KO‘PFAKTORLI REGRESSIYA MODELINING ISHLASH PRINSIPI VA AMALIY QO‘LLANILISHI

Xolmirzayeva O‘g‘ilshod O‘ktam qizi
Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universiteti
Raqamli texnologiyalar konvergensiyasi kafedrasi magistranti
Toshkent, O‘zbekiston
ogilshodxolmirzayeva@gmail.com

Annotatsiya. *Ushbu tezisda ko‘pfaktorli regressiya modelining mohiyati, ishlash prinsipi va amaliy masalalarni yechishda qo‘llash mexanizmi ko‘rib chiqiladi. Ko‘pfaktorli regressiya biror natijaviy ko‘rsatkichga bir vaqtning o‘zida ta’sir etuvchi bir nechta omillar o‘rtasidagi bog‘liqlikni matematik jihatdan ifodalash imkonini beradi. Tadqiqotda regressiya modelining umumiy ko‘rinishi, regressiya koeffitsiyentlarini aniqlash, eng kichik kvadratlar usuli, modelning statistik ahamiyatini baholash hamda prognozlash jarayoni yoritiladi. Shuningdek, determinatsiya koeffitsiyenti, o‘rtacha absolyut xatolik va ildiz-o‘rtacha-kvadrat xatolik kabi mezonlarning model sifatini baholashdagi ahamiyati ko‘rsatiladi. Ko‘pfaktorli regressiya usuli iqtisodiyot, texnologiya, ishlab chiqarish, ekologiya, energetika va boshqa ko‘plab sohalarda ma’lumotlar asosida qaror qabul qilish uchun samarali matematik vosita hisoblanadi.*

Kalit so‘zlar: *ko‘pfaktorli regressiya, regressiya modeli, statistik tahlil, regressiya koeffitsiyenti, eng kichik kvadratlar usuli, prognozlash, determinatsiya koeffitsiyenti, statistik model.*

Аннотация. *В данной тезисной работе рассматриваются сущность многофакторной регрессионной модели, принцип её функционирования и механизм применения для решения практических задач. Многофакторная регрессия позволяет математически описать взаимосвязь между несколькими факторами, одновременно влияющими на результирующий показатель. В исследовании рассматриваются общий вид регрессионной модели, определение коэффициентов регрессии, метод наименьших квадратов, оценка статистической значимости модели и процесс прогнозирования. Кроме того, раскрывается значение таких критериев, как коэффициент детерминации, средняя абсолютная ошибка и корень из средней квадратичной ошибки, для оценки качества модели. Метод многофакторной регрессии является эффективным ма-*

тематическим инструментом для принятия решений на основе данных в экономике, технологиях, производстве, экологии, энергетике и многих других областях.

Ключевые слова: *многофакторная регрессия, регрессионная модель, статистический анализ, коэффициент регрессии, метод наименьших квадратов, прогнозирование, коэффициент детерминации, статистическая модель.*

Abstract. *This paper examines the concept of a multiple regression model, its operating principles, and its application mechanisms for solving practical problems. Multiple regression makes it possible to mathematically describe the relationships between several factors that simultaneously influence a dependent variable. The study discusses the general form of a regression model, the estimation of regression coefficients, the least squares method, the assessment of statistical significance, and the forecasting process. Furthermore, it highlights the importance of evaluation metrics such as the coefficient of determination, mean absolute error, and root mean square error in assessing model performance. Multiple regression is an effective mathematical tool for data-driven decision-making in economics, technology, manufacturing, environmental science, energy, and many other fields.*

Keywords: *multiple regression, regression model, statistical analysis, regression coefficient, least squares method, forecasting, coefficient of determination, statistical model.*

Kirish

Zamonaviy axborot texnologiyalarining rivojlanishi natijasida turli jarayonlar haqida katta hajmdagi ma’lumotlarni yig‘ish va tahlil qilish imkoniyati kengaymoqda. Bunday ma’lumotlardan foydalanishda asosiy vazifalardan biri natijaviy ko‘rsatkichga ta’sir qiluvchi omillarni aniqlash va ular o‘rtasidagi bog‘liqlikni matematik jihatdan ifodalashdan iborat. Mazkur vazifani hal qilishda statistik modellashtirish usullaridan, xususan, regressiya tahlilidan keng foydalaniladi.

