Verified error bounds for multiple roots of systems of nonlinear equations

Stef Graillat

LIP6/PEQUAN, Sorbonne Universités, UPMC Univ Paris 06, CNRS

Joint work with Siegfried M. Rump

Journées méthodes de subdivisions pour les systèmes singuliers IRCCyN, Nantes, December 15-16, 2014







General motivations: self-validating methods

Verify assumptions of mathematical theorems on the computer

- Making mathematical proofs with computers
- Getting verified results :
 - \rightarrow an interval enclosure of the true result
 - \rightarrow an approximate result with a rigorous error bound
- Possibly with proof of uniqueness
- Being fast and accurate
- Dealing with "ill-posed problems"

Proofs with computers: how to do that ?

- with computer algebra systems: exact results but sometimes not efficient
- with floating-numbers: fast but often wrong results due to rounding errors

Possible solution: computing with floating-point but taking into account all the rounding errors !

Principle of self-validating methods

2 Multiple roots of polynomial systems



Principle of self-validating methods

2 Multiple roots of polynomial systems

3 Numerical experiments

Proving that a matrix is nonsingular

Theorem 1

Let A be a matrix and R another matrix such that ||I - RA|| < 1. Then A is nonsingular

Proof.

By contrapositive, if A is singular, there exists $x \neq 0$ such that Ax = 0. Then (I - RA)x = x and so $||I - RA|| \ge 1$.

On a computer, choose for $R \approx A^{-1}$ and then compute ||I - RA|| with interval arithmetic.

Let A be a matrix of dimension n

R = inv(A) C = eye(n) - R*intval(A) nonsingular = (norm(C,1) < 1)</pre>

If nonsingular = 1, then A is nonsingular. If nonsingular = 0, then we can say nothing Let $f : \mathbb{R}^n \to \mathbb{R}^n$ and $\widehat{x} \in \mathbb{R}^n$ unknown such that $f(\widehat{x}) = 0$

```
Let \widetilde{x} \approx \widehat{x} such that f(\widetilde{x}) \approx 0
```

Find a bound for \widetilde{x} : an interval X such that $\widehat{x} \in X$

We have

$$f(x) = 0 \qquad \Leftrightarrow \qquad g(x) = x$$

with $g(x) \coloneqq x - Rf(x)$ with $det(R) \neq 0$.

Theorem 2 (Brouwer, 1912)

Every continuous function from a closed ball of a Euclidean space to itself has a fixed point.

By Brouwer fixed point theorem,

 $X \in \mathbb{IR}^n$, $g(X) \subseteq X \Rightarrow \exists \widehat{x} \in X, g(\widehat{x}) = \widehat{x} \Rightarrow f(\widehat{x}) = 0$

We just have to check $g(X) \subseteq X$ and prove $det(R) \neq 0$.

But naive approach fails:

 $g(X) \subseteq X - Rf(X) \notin X$

Bounds for the solution of nonlinear systems

Mean Value Theorem :

if
$$f \in C^1$$
 then $f(x) = f(\widetilde{x}) + M(x - \widetilde{x})$ with $M = (\frac{\partial f}{\partial x}(\xi_i))_i$

Let $Y := X - \widetilde{x}$ and

$$\begin{array}{lll} x \in X & \Rightarrow & g(x) - \widetilde{x} & = & x - \widetilde{x} - Rf(x) \\ & & = & -Rf(\widetilde{x}) + (I - RM)(x - \widetilde{x}) \\ & & \in & -Rf(\widetilde{x}) + (I - RM)Y \end{array}$$

As a consequence

$$-Rf(\widetilde{x}) + (I - RM)Y \subseteq Y \quad \Rightarrow \quad g(X) - \widetilde{x} \subseteq Y \quad \Rightarrow \quad g(X) \subseteq X$$

Bounds for the solution of nonlinear systems (cont'd)

Theorem 3 (Rump, 1983)

