Calcul Numeric

Cursul 9

2020

Anca Ignat

Descompunerea după valori singulare
(**S**ingular **V**alue **D**ecomposition)

Teoremă

Fie $\boldsymbol{A} \in \mathbb{R}^{m\times n}$. Atunci există o matrice ortogonală pătratică de dimensiune $\boldsymbol{m}$, $\boldsymbol{U} \in \mathbb{R}^{m\times m}$, o matrice ortogonală pătratică de dimensiune $\boldsymbol{n}$, $\boldsymbol{V} \in \mathbb{R}^{n\times n}$ şi constantele pozitive:

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0, \qquad r \leq \min\{m,n\} \text{ a.î.}$$

$$A = U\Sigma V^T, \quad \Sigma \in \mathbb{R}^{m\times n}, \Sigma = \begin{bmatrix} D & \mathbf{0}_{r\times(n-r)} \\ \mathbf{0}_{(m-r)\times r} & \mathbf{0}_{(m-r)\times(n-r)} \end{bmatrix}, \qquad (1)$$

$$D \in \mathbb{R}^{r\times r}, D = \mathbf{diag}\left[\sigma_1,\ldots,\sigma_r\right]$$

Constanta ***r*** este chiar rangul matricii ***A***, ***r* = rang(*A*)**.
Constantele ***σ₁, σ₂, … , σᵣ*** poartă numele de ***valori singulare*** ale matricii ***A***.
Folosind relaţia (1) avem:

$$A^T = \left(U\Sigma V^T\right)^T = V\Sigma^T U^T,$$

$$AA^T = U\Sigma V^T V\Sigma^T U^T = U\Sigma\Sigma^T U^T = U\Lambda_m U^T ,$$

$$\Lambda_m = \Sigma\Sigma^T = \begin{bmatrix} D^2 & 0_{r\times(m-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(m-r)} \end{bmatrix} \in \mathbb{R}^{m\times m}$$

$$\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Sigma}^T\boldsymbol{U}^T\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T = \boldsymbol{V}\boldsymbol{\Sigma}^T\boldsymbol{\Sigma}\boldsymbol{V}^T = \boldsymbol{V}\boldsymbol{\Lambda}_n\boldsymbol{V}^T \quad ,$$

$$\boldsymbol{\Lambda}_n = \boldsymbol{\Sigma}^T\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{D}^2 & \mathbf{0}_{r\times(n-r)} \\ \mathbf{0}_{(n-r)\times r} & \mathbf{0}_{(n-r)\times(n-r)} \end{bmatrix} \in \mathbb{R}^{n\times n}$$

Ţinând cont de ortogonalitatea matricilor ***U*** şi ***V***, putem rescrie relaţiile de mai sus astfel:

$$(\boldsymbol{A}\boldsymbol{A}^T)\boldsymbol{U} = \boldsymbol{U}\,\boldsymbol{\Lambda}_m \quad , \ \boldsymbol{\Lambda}_m = \mathbf{diag}\left[\sigma_1^2, \sigma_2^2, \ldots, \sigma_r^2, 0, \ldots, 0\right] \in \mathbb{R}^{m\times m}$$

$$(\boldsymbol{A}^T\boldsymbol{A})\boldsymbol{V} = \boldsymbol{V}\,\boldsymbol{\Lambda}_n \quad , \ \boldsymbol{\Lambda}_n = \mathbf{diag}\left[\sigma_1^2, \sigma_2^2, \ldots, \sigma_r^2, 0, \ldots, 0\right] \in \mathbb{R}^{n\times n}$$

Din aceste relaţii deducem că $\boldsymbol{\sigma_1^2, \sigma_2^2, \ldots, \sigma_r^2}$ sunt valorile proprii strict pozitive ale matricilor $\boldsymbol{AA^T}$ şi/sau $\boldsymbol{A^TA}$ iar matricile $\boldsymbol{U}$ şi $\boldsymbol{V}$ sunt matrici ale căror coloane sunt vectorii proprii asociaţi (cei ce formează baze ortonormate). Matricile $\boldsymbol{AA^T}$ şi $\boldsymbol{A^TA}$ sunt matrici simetrice:

$$\boldsymbol{\left(AA^T\right)^T = \left(A^T\right)^T A^T = AA^T \quad , \quad \left(A^T A\right)^T = A^T \left(A^T\right)^T = A^T A}$$

şi au toate valorile proprii nenegative:

$$\left(AA^T\right)u = \lambda u \Rightarrow \left(AA^T u, u\right) = \left(\lambda u, u\right) \Rightarrow$$

$$\lambda = \frac{\left(A^T u, A^T u,\right)}{\left(u, u\right)} = \frac{\left\|A^T u\right\|_2^2}{\left\|u\right\|_2^2} \geq 0$$

Putem folosi descompunerea după valori singulare pentru a defini pseudo-inversa unei matrici oarecare $A \in \mathbb{R}^{m \times n}$ ($n \neq m$).

$$A = U\Sigma V^T \quad , \quad A^{-1} =_? \left(U\Sigma V^T\right)^{-1} = \left(V^T\right)^{-1} \Sigma_?^{-1} U^{-1} = V\Sigma_?^{-1} U^T$$

Rămâne de definit matricea $\Sigma_?^{-1}$. Urmând acest raţionament se defineşte ***pseudoinversa Moore-Penrose*** a unei matrici

