Non-Uniqueness in Inversion

John Doherty, Watermark Numerical Computing, Brisbane, Australia

Introduction

Inversion forms the basis of geophysical data interpretation. It is also important in other fields of applied science such as calibration of petroleum and geothermal reservoir models. Wherever it is used inversion faces the same challenge, namely how to extract as much information as possible from available data in order to best inform decisions.

The challenges that confront subsurface applications of inversion are great. Earth properties such as electrical conductivity and permeability can vary over orders of magnitude. Spatial heterogeneity of these properties can be high (particularly in hard rock environments), while constraints on the locations, disposition and properties attributable to anomalous areas forthcoming from direct outcrop or borehole measurements are normally few. Expert knowledge of the geological context can help; hence its role in the inversion process must be central. However such knowledge is not exact, being characterized more by certainty of what cannot exist beneath the surface than of what actually can exist. Hence expert knowledge is nonunique. It is, in essence, a probabilistic quantity which is best expressed through geostatistical methods, some very sophisticated forms of which have been developed to support petroleum reservoir modelling. See, for example, Strebelle (2002) and Journel (2002).

The purpose of this abstract is to present a short overview of how nonuniqueness arises in model-based data interpretation, and some of the ways in which it can be accommodated when modelling the response of the subsurface to natural and imposed excitations of various types. The premise of the approach taken herein is that inference of earth properties based on expert knowledge (often referred to as “soft data”), supplemented by point measurements of system state (often referred to as “hard data”) cannot support construction of a unique three-dimensional image of these properties. Rather the purpose of model-based data interpretation must be to encapsulate what we can know and quantify what we cannot know. This approach can then create a decision-making context in which risks are fully appraised before (often very expensive) investments are made.

The Null Space

Matrix theory provides a simple means of illustrating inversion nonuniqueness. See texts such as Menke et al (1989) and Aster et al (2013) for comprehensive treatments of this topic. Other approaches to inverse modelling do not rely on characterization of earth properties by a finite-sized vector. Many of these are based on the pioneering work of Backus and Gilbert (1968, 1970); see, for example, Press et al (2007).

Let the matrix **Z** represent earth processes of interest to us as they operate on earth properties comprising a vector **k**. We refer to the elements of **k** as “parameters”. **Z** thus represents a model which, as a matrix, is assumed to be a linear with respect to the parameters encapsulated in **k**. **Z** yields a set of outputs **o** (another vector) which describe the state of the system at points at which measurements of that state have been made. These measurements are encapsulated in the vector **h**; the noise associated with these measurements comprises the vector **ε**. The vector **k** may represent, for example, the electrical conductivity in every cell or element of the discretized domain of a numerical model which simulates generation of secondary electric and magnetic fields in response to primary excitation by one of these types of field types. The vector **o** may represent measurements of these fields at the surface of the earth. From the above it follows that:

$$\mathbf{h} = \mathbf{Zk} + \boldsymbol{\varepsilon} \qquad (1)$$

Suppose that the matrix **Z** has a null space (as it surely will if elements of **k** outnumber those of **h** – which is always the case). Then there exists one or more non-zero vectors **δk** for which the following equation holds:

$$\mathbf{0} = \mathbf{Z}\boldsymbol{\delta}\mathbf{k} \qquad (2)$$

The null space is composed of all vectors **δk**. The dimensionality of the null space is at least as high as the number by which observations outnumber parameters, and often far higher. By adding equations (1) and (2) together it is easily established that an immediate outcome of the existence of the null space is that the same observed system response to excitation can be obtained using different parameter sets. Hence uniqueness of inference of a parameter set which gives rise to a particular observed system response is impossible.

