Here's Some Math

Brynnydd Hamilton, Jake Gebbie

January 11, 2023

1 Single model "textbook" solution, no prior

This math follows Section 1.3.4. in Dynamical Insights from Data, and derives the equation behind the "solve_textbook" method in BLUEs.jl. We seek to solve the system

$$\mathbf{y} = \mathbf{E}\mathbf{x} \tag{1}$$

where $\mathbf{y}$ is some vector of observations, $\mathbf{x}$ is some vector of parameters and $\mathbf{E}$ is a model matrix that relates the two. Our observations $\mathbf{y}$ has some associated noise covariance matrix $\mathbf{C_{nn}}$ Here we solve an overdetermined problem, so length($\mathbf{y}$) $<$ length($\mathbf{x}$), and $\mathbf{E}$ has more rows than columns. We define a residual $\mathbf{n}$

$$\mathbf{n} = \mathbf{y} - \mathbf{E}\mathbf{x} \tag{2}$$

We seek to minimize a cost function defined by the weighted inner product of the residual

$$J = \mathbf{n^T C_{nn}^{-1} n} = (\mathbf{y} - \mathbf{E}\mathbf{x})^{\mathbf{T}} \mathbf{C_{nn}^{-1}} (\mathbf{y} - \mathbf{E}\mathbf{x}) \tag{3}$$

To minimize it, we want to minimize J with respect to $\mathbf{x}$. To accomplish this, we take the first partial derivative $\frac{\partial J}{\partial \mathbf{x}}$, set it equal to 0, and solve for our solution estimate $\tilde{\mathbf{x}}$

$$\frac{\partial J}{\partial \mathbf{x}} = \frac{\partial \mathbf{n}}{\partial \mathbf{x}} \frac{\partial J}{\partial \mathbf{n}} \tag{4}$$

$$\frac{\partial J}{\partial \mathbf{n}} = (\mathbf{C_{nn}}^{-1} + \mathbf{C_{nn}}^{-T})\mathbf{n} = 2\mathbf{C_{nn}}^{-1}\mathbf{n} \tag{5}$$

$$\frac{\partial \mathbf{n}}{\partial \mathbf{x}} = -\mathbf{E}^T \tag{6}$$

$$\frac{\partial J}{\partial \mathbf{x}} = \frac{\partial \mathbf{n}}{\partial \mathbf{x}} \frac{\partial J}{\partial \mathbf{n}} = -2\mathbf{E}^T \mathbf{C_{nn}}^{-1}\mathbf{n} = -2\mathbf{E}^T \mathbf{C_{nn}}^{-1}(\mathbf{y} - \mathbf{E}\mathbf{x}) = -2\mathbf{E}^T \mathbf{C_{nn}}^{-1}\mathbf{y} + 2\mathbf{E}^T \mathbf{C_{nn}}^{-1}\mathbf{E}\mathbf{x} \tag{7}$$

We then set the above equation equal to zero, and now solve for the solution estimate $\tilde{\mathbf{x}}$

$$-2\mathbf{E}^T \mathbf{C_{nn}}^{-1}\mathbf{y} + 2\mathbf{E}^T \mathbf{C_{nn}}^{-1}\mathbf{E}\tilde{\mathbf{x}} = 0 \tag{8}$$

$$\tilde{\mathbf{x}} = (\mathbf{E}^T \mathbf{C_{nn}}^{-1}\mathbf{E})^{-1}(\mathbf{E}^T \mathbf{C_{nn}}^{-1}\mathbf{y}) \tag{9}$$

2 Multiple model "textbook" solution, no prior

For our eventual application, we will have multiple, disparate observational vectors $\mathbf{y_i}$, that each have their own associated noise covariance matrix $\mathbf{C_{nn}}_{(i)}$, as well as their own model matrix $\mathbf{E}_i$. However, they will all be associated with the same parameter vector $\mathbf{x}$. Therefore, we are solving the simultaneous system

$$\mathbf{y}_i = \mathbf{E}_i \mathbf{x} \tag{10}$$

For each observational vector, we can define a corresponding noise vector

$$\mathbf{n}_i = \mathbf{y_i} - \mathbf{E}_i \mathbf{x} \tag{11}$$

Each system will have its own cost function, and we want to minimize all of them simultaneously, so we will minimize the sum of the cost functions

$$J = \sum_{i}^{N} \mathbf{n}_i^T \mathbf{C}_{\mathbf{nn}(i)} \mathbf{n}_i \tag{12}$$

As this is a simple sum, the math will remain relatively similar to Equations 4 - 7, with the new version of Equation 7 as follows

$$\frac{\partial J}{\partial \mathbf{x}} = \sum_{i}^{N} -2\mathbf{E}_i^T \mathbf{C}_{\mathbf{nn}(i)}^{-1} \mathbf{y}_i + 2\mathbf{E}_i^T \mathbf{C}_{\mathbf{nn}(i)}^{-1} \mathbf{E}_i \mathbf{x} \tag{13}$$

If we set Equation 13 equal to 0, and solve for $\tilde{\mathbf{x}}$, we would have to calculate

$$\tilde{\mathbf{x}} = \left(\sum_{i}^{N} \mathbf{E}_i^T \mathbf{C}_{\mathbf{nn}(i)}^{-1} \mathbf{E}_i \right)^{-1} \left(\sum_{i}^{N} \mathbf{E}_i^T \mathbf{C}_{\mathbf{nn}(i)}^{-1} \mathbf{y}_i \right) \tag{14}$$