І.Я. СПІВАК

ІНТЕРВАЛЬНІ ОБЧИСЛЕННЯ

ОПОРНИЙ КОНСПЕКТ ЛЕКЦІЙ

Тернопіль
ТНЕУ
2011

ЗМІСТ

ВСТУП

1. Місце інтервальних обчислень в інтервальному аналізі. Предмет дисципліни
2. Історія розвитку інтервальних обчислень

1. Місце інтервальних обчислень в інтервальному аналізі. Предмет дисципліни

У математичній енциклопедії інтервальний аналіз розглядається, як наукова дисципліна, що застосовується для врахування похибок заокруглень при розрахунках на ЕОМ. У довідковій літературі по математичних методах, поняття інтервального аналізу часто хибно трактується як "арифметика інтервалів". Однак розвиток методів інтервального аналізу призвів до певної еволюції його наукової термінології та самого трактування. Наприклад, у низці робіт інтервальний аналіз розглядається у більш ширшому трактуванні, – як теоретико-множинний підхід. На наш погляд, найбільш точніше визначення інтервального аналізу на сьогоднішній день наведене у працях новосибірської школи – як наукової дисципліни на стику інформатики та математики, предметом, якої є розв'язування задач з інтервальними невизначеностями в даних на вході, виході чи на проміжних етапах.

При розв'язуванні задач експериментального моделювання об'єктів на основі даних з інтервальною невизначеністю, оцінки параметрів моделей часто задаються множинами (областями) різної конфігурації: еліпсоїдні, многогранні області. Тим часом як в інтервальному аналізі традиційно застосовують "інтервальні методи оцінювання", коли область параметрів задається прямокутним гіперпаралелепіпедом. З іншого боку, запозичення методів та термінів з математичної статистики, зокрема, з регресійного аналізу, що стосуються інтервального аналізу, призвело до застосування та утвердження терміну "аналіз інтервальних даних". Тому у межах інтервального аналізу зручно виділити методи аналізу інтервальних даних, під якими будемо розуміти методи, напрямлені на розв'язування задач моделювання з інтервальними невизначеностями в експериментальних даних, дослідження механізмів впливу невизначеностей на їх формування, отримання та дослідження математичних моделей об’єктів з множинними оцінками параметрів.

Методи інтервального аналізу та їхній розвиток створили передумови розвитку трьох напрямків, наукової та практичної діяльності, пов’язаної з математичним моделюванням об’єктів на основі інтервальних даних. Це:

5.1. математичний, в межах якого проводяться дослідження математичних проблем інтервальних обчислень;

5.2. комп’ютерний, в межах якого досліджуються питання створення та використання інтервальних обчислень;

– прикладний, пов’язаний з використанням методів інтервального аналізу і відповідних комп’ютерних засобів для побудови та дослідження математичних моделей широкого класу об’єктів.

На сьогоднішній день опубліковано десятки тисяч праць, присвячених математичному моделюванню об’єктів на основі методів інтервального аналізу. В інституті прикладної математики Фрейбургського університету створена і постійно поповнюється “інтервальна бібліотека”. У періодичному виданні Freiburger Intervall-Berichte в міру накопичення матеріалу публікуються бібліографічні списки. Найбільше робіт опубліковано в сфері автоматики та управління, електротехніки та електроніки, екології та охорони здоров’я, технологічних дисциплін, економіки.

2. Історія розвитку інтервальних обчислень

Інтервальні обчислення є теоретичною основою усіх інтервальних методів моделювання. Своєю чергою інтервальні обчислення побудовані на базі методів інтервального аналізу.

Основи інтервального аналізу були розроблені на вимогу часу – як засіб урахування похибок при розрахунках на ЕОМ і описані у монографії Р. Мура

(R. Moore) за назвою “Interval analysis” у 1966 р. Перша праця серед вітчизняних вчених по даному напрямку була опублікована Канторовичем Л.В. у “Сибірському математичному журналі” у 1962 р.

Методи інтервального аналізу дозволяють враховувати похибки наведення вхідних даних, які у даному випадку набувають вигляду скінчених інтервалів, а також похибки заокруглень на ЕОМ. При цьому результат розрахунків наводиться також у інтервальному вигляді.

Усі задачі, які розв’язуються із застосуванням інтервальних обчислень можна розділити на дві групи: задачі із неточними – інтервальними даними та задачі із точними даними. Перші, пов’язані із обчисленням множини розв’язків, а другі із її поступовим уточненням. Прикладом задачі, яка відноситься до першої групи є система лінійних рівнянь з інтервальними коефіцієнтами. Для задач з точними даними інтервальний аналіз використовуються в методах, які породжують послідовність меж, що збігаються до розв’язку.

При реалізації інтервальних обчислень виникають значні проблеми, коли традиційні чисельні методи безпосередньо переносяться на інтервальні числа. В результаті такого перенесення відбувається значне розширення результуючого інтервалу. Інші проблеми, які стосуються інтервальних обчислень, пов’язані із виконанням операції ділення на інтервали, які включають нуль. Внаслідок розширення інтервалів при реалізації обчислень, такі ситуації виникають достатньо часто.

Проблеми, пов’язані із застосуванням класичної інтервальної арифметики часто вдається подолати застосуванням розширених інтервальних арифметик Ширина результуючого інтервалу в інтервальних обчисленнях залежить від порядку виконання операцій.

Розглянуто особливості інтервальних методів, для їх успішного застосування вимагають перегляду практично усіх розроблених на даний момент чисельних методів.

ТЕМА 1. ОСОБЛИВОСТІ ІНТЕРВАЛЬНОЇ АРИФМЕТИКИ ТА ІНТЕРВАЛЬНОГО ОЦІНЮВАННЯ

1.1.Дійсна інтервальна арифметика
1.2.Лінійне інтервальне рівняння та його алгебраїчний розв'язок
1.3.Відстань між двома інтервалами, абсолютна величина інтервалу та його ширина

1.1.Дійсна інтервальна арифметика

Основна ідея інтервальних обчислень полягає у наведені дійсного числа не одним машинним числом, а двома, які задають його гарантовані межі.

Для ілюстрації правил виконання операцій класичної інтервальної арифметики, розглянемо такі інтервальні числа: $[a]=[a^{-};a^{+}]$, $[b]=[b^{-};b^{+}]$, $[c]=[c^{-};c^{+}]$.

Позначимо за *a*, *b*, *c* дійсні числа, які належать до вказаних інтервалів, відповідно. Тоді правило виконання операцій класичної інтервальної арифметики таке: "якщо $[c^{-};c^{+}]=[a^{-};a^{+}]\circ[b^{-};b^{+}]$, $a\in[a^{-};a^{+}]$, $b\in[b^{-};b^{+}]$, то $a\circ b\in[c^{-};c^{+}]$, де символ $\circ$ – означає арифметичну операцію із набору {+, -, *, /}.

Результати операцій додавання, віднімання, ділення та множення над інтервалами можна записати явно за допомогою таких формул:

$$[a]+[b]=[a^{-}+b^{-};a^{+}+b^{+}],$$

- = ,

/ =

= ; .

Операції інтервальної арифметики поширюються на інші операції, наприклад на неперервні операції деяких функціональних перетворень *r(x)*:

exp, ln, sin, cos і т.д. В цьому випадку:

$$r(x) = [\min_{x \in X} r(x); \max_{x \in X} r(x)]$$

Властивості:

- $[a^-;a^+] + [b^-;b^+] = [b^-;b^+] + [a^-;a^+] \cdot$; $[a^-;a^+] \cdot * [b^-;b^+] = [b^-;b^+] * [a^-;a^+] \cdot$; $([a^-;a^+] \cdot + [b^-;b^+]) + [c^-;c^+] = [a^-;a^+] \cdot + ([b^-;b^+] + [c^-;c^+])$- комутативність
- $([a^-;a^+] \cdot * [b^-;b^+]) * [c^-;c^+] = [a^-;a^+] \cdot * ([b^-;b^+] * [c^-;c^+])$ – асоціативність
- $[a^-;a^+] \cdot * ([b^-;b^+] + [c^-;c^+]) \subseteq [a^-;a^+] \cdot * [b^-;b^+] + [a^-;a^+] \cdot * [c^-;c^+]$ –є лінійністысть
- Дистрибутивність виконується для випадків:

$a*([b^-;b^+] + [c^-;c^+]) = a*[b^-;b^+] + a*[c^-;c^+]$; $[a^-;a^+] \cdot * ([b^-;b^+] + [c^-;c^+]) = [a^-;a^+] \cdot * [b^-;b^+] + [a^-;a^+] \cdot * [c^-;c^+]$, якщо $b*c>=0$, для усіх $b \in [b^-;b^+]$; $c \in [c^-;c^+]$.

