New Model Order Selection in Large Dimension Regime for Complex Elliptically Symmetric Noise

Eugenie Terreaux [1] Frederic Pascal [2] Jean-Philippe Ovarlez [1]

[1] SONDRA, CentraleSupélec, University Paris-Saclay, 3 rue Joliot Curie, 91190 Gif-sur-Yvette, France

[2] L2S, CentraleSupélec, CNRS, University Paris-Saclay, 3 rue Joliot Curie, 91190 Gif-sur-Yvette, France

Introduction and Motivations

Let us consider a set of N **observations** $\{\mathbf{y}_i\}_{i\in[\![1,N[\![}$ where each $\mathbf{y}_i$ is a **multidimensional m-vector**.

GOAL : Estimate the model order for high dimensional and Complex Elliptically Symmetric (CES) distributed signal

- **Large number of data**: N and m are of same order with possibly $N > m$ $(N, m) \to \infty$ with $m/N \to c \neq 0$
- **CES noise**: [1] the noise $\mathbf{n}_i \sim \sqrt{\tau_i}\,\mathbf{C}^{1/2}\,\mathbf{x}_i\,,\quad i \in [\![1,N]\!]$

with:

. $\mathbf{x}_i$ a white noise

. $\mathbf{C}$ a Hermitian Toeplitz covariance matrix

Statistical Model:

$$\mathbf{y}_i = \sum_{j=1}^{p} s_{i,j}\,\mathbf{m}_j + \sqrt{\tau_i}\,\mathbf{C}^{1/2}\,\mathbf{x}_i\,,\quad i\in[\![1,N]\!]$$

. $\mathbf{m}_j$ p independent vectors the unknown m-steering vector of the j-th deterministic source

. $s_{i,j}$ the power of each source j in each wavelength i.

Proposed Algorithm:

. **Whitening** of the Received Signal

. **Robust Estimation** of the Covariance Matrix

. **Thresholding** of the Eigenvalues of the Estimated Covariance Matrix

Assumptions:

. $\sum_{i=1}^{N}\delta_{\tau_i}$ satisfies $\int \tau\mu_N(d\tau) \to 1$ almost surely

. $\max_i d_1(\lambda_i(\mathbf{C}), \text{supp}(\nu)) \to 0$

. $\{c_k\}_{k\in[0,m-1]}$ are absolutely summable coefficients, such that $c_0 \neq 0$.

. $\frac{1}{N}\sum \delta_{\lambda_i(\mathbf{C})}$ converges almost surely toward the true measure ν, λ_i the $i-th$ largest eigenvalue of $\mathbf{C}$

Signal Whitening

- Estimation of $\mathbf{C}$:

Notation :

. $\check{c}_k = \frac{1}{m\,N}\sum_{i=1}^{m}\sum_{j=1}^{N} y_{i,j}\,y^{\star}_{i+k,j}\,\mathbb{1}_{1\leq i+k\leq m}\,,\ k\in]\!]0m-1[\![$

. $\left([\mathcal{T}(\check{\mathbf{c}})]_{i,j}\right)_{i\leq j} = \check{\mathbf{c}}_{i-j}$ and $\left([\mathcal{T}(\check{\mathbf{c}})]_{i,j}\right)_{i>j} = \check{c}^{\star}_{i-j}$

Theorem 1: Consistent estimator of $\mathbf{C}$

$$\|\mathcal{T}(\check{\mathbf{c}}) - \mathbb{E}[\tau]\,\mathbf{C}\| \to 0$$

- Whitening: $\check{\mathbf{C}} = \mathcal{T}(\check{\mathbf{c}})$

$$\mathbf{Y}_w = \check{\mathbf{C}}^{-1/2}\,\mathbf{Y} = \check{\mathbf{C}}^{-1/2}\,\mathbf{M}\,\mathbf{S} + \check{\mathbf{C}}^{-1/2}\,\mathbf{C}^{1/2}\,\mathbf{X}\,\mathbf{T}^{1/2}$$

Signal Subspace Rank Estimation

- Robust Estimation of the white covariance matrix with a **M-estimator** [2] : $\check{\mathbf{\Sigma}}$ the unique solution if it exists of:

$$\mathbf{\Sigma} = \frac{1}{N}\sum_{i=1}^{N} u\left(\frac{1}{m}\,\mathbf{y}_{w_i}^{H}\,\mathbf{\Sigma}^{-1}\,\mathbf{y}_{w_i}\right)\,\mathbf{y}_{w_i}\,\mathbf{y}_{w_i}^{H}\,,$$

u : [0, $+\infty$) $\to$[0, $+\infty$) nonnegative, continue and non-increasing

- Thresholding of the eigenvalues of $\check{\mathbf{\Sigma}}$

Theorem 2: Convergence of $\check{\mathbf{\Sigma}}$

$$\left\|\check{\mathbf{\Sigma}} - \hat{\mathbf{S}}\right\| \to 0\ a.s.$$

Use of the **threshold** t of the eigenvalues of $\hat{\mathbf{S}}$ on $\check{\mathbf{\Sigma}}$.