Bayesian Approaches

So what can be inferred from discrete measurements of system state at a discrete number of locations? And what is the best way to infer it? A number of different approaches are available for solving this problem. The Bayesian approach embraces nonuniqueness as it exists in both expert knowledge and in the constraints that hard data imposes on this knowledge. It can be expressed as:

$$P(\mathbf{k} \mid \mathbf{h}) \propto P(\mathbf{h} \mid \mathbf{k}) P(\mathbf{k}) \qquad (3)$$

In equation (3) $P(\mathbf{k})$ is the prior probability distribution of parameters, this being expert knowledge expressed in probabilistic terms (which, in principle, is the only way to express it). $P(\mathbf{h}|\mathbf{k})$ is the likelihood function; this increases with goodness of fit between model outputs and field observations as parameters of the former are varied within the constraints set by $P(\mathbf{k})$. $P(\mathbf{k}|\mathbf{h})$ is the posterior probability distribution of parameters conditioned on the observation dataset **h**.

Methods have been developed (particularly in petroleum reservoir engineer) to generate realizations of $P(\mathbf{k}|\mathbf{h})$ for use in Monte Carlo uncertainty analysis. Ideally, this provides us with many different "pictures" of the subsurface that are simultaneously constrained by both hard and soft data; see for example Zhou et al (2012) and Caers and Hoffman (2006). Unfortunately, however, probability distributions are, in general, numerically "too hot to handle" so that Bayesean analysis can only be applied in its purest form in relatively simple cases. Nevertheless, this type of analysis is the subject of intensive research.

Regularized Inversion

As the term "inversion" is commonly applied, its goal is too seek uniqueness where none actually exists. Achievement of this goal requires that some special conditions be placed on the parameter field that emerges from this process. A logical condition is that the resulting parameter field be of minimized error variance. It is important to realize what this means. If its potential for wrongness has been minimized, this does not mean that its potential for wrongness is small; this is a function solely of the information content of hard and soft data. Another way to view the desired parameter field is that it be the average of the infinite number of parameter fields that would be generated using the posterior parameter distribution of equation (3). Once again, knowing the average of such fields says nothing about their propensity for variability about this average.

Two broad categories of regularisation exist. The first, Tikhonov regularisation (formalized in works such as Tikhonov and Arsenin, 1977), manufactures uniqueness by supplementing the observation dataset **h** with expert knowledge. Ideally the "observed values" of the new members of the augmented observation dataset are parameter values (or the values of relationships between parameters) of minimum error variance from the point of view of expert knowledge alone, the latter being expressed by *P*(**k**) of Bayes equation. Suppose that the prior parameter probability distribution can be characterised by the covariance matrix C(**k**). Then the Tikhonov-based solution to the inverse problem is calculated using a formula such as:

$$\underline{\mathbf{k}} = (\mathbf{Z}^t\mathbf{QZ} + \lambda C^{-1}(\mathbf{k}))^{-1}\mathbf{h} \tag{4}$$

where **Q** is a (normally diagonal) weight matrix that ideally should be related to the covariance matrix of measurement noise as:

$$\mathbf{Q} = C^{-1}(\boldsymbol{\varepsilon}) \tag{5}$$

In practice definition of λ, the factor that balances the worth of hard data against that of soft data in the above equation, may be difficult to determine. Definition of C(**k**) also presents a challenge. A number of innovative linear and nonlinear approaches to definition of this term have been pursued in the geophysical literature.

In contrast to Tikhonov methods, subspace methods seek uniqueness through reduction rather than augmentation of the inverse problem. With singular value decomposition (SVD) as the flagship of these methods they define, and then eliminate from consideration, combinations of parameters which comprise the null space (and are hence inestimable). Only parameter combinations lying within the solution space (the orthogonal compliment of this) are estimated through what is then guaranteed to be a well-posed inverse problem. Suppose that SVD of **Z** leads to:

$$\mathbf{Z} = \mathbf{USV}^t \tag{6}$$

U and **V** define orthogonal model output and parameter combinations which are linked to each other through the singular values comprising the diagonal matrix **S**. A unique parameter set $\underline{\mathbf{k}}$ can be calculated as:

$$\underline{\mathbf{k}} = \mathbf{V}_1\mathbf{S}^{-1}{}_1\mathbf{U}^t{}_1\mathbf{h} \tag{7}$$

where $\mathbf{V}_1$, $\mathbf{S}_1$ and $\mathbf{U}_1$ are obtained through partitioning of **V**, **S** and **U** according to the dimensions of the solution space (this being equal to the number of significantly non-zero elements of **S**). Ideally, SVD should not be applied directly to estimation of **k**. Instead it should seek to estimate principal components of C(**k**), this ensuring a minimized error variance status for estimated parameters.

In practice, use of SVD to obtain **k** is numerically intensive where parameters number more than a few thousand. Approximate, but very powerful methods such as LSQR (Paige and Saunders; 1982a, 1982b) are used in their place. One of the attractions of subspace methods is that they allow a pictorial representation to be made of the cost of uniqueness. The estimated parameter field $\underline{\mathbf{k}}$ can be considered to be an orthogonal projection of the unknown true parameter field **k** into a much smaller dimensional subspace. What emerges from inversion is thus a "shadow" of reality. This is the cost of uniqueness.

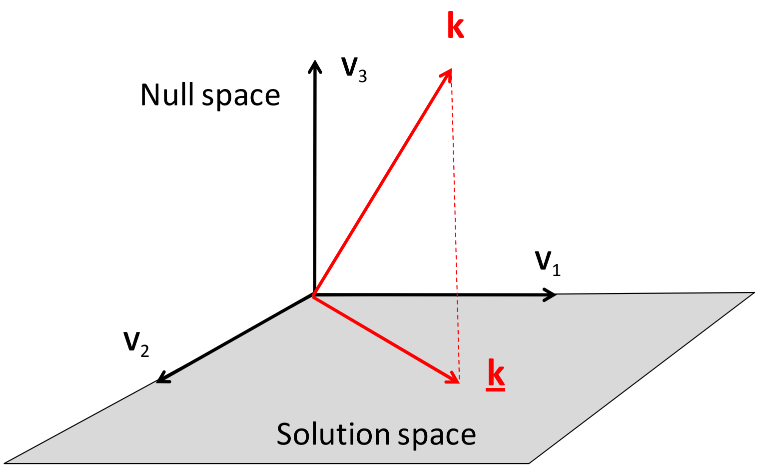

Figure 1. Subspace inversion estimates $\underline{\mathbf{k}}$, the projection of the real parameter vector **k** onto a smaller dimensional subspace.

Post-Inversion Uncertainty Analysis

Ideally, neither geophysical data analysis nor subsurface model calibration should stop at inversion, as the potential for error in $\underline{\mathbf{k}}$ may be very high. (It can be shown that if inversion is carried out wisely, this potential for error is almost equivalent to the posterior uncertainty of **k** as expressed by Bayes equation.) Expensive and high-risk exploration or management decisions should not be made without knowledge of how wrong an estimated $\underline{\mathbf{k}}$ can be. As stated above, this can be explored through Bayesian analysis. However the numerical cost of this is often too high. Approximate methods based on subspace analysis can achieve almost the same outcome at a much lower computational cost.

In subspace terms, post inversion uncertainty analysis (actually "error analysis" would be a better term to use for this type of analysis to avoid offending Bayesians) seeks to acquire a suite of parameter fields whose "shadow" on the solution space is $\underline{\mathbf{k}}$. This is illustrated in Figure 2. If parameter fields are assumed to be continuous rather than categorical in nature, this can be achieved using methods such as the Null space Monte Carlo method provided in the PEST suite of software (Doherty, 2013). Once the inverse problem has been solved once to obtain $\underline{\mathbf{k}}$, other parameter fields with the same shadow can then be obtained at relatively little numerical cost.

