A note on sharp identification regions

Definition

Let $P := \{\mathbb{P}_\theta \mid \theta \in \Theta\}$ *be a statsitical model and*

- $Y, \ldots$ *unobservable random variables,*
- $X, \underline{Y}, \overline{Y}, \ldots$ *observable random variables w.r.t an underlying probability space* $(\Omega, \mathcal{F}, \mathbb{P})$.
- *The joint distribution of the random Variables* $X, Y, \underline{Y}, \overline{Y}$ *under a model* P_θ *is denoted with* F_θ *and the joint distribution under the „true model"* $\mathbb{P}$ *is denoted with* $F^{X,Y,\underline{Y},\overline{Y}}$.
- *The unobserved variables fullfill a certain condition* $C(X, Y, \underline{Y}, \overline{Y}) = 1$.
  e.g. $\underline{Y} \leq Y \leq \overline{Y}$ *or* $\forall X : \mathbb{E}(\underline{Y} \mid X) \leq \mathbb{E}(Y \mid X) \leq \mathbb{E}(\overline{Y} \mid X)$.

Definition

- *Two parameters* θ_1 *and* θ_2 *are undistinguishable (i.e.* $\theta_1 \sim \theta_2$*) if the corresponding models* $\mathbb{P}_{\theta_1}$ *and* $\mathbb{P}_{\theta_2}$ *are empirically undistinguishable, which means, that the distributions of the observable variables are the same:*

$$F_{\theta_1}^{X,\underline{Y},\overline{Y}} = F_{\theta_2}^{X,\underline{Y},\overline{Y}}.$$

Definition

A statistical model P is called point-identified, if any two different parameters θ_1 *and* θ_2 *are empirically distinguishable, i.e.:*

$$\sim \quad = \quad \Delta_\Theta = \{(\theta, \theta) \mid \theta \in \Theta\}.$$

Otherwise it is called partially identified.

Example

The simple linear model

$$\Theta = B \times \mathbb{R}_{\geq 0} \times \mathcal{Z}(\mathbb{R}_{\geq 0}) \times \mathcal{Z}(\mathbb{R}_{\geq 0})$$

with $B = \mathbb{R}^2$. For $\theta = (\beta, \sigma^2, \sigma_l, \sigma_u) \in \Theta$, the random variables are defined as:

$$\begin{aligned} Y &= X\beta + \varepsilon \\ \underline{Y} &= X\beta + \varepsilon - \sigma_l \\ \overline{Y} &= X\beta + \varepsilon + \sigma_u \end{aligned}$$

with $\varepsilon \sim N(0, \sigma^2 I)$.

Example

The simple linear model

$$\Theta = B \times \mathbb{R}_{\geq 0} \times \mathcal{Z}(\mathbb{R}_{\geq 0}) \times \mathcal{Z}(\mathbb{R}_{\geq 0})$$

with $B = \mathbb{R}^2$. For $\theta = (\beta, \sigma^2, \sigma_l, \sigma_u) \in \Theta$, the random variables are defined as:

$$\begin{aligned} Y &= X\beta + \varepsilon \\ \underline{Y} &= X\beta + \varepsilon - \sigma_l \\ \overline{Y} &= X\beta + \varepsilon + \sigma_u \end{aligned}$$

with $\varepsilon \sim N(0, \sigma^2 I)$.

Here we are only interested in the values of $\beta \in B$.

This model is only partially identified. For example

$$((\beta_0, \beta_1), \sigma^2, 0, 1) \quad \sim \quad ((\beta_0 + 1, \beta_1), \sigma^2, 1, 0).$$

This model is only partially identified. For example

$$((\beta_0, \beta_1), \sigma^2, 0, 1) \quad \sim \quad ((\beta_0 + 1, \beta_1), \sigma^2, 1, 0).$$

Moreover, the quotient space $\Theta_{/\sim}$ ist not of the form

$$\Theta_{/\sim} \quad = \quad B_{/\approx} \times \text{„rest“},$$

so we must factorize the whole space Θ and not only the interesting B to make the model point-identified.

Estimation ⟵ *Model* ⟶ *Pediction*

Estimation ⟵ Model ⟶ Pediction

„model as a truth to be estimated “

„model as a tool to be applied“

Estimation ⟵ *Model* ⟶ *Pediction*

„model as a truth to be estimated “

„model as a tool to be applied“

e.g.: *least squares estimator* ⟵ *linear model* ⟶ *best linear predictor*

„Estimation“

Given distribution F^Y *of* Y *of the class* $\{F_\theta^Y \mid \theta \in \Theta\}$,

„Estimation“

Given distribution F^Y *of* Y *of the class* $\{F^Y_\theta \mid \theta \in \Theta\}$,
find (all) θ, *such that*

$$Y \sim F_\theta$$

„Estimation“

Given distribution F^Y *of* Y *of the class* $\{F_\theta^Y \mid \theta \in \Theta\}$,
find (all) θ, *such that*

$$
\begin{aligned}
& Y \sim F_\theta \\
\iff \quad & F^Y = F_\theta^Y
\end{aligned}
$$

„Estimation“

Given distribution F^Y *of* Y *of the class* $\{F_\theta^Y \mid \theta \in \Theta\}$,
find (all) θ, *such that*

$$
\begin{aligned}
& Y \sim F_\theta \\
\iff \quad & F^Y = F_\theta^Y \\
\iff \quad & L(F_\theta^Y, F) = 0
\end{aligned}
$$

for some distance-function $L(\cdot,\cdot)$.

„Prediction“

Given F^Y $\boxed{\textit{of the class } \{F^Y_\theta \mid \theta \in \Theta\}}$,

„Prediction"

Given F^Y *of the class* $\{F_\theta^Y \mid \theta \in \Theta\}$,
find (all) θ, *such that*

$$L(F_\theta, F^Y)$$

is minimal.

„Prediction“

Given F^Y $\boxed{\textit{of the class } \{F_\theta^Y \mid \theta \in \Theta\}}$*,*
find (all) θ*, such that*

$$L(F_\theta, F^Y)$$

is minimal.

- *also makes sense, if* $F^Y \notin \{F_\theta^Y \mid \theta \in \Theta\}$.

„Prediction“

Given F^Y *of the class* $\{F_\theta^Y \mid \theta \in \Theta\}$,
find (all) θ, *such that*

$$L(F_\theta, F^Y)$$

is minimal.

- *also makes sense, if* $F^Y \notin \{F_\theta^Y \mid \theta \in \Theta\}$.
- *if the model is correctly specified, then „prediction“ and „estimation“ are „nearly the same“.*

„Prediction“

Given F^Y *of the class* $\{F_\theta^Y \mid \theta \in \Theta\}$,
find (all) θ, *such that*

$$L(F_\theta, F^Y)$$

is minimal.

