Київський національний університет
імені Тараса Шевченка

ЧИСЕЛЬНИЙ АНАЛІЗ

(курс лекцій для студентів факультету комп’ютерних наук та кібернетики)

RAU

Київ 2023

УДК 528.11

ББК 22.19я73

Автор

Риженко А.І.

Чисельний аналіз (курс лекцій для студентів факультету комп’ютерних наук та кібернетики)
Риженко А.І. – Київ, 2023. - 82с.

RAI

Чисельне диференціювання

Побудова формул чисельного диференціювання

Задача чисельного диференціювання виникає у випадку коли необхідно обчислити похідну функції, значення якої задані таблицею. Нехай задано

$$f_i = f(x_i), \quad i = \overline{0,n}, \quad x_i \in [a,b].$$

Побудуємо інтерполяційний поліном для функції $f(x)$. Тоді

$$f(x) = L_n(x) + r_n(x), \tag{1}$$

де залишковий член запишемо у формі Ньютона і він буде мати вигляд:

$$r_n(x) = f(x, x_0, ..., x_n)\omega_n(x), \;\; \omega_n(x) = \prod_{i=0}^{n}(x - x_i).$$

Звідси

$$f^{(k)}(x) = L_n^{(k)}(x) + r_n^{(k)}(x), \; 0 \le k \le n. \tag{2}$$

За наближене значення похідної в точці x візьмемо

$$f^{(k)}(x) \approx L_n^{(k)}(x), x \in [a,b].$$

Оцінимо похибку наближення формули (2), тобто $r_n^{(k)}(x)$. За формулою Лейбніца:

$$r_n^{(k)}(x) = \sum_{j=0}^{k} C_k^j f^{(j)}(x, x_0, ..., x_n)\omega_n^{(k-j)}(x). \tag{3}$$

З властивості розділених різниць маємо для $f(x) \in C^{n+k+1}[a,b]$:

$$f^{(j)}(x, x_0, ..., x_n) = j!\, f(\underbrace{x, ... x}_{j+1}, x_0, \, ..., x_n) = \frac{j!}{(n+j+1)!} f^{(n+j+1)}(\xi_j), \; \xi_j \in [a,b].$$

Дійсно, маємо

$$f(x, x+\varepsilon, ..., x+j\varepsilon) = \frac{f^{(j)}(x_\varepsilon)}{j!},$$

де $x \le x_\varepsilon \le x + j\varepsilon$. Тоді

$$f^{(j)}(x) = j! \lim_{\varepsilon \to 0} f(x, x+\varepsilon, ..., x+j\varepsilon).$$

Звідси за визначенням розділених різниць за кратними вузлами будемо мати

$$f^{(j)}(x, x_0, \, ..., x_n) = j! \lim_{\varepsilon \to 0} f(x, x+\varepsilon, ..., x+j\varepsilon, x_0, \, ..., x_n) = j!\, f(\underbrace{x, ..., x}_{j+1}, x_0, ..., x_n).$$

Тоді остаточно вираз для похибки наближення похідної набуває вигляду

$$r_n^{(k)}(x) = \sum_{j=0}^{k} \frac{k!}{(k-j)!(n+j+1)!} f^{(n+j+1)}(\xi_j)\omega_n^{(k-j)}(x), \tag{4}$$

а оцінка похибки матиме вигляд

$$\left|f^{(k)}(x)-L_n^{(k)}(x)\right| \le M\sum_{j=0}^{k}\frac{k!}{(k-j)!(n+j+1)!}\left|\omega_n^{(k-j)}(x)\right|,$$

де $M=\max\limits_{0\le j\le k}\max\limits_{x\in[a,b]}\left|f^{(n+j+1)}(x)\right|$.

Зауважимо, що процес інтерполювання розбіжний. Окрім того, якщо $k>n$, то $L_n^{(k)}\equiv 0$. Тому не можна брати великими значення n та k. Як правило $k=1,2$. Відповідно, $n=k$, або $n=k+1$, або $n=k+2$.

Розглянемо, яким чином порядок збіжності процесу чисельного диференціювання залежить від кроку. Нехай $x_i=x_0+ih,\quad h>0-$ крок. Тоді за умови $x_i-x_0=O(h)$ будемо мати

$$\omega_n(x)=(x-x_0)\dots(x-x_n)=O\left(h^{n+1}\right),\quad x\in[x_0,x_n].$$

Перша похідна від $\omega_n(x)$ має порядок на одиницю менше, тобто

$$\omega_n'(x)=O\left(h^{n}\right).$$

Тоді

$$r_n^{(k)}(x)=O\left(h^{n+1-k}\right)\ \text{ і тому }\ f^{(k)}(x)-L_n^{(k)}(x)=O\left(h^{n+1-k}\right).$$

При умові $n\ge k$ останній вираз збігається до нуля, тобто

$$f^{(k)}(x)-L_n^{(k)}(x)\underset{h\to 0}{\to}0.$$

Формулу (3) можемо записати у вигляді

$$r_n^{(k)}(x)=\underbrace{f(x,x_0,\dots,x_n)\omega_n^{(k)}(x)}_{O(h^{n+1-k})}+\underbrace{\sum_{j=1}^{k}C_k^j f^{(j)}(x,x_0,\dots,x_n)\,\omega_n^{(k-j)}(x)}_{O(h^{n+2-k})},$$

або ж аналогічно з формули (4)

$$r_n^{(k)}(x)=\underbrace{\frac{1}{(n+1)!}f^{(n+1)}(\xi_0)\omega_n^{(k)}(x)}_{O(h^{n+1-k})}+\underbrace{\sum_{j=1}^{k}\frac{k!}{(k-j)!(n+j+1)!}f^{(n+j+1)}(\xi_j)\omega_n^{(k-j)}(x)}_{O(h^{n+2-k})}. \quad (5)$$

Якщо для деякої точки $\bar{x}$

$$\omega_n^{(k)}(\bar{x})=0,\ \text{то}\ r_n^{(k)}(\bar{x})=O\left(h^{n+2-k}\right). \quad (6)$$

Такі точки $\bar{x}$ називають точками підвищеної точності формул чисельного диференціювання.

<u>Приклад 1</u>. Побудувати формулу чисельного диференціювання для першої похідної по двох точках. Дати оцінку точності. Визначити точки підвищеної точності для цієї формули ?

Розв'язання. З умов задачі маємо, що $k=1, n=1$. Виберемо точки x_0, $x_1=x_0+h$, тоді інтерполяційний поліном має вигляд

$$L_1(x) = f_0 + (x - x_0)\frac{f_1 - f_0}{h}.$$

Для похідної отримаємо вираз

$$f'(x) \approx L_1'(x) = \frac{f_1 - f_0}{h},\ x \in [x_0, x_1].$$

З формули (5) отримаємо вираз для похибки

$$r_1'(x) = \frac{f^{(2)}(\xi_0)}{2!}(2x - x_1 - x_0) + \frac{f^{(3)}(\xi_1)}{3!}(x - x_0)(x - x_1) = O(h). \qquad (7)$$

Якщо $2x - x_1 - x_0 = 0$, то $\bar{x} = \frac{x_1 + x_0}{2}$ – точка підвищеної точності і з формули (7) будемо мати

$$\left| r_1'(\bar{x}) \right| \le \frac{h^2}{24} M_3, \text{ де } M_3 = \max_{x \in [x_0, x_1]} \left| f^{(3)}(x) \right|.$$

<u>Зауваження.</u> Виходячи з результатів прикладу 1 введемо наступні позначення:

при $x \in [x_i, x_{i+1}]$ позначимо $f_{x,i} = \frac{f_{i+1} - f_i}{h}$ (різницева похідна вперед, або права різницева похідна);

при $x \in [x_{i-1}, x_i]$ позначимо $f_{\bar{x},i} = \frac{f_i - f_{i-1}}{h}$ (різницева похідна назад, або ліва різницева похідна).

Будь яким з цих виразів можна наблизити $f'(x)$ принаймі з першим порядком точності.

<u>Приклад 2</u>. Побудувати формулу чисельного диференціювання для першої похідної по трьох рівновіддалених точках. Дати оцінку точності.

Розв'язання. З умов задачі маємо, що $k = 1, n = 2$. Виберемо точки $x_0,\ x_1 = x_0 + h,\ x_2 = x_0 + 2h$. Інтерполяційний поліном має вигляд

$$L_2(x) = f_0 + (x - x_0)\frac{f_1 - f_0}{h} + (x - x_0)(x - x_1)\frac{f_2 - 2f_1 + f_0}{2h^2}.$$

Тоді наблизимо

$f'(x) \approx L_2'(x) = \frac{f_1 - f_0}{h} + (2x - x_0 - x_1)\frac{f_2 - 2f_1 + f_0}{2h^2}$, $x \in [x_0, x_2]$. Якщо в цей вираз підставити $x = x_0$, то отримаємо $f'(x_0) \approx \frac{-f_2 + 4f_1 - 3f_0}{2h}$. Для точки $x = x_1$ маємо $f'(x_1) \approx \frac{f_2 - f_0}{2h} = f_{\underset{x,1}{0}}$, а для точки $x = x_2$:

$f'(x_2) \approx \frac{f_0 - 4f_1 + 3f_2}{2h}$, $x \in [x_0, x_2]$. З формули (5), для похибки цих формул отримаємо оцінку $r_2'(x) = O(h^2)$.

Зауваження. Введемо також наступне позначення

при $x \in [x_{i-1}, x_{i+1}]$ позначимо $f_{\substack{0\\x,i}} = \dfrac{f_{i+1} - f_{i-1}}{2h}$ (центральна різницева похідна).

Задача. Знайти точки підвищеної точності формули чисельного диференціювання для $k = 1, n = 2$ та оцінити похибку в цих точках.

Приклад 3. Для $n = 1,\ k = 1$ оцінити точність формул чисельного диференціювання методом розкладу в ряд Тейлора.
Розв'язання. В прикладі 1 була отримана формула наближеного обчислення першої похідної $f'(x) \approx L_1'(x) = \dfrac{f_1 - f_0}{h}$, $x \in [x_0, x_1]$. Тоді, позначивши $s = \dfrac{x - x_0}{h}$, $0 \le s \le 1$, матимемо

$$\begin{aligned}
&f'(x) - \frac{f(x_0 + h) - f(x_0)}{h} = f'(x) - \frac{1}{h}\left[f_0 + hf_0' + \frac{h^2}{2} f_0'' + \frac{h^3}{3!} f'''(\xi) - f_0 \right] = \\
&= [x - x_0 = sh] = f'(x_0 + sh) - f_0' - \frac{h}{2} f_0'' - \frac{h^2}{6} f'''(\xi) = \\
&= f_0' + shf_0'' + \frac{(sh)^2}{2} f'''(\xi_1) - f_0' - \frac{h}{2} f_0'' - \frac{h^2}{6} f'''(\xi) = \\
&= \left(s - \frac{1}{2} \right) h f_0'', + O(h^2),\ \xi, \xi_1 \in [x_0, x_1].
\end{aligned} \tag{8}$$

Отже

$$\left| f'(x) - \frac{f_1 - f_0}{h} \right| \le \left| s - \frac{1}{2} \right| hM_2, \text{ де } M_2 = \max_{[x_0, x_1]} |f''(x)|.$$

Тобто при $s \ne 0.5$ будемо мати $\left| f'(x) - \dfrac{f_1 - f_0}{h} \right| = O(h)$, зокрема при $s = 0$ отримаємо $\left| f'(x_0) - \dfrac{f_1 - f_0}{h} \right| \le \dfrac{1}{2} hM_2$ і при $s = 1$ буде $\left| f'(x_1) - \dfrac{f_1 - f_0}{h} \right| \le \dfrac{1}{2} hM_2$.

При $s = \dfrac{1}{2}$ додавши до розкладу в ряд Тейлора ще один доданок, отримаємо оцінку

$$\left| f'(\bar{x}) - \frac{f_1 - f_0}{h} \right| \le \frac{h^2 M_3}{24},\ \bar{x} = \frac{x_1 + x_0}{2} = x_{1/2},\ M_3 = \max_{[x_0, x_1]} |f'''(x)|. \ \square$$

Пояснення.

$$f'(x)-\frac{f(x_0+h)-f(x_0)}{h}=$$

$$=f'(x)-\frac{1}{h}\left[f_0+hf_0'+\frac{h^2}{2}f_0''+\frac{h^3}{3!}f_0^{(3)}+\frac{h^4}{4!}f^{(4)}(\xi)-f_0\right]=$$

$$=[x-x_0=sh]=f'(x_0+sh)-f_0'-\frac{h}{2}f_0''-\frac{h^2}{3!}f_0^{(3)}-\frac{h^3}{4!}f^{(4)}(\xi)=$$

$$=f_0'+shf_0''+\frac{(sh)^2}{2}f_0^{(3)}+\frac{(sh)^3}{6}f^{(4)}(\xi_1)-f_0'-\frac{h}{2}f_0''-\frac{h^2}{3!}f_0^{(3)}-\frac{h^3}{4!}f^{(4)}(\xi)=$$

$$=h(s-\frac{1}{2})f_0''+\left(\frac{(sh)^2}{2}-\frac{h^2}{3!}\right)f_0^{(3)}+\frac{(sh)^3}{6}f^{(4)}(\xi_1)-\frac{h^3}{4!}f^{(4)}(\xi),\ \xi,\xi_1\in[x_0,x_1],$$

що при $s=\frac{1}{2}$ дає оцінку

$$\left|f'(\bar{x})-\frac{f_1-f_0}{h}\right|\le\frac{h^2M_3}{24},\ \bar{x}=\frac{x_1+x_0}{2}=x_{1/2},\ M_3=\max_{[x_0,x_1]}|f'''(x)|.\ \square$$

Зауваження. За допомогою безпосереднього розкладу в ряд Тейлора для «фіксованих» точок, будемо мати

а) для $x=x_0$:

$$f'(x_0)-\frac{f(x_0+h)-f(x_0)}{h}=$$

$$f'(x_0)-\frac{1}{h}\left(f_0+hf_0'+\frac{h^2}{2}f''(\xi)-f_0\right)=-\frac{h}{2}f''(\xi),\ \xi\in[x_0,x_1].$$

Тоді

$$\left|f'(x_0)-\frac{f_1-f_0}{h}\right|\le\frac{M_2h}{2},\ M_2=\max_{[x_0,x_1]}|f''(x)|.$$

б) для $x=x_1$:

$$f'(x_1)-\frac{f(x_0+h)-f(x_0)}{h}=f'(x_1)-\frac{f(x_1)-f(x_1-h)}{h}=$$

$$=f'(x_1)-\frac{1}{h}\left(f_1-f_1+hf_1'-\frac{h^2}{2}f''(\xi)\right)=\frac{h}{2}f''(\xi),\ \xi\in[x_0,x_1].$$

Отже

$$\left|f'(x_1)-\frac{f_1-f_0}{h}\right|\le\frac{M_2h}{2},\ M_2=\max_{[x_0,x_1]}|f''(x)|.$$

в) для $x=\bar{x}=\dfrac{x_1+x_0}{2}=x_{1/2}$:

$$f'(\bar{x})-\frac{f(x_0+h)-f(x_0)}{h}=$$

$$=f'(\bar{x})-\frac{1}{h}[f(\bar{x})+\frac{h}{2}f'(\bar{x})+\frac{h^2}{8}f''(\bar{x})+\frac{h^3}{48}f'''(\bar{x})+\frac{h^4}{4!\cdot 2^4}f^{(4)}(\xi)-$$

$$-f(\bar{x})+\frac{h}{2}f'(\bar{x})-\frac{h^2}{8}f''(\bar{x})+\frac{h^3}{48}f'''(\bar{x})-\frac{h^4}{4!\cdot 2^4}f^{(4)}(\eta)]=$$

$$=-\frac{h^2}{24}f'''(\bar{x})+O(h^4),\ \xi,\eta\in[x_0,x_1].$$

Отже $\left|f'(\bar{x})-\dfrac{f_1-f_0}{h}\right|\le\dfrac{h^2M_3}{24},\bar{x}=\dfrac{x_1+x_0}{2},\ M_3=\max\limits_{[x_0,x_1]}|f'''(x)|$. □

В даному випадку розклад в ряд Тейлора виконувався відносно точки, в якій досліджувалася похибка.

<u>Задача.</u> Показати, що для $f(x)\in C^3[a,b]$ виконується оцінка

$$\left|f'(x_1)-\frac{f_2-f_0}{2h}\right|\le\frac{M_3h^2}{6}.$$

<u>Приклад 4.</u> Побудувати формулу чисельного диференціювання для другої похідної за трьома рівновіддаленими точками. Дати оцінку точності. Визначити точки підвищеної точності для цієї формули.

Розв'язання. З умов задачі маємо, що $k=2,n=2$. Виберемо точки $x_0,\ x_1=x_0+h,\ x_2=x_0+2h$. Інтерполяційний поліном другого степеня за рівновіддаленими точками має вигляд:

$$L_2(x)=f_0+(x-x_0)\frac{f_1-f_0}{h}+(x-x_0)(x-x_1)\frac{f_2-2f_1+f_0}{2h^2}.$$

Тоді наблизимо

$$f''(x)\approx L_2''(x)=\frac{f_2-2f_1+f_0}{h^2},\quad x\in[x_0,x_2].$$

Залишковий член відповідно до формули (5) буде мати вигляд

$$r_2^{(2)}(x)=\frac{1}{(2+1)!}f^{(2+1)}(\xi_0)\omega_2''(x)+\frac{2!}{1!(2+1+1)!}f^{(2+1+1)}(\xi_1)\omega_2'(x)+$$

$$+\frac{2!}{0!(2+2+1)!}f^{(2+2+1)}(\xi_2)\omega_2(x)=\frac{1}{3!}f^{(3)}(\xi_0)\omega_2''(x)+\frac{1}{12}f^{(4)}(\xi_1)\omega_2'(x)+$$

$$+\frac{1}{60}f^{(5)}(\xi_2)\omega_2(x).$$

Для $\forall\, x\in[x_0,x_2]$ маємо $r_2^{(2)}(x)=O\left(h^{n+1-k}\right)=O(h)$.

Знайдемо точки підвищеної точності формули чисельного диференціювання.

Маємо

$$\omega_2''(x) = (x - x_1) + (x - x_2) + (x - x_2) + (x - x_0) + (x - x_0) + (x - x_1) =$$
$$= 2(x - x_1 + x - x_0 + x - x_2) = 2(3x - x_0 - x_1 - x_2) = 0,$$

тоді $\bar{x} = \frac{1}{3}(x_0 + x_1 + x_2) = x_1$ - точка підвищеної точності формули чисельного диференціювання. Для залишкового члена будемо мати

$$r_2^{(2)}(x_1) = \frac{1}{12} f^{(4)}(\xi_1)\omega_2'(x_1) = -\frac{h^2}{12} f^{(4)}(\xi_1),$$
$$\left| r_2^{(2)}(x_1) \right| \le \frac{h^2}{12} M_4,$$

де $M_4 = \max\limits_{[x_0, x_2]} \left| f^{(4)}(x) \right|$, $\omega_2'(x_1) = (x - x_0)(x - x_2)\big|_{x=x_1} = -h^2$. □

<u>Задача.</u> Побудувати формулу чисельного диференціювання для $k = 2, n = 2$ у випадку нерівновіддалених вузлів: x_0, $x_1 = x_0 + h_1$, $x_2 = x_1 + h_2$. Оцінити точність формули. Знайти точки підвищеної точності.

<u>Зауваження.</u> Крім інтерполяційних формул для чисельного диференціювання можна застосовувати сплайни. Нехай $f_i = f(x_i)$. Побудуємо інтерполяційний сплайн першого степеня $s_1(x)$, для якого має місце оцінка $\left| f^{(k)}(x) - s_1^{(k)}(x) \right| = O\left(h^{2-k}\right), k = 0,1$. Звідси при $k = 1$ маємо $\left| f'(x) - s_1'(x) \right| = O(h)$.

Для кубічного інтерполяційного сплайну $s_3(x)$ маємо для першої та другої похідних

$$\left| f^{(k)}(x) - s_3^{(k)}(x) \right| = O\left(h^{4-k}\right), \quad k = 1,2.$$

Метод невизначених коефіцієнтів

Нехай потрібно знайти k - ту похідну функції f в точці x_j $(j = 0,1,2,...,n)$ використовуючи її значення $f_0,...,f_n$ в точках $x_0,...,x_n$. Розглянемо ще один спосіб побудови формул чисельногодиференціювання - метод невизначених коефіцієнтів. Він полягає в наступному: шукану похідну записуємо у вигляді

$$f^{(k)}(x_j)=\sum_{i=0}^{n}c_i f(x_i)+R(f),\ 0\le j\le n, \tag{9}$$

де $R(f)$ - залишковий член формули чисельного диференціювання. Будемо вимагати, щоб дана формула була точною для поліномів якомога вищого степеня. Тобто коефіцієнти c_i визначаємо із системи лінійних алгебраїчних рівнянь

$$R(x^{\alpha})=0,\ \alpha=0,1,2,...,n, \tag{10}$$

де

$$R(x^{\alpha})=\left(x^{\alpha}\right)^{(k)}\Big|_{x=x_j}-\sum_{i=0}^{n}c_i\,x_i^{\alpha}\ .$$

Зауваження. Для полегшення обчислень замість виразів x^{α}, $\alpha=0,1,2,...,n$ в (10), можна взяти, наприклад, поліноми

$$f=1,\ f=x-x_m,\ f=(x-x_m)^2,...,f=(x-x_m)^n, \tag{11}$$

де x_m одна з точок $x_0,x_1,...,x_n$. □

<u>Приклад</u>. Знайти $\dfrac{df}{dx}$ в точці $x=x_1$ за значеннями f_0,f_1,f_2,f_3 у рівновіддалених вузлах x_0,x_0+h,x_0+2h,x_0+3h.

<u>Розв'язання</u>. Маємо

$$f_1'\approx c_0f_0+c_1f_1+c_2f_2+c_3f_3. \tag{14}$$

Покладаючи послідовно $f(x)=1,\ f(x)=x-x_0,\ f(x)=(x-x_0)^2,\ f(x)=(x-x_0)^3$ та підставляючи в співвідношення $f'(x_1)=\sum_{i=0}^{3}c_i f(x_i)$ (тобто $R(f)=0$), отримаємо СЛАР

$$\begin{aligned}0&=c_0+c_1+c_2+c_3,\\ 1&=c_1h+c_2 2h+c_3 3h,\\ 2h&=c_1h^2+c_2(2h)^2+c_3(3h)^2,\\ 3h^2&=c_1h^3+c_2(2h)^3+c_3(3h)^3.\end{aligned} \tag{15}$$

Знайдемо розв'язок СЛАР (15). Від третього рівняння віднімемо друге рівняння помножене на h, отримаємо

$$h=2c_2h^2+6c_3h^2. \tag{16}$$

Тепер від четвертого рівняння віднімемо друге рівняння помножене на h^2, тоді будемо мати

$$2h^2=6c_2h^3+24c_3h^3. \tag{17}$$

Запишемо (16), (17) у вигляді

$$2c_2h + 6c_3h = 1,$$
$$6c_2h + 24c_3h = 2,$$

звідки $c_3 = -\frac{1}{6h}$, $c_2 = \frac{1}{h}$. Тоді з другого рівняння (15) отримаємо

$c_1 = \frac{1}{h}(1 + 2c_2h + 3c_3h) = -\frac{1}{2h}$, а з першого рівняння (15) $c_0 = -\frac{1}{3h}$.

Підставляючи отримані коефіцієнти в (14), отримаємо

$$\left.\frac{df}{dx}\right|_{x=x_1} \approx \frac{1}{6h}(-2f_0 - 3f_1 + 6f_2 - f_3).$$

Обчислювальна похибка чисельного диференціювання

В задачах чисельного диференціювання часто зменшення похибки формули чисельного диференціювання супроводжується ростом впливу похибки вхідних даних та обчислювальної похибки.

В формулах чисельного диференціювання з постійним кроком h значення функції f діляться на h^m, де m – порядок шуканої похідної. Тому при малому h неусувні похибки в значеннях функції f значно впливають на результат чисельного диференціювання. Таким чином, виникає задача вибору оптимального кроку h, оскільки похибка методу прямує до нуля при $h \to 0$, а неусувна похибка зростає.

Зробимо зауваження про вплив збурення функції на значення похідних.

Нехай $f(x) \in C^1[a,b]$, а збурена функція має вигляд

$$\tilde{f}(x) = f(x) + \frac{1}{n}\sin(\omega x).$$

При $n \to \infty$ маємо $\|\tilde{f}(x) - f(x)\|_{C[a,b]} = \frac{1}{n} \to 0$ звідси $\tilde{f}(x) \underset{n\to\infty}{\Rightarrow} f(x)$.

Для похідної функції маємо

$$\tilde{f}'(x) = f'(x) + \frac{\omega}{n}\cos(\omega x).$$

Нехай $\omega = n^2$, тоді

$$\|\tilde{f}' - f'\|_{C[a,b]} = \frac{|\omega|}{n} = n \xrightarrow[n\to\infty]{} \infty.$$

Отже, похибка диференціювання збуреної функції може бути як завгодно великою. Це є наслідком некоректності (нестійкості) оператора диференціювання. Тому є очікування (негативні), що ця нестійкість має місце і для чисельного диференціювання.

Припустимо, що значення функції обчисленні з деякою похибкою. З'ясуємо, як ці похибки впливають на значення похідних, обчислених за формулами чисельного диференціювання.

