

# A Note on Linear and Second Order Significance Testing in Nonlinear Models

Michael Brimacombe<sup>1</sup>

<sup>1</sup> University of Kansas Medical Center, Kansas City, Kansas, USA

Correspondence: Michael Brimacombe, Department of Biostatistics, University of Kansas Medical Center, Kansas City, KS, 66109, USA. Tel: 1-913-588-4703. E-mail: mbrimacombe@kumc.edu

Received: September 15, 2015 Accepted: October 23, 2015 Online Published: November 19, 2015

doi:10.5539/ijsp.v5n1p19 URL: <http://dx.doi.org/10.5539/ijsp.v5n1p19>

## Abstract

A modified ANOVA setting is used to develop a one degree of freedom test for intrinsic curvature effect. The approach provides clear cutoffs for the measurement of intrinsic curvature effect. It is shown that the significance of the correction factor itself and its usefulness in correcting a global test based on linear approximation are distinct elements. They depend on the local frame of reference provided by the model-data combination and null hypothesis considered.

**Keywords:** P-values, nonlinear regression, ANOVA, intrinsic curvature, Wald statistic

## 1. Introduction

Nonlinear models are applied in many areas of science, often reflecting scientific theory, for example in toxicology, cancer research and econometrics. Procedures for statistical inference are typically described in terms of Wald statistics, corresponding to the use of a local linear approximation of the regression surface, or likelihood based regions corresponding to use of the likelihood function and likelihood ratio (Seber and Wild, 1989). While likelihood ratio based inference procedures are preferred theoretically, use of Wald statistics in nonlinear regression is common as likelihood ratio based regions can be difficult to interpret (see Bates and Watts, 1988).

The presence of significant nonlinearity or local curvature in the regression surface when a linear approximation is used creates a type of model mis-specification and can affect the accuracy of p-values and confidence regions (Donaldson and Schnabel, 1987). This requires some correction to the relevant test statistics, typically based on squared lengths and related sums of squares. Note that nonlinear models often represent discipline specific theories or solutions to differential equations and have a specified structure and parameterization. They are often more difficult to analyze than linear models. A typical nonlinear regression model can be written;

$$y_i = \eta(\mathbf{x}_i, \beta) + \varepsilon_i \quad (1)$$

$i = 1, \dots, n$  where  $\mathbf{x}_i$  are fixed values of the explanatory variable  $\mathbf{x}$ , the model function  $\eta$  is of known form depending on the parameter vector  $\beta \in \mathbf{R}^p$  and  $\mathbf{x}_i$ , and the  $\varepsilon_i$  are independent error terms, each normally distributed with mean zero and variance element  $\sigma^2$ . Some typical examples are given in Table 1.

Table 1. Standard Nonlinear Models

Model	$\eta(\mathbf{x}_i, \beta)$
Logistic	$\beta_1 / (1 + \beta_2 \exp(\beta_3 x))$
Richards	$\beta_1 / [1 + \beta_2 \exp(-\beta_3 x)]^{1/\beta_4}$
Gompertz	$\beta_1 \exp[-\exp[\beta_2 - \beta_3 x]]$
Michaelis-Menten	$\beta_1 x / (\beta_2 + x)$

To assess possible approximation error, Bates and Watts (1980, 1988), following the work of Beale (1960) developed relative measures of the intrinsic curvature of the regression surface based on comparing differences between linear and second order approximations. Details can be found in Seber and Wild (1989), Chapter 4. The work of Beale can be viewed as developing an average measure of relative curvature, while the Bates and Watts measures reflect a measure of maximum relative curvature, assessed over all possible directions of the response vector  $y$  in relation to  $\eta(\mathbf{x}_i, \beta)$  evaluated at the *m.l.e.*. Note that intrinsic curvature effects in differential geometry average minimum and maximum projection lengths onto the principal axes at the point in question (Kreyszig, 1991), thus the values suggested by Bates and Watts are

not truly geometric curvatures. They applied their measure to many nonlinear models, finding their measure insignificant in almost all (Bates and Watts, 1988).

