Anchor Points Matter in ANOVA Decomposition

Zhongqiang Zhang, Minseok Choi and George Em Karniadakis

Abstract We focus on the analysis of variance (ANOVA) method for high dimensional approximations employing the Dirac measure. This anchored-ANOVA representation converges exponentially fast for certain classes of functions but the error depends strongly on the anchor points. We employ the concept of "weights per dimension" to construct a theory that leads to the optimal anchor points. We then present examples of a function approximation as well as numerical solutions of the stochastic advection equation up to 500 dimensions using a combination of anchored-ANOVA and polynomial chaos expansions.

1 Introduction

We consider an N-dimensional function f, which can be decomposed as

$$f(x_1,x_2,\cdots,x_N) = f_0 + \sum_{j_1=1}^{N} f_{j_1}(x_{j_1}) + \sum_{j_1<j_2}^{N} f_{j_1,j_2}(x_{j_1},x_{j_2}) + \cdots + f_{j_1,j_2,\cdots,j_N}(x_{j_1,j_2,\cdots,j_N}), \tag{1}$$

where f_0 is a constant, and f_S are $|S|$-dimensional functions, called the $|S|$–order terms. (Here $|S|$ denotes the cardinality of the index set S). This is the so-called ANOVA model. Here we consider the domain $I^N = [0,1]^N$, in a tensor-product form.

Zhongqiang Zhang
Division of Applied Mathematics, Brown University, Providence, RI 02912 e-mail: handyzang@gmail.com

Minseok Choi
Division of Applied Mathematics, Brown University, Providence, RI 02912 e-mail: minseok_choi@brown.edu

George Em Karniadakis
Division of Applied Mathematics, Brown University, Providence, RI 02912e-mail: gk@dam.brown.edu

The terms in the ANOVA decomposition are computed as follows

$$f_0 = \int_{[0,1]^N} f(\mathbf{x}) d\mu(\mathbf{x}). \tag{2a}$$

$$f_S = \int_{[0,1]^{N-|S|}} f(\mathbf{x}) d\mu(\mathbf{x}_{-S}) - \sum_{T \subset S} f_{\mathbf{T}}(\mathbf{x}_T). \tag{2b}$$

We note that there are several forms of ANOVA decomposition associated with different measures. Here we focus on the one using the Dirac measure, $d\mu(x) = \delta(x-c)\,dx$ ($c \in [0,1]$), which leads to the **anchored-ANOVA decomposition**. The point "c", which is often arbitrarily selected, is called the "anchor point". Another type is based on the Lebesgue measure, $d\mu(x) = \rho(x)\,dx$; this is the unanchored-ANOVA decomposition. See [5, 1] for details.

All the distinct ANOVA terms are mutually orthogonal with respect to the corresponding measure. Hence, for every term f_S with $S \subseteq \{1,2,\cdots,N\}$, we have

$$\int_{[0,1]} f_S(\mathbf{x}_S) d\mu(x_j) = 0, \qquad \text{if} \quad j \in S,$$

and

$$\int_{[0,1]^N} f_S(\mathbf{x}_S) f_T(\mathbf{x}_T) d\mu(\mathbf{x}) = 0, \qquad \text{if} \quad S \neq T.$$

The order at which we truncate the ANOVA model is called **effective dimension**, beyond which the difference between the ANOVA model and the truncated expansion in a certain measure is very small, see [10, 2, 12, 9]. It is not difficult to show that the variance of f can be a sum of variances of the ANOVA terms

$$\sigma^2(f) = \int_{[a,b]^N} f^2(\mathbf{x}) d\mathbf{x} - (\int_{[a,b]^N} f(\mathbf{x}) d\mathbf{x})^2 = \sum_{\emptyset \neq S \subseteq \{1,2,\cdots,N\}} \int_{[a,b]^{|S|}} f_S^2(\mathbf{x}_S) d\mathbf{x}_S. \tag{3}$$

or in compact form

$$\sigma^2(f) = \sum_{\emptyset \neq S \subseteq \{1,2,\cdots,N\}} \sigma_S^2(f). \tag{4}$$

The effective dimension of f in the superposition sense is the smallest integer d_s satisfying

$$\sum_{0<|S|\leq d_s} \sigma_S^2(f) \geq p\sigma^2(f), \tag{5}$$

where $S \subset \{1,2,\cdots,N\}$. This implies that we will ignore terms in the ANONA model corresponding to more than d_s interactions. The effective dimension is measured in the L^2–norm. Note that p is a proportionality constant with $0 < p < 1$ and close to 1, e.g., $p = 0.99$ in [2].

