

TO'G'RI GRADIENTLI OPTIMIZATSIYA ALGORITMI

Tojimamatov Israil Nurmamatovich*Farg'ona davlat universiteti Amaliy matematika va**informatika kafedrasi katta o'qituvchisi**israiltojiamamatov@gmail.com***Surayyo Turg'unova Ulug'bek qizi***surayyoturgunova16@gmail.com*

Annotatsiya. Ushbu maqolada gradientga asoslangan optimallashtirish usullarining eng asosiy va keng qo'llaniladigan turi — **to'g'ri gradientli optimallashtirish (Gradient Descent Optimization)** algoritmi batafsil yoritiladi. Algoritmning matematik mohiyati, uning ishlash tamoyillari, gradientni hisoblash mexanizmi hamda learning rate parametrining o'qitish jarayoniga ta'siri chuqur tahlil qilindi. Shu bilan birga, mashinaviy o'rganish va chuqur o'rganish modellarida parametrlar (vaznlar va biaslar)ni yangilashning nazariy asoslari ko'rib chiqiladi. Maqolada Gradient Descent algoritmining turli variantlari — Batch Gradient Descent, Stochastic Gradient Descent (SGD), Momentum, RMSProp va Adam optimallashtirish usullarining ishlash prinsiplari va ular o'rtasidagi asosiy farqlar tahlil qilindi. Gradient tushish algoritmi mashina va chuqur o'rganish tizimlarining markaziy komponenti bo'lib, neyron tarmoqlarni samarali o'qitishda muhim rol o'ynaydi. Ushbu maqola algoritmnı nazariy tushunishga yordam beradi, amaliy modellarni tuzishda va mavjud modellarni optimallashtirishda qo'llaniladigan asosiy tushunchalarni kengaytiradi.

Kalit so'zlar: Gradient descent, optimallashtirish algoritmi, learning rate, neyron tarmoq, SGD, Adam, RMSProp, yo'qotish funksiyasi, gradient, mashinaviy o'rganish.

Abstract. This article details the most basic and widely used type of gradient-based optimization method - the direct gradient optimization algorithm (Gradient Descent Optimization). The mathematical essence of the algorithm, its operational principles, the mechanism for calculating the gradient, and the impact of the learning rate parameter on the learning process were thoroughly analyzed. At the same time, the theoretical foundations of parameter updates (weights and biases) in machine learning and deep learning models are considered. The article analyzes the operating principles of various variants of the Gradient Descent algorithm - Batch Gradient Descent, Stochastic Gradient Descent (SGD), Momentum, RMSProp, and Adam optimization methods, and the main differences between them. The gradient descent algorithm is a crucial component of machine learning and deep learning systems, playing a vital role in the effective training of neural networks. This article helps to

understand the algorithm theoretically, expands the basic concepts used in constructing applied models, and optimizes existing models.

Keywords: Gradient descent, optimization algorithm, learning rate, neural network, SGD, Adam, RMSProp, loss function, gradient, machine learning.

Аннотация. В данной статье подробно рассматривается основной и широко используемый тип методов градиентной оптимизации - алгоритм прямой градиентной оптимизации (Gradient Descent Optimization). Глубоко проанализированы математическая сущность алгоритма, принципы его работы, механизм расчета градиента и влияние параметра learning rate на процесс обучения. В то же время рассматриваются теоретические основы обновления параметров (весов и ошибок) в моделях машинного обучения и глубокого обучения. В статье проанализированы принципы работы различных вариантов алгоритма Gradient Descent - Batch Gradient Descent, Stochastic Gradient Descent (SGD), Momentum, RMSProp и Adam, а также основные различия между ними. Алгоритм градиентного спуска является центральным компонентом систем машинного и глубокого обучения и играет важную роль в эффективном обучении нейронных сетей. Эта статья помогает теоретически понять алгоритм, расширяет основные понятия, используемые при построении практических моделей и оптимизации существующих моделей.

Ключевые слова: Gradient descent, алгоритм оптимизации, learning rate, нейронная сеть, SGD, Adam, RMSProp, функция потерь, градиент, машинное обучение.

