Second-order Correction to Linearized Resolution Matrix

Bill Menke, October 20, 2022

In a linear inverse problem, the resolution matrix, R_{ij}, defines unique weighted averages, $\langle m \rangle_i$, of model parameters, $\mathbf{m}$:

$$\langle m \rangle_i \equiv \sum_j R_{ij}\, m_j \tag{1}$$

When the model parameters possess some natural ordering, the i-th weighted average, $\langle m \rangle_i$, is typically chosen to be localized around m_i. The model parameters can be represented as deviations, $\Delta\mathbf{m}$, with respect to a reference value, $\mathbf{m}_0$; that is, $\mathbf{m} \equiv \mathbf{m}_0 + \Delta\mathbf{m}$. Then,

$$\langle m \rangle_i = \sum_j R_{ij}\ m_{0j} + \sum_j R_{ij}\ \Delta m_j = \langle m_0 \rangle_i + \langle \Delta m \rangle_i$$
$$\text{with}\quad \langle m_0 \rangle_i \equiv \sum_j R_{ij}\ m_{0j} \quad \text{and} \quad \langle \Delta m \rangle_i \equiv \sum_j R_{ij}\ \Delta m_j \tag{2}$$

As I have shown previously for the linearized problem, the resolution matrix varies with the choice of the reference value, $\mathbf{m}_0$; that is, $R_{ij}(\mathbf{m}_0)$. The expression,

$$\langle \Delta m \rangle_i \equiv \sum_j R_{ij}(\mathbf{m}_0)\ \Delta m_j \tag{3}$$

indicates that in the vicinity of $\mathbf{m}_0$, and for small perturbations, $\Delta\mathbf{m}$, the averages, $\langle \Delta m \rangle_i$, are unique. We now view $R_{ij}(\mathbf{m}_0)$ as the leading term of a Taylor series for $R_{ij}(\mathbf{m})\big|_{\mathbf{m}_0}$:

$$\langle \Delta m \rangle_i = \sum_j R_{ij}^{(0)} \Delta m_j + \frac{1}{2}\sum_j \sum_k R_{ijk}^{(0)} \Delta m_j \Delta m_j \quad + \cdots \quad \text{with} \quad R_{ij}^{(0)} \equiv R_{ij}(\mathbf{m}_0) \quad \text{and} \quad R_{ijk}^{(0)} \equiv \left.\frac{\partial R_{ij}}{\partial m_k}\right|_{\mathbf{m}_0} \tag{4}$$

The second term represents a weighted average of the elements of the dyad, ½$\Delta m_j \Delta m_j$, with weights, $R_{ijk}^{(0)}$. This modified version of the resolution equation is valid for larger deviations, but at the expense of adding complexity (Figure 1). The quantity, $\langle \Delta m \rangle_i$, is no longer a single weighted average of the model parameters, but the sum of two weighted averages, the second of which is quadratic in the model parameters. Furthermore, whether or not both $R_{ij}^{(0)}$ and $R_{ijk}^{(0)}$ simultaneously can be chosen (or even should to be chosen) to be localized is unclear.

As an example, consider $N = 2$, $M = 3$ case with data equations:

$$d_1 = m_1^2 + m_1^2 + (m_3 - 1)^2$$

$$d_2 = m_1^2 + m_1^2 + (m_3 + 1)^2 \tag{5}$$

Each data equation constrains the model parameters to lie on a sphere. Taken together, they imply:

$$d_2 - d_1 = 2m_3 \quad \text{or} \quad m_3 = \frac{d_2 - d_1}{2} \equiv A$$
$$m_1^2 + m_1^2 = d_2 - (m_3 + 1)^2 \equiv B^2 \tag{6}$$

The problem only has a solution when $B^2 > 0$. In that case, the model parameter, $m_3 = A$, is unique, but the model parameters m_1 and m_2 are nonunique, lying on a circle or radius, B, centered upon $(m_1, m_2, m_3) = (0,0,A)$ and lying in the $m_3 = A$ plane.

We focus on the plane containing the circle (Figure 2). The reference point, $\mathbf{m}_0$, is taken to be on the plane and near the circle, at radius, $b \approx B$, and making an angle, θ, with respect to the m_1 axis. We define rotated coordinates, $(\Delta m'_1, \Delta m'_2,)$ that are locally perpendicular and tangent, respectively, to the circle. The data constrain the radial component, $\Delta m'_1 = \cos(\theta)\,\Delta m_1 + \sin(\theta)\,\Delta m_2$, but not the tangential component, $\Delta m'_2 = -\sin(\theta)\,\Delta m_1 + \cos(\theta)\,\Delta m_2$. Thus, only one unique weighted average is possible, corresponding to a one-row resolution matrix:

$$\mathbf{R}(m_1, m_2) = [\cos(\theta) \quad \sin(\theta)] = \left[\frac{m_1}{[m_1^2 + m_2^2]^{½}} \quad \frac{m_2}{[m_1^2 + m_2^2]^{½}}\right] \tag{7}$$

