Linear inverse Gaussian theory and geostatistics – a tomography example

Thomas Mejer Hansen[1], Andre Journel[2], Albert Tarantola[3] and Klaus Mosegaard[1]

[1]Niels Bohr Institute, University of Copenhagen, Juliane Mariesvej 30,
2100 København Ø, E-mail: **tmh@gfy.ku.dk**
[2]Department of Petroleum Engineering, Stanford University, Stanford, CA 94305, USA
[3]IPGP, 4 place Jussieu, 75005 Paris, France

1. Abstract

Though inverse theory and geostatistics share the same goal: to estimate some parameters of a model space from (possibly indirect) observations, they differ significantly in their approach. When applying inverse theory one typically has some observations and some physical relation that maps model parameters into these observations. When applying geostatistics one typically has some direct observations, and an a priori covariance model to specify the spatial relations in the model space. In short, inverse theory focuses on the physical relation between model and data space, whereas geostatistics focuses on the spatial relation of model space. Here we combine linear inverse Gaussian theory and geostatistics, using sequential simulation to generate samples of the posterior probability distribution of any linear inverse Gaussian problem. In inverse theory the use of sequential simulation is new, and allows efficient simulation of models with relatively complex a priori information, through the a priori covariance model. In geostatistics, this is our first step into incorporating physical theory directly into the simulation process. We show an example applying sequential simulation to a tomography problem; a typical linear inverse problem.

2. Introduction

Consider the expression

$$\mathbf{d} = \mathbf{G}\,\mathbf{m}$$

where $\mathbf{G}$ is some forward mapping operator (linear or nonlinear) that maps the model parameters $\mathbf{m}$ into observations $\mathbf{d}$. Estimation of $\mathbf{m}$ using knowledge about the forward mapping operator $\mathbf{G}$ and the observed data $\mathbf{d}$ is referred to as solving the inverse problem.

Linear least squares based inversion, where the mapping operator $\mathbf{G}$ is linear, generally provide a smooth estimate of the properties being estimated. Monte Carlo methods can be applied to such linear (and nonlinear) problems, allowing complex a priori information to be included, Mosegaard and Tarantola (1995). For example, a covariance model can be used to describe the a priori spatial correlation of the model space, hereafter refereed to as $\mathbf{C_M}$. The use of standard Metropolis Sampler to generate models that honor data observations and the a priori information is, however, extremely inefficient. Such algorithms are asymptotic, `short memory' samplers, which do not take advantage of the linear and Gaussian nature of the problem.

Kriging is a geostatistical technique for interpolation of observed data values in space, given a covariance model specifying the spatial correlation of data, $\mathbf{C_M}$. The most simple form of kriging, a so-called simple kriging system, is in fact identical to a linear Gaussian inverse problem, as the one in Eq. (1) with direct observations of the model space. Interpolation based on kriging estimation, since it is least squares based, produce maps of $\mathbf{m}$ that are too smooth, as compared to the a priori assumed covariance model $\mathbf{C_M}$. See for example Journel and Huijbregts (1978) for a detailed description of the Kriging approach.

Geostaticians use the concept of 'sequential simulation' to reduce the kriging smoothing effect by simulating a number of realisations of the posterior probability function, $\mathbf{m}_{est}(\mathbf{x})$, honouring both the a priori information, in form of the covariance function $\mathbf{C_M}$, and the observed values $\mathbf{d}_{obs}(\mathbf{x})$. See for example Goovaerts (1997) for an introduction to sequential simulation.

Here follows methodology which applies the concept of sequential simulation to any linear Gaussian inverse problem with prior information given by a covariance function, describing the spatial correlation of the model space.

3. Sequential simulation and linear Gaussian inverse theory

Consider two types of observations, type A, $\mathbf{a}_0$: direct measurements of the model space, and type B, $\mathbf{b}_0$: linear averages of the model space. $\mathbf{C_M}$ is the a priori covariance model. The linear relation between these observations and the mode space, $\mathbf{m}$, is given by the following (similar) linear systems of equations :

$$\mathbf{a}_0 = \mathbf{A}\,\mathbf{m} \quad , \qquad \mathbf{b}_0 = \mathbf{B}\,\mathbf{m} \tag{1}$$

Data of type A is typical in geostatistics. $\mathbf{A}$ is but a simple linear transfer function that maps observed data directly into a model space parameter. Such data are not typical in geophysical applied inverse theory, as the operator is not linked to any physical law. On the other hand, data of type B is very typical in geophysical applied inverse theory, but seldom used directly in geostatistics.