- *also makes sense, if* $F^Y \notin \{F_\theta^Y \mid \theta \in \Theta\}$.
- *if the model is correctly specified, then „prediction“ and „estimation“ are „nearly the same“.*

The actual problem is, that F^Y *is unknown* $\Longrightarrow$ *later.*

Definition

Let $P = \{\mathbb{P}_\theta \mid \theta \in \Theta\}$ *be a statistical model with the corresponding joint distributions* $\{F_\theta^{X,Y,\underline{Y},\overline{Y}} \mid \theta \in \Theta\}$ *and* $X, \underline{Y}, \overline{Y}$ *random variables with the joint distribution* $F^{X,\underline{Y},\overline{Y}}$*. The* ***Sharp Estimation Region (SER)*** *is defined as:*

Definition

Let $P = \{\mathbb{P}_\theta \mid \theta \in \Theta\}$ be a statistical model with the corresponding joint distributions $\{F_\theta^{X,Y,\underline{Y},\overline{Y}} \mid \theta \in \Theta\}$ and $X, \underline{Y}, \overline{Y}$ random variables with the joint distribution $F^{X,\underline{Y},\overline{Y}}$. The ***Sharp Estimation Region (SER)*** *is defined as:*

$$SER(\underline{Y}, \overline{Y}) := \{\theta \in \Theta \mid C(X, Y, \underline{Y}, \overline{Y}) = 1\}.$$

Definition

Let $P = \{\mathbb{P}_\theta \mid \theta \in \Theta\}$ be a statistical model with the corresponding joint distributions $\{F_\theta^{X,Y,\underline{Y},\overline{Y}} \mid \theta \in \Theta\}$ and $X, \underline{Y}, \overline{Y}$ random variables with the joint distribution $F^{X,\underline{Y},\overline{Y}}$. The ***Sharp Estimation Region (SER)*** *is defined as:*

$$SER(\underline{Y}, \overline{Y}) := \{\theta \in \Theta \mid C(X, Y, \underline{Y}, \overline{Y}) = 1\}.$$

If the model is correctly specified, this region can also be written as:

$$SER(\underline{Y}, \overline{Y}) = \underset{\theta \in \Theta}{\operatorname{argmin}} \left(\inf_{Y s.t. C(X,Y,\underline{Y},\overline{Y})=1} L\left(F_\theta, F^{X,Y,\overline{Y},\underline{Y}}\right)\right).$$

Definition

Let $P = \{\mathbb{P}_\theta \mid \theta \in \Theta\}$ *be a statistical model with the corresponding joint distributions* $\{F_\theta^{X,Y,\underline{Y},\overline{Y}} \mid \theta \in \Theta\}$ *and* $X, \underline{Y}, \overline{Y}$ *random variables with the joint distribution* $F^{X,\underline{Y},\overline{Y}}$. *The* ***Sharp Estimation Region (SER)*** *is defined as:*

$$SER(\underline{Y}, \overline{Y}) := \{\theta \in \Theta \mid C(X, Y, \underline{Y}, \overline{Y}) = 1\}.$$

If the model is correctly specified, this region can also be written as:

$$SER(\underline{Y}, \overline{Y}) = \underset{\theta \in \Theta}{\operatorname{argmin}} \left(\inf_{Y s.t. C(X,Y,\underline{Y},\overline{Y})=1} L\left(F_\theta, F^{X,Y,\overline{Y},\underline{Y}}\right)\right).$$

The ***Sharp Prediction Region (SPR)*** *is defined as:*

Definition

*Let $P = \{\mathbb{P}_\theta \mid \theta \in \Theta\}$ be a statistical model with the corresponding joint distributions $\{F_\theta^{X,Y,\underline{Y},\overline{Y}} \mid \theta \in \Theta\}$ and $X, \underline{Y}, \overline{Y}$ random variables with the joint distribution $F^{X,\underline{Y},\overline{Y}}$. The **Sharp Estimation Region (SER)** is defined as:*

$$SER(\underline{Y}, \overline{Y}) := \{\theta \in \Theta \mid C(X, Y, \underline{Y}, \overline{Y}) = 1\}.$$

If the model is correctly specified, this region can also be written as:

$$SER(\underline{Y}, \overline{Y}) = \operatorname*{argmin}_{\theta \in \Theta} \left(\inf_{Y s.t. C(X,Y,\underline{Y},\overline{Y})=1} L\left(F_\theta, F^{X,Y,\overline{Y},\underline{Y}}\right)\right).$$

*The **Sharp Prediction Region (SPR)** is defined as:*

$$SPR(\underline{Y}, \overline{Y}) := \left\{ \operatorname*{argmin}_{\theta \in \Theta} L\left(F_\theta, F^{X,Y,\overline{Y},\underline{Y}}\right) \mid Y \text{ s.t. } C(X, Y, \underline{Y}, \overline{Y}) = 1 \right\}.$$

Now: Linear Model

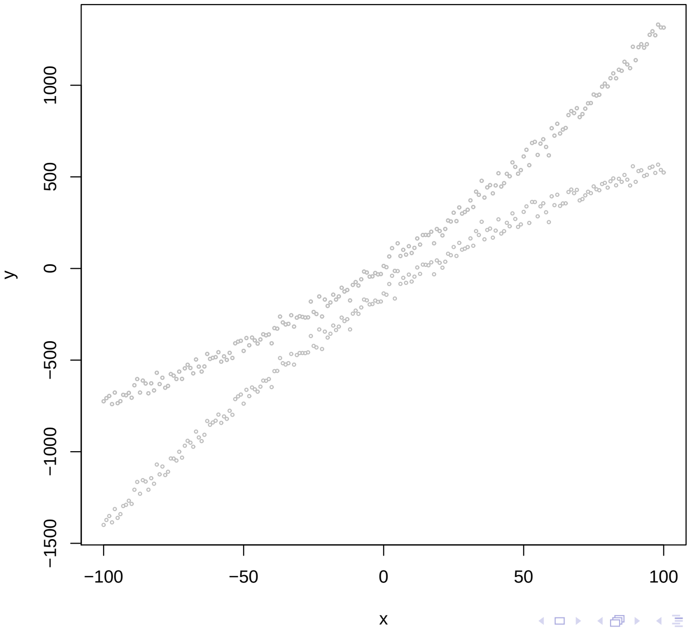
We are only interested in the components (β_0, β_1) of an element $\theta = ((\beta_0, \beta_1), \sigma^2, \sigma_l, \sigma_u) \in SER$ and denote the set

$$\{(\beta_0, \beta_1) \mid ((\beta_0, \beta_1), \sigma^2, \sigma_l, \sigma_u) \in SER\}$$

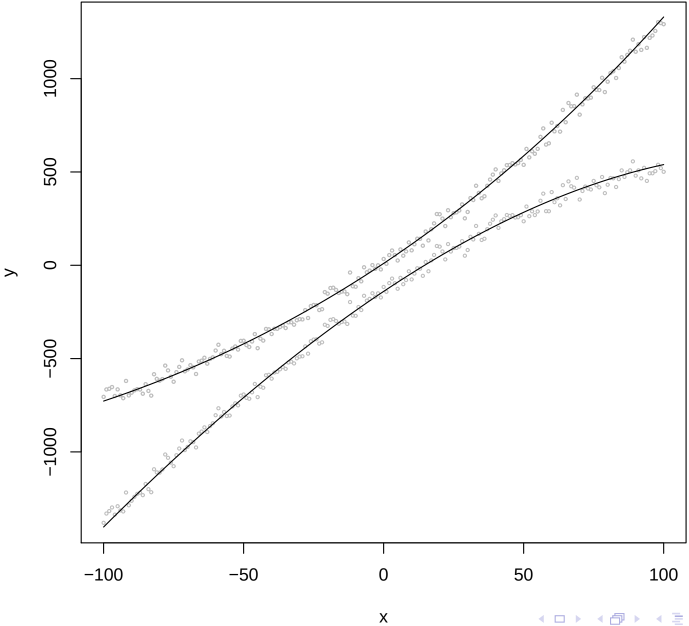
as the sharp estimation region (analogously for the sharp prediction region).

