ON SHARY'S ALGEBRAIC APPROACH FOR LINEAR INTERVAL EQUATIONS

ARNOLD NEUMAIER *

Abstract. A recent method by Shary for enclosing the solution set of a system of linear interval equations is derived in a new way. It is shown that the method converges to the fixed-point inverse, and that it has finite termination with probability 1.

Key words. linear interval equation, fixed point inverse, inner subtraction, H-matrix, finite termination

AMS subject classifications. 65G10

1. Introduction.

There are a variety of stationary iterative methods for enclosing the *solution set*

(1) $$\mathbf{A}^H\mathbf{b} = \square\{x \in \mathbb{R}^n \mid Ax = b \quad \text{for some } A \in \mathbf{A},\ b \in \mathbf{b}\}$$

of a system of linear equations with interval coefficient matrix $\mathbf{A} \in \mathbb{IR}^{n\times n}$ and interval right hand side $\mathbf{b} \in \mathbb{IR}^n$. By preconditioning, the system (1) is usually reduced to another one whose coefficient matrix is an H-matrix.

A detailed discussions of the enclosure methods and preconditioning techniques known in 1990 is in [2], together with an analysis of the approximation power of the methods. Other recent advances concern a method by Hansen [1], simplified and refined by Rohn [4] and Ning & Kearfott [3], giving optimal enclosures for problems involving an H-matrix with diagonal midpoint, arising by midpoint preconditioning.

Recently, Shary [5] introduced a new algorithm (called by him the "algebraic approach") for enclosing the solution set (1) when $\mathbf{A}$ is an H-matrix. While it is an iterative method, too, he empirically observed that, in exact arithmetic, the limit is usually achieved in a finite number of iterations. In this respect, the method resembles the conjugate gradient method for linear (noninterval) equations.

In the following, we rederive Shary's method in a way that makes the finite termination property explicit. We also show that the limit interval vector of Shary's algebraic approach is the fixed point inverse of $\mathbf{A}$ applied to $\mathbf{b}$, as defined in Neumaier [2].

In this paper, notation is as in [2], except that in interval quantities are in bold face.

2. A new derivation of Shary's method.

Shary's algebraic method is based on the fixed point equation

(2) $$\mathbf{z} = \mathbf{M}\mathbf{z} + G^{-1}\mathbf{b},$$

where $\mathbf{M} \in \mathbb{IR}^{n\times n}$ is an interval matrix, $\mathbf{b} \in \mathbb{IR}^n$ is an interval vector, and G is a real diagonal matrix with nonzero entries. To define Shary's algorithms in a way that makes the finite termination property apparent we introduce some machinery.

*Institut für Mathematik, Universität Wien, Strudlhofgasse 4, A-1090 Wien, Austria.
email: `neum@cma.univie.ac.at` WWW: `http://solon.cma.univie.ac.at/~neum/`

We shall need some obvious properties of the *inner subtraction* of two interval vectors $\mathbf{x}, \mathbf{y} \in \mathbb{IR}^n$, defined by

(3) $$\mathbf{x} \overset{o}{-} \mathbf{y} := [\underline{x} - \underline{y}, \bar{x} - \bar{y}].$$

This definition also makes sense for *improper interval vectors* where some lower bound exceeds the corresponding upper bound.

LEMMA 2.1.

(4) $$\mathbf{p} = \mathbf{x} \overset{o}{-} \mathbf{z} \;\Leftrightarrow\; \mathbf{z} = \mathbf{x} \overset{o}{-} \mathbf{p} \;\Leftrightarrow\; \mathbf{x} = \mathbf{z} + \mathbf{p},$$

(5) $$(\mathbf{z} + \mathbf{q}) \overset{o}{-} (\mathbf{z} + \mathbf{p}) = \mathbf{q} - \mathbf{p}.$$

Proof. Insert the definitions.
□

Shary's concept of immersion that identifies a (possibly improper) interval vector $\mathbf{x} = [\underline{x}, \bar{x}] \in \mathbb{IR}^n$ with the real vector $\begin{pmatrix} \underline{x} \\ \bar{x} \end{pmatrix} \in \mathbb{R}^{2n}$ can be dispensed with by defining instead the *extended product*

(6) $$\hat{B} * \mathbf{x} := [B^1\underline{x} + B^2\bar{x},\; B^3\underline{x} + B^4\bar{x}]$$

of a real $2n \times 2n$-matrix

$$\hat{B} = \begin{pmatrix} B^1 & B^2 \\ B^3 & B^4 \end{pmatrix}$$

with four $n \times n$ blocks $B^1, \dots, B^4$ with a (proper or improper) n-dimensional interval vector $\mathbf{x} = [\underline{x}, \bar{x}]$, emulating matrix vector multiplication in the immersed form.

