \infty \tag{5.8}$$

yielding $\|z_{n,k+1}^{\delta_i}\| \to \|\sigma_{n,k+1}\|$ in view of the definition of a Bregman distance. As $z_{n,k+1}^{\delta_i} \rightharpoonup \sigma_{n,k+1}$, we conclude that $z_{n,k+1}^{\delta_i} \to \sigma_{n,k+1}$ as $i \to \infty$, see Assumption 2.3 (b). So far, we have shown that each positive zero sequence $(\delta_i)_{i \in \mathbb{N}}$ contains a subsequence $(\delta_{i_j})_{j \in \mathbb{N}}$ such that

$$z_{n,k+1}^{\delta_{i_j}} \to \sigma_{n,k+1} \quad \text{as } j \to \infty$$

which proves (5.4). Consequently,

$$s_{n,k}^{\delta_i} \to \sigma_{n,k} - \xi$$

and

$$b_n^{\delta_i} - F'_{[n]}(x_n^{\delta_i}) s_{n,k}^{\delta_i} \to \tilde{b}_n - F'_{[n]}(\xi)(\sigma_{n,k} - \xi) \tag{5.9}$$

as $i \to \infty$ for all $k \le k_n(\xi)$. Now we have to differ three cases.

Case 1: $1 \le k_n(\xi) = k_{\mathrm{REG}}(\xi)$. From definition (5.2),

$$\|\tilde{b}_n - F'_{[n]}(\xi)(\sigma_{n,k_n(\xi)} - \xi)\| < \mu \|\tilde{b}_n\|.$$

It follows from (5.9) that for i large enough

$$\|b_n^{\delta_i} - F'_{[n]}(x_n^{\delta_i}) s_{n,k_n(\xi)}^{\delta_i}\| < \mu \|b_n^{\delta_i}\|$$

which in view of (3.5) implies $k_n^{\delta_i} \le k_n(\xi)$. Then $k_n^{\delta_i} \in \{0, \dots, k_n(\xi)\}$ and we conclude using (5.4) that $x_{n+1}^{\delta_i} = z_{n,k_n^{\delta_i}}^{\delta_i}$ splits into at most $k_n(\xi) + 1$ convergent subsequences, each one converging to an element of $\mathcal{X}_{n+1}$.

Brought to you by | De Gruyter / TCS
Authenticated
Download Date | 12/9/14 4:09 PM

Case 2: $k_n(\xi) = k_{\max}$. In this case,

$$k_n^{\delta_i} \le k_{\max} = k_n(\xi)$$

and we proceed as in Case 1.

Case 3: $k_n(\xi) = 0$. Then $\widetilde{b}_n = 0$, that is, $y_{[n]} = F_{[n]}(\xi)$ and $\xi \in \mathcal{X}_{n+1}$. We will prove that $x_{n+1}^{\delta_i} \to \xi$ as $i \to \infty$. Assume the contrary, then there exist an $\epsilon > 0$ and a subsequence $(\delta_{i_m})_{m\in\mathbb{N}}$ such that

$$\epsilon < \|\xi - x_{n+1}^{\delta_{i_m}}\|^s$$

and using Assumption 2.3 (c),

$$\epsilon \le \frac{1}{C}\Delta_p(\xi, x_{n+1}^{\delta_{i_m}}) \overset{(3.20)}{\le} \frac{1}{C}\Delta_p(\xi, x_n^{\delta_{i_m}}) \xrightarrow{i\to\infty} \frac{1}{C}\Delta_p(\xi,\xi) = 0,$$

contradicting $\epsilon > 0$.

