Algoritmi za prepoznavanje lica

Drmić, Dino

SVEUČILIŠTE U ZAGREBU
PRIRODOSLOVNO–MATEMATIČKI FAKULTET
MATEMATIČKI ODSJEK

Dino Drmić

ALGORITMI ZA PREPOZNAVANJE LICA

Diplomski rad

Voditelj rada:
izv. prof. dr. sc. Matija Kazalicki

Zagreb, lipanj 2022.

Ovaj diplomski rad obranjen je dana ____________________ pred ispitnim povjerenstvom u sastavu:

1. ______________________________, predsjednik
2. ______________________________, član
3. ______________________________, član

Povjerenstvo je rad ocijenilo ocjenom ________________.

Potpisi članova povjerenstva:

1. ______________________________
2. ______________________________
3. ______________________________

Sadržaj

Uvod

Numeričke metode imaju široku primjenu u području podatkovne znanosti i mogu imati prednost u odnosu na neke druge metode kao što su duboki modeli. Za dobar model neuronske mreže u klasifikaciji slika potreban je velik skup podataka za treniranje i to je nedostatak u odnosu na numeričke metode koje u kontekstu prepoznavanja lica trebaju svega nekoliko slika za treniranje. Cilj ovog rada je predstaviti te metode koje primjenjuju različite matrične i tenzorske dekompozicije i faktorizacije.

Za razumijevanje takvih metoda je potrebno razumjeti i koncept tenzora. Tenzori su višedimenzionalni nizovi prikladni za pohranjivanje podataka s više informacija. Vektori i matrice predstavljaju tenzore reda 1, odnosno tenzore reda 2 i oni nekad nisu idealni za učitavanje skupa podataka koji se sastoje više od dvije vrste informacija. Naprimjer, sliku u boji rezolucije 200×200 koju želimo učitati ne možemo pohraniti u obliku matrice, jer svaki piksel je predstavljen u obliku uređene trojke (R, G, B), odnosno boju svakog piksela možemo reprezentirati kombinacijom crvene, zelene i plave boje. U tom slučaju tenzor dimenzije $200 \times 200 \times 3$ predstavlja dobar izbor za pohranjivanje podataka. Sliku možemo matricizirati tako da ju prevedemo iz RGB koordinata u greyscale kanal i u tom slučaju dobivamo crno-bijelu sliku. No, ipak i kada radimo s matriciziranim slikama javlja se potreba za tenzorima višeg reda u slučajevima kada želimo pohraniti više slika u jedan niz podataka.

U ovom radu bavit ćemo se pohranjivanjem skupa podataka (slika) u matrice i tenzore te objasniti kako možemo primjenom njihovih različitih dekompozicija modelirati algoritme za prepoznavanje lica ljudi.

Prvu metodu koju obrađujemo je poznata pod imenom PCA ili Eigenfaces (mi ćemo ju zvati samo PCA). Svaka slika sadrži puno numeričkih podataka i računanje udaljenosti jedne od druge je skupa operacija. Metoda PCA je poznata metoda za redukciju prostora u kojem prikazujemo slike u kojoj ključnu ulogu igraju svojstveni vektori i svojstvene vrijednosti. U ovom poglavlju iskazati ćemo jedan od ključnih teorema numeričke matematike, a radi se o dekompoziciji matrice na singularne vrijednosti ili SVD (*eng.* Singular Value Decomposition) te ćemo vidjeti njegovu primjenu u algoritmu PCA.

Druga metoda se zove Tensorfaces i uvodi korištenje tenzora kao strukture za spremanje slika. Metoda Tensorfaces je multilinearni ekvivalent PCA metode i ustvari predstavlja bolju alternativu za prepoznavanje lica. Objasniti ćemo što su tenzori, definirati osnovne tenzorske operacije te iskazati generalizirani SVD kojeg zovemo singularna dekompozicija višeg reda ili HOSVD (*eng.* High Order Singular Value Decomposition) i primjeniti ga u algoritmu Tensorfaces.

Oba algoritma testiramo na istom skupu fotografija iz baze *Yale Face Database*. Na kraju je provedena usporedba i komentar zašto tenzorski pristup ima prednost nad matričnim PCA pristupom.

Poglavlje 1

Matrična PCA metoda

Jedna od najpoznatijih metoda u klasifikaciji slika je metoda analize glavnih komponenti ili skraćeno PCA (*eng.* Principal Component Analysis). Za ovu metodu ključno je razumjeti jedan moćan alat numeričke matematike koji će također biti temelj i tenzorskih metoda, a riječ je o dekompoziciji matrice na singularne vrijednosti kojeg ćemo odsada zvati SVD.

Prije objašnjenja SVD-a i PCA metode potrebno je navesti osnovne pojmove i rezultate iz vjerojatnosti i linearne algebre potrebne za razumijevanje daljnjeg teksta.

1.1 Osnovni pojmovi

Definicija 1.1.1. *Neka je* $(\Omega, \mathcal{F}, P)$ *vjerojatnosni prostor. Funkciju* $X : \Omega \to \mathbb{R}^p$ *za koju je* $X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{F}, \forall B \in \mathcal{B}(\mathbb{R}^p)$ *zovemo* p*-dimenzionalnim slučajnim vektorom na* $(\Omega, \mathcal{F}, P)$*. U slučaju* $p = 1$ *zovemo ju slučajnom varijablom na* $(\Omega, \mathcal{F}, P)$*.*

Slijedi jedan od osnovnih rezultata vezanih za slučajne vektore.

Propozicija 1.1.2. *Neka je* $X : \Omega \to \mathbb{R}^p, X = (X_1, X_2, \ldots, X_p)$*. Tada je* X *slučajni vektor ako i samo je* X_k *slučajna varijabla za svaki* $k = 1, 2, \ldots, p$*.*

Dokaz je dan u [6].

Iz prethodne propozicije slijedi da je p-dimenzionalan slučajan vektor uređena p-torka slučajnih varijabli.

Definicija 1.1.3. *Neka su* X_1, X_2 *slučajne varijable. Očekivanje*

$$\mathbb{E}[(X_1 - \mathbb{E}[X_1])(X_2 - \mathbb{E}[X_2])]$$

(ako postoji) zovemo kovarijancom slučajnih varijabli X_1 i X_2 i označavamo ju s $Cov(X_1, X_2)$. Specijalno, ako je $X_1 = X_2$, očekivanje

$$\mathbb{E}[(X_1 - \mathbb{E}[X_1])^2]$$

zovemo varijancom slučajne varijable X_1 i označavamo ju s $Var(X_1)$.

Primjetimo da vrijedi $Cov(X_1, X_2) = Cov(X_2, X_1)$ i $Cov(X_1, X_1) = Var(X_1)$.

Definicija 1.1.4. *Kažemo da su slučajne varijable X_1 i X_2 nekorelirane ako vrijedi $Cov(X_1, X_2) = 0$.*

Nakon definicija varijance i kovarijance sada možemo definirati matricu kovarijanci slučajnog vektora X.

Neka je $X = (X_1, \dots, X_p)$ p-dimenzionalan slučajni vektor takav da postoji $\mathbb{E}[X_i^2]$ za $i = 1, \dots, p$. Očekivanja slučajnih varijabli $\mathbb{E}[X_1], \dots, \mathbb{E}[X_p]$ možemo zapisati u vektorskom obliku $\mathbb{E}[X] = [\mathbb{E}[X_1], \dots, \mathbb{E}[X_p]]^T$. Matricu $Cov(X) = \mathbb{E}[(X - \mathbb{E}[X])(X - \mathbb{E}[X])^T]$ definiramo kao matricu kovarijanci slučajnog vektora X. Vrijedi

$$Cov(X) = \begin{bmatrix} \mathbb{E}[(X_1 - \mathbb{E}[X_1])^2] & \dots & \mathbb{E}[(X_1 - \mathbb{E}[X_1])(X_p - \mathbb{E}[X_p])] \\ \vdots & \ddots & \vdots \\ \mathbb{E}[(X_p - \mathbb{E}[X_p])(X_1 - \mathbb{E}[X_1])] & \dots & \mathbb{E}[(X_p - \mathbb{E}[X_p])^2] \end{bmatrix}$$

$$= \begin{bmatrix} Var(X_1) & \dots & Cov(X_1, X_p) \\ \vdots & \ddots & \vdots \\ Cov(X_1, X_p) & \dots & Var(X_p) \end{bmatrix}. \tag{1.1}$$

Primjetimo da je matrica kovarijanci simetrična matrica.