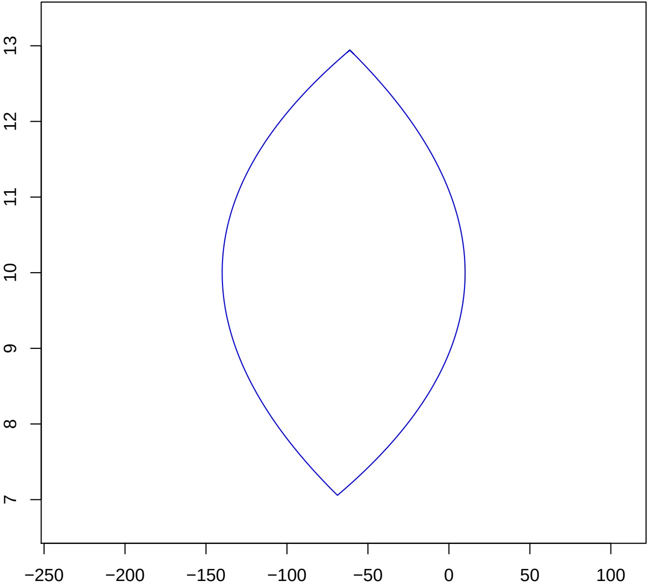
Linear Model

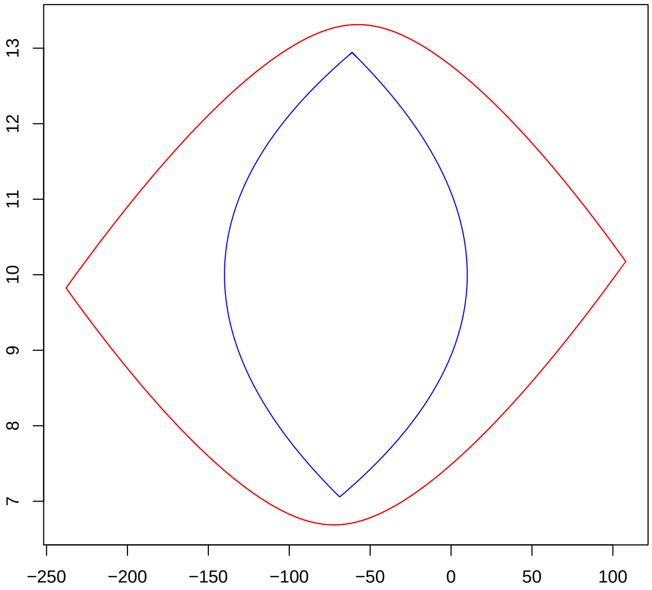
$$
\begin{aligned}
SER &= \{\beta \in B | \mathbb{E}(\underline{Y} \mid X) \leq X\beta \leq \mathbb{E}(\overline{Y} \mid X)\} \\
SPR &= \{\operatorname*{argmin}_{\beta \in B} \mathbb{E}((X\beta - Y)^2) \mid Y \in [\underline{Y}, \overline{Y}]\} \\
&= \{(X'X)^{-1}X'Y \mid Y \in [\underline{Y}, \overline{Y}]\}
\end{aligned}
$$

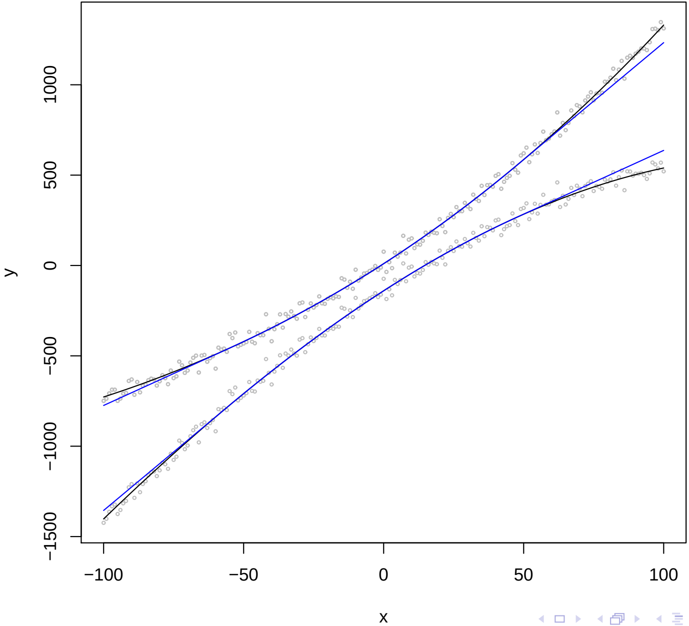


1000
500
0
−500
−1000
−1500
y
−100
−50
0
50
100
x

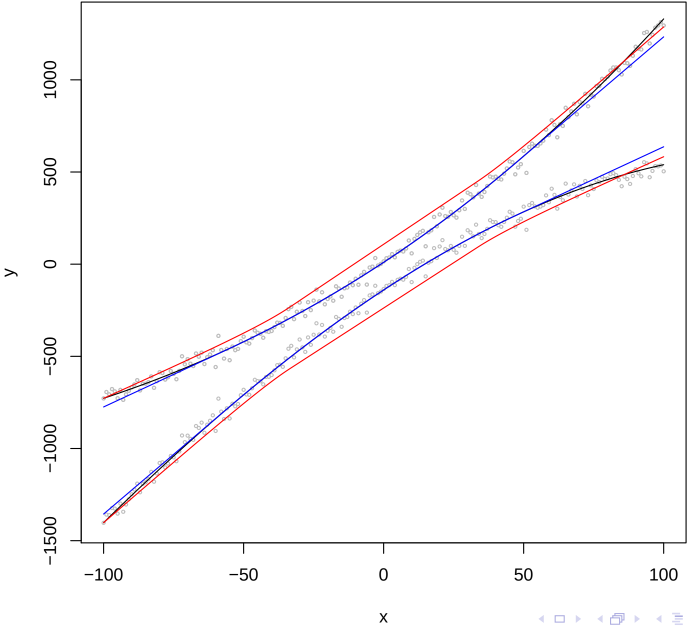

y
1000
500
0
−500
−1000
−100
−50
0
50
100
x


13
12
11
10
9
8
7
−250
−200
−150
−100
−50
0
50
100


13
12
11
10
9
8
7
−250
−200
−150
−100
−50
0
50
100



1000
500
0
−500
−1000
−1500
y
−100
−50
0
50
100
x


1000
500
0
−500
−1000
−1500
y
−100
−50
0
50
100
x

Theorem

Let $I \subset \mathbb{R}^2$ be a compact convex set. Then there exist random variables $X, \underline{Y}, \overline{Y}$ such that

$$SER(X, \underline{Y}, \overline{Y}) = I,$$

namely:

$$\begin{aligned} X &\sim N(0,1) \\ \underline{Y} &= \min\{\beta_0 + \beta_1 X \mid (\beta_0, \beta_1) \in I\} \\ \underline{Y} &= \max\{\beta_0 + \beta_1 X \mid (\beta_0, \beta_1) \in I\}. \end{aligned}$$

Definition

The Minkowski-Sum

$$M = \bigoplus_{i=1}^{n} l_i = \left\{ \sum_{i=1}^{n} p_i \mid p_i \in l_i \right\}$$

*of n line-segments $l_i \subseteq \mathbb{R}^d$ is called a **zonotope**.*
A zonotope is a convex, compact and centrally symmetric polytope with finite many extremepoints and central-symmetric facets.

