

NEYRON TARMOQLARINI O'QITISHDA ZANJIR QOIDASIGA
ASOSLANGAN HOSILA HISOBLASH USULLARI

Nuraliyeva Sabrina Shobek qizi

Shahrisabz davlat pedagogika instituti matematika yo'nalishi talabasi

Annotatsiya: Ushbu maqolada zamonaviy chuqur o'qitish modellarining asosi bo'lgan ko'p qatlamli neyron tarmoqlarini optimallashtirishda murakkab funksiyalarni differentsiallashtirish va zanjir qoidasining (chain rule) o'rni fundamental jihatdan tadqiq etiladi. Tadqiqotning dolzarbligi neyron tarmoqlarining o'ta murakkab iyerarxik strukturaga ega ekanligi va ulardagi millionlab parametrlarni (vaznlarni) xatolik funksiyasiga nisbatan minimal vaqt sarfi bilan optimallashtirish zarurati bilan belgilanadi. Maqolada teskari tarqalish (backpropagation) algoritmining matematik modeli ishlab chiqilgan bo'lib, unda xatolik signallarining chiqish qatlamidan kirish qatlamiga qarab rekursiv uzatilishi zanjir qoidasi orqali tahlil qilingan. Tadqiqot doirasida gradientlarni hisoblashning skalyar, vektor va tenzorli ko'rinishlari o'rtasidagi farqlar, shuningdek, Yakobian matritsalarining hisoblash graflaridagi o'rni yoritib berilgan. Alomatli differentsiallashtirish usullarining samaradorligi va gradientlarning yo'qolishi (vanishing gradient) hamda portlashi (exploding gradient) muammolari zanjir qoidasining matematik tabiati orqali asoslangan. Olingan natijalar shuni ko'rsatadiki, zanjir qoidasini to'g'ri qo'llash nafaqat algoritmlarning yaqinlashuvchanligini ta'minlaydi, balki hisoblash murakkabligini chiziqli darajaga tushirish imkonini beradi. Xulosa qismida ushbu matematik yondashuvning yangi avlod optimallashtirish algoritmlarini yaratishdagi ahamiyati bayon etilgan.

Kalit so'zlar: ko'p qatlamli neyron tarmoqlari, zanjir qoidasi (chain rule), teskari tarqalish algoritmi, gradient tushishi, xususiy hosila, hisoblash graflari, Yakobian matritsasi, gradientlarning yo'qolishi, tenzorli differentsiallashtirish, optimallashtirish samaradorligi, matematik tahlil.

Аннотация: В данной статье фундаментально исследуется роль цепного правила (chain rule) и дифференцирования сложных функций в оптимизации многослойных нейронных сетей, составляющих основу современных моделей глубокого обучения. Актуальность исследования обусловлена сложной иерархической структурой нейронных сетей и необходимостью оптимизации миллионов параметров (весов) относительно функции потерь с минимальными временными затратами. В работе разработана математическая модель алгоритма обратного распространения (backpropagation), в которой рекурсивная передача сигналов ошибки от выходного слоя к входному анализируется через призму цепного правила.

В рамках исследования освещены различия между скалярным, векторным и тензорным представлениями вычисления градиентов, а также роль матриц Якоби в вычислительных графах. Эффективность методов автоматического дифференцирования и проблемы затухания (vanishing gradient) и взрыва (exploding gradient) градиентов обоснованы через математическую природу цепного

правила. Полученные результаты показывают, что правильное применение цепного правила не только обеспечивает сходимость алгоритмов, но и позволяет снизить вычислительную сложность до линейного уровня. В заключении подчеркивается значимость данного математического подхода для разработки алгоритмов оптимизации нового поколения.

Ключевые слова: многослойные нейронные сети, цепное правило, алгоритм обратного распространения, градиентный спуск, частная производная, вычислительные графы, матрица Якоби, затухание градиента, тензорное дифференцирование, эффективность оптимизации, математический анализ.

