21st International Symposium on
Mathematical Theory of Networks and Systems
July 7-11, 2014. Groningen, The Netherlands

Numerical methods in higher dimensions using tensor factorizations

Ivan Oseledets[1]

***Abstract*— Numerical methods in higher dimensions using tensor factorizations I this talk I will collect recent advances in the solution of high-dimensional problems in different application areas: chemistry, biology, mathematics. The language of low-rank factorization gives a unified view on different algorithms for the solution of seemingly diverse and unconnected problems. Typical applications include quantum chemistry, chemical master equation in biology, latent variable models in data analyses.**

I. INTRODUCTION

Numerical tensor methods play increasingly important role in modern computational techniques and applications. A tensor is simply a **multidimensional array**. We will denote elements of an array $\mathbf{A}$ as $A(i_1, \ldots, i_d)$. A standard review for tensors and their decomposition is the paper by Kolda and Bader [1]. However, it is already outdated, since it does not cover many important concepts developed and rediscovered in the recent years. Several reviews have been put together to fill in this gap, see [2], a literature survey [3], a book [4]. The field is still developing quickly, with new insights coming at fast rate. The results in multilinear algebra community came from a general dissatisfaction with the established formats - *canonical polyadic* and *Tucker formats*. Both formats manifest different ways of **separation of variables** as a model reduction tool. The canonical format for a tensor $\mathbf{A}$ is defined as a representation of the form

$$A(i_1, \ldots, i_d) = \sum_{\alpha=1}^{r} U_1(i_1, \alpha) \ldots U_d(i_d, \alpha),$$

i.e. representing a tensor with a sum of r rank-1 tensors. Unfortunately, despite its simplicity, there are no robust algorithms to compute the decomposition due to the ill-conditioning and many local minima of the related optimization problem. The Tucker format alleviates this difficulty by introducing additional parameters:

$$A(i_1, \ldots, i_d) = \sum_{\alpha_1, \ldots, \alpha_d} G(\alpha_1, \ldots, \alpha_d) U_1(i_1, \alpha_1) \ldots U_d(i_d, \alpha_d).$$

This decomposition is stable in the sense that the optimal approximation always exists, and the quasi-optimal decomposition can be computed via High Order Singular Value Decomposition [5], [6]. The curse of decomposition still remains in the Tucker format. The Tree-Tucker [7] and Hierarchical Tucker [8] decompositions showed that it is possible to design a stable SVD-based format that does not intrinsically have the curse of dimensionality. Despite their similarity, these two formats were different and used different ideas of reducing the tensor to matrices, where the SVD can be used. Then it was realized that Tree-Tucker format is equivalent up to the permutation of indices to the Tensor Train (TT)-format, which, due to its simplicity and effectiveness became an important tool for solving high-dimensional problems. The TT-format has the form

$$A(i_1, \ldots, i_d) = G_1(i_1) G_2(i_2) \ldots G_d(i_d). \qquad (1)$$

where $G_k(i_k)$ is a $r_{k-1} \times r_k$ matrix for each fixed i_k, and $r_0 = r_d = 1$. It can be computed in a stable way using SVD-based techniques [9], [10]. Good news about the novel tensor formats is not only the representation formula, but the associated efficient and robust algorithms for solving important problems.

It is important to note, that the TT and HT formats were in fact known for a long time in other communities under different names. The Matrix Product State (MPS) representation [11], [12], [13] has been known for a while in statistical physics and solid state physics. It was used to model wavefunctions of spin systems. A comprehensive reviews are available, see [14]. Once this connection was introduced by T. Huckle in 2010 in the GAMM Seminar Leipzig, the mathematical community benefited a lot from a well-established algorithms in the field. One of the most prominent appoarches for computing with MPS, Density Matrix Renormalization Group (DMRG), introduced by White [15] was later adapted to the mathematical language, and even generalized to other problem which are not present in physics, like multidimensional interpolation [16], [17].

II. APPLICATIONS

High-dimensional problems appear in many applications. Tensor-related techniques were developed in many of them and there are several different languages. A general principle can be formulated as "if the high-dimensional problem is interesting, there is an algorithm for its solution". The problem is that such algorithms are problem-specific, and the development of new algorithms takes time. However, many approaches have common features. Novel tensor decompositions are a good candidate for such "mathematical apparatus". Typical applications include:

1) Quantum chemisty: the Schrödinger equation is a multidimensional PDE which is notoriously difficult. The established method can be considered as a separation

*This work was support by Dmitry Zimin Dynasty Foundation grant, by the grant of the Yandex company, by RFBR grant 12-01-00565-a

[1]Ivan Oseledets is with Skolkovo Institute of Science and Technology, Novaya Str. 100, 143025 Skolkovo, Moscow Region, Russia i.oseledets@skoltech.ru

ISBN: 978-90-367-6321-9

of variables with anti-symmetry constraints. Approximation of potential energy surfaces (PES) involves approximation of multivariate function. To summarize, quantum chemistry is a big source of high-dimensional problems.

