Seminar za optimizaciju i primjene
Odjel za matematiku, Sveučilište J.J.Strossmayera u Osijeku

Prepoznavanje nekih geometrijskih objekata u ravnini

Rudolf Scitovski, Odjel za matematiku, Sveučilište u Osijeku

Kristian Sabo, Odjel za matematiku, Sveučilište u Osijeku

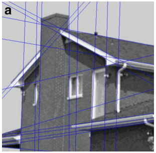
Keywords: Multiple circle detection; Multiple ellipse detection; Multiple generalized circle detection; DIRECT, k-means, Incremental algorithm.

Osijek, 12. prosinca, 2018.

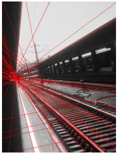
This work was supported by the Croatian Science Foundation through research grants: The optimization and statistical models and methods in recognizing properties of data sets measured with errors (OSMoMeSIP-IP-2016-06-6545)

Motivacija i primjene

Pravci kuće

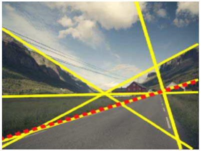


Kolodvor



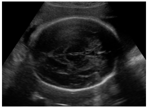
Perspektiva

Perspektiva

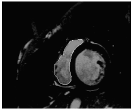


Biometric Measurements

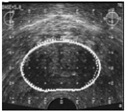
Glava fetusa



Rameni zglob

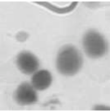


Prostata



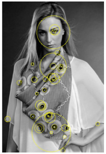
Object Detection

Rak dojke

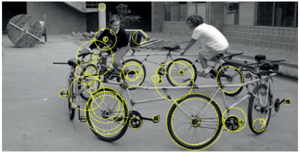


Više embrija

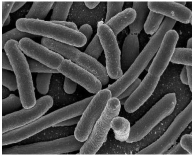
Kružnice



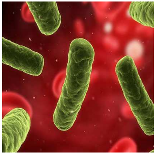
Elipse



Escherichia Coli



Enterobacter-cloaca

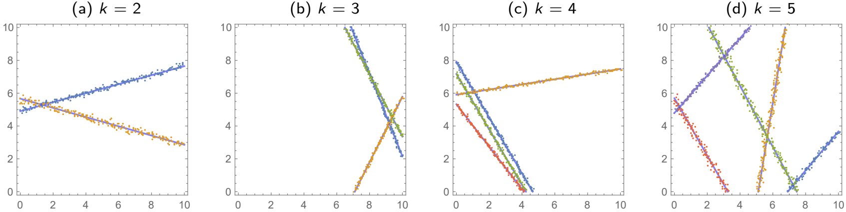


Prepoznavanje više pravaca u ravnini

R. Scitovski, K. Sabo, U. Radojičić, *A fast and efficient method for solving the multiple line detection problem* - na recenziji

$\mathcal{A} = \{a^i = (x_i, y_i)\colon\, i = 1, \dots, m\} \subset [0, a] \times [0, b] \subset \mathbb{R}^2$
(skup podataka koji dolazi od k unaprijed nepoznatih pravaca)

$$\ell_j \equiv u_j x + v_j y + z_j = 0, \quad u_j^2 + v_j^2 = 1, \quad j = 1, \dots, k,$$



$$\operatorname*{argmin}_{\mathbf{u},\mathbf{v},\mathbf{z}\in\mathbb{R}^k} F(\mathbf{u}, \mathbf{v}, \mathbf{z}), \qquad F(\mathbf{u}, \mathbf{v}, \mathbf{z}) = \sum_{i=1}^{m} \min_{1\le j\le k} \mathfrak{D}(\ell_j(u_j, v_j, z_j), a^i),$$

$$\mathfrak{D}(\ell, a^i) = \frac{(ux_i + vy_i + z)^2}{u^2 + v^2}.$$

Prepoznavanje više kružnica u ravnini

R. Scitovski, K. Sabo, *Application of the* `DIRECT` *algorithm to searching for an optimal k-partition of the set $\mathcal{A}\subset\mathbb{R}^n$ and its application to the multiple circle detection problem* - Revised Version na recenziji u `Journal of Global Optimization`

Problem 1: Skup $\mathcal{A}=\{a^i\in\mathbb{R}^n\colon i=1,\dots,m\}$ treba grupirati u k nepraznih disjunktnih klastera $\pi_1,\dots,\pi_k$, $1\leq k\leq m$ rješavanjem problema

$$\operatorname*{argmin}_{c\in\operatorname{conv}(\mathcal{A})^k} F(c), \qquad F(c)=\sum_{i=1}^{m}\min_{1\leq j\leq k}\|c_j-a^i\|^2$$

