706

Modeling a function with a change point from experimental data

T. I. Vassilev
Dept. of Computer Science, Queen Mary and Westfield College,
Mile End road, London E1 4NS, UK
A. S. Smrikarov
Dept. of Computer Systems, Technical University,
8 Studentska St, 7017 Rousse, Bulgaria

September 5, 1995

Abstract

This paper describes a method for constructing a function with a discontinuity in its first derivative from experimental data. The main idea is to use F test to find the exact position of the change point.

Keywords: Change Point Detection, F test, Regression Analysis, Least Squares Approximation.

1 Introduction

The characteristics of many industrial objects are evaluated by a function with a discontinuity in its first derivative. Such an example is shown in Figure 1. The change point **q** subdivides the characteristics into two parts, each with a different type of behaviour. Part 1 is a straight line and part 2 is an arbitrary function (a polynomial of higher degree or a spline function). The corner is of a great importance for the tested object but unfortunately it cannot always be measured. Practical examples for such characteristics are the diesel engine performance as a function of the power, the hydro-mechanical gear performance. Similar cases are described by Jain and Jain (1994), Smith and Scariano (1990) and Lombard et al. (1990).

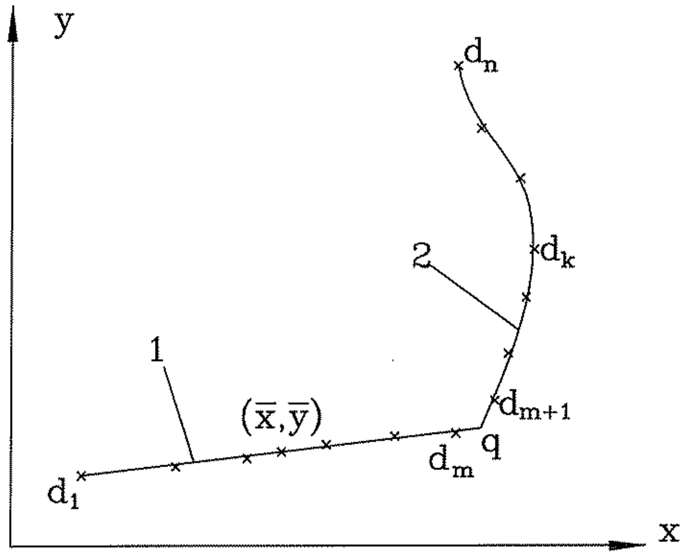

Figure 1: A function with a discontinuity in the first derivative

The objective of this paper is to develop a general method for subdividing the initial data set into two subsets, each one describing a different mathematical function. Then a least squares fit can be applied to the data sets and the change point $\mathbf{q}$ can be found as an intersection of the two functions.

2 Modeling the function

Let $\mathbf{D} = \{\mathbf{d}_i = (x_i, y_i), i = 1, \ldots, n\}$ be a noisy set of measured data points that describes an unknown function $y = g(x)$. The function $g(x)$ is continuous but not differentiable at the point $\mathbf{q}$ (see Figure 1) and consists of two different mathematical functions $g_1(x)$ and $g_2(x)$, where the type of $g_1(x)$ is known, e.g. polynomial of degree one or two. We want to find a continuous function $f(x)$ that is the best approximation from a certain class of functions to $g(x)$. The main difficulty here is to divide the set $\mathbf{D}$ into two subsets $\mathbf{D}_1 = \{\mathbf{d}_i, i = 1, \ldots, m\}$ and $\mathbf{D}_2 = \{\mathbf{d}_j, j = m+1, \ldots, n\}$, which describe the functions $g_1(x)$ and $g_2(x)$ respectively. We suggest the following algorithm:

1. Determine the type of the function $g_1(x)$. We recommend that it should be a polynomial of degree one or two but other functions could also be used.

2. Find a point $\mathbf{d}_k$ that is on $g_2(x)$. In Figure 1 this is the point with the biggest abscissa but the last point $\mathbf{d}_n$ will also serve well.

3. Approximate the set $\mathbf{d}_i, i = 1, \ldots, k$ with a polynomial $p_1(x)$ of degree chosen in step 1 using the least squares method, see Kennedy and Neville (1986).

