Properties of the Resolution Matrix in Generalized Least Squares

Bill Menke, October 12, 2022

Definitions

Let $\mathbf{m}$ be a length-M model parameter vector, $\mathbf{G}$ be an $N \times M$ data kernel matrix satisfying $\mathbf{Gm} = \mathbf{d}$, and $\mathbf{H}$ be an $K \times M$ prior information matrix satisfying $\mathbf{Hm} = 0$. The generalized least squares solution is $\mathbf{m} = \mathbf{G}^{-\mathrm{g}}\mathbf{d}$ where $\mathbf{G}^{-\mathrm{g}} \equiv [\mathbf{G}^{\mathrm{T}}\mathbf{G} + \varepsilon\mathbf{H}^{\mathrm{T}}\mathbf{H}]^{-1}\mathbf{G}^{\mathrm{T}}$ and where ε is a ratio of variances. The resolution matrix is $\mathbf{R} \equiv \mathbf{G}^{-\mathrm{g}}\mathbf{G}$.

Condition for which the resolution matrix has unit row sum

Let $\mathbf{m}$ be a length-M model parameter vector, $\mathbf{G}$ be an $N \times M$ data kernel matrix satisfying $\mathbf{Gm} = \mathbf{d}$, and $\mathbf{H}$ be an $K \times M$ prior information matrix satisfying $\mathbf{Hm} = 0$. The generalized least squares solution is $\mathbf{m} = \mathbf{G}^{-\mathrm{g}}\mathbf{d}$ where $\mathbf{G}^{-\mathrm{g}} \equiv [\mathbf{G}^{\mathrm{T}}\mathbf{G} + \varepsilon\mathbf{H}^{\mathrm{T}}\mathbf{H}]^{-1}\mathbf{G}^{\mathrm{T}}$ and where ε is a ratio of variances. The resolution matrix is $\mathbf{R} \equiv \mathbf{G}^{-\mathrm{g}}\mathbf{G}$. Let $\mathbf{w} = [1, 1, 1, \ \cdots, 1]^{\mathrm{T}}$ be a length-M vector of all ones, so that the row sum of $\mathbf{R}$ is $\mathbf{Rw}$. The resolution matrix, $\mathbf{R}$, has unit row sum when the prior information matrix, **H,** has zero row sum; that is $\mathbf{Hw} = \mathbf{0}$.

$$\mathbf{Rw} \stackrel{?}{=} \mathbf{w}$$

$$[\mathbf{G}^{\mathrm{T}}\mathbf{G} + \varepsilon\mathbf{H}^{\mathrm{T}}\mathbf{H}]^{-1}\mathbf{G}^{\mathrm{T}}\mathbf{Gw} \stackrel{?}{=} \mathbf{w}$$

$$\mathbf{G}^{\mathrm{T}}\mathbf{Gw} \stackrel{?}{=} [\mathbf{G}^{\mathrm{T}}\mathbf{G} + \varepsilon\mathbf{H}^{\mathrm{T}}\mathbf{H}]\mathbf{w} \stackrel{?}{=} \mathbf{G}^{\mathrm{T}}\mathbf{Gw} + \varepsilon\mathbf{H}^{\mathrm{T}}\mathbf{Hw}$$

$$\mathbf{G}^{\mathrm{T}}\mathbf{Gw} = \mathbf{G}^{\mathrm{T}}\mathbf{Gw} + 0$$

Except when $\varepsilon = 0$ or $K = 0$, this is a necessary as well as sufficient condition, (since $\mathbf{H}^{\mathrm{T}}\mathbf{H}$ cannot be zero.