Problem 2: Skup $\mathcal{A}=\{a^i=(x_i,y_i)\in\mathbb{R}^2\colon \alpha_1\leq x_i\leq\beta_1,\ \alpha_2\leq y_i\leq\beta_2, i=1,\dots,m\}$ dolazi od k unaprijed nepoznatih kružnica koje treba rekonstruirati

$$\operatorname*{argmin}_{(\mathbf{p},\mathbf{q})\in[\alpha,\beta]^k,\mathbf{r}\in[0,R]^k} F(\mathbf{p},\mathbf{q},\mathbf{r}), \quad F(\mathbf{p},\mathbf{q},\mathbf{r})=\sum_{i=1}^{m}\min_{1\leq j\leq k}\{\mathfrak{D}((p_j,q_j),r_j),a^i)\},$$

$$\mathfrak{D}((p_j,q_j),r_j),a^i)=(\|(p_j,q_j)-a^i\|^2-r_j^2)^2$$

O Problemu 1

Theorem 1

Neka je $\mathcal{A} = \{a^i = (a^i_1, \dots, a^i_n) \in [\alpha, \beta] \colon i = 1, \dots, m\} \subset \mathbb{R}^n$*, gdje je* $[\alpha, \beta] = \{x \in \mathbb{R}^n \colon \alpha_i \le x_i \le \beta_i\}$ *i* $\alpha = (\alpha_1, \dots, \alpha_n)^T$, $\beta = (\beta_1, \dots, \beta_n)^T \in \mathbb{R}^n$*. Funkcija* $F \colon [\alpha, \beta]^k \to \mathbb{R}_+$, $F(c) = \sum_{i=1}^m \min_{j=1,\dots,k} \|c_j - a^i\|^2$ *je Lipschitz neprekidna na* $[\alpha, \beta]^k$*.*

Algorithm 1 : GOPart$(\mathcal{A}, k)$

Input: $\mathcal{A} \subset [\alpha, \beta]^n$ {Set of data points}; $k \ge 2$ $\epsilon > 0$;

1: Define the mapping $T^{-1} \colon [0,1]^{kn} \to [\alpha, \beta]^k$, $T^{-1}(x) = D^{-1}x + u$ and the objective function $f = F \circ T^{-1}$, where $D = \operatorname{diag}\left(\frac{1}{\beta_1 - \alpha_1}, \dots, \frac{1}{\beta_n - \alpha_n}, \dots, \frac{1}{\beta_1 - \alpha_1}, \dots, \frac{1}{\beta_n - \alpha_n}\right) \in \mathbb{R}^{(kn)\times(kn)}$, and $u = (\alpha_1, \dots, \alpha_n, \dots, \alpha_1, \dots, \alpha_n) \in \mathbb{R}^{kn}$;

2: By using the DIRECT algorithm find the initial approximation $\hat{x} = (\hat{x}_1, \dots, \hat{x}_k) \in [0,1]^{kn}$;

3: By using the standard k-means algorithm with the initial approximation $\hat{x}$ determine cluster centers $x^\star = (x^\star_1, \dots, x^\star_k)$;

4: Calculate $c^\star = (c^\star_1, \dots, c^\star_k) = T^{-1}(x^\star) \in [\alpha, \beta]^k$;

Output: $\{c^\star, F(x^\star)\}$.

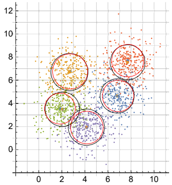
Testiranje Problema 1: $\mathcal{A} \subset \mathbb{R}^2$

Algorithm	CPU-time (sec.)	Detection of the best partition					Main circles recognized					
		6	5	4	3	2	5	4	3	2	1	0
`GOPart`	12.80	2	91	7	-	-	18	26	28	18	9	1
`Incremental`	8.41	3	88	9	-	-	15	28	18	25	11	3
k-`means`	98.13	-	92	8	-	-	18	17	33	19	8	5
`DIRECT`	688.21	-	100	-	-	-	21	31	25	23	-	-

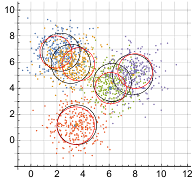
(Početna aproksimacija za `GOPart` dobivena u 5 iteracija `DIRECT` algoritma)

$$C_j(c_j, \sigma_j) = \{x \in \mathbb{R}^2 \colon \|c_j - x\| = \sigma_j\}, \quad \sigma_j^2 = \frac{1}{|\pi_j|} \sum_{a \in \pi_j} \|c_j - a\|^2$$ - main circle

