A Nonlinear Perturbative Scheme to Solve Weight Optimization Problem of High Dimensional Model Representation (HDMR)

BURCU TUNGA
İstanbul Technical University
Informatics Institute
Maslak, 34469, İstanbul
TÜRKİYE (TURKEY)
burcu@be.itu.edu.tr

METİN DEMİRALP
İstanbul Technical University
Informatics Institute
Maslak, 34469, İstanbul
TÜRKİYE (TURKEY)
demiralp@be.itu.edu.tr

Abstract: One way to increase the approximating quality of the High Dimensional Model Representation (HDMR) truncations is to increase the truncation order. However this is not generally desired for practical reasons if the order climbs to the multivariances beyond the bivariance. In these circumstances the other alternative is preferred. It is to change the structure of HDMR. This can be done either by using a different but again orthogonal geometry or by changing the structure of the weight function. Weight optimization is based on the constancy maximization and in fact gives different importances to the function values at different points of the HDMR domain. Weight function is considered as the square of a linear combination of certain basis functions and the linear combination coefficients are determined to maximize the constancy. The resulting equations are nonlinear. This work attempts to solve these equations by expanding unknowns around their certain nominal values.

Key–Words: Multivariate Approximation, High Dimensional Model Representation, Perturbation Expansion

1 Introduction

In last two decades there has been significant developments in High Dimensional Model Representation (HDMR) based algorithms. They were used to approximate a given mutivariate function in terms of some other less variate functions. The HDMR is based on a divide–and–conquer idea and was first designed by Sobol in 1993 [1].

After Sobol's epochal proposal, Rabitz and his group studied on this topic and they generalized the HDMR's domain from Sobol's unit cube to any, finite, semi-finite, or infinite orthogonal hyperprismatic regions. They also brought the weight function concept to the HDMR definition. In this perpective Sobol was using the constant weight function having the value of 1. The weight functions to be used in HDMR ought to be a product of univariate weight functions each of which depends on a different independent variable of HDMR [2–4]. Demiralp and his group developed various new methods based on the HDMR philosophy, each of which is designed to work well for a different specific purpose [5–9].

According to the Sobol's suggestion, a multivariate function of N independent variables can be expressed as follows

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

where $f(x_1,...,x_N)$ denotes the target multivariate function of HDMR. The functions appearing at the right hand side of the equation are called HDMR components. These are a constant term, univariate terms, bivariate terms and so on respectively.

The HDMR components at the right hand side of (1) are mutually orthogonal. That is, its right hand side is an orthogonal decomposition to the original function. The orthogonality is defined via the following inner product over $\mathcal{H}$, the Hilbert space of the functions square integrable over the orthogonal hyperprism defined as the direct product of the intervals $[\,a_i,b_i\,]$ $(i=1,2,...,N)$.

$$\begin{aligned} (g,h) \equiv & \int_{a_1}^{b_1} dx_1 ... \int_{a_N}^{b_N} dx_N W(x_1,...,x_N) \\ & \times g(x_1,...,x_N)\, h(x_1,...,x_N) \end{aligned} \quad (2)$$

where $g(x_1,...,x_N)$ and $h(x_1,...,x_N)$ are any two functions chosen from $\mathcal{H}$ while $W(x_1,...,x_N)$ stands for a given weight function.

The HDMR components in Sobol's case can be uniquely determined by using the following vanishing conditions first proposed by Sobol.

