MASHINALI O'QITISHDA GRADIENTLI TUSHUSH ALGORETMINI CHIZIQLI REGRESSIYA YORDAMIDA TADQIQ QILISH

Oromov Dilmurod Olim o‘g‘li

Sharof Rashidov nomidagi Samarqand davlat universiteti Intellektual tizimlar va kompyuter texnologiyalari fakulteti “Kompyuter ilmlari va texnologiyalari fanlari” kafedrasi talabasi

https://doi.org/10.5281/zenodo.12197415

Annotatsiya. *Ushbu maqolada, chiziqli regressiya modelini qurishda gradient tushish usulining qo’llanilishini tadqiq qilish uchun matematik analiz, ehtimollar nazaryasi va matematik statistika, mashinaviy o’qitish hamda python dasturlash tili kurslarining ba’zi tushuncha, usul va yondoshuvlaridan foydalanilgan.*

Kalit so’zlar: *chiziqli regressiya modelini qurish, matematik analiz, ehtimollar nazaryasi, matematik statistika.*

AN INVESTIGATION OF THE GRADIENT DESCENT ALGORITHM IN MACHINE LEARNING USING LINEAR REGRESSION

Abstract. *This paper uses some concepts, methods and approaches from mathematical analysis, probability theory and mathematical statistics, machine learning and Python programming language courses to investigate the application of gradient descent in linear regression model building.*

Key words: *building a linear regression model, mathematical analysis, probability theory, mathematical statistics.*

ИССЛЕДОВАНИЕ АЛГОРИТМА ГРАДИЕНТНОГО СПУСКА В МАШИННОМ ОБУЧЕНИИ С ИСПОЛЬЗОВАНИЕМ ЛИНЕЙНОЙ РЕГРЕССИИ

Аннотация. *В данной статье некоторые концепции, методы и подходы математического анализа, теории вероятностей и математической статистики, машинного обучения и курсов языка программирования Python используются для исследования применения метода градиентного спуска при построении модели линейной регрессии.*

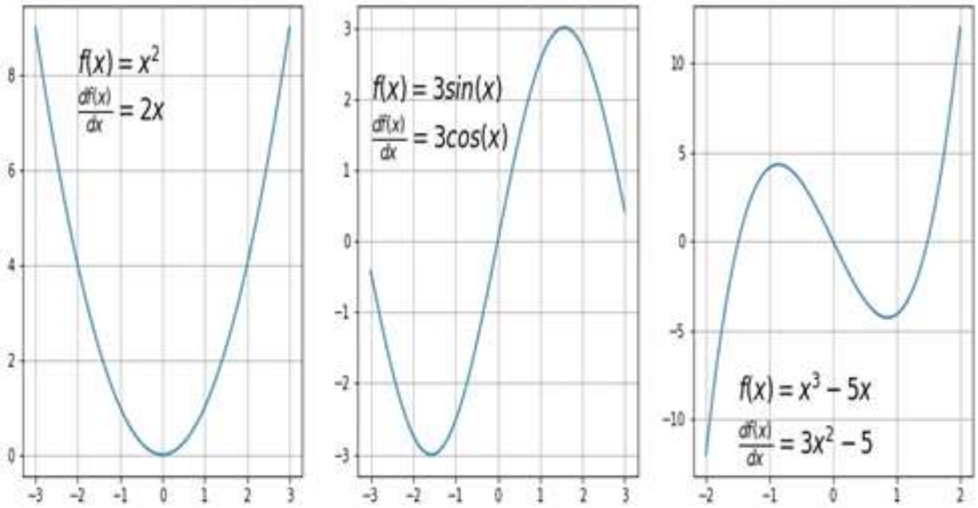
Ключевые слова: *построение модели линейной регрессии, математический анализ, теория вероятностей, математическая статистика.*

