Effect of Choice of Covariance Matrix on Curve Fitting

Bill Menke, with Zach Eilon

May 6, 2014

Question: In curve fitting, what is the difference between (A) minizing the second derivative and (B) enforcing an exponentially-decaying autocorrelation function.

Answer: Not much.

Part 1: Differences between three choices of H

Linear inverse problem for smoothing data, where data $\mathbf{d}$ and model parameters $\mathbf{m}$ are observed and true versions of a discretized function $m(x)$. The generalized least squares equation is:

$$\begin{bmatrix} \mathbf{I} \\ \mathbf{C}_\mathbf{m}^{-1/2}\mathbf{H} \end{bmatrix} \mathbf{m} = \begin{bmatrix} \mathbf{d} \\ 0 \end{bmatrix}$$

Green Case:

$\mathbf{H} = \mathbf{D}$ is second derivative operator and $\mathbf{C}_\mathbf{m}^{-1/2} = \varepsilon^2 \mathbf{I}$, with damping ε^2.

Interpretation: $\mathbf{m}$ has small curvature

Blue Case:

$\mathbf{H} = \mathbf{I}$ and $\mathbf{C}_\mathbf{m}^{-1/2}$ such that $C_{ij} = \varepsilon^2 \exp\left(-|i-j|/s\right)$, with s as scale factor.

Interpretation: Prior information has an decating-exponential autocorrelation function.

Red Case:

$$\mathbf{H} = \mathbf{S} - \mathrm{diag}(\mathbf{a})$$

$$S_{ij} = \exp\left(-|i-j|/s\right)$$

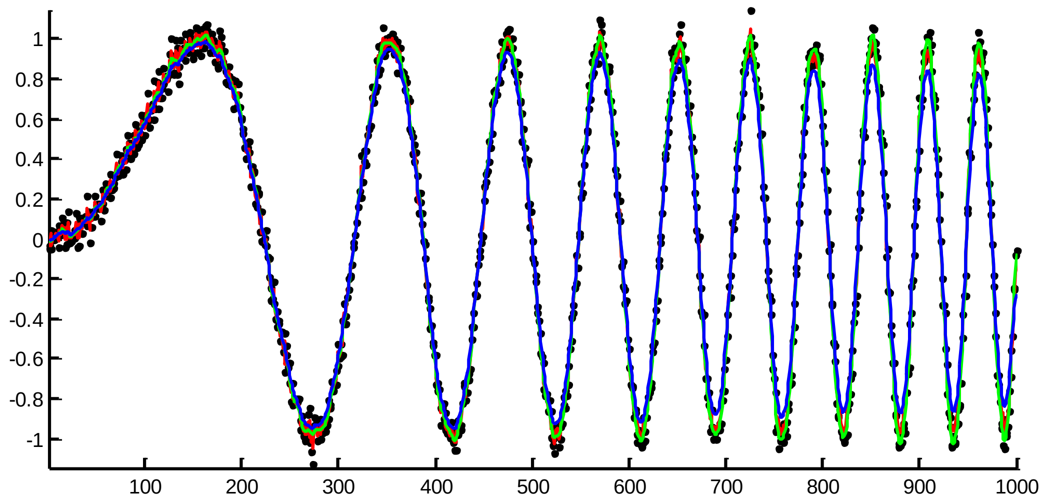
$$a_i = \sum_j S_{ij}$$

$$\mathbf{C}_\mathbf{m}^{-1/2} = \varepsilon^2 \mathbf{I}$$

Interpretation: $\mathrm{m_i}$ is close to its localized average, centered at i.

