

Aplikace matematiky

Milan Geryk

Algoritmy. 46. COR. Lineární regresní polynom s vedlejší podmínkou

Aplikace matematiky, Vol. 28 (1983), No. 2, 147--152

Persistent URL: <http://dml.cz/dmlcz/104017>

Terms of use:

© Institute of Mathematics AS CR, 1983

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* <http://project.dml.cz>

46. COR

LINEÁRNÍ REGRESNÍ POLYNOM S VEDLEJŠÍ PODMÍNKOU

Ing. RNDr. MILAN GERYK, CSc.

Přerovské strojírny, n. p., Kojetínská, 750 53 Přerov

V soustavě souřadnic x, y jsou dány 3 skupiny bodů. V první skupině je K bodů $[x_{1s}, y_{1s}]$, kterými hledaná křivka – odpovídající polynom – prochází; $K \geq 0$; žádné dva body této skupiny nemají stejnou souřadnici x . Ve druhé skupině je $L \geq 1$ bodů $[x_{2u}, y_{2u}]$, jejichž pořadnice byly získány empiricky a jsou zatíženy chybami; tyto je třeba vyrovnat metodou nejmenšího součtu čtverců; body, jejichž souřadnice x je shodná se souřadnicí x některého bodu první skupiny, neovlivní průběh křivky. Ve třetí skupině je $Q \geq 0$ bodů $[x_{3z}, y_{3z}]$, které se na konstrukci křivky nepodílejí, je však třeba i pro jejich souřadnice x vypočítat vyrovnaná Y . Celkový počet bodů je $N = K + L + Q$.

Řešení hledáme ve tvaru polynomu M -tého stupně; $M \geq 1$; $M \geq K$; $M \leq K + L - 1$. Vyjdeme z výrazu založeného na Lagrangeově interpolačním vzorci

$$(1) \quad Y(x) = \sum_{s=1}^K \left[y_{1s} \prod_{t=1, \dots, K, t \neq s} \frac{x - x_{1t}}{x_{1s} - x_{1t}} \right] + \left[\prod_{t=1}^K (x - x_{1t}) \right] \sum_{r=1}^{M-K+1} A_r \cdot x^{r-1}$$

Koeficienty $A_i, i = 1, 2, \dots, M - K + 1$ se mají určit z podmínky, aby součet čtverců odchylek vyrovnaných pořadnic v bodech druhé skupiny byl minimální. Položíme rovny nule první parciální derivace součtu čtverců odchylek podle jednotlivých A_i se zahrnutím vah W_u (W_u kladné) a dostaneme $M - K + 1$ normálních rovnic; v -tá rovnice má tvar

$$(2) \quad \sum_{u=1}^L \{ [y_{2u} - Y(x_{2u})] x_{2u}^{v-1} W_u \prod_{t=1}^K (x_{2u} - x_{1t}) \} = 0$$

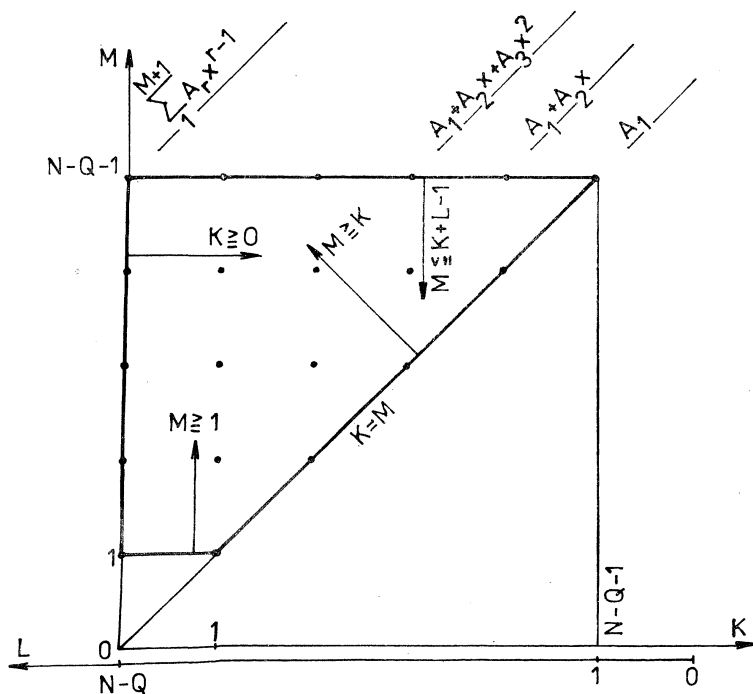
Po dosazení za $Y(x_{2s})$ z (1) a úpravě dostaneme soustavu lineárních rovnic s neznámými koeficienty A_r :

