Numer Algor (2010) 54:359–377
DOI 10.1007/s11075-009-9339-3

ORIGINAL PAPER

Verified error bounds for multiple roots of systems of nonlinear equations

Siegfried M. Rump · Stef Graillat

Received: 20 May 2009 / Accepted: 24 September 2009 /
Published online: 16 October 2009

Abstract It is well known that it is an ill-posed problem to decide whether a function has a multiple root. Even for a univariate polynomial an arbitrary small perturbation of a polynomial coefficient may change the answer from yes to no. Let a system of nonlinear equations be given. In this paper we describe an algorithm 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. For a univariate nonlinear function f we give a similar method also for a multiple root. A narrow error bound for the perturbation is computed as well. Computational results for systems with up to 1000 unknowns demonstrate the performance of the methods.

Keywords Nonlinear equations · Double roots · Multiple roots · Verification · Error bounds · INTLAB

Mathematics Subject Classifications (2000) 65H10 · 65G20 · 65H05 · 65-04

S. M. Rump (✉)
Institute for Reliable Computing, Hamburg University of Technology,
Schwarzenbergstraße 95, Hamburg 21071, Germany
e-mail: rump@tu-harburg.de

S. M. Rump
Faculty of Science and Engineering, Waseda University,
3–4–1 Okubo, Shinjuku-ku, Tokyo 169-8555, Japan

S. M. Rump · S. Graillat
Laboratoire LIP6, Département Calcul Scientifique, Université Pierre et Marie Curie
(Paris 6), 4 place Jussieu, 75252 Paris cedex 05, France

S. Graillat
e-mail: stef.graillat@lip6.fr

1 Introduction

It is well known that to decide whether a univariate polynomial has a multiple root is an ill-posed problem: An arbitrary small perturbation of a polynomial coefficient may change the answer from yes to no. In particular a real double root may change into two simple (real or complex) roots.

Therefore it is hardly possible to verify that a polynomial or a nonlinear function has a double root if not the entire computation is performed without any rounding error, i.e. using methods from Computer Algebra.

A typical so-called "verification method" is based on a theorem the assumptions of which are verified on the computer. Typically such theorems are in turn based on some kind of fixed point theorem (see Section 2). The verification of the assumptions is performed using floating-point arithmetic with rigorously estimating all intermediate rounding errors. The computed results have a mathematical certainty. Some of those methods are collected in INTLAB [18], the Matlab [10] Toolbox for Reliable Computing.

The computing time of such a verification method is often of the order of a comparable pure approximative (floating-point) algorithm, whereas the latter does not provide the kind of guaranty of the correctness of the result. A main reason is that verification methods use floating-point arithmetic as well, combined with suitable error estimations. Moreover, the input data may consist of intervals. In such a case, for example, it is possible to verify that all matrices within an interval matrix are nonsingular, an NP-hard problem [16]. Since the verification method is polynomially time bounded, this works only if the input intervals are not too wide; otherwise the verification fails, but no false answer is possible.

In case of an exactly given (real or complex floating-point) matrix, the verification of nonsingularity is, of course, possible as well. As a drawback however, in contrast to Computer Algebra methods, the verification of *singularity* is by principle outside the scope of verification methods because this is an ill-posed problem: An arbitrarily small perturbation of a singular matrix may produce a regular matrix changing the answer discontinuously from "yes" to "no".

For the same reason a verification method cannot prove that a polynomial has a double root or a matrix has a double eigenvalue unless all computations are performed without error; however, it is possible to verify that a (small) disc in the complex plane contains two roots or two eigenvalues. This is done without deciding whether it is a double root or not. Corresponding methods can be found in [1, 19, 20].

Other methods have been designed to deal with singularities. In [8, 22], the authors propose a statistic approach of the rounding errors. Using stochastic arithmetic, they make it possible to see if a result is significant and if rounding errors lead to a nonsignificant result. For example, if the computed determinant of a matrix has no significant digit, we can say that this matrix is numerically singular [9]. Using the same method one may compute the (numerical) multiplicity of a root. A root is considered as multiple if its exact significant digits are common with those of at least another one [2]. Note that a root may be proved to be multiple with a high probability, but not with certainty.

In this paper we describe a 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 $f : \mathbb{R} \to \mathbb{R}$ we compute two intervals $X, E \subseteq \mathbb{R}$ with the property that there exists $\hat{x} \in X$ and $\hat{e} \in E$ such that $\hat{x}$ is a double root of $\bar{f}(x) := f(x) - \hat{e}$. If the function f has a double root, typically the interval E is a very narrow interval around zero. For complex discs and system of equations assertions are similar.

Moreover, the computed inclusions are narrow, typically with relative error of the order of the relative rounding error. Note that the sensitivity of the problem prohibits such narrow error bounds for a pair of roots, see the next section.

The paper is organized as follows. In Section 2 we briefly summarize how to compute verified error bounds for a (simple) solution of a system of nonlinear equations. In Section 3 we develop methods to compute verified error bounds for a double or multiple root of a univariate nonlinear function, and in Section 4 we treat double roots of systems of nonlinear equations. In Section 5 we describe an application of the methods to multiple eigenvalues, and we close the paper with numerical results.

We mention that the results are formulated over the real numbers $\mathbb{R}$, but are valid mutatis mutandis over the complex numbers as well.

2 Verified solution of nonlinear systems

In the following we assume floating-point arithmetic performed in IEEE 754 double precision corresponding to a relative rounding error unit $\mathbf{u} = 2^{-53} \approx 1.11 \cdot 10^{-16}$. The i-th row or column of a matrix $A \in \mathbb{R}^{n \times n}$ are denoted by $A_{i,:}$ or $A_{:,i}$, respectively, similar to Matlab notation.