Assume now both data of types A and B are available as $\mathbf{a}_0$ and $\mathbf{b}_0$:

$$\mathbf{d}_0 = \begin{bmatrix} \mathbf{a}_0 \\ \mathbf{b}_0 \end{bmatrix}, \qquad \mathbf{C_D} = \begin{bmatrix} \mathbf{C_{aa}} & \mathbf{C_{ab}} \\ \mathbf{C_{ba}} & \mathbf{C_{bb}} \end{bmatrix}, \qquad \mathbf{G} = \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix} \tag{2}$$

where $\mathbf{C_{aa}}$ and $\mathbf{C_{bb}}$ are data covariances for the observed data $\mathbf{a}_0$ and $\mathbf{b}_0$, and $\mathbf{C_{ab}}$ is the cross data covariance between the two data types. The observations are linked to the model through the linear operator $\mathbf{G}$:

$$\mathbf{d}_0 = \mathbf{G}\,\mathbf{m} \tag{3}$$

The least squares solution to Eq. 3 then become a Gaussian probability density function with mean, $\tilde{\mathbf{m}}$:

$$\tilde{\mathbf{m}} = \mathbf{m}_0 + \mathbf{C_M}\left[\mathbf{A^t}\ \mathbf{B^t}\right]\mathbf{T}\left(\begin{bmatrix}\mathbf{a}_0\\ \mathbf{b}_0\end{bmatrix} - \begin{bmatrix}\mathbf{A}\\ \mathbf{B}\end{bmatrix}\mathbf{m}_0\right) \tag{4}$$

and with covariance, $\tilde{\mathbf{C}}_\mathbf{M}$:

$$\tilde{\mathbf{C}}_\mathbf{M} = \mathbf{C_M} - \mathbf{C_M}\left[\mathbf{A^t}\ \mathbf{B^t}\right]\begin{bmatrix}\mathbf{T_{aa}} & \mathbf{T_{ab}}\\ \mathbf{T_{ba}} & \mathbf{T_{bb}}\end{bmatrix}\begin{bmatrix}\mathbf{A}\\ \mathbf{B}\end{bmatrix}\mathbf{C_M} \tag{5}$$

where

$$\mathbf{S} = \begin{bmatrix}\mathbf{S_{aa}} & \mathbf{S_{ab}}\\ \mathbf{S_{ba}} & \mathbf{S_{bb}}\end{bmatrix} = \begin{bmatrix}\mathbf{C_{aa}} & \mathbf{C_{ab}}\\ \mathbf{C_{ba}} & \mathbf{C_{bb}}\end{bmatrix} + \begin{bmatrix}\mathbf{A}\\ \mathbf{B}\end{bmatrix}\mathbf{C_M}\left[\mathbf{A^t}\ \mathbf{B^t}\right] \tag{6}$$

and

$$\mathbf{T} = \begin{bmatrix}\mathbf{T_{aa}} & \mathbf{T_{ab}}\\ \mathbf{T_{ba}} & \mathbf{T_{bb}}\end{bmatrix} = \begin{bmatrix}\mathbf{S_{aa}} & \mathbf{S_{ab}}\\ \mathbf{S_{ba}} & \mathbf{S_{bb}}\end{bmatrix}^{-1} = \mathbf{S}^{-1} \tag{7}$$

Eq. (4) does not give an actual solution to Eq. (3), i.e. it is not a sample of the posterior probability density, but is simply the pointwise maximum likelihood point of the posterior probability density, sometimes referred to as the 'minimum error-variance estimate'. This is the conventional over smooth least squares result. For data of type A, geostaticians use sequential simulation to generate actual, unsmoothed, samples of the posterior probability density function. In a similar manner we make use of sequential simulation to generate actual samples of the posterior probability density.

