ELSEVIER

Nuclear Physics B (Proc. Suppl.) 63A-C (1998) 988–990

NUCLEAR PHYSICS B
PROCEEDINGS
SUPPLEMENTS

Gaussian limits for discrepancies.

André van Hameren[a], Ronald Kleiss[b] and Jiri Hoogland[c]

[a]University of Nijmegen, Nijmegen, the Netherlands. Email: andrevh@sci.kun.nl

[b]University of Nijmegen, Nijmegen, the Netherlands. Email: kleiss@sci.kun.nl

[c]CWI, Amsterdam, the Netherlands. Email: Jiri@cwi.nl

It is well known that in the Monte Carlo method of integration of an s-dimensional integral the convergence to zero of the expected error can be improved by using point sets that are distributed more uniformly than truly random point sets. For a mathematical analysis of these point sets measures of non-uniformity are needed. These measures are called *discrepancies*. We investigated the asymptotic behaviour, in the limit of a large number of points, of probability distributions of these discrepancies, and were able to formulate a 'Central Limit Theorem' for these distributions.

1. Monte Carlo integration

We look at the integration problem on the s-dimensional hypercube $K = [0,1)^s$. In the Monte Carlo method of integration the integral $J[f]$ of a function f, integrable on K, is estimated by

$$\hat{J}_{X_N}[f] = \frac{1}{N}\sum_{k=1}^{N} f(x_k)\,. \tag{1}$$

Here, the point set $X_N = \{x_1,\ldots,x_N\} \in K^N$ has to be truly random, that is, in the space of all possible point sets, those used in the estimation of J have to be distributed with probability density $P(X_N) = 1$. The integration error that comes with this method converges to zero as $\sqrt{V/N}$, where V is the variance of f.

2. Quasi Monte Carlo integration

The $1/\sqrt{N}$-behaviour of the error can be improved by not using truly random point sets, but point sets in which the points are distributed more uniformly. This is done in the so-called Quasi Monte Carlo method of integration [1]. To measure the non-uniformity of point sets, a function of these point sets, called the *discrepancy*, is used. In principle different definitions of this measure, different discrepancies, can be used. Depending on the function or the class of functions to be integrated, it can often be shown that the integration error is directly related to this discrepancy, such that a low discrepancy corresponds with a small error. However, it is often a problem to actually compute the discrepancy for a given point set and estimate the actual integration error. Therefore one may ask what is the probability distribution of a discrepancy under truly random point sets; what is the probability that a 'truly random point set' has a given value of the discrepancy?

3. Classical discrepancy

Originally the discrepancy of a point set was defined as follows. First the *local discrepancy* $g(y)$ in a point $y \in K$ is defined as

$$g(y) = \frac{1}{N}\sum_{k=1}^{N}\prod_{\nu=1}^{s}\theta(y^\nu - x_k^\nu) - \prod_{\nu=1}^{s} y^\nu\,. \tag{2}$$

It calculates the fraction of points inside the hyper-rectangle defined by y and the origin, and subtracts the volume of that rectangle. Then the global discrepancy is a power of $g(y)$ integrated over K:

$$D_m = \int_K g(y)^m\,dx\,. \tag{3}$$

In literature often used are the L_2 *star-discrepancy* D_2 and the the *extreme* or *Kolmogorov discrepancy* $D_\infty = \lim_{k\to\infty}(D_{2k})^{1/2k}$.

PII S0920-5632(97)00964-X

Especially for the extreme discrepancy, many rigorous statements can be made [2]. In connection with numerical integration there is, for example, the Koksma-Hlawka inequality, which states that

$$\left|\hat{J}_{X_N}[f] - J[f]\right| \leq \mathcal{V}[f] D_\infty(X_N) \,, \tag{4}$$

where $\mathcal{V}[f]$ is the variation of f, a function of f independent of X_N. This inequality is one of the relations that illustrate the explicit relation between discrepancy and integration error. There are also results known about bounds on D_∞, such as the Roth bound

$$D_\infty(X_N) \geq C_s \frac{(\log N)^s}{N} \,, \tag{5}$$

where C_s depends on the dimension s only. This bound clearly gives space for an improvement of the $1/\sqrt{N}$ behaviour of the conventional Monte Carlo error.

4. Induced discrepancy

H. Woźniakoski realized that the L_2 star-discrepancy D_2 is equal to an average case complexity on the space of quadratically integrable functions [3]:

$$D_2(X_N) = \int \left|\hat{J}_{X_N}[f] - J[f]\right|^2 d\mu_W(f) \,. \tag{6}$$

The measure $d\mu_W$ on the function space is a variation of the Wiener measure in which the functions are pinned down at $x = (1, 1, \ldots, 1)$ rather than at $(0, 0, \ldots, 0)$. It is a Gaussian measure with the 2-point correlation function given by

$$\prod_{\nu=1}^{s} \min(1 - x_1^\nu, 1 - x_2^\nu) \,. \tag{7}$$

R. Kleiss suggested that this can be used for more general measures $d\mu$ to define discrepancies [4]. With an integration problem often comes a whole class of typical functions to be integrated. If this class can be characterized by a certain measure $d\mu$, then the discrepancy D_N of a point set X_N can be defined as

$$\frac{D_N(X_N)}{N} = \int \left|\hat{J}_{X_N}[f] - J[f]\right|^2 d\mu(f) \,. \tag{8}$$

The factor N is divided out in our definition such that this discrepancy can go to a finite nonzero limit as N goes to infinity.