Основою усіх чисельних методів, реалізованих у інтервальній арифметиці, є властивість, яка називається “монотонністю включення”. Розглянемо суть цієї властивості

Нехай маємо інтервали $[a^-;a^+]$, $[b^-;b^+]$, $[c^-;c^+]$, $[d^-;d^+]$. Тоді справедлива така *властивість монотонності включення*:

,

Саме ця властивість дозволяє побудувати ітераційні процедури наближення множин розв’язків для задач із інтервальними (неточними) даними. При цьому розміри множини визначаються шириною інтервалів вхідних даних та можливостями ітераційної процедури.

1.2. Лінійне інтервальне рівняння та його алгебраїчний розв'язок

Нехай маємо лінійне рівняння:

$$[a^-;a^+]\cdot[x^-;x^+]=[b^-;b^+]$$

і при цьому $0 \notin [a^-;a^+]$. У якому випадку дане рівняння має розв'язок?

Відповідь на це запитання дає правило , яке побудоване на функції Ратшека:

$$\chi([a])=\begin{cases} a^-/a^+, якщо\_|a^-|\le|a^+| \\ a^+/a^-, в\_інших\_випадках \end{cases}.$$

Рівняння $[a^-;a^+]\cdot[x^-;x^+]=[b^-;b^+]$ має розв'язок відносно $[x^-;x^+]$, якщо $\chi([a])\ge\chi([b])$. При цьому розв'язок не єдиний, якщо $\chi([a])=\chi([b])\le 0$.

Що означає розв'язок $[x^-;x^+]$ рівняння $[a^-;a^+]\cdot[x^-;x^+]=[b^-;b^+]$?

Це означає, що:

$\min\{a^-\cdot x^-,a^-\cdot x^+,a^+\cdot x^-,a^+x^+\}=b^-$; $\max\{a^-\cdot x^-,a^-\cdot x^+,a^+\cdot x^-,a^+x^+\}=b^+$.

Проте в інтервальному аналізі є поняття алгебраїчного розв'язку у вигляді $[x^-;x^+]_{алг}=[b^-;b^+]/[a^-;a^+]$ і при цьому справедливим є:

$$[x^-;x^+]\subseteq[x^-;x^+]_{алг}.$$

Як видно, алгебраїчний розв'язок включає інтервал, який забезпечує виконання умови = .

1.3. Відстань між двома інтервалами, абсолютна величина інтервалу та його ширина

Відстань між двома інтервалами і :

(,)=

Важлива властивість введеної метрики

$\delta\,([a^-;a^+],[b^-;b^+]) \le \delta\,([a^-;a^+],[c^-;c^+]) + \delta\,([b^-;b^+],[c^-;c^+])$ нерівність трикутника.

Кожна послідовність інтервалів

$[a_1^-;a_1^+] \supseteq [a_2^-;a_2^+] \supseteq [a_3^-;a_3^+] \ldots \supseteq [a_k^-;a_k^+] \ldots$

збігається до інтервалу

$$[a^-;a^+] = \bigcap_{k=1}^{\infty} [a_k^-;a_k^+].$$

Операції інтервальної арифметики неперервні, тобто

$$\lim_{k\to\infty} ([a_k^-;a_k^+] + [b_k^-;b_k^+]) = \lim_{k\to\infty} [a_k^- + b_k^-; a_k^+ + b_k^+] =$$

$$[\lim_{k\to\infty}(a_k^- + b_k^-), \lim_{k\to\infty}(a_k^+ + b_k^+)] = [a^- + b^-; a^+ + b^+],$$

де

$$[a^-;a^+] = \bigcap_{k=1}^{\infty} [a_k^-;a_k^+],\quad [b^-;b^+] = \bigcap_{k=1}^{\infty} [b_k^-;b_k^+].$$

Аналогічне твердження для неперервних функцій $r([a^-;a^+])$.

Абсолютна величина інтервалу:

$|[a^-;a^+]| = \delta\,([a^-;a^+],[0;0])) = \max\{|a^-|,|a^+|\}$. Абсолютну величину можна записати і у вигляді:

$$[a^-;a^+] = \max_{a\in[a^-;a^+]} |a|.$$

Якщо , то .

Властивості введеної метрики. Нехай маємо інтервали , ,

, . Тоді:

1) (+ , +)= (,);

2) δ ($[a^-;a^+]+[b^-;b^+]$, $[c^-;c^+]+[d^-;d^+]$) ≤

$$\le \delta\,([a^-;a^+],[c^-;c^+]) + \delta\,([b^-;b^+],\ [d^-;d^+]);$$

3) δ (a$[b^-;b^+]$, a$[c^-;c^+]$)= $|a|$ δ ($[b^-;b^+]$, $[c^-;c^+]$), $a \in R$;

4) δ ($[a^-;a^+]\cdot[b^-;b^+]$, $[a^-;a^+]\cdot[c^-;c^+]$) ≤ $\left|[a^-;a^+]\right|$ δ ($[b^-;b^+]$, $[c^-;c^+]$).

Шириною інтервалу:

$$d([a^-;a^+]) = a^+ - a^- \ge 0$$

Із визначення ширини інтервалу відразу отримуємо наступні властивості:

1) якщо $[a^-;a^+] \subseteq [b^-;b^+]$, то $d([a^-;a^+]) \le d([b^-;b^+])$;

2) $d([a^-;a^+] \pm [b^-;b^+]) = d([a^-;a^+]) + d([b^-;b^+])$;

3) $d([a^-;a^+]\cdot[b^-;b^+]) \le d([a^-;a^+])\left|[b^-;b^+]\right| + \left|[a^-;a^+]\right| d([b^-;b^+])$;

4) $d([a^-;a^+]\cdot[b^-;b^+]) \ge \max\{\left|[a^-;a^+]\right| d([b^-;b^+]),\ \left|[b^-;b^+]\right| d([[a^-;a^+])$;

5) $d(a[b^-;b^+]) = |a| d([b^-;b^+]),\ a \in R$.

ТЕМА 2. ІНТЕРВАЛЬНЕ ОЦІНЮВАННЯ ТА ВЛАСТИВОСТІ ІНТЕРВАЛЬНИХ ФУНКЦІЙ

2.1.Множина значень у випадку дійсних функцій
2.2.Об'єднане інтервальне розширення неперервної дійсної функції
2.3.Інтервальна оцінка за умов апроксимації множини коефіцієнтів n-вимірними еліпсоїдами

2.1. Множина значень у випадку дійсних функцій

Розглянемо дійсні неперервні функції. Запис $y(\vec{x})$ означає обчислювальну процедуру значення функції y для заданих значень змінних $\vec{x} = (x_1,...,x_n)^T$. Будемо також розглядати функції $y = \eta\left(\vec{x}, \vec{b}, \vec{z}\right)$, які включають деякі константи, наприклад у такому вигляді $y = \vec{\varphi}^T(\vec{x}) \cdot \vec{\beta}$ чи у вигляді, коли константи $\beta_1,..., \beta_m$ замінені деякими множинними оцінками $\vec{b} \in \Omega$. Тоді оцінку діапазону зміни функції за умов $\vec{b} \in \Omega$ та $\vec{x} \in \chi$ отримаємо у вигляді:

$$[\hat{y}(x)] = [\hat{y}^-(x); \hat{y}^+(x)], \tag{1}$$

де $\hat{y}^-(\vec{x}) = \min\limits_{\vec{b}\in\Omega}\left(\vec{\varphi}^T(\vec{x}) \cdot \vec{b}\right)$ та $\hat{y}^+(\vec{x}) = \max\limits_{\vec{b}\in\Omega}\left(\vec{\varphi}^T(\vec{x}) \cdot \vec{b}\right)$.

