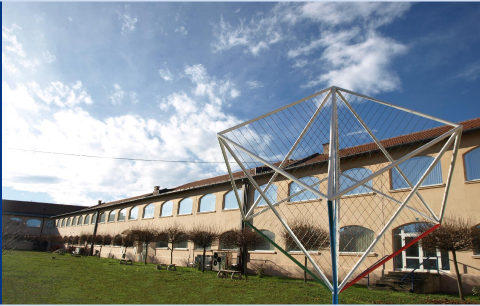
Incremental method for multiple line detection problem

Kristian Sabo, Rudolf Scitovski

UNIVERSITY J. J. STROSSMAYER OF OSIJEK
DEPARTMENT OF MATHEMATICS
Trg Ljudevita Gaja 6
31000 Osijek, Croatia
http://www.mathos.unios.hr

ksabo@mathos.hr, scitowsk@mathos.hr

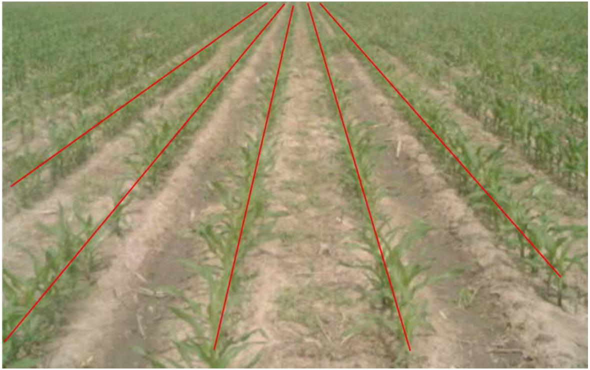


IP-2016-06-6545

[MOTOR 2020, 6TH-10TH JULY, 2020, NOVOSIBIRSK, RUSSIA]

Real world motivation example: recognizing maize rows

Real world motivation example: recognizing maize rows



Data points in the plane $\mathcal{A}$:

$$\mathcal{A} = \{a^i = (x_i, y_i) \colon i = 1, \ldots, m\} \subset \mathbb{R}^2$$

scattered along multiple lines, not known in advance.

MLD problem

Detect multiple lines on the basis of data points set $\mathcal{A}$

Data points in the plane $\mathcal{A}$:

$$\mathcal{A} = \{a^i = (x_i, y_i) \colon i = 1, \ldots, m\} \subset \mathbb{R}^2$$

scattered along multiple lines, not known in advance.

MLD problem

Detect multiple lines on the basis of data points set $\mathcal{A}$

- computer vision and image processing

  L. A. Fernandes, M. M. Oliveira, Real-time line detection through an improved Hough transform voting scheme, Pattern Recognition, 41(2008) 299–314

  A. Manzanera, T. P. Nguyen, X. Xu, Line and circle detection using dense one-to-one Hough transforms on greyscale images, EURASIP Journal on Image and Video Processing, (2016), DOI 10.1186/s13640-016-0149-y

- robotics, laser range measurements

  C. Fernández, V. Moreno, B. Curto, J. A. Vicente, Clustering and line detection in laser range measurements, Robotics and Autonomous Systems, 58(2010) 720–726

- civil engineering and geodesy

  A. Manzanera, T. P. Nguyen, X. Xu, Line and circle detection using dense one-to-one Hough transforms on greyscale images, EURASIP Journal on Image and Video Processing, (2016), DOI 10.1186/s13640-016-0149-y

- crop row detection in agriculture

  I. Vidović, R. Scitovski, Center-based clustering for line detection and application to crop rows detection, Computers and Electronics in Agriculture, 109(2014) 212–220

Methods for MLD problem

- Hough Transform (data without noise)
- Probabilistic Hough Transform and Randomized Hough Transform (data with noise)

L. A. Fernandes, M. M. Oliveira, Real-time line detection through an improved Hough transform voting scheme, Pattern Recognition, 41(2008) 299–314

A. Manzanera, T. P. Nguyen, X. Xu, Line and circle detection using dense one-to-one Hough transforms on greyscale images, EURASIP Journal on Image and Video Processing, (2016), DOI 10.1186/s13640-016-0149-y

P. Mukhopadhyay, B. B. Chaudhuri, A survey of Hough transform, Pattern Recognition, 48(2015) 993–1010

MLD problem as a clustering problem

Data points in the plane $\mathcal{A}$:

$$\mathcal{A} = \{a^i = (x_i, y_i) \colon i = 1, \ldots, m\} \subset \mathbb{R}^2$$

scattered along multiple lines, not known in advance.

Set of lines in the plane $\mathcal{L}$:

$$\mathcal{L} = \{\ell(\xi, \eta, \zeta) \equiv \xi x + \eta y + \zeta = 0,\ [\xi, \eta, \zeta]^T \in \mathcal{P}\}$$

Set of parameter-vectors $\mathcal{P}$:

$$\mathcal{P} = \{\mathbf{p} = [\xi, \eta, \zeta]^T \in \mathbb{R}^3 \colon \xi^2 + \eta^2 = 1\}$$

MLD problem as a clustering problem

Data points in the plane $\mathcal{A}$:

$$\mathcal{A} = \{a^i = (x_i, y_i) \colon i = 1, \ldots, m\} \subset \mathbb{R}^2$$

scattered along multiple lines, not known in advance.

Set of lines in the plane $\mathcal{L}$:

$$\mathcal{L} = \{\ell(\xi, \eta, \zeta) \equiv \xi x + \eta y + \zeta = 0,\ [\xi, \eta, \zeta]^T \in \mathcal{P}\}$$

Set of parameter-vectors $\mathcal{P}$:

$$\mathcal{P} = \{\mathbf{p} = [\xi, \eta, \zeta]^T \in \mathbb{R}^3 \colon \xi^2 + \eta^2 = 1\}$$

MLD problem as a clustering problem

Data points in the plane $\mathcal{A}$:

$$\mathcal{A} = \{a^i = (x_i, y_i) \colon i = 1, \ldots, m\} \subset \mathbb{R}^2$$

scattered along multiple lines, not known in advance.

Set of lines in the plane $\mathcal{L}$:

$$\mathcal{L} = \{\ell(\xi, \eta, \zeta) \equiv \xi x + \eta y + \zeta = 0, \, [\xi, \eta, \zeta]^T \in \mathcal{P}\}$$

Set of parameter-vectors $\mathcal{P}$:

$$\mathcal{P} = \{\mathbf{p} = [\xi, \eta, \zeta]^T \in \mathbb{R}^3 \colon \xi^2 + \eta^2 = 1\}$$

Data points in the plane $\mathcal{A}$:

$$\mathcal{A} = \{a^i = (x_i, y_i)\colon i = 1, \ldots, m\} \subset \mathbb{R}^2$$

scattered along multiple lines, not known in advance.

