Geometric Sampling: An Approach to Uncertainty in High Dimensional Spaces

Juan Luis Fernández-Martínez, Michael Tompkins,
Tapan Mukerji, and David Alumbaugh

Abstract. Uncertainty is always present in inverse problems. The main reasons for that are noise in data and measurement error, solution non-uniqueness, data coverage and bandwidth limitations, physical assumptions and numerical approximations. In the context of nonlinear inversion, the uncertainty problem is that of quantifying the variability in the model space supported by prior information and the observed data. In this paper we outline a general nonlinear inverse uncertainty estimation method that allows for the comprehensive search of model posterior space while maintaining computational efficiencies similar to deterministic inversions. Integral to this method is the combination of model reduction techniques, a constrained mapping approach and a sparse sampling scheme. This approach allows for uncertainty quantification in inverse problems in high dimensional spaces and very costly forward evaluations. We show some results in non linear geophysical inversion (electromagnetic data).

Keywords: Inverse Problems, Geometric sampling, High Dimensional Spaces, Uncertainty.

Juan Luis Fernández-Martínez and Tapan Mukerji
Energy Resources Department, Stanford University, Palo Alto, California, USA

Juan Luis Fernández-Martínez
Department of Mathematics, University of Oviedo, Oviedo, Spain

Juan Luis Fernández-Martínez
Department of Civil and Environmental Engineering, University of California Berkeley, Berkeley, USA

Michael Tompkins and David Alumbaugh
Schlumberger-EMI Technology Center, Richmond, CA 94804, USA
e-mail: jlfm@uniovi.es,mtompkins@richmond.oilfield.slb.com,
mukerji@stanford.edu,dalumbaugh@richmond.oilfield.slb.com

C. Borgelt et al. (Eds.): Combining Soft Computing & Stats. Methods, AISC 77, pp. 247–254.
springerlink.com

1 Inverse Problems and Uncertainty

Inverse problems can be written in discrete form as $\mathbf{F}(\mathbf{m}) = \mathbf{d}$, where $\mathbf{m} \in \mathbf{M} \subset \mathbf{R}^n$ are the model parameters, $\mathbf{d} \in \mathbf{R}^s$ the discrete observed data, and

$$\mathbf{F}(\mathbf{m}) = (f_1(\mathbf{m}), f_2(\mathbf{m}), \ldots, f_s(\mathbf{m}))$$

is the vector field representing the forward operator and $f_j(\mathbf{m})$ is the scalar field that accounts for the j-th data. Usually $s < n$, that is, the inverse problem has an underdetermined character. Furthermore, many geophysical problems are nonlinear and poorly sampled making the inverse ill-posed, non-unique, and ill-conditioned. Ill-conditioning is an important issue when solving the inverse problem as an optimization problem, because noise in data is amplified back to the model parameters through the inverse forward operator, $\mathbf{F}^{-1}$. In addition to these difficulties, we have measurement errors, data coverage and bandwidth limitations, and numerical approximations, which all contribute to uncertainty in our inverse solutions. In the context of nonlinear inversion, the uncertainty problem is that of quantifying the variability in the model space supported by prior information, the observed data, and the errors of the method.

Global optimization algorithms can be a good alternative to deterministic solutions, because they approach the nonlinear inverse problem as a sampling problem instead of looking for the inverse operator. Also, they only need as prior information the search space of possible solutions. Typically they use as a cost (or objective) function the data prediction misfit in a certain norm p:

$$\|\mathbf{F}(\mathbf{m}) - \mathbf{d}\|_p .$$

It is possible to show analytically that the models in the neighborhood of $\mathbf{m}_0$ that fit the data within the same tolerance, tol, belongs to the following hyperquadric:

$$(\mathbf{m} - \mathbf{m}_0)^T \mathbf{J}\mathbf{F}_{\mathbf{m}_0}^T \mathbf{J}\mathbf{F}_{\mathbf{m}_0} (\mathbf{m} - \mathbf{m}_0) + 2\boldsymbol{\Delta}\mathbf{d}^T (\mathbf{m} - \mathbf{m}_0) + \|\boldsymbol{\Delta}\mathbf{d}\|_2^2 = tol^2.$$

$\mathbf{J}\mathbf{F}_{\mathbf{m}_0}$ is the Jacobian matrix of the operator $\mathbf{F}$ in $\mathbf{m}_0$ and $\boldsymbol{\Delta}\mathbf{d} = \mathbf{F}(\mathbf{m}_0) - \mathbf{d}$. This means that the region of equivalent models locally in $\mathbf{m}_0$ have the direction of the vectors of the $\mathbf{V}$ base given by the singular value decomposition of $\mathbf{J}\mathbf{F}_{\mathbf{m}_0}$ and whose axes are proportional to the inverse of the singular values λ_k in each direction. Due to the continuity of the Jacobian operator, we finally conclude that with no regularization term the misfit function has a flat and elongated valley shape. This approach assumes derivability of $\mathbf{F}$ in $\mathbf{m}_0$, which is usually the case in most inverse problems.