Нехай $\tilde{f}_i = f_i + \delta_i, \quad f_i = f(x_i), i = \overline{0,n}$ $|\delta_i| \le \delta$. Розглянемо вплив похибок δ_i на конкретних формулах чисельного диференціювання.

<u>Приклад 1</u>. Оцінити вплив збурень на похибку наближеного обчислення першої похідної при $n = 1, k = 1$ за формулою лівої різницевої похідної.

Розв'язання. Маємо

$$f_i' - \frac{\tilde{f}_i - \tilde{f}_{i-1}}{h} = f_i' - \frac{f_i - f_{i-1}}{h} - \frac{\delta_i - \delta_{i-1}}{h},$$

$$\left| f_i' - \frac{\tilde{f}_i - \tilde{f}_{i-1}}{h} \right| \le \left| f_i' - \frac{f_i - f_{i-1}}{h} \right| + \left| \frac{\delta_i - \delta_{i-1}}{h} \right| \le \frac{M_2 h}{2} + \frac{2\delta}{h} \xrightarrow[h \to 0]{} \infty, \qquad (1)$$

де $M_2 = \max\limits_{[x_{i-1}, x_i]} |f''(x)|$.

Таким чином, як і для аналітичного диференціювання, маємо некоректність: при малих збуреннях $|\delta_i| \le \delta$ можуть бути як завгодно великі похибки, якщо $\frac{\delta}{h} \to \infty$ при $h \to 0$.

Мінімізуємо вплив цих збурень. Позначимо

$$\varphi(h) = \frac{M_2 h}{2} + \frac{2\delta}{h}.$$

Мінімум цієї функції досягається для наступного значення h:

$$\varphi'(h) = \frac{M_2}{2} - \frac{2\delta}{h^2} = 0 \Rightarrow h_0 = 2\sqrt{\frac{\delta}{M_2}}.$$

При такому значенні h оцінка похибка (1) буде наступною

$$\varphi(h_0) = 2\sqrt{M_2 \delta} = O\left(\delta^{\frac{1}{2}}\right) \underset{\delta \to 0}{\to} 0.$$

<u>Приклад 2.</u> Оцінити вплив збурень на похибку наближеного обчислення першої похідної при $n = 2, k = 1$ за формулою центральної різницевої похідної.

Розв'язання. Маємо

$$f_i' - \frac{\tilde{f}_{i+1} - \tilde{f}_{i-1}}{2h} = f_i' - \frac{f_{i+1} - f_{i-1}}{2h} - \frac{\delta_{i+1} - \delta_{i-1}}{2h}$$

$$\left| f_i' - \frac{\tilde{f}_{i+1} - \tilde{f}_{i-1}}{2h} \right| \le \left| f_i' - \frac{f_{i+1} - f_{i-1}}{2h} \right| + \left| \frac{\delta_{i+1} - \delta_{i-1}}{2h} \right| \le \frac{M_3 h^2}{6} + \frac{\delta}{h} = \varphi(h).$$

З рівняння $\varphi'(h) = \frac{M_3 h}{3} - \frac{\delta}{h^2} = 0$ отримаємо $h_0^3 = \frac{3\delta}{M_3}$, $h_0 = \sqrt[3]{\frac{3\delta}{M_3}}$. Отже

$$\varphi(h_0) = \frac{M_3}{6}\sqrt[3]{\frac{9\delta^2}{M_3^2}} + \frac{\delta}{\sqrt[3]{\frac{3\delta}{M_3}}} = \frac{1}{2}\sqrt[3]{\frac{M_3\delta^2}{3}} + \sqrt[3]{\frac{M_3\delta^2}{3}} = \frac{3}{2}\sqrt[3]{\frac{M_3\delta^2}{3}} = O\left(\delta^{\frac{2}{3}}\right),$$

де $M_3 = \max\limits_{[x_{i-1}, x_{i+1}]} \left| f^{(3)}(x) \right|$.

Таким чином швидкість збіжності при $\delta \to 0$ похибки формули чисельного диференціювання центральною похідною вища ніж для формули з прикладу 1 (похідна вперед або назад).

<u>Задача.</u> Дослідити похибку чисельного диференціювання для $n = 2, k = 2$, вибрати оптимальний крок h_0, дати оцінку $\varphi(h_0)$.□

У випадку застосування формул вищого порядку точності становище дещо поліпшується. Нехай похідна $f^{(k)}(x)$ обчислюється за формулою

$$f^{(k)}(x) \approx \frac{\sum\limits_{i=1}^{n} c_i f_i}{h^k}, \; c_i = O(1),$$

з залишковим членом $O(h^l)$.

Тоді повна похибка обчислюється за формулою

$$|R| \le g(h) = A_1 h^l + \frac{A_2 \delta}{h^k}.$$

Мінімум правої частини досягається при h порядку $\delta^{\frac{1}{l+k}}$, при цьому

$$g(h) \sim O\left(\delta^{\frac{l}{l+k}}\right).$$

Таким чином із зростанням l порядок похибки по відношенню до δ збільшується; при цьому значення кроку стає все більшим. Однак при цьому можуть зростати величини A_1 і A_2.

ЧИСЕЛЬНЕ ІНТЕГРУВАННЯ

1. Постановка задачі чисельного інтегрування

Розглянемо задачу знаходження визначеного інтегралу

$$I(f)=\int_a^b \rho(x)f(x)dx \tag{1.1}$$

де f – задана функція, $\rho(x)\geq 0$ – деякий ваговий множник ($\rho(x)=0$ на множині міри 0) . Поставлена задача часто вимагає чисельного розв'язання, оскільки значна кількість інтегралів типу (1.1) не може бути обчислена аналітично, або ж інформація про функцію f може бути задана у вигляді таблиці.

Нагадаємо, що за визначенням

$$I(f)=\lim_{\Delta\to 0}\sum_{i=1}^{n}\rho(\xi_i)f(\xi_i)\Delta x_i\ ,$$

де $\Delta x_i = x_i - x_{i-1}$, а $\{x_i\}_{i=0}^{n}$ – розбиття проміжку $[a,b]$, $x_i \in [a,b]$, $\xi_i \in [x_{i-1}, x_i]$. Тому наблизимо інтеграл скінченою сумою

$$I_n(f)=\sum_{k=0}^{n}c_k f(x_k), \tag{1.2}$$

яка називається *квадратурною сумою*. Тоді наближена рівність

$$\int_a^b \rho(x)f(x)dx \approx \sum_{k=0}^{n}c_k f(x_k) \tag{1.3}$$

називається *квадратурною формулою*. Точки x_k називаються вузлами квадратурної формули, а числа c_k – коефіцієнтами (ваговими множниками) квадратурної формули. Різниця

$$R_n(f)=I(f)-I_n(f) \tag{1.4}$$

називається похибкою квадратурної формули.

Задача чисельного інтегрування полягає в виборі таких $\{x_k, c_k\}_{k=0}^{n}$, щоб похибка була найменша: $R_n(f)=I(f)-I_n(f)\to \min$.

Визначення. Квадратурну формулу (1.3) називають квадратурною формулою замкненого типу, якщо $x_0 = a$ та $x_n = b$, і відкритого типу в протилежному випадку. □

Визначення. Квадратурна формула (1.3) має m-ий *степінь алгебраїчної точності*, якщо

$$R_n(f) = 0, \forall f \in \pi_m, \tag{1.5}$$

(π_m – множина поліномів m-го степеня) і $\exists P_{m+1}(x) \in \pi_{m+1}$, такий що $R_n(P_{m+1}) \neq 0$.□

Умову (1.5) можна замінити умовою

$$R_n(x^\alpha) = 0, \alpha = \overline{0,m}, R_n(x^{m+1}) \neq 0 \tag{1.6}$$

(вона більш зручна для перевірки).

2. Квадратурні формули прямокутників, трапецій, Сімпсона

Припустимо, що $\rho \equiv 1$. Тоді замінивши в інтегралі (1.1) функцію $f(x)$ її значенням в деякій точці x_0 можна побудувати наступні квадратурні формули при $n = 0$:

а) *лівих прямокутників:* $x_0 = a$: $I(f) = \int_a^b f(x)dx \approx I_0^{лп}(f) = (b-a)f(a)$;

б) *правих прямокутників:* $x_0 = b$: $I(f) = \int_a^b f(x)dx \approx I_0^{пп}(f) = (b-a)f(b)$;

в) *середніх прямокутників:* $x_0 = \dfrac{a+b}{2}$:

$$I(f) = \int_a^b f(x)dx \approx I_0^{сп}(f) = (b-a)f(x_0),\ x_0 = \frac{a+b}{2}. \tag{2.1}$$

Знайдемо алгебраїчний степінь точності цих квадратурних формул. Для формули лівих прямокутників:

$$I_0^{лп}(1) = b - a = I(1), \quad I_0^{лп}(x) = a(b-a) \neq I(x) = \int_a^b x dx = \left.\frac{x^2}{2}\right|_a^b = \frac{b^2 - a^2}{2}.$$

Отже алгебраїчний степінь точності $m = 0$. Таким же буде він і для квадратурної формули правих прямокутників $I_0^{пп}$. Для формули середніх прямокутників маємо

$$I_0^{сп}(1) = b - a = I(1), \;\; I_0^{сп}(x) = (b-a)\frac{a+b}{2} = I(x), \;\; I_0^{сп}(x^2) \neq I(x^2),$$

тому $m = 1$. Отже далі будемо користуватися формулою середніх прямокутників (2.1) і називати її просто формулою прямокутників

$$I_0(f) = I_0^{сп}(f).$$

Розглянемо сітку $\overline{\omega}_h = \{x_i = a + ih, i = 0,..,N, h = (b-a)/N\}$.

Подамо формулу (1) у вигляді $\int_a^b f(x)dx = \sum_{i=1}^{N} \int_{x_{i-1}}^{x_i} f(x)dx$ та розглянемо квадратурні формули на елементарному проміжку довжини h.

<u>Формула прямокутників (середніх).</u>

Розглянемо формулу прямокутників на проміжку довжини h :

$$\int_{x_{i-1}}^{x_i} f(x)dx \approx f(x_{i-1/2})h,$$

де $x_{i-1/2} = x_i - \frac{h}{2}$. Тоді

$$R_{0,i} = \int_{x_{i-1}}^{x_i} f(x)dx - f_{i-1/2}h = \int_{x_{i-1}}^{x_i} (f(x) - f_{i-1/2})\,dx.$$

Використовуючи розвинення в ряд Тейлора

$$f(x) = f_{i-1/2} + (x - x_{i-1/2})f'_{i-1/2} + \frac{(x - x_{i-1/2})^2}{2} f''(\xi_i), \; \xi_i \in [x_{i-1}, x_i],$$

отримаємо

$$R_{0,i} = \int_{x_{i-1}}^{x_i} \frac{(x - x_{i-1/2})^2}{2} f''(\xi_i)dx,$$

а отже

$$|R_{0,i}| \le M_{2,i} \int_{x_{i-1}}^{x_i} \frac{(x - x_{i-1/2})^2}{2} dx = \frac{h^3}{24} M_{2,i}$$

або

$$|R_{0,i}| \le \frac{h^3}{24} M_{2,i},$$

де $M_{2,i} = \max\limits_{x_{i-1} \le x \le x_i} |f''(x)|$.

Відзначимо, що для функції $f(x) = (x - x_{i-1/2})^2$ ця оцінка не може бути покращена.

Геометричний зміст формули прямокутників полягає в тому, що площа криволінійної трапеції замінюється на площу прямокутника.

На сітці $\overline{\omega}_h$ *складена квадратурна формула прямокутників* має вигляд

$$I(f) = \int_a^b f(x)dx = \sum_{i=1}^{N} \int_{x_{i-1}}^{x_i} f(x)dx \approx \sum_{i=1}^{N} hf(\overline{x}_i) = I_h^{пр}(f), \tag{2.4}$$

де $\overline{x}_i = x_{i-1/2} = x_i - \frac{h}{2}$. Оцінимо похибку цієї квадратурної формули:

$$R_h(f) = I(f) - I_h(f) = \sum_{i=1}^{N} \int_{x_{i-1}}^{x_i} \left(f(x) - f(\overline{x}_i)\right) dx = \sum_{i=1}^{N} f''(\xi_i) \frac{h^3}{24}, \xi_i \in [x_{i-1}, x_i],$$

$$|R_h(f)| \le \frac{M_2}{24} \sum_{i=1}^{N} h^3 = \frac{M_2 h^2}{24} \sum_{i=1}^{N} h = \frac{M_2 h^2 (b-a)}{24}, \tag{2.5}$$

де $M_2 = \max\limits_{a \le x \le b} |f''(x)|$, $f(x) \in C^2[a,b]$.

<u>Визначення.</u> Квадратурна формула складеного типу I_h має порядок (степінь) точності p по кроку h, якщо $R_h(f) = I(f) - I_h(f) = O(h^p)$. □

Тобто складена квадратурна формула прямокутників має степінь точності $p = 2$ по кроку h.

Якщо $f(x)\in C^4[a,b]$, то отримаємо

$$f(x)-f(\bar{x}_i)=f(\bar{x}_i)+(x-\bar{x}_i)f'(\bar{x}_i)+\frac{(x-\bar{x}_i)^2}{2}f''(\bar{x}_i)+$$
$$+\frac{(x-\bar{x}_i)^3}{6}f'''(\bar{x}_i)+\frac{(x-\bar{x}_i)^4}{24}f^{(4)}(\xi_i)-f(\bar{x}_i),\ x\in[x_{i-1},x_i]\ ,\ \xi_i\in[x_{i-1},x_i]$$

Тоді

$$R_h(f)=\sum_{i=1}^{N}\int_{x_{i-1}}^{x_i}\frac{(x-\bar{x}_i)^2}{2}f''(\bar{x}_i)dx+\sum_{i=1}^{N}\int_{x_{i-1}}^{x_i}\frac{(x-\bar{x}_i)^4}{24}f^{(4)}(\xi_i)dx=$$
$$=\frac{h^2}{24}\sum_{i=1}^{N}hf''(\bar{x}_i)+\sum_{i=1}^{N}\frac{h^5f^{(4)}(\eta_i)}{1920}=O(h^2),\quad \xi_i,\eta_i\in[x_{i-1},x_i]\,.$$

Оскільки $\sum_{i=1}^{N}hf''(\bar{x}_i)$ - це квадратурна формула середніх прямокутників для $f''(x)$ з похибкою $O(h^2)$, то

$$R_h(f)=\frac{h^2}{24}\int_a^b f''(x)dx+O(h^4)=O(h^4).$$

$$R_h(f)=\overset{0}{R}_h(f)+\alpha(h),\ \text{де}\ \overset{0}{R}_h(f)=\frac{h^2}{24}\int_a^b f''(x)dx,\ \alpha(h)=O(h^4). \qquad (2.6)$$

Такі оцінки використовуються для побудови програм з автоматичним вибором кроку інтегрування.

<u>Формула трапецій</u>

На проміжку $[x_{i-1},x_i]$ замінимо функцію $f(x)$ інтерполяційним поліномом Лагранжа 1-го степеня

$$L_{1,i}(x)=\frac{1}{h}((x-x_{i-1})f_i-(x-x_i)f_{i-1})$$

Тоді отримаємо формулу трапецій

$$\int_{x_{i-1}}^{x_i}f(x)dx\approx\int_{x_{i-1}}^{x_i}L_{1,i}(x)dx=h\frac{f(x_i)+f(x_{i-1})}{2}\,. \qquad (2.7)$$

Залишковий член інтерполяційного полінома иає вигляд

$f(x)-L_{1,i}(x)=\dfrac{(x-x_{i-1})(x-x_i)}{2}f''(\xi_i)$, тоді для залишкового члена квадратурної формули трапецій будемо мати

$$R_{1,i}(f)=\int_{x_{i-1}}^{x_i}(f(x)-L_{1,i}(x))dx=\int_{x_{i-1}}^{x_i}\frac{(x-x_{i-1})(x-x_i)}{2}f''(\xi_i)\,dx=$$
$$=-\frac{f''(\xi_i)h^3}{12},\quad \xi_i\in[x_{i-1},x_i], \tag{2.8}$$

а

$$|R_{1,i}|\le\frac{M_{2,i}h^3}{12},\quad \text{де } M_{2,i}=\max_{x_{i-1}\le x\le x_i}|f''(x)|. \tag{2.9}$$

Геометричний зміст формули трапецій полягає в тому, що площа криволінійної трапеції замінюється на площу прямокутної трапеції.

Складена квадратурна формула трапецій

$$I(f)\approx I_h^{mp}(f)=\sum_{i=1}^{N}\frac{h}{2}\big(f(x_{i-1})+f(x_i)\big)=\frac{h}{2}f(x_0)+\sum_{i=1}^{N-1}h\,f(x_i)+\frac{h}{2}f(x_N), \tag{2.10}$$

де $x_i=a+ih,\ h=(b-a)/N,\ i=\overline{0,N}$. Оцінка залишкового члена складеної квадратурної формули трапецій має наступний вигляд

$$\left|R_h(f)\right|\le M_2\frac{(b-a)}{12}h^2, \tag{2.11}$$

де $M_2=\max\limits_{a\le x\le b}|f''(x)|,\ f\in C^2[a,b]$.

Степінь точності по кроку h буде $p=2$.

Якщо $f\in C^4[a,b]$, то

$$R_h(f)=\overset{0}{R}_h(f)+\alpha(h),\ \text{де } \overset{0}{R}_h(f)=-\frac{h^2}{12}\int_a^b f''(x)dx,\ \alpha(h)=O(h^4). \tag{2.12}$$

<u>Формула Сімпсона (парабол*)*</u>

Замінимо $f(x)$ інтерполяційним поліномом $L_{2,i}(x)$ по вузлах $x_{i-1},x_{i-1/2},x_i$

$$L_{2,i}(x)=\frac{2}{h^2}\left((x-x_{i-1/2})(x-x_i)f_{i-1}-2(x-x_{i-1})(x-x_i)f_{i-1/2}+(x-x_i)(x-x_{i-1/2})f_i\right).$$

Тоді отримаємо квадратурну *формулу Сімпсона*

$$I(f)=\int_{x_{i-1}}^{x_i} f(x)\,dx \approx I_{2,i}=\int_{x_{i-1}}^{x_i} L_{2,i}(x)dx=\frac{h}{6}(f_{i-1}+4f_{i-1/2}+f_i),\ h=x_i-x_{i-1}. \quad (2.13)$$

Зауважимо, що формула Сімпсона точна на поліномах 3-го степеня.

Побудуємо поліном Ерміта, який задовольняє умовам

$$H_3(x_{i-1})=f_{i-1},\ H_3(x_i)=f_i, H_3^{(k)}(x_{i-1/2})=f_{i-1/2}^{(k)}, k=0,1.$$

Тоді $r_{3,i}(x)=f(x)-H_3(x)=\dfrac{f^{(4)}(\xi_i)}{4!}(x-x_i)(x-x_{i-1/2})^2(x-x_{i-1})$ - залишковий член полінома Ерміта. Для залишкового члена квадратурної формули Сімпсона будемо мати

$$R_{3,i}=\frac{1}{24}\int_{x_{i-1}}^{x_i}(x-x_i)(x-x_{i-1/2})^2(x-x_{i-1})f^{(4)}(\xi_i)dx=-\frac{f^{(4)}(\eta_i)}{2880}h^5 \quad (2.14)$$

та

$$|R_{3,i}|\le\frac{M_{4,i}}{24}\left|\int_{x_{i-1}}^{x_i}(x-x_i)(x-x_{i-1/2})^2(x-x_{i-1})dx\right|$$

або

$$|R_{3,i}^c|\le\frac{h^5}{2880}M_{4,i}\ , \quad (2.15)$$

де $M_{4,i}=\max\limits_{x_{i-1}\le x\le x_i}|f^{(4)}(x)|$.

Складена квадратурна формула Сімпсона має наступний вигляд

$$I(f)\approx I_h(f)=\sum_{i=1}^{N}\frac{h}{6}\left(f(x_{i-1})+4f(x_{i-1/2})+f(x_i)\right)=$$

$$=\frac{h}{6}\left(f(x_0)+f(x_N)+4(f(x_{1/2})+f(x_{3/2})+\ldots+f(x_{N-1/2}))+2(f(x_1)+f(x_2)\ldots+f(x_{N-1}))\right)=$$

$$=\frac{h}{6}\left(f(x_0)+f(x_N)+2\sum_{i=1}^{N-1}f(x_i)+4\sum_{i=1}^{N}f(x_{i-1/2})\right), \quad (2.16)$$

де $x_i=a+ih,\ h=\dfrac{b-a}{N},\ i=\overline{0,N}$.

Якщо $f\in C^4[a,b]$, то має місце оцінка

$$|R_h(f)|\le\frac{M_4}{2880}(b-a)h^4, \quad (2.17)$$

де $M_4=\max\limits_{a\le x\le b}|f^{(4)}(x)|$.

Степінь точності по кроку h буде $p=4$.

Якщо $f \in C^6[a,b]$, то

$$R_h(f) = \overset{0}{R}_h(f) + \alpha(h), \text{ де } \overset{0}{R}_h(f) = \frac{h^4}{2880}\int_a^b f^{(4)}(x)dx\ ,\ \ \alpha(h) = O(h^6). \quad (2.18)$$

<u>Зауваження:</u> щоб не використовувати дробових індексів в формулі (2.16), можна позначити $x_i = a + 0.5ih,\ h = \dfrac{b-a}{N},\ i = \overline{0,2N}$ та записати тоді складену квадратурну формулу Сімпсона у вигляді

$$I(f) \approx$$
$$\approx \frac{b-a}{6N}\left[f(x_0) + f(x_{2N}) + 4(f(x_1) + f(x_3) + \ldots + f(x_{2N-1})) + 2(f(x_2) + f(x_4)\ldots + f(x_{2N-2}))\right].$$

3. Квадратурні формули інтерполяційного типу

Будемо розглядати квадратурні формули виду (1.3) для інтегралу (1.1) одразу на всьому проміжку у випадку, коли функція $f(x)$ є досить гладкою. Замінівши в інтегралі (1.1) функцію $f(x)$ інтерполяційним поліномом Лагранжа

$$L_n(x) = \sum_{k=0}^{n} \frac{\omega_n(x)}{(x-x_k)\omega_n'(x_k)} f(x_k), \quad (3.1)$$

де $\omega_n = (x-x_0)(x-x_1)\ldots(x-x_n)$, отримаємо наближену формулу

$$\int_a^b \rho(x) f(x) dx \approx \sum_{k=0}^{n} c_k f(x_k), \quad (3.2)$$

з коефіцієнтами

$$c_k = \int_a^b \rho(x) \frac{\omega_n(x)}{(x-x_k)\,\omega_n'(x_k)} dx\ ,\ k = \overline{0,n}\ . \quad (3.3)$$

<u>Визначення.</u> Формула (3.2) називається квадратурною формулою інтерполяційного типу, якщо її коефіцієнти обчислюються за формулами (3.3). □

Оскільки вираз для похибки інтерполювання можемо подати у вигляді

$$r_n(x) = \frac{f^{(n+1)}(\xi(x))}{(n+1)!}\omega_n(x),$$

то для похибки квадратурної формули будемо мати

$$R_n(f) = \frac{1}{(n+1)!}\int_a^b \rho(x)\omega_n(x) f^{(n+1)}(\xi(x))\,dx.$$

(3.4)

Звідси отримаємо наступну оцінку похибки квадратурної формули інтерполяційного типу

$$|R_n(f)| \le \frac{M_{n+1}}{(n+1)!}\int_a^b \rho(x)\,|\omega_n(x)|\,dx, \qquad (3.5)$$

де $M_{n+1} = \max\limits_{a\le x\le b} |f^{(n+1)}(x)|$. З формули (3.5) видно, що квадратурна формула інтерполяційного типу, побудована за $n+1$ неспівпадаючими вузлами $x_0, x_1, ..., x_n$ буде точною для будь - якого полінома степеня n. Справедливо також обернене твердження:

<u>Теорема 3.1</u> Якщо квадратурна формула

$$\int_a^b \rho(x) f(x)\,dx \approx \sum_{k=0}^{n} d_k f(x_k) \qquad (3.6)$$

є точною для будь-якого полінома степеня n, то вона є квадратурною формулою інтерполяційного типу.

<u>Доведення.</u> Потрібно показати, що d_k визначаються за формулою (3.3). Розглянемо фундаментальні поліноми Лагранжа $\Phi_k(x) = \dfrac{\omega_n(x)}{(x-x_k)\omega_n'(x_k)}$, $k = 0...n$, що мають степінь n, та обчислимо інтеграли $I_k = \int_a^b \rho(x)\Phi_k(x)\,dx$. За умовою теореми справедливі точні рівності $I_k = \sum_{l=0}^{n} d_l \Phi_k(x_l)$. Оскільки $\Phi_k(x_l) = \delta_{lk}$, то маємо $I_k = d_k$, $k = 0,1,...,n$.

