Chin. Ann. of Math.
11B:1(1990), 51—65.

NUMBER THEORETIC METHOD IN APPLIED STATISTICS***

WANG YUAN(王　元)*　FANG KAITAI(方开泰)**

(*Dedicated to the Tenth Anniversary of CAM*)

Abstract

This paper gives some applications of number-theoretic method (or quasi Monte Carlo method) for numerical evaluation of probabilities and moments of a continuous multivariate distribution over a special domain such as cube, ball, sphere, simplex, etc., where the uniformly distributed sets of points in such domains, which are useful in experimental design, simulation, geometry probability, etc., are suggested. Some applications of number-theoretic method in optimization are discussed also.

§1. Introduction

The problem of numerical evaluation of probabilities and moments is really a problem of numerical integration. The number-theoretic method (or quasi Monte Carlo method) for numerical evaluation of multiple integrals and for optimization is based on the theory of uniform distribution (*u. d.*). Let $K=[a_1, b_1]\times\cdots\times[a_s, b_s]$ be a rectangle of R^s, $\boldsymbol{b}=(b_1, \cdots, b_s)'$, $\boldsymbol{x}=(x_1, \cdots, x_s)'$, and $F(\boldsymbol{x})$ be a continuous monotone distribution function on K, which satisfies $F(\boldsymbol{b})=1$ and $F(\boldsymbol{x})=0$ whenever at least one of the x_i is a_i. Note that a_i' s and b_i's may be defined to be $-\infty$ and ∞ respectively. We use $\boldsymbol{x}\leqslant\boldsymbol{b}$ to denote that $x_i\leqslant b_i(i=1, \cdots, s)$. For a set of points $P=(\boldsymbol{x}_k,\ k=1, \cdots, n)$ in K and a rectangle $G=[a_1, x_1]\times\cdots\times[a_s, x_s]$, where $\boldsymbol{x}\leqslant\boldsymbol{b}$, let $N(P, G)$ be the number of P satisfying $\boldsymbol{x}_k\in G$, and let

$$\sup_G\left|\frac{N(P, G)}{n}-F(\boldsymbol{x})\right|=D_F(n, P).$$

$D_F(n, P)$ is called the F-discrepancy of P with respect to $F(\boldsymbol{x})$. If $P_n=(\boldsymbol{x}_1^{(n)}, \cdots, \boldsymbol{x}_{k_n}^{(n)})$ is a sequence in K such that $k_n\to\infty$ as $n\to\infty$. and if $D_F(k_n, P)=o(1)$ as $n\to\infty$, then P_n is called an F-uniformly distributed sequence. If $K=I^s$, where $I=[0, 1]$ and $F(\boldsymbol{x})$ is uniform distribution on I^s, i. e., $F(\boldsymbol{x})=x_1\cdots x_s$, then we omit the F in the above notations (cf. Weyl [12], Hlawka and Muck [5], and Niderreiter [9]).

Manuscript received March 15, 1989.

* Institute of Mathematics, Academia Sinica, Beijing, China.

** Institute of Applied Mathematics, Academia Sinica, Beijing, China.

*** This Work is Supported by the Chinese National Science Foundation and the Science Foundation of the Chinese Academy of Sciences.

If $P=\{x_k,\ k=1,\ \cdots,\ n\}$ is a set I^s with decrepancy $D(n,\ P)$ or $D(n)$ for simplify and $f(x)$ is a function of bounded variation in the sense of Hardy and Krause with total variation $V(f)$, then it is known that

$$\left|\int_{I^s} f(x)dx-\frac{1}{n}\sum_{k=1}^{n}f(x_k)\right|\leqslant V(f)D(n) \tag{1.1}$$

(See Koksma [7], Hlawka [4], Hua and Wang [6]).

Let D be a domain (for example, ball, sphere, simplex, etc.) in R^s. In this paper we shall pay more attention to numerical evaluation of

$$I=\int_D f(x)dv, \tag{1.2}$$

where dv is the volume element of D and D has a parameter representation. First by using a transformation a quadrature formula over D can be transferred to a quadrature formula over I^t, where t is the dimension of D. Another approach to this problem is to use a u. d. sequence in D. We shall start from a u. d. sequence in I^t. and then derive the u. d. sequences with respect to certain distribution functions, in particular, to some uniform distribution functions in some special domains: ball, sphere, simplex, etc., which are often useful in simulation, geometry probability, experimental design and many problems in statistics. More details are given in our next paper with the same title.

Another application of the u. d. sequences in D is in optimization. Let $f(x)$ be a continuous function on D, we want to find its global maximum M in D. There are many gradient methods for this kind of optimization problems (cf. Avriel [2]). Unfortunately, there appear only few cases that the global maximum can be reached, and we can obtain in usual a local maximum if the function f is not unimodal, and the dimension of D is large, for example, dimension of $D\geqslant 5$, because the solution, in general, depends on the choice of initial point. Therefore, we use the following algorithm to find an approximate value of M:

$$m_1=f(x_1),$$

$$m_{k+1}=\begin{cases} m_k, & \text{if } f(x_{k+1})\leqslant m_k,\\ f(x_{k+1}), & \text{if } f(x_{k+1})\geqslant m_k,\end{cases}$$

where $(x_1,\ x_2,\ \cdots)$ is a u. d. sequene in D, i. e., $P_n=\{x_1,\ \cdots,\ x_n\}$ is a u. d. sequence. After a large number n of steps, we may reasonably expect that m_n is close to M, if $f(x)$ satisfies some regularity conditions. We often use the following quantity to measure the uniformity of distribution of these points

$$d(n,\ D)=\max_{x\in D}\min_{1\leqslant k\leqslant n} d(x,\ x_k), \tag{1.3}$$

where $d(x,\ x_k)$ denotes the Euclidan distance of x and x_k. $d(n,\ D)$ is called the dispersion of the set $\{x_k,\ k=1,\ \cdots,\ n\}$. However one can show that if $D=I^s$, then

$$\sqrt{s}\,n^{-1/s}\leqslant d(n,\ D)\leqslant 2\sqrt{s}\,(D(n))^{1/s} \tag{1.4}$$

(cf. Zielinski [13] and Niederreiter [10]). This means that it is true that m_n is closed to the globle maximum M if n is large. In Section 4, we shall generalize the above result to some kind of D's and give some applications in statistics.

