

Kvantil regressiya modeli

Annotatsiya

Mazkur metodologik asos O‘zbekiston ko‘chmas mulk bozori holatini tahlil qilishda uy-joy narxlarining fundamental qiymatini yanada aniqroq baholashda qo‘llaniladigan modellar qamrovini kengaytirish yuzasidan tayyorlangan. Bunda, ko‘rsatkichlar nisbiy xatoliklari normal taqsimlanmagan va kuzatuv davri davomida keskin farq qiladigan qiymatlar mavjud holatlarda ham ko‘rsatkichlar ta’sirini aniqroq baholashda foydalaniladigan kvantil regressiya modeli yoritilgan. Xususan, kvantil regressiya modeli parametrlari shartli kvantil funksiyasining noma’lum parametrlarini baholashda nisbiy xatoliklarning absolyut qiymatlarining vaznli yig‘indisini minimallashtirish orqali aniqlanadi.

Tayanch so‘zlar: absolyut, kvantil, nisbiy xatolik, regressiya, simpleks jadval, o‘zgaruvchi, shartli o‘rtacha.

Kvantil regressiya modeli

Regressiya tahlilining maqsadi erksiz o'zgaruvchi va erkli o'zgaruvchilar orasidagi bog'liqlikni ko'rsatishdan iborat. Klassik chiziqli regressiya modeli shartli o'rtachani modellashtiradi¹. Shartli o'rtachani modellashtirish jarayoni eng kichik kvadratlar usuliga olib kelinib, hisoblash jarayoni sodda ko'rinishga o'tkaziladi. Kichik kvadratlar usuli orqali regression tahlil o'tkazishda ko'rsatkichlarning nisbiy xatoliklari normal bo'lmashligi va kuzatuv davri davomida nisbatan keskin farq qiladigan davrlarning bo'lishi natijalar aniqligining pasayishiga, ya'ni nisbiy xatoliklar yuqori bo'lishiga olib kelishi mumkin. Shu sababli, regression tahlil o'tkazishda kichik kvadratlar modeli o'rniga kvantil regressiya (quantile regression) modelidan foydalanish natijalar aniqligini oshiradi².

Kvantil regressiya modeli parametrlari shartli kvantil funksiyasining noma'lum parametrlarini baholashda nisbiy xatoliklarning absolyut qiymatlari vaznli yig'indisining eng kichik qiymatini baholash orqali aniqlanadi. Kvantil regressiya modeli shartli kvantillardagi o'zgarishlarni baholaydi. Shartli median regressiyasi, kvantil regressiyaning alohida holati bo'lib, unda shartli markaziy, ya'ni 50-kvantil funksiyasi sifatida modellashtiriladi. Umuman olganda, taqsimotning markaziy bo'lmagan pozitsiyalarini tavsiflash uchun boshqa har qanday kvantillardan foydalanib, turli kvantillarda erkli va erksiz o'zgaruvchilar orasidagi bog'liqlik uchun tahlil o'tkazish mumkin.

Kvantil regressiya modeli tenglamasi oddiy chiziqli regressiya tenglamasidan farq qilmaydi. Kvantil regressiya modelining regressiya tenglamasi quyidagi ko'rinishda bo'ladi:

$$y_i = x_i\beta_q + \varepsilon_i$$

Bu yerda,

y_i – $n \times 1$ o'lchamli erksiz o'zgaruvchilar vektori;

x_i – $n \times k$ o'lchamli erkli o'zgaruvchilar matritsasi;

β_i – $k \times 1$ o'lchamli noma'lum parametrlar vektori;

ε_i – $n \times 1$ o'lchamli nisbiy xatoliklar vektori;

q – kvantil.

¹ Hastie, T., Tibshirani, R., & Friedman, J. (2017). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer Series in Statistics.

² Cameron, A., & Trivedi, P. (2005). Microeconometrics: Methods and Applications. Cambridge University Press.

Kvantil regressiya modelining parametrlarini baholashda musbat va manfiy xatoliklarga turli vaznlar beradi. Kvantil regressiya parametrlari nisbiy xatoliklar absolyut qiymatlari vaznli yig'indisini $(Q(\beta_q))$ minimallashtirish natijasida topilib, quyidagi ko'rinishda bo'ladi:

$$Q(\beta_q) = \sum_{i:y_i \geq x_i \beta_q} q|y_i - x_i \beta_q| + \sum_{i:y_i < x_i \beta_q} (1 - q)|y_i - x_i \beta_q|.$$

Tenglikdagi noma'lum parametrlarni topishda matematik dasturlashni optimallashtirish modelining chiziqli dasturlash turidan foydalanish talab qilinadi.