Kirish

Sun'iy intellekt, mashinaviy o'rganish va chuqur o'rganish texnologiyalari tez sur'atlar bilan rivojlanayotgan bir davrda optimallashtirish algoritmlarining ahamiyati yanada ortib bormoqda. Mashinaviy o'rganish modelining asosiy maqsadi — tanlangan yo'qotish funksiyasi qiymatini minimal darajaga tushirish orqali modelni yanada aniqroq bashoratlarni berishga moslashtirishdir. Bunda model parametrlarini to'g'ri yangilab berish jarayoni muhim o'rin tutadi. Aynan shu jarayon gradientga asoslangan optimallashtirish algoritmlarining vazifasidir. Gradient Descent algoritmi — optimallashtirishning eng sodda, tushunarli, shu bilan birga juda samarali usullaridan biri bo'lib, u ko'p o'zgaruvchili murakkab funksiyalarning minimal nuqtasiga yo'nalgan holda harakat qiladi. Model parametrlarining optimal to'plamini topish uchun gradient tushish yo'li bilan iterativ yangilash amalga oshiriladi. Har bir qadamda yo'qotish funksiyasining gradienti hisoblanib, funksiya qiymatini kamaytiruvchi yo'nalishda qadam tashlanadi. Gradient tushish algoritmi nafaqat matematik optimallashtirish masalalarida, balki amaliy sohalarda — kompyuter ko'rish, tabiiy tilni qayta ishlash, tasniflash, regressiya, tavsiya tizimlari, vaqt qatorlari

analizi kabi ko'plab yo'nalishlarda keng qo'llaniladi. Neyron tarmoqlarni o'qitish jarayonida gradientni teskari tarqalishi (Backpropagation) bilan birga ishlashi esa ushbu optimallashtirish usulining naqadar muhimligini yana bir bor isbotlaydi. Algoritmning muvaffaqiyatli ishlashi ko'plab omillarga bog'liq: learning rate'ning to'g'ri tanlanishi, gradientning to'g'ri hisoblanishi, funksiyaning shakli, mahalliy minimumlar mavjudligi va algoritmning varianti. Shu sababli turli vaziyatlarda Maxsus modifikatsiyalar — SGD, Momentum, RMSProp, Adam kabi optimallashtirish yondashuvlari qo'llaniladi.

Gradient tushish algoritmi funksiyaning minimal qiymatini topishga qaratilgan iterativ usul bo'lib, har bir qadamda funksiya yo'nalishidagi eng katta kamayish gradienti bo'yicha harakatlanadi. Agar funksiya ko'p o'zgaruvchili bo'lsa, gradient — bu barcha hosilalardan tashkil topgan vektordir. Algoritmning umumiy maqsadi yo'qotish funksiyasi $L(\theta)$ ni minimallashtirishdan iborat bo'lib, bunda θ model parametrlarini ifodalaydi. Optimizatsiya jarayoni quyidagi bosqichlardan tashkil topgan:

Birinchidan, boshlang'ich nuqta x_0 tasodifiy yoki oldindan belgilangan tarzda tanlanadi. Bu nuqta funksiyaning optimal qiymatidan uzoq bo'lishi mumkin va undan keyingi qadamlar shu nuqtadan boshlab amalga oshiriladi. Ikkinchidan, tanlangan nuqtadagi funksiya gradienti $\nabla L(x)$ hisoblanadi. Gradient — bu funksiya qiymatining vertikal o'zgarishi (dL) ning gorizontaal o'zgarishiga (dW) nisbati bo'lib, uning yordamida funksiya qanday o'zgarayotganligi, ya'ni eng steplik kamayish yo'nalishi aniqlanadi. Tik qiyaliklarda gradient qiymati katta bo'lib, yassi qiyaliklarda kichikroq bo'ladi.

Gradient hisoblash jarayoni quyidagi tenglama bilan ifodalanadi:

$$w_{t+1} = w_t - \alpha \frac{\partial L}{\partial w}$$

Bu yerda w_t — joriy vazn (parametr), w_{t+1} — yangilangan vazn, α — learning rate (o'rganish tezligi), $\partial L / \partial w$ esa vaznga nisbatan yo'qotish funksiyasining gradientidir. Learning rate — harakat qadamining uzunligini belgilovchi parametr bo'lib, u optimallashtirish jarayonining samaradorligini bevosita ta'sir qiladi.

