Beyond Whittle's likelihood - new Bayesian semiparametric approaches to multivariate time series analysis

Claudia Kirch, **Alexander Meier**, *Otto-von-Guericke University Magdeburg, Germany*
Renate Meyer, University of Auckland, New Zealand

Abstract

In nonparametric Bayesian time series analysis, Whittle's likelihood provides a well-established method to model a stationary time series. In general, Whittle's likelihood constitutes an approximation to the true likelihood. It often yields asymptotically correct inference results, however at the price of losses in efficiency. Recently, Whittle's likelihood was generalized [1] by first fitting *any* suitable parametric model (beyond Gaussian white noise) in the time domain and then applying a nonparametric correction in the frequency domain. This yields a pseudo likelihood that inherits the correct second order structure from the nonparametric spectral correction as well as the dependence between periodogram ordinates from the parametric fit. We present an extension of this approach to multivariate time series and give an outlook to upcoming tasks and challenges.

Whittle's likelihood: The multivariate case

Let $\{\underline{X}_n : n = 0, \pm 1, \dots\}$ be a stationary and centered real d-dimensional time series with absolute summable autocovariance matrix function $\mathbf{\Gamma}(h) = \mathbb{E}[\underline{X}_t \underline{X}_{t+h}^\top]$ and spectral density

$$\mathbf{f}(\lambda) = \frac{1}{2\pi} \sum_{h \in \mathbb{Z}} \mathbf{\Gamma}(h) e^{-ih\lambda}, \quad 0 \le \lambda \le 2\pi.$$

The Fourier coefficients of $\mathbf{X}_n = (\underline{X}_1, \dots, \underline{X}_n)^\top$ are given by

$$\tilde{\underline{X}}(\lambda_k) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \underline{X}_j e^{-ij\lambda_k}, \quad \lambda_k = 2\pi k/n$$

and the periodogram matrix is defined as

$$\mathbf{I}_{n,\lambda_k}(\mathbf{X}_n) = \frac{1}{2\pi} \tilde{\underline{X}}(\lambda_k) \tilde{\underline{X}}(\lambda_k)^*.$$

Asymptotically, the periodogram ordinates are independent Wishart distributed with asymptotic mean $\mathbf{f}(\lambda_k)$, motivating Whittle's likelihood:

$$p^W(\mathbf{x}_n|\mathbf{f}) \propto \prod_{k=1}^{n} |\det \mathbf{f}(\lambda_k)|^{-1/2} \exp\left\{ -\frac{1}{2} \sum_{k=1}^{n} \mathrm{tr}\left(\mathbf{I}_{n,\lambda_k}(\mathbf{x}_n) \mathbf{f}(\lambda_k)^{-1} \right) \right\}$$

- Periodogram covariances depend mainly on **excess kurtosis** of innovations [2].
- Whittle's likelihood **neglects this weak dependence** between periodogram ordinates.

Beyond Whittle: The corrected parametric likelihood

- **Idea:** Fit a parametric model that **mimics** [2] the weak periodogram dependence!
- **Problem:** If the parametric model is **misspecified**, spectral inference becomes wrong:

$$\mathbb{E}_{\text{param}} \mathbf{I}_{n,\lambda_k}(\mathbf{X}_n) = \frac{1}{2\pi} \mathbb{E}_{\text{param}} \left[\tilde{\underline{X}}(\lambda_k) \tilde{\underline{X}}(\lambda_k)^* \right] \to \mathbf{f}_{\text{param}}(\lambda_k) \neq \mathbf{f}(\lambda_k).$$

- **Solution: Correction** $\mathbf{Q}(\lambda) := \mathbf{f}(\lambda)^{1/2} \mathbf{f}_{\text{param}}(\lambda)^{-1/2}$ in frequency domain [1], [2], [3]:

$$\frac{1}{2\pi} \mathbb{E}_{\text{param}} \left[\left(\mathbf{Q}(\lambda_k) \tilde{\underline{X}}(\lambda_k) \right) \left(\mathbf{Q}(\lambda_k) \tilde{\underline{X}}(\lambda_k) \right)^* \right] \to \mathbf{Q}(\lambda_k) \mathbf{f}_{\text{param}}(\lambda_k) \mathbf{Q}(\lambda_k)^* = \mathbf{f}(\lambda_k).$$

time domain		**frequency domain**
$\mathbf{X}_n = (\underline{X}_1, \dots, \underline{X}_n) \sim p_{\text{param}}$	mDFT →	$(\tilde{\underline{X}}(\lambda_1), \dots, \tilde{\underline{X}}(\lambda_n))^\top =: \mathcal{F}_n \mathbf{X}_n$
		↓ Correction with $\mathbf{Q}(\lambda_k)$
$\mathcal{F}_n^{-1} \mathcal{Q}_n \mathcal{F}_n \mathbf{X}_n \sim p^C_{\text{param}}$	← mDFT^{-1}	$(\mathbf{Q}(\lambda_1)\tilde{\underline{X}}(\lambda_1), \dots, \mathbf{Q}(\lambda_n)\tilde{\underline{X}}(\lambda_n))^\top =: \mathcal{Q}_n \mathcal{F}_n \mathbf{X}_n$