LEMMA 2.2.

(7) $$\hat{B} * (\hat{C} * \mathbf{x}) = (\hat{B}\hat{C}) * \mathbf{x},$$

(8) $$\hat{B}^{-1} * (\hat{B} * \mathbf{x}) = \mathbf{x} \qquad \textit{if } \hat{B} \textit{ is nonsingular,}$$

(9) $$\hat{B} * \mathbf{x} \overset{o}{-} \hat{C} * \mathbf{x} = (\hat{B} - \hat{C}) * \mathbf{x}.$$

Proof. This follows immediately from corresponding matrix properties in dimension $2n$.
□

The importance of the extended product stems from the fact that it can be used to represent interval matrix-vector multiplication.

PROPOSITION 2.3. *For any $\mathbf{M} \in \mathbb{IR}^{n\times n}$ and $\mathbf{x} \in \mathbb{IR}^n$ there is a $2n \times 2n$ matrix*

$$\hat{M} = \begin{pmatrix} M^1 & M^2 \\ M^3 & M^4 \end{pmatrix} \qquad \textit{with } M^l_{ik} \in \{\underline{M}_{ik}, \overline{M}_{ik}, 0\} \textit{ for all } i,k,l$$

such that $\mathbf{M}\mathbf{x} = \hat{M} * \mathbf{x}$.

Proof. We have

$$\underline{(\hat{M}\mathbf{x})_i} = \sum_{k=1}^{n} \underline{\mathbf{M}_{ik}\mathbf{x}_k}$$

with

$$\begin{aligned} \underline{\mathbf{M}_{ik}\mathbf{x}_k} &= \min\{\underline{M}_{ik}\underline{x}_4, \underline{M}_{ik}\bar{x}_k, \overline{M}_{ik}\underline{x}_k, \overline{M}_{ik}\bar{x}_k\} \\ &= M^1_{ik}\underline{x}_k + M^2_{ik}\bar{x}_k \end{aligned}$$

where (M^1_{ik}, M^2_{ik}) is one of $(\underline{M}_{ik}, 0),\ (0, \underline{M}_{ik}),\ (\overline{M}_{ik}, 0),\ (0, \overline{M}_{ik})$, depending on which term in the min-expression is the smallest. Hence

$$(\underline{\mathbf{Mx}})_i = \sum_{n=1}^{n} (M^1_{ik}\underline{x}_k + M^2_{ik}\bar{x}_k) = (M^1\underline{x} + M^2\bar{x})_i$$

for all i, so that $\underline{\mathbf{Mx}} = M^1\underline{x} + M^2\bar{x}$. By a similar argument, $\overline{\mathbf{Mx}} = M^3\underline{x} + M^4\bar{x}$, where (M^3_{ik}, M^4_{ik}) also takes one of the four possibilites mentioned above.
□

Note that $\hat{M}$ depends on x and is not always unique; however, it is not difficult to extract from the proof an explicit algorithm for computing some $\hat{M}$ given $\mathbf{M}$ and $\mathbf{x}$. The fact that, independent of $\mathbf{x}$, there are only finitely many possible choices for $\hat{M}$ implies that interval matrix-vector multiplication is piecewise linear. In particular, *unless* $\mathbf{x}$ *happens to lie on one of the hypersurfaces where the linear pieces match,* $\hat{M}$ *is constant in a neighborhood of* $\mathbf{x}$. As we shall see, these observations explain the behavior of Shary's algorithm in practice.

We now use the extended product to give a new derivation of Shary's algorithm from which the finite termination property is apparent. Let $\mathbf{x}$ be an approximation to a solution $\mathbf{z}$ of the fixed point equation (2), and suppose that the matrix $\hat{M}$ of Proposition 2.3 satisfies both

$$\mathbf{Mx} = \hat{M} * \mathbf{x} \quad \text{and} \quad \mathbf{Mz} = \hat{M} * \mathbf{z}. \tag{10}$$

As mentioned above, this is the generic case when $\mathbf{x}$ and $\mathbf{z}$ are sufficiently close and on the same linear piece of the multiplication operator.

