Reduced Cost Construction of Jacobian Matrices for High-Resolution Inverse Modeling

Hannah Nesser[1] (hnesser@g.harvard.edu), Daniel Jacob[1], Joannes D. Maasakkers[2], Melissa Payer Sulprizio[1], Yuzhong Zhang[1], Tia Scarpelli[1]

[1]Harvard University, Cambridge, Massachusetts, United States, [2]SRON Netherlands Institute for Space Research, Utrecht, the Netherlands

The computational cost of analytic inversions can be decreased by optimizing only state vector elements with sufficient information content.

An analytic inversion of GOSAT observations over North America in July 2009 demonstrates the concept: the low-rank solution (right) accurately captures the true solution (center) in areas with high information content (left). The low-rank solution requires only ~300 model runs to optimize ~450 grid cells, while the standard solution requires ~2,000 model runs to optimize ~2,000 grid cells.

Introduction

- The Tropospheric Monitoring Instrument (TROPOMI) provides daily, global column methane concentration retrievals in 7 x 7 km^2 pixels with a ~5% success rate.
- Methane emissions can be inferred by inverting these observations. A Bayesian inversion accounts for errors. When the forward model F is linear, there is a closed-form solution for the posterior emissions estimate $\hat{\mathbf{x}}$, error $\hat{\mathbf{S}}$, and information content $\mathbf{A}$.
- The computational cost of the inverse solution is limited by the dimension n of the emission (state) vector because of the cost of characterizing the linear dependence of observations on emissions, given by the Jacobian $\mathbf{K}$.
- We propose an iterative method of constructing $\mathbf{K}$ that reduces computational cost by finding the model response to the dominant directions of the observing system's information content rather than individual state vector elements.

Method

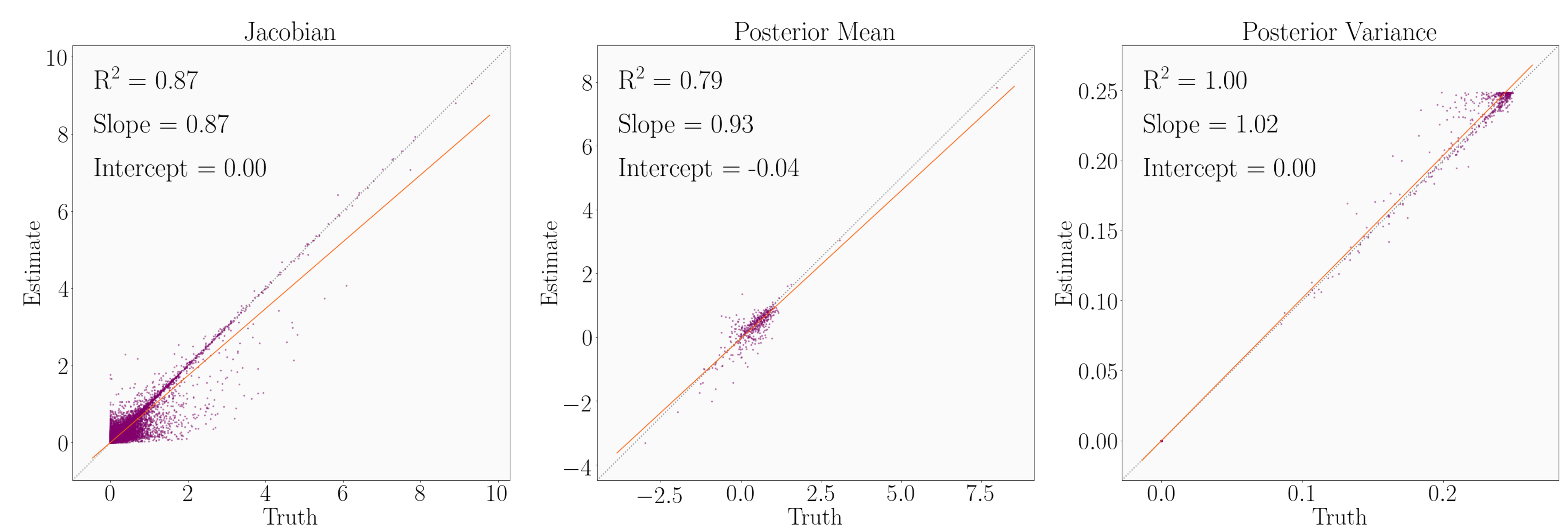
- $\mathbf{K}$ is typically built by perturbing each of the n state vector elements in the forward model F, requiring n+1 model runs.
- Not all state vector elements have equal information content. Perturbing the dominant directions of information content reduces the number of model runs needed to build $\mathbf{K}$.
- $\mathbf{A}$ is a function of $\mathbf{K}$, so we iteratively update $\mathbf{K}$ from an initial estimate.

Algorithm

1. Initialize the Jacobian $\mathbf{K}_0$.
2. Eigendecompose $\mathbf{S}_a^{-1/2}\mathbf{A}_0\mathbf{S}_a = \mathbf{W}^\top\mathbf{\Sigma}\mathbf{W}$, where $\mathbf{S}_a$ is the prior error covariance, $\mathbf{A}_0$ is the averaging kernel computed with $\mathbf{K}_0$, the columns of $\mathbf{W}$ are eigenvectors, and the diagonal of $\mathbf{\Sigma}$ contains the eigenvalues.
3. In the forward model F, perturb the first k eigenvectors of $\mathbf{A}_0$, given by the columns of $\mathbf{\Gamma}^\star = \mathbf{S}_a^{1/2}\mathbf{W}_k$, with $\mathbf{W}_k$ the first k columns of $\mathbf{W}$.
4. Transform the k-dimensional Jacobian given by $\mathrm{F}(\mathbf{\Gamma}^\star)$ to the original dimension with $\mathbf{\Gamma} = \mathbf{W}_k\mathbf{S}_a^{-1/2}$ so that $\mathbf{K}_1 = \mathrm{F}(\mathbf{\Gamma}^\star)\mathbf{\Gamma}$.
5. Repeat 2 - 4 with $\mathbf{K}_1$. Iterate until convergence.

Results

- An analytic inversion of GOSAT observations over North America in July 2009 is used as proof-of-concept.
- The Jacobian is initialized using a mass balance approach.
- The initial patterns of information content are improved by perturbing the patterns that correspond to 80% of the information content (~100 model runs).
- To capture the improved patterns, ~200 patterns are perturbed in the second iteration.
- Visual inspection shows that the first hundred patterns of the first and second estimate are similar and that the spectrum of information content becomes smooth after a second iteration, demonstrating convergence.
- Where the averaging kernel is less than 0.005, posterior values are set to the prior. The correspondence of the Jacobian and posterior to the truth where optimization occurs is shown above and below.



Conclusions

- Low-rank Jacobians decrease the computational cost of analytic inversions by decreasing the number of model runs by an order of magnitude, supporting high-resolution inversions.
- Low-rank Jacobians solve for emissions and error in areas with high information content. Areas with low averaging kernel values are set to prior estimates.

References and Acknowledgements

1. Maasakkers, J.D.; Jacob, D.J.; Sulprizio, M.P.; Scarpelli, T.R.; Nesser, H; Sheng, J; Zhang, Y; Hersher, M; Bloom, A.A.; Bowman, K.W.; Worden, J.R.; Janssens-Maenhout, G; Parker, R.J. *Atmos. Chem. Phys. Discuss.* In review. 2019.

2. N. Bousserez and Daven Henze. *Roy. Meteor. Soc.* 2017.