Structured covariance estimation in high resolution spectral analysis

Mattia Zorzi

Department of Information Engineering - University of Padova

February, 28th 2013

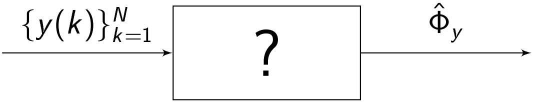
DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

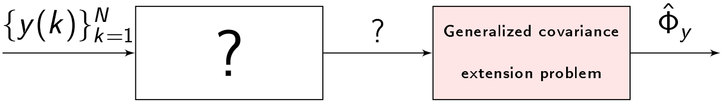
Outline

Spectral estimation problem

- $y = \{y_k \, ; \, k \in \mathbb{Z}\}$ zero mean, $\mathbb{C}^m-$valued, stationary and purely nondeterministic Gaussian process
- $\{y(k)\}_{k=1}^N$ available finite data sequence

TASK Estimate the power spectral density (psd) $\Phi_y(\mathrm{e}^{\mathrm{j}\vartheta})$ of y

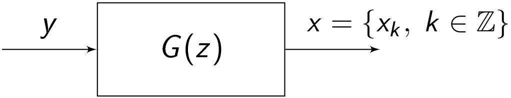




SPLIT the problem!

Generalized covariance extension problem

- $G(z) = (zI - A)^{-1}B$ *Bank of filters*
  - $A \in \mathbb{C}^{n\times n}$ strictly stable, $B \in \mathbb{C}^{n\times m}$ full column rank

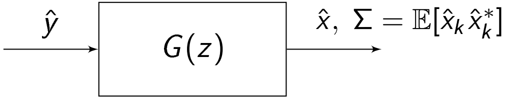


- $\Sigma = \mathbb{E}[x_k x_k^*] = \int G\Phi_y G^*$ Output covariance matrix

Generalized covariance extension problem (cont'd)

Problem statement

Given Σ, find a psd $\hat{\Phi}_y(\mathrm{e}^{\mathrm{j}\vartheta})$ such that $\int G\hat{\Phi}_y G^* = \Sigma$



Once given Σ, does $\hat{\Phi}_y$ exist?

Example: $G(z)$ bank of delays

- $$A = \underbrace{\begin{bmatrix} 0 & I_m & \dots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \dots & I_m \\ 0 & 0 & \dots & 0 \end{bmatrix}}_{l \times l \text{ blocks}}, \; B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ I_m \end{bmatrix} \Rightarrow x_k = \begin{bmatrix} y_{k-l+1} \\ y_{k-l+2} \\ \vdots \\ y_k \end{bmatrix}$$
- $$\Sigma = \begin{bmatrix} R_0 & R_1 & R_2 & \dots & R_{l-1} \\ R_1^* & R_0 & R_1 & \ddots & R_{l-2} \\ R_2^* & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \\ R_{l-1}^* & R_{l-2}^* & \ddots & \ddots & R_0 \end{bmatrix}, \quad R_l = \mathbb{E}[y_{k+l} y_k^*]$$

Covariance extension problem

How to choose $\hat{\Phi}_y$?

- Entropy rate of a stochastic process with psd Φ

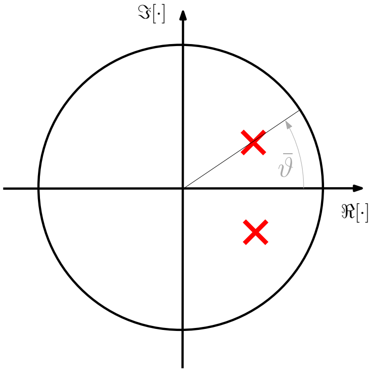
$$\mathbb{H}(\Phi) = \int \log \det \Phi(e^{j\vartheta})$$

Maximum entropy solution (Byrnes-Georgiou-Lindquist,2000)

$$\hat{\Phi}_{THREE}(e^{j\vartheta}) = \underset{\Phi \in \mathcal{S} \,:\, \int G\Phi G^* = \Sigma}{\text{argmax}} \; \mathbb{H}(\Phi)$$ RATIONAL!

- Special case: "Classical" maximum entropy solution when $G(z)$ bank of delays (Burg, 1967)

How to choose $G(z)$?



