

# Višeklasno spektralno klasteriranje

Marija Majda Perišić, Tomislav Droždjek, Josip Begić

19. siječnja 2013.

# 1 Uvod

Višeklasno spektralno klasteriranje uspješna je metoda za rješavanje problema poput segmentacije slike, balansiranja opterećenja ili nacрта sklopa. Metoda ne pretpostavlja ništa o globalnoj strukturi podataka, nego na lokalnoj razini prikuplja informacije o vjerojatnosti da dva podatka pripadaju nekoj klasi, a tek potom se donosi globalna odluka o dijeljenju podataka na disjunktne skupove ovisno o nekom kriteriju.

Spektralne metode su dobre i česte u upotrebi jer se globalni optimum u neprekidnoj domeni ostvaruje preko spektralne dekompozicije (eng. eigendecomposition). Međutim, dobiti diskretno rješenje od pripadnih svojstvenih vektora je novi problem klasteriranja, često u prostoru manje dimenzije. Kako bi dobili optimalno rješenje, koristimo činjenicu da se optimalno neprekidno rješenje sastoji od cijele familije razapete svojstvenim vektorima putem ortonormalnih transformacija. Naš cilj je naći pravu ortonormalnu transformaciju koja vodi do diskretizacije.

U našem konkretnom slučaju, bavimo se klasteriranjem skupa od  $N$  točaka u  $K$  klastera. Prvo rješavamo neprekidni problem optimizacije (eng. relaxed continuous optimization problem) spektralnom dekompozicijom. U radu ćemo objasniti ulogu svojstvenih vektora kao generatora svih optimalnih rješenja pomoću ortonormiranih transformacija (eng. orthonormal transformations). Nakon toga, iterativno tražimo diskretno rješenje najbliže neprekidnom pomoću izmjeničnog postupak optimizacije (eng. alternating optimization procedure): ponavljamo traženje neprekidnog optimuma najbližeg diskretnom rješenju tako da računamo najbolju ortonormiranu transformaciju te zatim, metodom nemaksimalne supresije, (eng. non-maximum suppression) tražimo najbliže diskretno rješenje onom neprekidnom. Ovakve iteracije, kao što ćemo pokazati, monotonno smanjuju razliku udaljenosti diskretnog i neprekidnog rješenja. Drugim riječima, diskretno rješenje konvergira ka neprekidnom, i to rješenje je, ispostavit će se, blizu globalnom optimumu.

## 2 Višeklasni normalizirani rezovi

Težinski graf  $G = (\mathbb{V}, \mathbb{E}, W)$  određen je skupom čvorova  $\mathbb{V}$ , bridova  $\mathbb{E}$  i težinama  $W$ . Težine spremamo u matricu  $W$ , čiji elementi  $w_{ij}$  određuju koliko su čvorovi  $i$  i  $j$  udaljeni jedan od drugog. Jasno, pretpostavlja se da je  $W$  simetrična i nenegativna. Ako sa  $[N] = \{1, 2, \dots, N\}$  označimo skup svih elemenata koje treba grupirati, onda problem klasteriranja  $N$  točaka u  $K$  klastera postaje problem dekomponiranja  $\mathbb{V}$  čvorova u  $K$  disjunktne skupa. Označavamo  $K$ -particiju sa  $\Gamma_{\mathbb{V}}^K = \{\mathbb{V}_1, \dots, \mathbb{V}_k\}$

### 2.1 Višeklasni kriterij particioniranja

Neka su  $\mathbb{A}, \mathbb{B} \subseteq \mathbb{V}$ . Definiramo

$$links(\mathbb{A}, \mathbb{B}) = \sum_{\substack{i \in \mathbb{A} \\ j \in \mathbb{B}}} W_{ij}$$

$$degree(\mathbb{A}) = links(\mathbb{A}, \mathbb{V})$$

$$\text{linkratio}(\mathbb{A}, \mathbb{B}) = \frac{\text{links}(A, B)}{\text{degree}(A)}$$

