

Л.Л. Гарт, А.С. Каманцев

Дніпровський національний університет імені Олеся Гончара

АСИМПТОТИЧНИЙ АНАЛІЗ СКЛАДНОСТІ ДЕЯКИХ АЛГОРИТМІВ НАБЛИЖЕННЯ ФУНКЦІЇ

Досліджено асимптотичну складність найбільш поширених алгоритмів наближення функції на відрізьку. Під час дослідження алгоритмізовано та програмно реалізовано чотири методи наближення функції за допомогою мови програмування Python. Проведено серію обчислювальних експериментів, виконано порівняльний аналіз результатів роботи алгоритмів на конкретному прикладі, надано науково-методичні рекомендації щодо підвищення точності наближення.

Ключові слова: наближення функції, інтерполяційні многочлени Лагранжа та Ньютона, многочлен найкращого середньоквадратичного наближення, інтерполяційні кубічні сплайни, асимптотичний аналіз складності, похибка, збіжність.

L.L. Hart, A.S. Kamantsev

Oles Honchar Dnipro National University

ASYMPTOTIC ANALYSIS OF THE COMPLEXITY OF SOME ALGORITHMS FOR THE APPROXIMATION OF A FUNCTION

The article is devoted to the study of the asymptotic complexity of algorithms of the wide known methods for approximating functions of one variable on a given interval: the interpolation method using the Lagrange polynomial, the Newton's polynomial, cubic splines and the least squares method of the best mean square approximation of a function. A software implementation of the named methods is developed using the Python programming language. A series of computational experiments is carried out, and a qualitative comparative analysis of the efficiency of the implemented algorithms is given on the example of approximating a specific function at a given interval.

Asymptotic analysis is one of the types of algorithms' analysis. The purpose of asymptotic analysis is to compare the time and other resources spent by various algorithms designed to solve the same problem on a large amount of input data. Estimating the complexity function used in asymptotic analysis, which is called the complexity of the algorithm, allows us to determine how fast the complexity of the algorithm increases with raising data volume.

The problem of finding an unknown function or simplifying a difficult function on the information of some its values has applications in many practical problems. As the volume of data required for processing increases with the development of technologies, the issue of achieving the minimum practical complexity while keeping the accuracy of the function approximation algorithms is very important.

Keywords: function approximation, Lagrange and Newton's interpolation polynomials, best mean square approximation polynomial, cubic interpolation splines, asymptotic analysis of complexity, error, convergence.

Л.Л. Гарт, А.С. Каманцев

Дніпровський національний університет імені Олеся Гончара

АСИМПТОТИЧЕСКИЙ АНАЛИЗ СЛОЖНОСТИ НЕКОТОРЫХ АЛГОРИТМОВ ПРИБЛИЖЕНИЯ ФУНКЦИИ

Исследована асимптотическая сложность наиболее используемых алгоритмов приближения функции на отрезке. В ходе исследования алгоритмизированы и программно реализованы четыре метода приближения функции с помощью языка программирования Python. Проведена серия вычислительных экспериментов, выполнен сравнительный анализ результатов работы алгоритмов на конкретном примере, даны научно-методические рекомендации по повышению точности приближения.

Ключевые слова: приближение функции, интерполяционные многочлены Лагранжа и Ньютона, многочлен наилучшего среднеквадратичного приближения, интерполяционные кубические сплайны, асимптотический анализ сложности, погрешность, сходимость.

Вступ. Під час дослідження різних процесів в природі і суспільстві інформація, яку отримують в результаті вимірів їх стану, має, як правило, дискретне представлення у вигляді масивів числових даних про функціональні залежності, що характеризують ці процеси. Труднощі, а часто і неможливість використання масивів виникають під час розв'язання таких прикладних задач, як математичне моделювання та прогнозування, визначення функціональної закономірності природи досліджуваного процесу (наприклад, у вигляді емпіричної формули), під час розв'язання проблем економного зберігання великих за обсягом масивів даних або швидкісною передачею їх по каналах зв'язку, а також під час відновлення значень функціональної залежності на «неосвітлених вимірами» ділянках та ін. [1]. Ці труднощі долають шляхом заміни (стиснення) великих масивів аналітичними виразами з невеликою кількістю параметрів. Для розв'язання цієї проблеми застосовують засоби аналітичної обробки даних на основі методів і алгоритмів апроксимації (наближення) функцій, таких як апроксимація поліномами, раціональними дробами, експонентами, сплайнами, сплесками (вейвелет-функціями), фрактальними методами та ін. В роботі [2] ефективність зазначених вище методів апроксимації продемонстровано на ряді прикладних задач: навігація автономно рухомих апаратів, неруйнівний контроль, реографія підшлункової залози, тепломасообмін, конструювання гібридних дзеркальних антен, апроксимація атмосферних характеристик та ін.

Апроксимація (лат. *proxima* – найближча) функції – це, як відомо, подання довільної складної функції $f(x)$ простою і зручною для практичного використання функцією $F(x)$ таким чином, щоб відхилення $F(x)$ від $f(x)$ в області її завдання було найменшим за певним критерієм наближення. Функції $F(x)$ мають при цьому назву функцій апроксимації.

Історично розкладання функцій в ряд Тейлора

$$f(x) \approx f(x_0) + \sum_{i=1}^N \frac{f^{(i)}(x_0)}{i!} (x - x_0)^i$$

стало одним з перших методів наближення функцій в околі точки x_0 . Перший член $f(x_0)$ ряду являє собою відлік функції в точці x_0 і грубе наближення до значень функції в околі цієї точки. Всі інші члени ряду деталізують значення функції в околі точки x_0 за інформацією в сусідніх точках і тим точніше наближають суму ряду до значень функції, чим більше членів суми беруть участь в наближенні, з одночасним розширенням інтервалу околу точного наближення. Наближені формули обчислення різноманітних функцій (таких, як корінь) або констант (таких, як π) були відомі ще в давнину. Початком сучасної теорії наближення прийнято вважати роботу П.Л. Чебишева 1857 року, присвячену поліномам, що найменш відхиляються від нуля (зараз їх називають поліномами Чебишева першого роду) [3].