Bugunguni kunda juda ko’plab sohalar (tavsiya tizimlari, tasvirlarni aniqlash, tabiiy tilni qayta ishlash va moliyaviy bashoratlar) dagi turli xil hayotiy masalalarni yechishda ularga oid tarixiy ma’lumotlar, mashinali o’qitish algoritmlari, jumladan chiziqli regressia algoritmi orqali tahlil qilinib, eng yaxshi yechim, tavsiya, bashorat yoki qarorlar qabul qilinmoqda. Chiziqli regressiya algoritmi bu nazoratli o'qitish algoritmi bo'lib, u kirish xususiyatlari va maqsadli o'zgaruvchi o'rtasidagi chiziqli munosabatni aniqlashdan iborat. Chiziqli resgressiya natijalarining aniqligi, chiziqli munosabat koeffisentlarining aniqligiga bo'gliq bo'lib, u odatda gradient tushish usullari orqali aniqlanadi. Shu nuqtai nazardan chiziqli regressiya modelini qurishda gradient tushish usulini tadqiq qilish va bu usulga asoslangan dasturiy vosita ishlab chiqish muhim hisoblanadi.

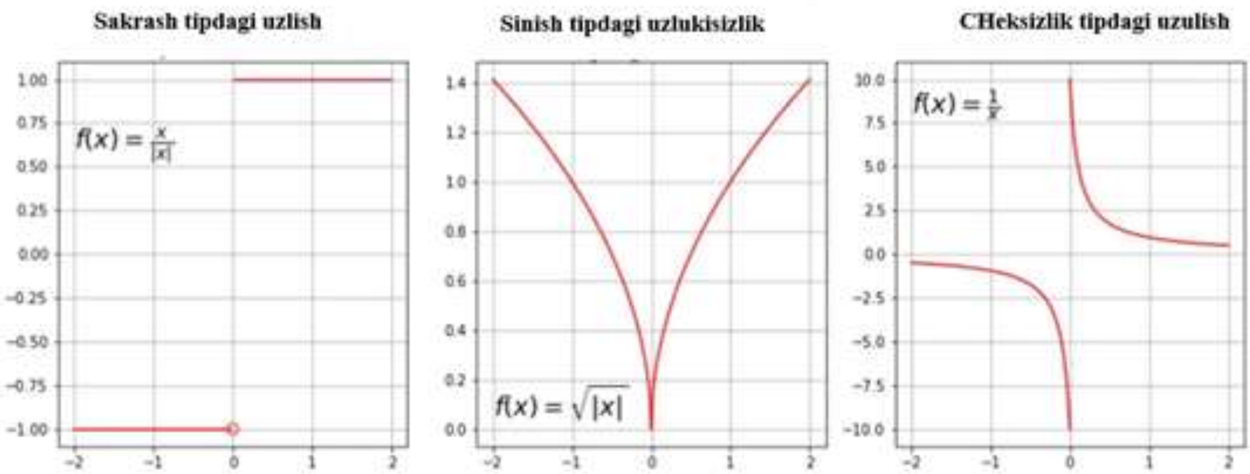
Mazkur ilmiy maqoladan olingan ilmiy natijalar amaliy-uslubiy xarakterga ega bo'lib, ishda ko'p o'zgaruvchili qavariq funksiyalarning ekstremumlari (maksimum yoki minimum) ni topish uchun gradient tushish usuliga asoslangan dasturiy ta'minot yaratligan. Bu dasturiy ta'minot Python dasturlash tilida yaratilgan bo'lib u vizuallashgan qulay interfeysga ega.