Definicija 1.1.5. *Za simetričnu matricu $A \in \mathbb{R}^{n \times n}$ kažemo da je pozitivno semidefinitna ako za sve $x \in \mathbb{R}^n$ vrijedi $x^T A x \geqslant 0$.*

Može se pokazati da je matrica kovarijanci pozitivno semidefinitna. Neka je $x \in \mathbb{R}^p$ proizvoljan vektor. Vrijedi:

$$x^T Cov(X) x = x^T \mathbb{E}[(X - \mathbb{E}[X])(X - \mathbb{E}[X])^T] x = \mathbb{E}[x^T (X - \mathbb{E}[X])]^2 \geqslant 0.$$

Navedimo još neke osnovne pojmove i rezultate iz linearne algebre.

Definicija 1.1.6. *Vektor $v \in \mathbb{R}^n$ je svojstveni vektor matrice $A \in \mathbb{R}^{n\times n}$ s pripadnom svojstvenom vrijednošću $\lambda \in \mathbb{R}$ ako je*

$$Av = \lambda v, \quad v \neq 0.$$

Propozicija 1.1.7. *Neka je $A \in \mathbb{R}^{n\times n}$ pozitivno semidefinitna matrica. Tada su njene svojstvene vrijednosti nenegativne. Svojstveni vektori koji pripadaju različitim svojstvenim vrijednostima matrice A su ortogonalni.*

Dokaz. Dokaz nejednakosti svojstvenih vrijednosti je dan u [1]. Pokažimo da vrijedi da su svojstveni vektori s različitim pripadnim svojstvenim vrijednostima međusobno ortogonalni.

Pretpostavimo da je matrica A pozitivno semidefinitna i ima dva svojstvena vektora u i v čije su pripadne svojstvene vrijednosti μ i λ različite. Tada vrijedi:

$$Au = \mu u, \quad Av = \lambda v.$$

Zbog simetričnosti matrice A vrijedi

$$u^T Av = u^T A^T v = (Au)^T v$$

Vrijedi i

$$u^T Av = u^T(\lambda v) = \lambda u^T v.$$

Iz prethodnih jednakosti slijedi $(\mu - \lambda)u^T v = 0$. S obzirom da smo pretpostavili da je $\mu \neq \lambda$, slijedi da je $u^T v = 0$, odnosno u i v su ortogonalni. □

Iz prethodne prepozicije slijedi bitan zaključak da ako matrica kovarijanci dana s (1.1) ima različite svojstvene vrijednosti, tada pripadni svojstveni vektori su ortogonalni.

Definirajmo još i Frobeniusovu normu koja će biti bitna kao uvjet kod klasifikacije slika.

Definicija 1.1.8. *Frobeniusova norma matrice $A \in \mathbb{R}^{m\times n}$ je matrična norma definirana kao korijen zbroja kvadrata apsolutnih vrijednosti elemenata matrice*

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{i,j}|^2}$$

1.2 Matrični SVD

Teorem 1.2.1. *(SVD) Neka su m i n prirodni brojevi takvi da je $m \geqslant n$ i $A \in \mathbb{R}^{m\times n}$ proizvoljna realna matrica. Tada postoji dekompozicija*

$$A = U\Sigma V^T \tag{1.2}$$

gdje su $U \in \mathbb{R}^{m\times m}$ i $V \in \mathbb{R}^{n\times n}$ ortogonalne matrice, a $\Sigma = diag(\sigma_1, \sigma_2, \ldots, \sigma_n)$ dijagonalna matrica tako da vrijedi $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_n \geqslant 0$. [7]

Dokaz. Ako je $A = 0$, tada je $\Sigma = 0$, a U i V su proizvoljne ortogonalne matrice. Prepostavimo da je $A \neq 0$.

Dokaz provodimo indukcijom po m i n. Iz pretpostavke o postojanju SVD-a za matrice dimenzija $(m-1)\times(n-1)$ dokazat ćemo postojanje SVD-a za $m\times n$ matrice.

Jer je $m \geqslant n$, baza indukcije je $n = 1$. Tada imamo jednostupčanu matricu A koju možemo zapisati u obliku:

$$A = U\Sigma V^T$$

gdje je

$$U = \frac{A}{\|A\|_2}, \quad \Sigma = \|A\|_2, \quad V = 1$$

pa tvrdnja vrijedi za $n = 1$ i bilo koji $m \geqslant 1$.

Prepostavimo da tvrdnja vrijedi za $A \in \mathbb{R}^{(m-1)\times(n-1)}$. Za korak indukcije izaberemo vektor v, takav da je $\|v\|_2 = 1$ i na njemu se dostiže maksimum 2-norme za A, tj. vrijedi:

$$\|A\|_2 = \max_{\|x\|_2=1} \|Ax\|_2 = \|Av\|_2.$$

Definiramo jedinični vektor

$$u = \frac{Av}{\|Av\|_2}.$$

Vektore u i v dopunimo matricama $\tilde{U}$ i $\tilde{V}$, tako da matrice

$$U_0 = [u, \tilde{U}], \qquad V_0 = [v, \tilde{V}]$$

budu ortogonalne matrice reda m, odnosno reda n. Sada računamo

$$U_0^T A V_0 = \begin{bmatrix} u^T \\ \tilde{T}^T \end{bmatrix} A[u, \tilde{V}] = \begin{bmatrix} u^T Av & u^T A\tilde{V} \\ \tilde{U}^T Av & \tilde{U}^T A\tilde{V} \end{bmatrix}.$$

Po definiciji vektora u i v vrijedi

$$u^T Av = \frac{v^T A^T}{\|Av\|_2} Av = \frac{\|Av\|_2^2}{\|Av\|_2} = \|Av\|_2 = \|A\|_2 \coloneqq \sigma.$$

Zbog ortogonalnosti matrice U_0 vrijedi $\tilde{U}^T u = 0$, pa vrijedi

$$\tilde{U}^T Av = \tilde{U}^T u \|Av\|_2 = 0.$$

Tvrdimo da je $u^T A\tilde{V} = 0$. Označimo $A_1 = U_0^T AV_0$, $w^T = u^T A\tilde{V}$ i $B = \tilde{U}^T A\tilde{V}$. Onda vrijedi

$$A_1 = U_0^T AV_0 = \begin{bmatrix} \sigma & w^T \\ 0 & B \end{bmatrix}.$$

Zbog invarijantnosti 2-norme na ortogonalnost dobivamo jednakost

$$\sigma = \|A\|_2 = \|U_0^T AV_0\|_2 = \|A_1\|_2.$$

Za proizvoljni vektor $z \neq 0$ vrijedi

$$\|A_1\|_2 = \max_{x \neq 0} \frac{\|A_1 x\|_2}{\|x\|_2} \geqslant \frac{\|A_1 z\|_2}{\|z\|_2}$$

odnosno $\|A_1\|_2 \|z\|_2 \geqslant \|A_1 z\|_2$. Izaberimo

$$z = \begin{bmatrix} \sigma \\ w \end{bmatrix}.$$

Tada vrijedi

$$\begin{aligned} \|A_1\|_2^2 \|z\|_2^2 = \|A_1\|_2^2 (\sigma^2 + \|w\|_2^2) \geqslant \|A_1 z\|_2^2 &= \left\| A_1 \begin{bmatrix} \sigma \\ w \end{bmatrix} \right\|_2^2 \\ &= \left\| \begin{bmatrix} \sigma & w^T \\ 0 & B \end{bmatrix} \begin{bmatrix} \sigma \\ w \end{bmatrix} \right\|_2^2 = (\sigma^2 + w^T w)^2 + \|Bw\|_2^2 \\ &\geqslant (\sigma^2 + \|w\|_2^2)^2. \end{aligned}$$