§ 2. Numerical Integration

Let D be a bounded domain in R^s. We are required to calculate the integral (1.2). Assume that tha dimension of D is s, $dv=\prod_{i=1}^{s} dx_i=dx$ and $D\subset I^s$. Then it may be simply suggested to use the following formula

$$I=\int_{I^s} f(x) I_D(x) dx,$$

where $I_D(x)$ is the index functisn of D (cf. Hua and Wang [6]). This will lead to a big error sometimes, since $f(x)I_D(x)$ may be discontinuous on the boundary of D. However, the domain D is often very special in statistics, so it is possible to reduce the integral over D to an integral over $I^t (t\leqslant s)$. More precisely, suppose that D has a representation

$$x_j=x_j(\varphi_1, \cdots, \varphi_t)=x_j(\boldsymbol{\varphi}),\ j=1, \cdots, s, \tag{2.1}$$

where $\boldsymbol{\varphi}\in I^t$, and that x_j, $j=1, \cdots, s$, have continuous derivatives with respect to φ_i, $i=1, \cdots, t$, over I. Let

$$\boldsymbol{T}=(\partial x_j/\partial\varphi_i),\ i=1,\cdots,t,\ j=1, \cdots, s,$$

and let

$$J(\boldsymbol{\varphi})=\det(\boldsymbol{T}\boldsymbol{T}')^{1/2}.$$

When $t=s$, $J(\boldsymbol{\varphi})$ is just the Jacobian of transfomation from x to $\boldsymbol{\varphi}$. Then we have

$$I=\int_D f(x)dv=\int_{I^t} f(x(\boldsymbol{\varphi}))J(\boldsymbol{\varphi})d\boldsymbol{\varphi}, \tag{2.2}$$

where $d\boldsymbol{\varphi}=\prod_{i=1}^{t} d\varphi_i$. Therefore a quadrature formula over I^t induces a quadrature formula over D. Denote by $v(D)$ the volume of D. Then

$$v(D)=\int_{I^t} J(\boldsymbol{\varphi})d(\boldsymbol{\varphi}). \tag{2.3}$$

Suppose further that $\varphi_1, \cdots, \varphi_t$ are independent, and

$$v(D)^{-1}J(\boldsymbol{\varphi})=\prod_{i=1}^{t} f_i(\varphi_i),$$

where $f_i(\varphi_i)$ is the density functions of φ_i, $i=1, \cdots, t$, and the corresponding distribution functions

$$F_i(x_i)=\int_0^{x_i} f_i(\varphi_i)d\varphi_i,\ i=1, \cdots, t$$

satisfying $F_i(0)=0$ and $F_i(1)=1$, $i=1, \cdots, t$. Let $F_i(x_i)=y_i$ and let $F_i^{-1}(y_i)$ denote the inverse function of $F_i(x_i)$, $i=1, \cdots, t$. Then

$$\int_{I^t} f(\varphi)J(\varphi)d\varphi = v(D)\int_{I^t} f(\boldsymbol{x}(\boldsymbol{F}^{-1}(\boldsymbol{y}))d\boldsymbol{y}), \tag{2.4}$$

where $\boldsymbol{F}^{-1}(\boldsymbol{y}) = (F_1^{-1}(y_1), \cdots, F_t^{-1}(y_t))'$, and $d\boldsymbol{y} = \prod_{i=1}^{t} dy_i$.

For a given set $\{\boldsymbol{b}_k = (b_{k1}, \cdots, b_{kt})', k=1, \cdots, n\}$ of I^t with discrepancy $D(n)$, we have a set $\{\boldsymbol{c}_k = F^{-1}(\boldsymbol{b}_k), k=1, \cdots, n\}$ which has F-discrepancy $D_F(n, \{c_k\}) = D(n)$ too, where $F(\boldsymbol{x}) = \prod_{i=1}^{t} F_i(x_i)$. Hence by (1.1), (2.2) and (2.4) we have

$$\begin{aligned}&\left|\int_D f(\boldsymbol{x})dv - v(D)\frac{1}{n}\sum_{k=1}^{n} f(\boldsymbol{x}(\boldsymbol{F}^{-1}(\boldsymbol{b}_k)))\right|\\ &\leqslant v(D)D_F(n, \{\mathbf{c}_k\})V(f(\boldsymbol{F}^{-1}))\\ &= v(D)D(n)V(f).\end{aligned} \tag{2.5}$$

By (2.2), (2.5), and the number theoretic method we thus may have the following two formulas for numerical evaluation of the multiple integral (1.2)

$$\int_D f(x)dv \cong \frac{1}{n}\sum_{k=1}^{n} f(\boldsymbol{x}(\boldsymbol{b}_k))J(\boldsymbol{b}_k) \tag{2.6}$$

and

$$\begin{aligned}\int_D f(\boldsymbol{x})dv &\cong \frac{1}{n}v(D)\sum_{k=1}^{n} f(\boldsymbol{x}(\mathbf{c}_k))\\ &= \frac{1}{n}v(D)\sum_{k=1}^{n} f(\boldsymbol{x}(\boldsymbol{F}^{-1}(\boldsymbol{b}_k))).\end{aligned} \tag{2.7}$$

Both formulas have the same order of accuracy.

Define a set of D by

$$P = \{\boldsymbol{x}_k = \boldsymbol{x}(\mathbf{c}_k), k=1, \cdots, n\}. \tag{2.8}$$

The volume $v(\varphi \leqslant \boldsymbol{y})$ of the domain in D defined by $\varphi \leqslant \boldsymbol{y}$ is equal to

$$v(\varphi \leqslant \boldsymbol{y}) = \int_{\varphi \leqslant \boldsymbol{y}} J(\varphi)d\varphi = v((D)\prod_{i=1}^{t} F_i(y_i),$$

so that

$$v(\varphi \leqslant \boldsymbol{y})/v(D) = \prod_{i=1}^{t} F_i(y_i).$$

Therefore, if we want the set P to be scattered uniformly over D, or that the ratio between the number $N(\varphi \leqslant \boldsymbol{y})$ of P lying in the domain defined by $\varphi \leqslant \boldsymbol{y}$ and n is approximately equal to the ratio between $v(\varphi \leqslant \boldsymbol{y})$ and $v(D)$, we should take the set $\{\mathbf{c}_k, k=1, \cdots, n\}$ with lower F-discrepancy in I^t. Since $\{\mathbf{c}_k, k=1, \cdots, n\}$ has F-discrepancy $D(n)$, we have

$$\begin{aligned}\sup_{\boldsymbol{y}\in I^t}\left|\frac{N(\varphi \leqslant \boldsymbol{y})}{n} - \frac{v(\varphi \leqslant \boldsymbol{y})}{v(D)}\right| &= \sup_{\boldsymbol{y}\in I^t}\left|\frac{N(\varphi < \boldsymbol{y})}{n} - F(\boldsymbol{y})\right|\\ &= D(n).\end{aligned} \tag{2.9}$$

We thus suggest an algorithm for obtaining a set P of D that is scattered uniformly in D from a known set with lower discrepancy in I^t. Now we give some examples.

