SIO 224

1. A brief look at resolution analysis

Here is some background for the Masters and Gubbins resolution paper.

Global Earth models are usually found iteratively by assuming a starting model and finding small perturbations to it. Usually, the "smallest" or "smoothest" perturbation is found that allows the data to be fit to some tolerance. This means that the final model can end up with features of the starting model which are not required by the data. These notes describe one way of assessing the true "resolution" of the data.

For concreteness, we consider the free oscillation inverse problem. Suppose we let $\mathbf{m}(r)$ be the true Earth and $\mathbf{m}_0(r)$ be a starting model where $\mathbf{m}$ is usually taken to be the triplet of functions:

$$\mathbf{m}(r) = (\rho(r), K_s(r), \mu(r))$$

then let

$$\delta\mathbf{m} = \mathbf{m} - \mathbf{m}_0$$

then we obtain for the free-oscillation problem

$$\delta\omega_i = \langle \mathbf{G}_i(\mathbf{m}_0), \delta\mathbf{m} \rangle + O|\delta\mathbf{m}|^2 \tag{1}$$

where, for the i'th mode

$$\delta\omega_i = \omega_{i_{obs}} - \omega_{i_{model}}$$

The braket notation is shorthand for:

$$\langle \mathbf{G}_i(\mathbf{m}_0), \delta\mathbf{m} \rangle = \int_0^R \left(G_{\rho_i}\delta\rho(r) + G_{K_i}\delta K_s(r) + G_{\mu_i}\delta\mu(r) \right) \, dr$$

and the G_i's can be computed from the eigenfunctions of the i'th mode for the starting model and so are implicitly a function of $\mathbf{m}_0$. Equation 1 is solved under the assumption that $|\delta\mathbf{m}|$ is small enough so that terms of order $|\delta\mathbf{m}|^2$ can be neglected.

Equation 1 is a linearization of the problem and so, in principle, it is possible to find many different $\mathbf{m}$ which satisfy the data and which may not be linearly close to one another. This occurs if the $\mathbf{G}$ are rapidly changing functions of m_0. For the case of free-oscillations, the linearization is probably valid since

$$\langle \mathbf{G}_i, \delta\mathbf{m} \rangle$$

correctly predicts the perturbation in frequency, $\delta\omega_i$, for perturbations up to several percent. (This isn't true for all modes but is true for the vast majority). A few percent is the typical uncertainty in the models of the spherically averaged Earth.

Suppose we have found a model which gives a satisfactory fit to the data, i.e. the residuals, $\delta\omega_i$, are normally distributed with zero mean. How do we find those features of the model which are truly resolved by the data? This question is addressed by the paper of Backus and Gilbert (1970). For simplicity, consider a one-dimensional model (the model consists of only one function of radius):

$$\delta\omega_i \pm \sigma_i = \int_0^R G_i(r)\delta m(r)\, dr$$

for $i = 1$ to N. Suppose we take a linear combination of data

$$\sum_i^N a_i \delta\omega_i = \int_0^R A(r)\delta m(r)\, dr$$

$$\text{where} \quad A(r) = \sum a_i G_i(r)$$

Suppose we choose the multipliers, a_i so that $A(r)$ approximates a δ-function peaked at a particular radius, r_0. If we achieved this perfectly, we would have

$$\sum_i^N a_i \delta\omega_i = \int_0^R \delta(r - r_0)\delta m(r)\, dr = \delta m(r_0)$$

With a finite amount of data we cannot make A(r) a perfect δ-function but we can try and make it as δ-like as possible. We then have

$$\begin{aligned}\sum_i^N a_i \delta\omega_i &= \int_0^R A(r)\delta m(r)\, dr \\ &= \int_0^R A(r)(m(r) - m_0(r))\, dr = 0\end{aligned}$$

where we have assumed that the model, $m_0(r)$ is linearly close to the real Earth, $m(r)$ and that the model fits the data so that the expected value of the residuals is zero. We thus obtain

$$\int_0^R A(r)m(r)\, dr = \int_0^R A(r)m_0(r)\, dr = \bar{m}(r_0) \quad \text{say}$$

where $\bar{m}(r_0)$ is an average of the *real* Earth (averaged with our approximation to a δ-function) and is identical to the same average of our model. We force the average to be unbiased by making $A(r)$ unimodular, i.e.

$$\int_0^R A(r)\, dr = 1$$

The data also have errors (σ_i) and we suppose that the uncertainties in the data are characterized by a covariance matrix, E_{ij}. We usually don't know what the covariances between our data are so we assume that the data are independent in which case E_{ij} is diagonal with elements along the diagonal which are the variances of the data: σ_i^2. The variance of out estimate, $\bar{m}$ is then given by

$$\bar{\sigma}^2 = \sum_{ij} a_i a_j E_{ij}$$

We would like this to be as small as possible. We want our multipliers to make $A(r)$ as δ-like as possible at a radius r_0 to localize information about $m(r)$ around r_0 and at the same time we want the localized information to be precise. Backus and Gilbert show that these aims are mutually exclusive. How do we choose the a_i's to make $A(r)$ peaked? Consider minimizing the form

$$S = \int_0^R f(r) A^2(r)\, dr$$

If $f(r)$ is dipped near r_0 then we would expect A(r) to be peaked at r_0. Backus and Gilbert suggest using a parabola:

$$f(r) = 12(r - r_0)^2$$

The factor of 12 is introduced to make S a measure of the peak width of A which we shall call the "spread". (If $A(r)$ is a boxcar of unit area centered at r_0 then S is exactly the width of the boxcar). We now have

$$S = \sum_{ij} a_i a_j S_{ij}$$

where

$$S_{ij} = 12 \int_0^R (r - r_0)^2 G_i(r) G_j(r)\, dr$$

If we define

$$u_i = \int_0^R G_i(r)\, dr$$

the unimodularity constraint reads

$$\sum a_i u_i = 1$$

Since $\bar{\sigma}^2$ and S cannot be minimized simultaneously, we consider the following combination:

$$M_{ij} = S_{ij} \cos\theta + w E_{ij} \sin\theta \qquad 0 \le \theta \le \pi/2$$

and minimise M where

$$M = \sum a_i a_j M_{ij} \quad \text{subject to} \quad \sum a_i u_i = 1$$

θ is called a "tradeoff parameter". When $\theta = 0$, we choose the a_i to minimize the spread. When $\theta = \pi/2$ we choose the a_i to minimize $\bar{\sigma}^2$. At intermediate values we compute a compromise between

spread and error. w is a weighting factor to make $w||\mathbf{E}||$ and $||\mathbf{S}||$ to be about the same – the tradeoff calculation will then be centered about $\theta \simeq 45°$. Written in vector form, our problem is

$$\text{minimize} \quad \mathbf{a} \cdot \mathbf{M} \cdot \mathbf{a} \quad \text{with} \quad \mathbf{a} \cdot \mathbf{u} = 1$$

The is a problem in calculus of variations and is solved by introducing a Lagrange multiplier, λ:

$$\text{minimize} \quad \mathbf{a} \cdot \mathbf{M} \cdot \mathbf{a} + \lambda(\mathbf{a} \cdot \mathbf{u} - 1)$$

Differentiating with respect to $\mathbf{a}$ and setting equal to zero gives

$$2\mathbf{M} \cdot \mathbf{a} + \lambda \mathbf{u} = 0$$

Thus

$$\mathbf{a} = -\frac{\lambda}{2}\mathbf{M}^{-1} \cdot \mathbf{u}$$

We can evaluate the Lagrange multiplier by dotting the above equation with $\mathbf{u}$ which gives

$$\mathbf{a} \cdot \mathbf{u} = 1 = -\frac{\lambda}{2}\mathbf{u} \cdot \mathbf{M}^{-1} \cdot \mathbf{u}$$

So elliminating $\lambda/2$ gives

$$\mathbf{a} = \frac{\mathbf{M}^{-1} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{M}^{-1} \cdot \mathbf{u}}$$

Note that $\mathbf{a}$ must be recalculated for each value of r_0 and θ and the calulation is made much more efficient if $\mathbf{M}$ is diagonal. This can be achieved (Gilbert 1971) but we don't consider numerical niceties any further here. Once $\mathbf{a}$ is computed, the spread: $\mathbf{a} \cdot \mathbf{S} \cdot \mathbf{a}$; the variance: $\mathbf{a} \cdot \mathbf{E} \cdot \mathbf{a}$, and the resolving kernel: $\mathbf{a} \cdot \mathbf{G}(r)$ can all be computed.

Some results of applying this technique to the mode problem are given in the accompanying PEPI article. In the example, we simultaneously try and peak information about density at some target radius while removing sensitivity to the elastic moduli. Rather than computing a tradeoff curve, we choose a specific error level (0.5% say) then adjust θ until the $\mathbf{a}$ that gives this error level is found.

N.B. The form chosen for S_{ij} above is relatively arbitrary. We might decide (as is done in the PEPI paper) that we want our resolving kernel to be a boxcar between radii r_1 and r_2 with unit area between these limits:

$$\int_{r_1}^{r_2} A(r)\, dr = 1$$

If $B(r)$ is the desired boxcar then we would minimize

$$\int_0^R (A(r) - B(r))^2\, dr$$

Substituting in $A = \sum a_i G_i$ and expanding the square gives

$$\mathbf{a}\cdot\mathbf{S}\cdot\mathbf{a} - 2\mathbf{a}\cdot\mathbf{u} + \int_0^R B(r)^2\,dr$$

where we have redfined $\mathbf{S}$ and $\mathbf{u}$:

$$S_{ij} = \int_0^R G_i(r)G_j(r)\,dr$$

and

$$u_i = \int_{r_1}^{r_2} G_i(r)\,dr$$

Since $\mathbf{a}\cdot\mathbf{u}$ is forced to be one, the only part of the above equation that depends on $\mathbf{a}$ is $\mathbf{a}\cdot\mathbf{S}\cdot\mathbf{a}$. We now form $\mathbf{M} = \mathbf{S}\cos\theta + w\mathbf{E}\sin\theta$ and find that we get the same answer for $\mathbf{a}$ as before:

$$\mathbf{a} = \frac{\mathbf{M}^{-1}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{M}^{-1}\cdot\mathbf{u}}$$

but with the redefined $\mathbf{M}$ and $\mathbf{u}$. This form is computationally efficient since $\mathbf{M}$ no longer depends on the target depth range for the boxcar. Only $\mathbf{u}$ has to be recomputed for new r_1, r_2.