7 МЕТОДИ ОБРОБКИ ЕКСПЕРИМЕНТАЛЬНИХ ДАНИХ

Дуже важливою задачею в процесі ідентифікації математичних моделей є обробка експериментальних даних, які отримані під час активних або пасивних ідентифікаційних експериментів. Для цього застосовуються різноманітні методи та алгоритми, вибір яких зумовлений видом об'єкта моделювання, моделі та наявними обчислювальними потужностями комп'ютерних засобів.

7.1 Інтерполяція

Мета *інтерполяції* (*interpolation*) – побудувати функцію $F(x)$, яка приймає в окремих точках $x_i[a,b]$ $(i = 0, 1, 2, ..., n)$ (вузли інтерполяції) значення

$$F(x_0) = y_0\,,\ \ F(x_1) = y_1,\ \ ..., F(x_i) = y_i,\ \ ..., F(x_n) = y_n, \tag{7.1}$$

що збігаються з раніше заданими (частіше за все, отриманими з експерименту) значеннями в цих точках невідомої функції $y = f(x)$. Геометрично це означає, що потрібно знайти криву $y = F(x)$ певного типу, яка проходить через систему точок $M(x_i, y_i)$ (i = 0, 1, 2,..., n) (рис. 7.1).

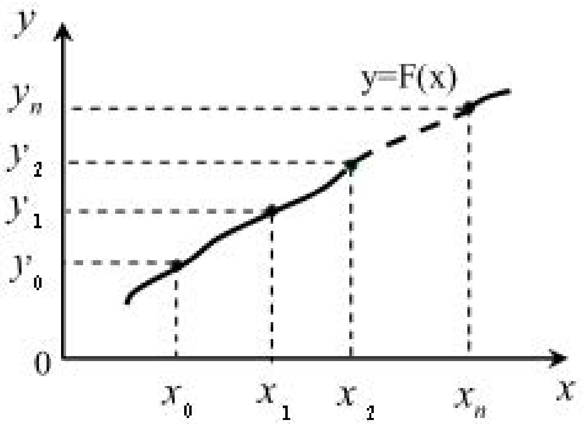

Рисунок 7.1 – Інтерполяція

В загальних випадках ця задача має нескінченну множину розв'язків чи зовсім не має розв'язку, але вона стає однозначною, якщо замість довільної функції $F(x)$ шукати поліном $P_n(x)$ степеня не вище n, який задовольняє умову (7.1), тобто

$$P_n(x_0) = y_0,\ \ P_n(x_1) = y_1\ ,...,\ \ P_n(x_i) = y_i\ ,...,\ \ P_n(x_n) = y_n.$$

Інтерполяційну формулу $y = F(x)$ використовують для наближеного обчислення значень функції $f(x)$ для $x \neq x_i$ $(i = 0, 1, 2, ..., n)$. Треба відзначити, що є інтерполяція в вузькому розумінні, коли $x \in [x_0, x_n]$, та *екстраполяція*, коли x знаходиться за межами інтервалу $[x_0, x_n]$ тобто $x < x_0$ чи $x > x_n$.

Говорячи про процедуру інтерполювання, обов'язково потрібно вказати обмеження, які накладаються на набір базових точок. Початкова сітка точок повинна описувати лише плавну функцію. Відповідно до умов конкретної задачі обов'язково повинні задаватися значення похідної функції у крайових точках вхідної сітки для отримання однозначного результату.

До основних методів інтерполяції належать:

1) *лінійна інтерполяція (linear interpolation)*. Найбільш простий і швидкий метод, в якому задані вузлові точки з'єднуються прямими;

2) *інтерполяція із застосуванням багаточлена (interpolation using polynomial)*. Використовується багаточлен n-го порядку, який в загальному випадку має вигляд

$$P(x) = P_n(x) = a_0 x^n + a_1 x^{n-1} + ... + a_{n-1} x + a_n,$$

де a_i, $i = \overline{0, n}$ – постійні коефіцієнти.

Всі методи знаходження інтерполяційного багаточлена зводяться до отримання постійних коефіцієнтів. До таких методів належать:

a. Інтерполяція різницевими методами;
b. Ермітова поліноміальна інтерполяція;
c. Інтерполяція Лагранжа;

3) *поліноміальна інтерполяція сплайнами (polynomial spline interpolation)*. Вузлові точки з'єднуються з використанням багаточлену заданого порядку, який обирається залежно від методу. До найбільш поширених методів інтерполяції сплайнами належать:

a. Класичні кубічні сплайни;
b. Ермітові сплайни;
c. B-сплайни;
d. Криві Без'є.

7.1.1 Різницеві методи

Існує багато кінцево-різницевих методів інтерполяції. Найбільш поширеним є метод Ньютона для інтерполяції "уперед" (метод Ньютона – Грегорі). Інтерполяційний поліном в цьому випадку має вигляд:

$$P_n(x) = C_0 + C_1(x - x_0) + C_2(x - x_0)(x - x_1) + ...$$

$$... + C_n(x - x_0)(x - x_1) \; ... \; (x - x_{n-1}).$$

Коефіцієнти C_i знаходять з рівнянь:

$$P_n(x_i) = y_i, \qquad i = 0, 1, 2, ..., n,$$

які дозволяють записати систему

$$\begin{cases} C_0 = y_0, \\ C_0 + C_1(x_1 - x_0) = y_1, \\ C_0 + C_1(x_2 - x_0) + C_2(x_2 - x_0)(x_2 - x_1) = y_2, \\ \ldots \\ C_0 + C_1(x_n - x_0) + ... + C_n(x_n - x_0)(x_n - x_1)...(x_n - x_{n-1}) = y_n. \end{cases} \tag{7.2}$$

Це лінійна система рівнянь з трикутною матрицею.

Якщо прийняти крок $x_{i+1} - x_i = h$, то в області змінювання $x \in [x_0, x_n]$ отримаємо одновимірну рівномірну сітку та зможемо скористатись різницевим зображенням системи (7.2), яке приводить до таких різницевих виразів для коефіцієнтів:

$$C_0 = y_0,$$

$$C_1 = \frac{y_1 - y_0}{h} = \frac{\Delta y_0}{h},$$

де Δy_0 – права різниця першого порядку в точці y_0;

$$C_2 = \frac{y_2 - 2y_1 + y_0}{2h^2} = \frac{\Delta^2 y_0}{2h^2},$$

де $\Delta^2 y_0$ – права різниця другого порядку;

$$C_j = \frac{\Delta^j y_0}{(j!)h^j},$$

де $\Delta^j y_0$ – права різниця j-го порядку.

