Pramāṇa - J. Phys., Vol. 28, No. 2, February 1987, pp. 97–107. © Printed in India.

Implementation and comparative study of random sequences for nonlinear least square data fitting

TAPAS BANDYOPADHYAY and P K SARKAR
Health Physics Unit, Variable Energy Cyclotron Centre, I/AF Bidhan Nagar, Calcutta 700 064, India

MS received 19 November 1986

Abstract. A numerical study of nonlinear least square data fitting using random numbers from the congruential generator and several quasi-random generators is presented. The results indicate that at least up to five dimensions some of the quasi-random sequences yield better accuracy than the congruential pseudo-random sequence. Some recommendations for selecting the generators of quasi-random sequences are also given.

Keywords. Least square data fitting; random search; quasi-random sequences; pseudo-random sequences; global optimization; Gaussian detector response.

PACS No. 02·60

1. Introduction

In an earlier work (Bandyopadhyay and Sarkar 1985) the authors have shown that the quasi-random search (QRS) technique can be used effectively to fit experimentally observed data (detector responses) with empirical expressions. It has also been shown that the QRS out-performs classical optimization techniques especially when the experimentally observed data are associated with poor statistics. This is because, with the introduction of statistical uncertainty the function to be optimized has a high probability of becoming multi-extremal defined over some multi-variate parameter space, and as such the traditional techniques can suffer severely from trapping in local minima for such problems. This limitation of the traditional techniques is crucial in experimental physics and astrophysics where observed data is often required to be fitted with empirical relations. The quasi-random search technique with importance sampling (Bandyopadhyay and Sarkar 1985) can avoid such trapping with proper choice of the importance function so that the localization of search is achieved rather slowly.

In the present work, we carry out numerical studies to investigate the effective implementation of the method using several quasi-random sequences (QRS) and one congruential pseudo-random sequence in dimensions three, four and five. Further, we compare these sequences to determine their relative merits for nonlinear least square fitting problems. Theoretical estimates of the error bounds of optimization can be obtained by using the extremal discrepancy of the sequence and the modulus of continuity of the function (Niederreiter 1983), but such estimates are not very precise. Furthermore, estimation of the extremal discrepancy of the sequences and the modulus of continuity of the functions is extremely difficult and time consuming especially at high dimensions. It is therefore practical to side step this problem heuristically by carrying out numerical experimentations with test functions of the type involved. For practical applications in solving nonlinear least square fitting problems, several sequences can be tried on test problems—giving perhaps relevant experience to choose with some confidence. It may be noted that quasi-random sequences have been used for optimization and random search in hypercubes by Neiderreiter and Mc Curley (1979), Neiderreiter and Peart (1982), Aird and Rice (1977), Sobol (1979) and Fox (1986). The present study though carried out with Gaussian-like functions, the results obtained and conclusions drawn therefrom can generally be believed to hold true for other nonlinear least square fitting problems also. It may be noted that in this study of Monte Carlo optimization we will only consider the case where we are interested in the global optimum of a function.

2. Computational aspects

2.1 *Description of the sequences*

Before describing the sequences used for the present study let us first define the radical inverse function $\phi_r(n)$.

Let n be a non-negative integer and r an integer with $r \geqslant 2$.

Let $$n = \sum_{i=0}^{k} b_i r^i \quad \text{with} \quad b_1 \in [0, 1, \ldots, g-1],$$

for $0 \leqslant i \leqslant k$ be the r-adic expansion of n. The radical inverse function $\phi_r(n)$ is then given by

$$\phi_r(n) = \sum_{i=0}^{k} b_i r^{-i-1}. \tag{1}$$

We now describe the individual sequences.

Hammersley sequence (HAM): The Hammersley sequence (Zaremba 1968) of order N in $I^s(I^s = 0, 1)^s$, the s-dimensional unit cube, $s \geqslant 2$ consists of

$$a_n = \left[\frac{n}{N}, \phi_r(n), \ \ldots \ \phi_{r_{s-1}}(n)\right], \quad n = 0, 1, \ldots N-1$$

where the bases $r_1, \ldots r_{s-1}$ are the pairwise relative prime. Usually, one takes $r_1, \ldots r_{s-1}$ as the first $s-1$ primes.

