Polynomial Regression & Interactions

1. Polynomial Regression with One Predictor Variable

1. Formulation of the Model

A nonlinear relationship between Y and X can often be approximately represented within the general linear model as a polynomial function of X.

Example:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \varepsilon_i$$

may be represented as a linear model

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \varepsilon_i$$

with the transformed variables $X_{i1} = X_i$ and $X_{i2} = X_i^2$.

The *order* of the polynomial function is the highest exponent of x; the model above is a *second-order* model.

To estimate a polynomial function x is often first deviated from its mean (or median) to reduce collinearity between X and higher powers of X. A variable deviated from its mean is called *centered*. The transformation is $x = X - \bar{X}$ where x (lower case) represents the centered variable and X (uppercase) the original (uncentered) variable.

More on why we mean center: http://www.ats.ucla.edu/stat/mult_pkg/faq/general/curves.htm

A polynomial function can be used when

- the true response function is polynomial
- the true response function is unknown but a polynomial is a good approximation of its shape

2. Graphic Representation of the Model

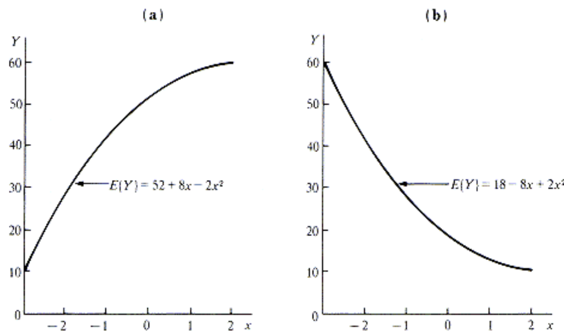
The response function E{y} for any polynomial model with one predictor variable can be represented on a 2-dimensional plot of Y against X.

A second degree polynomial implies a parabolic relationship. The signs of the coefficients determine the shape of the response function:

- when β_2 is positive, Y increases as the value of X increases
- when β_2 is negative, Y eventually decreases as the value of X increases

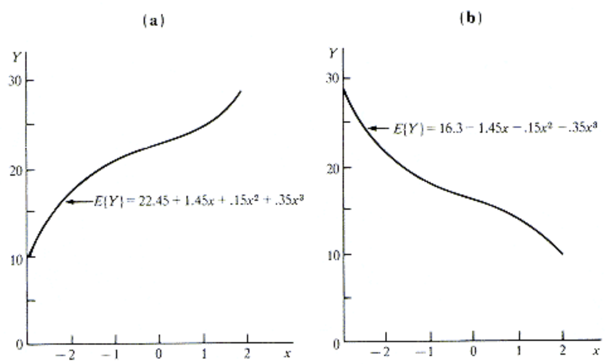
as shown in these graphs:

FIGURE 7.4 **Examples of Second-Order Polynomial Response Functions.**



Higher degree polynomials produce curves with more inflection points:

FIGURE 7.5 **Examples of Third-Order Polynomial Response Functions.**



When estimating a polynomial function, it is often useful to test for the joint significance of the coefficients of X, X^2, and higher powers of X, in addition to testing for the significance of each coefficient separately. In a joint test of significance one tests H_0: $\beta_1 = \beta_2 = 0$ against the alternative that at least one of the coefficient is not zero. Joint significance tests are explained in a later section.

NOTES

- when evaluating the shape of a polynomial response function, it is necessary to keep within the range of X in the data, as extrapolating beyond this range may lead to misleading predictions
- it is possible to convert from the coefficients of the centered model (involving x) to the non-centered model involving the original X (see page 301); however, the conversion is rarely needed for substantive purposes.

- fitting a polynomial regression with powers higher than three is rarely done as the interpretation of the coefficients becomes difficult and interpolation tends to become erratic. (A polynomial of order n - 1 can always be fitted exactly to n points.)
- polynomial regression models are often fitted with the *hierarchical approach* in which higher powers are introduced one at a time and tested for significance, and if a term of a high order is included (say, X^3) then all terms of lower order (X and X^2) are also included.

3. Interpretation of coefficients

- Interpretation of b_0 is predicted response at the mean of x (since x is mean centered)
- b_1 does not have a very helpful interpretation. It is the slope of the tangent line at the mean of x
- b_2 indicates the up/down direction of curve
  - $b_2 < 0$ means curve is concave down
  - $b_2 > 0$ means curve is concave up

3. Interactions

1. Formulation of the Model

Models that are not additive contain interaction effects. Interactions are commonly represented as cross-product terms called *interaction terms*.

The simplest interaction model is a special case (without the square terms) of the second-order polynomial model with two predictor variables with response function

$$E\{Y\} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2$$

The meaning of the regression coefficients β_1 and β_2 is not the same as it is in a model without interaction. In the interaction model, the change in $E\{Y\}$ with a unit increase in x_1 when x_2 is held constant is

$\beta_1 + \beta_3 x_2$

and the change in $E\{Y\}$ with a unit increase in x_2 when x_1 is held constant is

$\beta_2 + \beta_3 x_1$

Therefore in the interaction model the effect of both x_1 and x_2 depends on the level of the other variable. (So that the regression model is no longer additive.)

Interaction effects can be plotted (see SAS Code for details)