SPEKTRALNO K-PARTICIONIRANJE GRAFA

Seminarski rad

Ivančica Mirošević

Fakultet elektrotehnike, strojarstva i brodogradnje
Sveučilište u Splitu

Smjernice

- Ciljne funkcije particioniranja
- Diskretni optimizacijski problem
- Relaksirani optimizacijski problem
- Ky-Fanov teorem
- Primjeri

Ciljne funkcije particioniranja

Neka je zadan jednostavni konačni neusmjereni težinski graf $G = (V, B)$ i $C_i \subset V$, $C_i \neq \emptyset$, $i = 1, \cdots, k$.

$$\mathrm{rez}(C_1, C_2, ..., C_k) = \sum_{i<j} \mathrm{rez}(C_i, C_j)$$

Ciljne funkcije particioniranja

Neka je $\pi = \{C_1, C_2, ..., C_k\}$ k-particija skupa vrhova V grafa G.
Razmjerni rez

$$R(C_1, C_2, \cdots, C_k) = \sum_{\substack{i,j=1 \\ i<j}}^{k} \left(\frac{\mathrm{rez}(C_i, C_j)}{|C_i|} + \frac{\mathrm{rez}(C_i, C_j)}{|C_j|} \right) = \sum_{i=1}^{k} \frac{\mathrm{rez}(C_i, V \backslash C_i)}{|C_i|}$$

Normalizirani rez

$$N(C_1, C_2, \cdots, C_k) = \sum_{\substack{i,j=1 \\ i<j}}^{k} \left(\frac{\mathrm{rez}(C_i, C_j)}{t(C_i)} + \frac{\mathrm{rez}(C_i, C_j)}{t(C_j)} \right) = \sum_{i=1}^{k} \frac{\mathrm{rez}(C_i, V \backslash C_i)}{t(C_i)}$$

Particijski vektori

$$
\begin{aligned}
\mathbf{h}_1 &= [\overbrace{1, \cdots, 1}^{|C_1|}, 0, \cdots, 0, \cdots, 0, \cdots, 0]^T \\
\mathbf{h}_2 &= [0, \cdots, 0, \overbrace{1, \cdots, 1}^{|C_2|}, \cdots, 0, \cdots, 0]^T \\
&\ldots \\
\mathbf{h}_k &= [0, \cdots, 0, 0, \cdots, 0, \cdots, \overbrace{1, \cdots, 1}^{|C_k|}]^T
\end{aligned}
$$

Particijski vektori

$$\mathrm{rez}(C_i, V \backslash C_i) = \mathbf{h}_i^T \left(D - W\right) \mathbf{h}_i = \mathbf{h}_i^T L \mathbf{h}_i,$$

$$\mathrm{rez}(C_i, V) = t\left(C_i\right) = \mathbf{h}_i^T D \mathbf{h}_i \quad \text{i}$$

$$|C_i| = \mathbf{h}_i^T \mathbf{h}_i,$$

Razmjerni rez

$$R\left(C_1, C_2, \cdots, C_k\right) = \frac{\mathbf{h}_1^T L \mathbf{h}_1}{\mathbf{h}_1^T \mathbf{h}_1} + \cdots + \frac{\mathbf{h}_k^T L \mathbf{h}_k}{\mathbf{h}_k^T \mathbf{h}_k}$$

Stavimo li

$$\mathbf{x}_i = \frac{\mathbf{h}_i}{\left\|\mathbf{h}_i\right\|_2} \quad , \quad i = 1, \cdots, k \ \text{ i } \ X = \left[\mathbf{x}_1, \cdots, \mathbf{x}_k\right]$$

razmjerni rez je jednak

$$R\left(C_1, C_2, \cdots, C_k\right) = \mathbf{x}_1^T L \mathbf{x}_1 + \cdots + \mathbf{x}_k^T L \mathbf{x}_k = \mathrm{tr}\left(X^T L X\right).$$

Normalizirani rez

$$N\left(C_1, C_2, \cdots, C_k\right) = \frac{\mathbf{h}_1^T L \mathbf{h}_1}{\mathbf{h}_1^T D \mathbf{h}_1} + \cdots + \frac{\mathbf{h}_k^T L \mathbf{h}_k}{\mathbf{h}_k^T D \mathbf{h}_k}.$$

Stavimo li

$$\mathbf{y}_i = \frac{D^{\frac{1}{2}} \mathbf{h}_i}{\left\| D^{\frac{1}{2}} \mathbf{h}_i \right\|_2}, \ i = 1, \cdots, k, \ \text{ i } \ Y = [\mathbf{y}_1, \cdots, \mathbf{y}_k],$$

tada je

$$\begin{aligned} N\left(C_1, C_2, \cdots, C_k\right) &= \mathbf{y}_1^T D^{-\frac{1}{2}} L D^{-\frac{1}{2}} \mathbf{y}_1 + \cdots + \mathbf{y}_k^T D^{-\frac{1}{2}} L D^{-\frac{1}{2}} \mathbf{y}_k \\ &= \operatorname{tr}(Y^T D^{-\frac{1}{2}} L D^{-\frac{1}{2}} Y) \\ &= \operatorname{tr}(Y^T L_n Y) \end{aligned}$$