Here the intrinsic curvature of a regression surface is revisited within a local frame of reference; the observed direction of departure, null hypothesis of interest and data based properties of the nonlinear model in the direction of departure only. Within this context, an intrinsic curvature correction can be given precise calibration, a result lacking in earlier initial development (Seber and Wild (1989), p. 218), and the results organized in an ANOVA table. Significance testing of both the intrinsic curvature effect and related corrected global test are developed. It is seen that less than significant intrinsic curvature effects may still lead to significant adjusted global tests, depending on the local frame of reference considered. It is also the case that significant intrinsic curvature effect may only moderately correct the global test, implying that the intrinsic curvature effect need not be useful even when it exists and vice-versa. These elements are distinct in their significance, depending on the local frame of reference, model-data combination and null hypothesis considered.

Corrections for model mis-specification are often first stage adjustments of the assumed model using observed data. Such first stage adjustments in applied statistical modeling include propensity scores (Luellen et al., 2005), Box-Cox transformations (Box and Cox; 1964, 1982), the selection of link functions (McCullagh and Nelder, 1972), the selection of ridge regression constants (Hoerl and Kennard, 1970), variance stabilizing transformations and log transformations to remove interaction effects (Kutner et al., 2005). Here we locally adjust, using the differential geometric curvature of the surface, the sums of squares in the ANOVA table and adjusting the projection length related to the observed data only. There is no need to assess over all possible curvature related adjustments just as there is no need to average over all possible Box-Cox reparameterizations or variance stabilizing transformations.

Here intrinsic curvature effects are defined only in relation to the direction of the projection length of the observed response  $y$  onto the linear approximation by adding the projection onto the intrinsic curvature vector  $v$  defined at the point  $\eta(x; \beta_0)$ . An ANOVA based testing approach is then developed for tests of  $H_0 : \beta = \beta_0$ , using  $v$  to define an independent dimension in the table. It is argued here that first stage intrinsic curvature correction should reflect the observed direction of departure. It is only in this direction that the (intrinsic) curvature of the regression surface is relevant to modification of the relevant sums of squares and the determination of the Wald statistic value based on the assumed null value  $\eta(x; \beta_0)$  and the *m.l.e.*  $\eta(x; \hat{\beta})$ . The linear approximation here is taken at  $\beta = \beta_0$  not at the maximum likelihood value, giving the analysis a focused local aspect.

Note that a "parameter-effects" curvature measure (Bates and Watts, 1980; Cook and Witmer, 1985) is sometimes used as a formal diagnostic for indicating when re-parameterisation of the nonlinear model should be considered. This measure is technically the leftover second order effect after subtracting the Bates and Watts intrinsic curvature measure from the Hessian matrix evaluated at the *m.l.e.*. It can be interpreted as reflecting the degree of non-parallel parameter coordinates on the linear approximating surface examined over all directions (Cook and Witmer, 1985). It requires the assumption of normality and is useful only in smaller sample sizes, not agreeing very well with other measures of curvature defined for asymptotic likelihood calculations (Cook and Witmer, 1985). The standard errors in the ANOVA table here are based on the squared length of the projection orthogonal to the linear surface and are independent of parameter effects. As we do not examine all directions, the linearity of the parameter contours along the surface are not relevant to the analysis.

In this paper, a local intrinsic curvature measure is developed that can be viewed as a first stage modification and used to more accurately test global hypotheses. The inferential approach taken here is based on the Wald statistic, using linear approximation and squared projecting lengths with intrinsic curvature based correction factors. In the sections that follow, the basic geometry of the nonlinear regression model related to first and second order approximation is briefly reviewed. A breakdown of orthogonal projections available to assess significance of local surface curvature and related ANOVA tables is given. A corrected global test corrected for intrinsic curvature and one-degree-of-freedom test for curvature effect are then obtained in an easily interpreted ANOVA table. It is seen that less than significant intrinsic curvature effects may still lead to significant adjusted global tests, depending on the observed data  $y$  and null hypothesis in question. Locally defined intrinsic curvature effects are unique to the specific nonlinear model, null hypothesis and dataset considered. In this sense the significance of first and second order elements are distinct. Several examples are discussed.