$$\int_{a_i}^{b_i} dx_i f_i(x_i) = 0, \qquad i = 1, 2, ..., N \tag{3}$$

These conditions are not peculiar only to the univariate terms. As Sobol imposed, the integral of each HDMR component between 0 and 1 inclusive over anyone of its independent variables should vanish. These enable us to uniquely determine all HDMR components. (3) can be rewritten as follows

$$\int_{a_1}^{b_1} dx_N ... \int_{a_N}^{b_N} dx_N f_i(x_i) = 0, \qquad i = 1, 2, ..., N \tag{4}$$

through a tricky mathematical idea. This implies the orthogonalities of all univariate components to the constant function of 1 value and therefore to the constant component of HDMR. On the other hand (3) can also be rewritten as follows

$$\int_{a_1}^{b_1} dx_N ... \int_{a_N}^{b_N} dx_N f_i(x_i) f_j(x_j) = 0, \quad i, j = 1, 2, ..., N, \quad i \neq j \tag{5}$$

by using a similar consideration to the one we have used to create (4). These equalities prove the existence of the orthogonalities amongst the univariate components. Similar considerations allow us to show the existence of the orthogonality amongst the all HDMR components when the vanishing conditions are imposed.

The generalization of (4) to the case of nonunit weights can be expressed as follows

$$\int_{a_1}^{b_1} dx_N ... \int_{a_N}^{b_N} dx_N W(x_1, ..., x_N) f_i(x_i) = 0, \quad i = 1, 2, ..., N \tag{6}$$

The weight function appearing in these conditions is assumed to be a product of univariate functions each of which depends on a different independent variable. That is,

$$W(x_1, \ldots, x_N) \equiv \prod_{j=1}^{N} W_j(x_j), \quad x_j \in [a_j, b_j], \qquad 1 \leq j \leq N \tag{7}$$

where each univariate weight factor is assumed to satisfy the following integral normalization condition to facilitate the determination of the HDMR components

$$\int_{a_j}^{b_j} dx_j W_j(x_j) = 1, \qquad 1 \leq j \leq N \tag{8}$$

The generalized forms of the vanishing conditions for the other HDMR components are not given explicitly here since truncation at univariance is most preferred approximation as long as its quality is sufficiently high. However, it is not hard to guess the structure. The only thing to be changed is the addition of the weight function as a new factor. We find this information sufficient for our purpose in this work.

To facilitate the analysis we can use certain projection operators. We can start by defining the following projection operator to determine the constant HDMR component, f_0.

$$\mathcal{P}_0 g(x_1, ..., x_N) \equiv \int_{a_1}^{b_1} dx_1 ... \int_{a_N}^{b_N} dx_N W(x_1, ..., x_N) \times g(x_1, ..., x_N) \tag{9}$$

where $g(x_1, ..., x_N)$ can be any function in the Hilbert space $\mathcal{H}$. The orthogonality of all higher-than-zero-order multivariate components to f_0 implies that the integrals of those components over one of their independent variables over the related interval under the corresponding univariate weight function vanish (vanishing property proposed by Sobol) as we have mentioned above. If we now apply the projection operator $\mathcal{P}_0$ on both sides of (1) and then utilize the vanishing properties of the higher-than-zero-variate terms, and the normalized nature of the univariate weight factors then we can write

$$f_0 = \mathcal{P}_0 f(x_1, ..., x_N) \tag{10}$$

To determine the univariate terms, $f_i(x_i)$s, some other projection operators, $\mathcal{P}_i$s ($1 \leq i \leq N$) are defined by following the same philosophy of the constant term determination case. If we apply these projection operators on both sides of the equation (1), we obtain the univariate terms of HDMR as follows.

$$f_i(x_i) = \mathcal{P}_i f(x_1, ..., x_N) - f_0, \quad 1 \leq i \leq N \tag{11}$$

Bivariate terms and higher variate HDMR components can be found by defining some other projection operators in the same manner.

According to the HDMR algorithm, the N-dimensional multivariate function under consideration can be represented by a constant term, N univariate terms, $N(N-1)/2$ bivariate terms, $N(N-1)(N-2)/6$ trivariate terms and so on. Hence, the total number of HDMR components for a given N-variate function is 2^N. Although this number is finite, it may climb to a very high number as N increases. Hence, we intend to truncate HDMR at rather small multivariances as long as the truncation has a good representation quality. For this purpose, the HDMR approximants and the additivity measurers are defined

as follows.

