Robust Statistics for Signal Processing

Abdelhak M Zoubir

Signal Processing Group
Technische Universität Darmstadt
email: zoubir@spg.tu-darmstadt.de
URL: www.spg.tu-darmstadt.de

Motivation for Robust Statistics

TECHNISCHE UNIVERSITÄT DARMSTADT

- Common assumptions such as Gaussianity, linearity and independence are only *approximations* to reality.

  Robust statistics is a body of knowledge, partly formalised into 'theories of robustness' relating to deviations from idealised assumptions in statistics [Hampel et al. (1986)].

- Robust statistics may be separated into two distinct but related areas
  - Robust Estimation - A robustification of classical estimation theory (point estimation)
  - Robust Testing - A robustification of the classical theory of statistical hypothesis testing (interval estimation)

Motivation for Robust Statistics (Cont'd)

TECHNISCHE UNIVERSITÄT DARMSTADT

- Often engineering systems, such as in communication, are based on a parametric model where the observations are assumed to be Gaussian
- System performance is dependent on the accuracy of this model.
  - Classical parametric statistics are optimal under exact parametric models.
  - Performance becomes more uncertain the further we are from the assumed model.
  - Incorrect model leads to a performance decrease to an uncertain level.
- Systems designed using parametric models are very sensitive to deviations from the assumed model [Hampel *et al.* (1986), Huber (1981)].

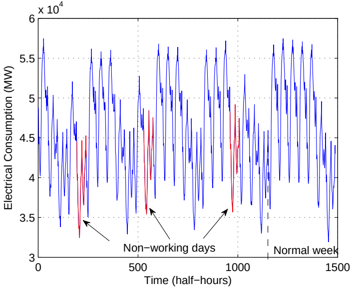
Solution: Robust estimation of parametric models

Motivation for Robust Statistics (Cont'd)

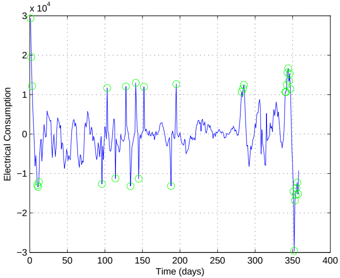
- In general, the aims of robust statistics are to:
  - describe (or fit a model to) the majority of the data
  - identify (and deal with) outliers or influential points
- How far away from Gaussianity are we in reality? Experience shows [Hampel *et al.* 1986] that
  - high quality data sets may contain up to 1% outliers,
  - low quality data sets may contain more than 10% outliers,
  - $1 - 10\%$ outliers is common.

Example: Electricity consumption data

TECHNISCHE UNIVERSITÄT DARMSTADT



Half-hourly daily French electricity consumption on April 27^{th} to June 1^{st}, 2007

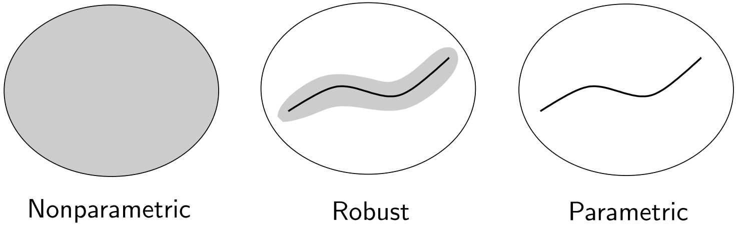


One-week differenced French load seasonal time series at 10:00, 2009

Motivation for Robust Statistics (Cont'd)

TECHNISCHE UNIVERSITÄT DARMSTADT

	Nonparametric	**Robust**	**Parametric**
Description	Model specified in terms of general properties	Parametric model allowing for deviations	Model completely specified by several parameters
Ideal Performance	Mediocre/Satisfactory	Good	Very Good/Excellent
Range of Validity	Large	Medium	Small



Robust is the most appropriate approach for real-life applications

Robust Estimation

Overview

TECHNISCHE UNIVERSITÄT DARMSTADT

- Measures of Robustness: Quantitative and Qualitative robustness
- Location Estimation
- Linear Regression Models
- Correlated Data
- Signal Processing Applications

Quantitative Robustness

- Define the 'neighborhood' using e.g. the ε contaminated mixture model (or 'gross error model')

$$\mathcal{F} = \{F \mid F = (1-\varepsilon)F_0 + \varepsilon H\},$$

  where F_0 is the nominal distribution and H is the contaminating distribution.
- Consider
  1. maximum bias
  $$b(\varepsilon) = \sup_{F \in \mathcal{F}} |\Theta(F) - \Theta(F_0)|$$
  2. maximum variance
  $$v(\varepsilon) = \sup_{F \in \mathcal{F}} AV(F, \Theta)$$