Let $f : \mathbb{R}^n \to \mathbb{R}^n$ with $f = (f_1, ..., f_n) \in C^1$, $\tilde{x} \in \mathbb{R}^n$, $X \in \mathbb{IR}^n$ with $0 \in X$ and $R \in \mathbb{R}^{n \times n}$ be given. Let $M \in \mathbb{IR}^{n \times n}$ be given such that

 $\{\nabla f_i(\zeta):\zeta\in\widetilde{x}+X\}\subseteq M_{i,:}.$

Denote by I the $n \times n$ identity matrix and assume

 $-Rf(\widetilde{x}) + (I - RM)X \subseteq int(X).$

Then there is a unique $\widehat{x} \in \widetilde{x} + X$ with $f(\widehat{x}) = 0$. Moreover, every matrix $\widetilde{M} \in M$ is nonsingular. In particular, the Jacobian $J_f(\widehat{x}) = \frac{\partial f}{\partial x}(\widehat{x})$ is nonsingular.

- Note that an inclusion of the range of the gradients ∇f_i over the set $\tilde{x} + X$ needs to be computed.
- A convenient way to do this in INTLAB is by interval arithmetic and the gradient toolbox. For a given (Matlab) function f, for xs = x and an interval vector X, the call

$$M = f(gradientinit(xs + X))$$

computes an inclusion M.

Principle of self-validating methods

2 Multiple roots of polynomial systems

3 Numerical experiments

Verification of multiple roots

- Verification method for computing guaranteed (real or complex) error bounds for double roots of systems of nonlinear equations.
- To circumvent the principle problem of ill-posedness we prove that a slightly perturbed system of nonlinear equations has a double root.
- For example, for a given univariate function *f* : ℝ → ℝ we compute two intervals *X*, *E* ⊆ ℝ with the property that there exists *x* ∈ *X* and *e* ∈ *E* such that *x* is a double root of *f*(*x*) := *f*(*x*) − *e*.
- If the function *f* has a double root, typically the interval *E* is a very narrow interval around zero.

Verification of multiple roots

The typical scenario in the univariate case is a function $f : \mathbb{R} \to \mathbb{R}$ with a double root \widehat{x} , i.e. $f(\widehat{x}) = f'(\widehat{x}) = 0$ and $f''(\widehat{x}) \neq 0$. Consider, for example,

$$f(x) = 18x^7 - 183x^6 + 764x^5 - 1675x^4 + 2040x^3 - 1336x^2 + 416x - 48$$

= $(3x - 1)^2(2x - 3)(x - 2)^4$



S. Graillat (Univ. Paris 6)

Verification of multiple roots

- Verification methods for multiple roots of polynomials already exist (Rump,2003). A set containing *k* roots of a polynomial is computed, but no information on the true multiplicity can be given.
- A hybrid algorithm based on the methods of (Rump,2003) is implemented in algorithm verifypoly in INTLAB. Computing inclusions X1, X2 and X3 of the simple root $x_1 = 1.5$, the double root $x_2 = 1/3$ and the quadruple root $x_3 = 2$ of f by algorithm verifypoly in INTLAB we obtain the following.

```
>> X1 = verifypoly(f,1.3), X2 = verifypoly(f,.3), X3 = verifypoly(f,2.1)
intval X1 =
[ 1.49999999999994, 1.500000000078]
intval X2 =
[ 0.33333316656015, 0.33333343640539]
intval X3 =
[ 1.99741678159164, 2.00363593397305]
```

Verification of multiple roots (cont'd)

- The accuracy of the inclusion of the double root $x_2 = 1/3$ is much less than that of the simple root $x_1 = 1.5$, and this is typical.
- If we perturb *f* into *f̃*(*x*) := *f*(*x*) − *ε* for some small real constant *ε* and look at a perturbed root *f̃*(*x̂* + *h*) of *f̃*, then

$$0 = \widetilde{f}(\widehat{x} + h) = -\varepsilon + \frac{1}{2}f''(\widehat{x})h^2 + \mathcal{O}(h^3)$$

implies

$$h \sim \sqrt{2\varepsilon/f''(\widehat{x})}.$$

In general floating-point computations are afflicted with a relative error of size ε ≈ 10⁻¹⁶. This has the same effect as a perturbation of the given function f into f̃. But for double roots, we cannot expect this inclusion to be of better relative accuracy than √ε ≈ 10⁻⁸.

