Geosci. Model Dev. Discuss., 5, C1304–C1311, 2013
www.geosci-model-dev-discuss.net/5/C1304/2013/

Geoscientific
Model Development
Discussions

Interactive comment on "Technical Note: Improving computational efficiency in large linear inverse problems: an example from carbon dioxide flux estimation" *by* V. Yadav and A. M. Michalak

V. Yadav and A. M. Michalak

michalak@stanford.edu

Received and published: 21 January 2013

We thank the reviewer for providing insightful comments on our paper. In light of the comments of the reviewer, we have revised the literature review sections of the paper and also decided to submit sample FORTRAN code substantiating the theoretical performance of the proposed algorithm for the matrix multiplication of **H** and **Q**. Responses to the comments of Reviewer are included below, with original comments in italicized text.

Overview

The manuscript by Yadav and Michalak presents algorithms for calculating matrix prod-

ucts that commonly arise in geophysical inverse problems. The subject matter is suitable for GMD, and the theoretical results and sample case appear to be of value in terms of significant computational advances over direct method. I do however have two main concerns. First, the presentation of the results could be better framed with regards to the computational literature. Second, based on tests using the sample code provided, the efficiency of the new algorithms over direct methods appears to occur under a more limited range of conditions that indicated by the theoretical analysis in the manuscript. Below I expand upon these points, as well as a few other minor comments, that would be best addressed prior to publication.

Response: We have modified the presentation of the literature review as described in the responses below. In addition, we want to clarify that the sample MATLAB code was only provided so that it can aid in improving the understanding of the two algorithms proposed in this research. The theoretical analysis cannot be replicated in MATLAB due to: (1) the use of highly optimized multithreaded external libraries (Basic Algebra Subroutines) for performing matrix multiplication, and (2) automatic memory management (e.g., allocation and reallocation of memory). This gives an impression of poor performance of the algorithm in certain practical situations. To allay these fears we have now also provided a completely serial FORTRAN code for performing matrix multiplication of **H** and **Q** matrix. The results using the FORTRAN code clearly align with the computational efficiencies described in the manuscript and we encourage readers to test the FORTRAN program with various dimensions of **H** and covariance matrices **D** and **E**.

General comments

1. *I appreciate the authors have provided MATLAB code to support their work. I tried to replicate the settings used for the case study based on my understanding of the variables as presented in the manuscript. For example, the manuscript states that* $r_t = m_s$*, and that* $m_s = 2635$ *for the case study. Similarly,* $pq = m_t$*. Assuming*

a square system, using r $= t = p = q =$ 52 would seem to be a reasonable approximation of their test case described in section 1.1. I then tested this with n $=$ 8503 for the HQ test in HQ_HQHt.m. Repeating the experiment several times showed that the indirect method was slightly faster on average, by one second out of ten. How is this consistent with the claims of several orders of magnitude increase in efficiency in the manuscript? Granted my tests conflate floating point efficiency with memory efficiency, but I would have expected greater speedups for the indirect approach given the conclusions of the manuscript.

Response: Please see response above. We encourage readers to use the same test case as described above to assess the performance of the algorithm using the newly provided FORTRAN code.

1. *Further, upon additional tests with the provided code, trying setups where r $=$ t and p $=$ q, it seems that if m$_s$ is smaller than m$_t$, then the indirect approach is less efficient, not more efficient, and vice versa. For example, if r $= t =$ 100 and p $= q =$ 10, then the indirect method was on average 30% faster, but if r $= t =$ 10 and p $= q =$ 100, then the direct method was faster by a factor of three. These were all using n $=$ 8503, although as mentioned in the manuscript, n does not impact the ratio of floating point calculations used in the direct relative to indirect approaches. Therefore, I encourage the authors to consider cases beyond their one test case, and to present the efficiency of their algorithm in terms of the magnitude of ms versus m$_t$. At present it appears that either their new approach is only of benefit when m$_s$ <m$_t$, or there is something that is not clear in the manuscript or provided code.*

Response: Please see responses above. Generally, we find a significant improvement in the time necessary for performing matrix multiplication of **H** and **Q** when using the indirect method for (p=q, r=t) p>3, r>10, and n>10.

1. *There isn't any background on literature regarding computational algorithms in the introduction. Instead, the introduction focuses exclusively on the literature related to applications. While the latter is of value given the GMD audience, the former should also be included. Therefore, discussion currently included on page 3334 of other algorithms such as Strassen's, Coppersmith-Winograd, and Saibaba and Kitanidis should be introduced and explained.*

Response: We agree with the reviewer, and have moved the literature review discussing the computational complexity of the existing matrix multiplication algorithms to the start of the manuscript.

1. *The claim that the matrix calculations HQ is the first bottleneck of inverse problems (p3328, line 8) also does not seem to be substantiated. The bottleneck in many approaches is the calculation of H itself, hence the development of gradient based approaches (e.g., 4D-Var) that do not require explicit construction of the full transport model Jacobian. The applicability of the results in terms of improving the computational efficiency of large linear inverse problems is thus perhaps more limited than the title would imply. For example, are results relevant most for specific approaches such as geo-statistical, Kalman filtering or variational methods?*

Response: We agree that for large linear inverse problems the computation of **H** itself is a bigger problem than the matrix multiplication **H** and **Q**, and did not intend to suggest otherwise. In the revised manuscript, we have mentioned this in the introduction. We have left the title of the paper unchanged, however, because in cases when **H** is available for a large inverse problem, the matrix multiplication of **H** and **Q** and computation of uncertainty are two major bottlenecks that can be overcome by the algorithms proposed in the manuscript.

