LOWESS (Theory)

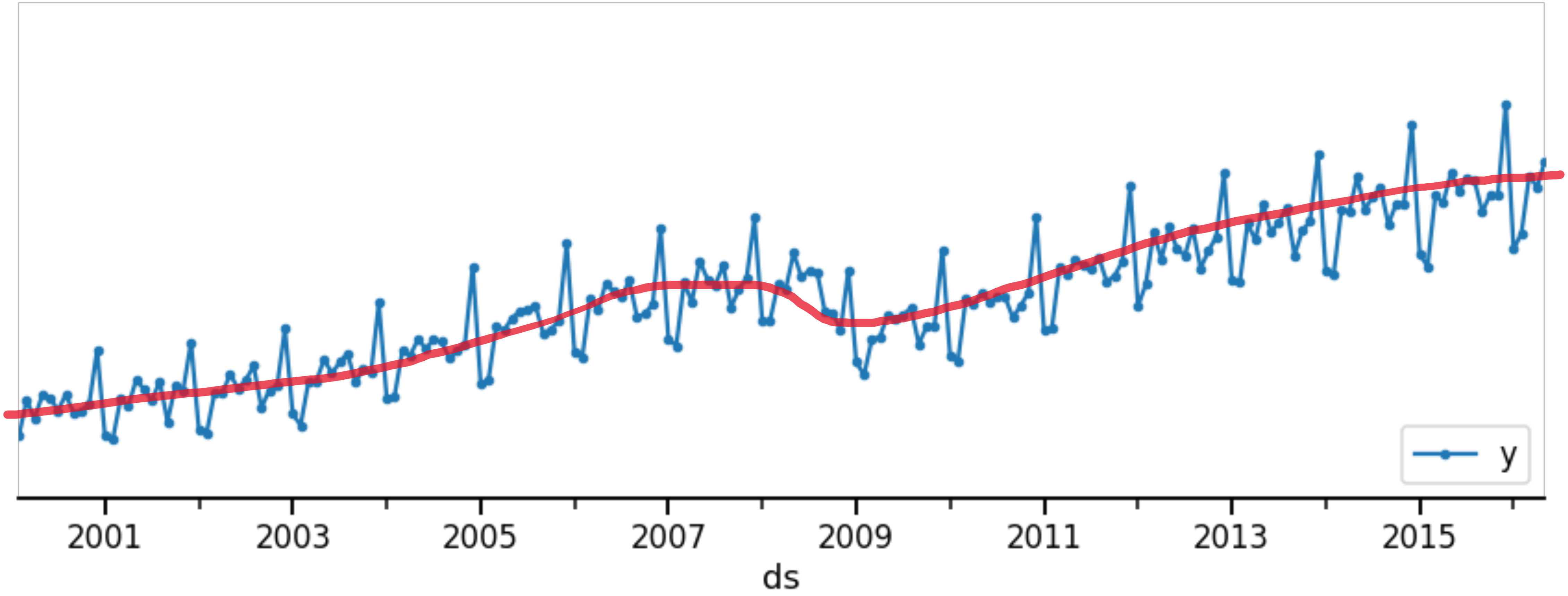
Time series decomposition

Contents

LOWESS FOR TREND EXTRACTION

DISCUSS LIMITATIONS

How can we extract the trend?



Limitations of moving averages

- Not robust to outliers
- Missing data at the edges
- Over-smooths rapid changes in trend
- Order of moving average was set by seasonal period

Date	y	mean
2020-02-12	23	
2020-02-13	30	41.0
2020-02-14	70	43.3
2020-02-15	**30**	41.7
2020-02-16	**25**	**25.7**
2020-02-17	**22**	

LOWESS

Robust Locally Weighted Regression and Smoothing Scatterplots

WILLIAM S. CLEVELAND*

The visual information on a scatterplot can be greatly enhanced, with little additional cost, by computing and plotting smoothed points. Robust locally weighted regression is a method for smoothing a scatterplot, (x_i, y_i), $i = 1, \ldots, n$, in which the fitted value at x_k is the value of a polynomial fit to the data using weighted least squares, where the weight for (x_i, y_i) is large if x_i is close to x_k and small if it is not. A robust fitting procedure is used that guards against deviant points distorting the smoothed points. Visual, computational, and statistical issues of robust locally weighted regression are discussed. Several examples, including data on lead intoxication, are used to illustrate the methodology.

KEY WORDS: Graphics; Scatterplots; Nonparametric regression; Smoothing; Robust estimation.

An early example of smoothing scatterplots is given by Ezekiel (1941, p. 51). The points are grouped according to x_i, and for each group the mean of the y_i is plotted against the mean of the x_i. More recently, Stone (1977) proves the consistency of a wide class of nonparametric regression estimates under very general conditions and presents a discussion and bibliography of methods that have appeared in the literature. Another method, which appeared after Stone's review, is that of Clark (1977), who proposes a technique for smoothing scatterplots in which the plot is interpolated by joining successive

Cleveland, William S. "Robust locally weighted regression and smoothing scatterplots." *Journal of the American statistical association* 74.368 (1979): 829-836.

The main idea

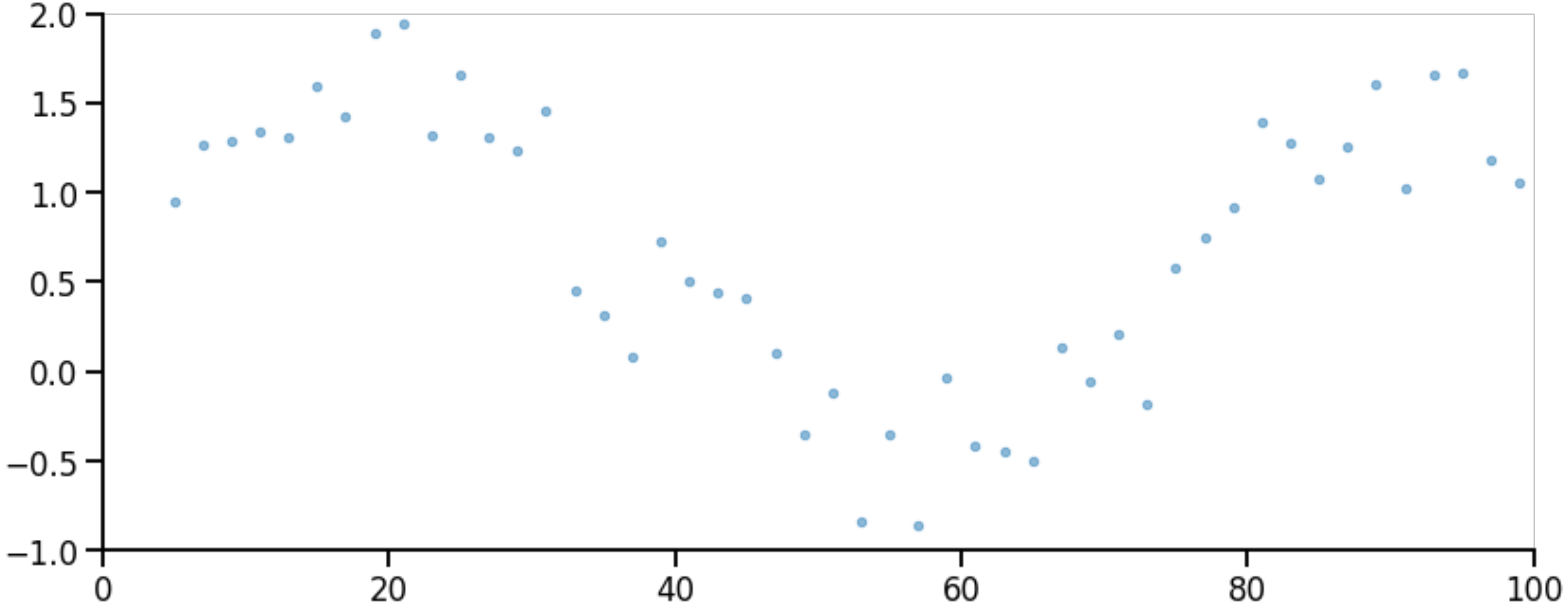
- Want to compute a smooth curve, $\hat{y}(x)$, to a scatterplot
- At x consider a window which captures a fraction f of the data
- Fit a weighted robust linear regression to this subset of the data
- The LOWESS curve at x is given by the below linear regression

$$\hat{y}(x) = \beta_0 + \beta_1 x;\ x_{train}, y_{train} \in \{(x_{j-n}, y_{j-n}), \dots, (x_{j+n}, y_{j+n})\}$$