To do this we need to be able to compute the local posterior probability density function at any one location, conditional to observed data of types A and B. This can be done simply by solving Eq. (4)-(5), for only one point in space, $\mathbf{x_i}$. This mean and covariance of the local posterior Gaussian probability density at location $\mathbf{x_i}$, conditioned to all observed data of types A and B, is referred to as $\hat{\mathbf{m}}(\mathbf{x_i})$ and $\hat{\mathbf{C}}_\mathbf{M}(\mathbf{x_i})$. To obtain samples of the posterior probability density function (realisations) we use the concept of sequential simulation:

1. Randomly visit a point in the model space, say $\mathbf{x_i}$.
2. Compute the posterior probability distribution, given by $\hat{\mathbf{m}}(\mathbf{x_i})$ and $\hat{\mathbf{C}}_\mathbf{M}(\mathbf{x_i})$, conditional to the known and previously simulated nodes ($\mathbf{a}_0$, data of type A) and observed linear averages ($\mathbf{b}_0$, data of type B).
3. Draw a (random) value from the found Gaussian conditional probability density function, say $\mathbf{m}_{\mathrm{draw}}(\mathbf{x}_\mathrm{i})$. This is a sample of the posterior Gaussian probability function.
4. Add $\mathbf{m}_{\mathrm{draw}}(\mathbf{x}_\mathrm{i})$ to the data of type A : $\mathbf{a}_{0,\mathrm{i+1}} = [\mathbf{a}_{0,\mathrm{i}}, \dots, \mathbf{m}_{\mathrm{draw}}(\mathbf{x}_\mathrm{i})]$.
5. Go to '1'. Continue until a desired part of the model space has been visited.

The collection of samples of the posterior Gaussian probability function is one realisation of the posterior Gaussian random field. Additional independent realisations of the posterior random Gaussian field can be obtained simply by rerunning the algorithm with a different random seed, following a new random path.

4. A synthetic tomography example

As an example of an application of the developed theory consider the following synthetic cross borehole tomography setup. A reference model grid of 60 (horizontal) by 80 (vertical) samples, with 250 meter between grid points, i.e. a model of 1500x2000 meters, is chosen. Using SGSIM from the GSLIB software (Deutsch and Journel, 1996), one realisation of a Gaussian random field is generated with a global mean and variance of V_{mean} = 5 km/s, σ=0.1 km/s and a spherical covariance model with range of 400m. This model is selected as a reference velocity field, see Fig. 1.

Fig. 2 show the ray coverage for three different cases resembling a cross borehole tomography setup. 2, 5 and 8 sources and receivers are located in the left- and right-most part of the model. All receivers measure the travel time delay observed along rays from each source which is transformed into an observed mean velocity along ray paths. This observed mean velocity along ray paths is our observed linear average data, $\mathbf{b}_0$. Thus, we consider three cases with 4, 25 and 64 rays respectively spanning the model space. We also consider the case where no rays are available. We refer to these cases as cases with 0, 4, 25 and 64 rays. For all cases the velocity in the left and right most column of the model is considered known, $\mathbf{a}_{0,}$ (indicated by circles in Fig. 2) as could be the case in for example a bore hole with a well log.. Thus both data of types A and B are available, $\mathbf{d}_0 = [\mathbf{a}_0 ; \mathbf{b}_0]$. $\mathbf{C_M}$ is assumed known, as the reference semivariogram model. $\mathbf{G}$ is obtained through simple ray tracing. Data are assumed noise free, $\mathbf{C_d}$=0. We now have a linear inverse problem as given by Eq. (3).

Figs. 3 and 4 show two independent realisations of the posterior probability density for 0, 4, 25 and 64 rays respectively. Comparing to Fig. 1a, it is clear that increasing the ray coverage results in realisations with more resemblance to the reference field.

Fig. 5a shows the estimated semivariograms from realisations of the 4 cases. It is apparent that the spatial variability in form of the semivariogram is quite well reproduced for all cases. The histogram of the generated realisations, Fig. 5b, also match the reference histogram, and the observed mean velocity along ray paths, $\mathbf{d}_{0,est}$, match the true mean velocities along the ray paths, Fig. 5c.

5. Conclusions

Linear inverse theory and geostatistics are closely related research fields, and we combine the sequential approach from geostatistics and data conditioning from inverse theory, bringing new theory and application to both research fields. The presented theory is closely related to the geostatistical method of sequential Gaussian simulation (SGS), and the presented theory can be seen as an extension to SGS to condition to linear averages, as well as direct observations, of the model space. The major application is non-iterative sequential 'simulation' that will generate actual samples from the posterior probability density function, consistent with data observations and a priori information in form of a prior covariance model describing the spatial variability. Compared to Monte Carlo based approaches, that can be used to generate samples of the posterior probability density function, the presented approach is non-iterative, and hence computationally very efficient.

6. References

Deutsch, C. V., Journel, A. G., 1997: GSLIB, Geostatistical software library and user's guide. Applied Geostatistics. Oxford University Press, 2nd edition.

