GRID WITH UNIFORMITY ADAPTED TO THE STRUCTURE OF A MULTIDIMENSIONAL PROBLEM

Vydu纳斯 ŠALTENIS

Institute of Mathematics and Informatics
2600 Vilnius, Akademijos 4, Lithuania
E-mail: saltenis@ktl.mii.lt

Abstract. In the present paper, a method of global optimisation (structure adapted search) is proposed. It uses the grid of trial points which are more uniformly distributed for the projections on variables or their groups that make more influence.

The paper uses a set of test models to demonstrate the merit of the approaches. The efficiency of structure adapted search as compared to the random search is investigated. The results of using the new approach may be treated as a success.

Key words: global optimisation, random search, uniformity, structure.

1. Introduction. Passive algorithms of global optimisation are of interest, especially at the beginning of optimisation, as their trial points do not depend on the known function values. One of the most important features of passive algorithms of global optimisation is the uniformity of trial points according to some criteria in the minimisation region – a multidimensional cube $K^n$.

The uniformity of trial points in a grid must be good not only for the entire optimisation region but also for the faces of smaller dimensionality. We can fix $s$ indices $1 \leq i_1 < \cdots < i_s \leq n$ ($1 \leq s \leq n$), and denote the $s$-dimensional face of $K^n$ as $K_{i_1 \ldots i_s}$. Then the projections of points on every face $K_{i_1 \ldots i_s}$ and for all $s$ ($1 \leq s \leq n$) must also be uniformly distributed. In a multidimensional case there are principle difficulties in achieving the uniformity for all projections simultaneously.

Random search is widely used in global optimisation as a passive algorithm. The main reason, explaining the popularity of random search is that it is simple to be realised on a computer. However, the realisation of uniform pseudo random points results in a nonuniform distribution of points. Special grids with a better uniformity were proposed (Halton, 1960; Niederreiter and Peart, 1982; 1986; Sobol', 1969; 1979).

In this paper we tackle the other end of the problem: we seek to adapt the uniformity of various projections to the degree of influence of the variables or their
groups. In practical problems the degree of influence is very different (Courtois, 1985; Šaltenis, 1989) so the way of adapting the uniformity may be efficient.

In Section 2 we consider the special grid used for $\text{LP}_r$ search. In Section 3 the structure characteristics of multidimensional problems are introduced. They characterise the influence of variables or their groups and may be used in constructing our grid. The algorithm of structure adapted search and some two-dimensional examples are presented in Section 4. In Section 5 we can find test problems which were used for experimental investigations. And the results of the investigations are presented in Section 6. Finally, Section 7 contains a brief discussion and conclusions of the paper.

2. $\text{LP}_r$ search and requirements for uniformity of a grid. The efficiency of passive global search using special grids must depend on the problem structure, on the degree of influence of the variables or their groups on the objective function. The $\text{LP}_r$ sequences (Sobol', 1969; 1979) have a better uniformity than pseudo random grids. Projections of the points on the coordinate axes (on one-dimensional faces $K_i$) are extremely uniform for all $i$, while the projections on the faces of greater dimensionality are not so uniform. The uniformity of projections on faces $K_{i,i+1}$ and $K_{i,i+1,i+2}$ depends on the number $i$ (Šaltenis, 1984). So the efficiency of $\text{LP}_r$ search mainly depends on the proper numbering order of variables, according to the structure of the problem.

3. Structure characteristics of a multidimensional problem. A decomposition of a multivariate function into the summands of different dimensionality (Cukier et al., 1975; Šaltenis, 1989; Šaltenis, 1996; Sobol', 1990) makes the base for the structure analysis which we use in structure adapted search.

Let a function

$$f(X) = f(x_1, \ldots, x_n),$$

be defined, for simplicity, on the unit cube $K^n (0 \leq x_1 \leq 1, \ldots, 0 \leq x_n \leq 1)$:

$$X \in K^n.$$

Sometimes a short notation $f$ will be used for $f(x_1, \ldots, x_n)$.

