Damped Least Squares Data and Model Resolution Equal for Symmetric Data Kernel
February 19, 2024

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 uniform, uncorrelated covariance \( \mathbf{C}_d \). Furthermore, suppose the prior information \( \mathbf{m} = \mathbf{0} \) with uniform, uncorrelated covariance \( \mathbf{C}_A = \sigma_A^2 \mathbf{I} \). Here, \( \sigma_d^2 \) and \( \sigma_A^2 \) are the variance of the observations and prior information, respectively. The Generalized least squares solution is achieved by solving

\[
\mathbf{Fm} = \mathbf{h} \quad \text{with} \quad \mathbf{F} = \begin{bmatrix} \sigma_d \mathbf{G} \\ \sigma_A \mathbf{I} \end{bmatrix} \quad \text{and} \quad \mathbf{f} = \begin{bmatrix} \sigma_d \mathbf{d}^{\text{obs}} \\ \mathbf{0} \end{bmatrix}
\]

by least squares. The result is called damped least squares, with damping parameter \( \varepsilon \)

\[
\mathbf{m}^{\text{est}} = \left[ \mathbf{F}^T \mathbf{F} \right]^{-1} \mathbf{F}^T \mathbf{f} = \left[ \mathbf{G}^T \mathbf{G} + \varepsilon^2 \mathbf{I} \right]^{-1} \mathbf{G}^T \mathbf{d}^{\text{obs}} = \mathbf{G}^{-g} \mathbf{d}^{\text{obs}} \quad \text{with} \quad \varepsilon = \frac{\sigma_d}{\sigma_A}
\]

Here, \( \mathbf{G}^{-g} \) is the generalized inverse. The data and model resolution matrices are defined as

\[
\mathbf{N}_G \equiv \mathbf{GG}^{-g} \quad \text{and} \quad \mathbf{R}_G \equiv \mathbf{G}^{-g} \mathbf{G}
\]

Now suppose that \( \mathbf{G} \) has singular value composition

\[
\mathbf{G} = \mathbf{UAV}^T
\]

with \( \mathbf{A} \) a diagonal matrix and \( \mathbf{U} \) and \( \mathbf{V} \) unary matrices (i.e., \( \mathbf{V}^{-1} = \mathbf{V}^T \) and \( \mathbf{U}^{-1} = \mathbf{U}^T \)). Then

\[
\mathbf{G}^T \mathbf{G} = \mathbf{VAU}^T \mathbf{UAV}^T = \mathbf{VA}^2 \mathbf{V}^T \quad \text{and} \quad \mathbf{I} = \mathbf{VV}^T = \mathbf{V}^T \mathbf{V}
\]

The data resolution is

\[
\mathbf{N}_G = \mathbf{G} \left[ \mathbf{G}^T \mathbf{G} + \varepsilon^2 \mathbf{I} \right]^{-1} \mathbf{G}^T = \mathbf{UA} \left[ \mathbf{VA}^2 \mathbf{V}^T + \varepsilon^2 \mathbf{VV}^T \right]^{-1} \mathbf{VAU}^T = \mathbf{UA} \left[ \mathbf{A}^2 + \varepsilon^2 \mathbf{I} \right]^{-1} \mathbf{AU}^T = \mathbf{U} \left[ \mathbf{A}^2 + \varepsilon^2 \mathbf{I} \right]^{-1} \mathbf{A}^2 \mathbf{U}^T
\]

Here, we have used the facts that for any invertible matrix \( \mathbf{M} \), \( \left( \mathbf{MVV}^T \right)^{-1} = \mathbf{VM}^{-1} \mathbf{V}^T \) (as can be seen from \( \mathbf{VMV}^T \mathbf{VM}^{-1} \mathbf{V}^T = \mathbf{VM} \mathbf{M}^{-1} \mathbf{V}^T = \mathbf{VV}^T \mathbf{V} \)) and that diagonal matrices commute. The model resolution is

\[
\mathbf{R}_G = \left[ \mathbf{G}^T \mathbf{G} + \varepsilon^2 \mathbf{I} \right]^{-1} \mathbf{G}^T \mathbf{G} = \left[ \mathbf{VA}^2 \mathbf{V}^T + \mathbf{V} \varepsilon^2 \mathbf{IV}^T \right]^{-1} \mathbf{VA}^2 \mathbf{V}^T = \mathbf{V} \left[ \mathbf{A}^2 + \varepsilon^2 \mathbf{I} \right]^{-1} \mathbf{A}^2 \mathbf{V}^T
\]

We note that both \( \mathbf{N}_G \) and \( \mathbf{R}_G \) are symmetric, as \( \left[ \mathbf{M}^T \mathbf{SM} \right]^T = \mathbf{M}^T \mathbf{SM} \) for any symmetric matrix \( \mathbf{S} \) (and diagonal matrices are symmetric). Furthermore, \( \mathbf{U} = \mathbf{V} \) when \( \mathbf{G} \) is symmetric, so the conditions that \( \mathbf{N}_G = \mathbf{R}_G \) are that \( N = M \) and \( \mathbf{G} \) is symmetric.

Reference