Derivation of Damped Least Squares with Change of Basis

Sean Swenson

Climate and Global Dynamics Laboratory, National Center for Atmospheric Research, Boulder, CO, USA

1 Abstract

This is a derivation of the equations describing a damped least squares solution subject to the following conditions: 1) the damping parameter is a vector rather than a constant and 2) the damping is best expressed in an alternate basis set.

The specific application is the regularization of GRACE spherical harmonic (Stokes) coefficients, using constraints defined in the geographic (latitude/longitude) coordinate system.

This can be combined with constraints in the spherical harmonic domain, although this is not described here (see Swenson and Wahr, Estimating Signal Loss in Regularized GRACE Gravity Field Solutions, *Geophys. J. Int.*, 185: 693-702. doi: 10.1111/j.1365-246X.2011.04977.x, 2011).

2 Derivation

2.1 Damping with Constant Weights

This follows *Varah* [1979], section 2(c). In the standard least squares problem, we wish to solve for x that best fits the following equation:

$$H\ x = y\ , \tag{1}$$

where the vector y represents the data, the vector x represents the model coefficients, and the matrix H contains the partial derivatives relating changes in the model to changes in the data.

In standard least-squares, the cost function is defined as the square of the difference between the data and the predictions

$$\epsilon = (y - H\ x)^T (y - H\ x)\ , \tag{2}$$

and a solution is obtained by minimizing ϵ

$$\frac{\partial \epsilon}{\partial x} = 0\ . \tag{3}$$

This leads to a solution of the form

$$x = [H^T H]^{-1} H^T y\ , \tag{4}$$

where $[H^T H]^{-1}$ is the model covariance matrix.

If some additional property of the solution is desired, the cost function can be modified. By minimizing the solution length, large model parameter values are discouraged and the solution is said to be "damped" toward the initial guess. In this case the cost function takes the form

$$\epsilon_{damped} = \epsilon + \alpha^2 x^T x\ , \tag{5}$$

where α is a constant that determines the relative weight of solution error relative to model error in the cost function. Solutions based on the cost function (5) are called damped least squares, or regularized, solutions and have the form

$$x_\alpha = [H^T H \ + \ \alpha^2 I]^{-1} H^T y \ , \tag{6}$$

where I is the identity matrix. α is often determined through a parameter grid-search, i.e. trial-and-error.

2.2 Varying Weights

If α is not a constant, but a matrix whose coefficients vary, then (5) becomes

$$\epsilon_{damped} = \epsilon + (\alpha x)^T (\alpha x) \ , \tag{7}$$

and (6) becomes

$$x_\alpha = [H^T H \ + \alpha^T \alpha]^{-1} H^T y \ . \tag{8}$$

3 Surface Mass Tiles

3.1 Relationship between Geoid and Surface Mass Anomalies

Given a set of spherical harmonic coefficients ("Stokes" coefficients) describing the Earth's geoid (in this case, geoid anomalies relative to the temporal mean geoid), one can estimate surface mass anomalies, $\Delta\sigma$, in a geographic coordinate system [*Wahr et al.*, 1998]

$$\Delta\sigma(\theta,\phi) = \sum_{l=0}^{\infty}\sum_{m=0}^{l} \tilde{P}_{lm}(cos\theta)\{\Delta C_{lm} cosm\phi + \Delta S_{lm} sinm\phi\}. \tag{9}$$

where $\Delta\sigma$ is the surface mass anomaly, θ is co-latitude, ϕ is longitude, $\tilde{P}_{lm}$ are associated Legendre functions (normalized according to geodesy convention) and ΔC_{lm} and ΔS_{lm} are Stokes coefficients converted to units of surface mass, i.e. mass per area. In the following, the Δ notation is ignored for clarity, and it should be assumed that these terms are anomalies.

The Stokes coefficients and the coefficients describing the corresponding changes in surface mass are related by the following expression

$$\left\{ \begin{array}{c} \Delta C_{lm} \\ \Delta S_{lm} \end{array} \right\} = \frac{a\rho_{earth}}{3} \frac{(2l+1)}{1+k_l} \left\{ \begin{array}{c} \Delta C'_{lm} \\ \Delta S'_{lm} \end{array} \right\} . \tag{10}$$

where ρ_{earth} is the average density of the Earth, k_l are the load Love numbers, and $\Delta C'_{lm}$ and $\Delta S'_{lm}$ are dimensionless Stokes coefficients.

3.2 Spatially Averaged Surface Mass Anomalies

The average surface mass value of a region can be obtained by multiplying the surface mass field by a mask, M, whose value is 1 inside the region and 0 outside the region and integrating over the sphere

$$\overline{\sigma} = \frac{1}{\Omega} \int d\Omega \; M(\theta,\phi) \; \sigma(\theta,\phi) \; , \tag{11}$$

$$\overline{\sigma} = \frac{1}{\Omega} \int d\Omega \; \sum_{l=0}^{\infty} \sum_{m=0}^{l} M(\theta,\phi) \; \tilde{P}_{lm}(cos\theta) \{\Delta C_{lm} cosm\phi + \Delta S_{lm} sinm\phi\}. \tag{12}$$

The sphere can be covered by a set of masks, or "tiles", of arbitrary shape, size, and number. Sometimes these tiles are referred to as "mascons", which is jargon for "mass concentrations" coined in the early days of satellite geodesy, when accuracy limited gravity field solutions to lumped coefficient sets. Because this term seems mysterious or technical to some people, the generic term "tile" is used here instead.