$$\begin{aligned} s_0(x_1,...,x_N) &\equiv f_0 \\ s_1(x_1,...,x_N) &\equiv s_0(x_1,...,x_N) + \sum_{i=1}^{N} f_i(x_i) \\ &\vdots \\ s_k(x_1,...,x_N) &\equiv s_{k-1}(x_1,...,x_N) \\ &\quad + \sum_{\substack{i_1,...,i_k=1 \\ i_1<\cdots<i_k}}^{N} f_{i_1...i_k}(x_{i_1},...,x_{i_k}) \\ &\quad 1 \le k \le N \end{aligned} \tag{12}$$

Additivity Measurers are defined for measuring the quality of these approximants for the characterization of the original function within a desired numerical precision.

$$\begin{aligned} \sigma_0 &\equiv \frac{1}{\|f\|^2}\|f_0\|^2 \\ \sigma_1 &\equiv \frac{1}{\|f\|^2}\sum_{i=1}^{N}\|f_i\|^2 + \sigma_0 \\ &\vdots \\ \sigma_N &\equiv \frac{1}{\|f\|^2}\|f_{12...N}\|^2 + \sigma_{N-1} \end{aligned} \tag{13}$$

Here, σ_0 is called Constancy Measurer and it defines the contribution percentage of the constant term to the HDMR expansion's norm square. As a generalization, σ_k called kth Order Additivity Measurer and it defines the contribution percentage of the all terms from constant term to kth order term inclusive to the HDMR expansion's norm.

As we have previously mentioned, we intend to truncate the HDMR terms at most bivariate terms, but we also intend to represent the original function as accurate as possible in this work. For this purpose, we can find an appropriate weight function by using optimization rules. This paper aims to realize an optimization via constancy measurer. After optimization we obtain a parametric linear equation together with a nonlinear algebraic equation, whose right hand side is a ratio of the quadratic forms over the unknown vector, to define the unknown parameter of the linear equation. These equations are attempted to be solved with the help of a perturbation expansion.

The rest of the paper is organized as follows. The second section covers the derivation of the optimization equations through constancy measurer. These equations can be solved by using perturbation expansion and this issue will be presented in the third section of the paper. The final section involves the concluding remarks about this work and the future steps to more maturize the issue.

2 The Weight Optimization

The weight function to be optimized could be structured as a linear combination of the certain elements from a complete orthonormal basis function set. However, this structure does not guarantee the nonnegative nature of the weight function. To provide the nonnegativity for the weight function and to use the benefits of linear algebraic facilities we use a squared linear combination of certain basis functions as follows

$$w(x) = \left(\sum_{j=1}^{m}\alpha_j w_j(x)\right)^2 \tag{14}$$

where α_js stand for certain arbitrary parameters. If this weight function is used to determine the constant HDMR term, the following structure is obtained

$$f_0 = \sum_{j=1}^{m}\sum_{k=1}^{m}\alpha_j\alpha_k a_{jk}^{(1)} \tag{15}$$

where $a_{jk}^{(1)}$ symbolizes the following integral

$$\begin{aligned} a_{jk}^{(1)} &\equiv \int_a^b dx w_j(x)w_k(x)f(x), \\ &\quad j,k = 1,2,...,m \end{aligned} \tag{16}$$

To determine the constancy measurer, σ_0, the expression $||f||^2$ has to be determined. It has the following expression

$$||f||^2 = \sum_{j=1}^{m}\sum_{k=1}^{m}\alpha_j\alpha_k a_{jk}^{(2)} \tag{17}$$

where $a_{jk}^{(2)}$ is defined as follows

$$\begin{aligned} a_{jk}^{(2)} &\equiv \int_a^b dx w_j(x)w_k(x)f(x)^2, \\ &\quad j,k = 1,2,...,m \end{aligned} \tag{18}$$

The general structure of the constancy measurer is obtained as follows after using what we have obtained until now

$$\sigma_0 = \frac{\left(\sum_{j=1}^{m}\sum_{k=1}^{m}\alpha_j\alpha_k a_{jk}^{(1)}\right)^2}{\sum_{j=1}^{m}\sum_{k=1}^{m}\alpha_j\alpha_k a_{jk}^{(2)}} \tag{19}$$