## 2. Method

Consider the standard nonlinear regression model given above  $y_i = \eta(\mathbf{x}_i, \beta) + \varepsilon_i$ . The set of possible mean values generated by considering all possible  $\beta$  values defines the surface  $\mathbf{N} = \{\eta(\beta) : \beta \in \Omega \subseteq \mathbf{R}^p\}$ , where  $\Omega$  is the parameter space, the explanatory variable  $\mathbf{x}_i$  is suppressed in the present notation and  $\eta(\beta)$  is the  $n \times 1$  column vector with  $i^{th}$  component given by  $\eta(\mathbf{x}_i, \beta)$ . This surface is assumed to be a locally smooth manifold, differentiable to the third order with respect to the coordinates  $\beta_j$ ,  $j = 1, \dots, p$  at each point on the surface. In particular it follows that the tangent plane at any point  $\eta(\beta_0)$  on  $\mathbf{N}$  is uniquely defined as well as second derivatives, necessary to calculate surface curvature.

## 2.1 Linear and Quadratic Approximation

The geometry of local linear and second order approximation are briefly reviewed. See for example Seber and Wild (1989).

### 2.1.1 Local Linear Approximation

The focus here is on approximating  $\eta(\beta)$  about the null value  $\eta(\beta_0)$ . Let  $F_0$  be the  $n \times p$  matrix with column elements given by  $\mathbf{f}_i = \partial\eta(\beta)/\partial\beta_i|_{\beta=\beta_0}$  for  $i = 1, \dots, p$ , and define  $\mathbf{L}(F_0)$  to be the linear span of  $F_0 = \{\mathbf{f}_1, \dots, \mathbf{f}_p\}$ . Thus  $\mathbf{L}(F_0)$  is the tangent plane to the manifold  $\mathbf{N}$  defined at  $\eta(\beta_0)$ . Further, let  $P_0 = F_0(F_0'F_0)^{-1}F_0'$  be the orthogonal projection matrix for  $\mathbf{L}(F_0)$ , evaluated at  $\beta = \beta_0$ , then  $P_0(\mathbf{y} - \eta(\beta_0))$  is the projection of  $(\mathbf{y} - \eta(\beta_0))$  onto the tangent plane at  $\beta = \beta_0$ . Let  $\mathbf{u} = P_0(\mathbf{y} - \eta(\beta_0))/\|P_0(\mathbf{y} - \eta(\beta_0))\|$ , where  $\|\cdot\|$  denotes length. This unit length vector, centered at  $\eta(\beta_0)$ , lies along the tangent plane, in the direction of the projected departure. It allows us to determine curvature in this direction.

### 2.1.2 Second Order Properties and Local Curvature

The quadratic approximation to  $\eta(\beta)$  at  $\beta = \beta_0$  using Taylor expansion is given by  $\eta(\beta) - \eta(\beta_0) = F_0\theta + (\frac{1}{2})\theta'H_0\theta$  where  $H_0$  is the Hessian  $p \times p$  matrix with vector elements  $\mathbf{h}^{ij} = \frac{\partial^2\eta(\beta)}{\partial\beta_i\partial\beta_j}$  evaluated at  $\beta = \beta_0$  and  $\theta = (\beta - \beta_0)$ . The local second order properties of  $\eta(\beta)$  in the specific direction  $\mathbf{u}$  are used here to define local relative curvature effects. The intrinsic acceleration vector in the direction  $\mathbf{u}$  can be defined as  $(I - \mathbf{u}\mathbf{u}')\eta''(\beta)$  and represented as  $-\mathbf{v}/\rho$ , where  $\mathbf{v}$  is the unit vector chosen perpendicular to the acceleration vector in the direction  $\mathbf{u}$  and  $\rho = \rho(\beta_0)$  is the radius of curvature at  $\eta(\beta_0)$ . Note that  $\rho$  depends by definition on  $\mathbf{u}$  and is defined by;

