Weight Parameters Optimization to Get Maximum Constancy in High Dimensional Model Representation

METİN DEMİRALP
Computational Science and Engineering Program, Informatics Institute,
İstanbul Technical University
İTÜ Bilişim Enstitüsü, İTÜ Ayazağa Yerleşkesi, Maslak, 34469, İstanbul, Türkiye (Turkey)
TURKEY (TÜRKİYE)

Abstract: - Since High Dimensional Model Representation (HDMR) uses a weight function it is possible to affect the contributions of the terms with different level of multivariance to the total norm square of HDMR. Of course, the most convenient HDMR is the one with the dominating constancy. This may not be achieved in all practically encountered cases. However, it is always better to try to get higher constancy in HDMR. This work focuses on the optimization of the HDMR's weight function to maximize the contribution of the HDMR's constant component to the total norm square of HDMR. To this end we parametrize the weight function factors by using appropriate basis functions. The resulting equation of optimization is a parametric matrix eigenvalue problem which can be easily treated by standing numerical methods.

Key-Words: - Multivariate Approximation, High Dimensional Model Representation, Optimization

1 Introduction

High Dimensional Model Representation and its certain varieties brought important contributions to multivariate analysis especially for approximating a multivariate function[1–11] in last fifteen years. Plain HDMR can be given through the following equality for a given multivariate function $f(x_1,...,x_N)$

$$f(x_1,...,x_N) = f_0 + \sum_{i_1=1}^{N} f_{i_1}(x_{i_1}) + \sum_{\substack{i_1,i_2=1\\ i_1<i_2}}^{N} f_{i_1,i_2}(x_{i_1},x_{i_2}) + \cdots \quad (1)$$

where the right hand side contains totally 2^N terms for the case of N independent variables. The right hand side components are determined by using the following conditions

$$\int_{a_i}^{b_i} dx_i W_i(x_i) f_{i_1,...,i_k}(x_{i_1},...,x_{i_k}) = 0, \quad i \in \{i_1,...,i_k\}, \quad 1 \le k \le N \quad (2)$$

where $W_i(x_i)$ $(1 \le i \le N)$ stand for appropriately chosen weight functions which are normalized as follow

$$\int_{a_i}^{b_i} dx_i W_i(x_i) = 1, \qquad 1 \le i \le N \quad (3)$$

and lead us to define the following product type multivariate weight function

$$W(x_1,...,x_N) \equiv \prod_{i=1}^{N} W_i(x_i) \quad (4)$$

These conditions correspond to the following orthogonality conditions

$$\int_{a_1}^{b_1} dx_1 W_1(x_1) \cdots \int_{a_N}^{b_N} dx_N W_N(x_N) \times f_{i_1,...,i_k}(x_{i_1},...,x_{i_k})\, f_{j_1,...,j_l}\left(x_{j_1},...,x_{j_l}\right) = 0,$$
$$\{x_{i_1},...,x_{i_k}\} \neq \left\{x_{j_1},...,x_{j_l}\right\}, \quad 1 \le k,l \le N \quad (5)$$

The multivariate function $f(x_1,...,x_N)$ is assumed to be known analytically in the above formulation of HDMR. This is required for uniqueness in the evaluation of the above integrals. Beyond this, the function $f(x_1,...,x_N)$ must be square integrable under the weight functions given above and over the hyperprism defined by the cartesian product of the closed intervals $[a_i,b_i]$ $(1 \le i \le N)$. This means that the function $f(x_1,...,x_N)$ must lie in an appropriately defined Hilbert space to mention about the orthogonality. This Hilbert space is defined over the abovementioned hyperprism with the following inner product for two arbitrary square integrable multivariate functions denoted by $g(x_1,...,x_N)$ and

$h(x_1,...,x_N)$ respectively in this space,

$$(g,h) \equiv \int_{a_1}^{b_1} dx_1 ... \int_{a_N}^{b_N} dx_N W(x_1,...,x_N) \times g(x_1,...,x_N)\, h(x_1,...,x_N) \qquad (6)$$