$\boldsymbol{A} \in \mathbb{R}^{m \times n}$ astfel:

$$\boldsymbol{A}^{I} = \boldsymbol{V}\boldsymbol{\Sigma}^{I}\boldsymbol{U}^{T} \quad , \; \boldsymbol{A}^{I} \in \mathbb{R}^{n \times m} \quad , \quad \boldsymbol{\Sigma}^{I} = \begin{bmatrix} \boldsymbol{D}^{-1} & 0_{r \times (m-r)} \\ \mathbf{0}_{(n-r) \times r} & 0_{(n-r) \times (m-r)} \end{bmatrix} \in \mathbb{R}^{n \times m} ,$$

$$\boldsymbol{D}^{-1} \in \mathbb{R}^{r \times r}, \boldsymbol{D}^{-1} = \mathbf{diag}\left[\frac{1}{\sigma_1}, \ldots, \frac{1}{\sigma_r}\right].$$

Pseudoinversa definită mai sus satisface următoarele proprietăţi:

$$\left(\boldsymbol{A}^{I}\right)^{I} = \boldsymbol{A} \quad , \; \forall \boldsymbol{A} \in \mathbb{R}^{m \times n} \quad ; \; \left(\boldsymbol{A}^{T}\right)^{I} = \left(\boldsymbol{A}^{I}\right)^{T} , \; \forall \boldsymbol{A} \in \mathbb{R}^{m \times n}$$

$$\boldsymbol{A}\boldsymbol{A}^{I}\boldsymbol{A} = \boldsymbol{A} \qquad , \qquad \boldsymbol{A}^{I}\boldsymbol{A}\boldsymbol{A}^{I} = \boldsymbol{A}^{I}$$

Există o proprietate care nu mai este satisfăcută de psudoinversă deşi este respectată de inversa clasică:

$$\exists\, \boldsymbol{A}, \boldsymbol{B} \ \textbf{a.}\text{î.} \quad \left(\boldsymbol{AB}\right)^{\boldsymbol{I}} \neq \boldsymbol{B}^{\boldsymbol{I}}\boldsymbol{A}^{\boldsymbol{I}}.$$

Descompunerea după valori singulare poate fi utilizată şi pentru rezolvarea sistemelor liniare cu matrici oarecare (***m≠n***)

$$\boldsymbol{Ax} = \boldsymbol{b} \quad , \quad \boldsymbol{A} \in \mathbb{R}^{\boldsymbol{m}\times\boldsymbol{n}} \ , \ \boldsymbol{b} \in \mathbb{R}^{\boldsymbol{m}} \quad , \quad \boldsymbol{x} := \boldsymbol{A}^{\boldsymbol{I}}\boldsymbol{b} \in \mathbb{R}^{\boldsymbol{n}}.$$

Problema celor mai mici pătrate

$$A \in \mathbb{R}^{m \times n}\ ,\ b \in \mathbb{R}^{m}\ ,\ Ax = b\ ,\ x \in \mathbb{R}^{n} \tag{1}$$

$$\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\
&a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\
&\vdots \\
&a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i \\
&\vdots \\
&a_{m1}x_1 + a_{m2}x_2 \cdots + a_{mn}x_n = b_m
\end{aligned}$$

Sistemul are soluţii clasice dacă:

$$\textbf{rang } \boldsymbol{A} = \textbf{rang } [\boldsymbol{A} \mid \boldsymbol{b}]$$

- $\boldsymbol{m} < \boldsymbol{n}$ - o infinitate de soluţii
- $\boldsymbol{m} \geq \boldsymbol{n}$
  - dacă **rang** $\boldsymbol{A}$ **= rang [**$\boldsymbol{A}$ **|** $\boldsymbol{b}$**]** soluţii clasice
  - dacă **rang** $\boldsymbol{A} \neq$ **rang [**$\boldsymbol{A}$ **|** $\boldsymbol{b}$**]** soluţii în sensul celor mai mici pătrate

Vectorul reziduu:

$$\boldsymbol{r}(\boldsymbol{x}) = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x} \in \mathbb{R}^{\boldsymbol{m}}$$

Vectorul $\boldsymbol{x} \in \mathbb{R}^{\boldsymbol{n}}$ se numeşte ***soluţie în sensul celor mai mici pătrate*** pentru sistemul (1) dacă este soluţia următoarei probleme de optimizare:

$$\mathbf{min}\{\|\boldsymbol{r}(\boldsymbol{x})\|_2^2 = \|\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}\|_2^2;\ \boldsymbol{x} \in \mathbb{R}^{\boldsymbol{n}}\} \qquad \textbf{\textit{(LSP)}}$$

$$\boldsymbol{A} = \begin{bmatrix} \mathbf{1} & \mathbf{4} \\ \mathbf{2} & \mathbf{5} \\ \mathbf{3} & \mathbf{6} \end{bmatrix} \in \mathbb{R}^{\mathbf{3\times2}} \quad , \quad \boldsymbol{b} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{1} \end{pmatrix},\ \boldsymbol{m} = \mathbf{3},\ \boldsymbol{n} = \mathbf{2}$$

$$\textbf{rang } \boldsymbol{A} = \mathbf{2} \neq \textbf{rang } [\boldsymbol{A} \mid \boldsymbol{b}] = \mathbf{3}$$

Sistemul:

$$\begin{aligned} \boldsymbol{x_1} + \mathbf{4}\boldsymbol{x_2} &= \mathbf{0} \\ \mathbf{2}\boldsymbol{x_1} + \mathbf{5}\boldsymbol{x_2} &= \mathbf{0} \\ \mathbf{3}\boldsymbol{x_1} + \mathbf{6}\boldsymbol{x_2} &= \mathbf{1} \end{aligned} \qquad \mathbf{(2)}$$

nu are soluţie clasică (nu există $\boldsymbol{x_1}$, $\boldsymbol{x_2}$ care să satisfacă toate cele 3 ecuaţii simultan). Vectorul reziduu are forma:

$$r(x)=b-Ax=\begin{pmatrix}0\\0\\1\end{pmatrix}-\begin{bmatrix}1&4\\2&5\\3&6\end{bmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}-x_1-4x_2\\-2x_1-5x_2\\1-3x_1-6x_2\end{pmatrix}$$

