Variational calculus with sums of elementary tensors of fixed rank

Mike Espig · Wolfgang Hackbusch · Thorsten Rohwedder · Reinhold Schneider

Received: 8 August 2009 / Revised: 22 October 2011
© Springer-Verlag 2012

Abstract In this article we introduce a calculus of variations for sums of elementary tensors and apply it to functionals of practical interest. The survey provides all necessary ingredients for applying minimization methods in a general setting. The important cases of target functionals which are linear and quadratic with respect to the tensor product are discussed, and combinations of these functionals are presented in detail. As an example, we consider the solution of a linear system in structured tensor format. Moreover, we discuss the solution of an eigenvalue problem with sums of elementary tensors. This example can be viewed as a prototype of a constrained minimization problem. For the numerical treatment, we suggest a method which has the same order of complexity as the popular alternating least square algorithm and demonstrate the rate of convergence in numerical tests.

Mathematics Subject Classification 15A69 · 90C06 · 65F10

1 Introduction

Approximation of solutions of high dimensional partial differential or integral equations by low rank tensors has yielded promising results, see e.g. [1,2,11,12,14,15]. A tensor \( u \in \mathbb{R}^{n^d} \) of order \( d \) requires in general a storage complexity of \( n^d \). If \( u \) can be approximated by a low rank tensor

M. Espig · W. Hackbusch
Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany
e-mail: Mike.Espig@mis.mpg.de

T. Rohwedder · R. Schneider
Technical University Berlin, Berlin, Germany
the memory requirement reduces to $d r n$, and the complexity of algebraic operations grows only linearly with respect to the order $d$. However, when using iterative methods for computing a low-rank tensor, one usually has to face the problem that the involved algebraic operations increase the tensor rank in each iteration step. To overcome this issue, efficient recompression methods have been developed in [1–3,7,9] to approximate a given sum of elementary tensors by low rank tensors. Moreover, the convergence of such approximate iterations is known, see [16] for an analysis. The subject of this article is in contrast to this approach. We will show that the representation of a tensor in low rank format allows in many cases of practical interest a direct optimization procedure on the set of tensors of a fixed rank $r$. Thus, we will solve the original problem directly in the low tensor rank format instead of solving a high dimensional problem indirectly by the use of approximate iterative schemes. This approach has some advantages: We can be sure that the solution is at least locally optimal with respect to the problem dependent functional, and we circumvent numerical problems that may arise during the compression step of approximate iterations even if the original task lacks this kind of approximation problems. An example where such problems may occur is the second numerical experiment discussed in this article, an eigenvalue problem for which we know a priori that there exists a low rank solution. If we would apply the indirect iterative methods discussed above, we have to approximate all iterands during the iterative process, while it is unclear that an iterand can be well approximated by low rank tensors.

In this work, we give some examples for a successful application of the above direct approach when combined with an accelerated gradient (AG) algorithm, which we consider a first step towards the application of the direct optimization concept to the canonical format. It suggests that the approach might be competitive with optimization algorithms for alternative tensor formats as the Tucker Format [30] or the HT [17] format (with the TT format [26–28] as special case), for which the development is also still in progress: For the Tucker format, one obtains an optimization problem on a suitable Grassmann manifold $G$, an approach used on approximation problems in [6,20,29]. Alternatively, the treatment of differential equations may be formulated via a stable characterization of the tangent space [22], and optimization problems may in this calculus be treated by following the gradient flow of the functional on $G$ (see [22] for the general approach). To our knowledge, there are no practical results available yet; nevertheless, the satisfactory convergence behavior of the related ALS procedure [21,24] gives rise to the hope that the Tucker format also provides a sensible alternative when the space dimension is small. As per the more recent TT format, a part of the authors recently obtained some promising practical results for optimization problems in [18]; from the theoretical side, this is complemented by the fact that the TT and HT format share many of the favorable properties of the Tucker format [10,19] while often allowing for a representation consuming less resources. The considerations in this article are closely related to the pioneer work of Beylkin and Mohlenkamp [1–3].