Тоді

$$P_n(x) = y_0 + \frac{\Delta y_0}{1!h}(x - x_0) + \frac{\Delta^2 y_0}{2!h^2}(x - x_0)(x - x_1) + ... \tag{7.3}$$

$$... + \frac{\Delta^n y_0}{n!h^n}(x - x_0)(x - x_1) \; ... \; (x - x_{n-1}).$$

З практичної точки зору для визначення різниць вищих порядків використовують вираз:

$$\Delta^j y_i = \Delta(\Delta^{j-1} y_i) = \Delta^{j-1} y_{i+1} - \Delta^{j-1} y_i.$$

При $n=1$ з (7.3) отримуємо формулу для лінійної інтерполяції

$$P_1(x) = y_0 + \frac{\Delta y_0}{h}(x - x_0),$$

а якщо $n = 2$ – формулу параболічної чи квадратичної інтерполяції

$$P_2(x) = y_0 + \frac{\Delta y_0}{h}(x - x_0) \ + \ \frac{\Delta^2 y_0}{2h^2}(x - x_0)(x - x_1).$$

Якщо дана необмежена кількість значень функції x, то кількість n може бути будь-якою. Практично в цьому випадку n вибирають так, щоб різниця $\Delta^n y_i$ була постійною з заданою точністю. За початкове значення x_0 можна приймати будь-яке табличне значення аргументу x. Коли кількість значень функції скінченна, то кількість n обмежена та не може бути більше кількості значень функції y, зменшеної на одиницю.

Формула (7.3) носить назву першої інтерполяційної формули Ньютона. Цей вираз незручний для інтерполяції в околі останніх значень y. В цьому випадку, як правило, використовують другу інтерполяційну формулу Ньютона, яка отримана при використанні лівих різниць від останнього значення (x_n, y_n) (інтерполяція “назад”). Тоді інтерполяційний поліном має вигляд:

$$\begin{aligned} P_n(x) = C_0 + C_1(x - x_n) + C_2(x - x_n)(x - x_{n-1}) + \\ + C_3(x - x_n)(x - x_{n-1})(x - x_{n-2}) + \ \dots \\ \dots \ + C_n(x - x_n)(x - x_{n-1}) \ \dots \ (x - x_1). \end{aligned}$$

Коефіцієнти C_j визначаються таким чином:

$$C_0 = y_n$$

$$C_1 = \frac{\Delta y_{n-1}}{h} = \frac{\nabla y_n}{h},$$

де ∇y_n – ліва різниця першого порядку в точці y_n,

$$C_2 = \frac{\Delta^2 y_{n-2}}{2!h^2} = \frac{\nabla^2 y_n}{2h^2},$$

де $\nabla^2 y_n$ – ліва різниця другого порядку,

$$C_j = \frac{\Delta^j y_{n-j}}{j!h^j} = \frac{\nabla^j y_n}{j!h^j},$$

де $\nabla^j y_n$ – ліва різниця j-го порядку.

Кінцевий вираз для другої інтерполяційної формули Ньютона:

$$P_n(x) = y_n + \frac{\Delta y_{n-1}}{1!h}(x - x_n) + \frac{\Delta y_{n-2}^2}{2!h^2}(x - x_n)(x - x_{n-1}) + ...$$

$$+ \frac{\Delta^n y_0}{n!h^n}(x - x_n)(x - x_{n-1})...(x - x_1).$$

Інтерполяційні формули Ньютона можуть бути використані для екстраполяції функції. Якщо $x < x_0$, то зручно використовувати першу інтерполяційну формулу Ньютона, причому

$$\frac{x - x_0}{h} < 0.$$

Якщо $x > x_n$, то використовують другу інтерполяційну формулу Ньютона, де

$$\frac{x - x_n}{h} > 0.$$

Таким чином, перша інтерполяційна формула Ньютона, як правило, використовується для інтерполяції “уперед” та екстраполяції “назад”, а друга – для інтерполяції “назад” та екстраполяції “уперед”.

В формулах Ньютона використовують ліві та праві різниці. Використання центральних різниць для отримання інтерполяційних формул приводить до формул Гаусса, Стірлінга та Бесселя.

Треба відзначити, що центральні різниці використовуються не в звичайному вигляді, а шляхом застосування правих різниць при поступовому зсуві індексів вліво.

Ці формули зручно розглядати на $(2n+1)$ рівновіддалених вузлах інтерполяції

$$x_{-n}, x_{-(n-1)}, ..., x_{-1}, x_0, x_1, ..., x_{n-1}, x_n,$$

причому

$$\Delta x_i = x_{i+1} - x_i = h \qquad (i = -n, -(n-1), ..., n-1),$$

а для функції $y = f(x)$ відомі її значення в цих вузлах $y_i = f(x_i)$.

Потрібно побудувати поліном $P(x)$ степеня не вище $2n$ такий, що

$$P(x_i) = y_i.$$

Поліном $P(x)$ відшукується у вигляді:

$$P(x) = C_0 + C_1(x - x_0) + C_2(x - x_0)(x - x_1) + C_3(x - x_{-1})$$
$$(x - x_0)(x - x_1) + ... + C_{2n-1}(x - x_{-(n-1)})... \quad (7.4)$$
$$...(x - x_{-1})(x - x_0)(x - x_1)...(x - x_{n-1})(x - x_n).$$

Аналогічно інтерполяційним формулам Ньютона знаходимо

$$C_0 = y_0;$$

$$C_1 = \frac{\Delta y_0}{h};$$

$$C_2 = \frac{\Delta^2 y_{-1}}{2! h^2};$$

$$......$$

$$C_{2n-1} = \frac{\Delta^{2n-1} y_{-(n-1)}}{(2n-1)! h^{2n-1}};$$

$$C_{2n} = \frac{\Delta^{2n} y_{-n}}{(2n)! h^{2n}}.$$

Підставляючи знайдені значення коефіцієнтів в (7.4), отримаємо першу інтерполяційну формулу Гаусса, яка містить різниці (таблиця 7.1):

$$\Delta y_0,\ \Delta^2 y_{-1},\ \Delta^3 y_{-1},\ \Delta^4 y_{-2},\ \Delta^5 y_{-2},\ \Delta^6 y_{-2},\ ...$$

Аналогічно можна отримати другу інтерполяційну формулу Гаусса, яка містить центральні різниці:

$$\Delta y_{-1},\ \Delta^2 y_{-1},\ \Delta^3 y_{-2},\ \Delta^4 y_{-2},\ \Delta^5 y_{-3},\ \Delta^6 y_{-3},\ ...$$

Використавши середнє арифметичне першої та другої інтерполяційних формул Гаусса, отримаємо формулу Стірлінга. Ці формули дозволяють вивести інтерполяційну формулу Бесселя. Взагалі, використання інтерполяційних формул з центральними різницями доцільно всередині інтервалу, тоді як на його краях, як правило, використовують формули Ньютона. Застосування цих формул розглядається в таблиці 7.1.

