1 BA2107-012: "Bayesian Functional PCA via Variational Message Passing"

Summary. This work studies functional principle component analysis (FPCA), a technique that, analogous to classical PCA, can be used to reduce dimensionality and identify modes of variation but, unlike classical PCA, is used when the dataset is composed of curves that are independent realizations from a stochastic process. A Bayesian hierarchical model for estimating the eigenvalues and eigenfunctions of the principle components approximation is proposed and these values are estimated through a variational message passing (VMP) algorithm. In particular, this work proposes an FPCA extension to the VMP framework for variational Bayesian inference introduced in (Wand, 2017). Compared to MCMC, variational Bayes for FPCA is much more computationally efficient, and generally Bayesian methods are more efficient than standard approaches that smooth and estimate a covariance surface.

Review. I believe the work to be sufficiently novel for publication and of interest to the Bayesian Analysis readership. The paper is very clearly written and the mathematical details all appear to be correct. I have a few comments about how the authors could improve the paper prior to publication. One weakness of the paper is that the scope of the paper appears a bit limited in that the work details an implementation of VMP for Bayesian inference on a specific hierarchical model for FPCA, instead of a more general VMP framework. I comment in more detail below about how I believe that the authors should include some additional discussions around how easy (or not) it would be to extend the applicability of the proposed ideas.

Detailed Comments. I enumerate my comments in the following.

1. **Choice of L.** Throughout, the authors assume that the number of principle components to be used in the approximation, L, is determined beforehand. In practice, how should this be done? Would a practitioner run the full procedure for many choices of L and afterwards compare the estimates in some way? Is it clear that if the procedure is run for both L and $L+1$ principle components that the top L eigenvalues and eigenfunctions found in both implementations would be approximately the same? I presume that the complexity analysis in Section 5.2 is for a fixed L, but how does the relative computational efficiency of VMP vs. MCMC depend on L? Presumably VMP would outperform MCMC more as L grows, and it may be worth commenting on this if it is true.

2. **Generality of results.** The VMP algorithm proposed is specifically approximating posteriors from the full Bayesian hierarchical model specified in Equation (2.6). Could the authors comment on how easy it would be to apply these results if one wanted to make reasonable modeling changes to (2.6). For example, how dependent are the results on the use of Gaussian and inverse-χ^2 distributions or how easy would it be to extend the results if one added additional levels int eh hierarchy? The authors include some small comments on this in Section 2.1 about how the specific construction facilitates arbitrary non-informative priors, but I think a larger discussion about the assumptions for the specific model in (2.6)

and the limitations that come with it, would be useful.

3. **Convergence.** It is assumed in the article that the message passing algorithm will converge. Perhaps the authors could add a few comments on why that is reasonable to assume.

4. **Code and software.** Do the authors plan to share the code for their simulations and/or provide software for the proposed method? Surely, this would be useful for the community.

5. **Some minor comments.**

   - If it is easy to do, I would suggest making Figure 2 and Figure 3 a bit larger.
   - There is a typo on page 18: "Each vector of principle component scores WAS simulated..."
   - I think something is wrong with the Section 7 sentence, “This study could be extended to other functional data models, such as function on scalar or vector regression models, that are yet to be treated under a VMP-based mean field variational Bayes approach."