

Zygmunt GARCZARZYK

O ZNAJDOWANIU WSZYSTKICH ROZWIĄZAŃ PEWNYCH NIELINIOWYCH OBWODÓW BEZINERCYJNYCH METODAMI INTERWAŁOWYMI

Streszczenie. Analiza interwałowa i uogólniona metoda bisekcji są podstawą algorytmów znajdowania wszystkich rozwiązań układu równań hybrydowych nieliniowych obwodów bezinercyjnych. W algorytmach opartych na tych technikach konieczna jest estymacja interwałowej reprezentacji macierzy Jacobiego lub funkcji nieliniowych. W pracy rozważa się ten problem wykorzystując rozwinięcia Taylora funkcji jednej i dwóch zmiennych oraz wielomiany Bernsteina. Do generacji rozwinięć Taylora zastosowano metody różniczkowania automatycznego.

ON FINDING ALL SOLUTIONS OF SOME NONLINEAR INERTIALESS CIRCUITS VIA INTERVAL METHODS

Summary. Interval analysis in conjunction with generalized bisection form the basis of algorithms that find all solutions of a system of hybrid equations of nonlinear inertialess circuits. In algorithms based on those techniques it is necessary to estimate an interval evaluation of Jacobian matrix or nonlinear functions. In the paper we study this problem using Taylor expansions of univariate and bivariate functions and Bernstein polynomials. Taylor expansions are generated with use of automatic differentiation methods.

ÜBER DIE FINDUNG ALLER LÖSUNGEN VON MANCHEN NICHTLINEAREN, NICHTINERTIALEN STROMKREISEN MITTELS INTERVALLMETHODEN

Zusammenfassung. Die Intervallanalyse und die verallgemeinerte Bisectionmethode sind die Grundlage für die Algorithmen bei der Findung aller Lösungen der hybriden Gleichungssystems von nichtlinearen, nichtinertialen Stromkreisen. Für diese Algorithmen ist die Schätzung von Intervallrepräsentation der Jakobische Matrix oder nichtlinearen Funktionen erforderlich. Es wird auf die Anwendung der Taylorentwicklung der Funktion einer und zweier Variablen sowie auf Bernstein-Polynome hingewiesen. Zur Erzeugung der Taylorentwicklung verwendete man automatische Differentiationsmethoden.

1. WPROWADZENIE

Metody interwałowe wraz z uogólnioną bisekcją mogą stanowić podstawę algorytmów znajdowania wszystkich rozwiązań w określonym obszarze $D \subseteq \mathbb{R}^n$ dla układu równań obwodu nieliniowego. W algorytmach opartych na tych metodach konieczna jest często estymacja macierzy Jacobiego równania obwodu. Problem ten sprowadza się do wyznaczenia granic zakresów funkcji, będących elementami macierzy, w zadanym interwale (przedziale) [4],[6],[7]. W pracy rozważa się to zagadnienie dla nieliniowych równań hybrydowych obwodu [1],[2]:

$$F(z) = f(z) - Hz - s = 0. \quad (1)$$

Równanie (1) opisuje obwody bezinercyjne zbudowane z liniowych, ewentualnie sprzężonych elementów rezystancyjnych (np. liniowych źródeł sterowanych, transformatorów, żyratorów itp.) i dwójników nieliniowych. W równaniu (1) $z \in \mathbb{R}^n$ jest nieznanym wektorem reprezentującym napięcia i prądy gałęzi z elementami nieliniowymi, H jest macierzą $n \times n$ liniowego n-wrotnika uzyskanego po usunięciu z obwodu wszystkich elementów nieliniowych, s oznacza wektor źródeł, który uwzględnia obecność w obwodzie źródeł niezależnych, $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$ jest wektorem charakterystyk elementów nieliniowych. Charakterystyki elementów są zadane jako dostatecznie gładkie funkcje $f_i: D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ lub $f_i: D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$, w szczególności jako wielomiany jednej oraz dwóch zmiennych. Pozwala to uwzględnić obecność w obwodzie nieliniowych elementów sterowanych. Elementy nieliniowe mogą mieć charakterystyki lokalnie pasywne lub aktywne, a więc obwód może posiadać wiele rozwiązań (punktów równowagi). Przyjmuje się, że rozwiązania tworzą zbiór ograniczony.

Podstawą przedstawionych w pracy rozważań jest analiza interwałowa. Podstawowe własności działań na interwałach można znaleźć w pracach [3],[4],[17],[18]. Zbiory wektorów rzeczywistych, macierzy rzeczywistych, wektorów interwałowych i macierzy interwałowych oznaczono odpowiednio małymi i dużymi literami oraz małymi i dużymi pogrubionymi literami.

Dla rzeczywistego interwału $z = [z, \bar{z}]$, $z \leq \bar{z}$, definiuje się

$$\text{moduł } |z| := \max(|z|, |\bar{z}|), \text{ } := \max(),$$

$$\text{szerokość } w(\mathbf{z}) := \bar{\mathbf{z}} - \mathbf{z}, \quad (2)$$

$$\text{środek } m(\mathbf{z}) := (\mathbf{z} + \bar{\mathbf{z}})/2.$$

Definiuje się także odległość między interwałami $\mathbf{x} = [\underline{x}, \bar{x}]$ i $\mathbf{y} = [\underline{y}, \bar{y}]$

$$q(\mathbf{x}, \mathbf{y}) := \max(|\underline{x} - \underline{y}|, |\bar{x} - \bar{y}|) \quad (3)$$

Dla wektorów i macierzy interwałowych pojęcia te są zdefiniowane przez ich elementy. Na

przykład, jeśli $\mathbf{A} = (\mathbf{a}_{ij})_{i,j=0}^n$ jest macierzą interwałową, to $w(\mathbf{A}) = (w(\mathbf{a}_{ij}))_{i,j=0}^n$.