Gradient tushish usuli bu iterativ birinchi darajali optimallashtirish algoritmi bo'lib, berilgan funksiyaning lokal minimum va maksimumlani topish uchun ishlatiladi. Ushbu usul odatda mashinali o'qitishda va chuqur o'qitishda maqsad/yo'qotish (loss function) funktsiyasini minimallashtirish uchun qo'llaniladi. Muhimligi va amalga oshirish qulayligi tufayli ushbu algoritm odatda deyarli barcha mashinali o'qitiish kurslarining boshida o'rganiladi. Biroq, undan foydalanish faqat mashimali o'qitish yoki chuqur o'qitish bilan cheklanmaydi, u quyidagi sohalarda ham keng qo'llaniladi: boshqaruv muhandisligi (robotexnika, kimyo va boshqalar) kompyuter o'yinlari, mashinasozlik va h.k.. Shuning uchun bu ishda biz birinchi darajali gradient tushish algoritmining matematikasi, amalga oshirilishi va xatti-harakati bilan chuqur tanishamiz. Biz to'g'ridan-to'g'ri maxsus maqsad funktsiyasini ko'rib chiqamiz va uning minimalini topamiz. Bu shuni anglatadiki, odatiy mashinali o'qitish darslarida bo'lgani kabi asosiy ma'lumotlar bo'lmaydi - biz funktsiya shakliga nisbatan moslashuvchanroq bo'lamiz. Bu usul zamonaviy kompyuterlar davridan ancha oldin 1847 yilda Avgustin-Lui Koshi tomonidan taklif qilingan. O'sha vaqtdan boshlab informatika va raqamli usullarda sezilarli rivojlanish bo'ldi. Bu esa gradient tushish usulining ko'plab takomillashtirilgan variantlari paydo bo'lishiga olib keldi. Biroq, ushbu ishda biz klassik gradient tushish usuli va unig chiziqli regressiyadagi tadbiqini o'rganish bilan chegaralanamiz.

Funksiyaga qo'yilgan talablar. Gradient tushish algoritmi barcha funktsiyalar uchun ishlamaydi. Ikkita asosiy maxsus talab mavjud. Funktsiya quyidagicha bo'lishi kerak: differensiallanuvchi va qavariq. Bu talablarni alohida-alohida ko'rib chiqamiz. Differensiallanuvchilik talabi, agar funktsiya differentsial bo'lsa, u o'z aniqlanish sohasidagi har bir nuqta uchun hosilaga ega - barcha funktsiyalar bu mezonlarga javob bermaydi. Birinchidan, ushbu mezonga javob beradigan ba'zi funktsiyalar misollarini ko'rib chiqaylik



Rasm 1. Uzluksiz differensiallanuvchi funksiyalar

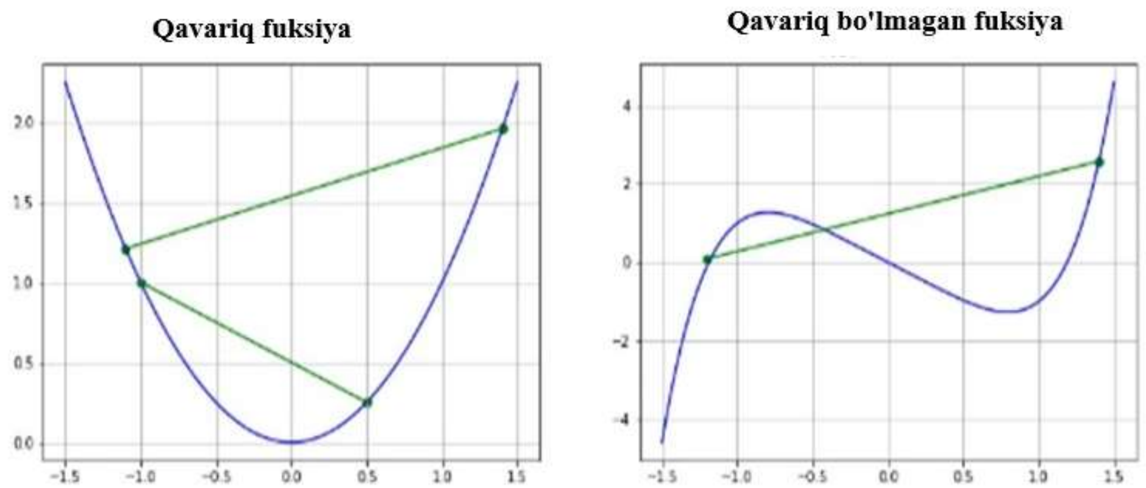


Rasm 2. Differensiallanuvchi bo'lmagan funksiyalar

Keyingi talab - funktsiya qavariq bo'lishi talabi. Bir o'zgaruvchan funktsiya uchun bu ikkita funktsiya nuqtasini bog'laydigan chiziq segmenti uning egri chizig'ida yoki uning ustida joylashganligini anglatadi (uni kesib o'tmaydi). Agar shunday bo'lsa, bu global bo'lmagan mahalliy minimumga ega ekanligini anglatadi. Matematik jihatdan funktsiya egri chizig'ida joylashgan ikkita x_1, x_2 nuqtalar uchun bu shart quyidagicha ifodalanadi:

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2)$$

Bu erda l nuqtaning kesma chizig'idagi joylashishini bildiradi va uning qiymati 0 (chap nuqta) va 1 (o'ng nuqta) orasida bo'lishi kerak, masalan. l=0,5 o'rtadagi joyni bildiradi. Quyida namunali bo'lim chiziqlari bilan ikkita funksiya mavjud.



Rasm 3. Qavariq va qavariq bo'lmagan funksiyalar

Bir o'zgaruvchan funksiya qavariq ekanligini matematik tarzda tekshirishning yana bir usuli ikkinchi hosilani hisoblash va uning qiymati noldan katta yoki yo'qligini tekshirishdir.

$$\frac{d^2f(x)}{dx^2} > 0$$

Keling, oddiy kvadrat funksiyani tekshirib ko'raylik:

$$f(x) = x^2 - x + 3$$

Uning birinchi va ikkinchi hosilalari quyidagilarga teng:

$$\frac{df(x)}{dx} = 2x - 1, \qquad \frac{d^2f(x)}{dx^2} = 2$$

Ikkinchi hosila har doim 0 dan katta bo'lgani uchun bizning funksiyamiz qat'iy ravishda qavariqdir. Shuningdek, gradientli algoritm bilan kvasi-konvex funksiyalardan foydalanish mumkin. Biroq, ko'pincha ular algoritm to'xtab qolishi mumkin bo'lgan taxminiy joylarga ega bo'ladi. Quasi-konveks funksiyaga misol.

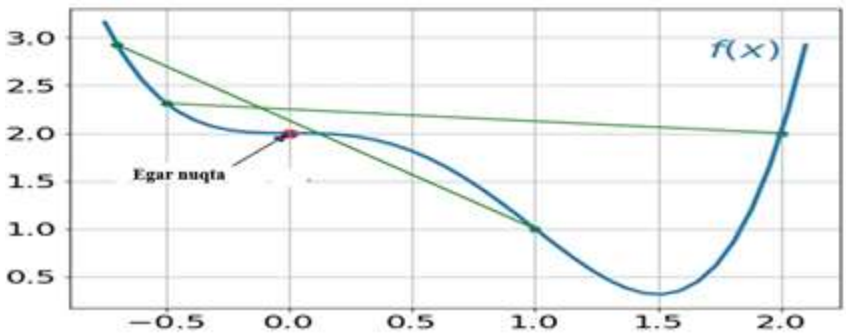
$$f(x) = x^4 - 2x^3 + 2$$

$$\frac{df(x)}{dx} = 4x^3 - 6x^2 = x^2(4x - 6)$$

Ko'rinib turibdiki, birinchi hosila *x=0* va *x=1.5* da nolga teng. Bu joylar funksiya uchun nomzodlar hisoblanadi (minimal yoki maksimal) burchak burchagi nol. Lekin avval ikkinchi hosilani tekshirib ko'rishimiz kerak.

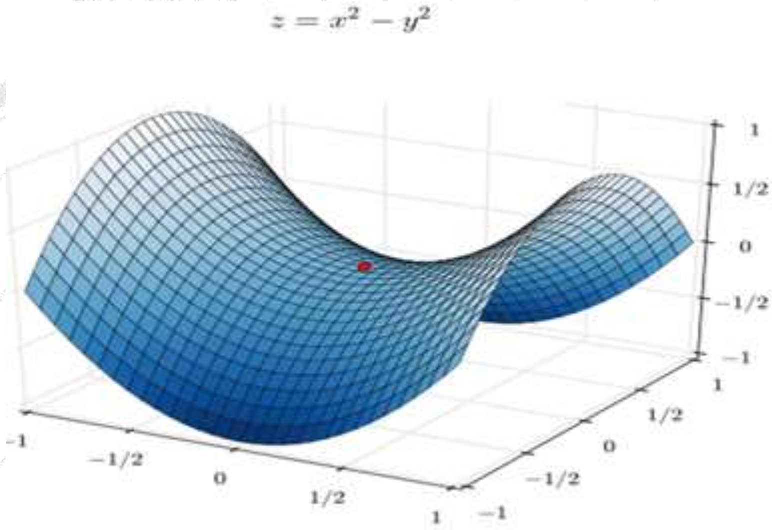
$$\frac{d^2f(x)}{dx^2} = 12x^2 + 12x = 12x(x - 1)$$

