Conjugate directions method for solving interval linear systems *

A.H. Bentbib

Faculté des Sciences et Techniques–Guéliz, Département de Mathématiques et Informatique, Université Cadi Ayyad de Marrakech, BP 618, Marrakech, Morocco

We propose the interval version of the conjugate directions method, to solve the problem of linear systems, with symmetric and positive definite interval matrix A, and a right-hand side interval vector b.

Keywords: interval, symmetric, positive definite, conjugate direction

1. Introduction

It is well known that in the formulae of the Gaussian algorithm, the Jacobi and Gauss–Seidel iterations can be used to bound the set S of solutions of linear systems for which the coefficient matrices and the right-hand sides are varying within given intervals [3,6,7], given by

$$S = \{x \in \mathbb{R}^n \mid \exists \widetilde{A} \in A,\ \exists \widetilde{b} \in b\colon\ \widetilde{A}x = \widetilde{b}\}.$$

Much work has been done to compute an enclosure interval vector of the set S, see, for example, [5,7–10,12,13].

We are interested here in solving the symmetric interval linear systems $Ax = b$, where A is an (n, n) symmetric interval matrix (i.e., $A_{i,j} = A_{j,i}$), and b is an interval vector. The set of symmetric solutions of such problems is given by

$$S_{\text{sym}} = \{x \in \mathbb{R}^n \mid \exists \widetilde{A} \in A,\ \widetilde{A}^{\mathrm{T}} = \widetilde{A},\ \exists \widetilde{b} \in b\colon\ \widetilde{A}x = \widetilde{b}\}.$$

The formulae of the Cholesky method can be used to solve the symmetric interval linear systems [1,2]. The purpose of the present paper is to apply the well-known conjugate directions method to compute an enclosure interval vector of S_{sym}. When the interval matrix A is not symmetric, we set $B = A^{\mathrm{T}}A$, where B is a symmetric interval matrix. Consider the interval linear system $Ax = b$, for which the set of solutions is given by

$$S = \{x \in \mathbb{R}^n \mid \exists \widetilde{A} \in A,\ \exists \widetilde{b} \in b\colon\ \widetilde{A}x = \widetilde{b}\}.$$

* This work was supported by the "Action Intégrée 1039/96" program, between the University Cadi Ayyad of Marrakech, Morocco, and the University Henri Poincaré of Nancy, France.

Consider now the symmetric interval linear system $Bx = c$, where $c = A^{\mathrm{T}}b$. The set of symmetric solutions is

$$S_{\text{sym}} = \{x \in \mathbb{R}^n \mid \exists \widetilde{B} \in B,\ \widetilde{B}^{\mathrm{T}} = \widetilde{B},\ \exists \widetilde{c} \in c\colon\ \widetilde{B}x = \widetilde{c}\}.$$

Proposition 1.1. We have the following inclusion:

$$S \subseteq S_{\text{sym}}.$$

Proof.

$$\begin{aligned} x \in S \implies & \exists \widetilde{A} \in A,\ \exists \widetilde{b} \in b\colon\ \widetilde{A}x = \widetilde{b} \implies \exists \widetilde{A} \in A,\ \exists \widetilde{b} \in b\colon\ \widetilde{A}^{\mathrm{T}}\widetilde{A}x = \widetilde{A}^{\mathrm{T}}\widetilde{b} \\ \implies & \exists \widetilde{B} \in B,\ \widetilde{B}^{\mathrm{T}} = \widetilde{B},\ \exists \widetilde{c} \in c\colon\ \widetilde{B}x = \widetilde{c} \implies x \in S_{\text{sym}}. \end{aligned}$$

□

2. Notations

By $\mathbb{IR}$, we denote the set of real compact intervals

$$[\alpha, \beta] = \{x \in \mathbb{R} \mid \alpha \leqslant x \leqslant \beta\}, \quad \text{for } \alpha \leqslant \beta;\ \alpha, \beta \in \mathbb{R}.$$

