Math 131/135/194, Fall 2004

Applied Calculus

Daniel Kaplan, Macalester College

Constructing Local Models

Our general purpose tool for constructing local models of functions of two variables is the polynomial. The point of constructing such models isn't to capture exactly every aspect of the relationship, but to build a scaffolding that can be used to analyze and interpret data, hopefully leading to a better description of the relationship.

We imagine that there is an output that is a function of two inputs: $f(x, y)$. The polynomial function that we will use will be

$$f(x, y) = a_0 + a_1 x + a_2 y + a_3 xy + a_4 x^2 + a_5 y^2$$

Depending on the values of the parameters $a_0, a_1, a_2, a_3, a_4, a_5$, this function can take on all sorts of shapes. But, in general, not all of the terms are needed.

a_0 The constant term. This sets a typical value of $f(x, y)$, but doesn't depend on either x or y. It is almost always included by default.

$a_1 x$ Produces a simple dependence on the input x; if the input x changes, then the output $f(x, y)$ will change.

$a_2 y$ Likewise, produces a simple dependence on the input y.

$a_4 x^2$ This quadratic term can do two things. It is absolutely needed in the model if there is a maximum or minimum with respect to x. But, even if there is no extremum, if there is an important change in $\frac{\nabla f}{\nabla x}$ as x changes, then there should be this quadratic term. Example: economists often speak of diminishing marginal returns — doubling the amount of investment doesn't lead to a doubling in output per dollar of investment.

$a_5 y^2$ Likewise, needed for there to be an extremum with respect to y, or a change in $\frac{\nabla f}{\nabla y}$.

$a_3 xy$ The interaction term. This term expresses how the inputs x and y interact: perhaps interfering with one another or reinforcing one another. Whenever the output will depend on x differently for different values of y, or vice versa, there should be an interaction term included in the model

Almost always, we include the constant and linear terms in a model, although we might discover that they are not needed if other terms are added. The question is generally whether to include the quadratic and bilinear terms. In order to decide which of these terms to include in a model $f(x, y)$, it helps to ask the following questions:

1. Is there an extremum with respect to x? That is, holding y fixed, is there a value of x at which $f(x, y)$ takes on a maximum or minimum value? If there is, you will want to include the quadratic term in x.
2. If there is an extremum with respect to x, does its position or magnitude depend on the value of y? If so, include the interaction term.
3. If there isn't an extremum with respect to x, does the slope with respect to x depend on y? If so, include the interaction term.
4. The same questions should be asked with respect to y.

Decide which terms should be included in local models in these situations:

Bicycle speed A bicycle's speed V depends on both the steepness S of the terrain and the gear ratio G for the bicycle. Assume that the gear ratio is a number between 1 and 6, and let the steepness be measured in percent (positive for uphill, negative for downhill). What terms should be included in $V(S, G)$?

Economic production The output of a factory, P, depends both on the amount of capital C and the amount of labor L. What terms should be included in $P(C, L)$?

Infectious disease The number of people N who get an illness such as the flu depends on both the number of people who already have the illness I, and the number who are susceptible S. What terms should be included in $N(S, I)$?

Survival of chicks The number of surviving fledglings F of a mother bird depends on the number of eggs N that are laid and the time that the mother spends collecting food T. What terms should be included in $F(N, T)$?

Day length The length of daylight D depends on both the time of the year M (for month) and the latitude L. What terms should be included in $D(M, L)$?

Growth of a crop The yield Y of a crop (bushels/acre) depends both on the amount of water applied (W, inches/acre) and the amount of fertilizer (F, lbs/acre). What terms should be included in $Y(W, F)$

Probability of admission to college The probability P that an applicant will be admitted to college depends on many things, but we will restrict consideration here to the math M and verbal V scores on an entrance examination such as the ACT or SAT. What terms should be included in $P(M, V)$?

School effectiveness The effectiveness E of an elementary school (perhaps as measured imperfectly by standardized tests) depends on both the qualifications of the teachers and the class size S. We'll crassly measure the across-the-district teacher qualification with the average teacher pay, P. What terms should be included in $E(S, P)$?

We can generalize this approach to more than two variables. Here is a function of (at least) three-variables.

Probability of a heart attach The probability of a heart attack p as a function of age A and amount of exercise E, and number of calories in the diet C. What terms should be included in $p(A, E, C)$