If the integral normalization condition of the weight function and the orthogonality amongst the basis functions of the linear combination appeared above are used then the α parameters are found to satisfy the following vector normalization condition

$$\sum_{j=1}^{m} \alpha_j^2 = 1 \tag{20}$$

We need to construct a cost functional to be extremized for finding the optimum values of α parameters by making constancy measurer maximum. Our needs lead us to define the following one

$$\begin{aligned} J(\alpha_1, \cdots, \alpha_m, \lambda) &= \frac{\left(\sum\limits_{j=1}^{m} \sum\limits_{k=1}^{m} \alpha_j \alpha_k a_{jk}^{(1)} \right)^2}{\sum\limits_{j=1}^{m} \sum\limits_{k=1}^{m} \alpha_j \alpha_k a_{jk}^{(2)}} \\ &+ \lambda \left(\sum_{j=1}^{m} \alpha_j^2 - 1 \right) \end{aligned} \tag{21}$$

each of whose first order derivatives with respect to independent variables should be set equal to zero to get the equations of the optimization. This produces the following equations from differentiations with respect to λ and α_i

$$\sum_{j=1}^{m} \alpha_j^2 = 1 \tag{22}$$

$$-2\mu \sum_{k=1}^{m} \alpha_k a_{ik}^{(1)} + \mu^2 \sum_{k=1}^{m} \alpha_k a_{ik}^{(2)} = \lambda \alpha_i \tag{23}$$

where μ is given as follows

$$\mu \equiv \frac{\sum\limits_{j=1}^{m} \sum\limits_{k=1}^{m} \alpha_j \alpha_k a_{jk}^{(1)}}{\sum\limits_{j=1}^{m} \sum\limits_{k=1}^{m} \alpha_j \alpha_k a_{jk}^{(2)}} \tag{24}$$

The last two equations can be put into concise forms by using certain matrix definitions. We can write

$$\left(-2\mu \mathbf{A_1} + \mu^2 \mathbf{A_2} \right) \boldsymbol{\alpha} = \lambda \boldsymbol{\alpha} \tag{25}$$

$$\mu = \frac{\boldsymbol{\alpha^T A_1 \alpha}}{\boldsymbol{\alpha^T A_2 \alpha}}, \qquad \mu_\ell \leq \mu \leq \mu_u \tag{26}$$

where $\mathbf{A}_1$ and $\mathbf{A}_2$ are the $m \times m$ type matrices whose elements at the intersection of jth row and kth column are $a_{jk}^{(1)}$ and $a_{jk}^{(2)}$ respectively. The lower and upper bounds for the parameter μ, μ_ℓ and μ_u, are the least and greatest eigenvalues of $\mathbf{A}_1$ under the weight matrix $\mathbf{A}_2$. $\boldsymbol{\alpha}$ stands for the m dimensional column vector whose elements are α_1,...,α_m. (25) is an eigenvalue problem where λ and $\boldsymbol{\alpha}$ symbolize the eigenvalue and the corresponding eigenvector of the matrix $\left(-2\mu \mathbf{A_1} + \mu^2 \mathbf{A_2} \right)$ as long as (26) is not imposed as an accompanying equation. Hence it is a linear eigenvalue problem as long as μ is considered as a given parameter. However the face of the problem changes when (26) accompanies (25). (26) brings nonlinearity to the problem and all beatiful properties of the linear algebraic eigenvalue problems may disappear depending on the natures of the given entities of the problem.

