IOP Publishing Inverse Problems

Inverse Problems **38** (2022) 059401 (3pp) https://doi.org/10.1088/1361-6420/ac612f

Addendum: On stochastic Kaczmarz type methods for solving large scale systems of ill-posed equations (2022 *Inverse Problems* 38 025003)

J C Rabelo[1] **and A Leitão**[2,*]

[1] Department of Mathematics, Federal University of Piaui, 64049-550 Teresina, Brazil

[2] Dep. of Mathematics, Federal Univ. of St. Catarina, PO Box 476, 88040-900 Florianópolis, Brazil

E-mail: joelrabelo@ufpi.edu.br and acgleitao@gmail.com

Received 27 January 2022
Accepted for publication 25 March 2022
Published 8 April 2022

CrossMark

Abstract

We address the convergence analysis derived in (Rabelo *et al* 2022 *Inverse Problems* **38** 025003) for the sPLWK method; a SGD type method for solving large scale systems of ill-posed equations. The assumption constraining the growth rate of the stopping index function $k^* : \delta \mapsto k^*(\delta) \in \mathbb{N}$ is removed; this assumption was needed in the proof of the semi-convergence result. The most important consequence of our findings is the fact that, what semi-convergence concerns, in the sPLWK method the growth rate of $k^*(\delta)$, as δ goes to zero, is independent of the decay rate of the noise level. This is in strong contrast to the deterministic theory, where one needs additional assumptions of the type $\lim_{\delta\to 0}\|\delta\|^2 k^*(\delta) = 0$ for many iterative schemes, i.e., the stopping index should not grow to fast.

Keywords: ill-posed problems, linear systems, SGD methods

In [2] the *stochastic projective Landweber–Kaczmarz* (sPLWK) method for solving large scale systems of linear ill-posed equations of the form $A_i\, x = y_i^\delta$, $i = 0, \ldots, N-1$, is proposed. A corresponding convergence analysis is provided, which includes a semi-convergence result [2, theorem 4.4]. The proof of [2, theorem 4.4] requires the following assumption concerning the *a priori* chosen stopping index function $k^* : \delta \ni (\mathbb{R}^+)^N \mapsto k^*(\delta) \in \mathbb{N}$.

(A5) The function $k^*(\delta)$ satisfies $\lim_{\delta\to 0} k^*(\delta) = \infty$ as well as $\lim_{\delta\to 0}\|\delta\|^2 k^*(\delta) = 0$.

*Author to whom any correspondence should be addressed.

The second part of assumption (A5) couples the growth rate of $k^*(\delta)$ with the decay rate of δ. This assumption is used only once in [2], namely, in the proof of the semi-convergence result.

In what follows we give a new proof of [2, theorem 4.4] that requires, instead of (A5), the weaker assumption.

(A5′) The stopping index function $k^*(\delta)$ satisfies $\lim_{\delta\to 0} k^*(\delta) = \infty$.

New proof of [2, theorem 4.4]. We claim that

$$\|x^\dagger - x_{k+1}^{\delta j}\|^2 \; - \; \|x^\dagger - x_k^{\delta j}\|^2 \; \leqslant \; 0. \tag{1}$$

Indeed, if $\|A_{I_k}^*(A_{I_k}x_k^{\delta j} - y_{I_k}^{\delta j})\| \leqslant \gamma\delta_{I_k}^j$ then $x_{k+1}^{\delta j} = x_k^{\delta j}$ (since $\lambda_{I_k} = 0$). Thus, (1) holds trivially. Otherwise, it follows from [2, (16)] and [2, remark 2.5]

$$\begin{aligned}
&\|x^\dagger - x_k^{\delta j}\|^2 - \|x^\dagger - x_{k+1}^{\delta j}\|^2 \\
&\quad \geqslant \; \theta_k(2-\theta_k)\,\lambda_{I_k}\|A_{I_k}x_k^{\delta j} - y_{I_k}^{\delta j}\|\left(\|A_{I_k}x_k^{\delta j} - y_{I_k}^{\delta j}\| - \delta_{I_k}^j\right) \\
&\quad = \theta_k(2-\theta_k)\,\lambda_{I_k}^2\|A_{I_k}^*(A_{I_k}x_k^{\delta j} - y_{I_k}^{\delta j})\|^2 \\
&\quad > \; \theta_k(2-\theta_k)\,\lambda_{\min}^2\,(\gamma\,\delta_{I_k}^j)^2 \; \geqslant \; 0
\end{aligned}$$