Relaksirani problem k-particioniranja

$$\min R\left(C_1, C_2, \cdots, C_k\right) \geq \min_{\substack{X^T X = I \\ X \in \mathbb{R}^{n \times k}}} \operatorname{tr}(X^T L X),$$

$$\min N\left(C_1, C_2, \cdots, C_k\right) \geq \min_{\substack{Y^T Y = I \\ Y \in \mathbb{R}^{n \times k}}} \operatorname{tr}(Y^T L_n Y).$$

Ky-Fanov teorem

Teorem 1 *Neka je* $A \in \mathbb{R}^{n \times n}$ *simetrična matrica sa svojstvenim vrijednostima* $\lambda_1 \leq \cdots \leq \lambda_n$*. Tada je*

$$\min_{\substack{Z \in \mathbb{R}^{n \times k} \\ Z^T Z = I}} \operatorname{tr}\left(Z^T A Z\right) = \sum_{i=1}^{k} \lambda_i$$

Dokaz Ky-Fanovog teorema

Neka su zadani skupovi

$$\Phi_{n,k} = \{U \in \mathcal{S}_n : 0 \preceq U \preceq I, \operatorname{tr}(U) = k\}$$

i

$$\varphi_{n,k} = \{\mathbf{u} \in \mathbb{R}^n : 0 \leq u_i \leq 1 \text{ za } i = 1, \cdots, n, \ \sum_{i=1}^{n} u_i = k\}.$$

Dokaz Ky-Fanovog teorema

Lema 1 *Skupovi* $\Phi_{n,k}$ *i* $\varphi_{n,k}$ *su kompaktni i konveksni. Vrijedi*

$$U \in \Phi_{n,k} \Leftrightarrow Z^T U Z \in \Phi_{n,k},$$

gdje je $Z \in \mathcal{O}_{n,n}$. *Veze između* $\Phi_{n,k}$ *i* $\varphi_{n,k}$ *dane su izrazima*

$$\Phi_{n,k} = \{U \in \mathcal{S}_n : U = Z^T D Z,\ \ Z \in \mathcal{O}_{nn},\ D = \mathrm{diag}\left(\mathbf{u}\right)\ ,\ \ \mathbf{u} \in \varphi_{n,k}\},$$

i

$$\varphi_{n,k} = \{\mathbf{u} \in \mathbb{R}^n : \mathbf{u} = [U_{11}, \cdots, U_{nn}]^T, \text{ gdje je } \ U \in \Phi_{n,k}\}.$$

Dokaz Ky-Fanovog teorema

Neka za svojstvene vrijednosti matrice $A \in \mathcal{S}_n$ vrijedi

$$
\begin{aligned}
\lambda_1 \leq \cdots \leq \lambda_r < \\
\lambda_{r+1} = \cdots = \lambda_k = \cdots = \lambda_{r+t} < \\
\lambda_{r+t+1} \leq \cdots \leq \lambda_n,
\end{aligned}
\tag{1}
$$

gdje su $t \geq 1$ i $r \geq 0$ cijeli brojevi.

Dokaz Ky-Fanovog teorema

Lema 2 *Vrijedi*

$$\min_{\mathbf{u}\in\varphi_{n,k}} \sum_{i=1}^{n} \lambda_i u_i = \sum_{i=1}^{k} \lambda_i$$

s

$$\arg\min\{\sum_{i=1}^{n} \lambda_i u_i : \mathbf{u} \in \varphi_{n,k}\} = \{\mathbf{u} \in \mathbb{R}^n : u_i = 1, \quad i = 1, \cdots, r,$$

$$0 \le u_i \le 1, \quad i = r+1, \cdots, r+t, \text{ i } \sum_{i=r+1}^{r+t} u_i = k - r \text{ ,}$$

$$u_i = 0, \quad i = r+t+1, \cdots, n\} \text{ .}$$

Dokaz Ky-Fanovog teorema

Lema 3 *Neka je* $\Lambda = \operatorname{diag}(\lambda_1, \cdots, \lambda_n)$ *gdje za* $\lambda_1, \cdots, \lambda_n$ *vrijedi (1). Tada je*

$$\min_{U \in \Phi_{n,k}} \operatorname{tr}\left(\Lambda^T U\right) = \min_{U \in \Phi_{n,k}} \operatorname{tr}(\Lambda U) = \sum_{i=1}^{k} \lambda_i$$

i

$$\arg\min\{\operatorname{tr}(\Lambda U) : U \in \Phi_{n,k}\} = \left\{U \in \mathcal{S}_n : U = \begin{bmatrix} I & & \\ & \widetilde{U} & \\ & & 0 \end{bmatrix}, \widetilde{U} \in \Phi_{t,k-r}\right\}$$

Blokovi na dijagonali matrice U *su dimenzija* r, t, *i* $n - r - t$ *respektivno.*

Dokaz Ky-Fanovog teorema

Neka je

$$V^T A V = \Lambda,$$

gdje je $\Lambda = \mathrm{diag}\,(\lambda_1, \cdots, \lambda_n)$, i $V = [\mathbf{v}^{[1]}, \cdots, \mathbf{v}^{[n]}] \in \mathcal{O}_{n,n}$.

