MascotNum2019 conference - Principal Component Analysis and "boosted" weighted least-squares method for training tree tensor networks

CECILE HABERSTICH
Ecole Centrale de Nantes

Supervisor(s): Anthony Nouy (Ecole Centrale de Nantes), Guillaume Perrin (CEA/DAM/DIF, F-91297, Arpajon)

Ph.D. expected duration: 2017 - 2020

Adress: CEA/DAM/DIF, F-91297, Arpajon

Email: cecile.haberstich@ec-nantes.fr

Abstract: One of the most challenging tasks in computational science is the approximation of high dimensional functions. Most of the time, only a few information on the functions is available, and approximating high-dimensional functions requires exploiting low-dimensional structures of these functions.
In this work, the approximation of a function u is built using point evaluations of the function, where the evaluations are selected adaptively. Such problems are encountered when the function represents the output of a black-box computer code, a system or a physical experiment for a given value of a set of input variables.
A multivariate function $u(x_1, ..., x_d)$ defined on a product set $\mathcal{X} = \mathcal{X}_1 \times ... \times \mathcal{X}_d$ can be identified with a tensor of order d.
Here, we present an algorithm for the construction of an approximation of a function u in tree-based tensor format (tree tensor networks whose graphs are dimension partition trees). A low-order tensor v_α, seen as a vector-valued map, is associated to each node α of the dimension partition tree T, and this set of tensors totally parameterizes the approximation.

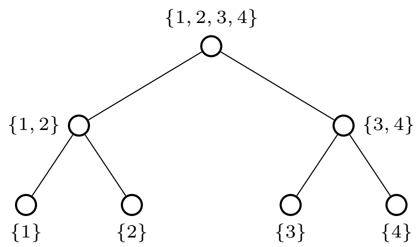

Figure 1: Example of a dimension partition tree T over $D = \{1, 2, 3, 4\}$

For example, an approximation v associated with the tree of the Figure 1 takes the form:

$$v = v_{1,2,3,4}(v_{1,2}(v_1(\phi_1(x_1)), v_2(\phi_2(x_2)), v_{3,4}(v_3(\phi_3(x_3)), v_4(\phi_4(x_4)))))$$

where the $\phi_\nu : \mathcal{X}_\nu \to \mathbb{R}^{n_\nu}, \nu \in \{1, 2, 3, 4\}$ are the feature maps.
The algorithm relies on an extension of principal component analysis (PCA) to multivariate functions in order to estimate the tensors. In practice, PCA is realized on sample-based projections of the function u, using interpolation or least-squares regression.
To provide a stable projection, least-squares regression usually requires a high number of evaluations of u, which is not affordable in our context. This number of evaluations can be decreased thanks to a so-called "boosted" weighted least-squares method. This method combines an optimal weighted least-squares method proposed in [1] and a re-sampling technique. With a particular choice of weights and samples and through re-sampling, an approximation error of the order of

the best approximation error is guaranteed using a moderate number of samples, of the order of the dimension of the approximation space.
We use this methodology in our algorithm and will compare it with strategies using standard least-squares method or interpolation (as proposed in [2]).

References

[1] A. Cohen and G. Migliorati. Optimal weighted least-squares methods. *SMAI Journal of Computational Mathematics*, 3:181203, 2017.

[2] A. Nouy. Higher-order principal component analysis for the approximation of tensors in tree-based low rank formats. *Numerische Mathematik*, 2019.

Short biography – I graduated from Centrale Nantes with a specialization in applied mathematics and from the Technical University of Munich with a specialization in computational mechanics in 2017. I began a PhD thesis whose subject is "Low-rank approximation methods for complex uncertainty quantification problems". This thesis is a joint work between Centrale Nantes and the CEA DAM.