2.2. Об'єднане інтервальне розширення неперервної дійсної функції

В інтервальній математиці, замість розв'язків інтервальної системи лінійних алгебраїчних рівнянь у вигляді множини Ω параметрів моделі, розглядають його локалізацію (наближення) за допомогою інтервального вектора   , який дозволяє знайти об'єднане інтервальне розширення неперервної дійсної функції. Очевидно властивості об'єднаного інтервального розширення матимуть певні відмінності в порівнянні з коридором інтервальних моделей, побудованим на основі розв'язків системи лінійних інтервальних алгебраїчних рівнянь.

Фактично, об’єднане інтервальне розширення, що являє множину інтервальних моделей, задається коридором $[\hat{y}(\vec{x})]$, коли взамін $\vec{b} \in \Omega$ можна записати $\vec{b} \in [\vec{b}]$. Надалі коридор інтервальних моделей при інтервальному задані їх параметрів (об’єднане інтервальне розширення) будемо позначати так: $[\hat{y}(\vec{x})]\Big|_{\vec{b}\in[\vec{b}]}$.

З іншого боку, замінивши у виразі для множини інтервальних моделей $\hat{y}(\vec{x}) = \vec{\varphi}^T(\vec{x}) \cdot \vec{b}$, $\vec{b} \in \Omega$, побудованій на основі множини розв’язків інтервальної системи лінійних алгебраїчних рівнянь, вектор оцінок $\vec{b}$ на інтервальний вектор $[\vec{b}]$ і виконавши операції інтервальної арифметики, отримаємо оцінку $[\hat{y}(\vec{x},[\vec{b}])]$ прогнозованого значення $\hat{y}(\vec{x})$ в інтервальній арифметиці.

Для загального випадку доведено наступне включення:

$$[\hat{y}(\vec{x})]\Big|_{\vec{b}\in[\vec{b}]} \subseteq [\hat{y}(\vec{x},[\vec{b}])], \forall \vec{x} \in \chi \quad . \qquad (2)$$

Процедури знаходження оцінок $[\hat{y}(\vec{x},[\vec{b}])]$, у яких застосовується інтервальна арифметика є достатньо надлишковими, а знайдені оцінки часто неточні. Крім того, при побудові інтервальних моделей статичних систем, не залежно від методу локалізації множини параметрів, важливим є не тільки можливість знаходження прогнозованого значення інтервалу виходу з мінімальними обчислювальними витратами, що в даному випадку забезпечується оцінками , але і забезпечення аналітичності заданя функціональних меж коридору моделей на усій області експерименту. Остання вимога зумовлює необхідність розгляду властивостей меж функціонального коридору лінійно-параметричних функцій.

Відмітимо, що для лінійно-параметричних функцій в силу властивості субдистрибутивності інтервальної арифметики, включення (2) перетворюється у рівність, тобто оцінки $[\hat{y}(\vec{x})]\big|_{\vec{b}\in[\vec{b}]}$ та $[\hat{y}(\vec{x},[\vec{b}])]$ співпадають для усіх значень $\vec{x}$.

Коридор інтервальних моделей у випадку локалізації області Ω оцінок параметрів інтервальним вектором $[\vec{b}]$, що в просторі параметрів є прямокутним паралелепіпедом

$$\Pi^{+}=\{\vec{b}\,\big|\,b_j^{-}\le b_j\le b_j^{+},\, j=1,...,m\}, \tag{3}$$

запишемо так:

$$[\hat{y}(\vec{x})]\big|_{\vec{b}\in[\vec{b}]}=[\min_{\vec{b}\in\Pi^{+}}\sum_{j=1}^{m}\varphi_j(\vec{x})\cdot b_j;\max_{\vec{b}\in\Pi^{+}}\sum_{j=1}^{m}\varphi_j(\vec{x})\cdot b_j]. \tag{4}$$

Зауважимо, що залежність між прогнозованим значенням виходу $\hat{y}$ та оцінками вектора параметрів $\vec{b}$ є лінійною. Тому при фіксованому векторі входів $\vec{x}$ розв'язками задач

$$\min_{\vec{b}\in\Pi^{+}}\sum_{j=1}^{m}\varphi_j(\vec{x})\cdot b_j,\quad \max_{\vec{b}\in\Pi^{+}}\sum_{j=1}^{m}\varphi_j(\vec{x})\cdot b_j \tag{5}$$

є вершини паралелепіпеда Π^{+}. Отже, межі коридору інтервальних моделей, як і у випадку застосування множини параметрів Ω, будуються на основі координат вершин області локалізації Π^{+}.

Введемо такі позначення:

, . (6)

З врахуванням вище зробленого зауваження перепишемо вираз (4) так:

. (7)

Перейдемо до аналізу виразу (7).

Як видно, для обчислення інтервалу прогнозування $[\hat{y}(\vec{x})]\big|_{\vec{b}\in[\vec{b}]}$ у фіксованій точці $\vec{x}_i$ необхідно визначити знаки значень усіх базових функцій $\varphi_j(\vec{x}_i),(j=1,...,m)$. Оскільки значення базових функцій в загальному випадку залежать від значень компонент вектора входів $\vec{x}$, то справедлива така властивість інтервальних моделей у випадку інтервальних оцінок їх параметрів.

Властивість 1. У загальному випадку межі коридору (7) інтервальних моделей є кусковими функціями.

Дану властивість можна проілюструвати прикладом.

Нехай маємо інтервальну модель $\hat{y}(x)=b_0+b_1\cdot x$, де $b_0\in[2,4], b_1\in[1,3]$, $x\in[-2,2], \varphi_1(x)=1, \varphi_2(x)=x$.

Коридор інтервальних моделей (7) для даного випадку буде таким:

$$[\hat{y}(\vec{x})]\big|_{\vec{b}\in[\vec{b}]}=[2,4]+[1,3]\cdot x.$$

Графічна ілюстрація коридору наведена на рисунку 1.

Рис. 1. Кусковість меж функціонального коридору

Як видно, на інтервалі [-2,0] значення базисної функції від’ємне, тобто

$\varphi_2(x) < 0$, а на інтервалі [0,2] $\varphi_2(x) > 0$, що є причиною кусковості меж функціонального коридору.

Властивість 2. Нехай на області експерименту χ усі базові функції $\varphi_j(\vec{x}),(j=1,...,m)$ у лінійно-параметричному рівнянні є неперервними та не кусковими і жодна з них не змінює свій знак на протилежний, тоді функціональні межі коридору інтервальних моделей (7) статичної системи на області χ є неперервними та не кусковими функціями.

Справедливість властивості 2 витікає з формул (6), та (7), оскільки за умов постійності знаків значень базових функцій $\varphi_j(\vec{x}),(j=1,...,m)$ на області експерименту χ, вибір значень $b_j^{\min}, b_j^{\max}$ не залежить від зміни значень вектора входів $\vec{x} \in \chi$.

Для поліноміальних моделей зміна знаків значень базових функцій пов'язана з їх "проходженням" через нульову точку $\vec{x}_0 = (0,...,0)^T$. Шляхом нормування змінних $\vec{x}$ центр $\vec{x}_0$ області експерименту χ можна вибрати так, щоб на області експерименту задовольнити умови справедливості властивості 2. У цьому випадку межі коридору інтервальних моделей (7) будуть неперервними та не кусковими функціями.

Однак таке нормування не завжди є зручним при розв'язуванні задач активної ідентифікації інтервальних моделей. Переважно для цих задач нормування проводять у такий спосіб, щоб центр $\vec{x}_0$ області експерименту саме співпадав з нульовою точкою .

Похибка прогнозування інтервальної моделі при інтервальних параметрах визначається шириною коридору (7):

. (8)

З урахуванням позначень (6), формулу (8) для визначення похибки прогнозування перепишемо так:

$$\Delta_{\hat{y}(\vec{x})}\Big|_{\vec{b}\in[\vec{b}]} = \left|\vec{\varphi}^{T}(\vec{x})\right| \cdot (\vec{b}^{+} - \vec{b}^{-}), \qquad (9)$$

де $\left|\vec{\varphi}^{T}(\vec{x})\right|$ – означає вектор абсолютних значень базових функцій у фіксованій точці $\vec{x}$;

$\vec{b}^{+}$, $\vec{b}^{-}$ – вектори, компонентами яких є b_j^{+} та b_j^{-}, відповідно.

