

CHIZIQLI REGRESSIYA MODELLARINING ANIQLIK DARAJASINI OSHIRISH USULLARI

Toshboltayev F.O¹., Xudoyberdiyeva Sh.I²., Nizomov A.B².

¹FarDU katta o'qituvchi, p.f.b.f.d(PhD). ²FarDU magistranti,

shahzodaxonxudoyberdiyeva@gmail.com

Annotatsiya: Mazkur ilmiy maqolada chiziqli regressiya modellarining aniqlik darajasini oshirish muammosi har tomonlama tahlil qilinadi. Tadqiqot jarayonida regressiya modellarining aniqligiga ta'sir etuvchi asosiy omillar, jumladan, ma'lumotlarni oldindan qayta ishlash, multikolinearlik muammosi, parametrlarni baholashning zamonaviy yondashuvlari hamda Ridge va Lasso kabi regularizatsiya usullarining samaradorligi o'rganildi. Tajriba natijalari shuni ko'rsatadiki, mazkur usullarni kompleks qo'llash regressiya modellarining bashorat aniqligini sezilarli darajada oshiradi.

Kalit so'zlar: chiziqli regressiya, aniqlik, bashoratlash, multikolinearlik, regularizatsiya, Ridge regressiya, Lasso regressiya.

KIRISH. Chiziqli regressiya matematik statistikaning asosiy vositalaridan biri bo'lib, mustaqil va bog'liq o'zgaruvchilar orasidagi chiziqli bog'lanishni modellashtirish imkonini beradi. Ushbu model soddaligi va tushunarli interpretatsiyasi sababli iqtisodiyot, moliya, muhandislik, tibbiyot hamda ijtimoiy fanlarda keng qo'llaniladi. Biroq real ma'lumotlar ko'pincha shovqinli, to'liq bo'lmagan yoki kuchli o'zaro bog'liq o'zgaruvchilardan iborat bo'ladi. Natijada regressiya modeli aniqligi pasayadi va noto'g'ri xulosalar chiqarilish xavfi yuzaga keladi. Shu sababli regressiya modellarining aniqligini oshirish masalasi dolzarb ilmiy-amaliy muammo hisoblanadi. Mazkur tadqiqotning asosiy maqsadi chiziqli regressiya modelining aniqligiga ta'sir etuvchi omillarni aniqlash va ularni yaxshilashga xizmat qiluvchi usullarni tahlil qilishdan iborat.

METODOLOGIYA. Chiziqli regressiya modelining matematik asosi : Chiziqli regressiya modeli quyidagi ko'rinishda ifodalanadi:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon$$

bu yerda: y -bog'liq o'zgaruvchi, x_i -mustaqil o'zgaruvchilar, β_i -regressiya koeffitsiyentlari, ε -tasodifiy xato. Parametrlar odatda eng kichik kvadratlar usuli

(OLS) yordamida baholanadi. Ushbu usul kuzatilgan va bashorat qilingan qiymatlar orasidagi kvadrat farqlar yig'indisini minimallashtiradi.

Aniqlikka ta'sir etuvchi asosiy omillar. Ma'lumotlar sifati Ma'lumotlarning noto'liq yoki noto'g'ri bo'lishi regressiya aniqligiga bevosita salbiy ta'sir ko'rsatadi. Shu bois yo'qolgan qiymatlarni to'ldirish, shovqinlarni kamaytirish va normallashtirish muhim hisoblanadi. Multikolinearlik muammosi. Agar mustaqil o'zgaruvchilar o'zaro kuchli bog'langan bo'lsa, regressiya koeffitsiyentlari beqaror bo'ladi. Bu holat VIF (Variance Inflation Factor) orqali aniqlanadi.

Modelning murakkabligi. Haddan tashqari murakkab model overfitting holatiga olib keladi, ya'ni model o'rganish ma'lumotlariga moslashib qoladi va yangi ma'lumotlarda yomon natija beradi. Aniqlikni oshirish usullari. Ma'lumotlarni oldindan qayta ishlash Standartlashtirish va normallashtirish regressiya modelining barqarorligini oshiradi.

Regularizatsiya usullari. Ridge regressiya L2-normani minimallashtirib, koeffitsiyentlarni kichraytiradi. Lasso regressiya L1-norma yordamida ayrim koeffitsiyentlarni nolga tenglashtiradi. Modelni baholash Model sifati MSE, MAE va R^2 mezonlari orqali baholanadi. Tajriba va natijalar (Results) Sun'iy yaratilgan ma'lumotlar ustida o'tkazilgan tajribalar natijasida Ridge va Lasso regressiya usullari oddiy OLS regressiyaga nisbatan yuqori aniqlik ko'rsatdi. Lasso regressiya modeli eng yuqori R^2 qiymatiga ega bo'ldi. Muhokama (Discussion) Natijalar shuni ko'rsatadiki, regressiya modellarining aniqligini oshirish uchun yagona usul yetarli emas. Eng yaxshi natija ma'lumotlarni to'g'ri tayyorlash va zamonaviy statistik usullarni kompleks qo'llash orqali erishiladi.

XULOSA. Mazkur tadqiqot chiziqli regressiya modellarining aniqligini oshirishda regularizatsiya usullari va ma'lumotlarni oldindan qayta ishlash muhim ahamiyatga ega ekanligini ko'rsatdi. Ushbu yondashuvlar real amaliy masalalarda yuqori sifatli bashoratlar olish imkonini beradi.

FOYDALANILGAN ADABIYOTLAR

1. References Montgomery D.C., Peck E.A., Vining G.G. Introduction to Linear Regression Analysis. 5th Edition. Wiley, 2012. Kutner M.H., Nachtsheim C.J., Neter J. Applied Linear Statistical Models. 5th Edition. McGraw-Hill, 2005. Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd Edition. Springer, 2009.
2. Draper N.R., Smith H. Applied Regression Analysis. 3rd Edition. Wiley, 1998. James G., Witten D., Hastie T., Tibshirani R. An Introduction to Statistical Learning: With Applications in R. Springer, 2013. Seber G.A.F., Lee A.J. Linear Regression Analysis. 2nd Edition. Wiley, 2012. Vittinghoff E., Glidden D.V., Shiboski S.C., McCulloch C.E. Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models. Springer, 2012.
3. Faraway J.J. Linear Models with R. 2nd Edition. Chapman & Hall/CRC, 2014. Harrell F.E. Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis. Springer, 2015. Wooldridge J.M. Introductory Econometrics: A Modern Approach. 7th Edition. Cengage, 2016.