Let us introduce some notation for the domain set of the function $f$, which is a Cartesian product of basic domains $\Omega_1, \ldots, \Omega_n$:

$$\Omega = \Omega_1 \times \cdots \times \Omega_n,$$
Grid with uniformity adapted

and for special domains:

\[ \Omega_{t_1 \ldots t_s} = \Omega_{t_1} \times \cdots \times \Omega_{t_s}, \quad 1 \leq t_1 < \ldots < t_s \leq n, \quad s = 1, \ldots, n, \]

\[ \Omega_{(i)} = \Omega_1 \times \cdots \times \Omega_{i-1} \times \Omega_{i+1} \times \cdots \times \Omega_n, \quad 1, \ldots, n, \]

\[ \Omega_{(ij)} = \Omega_1 \times \cdots \times \Omega_{i-1} \times \Omega_{i+1} \times \cdots \times \Omega_{j-1} \times \Omega_{j+1} \times \cdots \times \Omega_n, \]

\[ i, j = 1, \ldots, n, \quad i < j. \]

In the general case, the domain \( \Omega_{(t_1 \ldots t_s)} \) is defined in a similar way.

We use groups of indices \( t_1, \ldots, t_s \), where \( 1 \leq t_1 < \ldots < t_s \leq n \), \( s = 1, \ldots, n \), and denote the sum consisting of \( 2^n - 1 \) terms as:

\[ \sum_{t_1 \ldots t_s} T_{t_1 \ldots t_s} = \sum_{i=1}^{n} T_i + \sum_{1 \leq i < j \leq n} T_{ij} + \cdots + T_{12 \ldots n}. \]

The decomposition of the function \( f \) into the summands of different dimensionality

\[ f = f_0 + \sum f_{t_1 \ldots t_s}(x_{t_1}, \ldots, x_{t_s}) \quad (1) \]

is unique and orthogonal for each function \( f \) integrable on \( K^n \) (Sobol', 1990), if \( f_0 \) is constant and the integrals of summands (1) are equal to zero:

\[ \int_{0}^{1} f_{t_1 \ldots t_s}(x_{t_1}, \ldots, x_{t_s}) \, dx_{t_k} = 0, \quad 1 \leq k \leq s. \]

The summands of decomposition (1) may be found just like some integrals.

Let us introduce the following notation for the function of \( s \) variables:

\[ f^{t_1 \ldots t_s} = \int_{\Omega(t_1 \ldots t_s)} f, \]

where the integral is over all basic domains except \( \Omega_{t_1}, \ldots, \Omega_{t_s} \).

Then, after integrating (1) on \( \Omega \), the constant summand will be equal to

\[ f_0 = \int_{\Omega} f, \]

after integrating (1) on \( \Omega_{(i)} \), one-dimensional summands will be equal to

\[ f_i(x_i) = f^i - f_0, \quad i = 1, \ldots, n, \]
after integrating (1) on \( \Omega_{ij} \), two-dimensional summands will be equal to

\[
f_{ij}(x_i, x_j) = f^{ij} - f_0 - f_i(x_i) - f_j(x_j), \quad i, j = 1, \ldots, n, \ i < j,
\]

and so on.

The system of structure characteristics:

\[
D = \sum D_{i_1 \ldots i_s},
\]

\[
D_{i_1 \ldots i_s} = \int_{\Omega} (f_{i_1 \ldots i_s})^2,
\]

(3)

where

\[
D = \int_{\Omega} (f)^2 - (f_0)^2
\]

was proposed, investigated by (Saltenis, 1989) and applied in analysing the structure of optimisation problems. The structure characteristics \( D_{i_1 \ldots i_s} \) indicate the degree of influence of the respective variable groups in the approximation. Usually the characteristics are normalised (the sum of all characteristics is equal to one).

The normalised structure characteristics of some simple functions \((0 \leq x_1 \leq 1, \ 0 \leq x_2 \leq 1)\) are presented in Table 1.

<table>
<thead>
<tr>
<th>( f(X) )</th>
<th>( D_1 )</th>
<th>( D_2 )</th>
<th>( D_{12} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x_1 + x_2 )</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>( x_1 x_2 - \frac{x_1 + x_2}{2} )</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1. The structure characteristics of some simple functions

If we know the values of the function \( f(X) \) for some points \( X^j \) \((j = 1, \ldots, N)\), then the Monte-Carlo method is used for evaluations basing on (2), (3) and (4):

\[
f_0 \approx \frac{1}{N} \sum_{j=1}^{N} f(X^j),
\]

\[
D + (f_0)^2 \approx \frac{1}{N} \sum_{j=1}^{N} (f(X^j))^2,
\]

where \( N \) is the number of samples,
Grid with uniformity adapted

$X^j = (x_1^j, \ldots, x_n^j)$ are random points of dimensionality $n$, uniformly distributed in $\Omega$.