2 Weights and Effective Dimension

In order to obtain an estimate of the effective dimension, we adopt proper weights, which weight in some sense the contribution of each dimension. The concept of weights here is analogous to the concept employed in analyzing the Quasi Monte Carlo (QMC) method [11]. In particular, the idea is to define appropriate weights so that their minimization also leads to minimization of errors in QMC, see [8, 3]. In general, the weights should be in the interval of [0,1]. In addition, most of the weights should be less than one in order to have a low effective dimension for a nominally high-dimensional function.

Assuming a function in tensor product form, the weights in [11] were determined by the mean and the variance of the corresponding one-dimensional functions. This can be easily seen from the definition of the mean effective dimension [9]. Specifically, given a tensor product function

$$f(\mathbf{x}) = \prod_{k=1}^{N} f_k(x_k),$$

the mean and the variance of the function are

$$\begin{aligned}
\mu_k &= \int_0^1 f_k(x_k)\, dx_k < \infty, \quad k = 1, 2, \cdots, N, \\
\lambda_k^2 &= \int_0^1 \big(f_k(x_k) - \mu_k\big)^2\, dx_k < \infty, \quad k = 1, 2, \cdots, N.
\end{aligned}$$

The ANOVA terms and the corresponding variances are [9]:

$$\begin{aligned}
f_S &= \prod_{k \in S} \big(f_k(x_k) - \mu_k\big) \cdot \prod_{k \notin S} \mu_k, \\
\sigma_S^2(f_S) &= \prod_{k \in S} \lambda_k^2 \prod_{k \notin S} \mu_k^2.
\end{aligned} \tag{1}$$

Then, the weights γ_k's are defined as follows:

$$\gamma_k = \frac{\lambda_k^2}{\mu_k^2} \text{ if } \mu_k \neq 0 \text{ for } k = 1, 2, \cdots, N.$$

In the unanchored ANOVA (i.e., using the Lebesgue measure), the effective dimension has a more clear meaning. The truncation error, when the effective dimension is ν, by definition, is estimated as

$$\left\| f - \sum_{|S| \leq \nu} f_S \right\|_{L^2}^2 \leq (1-p)(\|f\|^2 - (\int_{I^N} f\, dx)^2),$$

where we use the equality $\|f\|^2 = (\int_{I^N} f\, dx)^2 + \sigma^2(f)$. Hence, we have

$$\left\| f - \sum_{|S|\le v} f_S \right\|^2 \le (1-p)(1-(\frac{\int_{I^N} f\,dx}{\|f\|})^2)\,\|f\|^2 = (1-p)(\int_{I^N} f\,dx)^2(\prod_{k=1}^{N}(1+\gamma_k)-1). \tag{2}$$

Remark 2.1 *From* (2)*, we have that*

$$\frac{\left\| f - \sum_{|S|\le v} f_S \right\|}{\|f\|} \le \sqrt{1-p}(1-\prod_{k=1}^{N}(1+\gamma_k)^{-1})^{\frac{1}{2}} < 0.1,$$

by choosing $p = 0.99$. In fact, when p is chosen as 0.99 the effective dimension is not always an integer. The estimate above corresponds to the worst case and, in fact, the error can be far better; see [9] for specific examples.