Визначитися з вибором ефективного алгоритму серед існуючих, що побудовані для розв'язання тої чи іншої проблеми, дозволяє складність алгоритму – кількісна характеристика, що відображує споживані алгоритмом ресурси під час його виконання (в основному, час виконання і простір пам'яті) [4]. Одним з видів аналізу складності алгоритмів є їх асимптотичний аналіз. Метою асимптотичного аналізу є порівняння витрат часу та інших ресурсів різноманітними алгоритмами, що призначені для вирішення однієї задачі при великому обсязі вхідних даних. Використовувана у асимптотичному аналізі оцінка функції трудомісткості, що називається складністю алгоритму, надає можливість визначити, як швидко зростає трудомісткість алгоритму зі збільшенням об'єму даних [5]. Оскільки з розвитком технологій збільшуються об'єми даних, що необхідно обробити, питання досягнення мінімальної практичної трудомісткості при збереженні точності алгоритмів наближення функції є дуже актуальною проблемою.

Постановка задачі. Задача апроксимації функцій, що розглядається в даній роботі, виникає під час розв'язання багатьох практичних задач, які передбачають обчислення значень деяких функціональних залежностей $y = f(x)$ на заданому відрізку $[a, b]$. При цьому, зазвичай, мають місце дві ситуації:

1) явна залежність між x та y на $[a, b]$ відсутня, а є тільки таблиця експериментальних даних $\{x_i; y_i\}$, $i=0, 1, \dots, n$ ($n \in \mathbb{N}$), і виникає необхідність визначення $y = f(x)$ на інтервалі $(x_i, x_{i+1}) \subset [a, b]$ ($i=0, 1, \dots, n-1$). До цієї задачі відноситься також уточнення таблиць експериментальних даних;

2) залежність $y = f(x)$ відома і неперервна, але настільки складна, що не придатна для практичних розрахунків. Ставиться задача спрощення обчислення значень функції $f(x)$ або її характеристик ($f'(x)$, $\max f(x)$, $\int_a^b f(x) dx$

та ін. [6]). Тому, з точки зору економії часу і матеріальних ресурсів, приходять до необхідності побудови деякої іншої функціональної залежності

$y = F(x)$, яка була б близька до $f(x)$ за основними її параметрами, але більш проста і зручна в реалізації під час подальших розрахунків, тобто ставиться задача про наближення (апроксимацію) функції $f(x)$ в області її визначення.

Метою даної роботи є алгоритмізація найбільш поширених методів наближення функції $f(x)$ (інтерполювання за допомогою многочленів Лагранжа, Ньютона, кубічних сплайнів, а також метод найменших квадратів найкращого середньоквадратичного наближення функції), програмна реалізація розроблених алгоритмів та їх порівняльний аналіз за точністю й асимптотичною складністю.

Методи розв'язання. Основний підхід до розв'язання задачі апроксимації функції $f(x)$ полягає в тому, що апроксимуючу функцію $F(x)$ вибирають залежною від деяких вільних параметрів, тобто $F(x) \equiv F(x, a_1, a_2, \dots, a_m)$, де $a_i = \text{const}$, $i = \overline{1, m}$ ($m \in \mathbb{N}$). Значення параметрів a_i , $i = \overline{1, m}$ вибирають з певних умов близькості для $f(x)$ і $F(x)$. В залежності від способу підбору параметрів, отримують різні види апроксимації.

В даний час на практиці добре вивчена і широко застосовується лінійна апроксимація, за якій функцію $F(x)$ вибирають лінійно залежною від числових параметрів a_i , $i = \overline{1, m}$ у вигляді так званого узагальненого многочлена

$$F(x) = a_1 \varphi_1(x) + a_2 \varphi_2(x) + \dots + a_m \varphi_m(x) = \sum_{i=1}^m a_i \varphi_i(x), \text{ де } \{\varphi_i(x)\}_{i=1, m} - \text{деяка вибрана}$$

лінійно-незалежна система базисних функцій. Важливим є те, щоб ця система була повною, тобто забезпечувала апроксимацію функції $f(x)$ за допомогою $F(x)$ із заданою точністю на всіх інтервалах $[a, b]$ визначення $y = f(x)$. Для більшості практичних задач найбільш зручною є така система базисних функцій $\{\varphi_i(x)\}_{i=1, m}$, що в підсумку приводить до апроксимації звичайними алгебраїчними многочленами.

Різні методи апроксимації функцій мають власні специфічні властивості і характеристики, такі як збіжність методу, стійкість обчислювального процесу, похибка апроксимації та необхідна кількість арифметичних дій, за якими оцінюють ефективність числового методу [7].

Більшість наближених методів розв'язання задач математичного аналізу, диференціальних рівнянь, математичної фізики так чи інакше пов'язана із апроксимацією функцій. Типовою серед вказаних задач є задача інтерполяції функцій. Нехай на відрізку $[a, b]$ задано $n+1$ впорядкованих не обов'язково рівновіддалених точок (вузли інтерполяції)

$$a \leq x_0 < x_1 < x_2 < \dots < x_n \leq b, \quad (1)$$

а також значення функції $f(x)$ у цих точках:

$$f(x_0) = f_0, \quad f(x_1) = f_1, \quad \dots, \quad f(x_n) = f_n.$$

Задача інтерполявання функції $f(x)$ на $[a, b]$ за системою вузлів $\{x_i\}_{i=0,n}$, як відомо, полягає у побудові такої функції $F(x)$, що належить відомому класу функцій і набуває у вузлах інтерполяції ті ж значення, що і $f(x)$:

$$F(x_0) = f_0, \quad F(x_1) = f_1, \quad \dots, \quad F(x_n) = f_n. \quad (2)$$

Основна мета інтерполяції – отримати швидкий (економічний) алгоритм обчислення значень $f(x)$ для точок $x \in [a, b]$, які не містяться в таблиці даних.

