Computation on partitionned matrices in chemometrics and sensometrics

Mohamed Hanafi[1], David Legland[2]

[1] *Unité de Sensométrie et Chimiométrie, ONIRIS-Nantes*
Rue de la Géraudière,CS 82225 44322 Nantes Cedex 03, France Code postale Ville, Pays
E-mail *: mohamed.hanafi@oniris-nantes.fr*

[2] *Unité BIA-BIBS, INRA Angers-Nantes*
Rue de la Géraudière BP 71627, 44316 Nantes Cedex 3
E-mail *:david.legland@nantes.inra.fr*

Abstract
The present talk introduces a formal framework and a software tool for the computation on partitioned matrices. The interest will be illustrated on the basis of some iterative algorithms encountered in chemometrics or sensometrics.

Keywords: matrix algebra, tensors, multibloc data, multiblock methods, Matlab, R

Résumé en Français
La présente communication introduit un cadre formel et un outil logiciel adapté pour le calcul sur des matrices partitionnées. L'intérêt sera illustré sur la base de quelques algorithmes itératifs rencontrés en chimiométrie et en sensométrie.

Mots-clés : algèbre matricielle, tenseurs, données multiblocs, méthodes multiblocs, Matlab, R

1. Problem statement.

It is well known that many tools in chemometrics (sensometrics) are based on matrix algebra. As an example the Singular Value Decomposition [1] is one of the most popular algorithms coming from matrix algebra and frequently used in chemometrics (sensometrics).
The connection between chemometrics (sensometrics) and matrix algebra is of major importance. Indeed, the matrix algebra contributed not only to consolidate the mathematical basis of the chemometric (sensometric) methods, but also to provide computational and geometric tools that facilitated the design of dedicated software for these methods. Just remember that the acronym of the most used software in chemometrics MatLab means Matrix Laboratory.

Matrices and tensors do not cover the full range of data acquired in chemometrics (sensometrics). Indeed, the continuous development of sensors allows today to acquire various kinds of measurement on one or more groups of individuals [2]. Mathematically, these data can be conceptualized as partitioned matrices.

Both Matrix algebra [1] and tensor algebra [3] appear inadequate to manipulate partitioned matrices. New specific rules for computation on partitioned matrices are needed. Indeed, it is always possible to manipulate partitioned matrices as simple matrices without considering the associated partition. This approach is especially inappropriate when the partition is an essential component of the problem to be solved. This is the case for the so called multiblock methods [4,5]. Furthermore, the partitioned matrices cannot be represented as tensors since the modes of blocks are all not identical.

2. Contributions

The aim of this talk is twofold. The first is conceptual, it consists to introduce a more appropriate vocabulary for partitioned matrices and to propose a standardization of notations used for these entities. The fundamental aspects, in particular different types of products between partitioned matrices, will be introduced and discussed.

The second aim of the present talk is practical. It consists to describe a computational kernel for handling partitioned matrices under MATLAB and R called respectively BlockMatrix [6] and BlockBerry [7]. The practical interest of this kernel lies in the fast prototyping algorithms involving partitioned matrices. This will be illustrated via iterative algorithms for multiblock methods [5].

References

[1] Golub, G.H.,. Van Loan, C.F(1996). Matrix Computations. Johns Hopkins University Press Baltimore. 3rd Edition. ISBN:0-8018-5414-8.

[2] Hanafi, M., G. Mazerolles, E. Dufour, E.M. Qannari (2006). Common components and specific weight analysis and multiple Co-inertia analysis applied to the coupling of several measurement techniques. Journal of Chemometrics, 20: 172-183.

[3] Kolda, T.G., Bader,W.B., 2009. Tensor decomposition and applications. SIAM Rev. 51, 455–500.

[4] Hanafi M., A. Kohler, E.M. Qannari (2011). Shedding new light on Hierarchical Principal Component Analysis. Journal of Chemometrics, 24 (11):703-709.

[5] Hanafi, M., H. A. L. Kiers (2006). Analysis of K sets of Data with differential Emphasis on agreement between and within sets. Computational Statistics and Data Analysis, 51 (3): 1491-1508.

[6] https://github.com/dlegland/BlockMatrixToolbox.

[7] https://github.com/gastonstat/blockberry.