INFORMATICA, 1990, Vol.1, No.1, 125–140

BRANCH AND PROBABILITY BOUND METHODS FOR GLOBAL OPTIMIZATION

Anatoly ZHIGLJAVSKY

Dep. of Mathematics, Leningrad University,
198904 Petrodvorets-Leningrad,
Bibliotechnaya-pl. 2, USSR

Abstract. The maximization problem for an objective function f given on a feasible region X is considered, where X is a compact subset of R^n and f belongs to a set of continuous multiextremal functions on X can be evaluated at any point x in X without error, and its maximum $M = \max_{x \in X} f(x)$ together with a maximizer x^*(a point x^* in X such that $M = f(x^*)$) are to be approximated. We consider a class of the global random search methods, underlying an apparatus of the mathematical statistics and generalizing the so-called branch and bound methods.

Key words: global optimization, random search, branch and bound, statistical inferences.

1.Gist of Branch and Bound Methods. Branch and bound methods used to advantage in various extremal problems (for instance, see Horst (1986)) consist in rejecting some of the subsets of X that do not for sure contain a maximizer and searching only among the rest of sets regarded as promising. Branch and bound methods may be schematized as a successive implementation at each iteration of the following three stages:

i) branching (decomposition) of the original set into the tree of subsets and evaluating the objective function values at some points in the subsets;

ii)estimation of functionals of the objective function over the obtained subsets (criteria of subset prospectiveness);

iii) selection of subsets that are promising for further branching.

In standard versions of branch and bound methods, deterministic upper bounds of the maximum of f on subsets are used as subset prospectiveness criteria. In doing so, all subsets Z whose upper bounds for $\sup\limits_{Z} f$ do not exceed the maximal of already obtained values of f are rejected. Prospectiveness criteria in these methods, thus, can be either 1 or 0,what means that branching of a subset should go on or be stopped.

Further, consideration will be given to non-standard variants of branch and bound methods that will be called branch and probability bound ones. Their distinction is that the maximum estimates on subsets are probabilistic (true with a close-to-one probability) rather than deterministic ones. The key notion arising under the description of the branch and probability bound methods is that of the set prospectiveness.

2. Prospectiveness Criteria. We shall call a subadditive set function $\varphi : B \to R^1$, resulting from processing the preceding values of the objective function and reflecting the possibility of locating the global maximizer in subsets, a prospectiveness criterion. Here B is the δ-algebra of the Borel subsets of X .If for two sets $Z_1, Z_2 \in B$ the inequality $\varphi(Z_1) \geqslant \varphi(Z_2)$ holds,then location of the global maximizer in Z_1 is according to the prospectiveness criterion φ at least as likely as in Z_2 .

Set functions φ assuming for $Z \in B$ the following values can be used as prospectiveness criteria:

a) $\varphi(Z)$ is an estimate for $M_Z = \sup\limits_Z f$,

b) $\varphi(Z)$ is a mean value $\int_Z f(z)\nu(dz)$ estimate,

c) $\varphi(Z)$ is a minimum $\min\limits_Z f$ estimate,

d) $\varphi(Z)$ is an upper confidence bound of a fixed confidence level for M_Z ,

e) $\varphi(Z)$ is an estimate of probability $Pr\{M_Z \geqslant f_k^*\}$,where f_k^* is the maximal of already determined values of f.

A lot of statistical procedures for construction of the above prospectiveness criteria are considered in Zhigljavsky (1985). The preference order is as follows: e), d), a), b) and c). Thus, criteria e) and d) should be used if possible. In the involved situations (for instance, in the case of a random noise presence under evaluation of f) they cannot be always constructed and one has to content oneself with b)-type criteria that can be constructed in general situations.

The estimates listed above can be constructed either via the results of evaluating the objective function or after the investigation of estimates or approximation of this function. As a rule, the estimates are probabilistic; deterministic estimates can be constructed mainly for a Lipschitz-type functional classes and correspond to the branch and bound approach.

3. Principal Construction of Branch and Probability Bound Methods. The methods under consideration are distinguished by (i) organization of the set branching, (ii) kinds of the prospectiveness criteria and (iii) rules for rejection of unpromising subsets.