This enables us to use an induced norm defined as follows for an arbitrary function $g(x_1,...,x_N)$ in the same space

$$\|g\| \equiv (g,g)^{\frac{1}{2}} \qquad (7)$$

Now we can write the following norm equality by using this norm definition and the mutual orthogonality of the HDMR components

$$\|f\|^2 = \|f_0\|^2 + \sum_{i_1=1}^{N} \left\| f_{i_1} \right\|^2 + \sum_{\substack{i_1,i_2=1\\ i_1<i_2}}^{N} \left\| f_{i_1,i_2} \right\|^2 + \cdots \qquad (8)$$

This brings the idea of defining certain parameters to measure the contributions of the HDMR truncations to the total norm square of HDMR into mind. We can do this as follows

$$\begin{aligned} \sigma_0(f) &\equiv \frac{\|f_0\|^2}{\|f\|^2} \\ \sigma_1(f) &\equiv \sigma_0(f) + \frac{1}{\|f\|^2} \sum_{i_1=1}^{N} \left\| f_{i_1} \right\|^2 \\ \sigma_2(f) &\equiv \sigma_1(f) + \frac{1}{\|f\|^2} \sum_{\substack{i_1,i_2=1\\ i_1<i_2}}^{N} \left\| f_{i_1,i_2} \right\|^2 \\ . \quad & \quad ... \end{aligned} \qquad (9)$$

where the function dependence of these parameters are explicitly shown as an argument. This formula obviously permits us to write

$$\sigma_0(f) \le \sigma_1(f) \le \sigma_2(f) \le ... \le \sigma_N(f) \equiv 1 \qquad (10)$$

which implies that $\{\sigma_0(f),...,\sigma_N(f)\}$ forms a monotonously increasing sequence towards to and including 1. σ_0 can achieve 1 if and only if the multivariate function at the focus remains constant everywhere in the domain of HDMR. As this function starts to deviate from its constant value, σ_0 starts to decrease from 1. Hence, σ_0 somehow measures the constancy of the target function of HDMR. Hence, we call it "Constancy Measurer". The other σ entities can be interpreted in similar ways and are generally called "Additivity Measurers".

The number of the HDMR components is 2^N for the case where the number of independent variables is N. Although HDMR contains a finite number of terms, the number of the HDMR components may climb to very high values as N grows. This may make impractical the utilization of the total expression of HDMR. Then not the whole expression but its certain truncation from the first term may be sought to be used as an approximation. Of course, the most preferable case is the truncation containing only the constant component of HDMR. If the function at the target of HDMR is somehow close to a constant then this truncation may be expected as a good approximation. Even this is not happening, the maximization of the HDMR's constant component can be considered as a reasonable action to increase the numerically expressing capability of the constant truncation of HDMR.

Before the completion of this section we need to give the explicit expression of the constant component of HDMR. To this end we can multiply the both sides of (1) by the weight function and integrate over the HDMR's domain and then take the orthogonality conditions into consideration. This gives

$$f_0 = \int_{a_1}^{b_1} dx_1 W_1(x_1) \cdots \int_{a_N}^{b_N} dx_N W_N(x_N) \times f(x_1,...,x_N) \qquad (11)$$

which is in fact the weighted mean value of the original function over the HDMR's hyperprismatic region.

The last equality explicitly shows the dependence of the constant component of HDMR on the weight function factors. Hence it seems to be possible to manipulate the the weight function factors to maximize the value of the HDMR's constant component, f_0. This can be done by using either an explicit variational scheme to get the functional structures of the weight function factors or by utilizing a linear algebraic optimization scheme with respect to certain scalar unknowns after appropriately parametrizing the weight function factors. Matrix algebraic tools will be used in this section and the existence and uniqueness of the solutions will be discussed in sufficient details. We will prefer the latter case in this work.

Paper is organized as follows. The second section presents the parametrization of the weight function. This puts all investigations into a more amenable form. Third section focuses on the optimization of the weight function parameters to get maximum constancy. The fourth section finalizes the paper by presenting concluding remarks.