Класичним методом розв'язання вказаної задачі є побудова інтерполяційного алгебраїчного многочлена, який можна подати у різних формах запису, наприклад [7]:

– форма Лагранжа

$$L_n(x) = \sum_{i=0}^n f(x_i) \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j}, \quad a \leq x \leq b; \quad (3)$$

– форма Ньютона

$$P_n(x) = f(x_0) + f(x_0, x_1)(x - x_0) + f(x_0, x_1, x_2)(x - x_0)(x - x_1) + \dots + f(x_0, x_1, \dots, x_n)(x - x_0)(x - x_1) \dots (x - x_{n-1}), \quad a \leq x \leq b, \quad (4)$$

де $f(x_0, x_1, \dots, x_k)$ – поділені різниці порядку k , що визначаються за допомогою поділених різниць попереднього, $(k-1)$ -го порядку за формулою

$$f(x_0, x_1, \dots, x_k) = \frac{f(x_1, x_2, \dots, x_k) - f(x_0, x_1, \dots, x_{k-1})}{x_k - x_0}, \quad k = 1, 2, \dots, n.$$

Даний підхід цілком виправданий, оскільки він ґрунтується на відомій теоремі Вейерштрасса, яка стверджує, що довільну неперервну на відрізку функцію можна рівномірно із довільною точністю наблизити многочленом. Проте виявляється, що на практиці межі застосування інтерполяційного многочлена на всьому відрізку $[a, b]$ досить обмежені. Справа в тому, що часто побудова інтерполяційного полінома n -го степеня за сукупністю всіх вузлів дає незадовільний результат. В цьому випадку весь відрізок інтерполявання можна розбити на часткові відрізки $[x_{i-1}, x_i]$, $i = 1, 2, \dots, N$ ($x_0 = a$, $x_N = b$) і на кожному з них побудувати свій інтерполяційний многочлен, наклавши умову, щоб у точках «зшивання» значення цих поліномів співпадали. Така інтерполяція називається кусково-поліноміальною і, зрозуміло, що вона використовується для побудови часткових поліномів невеликих степенів. У випадках, коли в результаті наближення необхідно отримати диференційовану функцію, висувають ще вимогу співпадання значень похідних у точках «зшивання» для кожного часткового многочлена. Найбільш широке практичне застосування, в силу їх простоти, знайшли інтерполяційні кубічні сплайни [8].

Кубічну сплайн-функцію $S(x)$ на кожному з часткових відрізків $[x_{i-1}, x_i]$ ($i = 1, 2, \dots, N$) шукають у вигляді полінома третього степеня

$$S_i(x) = a_i + b_i(x - x_i) + c_i \frac{(x - x_i)^2}{2} + d_i \frac{(x - x_i)^3}{6}, \quad x_{i-1} \leq x \leq x_i; \quad (5)$$

де a_i, b_i, c_i, d_i – коефіцієнти, що підлягають визначенню і мають такий зміст:

$$a_i = S_i(x_i), \quad b_i = S'_i(x_i), \quad c_i = S''_i(x_i), \quad d_i = S'''_i(x_i); \quad (6)$$

тут $S'_i(x) = b_i + c_i(x - x_i) + d_i \frac{(x - x_i)^2}{2}$, $S''_i(x) = c_i + d_i(x - x_i)$, $S'''_i(x) = d_i$.

З умов інтерполювання $S(x_i) = f(x_i)$, $i = \overline{1, N}$ відразу знаходять

$$a_i = f(x_i), \quad i = \overline{1, N}.$$

Умова неперервності сплайна $S(x)$ приводить до співвідношень

$$S_i(x_i) = S_{i+1}(x_i), \quad i = \overline{1, N-1},$$

або, з урахуванням (5), (6),

$$a_i = a_{i+1} + b_{i+1}(x_i - x_{i+1}) + \frac{c_{i+1}}{2}(x_i - x_{i+1})^2 + \frac{d_{i+1}}{6}(x_i - x_{i+1})^3, \quad i = \overline{1, N-1}, \quad (7)$$

де додатково позначено $a_0 = f(x_0)$. Якщо для скорочення запису ввести крок $h_i = x_i - x_{i-1}$ ($i = 1, 2, \dots, N$) між сусідніми вузлами, то рівняння (7) набудуть вигляду

$$h_i b_i - \frac{h_i^2}{2} c_i + \frac{h_i^3}{6} d_i = f_i - f_{i-1}, \quad i = \overline{1, N}. \quad (8)$$

Умова $S'_i(x) = S'_{i+1}(x)$, $i = \overline{1, N-1}$ неперервності першої похідної сплайна $S(x)$ приводить до рівнянь

$$h_i c_i - \frac{h_i^2}{2} d_i = b_i - b_{i-1}, \quad i = \overline{2, N}; \quad (9)$$

умова $S''_i(x) = S''_{i+1}(x)$, $i = \overline{1, N-1}$ неперервності другої похідної сплайна – до рівнянь

$$h_i d_i = c_i - c_{i-1}, \quad i = \overline{2, N}. \quad (10)$$

Отже, співвідношення (8)–(10) утворюють систему $3N - 2$ лінійних алгебраїчних рівнянь відносно $3N$ невідомих коефіцієнтів b_i, c_i, d_i , $i = \overline{1, N}$. Два додаткових рівняння задають з урахуванням обмежень у крайніх точках відрізка $[a, b]$. У роботі використано стратегію задання значень другої похідної $S''(a)$ і $S''(b)$, що приводить до побудови єдиного сплайна із заданою кривизною в крайніх точках [8]. Зокрема, якщо відомі значення $f''(a)$ і $f''(b)$, то природньо вимагати, щоб $S''(a) \equiv S''_1(x_0) = f''(a)$, $S''(b) \equiv S''_N(x_N) = f''(b)$, тобто

$$c_1 - h_1 d_1 = f''(a), \quad c_N = f''(b).$$

Умова $c_1 - h_1 d_1 = f''(a)$, вочевидь, співпадає з рівнянням (10) при $i = 1$, якщо покласти $c_0 = f''(a)$. Таким чином, для знаходження коефіцієнтів b_i, c_i, d_i , $i = \overline{1, N}$ сплайна $S(x)$ має місце замкнена система лінійних алгебраїчних рівнянь (СЛАР) (8), (9),

$$h_i d_i = c_i - c_{i-1}, \quad i = \overline{1, N}; \quad c_0 = f''(a), \quad c_N = f''(b). \quad (11)$$