Example 1. The domain D is a simplex $A_s = \{\boldsymbol{x}: 0 \leqslant x_s \leqslant x_{s-1} \leqslant \cdots \leqslant x_1 \leqslant 1\}$. Then D has a representation

$$x_j=\varphi_1\cdots\varphi_j,\ j=1,\ \cdots,\ s,$$

where $\boldsymbol{\varphi}\in I^s$. We have

$$J(\varphi)=\prod_{i=1}^{s-1}\varphi_i^{s-i},$$

and

$$v(A_s)=\int_{I^s}J(\boldsymbol{\varphi})d\boldsymbol{\varphi}=\prod_{i=1}^{s-1}\frac{1}{s-i+1}=1/s!.$$

Therefore

$$f_i(\varphi_i)=(s-i+1)\varphi_i^{s-i},\ i=1,\ \cdots,\ s$$

are density functions over I^s with corresponding distribution functions

$$F_i(x_i)=\int_0^{x_i}f_i(\varphi_i)d\varphi_i=x_i^{s-i+1}.$$

For a given set $\{\boldsymbol{b}_k,\ k=1,\ \cdots,\ n\}$ in I^s with discrepancy $D(n)$, we have a set

$$\boldsymbol{c}_k=\boldsymbol{F}^{-1}(\boldsymbol{b}_k)=(b_{k1}^{1/s},\ b_{k2}^{1/(s-1)},\ \cdots,\ b_{k,s-1}^{1/2},\ b_{ks}),\ k=1,\ \cdots,\ n$$

in I^s with F-discrepancy $D(n)$ too, and a set P of A_s:

$$\boldsymbol{x}_k=(x_{k1},\ \cdots,\ x_{ks})',\ k=1,\ \cdots,\ n, \tag{2.10}$$

where

$$x_{kj}=\prod_{i=1}^{j}b_{ki}^{1/(s-i+1)},\ k=1,\ \cdots,\ n,\ j=1,\ \cdots,\ s. \tag{2.11}$$

The set P satisfies (2.9).

Example 2. Let D be the s-dimensional unit ball

$$B_s=\{\boldsymbol{x}:\ x_1^2+\cdots+x_s^2\leqslant 1\}$$

which has a representation

$$x_j=\varphi_1S_2\cdots S_jC_{j+1},\quad j=1,\ \cdots,\ s-1,$$
$$x_s=\varphi_1S_2\cdots S_{s-1}S_s,$$

where $S_k=\sin(\pi\varphi_k)$, $C_k=\cos(\pi\varphi_k)$, $k=2,\ \cdots,\ s-1$, $S_s=\sin(2\pi\varphi_s)$ and $C_s=\cos(2\pi\varphi_s)$ in which $\boldsymbol{\varphi}\in I^s$. Then we have

$$J(\boldsymbol{\varphi})=2\pi^{s-1}\varphi_1^{s-1}\prod_{i=2}^{s}S^{s-i}$$

and

$$\begin{aligned}v(B_s)&=\int_{I^s}J(\boldsymbol{\varphi})d\boldsymbol{\varphi}\\&=2\pi^{s-1}\int_0^1\varphi_1^{s-1}d\varphi_1\prod_{i=2}^{s}\int_0^1S_i^{s-i}d\varphi_i\\&=\frac{2}{s}\prod_{i=2}^{s}B\left(\frac{1}{2},\ \frac{s-i+1}{2}\right),\end{aligned}$$

because for any integer $m>0$,

$$\int_0^1(\sin(\pi x))^m dx=\frac{1}{\pi}B\left(\frac{1}{2},\ \frac{m+1}{2}\right),$$

Therefore

$$f_i(\varphi_i)=\begin{cases}s\varphi_1^{s-1}, & \text{if } i=1,\\ \pi(\sin(\pi\varphi_i))^{s-i}\Big/B\left(\frac{1}{2},\ \frac{s-i+1}{2}\right), & \text{if } i=2,\ \cdots,\ s\end{cases}$$

are density functions over I^s with corresponfing distribution functions

$$F_1(x_1)=s\int_0^{x_1}\varphi_i^{s-1}d\varphi_1=x_1^s,$$

$$F_i(x_i)=\frac{\pi}{B(1/2,\ (s-i+1)/2)}\int_0^{x_i}(\sin\pi x)^{s-i}dx,\ i=2,\ \cdots,\ s.$$

For a given set $\{\boldsymbol{b}_k,\ k=1,\ \cdots,\ n\}$ in I^s with discrepancy $D(n)$, we have a set $\{\boldsymbol{c}_k,\ k=1,\ \cdots.\ n\}$, where

$$c_{k1}=b_{k1}^{1/s},$$

$$F_i(c_{ki})=b_{ki},\ i=2,\ \cdots,\ s,\ k=1,\ \cdots,\ n,$$

in I^s with F-discrepancy $D(n)$ too, and finally a set P of B_s:

$$\boldsymbol{x}_k=(x_{k1},\ \cdots,\ x_{ks})',\ k=1,\ \cdots,\ n, \tag{2.12}$$

where

$$\begin{aligned} x_{kj}&=b_{k1}^{1/s}\prod_{i=2}^{j}S_{ki}C_{k,j+1},\ j=1,\ \cdots,\ s-1,\\ x_{ks}&=b_{k1}^{1/s}\prod_{i=2}^{s}S_{ki}\end{aligned} \tag{2.13}$$

in which $S_{ki}=\sin(\pi c_{ki})$, $C_{ki}=\cos(\pi c_{ki})$, $i=1,\ \cdots,\ s-1$, $S_{ks}=\sin(2\pi c_{ks})$, $C_{ks}=\cos\ (2\pi c_{ks})$, $k=1,\ \cdots,\ n$.

Example 3. Let D be the $s-1$ dimensional unit sphere

$$S^{s-1}=\{\boldsymbol{x}:\ x_1^2+\cdots+x_s^2=1\}$$

which has a representation

$$x_j=\prod_{i=1}^{j-1}S_iC_j,\ j=1,\ \cdots,\ s-1,$$

$$x_s=\prod_{i=1}^{s-1}S_i,$$

where $S_i=\sin(\pi\varphi_i)$, $C_i=\cos(\pi\varphi_i)$, $i=1,\ \cdots,\ s-2$, $S_{s-1}=\sin(2\pi\varphi_{s-1})$ and $C_{s-1}=\cos(2\pi\varphi_{s-1})$ in which $\boldsymbol{\varphi}\in I^{s-1}$. Then

$$J(\boldsymbol{\varphi})=2\pi^{s-1}\prod_{i=1}^{s-2}S_i^{s-i-1}$$

and

$$v(S^{s-1})=\int_{I^{s-1}}J(\boldsymbol{\varphi})d\boldsymbol{\varphi}=2\pi\prod_{i=1}^{s-2}B\left(\frac{1}{2},\ \frac{s-i}{2}\right).$$

