

MOMENT MEXANIZMINING MATEMATIK ASOSLARI VA ADAM TURIDAGI ADAPTIV OPTIMIZATORLARDA QO'LLANISH SAMARADORLIGINI TAHLIL QILISH

Tojimatov Israil Nurmatovich

Farg'ona davlat universiteti Amaliy matematika
va informatika kafedrası katta o'qituvchisi

E-mail: israiltojimatov@gmail.com

Raxmonova Umidaxon Berdiqu qizi

Farg'ona davlat universiteti Amaliy matematika
yo'nalishi 3-bosqich 23.09-guruh talabasi

E-mail: umidaxonraxmonova2004@gmail.com

<https://doi.org/10.5281/zenodo.17999840>

ANNOTATSIYA

Ushbu tezisda mashinaviy o'rganish va sun'iy intellekt modellarini o'qitishda qo'llaniladigan optimallashtirish algoritmlarining muhim jihatlaridan biri bo'lgan moment mexanizmining matematik asoslari va uning Adam turidagi adaptiv optimizatorlarda qo'llanish samaradorligi tahlil qilinadi. Natijalar shuni ko'rsatadiki, moment mexanizmi va adaptiv optimallashtirish yondashuvlari zamonaviy chuqur o'rganish modellarini samarali o'qitishda muhim ahamiyat kasb etadi.

Kalit so'zlar: Moment mexanizmi, optimallashtirish algoritmlari, adaptiv o'rganish tezligi, Adam optimizatori, gradient tushish, chuqur o'rganish, sun'iy intellekt.

KIRISH

So'nggi yillarda sun'iy intellekt va mashinaviy o'rganish texnologiyalari juda tez sur'atlar bilan rivojlanib, kundalik hayotimizga sezilarli darajada kirib keldi. Bugungi kunda oddiy mobil ilovalardan tortib murakkab sanoat tizimlarigacha bo'lgan ko'plab yechimlar aynan o'rganishga qodir bo'lgan matematik modellar asosida ishlamoqda. Ushbu modellar qanchalik murakkab bo'lmasin, ularning samarali ishlashi bitta asosiy jarayonga – optimallashtirishga bevosita bog'liq. Modelni o'qitish jarayonida asosiy maqsad xatolik funksiyasini imkon qadar tez va barqaror ravishda minimumga olib kelishdan iborat bo'ladi. Amaliyotda bu vazifa ko'pincha gradient tushish algoritmlari yordamida bajariladi. Biroq oddiy gradient tushish usullari yuqori o'lchamli fazolarda, notekis va murakkab yo'qotish funksiyalarida sekin ishlashi, keskin tebranishlar yuzaga keltirishi yoki optimal nuqtaga yetib bormasdan jarayonda to'xtab qolishi mumkin. Aynan shu cheklovlar optimallashtirish algoritmlarini yanada "aqlli", moslashuvchan va barqaror qilish zaruratini yuzaga keltiradi. Shu nuqtayi nazardan moment mexanizmi va unga asoslangan adaptiv optimizatorlar zamonaviy mashinaviy o'rganish modellarida muhim ahamiyat kasb etadi.

Moment mexanizmi optimallashtirish algoritmlarini takomillashtirishda muhim matematik g'oya sifatida qaraladi va u fizikadagi impuls tushunchasiga bevosita o'xshashlik asosida ishlab chiqilgan. Klassik gradient tushish algoritmda parametrlar har bir iteratsiyada faqat joriy gradient qiymatiga tayanib yangilanadi. Bunday yondashuv keskin o'zgaruvchan gradientlarga ega bo'lgan yo'qotish funksiyalarida tebranishlarning kuchayishiga olib keladi. Moment mexanizmi esa bu muammoni bartaraf etish uchun oldingi gradientlar haqidagi ma'lumotni ham hisobga oladi. Matematik jihatdan, bu usul gradientlarning vaqt bo'yicha

eksponensial o'rtacha qiymatini hisoblash orqali amalga oshiriladi. Natijada parametrlarning yangilanish yo'nalishi faqat bitta qadamga emas, balki oldingi qadamlarning umumiy tendensiyasiga mos ravishda aniqlanadi. Bu yondashuv optimallashtirish trayektoriyasini silliqalashtiradi, tasodifiy tebranishlarni kamaytiradi va umumiy konvergentsiya jarayonini tezlashtiradi. Ayniqsa chuqur neyron tarmoqlarda gradientlar turli qatlamlarda turlicha miqyosga ega bo'lishi sababli moment mexanizmi modelni barqarorroq o'qitishga imkon beradi. Shunday qilib, moment usuli optimallashtirish jarayoniga "xotira" elementi qo'shib, modelni o'qitishda yanada ishonchli matematik asos yaratadi.