Set of lines in the plane $\mathcal{L}$:

$$\mathcal{L} = \{\ell(\xi, \eta, \zeta) \equiv \xi x + \eta y + \zeta = 0,\ [\xi, \eta, \zeta]^T \in \mathcal{P}\}$$

Set of parameter-vectors $\mathcal{P}$:

$$\mathcal{P} = \{\mathbf{p} = [\xi, \eta, \zeta]^T \in \mathbb{R}^3\colon \xi^2 + \eta^2 = 1\}$$

Euclidean distance from the point $a^i = (x_i, y_i) \in \mathcal{A}$ to the line $\ell(\mathbf{p}) \in \mathcal{L}$, $\mathbf{p} = [\xi, \eta, \zeta]^T \in \mathcal{P}$:

$$\mathfrak{D}(a^i, \ell(\mathbf{p})) = (\xi\, x_i + \eta\, y_i + \zeta)^2$$

Globally Optimal k-partition:

$$\operatorname*{argmin}_{\Pi \in Part(\mathcal{A};k)} \mathcal{F}(\Pi), \qquad \mathcal{F}(\Pi) = \sum_{j=1}^{k} \sum_{a^i \in \pi_j} \mathfrak{D}(a^i, \ell_j(\mathbf{p}_j)),$$

where $Part(\mathcal{A}; k)$ is the set of all k-partitions of the set $\mathcal{A}$ and

$$\mathbf{p}_j \in \operatorname*{argmin}_{\mathbf{p} \in \mathcal{P}} \sum_{a^i \in \pi_j} \mathfrak{D}(a^i, \ell(\mathbf{p})).$$

Euclidean distance from the point $a^i = (x_i, y_i) \in \mathcal{A}$ to the line $\ell(\mathbf{p}) \in \mathcal{L}$, $\mathbf{p} = [\xi, \eta, \zeta]^T \in \mathcal{P}$:

$$\mathfrak{D}(a^i, \ell(\mathbf{p})) = (\xi\, x_i + \eta\, y_i + \zeta)^2$$

Globally Optimal k-partition:

$$\underset{\Pi \in Part(\mathcal{A};k)}{\operatorname{argmin}} \mathcal{F}(\Pi), \qquad \mathcal{F}(\Pi) = \sum_{j=1}^{k} \sum_{a^i \in \pi_j} \mathfrak{D}(a^i, \ell_j(\mathbf{p}_j)),$$

where $Part(\mathcal{A}; k)$ is the set of all k-partitions of the set $\mathcal{A}$ and

$$\mathbf{p}_j \in \underset{\mathbf{p} \in \mathcal{P}}{\operatorname{argmin}} \sum_{a^i \in \pi_j} \mathfrak{D}(a^i, \ell(\mathbf{p})).$$

Globally Optimal k-partition:

$$\operatorname*{argmin}_{\mathbf{p}_j \in \mathcal{P}} F(\mathbf{p}_1 \ldots, \mathbf{p}_k), \quad F(\mathbf{p}_1 \ldots, \mathbf{p}_k) = \sum_{i=1}^{m} \min_{1 \le j \le k} \mathfrak{D}(a^i, \ell_j(\mathbf{p}_j)),$$

- F is nonconvex and nondifferentiable
- F is Lipschitz-continuous

Globally Optimal k-partition:

$$\operatorname*{argmin}_{\mathbf{p}_j \in \mathcal{P}} F(\mathbf{p}_1 \ldots, \mathbf{p}_k), \quad F(\mathbf{p}_1 \ldots, \mathbf{p}_k) = \sum_{i=1}^{m} \min_{1 \le j \le k} \mathfrak{D}(a^i, \ell_j(\mathbf{p}_j)),$$

- F is nonconvex and nondifferentiable
- F is Lipschitz-continuous

Globally Optimal k-partition:

$$\underset{\mathbf{p}_j \in \mathcal{P}}{\operatorname{argmin}} F(\mathbf{p}_1 \ldots, \mathbf{p}_k), \quad F(\mathbf{p}_1 \ldots, \mathbf{p}_k) = \sum_{i=1}^{m} \min_{1 \le j \le k} \mathfrak{D}(a^i, \ell_j(\mathbf{p}_j)),$$

- F is nonconvex and nondifferentiable
- F is Lipschitz-continuous

Special case $k = 1$

Total Least Squares (TLS) line

- $w_i > 0$ be the corresponding weights of the data points $a^i \in \mathcal{A}$.
- line $\tilde{\ell} \in \mathcal{L}$, with parameter-vector $[\tilde{\xi}, \tilde{\eta}, \tilde{\zeta}]^T \in \mathcal{P}$, passes through the centroid $(\bar{x}, \bar{y})$ of the weighted data point set $(w, \mathcal{A})$, it can be written in the form:

$$\tilde{\ell} \equiv \tilde{\xi}(x - \bar{x}) + \tilde{\eta}(y - \bar{y}) = 0.$$

$$\underset{[\xi,\eta]^T \in \mathbf{R}^2,\, \xi^2+\eta^2=1}{\operatorname{argmin}} G(\xi, \eta)$$

$$G(\xi, \eta) = \sum_{i=1}^{m} w_i(\xi(x_i - \overline{x}) + \eta(y_i - \overline{y}))^2$$

- The problem attains a unique global minimum if and only if at least one of the following two conditions is fulfilled:
  i) $\sum_{i=1}^{m} w_i(x_i - \overline{x})^2 \neq \sum_{i=1}^{m} w_i(y_i - \overline{y})^2$ and
  ii) $\sum_{i=1}^{m} w_i(x_i - \overline{x})(y_i - \overline{y}) \neq 0.$

Special case $k = 1$

Total Least Squares (TLS) line

- $w_i > 0$ be the corresponding weights of the data points $a^i \in \mathcal{A}$.
- line $\tilde{\ell} \in \mathcal{L}$, with parameter-vector $[\tilde{\xi}, \tilde{\eta}, \tilde{\zeta}]^T \in \mathcal{P}$, passes through the centroid $(\bar{x}, \bar{y})$ of the weighted data point set $(w, \mathcal{A})$, it can be written in the form:

$$\tilde{\ell} \equiv \tilde{\xi}(x - \bar{x}) + \tilde{\eta}(y - \bar{y}) = 0.$$

$$\underset{[\xi,\eta]^T \in \mathbf{R}^2,\, \xi^2+\eta^2=1}{\operatorname{argmin}} G(\xi, \eta)$$

$$G(\xi, \eta) = \sum_{i=1}^{m} w_i(\xi(x_i - \overline{x}) + \eta(y_i - \overline{y}))^2$$

- The problem attains a unique global minimum if and only if at least one of the following two conditions is fulfilled:
  - i) $\sum_{i=1}^{m} w_i(x_i - \overline{x})^2 \neq \sum_{i=1}^{m} w_i(y_i - \overline{y})^2$ and
  - ii) $\sum_{i=1}^{m} w_i(x_i - \overline{x})(y_i - \overline{y}) \neq 0$.

