Institute of Computer Science
Academy of Sciences of the Czech Republic

An Improvement of the Bauer-Skeel Bounds

Jiří Rohn

Technical report No. V-1065

22.02.2010

Pod Vodárenskou věží 2, 182 07 Prague 8, phone: +420 266 051 111, fax: +420 286 585 789, e-mail:e-mail:ics@cs.cas.cz

Institute of Computer Science
Academy of Sciences of the Czech Republic

An Improvement of the Bauer-Skeel Bounds

Jiří Rohn[1]

Technical report No. V-1065

22.02.2010

Abstract:

It is shown that the Hansen-Bliek-Rohn bounds are never worse, and "almost always" better, than the classical Bauer-Skeel bounds. Formulae for overestimation of the Hansen-Bliek-Rohn bounds are given.

Keywords:
Bauer-Skeel bounds, Hansen-Bliek-Rohn bounds, comparison, overestimation.

[1]Supported by the Czech Republic Grant Agency under grant 201/01/0343, and by the Institutional Research Plan AV0Z10300504.

1 Introduction

This is transcription of slides of a talk I delivered at the Bergische Universität Wuppertal, Germany on Dec. 9, 2003. It has never been made a regular paper. So after six years, I decided to make it at least a report written in a terse "slide-like" style. Theorems 3 through 6 are new.

2 Notations

- $A \leq B$,
- $|A|$,
- $\min\{A, B\}$, $\max\{A, B\}$

understood componentwise.
Especially,

$$|A - B| \leq C$$

is equivalent to

$$B - C \leq A \leq B + C.$$

3 The problem

Given a fixed system

$$A_c x_c = b_c$$

with A_c nonsingular, and a perturbed system

$$Ax = b$$

such that

$$\begin{aligned} |A - A_c| &\leq \Delta, \\ |b - b_c| &\leq \delta, \end{aligned}$$

estimate

$$\ldots \leq x \leq \ldots$$

in terms of x_c, A_c^{-1}, Δ and δ.

4 Assumption

We shall assume throughout that the data satisfy

$$\varrho(|A_c^{-1}|\Delta) < 1.$$

Under this *spectral condition* we have

$$M := (I - |A_c^{-1}|\Delta)^{-1} = \sum_{j=0}^{\infty}(|A_c^{-1}|\Delta)^j \geq I \geq 0.$$

In particular,

$$M_{ii} \geq 1$$

for each i (a property which will turn out extremely important).

5 The Bauer-Skeel bounds

Theorem 1. (Bauer 1966, Skeel 1979) *If*

$$\varrho(|A_c^{-1}|\Delta) < 1,$$

then for each A, b such that $|A - A_c| \leq \Delta$ and $|b - b_c| \leq \delta$, A is nonsingular and the solution of

$$Ax = b$$

satisfies

$$-x^* + x_c + |x_c| \leq x \leq x^* + x_c - |x_c|,$$

where

$$\begin{aligned} M &= (I - |A_c^{-1}|\Delta)^{-1}, \\ x^* &= M(|x_c| + |A_c^{-1}|\delta). \end{aligned}$$

Note. Usually presented as $|x - x_c| \leq x^* - |x_c|$, with $\delta = 0$ or in normwise setting. Two inversions needed.

6 Proof

We have

$$A_c^{-1}A = I - A_c^{-1}(A_c - A),$$

where

$$\varrho(A_c^{-1}(A_c - A)) \leq \varrho(|A_c^{-1}(A_c - A)|) \leq \varrho(|A_c^{-1}|\Delta) < 1,$$

hence $A_c^{-1}A$ is nonsingular and thus also A. If $Ax = b$, then

$$\begin{aligned} |x - x_c| &= |A_c^{-1}A_c(x - x_c)| \\ &\le |A_c^{-1}| \cdot |(A_c - A)x + (b - b_c)| \\ &\le |A_c^{-1}|(\Delta|x| + \delta). \end{aligned}$$

[Attention: This is the bifurcation point of the two proofs.]

