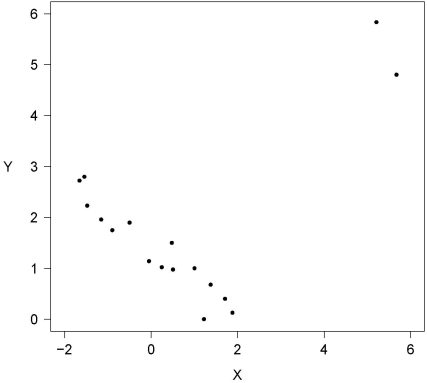
On the implementation of Likelihood-based Imprecise Regression

Marco Cattaneo and Andrea Wiencierz
Department of Statistics, LMU Munich

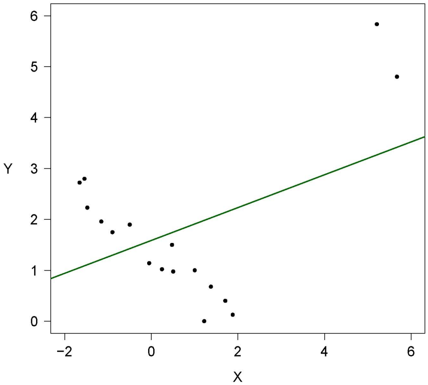
CFE-ERCIM 2011, London, UK
18 December 2011

simple linear regression



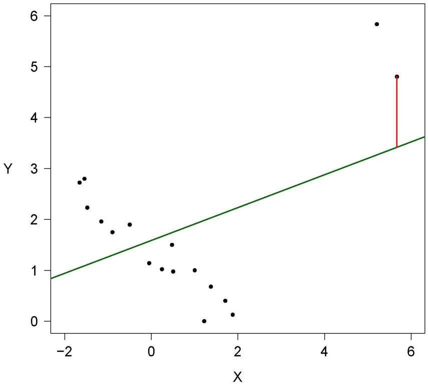
- precise data: $(x_i, y_i) \in \mathbb{R}^2$ for each $i \in \{1, \ldots, n\}$

simple linear regression



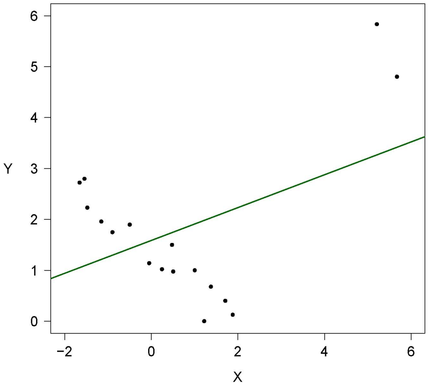
- precise data: $(x_i, y_i) \in \mathbb{R}^2$ for each $i \in \{1, \ldots, n\}$
- linear regression: $\mathcal{F} = \{f_{a,b} : a, b \in \mathbb{R}\}$ with $f_{a,b} : x \mapsto a + b\,x$

simple linear regression



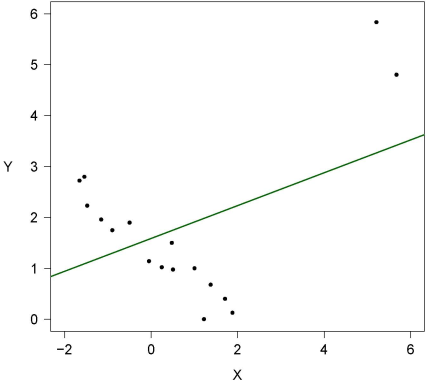
- precise data: $(x_i, y_i) \in \mathbb{R}^2$ for each $i \in \{1, \dots, n\}$
- linear regression: $\mathcal{F} = \{f_{a,b} : a, b \in \mathbb{R}\}$ with $f_{a,b} : x \mapsto a + b\,x$
- (absolute) residuals: $r_{f,i} = |y_i - f(x_i)|$ for each $f \in \mathcal{F}$, $i \in \{1, \dots, n\}$

simple linear regression



- precise data: $(x_i, y_i) \in \mathbb{R}^2$ for each $i \in \{1, \ldots, n\}$
- linear regression: $\mathcal{F} = \{f_{a,b} : a, b \in \mathbb{R}\}$ with $f_{a,b} : x \mapsto a + b\,x$
- (absolute) residuals: $r_{f,i} = |y_i - f(x_i)|$ for each $f \in \mathcal{F}$, $i \in \{1, \ldots, n\}$
- Least Squares: $f_{LS} = \arg\min_f \sum_i r_{f,i}^2$

