Introduction to this special section: AVO inversion

Vaughn Ball[1], Mosab Nasser[1], and Odd Kolbjørnsen[2]

The goal of AVO inversion is to estimate subsurface rock properties using seismic P-P reflection data as input. This article serves as a nonmathematical introduction to the topic. The reader will not find any equations here. Instead, we will use words and diagrams to introduce AVO inversion at a conceptual level and then set the context for the articles that have been contributed to this special section on AVO inversion.

The deterministic backstory to AVO inversion

The forward seismic problem can be represented in symbols as *seismic.model = function(rock.properties)*. As geophysicists, we are inclined to represent the inverse problem as the mathematical inverse of this: *rock.property_estimate = function^{-1}(seismic.model)*. In other words, our focus is on *function* and its mathematical inverse, *function*$^{-1}$. This is a *deterministic* inversion framework because it assumes that the inputs are known precisely and are not subject to errors, nor does it take into account the uncertainty in the match between the observed seismic data and the forward model. Unfortunately, the deterministic approach is not suitable for AVO inversion, and an underlying statistical model is required. We first explain why this is true before moving on to discuss the two major statistical frameworks used in AVO inversion.

Some input parameters to the AVO forward model have zero for an output response. Collectively, these parameters are called the *null space* of the forward model. This is illustrated in Figure 1. Most of the parameters input into the forward model will have some effect on the seismic response. However, this is not always the case. When an input parameter gives the output response of zero, we say this parameter is in the null space of the function. This situation arises in a variety of circumstances in AVO inversion. One example is an impedance profile that is constant with depth. There is no output response to a constant profile because the AVO forward model includes a derivative-like operator to compute the reflectivity. Even when an impedance profile changes with depth, the zero Hz component of the profile is in the null space of the derivative operator.

When a forward AVO model involves wavelets that are nonzero only within a certain frequency passband, then any elastic parameter component outside this passband is in the null space of the forward model. For example, if a thin bed is so thin that it is outside the passband of the wavelet, then the AVO response to the thin bed will be zero and the thin bed response is in the null space. Oftentimes, parameters are not in the null space but are near it. One example of this is the 5 Hz component of the elastic profile. Although the response to the 5 Hz component is not zero, it is nearly zero, so we can say that the 5 Hz component is near the null space. Similarly, if the seismic wavelet is nearly zero at some frequency, then the elastic profile at that frequency is near the null space of the forward model.

A related issue involves insensitivity to combinations of input parameters. Insensitivity implies that changing a parameter or combination of parameters has no effect on the AVO response. Take for example a binary sand/shale system. Changing the porosity of the sand component may not have an effect on the output so long as it is matched by just the right change in shale properties or the net-to-gross.

When a parameter is near the null space, or when the forward model is insensitive to a parameter or combination of parameters, then deterministic inversion fails. If a parameter is near the null space, then a large change in the parameter causes a small change in the modeled seismic data. When the inverse of the function is used to estimate parameters from seismic data, then the converse is true. That is, a small change in the measured seismic data results in a large change in the estimated parameter. This is a classic example of instability. Seismic noise can be thought of as random perturbations to the seismic data, and it results in large random perturbations to estimated properties when they are near the null space. A good example of this is that for fixed impedances, a large perturbation in density is required to make a visible difference on the seismic. Therefore, if a deterministic inverse is used, small amounts of noise on the seismic will result in large noise levels on the density estimate, even in the case when impedances are perfectly reconstructed.

The above paragraphs introduce some of the problematic issues associated with AVO inversion and explain why a deterministic inverse is not a suitable approach. The various real-world approaches to solving this kind of *ill-posed* inverse problem can be broadly divided into two categories based on their underlying statistical model. To many readers, it might come as a surprise that we need to discuss a statistical model in the inversion context, but the fact is that the statistical model plays an important role in AVO inversion, even in cases where it is not explicitly stated.