It remains to verify (5.8): Define

$$a_i := \Delta_p(z_{n,k+1}^{\delta_i}, z_{n,k}^{\delta_i}), \quad a := \limsup a_i, \quad c := \Delta_p(\sigma_{n,k+1}, \sigma_{n,k})$$

and

$$re_i := \frac{1}{r}\|b_n^{\delta_i} - F'_{[n]}(x_n^{\delta_i})s_{n,k+1}^{\delta_i}\|^r, \quad re := \liminf re_i.$$

In view of (5.7), it is enough to prove that $a \le c$. Suppose that $a > c$. From the definition of $\limsup$ there exists, for all $M \in \mathbb{N}$, some index $i > M$ such that

$$a_i > a - \frac{a-c}{4}. \tag{5.10}$$

From the definition of $\liminf$, there exists an $N_1 \in \mathbb{N}$ such that

$$re_i \ge re - \frac{\alpha_n(a-c)}{4} \tag{5.11}$$

for all $i \ge N_1$. As above, $\lim_{i\to\infty} T_{n,k}^{\delta_i}(\sigma_{n,k+1}) = W_{n,k}(\sigma_{n,k+1})$ and then there is an $N_2 \in \mathbb{N}$ such that

$$T_{n,k}^{\delta_i}(\sigma_{n,k+1}) < W_{n,k}(\sigma_{n,k+1}) + \frac{\alpha_n(a-c)}{2} \tag{5.12}$$

for all $i \ge N_2$. Using (5.6) and setting $M = N_1 \vee N_2$, there exists some index $i > M$ such that

$$\begin{aligned} W_{n,k}(\sigma_{n,k+1}) &\le re + \alpha_n c = re + \alpha_n a - \alpha_n(a-c) \\ &\le re_i + \frac{\alpha_n(a-c)}{4} + \alpha_n a_i + \frac{\alpha_n(a-c)}{4} - \alpha_n(a-c) \\ &= re_i + \alpha_n a_i - \frac{\alpha_n(a-c)}{2} \\ &= T_{n,k}^{\delta_i}(z_{n,k+1}^{\delta_i}) - \frac{\alpha_n(a-c)}{2} \\ &\le T_{n,k}^{\delta_i}(\sigma_{n,k+1}) - \frac{\alpha_n(a-c)}{2} \end{aligned}$$

where the second inequality comes from (5.11) and (5.10) and the last one, from the minimality of $z_{n,k+1}^{\delta_i}$. From (5.12) we obtain the contradiction $W_{n,k}(\sigma_{n,k+1}) < W_{n,k}(\sigma_{n,k+1})$. Thus, $a \le c$ and (5.8) holds. □

In the second step towards establishing the regularization property we provide a kind of uniform convergence of the set sequence $(\mathcal{X}_n)_n$ to solutions of (1.1). For the rigorous formulation in Lemma 5.3 below we need to introduce further notation: Let $l \in \mathbb{N}$ and set $\xi_0^{(l)} := x_0$. Now define

$$\xi_{n+1}^{(l)} := \sigma_{n,k_n^{(l)}}(\xi_n^{(l)})$$

by choosing $k_n^{(l)} \in \{1,\dots,k_n(\xi_n^{(l)})\}$ in case of $k_n(\xi_n^{(l)}) \ge 1$ and $k_n^{(l)} = 0$ otherwise. Then $\xi_{n+1}^{(l)}$ is a successor of $\xi_n^{(l)}$. Of course $\xi_n^{(l)} \in \mathcal{X}_n$ for all $n \in \mathbb{N}$ and reciprocally, each element in $\mathcal{X}_n$ can be written as $\xi_n^{(l)}$ for some $l \in \mathbb{N}$.

Brought to you by | De Gruyter / TCS
Authenticated
Download Date | 12/9/14 4:09 PM