Як видно з (9), при збільшені розмірів області Π^{+} локалізації Ω значення похибки прогнозування збільшується.

Врахуємо, що область Π^{+} є симетричною відносно центру

$$\vec{\bar{b}} = \frac{1}{2} \cdot (\vec{b}^{+} + \vec{b}^{-}). \qquad (10)$$

Тоді формула (7) для визначення коридору інтервальних моделей набуде такого вигляду

$$[\hat{y}(\vec{x})]\Big|_{\vec{b}\in[\vec{b}]} = [\vec{\varphi}^{T}(\vec{x}) \cdot \vec{\bar{b}} - \frac{1}{2} \cdot \Delta_{\hat{y}(\vec{x})}\Big|_{\vec{b}\in[\vec{b}]}; \vec{\varphi}^{T}(\vec{x}) \cdot \vec{\bar{b}} + \frac{1}{2} \cdot \Delta_{\hat{y}(\vec{x})}\Big|_{\vec{b}\in[\vec{b}]}]. \qquad (11)$$

Розглянемо тепер співвідношення між похибкою прогнозування (9) та похибкою прогнозування $\Delta_{\hat{y}(\vec{x})}$ інтервальної моделі при $\vec{b} \in \Omega$, тобто коли вона обчислюється як різниця меж коридору (6).

У силу виконання (2) – як рівності для лінійно-параметричних функцій у вигляді, а також монотонності включення інтервальних обчислень, справедлива така нерівність:

. (12)

Дійсно, оскільки , а , то оцінка виходу у інтервальній арифметиці (в силу монотонності включення інтервальних обчислень) включає коридор . Виконання рівності (2) у цьому випадку забезпечує справедливість включення:

$$[\hat{y}(\vec{x})] \subseteq [\hat{y}(\vec{x})]\Big|_{\vec{b}\in[\vec{b}]} . \qquad (13)$$

Звідси витікає справедливість нерівності (12).

Отже, нерівність (12) показує, що значення похибки прогнозування інтервальної моделі з параметрами, що належать множині розв'язків інтервальної системи лінійних алгебраїчних рівнянь ($\vec{b} \in \Omega$), у будь-якій точці області χ експерименту менше або дорівнює значенню похибки прогнозування інтервальної моделі при інтервальній локалізації вектора $\vec{b}$ її параметрів.

Для лінійної інтервальної моделі ($\vec{\varphi}^T(\vec{x}) = (x_1, x_2, ..., x_m)^T$) значення похибки $\Delta_{\hat{y}(\vec{x})}$ пропорційне відстані від центру області експерименту χ, а її максимальне значення досягається на межі області χ. Із формули (9) витікає, що дана властивість справедлива для інтервальних моделей з інтервальними параметрами $\vec{b} \in [\vec{b}]$, у яких усі базисні функції $\varphi_j(\vec{x})\,(j = 1,...,m)$ є монотонно зростаючими по модулю, коли відстань від центру $\vec{x}_0$ експерименту збільшувати у будь-якому напрямку. Якщо пронормувати незалежні змінні $x_1,...,x_n$ так, щоб центр експерименту збігався з нульовою точкою $\vec{x}_0 = (0,...,0)^T$, то розглянута властивість стає справедливою для поліноміальних моделей.

Для інтервальних моделей з інтервальними оцінками параметрів, заданих лінійно-параметричними функціями, у яких базові функції не відповідають розглянутій властивості, картина зміни похибки прогнозування на області експерименту як і у випадку інтервальних моделей при , може бути достатньо складною. Для підтвердження цього факту розглянемо графік залежності похибки прогнозування від вхідних змінних для інтервальних моделей з коридору

$$[\hat{y}(x_1, x_2)] = [4;3] \cdot x_1 + [1;3,5] \cdot \sin(x_2) + [1;2] \cdot x_2^2 + [-2;3] \cdot \cos(x_2).$$

На рисунку 2 наведена функція похибки прогнозування, отримана для інтервальних моделей розглянутого коридору.

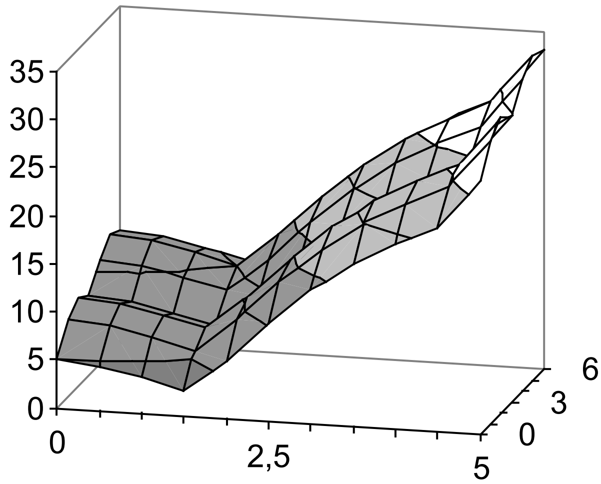

Рис.2. Функція похибки прогнозування

Отже, проведений порівняльний аналіз дозволяє зробити висновок, що за умови локалізації параметрів $\vec{b}$ інтервальної моделі у вигляді інтервального вектора $[\vec{b}]$, у загальному випадку значення похибки прогнозування отриманих інтервальних моделей збільшується, а межі функціонального коридору прогнозування (7) в просторі експерименту залишаються кусковими функціями. Виграш при цьому в порівнянні із інтервальними моделями полягає у суттєвому зменшенні обчислювальних витрат в зв'язку із спрощенням алгоритму розрахунку інтервальних значень виходу у фіксованих точках, а для інтервальних моделей, які задовольняють умовам справедливості властивості 2 – у гладкості границь функціонального коридору.

2.3. Інтервальна оцінка за умов апроксимації множини коефіцієнтів n-вимірними еліпсоїдами

$$[y(\vec{x})]\Big|_{\vec{b}\in\Omega} = [\min_{\vec{b}\in\Omega}\vec{\varphi}^T(\vec{x})\cdot\vec{b}, \max_{\vec{b}\in\Omega}\vec{\varphi}^T(\vec{x})\cdot\vec{b}]$$

Нехай знайдена деяка оптимальна еліпсоїдальна оцінка множини коефіцієнтів $\vec{b}$, тобто $\Omega\in Q_m$.

Шукаємо апроксимацію Ω так, щоб

Нехай для множини параметрів Ω інтервальних моделей отримана оптимальна еліпсоїдальна оцінка Q_m. Під оптимальною в даному випадку будемо розуміти оцінку:

$$Q_m = \{\vec{b}\in R^m \Big| (\vec{b}-\vec{\bar{b}})^T\cdot H\cdot(\vec{b}-\vec{\bar{b}}) = \gamma\},$$

яка в просторі параметрів $\vec{b}$ моделі є n-вимірним еліпсоїдом мінімального об'єму і при цьому справедливе таке включення:

$$\Omega\subseteq Q_m.$$

Як і у випадку інтервальної локалізації параметрів, для побудови моделей систем із застосуванням методів локалізації множини параметрів еліпсоїдальними множинами, важливим є вивчення та застосування властивостей отриманого при цьому коридору моделей.

Коридор інтервальних моделей $[\hat{y}(\vec{x})]\Big|_{\vec{b}\in Q_m}$, у випадку локалізації множини Ω n-вимірним еліпсоїдом, називається наближенням коридору

.

Враховуючи симетричність еліпсоїда, цей коридор матиме такий вигляд:

, (*)

де – похибка прогнозування.

Користуючись формулою $\Delta_{y(\vec{x})}\Big|_{\vec{b}\in Q_m} = 2\cdot\sqrt{\vec{\varphi}^T(\vec{x})\cdot(F^T\cdot E^{-2}\cdot F)^{-1}\cdot\vec{\varphi}(\vec{x})\cdot m}$ для знаходження верхньої оцінки $\Delta_{y(\vec{x})}\Big|_{\vec{b}\in Q_m}$ функції $\Delta_{y(\vec{x})}$ похибки прогнозування інтервальної моделі, а також з врахуванням заміни $F^T\cdot E^{-2}\cdot F = H$ та $m=\gamma$, для похибки прогнозування отримаємо наступну формулу:

$$\Delta_{y(\vec{x})}\Big|_{\vec{b}\in Q_m} = 2\cdot\sqrt{\vec{\varphi}^T(\vec{x})\cdot H^{-1}\cdot\vec{\varphi}(\vec{x})\cdot\gamma}\,. \qquad (**)$$

Як видно з виразів (*) та (**), функції похибки прогнозування та меж коридору інтервальних моделей для даного випадку є неперервними та не кусковими, що забезпечує аналітичність розрахунку інтервалу прогнозування у заданій точці області експерименту. Ця властивість є достатньо вагомою на користь застосування еліпсоїдальних методів локалізації.

