Abstract

In many areas of science and technology, there is a need for effective procedures for approximating multivariate functions. Sparse grids and cut-HDMR (High Dimensional Model Representation) are two alternative approaches to such multivariate approximations. It is therefore interesting to compare these two methods. Numerical experiments performed in this study indicate that the sparse grid approximation is more accurate than the cut-HDMR approximation that uses a comparable number of known values of the approximated function unless the approximated function can be expressed as a sum of high order polynomials of one or two variables.

Keywords: Sparse Grids, Approximation, Numerical experiments, Metamodelling, Curse of dimensionality

Streszczenie

W wielu obszarach nauki i technologii potrzebne są efektywne metody aproksymacji funkcji wielu zmiennych. Sieci rzadkie i cut-HDMR (High Dimensional Model Representation) są dwoma alternatywnymi podejściami do aproksymacji funkcji wielu zmiennych. Interesujące jest zatem porównanie tych dwóch metod. Eksperymenty numeryczne przeprowadzone w ramach niniejszych badań wskazują, że aproksymacja sieciami rzadkimi jest bardziej dokładna niż aproksymacja cut-HDMR wykorzystująca porównywalną liczbę znanych o ile aproksymowana funkcja nie może być wyrażona jako suma wielomianów wysokiego stopnia jednej lub dwóch zmiennych.

Słowa kluczowe: Sieci Rzadkie. Aproksymacja, Eksperymenty numeryczne, Metamodelowanie, Przekleństwo wymiarowości

* Institute of Teleinformatics, Faculty of Physics, Mathematics and Computer Science, Cracow University of Technology, Poland; mbaran@pk.edu.pl.
1. Introduction

The approximation of multivariate functions is a remarkably hard problem due to the so-called ‘curse of dimensionality’ [1]. However, the effective approximation of high-dimensional functions is the only solution in numerous practical problems from virtually all branches of science and technology. In particular such an approximation is an essential element of the so-called metamodelling [2, 3] (see also [4] for a discussion of metamodelling of high-dimensional problems, [5] for a comparison of a few metamodelling techniques and [6] for an example of the usage of sparse grids in metamodelling).

Examples of the application of high-dimensional approximation in science and engineering include ionosphere modelling [7], quantum mechanics [8], materials science [9], structural engineering [10], electrochemistry [11] and nuclear reactor modelling [12].

Sparse grids offer a method of function approximation where instead of one dense grid, we have a number of sparser grids and a linear combination is used [13]. The method is also known under other names such as the (discrete) blending method [14], the Boolean method [15] and hyperbolic cross approximation [16].

HDMR (High Dimensional Model Representation) approximation is a different approach that hinges on the fact that many high-dimensional functions can be efficiently approximated by sums of low-dimensional functions. The concept is attributed to Sobol [17]. The method is described in [18, 19]. Reference [20] describes many variants of HDMR approximation.

This paper compares the above two methods for the approximation of multivariate functions. In Section 2, a basic theory of sparse grid and cut-HDMR methods is described. Section 3 describes the performed numerical experiments. Section 4 summarizes the obtained results.

2. Theory of sparse grids and cut-HDMR

2.1. Sparse grids

Let us start with a one-dimensional interpolation. Consider a function \( f : [0,1]^M \to \mathbb{R} \). We need a sequence \( \{U^i\}_{i=1}^\infty \) of interpolating operators, each one providing a better approximation than the previous one. The formula for operator \( U^i \), which interpolates on nodes \( \{x_1, x_2, \ldots, x_m\} \), can be written as

\[
U^i(f)(x) = \sum_{j=1}^m f(x_j^i) a^i_j(x).
\]

Functions \( a^i_j(x) \) depend on the interpolation nodes and interpolation type. For Lagrange interpolation, the functions are given by [21]:

\[
a^i_j(x) = \prod_{1 \leq k \leq m, k \neq j} \frac{x-x_k}{x_j-x_k}.
\]
An $M$-dimensional approximation corresponds to the tensor product of operators $U^h_i, U^j_k, \ldots, U^M_p$ as follows:

$$U^h \otimes \cdots \otimes U^M_p \left( f \right)(x_1, x_2, \ldots, x_M) = \sum_{j_1=1}^{m_1} \cdots \sum_{j_M=1}^{m_M} f \left( x^h_{j_1}, \ldots, x^M_{j_M} \right) a^h_{j_1} \cdots a^M_{j_M} \left( x_M \right) \quad (3)$$

(see [22]). The interpolating operator for a $k$th variable is $U^k_k$. The calculation of the interpolant requires computing the interpolated function at $m_1 m_2 \cdots m_M$ nodes. In polynomial or spline interpolation, we select a single grid corresponding to a single tensor product of operators. In sparse grid approximation, we combine multiple grids.

