EAGE

26503

Convolved hidden Markov models for well-log inversion.

T. Fjeldstad* (NTNU) & H Omre (NTNU)

SUMMARY

Bayesian inversion of convolved data from discrete profiles, for example well-log data from lithofacies well profiles, is studied and found feasible to make. The projection approximation of the likelihood model provides reliable approximate posterior models, which can be used as proposal in an independent-proposal MCMC M-H algorithm, to generate realizations of lithofacies profiles from the correct posterior model.

Introduction

Bayesian inversion is commonly used to solve inverse problems in geophysics. Consider a profile through a reservoir unit discretized on the lattice $\mathscr{L}_{\mathscr{D}} = \{1, \ldots, n\}$. Given well-log observations $\boldsymbol{d} = \{d_t; t \in \mathscr{L}_{\mathscr{D}}\}$ our goal is to solve an inverse problem to predict the latent categorical variable $\boldsymbol{\kappa} = \{\kappa_t; t \in \mathscr{L}_{\mathscr{D}}\}$, with $\kappa_t \in \{1, \ldots, K\}$, describing the litholfacies-classes. We introduce a continuous latent variable $\boldsymbol{r} = \{r_t; t \in \mathscr{L}_{\mathscr{D}}\}$ describing the log-physics response. The response is often multimodal due to underlying lithofacies and the observations are usually convolved due to lack of resolution in the log-tool. We assume a Gaussian convolution kernel, as specified in Lindberg et al. (2014). We have a likelihood model such that the response is multimodal marginally. The well-log data $\boldsymbol{d}$ and the lithofacie profile $\boldsymbol{\kappa}$ typically look like the display next to top in Fig. 1. We extend the model studied by Rimstad and Omre (2013) to assess $\boldsymbol{\kappa}$ with spatial correlation. We consider two different k-th order approximations of the likelihood and show that the approximate posterior can be assessed by the Generalized Forward-Backward (GFB) algorithm, see Reeves and Pettitt (2004). As in Rimstad and Omre (2013) we use the approximate posterior as a proposal in an independent-proposal MCMC Metropolis-Hastings (M-H) algorithm and use the acceptance rate, α, as a measure of similarity between the approximate posterior and the true posterior. As an immediate consequence we are able to simulate from the true posterior model. A synthetic example comparing the two different likelihood approximation is included, a more thorough evaluation is found in Fjeldstad (2015).

Model

Assume that the prior model for $\boldsymbol{\kappa}$ is a first order stationary Markov chain,

$$p(\boldsymbol{\kappa}) = \prod_{t=1}^{n} p(\kappa_t | \kappa_{t-1}),$$

where $p(\kappa_1|\kappa_0) = p_s(\kappa)$ being the stationary distribution. We assume a Gauss-linear response likelihood

$$p(\boldsymbol{r}|\boldsymbol{\kappa}) = \phi_n\left(\boldsymbol{r}; \boldsymbol{\mu}_{r|\kappa}, \boldsymbol{\Sigma}_{r|\kappa}\right),$$

where $\boldsymbol{\mu}_{r|\kappa}$ is a $(n \times 1)$-vector of the means for r_t given κ_t for $t \in \mathscr{L}_{\mathscr{D}}$ and $\boldsymbol{\Sigma}_{r|\kappa}$ is a covariance $(n \times n)$-matrix. The covariance matrix is decomposed as $\boldsymbol{\Sigma}_{r|\kappa} = \boldsymbol{\Sigma}^{\sigma}_{r|\kappa} \boldsymbol{\Sigma}^{\rho}_{r} \boldsymbol{\Sigma}^{\sigma}_{r|\kappa}$, where $\boldsymbol{\Sigma}^{\sigma}_{r|\kappa}$ is a diagonal standard deviation $(n \times n)$-maxtrix and $\boldsymbol{\Sigma}^{\rho}_{r}$ is a spatial correlation $(n \times n)$-matrix dependent on $\rho(h) = \mathrm{Corr}(r_t, r_{t+h})$. We see that $p(\boldsymbol{r})$ is a Gaussian mixture by marginalizing with respect to $\boldsymbol{\kappa}$. We assume that the observations given the response, $\boldsymbol{d}|\boldsymbol{r}$, is represented by a Gauss-linear acquisition likelihood