By organizing the spherical harmonic basis into a one dimensional format, equation 12 can be expressed in matrix notation as

$$m = Tx \ , \tag{13}$$

where m is now a vector of N tile averaged surface mass anomalies ($\overline{\sigma}$), x is a vector of l_{tot} spherical harmonic coefficients, and T is a transfer matrix relating the two. An element of T can be obtained from

$$T_{ij} = \frac{1}{\Omega} \int d\Omega \ M_i(\theta, \phi) \ Y_j(\theta, \phi) \ , \tag{14}$$

where i is the index of tiles, j is the index of spherical harmonic coefficients, and Y_j is a spherical harmonic organized in vector format (i.e. Y_j corresponds to some $\tilde{P}_{lm}(cos\theta) \ cosm\phi$ or $\tilde{P}_{lm}(cos\theta) \ sinm\phi$). Thus, each row of T contains the spherical harmonic coefficients describing each tile.

4 Least Squares Tile Inversion

4.1 LSQ Tile Transformation

The original least squares problem relating the spacecraft data to the surface mass spherical harmonic coefficients is expressed as

$$H \ x = y \ , \tag{15}$$

and the surface mass tile averages are related to the same spherical harmonic set x by

$$m = Tx \ . \tag{16}$$

If H can be expressed as the product of T and another matrix A as

$$H = A \ T \ , \tag{17}$$

then

$$A\ T\ x = A\ m = y. \tag{18}$$

A can be expressed in terms of H and T according to equation 4 as

$$A = [T^T T]^{-1} T^T H\ , \tag{19}$$

The least squares solution of 18

$$m = [A^T A]^{-1} A^T y\ , \tag{20}$$

can be written in terms of H and T. Let

$$C = [T^T T]^{-1}\ , \tag{21}$$

and

$$D = T^T C\ . \tag{22}$$

Then

$$A^T A = D^T H^T H D\ , \tag{23}$$

and

$$A^T y = D^T H^T y\ , \tag{24}$$

which can be used in 20 to give

$$m = [D^T H^T H D]^{-1} D^T H^T y\ , \tag{25}$$

or

$$m = [D^T H^T H D]^{-1} D^T H^T H x \ , \tag{26}$$

which relates the tiles m to the original spherical harmonic coefficients x through the tile transfer matrix T (via D) and the original inverse covariance matrix $H^T H$.

The purpose of all this algebraic manipulation is to obtain an expression analogous to equation 4 which can be modified into the form of equation 8. This leads to the following equation, which has the desired property of having a damping matrix expressed in the coordinate domain, rather than the spherical harmonic domain

$$m = [D^T H^T H D \ + \alpha^T \alpha]^{-1} D^T H^T H x \ . \tag{27}$$

Because the damping matrix $\alpha^T \alpha$ is based in coordinate space, damping metrics defined in coordinate space can be employed to regularize the tile values.

4.2 Coordinate Domain Damping

The most basic form of damping uses non-uniform weights to constrain the amplitude of the model coefficients. In the spherical harmonic domain, this method has been used to reduce the large amplitude high degree GRACE Stokes coefficients [*Swenson and Wahr*, 2011]. However, this method basically damps both signal and noise impartially. Geographical information can provide more targeted constraints by using estimates of the spatial pattern of surface mass variability to damp the tile solutions. In addition to the pattern of variability, which can be described by a diagonal matrix, constraints on the relationship between tiles can be implemented using non-diagonal damping matrices. For example, a damping matrix can be created that gives greater weight to tiles that have greater departures from their neighbors. This type of constraint gives preference to solutions exhibiting spatial correlations.

A matrix equation quantifying the difference between a tile and the spatial average over its sourroundings can be written

$$\alpha = G - I \ , \tag{28}$$

where G is a non-diagonal matrix, and I is the identity matrix. The product Gm gives the weighted average of all the tiles. Adding a constraint on the magnitude of each tile's variability can be obtained by dividing by an estimate of the variance of each tile

$$\alpha = \frac{G - I}{\sigma_{tile}} \ , \tag{29}$$

where σ_{tile} can be based on a model of surface mass variability, e.g. a land surface model, or on a scaled initial solution from the GRACE data itself.

5 References

Varah, J. M., 1973, On the Numerical Solution of Ill-Conditioned Linear Systems with Applications to Ill-Posed Problems, SIAM Journal on Numerical Analysis, Vol. 10, No. 2., Apr., pp. 257-267.

Varah, J. M., 1979, A practical examination of some numerical methods for linear discrete ill-posed problems, SIAM Rev., 21, pp. 100-111.

Wahr, J., M. Molenaar, and F. Bryan, Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE, *J. Geophys. Res.*, 103(B12), 30205-30229, 1998.

Swenson, S., and J. Wahr, 2011, Estimating Signal Loss in Regularized GRACE Gravity Field Solutions, *Geophys. J. Int.*, 185: 693-702. doi: 10.1111/j.1365-246X.2011.04977.x.