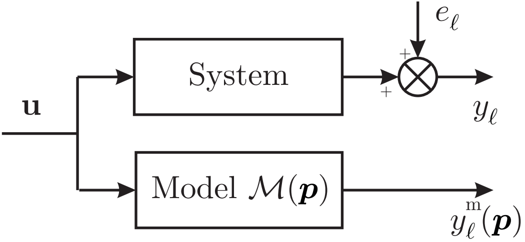
Guaranteed Characterization of Exact Non-Asymptotic Confidence Regions Using Interval Analysis

M. Kieffer

Université Paris-Saclay, CentraleSupélec, CNRS
Laboratoire des Signaux et Systèmes

24 november 2023

Introduction



Parameter identification: estimate value of parameter vector $\mathbf{p}$

- considering some model structure $\mathscr{M}(\cdot)$, with output $y_t^{\mathrm{m}}(\mathbf{p})$
- from noisy data vector $\mathbf{y} = (y_1, \ldots, y_n)^T$.

Introduction

Via minimization of cost function, for instance

$$J(\mathbf{p}) = \|\mathbf{y} - \mathbf{y}_{\mathrm{m}}(\mathbf{p})\|_2^2, \tag{1}$$

where

- $\mathbf{y}_{\mathrm{m}}(\mathbf{p}) = (y_1^{\mathrm{m}}(\mathbf{p}), \ldots, y_n^{\mathrm{m}}(\mathbf{p}))$ is vector of model outputs
- $\|\cdot\|_2$ is a (possibly weighted) ℓ_2 norm.

Then

$$\hat{\mathbf{p}} = \arg\min_{\mathbf{p}} J(\mathbf{p}). \tag{2}$$

Difficulties

Parameters of model may not be identifiable uniquely
↪ different values of $\hat{\mathbf{p}}$ may yield the same $\mathbf{y}_\mathrm{m}(\hat{\mathbf{p}})$

Numerical algorithm to compute $\hat{\mathbf{p}}$ may get trapped at local minimizer

Even if single $\hat{\mathbf{p}}$ is obtained and if $\mathbf{y} \simeq \mathbf{y}_\mathrm{m}(\hat{\mathbf{p}})$, $\hat{\mathbf{p}}$ cannot be considered as final answer to the estimation problem
↪ quality tag is missing.

$\hat{p}_i = 1.2345 \pm 10^{-4}$ is quite different of $\hat{p}_i = 1.2345 \pm 10^{3}$.

Difficulties

Parameters of model may not be identifiable uniquely
$\hookrightarrow$ different values of $\hat{\mathbf{p}}$ may yield the same $\mathbf{y}_{\mathrm{m}}(\hat{\mathbf{p}})$

Numerical algorithm to compute $\hat{\mathbf{p}}$ may get trapped at local minimizer

Even if single $\hat{\mathbf{p}}$ is obtained and if $\mathbf{y} \simeq \mathbf{y}_{\mathrm{m}}(\hat{\mathbf{p}})$, $\hat{\mathbf{p}}$ cannot be considered as final answer to the estimation problem
$\hookrightarrow$ quality tag is missing.

$$\hat{p}_i = 1.2345 \pm 10^{-4} \text{ is quite different of } \hat{p}_i = 1.2345 \pm 10^{3}.$$

Difficulties

Parameters of model may not be identifiable uniquely
$\hookrightarrow$ different values of $\hat{\mathbf{p}}$ may yield the same $\mathbf{y}_{\mathrm{m}}(\hat{\mathbf{p}})$

Numerical algorithm to compute $\hat{\mathbf{p}}$ may get trapped at local minimizer

Even if single $\hat{\mathbf{p}}$ is obtained and if $\mathbf{y} \simeq \mathbf{y}_{\mathrm{m}}(\hat{\mathbf{p}})$, $\hat{\mathbf{p}}$ cannot be considered as final answer to the estimation problem
$\hookrightarrow$ quality tag is missing.

$$\hat{p}_i = 1.2345 \pm 10^{-4} \text{ is quite different of } \hat{p}_i = 1.2345 \pm 10^{3}.$$

Outline

Classical approaches

Based on

- Level-set [WP97, SW03].
- Monte-Carlo techniques [WP97].
- Evaluation of the density of the estimator [Kay93].
- Bounded-error estimation [MNPLW96, JKDW01].

Characterization of parameter uncertainty via previous approaches relies on hypotheses on noise corrupting data

- difficult to verify from residuals $\mathbf{y} - \mathbf{y}_m(\hat{\mathbf{p}})$ when n_y is large,
- impossible to verify with only few data points.

Classical approaches

Based on

- Level-set [WP97, SW03].
- Monte-Carlo techniques [WP97].
- Evaluation of the density of the estimator [Kay93].
- Bounded-error estimation [MNPLW96, JKDW01].

Characterization of parameter uncertainty via previous approaches relies on hypotheses on noise corrupting data

- difficult to verify from residuals $\mathbf{y} - \mathbf{y}_{\mathrm{m}}(\hat{\mathbf{p}})$ when n_{y} is large,
- impossible to verify with only few data points.

White Gaussian noise with known variance

Assume that the prediction residuals

$$e^{\mathrm{p}}_{t_i}(\mathbf{p}^*) = y_{t_i} - y^{\mathrm{m}}_{t_i}(\mathbf{p}^*)$$

satisfy

$$e^{\mathrm{p}}_{t_i}(\mathbf{p}^*) = \varepsilon_{t_i}, \ \ i = 1, \dots, n_{\mathrm{t}}$$

with ε_{t_i} iid $\mathcal{N}\left(0, \sigma^2\right)$ with known σ^2.

Maximum likelihood estimation leads to minimization of

$$j(\mathbf{p}) = \sum_{i=1}^{n_{\mathrm{t}}} \left(e^{\mathrm{p}}_{t_i}(\mathbf{p})\right)^2.$$

White Gaussian noise with known variance

For the true value $\mathbf{p}^*$ of the vector of parameters

$$j(\mathbf{p}^*) = \sum_{i=1}^{n_\mathrm{t}} \left(e_{t_i}^{\mathrm{p}}(\mathbf{p}^*)\right)^2 = \sum_{i=1}^{n_\mathrm{t}} (\varepsilon(t_i))^2 \approx n_\mathrm{t}\sigma^2.$$

Useless to try to consider a criterion below $j_1 = n_\mathrm{t}\sigma^2$.

The larger σ^2, the higher the isocriterion to consider.

Noise raises the maximum acceptable value for the criterion.

White Gaussian noise with known variance

For the true value $\mathbf{p}^*$ of the vector of parameters

$$j(\mathbf{p}^*) = \sum_{i=1}^{n_\mathrm{t}} \left(e^{\mathrm{p}}_{t_i}(\mathbf{p}^*)\right)^2 = \sum_{i=1}^{n_\mathrm{t}} (\varepsilon(t_i))^2 \approx n_\mathrm{t}\sigma^2.$$

Useless to try to consider a criterion below $j_1 = n_\mathrm{t}\sigma^2$.

The larger σ^2, the higher the isocriterion to consider.

Noise raises the maximum acceptable value for the criterion.

White Gaussian noise with known variance

Prediction error $\mathbf{e}(\mathbf{p}) = \left(e^{\mathrm{p}}_{t_1}(\mathbf{p}), \ldots, e^{\mathrm{p}}_{t_{n_t}}(\mathbf{p})\right)^T$ evolves in space of dimension n_{t} and

$$j(\mathbf{p}^*)/\sigma^2 = \mathbf{e}^T(\mathbf{p}^*)\mathbf{e}(\mathbf{p}^*)/\sigma^2$$

distributed according to $\chi^2(n_{\mathrm{t}})$ law with n_{t} degrees of freedom.

Consider $X \sim \chi^2(n_{\mathrm{t}})$ and $\chi^2_\alpha(n_{\mathrm{t}})$ such that

$$\Pr\left(X \geqslant \chi^2_\alpha(n_{\mathrm{t}})\right) = \alpha$$

$\chi^2_\alpha(n_{\mathrm{t}})$ has a probability α to be exceeded by a random variable distributed according to $\chi^2(n_{\mathrm{t}})$ law.

Confidence region at $1-\alpha$ % is [SW03]

$$\mathbb{P}_{1-\alpha} = \left\{\mathbf{p} \,|\, j(\mathbf{p}) \leqslant \sigma^2\chi^2_\alpha(n_{\mathrm{t}})\right\}.$$

For example, $\alpha = 0.05$ leads to 95 % confidence region.

White Gaussian noise with known variance

Prediction error $\mathbf{e}(\mathbf{p}) = \left(e^{\mathrm{p}}_{t_1}(\mathbf{p}), \dots, e^{\mathrm{p}}_{t_{n_t}}(\mathbf{p})\right)^T$ evolves in space of dimension n_{t} and

$$j(\mathbf{p}^*)/\sigma^2 = \mathbf{e}^T(\mathbf{p}^*)\mathbf{e}(\mathbf{p}^*)/\sigma^2$$

distributed according to $\chi^2(n_{\mathrm{t}})$ law with n_{t} degrees of freedom.

Consider $X \sim \chi^2(n_{\mathrm{t}})$ and $\chi^2_\alpha(n_{\mathrm{t}})$ such that

$$\Pr\left(X \geqslant \chi^2_\alpha(n_{\mathrm{t}})\right) = \alpha$$

$\chi^2_\alpha(n_{\mathrm{t}})$ has a probability α to be exceeded by a random variable distributed according to $\chi^2(n_{\mathrm{t}})$ law.

Confidence region at $1-\alpha$ % is [SW03]

$$\mathbb{P}_{1-\alpha} = \left\{\mathbf{p} \,|\, j(\mathbf{p}) \leqslant \sigma^2\chi^2_\alpha(n_{\mathrm{t}})\right\}.$$

For example, $\alpha = 0.05$ leads to 95 % confidence region

White Gaussian noise with known variance

Prediction error $\mathbf{e}(\mathbf{p}) = \left(e^{\mathrm{p}}_{t_1}(\mathbf{p}), \ldots, e^{\mathrm{p}}_{t_{n_t}}(\mathbf{p})\right)^T$ evolves in space of dimension n_{t} and

$$j(\mathbf{p}^*)/\sigma^2 = \mathbf{e}^T(\mathbf{p}^*)\mathbf{e}(\mathbf{p}^*)/\sigma^2$$

distributed according to $\chi^2(n_{\mathrm{t}})$ law with n_{t} degrees of freedom.

Consider $X \sim \chi^2(n_{\mathrm{t}})$ and $\chi^2_\alpha(n_{\mathrm{t}})$ such that

$$\Pr\left(X \geqslant \chi^2_\alpha(n_{\mathrm{t}})\right) = \alpha$$

$\chi^2_\alpha(n_{\mathrm{t}})$ has a probability α to be exceeded by a random variable distributed according to $\chi^2(n_{\mathrm{t}})$ law.

Confidence region at $1-\alpha$ % is [SW03]

$$\mathbb{P}_{1-\alpha} = \left\{\mathbf{p} \,|\, j(\mathbf{p}) \leqslant \sigma^2\chi^2_\alpha(n_{\mathrm{t}})\right\}.$$

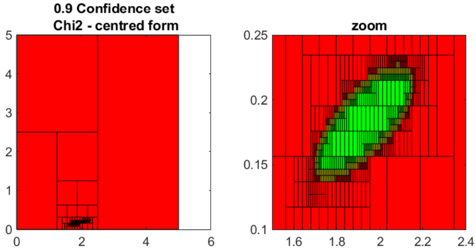
For example, $\alpha = 0.05$ leads to 95 % confidence region.