$\mathbb{IR}^{n\times m}$ is the set of (n, m) interval matrices A, whose elements $A_{i,j}$ belong to $\mathbb{IR}$. If $m = 1$, $\mathbb{IR}^{n\times 1}$ is denoted by $\mathbb{IR}^n$, and it represents the set of vectors with n interval components. For an interval $I = [\alpha, \beta] \in \mathbb{IR}$, we denote by

$$\begin{aligned} \check{I} &= \frac{\alpha + \beta}{2}: \quad \text{the midpoint of } I, \\ \rho(I) &= \frac{\beta - \alpha}{2}: \quad \text{the radius of } I. \end{aligned}$$

For $A \in \mathbb{IR}^{n\times m}$, $\check{A}$ is the real (n, m) matrix whose elements $\check{A}_{ij}$ are midpoints of corresponding elements A_{ij} of A. $\rho(A)$ is the positive real (n, m) matrix whose elements $\rho(A)_{i,j}$ are radii of corresponding elements A_{ij} of A.

3. Operations

If $*$ is one of the symbols $+$, $-$, $\cdot$, $/$, we define arithmetic operations on intervals by

$$[\alpha, \beta] * [\gamma, \delta] = \{x * y \mid \alpha \leqslant x \leqslant \beta,\ \gamma \leqslant y \leqslant \delta\}$$

except that we do not define $[\alpha, \beta]/[\gamma, \delta]$ if $0 \in [\gamma, \delta]$. $[\alpha, \beta] * [\gamma, \delta]$ is a real compact interval, and it is equal to

$$\bigl[\min\{\alpha * \gamma,\ \alpha * \delta,\ \beta * \gamma,\ \beta * \delta\}, \max\{\alpha * \gamma,\ \alpha * \delta,\ \beta * \gamma,\ \beta * \delta\}\bigr].$$

For $A, B \in \mathbb{IR}^{n\times m}$, $C = A \pm B$, is the (n, m) interval matrix whose elements are $C_{i,j} = A_{i,j} \pm B_{i,j}$. If $B \in \mathbb{IR}^{m\times s}$, $C = A \cdot B$, is the (n, s) interval matrix whose elements are

$$C_{i,j} = \sum_{k=1}^{m} A_{i,k} \cdot B_{k,j}.$$

If $V \in \mathbb{IR}^m$, $W = A \cdot V$ is the interval vector, whose components are given by

$$W_i = \sum_{k=1}^{m} A_{i,k} \cdot V_k.$$

$W = [\alpha, \beta] \cdot V$ is the interval vector with components $W_i = [\alpha, \beta] \cdot V_i$. If $0 \notin [\alpha, \beta]$,

$$\frac{V}{[\alpha, \beta]} = \frac{1}{[\alpha, \beta]} \cdot V.$$

4. Independence, norm and orthogonality in $\mathbb{IR}^n$

Definition 4.1. A set of interval vectors $\mathfrak{I} = \{U_1, U_2, \dots, U_p\}$ of $\mathbb{IR}^n$ is said to be linearly independent if each set of real vectors $\{u_1, u_2, \dots, u_p\}$, with $u_i \in U_i$, for $i = 1, \dots, p$, is linearly independent in $\mathbb{R}^n$.

Proposition 4.2. A set of interval vectors $\mathfrak{I} = \{U_1, U_2, \dots, U_p\}$ of $\mathbb{IR}^n$ is linearly independent if and only if

$$0 \in \alpha_1 U_1 + \alpha_2 U_2 + \cdots + \alpha_p U_p \Rightarrow \alpha_1 = \alpha_2 = \cdots = \alpha_p = 0.$$

Proof. Suppose that the set $\mathfrak{I} = \{U_1, U_2, \dots, U_p\}$ of $\mathbb{IR}^n$ is linearly independent. Let $\alpha_1, \alpha_2, \dots, \alpha_p \in \mathbb{R}$, such that

$$0 \in \alpha_1 U_1 + \alpha_2 U_2 + \cdots + \alpha_p U_p.$$

Then, there exists at least one set of real vectors $\{u_1, u_2, \dots, u_p\}$, $u_i \in U_i$, $i = 1, \dots, p$, which verify