Accordingly, the **corrected parametric likelihood** is given by

$$p^C_{\text{param}}(\mathbf{x}_n|\mathbf{f}) \propto \prod_{k=1}^{n} |\det \mathbf{Q}(\lambda_k)|^{-1} p_{\text{param}}(\mathcal{F}_n^{-1} \mathcal{Q}_n^{-1} \mathcal{F}_n \mathbf{x}_n).$$

Full Bayesian model

The **pseudo posterior distribution** of $\mathbf{f}$ is given by

$$p^C_{\text{post}}(\mathbf{f}|\mathbf{x}_n, \underline{\alpha}) \propto p(\mathbf{f}|\underline{\alpha}) \prod_{k=1}^{n} |\det \mathbf{Q}(\lambda_k)|^{-1} p_{\text{param}}(\mathcal{F}_n^{-1} \mathcal{Q}_n^{-1} \mathcal{F}_n \mathbf{x}_n).$$

- **Prior on f: Wavelet-based** [4] on **Cholesky components** [5].
- **Sample from** p^C_{post}: **MCMC scheme** similar to [4].

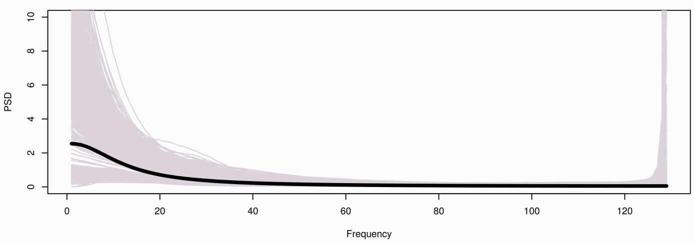
Next research steps

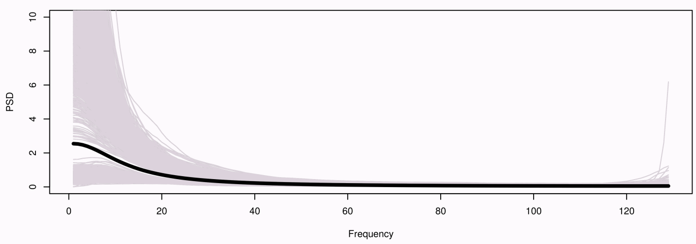
- **Show consistency:** For **Gaussian** time series [6], [1].
- **Other priors: Bernstein polynomials** [6], [1], **B-Splines**.
- **Simulation study:** Parametrize innovation **excess curtosis**.
- **Shrinkage approaches:** Cope with the course of dimensionality.

Simulations

- **Data:** Univariate AR(1)-process $X_t = aX_{t-1} + e_t$ with **Gaussian White noise** e_t.
- **Parametric working model:** $X_t = a_{\text{param}} X_{t-1} + e_t$ with prior $a_{\text{param}} \sim \mathcal{U}(-1, 1)$ and **Gaussian White noise** e_t.

Uniform Credible Intervals around pointwise median (pseudo) posterior spectral density with $a = 0.75$ for Whittle's likelihood (left) and the corrected Gaussian AR(1)-likelihood (right):





Integrated Absolute Error (IAE) of pointwise median pseudo posterior spectral density:

IAE	$a = 0.5$	$a = 0.75$	$a = 0.95$	AR(2), $a_1 = 0.75$, $a_2 = -0.2$
p^W	0.097	0.244	**3.532**	0.151
$p^C_{\text{AR}(1)}$	0.080	0.183	2.142	0.174

Uniform Credible Interval Coverage of pseudo posterior:

CI Coverage	$a = 0.5$	$a = 0.75$	$a = 0.95$	AR(2), $a_1 = 0.75$, $a_2 = -0.2$
p^W	0.986	0.895	**0.11**	0.991
$p^C_{\text{AR}(1)}$	0.992	0.999	0.912	0.996

References

[1] C. Kirch and R. Meyer. Beyond whittle: Nonparametric correction of a parametric likelihood with a focus on bayesian time series analysis. *Unpublished manuscript*, 2016.

[2] J.-P. Kreiss, E. Paparoditis, et al. Autoregressive-aided periodogram bootstrap for timeseries. *The Annals of Statistics*, 31(6):1923–1955, 2003.

[3] C. Jentsch and J.-P. Kreiss. The multiple hybrid bootstrap-resampling multivariate linear processes. *Journal of Multivariate Analysis*, 101(10):2320–2345, 2010.

[4] P. Müller and B. Vidakovic. Bayesian inference with wavelets: Density estimation. *Journal of Computational and Graphical Statistics*, 7(4):456–468, 1998.

[5] O. Rosen and D. S. Stoffer. Automatic estimation of multivariate spectra via smoothing splines. *Biometrika*, 94(2):335–345, 2007.

[6] N. Choudhuri, S. Ghosal, and A. Roy. Bayesian estimation of the spectral density of a time series. *Journal of the American Statistical Association*, 99(468):1050–1059, 2004.