Похибки інтерполяції для формул Ньютона можна оцінити відповідно для першої та другої формул як:

$$\Delta_n(x) = \frac{q(q-1)...(q-n)}{(n+1)!} \Delta^{n+1} y_0,$$

$$\Delta_n(x) = \frac{q(q+1)...(q+n)}{(n+1)!}\Delta^{n+1}y_n,$$

де $q = \dfrac{x - x_n}{h}$.

Таблиця 7.1 – Застосування різницевих формул інтерполяції

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$	**Примітки**
x_{-2}	y_{-2}		$\Delta^2 y_{-3}$		$\Delta^4 y_{-4}$	**Друга формула Ньютона**
		Δy_{-2}		$\Delta^3 y_{-3}$		
x_{-1}	y_{-1}		$\Delta^2 y_{-2}$		$\Delta^4 y_{-3}$	
		Δy_{-1}		$\Delta^3 y_{-2}$		**Формула Стірлінга**
x_0	y_0		$\Delta^2 y_{-1}$		$\Delta^4 y_{-2}$	**Формула Бесселя**
		Δy_0		$\Delta^3 y_{-1}$		
x_1	y_1		$\Delta^2 y_0$		$\Delta^4 y_{-1}$	
		Δy_1		$\Delta^3 y_0$		
x_2	y_2		$\Delta^2 y_1$		$\Delta^4 y_0$	**Перша формула Ньютона**
		Δy_2		$\Delta^3 y_1$		
x_3	y_3		$\Delta^2 y_2$		$\Delta^4 y_1$	

Для формули Стірлінга:

$$\Delta(x_n) = \frac{\Delta^{2n+1}y_{-(n-1)} + \Delta^{2n+1}y_{-n}}{2(2n+1)!}q(q^2-1)(q^2-2^2)\ \dots\ (q^2-n^2).$$

Для випадку нерівновіддалених значень аргументу можна отримати інтерполяційні формули, вживаючи визначення поділених різниць. Наприклад, відношення

$$[x_i, x_{i+1}] = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}$$

називається поділеною різницею першого порядку, а відношення

$$[x_i, x_{i+1}, x_{i+2}] = \frac{[x_{i+1}, x_{i+2}] \;-\; [x_i, x_{i+1}]}{x_{i+2} - x_i}$$

називається поділеною різницею другого порядку.

Поділені різниці порядку n отримуються з рекурентного відношення:

$$[x_i, x_{i+2},\; ...,\; x_{i+n}] = \frac{[x_{i+1},..., x_{i+n}] \;-\; [x_{i+1},..., x_{i+n-1}]}{x_{i+n} - x_i}.$$

Можна отримати інтерполяційну формулу Ньютона для нерівновіддалених значень аргументу:

$$P(x) = y_0 + [x_0, x_1](x - x_0) + [x_0, x_1, x_2](x - x_0)(x - x_1) + \;...$$

$$+ [x_0, x_1,..., x_n](x - x_0)(x - x_1)\; ...\; (x - x_{n-1}).$$

7.1.2 Інтерполяція за Лагранжем

Інтерполяція за Лагранжем (Lagrange interpolation) застосовується в загальному випадку для довільно розташованих вузлів.

Інтерполяційний поліном для методу Лагранжа поданий у вигляді:

$$P_n(x) = y_0 b_0(x) + y_1 b_1(x) + \;...\; + y_n b_n(x),$$

де всі $b_j(x)$ $(j=0,\dots, n)$ – поліноми степеня n, коефіцієнти яких можна знайти з допомогою $(n+1)$ рівняння:

$$P_n(x_i) = y_i,$$

внаслідок чого отримаємо систему рівнянь:

$$\begin{cases} y_0 b_0(x_0) + y_1 b_1(x_0) + \;...\; + y_n b_n(x_0) = y_0; \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ y_0 b_0(x_n) + y_1 b_1(x_n) + \;...\; + y_n b_n(x_n) = y_n. \end{cases}$$

Якщо значення $b_j(x_i)$ обирається так, що

$$b_j(x_i) = \begin{vmatrix} 1, & \text{при } i = j, \\ 0, & \text{при } i \neq j, \end{vmatrix}$$

то записана система рівнянь буде задовільна.

Ця умова означає, що будь-який поліном $b_j(x)$ дорівнює нулю при кожному x_i, крім такого, що дорівнює x_j. Тому в загальному випадку поліном $b_j(x)$ має такий вигляд:

$$b_j(x) = C_j(x - x_0)(x - x_1)...(x - x_{j-1})(x - x_{j+1})...(x - x_n).$$

Якщо $b_j(x_j)$ =1, то коефіцієнти C_j визначаються з виразу:

$$C_j = 1/(x_j - x_0)...(x_j - x_{j-1})(x_j - x_{j+1})...(x_j - x_n).$$

Для полінома, який шукаємо, отримаємо:

$$P_n(x) = \sum_{j=0}^{n} y_j \cdot$$

$$\cdot \frac{(x - x_0)(x - x_1)...(x - x_{j-1})(x - x_{j+1})...(x - x_n)}{(x_j - x_0)(x_j - x_1)...(x_j - x_{j-1})(x_j - x_{j+1})...(x_j - x_n)}.$$

Вводячи позначення

$$L_j(x) = (x - x_0)(x - x_1)...(x - x_{j-1})(x - x_{j+1})...(x - x_n),$$

записуємо

$$P_n(x) = \sum_{j=0}^{n} y_j \frac{L_j(x)}{L_j(x_j)}.$$

Треба відзначити дві головні властивості поліномів Лагранжа:

1) $\sum_{j=0}^{n} \frac{L_j(x)}{L_j(x_j)} = 1;$

2) якщо $P_n(x)$ лінійно залежить від y_j, то слушний принцип суперпозиції: інтерполяційний поліном суми декількох функцій дорівнює сумі інтерполяційних поліномів доданків.