White Gaussian noise with known variance

Example:

$$y^{\mathrm{m}}_{t_i}(\mathbf{p}^*) = p_1^* \exp(-p_2^* t_i),$$

with $p^* = (2, 0.2)$ and $t_i = 0.1i$, $i = 0, \ldots, 63$.



Noise with unknown variance

Assume now that the prediction error

$$e^{\mathrm{p}}_{t_i}(\mathbf{p}^*) = y_{t_i} - y^{\mathrm{m}}_{t_i}(\mathbf{p}^*)$$

satisfies

$$e^{\mathrm{p}}_{t_i}(\mathbf{p}^*) = \varepsilon_{t_i},\ \ i = 1, \ldots, n_{\mathrm{t}}$$

where ε_{t_i}s are iid random variables $\mathcal{N}(0, \sigma^2)$ with σ^2 unknown.

Confidence region at $1-\alpha$ % cannot be defined as

$$\mathbb{P}_{1-\alpha} = \left\{\mathbf{p} \,|\, j(\mathbf{p}) \leqslant \sigma^2 \chi^2_{\alpha}(n_{\mathrm{t}})\right\}.$$

since σ^2 is unknown.

Noise with unknown variance

Assume now that the prediction error

$$e^{\mathrm{p}}_{t_i}(\mathbf{p}^*) = y_{t_i} - y^{\mathrm{m}}_{t_i}(\mathbf{p}^*)$$

satisfies

$$e^{\mathrm{p}}_{t_i}(\mathbf{p}^*) = \varepsilon_{t_i},\ \ i = 1,\ldots,n_{\mathrm{t}}$$

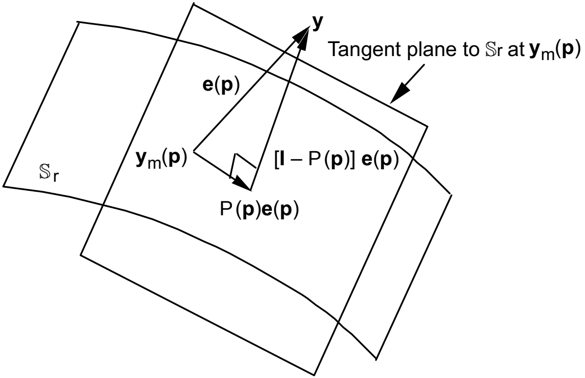
where ε_{t_i}s are iid random variables $\mathcal{N}\left(0,\sigma^2\right)$ with σ^2 unknown.

Confidence region at $1-\alpha$ % cannot be defined as

$$\mathbb{P}_{1-\alpha} = \left\{\mathbf{p} \,|\, j(\mathbf{p}) \leqslant \sigma^2 \chi^2_{\alpha}(n_{\mathrm{t}})\right\}.$$

since σ^2 is unknown.

Noise with unknown variance



In data space, $\mathbf{y}$ and $\mathbf{y}_m$ are points.

When $\mathbf{p}$ varies, $\mathbf{y}_m(\mathbf{p})$ describes surface response of model $\mathbb{S}_r$

- $\mathbb{S}_r$ hyperplane when model LP.
- $\mathbb{S}_r$ curved hypersurface when model NLP.

Noise with unknown variance

Consider

$$\boldsymbol{\Pi}(\mathbf{p}) = \frac{\partial \mathbf{y}_\mathrm{m}(\mathbf{p})}{\partial \mathbf{p}^T} \left(\left(\frac{\partial \mathbf{y}_\mathrm{m}(\mathbf{p})}{\partial \mathbf{p}^T} \right)^T \left(\frac{\partial \mathbf{y}_\mathrm{m}(\mathbf{p})}{\partial \mathbf{p}^T} \right) \right)^{-1} \left(\frac{\partial \mathbf{y}_\mathrm{m}(\mathbf{p})}{\partial \mathbf{p}^T} \right)^T$$

orthogonal projection matrix on hypersurface tangent to $\mathbb{S}_\mathrm{r}$ in $\mathbf{y}_\mathrm{m}(\mathbf{p})$

If $\dim(\mathbf{p}) = n_\mathrm{p}$, $\dim(\mathbf{y}) = n_\mathrm{t}$ and $\mathbf{e}(\mathbf{p}) = \mathbf{y} - \mathbf{y}_\mathrm{m}(\mathbf{p})$, then

- $j(\mathbf{p}^*) = \mathbf{e}^T(\mathbf{p}^*)\mathbf{e}(\mathbf{p}^*) \sim \sigma^2\chi^2(n_\mathrm{t})$, $\mathbf{e}(\mathbf{p})$ evolves in space of dimension n_t.
- $\mathbf{e}^T(\mathbf{p}^*)\boldsymbol{\Pi}(\mathbf{p}^*)\mathbf{e}(\mathbf{p}^*) \sim \sigma^2\chi^2(n_\mathrm{p})$, orthogonal projection of $\mathbf{e}(\mathbf{p})$ on tangent space evolves in space of dimension n_p.
- $\mathbf{e}^T(\mathbf{p}^*)(\mathbf{I}-\boldsymbol{\Pi}(\mathbf{p}^*))\mathbf{e}(\mathbf{p}^*) \sim \sigma^2\chi^2(n_\mathrm{t} - n_\mathrm{p})$, orthogonal complement of previous projection.

Noise with unknown variance

Consider

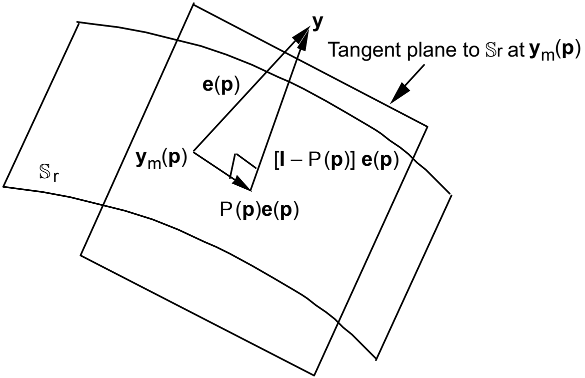
$$\mathbf{\Pi}(\mathbf{p}) = \frac{\partial \mathbf{y}_{\mathrm{m}}(\mathbf{p})}{\partial \mathbf{p}^T} \left(\left(\frac{\partial \mathbf{y}_{\mathrm{m}}(\mathbf{p})}{\partial \mathbf{p}^T} \right)^T \left(\frac{\partial \mathbf{y}_{\mathrm{m}}(\mathbf{p})}{\partial \mathbf{p}^T} \right) \right)^{-1} \left(\frac{\partial \mathbf{y}_{\mathrm{m}}(\mathbf{p})}{\partial \mathbf{p}^T} \right)^T$$

orthogonal projection matrix on hypersurface tangent to $\mathbb{S}_{\mathrm{r}}$ in $\mathbf{y}_{\mathrm{m}}(\mathbf{p})$

If $\dim(\mathbf{p}) = n_{\mathrm{p}}$, $\dim(\mathbf{y}) = n_{\mathrm{t}}$ and $\mathbf{e}(\mathbf{p}) = \mathbf{y} - \mathbf{y}_{\mathrm{m}}(\mathbf{p})$, then

- $j(\mathbf{p}^*) = \mathbf{e}^T(\mathbf{p}^*)\mathbf{e}(\mathbf{p}^*) \sim \sigma^2\chi^2(n_{\mathrm{t}})$, $\mathbf{e}(\mathbf{p})$ evolves in space of dimension n_{t}.
- $\mathbf{e}^T(\mathbf{p}^*)\mathbf{\Pi}(\mathbf{p}^*)\mathbf{e}(\mathbf{p}^*) \sim \sigma^2\chi^2(n_{\mathrm{p}})$, orthogonal projection of $\mathbf{e}(\mathbf{p})$ on tangent space evolves in space of dimension n_{p}.
- $\mathbf{e}^T(\mathbf{p}^*)(\mathbf{I}-\mathbf{\Pi}(\mathbf{p}^*))\mathbf{e}(\mathbf{p}^*) \sim \sigma^2\chi^2(n_{\mathrm{t}} - n_{\mathrm{p}})$, orthogonal complement of previous projection.

Noise with unknown variance



Moreover, $\mathbf{e}^T(\mathbf{p}^*)\mathbf{\Pi}(\mathbf{p}^*)\mathbf{e}(\mathbf{p}^*)$ and $\mathbf{e}^T(\mathbf{p}^*)(\mathbf{I}-\mathbf{\Pi}(\mathbf{p}^*))\mathbf{e}(\mathbf{p}^*)$ are independent, so that

$$\frac{\mathbf{e}^T(\mathbf{p}^*)\mathbf{\Pi}(\mathbf{p}^*)\mathbf{e}(\mathbf{p}^*)}{\mathbf{e}^T(\mathbf{p}^*)(\mathbf{I}-\mathbf{\Pi}(\mathbf{p}^*))\mathbf{e}(\mathbf{p}^*)}\,\frac{n_t - n_p}{n_p} \sim F(n_p, n_t - n_p)$$

where $F(n_p, n_t - n_p)$ Fisher-Snedecor law.

Noise with unknown variance

Computing quotient of two independent χ^2-distributed random variables, unknown σ^2 eliminated
$\hookrightarrow$ usable when σ^2 *a priori* unknown.

Consider $X \sim F(n_p, n_t - n_p)$ and $F_\alpha(n_p, n_t - n_p)$ such that

$$\Pr(X \geqslant F_\alpha(n_p, n_t - n_p)) = \alpha$$

$F_\alpha(n_p, n_t - n_p)$ is the value which has a probability α to be exceeded by a variable distributed according to a $F(n_p, n_t - n_p)$ law.

Confidence region with confidence level $1-\alpha$ % [SW03]

$$\mathbb{P}_{1-\alpha} = \left\{ \mathbf{p} \,\middle|\, \frac{\mathbf{e}^T(\mathbf{p})\boldsymbol{\Pi}(\mathbf{p})\mathbf{e}(\mathbf{p})}{\mathbf{e}^T(\mathbf{p})(\mathbf{I}-\boldsymbol{\Pi}(\mathbf{p}))\mathbf{e}(\mathbf{p})} \frac{n_t - n_p}{n_p} \leqslant F_\alpha(n_p, n_t - n_p) \right\}.$$

Much less amenable for characterization using interval analysis.

Noise with unknown variance

Computing quotient of two independent χ^2-distributed random variables, unknown σ^2 eliminated
$\hookrightarrow$ usable when σ^2 *a priori* unknown.

Consider $X \sim F(n_{\mathrm{p}}, n_{\mathrm{t}} - n_{\mathrm{p}})$ and $F_\alpha(n_{\mathrm{p}}, n_{\mathrm{t}} - n_{\mathrm{p}})$ such that

$$\Pr(X \geqslant F_\alpha(n_{\mathrm{p}}, n_{\mathrm{t}} - n_{\mathrm{p}})) = \alpha$$

$F_\alpha(n_{\mathrm{p}}, n_{\mathrm{t}} - n_{\mathrm{p}})$ is the value which has a probability α to be exceeded by a variable distributed according to a $F(n_{\mathrm{p}}, n_{\mathrm{t}} - n_{\mathrm{p}})$ law.