Halton sequence (HAL): The Halton sequence (Halton 1960) in s dimensions is defined by

$$a_n = [\phi_2(n), \phi_3(n), \ldots \phi_{p(s)}(n)]$$

where $p(s)$ are the first s primes.

Zaremba sequence (ZAR): The Zaremba sequence (Halton and Zaremba 1969) in s

dimensions is defined by

$$a_n[\psi_2(n), \ldots \psi_{p(s)}(n)]$$

where $\psi_r(n)$ is the folded radical inverse function and is given by:

$$\psi_r(n) = (b_0+0)_{\mathrm{mod}\, r} r^{-1} + (b_1+1)_{\mathrm{mod}\, r} r^{-2} + \ldots + (b_i+i)_{\mathrm{mod}\, r} r^{-i-1} + \ldots .$$

Haber sequence (*HAB*): The sequence suggested by Haber (1970) can be described as follows:

$$a_n = \left[\left\{\frac{n(n+1)}{2}\sqrt{2}\right\}, \cdots \left\{\frac{n(n+1)}{2}\sqrt{p_s}\right\}\right],$$

where $\{y\}$ denotes the fractional part of y.

Sequence 5 (*SEQ* 5): This is a sequence (Warnock 1972) defined by

$$a_n = [n/N, \psi_2(n), \ldots \psi_{p(s-1)}].$$

Sequence 6 (*SEQ* 6): This sequence (Warnock 1972) is defined by

$$a_n = [\{m\sqrt{2}\}, \ldots, \{m\sqrt{p(k)}\}].$$

Scrambled Halton sequence (*SCR-HAL*): This is closely related to the Halton sequence and is defined as follows:

$$x_n = \sum_{i=0}^{\infty} \sigma_i(b_i) r^{-i-1},$$

where $\Sigma = (\sigma_i)_{i \geqslant 0}$ are a set of permutations on the ensemble $(0, 1, 2, \ldots, r_3 - 1)$. The permutations Σ are obtained in such a manner as to reduce the one-dimensional r.m.s. discrepancy obtained with the integer r_s. In this work, we used the set of permutations suggested by Braaten and Weller (1979).

Pseudo-random sequence (*PRS*): A set of random numbers which pass some specified statistical test for randomness are known as pseudo-random numbers (Neiderreiter 1978). The commonly used sequence of pseudo-random numbers, called the congruential pseudo-random numbers in the unit interval [0, 1], may be generated as follows: let $m \geqslant 2$ be an integer. Generate a sequence (Y_n), $n = 0, 1, \ldots$ integers in the least residue system modulo m using the recursion $Y_{n+1} = \lambda Y_n + r \pmod m$, with Y_0 as integer satisfying $0 \leqslant Y_0 \leqslant m$ and λ is a positive integer relatively prime to m. Then the sequence (x_n), $n = 0, 1, \ldots$, where $x_n = Y_n/m$ is a sequence of congruential pseudo-random numbers provided the parameters m, λ, Y_0 and r are chosen so that the sequence will pass the test for randomness. In the present work we have chosen $m = 2^{31}$, $r = 0$, $\lambda = 65539$ and $Y_0 = 3115$.

In order to evaluate the performance of the sequences described, numerical results were obtained with the QRS technique as proposed by Bandyopadhyay and Sarkar (1985) which can be described briefly as follows.

2.2 *Description of the method*

Let $\chi^2 = F(x) = F(x_1, x_2, \ldots x_n)$ be the function to be minimized, where x_1, $x_2, \ldots$ etc denote the unknown parameters of χ^2. Let $F(x)$ be a bounded real valued function where the value of $M = \min F(x)$ is required. We select the points ${}^1x$, ${}^2x, \ldots {}^Nx$ on the hyperdimensional space on which F is defined and then estimate M and the optimized parameters x^* by the following algorithm:

$$ {}^1x^* = x \quad \text{and} \quad m_1 = F({}^1x), $$

$$ {}^2x^* = \begin{cases} {}^1x^* \text{ and } & m_2 = m_1 \quad \text{if} \quad F({}^2x) \geqslant m, \\ {}^2x \text{ and } & m_2 = F({}^2x) \quad \text{if} \quad F({}^2x) < m_1, \end{cases} $$

$$ {}^{i+1}x^* = \begin{cases} {}^ix^* \text{ and } & m_{i+1} = m_i \quad \text{if} \quad F({}^{i+1}x) \geqslant m_i, \\ {}^{i+1}x \text{ and } & m_{i+1} = F({}^{i+1}x) \quad \text{if} \quad F({}^{i+1}x) < m_i, \end{cases} $$

with ${}^ix = f'({}^ia)$, a random n-tuple

$$ {}^ia = {}^ia_1, {}^ia_2, \ldots, {}^ia_n $$

and
$$ {}^ix_j = f'_j({}^ia_j), \quad 0 \leqslant {}^ia_j \leqslant 1, \quad j = 1, 2, \ldots n. \tag{2} $$

$f' = (f'_j)$, $j = 1, 2, \ldots n$ is a mapping of an n-dimensional unit cube into another n-dimensional space of given shape (Gaussian in our case). The convergence of m_N to M as $N \to \infty$ is guaranteed whenever f is continuous and the sequence ${}^ix$, $i = 1, 2, \ldots$ is dense in the hyperspace.

The sampling scheme is incorporated into the above technique in the following manner.

Let $C = C(x^g, \sigma)$ denote a totally bounded space bounded by a n-dimensional Gaussian of standard deviation

$$ \sigma = [\sigma_1, \sigma_2, \ldots \sigma_n] \quad \text{and} \quad \text{mean } x^g = [x_1^g, x_2^g, \ldots x_n^g]. $$

Let $f_g^1(a, x^g, \sigma)$ be a function taking values in C such that for every

$$ t_i^2 = \frac{[f'_g(a_i, x_i, \sigma_i) - x_i^g]^2}{\sigma_i^2}, \quad i = 1, 2, \ldots n $$

$$ (2\pi)^{1/2} \int_0^\infty \exp(-t_i^2/2) \mathrm{d}t_i = Q_i $$

where
$$ \begin{aligned} Q_i &= a_i \text{ for } a_i \leqslant 0{\cdot}5, \\ &= 1 - a_i \text{ for } a_i > 0{\cdot}5. \end{aligned} \tag{3} $$

It is clear that if a_i is sampled from an uniform distribution between 0 and 1, $f'_g(a_i, x_i^g, \sigma_i)$ will be a member of a Gaussian population having a standard deviation σ_i and mean x_i^g. In other words f_g maps the n-dimensional unit hyper cube in C. Now for given N random n-tuples ${}^1a, {}^2a, \ldots {}^Na$ and a function F to be minimized, we set

$$ \begin{aligned} m_N(C) &= \min[F(x^g)], \min_{1 \leqslant l \leqslant N} F(f_g({}^la, x^g, \sigma) \qquad f'_g({}^la, x^g, \sigma) \in C \\ &= F[x^*(C)], \end{aligned} \tag{4} $$

where $x^*(C) \in C$ is one of the points at which F has been evaluated in this formula.

Further, we chose a sequence of n-tuples ${}^0\sigma, {}^1\sigma, {}^2\sigma, \ldots$ of positive numbers converging montonically to zero and define a sequence $C_0, C_1, \ldots$ of n-dimensional Gaussian bounded spaces as follows:

$$C_0 = C({}^0x^g, {}^0\sigma),\ C_1 = C(x^*(C_0), {}^1\sigma)$$

$$C_{i+1} \begin{cases} = C(x^*(C_i), {}^{i+1}\sigma); & \text{if} \quad m_N(C_1) < m_N(C_{i-1}) \\ = C(x^*(C_{i-1}), {}^{i+1}\sigma); & \text{otherwise.} \end{cases} \tag{5}$$

Thus the sequence $m_N(C_0)$, $m_N(C_1)$, $\ldots$ $m_N(C_i)$ is non-increasing and for suitably large i the value of $x^*(C_i)$ is taken an estimate of the optimized parameters. Here ${}^0x^g$ and ${}^0\sigma$ are the initial guesses of the respective parameters x^g and σ.