Denote by $\mathbb{IR}$ the set of real intervals, and by $\mathbb{IR}^n$ and $\mathbb{IR}^{n \times n}$ the set of real interval vectors and interval matrices, respectively. To understand the mathematical part of the following it suffices to think of operations between interval quantities as power set operations. Concerning an efficient implementation, interval arithmetic is used (see [14]). Using INTLAB [18] there is easy access to efficient interval operations.

Standard verification methods for systems of nonlinear equations are based on the following theorem [17].

Theorem 2.1 *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ with $f = (f_1, \ldots, f_n) \in \mathcal{C}^1$, $\tilde{x} \in \mathbb{R}^n$, $X \in \mathbb{IR}^n$ 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*

$$\{\nabla f_i(\zeta) : \zeta \in \tilde{x} + X\} \subseteq M_{i,:} \,. \tag{2.1}$$

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

$$-Rf(\tilde{x}) + (I - RM)X \subseteq int(X). \tag{2.2}$$

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

Remark Note that in (2.1) 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` $= \tilde{x}$ and an interval vector `X`, the call

$$\texttt{M} = \texttt{f(gradientinit(xs} + \texttt{X))} \tag{2.3}$$

computes an inclusion `M` satisfying (2.1). Note that (2.3) is executable Matlab/INTLAB code. Similarly Hessians can be computed using the Hessian toolbox in INTLAB.

The proof of Theorem 2.1 in [17] is based on Brouwer's Fixed Point Theorem and the fact that for every $x \in \tilde{x} + X$ there exists some $\tilde{M} \in M$ such that $f(x) = f(\tilde{x}) + \tilde{M}(x - \tilde{x})$. Note that there is no assumption on $\tilde{x}$, X or R. In order for (2.2) to be satisfied good choices are an approximation $\tilde{x}$ of a root of f, an approximation R of the inverse of the Jacobian $J_f(\tilde{x})$ and an interval vector X of small width containing zero. We stress, however, that independent of the quality of the approximations the assertions of Theorem 2.1 remain true. An implementation of a verification method for nonlinear systems based on Theorem 2.1 is algorithm `verifynlss` in INTLAB [18].

The theorem uses a modification of the Krawczyk operator introduced in [7] and the existence test of a root of a nonlinear system by Moore [11]. Whereas Krawczyk supposes an interval vector Y containing a unique root of $f(x)$ to be known already and Moore gives no clue what to do when the test fails, in [17] an iteration scheme is introduced for constructing an inclusion vector. The iteration uses the so-called epsilon-inflation, where under mild assumptions it is proved that the iteration constructs an inclusion if and only if there is a simple root of f near $\tilde{x}$. Moreover in [17] the inclusion of the error with respect to some approximate solution $\tilde{x}$ was introduced. All these techniques are today standard for many verification methods (see [3, 13, 15] and many others).

Part of the assertions of Theorem 2.1 is the nonsingularity of the Jacobian $J_f(\hat{x})$. Naturally this restricts the application to simple roots because it is proved that the root is simple. Next we will derive verification methods to prove existence of a truly multiple root of a slightly perturbed function.

3 The univariate case

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

$$\begin{aligned} f(x) &= 18x^7 - 183x^6 + 764x^5 - 1675x^4 + 2040x^3 - 1336x^2 + 416x - 48 \\ &= (3x-1)^2(2x-3)(x-2)^4 \end{aligned} \tag{3.1}$$

The graph of this function is shown in Fig. 1. In [20] verification methods for multiple roots of polynomials are presented. Here, for example, 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 in [20] is implemented in algorithm `verifypoly` in INTLAB.

To compute inclusions of the roots of f we need rough approximations. 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 in (3.1) by algorithm `verifypoly` in INTLAB we obtain the following (the polynomial is, of course, specified in expanded form, not the factored form). Note that only rough approximations of the roots are necessary.

```
>> X1 = verifypoly(f,1.3), X2 = verifypoly(f,.3),
   X3 = verifypoly(f,2.1)
intval X1 =
[   1.49999999999904,   1.50000000000078]
intval X2 =
[   0.33333316656015,   0.33333343640539]
intval X3 =
[    1.99741678159164,   2.00363593397305]
```

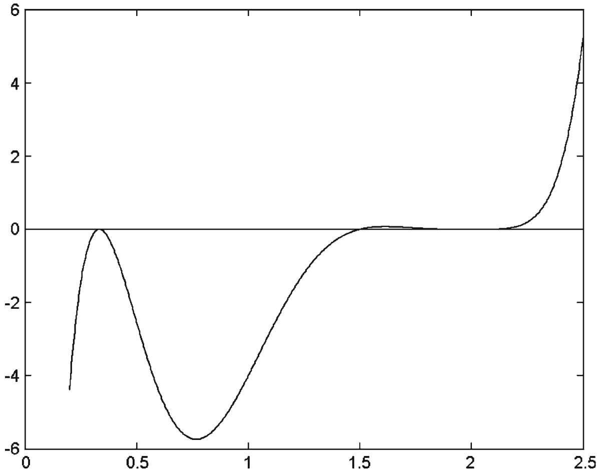

Fig. 1 Graph of $f(x) = (3x-1)^2(2x-3)(x-2)^4$

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 $\tilde{f}(x) := f(x) - \varepsilon$ for some small real constant ε and look at a perturbed root $\tilde{f}(\hat{x}+h)$ of $\tilde{f}$, then

$$0 = \tilde{f}(\hat{x}+h) = -\varepsilon + \frac{1}{2} f''(\hat{x})h^2 + \mathcal{O}(h^3) \tag{3.2}$$

implies

$$h \sim \sqrt{2\varepsilon/f''(\hat{x})}. \tag{3.3}$$

In general floating-point computations are afflicted with a relative error of size $\mathbf{u} \approx 10^{-16}$. This has the same effect as a perturbation of the given function f into $\tilde{f}$. Therefore we may compute an inclusion of two roots of a nonlinear function, but by (3.2) and (3.3) we cannot expect this inclusion to be of better relative accuracy than $\sqrt{\mathbf{u}} \approx 10^{-8}$. This corresponds to the inclusion `X2` above and to the results in [1, 19, 20].