Optimallashtirish jarayonida o'rganish tezligi tanlovi eng muhim masalalardan biri hisoblanadi, chunki noto'g'ri tanlangan o'rganish tezligi modelni sekinlashtirishi yoki umuman konvergentsiyaga erishmasligiga olib kelishi mumkin. Klassik yondashuvlarda barcha parametrlar uchun yagona va doimiy o'rganish tezligidan foydalaniladi, ammo bu usul murakkab modellar uchun har doim ham samarali bo'lavermaydi. Amaliyot shuni ko'rsatadiki, ba'zi parametrlar tezroq yangilanishni talab qilsa, boshqalari ehtiyotkor va kichik qadamlar bilan sozlanishi lozim. Adaptiv o'rganish tezligi g'oyasi aynan shu muammoni hal qilishga qaratilgan bo'lib, har bir parametr uchun o'rganish tezligini avtomatik ravishda moslashtirishni nazarda tutadi. Matematik nuqtayi nazardan, bu yondashuv gradientlarning tarqalishi va ularning kvadrat qiymatlaridan foydalanish orqali amalga oshiriladi. Natijada katta gradientlarga ega parametrlar uchun yangilanishlar sekinlashtiriladi, kichik gradientlarga ega parametrlar esa tezroq o'zgartiriladi. Bu esa optimallashtirish jarayonining barqarorligini oshirib, turli miqyosdagi parametrlar bilan samarali ishlash imkonini beradi. Adaptiv o'rganish tezligi konsepsiyasi keyinchalik zamonaviy optimizatorlarning, xususan Adam algoritmining nazariy poydevorini shakllantirdi.

Adam optimizatori zamonaviy mashinaviy o'rganish va chuqur neyron tarmoqlarda eng keng qo'llaniladigan optimallashtirish algoritmlaridan biri hisoblanadi. Uning asosiy ustunligi shundaki, u moment mexanizmi va adaptiv o'rganish tezligi g'oyalarini yagona matematik modelga muvaffaqiyatli tarzda birlashtiradi. Adam algoritmi parametrlarni yangilash jarayonida faqatgina joriy gradient qiymatlariga tayanib qolmay, balki gradientlarning vaqt bo'yicha birinchi va ikkinchi tartibli statistik xususiyatlarini ham hisobga oladi. Bu esa optimallashtirish jarayonini yanada barqaror, tezkor va ishonchli qiladi.

Adam optimizatorida birinchi navbatda gradientlarning eksponensial o'rtacha qiymati hisoblanadi, bu jarayon moment mexanizmining matematik ifodasiga mos keladi. Ushbu o'rtacha qiymat gradientlarning yo'nalishi bo'yicha umumiy tendensiyani aks ettiradi va keskin tebranishlarning oldini olishga xizmat qiladi. Shu bilan birga, Adam algoritmi gradientlarning kvadrat qiymatlarining eksponensial o'rtachasini ham hisoblab boradi. Bu ikkinchi komponent gradientlarning tarqalishini baholash imkonini beradi va adaptiv o'rganish tezligini shakllantirishda muhim rol o'ynaydi. Natijada har bir parametr uchun alohida va moslashuvchan yangilanish qadamini tanlash imkoniyati yuzaga keladi.

Adam optimizatorining muhim matematik jihatlaridan biri boshlang'ich iteratsiyalarda yuzaga keladigan statistik og'ishlarni tuzatish mexanizmidir. Dastlabki bosqichlarda eksponensial o'rtachalar nolga yaqin qiymatlarga moyil bo'lishi mumkin, bu esa parametrlarning yangilanishida noaniqliklarni keltirib chiqaradi. Adam algoritmi ushbu muammoni maxsus tuzatish koeffitsiyentlari yordamida hal qiladi. Ushbu tuzatishlar orqali birinchi va ikkinchi momentlarning bahosi yanada aniqroq bo'ladi va optimallashtirish jarayoni

boshidanoq barqaror yo'lga qo'yiladi. Aynan shu xususiyat Adam optimizatorini amaliy masalalarda juda samarali qiladi.

