INTERVAL PSEUDO-INVERSE MATRICES: COMPUTATION AND APPLICATIONS

Pavel Saraev

Lipetsk State Technical University, Lipetsk, Russia

SCAN'2012, Novosibirsk, Sep. 23-28, 2012

INTRODUCTION

Given: $A \in \mathbb{R}^{m \times n}$.

Pseudo-inverse (Moore-Penrose generalized inverse) matrix

$A^+ \in \mathbb{R}^{n \times m}$:

$$AA^+A = A;$$
$$A^+AA^+ = A^+;$$
$$(AA^+)^T = AA^+;$$
$$(A^+A)^T = A^+A$$

INTRODUCTION

Given: $A \in \mathbb{R}^{m\times n}$.

Pseudo-inverse (Moore-Penrose generalized inverse) matrix

$A^+ \in \mathbb{R}^{n\times m}$:

$$AA^+A = A;$$
$$A^+AA^+ = A^+;$$
$$(AA^+)^T = AA^+;$$
$$(A^+A)^T = A^+A$$

Advantages:

- always exists;
- unique;
- if nonsingular $A \in \mathbb{R}^{n\times n}$ $A^+ = A^{-1}$.

INTRODUCTION

Given: $A \in \mathbb{R}^{m\times n}$.

Pseudo-inverse (Moore-Penrose generalized inverse) matrix

$A^+ \in \mathbb{R}^{n\times m}$:

$$AA^+A = A;$$
$$A^+AA^+ = A^+;$$
$$(AA^+)^T = AA^+;$$
$$(A^+A)^T = A^+A$$

Advantages:

- always exists;
- unique;
- if nonsingular $A \in \mathbb{R}^{n\times n}$ $A^+ = A^{-1}$.

Disadvantage:

instability

PROBLEM STATEMENT

Given: $[A] \in \mathbb{IR}^{n\times n}$. Interval inverse matrix $[A]^{-1} \in \mathbb{IR}^{n\times n}$ such that $[A]^{-1} \supset \left\{A^{-1} : A \in [A]\right\}$

- G. Alefeld, J. Herzberger, *Introduction to Interval Computations*, Academic Press, New York, 1983.
- J. Rohn, Inverse Interval Matrix, *SIAM J. Num. Anal.*, V. 3 (1993), N 3, pp. 864–870.
- T. Nirmala, D. Datta, H.S. Kushwaha, K. Ganesan, Inverse Interval Matrix: A New Approach, *App. Math. Sci.*, V. 5 (2011), N 13, pp. 607–624.

PROBLEM STATEMENT

Given: $[A] \in \mathbb{IR}^{n\times n}$. Interval inverse matrix $[A]^{-1} \in \mathbb{IR}^{n\times n}$ such that $[A]^{-1} \supset \{A^{-1} : A \in [A]\}$

- G. Alefeld, J. Herzberger, *Introduction to Interval Computations*, Academic Press, New York, 1983.
- J. Rohn, Inverse Interval Matrix, *SIAM J. Num. Anal.*, V. 3 (1993), N 3, pp. 864–870.
- T. Nirmala, D. Datta, H.S. Kushwaha, K. Ganesan, Inverse Interval Matrix: A New Approach, *App. Math. Sci.*, V. 5 (2011), N 13, pp. 607–624.

Definition

Given: $[A] \in \mathbb{IR}^{m\times n}$. Interval pseudo-inverse matrix $[A]^{+} \in \mathbb{IR}^{n\times m}$ is the minimal interval matrix such that $[A]^{+} \supset \{A^{+} : A \in [A]\}$

$[A]^{+}$ is the extension of interval inverse matrix

Problem

Development of algorithm for computation of enclosure for $[A]^{+}$

PROPOSED ALGORITHM I

Interval Greville algorithm

Given

$[A] \in \mathbb{IR}^{m\times n}$

Notation

$[a_k]$ is k-th column of $[A]$, $k = 1, \ldots, n$.
$[A_k]$ are first k columns of $[A]$: $[A_k] = \big([a_1] \quad [a_2] \quad \ldots \quad [a_k]\big)$.
So,
$k = 1$: $[A_1] = [a_1]$
$k = 2, \ldots, n$: $[A_k] = \big([A_{k-1}] \quad [a_k]\big)$.