Keyingi bosqichda, har bir iteratsiyada parametrlar yangilanadi. Yangilanish formulasi quyidagicha ifodalanadi:

$$x_{new} = x_{old} - \alpha \Delta L(x)$$

Learning rate juda muhim bo'lib, agar u juda katta bo'lsa, algoritm optimal nuqtani "sakrab o'tishi" mumkin; juda kichik bo'lsa esa optimal nuqtaga yetishish juda sekin kechadi. Shu sababli, amaliyotda adaptiv learning ratega ega algoritmlar qo'llaniladi. Yakuniy bosqichda, parametrlar yangilanish jarayoni va gradient hisoblash iterativ tarzda takrorlanadi. Funksiya qiymati optimal nuqtaga yaqinlashguncha ushbu jarayon davom etadi.

Gradient descent algoritmining bir nechta variantlari mavjud. Ulardan biri — **Batch Gradient Descent**, bunda butun trening to'plami bo'yicha gradient hisoblanadi, bu esa aniqlikni oshiradi, ammo hisoblash sekinlashadi. Ikkinchisi — **Stochastic Gradient Descent (SGD)**, har bir namunaga asoslanib yangilanish olib boriladi, bu esa hisoblashni tezlashtiradi, ammo natija shovqinli bo'lishi mumkin. **Momentum** usuli oldingi gradientlarning ta'sirini hisobga olib, o'rganish jarayonida inersiya yaratadi va mahalliy minimumlardan chiqishga yordam beradi. **RMSProp** esa gradient kvadratlarining eksponensial o'rtachasini hisoblab, learning rate'ni moslashtiradi. Eng ommabopi — **Adam** algoritmi bo'lib, u Momentum va RMSProp usullarini birlashtirgan va juda samarali ishlaydi.

Optimizatorlarning Asosiy Vazifasi

Optimizatorlar neyron tarmoq modelining o'qitilishi jarayonida model parametrlarining optimal qiymatlarini aniqlash vazifasini bajaradi. Ushbu parametrlar — vaznlar, egilishlar yoki boshqa o'rganiladigan o'zgaruvchilar — tarmoqning har bir qatlami va birikmasida joylashgan bo'lib, ular modelning chiqishini belgilaydi. Optimizatorlar yordamida ushbu parametrlar tarmoq uchun eng yaxshi natijani beradigan tarzda yangilanadi. Masalan, tasniflash masalasida to'g'ri optimallashtirilgan model sinflar orasidagi chegarani aniq belgilaydi, regressiyada esa model haqiqiy qiymatlarga yaqin bashorat qiladi. Shu sababli optimallashtirish jarayoni butun mashinaviy o'rganish va chuqur o'rganish tizimlarining muvaffaqiyati uchun muhim hisoblanadi.

Gradient Hisoblash va Parametrlarni Yangilash

Optimizatsiya jarayonining yuragi — bu yo'qotish funksiyasining gradientini hisoblashdir. Yo'qotish funksiyasi modelning chiqish natijasini haqiqiy qiymat bilan taqqoslab, modelning qanchalik xato qilganini o'lchaydi. Gradient esa yo'qotish funksiyasining parametrlar bo'yicha qiyalik ko'rsatkichidir. Agar gradient ijobiy bo'lsa, demak yo'qotish funksiyasining qiymatini kamaytirish uchun parametрни kamaytirish kerak. Agar gradient manfiy bo'lsa — aksincha. Shu tarzda har bir parametr yangilanadi. Gradientni hisoblash uchun odatda **backpropagation** algoritmi qo'llaniladi, ayniqsa neyron tarmoqlar uchun. Backpropagation yordamida xatolik oxirgi qatlamdan boshlab tarmoq bo'ylab oldinga va orqaga tarqatiladi, natijada har bir parametr uchun gradient aniqlanadi.

