Aproximarea funcţiilor prin metoda celor mai mici pătrate

Prof.dr.ing. Gabriela Ciuprina

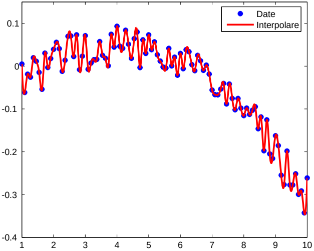
Universitatea "Politehnica" Bucureşti, Facultatea de Inginerie Electrică

Suport didactic pentru disciplina *Metode numerice*, 2017-2018

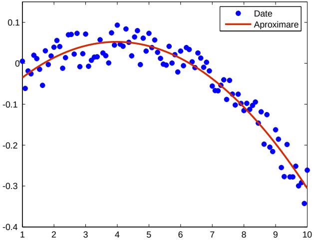
Cuprins

1. Introducere
   - Formularea problemei
   - Ideea metodei CMMP
2. CMMP prin rezolvarea sistemului normal
   - Funcţii de bază
   - Algoritm
   - Alte aspecte ale CMMP

Interpolare vs. aproximare



Interpolare



Aproximare

Se dă un set de date (x_k, y_k), $k = 0, \ldots, m$, unde $y_k = f(x_k)$.
Se caută o aproximare notată $g(x)$, unde $g(x_k) \approx y_k$.

Ideea metodei

Funcţia de aproximare se caută a.î:

$$\|g(x) - f(x)\| \quad \text{minimă}$$

Deoarece f este cunoscută într-un număr discret de puncte $\Rightarrow$ norma discretă (de exemplu norma Euclidiană discretă).
Metoda celor mai mici pătrate (CMMP) caută g a.î:

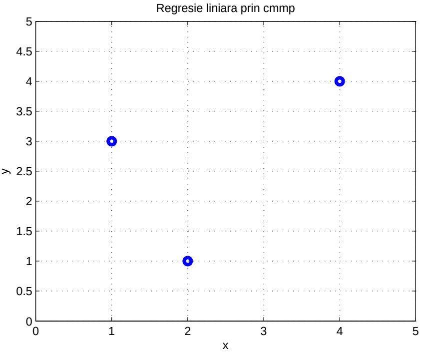
$$E = \sum_{k=0}^{m} (f(x_k) - g(x_k))^2 \quad \text{minimă} \tag{1}$$

Regresia liniară

$$g(x) = c_0 + c_1 x \quad c_0 = ?, \; c_1 = ?.$$

Exemplu: m = 3: (1,3), (2,1), (4,4).

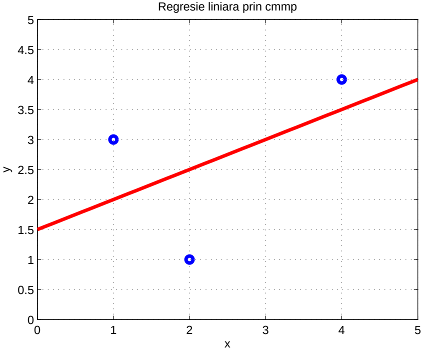
$$\begin{aligned} E(c_0, c_1) &= (y_0 - c_0 - c_1 x_0)^2 + \\ &+ (y_1 - c_0 - c_1 x_1)^2 + \\ &+ (y_2 - c_0 - c_1 x_2)^2 = \\ &= (3 - c_0 - c_1)^2 + \\ &+ (1 - c_0 - 2c_1)^2 + \\ &+ (4 - c_0 - 4c_1)^2. \end{aligned}$$



Regresia liniară

$$\begin{aligned} E(c_0, c_1) &= (y_0 - c_0 - c_1 x_0)^2 + (y_1 - c_0 - c_1 x_1)^2 + (y_2 - c_0 - c_1 x_2)^2 = \\ &= (3 - c_0 - c_1)^2 + (1 - c_0 - 2c_1)^2 + (4 - c_0 - 4c_1)^2. \end{aligned} \tag{2}$$

$$\left\{ \begin{array}{lcl} \frac{\partial E}{\partial c_0} & = & 0, \\ \frac{\partial E}{\partial c_1} & = & 0. \end{array} \right. \Rightarrow \left\{ \begin{array}{lcl} 3c_0 + 7c_1 & = & 8, \\ 7c_0 + 21c_1 & = & 21 \end{array} \right. \Rightarrow c_0 = 1.5, \quad c_1 = 0.5. \tag{3}$$