THEOREM 2.4. *If* (10) *holds and* $\hat{M} - I$ *is invertible then*

$$\mathbf{z} = \mathbf{x} \overset{\circ}{-} (\hat{M} - I)^{-1} * (\mathbf{Mx} + G^{-1}\mathbf{b} \overset{\circ}{-} \mathbf{x}). \tag{11}$$

Proof. Write $\mathbf{p} := \mathbf{x} \overset{\circ}{-} \mathbf{z}$, so that $\mathbf{x} = \mathbf{z} + \mathbf{p}$. Then

$$\begin{aligned} \mathbf{Mx} + G^{-1}\mathbf{b} \overset{\circ}{-} \mathbf{x} &= \hat{M} * (\mathbf{z} + \mathbf{p}) + G^{-1}\mathbf{b} \overset{\circ}{-} (\mathbf{z} + \mathbf{p}) && \text{by (10)}, \\ &= \hat{M} * \mathbf{z} + \hat{M} * \mathbf{p} + G^{-1}\mathbf{b} \overset{\circ}{-} (\mathbf{z} + \mathbf{p}) && \\ &= \mathbf{Mz} + G^{-1}\mathbf{b} + \hat{M} * \mathbf{p} \overset{\circ}{-} (\mathbf{z} + \mathbf{p}) && \text{by (10)}, \\ &= (\mathbf{z} + \hat{M} * \mathbf{p}) \overset{\circ}{-} (\mathbf{z} + \mathbf{p}) && \text{by (2)}, \\ &= \hat{M} * \mathbf{p} \overset{\circ}{-} \mathbf{p} && \text{by (5)}, \\ &= (\hat{M} - I) * \mathbf{p} && \text{by (9)}. \end{aligned}$$

Solving for $\mathbf{p}$ using (8) gives

$$\mathbf{p} = (\hat{M} - I)^{-1} * (\mathbf{Mx} + G^{-1}\mathbf{b} \overset{\circ}{-} \mathbf{x}),$$

and since $\mathbf{x} = \mathbf{z} + \mathbf{p}$ implies $\mathbf{z} = \mathbf{x} \ominus \mathbf{p}$, the assertion (11) follows.
□

Theorem 2.4 suggests the iteration

$$\begin{aligned} \mathbf{x}^{k+1} &= \mathbf{x}^k \ominus (\hat{M}_k - I)^{-1} * (\mathbf{M}\mathbf{x}^k + G^{-1}\mathbf{b} \ominus \mathbf{x}^k), \\ &\quad \text{with } \hat{M}_k * \mathbf{x}_k = \mathbf{M}\mathbf{x}_k \text{ from Proposition 2.3.} \end{aligned} \tag{12}$$

With the initialization

$$\mathbf{x}^0 = (\operatorname{mid} \mathbf{A})^{-1}\mathbf{b}, \tag{13}$$

this is just the algebraic method (with damping factor $\tau = 1$), as defined on p.129 of Shary [5]. In particular, Theorem 2.4 implies that as soon as $\mathbf{x}_k$ reaches the neighborhood of $\mathbf{z}$ that ensures (10), the method produces $\mathbf{x}^{k+1} = \mathbf{z}$ in the next step.

Shary proves convergence of his method under a technical assumption (6.1 in [5]) that is satisfied if the entries of $\mathbf{M}$ are sufficiently narrow together with Theorem 2.4, this gives finite termination with probability 1 (i.e., unless $\mathbf{z}$ lies on two linear pieces of the multiplication operator $\mathbf{M}$). We also see that one cannot expect finite termination when a damping factor $\tau < 1$ is used.

3. H-matrices and the fixed point inverse. A matrix $\mathbf{A} \in \mathbb{IR}^{n \times n}$ an *H-matrix* iff the comparison matrix $\langle \mathbf{A} \rangle$ defined by

$$\langle \mathbf{A} \rangle_{ii} = \langle \mathbf{A}_{ii} \rangle = \min\{|\alpha| \mid \alpha \in \mathbf{A}_{ii}\},$$

$$\langle \mathbf{A} \rangle_{ik} = -|\mathbf{A}_{ik}| = -\max\{|\alpha| \mid \alpha \in \mathbf{A}_{ik}\} \quad \text{for } i \neq k,$$

is nonsingular and its inverse is nonnegative, $\langle \mathbf{A} \rangle^{-1} \geq 0$. As a consequence, $0 \notin \mathbf{A}_{ii}$ for all i.