Залишковий член інтерполяційної формули

Припускаючи вузли інтерполяції відмінними один від другого, а функцію $f(x)$ такою, що має неперервну похідну порядку $n+1$ на проміжку $[a, b]$, де розміщені вузли інтерполяції, можна записати залишковий член інтерполяційної формули

$$R_n(x) = f(x) - P_n(x)$$

на цьому проміжку у вигляді

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

де $\xi \in [\alpha_1, \alpha_2], \alpha_1 = \min(x, x_0, x_1, ..., x_n), \alpha_2 = \max(x,\ x_0, x_1, ..., x_n)$.

Тоді

$$| R_n(x) | \le \frac{M_{n+1}}{(n+1)!} | \omega_{n+1}(x) |,$$

де $M_{n+1} = \max\limits_{x \in [\alpha_1, \alpha_2]} \left| f^{(n+1)}(x) \right|$.

7.1.3 Сплайн-інтерполяція

Інтерполяція багаточленом Лагранжа або Ньютона на всьому відрізку з використанням великого числа вузлів інтерполяції часто призводить до неточного наближення, що пояснюється значним накопиченням похибок в процесі обчислень. Крім того, через розбіжність процесу інтерполяції збільшення числа вузлів не завжди приводить до підвищення точності. Для того, щоб уникнути великих похибок, увесь відрізок розбивають на часткові відрізки і на кожному з часткових відрізків приблизно замінюють функцію *f(x)* багаточленом невисокого степеня (так звана кусково-поліноміальна інтерполяція).

Одним із способів інтерполяції на всьому відрізку є інтерполяція за допомогою сплайн-функцій. ***Сплайн-функцією*** або ***сплайном*** називають кусково-поліноміальну функцію, що визначена на відрізку і має на цьому відрізку деяке число неперервних похідних.

Слово “сплайн” (англійське “spline”) означає гнучку лінійку, що використовується для проведення гладких кривих через задані точки площини. Основною перевагою сплайнів є можливість локально змінювати форму кривої на виділеному проміжку значень.

Коли відрізок $[a;b]$ досить великий, то не можна підвищувати точність інтерполяції за рахунок збільшення порядку інтерполяційного полінома. Це пов’язано з тим, що у полінома n-го порядку може бути $n-1$ точка екстремуму. При $n \to \infty$ графік полінома починає сильно коливатись і $\left| f(x) - P_n(x) \right|$ не наближається до нуля. Таке явище називають *феноменом Рунге*.

Тому більш перспективним є застосування кусково-поліноміальної інтерполяції, при якій апроксимувальна функція складається з окремих поліномів, як правило, однакового невеликого порядку, визначених кожен на своїй частині відрізка $[a;b]$.

Розглянемо інтерполяцію *кусково-лінійною* та *кусково-квадратичною функціями.*

Нехай $f(x)$ задана на системі вузлів $a \le x_0 < x_1 < \ldots < x_n \le b$ й необхідно для розв'язання задачі інтерполяції побудувати кусково-лінійну функцію $\varphi(x)$, виходячи з умови критерію $\varphi(x_i) = y_i = f(x_i)$, $i = \overline{0,n}$.

Кусково-лінійною функцією (лінійним сплайном) називається функція:

$$\varphi(x) = \begin{cases} a_1 x + b_1, & x \in [x_0, x_1] \\ a_2 x + b_2, & x \in [x_1, x_2] \\ \ldots \\ a_n x + b_n, & x \in [x_{n-1}, x_n] \end{cases}.$$

Для знаходження n пар її коефіцієнтів маємо систему з $2n$ лінійних рівнянь:

$$\left[\begin{array}{l} \begin{cases} a_1 x_0 + b_1 = y_0 \\ a_1 x_1 + b_1 = y_1 \end{cases} \\ \begin{cases} a_2 x_1 + b_2 = y_1 \\ a_2 x_2 + b_2 = y_2 \end{cases} \\ \qquad \ldots \\ \begin{cases} a_n x_{n-1} + b_n = y_{n-1} \\ a_n x_n + b_n = y_n \end{cases} \end{array} \right.$$

Графік кусково-лінійної функції має вигляд (рис. 7.2):

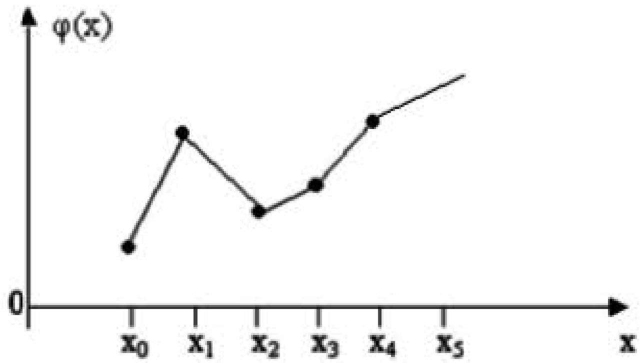

Рисунок 7.2 – Графік кусково-лінійної функції

Кусково-лінійна функція $\varphi(x)$ всередині інтервалу $(x_{i-1};x_i)$, $i = \overline{1,n}$ неперервна й диференційовна, а в точках x_i, $i = \overline{0,n-1}$ неперервна, але не диференційовна.

Розглянемо кусково-квадратичну функцію.

Нехай *f(x)* задана таблично на відрізку $[a;b]$, але $n=2m$ (парне) $a \le x_0 < x_1 < \ldots < x_n \le b$.