(notice that, due to [2, (A3)], we have $\theta_k(2-\theta_k) > 0$, for $k = 0, 1, \ldots$). Consequently, our claim (1) holds true. Next, taking the average in (1) over all possible realizations we conclude that

$$\mathbb{E}\left[\|x^\dagger - x_{k+1}^{\delta j}\|^2\right] \; - \; \mathbb{E}\left[\|x^\dagger - x_k^{\delta j}\|^2\right] \; \leqslant \; 0. \tag{2}$$

Due to (A5′) we may assume that $k_{\delta^j}^* = k^*(\delta^j)$ increases strictly monotonically with j. Given $m < n$, we add the above inequality, with $j = n$, from $k = k_{\delta^m}^*$ to $k_{\delta^n}^* - 1$, to obtain

$$\begin{aligned}
\mathbb{E}\left[\|x^\dagger - x_{k_n^*}^{\delta^n}\|^2\right] &\leqslant \; \mathbb{E}\left[\|x^\dagger - x_{k_m^*}^{\delta^n}\|^2\right] \\
&\leqslant \; 2\mathbb{E}\left[\|x^\dagger - x_{k_m^*}\|^2\right] + 2\mathbb{E}\left[\|x_{k_m^*} - x_{k_m}^{\delta^n}\|^2\right]
\end{aligned} \tag{3}$$

(we adopted the simplified notation $k_j^* = k_{\delta^j}^*$). Now, [2, theorem 3.6] guarantees the existence of a large enough m, s.t. the first term on the rhs of (3) is smaller then $\varepsilon/2$. Moreover, from [2, theorem 4.3] with $k = k_m^*$ we conclude that the second term on the rhs of (3) is smaller than $\varepsilon/2$ for large enough n, concluding the proof. □

Concluding remarks

—The main difference between the new proof of [2, theorem 4.4] and the former one is the rhs of the inequality (1), and consequently of (3).

—Due to the definition of the stepsize λ_{I_k} (see [2, (6b)]), inequality (1) holds true also for $k > k_\delta^*$ (if one continues to iterate after step k_δ^*). This fact is quintessential to derive the first inequality in (3) (using a telescopic-sum argument), in such a way that no k_δ^*-dependent term is present on the rhs of (3). This is the reason why no coupling assumption between the decay rate of $\|\delta\|$ and the growth rate of $k^*(\delta)$ is needed in the new proof of theorem 4.4.

—The above proof cannot be extended to deterministic iterative regularization methods of gradient type. Indeed, take for example the Landweber iteration with *a priori* chosen stopping index $k^*(\delta)$ and constant stepsize λ_k (see, e.g., [1, section 6.1]).

The monotonicity for the iteration error, inequality (1), holds true only for small iteration indexes $k \in \mathbb{N}$ (namely, before discrepancy takes place). Therefore, the telescopic-sum argument (used in the above proof) cannot be used for arbitrary $m < n \in \mathbb{N}$.

The standard semi-convergence proof for the Landweber iteration uses the stability estimate $\|x_k - x_k^\delta\| \leqslant \sqrt{k}\delta$ (see [1, lemma 6.2]). In this case, the additional assumption $\lim_{\delta\to 0}\|\delta\|^2 k^*(\delta) = 0$ is required to complete the proof.

Data availability statement

No new data were created or analysed in this study.

ORCID iDs

A Leitão https://orcid.org/0000-0001-6785-8835

References

[1] Engl H W, Hanke M and Neubauer A 1996 *Regularization of Inverse Problems* (Dordrecht: Kluwer)
[2] Rabelo J, Saporito Y and Leitão A 2022 On stochastic Kaczmarz type methods for solving large scale systems of ill-posed equations *Inverse Problems* **38** 025003