Oddiy regressiya modelida bitta natijaviy o‘zgaruvchi va bitta mustaqil o‘zgaruvchi o‘rtasidagi bog‘liqlik o‘rganiladi. Biroq real hayotdagi ko‘plab jarayonlarda natijaga bir vaqtning o‘zida bir nechta omillar ta’sir ko‘rsatadi. Masalan, mahsulotga bo‘lgan talab narx, reklama xarajatlari, iste’molchilar daromadi va mavsumiylik kabi bir nechta omillarga bog‘liq bo‘lishi mumkin. Shu sababli bunday jarayonlarni tahlil qilishda ko‘pfaktorli regressiya modelidan foydalanish maqsadga muvofiqdir.

Asosiy qism

Ko‘pfaktorli regressiya – natijaviy, ya’ni bog‘liq o‘zgaruvchi bilan ikki yoki undan ortiq mustaqil o‘zgaruvchilar o‘rtasidagi statistik bog‘liqlikni aniqlashga xizmat qiluvchi matematik-statistik modeldir. Modelning umumiy ko‘rinishi quyidagicha ifodalanadi:

$Y=\beta_0+\beta_1X_1+\beta_2X_2+\ldots+\beta_kX_k+\varepsilon$

bu yerda Y – natijaviy o‘zgaruvchi, $X_1,X_2,\ldots,X_k$ – natijaga ta’sir qiluvchi mustaqil omillar, β_0 – ozod had, $\beta_1,\beta_2,\ldots,\beta_k$ – regressiya koeffitsiyentlari, ε esa model tomonidan hisobga olinmagan tasodifiy xatolikni bildiradi.

Ko‘pfaktorli regressiyaning asosiy g‘oyasi shundan iboratki, natijaviy ko‘rsatkichni unga ta’sir qiluvchi omillarning chiziqli kombinatsiyasi orqali tushuntirish mumkin. Har bir regressiya koeffitsiyenti tegishli omilning natijaviy o‘zgaruvchiga ta’sirini ifodalaydi. Masalan, β_1 koeffitsiyenti X_1 bir birlikka o‘zgarganda, boshqa omillar o‘zgarmagan sharoitda Y ning o‘rtacha qancha miqdorga o‘zgarishini ko‘rsatadi. Shu jihatdan ko‘pfaktorli regressiya nafaqat prognozlash, balki omillarning natijaga ta’sirini tahlil qilish uchun ham ishlatiladi.

Modelni qurish jarayoni bir necha ketma-ket bosqichlardan tashkil topadi. Birinchi bosqichda tadqiqot obyekti aniqlanadi va natijaviy ko‘rsatkich Y belgilanadi. Keyingi bosqichda ushbu ko‘rsatkichga ta’sir qilishi mumkin bo‘lgan $X_1,X_2,\ldots,X_k$ omillar tanlanadi. Omillarni tanlashda ularning nazariy jihatdan asoslanganligi, ma’lumotlarning mavjudligi va natijaviy ko‘rsatkich bilan bog‘liqligi hisobga olinadi.

Keyingi bosqichda kuzatuvlar asosida regressiya koeffitsiyentlari hisoblanadi. Bunda eng kichik kvadratlar usuli keng qo‘llaniladi. Ushbu usulning asosiy maqsadi haqiqiy kuzatilgan qiymatlar bilan model tomonidan hisoblangan qiymatlar o‘rtasidagi farqlar kvadratlari yig‘indisini minimallashtirishdan iborat:

$S=\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2$

bu yerda Y_i – kuzatilgan haqiqiy qiymat $\hat{Y}_i$– regressiya modeli yordamida hisoblangan qiymat, n – kuzatuvlar soni.

Eng kichik kvadratlar usuli yordamida shunday regressiya koeffitsiyentlari aniqlanadiki, natijada model qiymatlari va haqiqiy qiymatlar o‘rtasidagi umumiy kvadratik farq eng kichik bo‘ladi. Ko‘pfaktorli regressiyaning matritsaviy ko‘rinishida koeffitsiyentlar quyidagi formula orqali aniqlanishi mumkin:

$\hat{\beta}=(X^TX)^{-1}X^TY$

Bu yerda X – mustaqil o‘zga ruvchilar matritsasi, Y – kuzatilgan natijalar vektori, $\hat{\beta}$ – baholangan regressiya koeffitsiyentlari vektori.

Model qurilgandan so‘ng uning qanchalik yaxshi ishlashini baholash zarur. Buning uchun bir nechta statistik mezonlardan foydalaniladi. Eng muhim ko‘rsatkichlardan biri determinatsiya koeffitsiyenti R^2 hisoblanadi. U natijaviy

o‘zgaruvchidagi o‘zgarishlarning qanchalik katta qismi regressiya modeli tarkibidagi omillar orqali tushuntirilishini ko‘rsatadi. Uning qiymati odatda 0 dan 1 gacha bo‘ladi. R2R^2 qiymatining yuqoriligi modelning ma’lumotlarga mosligini ko‘rsatishi mumkin, biroq faqat ushbu ko‘rsatkich asosida modelning mutlaq sifatiga xulosa qilish yetarli emas.

