Pertanika 13(3), 409-413 (1990)

The Determination of Sub-boxes Which Might Contain the Global Minimizer(s) Using Interval Arithmetic

ISMAIL BIN MOHD
Department of Mathematics
Universiti Pertanian Malaysia
43400 UPM Serdang, Selangor Darul Ehsan, Malaysia

Keywords: Nonlinear Programming Problems, Lagrange Multipliers, Interval Arithmetic.

ABSTRAK

Dalam kertas ini diperkenalkan satu kaedah memperoleh beberapa sub-kotak yang membatasi titik Kuhn-Tucker menggunakan aritmetik selang. Kaedah ini direka sehingga penjanaan beberapa sub-kotak seperti yang diterangkan dalam bahagian 12 (Hansen, 1980) boleh dielakkan dan pada ketika yang sama kaedah ini boleh juga digunakan untuk mengira pekali Lagrange seperti yang dihuraikan dalam (Robinson, 1973).

ABSTRACT

In this paper, a method for obtaining sub-boxes which bound the Kuhn-Tucker point using interval arithmetic is presented. This method is designed such that the generation of too many sub-boxes, as described in section 12 of (Hansen, 1980), is prevented while it is simultaneously used to compute the Lagrange multipliers as described in (Robinson, 1973).

1. INTRODUCTION

If each side of a box that is a parallelepiped with sides parallel to the co-ordinate axes is divided into two parts then this could give rise to 2^n sub-boxes where n is the number of co-ordinate axes. In order to prevent generation of too many sub-boxes, Hansen (Hansen 1980) has suggested that only one side is divided with largest width in half.

In this paper, we shall show how to derive a method for obtaining and computing 2^n sub-boxes $\hat{\underline{z}}^{(i)}$ (i = 0, ..., 2^n - 1) of $\hat{\underline{z}}$ which might contain the Kuhn-Tucker point (Fiacco 1968) corresponding to a global minimizer of the problems of the special form

$$\left.\begin{array}{c}\text{minimize } f(x)\\ \text{subject to } x \in D \in \mathrm{R}^n\end{array}\right\} \qquad (1.1)$$

where D is a parallelepiped with sides parallel to the coordinate axes, simultaneously avoiding the disadvantage mentioned by Hansen in section 12 of (Hansen, 1980).

In order to obtain such-boxes, we employ interval arithmetic as explained in Section 2.

2. INTERVAL ANALYSIS

Definition 2.1

An interval $\underline{x} \in \mathrm{I(R)}$ = {intervals on real line} is denoted by $[x_I, x_S]$ where x_I and x_S are called infimum and supremum respectively.

Definition 2.2

The binary arithmetic operations +, −, ., and / are defined on I(R) according to

$$\underline{x} * \underline{y} = \{x * y \,|\, x \in \underline{x},\ y \in \underline{y};\ \underline{x}, \underline{y} \in \mathrm{I(R)}\}$$

in which $* \in \{+, -, ., /\}$ save that $\underline{x}/\underline{y}$ is not defined if $0 \in \underline{y}$.■

Definition 2.3

An nxl interval vector which is called box $\underline{x} = (\underline{x}_i)_{nx1} \in \mathrm{I}(\mathrm{R}^n)$ has its element $\underline{x}_i = [x_{iI}, x_{iS}] \in \mathrm{I(R)}$.

More details can be found in (Alefeld-Herzberger 1983).

3. Preliminary Results

Consider the nonlinear programming problems of the form

$$\text{minimize } f(x) \quad (x \in \hat{D} \subseteq R^n) \tag{3.1}$$

subject to

$$c_i(x) \geq 0 \ (i = 1, \ldots, m) \tag{3.2}$$

and

$$h_j(x) = 0 \ (j = 1, \ldots, r) \tag{3.3}$$

where $f: D \subseteq R^n \rightarrow R^1, c_i: R^n \rightarrow R^1 (i = 1, \ldots, m)$, $h_j: R^n \rightarrow R^1$ ($j = 1, \ldots, r$) are given continuously differentiable functions and $\hat{D} \subset D$ is an open set containing the points which satisfy the constraints.

