CAPITULO IV

Cálculo de los valores de las funciones elementales

En la práctica de cálculos se necesita, bastante frecuentemente, resolver un problema de cálculo de cierta función en el punto dado. En este caso conviene tener en cuenta que matemáticamente las expresiones equivalentes no siempre resultan ser unívocas desde el punto de vista del cálculo de sus valores. Por ejemplo, para calcular el primer miembro de la identidad

$$a^2+2ab+b^2=(a+b)^2$$

es necesario cumplir cuatro operaciones de multiplicación y dos operaciones de adición, mientras que para calcular el segundo miembro hace falta hacer una sola adición y una sola multiplicación.

Ahora bien, llegamos a un problema importante referente a la forma óptima de representar la función para construir el algoritmo de cálculo de los valores de la misma. Lo óptimo puede entenderse como minimización de la cantidad total de las operaciones aritméticas o del tiempo indispensable para calcular los valores de la función.

§ 4.1. Cálculo de los valores de los polinomios algebraicos

Se llama *polinomio algebraico de n-ésimo grado* la expresión que tiene la forma

$$P(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n, \tag{1}$$

donde los coeficientes $a_0, a_1, \ldots, a_n$ son los números reales, con la particularidad de que a_0 se considera distinto del cero. El coeficiente a_n se denomina *término independiente* del polinomio (1).

El cálculo de los valores del polinomio algebraico es un ejemplo característico concerniente a la minimización de la cantidad de operaciones de un procedimiento de cálculo. Este problema resulta ser importante desde el punto de vista práctico no sólo en su sentido directo sino también porque está estrechamente vinculado con el problema de división del polinomio por un binomio lineal (polinomio de primer grado), es decir, con el problema de determinación de las raíces del polinomio.

Se denomina *raíz (cero) del polinomio* (1) todo número ξ que anula este polinomio al sustituir en el mismo ξ en vez de x, o sea, $P(\xi)=0$.

La raíz ξ del polinomio (1) se llama *raíz de multiplicidad s* (*o raíz de s-ésima multiplicidad*) si el mismo polinomio y todas sus derivadas hasta el orden $s-1$ se anulan en el punto $x=\xi$, mientras que la s-ésima derivada no se anula en este punto:

$$P(\xi)=P'(\xi)=\ldots=P^{(s-1)}(\xi)=0;\quad P^{(s)}(\xi)\neq 0.$$

Si la raíz de un polinomio tiene la multiplicidad $s=1$, esta raíz se considera *simple*.

Ejemplo 1. Determinar la multiplicidad de la raíz $x=-2$ del polinomio

$$P(x)=x^4+7x^3+18x^2+20x+8.$$

△ Hallamos la derivada primera $P'(x)=4x^3+21x^2+36x+20$; calculamos $P'(-2)=0$. Determinamos la segunda derivada $P''(x)=12x^2+42x+36$; tenemos $P''(-2)=0$. Encontramos la tercera derivada $P'''(x)=24x+42$; tenemos $P'''(-2)=-6\neq 0$.

Ahora bien, el mismo polinomio y sus derivadas primera y segunda se anulan en el punto $x=-2$, mientras que la tercera derivada es distinta del cero en este punto; por lo tanto, la raíz $\xi=-2$ tiene la multiplicidad $s=3$. ▲

Las propiedades fundamentales de las raíces del polinomio algebraico se dan en el § 5.8.

Consideremos el problema de cálculo del polinomio (1) en cierto punto $x=x^*$. Para resolver el problema planteado representemos este polinomio en la forma siguiente:

$$P_n(x)=a_n+x(a_{n-1}+x(a_{n-2}+\ldots+x(a_1+xa_0)\ldots)).\quad (2)$$

Para hallar el valor de $P_n(x^*)$ calculemos sucesivamente

$$\begin{aligned} b_0&=a_0,\\ b_1&=a_1+x^*b_0,\\ b_2&=a_2+x^*b_1,\\ &\ldots\ldots\ldots\\ b_n&=a_n+x^*b_{n-1}. \end{aligned}\qquad (3)$$

De las expresiones dadas para b_i se ve que cada coeficiente siguiente b_i se obtiene adicionando el coeficiente respectivo a_i al producto del coeficiente precedente b_{i-1} por x^*.

Para calcular $P_n(x^*)$ y los coeficientes b_i «a mano» es cómodo hacer uso de la tabla siguiente.

En la primera fila se escriben los coeficientes del polinomio $P_n(x)$ (los coeficientes negativos se toman con signo «—», delante de los coeficientes positivos el signo «+» puede omitirse). En la tercera fila se pone inmediatamente $b_0=a_0$. Luego cada coeficiente b_i se multiplica por x^* y se escribe debajo del coeficiente siguiente

a_0	a_1	a_2	a_3	...	a_{n-1}	a_n
	x^*b_0	x^*b_1	x^*b_2	...	x^*b_{n-2}	x^*b_{n-1}
b_0	b_1	b_2	b_3	...	b_{n-1}	b_n

a_{i+1}. Los números contenidos en las filas 1 y 2 se suman y el resultado b_{i+1} se escribe en la tercera fila.

En virtud de construcción del proceso de cálculo es evidente que $b_n = P_n(x^*)$.

No es difícil computar la cantidad de operaciones que ha de cumplirse para calcular $P_n(x^*)$: por la fórmula (1) hace falta efectuar $2n - 1$ multiplicaciones y n adiciones, mientras que por la fórmula 2 deben realizarse n adiciones y n multiplicaciones.

Ahora bien, la fórmula (2) permite economizar casi dos veces al realizar la operación de multiplicación. Además, si tenemos en cuenta que para la operación de multiplicación se gasta un tiempo varias veces mayor que es necesario para realizar la operación de adición, la aplicación del esquema de cálculo (2) resulta bastante eficaz.

El esquema considerado (2) . . . (3) lleva el nombre de **esquema de Horner**.

Ejemplo 1. Haciendo uso del esquema de Horner, calcular el valor del polinomio $P_5(x) = x^5 + 3x^4 - 2x^3 + x^2 - x + 1$ para $x = 3$.

△ Componemos el esquema de Horner para el polinomio dado

1	3	−2	1	−1	1
	3	18	48	147	438
1	6	16	49	146	$439 = P_5(3)$

Ahora bien, $P_5(3) = 439$. ▲

Resulta que al realizar el esquema de Horner no sólo calculamos $b_n = P_n(x^*)$, sino determinamos también los coeficientes b_i $(i = 0, 1, \ldots, n-1)$ del polinomio

$$P_{n-1}^{(1)}(x) = b_0 x^{n-1} + b_1 x^{n-2} + \ldots + b_{n-1} \tag{4}$$

que es el cociente obtenido de la división del polinomio $P_n(x)$ por el binomio $x - x^*$:

$$P_n(x) = (x - x^*) P^{(1)}_{n-1}(x) + P_n(x^*). \tag{5}$$

En efecto, abriendo paréntesis en el segundo miembro de la relación (5), reduciendo los términos semejantes y teniendo en cuenta que $P_n(x^*) = b_n$, obtenemos

$$P_n(x) = b_0x^n + (b_1 - x^*b_0)\, x^{n-1} + (b_2 - x^*b_1)\, x^{n-2} + \ldots$$
$$\ldots + (b_{n-1} - x^*b_{n-2})\, x + b_n - x^*b_{n-1}.$$

De aquí

$$a_0 = b_0,$$
$$a_1 = b_1 - x^*b_0,$$
$$a_2 = b_2 - x^*b_1,$$
$$a_n = b_n - x^*b_{n-1}$$

lo que está en plena concordancia con las igualdades (3) y de este modo demuestra la representación (5).

Ahora bien, llegamos al teorema siguiente.

Teorema del resto (de Bezout). *El resto quedado de la división del polinomio $P_n(x)$ por binomio $x - x^*$ es igual al valor de este polinomio par $x = x^*$.*

Ejemplo 2. Calcular el valor del polinomio $P_4(x) = 12x^4 + 19x^3 - 4$ en el punto $x^* = -2$ y determinar los coeficientes del polinomio $P^{(1)}_3(x)$ que es el cociente obtenido de la división de $P_4(x)$ por el binomio $x + 2$.

△ Hacemos el esquema de Horner para el polinomio dado:

12	19	0	0	−4
	−24	10	−20	40
12	−5	10	−20	36

Ahora bien, $P_4(-2) = 36$ y $P^{(1)}_3(x) = 12x^3 - 5x^2 + 10x - 20$. ▲

De otro ejemplo de utilización del esquema de Horner sirve el problema de cambio de la variable en el polinomio.

Supongamos que en el polinomio (1) se necesita pasar a una nueva variable y ligada con la variable x por la relación lineal $x = y + x^*$ o bien $y = x - x^*$.

Esto quiere decir que es necesario hallar los coeficientes del polinomio

$$P_n(y + x^*) = A_0y^n + A_1y^{n-1} + \ldots + A_{n-1}y + A_n. \tag{6}$$

Se puede mostrar que

$$A_n = P_n(x^*),$$
$$A_{n-1} = P^{(1)}_{n-1}(x^*),$$
$$A_{n-2} = P^{(2)}_{n-2}(x^*),$$
$$\cdots\cdots\cdots$$
$$A_0 = P^{(n)}_0,$$

donde $P_n(x)$ es el polinomio dado (1) y los demás polinomios

$$P^{(k)}_{n-k}(x) = b^{(k)}_0 x^{n-k} + b^{(k)}_1 x^{n-k-1} + \ldots + b^{(k)}_{n-k}$$
$$(k = 1, 2, \ldots, n)$$

se determinan por las relaciones

$$P_n(x) = (x - x^*) P^{(1)}_{n-1}(x) + P_n(x^*),$$
$$P^{(1)}_{n-1}(x) = (x - x^*) P^{(2)}_{n-2}(x) + P^{(1)}_{n-1}(x^*),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$P^{(n-1)}_1(x) = (x - x^*) P^{(n)}_0 + P^{(n-1)}_1(x^*).$$

Ahora bien, se obtiene el siguiente algoritmo simple de determinación de A_j ($j = n, n-1, \ldots, 1, 0$):

1°. Con ayuda del esquema de Horner se calcula el coeficiente $A_n = P_n(x^*)$ y junto con éste los coeficientes $b^{(1)}_i$ ($i = 0, 1, \ldots, n-1$) del polinomio $P^{(1)}_{n-1}(x)$.

2°. Aplicando el esquema de Horner a los polinomios $P^{(1)}_{n-1}(x)$, $P^{(2)}_{n-2}(x), \ldots, P^{(k)}_{n-k}(x)$, se calculan los coeficientes $A_{n-k} = P^{(k)}_{n-k}(x^*)$ y junto con éstos los coeficientes $b^{(k+1)}_i$ ($i = 0, 1, \ldots, n-k-1$) del polinomio $P^{(k+1)}_{n-k-1}(x)$. Este proceso se ejecuta hasta que k llegue a ser igual a $n-1$. Entonces se supone $A_0 = P^{(n)}_0 = b^{(n)}_0 = a_0$ y todos los coeficientes A_j están determinados.

El algoritmo considerado suele llamarse **esquema generalizado de Horner.**

Ejemplo 3. En el polinomio $P_4(x) = 12x^4 + 19x^3 - 4$ pasar a una nueva variable $y = x + 2$.

△ Utilizando las relaciones (5) y (6), hacemos el esquema

12	19	0	0	−4	
	−24	10	−20	40	
12	−5	10	−20	36	$= A_4$
	−24	58	−136		
12	−29	68	−156 $= A_3$		
	−24	106			
12	−53	174 $= A_2$			
	−24				
12	−77 $= A_1$				
12 $= A_0$					

Así, pues el polinomio $Q_4(y) = P_4(y - 2)$ tiene la forma

$$Q_4(y) = 12y^4 - 77y^3 + 174y^2 - 156y + 36. \blacktriangle$$

§ 4.2. Cálculo de los valores de las funciones analíticas

El cálculo de los valores de una función analítica se basa frecuentemente en su representación en forma de la serie de Taylor rápidamente convergente la cual en muchos casos es un instrumento cómodo para calcular los valores de esta función en los puntos pertenecientes al dominio de convergencia de la serie.

Supongamos que se necesita calcular el valor que la función $f(x)$, analítica sobre el segmento $[a, b]$, tiene en el punto $x = x^*$, $x^* \in [a, b]$ con el valor absoluto dado ε máximamente admisible.

La fórmula de Taylor con el término residual en la forma de Lagrange para la función $f(x)$ en el entorno del punto $x = c$, $c \in \in [a, b]$ tiene la forma

$$f(x) = f(c) + (x-c)\frac{f'(c)}{1!} + \ldots$$

$$\ldots + (x-c)^n \frac{f^{(n)}(c)}{n!} + (x-c)^{n+1} \frac{f^{(n+1)}(\bar{c})}{(n+1)!},$$

donde $\bar{c} = c + \theta(x - c)$; $0 < \theta < 1$.

Por lo tanto, el problema de cálculo de la $f(x^*)$ se reduce al cálculo de

$$f(x^*) = S_n(x^*) + R_n(x^*),$$

donde $S_n(x^*)$ es la n-ésima suma parcial de la serie:

$$S_n(x^*) = \sum_{i=0}^{n} (x^* - c)^i \frac{f^{(i)}(c)}{i!}$$

$$(0! = 1, \quad f^{(0)}(c) = f(c))$$

y $R_n(x^*)$, el valor del término residual $R_n(x)$ para $x = x^*$:

$$R_n(x^*) = (x^* - c)^{n+1} \frac{f^{(n+1)}(\bar{c}^*)}{(n+1)!},$$

$$\bar{c}^* = c + \theta^*(x^* - c), \quad 0 < \theta^* < 1.$$

El algoritmo de resolución del problema planteado consiste en lo siguiente:

1°. Sobre el segmento $[a, b]$ se elige un punto $x = c$ que sea, en la medida de lo posible, próximo al punto $x = x^*$ y tal que la misma función $f(x)$ y sus derivadas puedan ser calculadas fácilmente cuando $x = c$.

2°. Se representa ε en la forma de la suma:

$$\varepsilon = \varepsilon_1 + \varepsilon_2 + \varepsilon_3, \tag{1}$$

donde ε_1 es el error residual (error del método); ε_2, el error absoluto, máximamente admisible, de cálculo de $S_n(x^*)$; ε_3, el error absoluto, máximamente admisible, de redondeo del resultado. Hablando en general, ε_1, ε_2 y ε_3 pueden ser números positivos arbitrarios que satisfagan la condición (1).

Sin embargo, en la práctica ε se asigna de ordinario en la forma $\varepsilon = 10^{-m}$, donde m es un número entero. En este caso suelen tomarse $\varepsilon_3 = 0{,}5 \cdot 10^{-m}$ y $\varepsilon_1 = \varepsilon_2 = 0{,}25 \cdot 10^{-m}$. Si el error de redondeo final falta, se toma $\varepsilon_1 = \varepsilon_2 = 0{,}5 \cdot 10^{-m}$; $\varepsilon_3 = 0$.

3°. Se elige una cantidad de sumandos en S_n de un modo tal que se cumpla la desigualdad

$$| f(x^*) - S_n(x^*) | = | R_n(x^*) | \leqslant \varepsilon_1.$$

4°. Se calcula cada sumando de S_n de un modo tal que el valor aproximado $\overline{S}_n$ se distinga del valor exacto S_n no más que en ε_2. Para esto se calcula habitualmente cada sumando de $\overline{S}_n$ con error absoluto $\varepsilon_2/(n+1)$.

5°. Se redondea la suma aproximada $\overline{S}_n$, obtenida en el subp. 4°, (si $\varepsilon_3 \neq 0$) hasta la magnitud $\overline{\overline{S}}_n$.

6°. Se escribe la solución del problema planteado en la forma $f(x^*) = \overline{\overline{S}}_n \pm \varepsilon$.

Ejemplo. Calcular $e^{2,25}$ con exactitud hasta $\varepsilon = 0{,}01$.

△ La fórmula de Taylor con el término residual en la forma de Lagrange para la función e^x en el entorno del punto $x = 0$ tiene la forma

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + e^{\theta x} \frac{x^{n+1}}{(n+1)!}, \quad 0 < \theta < 1.$$

Puesto que al ser grandes x la serie de Taylor de la función e^x converge más lentamente, es racional calcular el valor de $e^{2,25}$ en la forma del producto $e^2 \cdot e^{0,25}$.

La magnitud e^2 puede ser fácilmente calculada con todo grado de exactitud y prácticamente consideramos que el error de su cálculo es igual a cero.

Tomemos ahora que los errores de redondeo y de cálculo de $e^2 \cdot e^{0,25}$ son iguales a 0,005. Entonces el error de cálculo de $e^{0,25}$ constituirá $\overline{\varepsilon} = 0{,}005/e^2 \approx 0{,}0006$.

Ahora bien, hemos reducido el cálculo de $e^{2,25}$ con la exactitud hasta $\varepsilon = 0{,}01$ al cálculo

$$e^{0,25} = 1 + 0{,}25^2 + \frac{0{,}25^2}{2!} + \ldots + \frac{0{,}25^n}{n!} + e^{0,25\theta} \frac{0{,}25^{n+1}}{(n+1)!}$$

con exactitud hasta $\overline{\varepsilon} = 0{,}0006$.

Sea luego $\varepsilon_1 = \varepsilon_2 = 0{,}0003$. Determinemos n, utilizando el término residual de la fórmula de Taylor para e^x si $x = 0{,}25$:

$$e^{0,25\theta} \frac{0{,}25^{n+1}}{(n+1)!} \leqslant 0{,}0003; \quad 0 < \theta < 1.$$

Teniendo en cuenta que $e^{0,25\theta} < e^{0,25} < 1{,}5$, obtenemos $n \geqslant 3$. Por consiguiente, debemos calcular la suma

$$S_3 = 1 + 0{,}25 + \frac{0{,}25^2}{2!} + \frac{0{,}25^3}{3!}$$

con error absoluto 0,0003. Puesto que los dos primeros sumandos ya están calculados con exactitud absoluta, es suficiente calcular los sumandos dos y tres con error absoluto máximo 0,0001, cada uno. Realizando los cálculos necesarios, hallamos

$$\overline{S}_3 = 1 + 0{,}25 + 0{,}0312 + 0{,}0026 = 1{,}2838.$$

Multiplicando el resultado obtenido por e^2, obtenemos $e^2\overline{S}_3 = 9{,}4860\ldots$ Por último, redondeando hasta las centésimas partes, tenemos $e^{2,25} = 9{,}49 \pm 0{,}01$. ▲

De un modo análogo se calculan los valores de las funciones trigonométricas (del seno y del coseno). Las fórmulas respectivas de Taylor con el término residual en la forma de Lagrange se escriben así:

$$\operatorname{sen} x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots$$
$$\ldots + (-1)^{n-1}\frac{x^{2n-1}}{(2n-1)!} + (-1)^n \frac{x^{2n+1}}{(2n+1)!}\cdot\theta,\ 0<\theta<1; \tag{2}$$
$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots$$
$$\ldots + (-1)^{n-1}\frac{x^{2n-2}}{(2n-2)!} + (-1)^n\frac{x^{2n}}{(2n)!}\cdot\theta,\ 0<\theta<1. \tag{3}$$

En las fórmulas (2) y (3) el argumento debe ser expresado en radianes y pertenecer al segmento $[0, \pi/4]$. En este caso la convergencia de las series respectivas será bastante rápida. Hacer que el argumento pertenezca al segmento indicado se puede con ayuda de las fórmulas de reducción y de la relación conocida entre las fórmulas trigonométricas.

Así, por ejemplo, para calcular el valor de sen 2,53 hace falta utilizar la fórmula de reducción $\operatorname{sen} 2{,}53 = \operatorname{sen}(\pi - 2{,}53)$; el argumento $\pi - 2{,}53$ pertenece al segmento $[0, \pi/4]$. En cambio, para calcular el valor de cos 1,27 conviene hacer uso de la fórmula $\cos 1{,}27 = \operatorname{sen}(\pi/2 - 1{,}27)$ y calcular el valor obtenido del seno, puesto que $(\pi/2 - 1{,}27) \in [0, \pi/4]$.

Al calcular el valor del seno, la condición para la elección de la cantidad de sumandos n se escribirá así:

$$\frac{x^{2n+1}}{(2n+1)!} \leqslant \varepsilon_1;\ x \in [0, \pi/4], \tag{4}$$

y al calcular el valor del coseno, dicha condición se escribirá:

$$\frac{x^{2n}}{(2n)!} \leqslant \varepsilon_1;\ x \in [0, \pi/4]. \tag{5}$$

Para calcular los valores de la función logarítmica la fórmula de desarrollo en serie

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \ldots +$$
$$+ (-1)^{n-1}\frac{x^n}{n} + \ldots;\ -1 < x \leqslant 1 \tag{6}$$

es poco útil debido a la convergencia lenta cuando $|x|$ es próximo a la unidad. Además, esta fórmula no permite calcular los logaritmos de números mayores que dos.

Para acelerar la convergencia y ampliar el dominio de aplicación transformemos la fórmula (6) del modo siguiente. Reemplazando en sus miembros segundo y primero x por $-x$, tenemos

$$\ln(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \ldots - \frac{x^n}{n} - \ldots$$

Sustrayendo de la igualdad obtenida la inicial, hallamos

$$\ln\frac{1-x}{1+x} = -2\left(x + \frac{x^3}{3} + \frac{x^5}{5} + \ldots + \frac{x^{2n-1}}{2n-1} + \ldots\right).$$

Suponiendo $(1-x)/(1+x) = z$ y teniendo en cuenta que en este caso $x = (1-z)/(1+z)$, obtenemos la fórmula inicial para calcular el logaritmo natural de todo número z contenido en el intervalo $(0, \infty)$:

$$\ln z = -2\left[\frac{1-z}{1+z} + \frac{1}{3}\left(\frac{1-z}{1+z}\right)^3 + \frac{1}{5}\left(\frac{1-z}{1+z}\right)^5 + \ldots\right]. \quad (7)$$

En los cálculos prácticos el número positivo x cuyo logaritmo natural ha de determinarse se representa cómodamente en la forma siguiente:

$$x = 2^m \cdot z, \quad (8)$$

donde m es cierto número entero y

$$0{,}5 \leqslant z < 1. \quad (9)$$

Entonces

$$\ln x = m \ln 2 + \ln z.$$

El primer sumando se calcula fácilmente si se conoce m y $\ln 2 = 0{,}69314718\ \ldots$, y el segundo sumando se puede hallar con ayuda de la fórmula (7). En este caso en virtud de la desigualdad (9) la magnitud

$$\xi = (1-z)/(1+z) \quad (10)$$

varía dentro de los límites

$$0 < \xi \leqslant 1/3 \quad (11)$$

lo que facilita la convergencia rápida de la serie (7).

Así pues, finalmente tenemos

$$\ln x = m\ln 2 - 2\left(\xi + \frac{\xi^3}{3} + \frac{\xi^5}{5} + \ldots + \frac{\xi^{2n-1}}{2n-1}\right) - R_n. \quad (12)$$

Estimemos el resto:

$$R_n = 2\left(\frac{\xi^{2n+1}}{2n+1} + \frac{\xi^{2n+3}}{2n+3} + \ldots\right) < \frac{2}{2n+1} \cdot \frac{\xi^{2n+1}}{1-\xi^2}.$$

Utilizando la desigualdad (11), esta estimación puede ser transformada en una forma cómoda para compararla con los términos correspondientes de la serie (7):

$$R_n < \frac{9}{4}\,\frac{\xi^{2n+1}}{2n+1} \leqslant \frac{\xi^{2n-1}}{4(2n+1)}. \tag{13}$$

Para calcular el valor del logaritmo decimal es necesario hacer uso de la fórmula

$$\log x = M \ln x,$$

donde $M = \log e = 0{,}434294481903252\ \ldots$

§ 4.3. Método iterativo de cálculo de los valores de las funciones

Consideremos un procedimiento artificial más para calcular el valor de una función continua sobre el segmento $[a, b]$

$$y = f(x) \tag{1}$$

en el punto $x = x^*$; $x^* \in [a, b]$.

Este procedimiento está basado en la utilización del método de Newton para resolver las ecuaciones algebraicas y trascendentes y consiste en lo siguiente:

1°. La función (1) se escribe en la forma implícita y se sustituye en vez de x su valor x^* en la expresión obtenida:

$$F(x^*, y) = 0. \tag{2}$$

De solución de esta ecuación sirve precisamente el valor buscado de la función $y^* = f(x^*)$.

2°. La ecuación (2) se resuelve por el método de Newton para lo cual se elige la aproximación inicial y_0 de un modo tal que se cumpla la condición

$$F(x^*, y_0)\, F''_{yy}(x^*, y_0) > 0 \tag{3}$$

y cada valor aproximado sucesivo y_n $(n = 1, 2, 3, \ldots)$ se calcula con ayuda de la fórmula

$$y_n = y_{n-1} - \frac{F(x^*, y_{n-1})}{F'_y(x^*, y_{n-1})} \quad (n = 1, 2, 3, \ldots). \tag{4}$$

La convergencia del proceso iterativo (4) quedará asegurada si la función $F(x^*, y)$ satisface las condiciones de convergencia del método de Newton.

Nótese que la función (1) puede representarse en la forma implícita (2) por un conjunto infinito de métodos. Entre todos estos métodos ha de elegirse tal que el proceso iterativo (4) sea convergente y la velocidad de convergencia sea bastante grande.

Como ejemplo importante de utilización del procedimiento indicado sirve el cálculo del valor de la función $y = \sqrt[m]{x}$ $(m = 2, 3, \ldots)$ en el intervalo $(0, \infty)$.

Para esta función en calidad de $F(x^*, y)$ es conveniente utilizar la expresión $y^m - x$. Entonces la condición (3) se transforma en la forma

$$y_0 > \sqrt[m]{x^*} \tag{3'}$$

y el proceso iterativo (4), en la forma

$$y_n = y_{n-1} - \frac{y_{n-1}^m - x^*}{m y_{n-1}^{m-1}}. \tag{4'}$$

Nótese que en virtud de las propiedades de la función $y^m - x^*$ el proceso iterativo (4′) converge no sólo en caso de cumplimiento de la condición (3′), sino también en caso de toda aproximación inicial positiva ($y_0 > 0$, con la particularidad de que para todas las aproximaciones sucesivas se cumplirá la condición $y_n > \sqrt[m]{x^*}$ ($n = 1, 2, \ldots$).

El error del valor aproximado y_n puede ser estimado del modo siguiente:

$$\Delta y = y_n - \sqrt[m]{x^*} < \sqrt{\frac{y_{n-1}^m}{x^*}}\,(y_{n-1} - y_n)$$

o bien

$$\Delta y_n = y_n - \sqrt[m]{x^*} < \frac{m-1}{2} \cdot \frac{y_{n-1}^{m-1}}{x^*}\,(y_{n-1} - y_n)^2. \tag{5}$$

Puesto que el caso $m = 2$ es mucho más frecuente que los demás, demos para él una estimación más exacto del error.

Representemos x^* en la forma $x^* = 2^k \bar{x}$, donde k es un número entero y $\bar{x}^* \in [0,5;\ 1)$. Entonces, suponiendo $y_0 = 2^{E(k/2)}$, obtenemos

$$\Delta y_n = y_n - \sqrt{x^*} < \frac{25}{12}\, y_1 \left(\frac{1}{5}\right)^{2^n} \leqslant \frac{25}{8}\, y_0 \left(\frac{1}{5}\right)^{2^n}.$$

Ejemplo. Calcular $\sqrt[3]{34}$ con exactitud hasta $\varepsilon = 10^{-4}$.

$\triangle$En consonancia con la desigualdad (3′) elegimos $y_0 = 3{,}4 > \sqrt[3]{34}$. Luego, utilizando la fórmula (4′), para $x^* = 34$ e $y_0 = 3{,}4$ hallamos sucesivamente y_n y con ayuda de la fórmula (5) calculamos Δy_n:

$$y_1 = 3{,}4 - \frac{3{,}4^3 - 34}{3 \cdot 3{,}4^2} = 3{,}247;\ \ \Delta y_1 = 0{,}01;$$

$$y_2 = 3{,}247 - \frac{3{,}247^3 - 34}{3 \cdot 3{,}247^2} = 3{,}23964;\ \ \Delta y_2 = 2 \cdot 10^{-5}.$$

Así pues, $\sqrt[3]{34} = 3{,}2396 \pm 0{,}0001$. ▲

Ejercicios

1. Haciendo uso del esquema de Horner, dividir el polinomio $P_5(x) = x^5 + 3x^4 - 2x^3 - x + 1$ por el binomio $x - 3$.
2. ¿Es $\xi = 1$ la raíz de la ecuación $x^3 + 2x^2 - 3 = 0$?
3. Haciendo uso del desarrollo en serie de potencias, calcular con exactitud hasta $\varepsilon = 0{,}001$: a) sen 25°; b) cos 20°; c) ln 4; d) $\sqrt[5]{e}$; e) cos 36°; f) sen 18°.
4. Haciendo uso del método iterativo, calcular con exactitud hasta $\varepsilon = 0{,}0005$: a) $\sqrt{12}$; b) $\sqrt{56}$; c) $\sqrt{42}$.

CAPITULO V

Métodos de resolución de las ecuaciones no lineales

§ 5.1. Ecuaciones algebraicas y trascendentes

En los problemas prácticos se necesita con frecuencia resolver las ecuaciones. Toda ecuación con una incógnita puede ser representada en la forma

$$\varphi(x) = g(x), \tag{1}$$

donde $\varphi(x)$ y $g(x)$ son las funciones dadas, definidas sobre cierto conjunto numérico X llamado *dominio de valores admisibles de la ecuación*.

La ecuación con una incógnita puede escribirse en la forma

$$f(x) = 0. \tag{2}$$

En efecto, transponiendo $g(x)$ al primer miembro de la ecuación (1), tenemos la ecuación $\varphi(x) - g(x) = 0$ equivalente a (1). Si designamos el primer miembro de la última ecuación con $f(x)$, obtenemos la ecuación (2).

El conjunto de los valores de la variable x con los cuales la ecuación (1) se transforma en identidad se llama *solución* de esta ecuación y cada valor x de este conjunto se denomina *raíz* de la ecuación.

Por ejemplo, la ecuación $x^2 = 2 - x$ tiene las raíces $x_1 = -2$ y $x_2 = 1$. Sustituyendo -2 y 1 en la ecuación dada en vez de x, obtenemos las identidades: $(-2)^2 = 2 - (-2)$, o sea, $4 \equiv 4$; $1^2 = 2 - 1$, o sea, $1 \equiv 1$.

Resolver una ecuación quiere decir hallar el conjunto de todas las raíces de esta ecuación. Este conjunto puede ser finito o infinito. Así, la ecuación recién considerada tiene dos raíces. La ecuación $\operatorname{sen} x = 0$ tiene la solución $x = \pi n$ $(n = 0, \pm 1, \pm 2, \ldots)$. Asignando a n distintos valores, obtenemos un conjunto infinito de raíces.

El conjunto de varias ecuaciones con varias incógnitas se llama *sistema de ecuaciones* (la incógnita designada con la misma letra en cada una de las ecuaciones debe denotar la misma magnitud incógnita).

Se llama *solución de un sistema de ecuaciones* con varias incógnitas el conjunto de los valores de estas incógnitas el cual convierte en identidad cada ecuación del sistema.

Por ejemplo, el sistema

$$\begin{cases} x^2 + y = 5, \\ x + y^2 = 3 \end{cases}$$

tiene la solución $x = 2$, $y = 1$, ya que con estos valores de las incógnitas las ecuaciones del sistema se convierten en identidades: $4 + 1 \equiv 5$, $2 + 1 \equiv 3$.

Resolver un sistema de ecuaciones significa hallar el conjunto de todas sus soluciones o mostrar que este sistema no tiene soluciones.

De todo el conjunto de las ecuaciones se destaca con frecuencia la clase de ecuaciones algebraicas que poseen varias particularidades características que se utilizan al resolverlas. Se denomina ecuación algebraica con una incógnita a la ecuación que tiene la forma

$$a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n = 0. \tag{3}$$

Aquí n es el número positivo entero llamado grado de la ecuación.

De otro tipo frecuente de ecuaciones sirven las ecuaciones irracionales en las cuales las variables incógnitas se hallan bajo el signo de radical.

Por ejemplo: $3x^2 - 4x + \sqrt[3]{x - 1} = 0$.

Al mismo tiempo la ecuación $\sqrt{1 + \sqrt{5}}\,x^2 + x\sqrt{2} - 3\sqrt{7} = 0$ no es irracional, puesto que x no está bajo el signo de radical.

Si la ecuación (3) está obtenida transformando cualquier otra ecuación, por ejemplo la irracional, es necesario tener en cuenta que las funciones que forman parte de la ecuación inicial pueden ser definidas no en el todo eje numérico como esto tiene lugar en la ecuación algebraica.

Por ejemplo, la ecuación

$$\sqrt{x - 2} + \sqrt{6 - x} = 3$$

después de suprimir la irracionalidad tomará la forma

$$4x^2 - 32x - 73 = 0.$$

Sin embargo, la ecuación inicial está definida no en todo el eje numérico sino para las x que pertenecen al segmento $[2, 6]$.

Los números $a_0, a_1, \ldots, a_n$ se denominan *coeficientes* de la ecuación (3), pueden ser tanto reales como complejos. A continuación se considerarán ecuaciones algebraicas de la forma (3) sólo en caso de tener coeficientes reales.

Las ecuaciones que contienen funciones trascendentes: función exponencial a^x, función logarítmica $\log_a x$, funciones trigonométricas $\operatorname{sen} x$, $\cos x$, $\operatorname{tg} x$, $\operatorname{ctg} x$, funciones trigonométricas inversas y otras, se reúnen bajo el nombre común de ecuaciones trascendentes las cuales junto con las ecuaciones algebraicas cuyo grado supera el primero suelen llamarse ecuaciones no lineales.

La resolución de ecuación con una incógnita consiste en la determinación de las raíces, o sea, los valores de x que convierten la ecuación en identidad. Las raíces de una ecuación pueden ser reales y no reales (complejas).

Se puede hallar los valores exactos de las raíces de una ecuación sólo en los casos excepcionales, por lo general, cuando existe cualquier fórmula simple para calcular los valores de las raíces, tal que permita expresarlos por las magnitudes conocidas.

Así, para determinar las raíces de la ecuación cuadrática que tiene la forma $x^2 + px + q = 0$ se utiliza la fórmula

$$x_{1,2} = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}. \qquad (4)$$

Para resolver la ecuación cúbica que tiene la forma $x^3 + px + q = 0$ se emplea la fórmula

$$x = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}. \qquad (5)$$

No obstante, la aplicación práctica de esta fórmula presenta muchas dificultades y se necesita utilizar números complejos.

Para resolver una ecuación de cuarto grado también existe el algoritmo, pero es tan complicado que prácticamente no se emplea y no vamos a considerarlo.

El matemático noruego Abel demostró que para $n \geqslant 5$ no existe ninguna fórmula que exprese la solución de la ecuación algebraica (3) a través de sus coeficientes con ayuda de las operaciones aritméticas y de la extracción de las raíces. Únicamente para algunos casos particulares de las ecuaciones algebraicas cuyo grado es más de cuatro pueden existir las fórmulas de solución.

Además, los coeficientes de algunas ecuaciones son números aproximados y, por consiguiente, la cuestión acerca de la determinación de las raíces exactas no puede ser planteada de ningún modo.

Por eso adquieren gran importancia los métodos de cálculo aproximado de las raíces de la ecuación $f(x) = 0$.

Al resolver muchos problemas prácticos la solución exacta de una ecuación no siempre es necesaria. El problema de determinación de las raíces se considera resuelto si las raíces están calculadas con el grado prefijado de precisión.

Entonces ¿cómo es necesario entender la afirmación «la raíz está calculada con el grado prefijado de precisión»? Sea ξ la raíz de la ecuación y $\bar{x}$, su valor aproximado con exactitud hasta ε; esto quiere decir que $|\xi - \bar{x}| \leqslant \varepsilon$. Si se ha determinado que la raíz buscada ξ está encerrada entre los números a y b, o sea $a < \xi < b$, con la particularidad de que $b - a \leqslant \varepsilon$, entonces los números a y b son valores aproximados de la raíz ξ, con defecto y con exceso, respectivamente, con exactitud hasta ε, ya que $|\xi - a| < b - a \leqslant \varepsilon$ y $|\xi - b| < b - a \leqslant \varepsilon$. Por valor aproximado de la raíz ξ con exactitud hasta ε se puede tomar todo número comprendido entre a y b.

En la práctica, siempre que esté determinado que la raíz ξ pertenece al segmento $[a, b]$, por su valor aproximado se toma el centro de este segmento y en este caso el error no superará la mitad de la longitud del segmento $\xi = \frac{a+b}{2} \pm \frac{b-a}{2}$.

Así, por ejemplo, si la raíz ξ está encerrada entre 23,07 y 23,10, o sea $23{,}07 \leqslant \xi \leqslant 23{,}10$, por valor aproximado de la raíz se toma $\frac{23{,}07+23{,}10}{2} = 23{,}085$ y el error de este valor aproximado $\frac{23{,}10-23{,}07}{2} = 0{,}015$. Así que $\xi = 23{,}085 \pm 0{,}015$.

En este capítulo consideraremos los métodos de resolución aproximada de las ecuaciones y de los sistemas de ecuaciones. Algunos de estos métodos son igualmente aplicable a la determinación de las raíces tanto de las ecuaciones trascendentes como algebraicas. Otros métodos son aplicables únicamente a las ecuaciones algebraicas.

§ 5.2. Separación de las raíces

El proceso de determinación de los valores aproximados de las raíces de ecuaciones se subdivide en dos etapas: 1) separación de las raíces; 2) determinación más exacta de las raíces hasta el grado prefijado de precisión.

En este párrafo consideramos la primera etapa, o sea, la separación de las raíces.

La raíz ξ de la ecuación $f(x) = 0$ se considera *separada* sobre el segmento $[a, b]$ si en este segmento la ecuación $f(x) = 0$ no tiene otras raíces.

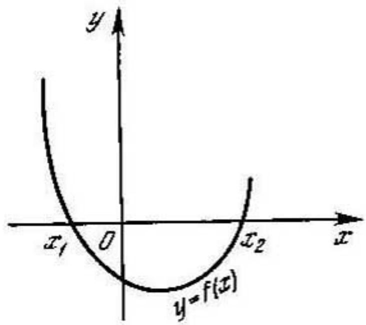

Fig. 5.1

Separar las raíces quiere decir partir todo el dominio de los valores admisibles en segmentos en cada uno de los cuales se contiene una sola raíz. La separación de las raíces puede realizarse por dos métodos: el gráfico y el analítico.

Método gráfico de separación de las raíces. *Procedimiento I.* Las raíces se separan fácilmente si está construido el gráfico de la función $y = f(x)$. Los puntos de intersección del gráfico con el eje Ox dan los valores de la raíz y por el gráfico no es difícil determinar dos números a y b entre los cuales está encerrada una sola raíz (fig. 5.1).

Procedimiento II. Todos los términos de la ecuación se dividen en dos grupos uno de los cuales se escribe en el primer miembro de la ecuación y el otro, en el segundo miembro, o sea, la representan en la forma de $\varphi(x) = g(x)$. Luego se trazan los gráficos de dos funciones $y = \varphi(x)$ e $y = g(x)$. Las abscisas de los puntos de intersección de estas dos funciones sirven precisamente de raíces de la ecuación dada. Supongamos que el punto de intersección de los gráficos tiene la abscisa x_0, las ordenadas de ambos gráficos en este punto son iguales entre sí, o sea, $\varphi(x_0) = g(x_0)$. De esta igualdad se deduce que x_0

es la raíz de la ecuación (fig. 5.2). Por el gráfico se determinan dos números a y b entre los cuales está encerrada la raíz.

Ejemplo 1. Determinar gráficamente entre qué números enteros están encerradas las raíces de la ecuación $x^3 - 3x - 1 = 0$.

△ *Procedimiento I.* Construyamos el gráfico de la función $y = x^3 - 3 \quad - 1$ (fig. 5.3) y determinemos las abscisas de los puntos de intersección de este gráfico con el eje Ox. La curva corta el eje Ox en tres puntos; por lo tanto, la ecuación tiene tres raíces reales (nótese que la ecuación algebraica de tercer grado tiene ora una raíz

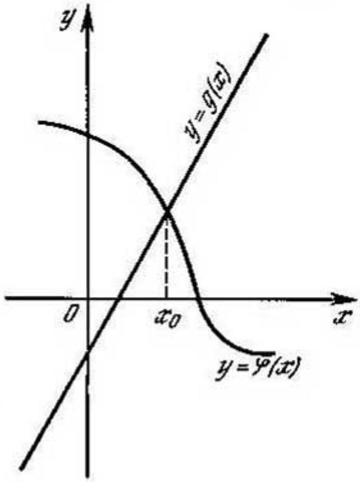

Fig. 5.2

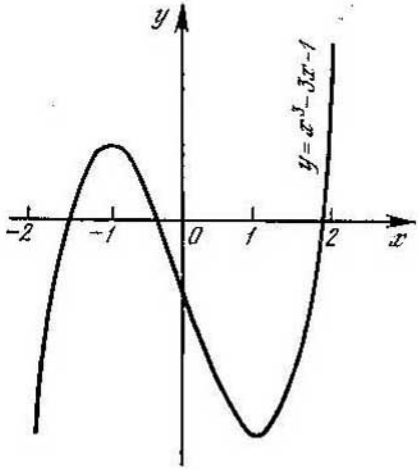

Fig. 5.3

real ora tres raíces reales). Del dibujo se ve que las raíces pertenecen a los segmentos $[-2, -1]$, $[-1, 0]$, $[1, 2]$.

Procedimiento II. Representemos la ecuación dada en la forma $x^3 = 3x + 1$ y tracemos los gráficos de las funciones $y = x^3$ e $y = 3x + 1$ (fig. 5.4). Las abscisas de los puntos de intersección de los gráficos de estas funciones son raíces de la ecuación dada. Del dibujo es fácil determinar los intervalos de aislamiento de las raíces. ▲

Ejemplo 2. Separar gráficamente las raíces de la ecuación $\log x - 3x + 5 = 0$.

△ Escribamos la ecuación del modo siguiente: $\log x = 3x - 5$. Las funciones en los miembros primero y segundo de la ecuación tienen el dominio común de definición: el intervalo $0 < x < +\infty$. Por eso buscaremos las raíces precisamente en este intervalo.

Trazamos los gráficos de las funciones $y = \log x$ e $y = 3x - 5$ (fig. 5.5). La recta $y = 3x - 5$ corta la curva logarítmica en dos puntos. En el dibujo es difícil mostrar la intersección de los gráficos de estas funciones en el primer punto; no obstante, teniendo en cuenta que la rama inferior de la curva logarítmica se aproxima ilimitadamente al eje Oy, se puede suponer que los gráficos se cortarán cerca del punto de intersección del gráfico de la función $y = 3x - 5$ y el eje Oy. Así pues, las raíces se sitúan sobre los segmentos [0; 0,5] y [1, 2]. ▲

Ejemplo 3. Separar gráficamente las raíces de la ecuación $x - \cos x = 0$.

△ Escribamos la ecuación en la forma $x = \cos x$ y construyamos los gráficos de las funciones $y = x$ e $y = \cos x$ en el intervalo

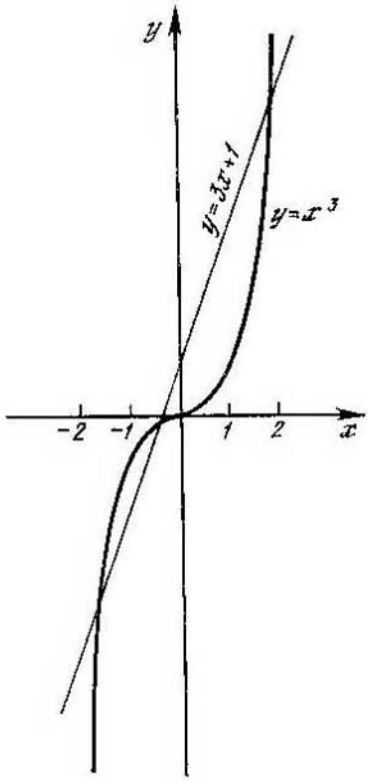

Fig. 5.4

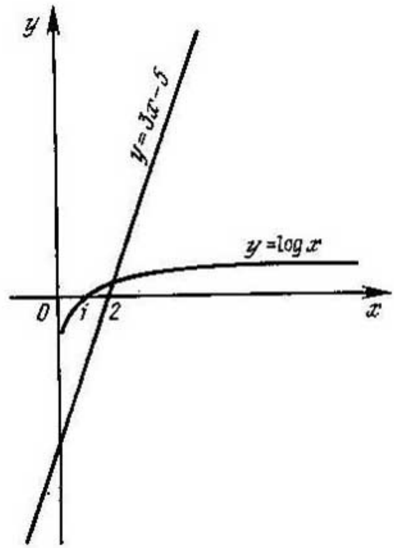

Fig. 5.5

$-\pi/2 \leqslant x \leqslant \pi/2$. Los gráficos de las funciones se cortan en un punto; en el intervalo dado la ecuación tiene la única raíz. Teniendo en cuenta las propiedades de las funciones $y = x$ e $y = \cos x$, se puede cerciorarse de que fuera de este intervalo la ecuación dada no tiene raíces. En caso de una construcción más exacta del dibujo determinamos que la raíz se sitúa en el segmento [0,6; 0,8] (fig. 5.6). ▲

Ejemplo 4. Separar gráficamente por dos procedimientos las raíces de la ecuación $2^x - 5x - 3 = 0$.

△ *Procedimiento I.* Construyamos el gráfico de la función $y = 2^x - 5x - 3$ y determinemos las abscisas de los puntos de intersección de este gráfico con el eje Ox. La curva corta el eje Ox en dos puntos; por consiguiente, la ecuación dada tiene dos raíces reales. Del dibujo se ve que las raíces se sitúan en los segmentos $[-1, 0]$ y $[4, 5]$ (fig. 5.7).

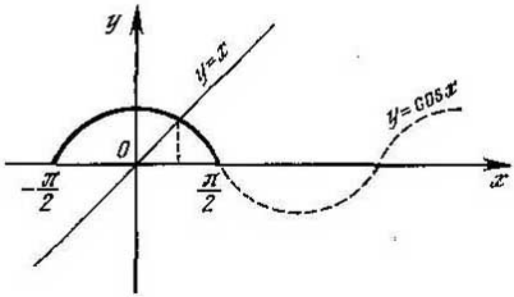

Fig. 5.6

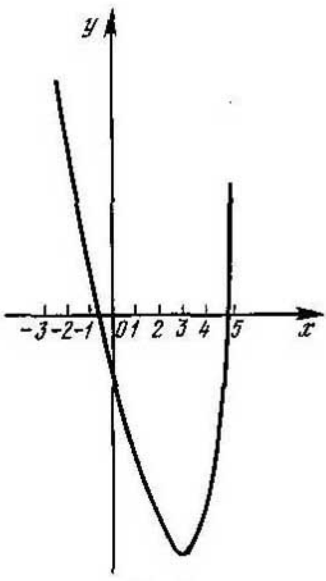

Fig. 5.7

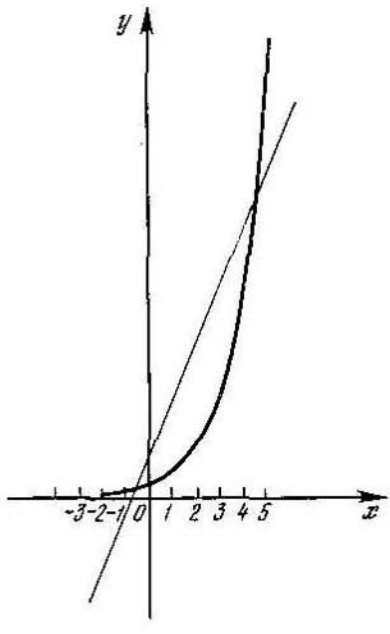

Fig. 5.8

Procedimiento II. Representemos la ecuación dada en la forma $2^x = 5x + 3$ y tracemos los gráficos de las funciones $y = 2^x$ e $y = 5x + 3$. Determinemos las abscisas de los puntos de intersección de los gráficos de estas funciones; estas abscisas son las raíces de la ecuación dada. Obtenemos los mismos intervalos de aislamiento de las raíces $[-1, 0]$ y $[4, 5]$ (fig. 5.8). ▲

Observación. Supongamos que el gráfico de la función $y = f(x)$ tiene la forma representada en la fig. 5.9. La curva corta tres veces el eje de abscisas; por lo tanto, la ecuación tiene tres raíces simples.

En cambio, si la curva toca el eje de abscisas (fig. 5.10), la ecuación tiene una raíz doble. Por ejemplo, la ecuación $x^3 - 3x + 2 = 0$ tiene tres raíces: $x_1 = -2$; $x_2 = x_3 = 1$ (fig. 5.11).

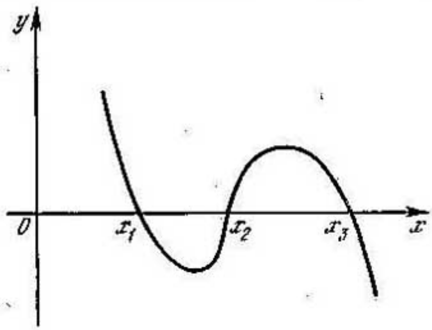

Fig. 5.9

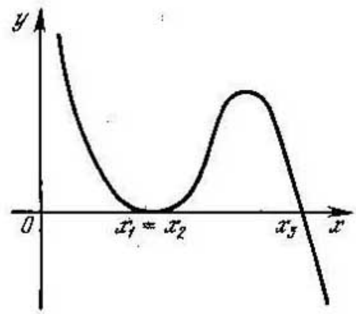

Fig. 5.10

No obstante, si la ecuación tiene una raíz real triple, en el lugar de contacto con el eje la curva $y = f(x)$ tiene un punto de inflexión (fig. 5.12). Por ejemplo, la ecuación $x^3 - 3x^2 + 3x - 1 = 0$ tiene la raíz triple igual a la unidad (fig. 5.13).

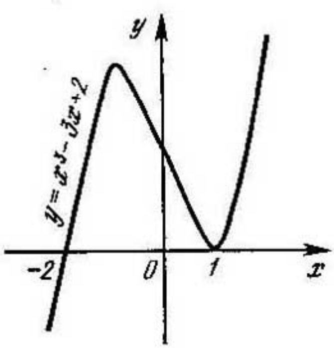

Fig. 5.11

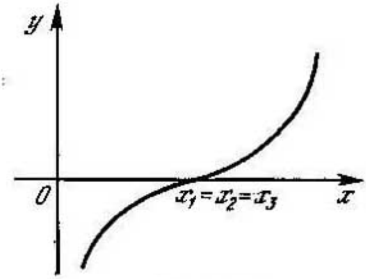

Fig. 5.12

El método gráfico de separación de las raíces no posee gran precisión. Ofrece la posibilidad de determinar aproximadamente los intervalos de aislamiento de las raíces. Luego las raíces se precisan, haciendo uso de los procedimientos indicados posteriormente.

Método analítico de separación de las raíces. Analíticamente las raíces de la ecuación $f(x) = 0$ pueden ser separadas utilizando algunas propiedades de las funciones, propiedades que se estudian en el curso de análisis matemático.

Exponemos ahora algunas nociones contenidas en el análisis matemático las cuales necesitaremos posteriormente.

Si la función $f(x)$ se da analíticamente, entonces se llama *dominio de existencia* (*dominio de definición*) *de la función* el conjunto de todos aquellos valores reales del argumento con los cuales la expresión analítica que define la función no pierde sentido numérico y toma sólo los valores reales.

La función $y = f(x)$ se denomina monótona en el intervalo dado si para todos valores $x_2 > x_1$ contenidos en este intervalo satisface la condición $f(x_2) \geqslant f(x_1)$ (función no decreciente) o bien $f(x_2) \leqslant f(x_1)$ (*función no creciente*). En cambio, si para todos valores $x_2 > x_1$ pertenecientes al intervalo dado $f(x_2) > f(x_1)$ (función creciente) o bien $f(x_2) f(x_1)$ (función decreciente), entonces la función $f(x)$ se denomina estrictamente monótona.

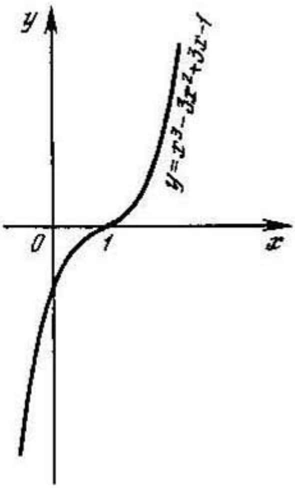

Fig. 5.13

Si la función continua $y = f(x)$ tiene la derivada en todos los puntos interiores del intervalo dado, la condición necesaria y suficiente de monotonía de la función en este intervalo es el cumplimiento de la desigualdad $f'(x) \geqslant 0$ o $f'(x) \leqslant 0$.

Supongamos que sobre el segmento $[a, b]$ la función $f(x)$ es continua y toma en los extremos del segmento los valores de signos contrarios, y la derivada $f'(x)$ mantiene un signo constante sobre el intervalo (a, b). Entonces *si en todos los puntos del intervalo* (a, b) *la primera derivada es positiva, o sea,* $f'(x) > 0$, *la función* $f(x)$ *en este intervalo crece*: $f(x_2) > f(x_1)$ *para* $x_2 > x_1$ (fig. 5.14, *a*, *c*).

En cambio, *si en todos los puntos del intervalo* (a, b) *la primera derivada es negativa*, o sea $f'(x) < 0$, *la función en este intervalo decrece*: $f(x_2) < f(x_1)$ para $x_2 > x_1$ (fig. 5.14, *b*, *d*). De raíz de la función sirve la abscisa del punto de intersección del gráfico de la función $f(x)$ con el eje Ox.

Supongamos que sobre el segmento $[a, b]$ la función $f(x)$ tiene la derivada de segundo orden que conserva un signo constante en todo el segmento. Entonces *si* $f''(x) > 0$, *el gráfico de la función es convexo hacia abajo* (fig. 5.14, *a*, *d*); en cambio, *si* $f''(x) < 0$, *el gráfico de la función es convexo hacia arriba* (fig. 5.14, *b*, *c*).

Los puntos en los cuales la primera derivada de la función es igual a cero, así como también tales en los cuales ella no existe (por ejemplo, se convierte en infinito), pero la función conserva su continuidad, se llaman *críticos* (este criterio es criterio necesario del extremo).

Si la función $f(x)$ es continua sobre el segmento $[a, b]$, entonces en este segmento siempre hay puntos en los cuales ella toma los valores máximo y mínimo. La función alcanza estos valores ora en los puntos críticos ora en los extremos del segmento. Por consiguiente,

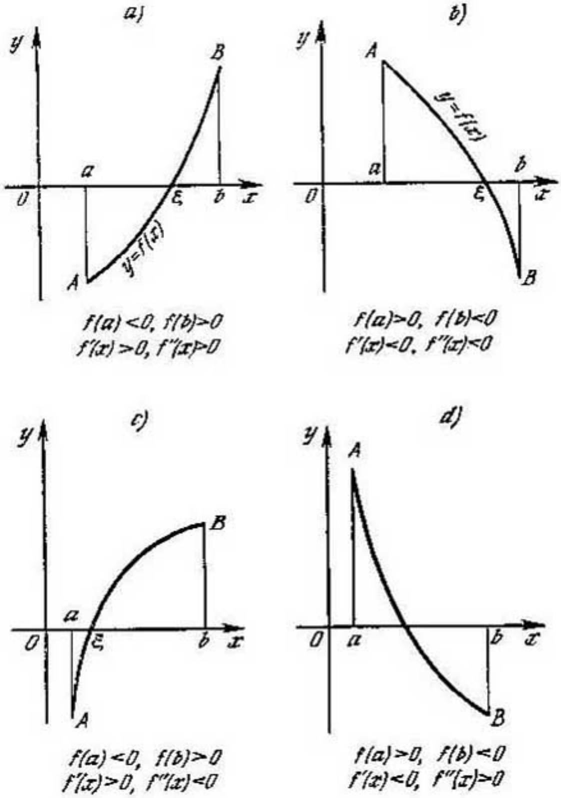

Fig. 5.14

para determinar los valores máximo y mínimo de la función sobre el segmento es necesario: 1) determinar los puntos críticos de la función; 2) calcular los valores de la función en los puntos críticos y en los extremos del segmento $[a, b]$; 3) el mayor de los valores hallados en el subpárr. 2 será valor máximo de la función en el segmento y el menor será valor mínimo de la misma.

Enunciemos, sin demostrarlos, los teoremas cuyo conocimiento es necesario al separar las raíces.

Teorema 1. *Si la función $f(x)$ es continua sobre el segmento $[a, b]$ y toma en los extremos de este segmento los valores de signos opuestos,* *dentro del segmento* $[a, b]$ *existe, al menos, una raíz de la ecuación* $f(x) = 0$.

Teorema 2. *Si la función* $f(x)$ *es continua y estrictamente monótona sobre el segmento* $[a, b]$ *y toma en los extremos del mismo los valores de signos opuestos, entonces dentro del segmento* $[a, b]$ *hay una raíz de la ecuación* $f(x) = 0$ *y esta raíz es la única.*

Teorema 3. *Si la función* $f(x)$ *es continua sobre el segmento* $[a, b]$ *y toma en sus extremos los valores de signos opuestos y la derivada* $f'(x)$ *existe y conserva un signo constante dentro del segmento, o sea* $f'(x) > 0$ *o bien* $f'(x) < 0$, *entonces dentro del segmento existe una raíz de la ecuación* $f(x) = 0$ *y esta raíz es la única.*

De acuerdo con lo expuesto anteriormente se puede recomendar el orden siguiente de operaciones para separar las raíces valiéndose del método analítico:

1) se halla $f'(x)$, o sea, la primera derivada;

2) se hace la tabla de signos de la función $f(x)$ suponiendo x igual: a) a los valores críticos (raíces) de la derivada o a los valores próximos a estos últimos; b) a los valores de frontera (partiendo del dominio de los valores admisibles de la incógnita);

3) se determinan los intervalos en cuyos extremos la función toma los valores de signos contrarios. Dentro de estos intervalos se contiene, cada vez, una y una sola raíz.

Ejemplo 5. Separar las raíces de la ecuación $2^x - 5x - 3 = 0$ con ayuda del método analítico.

△ Designemos $f(x) = 2^x - 5x - 3$. El dominio de definición de la función $f(x)$: todo el eje numérico. Determinemos la primera derivada $f'(x) = 2^x \ln 2 - 5$.

Igualamos a cero esta derivada y calculamos la raíz:

$$2^x \ln 2 - 5 = 0; \quad 2^x \ln 2 = 5; \quad 2^x = 5/\ln 2;$$

$$x \log 2 = \log 5 - \log \ln 2;$$

$$x = \frac{\log 5 - \log \ln 2}{\log 2} = \frac{0{,}6990 + 0{,}1592}{0{,}3010} = \frac{0{,}8582}{0{,}3010} = 2{,}85.$$

Hacemos la tabla de signos de la función $f(x)$ suponiendo x igual: a) a los valores críticos (raíces de la derivada) o a los valores próximos a estos últimos; b) a los valores de frontera (partiendo del dominio de los valores admisibles de la incógnita):

x	$-\infty$	2	3	$+\infty$
Signo de $f(x)$	+	−	−	+

La ecuación tiene dos raíces, ya que ocurren dos cambios de signo de la función. Hagamos una nueva tabla con intervalos más pequeños de aislamiento de la raíz:

x	-1	0	1	2	3	4	5
Signo de $f(x)$	+	−	−	−	−	−	+

Las raíces de la ecuación están encerradas en los intervalos $(-1, 0)$ y $(4, 5)$. ▲

§ 5.3. Determinación máx exacta de las raíces. Método de pruebas

El párrafo precedente ha sido dedicado a la primera de las etapas de resolución aproximada de las ecuaciones algebraicas y trascendentes, o sea, a la separación de las raíces.

La segunda etapa consiste en determinar más exactamente las raíces, es decir, en llevarlas al grado prefijado de precisión. En este párrafo consideramos algunos procedimientos para precisar las raíces los cuales se emplean al resolver las ecuaciones algebraicas y tras-

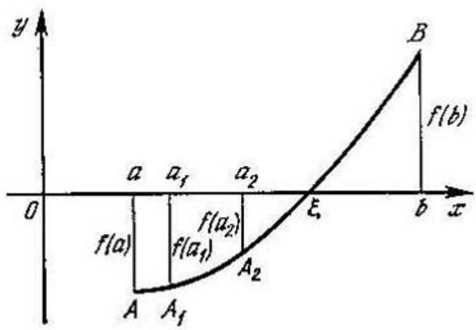

Fig. 5.15

cendentes. No obstante, existen los procedimientos que son aplicables sólo para la resolución de las ecuaciones algebraicas. Vamos a examinarlos a continuación.

Supongamos que se da la ecuación $f(x) = 0$, donde $f(x)$ es la función continua. Se necesita hallar la raíz de esta ecuación ξ con exactitud hasta ε, donde ε es cierto número positivo suficientemente pequeño.

