PETROLEUM 2015
GEOSTATISTICS

Th B04

An Approximate Bayesian Inversion Framework based on Local-gaussian Likelihoods

M. Jullum* (University of Oslo) & O. Kolbjørnsen (Lundin Norway)

SUMMARY

We derive a Bayesian statistical procedure for inversion of geophysical data to rock properties. The procedure is for simplicity presented in the seismic AVO setting where rock properties influence the data through elastic parameters. The framework may however easily be extended. The procedure combines sampling based techniques and a compound Gaussian approximation to assess local approximations to marginal posterior distributions of rock properties, which the inversion is based on. The framework offers a range of approximations where inversion speed and accuracy may be balanced. The approach is also well suited for parallelisation, making it attractive for large inversion problems. We apply the procedure to a 4D CO2 monitoring case with focus on predicting saturation content. Promising results are obtained for both synthetic and real data. Finally we compare our method with regular linear Gaussian inversion for density prediction, where our method gives an improved fit.

Introduction

Various types of inverse problems abound in the geophysical industry. These problems are typically hard to solve, being ill-posed without a unique mathematical solution. Within the petroleum industry, the final objective of these problems is often to predict rock properties such as porosity, permeability, lithology, saturation etc. based on geophysical data stemming from geophysical exploration methods such as seismology surveys. Bayesian statistical inversion approaches are attractive as they give the ability to incorporate additional knowledge of the problem through prior distributions, and solve the inversion problems in a unified way by simply consulting the resulting posterior distribution (Grana and Mukerji, 2015). Computing the posterior distribution is however terribly difficult in almost all relevant realistic cases. Short on computational advance, attempts to compute posterior distributions or related quantities typically involve some kind of simplification, approximation or intelligent elimination of redundancy. We present a general framework for Bayesian statistical inversion from geophysical data to rock properties by combining sampling based techniques with local compound Gaussian approximations and utilize the simplification this local focus allows for.

Methodology

Although the present methodology is not restricted to inversion of specific types of geophysical data, we shall for presentational simplicity be working with seismic AVO (Amplitude Versus Offset) data. Let us introduce the following global quantities: r denotes one or more rock properties as exemplified above; m denotes geophysical properties or elastic parameters like e.g. V_p, V_s, and ρ; and d denotes the geophysical seismic AVO data at a few different offsets. Let further $p(d|m)$ and $p(m|r)$ denote respectively the pre-defined global geophysical and rock physical likelihoods, and let $p(r)$ denote the prior for the rock property of interest. Denote by $\boldsymbol{R}$ a specified grid on the reservoir region of interest. Our approach consists of approximating the marginal posterior distribution of the rock properties $p(r_A|d)$ in each cell A in $\boldsymbol{R}$. Let us therefore first discuss inversion for *one* such cell A. Let then $B = B(A)$, $C = C(A)$, and $D = D(A)$ denote neighbourhood sets around A related to respectively r, m and d. The corresponding variable subsets r_B, m_C and d_D act as local *influence variables*. These are the subsets of their corresponding global variables which influence the target variable r_A (being the rock property in cell A) and each other the most. Hence, we will approximate models and perform inversion for r_A based on these influence variables solely.

Our approach relies on the following relation

$$p(r_A|d_D) = \int p(r_B|d_D)\, \mathrm{d}r_{B\backslash A} \propto \int p(d_D|r_B)p(r_B)\, \mathrm{d}r_{B\backslash A}, \qquad (1)$$

where $B\backslash A$ denotes the set that excludes A from B. We will approximate the local likelihood $p(d_D|r_B)$ by a (compound) Gaussian approximation $p^*(d_D|r_B)$ and compute quantities related to $p(r_A|d_D)$ by a weighted Monte Carlo sampling approach using samples from $p(r_B)$.

The Gaussian approximation $p^*(d_D|r_B)$ will be obtained by compounding an approximate local geophysical likelihood $p^*(d_D|m_C)$ and an approximate local rock physical likelihood $p^*(m_C|r_B)$, where the former will be required to have a mean which is linear in m_C and covariance structure independent of m_C.

In the seismic AVO setting, the global geophysical likelihood model may often be written as

$$d = Gm + \epsilon, \qquad (2)$$

where G is a matrix of appropriate dimension representing the linear dependence structure of the seismic data on the elastic properties, and ϵ is a zero-mean error term. Typically ϵ is assumed to be Gaussian with covariance matrix Σ_ϵ. This corresponds to $p(d|m) \sim N(Gm, \Sigma_\epsilon)$. The straight forward natural approach for approximating $p(d_D|r_B)$ is then to use

$$p^*(d_D|m_C) \sim N\big(G_{DC}m_C, \Sigma_{\epsilon,DD}\big), \qquad (3)$$

where G_{DC} is the submatrix of G containing the rows corresponding to D and columns corresponding to C, and similarly $\Sigma_{\epsilon,DD}$ is the submatrix of Σ_ϵ with columns and rows corresponding to D.