Soluţia în sensul celor mai mici pătrate a acestui sistem este definită ca soluţia problemei de optimizare:

$$\min\{(-x_1-4x_2)^2+(-2x_1-5x_2)^2+\left(1-3x_1-6x_2\right)^2;\ x_1,x_2\in\mathbb{R}\}$$

$$\min\{1-6x_1-12x_2+64x_1x_2+14x_1^2+77x_2^2;\ x_1,x_2\in\mathbb{R}\}$$

Această problemă de minimizare are soluţia:

$$x_1 = \frac{13}{18}, \qquad x_2 = -\frac{2}{9}, \qquad \| r(x) \|_2^2 = \frac{1}{6}$$

şi este soluţia în sensul celor mai mici pătrate a sistemului (2).

$$\textbf{range}\left(\mathbf{A}\right) = \{\, \mathbf{y} \in \mathbb{R}^m ; y = a_1 A^1 + a_2 A^2 + \cdots + a_n A^n ,\, a_i \in \mathbb{R},\, i = 1,n \}$$

$$A = \left[A^1 \; A^2 \cdots A^n \right],\; A^i \in \mathbb{R}^m \text{ sunt coloanele matricii } A$$

Teoremă

Fie $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ $\boldsymbol{(m \geq n)}$, $\boldsymbol{b} \in \mathbb{R}^{m}$. Un vector $\boldsymbol{x} \in \mathbb{R}^{n}$ minimizează norma euclidiană a vectorului reziduu $\|\boldsymbol{r(x)}\|_2 = \|\boldsymbol{b - Ax}\|_2$, rezolvând problema ***(LSP)***, dacă şi numai dacă:

$$\boldsymbol{r(x) \perp \mathrm{range}(A) \quad \Leftrightarrow \quad A^T r(x) = 0}$$

sau echivalent

$$\boldsymbol{A^T A x = A^T b} \tag{3}$$

Sistemul (3) poartă numele de sistemul de ***ecuaţii normale***.

Este un sistem pătratic de dimensiune $\boldsymbol{n}$, matricea sistemului $\boldsymbol{A}^T \boldsymbol{A} \in \mathbb{R}^{\boldsymbol{n} \times \boldsymbol{n}}$ este simetrică. Sistemul de ecuaţii normale (3) este nesingular dacă şi numai dacă **rang**$\boldsymbol{A} = \boldsymbol{n}$, în acest caz soluţia $\boldsymbol{x}$ a sistemului (3) este unică.

$$\mathbf{det}\, \boldsymbol{A}^T \boldsymbol{A} \neq \mathbf{0} \iff \mathbf{rang}\, \boldsymbol{A} = \boldsymbol{n}$$

$$\boldsymbol{A} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}, \quad \boldsymbol{A}^T \boldsymbol{A} = \begin{bmatrix} 14 & 32 \\ 32 & 77 \end{bmatrix}, \quad \boldsymbol{A}^T \boldsymbol{b} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}$$

$$\begin{aligned} 14x_1 + 32x_2 &= 3 \\ 32x_1 + 77x_2 &= 6 \end{aligned} \quad \Rightarrow \quad x_1 = \frac{13}{18}, \quad x_2 = -\frac{2}{9}$$

Pseudo-inversa matricii *A*

Presupunem că $\boldsymbol{A}$ are **rang** $\boldsymbol{A = n}$. Atunci pseudo-inversa poate fi definită ca:

$$\boldsymbol{A^{+} = \left(A^{T}A\right)^{-1}A^{T} \in \mathbb{R}^{n\times m}} \qquad \boldsymbol{(A^{+} = A^{I} \quad ?)}$$

$$\boldsymbol{A^{+} = \begin{bmatrix} 14 & 32 \\ 32 & 77 \end{bmatrix}^{-1} * \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}}$$

Rezolvarea sistemului de ecuaţii normale

1) Folosind factorizarea Cholesky (descompunere ***LU***) pentru matrici simetrice:

$$\boldsymbol{A^T A = LL^T}, \quad \boldsymbol{L} \in \mathbb{R}^{n \times n}$$ **matrice inferior triunghiulară**

- Se calculează matricea $\boldsymbol{A^T A}$ şi vectorul $\boldsymbol{A^T b}$;
- Se calculează factorizarea Cholesky a matricii $\boldsymbol{A^T A = LL^T}$;
- Se rezolvă sistemul inferior triunghiular $\boldsymbol{Ly = A^T b}$ pentru $\boldsymbol{y}$;
- Se rezolvă sistemul superior triunghiular $\boldsymbol{L^T x = y}$ pentru $\boldsymbol{x}$;

2) Se calculează descompunerea ***QR*** (cu algoritmul lui

Householder adaptat) pentru matricea $\boldsymbol{A}$:

$\boldsymbol{A} = \boldsymbol{QR}$, $\boldsymbol{Q} \in \mathbb{R}^{m \times m}$ **matrice ortogonală** , $\boldsymbol{R} \in \mathbb{R}^{m \times n}$,

$$\boldsymbol{R} = \begin{bmatrix} \bar{\boldsymbol{R}} \in \mathbb{R}^{n \times n} \\ \boldsymbol{0}_{(m-n) \times n} \end{bmatrix} \quad , \quad \bar{\boldsymbol{R}} - \textbf{matrice superior triunghiulară}$$

