УДК 517

ХААР ФУНКЦИЯСЫ ЖАҒДАЙЫНДАҒЫ РЕГРЕССИЯЛЫҚ АНАЛИЗ

Еркеғали Жансая Еркеғалиқызы
zhansayayerkegali@gmail.com
Л.Н.Гумилев атындағы Еуразия ұлттық университетінің магистранты, Нұр-Сұлтан, Қазақстан.
Ғылыми жетекшісі-А.А.Адамов

Абстракт

Базистік функция ретінде Хаар функциясын ала отырып, сызықты регрессиялық моделдің белгісіз параметрлері мен дисперсиясы бағаланған.

Кіріспе

Сызықты регрессиялық анализдеізделінетін функция мына түрде болады.

$$\delta(x) = \sum_{i=1}^{m} \theta_i \, \varphi_i(x) \qquad (1.1)$$

Мұндағы базистік функциялар $\varphi_i(x)$, $i=1 \dots m$ X компактысында үзіліссіз.[1]

Базистік функция ретінде Хаар функциясын алып, белгісіз параметрлер мен дисперсияны анықтайық.

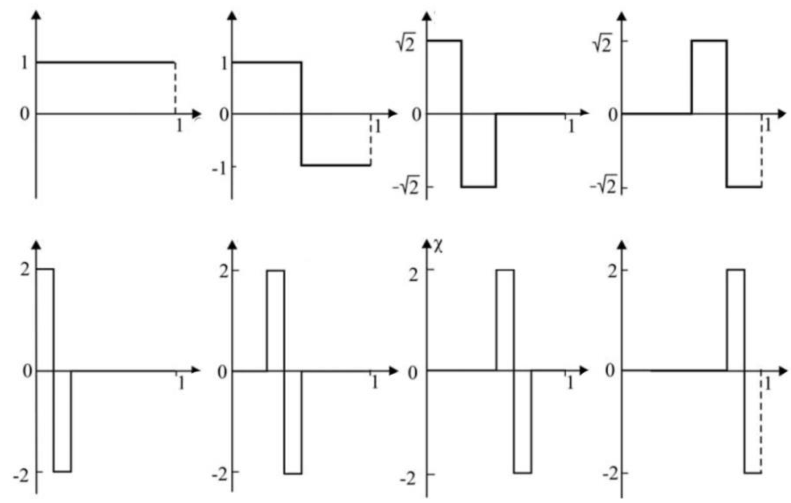
μ өлшемде X жиынында берілген Хаар функциясын келесі түрде анықтайық. $\mu(\mathrm{X}) = 1$ өлшемді X жиынын, өлшемі μ-ға тең қиылыспайтын $d(s, i_s)$ $(1 \leq s \leq m, i_s = 1 \dots 2^{s-1})$түріндегі 2^m жиыншаларына бөлеміз.$d(s, i_s), i_s = 1 \dots 2^{s-1}$ жиыншалары келесі теңдікпен анықталады.

$$d(s,i_s) = d(s+1, 2i_s - 1) \cup d(s+1, 2i_s)$$

$$\omega_{i_s}^{s} = \begin{cases} 2^s, \ x \in d(s+1, 2i_s - 1) \\ -2^{s/2}, \ x \in d(s+1, 2i_s) \quad (1.2) \\ 0, \qquad\qquad x \overline{\in} d(s, i_s) \end{cases}$$

$$\varphi_1(x) = 1 \quad \varphi_j(x) = \omega_{i_s}^{s}, \quad j = 2^s + i_s s = 1 \dots m$$

Егерμ [0,1] аралығында Лебег өлшемі болса, онда Хаар функциясы келесі түрде құрылады.[2] Алғашқы функциялар төмендегідей.



Сурет 1- Алғашқы сегіз Хаар функциясының графигі

Зерттеу қарапайым болу үшін, әуелі m=1,2 деп есептейік. Яғни, X жиыны 2^m жиыншаларға бөлінсін. Олай болса, сызықты регрессиялық модельде(1.1) сәйкесінше 2 және 4 белгісіз параметрлер болады. m=1 жағдайында Хаар функциясы келесідей жазылады

$$\varphi_1(x) = 1, \quad \varphi_2\,(x) = \begin{cases} 1, \ x \in [0, \frac{1}{2}] \\ -1, \ x \in [\frac{1}{2}, 1] \end{cases} (1.3)$$

$k_i (i = 1{,}2)$ бұл $[\frac{i-1}{n}, \frac{i}{n}]$ жиыншасынан алынған және $k_1 + k_2 = N$ болатын нүктелер саны болсын.