2 Weight Function Parametrization

Consider the weight function factor $W_i(x_i)$. This function must belong to the set of functions which are integrable over the interval $[a_i, b_i]$ and, beyond this, it must be nonnegative over the abovementioned interval. To be more practical we can consider the function $W_i(x_i)$ as the square of an arbitrary function belonging to the linear vector space of functions which are square integrable over the interval $[a_i, b_i]$. If we denote the elements of an orthonormal basis set by $\upsilon_j^{(i)}(x_i)$ $(1 \leq j < \infty)$ then we can write

$$W_i(x_i) = \left(\sum_{j=1}^{m_i} \alpha_j^{(i)} \upsilon_j^{(i)}(x_i) \right)^2, \quad 1 \leq i \leq N \tag{12}$$

which is m_i–term truncation of most general infinite series representation (we have truncated at m_i terms for practical reasons, however m_i can be increased as much as we want of course). (12) takes us to the following global weight formula

$$W(x_1, ..., x_N) = \prod_{i=1}^{N} \left[\left(\sum_{j_i=1}^{m_i} \alpha_{j_i}^{(i)} \upsilon_{j_i}^{(i)}(x_i) \right)^2 \right] \tag{13}$$

Let us define the following finite sets

$$\mathcal{S}_f \equiv \{u_1(x_1, ..., x_N), ..., u_{N_b}(x_1, ..., x_N)\} \tag{14}$$

$$\mathcal{S}_p \equiv \{a_1, ..., a_{N_b}\} \tag{15}$$

where

$$u_i(x_1, ..., x_N) \equiv \upsilon_{j_1}(x_1) \times \cdots \times \upsilon_{j_N}(x_N),$$

$$\begin{aligned} & 1 \leq i = i(j_1, ..., j_N) \leq N_b, \\ & N_b \equiv m_1 \times \cdots \times m_N \\ & 1 \leq j_1 \leq m_1, ..., 1 \leq j_N \leq m_N \end{aligned} \tag{16}$$

and

$$a_i \equiv \alpha_{j_1} \times \cdots \times \alpha_{j_N},$$

$$\begin{aligned} & 1 \leq i = i(j_1, ..., j_N) \leq N_b, \\ & N_b \equiv m_1 \times \cdots \times m_N \\ & 1 \leq j_1 \leq m_1, ..., 1 \leq j_N \leq m_N \end{aligned} \tag{17}$$

and subscripts f and p in set symbols and the subscript in N_b above imply the words “function”, “parameter” and “basis” respectively.

The recent definitions enable us to reexpress the overall weight function as follows

$$\begin{aligned} & W(x_1, ..., x_N) = \\ & = \sum_{i=1}^{N_b} \sum_{j=1}^{N_b} a_i a_j u_i(x_1, ..., x_N) u_j(x_1, ..., x_N) \end{aligned} \tag{18}$$

Let us use the following shorthand notations: (1) $\mathbf{x}$ for $x_1, ..., x_N$, (2) a single integral symbol with lower limit $\mathcal{V}$ standing for N–fold multivariate integration domain, the previously mentioned hyperprism which is in fact the cartesian product of the aforementioned intervals, that is,

$$\mathcal{V} \equiv [a_1, b_1] \times \cdots \times [a_N, b_N], \tag{19}$$

(3) $d\mathcal{V}$ for the volume element of the N–fold multivariate integration. That is,

$$d\mathcal{V} \equiv dx_1 ... dx_N. \tag{20}$$

Now the orthonormality of the N_b number of basis functions given above can be expressed as follows in terms of these shorthand notations

$$\int_{\mathcal{V}} d\mathcal{V} u_j(\mathbf{x}) u_k(\mathbf{x}) = \delta_{jk}, \quad 1 \leq j, k \leq N_b \tag{21}$$

where δ_{jk} stands for the Kroenecker’s delta symbol which vanishes when its two subscripts differ otherwise becomes 1. This equality and the normalization of the integral of the gloabal weight function to 1 reveal the following condition for the a_i $(1 \leq i \leq N_b)$ parameters

$$\sum_{i=1}^{N_b} a_i^2 = 1 \tag{22}$$

Now we have been sufficiently equipped to proceed for the optimization of weight parameters to get maximum constancy in HDMR.