Similarly it is known that the sensitivity of a k-fold root is of the order $\mathbf{u}^{1/k}$, so that for the quadruple root $x_3 = 2$ of f we cannot expect a better relative accuracy than $\sqrt[4]{\mathbf{u}} \sim 10^{-4}$. This corresponds to the accuracy of `X3`.

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

$$G(x, e) = \begin{pmatrix} f(x) - e \\ f'(x) \end{pmatrix} = 0 \tag{3.4}$$

in the two unknowns x and e. The Jacobian of this system is

$$J_G(x, e) = \begin{pmatrix} f'(x) & -1 \\ f''(x) & 0 \end{pmatrix}, \tag{3.5}$$

so that the nonlinear system (3.4) is well-conditioned for the double root $x_2 = 1/3$ of f in (3.1). Now we can apply a verification algorithm for solving general systems of nonlinear equations based on Theorem 2.1 such as algorithm `verifynlss` in INTLAB. Note that the system of nonlinear functions is provided by a Matlab subroutine for computing the function values. No more information is necessary; in particular derivatives are computed by means of automatic differentiation. Indeed, applying algorithm `verifynlss` to (3.4) we obtain

```
>> Y2 = verifynlss(G,[.3;0])
intval Y2 =
[  3.333333333333328e-001,  3.333333333333337e-001]
[ -2.131628207280424e-014,  2.131628207280420e-014]
```

This proves that there is a constant ε with $|\varepsilon| \leq 2.14 \cdot 10^{-14}$ such that the nonlinear equation $f(x) - \varepsilon = 0$ has a double root $\hat{x}$ with $0.3333333333333328 \leq \hat{x} \leq 0.3333333333333337$. In contrast to the previous inclusion `X2` the new inclusion `Y2` is very accurate. The reason is that only double roots are taken into account, and this removes the high sensitivity of the root. It is a kind of regularization.

We presented the approach (3.4) 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 based on Theorem 2.1, then $\hat{x}$ is a unique root of f' in X, and $\hat{x}$ 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 =
[  3.333333333333329e-001,  3.333333333333339e-001]
intval E =
[ -3.126388037344441e-013,  2.913225216616412e-013]
```

However, it is superior to expand f with respect to some point $m \in X$. For all $x \in X$ we have $f(x) \in f(m) + f'(X)(X - m) =: E1$, and in particular $f(\hat{x}) \in E1$. 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 in (3.4). 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. This seems superior to solving the bivariate system (3.4).

We note that one might use $f(x) \in f(m) + f'(m)(X - m) + \frac{1}{2} f''(X)(X - m)^2$ for all $x \in X$; however, we did not observe much advantage over using $E1$ as computed before.

For a k-fold multiple root we proceed similarly. Let X be an inclusion of a root $\hat{x}$ of $f^{(k-1)}$, choose $m \in X$ and compute successively for $j = 0, 1, \ldots, k-2$

$$E_j = f^{(k-2-j)}(m) + f^{(k-1-j)}(X)(X - m) - \sum_{\nu=0}^{j-1} \frac{E_\nu}{(j-\nu)!} X^{j-\nu} \,. \tag{3.6}$$

It follows

$$f^{(k-2-j)}(\hat{x}) = \hat{e}_j + \sum_{\nu=0}^{j-1} \frac{\hat{e}_\nu}{(j-\nu)!} \hat{x}^{j-\nu} \tag{3.7}$$

for some $\hat{e}_j \in E_j$ and $0 \le j \le k-2$. With a straightforward computation we obtain the following result.

Theorem 3.1 *Let $f : \mathbb{R} \to \mathbb{R}$ with $f \in \mathcal{C}^{k+1}$ be given. Assume $X \in \mathbb{IR}$ is an inclusion of a root $\hat{x}$ of $f^{(k-1)}$. Let $E_j \in \mathbb{IR}$ be computed by (3.6) for $j = 0, 1, \ldots, k-2$, so that there exist $\hat{e}_j \in E_j$ with (3.7). Define $g : \mathbb{R} \to \mathbb{R}$ by*

$$g(x) := f(x) - \sum_{\nu=0}^{k-2} \frac{\hat{e}_\nu}{(k-2-\nu)!}\, x^{k-2-\nu} \,. \tag{3.8}$$

Then $g^{(j)}(\hat{x}) = 0$ for $0 \le j \le k-1$. If the inclusion X is computed by a verification method based on Theorem 2.1, then the multiplicity of the the root $\hat{x}$ of g in X is exactly k-fold.

For the 4-fold root $\hat{x} = 2$ of f in (3.1) we obtain the following inclusions with the proof that the regularized equation $g(x) = f(x) - \frac{1}{2}\hat{e}_0 x^2 - \hat{e}_1 x - \hat{e}_2$ with $\hat{e}_j \in E_j$ has a quadruple root in X.

