Norsk Regnesentral
NORWEGIAN COMPUTING CENTER

Note

Interpretation of saturation and porosity from inverted 4D seismic parameters

Note no SAND/03/12

Authors Odd Kolbjørnsen

Date January 2012

About the authors

Dr.Ing. Odd Kolbjørnsen is a chief research scientist at the Norwegian Computing Center.

Norsk Regnesentral

Norsk Regnesentral (Norwegian Computing Center, NR) is a private, independent, non-profit foundation established in 1952. NR carries out contract research and development projects in the areas of information and communication technology and applied statistical modeling. The clients are a broad range of industrial, commercial and public service organizations in the national as well as the international market. Our scientific and technical capabilities are further developed in co-operation with The Research Council of Norway and key customers. The results of our projects may take the form of reports, software, prototypes, and short courses. A proof of the confidence and appreciation our clients have for us is given by the fact that most of our new contracts are signed with previous customers.

Norsk Regnesentral
Norwegian Computing Center
Postboks 114, Blindern
NO-0314 Oslo, Norway

Besøksadresse
Office address
Gaustadalléen 23
NO-0371 Oslo, Norway

Telefon · telephone
(+47) 22 85 25 00
Telefaks · telefax
(+47) 22 69 76 60

Internett · internet
www.nr.no
E-post · e-mail
nr@nr.no

Title	**Interpretation of saturation and porosity from inverted 4D seismic parameters**
Author	**Odd Kolbjørnsen**
Date	January
Year	2012
Publication number	SAND/03/12

Abstract

In this note we propose two alternative approaches for stochastic rock physics inversion of 4D seismic data. Both methods use the approach of approximate local marginal likelihood coming from a Bayesian inversion scheme. We describe the concept of approximate local marginal likelihood, and put forward arguments for the approach, using arguments from the linear estimation methods, bounding of error, and also a more heuristic argument through information theory. Within the framework we propose two alternative approximations, which differ with respect to which level we use the approximation. The fastest and simplest approach applies the concept of local approximate marginal likelihood directly to porosity and saturation, the slightly more time consuming approach applies the concept of approximate local marginal likelihood to the seismic parameters and will give more accurate results for porosity and saturation.

Keywords	Inverse problems, 4D seismic, rock physics
Target group	Project-group
Availability	Open
Project number	808007
Research field	Geophysical data
Number of pages	11

1 Introduction

As a part of the NFR funded MonCO2 project is to integrate seismic-travel time data, seismic amplitude data, and gravimetric measurements with a stochastic rock physics model in order to obtain a best possible picture of the spatial distribution of the injected CO2. Through the cooperation with UiB models for rock physics relations have been established. By assigning proper probability distributions to the rock parameters we establish a joint prior distribution of rock parameters (e.g. porosity and saturation), and 4D seismic parameters (i.e. pressure wave velocity, shear wave velocity, density, and changes in these). In the NR-Note Kjønsberg and Kolbjørnsen (2011), we discuss how to approximate the distribution of the seismic parameters using a Gaussian distribution. In additional NR-notes we discuss the inversion methodology for separate data types, and how we can use 3D inversion technologies to perform 4D inversion, see Kolbjørnsen and Kjønsberg (2011).

In this note we discuss how to interpret the results of Gaussian-linear inversion of seismic data in order to provide an approximate marginal likelihood for rock physics parameters such as porosity and saturation.

2 Methodology

We assume we have defined a joint probability distribution for rock physics parameters $\mathbf{r}$ and seismic parameters $\mathbf{m}$. We assume that this relation is valid for the full grid, i.e. that we have a stationary marginal distribution; we do not have any other local knowledge of the parameters under study. The target is to use results from Bayesian inversion methodologies to assess local information about rock-physics parameters $\mathbf{r}$. Define a grid covering the region, and let the variable $\mathbf{m}$ denote all the seismic parameters in the grid, i.e. $\mathbf{m}=[\mathbf{m}_1,\mathbf{m}_2,...,\mathbf{m}_N]$, where $\mathbf{m}_i$ is the seismic parameter in a specific location; and N is the number of grid cells. The seismic parameter is observed through an indirect measurement with a linear or linearized relation between the seismic parameter and data, i.e.