Observe that $(\xi_n^{(l)})_{n\in\mathbb{N}}$ represents a sequence generated by K-REGINN-IT with the inner iteration stopped with an arbitrary stop index $k_n^{(l)}$ less than or equal $k_n(\xi_n^{(l)})$. Due to this fact, we call the sequence $(k_n^{(l)})_{n\in\mathbb{N}}$ a *stop rule*. Then each element of $(\xi_n^{(l)})_{n\in\mathbb{N}}$ satisfies

$$\|y_{[n]} - F_{[n]}(\xi_n^{(l)}) - F'_{[n]}(\xi_n^{(l)})(\sigma_{n,k}(\xi_n^{(l)}) - \xi_n^{(l)})\| \geq \mu\|y_{[n]} - F_{[n]}(\xi_n^{(l)})\|, \quad k \leq k_n^{(l)} - 1,$$

see (5.2). Hence, Theorem 4.2 applies to $(\xi_n^{(l)})_{n\in\mathbb{N}}$, that is, the limit

$$x_\infty^{(l)} := \lim_{n\to\infty} \xi_n^{(l)} \tag{5.13}$$

exists and is a solution of (1.1) in $B_\rho(x^+, \Delta_p)$.

The following result is adapted from [6] and generalizes [13, Proposition 19].

Lemma 5.3. *Let all assumptions of Theorem 4.2 hold true and let $(\xi_n^{(l)})_{n\in\mathbb{N}}$ denote the sequence generated by the stop rule $(k_n^{(l)})_{n\in\mathbb{N}}$. Then, for each $\epsilon > 0$ there exists an $M = M(\epsilon) \in \mathbb{N}$ such that*

$$\|\xi_n^{(l)} - x_\infty^{(l)}\| < \epsilon \quad \text{for all } n \geq M \text{ and all } l \in \mathbb{N}.$$

In particular, if x^+ is the unique solution of (1.1) in $B_\rho(x^+, \Delta_p)$, then $\|\xi_n^{(l)} - x^+\| < \epsilon$ for all $n \geq M$ and all $l \in \mathbb{N}$.

Proof. Assume the statement is not true. Then, there exist an $\epsilon > 0$ and sequences $(n_j)_j, (l_j)_j \subset \mathbb{N}$ with $(n_j)_j$ strictly increasing such that

$$\|\xi_{n_j}^{(l_j)} - x_\infty^{(l_j)}\| > \epsilon \quad \text{for all } j \in \mathbb{N}$$

where $(\xi_n^{(l_j)})_n$ represents the sequence generated by the stop rule $(k_n^{(l_j)})_n$. We stress the fact that the iterates $\xi_{n_j}^{(l_j)}$ must be generated by infinitely many different sequences of stop rules (otherwise, as $\xi_{n_j}^{(l)} \to x_\infty^{(l)}$ as $j \to \infty$ for each l and as the l_j attain only a finite number of values, we would have $\|\xi_{n_j}^{(l_j)} - x_\infty^{(l_j)}\| < \epsilon$ for n_j large enough). Next we reorder the numbers l_j (excluding some iterates if necessary) such that

$$\|\xi_{n_l}^{(l)} - x_\infty^{(l)}\| > \epsilon \quad \text{for all } l \in \mathbb{N}. \tag{5.14}$$

Now we construct inductively an auxiliary sequence $(\hat{\xi}_n)_n$, which is generated by a stop rule $(\hat{k}_n)_n$, as well as a sequence of unbounded sets $(\mathcal{L}_n)_n$ such that $\mathcal{L}_n \subset \mathbb{N}\backslash\{1,\ldots,n\}$ with $\mathcal{L}_{n+1} \subset \mathcal{L}_n$, $n \in \mathbb{N}_0$, and

$$\hat{k}_n = k_n^{(l)} \quad \text{for all } l \in \mathcal{L}_n,\ n \in \mathbb{N}_0. \tag{5.15}$$

Set $\hat{\xi}_0 := x_0$. As $k_0(\hat{\xi}_0) < \infty$, there exists some $\hat{k}_0 \in \{0,\ldots,k_0(\hat{\xi}_0)\}$ such that $\hat{k}_0 = k_0^{(l)}$ for infinitely many $l \in \mathbb{N}$. Let $\mathcal{L}_0 \subset \mathbb{N}$ be the set of those indices l. Assume that $\hat{\xi}_0,\ldots,\hat{\xi}_n, \hat{k}_0,\ldots,\hat{k}_n$, and $\mathcal{L}_0,\ldots,\mathcal{L}_n$ have been constructed with the requested properties. Then, define