$\text{Links}(\mathbb{A}, \mathbb{B})$  je suma svih težina svih veza između čvorova iz  $\mathbb{A}$  i čvorova iz  $\mathbb{B}$ ,  $\text{degree}(\mathbb{A})$  su sve težine koje čvorovi iz  $\mathbb{A}$  ostvaruju sa svim ostalim čvorovima, a  $\text{linkratio}$  dobijemo normalizacijom  $\text{links}(\mathbb{A}, \mathbb{B})$ . Od posebne su važnosti  $\text{linkratio}(\mathbb{A}, \mathbb{A})$  (koliko veza ostaje u  $\mathbb{A}$ ) i  $\text{linkratio}(\mathbb{A}, \mathbb{V} \setminus \mathbb{A})$  (koliko veza ide van iz  $\mathbb{A}$ ). Očekujemo da dobro klasteriranje ima dobre veze unutar particije i slabe s ostalim particijama. Uvodimo oznake:

$$\text{knassoc}(\Gamma_{\mathbb{V}}^K) = \frac{1}{K} \sum_{l=1}^K \text{linkratio}(\mathbb{V}_l, \mathbb{V}_l)$$

$$\text{knacuts}(\Gamma_{\mathbb{V}}^K) = \frac{1}{K} \sum_{l=1}^K \text{linkratio}(\mathbb{V}_l, \mathbb{V} \setminus \mathbb{V}_l)$$

Primjetimo da je  $\text{knassoc} + \text{knacuts} = 1$  pa maksimiziranjem veza minimiziramo rezove. Stoga je naš problem ekvivalentan maksimiziranju:

$$\epsilon(\Gamma_{\mathbb{V}}^K) = \text{knassoc}(\Gamma_{\mathbb{V}}^K) \quad \epsilon \in (0, 1)$$

Pokazati ćemo da se gornja granica od  $\epsilon$  smanjuje monotono kako povećavamo  $K$ .

## 2.2 Reprezentacija

Koristimo  $N \times K$  particijsku matricu  $X$  kako bi reprezentirali  $\Gamma_{\mathbb{V}}^K$ . Neka je  $X = [X_1, \dots, X_K]$ , gdje je  $X_l$  binarni indikator za  $\mathbb{V}_l$ :

$$X(i, l) = \langle i \in \mathbb{V}_l \rangle, \quad i \in \mathbb{V}, l \in [K]$$

pri čemu je  $\langle \cdot \rangle = 1$  ukoliko je argument istinit, a 0 inače. Kako je svaki čvor pridružen jednoj i samo jednoj particiji, postoji ograničenje na retke od  $X : X1_K = 1_N$ , gdje je  $1_d$   $d \times 1$  jedinični vektor. Definiramo matricu stupnjeva (analogija sa definicijom od degree) za simetričnu matricu  $W$ :

$$D = \text{Diag}(W1_N)$$

pri čemu je  $\text{Diag}$  diagonalna matrica formirana vektorom kojeg prima kao argument. Sada možemo definirati  $\text{links}$  i  $\text{degree}$  kao

$$\text{links}(\mathbb{V}_l, \mathbb{V}_l) = X_l^T W X_l$$

$$\text{degree}(\mathbb{V}_l) = X_l^T D X_l$$

Trebamo, dakle riješiti program optimizacije varijable  $X$ , kojeg zovemo  $PNCX$ :

$$\text{maksimizirati } \epsilon(X) = \frac{1}{K} \sum_{l=1}^K \frac{X_l^T W X_l}{X_l^T D X_l}$$

$$\text{uz uvjet } X \in \{0, 1\}^{N \times K}$$

$$X1_K = 1_N$$

Ovaj problem je NP-težak već za  $K = 2$ , pa ćemo pokušati razviti algoritam za dobivanje skoro pa globalnog optimuma.

### 3 Rješavanje K-normaliziranih rezova

Program  $PNCX$  rješavamo u dva koraka. Prvo proširujemo transformiranu formulaciju na problem svojstvenih vektora. Pokazujemo da globalni optimum nije jedinstven, i da je jedno rješenje matrica svojstvenih vektora uređenog para matrica  $(W, D)$ . Transformiranjem te svojstvene matrice pomoću ortogonalnih transformacija, dobivamo skup neprekidnih globalnih rješenja. Zatim rješavamo problem diskretizacije, gdje tražimo diskretnu matricu particije najbližu neprekidnoj. To diskretno rješenje je skoro pa globalno optimalno.