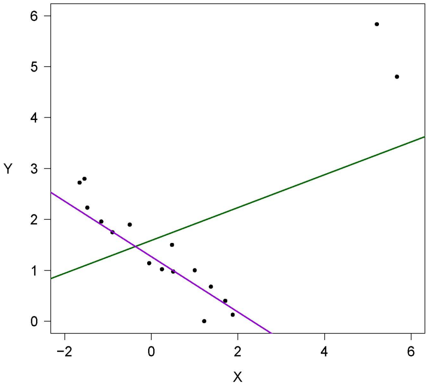
simple linear regression



breakdown point:
$\varepsilon^*_{LS} = 0$

- precise data: $(x_i, y_i) \in \mathbb{R}^2$ for each $i \in \{1, \ldots, n\}$
- linear regression: $\mathcal{F} = \{f_{a,b} : a, b \in \mathbb{R}\}$ with $f_{a,b} : x \mapsto a + b\,x$
- (absolute) residuals: $r_{f,i} = |y_i - f(x_i)|$ for each $f \in \mathcal{F}$, $i \in \{1, \ldots, n\}$
- Least Squares: $f_{LS} = \arg\min_f \sum_i r_{f,i}^2 = \arg\min_f \mathit{mean}_i\, r_{f,i}^2$

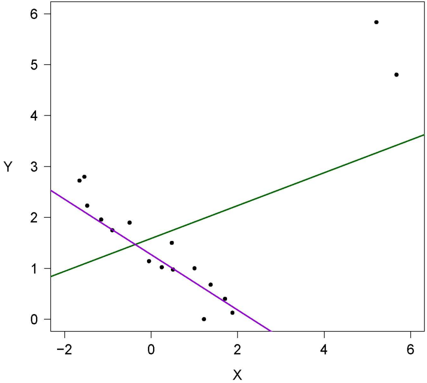
simple linear regression



breakdown point:
$\varepsilon^*_{LS} = 0$
$\varepsilon^*_{LMS} = \frac{1}{n} \left\lfloor \frac{n-1}{2} \right\rfloor \xrightarrow{n \to \infty} \frac{1}{2}$

- precise data: $(x_i, y_i) \in \mathbb{R}^2$ for each $i \in \{1, \dots, n\}$
- linear regression: $\mathcal{F} = \{f_{a,b} : a, b \in \mathbb{R}\}$ with $f_{a,b} : x \mapsto a + b\,x$
- (absolute) residuals: $r_{f,i} = |y_i - f(x_i)|$ for each $f \in \mathcal{F}$, $i \in \{1, \dots, n\}$
- Least Squares: $f_{LS} = \arg\min_f \sum_i r^2_{f,i} = \arg\min_f mean_i\, r^2_{f,i}$
- Least Median of Squares: $f_{LMS} = \arg\min_f med_i\, r^2_{f,i}$

simple linear regression



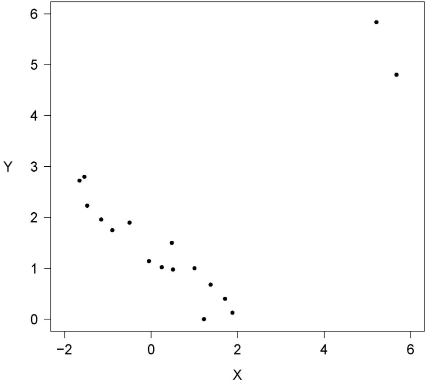
breakdown point:

$\varepsilon^*_{LS} = 0$

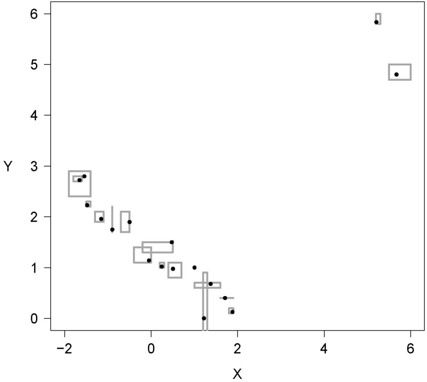
$\varepsilon^*_{LMS} = \frac{1}{n} \left\lfloor \frac{n-1}{2} \right\rfloor \xrightarrow{n\to\infty} \frac{1}{2}$

- precise data: $(x_i, y_i) \in \mathbb{R}^2$ for each $i \in \{1, \ldots, n\}$
- linear regression: $\mathcal{F} = \{f_{a,b} : a, b \in \mathbb{R}\}$ with $f_{a,b} : x \mapsto a + b\,x$
- (absolute) residuals: $r_{f,i} = |y_i - f(x_i)|$ for each $f \in \mathcal{F}$, $i \in \{1, \ldots, n\}$
- Least Squares: $f_{LS} = \arg\min_f \sum_i r^2_{f,i} = \arg\min_f \mathit{mean}_i\, r^2_{f,i}$
- Least Median of Squares: $f_{LMS} = \arg\min_f \mathit{med}_i\, r^2_{f,i} = \arg\min_f \mathit{med}_i\, r_{f,i}$