$s$ coordinates $Y^j = (x_{i_1}, \ldots, x_{i_s})$ must be identical for pairs of random points used for the evaluation of $D_{i_1\ldots i_s}$ (Sobol', 1990):

$$D_{i_1\ldots i_s} + (f_0)^2 \approx \frac{1}{N} \sum_{j=1}^{N} f(Y^j, Z^j) f(Y^j, U^j),$$

where $Y^j$ are random points of dimensionality $s$, uniformly distributed in $\Omega_{i_1\ldots i_s}$,

$Z^j$ and $U^j$ are random points of dimensionality $n - s$, uniformly distributed in $\Omega_{(i_1\ldots i_s)}$.

4. Algorithm for generating a structure adapted grid. One of the characteristics of uniformity of a set of points is the dispersion of points $X^1, X^2, \ldots, X^N$; $X^j = (x_1^j, \ldots, x_n^j)$:

$$d(X^1, X^2, \ldots, X^N) = \sup_{0 \leq X \leq 1} \min_{1 \leq j \leq N} |X - X^j|. \quad (5)$$

We find here the closest point to $X$ (the distance to the closest point is equal to $\min_{1 \leq j \leq N} |X - X^j|$) and choose $X$ in the worst way.

In a one-dimensional case, the lower bound for $d$ is equal to:

$$d(X^1, X^2, \ldots, X^N) \geq \frac{1}{2N}.$$  

The uniformity of projections of the points on faces is measured by the same dispersion of points (5). The difference consists only in calculating the distance $|X - X^j|$.

For example, in the case of projections on one-dimensional faces $K_i$, the distance will be equal to:

$$|X - X^j| = |x_i - x_i^j|.$$  

In the general case, the dispersions of projections on $s$-dimensional faces $K_{i_1, \ldots, i_s}$ are equal to:

$$d_{i_1, \ldots, i_s} = \sup_{0 \leq X \leq 1} \min_{1 \leq j \leq N} \sqrt{\sum_{i=i_1, \ldots, i_s} (x_i - x_i^j)^2}.$$
If we want to get a grid with the best uniformity of projections on the $s$-dimensional face $K_{i_1, \ldots, i_s}$, we have to minimise the dispersion:

$$\min_{X^1, X^2, \ldots, X^N} d_{i_1, \ldots, i_s} (X^1, X^2, \ldots, X^N).$$

The best uniformity for all projections simultaneously is a multicriteria optimisation problem. If we want to get the solution of the problem we have to minimise some weighted sum:

$$\min_{X^1, X^2, \ldots, X^N} \sum w_{i_1, \ldots, i_s} d_{i_1, \ldots, i_s} (X^1, X^2, \ldots, X^N),$$

where $\sum$ denotes the sum consisting of $2^n - 1$ terms for all possible indices $i_1, \ldots, i_s$ and $w_{i_1, \ldots, i_s}$ are weights ($\sum w_{i_1, \ldots, i_s} = 1$), used for multicriteria minimisation.

It is natural to use the system of normalised structure characteristics instead of weights:

$$w_{i_1, \ldots, i_s} = \frac{D_{i_1, \ldots, i_s}}{\sum D_{i_1, \ldots, i_s}}$$

The main steps of the algorithm for minimising the weighted sum (6) are:

1. Generate $N$ starting random points $X^1, X^2, \ldots, X^N$, uniformly distributed in the minimisation region.
2. Find the closest point to the point $X^1$ in one of possible projections on faces $K_{i_1, \ldots, i_s}$.
3. Change the coordinates of the point $X^1$ to increase the distance to the closest point. The amount of this change is proportional to the weight $w_{i_1, \ldots, i_s}$.
4. Repeat Step 3 for all projections on faces $K_{i_1, \ldots, i_s}$.
5. Repeat Steps 2, 3 and 4 for all points $X^1, X^2, \ldots, X^N$.
6. Repeat Steps 2, 3, 4 and 5 until is possible to reduce the weighted sum (6).

We use the term a "structure adapted" (SA) grid in this case. Thus, the points of the SA grid will be more uniformly distributed for the main variables and their groups.

Figs. 1–4 present simple illustrations of the SA grid in the case of $n = 2, N = 4$.

In Fig. 1 the weights were equal to $D_{12} = 1; D_1 = 0; D_2 = 0$. We can see four points uniformly distributed on the two-dimensional region, but their projections on axes are extremely nonuniform (some projections are even the same). In
Grid with uniformity adapted

Fig. 1. SA grid in the case $D_{12} = 1; D_1 = 0; D_2 = 0$.

Fig. 2. SA grid in the case $D_{12} = 0.8; D_1 = 0.1; D_2 = 0.1$.