Quantitative Robustness (Cont'd)

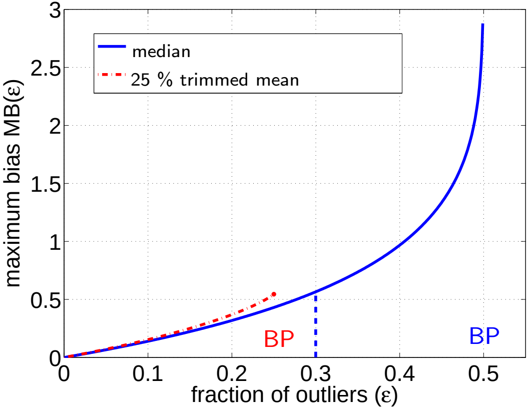
TECHNISCHE UNIVERSITÄT DARMSTADT

- Then the asymptotic breakdown point ε^* of an estimator at F_0 is

$$\varepsilon^*(F_0, \hat{\Theta}) = \sup\{\varepsilon | b(\varepsilon) < \infty\}.$$

  - Loosely speaking, it gives the limiting fraction of gross errors (outliers) the estimator can cope with (for details, see [Huber (1981) Section 1.4], [Hampel (1986) Section 2.2]).
  - In many cases ε^* does not depend on F_0 and is often the same for all the usual choices for $\mathcal{F}$.
- Maximum bias curve plots the maximum bias ($b(\varepsilon)$) of an estimator with respect to the fraction of contamination ε

Quantitative Robutsness: An Example



Maximum bias curves at $F_0 = \Phi$. Beyond the BP (ε^*), the maximum bias is infinite

Qualitative Robustness

TECHNISCHE UNIVERSITÄT DARMSTADT

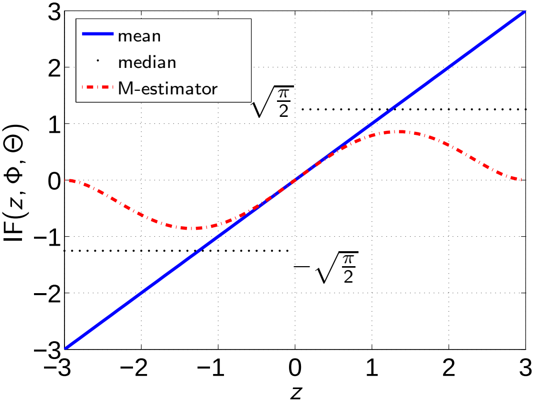
- The Influence function, introduced as influence curve [Hampel (1968,1974)], describes the effect (on an estimator $\hat{\Theta}$) of adding an observation of value x to a large sample; asymptotically, it is defined by

$$IF(x, F, \Theta) = \lim_{\Delta \to 0} \frac{\Theta((1-\Delta)F - \Delta\delta(x)) - \Theta(F)}{\Delta}$$

  where $\delta(x)$ denotes the point mass 1 at x.
- Roughly speaking, it is the first derivative of a statistic at F, where x plays the role of the contamination position.
- It measures the normalized asymptotic bias caused by an infinitesimal contamination at point x in the observations [Hampel (1986)].

Qualitative Robustness: An Example

TECHNISCHE UNIVERSITÄT DARMSTADT



Influence functions of three estimators for the standard normal distribution Φ

Trade-off Robustness vs. Efficiency

TECHNISCHE UNIVERSITÄT DARMSTADT

- Robustness: resistance of the estimator towards contamination quantified by: Breakdown Point ε^*, Maximum Bias Curve $b(\varepsilon)$, and Influence Function $IF(x, F, \Theta)$.
- Efficiency: the asymptotic behavior and variance of the estimator under the nominal model (clean data) quantified by $AV(F, \Theta)$ or $IF(x, F, \Theta)$.

$$AV(F, \Theta) = \int IF(x, F, \Theta)^2 dF(x)$$

Location Estimation

TECHNISCHE UNIVERSITÄT DARMSTADT

Consider the model

$$X_t = \mu + \varepsilon_t$$

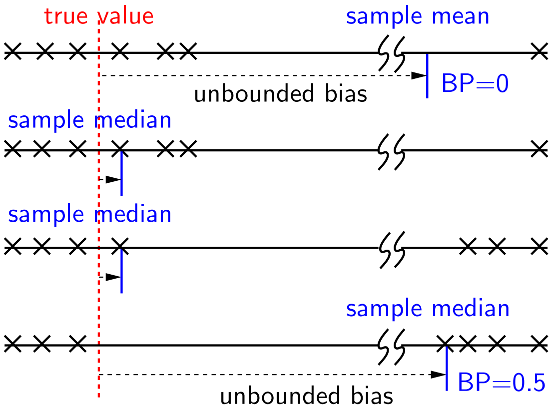
under a distribution F such that $X \sim F(x - \mu)$. We wish to estimate μ, given i.i.d $X_t,\ t = 1, \dots, n$. The Maximum likelihood estimate is