Poles of $G(z)$

If poles close to $|z|=1$, get higher resolution of $\hat{\Phi}_y(e^{j\vartheta})$ at $\bar{\vartheta}$

High resolution in prescribed frequency bands!

Other THREE-type solutions

- Kullback-Leibler relative entropy (Georgiou-Lindquist,2003)
- Hellinger distance (Ferrante-Pavon-Ramponi,2008)
- Relative entropy rate (Ferrante-Masiero-Pavon,2012)
- Beta divergence (Zorzi,submitted,2012)
- Alpha divergence (Zorzi,submitted,2013)

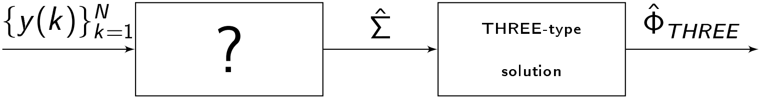
- $\hat{\Phi}_y = \frac{\Psi}{G^*\Lambda G}$
- $\hat{\Phi}_y = (I + G^*\Lambda G)^{-1}\Psi(I + G^*\Lambda G)^{-1}$
- $\hat{\Phi}_y = (\Psi^{-1} + G^*\Lambda G)^{-1}$
- $\hat{\Phi}_y = (\Psi^{-\frac{1}{\nu}} + G^*\Lambda G)^{-\nu}$
- $\hat{\Phi}_y = \frac{\Psi}{(1+\frac{1}{\nu} G^*\Lambda G)^{\nu}}$

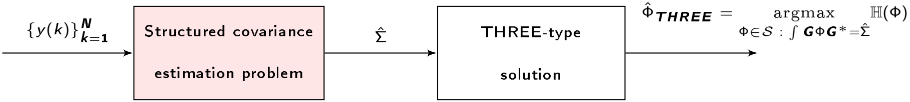
Ψ a priori spectral density

Λ Lagrange multiplier

Remarks

- Given entries for the spectral estimation problem (by choosing THREE-type solutions)
  - $\{y(k)\}_{k=1}^N \rightarrow$ Sample data
  - $G(z) \quad \rightarrow$ Fixed
- $\Sigma = \int G\Phi_y G^*$ is not given!




$\{y(k)\}_{k=1}^{N}$
Structured covariance estimation problem
$\hat{\Sigma}$
THREE-type solution
$\hat{\Phi}_{THREE} = \underset{\Phi \in \mathcal{S} \,:\, \int G \Phi G^* = \hat{\Sigma}}{\operatorname{argmax}} \mathbb{H}(\Phi)$

Structured covariance estimation problem

Problem statement

Given $G(z)$ and $\{y(k)\}_{k=1}^{N}$, find $\hat{\Sigma} > 0$ such that
$\hat{\Sigma} = \int G\Phi G^* \;\; \exists \Phi \in \mathcal{S}$

- $\Gamma : \mathcal{C} \rightarrow \mathcal{H}^{n\times n}, \quad \Delta \mapsto \int G\Delta G^*$
- $\mathrm{Range}\,\Gamma := \{M \in \mathcal{H}^{n\times n} \mid \exists\, \Delta \in \mathcal{C} \text{ s.t. } \int G\Delta G^* = M\}$
- $\hat{\Sigma} \in \mathrm{Range}\,\Gamma$ and positive definite iff there exists $\Phi \in \mathcal{S}$ such that $\int G\Phi G^* = \hat{\Sigma}$ (Georgiou, 2002)

Problem

Given $G(z)$ and $\{y(k)\}_{k=1}^{N}$, find a positive definite matrix
$\hat{\Sigma} \in \mathrm{Range}\,\Gamma$

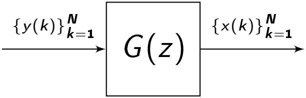
Example: $G(z)$ bank of delays

- $$\Sigma = \begin{bmatrix} R_0 & R_1 & \dots & R_{l-1} \\ R_1^* & R_0 & \ddots & R_{l-2} \\ \vdots & \ddots & \ddots & \ddots \\ R_{l-1}^* & \ddots & R_1^* & R_0 \end{bmatrix}$$

Range $\Gamma \equiv$ Toeplitz matrices

- $$\hat{\Sigma}_1 := \begin{bmatrix} \hat{R}_0 & \hat{R}_1 & \dots \\ \hat{R}_1^* & \hat{R}_0 & \ddots \\ \vdots & \ddots & \ddots \end{bmatrix}$$ not positive definite!