З іншого боку

$$I_k = \int_a^b \rho(x)\frac{\omega_n(x)}{(x-x_k)\omega_n'(x)}\,dx.$$

Тобто d_k визначаються за формулою (3.3). Отже формула (3.6) є інтерполяційною. □

<u>Симетричні формули</u>

<u>Визначення</u>. Квадратурна формула (3.2) називається симетричною, якщо

1) n-парне;

2) вузли квадратурної формули розташовані симетрично відносно середини проміжка $[a,b]$, тобто

$$\frac{a+b}{2}-x_k=x_{n-k}-\frac{a+b}{2}, k=0,1...,n/2 . \square$$

З 1), 2) випливає $c_k=c_{n-k}, k=0,1,...n/2$ (довести).
Має місце

<u>Теорема 3.2.</u> Нехай $\rho(x)$ парна функція відносно точки $\frac{a+b}{2}$, тобто $\rho(\frac{a+b}{2}+x)=\rho(\frac{a+b}{2}-x),\ x\in\left[0,\frac{a+b}{2}\right]$ і квадратурна формула симетрична.
Тоді, якщо квадратурна формула інтерполяційного типу точна для будь-якого полінома степеня n, то вона буде точною і для будь-якого полінома степеня $n+1$. □

Отримані в другому розділі квадратурні формули є частинними випадками формул інтерполяційного типу при $n=0,1,2,\ \rho(x)\equiv 1$.

4. Формули Ньютона-Котеса

Формулами Ньютона - Котеса називають формули інтерполяційного типу, побудовані на рівномірній сітці, коли $x_k-x_{k-1}=h,\ k=1,2,...,n$.
Розрізняють два види формул Ньютона-Котеса: замкненого типу і відкритого типу, коли хоча б один з вузлів x_0 або x_n не співпадає з відповідною граничною точкою.
Нехай $x_k=a+kh, k=0,...,n,\ nh=b-a$.
В (3.3) зробимо заміну $x=a+th,\ t=\frac{x-a}{h}, t\in[0,n]$. Оскільки $x-x_k=h(t-k)$, будемо мати

$$\omega_n(x)=h^{n+1}t(t-1)...(t-n),$$

$$\omega_n'(x_k)=(x_k-x_0)(x_k-x_1)...(x_k-x_{k-1})(x_k-x_{k+1})...(x_k-x_n)=(-1)^{n-k}h^n k!(n-k)! \ .$$

Тоді

$$c_k=h\frac{(-1)^{n-k}}{k!(n-k)!}\int_0^n\rho(a+th)\frac{t(t-1)...(t-n)}{t-k}dt=(b-a)B_k^{(n)},\quad k=\overline{0,n}$$

де $$B_k^{(n)}=\frac{(-1)^{n-k}}{k!(n-k)!n}\int_0^n\rho(a+th)t(t-1)...(t-k+1)(t-k-1)...(t-n)dt .$$

Видно, що коефіцієнти $B_k^{(n)}$ не залежать від проміжку $[a,b]$ і для них можуть бути складені таблиці .
При $n=8$ і $n=10$ серед $B_k^{(n)}$ є від'ємні числа, окрім того при $\rho(x)\equiv 1$ $\sum_{k=0}^{\infty}|B_k^{(n)}|\to\infty$ при $n\to\infty$. Тоді з рівності $\sum_{k=0}^{n}B_k^{(n)}=\frac{1}{b-a}\int_a^b 1dx=1$

випливає, що при великих n серед коефіцієнтів $B_k^{(n)}$ є як додатні так і від'ємні, що призводить до чисельної нестійкості формул чисельного інтегрування. А саме, нехай функція $f(x)$ обчислюється з похибкою, тобто замість $f(x_k)$ маємо $\widetilde{f}(x_k) = f(x_k) + \varepsilon_k$. Тоді замість I_n отримуємо

$$\widetilde{I}_n = \sum_{k=0}^{n} c_k (f(x_k) + \varepsilon_k) = \sum_{k=0}^{n} c_k f(x_k) + \sum_{k=0}^{n} c_k \varepsilon_k = I_n + \Delta I_n, \text{ де } \Delta I_n = \sum_{k=0}^{n} c_k \varepsilon_k .$$

Оскільки з того, що квадратурна формула точна для $f(x) \equiv 1$, то будемо мати

$$\sum_{k=0}^{n} c_k = \int_a^b \rho(x)\,dx = M > 0 .$$

Якщо всі коефіцієнти c_k невід'ємні, то отримаємо

$$|\Delta I_n| \le \sum_{k=0}^{n} |c_k|\,|\varepsilon_k| = \sum_{k=0}^{n} c_k\,|\varepsilon_k| \le (\max_{0 \le k \le n} |\varepsilon_k|) M .$$

Тобто похибка при обчисленні інтегралу має той же порядок, що і похибка при обчисленні функції. В цьому випадку говорять, що квадратурна формула стійка.

Якщо коефіцієнти c_k мають різні знаки то сума $\sum_{k=0}^{n} |c_k|$ не буде рівномірно обмеженою по n і відповідно з ростом n зростає вплив похибки при обчисленні I_n. Тому не доцільно використовувати великі степені n при побудові квадратурних формул інтерполяційного типу, а використовують складені квадратурні формули.

4. Метод (правило) Рунге оцінки похибки чисельного інтегрування

Нехай I - невідоме точне значення інтеграла, I_h - відоме його наближене значення, яке обчислене за допомогою деякої складеної квадратурної формули з кроком h (тобто залежить від кроку сітки h). Припустимо, що $I_h \approx I$ та $I_h \to I$ при $h \to 0$ та мають місце співвідношення

$$R_h = I - I_h = \overset{0}{R}_h + \alpha(h) = C \cdot h^m + \alpha(h), \tag{1}$$

де $\overset{0}{R}_h = C \cdot h^m$ - головний член похибки, невідома стала $C \ne 0$ не залежить від h, $\alpha(h) = o(h^m)$, а саме $\alpha(h) = O(h^{m+k})$. Обчислимо $I_{h/2}$. Тоді з (1) випливає, що

$$I = I_h + Ch^m + \alpha(h), \tag{2}$$

$$I = I_{h/2} + C\frac{h^m}{2^m} + \alpha(h). \tag{3}$$

Віднявши від (3) співвідношення (2), отримаємо

$$I_{h/2} - I_h = \frac{Ch^m}{2^m}\left(2^m - 1\right) + \alpha(h). \tag{4}$$

Звідси, порівнюючи останню формулу з (3), будемо мати

$$\overset{0}{R}_{h/2} = \frac{Ch^m}{2^m} = \frac{I_{h/2} - I_h}{2^m - 1} \tag{5}$$

та відповідно

$$\overset{0}{R}_h = \frac{2^m}{2^m - 1}\left(I_{h/2} - I_h\right).$$

Підставивши (5) в (3), отримаємо

$$I - I_{h/2} = \frac{I_{h/2} - I_h}{2^m - 1} + \alpha(h)\ . \tag{6}$$

Обчислення наближеної оцінки похибки за формулою (5) при виконанні умов (1) називається *правилом Рунге*.
Формула (5) називається *апостеріорною оцінкою* похибки наближеного обчислення інтеграла I за допомогою наближення $I_{h/2}$. (Апріорні оцінки - це оцінки отримані до обчислення величини I_h, апостеріорні оцінки – після її обчислення). Якщо $C \neq 0$, то величина $\frac{I_{h/2} - I_h}{2^m - 1}$ має m - тий порядок відносно h і відрізняється від головного члена похибки, тобто від $C\left(\frac{h}{2}\right)^m$, на величину більш високого порядку відносно h.

Використовуючи формулу (5), можна сформулювати наступний *алгоритм обчислення інтеграла із заданою точністю* ε:

1) обчислюємо I_h, $I_{h/2}$, $\overset{0}{R}_{h/2}$; перевіряємо виконання умови $\left|\overset{0}{R}_{h/2}\right| < \varepsilon$.

Якщо так, то $I \approx I_{h/2}$. Якщо ж ні, то

2) обчислюємо $I_{h/4}$, $\overset{0}{R}_{h/4}$; перевіряємо $\left|\overset{0}{R}_{h/4}\right| < \varepsilon$ і т.д.

Процес продовжуємо поки не буде виконана умова $\left|\overset{0}{R}_{h/2^j}\right| < \varepsilon \quad j = 1,2,\ldots$.

Зауваження. Правило Рунге дає змогу оцінити не похибку, а її головний член з точністю $\alpha(h)$, тому для такого підходу потрібно, щоб виконувалась умова $|\alpha(h)| << |\overset{0}{R}_h|$. □

За допомогою оцінки головного члена похибки можна отримати більш точне наближення для I. З формули (6) будемо мати

$$I = I_{h/2} + \frac{I_{h/2} - I_h}{2^m - 1} + \alpha(h) = \frac{2^m I_{h/2} - I_h}{2^m - 1} + \alpha(h) \ . \qquad (7)$$

Тоді величина

$$I^*_{h/2} = I_{h/2} + R^0_{h/2} = \frac{2^m I_{h/2} - I_h}{2^m - 1} = \frac{2^m}{2^m - 1} I_{h/2} - \frac{1}{2^m - 1} I_h \qquad (8)$$

називається *уточненим за Річардсоном* наближеним значенням величини I. Отже *екстраполяційна формула Річардсона* має вигляд

$$I^*_{h/2} = \frac{2^m I_{h/2} - I_h}{2^m - 1} = \frac{2^m}{2^m - 1} I_{h/2} - \frac{1}{2^m - 1} I_h \qquad (9)$$

і для неї відповідно до (7) при $C \neq 0$ справедлива оцінка

$$I - I^*_{h/2} = \alpha(h) = O\left(h^{m+k}\right).$$

Зауваження. Для розглянутих складених квадратурних формул інтерполяційного типу будемо мати

Квадратурна формула	m	k
Середніх прямокутників	2	2
Трапецій	2	2
Сімпсона	4	2

□

Наприклад, для квадратурної формули трапецій ($m = 2$) і

$$I - I_h = Ch^2 + O\left(h^4\right),\ \overset{0}{R}_{h/2} = \frac{I_{h/2} - I_h}{3},\ h_0 = b - a$$

маємо:

$$R_h = -\frac{h^2}{12}\int_a^b f''(x)dx + O\left(h^4\right) = O\left(h^2\right).$$

Отже, якщо застосовувати екстраполяційну формулу Річардсона, то

$$I^*_{h/2} = \frac{4}{3} I_{h/2} - \frac{1}{3} I_h \tag{10}$$

і

$$I - I^*_{h/2} = O\left(h^4\right).$$

<u>Задача.</u> Записати явний вигляд квадратурної формули, яка отримується екстраполяцією Річардсона з квадратурної формули трапецій. □

Розглянемо, чому формулу прямокутників не доцільно використовувати при застосуванні правила Рунге. Нехай точки, в яких обчислюється значення функції позначаються: в $I_h \leftrightarrow \circ$, в - $I_{h/2} \leftrightarrow *$.

Для формули трапецій використовуються такі точки:

$$\begin{array}{l} \circ - \quad - \circ - \quad - \circ - \quad - \circ - \quad - \circ - \quad - \circ - \quad - \circ - \quad - \circ - \quad - \circ \quad I_h \\ \circ - * - \circ - \quad * - \circ - * - \circ - * - \circ - * - \circ - * - \circ - * - \circ - * - \circ \quad I_{h/2} \end{array}.$$

Для формули прямокутників:

$$\begin{array}{l} - \circ - \quad - \circ - \quad - \circ - \quad - \circ - \quad - \circ - \quad - \circ - \quad - \circ - \quad - \circ - \quad I_h \\ - * - * - * - * - * - \quad * - * \quad - * - * \quad - * - \quad * - * - * \quad - * - * - * - \quad I_{h/2} \end{array}.$$

Видно, що для формули трапецій необхідно обчислити нові значення в N точках, а для формули прямокутників в $2N$ точках.

Для економного використання обчислених значень функції на сітці з кроком h для формули трапецій запишемо

$$I^{mp}_h = \frac{h}{2}\left(f(a) + 2\sum_{i=1}^{N-1} f(x_i) + f(b) \right),$$

$$I^{mp}_{h/2} = \frac{h}{4}\left(f(a) + 2\sum_{i=1}^{N-1} f(x_i) + 2\sum_{i=1}^{N} f(x_{i-1/2}) + f(b) \right) = \frac{1}{2} I^{mp}_h + \frac{h}{2}\sum_{i=1}^{N} f(x_{i-1/2}). \tag{11}$$

Отже, на одному кроці методу Рунге кількість обчислень значень функції $Q_{mp} = O(N)$, а для $Q_{np} = O(2N)$. Формула виду (11) справедлива також для квадратурних формул лівих та правих прямокутників.

Це правило застосовується і для формули Сімпсона ($m = 4$). Головна частина залишкового члена для цієї формули має вигляд

$$\overset{0}{R}_{h/2} = \frac{I_{h/2} - I_h}{15}.$$

Використання формули Річардсона дає

$$I_{h/2}^{*} = \frac{16}{15} I_{h/2} - \frac{1}{15} I_h,$$

$$I - I_{h/2}^{*} = O\left(h^6\right).$$

Якщо порядок точності m квадратурної формули невідомий, то для його оцінки можна використати правило Рунге. Для цього розглянемо величини I_h, $I_{h/2}$, $I_{h/4}$ і по аналогії з формулою (4), отримаємо

$$I_{h/2} - I_h = \frac{Ch^m}{2^m}\left(2^m - 1\right) + \alpha(h),$$

$$I_{h/4} - I_{h/2} = \frac{C\left(\frac{h}{2}\right)^m}{2^m}\left(2^m - 1\right) + \alpha(h).$$

Тоді з точністю $\alpha(h)$ будемо мати

$$2^m = \frac{I_{h/2} - I_h}{I_{h/4} - I_{h/2}},$$

звідки

$$m = \log_2 \left| \frac{I_{h/2} - I_h}{I_{h/4} - I_{h/2}} \right|.$$

Розглянемо використання так званих *адаптивних квадратурних формул*, в яких змінний крок вибирається за правилом Рунге. Для цього запишемо формулу трапецій із змінним кроком:

$$I_h^{mp}(f) = \sum_{i=1}^{N} \frac{h_i}{2}\left(f\left(x_{i-1}\right) + f\left(x_i\right)\right), \text{ де } h_i = x_i - x_{i-1}.$$

Оцінимо похибку на кожному з відрізків $[x_{i-1}, x_i]$:

$$R_{h_i} = I_i - I_{h_i} = \int_{x_{i-1}}^{x_i} f(x)\,dx - \frac{h_i}{2}\left(f(x_{i-1}) + f(x_i)\right) = -\frac{h_i^3}{12} f''(x_{i-1/2}) + O(h_i^5).$$

Таким чином $m = 3$ і головний член похибки оцінимо за формулою Рунге

$$\overset{0}{R}_{h_i/2} = \frac{I_{h_i/2} - I_{h_i}}{7}.$$

Кожен з відрізків $[x_{i-1}, x_i]$, $i = \overline{1,n}$ довжиною h_i розбиваємо навпіл, поки не буде виконуватися умова

$$\left| \overset{0}{R}_{h_i/2} \right| \le \frac{\varepsilon \cdot h_i}{b-a},$$

що забезпечує точність на всьому інтервалі $[a,b]$:

$$\left| \overset{0}{R}_{h/2} \right| = \left| \sum_{i=1}^{N} \overset{0}{R}_{h_i/2} \right| \le \sum_{i=1}^{N} \frac{\varepsilon\, h_i}{b-a} = \varepsilon \frac{b-a}{b-a} = \varepsilon .$$

Ще одне застосування правила Рунге – обчислення з високою точністю наближеного значення інтеграла від достатньо *гладкої функції* за допомогою *таблиць Ромберга*. Розглянемо побудову такої таблиці за допомогою складеної квадратурної формули трапецій із сталим кроком h. Обчислимо послідовність значень $I_h = I_h^{(0)}, I_{h/2} = I_{h/2}^{(0)}, I_{h/4} = I_{h/4}^{(0)}, I_{h/8} = I_{h/8}^{(0)}, \ldots$, які мають похибку $O(h^2)$. За допомогою екстраполяції Річардсона (8) з коефіцієнтами лінійної комбінації $\left(\frac{4}{3}, -\frac{1}{3}\right)$ уточнимо ці значення (формула (10)). Отримаємо $I_{h/2}^{(1)}, I_{h/4}^{(1)}, I_{h/8}^{(1)}, \ldots$. Вони мають похибку $O(h^4)$. Знову використовуємо екстраполяцію Річардсона з коефіцієнтами лінійної комбінації $\left(\frac{16}{15}, -\frac{1}{15}\right)$. Отримаємо значення $I_{h/4}^{(2)}, I_{h/8}^{(2)}, \ldots$, які мають точність $O(h^6)$ і т.д.. Отримані значення можна розмістити в наступній *таблиці Ромберга*

$O(h^2)$	$O(h^4)$	$O(h^6)$	$O(h^8)$
$I_h^{(0)}$			
$I_{h/2}^{(0)}$	$I_{h/2}^{(1)}$		
$I_{h/4}^{(0)}$	$I_{h/4}^{(1)}$	$I_{h/4}^{(2)}$	
$I_{h/8}^{(0)}$	$I_{h/8}^{(1)}$	$I_{h/8}^{(2)}$	$I_{h/8}^{(3)}$
$\left(\frac{4}{3}, -\frac{1}{3}\right)$	$\left(\frac{16}{15}, -\frac{1}{15}\right)$	$\left(\frac{64}{63}, -\frac{1}{63}\right)$	

Похибку всіх елементів отриманої таблиці (окрім розташованих на верхній діагоналі) можна оцінити за правилом Рунге (5), тобто

$$\left| I - I^{(l)}_{h/2^j} \right| \approx \left| \frac{I^{(l)}_{h/2^j} - I^{(l)}_{h/2^{j-1}}}{2^{m+kl}} \right|, \quad j = 1,2,... \quad .$$

Використання формули (11) для обчислення $I^{(0)}_{h/2^j}$ та лінійні комбінації виду (9) дають простий та економний алгоритм наближеного обчислення інтеграла I. Початкове значення h можна брати рівним $b-a$ або $\frac{b-a}{n}$, де n ціле. Зауважимо ще раз, що застосування правила Рунге при побудові таблиць Ромберга вимагає виконання умови $|\alpha(h)| << |\overset{0}{R}_h|$.

Квадратурні формули найвищого алгебраїчного степеня точності

При побудові квадратурні формули інтерполяційного типу підінтегральну функцію замінювали інтерполяційним поліномом, побудованим за $n+1$ вузлами, причому ці вузли були попередньо задані . В загальному випадку алгебраїчний степінь точності таких формул дорівнює n (у випадку симетричного відносно середньої точки відрізку розташування непарної кількості вузлів дорівнює $n+1$). Виявляється, що за рахунок вибору вузлів можна побудувати квадратурні фомули, які будуть точними для поліномів більш високого степеня.

Розглянемо інтеграл

$$I(f) = \int_a^b \rho(x) f(x) dx \ , \tag{1}$$

де

$$\rho(x) > 0, \quad x \in [a,b]; \quad \left| \int_a^b \rho(x) x^i dx \right| < \infty, \ (i = 0,1,2...).$$

Побудуємо квадратурну формулу

$$\int_a^b \rho(x) f(x) dx \approx \sum_{k=1}^{n} c_k f(x_k), \tag{2}$$

яка при заданому n була б точною для алгебраїчних поліномів якомога більш високого степеня. Такі квадратурні формули існують, вони називаються *квадратурними формулами найвищого алгебраїчного степеня точності* (або формулами Гауса). Зауважимо, що в формулах (2) нумерація вузлів починається з 1.

В (2) невідомими є c_k, x_k, $k = \overline{1,n}$. Їх обирають з умови, щоб формула (2) була точною для будь-якого полінома степеня m , а це еквівалентно умові, що формула (2) буде точною для функцій $f(x) = x^\alpha, \alpha = 0,1,...,m$. Тоді отримуємо умови

$$\int_a^b \rho(x) x^\alpha dx = \sum_{k=1}^{n} c_k x_k^\alpha, \ \alpha = \overline{0,m}. \qquad (3)$$

(3) – це система $m+1$ нелінійних алгебраїчних рівнянь відносно $2n$ невідомих $c_1, c_2, ..., c_n, x_1, x_2, ..., x_n$. При $m = 2n-1$ кількість рівнянь в (3) буде рівною кількості невідомих. Розглянемо кілька найпростіших випадків.

<u>Приклад 1.</u> Побудувати квадратурну формулу найвищого алгебраїчного степеня точності для $\rho(x) = 1, -1 \le x \le 1, n = 1$.
Розв'язання. При $n = 1$ будемо мати $m = 2n - 1 = 1$. З системи (3) отримаємо

$$\begin{cases} \int_{-1}^{1} x^0 dx = c_1, & (\alpha = 0), \\ \int_{-1}^{1} x dx = c_1 x_1, & (\alpha = 1), \end{cases} \Rightarrow \begin{cases} c_1 = 2 \\ x_1 = 0 \end{cases}.$$

Тоді, згідно формули (2), будемо мати

$$I(f) = \int_{-1}^{1} f(x) dx \approx 2 f(0).$$

Тобто отримали формулу середніх прямокутників для відрізку $[-1,1]$, яка є точною для поліномів першого степеня. □

<u>Приклад 2.</u> Побудувати квадратурну формулу найвищого алгебраїчного степеня точності для $\rho(x) = 1, -1 \le x \le 1, n = 2$.
Розв'язання. При $n = 2$ будемо мати $m = 2n - 1 = 3$. Із системи (3) отримаємо

$$\begin{cases} \int_{-1}^{1} x^0 dx = c_1 + c_2, \\ \int_{-1}^{1} x dx = c_1 x_1 + c_2 x_2, \\ \int_{-1}^{1} x^2 dx = c_1 x_1^2 + c_2 x_2^2, \\ \int_{-1}^{1} x^3 dx = c_1 x_1^3 + c_2 x_2^3. \end{cases} \Rightarrow \begin{cases} c_1 + c_2 = 2, \\ c_1 x_1 + c_2 x_2 = 0, \\ c_1 x_1^2 + c_2 x_2^2 = 2/3, \\ c_1 x_1^3 + c_2 x_2^3 = 0. \end{cases}$$

З першого рівняння останньої системи маємо $c_2 = 2 - c_1$. Підставляючи цей вираз в друге та четверте рівняння, отримаємо

$c_1x_1 = (c_1 - 2)x_2, \;\; c_1x_1^3 = (c_1 - 2)x_2^3.$

Звідси поділивши другий вираз на перший отримаємо $x_2 = -x_1$. Тоді з другого рівняння системи будемо мати $c_2 = c_1$. Це разом з першим рівнянням дає $c_1 = c_2 = 1$. Враховуючи знайдені значення, з третього рівняння маємо $x_1^2 = \frac{1}{3}$, звідки $x_1 = -\frac{1}{\sqrt{3}}, x_2 = \frac{1}{\sqrt{3}}$. Остаточно квадратурна формула буде мати вигляд

$$\int_{-1}^{1} f(x)dx \approx f\left(-\frac{1}{\sqrt{3}}\right) + f\left(\frac{1}{\sqrt{3}}\right). \;\; \square$$

Зауважимо, що розв'язання системи нелінійних рівнянь (3) є досить непростою задачею, тому розглянемо інший підхід для побудови квадратурних формул найвищого алгебраїчного степеня точності.

<u>Теорема 1</u> (Гауса). Квадратурна формула (2) буде точною для будь-якого полінома степеня $m \le 2n-1$ тоді і тільки тоді, коли виконуються умови:

1) поліном $\omega(x) = (x - x_1)(x - x_2)...(x - x_n)$ ортогональний на відрізку $[a, b]$ з ваговою функцією $\rho(x)$ до будь-якого полінома $Q(x)$ степеня менше n, тобто

$$\int_a^b \rho(x)\omega(x)Q(x)dx = 0, \;\; \forall Q(x) \in \pi_{n-1}, \tag{4}$$

де через π_{n-1} позначено множину всіх алгебраїчних поліномів до $(n-1)$-го степеня включно;

2) формула (2) є квадратурною формулою інтерполяційного типу, тобто її коефіцієнти обчислюються за формулою

$$c_k = \int_a^b \rho(x)\frac{\omega(x)}{(x - x_k)\omega'(x_k)}dx, \; k = \overline{1,n} \;\; . \tag{5}$$

Доведення.

<u>Необхідність</u>. Нехай формула (2) точна для поліномів степеня $m \le 2n-1$, тобто

$$I(f) = I_n(f), \;\; \forall f(x) \in \pi_{2n-1} \;. \tag{6}$$

Тоді $\forall Q(x) \in \pi_{n-1}$ маємо $Q(x)\omega(x) \in \pi_{2n-1}$ і з співвідношення (6) отримаємо

$$\int_a^b \rho(x)\omega(x)Q(x)dx = \sum_{k=1}^{n} c_k\omega(x_k)Q(x_k) = 0,$$

оскільки що $\omega(x_k) = 0, \; k = \overline{1,n}$. Тобто виконується (4).

Окрім того, якщо формула (2) записана за n вузлами точна для поліномів степеня $m \le 2n-1$, то вона буде точною і для поліномів степеня менше n, а отже буде формулою інтерполяційного типу і її коефіцієнти будуть обчислюватись за формулою (5).

