SOLVING NONLINEAR SYSTEMS VIA LINEAR PARAMETRIC MODELS

Lubomir Kolev, Georgi Tsenov

Department of Theoretical Electrical Engineering, Technical University of Sofia,
bd. Kl. Ohridski 8, 1000 Sofia, Bulgaria,
e-mail: lkolev@tu-sofia.bg, gogotzenov@tu-sofia.bg

Abstract: *Solving nonlinear systems of equations is a standard problem arising in the analysis of nonlinear dc circuits. The most popular method for global solution of nonlinear systems is the well-known interval Newton method using the interval Jacobean matrix of the system or the so-called interval slope matrix. In the present poster, a better approach is suggested which is based on the use of corresponding linear parametric models.*

Keywords: *Nonlinear DC circuits, Numerical analysis, Solving nonlinear systems*

1. INTRODUCTION

Solving nonlinear systems of equations is a standard problem arising in the global analysis of nonlinear dc circuits. More specifically, the problem considered is:
solve globally

$$f(x) = 0 \, , \tag{1.a}$$
$$x \in \boldsymbol{x}^0 \subset R^n. \tag{1.b}$$

Global nonlinear dc circuit analysis is guaranteed if interval methods are used [1].

The most popular interval method for global solution of nonlinear systems is the well-known interval Newton method (or its versions): an iterative method using interval extension $\boldsymbol{J}(\boldsymbol{x})$ of the Jacobean matrix (or some modifications) for the currant box $\boldsymbol{x}$ belonging to $\boldsymbol{x}^0$ and related to a given iteration. The following interval system

$$\boldsymbol{A}(\boldsymbol{x})\,y = f(x) \tag{2}$$

is to be solved at each iteration where $\boldsymbol{A}(\boldsymbol{x})$ is an interval matrix (standing for the interval Jacobean matrix, interval slope matrix, the Hansen-Sengupta operator or some other modification) while $f(x)$ is a real vector.

An alternative approach was suggested in [2] using the following approximation

$$f(y) \in A(\boldsymbol{x})x + \boldsymbol{b}(\boldsymbol{x}), \quad x \in \boldsymbol{x} \tag{3}$$

where $A(\boldsymbol{x})$ is a real matrix while $\boldsymbol{b}(\boldsymbol{x})$ is an interval vector. The right side in (3)

$$L(\boldsymbol{x}) = A(\boldsymbol{x})x + \boldsymbol{b}(\boldsymbol{x}), \qquad x \in \boldsymbol{x}, \tag{4}$$

is known as a *linear interval approximation* (LIA). In this case solving (1) reduces to repeatedly solving the interval system

$$A(\boldsymbol{x})y = -\boldsymbol{b}(\boldsymbol{x}) \tag{5}$$

where, unlike (2), $A(\boldsymbol{x})$ is a real matrix. This determines the better performance of the LIA approach as compared to the standard approach (see [3]).

A new approach for global nonlinear dc circuit analysis, i.e. solving (1), will be suggested here (Section 2). It is based on various alternative approximations of f in $\boldsymbol{x}$ which are in linear parametric form. Now, we obtain a linear parametric system

$$A(p)(y-x)=-f(x), \quad p \in \boldsymbol{p} \tag{6a}$$

or, equivalently

$$A(p)z=b(x), \quad p \in \boldsymbol{p} \tag{6b}$$

where $A(p)$ is a parametric matrix. It will be shown that (6) is a better way to bound the solutions of (1) than the interval Newton method. System (6) will be referred to as a *linear parametric model* of *f(x)* in $\boldsymbol{x}$.

The linear parametric model (6) will be extended to the analysis of nonlinear circuits containing resistors or other parameters given as intervals (Section 3).

2. SYSTEMS OF NONLINEAR (NONPARAMETRIC) EQUATIONS

2.1. Linear parametric approximation using slopes

This is yet another alternative linearization of nonlinear functions suggested in a different context for the first time in [3]. The novel approach is based on the use of the slope matrix *S(y,x)* and the equality

$$f(y)=f(x)+S(y,x)(y-x), \tag{7}$$

where *y* and *x* have some fixed values (typically, *x* is a known solution x^0 of (1)). We now "free" the components y_k of *y* and consider them as components of a parameter vector *p*, i.e.

$$p = (y_1,\ldots,y_n) \in \boldsymbol{x} = (\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n) . \tag{8}$$