PROPOSED ALGORITHM II

Step 1

Let $k = 1$. Assume $[d_1] = \|[a_1]\|^2 = \sum_{i=1}^{m} [a_{i1}]^2$.

$$[A_1]^+ = \begin{cases} [0], & \text{if } \overline{[d_1]} = 0, \\ \dfrac{[a_1]^T}{[d_1]}, & \text{if } \underline{[d_1]} > 0, \\ [0] \cup \dfrac{[a_1]^T}{[d_1]}, & \text{else}, \end{cases}$$

where $[0] \in \mathbb{IR}^m$ is the null interval vector, $\cup$ is the interval hull of union of interval vectors.

PROPOSED ALGORITHM III

Steps $2, \ldots, n$

$$[A_k]^+ = \begin{bmatrix} [A_{k-1}]^+(I - [a_k][f_k]) \\ [f_k] \end{bmatrix},$$

where I is the unitary matrix of the order m.
Let

$$[c_k] = (I - [A_{k-1}][A_{k-1}]^+)[a_k], \quad [d_k] = \|c_k\|^2,$$

$$[f_k] = \begin{cases} \dfrac{[c_k]^T}{[d_k]}, & \text{if } \underline{[d_k]} > 0, \\ \dfrac{[a_k]^T([A_{k-1}]^+)^T[A_{k-1}]^+}{1 + \|[A_{k-1}]^+[a_k]\|^2}, & \text{if } \overline{[d_k]} = 0, \\ \dfrac{[c_k]^T}{[d_k]} \cup \dfrac{[a_k]^T([A_{k-1}]^+)^T[A_{k-1}]^+}{1 + \|[A_{k-1}]^+[a_k]\|^2}, & \text{else.} \end{cases}$$

DISCUSSION OF THE ALGORITHM

- Proposed algorithm is extension of traditional Greville algorithm (if $\underline{[d_k]} = \overline{[d_k]}$).
- Pseudo-inversion is monotone operation: if $[A] \subset [B]$ then $[A]^+ \subset [B]^+$.
- It can be used union of finite number of interval matrices if software has this feature.
- Algorithm can obtain large overestimations (can contain infinite bounds) in some cases.
- Result can be interval matrix even for real matrix A.

DISCUSSION OF THE ALGORITHM

- Proposed algorithm is extension of traditional Greville algorithm (if $\underline{[d_k]} = \overline{[d_k]}$).
- Pseudo-inversion is monotone operation: if $[A] \subset [B]$ then $[A]^+ \subset [B]^+$.
- It can be used union of finite number of interval matrices if software has this feature.
- Algorithm can obtain large overestimations (can contain infinite bounds) in some cases.
- Result can be interval matrix even for real matrix A.

Accuracy criterion:

$$[t] = \|[A][A]^+[A] - [A]\| + \|[A]^+[A][A]^+ - [A]^+\| +$$
$$+\|([A][A]^+)^T - [A][A]^+\| + \|([A]^+[A])^T - [A]^+[A]\|.$$

Also can be used:

$$\mathrm{wid}\ \left([A]^+\right)$$

EXAMPLES I

Ex. 1. Real non-zero number

$$[A] = -2.5 \in \mathbb{R};$$

$$[A]^{+} = [-0.4; -0.4]$$

Ex. 2. Real zero number

$$[A] = 0 \in \mathbb{R};$$

$$[A]^{+} = [0; 0]$$

EXAMPLES II

Ex. 3. Interval number non-containing zero

$$[A] = [1; 2] \in \mathbb{IR};$$

$$[A]^{+} = [0,25; 2],$$

but true result is $[A]^{+} = [0,5; 1]$

Ex. 4. Interval number containing zero

$$[A] = [-1; 4] \in \mathbb{IR};$$

$$[A]^{+} = [\infty; +\infty],$$

which is interval hull of union $[-\infty; 1] \cup 0 \cup [0,25; +\infty]$

EXAMPLES III

Ex. 5. Real matrix

$$[A] = \begin{pmatrix} 1 & 3 \\ 0 & 0 \\ 1 & 3 \end{pmatrix} \in \mathbb{R}^{3\times 2};$$

$$[A]^+ = \begin{pmatrix} [0.049..9; 0.050..01] & [0;0] & [0.049..9; 0.050..01] \\ [0.15; 0.15] & [0;0] & [0.15; 0.15] \end{pmatrix},$$

$$[t] \approx [0; 6.334 \cdot 10^{-30}],$$

$$\mathrm{wid}\ \left([A]^+\right) \approx 1.43 \cdot 10^{-16},$$

$$[A]^+ \supset A^+ = \begin{pmatrix} 0.05 & 0 & 0.05 \\ 0.15 & 0 & 0.15 \end{pmatrix}$$