Достатність. Нехай виконуються умови (4), (5). Візьмемо $\forall f(x) \in \pi_{2n-1}$. Тоді її можна подати у вигляді
$f(x) = \omega(x)Q(x) + r(x)$, де $Q(x), r(x) \in \pi_{n-1}$.
Розглянемо

$$\int_a^b \rho(x) f(x) dx = \int_a^b \rho(x)\omega(x)Q(x)dx + \int_a^b \rho(x) r(x) dx = \int_a^b \rho(x) r(x) dx. \qquad (7)$$

Оскільки виконуються умови (5), то (2) буде квадратурною формулою інтерполяційного типу, а отже вона є точною для поліномів степеня менше n. З огляду на те, що $r(x) \in \pi_{n-1}$ та $r(x_k) = f(x_k) - \omega(x_k)Q(x_k) = f(x_k)$ (оскільки $\omega(x_k) = 0,\ k = \overline{1,n}$), отримаємо

$$\int_a^b \rho(x) r(x) dx = \sum_{k=1}^n c_k\, r(x_k) = \sum_{k=1}^n c_k \left(f(x_k) - \omega(x_k)Q(x_k)\right) = \sum_{k=1}^n c_k f(x_k). \qquad (8)$$

Отже з (7) і (8) будемо мати

$$\int_a^b \rho(x) f(x) dx = \sum_{k=1}^n c_k f(x_k),$$

тобто квадратурна формула (2) є точною для будь-якого полінома степеня $m = 2n-1$. □

Відзначимо, що теорема 1 значно спрощує побудову формул найвищого алгебраїчного степеня точності. Окрім того, зауважимо, що умови (4) еквівалентні наступним вимогам

$$\int_a^b \rho(x)\omega(x)x^\alpha dx = 0,\ \alpha = \overline{0, n-1}. \qquad (9)$$

Таким чином задача побудови формул Гауса призводить до необхідності дослідження існування та єдності полінома $\omega(x)$ степеня n ортогонального з вагою $\rho(x) > 0$ на відрізку $[a,b]$ до всіх поліномів степеня меншого n, а також потрібно переконатися, що всі його корені дійсні, різні і належать відрізку $[a,b]$.

Теорема 2. На відрізку $[a,b]$ існує та єдиний поліном $\omega(x)$ степеня n із старшим коефіцієнтом рівним 1, який буде ортогональним з ваговою функцією $\rho(x) > 0$ до будь-якого полінома степеня менше n на $[a,b]$.

Доведення. Будемо шукати поліном $\omega(x)$ у вигляді
$\omega(x) = a_0 + a_1 x + \ldots + a_{n-1}x^{n-1} + x^n$. Тоді з умов ортогональності (9) для визначення коефіцієнтів $a_0, a_2, \ldots, a_{n-1}$ отримаємо

$$\int_a^b \rho(x)\omega(x)x^\alpha dx = \int_a^b \rho(x)\left(a_0 + a_1 x + \ldots + a_{n-1}x^{n-1} + x^n\right)x^\alpha dx = 0,\ \alpha = \overline{0, n-1},$$

або

$$\int_a^b \rho(x)\left(a_0 + a_1 x + \ldots + a_{n-1}x^{n-1}\right)x^\alpha dx = -\int_a^b \rho(x) x^n x^\alpha dx.$$

Покажемо, що відповідна однорідна система рівнянь

$$\int_a^b \rho(x)\left(a_0 + a_1 x + ... + a_{n-1}x^{n-1}\right)x^\alpha dx = 0,\ \alpha = \overline{0, n-1} \tag{10}$$

має єдиний розв'язок $a_0 = a_1 = ... = a_{n-1} = 0$.

Помножимо систему (10) на a_α і просумуємо по всіх $\alpha = \overline{0, n-1}$:

$$\sum_{\alpha=0}^{n-1} a_\alpha \int_a^b \rho(x)\left(a_0 + a_1 x + ... + a_{n-1}x^{n-1}\right)x^\alpha dx = \int_a^b \rho(x)\sum_{\alpha=0}^{n-1}\sum_{j=0}^{n-1} a_j\, a_\alpha\, x^\alpha x^j dx =$$

$$= \int_a^b \rho(x)\left(\sum_{j=0}^{n-1} a_j x^j\right)^2 dx = 0.$$

Звідси і з умови $\rho(x) > 0$ випливає, що $a_0 = a_1 = ... = a_{n-1} = 0$. Тому відповідна неоднорідна система має єдиний розв'язок. Отже існує і єдиний поліном $\omega(x)$ степеня *n*, який ортогональний з вагою $\rho(x)$ до будь-якого полінома степеня менше n. □

<u>Теорема 3.</u> Якщо $\omega(x)$ - поліном степеня n із старшим коефіцієнтом рівним 1 ортогональний з ваговою функцією $\rho(x) > 0$ до будь-якого полінома степеня $m \le n-1$ на $[a,b]$, тоді при $n \ge 1$ всі корені полінома $\omega(x)$ дійсні, різні та розташовані всередині відрізку $[a,b]$.

Доведення. Розглянемо корені полінома $\omega(x)$, які розташовані всередині відрізка $[a,b]$ та мають непарну кратність (в цих точках $\omega(x)$ змінює знак). Нехай таких коренів m і позначимо їх $\xi_1, \xi_2, ..., \xi_m$. Для доведення теореми достатньо показати, що $m = n$, оскільки з цього випливає, що інших коренів у $\omega(x)$ немає і всі корені ξ_k - прості. Припустимо супротивне, тобто $m < n$, тоді при $n \ge 1$ запишемо $\omega(x) = (x-\xi_1)(x-\xi_2)...(x-\xi_m)q_{n-m}(x)$, причому поліном $q_{n-m}(x)$ має один і той же знак при всіх $x \in [a,b]$ (в протилежному випадку точка зміни знаку була б коренем $\omega(x)$ і повинна була б належати до множини ξ_i). Якщо $m < n$, то виходячи з умови ортогональності $\omega(x)$ до всіх поліномів степеня менше n повинна виконуватися рівність

$$\begin{aligned} &\int_a^b \rho(x)(x-\xi_1)^2(x-\xi_2)^2...(x-\xi_m)^2 q_{n-m}(x)dx = \\ &= \int_a^b \rho(x)\omega(x)(x-\xi_1)(x-\xi_2)...(x-\xi_m)dx = 0 \end{aligned} \tag{11}$$

З іншого боку видно, що добуток $(x-\xi_1)^2(x-\xi_2)^2...(x-\xi_m)^2 q_{n-m}(x)$ зберігає знак на відрізку $[a,b]$, окрім того він обертається в нуль в скінченій кількості точок, оскільки відмінний від тотожнього нуля. Маємо також, що $\rho(x) > 0$ на $[a,b]$, тому

$$\int_a^b \rho(x)(x-\xi_1)^2(x-\xi_2)^2...(x-\xi_m)^2 q_{n-m}(x)\,dx \ne 0,$$

а це суперечить (11). Отже припущення $m < n$ невірне і повинна виконуватись умова $m = n$. □

Доведені теореми дозволяють сформулювати наступне твердження: якщо $\rho(x) > 0,\ x \in [a,b]$, то квадратурна формула (2), яка є точною для поліномів степеня $2n-1$, існує для всіх $n = 1,2,...$ і є єдиною для кожного n. Окрім того вузли $x_k,\ k = \overline{1,n}$ квадратурної формули найвищого алгебраїчного степеня точності співпадають з нулями полінома $P_n(x)$, який ортогональний з ваговою функцією $\rho(x) > 0$ на відрізку $[a,b]$ до всіх поліномів меншого степеня (тобто належить до системи ортогональних поліномів на відрізку $[a,b]$ з вагою $\rho(x)$), а її коефіцієнти обчислюються за формулами (5). Такі формули називають ще формулами Гауса.

Розглянемо деякі **властивості квадратурних формул Гауса**.

1) Покажемо, що $m = 2n-1$ - найвищий алгебраїчний степінь точності квадратурних формул Гауса, тобто існує поліном степеня $2n$, для якого формула Гауса не є точною.

Для цього розглянемо функцію $f(x) = \omega^2(x)$, яка є поліномом степеня $2n$. Тоді

$$I(\omega^2) = \int_a^b \rho(x)\omega^2(x)dx > 0, \quad \rho(x) > 0,$$

але

$$I_n(\omega^2) = \sum_{k=1}^{n} c_k \omega^2(x_k) = 0.$$

Отже, $I(\omega^2) \neq I_n(\omega^2)$. Звідси $m = 2n-1$.

2) Для довільного n коефіцієнти c_k формул Гауса (2) додатні, тобто $c_k > 0,\ k = \overline{1,n}.$

Розглянемо поліноми $\varphi_i(x) = \left(\dfrac{\omega(x)}{(x - x_i)\omega'(x_i)} \right)^2,\ i = \overline{1,n}$,

які мають степінь $2n-2$ і задовольняють умовам $\varphi_i(x_k) = \delta_{ik}$, $i,k = \overline{1,n}$.

Оскільки для цих поліномів справедлива теорема Гауса, то

$$0 < \int_a^b \rho(x)\varphi_i(x)dx = I(\varphi_i) = I_n(\varphi_i) = \sum_{k=1}^{n} c_k \varphi_i(x_k) = c_i . \qquad (12)$$

Звідси випливає, що $c_i > 0,\ i = \overline{1,n}$.

Зауваження. При розгляді квадратурних формул інтерполяційного типу відзначалось, що додатність коефіцієнтів квадратурних формул дуже важлива для їх обчислювальної стійкості і дозволяє використовувати квадратурні формули з більшими значеннями n. □

Теорема 4. Нехай вагова функція $\rho(x)>0,\ x\in[a,b],\quad f(x)\in C^{2n}[a,b]$. Тоді існує така точка $\xi\in[a,b]$, що для похибки квадратурної формули найвищого алгебраїчного степеня точності виконується рівність

$$R_n(f)=\int_a^b \rho(x)f(x)dx-\sum_{k=1}^{n}c_k f(x_k)=\frac{f^{(2n)}(\xi)}{(2n)!}\int_a^b \rho(x)\omega^2(x)dx. \qquad (13)$$

Доведення. Розглянемо інтерполяційний поліном Ерміта $H(x)$ степеня $2n-1$, який задовольняє умовам: $f(x_i)=H(x_i)$, $f'(x_i)=H'(x_i),\ \ i=\overline{1,n}$. Для нього залишковий член інтерполювання має вигляд

$$r(x)=f(x)-H(x)=\frac{f^{(2n)}(\eta)}{(2n)!}\omega^2(x).$$

Оскільки $f(x)=H(x)+r(x)$, то будемо мати

$$\int_a^b \rho(x)f(x)dx=\int_a^b \rho(x)H(x)dx+\int_a^b \rho(x)\frac{f^{(2n)}(\eta)}{(2n)!}\omega^2(x)dx. \qquad (14)$$

Квадратурна формула Гауса точна для всіх поліномів до степеня $2n-1$ включно, тому $\int_a^b \rho(x)H(x)dx=\sum_{k=1}^{n}c_k H(x_k)=\sum_{k=1}^{n}c_k f(x_k)$, а тоді залишковий член $R_n(f)$ буде мати вигляд

$$R_n(f)=\int_a^b \rho(x)f(x)dx-\sum_{k=1}^{n}c_k f(x_k)=\frac{1}{(2n)!}\int_a^b \rho(x)\omega^2(x)f^{(2n)}(\eta)dx.$$

Оскільки $\rho(x)\omega^2(x)\geq 0$, то можна зробити висновок, що існує точка $\xi\in[a,b]$, для якої виконується співвідношення (13). □

Зауваження. Для обчислення коефіцієнтів c_k квадратурної формули (2) було отримано дві різних формули (5) та (12). Можна навести ще один більш зручний для обчислення вираз для знаходження c_k .

Розглянемо систему ортонормованих на $[a,b]$ з ваговою функцією $\rho(x)>0$ поліномів $P_n(x)=a_n x^n+a_{n-1}x^{n-1}+\ldots a_0$, $n=0,1,\ldots,\quad \int_a^b \rho(x)P_n^2(x)dx=1,\ a_n>0.$

Поліноми $P_n(x)$ відрізняються від $\omega(x)$ лише чисельним множником $P_n(x)=a_n\omega(x)$. Тоді, використовуючи відому з теорії ортогональних поліномів формулу Кристофеля-Дарбу, можна отримати наступний вираз

$$c_k=\frac{a_n}{a_{n-1}}\cdot\frac{1}{P_n'(x_k)P_{n-1}(x_k)},\ k=\overline{1,n}. \qquad (15)$$

Знайдені співвідношення більш зручні для обчислень, оскільки не вимагають інтегрування. □

Збіжність квадратурного процесу найвищого алгебраїчного степеня точності

Нехай $\rho(x) > 0,\ x \in [a,b]$. Квадратурні формулами найвищого алгебраїчного степеня точності можуть бути побудовані для довільного $n = 1,2,...$ Для кожного n вузли та коефіцієнти будуть мати свої значення, які будемо позначати відповідно x_k^n і c_k^n. Тоді

$$\int_a^b \rho(x) f(x) dx = \sum_{k=1}^{n} c_k^n f\left(x_k^n\right) + R_n(f) = I_n(f) + R_n(f). \qquad (16)$$

<u>Визначення.</u> говорять, що квадратурний процес збігається для функції f,

якщо $$\lim_{n\to\infty} I_n(f) = \int_a^b \rho(x) f(x) dx.$$

<u>Теорема 5.</u> Нехай $\rho(x) > 0,\ x \in [a,b]$ і функція $f(x) \in C[a,b]$. Тоді квадратурний процес найвищого алгебраїчного степеня точності збігається. □

<u>Зауваження.</u> Для квадратурних формул (16) загального вигляду має місце наступна

<u>Теорема 6.</u> Нехай $\rho(x) > 0,\ x \in [a,b]$ і функція $f(x) \in C[a,b]$. Тоді для того, щоб квадратурний процес (16) збігався необхідно та достатньо виконання умов :

а) квадратурний процес збігався для будь-якого алгебраїчного полінома;

б) існує число $M > 0$ таке, що виконується нерівність

$$\sum_{k=1}^{n} \left| c_k^n \right| \le M,\ n = 0,1,2,\ldots. \ \square$$

Частинні випадки квадратурних формул Гауса

1) Розглянемо відрізок $[-1,1]$ і ваговий множник $\rho(x) = 1$, тобто виведемо формули Гауса для обчислення інтегралу

$$\int_{-1}^{1} f(x) dx.$$

Для знаходження вузлів квадратурної формули розглянемо поліноми Лежандра, які мають вигляд

$$L_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n, \quad n = 0,1,...$$

Вони також задовольняють рекурентному співвідношенню

$$(n+1) L_{n+1}(x) = (2n+1) x L_n(x) - n L_{n-1}(x), \quad L_0 = 1,\ L_1(x) = x.$$

Поліноми Лежандра задовольняють умовам теореми 1, тому

$$\omega(x) = \frac{2^n (n!)^2}{(2n)!} L_n(x)$$

і вузлами квадратурної формули є корені цього полінома. Вагові множники обчислюються за формулою

$$c_k = \frac{2(1-x_k^2)}{n^2\, L_{n-1}^2(x_k)},\ k=\overline{1,n}\ ,$$

$$\left(\text{або } c_k = \frac{2}{(1-x_k^2)\,\left(L_n'(x_k)\right)^2},\ k=\overline{1,n}\right)$$

а залишковий член

$$R_n(f) = \frac{2^{2n+1}}{(2n+1)!}\frac{(n!)^4}{((2n)!)^3} f^{(2n)}(\xi)\ .$$

2) Розглянемо відрізок $[-1,1]$ і вагову функцію $\rho(x)=1/\sqrt{1-x^2}$, тобто розглянемо формули Гауса для обчислення інтегралу

$$I(f) = \int_{-1}^{1} \frac{f(x)dx}{\sqrt{1-x^2}}.$$

Поліноми Чебишова задовольняють умовам теореми 1, тому

$$\omega(x) = \overline{T}_n(x) = \frac{1}{2^{n-1}}\cos(n\arccos x).$$

Вузлами квадратурної формули є нулі полінома Чебишова, тобто корені рівняння $\cos(n\arccos x)=0$:

$$x_k = \cos\frac{(2k-1)\pi}{2n},\ k=\overline{1,n}.$$

Відповідні коефіцієнти обчислюються за формулами

$$c_k = \int_{-1}^{1} \frac{T_n(x)dx}{\sqrt{1-x^2}\ T_n'(x_k)(x-x_k)} = \frac{\pi}{n},\ k=\overline{1,n}.$$

Отже, квадратурні формули найвищого алгебраїчного степеня точності мають вигляд

$$\int_{-1}^{1} \frac{f(x)\,dx}{\sqrt{1-x^2}} \approx \frac{\pi}{n}\sum_{k=1}^{n} f(x_k)\ , \qquad (16)$$

де x_k – корені полінома Чебишова. Ці формули називають формулами Ерміта. Залишковий член має вигляд

$$R_n(f) = \frac{\pi}{2^{2n-1}(2n)!} f^{(2n)}(\xi).$$

<u>Зауваження.</u> Застосування квадратурних формул (16) є найбільш доцільним, коли значення функції $f(x_k)$ містять випадкові похибки.□

3) Розглянемо проміжок $(-\infty,\infty)$ і вагову функцію $\rho(x)=e^{-x^2}$, тобто розглянемо формули Гауса для обчислення інтеграла

$$\int_{-\infty}^{\infty} e^{-x^2} f(x)\,dx.$$

Тоді

$$\omega(x) = H_n(x)/2^n = \frac{(-1)^n}{2^n} e^{x^2} \frac{d^n}{dx^2} e^{-x^2},$$

де $H_n(x)$ поліноми Ерміта, які обчислюються за рекурентними формулами

$H_{n+1}(x) = 2xH_n(x) - 2n\,H_{n-1}(x)$, $H_{-1} = 0,\ H_0 = 1$.

Коефіцієнти квадратурної формули обчислюються наступним чином

$$c_k = \frac{2^{n+1} n! \sqrt{\pi}}{(H_n'(x_k))^2},\ k = \overline{1,n}.$$

Залишковий член має вигляд

$$R_n(f) = \frac{n!\sqrt{\pi}}{2^n (2n)!} f^{(2n)}(\xi).$$

Наприклад, при $n = 1$ $H_1(x) = 2x$. Корінь $x_0 = 0$, а

$$c_0 = \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}\ .$$

Квадратурна формула має вигляд

$I_1(x) = \sqrt{\pi} f(0)$.

4) Розглянемо відрізок $[0,\infty)$ і ваговий множник

$\rho(x) = x^\alpha e^{-x},\ (\alpha > -1)$, тобто отримаємо формули Гауса для обчислення інтегралу

$$\int_0^\infty x^\alpha e^{-x} f(x) dx.$$

Згідно теорії маємо

$$\omega(x) = L_n^\alpha(x) = (-1)^n x^{-\alpha} e^x \frac{d^n}{dx^n}\left(x^{\alpha+n} e^{-x}\right)$$

де $L_n^\alpha(x)$ поліноми Лаґерра, які задовольняють рекурентним співвідношенням

$(n+1)L_{n+1}^\alpha = (2n + \alpha + 1 - x)L_n^\alpha - (n+\alpha)L_{n-1}^\alpha$, де $L_0^\alpha = 1,\quad L_{-1}^\alpha = 0$.

Коефіцієнти квадратурної формули обчислюються за формулами

$$c_k = \frac{n!\,\Gamma(n+\alpha+1)}{x_k\,[(L_n^\alpha)'(x_k)]^2},\ k = \overline{1,n}.$$

Залишковий член має наступний вигляд

$$R_n(f) = \frac{n!\,\Gamma(n+\alpha+1)}{(2n)!} f^{(2n)}(\xi).$$

Інтегрування швидко змінних (осцилюючих) функцій.

Розглянемо інтеграл

$$I(f) = \int_a^b f(x) e^{i\omega x} dx,\ i^2 = -1. \qquad (1)$$

де $\omega(b-a)\geq 1$, $f(x)$ - гладка функція. Функції $\mathrm{Re}(f(x)\exp(i\omega x))$, $\mathrm{Im}(f(x)\exp(i\omega x))$ мають на проміжку $[a,b]$ приблизно $\frac{\omega(b-a)}{\pi}$ нулів. Оскільки поліном степені n має не більше n нулів, то такі функції можуть бути добре наближені поліномами n - го степеня лише при $n >> \frac{\omega(b-a)}{\pi}$.
Тому для наближеного обчислення інтегралів від таких функцій можна використовувати тільки інтерполяційні квадратурні формули, точні для алгебраїчних поліномів високого степеня. При великих ω застосування складених квадратурних формул інтерполяційного типу також приводить до великої похибки і при малих кроках h.
Більш прийнятним може бути шлях розгляду функції $\exp(i\omega x)$, як вагової. Функцію $f(x)$ замінюємо інтерполяційним поліномом. Для простоти візьмемо поліном першого степеня

$$f(x)=\frac{b-x}{h}f(a)+\frac{x-a}{h}f(b)+r_1(x), \tag{2}$$

де $h=b-a$.
Тоді

$$\int_a^b f(x)e^{i\omega x}dx \approx f(a)h^{-1}\int_a^b (b-x)e^{i\omega x}dx + f(b)h^{-1}\int_a^b (x-a)e^{i\omega x}dx =$$

$$=\frac{f(a)}{h\omega^2}\left(1-e^{i\omega h}+i\omega h\right)e^{i\omega a}+\frac{f(b)}{h\omega^2}\left(1-e^{-i\omega h}-i\omega h\right)e^{i\omega b}. \tag{3}$$

Сумуючи формулу (3) по елементарних відрізках довжини $h=x_i-x_{i-1}$ отримаємо складену формулу

$$\int_a^b f(x)e^{i\omega x}dx=\frac{f_0}{h\omega^2}\left(1-e^{i\omega h}+i\omega h\right)e^{i\omega x_0}+$$

$$+2\frac{1-\cos\omega h}{h\omega^2}\sum_{j=1}^{n-1}f_j e^{i\omega x_j}+\frac{f_n}{h\omega^2}\left(1-e^{-i\omega h}-i\alpha h\right)e^{i\omega x_n}, \tag{4}$$

$$x_j=a+jh, h=(b-a)/n,\ j=\overline{0,n}.$$

В загальному випадку замінимо відрізок $[a,b]$ на $[-1,1]$:
$x_j=\frac{a+b}{2}+\frac{b-a}{2}d_j,\ d_j\in[-1,1],\ j=\overline{1,n}$ (вузли можуть бути не рівновіддалені, якщо рівновіддалені, то $d_j=-1+j\frac{2}{n},\ j=\overline{1,n}$).

Замінимо $f(x)$ на інтерполяційний поліном Лагранжа $L_{n-1}(x)$ з вузлами x_j і отримаємо формулу

$$I(f) \approx I_n(f) = \int_a^b L_{n-1}(x) e^{i\omega x} dx, \qquad (5)$$

яка буде точною для всіх поліномів не вище $n-1$- го степеня . Тобто, якщо в (5) підставити поліном Лагранжа, то можна обчислити інтеграл та отримати квадратурну формулу

$$I(f) \approx I_n^{\omega}(f) = \frac{b-a}{2} \exp\left\{ i\omega \frac{a+b}{2} \right\} \sum_{j=1}^{n} D_j\left(\omega \frac{b-a}{2} \right) f(x_j) \ , \qquad (6)$$

де

$$D_j(p) = \int_{-1}^{1} \left(\prod_{\substack{k=1 \\ k \neq j}}^{n} \frac{\xi - d_k}{d_j - d_k} \right) \exp(ip\xi) d\xi, \ j = \overline{1,n} \ .$$

При $n = 3, d_1 = -1, d_2 = 0, d_3 = 1$ – це буде формула Філона. Можна брати і більше точок, наприклад $n = 5, d_1 = -1, d_2 = -\frac{1}{2}, d_3 = 0, d_4 = \frac{1}{2}, d_5 = 1$. Дані формули не слід застосовувати, якщо немає швидко осцилюючого множника.
<u>Зауваження.</u> Формулами Філона часто називають всі формули побудовані на основі розглянутого вище підходу.

Обчислення невласних інтегралів

Розглянемо обчислення інтегралів з такими особливостями :
а) інтеграли другого роду, тобто

$$I = \int_a^b f(x) dx, \quad f(x) \underset{x \to c, c \in [a,b]}{\to} \infty;$$

б) інтеграли першого роду

$$I = \int_a^{\infty} f(x) dx.$$

Розглянемо спочатку інтеграли другого роду, тобто

$$I = \int_a^b f(x) dx, \quad f(x) \underset{x \to c, c \in [a,b]}{\to} \infty.$$

1) <u>Мультиплікативний метод</u>.
Представимо підінтегральну функцію у вигляді $f(x) = \rho(x) g(x)$, причому $\rho(x)$– особлива, а $g(x)$– гладка. Далі для обчислення інтегралу

$$I = \tilde{I}(f) = \int_a^b \rho(x) g(x) dx$$

використовуємо відповідні квадратурні формули Гауса.
<u>Приклад 1</u>. Обчислити інтеграл

$$I=\int_{-1}^{1}\frac{dx}{\sqrt{1-x^4}}.$$

Розв'язання. Точки $x=\pm 1$ є особливими.

Представимо підінтегральну функцію у вигляді

$$f(x)=\underbrace{\frac{1}{\sqrt{1-x^2}}}_{\rho(x)}\underbrace{\frac{1}{\sqrt{1+x^2}}}_{g(x)},$$

тоді

$$I=\int_{-1}^{1}\frac{1}{\sqrt{1+x^2}}\frac{dx}{\sqrt{1-x^2}},$$

Функцію $\rho(x)=\dfrac{1}{\sqrt{1-x^2}}$ будемо розглядати в якості вагової і для обчислення інтеграла використовуємо квадратурну формулу Ерміта.

