ASUPRA PRECIZIEI CALCULULUI NUMERIC ÎN INTERPOLAREA PRIN POLINOAME DE DOUĂ VARIABILE

DE

TIBERIU POPOVICIU
MEMBRU CORESPONDENT AL ACADEMIEI R.P.R.
(Cluj)

Lucrare prezentată la Colocviul de teoria ecuațiilor cu derivate parțiale din 21—26 sept. 1959

1. Să considerăm $(m+1)(n+1)$ puncte (x_i, y_j), $i = 1, 2, \ldots, m+1$, $j = 1, 2, \ldots, n+1$ în plan, formînd o rețea dreptunghiulară de *noduri*. Pentru fixarea notațiilor vom presupune că $x_1 < x_2 < \ldots < x_{m+1}$, $y_1 < y_2 < \ldots < y_{n+1}$. Permutarea $P(r_1, r_2, \ldots, r_{m+1}; s_1, s_2, \ldots, s_{n+1})$ a rețelei de noduri este definită de o permutare $r_1, r_2, \ldots, r_{m+1}$ a indicilor $1, 2, \ldots, m+1$, de o permutare $s_1, s_2, \ldots, s_{n+1}$ a indicilor $1, 2, \ldots, n+1$ și de renumerotarea coordonatelor nodurilor rețelei astfel ca să avem $x'_i = x_{r_i}$, $i = 1, 2, \ldots, m+1$, $y'_j = y_{s_j}$, $j = 1, 2, \ldots, n+1$.

Să introducem notațiile

$$D^{i,j}_{\nu,\mu}[f] = \begin{bmatrix} x'_\nu, & x'_{\nu+1}, & \ldots, & x'_{\nu+i} \\ y'_\mu, & y'_{\mu+1}, & \ldots, & y'_{\mu+j} \end{bmatrix}; f \tag{1}$$

pentru diferențele divizate pe punctele $(x'_{\nu+\alpha}, y'_{\mu+\beta})$, $\alpha = 0, 1, \ldots, i$, $\beta = 0, 1, \ldots, j$, relative la funcția $f(x, y)$.

D.d. (= diferența divizată sau diferențele divizate)

$$D^{i,j}_{\nu,\mu}[f], \ \nu = 1, 2, \ldots, m+1-i, \ \mu = 1, 2, \ldots, n+1-j, \\ i = 0, 1, \ldots, m, \ j = 0, 1, \ldots, n \tag{2}$$

constituie *tabloul* d.d. corespunzător permutării

$$P(r_1, r_2, \ldots, r_{m+1}; s_1, s_2, \ldots, s_{n+1})$$

a rețelei de noduri.

Sistemul de $(m+1)\,(n+1)$ d.d.

$$D^{i,j}_{1,1}\,[f], \qquad i=0,1,\ldots,m,\ j=0,1,\ldots,n \tag{3}$$

formează un *sistem înlănţuit* de d.d.

Dacă fiecare termen al unui sistem înlănţuit de d.d. aparţine unui aceluiaşi tablou de d.d., vom zice că sistemul înlănţuit considerat aparţine acestui tablou de d.d. În particular, sistemul înlănţuit (3) aparţine tabloului (2). Bineînţeles, însă, sistemul înlănţuit (3) aparţine în acelaşi timp şi altor tablouri de d.d.

În particular tabloul (2) corespunzător permutării iniţiale $P(1, 2, \ldots, m+1;\ 1, 2, \ldots, n+1)$ a reţelei de noduri, se va numi tabloul *normal* de d.d. şi orice sistem înlănţuit de d.d. aparţinînd acestui tablou se va numi un sistem înlănţuit *normal* de d.d. În cazul permutării iniţiale avem

$$r_i = i,\ s_j = j,\ D^{i,j}_{\nu,\mu}\,[f] = \left[\begin{matrix} x_\nu, & x_{\nu+1}, & \ldots, & x_{\nu+i} \\ y_\mu, & y_{\mu+1}, & \ldots, & y_{\mu+j}\end{matrix};\ f\right]$$

pentru toate valorile posibile ale lui i, j, ν, μ.