Ushbu ifodaning qiymati *x=0* va *x=1* uchun nolga teng. Ushbu joylar egrilik nuqtasi deb ataladi, egrilik belgisi o'zgarib turadigan joy bu qavariqlikdan botiqlikga yoki aksinchaga o'zgarishini anglatadi. Ushbu tenglamani tahlil qilish orqali quyidagi xulosaga kelamiz: *x<0* bo'lganda funksiya qavariq bo'ladi,*0<x<1* bo'lganda funksiya botiq bo'ladi *x>1* bo'lganda funksiya yana qavariq bo'ladi. Endi biz *x=0* nuqtada funksiyaning birinchi va ikkinchi hosilalari ham nolga teng ekanligini ko'ramiz, ya'ni bu egar nuqta bo'ladi va *x=1.5* nuqtada funksiyaning birinchi hosilasi nolga teng va ikkinchi hosilasi musbat shuning uchun bu global minimum nuqta, bu nuqtada funksiya global minimumga erishadi. Keling, ushbu funksiyaning grafigiga qaraylik. Egar nuqta *x=0* va minimum nuqta *x=1,5* ekanligini grafikdan ko'rishimiz mumkin.



Rasm 4. Egar va minimum nuqta

Ko'p o'zgaruvchili funksiyalar uchun nuqta egar nuqtasi bo'ladimi yoki yo'qligini tekshirishning eng mos yo'li bu biroz murakkab hisob-kitoblarni o'z ichiga oldi va bu Hessa matritsasini hisoblashni o'z ichiga oladi. Ikki o'zgaruvchili funksiya uchun egar nuqta misoli quyidagi ko'rinishda bo'ladi.



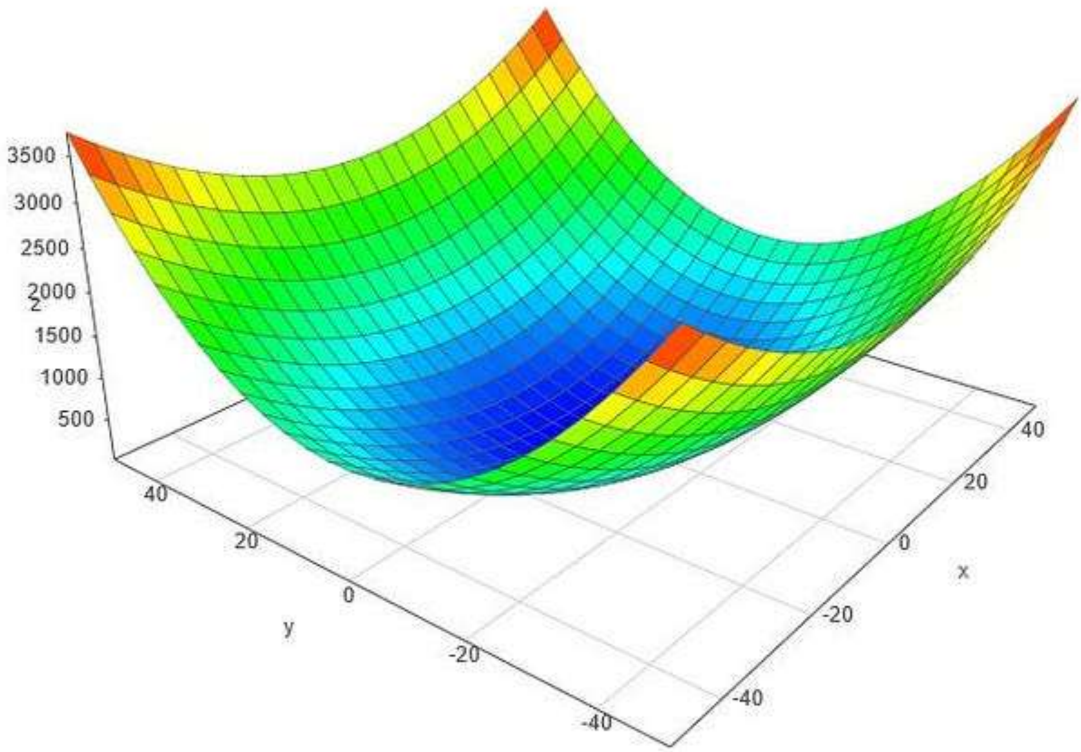
Rasm 5. Ikki o'zgaruvchili funksiyaning egar nuqtasi

