Damped Least Squares has a Symmetric Resolution Kernel

Bill Menke, February 23, 2018

When I first encountered a symmetric resolution kernel (that is, $\mathbf{R} = \mathbf{R}^{\mathrm{T}}$) in a demo I was preparing for class, I thought that it was something specific to the particular data kernel that I was using. But it turns out that it's a general property of damped least squares. We start with the damped least squares generalized inverse $\mathbf{G}^{-g}$, the definition of the resolution matrix $\mathbf{R}$ and the eigenvalue expansion of the symmetric matrix $\mathbf{G}^{\mathrm{T}}\mathbf{G}$:

$$\mathbf{G}^{-g} = [\mathbf{G}^{\mathrm{T}}\mathbf{G} + \varepsilon^2\mathbf{I}]^{-1}\mathbf{G}^{\mathrm{T}} \quad \text{and} \quad \mathbf{R} = \mathbf{G}^{-g}\mathbf{G} = [\mathbf{G}^{\mathrm{T}}\mathbf{G} + \varepsilon^2\mathbf{I}]^{-1}[\mathbf{G}^{\mathrm{T}}\mathbf{G}] \quad \text{and} \quad [\mathbf{G}^{\mathrm{T}}\mathbf{G}] = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{\mathrm{T}}$$

We then establish that $\mathbf{R}$ is symmetric, by simultaneously diagonalizing $[\mathbf{G}^{\mathrm{T}}\mathbf{G} + \varepsilon^2\mathbf{I}]^{-1}$ and $\mathbf{G}^{\mathrm{T}}\mathbf{G}$, and using the fact that matrices commute if and only if they are simultaneously diagonizable.

$$\mathbf{R}^{\mathrm{T}} = \mathbf{R}$$

$$[\mathbf{G}^{\mathrm{T}}\mathbf{G}][\mathbf{G}^{\mathrm{T}}\mathbf{G} + \varepsilon^2\mathbf{I}]^{-1} = [\mathbf{G}^{\mathrm{T}}\mathbf{G} + \varepsilon^2\mathbf{I}]^{-1}[\mathbf{G}^{\mathrm{T}}\mathbf{G}]$$

$$[\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{\mathrm{T}}][\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{\mathrm{T}} + \varepsilon^2\mathbf{V}\mathbf{I}\mathbf{V}^{\mathrm{T}}]^{-1} = [\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{\mathrm{T}} + \varepsilon^2\mathbf{V}\mathbf{I}\mathbf{V}^{\mathrm{T}}]^{-1}[\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{\mathrm{T}}]$$

$$[\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{\mathrm{T}}][\mathbf{V}(\boldsymbol{\Lambda} + \varepsilon^2\mathbf{I})\mathbf{V}^{\mathrm{T}}]^{-1} = [\mathbf{V}(\boldsymbol{\Lambda} + \varepsilon^2\mathbf{I})\mathbf{V}^{\mathrm{T}}]^{-1}[\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{\mathrm{T}}]$$

$$[\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{\mathrm{T}}][\mathbf{V}(\boldsymbol{\Lambda} + \varepsilon^2\mathbf{I})^{-1}\mathbf{V}^{\mathrm{T}}] = [\mathbf{V}(\boldsymbol{\Lambda} + \varepsilon^2\mathbf{I})^{-1}\mathbf{V}^{\mathrm{T}}][\mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{\mathrm{T}}]$$

$$[\mathbf{V}][\boldsymbol{\Lambda}(\boldsymbol{\Lambda} + \varepsilon^2\mathbf{I})^{-1}]\mathbf{V}^{\mathrm{T}} = \mathbf{V}[(\boldsymbol{\Lambda} + \varepsilon^2\mathbf{I})^{-1}\boldsymbol{\Lambda}]\mathbf{V}^{\mathrm{T}}$$

$$\boldsymbol{\Lambda}(\boldsymbol{\Lambda} + \varepsilon^2\mathbf{I})^{-1} = (\boldsymbol{\Lambda} + \varepsilon^2\mathbf{I})^{-1}\boldsymbol{\Lambda}$$

The last equation is symmetric since involves only the product of two diagonal matrices, and all diagonal matrices commute with each other.

Note that the resolution kernel of Generalized Least Squares (where $\mathbf{I}$ in the generalized inverse is replaced with a symmetric weight matrix $\mathbf{W}_{\mathrm{m}} \equiv \mathbf{D}^{\mathrm{T}}\mathbf{D}$) is *not* symmetric, since, in general, $\mathbf{W}_{\mathrm{m}}$ and $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ cannot be simultaneously diagonalized. However, as I have pointed out elsewhere[1], if both $\mathbf{G}$ and $\mathbf{D}$ correspond to convolutions, then $\mathbf{R}$ is approximately symmetric, since convolutions are represented by Toeplitz matrices, and Toeplitz matrices asymptotically commute.

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