Vamos a suponer que la raíz ξ está separada y se encuentra sobre el segmento $[a, b]$, o sea, tiene lugar la desigualdad $a \leqslant \xi \leqslant b$. En este caso por valor aproximado de la raíz se toma el centro del segmento $\frac{a+b}{2}$ y el error de este valor aproximado es igual a $\frac{b-a}{2}$. Por eso, si $b - a \leqslant 2\varepsilon$, la exactitud requerida está alcanzada y se toma $\xi = \frac{a+b}{2} \pm \varepsilon$. En cambio, si $b - a > 2\varepsilon$, es necesario estrechar el intervalo sobre el cual está la raíz ξ, o sea, elegir tales números $\bar{a}$ y $\bar{b}$ que se cumplan las desigualdades siguientes:

$$\bar{a} \leqslant \xi \leqslant b \quad \text{y} \quad \bar{b} - \bar{a} \leqslant 2\varepsilon.$$

Supongamos que la raíz ξ de la ecuación $f(x) = 0$ [$f(x)$ es una función continua] está separada sobre el segmento $[a, b]$, o sea $f(a) \cdot f(b) < 0$, con la particularidad de que $b - a < 2\varepsilon$. Se necesita hallar el valor de la raíz ξ con exactitud hasta ε (fig. 5.15).

Sobre el segmento $[a, b]$ elegimos de un modo arbitrario el punto a_1 que lo divide en dos segmentos $[a, a_1]$ y $[a_1, b]$. De estos dos segmentos hay que elegir aquél en cuyos extremos la función toma los valores contrarios en signo. En el caso en cuestión $f(a) \cdot f(a_1) > 0$, $f(a_1) \cdot f(b) < 0$, por eso conviene elegir el segmento $[a_1, b]$. Luego, sobre este segmento estrechado volvamos a tomar arbitrariamente el punto a_2 y determinemos los signos de los productos $f(a_1) \cdot f(a_2)$ y $f(a_2) \cdot f(b)$. Puesto que $f(a_2) \cdot f(b) < 0$, escogemos el segmento $[a_2, b]$. Continuamos este proceso hasta que la longitud del segmento, sobre el cual está la raíz, llegue a ser menor que 2ε. Obtenemos la

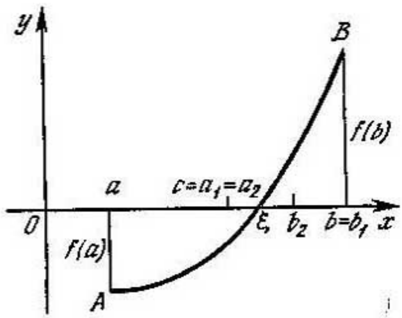

Fig. 5.16

raíz ξ como media aritmética de los extremos del segmento determinado, con la particularidad de que el error de la raíz no supera ε.

El método considerado se llama método de pruebas.

En la forma descrita anteriormente el método de pruebas no se emplea en los ordenadores. Para hacer programas y cálculos en los ordenadores éste se aplica en la forma del llamado **método de bisección**.

Supongamos que la raíz ξ de la ecuación $f(x) = 0$ está separada y se halla sobre el segmento $[a, b]$, o sea, $f(a) \cdot f(b) < 0$, con la particularidad de que $b - a > 2\varepsilon$ [aquí $f(x)$ es una función continua]. Al igual que antes, tomemos sobre el segmento $[a, b]$ un punto intermedio, pero no de un modo arbitrario sino así que éste sea el centro del segmento $[a, b]$, o sea, $c = (a + b)/2$. Entonces el segmento $[a, b]$ se divide por el punto c en dos segmentos iguales $[a, c]$ y $[c, b]$ cuyas longitudes son iguales a $(b - a)/2$ (fig. 5.16). Si $f(c) = 0$, entonces c es la raíz exacta de la ecuación $f(x) = 0$. En cambio, si $f(c) \neq 0$, entonces de dos segmentos obtenidos $[a, c]$ y $[c, b]$ elegimos aquél en cuyos extremos la función $f(x)$ toma los valores de signos opuestos. Designémoslo $[a_1, b_1]$. Luego bisecamos también el segmento $[a_1, b_1]$ y razonamos de un modo análogo. Obtenemos el segmento $[a_2, b_2]$ cuya longitud es igual a $(b - a)/2^2$. El proceso de bisección del segmento se lleva a cabo hasta que en cierta n-ésima etapa el centro del segmento resulte ser la raíz de la ecuación (caso muy raro en la práctica) o bien se obtenga el segmento $[a_n, b_n]$

tal que $b_n - a_n = (b-a)/2^n \leqslant 2\varepsilon$ y $a_n \leqslant \xi \leqslant b_n$ (el número n indica la cantidad de bisecciones efectuadas). Los números a_n y b_n son las raíces de la ecuación $f(x) = 0$ con exactitud hasta 2ε. Según hemos señalado anteriormente por valor aproximado de la raíz es necesario tomar $\xi = (a_n + b_n)/2$, con la particularidad de que el error no supera $(b-a)/2^{n+1} \leqslant \varepsilon$.

Ejemplo 1. Haciendo uso del método de pruebas, precisar hasta $\varepsilon = 10^{-3}$ la menor raíz de la ecuación $x^3 + 3x^2 - 3 = 0$.

△ Separamos las raíces de la ecuación analíticamente. La función $f(x) = x^3 + 3x^2 - 3$ está definida sobre todo el eje numérico. Igualamos a cero $f'(x)$ y calculamos la raíz de la derivada:

$$f'(x) = 3x^2 + 6x; \quad 3x^2 + 6x = 0; \quad x(x+2) = 0; \quad x_1 = 0; \quad x_2 = -2.$$

Hacemos la tabla de los signos de la función:

x	$-\infty$	-2	-1	0	1	$+\infty$
Signo de $f(x)$	$-$	$+$	$-$	$-$	$+$	$+$

Vemos que la primera raíz se contiene en el intervalo $(-\infty, -2)$. Tomemos para la prueba $x = -3$ y determinemos $f(-3) = -3$;

x	-3	-2	-1	0	1
Signo de $f(x)$	$-$	$+$	$-$	$-$	$+$

Por lo tanto, las raíces de la ecuación $x^3 + 3x^2 - 3 = 0$ se contienen en los intervalos $(-3; -2)$, $(-2, -1)$, $(0, 1)$.

Precisemos la menor raíz que está en el intervalo $(-3, -2)$ valiéndonos del método de bisección. Para la comodidad de los cálculos hagamos la tabla (véase la tabla 5.1). Los signos «$-$» y «$+$» presentes en los índices superiores a_n y b_n significan que $f(a_n) < 0$ y $f(b_n) > 0$.

Tabla 5.1

n	a_n^-	b_n^+	$x_n = \frac{a_n + b_n}{2}$	x_n^3	$3x_n^2$	$f(x_n)$
0	-3	-2	$-2{,}500$	$-15{,}625$	$18{,}750$	$0{,}125$
1	-3	$-2{,}500$	$-2{,}750$	$-20{,}800$	$22{,}689$	$-1{,}111$
2	$-2{,}750$	$-2{,}500$	$-2{,}625$	$-17{,}990$	$20{,}670$	$-0{,}320$
3	$-2{,}625$	$-2{,}500$	$-2{,}563$	$-16{,}840$	$19{,}701$	$-0{,}139$
4	$-2{,}563$	$-2{,}500$	$-2{,}532$	$-16{,}230$	$19{,}233$	$0{,}003$
5	$-2{,}563$	$-2{,}532$	$-2{,}548$	$-16{,}540$	$19{,}479$	$-0{,}071$
6	$-2{,}548$	$-2{,}532$	$-2{,}540$	$-16{,}390$	$19{,}356$	$-0{,}034$
7	$-2{,}540$	$-2{,}532$	$-2{,}536$	$-16{,}310$	$19{,}293$	$-0{,}014$
8	$-2{,}536$	$-2{,}532$	$-2{,}534$	$-16{,}270$	$19{,}263$	$-0{,}007$
9	$-2{,}534$	$-2{,}532$				

Así pues, la raíz de la ecuación $x \approx -2,533$. ▲

Ejemplo 2. Separar gráficamente la raíz de la ecuación $x^2 \times \times \log_{0,5}(x+1) = 1$. Haciendo uso del método de pruebas, calcular esta raíz con exactitud $\varepsilon = 10^{-2}$.

△ Representemos la ecuación en la forma $\log_{0,5}(x+1) = 1/x^2$ y construyamos los gráficos de las funciones $y = \log_{0,5}(x+1)$

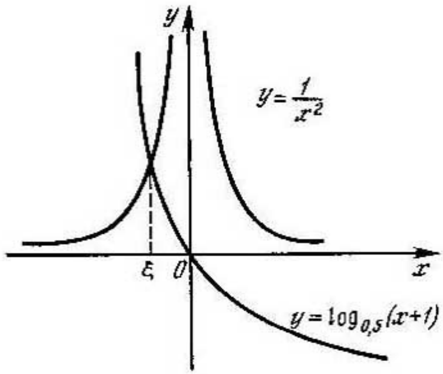

Fig. 5.17

e $y = 1/x^2$. De la fig. 5.17 se ve que la ecuación tiene una sola raíz $x_1 \approx -0,7$. Determinemos los signos de la función a la izquierda y a la derecha de x_1:

x	$-0,8$	$-0,5$
Signo de $f(x)$	$+$	$-$

Para la comodidad de los cálculos pasemos a los logaritmos decimales

$$f(x) = x^2 \frac{\log(x+1)}{\log 0,5} - 1 = x^2 \frac{\log(x+1)}{-0,301} - 1.$$

Hagamos la siguiente tabla (tabla 5.2):

Tabla 5.2

n	a_n^+	b_n^-	$x_n = \frac{a_n + b_n}{2}$	x_n^2	$\log(x_n+1)$	$f(x_n)$
0	$-0,8$	$-0,5$	$-0,65$	0,4225	$-0,4559$	$-0,360$
1	$-0,8$	$-0,65$	$-0,73$	0,5329	$-0,5686$	$-0,0067$
2	$-0,73$	$-0,65$	$-0,69$	0,4761	$-0,5086$	$-0,196$
3	$-0,73$	$-0,69$	$-0,71$	0,5041	$-0,5376$	$-0,099$
4	$-0,73$	$-0,71$				

Así pues, $x_1 \approx -0,72$. ▲

§ 5.4. Método de las cuerdas

El **método de las cuerdas** es uno de los más difundidos de resolución de las ecuaciones algebraicas y trascendentes. En los manuales se utilizan también otros sus nombres: «método de posición falsa» (*regula falsi*), «método de interpolación lineal» y «método de las partes proporcionales».

Supongamos que se da la ecuación $f(x) = 0$, donde $f(x)$ es una función continua que tiene en el intervalo (a, b) las derivadas de primer orden y de segundo orden. La raíz se considera separada y se halla sobre el segmento $[a, b]$, o sea, $f(a) \cdot f(b) < 0$.

La idea del método de las cuerdas consiste en que al ser suficientemente pequeño el intervalo $[a, b]$ el arco de la curva $y = f(x)$ se

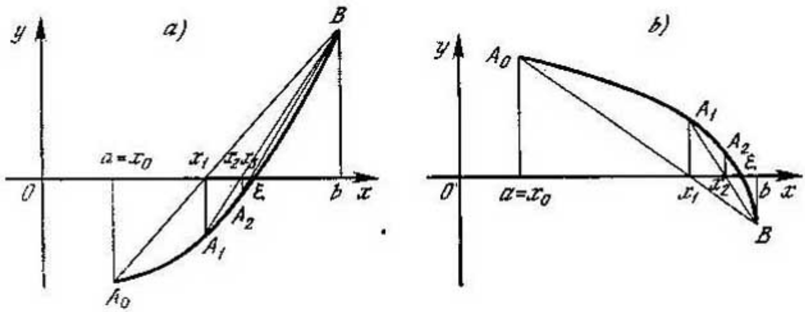

Fig. 5.18

reemplaza por la cuerda que lo subtiende. En calidad de valor aproximado de la raíz se toma el punto de intersección de la cuerda con el eje Ox.

Antes hemos examinado diferentes variantes de disposición del arco de la curva en función de los signos de las derivadas primera y segunda.

Caso I. Las derivadas primera y segunda tienen los mismos signos, o sea $f'(x) \cdot f''(x) > 0$. Supongamos, por ejemplo, que $f(a) < 0$, $f(b) > 0$, $f'(x) > 0$, $f''(x) > 0$ (fig. 5.18, *a*). El gráfico de la función pasa por los puntos $A_0\ (a;\ f(a))$, $B\ (b;\ f(b))$. La raíz buscada de la ecuación $f(x) = 0$ es la abscisa del punto de intersección del gráfico $y = f(x)$ con el eje Ox. Desconocemos este punto, pero en vez de él tomamos el punto x_1 de intersección de la cuerda A_0B con el eje Ox. Éste es precisamente el valor aproximado de la raíz.

La ecuación de la cuerda que atraviesa los puntos A_0 y B tiene la forma

$$\frac{y - f(a)}{f(b) - f(a)} = \frac{x - a}{b - a}.$$

Determinemos el valor $x = x_1$ para el cual $y = 0$:

$$x_1 = a - \frac{f(a)(b - a)}{f(b) - f(a)}. \tag{1}$$

Ahora la raíz ξ se halla dentro del segmento $[x_1, b]$. Si el valor de la raíz x_1 no nos satisface, podemos precisarlo, aplicando el método de las cuerdas al segmento $[x_1, b]$. Unimos el punto A_1 $(x_1; f(x_1))$ con el punto B $(b; f(b))$ y hallamos x_2, o sea, el punto de intersección de la cuerda A_1B con el eje Ox:

$$x_2 = x_1 - \frac{f(x_1)(b - x_1)}{f(b) - f(x_1)}.$$

Continuando este proceso, tenemos

$$x_3 = x_2 - \frac{f(x_2)(b - x_2)}{f(b) - f(x_2)}$$

y en general

$$x_{n+1} = x_n - \frac{f(x_n)(b - x_n)}{f(b) - f(x_n)}. \tag{2}$$

El proceso continúa hasta que obtengamos la raíz aproximada con el grado prefijado de precisión.

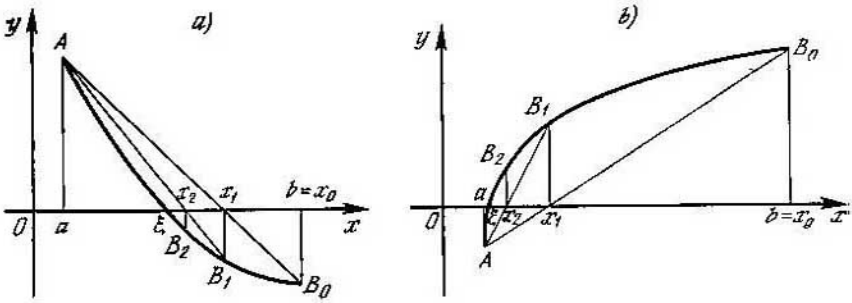

Fig. 5.19

Con ayuda de las fórmulas recién dadas se calculan también las raíces cuando $f(a) > 0$, $f(b) < 0$, $f'(x) < 0$, $f''(x) < 0$ (fig. 5.18,*b*).

Caso II. Las derivadas primera y segunda tienen los signos opuestos, o sea, $f'(x) \cdot f''(x) < 0$. Sea, por ejemplo, $f(a) > 0$, $f(b) < 0$, $f'(x) < 0$, $f''(x) > 0$ (fig. 5.19, *a*). Unimos los puntos A $(a; f(a))$ y B_0 $(b; f(b))$ y escribimos la ecuación de la cuerda que pasa por A y B_0:

$$\frac{y - f(b)}{f(b) - f(a)} = \frac{x - b}{b - a}.$$

Determinemos x_1 como punto de intersección de la cuerda con el eje Ox, suponiendo $y = 0$:

$$x_1 = b - \frac{f(b)(b - a)}{f(b) - f(a)}. \tag{3}$$

La raíz ξ está encerrada ahora dentro del segmento $[a, x_1]$. Aplicando el método de las cuerdas al segmento $[a, x_1]$, obtenemos

$$x_2 = x_1 - \frac{f(x_1)(x_1 - a)}{f(x_1) - f(a)}. \tag{4}$$

y en general

$$x_{n+1}=x_n-\frac{f(x_n)(x_n-a)}{f(x_n)-f(a)}. \qquad (4)$$

Con ayuda de estas mismas fórmulas se determina el valor aproximado de la raíz también cuando $f(a)<0$, $f(b)>0$, $f'(x)>0$, $f''(x)<0$ (fig. 5.19, b).

Por lo tanto, si $f'(x)\cdot f''(x)>0$, la raíz aproximada se calcula por las fórmulas (1) y (2); en cambio, si $f'(x)\cdot f''(x)<0$, se calcula por las fórmulas (3) y (4).

Sin embargo, la elección de unas u otras fórmulas se puede llevar a cabo valiéndose de una **regla** simple: *como extremo fijo del segmento sirve aquél para el cual el signo de la función coincide con el signo de la segunda derivada.*

Si $f(b)\cdot f''(x)>0$, es fijo el extremo b y en calidad de aproximación inicial hace falta tomar el extremo a [fórmulas (1) y (2)]. En cambio, si $f(a)\cdot f''(x)>0$, es fijo el extremo a y en calidad de aproximación inicial es necesario tomar el extremo b [fórmulas (3) y (4)].

Ahora bien, como resultado de aplicación reiterada de las fórmulas (2) ó (4) obtenemos la sucesión monótona $x_1, x_2, x_3, \ldots, x_n$ que converge hacia el valor de la raíz ξ.

Al estimar el error de aproximación se puede utilizar la fórmula

$$|\xi-x_n|<|x_n-x_{n-1}|, \qquad (5)$$

donde ξ es el valor exacto de la raíz y x_{n-1} y x_n son las aproximaciones a éste obtenidas en los pasos $n-1$ y n. Se puede hacer uso de ella si se cumple la condición

$$M\leqslant 2m, \quad \text{donde} \quad M=\max_{[a,\,b]}|f'(x)|, \quad m=\min_{[a,\,b]}|f'(x)|. \qquad (6)$$

Ejemplo 1. Con ayuda del método de las cuerdas precisar hasta $\varepsilon=0{,}001$ la menor raíz de la ecuación $x^3+3x^2-3=0$. Las raíces de la ecuación están separadas y la menor raíz se contiene sobre el segmento $[-3, -2]$ (véase el ejemplo 1, § 5.3).

△ Verificamos el cumplimiento de la condición (6):

$$|f'(x)|=|3x^2+6x|;$$

$$M=\max_{[-3,\,-2]}|f'(x)|=|27-18|=9;$$

$$m=\min_{[-3,\,-2]}|f'(x)|=|12-12|=0; \quad M \nless 2m.$$

Tomamos el centro del segmento $[-3, -2]$, o sea, el punto $x=-2{,}5$ y eligimos el intervalo $[-3; -2{,}5]$. Volvemos a verificar el cumplimiento de la condición (6):

$$M=\max_{[-3,\,-2{,}5]}|f'(x)|=9, \quad m=\min_{[-3\,-2{,}5]}|f'(x)|=3{,}75$$

$$M\nless 2m.$$

Tomamos ahora el centro del segmento $[-3;\ -2{,}5]$, o sea, el punto $x = -2{,}75$; tenemos $f(-2{,}75) < 0$, $f(-2{,}5) > 0$, $f(-3) < 0$; elegimos el segmento $[-2{,}75;\ -2{,}5]$. Encontramos $M = \max\limits_{[-2,75;\ -2,5]} |f'(x)| = 6{,}189$; $m = \min\limits_{[-2,75;\ -2,5]} |f'(x)| = 3{,}75$, o sea, en este caso está cumplida la condición $M < 2m$.

Ahora bien, para estimar el error de la raíz que está sobre el segmento $[-2{,}75;\ -2{,}5]$ se puede hacer uso de la fórmula (5): $|\xi - x_n| < |x_n - x_{n-1}|$, o sea, el proceso de aproximación sucesiva a la raíz ha de continuarse hasta que quede cumplida la condición $|x_n - x_{n-1}| \leqslant \varepsilon$.

Vamos a determinar el signo de la segunda derivada y con ayuda de qué fórmula hace falta llevar a cabo los cálculos. Hallamos $f''(x) = 6x + 6$; sobre el segmento $[-2{,}75;\ -2{,}5]$ tienen lugar las desigualdades $f(-2{,}75) < 0$, $f(-2{,}75) \cdot f''(x) > 0$. Esto quiere decir que por extremo fijo del segmento hay que tomar $x = -2{,}75$. Entonces es necesario realizar los cálculos valiéndose de las fórmulas (3) y (4):

$$x_1 = b - \frac{f(a)(b-a)}{f(b)-f(a)}; \quad x_{n+1} = x_n - \frac{f(x_n)(x_n - a)}{f(x_n) - f(a)},$$

donde $a = -2{,}75$ y $f(a) = -1{,}111$. Si la última expresión se representa en la forma

$$x_{n+1} - x_n = -\frac{f(x_n)(x_n - a)}{f(x_n) - f(a)},$$

inmediatamente se puede obtener la diferencia entre dos aproximaciones sucesivas y llevar a cabo la verificación del fin de los cálculos, o sea, comprobar el cumplimiento de la desigualdad $|x_{n+1} - x_n| \leqslant \varepsilon$.

Es cómodo realizar todos los cálculos en la tabla siguiente:

Tabla 5.3

n	x_n	x_n^3	x_n^2	$3x_n^2$	$f(x_n) = x_n^3 + 3x_n^2 - 3$	$x_n - a$	$-\frac{f(x)(x-a)}{f(x_n) - f(a)}$
0	$-2{,}5$	$-15{,}625$	$6{,}250$	$18{,}75$	$0{,}125$	$0{,}25$	$-0{,}025$
1	$-2{,}525$	$-16{,}098$	$6{,}3756$	$19{,}1268$	$0{,}0288$	$0{,}225$	$-0{,}006$
2	$-2{,}531$	$-16{,}213$	$6{,}4060$	$19{,}2180$	$0{,}0050$	$0{,}219$	$-0{,}0009$
3	$-2{,}5319$						

De la tabla 5.3. se ve que $|x_3 - x_2| = 0{,}0009$, por eso, redondeando hasta las milésimas partes, obtenemos $\xi = -2{,}532$. ▲.

Ejemplo 2. Haciendo uso del método de las cuerdas, precisar hasta $\varepsilon = 0{,}001$ la raíz de la ecuación $x - \operatorname{sen} x = 0{,}25$, encerrada en el segmento $[0,\ \pi/2]$.

△ Escribamos la ecuación en la forma $x - \operatorname{sen} x - 0{,}25 = 0$ y determinemos $f'(x) = 1 - \cos x$. Para verificar el cumplimiento de la condición (6) hagamos una tabla auxiliar en la cual las primeras dos columnas indican el comienzo del intervalo elegido de aislamiento de la raíz y su fin.

Tabla 5.4

a	b	Signos		M	m	Se cumple o no la condición $M \leqslant 2m$	$\frac{a+b}{2}$	$f\left(\frac{a+b}{2}\right)$
		$f(a)$	$f(b)$					
0,00	1,57	−	+	1,00	0	no	0,785	−
0,785	1,57	−	+	1,00	0,29251	no	1,178	+
0,785	1,178	−	+	0,6172	0,2925	no	0,982	−
0,982	1,178	−	+	0,6172	0,4446	sí		

De la última fila de la tabla 5.4 se ve que sobre el segmento [0,982; 1,178] la condición $M \leqslant 2m$ se cumple. Por consiguiente, al estimar el error del valor aproximado de la raíz con ayuda del método de las cuerdas se puede hacer uso de la desigualdad $|x_n - x_{n-1}| < \varepsilon$. La raíz de la ecuación $x - \operatorname{sen} x - 0,25 = 0$ está sobre el segmento [0,982; 1,178]. Vamos a determinar el signo de la segunda derivada dentro del segmento:

$$f'(x) = 1 - \cos x; \quad f''(x) = \operatorname{sen} x > 0.$$

Si retornamos a las anteriores designaciones, $a = 0,982$; $b = 1,178$. El signo de la segunda derivada coincide con el de la función en el punto b. Por lo tanto, este extremo del segmento es fijo y todas las aproximaciones a la raíz están por el lado del extremo a. Para calcular la raíz utilizamos las fórmulas (1) y (2):

$$x_1 = a - \frac{f(a)(b-a)}{f(b)-f(a)}; \quad x_{n+1} = x_n - \frac{f(x_n)(b-x_n)}{f(b)-f(x_n)},$$

donde $b = 1,178$; $f(b) = 0,00416$. Hagamos la tabla siguiente:

Tabla 5.5

n	x_n	$-\operatorname{sen} x_n$	$f(x_n) = x_n - \operatorname{sen} x_n - 0,25$	$b - x_n$	$-\frac{f(x_n)(b-x_n)}{f(b)-f(x_n)}$
0	0,982	−0,83161	−0,09961	0,196	0,189
1	1,171	−9,92114	−0,00014	0,007	0,0002
2	1,1712				

Así pues, $x = 1,171$ con exactitud hasta $\varepsilon = 0,001$. ▲

§ 5.5. Método de Newton (método de las tangentes)

De otro método iterativo sirve el **de Newton.**

Supongamos que la raíz de la ecuación $f(x) = 0$ está separada sobre el segmento $[a, b]$ con la particularidad de que $f'(x)$ y $f''(x)$

son continuas y conservan signos constantes en todo el segmento $[a, b]$.

El sentido geométrico del método de Newton consiste en que el arco de la curva $y = f(x)$ se reemplaza por la tangente a esta curva (de aquí proviene precisamente el segundo nombre de este método, es decir, el **método de las tangentes**).

Caso I. Las derivadas primera y segunda tienen los mismos signos. Sea $f(a) < 0$, $f(b) > 0$, $f'(x) > 0$, $f''(x) > 0$ (fig. 5.20, a) o bien $f(a) > 0$, $f(b) < 0$, $f'(x) < 0$, $f''(x) < 0$ (fig. 5.20, b).

Tracemos la tangente a la curva $y = f(x)$ en el punto B_0 $(b; f(b))$ y determinemos la abscisa del punto de intersección de la tangente

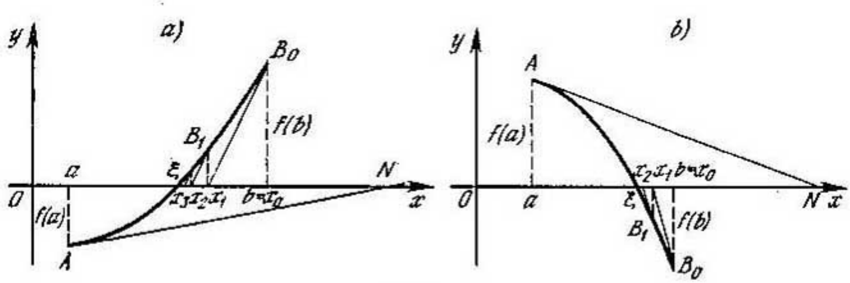

Fig. 5.20

con el eje Ox. Es sabido que la ecuación de la tangente en el punto B_0 $(b; f(b))$ tiene la forma

$$y - f(b) = f'(b)(x - b).$$

Suponiendo $y = 0$, $x = x_1$, obtenemos

$$x_1 = b - \frac{f(b)}{f'(b)}. \tag{1}$$

Ahora la raíz de la ecuación está sobre el segmento $[a, x_1]$. Volviendo a aplicar el método de Newton, tracemos la tangente a la curva en el punto B_1 $(x_1; f(x_1))$ y obtenemos

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$$

y en general

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}. \tag{2}$$

Obtenemos la sucesión de los valores aproximados $x_1, x_2, \ldots$ $\ldots, x_n, \ldots$, en la cual cada término sucesivo es más próximo a la raíz ξ que el precedente. No obstante, todas las x_n quedan mayores que la raíz verdadera ξ, o sea, x_n es el valor aproximado de la raíz ξ con exceso.

Caso II. Las derivadas primera y segunda tienen los signos contrarios. Sea $f(a) < 0$, $f(b) > 0$, $f'(x) > 0$, $f''(x) < 0$ (fig. 5.21, a) o bien $f(a) > 0$, $f(b) < 0$, $f'(x) < 0$, $f''(x) > 0$ (fig. 5.21, b). Si volvemos a trazar la tangente a la curva $y = f(x)$ en el punto B, ésta cortará el eje de abscisas en un punto que no pertenece al segmento $[a, b]$. Por eso tracemos la tangente en el punto A_0 $(a; f(a))$ y escribamos su ecuación para el caso dado:

$$y - f(a) = f'(a)(x - a).$$

Suponiendo $y = 0$, $x = x_1$, encontramos

$$x_1 = a \frac{f(a)}{f'(a)}. \tag{3}$$

Ahora la raíz ξ está sobre el segmento $[x_1, b]$. Volviendo a aplicar

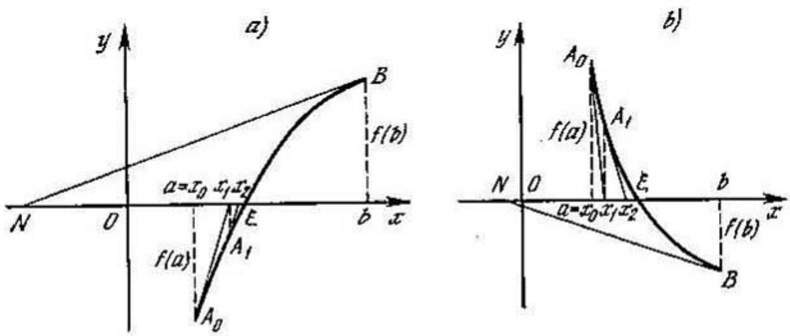

Fig. 5.21

el método de Newton, tracemos la tangente en el punto A_1 $(x_1; f(x_1))$ y obtenemos

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$$

y en general

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}. \tag{4}$$

Se obtiene la sucesión de los valores aproximados $x_1, x_2, \ldots, x_n, \ldots$ en la cual cada término sucesivo es más próximo a la raíz verdadera ξ que el precedente, o sea x_n es el valor aproximado de la raíz ξ con defecto.

Comparando estas fórmulas con las anteriormente deducidas, notamos que se distinguen unas de otras sólo por la elección de la aproximación inicial: en el primer caso por x_0 hemos tomado el extremo b del segmento y en el segundo, el extremo a.

Al elegir la aproximación inicial de la raíz es necesario guiarse por la **regla** siguiente: *por punto inicial ha de tomarse aquel extremo* *del segmento* $[a, b]$ *en el cual el signo de la función coincide con el de la segunda derivada.* En el primer caso $f(b)\cdot f''(x) > 0$ y el punto inicial $b = x_0$, en el segundo caso $f(a)\cdot f''(x) > 0$ y en calidad de la aproximación inicial tomamos $a = x_0$.

Para estimar el error se puede usar la fórmula general

$$|\xi - x_n| \leqslant \frac{|f(x_n)|}{m}, \quad \text{donde } m = \min_{[a,\, b]} |f'(x)| \tag{5}$$

(esta fórmula es útil también para el método de las cuerdas).

En el caso en que el segmento $[a, b]$ es tan pequeño que sobre él se cumple la condición $M_2 < 2m_1$, donde $M_2 = \max\limits_{[a,\, b]} |f''(x)|$

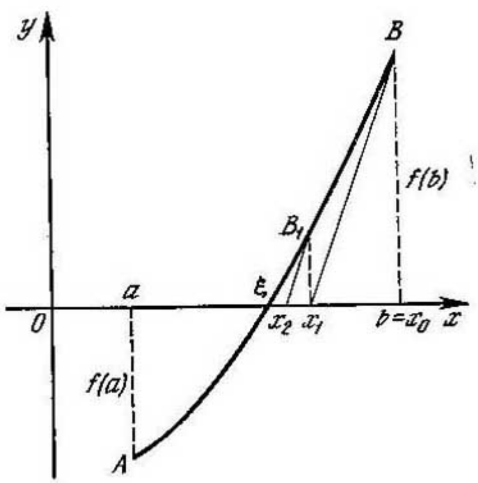

Fig. 5.22

y $m_1 = \min\limits_{[a,\, b]} |f'(x)|$, la exactitud de la aproximación en el n-ésimo paso se estima del modo siguiente: si $|x_n - x_{n-1}| < \varepsilon$, entonces $|\xi - x_n| < \varepsilon^2$.

Si la derivada $f'(x)$ poco varía sobre el segmento $[a, b]$, para simplificar los cálculos se puede utilizar la fórmula

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_0)}, \tag{6}$$

o sea, el valor de la derivada en el punto inicial basta calcularlo una sola vez. Geométricamente esto significa que las tangentes en los puntos $B_n(x_n;\ f(x_n))$ se reemplazan por las rectas paralelas a la tangente trazada a la curva $y = f(x)$ en el punto $B_0(x_0;\ f(x_0))$ (fig. 5.22).

Ejemplo 1. Haciendo uso del método de las tangentes, precisar hasta $\varepsilon = 0{,}001$ la raíz de la ecuación $x^3 + 3x^2 - 3 = 0$, situada sobre el segmento $[-2{,}75;\ -2{,}5]$.

△ Antes hemos determinado que $f(-2,75)\cdot f''(x) > 0$ (véase el ejemplo 1 del § 5.4). Por eso para servirse del método de las tangentes es necesario elegir $x_0 = -2,75$. Vamos a realizar los cálculos con ayuda de la fórmula (6). Hallamos

$$f'(x) = 3x^2 + 6x; \quad f'(x_0) = f'(-2,75) = 6,1875.$$

Para comodidad demos todos los cálculos en la tabla 5.6.

Tabla 5.6

n	x_n	x_n^3	x_n^2	$3x_n^2$	$f(x_n)$	$-\frac{f(x_n)}{6,1875}$
0	−2,75	−20,797	7,5625	22,6875	−1,111	0,179
1	−2,571	−16,994	6,6100	19,8300	−0,164	0,026
2	−2,545	−16,484	6,4770	19,431	−0,053	0,008
3	−2,537	−16,329	6,4364	19,309	0,020	0,003
4	−2,534	−16,271	6,4212	19,2636	0,007	0,001
5	−2,533					

De la tabla se ve que $|x_5 - x_4| = 0,001$; por eso $\xi = 2,533$. ▲

Ejemplo 2. Haciendo uso del método de las tangentes, precisar hasta $\varepsilon = 0,0001$ la raíz de la ecuación $x - \operatorname{sen} x = 0,25$, situada sobre el segmento $[0,982;\ 1,178]$.

△ Aquí $a = 0,982$; $b = 1,178$. Hallamos $f'(x) = 1 - \cos x$; $f''(x) = \operatorname{sen} x > 0$ sobre el segmento $[0,982;\ 1,178]$; $f(1,178) \times \times f''(x) > 0$. Por lo tanto, $x_0 = 1,178$. Vamos a realizar los cálculos con ayuda de las fórmulas (1) y (2) y démoslos en la tabla 5.7.

Tabla 5.7

n	x_n	$-\operatorname{sen} x_n$	$f(x_n) = x_n - \operatorname{sen} x_n - 0.25$	$f'(x_n) = 1 - \cos x_n$	$-\frac{f(x_n)}{f'(x_n)}$
0	1,178	−0,92384	0,00416	0,61723	−0,0065
1	1,1715	−0,92133	0,00017	0,61123	−0,0002
2	1,1713	−0,92127	0,00003	0,61110	−0,00005
3	1,17125				

De la tabla se ve que $|x_3 - x_2| < 0,0001$. Así pues, $\xi \approx 1,1712$. ▲

§ 5.6. Método combinado de las cuerdas y de las tangentes

El método de las cuerdas y el de las tangentes dan las aproximaciones de la raíz por los lados opuestos. Por eso se emplean frecuentemente combinados uno con otro y la raíz se precisa con mayor rapidez.

Supongamos que se da la ecuación $f(x) = 0$, la raíz ξ está separada y se halla sobre el segmento $[a, b]$. Apliquemos el **método combinado de las cuerdas y las tangentes** teniendo en cuenta el tipo del gráfico de la función.

Si $f'(x) \cdot f''(x) > 0$, el método de las cuerdas da las aproximaciones de la raíz con defecto y el de las tangentes, con exceso (fig. 5.23, *a*, *b*).

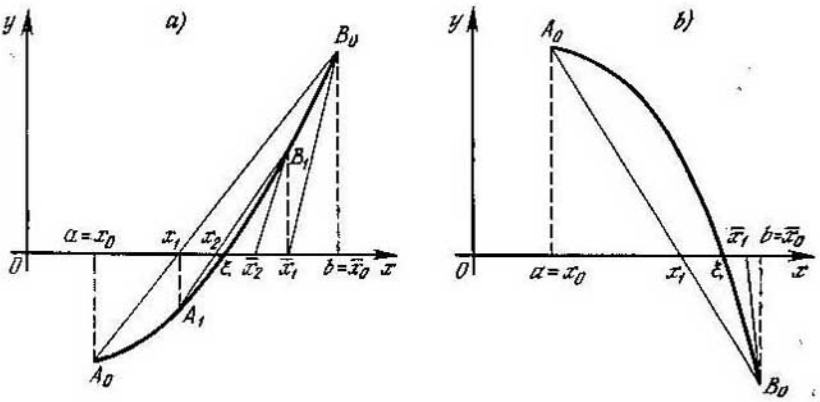

Fig. 5.23

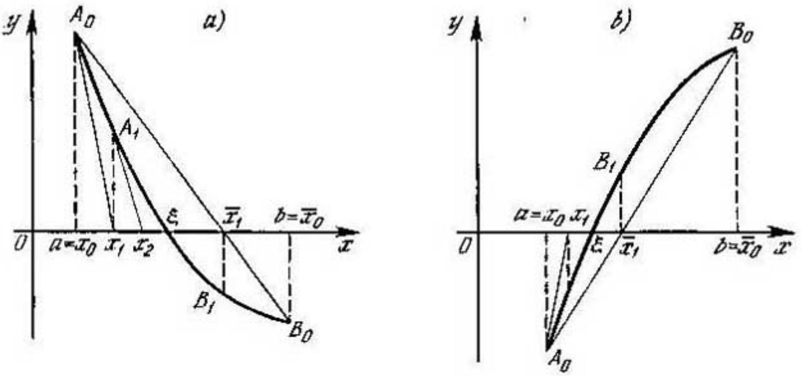

Fig. 5.24

En cambio, *si $f'(x) \cdot f''(x) < 0$, por el método de las cuerdas obtenemos los valores de la raíz con exceso y por el de las tangentes, con defecto* (fig. 5.24, *a*, *b*).

Sin embargo, en todos los casos la raíz verdadera está encerrada entre los valores aproximados de las raíces, obtenidos por el método de las cuerdas y por el de las tangentes, o sea, se cumple la desigualdad $a < \bar{x}_n < \xi < \bar{\bar{x}}_n < b$, donde $\bar{x}_n$ es el valor aproximado de la raíz tomado con defecto y $\bar{\bar{x}}_n$, tomado con exceso.

Los cálculos han de realizarse en el orden siguiente. Si $f'(x) \times \times f''(x) > 0$, en calidad de aproximación inicial para el método debe tomarse el extremo a y para el de las tangentes, el extremo b; entonces

$$a_1 = a - \frac{f(a)(b-a)}{f(b)-f}, \quad b_1 = b - \frac{f(b)}{f'(b)}. \tag{1}$$

Ahora la raíz verdadera está sobre el segmento $[a_1, b_1]$. Aplicando a este segmento el método combinado, obtenemos

$$a_2 = a_1 - \frac{f(a_1)(b_1 - a_1)}{f(b_1) - f(a_1)}, \quad b_2 = b_1 - \frac{f(b_1)}{f'(b_1)}$$

y en general

$$a_{n+1} = a_n - \frac{f(a_n)(b_n - a_n)}{f(b_n) - f(a_n)}, \quad b_{n+1} = b_n - \frac{f(b_n)}{f'(b_n)} \tag{2}$$

(véase la fig. 5.23, a, b).

En cambio, si $f'(x) \cdot f''(x) < 0$, en calidad de aproximación inicial para el método de las cuerdas es necesario tomar el extremo b y para el de las tangentes, el extremo a. Entonces

$$a_1 = a - \frac{f(a)}{f'(a)}, \quad b_1 = b - \frac{f(b)(b-a)}{f(b) - f(a)}. \tag{3}$$

Aplicando al segmento $[a_1, b_1]$ el método combinado, obtenemos

$$a_2 = a_1 - \frac{f(a_1)}{f'(a_1)}, \quad b_2 = b_1 - \frac{f(b_1)(b_1 - a_1)}{f(b_1) - f(a_1)}$$

y en general

$$a_{n+1} = a_n - \frac{f(a_n)}{f'(a_n)}, \quad b_{n+1} = b_n - \frac{f(b_n)(b_n - a_n)}{f(b_n) - f(a_n)} \tag{4}$$

(véase la fig. 5.24, a, b).

El método combinado es muy cómodo para estimar el error de cálculos. El proceso de cálculos cesa tan pronto como se cumpla la desigualdad $|\bar{\bar{x}}_n - \bar{x}_n| < 2\varepsilon$. Por valor aproximado de la raíz ha de tomarse

$$\xi = \frac{1}{2}(\bar{x}_n + \bar{\bar{x}}_n), \tag{5}$$

donde $\bar{x}_n$ y $\bar{\bar{x}}_n$ son los valores aproximados de la raíz con defecto y con exceso, respectivamente.

Ejemplo. Con ayuda del método combinado de las cuerdas y las tangentes precisar hasta 0,001 las raíces de la ecuación $x^3 + 3x^2 - - 24x + 1 = 0$.

$\triangle$ 1) Separamos las raíces analíticamente. Tenemos

$$f(x) = x^3 + 3x^2 - 24x + 1; \quad f'(x) = 3x^2 + 6x - 24,$$

o sea, las raíces de la derivada $x_1 = -4$; $x_2 = 2$. Hagamos la tabla de los signos de la función:

x	$-\infty$	-4	2	$+\infty$
Signo de $f(x)$	$-$	$+$	$-$	$+$

La ecuación dada tiene tres raíces reales: $x_1 \in (-\infty, -4)$, $x_2 \in (-4, 2)$, $x_3 \in (2, +\infty)$. Disminuyamos los intervalos de determinación de las raíces hasta la longitud igual a 1:

x	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4
Signo de $f(x)$	$-$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$-$	$-$	$-$	$+$

Así pues, $x_1 \in (-7, -6)$ $x_2 \in (0, 1)$, $x_3 \in (3, 4)$.

2) Utilizando el método combinado de las cuerdas y las tangentes precisemos la raíz que esta en el intervalo $(-7, -6)$. Tenemos $f(-7) = -27 < 0, f(-6) = 37 > 0, f'(x) = 3x^2 + 6x - 24 > 0$, $f''(x) = 6x + 6 < 0$, $f'(x) \cdot f''(x) < 0$.

Para los cálculos utilizamos las fórmulas (4):

$$a_{n+1} = a_n - \frac{f(a_n)}{f'(a_n)}, \quad b_{n+1} = b_n - \frac{f(b_n)(b_n - a_n)}{f(b_n) - f(a_n)},$$

o sea,

$$a_{n+1} = a_n + \Delta a_n, \text{ donde } \Delta a_n = -\frac{f(a_n)}{f'(a_n)};$$

$$b_{n+1} = b_n + \Delta b_n, \text{ donde } \Delta b_n = -\frac{f(b_n)(b_n - a_n)}{f(b_n) - f(a_n)}$$

(a_n y b_n son los valores aproximados de la raíz tomados con defecto y con exceso, respectivamente). Aquí $a_0 = a = -7$, $b_0 = b = -6$.

Es cómodo realizar los cálculos con ayuda de la tabla 5.8

Tabla 5.8

n	a_n	$b_n - a_n$	a_n^2	a_n^3	$f(a_n)$	$f'(a_n)$	$f(b_n) - f(a_n)$	Δa_n	a_{n+1}
	b_n		b_n^2	b_n^3	$f(b_n)$			Δb_n	b_{n+1}
0	-7	1	49	-343	-27	81	64	0,3333	$-6,6667$
	-6		36	-216	37			$-0,5781$	$-6,5781$
1	$-6,6667$	0,0886	44,4449	$-296,3007$	$-1,9652$	69,3345	6,01002	0,0283	$-6,6384$
	$-6,5781$		43,2714	$-284,6436$	4,0450			$-0,0596$	$-6,6377$
2	$-6,6384$	0,0007							
	$-0,6377$								

Teniendo en cuenta que $|b_2 - a_2| = 0{,}0007 < 0{,}002$, conviene terminar los cálculos y por valor aproximado de la raíz ξ_1 tomar

$$\xi_1 = \frac{1}{2}(-6{,}6384 - 6{,}6377) = -6{,}638.$$

3) Determinemos la raíz aproximada para el intervalo (0, 1). Tenemos $(0) > 0$, $f(1) < 0$, $f'(x) = 3x^2 + 6x - 24 < 0$, $f''(x) = 6x + 6 > 0$, $f'(x) \cdot f''(x) < 0$. Al igual que en el primer caso hagamos uso de las fórmulas 4) para $a_0 = a = 0$; $b_0 = b = 1$.

Hagamos la tabla siguiente

Tabla 5.9

n	a_n / b_n	$b_n - a_n$	a_n^2 / b_n^2	a_n^3 / b_n^3	$f(a_n)$ / $f(b_n)$	$f'(a_n)$	$f(b_n) - f(a_n)$	Δa_n / Δb_n	a_{n+1} / b_{n+1}
0	0	1	0	0	1	−24	−20	0,0417	0,0417
	1		1	1	−19			−0,9500	0,0500
1	0,0417	0,0083	0,0017	0,00007	0,0045	−23,7446	−0,1969	0,0002	0,0419
	0,0500		0,0025	0,0001	−0,1924			−0,0081	0,0419
2	0,0419	0,0000							
	0,0419								

Ahora bien, con la exactitud hasta 0,001 se puede tomar $\xi_2 = 0{,}042$.

4) Determinemos la raíz aproximada contenida en el intervalo (3, 4). Tenemos $f(3) = -17 < 0$, $f(4) = 17 > 0$, $f'(x) = 3x^2 + 6x - 24 > 0$, $f''(x) = 6x + 6 > 0$, $f'(x) \cdot f''(x) > 0$.

Vamos a realizar los cálculos con ayuda de las fórmulas (2):

$$a_{n+1} = a_n - \frac{f(a_n)(b_n - a_n)}{f(b_n) - f(a_n)}, \quad b_{n+1} = b_n - \frac{f(b_n)}{f'(b_n)},$$

o sea,

$$a_{n+1} = a_n + \Delta a_n, \text{ donde } \Delta a_n = -\frac{f(a_n)(b_n - a_n)}{f(b_n) - f(a_n)},$$

$$b_{n+1} = b_n + \Delta b_n, \text{ donde } \Delta b_n = -\frac{f(b_n)}{f'(b_n)}.$$

Aquí $a_0 = a = 3$, $b_0 = b = 4$.

Realicemos los cálculos con ayuda de la tabla (véase la tabla 5.10 en la pág. 190).

Ahora bien, con la exactitud hasta 0,001 se puede tomar $\xi_3 = 3{,}596$. ▲

Tabla 5.10

n	a_n / b_n	$b_n - a_n$	a_n^2 / b_n^2	a_n^3 / b_n^3	$f(a_n)$ / $f(b_n)$	$f'(b_n)$	$f(b_n) - f(a_n)$	Δa_n / Δb_n	a_{n+1} / b_{n+1}
0	3	1	9	27	−17	48	34	0,5000	3,5000
	4		16	64	17			−0,3542	3,6458
1	3,5000	0,1458	12,2500	42,875	−3,3750	37,7505	5,2110	0,0944	3,5944
	3,6458		13,2919	48,4595	1,8360			−0,0486	3,5972
2	3,5944	0,0028	12,9197	46,4386	−0,0679	36,4026	0,1017	0,0019	3,5963
	3,5972		12,9398	46,5472	0,0338			−0,0009	3,5963
3	3,5963	0,0000							
	3,5963								

§ 5.7. Método de iteraciones

Esencia del método. Uno de los métodos más importantes de la matemática de cálculos es el **método de iteraciones (método de aproximaciones sucesivas)**. La ventaja principal de este método consiste en que en cada paso se cumplen las operaciones del mismo tipo lo que facilita en grado considerable la programación para los ordenadores basada en los algoritmos iterativos.

La esencia del método de iteraciones consiste en lo siguiente. Consideremos la ecuación

$$f(x) = 0. \tag{1}$$

Supongamos que $f(x)$ es la función, continua sobre el segmento $[a, b]$, la cual se anula dentro de este intervalo, al menos, en un punto ξ. Se necesita precisar, como mínimo, una de sus raíces reales situados sobre $[a, b]$.

Reemplacemos la ecuación (1) por otra, equivalente, cuyas raíces son las mismas y que tiene la forma

$$x = \varphi(x). \tag{2}$$

Elijamos en calidad de aproximación inicial cualquier valor $x_0 \in [a, b]$, por ejemplo, $x_0 = (a + b)/2$. Luego calculemos $\varphi(x_0)$ y tomemos el número obtenido $x_1 = \varphi(x_0)$ por primer valor aproximado de la raíz ξ. Sustituyendo x_1 en vez de x en el segundo miembro de la ecuación (2), obtenemos un nuevo número $x_2 = \varphi(x_1)$. Continuando luego este proceso, llegamos a la sucesión de los números $x_0, x_1, x_2, \ldots, x_n, \ldots$, que se definen por las relaciones siguientes:

$$x_0 = (a + b)/2; \quad x_n = \varphi(x_{n-1}) \quad (n = 1, 2, \ldots). \tag{3}$$

Si existe $\lim\limits_{n\to\infty} x_n = \xi$ y la función $\varphi(x)$ es continua, entonces, pasando al límite en la igualdad (3), obtenemos

$$\xi = \varphi(\xi), \tag{4}$$

o sea, el límite ξ es la raíz de la ecuación (2) y, por consiguiente, de la ecuación (1). En virtud de la convergencia del proceso (3) esta raíz puede ser calculada, al ser bastante grande n, con toda exactitud prefijada.

Nótese que la representación (2) [o sea, la forma de la función $\varphi(x)$] puede realizarse por un conjunto infinito de procedimientos. Esto tiene gran importancia, ya que la forma de la función $\varphi(x)$ ejerce una influencia sustancial tanto en el mismo hecho de convergencia como también en la velocidad de esta última (si la convergencia está determinada).

Las condiciones de convergencia del proceso iterativo (3) se definen por el teorema siguiente.

Teorema. *Supongamos que se cumplen las condiciones:*

1°) *la función $\varphi(x)$ está definida y es derivable sobre el segmento $[a, b]$;*

2°) *todos los valores de $\varphi(x) \in [a, b]$ para $x \in [a, b]$;*

3°) *existe tal número $q < 1$ que para $x \in [a, b]$*

$$|\varphi'(x)| \leqslant q < 1. \tag{5}$$

Entonces el proceso iterativo (3) converge independientemente de la elección de la aproximación inicial $x_0 \in [a, b]$ y $\lim\limits_{n\to\infty} x_n$ es la raíz única y simple de la ecuación $x = \varphi(x)$ sobre el segmento $[a, b]$.

□ Planteemos las diferencias siguientes:

$$|x_1 - x_0| = |x_1 - x_0|,$$
$$|x_2 - x_1| = |\varphi(x_1) - \varphi(x_0)| = |\varphi'(c_1)|\,|x_1 - x_0| \leqslant \leqslant q\,|x_1 - x_0|,$$
$$|x_3 - x_2| = |\varphi(x_2) - \varphi(x_1)| = |\varphi'(c_2)| \cdot |x_2 - x_1| \leqslant \leqslant q^2\,|x_1 - x_0|,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$|x_n - x_{n-1}| = |\varphi(x_{n-1}) - \varphi(x_{n-2})| = = |\varphi'(c_n)| \cdot |x_{n-1} - x_{n-2}| \leqslant q^{n-1}\,|x_1 - x_0|.$$

Aquí $c_k \in [x_{k-1}, x_k]$ y $x_k \in [a, b]$ en virtud de la condición 2°.

Consideremos la serie con las sumas parciales siguientes: $S_{n+1} = x_0 + (x_1 - x_0) + (x_2 - x_1) + \ldots + (x_n - x_{n-1} \ldots) + + \ldots$ Es evidente que la suma $S_{n+1} = x_n$. La serie $x_0 + \sum\limits_{i=0}^{\infty} (x_{i+1} -$

$-x_i)$ es convergente, puesto que todos sus términos, comenzando con $x_1 - x_0$, no superan en valor absoluto los términos de la progresión geométrica con el denominador $q < 1$. Así pues, existe

$$\lim_{n\to\infty} S_{n+1} = \lim_{n\to\infty} x_n = \xi.$$

En virtud de la continuidad de la función $\varphi(x)$ tenemos $\lim_{n\to\infty} \varphi(x_{n-1}) = \varphi(\xi)$ y ya que $x_n = \varphi(x_{n-1})$, entonces $\varphi(\xi) = \xi$, o sea $\xi = \lim_{n\to\infty} x_n$ es la raíz de la ecuación (2).

Demostremos que esta raíz es la única. Sean ξ_1 y ξ_2 dos raíces de la ecuación (2), es decir, $\xi_1 = \varphi(\xi_1)$, $\xi_2 = \varphi(\xi_2)$. Entonces

$$|\xi_1 - \xi_2| = |\varphi(\xi_1) - \varphi(\xi_2)| = |\varphi'(c)| \cdot |\xi_1 - \xi_2|, \qquad (6)$$

donde $c \in (\xi_1, \xi_2)$. Transformemos la relación (6) de modo que tenga la forma

$$|\xi_1 - \xi_2| (1 - |\varphi'(c)|) = 0$$

y de la condición 3° se deduce que $\xi_1 = \xi_2$, o sea, no existen dos raíces diferentes.

Demostremos, por último, que la raíz obtenida es simple. Para esto basta demostrar que $x - \varphi(x)$ tiene la derivada que no se anula en ningún punto del segmento $[a, b]$. En efecto, $(x - \varphi(x))' = 1 - \varphi'(x)$ y en virtud de la condición 3° es evidente que esta expresión es positiva sobre el segmento $[a, b]$. ■

Estimación de error. Consideremos la diferencia entre los valores exacto y aproximado de la raíz ξ:

$$|\xi - x_n| = |\varphi(\xi) - \varphi(x_{n-1})| \leqslant q\,|\xi - x_{n-1}| =$$
$$= q\,|\xi - x_n + x_n - x_{n-1}| \leqslant q\,|\xi - x_n| + q\,|x_n - x_{n-1}|.$$

De aquí

$$|\xi - x_n| \leqslant \frac{q}{1-q}\,|x_n - x_{n-1}| \leqslant \frac{q^n}{1-q}\,|x_1 - x_0|. \qquad (7)$$

La relación (7) permite, inmediatamente después de la primera iteración (7) permite, inmediatamente después de la primera iteración, hallar el número máximo de iteraciones $N(\xi)$ necesario para calcular la raíz con el grado de precisión prefijado ε. En efecto, para que $|\xi - x_n|$ sea no más que ε es suficiente que

$$\frac{q^n}{1-q}\,|x_1 - x_0| \leqslant \varepsilon,$$

de donde

$$N(\varepsilon) \geqslant \frac{\log \dfrac{\varepsilon(1-q)}{|x_1 - x_0|}}{\log q}. \qquad (8)$$

Para $q \leqslant 1/2$ la estimación del error (7) se simplifica y toma la forma siguiente:

$$|\xi - x_n| \leqslant |x_n - x_{n-1}|. \qquad (9)$$

Según hemos señalado anteriormente, la forma de la ecuación $x = \varphi(x)$ no es indiferente para la convergencia del método de iteraciones. Ahora vamos a mostrar un procedimiento bastante general de construcción de la función $\varphi(x)$ para la cual está asegurado el cumplimiento de la condición 3°) del teorema.

Consideremos la ecuación inicial $f(x) = 0$. Supongamos que sobre el segmento $[a, b]$ existe la única raíz ξ de la ecuación, con la particularidad de que para $x \in [a, b]$ la derivada $f'(x)$ existe y conserva su signo así que

$$0 < m_1 \leqslant f'(x) \leqslant M_1, \quad \text{donde} \quad M_1 = \max_{[a,b]} f'(x); \quad m_1 = \min_{[a,b]} f'(x). \tag{10}$$

(sin alterar la generalidad, se puede considerar que $f'(x) > 0$).

Reemplacemos la ecuación $f(x) = 0$ por la ecuación equivalente

$$x = x - \lambda f(x) \tag{11}$$

y elijamos la constante λ de modo que se asegure el cumplimiento de la condición 3°).

Para la función $\varphi(x) \equiv x - \lambda f(x)$ la condición 3°) se escribe del modo siguiente:

$$|1 - \lambda f'(x)| < 1.$$

Vamos a resolver esta desigualdad; tenemos $-1 < 1 - \lambda f'(x) < 1$. de la segunda desigualdad obtenemos que $\lambda > 0$ y de la primera, que $\lambda < 2/f'(x)$, o sea,

$$0 < \lambda < 2/f'(x), \quad x \in [a, b],$$

o bien

$$0 < \lambda < 2/M_1.$$

De ordinario en calidad de λ se toma $1/M_1$. Ahora bien, obtenemos la ecuación

$$x = x - \frac{f(x)}{M_1} \tag{12}$$

y el proceso iterativo correspondiente tiene la forma

$$x_{n+1} = x_n - \frac{f(x_n)}{M_1}. \tag{13}$$

Con ayuda de los razonamientos análogos se puede mostrar fácilmente que en el caso de que $f'(x) < 0$ y $0 < m_1 \leqslant |f'(x)| \leqslant M_1$ obtenemos la ecuación

$$x = x + \frac{f(x)}{M_1} \tag{12'}$$

y el proceso iterativo correspondiente tomará la forma

$$x_{n+1} = x_n + \frac{f(x_n)}{M_1}. \tag{13'}$$

Supongamos ahora que, además de la condición (10) tiene lugar la relación

$$M_1 \leqslant 3m_1. \tag{14}$$

Entonces se puede exigir que se cumpla la desigualdad $|1 - \lambda f'(x)| \leqslant 1/2$. Resolviéndola, obtenemos para λ las restricciones siguientes:

$$\frac{1}{2m_1} \leqslant \lambda \leqslant \frac{3}{2M_1}. \tag{15}$$

Interpretación geométrica. Supongamos que se da la ecuación $f(x) = 0$ [$f(x)$ es la función continua]. Reduzcamos esta ecuación

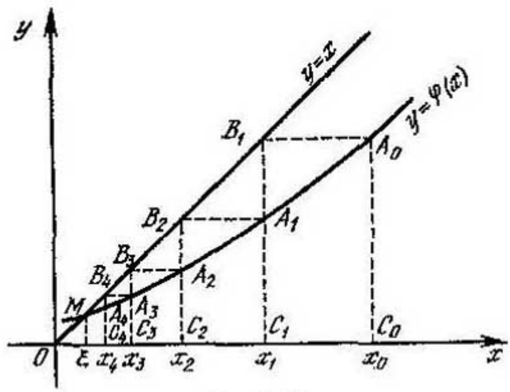

Fig. 5.25

a la forma $x = \varphi(x)$ y construyamos los gráficos de las funciones $y = x$ e $y = \varphi(x)$. La abscisa del punto de intersección de estas funciones es precisamente la raíz verdadera ξ (fig. 5.25).

Elijamos $x_0 \in [a, b]$ y determinemos $\varphi(x_0)$. Convengamos en designar la sucesión de los puntos que están sobre la curva $y = \varphi(x)$ con A_i ($i = 0, 1, 2, \ldots$) y la sucesión de los puntos que están sobre la recta $y = x$, con B_i ($i = 1, 2, 3, \ldots$). Del punto $A_0(x_0; \varphi(x_0))$ tracemos la recta paralela al eje Ox hasta que se interseque con la recta $y = x$; entonces obtenemos el punto $B_1(x_1; \varphi(x_0))$.

En efecto, $A_0C_0 = \varphi(x_0) = B_1C_1$, ya que $A_0B_1 \parallel OC_0$ y $B_1C_1 \parallel A_0C_0$. Pero $OC_1 = B_1C_1$ ($\triangle\, OC_1B_1$ es rectángulo e isósceles, puesto que la recta $y = x$ es la bisectriz del ángulo del cuadriculado). Por lo tanto, $x_1 = \varphi(x_0)$.

Tracemos $A_1B_2 \parallel OC_1$ y, repitiendo los razonamientos anteriormente expuestos, nos convencemos de que $x_2 = \varphi(x_1)$.

En la fig. 5.25 está representado el proceso iterativo convergente. La curva corta la bisectriz $y = x$ en el punto M con la abscisa ξ y para $x > \xi$ está debajo de la bisectriz; $\varphi'(x)$ satisface la condición

$0 < \varphi'(x) < 1$. Las aproximaciones sucesivas $x_0, x_1, \ldots, x_n, \ldots$ (las abscisas comunes de los puntos de los gráficos de ambas funciones) **decrecen monótonamente**. Cada aproximación sucesiva x_n es más próxima a la raíz verdadera que cada precedente x_{n-1}. La línea quebrada $A_0B_1A_1B_2A_2 \ldots$ tiene la forma de «escalera».

