

Aproximarea funcțiilor prin metoda celor mai mici pătrate

Prof.dr.ing. Gabriela Ciuprina

Universitatea "Politehnica" București, Facultatea de Inginerie Electrică

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

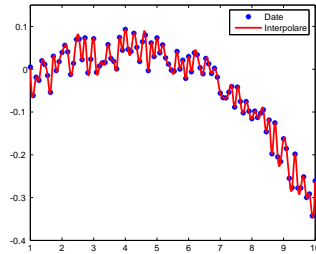
Notes

Cuprins

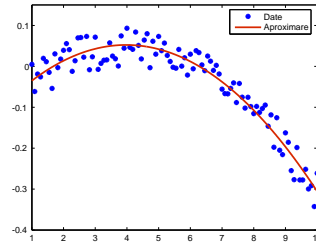
- 1 Introducere
 - Formularea problemei
 - Ideea metodei CMMP
- 2 CMMP prin rezolvarea sistemului normal
 - Funcții de bază
 - Algoritm
 - Alte aspecte ale CMMP

Notes

Interpolare vs. aproximare



Interpolare



Aproximare

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

Ideea metodei

Funcția de aproximare se caută a.î:

$$\|g(x) - f(x)\| \text{ minimă}$$

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:

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

Notes

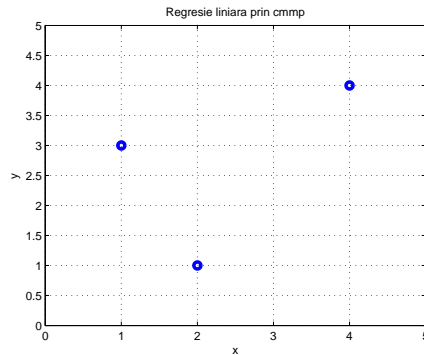
Notes

Regresia liniară

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

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

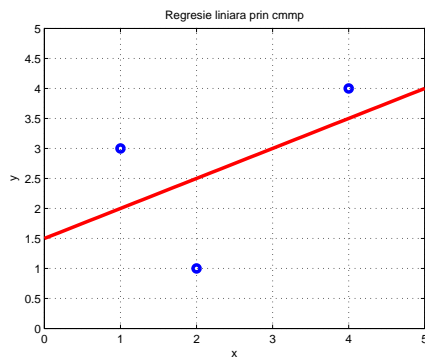
$$\begin{aligned} E(c_0, c_1) &= (y_0 - c_0 - c_1 x_0)^2 + \\ &+ (y_1 - c_0 - c_1 x_1)^2 + \\ &+ (y_2 - c_0 - c_1 x_2)^2 = \\ &= (3 - c_0 - c_1)^2 + \\ &+ (1 - c_0 - 2c_1)^2 + \\ &+ (4 - c_0 - 4c_1)^2. \end{aligned}$$



Regresia liniară

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

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



$$\begin{aligned} g(1) &= 2, \quad g(2) = 2.5, \quad g(3) = 3.5; \\ \min E &= (3 - 2)^2 + (1 - 2.5)^2 + (4 - 3.5)^2 \end{aligned}$$

Notes

Notes

Regresia liniară

Soluția se poate calcula explicit:

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

unde

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

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

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

8/30

Notes

Regresia liniară - alt raționament

Condițiile de interpolare pentru datele din tabel și

$g(x) = c_0 + c_1 x$:

$$c_0 + c_1 x_k = y_k, \quad k = 1, \dots, m, \quad (11)$$

⇒ sistem supradeterminat de ecuații. Dacă notăm

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

atunci (11) se scriu compact

$$\mathbf{A}\mathbf{c} = \mathbf{y}, \quad (13)$$

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

$$\mathbf{A}^T \mathbf{A} \mathbf{c} = \mathbf{A}^T \mathbf{y}. \quad (14)$$

9/30

Notes

Regresia liniară - alt raționament

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

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

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

$$\begin{cases} 3c_0 + 7c_1 = 8, \\ 7c_0 + 21c_1 = 21 \end{cases} \Rightarrow c_0 = 1.5, \quad c_1 = 0.5. \quad (17)$$

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 funcția definită de

$$E = \sum_{k=0}^m (f(x_k) - g(x_k))^2;$$

- 2 Se impun condițiile de minimizare;
- 3 Se rezolvă un sistem algebric liniar de dimensiune 3.