$$\widehat{\mu}_{ML} = \arg\max_{\mu} \sum_{t=1}^{n} \log f(x_t - \mu)$$

$$\Rightarrow \ \sum_{i=1}^{n} \psi(x_i - \widehat{\mu}_{ML}) = 0 \ \text{ where } \ \psi = f'/f$$

- F standard Gaussian: $\widehat{\mu}$ is the sample mean
- F double Exponential: $\widehat{\mu}$ is the sample median

Example of the Effect of Outliers

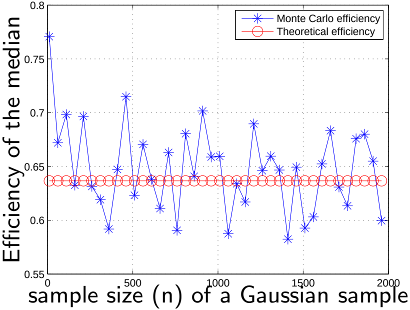


Effect of outliers on the bias of the sample mean and sample median

Example: Theoretical Robustness and Efficiency of the Sample Median

TECHNISCHE UNIVERSITÄT DARMSTADT

- IF(x,Φ,Med) bounded ⇒ sample median robust. Its BP $\varepsilon^*= 50$ %



- Median minimizes the maximum bias [Huber (1981)]. When $F_0 = \Phi$, efficiency of the Median is $2/\pi = 0.64$ ⇒ suggests M-estimators

M-Estimator

TECHNISCHE UNIVERSITÄT DARMSTADT

- The MLE is asymptotically optimal (unbiased, consistent, achieves CRB) *only* if model is correct. If model is incorrect, performance uncertain $\rightarrow$ possibly not robust.
- Huber's approach generalises the MLE of a parameter θ of interest for independent observations by replacing $f_{X_t}(x_t|\theta)$ by an arbitrary function $\rho(x_t, \theta)$.

$$arg \max_{\theta} \sum_{t=1}^{N} \log f_{X_t}(x_t|\theta) \qquad \rightarrow \qquad arg \max_{\theta} \sum_{t=1}^{N} \rho(x_t, \theta)$$

- This estimator is called M-estimator (ML-type estimator).

M-Estimator (Cont'd)

TECHNISCHE UNIVERSITÄT DARMSTADT

- Let ψ be the derivative of ρ, then

$$arg \max_{\theta} \sum_{t=1}^{N} \log f_{X_t}(x_t|\theta) \qquad \rightarrow \qquad arg \max_{\theta} \sum_{t=1}^{N} \rho(x_t, \theta)$$

leads to

$$\sum_{t=1}^{N} \frac{d \log f_{X_t}(x_t|\theta)}{d\theta} = 0 \qquad \rightarrow \qquad \sum_{t=1}^{N} \psi(x_t, \theta) = 0.$$

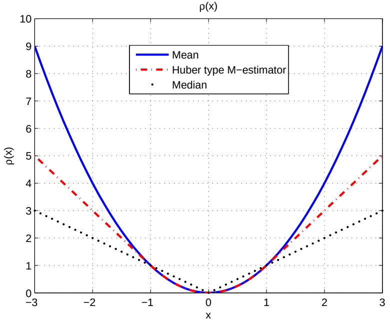
- ψ is called the score function since it 'scores' each observation x_t.

Robust M-estimators

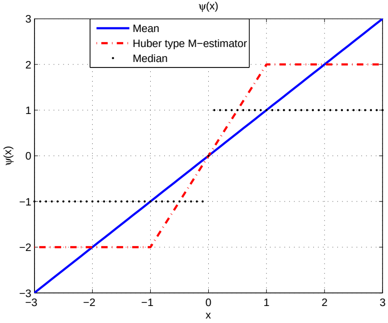
TECHNISCHE
UNIVERSITÄT
DARMSTADT

- The class of M-estimators contains in particular
  - The sample mean
  - The sample median
  - All maximum likelihood estimators
- Huber asked the following question
  How does one make an M-estimator robust?
- Implies determining f_X, or equivalently ψ, so that the M-estimator is robust
- An M-estimator is qualitatively robust if and only if ψ is bounded and continuous

Robust M-estimators



(a) ρ-function



(b) score function ψ

Huber M-estimator is qualitatively robust (ψ or IF bounded and continuous).