#### 3.1 Traženje optimalnih rješenja na proširenoj domeni

Pojednostavljujemo jednadžbu za  $PNCX$  kao  $\epsilon(X) = \frac{1}{K} \text{tr}(Z^T W Z)$ , pri čemu je  $\text{tr}$  trag matrice, a  $Z$  skalirana matrica particije:

$$Z = X(X^T D X)^{-\frac{1}{2}}$$

Kako je  $X^T D X$  diagonalna matrica, stupci od  $Z$  su oni od  $X$  podijeljeni drugim korjenom matrice stupnja odgovarajućeg klastera, pri čemu smo stupanj klastera  $V_K$  definirali kao sumu svih  $\text{degree}(i)$  za svaki čvor  $i \in V_K$ . Prirodno je ograničenje na  $Z$ :

$$Z^T D Z = (X^T D X)^{-\frac{1}{2}} X^T D X (X^T D X)^{-\frac{1}{2}} = I_K,$$

gdje je  $I_K$  jedinična  $K$  matrica. Ako ignoriramo ograničenja u  $PNCX$ , izveli smo novi program po varijabli  $Z$ , zovemo ga  $PNCZ$ :

$$\text{maksimizirati } \epsilon(Z) = \frac{1}{K} \text{tr}(Z^T W Z)$$

$$\text{uz uvjet } (Z^T W Z) = I_K$$

Proširenjem  $Z$  na neprekidnu domenu, od diskretnog problema došli smo do neprekidnog problema optimizacije. Taj problem je invarijantan na djelovanje ortogonalnih operatora o čemu govori sljedeća propozicija:

**Propozicija 1** (Invarijantnost na rotacije i refleksije). *Neka je  $R$   $K \times K$  matrica. Ako  $Z$  možemo dobiti kao rješenje programa  $PNCZ$ , tada je rješenje programa  $i \in \{ZR : R^T R = I_K\}$ . Nadalje, svi imaju istu ciljanu vrijednost (eng. objective value):  $\epsilon(ZR) = \epsilon(Z)$*

Program  $PNCZ$  je problem optimizacije Rayleightovog kvocijenta, čije je rješenje dano u Rayleight-Ritz teoremu. Propozicija 2 je ekvivalentna tom teoremu, ali je iskazana tako da je možemo primijentiti na naš problem. Ona pokazuje da su svojstveni vektori od  $(W, D)$  optimalno rješenje, ili, ekvivalentno, oni normalizirane težinske matrice  $P$ :

$$P = D^{-1}W.$$

Kako je  $P$  stohastička matrica, znamo da je  $1_N$  trivijalni svojstveni vektor od  $P$  koji korespondira najvećoj svojstvenoj vrijednosti 1.

**Propozicija 2** (Optimalno svojstveno rješenje). *Neka je  $(V, S)$  dekompozicija od  $P : PV = VS$ , pri čemu je  $V = [V_1, \dots, V_N]$  matrica koja sadrži svojstvene vektore, a  $S = \text{Diag}(s)$ , matrica koja na dijagonali sadrži svojstvene vrijednosti poredane u nerastućem poretku:  $s_1 \geq \dots \geq s_N$ .  $(V, S)$  je dobiven iz  $(\bar{V}, S)$  iz simetrične matrice  $D^{-\frac{1}{2}}WD^{-\frac{1}{2}}$ , gdje su  $\bar{V}_i$  i  $\bar{V}_j$  međusobno ortogonalni  $\forall i \neq j$ , te vrijedi:*

$$V = D^{-\frac{1}{2}}\bar{V}$$

$$D^{-\frac{1}{2}}WD^{-\frac{1}{2}}\bar{V} = \bar{V}S, \bar{V}^T\bar{V} = I_N$$

Stoga,  $V$  i  $S$  su svi realni i bilo kojih  $K$  svojstvenih vektora formira kandidata za lokalni optimum u programu PNCZ, uz

$$\epsilon([V_{\pi_1}, \dots, V_{\pi_K}]) = \frac{1}{K} \sum_{l=1}^K s_{\pi_l},$$

gdje je  $\pi$  vektor koji sadrži  $k$  prirodnih brojeva od 1 do  $N$ . Globalnim optimum programa PNCZ se postiže za  $\pi = [1, \dots, K]$ :

$$Z^* = [V_1, \dots, V_K],$$

$$\Lambda^* = \text{Diag}([s_1, \dots, s_K]),$$

$$\epsilon(Z^*) = \frac{1}{K} \text{tr}(\Lambda^*) = \max_{Z^T DZ = I_K} \epsilon(Z)$$