$$-\mathbf{v}/\rho = \frac{\theta'G_0\theta}{\|F_0\theta\|^2}$$

where  $G_0 = [(I - P_0)h_{ij}]$ . Taking the norm of this we have;

$$\kappa = 1/\rho = \frac{\|\theta'G_0\theta\|}{\|F_0\theta\|^2} \quad (2)$$

where all matrices are evaluated at  $\beta = \beta_0$  and the local intrinsic curvature  $\kappa$ , the inverse radius of  $\eta(\beta)$  at  $\beta = \beta_0$  in the direction  $\mathbf{u}$ , is given here by  $1/\rho$ . Note that for statistical interpretation these need to be compared to the underlying *SSE* in the fitted model and data. Note further that the intrinsic curvature is invariant to 1 – 1 reparameterizations of the parameter space (Seber and Wild, p. 692). The intrinsic relative curvature measure of Bates and Watts is similar in basic definition to (2), see for example Seber and Wild, p. 131, however the tangent to the solution locus is taken at  $\eta(\hat{\beta})$ , the *m.l.e.* for  $\eta(\beta)$  and their measure of curvature is a maximum measure of curvature measured over the entire set of directions out from  $\eta(\hat{\beta})$ .

Here the relative curvature reflects a localized frame of reference in the specific direction provided by the global null hypothesis and related linear approximation about  $\eta(\beta_0)$ . This better reflects the nature of the problem as  $\eta(\beta)$  is fixed,  $x$  is fixed,  $\beta = \beta_0$  and the tangent plane to be used to calculate the squared length of the orthogonal projection of  $(\mathbf{y}^{obs} - \eta(\beta_0))$  is also fixed once  $\mathbf{y}$  is observed. Averaging curvatures or considering them across many directions does not make sense in terms of the differential geometry nor the nature of these surfaces which can vary greatly in form and related curvature across different directions.

## 2.2 Modified Projection Lengths in ANOVA

As noted in Seber and Wild (1989) Chapter 2, inference based on the linear approximation where  $\sigma^2$  is unknown uses the following exact  $\alpha$ -level test procedure;

$$\|P(\mathbf{y} - \eta(\beta))\|^2 / \|(I - P)(\mathbf{y} - \eta(\beta))\|^2 \leq \frac{P}{n - p} F_{p, n-p, \alpha} \quad (3)$$

and the set of  $\beta$  values satisfying (3) yields a  $100(1 - \alpha)\%$  exact confidence region for  $\beta$  (with the underlying linear model itself being approximate). As noted in Section 2,  $P = F(F'F)^{-1}F'$  is the projection matrix to  $\mathbf{L}(F)$  and  $(I - P)$  is the projection matrix to  $\mathbf{L}^\perp(F)$ . With the assumption of normality, the derivation of (3) is formally a consequence of considering the efficient scores (Cox and Hinkley, 1974, p. 324) which are  $F'(\mathbf{y} - \eta(\beta))/\sigma^2$ . Geometrically, (3) is a ratio of squared projection lengths; the projection of  $(\mathbf{y} - \eta(\beta_0))$  onto respectively the tangent plane approximation  $\mathbf{L}(F)$  and  $\mathbf{L}^\perp(F)$  the space orthogonal to  $\mathbf{L}(F)$  in  $\mathbf{R}^n$ . See Table 1.