Як і у випадку інтервальної локалізації множини параметрів, похибка прогнозування при локалізації еліпсоїдом $\Delta_{\hat{y}(\vec{x})}\Big|_{\vec{b}\in Q_m}$ не менша (переважно більша) від похибки прогнозування $\Delta_{\hat{y}(\vec{x})}$ на усій області експерименту, тобто

$$\Delta_{\hat{y}(\vec{x})}\Big|_{\vec{b}\in Q_m} \geq \Delta_{\hat{y}(\vec{x})},\ \forall\vec{x}\in\chi\,.$$

Цей факт витікає із включення $\Omega\subseteq Q_m$ та лінійності за параметрами інтервальної моделі.

ТЕМА 3. МАШИННА ІНТЕРВАЛЬНА АРИФМЕТИКА

3.1. Машинні операції над дійсними числами
3.2. Машинні операції над інтервалами

3.1. Машинні операції над дійсними числами

Тепер зупинимось на питаннях реалізації інтервальних операцій на цифровій обчислювальній машині. Відомо, що на комп'ютері може бути представлено лише кінцева множина чисел. Частіше за все вони записуються в напівлогарифмічній формі, а точніше – у формі з плаваючою крапкою:

$$x = m \cdot b^{e},$$

де m - мантиса, b - основа степені, e - порядок. Як правило, для внутрімашинного представлення вибирають основу b рівну 2, а мантиса нормується, тобто її абсолютне значення належить інтервалу $[1/2;1)$. Ціле e належить інтервалу $[e_{\min}, e_{\max}]$.

Множина машинних чисел описаного типу позначається через R_m. Для апроксимації дійсних чисел, які належать інтервалу $[\min\limits_{y \in R_m} y; \max\limits_{y \in R_m} y]$, можна використовувати машинні числа $\{\tilde{x} | \tilde{x} \in R_m\}$. Апроксимація досягається застосуванням відображення:

$$fl : R \ni x \to \tilde{x} = fl(x) \in R_m.$$

Таке відображення називається заокругленням, якщо виконується властивість:

(монотонність).

Заокруглення, яке відображає в так, що

називається *оптимальним заокругленням*.

Особливий інтерес представляють так звані напрями заокруглення. Якщо для заокруглення $\downarrow$ справедлива імплікація

$$x \in R \Rightarrow \downarrow x \le x\ ,$$

то говорять про заокруглення вниз. Аналогічно,

$$\uparrow x := -(\downarrow (-x)),\ x \le R\ ,$$

визначається заокруглення вверх.

Подібно до того як дійсні числа наближаються за допомогою машинних, можна дійсні інтервали наближати машинними інтервалами. В цьому випадку інтервал $[x^{-};x^{+}]$ з $I(R)$, для якого справедливо співвідношення $[x^{-};x^{+}] \subseteq [\min\limits_{y \in R_m} y; \max\limits_{y \in R_m} y]$, заміняється відповідним машинним інтервалом із множини: $I(R_m) = \{[x^{-};x^{+} \mid x^{-}, x^{+} \in R_m, x^{-} \le x^{+}]\} \subset I(R)$.

Для того, щоб основні властивості інтервальних операцій виконувались і для їх машинних аналогів, застосовується заокруглення інтервалів (інтервальне заокруглення):

$$\updownarrow : I(R) \ni [x^{-};x^{+}] \rightarrow \updownarrow [x^{-};x^{+}] \in I(R_m)\ .$$

Заокруглення інтервалу $[x^{-};x^{+}]$ полягає у знаходженні $\updownarrow [x^{-};x^{+}]$ по правилу:

$$\updownarrow [x^{-};x^{+}] = [\downarrow x^{-}; \uparrow x^{+}]\ .$$

Якщо над двома машинними числами x та y із R_m відбувається машинна операція *, де $* \in \{+,-,\cdot,/\}$ з відповідним заокругленням fl, то її результатом буде нове число із діапазону :

.

Таким самим чином можна визначити результат машинної операції над інтервалами.

3.2. Машинні операції над інтервалами

Нехай $[a^-;a^+],[b^-;b^+]\in I(R_m)$, $*\in\{+,-,\cdot,/\}$, $\updownarrow$ - інтервальне заокруглення. Тоді результат операції *, виконаної над $[a^-;a^+]$ і $[b^-;b^+]$ із застосуванням $\updownarrow$ є:

$$[c^-;c^+]=\updownarrow([a^-;a^+]*[b^-;b^+])\in I(R_m).$$

При цьому основні властивості інтервальної арифметики зберігаються.

Для машинних інтервальних операцій $*\in\{+,-,\cdot,/\}$ справедливо наступне твердження:

$$\begin{aligned}&[a_k^-;a_k^+],[b_k^-;b_k^+]\in I(R_m),*\in\{+,-,\cdot,/\},[a_k^-;a_k^+]\subseteq[b_k^-;b_k^+],k=1,2,\\&\Rightarrow[c_1^-;c_1^+]=\updownarrow([a_1^-;a_1^+]*[a_2^-;a_2^+])\subseteq\updownarrow([b_1^-;b_1^+]*[b_2^-;b_2^+])\end{aligned}.$$

Нехай $\updownarrow$ - інтервальне заокруглення, яке зводиться до напрямленим заокругленням $\downarrow$ та $\uparrow$, і нехай $*\in\{+,-,\cdot,/\}$. Тоді

$$\begin{aligned}&[a^-;a^+],[b^-;b^+]\in I(R_m)\Rightarrow[a^-;a^+]*[b^-;b^+]\subseteq[c_1^-;c_1^+]=\updownarrow([a^-;a^+]*[b^-;b^+])\in I(R_m),\\&a\in[a^-;a^+],b\in[b^-;b^+]\Rightarrow a*b\in[c_1^-;c_1^+]=\updownarrow([a^-;a^+]*[b^-;b^+])\in I(R_m)\end{aligned}$$

Якщо є заокруглення fl, застосування якого приводить до виконання нерівності

$$\downarrow a\le fl(a)\le\uparrow a,\ a\in R,$$

то для x,y,z із R_m справедливо

$$z=fl(x*y)\in[z^-;z^+]=\updownarrow([x;x]*[y;y])\in I(R_m).$$

Інтервальне оцінювання аналітичного виразу функції, проведене з використанням операцій , дає інтервали, які включають значення оцінюваної функції. Серед цих інтервалів знаходяться і оцінки множини значень функції. Більш того, при виконанні подібних обчислень зберігається властивість монотонності включення.

Для двох машинних інтервалів і

$$\updownarrow([a^-;a^+]*[b^-;b^+]) = fl_1([a^-;a^+]*[b^-;b^+]) = [(1-\varepsilon_1)([a^-;a^+]*[b^-;b^+])_1,$$
$$(1+\varepsilon_2)([a^-;a^+]*[b^-;b^+])_2]$$
.

За допомогою $([a^-;a^+]*[b^-;b^+])_1$ та $([a^-;a^+]*[b^-;b^+])_2$ обчислюються межі точного значення результату, причому

$$-\varepsilon_1([a^-;a^+]*[b^-;b^+])_1 \le 0, \quad \varepsilon_2([a^-;a^+]*[b^-;b^+])_2 \ge 0$$

і

$$|\varepsilon_1|,|\varepsilon_2| \le b^{1-t_1},$$

де t_1 - t_1-розрядна мантиса.

Оцінкою ширини результату служить:

$$d(fl_1([a^-;a^+]*[b^-;b^+])) \le d([a^-;a^+]*[b^-;b^+]) + 2b^{1-t_1}\left|[a^-;a^+]*[b^-;b^+]\right|.$$

Ця оцінка показує, що коли використовується мантиса фіксованої довжини, то зростання ширини $d(fl_1([a^-;a^+]*[b^-;b^+]))$ визначається величиною $\left|[a^-;a^+]*[b^-;b^+]\right|$.