imprecisely observed data

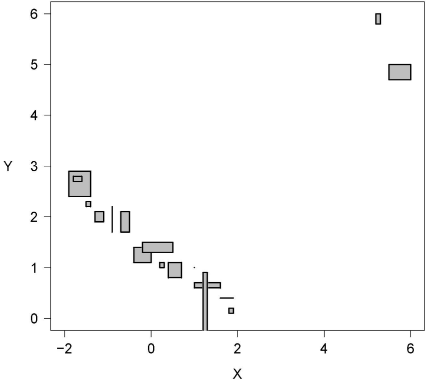


imprecisely observed data



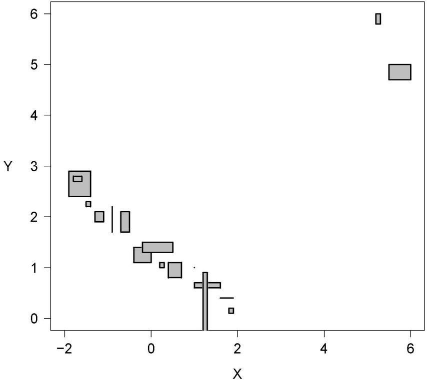
- imprecise data: $(\underline{x}_i, \overline{x}_i, \underline{y}_i, \overline{y}_i) \in \overline{\mathbb{R}}^4$ for each $i \in \{1, \ldots, n\}$

imprecisely observed data



- imprecise data: $(\underline{x}_i, \overline{x}_i, \underline{y}_i, \overline{y}_i) \in \overline{\mathbb{R}}^4$ for each $i \in \{1, \ldots, n\}$

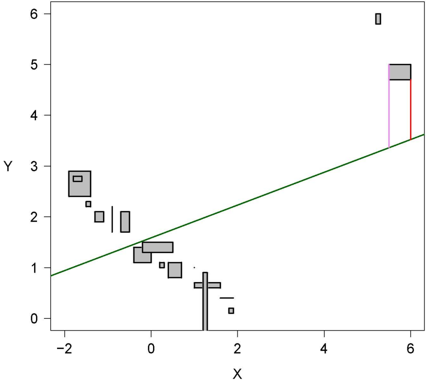
imprecisely observed data



- imprecise data: $(\underline{x}_i, \overline{x}_i, \underline{y}_i, \overline{y}_i) \in \overline{\mathbb{R}}^4$ for each $i \in \{1, \ldots, n\}$
- nonparametric statistical model: $\mathcal{P}$ is the set of all probability measures P such that $\underline{X}_i, X_i, \overline{X}_i, \underline{Y}_i, Y_i, \overline{Y}_i$ have a joint distribution satisfying

$$\underline{X}_i \leq X_i \leq \overline{X}_i \quad \text{and} \quad \underline{Y}_i \leq Y_i \leq \overline{Y}_i \quad P\text{-a.s.}$$

imprecisely observed data



imprecise residuals:

$$\underline{r}_{f,i} = \min_{(x,y)\in[\underline{x}_i,\overline{x}_i]\times[\underline{y}_i,\overline{y}_i]} |y - f(x)|$$

$$\overline{r}_{f,i} = \sup_{(x,y)\in[\underline{x}_i,\overline{x}_i]\times[\underline{y}_i,\overline{y}_i]} |y - f(x)|$$

for each $f \in \mathcal{F}$, $i \in \{1, \dots, n\}$

- imprecise data: $(\underline{x}_i, \overline{x}_i, \underline{y}_i, \overline{y}_i) \in \overline{\mathbb{R}}^4$ for each $i \in \{1, \dots, n\}$
- nonparametric statistical model: $\mathcal{P}$ is the set of all probability measures P such that $\underline{X}_i, X_i, \overline{X}_i, \underline{Y}_i, Y_i, \overline{Y}_i$ have a joint distribution satisfying

$$\underline{X}_i \leq X_i \leq \overline{X}_i \quad \text{and} \quad \underline{Y}_i \leq Y_i \leq \overline{Y}_i \quad P\text{-a.s.}$$

Likelihood-based Imprecise Regression

- imprecise probability models naturally appear with imprecise data:

Likelihood-based Imprecise Regression

- imprecise probability models naturally appear with imprecise data: for example, the empirical joint distribution $\widehat{P}_{\underline{X},\overline{X},\underline{Y},\overline{Y}}$ of the imprecise data corresponds to an imprecise joint distribution for the precise data: $\underline{\widehat{P}}_{X,Y}$ is a belief function with focal sets $[\underline{x}_i, \overline{x}_i] \times [\underline{y}_i, \overline{y}_i]$

Likelihood-based Imprecise Regression

- imprecise probability models naturally appear with imprecise data: for example, the empirical joint distribution $\widehat{P}_{\underline{X},\overline{X},\underline{Y},\overline{Y}}$ of the imprecise data corresponds to an imprecise joint distribution for the precise data: $\underline{\widehat{P}}_{X,Y}$ is a belief function with focal sets $[\underline{x}_i, \overline{x}_i] \times [\underline{y}_i, \overline{y}_i]$
- likelihood function: $lik : P \mapsto \prod_{i=1}^{n} \dfrac{P(\underline{X}_i = \underline{x}_i,\, \overline{X}_i = \overline{x}_i,\, \underline{Y}_i = \underline{y}_i,\, \overline{Y}_i = \overline{y}_i)}{\widehat{P}_{\underline{X},\overline{X},\underline{Y},\overline{Y}}(\underline{x}_i, \overline{x}_i, \underline{y}_i, \overline{y}_i)}$

