EAGE

26538

Framework for seismic inversion of full waveform data using sequential filtering

M. Gineste* (Norwegian University of Science and Technology) & J. Eidsvik (Norwegian University of Science and Technology)

SUMMARY

Subsurface velocity inversion using full waveform modelling continues to be a challenging problem. Instead of approaching it as a deterministic optimization problem, it is here formulated in a probabilistic framework as a filtering problem, using shot data sequentially to update the estimation procedure. Such an approach has the potential of being more robust to e.g. noise and starting guess, but comes at the cost of more forward model evaluations. We present a small-sized synthetic example of using a sequential filtering method known as the Ensemble Kalman Filter for seismic velocity inversion and conclude with an outline of direction for further investigations.

Introduction

Seismic data is extremely important for reliable petroleum reservoir assessment. It is used in several stages of reservoir exploration such as imaging and interpretation of geological horizons, as direct hydrocarbon indicators at large scale and for smaller scale reservoir characterization from amplitude data (post and pre-stack), as well as during reservoir development (time-lapse seismic). Here, the focus is on seismic velocity analysis or seismic inversion.

The seismic forward problem can be formulated as follows; knowing the seismic source signature, the source/receiver geometry, and the subsurface properties, what is the wavefield at receivers. The inverse problem is as follows; knowing the source, the source/receiver geometry, and observations of the wavefield at receivers, what are the subsurface medium properties. Whereas the forward problem is fairly straightforward, given today's numerical methods and computing power, the inverse problem is notably more difficult. As the impulse from the source propagates downwards through the subsurface, changes in medium properties cause reflections which are later registered in the reflection seismograms. However, many configurations of the subsurface properties could result in the same or similar observed wave which makes the determination of the subsurface very complicated.

Faster computers have made the methodology known as Full Waveform Inversion (FWI) a viable alternative for seismic inversion. Compared with more traditional solutions, it provides a more realistic modeling of the physics and as the full waveform signal has a larger bandwidth, resolution of smaller scales in the subsurface is supposed to be obtainable in the inversion. The FWI methodology is based on formulating an optimization problem, seeking to minimize the misfit between simulation and observation. For a good introduction to the topic and its components, see Fichtner (2010). Although this approach has matured over the last decade, with large amounts of research efforts, there are still challenges to overcome.

Instead of regarding FWI as an optimization problem, we will here phrase it as a statistical sampling problem, where the goal is to draw posterior samples of the subsurface parameters, given the seismic waveform amplitude data. We will integrate the data sequentially (by shots) and rely on a sequential filtering method for posterior assessment. The Ensemble Kalman filter (EnKF) is one of the most used such algorithms for complex large dimensional problems and an example of its use in seismic inversion is presented here.

Method

We seek to estimate the subsurface material properties in a target region, while the material properties in the remaining part of the domain are assumed known and fixed. The material properties are acoustic wave velocity v_p, shear wave velocity v_s and density ρ, but here only acoustic wave velocity is considered, so the shear wave velocity is set to zero throughout the domain and the density is assumed known and constant.

The random variable m_i describes the material property in a sub-region R_i so that the region of interest is parametrized by $\boldsymbol{m} = [\boldsymbol{m}_1, \boldsymbol{m}_2, \ldots, \boldsymbol{m}_{n_m}]^T$, where n_m is the number of variables. This vector constitutes the unknown states in the problem.

The predicted data is $\boldsymbol{d} = g(\boldsymbol{m})$, where the operator $g(\cdot)$ is the forward model operator. It implicitly holds the source and receiver locations, and processing of the measured wavefield. The processing of the recorded pressure field is minimal. A common muting period is applied to all common shot traces, the length of this period is determined by the source impulse to reach the sea floor. Tapering is applied to give smooth transitions and finally the signal is resampled. An example of such synthesized traces is shown in Figure 1, with 32 receiver locations and the shot location a little left of the middle. These traces contain the seafloor and water surface reflections, the water column waveguide effect and subsurface reflections. Seismic data processing tries to reduce the unwanted effects of the water and enhance the reflection information, but in this abstract such processing is not performed. The forward model is a numerical simulation of the acoustic wave equation, where the output consists of synthetic reflection seismograms at fixed receiver locations. To produce these synthesized seismograms the SPECFEM2D (see e.g. Komatitsch and Tromp (1999)) code is used, but the method is not restricted to this software. As the name implies, the 2D code uses a spectral element formulation. The software involves the necessary components to adequately reflect an offshore seismic exploration setup.