ТЕМА 4. ЛОКАЛІЗАЦІЯ НУЛІВ ФУНКЦІЙ ОДНОЇ ДІЙСНОЇ ЗМІННОЇ

4.1.Загальна постановка задачі
4.2.Інтервальні модифікації методу Ньютона
4.3.Оптимальний метод вибору умови поділу на е лінійність
4.4.Квадратично збіжні методи

4.1. Загальна постановка задачі

Розглядається задача локалізації нулів дійсної функції $y(x)$ із використанням інтервальних методів, які дозволять знайти множину інтервалів найменшої ширини, таких, що кожен інтервал включає один, або декілька інтервалів із заданого початкового інтервалу. Прості реалізації даних методів передбачають ділення початкового інтервалу. Це інтервальні варіанти методів двійкового пошуку, або інших методів пошуку.

Нехай відомий початковий інтервал X^0, який $x \in X^0$ включає хоча б один нуль функції $y(x)$. Щоб покращити локалізацію нулів (звузити інтервал X^0**)**, ми ділимо його пополам точкою:

$$m(X^0) = (x_1^0 + x_2^0)/2$$

на два інтервали U^0 та V^0, такі, що $X^0 = U^0 \bigcup V^0 = [x_1^0, m(X^0)] \bigcup [m(X^0), x_2^0]$.

Якщо $0 \in y(U^0)$, то U^0 включає нуль функції $y(x)$ і тому повторюємо процедуру половинного поділу для інтервалу . Або якщо 0 , то процедура поділу повторюється для . В протилежних випадках ігноруємо (відкидаємо) відповідні інтервали. Дана процедура породжує послідовність е лінійність, які містяться в і можуть мати нулі функції . Ширина цих інтервалів прямує до нуля, так як вона зменшується вдвічі на кожному кроці.

Дана процедура може призвести до породження великої кількості е лінійність, що є її основним недоліком.

Зауважимо, що процедура $m(X^0)$ в загальному випадку означає вибір будь-якої точки із інтервалу X^0.

4.2. Інтервальні модифікації методу Ньютона

Розглянемо модифікації методу Ньютона.

Нехай маємо неперервну функцію $y(x)$, яка має нуль в заданому інтервалі $X^0=[x_1^0;x_2^0]$, тобто $y(\xi)=0$ для деякого $\xi \in X^0$. Нехай $y(x_1^0)<0$ та $y(x_2^0)>0$ для межових точок інтервалу X^0. Нехай m_1 та m_2 – межі різницевих відношень

$$0<m_1 \le \frac{y(x)-y(\xi)}{x-\xi}=\frac{y(x)}{x-\xi} \le m_2 < \infty,\ \ \xi \ne x \in X^0 .$$

Ці межі визначають інтервал $M=[m_1;m_2]\in I(R)$. Аналогічне справджується для випадку $y(x_1^0)>0$, $y(x_2^0)<0$ та $m_2<0$.

Очевидно, що для заданих допущень функція $y(x)$ не має інших коренів в X^0.

Починаючи із вихідного інтервалу $\xi \in X^0$ обчислюємо ітераційно нові інтервали $X^k, k \ge 1$ відповідно до такої процедури:

, (4.1)

де .

Одна ітерація процесу проілюстрована на рисунку 4.1.

Рис. 4.1.

Формулу (4.1) можна описати і не в інтервальному вигляді:

$$x_{1k+1} = \begin{cases} \max\{x_{1k}, m(X_k) - y(m(X_k))/m_1\}, \text{якщо} \quad y(m(X_k)) \ge 0 \\ m(X_k) - y(m(X_k))/m_2, \quad \text{якщо} \quad y(m(X_k)) \le 0 \end{cases} \tag{4.2}$$

$$x_{2k+1} = \begin{cases} m(X_k) - y(m(X_k))/m_2, \quad \text{якщо} \quad y(m(X_k)) \ge 0 \\ \min\{x_{2k}, m(X_k) - y(m(X_k))/m_1, \text{якщо} \quad y(m(X_k)) \le 0 \end{cases} \tag{4.3}$$

Слід зауважити *важливі властивості послідовності ітерацій* $\{X_k\}_{k=0}^{\infty}$, яка обчислена по формулам (4.1)-(4.3):

1) $\xi \in X_k$, $k \ge 0$

2) $X^0 \supset X^1 \supset X^2 \supset \cdots$, де $\lim\limits_{k\to\infty} X_k = \xi$, або дана послідовність стабілізується за кінцеву кількість кроків на точці $[\xi;\xi]$];

3) .

4) Якщо , то для послідовності наближень вірна нерівність , яка уточнює властивість 3.

Четверта властивість показує, що якщо вибрати середину інтервалу за , то ширина локалізаційного інтервалу зменшуватиметься вдвічі на кожній ітерації.

Зауважимо, що інтервал $[m_1; m_2]$ можна локалізувати для випадку неперервно диференційованої функції $y(x)$ найменшим та найбільшим значенням її першої похідної.

4.3. Оптимальний метод вибору умови поділу на е лінійність

У попередньому методі ітерацій локалізації нулів дійсної функції, залежно від вибору точок $m(X_k)$ на інтервалах X_k отримуватимемо різні послідовності локалізаційних інтервалів $X^0 \supset X^1 \supset X^2 \supset \cdots$, тому тут поставимо задачу вибору $m(X_k)$ у такий спосіб, щоб ширина окремих елементів послідовності була найменшою.

Проведемо математичну формалізацію даного уточнення.

Нехай $y(x)$ належить класу функцій $\varphi[X]$, які володіють наступними властивостями:

а) $y(x_1) < 0$ та $y(x_2) > 0$;

б) для інтервалу $M = [m_1; m_2]$, такого, що $m_1 > 0$, справедливо

$$m_1 \le (y(x) - y(z))/(x - z) \le m_2 \text{ для усіх } x, z \in [X].$$

Очевидно, що така функція $y(x) \in \varphi[X]$, як і у попередньому методі, має тільки один корінь ξ в інтервалі $X = [x_1, x_2]$. Для його локалізації можна використати метод побудови послідовності за формулою (4.1).

Процес вибору розглядатимемо покроково.

Як видно із ітераційного методу локалізації нуля функції за формулою (4.1) чи формулами (4.2), (4.3), для обчислення нового наближення нам потрібно та . Якщо зафіксувати величину , то

X_{k+1} буде залежати тільки від $y(m(X_k))$. Оскільки значення $y(x)$ може змінюватися лише в межах від $y(x_{1k})$ до $y(x_{2k})$, то значення $y(m(X_k))$ також фіксовані. Це дозволяє визначити найбільшу можливу ширину для X_{k+1}:

$$\max\{d(X_{k+1})|m(X_k)=x, y_{1k}\le y(m(X_k))\le y_{2k}\}. \tag{4.4}$$

Це буде «найгіршим» випадком для функції $y(x)$, а «найкраще», тобто оптимальне значення $\tilde{x}=m(X_k)$ вибиратимемо за формулою:

$$\min_{x\in X_k}\{\max\{d(X_{k+1})|m(X_k)=x, y_{1k}\le y(m(X_k))\le y_{2k}\}, \tag{4.5}$$

тобто із умови мінімізації найбільшої ширини (4.4) (для найгіршого випадку).

Нехай метод ітерацій застосовується до функцій $y(x)\in\varphi[X]$. Якщо використовується правило

$$m(X_k)=1/2(x_{1k}+x_{2k}),\quad 0\le k\le i,\quad i\ge 0,$$

То максимальна ширина $d(X_{i+1})$ для функцій $y(x)\in\varphi[X]$ менше, ніж для будь-яких інших виборів точки $m(X_k)$. Якщо $y(x)\in\varphi[X]$, то маємо:

$$d(X_{i+1})\le 1/2^{i+1}(1-m_1/m_2)_{i+1}d(X^0).$$

4.4. Квадратично збіжні методи

В методі ітерацій використовується фіксована пара меж для різницевих відношень функції . Ця процедура відповідає інтервальному варіанту спрощеному методу Ньютона. Якщо ж допустити, що неперервно диференційована і для похідної є інтервальна оцінка , то можна визначити інтервальний варіант і для звичайного методу Ньютона.