Special case $k = 1$

Total Least Squares (TLS) line

- $w_i > 0$ be the corresponding weights of the data points $a^i \in \mathcal{A}$.
- line $\tilde{\ell} \in \mathcal{L}$, with parameter-vector $[\tilde{\xi}, \tilde{\eta}, \tilde{\zeta}]^T \in \mathcal{P}$, passes through the centroid $(\bar{x}, \bar{y})$ of the weighted data point set $(w, \mathcal{A})$, it can be written in the form:

$$\tilde{\ell} \equiv \tilde{\xi}(x - \bar{x}) + \tilde{\eta}(y - \bar{y}) = 0.$$

$$\underset{[\xi,\eta]^T \in \mathbf{R}^2,\, \xi^2+\eta^2=1}{\operatorname{argmin}} G(\xi, \eta)$$

$$G(\xi, \eta) = \sum_{i=1}^{m} w_i(\xi(x_i - \overline{x}) + \eta(y_i - \overline{y}))^2$$

- The problem attains a unique global minimum if and only if at least one of the following two conditions is fulfilled:
  - i) $\sum_{i=1}^{m} w_i(x_i - \overline{x})^2 \neq \sum_{i=1}^{m} w_i(y_i - \overline{y})^2$ and
  - ii) $\sum_{i=1}^{m} w_i(x_i - \overline{x})(y_i - \overline{y}) \neq 0.$

$$\mathbf{B} := \begin{bmatrix} x_1 - \overline{x} & y_1 - \overline{y} \\ \vdots & \vdots \\ x_m - \overline{x} & y_m - \overline{y} \end{bmatrix}, \quad \mathbf{D} := \mathsf{diag}(w_1 \ldots, w_m), \quad \mathbf{t} = [\xi, \eta]^T,$$

$$G(\xi, \eta) = \|\sqrt{\mathbf{D}}\mathbf{B}\mathbf{t}\|^2, \quad \|\mathbf{t}\| = 1$$

its global minimum is attained at every unit eigenvector $\hat{\mathbf{t}} = [\hat{\xi}, \hat{\eta}]^T$ corresponding to the smaller eigenvalue of the matrix $\mathbf{B}^T\mathbf{D}\mathbf{B}$.
$\hat{\ell}(\hat{\xi}, \hat{\eta}, \hat{\zeta}) \equiv \hat{\xi}(x - \overline{x}) + \hat{\eta}(y - \overline{y}) = 0$, $\hat{\xi}^2 + \hat{\eta}^2 = 1$, where $\hat{\zeta} = -\hat{\xi}\overline{x} - \hat{\eta}\overline{y}$.

$$\mathbf{B} := \begin{bmatrix} x_1 - \overline{x} & y_1 - \overline{y} \\ \vdots & \vdots \\ x_m - \overline{x} & y_m - \overline{y} \end{bmatrix}, \quad \mathbf{D} := \mathsf{diag}(w_1 \ldots, w_m), \quad \mathbf{t} = [\xi, \eta]^T,$$

$$G(\xi, \eta) = \|\sqrt{\mathbf{D}}\mathbf{B}\mathbf{t}\|^2, \quad \|\mathbf{t}\| = 1$$

its global minimum is attained at every unit eigenvector $\hat{\mathbf{t}} = [\hat{\xi}, \hat{\eta}]^T$ corresponding to the smaller eigenvalue of the matrix $\mathbf{B}^T\mathbf{D}\mathbf{B}$.
$\hat{\ell}(\hat{\xi}, \hat{\eta}, \hat{\zeta}) \equiv \hat{\xi}(x - \overline{x}) + \hat{\eta}(y - \overline{y}) = 0$, $\hat{\xi}^2 + \hat{\eta}^2 = 1$, where $\hat{\zeta} = -\hat{\xi}\overline{x} - \hat{\eta}\overline{y}$.

$$\mathbf{B} := \begin{bmatrix} x_1 - \overline{x} & y_1 - \overline{y} \\ \vdots & \vdots \\ x_m - \overline{x} & y_m - \overline{y} \end{bmatrix}, \quad \mathbf{D} := \mathsf{diag}(w_1 \ldots, w_m), \quad \mathbf{t} = [\xi, \eta]^T,$$

$$G(\xi, \eta) = \|\sqrt{\mathbf{D}}\mathbf{B}\mathbf{t}\|^2, \quad \|\mathbf{t}\| = 1$$

its global minimum is attained at every unit eigenvector $\hat{\mathbf{t}} = [\hat{\xi}, \hat{\eta}]^T$ corresponding to the smaller eigenvalue of the matrix $\mathbf{B}^T\mathbf{D}\mathbf{B}$.

$\hat{\ell}(\hat{\xi}, \hat{\eta}, \hat{\zeta}) \equiv \hat{\xi}(x - \overline{x}) + \hat{\eta}(y - \overline{y}) = 0$, $\hat{\xi}^2 + \hat{\eta}^2 = 1$, where $\hat{\zeta} = -\hat{\xi}\overline{x} - \hat{\eta}\overline{y}$.

$$\mathbf{B} := \begin{bmatrix} x_1 - \overline{x} & y_1 - \overline{y} \\ \vdots & \vdots \\ x_m - \overline{x} & y_m - \overline{y} \end{bmatrix}, \quad \mathbf{D} := \mathrm{diag}(w_1 \ldots, w_m), \quad \mathbf{t} = [\xi, \eta]^T,$$

$$G(\xi, \eta) = \|\sqrt{\mathbf{D}}\mathbf{B}\mathbf{t}\|^2, \quad \|\mathbf{t}\| = 1$$

its global minimum is attained at every unit eigenvector $\hat{\mathbf{t}} = [\hat{\xi}, \hat{\eta}]^T$ corresponding to the smaller eigenvalue of the matrix $\mathbf{B}^T\mathbf{D}\mathbf{B}$.
$\hat{\ell}(\hat{\xi}, \hat{\eta}, \hat{\zeta}) \equiv \hat{\xi}(x - \overline{x}) + \hat{\eta}(y - \overline{y}) = 0$, $\hat{\xi}^2 + \hat{\eta}^2 = 1$, where $\hat{\zeta} = -\hat{\xi}\overline{x} - \hat{\eta}\overline{y}$.