Likelihood-based Imprecise Regression

- imprecise probability models naturally appear with imprecise data: for example, the empirical joint distribution $\widehat{P}_{\underline{X},\overline{X},\underline{Y},\overline{Y}}$ of the imprecise data corresponds to an imprecise joint distribution for the precise data: $\underline{\widehat{P}}_{X,Y}$ is a belief function with focal sets $[\underline{x}_i, \overline{x}_i] \times [\underline{y}_i, \overline{y}_i]$
- likelihood function: $lik : P \mapsto \prod_{i=1}^{n} \dfrac{P(\underline{X}_i = \underline{x}_i,\, \overline{X}_i = \overline{x}_i,\, \underline{Y}_i = \underline{y}_i,\, \overline{Y}_i = \overline{y}_i)}{\widehat{P}_{\underline{X},\overline{X},\underline{Y},\overline{Y}}(\underline{x}_i, \overline{x}_i, \underline{y}_i, \overline{y}_i)}$
- likelihood-based learning of imprecise probability model: $\mathcal{P}_{>\beta} = \{P \in \mathcal{P} : lik(P) > \beta\}$ for some cutoff point $\beta \in (0, 1)$

Likelihood-based Imprecise Regression

- imprecise probability models naturally appear with imprecise data: for example, the empirical joint distribution $\widehat{P}_{\underline{X},\overline{X},\underline{Y},\overline{Y}}$ of the imprecise data corresponds to an imprecise joint distribution for the precise data: $\underline{\widehat{P}}_{X,Y}$ is a belief function with focal sets $[\underline{x}_i, \overline{x}_i] \times [\underline{y}_i, \overline{y}_i]$
- likelihood function: $\mathit{lik} : P \mapsto \prod_{i=1}^{n} \dfrac{P(\underline{X}_i = \underline{x}_i,\, \overline{X}_i = \overline{x}_i,\, \underline{Y}_i = \underline{y}_i,\, \overline{Y}_i = \overline{y}_i)}{\widehat{P}_{\underline{X},\overline{X},\underline{Y},\overline{Y}}(\underline{x}_i, \overline{x}_i, \underline{y}_i, \overline{y}_i)}$
- likelihood-based learning of imprecise probability model: $\mathcal{P}_{>\beta} = \{P \in \mathcal{P} : \mathit{lik}(P) > \beta\}$ for some cutoff point $\beta \in (0, 1)$
- if $\beta \geq 2^{-n}$, then for each $f \in \mathcal{F}$, the median of the distribution of the (precise) residuals is imprecise under the model $\mathcal{P}_{>\beta}$:

$$\underline{\mathit{med}} R_f = \underline{r}_{f,(\underline{k}+1)} \quad \text{and} \quad \overline{\mathit{med}} R_f = \overline{r}_{f,(\overline{k})},$$

where $\sqrt[n]{\beta} \mapsto \frac{\overline{k}}{n}$ is a decreasing bijection $[\frac{1}{2}, 1) \to (\frac{1}{2}, 1]$, and $\underline{k} = n - \overline{k}$

Likelihood-based Imprecise Regression

- imprecise probability models naturally appear with imprecise data: for example, the empirical joint distribution $\widehat{P}_{\underline{X},\overline{X},\underline{Y},\overline{Y}}$ of the imprecise data corresponds to an imprecise joint distribution for the precise data: $\underline{\widehat{P}}_{X,Y}$ is a belief function with focal sets $[\underline{x}_i, \overline{x}_i] \times [\underline{y}_i, \overline{y}_i]$
- likelihood function: $lik : P \mapsto \prod_{i=1}^{n} \frac{P(\underline{X}_i = \underline{x}_i,\, \overline{X}_i = \overline{x}_i,\, \underline{Y}_i = \underline{y}_i,\, \overline{Y}_i = \overline{y}_i)}{\widehat{P}_{\underline{X},\overline{X},\underline{Y},\overline{Y}}(\underline{x}_i, \overline{x}_i, \underline{y}_i, \overline{y}_i)}$
- likelihood-based learning of imprecise probability model: $\mathcal{P}_{>\beta} = \{P \in \mathcal{P} : lik(P) > \beta\}$ for some cutoff point $\beta \in (0, 1)$
- if $\beta \geq 2^{-n}$, then for each $f \in \mathcal{F}$, the median of the distribution of the (precise) residuals is imprecise under the model $\mathcal{P}_{>\beta}$:

$$\underline{med} R_f = \underline{r}_{f,(\underline{k}+1)} \quad \text{and} \quad \overline{med} R_f = \overline{r}_{f,(\overline{k})},$$

  where $\sqrt[n]{\beta} \mapsto \frac{\overline{k}}{n}$ is a decreasing bijection $[\frac{1}{2}, 1) \to (\frac{1}{2}, 1]$, and $\underline{k} = n - \overline{k}$
- Likelihood-based Region Minimax: $f_{LRM} = \arg\min_f \overline{med} R_f = \arg\min_f \overline{r}_{f,(\overline{k})}$

