Geosci. Model Dev. Discuss.,
doi:10.5194/gmd-2016-5-AC3, 2016
© Author(s) 2016. CC-BY 3.0 License.

Geoscientific Model Development Discussions

Interactive comment on "A tuning-free method for the linear inverse problem and its application to source term determination" *by* O. Tichý et al.

O. Tichý et al.

otichy@utia.cas.cz

Received and published: 9 September 2016

We would like to thank you for providing us with detailed reviews of our paper. We have considered all the comments and notes and we are glad that we can submit a revised version of our paper. In the following text, we will respond to all comments.

In the paper the authors propose to apply the Variational Bayesian methodology to estimate the tuning parameters of the objective function given in Eckhardt et al. (2008). The authors describe the method and the algorithm to compute such tuning parameters. Then they show the performance of the proposed algorithm using a synthetic dataset, and the ETEX dataset. Its performance in the ETEX dataset is compared with the performance of other state-of-the art algorithms.

1 Specific Comments:

1. page 3, line 20: For clarity, the optimization problem can be written including the non-negative constrain for x, x = argmin_x (J1 + J2 + J3) s.t. x >= 0

We agree that explicit statement is easier to follow and added a new equation (5) to emphasize it. It also makes comparison with the probabilistic approach easier.

2. page 4, line 16: The Gaussian assumption is a good choice if, in fact, the errors in the model are Gaussian. Otherwise, this can cause deviations in the estimation. The authors should justify why the Gaussian assumption is a reasonable one in this case. Also, this particular regularization enforces smoothness in the solution. It should be mentioned that it is not suitable for releases generated, for example, during explosions

Thanks for this suggestion. Indeed, suitability of the prior to abrupt changes was one of our primary motivation. It is now explicitly mentioned in the text. It is hard to justify the assumption of Gaussian noise since the residue is of complex nature here. We have added a sentence that this choice is motivated by tractability of its inference.

3. page 5, line 1: gamma(x) should be defined more precisely.

The γ term is now defined in text as logarithm of the characteristic function. Since this is only a minor illustrative point, we wanted to avoid the use of formal mathematical approach. The main point is that (9) and (5) are equivalent problems, which is now explicitly stated.

4. page 5, line 16: why do the authors assume that the variance for all the measurements is the same? Is it not more reasonable to define w as a vector instead of a scalar?

It is important to distinguish if the covariance matrix of the observations is known or not. If it is known, it can be used to transform the problem into isotropic noise. The

explicit equations for this transformation have been added to the paper.

A more demanding case is if we assume that the observation variance is unknown. Assuming completely independent variance for each observation leads to over-parametrization which can be addressed only by introduction of further restrictive assumptions. It is an interesting topic of future research, however, in our experiments the assumption of same variance was found to be quite reasonable and robust choice.

5. page 5, line 20: the authors should explain why the gamma distribution is chosen to model w.

The gamma distribution was chosen due to conjugacy to the Gaussian distribution so the prior and the posterior distribution have the same form in the Variational Bayes procedure which is beneficial for computational reasons. We commented this in the text a we added the citation for this model.

6. page 5, line 25: explain in more detail why this particular relaxation has been chosen.

This relaxation was made the preserve the tri-diagonality of the matrix $\Sigma_{\mathbf{x}}$ and using this model, each diagonal can be modeled separately. We added the comment into the text.

7. page 6, line 16: The authors should explain why they conclude that a wider range for that parameter l_j is not recommended. What are the effects if the range is wider?

We agree that this formulation need to be specified. We reformulated the paragraph and discussion on these prior constants was added to the paper.

8. page 6, line 23: does conditional independence make sense here? The authors should explain why they are making that assumption.

Indeed, this assumption seems arbitrary. Its motivation is primarily simple solution of

the implied variational problem. However, experience indicate that estimation of linear models under this assumption yields results very close to much more expensive MCMC approaches. We added few references to relevant literature.

9. page 7, line 11: the derivation of the parameters is not in the Appendix B. Only the definition of the parameters is given. An explanation on how the authors arrive there is recommended.

Indeed, teh parameers are not derived in Appendix B. Derivation of the parameters is long and routine procedure. We have now added reference to a book where it is described and replaced he word derived by 'given'.

10. page 8, line 1: since local minima exist, good initial points should be taken, or several initial points may be considered.

Indeed, the initialization of the LS-APC algorithm needs to be selected. We extended description of the proposed initialization in Algorithm 2. We have good experience with this choice in our experiments, however, it is not the only possibility. Much more advanced search strategies for global solution using any global optimization method are possible.

