Some Remarks on Monte Carlo Integration and the Curse of Dimensionality

Christian Lessig[1]

Otto-von-Guericke-Universität Magdeburg

Keywords: curse of dimensionality, Monte Carlo, convergence

2010 MSC: 42C15, 42C40

Image generation in computer graphics relies heavily on Monte Carlo integration,

$$I_n = \frac{|X|}{n} \sum_{i=1}^{n} f(x_i) \xrightarrow[n \to \infty]{} I = \int_X f(x) \, dx, \tag{1}$$

since the convergence rate is independent of the dimensionality (without loss of generality, we will assume a uniform probability density function). More formally, with the Chebychev inequality we have [1]

$$\mathrm{P} \left\{ |I_n - I| \geq \frac{1}{\sqrt{N}} \left(\frac{V(f)}{\delta} \right) \right\} \leq \delta \tag{2}$$

Hence, the convergence rate is $1/\sqrt{N}$ with no dependence on the dimensionality; in contrast to those for quadrature rules, e.g. for Bakhvalov's theorem.

As has been noted elsewhere before, however, this assumes that the variance $V(f)$ is independent of the dimensionality. To gain some intuition for what this implies, let us consider a simple function parametrized by dimension d,

$$f_d(x) = \prod_{i=1}^{d} a \, \sin(2\pi x_i), \quad x_i \in [0,1], a \in \mathbb{R}^+. \tag{3}$$

The functions $f_d(x)$ are obviously very well behaved, e.g. they are C^∞ and their modulus and that of each of its partial derivatives is bounded.

[1] lessig@isg.cs.uni-magdeburg.de

November 10, 2017

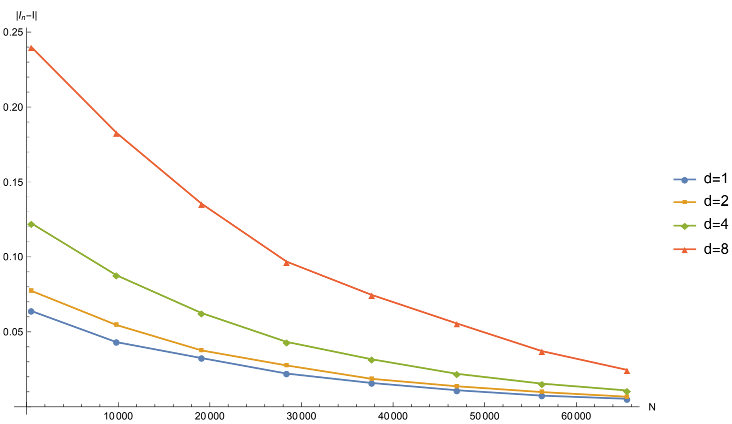

Figure 1: Curse of dimensionality for Monte Carlo integration of $f_d = \prod_{i=1}^{d} a\,\sin(2\pi x_i)$ with $a = 5/2$. Shown is the average error for 5,000 problem instances with N samples.

For a uniform probability density, the variance of a function $f(x)$ is

$$V(f) = \frac{1}{|X|} \int_X \big(f(x) - E[f]\big)^2 \, dx. \tag{4}$$

A little bit of algebra shows that for our family of functions we have

$$V(f_d) = \frac{a^{2d}}{2^d}. \tag{5}$$

Consequently, *for sufficiently large a the variance grows exponentially fast in the dimension*. Hence, also the convergence rate of Monte Carlo integration is no longer independent of d, as demonstrated in Fig. 1, and this despite $f_d(x)$ being a very well behaved function for any d.

The family $f_d(x)$ is the tensor product of $f(x) = a\sin(x)$, that is we use a quite simple model to describe the d-dependence. But, after all, our very model of d-dimensional space is typically as the product of 1-dimensional one. Eq. 5 shows, however, that the curse of dimensionality arises through this tensor product structure that leads to a scaling factor of a^d, that is the modulus of $f_d(x)$ grows with d. Using a scaling factor a_i that decays with d, such as by defining our family of functions as,

$$\bar{f}_d(x) = \prod_{i=1}^{d} \left(\frac{5}{2}\right)^{1/i} \sin(2\pi x_i), \qquad x_i \in [0,1] \tag{6}$$

enables one to again escape the curse of dimensionality, since in each additional dimension $\bar{f}_d(x)$ is more regular, in terms of its modulus, as in the previous one.

Although the above argument is based on a simple example, it demonstrates that also for Monte Carlo integration a convergence rate independent of the dimension is only attained when the function (family) becomes more and more regular as the dimension increases. But under this condition also quadrature rules can achieve a convergence rate independent of d. For classical C^r-smoothness this is shown by Bahvalov's theorem [2] and in the context of Quasi Monte Carlo methods this is described by weighted function spaces [3], which are conceptually similar to our modified family in Eq. 6, see for example [4] and references therein. Yet another avenue to escape the curse of dimensionality outside of Monte Carlo techniques is with sparse methods, either based on wavelets [5] or on other constructions [6], for which the sparsity grows sufficiently fast with the dimension. Compared to the deterministic bounds for quadrature rules, such as in [2], the analysis in Eq. 2 is also only probabilistic, i.e. without guarantees for an individual problem instance.

Our arguments are not new and similar ones have been made before, e.g. Donoho [7] and Bungartz and Griebel [6] discuss the assumptions that are required for $V(f)$ to be independent of the dimension.

References

[1] P. Hanrahan, E. Veach, D. Zorin, Mathematical Models for Computer Graphics Lecture Notes, Tech. rep., Stanford University (1997).
URL http://www-graphics.stanford.edu/courses/cs448-97-fall/notes.html

[2] N. S. Bahvalov, On Approximate Calculation of Multiple Integrals, Vestnik Moscow Univ. (1959) 3–18.

[3] I. H. Sloan, H. Woźniakowski, When Are Quasi-Monte Carlo Algorithms Efficient for High Dimensional Integrals?, Journal of Complexity 14 (1) (1998) 1–33. doi:10.1006/jcom.1997.0463.

[4] J. Dick, F. Pillichshammer, Digital Nets and Sequences: Discrepancy Theory and Quasi-Monte Carlo Integration, Cambridge University Press, 2010.

[5] W. Dahmen, C. Huang, G. Kutyniok, W.-Q. Lim, C. Schwab, G. Welper, Efficient Resolution of Anisotropic StructuresarXiv:1409.7908.
URL http://arxiv.org/abs/1409.7908

[6] H.-J. Bungartz, M. Griebel, Sparse Grids, Acta Numerica 13 (2004) 147–269.
URL http://journals.cambridge.org/abstract{_}S0962492904000182

[7] D. L. Donoho, High-Dimensional Data Analysis: The Curses and Blessings of Dimensionality, in: Mathematical Challenges of the 21st Century, American Mathematical Society, Los Angeles, CA, USA, 2000.
URL http://www-stat.stanford.edu/{~}donoho/Lectures/CBMS/Curses.pdf