W pracy [7] przedstawiono następującą metodę obliczania n-wymiarowego obszaru $D \subseteq \mathbb{R}^n$ zawierającego rozwiązanie \mathbf{z}^* równania (1):

$$\mathbf{b}^k := F(\mathbf{m}_k) + (\mathbf{A}_k - \mathbf{f}'(\mathbf{z}^k) + \mathbf{H})(\mathbf{m}_k - \mathbf{z}^k), \quad (4a)$$

$$\mathbf{z}^{k+1} := (\mathbf{m}_k - \mathbf{d}^k) \cap \mathbf{z}^k, \quad k = 0, 1, \dots \quad (4b)$$

gdzie $\mathbf{m} \in \mathbf{z}^k$ (np. $\mathbf{m} = m(\mathbf{z}^k)$), \mathbf{A}_k jest M-macierzą, a $\mathbf{f}'(\mathbf{z}^k)$ jest interwałową reprezentacją pochodnych f na zadanym wektorze interwałowym \mathbf{z}^k .

Składowe wektora interwałowego \mathbf{d}^k są równe:

$$d_i = [v_i - h_i/2, v_i + h_i/2], \quad i = 1, \dots, n, \quad (5)$$

gdzie v_i i h_i są odpowiednio rozwiązaniami układów równań liniowych

$$\mathbf{A}_k \mathbf{v} = \mathbf{m}(\mathbf{b}^k) \text{ i } \mathbf{A}_k \mathbf{h} = \mathbf{w}(\mathbf{b}^k) \quad (6)$$

Algorytm (4) w połączeniu z bisekcją wybranego (dopuszczalnego) obszaru $\mathbf{z}^0 \subseteq D \subseteq \mathbb{R}^n$ pozwala lokalizować rozwiązania równania (1).

2. APROKSYMACJA MACIERZY JACOBIEGO

Zasadniczym problemem przy obliczeniu \mathbf{z}^{k+1} jest szacowanie zakresów wartości pochodnych charakterystyk $f_i(\mathbf{z})$, $i = 1, 2, \dots, n$ w zadanym interwale $\mathbf{z} = \mathbf{z}^k$.

Dla dowolnej funkcji f , która posiada reprezentację interwałową, zakres wartości funkcji f w interwale \mathbf{z}

$$R(f, \mathbf{z}) := \{f(\mathbf{z}) \mid \mathbf{z} \in \mathbf{z}\} \quad (7)$$

jest zawarty w oszacowaniu funkcji $f(\mathbf{z})$ obliczonym w arytmetyce interwałowej

$$R(f, \mathbf{z}) \subseteq f(\mathbf{z}) \quad (8)$$

Na dodatek jest on mocno uzależniony od postaci wyrażenia arytmetycznego użytego do interwałowego szacowania wartości funkcji. Konieczne jest więc użycie bardziej wyszukanych metod.

Zagadnienie to przedstawiono szczegółowo w pracy [7]. Dla jasności wyводу przedstawimy tu podstawowe idee i rozszerzymy rozważania na zagadnienie obliczania wartości pochodnych wyższych rzędów charakterystyk rezystorów.

Dla charakterystyk rezystancyjnych dostatecznie gładkich w interwale z każdy element macierzy diagonalnej $f(z)$ ma rozwinięcie Taylora postaci

$$f_i(z_i) = u(z) = p(z) + r(z), \quad (9)$$

$$\text{gdzie } p(z) = \sum_{l=0}^{\sigma} \alpha_l z^l, \text{ przy czym } \alpha_l = \sum_{k=l}^{\sigma} (u)_k \binom{k}{l} (-z_0)^{k-l} \quad (10)$$

$$\text{oraz } r(z) = u^{(\sigma+1)}(z_0 + \xi(z-z_0))(z-z_0)^{\sigma+1} / (\sigma+1)! \quad (11)$$

$$\xi \in [0, 1], \quad z_0 \in z \text{ (np. } z_0 = m(z)).$$

Symbol $(u)_k$ we wzorze (10) oznacza współczynniki Taylora

$$(u)_k := \frac{1}{k!} \cdot \frac{d^k u}{dt^k}(z_0) := \frac{1}{k!} \cdot u^{(k)}(z_0), \quad \text{dla } k \geq 0. \quad (12)$$

Jeżeli $f_i(z_i) = u(z) : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ jest $\sigma+1$ krotnie różniczkowalna, to dla wszystkich $z \in D$ zachodzi [7]:

$$\text{(inkluzja)} \quad R(u, z) \subseteq V(u, z) := R(p, z) + u^{(\sigma+1)}(z)w(z)^{\sigma+1} / (\sigma+1)!, \quad (13)$$

$$\text{(odległość)} \quad q(R(u, z), V(u, z)) \leq \gamma w(z)^{\sigma+1}, \quad \gamma \geq 0, \quad (14)$$

gdzie $R(p, z)$ oznacza dokładny zakres wartości wielomianu $p(z)$ w interwale z .

Relacja (13) określa sposób obliczania zakresu; należy wyznaczyć zakres wielomianu oraz zakres składnika reszty. Z relacji (14) wynika, że przekroczenie zakresu $R(u, z)$ przez zakres $V(u, z)$ zmniejsza się wraz z potęgą $\sigma+1$ szerokości interwału $w(z)$, jeżeli $w(z) \leq 1$, zatem jeśli pochodna $u^{(\sigma+1)}(z)$ jest ograniczona, to można pominąć składnik reszty w $V(u, z)$ i wtedy $R(u, z) \approx R(p, z)$.