Linear Regression

Assume the model

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}$$

where $\mathbf{X}$ and $\mathbf{e}$ are independently distributed random variables. Let $\mathbf{Y}(t) = y_t$ and $\mathbf{x}_t'$ be the t^{th} row of the matrix $\mathbf{X}$.

Least-Squares Estimator (LSE)

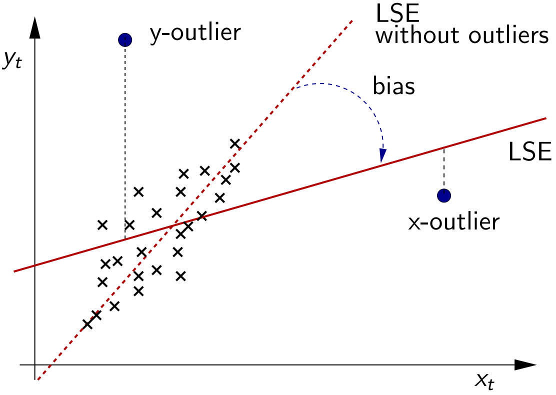
$$\widehat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} \sum_{t=1}^{n} r_t(\boldsymbol{\beta})^2$$

$r_t(\boldsymbol{\beta}) = y_t - \mathbf{x}_t'\boldsymbol{\beta}$, being the t^{th} residual. The LSE solves for

$$\sum_{t=1}^{n} r_t(\widehat{\boldsymbol{\beta}})\mathbf{x}_t = \mathbf{0}$$

Effect of Outliers

The regression line is tilted by the outliers



M-Estimation for Regression

TECHNISCHE UNIVERSITÄT DARMSTADT

M-Estimator:

$$\widehat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} \sum_{t=1}^{n} \rho\left(\frac{r_t(\boldsymbol{\beta})}{\widehat{\sigma}_r}\right) \Rightarrow \sum_{t=1}^{n} \psi\left(\frac{r_t(\widehat{\boldsymbol{\beta}})}{\widehat{\sigma}_r}\right) \mathbf{x}_t = \mathbf{0}$$

A leverage point ($\|x_t\|$ large) will dominate the equation. Here, $\widehat{\sigma}_r$ is a robust scale of the residuals r_t.

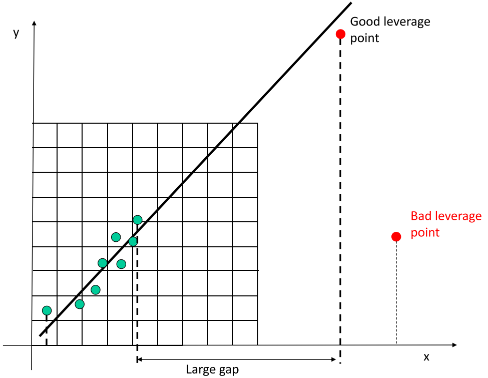
Solution:

- Redescending function ψ: MM estimator
- Down-weight leverage points: GM estimator
- Other robust estimators are: Residual Autocovariance (RA-), Least Median Squares (LMS), Least Trimmed Squares (LTS), S-, τ-Estimators.

Leverage Points in Regression

TECHNISCHE UNIVERSITÄT DARMSTADT

LSE and M-Estimator are not qualitatively robust



1

Robust Methods for Correlated Data

- Difficulty introduced by the correlation
- Several types of outliers: Additive, Innovative, patchy, isolated, ...
- Different definitions of robustness measures
- Estimators need to be adapted to deal with correlation

$$\Downarrow$$

Only a few methods exist

- 2 Definitions of influence function in the correlated data case ([Künsch (1981)] and [Martin (1981)])
- Breakdown Point (ε^*) still not clearly defined [Genton (2010)]

Some Robust Estimators for ARMA Models

TECHNISCHE UNIVERSITÄT DARMSTADT

- CML: Cleaned Maximum Likelihood after '3-σ' rejection (practical engineering method) $\Rightarrow$ it neglects the correlation, bad performance.
- Generalized-M (GM): used in Power Systems [Maronna *et al.* (2006)], not robust for ARMA; for AR(p), BP (ε^*) decreases with increasing p.
- Residual Autocovariance (RA) and truncated RA estimators (TRA) [Bustos and Yohai (1986)]. RA is not robust and TRA lacks efficiency for ARMA.
- Filtered-M: M-estimator combined with robust filter [Maronna (2006)]. Robust but lacks efficiency.
- Filtered-τ [Maronna (2006)], Ratios of Medians (RME), Medians of Ratios (MRE) and Filtered Hellinger based estimator [Chakhchoukh (2010)].