Після елементарних перетворень над рівняннями (8), (9), (11) одержують систему рівнянь відносно коефіцієнтів c_i ($i = \overline{1, N-1}$):

$$h_i c_{i-1} + 2(h_i + h_{i+1})c_i + h_{i+1}c_{i+1} = 6 \left(\frac{f_{i+1} - f_i}{h_{i+1}} - \frac{f_i - f_{i-1}}{h_i} \right), \quad i = \overline{1, N-1}; \quad (12)$$

$$c_0 = f''(a), \quad c_N = f''(b).$$

В силу діагонального переважання система (12) має єдиний розв'язок, який зазвичай одержують із використанням методу прогонки [7]. За знайденими коефіцієнтами c_i ($i = \overline{1, N-1}$) решту коефіцієнтів сплайна визначають так:

$$d_i = \frac{c_i - c_{i-1}}{h_i}, \quad b_i = \frac{h_i}{2} c_i - \frac{h_i^2}{3!} d_i + \frac{f_i - f_{i-1}}{h_i}, \quad i = \overline{1, N}.$$

Як зауваження слід відмітити, що серед всіх двічі диференційованих на відрізьку $[a, b]$ функцій, що інтерполюють функцію $f(x)$ за заданою сукупністю точок, кубічний сплайн $S(x)$ має найменшу осциляцію [8].

Інтерполяція таблично заданої функції $f(x)$ на відрізьку $[a, b]$ іноді буває практично неможлива, а у випадку слабо (стохастично або кореляційно) зв'язаних величин або ж неточних даних вона втрачає сенс. У таких випадках застосовують апроксимацію даних такою функцією простого вигляду, яка б вірно відображала характер залежності $y = f(x)$ в цілому і не обов'язково співпадала з нею на заданій сукупності точок. Найбільш ефективним методом апроксимації табличних даних є метод найменших квадратів, згідно з яким апроксимаційна функція вибирається так, щоб сума квадратів відхилень її від заданих значень була мінімальною. Під час поліноміальної апроксимації серед усіх алгебраїчних многочленів степеня $m \leq n$ вигляду

$$\Phi_m(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_{m-1} x^{m-1} + a_m x^m, \quad (13)$$

де $a_i \in \mathbb{R}$, $i = \overline{0, m}$, шукають такий, який надавав би мінімальне значення сумі

$$S = \sum_{i=0}^n [f(x_i) - \Phi_m(x_i)]^2. \quad (14)$$

Із використанням умов оптимальності для функції $S \equiv S(a_0, a_1, \dots, a_m)$ багатьох змінних отримують систему рівнянь відносно коефіцієнтів a_i ($i = \overline{0, m}$):

$$\begin{cases} c_0 a_0 + c_1 a_1 + c_2 a_2 + \dots + c_m a_m = d_0 \\ c_1 a_0 + c_2 a_1 + c_3 a_2 + \dots + c_{m+1} a_m = d_1, \\ c_2 a_0 + c_3 a_1 + c_4 a_2 + \dots + c_{m+2} a_m = d_2, \\ \dots \\ c_m a_0 + c_{m+1} a_1 + c_{m+2} a_2 + \dots + c_{2m} a_m = d_m. \end{cases} \quad (15)$$

де $c_j = \sum_{i=0}^n x_i^j$, $j = \overline{0, 2m}$; $d_k = \sum_{i=0}^n f(x_i)x_i^k$, $k = \overline{0, m}$. Матриця системи (15) симетрична і невироджена, отже система має єдиний розв'язок $a_0^*, a_1^*, \dots, a_m^*$, який доставляє мінімум величині (14) і утворює за формулою (13) алгебраїчний многочлен степеня m , що надає на відрізку $[a, b]$ найкраще середньоквадратичне наближення функції $f(x)$, заданої таблично.

Чисельний експеримент. Для проведення чисельного експерименту кожний із зазначених вище методів наближення функцій було реалізовано програмно мовою програмування Python з використанням бібліотек numpy і plotly. Розрахунки проводились на основі таких даних:

$$f(x) = 3 \cdot \cos x + 3,5 \cdot 3^{-x}, \quad [a, b] = [1,3; 5]; \quad (16)$$

$$x_k = a + k h, \quad k = 0, 1, \dots, n; \quad h = \frac{b-a}{n}; \quad (17)$$

$$\tilde{x}_j = a + (j-1) \frac{h}{5}, \quad j = 0, 1, \dots, 5n+3. \quad (18)$$

Задача чисельного експерименту полягала у наближенні функції (16) за системою заданих вузлів x_k ($k = 0, 1, \dots, n$) виду (17), де n – фіксоване натуральне число. Контроль якості наближення здійснювався шляхом обчислення значень функції $f(x)$ як у вузлових, так і деяких інших точках з області визначення функції, зокрема, у точках \tilde{x}_j ($j = 0, 1, \dots, 5n+3$) виду (18). Для досягнення найбільшої точності обчислень під час реалізації методу найкращого середньоквадратичного наближення степінь m апроксимуючого многочлена $\Phi_m(x)$ було обрано рівною степені n інтерполяційного многочлена функції $f(x)$. Результати застосування кожного з вищеозначених методів наближення функції $f(x)$ для випадку $n = 6$ подано на рисунках 1, 2.