To'g'ri gradientli optimizatsiya algoritmlari ko'plab sohalarda, ayniqsa quyidagi yo'nalishlarda muhim ahamiyat kasb etadi:

Mashina o'rganish: Katta hajmdagi ma'lumotlardan foydalangan holda statistik modellarni o'rganish va bashorat qilish uchun.

Neyron tarmoqlar va chuqur o'rganish: Murakkab arxitekturalar va ko'p qatlamli tarmoqlarni o'rgatishda parametrlarni optimallashtirish uchun.

Ma'lumot tahlili va regressiya: Ma'lumotlar orasidagi munosabatlarni aniqlash va prognoz qilishda.

Masofaviy ta'lim va maslahat tizimlari: Talabalar uchun individual o'quv yo'nalishini shakllantirishda.

Sun'iy intellekt sohasining boshqa sohalari: Masalan, tabiiy tilni qayta ishlash, kompyuter ko'rish, robototexnika va boshqalar.

Optimizatorlarning samarali ishlashi natijasida modellar yanada aniqroq va barqaror natijalar beradi, bu esa amaliyotda yuqori sifatli va foydali tizimlarni yaratishga olib keladi. Gradient tushish algoritmi va uning variatsiyalari samarali bo'lsa-da, ba'zi muammolarga duch keladi:

Lokal minimumlar: Funksiya ko'p lokal minimumlarga ega bo'lishi mumkin, va oddiy gradient tushish algoritmi global minimum o'rniga lokal minimumga tushib qolishi ehtimoli bor. Momentum va Adam kabi algoritmlar bunday muammoni yengillashtiradi.

O'rganish tezligining tanlanishi: Juda katta o'rganish tezligi algoritmnining barqaror ishlashini buzishi, juda kichik tezlik esa o'rganishni sekinlashtirishi mumkin. Shuning uchun adaptiv o'rganish tezligini sozlash muhim.

Gradientning yo'qolishi yoki portlashi: Ayniqsa chuqur tarmoqlarda gradient juda kichik yoki juda katta bo'lib, o'rganish jarayonida qiyinchiliklar yuzaga keladi. RMSProp va Adam algoritmlari ushbu muammoni kamaytirishga xizmat qiladi.

Xulosa

To'g'ri gradientli optimizatsiya algoritmlari sun'iy intellekt va mashinaviy o'rganishning asosiy tamoyillaridan biri bo'lib, ularning yordamida murakkab modellarning parametrlarini samarali tarzda o'rganish mumkin. Ushbu algoritmlar yordamida model o'zining xatolarini kamaytirib, natijada yanada aniq va ishonchli bashoratlar qilishi mumkin. Gradient tushish va uning variatsiyalari (SGD, Momentum, RMSProp, Adam va boshqalar) turli vazifalar va sharoitlar uchun moslashtirilgan, shuning uchun ular turli murakkablikdagi va hajmdagi modellarni o'rganishda keng qo'llaniladi. To'g'ri tanlangan optimizator va uning parametrlarini sozlash sizning modelingizning muvaffaqiyati va ishlash tezligida muhim rol o'ynaydi.

Foydalanilgan adabiyotlar

1. Bekmuratov Q.A. , Sun'iy intellekt va neyron tarmoqlar. O'quv qo'llanma, Samarqand – 2021.
2. Sadullayeva SH.A, Yusupov D.F., Yusupov F., Sun'iy intellect va neyronto'rli texnologiyalar. O'quv qo'llanma, Urganch – 2021.
3. H.N.Zayniddinov, T.A.Xo'jaqulov, M.P.Atadjanov, "Sun'iy intellekt" fanidan o'quv qo'llanma, Toshkent – 2018.
4. <http://www.raii.org/library> - Российская ассоциация искусственного интеллекта

5. <https://arxiv.org> Ilmiy maqolalar arxivi.
6. <https://www.coursera.org/learn/neural-networks-deep-learning>
7. **Stanford University CS231n: Convolutional Neural Networks for Visual Recognition (Lecture Notes)**
http://cs231n.stanford.edu/slides/2021/cs231n_2021_lecture_4.pdf
8. **Towards Data Science — “A Beginner’s Guide to Gradient Descent and Optimization Algorithms.”** <https://towardsdatascience.com/gradient-descent-and-its-variants-10f652806a3>