Dealing with double roots

• We consider for a double root the nonlinear system $G : \mathbb{R}^2 \to \mathbb{R}$ with

$$G(x,e) = \left(\begin{array}{c} f(x) - e \\ f'(x) \end{array}\right) = 0$$

in the two unknowns x and e.

• The Jacobian of this system is

$$J_G(x,e) = \begin{pmatrix} f'(x) & -1 \\ f''(x) & 0 \end{pmatrix},$$

so that the nonlinear system is well-conditioned for the double root $x_2 = 1/3$ of f.

Dealing with double roots (cont'd)

• Now we can apply a verification algorithm for solving general systems of nonlinear equation such as algorithm verifynlss in INTLAB. Indeed, applying algorithm verifynlss we obtain

intval Y2 =

- [3.33333333333328e-001, 3.3333333333333337e-001]
- [-2.131628207280424e-014, 2.131628207280420e-014]

Dealing with double roots (cont'd)

- We presented the previous approach in preparation for the multivariate case;
- However, for univariate nonlinear functions we may proceed more directly.

Suppose $X \in \mathbb{IR}$ is an inclusion of a root \hat{x} of f', and use the interval evaluation of f at X to compute $E \in \mathbb{IR}$ with $f(X) \subseteq E$. In particular $f(\hat{x}) \in E$, so that there exists $\hat{e} \in E$ such that the function $g(x) := f(x) - \hat{e}$ satisfies $g(\hat{x}) = g'(\hat{x}) = 0$.

• If, moreover, the inclusion *X* is computed by a verification method, then \widehat{x} is a unique root of f' in *X*, and \widehat{x} is proved to be a double root of *g*.

By this approach we obtain the inclusions for the double root \hat{x} are of the same quality, but the inclusion for the shift is a little weaker than in Y2:

intval X =
[3.33333333333329e-001, 3.333333333333339e-001]
intval E =
[-3.126388037344441e-013, 2.913225216616412e-013]

Dealing with double roots (cont'd)

However, it is superior to expand f with respect to some point $m \in X$. For all $x \in X$ we have $f(x) \in f(m) + f'(X)(X - m) =: E_1$, and in particular $f(\widehat{x}) \in E_1$.

Here *m* should be close to the midpoint of *X*, but need not to be equal to the midpoint. In this case we obtain with

```
intval E1 =
[ -2.131628207280369e-014, 2.131628207280378e-014]
```

an inclusion of the same quality as Y2 by solving *G*.

Note that we use only a univariate verification method to include a root of f', the shift E is obtained by a mere function evaluation.

- Let a suitably smooth function $f : \mathbb{R}^n \to \mathbb{R}^n$ and $\widehat{x} \in \mathbb{R}^n$ be given such that $f(\widehat{x}) = 0$ and the Jacobian of f at \widehat{x} is singular.
- A standard verification method such as verifynlss must fail because with an inclusion of a root the nonsingularity of the Jacobian at the root is proved as well.
- Again it is an ill-posed problem and we need some regularization technique.

Consider the model problem

$$f(x,y) = \begin{pmatrix} f_1(x,y) \\ f_2(x,y) \end{pmatrix} = \begin{pmatrix} x^2 + (x+1)(y-1)^2 - \operatorname{asinh}((x+3)^3 + y^2)\cos(x-xy) \\ (x+1.908718874061618)^2 - \sin(x)(y+1)^2 \end{pmatrix}$$



Figure : Contour lines of $f_1(x) = 0$ (solid) and $f_2(x) = 0$ (dashed)

S. Graillat (Univ. Paris 6)

• As a regularization we add, similar to the univariate case, a smoothing parameter *e* and rewrite into

$$F(x, y, e) = \begin{pmatrix} f_1(x, y) - e \\ f_2(x, y) \\ \det J_f(x, y) \end{pmatrix} = 0.$$

• The third equation forces the tangents of the zero contour lines to be parallel at the solution, whereas the first equation introduces a perturbation to *f*₁ so that the root becomes a double root.

