OMMABOP YUZNI TANISH ALGORITMLARI VA ULARNING ISHLASH PRINSIPLARI

Qodirov R. R

Muhammad al-Xorazmiy nomidagi TATU "Axborot texnologiyalari" kafedrasi stajyor-o'qituvchisi

Annotatsiya: *Respublikamizda axborot texnologiyalari (AT) dam foydalanish davlat va undagi tashkilotlarda qulay ishlash imkonini beradi va samaradorlikni maksimal darajada oshirish uchun xizmat qiladi. Hayotda odamlar qisqa vaqt ichida ko'proq ish qilish yo'llarini izlashadi va bu faqat axborot texnologiyalarini rivojlantirish orqali bo'lishi mumkin. Tezroq aloqa, elektron saqlash va axborotlarni himoya qilish - bu bizning yurtimizda mavjud bo'lishi mumkin bo'lgan afzalliklar. IT bir tomondan yangi, raqobatbardosh ishbilarmonlik bo'lsa va boshqa tomondan kompyuterlar tabiatidagi chuqur o'zgarishlar bilan bog'liq. IT deyarli har qanday ish muhiti bog'liq bo'lgan kompyuter dasturlari bilan bog'liq. Kompyuterlashtirilgan tizimlar juda keng qo'llanilganligi sababli, mamlakatimziga yangi dasturiy ta'minotlarni hamda axborot texnologiyalarini kiritish foydali.*

Kalit so'zlar: *Vektor, PCA, kovirant, dispersiya, Gauss, diskrekt.*

1. PCA - Principal component analysis

1-qadam:

$$\begin{bmatrix} Cov(x,x) & Cov(x,y) & Cov(x,z) \\ Cov(y,x) & Cov(y,y) & Cov(y,z) \\ Cov(z,x) & Cov(z,y) & Cov(z,z) \end{bmatrix}$$

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 muhimligining sababi shundaki, ikkinchisi boshlang'ich o'zgaruvchilarning o'zgarishiga nisbatan juda sezgir. Matematik jihatdan, buni har bir o'zgaruvchining har bir qiymati uchun o'rtacha qiymatni ayirish va standart og'ishga bo'lish orqali amalga oshirish mumkin.

$$z = \frac{value - mean}{standard\ deviation}$$

Standartlashtirish amalga oshirilgandan so'ng, barcha o'zgaruvchilar bir xil o'lchovga aylantiriladi.

2-qadam: Kovariant matritsaasini hisobi

Kovariatsiya matritsasi $p \times p$ nosimmetrik matritsa (bu erda p - o'lchamlar soni), u boshlang'ich o'zgaruvchilarning barcha mumkin bo'lgan juftliklari bilan bog'liq kovariatsiyalarga ega. Misol uchun, 3 o'zgaruvchiga ega 3 o'lchovli ma'lumotlar to'plami uchun x , y va z , kovariatsiya matritsasi buning 3 × 3 matritsasi:

3-qadam: Asosiy komponentlarni aniqlash uchun kovariansiya matritsasining vektorlarini hisoblash.

Xususiy vektorlar va xos qiymatlar chiziqli algebra tushunchalari bo'lib, ma'lumotlarning ***asosiy komponentlarini aniqlash uchun kovariatsiya matritsasidan hisoblashimiz kerak***. Asosiy komponentlar - bu chiziqli birikmalar yoki boshlang'ich o'zgaruvchilarning aralashmasi sifatida tuzilgan yangi o'zgaruvchilar. Bu 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 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 qolgan maksimal ma'lumotni ikkinchisiga va shunga o'xshash narsaga ega bo'lguncha joylashtirishga harakat qiladi.