Dakle, globalni optimum programa PNCZ nije jedinstven, već je potprostor razapet sa najvećih  $K$  svojstvenih vektora matrice  $P$

**Korolar 1** (Monotonost gornje međe). *Za svaki  $K$  je:*

$$\max \epsilon(\Gamma_{\mathbb{V}}^K) \leq \max_{Z^T DZ = I_K} \epsilon(Z) = \frac{1}{K} \sum_{l=1}^K s_l$$

$$\max_{Z^T DZ = I_{K+1}} \epsilon(Z) \leq \max_{Z^T DZ = I_K} \epsilon(Z)$$

Sada transformiramo  $Z$  natrag na prostor particijskih matrica. Ako je  $f$  preslikavanje koje spridružuje matricu  $X$  matrici  $Z$ , onda je  $f^{-1}$  normalizacija koja vraća  $Z$  natrag u  $X$ :

$$Z = f(X) = X(X^T DX)^{-\frac{1}{2}}$$

$$X = f^{-1}(Z) = \text{Diag}(\text{diag}^{-\frac{1}{2}}(ZZ^T))Z,$$

Ako retke od  $Z$  shvatimo kao koordinate  $K$ -dimenzionalnih točaka, onda  $f^{-1}$  normalizira njihove dužine tako da leže na jediničnoj hipersferi centriranoj u ishodištu. Sa  $f^{-1}$  transformiramo neprekidni optimum  $Z^*R$  sa  $Z$ -prostora u  $X$ -prostor: kako je  $R^T R = I_K$ ,

$$f^{-1}(Z^*R) = f^{-1}(Z^*)R$$

Pojednostavljenje nam dopušta da neprekidni optimum karakteriziramo pomoću  $f^{-1}(Z^*)$  u  $X$ -prostoru:

$$\{\tilde{X}^*R : \tilde{X}^* = f^{-1}(Z^*), R^T R = I_K\}$$

Sada se vidi da trebamo  $K$  i samo  $K$  svojstvenih vektora kako bi dobili  $K$  particija. Razlog je tomu to što su indikatori grupe ortogonalni - ne mogu biti odabrani po volji-

### 3.2 Traženje optimalnog diskretnog rješenja

U općenitom slučaju, optimalna rješenja programa  $PNCZ$  nisu ostvariva za  $PNCX$ . Ipak, možemo ih iskoristiti kako bi našli približno (skoro pa globalno optimalno) diskretno rješenje. U ovom odjeljku pokazat ćemo kako naći diskretno rješenje koje zadovoljava binarna ograničenja originalnog programa  $PNCX$  i najbliže je neprekidnom optimumu

**Teorem 1** (Optimalna diskretizacija). *Neka je  $\tilde{X}^* = f^{-1}(Z^*)$ . Optimalna diskretna particija  $X^*$  je ona koja zadovoljava sljedeći program, kojeg zovemo POD:*

$$\begin{aligned} \text{minimizirati } \phi(X, R) &= \|X - \tilde{X}^* R\|^2 \\ \text{uz uvjet } X &\in \{0, 1\}^{N \times K}, X 1_K = 1_N \\ R^T R &= I_K, \end{aligned}$$

gdje  $\|M\|$  označava Frobeniusovu normu matrice  $M$  :  $\|M\| = \sqrt{\text{tr}(MM^T)}$ . Lokalni optimum  $(X^*, R^*)$  ovog bilinearnog programa može se dobiti iterativno. Uz dani  $R^*$ , POD se svodi na PODX u  $X$ :

$$\begin{aligned} \text{minimizirati } \phi(X) &= \|X - \tilde{X}^* R\|^2 \\ \text{uz uvjet } X &\in \{0, 1\}^{N \times K}, X 1_K = 1_N \end{aligned}$$

