AN INTERVAL METHOD FOR SYSTEMS OF ODE

S.Markov and R.Angelov
Bulgarian Academy of Sciences, Sofia, and
High Institute for Economics, Varna, Bulgaria

Abstract. Considered is an interval algorithm producing bounds for the solution of the initial value problem for systems of ordinary differential equations $\dot{x}(t)=f(t,c,x(t))$, $x(t_0)=x_0$, involving inexact data c, x_0, taking values in given intervals $C=[\underline{c}, \overline{c}]$, resp. $X_0=[\underline{x}_0, \overline{x}_0]$. An estimate for the width of the computed inclusion of the solution set is given under the assumption that f is Lipschitzian. In addition, if f is quasi-isotone, the computed bounds converge to the interval hull of the solution set and the order of global convergence is $O(h)$.

1.Notations. As usually, we denote by $I(R)$ the set of all compact intervals on the real line R of the form $A=[\underline{a}, \overline{a}]$; $V_n(I(R))$ means the set of all n-dimentional interval vectors on R of the form $([\underline{a}_1,\overline{a}_1], \ldots, [\underline{a}_n, \overline{a}_n])$. The width of $A=[\underline{a}, \overline{a}]$, $B=[\underline{b}, \overline{b}] \in I(R)$ is denoted by $w(A) = \overline{a} - \underline{a}$, the joint of A and B is denoted by $A \vee B = [\min\{\underline{a}, \underline{b}\}, \max\{\overline{a}, \overline{b}\}]$.

2.Formulation of the problem. We consider the initial value problem for systems of n ODE's:

$$\text{(1a)}\qquad \begin{aligned} \dot{x} &= f(t,c,x(t)) \\ x(t_0)&=x_0 \end{aligned}$$

involving inexact (interval) data for the parameter vector c and the initial condition vector x_0, that is

$$\text{(1b)}\qquad \begin{aligned} c \in C &= [\underline{c}, \overline{c}] \in V_m(I(R)), \\ x_0 \in X_0 &= [\underline{x}_0, \overline{x}_0] \in V_n(I(R)). \end{aligned}$$

We shall seek an enclosure $[\underline{s}, \overline{s}]$ of the set $\{\check{x}\}$ of all solutions of (1) on an interval $T=[t_0,\overline{t}]$ (assuming that all solutions $\check{x}$ of (1) does exist on T), that is $\underline{s}(t) \leq \check{x}(t) \leq \overline{s}(t)$ for every solution $\check{x}$ of (1) and every $t \in T$.

We shall assume that f is an n-vector function defined on $T \times C \times D$, $D=([\underline{d}_1,\overline{d}_1],\ldots,[\underline{d}_n,\overline{d}_n])$, such that f is continuous with respect to c, Lipschitzian with respect to t and x, and quasiisotone with respect to x.

The algorithm described below requires the effective computation of:
a) the intervals $f_i(t,C,x) = \{f_i(t,c,x): c\in C\}$, $i=1,\dots,n$, for every $t\in T$, $x\in D$; the end-points of these intervals will be further denoted by $\underline{f}_i(t,x)$, resp. $\overline{f}_i(t,x)$;
b) the intervals $\underline{F}_i(\tilde{T},\tilde{D}) := \{\underline{f}_i(t,x): t\in\tilde{T}, x\in\tilde{D}\}$, $i=1,\dots,n$, and the intervals $\overline{F}_i(\tilde{T},\tilde{D}) := \{\overline{f}_i(t,x): t\in\tilde{T}, x\in\tilde{D}\}$, $i=1,\dots,n$, for every $\tilde{T}\subset T, \tilde{D}\subset D$.