$$\begin{aligned} |x - x_c| &\le |A_c^{-1}|(\Delta|x| + \delta) = |A_c^{-1}|(\Delta|x - x_c + x_c| + \delta) \\ &\le |A_c^{-1}|\Delta|x - x_c| + |A_c^{-1}|(\Delta|x_c| + \delta), \end{aligned}$$

hence

$$(I - |A_c^{-1}|\Delta)|x - x_c| \le |A_c^{-1}|(\Delta|x_c| + \delta).$$

Premultiplying by $M = (I - |A_c^{-1}|\Delta)^{-1} \ge 0$:

$$\begin{aligned} |x - x_c| &\le M|A_c^{-1}|(\Delta|x_c| + \delta) \\ &= (M - I)|x_c| + M|A_c^{-1}|\delta \\ &= x^* - |x_c| \end{aligned}$$

and equivalently

$$-x^* + x_c + |x_c| \le x \le x^* + x_c - |x_c|. \qquad \square$$

7 The HBR bounds

Theorem 2. (Hansen 1992, Bliek 1992, R. 1993) *Under the same assumption*

$$\varrho(|A_c^{-1}|\Delta) < 1$$

we have

$$\min\{\underset{\sim}{x}, T\underset{\sim}{x}\} \le x \le \max\{\tilde{x}, T\tilde{x}\},$$

where

$$\begin{aligned} M &= (I - |A_c^{-1}|\Delta)^{-1}, \\ D &= \mathrm{diag}(M_{11}, \ldots, M_{nn}), \\ T &= (2D - I)^{-1}, \\ x^* &= M(|x_c| + |A_c^{-1}|\delta), \\ \underset{\sim}{x} &= -x^* + D(x_c + |x_c|), \\ \tilde{x} &= x^* + D(x_c - |x_c|). \end{aligned}$$

8 Proof

As in the proof of the Bauer-Skeel bounds we proceed up to the "bifurcation point"

$$|x - x_c| \le |A_c^{-1}|(\Delta|x| + \delta),$$

but then we continue in another way: we have on one hand

$$x - x_c \le |x - x_c| \le |A_c^{-1}|(\Delta|x| + \delta) \tag{8.1}$$

and on the other hand

$$|x| - |x_c| \le |x - x_c| \le |A_c^{-1}|(\Delta|x| + \delta). \tag{8.2}$$

For i fixed, take the ith inequality from (8.1) and for $j \ne i$ from (8.2):

$$x_i \le (x_c)_i + (|A_c^{-1}|(\Delta|x| + \delta))_i$$

$$|x_j| \le |x_c|_j + (|A_c^{-1}|(\Delta|x| + \delta))_j, \quad j \ne i.$$

Since $x_i = |x_i| + (x_i - |x_i|)$ and the same holds for $(x_c)_i$, we can put them together as

$$|x| + (x_i - |x_i|)e_i \le |x_c| + ((x_c)_i - |x_c|_i)e_i + |A_c^{-1}|(\Delta|x| + \delta),$$

which implies

$$(I - |A_c^{-1}|\Delta)|x| + (x_i - |x_i|)e_i \le |x_c| + ((x_c)_i - |x_c|_i)e_i + |A_c^{-1}|\delta.$$

Again premultiplying by $M = (I - |A_c^{-1}|\Delta)^{-1} \ge 0$:

$$|x| + (x_i - |x_i|)Me_i \le x^* + ((x_c)_i - |x_c|_i)Me_i$$

and taking the ith inequality we get

$$|x_i| + (x_i - |x_i|)M_{ii} \le x_i^* + ((x_c)_i - |x_c|_i)M_{ii} = \tilde{x}_i,$$

an inequality containing x_i only. If $x_i \ge 0$, then this inequality becomes

$$x_i \le \tilde{x}_i,$$

and if $x_i < 0$, then it turns into

$$x_i \le \tilde{x}_i/(2M_{ii} - 1) = T_{ii}\tilde{x}_i,$$

in both cases

$$x_i \le \max\{\tilde{x}_i, T_{ii}\tilde{x}_i\}.$$

Since i was arbitrary, we conclude that

$$x \le \max\{\tilde{x}, T\tilde{x}\},$$

which is the upper bound. Similarly for the lower one. □

9 Comparison: preliminaries

For comparison, denote the Bauer-Skeel bounds by

$$\underline{x} \leq x \leq \overline{x}$$

and the HBR bounds by

$$\underline{\underline{x}} \leq x \leq \overline{\overline{x}},$$

i.e.

$$\begin{aligned}
\underline{x} &= -x^* + x_c + |x_c|, \\
\overline{x} &= x^* + x_c - |x_c|, \\
\underline{\underline{x}} &= \min\{\underset{\sim}{x}, T\underset{\sim}{x}\}, \\
\overline{\overline{x}} &= \max\{\tilde{x}, T\tilde{x}\}.
\end{aligned}$$