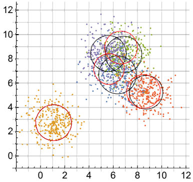
5-part: 5 main circles



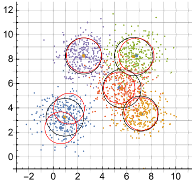
5-part: 1 main circle



4-part: 2 main circles



6-part: 2 main circles



Testiranje Problema 1: $\mathcal{A} \subset \mathbb{R}^5$

Problem 1: Skup $\mathcal{A} = \{a^i \in \mathbb{R}^5 \colon i = 1, \ldots, m\}$ treba grupirati u k nepraznih disjunktnih klastera $\pi_1, \ldots, \pi_k$, $1 \leq k \leq m$ rješavanjem problema

$$\operatorname*{argmin}_{c \in \mathrm{conv}(\mathcal{A})^k} F(c), \qquad F(c) = \sum_{i=1}^{m} \min_{1 \leq j \leq k} \|c_j - a^i\|^2$$

Algorithm	CPU-time (sec.)	Detection of the best partition					Main circles recognized					
		6	5	4	3	2	5	4	3	2	1	0
GOPart	22.32	4	94	2	-	-	82	2	6	10	-	-
Incremental	129.27	3	91	6	-	-	80	9	6	5	-	-
k-means	114.13	3	92	5	-	-	90	-	5	3	1	1
DIRECT	1490.83	-	100	-	-	-	81	-	12	7	-	-

(Početna aproksimacija za GOPart dobivena u 6 iteracija DIRECT algoritma)

O Problemu 2

$\mathcal{A} = \{a^i = (x_i, y_i) \in \mathbb{R}^2 : i = 1, \ldots, m\} \subset [\alpha, \beta], \qquad [\alpha, \beta] = [\alpha_1, \beta_1] \times [\alpha_2, \beta_2]$

(Skup podataka potječe od unaprijed nepoznatog broja kružnica u ravnini homogeno distribuiranog u okolini kružnica)

Metode: R. Scitovski, T. Marošević, *Pattern Recognition Letters*, **52**(2014)

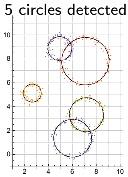
The k-closest circles algorithm (KCC)

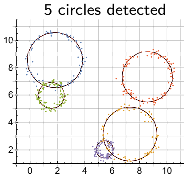
Multiple Circle Detection (MCD): DIRECT + KCC

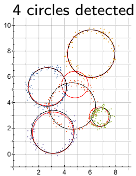
Modified multiple Circle Detection (MMCD)

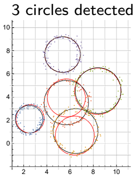
$$\underset{(\mathbf{p},\mathbf{q})\in[\alpha,\beta]^k}{\operatorname{argmin}} \tilde{F}(\mathbf{p},\mathbf{q}), \quad \tilde{F}(\mathbf{p},\mathbf{q}) = \sum_{i=1}^{m} \min_{1\le j\le k} \{\mathfrak{D}((p_j, q_j), r_j), a^i)\}, \quad r_j = \text{constant}$$

Algorithm	CPU (sec.)	No. of recognized circles 5	4	3	2	1	0
MCD	2.47	87	9	4	-	-	-
MMCD	1.96	88	9	2	1	-	-
Incremental	36.85	26	3	25	21	17	8
k-means	38.00	89	4	6	1	-	-
DIRECT	459.34	85	6	9	1	-	-









GO-index

$\Pi^\star = \{\pi_1^\star, \ldots, \pi_k^\star\}$ – globalno optimalna k-particija s centrima $\mathcal{C}_j^\star$

$\frac{1}{|\pi_j^\star|} \sum\limits_{a\in\pi_j^\star} \mathfrak{D}(\mathcal{C}_j^\star, a)$ – prosječno kvadratno odstupanje od $\mathcal{C}_j^\star$;

$\frac{|\pi_j^\star|}{|\mathcal{C}_j^\star|}$ – gustoča točaka u okolini $\mathcal{C}_j^\star$;