3. Description of the algorithm. Let $h>0$ be a sufficiently small step, defining a mesh $t_k = t_o + kh \in T$, $k=0,1,\dots,\tilde{k}$. The bounds $\underline{s}(t) = (\underline{s}_1(t),\dots,\underline{s}_n(t))$, $\overline{s}(t) = (\overline{s}_1(t),\dots,\overline{s}_n(t))$ for the solution set $\{\check{x}\}$ are seeked in the form of polygones with vertices at the mesh points t_k.
We set $\underline{s}(t_o) = \underline{x}_o$, $\overline{s}(t_o) = \overline{x}_o$. Assuming that $\underline{s}$, $\overline{s}$ are already computed at some t_k, that is $\underline{s}(t_k)$, $\overline{s}(t_k)$ are such that $\underline{s}_i(t_k) \le \check{x}_i(t_k) \le \overline{s}_i(t_k)$, $i=1,2,\dots,n$, we then compute $\underline{s}$, $\overline{s}$ in the interval $T_k = [t_k, t_{k+1}]$ by means of the following iteration procedure:
i) for the upper bound $\overline{s}$ we have for $r=0,1,2,\dots,\tilde{r}$

$$\overline{Z}_i^{(0)} = [\underline{d}_i, \overline{d}_i], \quad i = 1,\dots,n,$$

$$[\overline{p}_i^{(r)}, \overline{q}_i^{(r)}] = \overline{F}_i(T_k, \overline{Z}_1^{(r)}, \overline{Z}_2^{(r)},\dots,\overline{Z}_n^{(r)}), \quad i = 1,\dots,n,$$

$$\overline{Z}_i^{(r+1)} = \overline{s}_i(t_k) \vee (\overline{s}_i(t_k) + \overline{p}_i^{(r)}h) \vee (\overline{s}_i(t_k) + \overline{q}_i^{(r)}h), \quad i = 1,\dots,n,$$

$$\overline{s}_i(t) = \overline{s}_i(t_k) + \overline{q}_i^{(\tilde{r})}(t-t_k), \quad t\in T_k, \quad i = 1,\dots,n;$$

ii) for the lower bound $\underline{s}$ we compute for $r=0,1,\dots,\tilde{r}$

$$\underline{Z}_i^{(0)} = [\underline{d}_i, \overline{d}_i], \quad i = 1,\dots,n,$$

$$[\underline{p}_i^{(r)}, \underline{q}_i^{(r)}] = \underline{F}_i(T_k, \underline{Z}_1^{(r)}, \underline{Z}_2^{(r)},\dots,\underline{Z}_n^{(r)}), \quad i = 1,\dots,n,$$

$$\underline{Z}_i^{(r+1)} = \underline{s}_i(t_k) \vee (\underline{s}_i(t_k) + \underline{p}_i^{(r)}h) \vee (\underline{s}_i(t_k) + \underline{q}_i^{(r)}h), \quad i = 1,\dots,n,$$

$$\underline{s}_i(t) = \underline{s}_i(t_k) + \underline{p}_i^{(\tilde{r})}(t-t_k), \quad t\in T_k, \quad i = 1,\dots,n.$$

Theorem. For any nonnegative integer $r=0,1,2,\dots$

$$\underline{Z}_i^{(r+1)} \subset \underline{Z}_i^{(r)}, \quad \overline{Z}_i^{(r+1)} \subset \overline{Z}_i^{(r)}.$$

Proof. We have for $r=0$

$$\underline{Z}_i^{(1)} = \underline{s}_i(t_k) \vee (\underline{s}_i(t_k) + \underline{p}^{(0)}h) \vee (\underline{s}_i(t_k) + \underline{q}_i^{(0)}h), \quad i=1,\dots,n.$$

Since $\underline{s}_i(t_k) \in (\underline{d}_i, \overline{d}_i)$, we can take h sufficiently small so that $\underline{Z}_i^{(1)} \subset [\underline{d}_i, \overline{d}_i] = \underline{Z}_i^{(0)}$. Assume that $\underline{Z}_i^{(r)} \subset \underline{Z}_i^{(r-1)}$ for some $r\ge 2$. Then, since $\underline{F}_i$ is inclusion isotone, we have

$$[\underline{p}_i^{(r)}, \underline{q}_i^{(r)}] = \underline{F}_i(T_k, \underline{Z}_1^{(r)},\dots,\underline{Z}_n^{(r)}) \subset \underline{F}_i(T_k, \underline{Z}_1^{(r-1)},\dots,\underline{Z}_n^{(r-1)}) = [\underline{p}_i^{(r-1)}, \underline{q}_i^{(r-1)}].$$