Let

$$a_{ij}(p) = S_{ij}(p,x) \tag{9}$$

be the entries of the parametric matrix $A(p)$. On account of (7) to (9)

$$y \in f(x)+A(p)(y-x), \quad p \in \boldsymbol{x} . \tag{10}$$

The right-hand side of (10) is the novel *linear parametric approximation* (LPA) of *f* in $\boldsymbol{x}$. If *y* is a zero of *f*, then

$$A(p)(y-x)=-f(x), \quad p \in \boldsymbol{x} \tag{11a}$$

or, equivalently

$$A(p)z=b(x), \quad p \in \boldsymbol{x} . \tag{11b}$$

Thus, using the novel approximation, the *linear parametric model* (11) is obtained. Following [3], we show that (11*) is a better way than (2) to bound the solutions of (1)*. Indeed, introduce the solution sets

$$S_J = \{z:\ J\,z = b\ ,\ \ J \in \boldsymbol{J}(\boldsymbol{x})\}\ , \tag{12}$$

$$S_p = \{z:\ A(p)z = b\ ,\ \ p \in \boldsymbol{x}\}\ . \tag{13}$$

It is seen that while $\boldsymbol{J}(\boldsymbol{x})$ depends on n^2 independent entries, there are only n independent elements in $A(p)$. Moreover, the methods for enclosing S_p account for the interdependencies between the components of $A(p)$. Thus, it follows from (12) and (13) that

$$S_p \subset S_J\ . \tag{14}$$

If z_{out}^J and z_{out}^p [4] denote some outer interval solution of (2) and (11b), respectively, then we can expect that also

$$Z_{\text{out}}^p \subset Z_{\text{out}}^J \tag{15}$$

but (15) is, however, not guaranteed.

2.2. LPA using the Hansen-Sengupta operator in parametric form

The approach of § 2.1 is applicable only if $S(y, x)$ the slope matrix is available in analytical form. If this is not the case, then the Jacobian matrix in parametric form $J(p)$ can be used as suggested in [3]. Thus, (9) is replaced with

$$a_{ij}(p) = J_{ij}(p_1,,,p_n)\ ,\ \ p_i \in \boldsymbol{p}_i = \boldsymbol{x}_i,\ \ i = 1,,,n\ . \tag{16}$$

It is seen that each element $a_{ij}(p)$ depends on all n parameters p_i. A better LPA is suggested here which is based on the Hansen-Sengupta operator [5]. In its standard (nonparametric) form, it encloses each function $f_i(y)$ by the following expression

$$f_i(y) \in f_i(x) + \sum_{j=1}^{n}\left(y_j - x_j\right) g_{ij}\left(\boldsymbol{x}_1,...,\boldsymbol{x}_j;x_{j+1},...,x_n\right). \tag{17}$$

We now write (17) in parametric form

$$f_i(y) \in f_i(x) + \sum_{j=1}^{n}\left(y_j - x_j\right) g_{ij}\left(p_1,...,p_j;x_{j+1},...,x_n\right),\ \ p_1 \in \boldsymbol{x}_1,...,p_j \in \boldsymbol{x}_j. \tag{18}$$

Hence, using the Hansen-Sengupta operator, we have to replace (16) with

$$a_{ij}(p) = g_{ij}(p_1,,,p_j)\ ,\ \ p_k \in \boldsymbol{p}_k = \boldsymbol{x}_k,\ \ k = 1,..,j\ . \tag{19}$$

It is seen that, unlike (16) where all n parameters are intervals, now a fraction $(1/2)(1-1/n)$ in (18) are real parameters. This determines the better performance of the Hansen-Sengupta LPA as compared to the Jacobian LPA.

2.3. Checking uniqueness

Another advantage of the LPA strategy over the standard strategy is in checking the uniqueness of a solution in $\boldsymbol{x}$. The standard approach is to check if

$$N(\boldsymbol{x}) \in \text{int}(\boldsymbol{x}) \tag{20}$$

is valid ($N(\boldsymbol{x})$ can be the Newton, Gauss-Seidel, Krawczyk, Hansen-Sengupta operator). It is to be stressed that no additional assumption is needed. The new approach is based on the assumption that *we have already established the presence of a solution x* in x and it only remains to check for uniqueness of x* in x*. Very often in practice, this is really the case.

Theorem 1. Let f (x) be a continuously differentiable vector function. Let A(p) be the parametric matrix defined in x by (16). If x* is in x, f (x*) = 0 and A(p) is a regular matrix in $\boldsymbol{p} = \boldsymbol{x}$, then x* is the unique solution of f (x) = 0 in x.

The proof is based on Theorem 13.6.1 in [5].