Importance of Robust Filtering

For an AR(p), a residual r_t is evaluated by regressing y_t on the p variables $y_{t-1}, \dots, y_{t-p}$. An observation y_t is used in computing $p+1$ residuals: $r_t, \dots, r_{t+p}$
$\Rightarrow$ BP (ε^*) decreases significantly.

For an ARMA model, one observation will affect all the residuals $\Rightarrow$ BP $\varepsilon^* = 0$.

Problem: Propagation of outliers in the data used in the estimation
Solution: use robust residuals computed by a robust filter cleaner [Masreliez (2010)]:

$$\widetilde{r}_t = y_t - \phi_1 \widehat{y}_{t-1|t-1} - \dots - \phi_p \widehat{y}_{t-p|t-1}$$

Cleaning the data with a robust filter improves efficiency

Robust Filtering of an AR(1)

For an AR(1): $X_t = \phi X_{t-1} + \varepsilon_t$, we use the following robust filtering algorithm:

- estimate robustly $\widehat{\phi}$, e.g.: $\widehat{\phi} = \widehat{C}_r(1)/\widehat{C}_r(0)$ and run the recursions of the filter-cleaner; $\widehat{X}_{1|1} = X_1$, $P_{1|1} = MADN(X_t)$,

Prediction :

$$\widehat{X}_{2|1} = \widehat{\phi}\widehat{X}_{1|1};$$
$$\widehat{\varepsilon}_2 = Y_2 - \widehat{\phi}\widehat{X}_{1|1}$$
$$P_{2|1} = \widehat{\phi}2P_{1|1} + \sigma_\varepsilon 2$$

Correction :

$$\widehat{X}_{2|2} = \widehat{X}_{2|1} + \frac{1}{\sqrt{P_{2|1}}} P_{2|1}\psi\left(\frac{\widehat{\varepsilon}_2}{\sqrt{P_{2|1}}}\right)$$
$$P_{2|2} = P_{2|1} - \frac{1}{P_{2|1}} P_{2|1}^2 w\left(\frac{\widehat{\varepsilon}_2}{\sqrt{P_{2|1}}}\right)$$

Go to the next step of the recursion.

- Apply the ML to the filtered series $\{\widehat{X}_{t|t}\}$ or:
- Test to remove the outliers, e.g.: If $\widehat{\varepsilon}_2 > 3\sqrt{P_{2|1}}$ then X_2 is outlying
- Apply ML estimator that handles missing data.

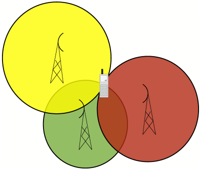
Applications

- Wireless Communications: Robust Geolocation, eg: [Hammes (2009)]
- Array Processing: Robust Direction of Arrival Estimation, eg: [Tsakalides (1995)]

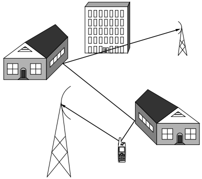
Wireless Communications: Robust geolocation

TECHNISCHE UNIVERSITÄT DARMSTADT

- Geolocation refers to identifying the position of a mobile terminal using a network of sensors.
- Applications for Geolocation arise e.g. in emergency call services, yellow page services and intelligent transport systems [Caffery (1999)].
- We consider wireless positioning of a stationary terminal based on TOA estimates.
- At least three sensors/BSs are needed to solve ambiguities



line-of-sight (LOS)



non-line-of-sight (NLOS)

Problem Statement

TECHNISCHE UNIVERSITÄT DARMSTADT

- Time of arrival (TOA) estimates are multiplied by the speed of light to obtain the measured distances

$$r_m = \underbrace{\sqrt{(x_m - x)^2 + (y - y_m)^2}}_{=h_m} + \tilde{v}_m, \qquad m = 1, \dots, M,$$

where x_m, y_m are the known coordinates of the BS and x, y describe the unknown location of the MT. The i.i.d. random variables $\tilde{v}_m$ have pdf

$$f_{\tilde{V}}(\tilde{v}) = (1-\varepsilon)\mathcal{N}(\tilde{v}; 0, \sigma_G) + \varepsilon\mathcal{H},$$

describing sensor noise and errors due to NLOS propagation ($\mathcal{H} = f_G * f_\eta$) where f_η may be any pdf with positive mean such that $\mathrm{E}\{\mathcal{H}\} > 0$.

