CDC special session on "Low-rank approximation"

Ivan Markovsky[1] and Konstantin Usevich[2]

1 — Electrical Engineering Department, Vrije Universiteit Brussel, Belgium
2 — Centre de Recherche en Automatique de Nancy, Université de Lorraine, CNRS, France

Abstract

Low-rank approximations play an important role in systems theory, signal processing, and machine learning. The problems of model reduction and system identification can be posed and solved as a low-rank approximation problem for structured matrices. On the other hand, signal source separation, nonlinear system identification, and multidimensional signal processing/data analysis lead to problems of low-rank tensor factorization. The proposed invited session is focused on theoretical and algorithmic aspects of matrix and tensor low-rank approximations, with applications in system identification and machine learning.

Motivation

The problem of approximating a matrix by a low-rank matrix has been extensively studied and well-understood. Low-rank approximations are widely used in signal processing, systems theory and machine learning as a tool for dimensionality reduction, feature extraction, and classification. The optimal solution can be obtained from the truncated singular value decomposition (SVD) [5].

However, in applications of systems, control and signal processing that involve dynamical systems, the data are structured [11, 10]. This structure should be preserved in the approximation. This leads to structured low-rank approximation. A classical example in systems theory is a Hankel structure, where low-rank approximation is used for model reduction and sum-of-damped-exponentials modelling. Another type of structure results from multi dimensionality of the data, e.g., video sequences (two spatial coordinates and time) [4].

The main challenge in solving structured and tensor low-rank approximations is that the underlying optimization problem is, in general, non-convex. In addition, there are other constraints, such as nonnegativity of the factorization, fixed values, and semidefinite low-rank matrices. Finally, the matrices/tensors may contain missing data.

For matrix structured low-rank approximation, the most popular approaches include local optimization methods [13]. For these methods, a good initial approximation is needed, which can be obtained from relaxation methods, such as unstructured SVD-based relaxations (subspace-based methods) or convex relaxations (mainly utilizing the nuclear norm heuristic) [3, 9, 12]. More recently, optimization on the manifold of low-rank matrices (using the Riemannian geometry on manifold) [1] and stochastic optimization algorithms [8] are used.

As for tensor decomposition, there is a variety of possibilities [7, 2]. In the sense of multilinear rank (generalization of column and row rank of a matrix to tensors), a generalization of the SVD is the higher-order singular value decomposition, (HOSVD), also called Tucker decomposition. The decomposition of a tensor in rank-1 terms (outer products of vectors), called canonical decomposition (or Parallel factor decomposition, PARAFAC), is another generalization of the SVD. Block term decompositions were proposed as the unifying framework for the mentioned decompositions. Other recent tensor network decompositions [6], which include tensor train and hierarchical Tucker decompositions, are meant for tensors of high order and large dimensions.

Summary of the papers

The session consists of six papers P1–6. The first four papers consider system identification problems while the last two papers consider state estimation and machine learning problems. From the identification problems, papers P1–3 deal with various aspects of linear time-invariant system identification: combining time and frequency domain data (P1), Frisch-Kalman problem (P2), and cointegrated systems (P3). Paper P4 considers a multiple-input-multiple-output Volterra system identification problem. Convex relaxations of the original rank minimization problem are used in papers P1 and P2, while a nonlinear local optimization approach is used in paper P3. The main result of paper P4 is reformulation of the Volterra identification problem into an extended Kalman filtering problem and using low-rank tensor networks for its solution. Paper P5 shows how low-rank properties can replace the traditional model-based observability property in state estimation problems. The problem considered in paper P6 is low-rank factorization of hyperbolic embeddings. This problem is used in machine learning and natural language processing. The main result of paper P6 is a computationally efficient algorithm for manifold optimization.

P1: A convex optimization approach to finding low rank mixed time/frequency domain interpolants with applications to control oriented identification by Rajiv Singh (The MathWorks, USA) and Mario Sznaier (Northeastern University, USA)

We consider the problem of finding a low order stable rational transfer function that interpolates a set of given noisy time and frequency domain data points. Our main result shows that exploiting results from rational interpolation theory allows for recasting this problem as minimizing the rank of a matrix constructed from the frequency domain data (the Loewner matrix), subject to a semidefinite constraints that enforces stability and consistency between the time and frequency domain data. These results are applied to a practical problem: identifying a system from noisy measurements of its time and frequency responses. As shown in the paper, the proposed method is able to obtain stable low order models using substantially less data points that those required by existing methods such as Hankel rank minimization.

P2: A convex approach to Frisch-Kalman problem by Di Zhao, Anders Rantzer, and Li Qiu (Lund University, Sweden)

This paper proposes a convex approach to the Frisch-Kalman problem that identifies the linear relations among variables from noisy observations. The problem was proposed by Ragnar Frisch in 1930s, and was promoted and further developed by Rudolf Kalman later in 1980s. It is essentially a rank minimization problem with convex constraints. Regarding this problem, analytical results and heuristic methods have been pursued over a half century. The proposed convex method in this paper is shown to be accurate and demonstrated to outperform several commonly adopted heuristics when the noisy components are well-bounded with respect to the true variables.