$$\hat{\xi}_{n+1} := \sigma_{n,\hat{k}_n}(\hat{\xi}_n),$$

see Definition 5.1. As $k_{n+1}(\hat{\xi}_{n+1}) < \infty$, we find some $\hat{k}_{n+1} \in \{0,\ldots,k_{n+1}(\hat{\xi}_{n+1})\}$ such that $\hat{k}_{n+1} = k_{n+1}^{(l)}$ for infinitely many $l \in \mathcal{L}_n\backslash\{1,\ldots,n+1\}$. Those indices l are collected in $\mathcal{L}_{n+1} \subset \mathcal{L}_n$ and the inductive construction of $(\hat{\xi}_n)_n$ is complete.

In view of (5.13) the limit $\hat{x}_\infty := \lim_{n\to\infty} \hat{\xi}_n$ exists in $B_\rho(x^+, \Delta_p)$ and solves (1.1). Observe that, if $l \in \mathcal{L}_0$, then

$$\xi_1^{(l)} = \sigma_{0,k_0^{(l)}}(\xi_0^{(l)}) \overset{(5.15)}{=} \sigma_{0,\hat{k}_0}(\xi_0^{(l)}) = \sigma_{0,\hat{k}_0}(\hat{\xi}_0) = \hat{\xi}_1.$$

By induction,

$$l \in \mathcal{L}_n \implies \xi_{n+1}^{(l)} = \hat{\xi}_{n+1} \quad \text{for all } n \in \mathbb{N}_0. \tag{5.16}$$

Since the "diagonal" sequence $\hat{\xi}_n$ converges to $\hat{x}_\infty$ as $n \to \infty$, there exists an $M = M(\epsilon) \in \mathbb{N}$ such that

$$\Delta_p(\hat{x}_\infty, \hat{\xi}_n) < \frac{C\epsilon^s}{2^{s+1}} \quad \text{for all } n > M, \tag{5.17}$$

Brought to you by | De Gruyter / TCS
Authenticated
Download Date | 12/9/14 4:09 PM

where $C > 0$ is the constant in Assumption 2.3 (c). We can additionally suppose that $\widehat{\xi}_n \in B_\rho(x^+, \Delta_p)$ for all $n > M$. In fact, as $\lim_{n\to\infty} \widehat{\xi}_n = \widehat{x}_\infty$ and the mappings J_p and $\Delta_p(\widehat{x}_\infty, \cdot)$ are continuous, we have that $\Delta_p(\widehat{x}_\infty, \widehat{\xi}_n)$ and $\langle J_p(\widehat{\xi}_n) - J_p(\widehat{x}_\infty), \widehat{x}_\infty - x^+\rangle$ converge to zero as $n \to \infty$. From (2.3),

$$\Delta_p(x^+, \widehat{\xi}_n) = \Delta_p(\widehat{x}_\infty, \widehat{\xi}_n) + \Delta_p(x^+, \widehat{x}_\infty) + \langle J_p(\widehat{\xi}_n) - J_p(\widehat{x}_\infty), \widehat{x}_\infty - x^+\rangle$$

and as $\Delta_p(x^+, \widehat{x}_\infty) < \rho$, we conclude that $\Delta_p(x^+, \widehat{\xi}_n) < \rho$ for n large enough.

Now, for $l_0 \in \pounds_M$ fixed,

$$\Delta_p(\widehat{x}_\infty, \xi_{M+1}^{(l_0)}) \overset{(5.16)}{=} \Delta_p(\widehat{x}_\infty, \widehat{\xi}_{M+1}) \overset{(5.17)}{<} \frac{C\epsilon^s}{2^{s+1}}.$$

