UNIVERSITÄT SALZBURG

Asymptotic Quadratic Convergence of the Two-Sided Serial and Parallel Block-Jacobi SVD Algorithm - Update

Gabriel Okša   Yusaku Yamamoto   Marián Vajteršic

Technical Report 2018-01   May 2018

Department of Computer Sciences

Jakob-Haringer-Straße 2
5020 Salzburg
Austria
www.cosy.sbg.ac.at

Technical Report Series

Asymptotic Quadratic Convergence of the Two-Sided Serial and Parallel Block-Jacobi SVD Algorithm - Update

Gabriel Okša* Yusaku Yamamoto† Marián Vajteršic‡

April 9, 2018

Abstract. *This report provides an update of the Salzburg report [4] from the year 2017. The upper bound in the proof of the global convergence of the parallel two-sided block-Jacobi SVD algorithm was improved. The scaled matrices in the Jacobi iteration process were introduced, and the asymptotic quadratic convergence of their off-diagonal Frobenius norm was proved both for the serial and parallel two-sided block-Jacobi SVD algorithm. Finally, the upper bound for the number of parallel iteration steps W, over which the asymptotic quadratic convergence can be observed, was improved.*

1 Serial two-sided SVD algorithm

The detailed description of the algorithm as well as the proof of the asymptotic quadratic convergence of the off-diagonal Frobenius norm can be found in [4].

1.1 Asymptotic quadratic convergence

We now introduce the *s*caling of iteration matrices $A^{(k)}$ in the process. Let $d^{(k)}$ be the vector of diagonal elements of $A^{(k)}$, and denote by $\mathrm{diag}(d^{(k)})$ the diagonal matrix of order n where its diagonal elements are the diagonal elements of $A^{(k)}$. Let $A^{(k)}$ be given by its rows, $A^{(k)} = (r_1^{(k)}, r_2^{(k)}, \ldots, r_n^{(k)})^T$, or by its columns, $A^{(k)} = (c_1^{(k)}, c_2^{(k)}, \ldots, c_n^{(k)})$. Define the diagonal left and

*Mathematical Institute, Slovak Academy of Sciences, Bratislava, Slovak Republic, email: gabriel.oksa@savba.sk

†Department of Communication Engineering and Informatics, University of Electro-Communications, Tokyo, Japan, email: yusaku.yamamoto@uec.ac.jp

‡Department of Computer Sciences, University of Salzburg, Austria, and Mathematical Institute, Slovak Academy of Sciences, Bratislava, Slovak Republic, email: marian@cosy.sbg.ac.at

right scaling matrix $D_L^{(k)}$ and $D_R^{(k)}$ at iteration step k, respectively, as follows:

$$D_L^{(k)} \equiv \text{diag}(\|r_1^{(k)}\|^{1/2}, \|r_2^{(k)}\|^{1/2}, \ldots, \|r_n^{(k)}\|^{1/2}),\ D_R^{(k)} \equiv \text{diag}(\|c_1^{(k)}\|^{1/2}, \|c_2^{(k)}\|^{1/2}, \ldots, \|c_n^{(k)}\|^{1/2}). \tag{1}$$

Now define the $s$caled iteration matrix $\tilde{A}^{(k)}$ by

$$\tilde{A}^{(k)} \equiv (D_L^{(k)})^{-1} A^{(k)} (D_R^{(k)})^{-1}. \tag{2}$$

Note that the off-diagonal Frobenius norm of the scaled iteration matrix $\tilde{A}^{(k)}$ is given by

$$\|\text{off}(\tilde{A}^{(k)})\|_F = \|(D_L^{(k)})^{-1} A^{(k)} (D_R^{(k)})^{-1} - (D_L^{(k)})^{-1} \text{diag}(d^{(k)}) (D_R^{(k)})^{-1}\|_F.$$

In [4], the asymptotic quadratic convergence (AQC) of the off-diagonal Frobenius norm was proved under the assumption of the existence of some positive constant δ–see [4] for details. The final estimate has the form:

$$\|\text{off}(A^{(W)})\|_F \leq \sqrt{2(w-2)}\, \frac{\|\text{off}(A^{(0)})\|_F^2}{\delta}, \tag{3}$$

where $W = w(w-1)/2$.