This approach may work for two or three unknowns, however, an explicit formula for the determinant of the Jacobian is prohibitive for larger dimensions. Consider the following way to ensure the Jacobian to be singular.

Let a function $f = (f_1, ..., f_n) : \mathbb{R}^n \to \mathbb{R}^n$ be given and let $\widehat{x} = (\widehat{x}_1, ..., \widehat{x}_n)$ be such that $f(\widehat{x}) = 0$ and the Jacobian $J_f(\widehat{x})$ of f at \widehat{x} is singular. Adding a smoothing parameter e we arrive with $g : \mathbb{R}^{n+1} \to \mathbb{R}^n$ and

$$g(x,e) = \begin{pmatrix} f_1(x) - e \\ f_2(x) \\ \dots \\ f_n(x) \end{pmatrix} = 0$$

at *n* equations in n + 1 unknowns. We force the Jacobian to be singular by

$$J_f(x)y=0$$

for some vector y in the kernel of J_f . In order to make y unique we normalize some component of y to 1.

Theorem 4

Let $f = (f_1, ..., f_n) : \mathbb{R}^n \to \mathbb{R}^n$ with $f \in C^2$ be given. Define $F : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ by

$$F(x, e, y) = \begin{pmatrix} g(x, e) \\ J_f(x)y \end{pmatrix} = 0,$$

where $x = (x_1, ..., x_n)$, $e \in \mathbb{R}$ and $y = (1, y_2, ..., y_n)$. Suppose F suitable assumptions and yields inclusions for $\widehat{x} \in \mathbb{R}^n$, $\widehat{e} \in \mathbb{R}$ and $\widehat{y} \in \mathbb{R}^{n-1}$ such that $F(\widehat{x}, \widehat{e}, \widehat{y}) = 0$. Then $g(\widehat{x}, \widehat{e}) = f(\widehat{x}) - (\widehat{e}, 0, ..., 0)^T = 0$, and the rank of the Jacobian $J_f(\widehat{x})$ of f at \widehat{x} is n - 1.

The system

$$f(x_1, x_2) = \begin{pmatrix} x_1^2 - x_2^2 \\ x_1 - x_2^2 \end{pmatrix} = 0$$

yields

$$J_F(x, e, y) = \begin{pmatrix} 2x_1 & -2x_2 & -1 & 0 \\ 1 & -2x_2 & 0 & 0 \\ 0 & -2 & 0 & 1 \\ 2y & -2 & 0 & 2x_1 \end{pmatrix},$$

as the Jacobian of the augmented system, which is nonsingular for $x_1 = x_2 = 0$. Thus an inclusion is in principle possible.

```
>> f=inline('[x(1)^2-x(2)^2;x(1)-x(2)^2]'),
   verifynlss2(f,[0.002;0.001])
f =
```

```
Inline function:
     f(x) = [x(1)^2 - x(2)^2; x(1) - x(2)^2]
intval ans =
  1.0e-323 *
```

- -0.666666666666666

- 0.66666666666666

Verified multiple eigenvalues

Computing eigenvalues can be viewed as solving the nonlinear system:

$$f(x,\lambda) = \begin{pmatrix} Ax - \lambda x \\ e_k^T x - 1 \end{pmatrix} = 0,$$

As before we regularize the system, but now not by shifting a whole partial function but by changing an individual component a_{ij} of A:

$$g(x,\lambda,\varepsilon,y) = \begin{pmatrix} Ax - \lambda x - \varepsilon x_j e_i \\ e_k^T x - 1 \\ J_f(x,\lambda)y \end{pmatrix} = 0.$$

Again an inclusion is calculated. In this case, the rank of the Jacobian

$$J_f(x,\lambda) = \left(egin{array}{cc} A - \lambda I & -x \ e_k^T & 0 \end{array}
ight)$$

is proved to be *n* and we can also prove that the eigenvalue is of geometric multiplicity one.