$$(3) \quad \sum_{r=1}^{M-K+1} \left\{ A_r \sum_{u=1}^L [x_{2u}^{r+v-2} W_u \prod_{t=1}^K (x_{2u} - x_{1t})^2] \right\} = \\ = \sum_{u=1}^L \left\{ x_{2u}^{v-1} W_u \left[\prod_{t=1}^K (x_{2u} - x_{1t}) \right] \left[y_{2u} - \sum_{s=1}^K \left(y_{1s} \prod_{\substack{t=1 \\ t \neq s}}^K \frac{x_{2u} - x_{1t}}{x_{1s} - x_{1t}} \right) \right] \right\},$$

$$v = 1, 2, \dots, M - K + 1$$

Subrutina COR napsaná v jazyku FORTRAN IV je založena na postupu výše uvedeném. Pracuje v dvojnásobné přesnosti po redukci x na interval $\langle 0; 1 \rangle$. Soustava rovnic je v našem příkladě řešena Gauss-Jordanovou eliminací pomocí subrutiny DJULIE, jejíž algolské znění je v [2]. Vyžaduje pomocné pole IPORAD, v jeho první buňce zanechává při výstupu informaci o úspěšném vyřešení. Počet $M - K + 1$ koeficientů A , hledaného polynomu je označen MK. Vektor zadaných x převezme subrutina v poli ZZ, vektor zadaných y v poli YY, informaci o příslušnosti ke skupině v poli IK, všechny o rozměru N. Bodům první skupiny odpovídá v poli IK záporné číslo, bodům třetí skupiny nula, bodům druhé skupiny kladné číslo, které navíc svou hodnotou udává váhu W_u (real). Pole XX slouží pro redukovaná x , pole UU pro vyrovnaná Y . Pole A obsahuje vektor pravých stran rovnic a při výstupu koeficienty A_r , $r = 1, 2, \dots, MK$ pro redukovaná x . Pole B je určeno pro matici soustavy; při výstupu nulová hodnota $B(1,1)$ signalizuje neúspěch řešení. Pole C, D, E jsou pomocná; E lze ekvivalencí ztotožnit s IPORAD. Pole XY a POLE obsahují ekvidistantní x od 0 do 1 a odpovídající Y a mohou sloužit např. k nakreslení grafu.

Zahrnutím vah lze body tří skupin učinit rovnoprávnými: body první skupiny mají nekonečnou váhu, body druhé skupiny váhu w_u , body třetí skupiny váhu nulovou. Před vstupem do COR uživatel obsadí N, K, MK, L, ZZ, YY, IK.



Obr. 1. Obor K, M, L při daném $N-Q$.

Výstup s nulovou hodnotou $B(1,1)$ nastává v těchto případech:

1. $M < K$ 2. $M < 1$ 3. $M \geq K + L$ 4. $N < K + L$
5. $K < 0$ 6. $L < 1$ 7. $x_{\max} - x_{\min} \leq 10^{-28}$

8. Některý dělitel při řešení soustavy lineárních rovnic je absolutně menší nebo roven 10^{-28} . V tomto případě bude při výstupu $IPORAD(1) = 0$.