Chiziqli dasturlash masalalarini hal qilish uchun avval chiziqli dasturlashning maqsad funksiyasi va chegaraviy shartlarini aniqlab olish zarur. Kvantil regressiyaning shartli kvantil funksiyasida musbat va manfiy xatoliklarga asimmetrik vazn³ beriladi:

$$Q(\beta_q) = \sum_{i:y_i \geq x_i \beta_q} q|y_i - x_i \beta_q| + \sum_{i:y_i < x_i \beta_q} (1 - q)|y_i - x_i \beta_q|.$$

Ushbu yig'indini minimallashtirish orqali noma'lum parametrlarning optimal bahosi topiladi. Chiziqli dasturlashning maqsad funksiyasining $(Q(\beta_q))$ ko'rinishini quyidagicha o'zgartirish natijasida hosil bo'ladi⁴:

$$Q(\beta_q) = \sum_{i:y_i \geq x_i \beta_q} q s_i^+ + \sum_{i:y_i < x_i \beta_q} (1 - q) s_i^-.$$

³ Asimmetrik vazn – bir-biriga teng bo'lmagan turli xil vaznlar.

⁴ Abdullahi, I. (2015). Analysis of quantile regression as alternative to ordinary least squares regression. Department of Mathematics Ahmadu Bello University.

Tenglikdagi s_i^+ va s_i^- lar mos ravishda musbat va manfiy xatoliklar bo'lib, ular quyidagi belgilashlar natijasida hosil qilinadi:

$$\begin{aligned} s_i^+ &= y_i - x_i \hat{\beta}_q \\ s_i^- &= -(y_i - x_i \hat{\beta}_q) \\ s_i^+ &\geq 0, s_i^- \geq 0. \end{aligned}$$

Bu yerda, $\hat{\beta}_q$ – model yordamida baholangan parametr.

Chiziqli dasturlashning chegaraviy shartlaridan kelib chiqqan holda quyidagi tenglik hosil qilinadi:

$$y_i = x_i \hat{\beta}_q + s_i^+ - s_i^-$$

Ushbu chegaraviy shartda musbat va manfiy xatoliklar bir vaqtning o'zida noldan farqli bo'lishi mumkin emas. Bunda parametrlarning kamida bittasi yoki ikkalasi ham nolga teng bo'lish holatlari yuzaga kelishi mumkin.

Barcha kvantillar uchun chegaraviy shartlar o'zgarmas bo'lib, faqatgina maqsad funksiya o'zgaruvchan bo'ladi. $Q(\beta_q)$ funksiyaning minimallashtirish chiziqli dasturlashning maqsad funksiyasini minimallashtirish muammosiga ekvivalent hisoblanadi. Shu sababli, chiziqli dasturlash masalasidagi β_q parametrlarning optimal qiymatini topish kifoya.

Simpleks usuli orqali noma'lum parametrlarni optimal qiymatini topishda dasturiy almashtirib borish jarayonlaridan foydalanadi. Ma'lumotlar qatorining kengayib borishi regressiya parametrlarining aniqligini oshiradi. Bunda, chiziqli dasturlash masalasining maqsad funksiyasidagi $(Q(\beta_q))$ yig'indini $Ax = B$ chegaraviy sharti ostida minimallashtiradi.

Simpleks usuli

Simpleks usuli chiziqli dasturlash masalasini yechishning asosiy usullaridan biri bo'lib, ketma-ket yaqinlashib borish orqali $x_1, x_2, \dots, x_n$ o'zgaruvchilarning shunday optimal qiymati topiladi va ushbu qiymatlar natijasida maqsad funksiyasi maksimal yoki minimal qiymatiga erishadi⁵.

⁵ Winston, W. (2004). Operations Research: Applications and Algorithms. Indiana University.