<u>Приклад 2.</u> Записати алгоритм наближеного обчислення інтеграла

$$I=\int_{0}^{\pi}\ln(\sin x)dx.$$

Розв'язання. Особливі точки $x=0, x=\pi$. Зведемо особливість до степеневої. Для цього подамо підінтегральну функцію у вигляді $f(x)=\rho(x)g(x)$, де $\rho(x)=\dfrac{1}{\sqrt{x(\pi-x)}}$ - вагова функція, а $g(x)=\sqrt{x(\pi-x)}\ln(\sin x)$. Відзначимо, що $g(x)\xrightarrow[x\to 0,\pi]{}0$. Використовуючи заміну $x=\dfrac{\pi}{2}(t+1),\ t=\dfrac{2x}{\pi}-1$ запишемо інтеграл у вигляді

$$I=\int_{0}^{\pi}\ln(\sin x)dx=\frac{\pi}{2}\int_{-1}^{1}\left(\frac{1}{\sqrt{(1-t^2)}}\right)\sqrt{(1-t^2)}\ln\left(\sin(\frac{\pi}{2}(t+1))\right)dt.$$

Для знаходження інтегралу з $\rho(t)=\dfrac{1}{\sqrt{(1-t^2)}}$ застосовуємо квадратурні формули Ерміта.

Однак, слід відзначити, що $g'(x)\xrightarrow[x\to 0,\pi]{}\infty$

($g'(x)=\dfrac{(1-x^2)ctg(x)-x\ln(\sin x)}{\sqrt{1-x^2}}$) (відповідно і $g'(t)\xrightarrow[t\to -1,1]{}\infty$,

($g'(t)=\dfrac{(t^2-1)tg(\frac{\pi t}{2})-2t\ln(\cos(\frac{\pi t}{2}))}{\sqrt{1-t^2}}$) хоча квадратурні формули даватимуть наближене значення). Тому застосовують інший спосіб розв'язання даної задачі, який буде розглянутий далі.

2) <u>Адитивний метод.</u>

Подамо підінтегральну функцію у вигляді

$f(x) = g(x) + \psi(x)$,

причому $g(x)$– має інтегровану особливість, $\psi(x)$– достатньо гладка функція, інтегрована чисельно. Тоді розбиваємо інтеграл на два: $I = I_1 + I_2$.

$I_1 = \int_a^b \psi(x)dx$ – обчислюють чисельно (за допомогою квадратурних формул),

$I_2 = \int_a^b g(x)dx$ – обчислюють аналітично (можливо апроксимувати функцію $g(x)$, наприклад, рядом).

<u>Приклад 3.</u> Записати алгоритм наближеного обчислення інтеграла (див. також приклад 2)

$$I = \int_0^{\pi} \ln(\sin x)dx.$$

Розв'язання. Маємо

$$I = 2\int_0^{\pi/2} \ln(\sin x)dx.$$

Підінтегральну функцію представимо наступним чином

$$\ln(\sin x) = \ln\frac{\sin x}{x} + \ln x.$$

Тоді отримаємо $I = 2(I_1 + I_2)$.

$$I_2 = \int_0^{\pi/2} \ln x dx = (x\ln x - x)\Big|_0^{\pi/2} = \frac{\pi}{2}\left(\ln\frac{\pi}{2} - 1\right),$$

а інтеграл $I_1 = \int_0^{\pi/2} \ln\frac{\sin x}{x}dx$ обчислюємо чисельно.

Зауважимо, що для функції $g(x) = \ln\frac{\sin x}{x}$ маємо $g'(x) = ctg(x) - \frac{1}{x}$ і $g'(x) \xrightarrow[x\to 0]{} 0$, а $g''(x) = \frac{1}{x^2} - ctg^2(x) - 1$ і $g''(x) \xrightarrow[x\to 0]{} -1/3$, отже для обчислення інтеграла I_1 можемо використати складену формулу прямокутників чи трапецій. □

2) <u>Метод виділення особливостей</u> (степеневий метод).

Цей метод базується на адитивному представленні інтеграла, тобто

$f(x) = g(x) + \psi(x)$,

причому $g(x)$– має таку ж особливість, як і функція $f(x)$, а $\psi(x) = f(x) - g(x)$– достатньо гладка функція, інтегрована чисельно. Тоді

розбиваємо інтеграл на два, один з яких обчислюють аналітично, а другий чисельно. Розглянемо застосування методу на наступному прикладі.

<u>Приклад 4.</u> Обчислити наближено інтеграл $I=\int\limits_0^{1/2}\frac{dx}{\sqrt{x(1-x)}}$.

Розв'язання. Функція має інтегровану особливість в т. $x=0$.
Згідно загальної теорії нехай функція $f(x)$ має в т. $x=x_0,\ x\in[a,b]$ інтегровану особливість виду $f(x)=(x-x_0)^{-\alpha}\varphi(x)$, де $0<\alpha<1$ а $\varphi(x)$ - гладка функція. Тоді

$$I=\int\limits_a^b f(x)dx=\int\limits_a^b (x-x_0)^{-\alpha}\varphi(x)dx=\int\limits_a^b (x-x_0)^{-\alpha}\left(c_0+c_1(x-x_0)+c_2(x-x_0)^2+\ldots\right)dx=$$

$$=\int\limits_a^b\left[(x-x_0)^{-\alpha}\left(c_0+c_1(x-x_0)+\ldots+c_n(x-x_0)^n\right)+(x-x_0)^{-\alpha}\left(c_{n+1}(x-x_0)^{n+1}+\ldots\right)\right]dx=$$

$$=\int\limits_a^b g(x)dx+\int\limits_a^b (x-x_0)^{n+1-\alpha}\left(c_{n+1}+c_{n+2}(x-x_0)\ldots\right)dx=\int\limits_a^b g(x)dx+\int\limits_a^b\left(f(x)-g(x)\right)dx=$$

$$=\int\limits_a^b\left(g(x)+\psi(x)\right)dx,$$

де $g(x)=(x-x_0)^{-\alpha}\left(c_0+c_1(x-x_0)+\ldots+c_n(x-x_0)^n\right)$,

$$\psi(x)=(x-x_0)^{-\alpha}\left(c_{n+1}(x-x_0)^{n+1}+\ldots\right)=$$

$$=(x-x_0)^{-\alpha}\varphi(x)-(x-x_0)^{-\alpha}\left(c_0+c_1(x-x_0)+\ldots+c_n(x-x_0)^n\right).$$

(вочевидь $\psi(x_0)=0$)

Інтеграл $\int\limits_a^b g(x)dx$ може бути обчислений аналітично, а для обчислення $\int\limits_a^b\psi(x)dx$ можна застосувати квадратурну формулу, оскільки $\psi(x)$ має n неперервних похідних на $[a,b]$.

Повернемося до роз'язання задачі. Маємо

$$\int\limits_0^{1/2}\frac{dx}{\sqrt{x(1-x)}}=\int\limits_0^{1/2}x^{-1/2}\left(1+\frac{1}{2}x+\frac{3}{8}x^2+\frac{5}{16}x^3+\frac{35}{128}x^4\right)dx+$$

$$+\int\limits_0^{1/2}x^{-1/2}\left((1-x)^{-\frac{1}{2}}-(1+\ldots+\frac{35}{128}x^4)\right)dx=\int\limits_0^{1/2}g(x)dx+\int\limits_0^{1/2}\left(\frac{63}{256}x^{\frac{9}{2}}+\ldots\right)dx=I_1+I_2$$

Обчислюючи аналітично, знайдемо $I_1=1.5692$. I_2 шукаємо за формулою Сімпсона з кроком $h=0.1$, отримаємо $I_2\approx 0.0013$. Тоді

$$I = I_1 + I_2 \approx 1.5705 \approx \frac{\pi}{2}.$$

3) <u>Метод обрізання границь</u>.

Припустимо, що $f(x)$ в точці $c \in [a,b]$ має інтегровану особливість (наприклад, в неї не існує границі в цій точці). Подамо інтеграл у вигляді

$$I = \int_a^b f(x)dx = \int_a^{c-\delta_1} f(x)dx + \int_{c-\delta_1}^{c+\delta_2} f(x)dx + \int_{c+\delta_2}^{b} f(x)dx.$$

Вибиремо величини δ_1 та δ_2 настільки малими, щоб

$\left|\int_{c-\delta_1}^{c+\delta_2} f(x)dx\right| < \frac{\varepsilon}{3}$ де ε - достатньо мала величина, а інтеграли $\int_a^{c-\delta_1} f(x)dx$ та $\int_{c+\delta_2}^{b} f(x)dx$ обчислюємо чисельно з точністю $\frac{\varepsilon}{3}$. Тоді отримаємо значення

$\mathrm{I} = \int_a^b f(x)dx \approx \int_a^{c-\delta_1} f(x)dx + \int_{c+\delta_2}^{b} f(x)dx$ з точністю ε.

Розглянемо тепер інтеграли першого роду

$$I = \int_a^{\infty} f(x)dx.$$

1) <u>Методи заміни змінної інтегрування</u>.

Нехай $a>0$.

а) Зробимо заміну $t = \frac{x-a}{x}$, $x = \frac{a}{1-t}$. Тоді

$$I = a\int_0^1 f\left(\frac{a}{1-t}\right)\frac{dt}{(1-t)^2},$$

а це інтеграл другого роду.

б) Можна використати також заміну $t = \frac{a}{x}$, $x = \frac{a}{t}$. Тоді отримаємо інтеграл другого роду

$$I = a\int_0^1 f\left(\frac{a}{t}\right)\frac{dt}{t^2}.$$

Нехай a=0.

а) зробимо заміну $t = e^{-x}$, $x = -\ln t$, тоді

$$I = \int_0^1 f(-\ln t)\frac{dt}{t},$$

тобто отримали інтеграл другого роду.

б) за допомогою заміни $t=\dfrac{x}{x+1}$, $x=\dfrac{t}{1-t}$ знову приходимо до інтегралу другого роду

$$I=\int_0^1 f\left(\frac{t}{1-t}\right)\frac{dt}{(1-t)^2}.$$

Нехай $a<0$ (в цьому випадку не можна робити заміну $t=\dfrac{x-a}{x}$, оскільки виникає особливість в точці $x=0$).

Тоді розбиваємо інтеграл на два

$$I=\int_a^0 f(x)dx+\int_0^\infty f(x)dx.$$

і обчислюємо за допомогою попередніх пунктів.

Розглянемо тепер інтеграл виду

$$I=\int_{-\infty}^{\infty} f(x)dx.$$

Записавши його у вигляді $I=\int_{-\infty}^{a} f(x)dx+\int_a^\infty f(x)dx$, можна використати один з попереднії способів його обчислення. Окрім того, за допомогою заміни $t=\dfrac{e^x-1}{e^x+1}$, $x=-\ln\dfrac{1+t}{1-t}$ шуканий інтеграл можна звести до інтегралу другого роду виду

$$I=2\int_{-1}^{1} f\left(\ln\frac{1+t}{1-t}\right)\frac{dt}{(1-t)^2}.$$

2) <u>Мультиплікативний спосіб</u> обчислення інтегралів першого роду ґрунтується на поданні підінтегральної функції у вигляді

$$f(x)=\rho(x)g(x),$$

де, наприклад,

$$\rho(x)=x^\alpha e^{-x},\ x\in[0,\infty).$$

Такий ваговий коефіцієнт відповідає поліномам Лагера. При $x\in(-\infty,\infty)$, $\rho(x)=e^{-x^2}$ приходимо до поліномів Ерміта.

3) <u>Метод обрізанні верхньої границі</u>.

Запишемо інтеграл у вигляді

$$I=\int_a^\infty f(x)dx=\int_a^A f(x)dx+\int_A^\infty f(x)dx.$$

Верхню границю A обчислюють виходячи з вимоги

$$\left|\int_A^\infty f(x)dx\right| < \frac{\varepsilon}{2},$$

де ε – задана точність обчислення інтеграла. Для обчислення інтеграла $\int_a^A f(x)dx$ з точністю $\frac{\varepsilon}{2}$ використовують квадратурні формули складеного типу.

<u>Приклад 1.</u> Записати алгорим обчислення невласного інтеграла $I = \int_0^\infty e^{-x^2}dx$ з точністю до 10^{-4}.

Розв'язання. Маємо

$$(x-A)^2 \ge 0;$$

$$x^2 - 2Ax + A^2 \ge 0;$$

$$-x^2 \le -2Ax + A^2.$$

Тоді

$$\int_A^\infty e^{-x^2}dx \le e^{A^2}\int_A^\infty e^{-2Ax}dx = \frac{1}{2A}e^{-A^2}.$$

Виберемо величину A так, щоб

$\int_A^\infty e^{-x^2}dx \le \frac{\varepsilon}{2} = \frac{1}{2}\cdot 10^{-4}$, тобто $\frac{e^{-A^2}}{A} < 10^{-4}$.

При $A = 3$ маємо

$$\int_3^\infty e^{-x^2}dx \le \frac{e^{-9}}{2\cdot 3} < \frac{1}{2}\cdot 10^{-4}.$$

Звідси

$$\int_0^\infty e^{-x^2}dx \le \int_0^3 e^{-x^2}dx + \frac{1}{2}\cdot 10^{-4}.$$

Перший інтеграл в правій частині останнього виразу можна обчислити за квадратурними формулами з точністю $\frac{1}{2}\cdot 10^{-4}$. Тоді

$\int_0^\infty e^{-x^2}dx \approx \int_0^3 e^{-x^2}dx$ з точністю 10^{-4}.

Обчислення кратних інтегралів

Розглянемо задачу наближеного обчислення інтегралу

$$I = \int_a^b \int_c^d f(x,y)dxdy \qquad (1)$$

Введемо позначення

$$G(x) = \int_c^d f(x,y)\,dy \qquad (2)$$

Тоді інтеграл (1) зводиться до повторного

$$I = \int_a^b G(x)dx. \qquad (3)$$

До однократних інтегралів можна застосувати, наприклад, квадратурну формулу середніх прямокутників. Тоді

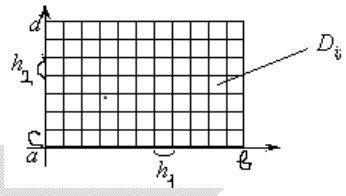
$$I \approx I_0 = (b-a)G(\bar{x}) = (b-a)\int_c^d f(\bar{x},y)dy \approx (b-a)(d-c)f(\bar{x},\bar{y}), \qquad (4)$$

де $\bar{x} = \dfrac{a+b}{2}$, $\bar{y} = \dfrac{c+d}{2}$.

Кубатурна формула (формули обчислення кратних інтегралів) у випадку, якщо використовується формула трапеції має вигляд:

$$I \approx I_1 = \frac{(b-a)(d-c)}{4}\big(f(a,c) + f(b,c) + f(a,d) + f(b,a)\big).$$

Точність залежить від поведінки функції та від довжини інтервалів $[a,b]$, $[c,d]$. Аналог формул складеного типу для (1) будується таким чином: розбиваємо D на комірки:



$$D = \bigcup_{i,j} D_{ij} \quad D_{ij} = \{x_{i-1} \le x \le x_i, y_{j-1} \le y \le y_j\}, x_i = a + ih_1,\ i = \overline{0,N_1},$$
$$h_1 = (b-a)/N_1;\ \ y_j = c + jh_2,\ j = \overline{0,N_2},\ h_2 = (d-c)/N_2.$$

Тоді для кожного інтегралу по комірці застосовуємо кубатурну формулу прямокутників (4):

$$I = \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\iint_{D_{ij}} f(x,y)dxdy \approx I_{0,h} = \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} h_1 h_2 f(\bar{x}_i, \bar{y}_j),$$

де $\bar{x}_i = x_i - \dfrac{h_1}{2}$, $\bar{y}_j = y_j - \dfrac{h_2}{2}$.

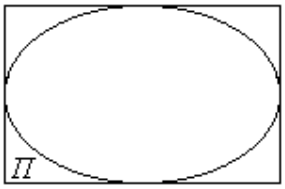
Якщо $f(x,y) \subset C^2(\bar{D})$, то $I - I_{0,h} = O(h_1^2 + h_2^2)$. Степінь точності цієї формули по кроках дорівнює двом. Складність методу пропорційна кількості комірок, тобто

$Q = O(N_1 \cdot N_2) = O(N^2)$,

при $N \approx N_1 \approx N_2$. В 3-вимірному просторі $f = f(x, y, z)$ складність буде $Q = O(N^3)$.

Якщо D не прямокутник, то замість f розглянемо функцію

$$\overline{f}(x, y) = \begin{cases} f(x, y), & x \in D, \\ 0, & x \in \Pi \setminus D. \end{cases}$$



де Π – найменший прямокутник, що містить D, тобто $D \subset \Pi$. Тоді $I = \iint\limits_{\Pi} \overline{f}(x, y)dxdy$.

Недолік такого підходу полягає в тому, що $\overline{f} = (x, y)$ може бути розривною функцією і тому буде низькою точність обчислення інтегралу.

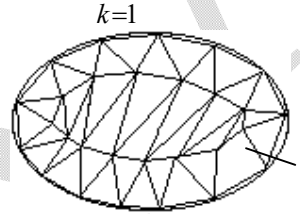
Наступний підхід базується на відповідній заміні змінних

$$x = x(\xi, \eta) \quad y = y(\xi, \eta) \quad D \to \Pi,$$

$$I = \iint\limits_{\Pi} f(x(\xi, \eta), y(\xi, \eta))J(\xi, \eta)d\xi d\eta,$$

де Π - прямокутник в площині (ξ, η); $J(\xi, \eta)$ – Якобіан переходу. Якщо область D гладка, то якобіан буде мати особливості, що також точність інтегрування.

Ще один підхід до обчислення інтегралу по довільній області D базується на триангуляції області. Якщо область довільного вигляду, то її можна розбити на трикутники $D = \bigcup\limits_{k=1}^{N} D_k$:



Тоді $I = \sum\limits_{k=1}^{N} I_k$, $I_k = \iint\limits_{D_k} f(x, y)dxdy$. Застосуємо кубатурні формули до кожного $I_k \approx I_k^h$. Для цього замінимо функцію поліномом першого степеня $f(x, y) \cong L_1(x, y) = A + Bx + Cy$.

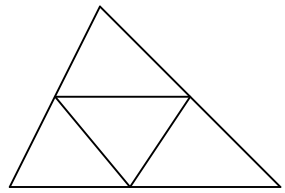
<u>Задача</u>. Побудувати явний вигляд кубатурної формули, яка дозволяє наближено обчислити I_k по трикутнику D_k, якщо замінити $f(x, y) \approx L_1(x, y)$ на інтерполяційний поліном 1-го степеня. □

Точність такого підходу

$$I - I^h = I - \sum_{k=1}^{N} I_k^h = O(h^2),$$

де $h = \max\limits_k diamD_k$.

Складність формули $Q = O(h^{-2})$ для $D \subset R_2$; $Q = O(h^{-3})$ $D \subset R_3$.



Можна згустити сітку, поділивши один трикутник D_k на чотири:

Розглянемо тепер ідею методу статистичних випробувань (метод Монте-Карло) для обчислення інтегралу $I = \iint\limits_D f(x,y)dxdy$. Замінимо наближено

$$I = \iint\limits_{\Pi} \overline{f}(x,y)dxdy \approx \frac{1}{N}\sum_{i=1}^{N} \overline{f}(\xi_i,\eta_i)\cdot mes\Pi = I_N,$$

де Π – найменший прямокутник, який містить D $D \subset \Pi$; $\overline{f}(x,y)$ - продовження функції (4); ξ_i, η_i - незалежні реалізації рівномірно розподілених на $[a,b]$ та $[c,d]$ випадкових величин ξ та η. Складність цього методу $Q = O(N)$. Оцінка точності $I - I_N = O(N^{-\frac{1}{2}})$ носить ймовірностний характер. Позитивною характеристикою методу є незалежна від розмірності складність, але він має досить низьку точність.

Наближені методи розв'язання задачі Коші для звичайних диференціальних рівнянь (ЗДР).

Розглянемо задачу Коші

$$\frac{d\overline{u}}{dt} = \overline{f}(t,\overline{u}), \qquad t \ge t_0, \qquad \overline{u}(t_0) = \overline{u}_0, \tag{1}$$

де $\overline{u} = (u_1, \dots, u_n)^T$, $\overline{f} = (f_1, \dots, f_n)^T$, $u_k = u_k(t)$, $f_k = f_k(t, u_1, \dots, u_n)$.

Або в розгорнутому вигляді

$$\begin{cases} \dfrac{du_1}{dt} = f_1(t, u_1, \dots, u_n), \\ \dfrac{du_2}{dt} = f_2(t, u_1, \dots, u_n), \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ \dfrac{du_n}{dt} = f_1(t, u_1, \dots, u_n), \\ u_1(t_0) = u_{01}, u_2(t_0) = u_{02}, \dots, u_n(t_0) = u_{0n}. \end{cases}$$

Якщо f_k неперервні по всіх аргументах, то розв'язок задачі (1) існує, якщо окрім того праві частини (1) задовольняють умовам Ліпшиця

$$\left| f_k(t, u_1, \dots, u_j, \dots, u_n) - f_k(t, u_1, \dots, v_j, \dots, u_n) \right| \le L_j \left| u_j - v_j \right|, \; 1 \le k, j \le n,$$

то розв'язок задачі (1) існує та єдиний.

До вигляду (1) зводяться задачі Коші для диференціальних рівнянь n-го порядку виду:

$$\begin{aligned} & v^{(n)} = F(t, v, v', \dots, v^{(n-1)}), \\ & v(t_0) = v_1, \dots, v^{(n-1)}(t_0) = v_n. \end{aligned} \tag{2}$$

Для цього введемо заміну

$$u_k = v^{(k-1)}, \; k = \overline{1, n},$$

тобто $u_1 = v, \; u_2 = v', \; u_3 = v'', \dots, \; u_n = v^{(n-1)}$.

Тоді задачу (2) можна записати у вигляді

$$\begin{cases} \dfrac{du_1}{dt} = u_2, \\ \dfrac{du_2}{dt} = u_3, \\ \quad \vdots \\ \dfrac{du_{n-1}}{dt} = u_n, \\ \dfrac{du_n}{dt} = F(t, u_1, \dots, u_n), \\ u_1(t_0) = v_1, u_2(t_0) = v_2, \dots, u_n(t_0) = v_n. \end{cases} \tag{3}$$

Тобто (3) має вигляд (1) з

$$\bar{f}(t,\bar{u}) = (u_2, u_3, \dots, u_n, F(t, u_1, \dots, u_n))^T, \;\; \bar{u}_0 = (v_1, \dots, v_n)^T.$$

Більшість алгоритмів розв'язання ЗК для одного рівняння можна застосувати для отримання методів розв'язання ЗК для систем диференціальних рівнянь. Тому надалі будемо розглядати методи розв'язання ЗК для одного рівняння.

Методи наближеного розв'язування ЗК можна розділити на дві групи:

1) наближені аналітичні (коли розв'язок подається у вигляді наближеної формули) і 2) чисельні методи (розв'язок подається у вигляді своїх наближених значень (чисел) у деяких точках (вузлах)).

<u>**Наближені аналітичні методи розв'язання ЗК.**</u>

а) <u>Метод послідовних наближень.</u>

Розглянемо задачу Коші

$$\begin{aligned} &\frac{du}{dt} = f(t,u), \\ &u(t_0) = u_0. \end{aligned} \tag{4}$$

Проінтегруємо рівняння (4) по проміжку $[t_0, t]$:

$$u(t) = u(t_0) + \int_{t_0}^{t} f(\xi, u(\xi))d\xi \tag{5}$$

Задамо початкове наближення $\overset{(0)}{u}(t)$ (наприклад, $\overset{(0)}{u}(t) = u_0$) та запишемо наступний ітераційний процес:

$$\overset{(k)}{u}(t) = u_0 + \int_{t_0}^{t} f(\xi, \overset{(k-1)}{u}(\xi))d\xi, \; k = 0,1,... \tag{6}$$

З теорії диференціальних рівнянь відомо, що $\exists T = T(L, u_0)$ (де $L = \max_j L_j$ - стала Ліпшиця), що $\overset{(k)}{u}(t) \underset{k\to\infty}{\to} u(t), \; t \in [t_0, T]$.

До недоліків методу можна віднести необхідність аналітичного інтегрування.

б) <u>Метод розкладання в ряд Тейлора.</u>

Нехай функція u - аналітична. Тоді її можна представити у вигляді ряду Тейлора

$$u(t)=\sum_{k=0}^{\infty}\frac{u^{(k)}(t_0)}{k!}(t-t_0)^k\ . \qquad (7)$$

Обмежимося скінченою кількістю членів ряду Тейлора:

$$u(t)\approx u_m(t)=\sum_{k=0}^{m}\frac{u^{(k)}(t_0)}{k!}(t-t_0)^k\ . \qquad (8)$$

Для того, щоб скористатися формулою (8), необхідно обчислити похідні до m -го порядку. З рівняння (4) маємо:

$$u'(t_0)=f(t_0,u_0)\,.$$

Далі диференціюємо рівняння по t:

$$u''(t_0)=\left.\frac{d^2u}{dt^2}\right|_{t=t_0}=\left.\frac{d}{dt}f(t,u)\right|_{t=t_0}=\left.\left(f_t'(t,u)+f_u'(t,u)f(t,u)\right)\right|_{t=t_0}=$$

$$=f_t'(t_0,u_0)+f_u'(t_0,u_0)f(t_0,u_0),$$

$$u'''(t_0)=f''_{tt,0}+2f''_{tu,0}f_0+f''_{uu,0}f_0^{\,2}+f'_{u,0}(f'_{t,0}+f'_{u,0}f_0)$$

і т.д.