Goovaerts, P., 1997: Geostatistics for natural resources evaluation. Applied Geostatistics. Oxford University Press.

Journel, A. G., Huijbregts, Ch. J., 1978: Mining geostatistics. Academic Press.

Mosegaard, K., Tarantola, A., 1995: Monte Carlo sampling of solutions to inverse problems, J. Geophys. Res., 100 (B7), 12431-12447.

Tarantola, A., 2005: Inverse problem theory, SIAM.

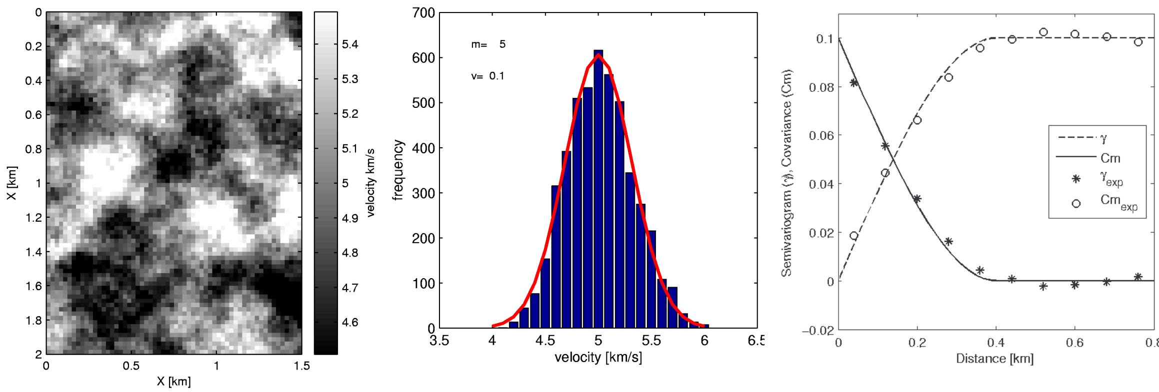

Fig. 1. Left: Reference model. Middle: Histogram of reference model. Right: A priori assumed semivariogram model (solid line) compared to the semivariogram calculated from the generated reference velocity field (open circles). The covariance is shown for comparison.

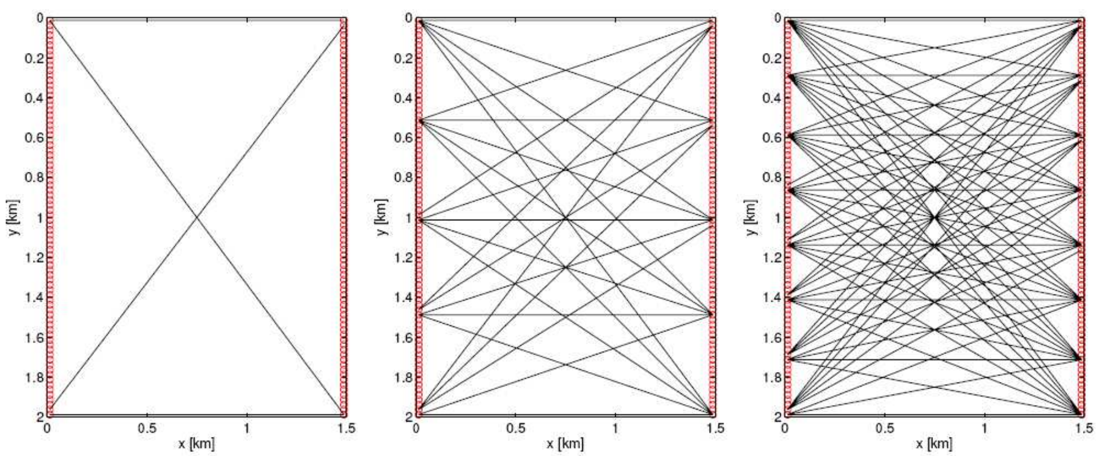

Fig. 2. Ray coverage for the synthetic tomography setup. Left: 4 rays, Middle: 25 Rays, Right: 64 Rays.