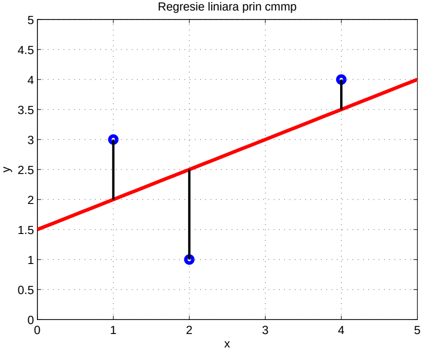
$g(1) = 2$, $g(2) = 2.5$, $g(3) = 3.5$;

$\min E = (3 - 2)^2 + (1 - 2.5)^2 + (4 - 3.5)^2$

Regresia liniară

$$\begin{aligned} E(c_0, c_1) &= (y_0 - c_0 - c_1 x_0)^2 + (y_1 - c_0 - c_1 x_1)^2 + (y_2 - c_0 - c_1 x_2)^2 = \\ &= (3 - c_0 - c_1)^2 + (1 - c_0 - 2c_1)^2 + (4 - c_0 - 4c_1)^2. \end{aligned} \tag{2}$$

$$\begin{cases} \frac{\partial E}{\partial c_0} = 0, \\ \frac{\partial E}{\partial c_1} = 0. \end{cases} \Rightarrow \begin{cases} 3c_0 + 7c_1 = 8, \\ 7c_0 + 21c_1 = 21 \end{cases} \Rightarrow c_0 = 1.5, \quad c_1 = 0.5. \tag{3}$$



$g(1) = 2$, $g(2) = 2.5$, $g(3) = 3.5$;

$\min E = (3 - 2)^2 + (1 - 2.5)^2 + (4 - 3.5)^2$

Regresia liniară

Set oarecare de date:

$$E(c_0, c_1) = \sum_{k=0}^{m} (y_k - c_0 - c_1 x_k)^2, \tag{4}$$

$$\begin{cases} \frac{\partial E}{\partial c_0} & = & 0, \\ \frac{\partial E}{\partial c_1} & = & 0. \end{cases} \quad \Rightarrow \quad \begin{cases} \sum_{k=0}^{m} -2(y_k - c_0 - c_1 x_k) & = & 0, \\ \sum_{k=0}^{m} -2x_k(y_k - c_0 - c_1 x_k) & = & 0. \end{cases} \tag{5}$$

$\Rightarrow$

$$\begin{cases} c_0(m+1) + c_1 \sum_{k=0}^{m} x_k & = & \sum_{k=0}^{m} y_k, \\ c_0 \sum_{k=0}^{m} x_k + c_1 \sum_{k=0}^{m} x_k^2 & = & \sum_{k=0}^{m} x_k y_k. \end{cases} \tag{6}$$

Sistem normal

Regresia liniară

Soluţia se poate calcula explicit:

$$c_0 = \frac{d_0}{d}, \quad c_1 = \frac{d_1}{d}, \tag{7}$$

unde

$$d = (m+1)\sum_{k=0}^{m} x_k^2 - \left(\sum_{k=0}^{m} x_k\right)^2, \tag{8}$$

$$d_0 = \left(\sum_{k=0}^{m} x_k^2\right)\left(\sum_{k=0}^{m} y_k\right) - \left(\sum_{k=0}^{m} x_k\right)\left(\sum_{k=0}^{m} x_k y_k\right), \tag{9}$$

$$d_1 = (m+1)\sum_{k=0}^{m} x_k y_k - \left(\sum_{k=0}^{m} x_k\right)\left(\sum_{k=0}^{m} y_k\right). \tag{10}$$

Regresia liniară - alt raţionament

Condiţiile de interpolare pentru datele din tabel şi $g(x) = c_0 + c_1 x$:

$$c_0 + c_1 x_k = y_k, \quad k = 1, \ldots, m, \tag{11}$$

$\Rightarrow$ sistem supradeterminat de ecuaţii. Dacă notăm

$$\mathbf{A} = \begin{bmatrix} 1 & x_0 \\ 1 & x_1 \\ \vdots & \vdots \\ 1 & x_m \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{bmatrix}, \tag{12}$$

atunci (11) se scriu compact

$$\mathbf{Ac} = \mathbf{y}, \tag{13}$$

unde $\mathbf{c} = [c_0 \ c_1]^T$, iar sistemul normal este

$$\mathbf{A}^T\mathbf{Ac} = \mathbf{A}^T\mathbf{y}. \tag{14}$$

Regresia liniară - alt raţionament

Reluăm exemplul simplu: $m = 3$: (1,3), (2,1), (4,4).