2. Permutării $P(r_1, r_2, \ldots, r_{m+1};\ s_1, s_2, \ldots, s_{n+1})$ a reţelei de noduri, corespunde polinomul de interpolare de două variabile, care sub forma ei generală este [2],

$$\begin{aligned} L(x,y) = \sum_{i=0}^{m}\sum_{j=0}^{n_i} (x-x_{r_1})\,(x-x_{r_2})\ldots(x-x_{r_i})\,(y-y_{s_1})\times \\ \times (y-y_{s_2})\ldots(y-y_{s_j})\,D^{i,j}_{1,1}\,[f], \end{aligned} \tag{4}$$

unde $0 \leqq n_i \leqq n$, $i=0, 1, \ldots, m$ iar pentru $i=0$ respectiv $j=0$ produsul $(x-x_{r_1})\,(x-x_{r_2})\ldots(x-x_{r_i})$ respectiv $(y-y_{s_1})\,(y-y_{s_2})\ldots\times$ $\times(y-y_{s_j})$ este înlocuit cu 1.

Pe un punct de interpolare dat (x, y), determinarea unei valori aproximative pentru $f(x, y)$ cu ajutorul polinomului (4), necesită calculul valorii numerice a acestui polinom. Acest calcul se face prin executarea, exact sau aproximativ, a unei anumite succesiuni de operaţii elementare: adunări, scăderi, înmulţiri şi împărţiri. Deoarece operaţiile se execută în general aproximativ, va rezulta o eroare de calcul care depinde de precizia cu care s-au executat calculele precum şi, bineînţeles, de ordinea efectuării acestor calcule, adică de *programul* pe baza căruia ele au fost executate. Nu vom lua în considerare erorile de care sînt eventual afectate datele problemei, deci coordonatele nodurilor şi valorile funcţiei $f(x, y)$ pe aceste noduri.

3. Pentru a calcula valoarea numerică a polinomului (4), poate fi utilă şi o modificare a expresiei lui, punînd

$$X_0 = 1,\ X_i = \frac{x - x_{r_i}}{k_i},\ Y_0 = 1,\ Y_j = \frac{y - y_{s_j}}{l_j},$$

$$E_{i,j} = k_0\,k_1 \ldots k_i\,l_0\,l_1 \ldots l_j\,D^{i,j}_{1,1}\,[f],\ i = 0, 1, \ldots, m,\ \ j = 0, 1, \ldots, n,$$

unde $k_0\,(=1)$, k_1, $k_2, \ldots, k_m$, $l_0\,(=1)$, l_1, $l_2, \ldots, l_n$ sînt nişte numere date, diferite de 0. Avem atunci

$$L(x,y) = \sum_{i=0}^{m}\sum_{j=0}^{n_i} X_0\,X_1 \ldots X_i\,Y_0\,Y_1 \ldots, Y_j\,E_{i,j},$$

care se calculează, de obicei, efectuînd suma

$$\sum_{i=0}^{m} X_0\,X_1 \ldots X_i\,A_i, \tag{5}$$

în prealabil fiind calculate sumele

$$A_i = \sum_{j=0}^{n_i} Y_0\,Y_1 \ldots Y_j\,E_{i,j}\ \ i = 0, 1, \ldots, m. \tag{6}$$

Fiecare dintre sumele (5), (6) este de forma

$$c_0\,C_0 + c_0\,c_1\,C_1 + \ldots + c_0\,c_1 \ldots c_\alpha\,C_\alpha. \tag{7}$$

Să presupunem că adunările şi scăderile nu comportă erori (ele se efectuează exact sau, în orice caz, cu nişte erori care se neglijează). Atunci dacă se calculează suma (7) pe baza schemei

$$c_0\big(C_0 + c_1\,(C_1 + \ldots + c_{\alpha-1}\,(C_{\alpha-1} + c_\alpha\,C_\alpha)\ldots)\big), \tag{8}$$

aşa cum se procedează de obicei, eroarea provine numai din efectuarea celor $\alpha+1$ înmulţiri succesive cu c_α, $c_{\alpha-1}, \ldots, c_0$ respectiv. Este uşor de văzut că dacă aceste înmulţiri se efectuează cu o eroare absolută, cel mult egală cu ε, rezultatul aproximativ obţinut va avea o eroare absolută $\leqq \varepsilon\,(1 + \sum_{i=0}^{\alpha-1} |\,c_0\,c_1 \ldots c_i\,|)$, ($= \varepsilon$ dacă $\alpha = 0$). În cazurile pe care le vom considera efectiv, înmulţirea prin c_0 nu comportă erori (vom avea totdeauna $c_0 = 1$). În acest caz eroarea absolută a rezultatului este $\leqq \varepsilon \sum_{i=0}^{\alpha-1} |\,c_0\,c_1 \ldots c_i\,|$, (0 dacă $\alpha = 0$).