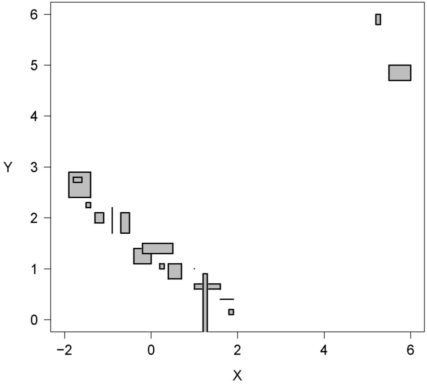
Likelihood-based Imprecise Regression

- imprecise probability models naturally appear with imprecise data: for example, the empirical joint distribution $\widehat{P}_{\underline{X},\overline{X},\underline{Y},\overline{Y}}$ of the imprecise data corresponds to an imprecise joint distribution for the precise data: $\underline{\widehat{P}}_{X,Y}$ is a belief function with focal sets $[\underline{x}_i, \overline{x}_i] \times [\underline{y}_i, \overline{y}_i]$
- likelihood function: $lik : P \mapsto \prod_{i=1}^{n} \frac{P(\underline{X}_i = \underline{x}_i,\, \overline{X}_i = \overline{x}_i,\, \underline{Y}_i = \underline{y}_i,\, \overline{Y}_i = \overline{y}_i)}{\widehat{P}_{\underline{X},\overline{X},\underline{Y},\overline{Y}}(\underline{x}_i, \overline{x}_i, \underline{y}_i, \overline{y}_i)}$
- likelihood-based learning of imprecise probability model: $\mathcal{P}_{>\beta} = \{P \in \mathcal{P} : lik(P) > \beta\}$ for some cutoff point $\beta \in (0,1)$
- if $\beta \geq 2^{-n}$, then for each $f \in \mathcal{F}$, the median of the distribution of the (precise) residuals is imprecise under the model $\mathcal{P}_{>\beta}$:

$$\underline{med}\, R_f = \underline{r}_{f,(\underline{k}+1)} \quad \text{and} \quad \overline{med}\, R_f = \overline{r}_{f,(\overline{k})},$$

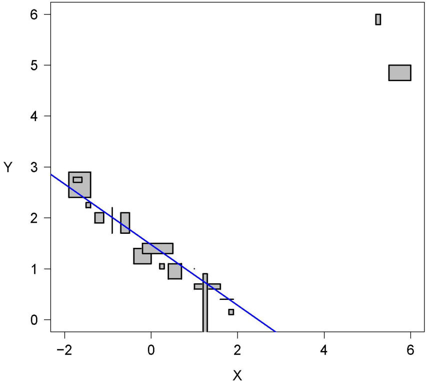
  where $\sqrt[n]{\beta} \mapsto \frac{\overline{k}}{n}$ is a decreasing bijection $[\frac{1}{2}, 1) \to (\frac{1}{2}, 1]$, and $\underline{k} = n - \overline{k}$
- Likelihood-based Region Minimax: $f_{LRM} = \arg\min_f \overline{med}\, R_f = \arg\min_f \overline{r}_{f,(\overline{k})}$
- interval dominance: $\mathcal{U} = \{f \in \mathcal{F} : \underline{med}\, R_f \leq \overline{med}\, R_{f_{LRM}}\}$ is the set of all undominated regression lines

algorithm for f_{LRM}



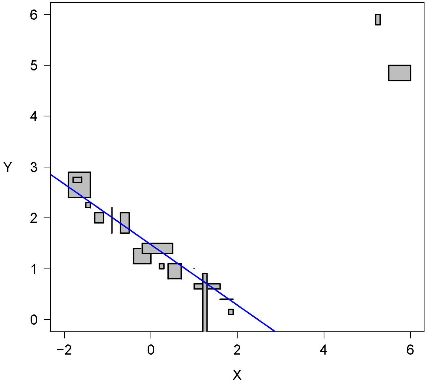
$n = 17$

algorithm for f_{LRM}



$n = 17$
$\beta = 0.8$
$\Rightarrow \overline{k} = 10$

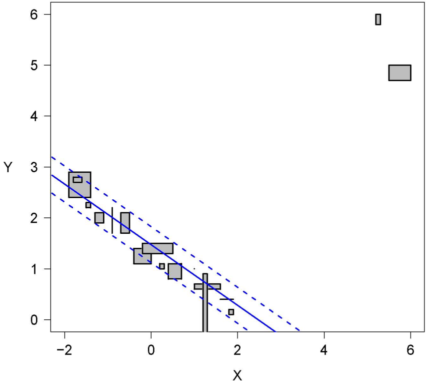
algorithm for f_{LRM}



$n = 17$
$\beta = 0.8$
$\Rightarrow \overline{k} = 10$

if less that $\overline{k}$ intervals $[\underline{y}_i, \overline{y}_i]$ are bounded, then $\overline{med R_f} = +\infty$ for each $f \in \mathcal{F}$

