Notes on data space inversion

Peter & Saeed

Fourth meeting: November 15, 2023

Paper: Pilkington, 2009, 3D magnetic data-space inversion with sparseness constraints.

First, we summarize the data space approach as outlined by Tarantola (2005) and compare to Pilkinton (2009):

The objective function: $S(\mathbf{m}) = (A\boldsymbol{m} - \boldsymbol{d}^{obs})^{\mathrm{T}} C_D^{-1} (A\boldsymbol{m} - \boldsymbol{d}^{obs}) + (\mathbf{m} - \mathbf{m}_0)^{\mathrm{T}} C_M^{-1} (\mathbf{m} - \mathbf{m}_0)$ (1)

Here C_D^{-1} and C_M^{-1} are inverse of data and model covariance matrices, respectively. The sensitivity matrix *A* is of dimension $N \times M$ (*M* is the number of model parameters and *N* is the number of data points). In potential field voxel inversion $M \gg N$.

Minimization of eq. (1) with respect to model parameters yields:

$$\boldsymbol{m} = \boldsymbol{m}_0 + (A^T C_D^{-1} A + C_M^{-1})^{-1} A^T C_D^{-1} (\boldsymbol{d}^{obs} - A\mathbf{m}_0) \quad (2)$$

Which effectively requires the inverse of matrix $A^T A$ that is of dimension $M \times M$. Using

$$(A^T C_D^{-1} A + C_M^{-1})^{-1} A^T C_D^{-1} = C_M A^T (A C_M A^T + C_D)^{-1} \quad (3)$$

Using a matrix identity, eq. (2) can be re-written as

$$\boldsymbol{m} = \boldsymbol{m}_0 + C_M A^T (A C_M A^T + C_D)^{-1} (\boldsymbol{d}^{obs} - A\mathbf{m}_0) \quad (4)$$

where the inverse of the matrix AA^T, that is of dimension $N \times N$, is needed. Then, it is more efficient to use eq. (4) in a CG implementation as applied by Pilkington (2009).

In Pilkington (2009), the final solution, in data space, was written as follows (equation (10) in the paper, omitting subscript here):

$$\boldsymbol{m} = \boldsymbol{m}_0 + \alpha Q S A^T (A S Q S^T A^T + D)^{-1} (\boldsymbol{d}^{obs} - A\boldsymbol{m} + AS[\boldsymbol{m} - \boldsymbol{m}_0]) \quad (5)$$

The matrix *S* is diagonal, resulting from the imposition of positivity, and is a function of the model: *S*= *S*(***m***). Their *Q* is also model dependent: *Q*=*Q*(***m***). We can simplify equation (5) by ignoring the matrix *S* (setting it to identity matrix):

$$\boldsymbol{m} = \boldsymbol{m}_0 + \alpha Q A^T (A Q A^T + D)^{-1} (\boldsymbol{d}^{obs} - A\boldsymbol{m}_0) \quad (6)$$

Here matrix *Q* is an $M \times M$ diagonal matrix including depth weighting and sparsity constraint. Comparison of equation (6) with equation (4) shows that $C_M = \alpha Q$ and $C_D = D$. In Pilkington's formulation of the inverse problem, both depth weighting and sparsity constraint matrices are diagonal, so combining them and inverting to get *Q* is trivial. Equation (6) can be written as

$$\Delta\boldsymbol{m} = \boldsymbol{m} - \boldsymbol{m}_0 = \alpha Q A^T (A Q A^T + D)^{-1} (\boldsymbol{d}^{obs} - A\boldsymbol{m}_0) = \alpha Q A^T \boldsymbol{b} \quad (7)$$

The CG algorithm can be used on the following system of equations to compute $\boldsymbol{b} = (AQA^T + D)^{-1} (\boldsymbol{d}^{obs} - A\boldsymbol{m}_0)$

$$\boldsymbol{f} = G\boldsymbol{b} \quad (8)$$

$$G = (AQA^T + D) \quad (9)$$

$$\boldsymbol{f} = (\boldsymbol{d}^{obs} - A\boldsymbol{m}_0) \quad (10)$$

Comparison with minimum structure inversion approach:

The final solution of minimum structure approach, as discussed by Li & Oldenburg (1996), is

$$\boldsymbol{m} = \boldsymbol{m}_0 + (A^T W_d^T W_d A + W_m^T W_m)^{-1} A^T W_d^T W_d (\boldsymbol{d}^{obs} - A\boldsymbol{m}_0) \quad (11)$$

(note that A is typically written G for potential field problems). It is possible to write this equation in data space as (using eq. (3))

$$\boldsymbol{m} = \boldsymbol{m}_0 + (W_m^T W_m)^{-1} A^T (A (W_m^T W_m)^{-1} A^T + (W_d^T W_d)^{-1})^{-1} (\boldsymbol{d}^{obs} - A\boldsymbol{m}_0) \quad (12)$$

Compare against equation (7). However, here, the matrix $W_m^T W_m$ is typically not diagonal, e.g. if using smoothing regularization, which makes computing its inverse time-consuming, thereby removing the efficiencies offered by the data space formulation.

Extensions:

If someone seeks to achieve a balance between the data misfit and the regularization term in this approach, they may introduce tradeoff parameter β on the model term and obtain:

$$\boldsymbol{m} = \boldsymbol{m}_0 + (\beta W_m^T W_m)^{-1} A^T (A (\beta W_m^T W_m)^{-1} A^T + (W_d^T W_d)^{-1})^{-1} (\boldsymbol{d}^{obs} - A\boldsymbol{m}_0) \quad (13)$$

or if a tradeoff parameter λ is instead introduced on the data misfit term we'd have

$$\boldsymbol{m} = \boldsymbol{m}_0 + (W_m^T W_m)^{-1} A^T (A (W_m^T W_m)^{-1} A^T + (\lambda W_d^T W_d)^{-1})^{-1} (\boldsymbol{d}^{obs} - A\boldsymbol{m}_0) \quad (14)$$

All of the discussion so far has been for L2 norms. If instead we wanted to use Lp norms, Farquharson & Oldenburg (1998) show that this can be posed in a way that leads to a further iterative procedure involving $W_m^T R W_m$ where R is a function of the model: $R=R(\boldsymbol{m})$.

Farquharson, C. G and Oldenburg, D. W, 1998, Non-linear inversion using general measures of data misfit and model structure, Geophysical Journal International, 134(1), 213–227.

Conclusion:

You can do most things in the data space formulation, including incorporation of a regularization trade-off parameter, and general Lp norms, but it only works if your regularization operator is diagonal.

Here there are some references that uses data space approach:

- Last, B. J. and Kubik, K., 1983, Compact gravity inversion, GEOPHYSICS, 48(6), 713-721.
- Guillen, A. and Menichetti, V., 1984, Gravity and magnetic inversion with minimization of a specific functional, GEOPHYSICS, 49(8), 1354-1360.
- Barbosa, V. C. F. and Silva, J. B. C., Generalized compact gravity inversion, 1994, GEOPHYSICS, 59(1), 57-68.
- Siripunvaraporn, W. and Egbert, G., 2000, An efficient data-subspace inversion method for 2-D magnetotelluric data, GEOPHYSICS, 65(3), 791–803.
- Boonchaisuk, S., Vachiratienchai , C. and Siripunvaraporn, W., 2008, Two-dimensional direct current (DC) resistivity inversion: Data space Occam's approach, Physics of the Earth and Planetary Interiors, 168, 204–211.