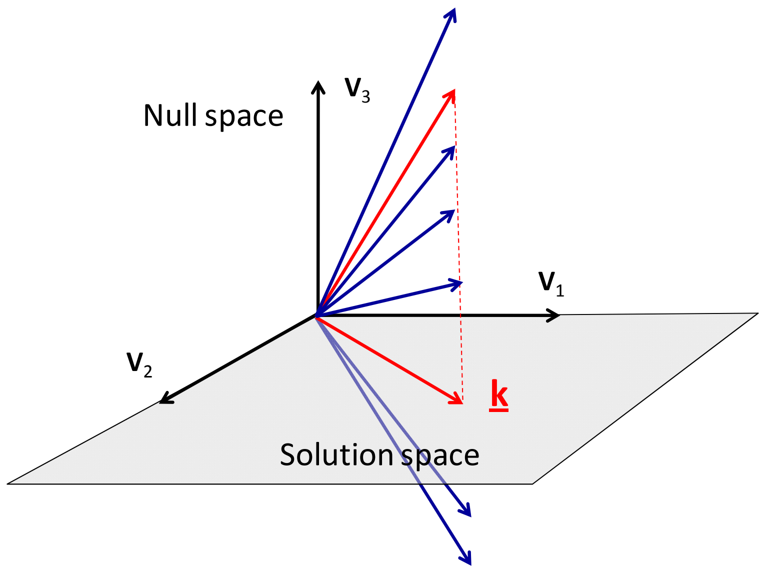

Figure 2. Many random parameter vectors which have the same solution space projection as **k** and hence all fit the observation dataset.

Unfortunately, the elusive goal of enforcing inversion constraints on complex realizations of geologically-based random fields involving discontinuous lithological categories has not yet been achieved in ways that are numerically tractable. However some progress has been made. Some fruitful was of combining the efficiencies of subspace methods with generation of geologically realistic categorical parameter fields are currently being pursued by the author.

References

Aster, R.C., Borchers, B. And Thurber, C.H., 2013. Parameter Estimation and Inverse Problems. Academic press.

Backus, G.E., and Gilbert, F. 1968. The resolving power of gross earth data.. Geophys. J. Roy. Astr. Soc., 16: 169-205.

Backus. G.E. and Gilbert. F., 1970. Uniqueness in the inversion of inaccurate gross earth data. Philo. Trans. Roy. Soc. Lond. A., 266: 123-192.

Caers, J. and Hoffman, T., 2006. The probability perturbation method: a new look at Bayesian inverse modelling. Math. Geol., 38 (1):81-100.

Doherty, J., 2013. PEST: Model Independent Parameter Estimation. Watermark Numerical Computing, Brisbane, Australia. Downloadable from: http://www.pesthomepage.org

Journel, A.G., 2002. Combining knowledge from diverse sources: an alternative to traditional data independence hypotheses. Math. Geol., 34 (5), 573-596.

Menke, W., 1989. Geophysical Data Analysis. Academic Press.

Paige, C.C. and Saunders, M.A., 1982a. LSQR: an algorithm for sparse linear equations and sparse least squares. ACM Trans. Math Softw, 8: 43-71.

Paige, C.C. and Saunders, M.A. 1982b. LSQR: an algorithm for spare linear equations and sparse least squares. ACM Trans. Math. Softw. Vol 8: 192-209.

Press. W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P., 2007. “Section 19.6. Backus-Gilbert Method”. Numerical Recipes: The Art of Scientific Computing (3rd ed.), Cambridge University Press, New York.

Strebelle, S., 2002. Conditional simulation of complex geological structures using multiple-point geostatistics. Math. Geol., 34, 1-22.

Tikhonov A.N., and Arsenin V.Y., 1977. Solution of Ill-Posed Problems. Washington DC: VH Winston and Sons..

Zhou, H., Li, L., Franssen, H,H., Gomez-Hernandez, J.J., 2012. Pattern recognition in bimodal aquifer using the normal score ensemble Kalman Filter. Math Geosci. 44, 169-185.