A comparison of spectral estimation methods for the analysis of chaotic and stochastic dynamical systems

Jared McBride and Kevin K. Lin

Department of Mathematics, University of Arizona, Tucson, AZ 85721 USA

Introduction

In the analysis of chaotic and/or spectral time series, estimating the power spectra is often a first step. Accurate spectral estimation across a range of scales is often required for, e.g., prediction and smoothing. However, for time series with multiple timescales, slow decay of correlations, or when the range of the power spectrum is large, as often occurs in chaotic dynamical systems, this can be difficult. Here, we compare a number of spectral estimators in current use on time series generated by stochastic and chaotic time series. We also propose a general variance reduction technique, based on the method of control variates, and test its performance.

Background

Spectral Estimation methods

We discuss two families of spectral estimators.

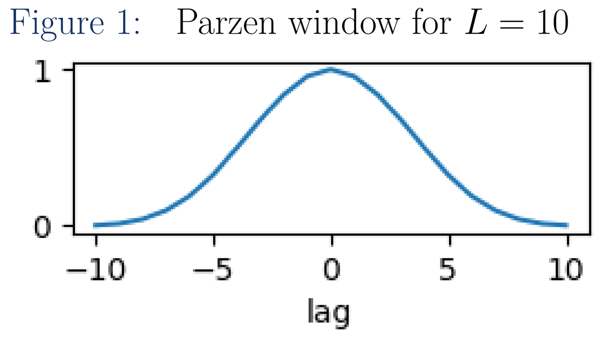
- **Bartlett-Parzen (BP).** This family improves on the *periodogram* estimate, which for a time series $X = (X_n,\ n = 1, \ldots, N)$ uses the discrete Fourier transform $\mathcal{F}$ and is $|\mathcal{F}(X)|^2/N$. It does so by applying a window function either by point-wise multiplication with the estimated autocovariance $\hat{C}_X(n)$ or (equivalently) convolution with the periodogram. The variance of the periodogram estimate does not decrease as N gets large. Windowing reduces the variance so that it vanishes with large N. The many choices of window functions are well studied (see, e.g. [Pri81]). We use the *Parzen window*, pictured in Figure 1, and a truncation parameter L that depends on the estimated autocorrelation time τ of the process ($L \approx 10\tau$). This spectral estimate for the data X is

$$\hat{S}_X^{\text{Parz}}(\omega) = \sum_{n=-L}^{L} \lambda_L^{\text{Parz}}(n) \cdot \hat{C}_X(n) e^{-2\pi i\omega n/N}$$

where

$$\lambda_L^{\text{Parz}}(n) = \begin{cases} 1 - 6(n/L)^2 + 6(|n|/L)^3, & |n| \le L/2, \\ 2(1 - |n|/L)^3, & L/2 \le |n| \le L, \\ 0, & |n| > L. \end{cases}$$

Figure 1: Parzen window for $L = 10$



- **Maximal Entropy Spectral Analysis (MESA).** This method, developed by Burg [Bur75], fits an *autoregressive model* of order p (AR(p)) to the data. A lattice of AR coefficients $a_p = (a_{0,p}, a_{1,p}, \ldots, a_{p,p})$ together with the error variance σ_p^2 are constructed iteratively as p increases from 1 to $p_{\max}$ (user specified). At each step coefficients and error variance are found that minimize the sum of the squares of the forward and backward prediction errors. From this lattice an optimal model order p is selected [MSDP21] and the spectral estimate becomes

$$\hat{S}_X^{\text{Burg}}(\omega) = \frac{\sigma_p^2}{|A(\omega)|^2} \quad \text{where} \quad A(\omega) = \sum_{k=0}^{p} a_{k,p} e^{-ik\omega}.$$

Background (continued)

Spectral Factorization, modeling and whitening

Whitening a time series, i.e. passing it through a filter to get a white noise process is a common step in optimal prediction, filtering, and smoothing.

Since the power spectrum $S_X(\omega)$ of some process X is positive-semi-definite, we can write $S(\omega) = L(\omega)L^*(\omega)$. If $L(\omega)$ is the frequency response to some linear time-invariant filter $\ell = (\ldots, \ell_{-1}, \ell_0, \ell_1, \ldots)$ then ℓ is a modeling filter for the process X, in the sense that passing a white noise process through this filter (convolution) gives a (stationary) process with spectrum $S_X(\omega)$. The inverse of this filter w is a *whitening filter*, in that passing X through it yields a white noise process. However, since

$$S_{w*X}(\omega) = L^{-1}(\omega) S_X(\omega) \left(L^{-1}(\omega)\right)^* = S(\omega)/\hat{S}_X(\omega), \qquad (1)$$

if $\hat{S}_X(\omega)$ differs from $S_X(\omega)$ the spectrum of the whitened process, the *whitened spectrum*, will *not* be flat.