The statistical backstory to AVO inversion

To set the background for understanding the two statistical frameworks for inversion, we pose this question: are earth properties random? The frequentist inversion framework answers no; there is a fixed but unknown set of rock properties at any given

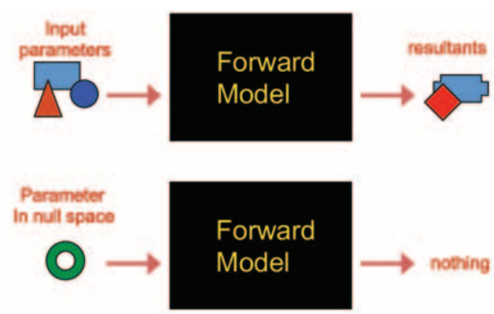

Figure 1. Under normal circumstances, an input parameter will have some effect on the output. When an input parameter is in the null space it will have no effect.

[1]Hess Corporation.
[2]Lundin Norway.

http://dx.doi.org/10.1190/tle35040399.1.

location in the subsurface. The task of inversion is to make an estimate of subsurface properties that is as close to this fixed truth as possible. By contrast, the Bayesian inversion framework represents the predrill subsurface properties using a probability distribution. This might sound unreasonable at first. After all, there is a truth about the subsurface; we just don't know what that truth is until a well is drilled. However, a simple analogy might help the reader understand why the Bayesian statistical framework has considerable merit.

Consider a flipped coin while it is still in the air. This is a classic example of a random variable where the probability of heads is 50/50. But what if the coin is caught and hidden until someone makes the call heads or tails? Is it still random? The answer is no; however, the state of knowledge is still 50/50, and so we would still say that the outcome is 50/50. Therefore, the Bayesian inversion framework recognizes that this is actually a statement about our state of knowledge, and not about the hidden coin itself. Analogously, the subsurface is already a flipped coin, and when we talk about probability distributions, it is actually a statement about our state of knowledge, not a statement about the subsurface.

To emphasize: the two broad approaches for dealing with the ill-posedness of the seismic forward problem are the frequentist and Bayesian approaches. Our starting point for discussing the differences between these two approaches is that the frequentist views the subsurface properties as fixed but unknown, while the Bayesian characterizes the prior understanding of the subsurface using a set of probability distributions. Deterministic inversion has no statistical model, and this is what differentiates it from both the frequentist and Bayesian approaches.

Most geophysicists feel more comfortable with the frequentist framework because it conforms to instincts about the subsurface and a dominant familiarity with the forward models. But there is significant value in the Bayesian approach, and it continues to gain popularity within the AVO-inversion community. We'll come to the Bayesian approach shortly, but first we will discuss the key elements of frequentist inversion.

The frequentist approach to AVO inversion

The frequentist approach to dealing with the issue of ill-posedness is to replace the original forward model with a new one, the inverse of which is *well-posed*. This process is called *regularization*, and it can appear in several forms. Take for example a three-term linear AVO equation that is parameterized in terms of intercept, slope, and curvature. One approach to regularizing the forward model is simply to drop the third term associated with curvature and perform a two-term inversion. By far, this is still the industry's most common approach to AVO regularization. It falls into a class of regularization approaches called *truncation*.

How does truncation work? If the original three-term model is more accurate, then why would we intentionally blunt that accuracy by throwing away one of the terms? We've actually already given the answer to this question above: when an input parameter or combination of input parameters imparts only a small change to the modeled seismic response, then small amounts of noise on the recorded data will result in very large changes to the low-order parameter estimate. This is truly a paradox! On the one hand, it seems desirable to have a forward model with higher precision, but, on the other hand, that extra precision causes the inverse problem to be more ill-posed. Take for example the Zoeppritz equation for P-P reflectivity. From a forward modeling perspective, Zoeppritz is presumably more accurate than one of the common three-term small-contrast approximations. And yet this precision is almost certainly blunted by some form of regularization when used in inversion. So, truncation works by eliminating those components of the forward model that are close to the null space.

Another important frequentist approach to regularization is to *augment* (or complement) the original forward model with other physical information. Instead of throwing out an unruly parameter from the forward model, we augment the forward model with constraint equations. For example, it is possible to augment the original AVO equations with additional equations that penalize the magnitude of the inverted parameters. This augmented structure translates directly into the familiar damped least-squares formulation. Sometimes, additional rock-physics information can be incorporated directly into the forward model. For example, if there is a known relationship between velocity and density, then this can be explicitly included in the forward model. Note the difference between this and the Bayesian approach: whereas the Bayesian approach employs a probability distribution to represent prior rock-physics information, the frequentist approach alters the forward model itself to incorporate rock-physics relationships.

