Теорія різницевих схем.
Чисельне моделювання процесів гідродинаміки

Оноцький В. В.

12 жовтня 2019 р.

Зміст

3 Точні та узагальнені розв'язки диференціальних та різницевих рівнянь у банахових просторах

Поняття точного та узагальненого розв'язків диференціальних та різницевих рівнянь у банахових просторах. Розв'язуючі оператори. Коректність постановок задач. Збіжність та стійкість. Теорема Лакса. [1, 2]

Нехай B — нормований банаховий простір. Функції, які залежать від просторових змінних x і координати часу t, при фіксованому t будемо тлумачити як точки цього функціонального простору і позначати їх символом u. Стан фізичної системи зображатимемо точкою функціонального простору, а її положення в часі відображатиме рух цієї точки у функціональному просторі B. Через $\tilde{\mathbb{R}}_n$ позначимо n-вимірний комплексний евклідів векторний простір зі скалярним добутком

$$\left(\vec{a}, \vec{b}\right) = \sum_{j=1}^{n} a_j \bar{b}_j$$

і нормою

$$\|\vec{a}\| = \sqrt{\sum_{j=1}^{n} a_j \bar{a}_j}.$$

Введемо також допоміжний нормований простір B', елементами якого є довільні неперервні криві, які при кожному $t \in [0, T]$, визначені як функції просторових змінних $w(t) \in B$. В B' введемо норму так:

$$\|w\|_{B'} = \max_{0 \le t \le T} \|w(t)\|_B.$$

3.1 Скінченно-різницеві задачі в абстрактних просторах Банаха

Розглянемо однопараметричне сімейство елементів $u(t) \in B$ з дійсним параметром t таких, що

$$\frac{\partial u(t)}{\partial t} = Au(t); \quad u(0) = u_0; \quad 0 \le t \le T, \tag{3.1}$$

де u_0 — заданий елемент з B, який характеризує початковий стан системи; A — лінійний оператор.

> **Означення 3.1.** *Точний розв'язок* задачі визначимо як однопараметричне сімейство $u(t)$, кожен елемент якого належить області визначення оператора A, $\forall t \in [0, T]$, $u(0) = u_0$ і
>
> $$\left\| \frac{u(t + \Delta t) - u(t)}{\Delta t} - Au(t) \right\| \to 0, \tag{3.2}$$
>
> при $\Delta t \to 0$, $\forall t \in [0, T]$.

Позначимо через D множину елементів $u_0 \in B$, для кожного з яких існує єдиний розв'язок задачі (3.1) при $u(0) = u_0$, а збіжність в (3.2) рівномірна за t. Нехай $E_0(T)$ відображає D в B і при фіксованому t встановлює відповідність між u_0 і $u(t)$. Тоді $u(t) = E_0(T)u_0$ є розв'язком задачі для тих $u_0 \in B$, для яких існує точний розв'язок.

> **Означення 3.2.** Задачі, визначені лінійним оператором A, назвемо *коректними за Адамаром*, якщо:
>
> 1) область визначення D перетворення $E_0(t)$ щільна в B;
>
> 2) сімейство перетворень $E_0(t)$ рівномірно обмежене, тобто існує така додатна стала $K \in \mathbb{R}$, що $\|E_0(t)\| \le K$ при $0 \le t \le T$.

Перша умова стверджує, що якщо для деякого початкового значення u_0 точний розв'язок не існує, то цей початковий елемент можна апроксимувати як завгодно точно з допомогою тих початкових елементів з D, для яких існує точний розв'язок. З другої умови випливає, що розв'язок задачі неперервно залежить від початкового значення.

3.2 Узагальнені розв'язки диференціальних рівнянь

Теорема 3.3 (теорема про розширення оператора)
Обмежений лінійний оператор T, область визначення якого щільна в B, має єдине лінійне обмежене розширення T', область визначення якого співпадає з B, і таке, що $\|T\| = \|T'\|$.

Означення 3.4. Обмежений лінійний оператор $E_0(t)$ з щільною в B областю визначення має єдине розширення $E(t)$, яке назвемо *узагальненим розв'язуючим оператором.*

Оператор $E(t)$ визначений на всьому просторі B, обмежений за нормою тим же числом K, що і оператор $E_0(t)$.

Означення 3.5. Рівність

$$u(t) = E(t)u_0$$

є *узагальненим розв'язком задачі* для довільного початкового елемента $u_0 \in B$.

Якщо оператор A явним чином залежить від часу, то узагальнений розв'язуючий оператор стає функцією двох змінних. Дійсно, оскільки в момент часу t_0 задано початковий стан u_0, то $u(t) = E(t, t_0)u_0$ і в силу напівгрупової властивості $E(t_2, t_0) = E(t_2, t_1)E(t_1, t_0)$ при $t_0 \leq t_1 \leq t_2$.