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_p u_p = 0.$$

This implies $\alpha_1 = \alpha_2 = \cdots = \alpha_p = 0$. Now, if $\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_p u_p = 0$, then $0 \in \alpha_1 U_1 + \alpha_2 U_2 + \cdots + \alpha_p U_p$, which implies $\alpha_1 = \alpha_2 = \cdots = \alpha_p = 0$. Thus, all sets $\{u_1, u_2, \dots, u_p\}$, $u_i \in U_i$, of real vectors are linearly independent. This means $\mathfrak{I} = \{U_1, U_2, \dots, U_p\}$ is linearly independent. $\square$

Example 4.3. Sets of interval vectors

$$\mathfrak{I}_1 = \left\{ \begin{pmatrix} [1,2] \\ [0,0] \end{pmatrix}, \begin{pmatrix} [0,0] \\ [1,2] \end{pmatrix} \right\} \quad \text{and} \quad \mathfrak{I}_2 = \left\{ \begin{pmatrix} [1,2] \\ [9,10] \end{pmatrix}, \begin{pmatrix} [1,1] \\ [1,2] \end{pmatrix} \right\}$$

of $\mathbb{IR}^2$ are linearly independent.

Let V be an interval vector belonging to $\mathbb{IR}^n$. The norm of V is designated by $\|V\|$, and computed as

$$\|V\| = \sum_{k=1}^{n} V_k^2,$$

where, for $X \in \mathbb{IR}$, $X^2 = \{x^2 \mid x \in X\}$ and, if $X \geqslant 0$ (that is, $\forall x \in X,\ x \geqslant 0$), $\sqrt{X} = \{\sqrt{x} \mid x \in X\}$. The square of the norm of V is given by

$$\|V\|^2 = \sum_{k=1}^{n} V_k^2.$$

Let now A be a symmetric positive definite interval (n,n) matrix, i.e., $A_{i,j} = A_{j,i}$ and for all real matrices $\widetilde{A}$ belonging to A, $\widetilde{A}$ is positive definite (for details, see [11]). The norm of V associated with A, denoted by $\|V\|_A$, is computed as

$$\|V\|_A = \sqrt{\sum_{k=1}^{n} \left(A_{k,k} \cdot V_k^2\right) + 2 \cdot \sum_{k=1}^{n-1} \left(V_k \cdot \sum_{j=k+1}^{n} A_{k,j} \cdot V_j\right)},$$

$$\|V\|_A^2 = \sum_{k=1}^{n} \left(A_{k,k} \cdot V_k^2\right) + 2 \cdot \sum_{k=1}^{n-1} \left(V_k \cdot \sum_{j=k+1}^{n} A_{k,j} \cdot V_j\right).$$

$\|V\|_A^2$ contains all $\widetilde{v}^{\mathrm{T}} \cdot (\widetilde{A} \cdot \widetilde{v})$, where $\widetilde{v} \in V$, and $\widetilde{A}$ is a symmetric real matrix belonging to A. Because of phenomenon of dependence of interval arithmetic, we have $\|V\|_A^2 \subseteq V^{\mathrm{T}} \cdot (A \cdot V)$.

Definition 4.4. Let A be a symmetric positive definite interval (n,n) matrix. A set of interval vectors $\{P_1, P_2, \ldots, P_m\}$ of $\mathbb{IR}^n$ is said to be A-orthogonal if for each symmetric real matrix $\widetilde{A} \in A$ and each real vector $\widetilde{p}_1 \in P_1$, there exists a set of real vectors $\{\widetilde{p}_2, \ldots, \widetilde{p}_m\}$, $\widetilde{p}_j \in P_j$, $j = 2, \ldots, m$, such that $\widetilde{p}_i^{\mathrm{T}} \cdot (\widetilde{A} \cdot \widetilde{p}_j) = 0$, for $1 \leqslant i \neq j \leqslant m$.

Such an A-orthogonal set of interval vectors can be found by

$$\begin{aligned}
P_1 &= e_1, \\
P_2 &= e_2 - \frac{P_1^{\mathrm{T}} \cdot (A \cdot e_2)}{\|P_1\|_A^2} \cdot P_1, \\
&\ \ \vdots \\
P_m &= e_m - \sum_{k=1}^{m-1} \frac{P_k^{\mathrm{T}} \cdot (A \cdot e_m)}{\|P_k\|_A^2} \cdot P_k,
\end{aligned}$$

where e_i is the real vector with one in the ith component, and zero otherwise. To compute P_{k+1}, we must have $0 \notin \|P_k\|_A$.