Linearization

TECHNISCHE UNIVERSITÄT DARMSTADT

- Squaring the nonlinear equation yields

$$r_m^2 = h_m^2 + \underbrace{2h_m\tilde{v}_m + \tilde{v}_m^2}_{=v_m(h_m)} \qquad m = 1, \dots, M$$

- For M BSs we have

$$\mathbf{r} = \mathbf{S}\boldsymbol{\theta} + \mathbf{v},$$

  where $\boldsymbol{\theta} = [x\ y\ R^2]^\mathsf{T}$ with $R^2 = x^2 + y^2$.
- Since $f_V(v)$ is non-Gaussian and contains outliers due to NLOS, least-squares estimation suffers from a performance loss.

$$\Downarrow$$

Robust methods

Iterative Robust Algorithm

TECHNISCHE UNIVERSITÄT DARMSTADT

1. **Initialisation:** Set $i = 0$. Obtain an initial estimate of $\boldsymbol{\theta}$, $\hat{\boldsymbol{\theta}}^0$.

2. **Determine residuals:** $\hat{\mathbf{v}} = \mathbf{r} - \mathbf{S}\hat{\boldsymbol{\theta}}^i$.

3. **Estimate λ, perform transformation KDE.**

4. **Estimate score function:** $\hat{\varphi} = -\frac{\hat{f}'_V(v)}{\hat{f}_V(v)}$.

5. **Update:** $\hat{\boldsymbol{\theta}}^{i+1} = \boxed{\hat{\boldsymbol{\theta}}^i + \mu(\mathbf{S}^\top\mathbf{S})^{-1}\mathbf{S}^\top\hat{\varphi}(\hat{\mathbf{v}})}$ or $\boxed{(\mathbf{S}^\top\boldsymbol{\Omega}\mathbf{S})^{-1}\mathbf{S}^\top\boldsymbol{\Omega}\mathbf{r},\ \boldsymbol{\Omega} = \mathrm{diag}(\boldsymbol{\omega}),\ \boldsymbol{\omega} = |\hat{\varphi}(\hat{\mathbf{v}})/\hat{\mathbf{v}}|}$

6. **Check for convergence:** If $\frac{\|\hat{\boldsymbol{\theta}}_{i+1} - \hat{\boldsymbol{\theta}}_i\|}{\|\hat{\boldsymbol{\theta}}_{i+1}\|} < \xi$ stop, otherwise set $i \to i+1$ and go to step 2.

Simulation Settings

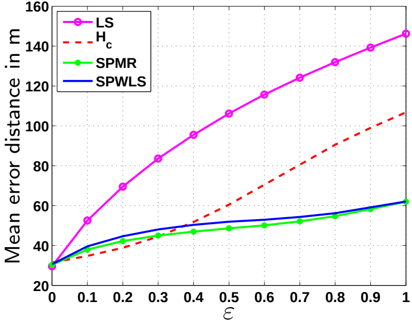
TECHNISCHE UNIVERSITÄT DARMSTADT

- Consider 10 BSs each of them collecting 10 measurements.
- We compare least-squares ('LS') with Huber's M-estimator ('H_c') where the clipping point $c = 0.6\hat{\sigma}_V$, where $\hat{\sigma}_V$ is estimated using the median absolute deviation.
- Semi-parametric estimators using Newton-Raphson algorithm labeled as 'SPMR' and the one based on weighted least-squares is 'SPWLS'.
- We average over $MC = 10,000$ Monte-Carlo runs and $\sigma_G = 150m$.
- Performance measure is the mean error distance, i.e.,

$$MED = \frac{1}{MC}\sum_{i=1}^{MC}\sqrt{(x-\hat{x}_i)^2 + (y-\hat{y}_i)^2}$$

Simulation results

TECHNISCHE UNIVERSITÄT DARMSTADT



MED vs. degree of NLOS contamination. f_η is an exponential distribution with $\sigma_\eta = 409m$

Array Processing: Robust Direction of Arrival Estimation

TECHNISCHE UNIVERSITÄT DARMSTADT

- Direction of Arrival (DOA) estimates are needed in array processing
  - Smart Antennas
  - Space-Time Adaptive Processing
  - Radar
- Classical methods for DOA estimation based on sample covariance matrix are not robust
  - Beamformer
  - Capon's minimum variance
  - ML techniques
  - Subspace methods MUSIC, ESPIRIT

