Appearance of the Data Resolution Matrix in Generalized Least Squares Formulas

Bill Menke during a discussion with Zach Eilon

February 17, 2024

Synopsis

I have previously shown for Generalized Least Squares (GLS) that formulas for the estimated model parameters $\mathbf{m}^{est}$ and their covariance $\mathbf{C}_{\mathbf{m}^{est}}$ can be reformulated to highlight their dependence on the model resolution matrix $\mathbf{R}_G \equiv \mathbf{G}^{-g}\mathbf{G}$:

$$\mathbf{m}^{est} = \mathbf{G}^{-g}\mathbf{d}^{obs} + (\mathbf{I} - \mathbf{R}_G)\mathbf{m}_A$$

$$\mathbf{C}_{\mathbf{m}^{est}} = \mathbf{G}^{-g}\mathbf{C}_d\mathbf{G}^{-gT} + (\mathbf{I} - \mathbf{R}_G)\mathbf{C}_{\mathbf{m}_A}(\mathbf{I} - \mathbf{R}_G)^T$$

Here, $\mathbf{G}^{-g}$ is the GLS generalized inverse, $\mathbf{d}^{obs}$ are the observed data and $\mathbf{C}_d$ is their covariance and $\mathbf{m}_A$ are the prior model parameters and $\mathbf{C}_{\mathbf{m}_A}$ is their covariance. In this note, I show that analogous formulas involving the data resolution matrix $\mathbf{N}_G \equiv \mathbf{G}\mathbf{G}^{-g}$ hold for the predicted data $\mathbf{d}^{pre}$ and its covariance $\mathbf{C}_{\mathbf{d}^{pre}}$:

$$\mathbf{d}^{pre} = \mathbf{N}_G\mathbf{d}^{obs} + (\mathbf{I} - \mathbf{N}_G)\mathbf{d}_A$$

$$\mathbf{C}_{\mathbf{d}^{pre}} = \mathbf{N}_G\mathbf{C}_d\mathbf{N}_G^T + (\mathbf{I} - \mathbf{N}_G)\mathbf{C}_{\mathbf{d}_A}(\mathbf{I} - \mathbf{N}_G)^T$$

Here, $\mathbf{d}_A \equiv \mathbf{G}\mathbf{m}_A$ are the data implied by the prior information and $\mathbf{C}_{\mathbf{d}_A}$ is their variance.

Derivation

The notation and general approach follow Menke (2014). The only the derivation of the equation relating $\mathbf{d}^{pre}$ and $\mathbf{N}_G$ are new (but the derivation of the equation relating $\mathbf{m}^{est}$ and $\mathbf{R}_G$ is streamlined a bit).

Let $\mathbf{m}$ be a length-M model parameter vector, $\mathbf{d}$ be a length-N data vector, $\mathbf{G}$ be a $N \times M$ data kernel matrix satisfying $\mathbf{Gm} = \mathbf{d}$ with covariance $\mathbf{C}_d$. Furthermore, suppose prior information $\mathbf{Hm} = \mathbf{h}_A$ with covariance $\mathbf{C}_h$ is available. The GLS solution is achieved by solving

$$\mathbf{Fm} = \mathbf{h} \quad \text{with} \quad \mathbf{F} \equiv \begin{bmatrix} \mathbf{C}_d^{-1/2}\mathbf{G} \\ \mathbf{C}_h^{-1/2}\mathbf{H} \end{bmatrix} \quad \text{and} \quad \mathbf{f} \equiv \begin{bmatrix} \mathbf{C}_d^{-1/2}\mathbf{d}^{obs} \\ \mathbf{C}_h^{-1/2}\mathbf{h}_A \end{bmatrix}$$

by least squares

$$\mathbf{m}^{est} \equiv [\mathbf{F}^T\mathbf{F}]^{-1}\mathbf{F}^T\mathbf{f} = [\mathbf{G}^T\mathbf{C}_d^{-1}\mathbf{G} + \mathbf{H}^T\mathbf{C}_h^{-1}\mathbf{H}]^{-1}[\mathbf{G}^T\mathbf{C}_d^{-1}\mathbf{d}^{obs} + \mathbf{H}^T\mathbf{C}_h^{-1}\mathbf{h}_A]$$

