

ASOSIY KOMPONENTLAR TAHLILI

Onarkulov Maqsadjon Karimberdiyevich

Farg'ona Davlat Universiteti Amaliy matematika va informatika

kafedrası dotsenti

maxmaqsad@gmail.com

Yusupov Mirsaid Abdulaziz o'g'li

Farg'ona Davlat Universiteti Amaliy matematika va informatika

kafedrası o'qituvchisi

mirsaidbek@gmail.com

Kimsanboyeva Shohsanam Mirzohid qizi

Farg'ona Davlat Universiteti Amaliy matematika yo'nalishi 3-kurs talabasi

kimsanboyevashohsanam@gmail.com

Annotatsiya: Ushbu maqolada biz o'lchamlarni kamaytirishning eng mashhur usullaridan birini, ya'ni asosiy komponentlar tahlilini (PCA) muhokama qilamiz. Asosiy komponentlar tahlili populyatsiya genetikasi, mikrobioma tadqiqotlari va atmosfera fanlari kabi ko'plab sohalarda qo'llaniladi.

Аннотация: Здесь мы обсудим один из самых популярных методов — анализ размерностей (PCA). Анализ главных компонент является эталоном во многих областях, включая популяционную генетику, характеристику микробиома и науку об атмосфере.

Abstract: Here we discuss one of the most popular methods, dimensional analysis (PCA). Principal component analysis is a benchmark in many fields, including population genetics, microbiome characterization, and atmospheric science.

Kalit so'zlar: Sun'iy intellektda asosiy komponentlar tahlilida ma'lumot to'plamlarini kamaytirish, asosiy komponentlarini qurish.

Keywords: reduction of datasets, construction of principal components in principal component analysis in artificial intelligence.

Ключевые слова: сокращение наборов данных, построение главных компонент в анализе главных компонент в искусственном интеллекте.

Asosiy komponentlar tahlili (PCA) - bu chiziqli o'lchamlarni qisqartirish usuli bo'lib, ma'lumotlarni tahlil qilish, vizualizatsiya qilish va ma'lumotlarni oldindan qayta ishlashda qo'llaniladi.

Ma'lumotlar chiziqli ravishda yangi koordinatalar tizimiga aylantiriladi, shunda ma'lumotlardagi eng katta o'zgarishlarni qamrab oladigan yo'nalishlarni osongina aniqlash mumkin.

TA'LIM, TARBIYA VA INNOVATSIYALAR

I son, Dekabr

Asosiy komponentlar tahlili populyatsiya genetikasi, mikrobioma tadqiqotlari va atmosfera fanlari kabi ko'plab sohalarida qo'llaniladi. PCA ni bajarayotganda, to'plamning birinchi asosiy komponenti .

O'zgaruvchilar - eng ko'p dispersiyani tushuntiruvchi asl o'zgaruvchilarning chiziqli birikmasi sifatida hosil qilingan hosila o'zgaruvchidir. Ikkinchi asosiy komponent birinchi komponentning ta'siri olib tashlanganidan keyin qolgan narsalardagi eng katta farqni tushuntiradi.

Ma'lumotlar to'plamidagi xususiyatlar yoki o'lchamlar soni ortishi bilan statistik ahamiyatga ega natijani olish uchun zarur bo'lgan ma'lumotlar miqdori eksponent ravishda oshadi. Bu haddan tashqari moslashish, hisoblash vaqtini ko'paytirish va mashinani o'rganish modellarining aniqligini pasaytirish kabi muammolarga olib kelishi mumkin, bu yuqori o'lchamli ma'lumotlar bilan ishlashda yuzaga keladigan o'lchovlilik muammolarining la'nati sifatida tanilgan. O'lchovlar soni ortib borishi bilan xususiyatlarning mumkin bo'lgan kombinatsiyalari soni eksponent ravishda oshadi, bu esa ma'lumotlarning vakillik namunasini olishni hisoblashda qiyinlashtiradi. Klasterlash yoki tasniflash kabi vazifalarni bajarish qimmatga tushadi, chunki algoritmlar ancha katta xususiyat maydonini qayta ishlashga muhtoj, bu esa hisoblash vaqtini va murakkabligini oshiradi. Bundan tashqari, ba'zi bir mashinani o'rganish algoritmlari o'lchovlar soniga sezgir bo'lishi mumkin, bu esa past o'lchamli ma'lumotlar bilan bir xil aniqlik darajasiga erishish uchun ko'proq ma'lumotlarni talab qiladi. O'lchovlilikning la'nati bilan kurashish uchun xususiyatni tanlash va xususiyatni ajratib olishni o'z ichiga olgan Xususiyat muhandislik usullari qo'llaniladi. O'lchovni qisqartirish - imkon qadar asl ma'lumotni saqlab qolgan holda kirish funktsiyalari sonini kamaytirishga qaratilgan xususiyatlarni ajratib olish texnikasining bir turi.