- Se calculează factorizarea $\boldsymbol{QR}$ modificată a matricii $\boldsymbol{A}$;
- Se calculează vectorul $\boldsymbol{Q}^T \boldsymbol{b}$ **;**
- Se rezolvă sistemul sup. triunghiular $\bar{\boldsymbol{R}} \boldsymbol{x} = \left(\boldsymbol{Q}^T \boldsymbol{b} \right)_{i=1,n}$;

3) Se foloseşte desc. după valori singulare a matricii $\boldsymbol{A}$

$$\boldsymbol{A} = \boldsymbol{U\Sigma V}^{\boldsymbol{T}}, \quad \boldsymbol{\Sigma} \in \mathbb{R}^{\boldsymbol{m \times n}}, \boldsymbol{U} \in \mathbb{R}^{\boldsymbol{m \times m}}, \boldsymbol{V} \in \mathbb{R}^{\boldsymbol{n \times n}}$$

- Se calculează SVD pentru matricea $\boldsymbol{A=U\Sigma V^T}$;
- Se calculează vectorul $\boldsymbol{U^T b}$;
- Se rezolvă sistemul diagonal $\boldsymbol{\Sigma w = U^T b}$ pentru $\boldsymbol{w}$;
- Soluţia este $\boldsymbol{x=Vw}$;

1), 2) sau 3) ? → se recomandă **2)**

Interpolare numerică

Presupunem că despre o funcţie ***f*** cunoaştem doar valorile într-un număr finit de puncte. Pornind de la aceste date, dorim să aproximăm funcţia ***f*** într-un alt punct.

x	x_0	x_1	x_2	...	x_{n-1}	x_n
f	y_0	y_1	y_2	...	y_{n-1}	y_n

În tabelul de mai sus $f(x_i) = y_i$, $i=0,1,\ldots,n$ şi $x_i \neq x_j$, $i\neq j$.

Dat un punct $x\neq x_i$, $i=0,1,\ldots,n$ dorim să aproximăm $f(x)$ cunoscând cele $(n+1)$ perechi (x_i ,y_i), $i=0,\ldots,n$. Punctele x_i se numesc ***noduri de interpolare***.

Polinomul de interpolare Lagrange

Notăm cu Π_n mulțimea polinoamelor de grad cel mult n. Dimensiunea acestui spațiu este $n+1$, baza uzuală fiind dată de polinoamele $1, x, x^2,\ldots, x^n$. Vom considera o altă bază în acest spațiu. Se consideră polinoamele p_i:

$$p_i \in \Pi_n \text{ astfel ca } p_i(x_j) = \begin{cases} 0 & \text{pentru } j \neq i \\ 1 & \text{pentru } j = i \end{cases}, \ j = 0,\ldots,n,\ i = 0,\ldots,n$$

Din relația $p_i(x_j)=0,\ \forall j \neq i$ şi faptul că p_i este de grad n rezultă că $x_0, x_1,\ldots,x_{i-1}, x_{i+1},\ldots,x_n$ sunt cele n rădăcini ale polinomului p_i.

Avem:

$$p_i(x) = c_i(x - x_0)\cdots(x - x_{i-1})(x - x_{i+1})\cdots(x - x_n),$$
$$c_i \in \mathbb{R},\ i = 0,\ldots,n$$

Constanta c_i se determină din relaţia $p_i(x_i) = 1$:

$$p_i(x_i) = 1 = c_i(x_i - x_0)\cdots(x_i - x_{i-1})(x_i - x_{i+1})\cdots(x_i - x_n) \Rightarrow$$
$$c_i = \frac{1}{(x_i - x_0)\cdots(x_i - x_{i-1})(x_i - x_{i+1})\cdots(x_i - x_n)}$$

Polinoamele p_i au forma:

$$p_i(x) = \frac{(x - x_0)\cdots(x - x_{i-1})(x - x_{i+1})\cdots(x - x_n)}{(x_i - x_0)\cdots(x_i - x_{i-1})(x_i - x_{i+1})\cdots(x_i - x_n)} = \prod_{\substack{j=0 \\ j\neq i}}^{n} \left(\frac{x - x_j}{x_i - x_j}\right)$$

$$i = 0, \dots, n$$

Propoziţie

Polinoamele $p_0, p_1, \dots, p_n$ formează o bază în Π_n.

Demonstraţie: Vom arăta că cele $n+1$ polinoame sunt liniar independente:

$$q(x) = a_0 p_0(x) + a_1 p_1(x) + \cdots + a_n p_n(x) = 0\,,\ \forall x$$

$$\Rightarrow a_0 = a_1 = \cdots = a_n = 0$$

Vom face pe rând ***x* = *x*₀, *x* = *x*₁,…, *x* = *x*ₙ** în polinomul ***q***:

$$x = x_0 \quad q(x_0) = a_0 p_0(x_0) + a_1 p_1(x_0) + \cdots + a_n p_n(x_0) =$$

$$= a_0 1 + a_1 0 + \cdots + a_n 0 = a_0 = 0 \quad \Rightarrow a_0 = 0$$

$$x = x_1 \quad q(x_1) = 0 \quad \Rightarrow \quad a_1 = 0$$

$$\vdots$$

$$x = x_k \quad q(x_k) = a_0 p_0(x_k) + \cdots + a_k p_k(x_k) + \cdots + a_n p_n(x_k) =$$

$$= a_0 0 + \cdots + a_k 1 + \cdots + a_n 0 = a_k = 0 \Rightarrow a_k = 0$$

$$\vdots$$

$$x = x_n \quad q(x_n) = 0 \quad \Rightarrow \quad a_n = 0$$

Toate constantele $\boldsymbol{a_i}$ sunt nule deci polinoamele $\{p_i; i=0,\dots,n\}$ formează o bază în Π_n.

