ДОПОВІДІ
НАЦІОНАЛЬНОЇ
АКАДЕМІЇ НАУК
УКРАЇНИ

8 • 2013

УДК 681.142

В. М. Заяць

Ітераційний підхід до мінімізації похибки числових методів другого порядку та їх застосування до аналізу нелінійних динамічних систем

(Представлено членом-кореспондентом НАН України О. Є. Божком)

Запропоновано ітераційний підхід до мінімізації похибки дискретизації числових методів другого порядку, який ґрунтується на модифікації методу трапецій і встановленні моменту часу, коли поправки явного і неявного методів Ейлера мають однаковий внесок до поправки для наступної точки дискретизації динамічної системи. Підтверджено доцільність його застосування до аналізу нелінійних динамічних систем коливної природи з високою добротністю та тривалими перехідними процесами.

1. При аналізі складних динамічних процесів та явищ, які можна описати системою неперервних диференційних рівнянь, поданих у нормальній формі Коші

$$\frac{dx}{dt} = f[x(t), t)],$$

де x — N-мірний вектор змінних стану; f — N-мірна вектор-функція, що описує динаміку фазових траєкторій системи, використовують числові (різницеві) методи для проведення дискретизації. Такі методи повинні бути збіжними та мати малу похибку дискретизації для забезпечення збереження якісної та кількісної відповідності між досліджуваним процесом або явищем та його дискретною моделлю [1–5]. Друга вимога до різницевих методів — це властивість A-стійкості. У протилежному випадку наявність незначної локальної похибки обмежень, допущеної на одному кроці, може призвести до нагромадження цієї похибки в процесі руху зображуючої точки вздовж фазової траєкторії і цілковитої непридатності для прикладних застосувань остаточного результату обчислень [6, 8].

У програмах комп'ютерного аналізу електронних схем [8], аналізі поведінки систем зі складною динамікою [3, 4], аналізі коливних систем з високою добротністю, для яких перехідні процеси є тривалими [6], виникає проблема між складністю різницевого алгоритму та його точністю. Як правило, застосовують методи не вище другого порядку складності або їх комбінації. Зокрема, часто використовується метод трапецій [6–8]. Різницева формула цього методу має вигляд:

$$x_{n+1} = x_n + \frac{h}{2}(\mathbf{f}_n + \mathbf{f}_{n+1}). \tag{1}$$

Ця формула є комбінацією двох методів: на першій половині кроку дискретизації застосовується явний метод Ейлера, а на другій половині — неявний метод Ейлера [8]. В результаті побудови такої комбінації, як засвідчують численні публікації, точність зростає більше ніж на порядок, порівняно з методами Ейлера. Крім того, для цього методу характерна властивість A-стійкості.

2. Спосіб мінімізації похибки дискретизації. У роботі [7] запропоновано враховувати поправки для наступної точки дискретизації не на середині кроку, а в той момент часу, коли внески явного і неявного методів Ейлера є еквівалентними. З цією метою різницеву формулу (1) подано у вигляді, запропонованому Лінігером–Уілабі:

$$x_{n+1} = x_n + h(1-\mu)\mathbf{f}_n + h \cdot \mu \cdot \mathbf{f}_{n+1}, \tag{2}$$

яка при $\mu = 0$ відповідає явному методу Ейлера; $\mu = 0{,}5$ — методу трапецій; $\mu = 1$ — неявному методу Ейлера. Прирівнявши другий і третій члени з правого боку у формулі (2), отримуємо значення параметра μ, при якому явний і неявний методи Ейлера вносять однаковий внесок у поправку до значення $\mathbf{x}_n$:

$$\mu = \frac{\mathbf{f}_n}{\mathbf{f}_n + \mathbf{f}_{n+1}}. \tag{3}$$

Після підстановки (3) в (2) одержуємо нову різницеву формулу:

$$x_{n+1} = x_n + \frac{2 \cdot h \cdot \mathbf{f}_n \cdot \mathbf{f}_{n+1}}{(\mathbf{f}_n + \mathbf{f}_{n+1})}. \tag{4}$$

Оскільки за побудовою формули (4) внесок кожного з методів Ейлера не перевищує половини віддалі між $\mathbf{x}_n$ і $\mathbf{x}_{n+1}$, то метод (4) дає гарантоване обмеження на величину похибки дискретизації на кожному кроці та забезпечує її додатність.

Геометрична ілюстрація запропонованого способу зменшення похибки дискретизації проілюстрована на рис. 1. Якщо поправки за явним та неявним методами Ейлера до наступної точки дискретизації враховувати в момент часу, що відповідає точці C, як показано на рис. 1, то отримаємо метод трапеції; в точці B маємо пропонований метод, який зрівноважує внески методів Ейлера; в точці A одержуємо оптимальну комбінацію, яка відповідає точці перетину дотичних до x_n та x_{n+1} точок дискретизації.