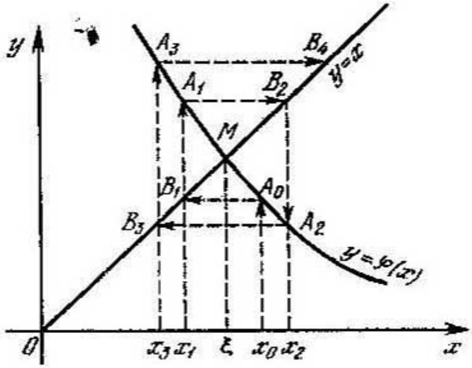

Fig. 5.26

En la fig. 5.26 la derivada $\varphi'(x) < 0$, pero en valor absoluto es menor que la unidad, o sea, $|\varphi'(x)| < 1$. El proceso iterativo converge, pero las aproximaciones **oscilan** cerca del valor exacto de la raíz. La quebrada $A_0B_1B_2A_2 \ldots$ tiene la forma de «espiral».

Ahora bien, si en cierto entorno (a, b) de la raíz ξ de la ecuación $x = \varphi(x)$ la derivada $\varphi'(x)$ conserva el signo constante y está cumplida la desigualdad $|\varphi'(x)| \leqslant q < 1$, con la particularidad de que $\varphi'(x) > 0$, entonces las aproximaciones sucesivas $x_n = \varphi(x_{n-1})$ $(n = 1, 2, \ldots)$, $x_0 \in [a\ b]$ convergen monótonamente a la raíz. En cambio, en el caso en que $\varphi'(x) < 0$, las aproximaciones sucesivas oscilan alrededor de la raíz ξ.

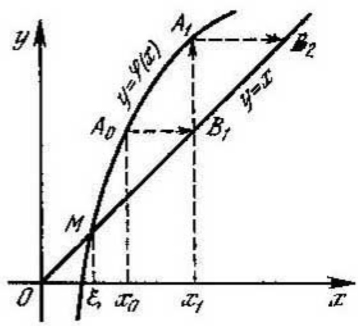

Fig. 5.27

En la fig. 5.27 se muestra el proceso iterativo divergente. Aquí $\varphi'(x) > 1$. La curva corta la bisectriz $y = x$ en el punto M y para $x > \xi$ está por encima de la bisectriz.

En la fig. 5.28 se representa el proceso iterativo divergente para el caso $|\varphi'(x)| > 1$. Las «aproximaciones» sucesivas se alejan del valor exacto de la raiz ξ.

Ejemplo 1. Con ayuda del método de iteraciones precisar hasta 10^{-4} la raíz de la ecuación $5x^3 - 20x + 3 = 0$, encerrada en el segmento $[0, 1]$.

△ La ecuación dada ha de reducirse a la forma $x = \varphi(x)$. Se puede hacer esto por varios procedimientos, por ejemplo:

$x = x + (5x^3 - 20x + 3)$, entonces $\varphi_1(x) = 5x^3 - 19x + 3$;
$x = \sqrt[3]{(20x - 3)/5}$, entonces $\varphi_2(x) = \sqrt[3]{(20x - 3)/5}$;
$x = (5x^3 + 3)/20$, entonces $\varphi_3(x) = (5x^3 + 3)/20$.

Determinemos cuál de las funciones obtenidas $\varphi(x)$ debemos utilizar para calcular las aproximaciones sucesivas. Recuérdese

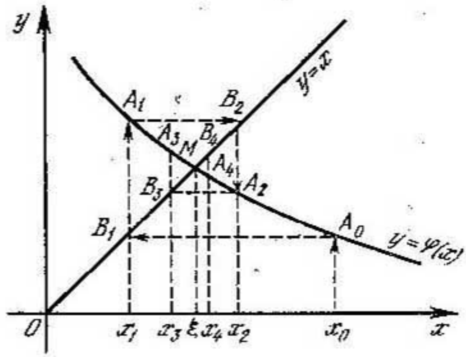

Fig. 5.28

que si $\varphi(x)$ satisface sobre el segmento $[a, b]$ la condición $|\varphi'(x)| \leqslant q < 1$, el proceso iterativo converge. Hallamos

$$|\varphi_1'(x)| = |15x^2 - 19| > 1 \quad \text{sobre} \quad [0, 1],$$
$$|\varphi_3'(x)| = 15x^2/20 = 3x^2/4 < 1 \quad \text{sobre} \quad [0, 1].$$

Por consiguiente, se puede utilizar la función $\varphi_3(x)$ y buscar las aproximaciones sucesivas por el método de iteraciones valiéndose de la fórmula $x_n = (5x_{n-1}^3 + 3)/20$. Tomemos por aproximación inicial máx $\varphi'(x)$ sobre $[0, 1]$, o sea, $x_0 = 0,75$. Haciendo uso de la fórmula (7) determinemos cuál ha de ser la diferencia entre dos aproximaciones sucesivas para que quede alcanzada la exactitud prefijada:

$$|x_n - x_{n-1}| \leqslant \frac{0,0001 \cdot (1 - 0,75)}{0,75} = \frac{0,0001 \cdot 0,25}{0,75} = 0,00003.$$

Ahora bien, cuando el valor absoluto de la diferencia $|x_n - x_{n-1}|$ no supere 0,00003, será necesario cesar el proceso iterativo y suponer que la exactitud prefijada está alcanzada.

Es cómodo realizar el cálculo con ayuda de la tabla 5.11

Aquí se puede terminar el proceso iterativo y considerar $\xi = 0,1514$. ▲

Ejemplo 2. Calcular con exactitud hasta $\varepsilon = 10^{-5}$ la raíz de la ecuación $e^x - x^2 = 0$.

Tabla 5.11

n	x_n	x_n^3	$\varphi(x_n) = x_{n+1}$
0	0,75	0,42188	0,25547
1	0,2555	0,016777	0,154144
2	0,1541	0,005652	0,151413
3	0,1514	0,005443	0,151361
4	0,15136	0,005442	0,151361

△ Escribamos la ecuación en la forma $e^x = x^2$ y separemos gráficamente las raíces. Construyamos los gráficos de las funciones $y = e^x$ e $y = x^2$ (fig. 5.29). Del dibujo se ve que la ecuación $e^x - x^2 = 0$ tiene una raíz real que está sobre el segmento $[-0{,}8;\ -0{,}7]$.

Verifiquemos si verdaderamente es así. Separemos $f(-0{,}8)$ y $f(-0{,}7)$; tenemos $f(-0{,}8) = 0{,}44933 - 0{,}64 = -0{,}19067 < 0$, $(f-0{,}7) = 0{,}49659 - 0{,}49 = 0{,}00659 > 0$. Puesto que los signos de la función $f(x) = e^x - x^2$ son opuestos en los extremos del segmento $[-0{,}8;\ -0{,}7]$, la raíz de la ecuación está dentro de este segmento.

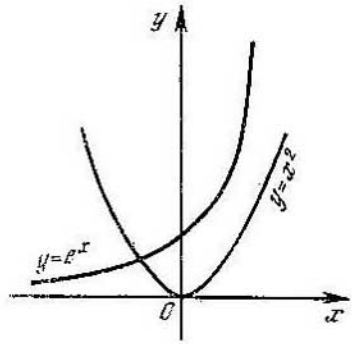

Fig. 5.29

Probemos hacer más estrecho el segmento, aplicando el método de pruebas. Encontramos $f(-0{,}75) = 0{,}49237 - 0{,}56250 < 0$ y $f(-0{,}7) > 0$. Por lo tanto, la raíz se halla sobre el segmento $[-0{,}75;\ 0{,}7]$. Estrechemos este segmento una vez más. Tenemos $f(-0{,}725) = 0{,}48432 - 0{,}52562 = -0{,}4130 < 0$ y $f(-0{,}7) > 0$. Por consiguiente la raíz está sobre el segmento $[-0{,}725;\ -0{,}7]$.

De la ecuación $e^x = x^2$ obtenemos que $x = -\sqrt{e^x}$ (delante del radical tomamos el signo «—», ya que conocemos que la raíz es negativa). Escribamos esta ecuación en la forma $x = e^{x/2}$ y comprobemos cuál será el proceso iterativo: convergente o divergente, o sea, se cumple o no la desigualdad $|\varphi'(x)| < 1$. En el ejemplo dado

$$\varphi(x) = -e^{x/2},\quad \varphi'(x) = (1/2)\,e^{x/2},\quad |\varphi'(-0{,}725)| = 0{,}34727;$$
$$|\varphi'(x)| = |\varphi'(-0{,}7)| = 0{,}35230.$$

Puesto que $|\varphi'(x)| < 1$, el proceso iterativo converge. En la fórmula (7) tomemos el número q igual a 0,36. Puesto que $\varepsilon = 10^{-5}$, entonces

$$|x_n - x_{n-1}| \leqslant \frac{0{,}00001\,(1-0{,}36)}{0{,}36} = 0{,}000018.$$

Ahora bien, la exactitud requerida puede ser alcanzada si se cumple la desigualdad $|x_n - x_{n-1}| \leq 0{,}000018$. Por aproximación nula se puede tomar cada uno de los extremos del segmento $[-0{,}725; -0{,}7]$ y todo punto dentro de este último. Tomemos $x_0 = -0{,}7$.

Hagamos los cálculos con ayuda de la tabla siguiente:

Tabla 5.12

n	x_n	$x_n/2$	$e^{x_n/2}$
0	−0,7	−0,35	−0,70460
1	−0,70460	−0,35230	−0,70307
2	−0,70307	−0,35154	−0,70360
3	−0,70360	−0,35180	−0,70342
4	−0,70342	−0,35171	−0,70348
5	−0,70348	−0,35174	−0,70346
6	−0,70346	−0,35173	−0,70347
7	−0,70347		

Puesto que $|x_7 - x_6| = |-0{,}70347 - (-0{,}70346)| = 0{,}00001$, la exactitud requerida de cálculos está alcanzada y $\xi \approx -70347$. ▲

Ejemplo 3. Haciendo uso del método de iteraciones, calcular con exactitud hasta 0,001 la raíz de la ecuación $x^3 + 3x^2 - 3 = 0$, situada sobre el segmento $[-2{,}75; -2{,}5]$ (véanse el ejemplo 1 del § 5.4 y el ejemplo 1 del § 5.5).

△ Encontramos $f'(x) = 3x^2 + 6x$. Por lo tanto,

$$M_1 = \max_{[-2,75;\ -2,5]} |f'(x)| = 6{,}189;\ m_1 = \min_{[-2,75;\ -2,5]} |f'(x)| = 3{,}75;$$

$$q = 1 - \frac{m_1}{M_1} = 1 - \frac{3{,}75}{6{,}189} < \frac{1}{2}.$$

Puesto que $q < 1/2$, para estimar el error se puede utilizar la fórmula (9). Tomemos $M_1 = 6$, entonces $\lambda = 1/6$ y

$$\varphi(x) = x - \lambda f(x) = x - \frac{1}{6}(x^3 + 3x^2 - 3).$$

El proceso iterativo correspondiente tiene la forma

$$x_{n+1} = x_n - \frac{1}{6}(x_n^3 + 3x_n^2 - 3);$$

entonces $|x_{n+1} - x_n| = \frac{1}{6}(x_n^3 + 3x_n^2 - 3)$. Los cálculos han de cesarse tan pronto como $|x_{n+1} - x_n| < \varepsilon$.

Realicemos los cálculos con ayuda de la tabla siguiente:

Tabla 5.13

n	x_n	x_n^3	$3x_n^2$	$\frac{1}{6}(x^3+3x^2-3)$
0	−2,5	−15,625	18,75	0,02
1	−2,52	−16,0030	19,0512	0,0080
2	−2,5280	−16,1559	19,1724	0,0028
3	−2,5308	−16,2096	19,2148	0,0008
4	−2,5316			

Así pues, por valor aproximado de la raíz con exactitud hasta 0,001 se puede tomar $\xi = -2{,}532$. ▲

§ 5.8. Propiedades generales de las ecuaciones algebraicas. Determinación de la cantidad de raíces reales de una ecuación algebraica

Propiedades generales de las ecuaciones algebraicas. Escribamos una ecuación algebraica de n-ésimo grado:

$$P_n(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n = 0, \qquad (1)$$

donde $P_n(x)$ es el polinomio de n-ésimo grado; n, el grado superior de la incógnita; $a_0, a_1, \ldots, a_n$ son los coeficientes reales.

Es sabido que todo número ξ que anula el polinomio, o sea tal que $P_n(\xi) = 0$, se llama raíz del polinomio.

El número ξ es la raíz del polinomio $P_n(x)$ si y sólo si $P_n(x)$ se divide exactamente por $x - \xi$. Recuérdese que si en este caso $P_n(x)$ se divide exactamente por $(x - \xi)^k$ $(k \geqslant 1)$, pero no se divide ya por $(x - \xi^{k+1})$, ξ se denomina raíz de k-ésimo grado de multiplicidad (o raíz de multiplicidad k) del polinomio $P_n(x)$. Las raíces de multiplicidad $k = 1$ se llaman raíces simples del polinomio.

Surge la pregunta ¿todo polinomio tiene obligatoriamente raíces? La respuesta a esta pregunta se da por el teorema siguiente que citamos sin demostración.

Teorema 1 (teorema fundamental del álgebra). *Todo polinomio con todos coeficientes reales el grado del cual no sea menor que la unidad tiene, al menos, una raíz que en el caso general es compleja.*

De este teorema se deduce un corolario importante que dice: *todo polinomio $P_n(x)$ de grado n $(n \geqslant 1)$ con cualesquiera coeficientes reales tiene exactamente n raíces, reales o complejas, si cada una de las raíces se cuenta tantas veces cuanta es su multiplicidad.*

Ahora bien, las raíces de la ecuación algebraica (1) pueden ser tanto reales como complejas.

Las raíces complejas de la ecuación (1) poseen la propiedad de *conjugación par*, o sea, si la ecuación (1) tiene una raíz compleja

$\xi = \alpha + \beta i$ (donde α y β son los números reales) de multiplicidad k, esta ecuación tiene asimismo la raíz compleja $\overline{\xi} = \alpha - \beta i$ también de multiplicidad k. Los módulos de estas raíces son iguales: $|\xi| = |\overline{\xi}| = \sqrt{\alpha^2 + \beta^2}$.

Si la ecuación (1) tiene raíces complejas, su número es par. Por eso toda ecuación algebraica de grado impar con coeficientes reales tiene, como mínimo, una raíz real.

Antes de calcular las raíces de una ecuación algebraica es necesario primeramente: a) determinar la cantidad de raíces que posee la ecuación dada; b) hallar el dominio de existencia de las raíces (determinar las cotas superior e inferior de situación de las raíces). Luego se puede comenzar a separar las raíces y precisarlas.

Determinación de la cantidad de raíces reales de las ecuaciones algebraicas. La cantidad de raíces reales positivas de la ecuación algebraica (1)

$$P_n(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n = 0$$

puede ser determinada aproximadamente con ayuda de la **regla de Descartes:** *la cantidad de raíces positivas reales de la ecuación algebraica $P_n(x) = 0$ con coeficientes reales (raíces cada una de las cuales se cuenta tantas veces cuanta es su multiplicidad) ora es igual al número de cambios del signo en la sucesión de los coeficientes de la ecuación $P_n(x) = 0$ ora es en un número par menor (los coeficientes iguales a cero no se tienen en cuenta).*

La cantidad de raíces negativas es igual al número de cambios del signo en la sucesión de los coeficientes $P_n(-x)$ o en un número par es menor.

Si la ecuación es completa, la cantidad de sus raíces positivas es igual al número de cambios del signo en la sucesión de los coeficientes o en un número par es menor y la cantidad de raíces negativas es igual al número de constancias del signo o en un número par es menor.

Ejemplo 1. Determinar la cantidad de raíces positivas y negativas de la ecuación $x^5 - 17x^4 + 12x^3 + 7x^2 - x + 1 = 0$.

△ Según el teorema fundamental del álgebra esta ecuación tiene cinco raíces (de ellas una raíz, como mínimo, es real).

La ecuación es completa, la sucesión de signos de los coeficientes es tal: $+, -, +, +, -, +$. El signo cambia cuatro veces, por consiguiente, hay cuatro o dos raíces positivas o bien ninguna raíz es positiva.

El número de constancias del signo es igual a 1, por lo tanto, la ecuación dada tiene una raíz negativa. ▲

Ejemplo 2. Determinar la cantidad de raíces reales positivas y negativas de la ecuación $x^6 - 3x^4 + x^3 + x^2 - 1 = 0$.

△ La ecuación dada tiene seis raíces; la sucesión de los signos $+, -, +, -$. Tiene lugar tres cambios del signo; por consiguiente, hay tres raíces positivas o una raíz positiva. Luego, para el polinomio

$$P_n(-x) = x^6 - 3x^4 - x^3 + x^2 - 1$$

la sucesión de los signos es tal: +, —, —, +, —. Aquí también tenemos tres cambios del signo, por eso hay tres raíces negativas o una raíz negativa. ▲

La cantidad de raíces de una ecuación algebraica se puede determinar más exactamente con ayuda del teorema de Sturm.

Puesto que las raíces múltiplas de la ecuación pueden siempre ser separadas como raíces generales de las ecuaciones $P_n(x) = 0$ y $P'_n(x) = 0$, sin limitar la generalidad se puede considerar que la ecuación dada $P_n(x) = 0$ tiene sólo raíces simples.

Supongamos que para la ecuación algebraica dada $P_n(x) = 0$ hemos determinado, por un procedimiento cualquiera, que todas sus raíces reales están en el intervalo (a, b) (a y b son los números reales y no son raíces de la ecuación; $a < b$). Determinemos la primera derivada $P'_n(x)$ y dividamos por ésta el polinomio $P_n(x)$. Tomemos el resto de la división de $P_n(x)$ por $P'_n(x)$ con el signo contrario y designémoslo con $R_1(x)$.

Luego, de un modo exactamente igual dividamos $P'_n(x)$ por $R_1(x)$; tomemos con el signo contrario el resto obtenido y designémoslo con $R_2(x)$. Al dividir $R_1(x)$ por $R_2(x)$ y al volver a tomar el resto con el signo opuesto, obtenemos $R_3(x)$. Continuamos el proceso de división hasta que se obtenga un resto que sea una magnitud constante. Tomemos esta magnitud también con el signo contrario.

Obtenemos una sucesión de funciones

$$P_n(x),\ P'_n(x),\ R_1(x),\ R_2(x),\ \ldots,\ R_{m-1}(x),\ R_m = \text{const}$$

que se llama *sistema de Sturm*. En esta sucesión sustituimos en vez de x primero a, luego b y contamos el número de cambios del signo en ambos casos [designemos los números obtenidos con $W(a)$ y $W(b)$, respectivamente].

Teorema 2 (teorema de Sturm). *Si los números reales a y b ($a < b$) no son raíces del polinomio $P_n(x)$, que no tiene raíces múltiplas, entonces $W(a) \geqslant W(b)$ y la diferencia $W(a) - W(b)$* es *igual a la cantidad de raíces reales del polinomio $P_n(x)$, encerradas entre a y b.*

Con ayuda del teorema de Sturm se puede determinar la cantidad de raíces negativas de la ecuación $P_n(x) = 0$, [o sea, la cantidad de raíces reales de la ecuación $P_n(x) = 0$ en el intervalo $(-\infty, 0)$], así como la cantidad de raíces positivas [en el intervalo $(0, +\infty)$]. El teorema de Sturm se aplica también para separar las raíces. Las funciones que forman parte del sistema de Sturm pueden multiplicarse y dividirse por números positivos arbitrarios. Esto simplifica mucho el trabajo cuando se realiza la división inexacta.

Ejemplo 3. Determinar la cantidad de raíces reales de la ecuación $5x^3 - 20x + 3 = 0$, así como separar estas raíces haciendo uso del teorema de Sturm.

△ Formamos el sistema de funciones de Sturm. Tenemos $P_n(x) = 5x^3 - 20x + 3$; $P'_n(x) = 15x^2 - 20$. Para determinar

$R_1(x)$ multipliquemos $P_n(x)$ por 3 y dividamos el resultado por $P'_n(x)$

$$\begin{array}{r|l} 15x^3-60x+9 & 15x^2-20 \\ \mp 15x^3 \pm 20x & \overline{\quad x \quad} \\ \hline -40x+9 & \end{array}$$

Por lo tanto, $R_1(x) = 40x - 9$ (hemos tomado el resto con el signo contrario). Multipliquemos $P'_n(x)$ por 8 y dividamos este producto por $R_1(x)$:

$$\begin{array}{r|l} 120x^2-160 & 40x-9 \\ \mp 120x^2 \pm 27x & \overline{3x+27} \\ \hline 40\,(27x-160) & \\ \hline 40\cdot 27x-40\cdot 160 & \\ \mp 40\cdot 27x \pm 9\cdot 27 & \\ \hline - & \end{array}$$

Puesto que el último resto es un número constante con el signo «—» (en este caso nos interesa precisamente el signo del resto constante), lo cambiamos por el opuesto, es decir, por «+».

Componemos la siguiente tabla de los signo de funciones que forman parte del sistema de Sturm:

x	$P_n(x)$	$P'_n(x)$	$R_1(x)$	R_2	$W(x)$
$-\infty$	$-$	$+$	$-$	$+$	3
0	$+$	$-$	$-$	$+$	2
$+\infty$	$+$	$+$	$+$	$+$	0

De la tabla se ve que en el intervalo $(-\infty, +\infty)$ hay tres raíces reales [ya que $W(-\infty) - W(+\infty) = 3 - 0 = 3$]. Una de ellas es negativa [$W(-\infty) - W(0) = 3 - 2 = 1$] y dos son positivas [$W(0) - W(+\infty) = 2 - 0 = 2$].

Utilizando este teorema de Sturm, separamos las raíces al abreviar los intervalos hasta la longitud igual a 1:

x	$P_n(x)$	$P'_n(x)$	$R_1(x)$	$R_2(x)$	$W(x)$
$-$	$-$	$+$	$-$	$+$	3
-3	$-$	$+$	$-$	$+$	3
-2	$+$	$+$	$-$	$+$	2
-1	$+$	$-$	$+$	$+$	2
0	$+$	$-$	$-$	$+$	2
1	$-$	$-$	$+$	$+$	1
2	$+$	$+$	$+$	$+$	0

De la última tabla se ve que las raíces están situadas en los intervalos $(-3, -2)$, $(0, 1)$, $(1, 2)$. ▲

§ 5.9. Determinación del dominio de existencia de las raíces de una ecuación algebraica

Regla del anillo. *Sea dada una ecuación algebraica* $P_n(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n = 0$, *donde* a_0, $a_1, \ldots, a_n$ *son los coeficientes reales y sea* $A = \text{máx}\{|a_1|, |a_2|, \ldots, |a_n|\}$, $B = \text{máx}\{|a_0|, |a_1|, \ldots, |a_{n-1}|\}$.
Entonces las raíces de la ecuación están encerradas en el anillo circular $r < |x| < R$, *donde*

$$r = \frac{1}{1 + B/|a_n|}; \quad R = 1 + \frac{A}{|a_0|}.$$

En este caso r es la cota inferior y R, la cota superior de las raíces positivas de la ecuación algebraica $P_n(x) = 0$ y $-R$, $-r$ son las cotas inferior y superior, respectivamente, de las raíces negativas (fig. 5.30).

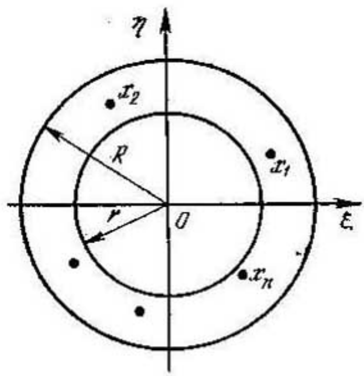

Fig. 5.30

Ejemplo 1. Determinar las cotas de las raíces de la ecuación $5x^3 - 20x + 3 = 0$.

△ Aquí $|a_0| = 5$, $A = 20$, $|a_n| = 3$, $B = 20$, o sea,

$$R = 1 + \frac{A}{|a_0|} = 1 + \frac{20}{5} = 5;$$

$$r = \frac{1}{1 + B/|a_n|} = \frac{1}{1 + 20/3} = \frac{3}{23} \approx 0{,}013.$$

Entonces si las raíces reales de la ecuación $5x^3 - 20x + 3 = 0$ existen (y éstas existen obligatoriamente, ya que la ecuación es de grado impar), se disponen en el intervalo $(-5;\ 5)$; en este caso las raíces negativas están en el intervalo $(-5;\ -0{,}013)$ y las positivas, en el intervalo $(0{,}013;\ 5)$. ▲

Al resolver las ecuaciones es cómodo determinar primero las cotas de las raíces y luego aplicar el teorema de Sturm. Según la regla del anillo estas cotas se determinan muy aproximadamente.

Vamos a mostrar el procedimiento de una determinación más exacta de las raíces reales de la ecuación algebraica $P_n(x) = 0$. Si R_1 es la cota superior de las raíces positivas de $P_n(x)$, R_2 es la cota superior de las raíces positivas de $P_n(-x)$, $R_3 > 0$ es la cota superior de las raíces positivas de $x^nP_n(1/x)$ y R_4 es la cota superior de las raíces positivas de $x^nP_n(-1/x)$, entonces todas las raíces reales, distintas del cero, de la ecuación $P_n(x) = 0$ (si éstas existen) están dentro de los intervalos $(-R_2, -1/R_4)$ y $(1/R_3, R_1)$.

Para determinar la cota superior de las raíces positivas de una ecuación algebraica se puede utilizar los métodos de Lagrange o de Newton.

Método de Lagrange. *Si los coeficientes del polinomio*

$$P_n(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_n$$

satisfacen las condiciones $a_0 > 0$, $a_1, a_2, \ldots, a_{m-1} \geqslant 0$, $a_m < 0$, *la cota superior de las raíces positivas de las ecuaciones* $P_n(x) = 0$ *se determina con ayuda de la fórmula* $R = 1 + \sqrt[m]{B/a_0}$, *donde* B *es el máximo entre los valores absolutos de los coeficientes negativos de* $P_n(x)$.

Ejemplo 2. Haciendo uso del método de Lagrange, determinar las cotas de las raíces positivas y negativas de la ecuación

$$8x^4 - 8x^2 - 32x + 1 = 0.$$

△ Aquí $a_0 = 8 > 0$; $a_1 = 0$; $a_2 = -8 < 0$; $a_3 = -32$; $a_4 = 1$, $m = 2$ (número del primero entre los coeficientes negativos), $B = 32$. Por consiguiente, $R_1 = 1 + \sqrt{32/8} = 3$.

Consideremos el polinomio

$$P_n(-x) = 8x^4 - 8x^2 + 32x + 1.$$

De un modo análogo encontramos que para este polinomio como cota superior de las raíces positivas sirve $R_2 = 1 + \sqrt{8/8} = 2$.

Luego, para el polinomio

$$x^4P_n(1/x) = x^4 - 32x^3 - 8x^2 + 8$$

tenemos $a_0 = 1 > 0$; $a_1 = -32 < 0$, o sea, $m = 1$; $B = 32$; $R_3 = 1 + 32 = 33$.

Por último, para el polinomio

$$x^4P_n(-1/x) = x^4 + 32x^3 - 8x^2 + 8$$

tenemos $a_0 = 1 > 0$; $a_1 = 32$; $a_2 = -8$; $a_3 = 0$; $a_4 = 8$, o sea, $m = 2$. Por eso $R_4 = 1 + \sqrt{8} = 1 + 2\sqrt{2} = 3{,}828$.

Por lo tanto, si la ecuación $8x^4 - 8x^2 - 32x + 1 = 0$ tiene raíces reales, éstas se hallan obligatoriamente en los intervalos $(-2, -1/3{,}828)$ y $(1/33, 3)$. ▲

Método de Newton. *Si para* $x = c$ *el polinomio*

$$P_n(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n$$

y sus derivadas $P_n'(x)$, $P_n''(x)$, ... *toman los valores positivos, entonces* c *es la cota superior de las raíces positivas de la ecuación* $P_n(x) = 0$.

Ejemplo 3. Haciendo uso del método de Newton, determinar la cota superior de las raíces positivas de la ecuación $8x^4 - 8x^2 - 32x + 1 = 0$.

△ Hallamos

$$P(x) = 8x^4 - 8x^2 - 32x + 1; \quad P'(x) = 32x^3 - 16x - 32;$$
$$P''(x) = 96x^2 - 16; \quad P'''(x) = 192x; \quad P^{IV}(x) = 192.$$

Han de verificarse los valores de $x > 0$. Para $x = c = 1$ tenemos $P(1) < 0$. Por consiguiente, luego no debe realizarse la verificación para $x = 1$. Comprobemos el valor $x = c = 2$: $P(2) > 0$; $P'(2) > 0$; $P''(2) > 0$; $P'''(2) > 0$; $P^{IV} > 0$. Ahora bien, de cota superior de las raíces positivas sirve el número 2, o sea, $R = 2$. En calidad de la cota inferior se puede tomar el número inverso a R, o sea, $r = 1/2$. ▲

§ 5.10. Método de Horner para precisar las raíces reales de una ecuación algebraica

Sea dada la ecuación algebraica

$$P(x) = x^n + a_1 x^{n-1} + a_2 x^{n-2} + \ldots + a_n = 0, \tag{1}$$

donde $a_1, a_2, \ldots, a_n$ son los coeficientes reales del polinomio.

Representemos la raíz buscada de la ecuación, escrita en el sistema decimal, en la forma

$$X = c_0 \cdot 10^m + c_1 \cdot 10^{m-1} + c_2 \cdot 10^{m-2} + \ldots + c_k \cdot 10^{m-k} + \ldots$$

El método de Horner consiste en determinar sucesivamente las cifras de la raíz $c_0, c_1, \ldots$ con ayuda de transformaciones especiales de la ecuación (1).

Si se emplea la sustitución $x = 10^m \xi$ (para $c_0 > 0$) o bien $x = -10^m \xi$ (para $c_0 < 0$), la ecuación (1) se transforma en ecuación

$$f_1(\xi) = \xi^n + a_1' \xi^{n-1} + a_2' \xi^{n-2} + \ldots + a_n' = 0$$

cuya raíz está encerrada en el intervalo $(0, 10)$. Por eso a continuación vamos a examinar precisamente este caso particular. Esto simplifica el proceso de cálculo de la raíz, aunque no es obligatorio para la aplicación del método de Horner.

Puesto que $0 < X < 10$, la raíz de la ecuación (1) se puede escribir en la forma

$$X = c_0 + \frac{c_1}{10} + \frac{c_2}{10^2} + \frac{c_3}{10^3} + \ldots = c_0,\ c_1 c_2 c_3 \ldots$$

Es fácil hallar la primera cifra de la raíz c_0 con ayuda del método de tabla o utilizando el intervalo $[a, b]$ del aislamiento de la raíz.

Luego, aplicando a la ecuación (1) la transformación

$$x - c_0 = y, \tag{2}$$

obtenemos $x = y + c_0$, de donde

$$P(x) = P(y + c_0) = -\varphi(y) = 0. \tag{3}$$

Como raíz de la última ecuación sirve, evidentemente, el número $y = 0,\ c_1 c_2 c_3 \ldots$ Aplicando a la ecuación (3) la sustitución

$$y = Y/10, \tag{4}$$

llegamos a la ecuación

$$P_1(Y) = 0, \tag{3'}$$

que tiene la raíz $Y = 10y = c_1, c_2c_3 \ldots$ De la ecuación (3') determinamos la cifra c_1 que es la segunda cifra de la raíz de la ecuación (1). Luego se puede volver a aplicar la sustitución del tipo (2) ó (4), pero ahora a la ecuación (3'). Como resultado obtenemos la ecuación

$$P_2(Z) = 0, \tag{5}$$

de cuya raíz sirve el número $z = c_2, c_3, \ldots$ Después se determina la cifra c_2.

El proceso descrito puede ser repetido hasta que se obtenga la cantidad requerida de cifras. El método de Horner puede aplicarse en combinación con cualquier otro método, por ejemplo, con el de las cuerdas, con el de las tangentes o bien con el combinado de las cuerdas y las tangentes. Primero se determinan algunas cifras de la raíz con ayuda del método de Horner y luego la raíz se precisa por otros métodos.

Cuando se necesita obtener una pequeña cantidad de cifras de la raíz, se puede utilizar el procedimiento siguiente.

Si una de las raíces de la ecuación es mucho menor que las demás, esta raíz puede ser calculada aproximadamente por dividir el término independiente, tomado con el signo «—», por el coeficiente de la x de primer grado.

¿Cómo realmente se realizan las sustituciones (2) y (4)?

Para ejecutar la sustitución (4) conviene multiplicar los coeficientes $a_1, a_2, \ldots, a_n$ de la ecuación inicial por $10, 10^2, 10^3, \ldots$ $\ldots, 10^n$, respectivamente.

La sustitución (2) se lleva a cabo después de desarrollar previamente el polinomio $P(x)$ según las potencias de $x - c_0$. Tal desarrollo puede ser efectuado por la fórmula de Taylor o bien utilizando la siguiente serie de igualdades:

$$\begin{aligned} P(x) &= q_1(x)(x - c_0) + r_0, \\ q_1(x) &= q_2(x)(x - c_0) + r_1, \\ q_2(x) &= q_3(x)(x - c_0) + r_2, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ q_{n-2}(x) &= q_{n-1}(x)(x - c_0) + r_{n-2}, \\ q_{n-1}(x) &= (x - c_0) + r_{n-1}, \end{aligned} \tag{6}$$

donde $q_i(x)$ y r_{i-1} son el cociente y el resto, respectivamente, que se obtienen al dividir $q_{i-1}(x)$ por $(x - c_0)$. Eliminando sucesivamente $q_i(x)$ de la igualdad (6), llegamos a la identidad

$$P(x) = (x - c_0)^n + r_{n-1}(x - c_0)^{n-1} + r_{n-2}(x - c_0)^{n-2} + \ldots$$
$$\ldots + r_1(x - c_0) + r_0, \tag{7}$$

o sea, obtenemos el desarrollo de $f(x)$ según las potencias de $x - c_0$. El cálculo de cada uno de los restos r_i $(i = 0, 1, \ldots, n - 1)$, o sea de los coeficientes del desarrollo, se lleva a cabo con ayuda del esquema de Horner.

El sistema de igualdades (6) corresponde a la tabla

c_0

1	a_1	a_2	a_3	$\ldots$	a_{n-1}	a_n
1	b_1	b_2	b_3	$\ldots$	b_{n-1}	r_0
1	e_1	e_2	e_3	$\ldots$	r_1	
.	..	..	..	..	..	..
1	r_{n-1}					

donde $b_1 = 1 \cdot c_0$, $b_2 = b_1 c_0$, $b_3 = b_2 c_0$, . . ., $e_1 = 1 \cdot c_0$, $e_2 = e_1 c_0$, $e_3 = e_2 c_0$, . . . Si el coeficiente de x^n es igual a a_0, en la primera columna en vez de la unidad debe ser a_0.

Ejemplo. Haciendo uso del método de Horner hallar la menor raíz de la ecuación $x^3 + 3x^2 - 3 = 0$ con seis cifras significativas; las raíces de la ecuación están separadas y la menor de ellas se halla sobre el segmento $[-3, -2]$.

△ 1) Puesto que la raíz es negativa, transformemos la ecuación inicial con ayuda de la sustitución $x = -\tilde{x}$:

$$-\tilde{x}^3 + 3\tilde{x}^2 - 3 = 0 \quad \text{o bien} \quad \tilde{x}^3 - 3\tilde{x}^2 + 3 = 0.$$

La raíz buscada de esta ecuación $\tilde{x} \in [2, 3]$. Por lo tanto, la primera cifra de la ecuación transformada es 2. Entonces las sustituciones (2) y (4) tienen la forma $\tilde{x} - 2 = y$ e $y = Y/10$. Realizamos la primera sustitución con ayuda del esquema de Horner:

$c_0 = 2$

1	−3	0	3
1	−1	−2	$\boxed{-1}$
1	1	$\boxed{0}$	
1	$\boxed{3}$		

De esta tabla según la fórmula (7) obtenemos $\varphi(y) = y^3 + 3y - 1$ y $P_1(Y) = Y^3 + 30Y^2 - 1000 = 0$.

A continuación vamos a escribir inmediatamente con ayuda de la tabla ambas transformaciones juntas, conservando la designación y, es decir, en vez de la última ecuación vamos a escribir $P_1(y) = y^3 + 30y^2 - 1000 = 0$.

Vamos a determinar los valores del polinomio $P_1(y)$ para ciertos valores de $y \in [0, 10]$ utilizando el esquema de Horner:

y	1	30	0	−1000
5	1	35	175	−125
6	1	36	216	296

Puesto que $P_1(5) < 0$ y $P_1(6) > 0$, entonces $y \in [5, 6]$. De aquí se deduce que $c_1 = 5$.

3) Para determinar la cifra siguiente planteemos la ecuación, cuyos coeficientes obtenemos con ayuda del esquema de Horner, haciendo uso de las sustituciones (2) y (4) para $c_1 = 5$:

$c_1 = 5$

1	30	0	−1000
1	35	175	$\boxed{-125}$
1	40	$\boxed{375}$	
1	$\boxed{45}$		

Ahora bien, $P_2(y) = y^3 + 450y^2 + 37\,500y - 125\,000 = 0$.

Determinemos los valores de $P_2(y)$ para ciertos valores de $y \in [0, 10]$ con ayuda del esquema de Horner:

y	1	450	37 500	−125 000
3	1	453	38 859	−8 423
4	1	454	39 316	32 264

Puesto que $P_2(3) < 0$ y $P_2(4) > 0$, entonces $y \in [3, 4]$. De aquí resulta que $c_2 = 3$.

3) Planteemos la ecuación para determinar las cifras sucesivas, en la cual se puede hallar los coeficientes utilizando el esquema de Horner:

$c_2 = 3$

1	450	37 500	−125 000
1	453	38 859	$\boxed{-8\,423}$
1	456		
1	$\boxed{459}$		

Por lo tanto, $P_3(y) = y^3 + 4590y^2 + 4\,022\,700y - 8\,423\,000 = 0$.

A la ecuación obtenida se puede aplicarle el procedimiento especial descrito anteriormente. Como resultado de la triple sustitución (4) todas las raíces, a excepción de la buscada, han aumentado 10^3 veces, aproximadamento. Entonces la raíz de la última ecuación es igual aproximadamente a $8\,423\,000/4\,022\,700 \approx 2{,}09$. Esto quiere decir que las cifras 2, 0 y 9 sirven de cifras decimales posteriores de la raíz de la ecuación inicial. Como resultado hallamos la raíz de la ecuación dada: $X = -2{,}53209$. ▲

Ejercicios

1. Separar las raíces analíticamente y precisarlas hasta 0.001 con ayuda del método de pruebas:

a) $x^3 - x + 1 = 0$; b) $x^3 + 2x - 4 = 0$; c) $x^4 + 5x - 3 = 0$; d) $2{,}2x - 2^x = 0$; e) $2^x - 2x^2 - 1 = 0$; f) $2^x - 4x = 0$.

2. Separar las raíces gráficamente y precisarlas hasta 0,001 con ayuda del método de las cuerdas:

a) $x^3 + x - 3 = 0$; b) $x^3 + 8x - 6 = 0$; c) $x^3 + 10x - 9 = 0$;

d) $x^2 - \cos \pi x = 0$; e) $x^2 - \operatorname{sen} \pi x = 0$; f) $\log x - \frac{1}{x^2} = 0$.

3. Haciendo uso del método de las tangentes, hallar con exactitud hasta 0,001 las raíces de las ecuaciones:

a) $x^3 - 6x^2 + 9x - 3 = 0$; b) $x^3 - 12x - 8 = 0$;

c) $x^3 + 4x^2 - 6 = 0$; d) $2 \log x - \frac{x}{2} + 1 = 0$;

e) $x^2 - 20 \operatorname{sen} x = 0$; e) $x - \cos x = 0$.

4. Utilizando el método combinado de las cuerdas y las tangentes determinar con exactitud hasta 0,001 las raíces de las ecuaciones:

a) $x^3 + 6x - 5 = 0$; b) $x^3 - 2x + 7 = 0$; c) $x^3 - 2x^2 + x + 1 = 0$; d) $1{,}8x^2 - \operatorname{sen} 10x = 0$; e) $\log x - \frac{7}{2x+6} = 0$; f) $2x \log x - 1 = 0$.

5. Haciendo uso del teorema de Sturm, separar las raíces de las ecuaciones y precisarlas hasta 0,001 con ayuda del método de iteraciones:

a) $x^3 + 4x - 3 = 0$; b) $x^4 - 2x - 1 = 0$;

c) $x^5 - 5x + 2 = 0$; d) $x^4 + x - 3 = 0$.

6. Utilizando el método de iteraciones, hallar las raíces de las ecuaciones con exactitud hasta 0,001:

a) $\ln x + (x + 1)^3 = 0$; b) $\sqrt{x + 1} = 1/x$; c) $x - \cos x = 0$;

d) $3x - \cos x - 1 = 0$; e) $x + \log x = 0{,}5$.

CAPÍTULO VI

Determinación de los valores propios de una matriz y de sus vectores propios

§ 6.1. Polinomio característico

Sean dadas la matriz cuadrada A y el vector columna no nulo $\mathbf{x}$

$$A=\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}, \quad x=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Multiplicando la matriz A por el vector $\mathbf{x}$, obtenemos el vector columna

$$\mathbf{y}=\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix},$$

o sea,

$$\mathbf{y} = A\mathbf{x}. \tag{1}$$

Si resulta que las coordenadas y_i ($i = 1, 2, \dots, n$) del vector $\mathbf{y}$ son proporcionales a las coordenadas correspondientes x_i del vector dado $\mathbf{x}$ con el coeficiente de proporcionalidad λ, o sea, si $y_i = \lambda x_i$ y, por consiguiente,

$$\mathbf{y} = \lambda\mathbf{x}, \tag{2}$$

el vector columna no nulo $\mathbf{x}$ se llama *vector propio* de la matriz A y el coeficiente de proporcionalidad λ, *valor propio* (o *número característico*) de la matriz A. Puesto que $\mathbf{y} = A\mathbf{x}$ e $\mathbf{y} = \lambda\mathbf{x}$, es evidente que

$$A\mathbf{x} = \lambda\mathbf{x}. \tag{3}$$

Ahora bien, si se cumple la condición (3), el vector $\mathbf{x}$ es el vector propio de la matriz A el cual corresponde al valor propio λ de esta última.

Ejemplo 1. Sea

$$A=\begin{bmatrix} 1 & 2 & 3 \\ 2 & 6 & -2 \\ 3 & 4 & -1 \end{bmatrix}, \quad \mathbf{x}=\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Entonces

$$Ax=\begin{bmatrix}1&2&3\\2&6&-2\\3&4&-1\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix}=\begin{bmatrix}6\\6\\6\end{bmatrix}=6\begin{bmatrix}1\\1\\1\end{bmatrix}.$$

Por lo tanto, el número $\lambda=6$ es el valor propio de la matriz A ya que se cumple la igualdad (3):

$$A_x=\lambda_x=6\begin{bmatrix}1\\1\\1\end{bmatrix},$$

y el vector $\mathbf{x}=\begin{bmatrix}1\\1\\1\end{bmatrix}$ es el vector propio de la matriz A el cual corresponde al valor propio $\lambda=6$.

Escribamos la relación (3) en la forma $A\mathbf{x}-\lambda\mathbf{x}=\mathbf{0}$ o bien

$$(A-\lambda E)\,\mathbf{x}=\mathbf{0}, \tag{4}$$

donde E es la matriz unidad del mismo tamaño que la matriz A y $\mathbf{0}$, el vector columna nulo. Es evidente que sin el factor E de λ la ecuación (4) no habría tenido sentido.

Puesto que

$$\lambda E=\begin{vmatrix}\lambda&0&0&\dots&0\\0&\lambda&0&\dots&0\\ \cdot&\cdot&\cdot&\cdot&\cdot\\0&0&0&\dots&\lambda\end{vmatrix},\quad \mathbf{0}=\begin{bmatrix}0\\0\\ \vdots\\0\end{bmatrix},$$

la ecuación (4) se puede escribir en la forma

$$\begin{vmatrix}a_{11}-\lambda&a_{12}&\dots&a_{1n}\\ v_{21}&a_{22}-\lambda&\dots&a_{2n}\\ \cdot&\cdot&\cdot&\cdot\\ a_{n1}&a_{n2}&\dots&a_{nn}-\lambda\end{vmatrix}\begin{bmatrix}x_1\\x_2\\ \vdots\\x_n\end{bmatrix}=\begin{bmatrix}0\\0\\ \vdots\\0\end{bmatrix}. \tag{5}$$

La relación (5) es un sistema homogéneo lineal de ecuaciones el cual tiene soluciones no nulas si y sólo si su determinante es igual a cero, o sea, cuando se cumple la condición

$$\det(A-\lambda E)=0. \tag{6}$$

La ecuación (6) se llama *ecuación característica* de la matriz A y su primer miembro, *polinomio característico* (o *determinante característico*) de la matriz A.

En la forma desarrollada la ecuación característica se escribe así:

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{vmatrix} = 0. \qquad (7)$$

Si se desarrolla el determinante en el primer miembro de la ecuación (7), se obtiene el polinomio de n-ésimo grado respecto a λ:

$$D(\lambda) = \det(A-\lambda E) =$$
$$= (-1)^n [\lambda^n - p_1\lambda^{n-1} + p_2\lambda^{n-2} - \dots + (-1)^{(n)} p_n]. \qquad (8)$$

La magnitud λ, que se determina de la ecuación (8), toma n valores $\lambda_1, \lambda_2, \dots, \lambda_n$ entre los cuales pueden ser también iguales. Para hallar los vectores propios $\mathbf{x}^{(i)}$ correspondientes a los valores propios λ_i ($i = 1, 2, \dots, n$) es necesario resolver el sistema homogéneo lineal de ecuaciones (5) para cada valor λ_i.

Ejemplo 2. Hallar los valores propios y los vectores propios de la matriz $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$.

△ 1) Escribamos el polinomio característico de la matriz A y determinemos λ. Tenemos

$$\det(A-\lambda E) = \begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = \lambda^2 - 4\lambda + 3.$$

La ecuación característica $\lambda^2 - 4\lambda + 3 = 0$ tiene dos raíces $\lambda_1 = 1$ y $\lambda_2 = 3$ que son precisamente los valores propios de la matriz A.

2) Determinemos el vector propio $\mathbf{x}^{(1)} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ correspondiente al valor $\lambda_1 = 1$. La ecuación matricial

$$(A-\lambda_1 E)\,\mathbf{x}^{(1)} = 0 \text{ o bien } \begin{bmatrix} 2-1 & 1 \\ 1 & 2-1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

es equivalente al sistema $\begin{cases} x_1 + x_2 = 0, \\ x_1 + x_2 = 0 \end{cases}$ el cual tiene un conjunto infinito de soluciones en la forma $x_1 = -x_2$. Suponiendo $x_1 = c$ (c es todo número), obtenemos $x_2 = -c$. Entonces el vector propio buscado se escribirá así: $\mathbf{x}^{(1)} = c\begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

3) Determinemos el vector propio $\mathbf{x}^{(2)} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ correspondiente al segundo valor propio $\lambda_2 = 3$. Tenemos

$$(A-\lambda_2 E)\,\mathbf{x}^{(2)} = 0 \text{ o bien } \begin{bmatrix} 2-3 & 1 \\ 1 & 2-3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Esta ecuación matricial conduce al sistema $\begin{cases} -x_1+x_2=0 \\ x_1-x_2=0 \end{cases}$ de donde $x_1 = x_2$. Suponiendo $x_1 = c$, obtenemos $x_2 = c$; por lo tanto, el segundo vector propio tiene la forma $\mathbf{x}^{(2)} = c\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. ▲

Al determinar los valores propios de las matrices y sus vectores propios se resuelve uno de dos problemas: 1) determinación de todos los valores propios y de los vectores propios de las matrices los cuales les pertenecen o bien 2) determinación de uno o de algunos valores propios y de los vectores propios que les pertenecen.

El primer problema consiste en desarrollar el determinante característico en polinomio de n-ésimo grado (o sea, en determinar los coeficientes $p_1, p_2, \ldots, p_n$), en calcular luego los valores propios $\lambda_1, \lambda_2, \ldots, \lambda_n$ y, por último, en determinar las coordenadas del vector propio $\mathbf{x}^{\top} = (x_1, x\ , \ldots, x_n)$.

El segundo problema consiste en determinar los valores propios de λ (uno o varios) con ayuda del método iterativo sin desarrollar preliminarmente el determinante característico.

Los métodos del primer problema son **exactos**, o sea, si se emplean para las matrices cuyos elementos se dan exactamente (por los números racionales) y si se realizan exactamente los cálculos (por las reglas de operaciones con las fracciones ordinarias), como resultado se obtendrá el valor exacto de los coeficientes del polinomio característico y las coordenadas de los vectores propios resultarán expresadas por las fórmulas exactas con ayuda de los valores propios.

De ordinario se logra determinar los vectores propios de la matriz, utilizando los resultados intermedios de cálculos realizados para determinar los coeficientes del polinomio característico. Desde luego, para determinar el vector propio que pertenece a uno u otro valor, este valor propio debe ya estar calculado.

Los métodos de resolución del segundo problema son **iterativos**, aquí los valores propios se obtienen como límites de ciertas sucesiones numéricas, al igual que las coordenadas de los vectores propios que les pertenecen. Puesto que estos métodos no requieren calcular los coeficientes del polinomio característico, son menos fatigosos. A continuación se consideran algunos métodos de desarrollo del determinante característico y los métodos iterativos de determinación de los valores propios de la matriz.

§ 6.2. Método de desarrollo inmediato

Tomando como ejemplo la matriz de tercer orden, consideremos cómo se determinan los coeficientes del polinomio característico por el desarrollo inmediato del determinante característico. Sea

$$A=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad \det(A-\lambda E)=\begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13} \\ a_{21} & a_{22}-\lambda & a_{23} \\ a_{31} & a_{33} & a_{33}-\lambda \end{vmatrix}$$

Calculemos el determinante por la regla de los triángulos:

$$\det(A-\lambda E)=\begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13} \\ a_{21} & a_{22}-\lambda & a_{23} \\ a_{31} & a_{32} & a_{33}-\lambda \end{vmatrix}=$$

$$=(a_{11}-\lambda)(a_{22}-\lambda)(a_{33}-\lambda)+a_{12}a_{23}a_{31}+$$
$$+a_{13}a_{32}a_{21}-a_{13}a_{31}(a_{22}-\lambda)-a_{12}a_{21}(a_{33}-\lambda)-$$
$$-a_{23}a_{32}(a_{11}-\lambda)=-\lambda^3+\lambda^2(a_{11}+a_{22}+a_{33})-$$
$$-\lambda[(a_{11}a_{22}-a_{12}a_{21})+(a_{22}a_{33}-a_{23}a_{32})+(a_{11}a_{33}-a_{13}a_{31})]+$$
$$+(a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{32}a_{21}-a_{13}a_{22}a_{31}-a_{12}a_{21}a_{33}-a_{23}a_{32}a_{11})=$$
$$=(-1)^3\cdot\left[\lambda^3-\lambda^2(a_{11}+a_{22}+a_{33})+\lambda\left(\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{vmatrix}+\right.\right.$$
$$\left.\left.+\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33}\end{vmatrix}+\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33}\end{vmatrix}\right)-\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33}\end{vmatrix}\right),$$

o bien

$$\det(A-\lambda E)=(-1)^3(\lambda^3-p_1\lambda^2+p_2\lambda-p_3)=0.$$

Aquí el coeficiente p_1 es la suma de los elementos diagonales de la matriz A; se llama *traza* de la matriz y se designa $\mathrm{Sp}\,A$:

$$p_1=\mathrm{Sp}\,A=a_{11}+a_{22}+a_{33};$$

el coeficiente p_2 es la suma de todos los menores diagonales de segundo orden de la matriz A:

$$p_2=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{vmatrix}+\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33}\end{vmatrix}+\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33}\end{vmatrix}$$

(recuérdese que se llaman menores diagonales de segundo, tercero, ..., n-ésimo orden los menores cuyos elementos de las diagonales principales son elementos de la diagonal principal del determinante $\det A$); el coeficiente

$$p_3=\det A=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33}\end{vmatrix}.$$

En general, si se necesita desarrollar el determinante $\det(A-\lambda E)$ en polinomio de n-ésimo grado:

$$D(\lambda)=(-1)^n[\lambda^n-p_1\lambda^{n-1}+p_2\lambda^{n-2}-\ldots+(-1)^n p_n],$$

los coeficientes $p_1, p_2, \ldots, p_n$ se calculan por las fórmulas siguientes:

$p_1=\sum\limits_{i=1}^{n} a_{ii}=\mathrm{Sp}\,A$ *es la suma de todos los elementos diagonales de la matriz A;*

$p_2 = \sum\limits_{\alpha<\beta} \begin{vmatrix} a_{\alpha\alpha} & a_{\alpha\beta} \\ a_{\beta\alpha} & a_{\beta\beta} \end{vmatrix}$ *es la suma de todos los menores diagonales de segundo orden de la matriz A;*

$p_3 = \sum\limits_{\alpha<\beta<\gamma} \begin{vmatrix} a_{\alpha\alpha} & a_{\alpha\beta} & a_{\alpha\gamma} \\ a_{\beta\alpha} & a_{\beta\beta} & a_{\beta\gamma} \\ a_{\gamma\alpha} & a_{\gamma\beta} & a_{\gamma\gamma} \end{vmatrix}$ *es la suma de todos los menores diagonales de tercer orden de la matriz A;*

. . . .

$p_n = \det A$ *es el determinante de la matriz A.*

La cantidad de menores diagonales de k-ésimo orden de la matriz es igual a

$$C_n^k = \frac{n(n-1)(n-2)\ldots(n-k+1)}{k!} \quad (k=1, 2, \ldots, n)$$

El método de desarrollo inmediato cuesta mucho trabajo y se emplea al determinar los polinomios característicos para las matrices de pequeños órdenes.

Ejemplo 1. Haciendo uso del método de desarrollo inmediato, hallar el polinomio característico de la matriz

$$A = \begin{bmatrix} -4 & -3 & 1 & 1 \\ 2 & 0 & 4 & -1 \\ 1 & 1 & 2 & -2 \\ 1 & 1 & -1 & -1 \end{bmatrix}.$$

1) Determinamos

$$p_1 = \mathrm{Sp}\, A = a_{11} + a_{22} + a_{33} + a_{44} = -4 + 0 + 2 - 1 = -3.$$

2) Tenemos $p_2 = \sum\limits_{\alpha<\beta} \begin{vmatrix} a_{\alpha\alpha} & a_{\alpha\beta} \\ a_{\beta\alpha} & a_{\beta\beta} \end{vmatrix}$. La cantidad de los menores diagonales de segundo orden en la matriz de cuarto orden es igual a $C_4^2 = \frac{43}{1\cdot 2} = 6$. Escribiendo todos estos menores y sumándolos, obtenemos

$$p_2 = \underbrace{\begin{vmatrix} -4 & -3 \\ 2 & 0 \end{vmatrix}}_{\alpha=1;\ \beta=2} + \underbrace{\begin{vmatrix} -4 & 1 \\ 1 & 2 \end{vmatrix}}_{\alpha=1;\ \beta=2} + \underbrace{\begin{vmatrix} -4 & 1 \\ 1 & -1 \end{vmatrix}}_{\alpha=1;\ \beta=4} + \underbrace{\begin{vmatrix} 0 & 4 \\ 1 & 2 \end{vmatrix}}_{\alpha=2;\ \beta=3} +$$

$$+ \underbrace{\begin{vmatrix} 0 & -1 \\ 1 & -1 \end{vmatrix}}_{\alpha=2;\ \beta=4} + \underbrace{\begin{vmatrix} 2 & -2 \\ -1 & -1 \end{vmatrix}}_{\alpha=3;\ \beta=4} = -7.$$

3) Tenemos $p_3 = \sum\limits_{\alpha<\beta<\gamma} \begin{vmatrix} a_{\alpha\alpha} & a_{\alpha\beta} & a_{\alpha\gamma} \\ a_{\beta\alpha} & a_{\beta\beta} & a_{\beta\gamma} \\ a_{\gamma\alpha} & a_{\gamma\beta} & a_{\gamma\gamma} \end{vmatrix}$. La cantidad de menores diagonales de tercer orden en la matriz de cuarto orden es igual a $C_4^3 = \frac{4\cdot3\cdot2}{1\cdot2\cdot3} = 4$. Por consiguiente,

$$p_3 = \underbrace{\begin{vmatrix} -4 & -3 & 1 \\ 2 & 0 & 4 \\ 1 & 1 & 2 \end{vmatrix}}_{\alpha=1;\ \beta=2;\ \gamma=3} + \underbrace{\begin{vmatrix} -4 & -3 & 1 \\ 2 & 0 & -1 \\ 1 & 1 & -1 \end{vmatrix}}_{\alpha=1;\ \beta=2;\ \gamma=4} + \underbrace{\begin{vmatrix} -4 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & -1 & -1 \end{vmatrix}}_{\alpha=1;\ \beta=3;\ \gamma=4} +$$

$$+ \underbrace{\begin{vmatrix} 0 & 4 & -1 \\ 1 & 2 & -2 \\ 1 & -1 & -1 \end{vmatrix}}_{\alpha=2;\ \beta=3;\ \gamma=4} = 24.$$

4) por último, hallamos

$$p_4 = \det A = \begin{vmatrix} -4 & -3 & 1 & 1 \\ 2 & 0 & 4 & -1 \\ 1 & 1 & 2 & -2 \\ 1 & 1 & -1 & -1 \end{vmatrix} = -15.$$

5) Ahora bien, finalmente obtenemos

$$D(\lambda) = \lambda^4 - p_1\lambda^3 + p_2\lambda^2 - p_3\lambda + p_4 =$$
$$= \lambda^4 + 3\lambda^3 - 7\lambda^2 - 24\lambda - 15. \ \blacktriangle$$

Ejemplo 2. **Desarrollar el polinomio característico de la matriz**

$$A = \begin{bmatrix} 3 & 1 & 0 \\ -4 & -1 & 0 \\ 4 & -8 & -2 \end{bmatrix}$$

y determinar cualquier valor propio y el vector propio correspondiente.

△ 1) Tenemos

$$p_1 = \operatorname{Sp} A = 3 - 1 - 2 = 0;$$

$$p_2 = \sum_{\alpha<\beta} \begin{vmatrix} a_{\alpha\alpha} & a_{\alpha\beta} \\ a_{\beta\alpha} & a_{\beta\beta} \end{vmatrix} = \underbrace{\begin{vmatrix} 3 & 1 \\ -4 & -1 \end{vmatrix}}_{\alpha=1;\ \beta=2} + \underbrace{\begin{vmatrix} 3 & 0 \\ 4 & -2 \end{vmatrix}}_{\alpha=1;\ \beta=3} + \underbrace{\begin{vmatrix} -1 & 0 \\ -8 & -2 \end{vmatrix}}_{\alpha=2;\ \beta=3} =$$
$$= 1 - 6 + 2 = -3;$$

$$p_3 = \det A = \begin{vmatrix} 3 & 1 & 0 \\ -4 & -1 & 0 \\ 4 & -8 & -2 \end{vmatrix} = -2.$$

Por lo tanto, $D(\lambda) = (-1)^3 (\lambda^3 - 3\lambda + 2)$; $\lambda^3 - 3\lambda + 2 = 0$; uno de los valores propios de la matriz A es $\lambda = 1$.

2) Hallamos el vector propio $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ correspondiente a $\lambda = 1$. La ecuación matricial

$$(A - \lambda E)\,\mathbf{x} = 0 \quad \text{o bien} \quad \begin{bmatrix} 3 - 1 & 1 & 0 \\ -4 & -1 - 1 & 0 \\ 4 & -8 & -2 - 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

es equivalente al sistema

$$\begin{cases} 2x_1 + x_2 = 0, \\ -4x_1 - 2x_2 = 0, \\ 4x_1 - 8x_2 - 3x_3 = 0. \end{cases}$$

Aquí las ecuaciones primera y segunda son proporcionales, por eso, suprimiendo la segunda ecuación, obtenemos el sistema

$$\begin{cases} 2x_1 + x_2 = 0, \\ 4x_1 - 8x_2 - 3x_3 = 0. \end{cases}$$

El menor $d = \begin{vmatrix} 2 & 1 \\ 4 & -8 \end{vmatrix} = -20$ es básico; x_1 y x_2 son las incógnitas básicas; x_3 es la incógnita independiente. Entonces con ayuda de las fórmulas de Cramer hallamos

$$x_1 = \frac{d_1}{d} = \frac{\begin{vmatrix} 0 & 1 \\ 3x_3 & -8 \end{vmatrix}}{-20} = \frac{3x_3}{20}.$$

$$x_2 = \frac{d_2}{d} = \frac{\begin{vmatrix} 2 & 0 \\ 4 & 3x_3 \end{vmatrix}}{-20} = \frac{6x_3}{-20}.$$

Suponiendo $x_3 = 20$, tenemos $x_1 = 3$; $x_2 = -6$. Así pues $\mathbf{x}^{(1)} = c \begin{bmatrix} 3 \\ -6 \\ 20 \end{bmatrix}$ es el vector propio buscado. ▲

§ 6.3. Método de Krylov para desarrollar el determinante característico

El método de Krylov se funda en propiedad de la matriz cuadrada de anular su polinomio característico.