Fig. 3. SA grid in the case $D_{12} = 0.6; D_1 = 0.2; D_2 = 0.2$.

Fig. 4. SA grid in the case $D_{12} = 0; D_1 = 0.5; D_2 = 0.5$.

Fig. 2 and Fig. 3 the uniformity of projections is introduced step by step. Fig. 4 illustrates the case of maximal uniformity of projections, where no attention is paid to the uniformity in the two-dimensional region ($D_{12} = 0; D_1 = 1; D_2 = 1$).

We can see that a cubic grid is the optimal grid in the case of maximal interaction ($D_{12} = 1$) while the grid with uniformly distributed projections on the coordinate axis is optimal, if the interaction is absent ($D_{12} = 0$).
Table 2. The structure characteristics of Branin test problem

<table>
<thead>
<tr>
<th></th>
<th>(D_1)</th>
<th>(D_2)</th>
<th>(D_{12})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>0.14</td>
<td>0.25</td>
<td>0.61</td>
</tr>
</tbody>
</table>

5. Basic test problems and their structure characteristics. We use a set of test models of various dimensionality \((n = 2, 6, 20)\) in our experiments to demonstrate the merit of the approaches.

The first two-dimensional multiextremal Branin test problem

\[
f(X) = \left( x_2 - \frac{5.1}{4\pi^2} + \frac{5}{\pi} x_1 - 6 \right)^2 + 10 \left( 1 - \frac{1}{8\pi} \right) \cos x_1 + 10,
\]

\(-5 \leq x_1 \leq 10, 0 \leq x_2 \leq 15,\)

was proposed in (Dixon and Czego, 1978) and is widely used. The values of its structure characteristics are investigated and presented in Table 2. One ought to pay attention to the relatively high interaction of variables \((D_{12} = 0.61)\).

The second test problem was proposed in Mathar and Žilinskas (1994). The objective function of the problem:

\[
f(X) = \sum_{i<j}^n (\delta_{ij} - d_{ij}(X))^2
\]

where \(X = (x_{11}, \ldots, x_{n1}, x_{12}, \ldots, x_{n2})\) and \(d_{ij}(X)\) denotes the Euclidean distance between the points \(x_{i1}, x_{i2}\) and \(x_{j1}, x_{j2}\). The matrix \((\delta_{ij})\) represents given pairwise dissimilarities among \(n\) objects. In multidimensional scaling the objective function (7) is called STRESS (Green et al., 1989). It formalises the problem to find \(n\) points in a metric space so that the interpoint distances fit the given dissimilarities. One of the proposed triangular matrixes of distances for \(n = 10\) was used in our test problem. It is presented in Table 3. The matrix corresponds to the experimental data on Cola testing in (Green et al. (1989)).

We use the test function in the case \(n = 10\) (dimensionality is equal to 20) with an abbreviated title "Mathar Žilinskas – 20". The optimisation region is the cube \([-1.2, 1.2]^2\).

The structure characteristics of the test function are symmetric. All first order characteristics \(D_i\) \((i = 1, \ldots, 20)\) are approximately equal, all second order characteristics \(D_{ij}\) \((i = 1, \ldots, 19; j = 2, \ldots, 20; i < j)\) are approximately equal and so on. Their values are presented in Table 4.
Table 3. The matrix of distances

<table>
<thead>
<tr>
<th></th>
<th>1.27</th>
<th>1.69</th>
<th>2.04</th>
<th>3.09</th>
<th>2.86</th>
<th>3.17</th>
<th>3.21</th>
<th>2.38</th>
</tr>
</thead>
<tbody>
<tr>
<td>**</td>
<td>1.43</td>
<td>2.35</td>
<td>3.18</td>
<td>3.22</td>
<td>2.56</td>
<td>3.18</td>
<td>3.18</td>
<td>2.31</td>
</tr>
<tr>
<td>**</td>
<td>2.43</td>
<td>3.26</td>
<td>3.27</td>
<td>2.58</td>
<td>3.18</td>
<td>3.18</td>
<td>2.42</td>
<td></td>
</tr>
<tr>
<td>**</td>
<td>2.85</td>
<td>2.88</td>
<td>2.59</td>
<td>3.12</td>
<td>3.17</td>
<td>1.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>**</td>
<td>1.55</td>
<td>3.12</td>
<td>1.31</td>
<td>1.70</td>
<td>2.85</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>**</td>
<td>3.06</td>
<td>1.64</td>
<td>1.36</td>
<td>2.81</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>**</td>
<td>3.00</td>
<td>2.95</td>
<td>2.56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>**</td>
<td>1.32</td>
<td>2.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>**</td>
<td>2.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4. The structure characteristics (in %) of the Mathar Žilinskas – 20 test function