Для оцінки похибки методу (4) проведено аналіз похибки дискретизації на прикладі моделі консервативної системи другого порядку

$$\frac{d^2x}{dt^2} = -\omega_0^2 x,$$

де ω_0 — частота коливань консервативної системи, який підтвердив, що похибка дискретизації методу (4) пропорційна до $h^2/24$, як і в методі трапецій, але має протилежний знак і в два рази меншу абсолютну величину. Дослідження показали, що метод (4), як і метод трапецій, має властивість A-стійкості. Цей результат підтверджено розрахунком генераторних схем з тривалими перехідними процесами.

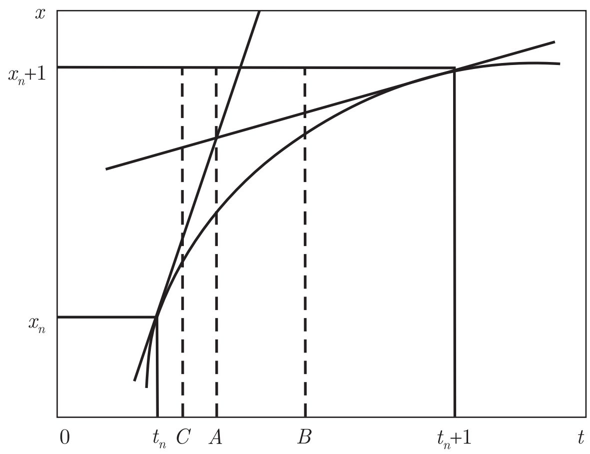

Рис. 1. Геометрична інтерпретація ітераційного підходу до мінімізації похибки дискретизації

3. Ітераційний підхід до мінімізації похибки дискретизації. Враховуючи, що похибка методу (4) і методу трапецій (формула (1)) мають протилежні знаки, можна провести їх арифметичне усереднення, тим самим зменшивши величину похибки. Застосовуючи на першій половині кроку формулу (4), а на другій — формулу (1), отримуємо різницеву формулу

$$x_{n+1} = x_n + \frac{h \cdot \mathbf{f}_n \cdot \mathbf{f}_{n+1}}{(\mathbf{f}_n + \mathbf{f}_{n+1})} + \frac{h}{4}(\mathbf{f}_n + \mathbf{f}_{n+1}), \tag{5}$$

яку назвемо різницевою комбінацією першого роду (К1Р). Похибка дискретизації при застосуванні (5) до консервативної системи виявилася вдвічі меншою, порівняно з методом (4) і протилежною за знаком по відношенню до методу трапеції. Тепер після усереднення (1) і (5) отримуємо різницеву комбінацію другого роду (К2Р):

$$x_{n+1} = x_n + \frac{h \cdot \mathbf{f}_n \cdot \mathbf{f}_{n+1}}{2(\mathbf{f}_n + \mathbf{f}_{n+1})} + \frac{3h}{8}(\mathbf{f}_n + \mathbf{f}_{n+1}). \tag{6}$$

Як засвідчили результати аналізу похибки дискретизації методу (6) при розгляді моделі без втрат, вона виявилася у чотири рази меншою за похибку методу трапецій і в два рази меншою, ніж похибка методу (5). При цьому знак похибки в К2Р збігається зі знаком похибки у методі трапецій і протилежний до похибки, який дає К1Р. Таким чином, можна очікувати подальшого зменшення величини похибки дискретизації комбінації методів (5) і (6), яка приводить до різницевої комбінації третього роду (К3Р):

$$x_{n+1} = x_n + \frac{3h \cdot \mathbf{f}_n \cdot \mathbf{f}_{n+1}}{4(\mathbf{f}_n + \mathbf{f}_{n+1})} + \frac{5h}{16}(\mathbf{f}_n + \mathbf{f}_{n+1}). \tag{7}$$

Зауважимо, що розглядати комбінацію (6) з (4) недоцільно (хоча вона й має право на існування), оскільки (5) має в чотири рази меншу похибку дискретизації, порівняно з (4). Крім того, знаки похибки в (4) і (6) збігаються.

4. Оптимальна комбінація для мінімізації похибки дискретизації. Запропоновані комбінації різницевих схем побудовано таким чином, що в комбінаціях непарного роду (К1Р, К3Р) більш істотним є внесок другого члена в отримані формули порівняно з третім, а в комбінаціях парного роду (К2Р) ці внески практично вирівнюються. Така побудова забезпечує зміну знака похибки при отриманні нової комбінації. Отже, можна сконструювати метод другого порядку, який забезпечить з точністю до членів другого порядку малості як завгодно малу похибку дискретизації. Після арифметичного усереднення (6) і (7) приходимо до різницевої схеми четвертого роду (К4Р):

$$x_{n+1} = x_n + \frac{5h \cdot \mathbf{f}_n \cdot \mathbf{f}_{n+1}}{8(\mathbf{f}_n + \mathbf{f}_{n+1})} + \frac{11h}{32}(\mathbf{f}_n + \mathbf{f}_{n+1}). \tag{8}$$