- Evaluate the same process across many x values to obtain a smooth fit

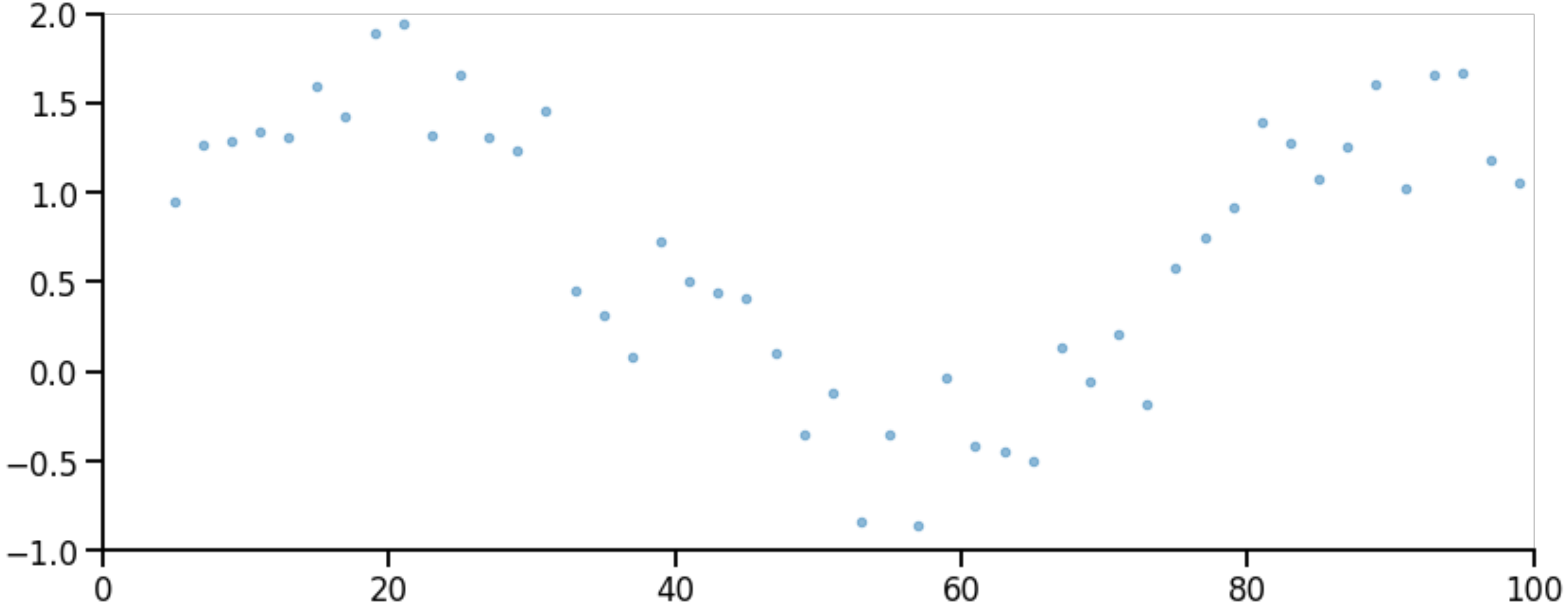
LOWESS

- Want to compute a smooth curve, $\hat{y}(x)$, to a scatterplot



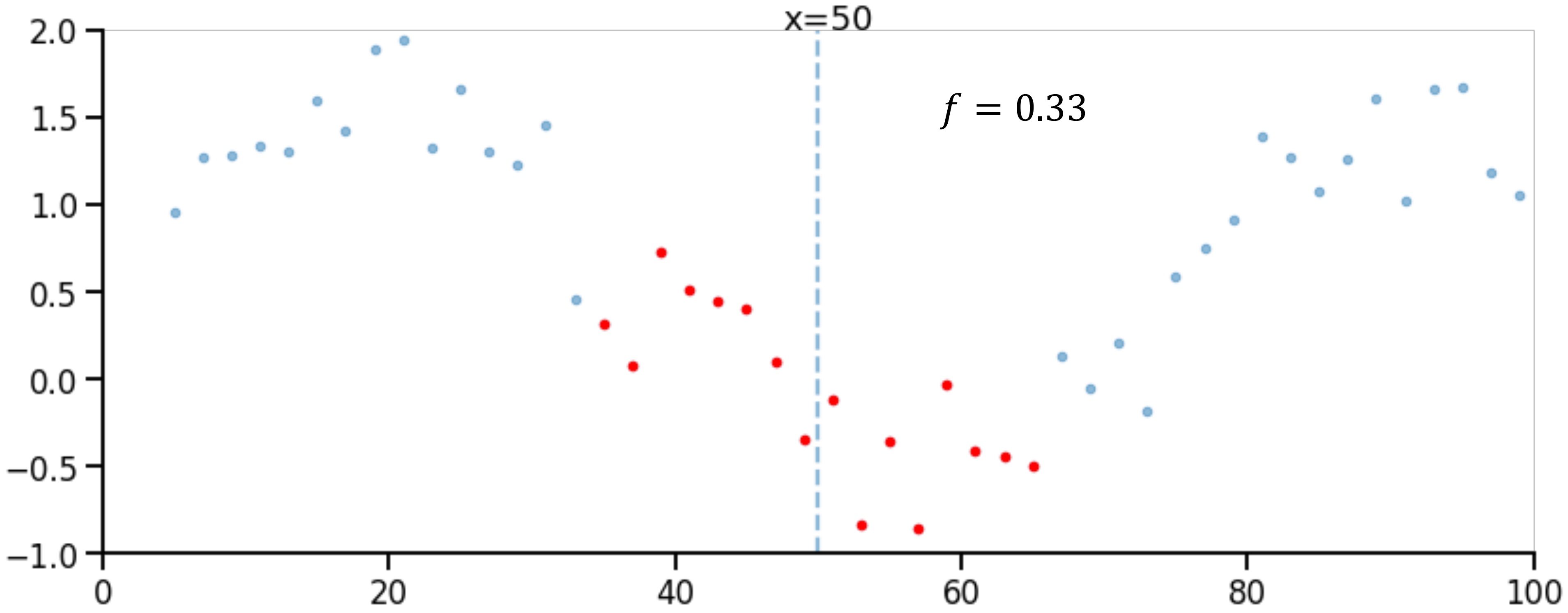
LOWESS

- At $x = 50$ consider a window which captures a fraction f of the data



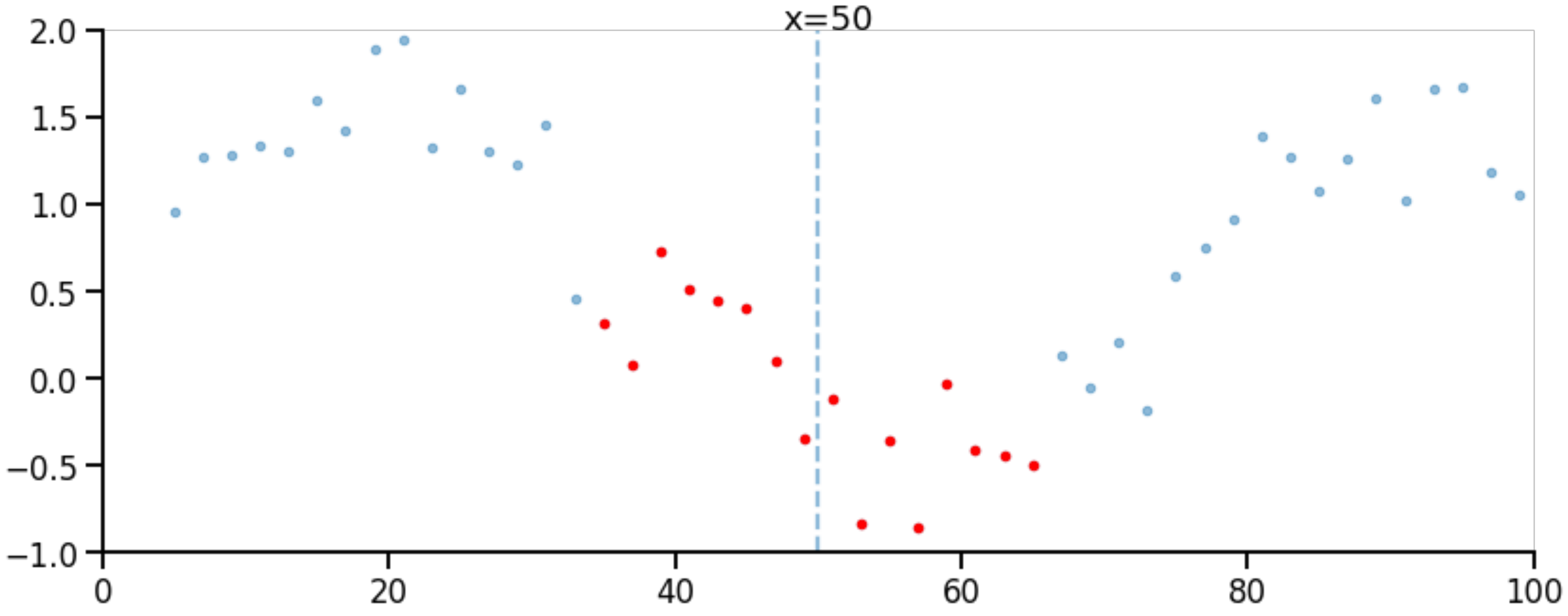
LOWESS

- At $x = 50$ consider a window which captures a fraction f of the data



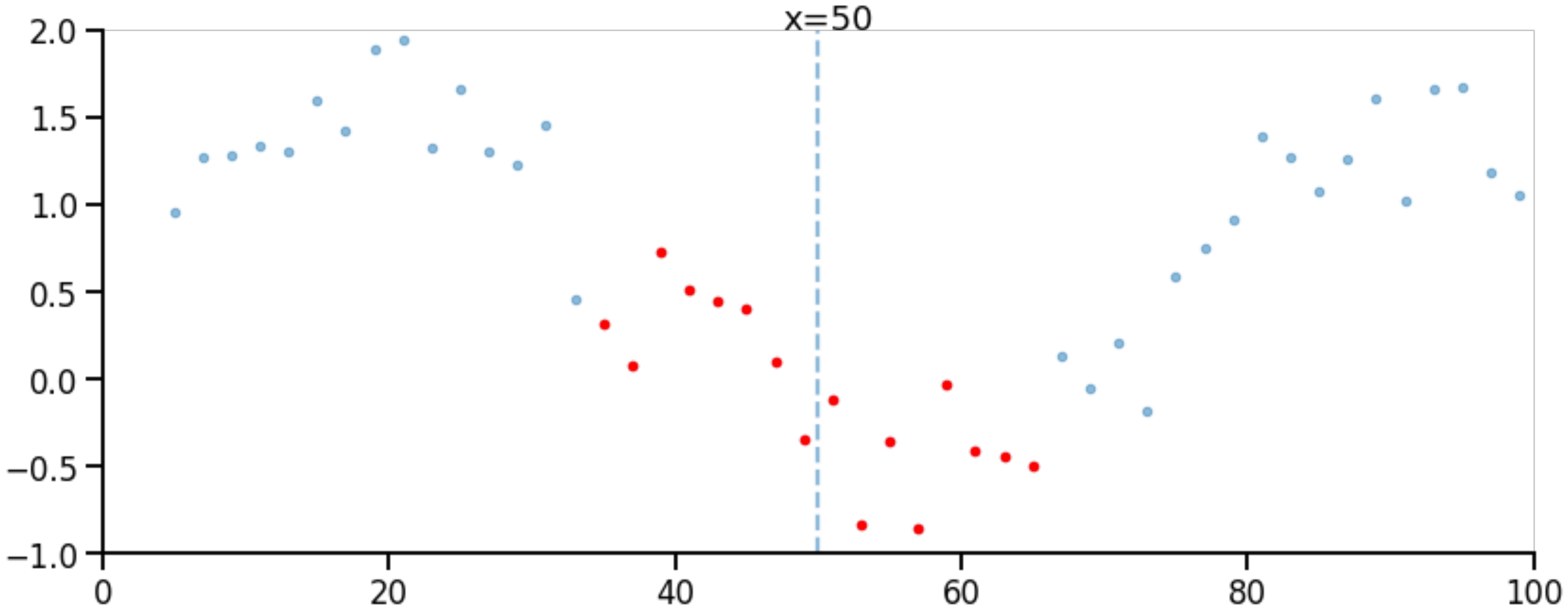
LOWESS

- Fit a weighted robust linear regression to this subset of the data



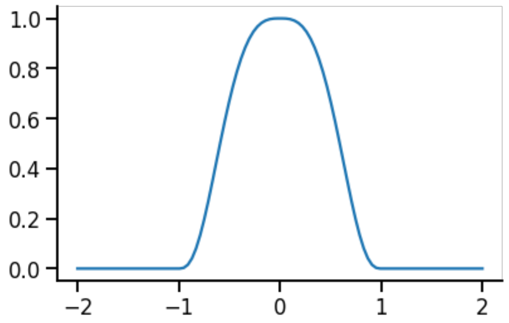
LOWESS

- Fit a **weighted** robust linear regression to this subset of the data



Weight function for LOWESS

Weights given by tricube function:



$$w(x) = \begin{cases} (1 - |x|^3)^3, & |x| < 1 \\ 0, & |x| \geq 1 \end{cases}$$

In practice $w(x)$ is rescaled when evaluating LOWESS at point x_i to fit the window size determined by f. So we compute:

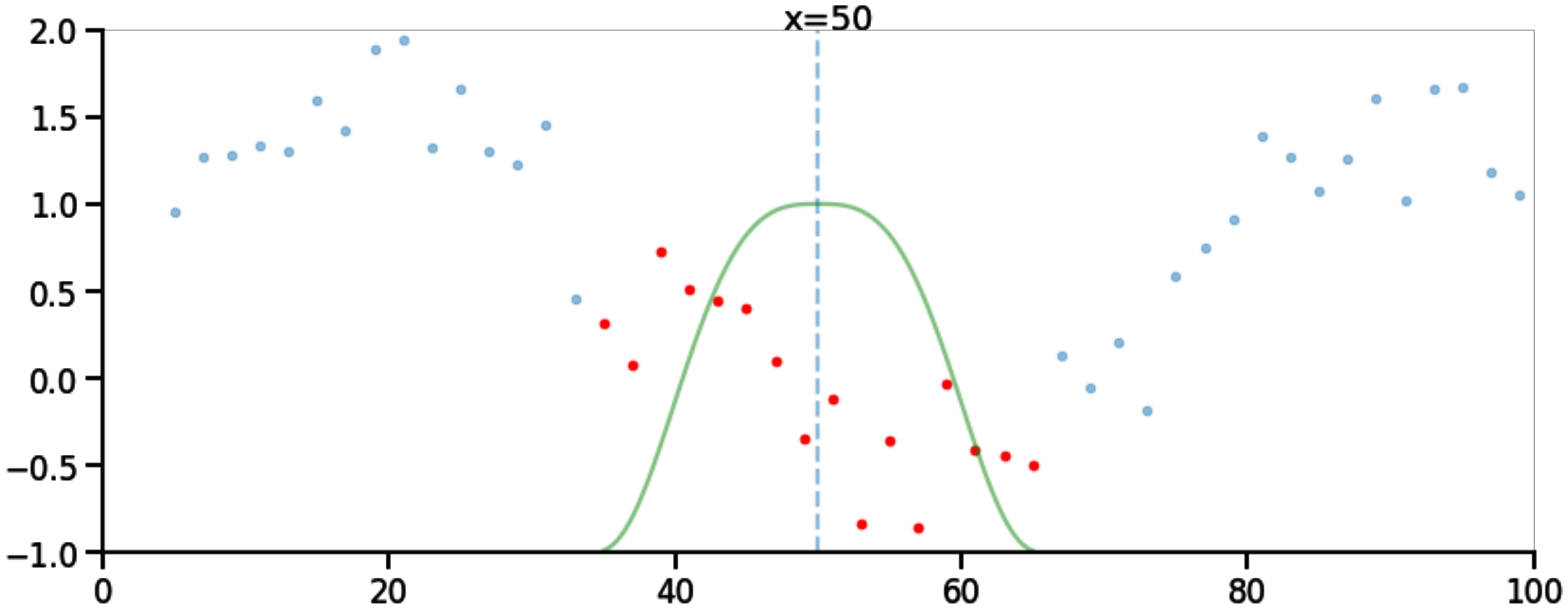
$$w^*(x) = w(\frac{1}{Nf}(x - x_i))$$

where N is the number of data points and f is the fraction of the dataset that determines the window size.

This ensures the weight function is centred at the point of interest x_i and the function goes to zero at the edge of the window.

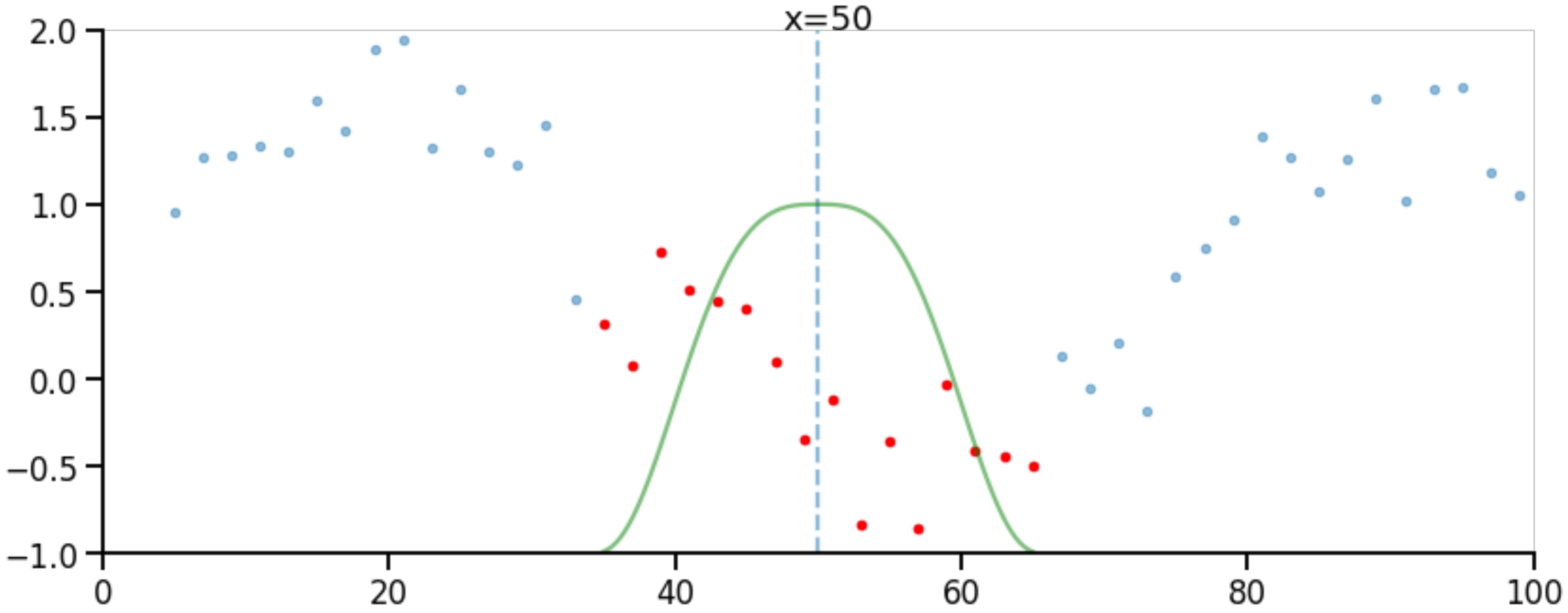
LOWESS

- Fit a **weighted** robust linear regression to this subset of the data



LOWESS

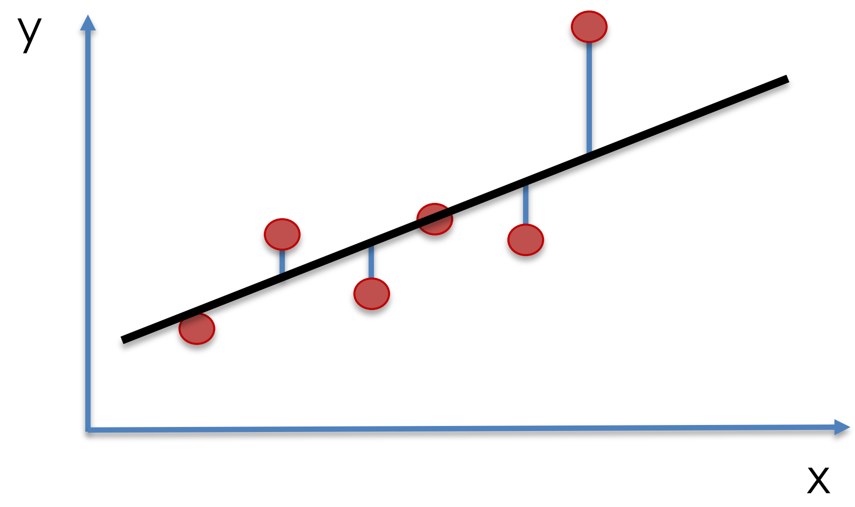
- Fit a weighted **robust** linear regression to this subset of the data



Robust regression

- Fit linear regression multiple times. On each iteration re-weight the data by residuals of the previous fit such that less weight is given to high residual data points.