Confidence region with confidence level $1-\alpha$ % [SW03]

$$\mathbb{P}_{1-\alpha} = \left\{ \mathbf{p} \,\middle|\, \frac{\mathbf{e}^T(\mathbf{p})\boldsymbol{\Pi}(\mathbf{p})\mathbf{e}(\mathbf{p})}{\mathbf{e}^T(\mathbf{p})(\mathbf{I}-\boldsymbol{\Pi}(\mathbf{p}))\mathbf{e}(\mathbf{p})} \frac{n_{\mathrm{t}} - n_{\mathrm{p}}}{n_{\mathrm{p}}} \leqslant F_\alpha(n_{\mathrm{p}}, n_{\mathrm{t}} - n_{\mathrm{p}}) \right\}.$$

Much less amenable for characterization using interval analysis.

Noise with unknown variance

Computing quotient of two independent χ^2-distributed random variables, unknown σ^2 eliminated
$\hookrightarrow$ usable when σ^2 *a priori* unknown.

Consider $X \sim F(n_\mathrm{p}, n_\mathrm{t} - n_\mathrm{p})$ and $F_\alpha(n_\mathrm{p}, n_\mathrm{t} - n_\mathrm{p})$ such that

$$\Pr(X \geqslant F_\alpha(n_\mathrm{p}, n_\mathrm{t} - n_\mathrm{p})) = \alpha$$

$F_\alpha(n_\mathrm{p}, n_\mathrm{t} - n_\mathrm{p})$ is the value which has a probability α to be exceeded by a variable distributed according to a $F(n_\mathrm{p}, n_\mathrm{t} - n_\mathrm{p})$ law.

Confidence region with confidence level $1 - \alpha$ % [SW03]

$$\mathbb{P}_{1-\alpha} = \left\{ \mathbf{p} \,\middle|\, \frac{\mathbf{e}^T(\mathbf{p})\boldsymbol{\Pi}(\mathbf{p})\mathbf{e}(\mathbf{p})}{\mathbf{e}^T(\mathbf{p})(\mathbf{I}-\boldsymbol{\Pi}(\mathbf{p}))\mathbf{e}(\mathbf{p})} \frac{n_\mathrm{t} - n_\mathrm{p}}{n_\mathrm{p}} \leqslant F_\alpha(n_\mathrm{p}, n_\mathrm{t} - n_\mathrm{p}) \right\}.$$

Much less amenable for characterization using interval analysis.

SPS and LSCR

Campi *et al.* [CW05, DWC07, CCW12] propose two approaches named LSCR and SPS

- *exact* characterization of parameter uncertainty
- in *non-asymptotic* conditions.

Hypotheses

1. System generating data must belong to model set (true value $\mathbf{p}^*$ should be meaningful)
2. Noise samples must be independently distributed with distributions symmetric with respect to zero.

SPS and LSCR

Campi *et al.* [CW05, DWC07, CCW12] propose two approaches named LSCR and SPS

- *exact* characterization of parameter uncertainty
- in *non-asymptotic* conditions.

Hypotheses

1. System generating data must belong to model set (true value $\mathbf{p}^*$ should be meaningful)
2. Noise samples must be independently distributed with distributions symmetric with respect to zero.

Outline

SPS

Introduction

SPS [CCW12]: *sign-perturbed sums*.

SPS is designed for linear regression, where

$$y_t = \varphi_t^\top \mathbf{p}^* + w_t, t = 1, \ldots, n, \tag{3}$$

with φ_t known regression vector.

SPS defines an exact confidence region for $\mathbf{p}^*$ around least-squares estimate $\hat{\mathbf{p}}$, which is solution to *normal equations*

$$\sum_{t=1}^{n} \varphi_t \left(y_t - \varphi_t^\top \hat{\mathbf{p}} \right) = \mathbf{0}. \tag{4}$$

SPS

Introduction

SPS [CCW12]: *sign-perturbed sums*.

SPS is designed for linear regression, where

$$y_t = \varphi_t^\top \mathbf{p}^* + w_t, t = 1, \ldots, n, \quad (3)$$

with φ_t known regression vector.

SPS defines an exact confidence region for $\mathbf{p}^*$ around least-squares estimate $\hat{\mathbf{p}}$, which is solution to *normal equations*

$$\sum_{t=1}^{n} \varphi_t \left(y_t - \varphi_t^\top \hat{\mathbf{p}} \right) = \mathbf{0}. \quad (4)$$

SPS

Description

For generic $\mathbf{p}$ consider

$$\mathbf{s}_0(\mathbf{p}) = \sum_{t=1}^{n} \varphi_t \left(y_t - \varphi_t^{\mathsf{T}} \mathbf{p} \right), \tag{5}$$

and sign-perturbed sums

$$\mathbf{s}_i(\mathbf{p}) = \sum_{t=1}^{n} \alpha_{i,t} \varphi_t \left(y_t - \varphi_t^{\mathsf{T}} \mathbf{p} \right), \tag{6}$$

where $i = 1, \ldots, m-1$ and $\alpha_{i,t} = \pm 1$ with equal probability, and

$$z_i(\mathbf{p}) = \|\mathbf{s}_i(\mathbf{p})\|_2^2, i = 0, \ldots, m-1. \tag{7}$$

When ordering $z_i(\mathbf{p}^*)$ in increasing order, rank of $z_0(\mathbf{p}^*)$ is uniformly distributed [CCW12].

SPS

Description

For generic **p** consider

$$\mathbf{s}_0(\mathbf{p}) = \sum_{t=1}^{n} \varphi_t \left(y_t - \varphi_t^\mathsf{T} \mathbf{p} \right), \tag{5}$$

and sign-perturbed sums

$$\mathbf{s}_i(\mathbf{p}) = \sum_{t=1}^{n} \alpha_{i,t} \varphi_t \left(y_t - \varphi_t^\mathsf{T} \mathbf{p} \right), \tag{6}$$

where $i = 1, \ldots, m-1$ and $\alpha_{i,t} = \pm 1$ with equal probability, and

$$z_i(\mathbf{p}) = \|\mathbf{s}_i(\mathbf{p})\|_2^2, i = 0, \ldots, m-1. \tag{7}$$

When ordering $z_i(\mathbf{p}^*)$ in increasing order, rank of $z_0(\mathbf{p}^*)$ is uniformly distributed [CCW12].

SPS

Description

For generic $\mathbf{p}$ consider

$$\mathbf{s}_0(\mathbf{p}) = \sum_{t=1}^{n} \varphi_t \left(y_t - \varphi_t^{\mathsf{T}} \mathbf{p} \right), \tag{5}$$

and sign-perturbed sums

$$\mathbf{s}_i(\mathbf{p}) = \sum_{t=1}^{n} \alpha_{i,t} \varphi_t \left(y_t - \varphi_t^{\mathsf{T}} \mathbf{p} \right), \tag{6}$$

where $i = 1, \ldots, m-1$ and $\alpha_{i,t} = \pm 1$ with equal probability, and

$$z_i(\mathbf{p}) = \|\mathbf{s}_i(\mathbf{p})\|_2^2, i = 0, \ldots, m-1. \tag{7}$$

When ordering $z_i(\mathbf{p}^*)$ in increasing order, rank of $z_0(\mathbf{p}^*)$ is uniformly distributed [CCW12].

SPS

Description

Consider set $\boldsymbol{\Sigma}_q$ of all $\mathbf{p}$ such that $z_0(\mathbf{p})$ is *not* among the q largest values of $(z_i(\mathbf{p}))_{i=0}^{m-1}$.

One has $\mathbf{p}^* \in \boldsymbol{\Sigma}_q$ with exact probability $1 - q/m$, see [CCW12].

$\boldsymbol{\Sigma}_q$ is confidence region with level $1 - q/m$.

SPS

Description

Consider set $\mathbf{\Sigma}_q$ of all $\mathbf{p}$ such that $z_0(\mathbf{p})$ is *not* among the q largest values of $(z_i(\mathbf{p}))_{i=0}^{m-1}$.

One has $\mathbf{p}^* \in \mathbf{\Sigma}_q$ with exact probability $1 - q/m$, see [CCW12].

$\mathbf{\Sigma}_q$ is confidence region with level $1 - q/m$.

SPS

Definition

$\boldsymbol{\Sigma}_q$ may be defined more formally as

$$\boldsymbol{\Sigma}_q = \left\{ \mathbf{p} \in \mathbb{P} \text{ such that } \sum_{i=1}^{m-1} \tau_i(\mathbf{p}) \geqslant q \right\} \tag{8}$$

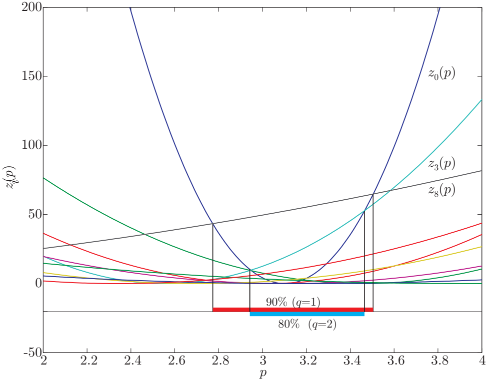
where

$$\tau_i(\mathbf{p}) = \begin{cases} 1 & \text{if } z_i(\mathbf{p}) - z_0(\mathbf{p}) > 0, \\ 0 & \text{else.} \end{cases} \tag{9}$$

SPS

Illustration

Model $y_t^{\mathrm{m}}(p) = p$, with 20 noisy data generated for $p^* = 3$.
We choose $m = 10$.



Outline

LSCR

Introduction - main idea

LSCR [CW05]: *leave-out sign-dominant correlated regions*

Independent estimates of the correlation of the prediction error

$$\varepsilon_t(\mathbf{p}) = y_t - y_t^{\mathrm{m}}(\mathbf{p})$$

should have random signs.

Leave out subset of parameter space where sign does not appear random (*i.e.* is sign dominant)

Defines, without any approximation,

region $\boldsymbol{\Theta}$ to which $\mathbf{p}^*$ belongs with specified probability.

LSCR

Introduction - main idea

LSCR [CW05]: *leave-out sign-dominant correlated regions*

Independent estimates of the correlation of the prediction error

$$\varepsilon_t(\mathbf{p}) = y_t - y_t^{\mathrm{m}}(\mathbf{p})$$

should have random signs.

Leave out subset of parameter space where sign does not appear random (*i.e.* is sign dominant)

Defines, without any approximation,

region Θ to which $\mathbf{p}^*$ belongs with specified probability.

LSCR

Introduction - main idea

LSCR [CW05]: *leave-out sign-dominant correlated regions*

Independent estimates of the correlation of the prediction error

$$\varepsilon_t(\mathbf{p}) = y_t - y_t^{\mathrm{m}}(\mathbf{p})$$

should have random signs.