$$p(\boldsymbol{d}|\boldsymbol{r}) = \phi_n\left(\boldsymbol{d}; \mathbf{W}\boldsymbol{r}, \sigma_d^2 \mathbf{I}\right),$$

where $\mathbf{W}$ is a convolution wavelet $(n \times n)$-matrix. Our approach is general and other linear operators such as selection and/or differential operators are also possible. The gross likelihood is

$$p(\boldsymbol{d}|\boldsymbol{\kappa}) = \phi_n\left(\boldsymbol{d}; \mathbf{W}\boldsymbol{\mu}_{r|\kappa}, \boldsymbol{\Sigma}^{\sigma}_{r|\kappa} \mathbf{W} \boldsymbol{\Sigma}^{\rho}_{r} \mathbf{W}^T \boldsymbol{\Sigma}^{\sigma}_{r|\kappa} + \sigma_d^2 \mathbf{I}\right),$$

which appear with two types of spatial effects, convolution $\mathbf{W}$ of lithofacies steps in $\boldsymbol{\mu}_{r|\kappa}$ and spatial correlations by $\mathbf{W}\boldsymbol{\Sigma}^{\rho}_{r}\mathbf{W}^T$ in the residual variable. In a Bayesian inverstion setting the goal is to assess the posterior of $[\boldsymbol{\kappa}|\boldsymbol{d}]$, which from Bayes' theorem is:

$$p(\boldsymbol{\kappa}|\boldsymbol{d}) = \mathrm{const} \times p(\boldsymbol{d}|\boldsymbol{\kappa}) p(\boldsymbol{\kappa})$$

The normalizing constant is infeasible as it requires K^n operations. We propose to approximate $p(\boldsymbol{d}|\boldsymbol{\kappa})$ on a k-th order factorial form. Let $\boldsymbol{\kappa}^{(k)}_t = (\kappa_{t-k+1}, \ldots, \kappa_t)$, $t = k, \ldots, n$, be a rephrasing of $\boldsymbol{\kappa}$. Hence a k-th order approximation of the true posterior is

$$p^{(k)}(\boldsymbol{\kappa}|\boldsymbol{d}) = \mathrm{const} \times \prod_{t=k}^{n} p^{(k)}\left(\boldsymbol{d}|\boldsymbol{\kappa}^{(k)}_t\right) p\left(\boldsymbol{\kappa}^{(k)}_t | \boldsymbol{\kappa}^{(k)}_{t-1}\right).$$

The approximate posterior model can be exactly assessed using the GFB-algorithm, with a computational cost of $(n-k-1) \times K^{k+1}$ operations. We propose two different approximations for the likelihood function $p^{(k)}\left(\boldsymbol{d}|\boldsymbol{\kappa}^{(k)}_t\right)$. The approximate posterior model, $p^{(k)}(\boldsymbol{\kappa}|\boldsymbol{d})$, may be used as a proposal in an independent-proposal MCMC M-H algorithm to generate realizations from the correct $p(\boldsymbol{\kappa}|\boldsymbol{d})$.

Likelihood Approximations

We present two different approximations for $p^{(k)}\left(\boldsymbol{d}|\boldsymbol{\kappa}_t^{(k)}\right)$, named the truncation and the projection approximation.

Truncation - The convolution $\mathbf{W}$ and spatial correlation $\boldsymbol{\Sigma}_r^{\rho}$ $(n \times n)$-matrices are both band-matrices, possibly broad-band. Define the k-th order band matrices $[\mathbf{W}]_k$ and $\left[\left(\mathbf{W}\boldsymbol{\Sigma}_r^{\rho}\mathbf{W}^T\right)^{1/2}\right]_k$ as the truncations of the respective matrices. The corresponding truncation k-th order approximate likelihood model can then be expressed as

$$p^{(k)}(\boldsymbol{d}|\boldsymbol{\kappa}) = \prod_{t=k}^{n} p^{(k)}\left(\boldsymbol{d}|\boldsymbol{\kappa}_t^{(k)}\right).$$