Set branching depends on a structure of X and on a researcher's software and computer resources. If X is a hyperrectangle,it is natural to choose the same form for branching sets Z_{kj} ,where k is an iteration number and j is a set number. In the general case, spheres, hyperrectangles and sometimes ellipsoids can be naturally chosen as Z_{kj}. Two conditions are

imposed on the choice of Z_{kj}: their union should cover the domain of search and the number of points where f is evaluated should be in each set sufficiently large for drawing statistical inferences. There is no need for Z_{kj} to be disjoint for any fixed k.

Branching (or decomposition) of the search domain can be carried out either a priori (i.e., independently of the values of f) or a posteriori. Numerical investigation shows that the second technique provides more economical algorithms. The following branching technique has been proved to be convenient and efficient. At each k–th iteration first isolate in the search domain X_k a subdomain Z_{k1} with the centre at the point,corresponding to the record value of f. The point corresponding to the record value of f over $X_k \setminus Z_{k1}$ is the centre of a subdomain Z_{k2}. Similar subdomains Z_{kj} $(j = 1, \ldots, J)$ are isolated until either X is not covered or the hypothesis is not rejected that the global maximum can occur in the residual set $X_k \setminus \bigcup_{j=1}^{J} Z_{kj}$ (the hypothesis can be verified by the procedure described below). The search domain X_{k+1} of the next $(k+1)$–th iteration is naturally to be either $Z^{(k)} = \bigcup_{j=1}^{J} Z_{kj}$ or a hyperrectangle covering $Z^{(k)}$,or a union of disjoint hyperrectangles covering $Z^{(k)}$. In multidimensional cases the last two ways induce more convenient variants for realization of the branch and probability bound methods than the first one.

Contrary to the standart branch and bound methods, in the methods in hand the prospectiveness criterion may assume any real value (for some criteria, the interval $[0, 1]$ can be the set of values) rather than two values (say, 0 and 1). An estimate of M_Z can be used as the prospectiveness criterion. It is more natural to take as the prospectiveness criterion of Z the upper confidence bound for M_Z of a fixed level $1 - \gamma$ (see Section 4). Finally, the prospectiveness criterion discussed at

the beginning of Section 5, which relies upon the procedure for testing statistical hypothesis about M_Z, is natural and easily computable.

Rules for rejecting unpromising sets also can differ. Under reasonable organization of branching and use of the mathematical statistics apparatus for construction of the prospectiveness criterion, there is usually no absolutely unpromising set Z_0 (i.e., such whose prospectiveness criterion is $\varphi(Z_0) = \inf_{Z \in B} \varphi(Z)$). Narrowing of the search domain (i.e., rejection of subsets), therefore, may occur only if the lower prospectiveness bound δ is defined; if $\varphi(Z) \leqslant \delta$, then the set Z may be regarded as unpromising and be rejected.

Further, it is intuitively evident that the more promising a subset, the greater number of points should be locted in it. This can be allowed for by taking, for instance, the number of points in a subset to be directly proportional to a value of the prospectiveness criterion. The extremal case where all points are always located in the most promising set is not efficient and reliable. The following can be also taken into consideration when constructing the rejection rule. A reasonable prospectiveness criterion depends not only on function values over a given subset, but also on those in the whole set X. A subset of a mean prospectiveness can, therefore, become the unpromising one (and be rejected) owing to the improved prospectiveness of other subsets. The following organization of the search algorithm is, thus, possible: if a subset Z was recognized at the k–th iteration as being of a mean prospectiveness (i.e. $\delta < \varphi(Z) \leqslant \delta_*$), one might not evaluate function f in Z at several subsequent iterations and wait for Z to become unpromising and be rejected.

4. Statistical Inferences about the Maximal Value of a Function. Let $\Xi = \{\xi_1, \ldots, \xi_n\}$ be an independent sample from a random vector ξ with a probability distribution

P given on a measurable set Z, N be large, $\eta_1 \leqslant \ldots \leqslant \eta_N$ be the order statistics,corresponding to the sample $Y = \{y_i = f(\xi_i),\ i = 1,\ldots,N\}$ of the objective function values at the points of Ξ,

$$F(t) = \int_{f(x)<t} P(dx) \tag{1}$$

be the cumulative distribution function ($c.d.f.$) of the random variable $y = f(\xi)$. Suppose that P is equivalent to the Lebesgue measure on Z, f is continuous and the function $V(v) = 1 - F(M - 1/v), v > 0$, regularly varies at infinity, i.e.,

$$\lim_{v\to\infty} V(tv)/V(v) = t^{-\alpha} \quad \text{for each } t > 0, \tag{2}$$

where

$$M = M_Z = \text{ess sup } y = \sup_{z\in Z} f(z)$$

and α, $0 < \alpha < \infty$ is some exponent called the tail index. We shall suppose α is known (for the condition on determining α see Haan (1981), Zhigljavsky (1985)). If it is not the case,then α can be estimated, see Hall (1982), Smith (1987), Zhigljavsky (1985).