As $\widehat{x}_\infty$ is a solution of (1.1) and $\xi_{M+1}^{(l_0)} = \widehat{\xi}_{M+1} \in B_\rho(x^+, \Delta_p)$, inequality (3.9) applies and the errors $\Delta_p(\widehat{x}_\infty, \xi_n^{(l_0)})$ are monotonically decreasing in n for all $n \geq M+1$. In particular, $n_{l_0} \geq l_0 \geq M+1$ (since $l_0 \in \pounds_M \subset \mathbb{N}\backslash\{1, \ldots, M\}$). Then

$$\Delta_p(\widehat{x}_\infty, \xi_{n_{l_0}}^{(l_0)}) \leq \Delta_p(\widehat{x}_\infty, \xi_{M+1}^{(l_0)}) < \frac{C\epsilon^s}{2^{s+1}}.$$

Since $\xi_n^{(l_0)} \to x_\infty^{(l_0)}$ as $n \to \infty$, we conclude that

$$\Delta_p(\widehat{x}_\infty, x_\infty^{(l_0)}) = \lim_{n\to\infty} \Delta_p(\widehat{x}_\infty, \xi_n^{(l_0)}) \leq \frac{C\epsilon^s}{2^{s+1}}.$$

From the inequality in Assumption 2.3 (c),

$$\|\xi_{n_{l_0}}^{(l_0)} - x_\infty^{(l_0)}\|^s \leq 2^s(\|\xi_{n_{l_0}}^{(l_0)} - \widehat{x}_\infty\|^s + \|\widehat{x}_\infty - x_\infty^{(l_0)}\|^s) \leq \frac{2^s}{C}(\Delta_p(\widehat{x}_\infty, \xi_{n_{l_0}}^{(l_0)}) + \Delta_p(\widehat{x}_\infty, x_\infty^{(l_0)})) < \epsilon^s,$$

contradicting (5.14). □

We are now well prepared to prove our main result.

Theorem 5.4 (Regularization property). *Let all assumptions of Theorem* 3.4 *hold true but with* $k_{\max} < \infty$ *for* $d > 1$. *If* $\delta_i \to 0$ *as* $i \to \infty$, *then the sequence* $(x_{N(\delta_i)}^{\delta_i})_{i\in\mathbb{N}}$ *splits into convergent subsequences, all of which converge strongly to solutions of* (1.1) *as* $i \to \infty$. *If* x^+ *is the unique solution of* (1.1) *in* $B_\rho(x^+, \Delta_p)$, *then*

$$\lim_{i\to\infty} \|x_{N(\delta_i)}^{\delta_i} - x^+\| = 0.$$

Proof. If $N(\delta_i) \leq I$ for some $I \in \mathbb{N}$ as $i \to \infty$, then $(x_{N(\delta_i)}^{\delta_i})_{i\in\mathbb{N}}$ splits into subsequences of the form $(x_n^{\delta_{i_j}})_{j\in\mathbb{N}}$ where n is an iteration index less than or equal to I. According to Lemma 5.2, each of these subsequences splits into convergent subsequences. Hence each limit of such a subsequence must be a solution of (1.1) due to the discrepancy principle (3.6). In fact, if $x_n^{\delta_i} \to a$ as $i \to \infty$, then using (3.1),

$$\|y_j - F_j(a)\| = \lim_{i\to\infty} \|y_j - F_j(x_n^{\delta_i})\| \leq \lim_{i\to\infty} (\tau + 1)\delta_i = 0, \quad j = 0, \ldots, d-1.$$

Suppose now that $N(\delta_i) \to \infty$ as $i \to \infty$ and let $\epsilon > 0$ be given. As the Bregman distance is a continuous function in both arguments, there exists some $\gamma = \gamma(\epsilon) > 0$ such that

$$\Delta_p(x, x_n^{\delta_i}) < C\epsilon^s \quad \text{whenever } \|x - x_n^{\delta_i}\| \leq \gamma \tag{5.18}$$

where $C > 0$ is the constant appearing in Assumption 2.3 (c). From Lemma 5.3, there is an $M \in \mathbb{N}$ such that, for each $\xi_M^{(l)} \in \mathcal{X}_M$, there exists a solution $x_\infty^{(l)}$ of (1.1) satisfying

$$\|x_\infty^{(l)} - \xi_M^{(l)}\| < \frac{\gamma}{2}.$$

According to Lemma 5.2, $(x_M^{\delta_i})_{i\in\mathbb{N}}$ splits into convergent subsequences, each one converging to an element of $\mathcal{X}_M$.