Remark 2.2 *From the definition of weights, we have that*

$$\left\| f - \sum_{|S|\le v} f_S \right\|^2 = (\int_{I^N} f\,dx)^2 \sum_{m=v+1}^{N} \sum_{|S|=m} \prod_{k\in S} \gamma_k.$$

According to (2)*,*

$$\sum_{m=v+1}^{N} \sum_{|S|=m} \prod_{k\in S} \gamma_k \le (1-p)(\prod_{k=1}^{N}(1+\gamma_k)-1). \tag{3}$$

As already mentioned, when a function is of low effective dimension, the dominating weights are much smaller than one. In fact, if $\mu_k \neq 0$ and $\gamma_k < 1$ for all $k = 1, 2, \cdots, N$, the mean effective dimension is [9]

$$d_s = \frac{\sum_{k=1}^{N} \frac{\gamma_k}{\gamma_k+1}}{1-\prod_{k=1}^{N} \frac{1}{\gamma_k+1}} = \frac{N - \sum_{k=1}^{N} \frac{1}{\gamma_k+1}}{1-\prod_{k=1}^{N} \frac{1}{\gamma_k+1}}. \tag{4}$$

While the previous discussion concerns the ANOVA version with Lebesque measure, it is by analogy that we can extend the concept of weights to the anchored-ANOVA as well. To this end, we define the weights using the L^∞–norm, as follows:

$$\gamma_k = \frac{\|f_k - f_k(c_k)\|_\infty}{|f_k(c_k)|}, \quad \text{when} \quad f(c) \neq 0. \tag{5}$$

Lemma 2.3 *Assuming that the anchored-ANOVA is truncated at the $\tilde{v}$–th order, and that $p_{\tilde{v}}$ satisfies*

$$\sum_{m=\tilde{v}+1}^{N} \sum_{|S|=m} \prod_{k\in S} \gamma_k = (1-p_{\tilde{v}})(\prod_{k=1}^{N}(1+\gamma_k)-1).$$

Then, the relative error in L^∞–norm can be estimated as

$$\frac{\left\| f-\sum_{|S|\le\tilde{\nu}} f_S\right\|_{L^\infty}}{\|f\|_{L^\infty}} \le (1-p_{\tilde{\nu}})(\prod_{k=1}^{N}(1+\gamma_k)-1)(\prod_{k=1}^{N}\frac{|f_k(c_k)|}{\|f_k\|_{L^\infty}}). \tag{6}$$

Also, for one-signed functions, if the anchored points $c=(c_1,c_2,\cdots,c_N)$ *are selected such that*

$$f_k(c_k)=\frac{1}{2}\max_{[0,1]} f_k(x_k)+\frac{1}{2}\min_{[0,1]} f_k(x_k).$$

Then, $\gamma_k=\left|\frac{\max_{[0,1]} f_k(x_k)-\min_{[0,1]} f_k(x_k)}{\max_{[0,1]} f_k(x_k)+\min_{[0,1]} f_k(x_k)}\right|$*, and it minimizes the weights defined in* (5)*.*

The minimized weights, in turn, minimize the error estimate in the last lemma.

Proof. Recalling the results from the ANOVA using Lebesgue measure with the same weights, we have

$$\begin{aligned}\frac{\left\| f-\sum_{|S|\le\nu} f_S\right\|_{L^\infty}}{\|f\|_{L^\infty}} &= \frac{\left\| f-\sum_{|S|\le\nu} f_S\right\|_{L^\infty}}{\prod_{k=1}^{N}|f_k(c_k)|}\frac{\prod_{k=1}^{N}|f_k(c_k)|}{\|f\|_{L^\infty}}\\ &\le \sum_{m=\tilde{\nu}+1}^{N}\sum_{|S|=m}\prod_{k\in S}\gamma_k(\prod_{k=1}^{N}\frac{|f_k(c_k)|}{\|f_k\|_{L^\infty}})\\ &\le (1-p_{\tilde{\nu}})(\prod_{k=1}^{N}(1+\gamma_k)-1)(\prod_{k=1}^{N}\frac{|f_k(c_k)|}{\|f_k\|_{L^\infty}}).\end{aligned}$$

This proves the error estimate. The following will complete the proof of how to minimize weights.