algorithm for f_{LRM}

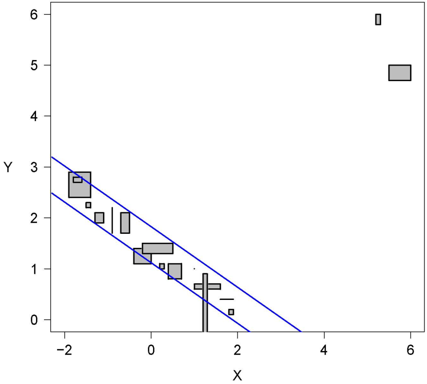


$n = 17$
$\beta = 0.8$
$\Rightarrow \overline{k} = 10$

if less that $\overline{k}$ intervals $[\underline{y}_i, \overline{y}_i]$ are bounded, then $\overline{med} R_f = +\infty$ for each $f \in \mathcal{F}$

otherwise, consider the strip $f_{LRM} \pm \overline{med} R_{f_{LRM}}$

algorithm for f_{LRM}

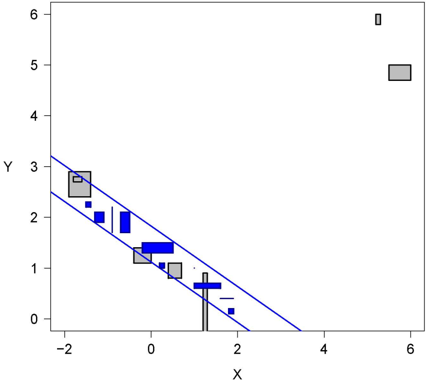


$n = 17$
$\beta = 0.8$
$\Rightarrow \overline{k} = 10$

if less that $\overline{k}$ intervals $[\underline{y}_i, \overline{y}_i]$ are bounded, then $\overline{med R_f} = +\infty$ for each $f \in \mathcal{F}$

otherwise, consider the strip
$f_{LRM} \pm \overline{med} R_{f_{LRM}} = f_{LRM} \pm \overline{r}_{f_{LRM},(\overline{k})}$

algorithm for f_{LRM}



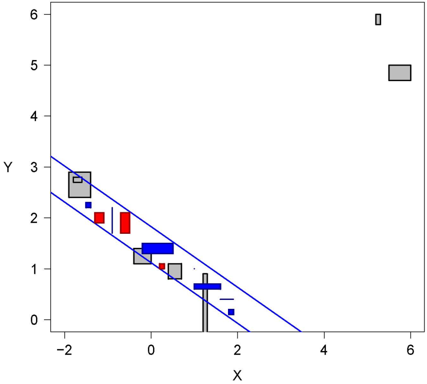
$n = 17$
$\beta = 0.8$
$\Rightarrow \overline{k} = 10$

if less that $\overline{k}$ intervals $[\underline{y}_i, \overline{y}_i]$ are bounded, then $\overline{med}\, R_f = +\infty$ for each $f \in \mathcal{F}$

otherwise, consider the strip $f_{LRM} \pm \overline{med}\, R_{f_{LRM}} = f_{LRM} \pm \overline{r}_{f_{LRM},(\overline{k})}$

- $f_{LRM} \pm \overline{med}\, R_{f_{LRM}}$ is the thinnest strip of the form $f \pm q$ containing (at least) $\overline{k}$ imprecise data $[\underline{x}_i, \overline{x}_i] \times [\underline{y}_i, \overline{y}_i]$, for all $f \in \mathcal{F}$, $q \in [0, +\infty)$

algorithm for f_{LRM}



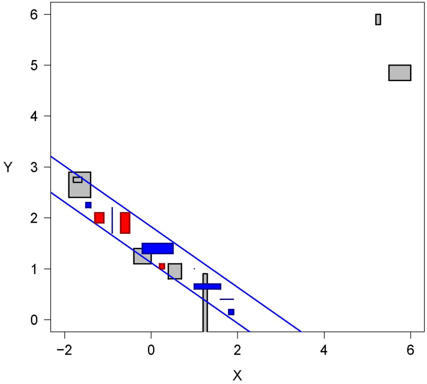
$n = 17$
$\beta = 0.8$
$\Rightarrow \overline{k} = 10$

if less that $\overline{k}$ intervals $[\underline{y}_i, \overline{y}_i]$ are bounded, then $\overline{med} R_f = +\infty$ for each $f \in \mathcal{F}$

otherwise, consider the strip $f_{LRM} \pm \overline{med} R_{f_{LRM}} = f_{LRM} \pm \overline{r}_{f_{LRM},(\overline{k})}$

- $f_{LRM} \pm \overline{med} R_{f_{LRM}}$ is the thinnest strip of the form $f \pm q$ containing (at least) $\overline{k}$ imprecise data $[\underline{x}_i, \overline{x}_i] \times [\underline{y}_i, \overline{y}_i]$, for all $f \in \mathcal{F}$, $q \in [0, +\infty)$
- if the slope $b_{LRM} \neq 0$, then the imprecise data contained in $f_{LRM} \pm \overline{med} R_{f_{LRM}}$ are bounded and (at least) 3 of them touch the boundary of the strip

algorithm for f_{LRM}



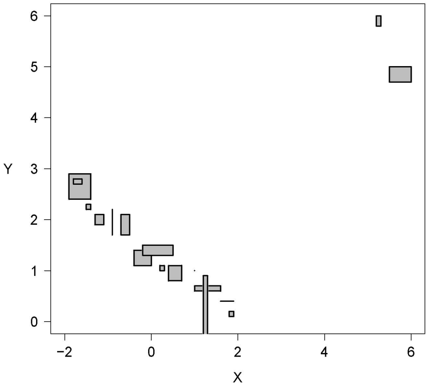
$n = 17$
$\beta = 0.8$
$\Rightarrow \overline{k} = 10$

if less that $\overline{k}$ intervals $[\underline{y}_i, \overline{y}_i]$ are bounded, then $\overline{med} R_f = +\infty$ for each $f \in \mathcal{F}$