Therefore

$$f_i(\varphi_i)=\pi S_i^{s-i-1}/B(1/2,\ (s-i)/2),\ i=1,\ \cdots,\ s-1$$

are density functions over I with corresponding distribution functions

$$F_i(x_i)=\frac{\pi}{B(1/2,\ (s-i)/2)}\int_0^{x_i}(\sin\pi t)^{s-i-1}\,dt,\ 0<i<s.$$

For a given set $\{\boldsymbol{b}_k,\ k=1,\ \cdots,\ n\}$ in I^{s-1} with discrepancy $D(n)$, we have a set $\{\boldsymbol{c}_k,\ k=1,\ \cdots,\ n\}$, where

$$F_i(c_{ki})=b_{ki},\ i=1,\ \cdots,\ s-1,\ k=1,\ \cdots,\ n,$$

in I^{s-1} with F-discrepancy $D(n)$ too, and finally a set P of S^{s-1}:

$$\boldsymbol{x}_k=(x_{k1},\ \cdots,\ x_{ks})',\ k=1,\ \cdots,\ n, \tag{2.14}$$

where

$$x_{kj}=\prod_{i=1}^{j-1} S_{ki}C_{kj},\ j=1,\ \cdots,\ s-1, \tag{2.15}$$

$$x_{ks}=\prod_{i=1}^{s-1} S_{ki}$$

in wihch $S_{ki}=\sin(\pi c_{ki})$, $C_{ki}=\cos(\pi_k c_i)$, $i=1,\ \cdots,\ s-2$, $S_{k,s-1}=\sin(2\pi c_{k,s-1})$ and $C_{k,s-1}=\cos(2\pi c_{k,s-1})$, $k=1,\ \cdots,\ n$.

Example 4. The domain D is a part of the boundary of s-dimensional unit simplex

$$T_{s-1}=\{\boldsymbol{x}:\ x_1+\cdots+x_s=1,\ x_i\geqslant 0,\ i=1,\ \cdots,\ s\}$$

which has a representation

$$x_i=(S_1\cdots S_{i-1}C_i)^2,\ i=1,\ \cdots,\ s-1,$$

$$x_s=(S_1\cdots S_{s-2}S_{s-1})^2,$$

where $S_i=\sin(\pi\varphi_i/2)$, $C_i=\cos(\pi\varphi_i/2)$, $i=1,\ \cdots,\ s-1$ and $\boldsymbol{\varphi}\in I^{s-1}$. We have

$$\det(\boldsymbol{TT}')=\left(\pi^{s-1}\prod_{i=1}^{s-1} S_i^{2(s-i)-1}C_i\right)^2\det(\boldsymbol{SS}'),$$

where

$$\boldsymbol{S}=\begin{pmatrix} -1 & C_2^2 & S_2^2C_3^2 & \cdots & S_2^2 & \cdots & S_{s-2}^2 & C_{s-1}^2 & S_2^2 & \cdots & S_{s-2}^2 & S_{s-1}^2 \\ 0 & -1 & C_3^2 & \cdots & S_3^2 & \cdots & S_{s-2}^2 & C_{s-1}^2 & S_3^2 & \cdots & S_{s-2}^2 & S_{s-1}^2 \\ \vdots & \vdots & \vdots & \cdots & & & \vdots & & & & & \vdots \\ 0 & 0 & 0 & \cdots & & & -1 & & & & & 1 \end{pmatrix}.$$

Note that $\det(\boldsymbol{SS}')$ is invarint if $\boldsymbol{S}$ is replaced by $\boldsymbol{AS}$, where $\boldsymbol{A}$ is an $(s-1)\times(s-1)$ matrix with $\det \boldsymbol{A}=\pm 1$. We now prove that there exists an $(s-1)\times(s-1)$ matrix $\boldsymbol{A}$ with $\det \boldsymbol{A}=\pm 1$ such that

$$\boldsymbol{AS}=\begin{pmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & -1 \end{pmatrix}=\boldsymbol{V}_{s-1},$$

say. In fact, if $s=2$, then $\boldsymbol{S}=(-1,\ 1)$, and the assertion is true. Suppose now that $s>2$ and the assertion holds for $s-1$. Then

$$\begin{pmatrix} 1 & -S_2^2 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}\boldsymbol{S}=\begin{pmatrix} -1 & 1 & 0 & \cdots & & 0 & \\ 0 & -1 & C_3^2 & \cdots & S_3^2 & \cdots & S_{s-1}^2 \\ \vdots & \vdots & \vdots & \cdots & & \vdots & \\ 0 & 0 & 0 & \cdots & -1 & 1 & \end{pmatrix}.$$

By induction hypothesis, there is an $(s-2)\times(s-2)$ matrix $\boldsymbol{A}_1$ with $\det \boldsymbol{A}_1=\pm 1$ and

$$A_1\begin{pmatrix} -1 & C_3^2 & \cdots & S_3^2 & \cdots & S_{s-1}^2 \\ 0 & -1 & \cdots & & & \vdots \\ \cdots & \vdots & \cdots & & & \vdots \\ 0 & 0 & \cdots & -1 & 1 & \end{pmatrix} = V_{s-2}.$$

Therefore

$$\begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & A_1 \end{pmatrix}\begin{pmatrix} 1 & -S_2^2 & \mathbf{0} \\ 0 & 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{pmatrix} S = V_{s-1},$$

and the assertion follows. We have

$$\det(SS') = \det(V_{s-1}V'_{s-1}) = \det\begin{pmatrix} 2 & -1 & \cdots & 0 & 0 \\ -1 & 2 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 2 & -1 \\ 0 & 0 & \cdots & -1 & 2 \end{pmatrix} = \Delta_{s-1},$$

say. Since $\Delta_1 = 2$ and $\Delta_t = 2\Delta_{t-1} - \Delta_{t-2}(t>2)$, we have $\Delta_{s-1} = s$. Hence

$$J(\varphi) = \pi^{s-1}s^{1/2}\prod_{i=1}^{s-1} S_i^{2(s-i)-1}C_i,$$

and

$$v(T_{s-1}) = \int_{I^{s-1}} J(\varphi)\,d\varphi = s^{1/2}/(s-1)\,!.$$

Therefore

$$f_i(\varphi_i) = (s-i)S_i^{2(s-i)-1}C_i,\ i=1,\ \cdots,\ s-1$$

are density functions over I with corresponding distribution functions

$$F_i(x_i) = \int_0^{x_i} f_i(\varphi_i)\,d\varphi_i = (\sin(\pi x_i/2))^{2(s-i)},\ i=1,\ \cdots,\ s-1.$$

For a given set $\{b_k,\ k=1,\ \cdots,\ n\}$ in I^{s-1} with discrepancy $D(n)$, we have a set $\{c_k,\ k=1,\ \cdots,\ n\}$, where

$$c_{ki} = (2/\pi)\arcsin(b_{ki}^{1/(2s-2i)}),\ i=1,\ \cdots,\ s-1,\ k=1,\ \cdots,\ n.$$