iz čega slijedi

$$\|A_1\|_2^2 (\sigma^2 + \|w\|_2^2) \geqslant (\sigma^2 + \|w\|_2^2)^2.$$

Dijeljenjem s $\sigma^2 + \|w\|_2^2$ dobivamo

$$\sigma^2 = \|A\|_2^2 = \|A_1\|_2^2 \geqslant \sigma^2 + \|w\|_2^2$$

što je moguće samo za $w = 0$. Drugim riječima, vrijedi

$$U_0^T A V_0 = \begin{bmatrix} \sigma & 0 \\ 0 & B \end{bmatrix}.$$

Po pretpostavci indukcije vrijedi

$$B = U_1 \Sigma_1 V_1^T$$

pa vrijedi

$$U_0 A V_0 = \begin{bmatrix} \sigma & 0 \\ 0 & U_1 \Sigma_1 V_1^T \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & U_1 \end{bmatrix} \begin{bmatrix} \sigma & 0 \\ 0 & \Sigma_1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & V_1 \end{bmatrix}^T$$

odakle odmah slijedi tvrdnja, jer ortogonalne matrice čine multiplikativnu grupu. Silazni poredak elemenata na glavnoj dijagonali od Σ_1 postižemo primjenom matrica permutacije koje su također ortogonalne. □

Definicija 1.2.2. *Stupce matrice $U = [u_1, u_2, \ldots, u_m]$ nazivamo lijevim singularnim vektorima, a stupce matrice $V = [v_1, v_2, \ldots, v_n]$ desnim singularnim vektorima. Brojevi σ_i zovu se singularne vrijednosti.*

Napomena 1.2.3. *U slučaju da matrica A ima veći broj stupaca od redaka, odnosno da vrijedi $m < n$, tada SVD definiramo za matricu A^T.*

1.3 Metoda PCA

Metoda analize glavnih komponenti (PCA) je jedna od standardnih metoda u klasifikaciji i kompresiji slika. Radi se o metodi koja smanjuje dimenziju skupova podataka čime proces klasifikacije postaje efikasniji i brži uz minimalne posljedice na točnost prepoznavanja. Dakle, transformiramo veliki skup varijabli u manji skup koji još uvijek sadrži većinu informacija, ali je lakši za analizu i vizualizaciju te ga algoritmi strojnog učenja brže procesiraju.

Bitna karakteristika skupa podataka je veza između podataka. Ponekad su varijable međusobno visoko korelirane te sadrže suvišne informacije. U takvoj situaciji korisno je pretvoriti taj skup varijabli u novi skup nekoreliranih varijabli koje zovemo glavne komponente. Glavne komponente su međusobno nekorelirane linearne kombinacije originalnih varijabli i to takve da je prva glavna komponenta linearna kombinacija slučajnih varijabli koja je najveće varijance, druga glavna komponenta linearna kombinacija slučajnih varijabli koja je druge najveće varijance i tako dalje. Dakle, prva glavna komponenta će objašnjavati što je više moguće varijacije podataka.

Ideja je reducirati početni skup podataka koji se sastoji od n mjerenja na p ulaznih varijabli na skup od n mjerenja na k glavnih komponenata. Glavne komponente reprezentiraju smjer maksimalne varijacije i omogućuju jednostavniji opis kovarijantne strukture. One ovise samo o matrici kovarijanci polaznih slučajnih varijabli $X_1, \ldots, X_p$. Neka je $X^T = [X_1, \ldots, X_p]$ slučajni vektor i neka je C pripadna matrica kovarijanci sa svojstvenim vrijednostima $\lambda_1 > \lambda_2 > \cdots > \lambda_p > 0$. U praksi je vrlo rijetko da su neke od svojstvenih vrijednosti jednake ili da su jednake nuli, pa ćemo odsada raditi s pretpostavkom da su sve svojstvene vrijednosti pozitivne i međusobno različite.

Promotrimo linearne kombinacije

$$\begin{aligned}
u_1 &= a_1^T X = a_{11}X_1 + a_{12}X_2 + \cdots + a_{1p}X_p \\
u_2 &= a_2^T X = a_{21}X_1 + a_{22}X_2 + \cdots + a_{2p}X_p \\
&\vdots \qquad\qquad\quad \vdots \\
u_p &= a_p^T X = a_{p1}X_1 + a_{p2}X_2 + \cdots + a_{pp}X_p.
\end{aligned}$$

Slijedi

$$\begin{aligned}
Var(u_i) = Var(a_i^T X) = \mathbb{E}[a_i^T X X^T a_i] = a_i^T C a_i, \qquad i = 1, 2, \ldots, p \\
Cov(u_i, u_k) = a_i^T C a_k, \qquad i, k = 1, 2, \ldots, p.
\end{aligned}$$

Izvedimo prvu glavnu komponentu. Želimo izabrati jedinični vektor a_1 i zamijeniti svaki podatak X s njegovom projekcijom na taj vektor, $a_1^T X$. Prilikom odabira vektora a_1 želimo zadržati što je više moguće varijacije podataka, odnosno želimo maksimizirati varijancu novih podataka $a_1^T X$. S obzirom da se varijanca $Var(u_i) = a_i^T C a_i$ može povećati množenjem vektora konstantom, pažnja se ograničava na vektore koeficijenata duljine jedan. Dakle, prva glavna komponenta je linearna kombinacija $u_1 = a_1^T X$ koja maksimizira

$$Var(a_1^T X) = a_1^T C a_1$$

uz dan uvjet

$$a_1^T a_1 = 1.$$

Izvedimo drugu glavnu komponentu. Tražimo smjer ortogonalan na prvu glavnu komponentu koji će zadržati što je više moguće varijacije podataka. Projekcije podataka na tu komponentu nemaju varijancu u smjeru prve glavne komponente, što znači da tražimo jedinični vektor a_2 takav da je $Cov(a_1^T X, a_2^T X) = 0$. Slijedi da je druga glavna komponenta linearna kombinacija $u_2 = a_2^T X$ koja maksimizira

$$Var(u_2) = a_2^T C a_2$$

uz uvjete

$$a_2^T a_2 = 1, \qquad Cov(a_1^T X, a_2^T X) = 0$$

Na isti način izvodimo i ostale glavne komponente. Općenito, i-ta glavna komponenta je linearna kombinacija $u_i = a_i^T X$ koja maksimizira

$$Var(u_i) = a_i^T C a_i$$

uz uvjete

$$a_i^T a_i = 1, \quad Cov(a_i^T X, a_k^T X) = 0, \quad k < i$$

Dakle, glavne komponente su one linearne kombinacije čije su varijance u padajućem poretku (prva glavna komponenta je linearna kombinacija s najvećom varijancom).

Dolazimo do sljedećeg teorema koji veže glavne komponente sa svojstvenim vektorima i pripadnim svojstvenim vrijednostima matrice kovarijanci C.