- $\hat{\Sigma}_2 := \frac{1}{N} \sum_{k=1}^{N} x(k)x(k)^* \notin \text{Range}\,\Gamma$



Approaches for estimating structured covariances

- Blackman-Tukey method (Blackman-Tukey,1959)
- Maximum likelihood method (Burg-Luenberger-Wenger,1982)
- Projection method (Ferrante-Pavon-Ramponi,2008)

Our contribution:

- Maximum entropy method
- Input covariance lags method

- Transportation distance method (Ning-Jiang-Georgiou, in press)

Maximum likelihood method (Burg *et al.*,1982)

- Information divergence between two Gaussian distribution p_{Q_1}, p_{Q_2} on $\mathbb{C}^n$ with zero mean and covariance Q_1, Q_2

$$\mathbb{D}(Q_1 \| Q_2) := \frac{1}{2}[\log\det(Q_1^{-1}Q_2) + \mathrm{Tr}(Q_1 Q_2^{-1}) - n]$$

ML method

Given $G(z)$ and $\hat{\Sigma}_C = \sum_{k=1}^{N} x(k)x(k)^* > 0$, compute

$$\hat{\Sigma}_{ML} := \underset{S>0,\ S\in \mathrm{Range}\,\Gamma}{\operatorname{argmin}} \ \mathbb{D}(\hat{\Sigma}_C \| S)$$

- Numerical method for finding a local minimum is presented

Maximum entropy method

ME method

Given $G(z)$ and $\hat{\Sigma}_C \sum_{k=1}^{N} x(k)x(k)^* > 0$, compute

$$\hat{\Sigma}_{ME} := \underset{S>0,\ S\in \text{Range}\,\Gamma}{\operatorname{argmin}} \mathbb{D}(S\|\hat{\Sigma}_C)$$

Constrained convex optimization problem!

- $S \in \mathrm{Range}\Gamma$ iff $\Pi_B^\perp(S - ASA^*)\Pi_B^\perp = 0$, $\Pi_B^\perp = I - B(B^*B)^{-1}B^*$
- $\mathcal{L}_{\hat{\Sigma}_C}(S,\Lambda) = \mathbb{D}(S\|\hat{\Sigma}_C) + \mathrm{Tr}\left[\Lambda\Pi_B^\perp(S - ASA^*)\Pi_B^\perp\right]$
- The optimal solution has the form $\hat{\Sigma}_{ME}(\Lambda) = (\hat{\Sigma}_C^{-1} + 2Q_\Lambda)^{-1}$

Maximum entropy method (cont'd)

Dual problem

Find Λ° by solving

$$\Lambda^\circ := \underset{\Lambda \in \mathcal{I}}{\operatorname{argmax}}\, \mathcal{L}_{\hat{\Sigma}_C}(\hat{\Sigma}_{ME}(\Lambda), \Lambda) = \underset{\Lambda \in \mathcal{I}}{\operatorname{argmax}} \left\{ \frac{1}{2} \operatorname{Tr} \log \left(\hat{\Sigma}_C^{-1} + 2 Q_\Lambda \right) \right\},$$

$$\mathcal{I} = \{\Lambda \in \mathcal{H}^{n \times n} \mid \hat{\Sigma}_{ME}(\Lambda) > 0\}$$

- It can be proved that Λ° exists
- Λ° can be computed via a globally convergent matricial Newton-like algorithm

Input covariance lags method

- $G(z) = (zI - A)^{-1}B$
- Σ is given by

$$\Sigma - A\Sigma A^* = AQ_R B^* + BQ_R A^* + BR_0 B^*$$

with $Q_R := \sum_{l=1}^{\infty} A^{l-1} B R_l^*$, $R_l := E[y_{k+l} y_k^*]$ (Georgiou,2002).