Defining $\mathbf{A} \equiv [\mathbf{G}^T\mathbf{C}_d\mathbf{G} + \mathbf{C}_A]$ and generalized inverses $\mathbf{G}^{-g} \equiv \mathbf{A}^{-1}\mathbf{G}^T\mathbf{C}_d^{-1}$ and $\mathbf{H}^{-g} \equiv \mathbf{A}^{-1}\mathbf{H}^T\mathbf{C}_h^{-1}$, the estimated model parameters are

$$\mathbf{m}^{est} = \mathbf{G}^{-g}\mathbf{d}^{obs} + \mathbf{H}^{-g}\mathbf{h}_A$$

Suppose that we define the prior model parameters $\mathbf{m}_A$ as those implied by the prior information acting alone

$$\mathbf{m}_A \equiv [\mathbf{H}^T\mathbf{C}_h^{-1}\mathbf{H}]^{-1}\mathbf{H}^T\mathbf{C}_h^{-1}\mathbf{h}_A$$

where it is understood that $\mathbf{H}$ is to be augmented as needed with very week smallness information, in order to prevent $[\mathbf{H}^T\mathbf{C}_h^{-1}\mathbf{H}]^{-1}$ from being singular. Note that the case $\mathbf{H} = \mathbf{I}$, for which case $\mathbf{h}_A$ are the prior model parameters, leads to no contradiction, as $\mathbf{m}_A = [\mathbf{C}_h^{-1}]^{-1}\mathbf{C}_h^{-1}\mathbf{h}_A = \mathbf{h}_A$. The definition of $\mathbf{m}_A$ implies

$$[\mathbf{H}^T\mathbf{C}_h^{-1}\mathbf{H}]\mathbf{m}_A \equiv \mathbf{H}^T\mathbf{C}_h^{-1}\mathbf{h}_A$$

Substituting this expression into the formula for the estimated model parameters yields

$$\mathbf{m}^{est} = \mathbf{A}^{-1}\mathbf{G}^T\mathbf{C}_d^{-1}\mathbf{d}^{obs} + \mathbf{A}^{-1}\mathbf{H}^T\mathbf{C}_h^{-1}\mathbf{h}_A$$

$$= \mathbf{A}^{-1}\mathbf{G}^T\mathbf{C}_d^{-1}\mathbf{d}^{obs} + \mathbf{A}^{-1}[\mathbf{H}^T\mathbf{C}_h^{-1}\mathbf{H}]\mathbf{m}_A$$

Adding and subtracting $\mathbf{A}^{-1}[\mathbf{G}^T\mathbf{C}_d^{-1}\mathbf{G}]\mathbf{m}_A$ yields

$$\mathbf{m}^{est} = \mathbf{A}^{-1}\mathbf{G}^T\mathbf{C}_d^{-1}\mathbf{d}^{obs} + \mathbf{A}^{-1}[\mathbf{H}^T\mathbf{C}_h^{-1}\mathbf{H}]\mathbf{m}_A + \mathbf{A}^{-1}[\mathbf{G}^T\mathbf{C}_d^{-1}\mathbf{G}]\mathbf{m}_A - \mathbf{A}^{-1}[\mathbf{G}^T\mathbf{C}_d^{-1}\mathbf{G}]\mathbf{m}_A$$

$$= \mathbf{A}^{-1}\mathbf{G}^T\mathbf{C}_d^{-1}\mathbf{d}^{obs} + \mathbf{A}^{-1}[\mathbf{G}^T\mathbf{C}_d^{-1}\mathbf{G} + \mathbf{H}^T\mathbf{C}_h^{-1}\mathbf{H} +]\mathbf{m}_A - \mathbf{A}^{-1}[\mathbf{G}^T\mathbf{C}_d^{-1}\mathbf{G}]\mathbf{m}_A$$

$$= \mathbf{G}^{-g}\mathbf{d}^{obs} + \mathbf{m}_A - \mathbf{G}^{-g}\mathbf{G}\mathbf{m}_A$$