Definition

The Minkowski-Sum

$$M = \bigoplus_{i=1}^{n} l_i = \left\{ \sum_{i=1}^{n} p_i \mid p_i \in l_i \right\}$$

of n line-segments $l_i \subseteq \mathbb{R}^d$ is called a ***zonotope****.*
A zonotope is a convex, compact and centrally symmetric polytope with finite many extremepoints and central-symmetric facets.

Definition

A closed, centrally symmetric convex set $Z \subseteq \mathbb{R}^d$ is called a ***zonoid****, if it can be approximated arbitrarily closely by zonotopes (w.r.t. a metric, e.g. the Hausdorff distance).*
For $d = 2$ the zonoids are exactly the closed, centrally symmetric convex sets.

Lemma

Let $I \subseteq \mathbb{R}^2$ be a zonoid in general position. Then there exists random variables $X, \underline{\mathrm{Y}}, \overline{\mathrm{Y}}$ such that

$$SPR(X, \underline{\mathrm{Y}}, \overline{\mathrm{Y}}) = I.$$

Lemma

Let $I \subseteq \mathbb{R}^2$ be a zonoid in general position. Then there exists random variables $X, \underline{Y}, \overline{Y}$ such that

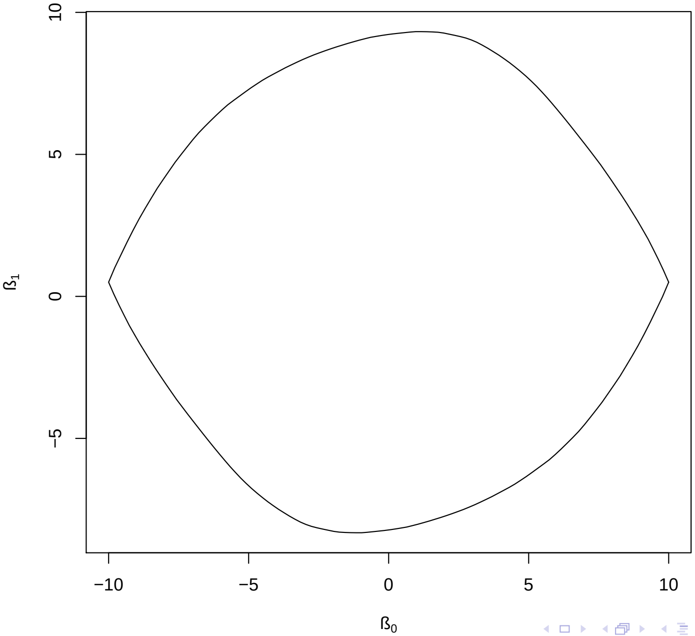
$$SPR(X, \underline{Y}, \overline{Y}) = I.$$

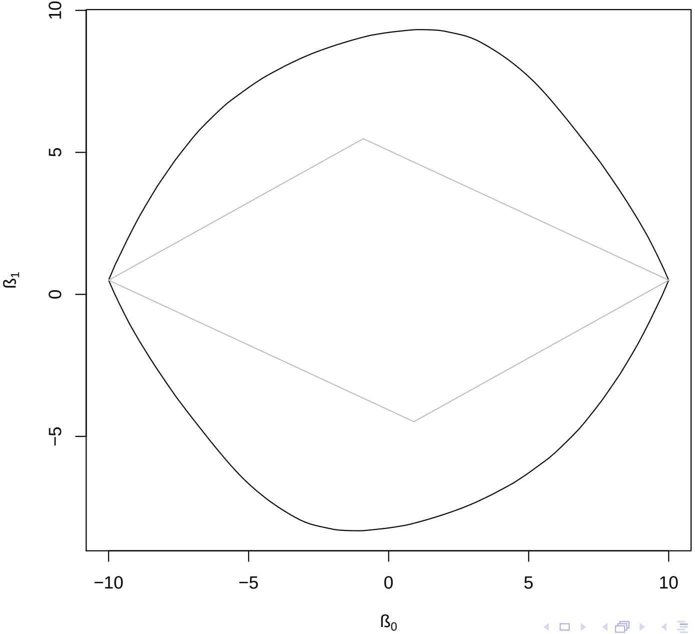
Lemma

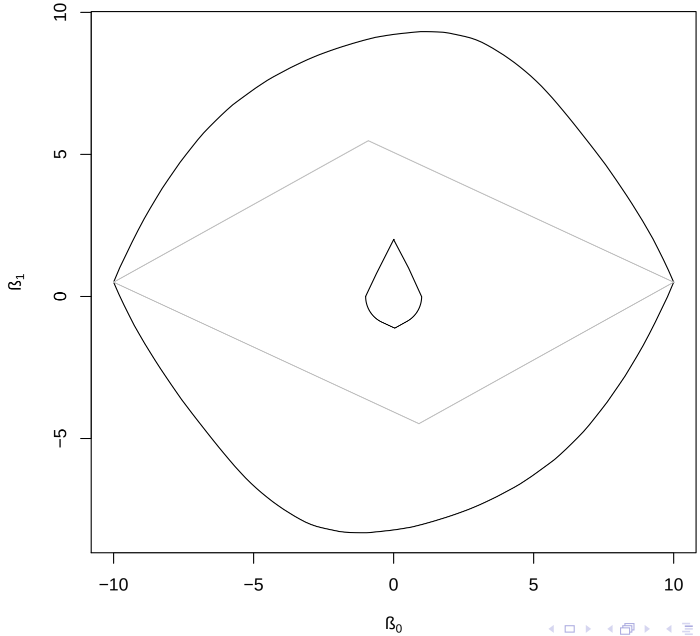
Let $I = SPR(X, \underline{Y}^*, \overline{Y}^*) \subseteq \mathbb{R}^2$ be a zonoid and $E \subseteq SER(X, \underline{Y}^*, \overline{Y}^*)$ an arbitrary compact convex set. Then for every $\varepsilon > 0$ there exist random variables $X, \underline{Y}, \overline{Y}$ such that:

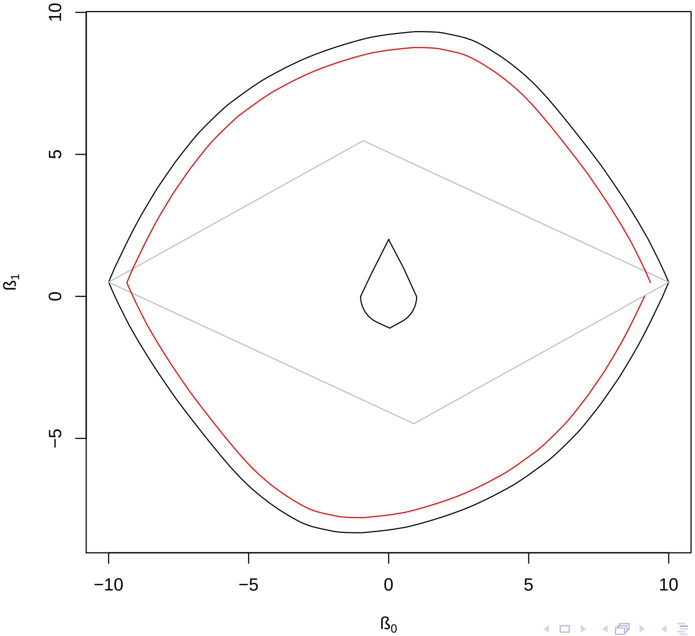
$$\begin{aligned} d_H(SPR(X, \underline{Y}, \overline{Y}), I) &\leq \varepsilon \\ d_H(SER(X, \underline{Y}, \overline{Y}), E) &\leq \varepsilon \end{aligned}$$

