Constraining an inversion to follow curving trends in an image

Andrew King
CSIRO
26 Dick Perry Ave, Kensington WA 6151
Andrew.King@csiro.au

SUMMARY

This paper addresses the question of how to include structural information, for example from a magnetic image, into an airborne electromagnetic (AEM) inversion. The kind of information we are interested in is the trend directions seen in the magnetic image, such as strike directions of dipping bodies, or the shape of palaeochannels.

A commonly-used technique for including prior information is to use a model covariance matrix, describing the spatial covariance between different model points. However, these covariances are usually constructed from a stationary covariance function which is dependent on the vector distance between two points, but is the same for the entire model. However, if a palaeochannel is visible in the magnetics, then we know that the AEM model is more likely to be similar along the channel than away from the channel. We therefore wish to construct a covariance matrix that can take curved and branching structure into account.

We construct an inhomogeneous covariance matrix from an image by breaking the image up into multiple windows, and then computing an elliptical distance metric in each window, such that distances in the direction of the features in that window are shorter than distances across those features. This collection of distance metrics then allows us to compute, between any two points in the image, a shortest path that curves to follow the directions of trends in the image. Using this curved-path distance allows us to generate a covariance matrix that encourages the inverted model to follow the trends in the image.

Key words: Bayesian Inversion, Prior Information

INTRODUCTION

Geophysical inversions are usually non-unique, meaning that many different models could equally well fit the observed data. In a Bayesian setting, in order to overcome the problem of non-uniqueness, we include prior information, such as knowledge of the spatial correlation properties of the physical property whose values we are trying to infer, which has the effect of imposing smoothness constraints on the model. Geostatistical methods can be used to estimate parameters for model priors, typically via prior covariance matrices (see e.g. Kitanidis, 2010; Oh and Kwon, 2014).

The goal of the work described here is to incorporate more complex spatial constraints on the model. In particular, we may have an image covering the same area as the model we are trying to invert, such as an airborne magnetic image, or a seismic section. We know that the model ought to exhibit similar spatial structure to what is visible in the image, simply because the same earth that our model is trying to describe has generated the image. We therefore seek a method of generating spatial constraints, in particular a prior covariance matrix, from an image, such that random models drawn from a probability distribution function described by the covariance matrix show similar spatial features to those in the image.

One method for imposing similarity of two models, used in joint inversion of two different data sets, is to use a cross gradient technique (Gallardo et al., 2005) where the inversion algorithm minimises an objective function including the cross-product of the gradients of each of the two models. Where the gradients are parallel, which happens when the spatial structure is the same, the cross-product goes to zero.

Our method, following work by Boisvert (2010) divides the image into cells, in each of which a local anisotropy is estimated. This allows a distance function to be defined, whose shortest paths follow the directions of spatial structure in the image. Using this distance function to compute a covariance matrix yields a prior probability distribution function, samples from which display the spatial patterns of the original image.

METHOD AND RESULTS

A commonly-used way of incorporating prior information into a Bayesian inversion is to express the prior as a multivariate Gaussian, with a mean vector and a covariance matrix (Mosegaard and Tarantola, 2002). The log posterior probability distribution for the model parameters, in the case of Gaussian noise in the data, then has the form,

$$\log P\left(\mathbf{m}\right) \sim \left(\mathbf{f}\left(\mathbf{m}\right) - \mathbf{d}_{obs}\right)^T \mathbf{C}_D^{-1}\left(\mathbf{f}\left(\mathbf{m}\right) - \mathbf{d}_{obs}\right) + \left(\mathbf{m} - \mathbf{m}_{prior}\right)^T \mathbf{C}_M^{-1}\left(\mathbf{m} - \mathbf{m}_{prior}\right).$$

Here **m** is the model parameter vector, **f**(**m**) is the predicted data generated by our forward model, $\mathbf{d}_{obs}$ is the observed data, $\mathbf{C}_D$ is the data covariance matrix, $\mathbf{m}_{prior}$ is a prior model parameter vector, and $\mathbf{C}_M$ is the prior model covariance matrix. When the model is in the form of a 2D grid, like the conductivity of a layer in an airborne electromagnetic inversion, then the covariance matrix $\mathbf{C}_D$ describes spatial smoothness constraints that are applied to the inverted model.

