Resolution of the Data Smoothing Problem

Bill Menke, February 5, 2014

In this note, we consider model parameters $\mathbf{m}$ and data $\mathbf{d}$ that are a discrete version of a continuous functions $m(x)$ and $d(x)$, say with spacing Δx. The data equation $\mathbf{Gm} = \mathbf{d}$ has unit variance and the prior constraint equation $\mathbf{Hm} = \mathbf{h} = \mathbf{0}$ has variance $\sigma^2 = \varepsilon^{-2}$. The Generalized Least Squares solution is:

$$\mathbf{m} = [\mathbf{G}^{\mathrm{T}}\mathbf{G} + \varepsilon^2\mathbf{H}^{\mathrm{T}}\mathbf{H}]^{-\mathbf{1}}\mathbf{G}^{\mathrm{T}}\mathbf{d} = \mathbf{G}^{-\mathrm{g}}\mathbf{d}$$

where $\mathbf{G}^{-\mathrm{g}}$ is the generalized inverse. The corresponding resolution matrix is

$$\mathbf{R} = \mathbf{G}^{-\mathrm{g}}\mathbf{G} = [\mathbf{G}^{\mathrm{T}}\mathbf{G} + \varepsilon^2\mathbf{H}^{\mathrm{T}}\mathbf{H}]^{-\mathbf{1}}\left(\mathbf{G}^{\mathrm{T}}\mathbf{G}\right)$$

Note that $\mathbf{R}$ is a symmetric matrix and that $\mathbf{R} = \mathbf{I}$ (perfect resolution) when $\varepsilon^2 = 0$. Now consider the case where $\mathbf{G} = \mathbf{I}$ (that is the data are direct estimates of the model parameters) and $\mathbf{H}$ is a second derivative operator, with rows like:

$$(\Delta x)^{-2}[0 \quad \cdots \quad 0 \quad 1 \quad -2 \quad 1 \quad 0 \quad \cdots \quad 0]$$

This inverse problem can be understood as finding model parameters that are approximately equal to the data, but which are smoother (that is, have smaller second derivative). The degree of smoothing increases with ε.

When the smoothness constraint is weak ($(\varepsilon/\Delta x)^2 \ll 1$), we can expand the resolution matrix in a Taylor Series, keeping only the first two terms:

$$\mathbf{R} = \mathbf{G}^{-\mathrm{g}}\mathbf{I} = [\mathbf{I} + \varepsilon^2\mathbf{H}^{\mathrm{T}}\mathbf{H}]^{-\mathbf{1}}\mathbf{II} \approx \mathbf{I} - \varepsilon^2\mathbf{H}^{\mathrm{T}}\mathbf{H}$$

This matrix has rows like:

$$[0 \quad \cdots \quad 0 \quad -(\varepsilon\Delta x)^{-2} \quad 4(\varepsilon\Delta x)^{-2} \quad (1-6(\varepsilon\Delta x)^{-2}) \quad 4(\varepsilon\Delta x)^{-2} \quad -(\varepsilon\Delta x)^{-2} \quad 0 \quad \cdots \quad 0]$$

In the absence of the constraint ($\varepsilon = 0$), the resolution is perfect; that is:

$$[0 \quad \cdots \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad \cdots \quad 0]$$

As the strength of the constraint is increased, the magnitude of the central value decreases and the nearest neighbor values become positive. For example, when $(\varepsilon\Delta x)^{-2} = 0.01$ we have:

$$[0 \quad \cdots \quad 0 \quad -0.01 \quad 0.04 \quad 0.94 \quad 0.04 \quad -0.01 \quad 0 \quad \cdots \quad 0]$$

The spread of the resolution has increased. However, the outermost non-zero values are negative, indicating that the smoothing cannot be interpreted as a weighted average in the normal sense.

The quantity $\mathbf{r} = \mathbf{Rm}_0$ represents a column of resolution matrix when $\mathbf{m}_0$ is set to a column of the identity matrix (say a column corresponding to position x_0), and the corresponding row is just $\mathbf{r}^{\mathrm{T}}$, since $\mathbf{R}$ is symmetric. But since $\mathbf{G} = \mathbf{I}$, the corresponding data $\mathbf{d}_0 = \mathbf{Gm}_0$ is also the same column of the identity matrix. Hence,

$$\mathbf{r} = \mathbf{Rm}_0 = \mathbf{G}^{-\mathbf{g}}(\mathbf{Gm}_0) = \mathbf{G}^{-\mathbf{g}}\mathbf{d}_0 = [\mathbf{I} + \varepsilon^2\mathbf{H}^{\mathrm{T}}\mathbf{H}]^{-\mathbf{1}}\mathbf{d}_0$$