Reducing the variance of spectral estimates by control variates

The method of control variates is a variance reduction technique from Monte Carlo theory. Suppose we want to estimate the expectation $\mu = \mathbb{E}X$ of some random variable X. Take n IID samples X_i of X and average

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

$\hat{\mu}$ is an unbiased estimator of μ with variance $\text{var}(\hat{\mu}) = \text{var}(X)/n$. A control variate is a mean zero random variable Y that is correlated with X. For n IID samples Y_i of Y,

$$\hat{\mu}^{\text{cv}} = \frac{1}{n}\sum_{i=1}^{n} X_i - \alpha Y_i.$$

will also be an unbiased estimator of μ, but here the variance can be controlled by α. If $\alpha = \text{cov}(X,Y)/\text{var}(Y)$ then

$$\text{var}(\hat{\mu}^{\text{cv}}) = \left(1 - |\rho_{XY}(0)|^2\right) \text{var}(\hat{\mu})$$

is minimized in α.

For spectral estimation, observe that at frequencies where the $S_X(\omega)$ is overestimated by $\hat{S}_X(\omega)$ the resulting whitening filter will *under compensate* the low power at those frequencies and the estimated whitened spectrum $\hat{S}_{w*X}(\omega)$ will be low. This indicates correlation between $\hat{S}_{w*X}(\omega)$ and $\hat{S}_X(\omega)$. So, we *take* $\log \hat{S}_{w*X}(\omega)$, which is reasonably assumed to have a small mean, *as a control variate for* $\log \hat{S}_X(\omega)$. This suggests the following procedure:

For the time series $X = (X_n,\ n = 1, \ldots, N)$,

1. Divide the full time series into K segments.
2. For each segment k, estimate the spectrum $\hat{S}^{(k)}$ and the whitened spectrum $\hat{W}^{(k)}\ \left(= \log \hat{S}_{w*X}(\omega)\right)$.
3. Take the logarithm $(\hat{S}^{(k)})_{k=1}^{K}$ and $(\log \hat{W}^{(k)})_{k=1}^{K}$.
4. Compute $\alpha = \dfrac{\text{cov}_k\left(\log \hat{S}^{(k)}, \log \hat{W}^{(k)}\right)}{\text{var}_k\left(\log \hat{W}^{(k)}\right)}$, at each frequency.
5. For $\hat{S}$ and $\hat{W}$, the spectrum and whitened spectrum of the full series, the final spectral estimate is then

$$\hat{S}^{\text{CV}} = \exp\left(\log \hat{S} - \alpha \log \hat{W}\right).$$

An Example

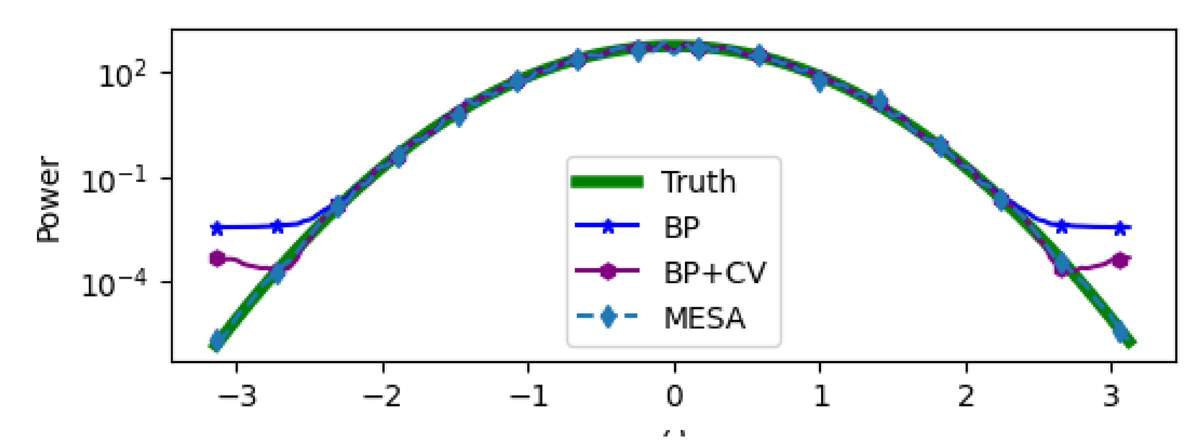
Gaussian power spectrum

Consider the process with spectrum given by the Gaussian

$$S(\omega) = \frac{1000}{\pi} e^{-2\omega^2}.$$

This process has an autocorrelation time of about $\tau \approx 5$ and the range of the spectrum is $3.7 \cdot 10^8 \approx \max_\omega S(\omega)/\min_\omega S(\omega)$. Using $3000\tau = 15{,}000$ steps we sample a realization and estimate it's power spectrum using BP, BP with control variate (BP+CV), and MESA, shown in Figure 2.