Thus, $\underline{p}_i^{(r-1)} \leq \underline{p}_i^{(r)} \leq \underline{q}_i^{(r)} \leq \underline{q}_i^{(r-1)}$ and therefore

$$\underline{s}_i(t_k)+\underline{p}_i^{(r-1)}h \leq \underline{s}_i(t_k)+\underline{p}_i^{(r)}h,$$

$$\underline{s}_i(t_k)+\underline{q}_i^{(r-1)}h \geq \underline{s}_i(t_k)+\underline{q}_i^{(r)}h,$$

that is, $\underline{Z}_i^{(r+1)} \subset \underline{Z}_i^{(r)}$.

The inclusion $\overline{Z}_i^{(r+1)} \subset \overline{Z}_i^{(r)}$ is proved analogously.

We shall now prove that $\underline{s}$, $\overline{s}$ are bounds for the solution set.

Theorem. $\underline{s}(t) \leq \{\check{x}(t)\} \leq \overline{s}(t), \quad t \in T_k$.

Proof. For any nonnegative integer r (and, in particular $r=\underset{\cdot}{r}$) we have

$$\underline{s}_i'(t)=\underline{p}_i^{(r)} \leq \underline{f}_i(t,x_1,\ldots,x_n), \; t \in T_k, \; x_j \in \underline{Z}_j^{(r)}, \; j=1,\ldots,n$$

Since $\underline{Z}_j^{(r+1)} \subset \underline{Z}_j^{(r)}$, we have for every $t \in T_k$

$$\underline{s}_j'(t)=\underline{s}_j(t_k)+\underline{p}_j^{(r)}(t-t_k) \in \underline{Z}_j^{(r+1)} \subset \underline{Z}_j^{(r)}, \; j=1,\ldots,n.$$

Therefore, $\underline{s}_i'(t) \leq \underline{f}_i(t,\underline{s}_1(t),\ldots,\underline{s}_n(t)), \quad i=1,\ldots,n.$

Analogously, it can be shown that

$$\overline{s}_i'(t) \geq \overline{f}_i(t,\overline{s}_1(t),\ldots,\overline{s}_n(t)), \quad i=1,\ldots,n.$$

Let $\check{x}(t)$ be be an arbitrary solution of (1) corresponding to some $c \in C$ and $x_0 \in X_0$. We have

$$\underline{s}'(t) \leq \underline{f}(t,\underline{s}(t)) \leq f(t,c,\underline{s}(t)),$$

$$\overline{s}'(t) \geq \overline{f}(t,\overline{s}(t)) \geq f(t,c,\overline{s}(t)),$$

$$\underline{s}(t_0)=\underline{x}_0 \leq x_0 \leq \overline{x}_0=\overline{s}(t).$$

From the relations

$$\underline{s}'(t) \leq f(t,c,\underline{s}(t)), \qquad \check{x}'(t)=f(t,c,\check{x}(t)), \qquad \overline{s}'(t) \geq f(t,c,\overline{s}(t)),$$
$$\underline{s}(t_0) \leq x_0, \qquad \check{x}(t_0)=x_0, \qquad \overline{s}(t_0) \geq x_0,$$

assuming that f is quasiisotone in x, we obtain $\underline{s}(t) \leq \check{x}(t) \leq \overline{s}(t)$, according to an well known theorem of M.Müller [2].