Según la **identidad de Hamilton—Cayley** toda la *matriz cuadrada es raíz de su polinomio característico y, por lo tanto, lo anula.*

Sea

$$D(\lambda) = \det(A - \lambda E) =$$
$$= (-1)^n (\lambda^n + p_1\lambda^{n-1} + p_2\lambda^{n-2} + \ldots + p_n) \qquad (1)$$

el polinomio característico de la matriz A. Reemplazando en la igualdad (1) la magnitud λ por $A = [a_{ij}]$ (donde $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, n$), obtenemos

$$A^n + p_1A^{n-1} + p_2A^{n-2} + \ldots + p_nE = 0. \qquad (2)$$

Tomemos un vector no nulo arbitrario

$$\mathbf{y}^{(0)} = \begin{bmatrix} y_1^{(0)} \\ y_2^{(0)} \\ \vdots \\ y_n^{(0)} \end{bmatrix} \qquad (3)$$

y multipliquemos ambos miembros de la igualdad (2) a la derecha por $\mathbf{y}^{(0)}$:

$$A^n\mathbf{y}^{(0)} + p_1A^{n-1}\mathbf{y}^{(0)} + p_2A^{n-2}\mathbf{y}^{(0)} + \ldots + p_n\mathbf{y}^{(0)} = 0. \qquad (4)$$

Pongamos ahora

$$A\mathbf{y}^{(k-1)} = \mathbf{y}^{(k)} \quad (k = 1, 2, \ldots, n), \qquad (5)$$

o sea,

$$\begin{aligned} \mathbf{y}^{(1)} &= A\mathbf{y}^{(0)}, \\ \mathbf{y}^{(2)} &= A\mathbf{y}^{(1)} = A^2\mathbf{y}^{(0)}, \\ &\cdots\cdots\cdots \\ \mathbf{y}^{(n)} &= A\mathbf{y}^{n-1} = A^n\mathbf{y}^{(0)}. \end{aligned}$$

Entonces la igualdad (4) se toma la forma siguiente:

$$\mathbf{y}^{(n)} + p_1\mathbf{y}^{(n-1)} + p_2\mathbf{y}^{(n-2)} + \ldots + p_n\mathbf{y}^{(0)} = 0 \qquad (6)$$

o bien

$$p_1\mathbf{y}^{(n-1)} + p_2\mathbf{y}^{(n-2)} + \ldots + p_n\mathbf{y}^{(0)} = -\mathbf{y}^{(n)},$$

o bien

$$p_1\begin{bmatrix} y_1^{(n-1)} \\ y_2^{(n-1)} \\ \vdots \\ y_n^{(n-1)} \end{bmatrix} + p_2\begin{bmatrix} y_1^{(n-2)} \\ y_2^{(n-2)} \\ \vdots \\ y_n^{(n-2)} \end{bmatrix} + \ldots + p_n\begin{bmatrix} y_1^{(0)} \\ y_2^{(0)} \\ \vdots \\ y_n^{(0)} \end{bmatrix} = -\begin{bmatrix} y_1^{(n)} \\ y_2^{(n)} \\ \vdots \\ y_n^{(n)} \end{bmatrix},$$

o sea,

$$\left\{ \begin{array}{l} p_1y_1^{(n-1)} + p_2y_1^{(n-2)} + \ldots + p_ny_1^{(0)} = -y_1^{(n)}, \\ p_1y_2^{(n-1)} + p_2y_2^{(n-2)} + \ldots + p_ny_2^{(0)} = -y_2^{(n)}, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ p_1y_n^{(n-1)} + p_2y_n^{(n-2)} + \ldots + p_ny_n^{(0)} = -y_n^{(n)} \end{array} \right. \qquad (7)$$

o bien, por último, en la forma matricial

$$\begin{bmatrix} y_1^{(n-1)} & y_1^{(n-2)} & \dots & y_1^{(0)} \\ y_2^{(n-1)} & y_2^{(n-2)} & \dots & y_2^{(0)} \\ \dots & \dots & \dots & \dots \\ y_n^{(n-1)} & y_n^{(n-2)} & \dots & y_n^{(0)} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix} = - \begin{bmatrix} y_1^{(n)} \\ y_2^{(n)} \\ \vdots \\ y_n^{(n)} \end{bmatrix}. \tag{8}$$

Los vectores $\mathbf{y}_i^{(1)}, \mathbf{y}_i^{(2)}, \dots, \mathbf{y}_i^{(n)}$ se calculan por las fórmulas

$$\begin{gathered} \mathbf{y}_i^{(1)} = \sum_{j=1}^{n} a_{ij}\mathbf{y}_j^{(0)} = A\mathbf{y}^{(0)}, \\ \mathbf{y}_i^{(2)} = \sum_{j=1}^{n} a_{ij}\mathbf{y}_j^{(1)} = A\mathbf{y}^{(1)}, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ \mathbf{y}_i^{(n)} = \sum_{j=1}^{n} a_{ij}\mathbf{y}_j^{(n-1)} = A\mathbf{y}^{(n-1)} \quad (i = 1, 2, \dots, n), \end{gathered} \tag{9}$$

con la particularidad de que las coordenadas del vector inicial (3) se toman arbitrariamente. Si el sistema lineal (7) tiene la única solución, sus raíces $p_1, p_2, \dots, p_n$ son coeficientes del polinomio característico (1). Esta solución puede ser hallada por el método de Gauss.

Ejemplo 1. Haciendo uso del método de Krylov, desarrollar el determinante característico de la matriz

$$A = \begin{bmatrix} -4 & -3 & 1 & 1 \\ 2 & 0 & 4 & -1 \\ 1 & 1 & 2 & -2 \\ 1 & 1 & -1 & -1 \end{bmatrix}.$$

△ 1) Elegimos el vector inicial $\mathbf{y}^{(0)} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$.

2) Con ayuda de las fórmulas (9) determinamos las coordenadas de los vectores $\mathbf{y}^{(k)} = A\,\mathbf{y}^{(k-1)}$ $(k = 1, 2, 3, 4)$:

$$\mathbf{y}^{(1)} = A\mathbf{y}^{(0)} = \begin{bmatrix} -4 & -3 & 1 & 1 \\ 2 & 0 & 4 & -1 \\ 1 & 1 & 2 & -2 \\ 1 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -4 \\ 2 \\ 1 \\ 1 \end{bmatrix},$$

$$\mathbf{y}^{(2)} = A\mathbf{y}^{(1)} = \begin{bmatrix} 4 & -3 & 1 & 1 \\ 2 & 0 & 4 & -1 \\ 1 & 1 & 2 & -2 \\ 1 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} -4 \\ 2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 12 \\ -5 \\ -2 \\ -4 \end{bmatrix},$$

$$
\mathbf{y}^{(3)} = A\mathbf{y}^{(2)} = \begin{bmatrix} -4 & -3 & 1 & 1 \\ 2 & 0 & 4 & -1 \\ 1 & 1 & 2 & -2 \\ 1 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 12 \\ -5 \\ -2 \\ -4 \end{bmatrix} = \begin{bmatrix} -39 \\ 20 \\ 11 \\ 13 \end{bmatrix},
$$

$$
\mathbf{y}^{(4)} = A\mathbf{y}^{(3)} = \begin{bmatrix} -4 & -3 & 1 & 1 \\ 2 & 0 & 4 & -1 \\ 1 & 1 & 2 & -2 \\ 1 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} -39 \\ 20 \\ 11 \\ 13 \end{bmatrix} = \begin{bmatrix} 120 \\ -47 \\ -23 \\ -43 \end{bmatrix}.
$$

3) Planteamos la ecuación matricial

$$
\begin{bmatrix} -39 & 12 & -4 & 1 \\ 20 & -5 & 2 & 0 \\ 11 & -2 & 1 & 0 \\ 13 & -4 & 1 & 0 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{bmatrix} = - \begin{bmatrix} 120 \\ -47 \\ -23 \\ -43 \end{bmatrix}.
$$

Tabla 6.1

p_1	p_2	p_3	p_4	Términos independientes	Σ_1	Σ_2
$\boxed{-39}$	12	−4	1	−120	−150	
20	−5	2	0	47	64	
11	−2	1	0	23	33	
13	−4	1	0	43	53	
1	−4/13	4/39	−1/39	40/13	50/13	
	$\boxed{15/13}$	−2/39	20/39	−189/13	−168/13	−168/13
	18/13	−5/39	11/39	−141/13	−121/13	−121/13
	0	−1/3	−1/3	3	3	3
	1	−2/45	4/9	−53/5	−56/5	−56/5
		$\boxed{-1/15}$	−1/3	33/5	31/5	31/5
		−1/3	1/3	3	3	3
		1	5	−99	−93	−93
			$\boxed{2}$	−30	−28	−28
			1	−15	−14	−14
			1	$p_4 = -15$	$\bar{p}_4 = -14$	−14
		1		$p_3 = -24$	$\bar{p}_3 = -23$	−23
	1			$p_2 = -7$	$\bar{p}_2 = -6$	−6
1				$p_1 = 3$	$\bar{p}_1 = 4$	4

Escribimos el sistema que tiene la forma (7):

$$\begin{cases} -39p_1+12p_2-4p_3+p_4=-120, \\ 20p_1-5p_2+2p_3 \qquad =47, \\ 11p_1-2p_2+p_3 \qquad =23, \\ 13p_1-4p_2+p_3 \qquad =43. \end{cases}$$

Resolvemos este sistema según el esquema de Gauss (véase la tabla 6.1).

Ahora bien,

$$D(\lambda) = \det(A-\lambda E) = \lambda^4+p_1\lambda^3+p_2\lambda^2+p_3\lambda+p_4 = \\ = \lambda^4+3\lambda^3-7\lambda^2-24\lambda-15. \blacktriangle$$

En cambio, si el sistema lineal obtenido (7) no tiene la solución única, es necesario cambiar el vector inicial.

Ejemplo 2. Haciendo uso del método de Krylov, desarrollar el determinante característico de la matriz

$$A=\begin{bmatrix} 1 & -1 & -1 & 2 \\ 2 & 3 & 0 & -4 \\ 1 & 1 & -2 & -2 \\ 1 & 1 & 0 & -1 \end{bmatrix}.$$

△ 1) Tomemos en calidad de vector inicial $\mathbf{y}^{(0)}=\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$.

Entonces obtenemos

$$\mathbf{y}^{(1)}=A\mathbf{y}^{(0)}=\begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix}; \quad \mathbf{y}^{(2)}=A\mathbf{y}^{(1)}=\begin{bmatrix} 0 \\ 4 \\ -1 \\ 2 \end{bmatrix};$$

$$\mathbf{y}^{(3)}=A\mathbf{y}^{(2)}=\begin{bmatrix} 1 \\ 4 \\ 2 \\ 2 \end{bmatrix}; \quad \mathbf{y}^{(4)}=A\mathbf{y}^{(3)}=\begin{bmatrix} -1 \\ 6 \\ -3 \\ 3 \end{bmatrix}.$$

2) Planteamos la ecuación matricial

$$\begin{bmatrix} 1 & 0 & 1 & 1 \\ 4 & 4 & 2 & 0 \\ 2 & -1 & 1 & 0 \\ 2 & 2 & 1 & 0 \end{bmatrix}\begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{bmatrix}=-\begin{bmatrix} 1 \\ 6 \\ -3 \\ 3 \end{bmatrix},$$

de donde obtenemos el sistema de ecuaciones

$$\begin{cases} p_1+p_3+p_4=1, \\ 4p_1+4p_2+2p_3=-6, \\ 2p_1-p_2+p_3=3, \\ 2p_1+2p_2+p_3=-3. \end{cases}$$

Resolvemos este sistema con ayuda del esquema de única división

Tabla 6.2

p_1	p_2	p_3	p_4	Términos independientes	Σ_1	Σ_2
$\boxed{1}$	0	1	1	1	4	
4	4	2	0	−6	4	
2	−1	1	0	3	5	
2	2	1	0	−3	2	
1	0	1	1	1	4	
	$\boxed{4}$	−2	−4	−10	−12	−12
	−1	−1	−2	1	−3	−3
	2	−1	−2	−5	−6	−6
	1	−0,5	−1	−3	−3,5	−3,5
		$\boxed{-1,5}$	−3	−2	−6,5	−6,5
		0	0	1	1	1
		1	2	1,333	4,333	4,333
			$\boxed{0}$	1	1	1

Puesto que el elemento guía es igual a cero, no es posible continuar los cálculos con ayuda de este esquema.

3) Para obtener la única solución cambiamos el vector inicial; suponiendo $\mathbf{y}^{(0)}=\begin{bmatrix}0\\0\\0\\1\end{bmatrix}$, hallamos

$$\mathbf{y}^{(1)}=A\mathbf{y}^{(0)}=\begin{bmatrix}2\\-4\\-2\\-1\end{bmatrix};\quad \mathbf{y}^{(2)}=A\mathbf{y}^{(1)}=\begin{bmatrix}6\\-4\\4\\-1\end{bmatrix};$$

$$\mathbf{y}^{(3)}=A\mathbf{y}^{(2)}=\begin{bmatrix}4\\4\\-4\\3\end{bmatrix};\quad \mathbf{y}^{(4)}=A\mathbf{y}^{(3)}=\begin{bmatrix}10\\8\\10\\5\end{bmatrix}.$$

Tabla 6.3

p_1	p_2	p_3	p_4	Términos independientes	Σ_1	Σ_2
$\boxed{2}$	3	1	0	−5	1	
1	−1	−1	0	2	−3	
−2	2	−1	0	−5	−6	
3	−1	−1	1	−5	−3	
1	1,5	0,5	0	−2,5	0,5	
	$\boxed{-2,5}$	−1,5	0	0,5	−3,5	−3,5
	5	0	0	−10	−5	−5
	−5,5	−2,5	1	2,5	−4,5	−4,5
	1	0,6	0	−0,2	1,4	1,4
		$\boxed{-3}$	0	−9	−12	−12
		0,8	1	1,4	3,2	3,2
		1	0	3	4	4
			$\boxed{1}$	−1	0	0
			1	$p_4=-1$	$\bar{p}_4=0$	0
		1		$p_3=3$	$\bar{p}_3=4$	4
	1			$p_2=-2$	$\bar{p}_2=-1$	−1
1				$p_1=-1$	$\bar{p}_1=0$	0

La ecuación matricial

$$\begin{bmatrix} 4 & 6 & 2 & 0 \\ 4 & -4 & -4 & 0 \\ -4 & 4 & -2 & 0 \\ 3 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{bmatrix} = - \begin{bmatrix} 10 \\ 8 \\ 10 \\ 5 \end{bmatrix}$$

conduce al sistema

$$\left\{ \begin{aligned} 4p_1+6p_2+2p_3&=-10, \\ 4p_1-4p_2-4p_3&=-8, \\ -4p_1+4p_2-2p_3&=-10, \\ 3p_1-p_2-p_3+p_4&=-5, \end{aligned} \right. \quad \text{o bien} \quad \left\{ \begin{aligned} 2p_1+3p_2+p_3&=-5, \\ p_1-p_2-p_3&=-2, \\ -2p_1+2p_2-p_3&=-5, \\ 3p_1-p_2-p_3+p_4&=-5, \end{aligned} \right.$$

el cual se resuelve con ayuda del esquema de única división (véase la tabla 6.3).

Por lo tanto,

$$D(\lambda) = \lambda^4 - \lambda^3 - 2\lambda^2 + 3\lambda - 1. \ \blacktriangle$$

§ 6.4. Cálculo de los vectores propios con ayuda del método de Krylov

Si se conocen los coeficientes $p_1, p_2, \ldots, p_n$ y las raíces $\lambda_1, \lambda_2, \ldots, \lambda_4$ del polinomio característico, el método de Krylov ofrece la posibilidad de determinar los vectores propios respectivos valiéndose de la fórmula siguiente:

$$\mathbf{x}^{(i)} = \mathbf{y}^{(n-1)} + q_{1i}\mathbf{y}^{(i-2)} + \ldots + q_{-1,i}\,\mathbf{y}^{(0)} \quad (i = 1, 2, \ldots, n). \tag{1}$$

Aquí $\mathbf{y}^{(n-1)}, \mathbf{y}^{(n-2)}, \ldots, \mathbf{y}^{(0)}$ son los vectores utilizados al hallar los coeficientes $p_1, p_2, \ldots, p_n$ con ayuda del método de Krylov y los coeficientes q_{ji} $(j = 1, 2, \ldots, n-1;\ i = 1, 2, \ldots, n)$ se determinan por el esquema de Horner:

$$q_{0i} = 1; \quad q_{ji} = \lambda_i q_{j-1,\,i} + p_j. \tag{2}$$

Ejemplo. Haciendo uso del método de Krylov, hallar los vectores propios de la matriz

$$A = \begin{bmatrix} 1 & -1 & -1 & 2 \\ 2 & 3 & 0 & -4 \\ 1 & 1 & -2 & -2 \\ 1 & 1 & 0 & -1 \end{bmatrix}.$$

△ El polinomio característico de la matriz A está conocido:

$$\det(A - \lambda E) = \lambda^4 - \lambda^3 - 2\lambda^2 + 3\lambda - 1$$

(véase el ejemplo 2 en el § 6.3), y los valores propios son tales: $\lambda_1 = \lambda_2 = 1$; $\lambda_3 = 0{,}618$; $\lambda_4 = -1{,}618$. Para determinar los vectores propios utilicemos la fórmula (1)

$$\mathbf{x}^{(i)} = \mathbf{y}^{(3)} + q_{1i}\mathbf{y}^{(2)} + q_{2i}\mathbf{y}^{(1)} + q_{3i}\mathbf{y}^{(0)}.$$

Aquí $q_{0i} = 1$ y los coeficientes q_{ji} $(j = 1, 2, 3;\ i = 1, 2, 3, 4)$ se calculan con ayuda del esquema de Horner (véase la tabla 6.4).

Tabla 6.4

λ_i	$p_0 = 1$	$p_1 = 1$	$p_2 = -2$	$p_3 = 3$
$\lambda_1 = 1$	$q_{01} = 1$	$q_{11} = 0$	$q_{21} = -2$	$q_{31} = 1$
$\lambda_2 = 1$	$q_{02} = 1$	$q_{12} = 0$	$q_{22} = -2$	$q_{32} = 1$
$\lambda_3 = 0{,}618$	$q_{03} = 1$	$q_{13} = -0{,}382$	$q_{23} = -2{,}236$	$q_{33} = 1{,}618$
$\lambda_4 = -1{,}618$	$q_{04} = 1$	$q_{14} = -2{,}618$	$q_{24} = 2{,}236$	$q_{34} = -0{,}618$

Utilizando las expresiones para los vectores $\mathbf{y}^{(0)}$, $\mathbf{y}^{(1)}$, $\mathbf{y}^{(2)}$, $\mathbf{y}^{(3)}$ hallados en el ejemplo 2 del § 6.3, obtenemos

$$\mathbf{x}^{(1)}=\mathbf{x}^{(2)}=\begin{bmatrix}4\\4\\-4\\3\end{bmatrix}+0\begin{bmatrix}6\\-4\\4\\-1\end{bmatrix}-2\begin{bmatrix}2\\-4\\-2\\-1\end{bmatrix}+1\begin{bmatrix}0\\0\\0\\1\end{bmatrix}=$$

$$=\begin{bmatrix}0\\12\\0\\6\end{bmatrix}=\begin{bmatrix}0\\1\\0\\0,5\end{bmatrix};$$

$$\mathbf{x}^{(3)}=\begin{bmatrix}4\\4\\-4\\3\end{bmatrix}-0,382\begin{bmatrix}6\\-4\\4\\-1\end{bmatrix}-2,236\begin{bmatrix}2\\-4\\-2\\-1\end{bmatrix}+$$

$$+1,618\begin{bmatrix}0\\0\\0\\1\end{bmatrix}=\begin{bmatrix}-2,764\\14,472\\-1,056\\7,236\end{bmatrix}=\begin{bmatrix}0,19\\1\\-0,07\\0,50\end{bmatrix};$$

$$\mathbf{x}^{(4)}=\begin{bmatrix}4\\4\\-4\\3\end{bmatrix}-2,618\begin{bmatrix}6\\-4\\4\\-1\end{bmatrix}+2,236\begin{bmatrix}2\\-4\\-2\\-1\end{bmatrix}-$$

$$-0,618\begin{bmatrix}0\\0\\0\\1\end{bmatrix}=\begin{bmatrix}-7,236\\5,528\\-18,944\\2,764\end{bmatrix}=\begin{bmatrix}-0,38\\0,29\\1\\0,15\end{bmatrix}. \ \blacktriangle$$

§ 6.5. Método de Le Verrier—Faddéev

Este método fue propuesto por Le Verrier y luego simplificado por el matemático soviético Faddéev. El **método de Le Verrier** se funda en las fórmulas de Newton para las sumas de los grados de las raíces de una ecuación algebraica y consiste en lo siguiente. Sea

$$\det(A-\lambda E)=\lambda^n+p_1\lambda^{n-1}+\ldots+p_n \tag{1}$$

el polinomio característico de la matriz $A=[a_{ij}]$ (donde $i=1, 2, \ldots, n$; $j=1, 2, \ldots, n$) y $\lambda_1, \lambda_2, \ldots, \lambda_n$ es el conjunto completo de las raíces del polinomio (1). Consideremos las sumas

$$S_k=\lambda_1^k+\lambda_2^k+\ldots+\lambda_n^k \ (k=1, 2, \ldots, n),$$

o sea,

$$S_1=\lambda_1+\lambda_2+\ldots+\lambda_n=\operatorname{Sp}A,$$
$$S_2=\lambda_1^2+\lambda_2^2+\ldots+\lambda_n^2=\operatorname{Sp}A^2,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$S_n=\lambda_1^n+\lambda_2^n+\ldots+\lambda_n^n=\operatorname{Sp}A^n$$

(cada suma S_k es la traza de la matriz A^k). Entonces para $k\leqslant n$ son válidas las *fórmulas de Newton*

$$S_k+p_1S_{k-1}+\ldots+p_{k-1}S_1=-kp_k,$$

de donde obtenemos

$$\begin{aligned}&\text{para } k=1\quad p_1=-S_1,\\ &\text{para } k=2\quad p_2=-\frac{1}{2}(S_2+p_1S_1),\\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ &\text{para } k=n\quad p_n=-\frac{1}{n}(S_n+p_1S_{n-1}+p_2S_{n-2}+\ldots+p_{n-1}S_1).\end{aligned}\tag{2}$$

Por consiguiente, los coeficientes del polinomio característico $p_1, p_2, \ldots, p_n$ pueden ser determinados fácilmente si están conocidas las sumas $S_1, S_2, \ldots, S_n$.

Ahora bien, el esquema de desarrollo del determinante característico por el método de Le Verrier consiste en lo siguiente:

1) se calculan los grados de $A^k=A^{k-1}A$ $(k=1, 2, \ldots, n)$;

2) se determinan S_k, o sea, las sumas de los elementos de las diagonales principales de las matrices A^k;

3) Haciendo uso de las fórmulas (2), se hallan los coeficientes p_i $(i=1, 2, \ldots, n)$.

El método modificado de Le Verrier, propuesto, por Faddéev, consiste en calcular la sucesión de las matrices $A_1, A_2, \ldots, A_n$ según el esquema siguiente:

$$\begin{aligned}&A_1=A;\quad \operatorname{Sp}A_1=q_1;\quad B_1=A_1-q_1E;\\ &A_1=AB_1;\quad \frac{\operatorname{Sp}A_2}{2}=q_2;\quad B_2=A_2-q_2E;\\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ &A_{n-1}=AB_{n-2};\quad \frac{\operatorname{Sp}A_{n-1}}{n-1}=q_{n-1};\quad B_{n-1}=A_{n-1}-q_{n-1}E;\end{aligned}\tag{3}$$

$$A_n=AB_{n-1};\quad \frac{\operatorname{Sp}A_n}{n}=q_n;\quad B_n=A_n-q_nE\ (B_n \text{ es la matriz nula});$$

$$q_1=-p_1;\ q_2=-p_2;\ \ldots;\ q_{n-1}=-p_{n-1},\ q_n=-p_n.$$

Ejemplo 1. Utilizando el método de Le Verrier—Faddéev, desarrollar el determinante característico de la matriz

$$A=\begin{bmatrix}1 & -1 & -1 & 2\\ 2 & 3 & 0 & -4\\ 1 & 1 & -2 & -2\\ 1 & 1 & 0 & -1\end{bmatrix}.$$

△ Sucesivamente hallamos

$$1)\ A_1=A=\begin{bmatrix}1 & -1 & -1 & 2\\ 2 & 3 & 0 & -4\\ 1 & 1 & -2 & -2\\ 1 & 1 & 0 & -1\end{bmatrix};\quad q_1=\operatorname{Sp}A_1=1+3-2-1=1;$$

$$B_1=A_1-q_1E=\begin{bmatrix}0 & -1 & -1 & 2\\ 2 & 2 & 0 & -4\\ 1 & 1 & -3 & -2\\ 1 & 1 & 0 & -2\end{bmatrix};$$

$$2)\ A_2=AB_1=\begin{bmatrix}1 & -1 & 1 & 2\\ 2 & 3 & 0 & -4\\ 1 & 1 & -2 & -2\\ 1 & 1 & 0 & -1\end{bmatrix}\begin{bmatrix}0 & -1 & -1 & 2\\ 2 & 2 & 0 & -4\\ 1 & 1 & 3 & -2\\ 1 & 1 & 0 & -2\end{bmatrix}=$$

$$=\begin{bmatrix}-1 & -2 & 2 & 4\\ 2 & 0 & -2 & 0\\ -2 & -3 & 5 & 6\\ 1 & 0 & -1 & 0\end{bmatrix};\quad q_2=\frac{\operatorname{Sp}A_2}{2}=\frac{-1+0+5+0}{2}=2;$$

$$B_2=A_2-q_2E=\begin{bmatrix}-3 & -2 & 2 & 4\\ 2 & -2 & -2 & 0\\ -2 & -3 & 3 & 6\\ 1 & 0 & -1 & -2\end{bmatrix}.$$

$$3)\ A_3=AB_2=\begin{bmatrix}1 & -1 & -1 & 2\\ 2 & 3 & 0 & -4\\ 1 & 1 & -2 & -2\\ 1 & 1 & 0 & -1\end{bmatrix}\begin{bmatrix}-3 & -2 & 2 & 4\\ 2 & -2 & -2 & 0\\ -2 & -3 & 3 & 6\\ 1 & 0 & -1 & -2\end{bmatrix}=$$

$$=\begin{bmatrix}-1 & 3 & -1 & -6\\ -4 & -10 & 2 & 16\\ 1 & 2 & -4 & -4\\ -2 & -4 & 1 & 6\end{bmatrix};\quad q_3=\frac{\operatorname{Sp}A_3}{3}=\frac{-1-10-4+6}{3}=-3;$$

$$B_3=A_3-q_3E=\begin{bmatrix}2 & 3 & -1 & -6\\ -4 & -7 & 2 & 16\\ 1 & 2 & -1 & -4\\ -2 & -4 & 1 & 9\end{bmatrix};$$

4) $$A_4 = AB_3 = \begin{bmatrix} 1 & -1 & -1 & 2 \\ 2 & 3 & 0 & -4 \\ 1 & 1 & -2 & -2 \\ 1 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 3 & -1 & -6 \\ -4 & -7 & 2 & 16 \\ 1 & 2 & -1 & -4 \\ -2 & -4 & 1 & 9 \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}; \quad q_4 = \frac{\mathrm{Sp}\, A_4}{4} = \frac{1+1+1+1}{4} = 1;$$

$$B_4 = A_4 - q_4 E = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = O_4.$$

5) Ahora bien, $p_1 = -q_1 = -1$; $p_2 = -q_2 = -2$; $p_3 = -q_3 = 3$; $p_4 = -q_4 = -1$ y $D(\lambda) = \lambda^4 - \lambda^3 - 2\lambda^2 + 3\lambda - 1$. ▲

La modificación del método de Le Verrier, propuesta por Faddéev, permite determinar la matriz inversa A^{-1}. De las fórmulas (3) tenemos $A_n = AB_{n-1}$, $B_n = A_n - q_nE = O_n$, de donde $A_n = q_nE$,

$$AB_{n-1} = q_nE. \tag{4}$$

Multiplicando a la izquierda la igualdad (4) por A^{-1}, obtenemos $A^{-1}AB_{n-1} = A^{-1}q_nE$, de donde

$$A^{-1} = \frac{B_{n-1}}{q_n} \text{ o bien } A^{-1} = \frac{B_{n-1}}{-p_n}. \tag{5}$$

Ejemplo 2. Calcular la matriz inversa para la matriz A dada en el ejemplo 1.

△ Utilizando la fórmula (5), obtenemos

$$A^{-1} = \frac{B_3}{-p_4} = 1 \begin{bmatrix} 2 & 3 & -1 & -6 \\ -4 & -7 & 2 & 16 \\ 1 & 2 & -1 & -4 \\ -2 & -4 & 1 & 9 \end{bmatrix}.$$

Verificación:

$$AA^{-1} = \begin{bmatrix} 1 & -1 & -1 & 2 \\ 2 & 3 & 0 & -4 \\ 1 & 1 & -2 & -2 \\ 1 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 3 & -1 & -6 \\ -4 & -7 & 2 & 16 \\ 1 & 2 & -1 & -4 \\ -2 & -4 & 1 & 9 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad ▲$$

§ 6.6. Cálculo de los vectores propios con ayuda del método de Le Verrier—Faddéev

Si se conocen las matrices $B_1, B_2, \ldots, B_{n-1}$, obtenidas por el método de Le Verrier — Faddéev, y las raíces $\lambda_1, \lambda_2, \ldots, \lambda_n$ del polinomio característico $D(\lambda)$, los vectores propios $\mathbf{x}^{(i)}$ se pueden calcular con ayuda de la fórmula

$$\mathbf{x}^{(i)} = \lambda_i^{(n-1)}\mathbf{e} + \lambda_i^{(n-2)}\mathbf{b}_1 + \lambda_i^{(n-3)}\mathbf{b}_2 + \ldots + \mathbf{b}_{n-1},$$

donde $\mathbf{e}$ es cualquier vector unitario y $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_{n-1}$ son los vectores columna de las matrices $B_1, B_2, \ldots, B_{n-1}$, homónimos con $\mathbf{e}$.

Ejemplo. Calcular los vectores propios de la matriz

$$A = \begin{bmatrix} 1 & -1 & -1 & 2 \\ 2 & 3 & 0 & -4 \\ 1 & 1 & -2 & -2 \\ 1 & 1 & 0 & -1 \end{bmatrix},$$

si se conocen las matrices B_1, B_2, B_3 (véase el ejemplo 1 en el § 6.5) y los valores propios $\lambda_1 = \lambda_2 = 1$; $\lambda_3 = 0{,}618$; $\lambda_4 = -1{,}168$ del polinomio característico $D(\lambda) = \lambda^4 - \lambda^3 - 2\lambda^2 + 3\lambda - 1$.

△ Tomemos $\mathbf{e} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$; entonces $\mathbf{b}_1 = \begin{bmatrix} 2 \\ -4 \\ -2 \\ -2 \end{bmatrix}$; $\mathbf{b}_2 = \begin{bmatrix} 4 \\ 0 \\ 6 \\ -2 \end{bmatrix}$;

$$\mathbf{b}_3 = \begin{bmatrix} -6 \\ 16 \\ -4 \\ 9 \end{bmatrix}$$

(las columnas 4 de las matrices B_1, B_2, B_3). Por la fórmula $\mathbf{x}^{(i)} = \lambda_i^3\mathbf{e} + \lambda_i^2\mathbf{b}_1 + \lambda_i\mathbf{b}_2 + \mathbf{b}_3$ hallamos:

$$\mathbf{x}^{(1)} = \mathbf{x}^{(2)} = 1 \cdot \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} + 1 \cdot \begin{bmatrix} 2 \\ -4 \\ -2 \\ -2 \end{bmatrix} + 1 \cdot \begin{bmatrix} 4 \\ 0 \\ 6 \\ -2 \end{bmatrix} + \begin{bmatrix} -6 \\ 16 \\ -4 \\ 9 \end{bmatrix} = \begin{bmatrix} 0 \\ 12 \\ 0 \\ 6 \end{bmatrix};$$

$$\mathbf{x}^{(3)} = 0{,}618^3 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} + 0{,}618^2 \begin{bmatrix} 2 \\ -4 \\ -2 \\ -2 \end{bmatrix} + 0{,}618 \begin{bmatrix} 4 \\ 0 \\ 6 \\ -2 \end{bmatrix} +$$

$$+ \begin{bmatrix} -6 \\ 16 \\ -4 \\ 9 \end{bmatrix} = \begin{bmatrix} -2{,}764 \\ 14{,}472 \\ -1{,}056 \\ 7{,}236 \end{bmatrix};$$

$$\mathbf{x}^{(4)} = (-1{,}618)^3 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} + (-1{,}618)^2 \begin{bmatrix} 2 \\ -4 \\ -2 \\ -2 \end{bmatrix} + (-1{,}618) \begin{bmatrix} 4 \\ 0 \\ 6 \\ -2 \end{bmatrix} +$$
$$+ \begin{bmatrix} -6 \\ 16 \\ -4 \\ 9 \end{bmatrix} = \begin{bmatrix} -7{,}236 \\ 5{,}528 \\ -18{,}944 \\ 2{,}764 \end{bmatrix}.$$

Escribimos los resultados de los cálculos en la tabla siguiente:

λ_i	I	II	III	IV	V	VI
$\lambda_1 = \lambda_2 = 1$	0 0 0 1	2 −4 −2 −2	4 0 6 −2	−6 16 −4 6	0 12 0 6	0 1 0 0,5
$\lambda_3 = 0{,}618$	0 0 0 0,236	0,764 −1,528 −0,764 −0,764	2,472 0 3,708 −1,236	−6 16 −4 9	−2,764 14,472 −1,056 7,236	0,19 1 −0,07 0,50
$\lambda_4 = -1{,}618$	0 0 0 −4,236	5,236 −10,472 −5,236 −5,236	−6,472 0 −9,708 3,236	−6 16 −4 9	−7,236 5,528 −18,944 2,764	−0,38 0,29 1 0,15

En las columnas II, III, IV están escritas las coordenadas de la columna 4 de las matrices B_i, multiplicadas por los grados correspondientes de λ_i y en la columna 1, las coordenadas del vector $\lambda_i^3 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$. La columna V contiene las coordenadas de los vectores $\mathbf{x}^{(i)}$ y la columna VI, sus coordenadas después de la normación.▲

§ 6.7. Método de Danilevski

Dos matrices A y B se llaman *semejantes* si una se obtiene de la otra por medio de la transformación con ayuda de cierta matriz regular, o sea, se cumple la igualdad

$$B = S^{-1}AS.$$

Si la matriz B es semejante a la A, se escribe $B \backsim A$.

Al construir el esquema de cálculo, en el **método de Danilevski** se utiliza la propiedad fundamental de las matrices semejantes: *las matrices semejantes tienen iguales polinomios característicos.*

Si se reduce la matriz dada

$$A=\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix} \tag{1}$$

con ayuda de las transformaciones de semejanza a la llamada *forma de Frobenius*

$$F=\begin{bmatrix} f_{11} & f_{12} & f_{13} & \dots & f_{1,n-1} & f_{1n} \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix} \tag{2}$$

y luego se desarrolla el determinante

$$\det(F-\lambda E)=\begin{vmatrix} f_{11}-\lambda & f_{12} & f_{13} & \dots & f_{1,n-1} & f_{1n} \\ 1 & -\lambda & 0 & \dots & 0 & 0 \\ 0 & 1 & -\lambda & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & -\lambda \end{vmatrix} \tag{3}$$

según los elementos de la primera fila, obtenemos

$$D(\lambda) = \det(F-\lambda E) = (f_{11}-\lambda)(-\lambda)^{n-1} - f_{12}(-\lambda)^{n-2} + \\ + f_{13}(-\lambda)^{n-3} - \dots + (-1)^{n-1} f_{1n},$$

o bien

$$D(\lambda) = \det(F-\lambda E) = (-1)^n (\lambda^n - p_1\lambda^{n-1} - p_2\lambda^{n-2} - \\ - p_3\lambda^{n-3} - \dots p_n). \tag{4}$$

Aquí $p_1 = f_{11}$, $p_2 = f_{12}$, $p_3 = f_{13}$, $\dots$, $p_n = f_{1n}$ son los coeficientes del polinomio característico de la matriz F los cuales en virtud de semejanza de las matrices F y A son coeficientes del polinomio característico de la matriz dada A.

Según el método de Danilevski el paso de la matriz A a la matriz semejante de Frobenius F se lleva a cabo con ayuda de $n-1$ transformaciones de semejanza que transforman sucesivamente las filas de la matriz A, comenzando con la última, en filas correspondientes de la matriz F.

Esquema de transformación de la matriz A en matriz semejante de Frobenius F. 1°. Supongamos que nos hace falta convertir la fila $a_{n1}a_{n2} \dots a_{n,n-1}a_{nn}$ en fila $0\ 0 \dots 1\ 0$. Admitiendo que $a_{n,n-1} \neq$

$\neq 0$, dividamos todos los elementos de la columna $n-1$ de la matriz A por $a_{n,\,n-1}$. Entonces su n-ésima fila toma la forma

$$a_{n1}a_{n2}\ldots\frac{a_{n,\,n-1}}{a_{n,\,n-1}}a_{nn}\quad\text{o bien}\ a_{n1}a_{n2}\ldots 1\ a_{nn}.$$

2°. Sustrayamos la columna $n-1$ de la matriz transformada, multiplicada, correspondientemente, por los números a_{n1}, a_{n2}, ..., a_{nn}, de todas las demás columnas. Para la n-ésima fila obtenemos

$$a_{n1}-a_{n1}a_{n2}-a_{n2}\ \ldots\ 1\ a_{nn}-a_{nn},\ \text{o bien}\ 0\ 0\ \ldots\ 1\ 0.$$

3°. En calidad de matriz regular tomamos la matriz M_{n-1} que se obtiene de la matriz unidad después de las mismas transformaciones:

$$M_{n-1}=\begin{bmatrix} 1 & 0 & \ldots & 0 & 0\\ 0 & 1 & \ldots & 0 & 0\\ \ldots & \ldots & \ldots & \ldots & \ldots\\ m_{n-1,\,1} & m_{n-1,\,2} & & m_{n-1,\,n-1} & m_{n-1,\,n}\\ 0 & 0 & \ldots & 0 & 1\end{bmatrix},$$

donde

$$m_{n-1,\,i}=-\frac{a_{ni}}{a_{n,\,n-1}},\quad m_{n-1,\,n-1}=\frac{1}{a_{n,\,n-1}}. \tag{5}$$

Las operaciones realizadas son equivalentes a la multiplicación, a la derecha, de la matriz M_{n-1} por la A:

$$B=AM_{n-1}=\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1,\,n-1} & a_{1n}\\ a_{21} & a_{22} & \ldots & a_{2,\,n-1} & a_{2n}\\ \ldots & \ldots & \ldots & \ldots & \ldots\\ a_{n-1,\,1} & a_{n-1,\,2} & \ldots & a_{n-1,\,n-1} & a_{n-1,\,n}\\ 0 & 0 & \ldots & 1 & 0\end{bmatrix}\times$$

$$\times\begin{bmatrix} 1 & 0 & \ldots & 0 & 0\\ 0 & 1 & \ldots & 0 & 0\\ \ldots & \ldots & \ldots & \ldots & \ldots\\ m_{n-1,\,1} & m_{n-1,\,2} & \ldots & m_{n-1,\,n-1} & m_{n-1,\,n}\\ 0 & 0 & \ldots & 0 & 1\end{bmatrix}=$$

$$=\begin{bmatrix} b_{11} & b_{12} & \ldots & b_{1,\,n-1} & b_{1n}\\ b_{21} & b_{22} & \ldots & b_{2,\,n-1} & b_{2n}\\ \ldots & \ldots & \ldots & \ldots & \ldots\\ b_{n-1,\,1} & b_{n-1,\,2} & \ldots & b_{n-1,\,n-1} & b_{n-1,\,n}\\ 0 & 0 & \ldots & 1 & 0\end{bmatrix}.$$

Los elementos de la matriz B se calculan con ayuda de las fórmulas

$$b_{ij}=a_{ij}+a_{i,\,n-1}m_{n-1,\,j};\quad b_{i,\,n-1}=a_{i,\,n-1}m_{n-1,\,n-1}. \tag{6}$$

Sin embargo, la matriz construida $B = AM_{n-1}$ no será semejante a la matriz A.

4°. Para obtener la transformación de semejanza es necesario multiplicar a la izquierda la matriz inversa M_{n-1}^{-1} por la matriz B:

$$M_{n-1}^{-1}AM_{n-1} = M_{n-1}^{-1}B.$$

La matriz inversa M_{n-1}^{-1} tiene la forma

$$M_{n-1}^{-1} = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{n,\,n-1} & a_{nn} \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix}.$$

Suponemos $M_{n-1}^{-1}AM_{n-1} = C$; por consiguiente, $C = M_{n-1}^{-1}B$. La multiplicación de la matriz M_{n-1}^{-1} a la izquierda por la matriz B no cambia la fila transformada de esta última y la matriz C tiene la forma

$$C = M_{n-1}^{-1}B = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{n,\,n-1} & a_{nn} \\ 0 & 0 & & 0 & 1 \end{bmatrix} \times$$

$$\times \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1,\,n-1} & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2,\,n-1} & b_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ b_{n-1,\,1} & b_{n-1,\,2} & \dots & b_{n-1,\,n-1} & b_{n-1,\,n} \\ 0 & 0 & & 1 & 0 \end{bmatrix} =$$

$$= \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1,\,n-1} & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2,\,n-1} & c_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ c_{n-1,\,1} & c_{n-1,\,2} & \dots & c_{n-1,\,n-1} & c_{n-1,\,n} \\ 0 & 0 & & 1 & 0 \end{bmatrix}.$$

En efecto, multiplicando las matrices M_{n-1}^{-1} y B, cambiamos sólo la fila $(n - 1)$-ésima de la matriz B, ya que $c_{ij} = b_{ij}$ para todas las demás filas. Los elementos de esta fila se determinan con ayuda de las fórmulas

$$c_{n-1,\,j} = \sum_{k=1}^{n} a_{nk}b_{kj} \quad (j = 1,\ 2,\ \dots,\ n). \tag{7}$$

La matriz obtenida C es semejante a la A y tiene una fila reducida.

5°. Luego, si $c_{n-1\ n-2} \neq 0$, entonces con la matriz C repetimos las operaciones análogas, tomando por fundamental la fila $n - 2$. En este caso, utilizando la matriz intermedia $D = CM_{n-2}$, como resultado obtenemos la matriz $E = M_{n-2}^{-1}D = M_{n-2}^{-1}CM_{n-2}$ con dos filas reducidas. Con la matriz E hacemos las mismas operaciones, etc. hasta que se obtenga la matriz de Frobenius.

Todas estas transformaciones se formalizan en forma esquemática de cálculo, el proceso de composición de la cual consideremos en un ejemplo.

Tabla 6.6

Filas	M^{-1}	Columnas 1	2	3	4	Σ	Σ'
1		−4	−3	1	1	−5	
2		2	0	4	−1	5	
3		1	1	2	−2	2	
4		1	1	$\boxed{-1}$	−1	0	
I	M_3 \ M_3^{-1}	1	1	−1	−1	0	
5	1	−3	−2	−1	0	−6	−5
6	1	6	4	−4	−5	1	5
7	−1	3	3	−2	−4	0	2
8	−1	0	0	1	0	1	0
7′		0	$\boxed{-1}$	−4	−1	−6	
II	M_2 \ M_2^{-1}	0	−1	−4	−1	−6	
9	0	−3	2	7	2	8	6
10	−1	6	−4	−20	−9	−27	−23
11	−4	0	1	0	0	1	0
12	−1	0	0	1	0	1	0
10′		$\boxed{-6}$	0	19	9	22	
III	M_1 \ M_1^{-1}	0,167−1	0	3,167	1,500	3,667	
13	−6	0,500	2,000	−2,500	−2,500	−2,500	−3
14	0	1	0	0	0	1	0
15	19	0	1	0	0	1	0
16	9	0	0	1	0	1	0
13′		−3	7	24	15	43	

Ejemplo 1. Haciendo uso del método de Danilevski, desarrollar el determinante característico

$$A=\begin{bmatrix} -4 & -3 & 1 & 1 \\ 2 & 0 & 4 & -1 \\ 1 & 1 & 2 & -2 \\ 1 & 1 & -1 & -1 \end{bmatrix}.$$

△ *Etapa I.* Reduzcamos la matriz A a la forma de Frobenius. Para los cálculos hacemos la tabla (véase la tabla 6.6).

1) En las filas 1 — 4 de la tabla de cálculo colocamos los elementos a_{ij} $(i, j = 1, 2, 3, 4)$ de la matriz dada A y las sumas de control $a_{i5} = \sum_{j=1}^{4} a_{ij}$ $(i = 1, 2, 3, 4)$, o sea, la columna Σ. Marcamos el elemento $a_{43} = -1$ perteneciente a la tercera columna (columna marcada).

2) En la fila 1 escribimos los elementos de la tercera fila de la matriz $M_{n-1} = M_3$ los cuales se calculan con ayuda de las fórmulas (5):

$$m_{31} = -\frac{a_{41}}{a_{43}} = -\frac{1}{-1} = 1; \quad m_{32} = -\frac{a_{42}}{a_{43}} = -\frac{1}{-1} = 1;$$

$$m_{33} = \frac{1}{a_{43}} = \frac{1}{-1} = -1; \quad m_{34} = -\frac{a_{44}}{a_{43}} = -\frac{-1}{-1} = -1;$$

$$m_{35} = -\frac{a_{45}}{a_{43}} = \frac{0}{-1} = 0.$$

El número 0 debe coincidir con la suma de los elementos de la fila 1 después de sustituir el valor obtenido del elemento m_{33} por -1, pero en el ejemplo dado $m_{33} = -1$. (Para comodidad el número -1 suele escribirse junto al elemento m_{33} y se separa de este último por la raya).

3) En las filas 5...8 en la columna M^{-1} escribimos la tercera fila de la matriz M_3^{-1} la cual debe coincidir con la cuarta fila de la matriz inicial A.

4) En las filas 5...8 en las columnas correspondientes escribimos los elementos de la matriz $B = A \cdot M_3$ los cuales se calculan por las fórmulas (6) para las columnas no marcadas.

La primera columna:

$$b_{11} = a_{11} + a_{13}m_{31} = -4 + 1\cdot 1 = -3;$$
$$b_{21} = a_{21} + a_{23}m_{31} = 2 + 4\cdot 1 = 6;$$
$$b_{31} = a_{31} + a_{33}m_{31} = 1 + 2\cdot 1 = 3;$$
$$b_{41} = a_{41} + a_{43}m_{31} = 1 + (-1)\cdot 1 = 0.$$

La segunda columna:

$$b_{12} = a_{12} + a_{13}m_{32} = -3 + 1\cdot 1 = -2;$$
$$b_{22} = a_{22} + a_{23}m_{32} = 0 + 4\cdot 1 = 4;$$
$$b_{32} = a_{32} + a_{33}m_{32} = 1 + 2\cdot 1 = 3;$$
$$b_{42} = a_{42} + a_{43}m_{32} = 1 + (-1)\cdot 1 = 0.$$

La cuarta columna:

$$b_{14} = a_{14} + a_{13} \cdot m_{34} = 1 + 1 \cdot (-1) = 0;$$
$$b_{24} = a_{24} + a_{23} m_{34} = -1 + 4 \cdot (-1) = -5;$$
$$b_{34} = a_{34} + a_{33} m_{34} = -2 + 2 \cdot (-1) = -4;$$
$$b_{44} = a_{44} + a_{43} m_{34} = -1 + (-1) \cdot (-1) = 0.$$

Los elementos transformados de la tercera columna (marcada) se obtienen con ayuda de la multiplicación de los elementos iniciales por $m_{33} = -1$.

La tercera columna:

$$b_{33} = a_{13} m_{33} = 1 \cdot (-1) = -1; \quad b_{23} = a_{23} \cdot m_{33} = 4 \cdot (-1) = -4;$$
$$b_{33} = a_{33} m_{33} = 2 \cdot (-1) = -2; \quad b_{43} = a_{43} m_{33} = (-1)(-1) = 1.$$

La última fila de la matriz B debe tener la forma 0 0 1 0.

Para el control completamos la matriz B por los elementos correspondientes de la columna Σ' transformados según las fórmulas análogas de dos términos para $m_{35} = 0$:

$$b_{16} = a_{15} + a_{13} m_{35} = -5 + 1 \cdot 0 = -5;$$
$$b_{26} = a_{25} + a_{23} m_{35} = 5 + 4 \cdot 0 = 5;$$
$$b_{36} = a_{35} + a_{33} m_{35} = 2 + 2 \cdot 0 = 2;$$
$$b_{46} = a_{45} + a_{43} m_{35} = 0 + (-1) \cdot 0 = 0.$$

Escribimos los resultados obtenidos en las filas correspondientes de la columna Σ'. Al adicionar a los elementos de la columna Σ' los elementos respectivos de la tercera columna (marcada), obtenemos las sumas de control para las filas 5...8 ($i = 1, 2, 3, 4$).

La columna Σ:

$$b_{15} = b_{16} + a_{13} = -5 - 1 = -6;$$
$$b_{25} = b_{26} + a_{23} = 5 - 4 = 1,$$
$$b_{35} = b_{36} + a_{33} = 2 - 2 = 0;$$
$$b_{45} = b_{46} + a_{43} = 0 + 1 = 1.$$

Además, los elementos de la columna Σ para el control se calculan con ayuda de la fórmula $b_{i5} = \sum_{j=1}^{4} b_{ij}$ $(i = 1, 2, 3, 4)$:

$$b_{15} = b_{11} + b_{12} + b_{13} + b_{14} = -3 - 2 - 1 + 0 = -6;$$
$$b_{25} = b_{21} + b_{22} + b_{23} + b_{24} = 6 + 4 - 4 - 5 = 1;$$
$$b_{35} = b_{31} + b_{32} + b_{33} + b_{34} = 3 + 3 - 2 - 4 = 0;$$
$$b_{45} = b_{41} + b_{42} + b_{43} + b_{44} = 0 + 0 + 1 + 0 = 1.$$

La matriz B tiene la forma siguiente:

$$B = \begin{bmatrix} -3 & -2 & -1 & 0 \\ 6 & 4 & -4 & -5 \\ 3 & 3 & -2 & -4 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

5) La transformación M_3^{-1} realizada en la matriz B, produciendo como resultado la matriz $C = M_3^{-1}B$, cambia sólo la tercera fila de la matriz B, o sea, la séptima fila de la tabla. Los elementos de esta fila transformada 7′ no son sino las sumas de los productos pares de los elementos de la columna M_3^{-1}, colocados en las filas 5...8, por los elementos correspondientes de cada una de las columnas de la matriz B [véanse las fórmulas (7)]:

$$c_{31} = 1\cdot(-3) + 1\cdot 6 + (-1)\cdot 3 + (-1)\cdot 0 = 0;$$
$$c_{32} = 1\cdot(-2) + 1\cdot 4 + (-1)\cdot 3 + (-1)\cdot 0 = -1;$$
$$c_{33} = 1\cdot(-1) + 1\cdot(-4) + (-1)\cdot(-2) + (-1)\cdot 1 = -4;$$
$$c_{34} = 1\cdot 0 + 1\cdot(-5) + (-1)(-4) + (-1)\cdot 0 = -1.$$

Realizamos las mismas transformaciones en la columna Σ:

$$c_{35} = 1\cdot(-6) + 1\cdot 1 + (-1)\cdot 0 + (-1)\cdot 1 = -6.$$

Como resultado obtenemos la matriz C compuesta por las filas 5, 6, 7′, 8 con las sumas de control en la columna Σ:

$$C = \begin{bmatrix} -3 & -2 & -1 & 0 \\ 6 & 4 & -4 & -5 \\ 0 & \boxed{-1} & -4 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

La matriz C es semejante a la matriz A y tiene una fila reducida. En esto se termina la construcción de la primera transformación semejante: $C = M_2^{-1}AM$.

Etapa II. Tomando por inicial la matriz C, separamos el elemento $c_{32} = -1$ (segunda columna) y continuamos procediendo de un modo análogo.

1) Hallamos los elementos de la matriz $M_{n-2} = M_2$ son ayuda de las fórmulas (5):

$$m_{21} = -\frac{c_{31}}{c_{32}} = -\frac{0}{-1} = 0; \quad m_{22} = \frac{1}{c_{32}} = \frac{1}{-1} = -1;$$
$$m_{23} = -\frac{c_{33}}{c_{32}} = -\frac{-4}{-1} = -4; \quad m_{24} = -\frac{c_{34}}{c_{32}} = -\frac{-1}{-1} = -1;$$
$$m_{25} = -\frac{c_{35}}{c_{32}} = -\frac{-6}{-1} = -6.$$

Vamos a sumar: $0 - 1 - 4 - 1 = -6$ ($m_{22} = -1$; si $m_{22} \neq -1$, sería necesario reemplazar m_{22} por -1).

2) En las filas 9...12 en la columna M^{-1} escribimos la segunda fila de la matriz M_2^{-1} la cual coincide con la tercera fila de la matriz C (véase la tabla 6.6). Hallamos los elementos de la matriz $D = CM_2$.

La primera columna:

$$d_{11} = -3 + (-2)\cdot 0 = -3; \quad d_{21} = 6 + 4\cdot 0 = -6;$$
$$d_{31} = 0 + 0 = 0.$$

La segunda columna (marcada) se obtiene multiplicando los elementos correspondientes de la matriz C por $m_{22} = -1$:

$$d_{12} = c_{12}m_{22} = (-2)\cdot(-1) = 2; \ d_{22} = c_{22}m_{22} = 4\cdot(-1) = -4;$$
$$d_{32} = c_{32}m_{22} = (-1)\cdot(-1) = 1.$$

La tercera columna:

$$\begin{aligned} d_{13} &= c_{13} + c_{12}m_{23} = -1 + (-2)\cdot(-4) = 7; \\ d_{23} &= c_{23} + c_{22}m_{23} = -4 + 4\cdot(-4) = -20; \\ d_{33} &= c_{33} + c_{32}m_{23} = -4 + (-1)\cdot(-4) = 0. \end{aligned}$$

La cuarta columna:

$$\begin{aligned} d_{14} &= c_{14} + c_{12}m_{24} = 0\cdot 1 + (-2)\cdot(-1) = 2; \\ d_{24} &= c_{24} + c_{22}m_{24} = -5 + 4\cdot(-1) = -9; \\ d_{34} &= c_{34} + c_{32}m_{24} = -1 + (-1)\cdot(-1) = 0. \end{aligned}$$

La columna Σ':

$$\begin{aligned} d_{16} &= c_{15} + c_{12}m_{25} = -6 + (-2)\cdot(-6) = 6; \\ d_{26} &= c_{25} + c_{22}m_{25} = 1 + 4\cdot(-6) = -23; \\ d_{36} &= c_{35} + c_{32}m_{25} = -6 + (-1)\,(-6) = 0. \end{aligned}$$

Los elementos de la matriz Σ se obtienen sumando los elementos de la columna Σ' con los elementos correspondientes de la columna marcada:

$$\begin{aligned} d_{15} &= d_{16} + d_{12} = 6 + 2 = 8; \\ d_{25} &= d_{26} + d_{22} = -23 - 4 = -27; \\ d_{36} &= d_{36} + d_{32} = 0 + 1 = 1. \end{aligned}$$

La matriz D tiene la forma

$$D = \begin{bmatrix} -3 & 2 & 7 & 2 \\ 6 & -4 & -20 & -9 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

La transformación M_2^{-1} que se realiza en la matriz D y ofrece la matriz $E = M_2^{-1}D$ cambia sólo la segunda fila de la matriz D, o sea, la décima fila de la tabla. Los elementos de esta fila transformada $10'$ no son sino las sumas de los productos pares de los elementos de la columna M_2^{-1} que están en las filas 9...12:

$$e_{21} = 0\cdot(-3) + (-1)\cdot 6 + (-4)\cdot 0 + (-1)\cdot 0 = -6;$$
$$e_{22} = 0\cdot 2 + (-1)\cdot(-4) + (-4)\cdot 1 + (-1)\cdot 0 = 0;$$
$$e_{23} = 0\cdot 7 + (-1)\cdot(-20) + (-4)\cdot 0 + (-1)\cdot 1 = 19;$$
$$e_{24} = 0\cdot 2 + (-1)\cdot(-9) + (-4)\cdot 0 + (-1)\cdot 0 = 9;$$
$$e_{25} = 0\cdot 8 + (-1)\cdot(-27) + (-4)\cdot 1 + (-1)\cdot 1 = 22;$$
$$\Sigma e_{2j} = -6 + 0 + 19 + 9 = 22.$$

En esto termina la construcción de la segunda transformación semejante: $E = M_2^{-1}CM_2$. La matriz $E \backsim C$ contiene dos filas reducidas:

$$E = \begin{bmatrix} -3 & 2 & 7 & 2 \\ \boxed{-6} & 0 & 19 & 9 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Etapa III. Tomamos por inicial la matriz E. En ésta separamos el elemento $e_{21} = -6$ (primera columna) y transformamos la matriz E en semejante matriz de Frobenius F. Continuando de un modo análogo el proceso, con ayuda de las fórmulas (5) determinemos los elementos de la matriz $M_{n-3} = M_1$:

$$m_{11} = \frac{1}{e_{21}} = \frac{1}{-6} = -0,167; \quad m_{12} = -\frac{e_{22}}{e_{21}} = -\frac{0}{-6} = 0;$$
$$m_{13} = -\frac{e_{23}}{e_{21}} = -\frac{19}{-6} = 3,167; \quad m_{14} = -\frac{e_{24}}{e_{21}} = -\frac{9}{-6} = 1,500;$$
$$m_{15} = -\frac{e_{25}}{e_{21}} = -\frac{22}{-6} = 3,667.$$

Para que se obtenga la suma $\Sigma = 3,667$ reemplazamos $m_{11} = -0,167$ por -1;

$$\Sigma = -1 + 0 + 3,167 + 1,500 = 3,667.$$

Vamos a escribir la matriz de Frobenius F en las filas 13...16. Primero construimos $G = EM_1$ y luego $F = M_1^{-1}G$. En la columna M_1^{-1} escribimos la fila 10′ de la matriz E (véase la tabla 6.6).