We now prove the AQC of the off-diagonal Frobenius norm of the scaled iteration matrices starting from Eq. (3). To this end, multiply both sides of Eq. (3) by two positive constants $\|(D_L^{(W)})^{-1}\|_2$ and $\|(D_R^{(W)})^{-1}\|_2$:

$$\|(D_L^{(W)})^{-1}\|_2\, \|\text{off}(A^{(W)})\|_F\, \|(D_R^{(W)})^{-1}\|_2 \leq \sqrt{2(w-2)} \|(D_L^{(W)})^{-1}\|_2\, \|(D_R^{(W)})^{-1}\|_2 \frac{\|\text{off}(A^{(0)})\|_F^2}{\delta}. \tag{4}$$

Let us bound the LHS (left-hand side) of Eq. (4) from below:

$$\begin{aligned} \|(D_L^{(W)})^{-1}\|_2\, \|\text{off}(A^{(W)})\|_F\, \|(D_R^{(W)})^{-1}\|_2 &= \|(D_L^{(W)})^{-1}\|_2\, \|A^{(W)} - \text{diag}(d^{(k)})\|_F\, \|(D_R^{(W)})^{-1}\|_2 \\ &\geq \|(D_L^{(W)})^{-1}\, A^{(W)}\, (D_R^{(W)})^{-1} - (D_L^{(W)})^{-1}\, \text{diag}(d^{(k)})\, (D_R^{(W)})^{-1}\|_F = \|\text{off}(\tilde{A}^{(W)})\|_F. \end{aligned}$$

Finally, we bound the RHS of Eq. (4) from above. The following auxiliary estimates will be useful:

$$\begin{aligned} \|(D_L^{(W)})^{-1}\|_2 &= \frac{1}{\min_{1\leq i\leq n}\{\|r_i^{(W)}\|^{1/2}\}} \leq \frac{1}{\sigma_n^{1/2}}, \\ \|(D_R^{(W)})^{-1}\|_2 &= \frac{1}{\min_{1\leq i\leq n}\{\|c_i^{(W)}\|^{1/2}\}} \leq \frac{1}{\sigma_n^{1/2}}, \\ \|(D_L^{(0)})\|_2^2 &= \left(\max_{1\leq i\leq n}\{\|r_i^{(0)}\|^{1/2}\}\right)^2 \leq \sigma_1, \\ \|(D_R^{(0)})\|_2^2 &= \left(\max_{1\leq i\leq n}\{\|c_i^{(0)}\|^{1/2}\}\right)^2 \leq \sigma_1. \end{aligned} \tag{5}$$

Then:

$$\begin{aligned}
&\sqrt{2(w-2)}\, \|(D_L^{(W)})^{-1}\|_2 \, \|(D_R^{(W)})^{-1}\|_2 \, \frac{\|\mathrm{off}(A^{(0)})\|_F^2}{\delta} \\
&= \sqrt{2(w-2)}\, \|(D_L^{(W)})^{-1}\|_2 \, \|(D_R^{(W)})^{-1}\|_2 \, \frac{\|A^{(0)} - \mathrm{diag}(d^{(0)})\|_F^2}{\delta} \\
&= \sqrt{2(w-2)}\, \|(D_L^{(W)})^{-1}\|_2 \, \|(D_R^{(W)})^{-1}\|_2 \cdot \\
&\quad \cdot \frac{\|D_L^{(0)}(D_L^{(0)})^{-1}A^{(0)}(D_R^{(0)})^{-1}D_R^{(0)} - D_L^{(0)}(D_L^{(0)})^{-1}\mathrm{diag}(d^{(0)})(D_R^{(0)})^{-1}D_R^{(0)}\|_F^2}{\delta} \\
&\leq \sqrt{2(w-2)}\, \|(D_L^{(W)})^{-1}\|_2 \, \|(D_R^{(W)})^{-1}\|_2 \, \|D_L^{(0)}\|_2^2 \, \|D_R^{(0)}\|_2^2 \, \frac{\|\mathrm{off}(\tilde{A}^{(0)})\|_F^2}{\delta} \\
&\leq \sqrt{2(w-2)}\, \frac{\sigma_1}{\sigma_n} \, \frac{\|\mathrm{off}(\tilde{A}^{(0)})\|_F^2}{(\delta/\sigma_1)}.
\end{aligned}$$