Approximating the local rock physics likelihood $p(m_C|r_B)$ is more involved as the global analogue is typically neither Gaussian nor linearly dependent on r. Thus, we suggest using a sampling based approach to estimate mean and covariance matrix functions of a Gaussian distribution on the form

$$p^*(m_C|r_B) \sim N\left(\mu_{m_C}(r_B), \Sigma_{\mathrm{m_C}}\left(k(r_B)\right)\right), \tag{4}$$

where $\mu_{m_C}(r_B)$ is some mean function depending on r_B in some way, and $\Sigma_{m_C}\left(k(r_B)\right)$ is some covariance function depending on a categorical variable $k(r_B) \in \{1, \dots, K\}$ for a fixed integer K. The mean and covariance matrix functions may be estimated from a large number M of samples $(m_C^{(i)}, r_B^{(i)})$, $i = 1, \dots, M$. Our suggested approach is to first split the set of samples into Ksubsets categorized by $k(r_B)$. Then for each of these categories, use a flexible (nonlinear) regression procedure to approximate the mean function $\mu_{k,m_C}(r_B) = \mu_{m_C}(k(r_B), r_B)$, compute the resulting residuals $\delta_i = m_C^{(i)} - \mu_{m,k}(r_B^{(i)})$ and use those to estimate the covariance matrix $\Sigma_{\mathrm{m_C}}(k(r_B))$. Finally we set $\mu_{m_C}(r_B)$ equal to $\mu_{k,m_C}(r_B)$ when $k(r_B) = k$. We tend to use multivariate adaptive regression splines (MARS) for the regression, and a range spanning covariance estimation routine based on the regular sample covariance to estimate the K covariance matrices, but other choices are of course possible. The main point here is to be flexible and use enough samples for the mean behaviour and dependence structure of $p(m_C|r_B)$ to be well approximated by those used in $p^*(m_C|r_B)$.

By compounding the two approximate Gaussian likelihoods $p^*(d_D|m_C)$ and $p^*(m_C|r_B)$, we obtain

$$p^*(d_D|r_B) \sim N\left(\mu_{d_D}(r_B), \Sigma_{d_D}(r_B)\right), \tag{5}$$

where $\mu_D(r_B) = G_{DC}\mu_{m_C}(r_B)$ and $\Sigma_{d_D}(r_B) = G_{DC}\Sigma_{m_C}\left(k(r_B)\right)G_{CD} + \Sigma_{\epsilon,DD}$ which may be inserted in equation (1) for an integral form approximation to the posterior of r_A. Finally a weighted Monte Carlo approach may be applied to approximate in principle any quantity related to this posterior distribution. That is, by sampling L times from $p(r_B)$, weighting $r_A^{(l)}$ in each sample $r_B^{(l)}$ by the normalized likelihood approximation $w^{(l)} = p^*(d_D|r_B^{(l)}) / \Sigma_j p^*(d_D|r_B^{(j)})$, for $l = 1, \dots L$ and finally aggregate these weighted samples properly in terms of the quantity of interest.

Methodology for inversion of a single cell A is now established. Inversion of the complete region $\boldsymbol{R}$ is then performed by simply repeating the procedure for each r_A with A in $\boldsymbol{R}$. Assume now that the influence variables shift with A (i.e. their location depends only on A) and that $p(d_{D(A)}, m_{C(A)}, r_{B(A)})$ is stationary with respect to A for all of (or at least larger parts of) $\boldsymbol{R}$. Then, when the compound Gaussian likelihood in equation (5) is built once, that approximation is valid for each A for which the stationarity assumption holds. Also, the same prior samples from $p(r_B)$ may be used in the final step. Those assumptions are not required for our methodology to work out, but give a computationally cheap algorithm as one may pre-evaluate parts of the likelihood $p^*(d_D|r_B)$ and heavily parallelise the Monte Carlo algorithm. The only requirement for our approach to be applicable is actually that $p(d|m)$ is (approximately) Gaussian and that one somehow may sample from $p(m_C, r_B)$. Specification of the neighbourhood sets B,C and D, and categorisation function $k(r_B)$, may typically be done using knowledge of the underlying features of the data and models, supported by performance achieved on synthetic test data. This should typically be done by balancing accuracy and inversion speed.