$$\mathbf{d}=\mathbf{G}\mathbf{m}+\boldsymbol{\varepsilon},$$

where $\mathbf{d}$ is the data; $\mathbf{G}$ is the forward map which describes the relation between the seismic parameter and the observation; and $\boldsymbol{\varepsilon}$ is the observation error. Below this will be assumed to have a Gaussian distribution, unless otherwise mentioned. We consider first the situation where our interest is in a linear combination of the parameters, $\mathbf{m}_L=\mathbf{A}\mathbf{m}$. For simplicity of the presentation (and because this will be of primary interest) we fill out the details only for the case where our parameter is the variable itself in one grid cell, say cell i. We separate the variable vector into two parts; the component in cell i and the rest, i.e. $\mathbf{m}=[\mathbf{m}_i,\mathbf{m}_{-i}]$. The latter component is $(N-1)$ times the size of the former.

Since our interest is in the local variable, we want to evaluate the local likelihood, i.e. $p(\mathbf{d}\,|\,\mathbf{m}_i)$ or the local posterior distribution $p(\mathbf{m}_i\,|\,\mathbf{d})$ keeping in mind that the dimension of data might

be several million, whereas $\mathbf{m}_i$ has low dimension typically less than 10. To solve this problem we will develop the concept of local marginal likelihood.

2.1 Approximate marginal likelihood for large-scale inverse problems

The marginal likelihood is obtained by marginalizing all but the target variable:

$$p(\mathbf{d}\,|\,\mathbf{m}_i) = \int p(\mathbf{d}\,|\,\mathbf{m})\,p(\mathbf{m}_{-i}\,|\,\mathbf{m}_i)d\mathbf{m}_{-i} = \frac{\int p(\mathbf{d}\,|\,\mathbf{m})\,p(\mathbf{m})d\mathbf{m}_{-i}}{p(\mathbf{m}_i)}.$$

The integral over $\mathbf{m}_{-i}$ being over all seismic parameters except the target variables has millions of dimensions. Below we first present an intuitive approximation of the marginal likelihood through the use of Gaussian-linear inversion. Next we argue that this approximation has a wider applicability than the framework of Gaussian-linear inversion.

2.1.1 An intuitive presentation of the approximate marginal likelihood

The marginalization in the integral of section 2.1 cannot in general be of performed analytically, but if we assume that the elastic parameters follow a Gaussian distribution we are able to compute the posterior analytically. Thus by using the approximation

$$p(\mathbf{m}) \approx p^*(\mathbf{m}),$$

where the star indicates a Gaussian distribution, we are able to compute the posterior distribution by standard Gaussian-linear relations. Thus $p^*(\mathbf{m}\,|\,\mathbf{d}),$ is a known quantity, and so is also $p^*(\mathbf{m}_i\,|\,\mathbf{d})$. This gives us the possibility to compute the approximation of the local marginal likelihood.

$$\begin{aligned} p\left(\mathbf{d}\middle|\mathbf{m}_i\right) &= \frac{\int p(\mathbf{d}\,|\,\mathbf{m})\,p(\mathbf{m})d\mathbf{m}_{-i}}{p(\mathbf{m}_i)} \\ &\approx \frac{\int p(\mathbf{d}\,|\,\mathbf{m})\,p^*(\mathbf{m})d\mathbf{m}_{-i}}{p^*(\mathbf{m}_i)} \\ &\propto \frac{\int p^*(\mathbf{m}\,|\,\mathbf{d})d\mathbf{m}_{-i}}{p^*(\mathbf{m}_i)} \\ &\propto \frac{p^*(\mathbf{m}_i\,|\,\mathbf{d})}{p^*(\mathbf{m}_i)}. \end{aligned}$$

Thus the approximate marginal likelihood is proportional to the ratio of the Gaussian posterior distribution to the Gaussian prior distribution for the local parameter, the resulting likelihood is hence Gaussian as well. The approximate likelihood should be used in the local computations, discussed in section 2.3. Note that the computations above are general, but in contrast to the situation for most other prior models, the Gaussian assumptions make it possible to evaluate both the prior and the posterior distribution in the ratio above.