Să aplicăm această schemă de calcul sumelor (5), (6). Suma (6) este de forma (7), unde punem

$$\alpha = n_i,\quad c_\nu = Y_\nu,\quad C_\nu = E_{i,\nu},\qquad \nu = 0, 1, \ldots, n_i.$$

Dacă deci efectuăm calculele pe baza schemei (8), înmulțirile fiind executate cu o eroare absolută $\leqq \varepsilon$, rezultatul obținut $\tilde{A}_i$ este cu o eroare absolută $\leqq \varepsilon \sum_{j=0}^{n_i-1} | Y_0 Y_1 \dots Y_j |$, $(= 0$, dacă $n_i = 0)$. Rezultă că aproximarea sumei (5),

$$\sum_{i=0}^{m} X_0 X_1 \dots X_i \tilde{A}_i, \tag{9}$$

are o eroare absolută $\leqq \varepsilon \sum_{i=0}^{m} | X_0 X_1 \dots X_i | \left(\sum_{j=0}^{n_i-1} | Y_0 Y_1 \dots Y_j | \right)$.

Suma (9) este de asemenea de forma

$$\alpha = m, \quad c_\nu = X_\nu, \quad C_\nu = \tilde{A}_\nu, \qquad \nu = 0, 1, \dots, m.$$

Dacă deci și aici efectuăm calculele pe baza schemei (8), înmulțirile fiind executate cu o eroare absolută $\leqq \varepsilon$, valoarea aproximativă obținută pentru (9) va avea o eroare absolută cel mult egală cu

$$\varepsilon \sum_{i=0}^{m-1} | X_0 X_1 \dots X_i | \ (= 0, \text{ dacă } m = 0).$$

În definitiv deci, făcînd calculele pe baza schemei indicate, obținem pentru $L(x, y)$ o valoare aproximativă cu o eroare absolută $\leqq \varepsilon M$, unde

$$M = \sum_{i=0}^{m} | X_0 X_1 \dots X_i | \left(\sum_{j=0}^{n_i-1} | Y_0 Y_1 \dots Y_j | \right) + \sum_{i=0}^{m-1} | X_0 X_1 \dots X_i |. \tag{10}$$

4. Ne putem pune problema de a găsi, pentru punctul de interpolare dat (x, y), acea permutare $P(r_1, r_2, \dots, r_{m+1}; s_1, s_2, \dots, s_{n+1})$ a rețelei de noduri pentru care numărul M dat de formula (10) este cel mai mic posibil. Dacă această condiție este îndeplinită vom considera că polinomul (4) este cel mai avantajos pentru calculele numerice (făcute pe baza schemei indicate).

Atît sumele

$$B_i = \sum_{j=0}^{n_i-1} | Y_0 Y_1 \dots Y_j |, \qquad i = 0, 1, \dots m \tag{11}$$

cît și suma (10) sînt de forma

$$S_0 + s_{\nu_1} S_1 + s_{\nu_1} s_{\nu_2} S_2 + \dots + s_{\nu_1} s_{\nu_2} \dots s_{\nu_\alpha} S_\alpha, \tag{12}$$

unde $S_0, S_1, \dots, S_\alpha$ sînt numere pozitive, iar $s_{\nu_1}, s_{\nu_2}, \dots, s_{\nu_\alpha}$ este o permutare a unui șir $s_1, s_2, \dots, s_\alpha$ de numere pozitive.