Geometrical Objects-index (`GO-index`):

$$\mathtt{GO}(k) := \mathtt{GO}(\Pi^\star(k)) = k \sum_{j=1}^{k} \frac{|\mathcal{C}_j^\star|}{|\pi_j^\star|^2} \sum_{a\in\pi_j^\star} \mathfrak{D}(\mathcal{C}_j^\star, a)$$

Najmanji $\mathtt{GO}(k)$ identificira paticiju s najprihvatljivijim brojem klastera

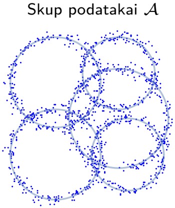
Density-Based Clustering index (DBC-index)

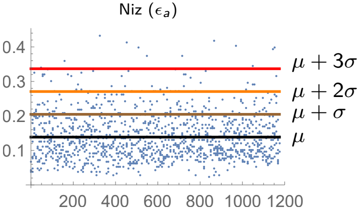
DBC-parametar $\epsilon > 0$
(M.Ester et all, 2nd International Conference on Knowledge Discovery and Data Mining, 1996)

Za svaki $a \in \mathcal{A}$ definiramo radijus ϵ_a kružića u kome se nalazi *inPts* > 2 elemenata skupa $\mathcal{A}$

$\mathcal{R} = \{\epsilon_a : a \in \mathcal{A}\}$

$$\epsilon = \mu + 3\sigma, \quad \mu = \mathrm{Mean}\,(\mathcal{R}), \quad \sigma^2 = \frac{1}{m-1}\sum_{a\in\mathcal{A}}(\epsilon_a - \mu)^2$$





$\Pi^\star = \{\pi_1^\star, \ldots, \pi_k^\star\}$ – globalno optimalna k-particija s centrima $\mathcal{C}_j^\star$

$V_j = \{\mathfrak{D}_1(\mathcal{C}_j^\star, a) : a \in \pi_j^\star\}$ – skup $\mathfrak{D}_1$-udaljenosti točaka iz $\pi_j^\star$ do $\mathcal{C}_j^\star$;

$$\mathtt{QAD[j]} := \mu_j + 2\sigma_j, \quad \mu_j = \mathrm{Mean}\,(V_j), \quad \sigma_j^2 = \frac{1}{|V_j|-1}\sum_{v_j\in V_j}(v_j - \mu_j)^2, \quad \text{j=1,\ldots,k}$$

$\mathtt{DBC}(k) := |\{j : \mathtt{QAD[j]} < \epsilon\}|$

Sve kružnice su prepoznate ako je $\mathtt{DBC}(k) = k$

Prepoznavanje više elipsi u ravnini

P. Nikić, R. Scitovski, K. Sabo, S. Majstorović, *A fast algorithm for solving the multiple ellipse detection problem* - KOI2018, Zadar, September 26–28, 2018

U. Radojičić, R. Scitovski, K. Sabo, *A fast and efficient method for solving the multiple closed curve detection problem* - 8^{th} International Conference on Pattern Recognition Applications and Methods, ICPRAM2019, Prague, 19-21- February, 2019

$\mathcal{A} = \{a^i = (x_i, y_i) \in \mathbb{R}^2 \colon i = 1, \ldots, m\} \subset [\alpha, \beta], \quad [\alpha, \beta] := [\alpha_1, \beta_1] \times [\alpha_2, \beta_2],$

Skup dolazi od k unaprijed nepoznatih elipsi koje treba rekonstruirati

$$\underset{\substack{S \in \Delta^k \\ \xi, \eta \in [0,R]^k, \vartheta \in [0,\pi]^k}}{\operatorname{argmin}} F(S, \xi, \eta, \vartheta), \quad F(S, \xi, \eta, \vartheta) = \sum_{i=1}^{m} \min_{1 \le j \le k} \mathfrak{D}(E_j(S_j, \xi_j, \eta_j, \vartheta_j), a^i),$$

$\mathfrak{D}(E_j(S_j, \xi_j, \eta_j, \vartheta_j), a^i)$ – udaljenost točke a^i do elipse E_j

[1] A.Y. Uteshev, M.V. Goncharova, *JCAM* **328**(2018)

[2] C. Akinlar, C. Topal, *Pattern Recognition* **46**(2013)

[3] D.K. Prasad, M.K.H. Leung, C. Quek, *Pattern Recognition* **46**(2013)

[4] T. Marošević, R. Scitovski, *Croatian Operational Research Review* **6**(2015)

[5] R. Grbić, D. Grahovac, R. Scitovski, *Pattern Recognition* **60**(2016)