These types of sampling methods can be useful, but they have limitations for large spaces, since they sample each model parameter as independent variables; a property that is not necessarily true for finite resolution methods (i.e, electromagnetic imaging). In contrast, local optimization methods are

not designed to approach this sampling problem, and they often fail to find a solution without regularization. These algorithms can very effectively handle inverse problems having thousands to millions of parameters. The main drawback of local methods is that they are highly dependent on the initial guess and the quality of the prior information that is built into the regularization term to achieve uniqueness and stability in the inverse solution. Furthermore they do not provide any measure of nonlinear uncertainty around the solution of the inverse problem.

An alternative to both these methods, presented by [7], is to solve the sampling problem in a bounded transformed space using optimally sparse grids. This method both accounts for the equivalence in our nonlinear inverse problem and allows for the inference of solution uncertainty by sampling the model posterior. In the following sections we review two important aspects of this technique including model parameter reduction using orthogonal transformations, and geometric sampling of the equivalent model space.

2 Model Reduction Techniques

The use of model reduction techniques act to decrease the dimension of the inverse problem. For an underdetermined linear inverse problem of the form

$$\mathbf{Gm} = \mathbf{d}$$

the method consists in expanding the solution $\mathbf{m}$ as a linear combination of a set of independent models, $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_q\}$:

$$\mathbf{m} \in \langle \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_q \rangle = \sum_{k=1}^{q} \alpha_k \mathbf{v}_k, \tag{1}$$

and to solve the linear system $\mathbf{B}\alpha = \mathbf{d}$ where $\mathbf{B} = \mathbf{GV}$ and $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_q]$. This methodology is related to subspace methods and can be easily generalized to nonlinear inverse problems, because once the base is determined, the search is performed on the α-space. The use of a reduced set of basis vectors that are consistent with our prior knowledge allows to regularize the inverse problem and to reduce the space of possible solutions. There are several ways of finding the base $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_q\}$ in order to reduce complexity. Principal Component Analysis (PCA), the Singular Value Decomposition, the Discrete Cosine Transform (DCT) and the Discrete Wavelet Transform have been presented in [2]. As stated in [2] the orthonormal base has to allow for classification of the model variability, and to be separable in order to expand this methodology to large parameterizations. Also one of these methods is covariance-based (PCA) while the other techniques are model-based, allowing for the use of different kinds of reduction techniques depending on the dimension of our inverse problem and the quality of the prior information.

3 Computing Uncertainty in High Dimensional Spaces

The methods to compute uncertainty in high dimensional spaces can be divided into two main groups depending on whether the forward problem is fast to solve or not. To the first category belong global optimization algorithms in a reduced model space [3]. Monte Carlo techniques can not address this kind of problem due to the dimensionality issue. Additionally they can be very inefficient, since they typically spend much effort sampling parts of the posterior that do not fit the observed data. Global optimization algorithms can address the non-convexity of the cost function by sampling the family of equivalent models. Nevertheless when the inverse problem has a very expensive forward problem these methods are not a good alternative. The main reason is that the tasks of sampling the posterior and the forward prediction are coupled.

3.1 The Geometric Sampling Approach

An alternative to stochastic sampling methods is geometric sampling as introduced by Tompkins and Fernández Martínez [7]. The methodology is composed of four steps: 1) parameter reduction 2) model constraint mapping by vertex enumeration, 3) sparse grid sampling, and 4) final forward evaluation. Model reduction techniques have been already presented, and we discuss sampling below.

Once we perform model reduction techniques on the model space we have at our disposal a set of orthonormal vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_q\}$ allowing the linear decomposition (1). Vectors $\mathbf{v}_i$ increase their frequency with their index. Thus this linear expansion has a regularization effect on the sampling. The next step is to provide to the model $\mathbf{m}$ some lower and upper bounds, $\mathbf{l}$ and $\mathbf{u}$:

$$\mathbf{l} \leq \sum_{k=1}^{q} \alpha_k \mathbf{v}_k \leq \mathbf{u}. \tag{2}$$