Projection - Let $p^{\star}(\boldsymbol{r})$ be a Gaussian approximation of $p(\boldsymbol{r}) = \sum_{\boldsymbol{\kappa}} p(\boldsymbol{r}|\boldsymbol{\kappa})\, p(\boldsymbol{\kappa})$, which actually is multimodal. Then $p^{\star}(\boldsymbol{d},\boldsymbol{r}) = p(\boldsymbol{d}|\boldsymbol{r})\, p^{\star}(\boldsymbol{r})$ is also Gaussian since $p(\boldsymbol{d}|\boldsymbol{r})$ is Gaussian, thus after marginalization and conditioning also $p^{\star}\left(\boldsymbol{d}|\boldsymbol{r}_t^{(k)}\right)$ is Gaussian. Since $p\left(\boldsymbol{r}_t^{(k)}|\boldsymbol{\kappa}_t^{(k)}\right)$ is Gaussian, we have that $p^{\star}\left(\boldsymbol{d}|\boldsymbol{\kappa}_t^{(k)}\right)$ is Gaussian. We propose the following likelihood approximation, termed the projection approximation

$$p^{(k)}\left(\boldsymbol{d}|\boldsymbol{\kappa}_t^{(k)}\right) = \begin{cases} \left[p^{\star}\left(\boldsymbol{d}|\boldsymbol{\kappa}_t^{(k)}\right)\right]^{1/k} \times \prod_{t=1}^{k-1} \left[p^{\star}\left(\boldsymbol{d}|\boldsymbol{\kappa}_{k-i}^{(k-i)}\right)\right]^{1/k} & \text{if } t = k \\ \left[p^{\star}\left(\boldsymbol{d}|\boldsymbol{\kappa}_t^{(k)}\right)\right]^{1/k} & \text{if } t = k+1,\ldots,n-1 \\ \left[p^{\star}\left(\boldsymbol{d}|\boldsymbol{\kappa}_t^{(k)}\right)\right]^{1/k} \times \prod_{t=1}^{k-1} \left[p^{\star}\left(\boldsymbol{d}|\boldsymbol{\kappa}_{k-i}^{(k-i)}\right)\right]^{1/k} & \text{if } t = k \end{cases}$$

where the k-th roots make sure that the observations are used only once and the second factors are boundary correction terms.

Examples

The examples are presented in Fig. 1, by a column of four displays. The top display presents the model parameters, with three different combinations of spatial correlation $\rho(\boldsymbol{r})$ and convolution wavelet $\boldsymbol{w}$. These combinations, numbered $1,2,3$, are constructed such that the band-matrices $\mathbf{W}\boldsymbol{\Sigma}_r^{\rho}\mathbf{W}^T$ have equal band-width. The results for parameter combination 1,2 and 3 are found in the three lower displays just below the parameter values. In the second display, the reference lithofacies profile, $\boldsymbol{\kappa}$, and the three reference responses, $\boldsymbol{r}$, and well-log observations, $\boldsymbol{d}$ are displayed. The proportion of lithofacies are $1/3$ for each and signal-to-noise-ratio, $\mathrm{S/N} \approx 1$. The observations appear with decreasing 'shoulder-effects' from left to right due to decreasing observation wavelet width and increasing response spatial correlation length. In the third display the truncation based MAP predictions of $\boldsymbol{\kappa}$ for varying order of k are given. The large features of the reference profile $\boldsymbol{\kappa}$ are reproduced, but the small-scale variability is lost, as can be expected in predictions. The α similarity measures to the correct posterior model $p(\boldsymbol{\kappa}|\boldsymbol{d})$ are also given, where α tend to increase with k. In the bottom display, the results for the projection based approximation are presented in the same format as for the truncation based one. The MAP predictions for $\boldsymbol{\kappa}$ are comparable, while the α-values are significantly higher for the projection approximation.

In Fig. 2, the α similarity measures to the correct $p(\boldsymbol{\kappa}|\boldsymbol{d})$ are summarized. For all parameter combinations the projection approximations are favorable to the truncation ones. Inversion of well-log data with relatively short-scale response variability in each lithofacies and wide observation convolution in the logging-tool is more reliable, than inversion with the opposite characteristics.