9. V poli IK je méně než K čísel záporných nebo méně než L čísel kladných.

Obor použitelnosti subroutiney COR v závislosti na K , L , M je silně vytažen na obr. 1. Do úvahy přicházejí jen zakroužkované celočíselné hodnoty, kterých je $(N - Q + 1) \cdot (N - Q) / 2 - 1$. Obor podle obr. 1 platí v případě x navzájem různých, jinak je nutno ho ještě zmenšit. Krajní případ obr. 1 představuje $M = K + L - 1 = N - Q - 1$ pro x navzájem různá; křivka prochází všemi body bez ohledu na váhy. Příklad je uveden v tab. I.

Tabulka I. – příklad použití

$N = 6, K = 2, L = 3, Q = 1$

ZZ	2	6	10	7	14	18
YY	100	200	0	300	250	330
IK	-1	1	0	1	1	-1

a) $M = 2, MK = 1$

$$A_1 = -339,869972488$$

UU 100,0000 207,8511 282,1348 229,5690 322,8511 330,0000

b) $M = 3, MK = 2$

$$A_1 = -1102,96806894 \quad A_2 = 1558,91903745$$

UU 100,0000 243,9190 261,7954 256,1951 256,2741 330,0000

c) $M = 4, MK = 3$

$$A_1 = 5656,96022727 \quad A_2 = -25\,896,6599026 \quad A_3 = 24\,195,4626623$$

UU 100,0000 200,0000 464,3506 300,000 250,0000 330,0000

ALGORITMUS

SUBROUTINE COR(N,K,MK,L,XX,YY,ZZ,UU,POLE,XY,IK,IPORAD,A,
/B,C,D,E)

- C N POCET BODU CELKEM VCETNE NEPOUZITYCH S IK=0.. $N=K+L+Q$
- C K POCET PEVNYCH BODU, GE.0, IK ZAPORNE
- C MK POCET NEZNAMÝCH KOEFICIENTU POLYNOMU, GT.0
- C L POCET VOLNYCH BODU, GT.0, IK Kladne znaci vahy
- C M STUPEN POLYNOMU, GT.0
- C Q POCET NEPOUZITYCH BODU, GE.0, IK=0.
- C XX POLE PRO REDUKOVANA X – NEZAVISLE PROMENNE, OD 0 DO 1
- C YY POLE ZADANYCH Y – ZAVISLE PROMENNYCH

```

C   ZZ POLE ZADANYCH X
C   UU POLE PRO VYROVNANA Y
C   POLE POLE PRO GRAF
C   XY POLE PRO TABULKU U GRAFU
C   IK POLE PRO VAHY A TRIDICI ZNAKY
C   IPORAD POLE PRO PODPROGRAM SOUSTAVY LIN. ROVNIC
C   A POLE PRAVYCH STRAN SOUSTAVY ROVNIC A PRO VYSLEDNE KOEF.
C   B POLE PRO LEVE STRANY SOUSTAVY ROVNIC
C   C,D,E POMOCNA POLE, V HLAV. PROGRAMU ZTOTOZNIT E S IPORAD
C   SUBROUTINA PRACUJE V DOUBLE PRECISION
      INTEGER IPORAD(MK)
      DOUBLE PRECISION XX(N),YY(N),ZZ(N),UU(N),A(MK,1),B(MK,MK),
/C(K),D(K),E(K),XS,YS,XLS,YLS,SX,SXX,DX
      REAL POLE(52),XY(51),IK(N)
      M=K+MK-1
      IF(MK.LE.0.OR.N.LT.L+K.OR.K.LT.0.OR.M.LT.1.OR.M.GE.L+K)
/GO TO 999
C   REDUKCE X DO INTERVALU 0, 1
      YS=ZZ(1)
      XS=YS
      DO 1000 I=2,N
      IF(ZZ(I).GT.XS)XS=ZZ(I)
      IF(ZZ(I).LT.YS)YS=ZZ(I)
1000 CONTINUE
      DX=XS-YS
      IF(DABS(DX).LE.1.D-28) GO TO 999
      DO 1001 I=1,N
1001 XX(I)=(ZZ(I)-YS)/DX
C   NULOVANI LEVYCH A PRAVYCH STRAN SOUSTAVY NORMALNICH ROVNIC
      DO 1002 I=1,MK
      A(I,1)=0.D0
      DO 1002 J=1,MK
1002 B(I,J)=0.D0
C   HLEDANI PEVNYCH BODU V POLI IK S PROMENNOU II
      IF(K) 999,4,1
      1 II=0
      DO 1003 I=1,K
      2 II=II+1
      IF(II.GT.N) GO TO 999
      IF(II) 3,2,2
      3 C(I)=XX(II)
1003 D(I)=YY(II)
C   JJ SLOUZI K HLEDANI VOLNYCH BODU V POLI IK,
C   SOUCET PODLE VOLNYCH BODU, ASPON 1 EXISTUJE
C   TVORENI NORMALNICH ROVNIC
      4 JJ=0
      DO 1004 J=1,L
      5 JJ=JJ+1
      IF(JJ.GT.N) GO TO 999
      IF(II(JJ)) 5,5,6