At this stage we suppose a uniform prior distribution of $\mathbf{m}$ in these bounds. Condition (2) is called the vertex enumeration problem [1] and generates in the reduced model space a polytope $P \subset \mathbb{R}^p$ for the coefficients on the reduced base α_k. This idea was first suggested by Ganapathysubramanian and Zabaras [5] for heat flow problems in random media using the PCA base. Nevertheless, in their work this idea was not used to compute uncertainty associated with to non-linear inverse problems through the geometric sampling approach. Geometric Sampling consists in sampling within the polytope using an uniform distribution and sparse grids [7]. This uniform sampling distribution induces a non-uniform prior distribution in the original space $\mathbf{M}$, since:

$$P(m_i < c_i) = P\left(\sum_{k=1}^{q} \alpha_k \mathbf{v}_{ki} < c_i\right),$$

where the vectors of the base are fixed.

This complete non-uniform distribution on **M**, the posterior, is not explicitly calculated, since we only have at our disposal some independent samples whose number depend on the sparsity of the sampling scheme that has been used. However, once we have sampled α_k over our reduced space, our approximation to this posterior is determined by mapping these samples back to our original model space, **M**.

Optimizing the sampling on the reduced space is very important in our methodology, since this allows us to tailor the sampling density to the cost and to the complexity of the forward evaluations. Of particular interest are sparse sampling techniques, such as the Smolyak grids [6], that can provide for adaptive sample refinement. That is, if we wish to extend the accuracy or breadth of our sampling from some initial set to some larger set, we simply need to evaluate samples at the additional nodes over the second set, which provides a means to optimize sampling based on some criteria (in our case, convergence of statistical moments of our posterior).

In order to sample inside the polytope, defined by Equation (2), we circumscribe a hypercube in our $q-$dimensional space and perform the sampling on a Cartesian grid in the reduced model space using Smolyak nested grids [8]. Once the sampling on the hypercube is performed, models, **M**, in our original model space are reconstructed using (1). The final step in uncertainty estimation is to evaluate the posterior model samples for their likelihood (i.e., data misfit). For this, forward simulations are performed and models are accepted or rejected based on a threshold misfit. The accepted models represent the equivalence space of equiprobable models. The uncertainty of the nonlinear inverse problem then follows from either the model ensemble itself or statistical measures (e.g., mean, covariance, percentile, interquartile range) computed from it.

4 Example: A Subsurface Resistivity Image

Over the past decade, marine controlled-source electromagnetic (CSEM) surveying has emerged as a useful technique for subsurface resistivity imaging. In this method, a deep-towed electric dipole source is used to excite a low-frequency (˜0.1–10 Hz) electromagnetic signal. This signal propagates through the seawater and subsurface and is perturbed by geologic variation to depths of several kilometers. Spatially-distributed, multi-component, seafloor receivers record this electromagnetic energy at offsets up to 20km. These electromagnetic data are typically interpreted using geophysical inverse methods that attempt to reconstruct subsurface resistivities from recorded fields [4]. Inversion of this data is nonlinear, and uncertainty comes from noise in the data and assumptions about the earth model.

To demonstrate the extension of our uncertainty method to large parameter spaces, we chose a marine CSEM field dataset collected by WesternGeco in the Potiguar basin in Brazil during 2009. The electromagnetic data consisted

of ∼3800 complex-valued fields at 4 frequencies (0.25, 0.50, 0.75, and 1.5Hz). For this problem, the original uniform pixel space had 33,280 parameters; however, we only considered the part of the final inverse model not occupied by air, seawater, or homogeneous resistive basement (>4500m depths). This left the inversion domain shown in Figure 1(A), which consisted of 5,461 parameters. Of particular interest, are the structures and magnitudes of the resistive features at ∼4000m depth. After subtracting a global mean from our inverse model, the SVD base was reduced to six terms that represented ∼ 81% of the variability in the residuals, following the methodology explained in [2]. We then performed optimal sampling over this six-dimensional reduced space to estimate the model posterior and solution uncertainty. The resulting posterior polytope was defined by 15,990 vertices, which, in this case, we approximated with a 6-D hypercube (i.e., 6 bases were chosen). Based on sparse sampling of this hypercube we evaluated 1942 equi-feasible models to generate the model posterior. The computational cost of evaluating these models was approximately 6 days using two 8-core workstations. With a threshold