Statistical inferences about M constructed through the sample Y are the background for the branch and probability bound methods. Their thorough description and study is considered in Zhigljavsky (1985) with the references contained there. We mention only three representatives.

The linear estimators for M have the form

$$M_{N,r} = \sum_{i=0}^{r} a_i \eta_{N-i} \ ,$$

where r is much smaller than N, $a_0,\ldots,a_r$ are some coefficients satisfying the condition $\sum\limits_{i=0}^{r} a_i = 1$. The optimal coefficients are expressed as follows:

$$(a_i)^* = c\nu_i \Gamma(i+1)/\Gamma(i+1+2/\alpha), \text{ where}$$

$$\nu_0 = \alpha + 1, \nu_r = -(\alpha r + 1),\ \nu_j = \alpha - 1 \text{ for } j = 1, \ldots, r-1,$$

$$c = \begin{cases} (\alpha - 2)/[\alpha\Gamma(r+2)/\Gamma(r+1+2/\alpha) - \\ \qquad -2/\Gamma(1+2/\alpha)] & \text{for } \alpha \neq 2, \\ 1/\sum_{i=0}^{r} 1/(i+1) & \text{for } \alpha = 2. \end{cases}$$

For $N \to \infty$ the corresponding estimators $M^*_{N,r}$ are asymptotically optimal in the class of linear estimates and asymptotically efficient (under $r \to \infty$, $r/N \to 0$ for $N \to \infty$).

A convenient one-sided confidence interval for M of a fixed asymptotical (for $N \to \infty$) confidence level $1-\gamma$ has the form

$$[\eta_N, q_{r,1-\gamma}(\eta_N - \eta_{N-r})] \ , \tag{3}$$

where

$$q_{r,\delta} = [(1-(1-\delta)^{1/r})^{-1/\alpha} - 1]^{-1}.$$

The corresponding test procedure for the hypothesis $H_o : M \geqslant M_*$,where M_* is a fixed number, $M_* \geqslant \eta_N$ is determined by the rejection region

$$\{Y : (M_* - \eta_N)/(\eta_N - \eta_{N-r}) \geqslant q_{r,1-\gamma}\} \tag{4}$$

5. Typical Variants of the Branch and Probability Bound Method. Consider one of the most natural and readily computable prospectiveness criterion.

Let f^*_k be the largest value of f obtained so far and the search domain X_k be decomposed at the k–th iteration into subsets $Z_{kj} (j = 1, \ldots, J_k) : X_k \subset \bigcup_j Z_{kj}$. Define the prospectiveness criterion of Z_{kj} as

$$\varphi_k(Z_{kj}) = p_{kj} = [1 - ((f^*_k - \eta_N)/(f^*_k - \eta_{N-r}))^{\alpha}]^r \ , \tag{5}$$

where η_N and η_{N-r} are the order statistics, corresponding to a sample $\{y_l = f_{(\xi_l)}, l = 1, \ldots, N\}$, r is much smaller than

$N, \xi_1, \ldots, \xi_N$ are independent realizations of a random vector on X got into Z_{kj}. The values p_{kj} can be treated in two ways: asymptotically (for r=const, $N \to \infty$) it is greater than or equal to the probability that

$$M_{Z_{kj}} = \sup_{x \in Z_{kj}} f(x) \geqslant f_k^*$$

and the probability of accepting the hypothesis $H_0 : M_{Z_{kj}} \geqslant f_k^*$,provided that the hypothesis is true and that the hypothesis testing procedure is determined via (4). To derive (5) from (4),it suffices to solve the inequality $(f_k^* - \eta_N)/(\eta_N - \eta_{N-r}) \geqslant q_{r,1-\gamma}$ with respect to γ and to substitute the corresponding equality for it.

In the algorithm below the number of points at each k-th iteration in promising subset Z_{kj} is assumed to be (in the probabilistic sense) proportional to the value of the prospectiveness criterion $\varphi_k(Z_{kj})$; a further branching is performed over those sets Z_{kj} ,whose values (5) are not less than given numbers δ_k.

Algorithm 1.