Unfortunately, changing the original forward model in the frequentist approach introduces a new problem called *bias*. The informal meaning of bias is that although the answer is less noisy, it suffers a shift away from the true answer. A simulation experiment can be performed to illustrate this effect. Remember that the frequentist assumption is that there is a single true answer; this gives the following setup for the experiment:

1) Assume some fixed value for rock properties.
2) Forward model the data using the full-forward model.
3) Add random noise to the simulated data.
4) Invert using the inverse of the *regularized* forward model and store the result.
5) Repeat from 3 for each realization of noise.

Figure 2 shows a comparison of the result where the three-term AVO equation is used as a forward model. The histogram obtained for the estimate of AVO gradient is displayed for a full three-term inversion and a two-term truncated model. The histogram obtained with the three-term expression is centered directly on the true value (green line) but has a considerably wider spread than the estimate based on two terms. The two-term histogram, however, is shifted away from the true value of the AVO gradient; the difference between the mean of this histogram (red line) and the true value is the bias. The solution to a regularized frequentist AVO inversion will always be biased. The only way to avoid bias is to use the full unregularized deterministic inverse. Note however that for a given data set, we do not know the actual noise. The result of a single inversion will be just a random sample from the histogram corresponding to the estimator we have chosen. Thus, it might be safer to select an estimator that is less affected by random noise. This is a familiar topic in statistical estimation theory, which is known as the *bias-variance tradeoff*. The task of frequentist inversion is to choose regularization parameters that minimize the combination of errors due to noise and errors due to bias. Failure to account for bias can lead to some common pitfalls in analysis of AVO-inversion products. A good summary of the frequentist approach to inversion can be found in Hansen (2010).

The Bayesian approach to AVO inversion

The Bayesian approach to AVO inversion has steadily gained favor in recent years. Unfortunately, it is somewhat harder for nonspecialists to understand. We have already emphasized that the frequentist assumes a single fixed truth about the subsurface and deals with the issue of ill-posedness by directly altering the forward model, a technique called regularization. By contrast, the Bayesian approach views the subsurface — or the understanding of the subsurface — as a probability distribution. Furthermore, the Bayesian approach leaves the forward model unchanged and instead uses the prior information as a way to estimate the solution near the null space. A good reference for Bayesian inference is Carlin and Louis (2008).

In the section on frequentist inversion, we outlined a simulation analog that helps understand the statistical model that underlies the frequentist framework. The statistical model is sometimes referred to as the *sampling model*. The Bayesian inversion approach also has a sampling model. The differences in the sampling models can be understood by simulation analogs. In the frequentist-sampling model, the properties are held constant for each realization, and only the noise is changed. By contrast, in Bayesian simulation, both the noise and model parameters change with each realization. Values for the rock properties in each realization are drawn randomly from the prior distribution, passed through the forward model, and added to a realization of noise. Note that the simulation analogs discussed here are not actually used in inversion. We mention the analogs only as an aid in understanding the respective sampling models.

We can divide the Bayesian inversion framework into four components. We have introduced two already. The first is the forward model that describes the physics of the forward problem. The second is the probability distribution that characterizes the prior knowledge of the subsurface. This is called the *prior probability distribution*, or simply the *prior*s. The third component is the *noise model*, and the fourth is the *posterior probability distribution*. This fourth component represents another important difference between the frequentist and Bayesian frameworks: frequentist inversion gives a "best" estimate of subsurface properties, along with error bars, while Bayesian inversion strives to describe our *understanding* of subsurface rock properties using a probability distribution. Note again that this does not imply that the rock properties are actually random, but rather it is a description of our understanding of the rock properties. In practice, Bayesian AVO inversion often delivers a single estimate of subsurface properties. Bayesian AVO inversion also provides a sound approach to sampling from the posterior to generate multiple realizations of subsurface properties, as in *geostatistical inversion*. We now consider the four components of Bayesian AVO inversion one by one.