Let $(x_M^{\delta_{i_j}})_{j\in\mathbb{N}}$ be a generic convergent subsequence, which converges to an element of $\mathcal{X}_M$, say

$$\lim_{j\to\infty} x_M^{\delta_{i_j}} = \xi_M^{(l_0)} \in \mathcal{X}_M.$$

We will prove that the subsequence $(x_{N(\delta_{i_j})}^{\delta_{i_j}})_{j\in\mathbb{N}}$ converges to the solution $x_\infty^{(l_0)}$. In fact, since

$$x_M^{\delta_{i_j}} \to \xi_M^{(l_0)} \quad \text{as } j \to \infty,$$

Brought to you by | De Gruyter / TCS
Authenticated
Download Date | 12/9/14 4:09 PM

there exists a $J_1 = J_1(\epsilon)$ such that

$$\|\xi_M^{(l_0)} - x_M^{\delta_{i_j}}\| < \frac{\gamma}{2} \quad \text{for all } j \geq J_1.$$

Since $N(\delta_{i_j}) \to \infty$ as $j \to \infty$, we have $N(\delta_{i_j}) \geq M$ for all $j \geq J$ where $J \geq J_1$ is a sufficiently large number. Then, for all $j \geq J$,

$$\|x_\infty^{(l_0)} - x_M^{\delta_{i_j}}\| \leq \|x_\infty^{(l_0)} - \xi_M^{(l_0)}\| + \|\xi_M^{(l_0)} - x_M^{\delta_{i_j}}\| \leq \gamma.$$

Finally, (5.18) and Assumption 2.3 (c) lead to

$$\|x_{N(\delta_{i_j})}^{\delta_{i_j}} - x_\infty^{(l_0)}\|^s \leq \frac{1}{C}\Delta_p(x_\infty^{(l_0)}, x_{N(\delta_{i_j})}^{\delta_{i_j}}) \overset{(3.9)}{\leq} \frac{1}{C}\Delta_p(x_\infty^{(l_0)}, x_M^{\delta_{i_j}}) \leq \epsilon^s. \qquad \square$$

We like to emphasize that the regularization property of K-REGINN-IT holds without any additional assumption on Y other than it is a general Banach space. However, convexity and smoothness properties of X and Y affect the convergence properties of the scheme, see Remark 3.2.

A Convergence of the fixed point iteration (3.8)

Let n and k be fixed. Let $v_0 \in X$ and consider

$$J_p(v_{m+1}) = J_p(z_{n,k}) + \frac{1}{\alpha_n} A_n^* j_r(b_n^\delta - A_n(v_m - x_n)), \quad m = 0, 1, 2, \dots, \tag{A.1}$$

which is a fixed point iteration to solve (3.8) as $J_p^{-1} = J_{p^*}^*$.

Proposition A.1. *Let all assumptions of Theorem 3.4 hold true and assume that $1 < p \leq 2 \leq r$, X is 2-convex and Y is 2-smooth. Additionally, let $\alpha_{\min} = \alpha_{\min}(p, r)$ be large enough and $v_0 \in X$ be chosen such that $\|v_0\| \leq K_0$ with $K_0 > 0$ independent of n and k. Then the fixed point iteration* (A.1) *converges to $z_{n,k+1}$, the unique minimizer of* (3.3).