La primera columna (marcada):

$$g_{11} = e_{11}m_{11} = (-3)\cdot(-0,167) = 0,500;$$
$$g_{21} = e_{21}m_{11} = (-6)\cdot(-0,167) = 1,000.$$

La segunda columna:

$$g_{12} = e_{12} + e_{11}m_{12} = 2 + (-3)\cdot 0 = 2,000;$$
$$g_{22} = e_{22} + e_{21}m_{12} = 0 + (-6)\cdot 0 = 0.$$

La tercera columna:

$$g_{13} = e_{13} + e_{11}m_{13} = 7 + (-3)\cdot 3{,}167 = -2{,}500;$$
$$g_{23} = e_{23} + e_{21}m_{13} = 19 + (-6)\cdot 3{,}167 = 0.$$

La cuarta columna:

$$g_{14} = e_{14} + e_{11}m_{14} = 2 + (-3)\cdot 1{,}500 = -2{,}500;$$
$$g_{24} = e_{24} + e_{21}m_{14} = 9 + (-6)\cdot 1{,}500 = 0.$$

La columna Σ':

$$g_{16} = e_{15} + e_{11}m_{15} = 8 + (-3)\cdot 3{,}667 = -3;$$
$$g_{26} = e_{25} + e_{21}m_{15} = 22 + (-6)\cdot 3{,}667 = 0.$$

La columna Σ:

$$g_{15} = g_{16} + g_{11} = -3 + 0{,}500 = -2{,}500;$$
$$g_{15} = g_{11} + g_{12} + g_{13} + g_{14} = 0{,}500 + 2{,}000 - 2{,}500 - 2{,}500 = = -2{,}500;$$
$$g_{25} = g_{26} + g_{21} = 0 + 1 = 1;$$
$$g_{25} = g_{21} + g_{22} + g_{23} + g_{24} = 1 + 0 + 0 + 0 = 1.$$

Los elementos de la fila transformada 13′ no son más que las sumas de los productos pares de la columna M_1^{-1} que están en las filas 13...16:

$$f_{11} = (-6)\cdot 0{,}500 + 0\cdot 1 + 19\cdot 0 + 9\cdot 0 = -3;$$
$$f_{12} = (-6)\cdot 2{,}000 + 0\cdot 0 + 19\cdot 1 + 9\cdot 0 = 7;$$
$$f_{13} = (-6)(-2{,}500) + 0\cdot 0 + 19\cdot 0 + 9\cdot 1 = 24;$$
$$f_{14} = (-6)(-2{,}500) + 0\cdot 0 + 19\cdot 0 + 9\cdot 0 = 15.$$
$$\Sigma = (-6)\cdot(-2{,}500) + 0\cdot 1 + 19\cdot 1 + 9\cdot 1 = 43;$$
$$\Sigma = -3 + 7 + 24 + 15 = 43.$$

Ahora bien, la matriz buscada de Frobenius F, semejante a la A, tiene la forma

$$F = \begin{bmatrix} -3 & 7 & 24 & 15 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Vamos a buscar el determinante característico de la matriz F:

$$D(\lambda) = \det(A - \lambda E) = \det(F - \lambda E) = \begin{vmatrix} -3-\lambda & 7 & 24 & 15 \\ 1 & -\lambda & 0 & 0 \\ 0 & 1 & -\lambda & 0 \\ 0 & 0 & 1 & -\lambda \end{vmatrix}.$$

De aquí, desarrollando el determinante $D(\lambda)$ según los elementos de la primera fila, obtenemos

$$D(\lambda)=\begin{vmatrix} -3-\lambda & 7 & 24 & 15 \\ 1 & -\lambda & 0 & 0 \\ 0 & 1 & -\lambda & 0 \\ 0 & 0 & 1 & \lambda \end{vmatrix} = (-3-\lambda)\begin{vmatrix} -\lambda & 0 & 0 \\ 1 & -\lambda & 0 \\ 0 & 1 & -\lambda \end{vmatrix} -$$

$$-7\begin{vmatrix} 1 & 0 & 0 \\ 0 & -\lambda & 0 \\ 0 & 1 & -\lambda \end{vmatrix} + 24\begin{vmatrix} 1 & -\lambda & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\lambda \end{vmatrix} - 15\begin{vmatrix} 1 & -\lambda & 0 \\ 0 & 1 & -\lambda \\ 0 & 0 & 1 \end{vmatrix} =$$

$$= (-3-\lambda)(-\lambda^3) - 7\lambda^2 + 24(-\lambda) - 15 =$$

$$= \lambda^4 + 3\lambda^3 - 7\lambda^2 - 24\lambda - 15. \quad \blacktriangle$$

Casos excepcionales en el método de Danilevski. Este método es aplicable sin complicaciones algunas si todos los elementos a separar son distintos del cero (al igual que en el ejemplo recién considerado). En cambio, si al transformar la matriz $A = [a_{ij}]$ ($i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, n$) en matriz de Frobenius F se obtiene la matriz en la forma

$$D=\begin{bmatrix} d_{11} & d_{12} & \ldots & d_{1l} & \ldots & d_{1,\,k-1} & d_{1k} & \ldots & d_{1,\,n-1} & d_{1n} \\ d_{21} & d_{22} & \ldots & d_{2l} & \ldots & d_{2,\,k-1} & d_{2k} & \ldots & d_{2,\,n-1} & d_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ d_{l1} & d_{l2} & \ldots & d_{ll} & \ldots & d_{l,\,k-1} & d_{lk} & \ldots & d_{l,\,n-1} & d_{ln} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ d_{k1} & d_{k2} & \ldots & d_{kl} & \ldots & d_{k,\,k-1} & d_{kk} & \ldots & d_{k,\,n-1} & d_{kn} \\ 0 & 0 & \ldots & 0 & \ldots & 0 & 1 & \ldots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & 0 & \ldots & 0 & 0 & \ldots & 1 & 0 \end{bmatrix},$$

con la particularidad de que resulta que $d_{k,\,k-1} = 0$, no se puede continuar las transformaciones siguiendo el método de Danilevski. Aquí son posibles dos casos.

Caso I. Supongamos que cualquier elemento de la matriz D, el cual está a la izquierda del elemento nulo $d_{k,\,k-1}$, es distinto del cero; por ejemplo, $d_{kl} \neq 0$, donde $l < k - 1$. Entonces promovemos este elemento al lugar del elemento nulo $d_{k,\,k-1}$, o sea, permutamos la columna $k - 1$ y la columna l de la matriz D y simultáneamente su fila $k - 1$ y su fila l. La nueva matriz será semejante a la dada y se puede continuar los cálculos según el método de Danilevski.

Caso II. Sea $d_{kl} = 0$ ($l = 1, 2, \ldots, k - 1$), es decir, el elemento separado, así como todos los elementos de la matriz que están a la izquierda del separado son iguales a cero. Entonces la matriz D tiene la forma

$$D=\left[\begin{array}{cccc|cccc}
d_{11} & d_{12} & \dots & d_{1,k-1} & d_{1k} & \dots & d_{1,n-1} & d_{1n}\\
d_{21} & d_{22} & \dots & d_{2,k-1} & d_{2k} & \dots & d_{2,n-1} & d_{2n}\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
d_{k-1,1} & d_{k-1,2} & \dots & d_{k-1,k-1} & d_{k-1,k} & \dots & d_{k-1,n-1} & d_{k-1,n}\\
\hline
0 & 0 & \dots & 0 & d_{kk} & \dots & d_{k,n-1} & d_{kn}\\
0 & 0 & \dots & 0 & 1 & \dots & 0 & 0\\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots\\
0 & 0 & \dots & 0 & 0 & \dots & 1 & 0
\end{array}\right]=$$

$$=\left[\begin{array}{c|c} D_1 & D_2\\ \hline 0 & D_3\end{array}\right].$$

Partimos la matriz D en cuatro células de un modo tal que una matriz sea nula. Entonces el determinante característico $\det(D-\lambda E)$ se descompone en dos determinantes:

$$\det(D-\lambda E)=\det(D_1-\lambda E)\cdot\det(D_3-\lambda E),$$

pero la matriz D_3 tiene ya la forma de Frobenius, por eso nos queda sólo por reducir a esta forma la matriz D_1.

Tabla 6.7

Filas	M_I	Columnas 1	2	3	4	Σ	Σ'
1		3	−2	1	−1	1	
2		3	−2	1	1	3	
3		5	−4	2	0	3	
4		−1	−1	$\boxed{1}$	1	0	
I	M_3 / M_3^{-1}	1	1	$1 \mid -1$	−1	0	
5	−1	4	−1	1	−2	2	1
6	−1	4	−1	1	0	4	3
7	1	7	−2	2	−2	5	3
8	1	0	0	1	0	1	0
7′		−1	$\boxed{0}$	1	0	0	

Ejemplo 2. Haciendo uso del método de Danilevski, desarrollar el determinante característico de la matriz

$$A=\begin{bmatrix} 3 & -2 & 1 & -1 \\ 3 & -2 & 1 & 1 \\ 5 & -4 & 2 & 0 \\ -1 & -1 & 1 & 1 \end{bmatrix}.$$

△ Escribimos los cálculos en la tabla (véase la tabla 6.7). El elemento a separar $c_{32}=0$; no se puede continuar los cálculos siguiendo el esquema de Danilevski. Puesto que $c_{31}\neq 0$, permutamos las columnas primera y segunda y las filas primera y segunda de la matriz C y continuamos los cálculos (véase la tabla 6.8).

Tabla 6.8

Filas	M^{-1}	Columnas				Σ	Σ'
		1	2	3	4		
5	-1	-1	4	1	0	4	
6	-1	-1	4	1	-2	2	
7′	1	0	$\boxed{-1}$	1	0	0	
8	1	0	0	1	0	1	
II	M_2 / M_2^{-1}	0	-1	1	0	0	
9	0	-1	-4	5	0	0	4
10	-1	-1	-4	5	-2	-2	2
11	1	0	1	0	0	1	0
12	0	0	0	1	0	1	0
10′		$\boxed{1}$	5	-5	2	3	
III	M_1 / M_1^{-1}	1 | -1	-5	5	-2	-3	
13	1	-1	1	0	2	2	3
14	5	1	0	0	0	1	0
15	-5	0	1	0	0	1	0
16	2	0	0	1	0	1	0
13′		4	-4	2	2	4	

Como resultado obtenemos la matriz de Frobenius $F \backsim A$:

$$F = \begin{bmatrix} 4 & -4 & 2 & 2 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

de donde hallamos

$$D(\lambda) = \det(A - \lambda E) = \det(F - \lambda E) = \lambda^4 - 4\lambda^3 + \\ + 4\lambda^2 - 2\lambda + 2. \blacktriangle$$

Ejemplo 3. Haciendo uso del método de Danilevski, desarrollar el determinante característico de la matriz

$$A = \begin{bmatrix} 0 & 1 & 3 & 2 \\ 1 & 4 & 5 & 0 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

△ Escribamos los resultados de cálculos en la tabla (véase la tabla 6.9).

Puesto que el elemento a separar es igual a cero, no se puede continuar los cálculos según el esquema de Danilevski.

La matriz $D \backsim C$ tiene la forma

$$D = \left[\begin{array}{c|ccc} 1 & 0{,}5 & -2{,}5 & 2{,}5 \\ \hline 0 & 6 & 4 & -7 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right].$$

Partimos la matriz D en cuatro células por orladura y calculamos $D(\lambda)$:

$$D(\lambda) = \det(D - \lambda E) = \left|\begin{array}{c|ccc} 1-\lambda & 0{,}5 & -2{,}5 & 2{,}5 \\ \hline 0 & 6-\lambda & 4 & -7 \\ 0 & 1 & -\lambda & 0 \\ 0 & 0 & 1 & -\lambda \end{array}\right| =$$

$$= (1-\lambda)\begin{vmatrix} 6-\lambda & 4 & -7 \\ 1 & -\lambda & 0 \\ 1 & 1 & -\lambda \end{vmatrix} = (1-\lambda)\left[(6-\lambda)\lambda^2 + 4\lambda - 7\right] =$$

$$= \lambda^4 - 7\lambda^3 + 2\lambda^2 + 11\lambda - 7. \blacktriangle$$

Tabla 6.9

Filas	M^{-1}	Columnas 1	2	3	4	Σ	Σ'
1		0	1	3	2	6	
2		1	4	5	0	10	
3		1	1	2	1	5	
4		1	1	$\boxed{1}$	1	4	
I	M_3 / M_3^{-1}	−1	−1	1 | −1	−1	−4	
5	1	−3	−2	3	−1	−3	−6
6	1	−4	−1	5	−5	−5	−10
7	1	−1	−1	2	−1	−1	−3
8	1	0	0	1	0	1	0
7′		−8	$\boxed{-4}$	11	−7	−8	
II	M_2 / M_2^{-1}	−2	−0,25 | −1	2,75	−1,75	−2	
9	−8	1	0,5	−2,5	2,5	1,5	1
10	−4	−2	0,25	2,25	−3,25	−2,75	−3
11	11	0	1	0	0	1	0
12	−7	0	0	1	0	1	0
10′		$\boxed{0}$	6	4	−7	3	

§ 6.8. Cálculo de los vectores propios con ayuda del método de Danilevski

Sea $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ el vector propio de la matriz de Frobenius F el cual corresponde al valor dado de λ. Entonces $F\mathbf{y} = \lambda\mathbf{y}$, de donde $(F - \lambda E)\,\mathbf{y} = 0$, o bien

$$\begin{bmatrix} f_{11}-\lambda & f_{12} & f_{13} & \cdots & f_{1n} \\ 1 & -\lambda & 0 & \cdots & -0 \\ 0 & 1 & -\lambda & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & & -\lambda \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = 0. \qquad (1)$$

Realizando la multiplicación, obtenemos el sistema para determinar las coordenadas $y_1, y_2, \ldots, y_n$ del vector propio $\mathbf{y}$:

$$\left.\begin{cases} (f_{11}-\lambda)\, y_1+f_{12}y_2+f_{13}y_3+\ldots+f_{1n}y_n=0, \\ \qquad y_1-\lambda y_2 \qquad\qquad\qquad\qquad =0, \\ \qquad\qquad y_2-\lambda y_3 \qquad\qquad\qquad =0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \qquad\qquad\qquad\qquad y_{n-1}-\lambda y_n=0. \end{cases}\right\} \tag{1'}$$

Este sistema de ecuaciones lineales es homogéneo. Con exactitud hasta el coeficiente de proporcionalidad sus soluciones pueden hallarse del modo siguiente. Pongamos $y_n = 1$; entonces obtenemos sucesivamente

$$y_{n-1} = \lambda, \quad y_{n-2} = \lambda y_{n-1} = \lambda^2, \ \ldots, \ y_1 = \lambda^{n-1}.$$

Ahora bien, el vector propio buscado es

$$\mathbf{y} = \begin{bmatrix} \lambda^{n-1} \\ \lambda^{n-2} \\ \vdots \\ 1 \end{bmatrix}. \tag{2}$$

Puesto que la matriz F es semejante a la matriz A, entonces λ es también el valor propio de la matriz A.

Designemos con $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ el vector propio de la matriz A el cual corresponde al valor de λ. Entonces obtenemos

$$\mathbf{x} = M_{n-1}M_{n-2} \ldots M_2M_1\mathbf{y}, \tag{3}$$

donde $M_{n-1}, M_{n-2}, \ldots, M_1$ son las matrices unidad transformadas según el método de Danilevski.

Por ejemplo, la transformación M_1 realizada en $\mathbf{y}$ ofrece

$$M_1\mathbf{y} = \begin{bmatrix} m_{11} & m_{12} & \ldots & m_{1n} \\ 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \sum\limits_{k=1}^{n} m_{1k}y_k \\ y_2 \\ \vdots \\ y_n \end{bmatrix} =$$

$$= \begin{bmatrix} \sum\limits_{k=1}^{n} m_{1k}y_k \\ \lambda^{n-2} \\ \vdots \\ 1 \end{bmatrix}.$$

Por consiguiente, la transformación M_1 cambia sólo la primera coordenada del vector $\mathbf{y}$. Análogamente, la transformación M_2 cambia sólo la segunda coordenada del vector $M_1\mathbf{y}$, etc. Repitiendo este proceso $n-1$ veces, obtenemos el valor propio buscado $\mathbf{x}$ de la matriz A.

Ejemplo. En el ejemplo 1 del § 6.7 hemos mostrado que la matriz

$$A=\begin{bmatrix} -4 & -3 & 1 & 1 \\ 2 & 0 & 4 & -1 \\ 1 & 1 & 2 & -2 \\ 1 & 1 & -1 & -1 \end{bmatrix}$$

con ayuda del método de Danilevski se reduce a la forma de Frrobenius

$$F=\begin{bmatrix} -3 & 7 & 24 & 15 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Calcular el vector propio $\mathbf{x}^{(1)} = (x_1, x_2, x_3, x_4)$ si $\lambda_1 = -1$.

△ Utilicemos la fórmula (3), o sea, $\mathbf{x} = M_3M_2M_1\mathbf{y}$, donde

$$\mathbf{y}=\begin{bmatrix} \lambda^3 \\ \lambda^2 \\ \lambda \\ 1 \end{bmatrix}=\begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix},$$

y tomemos las matrices M_3, M_2, M_1 de la tabla 6.6. Sucesivamente hallamos

$$M_1\mathbf{y}=\begin{bmatrix} -0,167 & 0 & 3,167 & 1,500 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix}=\begin{bmatrix} -1,500 \\ 1 \\ -1 \\ 1 \end{bmatrix},$$

$$M_2M_1\mathbf{y}=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & -4 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} -1,5 \\ 1 \\ -1 \\ 1 \end{bmatrix}=\begin{bmatrix} -1,5 \\ 2 \\ -1 \\ 1 \end{bmatrix};$$

$$\mathbf{x}^{(1)}=M_3M_2M_1\mathbf{y}=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & -1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} -1,5 \\ 2 \\ -1 \\ 1 \end{bmatrix}=\begin{bmatrix} -1,5 \\ 2 \\ 0,5 \\ 1 \end{bmatrix}.$$

De un modo análogo se puede hallar los vectores propios también para los valores λ_2, λ_3, λ_4. ▲

§ 6.9. Determinación del primer valor propio de la matriz con ayuda del método de iteraciones

Con ayuda de los métodos iterativos se puede determinar el primer valor propio (es decir, el mayor en módulo) de la matriz A sin desarrollar el determinante característico.

Sea

$$det\,(A - \lambda E) = 0 \tag{1}$$

la ecuación característica; $\lambda_1, \lambda_2, \ldots, \lambda_n$, *sus raíces que son los* valores propios de la matriz $A = [a_{ij}]$ (donde $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, n$). Supongamos que

$$|\lambda_1| > |\lambda_2| \geqslant |\lambda_3| \geqslant \ldots \geqslant |\lambda_n|,$$

o sea, λ_1 es el mayor en módulo número propio.

Entonces para hallar el valor aproximado de la raíz λ_1 se utiliza el esquema siguiente:

1) se elige arbitrariamente el vector inicial $\mathbf{y}$;

2) se plantean las iteraciones sucesivas:

$$\begin{aligned}
\mathbf{y}^{(1)} &= A\mathbf{y},\\
\mathbf{y}^{(2)} &= A\cdot A\mathbf{y} = A^2\mathbf{y},\\
\mathbf{y}^{(3)} &= A\cdot A^2\mathbf{y} = A^3\mathbf{y},\\
&\ldots\ldots\ldots\ldots\\
\mathbf{y}^{(m)} &= A\cdot A^{m-1}\mathbf{y} = A^{(m)}\mathbf{y},\\
\mathbf{y}^{(m+1)} &= A\cdot A^{(m)}\mathbf{y} = A^{(m+1)}\,\mathbf{y};
\end{aligned}$$

3) se eligen de la sucesión los dos últimos valores $\mathbf{y}^m = A^m\mathbf{y}$ e $\mathbf{y}^{(m+1)} = A^{m+1}\mathbf{y}$; entonces

$$\lambda_1 = \lim_{m\to\infty} \frac{y_i^{(m+1)}}{y_i^{(m)}} \text{ o bien } \lambda_1 \approx \frac{y_i^{(m+1)}}{y_i^{(m)}}\,, \tag{2}$$

donde $y_i^{(m+1)}$ e $y_i^{(m)}$ son las coordenadas correspondientes de los vectores $\mathbf{y}^{(m+1)}$ e $\mathbf{y}^{(m)}$ $(i = 1, 2, \ldots, n)$.

Ahora bien, tomando un número de iteración bastante grande m, se puede con todo grado de precisión calcular la raíz λ_1, mayor en módulo, de la ecuación característica de la matriz. Para determinar esta raíz puede ser utilizada toda coordenada del vector $\mathbf{y}^{(m)}$, en particular, se puede tomar la media aritmética de las relaciones correspondientes para diferentes coordenadas.

Si el vector inicial $\mathbf{y}$ está elegido con poca fortuna la fórmula (2) puede no dar la raíz necesaria e, incluso, en general no tener sentido, es decir, el límite de la relación $y_i^{(m+1)}/y_i^{(m)}$ puede no existir. Lo último se advierte fácilmente por los valores «saltantes» de esta relación. En tales casos es necesario cambiar el vector inicial. En calidad de primer vector propio se puede tomar $\mathbf{y}^{(m+1)}$.

Ejemplo. Hallar el primer valor propio de la matriz

$$A=\begin{bmatrix} 3 & 1 & 0 \\ 1 & 2 & 2 \\ 0 & 1 & 1 \end{bmatrix}$$

(con tres cifras decimales) y el vector propio correspondiente.

△ 1) Elegimos el vector inicial $\mathbf{y}=\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

2) Planteamos $m = 10$ iteraciones:

$$\mathbf{y}^{(1)} = A\mathbf{y}, \quad \mathbf{y}^{(2)} = A^2\mathbf{y}^{(1)}, \ \ldots, \ \mathbf{y}^{(10)} = A^{10}\mathbf{y}.$$

Colocamos los cálculos en la tabla 6.10.

Tabla 6.10

$\mathbf{y}$	$A\mathbf{y}$	$A^2\mathbf{y}$	$A^3\mathbf{y}$	$A^4\mathbf{y}$	$A^5\mathbf{y}$	$A^6\mathbf{y}$	$A^7\mathbf{y}$	$A^8\mathbf{y}$	$A^9\mathbf{y}$	$A^{10}\mathbf{y}$
1	4	17	69	274	1075	4189	16 260	62 973	243 569	941 370
1	5	18	67	253	964	3693	24 193	54 650	210 663	812 585
1	2	17	25	92	345	1309	5 002	19 195	73 845	284 508

3) Terminando las iteraciones en $\mathbf{y}^{(10)} = A^{10}\mathbf{y}$, tenemos para distintas coordenadas:

$$\lambda_1^{(1)} \approx \frac{y_1^{(10)}}{y_1^{(9)}} = \frac{941\,370}{243\,569} = 3{,}865; \quad \lambda_1^{(2)} \approx \frac{y_2^{(10)}}{y_2^{(9)}} = \frac{812\,585}{210\,663} = 3{,}857;$$

$$\lambda_1^{(3)} \approx \frac{y_3^{(10)}}{y_3^{(9)}} = \frac{284\,508}{73\,845} = 3{,}853.$$

4) Calculamos λ_1 como media aritmética $\lambda_1^{(1)}$, $\lambda_1^{(2)}$ y $\lambda_1^{(3)}$

$$\lambda_1 = \frac{\lambda_1^{(1)}+\lambda_1^{(2)}+\lambda_1^{(3)}}{3} = \frac{3{,}865+3{,}857+3{,}853}{3} = 3{,}858.$$

5) En calidad de primer vector propio de la matriz A se puede tomar el vector $\mathbf{y}^{(10)} = A^{10}\mathbf{y} = \begin{bmatrix} 941\ 370 \\ 812\ 585 \\ 284\ 508 \end{bmatrix}$. Normándolo, o sea, dividiendo todos sus coordenadas por la norma del vector, igual a

$$\|\mathbf{y}^{(10)}\|_3 = \sqrt{941\,370^2+812\,585^2+284\,508^2} = 1{,}28\cdot 10^6,$$

obtenemos el primer vector propio de la matriz A el cual pertenece al primer valor propio $\lambda_1 = 3{,}858$:

$$\mathbf{x}^{(1)} \begin{bmatrix} 0{,}74 \\ 0{,}64 \\ 0{,}22 \end{bmatrix} \ ▲$$

§ 6.10. Determinación de los valores propios sucesivos y de los vectores propios que les pertenecen

Supongamos que los valores propios de la matriz A son tales que

$$|\lambda_1| > |\lambda_2| > |\lambda_3| \geqslant \ldots \geqslant |\lambda_n|. \qquad (1)$$

Entonces para determinar λ_2 podemos utilizar las llamadas λ-*diferencias* haciendo uso el valor λ_1 que ya tenemos

$$\Delta_{\lambda_1} A^m \mathbf{y} = A^{m+1}\mathbf{y} - \lambda_1 A^m \mathbf{y}, \quad \Delta_{\lambda_1} A^{m-1}\mathbf{y} = A^m \mathbf{y} - \lambda_1 A^{m-1}\mathbf{y}$$

o bien

$$\Delta_{\lambda_1}\mathbf{y}^{(m)} = \mathbf{y}^{(m+1)} - \lambda_1 \mathbf{y}^{(m)}, \quad \Delta_{\lambda_1}\mathbf{y}^{(m-1)} = \mathbf{y}^{(m)} - \lambda_1 \mathbf{y}^{(m-1)}, \qquad (2)$$

de donde, pasando a las coordenadas de los vectores, obtenemos

$$\lambda_2 \approx \frac{\Delta_{\lambda_1} y_i^{(m)}}{\Delta_{\lambda_1} y_i^{(m-1)}} = \frac{y_i^{(m+1)} - \lambda_1 y_i^{(m)}}{y_i^{(m)} - \lambda_1 y_i^{(m-1)}} \quad (i = 1, 2, \ldots, n). \qquad (3)$$

La fórmula (3) ofrece los valores muy aproximados para λ_2, ya que λ_1 también fue determinado aproximadamente.

Si los módulos de todos los valores propios son distintos entre sí, con ayuda de las fórmulas, análogas a la (3), se puede calcular asimismo los demás valores propios, pero los resultados sucesivos serán todavía menos exactos.

Al calcular λ_2 el número de iteración m ha de tomarse menor que al calcular λ_1 para evitar que se pierda la exactitud durante la sustracción de los números próximos.

Ejemplo. Para la matriz

$$A = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 2 & 2 \\ 0 & 1 & 1 \end{bmatrix}$$

hallar el segundo valor propio λ_2 y el vector propio $\mathbf{x}^{(2)}$ que le pertenece. Realizar los cálculos para $m = 8$ iteraciones con tres cifras decimales.

$\triangle$ Utilicemos la tabla de los valores $A^m\mathbf{y}$ para $m = 7, 8, 9$ (véase la tabla 6.10)

$A^7\mathbf{y}$	$A^8\mathbf{y}$	$A^9\mathbf{y}$
16 260	62 973	243 569
14 193	54 650	210 663
5 002	19 195	73 845

Con ayuda de la fórmula (2) planteamos las λ-diferencias:

$$\Delta_{\lambda_1}\mathbf{y}_i^{(m)} = \mathbf{y}_i^{(m+1)} - \lambda_1 \mathbf{y}_i^{(m)} \quad (i = 1,\ 2,\ 3).$$

Para cada una de las filas se toma el propio valor λ_1. $\lambda_1^{(1)} = 3{,}865$; $\lambda_1^{(2)} = 3{,}857$; $\lambda_1^{(3)} = 3{,}853$. Obtenemos la tabla siguiente:

Tabla 6.11

$A^8\mathbf{y}$	$\lambda_1 A^7\mathbf{y}$	$\Delta_1 A^7\mathbf{y}$	$A^9\mathbf{y}$	$\lambda_1 A^8\mathbf{y}$	$\Delta_{y_1} A^8\mathbf{y}$
62 973	62 845	128	243 569	243 290	179
54 650	54 742	−92	210 663	210 785	−122
19 195	19 272	−77	73 845	73 958	−113

3) Con ayuda de la fórmula (3) para cada fila calculamos λ_2:

$$\lambda_2^{(1)} \approx \frac{\Delta_{\lambda_1} y_1^{(8)}}{\Delta_{\lambda_1}\mathbf{y}_1^{(7)}} = \frac{179}{128} = 1{,}400; \quad \lambda_2^{(2)} \approx \frac{\Delta_{\lambda_1} y_2^{(8)}}{\Delta_{\lambda_1} y_2^{(7)}} = \frac{-122}{-92} = 1{,}326;$$

$$\lambda_2^{(3)} \approx \frac{\Delta_{\lambda_1} y_3^{(8)}}{\Delta_{\lambda_1} y_3^{(7)}} = \frac{-113}{-77} = 1{,}468.$$

4) Determinamos λ_2 como media aritmética de $\lambda_2^{(1)}$, $\lambda_2^{(2)}$ y $\lambda_2^{(3)}$:

$$\lambda_2 = \frac{\lambda_2^{(1)} + \lambda_2^{(2)} + \lambda_2^{(3)}}{3} = \frac{1{,}400 + 1{,}326 + 1{,}468}{3} = 1{,}398.$$

5) En calidad de segundo vector propio tomamos

$$\mathbf{x}^{(2)} = \Delta_{\lambda_1} A^8\mathbf{y} = \begin{bmatrix} 179 \\ -122 \\ -113 \end{bmatrix}.$$

Normándolo, tenemos

$$\mathbf{x}^{(2)} = \begin{bmatrix} 0{,}73 \\ -0{,}50 \\ -0{,}46 \end{bmatrix}.$$

Para la matriz dada se puede hallar el tercer valor propio, conociendo la traza de la matriz: puesto que $\lambda_1 + \lambda_2 + \lambda_3 = \operatorname{Sp} A = 3 + 2 + 1 = 6$, entonces $\lambda_3 \approx 6 - 3{,}858 - 1{,}398 = 0{,}744$. ▲

Ejercicios

1. Utilizando el método de desarrollo inmediato, hallar los polinomios característicos para las matrices siguientes:

$$\text{a) } A=\begin{bmatrix} 2 & 5 & 7 \\ 6 & 3 & 4 \\ 5 & -2 & -3 \end{bmatrix}; \quad \text{b) } A=\begin{bmatrix} 2 & -1 & 3 & -4 \\ 3 & -2 & 4 & -3 \\ 5 & -3 & -2 & 1 \\ 3 & -3 & -1 & 2 \end{bmatrix};$$

$$\text{c) } A=\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}.$$

2. Determinar los números característicos y vectores propios de las matrices

$$\text{a) } A=\begin{bmatrix} 3 & 4 \\ 5 & 2 \end{bmatrix}, \quad \text{b) } A=\begin{bmatrix} 5 & 6 & -3 \\ -1 & 0 & 1 \\ 1 & 2 & -1 \end{bmatrix}.$$

3. Con ayuda del método de Krylov desarrollar los determinantes característicos de las matrices siguientes:

$$\text{a) } A=\begin{bmatrix} 1 & 2 & -4 \\ 3 & -1 & 2 \\ 1 & 0 & -1 \end{bmatrix}; \quad \text{b) } A=\begin{bmatrix} 2 & -2 & 1 & 3 \\ 4 & -1 & -2 & -3 \\ 1 & 2 & -5 & 4 \\ 1 & -1 & 2 & -4 \end{bmatrix};$$

$$\text{c) } A=\begin{bmatrix} 1 & -2 & 1 & -2 \\ 2 & -1 & 2 & -1 \\ 1 & 1 & -2 & -2 \\ -2 & -2 & 1 & 1 \end{bmatrix}.$$

4. Haciendo uso del método de Le Verrier — Faddéev, desarrollar los determinantes característicos para las matrices siguientes:

$$\text{a) } A=\begin{bmatrix} 5 & -4 & 3 & 0 \\ 2 & 1 & -2 & 1 \\ 1 & 1 & 1 & -1 \\ 2 & 1 & -2 & 0 \end{bmatrix}; \quad \text{b) } A=\begin{bmatrix} 2 & -4 & 3 & -2 \\ 1 & 0 & -5 & -3 \\ 1 & 1 & -2 & -2 \\ 0 & 1 & -1 & -1 \end{bmatrix}.$$

5. Aplicando el método de Le Verrier — Faddéev, hallar las matrices inversas para las dadas en el ejercicio 4.

6. Desarrollar el determinante característico de la matriz

$$A = \begin{bmatrix} 2 & -1 & 2 \\ 5 & -3 & 3 \\ -1 & 0 & -2 \end{bmatrix}$$

y calcular el vector propio con ayuda del método de Le Verrier — Faddéev.

7. Haciendo uso del método de iteraciones, calcular los valores propios primero y segundo de la matriz

$$A = \begin{bmatrix} 4 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

CAPÍTULO VII

Interpolación y extrapolación

§ 7.1. Función y métodos de su representación

En la actividad práctica nos encontramos constantemente con la necesidad de revelar las formas de relación en los procesos y fenómenos y con la necesidad de describirlas matemáticamente.

Consideremos las formas de relación para las cuales cierta magnitud y que caracteriza el proceso depende de un conjunto de magnitudes no relacionadas entre sí $x_1, x_2, \ldots, x_n$ de un modo tal que a cada colección $(x_1, x_2, \ldots, x_n)$ corresponda el único valor de la magnitud y. Tal correspondencia unívoca de la magnitud y al conjunto de las variables independientes $x_1, x_2, \ldots, x_n$ se llama *dependencia funcional* y la misma variable y, *función* de las variables $x_1, x_2, \ldots, x_n$ lo que se escribe formalmente así:

$$y = f(x_1, x_2, \ldots, x_n).$$

Ahora bien, la expresión $y = x_1^2 + 3\sqrt{x_2} + x_1x_3^2$ es función de tres variables.

Si la magnitud y es función de una sola variable independiente x, esta relación puede ser representada en la forma siguiente:

$$y = f(x).$$

Por ejemplo, el área de un círculo S es función de la variable independiente, o sea, del radio de círculo R, es decir, $S = f(R)$; la forma concreta de esta función $S = \pi R^2$. El volumen de una figura es ya la función de tres dimensiones: $V = f(x_1, x_2, x_3)$ y según la forma de la figura esta relación funcional se concretiza respectivamente.

Del curso de análisis matemático se conocen tres métodos de representación de las dependencias funcionales: 1) analítico; 2) gráfico; 3) tabular.

El método más cómodo de representar la dependencia funcional $y = f(x)$ es el **analítico**, ya que muestra directamente las operaciones, y la sucesión de realizarlas, con la variable independiente x para que se obtenga el valor correspondiente de la magnitud y.

Así, por ejemplo, como resultado del tratamiento matemático se puede obtener la siguiente dependencia analítica de los créditos monetarios en la agricultura para los valores mercantiles y materiales en función de los gastos para el ganado bovino:

$$y = 51,0203 + 0,1059x,$$

donde y son los créditos para los valores mercantiles y materiales; x, los gastos para el ganado bovino.

Otro ejemplo de la dependencia analítica: en el movimiento uniformemente acelerado el espacio recorrido y el tiempo están ligados por la relación

$$s = vt + 0,5at^2.$$

La ventaja del método analítico de representación consiste en la posibilidad de obtener los valores de y para todo argumento fijo x con toda precisión.

Entre los inconvenientes de este método figura la necesidad de realizar toda la sucesión de los cálculos; además, el método analítico no está visualizado.

Las desventajas indicadas del método analítico se eliminan al utilizar el **método gráfico** de representación de la función $y = f(x)$.

Se llama *gráfico* de la función dada $y = f(x)$ el conjunto de los puntos del plano xOy cuyas coordenadas satisfacen la ecuación $y = f(x)$.

El método tabular de representación de las funciones está difundido en la técnica, la física, la economía, las ciencias naturales (y con más frecuencia aparece como resultado de un experimento).

Por ejemplo, supongamos que como resultado de una prueba está obtenida la resistencia óhmica R de una barra de cobre en función de la temperatura $t°$ en forma de la tabla siguiente:

R	77,80	79,75	80,80	82,35	83,90	85,10
$t°$	25,0	30,1	36,0	40,0	45,1	50,0

En este experimento el valor de la resistencia óhmica de la barra de cobre cambia durante las oscilaciones de la temperatura y es una variable dependiente.

La ventaja del método tabular de representar una función consiste en que para cada valor de la variable independiente colocada en la tabla se puede hallar inmediatamente, sin mediciones o cálculos algunos, el valor correspondiente de la función.

El inconveniente del método tabular consiste en que no es posible representar continuamente toda la función, es decir, siempre hay tales valores de la variable independiente que faltan en la tabla.

§ 7.2. Tablas matemáticas

Entre las funciones que constantemente figuran en las matemáticas hay muchas tales cuyo cálculo cuesta bastante trabajo, a pesar de ser ellas sencillas. Tales funciones se someten a la tabulación, o sea, se representan en forma de las tablas matemáticas.

Tienen la más vasta aplicación las tablas de funciones de una variable. Entre ellas figuran las de los números inversos, de los cuadrados y cubos de los números, de las raíces cuadradas y cúbicas, las tablas de los logaritmos de las funciones trigonométricas, las tablas de la función exponencial y de otras funciones elementales. Hay tablas de las funciones de dos variables y de una cantidad mayor de estas últimas. Como ejemplo de la tabla de las funciones de dos variables puede servir la de los productos de dos números.

La tabla no es sino una colección de los valores de la función para una sucesión de los valores de los argumentos x_0, x_1, x_2, ..., x_n. Debe contener tal colección de los valores del argumento que para todos valores de éste, distintos de x_0, x_1, x_2, . . ., x_n se pueda obtener el valor de la función con el grado necesario de precisión.

Las características fundamentales de las tablas son: 1) denominación de las funciones cuyos valores ellas expresan; 2) volumen; 3) paso; 4) cantidad de cifras de la función a tabular; 5) cantidad de entradas.

En calidad de *denominación de la función* cuyos valores numéricos se dan en la tabla sirve la expresión analítica de esta función, por ejemplo, $\operatorname{sen} x$, $\log x$, e^x, etc.

El volumen de la tabla se predetermina por los valores inicial y final del argumento. Así, por ejemplo, el volumen de la tabla $y = \operatorname{sen} x$ contiene los valores del argumento entre $0°0'$ y $90°$.

Casi para todas las funciones a tabular los valores del argumento en la tabla forman la progresión aritmética cuya diferencia h se llama *paso de la tabla*. Ahora bien,

$$h = x_i - x_{i-1} = \text{const} \quad (i = 1, 2, \ldots, n).$$

Entonces

$$x_i = x_0 + ih \quad (i = 0, 1, \ldots, n).$$

A título de ilustración consideremos un fragmento de la tabla de conversión de los radianes en grados y de los grados en radianes (tabla 7.1)

En las primeras dos columnas de la tabla mencionada como variable independiente sirve la medida en radianes, mientras que los grados se consideran su función. Lo mismo es válido también para las columnas 3 y 4. En calidad del paso de la tabla aquí se elige $h = 0{,}01$ radianes.

Comenzando con la columna 5 se examina la función inversa a la dada, donde como variable independiente se eligen grados (o minutos y la medida en radianes es, respectivamente, la función de los grados

Tabla 7.1

Radianes	Grados	Radianes	Grados	Grados	Radianes
(1)	(2)	(3)	(4)	(5)	(6)
0,20	11,459	0,70	40,107	20	0,34907
21	12,032	71	40,680	21	36652
22	12,605	72	41,253	22	38397
23	13,178	73	41,826	23	40143
24	13,751	74	42,399	24	41888

Continuación

Grados	Radianes	Minutos	Radianes	Minutos	Radianes
(7)	(8)	(9)	(10)	(11)	(12)
70	1,22175	20	0,00582	50	0,01454
71	23918	21	00621	51	01484
72	25662	22	00640	52	01513
73	27409	23	00669	53	02542
74	29151	24	00698	54	01571

(o de los minutos). El paso de esta parte de la tabla es igual a un grado (en las columnas 5 y 7) y a un minuto (en las columnas 9 y 11).

En las tablas de guía se utiliza asimismo un paso compuesto de dos niveles. Por la vertical en la columna se dan los valores del argumento con un paso relativamente grande h^*: $x_i = x_0 + ih^*$ $(i = 0, 1, \ldots, n)$ y los valores correspondientes de la función $y_i = f_i = f(x_i)$. Por la horizontal en la primera fila se colocan los valores del argumento con un paso más pequeño h que de ordinario es igual a una décima parte del paso grande: $h = 0{,}1h^*$. En la segunda fila y en las sucesivas se ponen los valores de la función para el argumento que es igual a la suma de los valores x_i que están en la misma fila y en la misma columna en la intersección de las cuales se escribe el valor de la función. Así en la tabla 7.2 se da un fragmento de la tabla de las raíces cúbicas. Del fragmento mencionado de la tabla no es difícil determinar el paso grande por la vertical, igual a $h^* = 1$, y el paso pequeño por la horizontal, igual a $h = 0{,}1$.

Generalmente el paso de la tabla se expresa por una inidad de cualquier orden (más raramente por dos o cinco unidades de un orden determinado). Por ejemplo, en las tablas de los cuadrados y los cubos, en las de las raíces cuadradas y cúbicas y en las de los logaritmos el paso grande $h^* = 1$, en las tablas de los logaritmos naturales y en las de los números inversos el paso grande h^* es igual a 0,1 (véase la tabla 7.2).

Tabla 7.2

	0	1	2	3	4	5	6	7	8	9
60	8,43433	43901	44369	48836	45303	45769	46235	46700	47165	47629
61	8,48093	48556	49018	49481	49942	50403	50954	51324	51784	52243
62	8,52702	53160	53618	54075	54542	54988	55444	55899	56354	56808
63	8,57262	57715	58168	58620	59062	59524	59975	60425	60875	61325
64	8,61774	62222	62670	63118	63566	64012	64459	64904	55350	65795

Tabla 7.3

	0′	6′	12′	18′	24′	30′	36′	42′	48′	54′			1′	2′	3′	4′	5′
65°	0,9063	9070	9078	9085	9092	9100	9107	9114	9121	9128	0,9335	24°	1	2	4	5	6
66°	0,9135	9143	9150	9157	9164	9171	9178	9184	9191	9198	0,9205	23°	1	2	3	5	6
67°	0,9205	9212	9219	9225	9232	9239	9245	9252	9259	9265	0,9272	22°	1	2	3	4	6
68°	0,9272	9278	9285	9291	9298	9304	9311	9317	9323	9330	0,9336	21°	1	2	3	4	5
69°	0,9336	9342	9348	9354	9361	9367	9383	9379	9385	9391	0,9397	20°	1	2	3	4	5

Si examinamos la tabla de los senos (tabla 7.3), veremos que en esta tabla en calidad de paso grande se elige un grado y el paso pequeño es igual a seis minutos.

La siguiente característica de las tablas es la *cantidad de cifras* de la función a tabular, ya que los valores de la función $y = f(x)$ para los valores tabulados del argumento en las tablas matemáticas y los resultados de mediciones en las tablas técnicas son magnitudes aproximadas.

Para los cálculos prácticos manuales las más aplicables son «Tablas matemáticas de cinco cifras decimales» de B. I. Segal y K. A. Semendiáev, «Guía de matemáticas» de I. N. Bronshtein y K. A. Semendiáev, etc.

En las tablas se colocan sólo las cifras justas del valor numérico de la función: esto quiere decir que el error no supera cinco unidades del primer orden suprimido. En este caso los valores de la función para todos los valores de x dados en la tabla se determinan con error absoluto igual. La exactitud con que se dan en la tabla los valores de la función se llama *exactitud de la tabla*. A veces en distintas partes de la tabla la exactitud puede ser diferente.

En algunos casos al trabajar con las tablas es necesario conocer las diferencias de los valores vecinos de la función dados en la tabla, o sea, $y_{i+1} - y_i$. Estas diferencias se llaman *diferencias finitas de primer orden* y se escriben, a veces, en la columna de la función en el intervalo entre los valores de esta última que participan en la formación de la diferencia finita correspondiente. Las diferencias se escriben en las unidades del último orden sin ceros delante de las cifras significativas y sin coma. Por ejemplo, en la tabla

x	sen x	
1,000	0,84147	54
1,001	0,84201	

la diferencia finita 0,00054 = 0,84201 — 0,84147 está escrita entre los valores respectivos de la función.

La siguiente característica importante de las tablas es la *cantidad de entradas* en la misma. Es igual al número de argumentos de la función. Así, las tablas para las dependencias funcionales $y = f(x)$ son tablas con una entrada. Entre ellas figuran las tablas 7.1, 7.2 y 7.3 anteriormente dadas.

La tabulación de la función de dos variables $z = f(x, y)$ conduce a la tabla con dos entradas. Entre semejantes tablas tienen amplia aplicación práctica las tablas de multiplicación.

En la tabla 7.4 el multiplicando de tres órdenes está escrito en la columna izquierda y el multiplicador de un orden, en la fila superior de la tabla. El paso de ambas entradas en la tabla es igual a la unidad. Para obtener el producto de un número de tres cifras por un número de una cifra basta encontrar la fila en cuya primera columna está puesto el multiplicando y elegir la columna en la cual está situado el multiplicador. En la intersección de la fila y columna halladas está precisamente el producto buscado.

Tabla 7.4

1	2	3	4	5	6	7	8	9
541	1082	1623	2164	2705	3246	3787	4328	4869
542	1084	1626	2168	2710	3252	3794	4336	4878
543	1085	1629	2172	2715	3258	3801	4344	4887
544	1088	1632	2176	2720	3260	3808	4352	4896
545	1090	1635	2180	2725	3270	3815	4360	4905

Ejemplo 1. Supongamos que se necesita multiplicar 543 por 8. En la tabla 7.4 encontramos la fila que contiene 543 y la columna cuyo número es 8. En su intersección leemos el número 4344 lo que es precisamente el producto buscado.

Para multiplicar números polidígitos el multiplicando se divide en partes que contienen tres cifras, como máximo, y a cada una de estas partes se aplica el método indicado.

Ejemplo 2. Supongamos que se necesita multiplicar 541 544 por 37. Dividimos el número 541 544 en dos partes de tres órdenes:

541 y 544. Sucesivamente multiplicamos cada parte por tres decenas y por 7 unidades y sumamos los productos parciales obtenidos:

$$\begin{array}{r} 541\times 30=16\,230 \\ 541\times 7=\ \ 3\,787 \\ \hline 20\,017 \end{array} \qquad \begin{array}{r} 544\times 30=16\,320 \\ 544\times\ 7=\ \ 3\,808 \\ \hline 20\,128 \end{array}$$

El primer producto parcial se pone en tres órdenes a la izquierda respecto al segundo y se realiza la sumación:

$$\begin{array}{r} 20\,017 \\ +\quad 20\,128 \\ \hline 20\,037\,128 \end{array}$$

Éste es precisamente el resultado buscado.

§ 7.3. Conceptos principales de la teoría de aproximación de las funciones

La teoría y la práctica de *aproximación* se aplican al resolver muchos problemas prácticos.

Supongamos, por ejemplo, que en el proceso de cierto experimento en instantes discretos de tiempo $x_0, x_1, \ldots, x_N$ han sido obtenidos los valores $f_0, f_1, \ldots, f_N$ de cierta variable $f(x)$. Se necesita reconstruir la función $f(x)$ para otros $x \neq x_i$ $(i = 0, 1, \ldots, N)$. Un problema semejante surge al calcular muchas veces en el ordenador una misma función compuesta f en diferentes puntos. En vez de esto con frecuencia puede ser racional calcular la función f en un pequeño número de puntos característicos x_i y en los demás puntos calcular sus valores siguiendo una regla más sencilla al utilizar la información sobre los valores ya conocidos $f_i = f(x_i)$.

En calidad de otros ejemplos difundidos de aproximación de las funciones sirven los problemas de determinación de la derivada $f'(x)$ y de la integral $\int\limits_a^b f(x)\,dx$ según los valores dados f_i.

Por último, al componer los algoritmos de los programas estándares para el cálculo de las funciones elementales y especiales aparece una vez más el problema de aproximación de las funciones.

El enfoque clásico de resolución de semejantes problemas consiste en que, utilizando la información disponible sobre la función f, se considera otra función φ que es próxima, en cierto sentido, a f y permite realizar con ésta la operación correspondiente y obtener la estimación del error de tal «sustitución analítica».

En el proceso de realización numérica de este enfoque es necesario examinar las siguientes cuatro cuestiones fundamentales:

1. La cuestión acerca de la información disponible respecto a la función f, o sea, acerca de la forma en la cual está representada la función f.

2. La cuestión acerca de la clase de las funciones que aproximan, o sea, acerca de qué funciones φ será aproximada la función f.

3. La cuestión acerca de la proximidad de las funciones aproximable y aproximante, o sea, acerca de la elección del criterio de aceptación al cual debe satisfacer la función φ.

4. La cuestión acerca del error, o sea, acerca de la determinación de la diferencia entre los valores exacto y aproximado.

En la cuestión sobre la información respecto a la función f se distinguen dos casos principales: ora la función está representada analíticamente, ora en la forma de una tabla. El método gráfico de representar la función se refiere al primer caso o al segundo según el problema concreto. A continuación consideraremos sobre el segmento $[a, b]$ las funciones continuas $f(x)$, junto con una cantidad suficiente de sus derivadas, determinadas por sus valores $f_i = f(x_i)$ en los nodos x_i de la red dada

$$\Lambda_N: \{a \leqslant x_0 < x_1 < \ldots < x_N \leqslant b\}. \tag{1}$$

En la cuestión sobre la clase de las funciones aproximantes es necesario guiarse por dos factores principales. En primer lugar, la función aproximante debe reflejar las particularidades características de la aproximable y, en segundo lugar, debe ser bastante cómoda en el tratamiento, o sea, al realizar con ella las operaciones necesarias.

En el análisis numérico tienen gran aplicación tres grupos de las funciones aproximantes. El primero está formado por las funciones que tienen la forma $1, x, \ldots, x^n$ cuyas combinaciones lineales generan la clase de todos los polinomios de grado no superior a n. El segundo grupo está constituido por las funciones trigonométricas sen $a_i x$ y cos $a_i x$ que generan las series de Fourier y la integral de Fourier. Por último, el tercer grupo se compone de las funciones exponenciales $e^{a_i x}$ que determinan los fenómenos del tipo descomposición y acumulación los cuales se encuentran frecuentemente en situaciones reales.

A continuación examinaremos con más detalles la aproximación polinomial, o sea, tomaremos en calidad de la función aproximante el polinomio de cierto grado n.

En este caso la función aproximante suele designarse $P_n(x)$ y tiene la forma

$$\varphi(x) \equiv P_n(x) = \sum_{k=0}^{n} a_k x^k. \tag{2}$$

La cuestión sobre el criterio de aceptación consiste de hecho en determinar de cierto modo la «distancia» entre las funciones aproximable y aproximante y luego de toda la clase de las funciones aproximantes elegir aquélla para la cual esta «distancia» es mínima.

Uno de los criterios de aceptación difundidos es el de Chébyshev fundado en el concepto de distancia como magnitud máxima de desviación de la función φ respecto a la función f en los nodos x_i:

$$\rho_1 = \max_{0 \leqslant i \leqslant N} |f(x_i) - \varphi(x_i)|. \qquad (3)$$

Presenta mayor interés el caso particular cuando para la función aproximante la distancia $\rho_1 = 0$. Esto quiere decir que para la función tabulada $y = f(x)$ representada por sus valores $y_i = f_i = f(x_i)$:

x_0	x_1	$\dots$	x_N
y_0	y_1	$\dots$	y_N

se necesita (construir la función aproximante $\varphi(x)$ que coincida en los nodos x_i con los valores de la función dada $y = f(x)$, o sea tal que $\varphi(x_i) = y_i$.

Tal método de aproximación basado en el criterio de coincidencia de f y φ en los nodos x_i se llama *interpolación*. Si el argumento x,

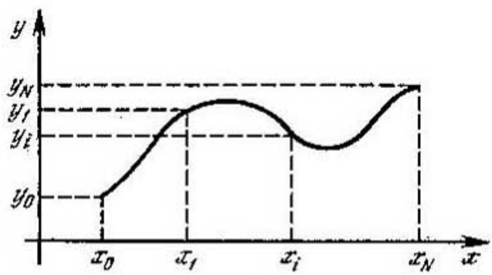

Fig. 7.1

para el cual se determina el valor aproximado de la función, pertenece al segmento $[x_0, x_N]$, el problema de determinación del valor de la función en el punto x se denomina *interpolación en sentido estrecho*. En cambio, si el argumento x está fuera del segmento $[x_0, x_N]$, el problema planteado se nombra *extrapolación*.

El problema de interpolación para la función de una variable $y = f(x)$ significa geométricamente la construcción de la curva que pasa por los puntos con coordenadas $(x_0; y_0)$, $(x_1; y_1)$, ..., $(x_N; y_N)$ sobre el plano xOy (fig. 7.1).

Vamos a citar un ejemplo más de criterio de aceptación. Introduzcamos el concepto de distancia entre las funciones f y φ como suma de cuadrados de sus desviaciones en los puntos nodales:

$$\rho = \sum_{i=0}^{N} [f(x_i) - \varphi(x_i)]^2. \qquad (6)$$

Elijamos ahora en calidad de función aproximante aquélla para la cual ρ es el mínimo. Es racional utilizar este criterio en caso de una gran cantidad de información dada con pequeña exactitud; el método de aproximación fundado en dicho criterio suele llamarse

método de cuadrados mínimos.. Entre las ventajas de este método conviene mencionar la sencillez y la armonía de su teoría matemática.

Por fin, la última cuestión — acerca de la exactitud de la solución que se obtiene — es, en muchos aspectos, la fundamental. En efecto en resumidas cuentas la calidad del método se determina, en primer lugar, por la rapidez con que se obtiene la solución con exactitud requerida o bien, como se dice de otra manera, por la velocidad de convergencia. Por eso es comprensible que la elección de los puntos nodales, de la clase de funciones aproximantes y del criterio de aceptación debe ser subordinada a la única cuestión, o sea, a la de exactitud requerida.

A primera vista la cuestión sobre la exactitud de la solución a obtener parece bastante simple: es necesario que la solución aproximada se distinga de la exacta en valor prefijado ε como máximo. Sin embargo, la cuestión acerca de la posibilidad de aproximar tan exactamente como se quiera la función f, cuestión que depende de los «parámetros» anteriormente citados (nodos x_i, clase de funciones φ, criterio de aceptación de f y φ), en el caso general, queda pendiente y se somete a la investigación para cada proceso de aproximación concreto.

§ 7.4. Interpolación con ayuda de los polinomios

Consideremos más detalladamente el problema de interpolación de la función f con ayuda de los polinomios algebraicos.

En este caso la función aproximante φ suele designarse $P_n(x)$ y tiene la forma

$$\varphi(x) \equiv P_n(x) = \sum_{k=0}^{n} a_k x^{n-k}. \tag{1}$$

La elección del valor concreto n se determina en gran parte por las propiedades de la función aproximable, por la exactitud requerida, así como por los nodos de interpolación. A continuación veremos que en la elección de la magnitud n ejerce influencia esencial también el proceso de cálculo que introduce en el resultado un error adicional.

En calidad de criterio de aceptación se toma, evidentemente, la condición de coincidencia de f y φ en los puntos nodales.

Es natural suponer que para determinar unívocamente $n+1$ coeficientes a_k del polinomio P_n es necesario exigir que coincidan f y P_n en el punto nodal $n+1$:

$$f(x_i) = P_n(x_i) \quad (i = 0, 1, \ldots, n). \tag{2}$$

El polinomio $P_n(x)$ que satisface las condiciones (2) se llama *polinomio interpolador*. Para subrayar que este polinomio depende de la función f, se designa frecuentemente $P_n(f, x)$.

En el caso en que es necesario calcular el valor de la función $f(x)$ en un punto x^*, por *error de interpolación* Δ_1 se entiende el valor absoluto de la diferencia entre los valores exacto y aproximado:

$$\Delta_1 = | f(x^*) - P_n(x^*) |. \tag{3}$$

En cambio, en el caso en que la interpolación se lleva a cabo sobre todo el segmento $[a, b]$, por error se toma la desviación máxima del polinomio P_n respecto a la función f en el segmento dado:

$$\Delta_1 = \max_{[a,b]} |f(x) - P_n(x)|.$$

Ahora bien, considermos el siguiente problema de interpolación En la red $\Lambda_n : a \leqslant x_0 < x_1 < \ldots < x_n \leqslant b$ en los nodos x_i se dan los valores $f_i = f(x_i)$ $(i = 0, 1, \ldots, n)$ de la función f. Se necesita construir el polinomio interpolador P_n que coincida con f en los nodos x_i y estimar el error Δ_1.

La existencia del polinomio interpolador y su unicidad se deducen del teorema siguiente.

Teorema. *Supongamos que*: 1°) *sobre el segmento* $[a, b]$ *se da la red* $\Lambda_n : a \leqslant x_0 \leqslant x_1 < \ldots < x_n \leqslant b$; 2°) *se dan los números arbitrarios* c_i $(i = 0, 1, \ldots, n)$. *Entonces existe el polinomio* P_n *de grado no superior a* n *que toma en los nodos* x_i *los valores dados de* c_i *y este polinomio es único.*

□ De las condiciones para determinar los coeficientes desconocidos a_k del polinomio P_n obtenemos el sistema de ecuaciones algebraicas

$$a_0 x_i^n + a_1 x_i^{n-1} + \ldots + a_n = c_i \ (i = 0, 1, \ldots, n). \tag{4}$$

El determinante de este sistema

$$W = \begin{vmatrix} x_0^n & x_0^{n-1} & \ldots & x_0 & 1 \\ x_1^n & x_1^{n-1} & \ldots & x_1 & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ x_n^n & x_n^{n-1} & \ldots & x_n & 1 \end{vmatrix} \tag{5}$$

es el determinante de Vandermonde que, como se sabe del álgebra, es distinto del cero si se cumple la condición $x_i \neq x_j$ para $i \neq j$. Esta condición se cumple, evidentemente, para la red en cuestión Λ_n. Por consiguiente, el sistema (4) tiene la única solución (la única colección de los coeficientes a_k). ■

Es evidente que si en calidad de números c_i se asignan los valores f_i de la función f en los nodos x_i, obtenemos la afirmación sobre la existencia del polinomio interpolador $P_n(f, x)$ y sobre su unicidad.

Los coeficientes a_k del polinomio interpolador (1) pueden ser determinados poniendo en el sistema (4) $c_i = f_i$ y resolviéndolo, por ejemplo, con ayuda de las fórmulas de Cramer

$$a_k = \Delta_k / W.$$

Aquí Δ_k es el determinante que se obtiene de W reemplazando la columna de los términos que contienen el grado $n - k$ de x_i $(i = 0,$

$1, \ldots, n$) por la columna f_i de los términos independientes del sistema (4):

$$\Delta_k = \begin{vmatrix} x_0^n & \ldots & x_0^{n-k+1} & f_0 & x_0^{n-k-1} & \ldots & 1 \\ x_1^n & \ldots & x_1^{n-k+1} & f_1 & x_1^{n-k-1} & \ldots & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ x_n^n & \ldots & x_n^{n-k+1} & f_n & x_n^{n-k-1} & \ldots & 1 \end{vmatrix}. \tag{7}$$

Sustituyendo los valores obtenidos de a_k en la igualdad (1), llegamos a una nueva forma de representación del polinomio interpolador $P_n(f, x)$:

$$\begin{vmatrix} P_n & 1 & x & \ldots & x^n \\ f_0 & 1 & x_0 & \ldots & x_0^n \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ f_n & 1 & x_n & \ldots & x_n^n \end{vmatrix} = 0. \tag{8}$$

Nótese que en la práctica suelen utilizarse los polinomios interpoladores de primer grado y de segundo grado. En este caso se trata de *interpolación lineal y cuadrática.*

Ejemplo. Con ayuda de los nodos x_0, x_1, x_2 y los valores correspondientes de la función f_0, f_1, f_2 construir el polinomio interpolador representándolo en la forma de la combinación lineal de los valores f_i ($i = 0, 1, 2$).

△ De acuerdo con la fórmula (8) tenemos

$$\begin{vmatrix} P_2 & 1 & x & x^2 \\ f_0 & 1 & x_0 & x_0^2 \\ f_1 & 1 & x_1 & x_1^2 \\ f_2 & 1 & x_2 & x_2^2 \end{vmatrix} = 0.$$

Desarrollando el determinante según los elementos de la primera columna, obtenemos

$$P_2 \begin{vmatrix} 1 & x_0 & x_0^2 \\ 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \end{vmatrix} - f_0 \begin{vmatrix} 1 & x & x^2 \\ 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \end{vmatrix} + f_1 \begin{vmatrix} 1 & x & x^2 \\ 1 & x_0 & x_0^2 \\ 1 & x_2 & x_2^2 \end{vmatrix} - f_2 \begin{vmatrix} 1 & x & x^2 \\ 1 & x_0 & x_0^2 \\ 1 & x_1 & x_1^2 \end{vmatrix} = 0.$$

Teniendo en cuenta que

$$\begin{vmatrix} 1 & x_i & x_i^2 \\ 1 & x_j & x_j^2 \\ 1 & x_k & x_k^2 \end{vmatrix} = (x_j - x_i)(x_k - x_i)(x_k - x_j),$$

finalmente hallamos

$$P_2(x) = f_0 \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} + f_1 \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} + \\ + f_2 \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)}. \quad ▲$$

§ 7.5. Error de los procesos de interpolación

Supongamos que la función f es aproximada por el polinimio interpolador, o sea,

$$f(x) = P_n(x) + R_n(x), \tag{1}$$

donde $R_n(x)$ es el término residual de la fórmula de interpolación

$$f(x) \approx P_n(x). \tag{2}$$

El término residual depende de muchos factores: de las propiedades de la función f, de los parámetros de interpolación y de la posición del punto de interpolación. Por eso el estudio de $R_n(x)$ es un problema difícil. En este caso es necesario, ante todo, responder a la pregunta: ¿qué es la medida numérica del error? Si el punto de interpolación x^* es fijo, por medida del error es natural tomar la magnitud $\Delta_1 = |R_n(x^*)|$. En cambio, si el punto x^* no está conocido de antemano y la interpolación se lleva a cado sobre el segmento $[a, b]$, es racional por medida del error tomar la magnitud

$$\Delta_1 = \max_{[a,b]} |R_n(x)|. \tag{3}$$

Según el problema concreto pueden ser elegidas también otras medidas del eror.

Por regla general, la estimación de la medida del error no se realiza para una función tomada por separado sino para toda una clase de funciones que poseen ciertas propiedades comunes.

Deduzcamos la expresión explícita para estimar el error (3) de la fórmula de interpolación (2) para la clase de funciones $C^{n+1}(a, b)$ que tienen sobre el segmento $[a, b]$ una derivada continua de orden $n + 1$.

Con este fin vamos a demostrar el teorema siguiente.