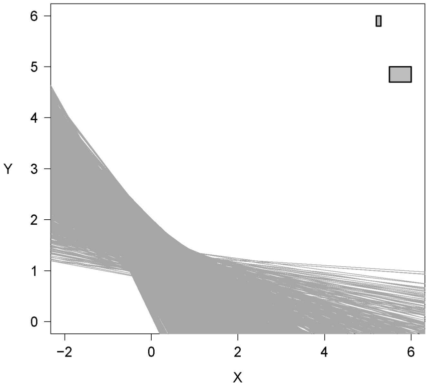
otherwise, consider the strip $f_{LRM} \pm \overline{med} R_{f_{LRM}} = f_{LRM} \pm \overline{r}_{f_{LRM},(\overline{k})}$

- $f_{LRM} \pm \overline{med} R_{f_{LRM}}$ is the thinnest strip of the form $f \pm q$ containing (at least) $\overline{k}$ imprecise data $[\underline{x}_i, \overline{x}_i] \times [\underline{y}_i, \overline{y}_i]$, for all $f \in \mathcal{F}$, $q \in [0, +\infty)$
- if the slope $b_{LRM} \neq 0$, then the imprecise data contained in $f_{LRM} \pm \overline{med} R_{f_{LRM}}$ are bounded and (at least) 3 of them touch the boundary of the strip
- therefore, b_{LRM} is either 0 or it is determined by a couple of bounded imprecise data, which gives us at most $4\binom{n}{2} + 1$ possible values for b_{LRM}

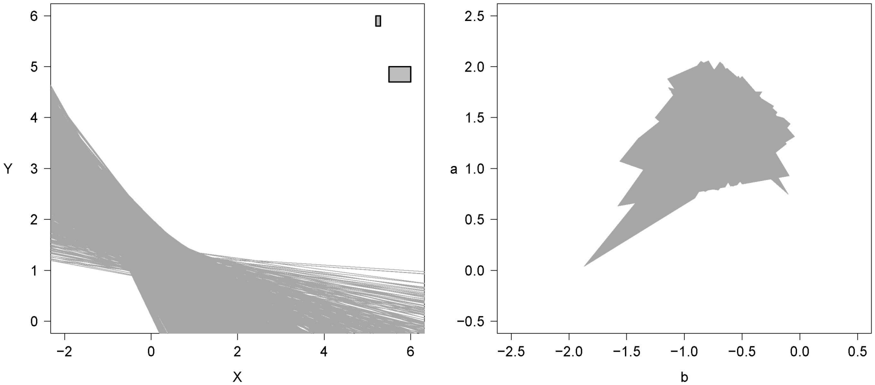
undominated regression lines



undominated regression lines

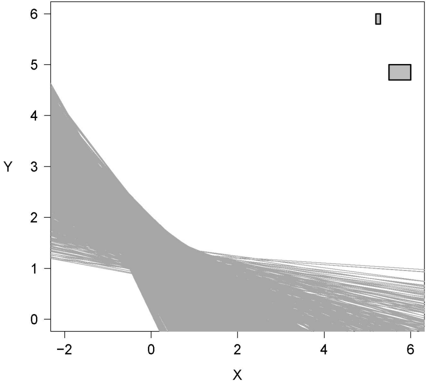


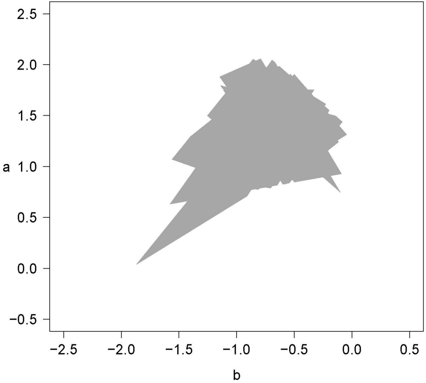
undominated regression lines



- set of undominated parameters: $\left\{(a, b) \in \mathbb{R}^2 : f_{a,b} \in \mathcal{U}\right\}$