Розглянемо неоднорідну крайову задачу

$$\frac{\partial u(t)}{\partial t} - Au(t) = g(t), \quad u(0) = u_0, \tag{3.3}$$

де u_0 і $g(t)$ задані, а $g(t)$ рівномірно або кусково рівномірно неперервна (в нормі простору B) за часом t функція на відрізку $0 \leq t \leq T$. Вважаємо, що оператор A, який визначає коректно поставлену однорідну задачу, замкнений, а області визначення всіх його степенів щільні в B.

Наступні твердження доводяться в [2]:

Твердження 3.6
Якщо

1) u_0 і $g(t) \in D(A)$;

2) $g(t) \in D(A^2)$;

3) функції $Ag(t)$ і $A^2g(t)$ неперервні,

то

$$u(t) = E(t)u_0 + \int_0^t E(t-s)g(s)\,\mathrm{d}s \tag{3.4}$$

є *точним розв'язком задачі* (3.3).

Твердження 3.7

Якщо на u_0 і $g(t)$ не накладено ніяких обмежень, крім неперервності $g(t)$, або ці умови зводяться тільки до умови існування інтегралу в (3.4), то (3.4) є узагальненим розв'язком задачі (3.3).

Твердження 3.8

При вказаних вимогах відносно оператора A і функції $g(t)$ узагальнений розв'язок існує.

Зауваження 3.9 — Єдиність розв'язку випливає з єдиності однорідної задачі, яка за припущенням поставлена коректно.

3.3 Скінченно-різницеві задачі

Розглянемо однорідне скінченно-різницеве рівняння

$$B_1 u^{n+1} = B_0 u^n, \tag{3.5}$$

де $B_0 = B_0(\Delta t, \Delta x_1, \Delta x_2, \ldots)$ і $B_1 = B_1(\Delta t, \Delta x_1, \Delta x_2, \ldots)$ — лінійні скінченно–різницеві оператори, залежні від приростів $\Delta t, \Delta x_1, \Delta x_2, \ldots$ і можливо від просторових змінних. Обидві частини рівняння є лінійними функціями значень u, визначених в точках з деякої множини (шаблону).

Нехай існує обернений оператор B_1^{-1}, $B_1^{-1} B_0$ обмежений і вони визначені на всьому B, а $\Delta x_i = q_i(\Delta t)$, $i = \overline{1, d}$, де d — розмірність простору. Позначимо

$$B_1^{-1}(\Delta t, \Delta x_1, \Delta x_2, \ldots) \cdot B_0(\Delta t, \Delta x_1, \Delta x_2, \ldots) = C(\Delta t)$$

тоді

$$u^{n+1} = C(\Delta t) u^n. \tag{3.6}$$

Означення 3.10. Сімейство операторів $C(\Delta t)$ *узгоджено апроксимує* крайову задачу (3.1), якщо для довільного $u(t)$ з деякого класу U точних розв'язків, початкові елементи яких утворюють в B щільну множину, справедлива *умова узгодження*:

$$\left\| \left(\frac{C(\Delta t) - I}{\Delta t} - A \right) u(t) \right\| \to 0, \tag{3.7}$$

при $\Delta t \to 0$, $0 \le t \le T$.

Тут I — одиничний оператор.

Означення 3.11. Враховуючи (3.2), маємо

$$\left\| \frac{u(t+\Delta t) - C(\Delta t)u(t)}{\Delta t} \right\| \to 0, \tag{3.8}$$

при $\Delta t \to 0$, $0 \leq t \leq T$, в якому вираз під знаком норми є *похибкою апроксимації.*

Якщо $\forall \varepsilon > 0\ \exists \delta > 0$ таке, що

$$\left\| \left(C(\Delta t) - E(\Delta t) \right) u(t) \right\| < \varepsilon \Delta t, \tag{3.9}$$

при $0 \leq t \leq T$, $0 < \Delta t < \delta$, то *збіжність рівномірна* за часом t.

Означення 3.12. Сімейство $C(\Delta t)$ забезпечує *збіжну апроксимацію* задачі, якщо для будь-якого фіксованого $t \in [0, T]$, для кожного $u_0 \in B$ і для кожної збіжної до нуля послідовності додатних приростів $\{\Delta_j t\}_{j=1}^{\infty}$, має місце граничне відношення

$$\| C(\Delta_j t)^{n_j} u_0 - E(t) u_0 \| \to 0 \tag{3.10}$$

при $j \to \infty$, де $n_j \in \mathbb{N}$ такі, що $n_j \Delta_j t \to t$ при $j \to \infty$.

Означення 3.13. Апроксимацію $C(\Delta t)$ назвемо *стійкою*, якщо для деякого $\tau > 0$ нескінченна множина операторів

$$C(\Delta t)^n, \quad 0 \leq n\Delta t \leq T, \quad 0 < \Delta t < \tau \tag{3.11}$$

рівномірно обмежена.