```
intval X =
[  1.999999999999963e+000,  2.000000000000040e+000]
intval E0 =
[ -4.547473510733392e-013, -4.547473506997975e-013]
intval E1 =
[ -1.364242053101423e-012, -1.364242052216876e-012]
intval E2 =
[  4.604316926434916e-012,  4.604316929015351e-012]
```

4 The multivariate case

Let a suitably smooth function $f : \mathbb{R}^n \to \mathbb{R}^n$ and $\hat{x} \in \mathbb{R}^n$ be given such that $f(\hat{x}) = 0$ and the Jacobian of f at $\hat{x}$ 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

$$\begin{aligned} f(x,y) &= \begin{pmatrix} f_1(x,y) \\ f_2(x,y) \end{pmatrix} \\ &= \begin{pmatrix} x^2 + (x+1)(y-1)^2 - \operatorname{asinh}((x+3)^3 + y^2)\cos(x - xy) \\ (x + 1.908718874061618)^2 - \sin(x)(y+1)^2 \end{pmatrix} = 0 \,. \end{aligned} \tag{4.1}$$

In Fig. 2 the zero contour lines of f are displayed. Near $(x, y) = (0.60, 2.34)$ the tangents of the contour lines are nearly parallel so that the Jacobian of f

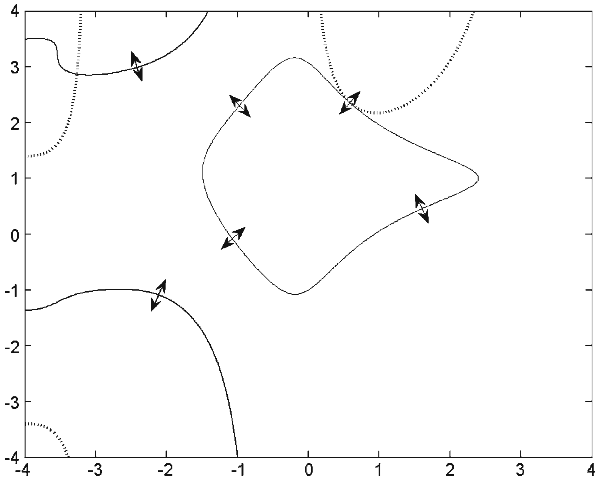

Fig. 2 Contour lines of $f_1(x) = 0$ (*solid*) and $f_2(x) = 0$ (*dashed*)

at the nearby root is nearly singular. As a regularization we add, similar to the univariate case, a smoothing parameter e and rewrite (4.1) into[1]

$$F(x, y, e) = \begin{pmatrix} f_1(x, y) - e \\ f_2(x, y) \\ \det J_f(x, y) \end{pmatrix} = 0\,. \tag{4.2}$$

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. Locally the zero contour lines behave linearly, so that the smoothing parameter expands or shrinks the zero line for f_1 as depicted by the double arrows in Fig. 2. Each point of the contour line moves locally normal to the contour line itself. Obviously this is optimal for the regularization.

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

$$g(x, e) = \begin{pmatrix} f_1(x) - e \\ f_2(x) \\ \cdots \\ f_n(x) \end{pmatrix} = 0 \tag{4.3}$$

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

$$J_f(x)y = 0 \tag{4.4}$$

[1] As has been pointed out by Prof. Oishi, some similar approach was followed in [5, 6].

for some vector y in the kernel of J_f. In order to make y unique we normalize some component of y to 1. For the moment we choose the first component so that $y = (1, y_2, \dots, y_n)$. In practice we have to choose a suitable component for normalization, see below. Now (4.4) adds another n equations in $n-1$ unknowns, so that we arrive at a system of $2n$ equations (4.3) and (4.4) in $2n$ unknowns $(x_1, \dots, x_n, e, y_2, \dots, y_n)$. Note that the new equations (4.4) only ensure the Jacobian to be singular and have no influence on the described regularization technique.

Theorem 4.1 *Let $f = (f_1, \dots, f_n) : \mathbb{R}^n \to \mathbb{R}^n$ with $f \in \mathcal{C}^2$ be given. Define $F : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ by (4.3) and*

$$F(x, e, y) = \begin{pmatrix} g(x, e) \\ J_f(x)y \end{pmatrix} = 0\,, \tag{4.5}$$

where $x = (x_1, \dots, x_n)$, $e \in \mathbb{R}$ and $y = (1, y_2, \dots, y_n)$. Suppose Theorem 2.1 is applicable to F and yields inclusions for $\hat{x} \in \mathbb{R}^n, \hat{e} \in \mathbb{R}$ and $\hat{y} \in \mathbb{R}^{n-1}$ such that $F(\hat{x}, \hat{e}, \hat{y}) = 0$. Then $g(\hat{x}, \hat{e}) = f(\hat{x}) - (\hat{e}, 0, \dots, 0)^T = 0$, and the rank of the Jacobian $J_f(\hat{x})$ of f at $\hat{x}$ is $n-1$.