Finally, we have a set P of T_{s-1}:

$$x_k = (x_{k1},\ \cdots,\ x_{ks})',\ k=1,\ \cdots,\ n, \tag{2.16}$$

where

$$\begin{cases} x_{kj} = \prod_{i=1}^{j-1} b_{ki}^{1/(s-i)}(1 - b_{kj}^{1/(s-j)}),\ j=1,\ \cdots,\ s-1, \\ x_{ks} = \prod_{i=1}^{s-1} b_{ki}^{1/(s-i)},\ k=1,\ \cdots,\ n. \end{cases} \tag{2.17}$$

§3. Some Applications

In this section we shall pay our attention to applications of number theoretic method in numerical evaluation of probabilities and moments of a continuous multivariate distributions. The basic quadrature formulas are given by (2.6) and (2.7). There are a number of methods to produce sets of points $\{b_k, k=1, \cdots, n\}$ in I^s (see Hua and Wang [6]). In view, of our experiences, we will recommend using the following algorithm: Let $(h_1, \cdots, h_s; n)$ be an integral vector, where $h_1=1$, $0<h_i<n$ and g. c. d. $(h_i, n)=1$, $i=1, \cdots, s$. Let

$$p_n(k)=(kh_1, \cdots, kh_s)=(q_{k1}, \cdots, q_{ks}) \pmod n, \; k=1, \cdots, n,$$

where $0<q_{ki}\leqslant n$. Set

$$b_{ki}=(2q_{ki}-1)/2n. \; i=1, \cdots, s, \; k=1, \cdots, n.$$

Then $\{b_k\}$ is a set of points in I^s with lower discrepancy if $(h_1, \cdots, h_s; n)$ are carefully selected. A table of $(h_1, \cdots, h_s; n)$ for $1<s<19$ was contained in [6] as an appendix.

Example 5. This problem was come from alloy steel industry (see Fang and Wu [3]in details). Let $\boldsymbol{x}$ be an $s\times 1$ vector which denotes the percentage of chemical elements in an slloy steel and let $(\boldsymbol{\mu x})=(\mu_1, \cdots, \mu_t)'$ be the corresponding vector which stands for the quality of the steel. The regression equation between $\boldsymbol{\mu}(\boldsymbol{x})$ and $\boldsymbol{x}$ is

$$\hat{\boldsymbol{\mu}}(\boldsymbol{x})=\boldsymbol{a}+\boldsymbol{B}\boldsymbol{x},$$

where $\boldsymbol{a}$ and $\boldsymbol{B}$ are $t\times 1$ and $t\times s$ matrices of regression coefficients and $\boldsymbol{x}$ belongs to a rectangle $K=[a_1, b_1]\times\cdots\times[a_s, b_s]$. Suppose that for each $\boldsymbol{x}\in K$, we have $\boldsymbol{\mu}(\boldsymbol{x})\sim N_t(\boldsymbol{a}+\boldsymbol{B}\boldsymbol{x}, \boldsymbol{\Sigma})$, the multivariate normal distribution, where $\boldsymbol{a}$, $\boldsymbol{B}$, and $\boldsymbol{\Sigma}$ can be used by their least square estimators. Let T_i, $i=1, \cdots, t$, be the constants such that the steel is said to be qualified if $\mu_i>T_i$, $i=1, \cdots, t$. Thus the probability that the alloy steel corresponding to $\boldsymbol{x}$ is qualified is equal to

$$p(\boldsymbol{x})=\int_{T_1}^{\infty}\cdots\int_{T_t}^{\infty} n_t\,(\boldsymbol{y}, \hat{\boldsymbol{\mu}}(\boldsymbol{x}), \boldsymbol{\Sigma})\; d\boldsymbol{y} \tag{3.1}$$

where $\boldsymbol{y}=(y_1, \cdots, y_t)$, and $n_t(\boldsymbol{y}, \boldsymbol{\mu}(\boldsymbol{x}), \boldsymbol{\Sigma})$ is the density of $N_t(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. The integral (3.1) can be evaluated by applying (2.6) if we choose suitable numbers A_i, $i=1, \cdots, t$, such that

$$\begin{aligned} p(\boldsymbol{x}) &\cong \int_{T_1}^{A_1}\cdots\int_{T_t}^{A_t} n_t(\boldsymbol{y}, (\hat{\boldsymbol{\mu}}\boldsymbol{x}), \boldsymbol{\Sigma})\; d\boldsymbol{y} \\ &= \prod_{i=1}^{t}(A_i-T_i)\frac{1}{n}\sum_{k=1}^{n} n_t(\boldsymbol{z}_k, \hat{\boldsymbol{\mu}}(\boldsymbol{x}), \boldsymbol{\Sigma}), \end{aligned} \tag{3.2}$$

whore

$$\boldsymbol{z}'_k=(z_{k1}, \cdots, z_{kt})=(T_1+(A_1-T_1)b_{k1}, \cdots, T_t+(A_t-T_t)\,b_{kt}),$$

$k=1, \cdots, n$ and $\{\boldsymbol{b}_k\}$ is a uniformly distributed set of points in $\boldsymbol{I}^t$. To illustrate the computational accuracy, set $\boldsymbol{\Sigma}=\boldsymbol{I}_5$, the 5×5 identity matrix and $\boldsymbol{\mu}(\boldsymbol{x})=\boldsymbol{0}$, $T_i=-1$, $A_i=1$, $i=1, \cdots, 5$ in (3.2). We have

$$p=\int_{1}^{1-}\cdots\int_{-1}^{1} n_5(\boldsymbol{y}, \boldsymbol{0}, \boldsymbol{I}_5)\; d\boldsymbol{y}.$$

Now we have by (3.2) the following:

Table 1 shows that 5-digit accuracy for numerical evaluating a 5-fold integral is

Table 1.

n	approximate values of p
1069	0.148299406
2129	0.148295351
5003	0.148291410
8191	0.148291358
∞	0.148291347

obtained by the use of 1069 points only.