Кусково-квадратичною функцією (квадратичним сплайном) називається функція:

$$\varphi(x)=\begin{cases} a_1x^2+b_1x+c_1, & x\in[x_0,x_2] \\ a_2x^2+b_2x+c_2, & x\in[x_2,x_4] \\ \ldots \\ a_mx^2+b_mx+c_m, & x\in[x_{2m-2},x_{2m}] \end{cases}.$$

Для пошуку невідомих коефіцієнтів a_k, b_k, c_k, $k=\overline{1,m}$ формується система рівнянь на основі критерію інтерполяції $\varphi(x_i)=y_i=f(x_i)$, $i=\overline{0,n}$. Ця система складається з m систем:

$$\begin{cases} a_kx_{2k-2}^2+b_kx_{2k-2}+c_k=y_{2k-2} \\ a_kx_{2k-1}^2+b_kx_{2k-1}+c_k=y_{2k-1} \\ a_kx_{2k}^2+b_kx_{2k}+c_k=y_{2k} \end{cases}.$$

Наприклад, при $k=1$:

$$\begin{cases} a_1x_0^2+b_1x_0+c_1=y_0 \\ a_1x_1^2+b_1x_1+c_1=y_1 \\ a_1x_2^2+b_1x_2+c_1=y_2 \end{cases}.$$

Графічно ця функція має вигляд (рис. 7.3):

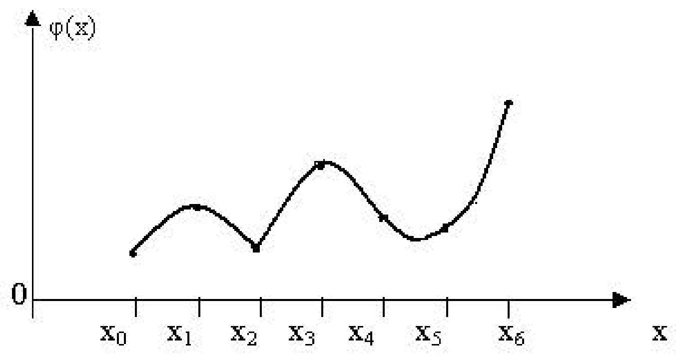

Рисунок 7.3 – Графік кусково-квадратичної функції

Кусково-квадратична функція всередині кожного інтервалу (x_{i-2},x_i) є неперервною та диференційовною двічі, а в точках x_i неперервна, але недиференційовна.

Узагальненням розглянутих функцій є сплайни.

Нехай на відрізку $[a;b]$ задана система вузлів $a_0 \le x_0 < x_1 < \ldots < x_n \le b$.

Сплайном $S_m(x)$ називається функція, яка визначена на відрізку $[a;b]$, належить класу l раз неперервно-диференційовних функцій й така, що на кожному відрізку $[x_{k-1}; x_k]$, $k=\overline{1,n}$ вона є поліномом порядку m.

Різницю $(m-l)$ називають *дефектом сплайна*, що показує різницю між порядком поліномів, які його складають, і ступенем гладкості загальної функції.

Якщо сплайн побудований за деякою функцією $f(x)$ таким чином, що $S_m(x_i)= f(x_i)$, то сплайн називають *інтерполяційним*. Вузли сплайна й вузли інтерполяції функції можуть не збігатися.

Очевидно, що кусково-лінійна функція є інтерполяційним сплайном порядку 1 й дефекту 1, а кусково-квадратична функція є інтерполяційним сплайном порядку 2, дефекту 2.

7.1.3.1 Класичний кубічний сплайн

Розглянемо найбільш відомий й поширений інтерполяційний сплайн порядку 3 дефекту 1. В машинобудівному кресленні ці сплайни застосовуються багато років, тому що це і є лекала чи гнучкі лінійки, які деформуються так, щоб з їх допомогою можна було провести криву через задані точки (x_i, y_i). Можна показати (вживаючи теорію згину брусу при малих деформаціях), що сплайн – це група поєднаних кубічних багаточленів, в місцях поєднання яких перша та друга похідні неперервні. Такі функції називаються **кубічними сплайнами**. Для їх побудови необхідно задати коефіцієнти, які однозначно визначають поліном у проміжку між двома точками.

Наведемо *математичний опис кубічних сплайнів*.

Нехай на відрізку $[a,b]$ дійсної осі Ox задана сітка $a = x_0 < x_1 < \ldots < x_n = b$, в вузлах якої визначені значення $y_i = f(x_i)$, $i=\overline{0,n}$ функції $f(x)$. Потрібно побудувати на відрізку $[a,b]$ неперервну функцію – сплайн $S(x)$, яка задовольняє такі вимоги:

1) На кожному відрізку $[x_{i-1}, x_i]$ сплайн $S(x)$ є багаточленом $S_i(x)$ третього степеня:

$$S_i(x) = a_i x^3 + b_i x^2 + c_i x + d_i. \tag{7.5}$$

2) У вузлах x_i сплайн $S_i(x)$ приймає задані значення y_i, $i=\overline{0,n}$, тобто

$$S_i(x_{i-1}) = y_{i-1}, i = \overline{1,n}$$

$$S_i(x_i) = y_i, i = \overline{1,n}. \quad (7.6)$$

Умови (7.6) потрібні для проходження сплайнів через вузли заданої сітки $a = x_0 < x_1 < \ldots < x_n = b$, $y_i = f(x_i)$, $i = \overline{0,n}$. Попередні дві умови утворюють $2n$ рівнянь.

3) У внутрішніх вузлах x_i, $i = \overline{1,n-1}$ сплайн має неперервну першу і другу похідні, тобто:

$$S_i'(x_i) = S_{i+1}'(x_i), i = \overline{1,n-1};$$

$$S_i''(x_i) = S_{i+1}''(x_i), i = \overline{1,n-1}.$$

В точках спряження сплайнів, їх перші та другі похідні повинні бути рівними. Таких умов $2n-2$. Для знаходження сплайна потрібно знайти коефіцієнти a_i, b_i, c_i, d_i багаточленів $S_i(x)$, $i = \overline{1,n-1}$, тобто $4n$ невідомих, які задовольняють $4n-2$ рівнянь.

Для отримання розв'язку системи потрібно два додаткових рівняння. Їх отримують, визначивши значення кривизни графіка сплайна на кінцях:

$$S''(x_0) = k_1,$$

$$S''(x_n) = k_2$$

Якщо $k_1 = k_2 = 0$, тоді такий сплайн називають природним. Коли є додаткові відомості про поведінку функції на кінцях інтервалу інтерполяції, то записуються інші крайові умови. Таким чином, кубічний сплайн, "склеєний" з кубічних парабол, проходить через задані точки, є гладеньким і має неперервну кривизну.

Для побудови кривої (7.5) знаходять чотири коефіцієнти. Запишемо вираз (7.5) у формі, що запропонована Чарльзом Ермітом і дозволяє зменшити кількість обчислювальних операцій.

Введемо позначення:

$$h_i = x_{i+1} - x_i; \quad \omega = \frac{x - x_i}{h_i}; \quad \overline{\omega} = 1 - \omega \ , \quad (7.7)$$

де h_{i+1} – довжина підінтервалу; ω і $\overline{\omega}$ – допоміжні змінні; x – проміжна точка на відрізку $[x_i, x_{i+1}]$.