In order to establish the first order necessary condition for (3.1) - (3.3) we need the following definition and theorems whereas its detail explanation can be found in (Fiacco, 1968).

Definition 3.1
The Lagrangean function $L: R^n \times R^m \times R^r \rightarrow R^1$ corresponding to (3.1) - (3.3) is defined by

$$L(x, u, w) = f(x) - \sum_{i=1}^{m} u_i c_i(x) + \sum_{i=1}^{r} w h_j(x_j) \cdot \blacksquare \tag{3.4}$$

Suppose that $x^* \in R^n$ is a feasible point for (3.1) - (3.3) and that $f: R^n \rightarrow R^1, c_i: R^n \rightarrow R^1 (i = 1, \ldots, m)$, and $h_j: R^n \rightarrow R^1$ ($j = 1, \ldots, r$) have first partial derivatives at x^*. Let $c_i^* = c_i(x^*)$ ($i = 1, \ldots, m$) etc, and let

$$B^* \equiv \{i \in N^+ | c_i(x^*) = 0\}, \tag{3.5}$$

$$Z_1^* \equiv \{z \in R^n | z^T \nabla c_i^* \geq 0 (\forall\, i \in B^*) \wedge z^T \nabla h_j^* = 0 \ (j = 1, \ldots r) \wedge z^T \nabla f^* \geq 0\}, \tag{3.6}$$

$$Z_2^* \equiv \{z \in R^n | z^T \nabla c_i^* \geq 0 (\forall\, i \in B^*) \wedge z^T \nabla h_j^* = 0 \ (j = 1, \ldots, r) \wedge z^T \nabla f^* < 0\}, \tag{3.7}$$

$$Z_3^* \equiv \{z \in R^n | (\exists\, i \in B^*, z^T \nabla c_i^* < 0) \vee (\exists\, j \in \{1, \ldots, r\}, z^T \nabla h_j^* \neq 0)\}, \tag{3.8}$$

where $\nabla c_i^* = \nabla c_i(x^*)(i = 1, \ldots, m)$, $\nabla h_j^* = \nabla h_j(x^*)$ ($j=1,\ldots,r$), $\nabla f^* = \nabla f(x^*)$ and N^+ is the set of positive integer. The sets Z_1^*, Z_2^*, and Z_3^* are disjoint and $Z_1^* \cup Z_2^* \cup Z_3^* = R^n$.

Observe that all feasible directions from x^* must be contained in $Z_1^* \cup Z_2^*$. Furthermore, $f(x)$ initially decreases along $z \in Z_2^*$ and initially increases or is constant along $z \in Z_1^*$. Thus if $Z_2^* \neq 0$ we would not expect x^* to be a local minimizer (Fiacco, 1968).

The following theorem says the existing of Lagrange multiplier generally (Fiacco, 1968).

Theorem 3.1
If (1) $x^* \in R^n$ is a feasible point for (3.1) - (3.3); (2) $f \in C^1(\hat{D})$, $c_i \in C^1(\hat{D})$ ($i = 1, \ldots, m$) and $h_j \in C^1(\hat{D})$ ($j = 1, \ldots, r$) where $\hat{D} \subset D$ is an open set containing the points which satisfy the constraints; (3) $Z_2^* = \varnothing$, then $\exists\, u^* \in R^m$ and $w^* \in R^r$ such that

$$u_i^* c_i(x^*) = 0 \quad (i = 1, \ldots, m), \tag{3.9}$$

$$u_i^* \geq 0 \quad (i = 1, \ldots, m) \tag{3.10}$$

and

$$\nabla_x L(x^*, u^*, w^*) = 0. \blacksquare \tag{3.10}$$

Definition 3.2
The point $(x^{*T}, u^{*T}, w^{*T})^T \in R^{n+m+r}$ is a Kuhn-Tucker (KT) point for (3.1) - (3.3) if and only if (3.9) - (3.11) hold. ■

In applying Theorem 3.1 one must be able to determine whether the set Z_2^* is empty or not. Clearly, assuming that the function are differentiable, $Z_2^* = \varnothing$ is a necessary and sufficient condition for the existence of the generalized Lagrange multipliers u^* and w^*.