1.Put $k = 1$, X_1, $f_0^* = -\infty$. Choose a distribution P_1.

2. Generate N_k times the distribution P_k, obtain a sample

$$\Xi_k = \{x_1^{(k)} \ldots, x_{N_k}^{(k)}\}.$$

3. Evaluate f at the points of Ξ_k and put

$$f_k^* = \max\{f_{k-1}^*, f(x_1^{(k)}), \ldots, f(x_{N_k}^{(k)})\}$$

4. Organize a branching of X_k by representing this set as $X_k \subset \bigcup_j Z_{kj}$,where Z_{kj} are measurable subsets of X ,having a sufficient number of points from Ξ_k for the statistical inferences drawing.

5. For each subset Z_{kj} compute (not necessarily through (5)) values of a prospectiveness criterion $\varphi_k(Z_{kj})$.

6. Put $X_{k+1} = \bigcup_j Z^*_{kj}$,where

$$Z^*_{kj} = \begin{cases} Z_{kj} & \text{if } \varphi_k(Z_{kj}) > \delta_k, \\ \varnothing & \text{otherwise,} \end{cases}$$

i.e., those subsets $Z_{kj} (j = 1, 2, \ldots)$ for which $\varphi(Z_{kj}) \leqslant \delta_k$ are rejected from the search domain X_k. Let I_k be a number of subsets Z^*_{kj}.

7. Put

$$P_{k+1}(dx) = \sum_{j=1}^{I_k} P_j^{(k)} \theta_{kj}(dx) \ , \tag{6}$$

where $P_j^{(k)} = 1/I_k$ and θ_{kj} are the uniform distributions over the sets Z^*_{kj} (j=1, ..., I_k). If φ_k is a nonnegative criterion,we may take

$$P_j^{(k)} = \varphi_k(Z^*_{kj}) / \sum_{l=1}^{I_k} \varphi_k(Z^*_{kl}). \tag{7}.$$

8. Return to Step 2 substituting $k + 1$ for k.

Closeness of f^*_k to an estimate of $M = M_X$ or to the upper bound of the confidence interval (3) are natural stopping rules for Algorithm1. Another type of the stopping rule is the reaching of rather a small volume of the search domain X_k.

Of course, after termination of the algorithm one may use a local ascent routine to determine a location of a global maximizer more precisely.

Distributions P_{k+1} in Algorithm 1 are sampled by means of the superposition method: discrete distribution concentrated at the points $\{1, 2, \ldots, I_k\}$ with probabilities $P_j^{(k)}$ is sampled first and is followed by the distribution $\theta_{k\tau}$ sampling, where τ is the realization obtained at sampling of the discrete

distribution. If the first procedure for choosing probabilities $P_j^{(k)}$ (i.e. $P_j^{(k)} = 1/I_k$) is used at Step 7 of Algorithm 1 and Z_{kj} $(j = 1, 2, \ldots)$ are disjoint,then (6) is nothing but a uniform distribution on the set X_{k+1}.

All the points obtained at the previous interations and occuring at the set X_k can be included into the collection Ξ_k at Step 2 of Algorithm 1, which improves the accuracy of statistical procedures for determination of the propectiveness criterion. If all distributions P_k are uniform on X_k ,the resulting samples are samples from the uniform on X_k distributions as well; if the form of P_{k+1} resembles (6), distributions of the resulting samples are not uniform on X_k but this fact is of no importance for drawing statistical inferences (see Zhigljavsky (1985)). It is not of course necessary to store all previous points and objective function values because statistical inferences are drawn only from the points, where f is relatively large; one even does not need to know the number of points within the domain.

After completion of Algorithm 1,one can apply (5) to the union of all rejected subsets in order to determine the probability of "not loosing" the global maximum. The following should be taken into consideration. Let $\gamma_{kj} = 1 - p_{kj}$ be the probabilities of loosing the global maximum in the rejected sets Z_{kj} as computed through (5). The total probability of loosing the global maximizer as determined through(5) is, then, $\max \gamma_{kj}$ at most. Indeed, let us take the set $Z_{kj}^* = Z_{kj}$ containing the point, corresponding to the maximal order statistic η_N of the set $Z = \bigcup\limits_{k,j} Z_{kj}$. Then η_{N-r} for Z is not less than η_{N-r} for Z_{kj}^* and $\max_k f_k^*$ (it plays the role of f_k^* for Z) is not less than f_k^*. But p_{kj} defined via (5) is an increasing function of both η_{N-r} and f_k^* for fixed α, r, r and η_N.