Șirul (11) este de forma (12), unde

$$\alpha = n_i - 1, \quad s_j = | y - y_j |, \quad \nu_j = s_j, \quad S_j = \frac{1}{| l_0 l_1 \dots l_j |},$$
$$j = 1, 2, \dots, n_i - 1, \quad S_0 = 1,$$

iar suma (10) este de forma (12), unde

$$\alpha = m, \ s_i = | x - x_i |, \ \nu_i = r_i, \ S_i = \frac{1}{| k_0 k_1 \dots k_i |} \left(1 + \sum_{j=0}^{n_i-1} | Y_0 Y_1 \dots Y_j | \right),$$
$$i = 0, 1, \dots, m-1, \ S_m = \frac{1}{| k_0 k_1 \dots k_m |} \left(\sum_{j=0}^{n_m-1} | Y_0 Y_1 \dots Y_j | \right), \ \nu_m = r_m.$$

Avem acum următoarea

LEMĂ. *Suma (12) își atinge cea mai mică valoare a sa dacă șirul* $s_{\nu_1}, s_{\nu_2}, \dots, s_{\nu_\alpha}$ *este nedescrescător* (cu alte cuvinte pentru o permutare nedescrescătoare a șirului $s_1, s_2, \dots, s_\alpha$).

Demonstrația este imediată. Se vede ușor că minimul este atins numai pentru permutările nedescrescătoare ale șirului $s_1, s_2, \dots, s_\alpha$.

Din lema precedentă rezultă atunci următoarea

TEOREMĂ. *Dacă permutarea* $P(r_1, r_2, \dots, r_{m+1}; s_1, s_2, \dots, s_{n+1})$ *este astfel încît șirurile*

$$| x - x_{r_1} |, | x - x_{r_2} |, \dots, | x - x_{r_m} |; | y - y_{s_1} |, | y - y_{s_2} |, \dots, | y - y_{s_{n-1}} |$$

să fie nedescrescătoare, numărul M își atinge cea mai mică valoare a sa.

5. Rezultă din cele ce preced că dacă șirurile

$$\begin{gathered} | x - x_{r_1} |, | x - x_{r_2} |, \dots, | x - x_{r_{m+1}} |; | y - y_{s_1} |, \\ | y - y_{s_2} |, \dots, | y - y_{s_{n+1}} | \end{gathered} \tag{13}$$

sînt nedescrescătoare, polinomul (4) este cel mai avantajos pentru calculul unei valori aproximative a lui $f(x, y)$. În acest caz, pe baza unui rezultat anterior [1], permutarea $r_1, r_2, \dots, r_{m+1}$ trebuie să fie consecutivă permutării $1, 2, \dots, m + 1$ iar permutarea $s_1, s_2, \dots, s_{n+1}$ consecutivă permutării $1, 2, \dots, n + 1$. Aceasta înseamnă că pentru orice t, $(t = 1, 2, \dots, m + 1$, respectiv $t = 1, 2, \dots, n + 1)$, șirurile $r_1, r_2, \dots, r_t$ și $s_1, s_2, \dots, s_t$ sînt formate din cîte t numere naturale consecutive (într-o anumită ordine).

Cînd șirurile (13) sînt nedescrescătoare, și în acest caz polinomul (4) este avantajos pentru calcule în sensul arătat mai sus, sistemul înlănțuit de d.d. corespunzător este un sistem înlănțuit normal. Este deci suficient să folosim tabloul normal de d.d., pentru a putea construi toate polinoamele (4) avantajoase, după diferitele poziții ale punctului de interpolare.

Putem, ca și în cazul unei singure variabile [1], să studiem diferite cazuri particulare. Dacă punctul de interpolare este în vecinătatea unuia din nodurile extreme (x_1, y_1), (x_1, y_{n+1}), (x_{m+1}, y_1), (x_{m+1}, y_{n+1}), găsim ca cele mai avantajoase polinoamele ordonate după diferențele ascendente și descendente în raport cu x, y, deci acelea pentru care sistemul înlănțuit (3) corespunde permutărilor $P(1, 2, \ldots, m+1; 1, 2, \ldots, n+1)$, $P(1, 2, \ldots, m+1; n+1, n, \ldots, 1)$, $P(m+1, m, \ldots, 1; 1, 2, \ldots, n+1)$, $P(m+1, m, \ldots, 1; n+1, n, \ldots, 1)$. Dacă punctele x_i și y_j sînt respectiv echidistante, deci dacă $x_i = x_1 + (i-1)h$, $i = 1, 2, \ldots, m+1$, $y_j = y_1 + (j-1)h'$, $j = 1, 2, \ldots, n+1$, $(h, h' > 0)$ și dacă punctul de interpolare este în vecinătatea centrului rețelei de noduri, găsim ca cele mai avantajoase polinoame (4), acelea ordonate după diferențele centrale în raport cu x și y.