Leave out subset of parameter space where sign does not appear random (*i.e.* is sign dominant)

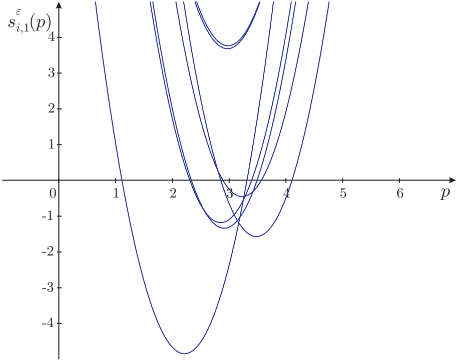
Defines, without any approximation,

region $\boldsymbol{\Theta}$ to which $\mathbf{p}^*$ belongs with specified probability.

LSCR

Example

Model $y_t^{\mathrm{m}}(p) = p$, with 8 noisy data generated with $p^* = 3$.



7 different empirical correlations as a function of p

LSCR

Description

Consider prediction error

$$\varepsilon_t(\mathbf{p}) = y_t - y_t^{\mathrm{m}}(\mathbf{p})$$

such that $\varepsilon_t(\mathbf{p}^*)$ is realization of noise corrupting data at time t.

- Select two integers $r \geqslant 0$ and $q \geqslant 0$.
- For $t = 1 + r, \ldots, k + r = n$, compute

$$c_{t-r,r}^{\varepsilon}(\mathbf{p}) = \varepsilon_{t-r}(\mathbf{p})\,\varepsilon_t(\mathbf{p}). \tag{10}$$

LSCR

Description

Consider prediction error

$$\varepsilon_t(\mathbf{p}) = y_t - y_t^{\mathrm{m}}(\mathbf{p})$$

such that $\varepsilon_t(\mathbf{p}^*)$ is realization of noise corrupting data at time t.

- Select two integers $r \geqslant 0$ and $q \geqslant 0$.
- For $t = 1 + r, \ldots, k + r = n$, compute

$$c_{t-r,r}^{\varepsilon}(\mathbf{p}) = \varepsilon_{t-r}(\mathbf{p})\,\varepsilon_t(\mathbf{p})\,. \tag{10}$$

LSCR

Description

- Compute

$$s^{\varepsilon}_{i,r}(\mathbf{p}) = \sum_{k \in \mathbb{I}_i} c^{\varepsilon}_{k,r}(\mathbf{p}), \; i = 1, \ldots, m. \tag{11}$$

  where $\mathbb{I}_i \subset \mathbb{I}$, set of indexes. Collection $\mathbb{G}$ of subsets $\mathbb{I}_i$, $i = 1, \ldots, m$, forms a group under the symmetric difference operation, *i.e.*, $(\mathbb{I}_i \cup \mathbb{I}_j) - (\mathbb{I}_i \cap \mathbb{I}_j) \in \mathbb{G}$.

  Then, from [CW05], the probability that less than q among the m $s^{\varepsilon}_{i,r}(\mathbf{p}^*)$s have different signs is exactly $2q/m$.

- Find $\mathbf{\Theta}^{\varepsilon}_{r,q}$ such that at least q of functions $s^{\varepsilon}_{i,r}(\mathbf{p})$ are larger than 0 and at least q are smaller than 0.

LSCR

Description

- Compute

$$s^{\varepsilon}_{i,r}(\mathbf{p}) = \sum_{k \in \mathbb{I}_i} c^{\varepsilon}_{k,r}(\mathbf{p}), \ i = 1, \dots, m. \tag{11}$$

  where $\mathbb{I}_i \subset \mathbb{I}$, set of indexes. Collection $\mathbb{G}$ of subsets $\mathbb{I}_i$, $i = 1, \dots, m$, forms a group under the symmetric difference operation, *i.e.*, $(\mathbb{I}_i \cup \mathbb{I}_j) - (\mathbb{I}_i \cap \mathbb{I}_j) \in \mathbb{G}$.

  Then, from [CW05], the probability that less than q among the m $s^{\varepsilon}_{i,r}(\mathbf{p}^*)$s have different signs is exactly $2q/m$.

- Find $\mathbf{\Theta}^{\varepsilon}_{r,q}$ such that at least q of functions $s^{\varepsilon}_{i,r}(\mathbf{p})$ are larger than 0 and at least q are smaller than 0.

LSCR

Description

Example of $\mathbb{G}$ st $\forall \mathbb{I}_i \in \mathbb{G}$,$\forall \mathbb{I}_i \in \mathbb{G}$ one has $(\mathbb{I}_i \cup \mathbb{I}_j) - (\mathbb{I}_i \cap \mathbb{I}_j) \in \mathbb{G}$

	1	2	3	4	5	6	7
$\mathbb{I}_1$	•	•		•	•		
$\mathbb{I}_2$	•		•	•		•	
$\mathbb{I}_3$		•	•		•	•	
$\mathbb{I}_4$	•	•				•	•
$\mathbb{I}_5$	•		•		•		•
$\mathbb{I}_6$		•	•	•			•
$\mathbb{I}_7$				•	•	•	•
$\mathbb{I}_8$							

$$s^{\varepsilon}_{1,1}(\mathbf{p}) = \varepsilon_1(\mathbf{p})\varepsilon_2(\mathbf{p}) + \varepsilon_2(\mathbf{p})\varepsilon_3(\mathbf{p}) + \varepsilon_3(\mathbf{p})\varepsilon_4(\mathbf{p}) + \varepsilon_4(\mathbf{p})\varepsilon_5(\mathbf{p}) + \varepsilon_5(\mathbf{p})\varepsilon_6(\mathbf{p}) + \varepsilon_6(\mathbf{p})\varepsilon_7(\mathbf{p}) + \varepsilon_7(\mathbf{p})\varepsilon_8(\mathbf{p})$$

$$s^{\varepsilon}_{2,1}(\mathbf{p}) = \varepsilon_1(\mathbf{p})\varepsilon_2(\mathbf{p}) + \varepsilon_2(\mathbf{p})\varepsilon_3(\mathbf{p}) + \varepsilon_3(\mathbf{p})\varepsilon_4(\mathbf{p}) + \varepsilon_4(\mathbf{p})\varepsilon_5(\mathbf{p}) + \varepsilon_5(\mathbf{p})\varepsilon_6(\mathbf{p}) + \varepsilon_6(\mathbf{p})\varepsilon_7(\mathbf{p}) + \varepsilon_7(\mathbf{p})\varepsilon_8(\mathbf{p})$$

$$\vdots$$

$$s^{\varepsilon}_{7,1}(\mathbf{p}) = \varepsilon_1(\mathbf{p})\varepsilon_2(\mathbf{p}) + \varepsilon_2(\mathbf{p})\varepsilon_3(\mathbf{p}) + \varepsilon_3(\mathbf{p})\varepsilon_4(\mathbf{p}) + \varepsilon_4(\mathbf{p})\varepsilon_5(\mathbf{p}) + \varepsilon_5(\mathbf{p})\varepsilon_6(\mathbf{p}) + \varepsilon_6(\mathbf{p})\varepsilon_7(\mathbf{p}) + \varepsilon_7(\mathbf{p})\varepsilon_8(\mathbf{p})$$

LSCR

Properties

The set $\boldsymbol{\Theta}^{\varepsilon}_{r,q}$ is such that [CW05]

$$\Pr\left(\mathbf{p}^* \in \boldsymbol{\Theta}^{\varepsilon}_{r,q}\right) = 1 - 2q/m.$$

Shape and size of $\boldsymbol{\Theta}^{\varepsilon}_{r,q}$ depend on

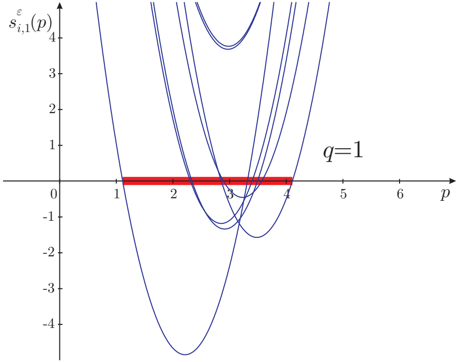
- values given to q and r
- group $\mathbb{G}$ and its number of elements m.

A procedure for generating $\mathbb{G}$ of appropriate size suggested in [Gor74].

LSCR

Example (continued)

Model $y_t^{\mathrm{m}}(p) = p$, with 8 noisy data generated with $p^* = 3$.



7 empirical correlations, and 71% confidence region

LSCR

More formal definition

The set $\mathbf{\Theta}^{\varepsilon}_{r,q}$ may be defined more formally as

$$\mathbf{\Theta}^{\varepsilon}_{r,q} = \mathbf{\Theta}^{\varepsilon,-}_{r,q} \cap \mathbf{\Theta}^{\varepsilon,+}_{r,q}, \tag{12}$$

with

$$\mathbf{\Theta}^{\varepsilon,-}_{r,q} = \left\{ \mathbf{p} \in \mathbb{P} \text{ such that } \sum_{i=1}^{m} \tau_i^{\varepsilon,-}(\mathbf{p}) \geqslant q \right\}, \tag{13}$$

$$\mathbf{\Theta}^{\varepsilon,+}_{r,q} = \left\{ \mathbf{p} \in \mathbb{P} \text{ such that } \sum_{i=1}^{m} \tau_i^{\varepsilon,+}(\mathbf{p}) \geqslant q \right\}, \tag{14}$$

where $\mathbb{P}$ is prior domain for $\mathbf{p}$.

LSCR

More formal definition

Moreover

$$\tau_i^{\varepsilon,-}(\mathbf{p}) = \begin{cases} 1 & \text{if } -s_{i,r}^{\varepsilon}(\mathbf{p}) \geqslant 0, \\ 0 & \text{else,} \end{cases} \tag{15}$$

and

$$\tau_i^{\varepsilon,+}(\mathbf{p}) = \begin{cases} 1 & \text{if } s_{i,r}^{\varepsilon}(\mathbf{p}) \geqslant 0, \\ 0 & \text{else.} \end{cases} \tag{16}$$

$\mathbf{\Theta}_{r,q}^{\varepsilon,-}$ contains all $\mathbf{p} \in \mathbb{P}$ such that at least q $s_{i,r}^{\varepsilon}(\mathbf{p})$s smaller than 0

$\mathbf{\Theta}_{r,q}^{\varepsilon,+}$ contains all $\mathbf{p} \in \mathbb{P}$ such that at least q $s_{i,r}^{\varepsilon}(\mathbf{p})$s larger than 0.

LSCR

More formal definition

Moreover

$$\tau_i^{\varepsilon,-}(\mathbf{p}) = \begin{cases} 1 & \text{if } -s_{i,r}^{\varepsilon}(\mathbf{p}) \geqslant 0, \\ 0 & \text{else,} \end{cases} \tag{15}$$

and

$$\tau_i^{\varepsilon,+}(\mathbf{p}) = \begin{cases} 1 & \text{if } s_{i,r}^{\varepsilon}(\mathbf{p}) \geqslant 0, \\ 0 & \text{else.} \end{cases} \tag{16}$$

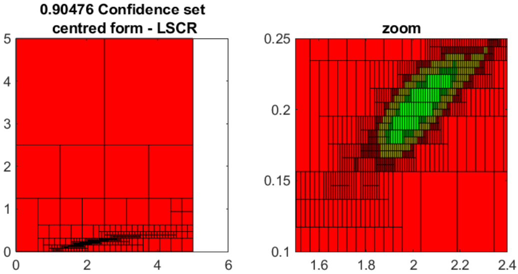
$\mathbf{\Theta}_{r,q}^{\varepsilon,-}$ contains all $\mathbf{p} \in \mathbb{P}$ such that at least q $s_{i,r}^{\varepsilon}(\mathbf{p})$s smaller than 0
$\mathbf{\Theta}_{r,q}^{\varepsilon,+}$ contains all $\mathbf{p} \in \mathbb{P}$ such that at least q $s_{i,r}^{\varepsilon}(\mathbf{p})$s larger than 0.