Estymatory dla maksimum i odpowiednio minimum wielomianu $p(z)$ w interwale z otrzymuje się obliczając współczynniki Bernsteina. Mamy bowiem

$$\min_{0 \leq j \leq v} b_j^{(v)} \leq \min_{z \in z} p(z) \leq \max_{z \in z} p(z) \leq \max_{0 \leq j \leq v} b_j^{(v)}, \quad (15)$$

$$\text{gdzie } v \geq \sigma, \text{ a } b_j^{(v)} = \sum_{s=0}^j \sum_{t=0}^{\sigma} \binom{t}{s} \alpha_t z^{t-s} w(z)^s \binom{j}{s} \binom{v-j}{s}, \quad j = 0, 1, \dots, v. \quad (16)$$

Estymatory te są zbieżne do górnej i dolnej granicy zakresu $p(z)$ w z dla $v \rightarrow \infty$ [5].

Estymatory te są zbieżne do górnej i dolnej granicy zakresu $p(z)$ w z dla $v \rightarrow \infty$ [5]. Współczynniki $b_j^{(v)}$ oblicza się rekursywnie korzystając z tablicy różnic skończonych [8]. Z powyższego widać, że wyznaczenie zakresu wartości $p(z)$ wymaga informacji o wartościach $(u)_k$. Stosując rekursywne zależności wynikające z metod różniczkowania automatycznego [10],[11], uzyskuje się szybko i dokładnie wartości współczynników Taylora.

Niech u, v, w oznaczają rzeczywiste funkcje, analityczne w otoczeniu z_0 ($\sigma+1$ -krotna różniczkowalność, dla dostatecznie dużego σ , także wystarcza).

Szereg Taylora funkcji u punkcie z_0 można zapisać następująco:

$$u(z) = \sum_{k=0}^{\infty} (u)_k (z - z_0)^k \quad (17)$$

Dla funkcji zdefiniowanych poprzez działania arytmetyczne na innych funkcjach, można znaleźć następujące reguły obliczania współczynników Taylora:

$$(u \pm v)_k = (u)_k \pm (v)_k \quad (16)$$

$$(uv)_k = \sum_{j=0}^k (u)_j (v)_{k-j} \quad (17)$$

$$(u/v)_k = (u)_k - \sum_{j=1}^k (v)_j (u/v)_{k-j} \quad (18)$$

Uwzględniając trywialne relacje

$$(c)_0 = c, \quad (c)_k = 0 \quad \text{dla } k \geq 1, \quad c \text{ jest stałą}$$

$$(z)_0 = z, \quad (z)_1 = 1, \quad (z)_k = 0, \quad \text{dla } k \geq 2, \quad \text{dla zmiennej niezależnej } z,$$

współczynniki Taylora dowolnego rzędu dla dowolnej funkcji można obliczać przez rekursywne obliczanie wartości wszystkich częściowych wyrażeń poczynając od współczynników dla $k=0$ potem dla $k=1$ itd.

Porównując współczynniki rozwinięcia

$$u'(z) = \sum_{k=1}^{\infty} k(u)_k (z - z_0)^{k-1} = \sum_{k=0}^{\infty} (k+1)(u)_{k+1} (z - z_0)^k \quad (19)$$

$$\text{oraz } u'(z) = \sum_{k=1}^{\infty} (u')_k (z - z_0)^k \quad (20)$$

$$\text{otrzymuje się } (u')_k = (k+1)(u)_{k+1} \text{ lub } (u)_{k+1} = \frac{1}{k+1} (u')_k. \quad (21)$$

Uwzględniając (21) i reguły różniczkowania funkcji złożonej można wyprowadzić formuły na obliczanie współczynników Taylora analogiczne do (16)–(18) dla szerszej klasy funkcji:

$$w(z) = v(u(z)) \rightarrow w'(z) = v'(u(z)) \cdot u'(z) \quad (22)$$

Zilustrujemy wyprowadzenie formuły rekursywnej na obliczanie $(w)_k$ na przykładzie funkcji wykładniczej

$$w = \exp(u) \rightarrow w' = \exp(u) \cdot u' \rightarrow (w')_{k-1} = (w \cdot u')_{k-1}, \quad k \geq 1.$$

Z zależności (17) i (21) wynika, że

$$k(w)_k = (w')_{k-1} = (w \cdot u')_{k-1} = \sum_{j=0}^{k-1} (w)_j (u')_{k-1-j} = \sum_{j=0}^{k-1} (w)_j (k-j)(u)_{k-j},$$

a ponieważ $(w) = (\exp(u))_k$, więc

$$(\exp(u))_k = \sum_{j=0}^{k-1} (1-j/k)(\exp(u))_j (u)_{k-j}, \quad k \geq 1. \quad (23)$$

W podobny sposób można wyprowadzić formuły rekursywne dla innych funkcji.

PRZYKŁAD 1. W celu praktycznego przedstawienia różniczkowania rekursywnego rozważmy następującą funkcję $u(z) = z \exp(-z^2)$. Chcemy obliczyć wartości współczynników Taylora w punkcie $z_0 = 0.5$. Funkcja $u(z)$ może być zdefiniowana także następująco:

$$u'(z) = (1-2z^2) \exp(-z^2) = 2(0.5-z^2) \exp(-z^2).$$

Przedstawimy $u'(z)$ za pomocą następującej listy jedno- i dwuargumentowych operacji:

$$T_1 = 0.5, \quad T_2 = z, \quad T_3 = -T_2^2, \quad T_4 = T_1 + T_3, \quad T_5 = \exp(T_3), \quad u'(z) = (u)_1 = 2T_4 T_5.$$