In selecting the sequence ${}^0\sigma$, ${}^1\sigma$, ${}^2\sigma$ it should be noted that if the sequence converges too quickly to zero, it may happen that the sequence of $m_N(C_i)$ does not converge to M. For the present study we have, by trial and error, found the following algorithm to be effective

$${}^1\sigma = {}^0\sigma \times {}^1t,$$

$${}^i\sigma = {}^{i-1}\sigma \times {}^it$$

where $${}^it = {}^{i-1}t \times 0{\cdot}95 \quad \text{and} \quad {}^1t = 0{\cdot}9 \tag{6}$$

The algorithm described above, when used with a random sequence, is a Monte Carlo method for minimizing F. The performance of such a sequence will be measured and compared with performance of other sequences in the following manner. We will use a number of carefully selected test functions. Let χ^2 be one of these functions and ${}^1x$, ${}^2x$, $\ldots$ ${}^Nx$ a sequence generated by a low discrepancy random sequence in I^s, with N a fixed integer. Let

$$m_N = \min_{1 \leqslant i \leqslant N} [\chi^2({}^ix)] = \chi^2({}^lx), \quad 1 \leqslant l \leqslant N,$$

where $$\chi^2({}^ix) = \frac{1}{L-S} \sum_{i=1}^{L} [y_i(x^*) - y_i({}^ix)]^2. \tag{7}$$

The function χ^2 attains its minimum at x^* and $\chi^2(x^*) = 0$. The quantity m_N is used as a measure of performance of the sequence with respect to the minimization of χ^2. Also, the quantity $\delta_i = [(x_i^* - {}^1x_i)/x_i^*] \times 100$ is used to show in certain selected cases the convergence of the variables in the minimization problem. The forms of the function $y_i(x)$ used in the present study depend on the dimensionality of the problem. We now describe these functions in dimensions 3, 4 and 5 and the numerical results obtained.

3. Results and discussion

We now present numerical results as values of χ^2 in the fourth and eighth iterations (i.e. $i = 4$ and $i = 8$ in equation (5)). Each iteration consists of a set of 500 random searches i.e. $N = 500$.

In three dimensions the test function chosen is of the following type:

$$Y_i(x) = {}^1x \exp(c_i - {}^2x)^2/{}^3x,$$

where c_i's represent the x-axis values of the distribution Y_i. Four different test sets y_i ($\mathbf{x}^*$) are generated and the random search is initiated with initial values of x different from x^*. Table 1 gives the results in three dimensions for following sets of x^* and the corresponding initial guesses:

(a) $x^* = [35600, 34{\cdot}263, 4{\cdot}964]$; initial guess = $[35197, 34{\cdot}0, 5{\cdot}2]$

(b) $x^* = [50291, 40{\cdot}016, 6{\cdot}066]$; initial guess = $[49963, 40{\cdot}0, 6{\cdot}98]$

(c) $x^* = [29128, 109{\cdot}65, 20{\cdot}137]$; initial guess = $[28901, 110{\cdot}0, 21{\cdot}173]$

(d) $x^* = [57187, 126{\cdot}16, 23{\cdot}775]$; initial guess = $[57034, 126{\cdot}0, 25{\cdot}44]$

From table 1, though it is difficult to judge precisely the relative merits of different sequences, it is evident that the pseudo-random sequence performs poorly compared to the quasi-random sequences. Among the quasi-random sequences, no sequence gives the best performance consistently for all the four cases; for example the Zaremba sequence gives the best performance for spectra number 4 whereas it is worst for spectra number 1. However judging from the overall performance it appears that the Haber sequence is best, the next is the scrambled Halton sequence and then the Zaremba sequence.

As for the performance of the method itself, the method works very well for this case. This is evident from table 2 which gives the values of δ_i for each variable at iteration steps 4 and 8 for spectra number 1. The table indicates good convergence even in the fourth iteration step and at the eight iteration step the convergence is almost absolute. The zeros in the table indicate the values less than 0·0005.