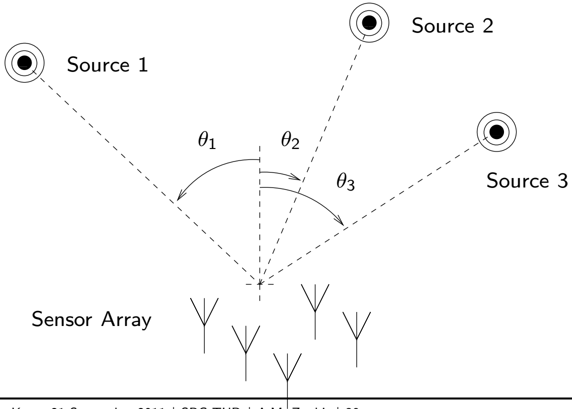
Array Processing: Robust Direction of Arrival Estimation

TECHNISCHE UNIVERSITÄT DARMSTADT

- Robust DOA methods exist based on
  - M-Estimation
  - Symmetric-Alpha-Stable (SaS) distributions and Fractional Lower Order Moments (FLOMs)
  - Gaussian mixture distributions and Space Alternating Generalised Expectation Maximisation (SAGE)
  - Nonparametric statistics using the spatial sign function
- Former three robust DOA estimators require knowledge of noise parameters/setting of thresholds/choice of weighting functions
- Last nonparametric estimator is simple and requires no prior knowledge or settings to be chosen

Array Processing: Robust Direction of Arrival Estimation

Problem: Estimate the direction-of-arrivals of the sources using the observations of a sensor array in an impulsive noise environment



Array Signal Model

TECHNISCHE UNIVERSITÄT DARMSTADT

Model

$$\mathbf{y}_n = \mathbf{A}\mathbf{s}_n + \mathbf{x}_n, \qquad n = 1, \dots, N$$

$\mathbf{y}_n$: p-dim snapshot from p array elements

$\mathbf{A} = (\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_q))$: $p \times q$-dim array steering matrix

$\mathbf{a}(\theta)$: p-dim array steering vector

$\theta_1, \dots, \theta_q$: directions to the q sources

$\mathbf{s}_n$: q-dim element source signal

$\mathbf{x}_n$: p-dim spherically symmetric noise

Assumptions

- snapshots $\mathbf{y}_n$, $n = 1, \dots, N$, are i.i.d.
- source signal $\mathbf{s}_n$ and noise $\mathbf{x}_m$ are independent for all $n, m = 1, \dots, N$
- $q < p$ to avoid identifiability problems
- $\mathbf{A}$ is of full rank q

Conventional DOA estimation: MUSIC

The spatial covariance matrix has the following structure

$$\mathbf{R} = \mathbf{A}\ \mathbf{R}_s \mathbf{A}^H + \sigma^2 \mathbf{I} \qquad (1)$$

From the eigendecomposition of R

$$\mathbf{R} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V} \qquad (2)$$

where $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_M]$ and $\boldsymbol{\Sigma} = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_M]$.

- Construct the signal and the noise subspace, $\mathbf{U}_s$ and $\mathbf{U}_n$, respectively
- Search for the peaks in the MUSIC pseudo-Spectrum $P(\theta)$

$$P(\theta) = \frac{1}{\|\mathbf{U}_n^H \mathbf{a}(\theta)\|^2} \qquad (3)$$

Conventional DOA Estimation

TECHNISCHE UNIVERSITÄT DARMSTADT

The estimated covariance matrix is given by

$$\hat{\mathbf{R}} = \frac{1}{N}\sum_{t=1}^{N}\mathbf{x}\mathbf{x}^H$$

$$\Downarrow$$

Robust Estimation:

- Normalize each of the snapshots, called spatial sign function
- Trim the corrupted observations, requires hypothesis testing
- Robust estimation of spatial covariance matrix, e.g. FLOM, maximum likelihood estimation. Some possible methods for robust covariance estimation MCD, MVE, MM-, ...

Spatial Sign Function & Robust Covariance

Spatial Sign Function

- Spatial sign function (SSF) of a *p*-variate complex vector **x**

$$\mathbf{u}(\mathbf{x}) = \begin{cases} \frac{x}{||x||} & \mathbf{x} \neq 0 \\ \mathbf{0} & \mathbf{x} = 0 \end{cases}$$

- **u** is a unit length direction vector
- Generalises the sign function $sgn(x)$ for 1-D to *p*-D