To obtain a more useful expression of D_N, the 2-point correlation function can be used. First of all we assume that it can be written as

$$\int_L h(y; x_1) h(y; x_2) \, d\mu_L(y) \,, \tag{9}$$

where L is some space with some measure $d\mu_L$. If L is countable, the integral becomes a sum. Now D_N can be written as

$$D_N = \frac{1}{N} \sum_{k,l}^{N} \beta(x_k, x_l) \,, \quad \text{where} \tag{10}$$

$$\beta(x_k, x_l) = \int_L \omega(y; x_k) \omega(y; x_l) \, d\mu_L(y) \,, \tag{11}$$

$$\omega(y; x_k) = h(y; x_k) - \int_K h(y; x) \, dx \,. \tag{12}$$

This way the D_N can be recognized as the error made when $h(y; x)$ is integrated in x by X_N, averaged over L.

5. Asymptotic distributions of D_N

The probability distribution of D_N can best be found by using the generating function $G(z) = \langle e^{zD_N} \rangle$, where the average is over the truly random point sets X_N. The probability distribution H is then given by its Laplace transform

$$H(D_N = t) = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} e^{-zt} G(z) \, dz \,. \tag{13}$$

The generating function can be calculated by the calculation of the momenta $\langle D_N^m \rangle$ of D_N. Using expression (10) of D_N, we can see that the m^{th} moment will have at first a factor $1/N^m$. Because taking the average over truly random point sets means integrating over all the N points, the m^{th} power of the sum over x_k and x_l will give terms with combinatorial factors $N^{\underline{n}}$, with $n = 1, \ldots, m$.[1] This gives the opportunity to make an expansion in $1/N$. To get an asymptotic result in the limit of large N, only terms of

[1] $N^{\underline{n}} = N(N-1)(N-2)\cdots(N-n+1)$

low order in $1/N$ have to be taken into account in the calculation of $\langle D_N^m \rangle$. The involved combinatorics can be handled with Feynman diagrams [5]. To zeroth order in $1/N$ the result is that

$$G(z) = \exp\left(\sum_{k=1}^{\infty} \frac{(2z)^k}{2k} R_k\right) \quad , \text{ where} \tag{14}$$

$$R_k = \int_K \beta(x_1, x_2) \cdots \beta(x_k, x_1)\, dx_1 \cdots dx_k \; . \tag{15}$$

5.1. A simple example

As an example we could take a Gaussian measure on the expansion coefficients v_n in an expansion of the square integrable functions f in terms of an orthonormal set of basis functions $\{u_n\}$:

$$f(x) = \sum_n v_n u_n(x) \; , \text{ and} \tag{16}$$

$$d\mu(f) = \prod_n \frac{\exp(-v_n^2/2\sigma_n^2)}{\sqrt{2\pi\sigma_n^2}} \; . \tag{17}$$

To keep it simple, we take the first M strengths σ_n equal to one value σ, and the others equal to infinity. In practice this means that in (16) and (17) the sum and the product only have to be taken from $n = 1$ to $n = M$. The index n can be taken for the variable y in (9), and $h(n;x)$ can now be taken equal to $\sigma u_n(x)$, if we additionally assume that each of the basis functions integrates to zero. The discrepancy is equal to $\frac{\sigma^2}{N} \sum_{k,l} \sum_n u_n(x_k) u_n(x_l)$, and the generating function is $1/(1-2z\sigma^2)^{M/2}$. The probability distribution finally is

$$H(t) = \frac{t^{\frac{M}{2}-1} \exp(-\frac{t}{2\sigma^2})}{(2\sigma^2)^{\frac{M}{2}} \Gamma(\frac{M}{2})} \; . \tag{18}$$

6. A Law of Large Number of Modes

A large number of function measures, including the Wiener measure, can be written as in (17). Differences are in the set of basis functions, which for example don't have to integrate to zero, and in the strengths. All these measures, however, have in common that the quantity

$$\Gamma_{n,m} = \int_K \omega(n;x)\omega(m;x)\, dx \tag{19}$$

is a real symmetric matrix. Therefore it can be diagonalized, giving eigenvalues λ_n. It can now be shown [6] that in all these cases *a necessary and sufficient condition for the discrepancy distribution to approach a Gaussian distribution is that*

$$\frac{\max_n \lambda_n^2}{\sum_n \lambda_n^2} \to 0 \; . \tag{20}$$

The arrow in (20) refers to some limit in the definition of the function measure. In the case of the simple example this limit is that $M \to \infty$. In the case of the Wiener measure, the limit is that the number of dimensions $s \to \infty$. In many cases as well, the limit $s \to \infty$ yields a Gaussian.

The 'law' (20) can be interpreted as follows: for a Gaussian distribution to appear in a certain limit, the number of 'modes' λ_n has to go to infinity or be infinite, such that they all go to zero about equally fast. If a finite number of modes becomes more important than the others, the Gaussian limit is spoiled.

7. Outlook

The result presented in the previous section has been derived in the limit of $N \to \infty$. In practice, however, one needs to know around which value of N the asymptotic regime sets in, that is, one needs to know the next-to-leading contributions. This will be the subject of [7].

REFERENCES

1. H. Niederreiter, *Random number generation and Quasi-Monte Carlo methods,* (SIAM, 1992).
2. L. Kuipers and H. Niederreiter, *Uniform distribution of Sequences,* (John Wiley, 1974).
3. H. Woźniakoski, *Bull. AMS* **24** (1991) 185-194.
4. R. Kleiss, *Comp. Phys. Comm.* **71** (1992) 39-53.
5. J. Hoogland and R. Kleiss, *Comp. Phys. Comm.* **98** (1996) 111-127,
6. A. v. Hameren, J. Hoogland and R. Kleiss, preprint physics/9708014.
7. A. v. Hameren, J. Hoogland and R. Kleiss, (in preparation).