In four dimensions the test function is of the type

$$y_i(x) = {}^1x\exp[(c_i - {}^2x)^2/{}^3x] + {}^4x|c_i - {}^2x|^4.$$

The values of x^* and initial guess values are as follows:

1. $x^* = [35600, 34{\cdot}263, 4{\cdot}964, 2{\cdot}517\ (-9)]$
   initial guess = $[35197, 34{\cdot}0, 5{\cdot}2, 1{\cdot}125\ (-8)]$
2. $x^* = [50291, 40{\cdot}016, 6{\cdot}066, 2{\cdot}611\ (-9)]$
   initial guess = $[49963, 40{\cdot}0, 6{\cdot}98, 2{\cdot}213\ (-10)]$
3. $x^* = [29128, 109{\cdot}65, 20{\cdot}137, 1{\cdot}351\ (-9)]$
   initial guess = $[28901, 110{\cdot}0, 21{\cdot}173, 7{\cdot}4563\ (-8)]$
4. $x^* = [57187, 126{\cdot}16, 23{\cdot}775, 1{\cdot}119\ (-9)]$
   initial guess = $[57034, 126{\cdot}0, 25{\cdot}44, 9{\cdot}873\ (-8)]$.

Table 3 gives the values of χ^2 for the four-dimensional study of the sequences. Here again the pseudo-random sequence performs poorly compared to the quasi-random sequences. However, for spectra number 1, it gives the best performance. Among the QRS, sequence 6 gives the best performance followed by the Haber and the Zaremba sequences. In this case, the performance of the scrambled Halton sequence has deteriorated. In fact, the Halton sequence is a shade better than the scrambled Halton sequence.

Table 4 gives the values of δ_i for spectra number 1 of this case. In this case also the convergence of the parameters is very good except for the parameter ${}^4x$. This is

Table 1. Values of χ^2 in three dimensional test cases with different random sequences.

Spectra number	Iteration number	HAM	HAL	ZAR	HAB	SEQ-5	SEQ-6	SCR-HAL	PRS
1	4	2·87(+4)*	8·95(+4)	1·17(+5)	2·86(+4)	5·02(+4)	1·73(+4)	3·37(+4)	7·39(+4)
	8	6·98(−1)	2·00(−2)	1·70(+0)	3·31(−1)	1·51(+0)	1·21(+0)	2·29(−1)	8·45(−1)
2	4	1·31(+5)	2·41(+5)	2·83(+5)	1·10(+5)	3·16(+5)	2·15(+5)	1·50(+5)	6·51(+5)
	8	4·81(+0)	2·04(+0)	7·41(−1)	1·41(+0)	1·90(+0)	3·99(+0)	2·64(+0)	5·86(+0)
3	4	3·11(+8)	2·72(+8)	1·19(+8)	2·33(+8)	2·09(+8)	3·07(+8)	2·85(+8)	9·53(+8)
	8	4·95(+3)	6·04(+3)	1·55(+3)	8·20(+3)	5·92(+3)	1·38(+5)	5·24(+3)	1·03(+7)
4	4	4·08(+9)	2·24(+9)	5·77(+8)	2·28(+8)	1·06(+9)	9·45(+8)	6·11(+8)	1·38(+9)
	8	3·73(+8)	5·38(+7)	8·23(+6)	2·47(+5)	1·05(+6)	5·20(+4)	9·72(+3)	6·53(+7)

* To be read as $2{\cdot}87 \times 10^{+4}$.

Table 2. Values of δ in three-dimensional test cases with different random sequences for spectra number 1.