Example 6. The moments of order statistics. Let $X_1, \cdots, X_s$ be a sample from the population with distribution function $F(x)$ and density $f(x)$. Let $Y_s=X_{(1)}\leqslant Y_{s-1}=X_{(2)}\leqslant \cdots Y_1=X_{(s)}$ be their order statistics. It is known that the joint density of $Y_1, \cdots, Y_s$ is given by

$$s!\prod_{i=1}^{s} f(y_i),\ y_s<y_{s-1}<\cdots<y_1.$$

Then the order $m_1, \cdots, m_s$ mixed moment of $X_{(i)}$, $i=1, \cdots, s$ is defined by

$$\mu(m_s, \cdots, m_1)=s!\int_{D^*}\prod_{j=1}^{s} y_j^{m_j}f(y_j)dv, \tag{3.3}$$

where $D^*=\{-\infty<y_s<y_{s-1}<\cdots<y_1<\infty\}$. There exist a and b such that

$$P(a<y_s,\ y_1<b)\cong 1.$$

Taking a transformation $z_i=(y_i-a)/(b-a)$, $i=1, \cdots, s$, we have

$$\mu(m_s, \cdots, m_1)=s!(b-a)^s\int_D\prod_{j=1}^{s}[(a+(b-a)z_j)^{m_j}f(a+(b-a)z_j)]dv$$

where D is defined in Example 1. By Example 1, (2.6) and (2.7) we suggest the following two formulas for the calculation of $\mu(m_s, \cdots, m_1)$:

$$\mu(m_s, \cdots, m_1)\cong s!(b-a)^s\frac{1}{n}\sum_{k=1}^{n}\prod_{j=1}^{s}\left[t_{kj}^{(m_j+1)-1}f\left(\prod_{i=1}^{j}t_{ki}\right)\right],$$

where $t_{kj}=a+(b-a)b_{kj}$ and $\{\boldsymbol{b}_k, k=1, \cdots, n\}$ is a uniformly distributed set of points in I^s, and

$$\mu(m_s, \cdots, m_1)\cong s!\ (b-a)^s\frac{1}{n}\sum_{k=1}^{n}\prod_{j=1}^{s}[a+(b-a)c_{kj}]^{m_j}f(a+(b-a)c_{kj})$$

$$=s!(b-a)^s\ \frac{1}{n}\sum_{k=1}^{n}\prod_{j=1}^{s}[a+(b-a)b_{kj}^{1/(s-j+1)}]^{m_j}f(a+(b-a)b_{kj}^{1/(s-j+1)})$$

where $\{\boldsymbol{c}_k\}$ and $\{\boldsymbol{b}_k\}$ are given in Example 1.

Since the mixed moments of order statiftics of uniform distribution $U(0, 1)$ on $I=[0, 1]$ can be formulated. We give an example in Table 2 which shows the (see Table 2) acc uracies.

Example 7. In his study of compositional data, Aitchison [1] introduced in 1986 a so-called additive logistic normal distribution. Let $\boldsymbol{T}_n$ be a domain defined in Example 4 with $n=N-1$. Any $\boldsymbol{x}$ in $\boldsymbol{T}_n$ is called a composition. For a given $\boldsymbol{x}\in T_n$, we denote by $\boldsymbol{x}_{-N}$ the n-dimensional vector formed by the first n components of $\boldsymbol{x}$.

Table 2. Mixed moments of order statistics of $U(0, 1)$, $s=7$

n	$E(X_{(1)}X_{(3)}X_{(5)})$	$E(X_{(1)}X_{(2)}^2X_{(3)}^3)$
418	0.03887518	0.00326433
597	0.03888518	0.00339034
828	0.03888511	0.00332116
1010	0.03887286	0.00332084
1220	0.03889159	0.00325168
∞	0.03888889	0.00326340

Let

$$\boldsymbol{y}=\log(\boldsymbol{x}_{-N}/X_N)=(\log(X_1/X_N),\ \cdots,\ \log(X_n/X_N))'. \tag{3.4}$$

The equation (3.4) yields an one to one mapping from T_n to R^n. A random vector $\boldsymbol{x}\in T_n$ is said to have an additive logistic normal distribution $AN_n\ (\boldsymbol{\mu},\boldsymbol{\Sigma})$ if its corresponding $\boldsymbol{y}\sim N_n\ (\boldsymbol{\mu},\boldsymbol{\Sigma})$.

Aitchison gave the formulas for $E(\log(X_i/X_j))$, $E(X_i/X_j)$, $\mathrm{Cov}(\log(X_i/X_j), \log(X_k/X_l))$, $\mathrm{Cov}(X_i/X_j, X_k/X_l)$. It seems difficulty for him to calculate $E(X_i)$ and $\mathrm{Cov}(X\mathrm{i}, X_j)$, which are required in many practictical problems.

The density function of $AN_n(\boldsymbol{\mu},\boldsymbol{\Sigma})$ is given by

$$(2\pi)^{-n/2}(\det\boldsymbol{\Sigma})^{-1/2}\Big(\prod_{i=1}^{N}X_i^{-1}\Big)\exp\Big\{-\frac{1}{2}\Big(\log\frac{\boldsymbol{x}_{-N}}{\boldsymbol{x}_N}-\boldsymbol{\mu}\Big)'\boldsymbol{\Sigma}^{-1}\Big(\log\frac{\boldsymbol{x}_{-N}}{\boldsymbol{x}_N}-\boldsymbol{\mu}\Big)\Big\} \tag{3.5}$$

and the mixed moment of $\boldsymbol{x}$ is

$$E(X_1^{t_1}\cdots X_N^{t_N})=(2\pi)^{-n/2}(\det\boldsymbol{\Sigma})^{-1/2}\int_{T_n}\prod_{i=1}^{N}x_i^{t_i-1}$$

$$\exp\Big\{-(1/2)\Big[\Big(\log\frac{\boldsymbol{x}_{-N}}{X_N}-\boldsymbol{\mu}\Big)'\boldsymbol{\Sigma}^{-1}\Big(\log\frac{\boldsymbol{x}_{-N}}{X_N}-\boldsymbol{\mu}\Big)\Big]\Big\}d\mathrm{v},$$

where $d\mathrm{v}$ is the volume element of T_n. By Example 4, we have

$$E(X_1^{t_1}\cdots X_N^{t_N})=C\int_{I^n}\prod_{i=1}^{n}(C_i^{2t_i-1}S_i^{2(t_{i+1}+\cdots+t_N)-3})Q(\boldsymbol{\varphi})d\boldsymbol{\varphi},$$

where $\boldsymbol{\varphi}\in I^n$, $d\boldsymbol{\varphi}=\prod_{i=1}^{n}d\varphi_i$, $C=(2/\pi)^{-n/2}(\det\boldsymbol{\Sigma})^{-1/2}N^{1/2}$ and

$$Q(\boldsymbol{\varphi})=\exp\{-(1/2)(\boldsymbol{g}(\boldsymbol{\varphi})-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{g}(\boldsymbol{\varphi})-\boldsymbol{\mu})\}$$

in which

$$g(\boldsymbol{\varphi})=(\log(C_1^2)-\log(S_1^2\cdots S_n^2),\ \cdots,\ \log(S_1^2\cdots S_{n-1}^2C_n^2)-\log(S_1^2\cdots S_n^2))',$$

$$=2\Big(\log C_1-\sum_{i=1}^{n}\log S_i,\ \log C_2-\sum_{i=2}^{n}\log S_i,\cdots,\log C_k-\sum_{i=n}^{n}\log S_i\Big)'.$$

By (2.6) or (2.7) we may obtain approximate values of any mixed moments of $AN_n(\boldsymbol{\mu},\ \boldsymbol{\Sigma})$.

