Sparse Precision (Interaction) Kernels for Machine Learning

Dionissios Hristopulos[1], Athanasios Liavas[1], Konstantinos Bekos[1], Nikolaos Paraskakis[1]
[1]Technical University Of Crete, Chania, Greece

Gaussian process regression (GPR) is a standard machine learning methods. The main component of GPR is the covariance kernel which carries information about the dependence across space and time. However, GPR-based predictive equations involve the inversion of the covariance kernel, which is a costly numerical operation since it scales as the third power of the number of sampling points. A different approach, inspired by statistical physics, relies on replacing the covariance kernel with the precision operator. The latter represents space-time interactions in the same spirit that is used in statistical physics for the construction of energy and action functionals for the purely spatial and space-time cases, respectively. Using the precision-operator formulation bypasses the inversion of the covariance kernel and can thus lead to significant computational gains. In the continuum case, it makes sense to construct precision operators which involve field derivatives. Such precision operators are localized by construction. In the case of discrete problems, that is, practically all cases that appear in the analysis of spatially and/or temporally distributed data, the discretization of the precision operator can be problematic. The problem is straightforwardly addressed in the case of regular sampling because the underlying lattice structure allows for replacement of the derivative operators with finite-difference approximations. In such cases one obtains Markov random fields, which are well studied in the scientific literature. However, for irregularly spaced samples (in space and/or time) the discretization of the precision operator is not straightforward. We will present two different discretization approaches. The first one builds the energy (action) functional in continuum space using a precision operator composed of polynomials of the Laplacian. The discrete version employs polynomials of the graph Laplacian (GLAP), in which weights are defined with the help of compactly supported kernel functions. Using properties of the graph Laplacian, it can be shown that such formulations are stable (positive definite), provided the respective coefficients are non-negative. The second approach involves an idea from the field of smoothed particle hydrodynamics (SPH): we spread out the influence of each sampling point using a kernel function with a free bandwidth parameter. This substitution replaces integrals over continuous space by weighted sums over particles, thus bridging the gap between the continuous and discrete cases. We will discuss the properties of the GLAP and SPH formulations focusing on their impact on estimation and prediction.

Acknowledgment: This research was funded by the research project SLIMNETS, which is implemented in the framework of H.F.R.I call “3rd Call for H.F.R.I.’s Research Projects to Support Faculty Members & Researchers” (H.F.R.I. Project Number: 25024.)