So that

$$\mathbf{m}^{est} = \mathbf{G}^{-g}\mathbf{d}^{obs} + (\mathbf{I} - \mathbf{R}_G)\mathbf{m}_A$$

where the model resolution matrix is defined as

$$\mathbf{R}_G \equiv \mathbf{G}^{-g}\mathbf{G}$$

Thus, the estimated model parameters depend in the prior information only to the degree that $\mathbf{R}_G$ departs from the identity matrix.

The predicted data is defined as

$$\mathbf{d}^{pre} \equiv \mathbf{G}\mathbf{m}^{est} = \mathbf{G}\mathbf{G}^{-g}\mathbf{d}^{obs} + \mathbf{G}(\mathbf{I} - \mathbf{R}_G)\mathbf{m}_A$$

The data implied by the prior information is defines as

$$\mathbf{d}_A \equiv \mathbf{G}\mathbf{m}_A$$

Thus

$$\mathbf{d}^{pre} = \mathbf{G}\mathbf{G}^{-g}\mathbf{d}^{obs} + \mathbf{G}(\mathbf{I} - \mathbf{R}_G)\mathbf{m}_A = \mathbf{G}\mathbf{G}^{-g}\mathbf{d}^{obs} + \mathbf{G}(\mathbf{I} - \mathbf{G}^{-g}\mathbf{G})\mathbf{m}_A$$

So

$$\mathbf{d}^{pre} = \mathbf{N}_G\mathbf{d}^{obs} + (\mathbf{I} - \mathbf{N}_G)\mathbf{d}_A$$

where the data resolution matrix is defined as

$$\mathbf{N}_G \equiv \mathbf{G}\mathbf{G}^{-g}$$

Thus, the predicted data depend on the data implied by the prior information only to the degree that $\mathbf{N}_G$ departs from the identity matrix.

Standard error propagation yields the following covariance formulas:

$$\mathbf{C}_{\mathbf{m}_A} = [\mathbf{H}^T\mathbf{C}_h^{-1}\mathbf{H}]^{-1}\mathbf{H}^T\mathbf{C}_h^{-1}\mathbf{C}_h\mathbf{C}_h^{-1}\mathbf{H}[\mathbf{H}^T\mathbf{C}_h^{-1}\mathbf{H}]^{-1} = [\mathbf{H}^T\mathbf{C}_h^{-1}\mathbf{H}]^{-1}$$

$$\mathbf{C}_{\mathbf{m}^{est}} = \mathbf{G}^{-g}\mathbf{C}_d\mathbf{G}^{-gT} + (\mathbf{I}-\mathbf{R}_G)\mathbf{C}_{m_A}(\mathbf{I}-\mathbf{R}_G)^T$$

$$\mathbf{C}_{\mathbf{d}_A} = \mathbf{G}\mathbf{C}_{\mathbf{m}_A}\mathbf{G}^T$$

$$\mathbf{C}_{\mathbf{d}^{pre}} = \mathbf{N}_G\mathbf{C}_d\mathbf{N}_G^T + (\mathbf{I}-\mathbf{N}_G)\mathbf{C}_{\mathbf{d}_A}(\mathbf{I}-\mathbf{N}_G)^T$$

Reference

Menke, W., Review of the Generalized Least Squares Method, Surveys in Geophysics 36, 1-25, 2014.

The generalized least squares solution is **m**=**G**−g**d** where **G**−g≡[**G**T**G**+ε**H**T**H**]−1**G**T and where ε is a ratio of variances. The resolution matrix is **R**≡**G**−g**G**.