We now differentiate R_{ij} to obtain the R_{ijk}. Noting that $m_1 = b\cos(\theta)$ and $m_2 = b\sin(\theta)$, we find

$$R_{111} = \frac{\partial \mathrm{R}_{11}}{\partial m_1} = [m_1^2 + m_2^2]^{-½} - {m_1}^2[m_1^2 + m_2^2]^{-3/2} = b^{-1} - b^{-1}\cos^2(\theta) = b^{-1}\sin^2(\theta)$$
$$R_{112} = \frac{\partial \mathrm{R}_{11}}{\partial m_2} = -m_1 m_2 [m_1^2 + m_2^2]^{-3/2} = -b^{-1}\sin(\theta)\cos(\theta)$$
$$R_{221} = \frac{\partial \mathrm{R}_{11}}{\partial m_1} = -b^{-1}\sin(\theta)\cos(\theta)$$
$$R_{222} = \frac{\partial R_{22}}{\partial m_2} = [m_1^2 + m_2^2]^{-½} - {m_2}^2[m_1^2 + m_2^2]^{-3/2} = b^{-1}\cos^2(\theta) \tag{8}$$

The overall scaling factor of b^{-1} implies $R_{ijk} \to 0$ as $b \to \infty$; that is, the problem becomes increasingly linear as the radius of the circle approaches infinity.

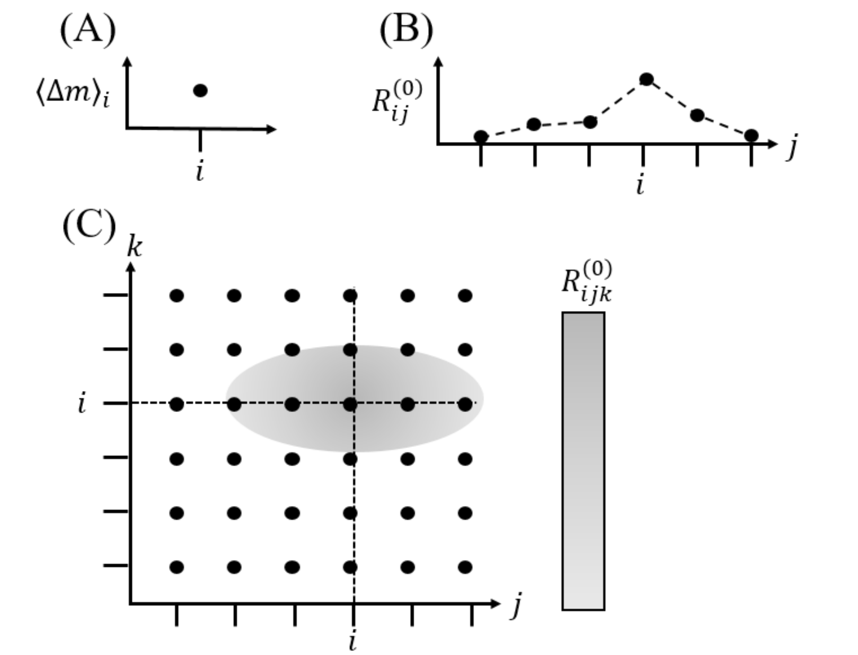

Figure 1. Graphical interpretation of Equation 3. (A) The left-hand side is the weighted average of the deviatoric model parameters, localized at position, i. (B) The first term on the right-hand side represents averaging the deviatoric model parameters, Δm_j, by applying the weighted averages, $R_{ij}^{(0)}$, depicted here as peaked at position, i. (C) The second term on the right-hand side represents averaging the dyad, $\Delta m_j \Delta m_k$, by applying the weighted averages, $R_{ijk}^{(0)}$, depicted here as peaked at position, i.

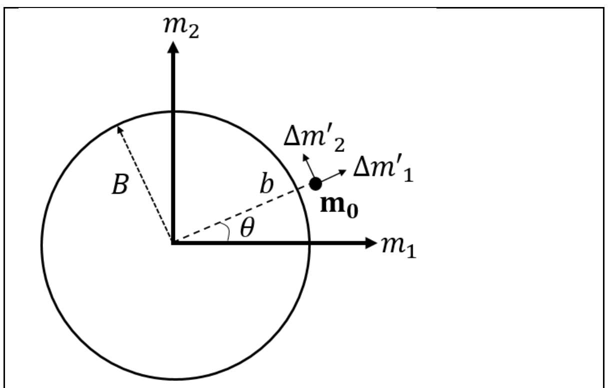

Figure 2. Geometry of the (m_1, m_2) plane in the example. See the text for further discussion.