The central idea in sparse grid approximation is Smolyak’s formula [13]. This formula represents a linear combination of interpolants on many grids. The sparse grid interpolation operator is defined by a linear combination of operators from Eq. (3):

$$A(q, M) = \sum_{q-M+1 \leq |i|=q} (-1)^{|i|} \binom{M-1}{q-|i|} U^h \otimes \cdots \otimes U^M_p. \quad (4)$$

The operator (4) has two arguments: the first ($q$) describes the density of the grid and the second ($M$) is the number of variables. The sum of components of index $i = \{i_1, i_2, \ldots, i_M\}$ is denoted by $|i| = \sum_{j=1}^{M} i_j$. Figure 1 shows two examples of two-dimensional sparse grids.

One can choose different approximation spaces for the sparse grid approximation. The simplest choice is a space constructed from so-called hat functions [23] but this space is rarely used in practice. Instead, commonly used sparse grids are based on polynomial or piecewise-polynomial functions.

One important feature of all sparse-grid methods is that one-dimensional basis functions $\psi_{ij} \left( x_j \right)$ are combined into $M$-dimensional basis functions $\psi_i \left( x \right)$ defined as:

$$\psi_i \left( x \right) = \prod_{j=1}^{M} \psi_{ij} \left( x_j \right). \quad (5)$$

In Eq. (5), $j$ is the number of independent variables and $i_j$ is a parameter to distinguish different one-dimensional functions of that variable. A sparse grid approximant is constructed as a linear combination of the multivariate basis functions $\psi_i \left( x \right)$ for different values of $i = \{i_1, i_2, \ldots, i_M\}$.

A motivating feature of sparse grids with polynomial interpolation is the fact that formula $A(M + k, M)$ exactly reproduces multivariate polynomials up to order $k$ [22]. Full grid approximation by operators from Eq. (3) on the other hand, exactly reproduces monomial $x_{1}^{m_1} x_{2}^{m_2} \cdots x_{M}^{m_M}$ of order $m_1 m_2 \cdots m_M$ but not monomial $x_{1}^{m_1+1}$ of order $m_1 + 1$. 
Fig. 1. Two-dimensional sparse grids with \( q \) equal to 3 (a) and 7 (b). Both examples use equidistant one-dimensional nodes.

### 2.2. Cut-HDMR approximation

An approximated \( M \)-variate function \( f(x) \) can be written as a sum of the constant term, functions of one variable, etc.

\[
f(x) = f_0 + \sum_{i=1}^{M} f_i(x_i) + \sum_{1 \leq i < j \leq M} f_{i,j}(x_i, x_j) + \ldots + f_{1,2,\ldots,M}(x_1, x_2, \ldots, x_M).
\]  

The choice of functions \( f_0, f_i, f_{i,j} \) etc. is not unique. By retaining only a few initial terms in the expansion (6) one obtains an HDMR approximation. The maximum dimension of the domain of functions used in the approximation is called the order of the approximation, for example, the first-order approximation consist only of a constant \( f_0 \) and one-dimensional functions \( f_i \).

In the cut-HDMR variant of the method, a reference point \( r \) is selected and used to determine the terms of expansion (6) in the following way:

\[
f_0 = f(r),
\]

\[
f_i(x_i) = f(x_1, \ldots, x_{i-1}, r_i, x_{i+1}, \ldots, x_M) - f_0,
\]

\[
f_{i,j}(x_i, x_j) = f(x_1, \ldots, x_{i-1}, r_i, x_{i+1}, \ldots, x_j, x_{j+1}, \ldots, x_M) - f_i(x_i) - f_j(x_j) - f_0,
\]

etc.

For analytic functions \( f \) this approximation can be compared to the multidimensional Taylor expansion at point \( r \):

\[
f(x) = f(r) + \sum_{i=1}^{M} \frac{\partial f(r)}{\partial x_i}(x_i - r_i) + \sum_{i=1}^{M} \sum_{j=i+1}^{M} \frac{\partial^2 f(r)}{\partial x_i \partial x_j}(x_i - r_i)(x_j - r_j) + \ldots.
\]
Regrouping the terms in Eq. (10) gives expressions for the cut-HDMR terms. As each cut-HDMR term corresponds to many Taylor expansion terms, cut-HDMR usually offers a better approximation than the Taylor expansion. The cut-HDMR approximation is, like the Taylor expansion, local. However, in certain subspaces, the approximation is exact (i.e. there is no error). They are called cut subspaces and depend solely on the expansion point (cut point, anchor point) and the order of the expansion. In agreement with Eqs. (7) to (9), first order cut subspaces are straight lines of the form
\[ \{(r_1, \ldots, r_{i-1}, x_i, r_{i+1}, \ldots, r_M) : x_i \in \mathbb{R} \}, 1 \leq i \leq M \] (11)
and second order cut subspaces are planes described by
\[ \{(r_1, \ldots, r_{i-1}, x_i, r_{i+1}, \ldots, r_j, x_j, r_{j+1}, \ldots, r_M) : x_i, x_j \in \mathbb{R} \}, 1 \leq i < j \leq M \] (12)