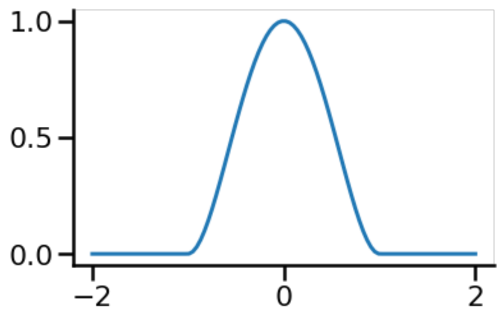
- Outliers produce large residuals. Hence, by re-weighting the data we can minimize their impact on the fit.



Robust regression

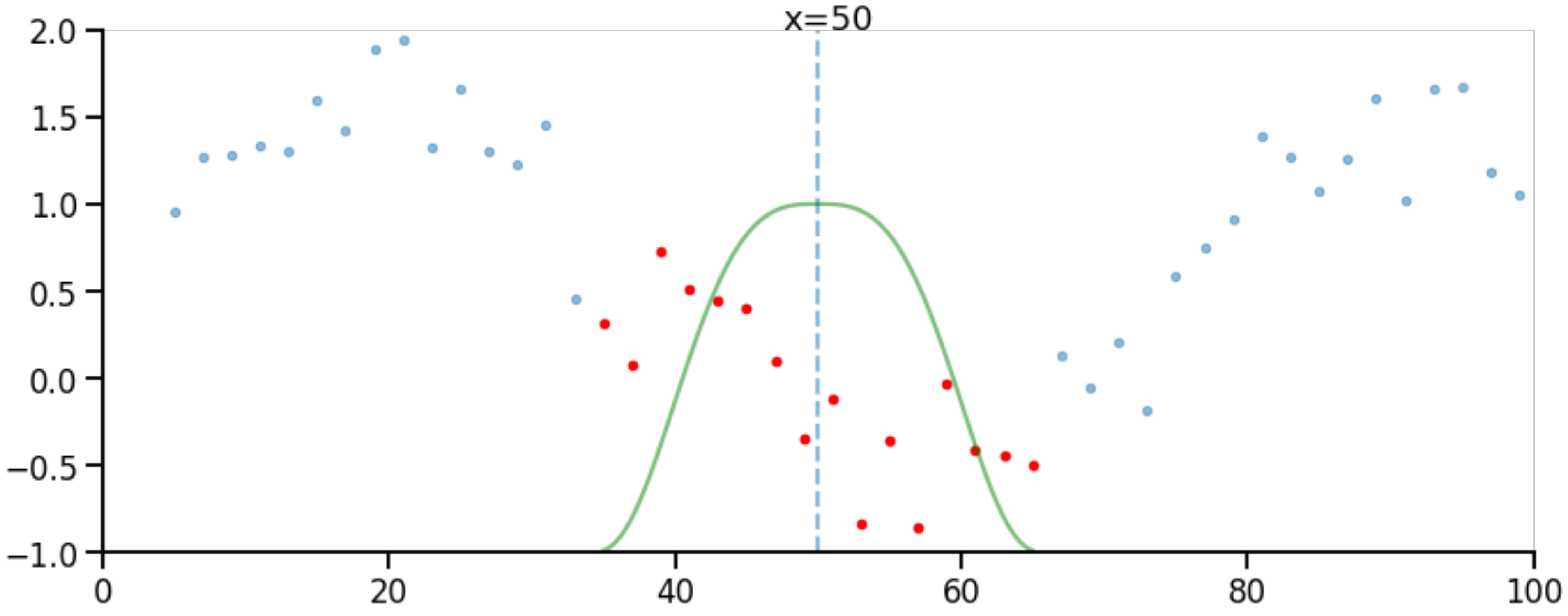
1. Fit weighted linear regression: $\hat{y} = \beta_0 + \beta_1 x; L = \sum_i w_i (y_i - \hat{y_i})^2$
2. Compute residuals: $e_i = y_i - \hat{y}_i$
3. Compute weights: $\delta_i = B(\frac{e_i}{6s})$, where s is median($|e_i|$)
4. Re-fit linear regression with weights: $\hat{y} = \beta_0 + \beta_1 x; L = \sum_i \delta_i w_i (y_i - \hat{y_i})^2$
5. Repeat t times