The resolution matrix can always be adjusted to have unit row sum

The key property of the resolution matrix, $\mathbf{R}$, is that it must be built from the rows of the data kernel, so that the predicted data, $\mathbf{m}^{est}$, in $\mathbf{m}^{est} = \mathbf{Rm}$, when viewed as localized average of $\mathbf{m}$, are unique. This condition implies that $\mathbf{R} = \mathbf{G}^{-g}\mathbf{G}$, where $\mathbf{G}^{-g}$ is some generalized inverse. If some particular generalized inverse does not lead to an $\mathbf{R}$ with unit row sum, one can always modify the definition of $\mathbf{G}^{-g}$ to some new generalized inverse, say $(\mathbf{G}^{-g})$, that it does. Suppose $\mathbf{Rw} = \mathbf{s}$, where $\mathbf{s}$ is a vector of row sums. Then, defining $\mathbf{S} \equiv \mathrm{diag}(\mathbf{s})$:

$$\mathbf{Rw} = \mathbf{G}^{-g}\mathbf{Gw} = \mathbf{s} = \mathbf{Sw}$$

$$\mathbf{S}^{-1}\mathbf{G}^{-g}\mathbf{Gw} = \mathbf{w}$$

$$\mathbf{R}'\mathbf{w} = (\mathbf{G}^{-g})'\mathbf{Gw} = \mathbf{w} \quad \text{with} \quad (\mathbf{G}^{-g})' = \mathbf{S}^{-1}\mathbf{G}^{-g}$$

The new localized averages (rows of $\mathbf{R}$) have the same shape as the old ones. However, depending on how the spread functions are define, the spread of resolution may be different. The Dirichlet spread function:

$$\text{spread}_D\left(\mathbf{r}'^{(i)}\right) = \sum_{j=1}^{M}\left(s_i^{-1}r_j - \delta_{ij}\right)^2 \neq \sum_{j=1}^{M}\left(r_i - \delta_{ij}\right)^2 = \text{spread}_D\left(\mathbf{r}^{(i)}\right)$$

and the Backus-Gilbert spread function:

$$\text{spread}_{BG}\left(\mathbf{r}'^{(i)}\right) = \sum_{j=1}^{M}\left(s_i^{-1}r_j\right)(i-j)^2 = s_i^{-1}\,\text{spread}_{BG}\left(\mathbf{r}^{(i)}\right)$$

are both different. The covariance changes, too:

$$\text{cov}(\mathbf{m}^{est}) = \sigma_d^2(\mathbf{G}^{-g})'(\mathbf{G}^{-g})'^{\boldsymbol{T}} = \sigma_d^2\mathbf{S}^{-1}\mathbf{G}^{-g}\mathbf{G}^{-gT}\mathbf{S}^{-1} \neq \sigma_d^2\mathbf{G}^{-g}\mathbf{G}^{-gT}$$

Condition for which the Resolution Matrix is Symmetric

A matrix is symmetric hen it is equal to its transpose. Note that $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ and $\mathbf{H}^{\mathrm{T}}\mathbf{H}$ are always symmetric.

$$\mathbf{R} \stackrel{?}{=} \mathbf{R}^{\boldsymbol{T}}$$

$$[\mathbf{G}^{\mathrm{T}}\mathbf{G} + \varepsilon\mathbf{H}^{\mathrm{T}}\mathbf{H}]^{-1}\mathbf{G}^{\mathrm{T}}\mathbf{G} \stackrel{?}{=} \mathbf{G}^{\mathrm{T}}\mathbf{G}[\mathbf{G}^{\mathrm{T}}\mathbf{G} + \varepsilon\mathbf{H}^{\mathrm{T}}\mathbf{H}]^{-1}$$

$$\mathbf{G}^{\mathrm{T}}\mathbf{G}[\mathbf{G}^{\mathrm{T}}\mathbf{G} + \varepsilon\mathbf{H}^{\mathrm{T}}\mathbf{H}] \stackrel{?}{=} [\mathbf{G}^{\mathrm{T}}\mathbf{G} + \varepsilon\mathbf{H}^{\mathrm{T}}\mathbf{H}]\mathbf{G}^{\mathrm{T}}\mathbf{G}$$

$$\mathbf{G}^{\mathrm{T}}\mathbf{G}\mathbf{G}^{\mathrm{T}}\mathbf{G} + \varepsilon\mathbf{G}^{\mathrm{T}}\mathbf{G}\mathbf{H}^{\mathrm{T}}\mathbf{H} = \mathbf{G}^{\mathrm{T}}\mathbf{G}\mathbf{G}^{\mathrm{T}}\mathbf{G} + \varepsilon\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{G}^{\mathrm{T}}\mathbf{G}$$

$$\mathbf{G}^{\mathrm{T}}\mathbf{G}\mathbf{H}^{\mathrm{T}}\mathbf{H} = \mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{G}^{\mathrm{T}}\mathbf{G}$$