3 A Perturbation Expansion Around a Chosen Value of the μ Parameter

It is numerically better to use a parameter between 0 and 1 inclusive instead of some other parameter between any two numbers. Hence we define

$$\overline{\mu} \equiv \frac{\mu - \mu_\ell}{\mu_u - \mu_\ell} \tag{27}$$

$$\overline{\lambda} \equiv \frac{\lambda}{\mu} \tag{28}$$

and

$$\begin{aligned} \overline{\mathbf{A}}_1 &\equiv (-2\mathbf{A}_1 + \mu_\ell \mathbf{A}_2) \\ \overline{\mathbf{A}}_2 &\equiv (\mu_u - \mu_\ell)\, \mathbf{A}_2 \end{aligned} \tag{29}$$

These convert the equations (25) and (26) to the following ones

$$\left(\overline{\mathbf{A}}_1 + \overline{\mu} \overline{\mathbf{A}}_2 \right) \boldsymbol{\alpha} = \overline{\lambda} \boldsymbol{\alpha} \tag{30}$$

$$\overline{\mu} = \frac{\boldsymbol{\alpha}^T \left(\mathbf{A}_1 - \mu_\ell \mathbf{A}_2 \right) \boldsymbol{\alpha}}{(\mu_u - \mu_\ell)\, \boldsymbol{\alpha}^T \mathbf{A}_2 \boldsymbol{\alpha}}, \qquad 0 \leq \overline{\mu} \leq 1 \tag{31}$$

If the right hand side of (31) were independent of $\boldsymbol{\alpha}$, that is, a constant then the last two equations would become just a linear algebraic eigenvalue problem which can be at least numerically solved by using one of well-known efficient methods. Hence, we can consider a couple of more general equations having a new parameter such that they match the last two equations when the new parameter becomes 1 and they become an algebraic eigenvalue problem when the new parameter vanishes. These are given as follows

$$\left(\overline{\mathbf{A}}_1 + \overline{\mu}\,(\epsilon)\, \overline{\mathbf{A}}_2 \right) \boldsymbol{\alpha}(\boldsymbol{\epsilon}) = \overline{\lambda}(\epsilon) \boldsymbol{\alpha}(\boldsymbol{\epsilon}) \tag{32}$$

$$\begin{aligned} \overline{\mu}(\epsilon) &= \overline{\mu}_c \\ &+ \epsilon \left(\frac{\boldsymbol{\alpha}^T(\epsilon)(\mathbf{A}_1 - \mu_\ell \mathbf{A}_2)\boldsymbol{\alpha}(\epsilon)}{(\mu_u - \mu_\ell)\, \boldsymbol{\alpha}^T(\epsilon) \mathbf{A}_2 \boldsymbol{\alpha}(\epsilon)} - \overline{\mu}_c \right) \end{aligned} \tag{33}$$

where we have three unknowns, $\overline{\mu}(\epsilon)$, $\overline{\lambda}(\epsilon)$, $\boldsymbol{\alpha}(\epsilon)$ and ϵ denotes the perturbation parameter. The parameter $\overline{\mu}_c$ is chosen in the domain of the parameter $\overline{\mu}$. To solve these equations, each unknown will be assumed to be expandable to a power series in the perturbation parameter. That is,

$$\overline{\mu}(\epsilon) \equiv \sum_{i=1}^{\infty} \overline{\mu}_i \epsilon^i \tag{34}$$

$$\overline{\lambda}(\epsilon) \equiv \sum_{i=1}^{\infty} \overline{\lambda}_i \epsilon^i \tag{35}$$

$$\boldsymbol{\alpha}(\epsilon) \equiv \sum_{i=1}^{\infty} \boldsymbol{\alpha}_i \epsilon^i \tag{36}$$

If we reorganize the resulting series in ascending powers of the perturbation parameter at the both sides of the resulting equations after we use these expansions in (32) and (33) then the coefficients of the same powers of the perturbation parameter at the both side of the equations should be equivalent. This produces a denumerably infinite number of equations to solve the indexed unknowns. The zeroth order unknowns satisfy the following equations

$$\overline{\mathbf{A}}_1 \boldsymbol{\alpha}_0 + \overline{\mu}_0 \overline{\mathbf{A}}_2 \boldsymbol{\alpha}_0 = \overline{\lambda}_0 \boldsymbol{\alpha}_0 \tag{37}$$

$$\overline{\mu}_0 = \overline{\mu}_c \tag{38}$$

which can be combined to the following single eigenvalue problem

$$\mathbf{A}\boldsymbol{\alpha}_0 = \overline{\lambda}_0 \boldsymbol{\alpha}_0 \tag{39}$$

where

$$\mathbf{A} \equiv \overline{\mathbf{A}}_1 + \overline{\mu}_c \overline{\mathbf{A}}_2 \tag{40}$$