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 4 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 1 & 3 \\ 1 & 1 \\ 1 & 4 \end{bmatrix}, \tag{15}$$

$$\mathbf{A}^T\mathbf{A} = \begin{bmatrix} 3 & 7 \\ 7 & 21 \end{bmatrix}, \quad \mathbf{A}^T\mathbf{y} = \begin{bmatrix} 8 \\ 21 \end{bmatrix}. \tag{16}$$

$$\left\{ \begin{array}{rcl} 3c_0 + 7c_1 & = & 8, \\ 7c_0 + 21c_1 & = & 21 \end{array} \right. \Rightarrow \quad c_0 = 1.5, \quad c_1 = 0.5. \tag{17}$$

Cazul general

Metoda celor mai mici pătrate nu este limitată la polinoame de gradul întâi.

Exemplu:

$g(x) = c_0 \ln(x) + c_1 \sin(x) + c_2 \sqrt{x}$.

1. Se minimizează funcţia definită de $E = \sum_{k=0}^{m} (f(x_k) - g(x_k))^2$;
2. Se impun condiţiile de minimizare;
3. Se rezolvă un sistem algebric liniar de dimensiune 3.

Cazul general

În general, se caută o aproximare de forma unui polinom generalizat

$$g(x) = \sum_{j=0}^{n} c_j \varphi_j(x), \tag{18}$$

unde $\varphi_i(x)$ se numesc *funcţii de bază*.
OBS: $n + 1 < m + 1$

$$E(c_0, c_1, \dots, c_n) = \sum_{k=0}^{m} (g(x_k) - f(x_k))^2 = \sum_{k=0}^{m} \left(\sum_{j=0}^{n} c_j \varphi_j(x_k) - y_k \right)^2. \tag{19}$$

Cazul general

$$E(c_0, c_1, \ldots, c_n) = \sum_{k=0}^{m} (g(x_k) - f(x_k))^2 = \sum_{k=0}^{m} \left(\sum_{j=0}^{n} c_j \varphi_j(x_k) - y_k \right)^2. \tag{20}$$

Punctul de minim al acestei funcţii este un punct critic:

$$\frac{\partial E}{\partial c_i} = 0, \quad i = 0, \ldots, n. \tag{21}$$

Înlocuind expresia (20) în (21) rezultă

$$\sum_{j=0}^{n} \left(\sum_{k=0}^{m} \varphi_i(x_k) \varphi_j(x_k) \right) c_j = \sum_{k=0}^{m} y_k \varphi_i(x_k), \quad i = 0, \ldots, n. \tag{22}$$

Sistem normal dificil de rezolvat dacă φ_k nu sunt alese potrivit.

Cazul general - alt raţionament

Condiţiile de interpolare pentru datele din tabel şi $g(x)$

$$g(x_k) = y_k, \quad k = 1, \ldots, m, \tag{23}$$

$\Rightarrow$ sistem supradeterminat de ecuaţii. Notăm

$$\mathbf{A} = \begin{bmatrix} \varphi_0(x_0) & \varphi_1(x_0) & \cdots & \varphi_n(x_0) \\ \varphi_0(x_1) & \varphi_1(x_1) & \cdots & \varphi_n(x_1) \\ \vdots & \vdots & & \vdots \\ \varphi_0(x_m) & \varphi_1(x_m) & \cdots & \varphi_n(x_m) \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{bmatrix}, \tag{24}$$

atunci (23) se scriu compact

$$\mathbf{Ac} = \mathbf{y}, \tag{25}$$

unde $\mathbf{c} = [c_0, \ c_1 \ \ldots \ c_{n+1}]^T$. Sistemul normal:

$$\mathbf{A}^T\mathbf{Ac} = \mathbf{A}^T\mathbf{y}. \tag{26}$$

Funcţii de bază clasice

Sistemul normal de rezolvat

$$\mathbf{Nc} = \mathbf{t} \tag{27}$$

are, în cazul *regresiei liniare* forma

$$\mathbf{N} = \begin{bmatrix} m+1 & \sum_{k=0}^{m} x_k \\ \sum_{k=0}^{m} x_k & \sum_{k=0}^{m} x_k^2 \end{bmatrix}, \quad \mathbf{t} = \begin{bmatrix} \sum_{k=0}^{m} y_k \\ \sum_{k=0}^{m} x_k y_k \end{bmatrix}. \tag{28}$$