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>First order</th>
<th>Second order</th>
<th>Third order</th>
<th>Characteristics of higher order</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_1$</td>
<td>2.1</td>
<td>1.1</td>
<td>0.1</td>
<td>0</td>
</tr>
<tr>
<td>$D_{ij}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D_{ijk}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The third test function of lower dimensionality 6 is based on the same problem of multidimensional scaling in case $n = 3$ (abbreviated title – "Mathar Žilinskas - 6"). It uses three first rows and three first columns of the matrix of distances in Table 3. The optimisation region is a cube $[-1.2,1.2]^6$.

The values of its structure characteristics are presented in Table 5.

The values of structure characteristics not included into the table (for example $D_3$, $D_4$) are approximately equal to zero.

6. Experiments and results. We investigate the efficiency of SA search as compared to the random search by analysing the optimisation results of some test function.

The experimental investigations of the efficiency are based on the averaged results of multiple optimisation. It is an obvious way to average the results of optimisation in the case of random search: we can repeat the optimisation by using different random trial points for the same test function.

It is difficult, however, to obtain some averaged results for SA search because
the trial points in this case may be treated as deterministic. The averaged results may be obtained, if we use the test functions of some class $f^{(i)}(X)$ ($i$ is the current number of the function in the class, $i = 1, \ldots, M$) generated on the base of some basic test function $f(X)$. Let us formulate the requirements to $f^{(i)}(X)$:

a) the minimal values for all the test functions of the class are equal:

$$\min_{X \in K^n} f^{(i)}(X) = \min_{X \in K^n} f(X), \quad (i = 1, \ldots, M);$$

b) the minimum points

$$\arg \min_{X \in K^n} f^{(i)}(X)$$

are different for $i = 1, \ldots, M$ and uniformly distributed at random in $K^n$;

### Table 5. The structure characteristics (in %) of the Mathar Žilinskas – 6 test problem

<table>
<thead>
<tr>
<th>Structure characteristic</th>
<th>The value of the characteristic</th>
<th>Structure characteristic</th>
<th>The value of the characteristic</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_1$</td>
<td>0.2</td>
<td>$D_{126}$</td>
<td>0.4</td>
</tr>
<tr>
<td>$D_2$</td>
<td>1.3</td>
<td>$D_{134}$</td>
<td>0.7</td>
</tr>
<tr>
<td>$D_3$</td>
<td>0.8</td>
<td>$D_{135}$</td>
<td>0.6</td>
</tr>
<tr>
<td>$D_{12}$</td>
<td>3.9</td>
<td>$D_{136}$</td>
<td>0.3</td>
</tr>
<tr>
<td>$D_{13}$</td>
<td>15.4</td>
<td>$D_{145}$</td>
<td>1.5</td>
</tr>
<tr>
<td>$D_{14}$</td>
<td>0.4</td>
<td>$D_{146}$</td>
<td>0.7</td>
</tr>
<tr>
<td>$D_{16}$</td>
<td>0.5</td>
<td>$D_{234}$</td>
<td>1.7</td>
</tr>
<tr>
<td>$D_{23}$</td>
<td>4.8</td>
<td>$D_{235}$</td>
<td>1.9</td>
</tr>
<tr>
<td>$D_{25}$</td>
<td>0.7</td>
<td>$D_{236}$</td>
<td>2.1</td>
</tr>
<tr>
<td>$D_{36}$</td>
<td>0.5</td>
<td>$D_{345}$</td>
<td>0.3</td>
</tr>
<tr>
<td>$D_{45}$</td>
<td>5.0</td>
<td>$D_{346}$</td>
<td>1.6</td>
</tr>
<tr>
<td>$D_{46}$</td>
<td>15.3</td>
<td>$D_{356}$</td>
<td>0.3</td>
</tr>
<tr>
<td>$D_{56}$</td>
<td>7.2</td>
<td>$D_{1245}$</td>
<td>6.2</td>
</tr>
<tr>
<td>$D_{123}$</td>
<td>3.4</td>
<td>$D_{1346}$</td>
<td>12.3</td>
</tr>
<tr>
<td>$D_{124}$</td>
<td>1.5</td>
<td>$D_{2235}$</td>
<td>6.9</td>
</tr>
<tr>
<td>$D_{125}$</td>
<td>1.8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Grid with uniformity adapted