Teorem 1.3.1. *Neka je dan slučajni vektor $X^T = [X_1, X_2, \ldots, X_p]$ i neka je C njegova matrica kovarijanci. Neka su*

$$(\lambda_1, e_1), (\lambda_2, e_2), \ldots, (\lambda_p, e_p)$$

parovi svojstvenih vrijednosti i svojstvenih vektora matrice C pri čemu vrijedi $\lambda_1 > \lambda_2 > \cdots > \lambda_p > 0$. Tada je i-ta glavna komponenta dana s

$$u_i = e_i^T X = e_{i1} X_1 + e_{i2} X_2 + \cdots + e_{ip} X_p, \quad i = 1, 2, \ldots, p.$$

Tada vrijedi i

$$\begin{aligned} Var(u_i) &= e_i^T C e_i = \lambda_i, \qquad i = 1, 2, \ldots, p \\ Cov(u_i, u_k) &= Cov(e_i^T C e_k) = 0, \qquad i \neq k. \end{aligned}$$

Dokaz. Izraz

$$\max_{a \neq 0} \frac{a^T C a}{a^T a}$$

je dostignut za $a = e_1$, tj. vrijedi

$$\max_{a \neq 0} \frac{a^T C a}{a^T a} = \frac{e_1^T C e_1}{e_1^T e_1} = \lambda_1.$$

Svojstveni vektori su normalizirani, odnosno vrijedi $e_1^T e_1 = 1$, pa slijedi

$$\lambda_1 = \frac{e_1^T C e_1}{e_1^T e_1} = e_1^T C e_1 = Var(u_1)$$

Slično

$$\max_{a \perp e_1, e_2, \ldots, e_k} \frac{a^T C a}{a^T a} = \frac{e_{k+1}^T C e_{k+1}}{e_{k+1}^T e_{k+1}} = \lambda_{k+1}.$$

uz uvjet $e_{k+1}^T e_i = 0$ za $i = 1, 2, \ldots, k$. Vrijedi

$$\lambda_{k+1} = \frac{e_{k+1}^T C e_{k+1}}{e_{k+1}^T e_{k+1}} = e_{k+1}^T C e_{k+1} = Var(u_{k+1}).$$

Jer je matrica C pozitivno semidefinitna, pripadni svojstveni vektori su okomiti.

Vrijedi

$$Ce_k = \lambda_k e_k.$$

Množenjem gornje jednadžbe s $e_i^T, i \neq k$ dobivamo

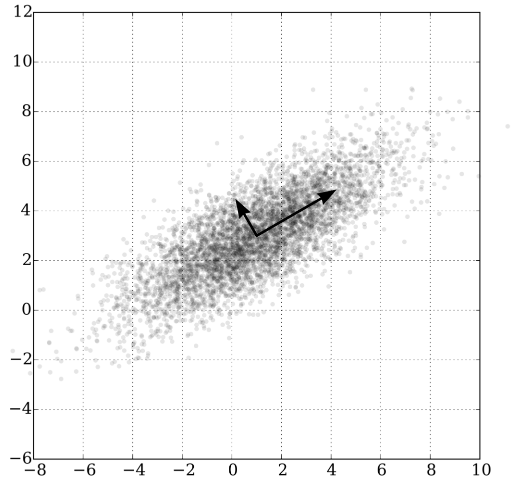
$$Cov(u_i, u_k) = e_i^T C e_k = e_i^T \lambda_k e_k = \lambda_k e_i^T e_k = 0.$$

$\square$

Napomena 1.3.2. *Često se pod i-tom glavnom komponentom misli na svojstveni vektor e_i, jer vektori e_i i $e_i^T X$ imaju isti smjer i orijentaciju (leže na istom pravcu) kojim se želi izraziti i-ta glavna komponenta. Normiranjem vektora $e_i^T X$ dobivamo vektor e_i.*

Već smo spomenuli algebarsku interpretaciju glavnih komponenti, a to je da su one linearne kombinacije p slučajnih varijabli. Geometrijski su te linearne kombinacije zapravo koordinatne osi novog koordinatnog sustava dobivenog rotacijom oko starog s glavnim komponentama kao koordinatnim osima.

Generalno nije praktično računati matricu kovarijanci skupa podataka kako bi našli svojstvene vektore. Ako je svaki podatak dimenzije $l \times n$, tada je veličina matrice kovarijanci $(l \times n) \times (l \times n)$. Postavlja se pitanje kako izračunati svojstvene



Slika 1.1: Glavne komponente skupa podataka u 2-dimenzionalnom prostoru

vektore matrice C bez da računamo matricu C? Odgovor ćemo dati u nastavku u kojem slijedi objašnjenje PCA metode u kontekstu prepoznavanja lica [3].

Neka je $\{X_1, \ldots, X_m\}$ kolekcija slika $l \times n$, gdje je l broj redova slike, a n broj stupaca slike. Svaka slika X_j je 2-D slika (matrica) koju možemo vektorizirati u 1-D sliku (vektor). Definiramo matricu A tako da vrijedi da je j-ti stupac $A(:j)$ jednak vektoriziranoj slici X_j, odnosno $A(:,j) = vec(X_j)$.

Uzoračku sredinu slika (prosječnu sliku) računamo kao

$$M = \frac{1}{m} \sum_{j=1}^{m} A(:,j).$$

Sada možemo matricu A prikazati u srednje-devijacijskoj formi ako svakom stupcu oduzmemo prosječnu sliku

$$\tilde{A}(:,j) = A(:,j) - M, \qquad j = 1, 2, \ldots, m. \tag{1.3}$$

Kovarijanca slučajnih varijabli x i y može se empirijski definirati kao

$$Cov(x,y) = \frac{1}{n-1} \sum_{i=1}^{m} (x_i - \overline{x})(y_i - \overline{y})$$

gdje je m veličina uzorka i $\overline{x}, \overline{y}$ su srednje vrijednosti observacija za varijable x i y. Slijedi da za matrično prikazan skup podataka A možemo izračunati njegovu matricu kovarijanci

$$C = \frac{1}{m-1} \sum_{i=1}^{m} (A(i,:) - M(i))(A(i,:) - M(i))^T. \tag{1.4}$$

Koristeći (1.3), jednakost (1.4) dobiva slijedeći oblik

$$C = \frac{1}{m-1} \tilde{A}\tilde{A}^T. \tag{1.5}$$

Matrica C je simetrična i pozitivno semidefinitna i može se ortogonalno dijagonalizirati, odnosno možemo ju zapisati kao

$$C = UDU^T$$

gdje je U ortogonalna matrica i $D = diag(\lambda_1, \lambda_2, \dots)$ dijagonalna matrica koja sadrži svojstvene vrijednosti matrice C i vrijedi $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant 0$. Stupci u_i od matrice U su svojstveni vektori od C i oni predstavljaju glavne komponente. S obzirom da je λ_1 najveća svojstvena vrijednost, vektor u_1 označava smjer u kojem skup podataka ima maksimalnu varijaciju.

Kao što je već rečeno, nepraktično je računati matricu kovarijanci. Praktičniji pristup je računanje SVD-a matrice

$$\tilde{A} = U\Sigma V^T.$$

Vrijedi

$$C = \frac{1}{m-1} \tilde{A}\tilde{A}^T = \frac{1}{m-1} U\Sigma^2 U^T.$$

Ortogonalna dijagonalizacija matrice C, gdje dijagonalna matrica D ima elemente na dijagonali u padajućem poretku, je jedinstvena. Iz prethodnog slijedi

$$D = \frac{1}{m-1} \Sigma^2,$$

odnosno singularne vrijednosti s_{ii} od $\tilde{A}$ povezane su sa svojstvenim vrijednostima od C

$$s_{ii} = \sqrt{(m-1)\lambda_i}.$$

Također slijedi i ključan zaključak da lijeva singularna matrica U od matrice $\tilde{A}$ sadrži svojstvene vektore matrice kovarijanci C.

Postavlja se pitanje kako odabrati koliko k glavnih komponenti želimo uzeti za redukciju prostora. Postoji više metoda odabira broja komponenti. Najpoznatija metoda je Kaiserov kriterij koji će zadržati sve one komponente čija je varijanca veća od 1.

Broj komponenti možemo odabrati i na drugi način. Može se uzeti onoliko komponenti čije će varijance činiti $80\% - 90\%$ ukupne varijance. Proporciju zbroja varijanci k komponenti definiramo izrazom $\frac{\lambda_1+\cdots+\lambda_k}{\lambda_1+\cdots+\lambda_p}, k < p$.

Iz ortogonalnosti matrice U slijedi

$$\tilde{A} = UU^T\tilde{A}.$$

Nakon što pažljivo odaberemo indeks smanjenja k, imamo

$$\tilde{A}(:,j) \approx U_k U_k^T \tilde{A}(:,j) \tag{1.6}$$

gdje je $U_k = U(:, 1:k)$ za neki $k < m$. Oznaka $U(:, 1:k)$ označava matricu koju dobijemo kada iz matrice U izostavimo zadnjih $m - k$ stupaca. Aproksimacija je dobra zato što su singularne vrijednosti $s_{ii}, i > k$, "male" u odnosu na m i većina varijacije podataka je dobro opisana u k smjerova koji odgovaraju prvim najvećim k singularnim vrijednostima. Dakle, iz matrice U možemo izostaviti zadnjih $m - k$ stupaca jer oni ne nose značajnu količinu informacija o slikama u odnosu na prvih k stupaca.