[6] B.K. Kwon et al., *International Journal of Control, Automation and Systems* **14**(2016)

Elipsa kao Mahalanobis kružnica

Elipsa $E(S,\xi,\eta,\vartheta)$, $S=(p,q)^T$:

$$\frac{[(x-p)\cos\vartheta+(y-q)\sin\vartheta]^2}{\xi^2}+\frac{[(x-p)\sin\vartheta-(y-q)\cos\vartheta]^2}{\eta^2}=1,$$

uz oznake: $\Sigma = U(\vartheta)\begin{bmatrix}\xi^2 & 0\\ 0 & \eta^2\end{bmatrix}U^T(\vartheta)$, $\quad U(\vartheta)=\begin{bmatrix}\cos\vartheta & -\sin\vartheta\\ \sin\vartheta & \cos\vartheta\end{bmatrix}$ može se zapisati kao

$$E=\{u\in\mathbb{R}^2\colon (S-u)^T\Sigma^{-1}(S-u)=1\}$$

Ako definiramo $r:=\sqrt{\xi\eta}$, tj. $r^2=\sqrt{\det\Sigma}$, onda

$$E=\{u\in\mathbb{R}^2\colon \sqrt{\xi\eta}(S-u)^T\Sigma^{-1}(S-u)=r^2\}$$

Definition 1

Mahalanobis circle (M-circle) $\mathfrak{M}(S,r;\Sigma)$ with the center at the point $S\in\mathbb{R}^2$, radius r and the covariance matrix Σ, is the set of points

$$\mathfrak{M}(S,r;\Sigma)=\{u\in\mathbb{R}^2\colon d_M(S,u;\Sigma)=r^2\},$$

where $d_M\colon \mathbb{R}^2\times\mathbb{R}^2\to\mathbb{R}_+$ given by

$$d_M(S,u;\Sigma)=\sqrt{\det\Sigma}(S-u)^T\Sigma^{-1}(S-u)=\|u-v\|_\Sigma^2$$

is the *Mahalanobis distance-like function*.

$\mathfrak{M}(S,1;\Sigma)$ - normalizirana Mahalanobis kružnica

Elipsa kao Mahalanobis kružnica

Lemma 1

Ellipse $E(S,\xi,\eta,\vartheta)$ *can be considered as an M-circle* $\mathfrak{M}(S,r;\Sigma)$*, with* $r=\sqrt{\xi\eta}$*. Conversely, an M-circle* $\mathfrak{M}(S,r;\Sigma)$ *is an ellipse* $E(S,\xi,\eta,\vartheta)$*, where*
$\mathrm{diag}(\xi^2,\eta^2)=U^T(\vartheta)\big(\frac{r^2}{\sqrt{\det\Sigma}}\Sigma\big)U(\vartheta)$, $\Sigma=\begin{bmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{12} & \sigma_{22}\end{bmatrix}$, $U(\vartheta)=\begin{bmatrix}\cos\vartheta & -\sin\vartheta\\ \sin\vartheta & \cos\vartheta\end{bmatrix}$ *and*

$$\vartheta=\tfrac{1}{2}\arctan\tfrac{2\sigma_{12}}{\sigma_{11}-\sigma_{22}}\in\begin{cases}[0,\pi/4), & \sigma_{12}\ge 0\ \&\ \sigma_{11}>\sigma_{22}\\ [\pi/4,\pi/2), & \sigma_{12}>0\ \&\ \sigma_{11}\le\sigma_{22}\\ [\pi/2,3\pi/4), & \sigma_{12}\le 0\ \&\ \sigma_{11}<\sigma_{22}\\ [3\pi/4,\pi), & \sigma_{12}<0\ \&\ \sigma_{11}\ge\sigma_{22}\end{cases}.$$

$$\mathfrak{D}_1(\mathfrak{M}(S,r,\Sigma),a)=|\|S-a\|_\Sigma-r|\quad\text{(LAD-distance)},$$
$$\mathfrak{D}(\mathfrak{M}(S,r,\Sigma),a)=(|\|S-a\|_\Sigma^2-r^2)^2\quad\text{(algebraic distance)}.$$

$$\underset{\substack{S\in\Delta^k\\ r\in[0,R]^k,\Sigma\in M_2^k}}{\operatorname{argmin}}\ F(S,r,\Sigma),\qquad F(S,r,\Sigma)=\sum_{i=1}^m\min_{1\le j\le k}\mathfrak{D}(\mathfrak{M}_j(S_j,r_j,\Sigma_j),a^i),$$