Ондай жағдайда ақпараттық матрицамыз төмендегідей болады.

$$F^T F = \begin{bmatrix} [\varphi_1, \varphi_1] & [\varphi_1, \varphi_2] \\ [\varphi_2, \varphi_1] & [\varphi_2, \varphi_2] \end{bmatrix} = \begin{bmatrix} k_1 + k_2 & k_1 - k_2 \\ k_1 - k_2 & k_1 + k_2 \end{bmatrix} \qquad (1.4)$$

Матрица анықтауышы $det\|F^T F\| = 4k_1 k_2 (1.5)$

$F^T F$ матрицасы ерекше болмауы үшін, әрбір $d(s, i_s)$ жиынында ең болмағанда бір x_j (j=1..N) нүктесі жату керек, яғни $k_j \geq 1, \ j = 1{,}2$ шарты орындалу қажет. Белгісіз параметрлерді ең кіші квадраттар әдісі формуласын қолданыпанықтай аламыз.[1]

$$\hat{\theta} = (F^T F)^{-1} F^T Y = \begin{bmatrix} \frac{1}{2}[\frac{1}{k_1}\beta_1 + \frac{1}{k_2}\beta_2] \\ \frac{1}{2}[\frac{1}{k_1}\beta_1 - \frac{1}{k_2}\beta_2] \end{bmatrix} (1.6)$$

$$(F^TF)^{-1} = \frac{1}{4}\begin{vmatrix} \frac{1}{k_1}+\frac{1}{k_2} & \frac{1}{k_1}-\frac{1}{k_2} \\ \frac{1}{k_1}-\frac{1}{k_2} & \frac{1}{k_1}+\frac{1}{k_2} \end{vmatrix} F^TY = \begin{bmatrix} \beta_1+\beta_2 \\ \beta_1-\beta_2 \end{bmatrix} \quad (1.7)$$

Ал дисперсиясы мына түрде өрнектеледі[1]

$$\mathrm{D}\hat{\theta} = \sigma^2(\mathrm{F^T F})^{-1}(1.8)$$

$$D\hat{\theta} = \sigma^2/4\begin{vmatrix} \frac{1}{k_1}+\frac{1}{k_2} & \frac{1}{k_1}-\frac{1}{k_2} \\ \frac{1}{k_1}-\frac{1}{k_2} & \frac{1}{k_1}+\frac{1}{k_2} \end{vmatrix}$$

Енді m=2 болсын. Онда біз жоғарғыдағы амалдарды орындай отырып, келесі нәтижелерге келеміз.

Хаар функциясы төмендегідей анықталады.

$$\varphi_1(x) = 1, \quad \varphi_2(x) = \begin{cases} 1, & x \in [0, \frac{1}{2}] \\ -1, & x \in [\frac{1}{2}, 1] \end{cases},$$

$$\varphi_3(x) = \begin{cases} \sqrt{2}, & x \in [0, \frac{1}{4}] \\ -\sqrt{2}, & x \in [\frac{1}{4}, \frac{1}{2}] \\ 0, & x \in [\frac{1}{2}, 1] \end{cases}, \varphi_4(x) = \begin{cases} \sqrt{2}, & x \in [\frac{1}{2}, \frac{3}{4}] \\ -\sqrt{2}, & x \in [\frac{3}{4}, 1] \\ 0, & x \in [0, \frac{1}{2}] \end{cases} \quad (1.9)$$

Ақпараттық матрица

$$F^TF = \begin{vmatrix} \sum_{i=1}^4 k_i & \sum_{i=1}^2 k_i - \sum_{i=3}^4 k_i & \sqrt{2}[k_1-k_2] & \sqrt{2}[k_3-k_4] \\ \sum_{i=1}^2 k_i - \sum_{i=3}^4 k_i & \sum_{i=1}^4 k_i & \sqrt{2}[k_1-k_2] & -\sqrt{2}[k_3-k_4] \\ \sqrt{2}[k_1-k_2] & \sqrt{2}[k_1-k_2] & 2[k_1+k_2] & 0 \\ \sqrt{2}[k_3-k_4] & -\sqrt{2}[k_3-k_4] & 0 & 2[k_3+k_4] \end{vmatrix} \quad (1.10)$$