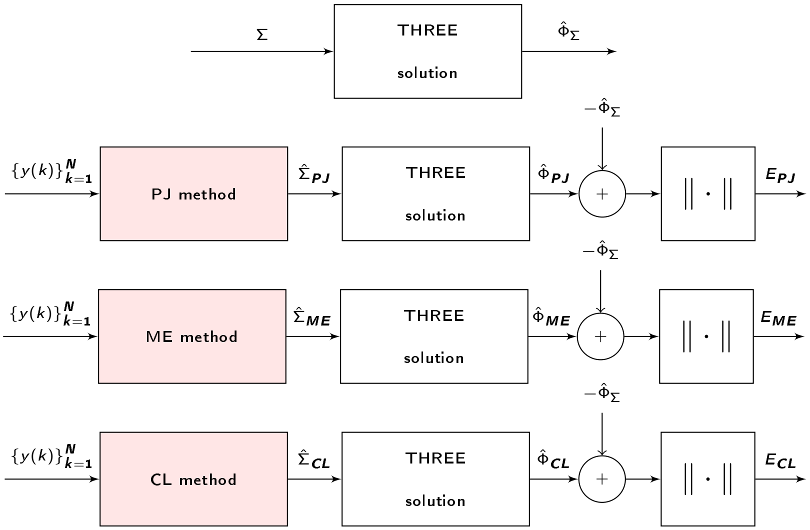
CL method

An estimate $\hat{\Sigma}_{CL} \geq 0$ of Σ is given by solving the Lyapunov equation

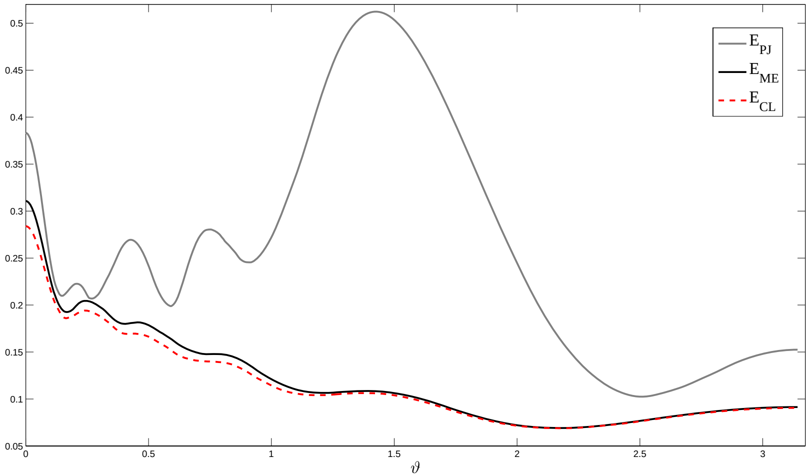
$$\hat{\Sigma}_{CL} - A\hat{\Sigma}_{CL} A^* = AQ_{\hat{R}} B^* + BQ_{\hat{R}} A^* + B\hat{R}_0 B^*$$

with $\{\hat{R}_l\}_{l=-\infty}^{\infty}$ input covariance lags sequence estimated by employing the Blackman-Tukey method

Estimated psd comparison

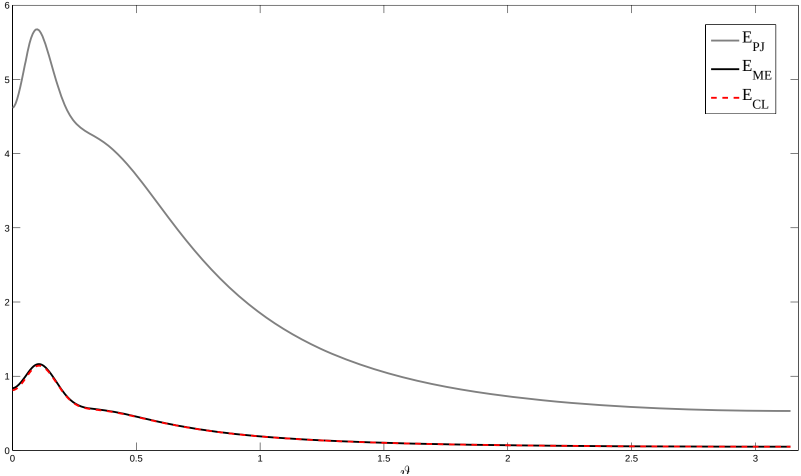


Scalar process (m=1) with THREE solution



Mean of the error norm comparison averaged over 500 experiments with $N = 500$

Bivariate process (m=2) with THREE solution



Mean of the error norm comparison averaged over 500 experiments with $N = 500$

Conclusions

- Psd estimation reliability strongly depends on the estimate of Σ (by employing THREE-like solutions)
- The ME and CL methods have been presented
- The ME and CL methods provide better performances than the PJ method in THREE-like spectral estimation paradigms