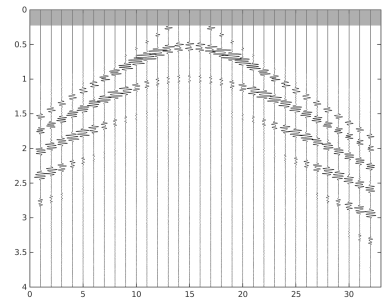


Figure 1: Example of synthesized seismic reflection recording. Shaded area indicates muted period.

The sequential filter

The core of the EnKF method is linear updating, where the approximation of the error covariance matrices uses Monte Carlo samples of an ensemble of states. The EnKF implementation of the FWI application differs from its more widespread use where the state variables are part of a dynamic system. In the current approach the state variables are assumed to be static. The sequential updating of the state variables is what Evensen (2009) refers to as stochastic optimization. An EnKF version with a perturbed measurement is used, having an additive error term with assumed known covariance in the measurement equation.

Consider a sequence of shots $s = 1,\ldots,n_s$ and from each of these there is measured data $\boldsymbol{d}_s^o$ available. For now, all recorded traces are concatenated in a single data vector $\boldsymbol{d}_s^o = \left[\boldsymbol{d}_{s,1}^o, \boldsymbol{d}_{s,2}^o, \ldots, \boldsymbol{d}_{s,n_r}^o\right]^T$ with n_r being the number of receiver locations. The EnKF runs over this sequence and updates the current state estimate by taking into account new observations.

The model formulation is

$$\begin{aligned} \boldsymbol{m}_s &= \boldsymbol{m}_{s-1} + \boldsymbol{e}_{1,s} \quad &, \quad \boldsymbol{e}_{1,s} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_{1,s}) \\ \boldsymbol{d}_s &= g(\boldsymbol{m}_s) + \boldsymbol{e}_{2,s} \quad &, \quad \boldsymbol{e}_{2,s} \sim \mathrm{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_{2,s}) \end{aligned}$$

where the error terms are independent, and the process noise terms are miniscule. The model specification is completed with a prior on $\boldsymbol{m}_0$. We assume a Gaussian multivariate prior model, so that the initial forecast ensemble $\boldsymbol{m}_0$ is simulated from a Gaussian random field distribution. The correlation structure can be formulated to give physical meaning and ensure reasonable input to the forward model. It further constrains the posterior solution in this complicated inverse problem.

Assume that n_e ensembles are provided at time step $s-1$. Initially, we generate all these from the Gaussian prior model. The EnKF builds a Monte Carlo approximation of the Kalman weight matrix based on the ensemble members $\boldsymbol{m}^i$ and $\boldsymbol{g}^{(i)} = g(\boldsymbol{m}^{(i)})$ at step s. The assimilated (a) ensemble members are obtained as follows:

$$\begin{aligned} \boldsymbol{m}_s^{a(i)} &= \boldsymbol{m}_s^{f(i)} + \widehat{\boldsymbol{K}}_s \left(\boldsymbol{d}_s^{o(i)} - \boldsymbol{g}^{(i)}\right) \quad , \quad \boldsymbol{d}_s^{o(i)} = \boldsymbol{d}_s^o + \boldsymbol{e}_{2,s}^{(i)} \\ \widehat{\boldsymbol{K}}_s &= \widehat{\boldsymbol{\Sigma}}_{md,s} \widehat{\boldsymbol{\Sigma}}_{d,s}^{-1} \\ \widehat{\boldsymbol{\Sigma}}_{md,s} &= \frac{1}{n_e - 1} \sum_{i=1}^{n_e} \left(\boldsymbol{m}_s^{(i)} - \bar{\boldsymbol{m}}_s\right)\left(\boldsymbol{g}_s^{(i)} - \bar{\boldsymbol{g}}_s\right)^T \\ \widehat{\boldsymbol{\Sigma}}_{d,s} &= \frac{1}{n_e - 1} \sum_{i=1}^{n_e} \left(\boldsymbol{g}_s^{(i)} - \bar{\boldsymbol{g}}_s\right)\left(\boldsymbol{g}_s^{(i)} - \bar{\boldsymbol{g}}_s\right)^T + \boldsymbol{\Sigma}_{2,s} \end{aligned}$$