### 2.2.1 Modified Orthogonal Decomposition

The usual orthogonal decomposition of regression and error can be replaced with a more detailed orthogonal decomposition  $\mathbf{y} = z_1\mathbf{u} + z_2\mathbf{v} + \mathbf{V}z_3$  where the residual space is viewed as being spanned by the curvature vector  $\mathbf{v}$  and the column

vectors of  $\mathbf{V}$ , which are any set of orthonormal vectors spanning the remaining dimensions orthogonal to both the tangent plane and the curvature vector  $\mathbf{v}$  evaluated at  $\beta = \beta_0$ . The curvature in the direction  $\mathbf{u}$  at  $\beta = \beta_0$  can be assessed by considering the orthogonal projection(s) of  $(\mathbf{y} - \eta(\beta_0))$  onto the linear and quadratic components in this direction. This is presented in Table 3 where  $P_v = \mathbf{v}\mathbf{v}'$  is the projection matrix onto the (normed) intrinsic curvature vector  $\mathbf{v}$ .

Table 2. Basic ANOVA Decomposition Using Linear Approximation

Subspace	Dimension	Squared Length of Projection
$\mathbf{L}(F)$ (Regression)	$p$	$\ P_0(\mathbf{y} - \eta(\beta_0))\ ^2$
$\mathbf{L}^\perp(F)$ (Residual)	$n - p$	$\ (I - P_0)(\mathbf{y} - \eta(\beta_0))\ ^2$
$\mathbf{R}^n$	$n$	$\ \mathbf{y} - \eta(\beta_0)\ ^2$

To test  $H_0 : \beta = \beta_0$  within the approximating linear model, without reference to curvature, the Regression component of Table 2 can be used to generate an F-test with  $p$  and  $(n - p)$  degrees of freedom. Normality is assumed here and as the projections are orthogonal by construction, the usual theorems regarding squared lengths and quadratic forms apply (see for example Seber and Wild, p. 24).

In particular, assuming  $\mathbf{y} \sim N((\eta(\beta_0), \sigma^2 I))$  where  $\sigma^2$  is unknown, we have under the null;

$$\frac{\|P_0(\mathbf{y} - \eta(\beta_0))\|^2 / p}{\|(I - P_0)(\mathbf{y} - \eta(\beta_0))\|^2 / (n - p)} \sim F_{p, n-p}. \quad (4)$$

with large values of the test statistic leading to rejection of  $H_0 : \beta = \beta_0$ . Note again that in the terminology of nonlinear models, this is an exact result, but its accuracy depends on the relevance of the local approximating tangent plane.

### 2.2.2 One-Degree-of-Freedom Test for Curvature

With moderate sample sizes, we can determine whether curvature adjustment for the test applied in Table 2 is necessary and the orthogonal decomposition given in Table 3 can be used to obtain a test of significance for local curvature defined in the direction  $\mathbf{u}$  based on the relative squared length of the orthogonal projection onto the curvature vector  $\mathbf{v}$ . The arguments underlying (3) give;

$$\frac{\|P_v(\mathbf{y} - \eta(\beta_0))\|^2}{(\|(I - P_0)(\mathbf{y} - \eta(\beta_0))\|^2 - \|P_v(\mathbf{y} - \eta(\beta_0))\|^2) / (n - (p + 1))} \sim F_{1, n-(p+1)} \quad (5)$$

where a large value for the test statistic reflects a pronounced projection length on the curvature vector  $\mathbf{v}$  in the direction  $\mathbf{u}$ . In the context of the approximating linear model, this is an exact test when the error are independent and normally distributed.

### 2.2.3 Adjusted Sums of Squares and F-Test

The linear regression component can be adjusted to account for second order effect by "borrowing" the squared length projection onto the local curvature vector  $\mathbf{v}$  from the residual space projection and adding it to the squared length projection of the linear component, giving an adjusted global test;

$$\frac{(\|P_0(\mathbf{y} - \eta(\beta_0))\|^2 + \|P_v(\mathbf{y} - \eta(\beta_0))\|^2) / (p + 1)}{(\|(I - P_0)(\mathbf{y} - \eta(\beta_0))\|^2 - \|P_v(\mathbf{y} - \eta(\beta_0))\|^2) / (n - (p + 1))} \sim F_{p+1, n-(p+1)}.$$

See Table 3. This uses the properties of normality, orthogonality implying independence and the sum of independent chi-square distributed variables having a chi-squared distribution.