Аналізуючи формули (5)–(8), на k-му кроці, застосовуючи півкроку парну комбінацію, а півкроку непарну, отримуємо різницеву схему для комбінації k-го роду (ККР):

$$x_{n+1} = x_n + \frac{a_k \cdot h \cdot \mathbf{f}_n \cdot \mathbf{f}_{n+1}}{(\mathbf{f}_n + \mathbf{f}_{n+1})} + a_{k+1} \cdot h \cdot (\mathbf{f}_n + \mathbf{f}_{n+1}), \tag{9}$$

де

$$a_k - \frac{2^k - (-1)^k}{3 \cdot 2^{k-1}}; \qquad a_{k+1} = \frac{2^{k+1} + (-1)^k}{3 \cdot 2^{k+1}}.$$

Очевидно, з ростом k величини коефіцієнтів a_k і a_{k+1} зменшуються, що приводить до зменшення похибки дискретизації.

При цьому похибка дискритизації будь-якої k-ї комбінації може бути обчислена за формулою

$$\delta = \frac{(-1)^k}{2^{k+1}}, \tag{10}$$

про що свідчить аналіз консервативних систем другого порядку та систем з високою добротністю високих порядків.

Для мінімізації похибки дискретизації в (9) здійснимо граничний перехід, спрямувавши k до безмежності. Отримуємо різницеву схему (11), для якої з точністю до членів другого порядку малості похибка дискретизації відсутня:

$$x_{n+1} = x_n + \frac{2h \cdot \mathbf{f}_n \cdot \mathbf{f}_{n+1}}{3(\mathbf{f}_n + \mathbf{f}_{n+1})} + \frac{1}{3}h(\mathbf{f}_n + \mathbf{f}_{n+1}). \tag{11}$$

Висновок про відсутність похибки дискретизації різницевої схеми (11) випливає з формули (10), якщо в ній спрямувати k до безмежності.

Зазначимо, що всі отримані різницеві формули (4)–(9), (11) для дискретизації неперервних систем мають властивість A-стійкості, що унеможливлює накопичення похибки дискретизації при тривалих перехідних процесах, які характерні для динамічних систем з високою добротністю. Цей результат підтверджено розрахунком кварцових генераторних пристроїв та високодобротних генераторних схем з тривалими перехідними процесами [6].

1. *Бондаренко В. М., Герасымив И. И., Мандзий Б. А., Маранов А. В.* Анализ точности и качественного соответствия дискретных моделей электрических цепей. – Киев, 1983. – 44 с. – (Препринт / НАН Украины, Ин-т электродинамики, № 307).
2. *Бутенин Н. В., Неймарк Ю. И., Фуфаев Н. А.* Введение в теорию нелинейных колебаний. – Москва: Наука, 1987. – 384 с.
3. *Васильев В. И., Шевченко А. И.* Комбинированный алгоритм оптимальной сложности // Праці Міжнар. конф. “Штучний інтелект”. – Т. 1. – Крим, 2002. – С. 308–310.

4. *Заяць В. М.* Построение и анализ дискретной модели дискретно-колебательной системы // Кибернетика и системный анализ. – 2000. – № 4. – С. 161–165.
5. *Заяць В. М.* Аналіз динаміки та умов стійкості дискретних моделей коливних систем // Вісн. НУ "Львівська політехніка". Інформаційні системи та мережі. – 2004. – № 519. – С. 132–142.
6. *Заяц В. М.* Ускоренный поиск установившихся режимов в высокочастотных автогенераторах с длительными переходными процессами // Изв. вузов. Радиоэлектроника. – 1993. – № 3. – С. 26–32.
7. *Заяць В. М.* Побудова комбінованих різницевих методів другого порядку // Зб. праць наук.-техн. конф. "Обчислювальні методи і системи перетворення інформації". – Львів, 7–8 жовтня 2011. – ФМІ НАНУ. – 2011. – С. 34–36.
8. *Чуа Л. О., Лин П.-М.* Машинный анализ электронных схем (алгоритмы и вычислительные методы). – Москва: Энергия, 1980. – 640 с.

НУ "Львівська політехніка"
Надійшло до редакції 07.12.2012

В. М. Заяц

Итерационный подход к минимизации погрешности численных методов второго порядка и их применение к анализу нелинейных динамических систем

Предложен итерационный подход к минимизации погрешности дискретизации численных методов второго порядка, основанный на модификации метода трапеций и установлении момента времени, когда поправки явного и неявного методов Эйлера имеют одинаковый вклад в поправку для следующей точки дискретизации динамической системы. Подтверждена целесообразность его применения к анализу нелинейных динамических систем колебательной природы с высокой добротностью и длительными переходными процессами.

V. M. Zayats

An iteration approach to the minimization of errors for second-order numerical methods and their application to the for analysis of nonlinear dinamical systems

An iteration approach to the minimization discretization errors for second-order numerical methods is proposed. It is based on a modification of the method of trapezoids and on setting the time when the contributions of the explicit and implicit Euler methods to the amendment to the next discretization point of a dynamical system are the same. The expediency of its application to the analysis of nonlinear dynamical systems of the oscillatory nature with a high quality factor and long transient processes is confirmed.