Studies in Computational Intelligence 539

Martine Ceberio
Vladik Kreinovich *Editors*

Constraint Programming and Decision Making

Springer

Editors
Martine Ceberio
Department of Computer Science
University of Texas at El Paso
El Paso Texas
USA

Vladik Kreinovich
Department of Computer Science
University of Texas at El Paso
El Paso Texas
USA

ISSN 1860-949X ISSN 1860-9503 (electronic)
ISBN 978-3-319-04279-4 ISBN 978-3-319-04280-0 (eBook)
DOI 10.1007/978-3-319-04280-0
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013957363

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Graph Subdivision Methods in Interval Global Optimization

Sergey P. Shary

Institute of Computational Technologies,
Novosibirsk, Russia

Abstract. The work advances a new class of global optimization methods, called *graph subdivision methods*, that are based on simultaneous adaptive subdivision of both the function's domain of definition and the range of values. An application to interval linear systems is given.

Keywords: global optimization, interval analysis, adaptive subdivision.

1 Introduction

The subject matter of our paper is the problem of global optimization of a real-valued function $f : \mathbb{R}^n \supseteq \boldsymbol{X} \to \mathbb{R}$ over an axis-aligned rectangular box $\boldsymbol{X}$ (i. e. over an interval vector):

$$\text{find} \quad \min_{x \in \boldsymbol{X}} f(x). \tag{1}$$

The problem (1) is known to be (more or less) successfully solved by various interval techniques [1, 3, 6], which enables one to reliably compute two-sided bounds for both the optimum value and the argument it is attained at. The basis of these methods is adaptive, according to the "branch-and-bound" strategy, subdivision of the domain of the function to be minimized combined with the interval evaluation of the ranges over the resulting subdomains.

The purpose of our work is to present a new promising interval approach for the solution of the problem (1) that relies upon joint adaptive subdivision of both the function's domain of definition and its range of values. For some classes of problems, the new approach is expected to turn out better than the traditional techniques from [1, 3, 6] in either implementation flexibility or computational efficacy and the quality of the results it produces. A shortened version of this article has been previously published as [9].

2 Idea of the New Approach

Notice that any function $f : \mathbb{R}^n \supseteq \boldsymbol{X} \to \mathbb{R}$, being by the very definition a special subset of the direct product $\mathbb{R}^n \times \mathbb{R}$, is an $(n+1)$-dimensional object. In connection with it, we usually use the concept of the *graph* of the function f:

$$\text{graph of } f = \left\{ (x,t) \in \mathbb{R}^{n+1} \mid x \in \mathbb{R}^n, t \in \mathbb{R}, f(x) = t \right\}.$$

M. Ceberio and V. Kreinovich (eds.), *Constraint Programming and Decision Making*, Studies in Computational Intelligence 539,
DOI: 10.1007/978-3-319-04280-0_18,

However, the interval global optimization methods that we mentioned in Introduction involve into active operation — adaptive subdivision — only the first n coordinates of this set. The last $(n+1)$-th coordinate of the function represented by its graph is processed in a substantially different manner, passively, and the same is true for the overwhelming majority of the classical optimization techniques. How could we correct the situation and what would be the result?

We start our consideration from the simplest case of a single-variable function $f : \mathbb{R} \supseteq \boldsymbol{X} \to \mathbb{R}$, defined on a closed interval $\boldsymbol{X}$, for which we have to solve the problem (1). In the plane $0xy$, let us construct a straight line parallel to the first axis, with the equation $y = l$, where l is a constant. We can ascertain whether the line intersects the graph of the function $y = f(x)$ after having solved the equation

$$f(x) - l = 0 \tag{2}$$

on $\boldsymbol{X}$ or, alternatively, making sure that it is incompatible (unsolvable). As is easily seen, the answer to the above question provides us with information on the minimum (1) under computation: if the straight line $y = l$ intersects the graph of the function $y = f(x)$, then

$$\min_{x \in \boldsymbol{X}} f(x) \le l.$$

Moreover, if $f(x)$ is continuous on $\boldsymbol{X}$, then

$$\min_{x \in \boldsymbol{X}} f(x) \;=\; \min\bigl\{\, l \in \mathbb{R} \mid \text{ the equation } f(x) - l = 0 \text{ is solvable} \,\bigr\}.$$

Therefore, varying the value of the "level" l and repeating the process of the solution of the equation (2), we can substantially improve the estimate for the sought-for minimum (1).

The procedure we have just described can be substantially modified by using the ideas and methods of the interval analysis:

First, the interval methods make it possible to easily compute estimates for the range of f over $\boldsymbol{X}$ from below and from above, which is necessary to determine the bounds of variation of the level l in the process of the correction of the minimum.

Second, it makes sense to examine the intersection of the graph of the function $y = f(x)$ not with single lines, but with the whole bundles of lines parallel to the $0x$ axis and defined by the equations $y = \boldsymbol{l}$, where $\boldsymbol{l}$ is an interval in $\mathbb{R}$. We will be able thereby to estimate the global minimum (1) both from below and from above, since $\min_{x \in \boldsymbol{X}} f(x)$ is not less than the minimum of the left endpoints and not greater than the minimum of the right endpoints of all the intervals $\boldsymbol{l}$ such that the bundle $y = \boldsymbol{l}$ intersects the graph of the function $y = f(x)$.

Third, the interval methods for the solution of equations (e.g. the interval Newton method and its modifications [1, 3, 5]) enable us, under very mild requirement on the smoothness of f, to examine solvability of both the point equation (2) and the interval equation $f(x) - \boldsymbol{l} = 0$. The latter is understood as the existence of some $l \in \boldsymbol{l}$ for which (2) is solvable.

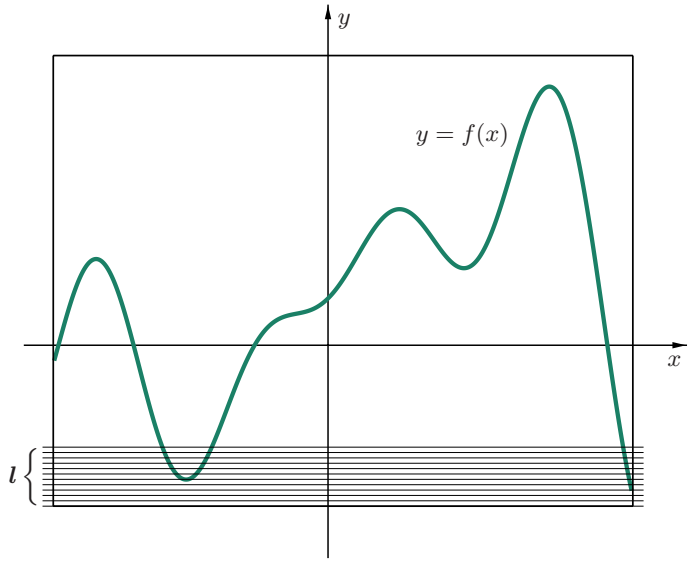

Fig. 1. Does the bundle of lines intersect the graph of the function?...