Штучною змінною $\omega \in [0,1]$ виконується нормалізація змінної x на кожному відрізку інтерполювання між двома вузловими точками сітки.

Якщо $x = x_i$, то $\omega = \dfrac{x_i - x_i}{h_{i+1}} = 0$; якщо $x = x_{i+1}$, то $\omega = \dfrac{x_{i+1} - x_i}{h_{i+1}} = 1$.

Тобто, коли x приймає всі значення в інтервалі $[x_i, x_{i+1}]$, змінна ω змінюється від 0 до 1, а змінна $\overline{\omega}=1-\omega$ змінюється від 1 до 0.

Поліноміальний сплайн третього степеня, що має неперервну першу та другу похідні на відрізку $[x_i, x_{i+1}]$, запишеться:

$$S_{i+1}(x)=\omega y_{i+1}+\overline{\omega}y_i+h_{i+1}^2((\omega^3-\omega)y'_{i+1}+(\overline{\omega}^3-\overline{\omega})y'_i). \qquad (7.8)$$

Номер сплайна збігається з індексом кінцевої точки відрізка $[x_i, x_{i+1}]$. Для запису виразу сплайна на сусідньому i-му відрізку достатньо в (7.8) зменшити всі індекси на одиницю:

$$S_i(x)=\omega y_i+\overline{\omega}y_{i-1}+h_i^2((\omega^3-\omega)y'_i+(\overline{\omega}^3-\overline{\omega})y'_{i-1}). \qquad (7.9)$$

Змінні ω і $\overline{\omega}$ визначаються відповідно до конкретного відрізка інтерполяції, тому вирази (7.8) та (7.9) містять фактично різні змінні ω і $\overline{\omega}$. Знайдемо значення сплайна $S_{i+1}(x)$ на кінцях відрізка $[x_i, x_{i+1}]$.

Маємо $x=x_i$ є початком для відрізка $[x_i, x_{i+1}]$, тому $\omega=0$, $\overline{\omega}=1$ і, відповідно до (7.6), $S_i(x_i)=y_i$.

На кінці відрізка $[x_i, x_{i+1}]$ $\omega=1$, $\overline{\omega}=0$, отримаємо $S_{i+1}(x_{i+1})=y_{i+1}$.

Для інтервалу $[x_{i-1}, x_i]$ точка x_i є кінцевою, тому $\omega=1$, $\overline{\omega}=0$, з формули (7.9) отримаємо $S_i(x_i)=y_i$.

Отже, виконується умова неперервності функції $S(x)$ у вузлах спряження кубічних багаточленів, і функція $S(x)$ інтерполює задані значення незалежно від вибору похідних y'_i. Для зручності запису в подальших розділах зробимо заміну $y'_i=m_i$.

Для того, щоб визначити коефіцієнти m_i, $i=\overline{0,n}$, продиференціюємо (7.8) двічі як складну функцію від x, враховуючи, що $\omega'=\left(\dfrac{x-x_i}{h_{i+1}}\right)'=\dfrac{1}{h_{i+1}}$ та $\overline{\omega}'=(1-\omega)'=-\dfrac{1}{h_{i+1}}$. Тоді

$$S'_{i+1}(x)=\omega' y_{i+1}+\overline{\omega}' y_i+h_{i+1}^2[(3\omega^2\omega'-\omega')m_{i+1}+(3\overline{\omega}^2\overline{\omega}-\overline{\omega}')m_i]=$$

$$=\frac{y_{i+1}}{h_{i+1}}-\frac{y_i}{h_{i+1}}+h_{i+1}^2[(3\omega^2-1)\frac{m_{i+1}}{h_{i+1}}-(3\overline{\omega}^2-1)\frac{m_i}{h_{i+1}}]=$$

$$= \Delta_{i+1} + h_{i+1}[(3\omega^2 - 1)m_{i+1} - (3\overline{\omega}^2 - 1)m_i] \ ,$$

де $\Delta_{i+1} = \dfrac{y_{i+1} - y_i}{h_{i+1}}$.

Зменшуючи індекс на одиницю, отримаємо:

$$S_i'(x) = \Delta_i + h_i((3\omega^2 - 1)m_i - (3\overline{\omega}^2 - 1)m_{i-1}) \, . \tag{7.10}$$

Для сплайна (7.9) виконується умова рівності другої похідної у внутрішніх точках інтерполяційної сітки. Тому для знаходження невідомих коефіцієнтів m_i записують систему рівнянь $S_i'(x) = S_{i+1}'(x)$, $i = \overline{0, n-1}$, використовуючи (7.10). Однозначний розв’язок системи отримаємо введенням додаткових крайових умов.

7.1.3.2 Створення сплайнових параметричних кривих

Особливість сплайнових параметричних кривих полягає у вираженні невідомих коефіцієнтів сплайна через похідні у визначених точках кривої. Такий прийом зменшує кількість невідомих полінома. Також потрібно відмітити простоту переходу запису параметричних кривих від одного до багатьох вимірів.

Криві і поверхні можуть бути подані явно, неявно і параметрично. Параметричне подання дуже широко застосовується в комп'ютерній графіці через простоту та універсальність. В наступних підрозділах наведено приклади знаходження невідомих коефіцієнтів та безпосередньо самих параметричних кривих.

Нехай необхідно побудувати кубічну параметричну криву, визначену трьома точками та тангенсом нахилу кривої в середній точці (рис. 7.4).

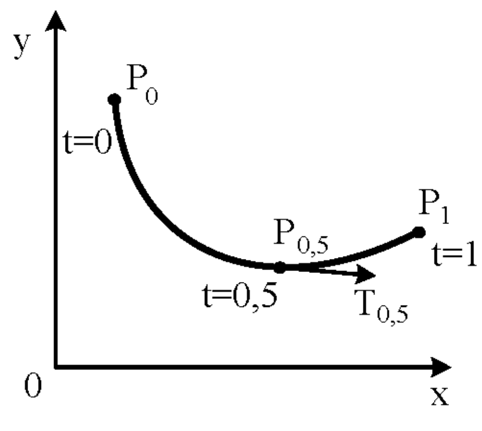

Рисунок 7.4 – Типова крива при побудові сплайна за трьома точками і значенням похідної в середній точці

Спершу побудуємо криву для змінної $x(t)$ (рис. 7.5). На криву накладається чотири обмеження для знаходження всіх чотирьох параметрів.