The covariance matrix is square, with an entry for each pair of model parameters, describing the covariance between them. In the case of a 2D-grid model, the covariance is then a function of the positions of the two model elements. The functional form of the covariance describes the smoothness that we expect to see in the model, and normally decreases with distance between the two model points. There are several different forms that are used in the geostatistical literature (Oliver et al., 2008) In this work, we use two functional forms, the Gaussian,

$$C_M\left(m_1, m_2\right) = \sigma_1 \sigma_2 \exp\left(-\frac{D\left(\mathbf{r}_1, \mathbf{r}_2\right)^2}{L^2}\right),$$

and the exponential,

$$C_M\left(m_1, m_2\right) = \sigma_1 \sigma_2 \exp\left(-\frac{D\left(\mathbf{r}_1, \mathbf{r}_2\right)}{L}\right).$$

In these expressions, σ_k is the standard deviation of the uncertainty on the k'th parameter, $D(\mathbf{r}_1, \mathbf{r}_2)$ is a distance (not necessarily Euclidean) between $\mathbf{r}_1$ and $\mathbf{r}_2$, and L is a scale parameter. **Error! Reference source not found.** illustrates the effect of different functional forms: the Gaussian covariance produces smoother models than the exponential covariance; and different scale lengths: the larger the scale parameter L,
the smoother the resulting image.

Apart from the model smoothness, additional information on the model spatial structure can be included by modifying the distance function. For example, in a sedimentary environment, rock physical properties are likely to be more similar in the plane of the bedding than perpendicular to bedding. This information can be included by setting a distance measure that makes points in the bedding directions closer together than points across the bedding. This is illustrated in **Error! Reference source not found.**, using an elliptical distance measure, with increasing anisotropy.

We would like to be able to impose more general geometrical constraints on a model. For example, it is very common to have airborne magnetic data available, and a magnetic image is often used for structural interpretation of the geology. Since the magnetic data is produced by the same geological features that produced the data we are trying to interpret (say AEM, for example) our model should have similar spatial features to those seen in the magnetic image (or a suitably-processed version of the magnetics.) The same would be true to some extent of radiometric images, or even aerial photographs, though of course these are only directly related to shallow features of our model.

We therefore wish to take an arbitrary image which we believe has spatial features that our model should also exhibit, and derive from the image a covariance matrix whose corresponding prior probability distribution will yield models with the desired spatial properties. To do this, we follow the method of Boisvert (2010; Boisvert and Deutsch, 2011) who defines a distance function, on an image, whose "shortest paths" between points follow curved lines parallel to any structure in the image.

The algorithm has three steps:

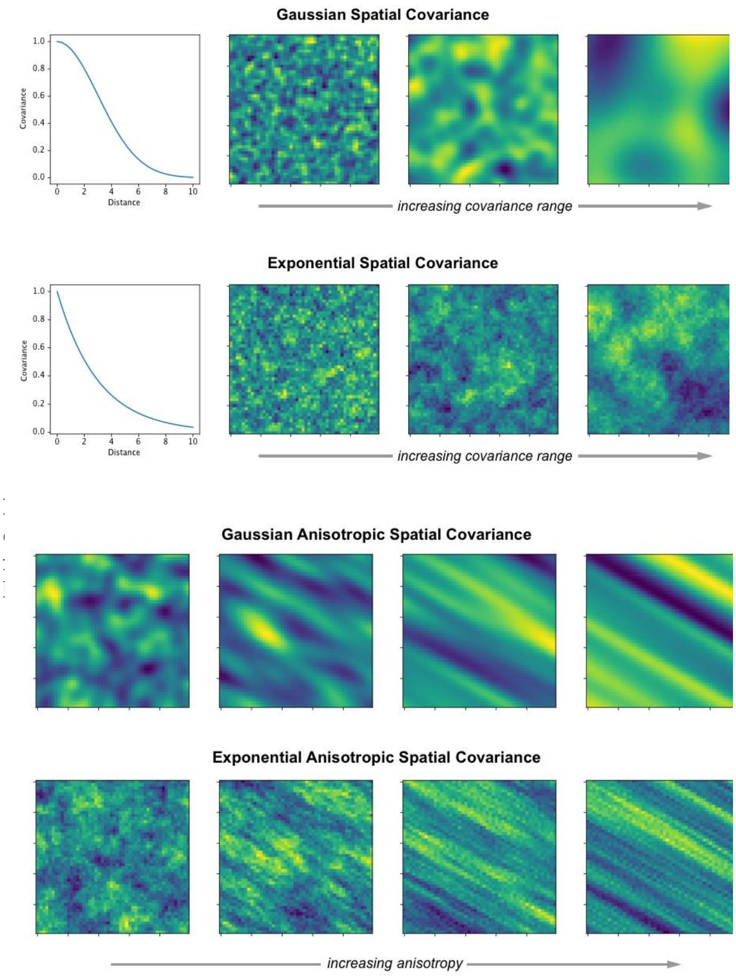

Figure 2. Examples of random images drawn from prior probability distributions using covariances with an anisotropic distance measure. (top) Gaussian, (bottom) Exponential.