The answer produced by the interval methods may have one of the following forms [1, 3, 5]:

1. The equation does not have solutions — *unsolvable* — within the interval under consideration, i. e. $0 \notin f(x) - \boldsymbol{l}$ for any $x \in \boldsymbol{X}$.
2. The equation has, with guarantee, a solution (or solutions) within the interval under consideration, i. e. there certainly exists $x^\star \in \boldsymbol{X}$ such that $f(x^\star) - l = 0$ for at least one $l \in \boldsymbol{l}$. We shall speak that the equation is just *solvable* then.
3. Applying the solution procedure does not allow us to speak, with certainty, that the equation is either has solutions or unsolvable on the given interval $\boldsymbol{X}$. In such cases, we shall speak that the equation is *possibly solvable*.

The third option is the most unfavorable algorithmically, but we should carefully take it into account in our reasoning since this kind of uncertainty is quite actual in computation often being the case when the equation (2) has multiple roots. Notice also that the interval methods never lose roots and cannot at all output the message "no solutions" if the equation really has them.

Finally, we will use the subdivision of the interval of the range of values instead of the "varying the level" l. The overall interval version of the procedure for finding the global minimum of the single-variable function $f(x)$ over the interval $\boldsymbol{X}$ can look as follows. It starts with computing a crude interval enclosure $\boldsymbol{Y}$ of the range of values $f(x)$ over $\boldsymbol{X}$ (for example, as the natural interval extension of f on $\boldsymbol{X}$). Further,

we bisect the interval $\boldsymbol{Y}$ to beget the subintervals $\boldsymbol{Y}' := [\underline{\boldsymbol{Y}}', \operatorname{mid} \boldsymbol{Y}]$ and $\boldsymbol{Y}'' := [\operatorname{mid} \boldsymbol{Y}, \overline{\boldsymbol{Y}}]$, where $\operatorname{mid} \boldsymbol{Y} = \frac{1}{2}(\overline{\boldsymbol{Y}} + \underline{\boldsymbol{Y}})$ is the midpoint of $\boldsymbol{Y}$;

we check the solvability of the interval equations $f(x) - \boldsymbol{Y}' = 0$ and $f(x) - \boldsymbol{Y}'' = 0$:

- if the equations is unsolvable, then we discard the respective interval, either $\boldsymbol{Y}'$ or $\boldsymbol{Y}''$, and never consider it;
- solvability or possible solvability of the equation implies that either lower or upper estimate of the global minimum can be corrected according to the prescription formulated in the item "Second" of the list at page 154.

The above procedure correcting the estimate of the minimum (1) may be repeated with respect to its descendants $\boldsymbol{Y}'$ and $\boldsymbol{Y}''$, after which the bisection-correction ought to be carried out again and so on unless the computed lower and upper bounds of the minimum are not sufficiently close to each other. Notice that, to maintain guarantee of our computation, in such a process we have to keep all the subintervals $\boldsymbol{y}$ of the initial interval $\boldsymbol{Y}$ for which the corresponding equations $f(x) - \boldsymbol{y} = 0$ are possibly solvable, since they may correspond to the bundles having nonempty intersection with the graph.

3 Multidimensional case

Theoretically, the computational scheme of the one-dimensional global optimization algorithm we have developed in the preceding section is completely applicable to the functions $f(x) := f(x_1, x_2, \ldots, x_n)$ of several variables. The only thing we should be able to do for that is to check intersection of the graph of the function $y = f(x)$ with the bundle of the hyperplanes $y = \boldsymbol{l}$ that are orthogonal to the $0y$ axis. Sometimes, that can be really done when we have a powerful equations solver and are able to apply it easily. In particular, Semenov [7] implemented a similar kind of procedure to refine the value of the optimum in some problems.

However, in most cases the practical implementation of our idea encounters big difficulties. The point is that, in the general multidimensional case, the solution of an equation — inquiring into its solvability — is in no ways easier problem than the global optimization. As opposed to the single-variable situation, we do not have simple and efficient techniques such as the interval Newton method and its modifications at our disposal. A way out of the difficulty may be subdivision of the domain of definition of f — the box $\boldsymbol{X}$ — along some (but not all!) selected coordinate directions, whose number and specific choice depend on the problem under solution and its objective function.

The coordinate directions along which the function's domain shall not be subdivided will be referred to as *mute*, and first we consider the simplest methods having only one mute direction with the number $\mu \in \{1, 2, \ldots, n\}$. Let, in the space $\mathbb{R}^{n+1}$, a line be given, parallel to the μ-th coordinate axis and having the

Table 1. The simplest graph subdivision method for global optimization (one mute variable)

Input

A box $\boldsymbol{X} \subseteq \mathbb{R}^n$ and a function $f : \boldsymbol{X} \to \mathbb{R}$. An accuracy $\epsilon > 0$.

A number μ of the mute component, $\mu \in \{1, 2, \ldots, n\}$.

A method for checking the solvability of the single-variable interval equation $\phi(\boldsymbol{Z}, t) = 0$ for ϕ and $\boldsymbol{Z}$ defined as (4)–(5).

Output

The lower $\underline{y}$ and upper $\overline{y}$ estimates, with the accuracy ϵ, for the global minimum of the function f over the box $\boldsymbol{X}$.

Algorithm

compute an enclosure $\boldsymbol{Y}$ of the range of f over $\boldsymbol{X}$;

assign $\boldsymbol{Z} := (\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{\mu-1}, \boldsymbol{X}_{\mu+1}, \ldots, \boldsymbol{X}_n, \boldsymbol{Y})$;

set $z := \underline{\boldsymbol{Y}}$ and $\overline{y} := \overline{\boldsymbol{Y}}$;

initialize the working list $\mathcal{L} := \{ (\boldsymbol{Z}, z) \}$;

