Note during discussion group meeting

First meeting: October 18, 2023

Paper: Li & Oldenburg, 1996, 3-D inversion of magnetic data

1. Why is the Lagrange parameter λ used as an inverse (λ^{-1}) in the global objective function?

The global objective function is: $\Phi(\mathbf{m}) = \|\mathrm{W_m}(\mathbf{m} - \mathbf{m}_0)\|_2^2 + \lambda^{-1}(\|\mathrm{W_d}(\mathrm{G}\mathbf{m} - \mathbf{d}^{obs})\|_2^2 - \varphi_d^*)$

Differentiating with respect to model parameters yields: $\left(\lambda^{-1}\mathrm{G}^T\mathrm{W}_d^T\mathrm{W_d}\mathrm{G} + \mathrm{W}_m^T\mathrm{W_m}\right)\mathbf{m} = \lambda^{-1}\mathrm{G}^T\mathrm{W}_d^T\mathrm{W_d}\mathbf{d}^{obs} + \mathrm{W}_m^T\mathrm{W_m}\mathbf{m}_0$

Now, when multiplied by λ, we obtain the following system of equations, which was used in the paper:

$$\left(\mathrm{G}^T\mathrm{W}_d^T\mathrm{W_d}\mathrm{G} + \lambda\mathrm{W}_m^T\mathrm{W_m}\right)\mathbf{m} = \mathrm{G}^T\mathrm{W}_d^T\mathrm{W_d}\mathbf{d}^{obs} + \lambda\,\mathrm{W}_m^T\mathrm{W_m}\mathbf{m}_0$$

On the other hand, this system of equations is equivalent to the case where we formulate the optimization problem as a general Tikhonov problem: $\Phi(\mathbf{m}) = \|\mathrm{W_d}(\mathrm{G}\mathbf{m} - \mathbf{d}^{obs})\|_2^2 + \lambda\,\|\mathrm{W_m}(\mathbf{m} - \mathbf{m}_0)\|_2^2$

This means that when they apply λ to the data misfit term, they use λ^{-1}, and when they apply it to the regularization term, they use λ.

2. Application of $\mathbf{m}$ versus $(\mathbf{m} - \mathbf{m}_0)$ in the smoothing terms:

$\emptyset_m = \alpha_s\|\mathrm{W_s}(\mathbf{m} - \mathbf{m}_0)\|_2^2 + \|\mathrm{W_v}\,\mathbf{m}\|_2^2,$ (1) $\qquad \mathrm{W}_v^T\mathrm{W}_v = \alpha_x\mathrm{W}_x^T\mathrm{W}_x + \alpha_y\mathrm{W}_y^T\mathrm{W}_y + \alpha_z\mathrm{W}_y^T\mathrm{W}_y$

Or

$\emptyset_m = \alpha_s\|\mathrm{W_s}(\mathbf{m} - \mathbf{m}_0)\|_2^2 + \|\mathrm{W_v}(\mathbf{m} - \mathbf{m}_0)\|_2^2,$ (2)

Both formulas are used in UBC codes.

In the first formulation, the gradient, in each direction, of the model parameters is panelized. Consequently, the boundary of the recovered model can have a smooth transition. In the second formulation, the differences in the gradient between constructed model and reference model is panelized (note the reference model is fixed during inversion). Then, the boundary characteristics of the reference model will be preserved in the recovered model. Typically, the reference model possesses sharp boundaries, and including it in the smoothing terms ensures the preservation of these sharp boundaries in the recovered model

3. How to balance smallness and smoothness terms:

Parameters α_s, α_x, α_y, and α_z, determine the relative importance of the components in the regularization term. In general, larger values of α_x, α_y, and α_z enhance the smoothing effect in the corresponding direction, whereas reducing parameter α_s results in an overall smoother model. The UBC website offers suggestions for determining these parameters. For example, to achieve equal contributions from all four components in the regularization term, one should set $\alpha_s = \frac{1}{h^2}$, where 'h' represents the cell size dimension. They also suggested the default values of these parameters in most inversion codes:

$$\left[\alpha_s = 1e-4\,, \quad \alpha_x = \alpha_y = \quad \alpha_z = 1\right]$$

Furthermore, to observe the effect of these parameters in a mixed L_p-norm inversion, I suggest looking at Figure 6 in the following paper (They used UBC SimPEG software, so the results should be extendable to our discussion in this paper):

Wei, X. and Sun, J., 2021, Uncertainty analysis of 3D potential-field deterministic inversion using mixed L_p norms, GEOPHYSICS, 86(6), G133-G158.