```

```

6  XLS=XX(JJ)
   YLS=YY(JJ)
   SX=1.D0
   SXX=SX
   IF(K)999,9,7
7  DO 1005 I=1,K
   E(I)=D(I)
1005 SX=SX*(XLS-C(I))
   IF(SX) 8,1004,8
8  SXX=SX*SX
   DO 1006 I=1,K
   E(I)=E(I)*SX/(XLS-C(I))
   DO 1007 II=1,K
   IF(I.NE.II) E(I)=E(I)/(C(I)-C(II))
1007 CONTINUE
1006 YLS=YLS-E(I)
   YLS=YLS*SX
9  DO 1008 I = 1, MK
   A(I,1)=A(I,1)+YLS*IK(JJ)
   SX=SXX
   SXX=SXX*XLS
   DO 1009 II=1,MK
   B(II,1)=B(II,1)+SX*IK(JJ)
1009 SX=SX*XLS
1008 YLS=YLS*XLS
1004 CONTINUE
C  (DJULIE NERESI 1 ROVNICI O 1 NEZNAME)
   IF(MK-1) 999,10,12
10  IF(DABS(B(1,1))-1.D-28)998,998,11
11  A(1,1)=A(1,1)/B(1,1)
   IPORAD(1)=1
   GO TO 13
C  RESENI SOUSTAVY NORMALNICH ROVNIC, VYSLEDKY V A
12  CALL DJULIE(MK,1,B,A,IPORAD)
   IF(IPORAD(1)) 999,999,13
13  B(1,1)=1.D0
C  DOSAZENI
   II=N+51
   DO 1010 J=1,II
   IF(J.GT.N) GO TO 18
   XLS=XX(J)
   GO TO 19
18  JN=J-N
   XLS=0.02*FLOAT(JN-1)
   XY(JN)=XLS*DX+YS
19  SX=0.D0
   SXX=A(MK,1)
   IF(MK-1)999,15,14
14  DO 1011 I=2,MK
   JJ=MK-I+1

```

```

1011 SXX=SXX*XLS+A(JJ,1)
15 IF(K) 999,20,16
16 DO 1012 I=1,K
    E(I)=D(I)
    SXX=SXX*(XLS-C(I))
    DO 1013 JJ=1,K
    IF(I.NE.JJ)E(I)=E(I)*(XLS-C(JJ))/(C(I)-C(JJ))
1013 CONTINUE
1012 SX=SX+E(I)
20 IF(J.LE.N) GO TO 21
    POLE(JN)=SX+SXX
    GO TO 1010
21 UU(J)=SX+SXX
1010 CONTINUE
    POLE(52)=0.
17 RETURN
998 IPORAD (1)=0
999 B(1,1)=0.D0
    GO TO 17
    END

```

Literatura

- [1] *Rektorys a kol.*: Přehled užití matematiky, Praha 1963.
- [2] *Pernica*, in: Výběr informací z organizační a výpočetní techniky, 1969, č. 6.
- [3] *Vogel*: Programování v jazyku FORTRAN, Praha 1971.