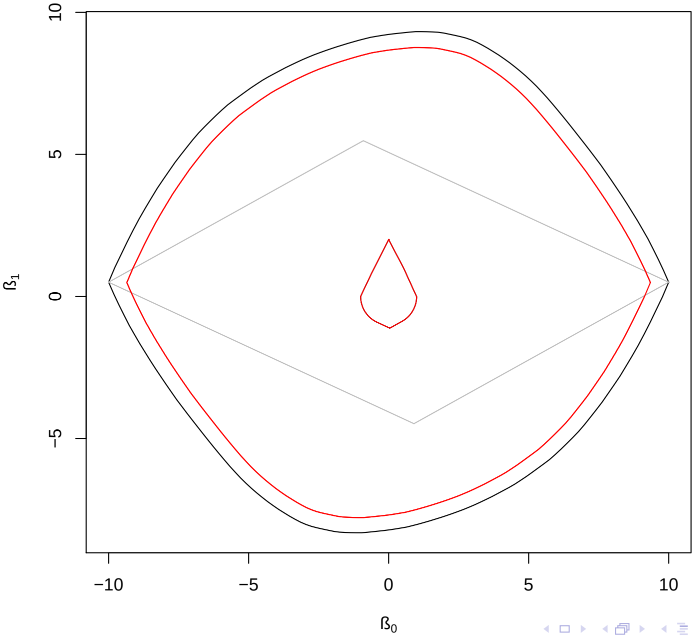
with the Hausdorff distance d_H.


10
5
0
−5
ß1
−10
−5
0
5
10
ß0


10
5
0
−5
ß$_1$
−10
−5
0
5
10
ß$_0$


10
5
0
−5
ß1
−10
−5
0
5
10
ß0


10
5
0
−5
ß1
−10
−5
0
5
10
ß0



10
5
0
−5
−10
−5
0
5
10
ß1
ß0

Mappings between ordered sets

Mappings between ordered sets

Definition

Let $(P, \leq)$ and $(Q, \sqsubseteq)$ be partially ordered sets. A pair (f, g) of mappings $f : P \longrightarrow Q$ and $g : Q \longrightarrow P$ is called ***adjunction****, if:*

$$\forall p \in P \forall q \in Q : \quad p \leq g(q) \iff f(p) \sqsubseteq q.$$

In this case, f is called ***left adjoint*** *and g is called* ***right adjoint****.*

Examples of adjunctions

Examples of adjunctions

- *Dempster-Shafer-Theory:*
  Multivalued mapping $\Gamma : X \longrightarrow 2^S$ *with corresponding*

  $\tilde{\Gamma} : (2^X, \subseteq) \longrightarrow (2^S, \subseteq) : A \mapsto \bigcup_{a \in A} \Gamma(a)$ *and the operator*

  $_* : (2^S, \subseteq) \longrightarrow (2^X, \subseteq) : T \mapsto \{x \in X \mid \Gamma(x) \subseteq T\}.$

  The pair $(\tilde{\Gamma}, {}_*)$ *is an adjunction.*
  From this, the ∞*-monotonicity of a Belief-function*

  $$Bel = P \circ {}_*$$

  with P *a probability-measure follows immediately, since* P *is* ∞*-monotone and* $_*$ *is meet-preserving. Furthermure it is clear, that also* $Bel \circ {}_*$ *is* ∞*-monotone.*

Examples of adjunctions

- *Lower coherent previsions:*
  $f : \underline{P} \mapsto \mathcal{M}(\underline{P}) = \{p \in \mathscr{P}(\Omega) \mid p \geq \underline{P}\}$ *and*

  $g : M \mapsto \underline{P}_M : X \mapsto \inf_{p \in M} p(X)$ *are an adjunction.*

Examples of adjunctions

- *Formal concept analysis:*
  Incidence structure $\mathbb{K} = (G, M, I)$
  with $G \ldots$ *objects,* $M \ldots$ *attributes and a relation* $I \subseteq G \times M$.
  $(g, m) \in I$ *means object* g *has attribute* m *(also denotad as* gIm*).*

  $f : (2^M, \subseteq) \longrightarrow (2^G, \subseteq) : X \mapsto \{g \in G | \forall m \in X : gIm\}$

  „The set of all objects having all attributes in X*"*

  $g : (2^G, \supseteq) \longrightarrow (2^M, \supseteq) : Y \mapsto \{m \in M \mid \forall g \in Y : gIm\}$

  „The set of all joint attributes of all objects in Y*".*

  The pair (f, g) *is an adjunction.*

Lemma

Let (f, g) be an adjunction. Then the following holds:

A1 $g \circ f$ *is extensive and* $f \circ g$ *is intensive.*

Lemma

Let (f, g) be an adjunction. Then the following holds:

A1 $g \circ f$ *is extensive and* $f \circ g$ *is intensive.*

A2 f *and* g *are order-preserving.*

Lemma

Let (f, g) be an adjunction. Then the following holds:

- A1 $g \circ f$ *is extensive and* $f \circ g$ *is intensive.*
- A2 f *and* g *are order-preserving.*
- A3 $f \circ g \circ f = f$ *and* $g \circ f \circ g = g$ *and thus* $f \circ g$ *and* $g \circ f$ *are idempotent.*

Lemma

Let (f, g) be an adjunction. Then the following holds:

- A1 $g \circ f$ *is extensive and* $f \circ g$ *is intensive.*
- A2 f *and* g *are order-preserving.*
- A3 $f \circ g \circ f = f$ *and* $g \circ f \circ g = g$ *and thus* $f \circ g$ *and* $g \circ f$ *are idempotent.*
- A4 *From A1 - A3 it follows, that* $g \circ f$ *is a hull operator and* $f \circ g$ *is a kernel operator.*

Lemma

Let (f, g) be an adjunction. Then the following holds:

A1 $g \circ f$ *is extensive and* $f \circ g$ *is intensive.*

A2 f *and* g *are order-preserving.*

A3 $f \circ g \circ f = f$ *and* $g \circ f \circ g = g$ *and thus* $f \circ g$ *and* $g \circ f$ *are idempotent.*

A4 *From A1 - A3 it follows, that* $g \circ f$ *is a hull operator and* $f \circ g$ *is a kernel operator.*

A5 *The adjoints* f *and* g *are determining each other unambiguously.*

Lemma

Let (f, g) be an adjunction. Then the following holds:

- A1 $g \circ f$ *is extensive and* $f \circ g$ *is intensive.*
- A2 f *and* g *are order-preserving.*
- A3 $f \circ g \circ f = f$ *and* $g \circ f \circ g = g$ *and thus* $f \circ g$ *and* $g \circ f$ *are idempotent.*
- A4 *From A1 - A3 it follows, that* $g \circ f$ *is a hull operator and* $f \circ g$ *is a kernel operator.*
- A5 *The adjoints* f *and* g *are determining each other unambiguously.*
- A6 f *is join-preserving and* g *is meet-preserving.*