The k-closest line algorithm (KCL)

Algorithm

Step A: *(Assignment step) For each set of mutually different lines $\ell_1, \ldots, \ell_k \in \mathcal{L}$, the set $\mathcal{A}$ should be divided into k disjoint unempty clusters $\pi_1, \ldots, \pi_k$ by using the minimal distance principle*

$$\pi_j := \{a \in \mathcal{A} \colon \mathfrak{D}(a, \ell_j) \le \mathfrak{D}(a, \ell_s), \, \forall s \ne j\};$$

Step B: *(Update step) Given a partition $\Pi = \{\pi_1, \ldots, \pi_k\}$ of the set $\mathcal{A}$, one can define the corresponding line cluster-centers $\hat{\ell}_1, \ldots, \hat{\ell}_k \in \mathcal{L}$ as corresponding* TLS*-lines. Set $\ell_j = \hat{\ell}_j$, $j = 1, \ldots, k$.*

KCL algorithm - usually initiated several times with different random initial lines and for the solution we will take the one yielding the smallest value of the function F.

The k-closest line algorithm (KCL)

Algorithm

Step A: *(Assignment step) For each set of mutually different lines $\ell_1, \ldots, \ell_k \in \mathcal{L}$, the set $\mathcal{A}$ should be divided into k disjoint unempty clusters $\pi_1, \ldots, \pi_k$ by using the minimal distance principle*

$$\pi_j := \{a \in \mathcal{A} \colon \mathfrak{D}(a, \ell_j) \leq \mathfrak{D}(a, \ell_s), \, \forall s \neq j\};$$

Step B: *(Update step) Given a partition $\Pi = \{\pi_1, \ldots, \pi_k\}$ of the set $\mathcal{A}$, one can define the corresponding line cluster-centers $\hat{\ell}_1, \ldots, \hat{\ell}_k \in \mathcal{L}$ as corresponding* TLS*-lines. Set $\ell_j = \hat{\ell}_j$, $j = 1, \ldots, k$.*

KCL algorithm - usually initiated several times with different random initial lines and for the solution we will take the one yielding the smallest value of the function F.

Incremental algorithm for multiple line detection

$\tilde{\ell}_1, \ldots, \tilde{\ell}_{k-1}$ are known lines, the next line $\tilde{\ell}_k$ will be obtained by solving the following GOP:

$$\underset{\mathbf{p}\in\mathcal{P}}{\operatorname{argmin}}\, \Phi_k(\mathbf{p}), \quad \Phi_k(\mathbf{p}) = \sum_{i=1}^{m} \min\{\delta_{k-1}^{(i)}, \mathfrak{D}(a^i, \ell(\mathbf{p}))\}, \qquad (1)$$

where

$$\delta_{k-1}^{(i)} = \min\{\mathfrak{D}(a^i, \tilde{\ell}_1), \ldots, \mathfrak{D}(a^i, \tilde{\ell}_{k-1})\}. \qquad (2)$$

After that, KCL algorithm is applied to the set of lines $\{\tilde{\ell}_1, \ldots, \tilde{\ell}_k\}$. Since by increasing the number of clusters in partition, the objective value decreases, it is reasonable to stop the iterative process when

$$\frac{F_k - F_{k-1}}{F_1} < \epsilon_B,$$

for some small $\epsilon_B > 0$ (say .005).

A. M. Bagirov, J. Ugon, H. Mirzayeva, Nonsmooth nonconvex optimization approach to clusterwise linear regression problems, European Journal of Operational Research, 229(2013) 132–142.

Incremental algorithm for multiple line detection

$\tilde{\ell}_1, \ldots, \tilde{\ell}_{k-1}$ are known lines, the next line $\tilde{\ell}_k$ will be obtained by solving the following GOP:

$$\operatorname*{argmin}_{\mathbf{p}\in\mathcal{P}} \Phi_k(\mathbf{p}), \quad \Phi_k(\mathbf{p}) = \sum_{i=1}^{m} \min\{\delta_{k-1}^{(i)}, \mathfrak{D}(a^i, \ell(\mathbf{p}))\}, \tag{1}$$

where

$$\delta_{k-1}^{(i)} = \min\{\mathfrak{D}(a^i, \tilde{\ell}_1), \ldots, \mathfrak{D}(a^i, \tilde{\ell}_{k-1})\}. \tag{2}$$

After that, KCL algorithm is applied to the set of lines $\{\tilde{\ell}_1, \ldots, \tilde{\ell}_k\}$. Since by increasing the number of clusters in partition, the objective value decreases, it is reasonable to stop the iterative process when

$$\frac{F_k - F_{k-1}}{F_1} < \epsilon_B,$$

for some small $\epsilon_B > 0$ (say .005).

A. M. Bagirov, J. Ugon, H. Mirzayeva, Nonsmooth nonconvex optimization approach to clusterwise linear regression problems, European Journal of Operational Research, 229(2013) 132–142.

Incremental algorithm for multiple line detection

$\tilde{\ell}_1, \ldots, \tilde{\ell}_{k-1}$ are known lines, the next line $\tilde{\ell}_k$ will be obtained by solving the following GOP:

$$\operatorname*{argmin}_{\mathbf{p}\in\mathcal{P}} \Phi_k(\mathbf{p}), \quad \Phi_k(\mathbf{p}) = \sum_{i=1}^{m} \min\{\delta_{k-1}^{(i)}, \mathfrak{D}(a^i, \ell(\mathbf{p}))\}, \qquad (1)$$

where

$$\delta_{k-1}^{(i)} = \min\{\mathfrak{D}(a^i, \tilde{\ell}_1), \ldots, \mathfrak{D}(a^i, \tilde{\ell}_{k-1})\}. \qquad (2)$$

After that, KCL algorithm is applied to the set of lines $\{\tilde{\ell}_1, \ldots, \tilde{\ell}_k\}$. Since by increasing the number of clusters in partition, the objective value decreases, it is reasonable to stop the iterative process when

$$\frac{F_k - F_{k-1}}{F_1} < \epsilon_B,$$

for some small $\epsilon_B > 0$ (say .005).

A. M. Bagirov, J. Ugon, H. Mirzayeva, Nonsmooth nonconvex optimization approach to clusterwise linear regression problems, European Journal of Operational Research, 229(2013) 132–142.

New smoothing method

Instead of solving GOP:

$$\operatorname*{argmin}_{\mathbf{p}\in\mathcal{P}} \Phi_k(\mathbf{p}), \quad \Phi_k(\mathbf{p}) = \sum_{i=1}^{m} \min\{\delta_{k-1}^{(i)}, \mathfrak{D}(a^i, \ell(\mathbf{p}))\},$$

we approximate Φ_k with differentiable function and we solve local optimization problem.