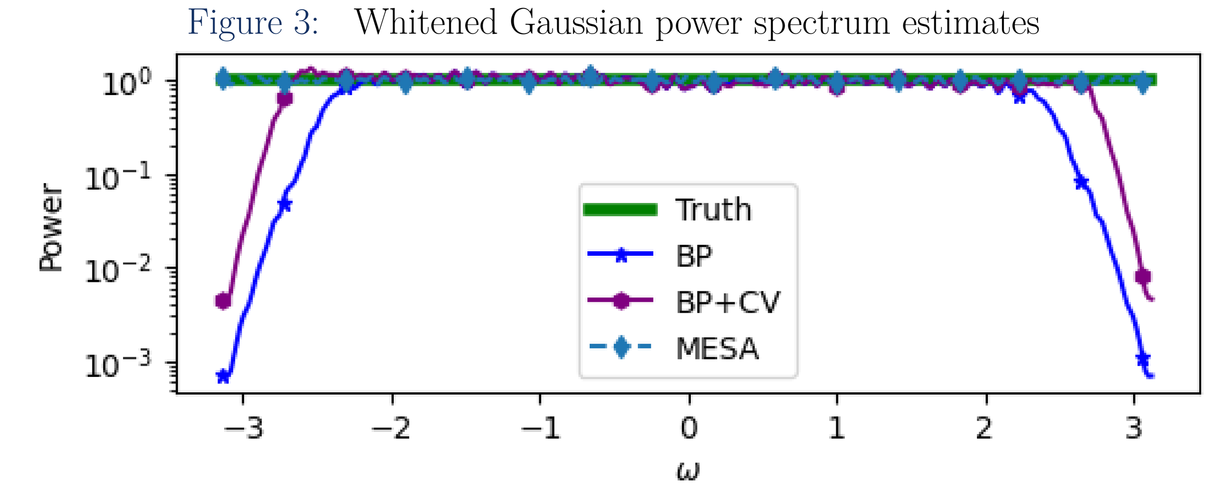
Figure 2: Gaussian power spectrum estimates



Observe that (1) MESA did well throughout the frequency range, (2) BP did poorly where $S(\omega)$ is small, and (3) BP+CV improved on BP at estimating some of the lower powers.

For each spectral estimate, we extract a whitening filter and pass the original data through each filter. MESA naturally produces a whitening filter form the AR coefficients, for the other methods we factor the spectrum. Figure 3 shows a plot of the whitened spectrum associated with each whitening filter. For comparison, each whitened spectrum is approximated using BP.

Figure 3: Whitened Gaussian power spectrum estimates



Note that MESA preforms both task well. Further note, as (1) suggests, BP and BP+CV fail to whiten in the frequencies their estimators are preform poorly.

An application

Kuramoto-Sivashinsky

The Kuramoto-Sivashinsky (KS) equation, given by $u_t + uu_x + u_{xx} + u_{xxxx} = 0$, $t \in [0, \infty)$, $x \in [0, L]$ with periodic boundary conditions, is a prototypical model of spatiotemporal chaos. Written in terms of its Fourier coefficients u_k, it becomes a system if ordinary differential equations which we solve using the usual method of fourth order exponential time differencing (ETDRK4) from [KT05]. For now we focus on the first Fourier mode u_1, whose autocorrelation time we estimate to be $\tau \approx 350$ steps. Figure 4 shows the three estimates of the power spectrum using 2000τ steps.

An application (continued)