4-qadam: Xususiy vektor

Oldingi bosqichda ko‘rganimizdek, xos vektorlarni hisoblash va ularni xos qiymatlari bo‘yicha kamayish tartibida tartiblash bizga asosiy komponentlarni ahamiyatlilik tartibida topish imkonini beradi. Ushbu bosqichda biz qiladigan narsa bu komponentlarning barchasini saqlab qolish yoki ahamiyati pastroq bo'lganlarini (past o'z qiymatlari) olib tashlashni tanlash. Shunday qilib, xususiy vektori oddiygina matritsa bo'lib, u ustunlar sifatida biz saqlashga qaror qilgan komponentlarning xos vektorlariga ega. Bu uni o'lchamlarni qisqartirish yo'lidagi birinchi qadamga aylantiradi, chunki agar biz ***n dan faqat p*** o'z vektorlarini (komponentlarini) saqlashni tanlasak, yakuniy ma'lumotlar to'plami faqat ***p*** o'lchamlarga ega bo'ladi.

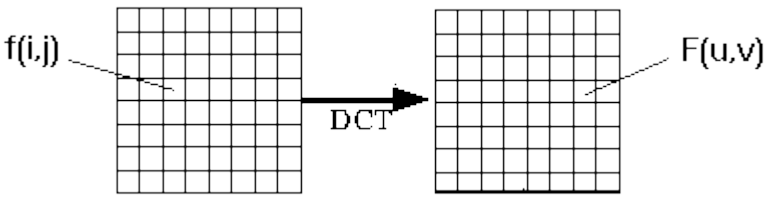
5-qadam: Asosiy komponentlar obyektlari bo'yicha ma'lumotlarni qayta ishlar

Oxirgi bosqich bo'lgan ushbu bosqichda maqsad kovariatsiya matritsasining xos vektorlari yordamida hosil qilingan xususiyat vektoridan foydalanish, ma'lumotlarni asl o'qlardan asosiy komponentlar (shuning uchun asosiy komponentlar tahlili nomi) bilan ifodalanganlarga qayta yo'naltirishdir.). Buni asl ma'lumotlar to'plamini ko'chirishni xususiyat vektorining ko'chirishiga ko'paytirish orqali amalga oshirish mumkin.

$$FinalDataSet = FeatureVector^T * StandardizedOriginalDataSet^T$$

2. DCT - Discrete Cosine Transform

Diskret kosinus o'zgarishi (DCT) tasvirni turli xil ahamiyatga ega bo'lgan qismlarga (yoki spektral pastki diapazonlarga) ajratishga yordam beradi (tasvirning vizual sifatiga nisbatan). DCT diskret Furye konvertatsiyasiga o'xshaydi: u signal yoki tasvirni fazoviy domendan chastota domeniga aylantiradi .



1-rasm. Signalning fazoviy domendan chastota damonga o’zgarishi

DCT ning asosiy bosqichi quyidagilardan iborat:

- Kirish tasviri N va M;
- f(i,j) - i qator va j ustundagi piksel intensivligi ;
- F(u,v) - DCT matritsasining k1 qatori va k2 ustunidagi DCT koeffitsienti.
- Ko'pgina tasvirlar uchun signal energiyasining katta qismi past chastotalarda yotadi; bular DCT ning yuqori chap burchagida ko'rinadi.
- Siqishga erishiladi, chunki pastki o'ng qiymatlar yuqori chastotalarni ifodalaydi va ko'pincha kichikdir - ozgina ko'rinadigan buzilishlar bilan e'tibordan chetda qoladigan darajada kichik.
- DCT kiritish 8 dan 8 gacha butun sonlar massividir. Bu massiv har bir pikselning kulrang shkala darajasini o'z ichiga oladi;
- 8 bitli piksellar 0 dan 255 gacha darajalarga ega.

3. LDA - Linear Discriminant Analysis

Birinchi qadam : **Turli** sinflar orasidagi alohida qobiliyatni, ya'ni turli sinflarning o'rtacha ko'rsatkichlari orasidagi masofani hisoblash, bu sinflar orasidagi dispersiya deb ham ataladi.

$$S_b = \sum_{i=1}^{g} N_i(\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})^T$$

Ikkinchi qadam **: Har bir** sinfning o'rtacha va namunasi orasidagi masofani hisoblash, bu sinf ichidagi dispersiya deb ham ataladi.

$$S_w = \sum_{i=1}^{g}(N_i - 1)S_i = \sum_{i=1}^{g}\sum_{j=1}^{N_i}(x_{i,j} - \bar{x}_i)(x_{i,j} - \bar{x}_i)^T$$

Uchinchi qadam: Sinflar orasidagi farqni maksimal darajada oshiradigan va sinf ichidagi tafovutni kamaytiradigan pastki o'lchamli bo'shliqni yaratish.