Анықтауыш $det\|F^TF\| = 16^2 k_1k_2k_3k_4(1.11)$

$k_i \geq 1,\ i = 1{,}2$ шартына сай θ белгісіз параметрлері үшін ең кіші квадраттар әдісі бағалауын аламыз.

$$\hat{\theta} = (F^TF)^{-1}F^TY = \begin{vmatrix} \frac{1}{4}\sum_{i=1}^4 \frac{1}{k_i}\beta_i \\ \frac{1}{4}[\sum_{i=1}^4 \frac{1}{k_i}\beta_i - \sum_{i=3}^4 \frac{1}{k_i}\beta_i] \\ \frac{\sqrt{2}}{4}[\frac{1}{k_1}\beta_1 - \frac{1}{k_2}\beta_2] \\ \frac{\sqrt{2}}{4}[\frac{1}{k_3}\beta_3 - \frac{1}{k_4}\beta_4] \end{vmatrix}(1.12)$$

Мұндағы $F^TY = \left\{\sum_{i=1}^4 \beta_i, \sum_{i=1}^4 \beta_i - \sum_{i=3}^4 \beta_i, \sqrt{2}[\beta_1-\beta_2], \sqrt{2}[\beta_3-\beta_4]\right\}^T$

Бұл бағалаудың дисперсиялық матрицасы келесідей болады

$$D\hat{\theta} = \sigma^2(F^TF)^{-1} = \sigma^2/16\begin{vmatrix} \sum_{i=1}^4 \frac{1}{k_i} & \sum_{i=1}^2 \frac{1}{k_i} - \sum_{i=3}^4 \frac{1}{k_i} & \sqrt{2}[\frac{1}{k_1}-\frac{1}{k_2}] & \sqrt{2}[\frac{1}{k_3}-\frac{1}{k_4}] \\ \sum_{i=1}^2 \frac{1}{k_i} - \sum_{i=3}^4 \frac{1}{k_i} & \sum_{i=1}^4 \frac{1}{k_i} & \sqrt{2}[\frac{1}{k_1}-\frac{1}{k_2}] & -\sqrt{2}[\frac{1}{k_3}-\frac{1}{k_4}] \\ \sqrt{2}[\frac{1}{k_1}-\frac{1}{k_2}] & \sqrt{2}[\frac{1}{k_1}-\frac{1}{k_2}] & 2[\frac{1}{k_1}+\frac{1}{k_2}] & 0 \\ \sqrt{2}[\frac{1}{k_3}-\frac{1}{k_4}] & -\sqrt{2}[\frac{1}{k_3}-\frac{1}{k_4}] & 0 & 2[\frac{1}{k_3}+\frac{1}{k_4}] \end{vmatrix} \quad (1.13)$$

Сонымен қатар алынған ε жоспар D оптималды болуы үшін төмендегі шарт орындалу қажет [1]

$$\varepsilon^{*} = \arg\max_{\varepsilon} \det M(\varepsilon) = \arg\min_{\varepsilon} \det D(\varepsilon) \qquad (1.14)$$

(1.5) және (1.11) формуласынан $k_1 = k_2 = \cdots = k_n$ болған жағдайда ақпараттық матрица анықтауышы өзінің максимум мәніне жетеді,алоған кері дисперсиялық матрица анықтауышы өзінің ең кіші мәніне жетеді, яғни N нүктелер саны бөліктеу санына еселі $N = Ln$ болу қажет.

Қорытынды және ұсыныстар:

Сызықты регрессиялық модельде,базистік функция ретінде Хаардың екі және төрт функциясы алынып, оның белгісіз параметрлері мен дисперсиясын бағалау зерттелді және Сонымен қатар D оптималдылық критериі бойынша ақпараттық матрица анықтауышы максимум (дисперсиялық матрица анықтауышы минимум) болатындай оптимал жоспардың k_i нүктелерін қалай таңдау қажет екендігі анықталды. Бұл тиімді жоспар құруға қатысты басқа да бағалау критерилерін зерттеуге мүмкіндік береді.

Қолданылған әдебиеттер тізімі

1. Ермаков С.М., Жиглявский А.А. Математическая теория оптимального эксперимента. М.,1987.

2. Ермаков С.М. Метод Монте-Карло и смежные вопросы. М., 1975.