Thank you for your attention

Other THREE-type solutions

- $G(z)$ Bank of filters
- Σ Output covariance matrix of G
- $\Psi(\mathrm{e}^{\mathrm{j}\vartheta})$ A priori power spectral density

Entropic-type solution (Georgiou-Lindquist, 2003)

$$\hat{\Phi}_{KL-THREE}(\mathrm{e}^{\mathrm{j}\vartheta}) = \underset{\Phi\in\mathcal{S}\,:\,\int G\Phi G^*}{\operatorname{argmin}} \mathbb{D}_{KL}(\Psi\|\Phi) \quad \text{RATIONAL!}$$

- Kullback-Leibler divergence between power spectral densities

$$\mathbb{D}_{KL}(\Psi\|\Phi) = \int \Psi \log\left(\frac{\Psi}{\Phi}\right)$$

Other THREE-type solutions (cont'd)

Other divergences employed

- Hellinger distance (Ferrante-Pavon-Ramponi,2008)

$$\mathbb{D}_H(\Psi,\Phi) \;=\; [\inf \|W_\Psi - W_\Phi\|_2^2 : \; W_\Psi, W_\Phi \in L_2^{m\times m}, \\ W_\Psi W_\Psi^* = \Psi, \; W_\Phi W_\Phi^* = \Phi]^{\frac{1}{2}}$$

- Relative entropy rate divergence (Ferrante-Masiero-Pavon,2012)

$$\mathbb{D}_{RER}(\Phi\|\Psi) = \frac{1}{2}\int \log\det(\Phi^{-1}\Psi) + \mathrm{Tr}[\Psi^{-1}(\Phi-\Psi)]$$

Other THREE-type solutions (cont'd)

Other divergences employed

- Beta divergence family (Zorzi,submitted,2012)

$$\mathbb{D}_\beta(\Phi\|\Psi) = \operatorname{Tr}\int \frac{1}{\beta-1}(\Phi^\beta - \Phi\Psi^{\beta-1}) - \frac{1}{\beta}(\Phi^\beta - \Psi^\beta)$$

- Alpha divergence family (Zorzi,submitted,2013)

$$\mathbb{D}_\alpha(\Phi\|\Psi) = \int \frac{1}{\alpha(\alpha-1)}\Phi^\alpha\Psi^{1-\alpha} - \frac{1}{\alpha-1}\Phi + \frac{1}{\alpha}\Psi$$

Projection method (Ferrante *et al*,2008)

- We can compute a basis of Range Γ (Georgiou,2002)

PJ method

Let $\hat{\Sigma}_\Gamma$ be the projection of $\hat{\Sigma}_C := \frac{1}{N}\sum_{k=1}^{N} x(k)x(k)^*$ onto Range Γ. Then,

$$\hat{\Sigma}_{PJ} := \hat{\Sigma}_\Gamma + \varepsilon \Sigma_+$$

with $\varepsilon \geq 0$ so large that $\hat{\Sigma}_{PJ} > 0$ and $\Sigma_+ \in$ Range Γ positive definite

Input covariance lags method

How to estimate Q_R and R_0?

- Blackman-Tukey correlogram

$$\hat{R}_l = \begin{cases} \frac{1}{N}\sum_{k=1}^{N-l} y(k+l)y(k)^*, & 0 \leq l < L \\ 0_{m\times m}, & l \geq L \end{cases}$$

$$\rightarrow \quad \hat{R}_0 = \frac{1}{N}\sum_{k=1}^{N} y(k)y(k)^*$$

$$\rightarrow \quad \hat{Q}_R = \frac{1}{N}\sum_{l=1}^{L-1}\sum_{k=1}^{N-l} A^{l-1}By(k)y(k+l)^*$$

The corresponding $\hat{\Sigma}_{CL}$ is positive semi-definite

- CL method can be generalized to $G(z) = (zI - A)^{-1}B + D$

Transportation distance method (Ning *et. al*, in press)