Several conditions have been imposed to ensure that the set Z_2^* be empty at a local minimizer and we have given one of them in the following discussion.

Definition 3.3
Let $x^* \in R^n$ be a feasible point for (3.1) - (3.3) and suppose that $c_i \in C^1(\hat{D})$ ($i = 1, \ldots, m$) and $h_j \in C^1(\hat{D})$ ($j = 1, \ldots, r$) where $\hat{D} \subset D$ is an open set containing the points which satisfy the constraints. The first order constraint qualification holds at x^* if and only if

$$((z \neq 0) \wedge (z^T \nabla c_i^* \geq 0 (\forall\, i \in B^*)) \wedge (z^T \nabla h_j^* = 0 \ (j = 1, \ldots, r)))$$

implying that z is tangenial to a once differentiable arc emanating from x^* and contained in the feasible region. ■

Theorem 3.2 (Fiacco 1968)
If (1) $x^* \in R^n$ is a feasible point for (3.1) - (3.3); (2) $f \in C^1(\hat{D})$, $c_i \in C^1(\hat{D})$ $(i = 1, ..., m)$ and $h_j \in C^1(\hat{D})$ $(j = 1, ..., r)$ where $\hat{D} \subset D$ is an open set containing the points which satisfy the constraints; (3) the first order constraint qualification holds at x^* then $\exists u^* \in R^m$ and $w^* \in R^r$ such that $(x^{*T}, u^{*T}, w^{*T})^T$ is a Kuhn-Tucker point for (3.1)–(3.3).

Theorem 3.3 (Fiacco, 1968)
If (1) $x^* \in R^n$ is a feasible point for (3.1) – (3.3); (2) $f \in C^1(\hat{D})$, $c_i \in C^1(\hat{D})$ $(i=1, ..., m)$ and $h_j \in C^1(\hat{D})$ $(j = 1, ..., r)$ where $\hat{D} \subset D$ is an open set containing the points which satisfy the constraints; (3) $\{\nabla c^*_i | i \in B^*\} \cup \{\nabla h^*_j | j = 1, .., r\}$ are linearly independent then the first order constraint qualification holds at x^*. ■

4. THE NONLINEAR PROGRAMMING PROBLEM EXPRESSED AS A SYSTEM OF NONLINEAR EQUATIONS AND INEQUALITIES

Consider the following special case of the problem (3.1) - (3.3).

$$\left.\begin{array}{l} \text{minimize} f(x) \ (x \in \hat{D} \subseteq R^n) \\ \text{subject to} \\ c_i(x) \geq 0 \ (i = 1, .., 2m) \end{array}\right\} \quad (4.1)$$

in which

$$c_i(x) = x_i - \underline{x}_{iI} \quad (i = 1, ..., n) \qquad (4.2)$$

and

$$c_i(x) = \underline{x}_{i-nS} - x_{i-n}, \ (i = n + 1, ..., 2n) \qquad (4.3)$$

where $\underline{x} \in I(\hat{D})$ is given, so that $m = 2n$. The problem of bounding the solutions of (4.1) is equivalent to problem (1.1).

Example 4.1
Suppose that $n = 2$. Then

$$c_1(x) = x_1 - \underline{x}_{1I}, \qquad (4.4)$$

$$c_2(x) = x_2 - \underline{x}_{2I}, \qquad (4.5)$$

$$c_3(x) = \underline{x}_{1S} - x_1, \qquad (4.6)$$

and

$$c_4(x) = \underline{x}_{2S} - x_2. \qquad (4.7)$$

Clearly that $c_i(x) \geq 0$ $(i=1,..., 4)$. ■

According to the discussion which has been presented in Section 3, we have the following theorem where its proof is similar to the Theorem 3.3. Therefore its proof is omitted from this paper.

