Finding Fault: Identifying and Testing 'Social Faultlines' in Surface Fitting Techniques

Chris Brunsdon[1]

[1]Department of Geography
Bennett Building
University of Leicester
University Road
Leicester LE1 7RH, UK
Telephone: 0116 252 3843
Fax: 0116 252 3854
Email: cb179@le.ac.uk

Introduction

Surface-based approaches have frequently been used to analyse social and economic data. Using approaches such as kernel regression it has been possible to fit continuous surfaces to spatially reference social and economic data, such as house prices. The technique has often proved a useful tool in identifying trends in the data - for example one can identify areas of town in which housing is generally more costly.

Thus, the idea of kernel regression is to estimate trend surfaces - so that if we have a set of points, say (x,y) and each one has a continuous scale attribute (say z) then the aim is to estimate the value of z at values of (x,y) other than those in the data set - essentially estimating a z-surface from a set of point observations of z. This is done by creating a *kernel* around the point (x,y) and taking a weighted mean of z-*values of* points in the vicinity of (x,y) - the weight decreasing the further the data points are away from (x,y). A typical kernel function might be

$$w = \exp\left(-\frac{d^2}{2k^2}\right) \tag{1}$$

where d is the distance from the point at which we estimate z and a point in the data set, and w is the weight given to the z value associated with that data point. k is a smoothing parameter - the larger its value the smoother the trend surface.
In a sense a regression surface passes smoothly through the *centre* of the observed z-values if viewed as a 3D point cloud - so observed and fitted values at points in the data set may differ. This is a sensible approach if the observed z-values may be subject to sampling variation or other random factors influencing the sale price of a house.

Here not only smooth regression surfaces are considered, but also surfaces in which there may be discontinuities. An example highlighting the differences is shown in fig. 1. Both surfaces were computed from the same set of point samples, but the RHS panel is the result of trend estimation with discontinuity detection, whereas the LHS panel uses a standard approach. Since the RHS approach applies smoothing over a window regardless of discontinuity, the effect is to smooth away this feature.

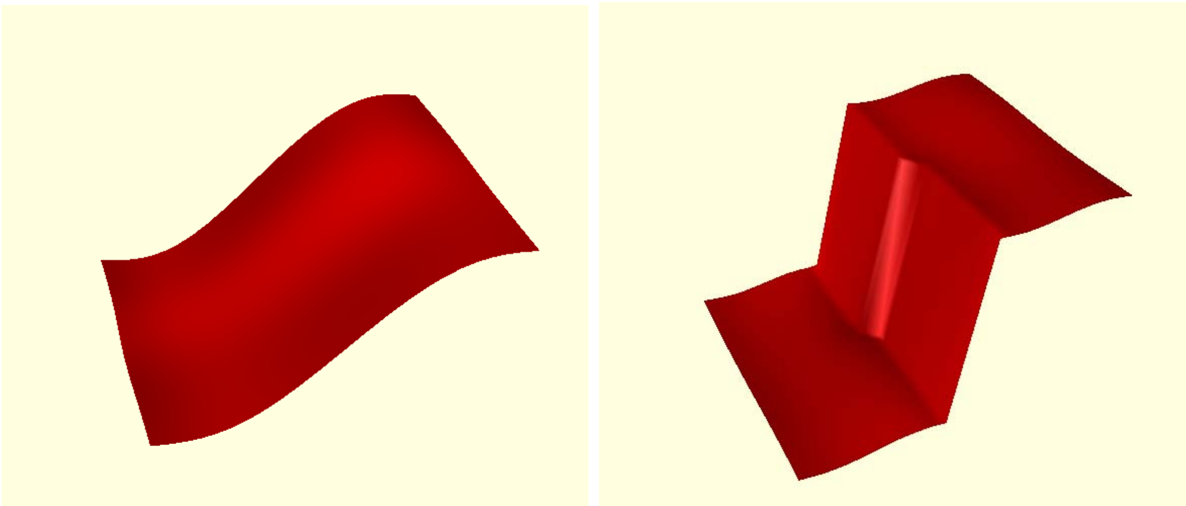

Figure 1. Two surfaces: LHS without discontinuities, RHS with discontinuities.