Proof By Theorem 2.1 we have $f(\hat{x}) = (\hat{e}, 0, \dots, 0)^T$ and $J_f(\hat{x})(1, \hat{y}_2, \dots, \hat{y}_n)^T = 0$, so that $J_f(\hat{x})$ has a nontrivial vector in its kernel. We have to show that the rank of the Jacobian $J_f(\hat{x})$ is not less than $n-1$. The Jacobian of F computes to

$$J_F(x, e, y) = \begin{pmatrix} J_f(x) & I_{:,1} & \mathcal{O}_{n,n-1} \\ H & \mathcal{O}_{n,1} & J_f(x)_{:,2..n} \end{pmatrix}\,, \tag{4.6}$$

where I denotes the $n \times n$ identity matrix and $\mathcal{O}_{k,l}$ the $k \times l$ zero matrix. The $i-th$ row of H computes to

$$H_{i,:} = (1, y_2, \dots, y_n) \cdot Hessian(f_i(x))\,.$$

By Theorem 2.1 we know that $J_F(\hat{x}, \hat{e}, \hat{y})$ is nonsingular. If the rank of $J_f(\hat{x})$ is less than $n-1$, then there is a vector $z \in \mathbb{R}^n$ in its kernel which is not a scalar multiple of $(1, \hat{y}_2, \dots, \hat{y}_n)^T$. If $z_1 = 0$, then $J_F(\hat{x}, \hat{e}, \hat{y})(0, \dots, 0, z)^T = 0$, a contradiction. If $z_1 \neq 0$, then a suitable linear combination of z and $(1, \hat{y}_2, \dots, \hat{y}_n)^T$ has a first component zero, which is again a contradiction to the nonsingularity of J_F. □

Two problems remain. The first is how to choose a suitable component for normalizing the vector in the kernel of J_f. For a given matrix $A \in \mathbb{R}^{n\times n}$, Gaussian elimination with partial pivoting yields LU-factors and a permutation matrix. Applying this to A^T yields $PA^T = LU$. Total pivoting guarantees that the rank of A is $n-1$ or less if and only if $U_{nn} = 0$ (cf. [4]), and except extraordinary circumstances this is also true for partial pivoting. Then $Ax = 0$ for $L^T Px = I_{:,n}$. Applying this to the Jacobian J_f and taking a component of x with largest absolute value is a suitable choice for the component to be normalized to 1.

The second problem is that an inclusion cannot be computed if the rank of the Jacobian J_f is less than $n-1$. More precisely, we proved that if an inclusion of a multiple root is computed, then the rank of the Jacobian is $n-1$, and it would be nice to have the converse, namely that for a root $f(\hat{x}) = 0$ and Jacobian $J_f(\hat{x})$ of rank $n-1$ an inclusion can be computed by applying Theorem 2.1 to (4.3) and (4.4). This is not true as by

$$f(x_1, x_2) = \begin{pmatrix} x_1 - x_2^2 \\ x_1^2 - x_2^2 \end{pmatrix} = 0\,. \tag{4.7}$$

Obviously the Jacobian has rank 1 at $x_1 = x_2 = 0$, but the Jacobian (4.6) of the augmented system (4.5) computes to

$$J_F(x, e, y) = \begin{pmatrix} 1 & -2x_2 & -1 & 0 \\ 2x_1 & -2x_2 & 0 & 0 \\ 0 & -2 & 0 & 1 \\ 2y & -2 & 0 & 2x_1 \end{pmatrix}, \tag{4.8}$$

which is singular for $x_1 = x_2 = 0$. This means that it is not possible to compute an inclusion of the multiple root (0, 0). However, in this case the reason is that the wrong equation was regularized. Exchanging the two equations in (4.7) into

$$f(x_1, x_2) = \begin{pmatrix} x_1^2 - x_2^2 \\ x_1 - x_2^2 \end{pmatrix} = 0 \tag{4.9}$$

yields

$$J_F(x, e, y) = \begin{pmatrix} 2x_1 & -2x_2 & -1 & 0 \\ 1 & -2x_2 & 0 & 0 \\ 0 & -2 & 0 & 1 \\ 2y & -2 & 0 & 2x_1 \end{pmatrix}, \tag{4.10}$$

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