Dalje, neka je $P_1 = [\mathbf{v}^{[1]}, \cdots, \mathbf{v}^{[r]}]$ i $Q_1 = [\mathbf{v}^{[r+1]}, \cdots, \mathbf{v}^{[r+t]}]$.

Vrijedi

$$P_1^T A P_1 = \mathrm{diag}\,(\lambda_1, \cdots, \lambda_r)\,; \qquad (2)$$

$$Q_1^T A Q_1 = \lambda_k I.$$

Dokaz Ky-Fanovog teorema

Teorem 2 *Neka $A \in \mathcal{S}_n$ ima svojstvene vrijednosti $\lambda_1 \leq \cdots \leq \lambda_n$. Ako svojstvene vrijednosti zadovoljavaju (1), onda je*

$$\min_{U \in \Phi_{n,k}} \operatorname{tr}\left(A^T U\right) = \sum_{i=1}^{k} \lambda_i .$$

i

$$\begin{aligned} \arg\min\{\operatorname{tr}\left(A^T U\right) : U \in \Phi_{n,k}\} &= \\ = \{U \in \mathcal{S}_n : U &= P_1 P_1^T + Q_1 \widetilde{U} Q_1^T;\ \widetilde{U} \in \Phi_{t,k-r}\} \end{aligned}$$

gdje P_1, Q_1 zadovoljavaju (2).

Dokaz Ky-Fanovog teorema

Lema 4 *Ekstremne točke skupa* $\varphi_{n,k}$ *su*

$$\{\mathbf{u} \in \mathbb{R}^n : \mathbf{u} = [u_1, \cdots u_n]^T, u_i \in \{0,1\}, \sum_{i=1}^{k} u_i = k\}.$$

Lema 5 *Ekstremne točke skupa* $\Phi_{n,k}$ *su elementi skupa* $\Phi_{n,k}$ *ranga* k *, a to je skup matrica* $U \in \Phi_{n,k}$ *s* k *svojstvenih vrijednosti jednakih* 1*, i* $n-k$ *svojstvenih vrijednosti jednakih nuli.*

Donje međe vrijednosti ciljnih funkcija

Rješenje relaksiranog problema minimuma razmjernog reza k-particije dano je s

$$\mathbf{v}^{[1]}, \cdots, \mathbf{v}^{[k]},$$

i vrijedi donja međa

$$\lambda_1 + \cdots + \lambda_k \leq \min_{\pi=\{C_1,\cdots,C_k\}} R\left(C_1, C_2, \cdots, C_k\right),$$

gdje su $\lambda_1, \cdots, \lambda_k$ svojstvene vrijednosti Laplacijana grafa.

Donje međe vrijednosti ciljnih funkcija

Rješenje relaksiranog problema minimuma normaliziranog reza k-particije dano je s

$$D^{-\frac{1}{2}}\mathbf{w}^{[1]}, D^{-\frac{1}{2}}\mathbf{w}^{[2]}, \cdots, D^{-\frac{1}{2}}\mathbf{w}^{[k]},$$

i vrijedi donja međa

$$\mu_1 + \cdots + \mu_k \leq \min_{\pi=\{C_1,\cdots,C_k\}} N\left(C_1, C_2, \cdots, C_k\right),$$

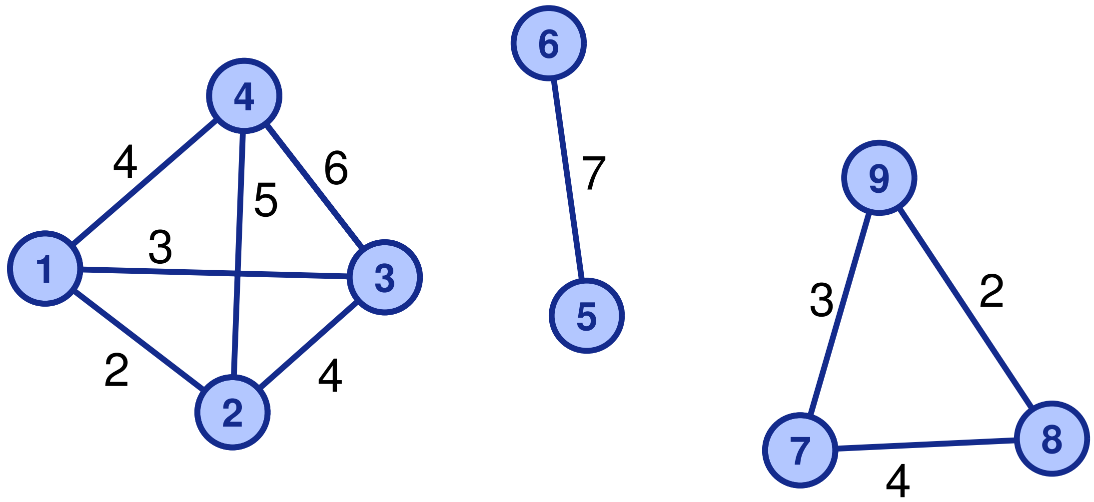
gdje su $\mu_1, \cdots, \mu_k$ svojstvene vrijednosti normaliziranog Laplacijana grafa.