Обчисливши похідні, знайдемо наближення (8). Якщо величина $\tau=t_1-t_0$ достатньо мала, то

$$\left|u(t)-u_m(t)\right|=O(\tau^{m+1})\,.$$

У випадку коли ряд (8) збігається тільки при $t_0\le t\le t_1$, тоді при необхідності обчислити розв'язок для $t>t_1$ використовуємо ряд (8), де t_0 замінюється на t_1 і т.д. На кожному проміжку $[t_n,\,t_{n+1}]$ маємо ряд із своїми коефіцієнтами.

До недоліків методу можна віднести зростання кількості доданків при обчисленні $u^{(k)}(t_0)$ та необхідність аналітичного диференціювання.

<u>**Чисельні методи розв'язання ЗК**</u>

Методи типу Ейлера

Розглянемо задачу Коші:

$$\frac{du}{dt} = f(t,u),\ u(t_0) = u_0 .$$

Розглянемо формулу (8) на $\forall$ проміжку $[t_n, t_{n+1}]$ при $m = 1$:

$$u(t) \approx u_n + u'(t_n)(t - t_n) . \tag{9}$$

Покладемо $t = t_{n+1}$, тоді отримаємо наближену формулу

$$y_{n+1} = y_n + \tau f(t_n,\ y_n) \tag{10}$$

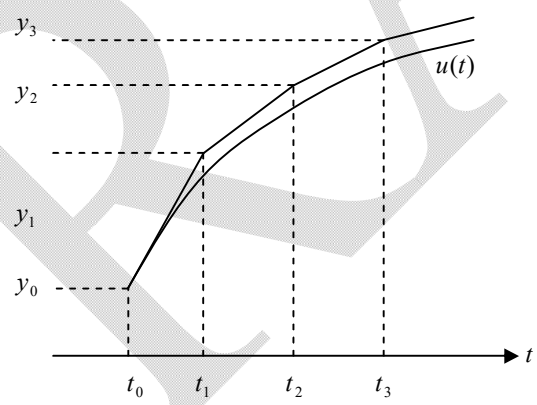
де $\tau = t_{n+1} - t_n$.

Це формула **явного методу Ейлера**. Метод належить до однокрокових методів, тобто розв'язок на наступному кроці обчислюється тільки за значеннями на одному попередньому кроці. Загальна формула однокрокових методів має вигляд

$$y_{n+1} = y_n + \tau\,\Phi(t_n, y_n) \tag{12}$$

Для методу Ейлера маємо $\Phi = f(t_n, y_n)$.

Геометрично метод Ейлера можна розглядати, як метод ламаних



Зауважимо, що крок τ може бути змінним, тобто $\tau_n = t_{n+1} - t_n$.

<u>Визначення</u>. Величина

$$R_n(\tau) = y_{n+1} - u(t_{n+1}) ,$$

де

$$y_{n+1} = \widetilde{\Phi}(t_n,\ y_n,\ f),\quad y_n = u(t_n) \tag{14}$$

називається **похибкою методу на кроці**, або локальною похибкою методу. □

Визначення: Метод має **похибку на кроці степеня (порядку)** m**, якщо**

$$\left|R_n(\tau)\right| = O(\tau^{m+1}).$$

Загальна похибка на $(n+1)$-му кроці складається з похибок на всіх попередніх кроках.

Визначення: Величину $z_{n+1} = y_{n+1} - u(t_{n+1})$ де y_{n+1} і всі попередні y_k, $k = 1,2,...,n$ також наближені, тобто обчислюються за формулами

$$y_{k+1} = \widetilde{\Phi}(t_k,\, y_k,\, f),\ k = 0,1,...,n\ ,\ \ y_0 = u(t_0)$$

називають **повною похибкою методу** (або глобальною похибкою). □

В даному випадку точний та наближений розв'язки співпадають тіільки в точці t_0 .

Нехай метод записаний у вигляді (12).

Визначення. Величина $\psi_n(\tau) = \dfrac{u_{n+1} - u_n}{\tau} - \Phi(t_n, u_n)$ називається **похибкою апроксимації методу**.□

Похибка апроксимації це нев'язка, коли замість y в різницеве рівняння підставляємо точний розв'язок задачі Коші u.

Для явного методу Ейлера

$$\psi_n(\tau) = \frac{u_{n+1} - u_n}{\tau} - f(t_n, u_n) = \frac{1}{\tau}\left(u_n + \tau\frac{du(t_n)}{dt} + O(\tau^2) - u_n\right) - f(t_n, u_n) =$$

$$= \frac{du(t_n)}{dt} - f(t_n, u_n) + O(\tau) = O(\tau) \xrightarrow[\tau \to 0]{} 0.$$

Визначення. Чисельний метод має **похибку апроксимації степеня (порядку)** p**,** якщо $\left|\psi_n(\tau)\right| = O(\tau^p)$. □

Отже порядок похибки апроксимації явного методу Ейлера дорівнює одиниці.

Для похибки на кроці методу Ейлера будемо мати

$$R_n(\tau) = y_{n+1} - u(t_{n+1}) = y_n + \tau\, f(t_n, y_n) - u(t_n + \tau) =$$
$$= u(t_n) + \tau\, f(t_n, u(t_n)) - u(t_n + \tau) = u(t_n) + \tau\, f(t_n, u(t_n)) -$$
$$- u(t_n) - \tau\, u'(t_n) + O(\tau^2) = O(\tau^2)$$

Отже $m = 1$. Маємо зв'язок похибок $\psi_n(\tau) = \frac{1}{\tau} R_n(\tau)$.

<u>Визначення</u>. Метод має **степінь точності** m (на всьому проміжку, глобальний), якщо

$$z_n = y_n - u(t_n) = O(\tau^m). \ \square$$

<u>Теорема</u>. Нехай $f(t,u) \in C^1(\overline{\Omega}_T)$, та $|f_u(t,u)| \le L$, де $\overline{\Omega}_T = \{t_0 < t \le T,\ |u(t)| \le K\}$. Тоді метод Ейлера (10) має степінь точності $m = 1$ і для нього має місце оцінка

$$|z_n| \le M \max_j |\psi_j| = O(\tau),$$

де $M = M(T, L)$.

Доведення. Підставляючи $y_n = z_n + u_n$ в (10), для z_n отримаємо задачу

$$z_{n+1} = z_n + \tau\psi_n + \tau[f(t_n, u_n + z_n) - f(t_n, u_n)].$$

Оцінимо z_{n+1}:

$$|z_{n+1}| \le |z_n| + \tau|\psi_n| + \tau L|z_n| \le (1 + \tau L)|z_n| + \tau|\psi_n| \le$$
$$\ldots \le (1 + \tau L)^n |z_0| + \sum_{j=0}^{n} \tau (1 + \tau L)^{n-j} |\psi_j| \le$$
$$\le (1 + \tau L)^n \max_j |\psi_j| \cdot \sum_{j=0}^{n} \tau \le T e^{LT} \max_j |\psi_j| \le M \max_j |\psi_j| = O(\tau),$$

оскільки $(1 + \tau L)^n = (1 + \tau L)^{\frac{1}{\tau} n\tau} = (1 + \tau L)^{\frac{1}{\tau} t_n} \le (1 + \tau L)^{\frac{1}{\tau} T} \le e^{LT}$. Позначивши $Te^{LT} = M$, отримаємо твердження теореми. □

Таким чином порядок точності методу Ейлера $m = 1$.

Методи типу Ейлера можна вивести з наступних міркувань:
Інтегруємо (4) по t, $t_n \le t \le t_{n+1}$:

$$u(t_{n+1}) = u(t_n) + \int_{t_n}^{t_{n+1}} f(t, u(t))dt. \qquad (15)$$

Застосуємо формулу лівих прямокутників для наближення інтеграла

$$\int_{t_n}^{t_{n+1}} f(t, u(t))dt \approx \tau f(t_n, u(t_n)).$$

Тоді з (15) отримаємо формулу **явного методу Ейлера**

$$y_{n+1} = y_n + \tau f(t_n, y_n).$$

Застосуємо формулу правих прямокутників

$$\int_{t_n}^{t_{n+1}} f(t,u(t))dt \approx \tau f(t_{n+1},u(t_{n+1})),$$

тоді отримаємо формулу **неявного методу Ейлера**

$$y_{n+1} = y_n + \tau f(t_{n+1}, y_{n+1}). \tag{16}$$

Ці формули 1-го степеня точності по кроку τ.

Замінюючи інтеграл квадратурною формулою трапеції

$$\int_{t_n}^{t_{n+1}} f(t,u(t))dt \approx \frac{\tau}{2}\left(f(t_n,u(t_n)) + f(t_{n+1},u(t_{n+1}))\right),$$

отримаємо формулу **методу трапецій** інтегрування задачі Коші

$$y_{n+1} = y_n + \frac{\tau}{2}\left(f(t_n, y_n) + f(t_{n+1}, y_{n+1})\right). \tag{17}$$

Це **неявний метод.**

Обчислимо похибку апроксимації цього методу

$$\psi_n = \frac{u_{n+1} - u_n}{\tau} - \frac{1}{2}\left(f(t_n,u_n) + f(t_{n+1},u_{n+1})\right) =$$

$$= \frac{1}{\tau}\left(u_{n+1/2} + \frac{\tau}{2}\overset{\circ}{u}_{n+1/2} + \frac{1}{2}\left(\frac{\tau}{2}\right)^2 \overset{\circ\circ}{u}_{n+1/2} - u_{n+1/2} + \frac{\tau}{2}\overset{\circ}{u}_{n+1/2} - \frac{1}{2}\left(\frac{\tau}{2}\right)^2 \overset{\circ\circ}{u}_{n+1/2} + O(\tau^3)\right) -$$

$$-\frac{1}{2}\left(f_{n+1/2} - \frac{\tau}{2}\overset{\circ}{f}_{n+1/2} + O(\tau^2) + f_{n+1/2} + \frac{\tau}{2}\overset{\circ}{f}_{n+1/2} + O(\tau^2)\right) =$$

$$= \overset{\circ}{u}_{n+1/2} + O(\tau^2) - f_{n+1/2} = O(\tau^2).$$

(використані наступні позначення $\overset{\circ}{u} = \frac{d}{dt}u(t)$, $\overset{\circ}{f} = \frac{d}{dt}f(t,u) = f_t' + f_u' f$).

Таким чином метод трапецій має другий порядок апроксимації, отже отримали більш точний метод. Недолік методу трапецій при реалізації – це те, що він неявний (потрібно розв'язувати нелінійне рівняння на кожному кроці). Для реалізації неявних методів застосовують алгоритми методів типу предиктор - коректор (PC – методи , де P- предиктор - прогнозоване значення, C- коректор – уточнене значення).

Обчислимо попередньо за явним методом Ейлера

$$\overline{y}_{n+1} = y_n + \tau f(t_n, y_n). \tag{18}$$

Уточнимо отримане значення за методом трапецій

$$y_{n+1} = y_n + \frac{\tau}{2}\left(f(t_n, y_n) + f(t_{n+1}, \overline{y}_{n+1})\right). \tag{19}$$

Формули (18) та (19) утворюють **метод Ейлера-Коші**.

Приклад. Визначити степінь точності на кроці методу Ейлера-Коші

$$\bar{y}_{n+1} = y_n + \tau f(t_n, y_n),$$

$$y_{n+1} = y_n + \frac{\tau}{2}\left(f(t_n, y_n) + f(t_{n+1}, \bar{y}_{n+1})\right).$$

Розв'язання. За визначенням похибка методу на кроці визначається формулою

$$R_n(\tau) = y_{n+1} - u(t_{n+1}) = y_n + \frac{\tau}{2}\left(f(t_n, y_n) + f(t_{n+1}, \bar{y}_{n+1})\right) - u(t_{n+1}) =$$

$$= u_n + \frac{\tau}{2}\left(f(t_n, u_n) + f(t_{n+1}, u_n + \tau f(t_n, u_n))\right) - u(t_{n+1}) =$$

Далі враховуючи (4) маємо:

$$A_1 = \frac{\tau}{2}\left(f(t_n, u_n) + f(t_{n+1}, u_n + \tau f(t_n, u_n))\right) =$$

$$= \frac{\tau}{2}\left(2f(t_n, u_n) + \tau f_t'(t_n, u_n) + \tau f(t_n, u_n) f_u'(t_n, u_n) + O(\tau^2)\right) =$$

$$= \tau f(t_n, u_n) + \frac{\tau^2}{2} f_t'(t_n, u_n) + \frac{\tau^2}{2} f(t_n, u_n) f_u'(t_n, u_n) + O(\tau^3)$$

Крім того

$$A_2 = u(t_{n+1}) - u(t_n) = \tau u'(t_n) + \frac{\tau^2}{2} u''(t_n) + O(\tau^3) =$$

$$= \tau f(x_n, u_n) + \frac{\tau^2}{2}(f_t'(t_n, u_n) + f(t_n, u_n) f_u'(t_n, u_n)) + O(\tau^3).$$

Тоді $R_n(\tau) = A_1 - A_2 = O(\tau^3)$, тобто степінь точності на кроці методу Ейлера-Коші $m = 2$. □

Застосуємо тепер в (15) формулу середніх прямокутників для наближеного обчислення інтегралу

$$\int_{t_n}^{t_{n+1}} f(t, u(t))dt \approx \tau f(t_{n+1/2}, u(t_{n+1/2})).$$

Тоді отримаємо формулу методу середніх прямокутників

$$y_{n+1} = y_n + \tau f(t_{n+1/2}, y_{n+1/2})$$

для реалізації якого застосуємо наступний алгоритм

$$y_{n+1/2} = y_n + \frac{\tau}{2} f(t_n, y_n), \tag{20}$$

$$y_{n+1} = y_n + \tau f(t_{n+1/2}, y_{n+1/2}). \tag{21}$$

Це **модифікований (або удосконалений) метод Ейлера**. На етапі предиктора (20) використовуємо явний метод Ейлера на відрізку $[t_n, t_{n+1/2}]$, а коректора (21) - формулу середніх прямокутників.

Задача. Показати, що для модифікованого методу Ейлера $R_n(\tau) = O(\tau^3)$, тобто $m = 2$.
Відзначимо,що методи типу Ейлера мають досить низьку точність.

Методи Рунге-Кутти .

Розглянемо задачу Коші

$$\frac{du}{dt} = f(t,u),\ u(t_0) = u_0 . \qquad (1)$$

Розглянемо побудову явного m**- стадійного методу Рунге-Кутти**. Нехай розв'язок $y_n = y(t_n)$ вже відомий. Будемо шукати наближений розв'язок задачі (1) у вигляді

$$y_{n+1} = y_n + \sum_{i=1}^{m} p_i k_i , \qquad (2)$$

де

$k_i = k_i(\tau) = \tau\, f(\xi_i, \eta_i)$,

$\xi_i = t_n + \alpha_i \tau,\ \ \alpha_1 = 0$,

$\eta_i = y_n + \beta_{i1} k_1 + \beta_{i2} k_2 + ... + \beta_{i,i-1} k_{i-1},\ i = \overline{1,m}$,

$p_i, \alpha_i, \beta_{ij}$ $(0 < j < i \le m)$ - сталі коефіцієнти.

Функція похибки на одному кроці має наступний вигляд

$$R(\tau) = y_{n+1} - u(t_n + \tau) = u_n + \sum_{i=1}^{m} p_i k_i(\tau) - u(t_n + \tau) =$$

$$= R(0) + \tau R'(0) + ... + \frac{\tau^p}{p!} R^{(p)}(0) + O(\tau^{p+1})$$

Коефіцієнти $p_i, \alpha_i, \beta_{ij}$ вибираються таким чином, щоб

$R(0) = R'(0) = ... = R^{(p)}(0) = 0$, тоді $R(\tau) = O(\tau^{p+1})$,

а отже метод буде мати p - й степінь точності на кроці.

Інформацію про m - стадійний метод можна записати у вигляді таблиці Батчера:

$$\begin{array}{c|cccc}
\alpha_1 = 0 & & & & \\
\alpha_2 & \beta_{21} & & & \\
\vdots & \vdots & & & \\
\alpha_m & \beta_{m1} & \cdots & \beta_{m\,m-1} & \\
\hline
 & p_1 & \cdots & p_{m-1} & p_m
\end{array}$$

Розглянемо частинні випадки .

1. Нехай $m=1$ (одностадійний метод).

Тоді з (2) будемо мати

$$y_{n+1}=y_n+p_1k_1(\tau).$$

p_1 - невідома величина. Маємо

$k_1(\tau)=\tau f(t_n,y_n)$, $R(\tau)=u_n+p_1k_1(\tau)-u(t_n+\tau)$, $R(0)=u_n-u(t_n)=0$,

$R'(\tau)=p_1k_1'(\tau)-u'(t_n+\tau)=p_1f(t_n,y_n)-u'(t_n+\tau)$,

$R'(0)=(p_1-1)f_n=0\Rightarrow p_1=1$.

Оскільки $R''(\tau)=u''(t_n+\tau)$, то $R''(0)\neq 0$. Тоді $R(\tau)=O(\tau^2)$. Таким чином отримали одностадійний метод Рунге-Кутти

$$y_{n+1}=y_n+\tau\, f(t_n,y_n).$$

Бачимо, що це явний метод Ейлера

2. Нехай $m=2$ (двохстадійні методи).

З формули (2) будемо мати

$$y_{n+1}=y_n+p_1k_1(\tau)+p_2k_2(\tau).$$

Тоді

$$R(\tau)=y_{n+1}-u(t_{n+1})=u_n+p_1k_1(\tau)+p_2k_2(\tau)-u(t_n+\tau)=$$
$$=u_n+p_1\tau\, f(t_n,u_n)+p_2\tau\, f(t_n+\alpha_2\tau,u_n+\beta_{21}\tau\, f(t_n,u_n))-u(t_n+\tau).$$

З умов

$R(0)=u_n-u(t_n)=0$,

$R'(0)=p_1k_1'(0)+p_2k_2'(0)-u'(t_n)=(p_1+p_2-1)f_n=0$,

$R''(0)=p_1k_1''(0)+p_2k_2''(0)-u''(t_n)=(2p_2\alpha_2-1)\dfrac{\partial f_n}{\partial\tau}+(2p_2\beta_{21}-1)\dfrac{\partial f_n}{\partial u}u_n'=0$

отримуємо систему трьох рівнянь з чотирма невідомими $p_1,p_2,\alpha_2,\beta_{21}$:

$$\begin{aligned} &p_1+p_2=1,\\ &2\alpha_2p_2=1,\\ &2\beta_{21}p_2=1\ . \end{aligned} \tag{3}$$

Оскільки $R(\tau)=O(\tau^3)$, то $p=2$ і отримуємо методи другого порядку точності.

З другого та третього рівняння (3) маємо $\alpha_2=\beta_{21}$, тоді

$$\begin{aligned} &p_1+p_2=1,\\ &2\alpha_2p_2=1,\\ &\beta_{21}=\alpha_2. \end{aligned} \tag{4}$$

Фіксуючи один з параметрів в системі (4), отримуємо конкретний метод.

а) Покладемо $\alpha_2 = \frac{1}{2}$, тоді з системи (4) отримаємо $\beta_{21} = \frac{1}{2},\ p_1 = 0,\ p_2 = 1$.

Отже

$$y(t_{n+1}) = y_n + k_2 = y_n + \tau f(t_n + \frac{\tau}{2},\ y_n + \frac{k_1}{2}) = y_n + \tau f(t_n + \frac{\tau}{2},\ y_n + \frac{\tau}{2} f(t_n, y_n))$$

Тобто отримали модифікований метод Ейлера

$$y_{n+1} = y_n + \tau f\left(t_{n+1/2}, y_n + \frac{\tau}{2} f_n \right).$$

б) Покладемо $\alpha_2 = 1$, тоді з системи (4) отримаємо $\beta_{21} = 1,\ p_1 = \frac{1}{2},\ p_2 = \frac{1}{2}$, а отже

$$y(t_{n+1}) = y_n + \frac{1}{2}(k_1 + k_2),\ k_1 = \tau\, f(t_n,\ y_n),\ k_2 = \tau\, f(t_n + \tau,\ y_n + k_1)\ .$$

Звідси

$$y_{n+1} = y_n + \frac{\tau}{2}\left(f\left(t_n, y_n\right) + f\left(t_{n+1}, y_n + \tau f(t_n, y_n)\right)\right)$$

Це формула методу Ейлера-Коші .

в) Покладемо $\alpha_2 = \frac{3}{4}$, тоді $\beta_{21} = \frac{3}{4}, p_1 = \frac{1}{3}, p_2 = \frac{2}{3}$ і отримаємо ще один метод другого порядку точності:

$$y_{n+1} = y_n + \frac{\tau}{3}\left(f\left(t_n, y_n\right) + 2 f\left(t_{n+3/4}, y_n + \frac{3}{4} f_n \right)\right).$$

г) Покладемо $\alpha_2 = \frac{2}{3}$, тоді $\beta_{21} = \frac{2}{3}, p_1 = \frac{1}{4}, p_2 = \frac{3}{4}$, тоді ще один метод другого порядку точності можна записати у вигляді:

$$y_{n+1} = y_n + \frac{\tau}{4}\left(f\left(t_n, y_n\right) + 3 f\left(t_{n+2/3}, y_n + \frac{2}{3} f_n \right)\right).$$

3. Трьохстадійні методи ($m = 3$) .

Для трьохстадійних методів $R(\tau) = O\left(\tau^4\right)$, отже $p = 3$ і отримуємо методи третього порядку точності.

Запишемо результат вибору параметрів у вигляді таблиці Батчера (один з частинних випадків):

$$\begin{array}{c|ccc} 0 & & & \\ \frac{1}{3} & \frac{1}{3} & & \\ \frac{2}{3} & 0 & \frac{2}{3} & \\ \hline & \frac{1}{4} & 0 & \frac{3}{4} \end{array}$$

4. Чотирьохстадійні методи ($m = 4$).

Для чотирьохстадійних методів $R(\tau) = O(\tau^5)$, отже $p = 4$ і отримуємо методи четвертого порядку точності. Найбільш поширені методи:

$$\begin{array}{c|cccc} 0 & & & & \\ \frac{1}{2} & \frac{1}{2} & & & \\ \frac{1}{2} & 0 & \frac{1}{2} & & \\ 1 & \frac{1}{2} & 0 & 1 & \\ \hline & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6} \end{array} \qquad \begin{array}{c|cccc} 0 & & & & \\ \frac{1}{4} & \frac{1}{4} & & & \\ \frac{1}{2} & 0 & \frac{1}{2} & & \\ 1 & 1 & -2 & 2 & \\ \hline & \frac{1}{6} & 0 & \frac{2}{3} & \frac{1}{6} \end{array}$$

Розглянемо, як пов'язаний порядок методу p і кількість стадій m. Маємо

кількість стадій m	1	2	3	4	5	6
точність методу p	1	2	3	4	4	5

З таблиці видно, що при $m \geq 5$ збільшення кількості стадій (а це означає збільшення обчислювальної роботи) не дає значного виграшу в точності, тому такі методи застосовуються досить рідко.

<u>Теорема</u>. Нехай m- стадійний метод Рунге-Кутти має p-й порядок апроксимації, а $f(t,u)$ задовольняє умові Ліпшиця по змінній u. Тоді метод має p-й степінь точності і для похибки методу Рунге-Кутти має місце така оцінка

$$|z_n| = |y_n - u(t_n)| \leq T \cdot e^{\alpha T} \cdot \max_{0 \leq j \leq n-1} |\psi_j| = O(\tau^p),$$

де $\psi_j = \frac{1}{\tau} R_j(\tau)$ - похибка апроксимації метода, $\alpha = \lambda L m (1 + BL\tau_0)^{m-1}$,

L – стала Ліпшиця, $B = \max\limits_{2\le i\le m, 1\le j\le i-1} |\beta_{ij}|$, $\quad T = \max\limits_{n} |t_n|$, $\lambda = \max\limits_{1\le i\le n} |p_i|$. □

Методи контролю точності на кроці

Під час розрахунків часто необхідно змінювати крок інтегрування (розв'язання задачі Коші) для контролю величини похибки методу на кроці. При практичній оцінці цієї величини можна застосувати такі підходи:

1) Використання правила Рунге.