Ushbu maqolada biz o'lchamlarni kamaytirishning eng mashhur usullaridan birini, ya'ni Asosiy komponentlar tahlilini (PCA) muhokama qilamiz.

Asosiy komponentlar tahlili yoki PCA - o'lchamlarni kamaytirish usuli bo'lib, u ko'pincha katta to'plamdagi ma'lumotlarning katta qismini o'z ichiga olgan katta o'zgaruvchilar to'plamini kichikroqqa aylantirish orqali katta ma'lumotlar to'plamlarining o'lchamlarini kamaytirish uchun ishlatiladi.

Ma'lumotlar to'plamining o'zgaruvchilari sonini kamaytirish, tabiiyki, aniqlik hisobiga keladi, ammo o'lchamlarni kamaytirishdagi hiyla oddiylik uchun biroz aniqlik bilan savdo qilishdir. Chunki kichikroq ma'lumotlar to'plamlarini o'rganish va vizualizatsiya qilish osonroq bo'ladi va shu bilan ma'lumotlar nuqtalarini tahlil qilishni begona o'zgaruvchilarsiz mashinani o'rganish algoritmlari uchun ancha oson va tezroq qiladi.

Xulosa qilib aytganda, PCA g'oyasi oddiy: iloji boricha ko'proq ma'lumotni saqlab, ma'lumotlar to'plamining o'zgaruvchilari sonini kamaytirish.

Asosiy komponentlar - bu chiziqli birikmalar yoki boshlang'ich o'zgaruvchilarning aralashmasi sifatida tuzilgan yangi o'zgaruvchilar. Ushbu kombinatsiyalar shunday amalga oshiriladiki, yangi o'zgaruvchilar (ya'ni, asosiy komponentlar) o'zaro bog'liq emas va boshlang'ich o'zgaruvchilar ichidagi ma'lumotlarning aksariyati birinchi komponentlarga

TA'LIM, TARBIYA VA INNOVATSIYALAR

I son, Dekabr

siqiladi yoki siqiladi. Shunday qilib, g'oya 10 o'lchovli ma'lumotlar sizga 10 ta asosiy komponentni beradi, ammo PCA birinchi komponentga maksimal mumkin bo'lgan ma'lumotni, so'ngra ikkinchi qismga maksimal qolgan ma'lumotni va shunga o'xshash narsaga ega bo'lmaguncha joylashtirishga harakat qiladi.

Ma'lumotni asosiy komponentlarda shu tarzda tashkil qilish sizga ko'p ma'lumotni yo'qotmasdan o'lchovni kamaytirishga imkon beradi, bu esa kam ma'lumotga ega bo'lgan komponentlarni tashlab, qolgan komponentlarni yangi o'zgaruvchilar sifatida ko'rib chiqish orqali. Bu yerda tushunish kerak bo'lgan muhim narsa shundaki, asosiy komponentlar kamroq izohlanadi va hech qanday haqiqiy ma'noga ega emas, chunki ular boshlang'ich o'zgaruvchilarning chiziqli kombinatsiyasi sifatida tuzilgan.

Geometrik nuqtai nazardan, asosiy komponentlar maksimal dispersiya miqdorini tushuntiruvchi ma'lumotlarning yo'nalishlarini, ya'ni ma'lumotlarning ko'p ma'lumotlarini qamrab oladigan chiziqlarni ifodalaydi. Bu erda dispersiya va ma'lumot o'rtasidagi bog'liqlik shundan iboratki, chiziq tomonidan olib boriladigan dispersiya qanchalik katta bo'lsa, uning bo'ylab ma'lumotlar nuqtalarining tarqalishi qanchalik katta bo'lsa va chiziq bo'ylab dispersiya qanchalik katta bo'lsa, u shunchalik ko'p ma'lumotga ega bo'ladi. Bularning barchasini sodda qilib aytadigan bo'lsak, asosiy komponentlarni ma'lumotlarni ko'rish va baholash uchun eng yaxshi burchakni ta'minlaydigan yangi o'qlar deb tasavvur qiling, shunda kuzatishlar orasidagi farqlar yaxshiroq ko'rinadi.