Idealan slučaj

Teorem 3 *Neka je G težinski graf s n vrhova. Algebarska višestrukost k svojstvene vrijednosti 0 Laplacijana grafa G jednaka je broju povezanih komponenti grafa G.*

Ako svaka komponenta ima n_i vrhova, $i = 1, \cdots, k$, onda $n \times k$ matrica Z, čiji su stupci linearno nezavisni svojstveni vektori koji odgovaraju svojstvenoj vrijednosti 0, sadrži k različitih redaka r^i, $i = 1, \cdots, k$, i redak r^i pojavljuje se n_i puta. Štoviše, indeksi redaka koji su međusobno jednaki odgovaraju vrhovima koji pripadaju istoj komponenti grafa G.

Primjer 1a



Primjer 1a

$$L = \begin{bmatrix} 9 & -2 & -3 & -4 & 0 & 0 & 0 & 0 & 0 \\ -2 & 11 & -4 & -5 & 0 & 0 & 0 & 0 & 0 \\ -3 & -4 & 13 & -6 & 0 & 0 & 0 & 0 & 0 \\ -4 & -5 & -6 & 15 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 7 & -7 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -7 & 7 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 7 & -4 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 & -4 & 6 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & -3 & -2 & 5 \end{bmatrix}$$

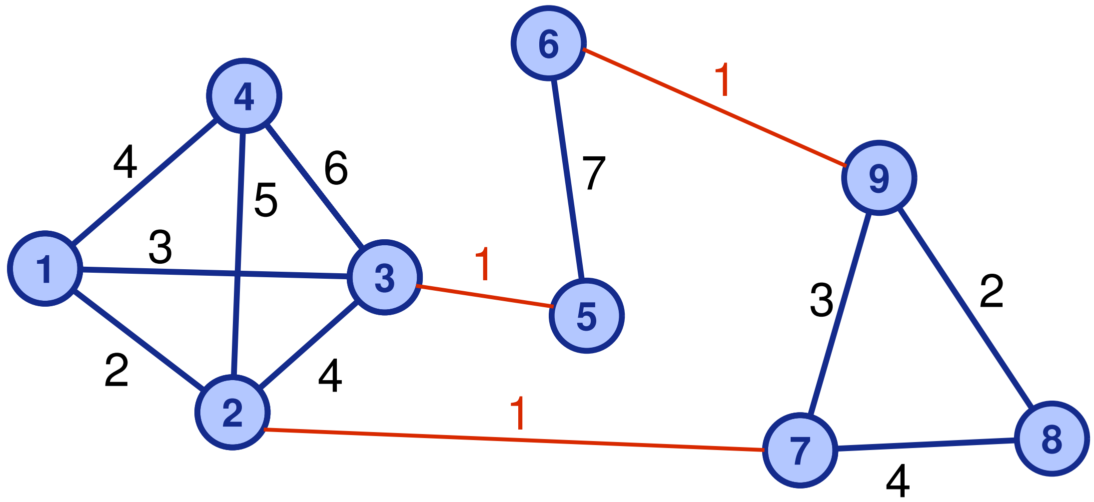
Primjer 1a

$$\Lambda = \mathrm{diag}(0, 0, 0, 7.27, 10.73, 11.52, 14, 16, 20.47).$$

$$\begin{aligned}
\mathbf{v}^{[1]} &= [-0.5, -0.5, -0.5, -0.5, 0, 0, 0, 0, 0]^T; \\
\mathbf{v}^{[2]} &= [0, 0, 0, 0, -0.71, -0.71, 0, 0, 0]^T; \\
\mathbf{v}^{[3]} &= [0, 0, 0, 0, 0, 0, -0.58, -0.58, -0.58]^T;
\end{aligned}$$

$$Z = \begin{bmatrix}
-0.5\alpha_1 & -0.5\alpha_2 & -0.5\alpha_3 \\
-0.5\alpha_1 & -0.5\alpha_2 & -0.5\alpha_3 \\
-0.5\alpha_1 & -0.5\alpha_2 & -0.5\alpha_3 \\
-0.5\alpha_1 & -0.5\alpha_2 & -0.5\alpha_3 \\
-0.71\beta_1 & -0.71\beta_2 & -0.71\beta_3 \\
-0.71\beta_1 & -0.71\beta_2 & -0.71\beta_3 \\
-0.58\gamma_1 & -0.58\gamma_2 & -0.58\gamma_3 \\
-0.58\gamma_1 & -0.58\gamma_2 & -0.58\gamma_3 \\
-0.58\gamma_1 & -0.58\gamma_2 & -0.58\gamma_3
\end{bmatrix}, \quad \alpha_i, \beta_i, \gamma_i \in \mathbb{R}.$$

Primjer 1b



Primjer 1b

$$L + E = \begin{bmatrix} 9 & -2 & -3 & -4 & 0 & 0 & 0 & 0 & 0 \\ -2 & 12 & -4 & -5 & 0 & 0 & -1 & 0 & 0 \\ -3 & -4 & 14 & -6 & -1 & 0 & 0 & 0 & 0 \\ -4 & -5 & -6 & 15 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 8 & -7 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -7 & 8 & 0 & 0 & -1 \\ 0 & -1 & 0 & 0 & 0 & 0 & 8 & -4 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 & -4 & 6 & -2 \\ 0 & 0 & 0 & 0 & 0 & -1 & -3 & -2 & 6 \end{bmatrix}$$

