Report on the article EJS2002-023RA0, titled:

Random weighting to approximate posterior inference in LASSO regression

by T. L. Ng & M. A. Newton

Firstly, I apologise for the delay in sending my report.

I am not an expert on Bayesian inference, so I may not appreciate all aspects of your article. All the results are based on randomly weighted version of the Lasso estimator, and specially I am interested in the sign-consistency results for $\widehat{\boldsymbol{\beta}}_n^w$. However, the presentation in this article makes it difficult for me to appreciate the use of your results. Some of my queries are stated below.

1. The article does not seem to have any numerical results, which I find a bit surprising. Can you elaborate?

   Your aim is to use this WBB scheme to approximate certain posterior laws, which itself can be found using more direct approaches (Gibbs sampling?). So, it seems reasonable to provide some numerical results.

2. If I consider the estimator in (1.4), and repeatedly compute it B times, it is likely that the each time the sign vector of $\widehat{\boldsymbol{\beta}}_n^w$ will vary. So, as an user, how do I interpret Theorem 3.3? Which sign vector does the user decide to use? It does not seem, that you aim to *average* out the B different estimators, as is done in bootstrapping. Please clarify this issue.

   I can imagine that in case of the asymptotic distributional result (Theorem 3.2), you can use *averaging* over B samples, to approximate the posterior limit law mentioned in Section 4.1. Is there a similar interpretation about model selection?

   I think, you should more clearly state the implications of random weighting in Theorem 3.3 and the earlier Proposition.

   It would be also helpful if you discuss the use of the consistency result (Theorem 3.1) in context of how it would be usable for posterior inference? I assume there is no fixed $\boldsymbol{\beta}_0$ in the Bayesian approach, so what are you trying to capture?

3. The Introduction is very complicated and does not speak of the main goal in simple language.

4. There are some papers on perturbation bootstrap for penalized regression estimators (mainly Lasso and variants), by Das & Lahiri (2019, Biometrika) and Das, Gregory & Lahiri (2019, Ann. of Stat). You may want to compare your methodology with theirs and see what is new in your case.

5. In Theorem 3.2, why do you center at the OLS? I expect to see centering at $\boldsymbol{\beta}_0$. Your result mimics a bootstrap distributional consistency result, but, the target here is $\boldsymbol{\beta}_0$ itself.

6. It is not clear to me how you choose W_i's? Obviously, I see $\mathbf{E}(W_i) = 1$ is a requirement while proving results, but other than that it seems any positive random variable with certain particular type of moments should do the job? Is it true? One needs a more refined analysis of the WBB procedure to judge its true benefits.

7. I do not understand the utility of introducing weights on the penalty term? Neither, I get the reason for writing Section 5 separately? The results in earlier sections can be a special case of the latter?

8. I have not checked the proofs, but while looking into them, I have some questions. In the middle of page 21, in the proof of Theorem 3.1, why do you write, *Conditional on the data*? The term on the next line $X'\epsilon/n$ does not involve weights. Similarly, in the penultimate convergence statement on that page, it seems that quantity is a joint function of both (ϵ_i, W_i) and hence you need a convergence under the joint probability distribution, not a conditional convergence statement. Could you please check.

9. Finally, the entire goal of the article is to approximate certain posterior laws as you mention briefly (in Section 4). So, your asymptotic results should be based on the same framework, and that means not assuming a *fixed* $\boldsymbol{\beta}_0$ parameter in (1.2). Neither you can assume classical sparsity settings? Probably, this issue needs explanation in depth.

   Otherwise, I do not think your results are very interesting, as simply putting random weights makes an easy problem more harder. I think the paper needs much more thought and clarity in presentation.