Interval Computations
No 1, 1994

Solving Systems of Ordinary Differential Equations Using Adams' Interpolation Method with Guaranteed Accuracy

Oleg B. Ermakov

A technique of constructing two-sided approximations for solution of a system of ordinary differential equations based on an implicit Adams' method is considered which takes account of truncation and rounding errors. Errors in the input data are taken into account at the stage of numerical solution using the PASCAL–XSC compiler, which implements a built-in interval vector-matrix arithmetic.

Решение системы обыкновенных дифференциальных уравнений интерполяционным методом Адамса с гарантированной точностью

О. Б. Ермаков

Рассматривается способ построения двусторонних приближений решения системы обыкновенных дифференциальных уравнений на основе неявного метода Адамса с учетом погрешности метода и погрешностей округлений. Погрешности входных данных учитываются на этапе численного решения посредством применения компилятора PASCAL–XSC со встроенной интервальной векторно-матричной арифметикой.

Let us consider the problem of obtaining a guaranteed two-sided estimate for solution of a system of ordinary differential equations (ODE) in the form of

$$y' = f(t,y), \tag{1}$$

$$y\left(t_0\right) \in \left[y_0\right] \tag{2}$$

where $y, f \in R^n, [y_0] := \left[\underline{y}_0, \overline{y}_0\right] \in I\left(R^n\right)$, with inaccurate initial data using an interpolative Adams' method. Let the right-hand sides of (1) be defined and continuous in the open domain $D \subset R^{n+1} = \{t, y_1, \ldots, y_n\}$ and satisfy the off-diagonal monotonicity condition in this domain [1, p. 235]: each function $f_i(t,y)$ $\left(i = \overline{1,n}\right)$ does not decrease on $y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_n$, i.e. from $y_1^{(1)} \leq y_1^{(2)}, \ldots, y_{i-1}^{(1)} \leq y_{i-1}^{(2)}$, $y_i^{(1)} = y_i^{(2)}, y_{i+1}^{(1)} \leq y_{i+1}^{(2)}, \ldots, y_n^{(1)} \leq y_n^{(2)}$ it follows that $f_i(t, y^{(1)}) \leq f_i(t, y^{(1)})$. This condition is satisfied if intersection of D and a plane $t = \tau$ is convex in R_y^n for any τ and $\frac{\partial f_i}{\partial y_j} \geq 0, \left(i \neq j; i,j = \overline{1,n}\right)$. As in [2], we append a constant interval vector $q := \left[\underline{q}, \overline{q}\right] \in I\left(R^n\right), \underline{q}_i < 0, 0 < \overline{q}_i$ to the right-hand side of the system (1) such that the error of Adams' method and rounding errors are overcompensated for upper functions and undercompensated for lower ones. The case of one equation is considered in [3]. Inaccurate input data of the problem (1)–(2) is taken into account during numerical solution using a PASCAL–XSC compiler that implements built-in interval arithmetic. In this context, in the nodes of the net $t_k \in [t_0, T]$ $\left(t_k := t_0 + kh, t_0 + ph = T, k = \overline{1,p}\right)$ the inclusion is guaranteed

$$y\left(t_k\right) \in \left[\underline{y}_k, \overline{y}_k\right] \tag{3}$$

where $\underline{y}_k, \overline{y}_k$ are the approximate values of lower and upper functions of the ODE system (1), respectively.

Thus, in order to approximate values of lower and upper functions $[y_k] := \left[\underline{y}_k, \overline{y}_k\right] \in I\left(R^n\right)$ of the system (1), where $k = \overline{m-1, p}$, we use the formula

$$[y_k] := [y_{k-1}] \oplus h \otimes \sum_{\nu=0}^{m-1} c_\nu^* \otimes \left(F\left(t_{k-\nu}, [y_{k-\nu}]\right) \oplus q\right) \tag{4}$$

where c_ν^* are the coefficients of the Adams' interpolation formula (see e.g. [4]), $F\left(t, [y]\right)$ is a natural interval extension [5] of the function $f\left(t, y\right)$, and

$\oplus, \otimes$ are the machine arithmetic operators that allow construction of two-sided approximations that take the rounding errors $\beta_{i,k}$ into account. Let us introduce the functions $\underline{y}_{k-1}(t)$, $\overline{y}_{k-1}(t) \in R^n$, $k = \overline{1,p}$, which are exact solutions of (1) given the initial conditions $\underline{y}_{k-1}(t_{k-1}) = \underline{y}_{k-1}$ and $\overline{y}_{k-1}(t_{k-1}) = \overline{y}_{k-1}$ for constructing lower and upper functions, respectively. Actually, the interval-valued function $[y_{k-1}(t)] := \left[\underline{y}_{k-1}(t), \overline{y}_{k-1}(t)\right]$ such that $[y_{k-1}(t_{k-1})] := \left[\underline{y}_{k-1}, \overline{y}_{k-1}\right]$ is thus introduced which is represented by its boundary real-valued functions. Below we will consider the case of constructing upper functions (if not stated otherwise). If ρ_k is the error of the underlying Adams' method, then

$$\overline{y}_{k-1}(t_k) = \overline{y}_{k-1}(t_{k-1}) + h\sum_{\nu=0}^{m-1} c_\nu^* f\left(t_{k-\nu}, \overline{y}_{k-1}(t_{k-\nu})\right) - \rho_k.$$