P3: Low-rank identification in the context of cointegrated systems by Ben Hanzon and M. Alqurashi (University College Cork, Ireland)

The concept of cointegration has been widely applied in econometrics since the seminal work of Granger and Engle. This paper focuses on a case of a continuous-time state space model with a finite number of observations taken at irregularly-spaced time points. We propose a parameterization and estimation method that translates the cointegration property into a low-rank constraint on a resulting likelihood optimization. First, we partially optimize the criterion using the likelihood function. The resulting criterion is a function of eigenvalues of a matrix due to the low-rank constraint. This raises a problem of calculating the derivatives of the parameterized matrix eigenvalues. We address this problem by applying the envelope theorem. We then derive a gradient method to optimize the remaining model parameters. The approach proposed in this paper has the advantage of explicitly incorporating the low-rank requirement while addressing the eigenvalue derivative issue through an application of the envelope theorem. Further details about the application to the particular cointegration model will be provided.

P4: Extended Kalman filtering with low-rank Tensor Networks for MIMO Volterra system identification by Ching-Yun Ko (The University of Hong Kong), Kim Batselier (TU-Delft, The Netherlands), and Ngai Wong (The University of Hong Kong)

This article reformulates the multiple-input-multiple-output Volterra system identification problem as an extended Kalman filtering problem. This reformulation has two advantages. First, it results in a simplification of the solution compared to the Tensor Network Kalman filter as no tensor filtering equations are required anymore. The second advantage is that the reformulation allows to model correlations between the parameters of different multiple-input-single-output Volterra systems, which might lead to better accuracy. The curse of dimensionality in the exponentially large parameter vector and covariance matrix is lifted through the use of low-rank tensor networks.

P5: An observer of a high-dimensional, low-rank state space by Kristiaan Pelckmans (Uppsala University, Sweden)

Assume the states evolve according to an unknown LTI system, but that they are measured through a known observation matrix. If the state is known to fit into a given ’ambient’ space, and we know that it varies within a low-rank space, then we can actually recover (under certain conditions) the states from measurements. This result indicates how low-rank properties could replace traditional model-based controllability/observability restrictions in state observers. This problem is motivated by a case study of a brain signal (EEG) reader using 8 sensors, but where the states take values in $O(10^7)$ dimensions.

P6: Low-rank hyperbolic embeddings by Bamdev Mishra (Microsoft, India)

The hyperbolic manifold is a smooth manifold of negative constant curvature. While the hyperbolic manifold is well-studied in the literature, it has gained interest in the machine learning and natural language processing communities lately due to its usefulness in modeling continuous hierarchies. Tasks with hierarchical structures are ubiquitous in those fields and a general interest there is to learning hyperbolic representations or embeddings of such tasks. Additionally, these embeddings of related tasks may also share a low-rank subspace. In this work, we propose to learn hyperbolic embeddings such that they also lie in a low-dimensional subspace. In particular, we consider the problem of learning a low-rank factorization of hyperbolic embeddings. We cast these problems as manifold optimization problems and propose computationally efficient algorithms. Empirical results illustrate the benefits of the proposed low-rank hyperbolic modeling approach.

References

[1] P.-A. Absil et al. “A Grassmann–Rayleigh Quotient Iteration for Computing Invariant Subspaces”. In: *SIAM Review* 44.1 (2002), pp. 57–73.

[2] L. De Lathauwer. “Signal processing based on multilinear algebra”. PhD thesis. K.U.Leuven, ESAT-SISTA, 1999.

[3] M. Fazel. “Matrix Rank Minimization with Applications”. PhD thesis. Elec. Eng. Dept., Stanford University, 2002.

[4] Y. Fu, ed. *Low-Rank and Sparse Modeling for Visual Analysis*. Springer, 2014.

[5] G. Golub and C. Van Loan. *Matrix Computations*. Third. Johns Hopkins University Press, 1996.

[6] Lars Grasedyck, Daniel Kressner, and Christine Tobler. “A literature survey of low-rank tensor approximation techniques”. In: *GAMM-Mitteilungen* 36.1 (2013), pp. 53–78.

[7] T. Kolda and B. Bader. “Tensor decompositions and applications”. In: *SIAM review* 51.3 (2009), pp. 455–500.

[8] E. Liberty et al. “Randomized algorithms for the low-rank approximation of matrices”. In: *Proceedings of the National Academy of Sciences* 104 (2007), pp. 20167–20172.

[9] Z. Liu and L. Vandenberghe. “Interior-point method for nuclear norm approximation with application to system identification”. In: *SIAM J. Matrix Anal. Appl.* 31.3 (2009), pp. 1235–1256.

[10] I. Markovsky. *Low-Rank Approximation: Algorithms, Implementation, Applications*. Second. Springer, 2019.

[11] I. Markovsky. “Structured low-rank approximation and its applications”. In: *Automatica* 44.4 (2008), pp. 891–909.

[12] M. Sznaier, M. Ayazoglu, and T. Inanc. “Fast Structured Nuclear Norm Minimization With Applications to Set Membership Systems Identification”. In: *IEEE Trans. Automat. Contr.* 59.10 (2014), pp. 2837–2842.

[13] K. Usevich and I. Markovsky. “Variable projection for affinely structured low-rank approximation in weighted 2-norms”. In: *J. Comput. Appl. Math.* 272 (2014), pp. 430–448.