Maksimal qiymat beruvchi chiziqli dasturlash masalasi quyidagi ko‘rinishda ifodalanadi:

$$\begin{aligned} \sum_{j=1}^n a_{ij}x_j &\leq y_i, \quad (i = 1, 2, \dots, m) \\ x_j &\geq 0, \quad (j = 1, 2, \dots, n) \\ Z = \sum_{j=1}^n c_{ij}x_j &\rightarrow \max. \end{aligned}$$

Masalani yechish uchun simpleks jadvali tuziladi. Simpleks jadvalini tuzishda avval berilgan masala kanonik shaklga keltiriladi. Bunda, kanonik shaklga keltirish uchun barcha chegaraviy shartlar tenglik ko‘rinishiga keltiriladi. Tengsizliklarni tenglik ko‘rinishiga keltirish uchun tengsizliklarga bazis o‘zgaruvchilar qo‘shiladi. Bazis o‘zgaruvchilarni qo‘shish natijasida $n - m$ ta o‘zgaruvchining qiymati nolga aylanib, qolgan m ta o‘zgaruvchi tenglamaning optimal qiymatini topish orqali yagona qiymati aniqlanadi. Bazis o‘zgaruvchilarni qo‘shgandan so‘ng tengsizliklar quyidagi tengliklar ko‘rinishiga keladi:

[illegible]

Bu yerda,

x_{n+i} – bazis o‘zgaruvchilar.

Bazis o'zgaruvchilarni boshqa o'zgaruvchilardan farqlash uchun mos ravishda $e_1, e_2, \dots, e_m$ deb belgilash kiritiladi. Maqsad funksiyasi tengligining o'ng tarafidagi ifoda esa, tenglikning chap tarafiga o'tkaziladi:

$$Z - c_1x_1 - c_2x_2 - \cdots - c_nx_n = 0.$$

O'zgartirishlardan so'ng simpleks jadvalini quyidagi ko'rinishda tuzish mumkin:

	x_1	x_2	$\dots$	x_s	$\dots$	x_n	b_i
e_1	a_{11}	a_{12}	$\dots$	a_{1s}	$\dots$	a_{1n}	b_1
e_2	a_{21}	a_{22}	$\dots$	a_{2s}	$\dots$	a_{2n}	b_2
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
e_r	a_{r1}	a_{r2}	$\dots$	a_{rs}	$\dots$	a_{rn}	b_r
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
e_m	a_{m1}	a_{m2}	$\dots$	a_{ms}	$\dots$	a_{mn}	b_m
Z	$-c_1$	$-c_2$	$\dots$	$-c_s$	$\dots$	$-c_n$	0

b_i ozod hadlarning barchasiga musbat bo'lish sharti beriladi. Agar ozod xad musbat bo'lmagan holatlarda, tenglikning o'ng va chap tarafi -1 ga ko'paytirilib musbat ko'rinishga keltiriladi.

Berilgan masala maksimallashtirish muammosi mavjudligi tufayli Z qatoridagi barcha parametrlar nomanfiy qiymatni qabul qilgunga qadar optimallashtirish jarayoni davom etadi.

Dastlab, mumkin bo'lgan asosiy yechim topiladi. Bunda, birinchi bazis o'zgaruvchilar ustuni asosiy yechim ustuni hisoblanadi va bu ustunda joylashmagan barcha o'zgaruvchilarning qiymati nolga teng hisoblanadi. Simpleks jadvalda mumkin bo'lgan e_i asosiy yechimlar aniqlanadi:

$$e_i = b_i \quad \text{va} \quad x_j = 0$$

Ushbu mumkin bo'lgan asosiy yechim Z qatoridagi barcha koeffitsiyentlar nomanfiy bo'lgan taqdirda optimal yechim hisoblanadi. Agar ushbu koeffitsiyentlardan birontasi manfiy bo'lsa keyingi simpleks jadvali tuziladi. Dastlab asosiy qator va asosiy ustunlar kesishmasidagi asosiy element topiladi. Asosiy ustunni topish uchun maqsad funksiyasi qatori manfiy parametrlari ichidan absolyut qiymati bo'yicha eng kattasi tanlanadi va o'sha ustun asosiy ustun hisoblanadi.