Let also

$$\Delta_k := \overline{y}_k - \overline{y}_{k-1}(t_k), \quad k = \overline{1, m-1} \tag{5}$$

be the actual local errors of computing the starting points $\overline{y}_0, \ldots, \overline{y}_{m-1}$ so that

$$\Delta_{i,k} \in [0, \Delta_i] \tag{6}$$

where $\Delta_i := \max_{1\le k\le m-1} \Delta_{i,k}$, $i = \overline{1,n}$. Estimating (5), we can establish conditions under which the inclusions (6) hold for $m-1 \le k \le p$ as well. In order to construct lower solutions we apply the same reasoning using functions $\underline{y}_{k-1}(t)$.

Theorem. *Let the following conditions be satisfied in the domain $D \subseteq R^{n+1}$:*

1) $f_i(t, y_1, \ldots, y_n), i = \overline{1,n}$ *are m times differentiable in D;*

2) $(t_k, [y_k]) \in D, k = \overline{m-1,p}$, *where $[y_k]$ are found in accordance with (4);*

3) $\beta_{i,k} + \rho_{i_k} \in \left[\underline{\gamma}_i, \overline{\gamma}_i\right] \subseteq \left[\underline{\gamma}, \overline{\gamma}\right], k = \overline{m-1,p},$

$$\frac{\partial f_i}{\partial y_j} \in \left[\underline{l}_i, \overline{l}_i\right] \subseteq \left[\underline{l}, \overline{l}\right],$$

$$\left|\frac{\partial^2 f_i}{\partial t \partial y_j} + \sum_{\sigma=1}^{n} \frac{\partial^2 f_i}{\partial y_j \partial y_\sigma} f_\sigma + \sum_{\tau=1}^{n} \frac{\partial f_i}{\partial y_\tau} \frac{\partial f_\tau}{\partial y_j}\right| \leq M, (i, j = \overline{1, n});$$

4) *the starting points* $\underline{y}_{i,k}, \overline{y}_{i,k}, k = \overline{1, m-1}$ *satisfy the inclusions*

$$\begin{aligned} \overline{y}_{i,k} - \overline{y}_{i,k-1}(t_k) &\in [0, \Delta_i] \subseteq [0, \Delta], \\ \underline{y}_{i,k} - \underline{y}_{i,k-1}(t_k) &\in [-\Delta_i, 0] \subseteq [-\Delta, 0]; \end{aligned}$$

5) *the parameters* $\Delta, \underline{q}_i, \overline{q}_i$ *are such that the following relations hold:*

$$\begin{aligned} \Delta &\geq \frac{\overline{\gamma} - \underline{\gamma}}{1 - nhc_0^* - nh\left(\overline{l} - \underline{l}\right)\left(\eta^+ - \eta^-\right)\exp\left(-nh\underline{l}\right) - 2nh^2 b}, \\ \overline{q}_i &= h^{-1}\left(-\frac{\underline{\gamma}_i + \overline{\gamma}_i}{2} + \frac{\Delta}{2}\left(1 - nhc_0^*\overline{l} + nh\left(\eta^+ + \eta^-\right)\left(\underline{l}_i + \overline{l}_i\right)\exp\left(-nh\underline{l}\right)\right)\right), \\ \underline{q}_i &= h^{-1}\left(-\frac{\underline{\gamma}_i + \overline{\gamma}_i}{2} - \frac{\Delta}{2}\left(1 - nhc_0^*\overline{l} + nh\left(\eta^+ + \eta^-\right)\left(\underline{l}_i + \overline{l}_i\right)\exp\left(-nh\underline{l}\right)\right)\right) \end{aligned}$$

where $\Delta := \max_{1 \leq i \leq n} \Delta_i$ *and the quantities* η^+, η^-, b *are determined as follows:* $\eta^+ := \eta_3 + \eta_5 + \cdots$, $\eta^- := \eta_2 + \eta_4 + \cdots$, $\eta_{\mu+1} := c_{\mu+1}^* + \cdots + c_m^*$, $\mu = 1, 2, \ldots$; $b := \sum_{\mu=1}^{m-2} b_\mu$, $b_\mu := M\Theta\Big(|\eta_{\mu+2}| \exp(-nh\underline{l}) + \cdots + |\eta_m| \exp\big(-nh\underline{l}(m - \mu - 1)\big)\Big)$, $\Theta := \big(1 - \exp(-nh\underline{l})\big)/(nh\underline{l})$.