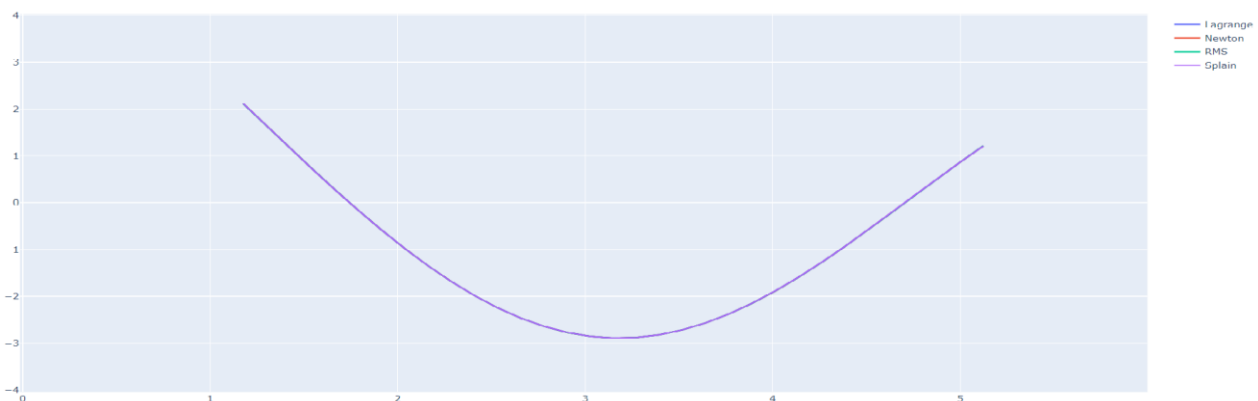


Рис. 1. Графічне зображення наближення функції $f(x)$ на проміжку $[1,17, 5,12]$ методами інтерполяції Лагранжа, Ньютона, найменших квадратів та кубічним сплайном.

Числові дані, за якими було побудовано наведені на рис. 1 графіки, подано

на рис. 2, де використано такі позначення: x – точки, в яких обчислено значення $f(x)$ та функцій, що наближають її за системою вузлів (17); **Original** – точні значення функції $f(x)$; **Splain** та **d Splain** – значення інтерполяційного кубічного сплайну та похибка наближення відповідно; **Lagrange** та **d Lagrange** – значення інтерполяційного многочлена Лагранжа та похибка наближення відповідно; **Newton** та **d Newton** – значення інтерполяційного многочлена Ньютона та похибка наближення відповідно; **RMS** та **d RMS** – значення многочлена найкращого середньоквадратичного наближення функції $f(x)$ та похибка такого наближення відповідно.

З отриманих результатів видно, що похибки наближення незначні і у вузлових точках майже дорівнюють нулю (з точністю не менше 9 десяткових знаків), що підтверджує коректність проведених розрахунків. Похибка наближення під час екстраполявання більша, порівняно з похибкою інтерполявання, для усіх розглянутих методів. Що стосується екстраполявання кубічним сплайном, то мала похибка у відповідних точках пояснюється вдалим збігом сплайну і наближуваної функції у розглядуваному прикладі.

X	Original	Splain	d Splain	RMS	d RMS	Lagrange	d Lagrange	Newton	d Newton
1.176667	2.112864	2.114380	-1.52e-03	2.110661	2.20e-03	2.113376	-5.12e-04	2.113376	-5.12e-04
1.300000	1.641590	1.641590	2.22e-16	1.641590	-3.00e-11	1.641590	0.00e+00	1.641590	-1.00e-14
1.423333	1.173554	1.171999	1.55e-03	1.174413	-8.59e-04	1.173400	1.54e-04	1.173400	1.54e-04
1.546667	0.712294	0.709781	2.51e-03	0.713248	-9.53e-04	0.712149	1.46e-04	0.712149	1.46e-04
1.670000	0.261702	0.259112	2.59e-03	0.262387	-6.86e-04	0.261614	8.74e-05	0.261614	8.74e-05
1.793333	-0.174102	-0.175831	1.73e-03	-0.173784	-3.18e-04	-0.174135	3.26e-05	-0.174135	3.26e-05
1.916667	-0.590874	-0.590874	0.00e+00	-0.590874	1.52e-10	-0.590874	0.00e+00	-0.590874	6.66e-16
2.040000	-0.984359	-0.984592	2.33e-04	-0.984558	1.99e-04	-0.984348	-1.09e-05	-0.984348	-1.09e-05
2.163333	-1.350393	-1.350572	1.80e-04	-1.350659	2.66e-04	-1.350384	-8.51e-06	-1.350384	-8.51e-06
2.286667	-1.684998	-1.685155	1.57e-04	-1.685225	2.27e-04	-1.684996	-2.22e-06	-1.684996	-2.22e-06
2.410000	-1.984474	-1.984678	2.03e-04	-1.984597	1.23e-04	-1.984476	1.43e-06	-1.984476	1.43e-06
2.533333	-2.245480	-2.245480	0.00e+00	-2.245480	-3.17e-10	-2.245480	0.00e+00	-2.245480	1.78e-15
2.656667	-2.465105	-2.464297	-8.08e-04	-2.465003	-1.01e-04	-2.465099	-5.34e-06	-2.465099	-5.34e-06
2.780000	-2.640933	-2.639442	-1.49e-03	-2.640778	-1.54e-04	-2.640922	-1.11e-05	-2.640922	-1.11e-05
2.903333	-2.771097	-2.769627	-1.47e-03	-2.770948	-1.49e-04	-2.771083	-1.35e-05	-2.771083	-1.35e-05
3.026667	-2.854323	-2.853563	-7.60e-04	-2.854232	-9.08e-05	-2.854313	-9.90e-06	-2.854313	-9.90e-06
3.150000	-2.889959	-2.889959	0.00e+00	-2.889959	3.52e-10	-2.889959	0.00e+00	-2.889959	1.33e-15
3.273333	-2.878000	-2.878032	3.26e-05	-2.878094	9.44e-05	-2.878013	1.35e-05	-2.878013	1.35e-05
3.396667	-2.819095	-2.819022	-7.34e-05	-2.819256	1.61e-04	-2.819121	2.57e-05	-2.819121	2.57e-05
3.520000	-2.714547	-2.714670	1.23e-04	-2.714721	1.74e-04	-2.714578	3.05e-05	-2.714578	3.05e-05
3.643333	-2.566303	-2.566723	4.21e-04	-2.566421	1.19e-04	-2.566325	2.26e-05	-2.566325	2.26e-05
3.766667	-2.376924	-2.376924	0.00e+00	-2.376924	-2.20e-10	-2.376924	0.00e+00	-2.376924	2.22e-15
3.890000	-2.149560	-2.147656	-1.90e-03	-2.149405	-1.55e-04	-2.149526	-3.40e-05	-2.149526	-3.40e-05
4.013333	-1.887904	-1.883859	-4.04e-03	-1.887605	-2.98e-04	-1.887835	-6.93e-05	-1.887835	-6.93e-05
4.136667	-1.596137	-1.591113	-5.02e-03	-1.595773	-3.64e-04	-1.596048	-8.92e-05	-1.596048	-8.92e-05
4.260000	-1.278873	-1.274997	-3.88e-03	-1.278591	-2.82e-04	-1.278800	-7.26e-05	-1.278800	-7.26e-05
4.383333	-0.941090	-0.941090	-1.11e-16	-0.941090	7.38e-11	-0.941090	-1.11e-16	-0.941090	6.00e-15
4.506667	-0.588058	-0.584304	-3.75e-03	-0.588545	4.87e-04	-0.588194	1.37e-04	-0.588194	1.37e-04
4.630000	-0.225260	-0.220285	-4.98e-03	-0.226352	1.09e-03	-0.225579	3.18e-04	-0.225579	3.18e-04
4.753333	0.141685	0.145990	-4.30e-03	0.140111	1.57e-03	0.141210	4.76e-04	0.141210	4.76e-04
4.876667	0.507113	0.509541	-2.43e-03	0.505641	1.47e-03	0.506653	4.59e-04	0.506653	4.59e-04
5.000000	0.865390	0.865390	0.00e+00	0.865390	-1.04e-11	0.865390	0.00e+00	0.865390	2.44e-15
5.123333	1.211004	1.208559	2.44e-03	1.215062	-4.06e-03	1.212347	-1.34e-03	1.212347	-1.34e-03