Defining $\mu \equiv \delta/\sigma_1$, using the condition number $\kappa(A) = \sigma_1/\sigma_n$ and applying the above estimates for the LHS and RHS of Eq. (4), one gets the AQC of the off-norm of scaled iteration matrices after W serial iteration steps:

$$\|\mathrm{off}(\tilde{A}^{(W)})\|_F \leq \sqrt{2(w-2)}\, \kappa(A) \frac{\|\mathrm{off}(\tilde{A}^{(0)})\|_F^2}{\mu}. \tag{6}$$

2 Parallel two-sided SVD algorithm with dynamic ordering

The detailed description of the algorithm as well as the proof of the asymptotic quadratic convergence of the off-diagonal Frobenius norm can be found in [4].

The global convergence of the parallel two-sided block-Jacobi SVD algorithm with the greedy implementation of the parallel dynamic ordering (GIPDO) was proved in [4] with the original upper bound

$$\|\mathrm{off}(A^{(k+1)})\|_F^2 \leq \left(1 - \frac{2}{w(w-1)}\right) \|\mathrm{off}(A^{(k)})\|_F^2.$$

In [1], this upper bound was improved to:

$$\|\mathrm{off}(A^{(k+1)})\|_F^2 \leq \left(1 - \frac{1}{2w-3}\right) \|\mathrm{off}(A^{(k)})\|_F^2. \tag{7}$$

Hence, $\|\mathrm{off}(A^{(k)})\|_F^2$ decreases at least as fast as the geometric sequence with the quotient $(W-1)/W$, $W = 2w-3$, and therefore converges to zero. Note that this proof does not depend on the distribution of singular values of A.

2.1 Asymptotic quadratic convergence

First, the better upper bound is derived for the number of parallel iteration steps W, after which the parallel block-Jacobi SVD algorithm with GIPDO converges quadratically.

We start with the definition of two auxiliary index sets. Without loss of generality, denote the parallel iteration steps by $k = 0, 1, \ldots$, and the off-diagonal blocks chosen for zeroing at step k by $A^{(k)}_{X_{k,1}Y_{k,1}}$, $A^{(k)}_{Y_{k,1}X_{k,1}}$, …, $A^{(k)}_{X_{k,p}Y_{k,p}}$, $A^{(k)}_{Y_{k,p}X_{k,p}}$, where $A^{(k)}_{X_{k,1}Y_{k,1}}$ and $A^{(k)}_{Y_{k,1}X_{k,1}}$ are the off-diagonal blocks that give the largest weight. Let $\mathcal{Q}_{k,\ell}$, ℓ even, be the index set of the $\ell/2$ pairs of off-diagonal blocks with the smallest weights at step k. Notice that they are chosen in a symmetric way–i.e., if $(I, J) \in Q_{k,\ell}$ then $(J, I) \in Q_{k,\ell}$. Consequently, $|\mathcal{Q}_{k,\ell}| = \ell$, where $|\mathcal{A}|$ denotes the number of elements in a set $\mathcal{A}$.

In addition, define the second symmetric index set $\mathcal{P}_k$ recursively as follows:

$$\begin{aligned} \mathcal{P}_0 &\equiv \emptyset, \\ \mathcal{P}_{k+1} &\equiv \mathcal{Q}_{k,|\mathcal{P}_k|} \cup \{(X_{k,1}, Y_{k,1}), (Y_{k,1}, X_{k,1}), \ldots, (X_{k,p}, Y_{k,p}), (Y_{k,p}, X_{k,p})\}. \end{aligned} \tag{8}$$

The set of block indices $\mathcal{P}_k$ is again symmetric, and $|\mathcal{P}_k| = |\mathcal{Q}_{k,|\mathcal{P}_k|}|$. Now we prove the following lemma.