Prognozlash aniqligini baholashda *MAE* va *RMSE* kabi xatolik mezonlaridan ham foydalaniladi. O‘rtacha absolyut xatolik quyidagicha aniqlanadi:

$$MAE=\frac{1}{n}\sum_{i=1}^{n} Y_i - \hat{Y}_i \vee$$

RMSE esa:

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}$$

ko‘rinishida hisoblanadi. Ushbu ko‘rsatkichlarning kichik qiymatlari model prognozlarining kuzatilgan qiymatlarga yaqinligini anglatadi.

Ko‘pfaktorli regressiya modelini qo‘llashda omillar o‘rtasidagi kuchli o‘zaro bog‘liqlik, ya’ni multikollinearlik masalasiga ham e’tibor berish kerak. Agar mustaqil o‘zgaruvchilar bir-biri bilan juda kuchli bog‘langan bo‘lsa, regressiya koeffitsiyentlarini ishonchli baholash qiyinlashishi mumkin. Bunday holatda korrelyatsiya matritsasi, VIF (Variance Inflation Factor) kabi statistik vositalardan foydalanish orqali omillar o‘rtasidagi bog‘liqlik tekshiriladi.

Shuningdek, regressiya modelining qoldiqlari ham tahlil qilinadi. Qoldiq – haqiqiy qiymat bilan model tomonidan hisoblangan qiymat o‘rtasidagi farq bo‘lib,

$$e_i=Y_i - \hat{Y}_i$$

formula orqali aniqlanadi. Qoldiqlarni tahlil qilish modelning asosiy statistik farazlari bajarilishini tekshirishga yordam beradi. Jumladan, qoldiqlarning mustaqilligi, dispersiyaning doimiyligi va taqsimot xususiyatlari modelning ishonchliligini baholashda muhim ahamiyatga ega.

Umuman olganda, ko‘pfaktorli regressiya modeli quyidagi mantiq asosida ishlaydi: avvalo natijaviy ko‘rsatkich va unga ta’sir qiluvchi omillar bo‘yicha ma’lumotlar yig‘iladi; keyin ma’lumotlar statistik jihatdan tayyorlanadi; regressiya koeffitsiyentlari hisoblanadi; modelning mosligi va xatoliklari baholanadi; shundan so‘ng model yangi ma’lumotlar uchun prognoz qiymatlarini hisoblashda ishlatiladi. Shu sababli ushbu usul katta hajmdagi ma’lumotlar bilan ishlashda jarayonlarni matematik ifodalash va kelajakdagi qiymatlarni prognozlashning muhim vositalaridan biri hisoblanadi.

Xulosa

Ko‘pfaktorli regressiya bir nechta mustaqil omillarning natijaviy ko‘rsatkichga ta’sirini bir vaqtning o‘zida baholash imkonini beruvchi samarali statistik modellashtirish usulidir. Uning asosida regressiya tenglamasini shakllantirish,

koeffitsiyentlarni eng kichik kvadratlar usuli yordamida aniqlash va model natijalarini statistik mezonlar orqali baholash jarayonlari yotadi. R^2, *MAE* va *RMSE* kabi ko‘rsatkichlardan foydalanish modelning ma’lumotlarga mosligi va prognozlash xatoligini baholash imkonini beradi. To‘g‘ri tanlangan omillar va sifatli ma’lumotlar asosida qurilgan ko‘pfaktorli regressiya modeli turli amaliy masalalarni tahlil qilish, mavjud jarayonlarni tushuntirish va kelajakdagi qiymatlarni prognozlashda qo‘llanilishi mumkin.

Foydalanilgan adabiyotlar

1. Montgomery D.C., Peck E.A., Vining G.G. *Introduction to Linear Regression Analysis*. Wiley, 2021.

2. James G., Witten D., Hastie T., Tibshirani R. *An Introduction to Statistical Learning*. Springer, 2021.

3. Wooldridge J.M. *Introductory Econometrics: A Modern Approach*. Cengage Learning, 2020.

4. Kutner M.H., Nachtsheim C.J., Neter J., Li W. *Applied Linear Statistical Models*. McGraw-Hill, 2005.

5. Gujarati D.N., Porter D.C. *Basic Econometrics*. McGraw-Hill, 2009.