Proof. As all estimates in this proof are uniform in the noise level, we suppress the superscript δ. We first prove that the sequence $(v_m)_{m\in\mathbb{N}}$ is uniformly bounded in n, k, and m. Under the assumptions of Theorem 3.4 the sequences $(z_{n,k})_{n,k}$, $(x_n)_n$ and $(b_n)_n$ are uniformly bounded, see (3.10) and (3.11). Then we can find constants $K_1, K_2 > 0$ independent of n, k, and m such that $\|z_{n,k}\|^{p-1} \leq K_1$ and $2^{r-2}M\|b_n + A_n x_n\|^{r-1} \leq K_2$. We prove by induction that $\|v_m\|^{p-1} \leq 4K_1 + K_0^{p-1}$ for all $m \in \mathbb{N}_0$ which is certainly true for $m = 0$. Assume the assertion holds for $v_0, \dots, v_m$ and define $K_3 := 2^{r-2}M^r(4K_1 + K_0^{p-1})^{(r-p)/(p-1)}$. By (A.1),

$$\begin{aligned}
\|v_{m+1}\|^{p-1} &\leq \|z_{n,k}\|^{p-1} + \frac{M}{\alpha_n}\|b_n + A_n x_n - A_n v_m\|^{r-1} \\
&\leq K_1 + \frac{2^{r-2}M}{\alpha_{\min}}[\|b_n + A_n x_n\|^{r-1} + (M\|v_m\|)^{r-1}] \\
&\leq K_1 + \frac{K_2}{\alpha_{\min}} + \frac{K_3}{\alpha_{\min}}\|v_m\|^{p-1}.
\end{aligned}$$

We have used the induction hypothesis $\|v_m\| \leq (4K_1 + K_0^{p-1})^{1/(p-1)}$ and $p \leq r$ in the last inequality. Now, for $\alpha_{\min} \geq \frac{K_2}{K_1} \vee 2K_3$,

$$\begin{aligned}
\|v_{m+1}\|^{p-1} &\leq 2K_1 + \frac{1}{2}\|v_m\|^{p-1} \\
&\leq 2K_1 + \frac{1}{2}\Big(2K_1 + \frac{1}{2}\|v_{m-1}\|^{p-1}\Big) \\
&\leq \dots \leq 2K_1 \sum_{j=0}^{m} \frac{1}{2^j} + \frac{1}{2^{m+1}}\|v_0\|^{p-1} \leq 4K_1 + K_0^{p-1},
\end{aligned}$$

which completes the proof by induction. Thus, there exists a positive constant independent of n, k, and m satisfying $\|b_n - A_n(v_m - x_n)\|^{r-2} \leq K_4$.

Brought to you by | De Gruyter / TCS
Authenticated
Download Date | 12/9/14 4:09 PM

We define $K_5 := C_{p^*,2}(4K_1 + K_0^{p-1})^{p^*-2} M^2 C_{r,2} K_4$ and show that the sequence $(v_m)_{m\in\mathbb{N}}$ converges to $z_{n,k+1}$. As X is 2-convex, X^* is 2-smooth, which implies that

$$\begin{aligned}
\|v_{m+1} - v_m\| &= \|J_{p^*}(J_p(v_{m+1})) - J_{p^*}(J_p(v_m))\| \\
&\le C_{p^*,2}(\|J_p(v_{m+1})\| \vee \|J_p(v_m)\|)^{p^*-2} \|J_p(v_{m+1}) - J_p(v_m)\| \\
&\le \frac{C_{p^*,2}(4K_1 + K_0^{p-1})^{p^*-2} M}{\alpha_{\min}} \|j_r(b_n - A_n(v_m - x_n)) - j_r(b_n - A_n(v_{m-1} - x_n))\| \\
&\le \frac{C_{p^*,2}(4K_1 + K_0^{p-1})^{p^*-2} M}{\alpha_{\min}} C_{r,2}(\|b_n - A_n(v_m - x_n)\| \vee \|b_n - A_n(v_{m-1} - x_n)\|)^{r-2} \\
&\qquad \times \|b_n - A_n(v_m - x_n) - (b_n - A_n(v_{m-1} - x_n))\| \\
&\le L\|v_m - v_{m-1}\|
\end{aligned}$$

with $L := K_5/\alpha_{\min} < 1$ for $\alpha_{\min} > K_5$. Hence, $(v_m)_{m\in\mathbb{N}}$ converges as it is a Cauchy sequence. Further, J_p and j_r are continuous, thus, its limit is the unique fixed point of (A.1), i.e., $v_m \to z_{n,k+1}$ as $m \to \infty$. □

Funding: The first author acknowledges support by the German Academic Exchange Service (DAAD).