4. Compute the variance

$$s_1^2 = \frac{1}{k-l} \sum_{i=1}^{k} (y_i - p_1(x_i))^2, \tag{1}$$

where l is the number of the coefficients of $p_1(x)$ ($l = 2$ for a first degree polynomial).

5. Approximate the set $\mathbf{d}_i, i = 1, \ldots, k-1$ with a polynomial $p_2(x)$ of the same degree using the least squares method.

6. Compute the variance

$$s_2^2 = \frac{1}{k-l-1} \sum_{i=1}^{k-1} (y_i - p_2(x_i))^2. \tag{2}$$

7. Compare the variances s_1^2 and s_2^2 using the F test.

   (a) Compute the distribution $F = s_1^2/s_2^2$.

   (b) Find the value $F_t = F_{k-l,k-l-1,\alpha}$ from the Table A.14, see Kennedy and Neville (1986). The level of significance α could be either of 0.05 or 0.01.

   (c) Compare F and F_t. If $F < F_t$ the null hypothesis $H_0 : s_1^2 = s_2^2$ cannot be rejected. This means that the two variances are statistically equal and the sets $\mathbf{d}_1, \ldots, \mathbf{d}_k$ and $\mathbf{d}_1, \ldots, \mathbf{d}_{k-1}$ are described equally badly by a polynomial chosen in step 1. Hence the point $\mathbf{d}_{k-1}$ lies also on the curve $g_2(x)$. Go to step 5 with $k := k-1$ and $s_1^2 := s_2^2$.

   (d) If $F > F_t$ the null hypothesis $H_0 : s_1^2 = s_2^2$ is rejected and the alternative hypothesis $H_a : s_1^2 > s_2^2$ is accepted. This means that a polynomial describes much better the data set $\mathbf{d}_1, \ldots, \mathbf{d}_{k-1}$ than the set $\mathbf{d}_1, \ldots, \mathbf{d}_k$. Hence the point $\mathbf{d}_{k-1}$ coincides with $\mathbf{d}_m$ and we have just performed an approximation to the function $g_1(x)$. Go to step 8 with $m := k - 1$.

Note: One cannot be sure that s_1^2/s_2^2 is really F distribution because the data sets are dependent. We use the F distribution only as an approximation to the real one, therefore it is a good idea to check that $F/F_t > 3$.

8. Find an approximation to $g_2(x)$ using the data set $\mathbf{D}_2$ and the least squares method. If the type of the function is not known then perform a B-spline least squares curve approximation described by de Boor (1978). In many practical cases the measurement of points in all intervals is not possible for physical reasons, e.g. unstable work of the tested object. This lack of data results in an approximating curve with unwanted undulations. In order to obtain a fair curve the approach of Vassilev (1995) can be applied. Its main idea is to insert additional data points that minimize the following integral

$$E = \int_{curve} [\alpha \dot{\mathbf{w}}(u) + \beta \ddot{\mathbf{w}}(u)] \, du, \tag{3}$$

where $\mathbf{w}(u)$ is the parametric representation of the curve and $\dot{\mathbf{w}}(u)$ and $\ddot{\mathbf{w}}(u)$ are its first and second parametric derivatives.

9. Find the corner $\mathbf{q}$ using some of the numerical methods, described by Press et al. (1989), for solving algebraic equations. If $g_1(x)$ is a straight line one can apply the approach of Farin (1993) for computing the intersection of a Bezier curve with a line. It can be adapted easily because a B-spline curve is actually a set of Bezier curves.

3 Error estimation

Since the corner **q** is very important for the quality of the tested object it is worth giving an estimation of the error in its prediction. It can be computed from

$$\varepsilon = \varepsilon_m + \varepsilon_a, \tag{4}$$

where ε_m is the error due to the measurement of the data points **D** and ε_a is the error due to the least squares approximation. The error due to the numerical solution of the algebraic equation is much smaller than the above two and it is neglected. Applying confidence interval estimation and regression analysis, see Kennedy and Neville (1986), one can write

$$\varepsilon_m = t_{N,\alpha} s_{max}, \tag{5}$$

where s_{max} is the maximal standard deviation of the data in **D** and N is the number of measurements for each point;

$$\varepsilon_a = t_{m,\alpha} s_y \sqrt{1 + \frac{1}{m} + \frac{(x_q - \bar{x})^2}{\sum (x_i - \bar{x})^2}}, \tag{6}$$

where $(\bar{x}, \bar{y})$ is the centroid point of the regression line (see Figure 1). When evaluating the overall error we considered that the error of the line approximation is much bigger than that of the B-spline approximation.

