Supplement to "Atmospheric inverse modeling with known physical bounds: an example from trace gas emissions"

Scot M. Miller, Anna M. Michalak, and Patricia J. Levi

This supplement describes in greater detail the multiple-try Metropolis-Hastings algorithm and the Gibbs sampler implementation.

1 The multiple-try Metropolis-Hastings

This algorithm (Liu et al., 2000) first requires the generation of an unconstrained unconditional realization, denoted $\boldsymbol{s}_{uu}$. The realization for step l, denoted $\boldsymbol{s}_{uu,l}$, is created by applying a modification to $\boldsymbol{s}_{uu,l-1}$. The modification to the previous realization is provided by what is known as the jumping distribution $T()$. This distribution should create new realizations that are sufficiently different from the previous one such that the algorithm effectively samples the entire probability space. However, the jumping distribution should avoid creating subsequent realizations that are so different such that $\boldsymbol{s}_{uu,l}$ gets rejected by the algorithm (e.g., Chib and Greenberg, 1995).

The jumping distribution used here requires taking the Cholesky decomposition of $\mathbf{Q}$:

$$\mathbf{Q} = \mathbf{C}\mathbf{C}^T \tag{1}$$

The distribution $T()$ can be chosen in any number of ways (e.g., Chib and Greenberg, 1995), but we generate new unconditional unconstrained realizations as follows (where $\boldsymbol{u}$ is a random vector with distribution $\mathcal{N}(0,1)$):

$$\begin{aligned} \boldsymbol{s}_{uu,0} &= \mathbf{C}\boldsymbol{u} \\ \boldsymbol{s}_{uu,l} &= 0.9\boldsymbol{s}_{uu,l-1} \ + \ 0.5\mathbf{C}\boldsymbol{u} \end{aligned} \tag{2}$$

The multiple-try Metropolis-Hasting with inequality constraints has the following steps:

1. Draw k trial proposals for $\boldsymbol{s}_{uu,l}$ from the jumping distribution described by Eq. 2.
2. Compute a conditional constrained realization ($\boldsymbol{s}_{cc,l}$) for each of the trial proposals by minimizing the posterior negative log-likelihood via Lagrange multipliers:

$$\begin{aligned} L_{\boldsymbol{s},\boldsymbol{\beta}} \ = \ & \tfrac{1}{2}(\boldsymbol{z}+\boldsymbol{v}-\mathbf{H}\boldsymbol{s}^*_{cc,l})^T\mathbf{R}^{-1}(\boldsymbol{z}+\boldsymbol{v}-\mathbf{H}\boldsymbol{s}^*_{cc,l}) + \\ & \tfrac{1}{2}(\boldsymbol{s}^*_{cc,l}-\boldsymbol{s}^*_{uu,l})^T\mathbf{G}(\boldsymbol{s}^*_{cc,l}-\boldsymbol{s}^*_{uu,l}) \end{aligned} \tag{3}$$

where $\boldsymbol{v}$ is a random vector with covariance $\mathbf{R}$. In this case, the asterisk ($*$) indicates that the candidate is one of k trial proposals for the realization.

3. Compute the weighting function for each trial proposal:

$$w(\boldsymbol{s}^*_{cc,l}|\boldsymbol{s}_{cc,l-1}) = \frac{p''(\boldsymbol{s}^*_{cc,l}|\boldsymbol{z},\mathbf{H},\mathbf{X})}{T(\boldsymbol{s}^*_{cc,l}|\boldsymbol{s}_{cc,l-1})} \tag{4}$$

where $p''(\boldsymbol{s}^*_{cc,l}|\boldsymbol{z},\mathbf{H},\mathbf{X})$ indicates the posterior probability of $\boldsymbol{s}^*_{cc,l}$, and $T(\boldsymbol{s}^*_{cc,l}|\boldsymbol{s}_{cc,l-1})$ is the jumping probability of $\boldsymbol{s}^*_{cc,l}$ given $\boldsymbol{s}_{cc,l-1}$. The posterior probability and approximate jumping probability can be calculated as follows (Michalak and Kitanidis, 2002):

$$\begin{aligned} p''(\boldsymbol{s}^*_{cc,l}|\boldsymbol{z},\mathbf{H},\mathbf{X}) \;&\propto\; \exp[-\tfrac{1}{2}(\boldsymbol{z}-\mathbf{H}\boldsymbol{s}^*_{cc,l})^T\mathbf{R}^{-1}(\boldsymbol{z}-\mathbf{H}\boldsymbol{s}^*_{cc,l}) \\ &\quad -\tfrac{1}{2}{\boldsymbol{s}^*_{cc,l}}^T\mathbf{G}\boldsymbol{s}^*_{cc,l}] \end{aligned} \tag{5}$$

$$\begin{aligned} T(\boldsymbol{s}^*_{cc,l}|\boldsymbol{s}_{cc,l-1}) \;&\sim\; \exp[-\tfrac{1}{2}(\boldsymbol{s}^*_{uu,l}-0.9\boldsymbol{s}_{uu,l-1})^T \\ &\quad \frac{\mathbf{Q}^{-1}}{0.5^2}(\boldsymbol{s}^*_{uu,l}-0.9\boldsymbol{s}_{uu,l-1})] \end{aligned} \tag{6}$$

4. Select $\boldsymbol{s}_{cc,l}$ from the trial proposals by individually, randomly drawing each element from $\boldsymbol{s}^*_{cc,l}$ with probability proportional to the weighting function $w(\boldsymbol{s}^*_{cc,l}|\boldsymbol{s}_{cc,l-1})$. Select the corresponding elements of $\boldsymbol{s}^*_{uu,l}$ to construct $\boldsymbol{s}_{uu,l}$.