Iteration number	HAM	HAL	ZAR	HAB	SEQ-5	SEQ-6	SCR-HAL	PRS
	2·92(−3)*	2·92(−3)	2·92(−3)	8·76(−3)	2·92(−3)	0·00(+0)	2·92(−3)	2·92(−3)
4	5·06(−2)	1·15(−1)	8·12(−2)	1·40(−2)	3·37(−2)	1·69(−2)	5·34(−2)	1·78(−2)
	1·17(−1)	1·11(−1)	1·71(−1)	2·03(−1)	1·41(−2)	3·83(−2)	6·85(−2)	5·64(−2)
	0·00(+0)	0·00(+0)	0·00(+0)	0·00(+0)	0·00(+0)	0·00(+0)	0·00(+0)	0·00(+0)
8	0·00(+0)	0·00(+0)	0·00(+0)	0·00(+0)	0·00(+0)	0·00(+0)	0·00(+0)	0·00(+0)
	2·01(−3)	0·00(+0)	0·00(+0)	0·00(+0)	0·00(+0)	0·00(+0)	0·00(+0)	0·00(+0)

* To be read as $2{\cdot}92 \times 10^{-3}$.

Table 3. Values of χ^2 in four-dimensional test cases with different random sequences.

Spectra number	Iteration number	HAM	HAL	ZAR	HAB	SEQ-5	SEQ-6	SCR-HAL	PRS
1	4	3·26(+4)*	8·77(+4)	1·13(+5)	2·97(+4)	6·10(+4)	1·77(+4)	3·79(+4)	7·99(+4)
	8	2·03(+2)	2·51(+3)	2·27(+3)	2·25(+2)	1·51(+3)	1·77(+2)	1·23(+3)	1·79(+1)
2	4	1·47(+5)	2·59(+5)	2·95(+5)	1·30(+5)	3·17(+5)	2·38(+5)	1·62(+5)	7·28(+5)
	8	1·27(+4)	1·19(+3)	1·00(+3)	1·07(+3)	1·26(+3)	1·34(+3)	1·20(+3)	1·29(+3)
3	4	1·30(+11)	6·72(+9)	5·41(+9)	1·11(+10)	3·64(+9)	1·75(+9)	2·33(+10)	6·89(+10)
	8	1·07(+11)	2·58(+9)	1·06(+9)	1·41(+9)	2·02(+9)	6·50(+8)	1·44(+10)	3·55(+10)
4	4	3·94(+11)	9·59(+10)	1·21(+11)	1·63(+11)	3·48(+11)	5·43(+10)	2·84(+11)	2·54(+11)
	8	3·32(+11)	5·46(+10)	9·44(+10)	9·80(+10)	3·06(+11)	5·05(+10)	2·63(+11)	2·07(+11)

* To be read as $3{\cdot}26 \times 10^{+4}$.

Table 4. Values of δ in four-dimensional test cases with different random sequences for spectra number 1.

Iteration number	HAM	HAL	ZAR	HAB	SEQ-5	SEQ-6	SCR-HAL	PRS
	2·92(−3)*	2·92(−3)	2·92(−3)	8·76(−3)	0·00(+0)	0·00(+0)	2·92(−3)	2·92(−3)
4	5·06(−2)	1·15(−1)	8·15(−2)	1·40(−2)	3·37(−2)	1·69(−2)	5·34(−2)	1·18(−2)
	1·17(−1)	1·11(−1)	1·71(−1)	2·03(−1)	1·41(−2)	3·83(−2)	6·85(−2)	5·64(−2)
	3·67(+3)	7·42(+2)	7·48(+2)	5·84(+2)	3·11(+3)	2·20(+1)	5·35(+2)	1·03(+2)
	0·00(+0)	0·00(+0)	0·00(+0)	0·00(+0)	0·00(+0)	0·00(+0)	0·00(+0)	0·00(+0)
8	2·81(−3)	2·81(−3)	2·81(−3)	2·81(−3)	0·00(+0)	0·00(+0)	2·81(−3)	0·00(+0)
	8·06(−3)	2·42(−2)	2·22(−2)	2·01(−2)	6·04(−3)	6·04(−3)	1·61(−2)	2·01(−3)
	2·66(+2)	7·67(+2)	7·30(+2)	5·96(+2)	2·29(+2)	3·42(+0)	5·36(+2)	6·23(+1)

* To be read as $2{\cdot}92 \times 10^{-3}$.

expected to be so because of the relatively low importance of this parameter in the function to be minimized.