Lemma 1 *There exists a step W, $w - 1 \le W < 2w(\log w + 1)$, for which $\mathcal{P}_W$ equals the set of indices of all off-diagonal blocks.*

Proof: We first show that $|\mathcal{P}_k|$ is a strictly increasing sequence and thereby prove the existence of W. $|\mathcal{P}_k|$ takes a value between 0 and $w(w - 1) = 2p(2p - 1)$. As will be shown below, the increase of $|\mathcal{P}_k|$ at each parallel step differs depending on the value of $|\mathcal{P}_k|$ itself. Accordingly, we define the following $p - 1$ sub-intervals of $[0, 2p(2p - 1)]$ and consider the increase of $|\mathcal{P}_k|$ in each sub-interval separately:

$$I_1 \equiv [0,\ 4 \cdot 3 - 2], \tag{9}$$

$$I_i \equiv (2i(2i - 1) - 2,\ (2i + 2)(2i + 1) - 2] \quad i = 2, 3, \ldots, p - 1. \tag{10}$$

Note that the intervals I_i, $2 \le i \le p-1$, are open from the left. Now consider the construction of $\mathcal{P}_{k+1}$ from $\mathcal{P}_k$. In GIPDO, the block pair $(A^{(k)}_{X_{k,\ell}Y_{k,\ell}}, A^{(k)}_{Y_{k,\ell}X_{k,\ell}})$, $1 \le \ell \le p$, is chosen to be the pair of maximal weight under the condition that $X_{k,\ell}, Y_{k,\ell} \notin \{X_{k,1}, Y_{k,1}, \ldots, X_{k,\ell-1}, Y_{k,\ell-1}\}$. Hence, it gives the maximal weight among $(w-(2\ell-2))(w-(2\ell-1))/2$ block pairs. In other words, among the $w(w-1)/2$ off-diagonal block pairs of $A^{(k)}$, there are at least $(w-(2\ell-2))(w-(2\ell-1))/2-1$ pairs whose weights are not larger than that of $(A^{(k)}_{X_{k,\ell}Y_{k,\ell}}, A^{(k)}_{Y_{k,\ell}X_{k,\ell}})$. Now assume that $|\mathcal{P}_k| \in I_i$. Consequently, $|\mathcal{Q}_{k,|\mathcal{P}_k|}| = |\mathcal{P}_k| \le (2i + 2)(2i + 1) - 2$. Hence, $\mathcal{Q}_{k,|\mathcal{P}_k|}$ consists of at most $(2i + 2)(2i + 1)/2 - 1$ pairs of block indices that identify the pairs of off-diagonal blocks with the smallest weights. On the other hand, as stated above for $\ell = p - i$, there are at least

$$(w - (2(p - i) - 2))(w - (2(p - i) - 1))/2 - 1 = (2i + 2)(2i + 1)/2 - 1$$

pairs of off-diagonal blocks whose weights are not larger than that of $(A^{(k)}_{X_{k,p-i}Y_{k,p-i}}, A^{(k)}_{Y_{k,p-i}X_{k,p-i}})$. Thus, by choosing the elements of $\mathcal{Q}_{k,|\mathcal{P}_k|}$ from them, we can ensure that the $p - i$ pairs of block indices $(X_{k,1}, Y_{k,1}), (Y_{k,1}, X_{k,1}), \ldots, (X_{k,p-i}, Y_{k,p-i}), (Y_{k,p-i}, X_{k,p-i})$ do not belong to $\mathcal{Q}_{k,|\mathcal{P}_k|}$. Combining this with Eq. (8), we know that $\mathcal{P}_{k+1}$ has at least $2(p - i)$ more elements than $\mathcal{P}_k$. Since the length of interval I_i is $8i + 2$, it takes at most $\lfloor \frac{8i+2}{2(p-i)} \rfloor + 1$ steps for $|\mathcal{P}_k|$ to increase from the left end of I_i to its right end, where $\lfloor \cdot \rfloor$ is the floor function. Thus, the number of steps required for $|\mathcal{P}_k|$ to increase from 0 to $w(w-1)-2$ can be bounded by

$$\begin{aligned}
\sum_{i=1}^{p-1}\left(\left\lfloor\frac{8i+2}{2(p-i)}\right\rfloor+1\right) &\le \sum_{i=1}^{p-1}\left(\frac{8i+2}{2(p-i)}+1\right) \\
&= (4p+1)\sum_{j=1}^{p-1}\frac{1}{j}-3(p-1) \\
&\le (4p+1)\left(1+\int_1^{p-1}\frac{1}{x}\,dx\right)-3(p-1) \\
&= (4p+1)(1+\log(p-1))-3(p-1) \\
&< 4p(1+\log p)-1 \\
&< 2w(1+\log w)-1. \qquad (11)
\end{aligned}$$