Example 8. We meet the problem of directional data in which some statistical distributions are defined over S^{s-1} (see Example 3). The so-called Langevin distribution which is an extension of Von Mises and Fisher's distribution has a density function

$$C\exp\{k\boldsymbol{\mu}'\boldsymbol{x}\}, \boldsymbol{x}\in S^{s-1},$$

where $\boldsymbol{\mu}\in S^{s-1}$, $k>0$ and C is the normaling constant (cf. Mardia [8]). Another distribution called Scheidegges-Watson distribution has a sensity functor

$$C\exp\{k(\boldsymbol{\mu}, \boldsymbol{x})^2\}, \boldsymbol{x}\in S^{s-1},$$

where C, k, and $\boldsymbol{\mu}$ have the similar meaning as before (cf. Watson [11]).

We can apply Example 3, (2.6) and (2.7) to calculate probabilities and mixed moments for these two kinds of distributions.

§4. Optimization

In this section we shall generalize inequlity (1.4) to some domains which have been discussed in the past sections. Then we give an example to show that the algorithm mentioned in Section 1 is powerful.

We suppose that the set of singularities of the transformation $\boldsymbol{x}=\boldsymbol{x}(\boldsymbol{\varphi})$, i. e., the set of solutions of $J(\boldsymbol{\varphi})=0$, is a set of D with dimension$<t$, where $\boldsymbol{x}$ is an s-dimensional vector, rnd t is the dimension of D. Hence for any given $\varepsilon>0$, there exist a domain $\mathfrak{C}$ and two positive constants c_1, c_2 depending only on ε such that $v(\mathfrak{C})<\varepsilon$ and

$$c_1<f_i(\varphi_i)<c_2, \quad i=1, \cdots, t,$$

where $\boldsymbol{\varphi}\in I^t\backslash\mathfrak{C}$. Then $dF_i^{-1}(\varphi_i)/d\varphi_i$, $i=1, \cdots, t$ are positive and bounded over $I^t\backslash\mathfrak{C}$ too.

First we take a set $\boldsymbol{b}_k=(b_{k1}, \cdots, b_{kt})'$, $k=1, \cdots, n$ in $I^t\backslash\mathfrak{C}$ with lower discrepancy $D(n)$. Then we have shown that

$$\boldsymbol{c}_k=\boldsymbol{F}^{-1}(\boldsymbol{b}_k)=(F_1^{-1}(b_{k1}), \cdots, F_t^{-1}(b_{kt}))', \quad k=1, \cdots, n$$

is a set in I^t with F-discrepancy $D(n)$ too, and finally we have a set in D: $\boldsymbol{x}_k=\boldsymbol{x}(\boldsymbol{c}_k)$, $k=1, \cdots, n$.

Let $\boldsymbol{x}=\boldsymbol{x}(\boldsymbol{\varphi})$, $\boldsymbol{x}^*=\boldsymbol{x}(\boldsymbol{\varphi}^*)$, $\boldsymbol{\varphi}=F^{-1}(\boldsymbol{\psi})$, $\boldsymbol{\varphi}^*=\boldsymbol{F}^{-1}(\boldsymbol{\psi}^*)$. $d\boldsymbol{x}=(dx_1, \cdots, dx_s)'$, $d\boldsymbol{\varphi}=(d\varphi_1, \cdots, d\varphi_t)'$, $d\boldsymbol{\psi}=(d\psi_1, \cdots, d\psi_t)'$, and S be the diagonal matrix

$$\boldsymbol{S}=\operatorname{diag}(d\varphi_1/d\psi_1, \cdots, d\varphi_t/d\psi_t).$$

Then

$$d\boldsymbol{x}'d\boldsymbol{x}=d\boldsymbol{\varphi}'\boldsymbol{T}\boldsymbol{T}\,d\boldsymbol{\varphi}=d\boldsymbol{\psi}'\boldsymbol{S}\boldsymbol{T}\boldsymbol{T}'\boldsymbol{S}d\boldsymbol{\psi}.$$

Since the elements of $\boldsymbol{S}$ and $\boldsymbol{T}$ are bounded in $I^t\backslash\mathfrak{C}$, we have

$$\begin{aligned} d(\boldsymbol{x}, \boldsymbol{x}^*) &= \int_{x(\varphi)}^{x(\varphi^*)} (d\boldsymbol{x}'d\boldsymbol{x})^{1/2} \\ &= \int_{\psi}^{\psi^*} (d\boldsymbol{\psi}'\boldsymbol{S}\boldsymbol{T}\boldsymbol{T}_{\prime}\boldsymbol{S}d\boldsymbol{\psi})^{1/2} < c(\varepsilon)\int_{\psi}^{\psi^*} (d\boldsymbol{\psi}'d\boldsymbol{\psi})^{1/2} \\ &= c(\varepsilon)d_t(\boldsymbol{\psi}, \boldsymbol{\psi}^*), \end{aligned}$$

where $d_t(\boldsymbol{y}, \boldsymbol{z})$ denotes the Euclidean distance in R^t. Hence by (1.4), we have

$$d(n, D)=\max_{\boldsymbol{x}\in D}\min_{1\leqslant k\leqslant n} d(\boldsymbol{x}, \boldsymbol{x}_k)\leqslant 2t^{1/2}c(\varepsilon)D(n)^{1/t}, \tag{4.1}$$

Thus if x_k, $k=1, \cdots, n$, are scattered uniformly in D, then the maximum value of a function on these points may be taken as an approximate value of the global maximum of the function on D.

Example 9. Additive logistic elliptical distributions. The so-called additive logistic elliptical distributions defined on T_{s-1} (see Example 4) are extensions of the additive logistic normal distributions mentioned in Example 7 and have density functions of the form

$$f(\boldsymbol{x}) = (\det \boldsymbol{\Sigma})^{-1/2} \prod_{i=1}^{s} x_i^{-1} g(\boldsymbol{x}), \tag{4.2}$$

where

$$g(\boldsymbol{x}) = g\left(\left(\log \frac{\boldsymbol{x}_{-s}}{\boldsymbol{x}_s} - \boldsymbol{\mu}\right)' \boldsymbol{\Sigma}^{-1} \left(\log \frac{\boldsymbol{x}_{-s}}{\boldsymbol{x}_s} - \boldsymbol{\mu}\right)\right) \tag{4.3}$$

and $\boldsymbol{x}_{-s}$ is an $(s-1)$-dimensional vector formed by the first $s-1$ components of $\boldsymbol{x}$. The mode of $f(\boldsymbol{x})$ can not be analytically formulated so far. However we may use uniformly distributed sets on T_{s-1} to calculate the approximate values of the mode.