By using

$$|x| = \lim_{\epsilon \to 0^+} \epsilon \log(2\cosh\tfrac{x}{\epsilon}),$$

$$0 \leq \epsilon \log(2\cosh\frac{x}{\epsilon}) - |x| \leq \epsilon \log 2,$$

and the identity

$$\min\{x,y\} = \frac{1}{2}\left(x + y - |x-y|\right),$$

we will determine a smooth approximation Φ_k^ϵ of the function Φ_k given in (3):

$$\Phi_k(\mathbf{p}) = \frac{1}{2}\sum_{i=1}^{m}\left(\delta_{k-1}^{(i)} + \mathfrak{D}(a^i, \ell(\mathbf{p})) - |\mathfrak{D}(a^i, \ell(\mathbf{p})) - \delta_{k-1}^{(i)}|\right)$$

$$\approx \tfrac{1}{2}\sum_{i=1}^{m}\left(\delta_{k-1}^{(i)} + \mathfrak{D}(a^i, \ell(\mathbf{p})) - \varepsilon \log\left(2\cosh\frac{\mathfrak{D}(a^i, \ell(\mathbf{p})) - \delta_{k-1}^{(i)}}{\varepsilon}\right)\right)$$

$$=: \Phi_k^\epsilon(\mathbf{p}).$$

Main idea

By using such smoothing of the function Φ_k, we propose a simple, efficient local optimization method for solving `MLD` problem and prove its convergence.

Lemma

Let $\epsilon > 0$ and suppose that $\Phi_k : \mathbb{R}^3 \to \mathbb{R}$ and $\Phi_k^\epsilon : \mathbb{R}^3 \to \mathbb{R}$. Then:

(i) $0 \leq \Phi_k(\mathbf{p}) - \Phi_k^\epsilon(\mathbf{p}) \leq \frac{m\epsilon}{2} \log 2$

(ii) $\Phi_k^\epsilon(\mathbf{p}) \geq -\frac{m\epsilon}{2} \log 2$, *i.e. function Φ_k^ϵ is bounded below.*

Main idea

By using such smoothing of the function Φ_k, we propose a simple, efficient local optimization method for solving MLD problem and prove its convergence.

Lemma

Let $\epsilon > 0$ *and suppose that* $\Phi_k : \mathbb{R}^3 \to \mathbb{R}$ *and* $\Phi_k^\epsilon : \mathbb{R}^3 \to \mathbb{R}$. *Then:*

(i) $0 \leq \Phi_k(\mathbf{p}) - \Phi_k^\epsilon(\mathbf{p}) \leq \frac{m\epsilon}{2}\log 2$

(ii) $\Phi_k^\epsilon(\mathbf{p}) \geq -\frac{m\epsilon}{2}\log 2$, *i.e. function* Φ_k^ϵ *is bounded below.*

Iterative procedure

Let $\mathbf{p}^{(0)} = [\xi^{(0)}, \eta^{(0)}, \zeta^{(0)}]^T \in \mathcal{P}$, and $\mathbf{s}^{(0)} = [\xi^{(0)}, \eta^{(0)}]^T$. Denoting $\tilde{\mathbf{a}}^i := [x_i, y_i, 1]^T$.

$$\mathbf{p}^{(n+1)} = \begin{bmatrix} \xi^{(n+1)} \\ \eta^{(n+1)} \\ \zeta^{(n+1)} \end{bmatrix} = \begin{bmatrix} \mathbf{s}^{(n+1)} \\ \zeta^{(n+1)} \end{bmatrix} = \begin{bmatrix} \operatorname*{argmin}_{\mathbf{s}, \|\mathbf{s}\|=1} \|\sqrt{\mathbf{D}^{(n)}}\mathbf{B}^{(n)}\mathbf{s}\|^2 \\ -\xi^{(n+1)}\overline{x}^{(n)} - \eta^{(n+1)}\overline{y}^{(n)} \end{bmatrix},$$

where

$$\mathbf{D}^{(n)} = \text{Diag}\left(w_1^\epsilon(\mathbf{p}^{(n)}), \ldots, w_m^\epsilon(\mathbf{p}^{(n)})\right)$$

and

$$\mathbf{B}^{(n)} = \begin{bmatrix} x_1 - \overline{x}^{(n)} & y_1 - \overline{y}^{(n)} \\ \vdots & \vdots \\ x_m - \overline{x}^{(n)} & y_m - \overline{y}^{(n)} \end{bmatrix},$$

$$w_i^\epsilon(\mathbf{p}) = \frac{1}{2}\left(1 - \tanh \frac{\left(\mathbf{p}^T\tilde{\mathbf{a}}^i\right)^2 - \delta_{k-1}^{(i)}}{\varepsilon}\right), \quad i = 1, \ldots, m,$$

Iterative procedure

Let $\mathbf{p}^{(0)} = [\xi^{(0)}, \eta^{(0)}, \zeta^{(0)}]^T \in \mathcal{P}$, and $\mathbf{s}^{(0)} = [\xi^{(0)}, \eta^{(0)}]^T$. Denoting $\tilde{\mathbf{a}}^i := [x_i, y_i, 1]^T$.

$$\mathbf{p}^{(n+1)} = \begin{bmatrix} \xi^{(n+1)} \\ \eta^{(n+1)} \\ \zeta^{(n+1)} \end{bmatrix} = \begin{bmatrix} \mathbf{s}^{(n+1)} \\ \zeta^{(n+1)} \end{bmatrix} = \begin{bmatrix} \underset{\mathbf{s}, \|\mathbf{s}\|=1}{\operatorname{argmin}} \|\sqrt{\mathbf{D}^{(n)}}\mathbf{B}^{(n)}\mathbf{s}\|^2 \\ -\xi^{(n+1)}\overline{x}^{(n)} - \eta^{(n+1)}\overline{y}^{(n)} \end{bmatrix},$$

where

$$\mathbf{D}^{(n)} = \text{Diag}\left(w_1^\epsilon(\mathbf{p}^{(n)}), \ldots, w_m^\epsilon(\mathbf{p}^{(n)})\right)$$

and

$$\mathbf{B}^{(n)} = \begin{bmatrix} x_1 - \overline{x}^{(n)} & y_1 - \overline{y}^{(n)} \\ \vdots & \vdots \\ x_m - \overline{x}^{(n)} & y_m - \overline{y}^{(n)} \end{bmatrix},$$

$$w_i^\epsilon(\mathbf{p}) = \frac{1}{2}\left(1 - \tanh \frac{\left(\mathbf{p}^T\tilde{\mathbf{a}}^i\right)^2 - \delta_{k-1}^{(i)}}{\varepsilon}\right), \quad i = 1, \ldots, m,$$

$$\overline{x}^{(n)} = \frac{\sum_{i=1}^m w_i^\epsilon(\mathbf{p}^{(n)})x_i}{\sum_{i=1}^m w_i^\epsilon(\mathbf{p}^{(n)})}, \quad \overline{y}^{(n)} = \frac{\sum_{i=1}^m w_i^\epsilon(\mathbf{p}^{(n)})y_i}{\sum_{i=1}^m w_i^\epsilon(\mathbf{p}^{(n)})}.$$