Masalan, manfiy elementlar ichida absolyut qiymati model jihatdan eng katta element s_s bo'lsa, ushbu s ustun asosiy ustun sifatida qaraladi. Asosiy ustun topilgandan so'ng asosiy qator aniqlanadi. Bu qatorni topish uchun ozod xadlar ustuni

elementlari mos ravishda asosiy ustun elementlariga bo‘lib chiqiladi va ulardan musbatlarining eng kichigi tanlanadi:

$$\min \left\{ \frac{b_1}{a_{1s}}; \frac{b_2}{a_{2s}}; \dots; \frac{b_r}{a_{rs}}; \dots; \frac{b_m}{a_{ms}} \right\}.$$

Ushbu nisbatlar ichida musbatlarning eng kichigi $\frac{b_r}{a_{rs}}$ bo‘lsa, a_{rs} element turgan qator asosiy qator deyiladi va ars elementning o‘zi esa asosiy element hisoblanadi. Asosiy element topilganidan so‘ng yangi simpleks jadvali bosqichma-bosqich to‘ldirib boriladi.

Ikkinchi simpleks jadvalini tuzish bir nechta bosqichlarda amalga oshiriladi. Dastlab, asosiy qator va asosiy ustun o‘zgaruvchilari (x_s va e_r) o‘rinlari almashtiriladi. So‘ngra, asosiy ustun elementlari asosiy elementga bo‘linib, yangi jadvalga manfiy qiymati bilan quyidagicha yoziladi:

$$a'_{is} = -\frac{a_{is}}{a_{rs}} \quad (i \neq r)$$

Asosiy qator elementlari ham asosiy elementga bo‘linib yangi jadvalga yoziladi:

$$a'_{rs} = \frac{a_{rj}}{a_{rs}} \quad (j \neq s)$$

Asosiy element teskari ko‘rinishga keltiriladi:

$$a'_{is} = \frac{1}{a_{rs}}$$

Yangi simpleks jadvalining qolgan elementlari esa, quyidagi formula orqali topiladi:

$$a'_{ij} = \frac{a_{ij}a_{rs} - a_{is}a_{rj}}{a_{rs}} = a_{ij} - \frac{a_{is}a_{rj}}{a_{rs}} \\ (j \neq s), (i \neq r).$$

Bu yerda,
 $a_{ij} - a'_{ij}$ elementning eski jadvaldagi mos elementi;
 $a_{is} - a_{ij}$ element turgan ustun bilan a_{rs} asosiy ustun kesishmasidagi element;
 $a_{rj} - a_{ij}$ element turgan ustun bilan a_{rs} asosiy element qatori kesishmasidagi element;
 a_{rs} – asosiy element.
 ars – asosiy element.

Asosiy yechimning simpleks jadvaldagi Z qator ozod xaddan tashqari barcha elementlari nomanfiy bo'lsa, bu olingan asosiy yechim yagona va u masalaning optimal yechimi bo'ladi. Agar Z qator elementlaridan kamida bittasi manfiy bo'lsa, optimallashtirish jarayoni yuqorida ko'rsatilgan tartibda optimal qiymatni topgunga qadar davom ettiriladi:

$$Z = \sum_{j=1}^n c_j x_j \longrightarrow \max.$$

Shuningdek, maksimallashtirish muammosini -1 ga ko'paytirib minimallashtirish muammosiga keltirish mumkin:

$$-Z = \sum_{j=1}^n -c_j x_j \longrightarrow \min.$$

Chiziqli dasturlashning simpleks usuli yordamida topilgan optimal parametrlar kvantil regressiya noma'lum parametrlarining optimal bahosi hisoblanadi. Bu parametrlar yordamida kvantil regressiyaning baholangan qiymatlari aniqlanadi.

Kvantil regressiya modeli orqali uy-joylarning 1 m^2 uchun fundamental narxini aniqlashda quyidagi kvantil regressiya tenglamasining 50-kvantilidan foydalanildi:

$$y_t = \beta_{1,q} y_{t-1} + \beta_{2,q} x_{1,t} + \beta_{3,q} x_{2,t} + \varepsilon_t.$$

Bu yerda,
 y_t – uy-joylarning t davrdagi bozor narxi;
 y_{t-1} – uy-joylarning bozor narxini $t - 1$ davrdagi qiymati;

$x_{1,t}$ – uy xo‘jaliklariga ajratilgan kreditlarning yillik o‘sinh sur‘ati;

$x_{2,t}$ – aholi ixtiyoridagi o‘rtacha daromad;

$\beta_{1,q}, \beta_{2,q}, \beta_{1,q}$ – noma‘lum koeffitsiyentlar;

q – kvantil qiymati;

ε_t – nisbiy xatolik.