Stosując rekursywne relacje podane wcześniej otrzymamy dla $k \geq 1$

$$(T_1)_k = 0, \quad k > 1, \quad (T_2)_k = \begin{cases} 1, & k = 1 \\ 0, & k > 1 \end{cases}, \quad (T_3)_k = - \sum_{j=0}^k (T_2)_j (T_2)_{k-j},$$

$$(T_4)_k = (T_3)_k, \quad k > 1, \quad (T_5)_k = \sum_{j=0}^{k-1} (1-j/k)(T_5)_j (T_3)_{k-j},$$

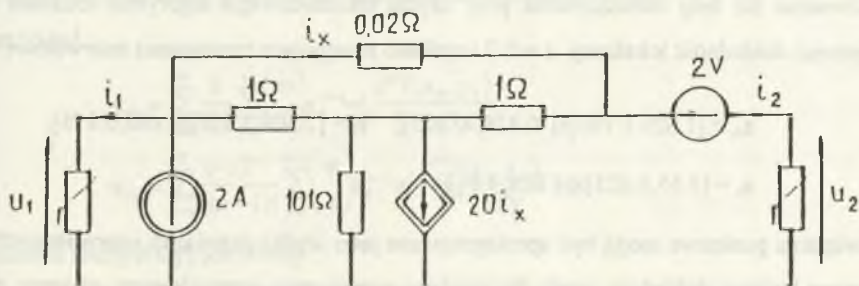
$$(u)_{k+1} = \frac{2}{k+1} \sum_{j=0}^k (T_4)_j (T_5)_{k-j}.$$

W tabelicy przedstawiono niektóre wartości związane z obliczeniami rekursywnymi, gdzie $e = \exp(1)$.

k	$(T_3)_k$	$(T_4)_k$	$(T_5)_k$	$(u)_{k+1}$
0	-1	-0.5	1/e	-1/e
1	-2	-2	-2/e	-1/e
2	-1	-1	1/e	5/3e
3	0	0	2/3e	-1/6e

PRZYKŁAD 2. Na rys.1 pokazano obwód nieliniowy, który jest opisany równaniem hybrydowym postaci:

$$F(z) = \begin{bmatrix} f_1(z_1) \\ f_2(z_2) \end{bmatrix} = \begin{bmatrix} 0.001113 & 1.225933 \\ 1.000036 & 0.001919 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} - \begin{bmatrix} 2.0 \\ -2.0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad z_1 = u_1, z_2 = i_2.$$



Rys. 1. Nieliniowy obwód rezystancyjny

Fig. 1. Nonlinear resistive circuit

Charakterystyki rezystorów nieliniowych są następujące :

$$f_1(z_1) = 10z_1(\exp(-z_1^2) + 0.02z_1^2), \quad f_2(z_2) = 3/(1 + \exp(-z_2)) - 1.5.$$

Jako obszar początkowy przyjęto wektor interwałowy $z^0 = [0.0, 4.0] \times [-1.0, 5.0]$.

Użyto stałej macierzy A_k

$$A_k = A = \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}.$$

W każdym kroku dokonywano bisekcji wektora interwałowego $z = (z_1, z_2)$ według reguły: znaleźć r takie, że $w(z_r) = \max w(z_i)$, $i = 1, 2$; przyjmując $m_r = m(z_r)$ i utworzyć dwa interwały z' i z'' tak, że np.

$$z' = [z_1, z_1] \times [m_r, z_2], \quad z'' = [z_1, z_1] \times [z_2, m_r].$$

Aproksymacja interwałowa $f'(z^k)$ była realizowana przy użyciu reprezentacji (10) dla charakterystyk nieliniowych, przy czym $z_0 = m(z)$. Przyjęto, że $\sigma = 4$ i $\nu = 8$ dla współczynników Bernsteina. Składnik reszty także aproksymowano. Zauważmy, że jeśli $m = m_k(z^k)$, to

$$m(b^k) = F(m_k), \quad w(b^k) = (A_k - f'(z^k) + H)w(z^k) \quad (24)$$

i wektor d^k może być obliczony bez konieczności użycia oprogramowania realizującego operacje interwałowe (por. np. [12]).

Równania (6) były rozwiązywane przy użyciu standardowego algorytmu rozkładu LU. Przyjmując dokładność lokalizacji $\epsilon = 0.2$ uzyskano następujące rozwiązania interwałowe:

$$z_A = [1.625, 1.750] \times [-0.438, -0.625], \quad z_B = [2.500, 2.625] [0.688, 0.875],$$

$$z_C = [3.55, 3.625] \times [4.625, 4.813]$$

Rozwiązania punktowe mogą być aproksymowane jako środki rozwiązań interwałowych lub obliczane bardziej dokładnie, jeżeli dla każdego rozwiązania interwałowego użyjemy $z^{(0)} = m(z)$ jako rozwiązania startowego dla algorytmu Newtona-Raphsona

$$F'(z^{(n)}) \delta^{(n)} = F(z^{(n)}), \quad z^{(n+1)} = z^{(n)} - \delta^{(n)}, \quad n = 0, 1, \dots$$

Uzyskuje się w ten sposób trzy rozwiązania punktowe: $z_A = (1.6721, -0.4455)$,

$z_B = (2.5997, 0.8495)$, $z_C = (3.4649, 4.7440)$.