Lemma

- *If P is a complete lattice, than f is a left adjoint, if and only if f is join-preserving.*
- *If Q is a complete lattice, than g is a right adjoint, if and only if g is meet-preserving.*

Lemma

The mapping

$$SER \quad : \quad (\mathcal{Z}(\Omega), \leq) \longrightarrow (2^B, \subseteq) : (X, \underline{Y}, \overline{Y}) \mapsto \{\beta \mid \mathbb{E}(\underline{Y}|X) \leq \beta X \leq \mathbb{E}(\overline{Y}|x)\}$$

with

$$\boxed{(X_1, \underline{Y}_1, \overline{Y}_1) \leq (X_2, \underline{Y}_2, \overline{Y}_2):} \quad \Longleftrightarrow \quad \boxed{\mathbb{E}(\underline{Y}_1 \mid X) \geq \mathbb{E}(\underline{Y}_2 \mid X) \ \&\ \mathbb{E}(\overline{Y}_1 \mid X) \leq \mathbb{E}(\overline{Y}_2 \mid X)}$$

i.e.: $(X_1, \underline{Y}_1, \overline{Y}_1)$ is more precise than $(X_2, \underline{Y}_2, \overline{Y}_2)$

is a right adjoint.

Lemma

The mapping

$$SER \quad : \quad (\mathcal{Z}(\Omega), \leq) \longrightarrow (2^B, \subseteq) : (X, \underline{Y}, \overline{Y}) \mapsto \{\beta \mid \mathbb{E}(\underline{Y}\,|X) \leq \beta X \leq \mathbb{E}(\overline{Y}\,|x)\}$$

with

$$\boxed{(X_1, \underline{Y}_1, \overline{Y}_1) \leq (X_2, \underline{Y}_2, \overline{Y}_2):} \quad \Longleftrightarrow \quad \boxed{\mathbb{E}(\underline{Y}_1 \mid X) \geq \mathbb{E}(\underline{Y}_2 \mid X) \;\&\; \mathbb{E}(\overline{Y}_1 \mid X) \leq \mathbb{E}(\overline{Y}_2 \mid X)}$$

i.e.: $(X_1, \underline{Y}_1, \overline{Y}_1)$ *is more precise than* $(X_2, \underline{Y}_2, \overline{Y}_2)$

is a right adjoint.
The corresponding left adjoint is the „prediction-operator":

$$PR : \qquad (2^B, \subseteq) \longrightarrow \left(\mathcal{Z}(\Omega), \leq) : M \mapsto (X, \min_{\beta \in M} X\beta, \max_{\beta \in M} X\beta\right).$$

Lemma

Thus, the following holds:

A1 $SER \circ PR$ *is extensive and* $PR \circ SER$ *is intensive.*

A2 PR *and* SER *are order-preserving.*

A3 $PR \circ SER \circ PR = PR$ *and* $SER \circ PR \circ SER = SER$ *and thus* $PR \circ SER$ *and* $SER \circ PR$ *are idempotent.*

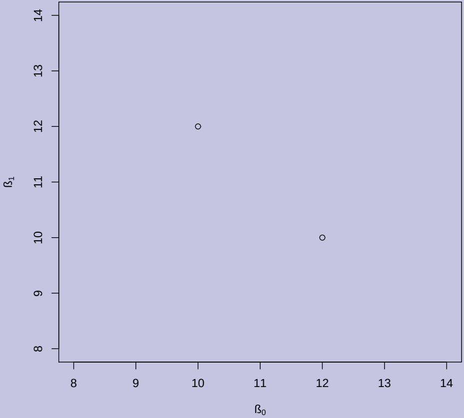
A4 *From A1 - A3 it follows, that* $SER \circ PR$ *is a hull operator and* $PR \circ SER$ *is a kernel operator.*

A5 *The adjoints* PR *and* SER *are determining each other unambiguously.*

A6 PR *is join-preserving and* SER *is meet-preserving.*

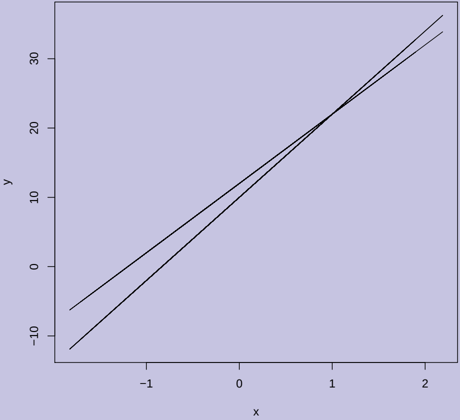
Lemma

A1 *$SER \circ PR$ is extensive and $PR \circ SER$ is intensive.*



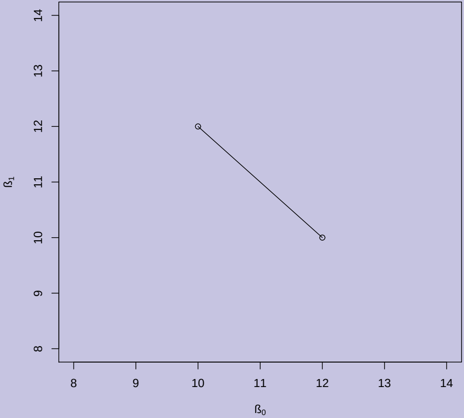
Lemma

A1 $SER \circ PR$ *is extensive and* $PR \circ SER$ *is intensive.*



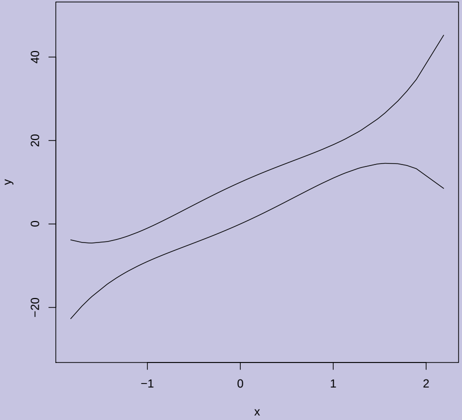
Lemma

A1 *$SER \circ PR$ is extensive and $PR \circ SER$ is intensive.*



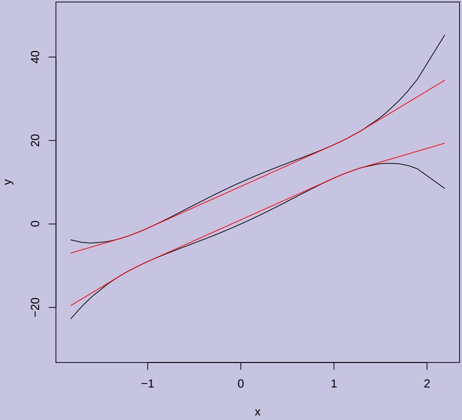
Lemma

A1 *SER* ∘ *PR is extensive and PR* ∘ *SER is intensive.*



Lemma

A1 *SER ∘ PR is extensive and PR ∘ SER is intensive.*



Lemma

Thus, the following holds:

A1 $SER \circ PR$ *is extensive and* $PR \circ SER$ *is intensive.*

A2 PR *and* SER *are order-preserving.*

A3 $PR \circ SER \circ PR = PR$ *and* $SER \circ PR \circ SER = SER$ *and thus* $PR \circ SER$ *and* $SER \circ PR$ *are idempotent.*

A4 *From A1 - A3 it follows, that* $SER \circ PR$ *is a hull operator and* $PR \circ SER$ *is a kernel operator.*

A5 *The adjoints* PR *and* SER *are determining each other unambiguously.*

A6 PR *is join-preserving and* SER *is meet-preserving.*

Lemma

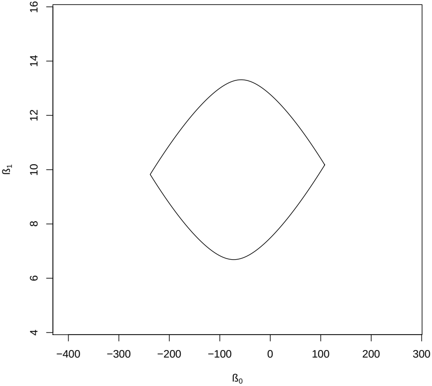
The mapping

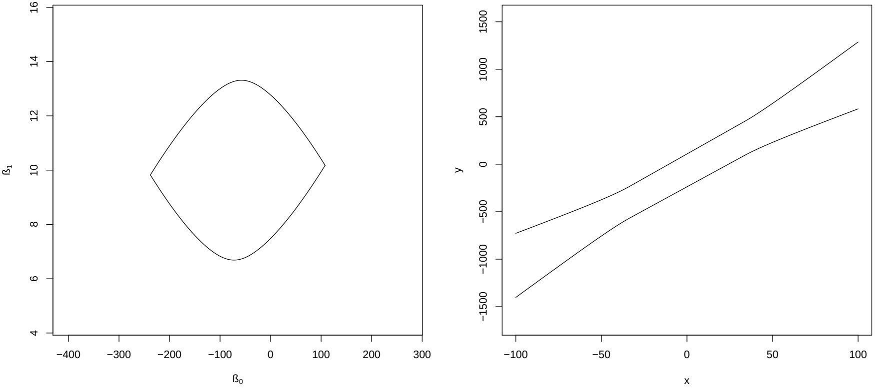
$$SPR : (\mathcal{Z}(\Omega), \leq) \longrightarrow (2^B, \subseteq) : (X, \underline{Y}, \overline{Y}) \mapsto \{(X'X)^{-1}X'Y \mid \underline{Y} \leq Y \leq \overline{Y}\}$$

is no right adjoint, since it is not meet-preserving.

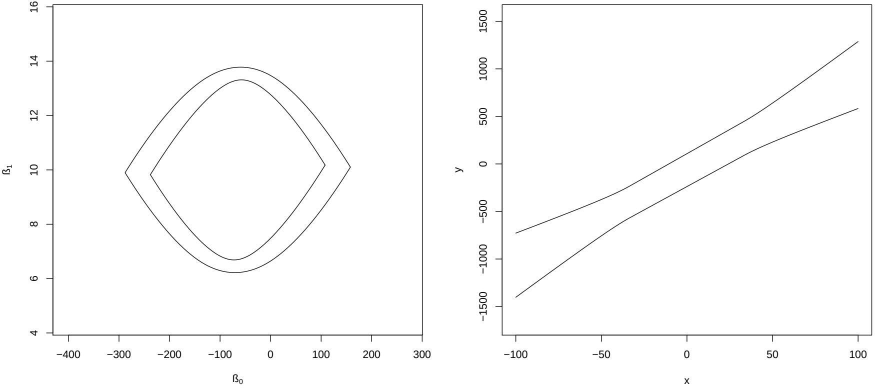
In general $SPR(Z_1 \wedge Z_2) \neq SPR(Z_1) \cap SPR(Z_2)$, *since the intersection of two zonoids is in general not a zonoid.*

Thus, in general, only $SPR \circ PR \circ SPR \supset SPR$ *holds.*

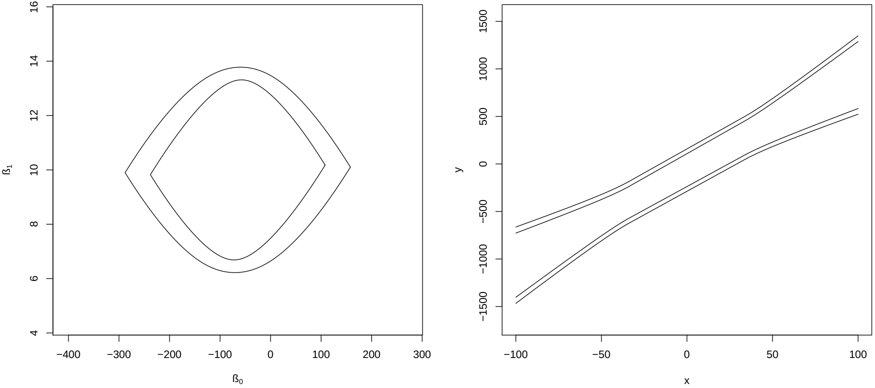

4
6
8
10
12
14
16
ß1
−400
−300
−200
−100
0
100
200
300
ß0



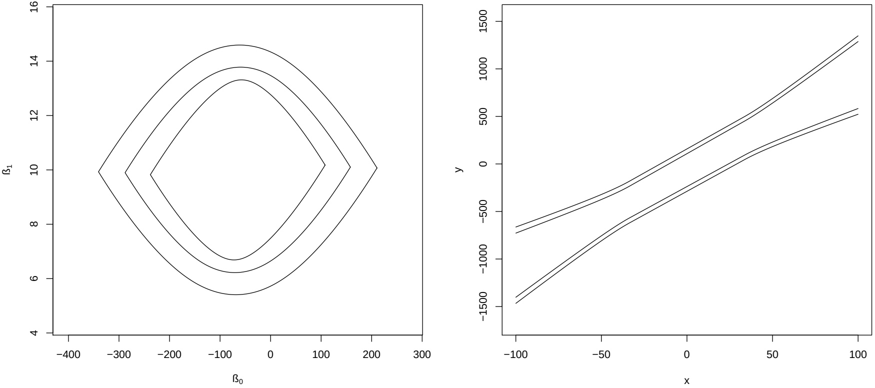
16
14
12
10
8
6
4
β_1
−400
−300
−200
−100
0
100
200
300
β_0
1500
1000
500
0
−500
−1000
−1500
y
−100
−50
0
50
100
x

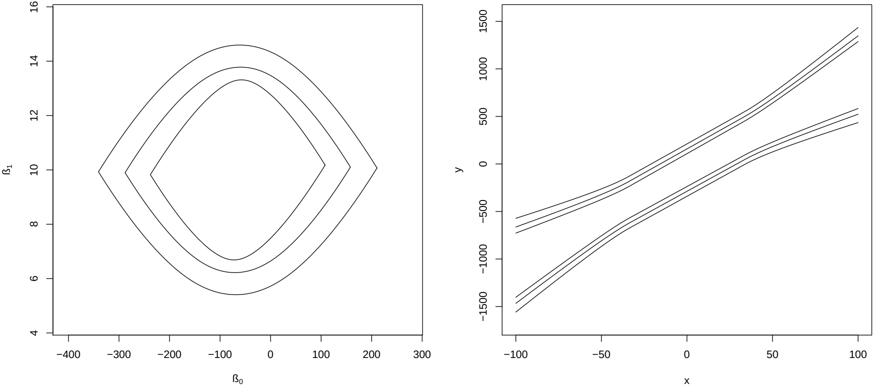


16
14
12
10
8
6
4
ß1
−400
−300
−200
−100
0
100
200
300
ß0
1500
1000
500
0
−500
−1000
−1500
y
−100
−50
0
50
100
x