Moving the matrix inverse to the l.h.s. of the equation yields:

$$\varepsilon^2\mathbf{H}^{\mathrm{T}}\mathbf{Hr} + \mathbf{r} = \mathbf{d}_0$$

Note that the second derivative operator is symmetric, so that $\mathbf{H}^{\mathrm{T}}\mathbf{H} = \mathbf{HH}$; that is, a second derivative operator applied twice to yield the fourth derivative operator. Except for the first and last row, where edge effects are important, the matrix equation is the discrete analogue to the differential equation:

$$\varepsilon^2 \frac{d^4 r}{dx^4} + r = \Delta x\, \delta(x - x_0)$$

The factor of Δx has been added so that the area under $\Delta x\, \delta(x - x_0)$ is the same as the area under $\mathbf{d}_0$. This well-known differential equation has solution:

$$r(x) = A \exp(-|x - x_0|/a)\,\{\cos(|x - x_0|/a) + \sin(|x - x_0|/a)\}$$

with

$$A = \frac{\Delta x a^3}{8\varepsilon^2} \text{ and } a = (2\varepsilon)^{1/2}$$

The differential equation can be interpreted as the deflection $r(x)$ of a elastic beam of flexural rigidity ε^2 floating on a fluid foundation, due to a point load at x_0. The beam will take on a shape that exactly mimics the load only in the case when it has no rigidity; that is, $\varepsilon^2 = 0$. For any finite rigidity, the beam will take on a shape that is a smoothed version of the load, where the amount of smoothing increases with ε^2. The parameter $a = (2\varepsilon)^{1/2}$ gives the scale length over which the smoothing occurs. Note that the formula for $r(x)$ contains oscillating trigonometric functions, so that the smoothing does not correspond to a weighted average in the usual sense; some weights are negative.

Example:

N=101 data and M=101 model parameters
$0 < x < 10$
Δx=0.1
$\varepsilon = 0.1$

$\mathbf{G} = \mathbf{I}$

$\mathbf{H}$ with rows $(\Delta x)^{-2}[0 \quad \cdots \quad 0 \quad 1 \quad -2 \quad 1 \quad 0 \quad \cdots \quad 0]$

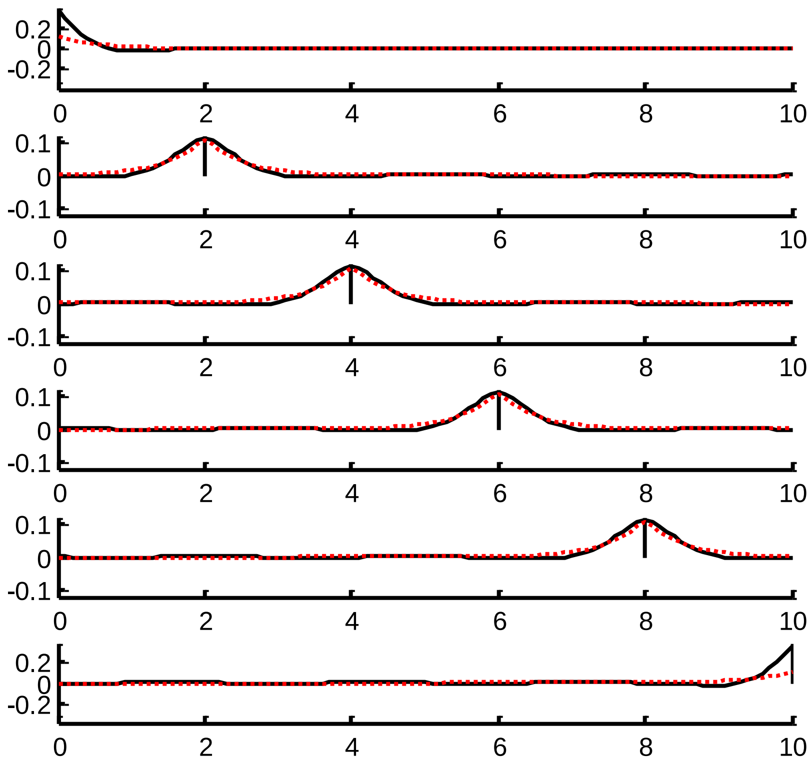

Figure: Selected rows of the resolution matrix as a function of x, as determined by the exact formula for the generalized inverse (black curves) and as determined by the flexural approximation (red curves).