Abstract: *This paper fundamentally investigates the role of the chain rule and the differentiation of composite functions in optimizing multi-layer neural networks, which form the bedrock of modern deep learning models. The relevance of this research is determined by the highly complex hierarchical structure of neural networks and the necessity to optimize millions of parameters (weights) relative to the loss function with minimal computational overhead. The study develops a mathematical model of the backpropagation algorithm, where the recursive transmission of error signals from the output layer to the input layer is analyzed through the lens of the chain rule. The research highlights the distinctions between scalar, vector, and tensor representations of gradient computations, as well as the role of Jacobian matrices in computational graphs. The efficiency of automatic differentiation methods and the issues of vanishing and exploding gradients are substantiated through the mathematical nature of the chain rule. The results indicate that the rigorous application of the chain rule not only ensures the convergence of algorithms but also enables the reduction of computational complexity to a linear scale. The paper concludes by emphasizing the significance of this mathematical approach in the development of next-generation optimization algorithms.*

Key words: *multi-layer neural networks, chain rule, backpropagation algorithm, gradient descent, partial derivative, computational graphs, Jacobian matrix, vanishing gradient, tensor differentiation, optimization efficiency, mathematical analysis.*

KIRISH

Zamonaviy hisoblash tizimlarida ko'p bosqichli differensiallanuvchi modellarni optimallashtirish masalasi fundamental matematik muammolardan biri hisoblanadi. Xususan, neyron tarmoqlari kabi murakkab iyerarxik tizimlarning o'rganish jarayoni xatolik funksiyasi sirtida global minimum nuqtasini izlashga tayanadi. Ushbu jarayonning asosi hisoblangan gradientli tushish usullari har bir vazn parametri bo'yicha xususiy hosilalarning aniq qiymatiga bog'liq. Tarmoq qatlamlari soni va parametrlar ko'lamini ortib borishi bilan hosila hisoblashning an'anaviy algoritmlari hisoblash resurslariga bo'lgan talabni eksponential darajada oshirib yuboradi. Ushbu muammoni hal qilishda murakkab funksiyalarni differensiallashning zanjir qoidasi (chain rule) hal qiluvchi rol o'ynaydi. Zanjir qoidasi murakkab tizimning umumiy hosilasini uning tarkibiy qismlari hosilalarining ko'paytmasi ko'rinishida ifodalash imkonini berib, teskari tarqalish (backpropagation) algoritmining matematik fundamentini tashkil etadi. Bu usul xatolik signallarining chiqish nuqtasidan

kirishga qarab rekursiv oqishini ta'minlaydi, natijada hisoblash murakkabligi parametrlar soniga nisbatan chiziqli ko'rinishga keladi. Tadqiqotning dolzarbligi shundaki, neyron tarmoqlarining chuqurlashishi gradientlarning so'nishi (vanishing gradient) yoki keskin ortib ketishi (exploding gradient) kabi destabilizatsiya jarayonlariga olib keladi. Mazkur maqolada zanjir qoidasiga asoslangan hosila hisoblash usullarining matematik mexanizmi, ularning skalyar va matritsaviy hisoblashlardagi samaradorligi hamda o'qitish jarayonining barqarorligini ta'minlashdagi deterministik ahamiyati atroflicha tahlil qilinadi. Tadqiqotning maqsadi — zanjir qoidasining neyron tarmoqlari iyerarxiyasidagi o'rnini nazariy jihatdan asoslash va optimallashtirish jarayonidagi samaradorlikni oshirish yo'llarini ko'rsatib berishdir.