For H-matrices, the theory in Neumaier [2, Chapter 4] shows that the best enclosure that can be achieved with stationary iterations based on triangular splitting is the *fixed point solution set* $\mathbf{A}^F\mathbf{b}$, defined as the unique solution z of the interval equations

$$\mathbf{z}_i = (\mathbf{b}_i - \sum_{k \neq i} \mathbf{A}_{ik}\mathbf{z}_k)/\mathbf{A}_{ii} \qquad (i = 1, \ldots, n) \tag{14}$$

(Theorem 4.4.4 in [2]). $\mathbf{A}^F\mathbf{b}$ is computable by means of the interval Gauss-Seidel iteration, but when $\langle \mathbf{A} \rangle$ is ill-conditioned, this iteration converges very slowly. It turns out that for a suitable choice of $\mathbf{M}$ and G in (2), Shary's algorithm also produces the fixed point solution set $\mathbf{z} = \mathbf{A}^F\mathbf{b}$, but usually much faster.

Shary defines the *deviation* $\operatorname{dev}(\mathbf{a})$ of an interval $\mathbf{a} = [\underline{a}, \bar{a}] \in \mathbb{IR}$ from 0 to be the number

$$\operatorname{dev}(\mathbf{a}) := \begin{cases} \underline{a} & \text{if } |\underline{a}| \geq |\bar{a}|, \\ \bar{a} & \text{otherwise.} \end{cases} \tag{15}$$

Using the deviation, the quotient of two intervals can be represented as the solution of a univariate fixed point equation:

Lemma 3.1. *Let* $\mathbf{r}, \mathbf{a} \in \mathbb{IR}, 0 \in \mathbf{a}$. *Then* $\mathbf{z} = \mathbf{r}/\mathbf{a}$ *satisfies the equation*

$$\mathbf{z} = \operatorname{dev}(\mathbf{a})^{-1}\left((\operatorname{dev}(\mathbf{a}) - \mathbf{a})\mathbf{z} + \mathbf{r}\right). \tag{16}$$

Proof. CASE 1. If $\mathbf{a} > 0,\ \mathbf{r} > 0$ then

$$\text{dev}(\mathbf{a}) = \bar{a}, \quad \mathbf{z} = [\underline{r}/\bar{a}, \bar{r}/\underline{a}],$$

and the right hand side equals

$$\begin{aligned} \bar{a}^{-1}\left([0, \bar{a} - \underline{a}][\underline{r}/\bar{a}, \bar{r}/\underline{a}] + [\underline{r}, \bar{r}]\right) &= \bar{a}^{-1}\left([0, (\bar{a} - \underline{a})\bar{r}/\underline{a}] + [\underline{r}, \bar{r}]\right) \\ &= \bar{a}^{-1}[\underline{r}, \bar{a}\bar{r}/\underline{a}] = [\underline{r}/\bar{a}, \bar{r}/\underline{a}] = \mathbf{z}. \end{aligned}$$

CASE 2. If $\mathbf{a} > 0,\ \mathbf{r} \ni 0$ then

$$\text{dev}(\mathbf{a}) = \bar{a}, \quad \mathbf{z} = [\underline{r}/\underline{a}, \bar{r}/\underline{a}],$$

and the right hand side equals

$$\begin{aligned} \bar{a}^{-1}\left([0, \bar{a} - \underline{a}][\underline{r}/\underline{a}, \bar{r}/\underline{a}] + [\underline{r}, \bar{r}]\right) &= \bar{a}^{-1}\left([(\bar{a} - \underline{a})\underline{r}/\underline{a}, (\bar{a} - \underline{a})\bar{r}/\underline{a}] + [\underline{r}, \bar{r}]\right) \\ &= \bar{a}^{-1}[\bar{a}\underline{r}/\underline{a}, \bar{a}\bar{r}/\underline{a}] = [\underline{r}/\underline{a}, \bar{r}/\underline{a}] = \mathbf{z}. \end{aligned}$$

The other cases can be reduced to one of these two by changing the signs of $\mathbf{r}$ and/or $\mathbf{a}$.
□

As explained in SHARY [5, pp. 129–130], Shary's algorithm for enclosing the solution set (1) is based on the fixed point equation (2), where

$$G := \text{Diag}\left(\text{dev}(\mathbf{A}_{ii})\right), \quad \mathbf{M} = G^{-1}(G - \mathbf{A}), \tag{17}$$

and the spectral radius $\rho(|\mathbf{M}|)$ of $|\mathbf{M}|$ is assumed to be less than one. The condition $\mathbf{A}_{ii} \neq 0$ is also needed to ensure that G is invertible.