Matrica $U_k U_k^T$ je ortogonalni projektor na $\{u_1, \ldots, u_k\}$, prostor razapet s prvih k glavnih komponenti. Primjetimo da se svaka slika (s oduzetom sredinom) $\tilde{A}(:,j)$ može aproksimirati kao linearna kombinacija glavnih komponenti, pa (1.6) možemo zapisati i kao:

$$\tilde{A}(:,j) \approx U_k(U_k^T\tilde{A}(:,j)) = \sum_{i=1}^{k} c_i u_i, \qquad c_i = u_i^T \tilde{A}(:,j).$$

Neka je X testna slika. Sliku moramo dovesti u istu formu kao što je kolekcija slika u matrici $\tilde{A}$, dakle potrebno ju je vektorizirati i oduzeti joj sredinu kolekcije slika $\tilde{X} = vec(X) - M$. Vektor $\tilde{X}$ želimo prikazati kao linearnu kombinaciju glavnih komponenti $u_1, \ldots, u_k$, a koeficijente možemo dobiti izrazom $U_k^T\tilde{X}$. Klasificiranje osobe na slici se radi tako da se nađe onaj stupac $\tilde{A}(:,j)$ čiji koeficijenti c_i su najbliži koeficijentima od $\tilde{X}$ u Frobeniusovoj normi. Ako je norma manja od danog iznosa tolerancije, osoba na testnoj slici X se klasificira kao osoba na slici $A(:,j)$. Vrijednost tolerancije namještamo eksperimentima.

Algoritam 1 - PCA metoda

Ulaz: Slike za treniranje: $X_i,\ i = 1, 2, ..., m$, Testna slika: X, Indeks smanjenja k, Tolerancija tol

Izlaz: 1 (ako je osoba u bazi), 0 (ako osoba nije u bazi)

for $i = 1$ **to** m **do**
  $A(:, i) \leftarrow$ vektorizirani X_i
end for
$M \leftarrow$ prosječna slika (vektorizirana)
$\tilde{A} \leftarrow$ srednje-devijacijska forma od A
$U \leftarrow$ lijevi singularni vektori od $\tilde{A}$
$G \leftarrow U(:, 1\!:\!k)^T \tilde{A}$
$X \leftarrow$ vektorizirani X
$\tilde{X} \leftarrow X - M$
$c \leftarrow U(:, 1\!:\!k)^T \tilde{X}$
for $j = 1$ **to** m **do**
  if $||c - G(:, j)||_F < tol$ **then**
    return 1
  end if
end for
return 0

Metodu PCA smo obradili s pretpostavkom da su svojstvene vrijednosti matrice kovarijanci različite od nule. Što ako je jedan ili više svojstvenih vrijednosti jednako nuli?

Ukoliko imamo da je q svojstvenih vrijednosti matrice C jednako nuli, onda je rang matrice $p - q$ umjesto q. Glavna komponenta koja ima varijancu 0 definira linearnu vezu između elemenata vektora X. Postojanost ovakve veze implicira da se vrijednost jedne varijable može izračunati pomoću drugih varijabli. Dakle, možemo smanjiti broj varijabli s p na $p - q$ bez gubitka informacija te za taj slučaj provesti izračun glavnih komponenti. U praksi je jednakost svojstvenih vrijednosti nuli vrlo rijedak slučaj.

1.4 Obrada slika i klasifikacija

Za testiranje svih algoritama uzete su slike iz baze podataka *Yale Face Database* [5]. Odabrano je 10 različitih osoba i za svaku osobu po 7 različitih ekspresija lica. Dakle, sveukupno imamo 70 slika.

Slika 1.2: Primjer slika svih ekspresija za osobu 1

Slike su izrezane te smanjene na rezoluciju 90×90. Također su prilikom učitavanja prevedene iz RGB koordinata u *greyscale* kanal. Posljednji korak je normalizacija vrijednosti piksela, tako da je minimalna vrijednost piksela 0 (za crnu boju), maksimalna vrijednost 1 (za bijelu boju), a sve ostalo između su nijanse sive boje.

Potrebno je određeni dio slika odabrati za bazu, odnosno za skup za treniranje. Za bazu uzete su prve tri ekspresije od prvih pet osoba (baza je označena zelenim pravokutnikom na slici 1.3). Slike je potrebno spremiti u trening matricu dimenzija 8100×15, gdje svaki stupac matrice je slika pohranjena u vektor.

Ostale slike su testne slike na kojima ispitujemo efikasnost algoritma. Cilj algoritma je vratiti vrijednost 1 ako danu testnu sliku klasificira kao jednu od osoba u bazi ili mora vratiti vrijednost 0 ako testna slika nije klasificirana kao jedna od osoba u bazi.

Do određenog iznosa tolerancije dolazimo eksperimentima, imajući na umu da je pogreška manja ako algoritam ne prepozna sliku osobe koja je u bazi u odnosu na pogrešku kada algoritam prepozna sliku osobe koja nije u bazi. Algoritam je bio najefikasniji za nivo tolerancije 10 i u tom slučaju je prepoznao sve testne slike koje su trebale biti prepoznate (osim jedne) te je odbacio sve testne slike koje su trebale biti odbačene.

Slika 1.3: Rezultati PCA algoritma - zelenom kvačicom su označene slike osoba koje su prepoznate da su u bazi; crvenim križićem su označene slike osoba koje nisu prepoznate da su u bazi

Poglavlje 2

Metoda Tensorfaces

Ovo poglavlje obrađuje metodu poznatu pod imenom Tensorfaces koja primjenjuje singularnu dekompoziciju višeg reda koju ćemo odsada zvati HOSVD (eng. High Order Singular Value Decomposition). Za razumijevanje navedene metode potrebno je objasniti što su tenzori, definirati njihove osnovne operacije te iskazati glavne rezultate.

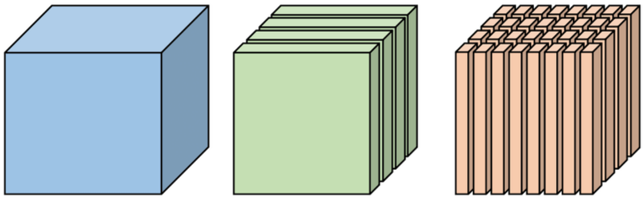
2.1 Tenzori i HOSVD

Tenzore shvaćamo kao nizove podataka [2]. Vektor je jednodimenzionalni niz podataka i smatramo ga tenzorom reda 1. Matrice su dvodimenzionalni nizovi podataka i smatramo ih tenzorima reda 2. Tenzore reda većeg ili jednakog 3 zovemo tenzorima višeg reda. U mnogim slučajevima je prikladno koristiti tenzore viših redova za spremanje podataka umjesto matriciziranih ekvivalenata.

Preciznije, tenzor reda N definiramo kao element tenzorskog produkta N vektorskih prostora. Red tenzora $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ je N. Elemente tenzora $\mathcal{A}$ označavamo s $\mathcal{A}_{i_1 i_2 \ldots i_3}$, gdje su $1 \leq i_n \leq I_n$, $n = 1, 2, \ldots N$. Indekse u tenzoru nazivamo modovima, a dimenzija pripadnog moda je broj različitih vrijednosti koje taj indeks može poprimiti.

U ovom radu ćemo se baviti isključivo tenzorima reda 3, tako da odsada podrazumijevamo tenzore u svim definicijama i rezultatima kao trodimenzionalne nizove podataka.