```
>> 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 *
[  -0.66666666666666,   0.66666666666666]
[  -1.00000000000000,   1.00000000000000]
[  -1.00000000000000,   1.00000000000000]
```

However, we mention that in this case the iteration is sensitive to the initial approximation as by

```
>> verifynlss2(f,[0.001;0.001])
intval ans =
[   0.49999999999999,   0.50000000000001]
[   0.70710678118654,   0.70710678118655]
[  -0.25000000000001,  -0.24999999999999]
```

which finds the double root $(0.5, 1/\sqrt{2})$ of $x^2 - y^2 + 0.25 = 0$ and $x - y^2 = 0$.

We might hope that there is always a renumbering of the equations such that for Jacobian $J_f(\hat{x})$ of rank $n-1$ an inclusion of the root $\hat{x}$ can be computed. Unfortunately this is not the case. Consider

$$f(x_1, x_2) = \begin{pmatrix} x_1^2 x_2 - x_1 x_2^2 \\ x_1 - x_2^2 \end{pmatrix} = 0\,. \tag{4.11}$$

The Jacobian of the augmented system computes to

$$J_F(x, e, y) = \begin{pmatrix} 2x_1x_2 - x_2^2 & x_1^2 - 2x_1x_2 & -1 & 0 \\ 1 & -2x_2 & 0 & 0 \\ 2(x_2 y + x_1 - x_2) & 2(x_1 - x_2)y - 2x_1 & 0 & 2x_1x_2 - x_2^2 \\ 0 & -2 & 0 & 1 \end{pmatrix}, \tag{4.12}$$

Obviously the third row is entirely zero for $x_1 = x_2 = 0$, and this does not change when interchanging the two equations in (4.11). Note that the Jacobian J_f at the root is $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ and forces the kernel vector to be $\begin{pmatrix} y_1 \\ 1 \end{pmatrix}$. Summarizing an inclusion for the root $(0, 0)$ of (4.11) is in principle not possible by our method. But this situation is rare and can be characterized as follows.

Theorem 4.2 *Let $f = (f_1, \ldots, f_n) : \mathbb{R}^n \to \mathbb{R}^n$ with $f \in \mathcal{C}^2$ be given. Define $D(x) : \mathbb{R}^n \to \mathbb{R}$ by $D(x) := det(J_f(x))$, and define the function $F^{[k]}(x, e) : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ for $x \in \mathbb{R}^n$ and $e \in \mathbb{R}$ by*

$$F^{[k]}(x, e) = \begin{pmatrix} g(x, e) \\ D(x) \end{pmatrix} = 0\,, \tag{4.13}$$

where the component functions of $g : \mathbb{R}^{n+1} \to \mathbb{R}^n$ are defined by

$$g_i(x, e) \quad = \quad \begin{cases} f_i(x) & \text{for } i \neq k \\ f_k(x) - e & \text{for } i = k \end{cases}. \tag{4.14}$$

Let some $\hat{x} \in \mathbb{R}^n$ be given such that the rank of the Jacobian $J_f(\hat{x})$ of f at $\hat{x}$ is $n - 1$. Then the following is equivalent.

i) For all k is the Jacobian $J_{F^{[k]}}(\hat{x})$ of $F^{[k]}$ at $\hat{x}$ singular.
ii) The $(n+1) \times n$ matrix $[J_f(\hat{x}); \nabla D(\hat{x})]$ of the Jacobian $J_f(\hat{x})$ of f at $\hat{x}$ appended by the gradient $\nabla D(\hat{x})$ of its determinant has rank $n - 1$.

Remark Note that the Jacobian of $F^{[k]}$ does not depend on e for all k.

Proof of Theorem 4.2 The last column of the Jacobian of $F^{[k]}$ is entirely zero except the k-th component, and the last row is the gradient of $D(x)$ appended by zero. Since the Jacobian $J_f(\hat{x})$ of f at $\hat{x}$ is singular, the result follows. □

It is easy to see that Theorem 4.2 applies to our formulation (4.2) as well. It means for Theorem 4.1 that in general and if the situation is numerically not too difficult we can expect to be able to compute an inclusion of a multiple root of the perturbed nonlinear system (4.3). Note that $\nabla \det J_f(\hat{x})$ is entirely zero in example (4.12), so that *ii*) in Theorem 4.2 is satisfied.

We can introduce a strategy for finding a suitable partial function f_k for regularization to make sure that for a system like (4.7) our method does not fail due to poor numbering of the equations. The strategy is much in the spirit of choosing a good component for normalization of a vector in the kernel of J_f. Suppose the first function f_1 is regularized. According to (4.6) it is necessary that the rows $2..n$ of J_f are linearly independent, otherwise the Jacobian $J_F(x, e, y)$ is singular. So this situation must be avoided.

Suppose the rank of J_f is $n-1$—otherwise by Theorem 4.1 a verification is not possible anyway. Then there exists a row K of J_f such that the rows $\{J_f(i,:) : i \neq K\}$ are linearly independent, which is equivalent to $x^T J_f \neq 0$ for nontrivial x with $x_K = 0$. Since the rank of J_f is $n-1$, its kernel is one-dimensional, i.e. a scalar multiple of some $z \in \mathbb{R}^n$. So any value of K is suitable with $z_K \neq 0$.

This shows that it is not likely to be trapped by this, and it also gives the clue to find a suitable K: As before Gaussian elimination with partial pivoting yields $PJ_f = LU$, and the solution of $L^T Px = I_{:,n}$ spans the kernel of J_f. Taking a component K of x of largest absolute value is a suitable choice. These strategies are implemented in algorithm `verifynlss2` in INTLAB.

5 Double eigenvalues

Finally we show an application of our method to double eigenvalues of a matrix. It has been mentioned that in [19] methods are given to calculate inclusions of a cluster of eigenvalues of a matrix. More precisely, for a given $n \times n$-matrix A inclusions of a $k \times k$-matrix M and an $n \times k$-matrix X are calculated such that the Jordan form of M is part of the Jordan form of A, and X is an invariant subspace of A. By calculating an inclusion $L \subseteq \mathbb{C}$ of the eigenvalues of M it follows that there are k eigenvalues of the original matrix A in L.

Note that it is not proved that A has a k-fold or even a double eigenvalue. Also note that an inclusion is only possible if the geometric multiplicity of all included eigenvalues is 1. The reason is again, as for multiple roots of polynomials, that the problem becomes ill-posed for geometric multiplicity greater that one [21].