Algorithm 2 (Optimal k-partition)

Input: $\mathcal{A} \subset \Delta$ {Set of data points}; $k \geq 1$;

1: By using the DIRECT algorithm solve simple GOP $\underset{S_1 \dots S_k \in \Delta}{\operatorname{argmin}} \sum_{i=1}^{m} \min_{1 \leq j \leq k} \mathfrak{D}(\mathcal{C}_j(S_j, 1), a^i)$ and denote the solution by S^0;

2: Define initial approximation $(S_j^0, 1, \mathbf{I})$, $j = 1, \dots, k$ and apply the KMCC algorithm.
Denote the solution by $\mathfrak{M}_j^\star(S_j^\star, r_j^\star, \Sigma_j^\star)$, $j = 1, \dots, k$;
Denote the corresponding optimal partition by $\Pi^\star$ and the objective function value by $F^\star$;

3: By using Lemma 1 determine the ellipses $E_j^\star(S_j^\star, \xi_j^\star, \eta_j^\star, \vartheta_j^\star)$, $j = 1, \dots, k$;

4: Compute value GO(k) (or QAD(k));

Output: $\{k, \Pi^\star, F^\star, \mathtt{GO}(k), E_j^\star(S_j^\star, \xi_j^\star, \eta_j^\star, \vartheta_j^\star), j = 1, \dots, k\}$.

Algorithm 3 (Optimal k-partition with the most appropriate number of clusters)

Input: $\mathcal{A} \subset \Delta$ {Set of data points}; $\epsilon > 0$;

1: Set $r = 1$ and call Algorithm 2;

2: Set $r = r + 1$ and call Algorithm 2;

3: **if** $(F_{r-1} - F_r)/F_1 > \epsilon$, **then**

4: Set $r = r + 1$ and call Algorithm 2 and go to Step 3;

5: **end if**

6: Find a k-partition $\Pi^\star$ $(k \leq r)$ with the smallest GO-index (or greatest DBC-index)
and denote corresponding cluster-centers by $E_1^\star, \dots, E_k^\star$;

Output: $\{k, \Pi^\star, \{E_1^\star, \dots, E_k^\star\}\}$.

R. Scitovski, K. Sabo, *A fast and efficient method for solving the multiple generalized circle detection problem* - Int. Conf. on Applied Mathematics & Computational Science, ICAMCS2018, Budapest, 6-8 October, 2018

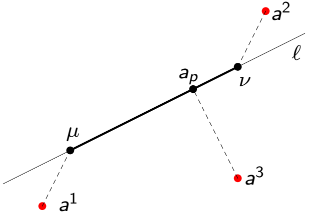
$\mathcal{A} = \{a^i = (x_i, y_i) \in \mathbb{R}^2 \colon i = 1, \dots, m\} \subset [\alpha, \beta], \qquad [\alpha, \beta] = [\alpha_1, \beta_1] \times [\alpha_2, \beta_2]$

(Skup podataka potječe od unaprijed nepoznatog broja kružnica u ravnini homogeno distribuiranog u okolini kružnica)

Generalizirana kružnica: $\mathcal{C}(\mu, \nu, r) = \{x \in \mathbb{R}^2 \colon D(\mathrm{seg}(\mu, \nu), x) = r\}$,

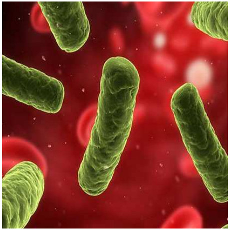
$$\mathrm{seg}(\mu, \nu) = \{x \in \mathbb{R}^2 \colon x = \lambda\mu + (1-\lambda)\nu,\ \lambda \in [0,1]\}$$

$$D(\mathrm{seg}(\mu,\nu), a) = \min_{t \in \mathrm{seg}[\mu,\nu]} \|t - a\|_2 = \begin{cases} \|a-\mu\|_2, & (\nu-\mu)^T(a-\mu) \le 0, \\ \|a-\nu\|_2, & (\mu-\nu)^T(a-\nu) < 0, \\ \left\| \mu \frac{(\nu-\mu)^T(\nu-a)}{\|\mu-\nu\|_2^2} + \left(1 - \frac{(\nu-\mu)^T(\nu-a)}{\|\mu-\nu\|_2^2}\right)\nu \right\|_2, & \text{else.} \end{cases}$$