2.2. Charakterystyki sterowane

Dla obwodów zawierających elementy nieliniowe, sterowane o charakterystykach zadanych dowolnymi, gładkimi funkcjami $f : D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ elementy macierzy $f'(x)$ są pochodnymi cząstkowymi. Niech każdy niezerowy element macierzy $f'(z)$ ma pochodne cząstkowe dostatecznie wysokiego rzędu w punkcie $z = (z_1, z_k) \in \mathbf{z}$, $i, k = 1, 2, \dots, n$, $i \neq k$ (dokładniej $(z_i, z_k) \in (\mathbf{z}_i, \mathbf{z}_k)$). Aby uprościć zapis, oznaczmy $(z_i, z_k) = (x, y)$ i $(\mathbf{z}_i, \mathbf{z}_k) = (\mathbf{x}, \mathbf{y})$. Każdy niezerowy element można przedstawić za pomocą szeregu Taylora

$$f'_{j,k}(x, y) = u(x, y) = p(x, y) + r(x, y) \quad (25)$$

gdzie

$$p(x, y) = \sum_{m=0}^{\sigma} \frac{1}{m!} \sum_{j=0}^m \binom{m}{j} \frac{\partial^m f(x_0, y_0)}{\partial^{m-j} x \partial^j y} (x - x_0)^{m-j} (y - y_0)^j = \sum_{s+t=0}^{\sigma} \alpha_{st} x^s y^t, \quad (26)$$

a symbol $\sum_{s+t=0}^{\sigma} \alpha_{st}$ oznacza sumę wszystkich α_{st} , gdzie $0 \leq s+t \leq \sigma$. Współczynniki α_{st} są liniowymi kombinacjami pochodnych cząstkowych w punkcie (x_0, y_0) .

Na przykład,

$$\alpha_{00} = \sum_{m=0}^{\sigma} \frac{1}{m!} \sum_{j=0}^m \binom{m}{j} x_0^{m-j} y_0^j \frac{\partial^m f(x_0, y_0)}{\partial^{m-j} x \partial^j y},$$

$$\alpha_{10} = \sum_{m=0}^{\sigma} \frac{(-1)^{k-1}}{(m-1)!} \sum_{j=0}^{m-1} \binom{m}{j} x_0^{m-1-j} y_0^j \frac{\partial^m f(x_0, y_0)}{\partial^{m-1} x \partial^j y}.$$

Składnik reszty $r(x, y)$ jest równy

$$r(x, y) = \frac{1}{(\sigma+1)!} \sum_{j=0}^{\sigma+1} \binom{\sigma+1}{j} \frac{\partial^{\sigma+1} u(\xi, \psi)}{\partial^{\sigma-j+1} x \partial^j y} (x - x_0)^{\sigma+1-j} (y - y_0)^j, \quad (27)$$

gdzie: $\xi = x_0 + \theta(x - x_0)$ oraz $\psi = y_0 + \theta(y - y_0)$, $\theta \in [0, 1]$

$(x_0, y_0) \in (\mathbf{x}, \mathbf{y})$ (np. $x_0 = m(\mathbf{x})$, $y_0 = m(\mathbf{y})$).

Aby oszacować granice zakresu wartości $u(x, y)$, wprowadzamy formę resztową rozwinięcia (25)

$$V(f'_{ij}, \mathbf{z}) = V(u, \mathbf{z}) := R(p, \mathbf{z}) + r(\mathbf{z}). \quad (28)$$

Zauważmy, że używamy tutaj dokładnego zakresu wielomianu p w \mathbf{z} , natomiast $r(\mathbf{z})$ może być oszacowaniem w arytmetyce interwałowej $r(\mathbf{z})$ lub innym estymatorem.

Dla $(x_0, y_0) \in (x, y)$

$$r(z) = r(x, y) \subseteq \frac{w(x)^\kappa w(y)^\lambda}{(\sigma+1)!} \sum_{j=0}^{\sigma+1} \binom{\sigma+1}{j} \frac{\partial^{\sigma+1} u(x, y)}{\partial^{\sigma-j+1} x \partial^j y}, \quad (29)$$

gdzie

$$\kappa = \begin{cases} 1 & \text{dla } w(x) \leq 1 \\ \sigma+1 & \text{dla } w(x) > 1 \end{cases}, \quad \lambda = \begin{cases} 1 & \text{dla } w(y) \leq 1 \\ \sigma+1 & \text{dla } w(y) > 1 \end{cases}.$$

Relacje (25)-(29) prowadzą do następującego twierdzenia:

TWIERDZENIE: Jeżeli $f'_{ij}(x, y) = u(x, y) : D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ jest $\sigma + 1$ - krotnie różniczkowalna cząstkowo, to dla wszystkich $z = (x, y) \subseteq D$ zachodzi:

1. $R(f'_{ij}, z) = R(u, z) \subseteq V(u, z)$,
2. $q(R(u, z), V(u, z)) \leq \gamma w(x)^\kappa w(y)^\lambda$, $\gamma \geq 0$.

Widać, że przekroczenie zakresu $R(u, z)$ przez $V(u, z)$ zależy od pewnych potęg $w(x)$ i $w(y)$. Przekroczenie to zmniejsza się z dowolną potęgą $\sigma+1$, jeżeli szerokości interwałów x i y są małe.

Forma resztowa funkcji $u(x, y)$ może być efektywnie wyznaczana, jeśli zastosujemy wielomiany Bernsteina dwóch zmiennych. Wielomian dwóch zmiennych $p(x, y)$ można przedstawić w postaci wielomianu Bernsteina stopnia $v \geq \sigma$ w interwale (prostokącie) $(x, y) = [\underline{x}, \bar{x}] \times [\underline{y}, \bar{y}]$ następująco:

$$p(x, y) = \sum_{s=0}^{\sigma} \sum_{t=0}^{\sigma} \alpha_{st} x^s x^t = \sum_{s=0}^v \sum_{j=0}^v b_{ij}^{(v)} \frac{\binom{v}{i} \binom{v}{j}}{(w(x)w(y))^m} (x - \underline{x})^i (\bar{x} - x)^{v-i} (y - \underline{y})^j (\bar{y} - y)^{v-j}, \quad (30)$$

gdzie $v \geq \sigma$, a współczynniki Bernsteina są równe

$$b_{ij}^{(v)} = \sum_{s=0}^i \sum_{t=0}^j \frac{\binom{i}{s} \binom{j}{t}}{\binom{v}{s} \binom{v}{t}} \sum_{u=s}^{\sigma} \sum_{v=t}^{\sigma} \binom{u}{s} \binom{v}{t} \alpha_{uv} \underline{x}^{u-s} \underline{y}^{v-t} (w(x))^i (w(y))^j \quad (31)$$

