Dimension Adaptive Sparse Grids for Machine Learning

J. Garcke[1]

Abstract: Sparse grids are an approach for efficient high dimensional function approximation. In the form of the combination technique sparse grids have successfully been applied to the machine learning problems of classification and regression using a regularisation network approach [GGT00,Gar06]. The solution, which describes the relationship between the features and the response variable, is constructed using a linear combination of solutions on small partial grids.

Typically the importance of and variance within a dimension vary in machine learning applications which can exploited by different mesh resolutions for each feature. The degree of interaction between different dimensions also varies; this makes the usage of all dimensions in each partial grid unnecessary.

We present here a dimension adaptive sparse grid combination technique for the machine learning problems of classification and regression [Gar07]. The partial functions are now adaptively chosen during the computational procedure and possibly depend only on a subset of all features. This approach (approximately) identifies the ANOVA-decomposition of the underlying problem: $f(x) = \sum_{\{j_1,\ldots,j_q\}} c_{j_1,\ldots,j_q} f_{j_1,\ldots,j_q}(x_{j_1},\ldots,x_{j_q})$, where each $f_{j_1,\ldots,j_q}(x_{j_1},\ldots,x_{j_q})$ depends only on a subset of size q of the dimensions and may have different refinement levels for each dimension. We assume here $q < d$, the computational complexity depends mainly on the so-called *superposition* (or *effective*) dimension q.

We introduce a new error indicator for the dimension adaptive procedure which is inspired by a hierarchical subspace decomposition. It provides, in a dimension adaptive sense, localised information of the contribution of a partial grid to the overall solution.

Gar06 Garcke, J. (2006). Regression with the optimised combination technique. *23rd Internationel Conference on Machine Learning '06* (pp. 321–328).

Gar07 Garcke, J. (2007). A dimension adaptive sparse grid combination technique for machine learning. *Proceedings of 13th CTAC-2006* (pp. C725–C740).

GGT00 Garcke, J., Griebel, M., & Thess, M. (2001). Data mining with sparse grids. *Computing, 67*, 225–253.

[1] Technische Universität Berlin
Institut für Mathematik
Sekretariat MA 3-3
Straße des 17. Juni 136
10623 Berlin, Germany
garcke@math.tu-berlin.de