The forward model. In AVO inversion, there are actually *two* forward models. The *seismic forward model* is usually based on the

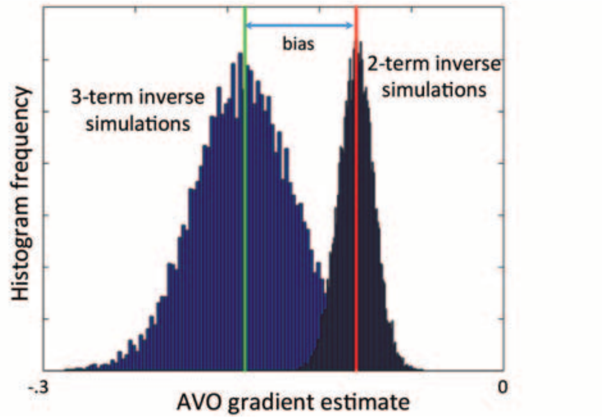

Figure 2. Three-term and two-term simulations of noisy estimation of gradient at a single interface. The two-term gradient estimate has much less noise, but the mean of all simulations is biased. The three-term estimate has almost no bias but carries more noise.

Zoeppritz P-P equation or one of its many linearized forms. The seismic forward model estimates seismic data given elastic rock properties. The *rock-physics forward model* estimates elastic properties given reservoir properties, such as porosity and mineralogy. In the context of AVO inversion, this raises an important question: is it best to invert directly to the reservoir properties, or is it best to first invert to the elastic properties and then follow this with a second inversion to the reservoir properties?

Two of our contributed papers follow a two-step approach. This approach has the distinct merit that it allows verification of the elastic results before going on to the rock-physics inversion. To be sure, the two-step approach will be most familiar to the quantitative interpretation community. For example, methods such as AVO crossplots, rock-physics templates, and seismic cokriging can be considered as a kind of interpretive rock-physics inversion step. More recently, this kind of interpretive approach to rock-physics inversion is being replaced by an actual inversion construct, such as the approach discussed in Nasser et al. in this issue.

By contrast, the contribution by Kolbjørnsen et al. advocates a one-step Bayesian approach. It can be readily shown in a Bayesian inversion context that the two-step approach will not give the same estimate of the posterior distribution as a one-step approach. In particular, the width of the posterior distribution associated with the two-step approach will be too narrow because it fails to account for the uncertainty in the estimates obtained in the first step. On the other hand, a good one-step workflow should include a verification component to ensure that the underlying elastic inference is reasonable. If the intent is to get a "best" estimate of rock properties with secondary emphasis on the distribution, then the two-step approach is a practical workflow that also facilitates verification.

The prior probability distribution (priors). Before an inversion is performed, we are not completely ignorant about rock properties. One of the more important types of prior information is the background elastic model that many AVO inversions use. In the case of P- and S-impedance inversion, these are the low-frequency models that are constructed from a combination of well control and seismic velocity data. Another type of prior information is shown in Figure 3. This is a crossplot of P- and S-impedance

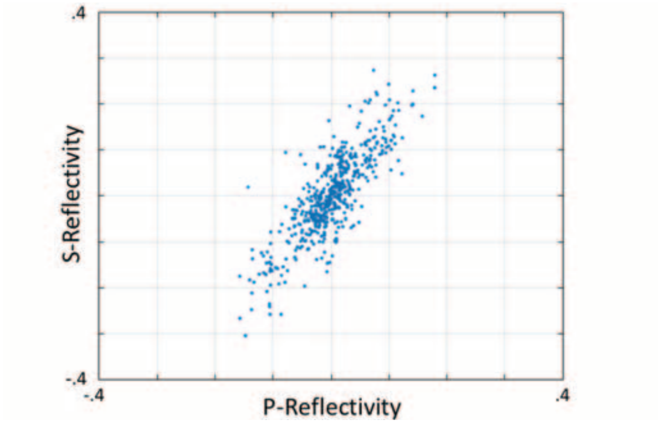

Figure 3. A crossplot of P- and S-reflectivity showing the expected positive correlation. Covariance of elastic properties and their reflectivities are examples of prior information, which can be incorporated in a Bayesian inversion.