Notation :

$\{\lambda_i(\check{\mathbf{\Sigma}})\}_{i=[1,N]}$ the eigenvalues of $\check{\mathbf{\Sigma}}$ sorted in descending order.

$\hat{p}$ the estimation of p

$$\hat{p} = \min_k(\lambda_k > t)$$

Results and Simulations

- Simulated Data

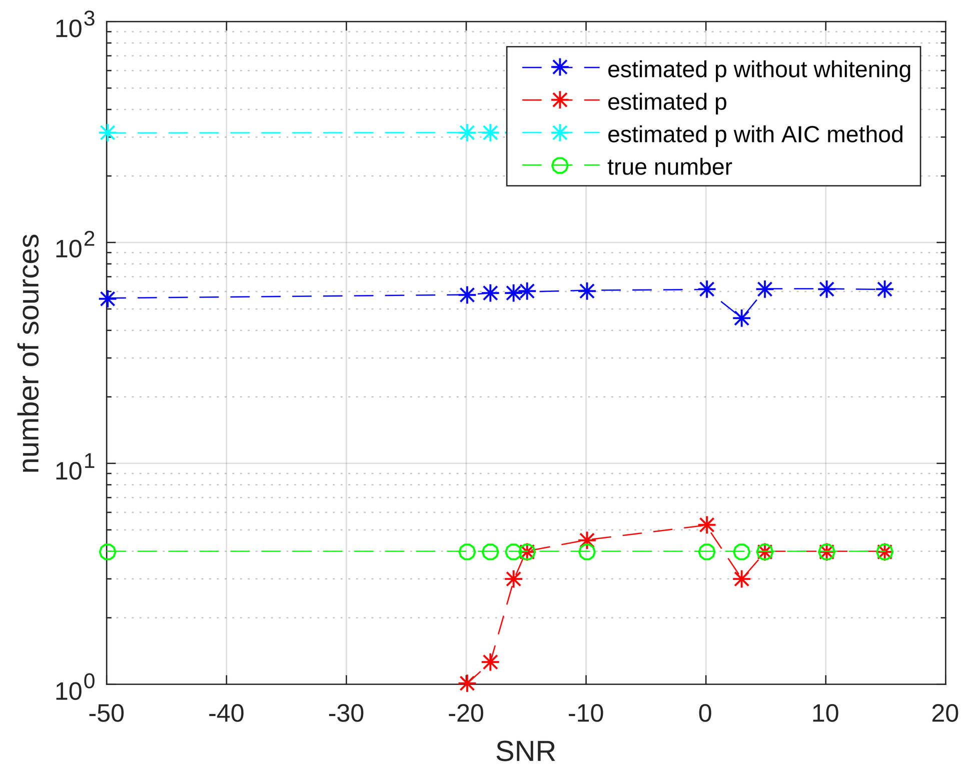
. Simulated and correlated CES noise

. $m = 400$ and $N = 2000$, $p = 4$

. $u : x \mapsto \dfrac{1+\nu}{\nu + x}$

. The threshold $t = \dfrac{(1+\nu)\,(1+\sqrt{c})^2}{\gamma_m\,(1 - c\,(1+\nu))}$ with γ_m the unique solution (if it exists) of $\sum_{i=1}^{N}\dfrac{\psi(\tau_i\,\gamma)}{1 + c_N\,\psi(\tau_i\,\gamma)} = 1$

. Comparison with two methods based on the information theory criterion: AIC and MDL



- Real Hyperspectral Images

Estimated p for different hyperspectral images. $M1$ is the proposed method, $M2$ is the same algorithm without the whitening step.

Images	Indian Pines	SalinasA	PaviaU	Cars
p	16	9	9	6
$\hat{p}$ M1	11	9	1	3
$\hat{p}$ M2	220	204	103	1
$\hat{p}$ AIC	219	203	102	143

- Our Algorithm gives interesting and encouraging results on simulated and on real hyperspectral images.

Conclusion

This method can be generally applied for any model order selection problems as radar clutter rank estimation, sources localization or any hyperspectral problems such as anomaly detection or linear or non-linear unmixing techniques.

[1] E. Ollila and D.E. Tyler and V. Koivunen, and H.V. "Poor Complex elliptically symmetric distributions: Survey, new results and applications", IEEE Transactions on signal processing, vol 60, no. 11, 5597-5625, 2012.

[2] R. Maronna and V. Yohai, "Robust Estimation of Multivariate Location and Scatter", John Wiley and Sons, Inc., 2004.