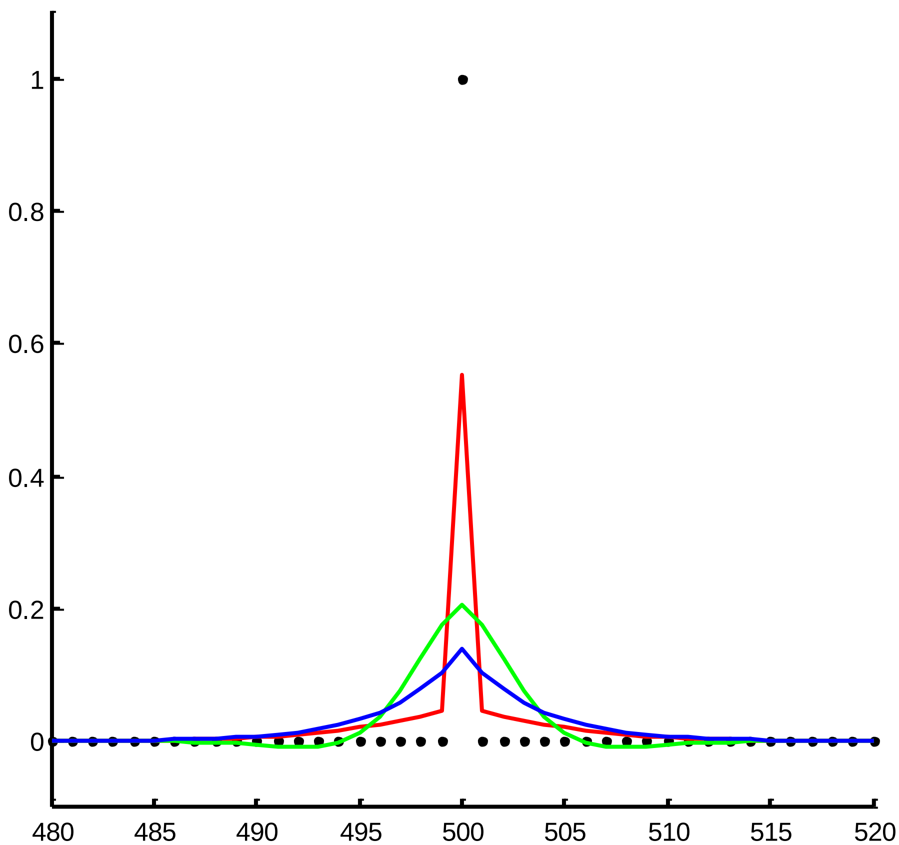
In all cases, I have adjusted the damping ε^2 and scale factor s sensibly.

Sample fit



Interpretation: not much difference between curves when damping & scale s are chosen propely

Impulse respomse



Notes:

Blue case: Impulse response is itself two-sided exponential. I have been able to prove this in the continuum limit. However, the scale length is not equal to s and the area under the curve is not unity.

Green Case has a bit of overshoot, which follows from the continuum limit being an :elastic plate over fluid foundation" problem.

Red Case has prominent central peak, which would seem undesirable.

Part 2: Shape of weighting function corresponding to exponential autocorrelation function

Black Case:

$\mathbf{C_m^{-1/2}H}$ with $\mathbf{H} = \mathbf{I}$ and $C_{ij} = \exp\left(-|i-j|/s\right)$, with s a scale factor. Note: I have worked out $\mathbf{C_m^{-1}}$ in the continuum limit; it corresponds to the linear operator $(1 - s^{-2}\, d^2/dx^2)$. However, I have not been able to compute the square root of this operator.

Red Case:

$\mathbf{C_m^{-1/2}H}$ with $\mathbf{C_m^{-1/2}} = \mathbf{I}$ and $\mathbf{H} = \mathbf{D}$, a second derivative operator

Figure shows result for scale factor $s = 25$.

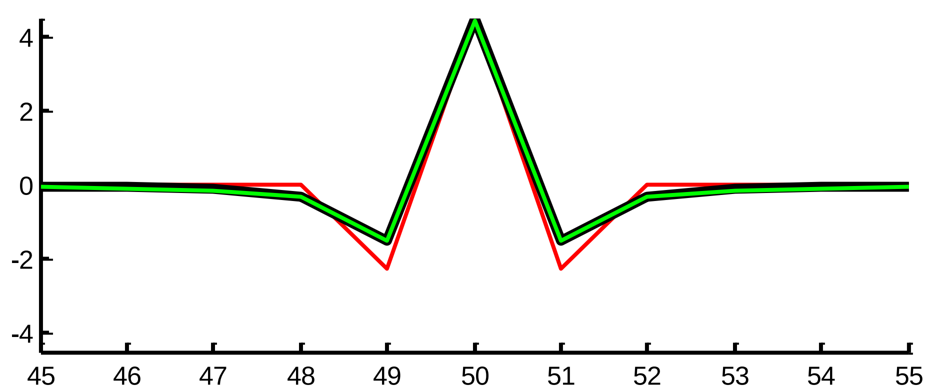
Green Case: Taylor series expansion of Black Case

$$\mathbf{M} = \mathbf{C_m^{-1/2}H} = m_1\mathbf{I} + m_2\mathbf{D} + m_3\mathbf{DD} + \cdots$$

(But note that constructing an $2n$-th order derivative by multiplying n second-order derivatives is not quite right, since it has edge effects). Then, determine the m's by least squares. The normlized error of the fit of E/E0= 0.07 when 4 terms are included, down from 0.16 when only the identity matrix and second derivative is included. The 10-term approximation (green curve, above) is excellent. Increasing the number of terms to 100 reduced the error only to 0.03, so perhaps the edge effects are leading to an imperfect fit.

Interpretation: $\mathbf{M}$ can be built up from sequence of derivatives. Second derivative makes a significant contribution but the fit is significantly improved when higher (but even-order) derivatives are included.

Note:



Interpretation: Black Case is similar in shape to a second derivitive operator, except the central down-up-down does not have exactly the ratio of $[-1 \quad 2 \quad 1]$ and also the function has leading a trailing tails. Increasing the scale length widens these tails slightly, but doesn't affect the central region much.