In Fig. 3, model parameter set 1 is used, and sequences of realizations from the independent-proposal MCMC M-H algorithm is displayed for the two approximations for varying order k. All displays contain realizations of the lithofacies from the correct posterior model, but we observe that the mixing of realizations improve with increasing k. The projection approximation provides better proposals than the truncation and higher order k tends to be favorable. Note, however, that the computational demands increase fast with order k. The realizations from the projection approximation with $k = 7$ do reproduce the hetergonity in the reference lithofacies profile.

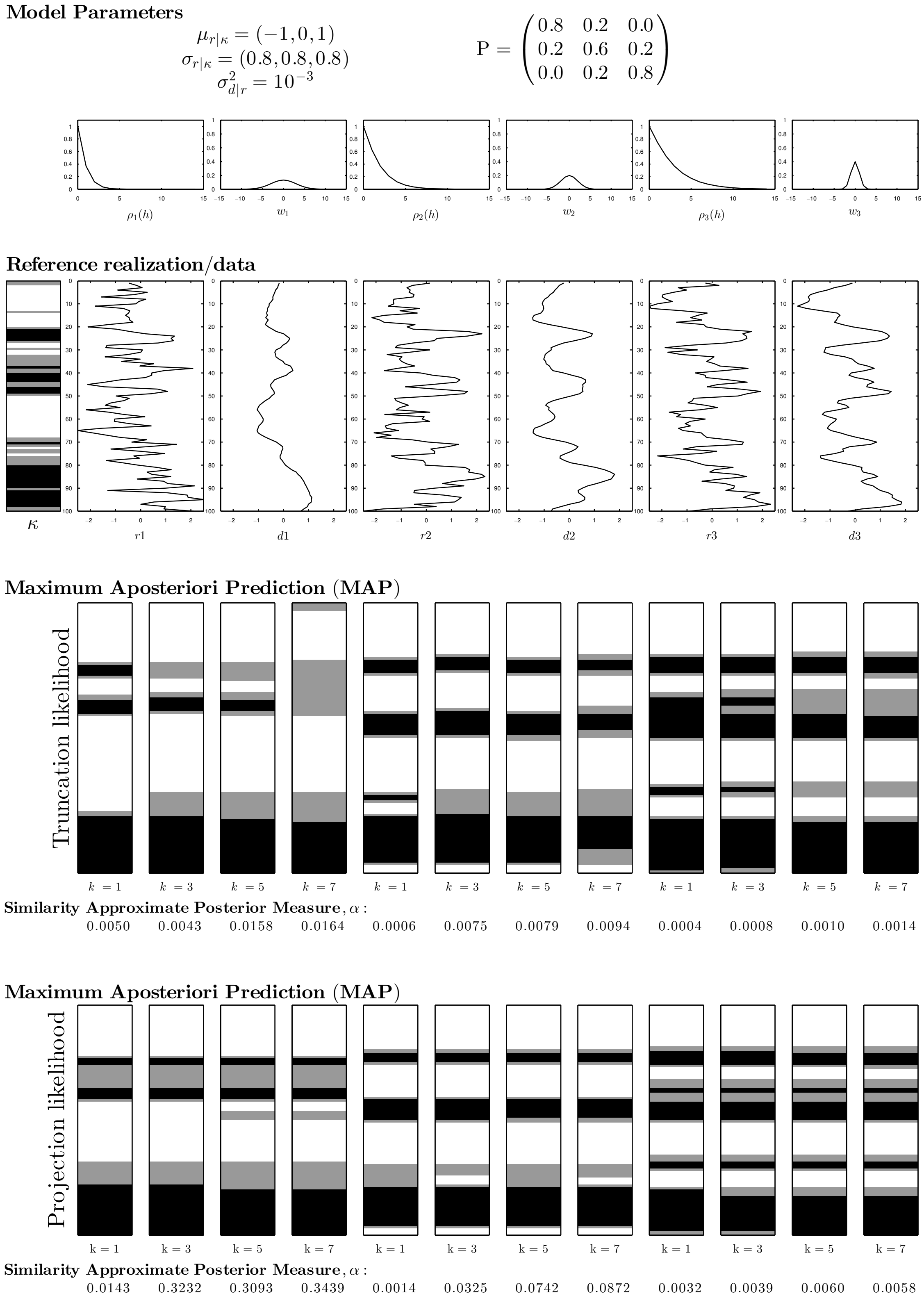

Figure 1: Model parameters, reference cases and MAP predictions/α-values for truncation and projection approximation for varying order k. Results based on 50,000 realizations.

