Project in the context of the course "High performance computing for numerical methods and data analysis"

1 Important information

Your project needs to be submitted by e-mail to *Laura.Grigori@inria.fr*, by December 18, 2021. The e-mail should contain the following:

- The pdf of your report, which should not be longer than 7 pages. Supplementary material, as extra figures, can be provided as an Appendix.
- The code that you have used for your project. Any language can be used for coding the randomized algorithms, as Matlab, Python, Julia, C, C++.
- Test matrices: the algorithms should be tested on about 4 matrices. You should pick matrices which have fast decay of their singular values, but also matrices for which the singular values decay slowly. You can find a set of matrices in the paper [1], Table 2, page 78, as well as their description. Some of the matrices come from the San Jose State Singular Matrix Database (example of usage in Matlab "Problem = SJget(135)) ". The paper can be downloaded from `https://who.rocq.inria.fr/Laura.Grigori/Papers/92157.pdf`
- In this project, studying the numerical behavior of a low rank approximation algorithm means that you should compare the approximated singular values by the randomized algorithms with the singular values as computed by the SVD.

2 Goal of the project

The goal of this project is to study randomized algorithms for computing a rank-k approximation of a matrix $A \in \mathbb{R}^{m \times n}$ as $A_k = \tilde{U}\tilde{W}^T$ where $\tilde{U} \in \mathbb{R}^{m \times k}$ and $\tilde{W}^T \in \mathbb{R}^{k \times n}$. The SVD provides the best rank-k approximation, however it is too expensive to compute.

We will study the randomized SVD algorithm, which computes first an approximate basis for the range of $A \in \mathbb{R}^{m \times n}$ by multiplying A with a random matrix $\Omega_1 \in \mathbb{R}^{n \times (k+p)}$, where the rank of the approximation is k and p is an oversampling parameter. We note $l = k + p$. Then it computes the QR factorization of $A\Omega_1$ to get Q_1 with orthonormal columns, computes the truncated SVD of $Q_1^T A$ and obtains the approximation A_k of A. The algorithm is described in the lecture of November 10, 2021 on randomized low rank approximation.

3 Content of the report (7 pages maximum)

The report must contain the following elements, which will guide also the approach to use in the project.

1. A description of the randomized SVD algorithm, followed by an experimental study of its numerical behavior. We consider first that Ω_1 is i.i.d. $N(0, 1)$ random matrix. To study the numerical behavior, you should compare your result with the result of the SVD algorithm. The rank k is a parameter that you should choose. The test matrices don't have to be big, they could be of dimension 250×250 and the rank k could be between 16 and 64.
2. A description of the parallelization of this algorithm, with the presentation of a pseudo-code algorithm using MPI routines for communication. We consider Ω_1 is given.
3. An analysis of the computation and communication costs of this algorithm, using the model α, β, γ that was studied during the course. For this analysis, the matrix A is considered to be dense.

4. Consider now the randomized SVD algorithm in which $\Omega_1 \in \mathbb{R}^{n \times l}$ is a subsampled random Hadamard transform (SRHT). Study experimentally its numerical behavior by comparing the approximation of the singular values obtained by this algorithm with the results obtained when using a i.i.d. N(0,1) random matrix, point 1. The bounds for this approximation are described in Corollary 9 from the lecture of November 10, 2021 on randomized low rank approximation. Discuss the choice of l.

5. Bonus: to go further, consider the randomized generalized LU algorithm, as described in the lecture on randomized algorithms, whose properties are given in Theorem 10 from the lecture of November 10, 2021 on randomized low rank approximation.

References

[1] J. Demmel, L. Grigori, M. Gu, and H. Xiang. Communication-avoiding rank-revealing qr decomposition. *SIAM Journal on Matrix Analysis and its Applications*, 36(1):55–89, 2015.

[2] J. Demmel, L. Grigori, and A. Rusciano. An improved analysis and unified perspective on deterministic and randomized low rank matrix approximations. Technical report, Inria, 2019. available at https://arxiv.org/abs/1910.00223.

[3] N. Halko, P.-G. Martinsson, and J. A. Tropp. Finding structure with randomness: probabilistic algorithms for constructing approximate matrix decompositions. *SIAM Review*, 53:217–288, 2011.