4. How to design the depth weighting in practice:

In Li & Oldenburg (1996), depth weighting inside the derivatives has been clearly used. This approach has also been adopted in many papers published by UBC. On page 401, after equation (20), they discussed the application of depth weighting outside the derivatives. They believe this approach can also yield reasonable results. In recent works by UBC, specifically on their website, they generally use depth weighting outside the derivative

I believe depth weighting should be applied outside the derivatives. The derivatives should be computed with respect to the model parameters themselves and not on transformed parameters. This is particularly significant for the depth direction.

5. Adjusting depth weighting for mixed L_p-norm regularization:

For this, I refer again to the paper by Wei & Sun (2021), Page G140, where they attempted to determine a range of optimal values for the parameter β in the depth weighting function $W(z) = \frac{1}{(z+z_0)^{\frac{\beta}{2}}}$ for a mixed L_p-norm inversion. Figures 8 and 9 in their paper show the recovered models for various β parameters, considering both L_{22}-norm and L_{12}-norm inversions. Generally, they observed that L_{12}-norm inversion exhibits greater robustness against variations in the beta parameter compared to L_{22}-norm inversion.

6. Utility of depth vs distance vs sensitivity weighting:

Distance weighting is a generalization of depth weighting. Unlike depth weighting, which emphasizes the sensitivity of the model on the depth below the surface, distance weighting takes into account the radial distance from the observation point to each cell, considering both depth and horizontal distance. Depth weighting is only used for surface or airborne data, while distance weighting can also be applied to down-hole measurements. It is particularly useful when dealing with highly irregular surfaces such as areas with extreme topography or rough airborne surveys. The sensitivity weighting is based on the sensitivity matrix, and it depends on the overall sensitivity of the entire dataset to a particular cell in the model (Li & Oldenburg, 2000). For surface data, both distance and sensitivity weightings yield results that are similar to those obtained through inversion using depth weighting. There is a comprehensive discussion about these methods in the following paper, on pages 543 and 544.

Li, Y. & Oldenburg, D. W., 2000, Joint inversion of surface and three-component borehole magnetic data, GEOPHYSICS, 65(2), 540-552.

7. Determining appropriate value of trade-off (regularization) parameter λ:

There are various approaches to determine the parameter λ, including the L-curve, GCV, MDP, and UPRE. In Oldenburg & Li (1994), they plotted the data misfit versus a range of parameter λ (in a one plot), as well as the regularization term versus that range of parameter λ (in another plot). The optimal parameter can be found approximately at the corner of these curves. It is important to note that their approach requires the introduction of three user-defined parameters and is not entirely automatic. It is indeed possible to combine these two plots into a single graph, where data misfit is plotted against the regularization term for a range of parameter λ. This results in an L-curve (or Tikhonov curve), and the location of maximum curvature in this plot gives the optimal regularization parameter. The advantage of this approach is that it can be automated, providing an automatic method of estimating the regularization parameter. However, it is worth noting that in some cases, the L-curve may not have the typical "L" shape, where it changes very smoothly. This can make it challenging to estimate the maximum curvature accurately, potentially leading to a large error in the estimated parameter.

There are indeed many papers addressing the estimation of the regularization parameter in inversion approaches. Specifically, for the minimum-structure inversion approach discussed in this meeting, there is a detailed discussion about L-curve and Generalized Cross-Validation (GCV) approaches in the following chapter:

Oldenburg, D., W. & Li, Y., 2005, "5. Inversion for Applied Geophysics: A Tutorial," Investigations in Geophysics: 89-150.

A comparison of L-curve, MDP, GCV, UPRE, and Chi^2 methods is provided for sparse inversion in

Vatankhah, S., Renaut, R. a. & Ardestani, V. E., 2014, Regularization parameter estimation for underdetermined problems by the chi^2 principle with application to 2D focusing gravity inversion, Inverse Problems, 30(8), 085002.

However, all these approaches are designed for single-data inversion. Estimating two or more regularization parameters in a joint inversion algorithm can be challenging.

8. The subspace approach and search directions used in this paper:

An efficient approach to solving moderate to large-scale inverse problems is to transform the original problem into a subspace, where a solution can be found. The solution can then be projected back into the original space. In this paper, the basis vectors for the subspace are generated using steepest descent vectors associated with the gradients of the data misfit and gradients of the model component of the objective function. The problem is then solved within this subspace, resulting in the solution (coefficients denoted as 'α'). Finally, this solution is projected back into the original space to compute the original model parameter '**m**'. Alternatively, someone can apply Golub-Kahan Bidiagonalization for the standard Tikhonov objective function or joint Golub-Kahan bidiagonalization for the general Tikhonov objective function to project the inverse problem into a smaller Krylov subspace. This is an efficient approach commonly used in mathematical and geophysical literature.