$$\begin{cases} x(0) = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix} \cdot A; \\ x(1) = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} \cdot A; \\ x(0{,}5) = \begin{bmatrix} 0{,}5^3 & 0{,}5^2 & 0{,}5 & 1 \end{bmatrix} \cdot A; \\ x'(0{,}5) = \begin{bmatrix} 3 \cdot 0{,}5^3 & 2 \cdot 0{,}5^2 & 1 & 0 \end{bmatrix} \cdot A. \end{cases} \qquad (7.11)$$

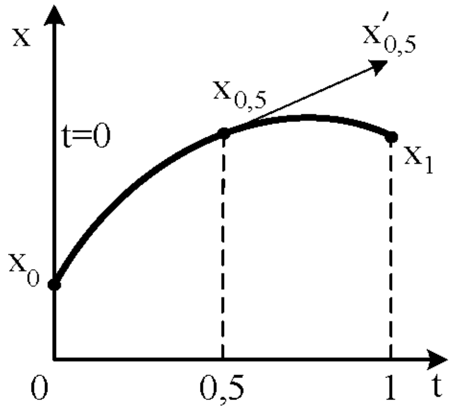

Рисунок 7.5 – Побудова кривої для змінної x(t)

Запишемо рівняння (7.11) в матричному вигляді.

$$G_x = Coef \cdot A;$$

$$G_x = \begin{bmatrix} x(0) \\ x(1) \\ x(0{,}5) \\ x'(0{,}5) \end{bmatrix} \quad Coef = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 0{,}125 & 0{,}25 & 0{,}5 & 1 \\ 0{,}75 & 1 & 1 & 0 \end{bmatrix}. \qquad (7.12)$$

З рівняння (7.12) досить просто знаходимо матрицю невідомих коефіцієнтів A:

$$A = G_x \cdot Coef^{-1}. \qquad (7.13)$$

Підставивши (7.12) у вираз (7.13), отримаємо:

$$x(t) = T \cdot G_x \cdot Coef^{-1}. \qquad (7.14)$$

Матрицю $Coef^{-1}$ зазвичай називають базисною. Для зручності запису виконаємо заміну $Coef^{-1} = M$. В нашому випадку матриця $\boldsymbol{M}$ дорівнює:

$$M = \begin{bmatrix} -4 & 4 & 0 & -4 \\ 8 & -4 & -4 & 6 \\ -5 & 1 & 4 & -2 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Помноживши матрицю T та матрицю M, отримаємо сім'ю базових функцій $f1(x),...,f4(x)$. Вони виконують роль основних складових компонентів для створення геометричних утворень з різними ваговими коефіцієнтами. Тобто,

$$\begin{aligned} & x(t) = \begin{bmatrix} f1(t) & f2(t) & f3(t) & f4(t) \end{bmatrix} \cdot G_x; \\ & f1(t) = -4t^3 + 8t^2 - 5t + 1; \\ & f2(t) = -4t^2 + 4t; \\ & f3(t) = -4t^3 + 6t^2 - 2t; \\ & f4(t) = 4t^3 - 4t^2 + 1. \end{aligned} \tag{7.15}$$

Зобразимо функції (7.15) на рис. 7.6.

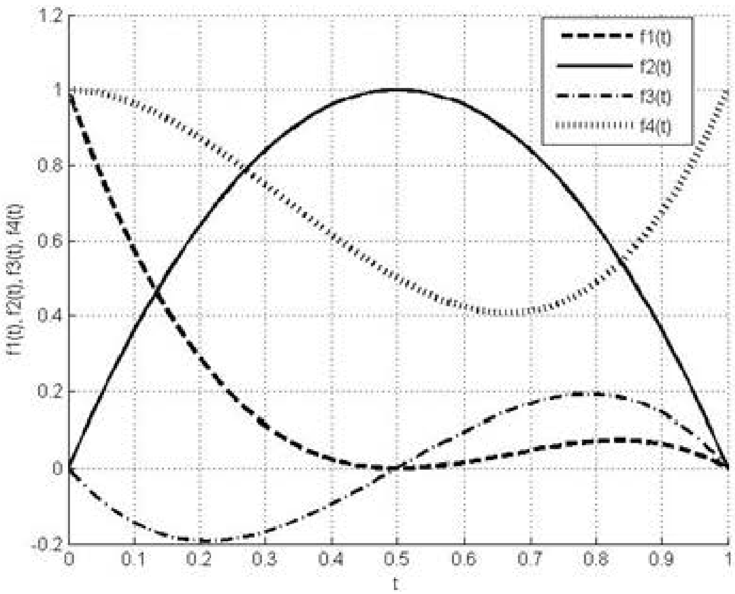

Рисунок 7.6 – Базові функції

Невідомі коефіцієнти для змінних $y(t)$ та $z(t)$ знаходяться аналогічно. Базова матриця та базові функції залишаються незмінними, змінюється лише геометричний вектор G. Тому загальне параметричне рівняння матиме вигляд:

$$z(t) = T \cdot M \cdot G_z;$$
$$P(x) = T \cdot M \cdot G.$$
$$x(t) = T \cdot M \cdot G_x;$$
$$y(t) = T \cdot M \cdot G_y;$$

В цьому прикладі розглядається побудова **ермітової кривої (Hermitian curve)** (рис. 7.7), яка задається координатами двох крайових точок та значеннями їх похідних (дотичних).

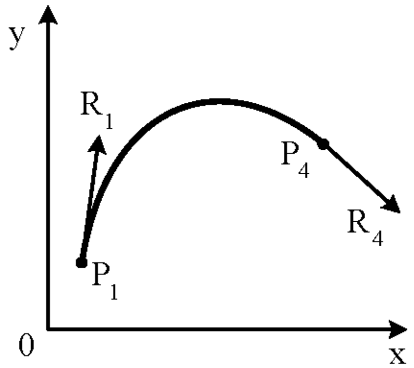

Рисунок 7.7 – Приклад побудови ермітової кривої

Визначимо рівняння для побудови ермітової кривої через геометричний вектор $G_h = [P_1 \quad P_4 \quad R_1 \quad R_4] \cdot T$. Виконаємо дії, аналогічні попередньому прикладу. Запишемо систему рівнянь для знаходження невідомих коефіцієнтів:

$$\begin{cases} x(0) = [0 \quad 0 \quad 0 \quad 1] \cdot A; \\ x(1) = [1 \quad 1 \quad 1 \quad 1] \cdot A; \\ x'(0) = [0 \quad 0 \quad 1 \quad 0] \cdot A; \\ x'(1) = [3 \quad 2 \quad 1 \quad 0] \cdot A. \end{cases} \tag{7.16}$$