ADABIYOTLAR TAHLILI VA METODOLOGIYA

Neyron tarmoqlarini o'qitishda gradientli optimallashtirish masalalari O'zbekistonlik olimlar tomonidan ham keng tadqiq etib kelinmoqda. Xususan, akademik S.S. Gulyamov va uning izdoshlari tomonidan intellektual tizimlarning matematik modellarini qurishda xatolik funksiyasini minimallashtirish usullari bo'yicha ko'plab izlanishlar olib borilgan. Shuningdek, murakkab tizimlarni boshqarishda differensial tenglamalar va zanjir qoidasini qo'llash masalalari Sh.H. Fazilov hamda O.R. Yusupovlarning ishlarida o'z aksini topgan. Ushbu manbalarda ta'kidlanishicha, ko'p o'lchovli sistemalarda gradientlarni hisoblashning analitik usullari algoritmlarning konvergensiya (yaqinlashuvchanlik) tezligini 30-40% gacha oshirish imkonini beradi. Tadqiqot metodologiyasi quyidagi bosqichma-bosqich matematik algoritmlarga asoslanadi: Neyron tarmog'ining har bir qatlamni kirish signallarining chiziqli transformatsiyasi va chiziqli bo'lmagan faollashtirish funksiyasi sifatida ifodalanadi. 1-chi qatlam uchun chiqish qiymati:

$$z^l = w^l \times a^{l-1} + b^l \quad a^l = v(z^l)$$

Bu yerda $v(\text{sigma})$ — faollashtirish funksiyasi (masalan, Sigmoid yoki ReLU).

Xatolik funksiyasi E ning w_{jk}^l vazniga nisbatan xususiy hosilasini hisoblashda zanjir qoidasining multiplikativ xususiyatidan foydalaniladi. Metodologiyaning asosi hosilani quyidagi qismlarga ajratishdan iborat:

$$\frac{\partial E}{\partial w_{jk}^l} = \frac{\partial E}{\partial a_j^l} \times \frac{\partial a_j^l}{\partial z_j^l} \times \frac{\partial z_j^l}{\partial w_{jk}^l}$$

Gradientni hisoblash "chiqishdan kirishga" (backward pass) tamoyili asosida amalga oshiriladi. Bunda har bir qatlam uchun "xatolik signali" (δ) aniqlanadi: $\delta_j^l = \frac{\partial E}{\partial z_j^l}$

Ushbu metodologiya hisoblash graflarida har bir oraliq natijani keshga (xotiraga) olish va ulardan keyingi hisob-kitoblarda qayta foydalanish imkonini beradi, bu esa takroriy hisoblashlarni cheklaydi. Tadqiqotda skalyar hisob-kitoblardan matritsaviy ko'rinishga o'tish metodikasi qo'llanilgan. Bu usul Yakobian matritsalar yordamida gradientlarni parallel hisoblash imkonini beradi:

NATIJALAR

Tadqiqotning ushbu qismida zanjir qoidasining ko'p qatlamli neyron tarmoqlarida gradientni aniqlashdagi aniqligi va samaradorligini tekshirish maqsadida sonli eksperiment o'tkazildi. Eksperiment uchun bitta yashirin qatlamga ega bo'lgan minimalistik model tanlandi. Modelda kirish signali $x = 1.0$, vazn koeffitsiyenti $w = 0.5$ va maqsadli qiymat $y_{\text{target}} =$

0.01 deb belgilandi. Faollashtirish funksiyasi sifatida sigmoid funksiyasi tanlanib, uning natijasi v harfi bilan ifodalandi. Neyronning umumiy kirish qiymati z va chiqish qiymati v quyidagicha aniqlandi:

$$z = w \times x = 0,5 \times 1,0 = 0,5 \quad v = v(z) = \frac{1}{1 + e^{-z}} \approx 0,622459$$

Olingan natijaga ko'ra, xatolik funksiyasining ($E = \frac{1}{2}(v - y_{\text{target}})^2$) joriy qiymati $E \approx 0.1875$ ga teng bo'ldi. Vaznni optimallashtirish uchun $\frac{\partial E}{\partial w}$ hosilasi zanjir qoidasiga ko'ra uchta komponentga ajratildi:

$$\frac{\partial E}{\partial w} = \frac{\partial E}{\partial v} \times \frac{\partial v}{\partial z} \times \frac{\partial z}{\partial w}$$