So $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ must commute with $\mathbf{H}^{\mathrm{T}}\mathbf{H}$. In general, for two arbitrary matrices $N \times M$ and $K \times M$ matrices, $\mathbf{G}$ and $\mathbf{H}$, $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ does not commute with $\mathbf{H}^{\mathrm{T}}\mathbf{H}$. Special cases in which they do commute include.

(A) When $\mathbf{G}$ and $\mathbf{H}$ are diagonal matrices, so that $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ and $\mathbf{H}^{\mathrm{T}}\mathbf{H}$ also are diagonal and commute.

(B) When $\mathbf{G}$ and $\mathbf{H}$ are orthogonal matrices, so that $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ and $\mathbf{H}^{\mathrm{T}}\mathbf{H}$ are diagonal and commute.

(C) When $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ and $\mathbf{H}^{\mathrm{T}}\mathbf{H}$ share the same eigenvectors, and so can be simultaneously diagonalized.

(D) When, $\mathbf{G}$ and $\mathbf{H}$ are discrete convolutions (Toeplitz matrices), then $\mathbf{G}$, $\mathbf{H}$, $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ and $\mathbf{H}^{\mathrm{T}}\mathbf{H}$

asymptotically commute (commute in the limit, $M \to \infty$).

Construction of the Resolution Matrix from Asserted-Predicted Solution Pairs.

The resolution matrix relates a predicted solution, $\mathbf{m}^{pre}$, to an asserted solution, $\mathbf{m}$, via $\mathbf{m}^{pre} = \mathbf{Rm}$. Suppose that M asserted solutions, $\mathbf{m}^{(i)}$, lead to M predicted solutions, $\mathbf{m}^{pre(i)}$. Then,

$$\begin{bmatrix} m_1^{pre(1)} & \cdots & m_1^{pre(M)} \\ \cdots & \ddots & \cdots \\ m_M^{pre(1)} & \cdots & m_M^{pre(M)} \end{bmatrix} = \mathbf{R} \begin{bmatrix} m_1^{(1)} & \cdots & m_1^{(M)} \\ \cdots & \ddots & \cdots \\ m_M^{(1)} & \cdots & m_M^{(M)} \end{bmatrix}$$

Then, as long as the asserted solutions are linearly independent, $\mathbf{R}$ can be constructed as:

$$\mathbf{R} = \begin{bmatrix} m_1^{pre(1)} & \cdots & m_1^{pre(M)} \\ \cdots & \ddots & \cdots \\ m_M^{pre(1)} & \cdots & m_M^{pre(M)} \end{bmatrix} \begin{bmatrix} m_1^{(1)} & \cdots & m_1^{(M)} \\ \cdots & \ddots & \cdots \\ m_M^{(1)} & \cdots & m_M^{(M)} \end{bmatrix}^{-1}$$

since $\mathbf{m}^{pre(i)} = \mathbf{G}^{-\mathrm{g}}\mathbf{d}^{(i)}$ and $\mathbf{d}^{(i)} = \mathbf{G}\mathbf{m}^{(i)}$, we can also write

$$\mathbf{R} = \mathbf{G}^{-\mathrm{g}}\mathbf{G} \begin{bmatrix} m_1^{(1)} & \cdots & m_1^{(M)} \\ \cdots & \ddots & \cdots \\ m_M^{(1)} & \cdots & m_M^{(M)} \end{bmatrix} \begin{bmatrix} m_1^{(1)} & \cdots & m_1^{(M)} \\ \cdots & \ddots & \cdots \\ m_M^{(1)} & \cdots & m_M^{(M)} \end{bmatrix}^{-1} = \mathbf{G}^{-\mathrm{g}}\mathbf{G}\mathbf{I} = \mathbf{G}^{-\mathrm{g}}\mathbf{G}$$

which recovers the usual formula for the resolution matrix.