Gradient. Algoritmga o'tishdan oldin yana bir narsani tushuntirish kerak - gradient nima. Intuitiv ravishda bu ma'lum bir yo'nalishdagi ma'lum nuqtada egri chiziqning qiyaligi (kamayib borishi). Bir o'zgaruvchili funksiya bo'lsa, u tanlangan nuqtadagi birinchi hosiladir. Ko'p o'zgaruvchan funktsiya bo'lsa, u har bir asosiy yo'nalishdagi (o'zgaruvchan o'qlar bo'ylab) hosilalar vektoridir. Bizni faqat bitta o'q bo'ylab qiyalik qiziqtiradi va boshqalarga ahamiyat bermaymiz, bu hosilalar qisman hosilalar deb ataladi. Berilgan *p* nuqtadagi *n* o'lchamli *f(x)* funksiya uchun gradient quyidagicha aniqlanadi:

$$\nabla f(p) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(p) \\ \vdots \\ \frac{\partial f}{\partial x_n}(P) \end{bmatrix}$$

Teskari uchburchak nabla deb ataladigan belgidir va siz uni "del" deb o'qiysiz. Uni qanday hisoblashni yaxshiroq tushunish uchun quyidagi ikki o'zgaruvchili funksiya uchun qo'lda hisob-kitob qilaylik.

$$f(x) = 0.5x^2 + y^2$$



Rasm 6. $f(x) = 0.5x^2 + y^2$ *funksiyaning 3D grafigi*

Faraz qilaylik, bizni p(10,10) nuqtadagi gradient qiziqtiradi:

$$\frac{\partial f(x,y)}{\partial_x} = x, \qquad \frac{\partial f(x,y)}{\partial_y} = 2y$$

shuning uchun

$$\nabla f(x,y) = \begin{bmatrix} x \\ 2y \end{bmatrix}, \qquad \nabla f(10,10) = \begin{bmatrix} 10 \\ 20 \end{bmatrix}$$

Ushbu qiymatlarga qarab, biz qiyalik *y* o'qi bo'ylab ikki baravar tikroq degan xulosaga kelamiz.

REFERENCES

1. Seber, G. A. F., & Lee, A. J. (2012). *Linear Regression Analysis* (2nd ed.). Wiley. ISBN: 978-1-118-28342-5.
2. Draper, N. R., & Smith, H. (1998). *Applied Regression Analysis* (3rd ed.). Wiley. ISBN: 978-0-471-17082-2.
3. Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to Linear Regression Analysis* (5th ed.). Wiley. ISBN: 978-1-118-38500-5.
4. Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). Springer. ISBN: 978-0-387-84858-7.
5. Kutner, M. H., Nachtsheim, C. J., & Neter, J. (2004). *Applied Linear Regression Models* (4th ed.). McGraw-Hill/Irwin. ISBN: 978-0-07-301344-2.
6. Weisberg, S. (2013). *Applied Linear Regression* (4th ed.). Wiley. ISBN: 978-1-118-38608-8.
7. Ruder, S. (2017). An Overview of Gradient Descent Optimization Algorithms. *arXiv preprint arXiv:1609.04747*.