New HDMR Based Methods for Multivariate Data Modelling

Assist. Prof. Dr. M. Alper TUNGA

Bahçeşehir University, Faculty of Engineering,
Software Engineering Department, İstanbul, TÜRKİYE (TURKEY)

e-mail: alper.tunga@bahcesehir.edu.tr

Abstract: High Dimensional Model Representation (HDMR), which was first proposed by Sobol in 1993, is a divide-and-conquer method with a finite expansion. This expansion has a constant value, a summation on some univariate functions, another summation on some bivariate functions and so on. There are totally 2^N components in this expansion where N is the number of independent variables or parameters of the given multivariate problem. The HDMR method is used to represent a given multivariate function in terms of less variate functions to reduce the mathematical and computational complexity coming from multivariance in especially the computer based applications. For instance, using only the constant value and the summation of univariate functions, a univariate approximation can be obtained to represent the given multivariate function through HDMR. The determination process of each component appearing in the HDMR expansion includes multituple integrals under a product type multivariate weight function.

In multivariate data modelling problems, the nodes with the associated function values are given and it is asked to construct a model to be able to estimate the function value of any node whose function value is unknown. A product type weight need in HDMR algorithm results in an orthogonal geometry prerequisite in the data structure of the given problem. However, a data modelling problem usually has a multivariate training data set which includes a number of arbitrarily distributed points with the associated function values. The function values at all possible nodes of the problem domain cannot be known which means that we cannot use a product type weight in HDMR for modelling this type of problems. In addition, we know that if we would have an orthogonal geometry in the given multivariate data set, then the approximation obtained for the given problem becomes of the highest quality.

The Indexing HDMR methods transform the given problem space having a nonorthogonal geometry to an index space which has an orthogonal structure and model the index space instead of original. This gives us the flexibility of using the HDMR philosophy with a product type weight and since we are dealing with modelling data, a Dirac delta type weight best fits our purpose. This talk will cover the details of the Indexing HDMR Method and the Matrix based Indexing Method which are developed to construct a model for a given multivariate data modelling problem.

Brief Biography of M. Alper TUNGA

M. Alper TUNGA was born in İstanbul, Turkey on 11th June 1975. He received a B.Sc. degree in Mathematics Engineering from İstanbul Technical University (İ.T.Ü.) in 1997. He got his M.Sc. degree in Systems Analysis from İstanbul Technical University in 1999. He got a PhD from İstanbul Technical University in 2006 with a thesis entitled "Data Partitioning and Multivariate Interpolation via Various High Dimensional Model Representations ". In 1998, he worked as a research assistant in Computational Science and Engineering Department of İ.T.Ü. Between the years 1999-2006 he worked as a research assistant in the Computer Engineering Department of Işık University of Turkey. Since 2007, he is Assistant Professor in Bahçeşehir University. He is also a member of Group for Science and Methods of Computing in Informatics Institute of Istanbul Technical University. He is working on methodology for computational sciences. His interests are HDMR, multivariate data modelling and data mining. M. Alper Tunga has 12 papers about these subjects in various scientific journals.