Funcţii de bază clasice

Sistemul normal de rezolvat

$$\mathbf{N}\mathbf{c} = \mathbf{t} \tag{29}$$

are, în cazul *regresiei parabolice* $g(x) = c_0 + c_1 x + c_2 x^2$, forma:

$$\mathbf{N} = \begin{bmatrix} m+1 & \sum_{k=0}^{m} x_k & \sum_{k=0}^{m} x_k^2 \\ \sum_{k=0}^{m} x_k & \sum_{k=0}^{m} x_k^2 & \sum_{k=0}^{m} x_k^3 \\ \sum_{k=0}^{m} x_k^2 & \sum_{k=0}^{m} x_k^3 & \sum_{k=0}^{m} x_k^4 \end{bmatrix}, \quad \mathbf{t} = \begin{bmatrix} \sum_{k=0}^{m} y_k \\ \sum_{k=0}^{m} x_k y_k \\ \sum_{k=0}^{m} x_k^2 y_k \end{bmatrix}. \tag{30}$$

OBS:
Generalizarea este uşoara pentru $n > 2$, dar sistemul este slab condiţionat. :(

Funcţii de bază clasice

Sistemul normal de rezolvat

$$\mathbf{N}\mathbf{c} = \mathbf{t} \tag{29}$$

are, în cazul *regresiei parabolice* $g(x) = c_0 + c_1 x + c_2 x^2$, forma:

$$\mathbf{N} = \begin{bmatrix} m+1 & \sum_{k=0}^{m} x_k & \sum_{k=0}^{m} x_k^2 \\ \sum_{k=0}^{m} x_k & \sum_{k=0}^{m} x_k^2 & \sum_{k=0}^{m} x_k^3 \\ \sum_{k=0}^{m} x_k^2 & \sum_{k=0}^{m} x_k^3 & \sum_{k=0}^{m} x_k^4 \end{bmatrix}, \quad \mathbf{t} = \begin{bmatrix} \sum_{k=0}^{m} y_k \\ \sum_{k=0}^{m} x_k y_k \\ \sum_{k=0}^{m} x_k^2 y_k \end{bmatrix}. \tag{30}$$

OBS:
Generalizarea este uşoara pentru $n > 2$, dar sistemul este slab condiţionat. :(Remedii ?

Funcţii de bază clasice

Remedii:

1. Folosirea unor alte funcţii de bază;
2. O altă strategie pentru CMMP, care nu rezolvă sistemul normal.

Funcţii de bază - polinoame Chebyshev

În cazul în care se caută polinoamelor algebrice, atunci folosirea polinoamelor Chebyshev, definite recursiv ca:

$$
\begin{aligned}
T_0(x) &= 1, && (31)\\
T_1(x) &= x, && (32)\\
T_j(x) &= 2xT_{j-1}(x) - T_{j-2}(x), && (33)
\end{aligned}
$$

generează un sistem de ecuaţii mult mai bine condiţionat.
OBS: Există un algoritm mai eficient de evaluare a aproximării $g(x)$ cu polinoame Chebyshev, decât rezolvarea sistemului normal [Cheney07].

Funcţii de bază ortonormale

Ideal (din p.d.v. al condiţionării) - varianta în care matricea coeficienţilor este matricea unitate, adică dacă

$$\sum_{k=0}^{m} \varphi_i(x_k)\varphi_j(x_k) = \delta_{ij}, \qquad \text{unde} \quad \delta_{ij} = \begin{cases} 1 & \text{dacă } i = j \\ 0 & \text{dacă } i \neq j \end{cases}, \tag{34}$$

ceea ce înseamnă că **funcţiile de bază sunt ortonormale**.

$$c_j = \sum_{k=0}^{m} y_k \varphi_j(x_k), \quad j = 0, \ldots, n. \tag{35}$$

Funcţii de bază ortonormale

Dacă notăm cu $\mathcal{G}$ spaţiul vectorial al funcţiilor generat de funcţii de bază $\varphi_j(x)$

$$\mathcal{G} = \left\{ g : \exists \ c_0, c_1, \ldots c_n, \text{ astfel încât } g(x) = \sum_{j=0}^{n} c_j \varphi_j(x) \right\}, \quad (36)$$

atunci o bază ortonormală se poate construi de exemplu printr-o procedură Gram-Schmidt.

Algoritm - paşi principali

Date:

- tabelul de valori;
- valorile în care se doreşte evaluarea polinomului de aproximare.

Se aleg:

- funcţiile de bază (în consecinţă şi gradul polinomului de aproximare).

Se calculează:

- valorile polinomului de aproximare în punctele dorite.