Головний член похибки на одному кроці $[t_n, t_{n+1}]$ при застосуванні, наприклад, методу Рунге–Кутти p-го степеня точності, має вигляд

$$\overset{0}{R}(\tau) = \frac{\varphi^{(p+1)}(0)\tau^{p+1}}{(p+1)!}.$$

В результаті проведення двох кроків τ отримаємо наближення $y^{\tau}(t_n + 2\tau)$ до значення $u(t_n + 2\tau)$ таке, що

$$y^{\tau}(t_n + 2\tau) - u(t_n + 2\tau) \approx 2\frac{\varphi^{(p+1)}(0)\tau^{p+1}}{(p+1)!} = 2\overset{0}{R}(\tau). \qquad (5)$$

Якщо тепер застосувати метод Рунге – Кутти p-го степеня точності один раз з кроком довжини 2τ на інтервалі $[t_n, t_{n+2}]$, то отримаємо наближене значення $y^{2\tau}(t_n + 2\tau)$, для якого

$$y^{2\tau}(t_n + 2\tau) - u(t_n + 2\tau) \approx \frac{\varphi^{(p+1)}(0)(2\tau)^{p+1}}{(p+1)!} = \overset{0}{R}(2\tau) = 2^{p+1}\overset{0}{R}(\tau). \qquad (6)$$

Віднявши від (6) відповідні частини співвідношення (5), будемо мати

$$y_{n+2}^{2\tau} - y_{n+2}^{\tau} \approx 2(2^p - 1)\overset{0}{R}(\tau),$$

звідки

$$2\overset{0}{R}(\tau) \approx \frac{y_{n+2}^{2\tau} - y_{n+2}^{\tau}}{2^p - 1}.$$

Тоді отримуємо оцінку головного члена похибки на кроці

$$y^{\tau}(t_n + 2\tau) - u(t_n + 2\tau) \approx 2\overset{0}{R}(\tau) \approx \frac{y_{n+2}^{2\tau} - y_{n+2}^{\tau}}{2^p - 1}.$$

При необхідності можна уточнити отримане наближене значення додавши до нього величину головного члена похибки, тобто

$$u(t_n + 2\tau) \approx y_{n+2}^{\tau} + \frac{y_{n+2}^{\tau} - y_{n+2}^{2\tau}}{2^p - 1}.$$

Позначимо $g(\tau) = \dfrac{y_{n+2}^{2\tau} - y_{n+2}^{\tau}}{2^p - 1}$. Якщо $|g(\tau)| \le \varepsilon$, де ε – деяка задана мала величина (похибка на одному кроці), то τ – успішний крок і $u(t_n + 2\tau) \approx y_{n+2}^{\tau}$.

Якщо $|g(\tau)| > \varepsilon$, то зменшуємо крок, наприклад вдвічі, поклавши $\tau := \frac{\tau}{2}$. Потрібно також передбачити умову збільшення кроку. Це можна зробити наступним чином: задаємо деяке $\delta << \varepsilon$ і, якщо $|g(\tau)| \le \delta$, то $u(t_n + 2\tau) \approx y_{n+2}^{\tau}$ і далі беремо $\tau := 2\tau$ та продовжуємо обчислення з більшим кроком. Параметр δ можна вибрати, наприклад, таким чином: $\delta = \alpha \cdot \varepsilon \cdot 2^{-p}$, $0 < \alpha < 1$, де p – порядок точності методу.

2) Другий підхід вибору кроку інтегрування полягає у використанні методів різного степеня точності. Тобто, якщо є два методи степеня точності p та $p+1$ на кроці, то будемо мати

$$y_{n+1}^{(1)} - u(t_{n+1}) = O\left(\tau^{p+1}\right),$$
$$y_{n+1}^{(2)} - u(t_{n+1}) = O\left(\tau^{p+2}\right).$$

Тоді можемо оцінити головний член похибки (першого методу)

$$g(\tau) = y_{n+1}^{(1)} - y_{n+1}^{(2)} = O\left(\tau^{p+1}\right).$$

Далі для головного членом похибки $g(\tau)$ використовуємо той же алгоритм, що і в першому підході.

Бажано мати можливість здійснювати крок інтегрування та оцінювати похибку при меншій кількості обчислення значень правих частин. Виграш досягається, якщо використовують методи, які називаються **вкладеними**. Таблиця Батчера для них має вигляд:

$$\begin{array}{c|cccc}
\alpha_2 & \beta_{21} & & & \\
\vdots & \vdots & & & \\
\alpha_m & \beta_{m1} & \cdots & \beta_{mm-1} & \\
\hline
 & p_1 & & p_{m-1} & p_m \\
 & \overline{p}_1 & & \overline{p}_{m-1} & \overline{p}_m
\end{array}$$

Метод з параметрами $p_1, ..., p_{m-1}, p_m$ має порядок точності p, а з параметрами $\overline{p}_1, ..., \overline{p}_{m-1}, \overline{p}_m$ - $p+1$. Коефіцієнти α_i, β_{ij} у обох таких методів однакові.

Найпростіший приклад вкладених методів для $m = 2$ має таблицю Батчера:

$$\begin{array}{c|cc}
0 & & \\
1 & 1 & \\
\hline
 & 1 & 0 \\
 & 1/2 & 1/2
\end{array}$$

Перший метод, якому відповідають коефіцієнти $p_1 = 1, p_2 = 0$, це метод Ейлера, $p = 1$. Другий - $p_1 = 1/2, p_2 = 1/2$ - метод Ейлера – Коші, $p = 2$. Обидва ці методи є методами типу Рунге - Кутти.

В загальному випадку будується метод Рунге -Кутти $p+1$ –го порядку точності для якого обчислюються всі необхідні коефіцієнти $k_i(\tau)$ та за їх допомогою підбирається новий наближений метод порядку *p,* який вже не є методом Рунге –Кутти (будується виходячи не з умов мінімізації похибки на кроці). Інколи беруть формули порядку менше ніж *p*, що спрощує обчислювальні формули.

Приклад - це сукупність формул шестистадійного ($m=6$) методу Рунге – Кутти – Фельберга п'ятого порядку точності :

$$y_{n+1} = y_n + \frac{1}{336}\left(14k_1 + 35k_4 + 162k_5 + 125k_6\right),$$

$$k_1 = \tau f\left(t_n, y_n\right),\ k_2 = \tau f\left(t_n + \frac{\tau}{2},\ y_n + \frac{k_1}{2}\right),\ k_3 = \tau f\left(t_n + \frac{\tau}{2},\ y_n + \frac{1}{4}\left(k_1 + k_2\right)\right),$$

$$k_4 = \tau f\left(t_n + \tau, y_n - k_2 + 2k_3\right),\ k_5 = \tau f\left(t_n + \frac{2\tau}{3}, y_n + \frac{1}{27}\left(7k_1 + 10k_2 + k_4\right)\right),$$

$$k_6 = \tau f\left(t_n + \frac{\tau}{5}, y_n + \frac{1}{625}\left(28k_1 - 125k_2 + 546k_3 + 54k_4 - 378k_5\right)\right)$$

дає оцінку для методу четвертого порядку

$$y_{n+1} = y_n + \frac{1}{6}\left(k_1 + 4k_3 + k_4\right),$$

з головним членом похибки

$$g(\tau) = -\frac{1}{366}\left(42k_1 + 224k_3 + 21k_4 - 162k_5 - 125k_6\right).$$

Методи Рунге – Кутти – Фельберга мають четвертий та п'ятий степінь точності. Цей метод дозволяє зменшити кількість обчислень правої частини вихідного рівняння до 6 в порівнянні з 11 звертаннями до правої частини при застосуванні правила Рунге.

<u>Зауваження.</u> Відзначимо, що в методах Рунге-Кутти для отримання досить високого степеня точності потрібно багато разів обчислювати значення функції, з іншого боку, є можливість зміни кроку інтегрування $\tau = \tau_n$ і за рахунок цього задовольняти умову точності на кроці.

Багатокрокові методи розв'язання задачі Коші

Методи Рунге-Кутти потребують великої кількості обчислень значень функцій на одному кроці (особливо чутливо це при розв'язанні систем ДР). Висока точність в цих методах досягається за рахунок обчислень для m-стадійних методів коефіцієнтів $k_i(\tau) = \tau f(\xi_i, \eta_i)$ в проміжних точках між t_n та $t_n + \tau$. Інший підхід полягає у використанні попередніх значень

$$f(t_n, y_n), f(t_{n-1}, y_{n-1}), ..., f(t_{n-m}, y_{n-m}).$$

Для розв'язання задачі Коші

$$\frac{du}{dt} = f(t,u),\ t > 0,\ u(0) = u_0 \qquad (1)$$

введемо сітку $\omega_\tau = \{t_n = n\tau,\ n = 0,1,...\}$ з постійним кроком $\tau > 0$. Позначимо через $y_n = y(t_n)$, $f_n = f(t_n, y_n)$ - функції, визначені на сітці ω_τ.

Лінійний m- кроковий метод розв'язання задачі Коші можна записати в наступному вигляді

$$\frac{a_0 y_n + a_1 y_{n-1} + ... + a_m y_{n-m}}{\tau} = b_0 f_n + b_1 f_{n-1} + ... + b_m f_{n-m},\ n = m, m+1, ..., \qquad (2)$$

де a_k, b_k – числові коефіцієнти, які не залежать від n, причому $a_0 \neq 0$,

або в іншому вигляді

$$\sum_{k=0}^{m} \frac{a_k}{\tau} y_{n-k} = \sum_{k=0}^{m} b_k f_{n-k}\ . \qquad (3)$$

Співвідношення (2) можна розглядати як рекурентне, яке виражає значення y_n через знайдені раніше значення $y_{n-1}, y_{n-2}, ..., y_{n-m}$. Розрахунок починається з $n = m$, тобто з рівняння

$$\frac{a_0 y_m + a_1 y_{m-1} + ... + a_m y_0}{\tau} = b_0 f_m + b_1 f_{m-1} + ... + b_m f_0\ .$$

Видно, що для початку розрахунків необхідно задати m початкових значень. Значення y_0 визначається початковою умовою задачі (1), а саме $y_0 = u_0$. Величини $y_1, ..., y_{m-1}$ можна обчислити, наприклад, за методом Рунге-Кутти, або за методом розкладання в ряд Тейлора. В подальшому будемо вважати, що початкові значення $y_0, y_1, ..., y_{m-1}$ задані.

З виразу (2) видно, що на відміну від методів Рунге-Кутти в багатокрокових методах праві частини обчислюються тільки в точках основної сітки ω_τ.

Метод (2) називається **явним**, якщо $b_0 = 0$, і відповідно шукане значення y_n виражається явно через $y_{n-1}, ..., y_{n-m}$.

$$y_n = \frac{\tau}{a_0}\left(-\sum_{k=1}^{m} \frac{a_k}{\tau} y_{n-k} + \sum_{k=1}^{m} b_k f_{n-k} \right).$$

Якщо $b_0 \neq 0$, то метод (2) називається **неявним**. В неявному методі для пошуку y_n потрібно розв'язувати нелінійне рівняння

$$\frac{a_0}{\tau} y_n - b_0 f(t_n, y_n) = F(y_{n-1}, y_{n-2}, ..., y_{n-m}),$$

де

$$F(y_{n-1}, y_{n-2}, ..., y_{n-m}) = \sum_{k=1}^{m} \left(b_k f_{n-k} - \frac{a_0}{\tau} y_{n-k} \right).$$

Як правило це рівняння розв'язують методом Ньютона, обираючи початкове наближення, наприклад, наступним чином : $\overset{(0)}{y_n} = y_{n-1}$.

При практичному використанні серед багатокрокових методів найбільш поширені **методи Адамса**.

Розглянемо побудову **інтерполяційного (неявного)** методу Адамса. Проінтегруємо рівняння (1) по $t \in [t_{n-1}, t_n]$:

$$\frac{u(t_n) - u(t_{n-1})}{\tau} = \frac{1}{\tau} \int_{t_{n-1}}^{t_n} f(t, u(t)) dt. \qquad (4)$$

Замінимо $f(t, u(t))$ на інтерполяційний поліном m - го степеня за вузлами інтерполювання $t_n, t_{n-1}, ..., t_{n-m}$. Тоді, використовуючи формулу Ньютона для рівновіддалених вузлів, отримаємо

$$f(t, u(t)) = L_m(t) + r_m(t) =$$

$$= f_n + s\Delta f_{n-1} + ... + \frac{s(s+1)\cdot ... \cdot (s+m-1)}{m!} \Delta^m f_{n-m} + r_m(t),$$

де $s = \dfrac{t - t_n}{\tau}$, τ - сталий крок, $r_m(t) = O(\tau^{m+1})$ - залишковий член інтерполяційної формули. В результаті підстановки в (4) $f(t, u(t)) \approx L_m(t)$ отримаємо **неявний метод Адамса (інтерполяційний метод Адамса)**

$$\frac{y_n - y_{n-1}}{\tau} = \sum_{k=0}^{m} p_k \Delta^k f_{n-k} = f_n + p_1 \Delta f_{n-1} + ... + p_m \Delta^m f_{n-m}, \qquad (5)$$

де

$$p_0 = 1, p_k = \int_{-1}^{0} \frac{s(s+1)\cdot ... \cdot (s+k-1)}{k!} ds, k = \overline{1, m}$$

$$\Delta f_{n-1} = f_n - f_{n-1}, \ \Delta^2 f_{n-2} = f_n - 2f_{n-1} + f_{n-2}, \dots .$$

Для похибки методу на кроці будемо мати

$$R_{n-1}(\tau) = \int_{t_{n-1}}^{t_n} r_m(t) dt = O(\tau^{m+2}),$$

а отже степінь точності методу на кроці $p = m + 1$.

Формула (4) називається формулою Адамса-Мултона.

Для побудови **екстраполяційного (явного)** методу Адамса виберемо наступні вузли інтерполювання $t_{n-1},...,t_{n-m}$. За цими вузлами побудуємо інтерполяційний поліном $(m-1)$ - го степеня, тоді

$$f(t,u(t)) = L_{m-1}(t) + r_{m-1}(t) =$$
$$= f_{n-1} + s\Delta f_{n-2} + ... + \frac{s(s+1)\cdot...\cdot(s+m-2)}{(m-1)!}\Delta^{m-1} f_{n-m} + r_{m-1}(t),$$

де $s = \frac{t - t_{n-1}}{\tau}$, $r_{m-1}(t) = O(\tau^m)$ - залишковий член інтерполяційної формули. В результаті підстановки в (4) $f(t,u(t)) \approx L_{m-1}(t)$ отримаємо **явний метод Адамса (екстраполяційний метод Адамса)**

$$\frac{y_n - y_{n-1}}{\tau} = \sum_{k=0}^{m-1} q_k \Delta^k f_{n-1-k} = f_{n-1} + q_1 \Delta f_{n-2} + ... + q_{m-1}\Delta^{m-1} f_{n-m}\ , \qquad (6)$$

де

$$q_0 = 1,\ q_k = \int_0^1 \frac{s(s+1)\cdot...\cdot(s+k-1)}{k!} ds\ ,\ k = \overline{1, m-1}.$$

Формула (6) називається формулою Адамса-Башфорта. Для похибки методу на кроці будемо мати

$$R_{n-1}(\tau) = \int_{t_{n-1}}^{t_n} r_{m-1}(t) dt = O(\tau^{m+1}),$$

а отже степінь точності методу на кроці $p = m$.

Таким чином методи Адамса є частинним випадком багатокрокових методів (2), коли похідна $u'(t)$ апроксимується тільки за двома точками t_n і t_{n-1}, тобто

$$a_0 = -a_1 = 1,\ a_k = 0,\ k = \overline{2, m}.$$

Отже методи Адамса мають вигляд

$$\frac{y_n - y_{n-1}}{\tau} = \sum_{k=0}^{m} b_k f_{n-k}\ .$$

У випадку, коли $b_0 = 0$, методи Адамса називають **явними**, а при $b_0 \neq 0$, методи Адамса називають **неявними.**

Метод невизначених коефіцієнтів побудови багатокрокових методів для розв'язання задачі Коші

Розглянемо задачу Коші

$$\frac{du}{dt} = f(t,u),\ t > 0,\ u(t_0) = u_0. \tag{1}$$

і багатокроковий метод

$$\sum_{k=0}^{m} \frac{a_k}{\tau} y_{n-k} = \sum_{k=0}^{m} b_k f_{n-k}\ . \tag{2}$$

Виберемо коефіцієнти a_k, b_k так, щоб досягти найвищої точності (найвищого порядку похибки апроксимації) методу (2).

Розглянемо похибку апроксимації $\psi(\tau)$ для формули (2). Похибкою апроксимації (або нев'язкою) різницевого методу (2) на розв'язку задачі (1) називається функція

$$\psi(\tau) = -\sum_{k=0}^{m} \frac{a_k}{\tau} u_{n-k} + \sum_{k=0}^{m} b_k f(t_{n-k}, u_{n-k}), \tag{3}$$

де $u_{n-k} = u(t_n - k\tau)$. Її отримують підстановкою точного розв'язку $u(t)$ задачі (1) в різницеве рівняння (2). Розглянемо питання про порядок похибки апроксимації при $\tau \to 0$ в залежності від вибору коефіцієнтів a_k, b_k, $k = \overline{0,m}$. Розкладемо функції $u_{n-k} = u(t_n - k\tau)$ в точці $t = t_n$ за формулою Тейлора

$$u(t_n - \tau k) = \sum_{j=0}^{p} u^{(j)}(t_n) \cdot \frac{(-k\tau)^j}{j!} + O(\tau^{p+1}), \tag{4}$$

де $u^{(j)}$ – j-та похідна. Відповідно

$$f(t_n - k\tau, u(t_n - k\tau)) = \frac{du(t_n - k\tau)}{dt} = \sum_{j=0}^{p-1} u^{(j+1)}(t_n) \frac{(-k\tau)^j}{j!} + O(\tau^p),\ k = \overline{1,m}. \tag{5}$$

Підставляючи ці вирази в формулу (3), отримаємо

$$\psi(\tau) = -\sum_{l=0}^{p} \varphi_l\, \tau^{l-1} u^{(l-1)}(t_n) + O(\tau^p).$$

Звідси видно, що похибка апроксимації має порядок p, якщо виконуються умови

$$\varphi_0 = \sum_{k=0}^{m} a_k = 0, \varphi_l = \sum_{k=0}^{m} \left[\frac{(-k)^l}{l!} a_k - \frac{(-k)^{l-1}}{(l-1)!} b_k \right] = 0,\ l = \overline{1,p}. \tag{6}$$

або в іншому вигляді

$$\varphi_0 = \sum_{k=0}^{m} a_k = 0,\ \ \varphi_l = \sum_{k=0}^{m} k^{l-1}(k a_k + l\, b_k) = 0,\ l = \overline{1,p}. \tag{6.1}$$

Формули (6) утворюють систему з $p+1$ лінійних алгебраїчних однорідних рівнянь відносно $2m+2$ невідомих $a_0, a_1, ..., a_m, b_0, b_1, ..., b_m$. Така система буде мати ненульовий розв'язок при умові $2m+2 > p+1$.
Окрім того, бачимо, що повинні виконуватися умови апроксимації лівих та правих частин (1) та (2):

$$\lim_{\tau\to 0}\sum_{k=0}^{m}\frac{a_k}{\tau}u_{n-k}=\frac{du(t_n)}{dt},\ \ \lim_{\tau\to 0}\sum_{k=0}^{m}b_k f\left(t_{n-k},u_{n-k}\right)=f\left(t_n,u_n\right). \tag{7}$$

Тоді, використовуючи розкладання в ряд Тейлора, отримаємо

$$\lim_{\tau\to 0}\left(\sum_{k=0}^{m}\frac{a_k}{\tau}u_n-\sum_{k=0}^{m}ka_k\frac{du(t_n)}{dt}+O(\tau)\right)=\frac{du(t_n)}{dt},$$

$$\lim_{\tau\to 0}\left(\sum_{k=0}^{m}b_k f\left(t_n,u_n\right)+O(\tau)\right)=f\left(t_n,u_n\right).$$

Звідки отримаємо, що умови (7) будуть виконані, якщо

$$\sum_{k=0}^{m}a_k=0,\ \ -\sum_{k=0}^{m}ka_k=1,\ \ \sum_{k=0}^{m}b_k=1. \tag{7.1}$$

З рівняння (6) при $l=0$ маємо

$$\sum_{k=0}^{m}a_k=0,$$

при $l=1$ отримуємо

$$-\sum_{k=0}^{m}ka_k-\sum_{k=0}^{m}b_k=0.$$

Тому умови (7.1) додатково до (6) дають тільки рівняння

$$\sum_{k=0}^{m}b_k=1. \tag{8}$$

Умови (8) називаються умовами нормування, оскільки коефіцієнти рівняння (2) визначені з точністю до множника.
Тобто a_k, b_k, ($k=\overline{0,m}$)) визначаються з (6) та (8) (або (6.1) та (8)).

Як уже відзначалося, система (6) буде мати ненульовий розв'язок при умові $p+1<2m+2$, або $p\le 2m$. Це означає, що порядок апроксимації лінійних m - крокових різницевих методів не може перевищувати $2m$ (неявні методи). Найвищий порядок апроксимації явних методів ($b_0=0$), буде дорівнювати $2m-1$. Якщо який-небудь параметр зафіксувати, то порядок апроксимації буде знижуватися (оскільки $p+1<2m+2$; тут $2m+2$ - кількість невідомих, а вона буде меншою).
Звідси бачимо, що найвищий порядок апроксимації неявного m- крокового методу Адамса дорівнює $m+1$, оскільки фіксовані m+1 коефіцієнтів a_k, але

в той же час автоматично виконуються умови $\varphi_0 = 0, \varphi_1 = 0$, тобто $p+1-2 < 2m+2-(m+1)$, або $p \le m+1$. Найвищий порядок апроксимації явного методу Адамса ($b_0 = 0$) дорівнює m.

Реалізація багатокрокових методів

1). Перша проблема, яка виникає при застосуванні багатокрокових методів це вибір додаткових початкових умов. Треба знайти $(m-1)$ додаткове початкове значення $y_1,...,y_{m-1}$. Шляхи вирішення проблеми наступні:

а) Використати методи Рунге-Кутти для пошуку цих початкових значень. Але потрібно, щоб ці методи мали або порядок похибки апроксимації p (або порядок точності методу на кроці p). Недоліком такого підходу є те, що сама процедура розрахунків за методами Рунге-Кутти потребують досить великої кількості проміжних обчислень (хоча й багатокрокові досить об'ємними відносно кількості обчислень).

б) Використати метод розкладання в ряд Тейлора. При цьому також потрібно узгоджувати точність p. Функцію розкладемо в ряд

$$\overset{(p-1)}{u}(t) = \sum_{k=0}^{p-1} \frac{u^{(k)}(t_0)}{k!}(t-t_0)^k,$$

де індекс $p-1$ над u вказує кількість членів ряду. Тоді похибка

$$u - \overset{(p-1)}{u} = O(\tau^p).$$

2). Друга проблема - реалізація неявних методів

$$y_n = \tau \frac{b_0}{a_0} f(t_n, y_n) + \Phi(y_{n-1},...,y_{n-m}).$$

Способи вирішення проблеми наступні:

а) Для знаходження розв'язку нелінійного рівняння можна використовувати метод простої ітерації

$$y_n^{(k+1)} = \varphi(y_n^{(k)}),$$

де $\varphi(y_n) = \tau \frac{b_0}{a_0} f(t_n, y_n) + \Phi(y_{n-1},...,y_{n-m})$, індекс k означає k-ту ітерацію.

Умова збіжності методу простої ітерації $|\varphi'(u)| \le q < 1$. Тоді маємо

$$|\varphi'(u)| = \left|\tau \frac{b_0}{a_0} f'_u(t,u)\right| \le q < 1 \;\Rightarrow\; \tau < \frac{a_0}{b_0 \cdot L}, \text{ де } |f'_u(t,u)| \le L,$$

де L – максимум похідної чи стала Ліпшиця. Тобто для збіжності методу необхідно, щоб крок τ був досить малим.

б) Для знаходження розв'язку нелінійного рівняння можна використовувати метод Ньютона. Умови збіжності методу Ньютона залежать

від вдалого вибору початкових умов. За початкове наближення можна вибрати, наприклад, $y_n^{(0)} = y_{n-1}$ або $y_n^{(0)}$ обчислюється за явною m-кроковою формулою.

В результаті умови збіжності методу Ньютона менш жорсткі, ніж у методу простої ітерації.

в) Для реалізація неявних методів можна використовувати методи типу предиктор-коректор. Запишемо формули неявного методу (C – коректор, P – предиктор):

P: $y_n = \overline{\Phi}(y_{n-1},...,y_{n-m}, f_{n-1},...,f_{n-m})$ – явний m- кроковий метод, $p = 2m-1$.

C: $y_n = \tau \dfrac{b_0}{a_0} f(t_n, y_n) + \Phi(y_{n-1},...,y_{n-m}, f_{n-1},...,f_{n-m})$ – неявний m-кроковий метод, $p = 2m$.

Далі виконується наступний алгоритм:

P: $\overline{y}_n = \overline{\Phi}(y_{n-1},...,y_{n-m}, f_{n-1},...,f_{n-m})$;

E: $f_n = f(t_n, \overline{y}_n)$;

C: $y_n = \tau \dfrac{b_0}{a_0} f(t_n, \overline{y}_n) + \Phi(y_{n-1},...,y_{n-m}, f_{n-1},...,f_{n-m})$;

$$\left\{ y_n = \tau \frac{b_0}{a_0} f(t_n, \overline{y}_n) + \Phi(y_{n-1},...,y_{n-m}, f(t_{n-1}, y_{n-1}),...,f(t_{n-m}, y_{n-m})) \right\};$$

E: $f_n = f(t_n, y_n)$,

де E – обчислення правої частини рівняння. Схема $PECE$ – це метод предиктор - коректор.

Іноді, щоб підвищити точність, використовують схему $PE(CE)^j$. (В цій схемі починаючи з другого кроку використовуємо $E: f_n = f(t_n, y_n); \overline{y}_n = y_n$). Ця схема аналогічна методу простої ітерації, де j – це максимальна кількість ітерацій.

Стійкість методів розв’язання задачі Коші

Розглянемо питання стійкості чисельних методів розв’язання задачі Коші, тобто неперервну залежність наближениих розв’язків від збурень вхідних даних, похибок заокруглення та ін.