Primjer 1b

$$\Lambda = \mathrm{diag}(0, 0.79, 1.21, 7.9211.15, 12.07, 14.95, 17.08, 20.83).$$

$$Z = [\mathbf{v}^{[1]}, \mathbf{v}^{[2]}, \mathbf{v}^{[3]}] = \begin{bmatrix} 0.3333 & 0.3836 & 0.1215 \\ 0.3333 & 0.3235 & 0.1300 \\ 0.3333 & 0.3446 & 0.0753 \\ 0.3333 & 0.3673 & 0.1151 \\ 0.3333 & -0.0953 & -0.6121 \\ 0.3333 & -0.1474 & -0.6044 \\ 0.3333 & -0.3555 & 0.2825 \\ 0.3333 & -0.4248 & 0.3114 \\ 0.3333 & -0.3960 & 0.1808 \end{bmatrix}, \quad \alpha_i, \beta_i, \gamma_i \in \mathbb{R}.$$

$$||E|| = 2.4495$$

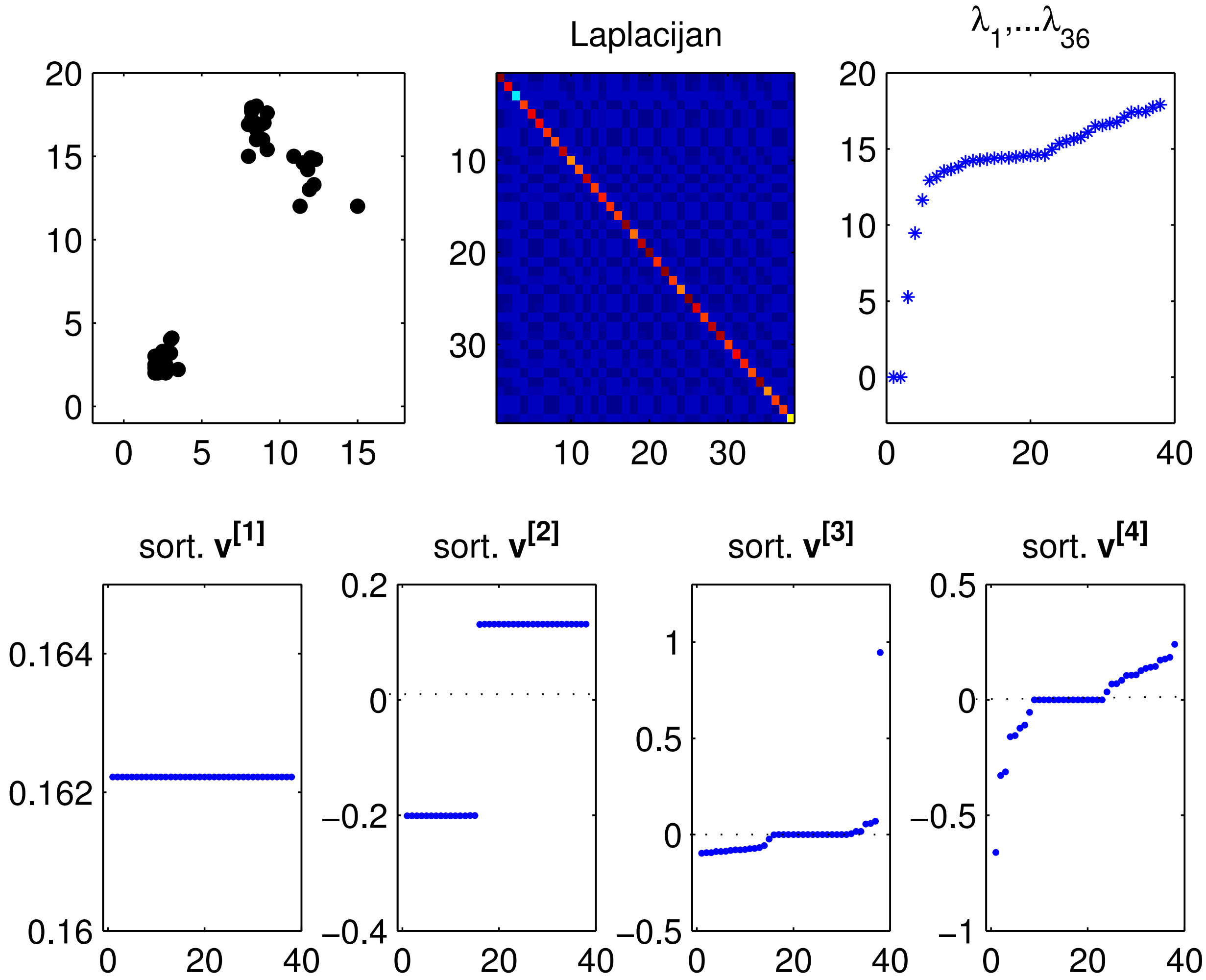
Stabilnost svojstvenih vrijednosti

Iz Weylovog teorema slijedi

Korolar 1 *Neka su A i E matrice u $\mathbb{R}^{n\times n}$ takve da su A i $A+E$ simetrične. Neka su $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ svojstvene vrijednosti od A, a $\lambda'_1 \leq \lambda'_2 \leq \cdots \leq \lambda'_n$ svojstvene vrijednosti od $A+E$. Tada za svaki $k = 1, 2, \cdots, n$ vrijedi*

