Fitting Data by Logistic Regression

Dr. Michael S. Pilant, Dept. of Mathematics, Texas A&M University

November 11, 2011

Abstract

A logistic model may fit observed data better than ordinary least squares, which is based on a linear model. In this paper, we look at conditions on data which indicate that a logistic model is appropriate.

1 Logistic Model

The two parameter logistic function is given by

\[ P(x) = \frac{e^{a(x-b)}}{1 + e^{a(x-b)}} = \frac{1}{1 + e^{-a(x-b)}} \]

\[ P'(x) = \frac{a e^{a(x-b)}}{(1 + e^{a(x-b)})^2} \]

We shall assume \( a > 0, b > 0 \). This function has several properties:

1. \( P(x) \) is defined for all values of \( x \).
2. \( 0 \leq P(x) \leq 1 \) for all \( x \).
3. \( \lim_{x \to -\infty} P(x) = 0 \)
4. \( \lim_{x \to \infty} P(x) = 1 \)
5. \( P(b) = 1/2 \)
6. \( P'(b) = a/4, P'(x) > 0 \) for all \( x \).
7. \( P''(b) = 0, P''(x) > 0 \) if \( x < b \) and \( P''(x) < 0 \) if \( x > b \)
2 Data Properties

If the data \((x_i, y_i)\) is to be fit by the logistic model, then it should have the following properties:

1. \(0 \leq y_i \leq 1\) on the interval \(\alpha \leq x_i \leq \beta\).
2. \(y_i\) should be nearly monotonic
3. max slope should occur near \(y_i = 1/2\)
4. \(y_i \to 0\) as \(x_i \to \alpha\)
5. \(y_i \to 1\) as \(x_i \to \beta\)
6. data should be concave up near \(\alpha\) and concave down near \(\beta\)

In other words, it should have the typical “S-curve” behavior.

The data will be bounded below by the green curve \((y = 0)\) on the left interval, and bounded above by the green curve \((y = 1)\) on the right interval. In fact this should reduce the distance between the logistic curve (red curve) and the actual data. For this particular problem, the relative error in approximating the green curve by the red curve is less than 7 percent!
In the next section, we will calculate the exact error for approximating the piecewise linear data curve. We will also generate random data by perturbing the theoretical data curve. Finally, we will take actual data from a number of MPE scenarios.

3 Examples