EXAMPLES IV

Ex. 6. Interval matrix

$$[A] = \begin{pmatrix} [1;3] & [-1;0] & [2;3] \\ [2;3] & [2;3] & [2;3] \end{pmatrix} \in \mathbb{IR}^{2\times 3};$$

$$\mathrm{wid}\ ([A]) = 2$$

$$[A]^{+} \approx \begin{pmatrix} [-7.45; 12.61] & [-3.04; 1.74] \\ [-9.47; 1.35] & [0.00; 2.58] \\ [-6.79; 12.50] & [-3.04; 1.55] \end{pmatrix},$$

$$[t] \approx [0; 4736161],$$

$$\mathrm{wid}\ \left([A]^{+}\right) \approx 20.06,$$

EXAMPLES V

Ex. 7. “Bad” real matrix

H. Minakuchi, H. Kai, K. Shirayanagi, M-T. Noday, Algorithm stabilization techniques and their application to symbolic computation of generalized inverses, *Electronic Proc. of the IMACS Conference on Applications of Computer Algebra (IMACS-ACA'97), 1997.*- http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.137.8604

$$[A] = \begin{pmatrix} 3 & \frac{211}{3} & \frac{15}{7} \\ 3 & 70.3333 & 2.14286 \end{pmatrix} \in \mathbb{R}^{2\times 3};$$

Specific of the matrix: $\frac{211}{3} \approx 70.3333$, $\frac{15}{7} \approx 2.14286$
For interval Greville algorithm applying: $\frac{211}{3} \subset [70.33..3; 70.33..4]$, $\frac{15}{7} \subset [2.142857..142857; 2.142857..142858]$

$$[A]^+ \approx \begin{pmatrix} [-\infty; +\infty] & [-\infty; +\infty] \\ [-\infty; +\infty] & [-\infty; +\infty] \\ [-\infty; +\infty] & [-\infty; +\infty] \end{pmatrix},$$

$$[t] \approx [0; +\infty], \quad \mathrm{wid}\ \left([A]^+\right) = +\infty$$

POSSIBLE APPLICATIONS I

1. Real case: unstable operation detection

$$A \in \mathbb{R}^{m\times n} \to [A_\varepsilon] \in \mathbb{IR}^{m\times n},$$

where $[A_\varepsilon]$ is epsilon-extension of A
$\overline{[t]}$ or $\mathrm{wid}\ ([A_\varepsilon]^+)$ is large – operation is unstable
$\overline{[t]}$ or $\mathrm{wid}\ ([A_\varepsilon]^+)$ is small – $A^+ \approx \mathrm{mid}\ ([A_\varepsilon]^+)$

POSSIBLE APPLICATIONS II

2. Linear least squares problem: analytic solution

Given: $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$

$$Ax = b \to \|Ax - b\|_2^2 \to \min$$

with respect to $x \in \mathbb{R}^n$
Normal pseudo-solution (solution of linear LSP):

$$x = A^+ b$$

Interval case ($[A][x] = [b]$):

$$[x] = [A]^+ [b]$$

POSSIBLE APPLICATIONS III

3. Guaranteed nonlinear least squares problem with separable variables

G.H. Golub, V. Pereyra. The Differentiation of Pseudo-Inverses and Nonlinear Least Squares Problems Whose Variables Separate *SIAM J. Num. Anal., 1973.– Vol. 10.– pp. 413–432*

S.L. Blyumin, P.V. Saraev. Reduction of adjusting weights space dimension in feedforward artificial neural networks training. *Proc. of IEEE Int. Conf. on Artificial Intelligence Syst.– 2002.– P. 242–247.*

Let u and v are linear and nonlinear parameter vectors (for example, feed-forward NN); y is known responses vector; $\Psi(v)$ is matrix of basis functions Linear-nonlinear equation:

$$u = \Psi(v)^+ y$$

Interval pseudo-inverse can be used to optimize modified function:

$$J(u,v) = \|\Psi(v)u - y\|_2^2 \to \min \to \hat{J}(v) = \|\Psi(v)\Psi(v)^+ y - y\|_2^2 \to \min$$

Interval case:

$$[u] = [\Psi([v])]^+ y$$

$$\hat{J}([v]) = \|\Psi([v])\Psi([v])^+ y - y\|_2^2 \to \min$$

CONCLUSION

Results

- Interval Greville algorithm for estimation of interval pseudo-inverse matrices is proposed and investigated
- Work of the algorithm is showed at numerical examples, some features of the algorithm are stated
- Possible applications of interval pseudo-inversion to detection of unstable operation and optimzation problems are showed

THANK YOU FOR ATTENTION!

Pavel Saraev, LSTU (psaraev@yandex.ru)