P ni Fisher mezoni deb ataladigan pastki o'lchamli fazo proyeksiyasi sifatida qabul qilish.

4. ICA - Independent Component Analysis

ICA da ikkita asosiy taxmin mavjud . Biz ochmoqchi bo'lgan yashirin mustaqil komponentlar bitta, statistik jihatdan mustaqil va ikkita Gauss bo'lmagan bo'lishi kerak. Semantik jihatdan, mustaqil deganda, x haqida ma'lumot sizga y haqida ma'lumot bermaydi va aksincha. Matematik jihatdan bu quyidagicha tarjima qilinadi:

$$p(x, y) = p(x)p(y)$$

Statistik mustaqillikning matematik ta'rifi. Muallif tomonidan tasvir. Bu yerda *p(x) x* ning ehtimollik taqsimotini ifodalaydi. *p(x,y)* x va y ning birgalikda taqsimlanishini ifodalaydi. Gauss bo'lmagan taxmin shunchaki mustaqil komponentlar Gauss bo'lmagan taqsimotlarga ega ekanligini anglatadi, ya'ni u qo'ng'iroq chizig'iga o'xshamaydi. Biz mustaqil omillar to'plamini olish uchun ma'lumotni ajratib olishni xohlaymiz. Agar ochish uchun bir nechta mustaqil ma'lumot generatorlari bo'lmasa, ICA ga ehtiyoj qolmaydi.

Ikkinchi taxminga ehtiyoj matematikada yotadi. ICA mustaqil komponentlarni ochish uchun gausslik **bo'lmaganlik g'oyasidan foydalanadi.** Gausslik emas **tasodifiy o'zgaruvchining taqsimlanishi Gauss bo'lishdan qanchalik *uzoqligini* miqdoriy beradi** . Gausslik bo'lmaganlikning misoli kurtoz va negentropiyadir. Nima uchun bunday o'lchov foydali bo'ladi **Markaziy chegara teoremasidan kelib chiqadi**.

Xususan, ikkita mustaqil tasodifiy o'zgaruvchining yig'indisi asl o'zgaruvchilardan biriga qaraganda Gaussga *yaqinroq* taqsimotga ega ekanligini ko'rsatadigan natija. ICA ma'lumotlarda yashiringan mustaqil komponentlarni ochish uchun ushbu g'oyani, Gauss bo'lmagan o'lchovlarni va Gauss bo'lmagan taxminni birlashtiradi. Buni tasvirlash uchun ikkita o'zgaruvchiga ega bo'lgan ma'lumotlar to'plamini ko'rib chiqing **x_1** va **x_2** . Ushbu o'zgaruvchilar bo'shliqni belgilaydigan asos bo'lib xizmat qiladi, ya'ni biz ularni 2 o'lchamdagi nuqtalarni chizish uchun ishlatishimiz mumkin. Aytaylik, biz ma'lumotlar asosidagi ikkita mustaqil komponentni bilamiz, **s_1** va **s_2** . Ushbu ikki komponent bir xil makonni tasvirlash uchun muqobil asos bo'lib xizmat qiladi. Shuning uchun, bu bo'shliqdagi har qanday **y nuqtasi x_1** va **x_2** o'zgaruvchilarning chiziqli birikmasi yoki s_1 va **s_2** komponentlari sifatida yozilishi mumkin.

$$\mathbf{y} = w_1\mathbf{x}_1 + w_2\mathbf{x}_2$$

O'lchangan signallarning chiziqli birikmasi, ya'ni kirish o'zgaruvchilari. Muallif tomonidan tasvir.

