Surrogate to Poincaré inequalities on manifolds for dimension reduction in nonlinear feature spaces

A. PASCO[†,1], A. NOUY[§,1]

[†] PhD student (presenting author). [§] PhD supervisor

PhD expected duration: Dec. 2022 – Nov. 2025

[1] Laboratoire de Mathématiques Jean Leray, Ecole Centrale de Nantes, Nantes Université
{alexandre.pasco,anthony.nouy}@ec-nantes.fr

Abstract

This work focuses on approximating a differentiable function $u : \mathbb{R}^d \to \mathbb{R}$ with $d \gg 1$ by a composition of functions $f \circ g$ where $g : \mathbb{R}^d \to \mathbb{R}^m$ and $f : \mathbb{R}^m \to \mathbb{R}$. The approximation error is assessed in the L^2_μ-norm where μ is some probability measure on $\mathbb{R}^d$. The approach considered is two-staged. Firstly the feature map g is selected among some prescribed functional class by minimizing some function $\mathcal{J}$ involved in the upper bound of the approximation error

$$\min_{f:\mathbb{R}^m\to\mathbb{R}} \mathbb{E}_\mu(|u - f \circ g|^2) \leq C_\mu \mathcal{J}(g), \tag{1}$$

which is based on Poincaré inequalities and requires evaluations of ∇u.

Secondly the function f is built using classical regression methods. Until recently, bounds of the form (1) were only available for linear feature maps g. This framework has been extensively studied under the name Active Subspace, see for example [2, 4], and the solution is given by the eigenvectors of the matrix $\mathbb{E}(\nabla u \nabla u^T) \in \mathbb{R}^d$. This approach is easy to implement, computationally efficient, has robust theoretical guarantees for some classical probability laws μ, and showed good performances in various numerical applications. However, there are many functions u for which such an approximation with $m < d$ is known to be not efficient.

Therefore, recent works consider non-linear feature maps in order to produce better dimension reduction. More especially, we will focus on the work from [1, 3] in which authors leverage Poincaré inequalities on smooth manifolds to obtain a bound of the form (1) for non-linear g. Although there are less theoretical guarantees, their numerical experiments showed improved performances compared with linear featuring. However, minimizing $\mathcal{J}$ is now much harder than finding eigenvectors of some matrix, and can only be done using iterative descent methods.

In this work we consider feature maps as in [1], of the form $g(x) = G^T \Phi(x)$ with $G \in \mathbb{R}^{K\times m}$ and where $\Phi : \mathbb{R}^d \to \mathbb{R}^K$, $K \geq d$, is fixed. We study a new quantity, denoted $\mathcal{L}(g)$, which can be expressed as the minimal singular value of some positive semi-definite matrix. We show that for a compact set of polynomial feature maps with $m = 1$, for some class of probability distributions, any minimizer g^* of $\mathcal{L}$ satisfies the sub-optimality result

$$\mathcal{J}(g^*) \lesssim \min_g \mathcal{J}(g)^\beta,$$

where $0 < \beta \leq 1$ is some constant which depends on the degree. We also extend this approach to the case $m > 1$ as well as for simultaneously learning a parametrized family of functions $u_y \in L^2_\mu$ by $y \in \mathcal{Y}$, although the theoretical results are weaker. Finally, we provide numerical examples to illustrate the performances of g^*, both as the feature map used in the regression step, or as the initializer for some iterative descent algorithm for minimizing $\mathcal{J}$.

Short biography (PhD student)

I graduated from Ecole Centrale de Nantes in 2022 as generalist engineer. It included 2 years focused on applied mathematics, from machine learning to numerical analysis. I continued at ECN by starting my PhD thesis on December 2022, funded by ANR COFNET, focusing on compositional function networks for nonlinear model reduction, for forward and inverse problems.

References

[1] Daniele Bigoni, Youssef Marzouk, Clémentine Prieur, and Olivier Zahm. Nonlinear dimension reduction for surrogate modeling using gradient information. *Information and Inference: A Journal of the IMA*, 11(4):1597–1639, December 2022.

[2] Paul G. Constantine, Eric Dow, and Qiqi Wang. Active Subspace Methods in Theory and Practice: Applications to Kriging Surfaces. *SIAM J. Sci. Comput.*, 36(4):A1500–A1524, January 2014.

[3] Romain Verdière, Clémentine Prieur, and Olivier Zahm. Diffeomorphism-based feature learning using Poincaré inequalities on augmented input space, December 2023.

[4] Olivier Zahm, Paul G. Constantine, Clémentine Prieur, and Youssef M. Marzouk. Gradient-Based Dimension Reduction of Multivariate Vector-Valued Functions. *SIAM J. Sci. Comput.*, 42(1):A534–A558, January 2020.