3.4 Теорема Лакса

Теорема 3.14 (Лакса, про еквівалентність)
Нехай задача (3.1)–(3.2) коректно поставлена та її скінченно-різницева апроксимація задовольняє умову узгодження. Тоді стійкість необхідна і достатня для збіжності.

Доведення. **Необхідність.** Спочатку ми покажемо, що збіжна схема необхідно є стійкою. Ми стверджуємо, що для всякої збіжної схеми й для довільного початкового фіксованого елемента $u_0 \in B$ величини $\|C(\Delta t)^n u_0\|$, $(0 < \Delta t < \tau$, $0 \leq n\Delta t \leq T)$, обмежені при деякому $\tau > 0$. Дійсно, якщо це не так, то знайдуться дві послідовності $\Delta_1 t, \Delta_2 t, \ldots, \Delta_j t, \ldots$ і $n_1, n_2, \ldots, n_j, \ldots$, для яких норми елементів $C(\Delta_1)^{n_1} t, C(\Delta_2)^{n_2} t, \ldots, C(\Delta_j)^{n_j} t, \ldots$ необмежено зростають (при цьому $\Delta_j t$ повинні прямувати до нуля в силу припущення про неперервну залежність $C(\Delta t)$ від додатних значень Δt); із цих елементів ми можемо вибрати підпослідовність, для якої величини $n_j \delta_j t$ збігаються до деякого t з відрізка $0 \leq t \leq T$; але це суперечить припущенню про збіжність схеми, оскільки при наявності збіжності норми елементів цієї підпослідовності повинні були б прямувати до кінцевої границі $\|E(t)u_0\|$. Отже, існує така функція $K_1(u)$, що неперервність $\|C(\Delta t)nu_0\| \le K_1(u)$ виконується для всіх операторів з множини (3.11) і всіх $u \in B$, отже множина (3.11) рівномірно обмежена. Таким чином, апроксимація стійка.

Достатність. Щоб довести зворотне твердження, припустимо, що $u(t) = E(t)u_0$ є точним розв'язком, що належать класу U, про який йшла мова при визначенні узгодженості. Нехай ε, δ ті ж, що й в умові узгодженості у формі (3.9), n_j і $\Delta_j t$ обрані так само, як і при визначенні збіжності, а ψ_j позначає різницю між обчисленим і точним значенням і в момент часу $n_j\Delta_j t$, тобто

$$\psi_j = \Big[C(\Delta_j t)^{n_j} - E(n_j\Delta_j t)\Big] = \\ = \sum_{k=0}^{n_j} C(\Delta_j t)^k \Big[C(\Delta_j t) - E(\Delta_j t)\Big] E((n_j - 1 - k)\Delta_j t)u_0. \quad (3.12)$$

Третя частина цієї рівності, співпадає з другою: після приведення подібних залишаються тільки перший і останній члени. Норму величини ψ_j можна оцінити за допомогою (3.9) і нерівності трикутника: якщо $0 < \Delta_j t < \delta$, то

$$\|\psi_j\| \le K_2 \sum_{k=0}^{n_j - 1} \varepsilon\Delta_j t = K_2\varepsilon n_j\Delta_j t \le K_2\varepsilon T, \quad (3.13)$$

де K_2 позначає рівномірну границю множини (3.11). Оскільки ε довільне, то $\|\psi_j\| \to 0$ при $\Delta_j t \to 0$, $n_j\Delta_j t \to t$. Для доведення збіжності покажемо, що в граничному переході при $j \to \infty$ в (3.12) можна замінити $E(n_j\Delta_j t)$ на $E(t)$. Якщо $s = |t - n_j\Delta_j t|$, $t' = \min\{t, n_j\Delta_j t\}$, то в силу напівгрупової властивості сім'ї $E(t)$ маємо $E(s + t') = E(s)E(t)$, тому $E(n_j\Delta_j t) - E(t) = \pm(E(s) - I)E(t')$, причому знак визначається знаком різниці $t - n_j\Delta_j t$. У будь-якому випадку

$$\|(E(n_j\Delta_j t) - E(t))u_0\| \le K_E\|(E(s) - I)u_0\|, \quad (3.14)$$

де K_E позначає границю для $\|E(t)\|$ при $0 \le t \le T$. Але права частина останньої нерівності прямує до нуля, якщо $s \to 0$, тобто якщо $j \to \infty$. Отже, величина

$$\|(E(s) - I)u_0\| \quad (3.15)$$

може бути зроблена як завгодно малою вибором достатньо малих $\Delta_j t$ і $|t - n_j\Delta_j t|$. Це справедливо для довільного u_0, що є початковим елементом точного розв'язку із класу U: але такі елементи щільні в B, тому для довільного $u \in B$ з них можна вибрати послідовність $u_1, u_2, \ldots$ збіжну до u. Тому

$$\Big(C(\Delta_j t)^{n_j} - E(t)\Big)u = \Big(C(\Delta_j t)^{n_j} - E(t)\Big)u_m + C(\Delta_j t)^{n_j}(u - u_m) - E(t)(u - u_m). \quad (3.16)$$