`DO WHILE` $(\overline{y} - z \geq \epsilon)$

choose the component k of the box $\boldsymbol{Z}$ having the largest length, i. e. such that $\operatorname{wid} \boldsymbol{Z}_k = \max_{1 \leq i \leq n} \operatorname{wid} \boldsymbol{Z}_i$;

bisect the box $\boldsymbol{Z}$ along the k-th coordinate to get the boxes $\boldsymbol{Z}'$ and $\boldsymbol{Z}''$ such that $\boldsymbol{Z}' := (\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_{k-1}, [\, \underline{\boldsymbol{Z}}_k, \operatorname{mid} \boldsymbol{Z}_k \,], \boldsymbol{Z}_{k+1}, \ldots, \boldsymbol{Z}_n)$,
$\boldsymbol{Z}'' := (\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_{k-1}, [\, \operatorname{mid} \boldsymbol{Z}_k, \overline{\boldsymbol{Z}}_k \,], \boldsymbol{Z}_{k+1}, \ldots, \boldsymbol{Z}_n)$;

if the equation $\phi(\boldsymbol{Z}', t) = 0$ is solvable or possibly solvable on $\boldsymbol{X}_\mu$, then assign $z' := \underline{\boldsymbol{Z}}'_n$ and put the record $(\boldsymbol{Z}', z')$ into $\mathcal{L}$ so that the second fields of the records in $\mathcal{L}$ increase ;

if the equation $\phi(\boldsymbol{Z}', t) = 0$ is solvable on $\boldsymbol{X}_\mu$, set $\overline{y} := \min\{ \overline{y}, \overline{\boldsymbol{Z}}'_n \}$;

if the equation $\phi(\boldsymbol{Z}'', t) = 0$ is solvable or possibly solvable on $\boldsymbol{X}_\mu$, then assign $z'' := \underline{\boldsymbol{Z}}''_n$ and put the record $(\boldsymbol{Z}'', z'')$ into $\mathcal{L}$ so that the second fields of the records in $\mathcal{L}$ increase ;

if the equation $\phi(\boldsymbol{Z}'', t) = 0$ is solvable on $\boldsymbol{X}_\mu$, set $\overline{y} := \min\{ \overline{y}, \overline{\boldsymbol{Z}}''_n \}$;

delete the former leading record $(\boldsymbol{Z}, z)$ from the list $\mathcal{L}$;

denote the new leading record of the list $\mathcal{L}$ by $(\boldsymbol{Z}, z)$;

`END DO`

$\underline{y} := z$;

parametric equation

$$\begin{cases} x_1 = r_1, \\ \quad\vdots \\ x_{\mu-1} = r_{\mu-1}, \\ x_\mu = t, \\ x_{\mu+1} = r_{\mu+1}, \\ \quad\vdots \\ x_n = r_n, \\ y = l, \end{cases} \tag{3}$$

where t is a parameter varying over the entire of $\mathbb{R}$ and $r_1, \ldots, r_{\mu-1}, r_{\mu+1}, \ldots, r_n$, l are some constants. Similar to the one-dimensional case,

$$\min_{x\in\boldsymbol{X}} f(x) \;=\; \min \left\{ l \in \mathbb{R} \;\middle|\; \begin{array}{c} \text{the line (3) defined by (3) intersects} \\ \text{the graph of the function } y = f(x) \end{array} \right\}$$

provided that f is continuous. Therefore, we can "feel about" the graph of the function to be minimized by the one-dimensional lines, making use of the efficient one-dimensional interval procedures (the famous interval Newton method and modifications) to check whether the elementary "level equations" $f(x) - l = 0$ are solvable or not.

Turning to the interval optimization procedure, we designate

$$\boldsymbol{Z} = (\,\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n) \;:=\; (\,\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{\mu-1}, \boldsymbol{X}_{\mu+1}, \ldots, \boldsymbol{X}_n, \boldsymbol{Y}\,), \tag{4}$$

$$\phi(\boldsymbol{Z}, t) \;:=\; f(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{\mu-1}, t, \boldsymbol{X}_{\mu+1}, \ldots, \boldsymbol{X}_n) - \boldsymbol{Y}. \tag{5}$$

The n-dimensional boxes $\boldsymbol{Z}$ represent the bundles of straight lines parallel to the μ-th coordinate direction and "groping" the graph of the function $y = f(x)$, while the result of either intersection or nonintersection of the bundle with the graph will be determined from the solution of the one-dimensional equation $\phi(\boldsymbol{Z}, t) = 0$ on $\boldsymbol{X}_\mu$ with respect to the unknown t. Keeping all the boxes that have nonempty intersection with the graph is the guarantee that the sought-for global minimum will not be lost.

To sum up, we organize the overall process of the successive improvement of the estimates for the minimum (1) similar to what has been done in the popular "branch-and-bound" based interval global optimization techniques from [1, 3, 6]:

- we arrange all the boxes, produced from the subdivision of the initial box $\boldsymbol{Z}$, as a *working list* $\mathcal{L}$;
- at each step of the algorithm, the bisected box is that from the list $\mathcal{L}$ having the smallest left endpoint of the last component, i. e. the one showing the smallest estimate of the range of f;
- we bisect only the longest component in the box to be subdivided.

Additionally, the boxes of the form (4) that the list $\mathcal{L}$ consists of will be ordered so that the values of the left endpoint of their last component (they represent the ranges of values) increase. The first record of the working list is, as usual, called *leading* for the current step of the algorithm. The overall pseudocode of the new method that we are going to call *graph subdivision method* is given in Table 1, where wid means the width of an interval.

Coming up next is a more general situation when s $(1 \le s \le n)$ coordinate directions are declared as mute, and without loss in generality we can take the numbers of these directions as 1, 2, ..., s. Let, in the space $\mathbb{R}^{n+1}$, a plane be defined, parallel to the mute coordinate directions and thus determined by the equation

$$\left\{\begin{array}{r} x_1 = t_1, \\ \vdots \\ x_{s-1} = t_{s-1}, \\ x_s = t_s, \\ x_{s+1} = r_{s+1}, \\ \vdots \\ x_n = r_n, \\ y = l, \end{array}\right. \tag{6}$$

where $t_1, \ldots, t_s$ are parameters varying over the whole of the real axis and $r_{s+1}, \ldots, r_n, l$ are some constants. Similar to the one-dimensional case, if f is continuous on $\boldsymbol{X}$,

$$\min_{x \in \boldsymbol{X}} f(x) \;=\; \min \left\{ l \in \mathbb{R} \;\middle|\; \begin{array}{l} \text{the plane defined by (6) intersects} \\ \text{the graph of the function } y = f(x) \end{array} \right\}.$$

We denote

$$\boldsymbol{Z} = (\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_{n-s+1}) \;:=\; (\boldsymbol{X}_{s+1}, \ldots, \boldsymbol{X}_n, \boldsymbol{Y}), \tag{7}$$

$$\varphi(\boldsymbol{Z}, t) \;:=\; f(t_1, \ldots, t_s, \boldsymbol{X}_{s+1}, \ldots, \boldsymbol{X}_n) - \boldsymbol{Y}, \tag{8}$$

so as the $(n-s)$-dimensional boxes $\boldsymbol{Z}$ are bundles of planes of the form (6), while either the intersection or nonintersection of such bundles with the graph will be determined from the result of the solution of the interval equation $\phi(\boldsymbol{Z}, t) = 0$ with respect to $t = (t_1, t_2, \ldots, t_s)$. Therefore, we can "grope" the graph of the objective function by the planes (6) provided that we are able to effectively check the solvability of these equations of s unknowns.