Teorema. *Supongamos que*: 1°) *los nodos* x_i $(i = 0, 1, \ldots, n)$ *son distintos y junto con* x^* *pertenecen al segmento* $[a, b]$; 2°) *la función* f *tiene sobre* $[a, b]$ *una derivada continua de orden* $n + 1$. *Entonces existe tal punto* $\xi \in (a, b)$ que

$$R_n(x^*) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{i=0}^{n} (x^* - x_i). \tag{4}$$

□ Nótese que si x^* coincide con uno de los nodos, la relación (4) se cumple, ya que su primer miembro y su segundo miembro son iguales a cero. Por eso suponemos luego que $x^* \neq x_i$ $(i = 0, 1, \ldots, n)$. Examinemos la función auxiliar

$$\Psi(x) = f(x) - P_n(x) - k \prod_{i=0}^{n} (x^* - x_i),$$

donde k es la constante elegida de un modo tal que la función Ψ se anule para $x = x^*$, o sea,

$$\Psi(x^*) = 0 = f(x^*) - P_n(x^*) - k \prod_{i=0}^{n} (x^* - x_i).$$

De aquí

$$k = \frac{f(x^*) - P_n(x^*)}{\prod_{i=0}^{n} (x^* - x_i)} = \frac{R_n(x^*)}{\prod_{i=0}^{n} (x^* - x_i)}. \tag{6}$$

En virtud de tal elección de k la función Ψ sobre el segmento $[a, b]$ se anula $n + 2$ veces, como mínimo, en los puntos $x_0, x_1, \ldots, x_n, x^*$. Entonces, utilizando el teorema de Rolle, se puede afirmar que la derivada Ψ' sobre el intervalo (a, b) se anula, al menos, $n + 1$ veces y la derivada Ψ'' se anula n veces, por lo menos, etc. hasta la derivada $\Psi^{(n+1)}$ que se anulará en un punto, como mínimo. Supongamos que éste es el punto $\xi \in (a, b)$.

Ahora, derivando los miembros segundo y primero de la relación (5) $n + 1$ veces respecto a x y luego suponiendo $x = \xi$, en el primer miembro se obtiene el cero, ya que $\Psi^{(n+1)}(\xi) = 0$. El primer sumando del segundo miembro ofrece el valor de la derivada en el punto ξ: $f^{(n+1)}(\xi)$. El segundo sumando del segundo miembro da el cero como derivada de orden $n + 1$ del polinomio cuyo grado no supera n. El tercer sumando es el producto de la constante k por el polinomio de grado $n + 1$ con el coeficiente superior igual a 1; la derivada de orden $n + 1$ de este polinomio, como se sabe, es igual a $(n + 1)!$. Ahora bien, sumando todo lo dicho, tenemos

$$0 = f^{(n+1)}(\xi) - k(n + 1)!$$

Sustituyendo aquí en vez de k su expresión (6), obtenemos la relación requerida (4). ■

Supongamos ahora que, para precisar,

$$|f^{(n+1)}(x)| \leqslant M_{n+1}; \quad x \in [a, b]. \tag{7}$$

Utilizando esta limitación y el teorema que acabamos de demostrar, llegamos a la siguiente estimación del error para el punto fijo x^*:

$$\Delta_1 = |R_n(x^*)| \leqslant \frac{M_{n+1}}{(n+1)!} \prod_{i=0}^{n} |x^* - x_i|. \tag{8}$$

Ahora no es difícil construir para la red fija Λ_n la estimación $|R_n|$ uniforme en todo el segmento $[a, b]$. Así pues,

$$\Delta_1 = \max_{[a, b]} |R_n(x)| \leqslant \frac{M_{n+1}}{(n+1)!} \max_{[a, b]} |\omega_n(x)|, \tag{9}$$

donde $\omega_n(x) = \prod_{i=0}^{n} (x - x_i)$.

Ejemplo 1. Obtener sobre el segmento $[-1, 1]$ la estimación uniforme de la desviación de la función $f = 1 - \cos(\pi x/2)$ respecto a su polinomio interpolador construido por los nodos $x_i = -1 + \frac{2i}{n}$ $(i = 0, 1, \ldots, n;\ n = 2, 3, 4)$.

△ Nótese, ante todo, que para la función en cuestión sobre el segmento dado $M_{n+1} = (\pi/2)^{n+1}$. Por eso en virtud de la estimación (9) la resolución del problema se reduce a la estimación de la magnitud $\max\limits_{[-1,\,1]} |\omega_n(x)|$ lo que se puede cumplir con ayuda de las reglas ordinarias de análisis matemático.

1. Consideremos el caso $n = 2$. Entonces

$$\omega_2(x) = (x+1)\,x\,(x-1); \quad \omega_2'(x) = 3x^2 - 1.$$

Las raíces del polinomio $\omega_2'(x)$ son $x_{1,2} = \pm 1/\sqrt{3} \approx \pm 0{,}5774$. Sustituyendo los valores obtenidos en ω_2, tenemos

$$\max_{[-1,\,1]} |\omega_2(x)| = |\omega_2(x_1)| = 2/\sqrt{27} \approx 0{,}3849$$

y, por lo tanto,

$$\Delta_1 \leqslant \left(\frac{\pi}{2}\right)^3 \cdot \frac{1}{3!} \cdot \frac{2}{\sqrt{27}} \approx 0{,}25.$$

2. Supongamos ahora que $n = 3$. En este caso

$$\omega_3(x) = (x+1)\left(x+\frac{1}{3}\right)\left(x-\frac{1}{3}\right)(x-1); \quad \omega_3'(x) = 4x^3 - \frac{20}{x}\,x.$$

De raíces del polinomio $\omega_3'(x)$ sirven $x_1 = 0$; $x_{2,\,3} = \pm\sqrt{5}/3 \approx \approx \pm 0{,}7454$. Es fácil comprobar que el valor máximo $|\omega_3(x)|$ se alcanza en los puntos x_1, x_3:

$$\max_{[-1,\,1]} |\omega_3(x)| = |\omega_3(x_2)| = 16/81 \approx 0{,}1975.$$

Por eso

$$\Delta_1 \leqslant \left(\frac{\pi}{2}\right)^4 \cdot \frac{1}{4!} \cdot \frac{16}{81} \approx 0{,}05.$$

3. Para $n = 4$ reduzcamos la estimación

$$\omega_4(x) = (x+1)\left(x+\frac{1}{2}\right)x\left(x-\frac{1}{2}\right)(x-1)$$

a la obtenida en el subp. 1. Efectivamente, en virtud de la imparidad de $\omega_4(x)$ se puede limitarse por la determinación del valor máximo de $|\omega_4(x)|$ sobre el segmento $[0, 1]$. En este caso

$$\max_{[0,\,1]} |\omega_4(x)| < 2 \cdot \frac{3}{2} \cdot \max_{[0,\,1]} \left| x\left(x-\frac{1}{2}\right)(x-1)\right|.$$

Reemplazando en el segundo miembro de la desigualdad obtenida la variable con ayuda de la fórmula $x = \frac{1}{2}(y+1)$ y tomando en cuenta los resultados del subp. 1, obtenemos

$$3\max_{[0,\,1]} \left| x\left(x-\frac{1}{2}\right)(x-1)\right| = \frac{3}{8}\max_{[-1,\,1]} |(y+1)\,y\,(y-1)| = \frac{1}{\sqrt{48}} \approx 0{,}1443.$$

Ahora bien,

$$\max_{[-1,\,1]} |\omega_4(x)| < 1/\sqrt{48} \approx 0{,}1443$$

y la estimación buscada

$$\Delta_1 < \left(\frac{\pi}{2}\right)^5 \cdot \frac{1}{5!} \cdot \frac{1}{\sqrt{48}} \approx 0{,}012. \ \blacktriangle$$

Parece que el ejemplo citado afirma la suposición acerca de que la estimación (9) es prácticamente conveniente y toma pequeños valores para la mayoría de las funciones al ser n suficientemente grandes. Sin embargo, en muchos casos esto no tiene lugar.

La cosa consiste en que sólo para una clase restringida de funciones (por ejemplo, para las funciones enteras) las derivadas de bastante alto orden son pequeñas. En cambio, para la mayoría de las funciones algunas de las derivadas de orden superior tienen la tendencia a crecer como $n!$. A título de ejemplo consideremos la función $y = \ln x$. Es evidente que para ella $y^{(n)} = (-1)^{n-1} \frac{(n-1)!}{x^n}$. Ahora bien, incluso cerca de los puntos donde la curva $y = \lg x$ parece suave, sus derivadas de órdenes bastante altos llegan a ser muy grandes y se comportan como $n!$

El inconveniente de la aproximación polinomial consiste en que, por lo general, falta el sentido físico el cual conduce de ordinario a generalizaciones útiles.

Por otro lado, la sencillez de la teoría de la aproximación polinomial y su desarrollo profundo en combinación con el mínimo de cálculos contribuyen a que este tipo de aproximación sea un instrumento cómodo al resolver distintos problemas, aún más porque, como muestra la experiencia, en los cálculos prácticos se obtienen buenos resultados en caso de la aproximación por polinomios, aunque el término residual ora, en general, es difícil de estimar, ora su estimación resulta demasiado exagerada.

Así pues, hemos examinado sólo un aspecto de la cuestión sobre el error, o sea, la influencia ejercida por las propiedadas de la función f en la magnitud Δ_1. La cuestión sobre el error en dependencia de la situación de los nodos de la red está estrechamente vinculada con las propiedades del polinomio de Chébyshev, y por eso conviene volver a estudiar con detalles este problema después de considerar estos polinomios. Aquí nos limitamos por una de las posibles estimaciones de la magnitud $|\omega_n(x)|$ sobre la red fija Λ_n. Sea que x esté entre x_k y x_{k+1}. Pongamos $\max\limits_{1\leqslant i\leqslant n} (x_i - x_{i-1}) = h$, entonces

$$|\omega_n(x)| = \prod_{i=0}^{n} |x - x_i| < (k+1)!\,(n-k)!\,h^{n+1} \leqslant n!\,h^{n+1}. \qquad (10)$$

Por eso la desigualdad (8) puede escribirse en la forma

$$\Delta_1 = \max_{[a;\,b]} |R_n(x)| < \frac{M_{n+1}}{n+1} h^{n+1}.$$

Nótese que la estimación (10) es bastante aproximada y no es difícil mejorarla (proponemos que el lector lo haga de por sí en calidad del ejercicio).

Ejemplo 2. ¿Con qué exactitud se puede calcular $\sqrt{117}$ con ayuda del polinomio interpolador para la función $y = \sqrt{x}$, eligiendo como nodos de interpolación $x_0 = 100$, $x_1 = 121$, $x_2 = 144$?

△ Ante todo, determinamos $M_3 = \max\limits_{[100;\ 144]} |(\sqrt{x})'''|$. Para esto hallamos

$$y' = \frac{1}{2} x^{-1/2};\ y'' = -\frac{1}{4} x^{-3/2};\ y''' = \frac{3}{8} x^{-5/2}.$$

De aquí $M_3 = \frac{3}{8} \cdot 100^{-5/2} = \frac{3}{8} \cdot 10^{-5}$. Por eso en virtud de la relación (8) tenemos

$$\Delta_1 \leqslant \frac{3}{8} \cdot 10^{-5} \cdot \frac{1}{3!} \, |(117-100)(117-121)(117-144)| \approx$$
$$\approx 0{,}12 \cdot 10^{-2}. \quad ▲$$

§ 7.6. Polinomio interpolador de Lagrange

La determinación inmediata de los coeficientes a_k del polinomio interpolador está ligada con ciertas dificultades de cálculo. Por eso al resolver los problemas prácticos se utilizan tipos especiales del polinomio interpolador.

En este párrafo vamos a examinar la forma del polinomio interpolador que se llama forma de Lagrange y suele designarse $L_n(x)$. Para construir L_n examinemos primero los polinomios auxiliares $l_i(x)$ de grado n que poseen las dos propiedades siguientes:

$$l_i(x_i) = 1 \quad (i = 0, 1, \ldots, n); \tag{1}$$
$$l_i(x_k) = 0 \quad (i \neq k;\ \mathrm{i},\ \mathrm{k} = 0, 1, \ldots, n). \tag{2}$$

Estas propiedades significan que, por ejemplo, el polinomio $l_0(x)$ toma en el punto x_0 el valor igual a la unidad y en los demás nodos se anula; el polinomio $l_1(x)$ toma en el nodo x_1 el valor igual a la unidad y en los demás puntos se anula, etc. En el caso general el polinomio $l_i(x)$ en el nodo x_i toma el valor igual a la unidad y en los demás nodos se anula. Ahora bien, en virtud de la propiedad (2) y del requisito de que el polinomio $l_i(x)$ tenga el grado n, obtenemos

$$l_i(x) = c_i(x-x_0)\ldots(x-x_{i-1})(x-x_{i+1})\ldots(x-x_n). \tag{3}$$

Luego, utilizando la propiedad (1), para determinar la constante c_i tenemos la ecuación

$$l_i(x_i) = c_i(x_i-x_0)\ldots(x_i-x_{i-1})(x_i-x_{i+1})\ldots$$
$$\ldots(x_i-x_n) = 1.$$

De aquí

$$c_i = \frac{1}{(x_i-x_0)\ldots(x_i-x_{i-1})(x_i-x_{i+1})\ldots(x_i-x_n)}. \tag{4}$$

Por eso la expresión explícita para $l_i(x)$ se puede representar del modo siguiente:

$$l_i(x)=\frac{(x-x_0)\ \ldots\ (x-x_{i-1})(x-x_{i+1})\ \ldots\ (x-x_n)}{(x_i-x_0)\ \ldots\ (x_i-x_{i-1})(x_i-x_{i+1})\ \ldots\ (x_i-x_n)}\,. \tag{5}$$

Hagamos ahora la siguiente combinación lineal de los polinomios l_i:

$$L_n(x)=\sum_{i=0}^{n} f_i l_i(x). \tag{6}$$

La expresión (6) es el polinomio de grado no superior a n. En el nodo x_i este polinomio toma el valor f_i, ya que el sumando correspondiente de la suma $f_i l_i(x_i)$ es igual a f_i y los demás sumandos $f_j l_j(x_i)$ son iguales al cero. Ahora bien, está construido el polinomio interpolador para la función $f(x)$. Esta forma precisamente se denomina *polinomio interpolador de Lagrange.*

Teniendo en cuenta que

$$\omega_n(x)=(x-x_0)(x-x_1)\ldots(x-x_n),$$

se puede considerar su derivada en el punto x_i:

$$\omega_n'(x_i)=(x_i-x_0)\ldots(x_i-x_{i-1})(x_i-x_{i+1})\ldots(x_i-x_n)$$

y escribir el polinomio de Lagrange en la forma

$$L_n(x)=\sum_{i=0}^{n} f_i\frac{\omega_n(x)}{(x-x_i)\,\omega_n'(x_i)}\,. \tag{7}$$

Las magnitudes $l_i(x)$ son, en cierto modo, polinomios ponderales de los nodos correspondientes y suelen llamarse *multiplicadores de Lagrange.* En adición a las propiedades (1) y (2) citemos una propiedad importante más de estos multiplicadores:

$$\sum_{i=0}^{n} l_i(x)=1. \tag{8}$$

En efecto, sea $f(x)\equiv 1$, entonces todas las $f_i=1$ $(i=0, 1, \ldots, n)$. Por otro lado, $f^{(n+1)}(x)\equiv 0$ y en virtud del teorema del § 7.5 $L_n(x)=$ $=f(x)=1$. Sustituyendo lo obtenido en la expresión (6), llegamos a la igualdad (8).

Ejemplo 1. Haciendo uso de los nodos $x_0=0$, $x_1=1/3$, $x_2=1$, construir el polinomio interpolador de Lagrange para la función $f=\operatorname{sen}(\pi x/2)$ y obtener una estimación uniforme del error sobre el segmento $[0, 1]$.

△ Nótese, ante todo, que $f(x_0)=0$, $f(x_1)=1/2$, $f(x_2)=1$. Luego, utilizando la expresión (7) para $n=2$, construimos el polinomio interpolador deseado:

$$L_2(x)=\frac{1}{2}\,\frac{x(x-1)}{\frac{1}{3}\left(\frac{1}{3}-1\right)}+1\cdot\frac{x\left(x-\frac{1}{3}\right)}{1\left(1-\frac{1}{3}\right)}\,.$$

La estimación de error se obtiene fácilmente de la relación (8) del § 7.5 para $n = 2$:

$$\Delta_1 \leqslant \frac{M_3}{3!} \max_{[0,1]} \left| x\left(x-\frac{1}{3}\right)(x-1) \right|.$$

En el caso dado es evidente que $M_3\ (\pi/2)^3$ y $\max\limits_{[0,1]} \left| x\left(x-\frac{1}{3}\right) \times \right.$ $\left. \times (x-1) \right| = 0{,}079$ y se determina de un modo análogo al hecho en el ejemplo 1 del § 7.5. Por eso tenemos finalmente

$$\Delta_1 \leqslant \left(\frac{\pi}{2}\right)^3 \cdot \frac{1}{3!} \cdot 0{,}079 \approx 0{,}05. \ \blacktriangle$$

Ejemplo 2. La función $f(x)$ se da en la forma tabular

x	0	1	2	6
y	-1	-3	3	1187

Valiéndose del polinomio interpolador de Lagrange, determinar el valor de esta función en el punto $x = 4$.

$\triangle$ Sustituyendo en la fórmula (7) los valores x_i y f_i para $n = 3$ y $x = 4$, obtenemos

$$L_3(4) = 1 \cdot \frac{(4-1)(4-2)(4-6)}{(-1)(-2)(-6)} - 3 \cdot \frac{4(4-2)(4-6)}{1(1-2)(1-6)} +$$
$$+ 3 \cdot \frac{4(4-1)(4-6)}{2(2-1)(2-6)} + 1187 \cdot \frac{4(4-1)(4-2)}{6(6-1)(6-2)} = 255. \ \blacktriangle$$

Si en el ejemplo considerado añadimos a la tabla un punto más, tendremos que volver a realizar los valores de la función para $x = 4$. Además, del mismo ejemplo se ve que el proceso de obtención del valor aproximado de la función ayuda de la fórmula de interpolación de Lagrange está vinculado con grandes cálculos. Esto origina la necesidad de simplificar el trabajo de cálculos.

Para la comodidad de los cálculos hagamos una tabla auxiliar

$x-x_0$	x_0-x_1	x_0-x_2	...	x_0-x_n	k_0
x_1-x_0	$x-x_1$	x_1-x_2	...	x_1-x_n	k_1
x_2-x_0	x_2-x_1	$x-x_2$	...	x_2-x_n	k_2
...	...	...	...	...	...
x_n-x_0	x_n-x_1	x_n-x_2	...	$x-x_n$	k_n

donde $x_0, x_1, \ldots, x_n$ son los nodos de interpolación y x es el valor del argumento para el cual se determina el valor aproximado con ayuda de la fórmula de interpolación de Lagrange. Designemos con k_0 el producto de los elementos de la primera fila

$$k_0 = (x - x_0)(x_0 - x_1) \ldots (x_0 - x_n).$$

En la forma general el producto de los elementos de la i-ésima fila es

$$k_i = (x_i - x_0) \ldots (x_i - x_{i-1})(x - x_i)(x_i - x_{i+1}) \ldots (x_i - x_n).$$

Coloquemos los números $k_0, k_1, \ldots, k_n$ en la columna derecha extrema de la tabla. Adicionalmente calculemos el producto de los elementos situados en la diagonal principal:

$$\omega_n(x) = (x - x_0)(x - x_1) \ldots (x - x_n).$$

Entonces el polinomio interpolador de Lagrange puede ser escrito en la forma

$$L_n(x) = \omega_n(x) \sum_{i=0}^{n} \frac{y_i}{k_i}. \qquad (9)$$

Haciendo uso de la fórmula (9), volvamos a resolver el ejemplo 2. Hacemos la tabla

4	−1	−2	−6	−48
1	4−1	1−2	1−6	15
2	2−1	4−2	2−6	−16
6	6−1	6−2	4−6	−240

y hallamos $\omega_3(4) = -48$. Determinamos el valor aproximado de la función en el punto $x = 4$, o sea, $f(4) \approx L_3(4)$, haciendo uso de la fórmula

$$L_3(x) = \omega_3(x) \sum_{i=0}^{3} \frac{y_i}{k_i},$$

o bien

$$L_3(4) = -48\left[\frac{-11}{-48} + \frac{-3}{15} + \frac{3}{-16} + \frac{1187}{-240}\right] = 255. \quad \blacktriangle$$

La fórmula de interpolación de Lagrange se simplifica considerablemente si los nodos de interpolación son *equidistantes*, o sea, $h = x_{i+1} - x_i = \text{const}$, donde h es el paso de interpolación. Introduzcamos la designación $t = (x - x_0)/h$. Según la fórmula (5) tenemos

$$l_i = (x) = \frac{(x - x_0) \ldots (x - x_{i-1})(x - x_{i+1}) \ldots (x - x_n)}{(x_i - x_0) \ldots (x_i - x_{i-1})(x_i - x_{i+1}) \ldots (x_i - x_n)}.$$

Puesto que

$$x-x_0=th,$$
$$x-x_1=th-h=h(t-1).$$
$$\cdots\cdots\cdots$$
$$x-x_i=th-ih=h(t-i),$$
$$\cdots\cdots\cdots$$
$$x-x_n=th-nh=h(t-n),$$

entonces

$$l_i=\frac{t(t-1)\ldots(t-i+1)(t-i-1)\ldots(t-n)h^n}{ih(i-1)h\ldots 1h(-1)h\ldots[-(n-i)h]}. \tag{10}$$

Nótese que una parte del producto en el denominador es igual a

$$ih(i-1)h\ldots h=i!h^i$$

y la otra parte es igual a

$$(-h)\ldots[-(n-i)h]=(-1)^{n-i}(n-i)!h^{n-i}.$$

Multiplicando el numerador y el denominador del segundo miembro de la igualdad (10) por $(-1)^{n-i}(t-i)$, obtenemos

$$l_i=\frac{t(t-1)\ldots(t-n)}{(t-i)i!(n-i)!}(-1)^{n-i}=$$
$$=(-1)^{n-i}\frac{C_n^i}{t-i}\frac{t(t-1)\ldots(t-n)}{n!},$$

donde $C_n^i=\dfrac{n!}{i!(n-i)!}$.

Entonces el polinomio interpolador de Lagrange para los nodos equidistantes de interpolación se puede escribir en la forma

$$L_n(x)=L_n(x_0+ht)=\frac{t(t-1)\ldots(t-n)}{n!}\sum_{i=0}^{n}(-1)^{n-i}\frac{C_n^i}{t-i}y_i. \tag{11}$$

Ejemplo 3. La función $y=\operatorname{sen} x$ se da en la forma tabular

x	0	$\pi/4$	$\pi/2$
y	0	0,707	1

Haciendo uso del polinomio interpolador de Lagrange, determinar el valor de esta función en el punto $x^*=\pi/6$. Estimar el error Δ_1.

$\triangle$ Ante todo, determinamos $t^*=\left(\frac{\pi}{6}-0\right)\left(\frac{\pi}{4}\right)^{-1}=\frac{2}{3}$. Sustituyendo en la fórmula (11) el valor obtenido de t^* y los valores y_i para $n=2$, tenemos

$$L_2\left(\frac{\pi}{6}\right)=\frac{(2/3)(2/3-1)(2/3-2)}{2!}\times$$
$$\times\left(\frac{2}{2/3-0}\cdot 0-\frac{2}{2/3-1}\cdot 0{,}707+\frac{1}{2/3-2}\cdot 1\right)=0{,}517.$$

Para estimar el error hacemos uso de la fórmula (8) del § 7.5. Aquí $M_3 = \max\limits_{[0,\, \pi/2]} |(\text{sen}\, x)'''| = 1$, $x^* = \pi/6$; por eso

$$\Delta_1 = \frac{1}{3!} \cdot \frac{\pi}{6} \cdot \frac{\pi}{12} \cdot \frac{\pi}{3} = 0{,}024.$$

Nótese que al calcular el error covniene convertir la medida en grados en medida en radianes.

Ahora bien, redondeando el resultado hasta dos cifras obtenemos sen $(\pi/6) = 0{,}52 \pm 0{,}03$. ▲

§ 7.7. Diferencias finitas

En los párrafos precedentes hemos considerado diferentes cuestiones de la teoría de interpolación sobre la red arbitraria de los nodos. En la práctica la información disponible sobre la función (en forma de sus valores) se da con frecuencia sobre una red uniforme, o sea, para los nodos equidistantes. En este caso no sólo se simplifican las formas de los polinomios interpoladores sino se reduce también el mismo proceso de cálculo lo que es un hecho no menos sustancial desde el punto de vista práctico. Al construir los polinomios interpoladores sobre un red uniforme se utilizan las magnitudes llamadas *diferencias finitas.*

Examinemos una red uniforme con paso h: $x_i = x_0 + ih$ $(i = 0, \pm 1, \pm 2, \ldots)$ en cuyos nodos se dan los valores $f_i = f(x_i)$ de la función $f(x)$.

En las obras matemáticas se utilizan tres tipos de diferencias finitas: *diferencias descendentes* $\Delta^k f_i$ para la interpolación hacia adelante; *diferencias ascendentes* $\nabla^k f_i$ para la interpolación hacia atrás; *diferencias centrales* $\delta^k f_i = f_i^k$ para construir las fórmulas de interpolación centrales.

Se llama *diferencia finita de primer orden* la dircríencia entre los valores de la función en el nodo dado y en el nodo precedente:

$$\begin{gathered} f_1 - f_0 = \Delta f_0 = \nabla f_1 = \delta f_{1/2} = f^1_{1/2}, \\ f_2 - f_1 = \Delta f_1 = \nabla f_2 = \delta f_{3/2} = f^1_{3/2}, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ f_{i+1} - f_i = \Delta f_i = \nabla f_{i+1} = \delta f_{i+1/2} = f^1_{i+1/2}. \end{gathered} \tag{1}$$

Esta definición puede escribirse en otra forma:

$$\Delta f_i = f_{i+1} - f_i; \ \nabla f_i = f_i - f_{i-1}; \quad \delta f_i = f_{i+1/2} - f_{i-1/2} \tag{2}$$

Se denomina *diferencia finita de segundo orden* la diferencia entre los valores de la diferencia finita primera en el nodo dado y en el precedente

$$\begin{gathered} \Delta^2 f_i = \Delta f_{i+1} - \Delta f_i; \quad \nabla^2 f_i = \nabla f_i - \nabla f_{i-1}; \\ \delta^2 f_i = f_i^2 = \delta f_{i+1/2} - \delta f_{i-1/2}. \end{gathered} \tag{3}$$

De un modo análogo se definen las diferencias finitas de orden arbitrario k;

$$\Delta^k f_i = \Delta^{k-1} F_{i+1} - \Delta^{k-1} f_i; \quad \nabla^k f_i = \nabla^{k-1} f_i - \nabla^{k-1} f_{i-1}; \quad (4)$$
$$\delta^k f_i = f_i^k = \delta^{k-1} f_{i+1/2} - \delta^{k-1} f_{i-1/2}.$$

En algunas fórmulas de interpolación que se examinan a continuación junto con diferoncias (4) se utilizan las medias aritméticas de las

Tabla 7.5

x	f	Δf	$\Delta^2 f$	$\Delta^3 f$	$\Delta^4 f$
x_0	f_0	Δf_0	$\Delta^2 f_0$	$\Delta^3 f_0$	$\Delta^4 f_0$
x_0+h	f_1	Δf_1	$\Delta^2 f_1$	$\Delta^3 f_1$	...
x_0+2h	f_2	Δf_2	$\Delta^2 f_2$	...	
x_0+3h	f_3	Δf_3	...		
x_0+4h	f_4	...			
...	...	...	...	...	...

Tabla 7.6

x	f	∇f	$\nabla^2 f$	$\nabla^3 f$	$\nabla^4 f$
...	...	...	...	...	...
x_0-4h	f_{-4}	...			
x_0-3h	f_{-3}	∇f_{-3}	...		
x_0-2h	f_{-2}	∇f_{-2}	$\nabla^2 f_{-2}$	...	
x_0-h	f_{-1}	∇f_{-1}	$\nabla^2 f_{-1}$	$\nabla^3 f_{-1}$	...
x_0	f_0	∇f_0	$\nabla^2 f_0$	$\Delta^3 f_0$	$\nabla^4 f_0$

diferencias finitas vecinas del mismo orden:

$$\mu f_i^k = \frac{1}{2}\left(f_{i+1/2}^k + f_{i-1/2}^k\right), \quad \mu f_{i+1/2}^k = \frac{1}{2}\left(f_{i+1}^k + f_i^k\right). \tag{5}$$

La primera de estas magnitudes se utiliza cuando k es impar y la segunda, cuando k es par.

Es cómodo escribir las diferencias finitas de la función f en la forma de tablas. En este caso las diferencias finitas ascendentes y descendentes se escriben en la forma de las tablas horizontales (tablas 7.5 y 7.6) y las diferencias finitas centrales, en la forma de las tablas centrales (tabla 7.7).

Tabla 7.7

x	f	δf	$\delta^2 f$	$\delta^3 f$	$\delta^4 f$
...	...	...	...	...	...
x_0-2h	f_{-2}				
		$\delta f_{-3/2}$			
x_0-h	f_{-1}		$\delta^2 f_{-1}$		
		$\delta f_{-1/2}$		$\delta^3 f_{-1/2}$	
x_0	f_0		$\delta^2 f_0$		$\delta^4 f_0$
		$\delta f_{1/2}$		$\delta^3 f_{1/2}$	
x_0+h	f_1		$\delta^2 f_1$		
		$\delta f_{3/2}$			
x_0+2h	f_2				
...	...	...	...	...	...

Consideremos algunas propiedades de las diferencias finitas.

1°. *Las diferencias descendentes, ascendentes y centrales están ligadas entre sí por las relaciones siguientes:*

$$\Delta^k f_i = \nabla^k f_{i+k} = \delta^k f_{i+k/2} \tag{6}$$

lo que se demuestra fácilmente con ayuda del método de inducción matemática partiendo de la definición de las diferencias finitas (4).

□ Para $k = 1$ las relaciones (6) son evidentes, ya que en virtud de las igualadades (4) tenemos

$$\Delta f_i = f_{i+1} - f_i; \ \nabla f_{i+1} = f_{i+1} - f_i; \ \delta f_{i+1/2} = f_{i+1} - f_i.$$

Supongamos ahora que las relaciones (6) son justas para todo $k \leqslant m - 1$. Mostremos que en este caso serán justas también para $k = m$ y, por consiguiente, para todos los k. Utilizando las igualdades (4) y la suposición sobre la validez de (6) para $k \leqslant m - 1$, tenemos

$$\Delta^m f_i = \Delta^{m-1} f_{i+1} - \Delta^{m-1} f_i = \nabla^{m-1} f_{i+m} - \nabla^{m-1} f_{i+m-1} = \nabla^m f_{i+m},$$

$$\Delta m f_i = \Delta^{m-1} f_{i+1} - \Delta^{m-1} f_i = \delta^{m-1} f_{i+\frac{m+1}{2}} - \delta^{m-1} f_{i+\frac{m-1}{2}} = \delta^m f_{i+\frac{m}{2}}. \ ■$$

2°. *La diferencia finita satisface la igualdad*

$$\Delta\,(af + bg)_i = a\Delta f_i + b\Delta g_i, \tag{7}$$

donde a y b son constantes.

□ En efecto,

$$\Delta\,(af + bg)_i = af_{i+1} + bg_{i+1} - (af_i + bg_i) = a\Delta f_i + b\Delta g_i. \ ■$$

Esta propiedad significa, en particular, que la diferencia finita de la suma de dos funciones o de la diferencia de las mismas es igual a la suma o diferencia de las diferencias finitas de estas funciones, respectivamente, así como que la diferencia finita del producto de la función por el multiplicador constante es igual al producto de este multiplicador por la diferencia finita de la función.

3°. *La diferencia finita está ligada con la derivada correspondiente por la relación*

$$\Delta^k f_i = h^k f^{(k)}(\xi); \quad \xi \in (x_i;\ x_i + kh). \tag{8}$$

Como consecuencia de la igualdad (8) obtenemos que las diferencias finitas de orden n del polinomio de grado n son constantes e iguales a $h^n n! a_0$ y las diferencias finitas de todo orden más alto son iguales a cero (a_0 es el coeficiente del polinomio para el grado superior de x).

4°. *La diferencia finita de orden k puede ser representada en forma de la siguiente combinación lineal de valores f_i:*

$$\Delta^k f_i = \sum_{j=0}^{k} (-1)^j C_k^j f_{i+k-j}, \tag{9}$$

donde $C_k^j = \frac{k!}{j!(k-j)!}$ es el número de combinaciones de k elementos dispuestos j a j (con la particularidad de que $0! = 1$).

□ Hagamos uso del método de inducción matemática. Para $k = 1$ esta relación es evidente, puesto que no es sino la definición de la diferencia finita primera: $\Delta f_i = f_{i+1} - f_i$.

Supongamos ahora que la igualdad (9) es justa para cierto $k = m$. Entonces

$$\Delta^{m+1} f_i = \Delta^m f_{i+1} - \Delta^m f_i =$$

$$= \sum_{j=0}^{m} (-1)^j C_m^j f_{i+1+m-j} - \sum_{j=0}^{m} (-1)^j C_m^j f_{i+m-j} =$$

$$= C_m^0 f_{i+1+m} + \sum_{j=1}^{m} (-1)^j (C_m^j + C_m^{j-1}) f_{i+1+m-j} +$$

$$+ (-1)^{m+1} C_{m+1}^{m+1} f_i = \sum_{j=0}^{m+1} (-1)^j C_{m+1}^j f_{i+m+1-j}.$$

Hemos utilizado las propiedades de las combinaciones: $C_m^{j-1} = C_{m+1}^j - C_m^j$ y $C_m^0 = C_{m+1}^0 = C_m^m = C_{m+1}^{m+1} = 1$. Ahora bien, si la igualdad (9) es justa para $k = m$, ésta será justa también para $k = m + 1$. ■

Ejemplo 1. Hacer la tabla horizontal de diferencias finitas de la función $y = x^3 + 3x^2 - x - 1$ a partir del valor inicial $x = 0$, tomando el paso $h = 1$.

△ Suponiendo $x_0 = 0$, $x_1 = 1$, $x_2 = 2$, . . ., hallamos los valores correspondientes:

x	0	1	2	3	4	5	...
y	-1	2	17	50	107	194	...

Hallamos las diferencias finitas de primer orden:

$$\Delta y_0 = y_1 - y_0 = 2 - (-1) = 3;$$
$$\Delta y_1 = y_2 - y_1 = 17 - 2 = 15;$$
$$\Delta y_2 = y_3 - y_2 = 50 - 17 = 33;$$
$$\Delta y_3 = y_4 - y_3 = 107 - 50 = 57;$$
$$\Delta y_4 = y_5 - y_4 = 194 - 107 = 87; \ldots$$

¿Cuales son las diferencias finitas de segundo orden?

$$\Delta^2 y_0 = \Delta y_1 - \Delta y_0 = 15 - 3 = 12;$$
$$\Delta^2 y_1 = \Delta y_2 - \Delta y_1 = 33 - 15 = 18;$$
$$\Delta^2 y_2 = \Delta y_3 - \Delta y_2 = 57 - 33 = 24;$$
$$\Delta^2 y_3 = \Delta y_4 - \Delta y_3 = 87 - 57 = 30; \ldots$$

Determinamos las diferencias finitas de tercer orden:

$$\Delta^3 y_0 = \Delta^2 y_1 - \Delta^2 y_0 = 18 - 12 = 6;$$
$$\Delta^3 y_1 = \Delta^2 y_2 - \Delta^2 y_1 = 24 - 18 = 6;$$
$$\Delta^3 y_2 = \Delta^2 y_3 - \Delta^2 y_2 = 30 - 24 = 6; \ldots$$

Vemos que las diferencias finitas terceras $\Delta^3 y_0$, $\Delta^3 y_1$, $\Delta^3 y_2$ son constantes. Esto se explica por el hecho de que la función $f(x)$ es un polinomio de tercer grado. La diferencia finita tercera puede calcularse también con ayuda de la fórmula

$$\Delta^n P_n(x) = n! a_0 h^n,$$

o sea, $\Delta^3 P_3(x) = 3! \cdot 1 \cdot 1^3 = 6$ y las diferencias finitas de cuarto orden son iguales a cero.

Hacemos la tabla de diferencias finitas:

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$
0	−1	3	12	6	0
1	2	15	18	6	0
2	17	33	24	6	
3	50	57	30		
4	107	87			
5	194				

Posteriormente durante los cálculos es conveniente apuntar inmediatamente las diferencias finitas en la tabla. ▲

Puesto que los valores iniciales de la función f_i se dan, por lo general, con cierto error ε, que representa en sí errores de redondeo o errores aleatorios, es racional considerar la influencia que estos factores ejercen en el error de las diferencias finitas de órdenes superiores.

Comencemos con la influencia ejercida por los errores aleatorios y los de redondeo. Supongamos que en vez de f_i^k hemos obtenido $f_i^k + \varepsilon$. Entonces la tabla de diferencias finitas tomará la forma representada en la pág. 281 (véase la tabla 7.8).

En virtud de la igualdad (9) esto quiere decir que el error ε en la diferencia de orden k se extiende luego a las diferencias de orden $k+m$ con los coeficientes $(-1)^j C_m^j$.

Sin embargo, si los valores iniciales f_i se dan con el mismo error ε, este error se extiende a las diferencias de orden m con el coeficiente 2^m y aumenta rápidamente con el crecimiento de m ($\Delta(f_i^m) = 2^m \varepsilon$).

Si las derivadas de órdenes bastante altos de la función f quedan restringidas, de la fórmula (8) se deduce que las diferencias finitas respectivas f_j^m decrecen con el aumento de m. Por eso llega naturalmente tal momento cuando los errores de diferencias finitas, introducidos debido al redondeo o debido a la inexactitud de la información inicial, llegarán a ser comparables con los mismos valores de las diferencias finitas o incluso los superarán. Por lo tanto, la información contenida en la tabla de estas diferencias, en realidad, resultará información sobre las diferencias de los errores y no sobre las de la función; en este caso la utilización de esta información se hará no racional. Entonces se dice que el orden de las últimas diferencias

Tabla 7.8

...	...	...	...
f_{i-2}^{k}			
	$f_{i-3/2}^{k+1}$		
f_{i-1}^{k}		$f_{i-1}^{k+2}+\varepsilon$	
	$f_{i-1/2}^{k+1}+\varepsilon$		$f_{i-1/2}^{k+3}-3\varepsilon$
$f_{i}^{k}+\varepsilon$		$f_{i}^{k+2}-2\varepsilon$	
	$f_{i+1/2}^{k+1}-\varepsilon$		$f_{i+1/2}^{k+3}+3\varepsilon$
f_{i+1}^{k}		$f_{i+1}^{k+2}+\varepsilon$	
	$f_{i+3/2}^{k+1}$		
f_{i+2}^{k}			
...			

finitas que todavía es conveniente usar en los cálculos es el *orden de exactitud de la tabla de diferencias finitas.*

Ejemplo 2. Se da la tabla de los valores de la función f = sen x:

x	46°	47°	48°	49°	50°	51°	52°
f	0,7198	0,7314	0,7431	0,7547	0,7660	0,7771	0,7880

Todas las cifras citadas son justas en el sentido estrecho. Hacer la tabla de diferencias finitas y determinar el orden de exactitud de la tabla.

△ Calculamos las diferencias finitas y hacemos la tabla:

Tabla 7.9

x	f	f^1	f^2	f^3	f^4	f^5	f^6
46°	0,7193						
		121					
47°	0,7314		−4				
		117		3			
48°	0,7431		−1		−5		
		116		−2		8	
49°	0,7547		−3		3		−12
		113		1		−4	
50°	0,7660		−2		−1		
		111		0			
51°	0,7771		−2				
		109					
52°	0,7880						
	0,00005	1	2	4	8	16	32

Nótese que las diferencias finitas suelen escribirse en las unidades del último orden de los valores de la función.

En la última fila de la tabla citada están colocados los errores absolutos correspondientes. Es evidente que los valores absolutos de las diferencias finitas terceras son comparables con su error y los valores absolutos de las diferencias sucesivas son considerablemente menores de sus errores. Por eso el orden de exactitud de la tabla 7.9 es igual a 2. ▲

Lo no racional de utilizar las diferencias finitas de un orden superior al segundo en el ejemplo citado puede ser confirmado también del modo siguiente. Puesto que los datos iniciales f_i están dados con error $\varepsilon = 0{,}5\cdot 10^{-4}$, las diferencias finitas que en valor absoluto son menores de este error no tiene sentido tomarlas en consideración. Por otro lado, utilizando la igualdad (8), obtenemos $\Delta^h f_i \approx$ $\approx (\pi/180)^h$, por eso el orden de exactitud de la tabla 7.9 se define por las desigualdades siguientes:

$$(\pi/180)^k > 0{,}5\cdot 10^{-4} \geqslant (\pi/180)^{k+1}$$

que se cumplen, evidentemente, para $k = 2$.

En el caso general, de orden de exactitud de la tabla de diferencias finitas en el sentido mencionado sirve el valor mínimo de k que satisfaga la desigualdad.

$$h^{k+1} M_{k+1} \leqslant \varepsilon, \text{ donde } M_\nu = \max_{[x_i,\, x_{i+\nu}]} |f^{(\nu)}(x)|. \tag{10}$$

Hemos considerado la influencia que un error de la información inicial ejerce en el grado del polinomio interpolador. Además, si los valores f_i se dan aproximadamente o por cualesquiera causas el cálculo del valor del polinomio $P_n(x^*)$ no puede ser realizado con exactitud absoluta, obtenemos de hecho sólo un valor aproximado $\overline{P}_n(x^*)$ para el exacto $P_n(x^*)$. En este caso el error de cálculo

$\Delta_2(\overline{P}_n) = | P_n(x^*) - \overline{P}_n(x^*) |$ se estima según las reglas generales de cálculo del error de una función.

Consideremos, por ejemplo, el polinomio de Lagrange $L_n(x) = \sum_{i=0}^{n} f_i l_i(x)$. Supongamos que se necesita calcular $L_n(x^*)$ para los errores dados f_i y sus errores ε_i. Las magnitudes de los coeficientes de Lagrange $l_i(x)$ están tabuladas para los nodos equidistantes y pueden considerarse números exactos, ya que están obtenidas a partir de los valores exactos de los nodos y a partir del x^* exacto. Por eso para el polinomio de Lagrange tenemos

$$\Delta_2(\overline{L}_n) \leqslant \sum_{i=0}^{n} \varepsilon_i \, | l_i(x^*) |.$$

En el caso en que todos los ε_i son los mismos e iguales a ε obtenemos

$$\Delta_2(\overline{L}_n) \leqslant \varepsilon \sum_{i=0}^{n} | f_i(x^*) |.$$

Análogamente se puede hallar el error de cálculo también para otras formas del polinomio interpolador.

Ejemplo 3. Sobre el segmento $[-1, 1]$ obtener la estimación uniforme del valor de cálculo de los valores del polinomio interpolador de Lagrange construido para la finción $f = \cos(\pi x/2)$ según los nodos $x_0 = -1/2$, $x_1 = 0$, $x_2 = 1/2$.

$\triangle$ Puesto que $f_0 = f_2 = 1/\sqrt{2} = 0{,}707 \pm 0{,}0002$ y $f_1 = 1$ es un número exacto, el error buscado de cálculo tiene la forma

$$\Delta_2(\overline{L}_2) = 0{,}0002 \left| \frac{x^*\left(x^* - \frac{1}{2}\right)}{-\frac{1}{2}\left(-\frac{1}{2} - \frac{1}{2}\right)} \right| + 0{,}0002 \left| \frac{\left(x^* + \frac{1}{2}\right) x^*}{\left(\frac{1}{2} + \frac{1}{2}\right) \cdot \frac{1}{2}} \right| =$$

$$= 0{,}0004 \left[\left| x^*\left(x^* - \frac{1}{2}\right) \right| + \left| \left(x^* + \frac{1}{2}\right) x^* \right| \right].$$

No es difícil mostrar que sobre el segmento $[-1, 1]$ $\Delta_2(\overline{L}_2)$ toma el valor máximo en los puntos $x^* = \pm 1$ y por eso la estimación buscada es $\Delta_2(\overline{L}_2) = 0{,}0008$. ▲

§ 7.8. Polinomios interpoladores de Stirling y de Bessel

Primero vamos a examinar una cuestión importante, desde el punto de vista del error de interpolación, concerniente a la elección de los nodos de interpolación para un grado fijo del polinomio. Consideremos la expresión para el término residual del polinomio interpolador:

$$\frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{i=0}^{n} (x^* - x_i); \ \xi \in (x_0, x_n).$$

Puesto que el segmento de interpolación es de ordinario no grande, la derivada $f^{n+1}(x)$ tiene una pequeña gama de medición. Por consiguiente, la gama de medición de la magnitud del error se determina, en lo fundamental, por el producto

$$|\omega_n(x^*)| = \prod_{i=0}^{n} |x^* - x_i|.$$

Esta magnitud será mínima si en calidad de nodos de interpolación se toman los nodos más próximos a x^*. Ahora bien, para un grado par $n = 2k$ del polinomio de interpolación es necesario tomar el nodo más próximo al punto x^* y k nodos a la izquierda y k nodos a la derecha de éste y, siempre que se trate de un grado impar $n = 2k + 1$, $k + 1$ nodos a la izquierda y $k + 1$ nodos a la derecha del punto x^*

Pasemos ahora al problema de construcción del polinomio interpolador para la función f asignada con sus valores f_i en los nodos x_i de la red uniforme con paso h.

Supongamos que el punto x^* está situado en la proximidad de cierto nodo que designamos x_0. Se necesita construir un polinomio interpolador de grado par. En virtud de lo dicho anteriormente en calidad de nodos de interpolación conviene elegir una red simétrica respecto al nodo x_0.

$$\ldots, x_{-k}, \ldots, x_{-1}, x_0, x_1, \ldots, x_k, \ldots$$

Introduzcamos una nueva variable t con cuya ayuda el origen se traslada al punto x_0:

$$t = (x - x_0)/h; \tag{1}$$

en este caso $t^* = (x^* - x_0)/h$.

Construyamos el polinomio interpolador en la forma siguiente:

$$P_{2k}(x) = P_{2k}(x_0 + th) = a_0 + \frac{a_1}{1!}t + \frac{a_2}{2!}t^2 + \ldots$$
$$\ldots + \frac{a_{2k-1}}{(2k-1)!}t(t^2 - 1^2)\ldots(t^2 - (k-1)^2) +$$
$$+ \frac{a_{2k}}{(2k)!}t^2(t^2 - 1^2)\ldots(t^2 - (k-1)^2). \tag{2}$$

Los coeficientes desconocidos a_i los determinemos de las condiciones de coincidencia del polinomio P_{2k} y de la función f en los nodos x_i. En este caso nótese que la relación (1) hace corresponder al nodo x_i la magnitud $t = i$. Por ejemplo, al nodo x_0 le corresponde $t = 0$ y al nodo x_{-3} le responde $t = -3$.

Ahora bien, para determinar los coeficientes a_i obtenemos el sistema de ecuaciones lineales

$$P_{2k}(x_0 + ih) = f_i \quad (i = 0, \pm 1, \ldots, \pm k). \tag{3}$$

La estructura de este sistema es tal que a_0 se determina inmediatamente de la primera escuación del sistema (3):

$$P_{2k}(x_0) = a_0 = f_0$$

y la determinación de los demás coeficientes se reduce a la resolución sucesiva de los sistemas de dos ecuaciones con dos incógnitas:

$$P_{2k}(x_0+h) = f_0 + a_1 + \frac{a_2}{2} = f_1,$$

$$P_{2k}(x_0-h) = f_0 - a_1 + \frac{a_2}{2} = f_{-1}.$$

De aquí (véanse las designaciones del § 7.7) $a_1 = \mu f_0^1$, $a_2 = f_0^2$.

Continuando este proceso, vemos que en el j — ésimo paso el determinante del sistema de ecuaciones respecto a los coeficientes a_{2j-1} y a_{2j} tiene la forma

$$\begin{vmatrix} 1 & 1/2 \\ -1 & 1/2 \end{vmatrix} = 1 \neq 0.$$

Por consiguiente, todos los coeficientes a_i del polinomio (2) se determinan unívocamente por el sistema de ecuaciones (3) y en virtud del teorema sobre la unicidad el polinomio (2) es interpolador para la función f.

Como resultado de transformaciones poco complicadas llegamos a las siguientes expresiones para los coeficientes:

$$a_0 = f_0, \; a_{2j-1} = \mu f_0^{2j-1}, \; a_{2j} = f_0^{2j} \; (j = 1, 2, \ldots). \qquad (4)$$

Sustituyendo los valores hallados de los coeficientes en la expresión (2), obtenemos el *polinomio interpolador de Stirling* que suele designarse S_{2k}:

$$S_{2k}(t) = f_0 = \mu f_0^1 t + \frac{1}{2!} f_0^2 t^2 + \ldots + \frac{\mu f_0^{2k-1}}{(2k-1)!} t(t^2 - 1^2) \ldots$$

$$\ldots (t^2 - (k-1)^2) + \frac{f_0^{2k}}{(2k)!} t^2 (t^2 - 1^2) \ldots (t^2 - (k-1)^2). \qquad (5)$$

Puesto que el polinomio de Stirling no es sino una nueva forma del polinomio interpolador de Lagrange construido sobre los nodos $x_{-k}, \ldots, x_k$, entonces, en virtud de la fórmula (4) del § 7.5 su termino residual respecto a la variable t se puede representar en la forma

$$R_{2k} = \frac{f^{(2k+1)}(\xi)}{(2k+1)!} h^{2k+1} \prod_{i=-k}^{k} (t-i); \; \xi \in (x_{-k}, x_k) \qquad (6)$$

y la estimación del error del valor aproximado de $P_{2k}(x) = S_{2k}(t)$ (del error del método), en la forma

$$\Delta_1 = |f(x) - S(t)| \leqslant \frac{M_{2k+1}}{(2k+1)!} h^{2k+1} |t(t^2-1^2) \dots (t^2-k^2)|, \quad (7)$$

donde $M_{2k+1} = \max\limits_{[x_{-k}, x_k]} |f^{(2k+1)}(x)|$.

Supongamos ahora que el punto de interpolación x^* está entre los nodos x_0 y x_1 cerca del punto $(x_0 + x_1)/2$. Se necesita construir el polinomio interpolador de grado impar. Entonces, según hemos señalado anteriormente, la red que minimiza el error es simétrica respecto al punto $(x_0 + x_1)/2$, o sea, respecto al punto $t = 1/2$.

Ahora bien, sobre la red

$$\dots, x_{-k}, \dots, x_{-1}, x_0, x_1, x_2, \dots, x_{k+1}, \dots$$

respecto a la variable t, definida por la fórmula (1), construimos el polinomio interpolador en la forma siguiente:

$$P_{2k+1}(x) = P_{2k+1}(x_0 + th) = b_0 + \frac{b_1}{1!}\left(t - \frac{1}{2}\right) + \frac{b_2}{2!} t(t-1) + \dots$$
$$\dots + \frac{b_{2k}}{(2k)!} t(t^2-1^2) \dots (t^2-(k-1)^2)(t-k) +$$
$$+ \frac{b_{2k+1}}{(2k+1)!} t(t^2-1^2) \dots (t^2-(k-1)^2)\left(t - \frac{1}{2}\right)(t-k). \quad (8)$$

Los coeficientes desconocidos b_i se determinan de las condiciones de coincidencia de los valores del polinomio P_{2k+1} con los valores de la función f en los nodos x_i.

Ahora bien, para determinar los coeficientes b_i tenemos un sistema de ecuaciones lineales

$$P_{2k+1}(x_0 + ih) = f_i \ (i = -k, \dots, 0, 1, \dots, k+1). \quad (9)$$

La estructura de este sistema es tal que su resolución se reduce a la resolución sucesiva del sistema de dos ecuaciones con dos incógnitas

$$P_{2k+1}(x_0) = b_0 - \frac{b_1}{2} = f_0,$$
$$P_{2k+1}(x_1) = b_0 + \frac{b_1}{2} = f_1.$$

De aquí $b_0 = \mu f_{1/2}$, $b_1 = f^1_{1/2}$.
Continuando este proceso, vemos que en el j-ésimo paso el determinante del sistema de ecuaciones respecto a los coeficientes b_{2j-2} y b_{2j-1} es distinto del cero lo que se puede mostrar fácilmente de un modo análogo al cómo esto hemos hecho al construir al polinomio de Stirling. Por lo tanto, todos los coeficientes b_i del polinomio (8) se definen unívocamente por el sistema de ecuaciones (9) y en virtud del teorema de unicidad el polinomio (8) es interpolador para la función f.

Las transformaciones poco complicadas ofrecen las siguientes expresiones para los coeficientes:

$$b_{2j-2} = \mu f_{1/2}^{2j-2}, \quad b_{2j-1} = f_{1/2}^{2j-1} \quad (j = 1, 2, \ldots). \tag{10}$$

Sustituyendo los valores hallados de los coeficientes en la expresión (8), obtenemos el *polinomio interpolador de Bessel* que suele designarse B_{2k+1}:

$$B_{2k+1}(t) = \mu f_{1/2} + \frac{f_{1/2}^{1}}{1!}\left(t - \frac{1}{2}\right) +$$

$$+ \frac{\mu f_{1/2}^{2}}{2!} t(t-1) + \frac{f_{1/2}^{3}}{3!} t(t-1)\left(t - \frac{1}{2}\right) + \ldots$$

$$\ldots + \frac{\mu f_{1/2}^{2k}}{(2k)!} t(t^2 - 1^2) \ldots (t^2 - (k-1)^2 (t-k) +$$

$$+ \frac{f_{1/2}^{2k+1}}{(2k+1)!} t(t^2 - 1^2) \ldots$$

$$\ldots (t^2 - (k-1)^2)(t-k)\left(t - \frac{1}{2}\right). \tag{11}$$

Puesto que el polinomio de Bessel es una forma más de representar el polinomio interpolador de Lagrange, construido sobre los nodos $x_{-k}, \ldots, x_{k+1}$, entonces, en virtud de la fórmula (4) del § 7.5, su término residual respecto a la variable t puede escribirse en la forma

$$R_{2k+1} = \frac{f^{(2k+2)}(\xi)}{(2k+2)!} h^{2k+2} \prod_{i=-k}^{k+1} (t - i); \; \xi \in (x_{-k}; \, x_{k+1}) \tag{12}$$

y la estimación del error del valor aproximado de $P_{2k+1}(x) = B_{2k+1}(t)$ (del error del método), en la forma

$$\Delta_1 = | f(x) - B_{2k+1}(t) | \leqslant \frac{M_{2k+2}}{(2k+2)!} h^{2k+2} \prod_{i=-k}^{k+1} | t - i |, \tag{13}$$

donde $M_{2k+2} = \max\limits_{[x_{-k}, x_{k+1}]} | f^{(2k+2)}(x) |$.

Ahora bien, hemos considerado dos polinomios interpoladores: el polinomio de Stirling que se utiliza al construir un polinomio de grado par y se construye según el número impar de nodos y el polinomio de Bessel que se utiliza al construir un polinomio de grado impar y se construye según un número par de nodos.

En cambio, si el grado del polinomio está fijado no rígidamente, o sea, puede ser tanto par como impar, es conveniente utilizar el polinomio de Stirling en el caso en que

$$| t^* | = | x^* - x_0 | / h \leqslant 0{,}25, \tag{14}$$

o sea cuando el punto de interpolación x^* está situado más próximo al nodo x_0 que al centro entre los nodos. El polinomio de Bessel ha de utilizarse en el caso en que

$$0{,}25 \leqslant t^* \leqslant 0{,}75, \tag{15}$$

o sea cuando el punto de interpolación x^* está situado más próximo al centro entre los nodos x_0 y x_1. Una de las condiciones (14) ó (15) puede ser siempre asegurada por la elección del nodo respectivo en calidad de x_0.

Ejemplo. Utilizando el polinomio interpolador correspondiente, calcular en los puntos $x_1^* = 48{,}63°$ y $x_2^* = 49{,}19°$ los valores de la función $f = \operatorname{sen} x$, representada en la forma de una tabla con paso igual a 1° (véase la tabla 7.9), que contiene los valores f_i con cuatro cifras justas en el sentido estrecho. Estimar el error del resultado.

△ En el ejemplo 2 del § 7.7 hemos determinado que el orden de exactitud de la tabla dada es 2. Por eso no es racional construir un polinomio interpolador cuyo grado sea superior al segundo, o sea, conviene construir el polinomio de Stirling de segundo grado o bien el polinomio de Bessel de primer grado.

Puesto que el punto $x_1^* = 48{,}63$ está situado más próximamente al centro entre los nodos de 48° y 49°, entonces, al calcular $\operatorname{sen} x_1^*$, en calidad de x_0 es necesario tomar el nodo de 48° y hacer uso del polinomio de Bessel; en este caso $t_1^* = (x_1^* - x_0) \times h^{-1} = 0{,}63$. El punto $x_2^* = 49{,}19°$ está cerca del ángulo 49°, por eso, al calcular $\operatorname{sen} x_2^*$, en calidad de central hay que elegir el nodo $x_0 = 49°$ y valerse del polinomio de Stirling; en este caso $t_2^* = (x_2^* - x_0)h^{-1} = 0{,}19$.

Ahora bien, utilizando lo dicho anteriormente, las fórmulas (11) y (5), así como los datos de la tabla 7.9, tenemos

$$B_1(0{,}63) = \frac{0{,}7431 + 0{,}7547}{2} + \frac{0{,}0116}{1!} \cdot 0{,}13 = 0{,}750408;$$

$$S_2(0{,}19) = 0{,}7547 + \frac{0{,}0116 + 0{,}0113}{2 \cdot 1!}\, 0{,}19 + \frac{-0{,}0003}{2!} \cdot 0{,}19^2 = \\ = 0{,}7568809.$$

Ahora vamos a estimar los errores del método con ayuda de las fórmulas (13) y (7), respectivamente:

$$\Delta_1(B_1) < \frac{0{,}76}{2!}\left(\frac{\pi}{180}\right)^2 \cdot 0{,}63 \cdot 0{,}37 = 0{,}27 \cdot 10^{-4};$$

$$\Delta_1(S_2) < \frac{1}{3!}\left(\frac{\pi}{180}\right)^3 \cdot 0{,}19\,|\,0{,}19^2 - 1^2\,| = 0{,}2 \cdot 10^{-6}.$$

Teniendo en cuenta errores en los valores de la función y de sus diferencias finitas, citados en la tabla 7.9, obtenemos las magnitudes de los errores de cálculo:

$$\Delta_2(B_1) = 0{,}5 \cdot 10^{-4} + 1 \cdot 10^{-4} \cdot 0{,}13 = 0{,}63 \cdot 10^{-4};$$

$$\Delta_2(S_2) = 0{,}5 \cdot 10^{-4} + 1 \cdot 10^{-4} \cdot 0{,}19 + 1 \cdot 10^{-4} \cdot 0{,}19^2 = \\ = 0{,}73 \cdot 10^{-4}.$$

Al redondear los valores de B_1 (0,63) y S_2 (0,19) hasta cuatro cifras, obtenemos los siguientes errores de redondeo: $\Delta_3(B_1) = 0{,}08 \times \times 10^{-4}$; $\Delta_3(S_2) = 0{,}11 \cdot 10^{-4}$.

Uniendo todos los errores hallados tenemos finalmente sen $48{,}63^\circ = 0{,}7509 \pm 0{,}0001$; sen $49{,}19^\circ = 0{,}7569 \pm 0{,}0001$. ▲

Observación. En algunos casos para minimizar el error total resulta racional la utilización de un polinomio interpolador cuyo grado sea superior al orden de exactitud de la tabla de diferencias finitas. Esto se explica por lo siguiente. Con el aumento del grado del polinomio el error de cálculo naturalmente crece. Al mismo tiempo la disminución del error del método puede ser tan brusco que originará también la disminución del error total.

§ 7.9. Primero y segundo polinomios interpoladores de Newton

Si el punto de interpolación x^* está al comienzo de la tabla o a su fin, no siempre es posible elegir una cantidad suficiente de nodos a la izquierda de x^* y a su derecha para construir las diferencias finitas necesarias. En este caso se utilizan formas especiales del polinomio interpolador.

Supongamos que el punto x^* está situado en la proximidad del primer nodo de la red $x_0, x_1, \ldots, x_k, \ldots$ Consideremos la variable t, definida por la relación (1) del § 7.8, y construyamos el polinomio interpolador en la forma siguiente:

$$P_k(x) = P_k(x_0 + th) = a_0 + \frac{a_1}{1!}t + \frac{a_2}{2!}t(t-1) + \ldots$$
$$\ldots + \frac{a_k}{k!}t \ldots (t-k+1). \tag{1}$$

Los coeficientes desconocidos a_i se determinan de las condiciones de coincidencia del polinomio P_k y de la función f en los nodos x_i. En este caso recuérdese que al nodo x_i le corresponde la magnitud $t = i$. Ahora bien, para determinar los coeficientes a_i obtenemos el sistema de ecuaciones lineales

$$P_k(x_0 + ih) = f_i \quad (i = 0, 1, \ldots, k). \tag{2}$$

La estructura de este sistema es tal que a_0 se determina inmediatamente de la primera ecuación del sistema (2), a_1 se determina de la segunda ecuación para a_0 ya determinado, etc. En efecto, suponiendo $i = 0$, de la primera ecuación del sistema (2) obtenemos

$$P_k(x_0) = a_0 = f_0$$

y de la segunda ecuación para $i = 1$ tenemos

$$P_k(x_0 + h) = f_0 + \frac{a_1}{1!} \cdot 1 = f_1; \quad a_1 = \Delta f_0.$$

Continuando este proceso, como resultado de transformaciones poco complicadas obtenemos las siguientes expresiones para los coeficientes:

$$a_0 = f_0;\ a_i = \Delta^i f_0\ (i = 1, 2, \ldots, k). \tag{3}$$

Sustituyendo los valores hallados de los coeficientes en la igualdad (1), obtenemos el *primer polinomio interpolador de Newton* que suele designarse N_k^1:

$$N_k^1(t) = f_0 + \frac{\Delta f_0}{1!} t + \frac{\Delta^2 f_0}{2!} t(t-1) + \ldots$$

$$\ldots + \frac{\Delta^k f_0}{k!} t(t-1) \ldots (t-k+1). \tag{4}$$

En virtud de la fórmula (4) del § 7.5 el término residual respecto a la variable t puede ser representado en la forma

$$R_k = \frac{f^{(k+1)}(\xi)}{(k+1)!} h^{k+1} t(t-1) \ldots (t-k);\ \xi \in (x_0, x_k) \tag{5}$$

y la estimación del error del valor aproximado de $N_k^1(t)$ (el error del método), en la forma

$$\Delta_1 = |f(x) - N_k^1(t)| \leqslant \frac{M_{k+1}}{(k+1)!} h^{k+1} |t(t-1) \ldots (t-k)|, \tag{6}$$

donde $M_{k+1} = \max\limits_{[x_0, x_k]} |f^{(k+1)}(x)|$.