В матричному вигляді рівняння (7.16) запишеться $G_x = Coef \cdot A$. Звідки коефіцієнт A розраховується як $A = G_x \cdot M$, $Coef^{-1} = M$. В нашому випадку матриця M дорівнює

$$M = \begin{bmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Подібно до попереднього прикладу розрахуємо параметричну криву, помноживши матрицю T на матрицю M. Отримаємо сім'ю базових функцій

$f1(x),...,f4(x)$:

$$\begin{aligned} &x(t) = \begin{bmatrix} f1(t) & f2(t) & f3(t) & f4(t) \end{bmatrix} \cdot G_x; \\ &f1(t) = 2t^3 - 3t^2 + 1; \\ &f2(t) = -2t^3 + 3t^2; \\ &f3(t) = t^3 - 2t^2 + t; \\ &f4(t) = t^3 - t^2. \end{aligned} \quad (7.17)$$

Зобразимо функції (7.17) на рис. 7.8.

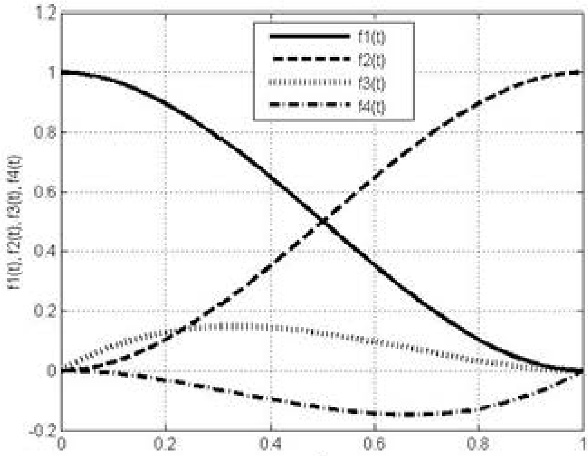

Рисунок 7.8 – Базові ермітові функції

Криві Без'є (Bezier Curves) – головний елемент побудови криволінійних форм у всіх без винятку програмах комп'ютерної графіки; за їх допомогою можна досить точно апроксимувати будь-яку лінію змінної кривизни.

Крива Без'є – це один з найпоширеніших видів гладких параметричних кривих. Ці криві знайшли широке застосування в сучасних CAD-системах. Також вони стали невід'ємною частиною операційної системи Windows (наприклад, вони використовуються для відображення шрифтів), і тому функція побудови кривих Без'є входить в стандартний набір функцій Windows GDI+. Криві Без'є за своєю сутністю є варіацією ермітових кривих. Вони вказані чотирма точками (рис. 7.9).

Визначимо рівняння для побудови ермітової кривої через геометричний вектор

$$G_h = \begin{bmatrix} Ph_1 & Ph_4 & Rh_1 & Rh_4 \end{bmatrix} \cdot T; \quad \begin{cases} Ph_1 = P_1; \\ Ph_4 = P_2; \\ Rh_1 = 3(P_2 - P_1); \\ Rh_4 = 3(P_4 - P_3). \end{cases} \quad (7.18)$$

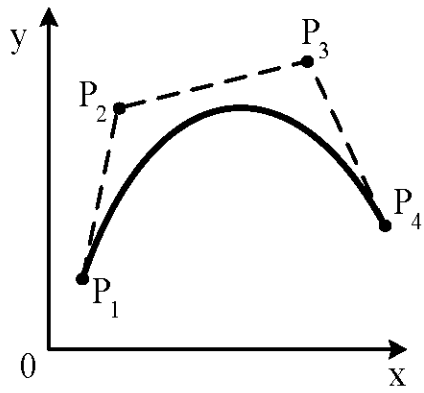

Рисунок 7.9 – Крива Без'є

Виконаємо дії, аналогічні попередньому прикладу. Напишемо систему рівнянь для розрахунку невідомих коефіцієнтів:

$$\begin{bmatrix} Ph_1 \\ Ph_4 \\ Rh_1 \\ Rh_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -3 & 3 & 0 & 0 \\ 0 & 0 & -3 & 3 \end{bmatrix} \cdot \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \end{bmatrix}.$$

Загальне параметричне рівняння запишеться у вигляді:

$$P(t) = T \cdot M \cdot Pb,$$

де $Pb = \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \end{bmatrix}$; $M = \begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 3 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$.

Сім'я базових функцій $f1(x),..., f4(x)$ для кривих Без'є записується:

$$P(t) = f1(t) \cdot P_1 + f2(t) \cdot P_2 + f3(t) \cdot P_3 + f4(t) \cdot P_4;$$

$$f1(t) = -t^3 + 3t^2 - 3t + 1;$$

$$f2(t) = 3t^3 - 6t^2 + 3t;$$

$$f3(t) = -3t^3 + 3t^2; \qquad (7.19)$$

$$f4(t) = t^3.$$

Зобразимо функції (7.19) на рис. 7.10.

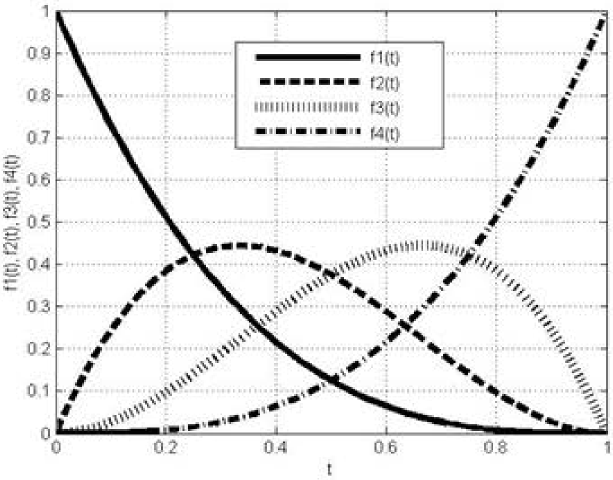

Рисунок 7.10 – Базові функції Без'є

Базові функції Без'є мають властивість опуклості. Крива ніколи не проходитиме через ламану, сформовану чотирма заданими вершинами. Функції Без'є задовольняють умови опуклості, а саме:

– $0 \le fi(t) \le 1;\ \ t \in [0;1]$;

– $f1(t) + \ldots + fn(t) = 1$.