I.M.SOBOL

AVERAGE DIMENSION

INTRODUCTION

Functions $f(x)$ defined in the d-dimensional unit hyper cube H are considered: $x=(x_1,\ldots,x_d)$. Average dimensions were introdused in 2006 by A.Owen (who called them *mean dimensions*). Various examples suggest that the efficiency of quasi-Monte Carlo integration depends on the average dimension of the integrand.

MONTE CARLO METHOD

$$a = \ ?$$

$$E\eta = a$$

$$\frac{1}{N}\sum_{k=1}^{N}\eta_k \xrightarrow{P} a$$

If $Var(\eta) < \infty$, the probable error $\delta_N^{prob} = 0.6745\sqrt{\frac{Var(\eta)}{N}}$

MONTE CARLO ALGORITHM

$$\left.\begin{array}{l} \frac{1}{N}\sum_{k=1}^{N}\eta_k \xrightarrow{P} a \\ \eta = g\left(\gamma_1, \gamma_2, \ldots\right) \end{array}\right\}$$

If $g = g\left(\gamma_1, \ldots, \gamma_d\right)$ we say that the constructive dimension of MC algorithm is d.

$$\Gamma = \left(\gamma_1, \ldots, \gamma_d\right) \text{ u.d. in } H$$

$$a = Eg\left(\Gamma\right) = \int_H g\left(x\right)dx$$

$$\frac{1}{N}\sum_{k=1}^{N} g\left(\Gamma_k\right) \xrightarrow{P} \int_H g\left(x\right)dx$$

QUASI-MONTE CARLO

In 1916 H.Weyl introdused u.d. sequences of non random points.

Weyl theorem

If $P_1,\ldots,P_k,\ldots$ is u.d. in H and $g(x)$ is integrable and bounded then

$$\frac{1}{N}\sum_{k=1}^{N} g\left(P_k\right) \to \int_H g\left(x\right) dx$$

In MC algorithms with constructive dimension d random points Γ_k can be replaced by P_k

This approach is called quasi-Monte Carlo

TWO PROBLEMS

that remain unsolved:

(a) Best sequence $P_1,\ldots,P_k,\ldots$ (?)

(b) MC or QMC (?)

My talk is on problem (b)

Several words on problem (a)

Many practitioners prefer Sobol sequences.

e.g. Monte carlo methods in finance

(P. Jaeckel, 2002, P. Glasserman, 2004)

My best version: SOBOLSEQ16384 in

I.M. Sobol', D. Asotsky, A. Kreinin and S. Kucherenko, Construction and comparison of high-dimensional Sobol' generators, Wilmott Mag. 2011 (2011), n.56, 64–79

ANOVA DECOMPOSITION

Let $f(x) \in L_2(H)$ and $f_0 = \int\limits_H f(x)\,dx$

The following identity is called decomposition of $f(x)$

$$f(x) = f_0 + \hat{\sum} f_{i_1,\dots,i_s}\left(x_{i_1,\dots,i_s}\right) \text{ where } 1 \le i_1 < \cdots < i_s \le d \text{ and } 1 \le s \le d$$

Main requirement $\int\limits_0^1 f_{i_1,\dots,i_s}\, d\left(x_{i_p}\right) = 0$ for $1 \le p \le s$

$$f(x) = f_0 + \sum_i f_i(x_i) + \sum_{i<j} f_{i,j}\left(x_i, x_j\right) + \dots + f_{1,2,\dots,d}\left(x_1, x_2, \dots, x_d\right)$$

$$D_{i_1,\dots,i_s} = \int\limits_H f^2_{i_1,\dots,i_s}\,dx \text{ - variances}$$

$$D = \int\limits_H f^2(x)\,dx - f_0^2 \text{ - total variance}$$

$$D = \hat{\sum} D_{i_1,\dots,i_s}$$

SENSITIVITY INDICES

$$S_{i_1,\dots,i_s} = D_{i_1,\dots,i_s} \big/ D$$

$$\hat{\Sigma}\, S_{i_1,\dots,i_s} = 1$$

Average dimension of $f(x)$

$$\hat{d} = \hat{\Sigma}\, sS_{i_1,\dots,i_s}\,.$$ Obviously, $1 \le \hat{d} \le d$.

Examples

$$f(x) = f_0 + c\prod_{i=1}^{d}\left(x_i - \frac{1}{2}\right) \qquad \hat{d} = d$$

$$f(x) = f_0 + c\sum_{i=1}^{d}\left(x_i - \frac{1}{2}\right) \qquad \hat{d} = 1$$

Various examples show that Q-MC outplays MC if for the integrand $\hat{d} < 3$.

THE OWEN THEOREM

S_i corresponds to $f_i\left(x_i\right)$

$S_i^{tot} = \sum S_{i_1,\dots,i_s}$ where one of the $i_1,\dots,i_s$ is i.

$$0 \le S_i \le S_i^{tot} \le 1$$

Let $x' = \left(x_1,\dots,x_{i-1},t,x_{i+1},\dots,x_d\right)$

$$S_i^{tot} = \frac{1}{2D}\int_H \int_0^1 \left(f\left(x\right) - f\left(x'\right)\right)^2 dxdt$$

Owen theorem: $\hat{d} = \sum_{i=1}^{d} S_i^{tot}$