If evaluations of f are costly, one has to be extremely cautious in planning computations. After completion of certain number of evaluations of f, one should extend the amount of auxiliary computations with the aim of extracting and exploiting as much information about the objective function as possible. To extract this information from a large union of sets Z_{kj},one should: compute probabilities (5); construct various estimates of $M_{Z_{kj}}$; check the hypothesis about the value of the tail index α which can provide information on whether a global maximizer vicinity is reached; estimate c.d.f. (1) for the values of t close to M (this will enable one to draw a conclusion about the expediency and approximate size of the remaining computations); and, in addition, one can estimate f and related functions in order to recheck and update the information. A decision about prospectiveness of subsets should be made through a large set of statistical procedures. It is natural that these procedures can be taken into consideration only if the algorithms are realized in the interactive fashion.

The major part of our assertions have an exact sense only if for the corresponding function $F(t)$ the condition (2) is met and the tail index α is known. As for the condition (2),the practice shows that it may be always regarded as met, if the problem at hand is not too exotic. In principle, a statistical inference about α can be always made via the procedures of Smith (1987) that are to be carried out successively as the points are accumulated. However, it is recommendable to use the results of Haan (1981) and Section 4.2 of Zhigljavsky (1985), if possible, since the accuracy of procedures for statistical inferences about α is high only under comparatively large samples. According to these results,for the case when $X \subset R^n$ and the objective function f is twice continuously differentiable and approximated by a non-degenerate quadratic form in a vicinity of a global maximizer x^*,one can always

set $\alpha = n/2$. While doing so, one may be confident that statistical inferences are asymptotically true for the subsets $Z \subset X$ containing x^* (together with subsets Z containing maximizers $x_Z^* = \arg\max_{x \in \overline{Z}} f(x)$ as interior points). The prospectiveness of other subsets Z may be lower than for the case of the usage of their true values of α but it is not the deficiency for the methods.

Let us finally describe the variant of the branch and probability bound method that is convenient for realization, uses most of the above recommendations and proved to be efficient for a wide range of practical problems.

Algorithm 2.

1. Set $k = 1$, $X_1 = X$, $f_0^* = -\infty$.

2. Generate a given number N times the uniform distribution on the search domain X_k, obtain $\Xi_k = \{x_1^{(k)}, \ldots, x_N^{(k)}\}$.

3. Evaluate $f(x_j^{(k)})$ for $j = 1, \ldots, N$. Put

$$f_k^* = \max\{f_{k-1}^*, f(x_1^{(k)}), \ldots, f(x_N^{(k)}\}.$$

Check the stopping criterion (closeness of f_k^* and the estimator $M_{N,r}^*$ for M).

4. Put $Y_{k,0} = X_k$, $j = 1$.

5. Put $Y_{k,j} = Y_{k,j-1} \backslash Z_{kj}$,where Z_{kj} is a cube (or a ball) of volume $\rho \mathrm{mes}(X_k)$ centered at the point having the maximal value of f among the points from the set $Y_{k,j-1}$ with evaluated objective function values.

6. If a number m of points in $Y_{k,j}$ with known values of the objective function is insufficient for drawing statistical inferences (i.e., $m \leqslant m_0$), then set $X_{k+1} = X_k$ and go to Step 9. If $m > m_0$,then put

$$\varphi_k(Y_{k,j}) = (1 - ((f_k^* - \eta_m)/(f_k^* - \eta_{m-r}))^\alpha)^r ,$$

where order statistics η_m, η_{m-r} correspond to the values of f at $Y_{k,j} \bigcap \bigcup_{l=1}^{k} \Xi_l$. If $\varphi_k(Y_{k,l}) \leqslant \delta$ then go to Step 8.

7. Substitute $j+1$ for j and go to Step 5.

8. Choose X_{k+1} as the union of disjoint hyperrectangles covering the set $Z_{k1} \bigcup \ldots \bigcup Z_{kj}$.

9. Substitute $k+1$ for k and return to Step 2.

The author proposes to use $N = 100, m_0 = 15, \rho = 0.1$, and $r = \min\{5, [m/10]\}$, where [.] is the integer part operation as a standart collection of parameters of Algorithm 2. Another choice of the parameters,corresponding to the recommendation of Zhigljavsky (1985), is possible as well.