2.1.2 Argument through bounding assumptions

The Gaussian assumption is not universally applicable, but by analyzing the approximation we might find features of distributions that make the approach suited. For the true local posterior distribution we find:

$$
\begin{aligned}
p(\mathbf{m}_i \mid \mathbf{d}) &= \int p(\mathbf{m} \mid \mathbf{d}) d\mathbf{m}_{-i} \\
&= \int \frac{p(\mathbf{d} \mid \mathbf{m}) p(\mathbf{m})}{p(\mathbf{d})} \frac{p^*(\mathbf{m})}{p^*(\mathbf{m})} d\mathbf{m}_{-i} \\
&= \int \frac{p(\mathbf{m})}{p^*(\mathbf{m})} \frac{p(\mathbf{d} \mid \mathbf{m}) p^*(\mathbf{m})}{p(\mathbf{d})} \mathrm{d}\mathbf{m}_{-i} \\
&= \int \frac{p(\mathbf{m})}{p^*(\mathbf{m})} \frac{p^*(\mathbf{m} \mid \mathbf{d}) p^*(\mathbf{d})}{p(\mathbf{d})} \mathrm{d}\mathbf{m}_{-i} \\
&= \frac{p^*(\mathbf{d})}{p(\mathbf{d})} \int \frac{p(\mathbf{m})}{p^*(\mathbf{m})} p^*(\mathbf{m} \mid \mathbf{d}) \mathrm{d}\mathbf{m}_{-i} \\
&= \frac{p^*(\mathbf{d})}{p(\mathbf{d})} \int \frac{p(\mathbf{m}_i) p(\mathbf{m}_{-i} \mid \mathbf{m}_i)}{p^*(\mathbf{m}_i) p^*(\mathbf{m}_{-i} \mid \mathbf{m}_i)} p^*(\mathbf{m}_i \mid \mathbf{d}) p^*(\mathbf{m}_{-i} \mid \mathbf{m}_i, \mathbf{d}) \mathrm{d}\mathbf{m}_{-i} \\
&= \frac{p^*(\mathbf{d})}{p(\mathbf{d})} p(\mathbf{m}_i) \frac{p^*(\mathbf{m}_i \mid \mathbf{d})}{p^*(\mathbf{m}_i)} \int \frac{p(\mathbf{m}_{-i} \mid \mathbf{m}_i)}{p^*(\mathbf{m}_{-i} \mid \mathbf{m}_i)} p^*(\mathbf{m}_{-i} \mid \mathbf{m}_i, \mathbf{d}) \mathrm{d}\mathbf{m}_{-i}.
\end{aligned}
$$

In this expression, we see that the ratio of the prior distributions for the target variable and data factors out, thus what remains is the local likelihood:

$$
p(\mathbf{d} \mid \mathbf{m}_i) = p^*(\mathbf{d}) \frac{p^*(\mathbf{m}_i \mid \mathbf{d})}{p^*(\mathbf{m}_i)} \int \frac{p(\mathbf{m}_{-i} \mid \mathbf{m}_i)}{p^*(\mathbf{m}_{-i} \mid \mathbf{m}_i)} p^*(\mathbf{m}_{-i} \mid \mathbf{m}_i, \mathbf{d}) \mathrm{d}\mathbf{m}_{-i}.
$$

The local likelihood consists of three factors: a normalizing constant; the ratio of the prior to the posterior Gaussian-linear approximation; and an additional integral factor which is disregarded in the Gaussian-linear approximation. To assure that the error in the approximation is bounded, we must control the last integral factor:

$$
\int \frac{p(\mathbf{m}_{-i} \mid \mathbf{m}_i)}{p^*(\mathbf{m}_{-i} \mid \mathbf{m}_i)} p^*(\mathbf{m}_{-i} \mid \mathbf{m}_i, \mathbf{d}) \mathrm{d}\mathbf{m}_{-i}.
$$

