M. SOMBRA, 2020

NLA III : Error analysis of GEPP

To control the error produced by GEPP, we want to apply the two steps:

(1) Analyse round offs to show that the matrix

$$\hat{A}_{GEPP} = A + \delta_{GEPP} A \quad \text{produced by GEPP}$$

has a small(*) relative error (<u>backward analysis</u>)

(2) Apply perturbation theory (condition numbers) to bound the error in the computation of x_{GEPP}:

$$A_{GEPP} \cdot x_{GEPP} = b$$

(*) What does "small" mean in this context?

Let ε be the machine epsilon & $\| \ \|$ a "standard" norm (like $\|\cdot\|_\infty$, $\|\cdot\|_1$, etc)

Rounding off the entries of A gives $\hat{A} = A + \delta A$ with

$$\frac{\|\delta A\|}{\|A\|} < \varepsilon$$

By perturbation theory, this error will be amplified to

$$\frac{\|\delta x\|}{\|x\|} \leq \kappa_{\|\cdot\|}(A) \cdot \varepsilon$$

To keep the quality of this bound, we want

$$\frac{\|\delta_{GEPP} A\|}{\|A\|} \leq C \cdot \varepsilon$$

with C as small as <u>possible</u>

To this end, we have to be careful about <u>pivoting</u>

<u>III. 1</u> <u>The need of pivoting</u> [D, §2.4.1]

We apply LU factorization without pivoting to

$$A = \begin{pmatrix} \eta & 1 \\ 1 & 1 \end{pmatrix}$$

with η a power of the base β that is smaller than ε. In the book

$$\beta = 10 \,, \qquad \varepsilon = 0.5 \times 10^{-3} \,, \qquad \eta = 10^{-4}$$

Hence

$$1 \oplus \eta = fl(1+\eta) = 1$$

η is "lost" when added to 1

Set $\qquad A = LU = \begin{pmatrix} 1 & 0 \\ \eta^{-1} & 1 \end{pmatrix} \cdot \begin{pmatrix} \eta & 1 \\ 0 & 1-\eta^{-1} \end{pmatrix}$

Gauss elimination without pivoting

Then $\qquad L_{GENP} = \begin{pmatrix} 1 & 0 \\ \eta^{-1} & 1 \end{pmatrix} = L$

<u>but</u> $\qquad U_{GENP} = \begin{pmatrix} \eta & 1 \\ 0 & \eta^{-1} \end{pmatrix}$

and $\qquad A_{GENP} = L_{GEPP} \cdot U_{GEPP} = \begin{pmatrix} \eta & 1 \\ 1 & 0 \end{pmatrix}$ not close to A!

$$\frac{\| \delta A_{GENP} \|_\infty}{\| A \|_\infty} = \frac{\left\| \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} \right\|_\infty}{\left\| \begin{pmatrix} \eta & 1 \\ 1 & 1 \end{pmatrix} \right\|_\infty} = \frac{1}{2}$$

(and <u>not</u> $< C \cdot \varepsilon$)

Hence GE without pivoting is <u>not</u> backward stable.

This is reflected in the loss of precision when applying this to linear equation solving: the equation

$$A\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix}1\\2\end{pmatrix}$$

has a correct answer close to $\begin{pmatrix}x_1\\x_2\end{pmatrix} \approx \begin{pmatrix}1\\1\end{pmatrix}$

Solving

$$L_{GEWP}\begin{pmatrix}y_1\\y_2\end{pmatrix} = \begin{pmatrix}1\\2\end{pmatrix}$$

gives $y_1 = 1$ and $y_2 = 2 \ominus \eta^{-1} = -\eta^{-1}$

Then

$$U_{GEWP}\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix}1\\-\eta^{-1}\end{pmatrix}$$

gives $x_2 = -\eta^{-1}/_{-\eta^{-1}} = 1$ and $x_1 = \frac{1 \ominus 1}{1 \ominus \eta} = 0$

Hence $x_{GEWP} = \begin{pmatrix}1\\0\end{pmatrix}$, not close to $x \approx \begin{pmatrix}1\\1\end{pmatrix}$

The instability is also reflected in the disparity between the condition ~~numbers~~ numbers of A and of L and L

$\|A\|_\infty \approx 4$ well-conditioned

$\|L\|_\infty, \|U\|_\infty \approx \eta^{-2}$ ill-conditioned

III.2 Formal error analysis of GEPP [D, §2.4.2]

When the intermediate quantities are too large, the information in A can be easily lost.
For simplicity, suppose that A is already pivoted.
Then studying how L and U are constructed and analysing round offs,

$$A_{GEPP} = L \cdot U + E$$

machine epsilon

with $|E| \leqslant n \cdot \varepsilon \, |L| \cdot |U|$

matrices of absolute values

see [D, pages 47–48 for details]

Hence

$$\|A - A_{GEPP}\|_\infty \leqslant n \cdot \varepsilon \cdot \| |L| \|_\infty \cdot \| |U| \|_\infty$$
$$\leqslant n^3 \cdot \varepsilon \cdot g_{GEPP} \cdot \|A\|_\infty$$

where

$$g_{GEPP} = \frac{\max |u_{ij}|}{\max |a_{ij}|} \quad \text{pivot growth factor,}$$

because $|l_{ij}| \leq 1$ and so $\|L\|_\infty \leqslant n$

and $|u_{ij}| \leqslant g_{GEPP} \|A\|_\infty$ and so $\|U\|_\infty \leqslant n \, g_{GEPP} \|A\|_\infty$

Thus

$$\boxed{\frac{\|\delta_{GEPP} A\|_\infty}{\|A\|_\infty} \leqslant n^3 \cdot \varepsilon \cdot g_{GEPP}} \quad (*)$$

III.2 Formal error analysis in GEPP [D, § 2.4.2]

When the intermediate quantities are too large, the information in A can be easily lost.

To make the analysis, suppose that A is already pivoted. Studying how L and U are constructed ~~xxxxx~~ we obtain that

$$A_{GEPP} = L \cdot U + E$$

with $|E| \leq n \cdot \varepsilon \cdot (|L| \cdot |U|)$, see [D, pages 47–48] for details

matrix of absolute values

Hence

$$\boxed{\| A - A_{GEPP} \|_\infty \leq n \cdot \varepsilon \cdot \| |L| \cdot |U|}$$

In general $g_{GEPP} \leq 2^{n-1}$ and this bound can be attained:

$$A = \begin{pmatrix} 1 & & & & 1 \\ & \ddots & & 0 & \vdots \\ & & \ddots & & \vdots \\ & -1 & & \ddots & \vdots \\ & & & & 1 \end{pmatrix} = \begin{pmatrix} 1 & & & \\ & \ddots & & 0 \\ & & \ddots & \\ -1 & & & 1 \end{pmatrix} \begin{pmatrix} 1 & & & & 1 \\ & \ddots & & & 2 \\ & & \ddots & & 4 \\ & 0 & & \ddots & \vdots \\ & & & 1 & 2^{n-2} \\ & & & & 2^{n-1} \end{pmatrix}$$

<u>Ex</u>: $\begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 4 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & -1 & 1 & 0 \\ -1 & -1 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & -1 & 1 & 0 \\ -1 & -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 8 \end{pmatrix}$

The bound (*) is too pesimistic in practice, since typically

$$\| |L| \|_\infty \, \| |U| \|_\infty \approx \|A\|_\infty$$

If this is the case

$$\frac{\|\delta_{GEPP} A\|}{\|A\|} \lesssim n \, \varepsilon$$

and GEPP would be ~~stable~~ stable.
We say that GEPP is backward stable in practice (whatever that means!)