Ograničavanjem pojedinih indeksa možemo indeksirati podtenzore. Ako imamo tenzor $A = (a_{i_1 i_2 i_3}) \in \mathbb{R}^{I_1 \times I_2 \times I_3}$, tada fiksiranjem indeksa u modu $1, 2$ ili 3 možemo definirati podtenzore $\mathcal{A}(k, :, :)$, $\mathcal{A}(:, k, :)$ ili $\mathcal{A}(:, :, k)$ (koristeći MATLAB-ovu notaciju). Takve podtenzore zovemo horizontalnim, lateralnim i frontalnim odsječcima, respektivno. Frontalne odsječke često označavamo i s $A^{(k)}$. U tom slučaju su podtenzori ustvari tenzori reda 2, odnosno matrice. Također možemo definirati i podtenzore $\mathcal{A}(:, i, j), \mathcal{A}(i, :, j)$ i $\mathcal{A}(i, j, :)$ fiksirajući sve indekse osim jednog. To su ustvari tenzori reda 1 odnosno vektori.



Slika 2.1: S lijeva na desno: prikaz tenzora $\mathcal{A}$, frontalnih odsječaka $\mathcal{A}(:, :, k)$ i podtenzora reda 1 (vektora) $\mathcal{A}(:, i, j)$. [8]

Definirajmo skalarni produkt dvaju tenzora reda 3 jednakih dimenzija.

Definicija 2.1.1. *Skalarni produkt dvaju tenzora istih dimenzija $\mathcal{A}, \mathcal{B} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ je suma umnožaka njihovih elemenata*

$$\langle \mathcal{A}, \mathcal{B} \rangle = \sum_{i=1}^{I_1} \sum_{j=1}^{I_2} \sum_{k=1}^{I_3} a_{ijk} b_{ijk}.$$

Pripadnu Frobeniusovu normu za tenzore reda 3 definiramo analogno definiciji Frobeniusove norme za matrice.

Definicija 2.1.2. *Frobeniusova norma tenzora reda 3 $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ je norma definirana kao korijen zbroja kvadrata svih elemenata tenzora*

$$\|\mathcal{A}\|_F = \sqrt{\langle \mathcal{A}, \mathcal{A} \rangle} = \sqrt{\sum_{i=1}^{I_1} \sum_{j=1}^{I_2} \sum_{k=1}^{I_3} a_{i,j,k}^2}$$

Definiramo množenje u modu 1 matrice s tenzorom.

Definicija 2.1.3. *Množenje u modu 1 tenzora $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ matricom $U \in \mathbb{R}^{l_0 \times I_1}$ (u oznaci $\mathcal{A} \times_1 U$), je $l_0 \times I_2 \times I_3$ tenzor dan s*

$$(\mathcal{A} \times_1 U)(j, i_2, i_3) = \sum_{k=1}^{I_1} u_{j,k} a_{k,i_2,i_3}.$$

U množenju tenzora $\mathcal{A}$ matricom U u modu 1 ustvari množimo stupčane vektore tenzora A s matricom U.

Slično, množenje u modu 2 tenzora A s matricom V je definirano kao

$$(\mathcal{A} \times_2 V)(i_1, j, i_3) = \sum_{k=1}^{I_2} v_{j,k} a_{i_1,k,i_3}.$$

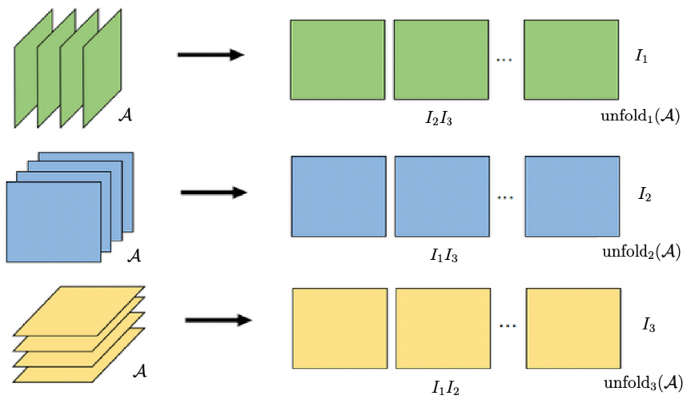
U ovom slučaju imamo da su svi retci tenzora A pomnoženi s matricom V.

Množenje tenzora matricom u redu 3 dobivamo analogno.

Često je korisno matricizirati tenzor u matricu. Matricizaciju tenzora $\mathcal{A}$ definiramo kroz tri moda:

$$\begin{aligned}
\text{unfold}_1(\mathcal{A}) &= (\mathcal{A}(:,1,:) \quad \mathcal{A}(:,2,:) \quad \dots \quad \mathcal{A}(:,I_2,:)) \quad \in \mathbb{R}^{I_1 \times I_2 I_3},\\
\text{unfold}_2(\mathcal{A}) &= (\mathcal{A}(:,:,1)^T \quad \mathcal{A}(:,:,2)^T \quad \dots \quad \mathcal{A}(:,:,I_3)^T) \quad \in \mathbb{R}^{I_2 \times I_1 I_3},\\
\text{unfold}_3(\mathcal{A}) &= (\mathcal{A}(1,:,:)^T \quad \mathcal{A}(2,:,:)^T \quad \dots \quad \mathcal{A}(I_1,:,:)^T) \quad \in \mathbb{R}^{I_3 \times I_1 I_2}.
\end{aligned}$$

Matricizacije tenzora u modovima 1, 2 i 3 ćemo još označavati s $A_{(1)}$, $A_{(2)}$ i $A_{(3)}$, respektivno. Primjetimo da matrice $A_{(n)}$ dobijemo tako da vektore iz moda n tenzora $\mathcal{A}$ složimo kao stupce u matricu $A_{(n)}$.



Slika 2.2: Matricizacije tenzora $\mathcal{A}$. [8]

Takve matrice možemo vratiti u originalan tenzor $\mathcal{A}$ koristeći operator fold koji je definiran kao inverz operatora unfold

$$\text{fold}_i(\text{unfold}_i(\mathcal{A})) = \mathcal{A}.$$

Sada pomoću fold-unfold operacija možemo zapisati množenje tenzora u modu i matricom na slijedeći način

$$\mathcal{A} \times_i U = \text{fold}_i(U \ \text{unfold}_i(\mathcal{A})) = \text{fold}_i(UA_{(i)}).$$

Singularnu dekompoziciju matrice možemo generalizirati na singularnu dekompoziciju tenzora višeg reda.

Prisjetimo se matričnog SVD-a. Izraz (1.2) možemo zapisati koristeći tenzorsku notaciju

$$A = U\Sigma V^T = \Sigma \times_1 U \times_2 V$$

Na isti način možemo dekomponirati tenzore reda jednakog ili većeg od 3, o čemu nam govori sljedeći teorem.

Teorem 2.1.4. *(HOSVD) Tenzor $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \cdots \times I_N}$ možemo zapisati kao*

$$\mathcal{A} = \mathcal{S} \times_1 U^{(1)} \times_2 U^{(2)} \cdots \times_N U^{(N)} \tag{2.1}$$

gdje su $U^{(n)} \in \mathbb{R}^{I_n \times I_n}$ ortogonalne matrice i $\mathcal{S}$ je tenzor istih dimenzija kao i $\mathcal{A}$ te ga zovemo još i jezgrenim tenzorom. Svaka dva odsječka tenzora $\mathcal{S}$ su ortogonalna u smislu skalarnog produkta

$$\langle \mathcal{S}(i,:,\ldots,:), \mathcal{S}(j,:,\ldots,:)\rangle = \langle \mathcal{S}(:,i,:,\ldots,:), \mathcal{S}(:,j,:,\ldots,:)\rangle = \cdots = 0$$

za $i \neq j$. Singularne vrijednosti u modu 1 su definirane s

$$\sigma_j^{(1)} = \|\mathcal{S}(j,:,\ldots,:)\|_F, \quad j = 1,\ldots,I_1$$

i sortirane su silazno

$$\sigma_1^{(1)} \geq \sigma_2^{(1)} \geq \cdots \geq \sigma_{I_1}^{(1)}.$$

Analogno vrijedi i za singularne vrijednosti u ostalim modovima. [4]

Dokaz. Odredimo SVD matrica koje dobijemo matricizacijom tenzora $\mathcal{A}$ u svakom od N modova

$$A_{(i)} = U^{(i)}\Sigma^{(i)}(V^{(i)})^T, \quad i = 1,2,\ldots,N. \tag{2.2}$$

Označimo

$$\mathcal{S} = \mathcal{A} \times_1 (U^{(1)})^T \times_2 (U^{(2)})^T \cdots \times_N (U^{(N)})^T.$$

Ostalo je za pokazati da su odsječci tenzora $\mathcal{S}$ međusobno okomiti i da su singularne vrijednosti u silaznom poretku u svakom modu. U tu svrhu nam treba Kroneckerov produkt matrica. Za matrice $A \in \mathbb{R}^{i\times j}$ i $B \in \mathbb{R}^{k\times l}$ definiramo Kroneckerov produkt kao

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1j}B \\ a_{21}B & a_{22}B & \dots & a_{2j}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1}B & a_{i2}B & \dots & a_{ij}B \end{bmatrix}.$$

Primjetimo da je $A \otimes B \in \mathbb{R}^{(ik)\times(jl)}$.