Example

Example:

$$y^{\mathrm{m}}_{t_i}(\mathbf{p}^*) = p_1^* \exp(-p_2^* t_i),$$

with $p^* = (2, 0.2)$ and $t_i = 0.1i$, $i = 0, \ldots, 63$.

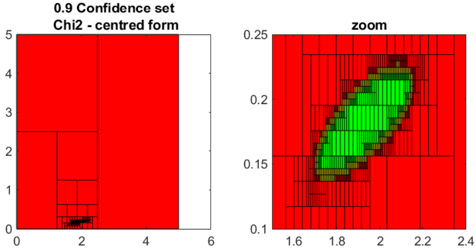


White Gaussian noise with known variance

Back to previous result (using χ^2 distribution:

$$y^{\mathrm{m}}_{t_i}(\mathbf{p}^*) = p_1^* \exp(-p_2^* t_i),$$

with $p^* = (2, 0.2)$ and $t_i = 0.1i$, $i = 0, \ldots, 63$.



Outline

Guaranteed characterization

In SPS (and LSCR), one has to characterize

$$\boldsymbol{\Psi}_q = \left\{ \mathbf{p} \in \mathbb{P} \text{ such that } \sum_{i=1}^{m} \tau_i(\mathbf{p}) \geqslant q \right\}, \tag{17}$$

where $\tau_i(\mathbf{p})$ is some indicator function

$$\tau_i(\mathbf{p}) = \begin{cases} 1 & \text{if } f_i(\mathbf{p}) \geqslant 0, \\ 0 & \text{else,} \end{cases} \tag{18}$$

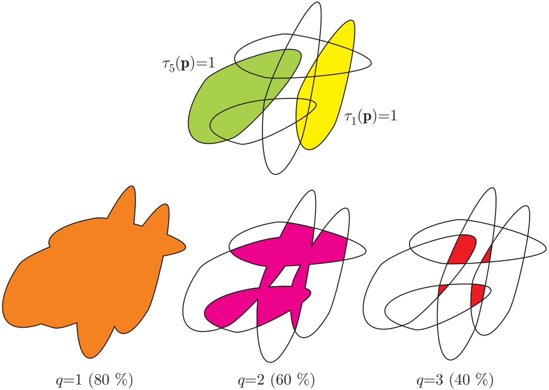
and where $f_i(\mathbf{p})$ depends on

- model structure,
- measurements,
- parameter vector $\mathbf{p}$.

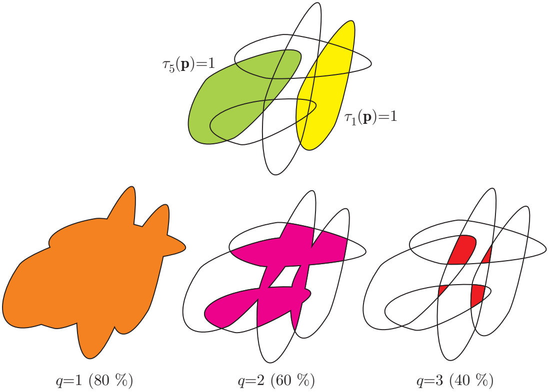
Guaranteed characterization

In SPS (and LSCR), one has to characterize

$$\boldsymbol{\Psi}_q = \left\{ \mathbf{p} \in \mathbb{P} \text{ such that } \sum_{i=1}^{m} \tau_i(\mathbf{p}) \geqslant q \right\}, \tag{19}$$



Guaranteed characterization



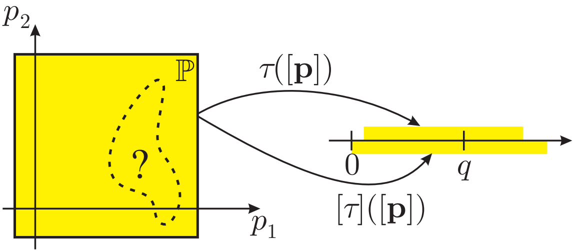
Characterization

- approximate using gridding in [CW05, DWC07, CCW12].
- guaranteed using interval analysis here [KW14].

SIVIA

To characterize $\mathbf{\Psi}_q = \{\mathbf{p} \in \mathbb{P} \text{ such that } \sum_{i=1}^{m} \tau_i(\mathbf{p}) \geqslant q\}$, one uses SIVIA and an inclusion function [Moo66, JKDW01] $[\tau]([\mathbf{p}])$ of

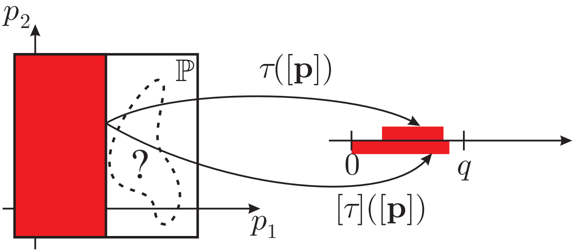
$$\tau(\mathbf{p}) = \sum_{i=1}^{m} \tau_i(\mathbf{p}).$$



SIVIA

To characterize $\boldsymbol{\Psi}_q = \{\mathbf{p} \in \mathbb{P} \text{ such that } \sum_{i=1}^m \tau_i(\mathbf{p}) \geqslant q\}$, one uses an inclusion function $[\tau]([\mathbf{p}])$ of

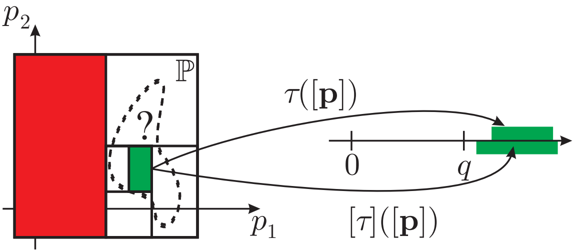
$$\tau(\mathbf{p}) = \sum_{i=1}^{m} \tau_i(\mathbf{p}).$$



SIVIA

To characterize $\mathbf{\Psi}_q = \{\mathbf{p} \in \mathbb{P} \text{ such that } \sum_{i=1}^{m} \tau_i(\mathbf{p}) \geqslant q\}$, one uses an inclusion function $[\tau]([\mathbf{p}])$ of

$$\tau(\mathbf{p}) = \sum_{i=1}^{m} \tau_i(\mathbf{p}).$$



Contractors may also be used, see [KW14].

Outline

Source localization problem

Problem encountered in

- indoor localization
- localization of electromagnetic source
- ...

Various approaches use measurements of wave emitted by object

- Time of arrival
- Difference of time of arrivals
- Received signal strength
- ...

Here, comparison of bounded-error approaches and LSCR [HKL18]

Source localization problem

Problem encountered in

- indoor localization
- localization of electromagnetic source
- ...

Various approaches use measurements of wave emitted by object

- Time of arrival
- Difference of time of arrivals
- Received signal strength
- ...

Here, comparison of bounded-error approaches and LSCR [HKL18]

Source localization problem

Problem encountered in

- indoor localization
- localization of electromagnetic source
- ...

Various approaches use measurements of wave emitted by object

- Time of arrival
- Difference of time of arrivals
- Received signal strength
- ...

Here, comparison of bounded-error approaches and LSCR [HKL18]

System model

n_a anchor nodes, with fixed and known locations θ_i, $i = 1, \ldots, n_a$,

Agent with unknown location θ_0

- emits electromagnetic/acoustic signal, received by anchors.
- $y(i,k)$: k-th RSS measurement by anchor node i

Anchor nodes transmit RSS measurements to central processing unit.

Confidence region for estimator of θ_0 to be derived from $y(i,k)$, $i = 1, \ldots, n_a$, $k = 1, \ldots, n$..

System model

n_a anchor nodes, with fixed and known locations θ_i, $i = 1, \ldots, n_a$,

Agent with unknown location θ_0

- emits electromagnetic/acoustic signal, received by anchors.
- $y(i,k)$: k-th RSS measurement by anchor node i

Anchor nodes transmit RSS measurements to central processing unit.

Confidence region for estimator of θ_0 to be derived from $y(i,k)$, $i = 1, \ldots, n_a$, $k = 1, \ldots, n$..

System model

n_a anchor nodes, with fixed and known locations θ_i, $i = 1, \dots, n_a$,

Agent with unknown location θ_0

- emits electromagnetic/acoustic signal, received by anchors.
- $y(i,k)$: k-th RSS measurement by anchor node i

Anchor nodes transmit RSS measurements to central processing unit.

Confidence region for estimator of θ_0 to be derived from $y(i,k)$, $i = 1, \dots, n_\mathrm{a}$, $k = 1, \dots, n$..

System model

k-th measurement by anchor node i described by Okumura-Hata model [Hat80]

$$y(i,k) = P_0 - 10\gamma_{\mathrm{P}} \log_{10} \frac{\|\theta_0 - \theta_i\|}{d_0} + \varepsilon(i,k),$$

where

- P_0 signal power at reference distance d_0,
- γ_{P} path-loss exponent,
- $\varepsilon(i,k)$ measurement noise.

One assumes that

- γ_{P} is the same for all anchors.
- $\varepsilon(i,k)$s, independently, not necessarily identically distributed random variables with distribution symmetric around zero.

System model

k-th measurement by anchor node i described by Okumura-Hata model [Hat80]

$$y(i,k) = P_0 - 10\gamma_{\mathrm{P}} \log_{10} \frac{\|\theta_0 - \theta_i\|}{d_0} + \varepsilon(i,k),$$

where

- P_0 signal power at reference distance d_0,
- γ_{P} path-loss exponent,
- $\varepsilon(i,k)$ measurement noise.

One assumes that

- γ_{P} is the same for all anchors.
- $\varepsilon(i,k)$s, independently, not necessarily identically distributed random variables with distribution symmetric around zero.

Parameters to estimate

Parameter vector

$$\mathbf{p} = \left[\theta_0^T, P_0, \gamma_{\mathrm{P}}\right]^T$$

True value $\mathbf{p}^*$ of parameter vector, then

$$y(i,k) = y^{\mathrm{m}}(i, \mathbf{p}^*) + \varepsilon(i,k) \tag{20}$$

with

$$y^{\mathrm{m}}(i, \mathbf{p}^*) = P_0^* - 10\gamma_{\mathrm{P}}^* \log_{10} \frac{\|\theta_0^* - \theta_i\|}{d_0}. \tag{21}$$

Search for $\mathbf{p}^*$ in $\mathbb{P}_0$.

Parameters to estimate

Parameter vector

$$\mathbf{p} = \left[\theta_0^T, P_0, \gamma_\mathrm{P}\right]^T$$

True value $\mathbf{p}^*$ of parameter vector, then

$$y(i,k) = y^\mathrm{m}(i,\mathbf{p}^*) + \varepsilon(i,k) \tag{20}$$

with

$$y^\mathrm{m}(i,\mathbf{p}^*) = P_0^* - 10\gamma_\mathrm{P}^* \log_{10} \frac{\|\theta_0^* - \theta_i\|}{d_0}. \tag{21}$$

Search for $\mathbf{p}^*$ in $\mathbb{P}_0$.