- Transportation distance between two Gaussian distribution p_{Q_1}, p_{Q_2} on $\mathbb{R}^n$ with zero mean and covariance Q_1, Q_2

$$\mathbb{D}_{TD}(Q_1, Q_2) = \min_{T} \left\{ \mathrm{Tr}(Q_1 + Q_2 - T - T^*) \mid \begin{bmatrix} Q_1 & T \\ T^* & Q_2 \end{bmatrix} \geq 0 \right\}$$

TD method

Given $G(z)$ and $\hat{\Sigma}_C > 0$, compute

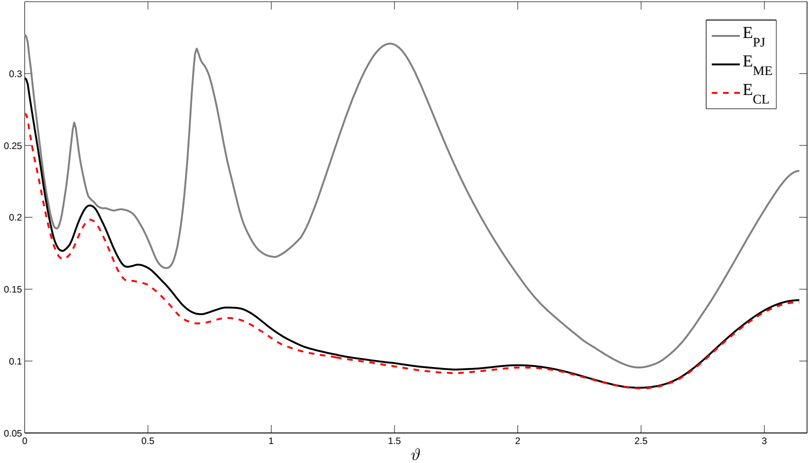
$$\hat{\Sigma}_{TD} := \underset{S>0,\, S \in \mathrm{Range}\,\Gamma}{\operatorname{argmin}} \mathbb{D}_{TD}(\hat{\Sigma}_C, S)$$

Relative matrix error norm comparison

- N sample length
- μ mean of the relative matrix error norm $e := \frac{\|\hat{\Sigma}-\Sigma\|}{\|\Sigma\|}$
- σ variance of the relative error norm e
- $\sharp F$ times that the projection $\Sigma_\Gamma \not> 0$

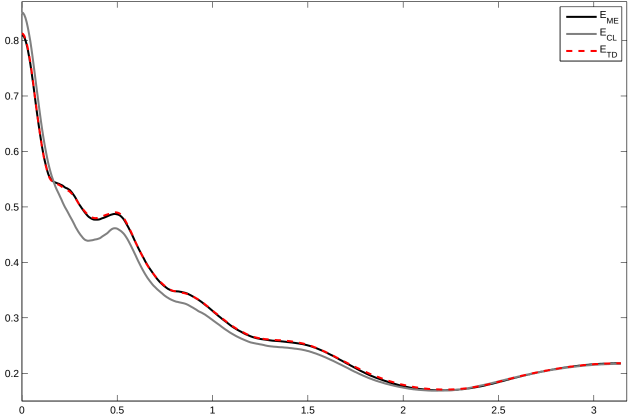
N	μ_{CL}	μ_{PJ}	μ_{ME}	σ^2_{CL}	σ^2_{PJ}	σ^2_{ME}	$\sharp F$
300	0.18	0.81	0.18	0.018	2.65	0.02	73
500	0.16	0.47	0.15	0.013	1.37	0.013	37
700	0.13	0.29	0.13	0.001	0.74	0.009	18

Scalar process (m=1) with KL-THREE solution



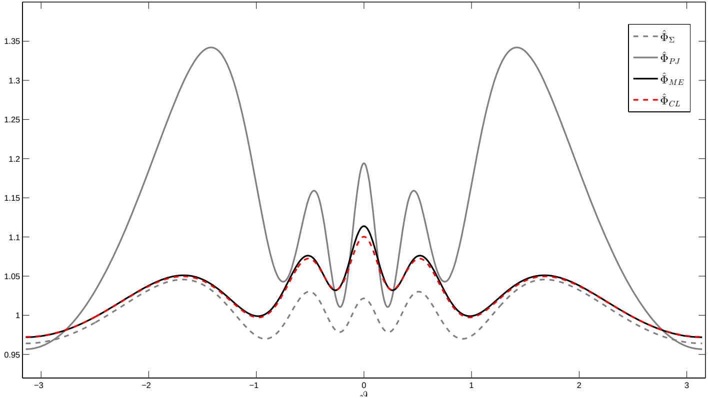
Mean of the error norm comparison averaged over 500 experiments with $N = 500$