2) Parametric and stochastic problem in engineering sciences. Several successful applications of tensor techniques for parametric problems have already been demonstrated [18], [19]. Discretization of parameter space naturally leads to high dimensionsional problems.
3) Data compression and data mining. Another field that shows deep connection with tensor methods are the *graphical models* and *latent variable models*. The multivariate function is the joint probability distribution $p(x_1, \dots, x_d)$, where $x_1, \dots, x_d$ are some random variables. Algebraically, the well-known hidden Markov Chain (HMC) is very similar to the TT-format, and common problems and methods can be derived. One of the successful applications of tensor methods to this class of problems was described in [20]. There is a deep relation between tensors and wavelets throught the **Quantized TT** format [21], [22]. This relation is based on the wavelet tensor train (WTT) representation [23]. It is also worth note mention the tensor properties of the wavelet matrices [24].

There are more applications to come. For example an important field to study is the relation between novel tensor formats and convolutional neural networks, which have, in some sense, similar structure but again different focus and purpose.

III. Algorithms

There are several basic problems that need to be solved. Typically, a high-dimensional problem reduces to a certain combination of those. They include:

1) Solution of linear systems $Ax = f$
2) Solution of eigenvalue problems $Ax = \lambda x$
3) Solution of non-stationary problems $\frac{dx}{dt} = Ax + f$
4) Interpolation of a tensor from its entries

At the moment, there are efficient algorithms for all of these tasks, see [25], [26] for linear systems, [27] for the eigenvalue problems (including block eigenvalue problems) [28], [29] for the dynamical problems and [30], [31], [32] for the interpolation. There is still a lot of space for the improvement, but these algorithms already work quite well. The next steps would be the development of approaches for the selection of right variables to separate, maybe some tools for selecting the right tree / order of the indices and so on. The main question is then how to select variables in such a way that they separate well.

IV. Conclusion

In this paper we presented a brief report on the current state of art for the numerical tensor methods. There is still a lot to be done to find the interconnection between these techniques and approaches developed in other fields in physics, chemistry and data analysis. This will help to develop a set of universal tools for solving high-dimensional problems in the efficient way.