Thus, we can *prove uniqueness by checking regularity (strong regularity) of A(p)* in $\boldsymbol{x}$. A better choice is to check regularity of the slope matrix $S(p, x^*)$, $p \in \boldsymbol{x}$ by some method (a simple such test for regularity is given by Theorem 4, Section 3.2 in [3] but any better check for regularity of parametric matrices could be used.

3. SYSTEMS OF NONLINEAR PARAMETRIC EQUATIONS

We now extend some of the ideas considered earlier to the parametric case

$$f(x, p) = 0 , \tag{21a}$$

$$p \in \boldsymbol{p} \subset R^m , \tag{21b}$$

$$x \in \boldsymbol{x}^0 \subset R^n . \tag{21c}$$

Assumption 1. A pair (x^0, p^0) satisfying (21b), (21c) is known which is a zero of f in (21a).

The vector p^0 is usually the centre of $\boldsymbol{p}$ and x^0 is the solution of (21a) for $p = p^0$.

The solution set of (21a), (21b) is the set

$$S_f(\boldsymbol{p}) := \{x : f(x, p) = 0, \ p \in \boldsymbol{p}\} . \tag{22}$$

The interval hull of $S_f(\boldsymbol{p})$ will be denoted $\boldsymbol{x}^*$; any other interval $\boldsymbol{x}$ such that $\boldsymbol{x}^* \in \boldsymbol{x}$ is referred to as an interval (outer) bound on $S_f(\boldsymbol{p})$.

3.1. A basic problem

A basic problem is to determine $\boldsymbol{x}$ for a given f and $\boldsymbol{p}$. Finding $\boldsymbol{x}$ has been considered as a sensitivity analysis problem associated with (19a), (19b) (e.g., [5]). Various methods and algorithms are based on the parameterized versions of the Newton method and its variants. Now

$$f(y, p) \in f(x) + \boldsymbol{J}(\boldsymbol{x}, \boldsymbol{p})(y - x) \tag{23}$$

is used which becomes

$$\boldsymbol{J}(\boldsymbol{x}, \boldsymbol{p})(y - x) = -f(x, p) \tag{24}$$

if (y, p) is a zero of f. Another alternative idea is based on the use of (3) and (4) for linearizing non-linear functions applied to the function $f(u)$ when $u = (x, p)$ as well as on Assumption 1. Reference [6] seems to be the only paper where this approach has been developed to offer a method for determining $\boldsymbol{x}$.

3.2. A new linear parametric approximation (LPA) approach

This approach to solving the basic problem is reported for the first time here. In the case of the parametric equation (19a), for a fixed p formula (7) becomes

$$f(y, p) = f(x, p) + S(y, x, p)(y - x) \tag{25}$$

where most often x is the centre of $\boldsymbol{x}$. Once again, we "free" the components y_k of y to take on values in $\boldsymbol{x}$. Thus, we introduce the additional parameter vector

$$q = (y_1,\ldots,y_n) \in \boldsymbol{x} = (\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n). \tag{26a}$$

and let

$$a_{ij}(q, p) = S_{ij}(q, x, p) \tag{26b}$$

be the entries of the parametric matrix $A(q, p)$, $q \in \boldsymbol{x}$, $p \in \boldsymbol{p}$ (x is fixed). For each $p \in \boldsymbol{p}$

$$f(y, p) \in f(x) + A(q, p)(y - x), q \in \boldsymbol{x}, p \in \boldsymbol{p}. \tag{27}$$

If (y, p) is zero of f

$$A(q, p)z = b(p), \quad q \in \boldsymbol{x}, \quad p \in \boldsymbol{p} \tag{28}$$

where $z = y - x$ and $b(p) = -f(x,p)$. The linear parametric system (28) is the new LPA model suggested here to tackle the problem of obtaining an outer approximation $\boldsymbol{y}$ of the solution set $S_f(\boldsymbol{p})$. Indeed, consider the sets

$$S_{JP} = \{z: \quad J\,z = b, \quad J \in \boldsymbol{J}(\boldsymbol{x}, \boldsymbol{p}\}, \tag{29}$$

$$S_{pq} = \{z: A(q, p)z = b(p), q \in \boldsymbol{x}\ p \in \boldsymbol{p}\}. \tag{30}$$

Clearly,

$$S_{pq} \subset S_{JP} \tag{31}$$

($\boldsymbol{J}(\boldsymbol{x}, \boldsymbol{p}$) has n^2 independent entries, each being an interval extension of the function $J(x, p)$ of $n + m$ arguments, while there are only $n + m$ dependent elements in $A(q, p)$ and m dependent elements in $b(p)$.) Hence we can expect (28) to be a better model than (24).