Analyzing this factor we see that it consist of a ratio between the true and the Gaussian prior distribution given the target parameter integrated over the Gaussian approximation of the posterior for the non-target variables given the target. It is thus sufficient to put an upper bound of the ratio in order to limit the error in the local likelihood. This simple expression discloses some situations where the approximation has limited value. If the prior distribution has heavier tails (or rather skirts) than the Gaussian approximation, the integral can be unbounded. On the other side, if the distribution of the parameters have compact support (i.e. there exist upper and lower limits for the parameters), the approximation will have a bounded error, in the sense that there exist lower and upper limits c_L and c_U such that:

$$
c_L p^*(\mathbf{d} \mid \mathbf{m}_i) < p(\mathbf{d} \mid \mathbf{m}_i) < c_U p^*(\mathbf{d} \mid \mathbf{m}_i), \forall \mathbf{m}_i .
$$

Independence of the target variable and non-target variables in the prior distribution are in general not sufficient for the integral factor to collapse to a constant, since the non-target variables and target variables in general are dependent in the posterior distribution;

$$
p^*(\mathbf{m}_{-i} \mid \mathbf{m}_i) = p^*(\mathbf{m}_{-i}) \not\Rightarrow p^*(\mathbf{m}_{-i} \mid \mathbf{m}_i, \mathbf{d}) = p^*(\mathbf{m}_{-i} \mid \mathbf{d})
$$

2.1.3 Argument through linear estimation

We assume that the observations have the form:

$$\mathbf{d} = \mathbf{G}\mathbf{m} + \boldsymbol{\varepsilon},$$

and use linear predictions to assess the influence of the target variable on data. From linear prediction theory (i.e. kriging) we know that the full parameter can be written as:

$$\mathbf{m} = \boldsymbol{\mu}_m + \boldsymbol{\Sigma}_{m,i}\boldsymbol{\Sigma}_i^{-1}\left(\mathbf{m}_i - \boldsymbol{\mu}_i\right) + \boldsymbol{\varepsilon}_{m|i},$$

where $\boldsymbol{\mu}_m$ is the expected value of $\mathbf{m}$; $\boldsymbol{\Sigma}_{m,i}$ is the covariance between $\mathbf{m}$ and $\mathbf{m}_i$; $\boldsymbol{\Sigma}_i$ is the covariance of $\mathbf{m}_i$; $\boldsymbol{\mu}_i$ is the expected value of $\mathbf{m}_i$; and $\boldsymbol{\varepsilon}_{m|i}$ is a random variable which is uncorrelated with $\mathbf{m}_i$. Inserting this into the data expression we find that:

$$\mathbf{d} = \mathbf{G}\boldsymbol{\Sigma}_{m,i}\boldsymbol{\Sigma}_i^{-1}\mathbf{m}_i + \mathbf{G}\left(\boldsymbol{\mu}_m - \boldsymbol{\Sigma}_{m,i}\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\mu}_i\right) + \mathbf{G}\boldsymbol{\varepsilon}_{m|i} + \boldsymbol{\varepsilon},$$

The observed data has four terms. The first shows a linear dependence to the target parameter, the second is constant, the third and fourth are terms that are uncorrelated to the target parameter. Thus we may write

$$\mathbf{d}^* = \mathbf{G}_i^*\mathbf{m}_i + \boldsymbol{\varepsilon}^*,$$

where:

$$\mathbf{G}_i^* = \mathbf{G}\boldsymbol{\Sigma}_{m,i}\boldsymbol{\Sigma}_i^{-1},$$

$$\mathbf{d}^* = \mathbf{d} - \mathbf{G}\boldsymbol{\mu}_m + \mathbf{G}_i^*\boldsymbol{\mu}_i,$$

$$\mathrm{Cov}(\boldsymbol{\varepsilon}^*) = \boldsymbol{\Sigma}_{\varepsilon^*} = \mathbf{G}\left(\boldsymbol{\Sigma}_m - \boldsymbol{\Sigma}_{m,i}\boldsymbol{\Sigma}_i^{-1}\boldsymbol{\Sigma}_{i,m}\right)\mathbf{G}^T + \boldsymbol{\Sigma}_\varepsilon.$$