Figure 4: KS power spectrum

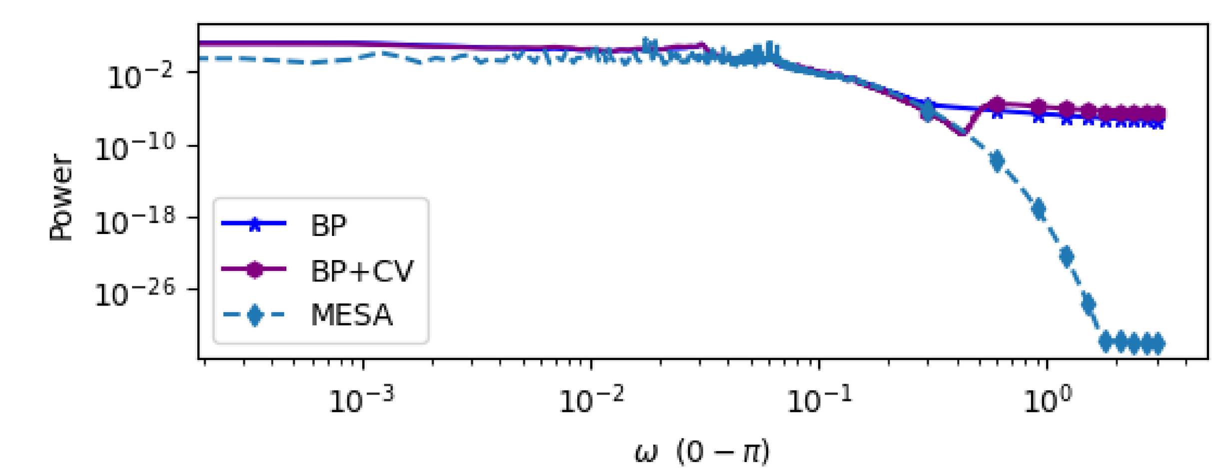
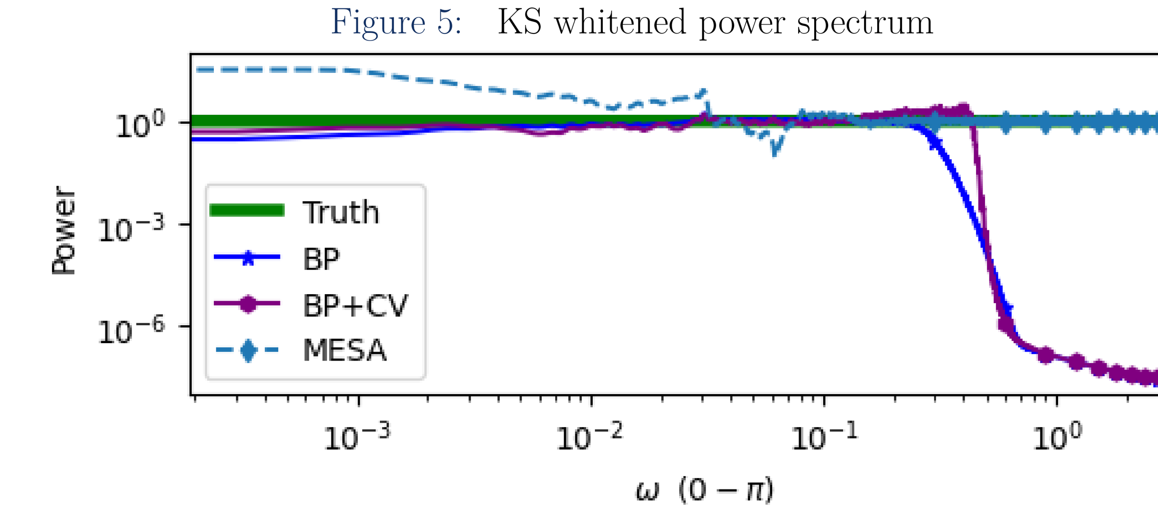


Figure 5 shows the whitened spectral estimates, as in Figure 3. We again use PB to estimate the spectra.

Figure 5: KS whitened power spectrum



MESA estimates the spectrum to have very low power ($\approx 10^{-32}$) in the high frequencies. The accuracy of this is suggested by the excellent whitening that MESA effects over those frequencies. In the low frequencies, however, Mesa does poorly. BP+CV does improve on BP in both spectral estimation and whitening.

Conclusions

In our experiments, MESA performs very well both in spectral estimation and whitening, even in the presence of very low power. We found control variates to be a simple way to improve the performance of periodogram based estimators. But overall MESA was out preformed both both PB and PB+CV. [LM22].

References

[Bur75] John Parker Burg, *Maximum entropy spectral analysis.*, Stanford University, 1975.

[KT05] Aly-Khan Kassam and Lloyd N Trefethen, *Fourth-order time-stepping for stiff pdes*, SIAM Journal on Scientific Computing **26** (2005), no. 4, 1214–1233.

[LM22] Kevin K. Lin and Jared McBride, *A comparison of spectral estimation methods for the analysis of chaotic and stochastic dynamical systems*, In preparation, 2022.

[MSDP21] Alessandro Martini, Stefano Schmidt, and Walter Del Pozzo, *Maximum entropy spectral analysis: a case study*, 2021.

[Pri81] Maurice Bertram Priestley, *Spectral analysis and time series: Univariate series*, vol. 1, Academic press, 1981.

This research was supported in part by the National Science Foundation through grants DMS-1937229 and DMS-1821286.