Then at the points $t_k \in [t_0, T]$, $k = \overline{m-1, p}$, *we have the solution of the problem (1)–(2):*

$$y(t_k) \in [y_k].$$

Comment. One can estimate the widths of the intervals obtained, for instance, by solving the corresponding differential equation for the difference of upper and lower functions with already known corrections $\overline{q}$ and $\underline{q}$. Another method to get this estimate is to directly measure the intervals $[y_k]$ of (4) with $q := [\underline{q}, \overline{q}]$, taking into account item (5) of the Theorem.

Based on this Theorem, we can propose the following algorithm for constructing upper (lower) functions of the solution of (1)–(2) using Adams' interpolation method.

To get the values of the upper (lower) functions by (4), one must have certain starting points (intervals). These starting values $[y_k], k = \overline{1, m-1}$

can be obtained, for instance, by Euler's method. In this case the actual local error is equal to the sum of the errors of the method and rounding errors. Indeed,

$$[y_k] = [y_{k-1}] \oplus h \otimes \big(F\big(t_{k-1}, [y_{k-1}]\big) \oplus q^*\big) \tag{7}$$

where $q^* := \left[\underline{q}^*, \overline{q}^*\right] \in I\left(R^n\right)$ is the correction constant for the right-hand side of (1) for Euler's method. Analogously, we can find the value of the actual local error for upper functions:

$$\begin{aligned}
\Delta_{i,k} &:= \overline{y}_{i,k} - \overline{y}_{i,k-1}\left(t_k\right) \\
&= h\Big(f(t_{k-1}, \overline{y}_{k-1}) - f\big(t_{k-1}, \overline{y}_{k-1}\left(t_{k-1}\right)\big)\Big) + \beta_{i,k} + \rho_{i,k} + h\overline{q}_i^* \\
&= \beta_{i,k} + \rho_{i,k} + h\overline{q}_i^* \qquad (8)
\end{aligned}$$

Let $\beta_{i,k} + \rho_{i,k} \in \left[\underline{\gamma}_i^*, \overline{\gamma}_i^*\right], i = \overline{1,n}$. Using the theorem on monotonicity with respect to inclusion [5], we require that the correction constants in (8) for obtaining starting points by Euler's method $\overline{q}_i^*, i = \overline{1,n}$ satisfy the condition

$$\left[\underline{\gamma}_i^*, \overline{\gamma}_i^*\right] + h\overline{q}_i^* \subseteq [0, \Delta_i].$$

Consequently, for any value of the actual local error in the starting points,

$$\Delta_i \geq \overline{\gamma}_i^* - \underline{\gamma}_i^*, \quad i = \overline{1,n} \tag{9}$$

it is possible to determine the correction vector

$$\overline{q}_i^* = h^{-1}\left(-\frac{\underline{\gamma}_i^* + \overline{\gamma}_i^*}{2} + \frac{\Delta_i}{2}\right).$$

The reasoning is analogous in the case of constructing lower functions.

After we have calculated starting values of the upper function by Euler's method with the actual local error as defined in (9), we must estimate the values in item (3) of the theorem and then inspect the inclusions

$$[0, \Delta_i] \subseteq [0, \Delta] \tag{10}$$

where Δ is determined from item (5) of the Theorem. If (10) proves true, then the solution is sought using (4). Otherwise, the step h must be decreased and the procedure repeated.

In conclusion, we would like to note that the whole process of constructing upper (lower) solutions of the problem (1)–(2) includes three major steps: choosing starting points for obtaining two-sided approximations as described in [2], computing guaranteed solutions using an extrapolation method [2], and refinement of the estimate obtained by the interpolation method described in the present paper.

References

[1] Rumyantsev, V. V. and Oziraner, A. S. *Steadiness and stabilization of motion with respect to a subset of variables.* Nauka, Moscow, 1987 (in Russian).

[2] Ermakov, O. B. *Two-sided method for solving system of ordinary differential equations with automatic determination of guaranteed estimates.* Interval Computations 3 (5) (1992), pp. 63–69.

[3] Filippov, A. F. *Obtaining upper and lower estimates for solutions of differential equations.* In: "Proceedings of the Seminar on Interval Mathematics, Saratov, May 29–31, 1990", pp. 105–110 (in Russian).

[4] Hairer, E., Nersett, S., and Vanner, G. *Solving ordinary differential equations.* Moscow, Mir, 1990 (in Russian).

[5] Kalmykov, S. A., Shokin, Yu. I., and Yuldashev, Z. Kh. *Methods of interval analysis.* Novosibirsk, Nauka, 1986 (in Russian).

Received: February 17, 1994
Revised version: July 6, 1994

Department of Mechanics and Mathematics
Saratov State University
Astrakhanskaya str., 83
410071 Saratov
Russia