Sada možemo jednadžbu (2.1) zapisati u matričnom obliku

$$A_{(n)} = U^{(n)} S_{(n)} (U^{(n+1)} \otimes U^{(n+2)} \otimes \cdots \otimes U^{(N)} \otimes U^{(1)} \otimes U^{(2)} \otimes \cdots \otimes U^{(n-1)})^T. \quad (2.3)$$

Označimo s r_n najveći indeks za koji je $\sigma^{(n)}_{r_n} > 0$. Usporedbom (2.2) i (2.3) uz ortogonalnost Kroneckerovog faktora u (2.3) slijedi

$$S_{(n)} = \Sigma^{(n)} (V^{(n)})^T (U^{(n+1)} \otimes U^{(n+2)} \otimes \cdots \otimes U^{(N)} \otimes U^{(1)} \otimes U^{(2)} \otimes \cdots \otimes U^{(n-1)}). \quad (2.4)$$

Iz (2.4) slijedi da za proizvoljne ortogonalne matrice $U^{(1)}, U^{(2)}, \dots, U^{(n-1)}, U^{(n+1)}, \dots, U^{(N)}$ vrijedi da je skalarni produkt bilo koja dva različita odsječka (u istom modu) tenzora $\mathcal{S}$ jednak 0 te vrijedi

$$\sigma^{(n)}_1 \geq \sigma^{(n)}_2 \geq \cdots \geq \sigma^{(n)}_{I_n} \geq 0,$$

te za $r_n < I_n$ vrijedi

$$\sigma^{(n)}_{r_n+1} = \cdots = \sigma^{(n)}_{I_n} = 0.$$

□

Jezgreni tenzor $\mathcal{S}$ u HOSVD dekompoziciji ima masu koncentriranu u prednjem gornjem lijevom kutu ako tenzor zamislimo kao trodimenzionalni kvadar. Što se više odmičemo u bilo kojem smjeru njegovi elementi postaju sve manji.

Često se događa (primjerice, kod slika) da je dimenzija jednog moda veća od umnoška dimenzija drugih modova. Ako imamo tenzor $\mathcal{A} \in \mathbb{R}^{l\times m\times n}$ reda 3 za koji vrijedi $l > mn$, može se pokazati da za jezgreni tenzor $\mathcal{S}$ vrijedi

$$\mathcal{S}(i, :, :) = 0, \quad i > mn.$$

U tom slučaju možemo zanemariti dio s nulama jezgrenog tenzera te napisati tanku verziju HOSVD-a

$$\tilde{\mathcal{A}} = \tilde{\mathcal{S}} \times_1 \tilde{U}^{(1)} \times_2 U^{(2)} \times_3 U^{(3)} \quad (2.5)$$

gdje su $\tilde{\mathcal{S}} \in \mathbb{R}^{mn\times m\times n}$ i $\tilde{U}^{(1)} \in \mathbb{R}^{l\times mn}$.

2.2 Metoda Tensorfaces

Tensorfaces je bio prvi algoritam za prepoznavanje lica koji se bazirao na tenzorskim operacijama. Za razliku od PCA metode gdje smo kolekciju slika prikazali u obliku matrice, u Tensorfaces metodi reprezentiramo podatke u obliku tenzora. Sličnost ovih dviju metoda je što i dalje radimo sa vektoriziranim slikama.

Pretpostavimo da imamo kolekciju slika od n_3 osoba, za svaku osobu po n_2 slika (za sve različite izraze lica). Neka su slike rezolucije $m_1 \times m_2$. S obzirom da radimo s greyscale slikama, slike možemo shvatiti kao matrice s $m_1 m_2 = n_1$ elemenata. Tada vektorizirane slike predstavljaju vektore u $\mathbb{R}^{n_1}$.

Takve vektorizirane slike spremamo u stupce frontalnih odsječaka tenzora $\mathcal{A}$, i to na taj način da j-ti frontalni odsječak reprezentira kolekciju slika za j-tu osobu. Tada cijelu kolekciju imamo spremljenu u tenzor

$$\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}.$$

Pretpostavljamo da je $n_1 > n_2 n_3$. Za tenzor $\mathcal{A}$ možemo pisati tanki HOSVD (2.5)

$$\mathcal{A} = \mathcal{S} \times_1 F \times_2 G \times_3 H,$$

gdje je $\mathcal{S} \in \mathbb{R}^{n_2 n_3 \times n_2 \times n_3}$ reducirani jezgreni tenzor, $F \in \mathbb{R}^{n_1 \times n_2 n_3}$ matrica s ortogonalnim stupcima, te $G \in \mathbb{R}^{n_2 \times n_2}$ i $H \in \mathbb{R}^{n_3 \times n_3}$ su ortogonalne matrice.

Označimo

$$\mathcal{C} = \mathcal{S} \times_1 F \times_2 G.$$

Sada tanki HOSVD tenzora $\mathcal{A}$ možemo zapisati kao

$$\mathcal{A} = \mathcal{C} \times_3 H.$$

Lateralni odsječci $\mathcal{A}(:, i, :)$ predstavljaju kolekciju fotografija po izrazima lica. Za određeni izraz $\mathcal{A}(:, e, :)$ imamo

$$\mathcal{A}(:, e, :) = \mathcal{C}(:, e, :) \times_3 H.$$

Označimo podtenzore $\mathcal{A}(:, e, :)$ i $\mathcal{C}(:, e, :)$ kao matrice A_e i C_e. Sada imamo

$$A_e = C_e H^T, \quad e = 1, 2, \ldots, n_2.$$

Ista ortogonalna matrica H će se pojaviti u svih n_2 relacija. Ako zapišemo H^T kao $(h_1, h_2, \ldots, h_{n_3})$, stupac p matrice A_e možemo zapisati kao

$$a_p^{(e)} = C_e h_p.$$

gdje je $a_p^{(e)}$ slika osobe p u izrazu e, C_e matrica čiji stupci čine bazne vektore za izraz e, te h_p red p matrice H koji sadrži koordinate fotografija osobe p u toj bazi.

Neka je $X \in \mathbb{R}^{n_1}$ vektorizirana testna fotografija. Klasificiranje osobe na testnoj fotografiji se radi tako da koordinate vektora X izračunamo u svim bazama za izraze i provjeravajući za svaki izraz podudaraju li se koordinate X s elementima bilo kojeg reda matrice H. Koordinate testne fotografije X u izraznoj bazi e nalazimo rješavajući problem najmanjih kvadrata

$$\min_{\alpha_e} \|C_e \alpha_e - X\|_2.$$

Možemo pretpostaviti da je C_e matrica punog stupčastog ranga. U tom slučaju rješenje možemo izračunati kao

$$\alpha_e = (C_e^T C_e)^{-1} C_e^T X.$$

Podudarnost testne fotografije X sa slikom $a_p^{(e)}$ izračunavamo koristeći Frobeniusovu normu razlike između α_e (fotografije X u izraznoj bazi e) i svih vektora h_p, $p = 1, 2, \ldots, n_3$.