$\frac{\partial E}{\partial v} = v - y_{\text{target}} = 0,622459 \times 0,01 = 0.612459$ sigmoid hosilasi $v'(z) = v(1-v)$ xossasiga ega bulgani uchun $\frac{\partial v}{\partial z} = v(1-v) = 0,622459 \times (1 - 0,622459) \approx 0.23499$

$\frac{\partial z}{\partial w} = x = 1.0$ endi gradiyentni aniqlaymiz: $\frac{\partial E}{\partial w} = 0.612459 \times 0.23499 \times 1.0 \approx 0.14392$ Olingan $\frac{\partial E}{\partial w} = 0.14392$ qiymati xatolik funksiyasining vazn o'zgarishiga nisbatan sezgirligini ko'rsatadi. Ushbu sonli eksperiment shuni isbotlaydiki, zanjir qoidasi yordamida hatto murakkab iyerarxik bog'lanishlarda ham har bir parametrlarning xatolikka qo'shgan "ulushini" aniq hisoblash mumkin. Bu esa o'z navbatida, neyron tarmoqlarini o'qitish jarayonida "ko'r-ko'rona" qidiruv o'rniga, matematik asoslangan deterministik optimallashtirishni ta'minlaydi.

MUHOKAMA

Olingan natijalar ko'p qatlamli neyron tarmoqlarini optimallashtirishda zanjir qoidasining nafaqat matematik vosita, balki tizimli samaradorlikni ta'minlovchi asosiy arxitektura ekanligini tasdiqlaydi.

O'tkazilgan sonli eksperimentda gradient qiymatining $\frac{\partial E}{\partial w} = 0.14392$ bo'lib chiqishi, xatolik signalining tarqalishida faollashtirish funksiyasi hosilasi (≈ 0.23499) hal qiluvchi to'siq yoki o'tkazgich vazifasini bajarishini ko'rsatdi. Bu holat neyron tarmoqlari chuqurlashgani sayin gradientning nima sababdan nolga yaqinlashib borishini (so'nishini) aniq matematik ifodalarda tushuntirib beradi.

Tadqiqotda qo'llanilgan teskari tarqalish metodikasi Sh.H. Fazilov va O.R. Yusupovlar tomonidan ilgari surilgan ko'p o'lchovli tizimlarni optimallashtirish nazariyasini amaliy jihatdan boyitadi. Ayniqsa, metodologiyada keltirilgan hosilani komponentlarga ajratish usuli, murakkab tizimlarni qismlarga bo'lib tahlil qilish imkonini beradi.

Eksperiment natijalari shuni ko'rsatadiki, vaznlarni yangilashda gradientning aniq hisoblanishi algoritmlarning yaqinlashuvchanligini (konvergentsiyasini) ta'minlaydi, bu esa S.S. Gulyamov tadqiqotlarida ta'kidlanganidek, intellektual tizimlarning aniqlik darajasini 30-40% gacha oshirishning muhim omilidir. Shuni ham alohida ta'kidlash joizki, zanjir qoidasining skalyar ko'rinishdan matritsaviy ko'rinishga o'tkazilishi zamonaviy GPU protsessorlarining parallellashtirish imkoniyatlaridan to'liq foydalanishga yo'l ochadi. Garchi biz bitta neyron misolida sodda hisob-kitobni amalga oshirgan bo'lsak-da, ushbu tamoyil millionlab neyronlardan iborat tizimlarda ham o'zgarishsiz qoladi.

Biroq, muhokama qilinayotgan usulning bir kamchiligi shundaki, qatlamlar soni haddan tashqari ortib ketganda, zanjir qoidasi multiplikativ xarakteri tufayli gradientning destabilizatsiyasiga (so'nishi yoki portlashi) to'liq to'sqinlik qila olmaydi.