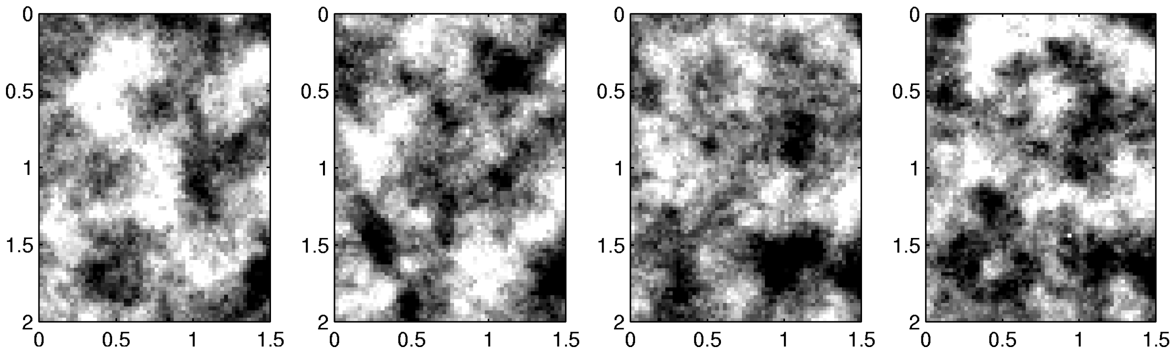

Fig. 3. Conditional simulation using 0, 4, 25 and 64 rays (from left to right). Color scale as in Fig. 1a.

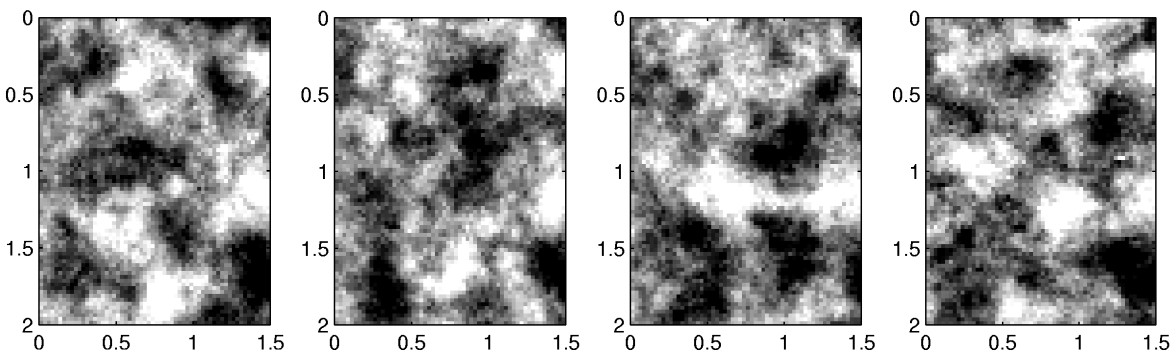

Fig. 4. Another (independent) conditional simulation using 0, 4, 25 and 64 rays (from left to right). Color scale as in Fig. 1a.

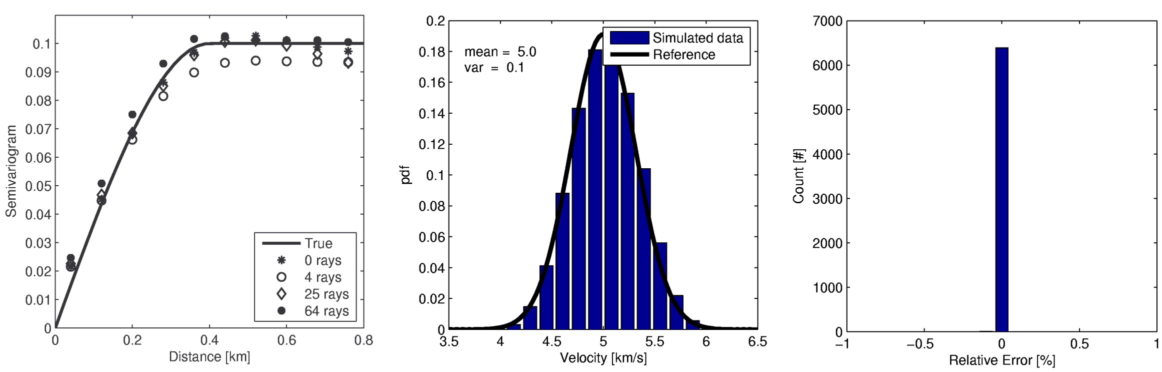

Fig. 5. Left: Estimated semivariogram using 0, 4, 25 and 64 rays respectively. Middle: Histogram of the simulated data compared to the reference probability density function. Right: Relative error of the estimated $\mathbf{d}_{0,\mathbf{est}}$. Plots generated using 100 realisations.