Notes

Notes

Cazul general

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

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

unde $\varphi_i(x)$ se numesc *funcții de bază*.

OBS: $n + 1 < m + 1$

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

Notes

Cazul general

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

Punctul de minim al acestei funcții este un punct critic:

$$\frac{\partial E}{\partial c_i} = 0, \quad i = 0, \dots, n. \quad (21)$$

Înlocuind expresia (20) în (21) rezultă

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

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

Notes

Cazul general - alt raționament

Condițiile de interpolare pentru datele din tabel și $g(x)$

$$g(x_k) = y_k, \quad k = 1, \dots, m, \quad (23)$$

⇒ sistem supradeterminat de ecuații. Notăm

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

atunci (23) se scriu compact

$$\mathbf{Ac} = \mathbf{y}, \quad (25)$$

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

$$\mathbf{A}^T \mathbf{Ac} = \mathbf{A}^T \mathbf{y}. \quad (26)$$

Notes

Funcții de bază clasice

Sistemul normal de rezolvat

$$\mathbf{Nc} = \mathbf{t} \quad (27)$$

are, în cazul *regresiei liniare* forma

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

Notes

Funcții de bază clasice

Sistemul normal de rezolvat

$$\mathbf{Nc} = \mathbf{t} \quad (29)$$

are, în cazul *regresiei parabolice* $g(x) = c_0 + c_1x + c_2x^2$,
forma:

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

OBS:

Generalizarea este ușoară pentru $n > 2$, dar sistemul este slab condiționat. :(

Notes

Funcții de bază clasice

Sistemul normal de rezolvat

$$\mathbf{Nc} = \mathbf{t} \quad (29)$$

are, în cazul *regresiei parabolice* $g(x) = c_0 + c_1x + c_2x^2$,
forma:

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

OBS:

Generalizarea este ușoară pentru $n > 2$, dar sistemul este slab condiționat. :([Remedii ?](#)

Notes

Funcții de bază clasice

Remedii:

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

Funcții de bază - polinoame Chebyshev

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

$$T_0(x) = 1, \quad (31)$$

$$T_1(x) = x, \quad (32)$$

$$T_j(x) = 2xT_{j-1}(x) - T_{j-2}(x), \quad (33)$$

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

Notes

Notes

Funcții de bază ortonormale

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

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

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

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

Funcții de bază ortonormale

Dacă notăm cu \mathcal{G} spațiul vectorial al funcțiilor generat de funcții de bază $\varphi_j(x)$

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

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

Notes

Notes

Algoritm - pași principali

Date:

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

Se aleg:

- funcțiile de bază (în consecință și gradul polinomului de aproximare).

Se calculează:

- valorile polinomului de aproximare în punctele dorite.

Algoritm - pași principali

CMMP_SystemNormal

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

Notes

Notes

Probleme de CMMP neliniare

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

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

ecuație neliniară în $c \Rightarrow$

- rezolvare de ecuații neliniare;
- metode de minimizare.

Probleme de CMMP neliniare

Pentru setul de date (x_k, y_k) , $k = 0, m$ se caută $g(x) = e^{cx}$.
Acestea sunt **probleme de CMMP neliniare**.

OBS:

Reformulare ca o problemă de CMMP liniară:

Pentru setul de date $(x_k, \ln(y_k))$, $k = 0, m$ se caută $h(x) = cx$.
Valoarea c din problema reformulată nu este soluția ecuației neliniare (39), dar aproximarea $e^{h(x)}$ s-ar putea să fie satisfăcătoare.

Notes

Notes

Concluzii

- Se caută aproximarea unor date dintr-un tabel de valori;
- CMMP găsește o funcție a.î. distanța dintre ea și tabelul de date să fie minimă (în normă Euclidiană);
- Aceasta problemă este echivalentă cu rezolvarea unui sistem de ecuații supradimensionat $\mathbf{Ac} = \mathbf{y}$;
- Acesta se poate aduce la un sistem de ecuații normal, cu matricea coeficienților pătrată;
- Condiționarea sistemului normal depinde de alegerea funcțiilor 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.

Notes

Notes

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 reziduuului

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

Există și alți algoritmi CMMP, care nu assemblează sistemul normal (bazați 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 găsește la biblioteca UPB, puteți verifica accesând catalogul <http://www.library.pub.ro/>.

Notes

Notes