Finally, we arrange the overall process of the successive improvement of the estimates for the global minimum according to the "branch-and-bound" strategy, and the pseudocode of the resulting new algorithm presented in Table 2 is quite similar to the previous case of only one mute direction.

The two above pseudocodes are evidently intended for the computation of the function's minimum (1) only, but a straightforward modification may adjust the algorithm in order to also find the values of the variables where f takes its

Table 2. The simplest graph subdivision method for global optimization (several mute variables)

Input

A box $\boldsymbol{X} \subseteq \mathbb{R}^n$ and a function $f : \boldsymbol{X} \to \mathbb{R}$. An accuracy $\epsilon > 0$.

A method for checking the solvability of the interval equation $\varphi(\boldsymbol{Z}, t) = 0$ for $t = (t_1, \ldots, t_s)$ and φ, $\boldsymbol{Z}$, defined in (7)–(8).

Output

The lower $\underline{y}$ and upper $\overline{y}$ estimates, with the accuracy ϵ, for the global minimum of the function f over the box $\boldsymbol{X}$.

Algorithm

compute an enclosure $\boldsymbol{Y}$ of the range of f over $\boldsymbol{X}$;

assign $\boldsymbol{Z} := (\boldsymbol{X}_{s+1}, \ldots, \boldsymbol{X}_n, \boldsymbol{Y})$;

set $z := \underline{\boldsymbol{Y}}$ and $\overline{y} := \overline{\boldsymbol{Y}}$;

initialize the working list $\mathcal{L} := \{ (\boldsymbol{Z}, z) \}$;

`DO WHILE` $(\overline{y} - z \geq \epsilon)$

choose the component k of the box $\boldsymbol{Z}$ having the largest length, i. e. such that $\operatorname{wid} \boldsymbol{Z}_k = \max_{1 \leq i \leq (n-s+1)} \operatorname{wid} \boldsymbol{Z}_i$;

bisect the box $\boldsymbol{Z}$ along the k-th coordinate to half-boxes $\boldsymbol{Z}'$ and $\boldsymbol{Z}''$ such that $\boldsymbol{Z}' := (\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_{k-1}, [\underline{\boldsymbol{Z}}_k, \operatorname{mid} \boldsymbol{Z}_k], \boldsymbol{Z}_{k+1}, \ldots, \boldsymbol{Z}_{n-s+1})$, $\boldsymbol{Z}'' := (\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_{k-1}, [\operatorname{mid} \boldsymbol{Z}_k, \overline{\boldsymbol{Z}}_k], \boldsymbol{Z}_{k+1}, \ldots, \boldsymbol{Z}_{n-s+1})$;

if the equation $\phi(\boldsymbol{Z}', t) = 0$ on $(\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_s)$ is solvable or possibly solvable, then assign $z' := \underline{\boldsymbol{Z}}'_{n-s+1}$ and put the record $(\boldsymbol{Z}', z')$ into $\mathcal{L}$ so that the second fields of the records in $\mathcal{L}$ increase ;

if the equation $\phi(\boldsymbol{Z}', t) = 0$ on $(\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_s)$ is solvable, then set $\overline{y} := \min\{ \overline{y}, \overline{\boldsymbol{Z}}'_{n-s+1} \}$;

if the equation $\phi(\boldsymbol{Z}'', t) = 0$ on $(\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_s)$ is solvable or possibly solvable, then assign $z'' := \underline{\boldsymbol{Z}}''_{n-s+1}$ and put the record $(\boldsymbol{Z}'', z'')$ into $\mathcal{L}$ so that the second fields of the records in $\mathcal{L}$ increase ;

if the equation $\phi(\boldsymbol{Z}'', t) = 0$ on $(\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_s)$ is solvable, then set $\overline{y} := \min\{ \overline{y}, \overline{\boldsymbol{Z}}''_{n-s+1} \}$;

delete the former leading record $(\boldsymbol{Z}, z)$ from the list $\mathcal{L}$;

denote the new leading record of the list $\mathcal{L}$ by $(\boldsymbol{Z}, z)$;

`END DO`

$\underline{y} := z$;

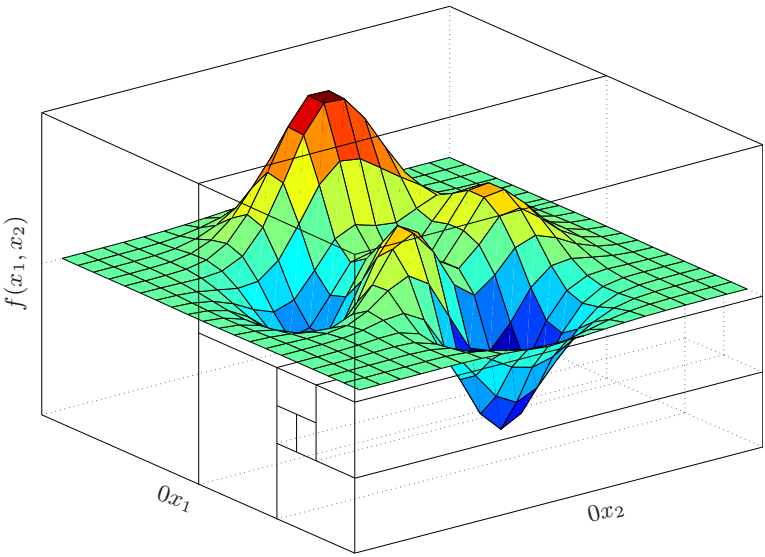

Fig. 2. A global minimization process via graph subdivision technique for an objective function $f : \mathbb{R}^2 \to \mathbb{R}$.

global minimums. Namely, we should trace and store all the roots (either certain or possible) of the "level equations" $\phi(\boldsymbol{Z}, t) = 0$ apart from the information on their solvability. This will require extending the records that compose the working list $\mathcal{L}$ to incorporate the root enclosures into them.

What can be said about the convergence of the graph subdivision methods? In the traditional interval global optimization algorithms from [1, 3, 6], the diameters of the leading boxes are well-known to tend to zero, and this should be also valid for the graph subdivision methods inasmuch as their logical structure coincides with that of the traditional methods. Therefore, the "level equations" $\phi(\boldsymbol{Z}, t) = 0$ defined by (5) and (8) tend to point (noninterval) equations. If the objective function f is such that the roots of $\phi(Z, t) = 0$ depend continuously on the parameter Z, then we can expect that the graph subdivision method converges to global optimums.