References

[1] T. Bonesky, K. S. Kazimierski, P. Maass, F. Schöpfer and T. Schuster, Minimization of Tikhonov functionals in Banach spaces, *Abstr. Appl. Anal.* (2008), Article ID 192679.

[2] M. Burger and B. Kaltenbacher, Regularizing Newton-Kaczmarz methods for nonlinear ill-posed problems, *SIAM J. Numer. Anal.* **44** (2006), no. 1, 153–182 (electronic).

[3] I. Cioranescu, *Geometry of Banach Spaces, Duality Mappings and Nonlinear Problems*, *Math. Appl.* 62, Kluwer Academic Publishers, Dordrecht, 1990.

[4] M. Haltmeier, A. Leitão and O. Scherzer, Kaczmarz methods for regularizing nonlinear ill-posed equations. I. Convergence analysis, *Inverse Probl. Imaging* **1** (2007), no. 2, 289–298.

[5] M. Hanke, A regularizing Levenberg–Marquardt scheme, with applications to inverse groundwater filtration problems, *Inverse Problems* **13** (1997), no. 1, 79–95.

[6] M. Hanke, Regularizing properties of a truncated Newton-CG algorithm for nonlinear inverse problems, *Numer. Funct. Anal. Optim.* **18** (1997), no. 9–10, 971–993.

[7] Q. Jin and L. Stals, Nonstationary iterated Tikhonov regularization for ill-posed problems in Banach spaces, *Inverse Problems* **28** (2012), Article ID 104011.

[8] Q. Jin and W. Wang, Landweber iteration of Kaczmarz type with general non-smooth convex penalty functionals, *Inverse Problems* **29** (2013), no. 8, Article ID 085011.

[9] S. Kaczmarz, Approximate solution of systems of linear equations, *Internat. J. Control* **57** (1993), no. 6, 1269–1271.

[10] R. Kowar and O. Scherzer, Convergence analysis of a Landweber–Kaczmarz method for solving nonlinear ill-posed problems, in: *Ill-Posed and Inverse Problems*, VSP, Zeist (2002), 253–270.

[11] A. Lechleiter and A. Rieder, Towards a general convergence theory for inexact Newton regularizations, *Numer. Math.* **114** (2010), no. 3, 521–548.

[12] A. Leitão and M. Marques Alves, On Landweber-Kaczmarz methods for regularizing systems of ill-posed equations in Banach spaces, *Inverse Problems* **28** (2012), no. 10, Article ID 104008.

[13] F. Margotti, A. Rieder and A. Leitão, A Kaczmarz version of the REGINN-Landweber iteration for ill-posed problems in Banach spaces, *SIAM J. Numer. Anal.* **52** (2014), no. 3, 1439–1465.

[14] A. Rieder, On the regularization of nonlinear ill-posed problems via inexact Newton iterations, *Inverse Problems* **15** (1999), no. 1, 309–327.

[15] F. Schöpfer, A. K. Louis and T. Schuster, Nonlinear iterative methods for linear ill-posed problems in Banach spaces, *Inverse Problems* **22** (2006), no. 1, 311–329.

[16] T. Schuster, B. Kaltenbacher, B. Hofmann and K. S. Kazimierski, *Regularization Methods in Banach Spaces*, Radon Ser. Comput. Appl. Math. 10, Walter de Gruyter, Berlin, 2012.

[17] Z. B. Xu and G. F. Roach, Characteristic inequalities of uniformly convex and uniformly smooth Banach spaces, *J. Math. Anal. Appl.* **157** (1991), no. 1, 189–210.

Received May 2, 2014; revised August 27, 2014; accepted September 21, 2014.

Brought to you by | De Gruyter / TCS
Authenticated
Download Date | 12/9/14 4:09 PM