In each iteration of the optimization process, the solution $\mathbf{s}^{(n+1)} = [\xi^{(n+1)}, \eta^{(n+1)}]^T$ is equal to the eigenvector corresponding to the smaller eigenvalue of the matrix $\left(\mathbf{B}^{(n)}\right)^T \mathbf{D}^{(n)}\mathbf{B}^{(n)}$.

$$\overline{x}^{(n)} = \frac{\sum_{i=1}^m w_i^\epsilon(\mathbf{p}^{(n)}) x_i}{\sum_{i=1}^m w_i^\epsilon(\mathbf{p}^{(n)})}, \quad \overline{y}^{(n)} = \frac{\sum_{i=1}^m w_i^\epsilon(\mathbf{p}^{(n)}) y_i}{\sum_{i=1}^m w_i^\epsilon(\mathbf{p}^{(n)})}.$$

In each iteration of the optimization process, the solution $\mathbf{s}^{(n+1)} = [\xi^{(n+1)}, \eta^{(n+1)}]^T$ is equal to the eigenvector corresponding to the smaller eigenvalue of the matrix $\left(\mathbf{B}^{(n)}\right)^T \mathbf{D}^{(n)} \mathbf{B}^{(n)}$.

Converegence of iterative procedure

Proposition

Let $\mathbf{p}^{(0)} = [\xi^{(0)}, \eta^{(0)}, \zeta^{(0)}]^T \in \mathcal{P}$ *and* $\mathbf{s}^{(0)} = [\xi^{(0)}, \eta^{(0)}]^T$*, and let the sequence* $\left(\mathbf{p}^{(n)}\right)$ *be given by the previous iterative process.*
If $\mathbf{p}^{(n+1)} \neq \mathbf{p}^{(n)}$*, then* $\Phi_k^{\epsilon}(\mathbf{p}^{(n+1)}) < \Phi_k^{\epsilon}(\mathbf{p}^{(n)})$

Theorem

Let $\mathbf{p}^{(0)} = [\xi^{(0)}, \eta^{(0)}, \zeta^{(0)}]^T \in \mathcal{P}$*, and* $\mathbf{s}^{(0)} = [\xi^{(0)}, \eta^{(0)}]^T$*, and let the sequence* $\left(\mathbf{p}^{(n)}\right)$ *be given by the previous iterative process. Then*

(i) *The sequence* $\left(\mathbf{p}^{(n)}\right)$ *has an accumulation point.*

(ii) *The sequence* $\left((\Phi_k^{\epsilon})^{(n)}\right)$*, where* $(\Phi_k^{\epsilon})^{(n)} := \Phi_k^{\epsilon}(\mathbf{p}^{(n)})$*, converges.*

(iii) *Every accumulation point* $\hat{\mathbf{p}}$ *of the sequence* $\left(\mathbf{p}^{(n)}\right)$ *is a stationary point of the functional* Φ_k^{ϵ}*.*

(iv) *If* $\hat{\mathbf{p}}_1$ *and* $\hat{\mathbf{p}}_2$ *are two accumulation points of the sequence* $\left(\mathbf{p}^{(n)}\right)$*, then* $\Phi_k^{\epsilon}(\hat{\mathbf{p}}_1) = \Phi_k^{\epsilon}(\hat{\mathbf{p}}_2)$*.*

Converegence of iterative procedure

Proposition

Let $\mathbf{p}^{(0)} = [\xi^{(0)}, \eta^{(0)}, \zeta^{(0)}]^T \in \mathcal{P}$ *and* $\mathbf{s}^{(0)} = [\xi^{(0)}, \eta^{(0)}]^T$*, and let the sequence* $\left(\mathbf{p}^{(n)}\right)$ *be given by the previous iterative process.*
If $\mathbf{p}^{(n+1)} \neq \mathbf{p}^{(n)}$*, then* $\Phi_k^\epsilon(\mathbf{p}^{(n+1)}) < \Phi_k^\epsilon(\mathbf{p}^{(n)})$

Theorem

Let $\mathbf{p}^{(0)} = [\xi^{(0)}, \eta^{(0)}, \zeta^{(0)}]^T \in \mathcal{P}$*, and* $\mathbf{s}^{(0)} = [\xi^{(0)}, \eta^{(0)}]^T$*, and let the sequence* $\left(\mathbf{p}^{(n)}\right)$ *be given by the previous iterative process. Then*

(i) *The sequence* $\left(\mathbf{p}^{(n)}\right)$ *has an accumulation point.*

(ii) *The sequence* $\left((\Phi_k^\epsilon)^{(n)}\right)$*, where* $(\Phi_k^\epsilon)^{(n)} := \Phi_k^\epsilon(\mathbf{p}^{(n)})$*, converges.*

(iii) *Every accumulation point* $\hat{\mathbf{p}}$ *of the sequence* $\left(\mathbf{p}^{(n)}\right)$ *is a stationary point of the functional* Φ_k^ϵ*.*

(iv) *If* $\hat{\mathbf{p}}_1$ *and* $\hat{\mathbf{p}}_2$ *are two accumulation points of the sequence* $\left(\mathbf{p}^{(n)}\right)$*, then* $\Phi_k^\epsilon(\hat{\mathbf{p}}_1) = \Phi_k^\epsilon(\hat{\mathbf{p}}_2)$*.*

usually initiated several times with different random initial approximation and for the solution we will take the one yielding the smallest value of the function Φ_k^{ϵ}.

