Linear convergence of a Sparse Randomized Extended Kaczmarz method

Dirk A. Lorenz
TU Braunschweig
d.lorenz@tu-braunschweig.de

Frank Schöpfer
Carl von Ossietzky Universität Oldenburg
frank.schoepfer@uol.de

Lionel Ngoupeyou Tondji
TU Braunschweig
l.ngoupeyou-tondji@tu-braunschweig.de

Maximilian Winkler
TU Braunschweig
maximilian.winkler@tu-braunschweig.de

***Abstract*—The Randomized Extended Kaczmarz method has attracted attention in the last years as an iterative method which converges to the minimum l_2-norm least squares solution of a linear system with a linear expected rate. Also, the Randomized Sparse Kaczmarz method has been shown to converge linearly to a sparse solution of a consistent linear system. Here, we combine both ideas and propose an Extended Sparse Randomized Kaczmarz method. We show linear expected convergence to a sparse least squares solution in the sense that an extended kind of the regularized basis pursuit problem is solved.**

I. Introduction

The *Randomized Sparse Kaczmarz method* uses two variables and reads as

$$x_{k+1}^* = x_k^* - \frac{\langle a_i \,,\, x_k\rangle - b_i}{\|a_i\|_2^2} \cdot a_i$$
$$x_{k+1} = S_\lambda(x_{k+1}^*)$$

with $\lambda > 0$ and the soft shrinkage function

$$S_\lambda(x) = \max\{|x| - \lambda, 0\} \cdot \operatorname{sign}(x).$$

It has been shown in [2] that for a consistent system $Ax = b$ with an arbitrary matrix A the iterates x_k converge linearly in expectation to the unique solution $\hat{x}$ of the *regularized Basis Pursuit problem*,

$$\min_{x\in\mathbb{R}^n} f(x) = \lambda\|x\|_1 + \tfrac{1}{2}\|x\|_2^2 \quad s.t. \quad Ax = b\,.$$

For an inconsistent system, the Kaczmarz method was extended in [1] and [3] to the *Randomized Extended Kaczmarz* method which finds the Moore-Penrose pseudoinverse solution, also with an expected linear rate. The refined method by [3] is given by Algorithm 1.

Algorithm 1 Extended Kaczmarz method (REK-S)

Input: starting points $x_0 \in \mathcal{R}(A^T)$ and $z_0 \in b+\mathcal{R}(A)$, $b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m\times n}$ with rows $a_i^T \in \mathbb{R}^n\backslash\{0\}$ and columns $\tilde{a}_j \in \mathbb{R}^m\backslash\{0\}$
Output: least-squares solution $x^+ = A^+b$
initialize $k = 0$
repeat
  choose a column index $j_k = j \in \{1,\dots,n\}$ with probability $\tilde{p}_j = \|\tilde{a}_j\|_2^2/\|A\|_F^2$
  update $z_{k+1} = z_k - \frac{\langle \tilde{a}_{j_k}\,,\,z_k\rangle}{\|\tilde{a}_{j_k}\|_2^2} \cdot \tilde{a}_{j_k}$
  choose a row index $i_k = i \in \{1,\dots,m\}$ with probability $p_i = \|a_i\|_2^2/\|A\|_F^2$
  update $x_{k+1} = x_k - \frac{\langle a_{i_k}\,,\,x_k\rangle - b_{i_k} + z_{k+1,i_k}}{\|a_{i_k}\|_2^2} \cdot a_{i_k}$
  increment $k = k+1$
until a stopping criterion is satisfied

II. Main Result

Combining both ideas, we propose the following method.

Algorithm 2 Extended Sparse Randomized Kaczmarz (ExSRK)

Input: starting points $x_0 = x_0^* = 0 \in \mathbb{R}^n$ and $z_0 = b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m\times n}$ with rows $a_i^T \in \mathbb{R}^n\backslash\{0\}$ and columns $\tilde{a}_j \in \mathbb{R}^m\backslash\{0\}$
Output: solution of $\min_{x\in\mathbb{R}^n} \lambda\|x\|_1 + \frac{1}{2}\|x\|_2^2$ s.t. $Ax = \hat{y}$, where $\hat{y} = \operatorname{argmin}_{y\in\mathbb{R}^m} \|b - y\|_2^2$ s.t. $y \in \mathcal{R}(A)$
initialize $k = 0$
repeat
  choose a column index $j_k = j \in \{1,\dots,n\}$ with probability $\tilde{p}_j = \|\tilde{a}_j\|_2^2/\|A\|_F^2$
  update $z_{k+1} = z_k - \frac{\langle \tilde{a}_{j_k}\,,\,z_k\rangle}{\|\tilde{a}_{j_k}\|_2^2} \cdot \tilde{a}_{j_k}$
  choose a row index $i_k = i \in \{1,\dots,m\}$ with probability $p_i = \|a_i\|_2^2/\|A\|_F^2$
  update $x_{k+1}^* = x_k^* - \frac{\langle a_{i_k}\,,\,x_k\rangle - b_{i_k} + z_{k+1,i_k}}{\|a_{i_k}\|_2^2} \cdot a_{i_k}$
  update $x_{k+1} = S_\lambda(x_{k+1}^*)$
  increment $k = k+1$
until a stopping criterion is satisfied

Theorem 2.1: The iterates x_k of Algorithm 2 converge in expectation to the unique solution $\hat{x}$ of

$$\min_{x\in\mathbb{R}^n} f(x) := \lambda\|x\|_1 + \tfrac{1}{2}\|x\|_2^2 \quad \text{s.t.} \quad Ax = \hat{y}, \tag{1}$$
$$\text{where} \quad \hat{y} = \operatorname*{argmin}_{y\in\mathbb{R}^m} \|b - y\|_2^2 \quad \text{s.t.} \quad y \in \mathcal{R}(A)\,.$$

For some $q \in (0,1)$ the expected rate of convergence is

$$\mathbb{E}\left[\|x_k - \hat{x}\|_2^2\right] \le q^k \cdot 2 \cdot f(\hat{x}) + k \cdot q^k \cdot \frac{\|\hat{y}\|_2^2}{\|A\|_F^2}\,,$$

where the expectation is taken with respect to the joint probability distribution $p_i = \|a_i\|_2^2/\|A\|_F^2$ and $\tilde{p}_j = \|\tilde{a}_j\|_2^2/\|A\|_F^2$.

Remark 2.2: Also, we are able to show linear convergence for different strongly convex functions f under certain regularity assumptions by updating

$$x_{k+1} = \nabla f^*(x_{k+1}^*).$$

References

[1] A. Zouzias and N. M. Freris, *Randomized extended Kaczmarz for solving least squares*. SIAM Journal on Matrix Analysis and Applications 34.2: 773-793, 2013.

[2] F. Schöpfer, and D. Lorenz, *Linear convergence of the randomized sparse Kaczmarz method*. Mathematical Programming 173.1: 509-536, 2019.

[3] Du, K, *Tight upper bounds for the convergence of the randomized extended Kaczmarz and Gauss–Seidel algorithms*. Numerical Linear Algebra with Applications 26.3: e2233, 2019.