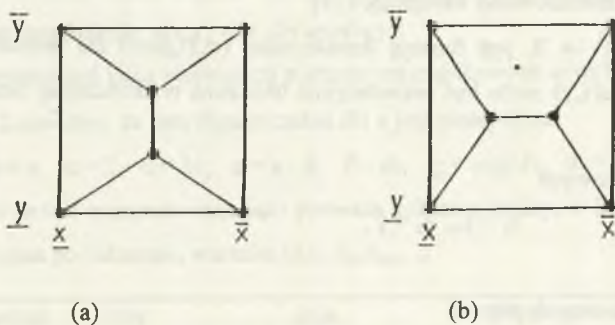
Współczynniki Bernsteina przybliżają zakres wartości $p(x, y)$, zachodzi bowiem

$$g = \min_{0 \leq i, j \leq v} b_{ij}^{(v)} \leq p(x, y) \leq \max_{0 \leq i, j \leq v} b_{ij}^{(v)} = G. \quad (32)$$

$$0 \leq i, j \leq v \quad (x, y) \in (x, y) \quad 0 \leq i, j \leq v$$

Równość zachodzi w lewej (odpowiednio prawej) nierówności wtedy i tylko wtedy, gdy g (odpowiednio G) jest jednym ze współczynników $b_{00}^{(v)}, b_{0v}^{(v)}, b_{v0}^{(v)}, b_{vv}^{(v)}$. Podobnie jak w przypadku wielomianów jednej zmiennej granice są bliższe rzeczywistego zakresu, gdy v jest większe.

Współczynniki b_{ij} mogą być obliczane bardziej efektywnie niż bezpośrednio ze wzoru (31). Współczynniki Bernsteina są także użyteczne przy estymacji zakresu składnika reszty $r(x,y)$. W tym celu interpoluje się pochodne cząstkowe wielomianami dwóch zmiennych stopnia drugiego. Interpolacja wymaga obliczania wartości pochodnych cząstkowych na pewnej siatce, np. na jednej z pokazanych na rys.2(a) lub 2(b).



Rys.2. Siatki interpolacji

Fig.2. Interpolation grids

Prowadzi to do układu sześciu liniowych równań, które można rozwiązać dowolnym standardowym algorytmem. Każda pochodna cząstkowa ma postać:

$$\frac{\partial^m u(x,y)}{\partial^{m-j} x \partial^j y} = \sum_{s=0}^2 \sum_{t=0}^2 \beta_{st} x^s y^t, \quad (x,y) \in (x,y). \quad (33)$$

$$m = \sigma + 1, \quad j = 1, 2, \dots, m.$$

Zakres wartości składnika reszty oblicza się następująco:

$$r(z) = r(x,y) \subseteq \frac{w(x)^{\sigma} w(y)^{\lambda}}{(\sigma+1)!} \sum_{j=0}^{\sigma+1} \binom{\sigma+1}{j} R \left(\frac{\partial^m u(x,y)}{\partial^{m-j} x \partial^j y}, z \right). \quad (34)$$

W przypadku obszarów startowych (początkowych) szerokości niektórych interwałów mogą być duże i podejście to może prowadzić do nadestymacji zakresu. Można wtedy rozdzielić obszar na mniejsze podobszary i obliczać zakres jako sumę zakresów (w sensie sumy zbiorów), podobnie jak to pokazano w pracy [7].

Przedstawimy teraz podstawowe idee stanowiące podstawę systematycznego i dokładnego wyznaczania wartości pochodnych cząstkowych wyższych rzędów. Podstawowa idea to dekompozycja obliczania wartości złożonej funkcji dwóch zmiennych w sekwencję prostszych obliczeń szczególnych funkcji jednej i dwóch zmiennych. W celu skonstruowania algorytmu obliczania pochodnych cząstkowych rzędu s wprowadza się pojęcie funkcji dopuszczalnej $(\sigma, 2)$ zdefiniowanej następująco [9]:

Funkcja $u: D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$, jest funkcją dopuszczalną $(\sigma, 2)$, jeżeli dla dowolnego punktu $(x_0, y_0) \in D$ wartość $u(x, y)$ może być sekwencyjnie obliczona w skończonej liczbie kroków przy wykorzystaniu:

a) warunków początkowych

$$s_1 = x_0, \quad s_2 = y_0, \quad (35)$$

b) funkcji dwóch zmiennych postaci

$$w = p + q, \quad w = p - q, \quad w = pq, \quad w = p/q, \quad w = p^q \quad (36)$$

c) dowolnej funkcji jednej zmiennej, σ -krotnie różniczkowalnej

$$h: M \subseteq \mathbb{R} \rightarrow \mathbb{R} \quad (37)$$

Łatwo spostrzec, że jeżeli potrafimy podać zapis analityczny funkcji u i jej pochodnych do rzędu σ , to u jest funkcją dopuszczalną $(\sigma, 2)$.

Dla dowolnej funkcji dopuszczalnej $(\sigma, 2)$ algorytm automatycznego różniczkowania w punkcie (x, y) należącym do dziedziny może być zbudowany następująco:

Krok 1. Utworzyć listę dopuszczalną dla funkcji $u(x, y)$, tj. listę warunków początkowych

(35) oraz funkcji (36) i (37), które wyznaczane kolejno dają $u(x, y)$.

Krok 2. Dla warunków początkowych $s_1 = x_0, s_2 = y_0$ utworzyć jednowymiarową tablicę (wektor) zawierającą wartości pochodnych cząstkowych do rzędu σ włącznie funkcji $g(x, y) \equiv x$ i $g(x, y) \equiv y$.