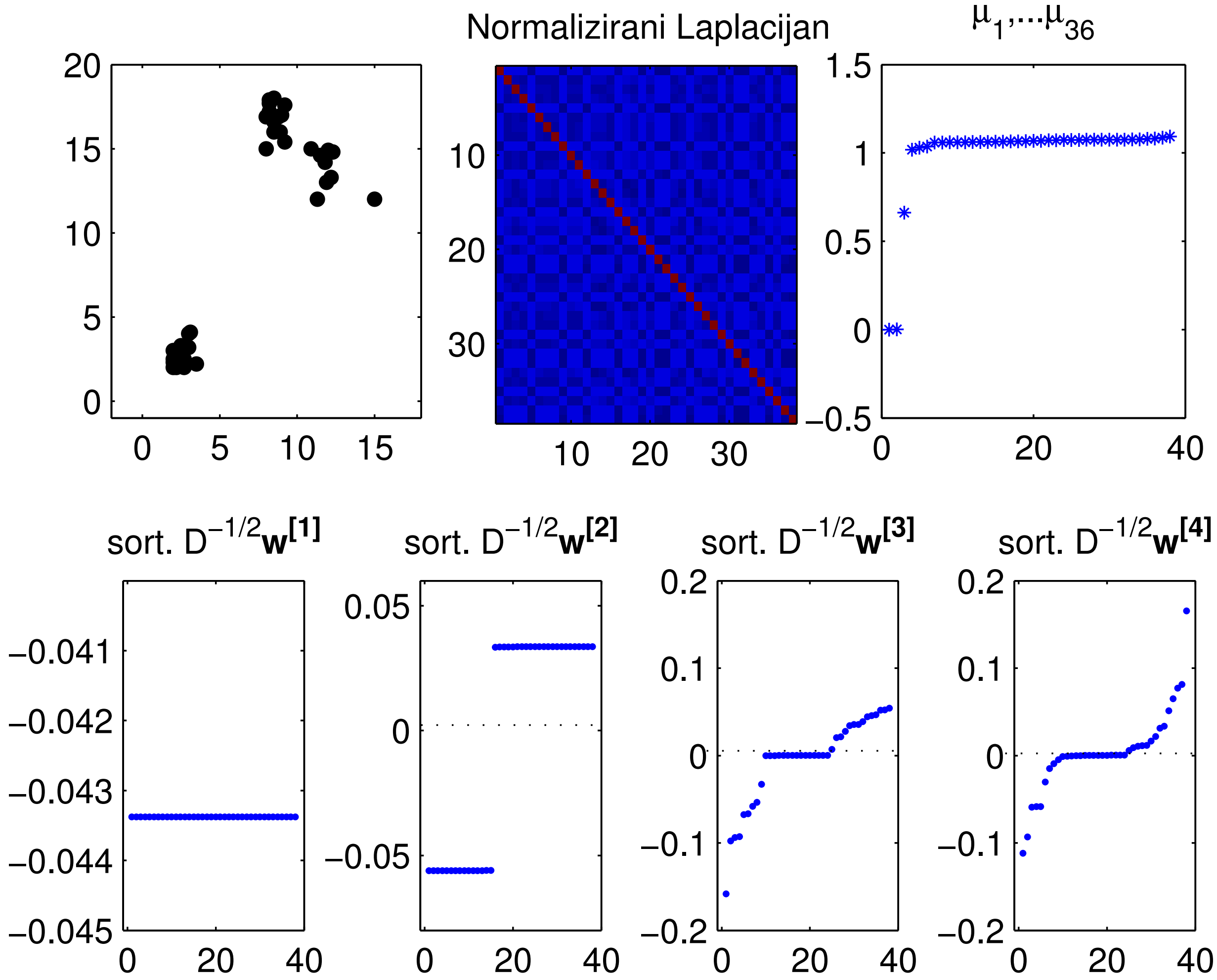
$$|\lambda'_k - \lambda_k| \leq ||E||_2.$$