Adam algoritmi yuqori o'lchamli parametrlar fazosida, ayniqsa chuqur neyron tarmoqlarda muvaffaqiyatli ishlaydi. Har bir qatlamdagi parametrlar turli masshtab va sezgirlikka ega bo'lishi mumkinligi sababli, adaptiv yangilanish mexanizmi modelni o'qitishda katta ustunlik beradi. Adam optimizatori shuningdek, shovqinli gradientlar bilan ishlashda ham samarali bo'lib, katta hajmdagi ma'lumotlar to'plamida mini-batch usuli bilan o'qitiladigan modellar uchun qulay hisoblanadi. Shu jihatlari tufayli Adam bugungi kunda tasvirlarni aniqlash, tabiiy tilni qayta ishlash va vaqt qatorlarini tahlil qilish kabi ko'plab sohalarda standart optimallashtirish usuli sifatida qo'llanilmoqda.

Adam optimizatorining samaradorligi bir nechta asosiy mezonlar orqali baholanadi, jumladan konvergensiya tezligi, barqarorlik va hisoblash samaradorligi. Konvergensiya tezligi nuqtayi nazaridan Adam algoritmi klassik gradient tushish va oddiy moment usullariga nisbatan sezilarli ustunlikka ega. Adaptiv o'rganish tezligi va moment mexanizmining birgalikdagi ta'siri tufayli Adam parametrlarni optimal nuqtaga tezroq olib keladi. Ayniqsa murakkab va notekis yo'qotish funksiyalarida Adam algoritmi lokal minimumlar va tekis sohalardan samarali o'tib ketish imkonini beradi.

Barqarorlik jihatidan Adam optimizatori optimallashtirish jarayonida yuzaga keladigan tebranishlarni sezilarli darajada kamaytiradi. Gradientlarning birinchi va ikkinchi momentlari hisobga olingani sababli, parametrlar yangilanishi keskin sakrashlarsiz amalga oshiriladi. Bu holat modelni o'qitish jarayonida xatolik funksiyasining silliq pasayishini ta'minlaydi. Natijada o'qitish jarayoni davomida modelning o'rganishi izchil va nazorat qilinadigan bo'ladi. Bu xususiyat Adam optimizatorini real amaliy vazifalarda, ayniqsa katta va shovqinli ma'lumotlar bilan ishlashda juda foydali qiladi.

Hisoblash samaradorligi nuqtayi nazaridan Adam algoritmi qo'shimcha hisoblash xarajatlarini talab qilsa-da, ushbu xarajatlar amaliy jihatdan o'zini oqlaydi. Har bir parametr uchun qo'shimcha o'rtacha qiymatlarni saqlash zarur bo'lsa-da, bu xotira talabi zamonaviy hisoblash tizimlari uchun muammo tug'dirmaydi. Aksincha, tezroq konvergensiyaga erishilishi umumiy o'qitish vaqtini qisqartiradi. Shu sababli Adam optimizatori katta hajmdagi modellarda ham samarali ishlash imkonini beradi.

Shu bilan birga, Adam optimizatorining ayrim cheklovlari ham mavjud. Ba'zi holatlarda u umumlashtirish qobiliyati jihatidan oddiy stochastic gradient tushish usullaridan ortda qolishi mumkin. Tadqiqotlar shuni ko'rsatadiki, Adam juda tez konvergensiyaga erishgani sababli, ayrim modellar optimal bo'lmagan minimumlarda to'xtab qolishi ehtimoli mavjud. Shu sababli amaliyotda Adam optimizatoridan foydalanishda uning giparparametrlarini ehtiyotkorlik bilan tanlash muhim ahamiyatga ega. Ushbu cheklovlarni bartaraf etish maqsadida AdamW, AMSGrad kabi modifikatsiyalar ishlab chiqilgan bo'lib, ular Adam algoritmining barqarorligi va umumlashtirish qobiliyatini yanada oshirishga qaratilgan.