(39) represents an matrix eigenvalue problem whose solutions can be obtained at least numerically by using one of the efficient standard methods. If we denote the eigenvalues and the normalized eigenvectors of the matrix $\mathbf{A}$ by ξ_i and $\mathbf{x}_i$ $(i = 1, ..., m)$ respectively then $\overline{\lambda}_0$ and $\boldsymbol{\alpha}_0$ can take a pair from these values, say ξ_k and $\mathbf{x}_k$. Therefore we write

$$\overline{\lambda}_0 = \xi_k, \qquad \boldsymbol{\alpha}_0 = \mathbf{x}_k \tag{41}$$

The first order perturbation equations can be obtained by setting the coefficients of the first power of ϵ at the both sides of each perturbation equation equal. The results are given below

$$(\mathbf{A} - \xi_k \mathbf{I})\boldsymbol{\alpha}_1 = \left(\overline{\lambda}_1 \mathbf{I} - \overline{\mu}_1 \overline{\mathbf{A}}_2\right)\mathbf{x}_k \tag{42}$$

$$\overline{\mu}_1 = \frac{1}{\mathbf{x}_k^T \mathbf{A}_2 \mathbf{x}_k} \mathbf{x}_k^T \left[\, \mathbf{A}_1 - \mu_u \mathbf{A}_2 - \overline{\mu}_c \mathbf{A}_2 \,\right] \mathbf{x}_k \tag{43}$$

where we have used (29), (31), (38), (40), and (41). The coefficient matrix in the left hand side of (42) is not invertible since its common left and right nullspaces are spanned by ξ_k and therefore not empty. This enforces the orthogonality of the right hand side to this nullspace. That is, the right hand side should be orthogonal to the eigenvector ξ_k. Thus we obtain the following new algebraic equation over two unknowns $\overline{\lambda}_1$ and $\overline{\mu}_1$

$$\overline{\lambda}_1 = \overline{\mu}_1 \mathbf{x}_k^T \overline{\mathbf{A}}_2 \mathbf{x}_k \tag{44}$$

where we have use the fact that $\mathbf{x}_k^T \mathbf{x}_k = 1$. The reorganization of the equation obtained after using (44) in (42) enables us to write

$$(\mathbf{A} - \xi_k \mathbf{I})\boldsymbol{\alpha}_1 = -\overline{\mu}_1 \left(\mathbf{I} - \mathbf{x}_k \mathbf{x}_k^T\right) \overline{\mathbf{A}}_2 \mathbf{x}_k \tag{45}$$

which neatly shows that the right hand side is orthogonal to the left hand side's nullspace. We can express the solution to this equation can be written as follows

$$\begin{aligned} \boldsymbol{\alpha}_1 \quad = \quad & -\overline{\mu}_1 (\mathbf{A} - \xi_k \mathbf{I})^{(-1,r)} \left(\mathbf{I} - \mathbf{x}_k \mathbf{x}_k^T\right) \overline{\mathbf{A}}_2 \mathbf{x}_k \\ & + c_1 \mathbf{x}_k \end{aligned} \tag{46}$$

where the superscript of the matrix $(\mathbf{A} - \xi_k \mathbf{I})^{(-1,r)}$ implies that the inverse of the matrix $(\mathbf{A} - \xi_k \mathbf{I})$ on its range is taken. This inverse is explicitly expressed as follows

$$(\mathbf{A} - \xi_k \mathbf{I})^{(-1,r)} \equiv \sum_{i=1, i \neq k}^{m} \frac{1}{\xi_i - \xi_k} \mathbf{x}_i \mathbf{x}_i^T \tag{47}$$