Primjer 2



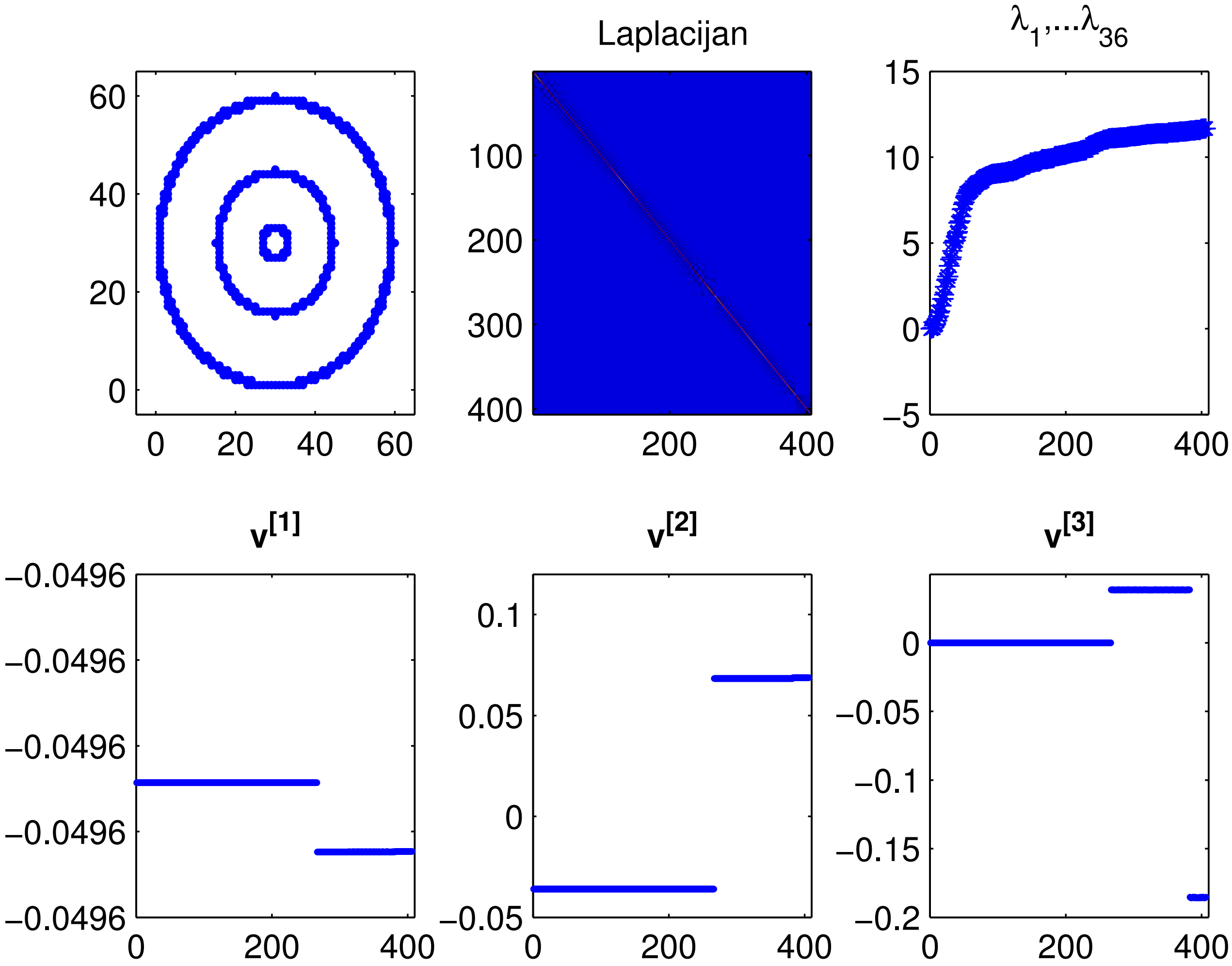
Laplacijan
λ1,...λ36
20
15
10
5
0
0
5
10
15
10
20
30
10
20
30
20
15
10
5
0
0
20
40
sort. v[1]
sort. v[2]
sort. v[3]
sort. v[4]
0.164
0.162
0.16
0
20
40
0.2
0
−0.2
−0.4
0
20
40
1
0.5
0
−0.5
0
20
40
0.5
0
−0.5
−1
0
20
40

Primjer 3



Normalizirani Laplacijan
$\mu_1,\ldots\mu_{36}$
sort. $D^{-1/2}\mathbf{w}^{[1]}$
sort. $D^{-1/2}\mathbf{w}^{[2]}$
sort. $D^{-1/2}\mathbf{w}^{[3]}$
sort. $D^{-1/2}\mathbf{w}^{[4]}$

Primjer 4



Laplacijan
λ1,...λ36
60
40
20
0
0 20 40 60
100
200
300
400
200 400
15
10
5
0
−5
0 200 400
v[1]
v[2]
v[3]
−0.0496
−0.0496
−0.0496
−0.0496
−0.0496
0 200 400
0.1
0.05
0
−0.05
0 200 400
0
−0.05
−0.1
−0.15
−0.2
0 200 400

k-particijski algoritam za minimiziranje razmjernog reza

1. Za zadanu matricu susjedstva W izračunaj $L = D - W$;
2. Izračunaj k svojstvenih vektora matrice L, $\mathbf{v}^{[1]}, ..., \mathbf{v}^{[k]}$, i formiraj matricu Z

$$Z = [\mathbf{v}^{[1]}, ..., \mathbf{v}^{[k]}];$$

3. Pokreni algoritam k-means nad retcima matrice Z.

K-means algoritam će na izlazu dati središta $\mathbf{c}_1, ..., \mathbf{c}_k \in \mathbb{R}^k$ skupina, te vektor $[sk_1, sk_2, \cdots, sk_n]$ gdje $sk_i \in \{1, 2, \cdots, k\}$, $i = 1, \cdots, n$, označava redni broj skupine kojoj i-ti redak, odnosno i -ti podatak, pripada.