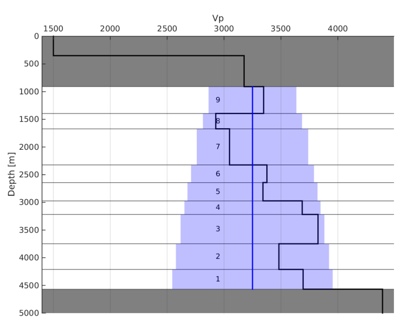
and the filter continues with the forecast (f) defined by $m_{s+1}^{f(i)} = m_s^{a(i)}$ in our case.

The EnKF method is known to suffer from ensemble collapse for certain problems. For the static problem considered here, this occurs almost immediately. Without the use of inflation, the ensemble will have zero spread after a few steps and the updates add nothing. The result is bias and underestimation of uncertainty. Ensemble inflation can be performed as $m^{a(i)} = \alpha m^{a(i)} + (1-\alpha)\overline{m}$, with a factor $\alpha > 1$. This kind of inflation ensures that the ensemble mean is maintained and the members are spread out. There are many approaches for determining the factor α in the literature, but none seems to handle all problems equally well, and tuning the parameter is a case-specific problem. In this presentation the methodology suggested by Saetrom and Omre (2013) is used to inflate (de-couple) the ensemble members.

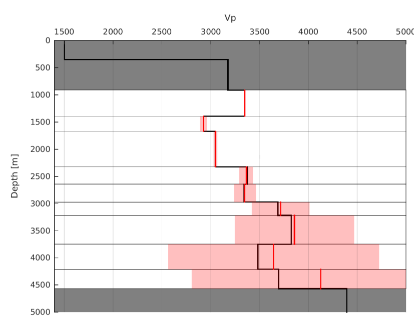
Example

A small example is studied to illustrate the EnKF for subsurface velocity inversion. A domain of $5\,\text{km} \times 5\,\text{km}$ is used and simulation starts with source impulse at time $t = 0$ and recording at receiver locations stops at $t = 4\,\text{s}$. The receiver locations have spacing 150 m and they are positioned at 50 m depth, in a water layer of 350 m depth. This entails 32 receiver positions. Sources are placed sequentially above each receiver at depth 10 m.
The subsurface velocity distribution is one-dimensional, so that there is no complicated horizontal variation in medium properties. The subsurface is divided into 11 layers with thickness' in the range 300-500 m, thick enough that tuning effects should not be a concern.
White noise were added to the simulations to generate the synthetic observations. The added noise terms have a variance determined by the signal of the particular trace, so that a SNR of 13 dB is kept for all observations traces. These variances in turn defines $\mathbf{\Sigma}_{2,s}$, $s = 1, 2, \ldots, n_s$.

Over 20 shots, the filtering takes place using an ensemble size $n_e = 50$. Variable vector $\boldsymbol{m}$ of size $n_m = 9$ indicates the layer properties, with interface locations identical to the ones used to generate the observations. Thus, the goal is to study whether the EnKF can estimate the right combinations of material properties through their observed noisy seismic contrasts.
The true model and the properties for the initial ensemble generation are depicted in Figure 2a, resulting in an ensemble with the same mean throughout the domain and a wider spread towards the bottom. The results of the filtering is shown in Figure 2b with indicated 95% confidence bands. Had the filter been allowed further steps, the bias would have decreased and the confidence bands narrowed in.



