STL Quadratic Interpolation Implementation Notes

James A. Crotinger
email: jim.crotinger@servicenow.com

April 14, 2017

Implementation

The algorithm used by `stl-decomp-4j` assumes that the input data is regularly spaced with no missing values, allowing for a very efficient implementation of the underlying Loess smoother. Part of this formulation is explained here in order to extend the formulation to quadratic interpolation.

Local Linear Interpolation

The local (weighted) linear interpolation in `LinearLoessInterpolator` is a straight port of the code in the original Ratfor function `stl.r:est`. The input data is a sequence of data points $\{x_i, y_i\}$, where the x_i are the regularly spaced grid points. The Loess interpolation of the data set at an arbitrary point x can be expressed as

$$y(x) = \sum_{i=1}^{m} \hat{w}_i(x) y_i \tag{1}$$

i.e. the interpolation can be re-cast as a linear operation on the input y-values. The weights $\hat{w}_i(x)$ depend only on the original weights, w_i, and on geometric factors.

For linear interpolation, we desire coefficients α and β such that the line

$$y(x) = \alpha + \beta x \tag{2}$$

is the best fit to our set of points. The square error from a weighted least-squares fit of this curve to the training data is

$$E = \frac{1}{2} \sum_{i=1}^{m} (y_i - \alpha - \beta x_i)^2 \cdot w_i \tag{3}$$

where $\sum w_i = 1$ are external weights (in Loess these come from the implementation of the locality window).

Finding α

The optimal choices of α and β are found by differentiating Eq. (3) with respect to each of these, setting to zero and solving:

$$\frac{\partial E}{\partial \alpha} = -\sum_{i=1}^{m} w_i (y_i - \alpha - \beta x_i) = 0 \tag{4}$$

Then

$$\sum w_i y_i - \alpha \sum w_i - \beta \sum w_i x_i = 0 \tag{5}$$

For a given sequence z_i we define

$$\langle z \rangle \equiv \sum_i w_i z_i \tag{6}$$

Then we can rewrite the Eq. (5) as

$$\alpha = \langle y \rangle - \beta \langle x \rangle \tag{7}$$

Finding β

Repeating this exercise for β, skipping the intermediate details, gives

$$\beta = \frac{\langle xy \rangle - \langle x \rangle \langle y \rangle}{\langle x^2 \rangle - \langle x \rangle^2} \tag{8}$$

The Weight

Given Eqs. (7) and (8), Eq. (2) becomes

$$\begin{aligned} y(x) &= \alpha + \beta x & (9) \\ &= \langle y \rangle + \beta(x - \langle x \rangle) & (10) \\ &= \langle y \rangle + \frac{\langle xy \rangle - \langle x \rangle \langle y \rangle}{\langle x^2 \rangle - \langle x \rangle^2}(x - \langle x \rangle) & (11) \end{aligned}$$

Writing out the averages that involve y, we have:

$$\begin{aligned} y(x) &= \sum_j w_j y_j + \sum_j \frac{x - \langle x \rangle}{\langle x^2 \rangle - \langle x \rangle^2} w_j (x_j - \langle x \rangle) y_j & (12) \\ &= \sum_j w_j \left[1 + \frac{x - \langle x \rangle}{\langle x^2 \rangle - \langle x \rangle^2}(x_j - \langle x \rangle) \right] y_j & (13) \\ &= \sum_j \hat{w}_j(x) y_j & (14) \end{aligned}$$

where:

$$\hat{w}_j(x) \equiv w_j \left[1 + \frac{(x - \langle x \rangle)(x_j - \langle x \rangle)}{\langle x^2 \rangle - \langle x \rangle^2} \right] \tag{15}$$

So, given the point x at which we want to perform the interpolation (or extrapolation - x is not limited to being in the range of the set of grid points, $\{x_i\}$), we just calculate a geometric adjustment to the original weights, w_i. These various averages can be computed efficiently since the weights w_i are non-zero only in the Loess window near the point x.