Suppose that f_k does not change sign over the interval $[0,1]$. Without loss of generality, let $f_k>0$. Denote the maximum and the minimum of f_k by M_k and m_k, respectively, and assume that $f_k(c_k)=\alpha_k M_k+(1-\alpha_k)m_k$ where $\alpha_k\in[0,1]$. Then

$$\|f_k-f_k(c_k)\|_\infty=\max(M_k-f_k(c_k),f_k(c_k)-m_k)=(M_k-m_k)\max(1-\alpha_k,\alpha_k),$$

and the weight γ_k is

$$\frac{\|f_k-f_k(c_k)\|_\infty}{|f_k(c_k)|}=\frac{(M_k-m_k)\max(1-\alpha_k,\alpha_k)}{\alpha_k M_k+(1-\alpha_k)m_k}.$$

Let us consider the function of $g(\alpha_k)=\frac{(1-y)\max(1-\alpha_k,\alpha_k)}{\alpha_k+(1-\alpha_k)y}$, where $\alpha_k\in[0,1]$, $y=\frac{m_k}{M_k}\in(0,1)$ and see how to choose α_k. Notice that

$$g'(\alpha_k)=\begin{cases}\frac{y-1}{(\alpha_k+(1-\alpha_k)y)^2}<0 & \text{if } \alpha_k\in(0,\frac{1}{2}),\\ \frac{(1-y)y}{(\alpha_k+(1-\alpha_k)y)^2}>0 & \text{if } \alpha_k\in(\frac{1}{2},1).\end{cases}$$

From this we know that $g(\frac{1}{2})$ reaches the minimum of $g(\alpha_k)$ with $\alpha_k\in(0,1)$. Then, $\alpha_k=\frac{1}{2},\quad \gamma_k=g(\frac{1}{2})=\frac{1-\frac{m_k}{M_k}}{1+\frac{m_k}{M_k}}<1.$

Actually, according to the definition of weights,

$$(\prod_{k=1}^{N}(1+\gamma_k)-1)(\prod_{k=1}^{N}\frac{|f_k(c_k)|}{\|f_k\|_{L^\infty}}) = \prod_{k=1}^{N}\frac{|f_k(c_k)|+\|f_k-f_k(c_k)\|_{L^\infty}}{\|f_k\|_{L^\infty}} - \prod_{k=1}^{N}\frac{|f_k(c_k)|}{\|f_k\|_{L^\infty}}.$$

If $\alpha_k > \frac{1}{2}$,

$$\begin{aligned}
&\prod_{k=1}^{N}(1+\gamma_k)-1)(\prod_{k=1}^{N}\frac{|f_k(c_k)|}{\|f_k\|_{L^\infty}}) \\
&= \prod_{k=1}^{N}\frac{\alpha_k M_k+(1-\alpha_k)m_k+(M_k-m_k)\max(1-\alpha_k,\alpha_k)}{M_k} - \prod_{k=1}^{N}\frac{\alpha_k M_k+(1-\alpha_k)m_k}{M_k} \\
&= \prod_{k=1}^{N}\big(2\alpha_k(1-\frac{m_k}{M_k})+\frac{m_k}{M_k}\big) - \prod_{k=1}^{N}\big(\alpha_k(1-\frac{m_k}{M_k})+\frac{m_k}{M_k}\big).
\end{aligned}$$

Hence, the first term in the last inequality increases faster than the last term, since $2\alpha_k(1-\frac{m_k}{M_k})+\frac{m_k}{M_k} > \alpha_k(1-\frac{m_k}{M_k})+\frac{m_k}{M_k}$ for $\alpha_k > \frac{1}{2}$. If $\alpha_k < \frac{1}{2}$,

$$\prod_{k=1}^{N}(1+\gamma_k)-1)(\prod_{k=1}^{N}\frac{|f_k(c_k)|}{\|f_k\|_{L^\infty}}) = 1-\prod_{k=1}^{N}\big(\alpha_k(1-\frac{m_k}{M_k})+\frac{m_k}{M_k}\big).$$

Thus $\alpha_k = \frac{1}{2}$ is the best choice when it minimizes the error estimate. Notice here the choice of $\alpha_k = \frac{1}{2}$ also minimizes the weight. This ends the proof.