Although the graph subdivision methods may appear unnecessarily complex in comparison with the traditional ("direct") interval global optimization methods based on adaptive subdivision of the domain of the objective function, there exists a large realm of problems where both approaches have equal practicalities. These are optimization problems with implicitly defined objective functions. In such problems, evaluation of the objective fucntion requires solving an equation or a system of equations anyway.

Yet another idea that can make the graph subdivision methods much more attractive and practical is the use of constraint propagation techniques for the

solution of "level equations". This is one of the main reasons why the article is published among this collection of constraint propagation papers.

The simplest graph subdivision methods have been implemented using Sun Microsystems' Fortran 95 (also known as FORTE Fortran) and, for a number of the standard test problems, demonstrated very high sharpness of enclosing the global optimums, although achieved at the price of relatively large labor consumption. Much is to be done to modify and tune up the new idea.

4 Gradient tests

We use the term "gradient tests" to denote procedures that involve gradient of the objective function and help to discard unpromising boxes from the working list maintained by our algorithm.

If f is a continuously differentiable function then its gradient vanishes in the global minimum point which are interior in the domain $\boldsymbol{X}$. Therefore, if an enclosure of the gradient over a box $\boldsymbol{x}$ lying in the interior of $\boldsymbol{X}$ does not contain zero, then there are no extrema within $\boldsymbol{x}$. Deleting the box $\boldsymbol{x}$ from the domain of the objective function (and the corresponding record from the working list of the algorithm) will not affect the results of the global optimization process.

If the subbox $\boldsymbol{x}$ is not interior for $\boldsymbol{X}$, then we cannot discard it so simlply. Although the interior of $\boldsymbol{x}$ really cannot have extremums of f, one need to additionally investigate the part of $\boldsymbol{x}$ that shows up the boundary of the entire domain box $\boldsymbol{X}$. The techniques using gradients enclosures are very popular in the interval global optimization methods (see [1, 3, 6]), but application of the above idea in the graph subdivision methods has specific character.

In the graph subdivision methods, we do not subdivide the domain along the mute coordinate directions. As the result, all the boxes from the working list intersect the boundary of the initial domain $\boldsymbol{X}$ and never become interior subboxes. We have to take this fact into account when processing new sub-boxes during the execution of the algorithm. Let, for example, the algorithm of Table 1, with the mute direction μ, has generated a record $(\boldsymbol{Z}, z)$, $\boldsymbol{Z} = (\,\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{\mu-1}, \boldsymbol{x}_{\mu+1}, \ldots, \boldsymbol{x}_n, \boldsymbol{y}\,)$, such that within the box $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{\mu-1}, \boldsymbol{X}_\mu, \boldsymbol{x}_{\mu+1}, \ldots, \boldsymbol{x}_n\,) \subseteq \boldsymbol{X}$ the gradient of the objective function does not contain zero. Hence, the sought-for extremum can be attained only at the points from the box $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{\mu-1}, \boldsymbol{X}_\mu, \boldsymbol{x}_{\mu+1}, \ldots, \boldsymbol{x}_n)$ that goes out to the boundary $\partial\boldsymbol{X}$ of the initial box $\boldsymbol{X}$, i. e. they are in the intersection

$$\boldsymbol{X} \cap (\,\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{\mu-1}, \boldsymbol{X}_\mu, \boldsymbol{x}_{\mu+1}, \ldots, \boldsymbol{x}_n\,). \tag{9}$$

Therefore, at best we have to retain for further processing only two $(n-1)$-dimensional subboxes of $\boldsymbol{X}$, i. e.

$$(\,\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{\mu-1}, \underline{\boldsymbol{X}}_\mu, \boldsymbol{x}_{\mu+1}, \ldots, \boldsymbol{x}_n\,),$$
$$(\,\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{\mu-1}, \overline{\boldsymbol{X}}_\mu, \boldsymbol{x}_{\mu+1}, \ldots, \boldsymbol{x}_n\,),$$

obtained from $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{\mu-1}, \boldsymbol{X}_\mu, \boldsymbol{x}_{\mu+1}, \ldots, \boldsymbol{x}_n)$ by throwing away the points of the interior of $\boldsymbol{X}$, and at worst we have to retain $(2n-1)$ or even $2n$ faces of the intersection box (9) (in case $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{\mu-1}, \boldsymbol{X}_\mu, \boldsymbol{x}_{\mu+1}, \ldots, \boldsymbol{x}_n) = \boldsymbol{X}$).

The same happens to the graph subdivision methods with several mute variables, with the sole difference that the number of pieces of the boundary that we have to retain may only increase.

To sum up, in the graph subdivision methods, we never discard the subboxes entirely, but always retain parts of their boundaries. In the traditional "direct" interval global optimization methods, whole subboxes may be in the interior of the initial domain $\boldsymbol{X}$, and we entirely discard them.

5 Application to interval band linear systems

The most interesting implementations of the idea of graph subdivision methods are those where one can take several mute variables and thus substantially decrease the dimension of the argument of the original optimization problem. As a practical example of such an application, we consider the problem of outer component-wise estimation of the solution set to interval linear systems with band matrices.

Solution set to the interval linear system

$$\boldsymbol{A}x = \boldsymbol{b} \tag{10}$$

with an interval $m \times n$-matrix $\boldsymbol{A}$ and interval m-vector $\boldsymbol{b}$ is the set

$$\Xi(\boldsymbol{A}, \boldsymbol{b}) = \{\, x \in \mathbb{R}^n \mid (\exists A \in \boldsymbol{A})(\exists b \in \boldsymbol{b})(Ax = b) \,\},$$

formed by all solutions to the point linear systems $Ax = b$ for $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$. Structure of the solution set is quite complex, and usually we confine ourselves to the problems of approximate description (estimation) of the solution set according to this or that sense. For simplicity, we consider only square systems of equations with an $n \times n$-matrix $\boldsymbol{A}$.