In five dimensions we have the test function as

$$Y_i(\mathbf{x}) = {}^1x\exp[(c_i - {}^2x)^2/{}^3x] + {}^4x|c_i - {}^2x|^4 + {}^5x|c_i - {}^2x|^{12}.$$

The values of x^* and the initial guess values are as follows:

1. x^* = [35600, 34·263, 4·964, 2·517 (−4), 3·911 (−10)]
   initial guess = [35197, 34·0, 5·2, 1·125 (−3), 1·526 (−11)]
2. x^* = [50291, 40·016, 6·066, 2·611 (−4), 1·595 (−10)]
   initial guess = [49963, 40·0, 6·98, 2·213 (−3), 1·111 (−11)]
3. x^* = [29128, 109·65, 20·137, 1·351 (−4), 7·129 (−13)]
   initial guess = [28901, 110·0, 7·4563 (−5), 8·9235 (−12)]
4. x^* = [57187, 126·16, 23·775, 1·119 (−4), 3·205 (−13)]
   initial guess = [57034, 126·0, 25·44, 9·873 (−3), 1·2345 (−11)].

Table 5 gives the values of χ^2 for five-dimensional cases. It is seen that the overall performance of the PRS is better than some of the QRS. Among the QRS, SEQ-6 gives the best performance followed by the Zaremba sequence and the Hammersley sequence. Here also the scrambled Halton sequence does not show any improvement over the Halton sequence, and both perform poorly compared to the PRS.

Table 6 gives the values of δ_i for the five-dimensional case for spectra number 1. Here again the overall convergence is good except for the parameters ${}^4x$ and ${}^5x$ which is expected.

From the numerical study carried out it is evident that the quasi-random search technique with the Gaussian importance sampling scheme works well for the type of least square data fitting described in the text, even with 500 random searches per iterative step at least up to 5-dimensions. The PRS performs poorly, in general, compared to the QRS, at least to some of them. However, there is an indication that with the increase in the dimensionality, the relative performance of the PRS is improving and at 5-dimensions it has outperformed some of the QRS on an overall basis. Among the QRS, the Zaremba sequence has shown consistently good performance over the range of dimensions considered here. At dimensions 3 and 4 the Haber sequence gives good performance whereas at dimensions 4 and 5 sequence 6 gives good performance. There is no advantage gained by scrambling the Halton sequence—a result which contradicts those obtained for multidimensional integrals (Braaten and Weller 1979), but supports what is obtained for integral equations (Sarkar and Prasad 1987).

In the present study no attempt was made to compare the efficiencies of the different sequences which would also have involved the estimation of the generation time for each sequence. This was because the relative importance of the generation time for any sequence will greatly depend on the type of problem studied. Secondly, and more importantly, in the present scheme a set of 500 random numbers can be generated and stored once for all at the beginning of the computation or they may be given as an input data. However, as one would except, the QRS takes much more computing time than the PRS.

Table 5. Values of χ^2 in five-dimensional test cases with different random sequences.

Spectra number	Iteration number	HAM	HAL	ZAR	HAB	SEQ-5	SEQ-6	SCR-HAL	PRS
1	4	8·68(+5)*	1·24(+6)	8·38(+5)	8·20(+5)	8·24(+5)	9·06(+6)	1·09(+6)	8·24(+5)
	8	8·14(+5)	7·77(+5)	6·76(+5)	7·29(+5)	7·17(+5)	7·67(+5)	7·95(+5)	7·11(+5)
2	4	9·23(+6)	3·15(+7)	1·05(+7)	9·45(+6)	9·91(+6)	8·82(+6)	1·32(+7)	9·60(+6)
	8	8·47(+6)	1·90(+7)	9·36(+6)	8·82(+6)	9·38(+6)	8·36(+6)	1·01(+7)	8·75(+6)
3	4	8·61(+10)	5·83(+10)	1·05(+11)	2·23(+11)	2·18(+11)	1·84(+10)	7·18(+10)	1·90(+11)
	8	5·62(+10)	4·97(+10)	9·41(+10)	1·72(+11)	1·23(+11)	5·91(+9)	2·39(+10)	1·58(+11)
4	4	1·01(+13)	3·93(+13)	2·53(+12)	3·40(+13)	4·38(+12)	3·46(+12)	3·66(+13)	3·46(+13)
	8	5·79(+11)	3·21(+13)	8·47(+11)	3·33(+13)	3·47(+11)	7·75(+11)	3·56(+13)	3·38(+13)

*To be read as $8{\cdot}68 \times 10^{+5}$.