3 Weight Optimization for Maximum Constancy

We can start with f_0 to construct a cost functional for the optimization of the weight parameters to get maximum constancy. We can reexpress f_0 as follows

$$f_0 = \mathbf{a}^T \mathbf{M}(f) \mathbf{a} \tag{23}$$

where the elements of $N_b \times N_b$ matrix $\mathbf{M}(f)$ and the N_b element column vector $\mathbf{a}$ are defined as follows

$$\begin{aligned} M_{ij}(f) &\equiv \int_{\mathcal{V}} d\mathcal{V} u_i(\mathbf{x}) f(\mathbf{x}) u_j(\mathbf{x}) \\ & \quad 1 \le i, j \le N_b \end{aligned} \tag{24}$$

$$\mathbf{a}^T \equiv [\, a_1\, a_2 \ldots a_N \,] \tag{25}$$

The last three formulae enable us to write

$$\|f_0\|^2 = \left(\mathbf{a}^T \mathbf{M}(f)\,\mathbf{a}\right)^2 \tag{26}$$

On the other hand the above considerations lead us to write

$$\|f\|^2 = \mathbf{a}^T \mathbf{M}\left(f^2\right)\mathbf{a} \tag{27}$$

where the elements of $N_b \times N_b$ matrix $\mathbf{M}\left(f^2\right)$ are defined as follows

$$\begin{aligned} M_{ij}\left(f^2\right) &\equiv \int_{\mathcal{V}} d\mathcal{V} u_i(\mathbf{x}) f(\mathbf{x})^2 u_j(\mathbf{x}), \\ & \quad 1 \le i, j \le N_b \end{aligned} \tag{28}$$

Now we can define the following cost functional for the optimization of the weight function parameters

$$\mathcal{J}(f, \mathbf{a}) \equiv \|f\|^2 - \|f_0\|^2 + \lambda\left(\|\mathbf{a}\|^2 - 1\right) \tag{29}$$

where the Frobenius norm of the matrices has been used here and λ stands for the Lagrange parameter to imbed the normalization condition of $\mathbf{a}$. We can rewrite (29) more explicitly as follows

$$\begin{aligned} \mathcal{J}(f, \mathbf{a}) &\equiv \mathbf{a}^T \mathbf{M}\left(f^2\right)\mathbf{a} - \left(\mathbf{a}^T \mathbf{M}(f)\,\mathbf{a}\right)^2 \\ & \quad + \lambda\left(\mathbf{a}^T \mathbf{a} - 1\right) \end{aligned} \tag{30}$$

The optimization procedure requires the setting of this cost functional's first order partial derivatives with respect to the elements of vector $\mathbf{a}$ equal to zero. That is,

$$\frac{\partial \mathcal{J}(f, \mathbf{a})}{\partial a_i} = 0, \qquad 1 \le i \le N_b \tag{31}$$

which produces the following parametric matrix eigenvalue problem

$$2 f_0 \mathbf{M}(f)\,\mathbf{a} - \mathbf{M}\left(f^2\right)\mathbf{a} = \lambda \mathbf{a} \tag{32}$$

This problem can be solved at least numerically and the eigenvalue and eigenvectors can be evaluated as functions of f_0. If we denote the resulting f_0–dependent eigenvalues together with the f_0–dependent eigenvectors whose norms are 1 with respect to square (Frobenius) norm by $\lambda_1(f_0),\ldots,\lambda_{N_b}(f_0)$ and $\mathbf{a}_1(f_0),\ldots,\mathbf{a}_{N_b}(f_0)$ respectively then we can make the following definitions

$$\mu_i(f_0) \equiv \mathbf{a}_i^T(f_0)\,\mathbf{M}(f)\,\mathbf{a}_i(f_0), \quad 1 \le i \le N_b \tag{33}$$

which take us to the following algebraic equations

$$\mu_i(f_0) = f_0, \quad 1 \le i \le N_b \tag{34}$$

Some of the f_0 value obtained from all of these equations will make the constancy measurer σ_0 maximum. That or those values and corresponding $\mathbf{a}$ vectors should be taken as the solution of the optimization problem to get maximum constancy in HDMR.