Supongamos ahora que el punto x^* está situado cerca del último nodo de la red $\ldots x_{-k}, \ldots, x_{-1}, x_0$. Volviendo a utilizar la variable t, definida por la relación (1) del § 7.8, para esta red construimos el polinomio interpolador en la forma siguiente:

$$P_k(x) = P_k(x_0 + th) = a_0 + \frac{a_1}{1!} t +$$

$$+ \frac{a_2}{2!} t(t+1) + \ldots + \frac{a_k}{k!} t \ldots (t+k-1). \tag{7}$$

Los coeficientes desconocidos a_i se determinan de las condiciones de coincidencia del polinomio P_k y de la función f en los nodos x_i. En este caso obsérvese que al nodo x_{-i} le corresponde el valor de $t = -i$. Ahora bien, para hallar los coeficientes obtenemos el sistema de ecuaciones lineales

$$P_k(x_0 - ih) = f_i \quad (i = 0, 1, \ldots, k). \tag{8}$$

La estructura de este sistema es tal que a_0 se calcula inmediatamente de la primera ecuación del sistema (8) y a_1, de la segunda ecuación para a_0 ya determinado, etc. En efecto, suponiendo $i = 0$, de la primera ecuación del sistema (8) hallamos

$$P_k(x_0) = a_0 = f_0$$

y de la segunda ecuación para $i = 1$ tenemos

$$P_k(x_0 - h) = f_0 + \frac{a_1}{1!} 1 = f_{-1};\ a_1 = \nabla f_0.$$

Continuando este proceso, como resultado de transformaciones poco complicadas obtenemos las siguientes expresiones para los coeficientes:

$$a_0 = f_0;\ a_i = \nabla^i f_0 \ (i = 1, 2, \ldots, k). \tag{9}$$

Sustituyendo los valores hallados de los coeficientes en la igualdad (7), obtenemos el *segundo polinomio interpolador de Newton* que suele designarse N_k^{II}:

$$N_k^{II}(t) = f_0 + \frac{\nabla f_0}{1!} t + \frac{\nabla^2 f_0}{2!} t(t+1) + \ldots$$
$$\ldots + \frac{\nabla^k f_0}{k!} t(t+1) \ldots (t+k-1) \tag{10}$$

con el término residual

$$R_k = \frac{f^{(k+1)}(\xi)}{(k+1)!} h^{k+1} t(t+1) \ldots (t+k); \tag{11}$$
$$\xi \in (x_{-k},\ x_0)$$

y con la estimación del error del valor aproximado

$$\Delta_1 = | f(x) - N^{II}(t) | \leqslant \frac{M_{k+1}}{(k+1)!} h^{k+1} | t(t+1) \ldots (t+k) |, \tag{12}$$

donde $M_{k+1} = \max\limits_{[x_{-k}, x_0]} | f^{(k+1)}(x) |$.

Las fórmulas (4) y (10) se llaman con frecuencia *fórmulas de interpolación de Newton para la interpolación hacia adelante y hacia atrás*, respectivamente.

Ejemplo. Plantear los polinomios interpoladores respectivos y calcular en los puntos $x^* = 0{,}63$ y $x_2^* = 1{,}35$ los valores de la función $f = 3^x$ representada en forma de la tabla siguiente que contiene los valores f_i con cuatro cifras justas en amplio sentido:

x	0,50	0,75	1,00	1,25	1,50
f	1,732	2,280	3,000	3,948	5,196

Estimar el error del resultado.

△ Vamos a completar la tabla dada por los valores de las diferencias finitas colocando en la última fila los valores de los errores de las diferencias finitas correspondientes.

Puesto que la cuarta diferencia finita coincide con el propio error no conviene aproximar, desde el punto de vista del error de cálculo, la función representada haciendo uso de un polinomio cuyo grado sea superior al tercero.

Luego, puesto que $x_1^* = 0{,}63$ está situado al inicio de la tabla y $x_2^* = 1{,}35$ al fin de ésta, para calcular $f_1^* = 3^{0,63}$ es necesario utilizar el primer polinomio interpolador de Newton y para calcular $f_2^* = 3^{1,35}$, el segundo.

Tabla 7.10

x	f	f^1	f^2	f^3	f^4
0,50 0,75 1,00 1,25 1,50	1,732 2,280 3,000 3,948 5,196	548 720 948 1248	172 228 300	56 72	16
	0,001	2	4	8	16

Así pues, suponiendo $x_0 = 0{,}5$, calculamos $t_1^* = (x_0^* - x_0)/h = (0{,}63 - 0{,}5)/0{,}25 = 0{,}52$. Sustituyendo el valor obtenido de t_1^* en la expresión (4) para el primer polinomio interpolador de Newton y utilizando las magnitudes de las diferencias finitas dadas en la tabla 7.10, tenemos

$$N_3^{\mathrm{I}}(0{,}52) = 1{,}732 + \frac{0{,}548}{1!}\cdot 0{,}52 + \frac{0{,}172}{2!}\cdot 0{,}52\,(-0{,}48) + \\ + \frac{0{,}056}{3!}\cdot 0{,}52\,(-0{,}48)\,(-1{,}48) = 1{,}9989420.$$

Análogamente, suponiendo $x_0 = 1{,}50$, calculamos $t_2^* = (1{,}35 - 1{,}50)/0{,}25 = -0{,}60$ y utilizando la expresión (10) para el segundo polinomio interpolador de Newton, obtenemos

$$N_3^{\mathrm{II}}(-0{,}60) = 5{,}196 + \frac{1{,}248}{1!}\cdot(-0{,}60) + \frac{0{,}300}{2!}\cdot(-0{,}60)\cdot 0{,}40 + \\ + \frac{0{,}072}{3!}\cdot(-0{,}60)\cdot 0{,}40\cdot 1{,}40 = 4{,}407168.$$

Vamos a estimar el error del método con ayuda de las fórmulas (6) y (12), respectivamente:

$$\Delta_1(N_3^{\mathrm{I}}) < \frac{4\ln^4 3}{4!}\cdot 0{,}25^4\cdot 0{,}52\cdot 0{,}48\cdot 1{,}48\cdot 2{,}48 = 0{,}0009;$$

$$\Delta_1(N_3^{\mathrm{II}}) < \frac{5{,}2\ln^4 3}{4!}\cdot 0{,}25^4\cdot 0{,}60\cdot 0{,}40\cdot 1{,}40\cdot 2{,}40 = 0{,}001.$$

Teniendo en cuenta los errores de f^k dados en la tabla 7.10, estimamos los errores de cálculo:

$\Delta_2(N_3^{\mathrm{I}}) < 0{,}001 + 0{,}0011 + 0{,}0005 + 0{,}0005 = 0{,}0031;$
$\Delta_2(N_3^{\mathrm{II}}) < 0{,}001 + 0{,}0012 + 0{,}0005 + 0{,}0005 = 0{,}0032.$

Redondeando los valores de $N_3^{\rm I}$ (0,52) y $N_3^{\rm II}$ (—0,60) hasta cuatro cifras, obtenemos los siguientes errores de redondeo:

$$\Delta_3 (N_3^{\rm I}) = 0{,}06\cdot 10^{-3}; \quad \Delta_3 (N_3^{\rm II}) = 0{,}2\cdot 10^{-3}.$$

Uniendo todos los errores hallados, tenemos finalmente $3^{0,63} = 1{,}999 \pm 0{,}005$; $3^{1,35} = 4{,}407 \pm 0{,}005$. ▲

§ 7.10. Diferencias divididas

En los párrafos precedentes hemos considerado diferentes formas del polinomio interpolador para la red uniforme de nodos. Los coeficientes de los polinomios construidos se han determinado con ayuda de las diferencias finitas. En este párrafo volvemos a examinar el caso en que los valores de la función se dan en los nodos no equidistantes. En este caso en vez de diferencias finitas se consideran las *diferencias divididas* que son, en cierto sentido, un análogo del concepto de derivada y se definen del modo siguiente.

Supongamos que la función $y = f(x)$ se da por sus valores $y_0 = f_0 = f(x_0)$, $y_1 = f_1 = f(x_1)$, ..., $y_k = f_k = f(x_k)$, ... en los nodos x_i de una red arbitraria Λ_n. *Las diferencias divididas de orden nulo de* $f(x_i)$ coinciden con los valores que la función tiene en los puntos x_i. Se llaman *diferencias divididas de primer orden* las relaciones

$$f(x_0;\ x_1) = \frac{f_1 - f_0}{x_1 - x_0}; \quad f(x_1;\ x_2) = \frac{f_2 - f_1}{x_2 - x_1}; \quad \ldots$$

$$\ldots; \quad f(x_i;\ x_{i+1}) = \frac{f_{i+1} - f_i}{x_{i+1} - x_i}; \quad \ldots$$

Se denominan *diferencias divididas de segundo orden* las relacionese

$$f(x_0;\ x_1;\ x_2) = \frac{f(x_1;\ x_2) - f(x_0;\ x_1)}{x_2 - x_0};$$

$$f(x_1;\ x_2;\ x_3) = \frac{f(x_2;\ x_3) - f(x_1;\ x_2)}{x_3 - x_1};$$

$$f(x_i;\ x_{i+1};\ x_{i+2}) = \frac{f(x_{i+1};\ x_{i+2}) - f(x_i;\ x_{i+1})}{x_{i+1} - x_i}.$$

En el caso general la *diferencia dividida de k-ésimo orden* se define por la diferencia dividida de $(k-1)$-ésimo orden con ayuda de la fórmula

$$f(x_i;\ x_{i+1};\ \ldots;\ x_{i+k}) = \frac{f(x_{i+1};\ \ldots\ x_{i+k}) - f(x_i;\ \ldots,\ x_{i+k-1})}{x_{i+k} - x_i}.$$

Para las relaciones de diferencia se utiliza también otra designación:

$$f(x_i;\ x_{i+1};\ \ldots;\ x_{i+k}) \equiv [x_i,\ x_{i+1};\ \ldots;\ x_{i+k}].$$

Es cómodo escribir las diferencias divididas en forma de una tabla, análogamente a como lo hemos hecho para las diferencias finitas centrales (véase la tabla 7.11).

Tabla 7.11

x_i	$y_i=f(x_i)$	$[x_i;\ x_{i+1}]$	$[x_i;\ x_{i+1};\ x_{i+2}]$	$[x_i;\ x_{i+1};\ x_{i+2};\ x_{i+3}]$	$[x_i;\ x_{i+1};\ x_{i+2};\ x_{i+3};\ x_{i+4}]$
x_0	y_0				
		$[x_0;\ x_1]$			
x_1	y_1		$[x_0;\ x_1;\ x_2]$		
		$[x_1;\ x_2]$		$[x_0;\ x_1;\ x_2;\ x_3]$	
x_2	y_2		$[x_1;\ x_2;\ x_3]$		$[x_0;\ x_1;\ x_2;\ x_3;\ x_4]$
		$[x_2;\ x_3]$		$[x_1;\ x_2;\ x_3;\ x_4]$	
x_3	y_3		$[x_2;\ x_3;\ x_4]$		
		$[x_3;\ x_4]$			
x_4	y_4				

Ejemplo 1. Hacer la tabla de diferencias divididas para la función $y = f(x)$ representada en forma de la tabla siguiente:

x	0	1	5	10
y	10	20	100	1100

Haciendo uso inmediatamente de la definición hallamos las diferencias divididas de primer orden:

$$[x_0;\ x_1] = \frac{y_1 - y_0}{x_1 - x_0} = \frac{20-10}{1-0} = 10;$$

$$[x_1;\ x_2] = \frac{y_2 - y_1}{x_2 - x_1} = \frac{100-20}{5-1} = \frac{80}{4} = 20;$$

$$[x_2;\ x_3] = \frac{y_3 - y_2}{x_3 - x_2} = \frac{1100-100}{10-5} = \frac{1000}{5} = 200.$$

Análogamente obtenemos las diferencias divididas de segundo orden

$$[x_0;\ x_1;\ x_2]=\frac{[x_1;\ x_2]-[x_0;\ x_1]}{x_2-x_0}=\frac{20-10}{5-0}=2;$$

$$[x_1;\ x_2;\ x_3]=\frac{[x_2;\ x_3]-[x_1;\ x_2]}{x_3-x_1}=\frac{200-20}{10-1}=20.$$

La diferencia dividida de tercer orden

$$[x_0;\ x_1;\ x_2;\ x_3]=\frac{[x_1;\ x_2;\ x_3]-[x_0;\ x_1;\ x_2]}{x_3-x_0}=\frac{20-2}{10}=1,8.$$

Representemos los resultados de cálculos en forma de la tabla de diferencias divididas:

x	y	$[x_i;\ x_{i+1}]$	$[x_i;\ x_{i+1};\ x_{i+2}]$	$[x_i;\ x_{i+1};\ x_{i+2};\ x_{i+3}]$
0	10			
1	20	10	2	1,8
5	100	20	20	
10	1100	200		

Vamos a citar algunas propiedades de las diferencias divididas.

1°. *Las diferencias divididas de todos los órdenes son una combinación lineal de los valores $f_i = f(x_i)$, es decir, es válida la fórmula siguiente:*

$$f(x_0;\ \ldots,\ x_k)=\sum_{i=0}^{k}\frac{f_i}{\prod\limits_{\substack{j=0\\ j\neq i}}^{k}(x_i-x_j)}\ . \tag{2}$$

□ Para $k=0$ obtenemos la identidad $f(x_0)=f_0$; para $k=1$ tenemos $f(x_0;\ x_1)=\dfrac{f(x_0)-f(x_1)}{x_0-x_1}$ lo que es la definición de la diferencia dividida $f(x_0;\ x_1)$.

Realicemos la demostración posterior haciendo uso de la inducción. Supongamos que la igualdad (2) es válida para todos los $k\leqslant n$. Entonces

$$f(x_0;\ \ldots;\ x_{n+1})=\frac{f(x_0;\ \ldots;\ x_n)-f(x_1;\ \ldots;\ x_{n+1})}{x_0-x_{n+1}}=$$

$$=\frac{1}{x_0-x_{n+1}}\left(\sum_{i=0}^{n}\frac{f_i}{\prod\limits_{\substack{j=0\\ j\neq i}}^{n}(x_i-x_j)}-\sum_{i=1}^{n+1}\frac{f_i}{\prod\limits_{\substack{j=1\\ j\neq i}}^{n+1}(x_i-x_j)}\right)=$$

$$= \frac{f_0}{\prod\limits_{j=1}^{n+1}(x_0-x_j)} + \sum_{i=1}^{n} \frac{f_i}{\prod\limits_{\substack{j=0\\j\neq i}}^{n+1}(x_i-x_j)} +$$

$$+ \frac{f_{n+1}}{\prod\limits_{j=0}^{n}(x_{n+1}-x_j)} = \sum_{i=0}^{n+1} \frac{f_i}{\prod\limits_{\substack{j=0\\j\neq i}}^{n+1}(x_i-x_j)}.$$

Ahora bien, la relación (2) es justa también para $k = n+1$. ■

2°. *La diferencia dividida es la función simétrica de los propios argumentos, o sea, no cambia para toda permutación de los mismos.*

□ Para toda permutación de los argumentos $x_0, \ldots, x_k$ en el segundo miembro de la relación (2) los sumandos correspondientes cambian sólo de lugar y la suma, como es evidente, no cambia. Por consiguiente, tampoco cambia el primer miembro de la relación (2), o sea, $f(x_0; \ldots; x_k)$. ■

3°. *La diferencia dividida satisface la igualdad*

$$(af+bg)(x_0; \ldots; x_k) = af(x_0; \ldots; x_k) + bf(x_0; \ldots; x_k), \quad (3)$$

donde a y b son constantes.

Esto se deduce inmediatamente de la relación (2) en virtud de que su segundo miembro es lineal respecto a f_i.

La segunda propiedad expresa la relación existente entre la diferencia dividida $f(x_0; \ldots; x_k)$ y la derivada $f^{(k)}(x)$.

4°. *Si los nodos $x_0, \ldots, x_k$ pertenecen al segmento $[a, b]$ y la función $f(x)$ tiene sobre $[a, b]$ una derivada continua de k-ésimo orden, entonces existe tal punto $\xi \in (a, b)$ que*

$$f(x_0; \ldots; x_k) = \frac{1}{k!} f^{(k)}(\xi). \quad (4)$$

De esta propiedad se desprende un corolario simple. Sea $f(x) = a_0x^k + a_1x^{k-1} + \ldots + a_k$ el polinomio de k-ésimo grado. Entonces, evidentemente, $f^{(k)}(x) = a_0k!$ y la relación (4) da para la diferencia dividida el valor

$$f(x_0; \ldots x_k) = \frac{1}{k!} a_0k! = a_0. \quad (5)$$

Así pues, en todo polinomio de k-ésimo grado las diferencias divididas de k-ésimo orden son iguales a una magnitud constante, es decir, al coeficiente del grado superior del polinomio. Las diferencias divididas de órdenes superiores (mayores que k) son, evidentemente, iguales a cero.

Ejemplo 2. Verifiquemos la validez de la propiedad 1° utilizando cuatro valores de la función $y = f(x)$: $y_i = f(x_i)$ $(i = 0, 1, 2, 3)$.

Las diferencias divididas primeras las determinamos así:

$$f(x_0;\ x_1)=\frac{y_0-y_1}{x_0-x_1}\ ;\ f(x_1;\ x_2)=\frac{y_1-y_2}{x_1-x_2}\ ;$$

$$f(x_2;\ x_3)=\frac{y_2-y_3}{x_2-x_3}.$$

Hallamos las diferencias divididas segundas:

$$f(x_0;\ x_1;\ x_2)=\frac{1}{x_0-x_2}\left(\frac{y_0-y_1}{x_0-x_1}-\frac{y_1-y_2}{x_1-x_2}\right)=$$

$$=\frac{y_0}{(x_0-x_1)(x_0-x_2)}+\frac{y_1}{(x_1-x_0)(x_1-x_2)}+\frac{y_2}{(x_2-x_0)(x_2-x_1)}\ ;$$

$$f(x_1;\ x_2;\ x_3)=\frac{1}{x_1-x_3}\left(\frac{y_1-y_2}{x_1-x_2}-\frac{y_2-y_3}{x_2-x_3}\right)=$$

$$=\frac{y_1}{(x_1-x_2)(x_1-x_3)}+\frac{y_2}{(x_2-x_1)(x_2-x_3)}+\frac{y_3}{(x_3-x_1)(x_3-x_2)}$$

lo que está en plena correspondencia con la propiedad 1°. Análogamente se puede mostrar que la diferencia dividida tercera

$$f(x_0;\ x_1;\ x_2;\ x_3)=\frac{y_0}{(x_0-x_1)(x_0-x_2)(x_0-x_3)}+$$

$$+\frac{y_1}{(x_1-x_0)(x_1-x_2)(x_1-x_3)}+\frac{y_2}{(x_2-x_0)(x_2-x_1)(x_2-x_3)}+$$

$$+\frac{y_3}{(x_3-x_0)(x_3-x_1)(x_3-x_2)}.$$

Calculemos la diferencia finita tercera, utilizando los valores de la función y del argumento, dados en el ejemplo 1.

$$[x_0;\ x_1;\ x_2;\ x_3]=\frac{10}{(-1)(-5)(-10)}+\frac{20}{1\cdot(-4)(-9)}+$$

$$+\frac{100}{5\cdot 4\cdot(-5)}+\frac{1100}{10\cdot 9\cdot 5}=-\frac{1}{5}+\frac{5}{9}-1+\frac{22}{9}=1{,}8$$

lo que coincide con el valor obtenido en el ejemplo 1.

Ejemplo 3. Verifiquemos ahora para la función $y=f(x)$ la validez de la propiedad 2°. Mostremos, por ejemplo, que $f(x_0;\ x_1;\ x_2)=$ $=f(x_1;\ x_0;\ x_2)$.

En efecto, según la definición de la diferencia dividida tenemos

$$f(x_1;\ x_0;\ x_2)=\frac{1}{x_1-x_2}\left(\frac{y_1-y_0}{x_1-x_0}-\frac{y_0-y_2}{x_0-x_2}\right)=$$

$$=\frac{y_1}{(x_1-x_0)(x_1-x_2)}+\frac{y_0}{(x_0-x_1)(x_0-x_2)}+\frac{y_2}{(x_2-x_0)(x_2-x_1)}.$$

Hemos obtenido el mismo resultado que en el ejemplo 2 (con el cambio del orden de los sumandos).

Ejemplo 4. Vamos a ilustrar la propiedad 3°. Junto con la función $y=f(x)$ (véase el ejemplo 2) consideremos la función $z=g(x)$ y sus valores $z_i=g(x_i)$ $(i=0,\ 1,\ 2,\ 3)$ dados en los mismos nodos que y_i. Planteemos la combinación lineal $u(x) = af(x) + bg(x)$, donde a y b son las constantes. Calculemos ahora $u(x_0;\ x_1;\ x_2)$. Al igual que en el ejemplo 2, hallamos

$$u(x_0;\ x_1;\ x_2) = \frac{ay_0 + bz_0}{(x_0 - x_1)(x_0 - x_2)} + \frac{ay_1 + bz_1}{(y_1 - x_0)(x_1 - x_2)} + \\ + \frac{ay_2 + bz_2}{(x_2 - x_0)(x_2 - x_1)}.$$

Agrupando por separado los primeros y segundos sumandos de cada uno de tres términos de la suma obtenemos

$$u(x_0;\ x_1;\ x_2) = af(x_0;\ x_1;\ x_2) + bg(x_0;\ x_1;\ x_2)$$

que es lo que se necesitaba mostrar.

§ 7.11. Polinomio interpolador de Newton para una red arbitraria de nodos

Utilizando la forma de Lagrange, representemos el polinomio interpolador en la forma siguiente:

$$L_0(x) + (L_1(x) - L_0(x)) + \ldots + (L_n(x) - L_{n-1}(x)).$$

Aquí $L_0(x) = f(x_0)$; $L_k(x)$ $(k = 1, 2, \ldots, n)$ son los polinomios interpoladores en la forma de Lagrange construidos sobre los nodos $x_0, x_1, \ldots, x_k$. Consideremos las diferencias

$$L_k - L_{k-1} = \sum_{i=0}^{k} f_i \frac{\omega_k(x)}{(x - x_i)\,\omega'_k(x_i)} - \\ - \sum_{i=0}^{k-1} f_i \frac{\omega_{k-1}(x)}{(x - x_i)\,\omega'_{k-1}(x_i)} = \sum_{i=0}^{k-1} f_i \frac{\omega_{k-1}(x)}{\omega'_k(x_i)} + \\ + f_k \frac{\omega_{k-1}(x)}{\omega'_k(x_k)} = \omega_{k-1}(x) \sum_{i=0}^{k} \frac{f_i}{\omega'_k(x_i)}.$$

Ahora bien, utilizando la fórmula (2) del § 7.10, obtenemos

$$L_k - L_{k-1} = \varphi_{k-1}(x)\, f(x_0;\ \ldots;\ x_k) \tag{1}$$

y el polinomio interpolador toma la forma

$$N_n(x) = f_0 + (x - x_0)\, f(x_0;\ x_1) + \ldots \\ \ldots + (x - x_0) \ldots (x - x_{n-1})\, f(x_0;\ \ldots;\ x_n). \tag{2}$$

Esta forma se llama *polinomio interpolador de Newton con diferencias divididas.*

La expresión para el error tiene, evidentemente, la misma forma que en el caso del polinomio de Lagrange [véanse las fórmulas (8) y (9) del § 7.5].

Ejemplo 1. Construir sobre los nodos $x_0 = 0$, $x_1 = 1/3$, $x_2 = 1$ el polinomio interpolador de Newton para la función $f = \text{sen}\,(\pi x/2)$.

△ Teniendo en cuenta que $f_0 = 0$; $f_1 = 0{,}5$; $f_2 = 1$ planteemos las diferencias divididas necesarias:

$$f(x_0;\ x_1) = \frac{0-\frac{1}{2}}{0-\frac{1}{3}} = \frac{3}{2};\ f(x_1,\ x_2) = \frac{\frac{1}{2}-1}{\frac{1}{3}-1} = \frac{3}{4};$$

$$f(x_0;\ x_1;\ x_2) = \frac{\frac{3}{2}-\frac{3}{4}}{0-1} = -\frac{3}{4}.$$

Sustituyendo los valores obtenidos en la fórmula (2), para $n = 2$ tenemos

$$N_2(x) = 0 + \frac{3}{2}(x-0) - \frac{3}{4}(x-0)\left(x-\frac{1}{3}\right).$$

El polinomio construido debe ser, evidentemente, idéntico al de Lagrange construido en el ejemplo 1 del § 7.6; proponemos que el lector se cerciore de ello de por sí. ▲

Nótese que en la forma (2) del polinomio interpolador los nodos x_i llevan impuesta la única condición, es decir, la falta de coincidencia de los mismos. Por eso los nodos pueden ser numerados en un orden arbitrario. Por ejemplo, con índice «0» se designa frecuentemente el último nodo de la tabla, por x_1 se toma el penúltimo nodo designándolo x_{-1}, etc. En este caso el polinomio (2) toma la forma

$$\begin{aligned} N_n(x) &= f_0 + (x - x_0)\, f(x_0;\ x_{-1}) + \ldots \\ \ldots &+ (x - x_0) \ldots (x - x_{-n+1})\, f(x_0;\ \ldots\ ;\ x_{-n}) \end{aligned} \tag{3}$$

y se llama *polinomio de Newton para interpolar hacia atrás.*

A título de ilustración de lo dicho vamos a resolver el ejemplo 1 para $x_0 = 1$, $x_{-1} = 1/3$; $x_{-2} = 0$. Según la fórmula (3) obtenemos

$$N_2(x) = 1 + \frac{3}{4}(x-1) - \frac{3}{4}(x-1)\left(x-\frac{1}{3}\right).$$

Nótese que las diferencias divididas necesarias han sido tomadas del ejemplo 1, con la particularidad de que hemos utilizado las propiedades de su simetría respecto a los propios argumentos.

De la comparación de las formas de Lagrange y de Newton para el polinomio interpolador se deduce que la utilización de la representación en la forma de Lagrange puede recomendarse, en primer lugar, en las investigaciones teóricas, por ejemplo al estudiar la cuestión acerca de la convergencia de $P_n\ (f, x)$ hacia f cuando $n \to \infty$; en segundo lugar, al interpolar varias funciones sobre una misma red de nodos, puesto que en este caso se puede calcular una vez dos multiplicadores de Lagrange l_i y emplearlos para la interpolación de todas las funciones.

La representación en la forma de Newton resulta más cómoda en los cálculos prácticos. En efecto, el número de nodos a utilizar y el grado del polinomio interpolador con frecuencia no están conocidos de antemano y al pasar de n nodos a $n+1$ nodos en la forma de Newton se añade sólo un término que no es sino una corrección para el valor ya calculado. En cambio, la adición de un sumando más en la forma de Lagrange se acompaña del recálculo completo del resultado antes obtenido. Además, en la práctica de cálculos la interpolación suele realizarse sobre un pequeño segmento cuya longitud $h < 1$. En este caso los sumandos de la forma de Newton tienen el orden $h^0, h^1, h^2, \ldots$, o sea, están situados en el orden de decrecimiento lo que es útil al determinar la exactitud del resultado de interpolación.

§ 7.12. Interpolación práctica en las tablas

Por regla general, al interpolar en las tablas se utiliza la interpolación lineal o la cuadrática.

En caso de la interpolación lineal el valor que la función tiene en un punto, distinto de los nodos de interpolación, se determina por dos valores conocidos de la función a tabular $y_i = f(x_i)$, $y_{i+1} = f(x_{i+1})$ en los nodos de interpolación x_i y x_{i+1} entre los cuales está situado el valor del argumento x que nos interesa.

La fórmula de interpolación de Lagrange para la interpolación lineal tomará la forma

$$L_1(x) = y_i \frac{x - x_{i+1}}{x_i - x_{i+1}} + y_{i+1} \frac{x - x_i}{x_{i+1} - x_i}$$

y la primera fórmula de interpolación de Newton

$$N_1(x) = y_i + \frac{\Delta y_i}{h}(x - x_i),$$

donde $\Delta y_i = y_{i+1} - y_i$ es la diferencia finita primera en el punto x_i y $h = x_{i+1} - x_i$, el paso de interpolación.

Así pues, para obtener el valor aproximado de la función $y(x)$ con ayuda de la fórmula de Newton es suficiente añadir una corrección, igual a $\Delta y_i (x - x_i) h^{-1}$ al valor tabular y_i.

Ejemplo 1. Calcular cuántos grados se contiene en una medida en radianes igual a 0,222.

△ Hagamos uso de la tabla

Radianes	Grados
0,22	12,605
23	13,178
24	13,751

Para la interpolación lineal basta considerar los datos contenidos en las primeras dos filas. Planteemos la diferencia tabular

$$\Delta y_i = y_{i+1} - y_i = 13,178 - 12,605 = 0,573.$$

El paso de la tabla $h = 0,01$; $x - x_i = 0,222 - 0,220 = 0,002$. Calculamos la corrección

$$\frac{\Delta y_i}{h}(x - x_i) = \frac{0,573}{0,01} \cdot 0,002 = 0,1146$$

y añadímosla al valor tabular

$$y = 12,605 + 0,1146 = 12,7196.$$

Vamos a estimar el error del valor obtenido. Utilizando la fórmula (8) del § 7,5 para el error del método, tenemos

$$\Delta_1 = \frac{M_2}{2!} |(0,222 - 0,22)(0,222 - 0,23)|.$$

Puesto que la función a interpolar es $y = (180x/\pi)$, entonces $M_2 = 0$ y $\Delta_1 = 0$.

Hallamos el error de cálculo, teniendo en cuenta que el error de los datos iniciales es de 0,0005:

$$\Delta_2 = 0,0005 + \frac{0,001}{0,01} \cdot 0,002 = 0,0007.$$

Redondeando el valor $y\,(0,222)$ hasta tres cifras después de la coma, obtenemos $\Delta_3 = 0,0004$.

Sumando todos los errores hallados, tenemos finalmente $y\,(0,222) = 12,720 \pm 0,0011$. ▲

En caso de la interpolación cuadrática se utilizan tres valores de la función tabulada $y_{-1} = f(x_{-1})$, $y_0 = f(x_0)$, $y_1 = f(x_1)$. En este caso el polinomio interpolador se construye en la forma de Lagrange (para los nodos no equidistantes) o bien en forma de Newton, cuando el punto de interpolación está situado más próximo a x_{-1} o a x_1 que a x_0, o bien en la forma de Stirling, cuando el punto de interpolación está cerca del nodo x_0.

Ejemplo 2. Para la función $y = \ln x$ según la tabla dada de sus valores

x	2	3	5
y	0,6931	1,0986	1,6094

construir el polinomio interpolador de Newton y obtener la estimación uniforme del error sobre el segmento [2, 5].

△ Ante todo, obtenemos las diferencias divididas:

$$f(2;\ 3) = \frac{1,0986 - 0,6931}{3 - 2} = 0,4055;$$

$$f(3;\ 5) = \frac{1,6092 - 1,0986}{5 - 3} = 0,2554;$$

$$f(2;\ 3;\ 5) = \frac{0,2554 - 0,4055}{5 - 2} = -0,0500.$$

Sustituyendo los valores hallados en la fórmula (2) del § 7.11, para $n = 2$, tenemos

$$N_2(x) = 0{,}6931 + 0{,}4055\,(x-2) - 0{,}0500\,(x-2)(x-3).$$

Utilizando la estimación (8) del § 7.5, para el error del método tenemos

$$\Delta_1 = \frac{M_3}{3!} \max_{[2,\,5]} |(x-2)(x-3)(x-5)|.$$

Luego hallamos

$$M_3 = \max_{[2,\,5]} \left|\frac{2}{x^3}\right| = \frac{1}{4}; \quad \max_{[2,\,5]} |(x-2)(x-3)(x-5)| \approx 2{,}2.$$

Ahora bien, $\Delta_1 = 0{,}1$.

Es evidente que el error de cálculo será despreciablemente pequeño en comparación con el del método. Por eso el error de interpolación máximo posible es de 0,1. ▲

§ 7.13. Método de iteración-interpolación de Aitken

En los casos en que falta la necesidad de obtener una expresión analítica aproximada de la función $f(x)$, representada tabularmente y se requiere sólo determinar el valor de esta función en cierto punto x^* distinto de los nodos de interpolación es conveniente utilizar el **método de iteración-interpolación de Aitken**. Dicho con propiedad este método consiste en interpolación lineal sucesiva. El proceso de cálculo de $f(x^*)$ radica en lo siguiente. Numeremos los nodos de interpolación, por ejemplo, en el orden en que éstos se alejan de x^*, y hagamos la tabla siguiente.

Tabla 7.12

$\vdots$	$\vdots$				
x_4	P_0^4	$P_1^{2,\,4}$			
x_2	P_0^2	$P_1^{0,\,2}$	$P_2^{0\,24}$	$P_3^{0\,124}$	$\cdots$
x_0	P_0^0	$P_1^{1,\,0}$	$P_2^{0\,12}$	$P_3^{0\,123}$	
x_1	P_0^1	$P_1^{1,\,3}$	$P_2^{0\,13}$		
x_3	P_0^3				
$\vdots$	$\vdots$				

Aquí:

$$P_0^k = f(x_k);$$

$$P_1^{ij}(x) = f_i\,\frac{x-x_j}{x_i-x_j} + f_j\,\frac{x-x_i}{x_j-x_i} = \frac{1}{x_j-x_i}\begin{vmatrix} x-x_i & P_0^i \\ x-x_j & P_0^j \end{vmatrix}$$

es un polinomio interpolador de grado no superior al primero, construido sobre los nodos x_i y x_j:

$$P_2^{ijk}(x)=\frac{1}{x_k-x_i}\begin{vmatrix} x-x_i & P_1^{ij} \\ x-x_k & P_1^{jk}\end{vmatrix}$$

es un polinomio interpolador de grado no superior al segundo, construido sobre los nodos x_i, x_j, x_k. Continuando este proceso, construimos el polinomio

$$P_n^{ij\ldots km}(x)=\frac{1}{x_m-x_i}\begin{vmatrix} x-x_i & P_{n-1}^{ij\ldots k}(x) \\ x-x_m & P_{n-1}^{j\ldots km}(x)\end{vmatrix}. \qquad (1)$$

Mostremos que si $P_{n-1}^{ij\ldots k}(x)$ y $P_{n-1}^{j\ldots km}(x)$ son los polinomios interpoladores construidos sobre los nodos $x_i, x_j, \ldots, x_k$ y $x_j, \ldots, x_k, x_m$, respectivamente, entonces $P_n^{ij\ldots km}(x)$ es el polinomio interpolador construido sobre los nodos $x_i, x_j, \ldots, x_k, x_m$.

Efectivamente, en primer lugar, $P_n^{ij\ldots km}(x)$ es un polinomio cuyo grado no supera n, lo que es evidente de la construcción de la fórmula (1). En segundo lugar, en todos los nodos x_p el polinomio $P_n^{ij\ldots km}(x)$ toma el valor correspondiente:

$$P_n^{ij\ldots km}(x_i)=\frac{-(x_i-x_m)f_i}{x_m-x_i}=f_i \quad (x_p=x_i);$$

$$P_n^{ij\ldots km}(x_m)=\frac{(x_m-x_i)f_m}{x_m-x_i}=f_m \quad (x_p=x_m);$$

$$P_n^{ij\ldots km}(x_p)=\frac{1}{(x_m-x_i)}((x_p-x_i)f_p-(x_p-x_m)f_p)=f_p.$$

Calculando sucesivamente con ayuda de la fórmula (1) los valores de $P_n^{01\ldots n}(x^*)$, los toman por las aproximaciones sucesivas de $f(x^*)$. El proceso de cálculo termina prácticamente cuando el valor absoluto de la diferencia de dos aproximaciones sucesivas llega a ser suficientemente pequeño.

Ejemplo 1. Calcular en el punto $x^*=6$ con exactitud hasta $\varepsilon=0{,}05$ el valor de la función $f=\ln x$ representada en la forma de la tabla:

x	1	2	4	5	8	10
y	0,00	0,69	1,39	1,61	2,08	2,30

△ Numeremos los nodos en el orden siguiente: $x_0=5$, $x_1=8$, $x_2=4$, $x_3=10$, $x_4=2$, $x_5=1$. Con ayuda de la fórmula (1) calculamos los valores de los polinomios interpoladores P_n (6):

$$P_1^{01}=\frac{1}{8-5}\cdot\begin{vmatrix} 6-5 & 1{,}61 \\ 6-8 & 2{,}08\end{vmatrix}=1{,}77;$$

$$P_1^{02}=\frac{1}{5-4}\cdot\begin{vmatrix} 6-4 & 1{,}39 \\ 6-5 & 1{,}61\end{vmatrix}=1{,}83;$$

$$P_2^{012}=\frac{1}{8-4}\cdot\begin{vmatrix} 6-4 & 1{,}83 \\ 6-8 & 1{,}77\end{vmatrix}=1{,}80.$$

Puesto que $|P_1^{02} - P_2^{012}| = 0{,}03 < 0{,}05$, cesamos los cálculos y suponemos $\ln 6 = 1{,}80 \pm 0{,}03$. ▲

Ejemplo 2. Haciendo uso del esquema de Aitken calcular con exactitud de hasta $0{,}5\cdot 10^{-4}$ el valor de sen 0,674 para la función $y = \operatorname{sen} x$ representada en la forma tabular:

$x_0 = 0{,}66$	$x_1 = 0{,}67$	$x_2 = 0{,}68$
$y_0 = 0{,}61312$	$y_1 = 0{,}62099$	$y_2 = 0{,}62879$

△ Según la fórmula (1) tenemos

$$P_1^{01}(0{,}674) = \frac{\begin{vmatrix} 0{,}674 - 0{,}68 & 0{,}61312 \\ 0{,}674 - 0{,}67 & 0{,}62090 \end{vmatrix}}{0{,}67 - 0{,}66} = 0{,}625730;$$

$$P_1^{12} = \frac{1}{0{,}68 - 0{,}67}\cdot\begin{vmatrix} 0{,}674 - 0{,}67 & 0{,}62099 \\ 0{,}674 - 0{,}68 & 0{,}625643 \end{vmatrix} = 0{,}625643;$$

$$P_2^{012} = \frac{1}{0{,}68 - 0{,}66}\cdot\begin{vmatrix} 0{,}674 - 0{,}66 & 0{,}625730 \\ 0{,}674 - 0{,}68 & 0{,}625643 \end{vmatrix} = 0{,}625676.$$

Por lo tanto, sen 0,674 = 0,62568 + 0,00004. ▲

§ 7.14. Optimización de los nodos de interpolación

Volvamos a considerar la estimación del error, expresada por la fórmula (8) del § 7.5. Supongamos ahora que no hay restricciones en la elección de la red Λ_n. Planteemos el problema acerca de la mejor elección de los nodos de interpolación. En la clase dada de funciones $C^{n+1}([a, b])$, partiendo de la estimación (8) del § 7.5, como mejores nodos de interpolación han de reconocerse los x_i para los cuales la expresión $\max\limits_{[a,b]} |\omega_n(x)|$ es mínima. La determinación de estos nodos se reduce de hecho a la búsqueda de las raíces del polinomio que se desvía mínimamente del cero sobre el segmento $[a, b]$. Según se sabe de la teoría de las funciones, tal polinomio se genera por *polinomios de Chébyshev de primer género* que se definen por las relaciones recurrentes siguientes:

$$T_0 = 1;\quad T_1 = x;\quad T_{n+1} = 2xT_n - T_{n-1},\ n > 0. \tag{1}$$

Consideremos las propiedades fundamentales de los polinomios de Chébyshev.

1°. *El término de mayor grado del polinomio T_{n+1} se obtiene del término de mayor grado del polinomio T_n ($n = 1, 2, \ldots$) multiplicando este último por $2x$.*

□ Tenemos:

$$T_2 = 2x\cdot x - 1 = 2x^2 - 1,$$
$$T_3 = 2x(2x^2 - 1) - x = 4x^3 - 3x.$$

Por eso el término de mayor grado del polinomio T_{n+1} es igual a $2^n x^{n+1}$. La forma general del polinomio T_{n+1} es tal:

$$T_{n+1} = 2^n x^{n+1} + \ldots \ ■$$

2°. Todos los *polinomios* $T_{2n}(x)$ *son funciones pares y* $T_{2n+1}(x)$, *funciones impares.*

□ Para $n = 0$ esto es evidente. Supongamos que esto es válido para cierto n. Entonces la función $2xT_{2n+1}(x)$ es par y, por lo tanto, $T_{2n+2} = 2xT_{2n+1}(x) - T_{2n}(x)$ también es una función par. Luego, la función $2xT_{2n+2}(x)$ es impar y por eso $T_{2n+3}(x) = 2xT_{2n+2}(x) - T_{2n+1}(x)$ es también una función impar. ■

3°. *Si* $x \in]-1, 1[$, *entonces los polinomios de Chébyshev tienen la expresión explícita siguiente*:

$$T_{n+1}(x) = \cos[(n+1)\arccos x]; \quad n \geqslant -1. \tag{2}$$

□ Mostremos que el segundo miembro de la igualdad (2) satisface la definición (1) del polinomio de Chébyshev. En efecto,

$$\{\cos[(n+1)\arccos x]\}_{n=-1} = 1 = T_0;$$
$$\{\cos[(n+1)\arccos x]\}_{n=0} = x = T_1.$$

Para demostrar el cumplimiento de la fórmula recurrente consideremos la relación trigonométrica evidente

$$\cos(n+1)\theta = 2\cos\theta\cos n\theta - \cos(n-1)\theta.$$

Suponiendo $\theta = \arccos x$, de donde $x = \cos\theta$, obtenemos $\cos[(n+1)\arccos x] = 2x\cos[n\arccos x] - \cos[(n-1)\arccos x]$. ■

4°. *Sobre el segmento* $[-1, 1]$ *los polinomios* $T_{n+1}(x)$ *tienen* $n+1$ *raíces diferentes*:

$$x_k = \cos\frac{2k+1}{2(n+1)}\pi \quad (k = 0, 1, \ldots, n). \tag{3}$$

□ Utilizando la expresión (2), para determinar las raíces del polinomio $T_{n+1}(x)$ obtenemos la ecuación

$$(n+1)\arccos x_k = \frac{\pi}{2} + \pi k \ (k = 0, 1, \ldots, n).$$

Despejando en esta ecuación x_k, llegamos a la relación (3). ■

5°. *Sobre el segmento* $[-1, 1]$ *es válida la desigualdad*

$$[T_{n+1}(x)] \leqslant 1. \tag{4}$$

Esto se desprende inmediatamente de la relación (2).

De la misma relación (2) determinemos todos los puntos x_m en los cuales el polinomio $T_{n+1}(x)$ alcanza sus valores extremos ± 1. Para esto es necesario que

$$(n+1)\arccos x_m = \pi m \ (m = 0, 1, \ldots, n+1)$$

y, por consiguiente,

$$x_m = \cos\frac{m}{n+1}\pi \ (m = 0, 1, \ldots, n+1). \tag{5}$$

Sustituyendo estos valores en la igualdad (2), obtenemos

$$T_{n+1}(x_m) - \cos m\pi = (-1)^m. \tag{6}$$

Esto quiere decir que los puntos en los cuales $T_{n+1} = 1$ y $T_{n+1} = -1$ alternan comenzando con $x_0 = 1$, donde $T_{n+1}(1) = 1$. Nótese una vez más que la desigualdad (4) es válida no para todas las x. Si $|x| > 1$, entonces $\arccos x$ no existe sobre el conjunto de los números reales.

Examinemos ahora los polinomios

$$T_{n+1}(x) = 2^{-n} T_{n+1}(x) = x^{n+1} + \dots \tag{7}$$

Son los polinomios que se desvían mínimamente del cero sobre el segmento $[-1, 1]$ lo que confirma el teorema siguiente.

Teorema. *Sea $P_{n+1}(x)$ el polinomio de grado $n+1$ con el coeficiente superior igual a 1. Entonces*

$$\max_{[-1,1]} |P_{n+1}(x)| \geqslant \max_{[-1,1]} |\overline{T}_{n+1}(x)| = 2^{-n}. \tag{8}$$

□ Supongamos que la desigualdad (8) no se cumple. Entonces el polinomio $Q_n(x) = \overline{T}_{n+1} - P_{n+1}$, cuyo grado no supera n, habría coincidido en todos $n+2$ puntos extremos x_m del polinomio T_{n+1} con este último en signo y por lo tanto, habría tomado alternadamente en estos puntos una vez el valor positivo, otra vez el valor negativo. Por eso $Q_n(x)$ debería tener $n+1$ raíces diferentes lo que es imposible para un polinomio cuyo grado no supera n. La contradicción obtenida demuestra el teorema. ■

Todo segmento $[a, b]$ puede ser obtenido del segmento $[-1, 1]$ por el reemplazo lineal de las variables

$$x' = \frac{b+a}{2} + \frac{b-a}{2} x. \tag{9}$$

En este caso el polinomio $\overline{T}_{n+1}(x)$ se transforma en polinomio $\overline{T}_{n+1}\left(\frac{2x-(b+a)}{b-a}\right)$ con el coeficiente superior $\left(\frac{2}{b-a}\right)^{n+1}$.

Por consiguiente,

$$\overline{T}^{[a,b]}_{n+1}(x) = (b-a)^{n+1} 2^{-2n-1} T_{n+1}\left(\frac{2x-(b+a)}{b-a}\right) \tag{10}$$

es un polinomio, con el coeficiente superior igual a 1, el cual se desvía mínimamente del cero sobre $[a, b]$ y para todo polinomio $P_{n+1}(x)$ de grado $n+1$, con el coeficiente superior igual a 1, se cumple la desigualdad

$$\max_{[a,b]} |P_{n+1}(x)| \geqslant \max_{[a,b]} (x) |\overline{T}^{[a,b]}_{n+1}(x)| = 2^{-n}\left(\frac{b-a}{2}\right)^{n+1}. \tag{11}$$

En virtud del reemplazo lineal (9) las raíces del polinomio $\overline{T}^{[a,b]}_{n+1}(x)$ tienen la forma

$$x_k = \frac{b+a}{2} + \frac{b-a}{2} \cos \frac{2k+1}{2(n+1)} \pi \quad (k = 0, \dots, n) \tag{12}$$

y los puntos extremos tienen la forma

$$x_m = \frac{b+a}{2} + \frac{b-a}{2} \cos \frac{m}{n+1} \pi \quad (m = 0, 1, \ldots, n+1). \qquad (13)$$

Volvamos ahora a considerar el problema de minimización del error de interpolación Δ_1 sobre el segmento $[a, b]$ para una red arbitraria en la clase de funciones continuamente derivables $n+1$ veces que satisfacen la condición (7) del § 7.5. Vamos a designar esta clase de funciones $C^{n+1}(M_{n+1}, [a, b])$. Para resolver el problema planteado en virtud de la fórmula (8) del § 7.5 es necesario minimizar la magnitud $\max\limits_{[a,b]} |\omega_n(x)|$. Puesto que $\omega_n(x)$ es el polinomio de grado $n+1$ con el coeficiente superior igual a 1, es evidente que la magnitud $\max\limits_{[a,b]} |\omega_n(x)|$ alcanza su valor mínimo para los polinomios de Chébyshev $\overline{T}_{n+1}^{[a,b]}(x)$. Por lo tanto, en calidad de nodos de interpolación conviene elegir los puntos x_k definidos por la expresión (12). En este caso

$$\max_{[a,b]} |\omega_n(x)| = \max_{[a,b]} |\overline{T}_{n+1}^{[a,b]}(x)| = 2^{-n} \left(\frac{b-a}{2}\right)^{n+1} \qquad (14)$$

y la estimación (8) del § 7.5 toma la forma

$$\Delta_1 \leqslant \frac{M_{n+1}}{(n+1)!} 2^{-n} \left(\frac{b-a}{2}\right)^{n+1}. \qquad (15)$$

Esta estimación es no mejorable, ya que en ella tiene lugar el signo de igualdad si en calidad de función $f(x)$ se elige el siguiente polinomio de grado $n+1$:

$$f(x) = \frac{M_{n+1}}{(n+1)!} x^{n+1} + a_n x^n + \ldots$$

y en calidad de nodos de interpolación, los puntos x_k definidos por la expresión (12).

Ejemplo 1. Sobre el segmento $[-1, 1]$ obtener la estimación uniforme de desviación de la función $f(x) = 1 - \cos(\pi x/2)$ respecto a su polinomio interpolador construido sobre los nodos de Chébyshev (3) para $n = 2, 3, 4$.

$\triangle$ Nótese, ante todo, que para la función en cuestión sobre el intervalo dado $M_{n+1} = (\pi/2)^{n+1}$, $b - a = 2$. Por eso en virtud de la estimación (15) tenemos:

$$n=2,\ \Delta_1 \leqslant \left(\frac{\pi}{2}\right)^3 \frac{1}{3!} \left(\frac{1}{2}\right)^2 \approx 0{,}17;$$

$$n=3,\ \Delta_1 \leqslant \left(\frac{\pi}{2}\right)^4 \frac{1}{4!} \left(\frac{1}{2}\right)^3 \approx 0{,}032;$$

$$n=4,\ \Delta_1 \leqslant \left(\frac{\pi}{2}\right)^5 \frac{1}{5!} \left(\frac{1}{2}\right)^4 \approx 0{,}005.$$

Recomendamos comparar la solución obtenida con la del ejemplo 1 del § 7.5. $\blacktriangle$

§ 7.15. Interpolación con nodos múltiplos

Hasta ahora hemos considerado el problema en el cual los parámetros de interpolación—los coeficientes del polinomio interpolador—se determinaban sólo por los valores de la función a interpolar. Tal problema suele llamarse *problema de interpolación según Lagrange* y el mismo proceso de construcción del polinomio interpolador, *proceso de Lagrange.*

Examinemos ahora un problema algo más amplio: el problema de interpolación respecto a los valores de la función y de sus derivadas o, como se dice de otro modo, el *problema de interpolación múltipla.*

Supongamos que sobre la red Λ_m: $a \leqslant x_0 < x_1 < \ldots < x_m \leqslant b$ en los nodos x_i se dan los valores f_i de cierta función f y las derivadas de esta última $f_i^{(k)}$ $(i = 0, 1, \ldots, m;\ k = 0, 1, \ldots, \alpha_i - 1)$, con la particularidad de que $\sum\limits_{i=0}^{m} \alpha_i = n + 1$. Se necesita construir el polinomio H_n cuyos valores y derivadas hasta el orden $\alpha_i - 1$ en los nodos x_i $(i = 0, 1, \ldots, m)$ coinciden con los valores de f y con las derivadas correspondientes de esta última, así como estimar el error.

Tal tipo de interpolación se llama *interpolación según Hermite* y el polinomio respectivo H_n, *polinomio de Hermite*. Los números α_i se denominan *multiplicidades de los nodos* x_i. En este caso se puede demostrar que el polinomio de Hermite existe y es único.

El término residual de la fórmula de interpolación $f(x) \approx H_n(x)$ se puede representar en la forma siguiente:

$$R_n(x) = f(x) - H_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{i=0}^{m} (x - x_i)^{\alpha_i};\ \xi \in (a, b). \qquad (1)$$

Sea ahora, para precisar,

$$|f^{(n+1)}(x)| \leqslant M_{n+1},\ x \in [a, b]. \qquad (2)$$

Utilizando esta restricción y la fórmula (1), obtenemos la estimación del error para el punto fijo x:

$$\Delta_1 = |f(x) - H_n(x)| \leqslant \frac{M_{n+1}}{(n+1)!} \prod_{i=0}^{m} |x - x_i|^{\alpha_i}. \qquad (3)$$

Ahora no es difícil construir una estimación uniforme sobre todo el segmento $[a, b]$ para la red fija Λ_m. En efecto,

$$\Delta_1 = \max_{[a,\, b]} |R_n(x)| \leqslant \frac{M_{n+1}}{(n+1)!} \max_{[a,\, b]} |\Omega_n(x)|, \qquad (4)$$

donde

$$\Omega_n(x) = \prod_{i=0}^{m} (x - x_i)^{\alpha_i}. \qquad (5)$$

Ejemplo. Construir el polinomio interpolador de Hermite para la función $f = 1 - \cos(\pi x/2)$ sobre los nodos $x_0 = -1$, $x_1 = 0$, $x_2 = 1$ con las multiplicidades $\alpha_0 = 1$, $\alpha_1 = 2$, $\alpha_2 = 1$, respectivamente. Obtener una estimación uniforme del error sobre el segmento $[-1, 1]$.

$\triangle$ Calculemos en los nodos dados los valores de la función y de su derivada: $f(x_0) = f(x_2) = 1$; $f(x_1) = f'(x_1) = 0$. Ahora vamos a construir el polinomio de Hermite teniendo en cuenta las multiplicidades de los nodos:

$$H_3(x) = 1 \cdot \frac{x^2(x-1)}{-2} + 0 \cdot \frac{(x-1)(x+1)}{-1} + 1 \cdot \frac{(x+1)x^2}{2} = x^2.$$

Nótese que en vez del polinomio de tercer grado hemos obtenido el de segundo grado lo que es consecuencia de la simetría de la información inicial (pero no de la función f).

Hallamos ahora la estimación del error. Utilizando la fórmula (4) y teniendo en cuenta que para la función dada $M_4 = (\pi/2)^4$, obtenemos

$$\Delta_1 \leqslant \left(\frac{\pi}{2}\right)^4 \cdot \frac{1}{4!} \cdot \max_{[-1,1]} |(x+1)x^2(x-1)|.$$

No es difícil mostrar que $\max\limits_{[-1,1]} |(x+1)x^2(x-1)| = 0{,}25$. Por eso finalmente tenemos $\Delta_1 \leqslant 0{,}065$. ▲

§ 7.16. Aparato matemático de la interpolación trigonométrica

En los párrafos siguientes consideraremos el problema de aproximación de las funciones con ayuda del polinomio trigonométrico. Esto quiere decir que en calidad de función aproximante se toma la combinación lineal de las funciones trigonométricas sen nx y cos nx.

Para fundamentar formalmente la elección de las funciones trigonométricas en calidad de aproximantes necesitaremos algunas nociones del curso de análisis matemático que se refieren a las series de Fourier. A continuación citamos estas nociones omitiendo las demostraciones de las afirmaciones respectivas.

Sucesiones. Series. Consideremos cierta función $f(x)$. En calidad de dominio de definición de esta función tomemos el conjunto de los números naturales, o sea, el argumento x toma los valores $1, 2, \ldots, n$. Tal función se llama *sucesión*.

La sucesión se escribe en la forma

$$a_1, a_2, \ldots, a_{n-1}, a_n, a_{n+1}, \ldots, \text{ o bien } \{a_n\},$$

con la particularidad de que a_n se llama *término general*; a_{n-1}, término precedente a a_n; a_{n+1}, término siguiente a a_n.

Citemos algunos ejemplos de las sucesiones.

1) La sucesión 1, 2, 3, ..., n, ..., cuyo término general es $a_n = n$ se llama *serie de los números naturales.*

2) La sucesión $a_1, a_2, \ldots, a_{n-1}, a_n, \ldots$ para la cual $a_n - a_{n-1} = d$, donde d es una magnitud constante, se denomina *progresión aritmética*. Para definir la progresión aritmética basta conocer su primer término a_1 y la diferencia de la progresión d. En efecto, el término general se expresa por la fórmula

$$a_n = a_1 + d\,(n-1).$$

Puesto que con ayuda de la fórmula citada se puede hallar todo término de la sucesión, sustituyendo los valores $n = 1, 2, 3, \ldots$ la sucesión se considera definida.

3) La sucesión $b_1, b_2, \ldots, b_{n-1}, b_n, \ldots$ para la cual $b_n = b_{n-1}q$, donde q es una magnitud constante, se llama *progresión geométrica*. Para definir una progresión geométrica basta conocer su primer término b_1 y la razón q. El término general se expresa por la fórmula

$$b_n = b_1 q^{n-1}.$$

4) La sucesión $c_1, c_2, \ldots, c_n, \ldots$ para la cual $c_n = c$, donde c una magnitud constante, se denomina *sucesión constante.*

5) Consideremos un ejemplo más de la sucesión. Vamos a calcular el número e sucesivamente con una cifra, con dos cifras, con tres cifras, etc. Los resultados de cálculos pueden ser representados en la forma siguiente:

$$\begin{array}{c} 2;\ 2{,}7;\ 2{,}71;\ 2{,}718;\ \ldots; \\ (1);\ (2);\ (3);\ (4);\ \ldots \end{array}$$

Al numerar los valores obtenidos por los números naturales de la serie, según se muestra entre paréntesis, obtenemos una sucesión.

Además de las sucesiones numéricas, consideraremos también las *sucesiones funcionales.*

Vamos a citar algunos ejemplos de sucesiones funcionales:

$$1)\ a_0,\ a_1x,\ a_2x^2,\ \ldots,\ a_{n-1}x^{n-1},\ \ldots;$$
$$2)\ \operatorname{sen} x,\ \operatorname{sen} 2x,\ \operatorname{sen} 3x,\ \ldots,\ \operatorname{sen} nx,\ \ldots$$

Recuérdese la definición del límite de una sucesión numérica. El número A se llama *límite de la sucesión* $\{a_n\}$ si para todo $\varepsilon > 0$ existe un tal número N que para todos los $n > N$ se cumpla la desigualdad $|a_n - A| < \varepsilon$. En este caso se escribe $\lim\limits_{n\to\infty} a_n = A$.

La sucesión que tiene el límite se llama *convergente* y en el caso contrario, *divergente.*

Ejemplo 1. Mostrar que la progresión geométrica $b, bq, bq^2, \ldots, bq^{n-1}, \ldots$ para $|q| < 1$ es una sucesión convergente y para $|q| \geqslant 1$, una sucesión divergente.

△ 1) Primero consideremos el caso cuando $|q| < 1$. Mostremos que

$$\lim_{n\to\infty} bq^{n-1} = 0,$$

o sea, según el dado $\varepsilon > 0$ hallemos un N tal que para $n > N$ se cumpla la desigualdad $|bq^{n-1} - 0| < \varepsilon$. Para esto resolvamos la desigualdad respecto a n. Escribámosla en la forma

$$|b|\,|q|^{n-1} < \varepsilon \text{ o bien } |q|^{n-1} \leqslant \varepsilon/|b|.$$

Logaritmemos la última desigualdad

$$(n-1)\ln|q| < \ln(\varepsilon/|b|).$$

Dividiendo ambos sus miembros por el número negativo $\ln|q|$, obtenemos

$$n-1 > \frac{\ln(\varepsilon/|b|)}{\ln|q|}, \text{ o sea, } n > 1 + \frac{\ln(\varepsilon/|b|)}{\ln|q|}.$$

Es evidente que en calidad de N es suficiente tomar

$$N = E\left(1 + \frac{\ln(\varepsilon/|b|)}{\ln|q|}\right),$$

donde $E(x)$ designa el número entero máximo que no supere x.

2) Examinemos ahora el caso cuando $q > 1$, $b > 0$. Mostremos que la sucesión es divergente. Para esto basta mostrar que para todo M grande como se quiera existe un N tal que para todos los $n > N$ se cumpla la desigualdad $bq^{n-1} > M$. Despejando en la última desigualdad n, obtenemos

$$n > 1 + \frac{\ln(M/b)}{\ln q}.$$

En calidad de N tomamos

$$N = E\left(1 + \frac{\ln(M/b)}{\ln q}\right).$$

Nótese que para $q = 1$ la sucesión es constante y $\lim\limits_{n\to\infty} bq^{n-1} = b$.

Se puede mostrar que en todos los demás casos la sucesión se diverge. ▲

Supongamos que hay una sucesión $a_1, a_2, \ldots, a_n, \ldots$. La expresión que tiene la forma

$$a_1 + a_2 + \ldots + a_n + \ldots = \sum_{n=1}^{\infty} a_n$$

se llama *serie*; aquí a_n es el n-ésimo término de la serie.