undominated regression lines



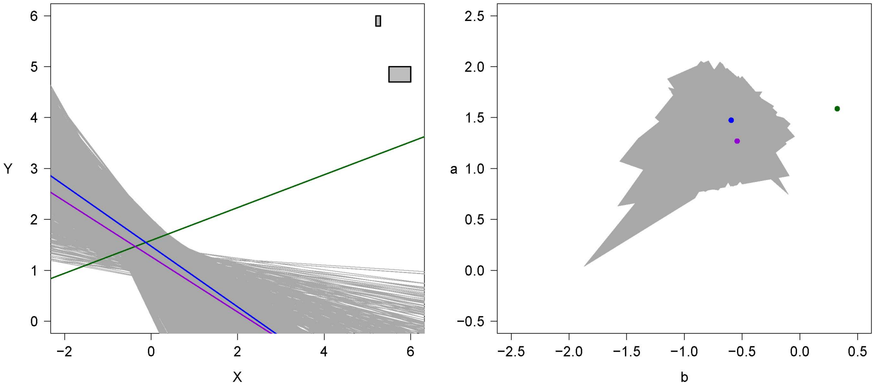


- set of undominated parameters: $\left\{(a,b) \in \mathbb{R}^2 : f_{a,b} \in \mathcal{U}\right\} =$

$$= \bigcup_{i=1}^{\overline{k}} \left\{(a,b) \in \mathbb{R}^2 : \underline{d}_{b,(i+\underline{k})} - \overline{med} R_{f_{LRM}} \leq a \leq \overline{d}_{b,(i)} + \overline{med} R_{f_{LRM}} \right\},$$

where $\underline{d}_{b,i} = \inf_{x \in [\underline{x}_i, \overline{x}_i]} (\underline{y}_i - b\,x)$ and $\overline{d}_{b,i} = \sup_{x \in [\underline{x}_i, \overline{x}_i]} (\overline{y}_i - b\,x)$

undominated regression lines



- set of undominated parameters: $\left\{(a, b) \in \mathbb{R}^2 : f_{a,b} \in \mathcal{U}\right\} =$

$$= \bigcup_{i=1}^{\overline{k}} \left\{(a, b) \in \mathbb{R}^2 : \underline{d}_{b,(i+\underline{k})} - \overline{med}\,R_{f_{LRM}} \leq a \leq \overline{d}_{b,(i)} + \overline{med}\,R_{f_{LRM}}\right\},$$

where $\underline{d}_{b,i} = \inf_{x \in [\underline{x}_i, \overline{x}_i]} (\underline{y}_i - b\,x)$ and $\overline{d}_{b,i} = \sup_{x \in [\underline{x}_i, \overline{x}_i]} (\overline{y}_i - b\,x)$

- for example: $\overline{med}\,R_{f_{LRM}} \approx 0.354$, $\underline{med}\,R_{f_{LMS}} \approx 0.002$, $\underline{med}\,R_{f_{LS}} \approx 0.909$

statistical properties of LIR

- breakdown point: $\varepsilon^*_{LIR} = \frac{k}{n} \xrightarrow{n \to \infty} \frac{1}{2}$

statistical properties of LIR

- breakdown point: $\varepsilon^*_{LIR} = \frac{k}{n} \xrightarrow{n \to \infty} \frac{1}{2}$
- coverage probability of $\mathcal{U}$: $Y_i = a_0 + b_0\, X_i + \varepsilon_i$ with $X_i, \varepsilon_i \overset{i.i.d.}{\sim} F_0$

statistical properties of LIR

- breakdown point: $\varepsilon^*_{LIR} = \frac{k}{n} \xrightarrow{n \to \infty} \frac{1}{2}$
- coverage probability of $\mathcal{U}$: $Y_i = a_0 + b_0\, X_i + \varepsilon_i$ with $X_i, \varepsilon_i \overset{i.i.d.}{\sim} F_0$

β	n	$\underline{P}(\underline{med}\, R_f \leq med\, R_f \leq \overline{med}\, R_f)$	F_0	$\underline{P}(f_{a_0,b_0} \in \mathcal{U})$
0.5	20	0.737	*Normal*	0.83
			Cauchy	0.97
	1000	0.758	*Normal*	1.00
			Cauchy	1.00
0.75	20	0.497	*Normal*	0.39
			Cauchy	0.72
	1000	0.533	*Normal*	0.91
			Cauchy	1.00
0.999	20	0.176	*Normal*	0.03
			Cauchy	0.11
	1000	0.025	*Normal*	0.00
			Cauchy	0.01

references

Cattaneo, M., and Wiencierz, A. (2011*a*). Likelihood-based Imprecise Regression. Technical Report 116. Department of Statistics, LMU Munich.

Cattaneo, M., and Wiencierz, A. (2011*b*). Regression with Imprecise Data: A Robust Approach. In *ISIPTA '11, Proceedings of the Seventh International Symposium on Imprecise Probability: Theories and Applications*. SIPTA, 119–128.

Cattaneo, M., and Wiencierz, A. (2011*c*). Robust regression with imprecise data. Technical Report 114. Department of Statistics, LMU Munich.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J., and Stahel, W. A. (1986). *Robust Statistics: The Approach Based on Influence Functions*. Wiley.

Rousseeuw, P. J., and Leroy, A. M. (1987). *Robust Regression and Outlier Detection*. Wiley.

Steele, J., and Steiger, W. (1986). Algorithms and complexity for least median of squares regression. *Discrete Appl. Math.* 14, 93–100.