$$B(x) = \begin{cases} (1 - x^2)^2, & |x| < 1 \\ 0, & |x| \geq 1 \end{cases}$$



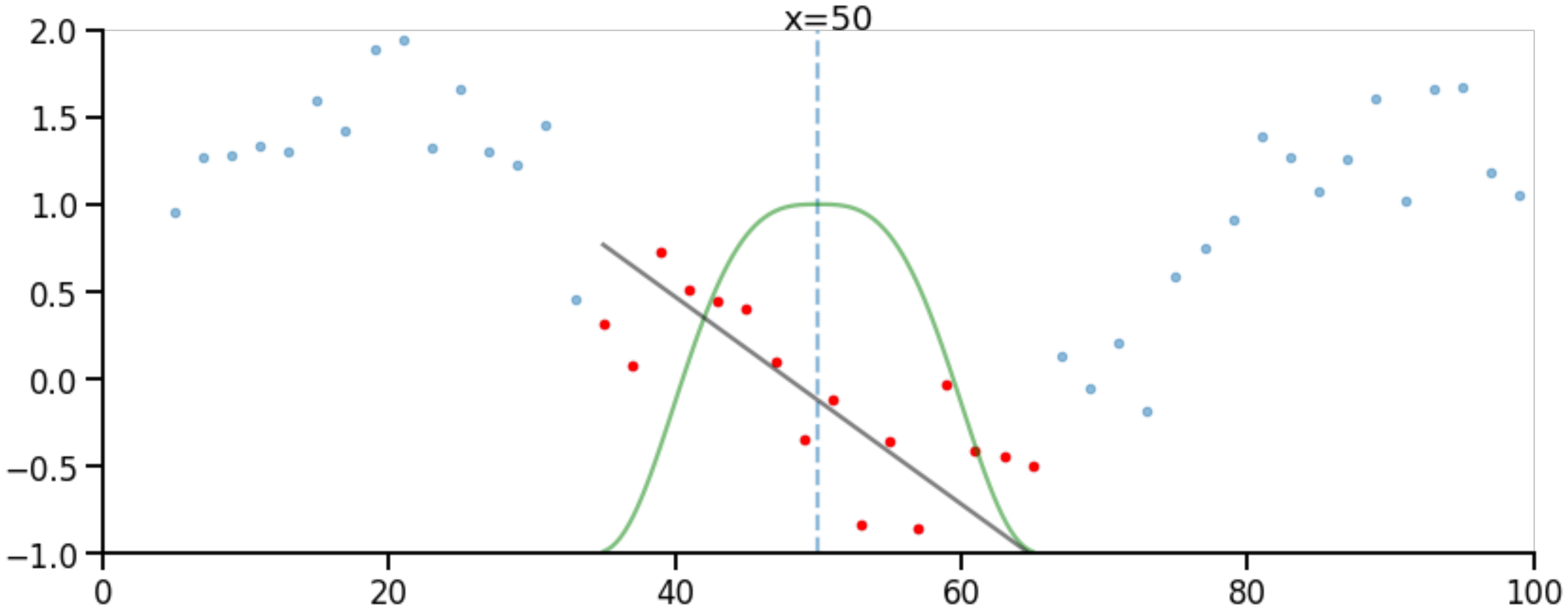
LOWESS

- Fit a weighted **robust** linear regression to this subset of the data



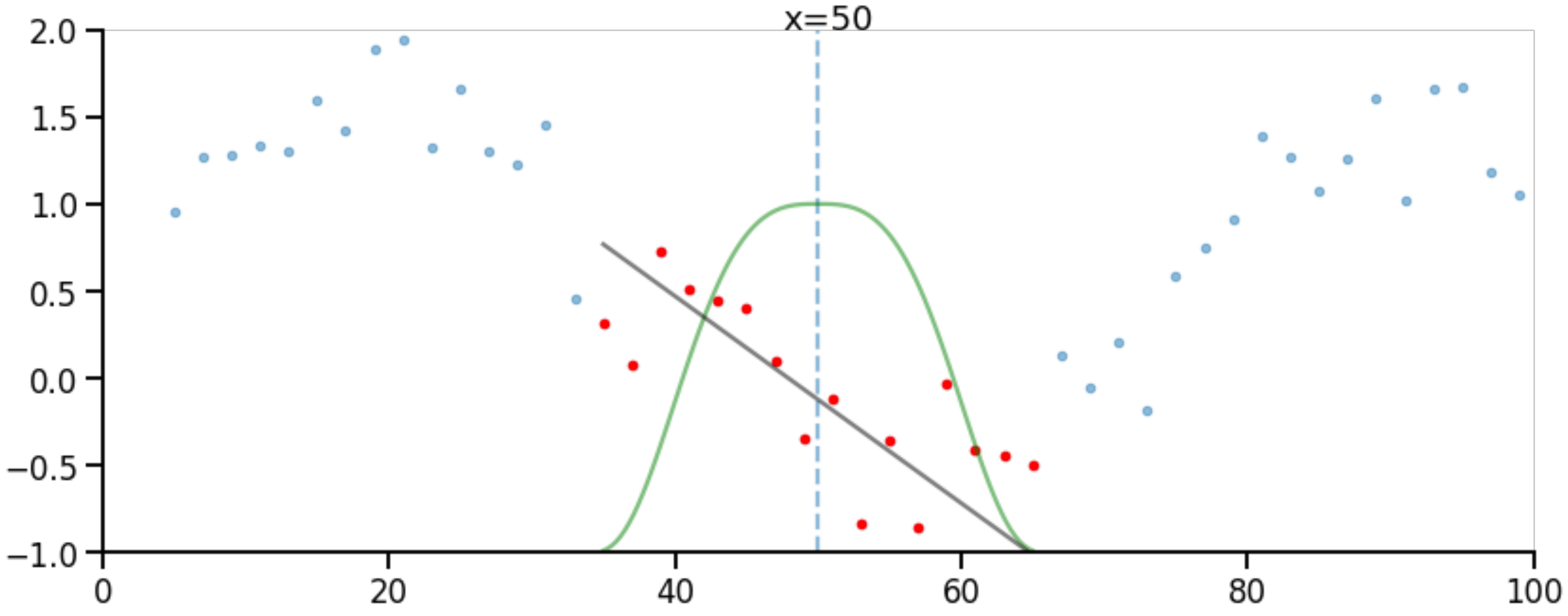
LOWESS

- Fit a weighted **robust** linear regression to this subset of the data



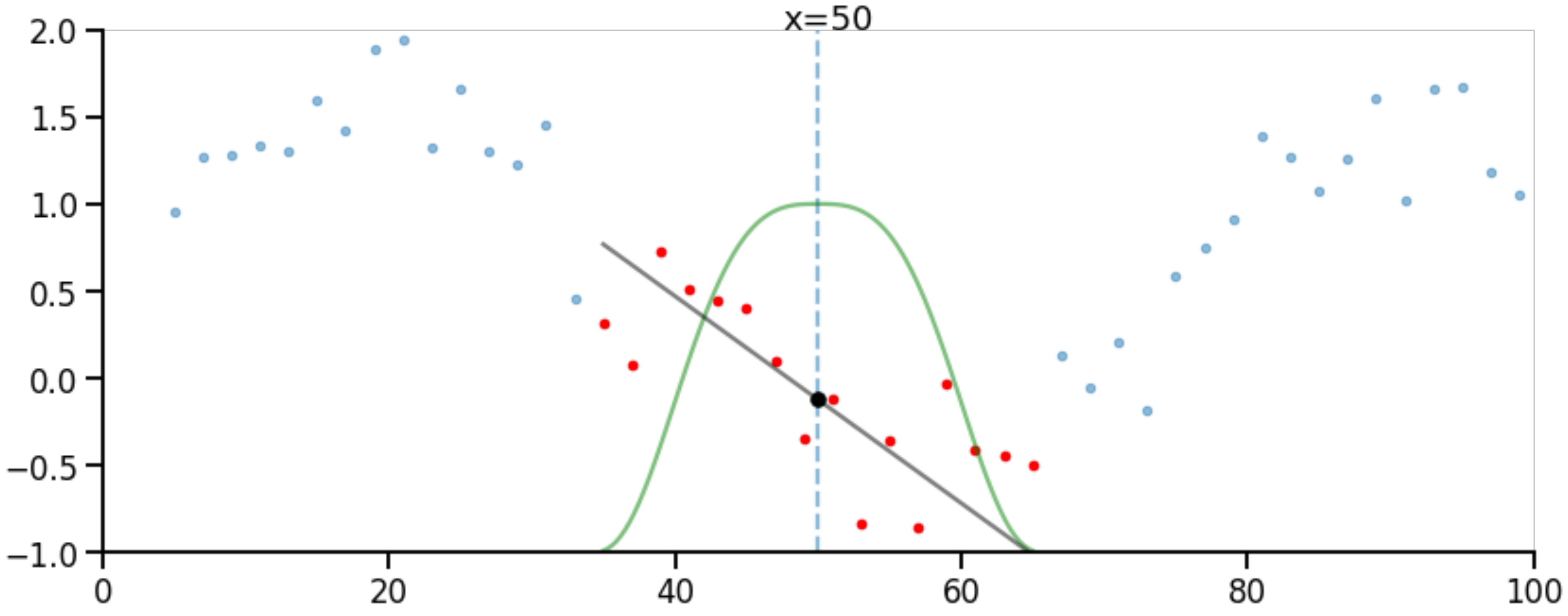
LOWESS