Theorem 4.1
If (1) $x^* \in R^n$ is a feasible solution of (4.1); (2) $f \in C^1(\hat{D})$, $c_i \in C^1(\hat{D})$ $(i = 1, .., 2n)$ where $\hat{D} \subset D$ is an open set containing the points which satisfy the constraints; (3) $\{\nabla c^*_i | i \in B^*\}$ are linearly independent, then the first order constraint qualification holds at x^*. ■

In order to obtain the computable bounds of Lagrange multipliers for the minimizer of (4.1) we need the following definition and theorem.

Definition 4.1
Strict complementary slackness is said to hold at a Kuhn-Tucker point $z^* = (x^{*T}, u^{*T})^T$ if and only if for $i = 1, .., 2n$, $u^*_i > 0$ if $c_i(x^*) = 0$ and $u^*_i = 0$ if $c_i(x^*) > 0$. ■

Theorem 4.2 (Mohd, 1990)
(a) If $x^* \in \text{int}(\underline{x})$ is a solution of (4.1) then strict complementary slackness holds at x^*; (b) if $x^* \in \beta(\underline{x})$ (Boundary of $\underline{x}$) is a solution of (4.1) then strict complementary slackness holds at x^* if and only if $\partial_i f(x^*) > 0 \ \forall i \in \{1, ..., n\}$ such that $x^*_i = \underline{x}_{iI}$ and $\partial_i f(x^*) < 0 \ \forall i \in \{1,..., n\}$ such that $x^*_i = \underline{x}_{iS}$. ■

5. THE DETERMINATION OF SUB-BOXES

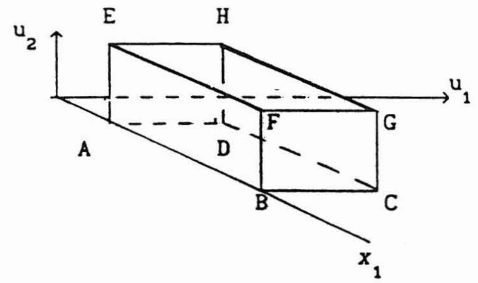
In this section, we shall derive a method to split $\underline{z}$ into 2^n sub-boxes which might contain Kuhn-Tucker point $z^* = (x^{*T}, u^{*T})^T$ corresponding to a global minimizer x^* of f in $\underline{x}$ where n is the number of components in interval vector $\underline{x}$.

Suppose that $\underline{z} = (\underline{x}^T, \underline{u}^T)^T$ is a box which is assumed to contain a Kuhn-Tucker point $z^* = (x^{*T}, y^{*T})^T$ for (4.1) where

$$\underline{x} = (\underline{x}_1, ..., \underline{x}_n)^T \text{ and } \underline{u}_1, ..., \underline{u}_{2n})^T.$$

We are given $\underline{x}$ but not $\underline{u}$. However we can determine $\underline{u}_1, ..., \underline{u}_{2n}$ which contain the corresponding Lagrange multipliers as shown in (Mohd, 1990).

Let us consider the case corresponding to n=1. So $\underline{z} = (\underline{x}, \underline{u}_1, \underline{u}_2)^T$.



In Fig 5.1, A - H are cornes of the box $\underline{z} = (\underline{x}, \underline{u}_1, \underline{u}_2)^T$. . As discussed in the previous section that

$$\underline{z} = \begin{cases} (\underline{x}, \underline{0}, \underline{0})^T \ (x^* \in \mathrm{int}(\underline{x})), \\ ([x_{1I}, x_{1I}], \underline{u}_1, \underline{0})^T \ (x^* = x_{1I}), \\ ([x_{1S}, x_{1S}], \underline{0}, \underline{u}_2)^T \ (x^* = x_{1S}). \end{cases}$$

Therefore the Kuhn-Tucker point might lie on AB, AD or BF. So $\underline{z}$ can be split into 2^1 sub-boxes $\underline{z}^{(0)}$, and $\underline{z}^{(1)}$ which contain the Kuhn-Tucker point z* and these are given by

$$\underline{z}^{(0)} = ([m(\underline{x}_1), x_{1S}], \underline{0}\, \underline{u}_2)^T \qquad (5.1)$$

and

$$\underline{z}^{(1)} = ([\underline{x}_{1I}, m(x_1)], \underline{u}_1, \underline{0})^T. \qquad (5.2)$$