În încheiere observăm că, în loc de tabloul normal de d.d., putem folosi tabloul format cu numerele $k_0 k_1 \ldots k_j\, l_0 l_1 \ldots l_j\, D^{i,j}_{\nu,\mu}[f]$, d.d. fiind referitoare la permutarea inițială a rețelei de noduri. Dacă luăm $k_i = ih$, $l_j = jh'$, pentru orice $i, j \geqq 1$, numerele considerate se reduc la diferențele $\Delta^{i,j}_{h,h'} f(x, y)$ ale funcției $f(x, y)$. Tabloul normal de d.d. se poate atunci înlocui cu tabloul diferențelor, a cărui formare este deosebit de simplă deoarece nu necesită decît scăderi succesive.

În sfîrșit, considerațiile precedente se pot extinde la cazul a mai mult de două variabile independente.

О ТОЧНОСТИ ЧИСЛОВОГО РАСЧЕТА ПРИ ИНТЕРПОЛИРОВАНИИ МНОГОЧЛЕНАМИ ОТ ДВУХ ПЕРЕМЕННЫХ

РЕЗЮМЕ

Вычисление значения интерполяционного многочлена (4) при заданных x, y проводится путем вычислений суммы (5), где числа A_i даны соотношением (6). Здесь числа X_i, Y_j, $E_{i,j}$ даны в работе и связаны с функцией $f(x, y)$ через разделенные разности (3), соответствующие перестановке (x_{r_i}, y_{s_j}) узлов интерполяции. Суммы (5), (6) имеют вид (7). Вычисление такой суммы проводится при помощи схемы (8), где только умножения на $c_\alpha, c_{\alpha-1}, \ldots, c_0$ сопряжены с ошибками, меньшими числа ε по абсолютной величине. Тогда наиболее выгодные для вычислений многочлены (4) — это те, для которых последовательности (13) являются неубывающими. Таким образом, мы приходим к теоретическому обоснованию практического использования наиболее известных интерполяционных формул, как например, формулы Ньютона, Стирлинга, Бесселя и т.д.

SUR LA PRÉCISION DU CALCUL NUMÉRIQUE DANS L'INTERPOLATION PAR POLYNÔMES À DEUX VARIABLES

RÉSUMÉ

Le calcul de la valeur du polynôme d'interpolation (4), dans le cas de x, y donnés, se fait en calculant la somme (5) où les nombres A_i sont donnés par la relation (6). Les nombres X_i, Y_j, $E_{i,j}$ sont donnés dans le texte et sont liés à la fonction $f(x, y)$ par l'intermédiaire des différences divisées (3) correspondant à une permutation (x_{r_i}, y_{s_j}) des nœuds d'interpolation. Les sommes (5), (6) sont de la forme (7). Le calcul d'une telle somme se fait au moyen du schéma (8) où seules les multiplications par $c_\alpha, c_{\alpha-1}, \ldots, c_0$ sont affectées d'erreurs inférieures en valeur absolue au nombre ε. Dans ce cas les polynômes (4) les plus avantageux pour les calculs sont ceux pour lesquels les suites (13) sont non décroissantes. On arrive ainsi à justifier théoriquement l'emploi dans la pratique des formules d'interpolation bien connues, telles que celles de Newton, Stirling, Bessel, etc.

BIBLIOGRAFIE

1. POPOVICIU T., *Considerații teoretice asupra utilizării practice a unor formule de interpolare.* Bul. științ. Acad. R.P.R., Secțiunea de științe matematice și fizice, **III**, 441–449 (1951).
2. STEFFENSEN J. F., *Interpolation.* New York, 1950.

Primit la 3.XII.1959