Fig. 5. Investigation by the Branin test function.

c) the structure characteristics are equal to the characteristics of the basic problem:
\[ D_{\text{structure}}^{(i)} = D_{\text{basic}}^{(i)}, \quad i = 1, \ldots, M. \]

The next procedure satisfies the requirements formulated above.
\[ f^{(i)}(X) = f(X^*), \]

where

\[ X = (x_1, \ldots, x_n), \quad X^* = (x^*_1, \ldots, x^*_n), \]
\[ x^*_j = (x_j - \xi^{(i)}_j) \quad (\text{mod } 1), \quad i = 1, \ldots, M, \ j = 1, \ldots, n, \]

\( \xi^{(i)}_j \) is a random value, uniformly distributed in the interval [0, 1].

The results of experimental investigation are presented in Figs. 5–7 for three classes of test functions.

The averaged relationships \( \min f(N) \) are approximately linear, because the logarithmic scale was used for \( N \) and the values of test functions were transformed by a convenient monotonous transformation (usually by logarithmic). The linearity of the dependencies is useful while extrapolating our experimental results for greater values of \( N \).

We can see two lines in each figure (except for Fig. 5): one for random search, the other for SA search. Let us denote the minimal function value
achieved after \( N \) trial points of SA search as \( \min f(N) \) and the minimal function value achieved after \( N' \) trial points of random search as \( \min f'(N') \). Then
the efficiency of SA search can be evaluated by the ratio:

$$E = \frac{N'}{N},$$  \quad (8)

where the condition

$$\min f(N) = \min f'(N')$$  \quad (9)

must be satisfied.

The ratio $E$ shows how many times the number of trial points $N$ is reduced if we use the SA search and the minimal function values are the same in both cases.

In the linear case and for logarithmic scale

$$\min f(N) = k \log(N),$$  \quad (10)

and

$$\min f'(N') = k' \log(N').$$  \quad (11)

Then after substituting (10) and (11) into (8) and (9) we can get:

$$E = N^{\left(\frac{1}{k'} - 1\right)}. \quad (12)$$

Coefficients $k$ and $k'$ may be calculated directly from linear dependencies illustrated by the graphs in Figs. 5–7, so we can simply evaluate the forecast of the efficiency $E$ of SA search as compared to random search by (12).

Fig. 8 illustrates the efficiency for various test functions and various $N$.

The forecast of the efficiency is higher for the Branin test function because its structure is less complicated in comparison to the Mathar Žilinskas test functions. The SA search may be 2–6 times more efficient for the Branin function while for other test functions the efficiency amounts 1.8–2.2 only for extremely large numbers of $N$.

The third graph “Bad SA search” is presented for the Branin function in Fig. 5. It illustrates the efficiency of SA search with especially changed weights of uniformity. In this case, the uniformity was maximised only for two-dimensional projections while the uniformity of one-dimensional projections remained bad. In other words, the “bad SA search” assumed that structure characteristics of the test function are equal to $D_1 = 0; \ D_2 = 0; \ D_{12} = 1$, while the true values were equal to: $D_1 = 0.14; \ D_2 = 0.25; \ D_{12} = 0.61$. We can see that this bad SA search is considerably worse than the pure random search.
7. Conclusions. The main point of this paper was to show that the SA search can be more efficient than other passive global search algorithms, especially in the cases where the first order structure characteristics are relatively small.

At the same time we must bear in mind that the SA search uses structure characteristics of the problem. The statistical evaluation of the characteristics is relatively expensive, so the general efficiency in some cases may be reduced. In the cases, where the structure is approximately known a priori (for example, the structure is common for all problems from some class), the SA search may be used especially in a proper way.

REFERENCES


Received November 1997
V. Šaltenis graduated from the Kaunas Technological Institute, Lithuania, in 1959. He received the Ph.D. degree from the Moscow Energy Institute of the USSR Academy of Sciences in 1966. He is a senior researcher of the Department of optimisation at the Institute of Mathematics and Informatics, Lithuania. His present research interests include both theory and applications of the structure of optimisation problems, multicriteria decision support systems.

TAŠKAI, KURIŲ IŠDĖSTYMO TOLYDUMAS ATITINKA DAUGIAMAČIO UŽDavinio Struktūrą

Vydūnas ŠALTENIS


Atlikti eksperimentai, kurie įvertina tokio metodo efektyvumą, lyginant ji su atsitikin-}

tine pateikta.