For n=2 we obtain

$$\underline{z}^{(0)} = ([m(\underline{x}_1), x_{1S}), m(x_2), x_{2S}], \underline{0}, \underline{0}, \underline{u}_3, \underline{u}_4)^T. \qquad (5.3)$$

$$\underline{z}^{(1)} = ([m(\underline{x}_1), x_{1S}), (x_{2I}, m(x_2)], \underline{0}, \underline{u}_2, \underline{u}_3, \underline{0})^T. \qquad (5.4)$$

$$\underline{z}^{(2)} = ([(x_{1I}, m(\underline{x}_1)], [m(\underline{x}_2), x_{2S}], \underline{u}_1, \underline{0}, \underline{0}, \underline{u}_4)^T \qquad (5.5)$$

and

$$\underline{z}^{(3)} = ([(x_{1I}, m(\underline{x}_1)], [x_{2I}, m(\underline{x}_2)], \underline{u}_1, \underline{u}_4 \underline{0}, \underline{0},)^T \qquad (5.6)$$

Arrange the lagrange multiplier intervals given in (5.1) - (5.2) and (5.3) - (5.6) in the form

$$\begin{bmatrix} \underline{0} & \underline{u}_2 \\ \underline{u}_1 & \underline{0} \end{bmatrix} \qquad (5.7)$$

and

$$\begin{bmatrix} \underline{0} & \underline{0} & \underline{u}_3 & \underline{u}_4 \\ \underline{0} & \underline{u}_2 & \underline{u}_3 & \underline{0} \\ \underline{u}_1 & \underline{0} & \underline{0} & \underline{u}_4 \\ \underline{u}_1 & \underline{u}_2 & \underline{0} & \underline{0} \end{bmatrix} \qquad (5.8)$$

with dimension 2^2 x 2n for n = 1, and 2 respectively. For i=1, ..., n, if $\underline{u}_i \neq \underline{0}$ then $\underline{u}_{i+n} = \underline{0}$ and vice versa (See Definition 4.1.). Therefore if the first n columns of the matrices are known then the last n colums can be determined and vice versa. Also for i=1, ..., n if $\underline{u}_i \neq \underline{0}$ then the corresponding interval is$[x_{iI}, m(\underline{x}_i)]$ while if $\underline{u}_i = \underline{0}$ then` the corresponding interval is $[m(\underline{x}_i), x_{iS}]$.

From the preceding remarks, we need either the first n columns of the matrix or the last n columns of the matrix only in order to determine 2^n sub-boxes of $\underline{z}$ which contain the Kuhn-Tucker point z*. If we replace all the $\underline{u}_i \neq \underline{0}$ (i= 1,...,n) in the first n columns of (5.7) and (5.8) with unity then we obtain the matrices

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad (5.9)$$

and

$$\begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \qquad (5.10)$$

corresponding to n=1, and 2 respectively.

The rows of (5.9) and (5.10) can be considered as binary numbers which are the components of the vectors

$$(0\ 1)^T \qquad (5.11)$$

and

$$(0\ 1\ 2\ 3)^T \qquad (5.12)$$

corresponding to (5.9) and (5.10) respectively. Therefore the box $\underline{z}^{(4)}$ for n = 4 has corresponding binary digits (0 1 0 0) which give the lagrange multiplier pattern (0 1 0 0 1 0 1 1), corresponding to the lagrange multipliers

$$(0,\ u_2,\ 0,\ 0,\ u_5,\ 0,\ u_7,\ u_8).$$

So

$$\underline{Z}^{(4)} = ([m(\underline{X}_{1S}), X_{1S}], [X_{2I}, m(\underline{X}_2)],$$
$$[m(\underline{X}_3), X_{3S}], [m(\underline{X}_4), X_{4S}],$$
$$\underline{0}, \underline{U}_2, \underline{0}, \underline{0}, \underline{U}_5, \underline{0}, \underline{U}_7, \underline{U}_8)^T.$$

6. ALGORITHM

In this section, we derive an algorithm written in Pascal pseudo-code for construction the sub-boxes which have been discussed in the previous sections. In the same algorithm we can see that such sub-boxes can be called without keeping them on the stack or computer memory.