Pentru a aproxima funcţia $\boldsymbol{f}$ pornind de la tabelul de mai sus, vom construi un polinom $l_n \in \Pi_n$ a.î. să satisfacă ***condiţiile de interpolare***:

$$l_n \in \Pi_n \quad , \quad l_n(x_i) = y_i \quad , \quad \forall i = 0,\dots,n \tag{1}$$

Odată construit acest polinom, vom aproxima $\boldsymbol{f(x)}$ prin $\boldsymbol{l_n(x)}$, $f(x) \approx l_n(x)$

Vom scrie polinomul $\boldsymbol{l_n}$ în raport cu noua bază $\{p_i; i=0,\dots,n\}$, deci există constantele reale $\boldsymbol{a_0, a_1,\dots,a_n}$ astfel ca:

$$l_n(x) = \sum_{i=0}^{n} a_i p_i(x)$$

Constantele a_k se determină astfel:

$$y_k = l_n(x_k) = a_0 p_0(x_k) + \cdots + a_k p_k(x_k) + \cdots + a_n p_n(x_k) =$$

$$= a_0 0 + \cdots + a_k 1 + \cdots + a_n 0 = a_k \quad \Rightarrow \; a_k = y_k$$

Prin urmare un polinom de grad n care îndeplinesc condiţiile de interpolare (1) este:

$$l_n(x) = \sum_{i=0}^{n} y_i p_i(x) = \sum_{i=0}^{n} y_i \frac{(x - x_0)\cdots(x - x_{i-1})(x - x_{i+1})\cdots(x - x_n)}{(x_i - x_0)\cdots(x_i - x_{i-1})(x_i - x_{i+1})\cdots(x_i - x_n)}$$

$$= \sum_{i=0}^{n} y_i \prod_{\substack{j=0 \\ j \neq i}}^{n} \left(\frac{x - x_j}{x_i - x_j}\right) \qquad \textbf{(2)}$$

Polinomul din formula (2) se numeşte ***polinomul de interpolare Lagrange***.

Propoziţie

Polinomul $\boldsymbol{l_n}$ dat de formula (2) este unicul polinom de grad $\boldsymbol{n}$ care îndeplineşte condiţiile de interpolare (1).

Demonstraţie: Presupunem că mai există un polinom $\boldsymbol{q \in \Pi_n}$ care îndeplineşte condiţiile (1):

$$\boldsymbol{q \in \Pi_n \quad , \; q(x_i) = y_i \quad , \quad \forall i = 0, \ldots, n}$$

Fie polinomul $\boldsymbol{p(x)=l_n(x)-q(x) \in \Pi_n}$.

$$p(x_k) = l_n(x_k) - q(x_k) = y_k - y_k = 0 \ , \forall k = 0, \dots, n$$

Polinomul $\boldsymbol{p}$ are ca rădăcini toate nodurile de interpolare. Polinomul $\boldsymbol{p}$ este polinom de grad cel mult $\boldsymbol{n}$ şi are $\boldsymbol{(n+1)}$ rădăcini distincte ($\boldsymbol{x_i \neq x_j}, \forall \ \boldsymbol{i \neq j}$). Acest polinom nu poate fi decât polinomul identic nul:

$$p(x) = l_n(x) - q(x) \equiv 0 \quad \forall x \,, \quad l_n(x) = q(x) \quad \forall x$$

Polinomul $\boldsymbol{l_n}$ este unicul care satisface (2).

Fie w_{n+1} polinomul de grand $(n+1)$ care are ca rădăcini nodurile de interpolare:

$$w_{n+1}(x) = (x - x_0)(x - x_1)\cdots(x - x_n) \in \Pi_{n+1}$$

Fie $a = \min\{x_0, x_1, \ldots, x_n\}$, $b = \max\{x_0, x_1, \ldots, x_n\}$.

Teorema restului (eroarea la interpolarea Lagrange_

Fie $f \in C^{n+1}[a,b]$ şi $\bar{x} \in [a,b],\ \bar{x} \neq x_i\ , \forall i = 0,\ldots,n.$
Atunci există un punct $y \in [a,b],\ y = y(x_0, x_1, \ldots, x_n, \bar{x})$ (punctul y depinde de nodurile de interpolare x_i şi de punctul $\bar{x}$) astfel că eroarea la interpolarea numerică este dată de:

$$f(\bar{x}) - l_n(\bar{x}) = \frac{f^{(n+1)}(y)}{(n+1)!} w_{n+1}(\bar{x}) \qquad (3)$$

Demonstraţie: Considerăm funcţia F:

$$F(x) := f(x) - l_n(x) - c w_{n+1}(x)$$

Constanta reală c este aleasă astfel ca $F(\bar{x}) = 0$ adică:

$$c = \frac{f(\overline{x}) - l_n(\overline{x})}{w_{n+1}(\overline{x})}, \; (\, x \neq x_i \;\; \forall i \,) \Rightarrow \; w_{n+1}(\overline{x}) \neq 0 \,) \qquad (4)$$

Funcţia f fiind de clasă C^{n+1} pe intervalul $[a,b]$ rezultă că şi funcţia F este din $C^{n+1}[a,b]$. Avem:

$$F(x_i) = f(x_i) - l_n(x_i) - cw_{n+1}(x_i) = y_i - y_i - c\,0 = 0 \;\; , \forall i = 0, \ldots, n$$

Funcţia F are *(n+2)* zerouri, $x_0, x_1, \ldots, x_n, \overline{x}$. Aplicând succesiv Teorema lui Rolle rezultă că F' are *(n+1)* zerouri, F'' are n zerouri,…, $F^{(n+1)}$ are *1* zero în intervalul $[a,b]$. Vom nota această rădăcină a lui $F^{(n+1)}$ cu y. Punctul y depinde de zerourile iniţiale $x_0, x_1, \ldots, x_n, \overline{x}$ şi:

$$y = y(x_0, x_1, \ldots, x_n, \overline{x}) \in [a,b] \quad \text{a.î.} \quad F^{(n+1)}(y) = 0. \qquad (5)$$