References

[1] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Review*, vol. 51, no. 3, pp. 455–500, 2009.
[2] B. N. Khoromskij, "Tensor-structured numerical methods in scientific computing: Survey on recent advances," *Chemometr. Intell. Lab. Syst.*, vol. 110, no. 1, pp. 1–19, 2012.
[3] L. Grasedyck, D. Kressner, and C. Tobler, "A literature survey of low-rank tensor approximation techniques," arXiv preprint 1302.7121, 2013. [Online]. Available: http://arxiv.org/abs/1302.7121
[4] W. Hackbusch, *Tensor spaces and numerical tensor calculus*. Springer–Verlag, Berlin, 2012.
[5] L. de Lathauwer, B. de Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 21, pp. 1253–1278, 2000.
[6] ——, "On best rank-1 and rank-$(R_1, R_2, ..., R_N)$ approximation of high-order tensors," *SIAM J. Matrix Anal. Appl.*, vol. 21, pp. 1324–1342, 2000.
[7] I. V. Oseledets and E. E. Tyrtyshnikov, "Breaking the curse of dimensionality, or how to use SVD in many dimensions," *SIAM J. Sci. Comput.*, vol. 31, no. 5, pp. 3744–3759, 2009.
[8] W. Hackbusch and S. Kühn, "A new scheme for the tensor representation," *J. Fourier Anal. Appl.*, vol. 15, no. 5, pp. 706–722, 2009.
[9] I. V. Oseledets, "A new tensor decomposition," *Doklady Math.*, vol. 80, no. 1, pp. 495–496, 2009.
[10] ——, "Tensor-train decomposition," *SIAM J. Sci. Comput.*, vol. 33, no. 5, pp. 2295–2317, 2011.
[11] A. Klümper, A. Schadschneider, and J. Zittartz, "Matrix product ground states for one-dimensional spin-1 quantum antiferromagnets," *Europhys. Lett.*, vol. 24, no. 4, pp. 293–297, 1993.
[12] M. Fannes, B. Nachtergaele, and R. Werner, "Finitely correlated states on quantum spin chains," *Communications in Mathematical Physics*, vol. 144, no. 3, pp. 443–490, 1992.
[13] I. Affleck, T. Kennedy, E. H. Lieb, and H. Tasaki, "Rigorous results on valence-bond ground states in antiferromagnets," *Phys. Rev. Lett.*, vol. 59, no. 7, pp. 799–802, 1987.
[14] U. Schollwöck, "The density-matrix renormalization group in the age of matrix product states," *Annals of Physics*, vol. 326, no. 1, pp. 96–192, 2011.
[15] S. R. White, "Density-matrix algorithms for quantum renormalization groups," *Phys. Rev. B*, vol. 48, no. 14, pp. 10 345–10 356, 1993.
[16] D. V. Savostyanov and I. V. Oseledets, "Fast adaptive interpolation of multi-dimensional arrays in tensor train format," in *Proceedings of 7th International Workshop on Multidimensional Systems (nDS)*. IEEE, 2011.
[17] S. Holtz, T. Rohwedder, and R. Schneider, "The alternating linear scheme for tensor optimization in the tensor train format," *SIAM J. Sci. Comput.*, vol. 34, no. 2, pp. A683–A713, 2012.
[18] B. N. Khoromskij and I. V. Oseledets, "Quantics-TT collocation approximation of parameter-dependent and stochastic elliptic PDEs," *Comput. Meth. Appl. Math.*, vol. 10, no. 4, pp. 376–394, 2010.
[19] D. Kressner and C. Tobler, "Low-rank tensor Krylov subspace methods for parametrized linear systems," *SIAM J. Matrix Anal. Appl.*, vol. 32, no. 4, pp. 273–290, 2011.
[20] L. Song, M. Ishteva, A. Parikh, E. Xing, and H. Park, "Hierarchical tensor decomposition of latent tree graphical models," in *Proceedings of the 30th International Conference on Machine Learning (ICML-13)*, 2013, pp. 334–342.
[21] B. N. Khoromskij, "$\mathcal{O}(d \log n)$–Quantics approximation of N–d tensors in high-dimensional numerical modeling," *Constr. Appr.*, vol. 34, no. 2, pp. 257–280, 2011.
[22] I. V. Oseledets, "Approximation of $2^d \times 2^d$ matrices using tensor decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 31, no. 4, pp. 2130–2145, 2010.
[23] I. V. Oseledets and E. E. Tyrtyshnikov, "Algebraic wavelet transform via quantics tensor train decomposition," *SIAM J. Sci. Comput.*, vol. 33, no. 3, pp. 1315–1328, 2011.

[24] V. A. Kazeev and I. V. Oseledets, "The tensor structure of a class of adaptive algebraic wavelet transforms," ETH SAM, Zürich, Preprint 2013-28, 2013. [Online]. Available: http://www.sam.math.ethz.ch/sam_reports/reports_final/reports2013/2013-28.pdf

[25] S. V. Dolgov and I. V. Oseledets, "Solution of linear systems and matrix inversion in the TT-format," *SIAM J. Sci. Comput.*, vol. 34, no. 5, pp. A2718–A2739, 2012.

[26] S. V. Dolgov and D. V. Savostyanov, "Alternating minimal energy methods for linear systems in higher dimensions. Part I: SPD systems," arXiv preprint 1301.6068, 2013. [Online]. Available: http://arxiv.org/abs/1301.6068

[27] S. V. Dolgov, B. N. Khoromskij, I. V. Oseledets, and D. V. Savostyanov, "Computation of extreme eigenvalues in higher dimensions using block tensor train format," arXiv preprint 1306.2269, 2013. [Online]. Available: http://arxiv.org/abs/1306.2269

[28] C. Lubich, T. Rohwedder, R. Schneider, and B. Vandreycken, "Dynamical approximation of hierarchical Tucker and tensor-train tensors," University of Tübingen, Tech. Rep., 2012.

[29] C. Lubich and I. V. Oseledets, "A projector-splitting integrator for dynamical low-rank approximation," *BIT*, pp. 1–18, 2013.

[30] I. V. Oseledets and E. E. Tyrtyshnikov, "TT-cross approximation for multidimensional arrays," *Linear Algebra Appl.*, vol. 432, no. 1, pp. 70–88, 2010.

[31] D. V. Savostyanov, "Quasioptimality of maximum–volume cross interpolation of tensors," arXiv preprint 1305.1818, 2013. [Online]. Available: http://arxiv.org/abs/1305.1818

[32] J. Ballani, L. Grasedyck, and M. Kluge, "Black box approximation of tensors in hierarchical Tucker format," *Linear Alg. Appl.*, vol. 428, pp. 639–657, 2013.