The algorithm is as follows.

Data : $\underline{X} \in I(R^n)$, $\underline{U} \in I(R^{2n})$, and $n \in N^+$

1. for i : = 0 to 2^n - 1 do
   (* Clearly we do not need to store the sub-boxes $\underline{Z}^{(i)}$ (i=0,...., 2^n - 1). *)
   1.1. $\underline{Z}^{(i)} := (\underline{0})_{3n\times 1}$
   1.2. converts a decimal integer i into the corresponding n-digit binary number b= $(b_1,....,b_n)$.
   1.3 for j : = 1 to n do
   (* Construct the sub-box $\underline{Z}^{(i)}$ *)
   1.3.1. if $b_j = 0$
   then

   1.3.1.1. $\underline{Z}^{(i)}_j := [m(\underline{X}_j), X_{jS}]$

   1.3.1.2. if $x^* \in \underline{X}$ then

   1.3.1.2.1. $\underline{Z}^{(i)}_{2n+j} := \underline{U}_{n+j}$

   else

   1.3.1.3. $\underline{Z}^{(i)}_j\ \underline{Z}^{(i)} = [X_{jI}, m(\underline{X}_j)]$

   1.3.1.4. if $x^* \in \underline{X}$ then

   1.3.1.4.1. $\underline{Z}^{(i)}_{n+j} := \underline{U}_j$

   1.4. the box $\underline{Z}^{(i)}$ is processed for example by Hansen's global optimization algorithm (Hansen, 1980)
2. return

7. CONCLUSION

In this paper, we have shown that $\underline{Z} = (\underline{X}^T, \underline{U}^T)^T$ can be split into 2^{2n} sub-boxes and given the name $\underline{Z}^{(i)}$ for i= 0,....., 2^n-1. These sub-boxes can be processed for example to compute and bound the global minimizer (s) of (1.1) by the methods given in (Robinson 1973), and (Shearer 1985), without keeping them in the computer memory.

An algorithm given in Section 6 shows us that we do not need computer memory to keep all sub-boxes. Therefore we can compute and bound the global minimizer (s) of (1.1) even though n is large. This idea can be employed for avoiding the disadvantage which is hightlight by Hanses (Hansen 1980).

REFERENCES

ALEFELD, G. and J. HERZBERGER. 1983. Introduction to Interval Computations, New York: Academic Press.

FIACCO A.V. and G. P MCCORMICK. 1968. Nonlinear Programming: Sequential Unconstrained Minimization Techniques. John Wiley and Sons.

HANSEN, E. R. 1980. Global Optimization Using Interval Analysis-The Multi-Dimensional Case. *Numer. Math* **34**: 247-270.

MOHD, I. B. 1990. Unconstrained Global Optimization Using Strict Complementary Slackness. *Applied Mathematics and Computation*, (In press).

NICKEL, K. L. E. 1971. On The Newton Method in Interval Analysis. Tecnical Summary Report No. 1136, Mathematics Research Center, University of Wisconsin, Madison, Wisconsin.

RATSCHEK H. and J. ROKNE. 1988. New Methods for Global Optimization, Chichester: Ellis Horwood.

ROBINSON, S. M. 1973. Computable Error Bounds for Nonlinear Programming. *Mathematical Programming*, 5: 235-242.

SHEARER J. M. and WOLFE, M. A. 1985. Some Computable Existence, Uniqueness, and Convergence Tests for Nonlinear Systems. *SIAM Journal on Numerical Analysis* **22**: 1200-1207.

SHEARER J. M. and WOLFE M. A. 1985. An Improved Form of the Krawczyk-Moore Algorithm. *Applied Mathematics and Computation* **17**: 229-239.

(*Received 24 March, 1990*)