Smoothed incremental algorithm for multiple line detection

$\tilde{\ell}_1, \ldots, \tilde{\ell}_{k-1}$ are known lines, the next line $\tilde{\ell}_k$ will be obtained by solving the following GOP:

$$\operatorname*{argmin}_{\mathbf{p}\in\mathcal{P}} \Phi_k(\mathbf{p}), \quad \Phi_k^{\epsilon}(\mathbf{p}) = \sum_{i=1}^{m} \min\{\delta_{k-1}^{(i)}, \mathfrak{D}(a^i, \ell(\mathbf{p}))\}, \qquad (3)$$

where

$$\delta_{k-1}^{(i)} = \min\{\mathfrak{D}(a^i, \tilde{\ell}_1), \ldots, \mathfrak{D}(a^i, \tilde{\ell}_{k-1})\}. \qquad (4)$$

After that, KCL algorithm is applied to the set of lines $\{\tilde{\ell}_1, \ldots, \tilde{\ell}_k\}$. Since by increasing the number of clusters in partition, the objective value decreases, it is reasonable to stop the incremental iterative process when

$$\frac{F_k - F_{k-1}}{F_1} < \epsilon_B,$$

for some small $\epsilon_B > 0$ (say .005).

Smoothed incremental algorithm for multiple line detection

$\tilde{\ell}_1, \ldots, \tilde{\ell}_{k-1}$ are known lines, the next line $\tilde{\ell}_k$ will be obtained by solving the following GOP:

$$\operatorname*{argmin}_{\mathbf{p}\in\mathcal{P}} \Phi_k(\mathbf{p}), \quad \Phi_k^{\epsilon}(\mathbf{p}) = \sum_{i=1}^{m} \min\{\delta_{k-1}^{(i)}, \mathfrak{D}(a^i, \ell(\mathbf{p}))\}, \tag{3}$$

where

$$\delta_{k-1}^{(i)} = \min\{\mathfrak{D}(a^i, \tilde{\ell}_1), \ldots, \mathfrak{D}(a^i, \tilde{\ell}_{k-1})\}. \tag{4}$$

After that, KCL algorithm is applied to the set of lines $\{\tilde{\ell}_1, \ldots, \tilde{\ell}_k\}$.

Since by increasing the number of clusters in partition, the objective value decreases, it is reasonable to stop the incremental iterative process when

$$\frac{F_k - F_{k-1}}{F_1} < \epsilon_B,$$

for some small $\epsilon_B > 0$ (say .005).

Smoothed incremental algorithm for multiple line detection

$\tilde{\ell}_1, \ldots, \tilde{\ell}_{k-1}$ are known lines, the next line $\tilde{\ell}_k$ will be obtained by solving the following GOP:

$$\operatorname*{argmin}_{\mathbf{p}\in\mathcal{P}} \Phi_k(\mathbf{p}), \quad \Phi_k^{\epsilon}(\mathbf{p}) = \sum_{i=1}^{m} \min\{\delta_{k-1}^{(i)}, \mathfrak{D}(a^i, \ell(\mathbf{p}))\}, \qquad (3)$$

where

$$\delta_{k-1}^{(i)} = \min\{\mathfrak{D}(a^i, \tilde{\ell}_1), \ldots, \mathfrak{D}(a^i, \tilde{\ell}_{k-1})\}. \qquad (4)$$

After that, KCL algorithm is applied to the set of lines $\{\tilde{\ell}_1, \ldots, \tilde{\ell}_k\}$.
Since by increasing the number of clusters in partition, the objective value decreases, it is reasonable to stop the incremental iterative process when

$$\tfrac{F_k - F_{k-1}}{F_1} < \epsilon_B,$$

for some small $\epsilon_B > 0$ (say .005).

Improving Stoping criteria

Motivated by `DBSCAN` algorithm:

M. Ester, H. Kriegel, J. Sander, A density-based algorithm for discovering clusters in large spatial databases with noise, In: 2nd

International Conference on Knowledge Discovery and Data Mining (KDD-96), Portland, 1996, 226–231.

Let $MinPts(\mathcal{A}) > 2$. For every $a \in \mathcal{A}$, let $\epsilon_a > 0$ be the radius of the smallest disc centered at a and containing $MinPts$ elements of the set $\mathcal{A}$

$MinPts(\mathcal{A}) := \lfloor \log |\mathcal{A}| \rfloor$

R. Scitovski, K. Sabo, DBSCAN-like clustering method for various data densities, Pattern Analysis and Applications 23 (2020), 541-554

Let $\mathcal{R}(\mathcal{A}) = \{\epsilon_a : a \in \mathcal{A}\}$. We define ϵ-*density* of the set $\mathcal{A}$ to be the 99% quantile of the set $\mathcal{R}(\mathcal{A})$ and denote it by $\epsilon(\mathcal{A})$.

Motivated by DBSCAN algorithm:

M. Ester, H. Kriegel, J. Sander, A density-based algorithm for discovering clusters in large spatial databases with noise, In: 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96), Portland, 1996, 226–231.

Let $MinPts(\mathcal{A}) > 2$. For every $a \in \mathcal{A}$, let $\epsilon_a > 0$ be the radius of the smallest disc centered at a and containing $MinPts$ elements of the set $\mathcal{A}$

$MinPts(\mathcal{A}) := \lfloor \log |\mathcal{A}| \rfloor$

R. Scitovski, K. Sabo, DBSCAN-like clustering method for various data densities, Pattern Analysis and Applications 23 (2020), 541-554

Let $\mathcal{R}(\mathcal{A}) = \{\epsilon_a \colon a \in \mathcal{A}\}$. We define ϵ-*density* of the set $\mathcal{A}$ to be the 99% quantile of the set $\mathcal{R}(\mathcal{A})$ and denote it by $\epsilon(\mathcal{A})$.

Improving Stoping criteria

Motivated by DBSCAN algorithm:

M. Ester, H. Kriegel, J. Sander, A density-based algorithm for discovering clusters in large spatial databases with noise, In: 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96), Portland, 1996, 226–231.

Let $MinPts(\mathcal{A}) > 2$. For every $a \in \mathcal{A}$, let $\epsilon_a > 0$ be the radius of the smallest disc centered at a and containing $MinPts$ elements of the set $\mathcal{A}$

$MinPts(\mathcal{A}) := \lfloor \log |\mathcal{A}| \rfloor$

R. Scitovski, K. Sabo, DBSCAN-like clustering method for various data densities, Pattern Analysis and Applications 23 (2020), 541-554

Let $\mathcal{R}(\mathcal{A}) = \{\epsilon_a : a \in \mathcal{A}\}$. We define ϵ-*density* of the set $\mathcal{A}$ to be the 99% quantile of the set $\mathcal{R}(\mathcal{A})$ and denote it by $\epsilon(\mathcal{A})$.

Improving Stoping criteria

Motivated by DBSCAN algorithm:

M. Ester, H. Kriegel, J. Sander, A density-based algorithm for discovering clusters in large spatial databases with noise, In: 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96), Portland, 1996, 226–231.