Neka je  $\tilde{X} = \tilde{X}^* R^*$ . Optimalno rješenje dano je metodom nemaksimalne supresije (ako je više maksimuma uzimamo samo jedan, i to bilo koji, kako bi vrijedilo ograničenje na matricu particije):

$$X^*(i, l) = \arg \max_{k \in [K]} \tilde{X}(i, k), \quad i \in \mathbb{V}.$$

Uz dani  $X^*$ , POD se svodi na PODR u  $R$ :

$$\begin{aligned} \text{minimizirati } \phi(R) &= \|X^* - \tilde{X}^* R\|^2 \\ \text{uz uvjet } R^T R &= I_K, \end{aligned}$$

i rješenje je dano vektorima dobivenim iz SVD dekompozicije:

$$R^* = \tilde{U} U^T,$$

$$X^{*T} \tilde{X}^* = U \Omega \tilde{U}^T, \quad \Omega = \text{Diag}(\omega),$$

gdje je  $(U, \Omega, \tilde{U})$  SVD dekompozicija matrice  $X^{*T} \tilde{X}^*$ , uz  $U^T U = I_K, \tilde{U}^T \tilde{U} = I_K$  i  $\omega_1 \geq \dots \geq \omega_K$

*Dokaz.* Prvo primjetimo da je  $\phi(X, R) = \|X\|^2 + \|\tilde{X}^*\|^2 - \text{tr}(X R^T \tilde{X}^{*T} + X^T \tilde{X}^* R) = 2N - 2\text{tr}(X R^T \tilde{X}^{*T})$ . Dakle, minimiziranje  $\phi(X, R)$  je ekvivalentno maksimiziranju  $\text{tr}(X R^T \tilde{X}^{*T})$ . Za PODX, uz dano  $R = R^*$ , budući da svaki element od  $\text{diag}(X R^* \tilde{X}^{*T})$  može biti optimiziran nezavisno, slijedi tvrdnja (36). Za PODR, uz dano  $X = X^*$  konstruiramo Lagrangian koristeći simetričnu matricu  $\Lambda$ :

$$L(R, \Lambda) = \text{tr}(X^* R^T \tilde{X}^{*T}) - \frac{1}{2} \text{tr}(\Lambda^T (R^T R - I_K))$$

Optimum  $(R^*, \Lambda^*)$  mora zadovoljavati:

$$L_R = \tilde{X}^{*T} X^* - R\Lambda = 0, \quad \text{tj.} \quad \Lambda^* = R^{*T} \tilde{X}^{*T} X^*.$$

Stoga,  $\Lambda^{*T} \Lambda^* = U\Omega^2 U^T$ , a kako je  $\Lambda = \Lambda^T$ , onda slijedi i  $\Lambda^* = U\Omega U^T$ . Iz (41) imamo:  $R^* = \tilde{U}U^T$  i  $\phi(R^*) = 2N - 2\text{tr}(\Omega)$ . Što je veći  $\text{tr}(\Omega)$ ,  $X^*$  je bliže  $\tilde{X}^* R^*$ .  $\square$

Zbog invarijantnosti neprekidnog optimuma na množenje ortonormiranim matricama, naša metoda je robusna s obzirom na proizvoljnu inicijalizaciju, bilo iz  $X$  ili  $R$ . No, treba naglasiti da će dobra inicijalizacija ubrzati konvergenciju. Ispostavilo se da je metoda grupiranja pomoću  $K$ -sredina, gdje je centar svake grupe jedna od  $K$  skoro ortogonalnih točaka relativno dobra i brza. Računski, to je ekvivalentno inicijalizaciji  $R^*$  odabirom  $K$  redova iz  $\tilde{X}^*$ , koji su međusobno najviše moguće ortogonalni. Štoviše, dobivanje  $X^*$  kao u formuli  $X^*(i, l) = \langle l = \arg \max_{k \in [K]} \tilde{X}(i, k) \rangle$ ,  $i \in \mathbb{V}$ . na tom neortogonalnom  $R^*$  je upravo grupiranje pomoću  $K$ -sredina, gdje su centri svih grupa norme 1. Za dani  $X^*$ , rješavamo PODR kako bi pronašli njemu najbliži neprekidni optimum  $\tilde{X}^* R^*$ . Nakon toga, za taj neprekidni optimum, rješavamo  $PODX$  kako bi pronašli najbliže diskretno rješenje. Svakim korakom smanjujemo isti  $\phi$ , pri čemu  $\phi$  mjeri razliku svojstvenih vrijednosti matrica dobivenih u  $k$ -toj i  $k+1$ -voj iteraciji. Te iteracije teže prema lokalnom optimumu, koji nije jednak za svaku početnu iteraciju. No, svi su  $\tilde{X}^* R^*$  globalni optimumi, bez obzira na  $R^*$ , pa kojem god  $\tilde{X}^* R^*$  program POD konvergira, njegovo najbliže diskretno rješenje  $X^*$  neće biti predaleko od optimalnog.