Using our approach we may prove existence of a double eigenvalue, but of a slightly perturbed matrix. In this particular case we may estimate how much an individual component of the input matrix A has to be perturbed such that a true double eigenvalue appears. As in [17, 19] consider

$$f(x,\lambda) = \begin{pmatrix} Ax - \lambda x \\ e_k^T x - 1 \end{pmatrix} = 0\,, \tag{5.1}$$

where e_k denotes the k-th column of the identity matrix. Obviously a solution (x,λ) is an eigenvector/eigenvalue pair of A. Here k is a suitably chosen component normalizing the eigenvector x. As before we regularize the system (5.1), but now not by shifting a whole partial function but by changing an individual component a_{ij} of A:

$$g(x,\lambda,\varepsilon,y) = \begin{pmatrix} Ax - \lambda x - \varepsilon x_j e_i \\ e_k^T x - 1 \\ J_f(x,\lambda)y \end{pmatrix} = 0\,. \tag{5.2}$$

Again an inclusion is calculated using Theorem 2.1. In this case, as by Theorem 4.1, the rank of the Jacobian

$$J_f(x,\lambda) = \begin{pmatrix} A - \lambda I & -x \\ e_k^T & 0 \end{pmatrix}$$

is proved to be n, and it is easy to see [21] that the eigenvalue λ must be of geometric multiplicity 1. Computational tests show that for dimensions over $n = 200$ of the matrix inclusions deteriorate. This means that in general only some 6 to 10 decimal places of the inclusion can be guaranteed. The inclusion of ε is always not far from $\mathbf{u} \cdot \mathrm{norm}(A)$. It proves that changing a_{ij} into $a_{ij} + \varepsilon$ produces a matrix with a double eigenvalue. So there is the choice to prove that the *original* matrix A has two (possibly separated) eigenvalues within some computed narrow bounds, or that a slightly *perturbed* matrix has a true double eigenvalue.

6 Numerical results

We add some numerical examples for the univariate and the multivariate case. We implemented the methods using (3.4) and (4.3), (4.4) in Algorithm `verifynlss2` in INTLAB, see http://www.ti3.tu-harburg.de/rump. Following we display results of this algorithm.

Table 1 Inclusions for the double root $\hat{x} = \pi/2$ and a nearby simple root α for f as in (6.1)

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 2 Inclusions for the double root $\hat{x} = \pi/2$ and a nearby double root α for f as in (6.2)

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	

First consider

$$f(x) = (\sin(x) - 1)(x - \alpha) \quad \text{for } \alpha := \frac{\pi}{2}(1 + \varepsilon)\ . \tag{6.1}$$

The function f has a double root $\hat{x} = \pi/2$ with another simple root α of relative distance ε to $\pi/2$. Hence in any case we expect the inclusion E of the offset e for regularization to be a narrow inclusion of zero. Table 1 displays the results for different values of ε.

As can be seen for decreasing relative distance of α to the double root $\hat{x}$ the quality of the inclusion decreases. An inclusion is possible until about a relative error $10^{-8} \sim \sqrt{\mathbf{u}}$. This corresponds to the sensitivity of the double root $\hat{x}$: If there is another root α of relative distance $\sqrt{\mathbf{u}}$, then numerically the three roots cannot be distinguished in a floating-point arithmetic with relative rounding error unit $\mathbf{u}$. This effect can also be observed when changing f into

$$f(x) = (\sin(x) - 1)(x - \alpha)^2 \quad \text{for } \alpha := \frac{\pi}{2}(1 + \varepsilon)\ , \tag{6.2}$$

so that now there is a double root α near the double root $\hat{x}$. For a relative distance ε of about $\sqrt[4]{\mathbf{u}} \sim 10^{-4}$ the four roots behave like a quadruple root. This is confirmed by the results in Table 2.

Note that in both cases the inclusions of the offset for the regularization are very accurate inclusions of zero.

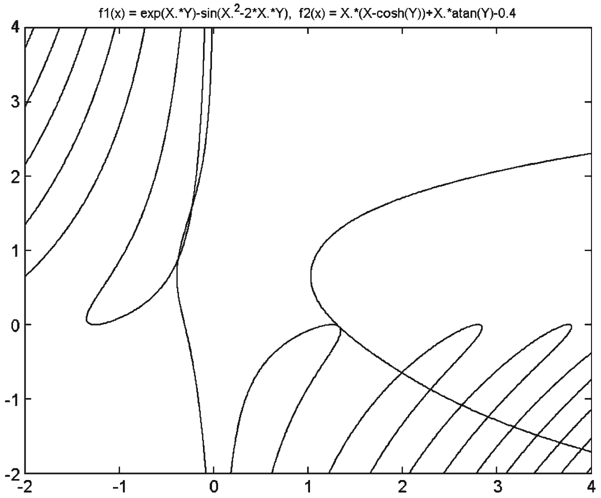

Fig. 3 Zero contour lines of $f(x_1, x_2)$ as defined in (6.3) for $\alpha = 0.4$

Table 3 Inclusions X_1, X_2 for two single roots and X for a nearly double root for f as in (6.3) and $\alpha = 0.4003120447407$

X_1	X_2	X	E
1.328899621^{86}_{28}	1.32889951^{57}_{48}	1.32889956839071^{6}_{5}	
-0.02729805^{67}_{59}	-0.02729792^{98}_{88}	$-0.02729799275879^{41}_{34}$	$[-5.2, -5.0] \cdot 10^{-14}$

Next we test some systems of nonlinear equations. The first test function is

$$f(x_1, x_2) = \begin{pmatrix} e^{x_1 x_2} - \sin\left(x_1^2 - 2x_1x_2\right) \\ x_1(x_1 - \cosh(x_2)) + x_1 \operatorname{atan}(x_2) - \alpha \end{pmatrix} = 0\,, \tag{6.3}$$

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 Fig. 3. For $\alpha = 0.40031204474074$ there is an almost double root near (1.329, −0.0273). The parameter α is chosen such that we can just separate the nearly double root into two single roots. The results are displayed in Table 3.

The inclusions of the two simple roots are separated by about 10^{-7} which is almost $\sqrt{\mathbf{u}}$, and the quality of the inclusion is about $\sqrt{\mathbf{u}}$ as well. Subtracting a constant $\varepsilon \in E$ from the first equation generates a truly double root. Note that $|\varepsilon| < 6 \cdot 10^{-14}$. As expected the quality of the inclusion of the double root is much better than those of the separated simple roots, almost of maximum accuracy.

For $\alpha = 0.35653033083794$ there is another almost double root of f as in (6.3) near (−0.292, 1.195). Again the parameter α is chosen such that we can just separate the nearly double root into to two single roots. The results are displayed in Table 4. The quality of the results is very similar to the previous example, see Figure 4.