Рис. 2. Числові результати наближення функції $f(x)$ на проміжку $[1.17, 5.12]$.

При змінненні числа n ($3 \leq n \leq 10$), що регламентує кількість вузлів інтерполяції на відрізку $[a, b]$, була отримана наступна залежність середніх квадратів відхилень наближених значень функції кожним з методів.

Середні квадрати відхилень наближених значень функції

n	ІМ Ньютона	ІМ Лагранжа	МНСН	Кубічний сплайн
3	0.3329	0.3329	0.2575	0.0004
4	5.96E-3	5.96E-3	1.17E-2	5.58E-05
5	1.61E-3	1.61E-3	1.54E-3	1.55E-05
6	3.37E-06	3.37E-06	3.10E-05	5.80E-06
7	9.80E-07	9.80E-07	5.82E-07	2.62E-06
8	1.63E-09	1.63E-09	1.48E-09	1.34E-06
9	5.31E-10	5.31E-10	9.21E-11	7.43E-07
10	2.36E-13	2.36E-13	2.72E-10	4.41E-07

У табл. 1 використані такі скорочення: ІМ – інтерполяційний многочлен, МНСН – многочлен найкращого середньоквадратичного наближення. Отже, для розглядуваних функції $f(x)$ і сіток вузлів на $[a, b]$ можна стверджувати:

1. Метод найкращого середньоквадратичного наближення виявив найгіршу точність при $n < 7$.

2. При $n \geq 7$ точність методу найкращого середньоквадратичного наближення виявилась кращою, ніж точність наближення кубічними сплайнами.

3. Метод інтерполяції кубічними сплайнами виявся точнішим, ніж многочлени Лагранжа і Ньютона при $n < 6$.

4. При $n \geq 6$ многочлени Ньютона і Лагранжа виявили кращу точність, ніж метод інтерполяції кубічними сплайнами.

Важливим питанням, що впливає на можливість алгоритмів бути застосованими на практиці, є їх асимптотична складність. Проаналізуємо кількість виконаних операцій кожним з розглянутих методів наближення у залежності від значення n .

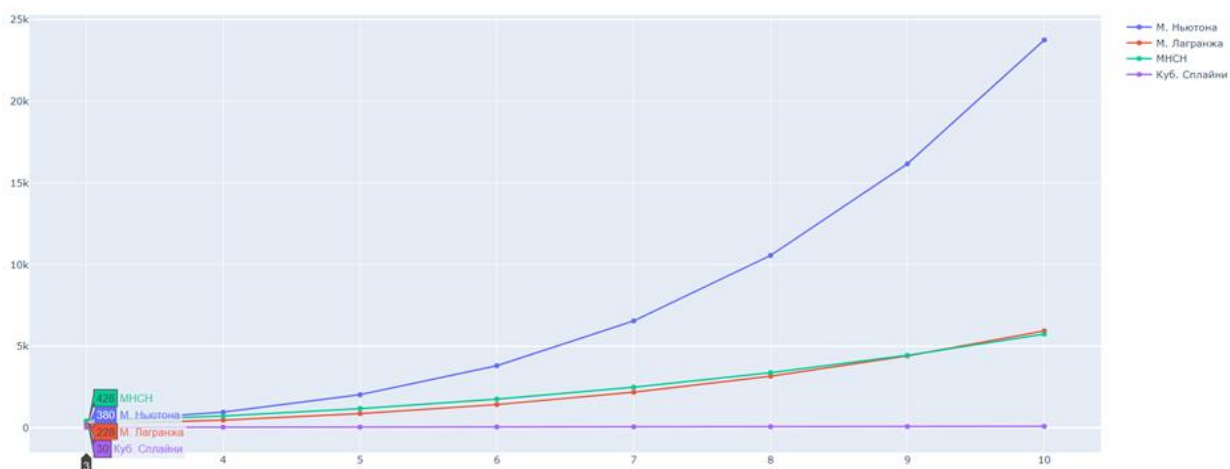


Рис. 3. Графічне зображення залежності від n кількості виконаних операцій розглянутими методами наближення функції $f(x)$.

Числові дані, за якими побудовано графіки, наведені на рис. 3, подано в табл. 2.

Таблиця 2

Кількість виконаних операцій

n	ІМ Ньютона	ІМ Лагранжа	МНСН	Кубічний сплайн
3	380	228	428	30
4	960	480	735	39
5	2030	870	1183	48
6	3808	1428	1764	57
7	6552	2184	2492	67
8	10560	3168	3381	76
9	16170	4410	4444	85
10	23760	5940	5746	94

Кількість точок, у яких проводився контроль якості наближення функції, було обрано близьку до $5n$ з метою демонстрації зміни кількості операцій, виконаних під час реалізації методу найменших квадратів, відносно операцій, виконаних під час розрахунків за допомогою многочленів Ньютона та Лагранжа при змінній n від 3 до 10.