## 4 Algoritam i primjeri

### 4.1 Algoritam

U ovom poglavlju prikazujemo algoritam koji smo koristili za rješavanje našeg problema, te prilažemo nekoliko primjera sa rezultatima

```
%n——>broj tocaka
%k——>broj particija

%računamo težinsku matricu
W = zeros( n, n );
sigma=1;

for i = 1 : n
    for j = 1 : n
        norma = norm( mat(i,:) - mat(j,:) );
        W( i, j ) = exp( - norma*norma/sigma );
    end
end

%algoritam

%1. korak
D=diag(W*ones(n,1));

%2. korak
%trazimo prvih K eigenvektora od  $(D^{-1/2})W(D^{-1/2})$ , nakon toga racunamo Z
[V,S]=eig(sqrt(D)\W/(sqrt(D)));

[d,order] = sort(diag(S),'descend');
V = V(:,order);

Z=sqrt(D)\V(:,1:k);

%3. korak
XZT=diag((diag(Z*Z')).^(-1/2))*Z;

%4. korak
i=randint(1,1,n);
R(:,1)=XZT(i,:);
c=zeros(n,1);

for j=2:k
    c=c+abs(XZT*R(:,j-1));
    [Y,I] = min(c);
    R(:,j)=XZT(I,:);
end

%5. korak
FIz=0;
p=0; %p je brojač iteracija
```



```

%ulazimo u petlju kojom tražimo najbliže diskretno
%rješenje optimalnom rješenju neprekidnog problema

while(1)
    p=p+1;

%6.korak – tražimo optimalno diskretno rješenje X*
    XT=XZT*R;
    XZ=zeros(n,k);

        for i=1:n
            for l=1:k
                if (XT(i,l)==max(XT(i,:)))
                    XZ(i,l)=1;
                    break
                end
            end
        end

%7.korak – tražimo optimalnu ortogonalnu transformaciju (rotaciju) R*
    [U,O,UT]=svd(XZ'*XZT);
    FI=trace(O);

    if abs(FI==FIz)
        break
    end

    FIz=FI;
    R=UT*U';

end

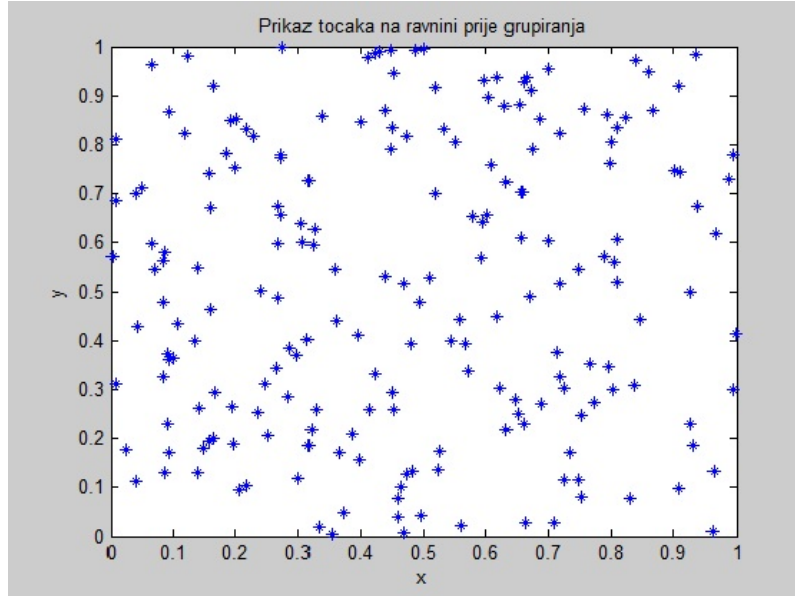
```