If the slopes $S_{ij}(q, x, p)$ cannot be found in analytical form, then they should be replaced with the components $g_{ij}(q, p)$. In that case, the elements of the parametric matrix $A(q, p)$ are given not by (27) but as follows $a_{ij}(q, p) = g_{ij}(q_1,\ldots,q_j; p)$.

3.3. Determining an outer solution

Let $N_f(x, p)$ denote an outer interval solution of

$$A(q, p)(y - x) = b(p), \quad q \in x, \quad p \in p. \tag{32}$$

Consider the following iteration process

$$x^{(k+1)} = N(x^{(k)}, p; \varepsilon_x^{(k)}, \varepsilon_p^{(k)}), \quad x^{(0)} = x^0, \quad k \geq 0 \tag{33}$$

where x^0 is the solution of f corresponding to p^0 (the centre of p) while $\varepsilon_x^{(k)}, \varepsilon_p^{(k)}$ are some parameters introduced to accelerate the convergence of the iterations. The iterations are terminated (if possible) as soon as

$$x^{(k+1)} \subset \mathrm{int}(x^{(k)}). \tag{34}$$

The outer solution x^b (containg $S_f(p)$) is given by $x^{(k+1)}$. This is a rather general algorithm. Thus, according to [6] $\varepsilon_p^{(k)} = 1$, $\varepsilon_x^{(k)} = 1$ until stationarity $x^{(k+1)} = x^{(k)}$ is reached for some iteration number $\overline{k}$ and only then $\varepsilon_x^{(\overline{k}+1)} = 1 + \varepsilon$ with $\varepsilon = 0.05$ for the examples considered. Perhaps, it would be useful to try with

$$\varepsilon_x^{(k+1)} = 1 + [1/(1 + \varepsilon_1 k)]\varepsilon_2 \tag{35}$$

inflating $x^{(k+1)}$ more drastically for the initial iterates.

A similar approach would be to vary $\varepsilon_p^{(k+1)}$ as in (35) (keeping $\varepsilon_x^{(k)} = 1$) or to let both $\varepsilon_x^{(k+1)}$ and $\varepsilon_p^{(k+1)}$ change.

3.4. Uniqueness

Some of the methods for determining x^b guarantees uniqueness (i.e., $x(p)$ is unique in x^b for any $p \in p$); others do not. In the latter case an additional test for uniqueness is needed. Here it is assumed that (i) Assumption 1 holds, (ii) an interval vector x is known to contain the interval hull x^* of (21). The problem is to establish whether any other zero (x, p), $p \neq p^0$ is unique in x for each $p \in p$.

Theorem 2. Let the above assumptions be valid. Assume additionally that $f(x, p)$ is continuously differentiable in both $x \in x$ and $p \in p$. Let $J(x, p)$ denote the derivation of f wrt x (the Jacobian) while $A(q, p)$ is the parametric matrix which elements are defined by $A(q, p) = J_{ij}(q, p)$. If $A(q, p)$, $q \in x$, $p \in p$ is a regular interval parametric matrix, then:

a) there exists a continuous function $x = g(p)$, $p \in p$ such that each $x \in S_f(p)$;

b) $S_f(p)$ is a connected set in x.

The theorem improves on a result due to Hansen: Theorem 17.6.1 and Corollary 17.6.2 in [5].

Remark. The results obtained in § 2 and § 3 could be applied to the case of underdetermined systems (more variables than equations).

References

[1] L. Kolev, Interval Methods for Circuit Analysis, World Scientific, Singapore, New Jersey, London, 1993.

[2] L. Kolev. A general interval method for global nonlinear dc analysis. *Proc. of the 1997 European Conf. on Circuit Theory and Design, ECCD'97*, Technical University of Budapest, 30th August - 3rd Sept., 1997, Budapest, Hungary, volume 3, pp.1460-1462.

[3] L. Kolev. Some ideas towards global optimization of improved efficiency. Talk at the GICOLAG workshop, 4-9 November, Wien, Austria, 2006.

[4] L. Kolev. Componentwise Determination of the Interval Hull Solution for Linear Interval Parameter Systems. Reliable Computing 20 (2014), pp. 1-24.

[5] E. Hansen. Global optimization using interval analysis. Marcel Dekker Inc. (1992). L. Kolev and I. Nenov.

[6] L. Kolev and I. Nenov. Cheap and tight bounds on the solution set of perturbed systems. *Reliable Computing*, vol.7, 2001, pp. 399-408.

Reviewer: Assoc. Prof. PhD S. Petrakieva