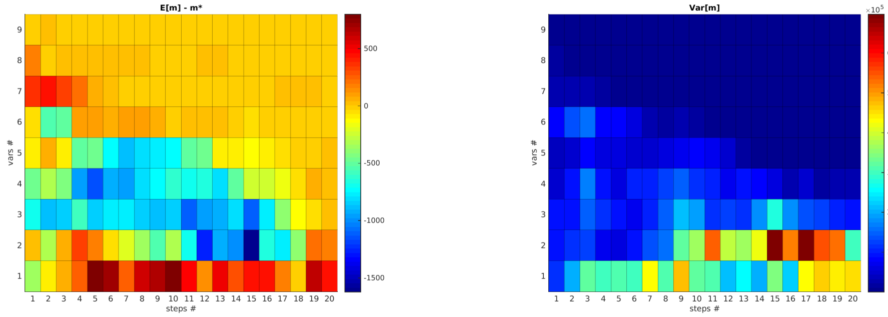
(a) True model (black line), initial mean and 95% confidence interval (blue line and blue shaded area) for ensemble generation. Gray shaded areas are fixed. Numbers are variable indices.



(b) True model (black line), parameter estimates and 95% confidence interval (red line and red shaded area). Gray shaded area are fixed.

Figure 2: Model used in example, initial mean for material distribution and its $\pm 1.96\sigma$ band (left), and estimated material and $\pm 1.96\hat{\sigma}$ band (right).

In Figure 3 is displayed (a) how the filter estimates the velocities top-down and (b) with diminishing variance. This is promising as if the top layer velocities does not get close to their true value, it will be



(a) Bias in ensemble mean after updates.

(b) Variance of ensemble after updates.

Figure 3: Results after EnKF update steps. Bias in variable estimates (left) and variance (right).

increasingly difficult to estimate the lower layers.

Conclusions

As the small-size example presented here shows, it is principally possible to perform seismic waveform inversion using a sequential filtering approach. The stochastic sampling comes at a cost of an increased number of forward model evaluations which might be outweighted by an increased robustness to noise and starting point for the estimation. Moreover, sampling approaches naturally represent the uncertainty of the solution, allowing probabilistic statements from the analysis.
Some of the future areas to be investigated are:

- EnKF is simple to implement and suitable for problems with high dimensional state vector, but the seismic forward model is highly nonlinear, and linearized updating might limit applicability. Modifications to the EnKF or the model formulations could be more suitable for this particular type of problem.
- With the need for inflation to prevent ensemble collapse, the approach for determining the inflation factor is also a ”tuning parameter”. One could imagine other approaches being more suitable for this specific type of problem.
- Data could be used differently than in this example, as the time signal might not be the best way to extract the information carried in the reflected wave on variation in medium.

With these challenges ahead, it is still expected that the use of (sequential) probabilistic approaches can bring new and useful elements to the field of full waveform inversion.

Acknowledgments

We thank the sponsors of the Uncertainty in Reservoir Evaluation (URE) consortium at the Norwegian University of Science and Technology.

References

Evensen, G. [2009] *Data Assimilation*. Springer, ISBN http://id.crossref.org/isbn/978-3-642-03711-5, doi:10.1007/978-3-642-03711-5.

Fichtner, A. [2010] *Full Seismic Waveform Modelling and Inversion (Advances in Geophysical and Environmental Mechanics and Mathematics)*. Springer, ISBN 3642158064.

Komatitsch, D. and Tromp, J. [1999] Introduction to the spectral element method for three-dimensional seismic wave propagation. *Geophysical Journal International*, **139**(3), 806–822, ISSN 1365-246X, doi:10.1046/j.1365-246x.1999.00967.x.

Saetrom, J. and Omre, H. [2013] Uncertainty quantification in the ensemble Kalman filter. *Scandinavian Journal of Statistics*, **40**(4), 868–885, ISSN 0303-6898, doi:10.1111/sjos.12039.