Local Quadratic Interpolation

As before, given data points $\{x_i, y_i\}, i = 1, ..., m$, and externally supplied weights $w_i, \sum_i w_i = 1$, we want to find a set of modified weights $\hat{w}_i$ such that the interpolation at a value x can be written as

$$y(x) = \sum_{i=1}^{m} \hat{w}_i(x) y_i \tag{16}$$

Now we model the data as a local quadratic:

$$y(x) = a_0 + a_1 x + a_2 x^2 \tag{17}$$

Finding a_0

The square error in the local interpolation of the training data is:

$$E = \frac{1}{2}\sum_i w_i(y_i - a_0 - a_1 x_i - a_2 x_i^2) \tag{18}$$

Minimizing with respect to a_0,

$$\frac{\partial E}{\partial a_0} = -\sum_i w_i(y_i - a_0 - a_1 x_i - a_2 x_i^2) = 0 \tag{19}$$

This yields the obvious extension of the linear result:

$$a_0 = \langle y \rangle - a_1\langle x \rangle - a_2\langle x^2 \rangle \tag{20}$$

where the averages $\langle \cdot \rangle$ are defined as before.

Finding a_1

Similarly, for a_1 we have

$$\frac{\partial E}{\partial a_1} = -\sum_i w_i x_i(y_i - a_0 - a_1 x_i - a_2 x_i^2) = 0 \tag{21}$$

$$= -\langle xy \rangle + a_0\langle x \rangle + a_1\langle x^2 \rangle + a_2\langle x^3 \rangle = 0 \tag{22}$$

Solving for $\langle xy \rangle$ and substituting a_0 from Eq. (20):

$$\langle xy \rangle = \left[\langle y \rangle - a_1\langle x \rangle - a_2\langle x^2 \rangle\right]\langle x \rangle + a_1\langle x^2 \rangle + a_2\langle x^3 \rangle \tag{23}$$

$$= \langle x \rangle\langle y \rangle - a_1\langle x \rangle^2 - a_2\langle x \rangle\langle x^2 \rangle + a_1\langle x^2 \rangle + a_2\langle x^3 \rangle \tag{24}$$

Gathering terms

$$\langle xy \rangle - \langle x \rangle\langle y \rangle = a_1(\langle x^2 \rangle - \langle x \rangle^2) + a_2(\langle x^3 \rangle - \langle x^2 \rangle\langle x \rangle) \tag{25}$$

We define the following geometric factors M_2 and M_3, and a correlation factor C_{xy}:

$$M_2 = \langle x^2 \rangle - \langle x \rangle^2 \tag{26}$$

$$M_3 = \langle x^3 \rangle - \langle x^2 \rangle\langle x \rangle \tag{27}$$

$$C_{xy} = \langle xy \rangle - \langle x \rangle\langle y \rangle \tag{28}$$

Then Eq. (25) can be rewritten as

$$C_{xy} = a_1 M_2 + a_2 M_3 \tag{29}$$

Solving for a_1

$$a_1 = \frac{C_{xy}}{M_2} - a_2\frac{M_3}{M_2} \tag{30}$$

Finding a_2

The result of minimizing Eq. (18) with respect to α_2 results in an expression similar to Eq. (25), just adding another x in the appropriate averages, leading to

$$\langle x^2 y\rangle - \langle x^2\rangle\langle y\rangle = a_1(\langle x^3\rangle - \langle x^2\rangle\langle x\rangle) + a_2(\langle x^4\rangle - \langle x^2\rangle^2) \tag{31}$$

We define the following geometric and correlation factors:

$$M_4 = \langle x^4\rangle - \langle x^2\rangle^2 \tag{32}$$
$$C_{x^2y} = \langle x^2 y\rangle - \langle x^2\rangle\langle y\rangle \tag{33}$$

Using Eq. (27) and Eqs. (32-33), Eq. (31) becomes

$$C_{x^2y} = a_1 M_3 + a_2 M_4$$

Substituting our expression for a_1 from Eq. (30),

$$C_{x^2y} = \left[\frac{C_{xy}}{M_2} - a_2\frac{M_3}{M_2}\right] M_3 + a_2 M_4 \tag{34}$$
$$= \frac{M_3}{M_2}C_{xy} + a_2(M_4 - \frac{M_3^2}{M_2}) \tag{35}$$

Solving for a_2 and simplifying

$$a_2 = \frac{C_{x^2y} - \frac{M_3}{M_2}C_{xy}}{M_4 - \frac{M_3^2}{M_2}} \tag{36}$$
$$= \frac{M_2 C_{x^2y} - M_3 C_{xy}}{M_2 M_4 - M_3^2} \tag{37}$$