It turns out that crucial for the comparison is the fact that

$$M_{ii} \geq 1 \text{ for each } i.$$

10 Main result

Theorem 3. *Under the common assumption $\varrho(|A_c^{-1}|\Delta) < 1$, for each i we have*

$$\overline{x}_i - \overline{\overline{x}}_i \geq \min\left\{(M_{ii}-1)(|x_c|_i - (x_c)_i),\ \tfrac{2(M_{ii}-1)}{2M_{ii}-1}(x_i^* - |x_c|_i)\right\} \geq 0,$$

$$\underline{\underline{x}}_i - \underline{x}_i \geq \min\left\{(M_{ii}-1)(|x_c|_i + (x_c)_i),\ \tfrac{2(M_{ii}-1)}{2M_{ii}-1}(x_i^* - |x_c|_i)\right\} \geq 0.$$

In particular,

$$\underline{x} \leq \underline{\underline{x}} \leq \overline{\overline{x}} \leq \overline{x},$$

i.e. the HBR bounds are never worse than the Bauer-Skeel bounds.

Remark. Nonnegativity follows from the facts that $M \geq I$ and $x^* = M(|x_c| + |A_c^{-1}|\delta) \geq |x_c|$.

11 Refinement

Theorem 4. *Let the spectral condition hold. Then for each i such that $M_{ii} > 1$ and $(x_c)_i \neq 0$ we have*

$$(\overline{x}_i - \underline{x}_i) - (\overline{\overline{x}}_i - \underline{\underline{x}}_i) \geq \frac{2(M_{ii}-1)^2}{2M_{ii}-1}|x_c|_i > 0,$$

hence

$$\overline{\overline{x}}_i - \underline{\underline{x}}_i < \overline{x}_i - \underline{x}_i,$$

i.e., the ith HBR bound is better than the Bauer-Skeel bound.

Remark. Recall that $M = (I - |A_c^{-1}|\Delta)^{-1} = \sum_{j=0}^{\infty}(|A_c^{-1}|\Delta)^j \geq I$. Hence $M_{ii} > 1$ e.g. if $(|A_c^{-1}|\Delta)_{ii} > 0$.

12 Partial conclusion

We can conclude that the HBR bounds are "almost always" better than the Bauer-Skeel bounds. Still, how good are the HBR bounds themselves?

13 Exact bounds

For each i define

$$\begin{aligned} x_i^e &= \min\{x_i\,;\; Ax = b,\ |A - A_c| \le \Delta,\ |b - b_c| \le \delta\}, \\ x_i^E &= \max\{x_i\,;\; Ax = b,\ |A - A_c| \le \Delta,\ |b - b_c| \le \delta\}. \end{aligned}$$

Obviously, x^e and x^E are exact componentwise bounds, so that they satisfy

$$\underline{\underline{x}} \le x^e \le x^E \le \overline{\overline{x}}$$

(x^e, x^E are NP-hard to compute). Now, what is the amount of overestimation?

14 Overestimation of the HBR bounds

Theorem 5. (2000, not yet published) *Let the spectral condition hold. Then for each $i \in \{1, \dots, n\}$ we have*

$$\begin{aligned} \underline{\underline{x}}_i &\le x_i^e \le \underline{\underline{x}}_i + \underline{d}_i, \\ \overline{\overline{x}}_i - \overline{d}_i &\le x_i^E \le \overline{\overline{x}}_i, \end{aligned}$$

where

$$\begin{aligned} \underline{d}_i &= (M|(\mathrm{diag}(\underline{z})A_c^{-1}\mathrm{diag}(\underline{z}) - |A_c^{-1}|)(\underline{\xi}_i \Delta M e_i + \Delta x^* + \delta)|)_i, \\ \overline{d}_i &= (M|(\mathrm{diag}(\overline{z})A_c^{-1}\mathrm{diag}(\overline{z}) - |A_c^{-1}|)(\overline{\xi}_i \Delta M e_i + \Delta x^* + \delta)|)_i, \\ \underline{\xi}_i &= (|\underline{\underline{x}}| + \underline{\underline{x}} - x_c - |x_c|)_i, \\ \overline{\xi}_i &= (|\overline{\overline{x}}| - \overline{\overline{x}} + x_c - |x_c|)_i \end{aligned}$$

and $\underline{z}$, $\overline{z}$ are given by

$$\underline{z}_j = \begin{cases} \mathrm{sgn}\,(x_c)_j & \text{if } j \ne i, \\ -1 & \text{if } j = i, \end{cases}, \qquad \overline{z}_j = \begin{cases} \mathrm{sgn}\,(x_c)_j & \text{if } j \ne i, \\ 1 & \text{if } j = i \end{cases}, \qquad (j = 1, \dots, n).$$