The expression is merely a rewrite of the original equation, but where the prior distribution of the non- target parameters are interpreted as a part of the likelihood model, and not as a part of the prior model. The system above is a standard regression setting where the error covariance is known. Following the regression approach the solution to the problem is given by the unbiased estimator:

$$\hat{\mathbf{m}}_i = \left(\mathbf{G}_i^{*T}\boldsymbol{\Sigma}_{\varepsilon^*}^{-1}\mathbf{G}_i^*\right)^{-1}\mathbf{G}_i^{*T}\boldsymbol{\Sigma}_{\varepsilon^*}^{-1}\mathbf{d}^*,$$

which has the variance:

$$\mathrm{Cov}\left\{\hat{\mathbf{m}}_i\right\} = \left(\mathbf{G}_i^{*T}\boldsymbol{\Sigma}_{\varepsilon^*}^{-1}\mathbf{G}_i^*\right)^{-1}$$

Thus defining a projection matrix:

$$\mathbf{P} = \left(\mathbf{G}_i^{*T}\boldsymbol{\Sigma}_{\varepsilon^*}^{-1}\mathbf{G}_i^*\right)^{-1}\mathbf{G}_i^{*T}\boldsymbol{\Sigma}_{\varepsilon^*}^{-1}$$

we may transform the problem to a system of equations having a dramatically reduced dimension:

$$\mathbf{d}^{**} = \mathbf{m}_i + \boldsymbol{\varepsilon}^{**},$$

where:

$$\mathbf{d}^{**} = \mathbf{P}\mathbf{d}^*,$$

$$\boldsymbol{\Sigma}_{\varepsilon^{**}} = \left(\mathbf{G}_i^{*T}\boldsymbol{\Sigma}_{\varepsilon^*}^{-1}\mathbf{G}_i^*\right)^{-1}$$

The matrix $\mathbf{G}_i^{*T}\boldsymbol{\Sigma}_{\varepsilon^*}^{-1}\mathbf{G}_i^*$ is central in the projection which reduces the data dimension. This is the precision matrix of the projected observations, as seen above. The matrix might however not be invertible, which in turn imply that the covariance matrix of the data does not exist. If the

matrix $\mathbf{G}_i^{*T}\boldsymbol{\Sigma}_{\varepsilon^*}^{-1}\mathbf{G}_i^*$ has reduced rank this means that there are linear features in $\mathbf{m}_i$ which are unidentified by the data. By factoring out these features, we might proceed as if the matrix has full rank, keeping in mind that some features of our parameter are unidentifiable. Section 2.2 discusses this situation.

The assumption of the approximate local marginal likelihood through the Gaussian approximation is equivalent with assuming a multi-normal distribution for the error

$$\boldsymbol{\varepsilon}^{**} = \mathbf{PG}\boldsymbol{\varepsilon}_{m|i} + \mathbf{P}\boldsymbol{\varepsilon}.$$

The error $\boldsymbol{\varepsilon}^{**}$ consists of a sum of multiple errors sources, thus there is an averaging effect which makes the multi normal distribution plausible. Limiting properties are however not available in the general case. The computations show that the marginal approximate likelihood approach which we propose preserves the linear information of the target-parameter in the data.

2.1.4 Argument through information theory

It is since long established that the nature does not comply with a Gaussian assumption, more over the parameters of the Gaussian- distribution is hard to assess. We still argue that the Gaussian distribution is a reasonable choice for the marginalization. First it is easy to claim that the nature is non-Gaussian, but it is hard to come up with an alternative. In spatial statistics and through geostatistical approaches there exist multiple alternatives to the Gaussian approach, but many alternative models can be discredited for the same reasons as the Gaussian, other models display specific characteristics that are desirable, but it is often hard to assess which types of structures that are excluded by these models. Also parameterization of distributions which are alternatives to the Gaussian is even harder than for the Gaussian distribution.