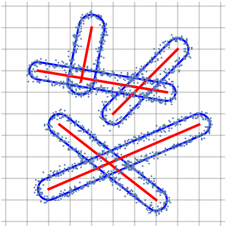


Prepoznavanje više generaliziranih kružnica u ravnini

Enterobacter-cloaca



Generirani podaci



$$\underset{\mu,\nu\in\Delta^k,\, r\in(0,R]^k}{\operatorname{argmin}} F(\mu,\nu,r), \quad F(\mu,\nu,r) = \sum_{i=1}^{m} \min_{1\le j\le k} \mathfrak{D}(\mathcal{C}_j(\mu_j,\nu_j,r_j), a^i),$$

$$\mathfrak{D}(\mathcal{C}(\mu,\nu,r),a) = (D(\mathrm{seg}(\mu,\nu),a)^2 - r^2)^2 \quad \text{(algebraic distance)}$$

Metda:

1. pronaći dobru početnu aproksimaciju (particiju ili $\mathcal{C}$-klaster centre);
2. pronaći optimalno rješenje primjenom k-means algoritma modificiranog za $\mathcal{C}$-klaster centre.

Modifikacija k-means algoritma za $\mathcal{C}$-klaster-centre

Algorithm 1

A: (Assignment step) For each set of mutually different generalized circles $\mathcal{C}_1(\mu_1, \nu_1, r_1)$, $\ldots$, $\mathcal{C}_k(\mu_k, \nu_k, r_k)$, the data set $\mathcal{A}$ should be partitioned into k disjoint nonempty clusters $\pi_1, \ldots, \pi_k$ by using the minimal distance principle for $j = 1, \ldots, k$

$$\pi_j := \{a \in \mathcal{A} \colon \mathfrak{D}(\mathcal{C}_j(\mu_j, \nu_j, r_j), a) \leq \mathfrak{D}(\mathcal{C}_s(\mu_s, \nu_s, r_s), a),\ s \neq j\};$$

B: (Update step) For the given partition $\Pi = \{\pi_1, \ldots, \pi_k\}$ of the set $\mathcal{A}$, we should determine the corresponding $\mathcal{C}$-cluster-centers $\hat{\mathcal{C}}_j(\hat{\mu}_j, \hat{\nu}_j, \hat{r}_j)$, $j = 1, \ldots, k$ by solving the following GOPs:

$$\underset{\mu,\nu \in \Delta,\ r \in [0,R]}{\operatorname{argmin}} \sum_{a \in \pi_j} \mathfrak{D}(\mathcal{C}(\mu, \nu, r), a), \quad j = 1, \ldots, k.$$

Theorem 2

The KMCC*-algorithm gives an accumulation point of the sequence (F_n) of the objective function values for the function F in finitely many steps. Moreover, the sequence (F_n) decreases monotonically.*

Kriterij zaustavljanja: $\frac{F_{k-1} - F_k}{F_1} < \epsilon_B, \quad \epsilon_B > 0$

Input: $\pi_j \in \Pi$;

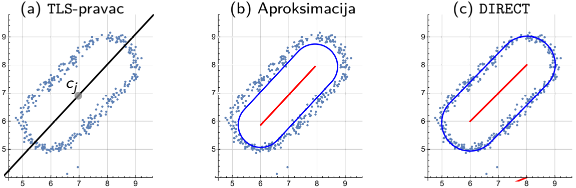
1. Odrediti centroid c_j i kovarijacijsku matricu i svojstvene vektore $v_1^{(j)}, v_2^{(j)}$;
2. Prema [1] procijeniti duljinu segmenta $\delta_j = \frac{2}{|\pi_j|} \sum\limits_{s\in\pi_j} |(a^i - c_j)\cdot v_1^{(j)}|$;
3. Odrediti aproksimaciju segmenta $\mathrm{seg}(\mu_j, \nu_j)$ i radijusa r_j kružnice:

$$\mathrm{seg}(\mu_j, \nu_j) \approx [c_j - \tfrac{1}{2}\delta_j v_0^{(j)}, c_j + \tfrac{1}{2}\delta_j v_0^{(j)}],$$

$$r_j \approx \frac{1}{|\pi_j|} \sum_{s\in\pi_j} |(a^i - c_j)\cdot v_2^{(j)}|.$$

4. Korištenjem `DIRECT` algoritma odrediti traženi $\mathcal{C}$-klaster-centar;

Output: $\{\hat{\mathcal{C}}_j\}$.