15 Example (J. Albrecht 1961)

Here $A_c x = b_c$ reads

$$\begin{aligned} 4.33x_1 - 1.12x_2 - 1.08x_3 + 1.14x_4 &= 3.52 \\ -1.12x_1 + 4.33x_2 + 0.24x_3 - 1.22x_4 &= 1.57 \\ -1.08x_1 + 0.24x_2 + 7.21x_3 - 3.22x_4 &= 0.54 \\ 1.14x_1 - 1.22x_2 - 3.22x_3 + 5.43x_4 &= -1.09 \end{aligned}$$

and

$$\Delta_{ij} = \delta_i = 0.005$$

for each i, j,

$$\varrho(|A_c^{-1}|\Delta) = 0.008.$$

16 Results

(rounded to four decimal digits)

$$[\underline{\underline{x}}, \underline{\underline{x}} + \underline{d}] = \begin{pmatrix} [1.0408, 1.0441] \\ [0.5567, 0.5593] \\ [0.1056, 0.1072] \\ [-0.2352, -0.2299] \end{pmatrix}, \quad [\overline{\overline{x}} - \overline{d}, \overline{\overline{x}}] = \begin{pmatrix} [1.0517, 1.0517] \\ [0.5670, 0.5689] \\ [0.1129, 0.1164] \\ [-0.2218, -0.2210] \end{pmatrix}.$$

17 Unsatisfactory result (1997)

Theorem 5 applied to the system

$$\begin{pmatrix} \varepsilon^2 & [-\varepsilon, \varepsilon] & [-\varepsilon, \varepsilon] & [-\varepsilon, \varepsilon] \\ 0 & 1.1 & 1 & 1 \\ 0 & 1 & 1.1 & 1 \\ 0 & 1 & 1 & 1.1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ [-\varepsilon, \varepsilon] \\ [-\varepsilon, \varepsilon] \\ [-\varepsilon, \varepsilon] \end{pmatrix}$$

works for each $\varepsilon > 0$ (since $\varrho(|A_c^{-1}|\Delta) = 0$) and yields independently of ε

$$[\overline{\overline{x}}_1 - \overline{d}_1, \overline{\overline{x}}_1] = \left[\frac{30}{31}, \frac{1230}{31}\right] = [0.97, 39.68]$$

whereas $x_1^E = \frac{830}{31} = 26.77$, i.e., $\frac{\overline{\overline{x}}_1 - x_1^E}{x_1^E} = \frac{40}{83} = 0.48$ (rounded to two decimal digits).

18 Zero overestimation cases

Theorem 6. *Let the spectral condition hold. Then we have:*

(i) $x^e = \underline{\underline{x}}$, $x^E = \overline{\overline{x}}$ *if* A_c *is a diagonal matrix with positive diagonal entries,*

(ii) $x^e = \underline{\underline{x}}$ *if* $A_c^{-1} \geq 0$ *and* $A_c^{-1} b_c \leq 0$,

(iii) $x^E = \overline{\overline{x}}$ *if* $A_c^{-1} \geq 0$ *and* $A_c^{-1} b_c \geq 0$.

19 Conclusions

- both the Bauer-Skeel bounds and the HBR bounds require solving $A_c x = b_c$ and computing A_c^{-1} and $(I - |A_c^{-1}|\Delta)^{-1}$,
- the HBR bounds are never worse, and "almost always" better, than the Bauer-Skeel bounds,
- overestimation of the HBR bounds can be computed at almost no additional cost.

References

[1] F. L. Bauer. Genauigkeitsfragen bei der Lösung linearer Gleichungssysteme. *ZAMM*, 46:409–421, 1966.

[2] R. D. Skeel. Scaling for numerical stability in Gaussian elimination. *Journal of the ACM*, 26:494–526, 1979.

[3] E. R. Hansen. Bounding the solution of interval linear equations. *SIAM Journal on Numerical Analysis*, 29:1493–1503, 1992.

[4] C. Bliek. Computer Methods for Design Automation. PhD thesis, MIT, Cambridge, MA, July 1992.

[5] J. Rohn. Cheap and tight bounds: The recent result by E. Hansen can be made more efficient. *Interval Computations*, 4:13–21, 1993.