CEMRACS 2016: project ALORA
Arhictecture aware LOw RAnk approximations

Leader: Laura Grigori, Inria Paris, Alpines

Participants:

- Alan Ayala, Inria Alpines
- Sebastien Cayrols, Inria Alpines/Maison de la Simulation
- Xavier Claeys, UPMC/Inria Alpines
- Laura Grigori, Inria Alpines
- Maxim Naumov, Nvidia
- Benoit Pelleteir, ATOS/Bull
- Stéphane Pralet, ATOS/Bull

Funding: Inria Alpines, NLAFET FETHPC H2020 project, ATOS/Bull, Nvidia

Abstract: In this project we consider the problem of computing a low rank approximation of a large matrix on large scale computers. This problem has many different applications, ranging from scientific computing problems such as fast solvers for integral equations to data analytics problems such as principal component analysis (PCA) or image processing. The singular value decomposition produces the best rank-k approximation, however it is expensive to compute. We focus in this project on rank revealing LU factorization with tournament pivoting, a technique that was introduced recently. We will study this algorithm for both dense and sparse matrices and its adaptability on current and future architectures of interest for Bull.

Presentation of the project In this project we consider the problem of computing a low rank approximation of a large sparse matrix. This problem has many different applications, ranging from scientific computing problems such as fast solvers for integral equations to data analytics problems such as principal component analysis (PCA) or image processing. The singular value decomposition produces the best rank-k approximation, however it is expensive to compute. Several different less expensive approaches exist in the literature that approximate the singular value decomposition of a matrix, such as rank revealing QR and LU factorizations, or the Lanczos algorithm. In the recent years, several randomized algorithms have been introduced for this problem that aim at further decreasing the computational cost while obtaining accurate results with high probability. We focus in this project on rank revealing LU factorization with tournament pivoting, a technique that was introduced in [1, 2].

Objectives The goal of this project is to study the low rank approximation algorithm based on tournament pivoting and its performance on current parallel computers as those from Bull or Nvidia.

The current algorithm is efficient on computing low rank approximations when the rank is known in advance. However in many cases the rank needs to be discovered during the factorization. One goal is to extend tournament pivoting such that the rank can be discovered during the factorization while maintaining a near linear complexity. Another goal is to study its numerical

accuracy in theory and on practical applications provided by the industrial partners of the prohect. And finally, the performance of the algorithm on architectures of interest for Bull and Nvidia will be tested and improved.

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] L. Grigori, S. Cayrols, and J. W. Demmel. Low rank approximation of a sparse matrix based on lu factorization with column and row tournament pivoting. Technical Report 8910, Inria, 2016.