Application

To illustrate our approach, consider a 4D survey of the Utsira formation from the Sleipner field offshore Norway. Here CO_2 has been injected for storage, and base and monitor surveys has been conducted. The CO_2 is typically trapped beneath thin layers of shale within the formation. The main

objective is to 'map' the CO_2 through the saturation at different positions in the formation. The saturation was in principle zero everywhere at the time of the base survey. We focus on the saturation (r) at the time of one monitor survey, and use data and models working on the difference in seismic data, elastic parameters and saturation between those two surveys. Let d denote the change in aligned seismic angle gathers from base to monitor time, let m denote the change in the logarithm of P-wave and S-wave velocity and density from base to monitor time.

The model setup here possesses only vertical spatial dependencies. Adjacent cells are generally more correlated than distant ones and stationarity with respect to A is assumed. The model parts are briefly as follows: $p(r_A) = 0$ with probability 0.99 and otherwise Beta(6,1.5)-distributed. The global rock physics model $p(m|r)$ has a rock matrix consistent with loose sand, and the effect of the saturation is modelled by fluid substitution with a homogenous fluid mix (Gassmann's equations). The global geophysical model $p(d|m)$ is on the form of (2) with G corresponding to the difference effect of 25, 30 and 35Hz Ricker wavelets for the weak-contrast coefficients with offsets 12.5°, 25° and 40°. Σ_ϵ specifies standard deviations of 0.5, 0.5 and 0.2 for the 3 offsets with independence between them.

For the real data we concentrate on a 2D region positioned 800m below the sea level and intersecting the injection well and spanning 5550m × 334ms. Each cell is of size 50m × 2ms. Figure 1 show the seismic difference data for that region (left panel), the resulting approximate cell-wise marginal posterior means (upper right panel) and more detailed inversion results for a vertical trace lateral position 2000-2050m, close to the injection well (lower right panels). The resulting cell-wise posterior means, which we take as the pointwise predictions of the saturation, match well with the reflections in the seismic data. The vertical plot shows 5 main layers of increased saturation at that trace. The two shallowest and the deepest are fairly certain, while the other two are although predicted, quite uncertain. The predictions are consistent with qualitative interpretations previously conducted in the formation. A total of 2×10^5 samples from the prior $p(r_B)$ were used in the inversion.

In tests with synthetic data, the selected method achieved performance as follows: About 98.3% of the cells with a saturation level of less than 0.1 did 'correctly' have a posterior expectation less than 0.1. Conversely 97% of the cells which had saturation level above 0.1 did 'correctly' have a posterior expectation above 0.1. That is, the proportions of cells which are false positives and false negatives are about 1.7% and 3% when using the resulting approximate posterior mean (with boundary level at 0.1) as a classifier of significant saturation presence.

Finally we leave the goal of prediction saturation and turn to prediction of density (ρ) at monitor time instead. With this focus, our method may easily be compared to regular 'Gaussian inversion' (Buland and Omre, 2003) which in principle assumes everything is Gaussian and linear. For a synthetic 2D region similar to the above, Figure 2 shows the resulting posterior means of log-densities for our method (upper left panel) and the regular Gaussian inversion (lower left panel), in addition to the true synthetic log-densities (upper right panel). Our method clearly outperforms regular Gaussian inversion. In the latter the posterior means are less clear and distinct compared to those based on our method. Many low density areas are missed and many high density areas are falsely predicted. The general mean picture is also noisy.

Conclusion

We presented a general framework for inversion of geophysical data to rock properties based on approximating local Gaussian likelihoods followed by a weighted Monte Carlo sampling scheme. The approach is tested on a 4D CO_2-monitoring case with promising results both for synthetic and real data. The approach is well suited for parallelisation, making it attractive for large inversion problems. Specifically, inversion of the real data 2D region, having more than 18000 cells, took less than half an hour on a Windows based laptop without m speed boosting attempts. A final conceptual advantage of our general approach, as opposed to other alternatives, is that we may start from realistic prior distributions and models which we approximate directly – without having the full procedure being limited by a specific assumption for the spatial relations.