An important problem arising in connection with the interval linear systems (10) is that of computing outer component-wise estimates of the solution set:

For an interval system of linear algebraic equations $\boldsymbol{A}x = \boldsymbol{b}$ find the estimates for $\min\{\, x_\nu \mid x \in \Xi(\boldsymbol{A}, \boldsymbol{b}) \,\}$ from below and for $\max\{\, x_\nu \mid x \in \Xi(\boldsymbol{A}, \boldsymbol{b}) \,\}$ from above, $\nu = 1, 2, \ldots, n$.	(11)

When speaking of the "solution of interval linear systems of equations" one often means the problem (11). In our work, we fix the index ν and concentrate on computing only $\min\{\, x_\nu \mid x \in \Xi(\boldsymbol{A}, \boldsymbol{b}) \,\}$, since

$$\max\{\, x_\nu \mid x \in \Xi(\boldsymbol{A}, \boldsymbol{b}) \,\} = -\min\{\, x_\nu \mid x \in \Xi(\boldsymbol{A}, -\boldsymbol{b}) \,\}.$$

The matrix $\boldsymbol{A} = (\,\boldsymbol{a}_{ij})$ (either point or interval) is called *band*, if there exists nonnegative integers p and q, such that $a_{ij} = 0$ for $j > i + p$ and $i > j + q$. Then

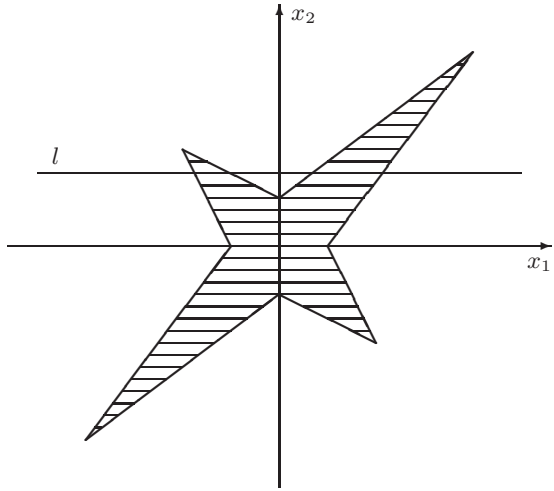

Fig. 3. Solution set and crossing it by a line.

the value of $(p+q+1)$ is the *width* of the band in the matrix $\boldsymbol{A}$. Below, we take the assumption that the band is not "too wide", namely

$$p+q \le \frac{n}{2}. \tag{12}$$

When solving the problem (11), we suppose that an interval box $\boldsymbol{V}$ is known that contains the solution set estimated, that is, $\boldsymbol{V} \supseteq \Xi(\boldsymbol{A},\boldsymbol{b})$. The box $\boldsymbol{V}$ can be found by any of the methods described e. g. in [3–5], and its size is not a big part of the entire technique.

The fact of fundamental importance is that the problem (11) of outer interval estimation of the solution set is, in essense, an optimization problem. The corresponding reformulation can be given, for example, in the following way [8]. If $\nu \in \{\,1,2,\ldots,n\,\}$ is a fixed index, then, through l, we denote a straight line in $\mathbb{R}^n$ that is parallel to the ν-th coordinate axis and has the parametric equation

$$\left\{\begin{array}{rcl} x_1 &=& r_1, \\ &\vdots& \\ x_{\nu-1} &=& r_{\nu-1}, \\ x_\nu &=& t\,, \\ x_{\nu+1} &=& r_{\nu+1}, \\ &\vdots& \\ x_n &=& r_n, \end{array}\right. \tag{13}$$

where t is a real parameter. Every such line is entirely determined by an $(n-1)$-dimensional vector $r = (r_1,\ldots,r_{\nu-1},r_{\nu+1},\ldots,r_n)^\top$, and, to explicitly show its

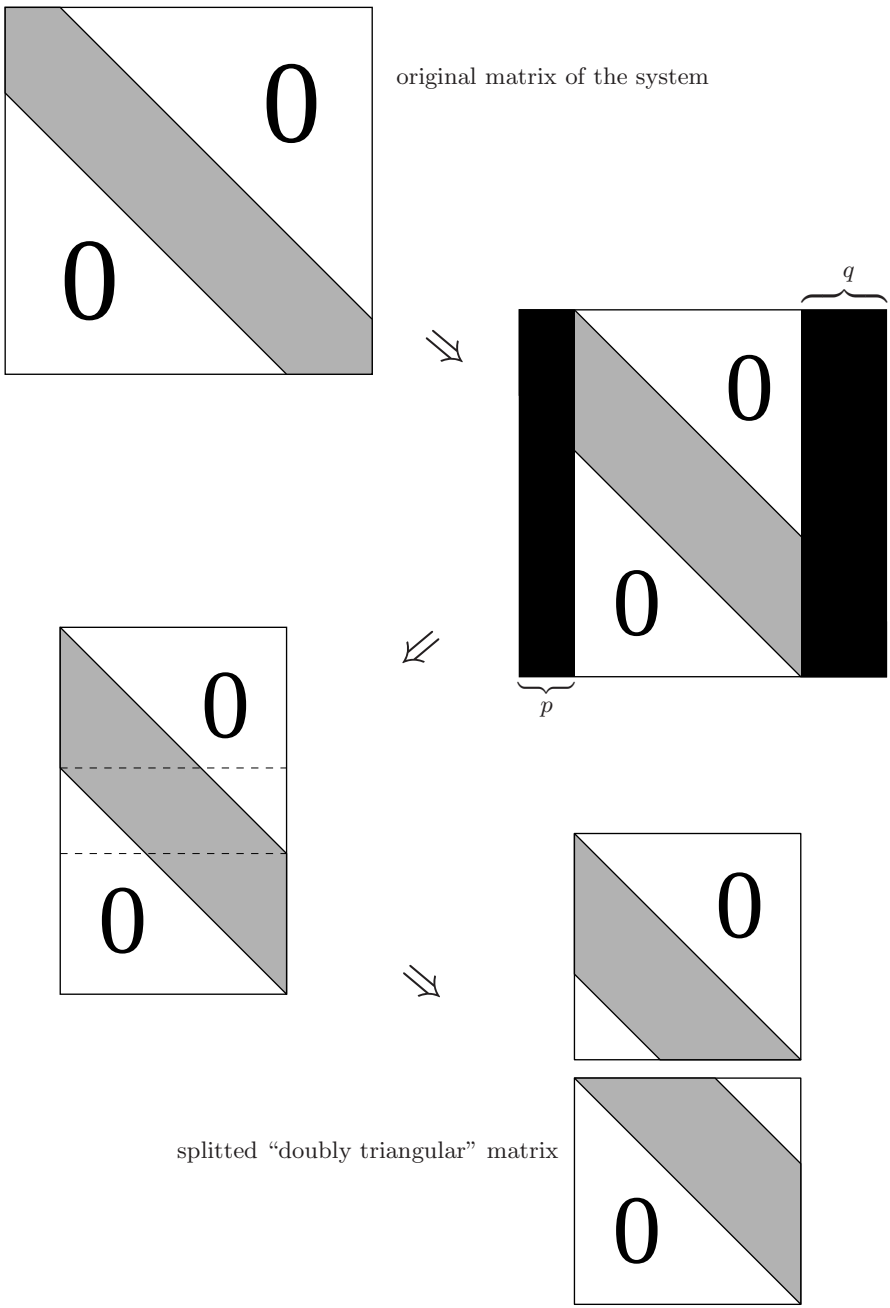

Fig. 4. How the band matrix is transformed.