reflectivities taken from a well log. It contains important prior information that could be included in an AVO inversion. It says that an increase in P-impedance *usually* will be accompanied by an increase in S-impedance. But what if the seismic is telling us that, at a particular interface, the S-impedance is increasing, while the P-impedance is decreasing? While there could be physical reasons for this — such as at the top of a gas sand — it could also be a mistake caused by noise in the seismic data. Bayesian AVO inversion is a formalism that merges the prior information and the seismic data to describe the posterior distribution. Thus, an essential part of the Bayesian analysis is to estimate the prior distribution by integrating additional information, creating an *informative prior*. Informative (or strong) priors are good if they are right but are a pitfall when they are wrong. Avseth et al., in this issue, give a good example of this. When we impose a prior distribution that is too strong, we claim to know more than we actually do. Note that the frequentist approach does not avoid the pitfalls associated with priors that are too strong. For example, a frequentist AVO inversion also employs a background model. If too much character is put into the background model, then this character will pass through to the inverted solution.

Other types of prior information can be considered. For example, we usually know something about the spatial juxtaposition of earth materials and that there are patterns about how facies are associated with each other in the subsurface. New approaches to using facies priors are actively being developed. One of the computational challenges with priors that describe spatial relationships is that the entire prior model becomes a single interconnected entity. Therefore, assumptions at one lattice point can influence the inference at other lattice points throughout the volume.

The noise model. From an inversion perspective, noise is a random quantity that is added to the seismic data. This is important; noise is something added to the data, not the elastic properties. It is not uncommon for analysts to note the scatter in the crossplot shown in Figure 3 and refer to this as noise. But this is not correct. Such scatter is the uncertainty inherent in elastic-property relationships. Even if we knew P-impedance exactly, we would still not know S-impedance. This uncertain relationship of rock properties is captured in the prior information as described above. By contrast, *noise* is something random and unpredictable that appears on the recorded seismic data. When an AVO inversion is performed, this noise is mapped back to the rock-property estimates. So, while we do not talk about noise of the true rock properties, we do talk about noise on the data and on estimates of the inverted rock properties. Confusion between noise and the variability implied by the prior distribution can lead to such incorrect heuristic statements as, "The density is noisier in AVO inversion because the density log data has more scatter." In fact, scatter in the density has nothing to do with why inverted density is often noisier than the other two properties. Instead, it is related to the forward model and the stability of the inverse.

In both the frequentist and Bayesian frameworks, the combination of the forward model and the noise model taken together represent the physics of the forward problem. Thus, both approaches require an estimate of the noise structure. (By noise structure, we mean the noise levels on the angle stacks and noise correlation between angle stacks.)

The posterior probability distribution. In Bayesian inversion, the entire posterior probability distribution is the "answer." To the extent the prior and the noise structure are correctly estimated, the Bayesian posterior distribution is an optimal trade-off between information from the seismic data and information from the priors. In some cases, we don't care about the whole distribution, and instead just want to know what the *best* single answer is. The mean is often chosen for this purpose because it minimizes a quantity called the *Bayes risk*, which can be thought of as the Bayesian equivalent of the frequentist concept of the mean-square error. The mean may often be difficult to compute, in which case the maximum of the posterior distribution can be a more accessible point estimator. This is called the maximum a posteriori solution, or MAP. When the priors and noise model are assumed to be Gaussian, the posterior is also Gaussian, and its mean and variance can be given in closed form. Furthermore, the mean corresponds to the MAP estimator. In practice, an assumption of a Gaussian prior may be unsatisfactory because the presence of different facies can imply a multimodal prior.

When priors associated with each facies can be adequately represented as Gaussian, then the composite prior is a *Gaussian mixture model*. In this case, and assuming a Gaussian noise model, the posterior will also be a Gaussian mixture model where each component of the mixture is associated with a particular facies. These sorts of considerations lie at the core of recent efforts to more reliably identify different facies using AVO inversion.