1. The image is split into a grid of rectangular cells, and an elliptical locally-varying anisotropy (LVA) is estimated within each cell; the anisotropy is characterised by the direction and lengths of major and minor axes. This anisotropy defines a quadratic metric for measuring distance between any two points.
2. The distance between pairs of points in the desired model is computed, using a distance function where shortest-path curves follow the anisotropic metric.
3. A covariance matrix is computed from these distance pairs. This non-Euclidean distance often results in covariance matrices which are not positive definite. Since a covariance matrix has to be positive definite to make sense, we find a nearby matrix which is positive definite. See Boisvert (2010) for a discussion of methods for correcting the covariance matrix.

For the first step, Boisvert (2010) uses a moment of inertia computation to estimate the local anisotropy within a cell. We have found this method not be effective, so we instead compute a finite-difference gradient over the pixels in the cell, and use the variance to estimate the degree of anisotropy, and the mean strike direction to estimate the direction of the major axis. This is an aspect that requires further work, however. The top panel in **Error! Reference source not found.** shows an example image along with the estimated locally-varying

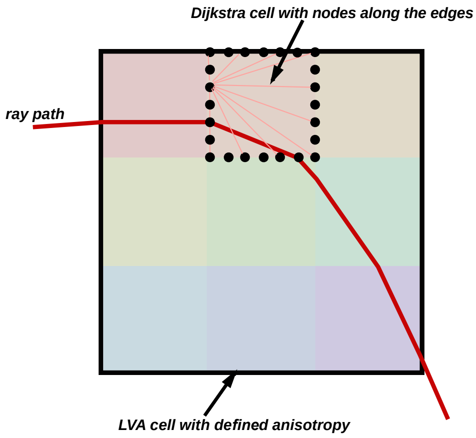

Figure 3. Each LVA cell has a single defined anisotropy, and is split into multiple Dijkstra cells with edges joining nodes on the cell sides. The shortest path follows these edges, and can bend within the LVA cell.

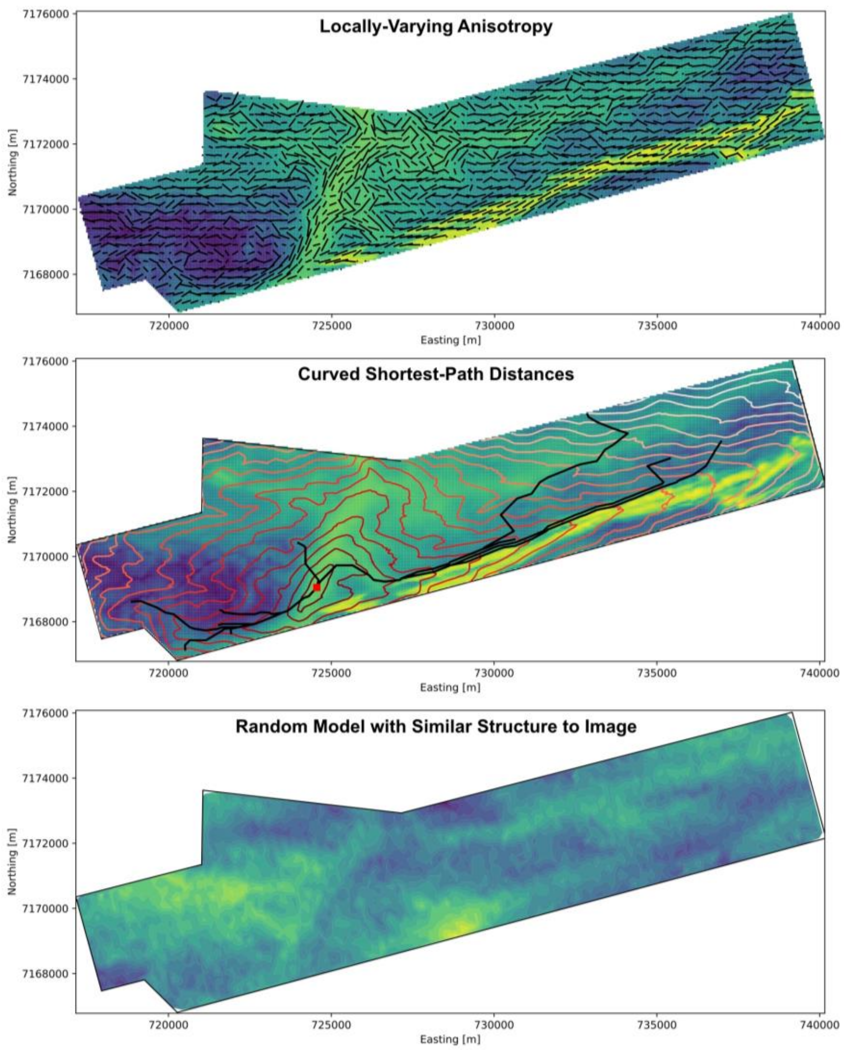

Figure 4. Deriving a spatially-structured covariance matrix from an image. (top) An AEM image with locally-varying anisotropy (black lines) superimposed. The line direction and length represent the direction of the major axis and the degree of anisotropy. (middle) The resulting distance function, represented as distance contours from a selected point, along with shortest-path curves for a few chosen points. (bottom) An example random image drawn from the resulting prior probability distribution function.

anisotropy. (This image is the time-constant of a best-fit single exponential decay to each sounding of an AEM survey. (data courtesy Sandfire Resources Ltd.))