Umuman olganda, Adam turidagi optimizatorlar moment mexanizmi va adaptiv o'rganish tezligi g'oyalarining muvaffaqiyatli uyg'unlashuvi natijasida yuzaga kelgan bo'lib, ular zamonaviy sun'iy intellekt modellarini o'qitishda muhim o'rin egallaydi. Ularning samaradorligi nafaqat nazariy tahlillar, balki ko'plab amaliy tajribalar orqali ham tasdiqlangan. Shu sababli Adam va uning modifikatsiyalari bugungi kunda chuqur o'rganish tizimlarida eng ishonchli optimallashtirish usullaridan biri sifatida qaraladi.

Xulosa

Olib borilgan tahlillar shuni ko'rsatadiki, optimallashtirish algoritmlarining samaradorligi mashinaviy o'rganish modellarining umumiy ishlash sifatiga bevosita ta'sir ko'rsatadi. Klassik gradient tushish usullari sodda bo'lishiga qaramay, murakkab va yuqori o'lchamli masalalarda yetarli darajada samarali emas. Moment mexanizmining joriy etilishi optimallashtirish jarayoniga vaqt bo'yicha xotira xususiyatini kiritib, tebranishlarni kamaytirish va konvergensiyani tezlashtirish imkonini beradi. Adaptiv o'rganish tezligi g'oyasi esa parametrlarning turli miqyosda yangilanishiga moslashuvchan yondashuvni ta'minlaydi.

Adam optimizatori ushbu ikki muhim yondashuvni yagona matematik modelda birlashtirgan holda, zamonaviy sun'iy intellekt modellarini o'qitishda yuqori samaradorlikni namoyon etadi. Uning tez konvergensiya qilishi, barqaror ishlashi va shovqinli gradientlarga nisbatan chidamliligi amaliy masalalarda katta ustunlik hisoblanadi. Shu bilan birga, Adam algoritmining ayrim holatlarda umumlashtirish qobiliyatiga salbiy ta'sir ko'rsatishi mumkinligi ham e'tibordan chetda qolmasligi lozim. Ushbu cheklovlar AdamW va AMSGrad kabi modifikatsiyalar orqali qisman bartaraf etilgan.

Umuman olganda, moment mexanizmi va Adam turidagi adaptiv optimizatorlar chuqur o'rganish sohasida optimallashtirish muammolarini samarali hal etishda muhim nazariy va amaliy ahamiyatga ega. Ularning to'g'ri tanlanishi va mos sozlanishi zamonaviy sun'iy intellekt tizimlarining aniqligi va barqarorligini oshirishda muhim omil bo'lib xizmat qiladi.

Adabiyotlar, References, Литературы:

1. Kingma D. P., Ba J. L. **Adam: A Method for Stochastic Optimization** // Proceedings of the 3rd International Conference on Learning Representations (ICLR). – San Diego, 2015. – 15 p.
2. Goodfellow I., Bengio Y., Courville A. **Deep Learning**. – Cambridge: MIT Press, 2016. – 775 p.
3. Ruder S. **An overview of gradient descent optimization algorithms** // arXiv preprint arXiv:1609.04747. – 2016. – 23 p.
4. Bottou L., Curtis F. E., Nocedal J. **Optimization Methods for Large-Scale Machine Learning** // SIAM Review. – 2018. – Vol. 60, No. 2. – P. 223–311.
5. Duchi J., Hazan E., Singer Y. **Adaptive Subgradient Methods for Online Learning and Stochastic Optimization** // Journal of Machine Learning Research. – 2011. – Vol. 12. – P. 2121–2159.
6. Reddi S. J., Kale S., Kumar S. **On the Convergence of Adam and Beyond** // Proceedings of the International Conference on Learning Representations (ICLR). – Vancouver, 2018. – 19 p.
7. Nesterov Yu. **Introductory Lectures on Convex Optimization: A Basic Course**. – Boston: Springer, 2004. – 236 p.
8. Bishop C. M. **Pattern Recognition and Machine Learning**. – New York: Springer, 2006. – 738 p.
9. Polyak B. T. **Introduction to Optimization**. – New York: Optimization Software Inc., 1987. – 188 p.
10. Sutton R. S., Barto A. G. **Reinforcement Learning: An Introduction**. – 2nd ed. – Cambridge: MIT Press, 2018. – 552 p.