Once $|\mathcal{P}_k|$ reaches $w(w-1)-2$, one more step is sufficient for it to reach $w(w-1)$. This follows from the fact that the pair of off-diagonal blocks with block indices $(X_{k,1}, Y_{k,1})$ and $(Y_{k,1}, X_{k,1})$ has the largest weight among all pairs, and therefore these block indices do not belong to $\mathcal{Q}_{k,|\mathcal{P}_k|}$ whenever $|\mathcal{Q}_{k,|\mathcal{P}_k|}| = |\mathcal{P}_k| \le w(w-1)-2$. This establishes the existence of W and, at the same time, proves its upper bound.

To prove the lower bound of W, notice that $|\mathcal{P}_k|$ can be increased by at most $w = 2p$ at each step (see Eq. (8)). Thus, it requires at least $w-1$ steps to increase $|\mathcal{P}_k|$ from 0 to $w(w-1)$. $\square$

Remark Theorem 1 in [4] shows that the serial block-Jacobi SVD method converges quadratically after $w(w-1)/2$ iterations. On the other hand, the above Lemma 1 claims that the parallel block Jacobi method with GIPDO converges quadratically after at most $2w(\log w+1)$ parallel steps. These results indicate the ability of the parallel method to reduce the number of iteration steps by a factor of $O(w/\log w)$ and $O(w)$ in the worst and best case, respectively. $\square$

Next, the AQC of the off-diagonal Frobenius norm for scaled iteration matrices will be proved. The AQC of the off-diagonal Frobenius norm was proved in [4]. The final estimate has the following form:

$$\|\mathrm{off}(A^{(W)})\|_F \le \sqrt{12(w-2)}\,\frac{\|\mathrm{off}(A^{(0)})\|_F^2}{\delta}. \qquad (12)$$

Starting from Eq. (12) and using the same approach as in Section 1.1, one gets the AQC of the off-diagonal Frobenius norm for scaled matrices in the parallel two-sided SVD algorithm with GIPDO after W parallel iteration steps:

$$\|\mathrm{off}(\tilde{A}^{(W)})\|_F \le \sqrt{12(w-2)}\,\kappa(A)\frac{\|\mathrm{off}(\tilde{A}^{(0)})\|_F^2}{\mu}, \qquad (13)$$

where $\mu = \delta/\sigma_1$.

3 Conclusion

Note that all above mentioned updates, and especially the proof of the AQC of the off-diagonal Frobenius norm for scaled iteration matrices, can be applied to the classical serial block-Jacobi EVD algorithm in [2], and to the parallel block-Jacobi EVD algorithm with GIPDO in [3].

References

[1] S. Kudo, K. Yasuda, Y. Yamamoto, *Performance of the parallel block Jacobi method with dynamic ordering for the symmetric eigenvalue problem*, submitted to JSIAM Letters.

[2] G. Okša, Y. Yamamoto, M. Vajteršic, *Asymptotic quadratic convergence of the serial block-Jacobi EVD algorithm for Hermitian matrices*, Numer. Math. (2016) DOI: 10.1007/s00211-016-0863-5.

[3] G. Okša, Y. Yamamoto, M. Vajteršic, *Asymptotic Quadratic Convergence of the Parallel Block-Jacobi EVD Algorithm*, Research Report, Salzburg, 2016.

[4] G. Okša, Y. Yamamoto, M. Vajteršic, *Asymptotic Quadratic Convergence of the Two-Sided Serial and Parallel Block-Jacobi SVD Algorithm*, Research Report, Salzburg, 2017.