Substituting Eq. (37) back into Eq. (30) gives

$$\begin{aligned} a_1 &= \frac{C_{xy}}{M_2} - \frac{M_2 C_{x^2y} - M_3 C_{xy}}{M_2 M_4 - M_3^2}\frac{M_3}{M_2} \\ &= \frac{C_{xy}}{M_2} - \frac{M_2 C_{x^2y}}{M_2 M_4 - M_3^2}\frac{M_3}{M_2} + \frac{M_3^2}{M_2 M_4 - M_3^2}\frac{C_{xy}}{M_2} \\ &= \frac{C_{xy}}{M_2}\frac{M_2 M_4 - M_3^2}{M_2 M_4 - M_3^2} + \frac{M_3^2}{M_2 M_4 - M_3^2}\frac{C_{xy}}{M_2} - \frac{M_3 C_{x^2y}}{M_2 M_4 - M_3^2} \end{aligned}$$

So our final expression for a_1 is

$$a_1 = \frac{M_4 C_{xy} - M_3 C_{x^2y}}{M_2 M_4 - M_3^2} \tag{38}$$

The Weight

Returning to the interpolation expression, Eq. (17), we start by substituting back a_0:

$$y(x) = a_0 + a_1 x + a_2 x^2 \tag{39}$$
$$= \langle y\rangle - a_1\langle x\rangle - a_2\langle x^2\rangle + a_1 x + a_2 x^2 \tag{40}$$
$$= \langle y\rangle + a_1(x - \langle x\rangle) + a_2(x^2 - \langle x^2\rangle) \tag{41}$$

Defining the following geometric terms

$$\beta_2 = \frac{M_4}{M_2 M_4 - M_3^2} \tag{42}$$

$$\beta_3 = \frac{M_3}{M_2 M_4 - M_3^2} \tag{43}$$

$$\beta_4 = \frac{M_2}{M_2 M_4 - M_3^2} \tag{44}$$

We can rewrite Eqs. (37) and (38) as

$$a_1 = \beta_2 C_{xy} - \beta_3 C_{x^2y} \tag{45}$$

$$a_2 = \beta_4 C_{x^2y} - \beta_3 C_{xy} \tag{46}$$

Then Eq. (41) becomes

$$y(x) = \langle y \rangle + (x - \langle x \rangle)(\beta_2 C_{xy} - \beta_3 C_{x^2y}) + (x^2 - \langle x^2 \rangle)(\beta_4 C_{x^2y} - \beta_3 C_{xy}) \tag{47}$$

$$= \langle y \rangle + [\beta_2(x - \langle x \rangle) - \beta_3(x^2 - \langle x^2 \rangle)]C_{xy} + [\beta_4(x^2 - \langle x^2 \rangle) - \beta_3(x - \langle x \rangle)]C_{x^2y} \tag{48}$$

Defining the following functions of x

$$\hat{a}_1(x) \equiv \beta_2(x - \langle x \rangle) - \beta_3(x^2 - \langle x^2 \rangle) \tag{49}$$

$$\hat{a}_2(x) \equiv \beta_4(x^2 - \langle x^2 \rangle) - \beta_3(x - \langle x \rangle) \tag{50}$$

we can write, using the definitions of C_{xy} and C_{x^2y}, Eqs. (28) and (33):

$$y(x) = \langle y \rangle + \hat{a}_1(x) C_{xy} + \hat{a}_2(x) C_{x^2y} \tag{51}$$

$$= \sum_i w_i y_i + \hat{a}_1(x) \sum_i w_i (x_i - \langle x \rangle) y_i + \hat{a}_2(x) \sum_i w_i (x_i^2 - \langle x^2 \rangle) y_i \tag{52}$$

$$= \sum_i w_i y_i \left[1 + \hat{a}_1(x)(x_i - \langle x \rangle) + \hat{a}_2(x)(x_i^2 - \langle x^2 \rangle)\right] \tag{53}$$

So finally we arrive at our goal

$$y(x) = \sum_i \hat{w}_i(x) y_i \tag{54}$$

where

$$\hat{w}_i(x) \equiv w_i \left[1 + \hat{a}_1(x)(x_i - \langle x \rangle) + \hat{a}_2(x)(x_i^2 - \langle x^2 \rangle)\right] \tag{55}$$

The code in `QuadraticLoessInterpolator.updateWeights` in `LoessInterpolator.java` is very close to a literal transcription of the above math.