parameters, we will designate this line as $l(r)$. Also, let

$$\Omega(r) \;=\; \min\{\, x_\nu \mid x \in \Xi(\boldsymbol{A},\boldsymbol{b}) \cap l(r)\,\}$$

be the smallest value of the ν-th coordinate of the points from the intersection of $l(r)$ with the solution set (10) (see Fig. 3). If $\Xi(\boldsymbol{A},\boldsymbol{b}) \cap l(r) = \varnothing$, then we set $\Omega(r) = +\infty$. Then we have

$$\min\{\, x_\nu \mid x \in \Xi(\boldsymbol{A},\boldsymbol{b})\,\} \;=\; \min\left\{\; x_\nu \;\middle|\; x \in \bigcup_{l \cap \boldsymbol{V} \neq \varnothing} \big(\Xi(\boldsymbol{A},\boldsymbol{b}) \cap l\big)\;\right\}$$

$$=\; \min\left\{ \min\{\, x_\nu \mid x \in \Xi(\boldsymbol{A},\boldsymbol{b}) \cap l(r)\} \;\middle|\; r \in (\boldsymbol{V}_1,\ldots,\boldsymbol{V}_{\nu-1},\boldsymbol{V}_{\nu+1},\ldots,\boldsymbol{V}_n)\;\right\}$$

$$=\; \min\{\, \Omega(r) \mid r \in (\boldsymbol{V}_1,\ldots,\boldsymbol{V}_{\nu-1},\boldsymbol{V}_{\nu+1},\ldots,\boldsymbol{V}_n)\,\}, \quad (14)$$

i. e. finding the ν-the coordinate estimate of the points from the solution set $\Xi(\boldsymbol{A},\boldsymbol{b})$ reduces to optimization of the objective function $\Omega(r)$ over an $(n-1)$-dimensional interval box.

The main idea of the section is to apply the technique developed in §3 to the solution of the optimization problem (14). If, in doing this, we take s mute coordinate directions, then examining, at each algorithm step, intersections of the bundles of s-dimensional planes with the solution set will require solving interval $n \times s$-systems of linear equations whose matrices are made up of the columns of $\boldsymbol{A}$ having the numbers of mute components.

Specifically, we employ the general scheme of the graph subdivision methods when the mute variables have the indices $p+1,\ \ldots,\ n-q$, so that there are $(n-p-q)$ of them in total. Let us consider in detail the situation when the number ν of estimated component satisfies $1 \le \nu \le p$ or $n-q+1 \le \nu \le n$.

If the first variables $x_1,\ x_2,\ \ldots,\ x_p$ and the last variables $x_{n-q},\ \ldots,\ x_n$ are assigned specific values, then the initial band system (10) turns into an interval linear system with an $m \times (n-p-q)$-matrix that has a special form. Namely, the boundaries of the band of its nonzero elements are diagonals of the new matrix (see Fig. 3). Hence, such a matrix is represented as the union of two (overlapping) triangular matrices, lower and upper ones. Each of the resulting triangular interval linear systems can be solved by either forward or back substitution respectively, and then the solvability of the entire interval $n \times (n-p-q)$-system can be revealed through intersecting of the enclosures for the solution sets to the upper and lower susbsystems obtained.

Therefore, the dimension of the global optimization problem (14) that we have to solve in connection with outer estimation of the solution set dicreases to just $(p+q)$, no matter what is the dimension of the initial system. For example, for tridiagonal interval linear systems this amounts to only 2.

In Tables 3–4, the overall algorithm for solving interval linear band systems is presented. Table 3 shows how checking solvability of the interval subsystems can be organized, while Table 4 gives the general algorithm. In Tables 3–4, we

Table 3. Checking solvability of subsystem generated by the algorithm of Table 4

```
DO  i = 1  TO  n
      b̆_i(Z) := b_i − Σ_{j=1}^{p} a_ij Z_j − Σ_{j=p+1}^{p+q} a_{i,j+n−p−q} Z_j
END DO
G_{p+1} := b̆_1(Z) / a_{1,p+1} ;
DO  i = p + 2  TO  n − q
      G_i := ( b̃_{i−p}(Z) − Σ_{j=p+1}^{i−1} a_{i−p,j} G_j ) / a_{i−p,i}
END DO
H_{n−q} := b̆_n(Z) / a_{n,n−q} ;
DO  i = n − q − 1  DOWNTO  p + 1
      H_i := ( b̃_{i+q}(Z) − Σ_{j=i+1}^{n−q} a_{i+q,j} H_j ) / a_{i+q,i}
END DO
IF ( G ∩ H ≠ ∅ ) THEN
      the system Ă x = b̆(Z) is solvable
ELSE
      the system Ă x = b̆(Z) is not solvable
END IF
```

denote $\boldsymbol{Z} = \big(\, \boldsymbol{Z}_1, \boldsymbol{Z}_2, \ldots, \boldsymbol{Z}_{p+q} \big)^\top \in \mathbb{IR}^{p+q}$, and

$$\breve{\boldsymbol{A}} = \big(\boldsymbol{a}_{ij} \big)_{j=p+1}^{n-q}, \qquad \breve{\boldsymbol{b}}(\boldsymbol{Z}) = \big(\breve{\boldsymbol{b}}_1(\boldsymbol{Z}), \breve{\boldsymbol{b}}_2(\boldsymbol{Z}), \ldots, \breve{\boldsymbol{b}}_n(\boldsymbol{Z}) \big)^\top,$$

$$\breve{\boldsymbol{b}}_i(\boldsymbol{Z}) \;=\; \boldsymbol{b}_i - \sum_{j=1}^{p} \boldsymbol{a}_{ij} \boldsymbol{Z}_j - \sum_{j=p+1}^{p+q} \boldsymbol{a}_{i,j+n-p-q} \boldsymbol{Z}_j .$$

The interval linear system $\breve{\boldsymbol{A}} x = \breve{\boldsymbol{b}}(\boldsymbol{Z})$ is an analogue of the "level equation" from §§2–3. To examine its solvability, we split it as shown in Fig. 4, and then compute the interval vectors $\boldsymbol{G} = (\boldsymbol{G}_1, \boldsymbol{G}_2, \ldots, \boldsymbol{G}_n)^\top$ and $\boldsymbol{H} = (\boldsymbol{H}_1, \boldsymbol{H}_2, \ldots, \boldsymbol{H}_n)^\top$, interval hulls of the solution sets to the lower and upper triangular interval linear systems obtained from $\breve{\boldsymbol{A}} x = \breve{\boldsymbol{b}}(\boldsymbol{Z})$. In the pseudocode of Table 3, $\boldsymbol{G}$ and $\boldsymbol{H}$ are found by forward substitution and back substitution respectively.