- LOWESS curve at x is given by the weighted robust linear regression



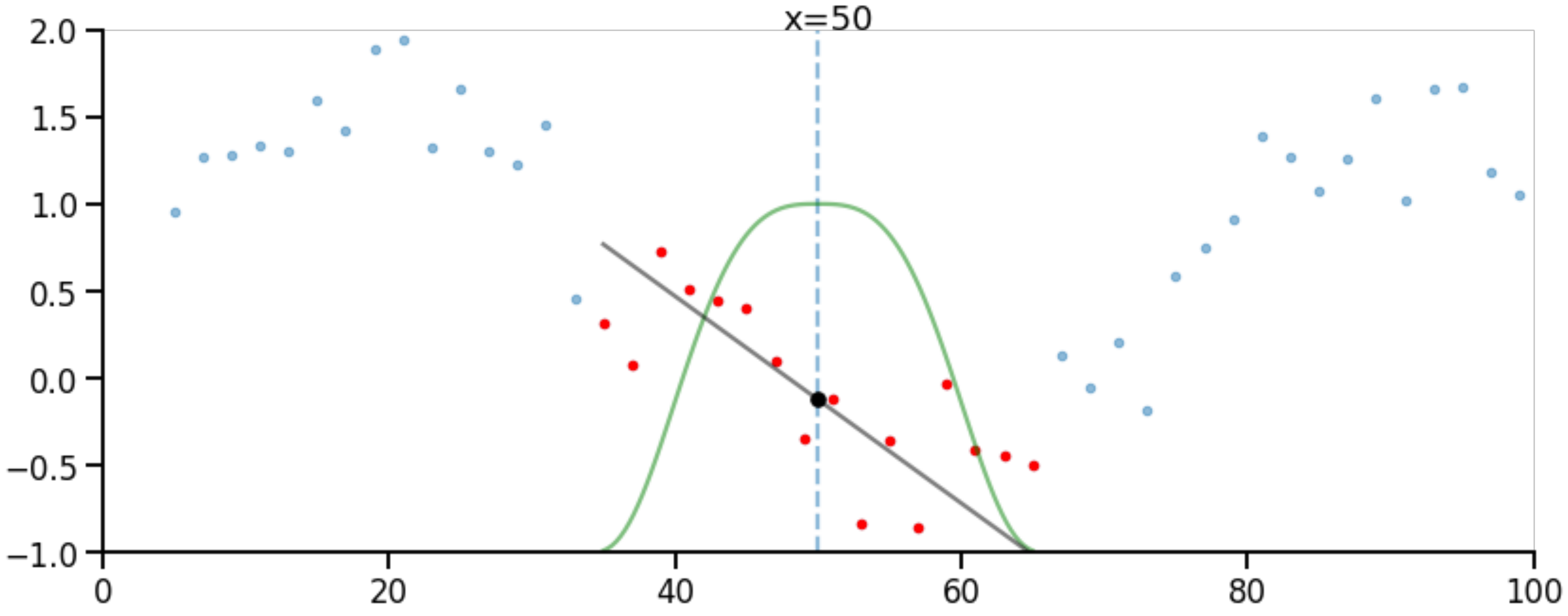
LOWESS

- LOWESS curve at x is given by the weighted robust linear regression



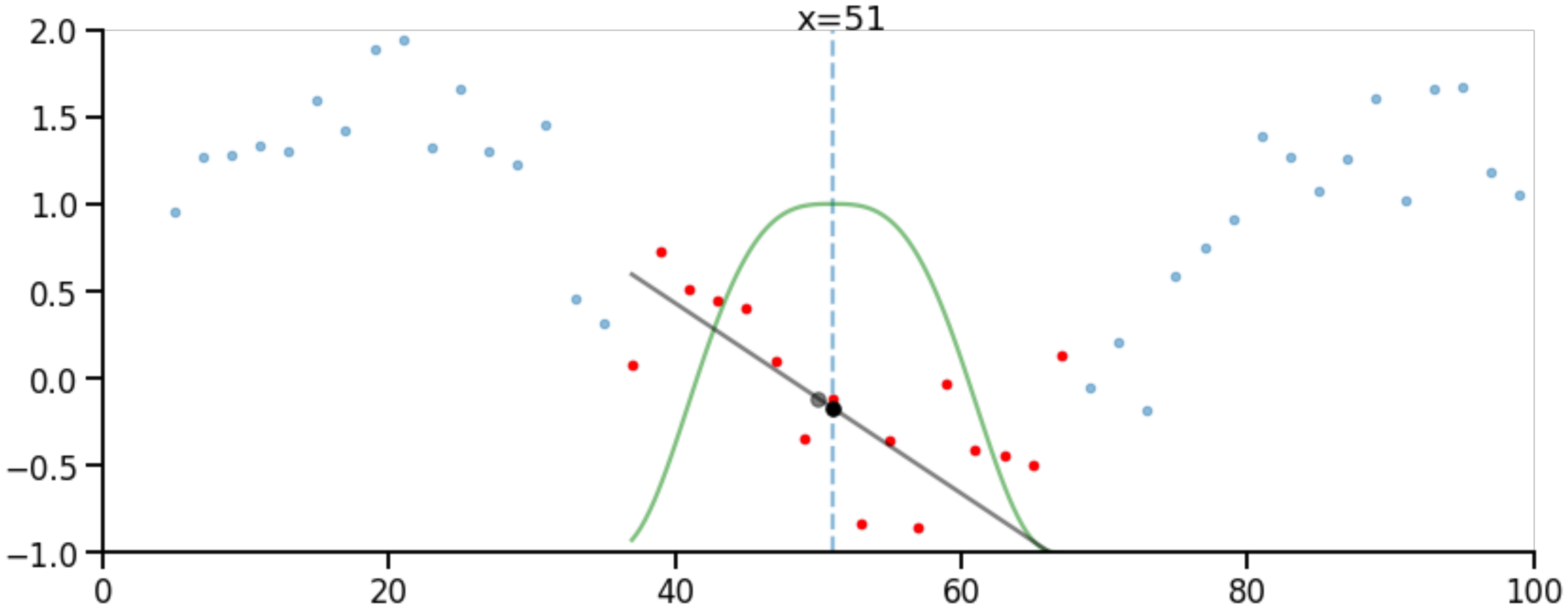
LOWESS

- Evaluate the same process across many x values to obtain a smooth fit



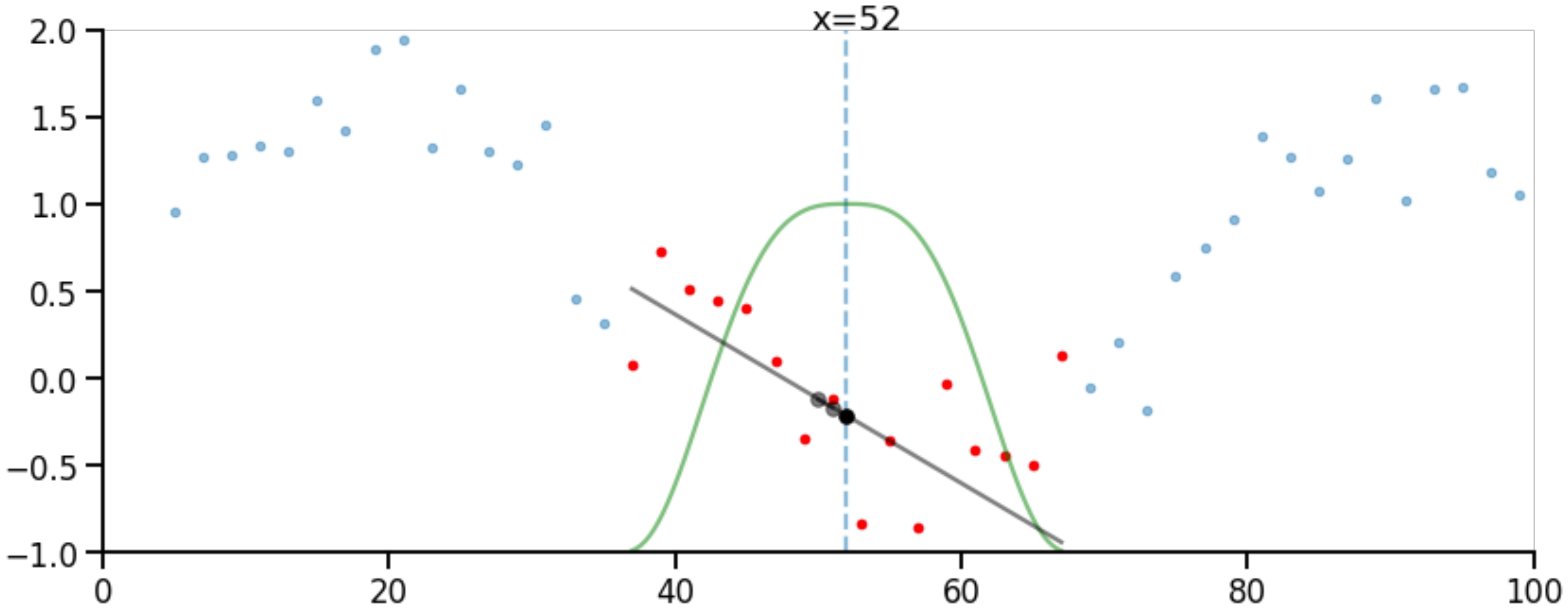
LOWESS

- Evaluate the same process across many x values to obtain a smooth fit