k-particijski algoritam za minimiz. normaliziranog reza

1. Za zadanu matricu susjedstva W izračunaj $L_n = D^{-\frac{1}{2}} \left(D - W\right) D^{-\frac{1}{2}}$;
2. Izračunaj k svojstvenih vektora matrice L_n, $\mathbf{w}^{[1]}, \ldots, \mathbf{w}^{[k]}$, i formiraj matricu

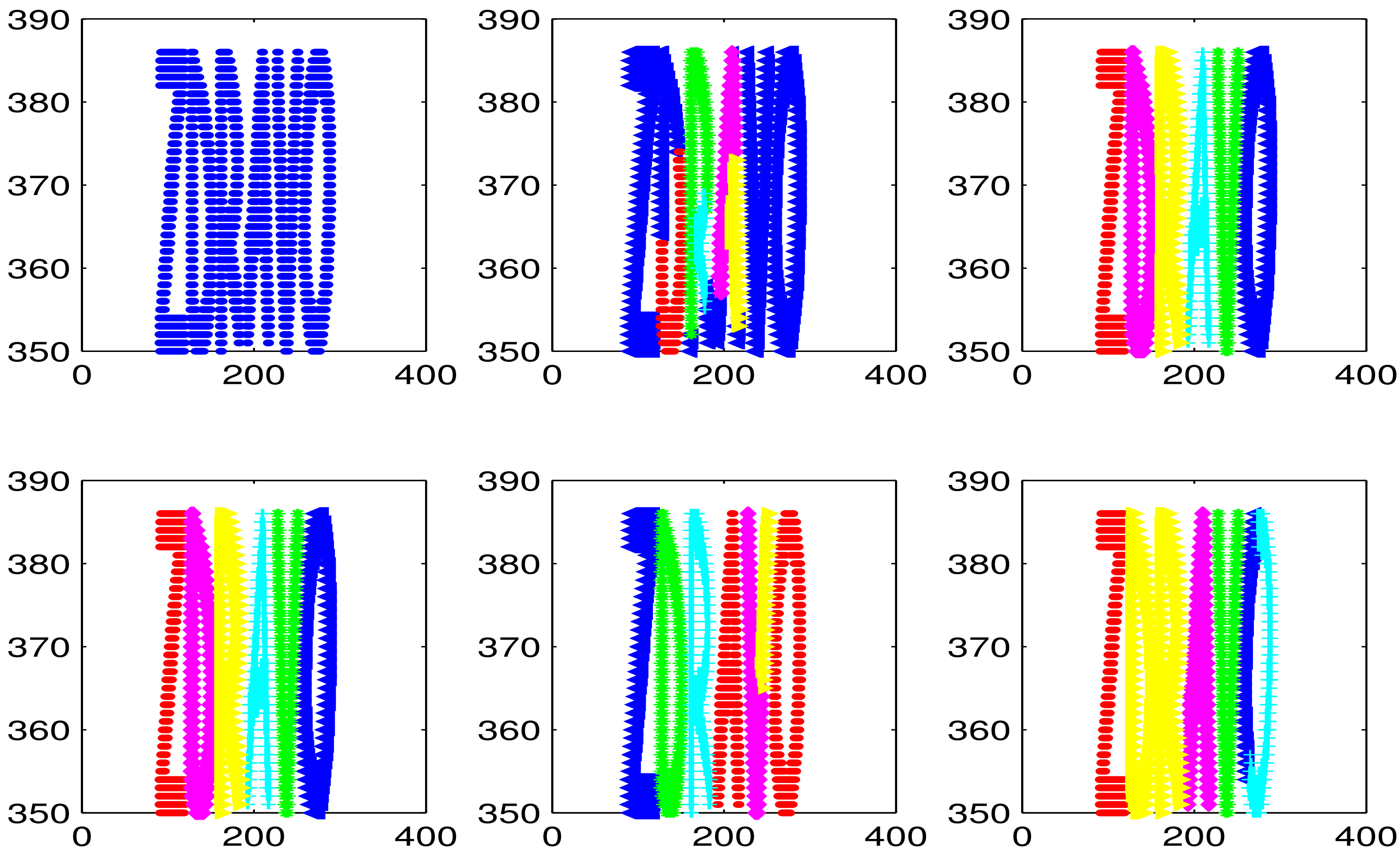
$$Z = D^{-\frac{1}{2}} U,$$

gdje je

$$U = \left[\mathbf{w}^{[1]}, \ldots, \mathbf{w}^{[k]}\right];$$

3. Pokreni algoritam k-means nad retcima matrice Z.

Primjer 5



Usporedba algoritama

Primjer	n	k		km	rbp-rr	rbp-nr	mp-rr	mp-nr
4	406	3	t	0.7475s	0.256s	0.265s	0.15s	0.164s
			rr	1.613	0.2	0.2	0.827	0.994
			nr	0.141	0.018	0.018	0.07	0.087
5	2829	6	t	83.45s	68.27s	73.84s	33.215s	51.87s
			rr	30.9181	0.7732	0.7732	5.1882	6.039
			nr	0.72	0.0967	0.0967	0.2458	0.2697