Updating the Diagonalization of a Symmetric Matrix [1]

Fabienne Jézéquel
LIP6, 75252 Paris Cedex 05, France
Fabienne.Jezequel@lip6.fr

and

Bernard Philippe
INRIA, Campus Beaulieu, 35042 Rennes Cedex, France
Bernard.Philippe@inria.fr

Problem definition. It is not rare that an application gives rise to a sequence of symmetric matrices $(A_k)_{k\geq 0} \subset \mathbb{R}^{n\times n}$ which represents the evolution of an operator with respect to a parameter such as the time. In this talk, we assume that a full diagonalization of each matrix of the sequence is required. When the dependency on the parameter is not too stiff, one may expect an efficient updating process to get the spectral decomposition of A_{k+1} from the decomposition of A_k.

Let us assume that $A_k = U_k D_k U_k^T$, where U_k is an orthogonal matrix of a full set of eigenvectors, and D_k a diagonal matrix, of which the main diagonal entries are the eigenvalues. By computing the matrix

$$B_{k+1} \quad = \quad U_k^T A_{k+1} U_k, \tag{1}$$

it may be expected that B_{k+1} is close to a diagonal matrix. If a process may compute the eigenvector matrix V_{k+1} of B_{k+1} at the cost of a small number of operations, the spectral decomposition may be retrieved by :

$$\begin{aligned} &V_{k+1}^T B_{k+1} V_{k+1} = D_{k+1}, \\ &\qquad \text{and therefore } A_{k+1} = U_{k+1} D_{k+1} U_{k+1}^T, \text{ with } U_{k+1} = U_k V_{k+1}. \end{aligned} \tag{2}$$

The spectral decomposition of a symmetric matrix of order n can be directly obtained with $9n^3 + O(n^2)$ operations (tridiagonalization followed by a QR method, see [2]). Therefore, to introduce a more efficient method with the approach described in equation (2), it is necessary to devote a maximum of $\mathcal{C}_{k+1} \leq 4n^3 + O(n^2)$ operations to the diagonalization of B_{k+1}.

[1] This research is supported by the French program PEPS *Intéractions Maths-Informatique-Ingénierie (CNRS-INSMI)* under the action *Diagonalisation de matrices denses symétriques sur architecture hybride de type GPU (DIAM-GPU).*

This conclusion must be adapted in relation the architecture under target: for instance the performances of a parallel or GPU computer can behave at very different speed depending on the computing kernels.

Two methods for updating a spectral decomposition The Jacobi method [2] iteratively applies 2×2-rotations on both sides of the matrix B_{k+1}, to converge to a diagonal shape. In the present context, the goal is to involve a limited number of rotations in order to be more efficient than a direct computation. It can be estimated that the number of necessary rotations must not exceed $2/3$ of a sweep. For this approach, it might be of interest to look at the block version of the Jacobi algorithms [1] to perform more efficiently the necessary rotations.

The second approach considers the splitting $B_{k+1} = D + \Delta$ where D is a diagonal matrix and Δ is viewed as a perturbation. In [4], a Taylor expansion of the spectral projectors of $D + \Delta$ from D is defined. The expansion is provided by the expansion pf the resolvent in the spectral theorem (see [3]) which defines the spectral projector of a matrix $D + \Delta$ which corresponds to the eigenvalues included in an interval J, diameter of a circle Γ:

$$P_{D+\Delta,J} = \frac{1}{2\pi i}\int_{\Gamma}(zI - D - \Delta)^{-1}dz = P_{D,J} + T^{(1)} + T^{(2)} + R, \quad (3)$$

where $P_{D,J}$ is a part of $\mathbb{R}^n$ abd where $\|R\| = O(\|\Delta\|^3)$. The first order and second order terms can be explicitely derived.

Numerical tests illustrate that the two approaches compete with the straight diagonalization only when the quantity $\|A_{k+1} - A_k\|$ remains small.

References

[1] M. Becka, Gabriel Oksa and M. Vajtersic. *Dynamic ordering for a parallel block-Jacobi SVD algorithm*, Parallel Computing, 28, (2002), pp. 243-262.