5. Create $(k-1)$ new trial proposals for $\boldsymbol{s}_{cc,l-1}$. To do this, draw samples from the jumping distribution $T(\boldsymbol{s}^*_{uu,l-1}|\boldsymbol{s}_{uu,l})$ (i.e., $\boldsymbol{s}^*_{uu,l-1} = 0.9\boldsymbol{s}_{uu,l} + 0.5\mathbf{C}\boldsymbol{u}$). Calculate the trial conditional constrained realizations $\boldsymbol{s}^*_{cc,l-1}$ using the procedure outlined in step 2. Set trail proposal k to $\boldsymbol{s}_{cc,l-1}$. Finally, calculate the weighting function for each trial conditional constrained realization, $w(\boldsymbol{s}^*_{cc,l-1}|\boldsymbol{s}_{cc,l})$.

6. Calculate the acceptance/rejection probability (Liu et al., 2000):

$$\xi = \min\left\{1, \frac{\sum^k w(\boldsymbol{s}^*_{cc,l}|\boldsymbol{s}_{cc,l-1})}{\sum^k w(\boldsymbol{s}^*_{cc,l-1}|\boldsymbol{s}_{cc,l})}\right\} \tag{7}$$

Accept $\boldsymbol{s}_{cc,l}$ if $\xi > \mathcal{U}(0,1)$. Otherwise, set $\boldsymbol{s}_{cc,l} = \boldsymbol{s}_{cc,l-1}$.

Repeat steps 1 – 6 until a sufficient number of realizations have been generated to sample across the entire posterior probability space. Note that unlike the Gibbs sampler, this multiple-try Metropolis-Hastings algorithm does not require discarding realizations from an initial spin-up period. For this application, we choose $k = 8$. Larger values for k can lead to greater acceptance rates but higher computational cost. Liu et al. (2000) note that an acceptance rate of $0.4 - 0.5$ is ideal for a multiple-try Metropolis-Hastings algorithm.

2 The Gibbs sampler implementation

The Gibbs sampler requires generating the conditional probability density, the probability of any individual element in $\boldsymbol{s}$ given an estimate of all other elements in $\boldsymbol{s}$. This conditional density is denoted $p(s_i|\boldsymbol{s},\boldsymbol{z})$ where i is one of m elements in $\boldsymbol{s}$. The equations for $p(s_i|\boldsymbol{s},\boldsymbol{z})$ can be found in Michalak (2008) for the inversion setup discussed in this paper.

The Gibbs sampler has the following steps:

1. Make an initial guess for $\boldsymbol{s}_1$ where the subscript '1' denotes the first realization of $\boldsymbol{s}$.

2. Obtain a new realization, $\boldsymbol{s}_l$, from the previous realization, $\boldsymbol{s}_{l-1}$. To do this, successively generate a conditional probability for each element in $\boldsymbol{s}$, and draw a random sample from each one:

$$\begin{aligned} p(s_{1,l}) &= p(s_{1,l}|s_{2,l-1}, ..., s_{m,l-1}) \\ p(s_{2,l}) &= p(s_{2,l}|s_{1,l}, s_{3,l-1}, ..., s_{m,l-1}) \\ p(s_{i,l}) &= p(s_{i,l}|s_{1,l}, ..., s_{i-1,l}, s_{i+1,l-1}, ..., s_{m,l-1}) \\ p(s_{m,l}) &= p(s_{m,l}|s_{1,l}, ..., s_{m-1,l}) \end{aligned} \tag{8}$$

3. Update l to $l+1$ and continue generating realizations.

Create a large number of realizations (in this case 1200) to fully sample across the posterior probability space. The initial realizations are usually discarded as a "spin-up" period.

In this implementation $p(s_i|\boldsymbol{s}, \boldsymbol{z})$ is Gaussian. To enforce the inequality constraints, Michalak (2008) draws a random sample from $p(s_i|\boldsymbol{s}, \boldsymbol{z})$ until the random draw falls within the bounds. This draw becomes the estimate for $s_{i,l}$. The approach is equivalent to sampling from a truncated normal distribution.

This study uses a modified approach for the methane case study. If the random sample from $p(s_i|\boldsymbol{s}, \boldsymbol{z})$ is positive, it becomes the estimate for $s_{i,l}$. If the random sample is negative, set $s_{i,l} = 0$. This approach is equivalent to sampling from a truncated normal distribution with an added Dirac delta function. The method adapted here increases the probability of estimating zero emissions for a given flux or emissions location.

References

Chib, S. and Greenberg, E.: Understanding the Metropolis-Hastings Algorithm, Am. Stat., 49, pp. 327–335, 1995.

Liu, J. S., Liang, F., and Wong, W. H.: The Multiple-Try Method and Local Optimization in Metropolis Sampling, J. Am. Stat. Assoc., 95, 121–134, doi:10.1080/01621459.2000.10473908, 2000.

Michalak, A. and Kitanidis, P.: Application of Geostatistical Inverse Modeling to Contaminant Source Identification at Dover AFB, Delaware, in: International Groundwater Symposium: Bridging the Gap between Measurement and Modeling in Heterogeneous Media, edited by Findikakis, A., pp. 137–139, 2002.

Michalak, A. M.: A Gibbs sampler for inequality-constrained geostatistical interpolation and inverse modeling, Water Resour. Res., 44, doi:10.1029/2007WR006645, 2008.