XULOSA

Ushbu tadqiqot natijalari shuni ko'rsatadiki, zamonaviy intellektual tizimlarning matematik asosi bo'lgan ko'p bosqichli differensiallanuvchi modellarni optimallashtirishda zanjir qoidasi fundamental ahamiyatga ega. O'tkazilgan nazariy tahlil va sonli eksperimentlar orqali isbotlandiki, ushbu qoida murakkab iyerarxik tizimlarda gradientlarni hisoblash murakkabligini eksponensial o'sishdan saqlab, uni chiziqli ko'rinishda ushlab turuvchi eng optimal mexanizmdir. Tadqiqot davomida sigmoid funksiyasi misolida ko'rib chiqilgan amaliy misol xatolik signallarining qatlamlararo tarqalishi nafaqat funksional bog'liqlik, balki deterministik matematik qonuniyatlarga tayanishini tasdiqladi. Shu bilan birga, zanjir qoidasining multiplikativ tabiati chuqur neyron tarmoqlarida gradientning so'nishi kabi destabilizatsiya jarayonlarining asosiy sababi ekanligi nazariy jihatdan asoslandi. Olingan natijalar neyron tarmoqlarini "qora quti" sifatida emas, balki aniq hisob-kitoblarga asoslangan tizim sifatida talqin qilish imkonini beradi va algoritmlarni parallellashtirish orqali hisoblash unumdorligini oshirish uchun zamin yaratadi. Kelgusida ushbu matematik apparatni o'ta chuqur modellar sharoitida yanada takomillashtirish va gradientlarni approksimatsiyalashning yangi usullarini joriy etish sun'iy intellekt texnologiyalarining energiya samaradorligini oshirishda hal qiluvchi rol o'ynaydi. Umumiy xulosa qilib aytganda, zanjir qoidasining neyron tarmoqlari iyerarxiyasidagi o'rni nazariy jihatdan o'zgarmas bo'lib, u murakkab tizimlarni o'qitish va optimallashtirish jarayonlarining eng samarali vositasi bo'lib qolmoqda.

FOYDALANILGAN ADABIYOTLAR RO'YXATI:

1. Gulyamov S.S. va boshqalar. "Sun'iy intellekt tizimlari va raqamli iqtisodiyot: matematik modellar va algoritmlar". Toshkent, "Iqtisodiyot" nashriyoti, 2021-yil.
2. Fazilov Sh.H., Mirzaev N.M. "Ob'ektlarni aniqlash va klassifikatsiyalashning matematik usullari va algoritmlari". O'zbekiston Fanlar Akademiyasi axborotnomasi, Toshkent, 2019-yil.
3. Yusupov O.R. "Ko'p o'lchovli murakkab tizimlarni boshqarishda differensial tenglamalar va optimallashtirish usullari". O'zbekiston Milliy universiteti ilmiy jurnali, 2020-yil.
4. Zaynutdinova M.B. "Neyron tarmoqlari: arxitekturasini va o'qitish algoritmlari". O'quv qo'llanma. Toshkent Axborot texnologiyalari universiteti (TATU), 2022-yil.
5. Kabulov A.V., Normatov A.M. "Intellektual tahlil va mantiqiy modelashtirishning dolzarb masalalari". "Algoritmlar" ilmiy-texnik markazi to'plami, 2018-yil.
6. Xaydarov A.A. "Gradientli tushish usullarining yaqinlashuvchanligini oshirish mexanizmlari". O'zbekiston Fanlar Akademiyasi "Muammolar va yechimlar" jurnali, 2021-yil.
7. Bekmurodov A.A. "Hisoblash tizimlarida ko'p qatlamli modellarni optimallashtirishning iqtisodiy va matematik samaradorligi". Axborot texnologiyalari va kommunikatsiyalar jurnali, 2023-yil.