5. Steepest descent in interval arithmetic

Let $A \in \mathbb{IR}^{n\times n}$ be a symmetric positive definite (s.p.d.) interval matrix, i.e., each real matrix $\widetilde{A} \in A$ is positive definite [11], $b \in \mathbb{IR}^n$ an interval vector. We consider the formal minimization problem

$$\min\left\{\Phi(x) = \frac{1}{2}x^{\mathrm{T}}\cdot(A\cdot x) - x^{\mathrm{T}}\cdot b\right\},$$

which by definition means the minimization of all functions

$$\phi_{\tilde{A},\tilde{b}}(x) = \frac{1}{2}x^{\mathrm{T}}\cdot\big(\widetilde{A}\cdot x\big) - x^{\mathrm{T}}\cdot\widetilde{b},$$

where $\widetilde{A}$ is a real s.p.d. matrix which belongs to A, and $\widetilde{b}$ is real vector varying in the interval vector b. It is well known that the minimum value of $\phi_{\tilde{A},\tilde{b}}(x)$ is achieved by setting $\widetilde{x} = \widetilde{A}^{-1}\cdot\widetilde{b}$ [4]. Let Σ be the set

$$\big\{\widetilde{x} \in \mathbb{R}^n \mid \exists\widetilde{A} \in A,\ \widetilde{A}\ \text{s.p.d.},\ \exists\widetilde{b} \in b\colon\ \phi_{\tilde{A},\tilde{b}}\big(\widetilde{x}\big) = \min\phi_{\tilde{A},\tilde{b}}(x)\big\},$$

then it is obvious that $\Sigma = S_{\mathrm{sym}}$. Thus, solving the interval minimization problems and the symmetric, positive definite interval linear systems are equivalent. The method of steepest descent consists in minimizing each functional $\phi_{\tilde{A},\tilde{b}}$, at a current point x, in the direction of the negative gradient

$$-\nabla\phi_{\tilde{A},\tilde{b}}(x) = \widetilde{b} - \widetilde{A}\cdot x.$$

We call $\widetilde{r} = \widetilde{b} - \widetilde{A}\cdot x$ the residual of x associated with $\widetilde{A}$ s.p.d. $\in A$ and $\widetilde{b} \in b$. If the residual $\widetilde{r}$ is nonzero, then

$$\widetilde{\alpha} = \frac{\widetilde{r}^{\mathrm{T}}\cdot\widetilde{r}}{\widetilde{r}^{\mathrm{T}}\cdot(\widetilde{A}\cdot\widetilde{r})}$$

minimizes $\phi_{\tilde{A},\tilde{b}}(x + \alpha\cdot\widetilde{r})$. This gives the algorithm

$X_0 = 0,\ R_0 = b$
for $k = 1, 2, \ldots$
  if $0 \notin R_{k-1}$, then $\alpha_{k-1} = \|R_{k-1}\|^2/\|R_{k-1}\|_A^2$
  $X_k = X_{k-1} + \alpha_{k-1}\cdot R_{k-1}$
  $R_k = b - A\cdot X_k$

If $0 \notin R_{k-1}$, then all residuals $\widetilde{r}_{k-1} = \widetilde{b} - \widetilde{A}\cdot\widetilde{x}_{k-1}$, where $\widetilde{A}$ s.p.d. $\in A$, $\widetilde{b} \in b$ and $\widetilde{x}_{k-1} \in X_{k-1}$, are nonzero.

$$\widetilde{\alpha}_{k-1} = \frac{\widetilde{r}_{k-1}^{\mathrm{T}}\cdot\widetilde{r}_{k-1}}{\widetilde{r}_{k-1}^{\mathrm{T}}\cdot(\widetilde{A}\cdot\widetilde{r}_{k-1})}\quad\text{is in the interval}\quad\alpha_{k-1} = \frac{\|R_{k-1}\|^2}{\|R_{k-1}\|_A^2},$$

thus $\widetilde{x}_k = \widetilde{x}_{k-1} + \widetilde{\alpha}_{k-1}\cdot\widetilde{r}_{k-1} \in X_k$. Unfortunately, this algorithm fails when the interval residual vector R_{k-1} contains zero, which is possible when X_{k-1} contains at least one solution $\widetilde{x}$ for a given minimization problem associated with one real s.p.d. matrix $\widetilde{A} \in A$ and one real vector $\widetilde{b} \in b$.