Algoritam 2 - Metoda Tensorfaces

Ulaz: Slike za treniranje: $X_{i,e,p}$, $i = 1, 2, ..., n_1$, $e = 1, 2, ..., n_2$, $p = 1, 2, ..., n_3$, Testna slika: X, Tolerancija *tol*
Izlaz: 1 (ako je osoba u bazi), 0 (ako osoba nije u bazi)

for $e = 1, 2, ..., n_2$, $p = 1, 2, ..., n_3$ **do**
  $\mathcal{A}(:, e, p) \leftarrow$ vektorizirani $X_{i,e,p}$
end for
HOSVD tenzora $\mathcal{A}$ kako bi dobili jezgreni tenzor $\mathcal{S}$ i ortogonalne matrice F, G, H
$\mathcal{C} \leftarrow \mathcal{S} \times_1 F \times_2 G$
$X \leftarrow$ vektorizirani X
for $e = 1, 2, \ldots, n_2$ **do**
  $C_e \leftarrow \mathcal{C}(:, e, :)$
  $\alpha_e \leftarrow (C_e^T C_e)^{-1} C_e^T X$
  for $p = 1, 2, \ldots, n_3$ **do**
    if $||\alpha_e - H(p, :)||_F < tol$ **then**
      return 1
    end if
  end for
end for
return 0

2.3 Klasifikacija i usporedba s PCA algoritmom

Ponovno koristimo slike iz baze podataka *Yale Face Database* za testiranje algoritma. Imamo 10 osoba i 7 izraza lica za svaku osobu. Slike su obrađene na isti način: smanjene su na rezoluciju 90×90, prevedene u *greyscale* kanal i vrijednosti piksela su normalizirane.

Za bazu uzimamo prve tri ekspresije lica za prvih pet osoba (slike u zelenom pravokutniku na slici 2.3), a ostatak slika služe za testiranje. Slike vektoriziramo i zatim ih slažemo u tenzor dimenzija $8100 \times 3 \times 5$. Slike će biti posložene u stupce, i to tako da prvi frontalni odsječak baznog tenzora dimenzije 8100×3 čine slike prve osobe, drugi frontalni odsječak čine slike druge osobe itd...

Algoritam vraća vrijednost 1 ako je osoba prepoznata da je u bazi, u protivnom vraća 0. Do iznosa tolerancije ponovno dolazimo eksperimentima. U ovom testiranju najefektivnija tolerancija je bila 0.41 koja je prepoznala sve slike na kojima su osobe u bazi te je odbacila sve slike onih osoba koje nisu u bazi (slika 2.3).

Konkretno na ovoj bazi slika testirali smo algoritme PCA i Tensorfaces. Naštimavanjem tolerancije vidimo da smo dobili bolju klasifikaciju s Tensorfaces algoritmom, jer je PCA odbacio jednu sliku osobe koja je u bazi. Povećanjem tolerancije u PCA možemo dobiti da algoritam prepozna i tu sliku, ali je tada algoritam počeo prepoznavati i druge testne slike da su u bazi iako ustvari nebi trebale biti prepoznate. Razlog zašto ne uzimamo tu toleranciju je taj da je greška neprepoznavanja osobe u bazi manja nego greška prepoznavanja osobe koja pak nije u bazi.

Još veću prednost algoritma Tensorfaces možemo vidjeti ako imamo više parametara koji variraju. Naime slike nad kojima su testirani algoritmi imaju iste uvjete poput kuta gledanja, osvijetljenosti i sl. Kada bi imali slike gdje bi naprimjer parametar kuta gledanja varirao, pokazano je da PCA pokazuje puno lošije rezultate. U Tensorfaces algoritmu tada bi "pakirali" slike u tenzor reda 4, gdje bi 4. mod označavao mod kuta gledanja. Tada bi imali dekompoziciju tenzora reda 4, odnosno imali bi jezgreni tenzor reda 4 i 4 kvadratne matrice. Prednost multilinearnog tenzorskog pristupa u odnosu na linearnu PCA metodu se očitava u činjenici da Tensorfaces preslikava sve slike jedne osobe u isti vektor koeficijanata h_p bez obzira na izraz lica, kut gledanja, osvjetljenost itd... U PCA pristupu svaka slika osobe odgovarala je jednom vektoru koeficijanata. Iz tog razloga Tensorfaces algoritam stvara dobro opredijeljene klase ljudi jer maksimiziramo omjer raspršenosti među klasama i raspršenosti unutar jedne klase.

Slika 2.3: Rezultati Tensorfaces algoritma - zelenom kvačicom su označene slike osoba koje su prepoznate da su u bazi; crvenim križićem su označene slike osoba koje nisu prepoznate da su u bazi

Bibliografija

[1] D. Bakić, *Linearna algebra*, Školska knjiga, 2008.

[2] N. Bosner I. Zirdum, *Prepoznavanje lica pomoću tenzorske dekompozicije singularnih vrijednosti*, Hrvatski matematički elektronički časopis (2019).

[3] R. C. Hoover K. Braman, *Facial Recognition Using Tensor-Tensor Decompositions*, SIAM, Journal on Imaging Sciences (2013).

[4] J. Vandewalle L. De Lathauwer, B. De Moor, *Multilinear Singular Value Tensor Decompositions*, SIAM, Journal on Matrix Analysis and Applications (2000).

[5] D. Kriegman P. Belhumer, J. Hespanha, *Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection*, IEEE Transactions on Pattern Analysis and Machine Intelligence (1997), 711–720.

[6] N. Sarapa, *Teorija vjerojatnosti*, Školska knjiga, 2002.

[7] S. Singer, *Numerička analiza, 25. predavanje*, (2009), `https://web.math.pmf.unizg.hr/~singer/num_anal/NA_0910/25.pdf`.

[8] Z. Zhang, *Tucker Tensor Decomposition on FPGA*, ResearchGate (2019).

Sažetak

Tema ovog diplomskog rada je prepoznavanje lica algoritmima PCA i Tensorfaces. Rad prikazuje primjenu bitnih teorema o singularnim dekompozicijama matrice i tenzora u algoritmima za prepoznavanje lica.

U početku su prikazani osnovni pojmovi i rezultati iz linearne algebre i vjerojatnosti. Nakon toga slijedi iskaz singularne dekompozicije matrice i opis poznate metode za redukciju prostora, PCA, u kontekstu prepoznavanja lica, kao i primjena samog algoritma na određenom skupu slika. U sljedećem poglavlju uvedeni su pojam tenzora, osnovne tenzorske operacije i singularna dekompozicija tenzora. Nakon toga slijedi opis metode Tensorfaces i primjena algoritma na istom skupu slika.

Kodovi samih algoritama napisani su u programskom jeziku MATLAB.

Summary

The topic of this thesis is the Face Recognition using PCA and Tensorfaces algorithms. This thesis shows the application of important theorems of singular matrix and tensor decompositions in face recognition algorithms.

The introduction contains basic terms and results in the field of linear algebra and probability. It is followed by the statement of the matrix singular value decomposition theorem, the description of the PCA algorithm in the context of face recognition and its application on the image dataset. Next chapter introduces tensors, basic tensor operations and the singular decomposition of a tensor. It is followed by the description of the Tensorfaces method and the application of the algorithm on the same image dataset.

All of the algorithm codes are written in MATLAB programming language.

Životopis

Rođen sam 02. srpnja 1994. godine u Zagrebu. Svoje školovanje započinjem u Osnovnoj školi Malešnica koju završavam 2009. godine i nakon koje upisujem Gimnaziju Lucijana Vranjanina u Zagrebu, smjer prirodoslovno-matematički. Godine 2015. upisujem preddiplomski studij matematike, smjer nastavnički, na Prirodoslovno-matematičkom fakultetu u Zagrebu. Po završetku preddiplomskog studija upisujem diplomski sveučilišni studij Matematička statistika na istom sveučilištu. Trenutno radim kao konzultant u odjelu za podatkovnu analitiku u firmi NEOS.