Spectral density estimation using P-spline priors

Response letter to Statistical Papers

Patricio Maturana-Russel, Renate Meyer

March 28, 2020

Referee #1

We would like to thank the referee for a thorough review of our paper. Below, you find a point-by-point response to your points raised. We have now carefully revised the paper taking all of your – and the other referees' – comments into account and hope that you will now find the paper acceptable for publication. For your convenience, all changes made are marked in red in the manuscript.

1. *The authors' terminology is surprising. They compute a P-spline fit to a power spectrum, but they call it a P-spline prior. The P-spline coefficients are the parameters of the model and they have a prior.*

   We have edited the manuscript and use *P-spline prior* only when referring to the prior but use *posterior distribution based on the P-spline prior* when referring to the posterior distribution or Bayesian estimate of the spectral density. We use *P-spline prior* in analogy to *B-spline Dirichlet process prior* and *Bernstein-Dirichlet process prior*. However, the referee raises a subtle point here in what frequentist and Bayesian regard as unknown parameters. In our view, the unknown parameter in this model is the spectral density function. As a prior on this spectral density function, we use a mixture of P-spline distributions. The mixture weights are hyperparameters with their own hyperprior distribution.

2. *To get a better fit to sharp peaks in the spectrum, the authors propose an adaptive scheme for knot placement. Apparently, they forgot that the differences in the penalty must change too, to correct for the variations in the distance between knots. One solution could be to use divided differences, another to use the algorithm of Wood for a derivative-based penalty.*

   Thank you very much for pointing this out. You are absolutely correct, we did use the same penalty for the unequally spaced knots as for the equidistant knots. We have rectified this and now use derivative-based penalties for the unequally spaced knots. We describe the prior specification in paragraph 4 of the section "adaptive knots" in Chapter 3. As third order derivative-based penalties do not make sense for cubic splines and also because we want to compare the results for the equally-spaced knots and adaptive knots in the simulation study, we only compare results based on first and second order difference and derivative-based penalties, respectively. Therefore we deleted the definition of the third order difference matrix in Chapter 3.

3. *A more principle approach is to model the logarithm of the spectrum. That can be combined with adaptive smoothing, as described by Wood and Fasiolo, or Rodrigues-Alvarez et al. R packages are available.*

It is debatable whether putting a prior on the log spectral density is more *principled* than putting a prior on the spectral density itself as both terms enter the Whittle likelihood function (1) and thus the prior can be put on either scale. Our P-spline prior is based on the idea of approximating a *density* function as a mixture of basis distributions on the unit interval. Thus, the original scale and not the log scale is appropriate for this particular approach even though for many applications, the log-scale might be the more natural one and we cited many papers that put a prior on the log spectral density, see our paragraph 5 of the introduction. This could be an interesting avenue to pursue in future research and see whether putting a P-spline based prior on the log spectral density has advantages over putting a prior on the spectral density. We added a paragraph in the Discussion. We have now also added the references to the alternative frequentist approaches that you suggested. Thank you for pointing these out.

We were wondering whether the referee meant "model the logarithm of the periodogram". One might consider to model the log-periodogram instead of the periodogram itself which would require using the logarithm of an exponential distribution or approximating this via a mixture of Gaussian distributions as in Carter and Kohn (1997). However, here we stick to the traditional Whittle likelihood which is contiguous to true likelihood for Gaussian stationary time series (Kirch et al 2019).

4. *The plots of spectra and fit with logarithmic are popular in spectrum analysis, but they are a bit misleading as the fitting is done on the linear scale.*

Please see our response to the previous point. Even though we fit a mixture of B-spline densities to the normalized spectral density, it is still possible to show the fit on the log scale. Large peaks are better visible using the log scale. We also want to have a direct comparison of our model fits in Figure 1 and Figure 2 to the corresponding ones shown in Edwards et al. (2019) so it makes sense to use the same scale.

5. *I took a short look at the repository. It could be made more valuable by adding the data and some example scripts.*

Data and example scripts are included in the R package `psplinePsd`.

6. *Sunspots counts are relatively familiar, but very few people will have seen a time the light curve of S. Carinae. I propose to add a plot of (a part of) the time series.*

We have added a time series plot of the first 150 observations of the *S. Carinae* variable star light intensities data in Figure 4.

7. *The term 'P-spaced knots' is confusing in the context of P-splines, P is already used for penalties. An adaptive scheme is proposed that is not exclusive to periodograms. Simply calling them adaptive knot looks adequate to me.*

This could indeed be confusing. We have changed 'P-spaced' to 'adaptive' knots.

8. *That there is no consensus on the optimal number of knots for P-splines (page 10) has a simple explanation: it is irrelevant. Ruppert's recommendations have little value.*

   We see this as a rough guideline only.

9. *I noticed some language problems*

   Thank you for pointing these out. These have been corrected.

Referee #2

We would like to thank the referee for a thorough review of our paper. Below, you find a point-by-point response to your points raised. We have now carefully revised the paper taking all of your – and the other referees' – comments into account and hope that you will now find the paper acceptable for publication. For your convenience changes are marked in red in the manuscript.

1. *The entire Section 2 is dedicated to the B-spline prior, which were already effectively introduced in Edwards et al. (2019). I would significantly reduce the length of this section.*

   We have deleted the detailed description of the BspDP prior but kept the definition of the B-splines as these are required for the definition of the prior based on the P-splines.

2. *The proposed method deals with estimating one single spectral density. However, many existing paper already deal with this. Can the proposed methodology be extended to multiple or multivariate time series?*

   This could be a very interesting avenue for future research and we have added a paragraph in the Discussion suggesting a potential approach. However, a multivariate extension is outside the scope of this paper.

3. *A Metropolis step is used for v, but a data augmentation step based on Polya-Gamma auxiliary variables will allow for a full Gibbs sampler. See "Bayesian Inference for Logistic Models Using PolyaGamma Latent Variables" (2013) of Polson, Scott and Windle.*

   We had already pointed this out in section "posterior computation" and had cited the paper by Polson et al. (2013).

Further Changes

We have changed the title from *Spectral density estimation using P-spline priors* to *Bayesian spectral density estimation using P-splines*.

We have added the keyword *Bernstein-Dirichlet process prior*.

In the introduction, we have added references to relevant papers published in *Statistical Papers* that deal with the penalized spline approach for time series (these are Krivoboka et al (2006) and Wegener and Kauermann (2017)).