Interval Computations
No 4(6), 1992

USING INTERVAL METHODS IN CLUSTER ANALYSIS AND VERIFIED REPRESENTATION OF CONNECTED SETS

Pavel S.Pankov and Batyigul D.Bayachorova

The well-known problem of effective (outer) representation of range S of mutually dependent variables is considered. By modifying the known algorithm of generalized bisection with constrained depth of subdividing the algorithm is described that finds collection of (narrow) interval vectors (i.e. clusters) such that their union contains S and each connected subset of S is contained in some cluster.

ПРИМЕНЕНИЕ ИНТЕРВАЛЬНЫХ МЕТОДОВ В КЛАСТЕРНОМ АНАЛИЗЕ И ДОКАЗАТЕЛЬНОЕ ПРЕДСТАВЛЕНИЕ СВЯЗНЫХ МНОЖЕСТВ

П.С.Панков, Б.Д.Баячорова

Рассматривается известная проблема эффективного внешнего представления области S изменения переменных, связанных взаимной зависимостью. Путем модификации известного алгоритма обобщенной бисекции с ограничением на глубину дробления построен алгоритм, который находит набор (узких) интервальных векторов (т.е. кластеров), объединение которых содержит S, и каждая связная компонента S содержится в одном из кластеров.

This paper is an extended translation of [1].

The search of more effective representations of a range of values of mutually dependent variables (if a priori range for each of them is given) is the important problem of interval mathematics.

We need, as usually, the following initial data: the interval vector (box) $X \subseteq R^n$ whose components are a priori ranges for variables and the interval indicator function $I(X')$ defining the set $S \subseteq X$. We shall use the following three-valued logical indexation [3]:

> If it is proved that all points of a box X' belong to S ($X' \subseteq S$) then $I(X') := -1$;
> If it is proved that $X' \cap S = \emptyset$ then $I(X') := 1$;
> Otherwise $I(X') := 0$.

(If such indicator function is calculated by means of common interval operations then it is monotomic: if $X' \subseteq X''$ and $I(X'') \neq 0$ then $I(X') = I(X'')$).

The use of any generalized bisection algorithm, as a rule, yields a list of too many boxes whose union contains S.

Remark 1. Here we mean the union itself (not the "representation by the outer interval").

The problem is to obtain any (sufficitently narrow) outer representation of S as the union of some (large) boxes and we see that such formulation may be considered within the cluster analysis (see for instance [4]) and the interval (verified) approach yields strict formulation of problems which may be also effective.

(The cluster approach for the close problem of optimization was considered in [2]).

We shall require that any connected subset of S belongs to any only cluster (one may see that the reverse demand is not feasible).

We describe the algorithm of searching such cluster representation being the modification of the algorithm with the constrained depth of bisection [3], which does not demand vast memory.

Denote:

$X_1 \ldots X_k$ are sub-boxes under consideration; n_j is the index of subdivision depth (inverse of volume) of X_j; $W_1 \ldots W_M$ are candidates for cluster-boxes.

Algorithm. *Let the box X, the indicator function I and the constraint N of subdivision depth be given.*

Step 1. Let $M := 0,\ k := 1,\ X_1 := X,\ n_1 := 1$.
Step 2. Calculate $J := I(X_k)$. *If* $J = 1$ *then go to step 8.*
Step 3. If $J = -1$ *or* ($J = 0$ *and* $n_k = N$) *then go to step 4 else* ($J = 0$ *and* $n_k < N$) *bisect* X_k *into two sub-boxes, denote them as* X_k *and* X_{k+1}, *let* $n_{k+1} := n_k := n_k + 1,\ k := k + 1$ *and go to step 2.*
Step 4. If $M = 0$ *then let* $W_1 := X_k$ *and go to step 8.*
Step 5. Find the set (list) $T := \{j \in [1 \dots M] \mid X_k \cap W_j \neq \emptyset\}$.
Step 6. If $T = \emptyset$ *then let* $W_{M+1} := X_k$ *else diminish* M *by the number of elements of* T *and calculate the "representation by the outer interval"* $W_{M+1} := \overline{\cup}\{W_j \mid j \in T\}\overline{\cup}X_k$.
Step 7. Let $M := M + 1$.
Step 8. Let $k := k - 1$. *If* $k > 0$ *then go to step 2 else STOP.*

The list $W_1 \dots W_M$ is the result of execultion of this algorithm.