Spatial Sign Function & Robust Covariance

- Sample spatial sign covariance matrix (SCM) of

$$R_1 = E[\mathbf{u}(\mathbf{x})\mathbf{u}^H(\mathbf{x})]$$

$$\hat{R}_1 = \frac{1}{N}\sum_{n=1}^{N}\mathbf{u}(\mathbf{x}_n)\mathbf{u}^H(\mathbf{x}_n)$$

also known as quadrant correlation

- Sample spatial tau covariance matrix (TCM) of

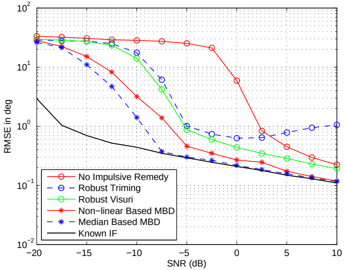
$$R_2 = E[\mathbf{u}(\mathbf{x}-\mathbf{y})\mathbf{u}^H(\mathbf{x}-\mathbf{y})]$$

$$\hat{R}_2 = \frac{1}{N(N-1)}\sum_{n=1}^{N}\sum_{m=1}^{N}\mathbf{u}(\mathbf{x}_n-\mathbf{x}_m)\mathbf{u}^H(\mathbf{x}_n-\mathbf{x}_m)$$

Simulation results

Setup

- Two linear FM signals impinging on an array of $m = 8$ sensors in ULA geometry
- DOAs are $[-3^o \ 2^o]$
- Total number of snapshots $N = 128$
- ϵ-contaminated mixture, $\epsilon = 0.2$ and $\kappa = 20$



RMSE DOA Estimation

Conclusions

TECHNISCHE UNIVERSITÄT DARMSTADT

- Robust methods are useful tools for estimation in many real world applications.
- Robust statistics for independently and identically distributed data are well-established.
- There exists a need for robust techniques for correlated data: the more interesting case for a signal processing practitioner.
- Optimality has its advantage, but *Robustness* is the engineer's interest.

References

References

D.F. Andrews, P.J. Bickel, F.R. Hampel, P.J. Huber, W.H. Rogers and J.W. Tuckey.
Robust Estimates of Location: Survey and Advances.
In *Princeton University Press*, Princeton, New York, 1972.

F. Hampel, E. Ronchetti, P. Rousseeuw, and W. Stahel.
Robust Statistics, The Approach Based on Influence Functions.
John Wiley, 1986.

P. Huber.
Robust Statistics 1st edition.
Wiley, 1981.

References

F. Hampel, E. Ronchetti, P. Rousseeuw, and W. Stahel.
Robust Statistics, The Approach Based on Influence Functions.
John Wiley, 1986.

P. Huber and E.M. Ronchetti.
Robust Statistics, 2nd edition.
Wiley, 2009.

R.A. Maronna, R.D. Martin and V.J. Yohai.
Robust Statistics Theory and Methods.
John Wiley, 2006.

P. Rousseeuw and A. Leroy.
Robust Regression and Outlier Detection.
John Wiley, 1987.

References

R. Staudte and S. Sheather.
Robust Estimation and Testing.
John Wiley, 1990.

R. Wilcox.
Introduction to Robust Estimation and Hypothesis Testing.
Academic Press, 1997.

References

O.H. Bustos and V.J. Yohai
Robust estimates for ARMA models.
J. Amer. Statist. Assoc., 81:155-168, 1986.

A.J. Fox
Outliers in Time Series.
J. R. Stat. Soc. Ser. B, 34:350-363, 1972.

M.G. Genton and A. Lucas
Comprehensive definitions of breakdown-points for independent and dependent observations.
J. R. Stat. Soc. Ser. B, 65:81-94, 2003.

References

R.D. Martin and V.J. Yohai
Influence functionals for time series.
Ann. Statist., 14:781-818, 1986.

N. Muler, D-Peña and V.J. Yohai
Robust Estimation for ARMA models.
Ann. Statist., 37(2):816-840, 2009.

H. Künsch,
Infinitesimal robustness for autoregressive processes.
Ann. Statist., 12:843–863, 1984.

References

C. J. Masreliez, and R. D. Martin.
Robust Bayesian estimation for the linear model and robustifying the Kalman filter.
IEEE Transactions on Automatic Control, 22(3):361–371, 1977.

Y. Chakhchoukh.
Contribution to the robust estimation of SARIMA models. Application to short term load forecasting of electricity consumption
In *Ph.D. Thesis*, University of Paris XI, Paris, 2010.

Yin et. al.
Weighted median filters: A tutorial.
IEEE Transactions on Circuits and Systems-II: Analog and Digital Signal Processing, 43(3):157–192, March 1996.

References

F. Gustafsson and F. Gunnarsson.
Mobile positioning using wireless networks: possibilities and fundamental limitations based on available wireless network measurements.
IEEE Signal Processing Magazine, 22(4):41–53, 2005.

U. Hammes.
Robust positioning algorithms for wireless networks
In *Ph.D. Thesis*, TU Darmstadt, Darmstadt, 2009.

P. Tsakalides.
Subspace-based direction finding in alpha-stable noise environments.
In *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 2092–5, Detroit, USA, 1995.