4 Concluding Remarks

The main goal of this paper has been to get the mathematical equations for obtaining the global (multivariate) weight function to maximize the constancy of the High Dimensional Model Representation of an analytically given multivariate function. The basic idea to this end is to parametrize the univariate factors of the global weight function which is assumed to be product of univariate functions each of which depends on a different independent variable. We have chosen the way of parametrization in such a way that each univariate factor of the weight function can be expressed as the square of certain finite linear combinations of the basis functions (which are infinite in number since they lie in an Hilbert space) of the space of squarely integrable functions. This choice brought the possibility of expressing the global weight function as a quadratic form in terms of the linear combination coefficients which are in fact the optimization parameters. This enabled us to construct a cost functional which is basically the sum of the deviation of the norm square of the constant component of HDMR from the norm square of the original function and a Lagrange multiplier including constraint term about the normalization of the optimization parameters. We have obtained a parametric matrix eigenvalue problem such that it can be solved at least numerically without any serious problem because of the symmetry in the matrices.

This idea can be used for all continuous HDMR problems where the function under consideration is given analytically. The cases where the functions are given as certain numerical values at certain specific points of the domain of HDMR can not be treated in this way since the Dirac's delta function's square is undefined and we need it in the construction. Hence some other ways of parametrization should be considered there.

One can of course consider some other ways of parametrization even in this case of contionuous function's HDMR. We tried to use possibly most easiest way of constructing an optimization technique to get best value of constancy.

We have not given any numerical application here since our basic aim was just methodology without any doubt on numerical instability or applicability. What we have obtained here can be treated by the standard methods of solving eqigenvalue problems.

Acknowledgement:
Author is grateful to Turkish Academy of Sciences for its support.

References:

[1] I.M. Sobol, Sensitivity Estimates for Nonlinear Mathematical Models, *Mathematical Modeling and Computational Experiments*, 1993, VOL. 1, pp 407-414.

[2] J.A. Shorter, P.C. Ip, and H. Rabitz, An Efficient Chemical Kinetics Solver Using High Dimensional Model Representation, txitext J. Phys. Chem. A, 1999, VOL 103, pp 7192-7198.

[3] Ö. Alış and H. Rabitz, Efficient Implementation of High Dimensional Model Representations, *Journal of Mathematical Chemistry*, 2001, VOL 29, pp 127-142.

[4] G. Li, S. W. Wang, H. Rabitz, S. Wang and P. Jaffe, Global uncertainty assessments by high dimensional model representation (HDMR). *Chemical Engineering Science*, 2002, VOL 57, pp 4445-4460.

[5] G. Li, M. Artamonov, H. Rabitz, S. Wang, P.G. Georgopoulos and M. Demiralp, High Dimensional Model Representations Generated from Low Order Terms -lp-RS-HDMR, *Journal of Computational Chemistry*, 2003, VOL 24, pp 647-656.

[6] İ. Yaman and M. Demiralp, High Dimensional Model Representation Approximation of an Evolution Operator with a First Order Partial Differential Operator Argument, *Applied and Numerical Analysis and Computational Mathematics*, 2003, VOL 1, pp 287-296.

[7] M.A. Tunga and M. Demiralp, A Factorized High Dimensional Model Representation on the Partitioned Random Discrete Data, *Applied and Numerical Analysis and Computational Mathematics*, 2003, VOL 1, pp 233-243.

[8] M.A. Tunga and M. Demiralp, A factorized high dimensional model representation on the nodes of a finite hyperprismatic regular grid, *Applied Mathematics and Computation*, 2005, VOL 164, pp 865-883.

[9] M.A. Tunga and M. Demiralp, Hybrid high dimensional model representation (HHDMR) on the partitioned data, *Journal of Computational and Applied Mathematics*, 2006, VOL 185, pp 107-132.

[10] M. Demiralp, High Dimensional Model Representation and its application varieties, *Tools for Mathematical Methods, Mathematical Research, St-Petersburg*, 2003, VOL 9, pp 146-159.

[11] M. Demiralp, A Univariately Approximate Data Partitioning Algorithm for a Dominantly Multiplicative Multivariate Function, *Proceedings, WSEAS 5. International Conference on Applied Computer Science (ACOS'06), Hangzhou-China*, 2006.