PCA asosiy komponentlarni qanday quradi:

Ma'lumotlarda qancha o'zgaruvchi bo'lsa, shuncha asosiy komponentlar mavjudligi sababli, asosiy komponentlar shunday tuzilganki, birinchi asosiy komponent ma'lumotlar to'plamidagi mumkin bo'lgan eng katta farqni hisobga oladi. Masalan, ma'lumotlar to'plamimizning tarqalish grafigi quyida ko'rsatilganidek, deb faraz qilaylik, biz birinchi asosiy komponentni taxmin qila olamizmi? Ha, bu taxminan binafsha rangli belgilarga mos keladigan chiziq, chunki u boshlang'ichdan o'tadi va bu nuqtalarning proyeksiyasi (qizil nuqta) eng ko'p tarqalgan chiziqdir. Yoki matematik jihatdan aytganda, bu dispersiyani maksimal darajada oshiradigan chiziq (prognoz qilingan nuqtalardan (qizil nuqta) boshlang'ichgacha bo'lgan kvadratik masofalarning o'rtacha qiymati).

Asosiy komponent tahlili ikkinchi asosi ham xuddi shu tarzda, birinchi asosiy komponent bilan korrelyatsiyasiz (ya'ni, perpendikulyar) va keyingi eng yuqori dispersiyani hisobga olgan holda hisoblanadi. Bu o'zgaruvchilarning dastlabki soniga teng bo'lgan p asosiy komponentlar jami hisoblanmaguncha davom etadi.

PCA ni bosqichma-bosqich tushuntirish:

Standartlashtirish- ushbu bosqichning maqsadi doimiy boshlang'ich o'zgaruvchilar diapazonini standartlashtirishdir, shunda ularning har biri tahlilga teng hissa qo'shadi. Aniqrog'i, PCA dan oldin standartlashtirishni amalga oshirish juda muhim bo'lgan sabab, ikkinchisi boshlang'ich o'zgaruvchilarning o'zgarishiga nisbatan juda sezgir. Ya'ni, agar boshlang'ich o'zgaruvchilar diapazonlari o'rtasida katta farqlar mavjud bo'lsa, kattaroq diapazonli o'zgaruvchilar kichik diapazonli o'zgaruvchilar ustidan ustunlik qiladi (masalan,

0 dan 100 gacha bo'lgan o'zgaruvchi 0 dan 1 gacha bo'lgan o'zgaruvchidan ustun bo'ladi.), bu noxolis natijalarga olib keladi. Shunday qilib, ma'lumotlarni taqqoslanadigan o'lchovlarga aylantirish bu muammoni oldini oladi.

FOYDALANILGAN ADABIYOTLAR:

1. Markram, H. (2012). The Human Brain Project. *Neuron*, 74(1), 28-45.
2. Kandel, E. (2000). *Principles of Neural Science*. McGraw Hill.
3. Poon, C.-S., & Zhou, K. (2011). Neuromorphic Silicon Neurons and Large-Scale Neural Networks: Challenges and Opportunities. *Frontiers in Neuroscience*, 5(108).
4. Furber, S. (2016). Large-Scale Neuromorphic Computing Systems. *Journal of Neural Engineering*, 13(5), 051001.
5. Tojimamatov, I., & G'ulomjonova, S. (2024). NEYRO KOMPYUTERLAR VA ULARNING ARXITEKTURASI. Development of pedagogical technologies in modern sciences, 3(6), 10-16.
6. Nurmamatovich, T. I. (2024). NEYROKOMPYUTERLARNING AMALIYOTDA QO'LLANILISHI. *worldly knowledge conferens*, 8(1), 287-291.
7. Mirzaakbarov D. D., G'oyibjonova G. KOMPIYUTERLARNING MA'LUMOTLARNI QAYTA ISHLASH BILAN BO'G'LIQ BIROR MASALANI YANGI AXBOROT TEXNOLOGIYASI DOIRASIDA SAMARALI BAJARISH UCHUN QO'LLANILADIGAN IMKONIYATLARI //Models and methods in modern science. – 2023. – T. 2. – №. 12. – C. 90-93.
8. Tojimamatov I., G'ulomjonova S. NEYRO KOMPYUTERLAR VA ULARNING ARXITEKTURASI //Development of pedagogical technologies in modern sciences. – 2024. – T. 3. – №. 6. – C. 10-16.
9. Abdullayev S., Xomidova M. Radial neyron tarmoqlari //Журнал академических исследований нового Узбекистана. – 2024. – Т. 1. – №. 6. – С. 49-52.
10. Yoshua Bengio, Ian Goodfellow, Aaron Courville, "Deep Learning" - (2016)