S. Graillat (Univ. Paris 6)

Principle of self-validating methods

2 Multiple roots of polynomial systems



First example

Consider

$$f(x) = (\sin(x) - 1)(x - \alpha)$$
 for $\alpha \coloneqq \frac{\pi}{2}(1 + \varepsilon)$.

The function *f* has a double root $\hat{x} = \pi/2$ with another simple root α of relative distance ε to $\pi/2$. Hence we expect the inclusion *E* of the offset *e* for regularization to be a narrow inclusion of zero.

| ε | X | E |
|-----------|--|------------------------------|
| 10^{-2} | $1.5707963267949 \pm 1.8 \cdot 10^{-14}$ | $[-3.5, 1.8] \cdot 10^{-18}$ |
| 10^{-3} | $1.5707963267948 \pm 1.7 \cdot 10^{-13}$ | $[-3.5, 1.8] \cdot 10^{-19}$ |
| 10^{-4} | $1.570796326795 \pm 1.6 \cdot 10^{-12}$ | $[-3.5, 1.8] \cdot 10^{-20}$ |
| 10^{-5} | $1.57079632679 \pm 1.2 \cdot 10^{-10}$ | $[-3.5, 1.8] \cdot 10^{-21}$ |
| 10^{-6} | $1.5707963268 \pm 1.5 \cdot 10^{-9}$ | $[-3.5, 1.8] \cdot 10^{-22}$ |
| 10^{-7} | $1.570796327 \pm 1.6 \cdot 10^{-8}$ | $[-3.5, 1.8] \cdot 10^{-23}$ |
| 10^{-8} | failed | |

Table : Inclusions for the double root $\hat{x} = \pi/2$ and a nearby simple root α for *f*

S. Graillat (Univ. Paris 6)

Second example

Consider now

$$f(x) = (\sin(x) - 1)(x - \alpha)^2$$
 for $\alpha \coloneqq \frac{\pi}{2}(1 + \varepsilon)$,

so that there is a double root α near the double root \hat{x} . For a relative distance ε of about $\sqrt[4]{\varepsilon} \sim 10^{-4}$ the four roots behave like a quadruple root. This is confirmed by the results in the Table.

| ε | X | E |
|-----------|---|------------------------------|
| 10^{-2} | $1.57079632679488 \pm 1.2 \cdot 10^{-14}$ | $[-2.8, 5.5] \cdot 10^{-20}$ |
| 10^{-3} | $1.5707963267948 \pm 2.4 \cdot 10^{-13}$ | $[-2.8, 5.5] \cdot 10^{-22}$ |
| 10^{-4} | $1.570796326794 \pm 2.8 \cdot 10^{-12}$ | $[-2.8, 5.5] \cdot 10^{-24}$ |
| 10^{-5} | failed | |

Table : Inclusions for the double root $\widehat{x} = \pi/2$ and a nearby double root α for *f*

Some systems of nonlinear equations

The first test function is

$$f(x_1, x_2) = \begin{pmatrix} e^{x_1 x_2} - \sin(x_1^2 - 2x_1 x_2) \\ x_1(x_1 - \cosh(x_2)) + x_1 \operatorname{atan}(x_2) - \alpha \end{pmatrix} = 0,$$

where we choose the parameter α such that the system has a nearly double root. For example, for $\alpha = 0.4$ the zero contour lines look like in Figure.



Some systems of nonlinear equations (cont'd)



| X_1 | X_2 | X | E |
|-------------------------|-------------------------|-------------------------------|-------------------------------|
| 1.328899621_{28}^{86} | 1.32889951_{48}^{57} | 1.32889956839071_5^6 | |
| -0.02729805_{59}^{67} | -0.02729792_{88}^{98} | $-0.02729799275879_{34}^{41}$ | $[-5.2, -5.0] \cdot 10^{-14}$ |

Table : Inclusions X_1 , X_2 for two single roots and X for a nearly double root for f and $\alpha = 0.4003120447407$.

Some systems of nonlinear equations (cont'd)



Table : Inclusions X_1 , X_2 for two single roots and X for a nearly double root for f and $\alpha = 0.35653033083794$.