4 Results

The method was applied to a real data set acquired using the microcomputer system of Vassilev et al. (1990). The result is shown in Figure 2. This is a diesel engine performance visualizing the relation: speed (n), torque (M), fuel consumption (B) and specific fuel consumption (b_e) versus power (P). The change point here is the point where the speed controller turns on. Since it is well known that part 1 of the function $n = g(P)$ is a straight line, we can apply the above described method. The results are as follows:

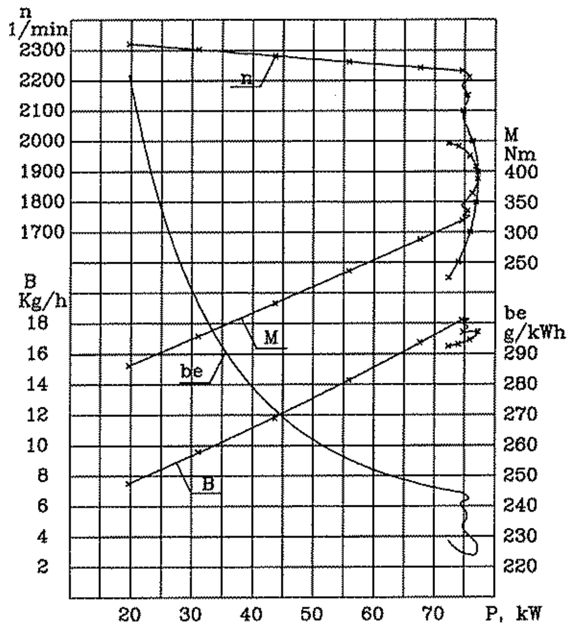

Figure 2: Diesel engine performance

1. A line describes best the first $m = 6$ points. The ratio F/F_t at level of significance $\alpha = 0.05$ is 21.3 which is much bigger than 3. The line equation is $n = 2349.6 - 1.3045P$.

2. The fair B-spline approximation of Vassilev (1995) is applied for part 2 of the performance.

3. The coordinates of the point **q** are $(74.5, 2232)$, which (with probability 0.95) are with a relative error less than 1.5%.

5 Discussion

A statistical method for reconstructing a function with a discontinuity in the first derivative from noisy data was described. Although the technique was presented only for one change point it can be applied without any limitations to functions with n change points in case we know the type of the relation in the first n intervals.

References

de Boor, C. (1978), *A Practical Guide to Splines*, Springer-Verlag, New York.

Farin, G. (1993), *Curves and Surfaces for Computer Aided Geometric Design*, Academic Press, Inc.

Jain, S. and Jain, R. K. (1994), An application of SPRT for Detecting Change Point in a Reliability Growth Model, *Microelectronics and Reliability*, Vol. 34(5), pp 811-814.

Kennedy, J. B. and Neville A. M. (1986), *Basic Statistical Methods for Engineers and Scientists*, Harper and Row, Publishers, New York.

Lombard, F., de Jager, O. C. and Schultz, D. M. (1990), The Detection of a Change Point in Periodic Gamma Ray Data, *Nuclear Physics B, Proceedings Supplements*, Vol. 14A, pp 285-290.

Press, W. H., Flannery, B. P., Teukolsky, S. A. and Vetterling W. T. (1989), *Numerical Recipes, The Art of Scientific Computations*, Cambridge University Press, New York.

Smith, T. M. and Scariano, S. M. (1990), Change-point Estimation Using Sample Means, *Mathematics and Computer Education*, Vol. 24(3), pp 217-224.

Vassilev, T. I., Petkov, G. B. and Smrikarov, A. S. (1990), Microcomputer System for Testing Internal Combustion Engines, *Masinostroene*, Vol. 4, Sofia.

Vassilev, T. I. (1995), A Method for Fair B-spline Interpolation and Approximation by Insertion of Additional Data Points, Submitted for publication in CAD.