A careful investigation reveals the validity of the following identity

$$(\mathbf{A} - \xi_k \mathbf{I})^{(-1,r)} \left(\mathbf{I} - \mathbf{x}_k \mathbf{x}_k^T\right) \equiv (\mathbf{A} - \xi_k \mathbf{I})^{(-1,r)} \tag{48}$$

which allows us to rewrite (46) as follows

$$\boldsymbol{\alpha}_1 = -\overline{\mu}_1 (\mathbf{A} - \xi_k \mathbf{I})^{(-1,r)} \overline{\mathbf{A}}_2 \mathbf{x}_k + c_1 \mathbf{x}_k \tag{49}$$

c_1 appearing here and in (46) stands for an arbitrary constant and can be determined with respect to a given criterion which is generally the normalization of the vector whose perturbation series is under consideration. It can be taken just 0 for simplicity and the vector $\boldsymbol{\alpha}$ can be normalized via an appropriate scaling after the perturbation expansion is constructed. We follow this approach here. The choice of c_1 can also be realized in such a way that the total perturbative scheme's convergence and therefore its approximation quality

can be controlled. That procedure should be based on optimization. We keep it outside the goal of this paper.

After the determination, or more realistically, the choosing an appropriate value for c_1 everything becomes uniquely determined. We can proceed to the higher order exactly in the same way. If we are at the nth order perturbation terms what we need to determine is $\overline{\mu}_n$, $\overline{\lambda}_n$, and $\boldsymbol{\alpha}_n$. It is not hard to see that the equation for $\overline{\mu}_n$ depends on predetermined entities only, that is, its expression does not contain none of the unknowns undetermined yet. So it can be immediately evaluated in terms of the predetermined entities.

The equation derived from the perturbative expansion of the matrix eigenvalue problem produces an equation whose left hand side is $(\mathbf{A} - \xi_k \mathbf{I})\,\boldsymbol{\alpha}_n$ and its right hand side contains all predetermined entities together with the unknown $\overline{\lambda}_n$. As it has happened to be in the case of first order perturbation equations the left hand side matrix of this equation has the same nullspace as the first order case and enforces the right hand side to be orthogonal to this nullspace, that is, to the eigenvector $\mathbf{x}_k$. This determines the value of the unknown $\overline{\lambda}_n$ in terms of predetermined entities. The next step after this action is the inversion of the matrix $(\mathbf{A} - \xi_k \mathbf{I})$ on its range and to obtain the unknown vector $\boldsymbol{\alpha}_n$ within an uncertainty reflected by an additive term $c_n \mathbf{x}_k$ where the arbitrary constant c_n should be determined via same steps we mentioned about the determination of c_1 above. We do not intend to give the explicit expressions of higher order perturbation terms despite we use the second order terms in our implementations.

4 Implementations

In this part, using various types of functions, several numerical implementations are constructed to test the performance of our new algorithm and all the computations are done by using MuPAD [10] Computer Algebra System with 10-digit precision. The program codes are run under Linux (Ubuntu 7.10) Operating System.

In Table 1, the zeroth, first, and second order pertubative results for the constancy measurer are presented. The constant and therefore univariate truncation of HDMR increase in quality as the constancy measurer gets very close to 1. This is an expected result coming from the optimization on the weight function of the HDMR method. However, it can be clearly seen that the abovementioned convergence of the constancy measurer value is not enough to obtain a reasonably acceptable representation for the original function. Higher order perturbation terms are needed in the algorithm to get higher quality results. This undesiredly increases the cost of the algorithm.