Remark. The inclusions $[\overline{p}^{(r+1)}, \overline{q}^{(r+1)}] \subset [\overline{p}^{(r)}, \overline{q}^{(r)}]$, $[\underline{p}^{(r+1)}, \underline{q}^{(r+1)}] \subset [\underline{p}^{(r)}, \underline{q}^{(r)}]$ show that the the computed bounds of the solution set are improved at each step of the iteration procedure. Each step produces a local (that is in the interval T_k) approximation of the solution set of order $O(h^2)$; thus $\underset{\cdot}{r}=2$ is a suitable choice for practical applications. If computer arithmetic with directed roundings is available, then the effect of finite convergence can be recommended as stopping criteria of the local iteration procedure [1].

4. An estimate for the width of the obtained inclusion. For any $\tilde{T} \subset T$, $X_i \subset [\underline{d}_i, \overline{d}_i]$, $i=1,\dots,n$, we have

$$w(\underline{F}_j(\tilde{T},X_1,\dots,X_n)) = \max_{\substack{t\in\tilde{T},\\ x_i\in X_i,\, i=1,\dots,n}} \underline{f}_j(t,x_1,\dots,x_n) - \min_{\substack{t\in\tilde{T},\\ x_i\in X_i,\, i=1,\dots,n}} \underline{f}_j(t,x_1,\dots,x_n)$$

$$= \underline{f}_j(t',x_1',\dots,x_n') - \underline{f}_j(t'',x_1'',\dots,x_n'') \le$$

(assuming that f is Lipschitzian in t and x)

$$\le \tilde{L}\,|t'-t''| + \sum_{i=1}^{n} L_i\,|x_i'-x_i''|$$

$$\le \tilde{L}\,w(\tilde{T}) + \sum_{i=1}^{n} L_i w(X_i)$$

where $\tilde{L},L_1,\dots,L_n$ are some constants. Analogously we have

$$w(\overline{F}_j(\tilde{T},X_1,\dots,X_n)) \le \tilde{L}\,w(\tilde{T}) + \sum_{i=1}^{n} L_i w(X_i)$$

for any $\tilde{T}\subset T$, $X_i \subset [\underline{d}_i, \overline{d}_i]$, $i=1,\dots,n$.

The above estimates are used in the proof of the following

Theorem. The bounds $\underline{s}$, $\overline{s}$ for the solution set $\{\check{x}\}$ of problem (1) satisfy the inequality

$$\overline{s}_i(t)-\underline{s}_i(t) \le A_1w_0+A_2M+A_3h,\quad i=1,\dots,n,\quad t\in T,$$

wherein

$$w_0=w(X_0)=\max_{i=1,\dots,n} |\overline{x}_{0i}-\underline{x}_{0i}|,\qquad M=\max_{\substack{t\in T,x\in D,\\ i=1,\dots,n}} (\overline{f}_i(t,x)-\underline{f}_i(t,x)),$$

and the constants A_1, A_2, A_3 do not depend on w_0, M and h.

Proof. Let $\overline{s}_i(t_k)-\underline{s}_i(t_k) \le w_k$, $i=1,\dots,n$, $k=0,1,\dots,\tilde{k}$. We have

$$\overline{s}_i(t_{k+1})-\underline{s}_i(t_{k+1})=\overline{s}_i(t_k)+\overline{q}_i^{(\mathfrak{r})}h-\underline{s}_i(t_k)-\underline{p}_i^{(\mathfrak{r})}h$$

$$=\overline{s}_i(t_k)-\underline{s}_i(t_k) + (\overline{q}_i^{(\mathfrak{r})}-\underline{p}_i^{(\mathfrak{r})})h$$

$$\le w_k+hw(\overline{F}_i(T_k,\overline{Z}_1^{(\mathfrak{r})},\dots,\overline{Z}_n^{(\mathfrak{r})})\vee\underline{F}_i(T_k,\underline{Z}_1^{(\mathfrak{r})},\dots,\underline{Z}_n^{(\mathfrak{r})}))$$