1. *Following on the previous comments, since it appears from the structure of the original manuscript that consideration of other methods in the literature was somewhat of an afterthought, it isn't evident that the authors have done a thorough literature review of the computational aspects of their problem (perhaps they have, it just does not come across in the paper). For example, a recent paper by Singh et al. (GMD, 2011) presents work on efficient representation of covariance matrices using Kronecker products, including common linear algebra manipulations such as matrix-vector products and matrix inverse. How is the work of Singh related to the approaches presented here? At the very least it seems worth mentioning, particularly given the concluding sentence of the manuscript regarding the need for advances in the numerics of other aspects of geophysical inverse problems. Another work to consider that used Krockecker product representation of covariance matrices is Meirink et al. (2008). Another reference that may be of interest is: Sun, Y., Li, B., and Genton, M. G. (2012), "Geostatistics for large datasets," in Advances And Challenges In Space-time Modelling Of Natural Events, Springer, Vol. 207, Chapter 3, 55-77.*

Response: We had reviewed the paper by Singh et al. before submitting our paper for publication in GMD. However, since there were significant differences between our paper and Singh et al. we did not cite their work.

Our paper presents: (a) an efficient algorithm for performing matrix multiplication between an arbitrary matrix and a matrix expressed as a Kronecker product with application in inverse problems, and (2) an algorithm to efficiently compute uncertainty at aggregated scales in inverse problems where **H** has been pre-computed. On the other hand, the paper of Singh et al. focuses on methods to represent a background spatial covariance in a 4D-VAR data assimilation method (not spatio-temporal!) using a Kronecker product. On the basis of reviewer's comment, we once again looked at Singh et al. and especially looked at the method for computing matrix vector products in section 4.4. After thoroughly reading the paper we understand that the paper of Singh et al.

discusses efficient ways to compute matrix-vector products in the context of the matrix structure presented in their paper (see step 1 and step 4 in equation 29 on page 306), but their method cannot be generalized for the multiplication of any two arbitrary matrices where one can be expressed as a Kronecker product of two smaller matrices. Additionally, they do not provide any comparison of the theoretical performance of their method to other matrix multiplication methods. However, because they do use Kronecker product properties to efficiently compute matrix-vector products, we have cited their work in the revised manuscript in response to the reviewer's suggestion. Due to some apparent inconsistencies in the presented equations in Singh et al. we were not able to thoroughly analyze the computational efficiency of their method. For example, we do not know whether [.] in their paper represents a matrix multiplication or a Hadamard product, as both would lead to positive semi-definite covariance matrices (compare equations 19 and 20 and equation 7 where space is used to show matrix multiplication).

Based on the reviewer's suggestion, we have also cited the work of Meirink et al. in the revised manuscript, and describe in the introductory section how Kronecker product formulations have been used to construct covariance matrices in inverse problems. Lastly, we were aware of the computational methods mentioned in Sun et al., but had decided not to refer to their work as the focus of our work is on developing efficient algorithms for matrix multiplication and uncertainty quantification, whereas their work primarily discuses methods for modifying covariance matrices for solving large-scale Kriging problems. However, on the advice of the reviewer, we have also added a citation to this work in the concluding section of the revised manuscript.

Specific comments

1. *It is a bit lazy to reproduce verbatim the sentences from the introduction in the abstract*

Response: We have made the necessary edits.

1. 3327.5 I believe the 2nd edition of Aster is copyright 2013. Also, it is a bit odd to use this as a reference for hydraulics and remote sensing; as I recall most of the examples are taken from tomography and seismology.

Response: We have fixed the copyright date for the Aster reference. We had listed Aster as a reference for approaches aimed at estimating subsurface structure, but have made the data used to do so clearer in the revised manuscript.

1. *3327.10 Check the year on the Ciais citation.*

Response: We have corrected the citation.

1. *3328 It could be useful to more clearly explain what is meant by aggregation error; Meirink et al. (2008) would be a suitable reference.*

Response: In the revised manuscript, we have clearly explained the meaning of aggregation error and also referenced the work of Meirink et al. (2008).

1. *3328.15 Not a big deal, but curious why here the penalty term in the objective function is described as a "prior sp" whereas the previous works by Gourdji emphasized that explicit prior fluxes were not used, and that this term was X_.*

Response: We used the more common Bayesian Inversion approach (which involves sp) to demonstrate the applicability and computational performance of the algorithms proposed in this research, whereas the Gourdji et al. paper used a geostatistical inversion framework (which involves Xβ). The computational gains are equivalent for both

approaches, but we chose to present the algorithms using a more common approach in order to increase the paper's accessibility to readers.

1. 3329.23 *Could be useful to mention here the coding language.*

Response: We want to avoid mentioning the coding language as the MATLAB scripts and the FORTRAN program submitted with the revised manuscript are just a medium to demonstrate the application of the proposed algorithms. The algorithms themselves are independent of programming languages, and we are of the view that the readers of the manuscript should not get any impression that these algorithms can only be implemented in MATLAB or FORTRAN.

1. *3333 It is a bit odd to present the efficiency for the specific case as a ratio (eq 13), but later for the generic case as two equations (14 and 15). Why not just say for the latter that the indirect approach is a factor of 2n1/2 faster for large n? Overall, considering how the efficiencies are presented in both the results and abstract, at the moment there are ratios, percentages, order of magnitude comparisons,. . . the paper may befit from picking a single metric for comparison and sticking without throughout.*

Response: We agree with the reviewer. In the revised manuscript we have only presented the ratios through equations and the computational savings in floating point operations for the example case study.

Interactive comment on Geosci. Model Dev. Discuss., 5, 3325, 2012.