This contrasts with the approach of *Bayesian classification* of facies based on elastic-inversion results. Using Bayesian classification, the prior probabilities of the rock properties for each facies are used to identify facies based on elastic inversion. From a Bayesian perspective, this is not quite right, and developing a full facies-dependent posterior distribution would be considered more sound. As noted previously, sampling from the posterior distribution is one way to characterize the entire distribution. Such realizations represent possible subsurface models that are consistent with both the seismic and the prior rock-physics understanding, and are a good way to generate ranges in geostatic reservoir models. Note that these realizations are consistent with the seismic in a statistical sense. Moreover, the forward model of each realization is not required to exactly match the seismic, but will be consistent within the range allowed by estimates of noise and the prior distribution.

Finally, we return again to the topic of bias. As noted previously, bias is a frequentist concept and refers to the difference between the true value and the expected value of the inversion solution. The fact that the Bayesian approach does not include the notion of a true value means that the formal concept of bias isn't defined. But this should not be taken to imply that the Bayesian approach somehow avoids the underlying issue. The Bayesian inversion is always influenced by the prior distribution, in the same way the augmented forward model creates a bias in the frequentist approach. Indeed, the MAP obtained from a Gaussian prior is nothing but the frequentist damped least-squares estimate. Thus, if the Bayesian posterior mean is not interpreted in light of the prior, then this implies the same need for caution as in the frequentist approach.

Summary

AVO inversion is a core component of seismic rock physics and quantitative analysis. There are two major statistical frameworks for AVO inversion. The frequentist framework assumes that there is a fixed but unknown truth about the subsurface and seeks to estimate that truth as closely as possible. To achieve this, the frequentist approach modifies the original forward model such that its inverse is more stable. The goal of such regularization is to minimize the combination of bias in the inverted property and noise. A classic example of this is two-term AVO inversion, which is one type of truncation.

By contrast, the Bayesian framework describes understanding of the subsurface as a prior probability distribution. The inversion approach does not involve modification of the forward model to make it more stable. Instead, those data are used in combination with the forward model to produce a posterior probability distribution that characterizes understanding of the subsurface given the seismic data. Careful implementation in either framework will yield good results. The Bayesian framework is steadily gaining adherents because of the ease with which prior information can be included in the inversion and because of its ability to facilitate the idea of multiple solutions that are consistent with both the seismic data and the prior probability distribution.

This month's special section on AVO inversion includes papers that address a variety of issues encountered in the AVO-inversion framework discussed above. Here we give a brief summary of each.

Avseth et al. focus on the use of AVO methodology in the exploration setting and show key learnings from evaluating multiple prospects in the Norwegian Sea. The main message is that one should always have a critical view of the assumptions imposed by the methodology. On the data side, the authors show that the cutoff angle for when refraction energy dominates the reflection energy becomes essential. On the modeling side, the authors show the pitfalls of excessive confidence in the background model in regions far from well control.

Nasser et al. show a full-scale 4D case study where different types of data from geology, geophysics, and rock physics are integrated to provide a unified interpretation of production effects. In particular, the authors have used facies-specific low-frequency models as priors for the 3D inversion and velocity changes from time-lapse time shifts as priors for the 4D inversion.

Humberto and Dvorkin interpret inverted P- and S-impedances in terms of parameters of a quartz-clay system. Initially, they classify the fluid content based on the ratio of P- and S-impedance, then they interpret the total porosity and clay content assuming the fluid content is as classified. They emphasize the importance of the rock-physics model and the need for high-quality well data and inversion results.

Kolbjørnsen et al. emphasize that the spatial structure related to rock properties can improve the inversion results. Integration of spatial structure is shown to improve prediction capabilities and increase resolution compared to a standard inversion approach. The integration is done in a Bayesian framework, and results are obtained by a fast parallelizable algorithm.

Thomas et al. highlight the fundamental difference between the forward problem and the inverse problem in AVO inversion. They analyze the situation of a two-term AVO reflectivity inversion in detail and show that an approach based on forward-model reasoning introduces inversion bias. They further show how this error can be corrected using a debiased mapping from intercept and gradient into elastic properties. TLE

Corresponding author: mnasser@hess.com

References

Carlin, B. P., and T. A. Louis, 2008, Bayesian methods for data analysis, 3rd ed.: Chapman and Hall/CRC Texts in Statistical Science.

Hansen, P. C., 2010, Discrete inverse problems: Insight and algorithms: Society for Industrial and Applied Mathematics, http://dx.doi.org/10.1137/1.9780898718836.fm.