Parameters to estimate

Parameter vector

$$\mathbf{p} = \left[\theta_0^T, P_0, \gamma_\mathrm{P}\right]^T$$

True value $\mathbf{p}^*$ of parameter vector, then

$$y(i,k) = y^\mathrm{m}(i,\mathbf{p}^*) + \varepsilon(i,k) \quad (20)$$

with

$$y^\mathrm{m}(i,\mathbf{p}^*) = P_0^* - 10\gamma_\mathrm{P}^* \log_{10} \frac{\|\theta_0^* - \theta_i\|}{d_0}. \quad (21)$$

Search for $\mathbf{p}^*$ in $\mathbb{P}_0$.

Bounded-error estimation

Noise samples $\varepsilon(i,k)$ assumed bounded with known bounds

$$\varepsilon(i,k) \in [\underline{e}(i,k), \overline{e}(i,k)], \ i = 1 \dots n_a, k = 1 \dots n$$

Set of all $\mathbf{p} \in \mathbb{P}_0$ consistent with

- system model,
- measurements,
- noise bounds

defined as

$$\mathbb{P}_{\mathrm{BE}} = \{\mathbf{p} \in \mathbb{P}_0 | y_{\mathrm{m}}(i,\mathbf{p}) \in y(i,k) - [\underline{e}(i,k), \overline{e}(i,k)], i = 1 \dots n_a, k = 1 \dots n\}.$$

Difficulty:

How should $\underline{e}(i,k)$ and $\overline{e}(i,k)$ be chosen?

Bounded-error estimation

Noise samples $\varepsilon(i,k)$ assumed bounded with known bounds

$$\varepsilon(i,k) \in [\underline{e}(i,k), \overline{e}(i,k)], \ i = 1 \dots n_a, k = 1 \dots n$$

Set of all $\mathbf{p} \in \mathbb{P}_0$ consistent with

- system model,
- measurements,
- noise bounds

defined as

$$\mathbb{P}_{\text{BE}} = \{\mathbf{p} \in \mathbb{P}_0 | y_{\text{m}}(i,\mathbf{p}) \in y(i,k) - [\underline{e}(i,k), \overline{e}(i,k)], i = 1 \dots n_a, k = 1 \dots n\}.$$

Difficulty:

How should $\underline{e}(i,k)$ and $\overline{e}(i,k)$ be chosen?

Bounded-error estimation

Noise samples $\varepsilon(i,k)$ assumed bounded with known bounds

$$\varepsilon(i,k) \in [\underline{e}(i,k), \overline{e}(i,k)],\ i = 1 \ldots n_a, k = 1 \ldots n$$

Set of all $\mathbf{p} \in \mathbb{P}_0$ consistent with

- system model,
- measurements,
- noise bounds

defined as

$$\mathbb{P}_{\text{BE}} = \{\mathbf{p} \in \mathbb{P}_0 | y_{\text{m}}(i,\mathbf{p}) \in y(i,k) - [\underline{e}(i,k), \overline{e}(i,k)], i = 1 \ldots n_a, k = 1 \ldots n\}.$$

Difficulty:

How should $\underline{e}(i,k)$ and $\overline{e}(i,k)$ be chosen?

Robust bounded-error estimation

BE approaches may provide $\mathbb{P}_{\mathrm{BE}} = \emptyset$ as a result:

- noise bounds too optimistic
- inappropriate system model
- initial search box too small.

Robust bounded-error estimation

RBE estimation methods: find set of $\mathbf{p}$ consistent with all but q measurements and related noise bounds

$$\mathbb{P}_{\text{RBE},\xi} = \{\mathbf{p} \in \mathbb{P}_0 | \tau(\mathbf{p}) \in \mathbb{Y}_q\},$$

where

$$\tau(\mathbf{p}) = \sum_{i=1}^{n_a} \sum_{k=1}^{n} \tau_{i,k}(\mathbf{p}),$$

$$\tau_{i,k}(\mathbf{p}) = \begin{cases} 1 & y_{\text{m}}(i,\mathbf{p}) \in y(i,k) - [\underline{e}(i,k), \overline{e}(i,k)], \\ 0 & \text{else} \end{cases}$$

and $\mathbb{Y}_\xi = [n_a n - q, n_a n]$.

Problems:

- how should $\underline{e}(i,k)$ and $\overline{e}(i,k)$ be chosen?
- how should q be chosen?

Robust bounded-error estimation

RBE estimation methods: find set of **p** consistent with all but q measurements and related noise bounds

$$\mathbb{P}_{\mathrm{RBE},\xi} = \{\mathbf{p} \in \mathbb{P}_0 | \tau(\mathbf{p}) \in \mathbb{Y}_q\},$$

where

$$\tau(\mathbf{p}) = \sum_{i=1}^{n_a} \sum_{k=1}^{n} \tau_{i,k}(\mathbf{p}),$$

$$\tau_{i,k}(\mathbf{p}) = \begin{cases} 1 & y_{\mathrm{m}}(i,\mathbf{p}) \in y(i,k) - [\underline{e}(i,k), \overline{e}(i,k)], \\ 0 & \text{else} \end{cases}$$

and $\mathbb{Y}_\xi = [n_a n - q, n_a n]$.
Problems:

- how should $\underline{e}(i,k)$ and $\overline{e}(i,k)$ be chosen?
- how should q be chosen?

Simulation conditions

- Five anchor nodes ($n_a = 5$) are placed in the corners and in the center of a square of $20\,\mathrm{m} \times 20\,\mathrm{m}$.
- $N = 32$ agents are regularly placed in the square
- Each agent broadcasts $n = 10$ times message containing its identifier.
- $P_0 = 30$ dBm at $d_0 = 1$ m is the same for all agents.
- $\gamma_{\mathrm{P}} = 4$.

Simulation conditions

- Five anchor nodes ($n_a = 5$) are placed in the corners and in the center of a square of $20\,\mathrm{m} \times 20\,\mathrm{m}$.
- $N = 32$ agents are regularly placed in the square
- Each agent broadcasts $n = 10$ times message containing its identifier.
- $P_0 = 30$ dBm at $d_0 = 1$ m is the same for all agents.
- $\gamma_{\mathrm{P}} = 4$.

Simulation conditions

Data are corrupted by two types of noise samples

- iid zero-mean Gaussian noise with $\sigma_0 = 2$ dBm.
- iid Gaussian-Bernoulli-Gaussian variables
  - with a probability $p_0 = 0.9$, $\sigma_0 = 2$ dBm
  - with a probability $p_1 = 0.1$, $\sigma_1 = 5$ dBm.

Simulation conditions

Data are corrupted by two types of noise samples

- iid zero-mean Gaussian noise with $\sigma_0 = 2$ dBm.
- iid Gaussian-Bernoulli-Gaussian variables
    - with a probability $p_0 = 0.9$, $\sigma_0 = 2$ dBm
    - with a probability $p_1 = 0.1$, $\sigma_1 = 5$ dBm.

Simulation conditions

Three estimation problems are considered:

1. Only the location $\theta_{0,i}$, $i = 1, \ldots, N$ of each agent has to be estimated, γ and P_0 are assumed to be known.
2. $\theta_{0,i}$ and $P_{0,i}$, $i = 1, \ldots, N$ have to be determined for each agent.
3. $\theta_{0,i}$ and γ, $i = 1, \ldots, N$ have to be determined for each agent.

Simulation conditions

Three estimation problems are considered:

1. Only the location $\theta_{0,i}$, $i = 1, \ldots, N$ of each agent has to be estimated, γ and P_0 are assumed to be known.
2. $\theta_{0,i}$ and $P_{0,i}$, $i = 1, \ldots, N$ have to be determined for each agent.
3. $\theta_{0,i}$ and γ, $i = 1, \ldots, N$ have to be determined for each agent.

Simulation conditions

Three estimation problems are considered:

1. Only the location $\theta_{0,i}$, $i = 1, \ldots, N$ of each agent has to be estimated, γ and P_0 are assumed to be known.
2. $\theta_{0,i}$ and $P_{0,i}$, $i = 1, \ldots, N$ have to be determined for each agent.
3. $\theta_{0,i}$ and γ, $i = 1, \ldots, N$ have to be determined for each agent.

Selection of the parameters of LSCR

Different ways to organize the measurements are considered

$$\mathbf{y} = (y(1,1), y(1,2), y(1,3), y(2,1), y(2,2), y(2,3), \ldots, y(n_\mathrm{a},1), y(n_\mathrm{a},2), y(n_\mathrm{a},3))^T$$

or

$$\mathbf{y} = (y(1,1), y(2,1), \ldots, y(n_\mathrm{a},1), y(1,2), y(2,2), \ldots, y(n_\mathrm{a},2), y(1,3), y(2,3), \ldots, y(n_\mathrm{a},3))^T$$

Selection of the parameters of LSCR

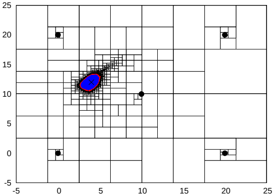
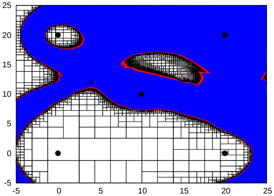
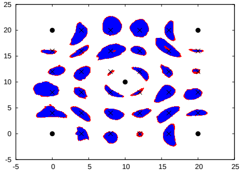
$\mathrm{y}=(y(1,1),y(1,2),y(1,3),y(2,1),y(2,2),y(2,3),\dots,y(n_{\mathrm{a}},1),y(n_{\mathrm{a}},2),y(n_{\mathrm{a}},3))^T$

$\mathrm{y}=(y(1,1),y(2,1),\dots,y(n_{\mathrm{a}},1),y(1,2),y(2,2),\dots,y(n_{\mathrm{a}},2),y(1,3),y(2,3),\dots,y(n_{\mathrm{a}},3))^T$


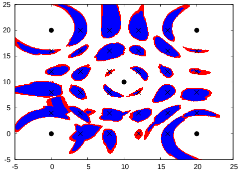
Table: Confidence regions as defined by LSCR obtained for different organizations of the measurement vector

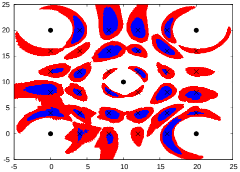
Comparison with alternative techniques

θ_0	θ_0, P_0	θ_0, γ_P

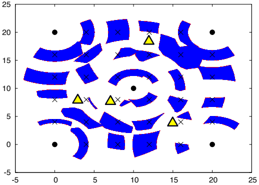
LSCR

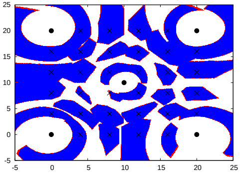






BE





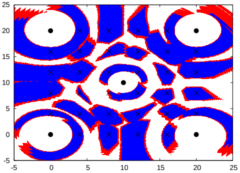


Table: Measurements corrupted by Gaussian noise

Comparison with alternative techniques

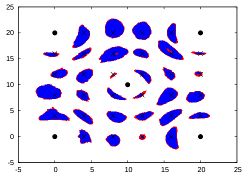
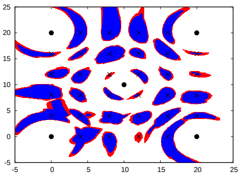
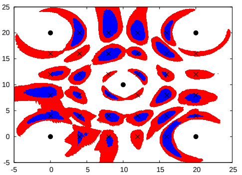
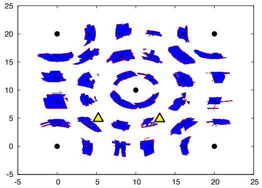
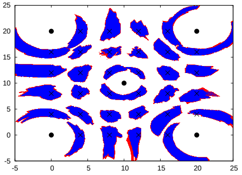
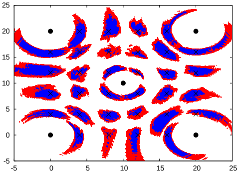
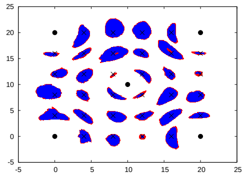
θ_0	θ_0, P_0	θ_0, γ_P
	LSCR	
		