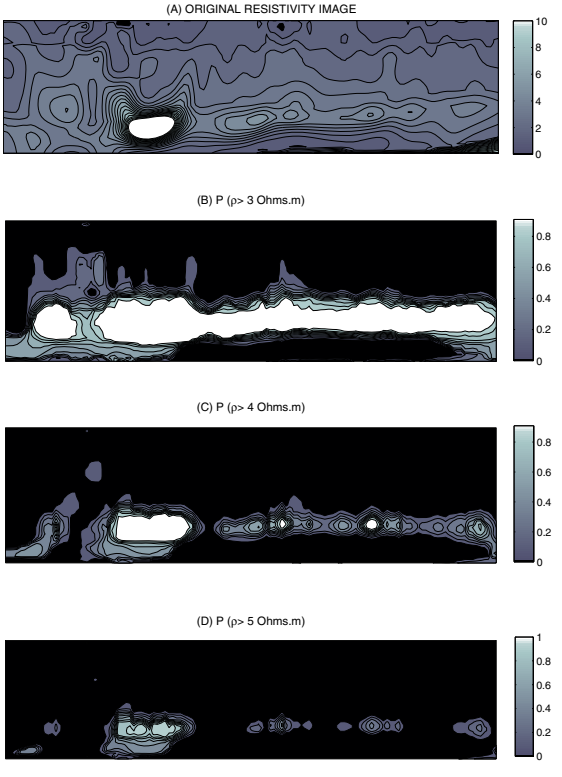

Fig. 1 Original resistivity field and probability maps for different cut-offs (3, 4 and $5\,\Omega.m$).

misfit, RMS<15%, we generated the final equivalent model set (283 models). Once we have this ensemble, we can compute statistical properties from it as well, for example, e-types, variances, or indicator probabilities. Probability (normalized frequency) maps are a useful way to visualize uncertainty. Figures 1(B)–(D) show the probability maps for different cut-offs (3, 4 and 5 Ω.m) deduced from the approximation of the model posterior over the low misfit region. These probability maps represent the probability of occurrence of a resistivity of at least 3, 4 and 5 Ω.m in our model space, and quantifies some aspects of the uncertainty in our subsurface resistivity image inverse problem. Additional measures of uncertainty are possible using any number of statistical properties of the model posterior.

5 Conclusions

The combined use of model reduction techniques, and sparse sampling allows us to approach efficiently the uncertainty problem in high dimensional spaces. This methodology can be efficiently applied to estimate nonlinear inverse model uncertainty in any kind of inverse problem. The combination of these methods can reduce the nonlinear uncertainty problem to a deterministic sampling problem in only a few dimensions, requiring only limited forward solves, and resulting in an optimally sparse representation of the posterior model space. While forward solves are required to evaluate the sampled models, our scheme optimizes sample size by iteratively increasing sampling complexity until uncertainty measures converge or a maximum number of forward solves is completed.

Acknowledgements. We acknowledge WesternGeco for allowing us to use and publish the resistivity image data set.

References

1. Avis, D., Fukuda, K.: A pivoting algorithm for convex bulls and vertex enumeration of arrangements and polyhedra. J. Discrete Comp. Geometry 8, 295–313 (1992)
2. Fernández-Martínez, J.L., Tompkins, M., Fernández-Muñiz, Z., Mukerji, T.: Inverse problems and model reduction techniques. In: Borgelt, C., González-Rodríguez, G., Trutschnig, W., Lubiano, M.A., Gil, M.A., Grzegorzewski, P., Hryniewicz, O. (eds.) Combining Soft Computing and Statistical Methods in Data Analysis. Advances in Soft Computing. Springer, Berlin (in this book, 2010)
3. Fernández-Martínez, J.L., Mukerji, T., García-Gonzalo, E.: Particle Swarm Optimization in high dimensional spaces. In: Proceedings of the Seventh International Conference on Swarm Intelligence, ANTS 2010, Bruxelles, Belgium (2010)

4. MacGregor, L.M., Sinha, M.C., Constable, S.: Electrical resistivity structure of the Valu Fa Ridge, Lau Basin, from marine controlled-source electromagnetic sounding. Geophys. J. Int. 146, 217–236 (2001)
5. Ganapathysubramanian, B., Zabaras, N.: Modeling diffusion in random heterogeneous media: Data-driven models, stochastic collocation and the variational multiscale method. J. Comput. Phys. 226, 326–353 (2007)
6. Smolyak, S.: Quadrature and interpolation formulas for tensor products of certain classes of functions. Dokl. Math. 4, 240–243 (1963)
7. Tompkins, M.J., Fernández-Martínez, J.L.: Scalable Solutions for Nonlinear Inverse Uncertainty Using Model Reduction, Constraint Mapping, and Sparse Sampling. In: Proceedings of the 72nd EAGE Conference & Exhibition, Barcelona, Spain (2010)
8. Xiu, D., Hesthaven, J.S.: High-order collocation methods for differential equations with random inputs. SIAM J. Sci. Comput. 27, 1118–1139 (2005)