Remark 2.4 *Weights and corresponding ancor points can also be defined in the L^1-norm using appropriate quadrature formulas, e.g. see [6].*

3 Numerical Examples

Here we present two examples, first in approximating a high-dimensional function and subsequently in solving the stochastic advection equation.

Example 1: We consider the Genz function [4] $f_5 = \prod_{j=1}^{N}\exp(-c_j|x_j-w_j|)$ with the parameters $c_j = \exp(-0.2j)$ and w_j following a uniform distribution.

$$\begin{aligned}
w = \ &(0.695106, 0.851463, 0.413355, 0.410178, 0.226185, \\
&0.7078, 0.478756, 0.183078, 0.0724332, 0.483279)
\end{aligned}$$

The centered point refers to $(\frac{1}{2},\frac{1}{2},\cdots,\frac{1}{2})$, while the optimal point is the point chosen according to the Lemma 2.3. Both results in table 1 demonstrate exponential accuracy in terms of the truncation dimension but using the optimal anchor points leads to accuracy close to three orders better than using the centered point.

Example 2: Next we consider the stochastic advection equation

truncation order	centered point	optimal point
1	6.6207×10^{-2}	3.7949×10^{-3}
2	5.2552×10^{-3}	8.8265×10^{-5}
3	2.3796×10^{-4}	1.2680×10^{-6}
4	6.2412×10^{-6}	1.1568×10^{-8}
5	9.0972×10^{-8}	6.6648×10^{-11}

Table 1 Error in the mean: $N = 10$.

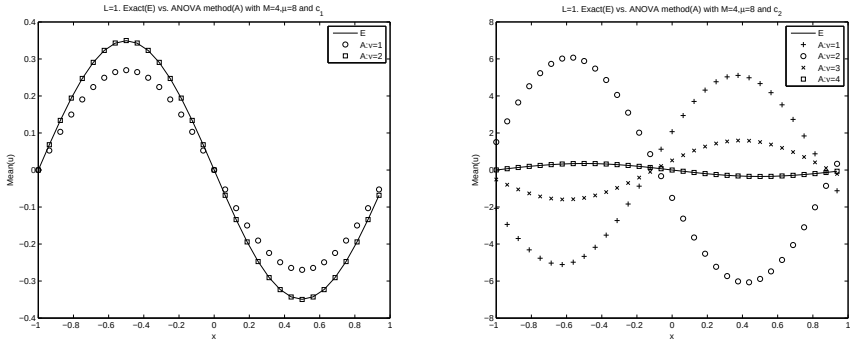

Fig. 1 Mean solution using the optimum anchor point c_1 (left) and a different point c_2 (right). Here $M = 4; L = 1$.

$$\frac{\partial u}{\partial t} + V(t;\xi)\frac{\partial u}{\partial x} = 0$$

in the interval $[-1,1]$ with periodic boundary conditions and initial condition $u(x,t=0) = sin(\pi(x+1))$. The advection velocity is a stochastic process with zero mean and is represented using a Karhunen-Loeve expansion, i.e. $V(t,\xi) = \sum_{k=0}^{M}\sqrt{\lambda_k}\phi_k(t)\xi_k$, with ξ_k being uncorrelated and also independent variables following a uniform distribution. The eigenpairs (λ_k,ϕ_k) are derived from the covariance kernel of the form $\exp[-|t_1-t_2|/L]$, where L is the correlation length. Here we consider three values of L corresponding to different truncations, i.e., $(L,M) = (1,4);(0.1,10);(0.005;500)$ selected so that 90% of the energy is captured by the coefficients of the truncated expansion. In the simulations we employ a Fourier-collocation in space and a probabilistic collocation method in random space using Legendre-chaos (8th-order).