Derivata de ordinul ***(n+1)*** a funcţiei ***F*** se calculează astfel:

$$F^{(n+1)}(x) = f^{(n+1)}(x) - l_n^{(n+1)}(x) - c\, w_{n+1}^{(n+1)}(x) =$$

$$= f^{(n+1)}(x) - 0 - c(n+1)! = f^{(n+1)}(x) - c(n+1)! \qquad (6)$$

(derivata de ordin ***(n+1)*** a polinomului de grad ***n*** l_n este ***0***).
Din relaţiile (4), (5) şi (6) rezultă că:

$$c = \frac{f^{(n+1)}(y)}{(n+1)!} = \frac{f(\overline{x}) - l_n(\overline{x})}{w_{n+1}(\overline{x})} \Rightarrow f(\overline{x}) - l_n(\overline{x}) = \frac{f^{(n+1)}(y)}{(n+1)!} w_{n+1}(\overline{x})$$

Forma Newton a polinomului de interpolare Lagrange

Fie $l_k(x, x_0, x_1,\ldots, x_k, f)$ polinomul de interpolare Lagrange pentru funcţia f pe sistemul de noduri distincte $\{x_0, x_1,\ldots, x_k\}$.

Propoziţie

Fie $l_{k-1}(x, x_0, x_1,\ldots, x_{k-1}, f)$, $l_{k-1}(x, x_1, x_2,\ldots, x_k, f) \in \Pi_{k-1}$ polinoamele de interpolare Lagrange pentru funcţia f pe sistemele de noduri $\{x_0, x_1,\ldots, x_{k-1}\}$ şi respectiv $\{x_1, x_2,\ldots, x_k\}$. Atunci:

$$l_k(x, x_0, x_1, \ldots, x_k, f) =$$

$$= \frac{(x - x_0) l_{k-1}(x, x_1, x_2, \ldots, x_k, f) - (x - x_k) l_{k-1}(x, x_0, x_1, \ldots, x_{k-1}, f)}{x_k - x_0} \tag{1}$$

Demonstraţie: Exerciţiu.

Considerăm următoarele probleme de interpolare pentru f:

$$\{(x_0, y_0),(x_1, y_1),\ldots,(x_{k-1}, y_{k-1})\} \;\to\; l_{k-1}(x, x_0, x_1,\ldots, x_{k-1}, f)$$

$$\{(x_0, y_0),(x_1, y_1),\ldots,(x_k, y_k)\} \;\to\; l_k(x, x_0, x_1,\ldots, x_k, f)$$

Ne interesează să găsim o formulă de trecere rapidă de la polinomul de interpolare pe k noduri la cel care are un nod în plus. Deoarece polinomul de grad cel mult k:

$$q(x) = l_k(x, x_0, x_1,\ldots, x_k, f) - l_{k-1}(x, x_0, x_1,\ldots, x_{k-1}, f) \in \Pi_k$$

are ca rădăcini punctele $x_0, x_1, \ldots, x_{k-1}$ ($q(x_i) = y_i - y_i = 0$, $i=0,...,k-1$) avem relaţia:

$$l_k(x, x_0, x_1, \ldots, x_k, f) = l_{k-1}(x, x_0, x_1, \ldots, x_{k-1}, f) + A\prod_{j=0}^{k-1}(x - x_j) \quad (2)$$

în care ***A*** este dat de relaţia:

$$A = \frac{l_k(x_k, x_0, x_1, \ldots, x_k, f) - l_{k-1}(x_k, x_0, x_1, \ldots, x_{k-1}, f)}{\prod_{j=0}^{k-1}(x_k - x_j)} \quad (3)$$

$$A = \frac{y_k}{\prod\limits_{j=0}^{k-1}(x_k - x_j)} - \frac{\sum\limits_{i=0}^{k-1} y_i \prod\limits_{\substack{j=0 \\ j\neq i}}^{k-1}\left(\frac{x_k - x_j}{x_i - x_j}\right)}{\prod\limits_{j=0}^{k-1}(x_k - x_j)} =$$

$$= \frac{y_k}{\prod\limits_{j=0}^{k-1}(x_k - x_j)} - \sum_{i=0}^{k-1} \frac{y_i}{(x_k - x_i)\prod\limits_{\substack{j=0 \\ j\neq i}}^{k-1}(x_i - x_j)}$$

$$A=\sum_{i=0}^{k}\frac{y_i}{\prod_{\substack{j=0\\ j\neq i}}^{k}(x_i-x_j)} \qquad (4)$$

Considerăm următoarele problemele de interpolare pentru f:

$$\{(x_1,y_1),(x_2,y_2),\ldots,(x_k,y_k)\} \to l_{k-1}(x,x_1,x_2,\ldots,x_k,f)$$

$$\{(x_0,y_0),(x_1,y_1),\ldots,(x_k,y_k)\} \to l_k(x,x_0,x_1,\ldots,x_k,f)$$

Vom avea, analog ca mai sus:

$$l_k(x, x_0, x_1, \ldots, x_k, f) = l_{k-1}(x, x_1, x_2, \ldots, x_k, f) + B\prod_{j=1}^{k}(x - x_j) \tag{5}$$

Dacă înmulţim relaţia (2) cu ***(x-x_k)*** iar relaţia (5) cu ***(x-x₀)*** şi scădem aceste relaţii obţinem:

$$(x_0 - x_k)l_k(x, x_0, x_1, \ldots, x_k, f) = (x - x_k)l_{k-1}(x, x_0, x_1, \ldots, x_{k-1}, f) -$$
$$- (x - x_0)l_{k-1}(x, x_1, x_2, \ldots, x_k, f) + (A - B)\prod_{j=0}^{k}(x - x_j)$$

Ţinând seama de relaţia (1) rezultă că:

$$(A-B)\prod_{j=0}^{k}(x-x_j)=0 \text{ adică } A=B$$

Vom nota în cele ce urmează:

$$A=\left[x_0,x_1,\ldots,x_k\right]_f$$

numită ***diferenţă divizată de ordin k a funcţiei f pe nodurile*** $\left\{x_0,x_1,\ldots,x_k\right\}$.