	RBE	
		

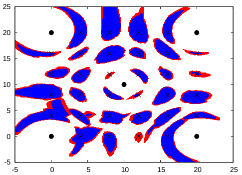
Table: Measurements corrupted by Gaussian noise

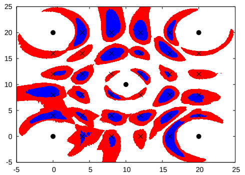
Comparison with alternative techniques

θ_0	θ_0, P_0	θ_0, γ_P

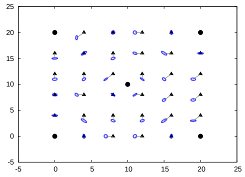
LSCR

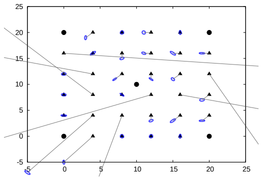






MLE





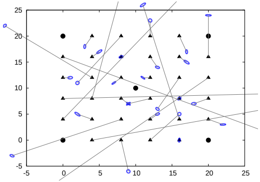


Table: Measurements corrupted by Gaussian noise

Comparison with alternative techniques

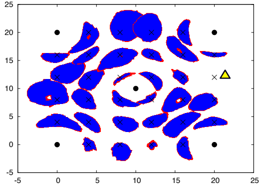
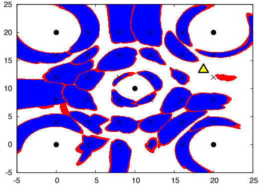
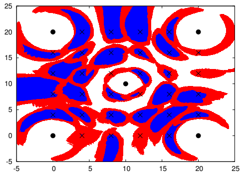
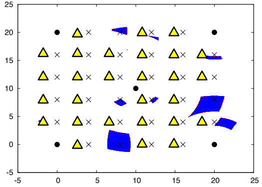
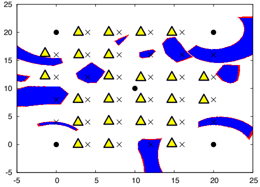
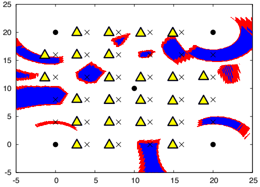
	θ_0	θ_0, P_0	θ_0, γ_P
LSCR			
BE			

Table: Measurements corrupted by Gaussian-Bernoulli-Gaussian noise

Comparison with alternative techniques

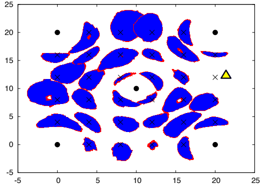
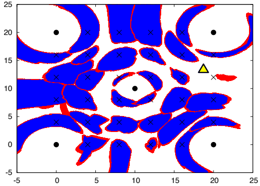
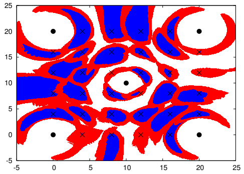
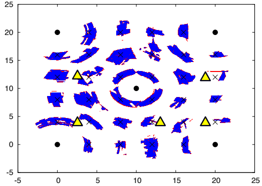
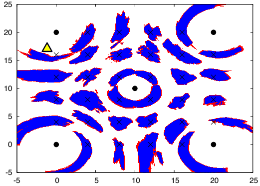
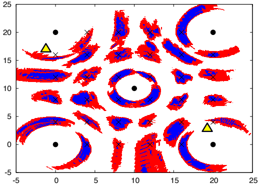
θ_0	θ_0, P_0	θ_0, γ_P
	LSCR	
		
	RBE	
		

Table: Measurements corrupted by Gaussian-Bernoulli-Gaussian noise

Comparison with alternative techniques

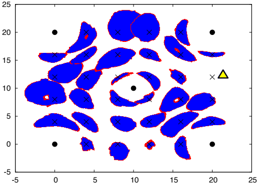
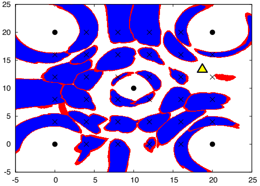
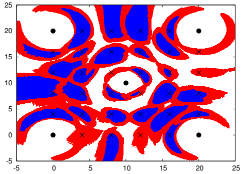
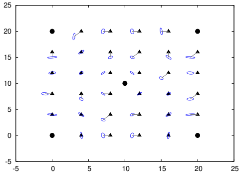
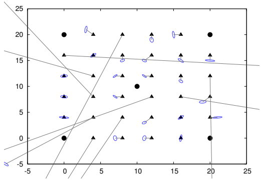
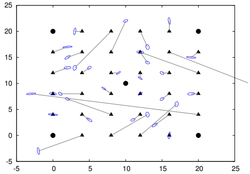
θ_0	θ_0, P_0	θ_0, γ_P
LSCR		
		
MLE		
		

Table: Simulations considering Gaussian-Bernoulli-Gaussian noise

Comparison with alternative techniques

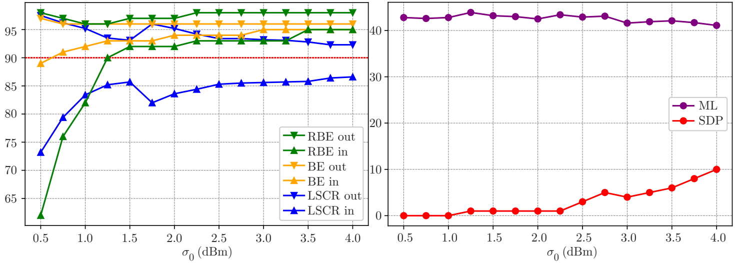


Figure: Proportions of agents for which the true value of the agent location is contained in the projection on the (θ_1, θ_2)-plane of the NACRs, the set estimates or the confidence region derived from the CRLB

Comparison with alternative techniques

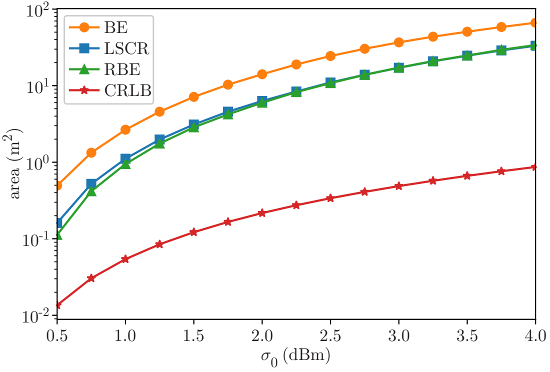


Figure: Evolution of the average surface of the projection on the (θ_1, θ_2)-plane of the NACRs, the set estimates or the confidence region derived from the CRLB

Comparison with alternative techniques

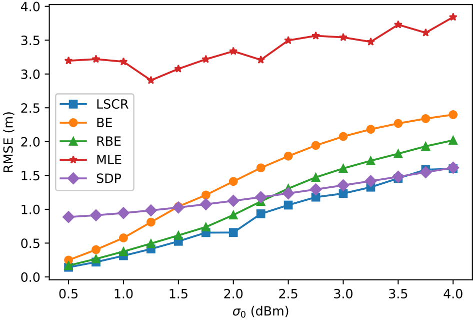


Figure: RMS localization error as a function of σ_0

Outline

Conclusions

Several ways to obtain non-asymptotic confidence regions

- Gaussian noise with known variance: adapted to characterization using IA
- Gaussian noise with unknown variance: difficult for IA

SPS

- Suited for model linear in the parameters
- Efficient characterization, efficient contractors, see [KW14]

LSCR

- Applies to linear and nonlinear models
- Efficient characterization, less efficient contractors, see [KW14]

Conclusions

Several ways to obtain non-asymptotic confidence regions

- Gaussian noise with known variance: adapted to characterization using IA
- Gaussian noise with unknown variance: difficult for IA

SPS

- Suited for model linear in the parameters
- Efficient characterization, efficient contractors, see [KW14]

LSCR

- Applies to linear and nonlinear models
- Efficient characterization, less efficient contractors, see [KW14]

Conclusions

Several ways to obtain non-asymptotic confidence regions

- Gaussian noise with known variance: adapted to characterization using IA
- Gaussian noise with unknown variance: difficult for IA

SPS

- Suited for model linear in the parameters
- Efficient characterization, efficient contractors, see [KW14]

LSCR

- Applies to linear and nonlinear models
- Efficient characterization, less efficient contractors, see [KW14]

References I

B. C. Csáji, M. C. Campi, and E. Weyer.
Non-asymptotic confidence regions for the least-squares estimate.
In *Proc. IFAC Symposium on System Identification*, pages 227–232, Brussels, Belgium, 2012.

M. C. Campi and E. Weyer.
Guaranteed non-asymptotic confidence regions in system identification.
Automatica, 41(10):1751–1764, 2005.

M. Dalai, E. Weyer, and M. C. Campi.
Parameter identification for nonlinear systems: Guaranteed confidence regions through LSCR.
Automatica, 43:1418 – 1425, 2007.

References II

L. Gordon.
Completely separating groups in subsampling.
Annals of Statistics, 2(3):572–578, 1974.

M. Hata.
Empirical formula for propagation loss in land mobile radio services.
IEEE Transactions on Vehicular Technology, 29(3):317–325, 1980.

C.Y. Han, M. Kieffer, and A. Lambert.
Guaranteed confidence region characterization for source localization using rss measurements.
Signal Processing, 2018.

References III

L. Jaulin.
Robust set membership state estimation; application to underwater robotics.
Automatica, 45(1):202–206, 2009.

L. Jaulin.
Set-membership localization with probabilistic errors.
Robotics and Autonomous Systems, 59(6):489–495, 2011.

L. Jaulin, M. Kieffer, O. Didrit, and E. Walter.
Applied Interval Analysis.
Springer-Verlag, London, 2001.

L. Jaulin and E. Walter.
Guaranteed robust nonlinear minimax estimation.
IEEE Transaction on Automatic Control, 47(11):1857–1864, 2002.

References IV

S. M. Kay.
Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory.
Prentice Hall, 1993.

M. Kieffer and E. Walter.
Guaranteed characterization of exact non-asymptotic confidence regions as defined by SPS and LSCR.
Automatica, 50(2):507–512, 2014.

M. Milanese, J. Norton, H. Piet-Lahanier, and E. Walter, editors.
Bounding Approaches to System Identification.
Plenum Press, New York, NY, 1996.

