Spectral tensor-train decomposition for low-rank surrogate models

Daniele Bigoni¹, Allan P. Engsig-Karup¹, Youssef M. Marzouk²

¹Department of Applied Mathematics and Computer Science, Technical University of Denmark
²Department of Aeronautics and Astronautics, Massachusetts Institute of Technology

Corresponding author: dabi@dtu.dk

Introduction

The construction of surrogate models is very important as a mean of acceleration in computational methods for uncertainty quantification (UQ). When the forward model is particularly expensive compared to the accuracy loss due to the use of a surrogate – as for example in computational fluid dynamics (CFD) – the latter can be used for the forward propagation of uncertainty [7] and the solution of inference problems [4].

Problem setting

We consider \( f \in L^2(\mathbb{R}^d) \), where \( d \gg 1 \) and assume \( f \) is a computationally expensive function. Let \( \xi \in [0, b) \) be random variables entering the formulation of a parametric problem. In the context of UQ, we might want to:

- Compute relevant statistics
- Inquire the sensitivity of \( f \) to \( \xi \)
- Infer the distribution of \( \xi \)

In most real problems, these goals require an high number of evaluations of \( f \). Often the construction of the surrogate and its evaluation in place of the original \( f \) provides a good payoff.

Tensor-train decomposition

Let \( f \) be evaluated at all points on a tensor grid \( \mathcal{X} = \bigotimes_{j=1}^{d} x_j \), where \( x_j = (r_j)^0_1 \) for \( j \in [1, d] \). Let \( \mathcal{A} = f(\mathcal{X}) \).

Discrete tensor-train approximation [5]

Let \( r = (r_1, \ldots, r_d) \), let \( \mathcal{A}_{TT} \) be s.t.

\[
\mathcal{A}_{TT} = \bigotimes_{i=1}^{d} G_i(a_i, \alpha_i, \sigma_i) = \bigotimes_{i=1}^{d} \mathcal{G}_i(\alpha_i, \sigma_i)
\]

The construction can be built through the evaluation of \( f \) on the most important fibers (Fig. 1), detected using the TT-Cross algorithm [6]. For example, let \( f(x,y) = \frac{1}{\pi} \arctan(xy) \).

Figure 1: TT-cross

Functional TT-decomposition

Using the spectral theory on (non-symmetric) Hilbert-Schmidt kernels, we can construct a functional counterpart of the discrete TT-approximation.

Functional tensor-train approximation [1]

For \( r = (r_1, \ldots, r_d) \), let \( f_{TT} \) be s.t.

\[
f(x) = f_{TT}(x) = \sum_{\alpha_i=0}^{n_i} \gamma_i(a_i, \alpha_i, x_i) \quad \text{for all } x \in [0, b)^d \]

where \( \gamma_i(a_i, \alpha_i, x_i) \) are orthogonal (see [1]).

\( f_{TT} \) is constructed through the eigenvalue decomposition of Hermitian integral operators defined in terms of \( f \). It can be proved that [1]:

- for fixed \( r \), \( f_{TT} \) is optimal
- \( \gamma_i(a_i, \alpha_i, x_i) \in C^1(\mathbb{R}) \) for all \( i, \alpha_i \).

The latter statement can be relaxed:

Spectral TT-decomposition

Let \( P : L^2(\mathbb{R}^d) \rightarrow \text{span} (\phi_i^{(N)}) \) where \( \phi_i^{(N)} \) are orthonormal polynomials:

**STT-Projection**

\[
P_{STT} = \sum_{i=1}^{N} \phi_i^{(N)} f_{TT}(x) \quad \text{for all } x \in [0, b)^d
\]

**STT-Interpolation**

\[
P_{STT} = \sum_{i=1}^{N} \phi_i^{(N)} f_{TT}(x) \quad \text{for all } x \in [0, b)^d
\]

where \( L^2 \) is the Lagrange interpolation matrix.

Conclusions

- Tackles the curse of dimensionality.
- Spectral convergence on smooth functions.

Ongoing works

- Anisotropic heterogeneous adaptivity.
- Ordering problem.
- Application in the fields of coastal engineering [2, 3] and geoscience.

Numerical Examples

**Ordering problem**

TT and STT are negatively affected by the wrong ordering of the dimensions, leading to an increased computational cost and severe loss of accuracy. We propose a strategy to find a good ordering.

Gonz functions: \( f_{TT}(x) = \cos(\pi x) \), \( f_{TT}(x) = \sin(\pi x) \), \( f_{TT}(x) = \cos(\pi x) + \sin(\pi x) \)

The method shows spectral convergence on both the tests, even on \( f_{TT} \), there is no analytical low-rank representation.

For \( d = 5 \), we compare the non-adaptive STT-Projection with the anisotropically adaptive Smolyak Sparse Grid.

References