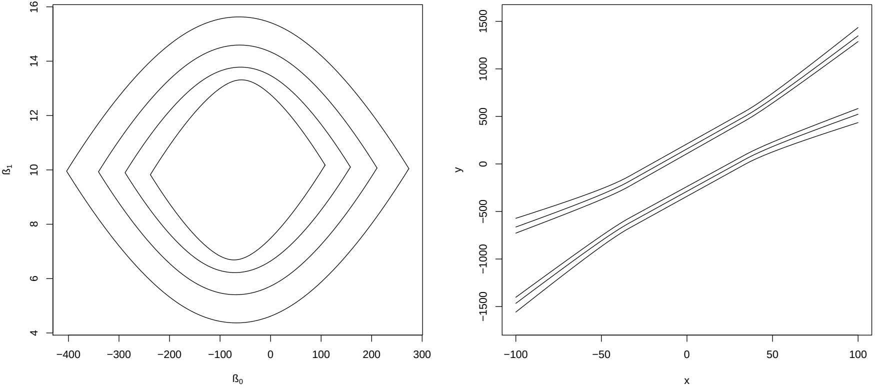


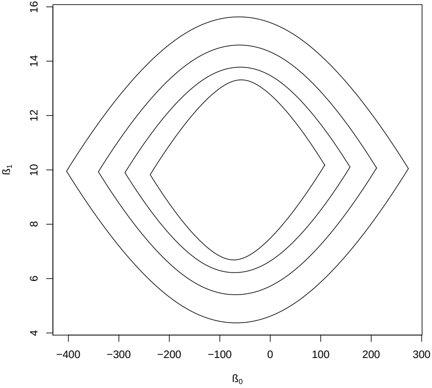
16
14
12
10
8
6
4
ß1
−400
−300
−200
−100
0
100
200
300
ß0
1500
1000
500
0
−500
−1000
−1500
y
−100
−50
0
50
100
x


16
14
12
10
8
6
4
ß1
−400
−300
−200
−100
0
100
200
300
ß0
1500
1000
500
0
−500
−1000
−1500
y
−100
−50
0
50
100
x

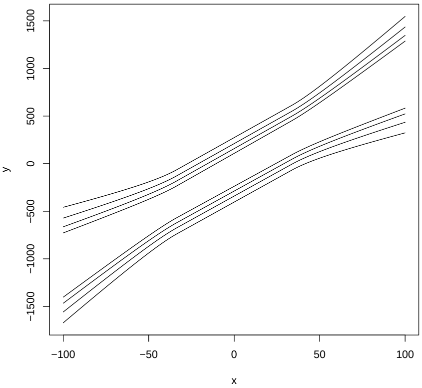


16
14
12
10
8
6
4
ß1
−400
−300
−200
−100
0
100
200
300
ß0
1500
1000
500
0
−500
−1000
−1500
y
−100
−50
0
50
100
x


16
14
12
10
8
6
4
$ß_1$
−400
−300
−200
−100
0
100
200
300
$ß_0$
1500
1000
500
0
−500
−1000
−1500
y
−100
−50
0
50
100
x



16
14
12
10
8
6
4
ß1
−400
−300
−200
−100
0
100
200
300
ß0



1500
1000
500
0
−500
−1000
−1500
y
−100
−50
0
50
100
x

Definition

Let $E : (P, \leq) \longrightarrow (Q, \sqsubseteq)$ *be a mapping.*
The monotone hull of E *is defined as:*

$$H(E) \quad : \quad (P, \leq) \longrightarrow (Q, \sqsubseteq) : X \mapsto \bigvee_{Y \leq X} E(Y).$$

The monotone kernel of E *is defined as:*

$$K(E) \quad : \quad (P, \leq) \longrightarrow (Q, \sqsubseteq) : X \mapsto \bigwedge_{Y \geq X} E(Y).$$

These set-valued mappings are both order-preserving
(i.e: $X \leq Y \Longrightarrow (H(E))(X) \sqsubseteq (H(E))(Y) \quad \& \quad (K(E))(X) \sqsubseteq (K(E))(Y)$*).*

A criterion-function-based mapping

A criterion-function-based mapping

Lemma

Let the criterion-function $Q : B \longrightarrow \mathbb{R}$ *be defined as:*

$$Q(\beta) = \int \left\{ (\mathbb{E}(\underline{Y}\,|x) - x\beta)_+^2 + \left(\mathbb{E}(\overline{Y}\,|x) - x\beta\right)_-^2 \right\} d\mathbb{P}(x).$$

A criterion-function-based mapping

Lemma

Let the criterion-function $Q : B \longrightarrow \mathbb{R}$ be defined as:

$$Q(\beta) = \int \left\{ (\mathbb{E}(\underline{Y}\,|x) - x\beta)_+^2 + \left(\mathbb{E}(\overline{Y}\,|x) - x\beta\right)_-^2 \right\} d\mathbb{P}(x).$$

Then the criterion-based mapping

$$E_Q : \mathcal{Z}(\Omega) \longrightarrow 2^B : (X, \underline{Y}, \overline{Y}) \mapsto \operatorname*{argmin}_{\beta \in B} Q(\beta)$$

is a source of SER and SPR:

A criterion-function-based mapping

Lemma

Let the criterion-function $Q : B \longrightarrow \mathbb{R}$ *be defined as:*

$$Q(\beta) = \int \left\{ \left(\mathbb{E}(\underline{Y}\,|x) - x\beta\right)_+^2 + \left(\mathbb{E}(\overline{Y}\,|x) - x\beta\right)_-^2 \right\} d\mathbb{P}(x).$$

Then the criterion-based mapping

$$E_Q : \mathcal{Z}(\Omega) \longrightarrow 2^B : (X, \underline{Y}, \overline{Y}) \mapsto \operatorname*{argmin}_{\beta \in B} Q(\beta)$$

is a source of SER and SPR:

$$\begin{aligned} SPR &= H(E_Q) \\ SER &= K(E_Q). \end{aligned}$$

Estimation of SER and SPR

Estimation of SER and SPR

Lemma

In general, there is no monotone, nonpartial, consistent estimator of SER.

Estimation of SER and SPR

Lemma

In general, there is no monotone, nonpartial, consistent estimator of SER.

Lemma

In general, there is no consistent and (in a certain sense) robust estimator of SER.

- Beresteanu, A., Molinari, F. (2008) Asymptotic Properties for a Class of Partially Identified Models, Econometrica, vol. 76, issue 4, pages 763-814.
- Chernozhukov, V., Hong, H., Tamer, E. (2007) Estimation and Confidence Regions for Parameter Sets in Econometric Models, Econometrica, vol. 75, issue 5, pages 1243-1284.
- Bolker, E.D. (1971) The Zonoid Problem, The American Mathematical Monthly, vol. 78, no. 5, pages 529-531