Vom înlocui în formula (2) $\boldsymbol{l_{k-1}(x, x_0,\ldots, x_{k-1}, f)}$ cu:

$$l_{k-1}(x,x_0,\ldots,x_{k-1},f)=l_{k-2}(x,x_1,\ldots,x_{k-1},f)+$$

$$+\left[x_0,x_1,\ldots,x_{k-1}\right]_f\prod_{j=1}^{k-1}(x-x_j)$$

iar în formula (5) $\boldsymbol{l_{k-1}(x, x_1,\ldots, x_k, f)}$ cu:

$$l_{k-1}(x, x_1, \ldots, x_k, f) = l_{k-2}(x, x_1, \ldots, x_{k-1}, f) + [x_1, x_2, \ldots, x_k]_f \prod_{l=1}^{k-1}(x - x_l)$$

şi apoi scădem membru cu memebru cele două relaţii. Obţinem:

$$[x_0, \ldots, x_{k-1}]_f \prod_{j=1}^{k-1}(x - x_j) + [x_0, \ldots, x_k]_f \prod_{j=0}^{k-1}(x - x_j) -$$

$$-[x_1, \ldots, x_k]_f \prod_{l=1}^{k-1}(x - x_l) - [x_0, \ldots, x_k]_f \prod_{l=1}^{k}(x - x_l) = 0$$

Putem scrie:

$$\prod_{j=1}^{k-1}(x-x_j)\ \left\{\left[x_0,\dots,x_{k-1}\right]_f-\left[x_1,\dots,x_k\right]_f\right\}+$$

$$+\left[x_0,\dots,x_k\right]_f\prod_{j=0}^{k-1}(x-x_j)\left[x-x_0-x+x_n\right]=0$$

relaţie din care obţinem:

$$\left[x_0,x_1,\dots,x_k\right]_f=\frac{\left[x_1,x_2,\dots,x_k\right]_f-\left[x_0,x_1,\dots,x_{k-1}\right]_f}{x_k-x_0} \qquad (6)$$

Relaţia (6) justifică denumirea de diferenţă divizată.

Se introduce şi noţiunea de diferenţă divizată de ordinul 0:

$$\left[x_k \right]_f = y_k = f(x_k) \; , \tag{7}$$

Diferenţele divizate se pot obţine folosind definiţia directă (4) sau folosind definiţia recursivă (7), (6). Cele 2 definiţii sunt echivalente:

Propoziție

$$\left[x_0, x_1, \ldots, x_k \right]_f = \sum_{i=0}^{k} \frac{y_i}{\prod\limits_{\substack{j=0 \\ j \neq i}}^{k} (x_i - x_j)} = \sum_{i=0}^{k} \frac{y_i}{\left(w_{n+1}(x_k) \right)'} \tag{8}$$

pentru orice sistem de noduri $\left\{ x_0, x_1, \ldots, x_k \right\}$ şi orice $\boldsymbol{k}$.

Demonstraţie: Se face prin inducţie. Pentru ***k=1*** avem:

$$\left[x_0, x_1\right]_f = \frac{y_0}{x_0 - x_1} + \frac{y_1}{x_1 - x_0} = \frac{\left[x_1\right]_f - \left[x_0\right]_f}{x_1 - x_0}$$

Presupunem că relaţia (8) este valabilă pentru orice ***k*** şi pentru orice sistem de noduri $\left\{x_0, x_1, \ldots, x_k\right\}$. Pentru ***k+1*** folosim relaţia de recurenţă şi apoi aplicăm ipoteza inductivă:

$$\left[x_0,x_1,\ldots,x_{k+1}\right]_f=\frac{\left[x_1,x_1,\ldots,x_{k+1}\right]_f-\left[x_0,x_2,\ldots,x_k\right]_f}{x_{k+1}-x_0}=$$

$$=\frac{1}{x_{k+1}-x_0}\left(\sum_{i=1}^{k+1}\frac{y_i}{\prod\limits_{\substack{j=1\\j\neq i}}^{k+1}(x_i-x_j)}-\sum_{i=0}^{k}\frac{y_i}{\prod\limits_{\substack{j=0\\j\neq i}}^{k}(x_i-x_j)}\right)=$$

$$=\frac{1}{x_{k+1}-x_0}\left\{-\frac{y_0}{\prod\limits_{\substack{j=0\\j\neq 0}}^{k}(x_0-x_j)}+\frac{y_{k+1}}{\prod\limits_{\substack{j=1\\j\neq k+1}}^{k+1}(x_{k+1}-x_j)}+\right.$$

$$\left.+\sum_{i=1}^{k}\left[\frac{y_i}{\prod\limits_{\substack{j=1\\j\neq i}}^{k}(x_i-x_j)}\left(\frac{1}{x_i-x_{k+1}}-\frac{1}{x_i-x_0}\right)\right]\right\}=$$

$$= \frac{y_0}{\prod_{\substack{j=0 \\ j \neq 0}}^{k+1} (x_0 - x_j)} + \frac{y_{k+1}}{\prod_{\substack{j=0 \\ j \neq k+1}}^{k+1} (x_{k+1} - x_j)} + \sum_{i=1}^{k} \frac{y_i}{\prod_{\substack{j=0 \\ j \neq i}}^{k+1} (x_i - x_j)} =$$

$$= \sum_{i=0}^{k+1} \frac{y_i}{\prod_{\substack{j=0 \\ j \neq i}}^{k+1} (x_i - x_j)}$$

Inducţia este completă.