Чисельний метод розв'язання задачі Коші

$$\frac{du}{dt} = f(t,u),\ u(t_0) = u_0 \tag{1}$$

називається стійким за початковими даними, якщо для довільної задачі Коші (тобто $\forall f(t,u)$) для $\forall\, \varepsilon > 0\, \exists\, \sigma = \sigma(\varepsilon) > 0$ таке, що при $\left|\overline{u}_0 - u_0\right| < \sigma$ будемо мати $\left|\overline{y}_n - y_n\right| < \varepsilon,\quad \forall n$.

Чисельний метод називається нестійким, якщо $\exists\, \varepsilon > 0$ таке, що $\forall\, \sigma = \sigma(\varepsilon) > 0\ \ \exists\, f(t,u)$, що при $\left|\overline{u}_0 - u_0\right| < \sigma$ будемо мати $\left|\overline{y}_n - y_n\right| > \varepsilon,\ \ n \ge N(\varepsilon)$.

Перше визначення складніше перевіряти ніж друге. Перевіркою умови нестійкості можна відокремити нестійкі методи на наступних класах правих частин:

а) $f \equiv 0$; б) $f = \lambda u, \lambda < 0$; в) $f = \lambda u, \operatorname{Re}\lambda < 0$.

Розглянемо ці випадки:

а) $f \equiv 0$. Якщо застосувати метод Рунге-Кутти, то $y_{n+1} = y_n = \ldots = y_0$, а це означає, що малі збурення y_0 приводять до таких самих збурень в y_{n+1}: $\delta = \varepsilon$.

Розглянемо тепер багатокроковий метод, що приводить до різницевого рівняння n-го порядку

$$\sum_{k=0}^{m} \frac{a_k}{\tau} y_{n-k} = 0 \tag{2}$$

Часткові розв'язки будемо шукати у вигляді $y_n = q^n$, ... $, y_{n-k} = q^{n-k}$. Підставимо їх в (2), тоді отримаємо характеристичне рівняння

$$\sum_{k=0}^{m} a_k q^{m-k} = 0 \tag{3}$$

корені якого позначимо через $q_1, \ldots, q_m$.

Розв'язок рівняння (2) будемо шукати у вигляді:

а) якщо q – дійсні та різні, то $y_n = \sum_{k=0}^{m} c_k q_k^n$;

б) якщо серед коренів є дійсні кратні (наприклад, корінь q_j має кратність 2), тоді $y_n = \ldots + c_j q_j^n + \tilde{c}_j n q_j^n + \ldots,$;

в) якщо якщо серед коренів є комплексні (наприклад, корінь q_j - комплексний), то $y_n = .. + c_j \rho_j^n \cos(n \operatorname{Im} q_j) + \tilde{c}_j \rho_j^n \sin(n \operatorname{Im} q_j) + ..$,

де $\rho_j = \left|q_j\right|$. Якщо $\exists\, \rho_j > 1$, то $y_n \to \infty$ при $n \to \infty$.

Величини c_k знаходять виходячи з початкових умов.

Якщо $\left|q_k\right| > 1$, то багатокроковий метод – нестійкий.

Визначення. Багатокроковий метод розв'язання задачі Коші задовольняє кореневій умові, якщо всі корені характеристичного рівняння (3) лежать в одиничному колі, а на границі кола немає кратних коренів, тобто $|q_j| \le 1$, для $q_j = q_m$, $|q_j| = |q_m| < 1$.
Якщо метод не задовольняє кореневій умові, то він нестійкий.

Приклад. Розглянемо явний двокроковий метод найвищого степеня точності ($m = 2$, $p = 3$):

$$\frac{y_n + 4y_{n-1} - 5y_{n-2}}{6\tau} = \frac{2}{3} f_{n-1} + \frac{1}{3} f_{n-2}.$$

Характеристичне рівняння для нього буде мати вигляд ($f = 0$)

$$q^n + 4q^{n-1} - 5q^{n-2} = 0.$$

Тоді

$$q^2 + 4q - 5 = 0,\ q_1 = 1, q_2 = -5.$$

Отже явний двокроковий метод найвищого степеня точності – нестійкий метод оскільки $q_2 = -5$, тобто $|q_2| > 1$.

Теорема 1. Якщо багатокроковий метод задовольняє кореневій умові та права частина задовольняє умові Ліпшиця, $|f(t,u) - f(t,v)| \le L|u - v|$, тоді має місце оцінка точності

$$|z_n| = |y_n - u(t_n)||z_n| \le M\left(\max_{k=\overline{0,n-1}} |z_k| + \max_{k=\overline{0,n-1}} |\psi_k|\right), \tag{4}$$

де ψ_k – похибка апроксимації. □

б) $f = \lambda u$, $\lambda < 0$. В цьому випадку розв'язок задачі Коші (1) $u(t) = u_0 e^{\lambda(t-t_0)} \underset{t\to\infty}{\to} 0$ і він є асимптотично стійким. Слід очікувати, що чисельний метод буде мати таку ж властивість. Тобто

$$|y_n| < |y_{n-1}| < \ldots < |y_0| = |u_0|. \tag{5}$$

Для явного методу Ейлера будемо мати

$$y_{n+1} = y_n + \tau f(t_n, y_n) = y_n + \tau \lambda y_n = (1 + \tau \lambda) y_n = q y_n,\ q = 1 + \tau\lambda .$$

Тому (5) має місце, якщо $|q| < 1$ або $-1 < q < 1$ (корінь характеристичного рівняння при $m = 1$). Права нерівність виконується для довільного $\tau > 0$, а ліва тільки для $\tau < \frac{2}{|\lambda|}$.

Визначення. Метод розв'язання задачі Коші називається **умовно стійким**, якщо він стійкий при $\tau \le \tau_0$. Якщо ж він стійкий для $\forall \tau > 0$, то такий метод називається **абсолютно стійким**.

Таким чином явний метод Ейлера умовно стійкий.

Для неявного методу Ейлера маємо

$$y_{n+1} = y_n + \tau f(t_{n+1}, y_{n+1}) = y_n + \tau\lambda y_{n+1}.$$

Звідси маємо $y_{n+1} = \dfrac{1}{1-\tau\lambda} y_n$, $0 < q = \dfrac{1}{1-\tau\lambda} < 1, \lambda < 0$. Отже неявний метод Ейлера абсолютно стійкий на тестовому рівнянні.

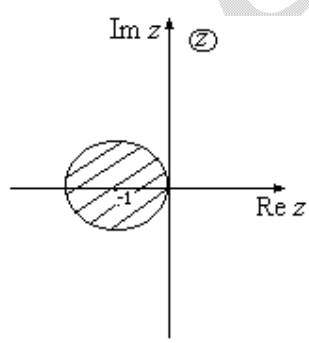
в) $f = \lambda u$, $\operatorname{Re}\lambda < 0$. Розглянемо систему

$$\frac{d\vec{u}}{dt} = A\vec{u},$$

де A – матриця простої структури, тобто $\exists H: \ \Lambda = HAH^{-1} = diag(\lambda_i)$, $\lambda_i = \lambda_i(A)$. Попередня система зводиться до наступної:

$$\frac{d\vec{v}}{dt} = \Lambda\vec{v}, \text{ або } \frac{dv_i}{dt} = \lambda_i v_i, \ i = \overline{1,n}.$$

Тому $v_i = v_i^0 e^{\lambda_i t} = v_i^0 e^{\operatorname{Re}\lambda_i t}\left(\cos\varphi(t) + j\sin\varphi(t)\right)$, де φ – кут, для якого $tg\varphi = \dfrac{\operatorname{Im}\lambda_i}{\operatorname{Re}\lambda_i}$. Дослідимо явний та неявний методи Ейлера на цьому тестовому прикладі.



Для явниго методу Ейлера маємо $y_{n+1} = qy_n$, $q = 1 + \tau\lambda = 1 + z$, $z = \tau\lambda = \operatorname{Re} z + j\operatorname{Im} z$. Умова $|q| \le 1$ виконується для внутрішності кола радіуса 1 з центром в точці (-1, 0)

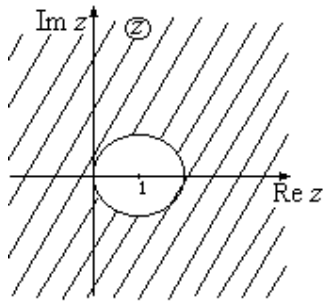
$$|q_i| = |1 + \tau\lambda_i| < 1, \ \forall i.$$

З останнього виразу отримаємо

$$\tau \le \frac{2\operatorname{Re}\lambda_i}{|\lambda_i|^2} = \tau_0,$$

Таким чином явний метод Ейлера умовно стійкий і для тестового рівняння з $\operatorname{Re}\lambda < 0$.

Для неявного методу Ейлера $q = \dfrac{1}{1-z}$.



Умова $|q| \le 1$ виконується для зовнішності кола радіуса $R = 1$, з центром в точці $z_0 = (1,0)$, в той час, як точки $z_i = \tau\lambda_i$ лежать в лівій півплощині, отже умова стійкості виконується для довільних τ. Таким чином неявний метод Ейлера є абсолютно стійким. Недоліком неявного методу Ейлера є низька точність: $p = 1$.

<u>Визначення</u>. Метод розв'язання задачі Коші є A- стійким, якщо область його стійкості для тестового рівняння з $f = \lambda u$, $\operatorname{Re}\lambda < 0$ включає всю півплощину $\operatorname{Re} z < 0$, $z = \tau\lambda$.

<u>Задача.</u> Показати , що метод трапецій розв'язання задач Коші є A-стійким.

<u>Теорема 2</u> (Далквіст). Явний лінійний багатокроковий метод не може бути A-стійким. Порядок неявного A- стійкого методу не може бути вищим двох. □

Найбільш точним серед цих методів є метод трапецій.

<u>Визначення.</u> Задача Коші (1) називається жорсткою для $t \in [t_0, T]$, якщо $\operatorname{Re}\lambda_i < 0$ та

$$S(t) = \frac{\max\limits_i |\operatorname{Re}\lambda_i(t)|}{\min\limits_i |\operatorname{Re}\lambda_i(t)|} >> 1,$$

де $\lambda_i = \lambda_i(t)$ власні значення матриці Якобі правої частини рівняння, тобто

$$A_{ij} = \left(\frac{\partial f_i}{\partial u_j}(t_n, \bar{u}) \right)_{i,j=1}^{n}.$$

Характерною особливістю розв'язків жорсткої задачі Коші є наявність компонент $e^{\lambda_i t}$, які повільно змінюються (коли $|\operatorname{Re}\lambda_i| \approx \min\limits_i |\operatorname{Re}\lambda_i|$) та швидко спадають (коли $|\operatorname{Re}\lambda_i| \approx \max\limits_i |\operatorname{Re}\lambda_i|$). Стійкість явних методів (Рунге-Кутта,

явного методу Ейлера) визначається компонентами, які швидко змінюються, в той час як в точному розв'язку вони швидко прямують до нуля і дають малий вклад в розв'язок. Сам розв'язок змінюється повільно разом з «повільними» компонентами. Але умова стійкості явних методів на всьому проміжку $[t_0, T]$ повинна бути орієнтована на крок $\tau < \dfrac{2}{\max\limits_i |\lambda_i|}$. Це дуже жорстка умова. Тому такі системи називаються жорсткими.

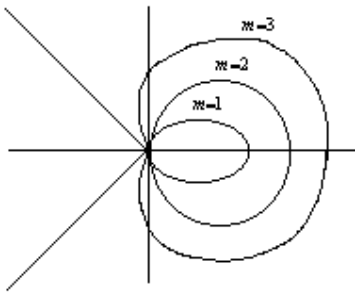
Застосування A- стійких методів знімає проблему стійкості: крок τ можна міняти тільки з умови виконання малої похибки на одному кроці. Але точність A- стійких методів не більше 2, тому можна послабити означення A-стійкості.

<u>Визначення</u>. Метод розв'язання задачі Коші називають $A(\alpha)$- *стійким* $0 \le \alpha \le \dfrac{\pi}{2}$, якщо область стійкості методу для тестового рівняння з $f = \lambda u$, $\operatorname{Re}\lambda < 0$, включає частину півплощини $\operatorname{Re} z$ ($z = \tau\lambda$), що лежить в середині між променями, які утворюють з віссю $\operatorname{Im} z$ кути α. Зауважимо, що A - стійкість це є $A(0)$- стійкість.

Найбільш поширеними для розв'язування жорстких систем є методи Куртіса-Хершфільта:

$$\sum_{k=0}^{m} \frac{a_k}{\tau} y_{n-k} = f(t_n, y_n).$$

Області їх стійкості для різних m приведена на наступному рис.:



Це неявні m- крокові методи з $b_0 = 1$, $b_k = 0$ $k = \overline{1, m}$ (інші назви: методи диференціювання назад, повністю неявні методи).

Для жорстких задач треба використовувати реалізацію неявних методів за допомогою ітераційного методу Ньютона з вибором початкового наближення за явною схемою.

Для жорстких задач також застосовують неявні методи Рунге-Кутти. Але для їх реалізації треба розв'язувати систему з $m \times N$ нелінійних рівнянь, де m–кількість стадій, N – розмірність системи.

Приклад 1. Дослідити стійкість явного методу Ейлера на тестовому рівнянні з $f = \lambda u, \operatorname{Re}\lambda < 0$.

Розв‘язання. Підставимо у формулу методу Ейлера праву частину $f = \lambda u, \operatorname{Re}\lambda < 0$. Отримуємо

$$y_{n+1} = (1 + \tau\lambda) y_n = q y_n .$$

Метод стійкий, якщо

$|q| = \sqrt{\operatorname{Re}^2 q + \operatorname{Im}^2 q} = \sqrt{(1 + \tau \operatorname{Re}\lambda)^2 + \tau \operatorname{Im}^2 \lambda} \le 1, \operatorname{Re}\lambda < 0$. В комплексній площині $z = \tau\lambda$ область стійкості методу Ейлера є колом радіуса 1 з центром в точці $(-1,0)$.

Якщо $\operatorname{Im}\lambda = 0$, то умова має вигляд $|1 + \tau\lambda| \le 1, \lambda < 0$. З цього випливає умова на крок інтегрування $\tau \le \dfrac{2}{|\lambda|}$.

Приклад 2. Дослідити стійкість методу Ейлера-Коші на тестовому рівнянні з $f = \lambda u, \lambda < 0$.

Розв‘язання. Підставимо в формули методу Ейлера-Коші праву частину $f = \lambda u, \lambda < 0$. Отримаємо

$$\bar{y}_{n+1} = y_n + \tau\lambda y_n ,$$

$$y_{n+1} = y_n + \frac{\lambda\tau}{2}(\bar{y}_{n+1} + y_n) = y_n + \frac{\lambda\tau}{2}(y_n + \tau\lambda y_n + y_n),$$

звідки

$$y_{n+1} = \left(1 + z + \frac{z^2}{2}\right) y_n, \; z = \tau\lambda .$$

Умова стійкості $|q| = \left|1 + z + \dfrac{z^2}{2}\right| \le 1$. Вона еквівалентна двом нерівностям

$$1 + z + \frac{z^2}{2} \le 1, \; -1 \le 1 + z + \frac{z^2}{2} .$$

Перша дає умову $z + \dfrac{z^2}{2} \le 0$. Оскільки за умовою задачі $z < 0$, то остаточно перша нерівність еквівалентна умові $|z| \le 2$ або $\tau \le \dfrac{2}{|\lambda|}$.

Друга нерівність $2 + z + \dfrac{z^2}{2} \ge 0$ виконується завжди, оскільки дискримінант квадратного тричлена $D = 1 - 4 \cdot 2 \cdot \dfrac{1}{2} = -3 < 0$.

ДОДАТОК

Системи ортогональних функцій

1). $H = L_2(-1,1)$; $\rho \equiv 1$ (ваговий множник), тоді $\varphi_i(x) = L_i(x)$ - система поліномів **Лежандра**, які мають вигляд

$$L_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n .$$

Система поліномів Лежандра задовольняє рекурентним формулам

$$(n+1)L_{n+1}(x) = (2n+1)xL_n(x) - nL_{n-1}(x),$$

$$L_0(x) = 1,\ L_1(x) = x, L_2(x) = \frac{1}{2}(3x^2 - 1), L_3(x) = \frac{1}{2}(5x^3 - 3x),$$

$$L_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3), L_5(x) = \frac{1}{8}(63x^5 - 70x^3 + 15x),$$

$$L_6(x) = \frac{1}{16}(231x^6 - 315x^4 + 105x^2 - 5), \ldots \ .$$

Це ортогональна система, тобто

$$(L_i, L_k) = \int_{-1}^{1} L_i(x)L_k(x)dx = \delta_{ik} \|L_i(x)\|^2 , \ \ \|L_i(x)\|^2 = \frac{2}{2i+1}, i,k = \overline{0,n}.$$

2). $H = L_{2,\rho}(-1,1)$, $\rho(x) = \dfrac{1}{\sqrt{1-x^2}}$, скалярний добуток $(u,v) = \int_{-1}^{1} \dfrac{u(x)v(x)}{\sqrt{1-x^2}} dx$, тоді $\varphi_i(x) = T_i(x)$, де $\{T_i(x)\}$ система ортогональних поліномів **Чебишова** 1-го роду, які мають вигляд

$$T_n(x) = \cos(n \arccos x).$$

Рекурентна формула для цих поліномів:

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x),$$

$$T_0(x) = 1,\ T_1(x) = x,\ T_2(x) = 2x^2 - 1, T_3(x) = 4x^3 - 3x, T_4(x) = 8x^4 - 8x^2 + 1,$$

$$T_5(x) = 16x^5 - 20x^3 + 5x, T_6(x) = 32x^6 - 48x^4 + 18x^2 - 1,$$

$$T_7(x) = 64x^7 - 112x^5 + 56x^3 - 7x,$$

$$\|T_n\|^2 = \begin{cases} \pi, & n = 0 \\ \dfrac{\pi}{2}, & n = 1,2... \end{cases}$$

3). $H = L_{2,\rho}(-1,1)$, $\rho(x) = (1-x)^{\alpha}(1+x)^{\beta}$. Система $\varphi_i(x) = P_i^{(\alpha,\beta)}(x)$ - поліномів **Якобі** (α, β – числові параметри, причому $\alpha, \beta > -1$) ортогональна в сенсі скалярного добутку $(u,v) = \int_{-1}^{1} (1-x)^{\alpha}(1+x)^{\beta} u(x)v(x)dx$. Ця система є узагальненням випадків 1) та 2). Диференціальна формула для поліномів:

$$P_n^{(\alpha,\beta)}(x) = \frac{(-1)^n}{2^n n!}(1-x)^{-\alpha}(1+x)^{-\beta} \frac{d^n}{dx^n}\left[(1-x)^{n+\alpha}(1+x)^{n+\beta}\right];$$

Рекурентна формула:

$$2(n+1)(n+\alpha+\beta+1)(2n+\alpha+\beta)P_{n+1}^{(\alpha,\beta)}(x) =$$
$$= (2n+\alpha+\beta+1)[(2n+\alpha+\beta)(2n+\alpha+\beta+2)x + \alpha^2 - \beta^2]P_n^{(\alpha,\beta)}(x) -$$
$$- 2(n+\alpha)(n+\beta)(2n+\alpha+\beta+2)P_{n-1}^{(\alpha,\beta)}(x),\quad n = 1,2,...$$

де

$$P_0^{(\alpha,\beta)} = 1,\quad P_{-1}^{(\alpha,\beta)} = 0,\quad \left\|P_n^{(\alpha,\beta)}\right\|^2 = \frac{2^{\alpha+\beta+1}\Gamma(\alpha+n+1)\Gamma(\beta+n+1)}{n!(\alpha+\beta+2n+1)\Gamma(\alpha+\beta+n+1)},$$

та

$$\Gamma(z) = \int_0^{\infty} e^{-t} t^{z-1} dt,\quad \Gamma(z+1) = z\Gamma(z),\quad \Gamma(n+1) = n!$$

Коли $\alpha = \beta = 0$: $P_n^{(0,0)}(x) = L_n(x)$, а для $\alpha = \beta = -\dfrac{1}{2}$: $P_n^{(-\frac{1}{2},-\frac{1}{2})}(x) = T_n(x)$.

4). $H = L_{2,\rho}[0,\infty)$, $\rho(x) = x^{\alpha}e^{-x}$, $\alpha > -1$. Система $\varphi_i(x) = L_i^{\alpha}(x)$ - поліномів **Лагерра** ортогональна в сенсі скалярного добутку $(u,v) = \int_0^{\infty} x^{\alpha}e^{-x}u(x)v(x)dx$. Диференціальна формула для поліномів:

$$L_n^\alpha(x) = (-1)^n x^{-\alpha} e^x \frac{d^n}{dx^n}\left[x^{\alpha+n} e^{-x}\right],$$

зокрема,

$L_0^\alpha(x) = 1,\ L_1^\alpha(x) = -x + \alpha + 1,\ L_2^\alpha(x) = x^2 - 2(\alpha+2)x + (\alpha+1)(\alpha+2),$

$L_3^\alpha(x) = -x^3 + 3(\alpha+3)x^2 - 3(\alpha+2)(\alpha+3)x + (\alpha+1)(\alpha+2)(\alpha+3),$

$L_4^\alpha(x) = x^4 - 4(\alpha+4)x^3 + 6(\alpha+3)(\alpha+4)x^2 - 4(\alpha+2)(\alpha+3)(\alpha+4)x +$

$+ (\alpha+1)(\alpha+2)(\alpha+3)(\alpha+4),$

$L_5^\alpha(x) = -x^5 + 5(\alpha+5)x^4 - 10(\alpha+4)(\alpha+5)x^3 + 10(\alpha+3)(\alpha+4)(\alpha+5)x^2 -$

$- 5(\alpha+2)(\alpha+3)(\alpha+4)(\alpha+5)x + (\alpha+1)(\alpha+2)(\alpha+3)(\alpha+4)(\alpha+5), \ldots$.

Рекурентна формула:

$$(n+1)L_{n+1}^\alpha = (2n + \alpha + 1 - x)L_n^\alpha - (n+\alpha)L_{n-1}^\alpha,$$

де $L_0^\alpha = 1,\quad L_{-1}^\alpha = 0$ та з нормою $\left\|L_n^\alpha\right\|^2 = n!\Gamma(\alpha + n + 1)$.

5). $H = L_{2,\alpha}(-\infty,\infty),\quad \rho(x) = e^{-x^2}$. Система $\varphi_i(x) = H_i(x)$ - поліномів **Ерміта** ортогональна в сенсі скалярного добутку $(u,v) = \int_{-\infty}^{\infty} e^{-x^2} u(x)v(x)dx$.

Диференціальна формула для поліномів:

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n}(e^{-x^2}),$$

зокрема

$$H_0(x) = 1,\ H_1(x) = 2x,\ H_2(x) = 4x^2 - 2,\ H_3(x) = 8x^3 - 12x,$$
$$H_4(x) = 16x^4 - 48x^2 + 12,\ H_5(x) = 32x^5 - 160x^3 + 120x, \ldots .$$

Рекурентна формула:

$$H_{n+1} = 2xH_n - 2nH_{n-1},$$

де $H_0 = 1,\quad H_{-1} = 0$ та $\left\|H_n\right\|^2 = 2^n n!\sqrt{\pi}$.

6). $H = L_2(0,2\pi),\quad \rho(x) \equiv 1\quad f(x) = f(x+2\pi)$ ($f(x)$ – 2π - періодичні функції), скалярний добуток $(u,v) = \int\limits_0^{2\pi} u(x)v(x)dx$. За систему ортогональних функцій вибираємо **тригонометричну систему**

$$\varphi_0 = \frac{1}{\sqrt{2\pi}},\qquad \varphi_{2k-1}(x) = \frac{1}{\sqrt{\pi}}\cos(kx),\qquad \varphi_{2k}(x) = \frac{1}{\sqrt{\pi}}\sin(kx),\ k = \overline{0,n},$$

де

$$\left\|\varphi_m\right\|^2 = 1,\ m = \overline{0,2n}.$$

Елемент найкращого середньоквадратичного наближення є тригонометричним поліномом n – го степеня

$$\Phi_0(x) = T_n(x) = a_0 + \sum_{k=1}^{n}(a_k \cos(kx) + b_k \sin(kx)).$$

7). $H = l_2, \quad x_i = i,\ i = \overline{0,N}, \quad (u,v) = \dfrac{1}{N+1}\sum_{i=0}^{N} u_i v_i$, (функція задана таблично).

За систему ортогональних функцій вибираємо систему поліномів **Чебишова дискретного аргументу** $\varphi_k(x) = p_k^{(N)}(x), \quad k = \overline{0,m} \quad (m \le N)$, яка задається формулою

$$p_k^{(N)}(x) = \sum_{j=0}^{k} \frac{(-1)^j C_k^j C_{k+j}^j}{N^{(j)}} x^{(j)},$$

де $x^{(j)} = x(x-1)...(x-j+1)$ - факторіальний поліном.

Рекурентна формула:

$$\frac{(m+1)(N-m)}{2(2m+1)} p_{m+1}^{(N)} = \left(\frac{N}{2} - x\right) p_m^{(N)} - \frac{m(N+m+1)}{2(2m+1)} p_{m-1}^{(N)}, \quad p_0^{(N)} = 1, \quad p_{-1}^{(N)} = 0.$$

$$\left\| p_k^{(N)} \right\|^2 = \frac{(N-k)!(N+k+1)!}{(N!)^2(2k+1)}.$$

(У випадку, якщо задані вузли $t_i = t_0 + ih, i = \overline{0,N}$, то робимо заміну $x_i = \dfrac{t_i - t_0}{h} = i$).