Нову процедуру можна отримати шляхом модифікації методу ітерацій, якщо замінити інтервал M на інтервал

$$M_k = y'(X_k) \tag{4.5}$$

на k-ому кроці ітерації. Якщо відомі апріорні оцінки

$$0 < l_1 \le y'(x) \le l_2 ,$$

то можна гарантувати оцінку $m_1 > 0$ і використовувати вираз

$$M_k = [m_{1k}, m_{2k}] = y'(X_k) \cap L , \quad L = [l_1, l_2] . \tag{4.6}$$

Таким чином, ми отримуємо

$$X_{k+1} = \{m(X_k) - y(m(X_k)) / M_k\} \cap X_k . \tag{4.7}$$

ТЕМА 5. ІНТЕРВАЛЬНА СИСТЕМА ЛІНІЙНИХ АЛГЕБРАЇЧНИХ РІВНЯНЬ

5.1.Особливості загальної постановки задачі

Розглянемо основні припущення, на яких базуються методи аналізу інтервальних даних у випадку побудови моделей статичних систем.

У вітчизняній літературі ці гіпотези вперше були сформульовані в рамках теоретико-множинного підходу до задач параметричної ідентифікації:

Н1. Статична система (об'єкт) описується лінійно-параметричним рівнянням

$$y_o = \beta_1 \cdot \varphi_1(\vec{x}) + \ldots + \beta_m \cdot \varphi_m(\vec{x}), \tag{5.1}$$

де y_o – істинне невідоме значення виходу системи;

$\vec{x} \in R^n$ – вектор вхідних змінних;

$\vec{\beta} = (\beta_1, \ldots, \beta_m)^T$ – вектор невідомих параметрів;

$\vec{\varphi}^T(\vec{x}) = (\varphi_1(\vec{x}), \ldots, \varphi_m(\vec{x}))^T$ – вектор відомих базисних функцій.

Н2. Результати експерименту представлені у вигляді матриці X значень вхідних змінних і відповідних інтервальних значень вихідної змінної y :

(5.2)

Припускають, що в довільному -у спостереженні істинне значення виходу $y_{oi} = \vec{\varphi}^T(\vec{x}_i) \cdot \vec{\beta}$ належить інтервалу $[y_i^-, y_i^+]$, тобто $y_i^- \le y_{oi} \le y_i^+$.

Інтервальні значення виходу $[y_i^-, y_i^+]$ можуть бути одержані на основі моделей інтервальних похибок, розглянутих у попередньому підрозділі.

Завданням аналізу інтервальних даних є оцінювання невідомого вектора $\vec{\beta}$ так, щоб значення функції $y = \vec{\varphi}^T(\vec{x}) \cdot \vec{\beta}$ в точках експерименту належали відповідним інтервалам виходу. Якщо оцінка $\vec{b}$ вектора $\vec{\beta}$ існує, то одержану функцію $\hat{y}(\vec{x}) = \vec{\varphi}^T(\vec{x}) \cdot \vec{b}$ називатимемо моделлю статичної системи.

Згідно сформульованих гіпотез, шуканий вектор $\vec{b}$ повинен задовольняти таку систему N нерівностей з m невідомими:

$$\begin{cases} y_1^- \le b_1\varphi_1(\vec{x}_1) + \ldots + b_m\varphi_m(\vec{x}_1) \le y_1^+; \\ \vdots \\ y_i^- \le b_1\varphi_1(\vec{x}_i) + \ldots + b_m\varphi_m(\vec{x}_i) \le y_i^+; \\ \vdots \\ y_N^- \le b_1\varphi_1(\vec{x}_N) + \ldots + b_m\varphi_m(\vec{x}_N) \le y_N^+; \end{cases} \tag{5.3}$$

Оскільки кожна i-та нерівність у системі (5.3) забезпечує належність значення функції $\hat{y}(\vec{x})$ в i-тій точці експерименту, відповідному i-тому інтервалу виходу, то одночасне виконання умов, заданих нерівностями системи, означає існування розв’язку задачі, тобто “проходження” функції через усі інтервали.

Розглянемо деякі важливі властивості системи (5.3) та її розв’язків.

Система (5.3) є системою лінійних нерівностей відносно невідомих

.

е лінійність функцій в (5.3) не суперечить попередньому твердженню, тому, що при відомому аргументі $\vec{x}_i$ вони стають відомими коефіцієнтами.

Якщо згадані коефіцієнти позначити через $\phi_{ij} = \varphi_j(\vec{x}_i)$, то систему (5.3) можна переписати у такому вигляді:

$$\begin{cases} y_1^- \le \phi_{11} b_1 + \ldots + \phi_{1m} b_m \le y_1^+ \\ \vdots \\ y_N^- \le \phi_{N1} b_1 + \ldots + \phi_{Nm} b_m \le y_N^+, \end{cases}$$

звідки очевидна її лінійність. В майбутньому нам зручно буде розглядати систему (5.3) в матричному вигляді

$$\vec{Y}^- \le F \cdot \vec{b} \le \vec{Y}^+, \tag{5.4}$$

де $\vec{Y}^- = \{y_i^-, i = 1, \ldots, N\}, \vec{Y}^+ = \{y_i^+, i = 1, \ldots, N\}$ – вектори, складені із верхніх та нижніх меж інтервалів $[y_i^-, y_i^+]$, відповідно;
$F = \{\phi_{ij}, i = \overline{1, N}, j = \overline{1, m}\}$ – відома матриця значень базисних функцій.

Система (5.4) може не мати жодного розв’язку, тобто бути несумісною або мати багато розв’язків.

Стосовно задач аналізу інтервальних даних, несумісність системи (5.4) означає, що не виконуються припущення методу, тобто або невірно задано вигляд функції (5.1), або невірно визначені інтервали $[y_i^-, y_i^+]$. Обидва порушення гіпотез не забезпечують належність значень функції в точках експерименту до відповідних інтервалів виходу.

Нехай система (5.4) є сумісною. Позначимо через множину її розв’язків, тобто

(5.5)

Властивості цієї множини, розглядалися у багатьох працях. Наведемо основні із них, важливі для моделювання статичних систем.

1. У просторі параметрів $\beta_1,...,\beta_m$ множина Ω є опуклий многогранник. Це означає, що довільна точка множини Ω є розв'язком системи (5.4).

2. Довільний розв'язок $\vec{b} \in \Omega$ системи породжує модель $\hat{y}(\vec{x}) = \vec{\varphi}^T(\vec{x}) \cdot \vec{b}$, що "проходить" через усі інтервали $[y_i^-, y_i^+]$, яку надалі називатимемо інтервальною моделлю.

3. Множина розв'язків Ω породжує множину рівнозначних (з точки зору наявної інтервальної невизначеності) інтервальних моделей, кожна з яких задовольняє умовам задачі. При цьому, всі інтервальні моделі знаходяться у коридорі:

$$[\hat{y}(x)] = [\hat{y}^-(x); \hat{y}^+(x)], \tag{5.6}$$

де $\hat{y}^-(\vec{x}) = \min\limits_{\vec{b}\in\Omega}\left(\vec{\varphi}^T(\vec{x}) \cdot \vec{b}\right)$ та $\hat{y}^+(\vec{x}) = \max\limits_{\vec{b}\in\Omega}\left(\vec{\varphi}^T(\vec{x}) \cdot \vec{b}\right)$ – нижня та верхня межі функціонального коридору.

4. Істинний невідомий вектор $\vec{\beta}$ є одним із розв'язків системи (5.4), тобто $\vec{\beta} \in \Omega$. Тому можна стверджувати, що довільна точка множини Ω може бути істинним вектором параметрів. Ця властивість множини розв'язків Ω дозволяє трактувати її як множину можливих значень невідомих параметрів $\beta_1,...,\beta_m$.

Звідси випливає, що чим "ширша" множина , тим більша невизначеність відносно істинних параметрів статичної системи.

Розмір множини характеризується діаметром , який визначається як відстань між двома найбільш віддаленими точками множини:

, (5.7)

де $\vec{b}_p$, $\vec{b}_s$ – відповідні вершини області Ω.