For the purpose of representing a cut-HDMR expansion on a computer, a method of interpolating the low-variate functions is necessary. Typically, first degree spline interpolation is used. The approximated function is calculated on grids spanning the cut subspaces. This is an improvement over interpolation in the whole \( \mathbb{R}^M \), as fewer values of the approximated function need to be known. The number \( N \) of values of function \( f \) that need to be stored decreases from \( K^M \) (assuming \( K \) interpolation nodes for each independent variable), to the value
\[ N = \binom{M}{l} (K-1)^{l-1} + \binom{M}{l-1} (K-1)^{l-2} + \ldots + \binom{M}{1} (K-1) + 1 \] (13)
for an \( l \)th order expansion. For each \( l \) the value \( N \) given by Eq. (13) is a polynomial in \( M \).

Equation (13) can be obtained by counting the number of points that lie in an \( l \)th order cut subspace, but not in an \( l-1 \)st order cut subspace, adding points that lie in an \( l-1 \)st order cut subspace, but not in an \( l-2 \)nd order cut subspace etc. In this way no point is counted twice.

The choice of the cut point \( r \) is an important issue. Wang [24] proposed an automatic method of selecting the best cut point. A low-discrepancy sequence of points \( \{x^i : x^i \in [0,1]^M \}^P_{i=1} \) is selected and each point is taken tentatively as a cut point of the HDMR decomposition. The error of the expansion is calculated as
\[ e(f, f_L) = \frac{1}{\sigma^2(f)} e_2(f, f_L) \] (14)
where:
\[ f \] – the approximated function,
\[ f_L \] – the approximant.

The point with the lowest error is finally selected as the cut point. The variance \( \sigma^2(f) \) of function \( f \) is defined as [25]:
\[ \sigma^2(f) = \int_{[0,1]^M} (f(x))^2 \, dx - \left( \int_{[0,1]^M} f(x) \, dx \right)^2 \] (15)
and the function \( e_2 \) is given by
\[ e_2(f, f_L) = \int_{[0,1]^M} (f(x) - f_L(x))^2 \, dx. \] (16)

It can be observed that since the variance is a constant positive number that does not depend on the approximation, the function \( f_L \) minimizing \( e_2(f, f_L) \) minimizes \( e(f, f_L) \) as well.

### 2.3. Theoretical comparison of sparse grids and cut-HDMR

Sparse grids and cut-HDMR method use two different approaches for the approximation of multivariate functions. Both of these methods try to overcome the curse of dimensionality. The cut-HDMR method approximates a given function by a number of low-dimensional functions. The sparse grids method combines the results of interpolation using a number of grids to give a better approximation.

The methods utilize different assumptions about an approximated function. The sparse grids method assumes that high-order terms of the Taylor expansion of an approximated function are negligible. Cut-HDMR approximation assumes that terms of the Taylor expansion involving more than a few (typically one or two) variables are negligible.

### 3. Numerical experiments

A number of numerical experiments were carried out to compare the sparse grid approximation with the cut-HDMR approximation. Six functions defined on the cube \([0, 1]^M\) were selected for the experiments. Results presented below refer to the following functions \( f_1 \) to \( f_4 \):

\[ f_1(x) = \frac{\log\left(\sum_{i=1}^{M} 2x_i + 1\right)}{\log(2M + 1)}, \] (17)

\[ f_2(x) = \sum_{i=1}^{M} (2x_i - 1)^{20}, \] (18)

\[ f_3(x) = \sum_{i=1}^{M} \sum_{j=1}^{M} (2x_i - 1)^{10}(2x_j - 1)^{10}, \] (19)

\[ f_4(x) = e^{2x_{M-1}}x_1 + \sum_{i=1}^{M-1} e^{2x_{i-1}}x_{i+1}. \] (20)

For other functions, similar results were obtained.

The above test functions have been selected with the aim of objectively test both approximation methods, without favouring any one of them. Assuming the approximated function is analytical, sparse grids behave poorly when there are high degree terms in the Taylor expansion of the approximated function at a given point. The functions \( f_1 \) and \( f_4 \) fulfill
these requirements. On the other hand, cut-HDMR effectively approximates high degree terms as long as they are monomials of no more variables than the order of the expansion (functions $f_2, f_3,$ and $f_4$).

A few different one-dimensional node placements were tested for sparse grids:

- equidistant nodes $x_k = \frac{k}{m-1}, k = 0, 1, \ldots, m-1$,
- extrema of the Chebyshev polynomials together with endpoints
  \[ x_k = 0.5 - 0.5 \cos\left(\frac{\pi k}{m-1}\right), k = 0, 1, \ldots, m-1, \]
- roots of the Legendre polynomials $P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n}\left[(x^2 - 1)^n\right]$ mapped to the interval $[0, 1]$.