$$\le w_k+h(w(\overline{F}_i(T_k,\overline{Z}_1^{(\mathfrak{r})},\dots,\overline{Z}_n^{(\mathfrak{r})})+w(\underline{F}_i(T_k,\underline{Z}_1^{(\mathfrak{r})},\dots,\underline{Z}_n^{(\mathfrak{r})}))$$

$$+h\,|\overline{f}_i(t_k,\overline{s}(t_k)) - \underline{f}_i(t_k,\underline{s}(t_k))|$$

$$\leq w_k+h(\tilde{L}(t_{k+1}-t_k)+\sum_{j=1}^{n} L_j w(\overline{Z}_j^{(\mathfrak{r})})+\tilde{L}(t_{k+1}-t_k)+\sum_{j=1}^{n} L_j w(\underline{Z}_j^{(\mathfrak{r})})$$

$$+h(\left|\overline{f}_i(t_k,\overline{s}(t_k))-\overline{f}_i(t_k,\underline{s}(t_k)\right| + \left|\overline{f}_i(t_k,\underline{s}(t_k))-\underline{f}_i(t_k,\underline{s}(t_k))\right|)$$

$$\leq w_k+h(2\tilde{L}h+\sum_{j=1}^{n} L_j(w(\overline{Z}_j^{(\mathfrak{r})})+w(\underline{Z}_j^{(\mathfrak{r})})))$$

$$+hw(\overline{F}_i(t_k,[\overline{s}(t_k)\vee\underline{s}(t_k)]))+hM$$

$$\leq w_k+2\tilde{L}h^2+h\sum_{j=1}^{n} L_j(w(\overline{Z}_j^{(\mathfrak{r})})+w(\underline{Z}_j^{(\mathfrak{r})}))+h\sum_{j=1}^{n} L_j w_k+Mh$$

$$=(1+hL)w_k+2\tilde{L}h^2+h\sum_{j=1}^{n} L_j(w(\overline{Z}_j^{(\mathfrak{r})})+w(\underline{Z}_j^{(\mathfrak{r})}))+Mh, \text{ where } L=\sum_{i=1}^{n} L_i.$$

We have $w(\underline{Z}_j^{(\mathfrak{r})}) \leq (\underline{s}_j(t_k)+|\underline{q}_j^{(\mathfrak{r}-1)}|h)-(\underline{s}_j(t_k)-|\underline{p}_j^{(\mathfrak{r}-1)}|h)$

$$=(|\underline{q}_j^{(\mathfrak{r}-1)}|+|\underline{p}_j^{(\mathfrak{r}-1)}|)h \leq 2Gh, \text{ wherein } G=\max_{\substack{t\in T,x\in D,\\ c\in C,i=1,\ldots,n}} |f_i(t,c,x)|.$$

Similarly, we obtain $w(\overline{Z}_j^{(\mathfrak{r})}) \leq 2Gh$.

We thus have $\overline{s}_i(t_{k+1})-\underline{s}_i(t_{k+1})\leq w_{k+1}=(1+hL)w_k+Mh+(2\tilde{L}+4GL)h^2, i=1,\ldots,n$.

From the equalities

$$w_1=(1+hL)w_0+Mh+(2\tilde{L}+4GL)h^2,$$
$$w_2=(1+hL)w_1+Mh+(2\tilde{L}+4GL)h^2,$$
$$\cdots\cdots\cdots\cdots$$
$$w_k=(1+hL)w_{k-1}+Mh+(2\tilde{L}+4GL)h^2$$

we obtain

$$w_k=(1+hL)^k w_0+hM\sum_{j=0}^{k-1}(1+hL)^j+h^2(2\tilde{L}+4GL)\sum_{j=0}^{k-1}(1+hL)^j$$

We have $h \leq (\overline{t}-t_0)/\tilde{k} \leq (\overline{t}-t_0)/k$, $k=1,\ldots,\tilde{k}$, and, therefore

$$(1+hL)^k \leq (1+(\overline{t}-t_o)L/k)^k \leq e^{(\overline{t}-t_o)L}=A_1,$$

$$h\sum_{j=0}^{k-1}(1+hL)^j=((1+hL)^k-1)/L \leq (e^{(\overline{t}-t_o)L}-1)/L=A_2.$$