Під час підрахунку виконаних арифметичних операцій були враховані такі особливості:

– операція піднесення до степеню не може бути виконана за сталий час, тому кожне піднесення до степеню k дає $\log_2 k$ операцій (за алгоритмом бінарного піднесення до степеню);

– СЛАР у методі найменших квадратів розв'язується за допомогою методу Гауса, що потребує порядку $(m+1)^3$ операцій;

– СЛАР під час інтерполювання кубічними сплайнами розв'язується за допомогою методу прогонки, що дає $2n-3$ операцій.

Виходячи з формул, за якими відбувається наближення функції кожним із реалізованих методів, можна надати такі асимптотичні оцінки трудомісткості обчислень у фіксованій точці $x \in [a, b]$:

– інтерполяційний многочлен Лагранжа – $O(n^2)$;

– інтерполяційний многочлен Ньютона – $O(n^3)$;

– многочлен найкращого середньоквадратичного наближення має складність попереднього обчислення коефіцієнтів $O(n^3)$ і $O(n \cdot \log n)$ для кожної обчислюваної точки;

– наближення інтерполяційним кубічним сплайном має складність попереднього обчислення коефіцієнтів $O(n)$; складність обчислення значення на проміжку є константною – $O(1)$; визначити, якому частковому проміжку належить задана точка, можна за $O(\log n)$ операцій, отже складність обчислення значення сплайну у кожній новій точці складає $O(\log n)$.

З цього випливає, що:

– найменшу асимптотичну складність має наближення функції інтерполяційним кубічним сплайном;

– найбільшу асимптотичну складність має інтерполяційний многочлен Ньютона;

– метод найменших квадратів має кращу асимптотичну складність, ніж методи інтерполювання за допомогою многочленів Лагранжа та Ньютона. Проте, якщо кількість точок, у яких необхідно наблизити функцію, не на багато перевищує n , то за малих n многочлени Лагранжа та Ньютона можуть дати кращі результати. Наприклад, якщо значення функції наближується у $5n$ точках, многочлен Лагранжа виявляється ефективнішим при $n < 9$, а многочлен Ньютона виявляється ефективнішим при $n \leq 3$. При збільшенні кількості точок для наближення значень $f(x)$, обмеження для n зменшується.

Отримана практично алгоритмічна складність узгоджується з наведеними асимптотичними оцінками.

Прикладна задача. Аналіз ефективності розглянутих алгоритмів наближення функції було проведено на прикладі розв'язання задачі відновлення функціональної залежності $b(x)$ питомої витрати палива від швидкості дизельного двигуна за результатами експериментальних вимірювань [9]. Результати вимірювань подані у табл. 3.

Таблиця 3

Питома витрата палива в залежності від швидкості двигуна

№	$x, \frac{\text{об.}}{\text{хв.}}$	$b, \frac{\text{г}}{\text{кВт} \cdot \text{год}}$	$r(x), \frac{\text{г}}{\text{кВт} \cdot \text{год}}$
1	1400	279,00	$\pm 3,73$
2	1600	265,03	$\pm 2,07$
3	1800	257,42	$\pm 2,55$
4	2000	255,50	$\pm 2,09$
5	2200	259,35	$\pm 4,24$
6	2400	264,22	$\pm 2,52$
7	2600	267,16	$\pm 1,80$
8	2800	268,60	$\pm 2,66$
9	3000	270,00	$\pm 2,97$
10	3200	273,49	$\pm 1,94$
11	3400	281,25	$\pm 3,16$

У табл. 3 використані такі позначення: x – швидкість двигуна (чотиритактний дизельний двигун польського виробництва типу ANDORIA 4СТi90-1BE6 з наддувом та проміжним охолоджувачем); b – питома витрата палива (тип масового датчика палива AMX 212F); $r(x)$ – похибка вимірювання. Експериментальні дані з номерами 1, 3, 5, 7, 9, 11 розглядалися як вихідні дані (вузли $\{x_i\}$ та значення $\{b_i\}$ наближуваної функції) в задачі інтерполювання. Результати застосування розробленого програмного продукту до наближення функції $b(x)$ за такими вихідними даними подано на рисунках 4, 5.

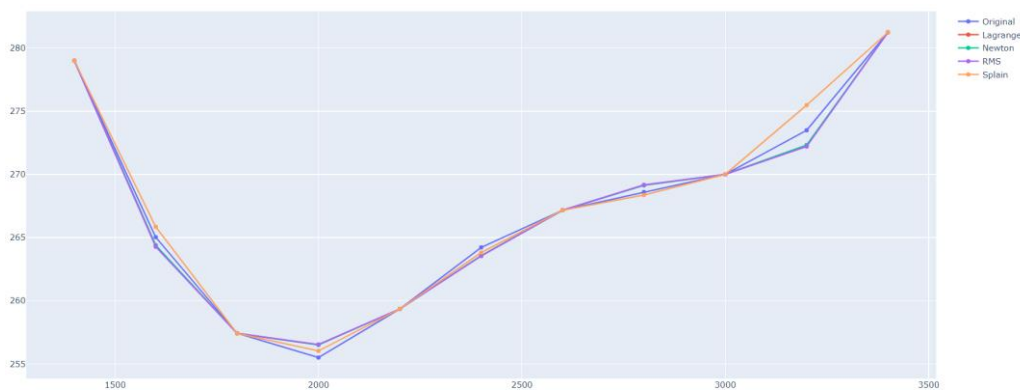


Рис. 4. Графічне зображення наближення функції $b(x)$ на проміжку $[1400, 3400]$ методами інтерполяції Лагранжа, Ньютона, найменших квадратів та кубічним сплайном.

Числові дані, за якими було побудовано наведені на рис. 4 графіки, подано на рис. 5.