(Some of Shary's theory is more general, but the only situation worked out in the algorithmic stage is the one stated here.)

THEOREM 3.2. *$\mathbf{A} \in \mathbb{IR}^{n\times n}$ is an H-matrix iff G is invertible and the spectral radius of $|\mathbf{M}|$ is less than one. In this case, $\mathbf{z} = \mathbf{A}^F\mathbf{b}$ is the unique solution of the fixed point equation* (2).

Proof. Suppose first that G is invertible and $\rho(|\mathbf{M}|) < 1$. By Perron-Frobenius theory, $I - |\mathbf{M}|$ is invertible and its inverse is nonnegative. Since G is diagonal and $\mathbf{A} = G(I - \mathbf{M})$ we have $\langle\mathbf{A}\rangle = \langle G\rangle(I - |\mathbf{M}|)$ and $\langle\mathbf{A}\rangle^{-1} = (I - |\mathbf{M}|)^{-1}\langle G\rangle^{-1} \geq 0$. Hence $\mathbf{A}$ is an H-matrix.

Conversely, suppose that $\mathbf{A}$ is an H-matrix. Then $0 \neq \mathbf{A}_{ii}$ whence $\text{dev}(\mathbf{A}_{ii}) \neq 0$ and G is invertible. Moreover, $I - |\mathbf{M}| = \langle G\rangle^{-1}\langle\mathbf{A}\rangle$ is invertible, with inverse $(I - |\mathbf{M}|)^{-1} = \langle\mathbf{A}\rangle^{-1}\langle G\rangle \geq 0$. Again by Perron-Frobenius theory, this implies that $\rho(|\mathbf{M}|) < 1$. Now let $\mathbf{z} = \mathbf{A}^F\mathbf{b}$. By (14) we have $\mathbf{z}_i = \mathbf{r}_i/\mathbf{A}_{ii}$, where

$$\mathbf{r}_i = \mathbf{b}_i - \sum_{k\neq i}\mathbf{A}_{ik}\mathbf{z}_k. \tag{18}$$

Using $\text{dev}(\mathbf{A}_{ii}) = G_{ii}$ we find

$$\begin{aligned} \mathbf{z}_{ii} &= G_{ii}^{-1}\left((G_{ii} - \mathbf{A}_{ii})\mathbf{z}_i + \mathbf{r}_i\right) && \text{by Lemma 3.1,} \\ &= G_{ii}^{-1}\left((G_{ii} - \mathbf{A}_{ii})\mathbf{z}_i - \sum_{k\neq i}\mathbf{A}_{ik}\mathbf{z}_k + \mathbf{b}_i\right) && \text{by (18),} \\ &= G_{ii}^{-1}\left((G - \mathbf{A})\mathbf{z} + \mathbf{b}\right)_i && \text{since } G \text{ is diagonal,} \\ &= (\mathbf{M}\mathbf{z} + G^{-1}\mathbf{b})_i && \text{by (17).} \end{aligned}$$

Hence $\mathbf{z}$ is a solution of the fixed point equation (2). Since $|\mathbf{M}|$ has spectral radius <1, this equation has a unique solution, so $\mathbf{z} = \mathbf{A}^F\mathbf{b}$ is the only solution of (2). □

Together with the good overestimation properties of $\mathbf{A}^F\mathbf{b}$ (derived in NEUMAIER [2]) when $\mathbf{A}$ is strictly diagonally dominant, this result explains the good numerical properties of the enclosures computed in SHARY [5].

REFERENCES

[1] E. Hansen, Bounding the solution of interval linear equations, SIAM J. Numer. Anal. 29 (1992), 1493-1503.

[2] A. NEUMAIER, *Interval Methods for Systems of Equations*, Cambridge Univ. Press, Cambridge 1990.

[3] S. Ning and R. B. Kearfott, A comparison of some methods for solving linear interval equations, SIAM J. Numer. Anal. 34 (1997), 1289-1305.

[4] J. Rohn, Cheap and tight bounds: the recent result by E. Hansen can be made more efficient, Interval Computations 4 (1993), 13-21.

[5] S.P. SHARY, *Algebraic approach in the "outer problem" for interval linear equations*, Reliable Computing 3 (1997), pp. 103–135.