[1] J.C.R. Thomas, Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications, Springer, 2011

Traženje optimalne particije korištenjem povoljne početne particije

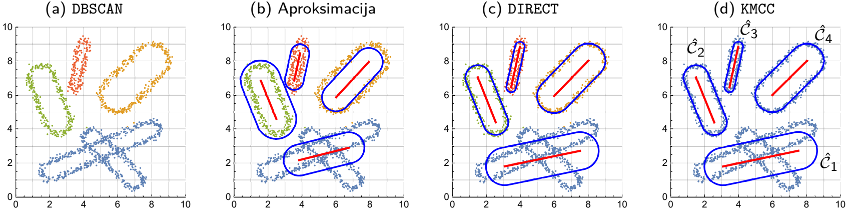
Input: $\mathcal{A} \subset \mathbb{R}^2$; $\quad MinPts = 3$;

1. Odrediti DBSCAN-parametar ϵ;
2. Korištenjem DBSCAN algoritma odrediti početnu particiju $\Pi^0 = \{\pi_1, \ldots, \pi_\kappa\}$;
3. Korištenjem KMCC algoritma uz početnu particiju Π^0 odrediti generalizirane kružnice $\mathcal{C}_1^\star, \ldots, \mathcal{C}_\kappa^\star$ i odgovarajuće vrijednosti QAD-indeksa

$$\texttt{QAD[j]} := \mu_j + 2\sigma_j, \quad \mu_j = \mathrm{Mean}\,(V_j), \quad \sigma_j^2 = \frac{1}{|V_j|-1} \sum_{v_j \in V_j} (v_j - \mu_j)^2$$

4. Odrediti QAD[j], $j = 1, \ldots, \kappa$;

Output: $\{(\mathcal{C}_1^\star, \texttt{QAD[1]}), \ldots, (\mathcal{C}_\kappa^\star, \texttt{QAD}[\kappa])\}$.



QAD=(.831, .139, .174, .155); DBSCAN-parametar $\epsilon = .203$

Traženje optimalne particije korištenjem povoljnih početnih $\mathcal{C}$-klaster-centara (u slučaju presjecajućih generaliziranih kružnica)

Input: $\mathcal{A} \subset \mathbb{R}^2$;

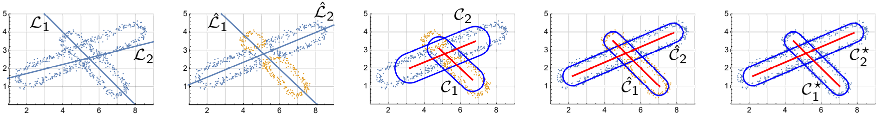
1. Odrediti najbolji TLS-pravac $\mathcal{C}_1^\star$ skupa $\mathcal{A}$;
2. Riješiti GOP: $\mathcal{L}_2 \in \underset{\substack{\alpha\in[-\pi/2,\pi/2],\\ \delta\in[0,\beta_2]}}{\operatorname{argmin}} \sum_{i=1}^{m} \min\{d_1^{(i)}, \mathfrak{D}_H(\ell(\alpha,\delta), a^i)\}$, gdje je

$$d_1^{(i)} = \frac{(u_1 x_i + v_1 y_i + z_1)^2}{u_1^2 + v_1^2}, \quad \mathfrak{D}_H(\ell(\alpha,\delta), a^i) = (x_i \cos\alpha + y_i \sin\alpha - \delta)^2, \quad \ell(\alpha,\delta) \equiv x\cos\alpha + y\sin\alpha - \delta = 0;$$

3. Korištenjem KMCL-algoritma odrediti $\hat{\mathcal{L}}_1, \hat{\mathcal{L}}_2$;
4. Korištenjem algoritma za traženje najboljeg $\mathcal{C}$-klaster-centra, odrediti $\hat{\mathcal{C}}_1, \hat{\mathcal{C}}_2$;
5. Korištenjem KMCC-algoritma uz početne $\mathcal{C}$-klaster-centre $\hat{\mathcal{C}}_1, \hat{\mathcal{C}}_2$ odrediti generalizirane kružnice $\mathcal{C}_1^\star, \mathcal{C}_2^\star$ i indekse QAD[j], $j = 1, 2$;

Output: $\{(\mathcal{C}_1^\star, \texttt{QAD[1]}), (\mathcal{C}_2^\star, \texttt{QAD[2]})\}$.

(1-2) Inkrementalni alg. (3) KMCL-algoritam (4) Aproksimacije $\mathcal{C}$-klaster-centara (5) KMCC-algoritam



QAD=(.16 .13); DBSCAN-parametar $\epsilon = .203$