$$\mathbf{y} = a_1\mathbf{s}_1 + a_2\mathbf{s}_2$$

Mustaqil komponentlarning chiziqli birikmasi. Muallif tomonidan tasvir. Markaziy chegara teoremasiga qaytadigan bo'lsak, ikkita tasodifiy o'zgaruvchining yig'indisini taqsimlash har ikkala individual o'zgaruvchiga qaraganda *ko'proq Gauss bo'ladi.* Shunday qilib, a_1 va a_2 nolga teng bo'lmaganda, **y** ning taqsimlanishi **s_1** yoki **s_2 ga** qaraganda *ko'proq Gauss* bo'ladi . Buning teskarisi, agar a_1 yoki a_2 nolga teng bo'lsa, u holda **y ning taqsimoti** avvalgi holatdan *kamroq Gauss* bo'ladi. Va agar s_1 va **s_2 ning Gauss bo'lmagan taxminlari bajarilsa**, u umuman Gauss bo'lmaydi, chunki **y** mustaqil komponentlardan biriga to'liq teng bo'ladi! Boshqacha qilib aytganda, y **ning** nogaussligi , agar u mustaqil komponentlardan biriga to'g'ridan-to'g'ri proportsional bo'lsa, maksimal bo'ladi. Bu bizga ICA ni optimallashtirish muammosi sifatida belgilash imkonini beradi.

$$\max \quad kurt(w_1\mathbf{x}_1 + w_2\mathbf{x}_2)$$

ICAni yagona mustaqil komponent uchun optimallashtirish muammosi sifatida shakllantirish. Muallif tomonidan tasvir. Biz ma'lum bo'lgan kirish o'zgaruvchilarimizning chiziqli birikmasining kurtozini maksimal darajada oshiradigan w_1 va w_2 qiymatlarini topmoqchi bo'lgan joy. Bu w_1 va w_2 optimal qiymatlari mustaqil komponentni belgilaydi.

$$\mathbf{s} = w_1^*\mathbf{x}_1 + w_2^*\mathbf{x}_2$$

ICA optimallashtirish muammosining yechimlari mustaqil komponentlarni belgilaydi. Umuman olganda, biz og'irliklar matritsasi W ni hal qilishimiz mumkin, bu W matritsaning

ko'paytmasining gauss bo'lmaganligini va ma'lumotlar matritsasi X ni maksimal darajada oshiradi.

XULOSA

Biz amal qilgan barcha qadamlarni qisqacha ko'rib chiqamiz:

1. Tasvirning soddalashtirilgan versiyasini yaratish uchun HOG algoritmi yordamida rasmni kodlash. Ushbu soddalashtirilgan tasvirdan foydalanib, tasvirning yuzning umumiy HOG kodlashiga o'xshash qismini topih.

2. Yuzdagi asosiy belgilarni topib, yuzning pozasini aniqlash. Ushbu belgilarni topganimizdan so'ng, ko'zlar va og'iz markazda bo'lishi uchun tasvirni burish uchun ulardan foydalanish.

3. Markazlangan yuz tasvirini yuz xususiyatlarini qanday o'lchashni biladigan neyron tarmoq orqali o'tkazish. Ushbu 128 o'lchovni saqlash.

4. O'tmishda biz o'lchagan barcha yuzlarni ko'rib chiqsak, qaysi odamning o'lchovlari bizning yuzimiz o'lchovlariga eng yaqin ekanligini bilib olish.

FOYDALANILGAN ADABIYOTLAR:

1. H.N. Zaynidinov, O.U. Mallayev. Definition of synchronization processes during parallel signal processing in multicore processors // International 199 Conference on Information Science and Communications Technologies: Applications, Trends and Opportunities, ICISCT 2019 (2019).

2. https://ieeexplore.ieee.org/document/9012006

3. H.N. Zaynidinov, O.U Mallaev and B.B. Anvarjonov. A parallel algorithm for finding the human face in the image // IOP Conf. Series: Materials Science and Engineering, 862 Volume-5, May 2020, 052004,

4. H.N. Zaynidinov, O.U. Mallayev, I. Yusupov. Cubic Basic Splines and Parallel Algorithms // International Journal of Advanced Trends in Computer Science and Engineering (IJATCSE), (Indexed by SCOPUS), ISSN: 2278-3091,9(3).