The first node placement was selected for its simplicity; the second, because it eliminates the Runge effect [26]. Their implementation was obtained from the code of the TASMANIAN sparse grid package [27]. The values of the parameter $q$ in Eq. (4) that were used are 3, 4, 5, 6, 7, 8 and 9.

Figure 2 shows the relationship between the absolute approximation error calculated as:

\[ ERR(f, f_L, G) = \max_{x \in G} \left| f(x) - f_L(x) \right|, \tag{21} \]

and the number $N$ of points at which the function needs to be calculated. Grid $G$, which is the third argument of the error function (21), is a Cartesian product of $M$ sets of $k$ equally spaced points between 0 and 1:

\[ G(M, k) = \prod_{n=1}^{M} \left\{ 0, \frac{1}{k-1}, \ldots, \frac{k-2}{k-1}, \frac{k-1}{k-1} \right\}, \tag{22} \]

In the experiments, $M = 4$ and $k = 19$ were assumed. The first and second order cut-HDMR approximations were used. The error (21) was estimated on grid $G(4, 8)$. The cut-HDMR approximation used a first order spline interpolation or a Hermitian piecewise cubic interpolation employing first function derivatives approximated by three-point finite differences to represent functions of one or two arguments.

The method proposed by Wang [24] was used to determine the best placement of the cut points by minimizing expression (16). Sobol sequence [28] was used both for picking up candidates for the cut points and for the Quasi-Monte Carlo integration needed in Eq. (16). In both cases, 1000 points from the Sobol Sequence were taken. The coefficients needed for calculation of the Sobol sequence were obtained from the web page http://web.maths.unsw.edu.au/~fkuo/sobol/new-joe-kuo-6.21201 (accessed 2014-04-01). The details of the Sobol sequence generation can be found in [29].

Plot (a) in Fig. 2 shows errors for a function that cannot be expressed as a sum of functions of at most two arguments. As a result, the error of the cut-HDMR expansion reaches a minimum of approximately 0.11 and cannot become lower. At the same time, the sparse grids can achieve much lower errors.
Fig. 1. Absolute approximation error as dependent on the number $N$ of known values of function $f$ employed by the approximation. Plot (a) – function $f_1$, (b) – function $f_2$, (c) – function $f_3$, (d) – function $f_4$. Notation: (■) – sparse grids with equidistant nodes and first-order spline interpolation, (◊) – sparse grids with equidistant nodes and third-order spline interpolation, (○) – sparse grids with nodes at the extrema of Chebyshev polynomials and with Lagrange interpolation, (●) – sparse grids with nodes at the roots of Legendre polynomials and with Lagrange interpolation, (Δ) – second order cut-HDMR with first-order spline interpolation, (▲) – second order cut-HDMR with piecewise third-order polynomial interpolation, (♦) – first order cut-HDMR with piecewise third-order polynomial interpolation.
Plot (b) in Fig. 2 shows errors for a function that can be expressed as a sum of functions of one argument. In this case, the error of the expansion (6) truncated to order 2 is zero so that only the interpolation error remains and cut-HDMR approximation can be arbitrarily accurate. The variant with a higher order of interpolating polynomials shows significantly lower errors than the one that uses only first-order spline interpolation. The variant of sparse grids using third-degree piecewise polynomial interpolation has the lowest error of all sparse grid methods. The cut-HDMR decomposing only to functions of 0 and 1 variable (using third-degree piecewise polynomial interpolation) has the lowest error among all other methods.

Plot (c) in Fig. 2 shows errors for a function that can be expressed as a sum of functions of at most two variables. This function is characterized by high correlations between variables and is relatively fast-changing so that the error of the expansion (6) truncated to order 2 is zero again. The second-order cut-HDMR has the lowest errors in the range of $N$ where the logarithm of the error is lower than 0.

Plot (d) in Fig. 2 shows errors for a function that can be represented as a sum of functions of two arguments but is relatively slow-changing. In this case, two types of sparse grid approximation have the lowest errors.

4. Conclusions

In this study, two methods of approximating multivariate functions were compared – sparse grids and cut-HDMR. The comparison regarded the accuracy of both methods defined as the maximum absolute error of approximation. In most cases, sparse grids appear more accurate than the cut-HDMR method that employs a comparable number of known values of the approximated function. However, the cut-HDMR approximation has lower errors when the function is fast-changing and can be exactly represented as a sum of functions of at most one or two variables.

In conclusion, sparse grids are recommended over cut-HDMR for approximating multivariate functions. Cut-HDMR approximation should only be used when the sparse grids method cannot achieve desired accuracy using the assumed number of known values of the approximated function.

References