[2] G. H. Golub and C. F. Van Loan. Matrix computations. The John Hopkins University Press, Baltimore and London, 3rd edition, 1997.

[3] T. Kato. The perturbation theory for linear operator. Springer-Verlag, Berlin, Heidelberg, New-York, 1980.

[4] B. Philippe. *Perturbation de la décomposition spectrale d'une matrice hermitienne*, Research Report INRIA n°269, Rocquencourt, 1984.

New Ordering for the Parallel One-Sided Block-Jacobi SVD Algorithm

Gabriel Okša
Mathematical Institute, Department of Informatics,
Slovak Academy of Sciences, Bratislava, Slovak Republic
Gabriel.Oksa@savba.sk

We have designed, implemented and tested the new dynamic ordering for the parallel one-sided block-Jacobi SVD algorithm. Our idea is based on the estimation of the cosines of principal angles between two block columns X and Y of the same width without explicitly forming the matrix product X^TY (or Y^TX) and computing its SVD. Instead, we propose to use a fixed number k of iterations in the Lanczos algorithm applied to the symmetric Jordan-Wielandt matrix with zero diagonal blocks, $(2,1)$-block X^TY and $(1,2)$-block Y^TX. When the width of block columns is L, the order of the Jordan-Wielandt matrix is $2L$. However, the matrix blocks X^TY and Y^TX are never formed explicitly; the needed matrix-vector multiplications are computed in two steps (e.g., $X^T(Yv)$) by sending/receiving the vector y (or the preliminary result Yv) between two processors that host the block column X and Y. After computing k iterations, the Frobenius norm of an auxiliary tridiagonal matrix of order k estimates the square root of twice the sum of squares of k largest cosines (representing k smallest principal angles) between X and Y. In the parallel algorithm using p processors, these weights can be used for choosing p pairs of block columns, which are far from orthogonality with respect to those k smallest angles. We show how to implement this new parallel ordering in the distributed paradigm of parallel computing using the Message Passing Interface (MPI). First numerical results show that the one-sided parallel dynamic ordering can lead to a substantial decrease of the number of parallel iteration steps needed for the convergence as compared to a cyclic ordering.

Nonnegative Matrix and Tensor Factorizations for Knowledge Discovery from Textual Media

Michael W. Berry
Center for Intelligent Systems and Machine Learning,
University of Tennessee, Knoxville, USA
berry@eecs.utk.edu

The development and use of low-rank approximate nonnegative matrix factorization (NMF) algorithms are highly needed for effective feature and scenario extraction and identification from digital media. Using these factorizations to retain natural data nonnegativity, one can eliminate subtractive basis vector and encoding calculations present in techniques such as principal component analysis for semantic feature abstraction. Moving beyond two-way factorizations, we demonstrate how nonnegative tensor factorizations (NTFs) can be used to capture temporal and semantic proximity from documents and thereby enable the tracking of both targeted and latent (previously unknown) discussions or communication patterns. Contributions by Professor Marian Vajtersic and his colleagues on the parallelization of NMF for distributed parallel architectures will be highlighted along with demonstrations of NMF/NTF algorithms for topic (or discussion) detection and tracking using the Enron Email Collection and documents from the Airline Safety Reporting System (ASRS).

Energy Efficient Computing: A New Challenge for Computer Architecture

Arndt Bode
Leibniz Supercomputing Centre and TU Munich, Germany
bode@informatik.tu-muenchen.de

High performance computers have reached the situation that cost of electricity over a 5-year lifetime of the system exceeds the price for the investment. Energy efficient computing covers a large area of activities from building infrastructure, cooling technologies up to the use of software control of the processor clock. The talk gives an overview on the challenges and opportunities of the different approaches. It mentions technologies realized in the context of the German National Supercomputer SuperMUC, a system available for science in PRACE (Partnership for Advanced Computing in Europe).