The argument from information theory is that the Gaussian distribution is the least structured distribution among any distributions which have the same first two moments. We make the least amount of additional assumptions regarding the structure in the problem, (i.e. the Gaussian distribution has the maximum entropy among all distributions.) This off course has the drawback that sharper conclusions can be drawn if more information is put into the prior distribution. But one might claim that the Gaussian choice gives a conservative estimate of the local information content.

2.2 Influence of information on local variables

The previous section gives the argument for the marginal local likelihood approximation. In this section we discuss how we should parameterize this information. Thus we assume the prior and posterior distributions of seismic parameters are given under the Gaussian approximation.

For any selection of linear combinations of seismic parameters, we consider the prior and posterior distribution of the Gaussian approximation, given by:

$$p(\mathbf{m}_L) \approx N\left(\boldsymbol{\mu}_L, \boldsymbol{\Sigma}_L\right)$$
$$p(\mathbf{m}_L \mid \mathbf{d}) \approx N\left(\boldsymbol{\mu}_{L|d}, \boldsymbol{\Sigma}_{L|d}\right)$$

Through these relations we find that the likelihood is given by the relations:

$$\mathbf{d} = \mathbf{m}_L + \boldsymbol{\varepsilon}$$
$$\mathrm{Cov}\{\boldsymbol{\varepsilon}\} = \left(\boldsymbol{\Sigma}_{L|d}^{-1} - \boldsymbol{\Sigma}_L^{-1}\right)^{-1}$$
$$\mathbf{d}_{\text{obs}} = \left(\boldsymbol{\Sigma}_{L|d}^{-1} - \boldsymbol{\Sigma}_L^{-1}\right)^{-1}\left(\boldsymbol{\Sigma}_{L|d}^{-1}\boldsymbol{\mu}_{L|d} - \boldsymbol{\Sigma}_L^{-1}\boldsymbol{\mu}_L\right).$$

However as commented in section 2.1.3 the matrix $\boldsymbol{\Sigma}_{L|d}^{-1} - \boldsymbol{\Sigma}_L^{-1}$ need not be invertible. To better investigate the information present in the data with respect to the target parameter, we identify linear combinations of the target parameter, say $f_j = \mathbf{v}_j{}^T\mathbf{m}_L$, which are statistically independent both in prior and posterior distribution. This is obtained by solving a generalized eigenvalue problem. That is by identifying $\mathbf{v}_j$ and λ_j such that:

$$\mathbf{v}_j^T \cdot \boldsymbol{\Sigma}_{L|d} = \lambda_j \cdot \mathbf{v}_j^T \boldsymbol{\Sigma}_L.$$

The factors $f_j = \mathbf{v}_j{}^T\mathbf{m}_L$ are independent in both the prior and posterior distribution, but the variance is a factor $0 \le \lambda_j \le 1$ lower than the prior. Now note that if $\lambda_j = 1$ this means that the prior and posterior distribution has identical variance, thus a feature corresponding to a generalized eigenvalue of 1 is unidentifiable and should be left out in the computations. On the other side if $\lambda_j = 0$ this means that the feature is fully identified by the data. The linear features better control of the observations, and gives an equivalent set of independent observations. We might identify the local likelihood approximation as independent observations of features f_j. When the prior and posterior distribution of f_i is denoted:

$$p(f_j) \approx N\left(\mu_j, \sigma_j{}^2\right),$$
$$p(f_j \mid \mathbf{d}) \approx N\left(\mu_{j|d}, \sigma_{j|d}{}^2\right) = N\left(\mu_{j|d}, \lambda_j \sigma_j{}^2\right),$$

