Nonparametric Regression: Lowess/Loess

GEOG 414/514:

Advanced Geographic Data Analysis Scatter-diagram smoothing.

Adapted by Ronaldo Dias

1 Introduction

Scatter-diagram smoothing involves drawing a smooth curve on a scatter diagram to summarize a relationship, in a fashion that makes few assumptions initially about the form or strength of the relationship. It is related to (and is a special case of) nonparametric regression, in which the objective is to represent the relationship between a response variable and one or more predictor variables, again in way that makes few assumptions about the form of the relationship. In other words, in contrast to "standard" linear regression analysis, no assumption is made that the relationship is represented by a straight line (although one could certainly think of a straight line as a special case of nonparametric regression).

Another way of looking at scatter diagram smoothing is as a way of depicting the "local" relationship between a response variable and a predictor variable over parts of their ranges, which may differ from a "global" relationship determined using the whole data set. (And again, the idea of "local" as opposed to "global" relationships has an obvious geographical analogy.)

2 Loess

A bivariate smoother is a function or procedure for drawing a smooth curve through a scatter diagram. Like linear regression (in which the "curve" is a straight line), the smooth curve is drawn in such a way as to have some desirable properties. In general, the properties are that the curve indeed be smooth, and that locally, the curve minimize the variance of the residuals or prediction error.

The bivariate smoother used most frequently in practice is known as a "lowess" or "loess" curve. The acronyms are meant to represent the notion of locally weighted regression–a curve- or function-fitting technique that provides a generally smooth curve, the value of which at a particular location along the x-axis is determined only by the points in that vicinity. The method consequently makes no assumptions about the form of the relationship, and allows the form to be discovered using the data itself. (The difference between the two acronyms or names is mostly superficial, but there is an actual difference in R–there are two different functions, lowess() and loess(), which will be explained below.)

2.1 Robust Loess

Cleveland (1979) proposed the algorithm LOWESS, locally weighted scatter plot smoothing, as an outlier resistant method based on local polynomial fits. The basic idea is to start with a local polynomial (a k-NN type fitting) least squares fit and then to use robust methods to obtain the final fit. Specifically, one can first fit a polynomial regression in a neighborhood of x, that is, find $\beta \in \mathbb{R}^{p+1}$ which minimize

$$n^{-1} \sum_{i=1}^{n} W_{ki}(x) \Big(y_i - \sum_{j=0}^{p} \beta_j x^j \Big)^2, \tag{2.1}$$

where $W_{ki}(x)$ denote k-NN weights. Compute the residuals $\hat{\epsilon}_i$ and the scale parameter $\hat{\sigma} = median(\hat{\epsilon}_i)$. Define robustness weights $\delta_i = K(\hat{\epsilon}_i/6\hat{\sigma})$, where $K(u) = (15/16)(1-u)^2$, if $|u| \leq 1$ and $K(u) = 0$, if otherwise. Then, fit a polynomial regression as in (2.1) but with weights $(\delta_i W_{ki}(x))$. Cleveland suggests that $p = 1$ provides good balance between computational ease and the need for flexibility to reproduce patterns in the data. The smoothing parameter can be determined by cross-validation.

2.2 Lowess/Loess in R

Note that there are actually two versions of the lowess or loess scatter-diagram smoothing approach implemented in R. The former (lowess) was implemented first, while the latter (loess) is more flexible and powerful. Example of lowess:

`lowess(x, y, f=2/3, iter=3, delta=.01*diff(range(x)))`. Where we suppose the following model

$$y = g(x) + \varepsilon$$

f: `the smoother span. This gives the proportion of points in the plot which influence the smooth at each value. Larger values give more smoothness.`

iter: `the number of robustifying iterations which should be performed. Using smaller values of ‘iter’ will make ‘lowess’ run faster.`

delta: `values of ‘x’ which lie within ‘delta’ of each other replaced by a single value in the output from ‘lowess’.`

```
data(cars)
plot(cars, main = "lowess(cars)")
lines(lowess(cars), col = 2)
lines(lowess(cars, f=.2), col = 3)
legend(5, 120, c(paste("f = ", c("2/3", ".2"))), lty = 1, col = 2:3)
```

Observe that, The newer loess() function uses a formula to specify the response (and in its application as a scatter-diagram smoother) a single predictor variable. The loess() function creates an object that contains the results, and the predict() function retrieves the fitted values. These can then be plotted along with the response variable. However, the points must be plotted in increasing order of the predictor variable in order for the lines() function to draw the line in an appropriate fashion this is done by using the results of the order() function applied to the predictor variable values and the explicit subscripting (in square brackets [ ]) to arrange the observations in ascending order.

Example of Loess: `data(cars)`

```
cars.lo <- loess(dist   speed, cars)
predict(cars.lo, data.frame(speed = seq(5, 30, 1)), se = TRUE)
# to allow extrapolation
lines(speed,cars.lo$fit)
cars.lo1<-loess(dist   speed, cars,span=1.5)
lines(speed,cars.lo$fit,col=2)
cars.lo2 <- loess(dist   speed, cars, control = loess.control(surface =
"direct"))
predict(cars.lo2, data.frame(speed = seq(5, 30, 1)), se = TRUE)
lines(speed,cars.lo2$fit)
```

References

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots, *J. Amer. Statist. Assoc.* **74**(368): 829–836.