6. A-conjugate direction method in interval arithmetic

To avoid the problem of the interval steepest descent method, when the residual contains zero, we consider the successive minimization of Φ along a set of A-orthogonal interval vector directions $\{P_1, P_2, \ldots, P_n\}$ of $\mathbb{IR}^n$. Thus, we obtain the following algorithm:

$X_0 = 0,\ R_0 = b$
for $k = 1, \ldots, n$
  $\alpha_{k-1} = (P_k^{\mathrm{T}} \cdot b)/\|P_k\|_A^2$
  $X_k = X_{k-1} + \alpha_{k-1} \cdot P_k$
  $R_k = b - A \cdot X_k$

Proposition 6.5. At each step k of the algorithm given above, $\rho(b) \leqslant \rho(R_k)$, so the radius of the interval $P_k^{\mathrm{T}} \cdot b$ is less than the radius of $P_k^{\mathrm{T}} \cdot R_{k-1}$.

Proof. The proof is based on a simple remark: for $I, J \in \mathbb{IR}$ we have $\rho(I \pm J) = \rho(I) + \rho(J) \leqslant \rho(I)$. □

Theorem 6.6. Let $\{P_1, P_2, \ldots, P_n\}$ be a set of A-orthogonal interval vector directions of $\mathbb{IR}^n$ that verify $0 \notin P_k$, for $k = 1, \ldots, n$. The interval vector sequences X_k constructed by the algorithm given above verify $\boldsymbol{\Sigma} \subseteq X_n$.

Proof. Let $\widetilde{A}$ be a real s.p.d. matrix belonging to A, $\widetilde{b}$ a real vector which belongs to b. There exists a set of real vectors $\{\widetilde{p}_1, \ldots, \widetilde{p}_n\}$, $\widetilde{p}_k \in P_k$, $k = 1, \ldots, n$, such that $\widetilde{p}_i^{\mathrm{T}} \cdot (\widetilde{A} \cdot \widetilde{p}_j) = 0$, for $1 \leqslant i \neq j \leqslant n$. The real vector sequences $\widetilde{x}_k$ given by

$\widetilde{x}_0 = 0,\ \widetilde{r}_0 = \widetilde{b}$
for $k = 1, \ldots, n$
  $\widetilde{\alpha}_{k-1} = \dfrac{\widetilde{p}_k^{\mathrm{T}} \cdot \widetilde{r}_{k-1}}{\widetilde{p}_k^{\mathrm{T}} \cdot (\widetilde{A} \cdot \widetilde{p}_k)} = \dfrac{\widetilde{p}_k^{\mathrm{T}} \cdot \widetilde{b}}{\widetilde{p}_k^{\mathrm{T}} \cdot (\widetilde{A} \cdot \widetilde{p}_k)}$
  $\widetilde{x}_k = \widetilde{x}_{k-1} + \widetilde{\alpha}_{k-1} \cdot \widetilde{p}_k$
  $\widetilde{r}_k = \widetilde{b} - \widetilde{A} \cdot \widetilde{x}_k$

verify $\widetilde{x}_k \in X_k$, for $k = 1, \ldots, n$. The real vector $\widetilde{x}_n$ minimizes $\phi_{\tilde{A},\tilde{b}}(x)$. From the equality

$$\frac{\widetilde{p}_k^{\mathrm{T}} \cdot \widetilde{r}_{k-1}}{\widetilde{p}_k^{\mathrm{T}} \cdot (\widetilde{A} \cdot \widetilde{p}_k)} = \frac{\widetilde{p}_k^{\mathrm{T}} \cdot \widetilde{b}}{\widetilde{p}_k^{\mathrm{T}} \cdot (\widetilde{A} \cdot \widetilde{p}_k)}$$

and by proposition 6.5, in order to construct $\alpha_{k-1} = (P_k^{\mathrm{T}} \cdot R_{k-1})/\|P_k\|_A^2$, it will be better to take $\alpha_{k-1} = (P_k^{\mathrm{T}} \cdot b)/\|P_k\|_A^2$. □