where $\mu_j = \mathbf{v}_j^T\boldsymbol{\mu}_L, \sigma_j{}^2 = \mathbf{v}_j^T\boldsymbol{\Sigma}_L\mathbf{v}_j$, and $\mu_{j|d} = \mathbf{v}_j^T\boldsymbol{\mu}_{L|d}, \sigma_{j|d}{}^2 = \mathbf{v}_j^T\boldsymbol{\Sigma}_{L|d}\mathbf{v}_j$, we find the equivalent observation to be:

$$f_j^{\text{obs}} = \frac{\mu_{j|d} - \lambda_j\mu_j}{1-\lambda_j}.$$

where

$$f_j^{\text{obs}} = f_j + \varepsilon_j,$$

and

$$\varepsilon_j = N\left(0, \sigma_j{}^2\left(\frac{\lambda_j}{1-\lambda_j}\right)\right).$$

From this expression we also see the interpretation of the observation of a feature with $\lambda_j = 1$, which means that is has zero precision (the uncertainty of the observation is infinite). By only including features with $\lambda_j < 1$ we avoid problems with precision matrix of the likelihood being of reduced rank.

In seismic inversion the information content in some factors will be very low. If $\lambda_j \approx 1$, this means that the data carries marginal information about the factor (the error variance tends to infinity as lambda goes to one). It is natural to exclude factors which correspond to a lambda larger than a threshold, e.g. 0.99. When a limited number of parameters carry the information, this will simplify the complexity of the computations.

2.3 Approximations for rock physics

Assume the local likelihood is given $p(\mathbf{d}\,|\,\mathbf{m}_i)$, and that we have samples from the joint distribution $p(\mathbf{m}_i, \mathbf{r}_i)$. The simplest approximation is to include $\mathbf{r}_i$ in the Gaussian approximation, whereas a more detailed approach uses importance sapling. There is also a third option which is used in Buland et al (2008) where some of the calculations can be done in a pre-process, this does however require that the dimension of the factors that carry information is small (preferably three or less) .

2.3.1 Full linearization

In this approximation we make a Gaussian distribution

$$p(\mathbf{m}_i, \mathbf{r}_i) \approx N\left(\begin{bmatrix}\boldsymbol{\mu}_m \\ \boldsymbol{\mu}_r\end{bmatrix}, \begin{bmatrix}\boldsymbol{\Sigma}_m & \boldsymbol{\Sigma}_{m,r} \\ \boldsymbol{\Sigma}_{r,m} & \boldsymbol{\Sigma}_r\end{bmatrix}\right)$$

Thus we have:

$$\begin{aligned}\mathbf{r}_i &\approx \boldsymbol{\mu}_r - \boldsymbol{\Sigma}_{r,m}\boldsymbol{\Sigma}_m^{-1}\boldsymbol{\mu}_m + \boldsymbol{\Sigma}_{r,m}\boldsymbol{\Sigma}_m^{-1}\mathbf{m}_i + \boldsymbol{\varepsilon}_r \\ \boldsymbol{\varepsilon}_r &= N\left(\mathbf{0}, \boldsymbol{\Sigma}_r - \boldsymbol{\Sigma}_{r,m}\boldsymbol{\Sigma}_m^{-1}\boldsymbol{\Sigma}_{m,r}\right).\end{aligned}$$

Since the posterior distribution of the elastic parameters in the target location is given in the inversion we get:

$$\begin{aligned}\mathrm{E}\left\{\mathbf{r}_i \,|\, \mathbf{d}\right\} &= \boldsymbol{\mu}_{r|d} \approx \boldsymbol{\mu}_r - \boldsymbol{\Sigma}_{r,m}\boldsymbol{\Sigma}_m^{-1}\left(\boldsymbol{\mu}_m - \boldsymbol{\mu}_{m|d}\right) \\ \mathrm{Var}\left\{\mathbf{r}_i \,|\, \mathbf{d}\right\} &= \boldsymbol{\Sigma}_{r|d} \approx \boldsymbol{\Sigma}_r - \boldsymbol{\Sigma}_{r,m}\boldsymbol{\Sigma}_m^{-1}\left(\boldsymbol{\Sigma}_m - \boldsymbol{\Sigma}_{m|d}\right)\boldsymbol{\Sigma}_m^{-1}\boldsymbol{\Sigma}_{m,r}\end{aligned}$$