Further simulation results: scalar process (m=1) with THREE solution



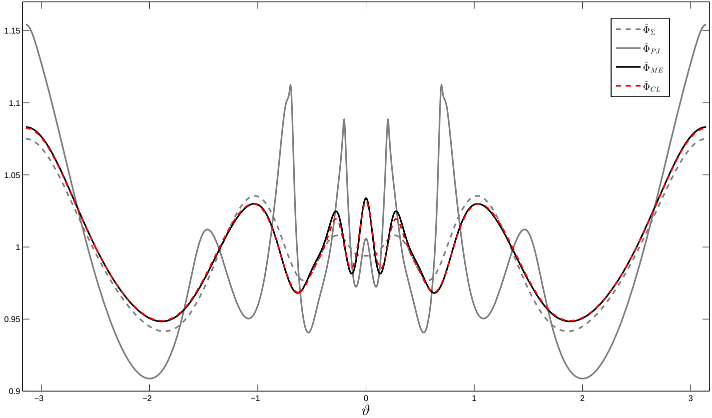
Mean of the error norm comparison averaged over 500 experiments with $N = 100$

Further simulation results: Scalar process (m=1) with THREE solution



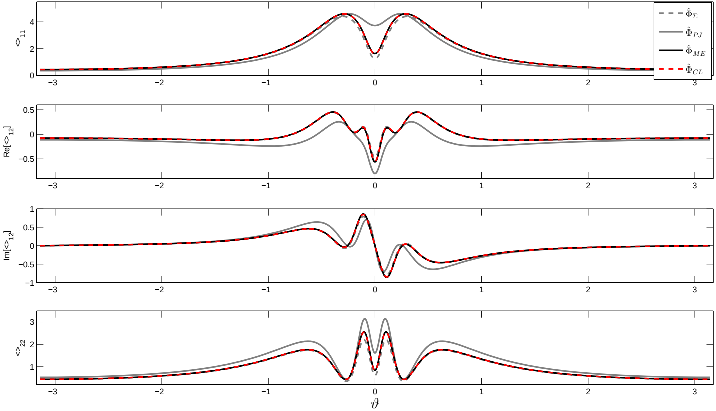
Mean spectra comparison averaged over 500 experiments with $N = 500$

Further simulation results: Scalar process (m=1) with KL-THREE solution



Mean spectra comparison averaged over 500 experiments with $N = 500$

Further simulation results: Bivariate process (m=2) with THREE solution



Mean spectra comparison averaged over 500 experiments

References

A. Ferrante, M. Pavon and M. Zorzi
A maximum entropy enhancement for a family of high-resolution spectral estimators
IEEE Trans. Aut. Control, Vol. 57(2):318–329, 2012.

M. Zorzi and A. Ferrante
On the estimation of structured covariance matrices
Automatica, Vol. 48(9):2145–2151, 2012.

J. P. Burg, D. G. Luenberger and D. L. Wenger
Estimation structured Covariance Matrices
Proceedings of the IEEE, Vol. 70(9):963–974, 1982.

C. I Byrnes, T. Georgiou and A. Lindquist
A new approach to spectral estimation: A tunable high-resolution spectral estimator
IEEE Transaction on Signal Processing, Vol. 49(11):3189–3205, 2000.

A. Ferrante, M. Pavon and F. Ramponi
Hellinger vs. Kullback-Leibler multivariable spectrum approximation
IEEE Trans. Aut. Control, Vol. 53(4):954–967, 2008.

A. Ferrante, C. Masiero and M. Pavon
Time and spectral domain relative entropy: A new approach to multivariate spectral estimation
IEEE Trans. Aut. Control, Vol. 57(10):2561–2575, 2012.

M. Zorzi
A new family of high-resolution multivariate spectral estimators
Submitted, 2012.

M. Zorzi
Rational approximations of spectral densities based on the Alpha divergence
Submitted, 2013.