Finally we show examples of higher dimensions. Consider Brown's almost linear function $f : \mathbb{R}^n \to \mathbb{R}^n$ with [12]

$$\begin{aligned} f_k(x) &= x_k + \textstyle\sum_{j=1}^n x_j - (n+1) \quad \text{for} \quad 1 \le k \le n-1\,, \\ f_n(x) &= \left(\textstyle\prod_{j=1}^n x_j\right) - 1 - e\,, \end{aligned} \tag{6.4}$$

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

$$e = \left(1 - \frac{1}{n^2}\right)^{n-1} \left(1 + \frac{1}{n}\right) - 1 \tag{6.5}$$

Table 4 Inclusions X_1, X_2 for two single roots and X for a nearly double root for f as in (6.3) and $\alpha = 0.35653033083794$

X_1	X_2	X	E
-0.29197330^{91}_{44}	-0.2919733^{61}_{57}	$-0.29197333312764^{41}_{29}$	
1.1950051^{23}_{00}	1.1950048^{69}_{53}	$1.1950049857509^{92}_{87}$	$[-1.17, -0.96] \cdot 10^{-14}$

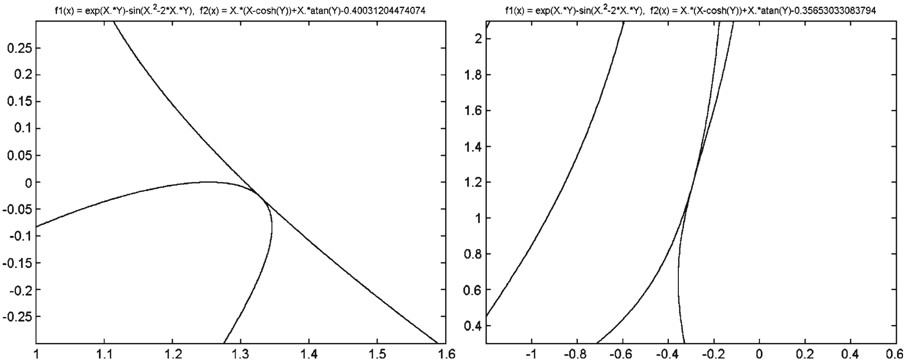

Fig. 4 Zero contour lines of $f(x_1, x_2)$ as defined in (6.3) for two different parameter values α

and

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

the vector $(1, \ldots, 1, -n)$ is in the kernel of the Jacobian of f as in (6.4). Thus $\bar{x}$ is not a simple root of f. Inclusions for different dimensions n are displayed in Table 5. More precisely it is verified that there exists $\hat{x} \in X$ and $\hat{\varepsilon} \in E$ such that $f(\hat{x}) - (\hat{\varepsilon}, \ldots, 0) = 0$ and the Jacobian $J_f(\hat{x})$ of f at $\hat{x}$ is singular.

The inclusions for the first $n-1$ components of X are identical. As can be seen the accuracy of the inclusions decreases slowly with increasing dimension. All inclusions including that of the regularization parameter ε are of remarkable quality. Note that for all computations the same algorithm `verifynlss2` in INTLAB has been used without change. This code is available in INTLAB, see http://www.ti3.tu-harburg.de/rump.

Table 5 Inclusions of a double root of (6.4) for different dimensions

n	$X_{1\cdots n-1}$	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	$0.999999 \pm 7.5 \cdot 10^{-11}$	$1.001000 \pm 2 \cdot 10^{-13}$	$[-1.1, 6.4] \cdot 10^{-11}$

7 Conclusion

In this paper we provided 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. We have applied those to univariate polynomials, to multivariate polynomials and also to eigenvalue problems. Numerical experiments have confirmed the performance of our algorithms.

Acknowledgement The authors wish to thank Prof. Jean Vignes from Paris VI for his constructive comments.

References

1. Alefeld, G., Spreuer, H.: Iterative improvement of componentwise errorbounds for invariant subspaces belonging to a double or nearly double eigenvalue. Computing **36**, 321–334 (1986)
2. Alt, R., Vignes, J.: Stabilizing Bairstow's method. Comput. Math. Appl. **8**(5), 379–387 (1982)
3. Frommer, A., Lang, B., Schnurr, M.: A comparison of the moore and miranda existence test. Computing **72**(3–4), 349–354 (2004)
4. Higham, N.J.: Accuracy and Stability of Numerical Algorithms, 2nd edn. SIAM, Philadelphia (2002)
5. Kanzawa, Y., Oishi, S.: Calculating bifurcation points with guaranteed accuracy. IEICE Trans. Fundam. **E82-A**(6), 1055–1061 (1999)
6. Kanzawa, Y., Oishi, S.: Imperfect singular solutions of nonlinear equations and a numerical method of proving their existence. IEICE Trans. Fundam. **E82-A**(6), 1062–1069 (1999)
7. Krawczyk, R.: Newton-algorithmen zur bestimmung von nullstellen mit fehlerschranken. Computing **4**, 187–201 (1969)
8. La Porte, M., Vignes, J.: Étude statistique des erreurs dans l'arithmétique des ordinateurs; application au controle des resultats d'algorithmes numériques. Numer. Math. **23**, 63–72 (1974)
9. La Porte, M., Vignes, J.: Méthode numérique de détection de la singularité d'une matrice. Numer. Math. **23**, 73–81 (1974)
10. MATLAB User's Guide, Version 7. The MathWorks Inc (2004)
11. Moore, R.E.: A Test for Existence of Solutions for Non-Linear Systems. SIAM J. Numer. Anal. (SINUM) **4**, 611–615 (1977)
12. Moré, J.J., Cosnard, M.Y.: Numerical solution of non-linear equations. ACM Trans. Math. Software **5**, 64–85 (1979)
13. Nakao, M.R.: Numerical verification methods for solutions of ordinary and partial differential equations. Numer. Funct. Anal. Optim. **33**(3/4), 321–356 (2001)
14. Neumaier, A.: Introduction to Numerical Analysis. Cambridge University Press, Cambridge (2001)
15. Plum, M., Wieners, Ch.: New solutions of the gelfand problem. J. Math. Anal. Appl. **269**, 588–606 (2002)
16. Poljak, S., Rohn, J.: Checking robust nonsingularity is np-hard. Math. of Control, Signals, and Systems **6**, 1–9 (1993)
17. Rump, S.M.: Solving algebraic problems with high accuracy. Habilitationsschrift. In: Kulisch, U.W., Miranker, W.L. (eds.) A New Approach to Scientific Computation, pp. 51–120. Academic, New York (1983)
18. Rump, S.M.: INTLAB - INTerval LABoratory. In: Csendes, T. (ed.). Developments in Reliable Computing. pp. 77–104. Kluwer, Dordrecht (1999)
19. Rump, S.M.: Computational Error Bounds for Multiple or Nearly Multiple Eigenvalues. Linear Algebra Appl. (LAA) **324**, 209–226 (2001)

20. Rump, S.M. Ten methods to bound multiple roots of polynomials. J. Comput. Appl. Math. (JCAM) **156**, 403–432 (2003)
21. Rump, S.M., Zemke, J.: On eigenvector bounds. BIT Numer. Math. **43**, 823–837 (2004)
22. Vignes, J.: Algorithmes numériques, analyse et mise en œuvre. 2. Éditions Technip, Paris, 1980. Équations et systèmes non linéaires. [Nonlinear equations and systems], With the collaboration of René Alt and Michèle Pichat, Collection Langages et Algorithmes de l'Informatique.