Algoritm - paşi principali

CMMP_SistemNormal

1. Asamblează $\mathbf{A}$ şi $\mathbf{y}$
2. Calculează $\mathbf{N} = \mathbf{A}^T\mathbf{A}$ şi $\mathbf{t} = \mathbf{A}^T\mathbf{y}$
3. Rezolvă sistemul pătrat $\mathbf{Nc} = \mathbf{t} \Rightarrow \mathbf{c}$
4. Evaluează polinomul de aproximare $g(x)$ în punctul dorit

Algoritm - paşi principali

Obs:

- În cazuri particulare, se poate evita calculul explicit al coeficienţilor $\mathbf{c}$.
- Se poate ca tabelul de valori să includă numere complexe. În acest caz:
$$\mathbf{N} = \mathbf{A}^H\mathbf{A}, \qquad \mathbf{t} = \mathbf{A}^H\mathbf{y}$$
unde $\mathbf{A}^H$ (notată uneori şi cu $\mathbf{A}^*$) este *transpusa hermitiană* adica $(a^H)_{ij} = \overline{a_{ji}}$.
- Sistemul normal $\mathbf{A}^H\mathbf{A}\mathbf{c} = \mathbf{A}^H\mathbf{y}$ este nesingular dcaă şi numai dacă $\mathbf{A}$ este de rang complet [Trefethen97].

Probleme de CMMP neliniare

Până acum polinoamele de interpolare = combinaţii liniare de coeficienţi necunoscuţi.

Dar dacă:

Pentru setul de date (x_k, y_k), $k = 0, m$ se caută $g(x) = e^{cx}$?

Metoda CMMP minimizează funcţia

$$E(c) = \sum_{k=0}^{m} (e^{cx_k} - y_k)^2 \tag{37}$$

Minimul are loc pentru $E'(c) = 0$, deci

$$2 \sum_{k=0}^{m} (e^{cx_k} - y_k)\, e^{cx_k} x_k = 0, \tag{38}$$

Probleme de CMMP neliniare

Pentru setul de date (x_k, y_k), $k = 0, m$ se caută $g(x) = e^{cx}$. CMMP duce la

$$\sum_{k=0}^{m} \left(e^{cx_k} - y_k\right) e^{cx_k} x_k = 0, \tag{39}$$

ecuaţie neliniară în $c \Rightarrow$

- rezolvare de ecuaţii neliniare;
- metode de minimizare.

Probleme de CMMP neliniare

Pentru setul de date (x_k, y_k), $k = \overline{0, m}$ se caută $g(x) = e^{cx}$.
Acestea sunt **probleme de CMMP neliniare**.
OBS:
Reformulare ca o problemă de CMMP liniară:
Pentru setul de date $(x_k, \ln(y_k))$, $k = \overline{0, m}$ se caută $h(x) = cx$.
Valoarea c din problema reformulată nu este soluţia ecuaţiei neliniare (39), dar aproximarea $e^{h(x)}$ s-ar putea să fie satisfăcătoare.

Concluzii

- Se caută aproximarea unor date dintr-un tabel de valori;
- CMMP găseşte o funcţie a.î. distanţa dintre ea şi tabelul de date să fie minimă (în normă Euclidiană);
- Aceasta problemă este echivalentă cu rezolvarea unui sistem de ecuaţii supradimensionat $\mathbf{Ac} = \mathbf{y}$;
- Acesta se poate aduce la un sistem de ecuaţii normal, cu matricea coeficienţilor pătrată;
- Condiţionarea sistemului normal depinde de alegerea funcţiilor de bază;

Concluzii

- Polinoamele algebrice clasice se pot folosi pentru regresii liniare sau parabolice;
- Pentru ordine mai mari e mai bine să se folosească polinoame Chebyshev;
- O procedură eficientă poate folosi polinoame ortogonale adaptate setului de date;
- Găsirea ordinului optim se poate facând o analiză statistică.

Concluzii

CMMP poate fi privită ca o metodă de rezolvare a unui sistem supradeterminat, care a fost notat

$$\mathbf{Ac} = \mathbf{y}.$$

Minimizarea normei Euclidiene discrete este echivalentă cu minimizarea normei reziduului

$$\mathbf{r} = \mathbf{Ac} - \mathbf{y}.$$

Există şi alţi algoritmi CMMP, care nu asamblează sistemul normal (bazaţi pe factorizare QR sau SVD).

Lectură obligatorie

[Ioan98] D. Ioan et al., *Metode numerice in ingineria electrica*, Ed. Matrix Rom, Bucuresti, 1998. (Capitolul 13)

Cartea se găseşte la biblioteca UPB, puteţi verifica accesând catalogul http://www.library.pub.ro/.