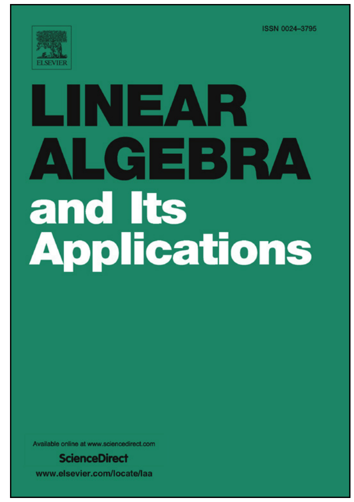
Accepted Manuscript



Tangent Cones to Tensor Train Varieties

Benjamin Kutschan

PII: S0024-3795(18)30018-1
DOI: https://doi.org/10.1016/j.laa.2018.01.012
Reference: LAA 14435

To appear in: *Linear Algebra and its Applications*

Received date: 6 June 2017
Accepted date: 8 January 2018

Please cite this article in press as: B. Kutschan, Tangent Cones to Tensor Train Varieties, *Linear Algebra Appl.* (2018), https://doi.org/10.1016/j.laa.2018.01.012

Tangent Cones to Tensor Train Varieties

Benjamin Kutschan

TU Berlin, Straße des 17. Juni 136, MA5-3, 10623 Germany

Abstract

As already done for the matrix case in [1, p.256] [2, Thm. 6.1, p.1872] and [3, Thm. 3.2], we give a parametrization of the Bouligand tangent cone of the variety of tensors of bounded tensor train (TT) rank. We discuss how the proof generalizes to any binary hierarchical format. The parametrization can be rewritten as an orthogonal sum of TT tensors. Its retraction onto the variety is particularly easy to compose. We also give an implicit description of the tangent cone as the solution of a system of polynomial equations.

Keywords: Riemannian optimization, tangent cone, algebraic variety, low-rank matrices, hierarchical tensor decomposition, tensor train, multilinear algebra, 15A69, 65K10, 62Fxx

1. Introduction

For the purposes of this work we define a tensor as a basis representation of an element from a finite-dimensional tensor space.

Definition 1.1. A *tensor* A is an element of the tensor space $\mathbb{R}^{n_1 \times \dots \times n_d}$ where d is called the *order* and n_i is called the *dimension* (in the direction) of order i.

Tensor spaces are usually so high-dimensional, that it is impossible to save all entries of a tensor directly in computer memory (for example a tensor in $\mathbb{R}^{10 \times \dots \times 10}$ has 10^d entries, if the 10 appears d times in the exponent). The tensor train and hierarchical Tucker varieties are parametrizable subvarieties of tensor spaces. They can be chosen such that their elements fit in memory. Riemannian optimization techniques aim to find a local minimum of a real function $f : \mathbb{R}^{n_1 \times \dots \times n_d} \to \mathbb{R}$ defined on the tensor space. They have been successfully applied by [3] in the special case of low-rank matrix varieties. Riemannian optimization starts with some point on the variety. From this point it generates a new one that also lies on the variety but is better in the sense of the cost functional f. The tangent cone at a point is the set of all possible directions of curves that stay on the variety. So the first step in a Riemannian optimization algorithms is to find a direction in the tangent cone, along which f decreases. It is thus essential to know the structure of the tangent cone.

Email address: `kutschan@math.tu-berlin.de` (Benjamin Kutschan)