Let $MinPts(\mathcal{A}) > 2$. For every $a \in \mathcal{A}$, let $\epsilon_a > 0$ be the radius of the smallest disc centered at a and containing $MinPts$ elements of the set $\mathcal{A}$
$MinPts(\mathcal{A}) := \lfloor \log |\mathcal{A}| \rfloor$

R. Scitovski, K. Sabo, DBSCAN-like clustering method for various data densities, Pattern Analysis and Applications 23 (2020), 541-554

Let $\mathcal{R}(\mathcal{A}) = \{\epsilon_a \colon a \in \mathcal{A}\}$. We define ϵ-*density* of the set $\mathcal{A}$ to be the 99% quantile of the set $\mathcal{R}(\mathcal{A})$ and denote it by $\epsilon(\mathcal{A})$.

Note that for almost all points $a \in \mathcal{A}$, the corresponding disc with center a and radius $\epsilon(\mathcal{A})$ contains at least $MinPts$ elements from the set $\mathcal{A}$.
Let $\Pi^{(k)}$ be the partition obtained in step k of the iterative process. For each cluster $\pi \in \Pi^{(k)}$ with line-center $\ell(\mathbf{p})$ let

$$V(\pi) := \{\mathfrak{D}_1(a^i, \ell(\mathbf{p}))\colon a^i \in \pi\},$$

where $\mathfrak{D}_1(a^i, \ell(\mathbf{p})) = |\xi x_i + \eta y_i + \zeta|$ is the ordinary Euclidean distance from the point $a^i = (x_i, y_i) \in \pi$ to the line $\ell(\mathbf{p})$. We define *Quantile of the Data to Line Deviations* (QD) of cluster π as 90% quantile of the set $V(\pi)$.
We expect that the partition $\Pi^{(k)}$ is near k-GOPart if:

$$\mathtt{QD}[\pi] < \epsilon(\mathcal{A}), \quad \forall \pi \in \Pi^{(k)}.$$

Therefore, stopping criterion can be complemented with this condition.

Numerical experiments

The method has been tested on numerous artificial data sets, which have been constructed in the following way: $n \in \{5, 10\}$ lines, which intersect rectangle $\Delta = [0, 10] \times [0, 10]$ and whose mutual Hausdorff distances in rectangle Δ are at least 1, were chosen randomly. To each point on the line noise was added by generating pseudorandom numbers from bivariate normal distribution with mean zero and covariance matrix $\sigma^2 I$, $\sigma^2 \in \{.05, .1\}$, where I is the 2×2 identity matrix. For each pair (n, σ^2), 100 examples were generated.

Results - percent of recognition

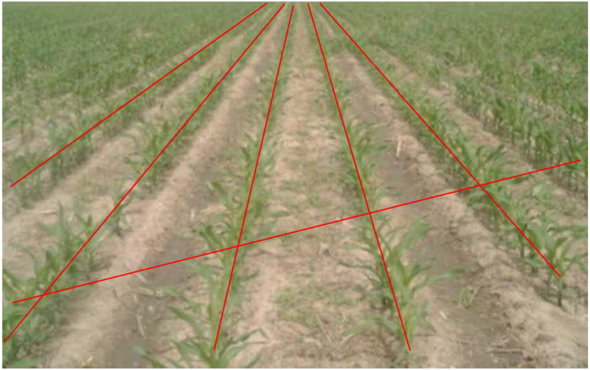
n	$\sigma^2 = 0.05$	$\sigma^2 = 0.1$
5	100%	100%
10	96%	88%

Numerical experiments

The method has been tested on numerous artificial data sets, which have been constructed in the following way: $n \in \{5, 10\}$ lines, which intersect rectangle $\Delta = [0, 10] \times [0, 10]$ and whose mutual Hausdorff distances in rectangle Δ are at least 1, were chosen randomly. To each point on the line noise was added by generating pseudorandom numbers from bivariate normal distribution with mean zero and covariance matrix $\sigma^2 I$, $\sigma^2 \in \{.05, .1\}$, where I is the 2×2 identity matrix. For each pair (n, σ^2), 100 examples were generated.

Results - percent of recognition

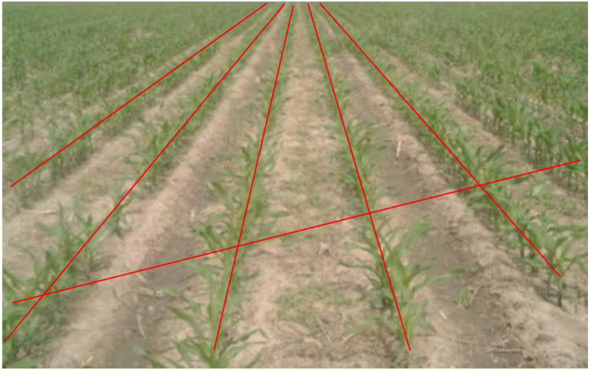
n	$\sigma^2 = 0.05$	$\sigma^2 = 0.1$
5	100%	100%
10	96%	88%

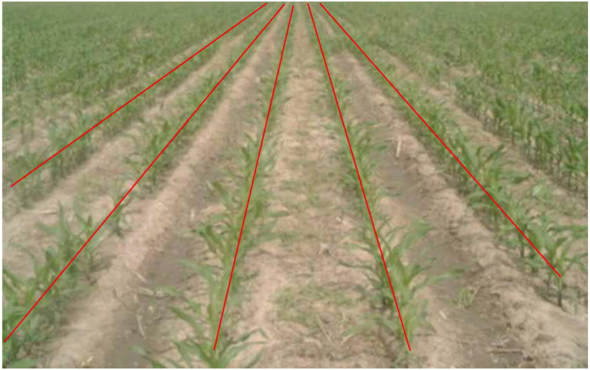


For each cluster $\pi = \pi(\ell) \in \Pi^{(k)}$ we will determine $MinPts(\pi) := \lfloor \log |\pi| \rfloor$ and ϵ-density $\epsilon(\pi)$ of the cluster π. In the sequence $(\epsilon(\pi))_{\pi \in \Pi^{(k)}}$, we will try to identify outliers, so that this sequence is standardized with

$$\nu_\pi := \left|\epsilon(\pi) - \operatorname*{med}_{\pi \in \Pi^{(k)}} \epsilon(\pi)\right| / \mathtt{MAD}, \qquad \mathtt{MAD} = 1.483 \operatorname*{med}_{\pi \in \Pi^{(k)}} \left|\epsilon(\pi) - \operatorname*{med}_{\pi \in \Pi^{(k)}}\right|. \tag{5}$$

The center-line ℓ of a cluster π will be dropped if $\nu_\pi > 3.5$. This would mean that such cluster is more sparse around its center-line than other clusters.





Method weakness

- Problems when data are in noisy environment $\Rightarrow$ Further research!