U koraku 2, koristimo  $\tilde{V}_{[K]}$  kao oznaku za  $[\tilde{V}_1, \dots, \tilde{V}_K]$ , te analogno za  $\tilde{S}_{[K]}$ . U koraku 4,  $B = \text{abs}(A)$  označava apsolutnu vrijednost elemenata od  $A$ . U koraku 3, kako  $\tilde{X}^* = \text{Diag}(\text{diag}^{-\frac{1}{2}}(Z^*Z^{*T}))Z^*$  skalira dužine svakog retka na 1, možemo preskočiti skaliranje  $\tilde{V}$  kako bi dobili  $V$ . Korak 2 rješava prvih  $K$  svojstvenih vektora  $N \times K$  obično slabo popunjene (eng. sparse) matrice. Nadalje, to uzima najviše vremena u algoritmu. Korak 4 ima  $NK(K-1)$  množenja u odabiru  $K$  centara. Korak 6 uključuje  $NK^2$  množenja da izračuna  $\tilde{X}^*R^*$ . Korak 7 uključuje SVD od  $K \times K$  matrice i  $K^3$  množenja da izračuna  $R^*$ . Kako je  $X^*$  binarna matrica, računanje  $X^{*T}\tilde{X}^*$  može se svesti samo na zbrajanje. Sve zajedno, vremenska složenost algoritma je  $O(N^{\frac{3}{2}}K + NK^2)$ .

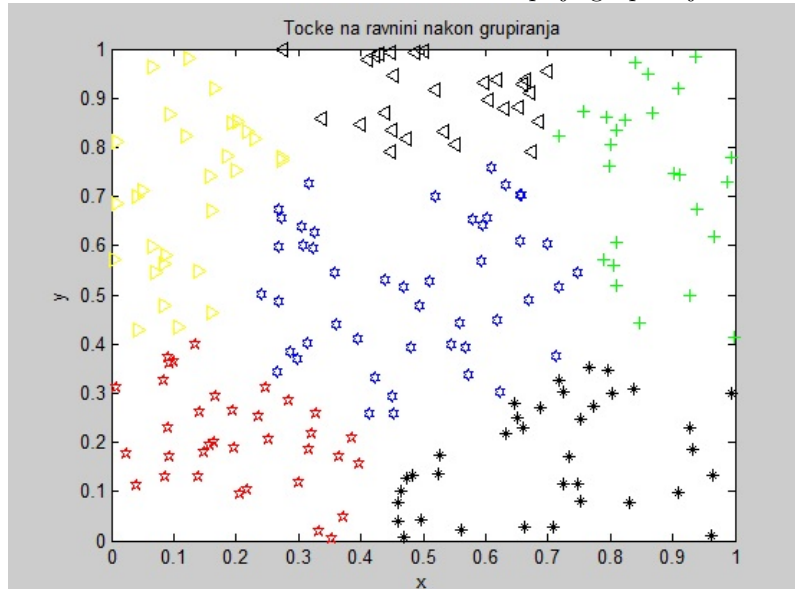
## 4.2 Primjeri

Promotrimo sada neke primjere na kojima je testiran naš program. Sve točke su nasumično odabrane. Težinska matrica zadana je formulom:

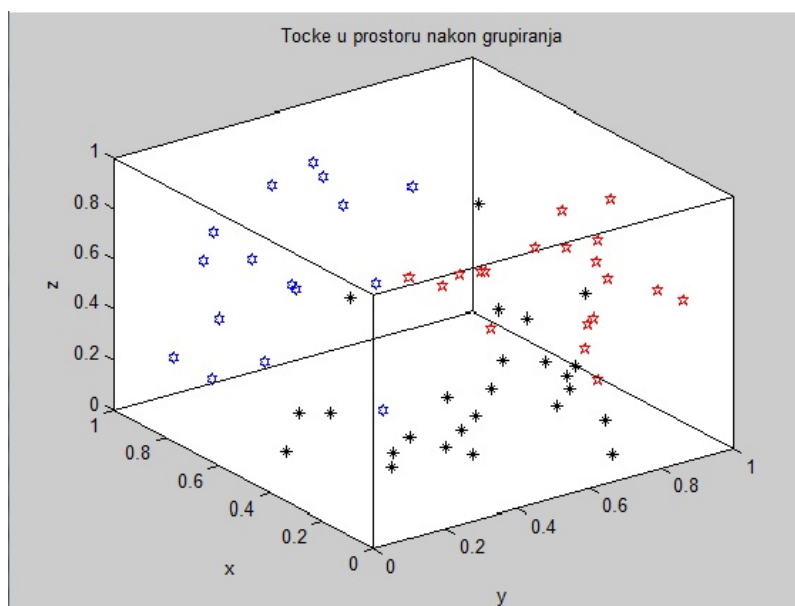
$$W(i, j) = \exp\left(\frac{-\|x_i - x_j\|^2}{\sigma}\right) \text{ , uz } \sigma = 1$$



Slika 1: Prikaz 600 tocaka u ravnini prije grupiranja



Slika 2: Prikaz 600 tocaka u ravnini grupiranih u 6 klastera



Slika 3: Prikaz 60 tocaka u prostoru grupiranih u 3 klastera

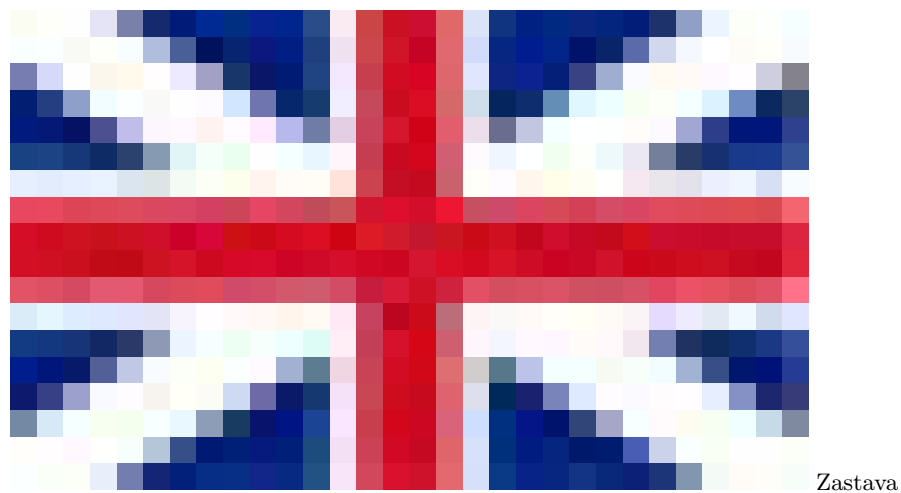
## 5 Zaključak

Dolazimo do zaključka da je višeklasno spektralno klasteriranje vrlo efikasna metoda za rješavanje problema grupiranja podataka. A.Y. Ng, M. Jordan i Y. Weiss su pokazali u svojem radu "On spectral clustering: Analysis and an algorithm" (2002.g) da K-sredine mogu dati iste rezultate kao i višeklasno spektralno klasteriranje, ali treba im čak i dvostruko duže vremena da konvergiraju ka rješenju. Nadalje, dok je u K-sredinama bitna dobra početna procjena, naša je metoda robusna na nasumičnu inicijalizaciju. Ipak, još jednom naglašavamo, dobra inicijalizacija može ubrzati konvergenciju. Kao što smo pokazali, konačno rješenje je skoro pa globalno optimalno.

U dodatku prilažemo rezultate koje smo dobili kada smo u program ubacili slike. Nažalost, naše računalo nije bilo u mogućnosti dobiti rješenje već za slike veličine  $120 \times 120$  - kao glavni problem pokazalo se računanje težinske matrice  $W$ .

## 6 Dodatak

Prikazujemo rezultate dobivene pri pokušaju grupiranja slike zastave Velike Britanije. Slika je veličine 18x30.



Velike Britanije