Thus we find the marginal Gaussian likelihood for $\mathbf{r}_i$:

$$\begin{aligned}l(\mathbf{r}_i) &\propto N\left(\mathbf{r}, \boldsymbol{\Sigma}_{d|r}\right) \\ \boldsymbol{\Sigma}_{d|r} &= \left(\boldsymbol{\Sigma}_{r|d}^{-1} - \boldsymbol{\Sigma}_r^{-1}\right)^{-1} \\ \mathbf{r}_{\text{obs}} &= \boldsymbol{\Sigma}_{d|r}\left(\boldsymbol{\Sigma}_{r|d}^{-1}\boldsymbol{\mu}_{r|d} - \boldsymbol{\Sigma}_r^{-1}\boldsymbol{\mu}_r\right).\end{aligned}$$

This likelihood is identical to the one in section 2.2, and can be interpreted similarly.

2.3.2 Importance weighting

An alternative approach which only uses the linearization for the elastic parameters is the importance weighting. This is in particular of interest if one is interested in detailed analysis of the parameters. It is however more time consuming to compute. In this approach we use the local likelihood of the elastic parameter to compute the importance weight of samples from the prior distribution. Assume that the local marginal likelihood is given by the relations above, and assume we have S samples from the prior distribution.

$$\left(\mathbf{m}_i^1, \mathbf{r}_i^1\right), \left(\mathbf{m}_i^2, \mathbf{r}_i^2\right), \ldots, \left(\mathbf{m}_i^S, \mathbf{r}_i^S\right)$$

For simplicity we further compute factors from section 2.2 giving:

$$\left(\mathbf{f}^1,\mathbf{r}_i^1\right),\left(\mathbf{f}^2,\mathbf{r}_i^2\right),...,\left(\mathbf{f}^S,\mathbf{r}_i^S\right)$$

We now assign a weight to each sample being, proportional to the local marginal likelihood. That is the weight of $\left(\mathbf{m}_i^k,\mathbf{r}_i^k\right)\Leftrightarrow\left(\mathbf{f}_i^k,\mathbf{r}_i^k\right)$ is:

$$w_k=\prod_{j=1}^{L} l(f_i^k)$$

Thus properties of the local posterior distribution of $\mathbf{r}_i$ can be evaluated using the importance weight of the sample. Thus if we want to find the expected value of a function, say $g(\cdot)$, of the rock physics parameters $\mathbf{r}_i$ this is approximated by the standard weighted sum:

$$\mathrm{E}\left\{g\left(\mathbf{r}_i\right)\right\}\approx\sum_{k=1}^{S} w_k g\left(\mathbf{r}_i^k\right)\Big/\sum_{k=1}^{S} w_k .$$

The local posterior distribution itself can be approximated with a kernel estimate:

$$p\left(\mathbf{r}_i\right)\approx\sum_{k=1}^{S} w_k \tfrac{1}{h} K\left(\tfrac{\mathbf{r}_i-\mathbf{r}_i^k}{h}\right)\Big/\sum_{k=1}^{S} w_k .$$

3 Conclusions

The proposed approach for retrieving the local likelihood in linear inversion is simply to use a Gaussian- linear inversion approach, and derive the local likelihood as the ratio of the prior distribution to that of the posterior.

References:

Kolbjørnsen, Odd and Heidi Kjønsberg (2011), "Joint 4D inversion of multiple data sources for CO_2 monitoring", NR-Note: SAND/20/11

Kjønsberg, Heidi and Odd Kolbjørnsen (2011). "Making a 4D seismic prior model from rock physics relations", NR-Note: SAND/22/11

Buland A., O. Kolbjørnsen, R. Hauge, Ø. Skjæveland, and K. Duffaut (2008), "Bayesian lithology and fluid prediction from seismic prestack data", Geophysics, 73 NO. 3. MAY-JUNE 2008; P. C13–C21