Тут два останні члени в правій частині можуть бути зроблені як завгодно малими за допомогою вибору досить великого m, оскільки клас (3.11) і множина операторів $E(t)$ рівномірно обмежені, а малість першого члена може бути забезпечена за рахунок вибору достатньо малих $\Delta_j t$ і $|t - n_j\Delta_j t|$. Оскільки u — довільний елемент із B, то збіжність встановлена й теорема про еквівалентність доведена. □

Відзначимо в якості природнього наслідку цієї теореми, що для даного початкового елемента u_0 збіжність рівномірна по t на відрізку $[0, T]$ у тому розумінні, що обмеження, які потрібно накладати на $\Delta_j t$ і $|t - n_j \Delta_j t|$, щоб зробити (3.16) нескінченно малим, не залежать ні від вибору t, ні від вибору послідовності $\Delta_j t$. Ця обставина має велике практичне значення, тому що дозволяє при чисельному інтегруванні знаходити такий крок Δt, при якому наближений розв'язок виявляється досить точним на всьому відрізку $[0, T]$ одночасно. Часто для досягнення потрібної точності Δt варіюють у процесі обчислень, але існує таке граничне додатнє значення Δt, нижче якого заходити нема рації.

Викладену теорему Лакса можна дуже просто застосувати до неявного різницевого рівняння для одномірного завдання дифузії

$$\frac{u_j^n - u_j^{n-1}}{\Delta t} = \sigma \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{(\Delta x)^2}. \tag{3.17}$$

Насамперед покажемо, що розв'язок цього рівняння задовольняє наступному принципу максимуму. Припустимо, що рівняння розглядається в прямокутнику $0 \leq x \leq a$, $0 \leq t \leq T$ і що Δx вибирається рівним a/J, де J — натуральне число. Тоді максимальне значення M_1, що досягається величиною u_j^n усередині цього прямокутника, не може перевищувати максимального значення M_2, що досягається початковими й граничними значеннями (тобто значеннями на відрізках прямих $t = 0$, $x = 0$, $x = a$). Дійсно, припустимо протилежне, а саме що $M_1 > M_2$, і нехай (n, j) — перша внутрішня точка сітки, у якій $u_j^n = M_1$ (перша у тому розумінні, що індекси n і j мають найменші значення). Тоді в цій точці сітки ліва частина наведеного вище рівняння повинна бути додатна, а права від'ємна, тому що u_j^n за припущенням перевершує сусіднє значення u_{j-1}^n ліворуч і сусіднє значення u_j^{n-1} знизу й щонайменше дорівнює сусідньому значенню u_{j+1}^n праворуч. Отже, наше припущення неправильне, і принцип максимуму встановлений. Очевидно, що ці міркування можна застосувати й до $-u_j^n$ і тим самим установити, що $|u_j^n|$ обмежені при будь-якому виборі сітки, тобто що розглянуте різницеве рівняння стійке.

Система різницевих рівнянь для задачі (3.3) має вигляд

$$B_1 u^{n+1} - B_0 u^n = g^{n+1}, \quad u^0 = u_0, \tag{3.18}$$

де B_1 і B_0 — розглянуті раніше різницеві оператори, а g^n апроксимує $g(n\Delta t)$. Якщо величину $\|g^n - g(n\Delta t)\|$ можна зробити як завгодно малою рівномірно за n при $0 \leq n\Delta t \leq T$ за допомогою вибору Δt, то останнє різницеве рівняння можна подати у вигляді

$$u^{n+1} - C(\Delta t)u^n = D(\Delta t)g^{n+1}, \tag{3.19}$$

де $C(\Delta t) = B_1^{-1}B_0$ і $D(\Delta t) = B_1^{-1}$ обмежені і залежать тільки від Δt і, можливо, координат, а Δx_i можуть бути виражені через Δt.

З (3.19) випливає

$$u^n = C(\Delta t)^n u_0 + \Delta t \sum_{j=1}^{n} C(\Delta t)^{n-j} g^j. \tag{3.20}$$

Якщо виконуються умови теореми Лакса, то остання сума апроксимує інтеграл з (3.4), звідки випливає збіжність наближеного розв'язку до точного.

Література

[1] Грищенко and Ляшко, *Методи Фур'є та першого диференціального наближення в теорії різницевих схем.* ВПЦ Київський університет, 2005.

[2] Рихтмайер and Мортон, *Разностные методы решения краевых задач.* Мир, 1972.