STL Quadratic Interpolation Implementation Notes

James A. Crotinger

email: `jim.crotinger@servicenow.com`

April 8, 2017

Implementation

The algorithm used by `stl-decomp-4j` takes advantage of the fact that the data that is being decomposed is regularly spaced with no missing values to implement the underlying Loess smoothers in an efficient fashion. This document walks through the formulation of local linear interpolation used by the Loess interpolator to explain the flavor of the implementation, and then extends this analysis to the quadratic interpolation that was added to the Java implementation.

Local Linear Interpolation

This section walks through the derivation of the code used to perform local (weighted) linear interpolation in `LinearLoessInterpolator`, which 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 goal is to derive a set of weights w_i such that the Loess interpolation of the data set at an arbitrary point x can be written as

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

i.e. the interpolation is re-cast as a linear operation on the input y-values.

We can write a measure of the square error from weighted least-squares interpolation as

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

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 with respect to these, setting to 0 and solving:

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

Then

$$\sum w_i y_i - \alpha \sum w_i - \beta \sum w_i x_i = 0$$

Or, defining $\langle y \rangle = \sum w_i y_i$, $\langle x \rangle = \sum w_i x_i$ and recalling that the w_i are normalized, this can be rewritten as:

$$\alpha = \langle y \rangle - \beta \langle x \rangle$$

Finding β

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

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

The Weight

Given the expressions above, we can write our expression for the interpolant at a point x as

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

Writing out the averages that involve y we have:

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

where we define:

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

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 w_i are non-zero only in the Loess window around 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

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

Now we model the data as a local quadratic:

$$y_i = a_0 + a_1 x_i + a_2 x_i^2 + \epsilon_i$$

Finding a_0

As before, we define an error:

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

Minimizing

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

$$\Rightarrow a_0 = \langle y \rangle - a_1 \langle x \rangle - a_2 \langle x^2 \rangle$$

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

Finding a_1

Simiarly, for a_1 we have

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

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

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

$$= \langle x \rangle \langle y \rangle - a_1 \langle x^2 \rangle - a_2 \langle x \rangle \langle x^2 \rangle + a_1 \langle x^2 \rangle + a_2 \langle x^3 \rangle$$

Gathering terms

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

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

$$M_2 = \langle x^2 \rangle - \langle x \rangle^2$$
$$M_3 = \langle x^3 \rangle - \langle x^2 \rangle \langle x \rangle$$
$$C_{xy} = \langle xy \rangle - \langle x \rangle \langle y \rangle$$
$$= a_1 M_2 + a_2 M_3$$

Then we can write:

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

Finding a_2

For α_2 the result of minimizing the error with respect to α_2 is similar to the above, just adding another x in the averages, leading to

$$\langle x^2 y \rangle - \langle x^2 \rangle \langle y \rangle = a_1 (\langle x^3 \rangle - \langle x^2 \rangle \langle x \rangle) + a_2 (\langle x^4 \rangle - \langle x^2 \rangle^2) \tag{1}$$

We define the following geometric and correlation factors:

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

Substituting these into the above gives

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

Substituting our expression for a_1 from above

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

$$= \frac{M_3}{M_2} C_{xy} + a_2 (M_4 - \frac{M_3^2}{M_2})$$

Solving for a_2...

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

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

Substituting back into the expression for a_1 gives

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

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

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

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

The Weight

Returning to the interpolation expression, we start by substituting back a_0:

$$y(x) = a_0 + a_1 x + a_2 x^2$$

$$= \langle y \rangle - a_1 \langle x \rangle - a_2 \langle x^2 \rangle + a_1 x + a_2 x^2$$

$$= \langle y \rangle + a_1 (x - \langle x \rangle) + a_2 (x^2 - \langle x^2 \rangle)$$

Defining

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

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

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

We can write

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

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

Then

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

Defining

$$\hat{a}_1(x) \equiv \beta_2 (x - \langle x \rangle) - \beta_3 (x^2 - \langle x^2 \rangle)$$
$$\hat{a}_2(x) \equiv \beta_4 (x^2 - \langle x^2 \rangle) - \beta_3 (x - \langle x \rangle)$$

we can write, using the definitions of C_{xy} and $C_{x^2 y}$

$$y(x) = \langle y \rangle + \hat{a}_1(x) C_{xy} + \hat{a}_2(x) C_{x^2 y}$$
$$= \sum_i w_i y_i + \hat{a}_1(x) \sum_i w_i (x_i - \langle x \rangle) y_i + \hat{a}_2(x) \sum_i w_i (x_i^2 - \langle x^2 \rangle) y_i$$
$$= \sum_i w_i y_i \left[1 + \hat{a}_1(x)(x_i - \langle x \rangle) + \hat{a}_2(x)(x_i^2 - \langle x^2 \rangle) \right]$$
$$= \sum_i \hat{w}_i(x) y_i$$

where

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

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