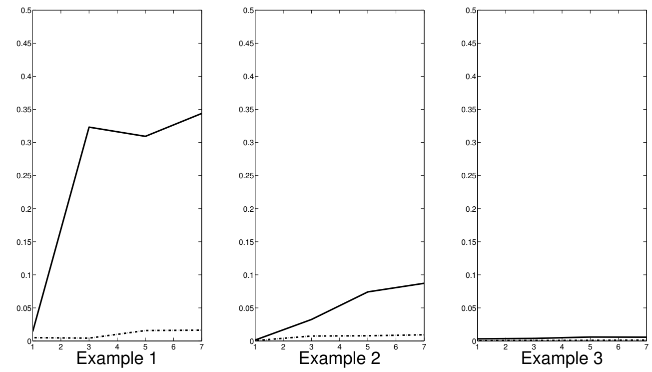

Figure 2: Similarity measure α for approximate vs correct posterior model, for truncation (hatch) and projection (solid) approximation for varying order k.

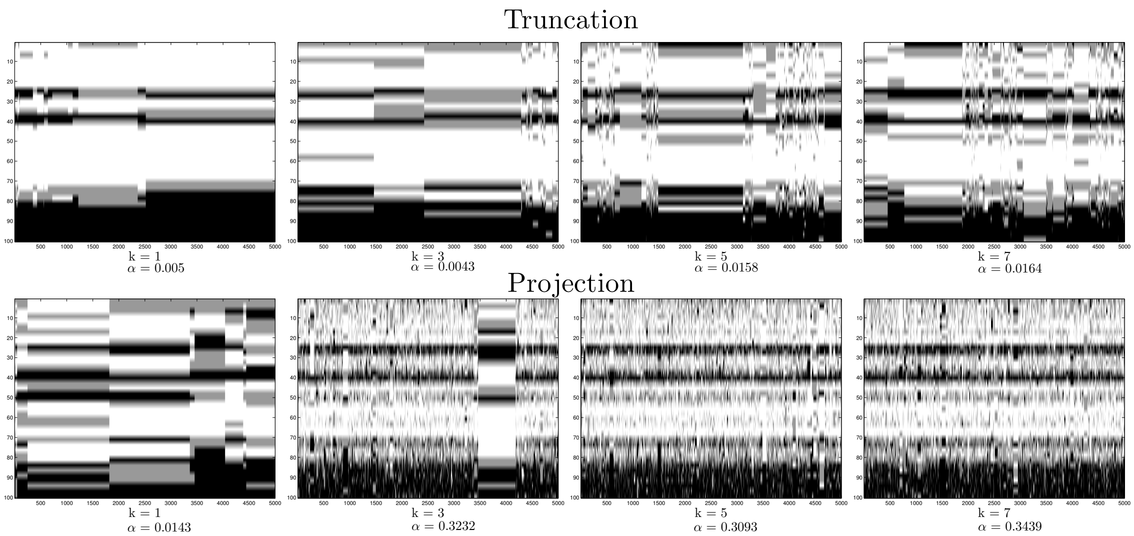

Figure 3: Realizations from the correct posterior model from MCMC-algorithm - 5,000 last realizations out of 50,000 - for model parameter set 1.

Conclusion

Bayesian inversion of convolved data from discrete profiles, for example well-log data from lithofacies well profiles, is feasible. The projection approximation of the likelihood model provides realiable approximate posterior models, which can be used as proposal in an independent-proposal MCMC M-H algorithm, to generate realizations from the correct posterior model.

References

Fjeldstad, T. [2015] *Convolved Hidden Markov Models*. Master's thesis, Norwegian University of Science and Technology.

Lindberg, D.V., Rimstad, E. and Omre, H. [2014] Inversion of well logs into facies accounting for spatial dependencies and convolution effects., paper submitted for publication.

Reeves, R. and Pettitt, A. [2004] Efficient recursions for general factorisable models. *Biometrika*, **91**(3), 751–757.

Rimstad, K. and Omre, H. [2013] Approximate posterior distributions for convolutional two-level hidden markov models. *Computational Statistics & Data Analysis*, **58**(C), 187–200.