La suma de los primeros n términos de la serie se denomina *n-ésima suma parcial* de esta última:

$$S_n = \sum_{i=1}^{n} a_i.$$

Se llama *suma de la serie* el límite de la sucesión de sus sumas parciales

$$\lim_{n\to\infty} S_n = S.$$

Si la serie tiene la suma, se llama *convergente*, en el caso contrario se dice que la serie *diverge.*

Citemos algunos ejemplos de las series convergentes y divergentes.

Ejemplo 2. Consideremos una serie que se obtiene de la progresión aritmética y escribamos su n-ésima suma parcial

$$S_n = \frac{a_1+a_n}{2}\cdot n = \frac{2a_1+d(n-1)}{2}\cdot n = \frac{d}{2}n^2 + \frac{2a_1-d}{2}\, n.$$

Es evidente que cuando n tiende hacia el infinito esta suma parcial crece indefinidamente en valor absoluto y, por consiguiente, la serie dada diverge.

Ejemplo 3. Consideremos una serie formada por la progresión geométrica para $|q| < 1$ y determinemos la suma de aquélla. Hagamos uso de la fórmula de la suma de n términos de la progresión geométrica:

$$S_n = \frac{b_1(1-q^n)}{1-q} = \frac{b_1}{1-q} - \frac{b_1}{1-q}\, q^n.$$

Antes hemos demostrado (véase el ejemplo 1) que $\lim\limits_{n\to\infty} q^n = 0$ ($|q| < 1$); por lo tanto,

$$S = \lim_{n\to\infty} S_n = \frac{b_1}{1-q}.$$

Las series cuyos términos son funciones se llaman *series funcionales.* Tales son, por ejemplo, las series

$$a_0 + a_1x + a_2x^2 + \ldots + a_{n-1}x^{n-1} + \ldots,$$
$$b_0 + b_1\cos x + b_2\cos 2x + \ldots + b_{n-1}\cos(n-1)x + \ldots.$$

La primera de ellas se denomina serie *de potencias* y la segunda, serie *trigonométrica.*

Consideremos la serie funcional

$$u_1(x) + u_2(x) + \ldots + u_n(x) + \ldots,$$

donde $u_n(x)$ son las funciones definides sobre el segmento $[a, b]$. Sea $x_0 \in [a, b]$; entonces la serie

$$u_1(x_0) + u_2(x_0) + \ldots + u_n(x_0) + \ldots$$

es una serie numérica y puede resultar convergente o divergente.

El conjunto de todos los valores de $x \in [a, b]$ para los cuales converge la serie numérica correspondiente se denomina *dominio de convergencia* de la serie funcional.

Es evidente que

$$S(x) = \lim_{n\to\infty} S_n(x),$$

donde $S_n(x) = \sum_{i=1}^{n} u_i(x)$ depende de la elección de la variable x, o sea, la suma $S(x)$ de la serie funcional es función del punto x.

Sea $\{S_n(x)\}$ la sucesión de las sumas parciales de la serie funcional definidas sobre el mismo intervalo cerrado $[a, b]$. La serie funcional se denomina *uniformemente convergente* hacia la función $S(x)$ definida sobre $[a, b]$ si a todo $\varepsilon > 0$ se puede hacer corresponder un tal número N, no dependiente de $x \in [a, b]$, que para todo $n > N$ se cumpla la desigualdad $|S_n(x) - S(x)| < \varepsilon$.

Examinemos un ejemplo que ilustra la diferencia existente entre los conceptos de convergencia de la serie y de convergencia uniforme de la misma.

Ejemplo 4. Para la serie $\sum_{n=1}^{\infty} \frac{x}{(1+x)^n}$, donde $0 \leqslant x \leqslant 1$, la suma parcial es

$$S_n(x) = \sum_{i=1}^{n} \frac{x}{(1+x)^i} = 1 - \frac{1}{(1+x)^n}.$$

Mostremos que en el segmento dado la serie converge.

En efecto, si $x = 0$, entonces $S_n(0) = S(0) = 0$. Sea ahora $x > 0$. Demostremos que en este caso $S(x) = 1$, o sea, según el dado $\varepsilon > 0$ determinemos un N tal que para $n > N$ se cumpla la desigualdad

$$\left| 1 - \frac{1}{(1+x)^n} - 1 \right| < \varepsilon.$$

Resolviendo esta desigualdad, obtenemos

$$N = E\left(-\frac{\ln \varepsilon}{\ln(1+x)} \right).$$

De la última igualdad se ve que, hablando en general, N depende no sólo de ε sino también de x, con la particularidad de que para un mismo ε, si x se aproxima a cero, N crece indefinidamente. Esto quiere decir que según el ε dado no podemos elegir el N *único para* todos los $x \in [0, 1]$, con otras palabras, la serie *no es uniformemente convergente sobre el segmento señalado.*

El ejemplo citado muestra que la sucesión de las sumas parciales, continuas sobre cierto segmento, puede converger hacia una función discontinua sobre este segmento. Una de las causas que originan la introducción del concepto de convergencia uniforme de una serie consiste en que una serie uniformemente convergente de las funciones continuas tiene como su suma también una función continua.

Desarrollo de las funciones en serie de Fourier. Muchos problemas de la ciencia y la técnica están vinculados con las funciones periódicas que reflejan procesos cíclicos.

La función $f(x)$ se llama *periódica* con el período $T > 0$ si satisface la igualdad

$$f(x) = f(x + T). \tag{1}$$

Partiendo de las consideraciones prácticas es cómodo, con un grado suficiente de precisión, representar tales funciones en forma de una serie trigonométrica o de su suma parcial.

La serie funcional que tiene la forma

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \operatorname{sen} nx) \tag{2}$$

se llama *trigonométrica*, con la particularidad de que a_n y b_n son los números reales que no dependen de x.

Supongamos que esta serie converge para todo x que está en el intervalo $[-\pi, \pi]$; entonces ella define la función periódica $f(x)$ de período $T = 2\pi$.

La serie que tiene la forma (2) se denomina *serie de Fourier* para la función $f(x)$, integrable sobre el segmento $[-\pi, \pi]$, si los coeficientes de esta serie se calculan con ayuda de las fórmulas siguientes:

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)\, dx, \tag{3}$$

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx\, dx \; (n = 1, 2, \ldots). \tag{4}$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \operatorname{sen} nx\, dx \; (n = 1, 2, \ldots). \tag{5}$$

Ahora bien, podemos formalmente considerar la serie de Fourier para la función dada $f(x)$. No obstante, en este caso surgen las preguntas siguientes: 1) ¿converge o no la serie de Fourier de la función $f(x)$? 2) si la serie converge ¿tendrá ésta como su suma $f(x)$? Es el teorema de Dirichlet que da las respuestas a estas preguntas. Antes de pasar al enunciado del mismo teorema, recordemos algunos conceptos.

La función $f(x)$ se considera *monótona* sobre un intervalo si para todos x_1 y x_2, pertenecientes a este intervalo y tales que $x_1 < x_2$, se cumple sólo una de las desigualdades $f(x_1) \leqslant f(x_2)$ o bien $f(x_1) \geqslant f(x_2)$.

La función $f(x)$ se dice *monótona a trozos* sobre un intervalo si éste puede ser dividido en un número finito de intervalos abiertos en cada uno de los cuales la función es monótona.

La función $f(x)$ se considera *continua a trozos* sobre un intervalo si tiene sobre éste un número finito de puntos de discontinuidad.

Designemos con $f(a+0)$ el límite de la función $f(x)$ si x tiende a a del lado derecho (límite por la derecha) y, respectivamente, designemos con $f(a-0)$ el límite por la izquierda.

Teorema de Dirichlet. *Si una función $f(x)$, dada sobre el segmento $[-\pi, \pi]$, es monótona a trozos y continua a trozos, la serie de Fourier de esta función converge sobre todo el segmento $[-\pi, \pi]$ y la suma de la serie es igual:*

1) a $f(x)$ *en todos los puntos de continuidad pertenecientes a* $(-\pi, \pi)$;

2) a $\frac{1}{2}[f(x-0)+f(x+0)]$ *en todos los puntos de discontinuidad pertenecientes a* $(-\pi; \pi)$;

3) a $\frac{1}{2}[f(-\pi+0)+f(\pi-0)]$ *en los extremos del segmento, es decir, en los puntos* $x=-\pi$ *y* $x=\pi$.

A continuación escribiremos que

$$f(x)=\frac{a_0}{2}+\sum_{n=1}^{\infty}(a_n \cos nx+b_n \operatorname{sen} nx) \tag{6}$$

según el sentido del teorema de Dirichlet recién enunciado.

El teorema de Dirichlet no afirma que la serie de Fourier converge uniformemente hacia la función $f(x)$. No obstante, si se refuerzan las propiedades a las cuales debe satisfacer la función, es decir, exigir de ella que sea continua sobre todo el segmento $[-\pi, \pi]$ y sea monótona a trozos sobre éste, así como que se cumpla la igualdad $f(-\pi)=f(\pi)$, entonces para tal función la serie de Fourier convergirá uniformemente hacia la función $f(x)$ sobre todo el segmento $[-\pi, \pi]$.

Se puede mostrar que *para la función par todos los coeficientes b_n son iguales a cero y la serie de Fourier correspondiente no contiene senos*:

$$f(x)=\frac{a_0}{2}+\sum_{n=1}^{\infty} a_n \cos nx, \tag{7}$$

donde

$$a_n=\frac{2}{\pi}\int_0^{\pi} f(x)\cos nx\,dx \;(n=0,\ 1,\ 2,\ \ldots). \tag{8}$$

Análogamente *para la función impar todos los coeficientes a_n son iguales a cero y la serie de Fourier correspondiente no contiene cosenos*:

$$f(x)=\sum_{n=1}^{\infty} b_n \operatorname{sen} nx, \tag{9}$$

donde

$$b_n=\frac{2}{\pi}\int_0^{\pi} f(x)\ \text{sen}\ nx\, dx\ (n=1,\ 2,\ \ldots). \qquad (10)$$

Ejemplo 1. Desarrollar en serie de Fourier la función

$$f(x)=\begin{cases} -x & \text{para} \quad -\pi\leqslant x\leqslant 0, \\ 2x & \text{para} \quad 0<x\leqslant \pi. \end{cases}$$

△ Determinemos los coeficientes de Fourier de la función $f(x)$. Con ayuda de las fórmulas (3) y (4) hallamos los coeficientes a_0 y a_n:

$$a_0=\frac{1}{\pi}\int_{-\pi}^{0}(-x)\, dx+\frac{1}{\pi}\int_0^{\pi} 2x\, dx=\frac{1}{\pi}\cdot\frac{\pi^2}{2}+\frac{1}{\pi}\cdot\pi^2=\frac{3}{2}\pi;$$

$$a_n=\frac{1}{\pi}\int_{-\pi}^{0}(-x)\cos nx\, dx+\frac{1}{\pi}\int_0^{\pi} 2x\cos nx\, dx.$$

Integrando por partes obtenemos

$$a_n=-\frac{1}{\pi}\left(\frac{1}{n}x\ \text{sen}\ nx\Big|_{-\pi}^{0}-\frac{1}{n}\int_{-\pi}^{0}\text{sen}\ nx\, dx-\right.$$

$$\left.-\frac{2}{n}x\ \text{sen}\ nx\Big|_0^{\pi}+\frac{2}{n}\int_0^{\pi}\text{sen}\ nx\, dx\right)=-$$

$$-\frac{1}{\pi n}\left(\frac{1}{n}\cos nx\Big|_{-\pi}^{0}-\frac{2}{n}\cos nx\Big|_0^{\pi}\right)=-\frac{3}{\pi n^2}[1-(-1)^n],$$

o sea,

$$a_{2k}=0,\ a_{2k-1}=-\frac{6}{\pi(2k-1)^2}\ (k=1,\ 2,\ 3,\ \ldots).$$

Obtenemos los coeficientes b_n con ayuda de la fórmula (5):

$$b_n=\frac{1}{\pi}\int_{-\pi}^{0}(-x)\ \text{sen}\ nx\, dx+\frac{1}{\pi}\int_0^{\pi} 2x\ \text{sen}\ nx\, dx=$$

$$=-\frac{1}{\pi}\left(-\frac{x}{n}\cos nx\Big|_{-\pi}^{0}+\frac{1}{n}\int_0^{\pi}\cos nx\, dx+\frac{2x}{n}\cos nx\Big|_0^{\pi}-\right.$$

$$\left.-\frac{2}{n}\int_0^{\pi}\cos nx\, dx\right)=-\frac{1}{\pi}\left[-\frac{\pi}{n}(-1)^n+2\frac{\pi}{n}(-1)^n\right]=\frac{(-1)^{n+1}}{n},$$

o sea,

$$b_{2k}=-\frac{1}{2k},\ b_{2k-1}=\frac{1}{2k-1}\ (k=1,\ 2,\ 3,\ \ldots).$$

Ahora bien, la serie de Fourier de la función dada tiene la forma

$$f(x)=\frac{3}{4}\pi-\frac{6}{\pi}\sum_{k=1}^{\infty}\frac{1}{(2k-1)^2}\cos(2k-1)x+$$
$$+\sum_{k=1}^{\infty}\frac{1}{2k-1}\operatorname{sen}(2k-1)x-\frac{1}{2}\sum_{k=1}^{\infty}\frac{1}{k}\operatorname{sen}2kx.$$

En el intervalo $(-\pi, \pi)$ la serie converge hacia la función $f(x)$ y en los puntos $x=\pm\pi$, hacia el número $\frac{1}{2}[f(-\pi+0)+$ $+f(\pi-0)]=\frac{3}{2}\pi$. ▲

Ejemplo 2. Desarrollar en serie de Fourier la función

$$f(x)=\begin{cases}-\operatorname{sen}x & \text{para } -\pi\leqslant x\leqslant 0,\\ \operatorname{sen}x & \text{para } \quad 0<x\leqslant\pi.\end{cases}$$

△ La función dada es par, por consiguiente, todos los coeficientes $b_n=0$ y a_n se determinan con ayuda de la fórmula (8):

$$a_n=\frac{2}{\pi}\int_0^{\pi}\operatorname{sen}x\cos nx\,dx\quad(n=0,\ 1,\ 2,\ \ldots).$$

De aquí

$$a_0=\frac{2}{\pi}\int_0^{\pi}\operatorname{sen}x\,dx=\frac{4}{\pi}\ .$$

Luego,

$$a_n=\frac{2}{\pi}\int_0^{\pi}\operatorname{sen}x\cos nx dx=\frac{1}{\pi}\int_0^{\pi}[\operatorname{sen}(n+1)x-\operatorname{sen}(n-1)x]\,dx=$$
$$=\begin{cases}0 & \text{para } n=2k-1,\\ -\dfrac{4}{\pi(n^2-1)} & \text{para } n=2k\ (k=1,\ 2,\ 3,\ \ldots).\end{cases}$$

Por lo tanto,

$$f(x)=|\operatorname{sen}x|=\frac{2}{\pi}-\frac{4}{\pi}\sum_{k=1}^{\infty}\frac{1}{4k^2-1}\cos 2kx.$$

Nótese que la serie obtenida converge hacia la función $|\operatorname{sen}x|$ sobre todo el segmento $[-\pi, \pi]$. ▲

Ejemplo 3. Desarrollar en serie de Fourier la función

$$f(x)=\begin{cases}-1 & \text{para } -\pi<x<0,\\ 0 & \text{para } \quad x=0,\\ 1 & \text{para } \quad 0<x<\pi.\end{cases}$$

△ La función dada es impar, por lo tanto, todos los coeficientes $a_n = 0$ y b_n se hallan con ayuda de la fórmula (10):

$$b_n = \frac{2}{\pi} \int_0^{\pi} 1 \cdot \operatorname{sen} nx \, dx = -\frac{2}{\pi n} \cos nx \Big|_0^{\pi} = \frac{2}{\pi n} [1 - (-1)^n].$$

Ahora bien, todos los coeficientes pares b_n son iguales a cero y los impares tienen la forma $b_{2k-1} = \frac{4}{\pi (2k-1)}$. Por consiguiente,

$$f(x) = \frac{4}{\pi} \sum_{k=1}^{\infty} \frac{1}{2k-1} \operatorname{sen} (2k-1) x.$$

Es evidente que la suma de la serie en los puntos $x = 0$ y $x = \pm\pi$ es igual a cero. ▲

Ejemplo 4. Desarrollar en serie de Fourier la función

$$f(x) = \begin{cases} x & \text{para } 0 < x \leqslant 1, \\ 1 & \text{para } 1 < x < 2. \end{cases}$$

△ Puesto que la función se da en un intervalo distinto de $(-\pi, \pi)$, sustituyamos la variable independiente con ayuda de la fórmula $x = (x' + \pi)/\pi$ o bien $x' = \pi (x - 1)$. Ahora bien, obtenemos la función siguiente:

$$f(x') = \begin{cases} \frac{1}{\pi} (x' + \pi) & \text{para } -\pi < x' \leqslant 0, \\ 1 & \text{para } \quad 0 < x' < \pi. \end{cases}$$

Puesto que esta serie está definida sobre el intervalo $(-\pi, \pi)$, para ella se puede escribir la serie de Fourier. Calculemos los coeficientes de esta serie:

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{0} \frac{x' + \pi}{\pi} \, dx' + \frac{1}{\pi} \int_0^{\pi} 1 \cdot dx' = \frac{3}{2},$$

$$a_n = \frac{1}{\pi} \int_{-\pi}^{0} \frac{x' + \pi}{\pi} \cos nx' \, dx' + \frac{1}{\pi} \int_0^{\pi} 1 \cdot \cos nx' \, dx' =$$

$$= \frac{1}{n\pi^2} \int_{-\pi}^{0} \operatorname{sen} nx' \, dx' = \frac{1}{\pi^2 n^2} [1 - (-1)^n];$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{0} \frac{x' + \pi}{\pi} \operatorname{sen} nx' \, dx' + \frac{1}{\pi} \int_0^{\pi} 1 \cdot \operatorname{sen} nx' \, dx' =$$

$$= \frac{1}{\pi^2} \int_{-\pi}^{0} x' \operatorname{sen} nx' \, dx' = -\frac{(-1)^n}{\pi n}.$$

Por lo tanto,

$$f(x)=\frac{3}{4}+\frac{2}{\pi^2}\sum_{k=1}^{\infty}\frac{1}{(2k-1)^2}\cos[(2k-1)\pi(x-1)]-$$
$$-\frac{1}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^n}{n}\operatorname{sen}[n\pi(x-1)].$$

Calculemos los valores de la suma de la serie en los extremos del intervalo:

$$\frac{1}{2}[f(0+0)+f(2-0)]=\frac{0+1}{2}=\frac{1}{2}.$$

El resultado obtenido ofrece la posibilidad de hallar la suma de la serie numérica

$$S=\sum_{k=1}^{\infty}\frac{1}{(2k-1)^2}=1+\frac{1}{3^2}+\frac{1}{5^2}+\ldots+\frac{1}{(2k-1)^2}+\ldots$$

En efecto, en virtud del teorema de Dirichlet para $x=0$ o para $x=2$ tiene lugar la igualdad

$$\frac{1}{2}=\frac{3}{4}-\frac{2}{\pi^2}\sum_{k=1}^{\infty}\frac{1}{(2k-1)^2},$$

de donde

$$\sum_{k=1}^{\infty}\frac{1}{(2k-1)^2}=\frac{\pi^2}{8}. \quad \blacktriangle$$

En conclusión nótese que la integral de la función $f(x)$ se obtiene integrando por términos la serie de Fourier que le corresponde y la derivada $f'(x)$, derivándola por términos. En caso de la derivación la condición $f(-\pi)=f(\pi)$ es necesaria.

§ 7.17. Interpolación trigonométrica

La operación de representar la función $f(x)$ por la serie de Fourier se llama *análisis armónico*. En los cálculos prácticos nos vemos obligados a considerar sólo algunos primeros términos de la serie de Fourier. Como resultado se obtiene únicamente una expresión analítica aproximada para la función $f(x)$ en la forma del polinomio trigonométrico de N-ésimo orden:

$$Q_N(x)=\frac{a_0}{2}+\sum_{n=1}^{N}(a_n\cos nx+b_n\operatorname{sen}nx)\ (-\pi\leqslant x\leqslant\pi). \qquad (1)$$

Además, las fórmulas (3) ... (5) del § 7.16 para calcular los coeficientes de Fourier son útiles exclusivamente en caso de la representación analítica de la función. En la práctica la función $f(x)$ se representa, por regla general, en forma de una tabla o gráfico. Por eso surge el problema de determinar aproximadamente los coeficientes de Fourier valiéndose del número finito de los valores disponibles de la función.

Generalizando lo dicho, enunciemos el problema siguiente del análisis armónico numérico o bien, como con frecuencia se lo llama de otro modo, del análisis armónico práctico: aproximar sobre el intervalo $(0, T)$ por un polinomio trigonométrico de N-ésimo grado la función $y = f(x)$ para la cual se conocen m valores suyos $y_k = f(x_k)$ para $x_k = kT/m$ $(k = 0, 1, 2, \ldots, m-1)$.

El polinomio trigonométrico para la función definida sobre el intervalo $(0, T)$ tiene la forma

$$Q_N(x) = \frac{a_0}{2} + \sum_{n=1}^{N} \left(a_n \cos n \frac{2\pi}{T} x + b_n \operatorname{sen} n \frac{2\pi}{T} x\right). \tag{2}$$

Los coeficientes a_n y b_n se definen por las relaciones siguientes:

$$a_n = \frac{2}{T} \int_0^T f(x) \cos n \frac{2\pi}{T} x \, dx, \tag{3}$$

$$(n = 0, 1, 2, \ldots, N).$$

$$b_n = \frac{2}{T} \int_0^T f(x) \operatorname{sen} n \frac{2\pi}{T} x \, dx. \tag{4}$$

Aplicando en las relaciones (3) y (4) la fórmula de los rectángulos para calcular los integrales según los valores de las expresiones subintegrales en los puntos $x_k = kT/m$ $(k = 0, 1, 2, \ldots, m-1)$, tenemos

$$a_n = \frac{2}{m} \sum_{k=0}^{m-1} y_k \cos n \frac{2\pi k}{m}, \tag{5}$$

$$(n = 0, 1, 2, \ldots, \ldots, N).$$

$$b_n = \frac{2}{m} \sum_{k=0}^{m-1} y_k \operatorname{sen} n \frac{2\pi k}{m}. \tag{6}$$

Ahora bien, el polinomio trigonométrico (2), cuyos coeficientes a_n y b_n se determinan con ayuda de las fórmulas (5) y (6), sirve de solución del problema planteado.

Se puede mostrar que para $m > 2N$ el polinomio (2) da la mejor aproximación a la función $f(x)$, en lo que se refiere al método de los cuadrados mínimos, si los coeficientes de aquél se calculan haciendo uso de las fórmulas (5) y (6). Con otras palabras, los coeficientes (5)

Tabla 7.13

Meses	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	$\frac{1}{6}\Sigma$
x_k	0	1	2	3	4	5	6	7	8	9	10	11	
Volumen de producción en unidades conv.	95	71	55	43	36	31	28	26	25	45	91	102	108
$\cos\frac{\pi}{6}x_k$	1	0,866	0,5	0	−0,5	−0,866	−1	−0,866	−0,5	0	0,5	0,866	
$\operatorname{sen}\frac{\pi}{6}x_k$	0	0,5	0,866	1	0,866	0,5	0	−0,5	−0,866	−1	−0,866	−0,5	
$\cos\frac{\pi}{3}x_k$	1	0,5	−0,5	−1	−0,5	0,5	1	0,5	−0,5	−1	−0,5	0,5	
$\operatorname{sen}\frac{\pi}{3}x_k$	0	0,866	0,866	0	−0,866	−0,866	0	0,866	0,866	0	−0,866	−0,866	
$\cos\frac{\pi}{2}x_k$	1	0	−1	0	1	0	−1	0	1	0	−1	0	
$\operatorname{sen}\frac{\pi}{2}x_k$	0	1	0	−1	0	1	0	−1	0	1	0	−1	
$\cos\frac{2\pi}{3}x_k$	1	−0,5	−0,5	1	−0,5	−0,5	1	−0,5	−0,5	1	−0,5	−0,5	
$\operatorname{sen}\frac{2\pi}{3}x_k$	0	0,866	−0,866	0	0,866	−0,866	0	0,866	−0,866	0	0,866	−0,866	

Continuación de la tabla 7.13

Meses	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	$\frac{1}{6}\Sigma$
x_k	0	1	2	3	4	5	6	7	8	9	10	11	
Volumen de producción en unidades conv.	95	71	55	43	36	31	28	26	25	45	91	102	108
$y_k \cos \frac{\pi}{6} y_k$	95	61,49	27,5	0	−18	−26,85	−28	−25,52	−12,5	0	45,5	83,33	34,99
$y_k \operatorname{sen} \frac{\pi}{6} y_k$	0	35,5	47,63	43	31,18	15,5	0	−13	21,65	−45	−78,81	−51	−6,11
$y_k \cos \frac{\pi}{3} y_k$	95	35,5	−27,5	−43	−18	15,5	28	13	−12,5	−45	−45,5	51	7,75
$y_k \operatorname{sen} \frac{\pi}{3} y_k$	0	61,49	47,63	0	31,18	−28,85	0	22,52	21,65	0	−78,81	−88,33	−11,98
$y_k \cos \frac{\pi}{2} y_k$	95	0	−55	0	36	0	−28	0	25	0	−91	0	−3
$y_k \operatorname{sen} \frac{\pi}{2} y_k$	0	71	0	−43	0	31	0	−26	0	45	0	−102	−4
$y_k \cos \frac{2\pi}{3} y_k$	95	−35,5	−27,5	43	−18	−15,5	28	−13	−12,5	45	−45,5	−51	−1,25
$y_k \operatorname{sen} \frac{2\pi}{3} y_k$	0	61.49	−47,63	0	31,18	−26,85	0	22,52	−21,65	0	78,81	−88,33	1,59

y (6) minimizan la suma de los cuadrados de desviaciones

$$\delta_N^2 = \sum_{k=0}^{m-1} [Q_N(x_k) - y_k]^2. \qquad (7)$$

En el caso particular para $m = 2N$ los coeficientes a_n y b_n ($n = 0, 1, 2, \ldots, N-1$) se definen por las relaciones (5) y (6) y el coeficiente a_N es

$$a_N = \frac{1}{m} \sum_{k=0}^{m-1} (-1)^k y_k. \qquad (8)$$

El mismo polinomio $Q_N(x)$ llega a ser un polinomio interpolador, ya que en este caso para todo b_N se cumplen las relaciones $Q_N(x_k) = y_k$ para todos los $x_k = kT/m$ ($k = 0, 1, 2, \ldots, m-1$).

Ejemplo. Investiguemos la dinámica de producción del azúcar de remolacha. Esta producción lleva un carácter periódico condicionado por la periodicidad del cultivo y por las condiciones de almacenamiento de las materias primas. Por eso en calidad de función que aproxima la dinámica de fabricación del azucar se puede tomar el polinomio trigonométrico (2) para $m = 12$ (esto corresponde al número de meses en el ciclo anual y permite revelar una particularidad específica, o sea, la temporalidad de la fabricación). Por consiguiente,

$$Q_N(x) = \frac{a_0}{2} + \sum_{n=1}^{N} \left(a_n \cos n\frac{\pi}{6}x + b_n \operatorname{sen} n\frac{\pi}{6}x\right) \ (0 \leqslant x \leqslant 11).$$

En las investigaciones económicas para una buena aproximación de la serie dinámica periódica suelen elegirse cuatro armónicos, como máximo. Los coeficientes a_n y b_n se expresan en la forma siguiente:

$$a_n = \frac{1}{6}\sum_{k=0}^{11} y_k \cos n\frac{\pi}{6}x_k, \quad b_n = \frac{1}{6}\sum_{k=0}^{11} y_k \operatorname{sen} n\frac{\pi}{6}x_k.$$

Calculemos estos coeficientes para los primeros cuatro armónicos del polinomio $Q_N(x)$. Los cálculos necesarios se formalizan tabularmente (véase la tabla 7.13 en las págs 321 y 322).

De la tabla dada obtenemos que $a_0 = 108$; $a_1 = 34{,}99$; $a_2 = 7{,}75$; $a_3 = -3$; $a_4 = -1{,}25$; $b_1 = -6{,}11$; $b_2 = -11{,}98$; $b_3 = -4$; $b_4 = 1{,}59$. Ahora bien, tenemos los siguientes cuatro modelos matemáticos de temporalidad para la fabricación del azucar:

$$Q_1(x) = 54 + 34{,}99\cos\frac{\pi}{6}x - 6{,}11\operatorname{sen}\frac{\pi}{6}x;$$

$$Q_2(x) = 54 + 34{,}99\cos\frac{\pi}{6}x - 6{,}11\operatorname{sen}\frac{\pi}{6}x + 7{,}75\cos\frac{\pi}{3}x - {}$$
$$- 11{,}98\operatorname{sen}\frac{\pi}{3}x;$$

$$Q_3(x) = 54 + 34{,}99\cos\frac{\pi}{6}x - 6{,}11\operatorname{sen}\frac{\pi}{6}x + 7{,}75\cos\frac{\pi}{3}x -$$

$$-11,98 \operatorname{sen} \frac{\pi}{3} x - 3 \cos \frac{\pi}{2} x - 4 \operatorname{sen} \frac{\pi}{2} x;$$

$$Q_4 = 4 + 34,99 \cos \frac{\pi}{6} x - 6,11 \operatorname{sen} \frac{\pi}{6} x + 7,75 \cos \frac{\pi}{3} x -$$
$$- 11,98 \operatorname{sen} \frac{\pi}{3} x - 3 \cos \frac{\pi}{2} x - 4 \operatorname{sen} \frac{\pi}{2} x - 1,25 \cos \frac{2\pi}{3} x +$$
$$+ 1,59 \operatorname{sen} \frac{2\pi}{3} x.$$

La comparación de $Q_1(x_k)$ con los valores correspondientes de y_k muestra que ya el primer armónico da, en general, el modelo correcto de la dinámica de la fabricación del azúcar, reflejando su temporalidad.

Calculemos las desviaciones medias cuadráticas, o sea, las desviaciones estándar $\delta_i = \sqrt{\sum_{k=0}^{11} [Q_i(x_k) - y_k]^2}$ para todos los $Q_i(x)$; encontramos $\delta_1 = 37,80$; $\delta_2 = 14,40$; $\delta_3 = 7,59$; $\delta_4 = 5,75$. Como era de esperar, los valores δ_i decrecen monótonamente con el crecimiento de i, con la particularidad de que δ_4 poco se distingue de δ_3. Además, los mismos valores δ_3 y δ_4 son bastante pequeños, así que el polinomio $Q_3(x)$ es una buena aproximación de la serie que caracteriza la dinámica anual de la fabricación del azúcar.

§ 7.18. Métodos numéricos de determinación de los coeficientes de Fourier

Supongamos que se da una serie de Fourier que converge hacia la función periódica $f(x)$:

$$f(x) = \frac{a_0}{2} + \sum_{m=1}^{\infty} (a_m \cos mx + b_m \operatorname{sen} mx), \tag{1}$$

donde

$$a_m = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos mx \, dx \quad (m = 0, 1, 2, \ldots). \tag{2}$$

$$b_m = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \operatorname{sen} mx \, dx \quad (m = 1, 2, 3, \ldots). \tag{3}$$

En el párrafo precedente hemos enunciado el problema de aproximación de la función $f(x)$ por el polinomio trigonométrico $Q_N(x)$. Allí mismo al calcular los coeficientes a_m y b_m con ayuda de los integrales hemos utilizado la fórmula de los rectángulos.

En el caso general el cálculo aproximado de los coeficientes a_m y b_m se funda en la sustitución de los integrales de las fórmulas (3) y (4) del § 7.17 por sus valores, obtenidos con ayuda de una de las fórmulas de integración aproximada. En este párrafo hacemos uso de la fórmula de los trapecios.

Vamos a suponer que la función $f(x)$ es peri ódica con le período de 2π. Nótese que en vez de los límites us uales de integración entre $-\pi$ y π, al determinar los coefici entes a_m y b_m se puede considerar todo segmento de integración de 2π de largo. Para comodidad de los cálculos tomemos un segmento de 0 a 2π, así que

$$a_m = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos mx\,dx \quad (m = 0,\ 1,\ 2,\ \ldots), \tag{4}$$

$$b_m = \frac{1}{\pi}\int_0^{2\pi} f(x)\,\mathrm{sen}\,mx\,dx \quad (m = 1,\ 2,\ 3,\ \ldots). \tag{5}$$

Al dividir el segmento $[0, 2\pi]$ en N partes iguales, como resultado obtenemos los puntos de división $0,\ 1\cdot\frac{2\pi}{N},\ 2\cdot\frac{2\pi}{N},\ \ldots\ (N-1)\frac{2\pi}{N}$, 2π. En los p untos de división designemos los valores respectivos de N la N funci ón $f(x)$ con $y_0,\ y_1,\ y_2,\ \ldots,\ y_{N-1},\ y_N = y_0$. Aplicando la fórmula d e los trapecios, obtenemos las siguientes fórmulas aproximadas para calcular los coeficientes a_m y b_m:

$$\frac{N}{2}a_0 = \sum_{k=0}^{N-1} y_k = y_0 + y_1 + \ldots + y_{N-1};$$

$$\frac{N}{2}a_m = \sum_{k=0}^{N-1} y_k \cos k\frac{2m\pi}{N} = y_0 + y_1\cos\frac{2m\pi}{N} + \ldots$$

$$\ldots\ y_{N-1}\cos(N-1)\frac{2m\pi}{N};$$

$$\frac{N}{2}b_m = \sum_{k=1}^{N-1} y_k\,\mathrm{sen}\,k\frac{2m\pi}{N} = y_1\,\mathrm{sen}\frac{2m\pi}{N} + \ldots$$

$$\ldots + y_{N-1}\,\mathrm{sen}\,(N-1)\frac{2m\pi}{N}.$$

Supongamos que $N = 12$, o sea, el segmento $[0, 2\pi]$ está dividido en 12 partes iguales, así que se utilizan los valores del argumento $0,\ \pi/6,\ \pi/3,\ \ldots,\ 11\pi/6$; a ellos les corresponden los valores de la función $y_0,\ y_1,\ y_2,\ \ldots,\ y_{11}$, y las magnitudes, por las cuales se multiplican estos valores, son tales: ± 1; $\pm\mathrm{sen}\,(\pi/6) = \pm 0{,}5$; $\pm\mathrm{sen}\,(\pi/3) = \pm 0{,}866$. De aquí, omitiendo los cálculos voluminosos, obtenemos

$$6a_0 = y_0 + y_1 + y_2 + y_3 + y_4 + y_5 + y_6 + y_7 + y_8 + y_9 + y_{10} + y_{11},$$

$$6a_1 = (y_2 + y_{10} - y_4 - y_8)\,\mathrm{sen}\frac{\pi}{6} +$$

$$+ (y_1 + y_{11} - y_5 - y_7)\,\mathrm{sen}\frac{\pi}{3} + (y_0 - y_6),$$

$$6a_2 = (y_1 + y_5 + y_7 + y_{11} - y_2 - y_4 - y_8 - y_{10}) \operatorname{sen} \frac{\pi}{6} +$$
$$+ (y_0 + y_6 - y_3 - y_9),$$
$$6a_3 = y_0 + y_4 + y_8 - y_2 - y_6 - y_{10},$$
$$6b_1 = (y_1 + y_5 - y_7 - y_{11}) \operatorname{sen} \frac{\pi}{6} +$$
$$+ (y_2 + y_4 - y_8 - y_{10}) \operatorname{sen} \frac{\pi}{3} + (y_3 - y_9),$$
$$6b_2 = (y_1 + y_2 + y_7 + y_8 - y_4 - y_5 - y_{10} - y_{11}) \operatorname{sen} \frac{\pi}{3},$$
$$6b_3 = y_1 + y_5 + y_9 - y_3 - y_7 - y_{11},$$

etc.

Para reducir al mínimo el número de operaciones aritméticas necesarias con el fin de obtener los valores a_m y b_m, se utiliza un esquema de cálculos especial llamado **esquema de Runge.**

Paso I. Se escriben los valores de la función $f(x)$ observando el orden siguiente:

$$\begin{array}{c} y_0 y_1 y_2 y_3 y_4 y_5 y_6 \\ y_{11} y_{10} y_9 y_8 y_7 \end{array}$$

Paso II. Se cuentan las sumas y diferencias de cada par de los valores que están uno debajo del otro. Las sumas y diferencias obtenidas se escriben del modo siguiente:

$$\begin{array}{ll} & y_0 y_1 y_2 y_3 y_4 y_5 y_6 \\ & \underline{y_{11} y_{10} y_9 y_8 y_7} \\ \text{Sumas} & u_0 u_1 u_2 u_3 u_4 u_5 u_6 \\ \text{Diferencias} & v_1 v_2 v_3 v_4 v_5 \end{array} \qquad (6)$$

Paso III. Las operaciones análogas se realizan con las sumas y diferencias (6)

$$\begin{array}{lll} & u_0 u_1 u_2 u_3 & v_1 v_2 v_3 \\ & \underline{u_6 u_5 u_4} & \underline{v_5 v_4} \\ \text{Sumas} & c_0 c_1 c_2 c_3 & g_1 g_2 g_3 \\ \text{Diferencias} & d_0 d_1 d_2 & h_1 h_2 \end{array} \qquad (7)$$

Paso IV. Se calculan los valores a_m y b_m con ayuda de las fórmulas aproximadas

$$\left.\begin{array}{l} 6a_0 = c_0 \;\; c_1 + c_2 + c_3, \\ 6a_1 = d_0 + 0{,}866 d_1 + 0{,}5 d_2, \\ 6a_2 = (c_0 - c_3) + 0{,}5\,(c_1 - c_2), \\ 6a_3 = d_0 - d_2, \\ 6b_1 = 0{,}5 g_1 + 0{,}866 g_2 + g_3, \\ 6b_2 = 0{,}866\,(h_1 - h_2), \\ 6b_3 = g_1 - g_3, \end{array}\right\} \qquad (8)$$

etc.

Para comparar de un modo evidente los coeficientes a_m y b_m obtenidos por las fórmulas aproximadas con los valores exactos de los mismos, citemos un ejemplo en el cual la función está representada analíticamente.

Ejemplo. Consideremos la función periódica con el período de 2π:

$$y = f(x) \begin{cases} x/\pi & \text{para } 0 \leqslant x \leqslant \pi, \\ 1 & \text{para } \pi < x < 2\pi, \\ 0 & \text{para } \qquad x = 2\pi. \end{cases}$$

Hagamos la tabla:

x_k	0	$\frac{\pi}{6}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$	$\frac{2\pi}{3}$	$\frac{5\pi}{6}$	π	$\frac{7\pi}{6}$	$\frac{4\pi}{3}$	$\frac{3\pi}{2}$	$\frac{5\pi}{3}$	$\frac{11\pi}{6}$	2π
y_k	0	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{5}{6}$	1	1	1	1	1	1	0

Escribamos según el esquema de Runge los valores y_k y realicemos las adiciones y sustracciones indicadas en éste (véanse las fórmulas (6)]

$$\begin{array}{lccccccc} & 0 & \frac{1}{6} & \frac{1}{3} & \frac{1}{2} & \frac{2}{3} & \frac{5}{6} & 1 \\ & & 1 & 1 & 1 & 1 & 1 & \\ \hline \text{Sumas} & 0 & \frac{7}{6} & \frac{4}{3} & \frac{3}{2} & \frac{5}{3} & \frac{11}{6} & 1 \\ \text{Diferencias} & & -\frac{5}{6} & -\frac{2}{3} & -\frac{1}{2} & -\frac{1}{3} & -\frac{1}{6} & \end{array}$$

Ahora realicemos las sustracciones y adiciones respecto a las sumas y diferencias halladas [véanse las fórmulas (7)]:

$$\begin{array}{lcccc} & & \text{sumas} & & \\ & 0 & \frac{7}{6} & \frac{4}{3} & \frac{3}{2} \\ & 1 & \frac{11}{6} & \frac{5}{3} & \\ \hline \text{sumas} & 1 & 3 & 3 & \frac{3}{2} \\ \text{diferencias} & -1 & -\frac{2}{3} & -\frac{1}{3} & \end{array} \qquad \begin{array}{lccc} & & \text{diferencias} & \\ & -\frac{5}{6} & -\frac{2}{3} & -\frac{1}{2} \\ & -\frac{1}{6} & -\frac{1}{3} & \\ \hline \text{sumas} & -1 & -1 & -\frac{1}{2} \\ \text{diferencias} & -\frac{2}{3} & -\frac{1}{3} & \end{array}$$

Escribamos ahora las expresiones para a_m y b_m;

$6a_0 = 1+3+3+\frac{3}{2}$; $\qquad 6b_1 = 0{,}5(-1)+0{,}866(-1)-\frac{1}{2}$;

$6a_1 = -1-\frac{2}{3}\cdot 0{,}866-0{,}5\cdot\frac{1}{3}$; $\qquad 6b_2 = 0{,}866\left(-\frac{2}{3}-\frac{1}{3}\right)$;

$6a_2 = \left(1-\frac{3}{2}\right)+0{,}5\,(3-3)$; $\qquad 6b_3 = -1-\left(-\frac{1}{2}\right)$;

$6a_3 = -1-\left(-\frac{1}{3}\right)$;

De aquí

$$a_0 = 1{,}417;\ a_1 = -0{,}291;\ a_2 = -0{,}083;\ a_3 = -0{,}111,$$
$$b_1 = -0{,}311;\ b_2 = -0{,}144;\ b_3 = -0{,}083.$$

Para comparar damos los valores exactos de los coeficientes:

$$a_0 = 1{,}500;\ a_1 = -0{,}203;\ a_2 = 0{,}000;\ a_3 = -0{,}022;$$
$$b_1 = -0{,}318;\ b_2 = -0{,}159;\ b_3 = -0{,}106.$$

Para obtener con una exactitud más alta los valores de los coeficientes haciendo uso de las fórmulas aproximadas, se pueden utilizar los esquemas provistos de un mayor número de ordenadas.

Cabe señalar que el análisis armónico práctico ofrece la posibilidad de obtener expresiones analíticas que aproximan las funciones dadas con un error estándar mínimo.

§ 7.19. Interpolación inversa

La interpolación de las funciones resulta un aparato útil para resolver varios problemas. El problema de determinación de la raíz de una ecuación o de la de una función es un ejemplo característico de utilización del polinomio interpolador.

Consideremos el siguiente problema de interpolación inversa. Sobre la red Λ_{2k}: $a \leqslant \ldots,\ x_{-k} < \ldots < x_0 < \ \ldots < x_k < \ldots$ $\ldots \leqslant b$ está definida por sus valores f_i $(i = f_1, \pm 1, \ldots, \pm k)$ la función $f(x)$ que es continua sobre el segmento $[a, b]$. Se necesita determinar el valor del argumento $x^* \in (x_0, x_1)$ que corresponde al valor dado de la función f, $f^* = f_0 + \theta\,(f_1 - f_0)$; $\theta \in (0, 1)$. Se supone que el intervalo (x_0, x_1) es tan pequeño que x^* es único.

En realidad, aquí se necesita determinar la raíz de la ecuación

$$f(x) = f^*. \tag{1}$$

Una de las vías posibles de resolver este problema consiste en lo siguiente. Aproximemos la función $f(x)$ por su polinomio interpolador $P_n(f, x)$ y reemplecemos la ecuación (1) por la siguiente:

$$P_n(f, x) = f^*. \tag{2}$$

Encontramos la raíz real $\overline{x^*}$ de la ecuación (2), perteneciente al intervalo (x_0, x_1). Prácticamente obtenemos sólo una solución aproximada de la ecuación (2), o sea, el valor $\bar{\bar{x}}^*$. Ahora suponemos que $x^* = \bar{\bar{x}}^*$.

Vamos a estimar el error de tal solución. Supongamos que el error total de interpolación es igual a Δ, o sea,

$$| P_n (f, x) - f(x) | \leqslant \Delta \tag{3}$$

y el error de la solución de la ecuación (2) es igual a ε, o sea,

$$| \bar{x}^* - \bar{\bar{x}}^* | \leqslant \varepsilon \tag{4}$$

Entonces el incremento de la función f en el punto $\bar{x}^*$ puede ser representado en la forma

$$f^* - f(\bar{x}^*) = (x^* - \bar{x}^*) f'(\xi); \; \xi = \bar{x}^* + \theta (x^* - \bar{x}^*), \; \theta \in (0, 1).$$

De aquí, tomando en consideración que $f^* = P_n (f, \bar{x}^*)$, tenemos

$$P_n (f \; \bar{x}^*) - f(x^*) = (x^* - \bar{x}^*) f'(\xi).$$

Suponiendo ahora que $\min\limits_{[x_0, x_1]} | f'(x) | = m_1 > 0$ y utilizando la estimación (3), obtenemos

$$| x^* - \bar{x}^* | \leqslant \Delta / m_1. \tag{5}$$

Luego,

$$| x^* - \bar{\bar{x}}^* | = | x^* - \bar{x}^* + \bar{x}^* - \bar{\bar{x}}^* | \leqslant | x^* - \bar{x}^* | + | \bar{x}^* - \bar{\bar{x}}^* |$$

y en virtud de las desigualdades (4) y (5) encontramos finalmente

$$|x^* - \bar{\bar{x}}^*| \leqslant \frac{\Delta}{m_1} + \varepsilon \tag{6}$$

Ahora bien, tanto la solución del problema planteado como el error (6) se determinan por dos procesos: por la construcción del polinomio interpolador y por la resolución de la ecuación (2), o sea, por la determinación de las raíces del polinomio interpolador.

Puede parecer que estos dos momentos de ningún modo están relacionados entre sí. Sin embargo, esto no es completamente así.

Hay que tener presente que el aumento del grado del polinomio, por un lado, disminuye el error Δ, pero, por el otro lado, aumenta el volumen de trabajo necesario para resolver la ecuación (2).

Por eso el grado del polinomio interpolador debe ser mínimo a condición de que se alcance la precisión requerida.

Al resolver prácticamente el problema de interpolación inversa sobre una red uniforme en calidad de polinomios interpoladores, suelen utilizarse los polinomios de Stringer y de Bessel. En este caso la ecuación (2) escrita respecto a la variable $t = (x - x_0)/h$ se reduce a la forma $t = \varphi(t)$ y se resuelve por el método de iteraciones.

Al utilizar el polinomio de Stringer tenemos

$$t = \frac{1}{\mu f_0^1} \left(f^* - f_0 - \frac{f_0^2}{2!} t^2 - \frac{\mu f_0^3}{3!} t (t^2 - 1^2) - \ldots \right). \tag{7}$$

La utilización del polinomio de Bessel da

$$t = \frac{1}{2} + \frac{1}{f'_{1/2}}\left(f^* - \mu f_{1/2} - \frac{\mu f^{2}_{1/2}}{2!} t(t-1) - \right.$$
$$\left. - \frac{f^{3}_{1/2}}{3!} t\left(t-\frac{1}{2}\right)(t-1) - .\right). \tag{8}$$

En el primer caso como aproximación inicial t^0 se toma $\frac{1}{\mu f'_0}(f^* - f_0)$ y en el segundo, 0,5 ó bien $\frac{1}{2} + \frac{1}{f'_{1/2}}(f^* - \mu f_{1/2})$. Obtenida t^*, o sea la solución de la ecuación (7) ó (8), x^* se determina con ayuda de la fórmula $x^* = x_0 + t^*h$.

Análogamente, en caso de necesidad, se puede obtener la solución del problema planteado haciendo uso del primero o segundo polinomio interpolador de Newton.

Consideremos el segundo método de resolución del problema de interpolación inversa, método fundado en la existencia de la función $g(y)$, inversa a $f(x)$.

Supongamos que la función $g(y)$ es continua junto con una cantidad suficiente de sus derivadas sobre el intervalo mínimo, que contiene los valores $y_i = f_i$ $(i = 0, \pm 1, \ldots)$ e $y^* = f^*$. En este caso la determinación de x es equivalente al cálculo de la función inversa $g(y)$, definida por sus valores x_i en los nodos y_i y en el punto $y = f^*$, ya que $x^* = g(f^*)$.

Ahora bien, el problema dado está reducido al de interpolación de la función inversa $g(y)$ al cálculo de $g(f^*)$.

El método mencionado de resolución del problema de interpolación inversa es más eficaz que el que contiene la resolución de una ecuación en calidad de una de las etapas. El método en cuestión es sobre todo cómodo si se necesita hallar la solución para una cantidad bastante grande de valores f^* o bien si se necesita obtener la expresión explícita para la raíz de la ecuación (1). El inconveniente del segundo método consiste en la exigencia de que exista una función inversa suave, lo que no siempre puede ser cumplido (por ejemplo, esta exigencia no se cumple para las funciones no monótonas).

En conclusión nótese que para calcular x^* con ayuda de la función inversa el más cómodo es el método de iteración-interpolación de Aitken expuesto en el § 7.13.

Ejemplo 1. Valiéndose de la tabla de valores de la función $f = 3^x$ (véase la tabla del ejemplo dado en el § 7.9,), determinar a qué valor del argumento x^* corresponde el valor de la función $f^* = 5$. Estimar el error.

△ En el ejemplo § 7.9 hemos determinado que el orden de exactitud de la tabla dada es igual a 3. Puesto que el valor dado f^* está situado al fin de la tabla, para calcular x^* es necesario utilizar el segundo polinomio interpolador de Newton que es de tercer grado. Suponiendo $x_0 = 1{,}50$ y $t = (x - x_0)/h$, así como utilizando la fórmula (10) del § 7.9, obtenemos la ecuación para determinar t^*:

$$5 = 5{,}196 + \frac{1{,}248}{1!} t + \frac{0{,}300}{2!} t(t+1) + \frac{0{,}072}{3!} t(t+1)(t+2).$$

En este caso, teniendo en cuenta los resultados del § 7.9, calculemos el error del valor de $\bar{x}^*$ con ayuda de la fórmula (5). Puesto que $\Delta = \Delta_1 + \Delta_2 = 0{,}004$ y $m_1 = \min\limits_{[0,75;\,1,5]} |3^x \ln 3| \approx 2{,}5$, el error buscado constituye $|x^* - \bar{x}^*| \leqslant \Delta/m_1 = 0{,}0016$.

Vamos a transformar la ecuación para determinar t^* en forma cómoda para la aplicación del método de iteraciones

$$t = \frac{5 - 5{,}196}{1{,}248} - \frac{1}{1{,}248}\left[\frac{0{,}300}{2!} t(t+1) + \frac{0{,}072}{3!} t(t+1)(t+2)\right].$$

Resolvamos esta ecuación, tomando por aproximación inicial $t_0 = (5 - 5{,}196) : 1{,}248 \approx -0{,}16$. Entonces

$$t_1 = -0{,}16 - \frac{1}{1{,}248}\left[\frac{0{,}300}{2!}(-0{,}16)(0{,}84) + \right.$$

$$\left. + \frac{0{,}072}{3!}(-0{,}16)\cdot(0{,}84)(1{,}84)\right] = -0{,}141;$$

$$t_2 = -0{,}16 - \frac{1}{1{,}248}\left[\frac{0{,}300}{2!}(-0{,}141)(0{,}859) + \right.$$

$$\left. + \frac{0{,}072}{3!}(-0{,}141)(0{,}859)(1{,}859)\right] = -0{,}143.$$

Ahora bien, en calidad de solución aproximada de la ecuación se puede tomar el valor $t^* = -0{,}14 \pm 0{,}003$. De aquí

$$\bar{\bar{x}}^* = x_0 + t^*h = 1{,}465$$

y el error de la solución de la ecuación

$$|\bar{x}^* - \bar{\bar{x}}^*| = \varepsilon < 0{,}0008,$$

Ahora bien, la solución final $x^* = 1{,}465 \pm 0{,}003$. ▲

Ejemplo 2. Haciendo uso de la tabla de valores de la función $f = \ln x$ (véase el ejemplo 1 del § 7.13), calcular e^2 con exactitud de hasta 0,01.

△ La función dada f tiene la función inversa $g(y) = e^y$ que es continua junto con sus derivadas sobre el intervalo $(-\infty, +\infty)$. Por eso el cálculo de e^y puede ser reducido al cálculo en el punto $y = y^* = 2$ de la función e^y representada en forma de la tabla:

y	0,00	0,69	1,39	1,61	2,08	2,30
g	1	2	4	5	8	10

Vamos a resolver el problema de interpolación obtenido con ayuda del método de Aitken. Numeremos los nodos y_i en el orden siguiente: $y_0 = 2{,}08$; $y_1 = 2{,}30$; $y_2 = 1{,}61$; $y_3 = 1{,}39$; $y_4 = 0{,}69$; $y_5 = 0{,}00$. Utilizando ahora la fórmula (1) del § 7.13 y reemplazando x por y y x_m por y_m, calculamos los valores de los polinomios interpoladores P_n (y^*):

$$P_1^{01}(2)=\frac{1}{0,22}\cdot\begin{vmatrix}-0,08 & 8\\ -0,30 & 10\end{vmatrix}=7,27;$$

$$P_1^{02}(2)=\frac{1}{0,47}\cdot\begin{vmatrix}0,39 & 5\\ -0,08 & 8\end{vmatrix}=7,49;$$

$$P_2^{012}(2)=\frac{1}{0,69}\cdot\begin{vmatrix}0,39 & 7,49\\ -0,30 & 7,27\end{vmatrix}=7,37;$$

$$P_1^{23}(2)=\frac{1}{0,22}\cdot\begin{vmatrix}0,61 & 4\\ 0,39 & 5\end{vmatrix}=6,77;$$

$$P_2^{023}(2)=\frac{1}{0,69}\cdot\begin{vmatrix}0,61 & 6,77\\ 0,08 & 7,49\end{vmatrix}=7,41;$$

$$P_3^{0123}(x)=\frac{1}{0,91}\cdot\begin{vmatrix}0,61 & 7,41\\ -0,30 & 7,37\end{vmatrix}=7,38.$$

Puesto que $|P_3^{0123}(2)-P_2^{012}(2)|=0,01$ y la exactitud requerida está alcanzada, cesamos los cálculos y suponemos $e^2=7,38\pm$ $\pm\,0,01$. ▲

Ejercicios

1. La función $y=f(x)$ se da tabularmente:

x	1,522	1,523	1,524
y	20,477	20,906	21,354

Determinar su valor en el punto $x=1,5228$ con ayuda de la primera fórmula de interpolación de Newton.

2. La función $y=f(x)$ se da tabularmente:

x	1,529	1,530	1,531
y	23,911	24,498	25,115

Determinar su valor en el punto $x=1,5303$, haciendo uso de la segunda fórmula de interpolación de Newton.

3. Construir el polinomio interpolador de Lagrange para la función representada tabularmente:

x	−2	−1	2	3
y	−12	−8	3	5

4. Construir el polinomio interpolador de Lagrange para la función $f(x)=e^{-x}$ si de nodos de interpolación sirven los puntos $x_0=1$, $x_1=2$, $x_2=3$. Estimar el error para $x=1,5$.

5. Hacer la tabla de diferencias finitas para la función representada tabularmente:

x	-2	-1	0	1	2	3
y	10	5	1	-15	-20	-100

6. Hacer la tabla de diferencias divididas para la función representada tabularmente:

x	-3	1	0	2	3
y	-15	-7	1	25	47

7. Para la función $y = f(x)$ representada en la forma tabular:

x	1,03	1,08	1,016	1,23	1,26	1,33	1,39
y	2,80107	2,94468	3,18993	3,42123	3,52542	3,78104	4,01485

calcular el valor en el punto $x = 1{,}21555$ con exactitud de hasta 10^{-5}, haciendo uso de la fórmula de Aitken.

8. Calcular el valor de la función en el punto $x = 1{,}34627$ utilizando la fórmula de Stirling si la función $y = f(x)$ está representada en la forma tabular:

x	1,355	1,340	1,345	1,350	1,355	1,360
y	4,16206	4,25562	4,35325	4,45522	4,56184	4,67433

9. Para la función representada en la forma tabular:

x	1,435	1,440	1,445
y	0,892687	0,893698	0,894700

determinar el valor del argumento, correspondiente al de la función 0,892914.

10. Haciendo uso del método de interpolación inversa, hallar con la exactitud de hasta 10^{-5} la raíz de la ecuación $\sqrt{x+1} - \frac{1}{x} = 0$, que está sobre el segmento $[0{,}7;\ 0{,}8]$.

11. Construir un polinomio interpolador que refleje la variación de la exportación realizada por la industria de construcción durante el período de 1981 a 1983, utilizando los datos siguientes:

Años	1981	1982	1983
Exportación, en millones de rublos convertibles	384,6	507,9	477,9

12. Construir un polinomio interpolador que refleje la variación de la importación realizada por la industria de materiales de construcción durante el período de 1981 a 1983, utilizando los datos

Años	1981	1982	1983
Importación, en millones de rublos convertibles	349,1	416,9	430,6

13. Construir un polinomio interpolador que refleje la variación de la importación realizada por la industria de construcción durante el período de 1981 a 1983, utilizando los datos siguientes:

Años	1981	1982	1983
Importación, en millones de rublos convertibles	657,8	980,6	949,6

14. Construir un polinomio interpolador que refleje la variación de la exportación realizada por la energética eléctrica durante el período de 1980 a 1983, utilizando los datos siguientes:

Años	1980	1981	1982	1983
Exportaciones, en millones de rublos convertibles	421,1	469,4	544,6	619,2

15. La importación de los productos de las industrias forestal, papelera y maderera constituyó según los año

Años	1980	1981	1982	1983
Importación, en millones de rublos convertibles	889,2	938,7	883,6	796,8

Construir un polinomio interpolador que refleje la variación de la importación durante el período de 1980 a 1983.

16. Construir un polinomio interpolador que refleje la variación de la exportación realizada por las industrias forestal, papelera y maderera en el período de 1979 a 1981, utilizando los datos siguientes:

Años	1979	1980	1981
Exportación, en millones de rublos convertibles	1742,7	2008,5	1893,4

17. Costruir un polinomio interpolador que refleje la variación de la importación realizada por la industria ligera en el período de 1980 a 1983, utilizando los datos siguientes:

Años	1980	1981	1982	1983
Importación, en millones de rublos convertibles	7085,3	8454,4	8821,1	8760,8

18. El nivel de instrucción de la población urbana ocupada está representado en la tabla:

Años	1939	1959	1979	1983
Tienen instrucción superior por 1000 personas de la población urbana ocupada	32	59	130	142

¿Cuál fue el nivel de instrucción de la población urbana ocupada en 1982?

19. El nivel de instrucción de la población rural ocupada está representado en la tabla:

Años	1939	1959	1979	1983
Tienen instrucción superior por 1000 personas de la población rural ocupada	3	11	42	50

¿Cuál fue el nivel de instrucción en los lugares rurales en 1982?

20. La cantidad de egresados de los centros de enseñanza superior, especialidad de geología, está representada en la tabla:

Años	1965	1970	1975	1980
Cantidad de especialistas, en miles de personas	3,2	5,1	5,9	6,2

Determinar la cantidad de egresados, especialidad de geología, en 1979.

21. La renta nacional utilizada para la acumulación y otros gastos constituyó

Años	1965	1970	1975	1980
Renta nacional, en miles de millones de rublos	50,2	84,2	96,6	107,2

Determinar la renta nacional en 1979.

22. El consumo de energía eléctrica en la agricultura constituyó:

Años	1965	1970	1975	1980
Consumo de energía eléctrica, en miles de millones de kWh	21,1	38,6	73,8	111

¿Cuál fue el consumo de energía eléctrica en 1979?

23. La fabricación de los motores eléctricos de corriente alterna constituyó:

Años	1970	1975	1980	1982
Fabricación de motores eléctricos con consumo de energía, en millones de kWh	36,3	45,9	51,8	53,3

¿Cuál fue la fabricación de los motores eléctricos (según el consumo de energía) en 1981?

24. Calcular el valor de la función $f(x)$ en el punto $x = x_1$ con ayuda del polinomio interpolador correspondiente, utilizando las tablas de cuatro cifras de funciones trigonométricas con el paso igual a 1°. Determinar el error absoluto del resultado:

a) $f(x) = \operatorname{sen} x$, $x_1 = 37{,}7°$; b) $f(x) = \cos x$, $x_1 = 19°48'$;
c) $f(x) = \operatorname{sen} x$, $x = 53°12'$; d) $f(x) = \cos x$, $x_1 = 36°48'$;
e) $f(x) = \cos x$, $x_1 = 71{,}6°$; f) $f(x) = \operatorname{tg} x$, $x_1 = 67°48'$.

25. Para la función $y = f(x)$ definida sobre el intervalo $(0; \pi)$ y representada en forma de la tabla de valores $y_k = f(x_k)$:

k	0	1	2	3
x_k	0	$\pi/4$	$\pi/2$	$3\pi/4$
y_k	1	2	2,4	2,6

formar el polinomio interpolador trigonométrico.

26. Para la función $y = f(x)$ definida sobre el intervalo $(0, 1)$ y representada en forma de la tabla de valores $y_k = f(x_k)$:

k	0	1	2	3	4	5
x_k	0	1/6	1/3	1/2	2/3	5/6
y_k	1	0	−2	−3	0	2

formar un polinomio trigonométrico cuyo orden no sea inferior al segundo.