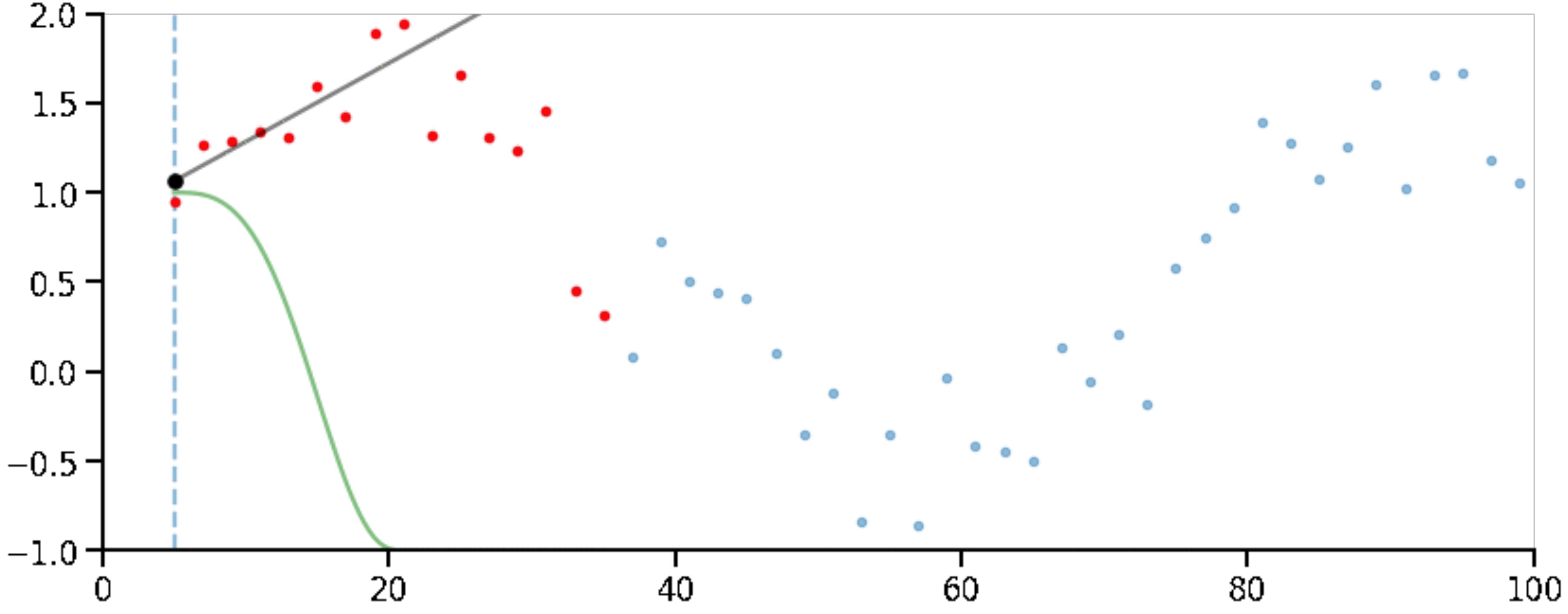
LOWESS

- Evaluate the same process across many x values to obtain a smooth fit



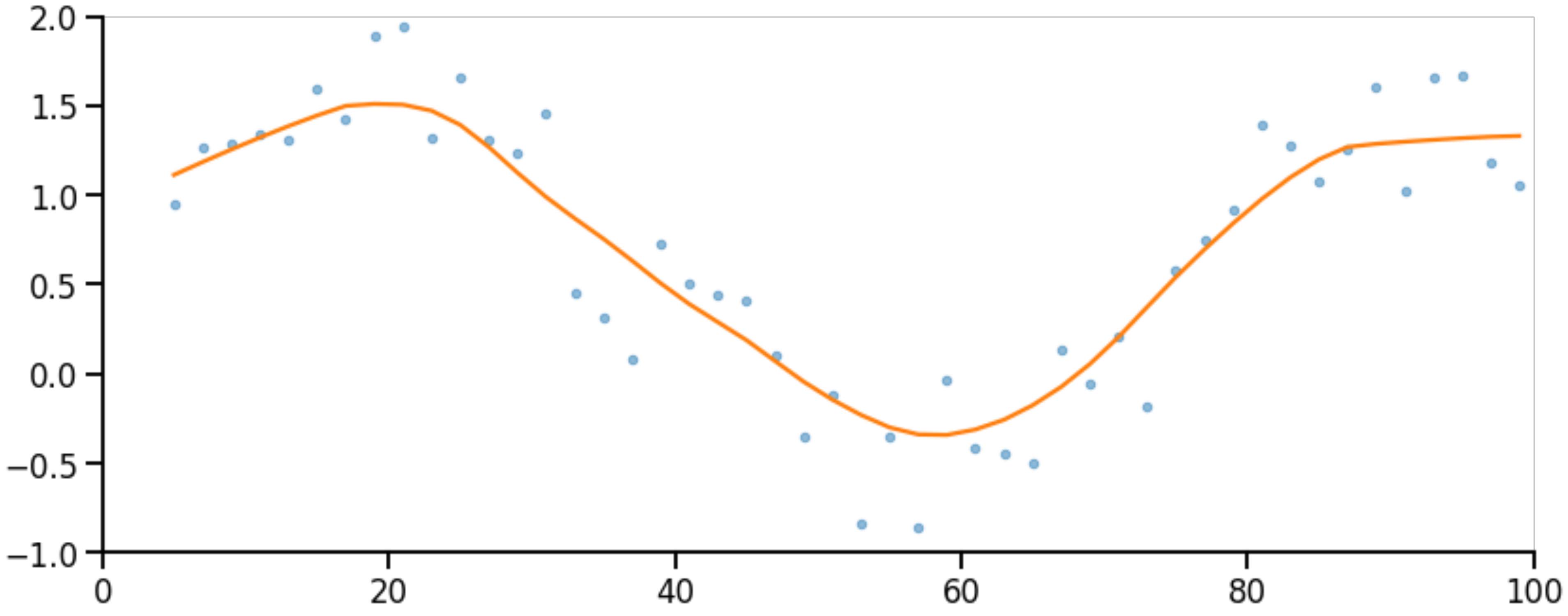
LOWESS

- Evaluate the same process across many x values to obtain a smooth fit



LOWESS

- Evaluate the same process across many x values to obtain a smooth fit



LOWESS

- The LOWESS fit can be used as an estimate of the trend of a time series