Available at www.ElsevierMathematics.com
POWERED BY SCIENCE DIRECT

Journal of Statistical Planning and Inference 127 (2005) 143–156

journal of statistical planning and inference

www.elsevier.com/locate/jspi

Asymptotic theory for bent-cable regression—the basic case☆

Grace Chiu[a,b,*], Richard Lockhart[a], Richard Routledge[a]

[a] *Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6*

[b] *Pacific Institute for the Mathematical Sciences, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6*

Accepted 6 September 2003

Abstract

We use what we call the bent-cable model to describe potential change-point phenomena. The class of bent cables includes the commonly used broken stick (a bent cable without a bend segment). Theory for least-squares (LS) estimation is developed for the basic bent cable, whose incoming and outgoing linear phases have slopes 0 and 1, respectively, and are joined smoothly by a quadratic bend. Conditions on the design are given to ensure regularity of the estimation problem, despite non-differentiability of the model's first partial derivatives (with respect to the covariate and model parameters). Under such conditions, we show that the LS estimators (i) are consistent, regardless of a zero or positive true bend width; and (ii) asymptotically follow a bivariate normal distribution, if the underlying cable has all three segments. In the latter case, we show that the deviance statistic has an asymptotic chi-squared distribution with two degrees of freedom.

MSC: 62F12 (62J02)

Keywords: Change points; Segmented regression; Least-squares estimation; Asymptotic theory

☆ This research has been funded by the Natural Sciences and Engineering Research Council of Canada through a Postgraduate Scholarship to G. Chiu during her graduate studies at SFU, and Research Grants to R. Lockhart and R. Routledge.

* Corresponding author.

E-mail addresses: sgchiu@stat.sfu.ca, grace@pims.math.ca (G. Chiu), lockhart@stat.sfu.ca (R. Lockhart), routledg@stat.sfu.ca (R. Routledge).

0378-3758/$ - see front matter
doi:10.1016/j.jspi.2003.09.007

1. Introduction

Given known design points $x_1,\ldots,x_n$, we consider random responses $Y_1,\ldots,Y_n$ generated by the regression model

$$Y_i = q(x_i;\tau_0,\gamma_0) + \varepsilon_i, \qquad i=1,\ldots,n, \tag{1}$$

where

$$q(x;\tau,\gamma) = \frac{(x-\tau+\gamma)^2}{4\gamma}\mathbf{1}\{|x-\tau|\leqslant\gamma\} + (x-\tau)\mathbf{1}\{x-\tau>\gamma\} \tag{2}$$

is referred to as the *basic bent cable* (Fig. 1), and ε_i's are i.i.d. random errors with mean 0 and known, constant standard deviation σ.[1] We write $\boldsymbol{\theta}_0=(\tau_0,\gamma_0)$.

Least-squares (LS) estimation of $\boldsymbol{\theta}_0\in\Omega=(-\infty,M]\times[0,\infty)$ on the open regression domain $\mathscr{X}=\mathbb{R}$ is considered. Here, M is some finite positive upper bound (large) for the candidate τ-values. Zero is the natural lower bound for candidate γ-values. Any basic bent cable $q(x;\boldsymbol{\theta})$ for $\boldsymbol{\theta}\in\Omega$ is a candidate model. In this article, we prove, given a set of conditions on the location of the design points $x_1,\ldots,x_n$ (Section 3), that the least-squares estimators (LSEs) for τ_0 and γ_0 are consistent when $\gamma_0\geqslant 0$, and asymptotically follow a bivariate normal distribution when $\gamma_0>0$. Asymptotic distributional properties for the case of $\gamma_0=0$ appear in Chiu et al. (2002a). A bent cable with free slope parameters is required in practice. The full bent-cable model can be written as $f(x;\beta_0,\beta_1,\beta_2,\tau,\gamma)=\beta_0+\beta_1 x+\beta_2 q(x;\tau,\gamma)$. This article is intended to provide a framework for the complex estimation theory associated with the full model.

Seber and Wild (1989, Chapter 9) have suggested employing the class of bent-cable models—which includes the piecewise-linear "broken-stick" model when $\gamma=0$—in situations where both smooth and sharp transitions are plausible. However, modeling change phenomena by the broken stick remains common (Barrowman and Myers, 2000; Naylor and Su, 1998; Neuman et al., 2001). Numerical instability due to the non-differentiability of this model prompted Tishler and Zang (1981) to develop (2). Their introduction of a "phoney" bend of fixed, non-trivial width γ to replace the kink at τ was a computational tactic. Upon numerical convergence, γ would be ignored.

However, when no law of nature or auxiliary knowledge is available to support an abruptness notion, a broken-stick fit would encourage the investigator to look for sources of change associated with the sole value of $\hat{\tau}$. In contrast, the bent cable incorporates γ as part of the parametric model. It generalizes the broken stick by removing the a priori assumption of a sharp threshold, allowing for a possibly gradual transition. A bent-cable fit would point to one or more sources of change whose influence took hold gradually over a certain covariate range. Thus, it helps to avoid data misinterpretation due to possible over-simplification of the nature of change. We call it the "bent cable" due to the smooth bend as opposed to a sharp break in a snapped stick. The performance of bent-cable regression for assessing the abruptness of change is discussed in Chiu et al. (2002b).

[1] In practice, estimation of σ may be required. Chiu (2002) shows that the results of this article extend to LS estimation assuming unknown σ, and that the LSE of σ is consistent.