2018-00532 - [Campagne CORDI-S-CRI Paris] Large scale compression and approximation techniques for tensor computations

Level of qualifications required: Graduate degree or equivalent
Fonction: PhD Position

Context
The research will be performed in the context of Alpines team. Alpines is a joint research group between INRIA Paris and J.L. Lions Laboratory, Sorbonne University, which focuses on scientific computing. The topics studied in the group are mesh generation for parallel computation, linear solvers based on domain decomposition methods and incomplete factorizations, and computational kernels for sparse and dense linear algebra. The developed numerical algorithms are regularly integrated into FreeFem++ package and are validated on challenging numerical simulations in collaboration with our academic and industrial partners.

In summary, the topics studied in our group are the following:

- Numerical methods and algorithms
  - Computational kernels for numerical linear algebra
  - Linear solvers
  - Middle-layer for numerical simulations – parallel FreeFem++
  - Modelisation and numerical simulations

Assignment
This PhD thesis focuses on large scale compression and approximation techniques for tensor computations. Tensor computations are a very active area of research, and are used in both scientific computing and data analytics applications. Indeed, tensor numerical methods can be used to solve multidimensional PDEs, parametric PDEs, uncertainty quantification problems. The large scale nature of these problems can be tackled by using compression and approximation techniques which result in using tensors with low dimensional linear multilinear structure. In addition, the usage of large scale computers is very often mandatory in this context. One of the challenges here is the fact that massively parallel computers have extremely complex hardware architectures, formed by thousands of multicore processors, and most of the current algorithms are not able to efficiently exploit these architectures. This is due to the exponentially increasing gap between the time required to perform floating-point operations by one of the processors and the time it takes to communicate its result to another. This challenging question will be addressed by designing algorithms that minimize communication.

Main activities
In this research context, we consider a class of tensor algorithms that rely on compression and approximation techniques. Our goal is to introduce efficient and robust tensor algorithms that are also able to exploit the current peta- and future exa-scale machines by minimizing communication.

The first goal of this PhD is to study communication lower bounds for tensor operations. In the case of linear algebra lower bounds on communication for direct methods of factorizations were introduced in the recent years (see e.g. [4]). Based on the lower bounds, the second goal of this thesis is to develop algorithms that are able to drastically reduce the communication cost with respect to previous approaches, or even attain the lower bounds whenever possible. We will first consider compression techniques for tensors. While in two dimensions the truncated singular value decomposition (SVD) provides the best low rank approximation of a matrix, it is still an open question to define the equivalent of the SVD in case of tensors. We will study here a definition of an analogous singular value and singular vector of a tensor and use this to define low rank approximations of a tensor (by extending tournament pivoting techniques introduced in [4]). We will then consider solving systems of equations that involve an operator with a tensor structure by considering hierarchical decompositions, in the spirit of hierarchical matrices. The new algorithms will be implemented on large scale machines and their impact will be evaluated on several real applications.

Skills

General Information
- Theme/Domain: Distributed and High Performance Computing
  Scientific computing (BAP E)
- Town/city: Paris
- Inria Center: CRI de Paris
- Starting date: 2018-10-01
- Duration of contract: 3 years
- Deadline to apply: 2018-04-23

Contacts
- Inria Team: ALPINES
- Recruiter: Grigori Laura / laura.grigori@inria.fr

The keys to success

Conditions for application

Defence Security:
This position is likely to be situated in a restricted area (ZRR), as defined in Decree No. 2011-1425 relating to the protection of national scientific and technical potential (PPST). Authorisation to enter an area is granted by the director of the unit, following a favourable Ministerial decision, as defined in the decree of 3 July 2012 relating to the PPST. An unfavourable Ministerial decision in respect of a position situated in a ZRR would result in the cancellation of the appointment.

Recruitment Policy:
As part of its diversity policy, all Inria positions are accessible to people with disabilities.

Warning: you must enter your e-mail address in order to save your application to Inria. Applications must be submitted online on the Inria website. Processing of applications sent from other channels is not guaranteed.
Master in applied mathematics or computer science. Good background in numerical linear algebra. Knowledge of C coding. Knowledge of parallel computing or parallel programming is a plus.

Benefits package
- Subsidised catering service
- Partially-reimbursed public transport

Remuneration
Gross Salary per month: 1 982 € the first 2 years and 2 085 € the last year