Table 1: Obtained $\sigma_0^{(0)}$, $\sigma_0^{(1)}$, $\sigma_0^{(2)}$ values for different structure of functions

	$1-x^2$	$\sqrt{1-x^2}$	$ln(1-x^2)$	$sin(1-x^2)$
$\sigma_0^{(0)}$	0.5122	0.7870	0.2553	0.5391
$\sigma_0^{(1)}$	0.4627	0.7718	0.2468	0.4986
$\sigma_0^{(2)}$	0.7707	0.9163	0.2683	0.7104

5 Concluding Remarks

The basic philosophy of this work is to construct a new algorithm to represent a multivariate function by using rather low-variate HDMR components. For this purpose, HDMR expansion's weight function is optimized and the equations obtained through the optimization are solved by a perturbation expansion method.

We take the first three terms in perturbation expansion. The convergence of the expansion depends on the inputs of the scheme and on the value of $\overline{\mu}_e$. In many cases, the truncation level of the perturbation expansion may climb to very high numbers or the expansion may diverge while certain cases enable us to use only first few terms of the perturbation expansion. To understand how perturbation expansion converges we need to estimate the convergence radius of the perturbation series, that is, the least bound to the perturbation parameter ϵ to get a convergence with a desired speed.

We need computer based programming scripts and applications because of the usage of many terms in our algorithm. MuPAD can be used in these types of calculations and this allows us to take 100 or 200 number of expansion terms into consideration in this new method.

The need for the utilization of many expansion terms in the application of the algorithm brings the doubt to the convergence of the perturbation series. After certain numerical and theoretical investigations, we have seen that the perturbation series solutions to the matrix eigenvalue problem portion of the weight optimization problem may not converge even (26) is not taken into consideration and the μ value in (25) is assumed to be given.

There exists a weighted Rayleigh ratio in the structure given in (26). $\mathbf{A}_2$ is a positive definite weight matrix. Hence, the relation (26) bounds the possible values of μ both from below and from top. We could use μ as the perturbation parameter also. The investigations show us that the employment of μ as a pertur-

bation parameter does not guarantee the convergence in that perturbation series for all possible μ values permitted by the bounds of (26). This situation enforces us to perturb the matrix of the eigenvalue problem via different ways and gives different perturbation expansions for different subintervals of μ. These expressions depending on μ are then inserted into the equation (26) to determine algebraic equations for μ. The piecewise nature of the results are unpleasent things for continuity and therefore analyticity. One other important thing is the convergence analysis which may be quite difficult because of the nonlinearity in the equations.

The perturbation expansion we have proposed in this work has not a linear structure due to the Rayleigh quotient in the equation for μ and it is quite hard to prove whether the convergence domain for ϵ includes the value 1 or not. The absence of the convergence warranty for $\epsilon = 1$ makes the method quite unpleasent. Hence, we have come to the decision that an attractive method based on the perturbation expansion cannot be developed in this manner. We need some other considerations to develop convergent and efficient perturbation expansion.

Acknowledgements: The second author is grateful to Turkish Academy of Sciences and both authors thank to WSEAS, for their supports.

References:

[1] I.M. Sobol, Sensitivity Estimates for Nonlinear Mathematical Models, *Mathematical Modelling and Computational Experiments (MMCE)*, 1, 1993, No.4.407.

[2] H. Rabitz and Ö. Alış, General Foundations of High Dimensional Model Representations, *J. Math. Chem.* 25, 1999, pp. 197–233.

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

[4] 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.

[5] M. Demiralp, High Dimensional Model Representation and its Application varieties, *Mathematical Research* 9, 2003, pp. 146–159.

[6] 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.

[7] 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.

[8] B. Tunga and M. Demiralp, Hybrid High DimensionalModel Representation Approximants and Their Utilization in Applications, Mathematical Research, 2003, VOL 9, pp 438-446.

[9] B. Tunga and M. Demiralp, A Novel Hybrid High Dimensional Model Representation HHDMR) Based on the Combination of Plain and Logarithmic High Dimensional Model Representations, 157-161, 29-31 December, 2007, Cairo, Egypt 2007.

[10] W. Oevel, F. Postel, S. Wehmeier and J. Gerhard, The MuPAD Tutorial, Springer, 2000.