Krok 3. Dla każdej funkcji postaci (36) i (37) umieszczonej na liście dopuszczalnej dla $u(x,y)$ podać formuły obliczania wartości tej funkcji oraz wszystkich jej pochodnych cząstkowych do rzędu σ . Dla każdej funkcji postaci $a = f(b,c)$ wielkościami wejściowymi będą elementy uprzednio wyznaczonych wektorów $B = (b, b_x, b_y, b_{xx}, \dots)$, $C = (c, c_x, c_y, c_{xx}, \dots)$, a wielkości wyjściowe będą zawarte w wektorze $A = (a, a_x, a_y, a_{xx}, \dots)$. Dla dowolnej funkcji jednej zmiennej $d = h(e)$ wielkością wejściową do formuł obliczeniowych będzie wcześniej wyznaczony wektor $E = (e, e_x, e_y, e_{xx}, \dots)$, a wielkością wyjściową będzie wektor $D = (d, d_x, d_y, d_{xx}, \dots)$.

PRZYKŁAD 3. Aby bliżej objaśnić przedstawiony algorytm automatycznego różniczkowania, rozpatrzmy funkcję $u(x,y) = (x - 2y)\exp(-xy)$.

Chcemy wyznaczyć kilka pierwszych pochodnych cząstkowych w wybranym punkcie $(x_0, y_0) \in D$. Zauważmy, że lista dopuszczalna dla u jest następująca:

$$a := x, \quad b := y, \quad c := 2, \quad d := bc, \quad e := a - d, \quad f := ab, \quad g := \exp(-f), \quad u := e/g.$$

Lista ta może być interpretowana jako pierwsza kolumna tablicy, w której można obliczać kolejno, kolumna po kolumnie, wartości $(u, u_x, u_y, u_{xx}, \dots)$.

Krok	Funkcja	$\partial/\partial y$	$\partial/\partial x$	$\partial^2/\partial x^2$
(1)	$a=x$	$a_x=1$	$a_y=0$	$a_{xx}=0$
(2)	$b=y$	$b_x=0$	$b_y=1$	$b_{xx}=0$
(3)	$c=2$	$c_x=0$	$c_y=0$	$c_{xx}=0$
(4)	$d=bc$	$d_x=b_x c + b c_x$	$d_y=b_y c + b c_y$	$d_{xx}=b_{xx} c + 2b_x c_x + b c_{xx}$
(5)	$e=a-d$	$e_x=a_x - d_x$	$e_y=a_y - d_y$	$e_{xx}=a_{xx} - d_{xx}$
(6)	$f=ab$	$f_x=a_x b + a b_x$	$f_y=a_y b + a b_y$	$f_{xx}=a_{xx} b + 2a_x b_x + a b_{xx}$
(7)	$g=\exp(f)$	$g_x=g f_x$	$g_y=g f_y$	$g_{xx}=g_x f_x + g f_{xx}$
(8)	$u=e/g$	$u_x = \frac{e_x g - e g_x}{g^2}$	$u_y = \frac{e_y g - e g_y}{g^2}$	$u_{xx} = (g e_{xx} - e g_{xx}) g^{-2} - 2g^{-1} g_x u_x$

PRZYKŁAD 4. Na rys.3. przedstawiono obwód, który ma równanie hybrydowe postaci

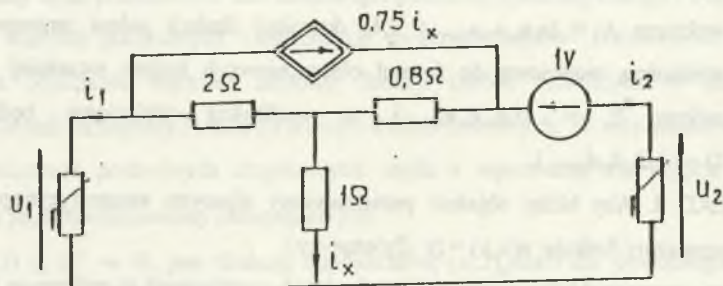
$$F(z) = F(x, y) = \begin{bmatrix} f_1(x, y) \\ f_2(x, y) \end{bmatrix} - \begin{bmatrix} 0.5 & -0.25 \\ -0.8 & 2.8 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad x = u_1, \quad y = i_2.$$

Charakterystyki elementów nieliniowych są następujące

$$f_1(x,y) = (y-2x)\exp(-xy), \quad f_2(x,y) = -0.09x^2 + 0.63xy - 1.1025y^2.$$

Chcemy znaleźć rozwiązania równania w obszarze

$$z = [-16,16] \times [-16,16].$$



Rys.3. Nieliniowy obwód bezinercyjny

Fig.3. Nonlinear inertialess circuits

Algorytm obliczeniowy wymaga wyznaczenia interwałowej reprezentacji macierzy Jacobiego

$$F'(z) = \frac{\partial f(z)}{\partial z} - H = [J, \bar{J}],$$

gdzie J i \bar{J} oznaczają dolną i górną granicę macierzy interwałowej.

Przy wyznaczaniu tej reprezentacji uwzględniono w rozwinięciu (26) $\sigma = 3$, a składnik reszty (27) wyznaczano na siatkach pokazanych na rys.2(a) i 2(b) stosowanych na przemian w kolejnych iteracjach. Wymagało to wyznaczenia 10 wartości pochodnych dla wielomianowej części rozwinięcia i 30 wartości (ze względu na interpolację) dla składnika reszty, w każdej iteracji.