Діаметр множини Ω тісно пов'язаний з матрицею F системи (5.4). Зокрема, якщо кількість різних точок $\vec{x}_i$ спостережень у матриці F буде менша від кількості невідомих параметрів m, то множина Ω буде "розірвана". Тобто, якщо $rang(F) < m$, то $d \to \infty$. З іншого боку, якщо $rang(F) = m$, то діаметр d обмежений.

5.2.Особливості застосування традиційних алгоритмів розв'язку СЛАР

Нехай $[F]$ - інтервальна матриця, $[\vec{Y}]$ - інтервальний вектор, обернення F_k^{-1} існує для всіх $F_k \in [F]$. Необхідно знайти множину:

$$\Omega = \left\{ \vec{b} \in R^m \,\middle|\, [F] \cdot \vec{b} = [\vec{Y}] \right\}$$

Дана множина в загальному випадку не має простого опису. Тому обмежуються її локалізацією за допомогою інтервального вектора. Один із способів знаходження такого інтервального вектора – застосування безпосереднього узагальнення методу Гауса на системи з інтервальними коефіцієнтами.

Нехай нам відома таблиця коефіцієнтів

$$\begin{matrix} F_{11} & \cdots & F_{1n} & Y_1 \\ \vdots & & \vdots & \vdots \\ F_{n1} & \cdots & F_{nn} & Y_n \end{matrix}.$$

Застосовуючи формули

, , ,

, ,

, ,

$F_{i1}^{'} = 0$, $2 \le i \le n$,

із врахуванням того, що $0 \notin F_{11}$, обчислюємо нову таблицю коефіцієнтів:

$$\begin{matrix} F_{11}^{'} & F_{12}^{'} & \cdots & F_{1n}^{'} & Y_1^{'} \\ 0 & F_{22}^{'} & \cdots & F_{2n}^{'} & Y_2^{'} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & F_{n2}^{'} & \cdots & F_{nn}^{'} & Y_n^{'} \end{matrix}.$$

З теорії лінійних рівнянь відомо, що система рівнянь $[F_k^{'}] \cdot \vec{b} = [\vec{Y}_k^{'}]$ має ті ж рішення, що і $[F_k\ ] \cdot \vec{b} = [\vec{Y}_k\ ]$.

Якщо описаний вище крок провести $n-1$ раз, то початкова таблиця коефіцієнтів перетворюється у верхню трикутну матрицю:

$$\begin{matrix} \widetilde{F}_{11} & \widetilde{F}_{12} & \cdots & \widetilde{F}_{1n} & \widetilde{Y}_1 \\ & \widetilde{F}_{22} & & \vdots & \vdots \\ & & \ddots & & \\ & & & \widetilde{F}_{nn} & \widetilde{Y}_n \end{matrix},$$

для якої виконується включення:

$$\left\{\vec{b}\,\middle|\,[F] \cdot \vec{b} = [\vec{Y}]\right\} \subseteq \left\{\vec{b}\,\middle|\,[\widetilde{F}] \cdot \vec{b} = [\widetilde{\vec{Y}}]\right\}.$$

Використовуючи формули:

$b_n = \widetilde{Y}_n / \widetilde{F}_{nn}$,

$b_i = (\widetilde{Y}_i - \sum_{j=i+1}^{n} \widetilde{F}_{ij} X_j) / \widetilde{F}_{ii}$, ,

ми отримуємо інтервальний вектор , який задовольняє умові:

.

5.3.Ітераційні алгоритми знаходження розв’язків ІСЛАР

Залежно від цілей моделювання та невизначеностей у вхідних та вихідних змінних моделі динамічної системи, в інтервальному аналізі розглядають узагальнені множини розв'язків лінійної інтервальної алгебраїчної системи рівнянь. Серед них на практиці найбільш частіше зустрічаються задачі знаходження:

- об'єднаної множини розв'язків:

$$\Omega_{gar} = \left\{ \vec{b} \in R^m \middle| (\exists F \in [F]), (\exists \vec{Y} \in [\vec{Y}]), (\vec{b}^T \cdot F = \vec{Y}) \right\};$$

- допустимої множини розв'язків:

$$\Omega_{dop} = \left\{ \vec{b} \in R^m \middle| (\forall F \in [F]), (\forall \vec{Y} \in [\vec{Y}]), (\vec{b}^T \cdot F = \vec{Y}) \right\};$$

- алгебраїчного інтервального розв'язку:

$$\vec{b}^T \cdot [F] = [\vec{Y}].$$

Об'єднану і допустиму множини розв'язків переважно використовують у випадку розв'язування задач параметричної ідентифікації та наближення складних моделей систем простішими.

У загальному випадку, для задач високої розмірності інтервальної системи алгебраїчних рівнянь чисельно визначити узагальнені множини її розв'язків неможливо. Тому в інтервальному аналізі застосовують ітераційні процедури зовнішнього та внутрішнього наближення, які називаються інтервальною локалізацією чи допустимим оцінюванням.

У випадку локалізації множини розв'язків ІСЛАР, наближення проводиться описаним прямокутним паралелепіпедом (зовнішнє оцінювання). У результаті отримують інтервальний вектор , який включає множину .

При допустимому оцінюванні знаходять вписаний прямокутний паралелепіпед (внутрішнє оцінювання). У результаті отриманий вектор

$[\vec{b}]$ належить множині Ω_{dop}. Зауважимо, що у випадку допустимого інтервального оцінювання множини Ω_{dop} задачу оцінювання називають «задачею про допуски» і з метою забезпечення оптимальності розв'язків $[\vec{b}]$, у випадку наближення паралелепіпедом Π^{+} мінімізують його об'єм, а у випадку допустимого оцінювання за допомогою паралелепіпеда Π^{-} знаходять його максимальний об'єм.

В інтервальному аналізі значна частина ітераційних алгоритмів локалізації розв'язків системи ІСЛАР побудована на такій властивості системи, що точні значення вектора $[\vec{b}]$ досягаються на межових значеннях матриці $[F]$ та вектора $[\vec{Y}]$.

На першій ітерації даних алгоритмів знаходиться Π- множина – як наближення розв'язку Ω_m для m інтервальних рівнянь. В результаті отримується інтервальний вектор оцінок параметрів: $[\vec{b}_i] = ([b_{i1,}],...,[b_{ij,}],...,[b_{im}])^T$. Наступними $k = m+1,...,N-m$ ітераціями алгоритму є уточнення Π- множини оцінок параметрів ІСЛАР шляхом послідовного додавання решти $N-m$ інтервальних рівнянь із ІСЛАР і знаходження розв'язку отриманої системи на $k = m+l$ -ій ітерації.

Отримана на $k = m+l$ -ій ітерації система суттєво спрощується внаслідок послідовного наближення множини загального розв'язку не усіх рівнянь ІСЛАР, а лише частини. На кожній ітерації алгоритму система включає одне інтервальне рівняння із загальної ІСЛАР та m рівнянь з попередньої ітерації наближення - множиною. Отже, після реалізації алгоритмів отримують наближення загального розв'язку ІСЛАР у вигляді множин та .

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Алефельд Г., Херцбергер Ю. Введение в интервальные вычисления: Пер. с англ.- М.: Мир, 1987. – 360с.
2. Астафьев Н.М. Линейные неравенства и выпуклость. – М.: Наука, 1982. – 153 с.
3. Бакан Г.М., Куссуль Н.Н. Теоретико-множественная идентификация линейных объектов в классе размытых эллипсоидальных множеств // Автоматика.- 1990. - №3.- С. 29-40..
4. Бахвалов Н.С. Численные методы. – М.: Наука, 1975. – 631 с.
5. Бахвалов Н.С., Жидков Н.П., Кобельков Г.М. Численные методы. – М.: Мир, 1986. – 502 с.
6. Бочков А. Ф., Милевский М. В. Оценивание параметров модели для объектов с интервальной неопределенностью в выходных параметрах. - Москва, 1988. - 23 с. – Деп. в ВИНИТИ, №92б-В88.
7. Бочков А. Ф., Милевский М. В. Интервальные модели – один подход к описанию неопределенностей. – В кн.: Вопросы кибернетики. Устройства и системы. Сб. научных трудов / Под ред. Евтихиева Н.Н. – М.: 1987. – С. 19-27.