Parallel Multilevel Schwarz Preconditioners for Elliptic Optimal Control Problems

Daniela di Serafino
Second University of Naples and ICAR-CNR Naples Branch, Italy
daniela.diserafino@unina2.it
Joint work with Alfio Borzì, University of Würzburg, Germany,
and Valentina De Simone, Second University of Naples, Italy

The solution of optimal control problems involving partial differential equations (PDEs) is a computationally demanding task, requiring accurate and efficient numerical methods as well as high-performance computing resources. We focus on the application of multilevel methods for preconditioning linear systems that arise from the discretization of the optimality conditions associated with optimal control problems governed by elliptic PDEs. The interest for such methods is motivated by their optimal convergence properties and their robustness with respect to optimization parameters; furthermore, they have shown good efficiency in the context of parallel computing. We present parallel algebraic multilevel preconditioners based on the additive Schwarz methods and the smoothed aggregation coarsening technique. The classical smoothed aggregation has been modified, obtaining restriction and prolongation operators that are independent of the regularization parameter in the cost functional and are cheaper to compute. The resulting preconditioners have been implemented within the framework of MLD2P4 (MultiLevel Domain Decomposition Parallel Preconditioners Package based on PSBLAS). Results of numerical experiments, showing the effectiveness of our approach, are reported.

Local Meshless Methods for the Numerical Solution of Partial Differential Equations

Roman Trobec
Jožef Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia
roman.trobec@ijs.si

The foundations of the meshless methods will be described including local weak formulation, computational complexity and accuracy, and their potential for parallelization. The moving least squares (MLS) approximation is an essential approach that guarantees the locality of many meshless methods. It is an extension of the classical least squares in the way that just a few nearest discretization points influence the approximated value in a selected point. The locality is implemented by hat-shaped MLS weight functions applied on local support domains. The character of weight functions and the dimension of MLS support domains have a significant impact on the accuracy of the MLS approximation, particularly in case of non-uniform distribution of points. Various methods for the determination of the MLS support domain will be analyzed experimentally: constant radius, sampled and interpolated radius and continuously variable radius with selected numbers of support points, for uniform and non-uniform point distributions.

A typical representative of local meshless methods is meshless local Petrov-Galerkin method with MLS weight functions as test functions (MLPG1). MLS weight functions also determine the quadrature domains and therefore have an important impact on the calculation complexity and accuracy of MLPG1 solutions. We confirmed with experimental analysis that there is only a short interval of MLS support radii that provides acceptable MLPG1 solutions. The accuracy of the MLPG solution was compared with that of standard methods (FDM and FEM) on a unit square and on an irregularly shaped test domain. We further compare computational complexity of the MLPG with FDM and FEM from the user point of view, who need an accurate and fast solution of PDE. Theoretically, MLPG is the most complex of the three methods. Experimental results show that MLPG, with appropriately selected integration order and dimensions of support and quadrature domains, achieves similar accuracy to that of FEM. MLPG remains competitive if human assistance is needed for meshing.

Analysis of Implants and Wood Based Products from CT-Data

Karl Entacher
Salzburg University of Applied Sciences, Puch, Austria
karl.entacher@fh-salzburg.ac.at [2]

In the presentation we demonstrate different scenarios for an analysis of material based on data from computed tomography (CT).

Analysis of Implants Two case studies will be shown. Starting with an analysis of dental implants using a finite element (FE) study which is a common method for biomechanical investigations. A 3D-FE-model simulates the total complex anatomy of the left half maxilla. The model was segmented and designed from CT-data of a patient provided with virtually inserted dental implants. Different boundary conditions and material parameters have been used throughout the study. Central goal is to analyze the influence of an inserted cortical bone graft block, and the sensitivity of varying material parameters for the bone augmentation. In a further FE study we demonstrate a patient specific analysis of special shoulder implants.

Analysis of Wood Based Products Using computed tomography in combination with image analyses has been shown to be a valid technique for determining the microstructure of selected wood based products. In particular images from 3D sub-micrometer-CT and a combination of erosion and dilation procedures using arbitrary shaped structuring elements where used to derive the number of microstructures in classes equivalent to the size of the structuring elements. It has turned out that hardware complexity of the suggested algorithm increases significantly with the size of the used structuring elements. Strategies for speed optimization are shown.

[2]The presented projects are supported by the Austrian Science Fund FWF L526-B05