When the function g in (4.3) has the form

$$g(u) = C\ (1+u/m)^{-p},\ p>s/2,\ m>0, \tag{4.4}$$

where

$$C = (\pi m)^{-s/2} \Gamma(p) / \Gamma(p-s/2),$$

the corresponding distribution is called additive logistic elliptical Pearson Type VII distribution. In this case to find the mode of $f(\boldsymbol{x})$ is equivalent to obtain the maximum of

$$h(\boldsymbol{x}) = \prod_{i=1}^{s} x_i^{-1} \left[1+\left(\log \frac{\boldsymbol{x}_{-s}}{\boldsymbol{x}_s} - \boldsymbol{\mu}\right)' \boldsymbol{\Sigma}^{-1} \left(\log \frac{\boldsymbol{x}_{-s}}{\boldsymbol{x}_s} - \boldsymbol{\mu}\right)/m\right]^{-p}$$

over T_{s-1}. The results in Table 3 show that the approximate values of the mode in the case of $s=3$, $p=9$, $m=5.5$ and

$$\boldsymbol{\mu} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} 1 & -0.7 \\ -0.7 & 1 \end{pmatrix}$$

are closed to those of the mode $(1/3, 1/3, 1/3)'$.

Table 3 Approximate values of the mode

n	M_n	x_1	x_2	x_3
233	26.55203	0.3271035	0.3306723	0.3422242
377	26.82553	0.3061467	0.3598099	0.3340434
610	26.48600	0.3604561	0.3150540	0.3244899
4181	26.89095	0.3256147	0.3368701	0.3375152
10946	26.99783	0.3334971	0.3337690	0.3827339
17711	26.98296	0.3292827	0.3388427	0.3318746

We note in Table 3 that when n is increasing, the corresponding M_n, in principle, is increasing also. But sometimes $M_n < M_{n'}$, where $n > n'$, because the sets $\{\boldsymbol{c}_k\}$ for different n may be completely distinct. Hence we suggest using a sequential method to improve the above result.

The following program is designed for our presented problem,

Step 1. Choose a uniformly distributed set of points $\{x_k, k=1, \cdots, n_0\}$ in $D_0=T_{s-1}$ with suitable n_0. Find the maximun M_0 of the function among these points, and assume that it is attained at $x_0^*=(x_{01}, \cdots, x_{0s})'$.

Step 2. Find a small domain D_1 of D_0 such that $D_1 \subset D_0$ and $x_0^* \subset D_1$. For instance, D_1 is a domain with x_0^* located near the gravity of D_1. More precisely, we cdoose a_i, $i=1, \cdots, s$, such that

$$0 \leqslant a_i < x_{0i}, \quad i=1, \cdots, s.$$

Set $a=a_1+\cdots+a_s$ and $b_i=a_i+1-a$, $i=1, \cdots, s$. Then

$$1 \geqslant b_i \geqslant a_i + \sum_{j=1}^{s} x_{0j} - \sum_{k=1}^{s} a_k = x_{0i} + \sum_{\substack{j=1 \\ i \neq j}}^{s} (x_{0j} - a_j) \geqslant x_{0i},$$

$i=1, \cdots, s$. Denote

$$D_1 = \{x = (x_1, \cdots, x_s)' : a_i \leqslant x_i \leqslant b_i, \ i==1, \cdots, s, \ x \in D_0\}.$$

Let z_k, $k=1, \cdots, n_1$, be an uniformly distributed set of points in T_{s-1}. Then we have a set $\{x_k, k=1, \cdots, n_1\}$, where

$$x_{ki} = a_i + (1-a) z_{ki}, \quad i=1, \cdots, s, \ k=1, \cdots, n_1,$$

which is uniformly distributed over D_1. Denote by M_1 the maximum of the function on x_k' s which is attained at the point x_1^*.

Step 3. Suppose that in the jth step we have found the maximum M_j of the function and the correrponding point x_j^*. By a similar method we can reduce the domain D_j to D_{j+1}, and make a set of points on D_{j+1}, by which we can find another maximum M_{j+1} of the function and the corresponding point x_{j+1}^*.

Repeat Step 3 until the search domain is smaller. The last maximum M_{j+1} is expected to be closed to the global maximum M of the function.

Applying the above program to our problem, we set $n_0=n_1=\cdots=223$ for each step, and the results are given in Table 4, which improve those in Table 3.

Table 4. The sequencial method for optimization

No	a_i	b_i	M_i	x_1^*	x_2^*	x_3^*
1	0.0000	1.0000	26.55203	0.3271035	0.3306723	0.3422242
2	0.3000	0.4000	26.97836	0.3304331	0.3343395	0.3352274
3	0.3300	0.3400	26.99543	0.3327104	0.3330673	0.3342224
4	0.3330	0.3340	26.99994	0.3332711	0.3333067	0.3334223
the	global	maximum	27.00000	0.3333333	0.3333333	0.33333333

We have done many examples which all show that the present sequential method is advantageous.

References

[1] Aitchison, J., The statistical Analysis of Compositional Data, Chaman and Hall, London/New York, (1986).

[2] Avriel, M., Nonlinear Programming, Analysis and Methods Prentice-Hall, Inc. Englewood Cliffs, New Jersey, 1976.

[3] Fang, K. T. & Wu, C. Y., The extrime value problem of some probability function, *Acta Math. Appl. Sinica*, 2 (1979), 132—148.

[4] Hlawka, E., Funktionen von beschrankter Variation in der Theorie Gleichverteilnug, *Ann. Mat. pure Appl.*, **54**, (1961), 325—333.

[5] Hlawka, E. & Muck, R., A transformations of equidistributed sequences, in "Applications of Number Theory to Numerical Analyssi" (Zaremba.S. K. ed). Acad. Press, New York, 1972, 371—388.

[6] Hua, L. K. & Wang. Y., Applications of Number Theory to Numerical Analysis, Springer-Verlag (Heidelberg) and Science Press (Beijing), 1981.

[7] Koksma, I, F., Een algemeene stelling uit de theorie der gelijkmatige Verdeeling modulo 1, *Math. B (Zutphen)*, **11**, (1942—1943), 7—11.

[8] Madia, K. V., Statistcs of Directional Data, Acad. Press, New York, (1972).

[9] Niederreiter, H., Metric theorems on the distribution of sequenes, *Proc. Symp. Pure Math.*, 24, *AMS Pro. R. I.*, (1973), 195—212.

[10] Niederreiter, H., A quasi-Monte Carlo method for the approximate computation of the extreme values of a function, *Studies in Pure Math, Birkhauser, Basel*, (1983), 523-529.

[11] Watson, G. S., Statistics on Sphere, Wiley, New York, (1983).

[12] Weyl, H., Uber die Gleichverteilurg der Zahlen mod Eins, *Math. Ann.*, **77**, (1916), 313—352.

[13] Zielinski, R., On the Monte Carlo evaluaton of the extreme values of a function, *Algorytncy*, 2 (1965), 7—13.