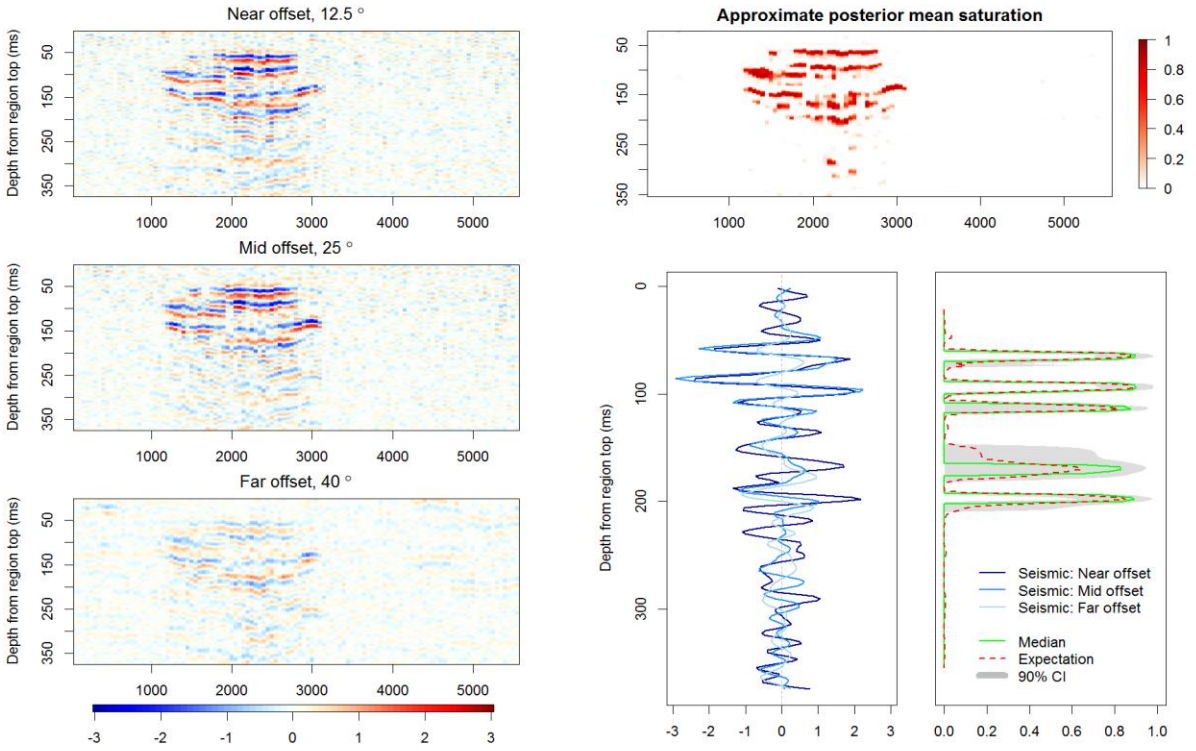

Figure 1 *Seismic difference data and resulting approximate inversion results for the real data case.*

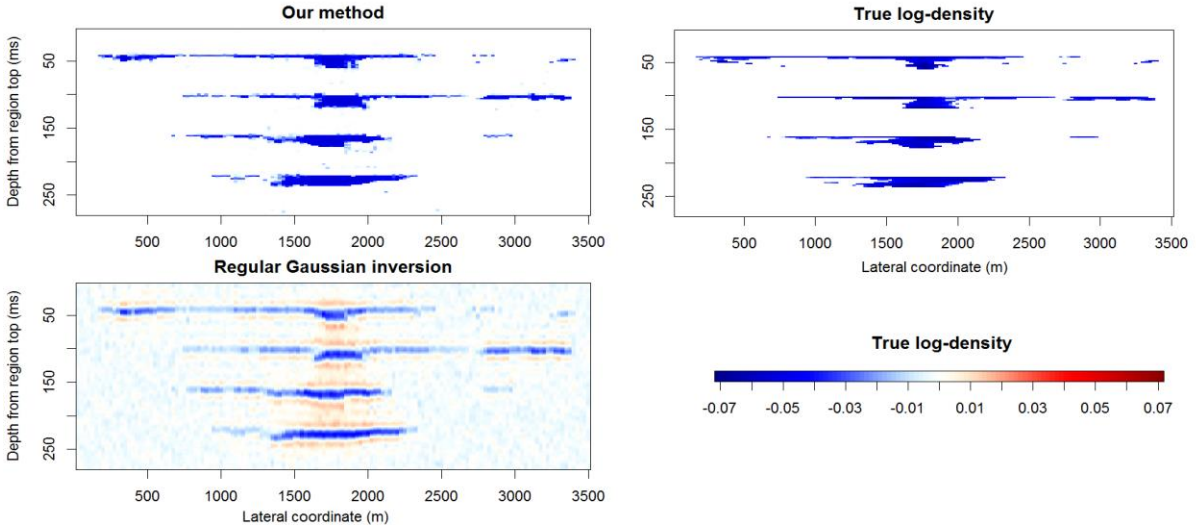

Figure 2 *Comparing our method and standard Gaussian inversion to density (ρ) for synthetic data.*

Acknowledgements

Thanks to the Sleipner license partners Statoil, ExxonMobil, and Total for permission to use the Sleipner data.

References

Buland, A., and H. Omre [2003] Bayesian linearized AVO inversion. *Geophysics*, **68**, 185-198.

Friedman, J.H. [1991] Multivariate adaptive regression splines. *The Annals of Statistics*, **19**, 1-67.

Grana, D., and T. Mukerji, [2015] Bayesian inversion of time-lapse seismic data for the estimation of static reservoir properties and dynamic property changes. *Geophysical Prospecting*, **63**, 637-655.