Example 6.7.

$$A = \begin{pmatrix} [1.9900,\ 2.0100] & [0.4900,\ 0.5100] & [0.3333,\ 0.3333] \\ [0.4900,\ 0.5100] & [1.3233,\ 1.3433] & [0.2400,\ 0.2600] \\ [0.3333,\ 0.3333] & [0.2400,\ 0.2600] & [1.1900,\ 1.2100] \end{pmatrix}.$$

b is given by

$$\check{b} = \check{A} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad \text{and} \quad \varrho(b) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad b = \begin{pmatrix} [1.8333,\ 3.8333] \\ [1.0833,\ 3.0833] \\ [0.7833,\ 2.7833] \end{pmatrix}.$$

By using the algorithm given in section 4, we give a set of A-orthogonal interval directions:

$$\left\{ P_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},\ P_2 = \begin{pmatrix} [-0.2563,\ -0.2438] \\ 1 \\ 0 \end{pmatrix},\ P_3 = \begin{pmatrix} [-0.1376,\ -0.1266] \\ [-0.1499,\ -0.1263] \\ 1 \end{pmatrix} \right\}.$$

Denote by X^{CD} and X^{Ch}, the interval solution given, respectively, by interval conjugate directions and interval Cholesky method. We have the following results:

$$X^{\mathrm{CD}} = \begin{pmatrix} [\ \ 0.0456,\ 1.9318] \\ [-0.2443,\ 2.2385] \\ [-0.1862,\ 2.1804] \end{pmatrix} \quad \text{and} \quad X^{\mathrm{Ch}} = \begin{pmatrix} [-0.0318,\ 2.0288] \\ [-0.2503,\ 2.2449] \\ [-0.2293,\ 2.2274] \end{pmatrix}.$$

Σ denotes the set of solutions of the interval linear system $A \cdot x = b$. By the numerical results given above, we have $\Sigma \subseteq X^{\mathrm{CD}} \subseteq X^{\mathrm{Ch}}$.

Acknowledgement

The author thanks Professor Bernard Germain-Bonne for valuable discussions during his stay in Marrakech.

References

[1] G. Alefeld and G. Mayer, The Cholesky method for interval data, Linear Algebra Appl. 194 (1993) 161–182.
[2] G. Alefeld and G. Mayer, On the symmetric and unsymmetric solution set of interval systems, SIAM J. Matrix Anal. Appl. 16(4) (1995) 1223–1240.
[3] A. Frommer and G. Mayer, A new criterion to guarantee the feasability of the interval Gaussian algorithm, SIAM J. Matrix Anal. Appl. 14(2) (1993) 408–419.
[4] G.H. Golub and C.F. Van Loan, *Matrix Computations* (North Oxford Academic, Oxford, 1986).
[5] E. Hansen, Interval arithmetic in matrix computations, SIAM J. Ser. B 2(2) 1965.
[6] R.E. Moore, *Interval Analysis* (Prentice-Hall, Englewood Cliffs, NJ, 1966).
[7] A. Neumaier, New techniques for the analysis of linear interval equations, Linear Algebra Appl. 58 (1984) 273–325.
[8] K. Nickel, *Interval Mathematics*, Lecture Notes in Computer Science (Springer, Berlin, 1975).

[9] W. Oettli and W. Prager, Compatibility of approximate solution of linear equations with given error bounds for coefficients and right-hand sides, Numer. Math. 6 (1964) 405–409.
[10] J. Rohn, Systems of linear interval equations, Linear Algebra Appl. 126 (1989) 39–78.
[11] J. Rohn, Positive definiteness and stability of interval matrices, SIAM J. Matrix Anal. Appl. 15(1) (1994) 175–184.
[12] S.P. Shary, On optimal solution of interval linear equations, SIAM J. Numer. Anal. 32(2) (1995) 610–630.
[13] S.P. Shary, Solving the linear interval tolerance problem, Math. Comput. Simulation 39 (1995) 53–85.