References V

R. E. Moore.
Interval Analysis.
Prentice-Hall, Englewood Cliffs, NJ, 1966.

G. A. F. Seber and C. J. Wild.
Nonlinear Regression.
Wiley-Interscience, 2003.

E. Walter and L. Pronzato.
Identification of Parametric Models from Experimental Data.
Springer-Verlag, London, 1997.

Contractors

Introduction

Contractor $\mathscr{C}_{\mathbf{f},\mathbb{Y}}$ associated with generic set-inversion problem

$$\mathbb{X} = [\mathbf{x}] \cap \mathbf{f}^{-1}(\mathbb{Y}), \tag{22}$$

takes $[\mathbf{x}]$ as input and returns

$$\mathscr{C}_{\mathbf{f},\mathbb{Y}}([\mathbf{x}]) \subset [\mathbf{x}] \tag{23}$$

such that

$$[\mathbf{x}] \cap \mathbb{X} = \mathscr{C}_{\mathbf{f},\mathbb{Y}}([\mathbf{x}]) \cap \mathbb{X}, \tag{24}$$

so no part of $\mathbb{X}$ in $[\mathbf{x}]$ is lost.

Contractors

Examples

Various types of contractors

- by interval constraint propagation,
- by parallel linearization,
- the Newton contractor,
- the Krawczyk contractor, *etc*.

Contractors

With LSCR and SPS

$$\mathbf{\Psi}_q = \mathbb{P} \cap \tau^{-1}([q, m]), \tag{25}$$

The τs are not differentiable and forbid use of classic contractors.

Proposed contractor assumes f_is differentiable.

1. build set of m possibly overlapping subboxes of $[\mathbf{p}]$, trying to remove all values of $\mathbf{p} \in [\mathbf{p}]$ such that $f_i(\mathbf{p}) < 0$, $i = 1, \ldots, m$.
2. compute union of all non-empty intersections of at least q of these boxes.

Box contraction using the f_i's, suitable for LSCR and SPS

First step: Centered inclusion function of f_i, for some $\mathbf{m} \in [\mathbf{p}]$,

$$[f_{i,\mathrm{c}}]([\mathbf{p}]) = f_i(\mathbf{m}) + ([\mathbf{p}] - \mathbf{m})^{\mathsf{T}} [\mathbf{g}_i]([\mathbf{p}]) \quad (26)$$

$$= f_i(\mathbf{m}) + \sum_{j=1}^{n_\mathrm{p}} ([p_j] - m_j) [g_{i,j}]([\mathbf{p}]), \quad (27)$$

where $\mathbf{g}_i$ is gradient of f_i.

Box contraction using the f_i's, suitable for LSCR and SPS

For k-th component $[p_k]$ of $[\mathbf{p}]$, when $0 \notin [g_{i,k}]([\mathbf{p}])$, $\mathscr{C}_{f_i,[0,\infty[}$ associates the contracted interval

$$\left[p'_{i,k}\right] = [p_k] \cap \left(\left(([f_{i,\mathrm{c}}]([\mathbf{p}]) \cap [0,\infty[) - f_i(\mathbf{m}) - \sum_{j=1, j\neq k}^{n_\mathrm{p}} ([p_j] - m_j)[g_{i,j}]([\mathbf{p}]) \right) / [g_{i,k}]([\mathbf{p}]) + m_k \right). \quad (28)$$

When $0 \in [g_{i,k}]([\mathbf{p}])$, $\mathscr{C}_{f_i,[0,\infty[}$ leaves $[p_k]$ unchanged, *i.e.*,

$$\left[p'_{i,k}\right] = [p_k]. \quad (29)$$

Box contraction using the f_i's, suitable for LSCR and SPS

Considering m functions f_i and applying all the contractors $\mathscr{C}_{f_i,[0,\infty[}$, $i = 1, \ldots, n$, to $[\mathbf{p}]$, one obtains

$$\mathscr{L} = \left\{\mathscr{C}_{f_1,[0,\infty[}([\mathbf{p}]), \ldots, \mathscr{C}_{f_m,[0,\infty[}([\mathbf{p}])\right\} \quad (30)$$

$$= \left\{[\mathbf{p}'_1], \ldots, [\mathbf{p}'_m]\right\}. \quad (31)$$

$[\mathbf{p}'_i] = \emptyset$ indicates that there is no $\mathbf{p} \in [\mathbf{p}]$ such that $f_i(\mathbf{p}) \geqslant 0$.

Box contraction suitable for SPS only

Main idea

Takes advantage of $\mathbf{s}_i(\mathbf{p})$, $i = 0, \ldots, m$ affine in $\mathbf{p}$ to

- reduce number of occurrences of $\mathbf{p}$ in $\mathbf{s}_i(\mathbf{p})$,
- reduce pessimism of corresponding inclusion functions.

Box contraction suitable for SPS only

One may rewrite

$$\mathbf{s}_0(\mathbf{p}) = \sum_{t=1}^{n} \varphi_t \left(y_t - \varphi_t^{\mathsf{T}} \mathbf{p} \right),$$

as

$$\begin{aligned} \mathbf{s}_0(\mathbf{p}) &= \sum_{t=1}^{n} y_t \varphi_t - \left(\sum_{t=1}^{n} \varphi_t \varphi_t^{\mathsf{T}} \right) \mathbf{p} && (32) \\ &= \mathbf{b}_0 - \mathbf{A}_0 \mathbf{p} && (33) \end{aligned}$$

with $\mathbf{b}_0 = \sum_{t=1}^{n} y_t \varphi_t$ and $\mathbf{A}_0 = \sum_{t=1}^{n} \varphi_t \varphi_t^{\mathsf{T}}$.

Box contraction suitable for SPS only

Similarly,

$$\mathbf{s}_i(\mathbf{p}) = \sum_{t=1}^{n} \alpha_{i,t} \varphi_t \left(y_t - \varphi_t^\mathsf{T} \mathbf{p} \right), \tag{34}$$

may be rewritten as

$$\mathbf{s}_i(\mathbf{p}) = \mathbf{b}_i - \mathbf{A}_i \mathbf{p} \tag{35}$$

with $\mathbf{b}_i = \sum_{t=1}^{n} \alpha_{i,t} y_t \varphi_t$ and $\mathbf{A}_i = \sum_{t=1}^{n} \alpha_{i,t} \varphi_t \varphi_t^\mathsf{T}$.

Box contraction suitable for SPS only

One gets then

$$\begin{aligned} z_i(\mathbf{p}) - z_0(\mathbf{p}) &= (\mathbf{b}_i - \mathbf{A}_i\mathbf{p})^{\mathsf{T}}(\mathbf{b}_i - \mathbf{A}_i\mathbf{p}) \\ &- (\mathbf{b}_0 - \mathbf{A}_0\mathbf{p})^{\mathsf{T}}(\mathbf{b}_0 - \mathbf{A}_0\mathbf{p}) \end{aligned} \quad (36)$$

The matrices $\mathbf{A}_i^2 - \mathbf{A}_0^2$ are symmetric

$$\mathbf{A}_i^2 - \mathbf{A}_0^2 = \mathbf{U}_i^{\mathsf{T}}\mathbf{D}_i\mathbf{U}_i. \quad (37)$$

Using the change of variables $\pi = \mathbf{U}_i\mathbf{p}$, $z_i(\mathbf{p}) - z_0(\mathbf{p})$ becomes

$$z_i(\mathbf{p}) - z_0(\mathbf{p}) = \pi^{\mathsf{T}}\mathbf{D}_i\pi - 2\beta_i^{\mathsf{T}}\pi + \gamma_i, \quad (38)$$

with $\beta_i^{\mathsf{T}} = \left(\mathbf{b}_i^{\mathsf{T}}\mathbf{A}_i - \mathbf{b}_0^{\mathsf{T}}\mathbf{A}_0\right)\mathbf{U}_i^{\mathsf{T}}$ and $\gamma_i = \mathbf{b}_i^{\mathsf{T}}\mathbf{b}_i - \mathbf{b}_0^{\mathsf{T}}\mathbf{b}_0$.

Box contraction suitable for SPS only

One then obtains

$$z_i(\mathbf{p}) - z_0(\mathbf{p}) = \sum_{j=1}^{n_\mathrm{p}} d_{i,j} \left(\pi_j - \frac{\beta_{i,j}}{d_{i,j}} \right)^2 + \gamma_i - \sum_{j=1}^{n_\mathrm{p}} \frac{\beta_{i,j}^2}{d_{i,j}}. \tag{39}$$

If $\mathbf{p} \in [\underline{\mathbf{p}}, \overline{\mathbf{p}}]$, one is able to get $\pi \in [\underline{\pi}, \overline{\pi}] = \mathbf{U} [\underline{\mathbf{p}}, \overline{\mathbf{p}}]$.
Whenever $d_k \neq 0$, a contractor for $[\pi_k]$ is obtained from (39) as follows

$$[\pi_k'] = [\pi_k] \cap \left(\pm \sqrt{\left(([z_i - z_0]([\mathbf{p}]) \cap [0, \infty[) - \sum_{j=1, j\neq i}^{n_\mathrm{p}} d_{i,j} \left(\pi_j - \frac{\beta_{i,j}}{d_{i,j}} \right)^2 - \gamma_i + \sum_{j=1}^{n_\mathrm{p}} \frac{\beta_{i,j}^2}{d_{i,j}} \right) / d_{i,k}} + \frac{\beta_{i,k}}{d_{i,k}} \right). \tag{40}$$

If $d_{i,k} = 0$, $[\pi_k]$ is left unchanged. From (40), a contractor for $[\mathbf{p}]$ is obtained as

$$[\mathbf{p}'] = [\mathbf{p}] \cap \left(\mathbf{U}_i^\mathrm{T} [\pi'] \right). \tag{41}$$

Building a q-relaxed intersection

Second step: contractor builds a box $[\mathbf{p}']$ enclosing the q-relaxed intersection $\mathscr{P}$ [JW02, Jau09, Jau11] of the boxes in $\mathscr{L} = \{[\mathbf{p}'_1], \ldots, [\mathbf{p}'_m]\}$

$$\mathscr{P} = \bigcap_{j \in \{1,\ldots,m\}}^{q} [\mathbf{p}'_j] . \tag{42}$$

$$= \bigcup_{\substack{J \subset [1,\ldots,m] \\ \mathrm{card}(J) \geqslant q}} \bigcap_{j \in J} [\mathbf{p}'_j] , \tag{43}$$

and satisfying

$$\mathscr{P} \subset [\mathbf{p}'] \subset [\mathbf{p}] . \tag{44}$$

Evaluating the q-relaxed intersection

Consider a list $\mathscr{L} = \{[p_1], \dots, [p_m]\}$ of m intervals.

```
[p] = ∅;
Reindex the boxes [p_i] in such a way that
    p̲_1 ⩽ p̲_2 ⩽ ··· ⩽ p̲_n;
For i = q to n
    if Σ_{j=1}^{n} (p̲_i ∈ [p_j]) ⩾ q
        p̲ = p̲_i; break;
Reindex the boxes [p_i] in such a way that
    p̄_1 ⩾ p̄_2 ⩾ ··· ⩾ p̄_n
For i = q to n
    if Σ_{j=1}^{n} (p̄_i ∈ [p_j]) ⩾ q
        p̄ = p̄_i; break;
```

$[p] = q\text{-relaxed intersection}([p_1], \dots, [p_n])$