Finally, we obtain for w_k

$$w_k \leq A_1 w_0+A_2 M+A_3 h,\ A_3=A_2(2\tilde{L}+4GL),\ k=1,2,\ldots,\tilde{k}.$$

From the relation $\overline{s}_i(t_k)-\underline{s}_i(t_k) \leq w_k \leq A_1 w_0+A_2 M+A_3 h$, $i=i,\ldots,n$, for $k=1,2,\ldots,\tilde{k}$ and the fact that $\overline{s}_i(t)$ and $\underline{s}_i(t)$ are polygones on T we may conclude that $\overline{s}_i(t)-\underline{s}(t) \le A_1 w_0+A_2M+A_3h$, $i=1,\dots,n, t \in T$, which proves the theorem. Let us remark that in the above estimate w_0 and M can be considered as measures for inexactness of the initial condition data, resp. of the right-hand side f of problem (1).

5.Convergence. Convergence of the enclosure $[\underline{s},\overline{s}]$ to the interval hull of the solution set $\mathrm{hull}\{\check{x}\} = [\inf\{\check{x}\}, \sup\{\check{x}\}]$ can be demonstrated under the assumption that there exist $c_1, c_2 \in C$, such that

$$\underline{f}(t,x)=f(t,c_1,x), \qquad \overline{f}(t,x)=f(t,c_2,x), \qquad t \in T,\ x \in D.$$

Denote by $\underline{x}$, $\overline{x}$ the solutions of the initialvalue problems(2), resp.(3)

(2) $\quad \dot{x} = \underline{f}(t,x), \quad x(t_o) = \underline{x}_o,$

(3) $\quad \dot{x} = \overline{f}(t,x), \quad x(t_o) = \overline{x}_o$

Let $\check{x}$ be an arbitrary solution of (1) corresponding to some $c \in C$, $x_o \in X_o$. We then have $\underline{x}' = \underline{f}(t,\underline{x}) \le f(t,c,\underline{x})$, $\overline{x}' = \overline{f}(t,\overline{x}) \ge f(t,c,\overline{x})$, $\underline{x}_o \le x_o \le \overline{x}_o$, and, by the quasi-isotonicity of f, $\underline{x}(t) \le \check{x}(t) \le \overline{x}(t)$. Since $\underline{x}(t)$, $\overline{x}(t)$ are solutions (belong to $\{\check{x}\}$), we have

$$\mathrm{hull}\{\check{x}\} = [\underline{x}(t), \overline{x}(t)].$$

Apply the algorithm to problems (2) and (3) and denote by $\underline{u}_h, \overline{u}_h$, and $\underline{v}_h$, $\overline{v}_h$ the corresponding bounds for $\underline{x}$, resp. $\overline{x}$, produced by the algorithm. Since $\underline{x}$, $\overline{x}$ are solutions of particular exact problems, we have $\overline{u}_h-\underline{u}_h \to 0$, $\overline{v}_h-\underline{v}_h \to 0$ with $O(h)$. The functions $\underline{u}_h$ and $\overline{v}_h$ are bounds for $\mathrm{hull}\{\check{x}\}$. The relations $\underline{x} - \underline{u}_h \le \overline{u}_h - \underline{u}_h \to 0$, $\overline{v}_h - \overline{x} \le \overline{v}_h - \underline{v}_h \to 0$ show that the computed bounds $\underline{u}_h$, $\overline{v}_h$ tend to $\mathrm{hull}\{\check{x}\}$ with $O(h)$.

A FORTRAN program realizing the above algorithm is available (as part of a program package called RINA).

REFERENCES

1.R.E.Moore.Methods and Applications of Interval Analysis. SIAM Studies in Applied Mathematics. 1979.

2.W.Walter.Differential and Integral Ineqalities. Springer, 1970.