Din definiţie se observă că diferenţa divizată $\left[x_0, x_1, \ldots, x_k\right]_f$ nu depinde de ordinea nodurilor $\left\{x_0, x_1, \ldots, x_k\right\}$.

Vom nota în continuare cu $l_k(x)$ polinomul de interpolare Lagrange pe nodurile $\left\{x_0, x_1, \ldots, x_k\right\}$ pentru funcţia f. Avem:

$$
\begin{aligned}
l_n(x) &= l_0(x) + [l_1(x) - l_0(x)] + \cdots + [l_k(x) - l_{k-1}(x)] + \cdots + [l_n(x) - l_{n-1}(x)] = \\
&= y_0 + \left[x_0, x_1\right]_f (x - x_0) + \cdots + \left[x_0, x_1, \ldots, x_k\right]_f (x - x_0)\cdots(x - x_{k-1}) + \cdots \\
&\quad + \left[x_0, x_1, \ldots, x_n\right]_f (x - x_0)\cdots(x - x_{n-1})
\end{aligned}
$$

Am obţinut ***forma Newton*** a polinomului de interpolare Lagrange:

$$l_n(x) = y_0 + [x_0, x_1]_f (x - x_0) + [x_0, x_1, x_2]_f (x - x_0)(x - x_1) + \cdots +$$
$$+ [x_0, x_1, \ldots, x_n]_f (x - x_0) \cdots (x - x_{n-1})$$

Schema lui Aitken de calcul a diferențelor divizate

Ne propunem să calculăm diferențele divizate

$$[x_0, x_1]_f, [x_0, x_1, x_2]_f, \ldots, [x_0, x_1, \ldots, x_n]_f$$

necesare construirii polinomului de interpolare Lagrange în forma Newton. Procedeul folosește definiția recursivă a diferențelor divizate și se desfășoară în n pași. La pasul 1 se calculează numai diferențe divizate de ordinul 1:

$$\left[x_0, x_1\right]_f, \left[x_1, x_2\right]_f, \ldots, \left[x_{n-1}, x_n\right]_f.$$

În general, la pasul $\boldsymbol{k}$ se calculează diferențe divizate de ordin $\boldsymbol{k}$:

$$\left[x_0, x_1, \ldots, x_k\right]_f, \left[x_1, x_2, \ldots, x_{k+1}\right]_f, \ldots, \left[x_{n-k}, x_{n-k+1}, \ldots, x_n\right]_f.$$

La pasul $\boldsymbol{n}$ se calculează o singură diferență divizată de ordin $\boldsymbol{n}$ și anume $\left[x_0, x_1, \ldots, x_n\right]_f$.

$$
\begin{array}{lllll}
 & & \textbf{Pas } 1 \quad \cdots & \textbf{Pas } k & \cdots \quad \textbf{Pas } n \\
x_0 & y_0 & & & \\
x_1 & y_1 & [x_0, x_1]_f & & \\
x_2 & y_2 & [x_1, x_2]_f & & \\
\vdots & & & & \\
x_k & y_k & [x_{k-1}, x_k]_f & [x_0, x_1, \ldots, x_k]_f & \\
\vdots & & & \vdots & \ddots \\
x_{n-1} & y_{n-1} & [x_{n-2}, x_{n-1}]_f & [x_{n-k-1}, \ldots, x_{n-1}]_f & \\
x_n & y_n & [x_{n-1}, x_n]_f & [x_{n-k}, \ldots, x_{n-1}]_f & \cdots \quad [x_0, x_1 \ldots, x_n]_f
\end{array}
$$

Notăm **dd[i,k]=** $\left[x_i, x_{i+1}, \ldots, x_{i+k}\right]_f$ diferenţa divizată de ordin ***k***, pe nodurile consecutive $\left\{x_i, x_{i+1}, \ldots, x_{i+k}\right\}$ ***i=0,…,n-k, k=1,…,n,*** cu ***dd[i,0]=y_i, i=0,…,n***. Schema lui Aitken se implementează astfel:

$$\boldsymbol{dd[i,0] = y_i,\ i = 0,\ldots,n;}$$

$$\textbf{for } \boldsymbol{k = 1,\ldots,n}$$

$$\quad \textbf{for } \boldsymbol{i = 0,\ldots,n-k}$$

$$\boldsymbol{dd[i,k] = \frac{dd[i+1,k-1] - dd[i,k-1]}{x_{i+k} - x_i}}$$

Putem face aceleași calcule folosind un singur vector, de exemplu rescriind vectorul $\boldsymbol{y}$ astfel:

$$\mathbf{for}\ \boldsymbol{k = 1,\ldots,n}$$
$$\quad \mathbf{for}\ \boldsymbol{i = n,\ldots,k}$$
$$\qquad \boldsymbol{y_i = \frac{y_i - y_{i-1}}{x_i - x_{i-k}}}$$

La finalul acestei secvențe de program, vectorul $\boldsymbol{y}$ va conține elementele:

$$\boldsymbol{y_0, [x_0, x_1]_f, [x_0, x_1, x_2]_f, \ldots, [x_0, x_1, \ldots, x_n]_f}$$
$$(\boldsymbol{y_k = [x_0, x_1, \ldots, x_k]_f,\ k=0,\ldots,n}).$$