In this paper, two approaches to kernel smoothing - both modifying the basic kernel smoothing idea in different ways, will be outlined. The methods outlined are relevant to both physical and human geography data - in physical terms it allows terrain modeling with cliff edges, for example, but in this instance focus will be given to the fact that it can also be used to detect metaphorical fault lines in terms of social data - situations in which relatively affluent regions lie beside areas of high deprivation, or where house prices suddenly increase when a 'golden postcode' boundary is crossed. It is also worth noting that both methods *detect* discontinuities rather than work with locations deemed discontinous on an *a priori* basis.

The two methods will now be outlined:

Anisotropic Smoothing

This method is due to Perona and Malik (1990). This operates on regular grid data. One way of smoothing grid data (without detecting discontinuities) is to replace the value of each pixel in the grid with a weighted average its immediate neighbours:

$$z^*_{i,j} = \frac{w_{1,0}z_{i+1,j} + w_{-1,0}z_{i-1,j} + w_{0,-1}z_{i,j-1} + w_{0,1}z_{i,j+1}}{w_{1,0} + w_{-1,0} + w_{0,-1} + w_{0,1}} \qquad (2)$$

The asterisk on the z denotes an updated value, and this computation is applied to each z in the 2D array. This only applies a small amount of smoothing - it is rather like a moving window approach where the window is just one pixel wide. An effect equivalent to a larger window can be achieved by repeatedly applying the operation just described.

Here the weights do not depend on the indices i and j - they are stationary - and the overall effect is similar to moving window smoothing. Unfortunately, for that reason, this approach also does not work well with discontinuities. However, suppose that the weights used were reduced in situations where the values of adjacent pixels were very different, or were on a rapidly changing part of the surface. This would reduce the 'smoothing off of edges' problem outlined in the last section.

One way to do this would be to make the weights depend on the slope estimates at each of the pixels - so that the influence of pixels on steeply sloping parts of the surface would be downgraded in the smoothing process. This could be achieved by a minor modification of the smoothing approach outlined above:

$$z^*_{i,j} = \frac{w_{1,0}z_{i+1,j} + w_{-1,0}z_{i-1,j} + w_{0,-1}z_{i,j-1} + w_{0,1}z_{i,j+1}}{w_{1,0} + w_{-1,0} + w_{0,-1} + w_{0,1}}$$

$$\text{where } w_{i,j} = f(s_{i,j}) \quad (3)$$

and $s_{i,j}$ is a slope estimate at pixel (*i,j*). Typically, *f* is a decreasing function, for example

$$\exp\left(-\frac{s_{i,j}^2}{h_2^2}\right) \quad (4)$$

so that the influence of pixels on a steep slope is downgraded.

As with the straightforward pixel based smoothing a single smooth operation takes place over a very tight window, but this time it does not over smooth when the surface changes rapidly. As before, the effect of using larger smoothing windows is achieved by repeated application. Typically, 20 or 30 applications are used.

A final issue here is how the slope estimation is carried out. There are a number of possibilities - for example using Horn's method - however here, a fairly simple approach appears to be effective. Assuming the grid spacing of the pixels is the same in the *x* and *y* directions a reasonable estimator of slope is:

$$s_{i,j} = \left\{\frac{(z_{i+1,j} - z_{i-1,j})^2 + (z_{i,j+1} - z_{i,j-1})^2}{4}\right\}^{\frac{1}{2}} \quad (5)$$

This only applies to gridded data, and in social and economic data, irregular points are often found. This technique requires the *z*-values to be known at the points where the trend is to be fitted, and using regular gridded data is a convenient way of ensuring this that will be the case. However, the question remains: "how can one apply this approach to irregular data?". The simple answer here is that a standard moving window smoother is first applied to the irregular data, resulting in a grid of values and subsequently the anisotropic diffusion filter is applied to this grid. One issue here is that this involves the data undergoing *two* smoothing processes - and so there is a danger of oversmoothing. Usually this can be solved by carrying out the first standard smoothing with a very small window - the aim here is really just to transform the data into regular grid format - and then to apply anisotropic smoothing.

Bilateral Filtering

This method is due to Tomasi and Manduchi (1998). This method also involves controlling the degree of smoothing when there is a large difference between the *z*-values of nearby points. To do this a kernel function in terms of *x,y and z* is used - typically kernel functions only depend on nearness in geographical space (x,y), but in this case nearness in attribute space is also considered. A typical function may be:

$$w = \exp(-\frac{d_1^2}{2k_1^2})\mathrm{I}(d_2 < k_2) \quad (6)$$

where d_1 is a geographical distance as before and d_2 is the absolute difference in *z* values between a pair of points. k_1 and k_2 are now both smoothing parameters - k_1 functions as *k* does in a standard kernel smoother. For the term involving d_2 and k_2, *I* is an indicator function. The effect is that standard kernel weighting occurs when two *z*-values are within d_2 of one another, otherwise no weighting occurs. Thus if d_2 is regarded as a significant enough difference to suggest a 'cliff edge' is between the points, no smoothing occurs. Since this method requires a z value at each location a two stage smoothing process is required as before.