Of course, in such testing solvability of the interval system, we allow some coarsening, since we intersect not the solution sets of the subsystems, but their interval hulls (i. e., the tightest enclosures), that is, wider sets. Let us consider,

Table 4. Estimating the solution set for band interval linear systems

Input

An interval linear system $\boldsymbol{A}x = \boldsymbol{b}$ with a band matrix $\boldsymbol{A}$.

A number ν of the estimated component of the solution set.

An interval enclosure $\boldsymbol{V} \supseteq \Xi(\boldsymbol{A}, \boldsymbol{b})$ for the solution set estimated.

Output

An estimate $\underline{y}$ for $\min\{\, x_\nu \mid x \in \Xi(\boldsymbol{A}, \boldsymbol{b}) \,\}$ from below.

Algorithm

assign $\boldsymbol{Z} := (\, \boldsymbol{V}_1, \ldots, \boldsymbol{V}_p, \boldsymbol{V}_{n-q+1}, \ldots, \boldsymbol{V}_n \,)$;

set $z := \underline{\boldsymbol{V}}_\nu$;

initialize the working list $\mathcal{L} := \{\, (\boldsymbol{Z}, z) \,\}$;

`DO WHILE` (the box $\boldsymbol{Z}$ is wide)

choose the component k along which the box $\boldsymbol{Z}$ has the largest width, i. e. wid $\boldsymbol{Z}_k = \max_{1 \le i \le (p+q)}$ wid $\boldsymbol{Z}_i$;

bisect the box $\boldsymbol{Z}$ along its k-th coordinate direction to such boxes $\boldsymbol{Z}'$ and $\boldsymbol{Z}''$ that

$$\boldsymbol{Z}' := (\, \boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_{k-1}, [\, \underline{\boldsymbol{Z}}_k, \operatorname{mid} \boldsymbol{Z}_k \,], \boldsymbol{Z}_{k+1}, \ldots, \boldsymbol{Z}_{p+q}) \,,$$
$$\boldsymbol{Z}'' := (\, \boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_{k-1}, [\, \operatorname{mid} \boldsymbol{Z}_k, \overline{\boldsymbol{Z}}_k \,], \boldsymbol{Z}_{k+1}, \ldots, \boldsymbol{Z}_{p+q}) \,;$$

if the system $\breve{\boldsymbol{A}}x = \breve{\boldsymbol{b}}(\boldsymbol{Z}')$ is solvable, then assign $z' := \underline{\boldsymbol{Z}'}_\nu$ and put the pair $(\boldsymbol{Z}', z')$ into $\mathcal{L}$ so that the second field of the pairs in $\mathcal{L}$ increase;

if the system $\breve{\boldsymbol{A}}x = \breve{\boldsymbol{b}}(\boldsymbol{Z}'')$ is solvable, then assign $z'' := \underline{\boldsymbol{Z}''}_\nu$ and put the pair $(\boldsymbol{Z}'', z'')$ into $\mathcal{L}$ so that the second field of the pairs in $\mathcal{L}$ increase;

delete the former leading box $(\boldsymbol{Z}, z)$ from the list $\mathcal{L}$;

denote the current leading box of the list $\mathcal{L}$ through $(\boldsymbol{Z}, z)$;

`END DO`

$\underline{y} := z$;

as an example, the interval linear system

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} [-1, 1] \\ [-1, 1] \end{pmatrix}. \tag{15}$$

Its matrix is an upper triangular point matrix, and the solution set is depicted at Fig. 5. As the result, we can compute an estimate of the solution set to the band interval linear system which is not optimal. But our computational experience

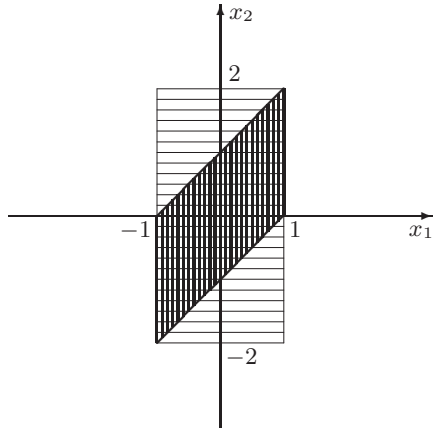

Fig. 5. Solution set for the system (15) and its interval hull.

shows that they are quite sharp providing that the band width is small and intervals in the matrix are not too wide.

As distinct from the graph subdivision methods from §§3–4 designed for the solution of general optimization problems, we do not have to involve interval bounds on the mute variables, that is, $\boldsymbol{V}_{p+1}$, $\boldsymbol{V}_{p+2}$, ..., $\boldsymbol{V}_{n-q}$. The point is that the procedure for testing solvability of the "level equations" used in the algorithm (Table 3) can spare these values. Hence, initially we suffice to know not the entire box $\boldsymbol{V} \supseteq \Xi(\boldsymbol{A}, \boldsymbol{b})$, but only its components $\boldsymbol{V}_1$, ..., $\boldsymbol{V}_p$ and $\boldsymbol{V}_{n-q+1}$, ..., $\boldsymbol{V}_n$.

There is room for further improvement of our algorithm through taking into account fine geometric structure of the solution set to triangular interval linear systems. For example, we can use prisms for enclosing them rather than axis aligned boxes. This will require additional efforts to reveal their intersection, but results in sharper estimates of the solution set.

References

1. Hansen, E., Walster G.W.: Global Optimization Using Interval Analysis. Marcel Dekker, New York (2004)
2. Kearfott, R.B., Nakao, M.T., Neumaier, A., Rump, S.M., Shary, S.P., van Hentenryck, P.: Standardized Notation in Interval Analysis. Computational Technologies 15 (1), pp. 7–13 (2010)
3. Kearfott, R.B.: Rigorous Global Search. Continuous Problems. Kluwer, Dordrecht (1996)
4. Moore, R.E., Kearfott, R.B., Cloud, M.J.: Introduction to Interval Analysis. SIAM, Philadelphia (2009)
5. Neumaier, A.: Interval Methods for Systems of Equations. Cambridge University Press, Cambridge (1990)

6. Ratschek, H. and Rokne, J.: New Computer Methods for Global Optimization. Ellis Horwood, Halsted Press, Chichester, New York (1988)
7. Semenov, A.L.: Solving optimization problems with help of the UniCalc solver. In: Kearfott, R.B. and Kreinovich, V. (eds.), Applications of Interval Computations, pp. 211–225. Kluwer, Dordrecht (1996)
8. Shary, S.P.: On Optimal Solution of Interval Linear Equations. SIAM Journal on Numerical Analysis 32 (2), pp. 610–630 (1995)
9. Shary, S.P.: A Surprising Approach in Interval Global Optimization. Reliable Computing 7 (6), pp. 497–505 (2001)