For computing a curved shortest path, we use a Dijkstra algorithm (Dijkstra, 1959) as implemented in SciPy (Jones et al., 2001). We split each LVA cell into several Dijkstra cells with nodes along the boundaries, and edges between the nodes in one cell. The Dijkstra algorithm computes the shortest path between two end points, utilising the edges between the Dijkstra nodes (see Figure 3). The length of any edge is defined by the quadratic metric,

$$D\left(\mathbf{r}_1, \mathbf{r}_2\right) = \left(\mathbf{r}_1, \mathbf{r}_2\right)^T \mathbf{R}^T \mathbf{\Lambda}^{-1} \mathbf{R} \left(\mathbf{r}_1, \mathbf{r}_2\right),$$

where $\mathbf{\Lambda}$ contains the semi-major and semi-minor axis lengths on the diagonal, and $\mathbf{R}$ is a rotation matrix, specifying the direction of anisotropy. The middle panel in Figure 4 shows an example of the result, as the contoured distance from a given point, along with curved shortest paths to a few selected points for illustration. The rather jagged appearance of these paths is caused by the initial local anisotropy determination, which could be improved.

Finally, the prior covariance matrix is constructed using either the Gaussian or exponential formula above. A random image can be simulated by pre-multiplying an image drawn from an isotropic, unit-variance probability function by the Cholesky decomposition of this covariance matrix. An example image is shown in the bottom panel of Figure 4, and it can be seen that, as desired, it shows similar spatial structure to the original image.

CONCLUSIONS

We have developed a method for generating a prior probability distribution which produces samples that have similar spatial structure to the structure in a given image. This was done by splitting the image into cells, and computing a quadratic measure of the locally-varying anisotropy within each cell. This quadratic function was used as a distance metric, measuring distances depending on the angle as well as separation of two points inside the cell. Combining all the cells, a Dijkstra algorithm was used to compute a curved-path distance function between any two points in the image, whose shortest distance path followed the image's spatial structure. Finally, this distance measure was used to define the covariance matrix of a 2D Gaussian prior probability distribution function. Models drawn from this probability distribution have spatial features similar to the ones in the image.

The most important remaining work is to improve the local estimates of anisotropy.

ACKNOWLEDGMENTS

This work was supported by the Science and Industry Endowment Fund. We would like to thank Sandfire Resources Ltd. for giving us access to the AEM dataset used as an example image in this study.

REFERENCES

BOISVERT, J. 2010. Geostatistics with locally varying anisotropy. Ph.D. Thesis, University of Alberta

PBOISVERT, J.B., AND DEUTSCH, C.V. 2011. Programs for kriging and sequential Gaussian simulation with locally varying anisotropy using non-Euclidean distances. *Computers and Geosciences* **37**, 495–510.

DIJKSTRA, E.W. 1959. A note on two problems in connexion with graphs. *Numerische Mathematik* **1**, 269–271.

GALLARDO, L.A., MEJU, M.A., AND PÈREZ-FLORES, M.A. 2005. A Quadratic Programming Approach for Joint Image Reconstruction: Mathematical and Geophysical Examples. *Inverse Problems* **21**, 435–452–.

JONES, E., OLIPHANT, E., AND PETERSON, P. 2001. SciPy: Open Source Scientific Tools for Python. *Www.Scipy.org*.

KITANIDIS, P.K. 2010. Bayesian and Geostatistical Approaches to Inverse Problems. In Large-Scale Inverse Problems and Quantification of Uncertainty, (John Wiley & Sons, Ltd), pp. 71–85.

MOSEGAARD, K., AND TARANTOLA, A. 2002. Probabilistic approach to inverse problems. *International Geophysics Series*.

OH, S.-H., AND KWON, B.-D. 2014. Geostatistical approach to bayesian inversion of geophysical data: Markov chain Monte Carlo method. *Earth Planets and Space* **53**, 777–791.

OLIVER, D.S., REYNOLDS, A.C., AND LIU, N. 2008. Inverse Theory for Petroleum Reservoir Characterization and History Matching (Cambridge University Press).