11. page 8, line 8: The authors should also comment on the convergence guarantees of the algorithm.

Actually, the VB algorithm converges only to local extreme as it is now cited in the text.

12. algorithm 2: Step 2 is not clear. Is <x>^(i) equal to <x>? What is exactly the analytic expression for <x>^(i)?

Thank you for this remark, the description of the iteration indexing $< \mathbf{x} >^{(i)}$ is certainly needed. We clarify this when introducing the initialization of the algorithm. All required moments are given in Appendix A and detailed references to particular equations have been added.

13. page 8 , line 12: The authors should provide the condition number of the matrix M.

We agree and we added the condition number of the matrix M into the text. Since the matrix is not ill conditioned, we also reformulated the sentence.

14. Figure 2: x-axis labels are missing

The labels are added now.

15. page 8, line 14. It is not clear if the three 'sets' refer to three different synthetic experiments, or not.

In the synthetic experiment, the same matrix M was used. The difference is in the realization of noise, i.e. in vector $\mathbf{y}$. We added a comment into the description of the data.

16. page 8 : in the experiments with the synthetic dataset, it would be interesting to compare the estimated parameters w.r.t their truth value, i.e. as in the case for the source term, a red line representing the ground truth could be added in the other plots. It will give a more precise idea of the quality of the estimate of the parameters.

Derivation of the ground truth for these parameter is not very clear. First, there is more ways how to define the "true prior covariance matrix". One possible choice is the empirical covariance from the single realization of the parameter. Then, the matrix has rank=1 and decomposition into choleski factors is not unique. There is a number of matrices giving the same covariance. Probably the most illustrative example could be display of the implied covariance as an image, such as in Fig. 1, bottom row. However, we are not sure that it would be understandable.

17. Matlab code: the Matlab code provided reproduces the results given by the LS APC algorithm. To facilitate the reproducibility of all the results in the paper, the authors should include the code used to generate the results given by the

other algorithms as well.

We agree that it will be beneficial to include also other algorithms into the code posted online. We include our implemnetation of the LASSO and Tikhonov algorithm into the MATLAB package provided online. Evaluation of the remaining methods was done using their code. The RegClean algorithm by Martinez-Camara et al., (2014), can be downloaded as a supplement to their paper (published in GMD). The algorithm from Eckhardt et al., (2008), is not implemented in MATLAB (see Eckhardt paper for details) so it would not be consistent to provide it together with the LS-APC algorithm. Therefore, we are not redistributing these algorithms.

Interactive comment on Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-5, 2016.

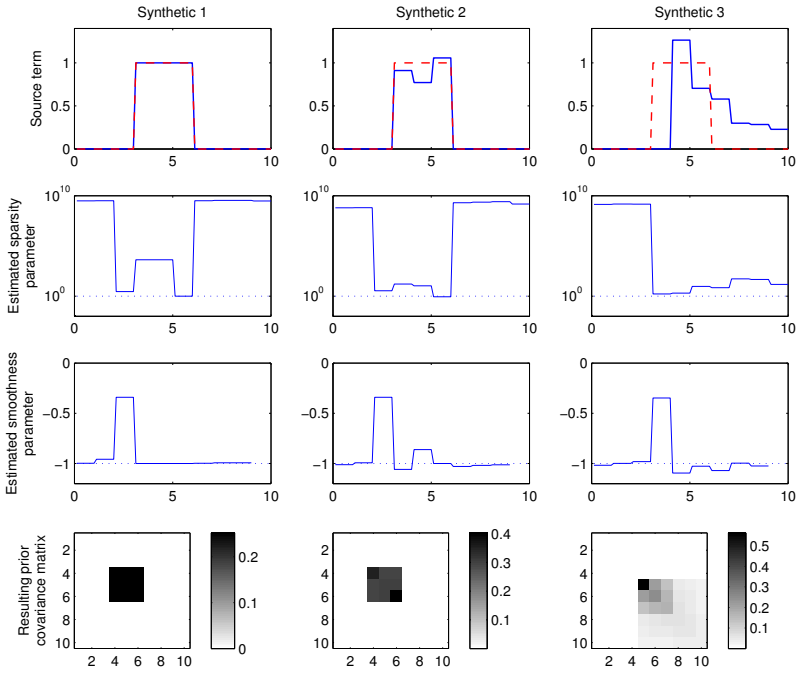
Synthetic 1
Synthetic 2
Synthetic 3
Source term
Estimated sparsity parameter
Estimated smoothness parameter
Resulting prior covariance matrix

Fig. 1.