Remark 2. As in [3], we may apply this algorithm with the increased depth of subdivision not to the whole box X but to the obtained set $\{W_j \mid j = 1 \dots M\}$.

Remark 3. We may choose the boxes to be united in step 6 by another criterion, for instance, "if the distance between X_k and W_j is sufficiently small."

We note that the outer representation of S by the most narrow single box (i.e. "by one cluster") was realized in [3].

The verified searching of connected subsets of a plane set (branches of curve) $S := \{(x_1, x_2) \mid F(x_1, x_2) = 0\}$ by two-sided bounding by broken lines was supposed in [5].

Describe another approach to searching an effective representation of mathematical objects. To optimize the running of the algorithm [3] it is desirable to choose arguments of the indicator function I to be independent variables. But in such case the outer representation of all sub-boxes with $I \leq 0$ does not take into account the actual relations between the arguments and may be too wide.

Because of it we suggested [6] to calculate simultaneously the outer representation within the original space.

Such approach was successfully applied [7] to search the solution of the integral equation if the estimation of its derivative is given: Suppose that there exists any function y fulfilling

$$\int_a^b K(t,s)y(s)ds = f(t), \ | y'(t) | \leq c.$$

To find any (narrow) interval set of functions containing y let us choose a natural m and denote $h := (b-a)/(2m)$, $t_j := a + hj$, $j = 0 \dots 2m$.

The "natural" variables are $y_j := y(t_j)$, $j = 0 \dots 2m$, the "independent" ones for the algorithm of global search were y_m and $z_j := y_{j+1} - y_j$, $j = 0, \dots 2m-1$ ($z \in [-ch, \ ch]$).

The set of cluster-boxes may be considered as a generalization of the concept of multi-interval [6] to the multi-dimensional case.

References

1. Pankov, P. S. and Bayachorova, B. D. *Using interval methods in cluster analysis and verified representation of connected sets.* In: "International conference on interval and stochastic methods in science and engineering, Moscow, September 22–26, 1992, vol. 1", Moscow, 1992, pp. 134–136 (in Russian).

2. Kearfott, R. B. *Problems of cluster in global optimization.* In: "International conference on interval and stochastic methods in science and engineering. Moscow, September 22–26, 1992, vol 2", Moscow, 1992, pp. 42–44 (in Russian).

3. Pankov, P. S., Kuznetzov, V. A. and Kenenbaeva, G. M. *Global search algorithm and software package to approximate kinetic experimental data with known error of measurement.* Preprint, Institute of Chemical Physics of the USSR Academy of Sciences, Chernogolovka (1990) (in Russian).

4. *Classification and Clusters.* Proc. of Sci. Seminar, Madison, 1976.

5. Kenenbaeva, G. M. *The verified approximation of curves and boundaries of two-dimensional domains.* Inf.-operat. material (interval analysis), preprint 16, Computer Center, Siberian Branch of the USSR Academy of Sciences, Krasnoyarsk (1990), pp. 15–18 (in Russian).

6. Bayachorova, B. D. *The improved algorithm of global search and its application to solve the integral equation of the first kind in comcompact space.* To be published (in Russian).

7. Bayachorova, B. D. *Algorithm of verified search in compact functional space.* In: "Proceedings of all-union conference. Asymptotical methods of the theory of

singularly perturbed equations and improperly posed problems", Bishkek, 1991, pp. 23 (in Russian).

8. Yakovlev, A. G. *Computer arithmetic of multi-intervals.* In: "Problems of cybernetics", **125** (1987) (in Russian).

Interval Computatio
No 4(6), 1992

**SOLV
NON
OF SOM
IN**

An interval
developed and
account the sp
convergence th

**РЕШЕНИЕ
СПЕЦ
МОДИ
ИТЕРА**

Развит и и
нивания обр
вальные мет
мой частной
предложенны

Solving the sy

where $f : \mathbb{R}^n \to$
of equations obta