Przyjmując dokładność $\varepsilon = 0.3$ uzyskano następujące rozwiązania interwałowe:

$$\begin{aligned} z_A &= [-0.25, 0.0] \times [-0.5, -0.25], & z_B &= [1.5, 1.75] \times [3.2, 3.45], \\ z_D &= [-2.0, -1.75] \times [-1.0, -0.75], & z_C &= [2.85, 3.1] \times [0.45, 0.7], \\ z_E &= [0.45, 0.7] \times [2.85, 3.1], & z_F &= [-10.7, -10.45] \times [-0.25, 0.0]. \end{aligned}$$

Obwód posiada sześć rozwiązań punktowych:

$$\mathbf{z}_A = (-0.185, -0.370), \quad \mathbf{z}_B = (1.667, 3.333), \quad \mathbf{z}_C = (-1.882, -0.885),$$

$$\mathbf{z}_D = (2.993, 0.538), \quad \mathbf{z}_E = (0.535, 3.010), \quad \mathbf{z}_F = (-10.535, -0.153).$$

3. UWAGI KOŃCOWE

Rezultaty tutaj przedstawione mogą być także wykorzystane w innych algorytmach opartych na analizie interwałowej, na przykład przedstawionych w pracach [6],[16].

Efektywność obliczeń i dokładność w proponowanym podejściu zależą od możliwości obliczania wartości pochodnych (współczynników Taylora) dostatecznie wysokiego rzędu. Przedstawione tu metody różniczkowania rekursywnego dają taką możliwość. Odpowiednia adaptacja reguł różniczkowania funkcji elementarnych prowadzi do procedur, które pozwalają obliczać pochodne wybranego rzędu z dokładnością wynikającą jedynie z arytmetyki komputera. Prowadzi to do zwartego i ogólnego sposobu wyznaczania estymatora interwałowej reprezentacji macierzy Jacobiego równania hybrydowego. Podejście to może być interesujące w praktyce obliczeniowej. Języki programowania najnowszej generacji [12] - [15] pozwalają na efektywne programowanie operacji w arytmetyce interwałowej oraz procedur automatycznego różniczkowania.

LITERATURA

1. Chua L.O., Lin P.M.: Computer-Aided Analysis of Electronic Circuits, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1975.
2. Chua L.O., Desoer C.A., Kuh E.S.: Linear and Nonlinear Networks, McGraw-Hill, New York 1987.
3. Alefeld G., Herzberger J.: Introduction to Interval Computations, Academic Press, New York 1983.

4. Moore R.E.: *Methods and Applications of Interval Analysis*, SIAM, Philadelphia 1979.
5. Ratschek H., Rokne J.: *Computer Methods for the Range of Functions*, Horwood, Chichester 1984.
6. Kolev L.V., Mladenov V.M.: An interval method for finding all operating points of nonlinear resistive circuits, *Int.J.Cir.Theor.Appl.*,18, 257-267,1990.
7. Garczarczyk Z.: An interval approach to finding all equilibrium points of some nonlinear resistive circuits, *Proc. European Conference on Circuit Theory and Design - ECCTD'93*,1281-1286,Davos 1993.
8. Garczarczyk Z.: An interval approach to finding all solutions of some nonlinear inertialess circuits, *International AMSE Conference "Signal & Systems"*, Warszawa 1991.
9. Kalaba R., Tishler A.: Automatic derivative evaluation in optimisation of nonlinear models, *Rev.Econom.Stat*.66,1984.
10. Rall L.B.: *Automatic Differentiation: Techniques and Applications*, Springer-Verlag, New York 1981.
11. Griewank A., Corliss G.F. (eds): *Automatic Differentiation of Algorithms*, SIAM, Philadelphia 1991.
12. Klatte R. i inni: *Pascal-XSC. Language Reference with Examples*, Springer-Verlag, Berlin 1992.
13. Davidenkoff A.: ACRITH-XSC. A programming language for scientific engineering computation (A survey), *ZAMM* 72/6, 465-467, 1992.
14. Reid J.: The advantages of Fortran 90, *Computing* 48, 219-238, 1992.
15. Walter W.V.: FORTRAN-XSC. A portable Fortran 90 module library for accurate and reliable scientific computing, *Computing, Suppl.*9,265-285, 1993.
16. Gan J., Song Y.M.: All DC solutions of nonlinear circuits, *Proc. International Symposium on Circuits and Systems*, New Orleans 1990.
17. Oppenheimer E.P., Michel A.N.: Application of interval analysis techniques to linear systems: Part I - Fundamental results, *IEEE Trans. Circuit. Syst.*, vol.35, no.9, 1988.

18. Neumaier A.: Interval Methods for Systems of Equations, Cambridge University Press, Cambridge 1990.

Recenzent: Prof.dr hab.inż. Stanisław Osowski

Wpłynęło do Redakcji 11 kwietnia 1994 r.

Abstract

Interval analysis in conjunction with generalized bisection form the basis of algorithms that find all solutions of a system of hybrid equations of nonlinear resistive circuits.

In algorithms based on those techniques it is often necessary to estimate an interval evaluation of the Jacobian matrix of the equations or a nonlinear functions. The estimating problem is equivalent to that of bounding ranges of functions over an interval. In the paper we study this problem for a nonlinear circuits containing elements with a characteristics described by any sufficiently smooth function $f:D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ and $f:D \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ especially by univariate or bivariate polynomials. It permits to take into account presence of nonlinear controlled elements in the circuit. Approximation of the interval extension of Jacobian matrix is done by univariate and bivariate Taylor expansion of the functions concerned with nonlinear characteristics. Ranges of values of the expansion are estimated with use of Bernstein polynomials. It is shown that overestimation depends on the order of expansion. Evaluation of the ranges needs information about values of higher order Taylor coefficients or partial derivatives in univariate or bivariate case, respectively. Employing recursive computations, the method of automatic differentiation provides effectively values of derivatives. Presented methods are illustrated by numerical examples.