In figure 1 we plot the mean solution at $t = 0.5$ in order to compare the effect of the anchor point on the convergence of the ANOVA expansion. We see that for the optimum point $c_1 = (0,0,\ldots,0)$ the solution converges to the exact solution when $v = 2$ but for another point $c_2 = (1,1,\ldots,1)$ the solution converges to the exact solution only if $v = M = 4$, i.e., for the full expansion. Here the exact solution is computed as in [7]. Using the optimum point we can now vary the correlation length L and produce accurate solutions in the high-dimensional space for small values of L and up to $M = 500$ dimensions as shown in figure 2.

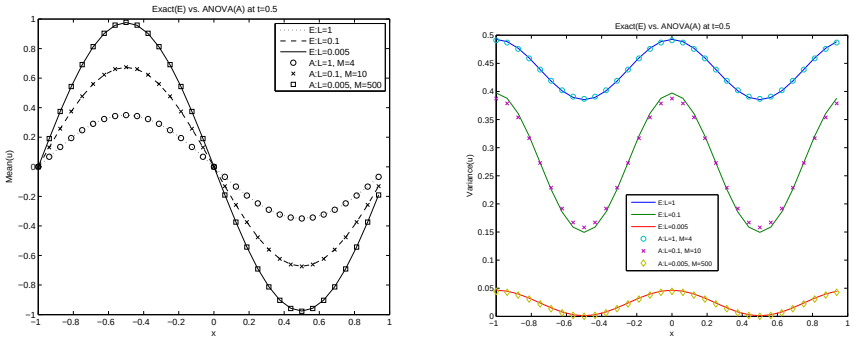

Fig. 2 Mean solution (left) and Variance (right) using the optimum anchor point c_1 for different values of the correlation length ($L = 1, 0.1, 0.005$) and corresponding truncation dimension ($\nu = 2, 2, 1$).

Acknowledgements This work was supported by an OSD/AFOSR MURI grant and also by a DOE Computational Mathematics grant.

References

1. Bieri, M. and Schwab, C.: Sparse high order FEM for elliptic sPDEs, Tech. Report 22, ETH, Switzerland, May 2008.
2. Caflisch, R.E., Morokoff, W. and Owen, A.: Valuation of mortgage-backed securities using Brownian bridges to reduce effective dimension, J. Comput. Finance, **1**, 27–46 (1997).
3. Dick, J., Sloan, I. H., Wang, X. and Wozniakowski, H.: Liberating the weights, Journal of Complexity, **20**, 593–623 (2004).
4. Genz, A.: A package for testing multiple integration subroutines, Numerical Integration: Recent developments, software and applications, Ed. Reidel, pp. 337–340, (1987).
5. Griebel, M.: Sparse grids and related approximation schemes for higher dimensional problems, In: Foundations of Computational Mathematics (FoCM05), Santander, L. Pardo, A. Pinkus, E. Suli, and M. Todd, eds., Cambridge University Press, 106–161, (2006).
6. Griebel, M. and Holtz, M.: Dimension-wise integration of high-dimensional functions with applications to finance, INS Preprint No. 0809, University of Bonn, Germany, (2009).
7. Jardak, M., Su, C.H. and Karniadakis, G.E.: Spectral polynomial chaos solutions of the stochastic advection equation, Journal of Scientific Computing, **17** , 319–338 (2002).
8. Larcher, G.,Leobacher, G. and Scheicher, K.: On the tractability of the Brownian bridge algorithm, Journal of Complexity, **19**, 511–528 (2003).
9. ———, The dimension distribution and quadrature test functions, Statistica Sinica, **13**, 1–17 (2003).
10. Paskov, S.H. and Traub, J.F.: Faster valuation of financial derivatives, Journal of Portfolio Management, **22**, 113–120 (1995).
11. Sloan, I.H. and Wozniakowski, H.: When are Quasi-Monte Carlo algorithms efficient for high dimensional integrals?, Journal of Complexity, **14** 1–33, (1998).
12. Wang, X. and Fang, K.-T.: The effective dimension and quasi-Monte Carlo integration, Journal of Complexity, **19**, 101–124 (2003).