Under the condition $m_0 \to \infty$, Algorithm 2 converges with probability $1-\gamma$, i.e., it looses a global maximizer with a probability not larger than $1-\gamma$. It follows from the results of Section 4.2 of Zhigljavsky (1985).

6. Numerical Results. We present here some numerical results for Algorithm 2 and compare them with results of some other authors basing on Žilinskas (1986). We use Hartman and Shekel test functions with the coefficients given in Žilinskas (1986) pp. 146, 147 as the test ones . The numerical results for the P^*-algorithm of Žilinskas , some variant of Algorithm 2 (branch and probability bound (BPB) algorithm) added both with local stages to obtain more precise approximations for x^* (we use Nelder-Meed local algorithm) and the best results in Dixon and Szegö (1978) are given in Tables 1 and 2 for Hartman and Shekel test functions,correspondingly. In most cases our algorithm has won. Of course, it is not meant that in any case it is better than the above mentioned P^*-algorithm or some other. One of the reasons for a good feature of our algorithm is a suitable choice of the algorithm parameters.

Table 1. Test results for Shekel test function

	Global stage only		Local stage		Total number of evaluation		
	P^*	BPB	P^*	BPB	P^*	BPB	Best in DS
4 terms	2.289	4.726	10.145	10.148			
	4.531	4.042	4.007	4.003			
	4.004	3.705	4.000	4.001			
	4.276	3.833	4.005	3.997			
	3.893	4.04	3.998	3.994			
Evaluation numbers	386	57	53	262	439	319	620
6 terms	2.716	3.002	10.380	10.398			
	4.510	3.545	3.995	3.995			
	4.092	4.264	4.013	4.003			
	3.726	4.035	3.988	4.001			
	4.089	3.946	3.989	3.996			
Evaluation numbers	356	150	53	92	409	242	788
9 terms	2.004	4.468	10.469	10.531			
	3.557	3.842	4.018	4.004			
	4.605	4.144	3.982	4.000			
	4.020	3.669	4.002	4.008			
	3.895	3.999	4.005	3.999			
Evaluation numbers	400	300	49	119	449	419	1160

Table 2. Test results for Hartman test function

	Global stage only		Local stage		Total number of evaluation		
	P^*	BPB	P^*	BPB	P^*	BPB	Best in DS
	3.778	3.856	3.763	3.864			
Dimension 3	0.239	0.237	0.114	0.153			
	0.524	0.557	0.556	0.554			
	0.830	0.857	0.852	0.853			
Evaluation numbers	197	200	50	52	247	252	513
	2.093	3.630	3.322	3.946			
	0.009	0.638	0.201	0.656			
	0.123	0.365	0.150	0.418			
Dimension 6	0.687	0.507	0.477	0.439			
	0.177	0.879	0.275	0.921			
	0.375	0.602	0.311	0.417			
	0.595	0.140	0.657	0.102			
Evaluation numbers	276	252	122	120	398	372	515

REFERENCES

Dixon, L.C.W., and G.P.Szegö (1978). The global optimization problem: an introduction. In L.C.W. Dixon and G.P. Szegö (Ed.), *Towards Global Optimization 2.* North Holland, Amsterdam. pp. 1–15.

Haan, L. (1981). Estimation of the minimum of a function using order statistics. *J. Amer. Statist. Assoc.*, **76**(374), 467–475.

Hall, P. (1982). On estimating the endpoint of a distribution. *Ann. Statist.*, **10**(2), 556–568.

Horst, P. (1986). A general class of branch-and-bound methods in global optimization with some new approaches for concave minimization. *J. Optimiz. Theory and Applic.*, **51**(2), 271–291.

Smith, R.L. (1987). Estimating tails of probability distributions. *Ann. Statist.*, **15**(3), 1174–1207.

Zhigljavsky, A.A.(1985). *Mathematical Theory of the Global Random Search.* Leningrad University Press, Leningrad (in Russian, to be published in English by Kluwer Academic Publishers).

Žilinskas, A.G. (1986). *Global Optimization: Axiomatics of Statistical Models, Algorithms, Applications.* Mokslas, Vilnius (in Russian).

Received May 1989

A. Zhigljavsky received the Degrees of Candidate and Doctor of Physical and Mathematical Sciences at Leningrad University in 1981 and 1987, respectively. He is a professor at the Mathematical Department of the Leningrad University. His research interests include global optimization, experimental design, statistical modelling, sequential statistical analysis.