Table 6. Values of δ in five-dimensional test cases with different random sequences for spectra number 1.

Iteration number	HAM	HAL	ZAR	HAB	SEQ-5	SEQ-6	SCR-HAL	PRS
	2·92(−3)*	1·75(−2)	8·76(−3)	5·84(−3)	2·92(−3)	2·92(−3)	8·76(−3)	5·84(−3)
	9·55(−2)	1·25(−1)	6·74(−2)	3·65(−2)	1·85(−1)	4·49(−2)	8·99(−2)	0·00(+0)
4	4·25(−1)	4·37(−1)	4·43(−1)	2·84(−1)	5·90(−1)	7·29(−1)	3·44(−1)	3·32(−1)
	5·80(+1)	1·63(+0)	4·53(+1)	4·08(+1)	5·08(+1)	7·20(+1)	2·15(+1)	3·97(+1)
	9·90(+1)	9·82(+1)	9·20(+1)	9·46(+1)	9·60(+1)	9·79(+1)	9·79(+1)	9·43(+1)
	0·00(+0)	0·00(+0)	0·00(+0)	0·00(+0)	0·00(+0)	2·92(−3)	0·00(+0)	0·00(+0)
	6·46(−2)	4·21(−2)	5·62(−2)	6·18(−2)	6·46(−2)	6·46(−2)	1·40(−2)	5·62(−2)
8	5·10(−1)	3·53(−1)	4·43(−1)	4·63(−1)	4·92(−1)	5·14(−1)	1·99(−1)	4·41(−1)
	5·03(+1)	3·93(+1)	4·43(+1)	4·63(+1)	4·82(+1)	4·99(+1)	3·08(+1)	4·44(+1)
	9·89(+1)	9·82(+1)	9·19(+1)	9·47(+1)	9·55(+1)	9·81(+1)	9·79(+1)	9·43(+1)

*To be read as $2{\cdot}92 \times 10^{-3}$.

4. Conclusions

(i) The quasi-random search technique with importance sampling can be effectively used for nonlinear least square data fitting.

(ii) The pseudo-random sequence from the congruential generator does not perform well as compared to the quasi-random sequences at least up to five dimensions.

(iii) The Zaremba sequence shows consistently good performance. The Haber sequence and sequence 6 also performs well. No advantage was obtained by scrambling the Halton sequence.

Acknowledgement

We thank Dr Harald Niederreiter for useful suggestions and correspondence.

References

Aird T J and Rice J R 1977 *SIAM J. Numer. Anal.* **14** 296
Bandyopadhyay T and Sarkar P K 1985 *Pramana–J. Phys.* **24** 643
Braaten E and Weller G 1979 *J. Comp. Phys.* **33** 249
Fox B L 1986 Implementation and relative efficiency of quasi random sequence generators (to be published)
Haber S 1970 *SIAM Rev.* **12** 481
Halton J H 1960 *Numer. Math.* **2** 84
Halton J H and Zaremba S K 1969 *Montash. Math.* **73** 316
Niederreiter H 1978 *Bull. Amer. Math. Soc.* **84** 957
Niederreiter H 1983 in *Studies in pure mathematics* (Budapest: Akademiaikiado) p. 523
Niederreiter H and McCurley K 1979 *Computing* **22** 119
Niederreiter H and Peart P 1982 *Caribbean J. Math.* **1** 27
Sarkar P K and Prasad M A 1987 *J. Comp. Phys.* **67** 257
Sobol I M 1979 *SIAM J. Numer. Anal.* **16** 790
Warnock T T 1972 *Application of number theory to numerical analysis* (ed.) S K Zaremba (New York: Academic Press) p. 319
Zaremba S K 1968 *Ann. Polon. Math.* **21** 85