Example of higher dimensions

Consider Brown's almost linear function $f : \mathbb{R}^n \to \mathbb{R}^n$ with

$$f_k(x) = x_k + \sum_{j=1}^n x_j - (n+1) \text{ for } 1 \le k \le n-1,$$

$$f_n(x) = \left(\prod_{j=1}^n x_j\right) - 1 - e,$$

where the last function is shifted by some e. One verifies that for

$$e = \left(1 - \frac{1}{n^2}\right)^{n-1} \left(1 + \frac{1}{n}\right) - 1$$

and

$$\overline{x}_k = 1 - \frac{1}{n^2}$$
 for $1 \le k \le n - 1$,
 $\overline{x}_n = 1 + \frac{1}{n}$

the vector $(1, \ldots, 1, -n)$ is in the kernel of the Jacobian of f.

Example of higher dimensions (cont'd)

Thus \overline{x} is not a simple root of f. More precisely it is verified that there exists $\widehat{x} \in X$ and $\widehat{\varepsilon} \in E$ such that $f(\widehat{x}) - (\widehat{\varepsilon}, ..., 0) = 0$ and the Jacobian $J_f(\widehat{x})$ of f at \widehat{x} is singular.

| п | $X_{1\cdots n-1}$ | X_n | E |
|------|------------------------------------|---------------------------------|------------------------------|
| 10 | $0.990000 \pm 1.0 \cdot 10^{-14}$ | $1.100000 \pm 1 \cdot 10^{-14}$ | $[-3.5, 5.8] \cdot 10^{-15}$ |
| 20 | $0.997500 \pm 4.0 \cdot 10^{-14}$ | $1.050000 \pm 1 \cdot 10^{-14}$ | $[-1.4, 2.2] \cdot 10^{-14}$ |
| 50 | $0.996000 \pm 2.1 \cdot 10^{-13}$ | $1.020000 \pm 2 \cdot 10^{-14}$ | $[-0.1, 1.9] \cdot 10^{-13}$ |
| 100 | $0.999900 \pm 8.2 \cdot 10^{-13}$ | $1.010000 \pm 2 \cdot 10^{-14}$ | $[-5.4, 2.9] \cdot 10^{-13}$ |
| 200 | $0.999975 \pm 3.3 \cdot 10^{-12}$ | $1.005000 \pm 5 \cdot 10^{-14}$ | $[-1.3, 2.0] \cdot 10^{-12}$ |
| 500 | $0.999996 \pm 1.9 \cdot 10^{-11}$ | $1.002000 \pm 1 \cdot 10^{-13}$ | $[-0.6, 1.3] \cdot 10^{-11}$ |
| 1000 | $0.9999999 \pm 7.5 \cdot 10^{-11}$ | $1.001000 \pm 2 \cdot 10^{-13}$ | $[-1.1, 6.4] \cdot 10^{-11}$ |

Table : Inclusions of a double root for different dimensions.

Conclusion and future work

Conclusion:

- Efficient algorithms for computing verified and narrow error bounds with the property that a slightly perturbed system is proved to have a double root within the computed bounds
- Applied those to univariate polynomials, to multivariate polynomials and also to eigenvalue problems
- Numerical experiments have confirmed the performance of our algorithms

Future work:

- Detecting singular matrices
- Applications to approximate coprimeness

Bibliography I

Siegfried M. Rump and Stef Graillat.
 Verified error bounds for multiple roots of systems of nonlinear equations.

Numer. Algorithms, 54 (2010), no. 3, 359-377.

Siegfried M. Rump.

Verification methods: Rigorous results using floating-point arithmetic.

Acta Numerica (2010), pp. 287-449.

Bo Einarsson.

Accuracy and Reliability in Scientific Computing. Software-Environments-Tools. SIAM, Philadelphia, PA, 2005.

Bibliography II

Nicholas J. Higham.

Accuracy and stability of numerical algorithms. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 2002.

- Jean-Michel Muller et al. Handbook of Floating-Point Arithmetic. Birkhäuser, 2010.
- R. Moore, R. Kearfott et M. Cloud. Introduction to Interval Analysis. SIAM, 2009.

Thank you for your attention