X	Original	Splain	d Splain	RMS	d RMS	Lagrange	d Lagrange	Newton	d Newton
1400	279.000000	279.000000	0.00e+00	279.000000	-4.13e-10	279.000000	0.00e+00	279.000000	-5.68e-14
1600	265.030000	265.846518	-8.17e-01	264.272794	7.57e-01	264.381680	6.48e-01	264.381680	6.48e-01
1800	257.420000	257.420000	0.00e+00	257.420000	1.89e-09	257.420000	0.00e+00	257.420000	0.00e+00
2000	255.500000	256.028457	-5.28e-01	256.538834	-1.04e+00	256.502539	-1.00e+00	256.502539	-1.00e+00
2200	259.350000	259.350000	0.00e+00	259.350000	-3.46e-09	259.350000	0.00e+00	259.350000	0.00e+00
2400	264.220000	263.813158	4.07e-01	263.535598	6.84e-01	263.561523	6.58e-01	263.561523	6.58e-01
2600	267.160000	267.160000	0.00e+00	267.160000	3.16e-09	267.160000	0.00e+00	267.160000	5.68e-14
2800	268.600000	268.362661	2.37e-01	269.173678	-5.74e-01	269.137383	-5.37e-01	269.137383	-5.37e-01
3000	270.000000	270.000000	0.00e+00	270.000000	-1.43e-09	270.000000	0.00e+00	270.000000	0.00e+00
3200	273.490000	275.479799	-1.99e+00	272.204981	1.29e+00	272.313867	1.18e+00	272.313867	1.18e+00
3400	281.250000	281.250000	0.00e+00	281.250000	2.61e-10	281.250000	0.00e+00	281.250000	-2.27e-13

Рис. 5. Числові результати наближення функції $b(x)$ на проміжку $[1400, 3400]$.

З результатів роботи програми бачимо, що похибки отриманих наближень незначні і не перевищують похибок експериментальних вимірювань, наведених у табл. 1, тому якість наближення функціональної залежності питомої витрати палива $b(x)$ від швидкості x дизельного двигуна можна вважати задовільною [9]. Кількість виконаних арифметичних операцій і середній квадрат відхилень для кожного з реалізованих методів наближення функції $b(x)$ подані у табл. 4.

Таблиця 4

Показники ефективності та якості наближення функції $b(x)$

	ІМ Ньютона	ІМ Лагранжа	МНСН	Кубічний сплайн
N	840	360	996	53
ε	0.3210	0.3210	0.3728	0.4661

Тут N – кількість виконаних операцій, ε – середній квадрат відхилень розрахункових значень функції $b(x)$ від експериментальних даних. З отриманих результатів випливає, що для розв’язання цієї практичної задачі найкращу точність проявили методи інтерполяції Лагранжа і Ньютона, а найгіршу – інтерполювання кубічними сплайнами. Найбільшу кількість операцій зажадав

метод найкращого середньоквадратичного наближення, найменшу – метод інтерполювання кубічними сплайнами. Одержані висновки щодо показників ефективності та якості наближення функції $b(x)$ аналогічні висновкам щодо відповідних показників під час наближення аналітично заданої функції (16).

Аналіз отриманих результатів та висновки. Під час виконання роботи проведено алгоритмізацію та програмну реалізацію найбільш поширених методів наближення функції $f(x)$ на відрізку (інтерполювання за допомогою многочленів Лагранжа, Ньютона, кубічних сплайнів, а також методу найменших квадратів найкращого середньоквадратичного наближення функції), виконано порівняльний аналіз розроблених алгоритмів за точністю і асимптотичною складністю.

Підсумовуючи отримані результати, можна навести наступний алгоритм вибору методу інтерполювання функції $f(x)$ на заданому відрізку:

1. Якщо висока точність наближення не вимагається або n є малим, то доцільно використовувати інтерполювання за допомогою кубічних сплайнів;
2. Якщо n не дуже мале, алгоритмічна складність алгоритму не має суттєвого значення і вимагається висока точність наближення, то найкращим вибором буде інтерполяційний многочлен у формі Лагранжа;
3. Якщо n досить велике, алгоритмічна складність має значення і вимагається висока точність наближення, то доцільно використовувати метод найменших квадратів.

Бібліографічні посилання

1. **Каленчук-Порханова, А.А.** Наилучшая чебышёвская аппроксимация как эффективный способ обработки информации [Текст] / А.А. Каленчук-Порханова // Штучний інтелект, 2009, № 3, с. 53–63.
2. **Бердышев, В.И., Петрак, Л.В.** Аппроксимация функций, сжатие численной информации, приложения [Текст] / В.И. Бердышев, Л.В. Петрак. – Екатеринбург: УрОран, 1999. – 296 с.
3. **Rivlin, T.J.** Introduction to the Approximation of Functions [Text] / Theodor J. Rivlin. – Dover Publications, 2010. – 160 p.
4. **Wilf, H.S.** Algorithms and Complexity [Text] / Herbert S. Wilf. – A K Peters/CRC Press, 2002. – 219 p.
5. **Кормен, Т.Х.** Алгоритмы: построение и анализ [Текст] / Т.Х. Кормен, Ч.И. Лейзерсон, Р.Л. Ривест, К. Штайн. – М.: ООО “И.Д. Вильямс”, 2013. – 1328 с.
6. **Kiseleva, E.M.** On a problem of numerical simulating the derivative of discrete time series with approximate values [Text] / E.M. Kiseleva, L.L. Hart, P.A. Dovgay // Journal of Automation and Information Sciences, 2015, Vol. 47, Issue 12, p. 1–17.
7. **Самарский, А.А., Гулин, А.В.** Численные методы [Текст] / А.А. Самарский, А.В. Гулин. – М.: Наука, 1989. – 430 с.
8. **Довгий, Б.П.** Сплайн-функції та їх застосування [Текст] / Б.П. Довгий, А.В. Ловейкін, Є.С. Вакал, Ю.Є. Вакал. – К.: ВПЦ “Київський університет”, 2016. – 117 с.
9. **Stoeck, T.** Application of the polynomial interpolation method for determining performance characteristics of a diesel engine [Text] / T. Stoeck, K.F. Abamek // Metrology and Measurement Systems, 2014, Vol. 21, № 1, pp. 157–168.