The Presentation

In the proposed presentation these two smoothing methods will be introduced, and then illustrated with an example using Townsend scores of deprivation (Townsend et al 1988) in Leicestershire, UK. These identify some potential 'faultlines' - see fig. 2:

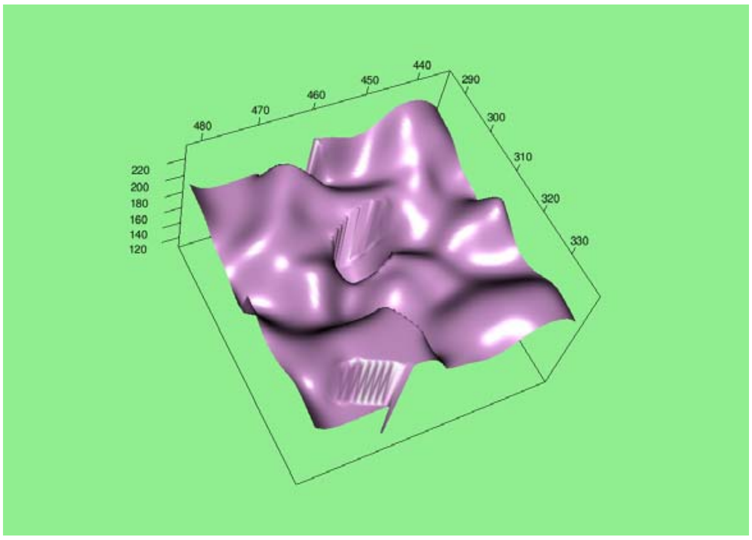

Figure 2. Surface of Townsend Score for Leicestershire with faultline detection

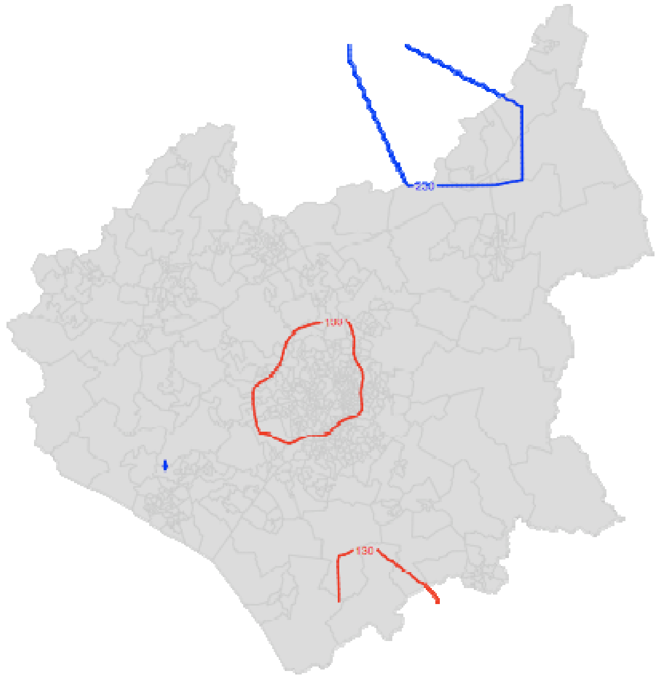

Figure 3. 'Faultlines' illustrated on a map of Leicestershire

This is shown in a more familiar map-based form in fig. 3 (above) – the contour lines here mark the faultlines in the 3D plot in fig 2.

Finally, to assess the reliability of these 'faultlines' a bootstrap analysis is carried out to see whether they could have occurred as an artifact of the methodology even in a situation where no sudden changes in the surfaces exist in reality. This will be shown in the presentation.

References

Townsend P, Phillimore P, Beattie A. 1988 *Health and Deprivation: Inequality and the North*. Routledge: London

Perona P. and Malik, J. 1990 *Scale-space and edge detection using anisotropic diffusion*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(7):629-639.

Tomasi, C. and Manduchi, R, 1988 *Bilateral Filtering for Gray and Color Images*, Proceedings of the 1998 IEEE International Conference on Computer Vision, Bombay, India