Both the curvature test and adjusted global test can be carried out as their respective projection lengths and significance levels are not directly correlated. This implies that even if not significant, the curvature related projection used as a correction factor may affect the significance of the linear regression component. If a specific sequence of global tests is to be examined, for example to study the curvature properties of the nonlinear regression surface over a specific region, then for each respective null hypothesis a new ANOVA table should be generated as the correction to the projected sum of squares components will alter depending on the direction  $\mathbf{u}$  and associated intrinsic curvature vector  $\mathbf{v}$ . This will depend where on the actual regression surface the linear approximation is developed and the relative position of the response  $y$  in relation to the linear approximation. Here we work in relation to local properties defined about  $\eta(\beta_0)$ .

Table 3. Anova with Local Second Order Adjusted Sum of Squares

Subspace	Dimension	Squared Length of Projection
Regression	$p$	$\ P_0(\mathbf{y} - \eta(\beta_0))\ ^2$
Residual	$n - p$	$\ (I - P_0)(\mathbf{y} - \eta(\beta_0))\ ^2$
Curvature	1	$\ P_v(\mathbf{y} - \eta(\beta_0))\ ^2$
Regression + Curvature	$p + 1$	$\ P_0(\mathbf{y} - \eta(\beta_0))\ ^2 + \ P_v(\mathbf{y} - \eta(\beta_0))\ ^2$
Residual*(w/o Curvature)	$n - (p + 1)$	$\ (I - P_0)(\mathbf{y} - \eta(\beta_0))\ ^2 - \ P_v(\mathbf{y} - \eta(\beta_0))\ ^2$
$\mathbf{R}^n$	$n$	$\ \mathbf{y} - \eta(\beta_0)\ ^2$

### 3. Results

#### 3.1 Example 1

We apply these concepts in the context of an asymptotic growth model applied to the BOD dataset found in Bates and Watts (1988). This displays curvature related properties and is given by;

$$y_i = \beta_1(1 - \exp(-\beta_2 x_i)) + \varepsilon_i.$$

The dataset is given in Table 4.

Table 4. BOD Data

Demand	8.3	10.3	19	16	15.6	19.8
Time	1	2	3	4	5	7

A plot of the original data is shown in Diagram 1.

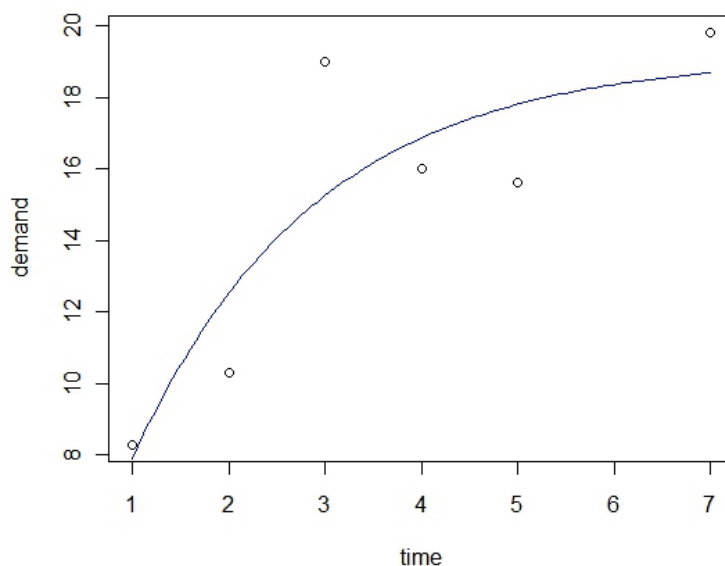


Diagram 1. Plot of Data and Fitted Nonlinear Regression Line

The log-likelihood is plotted in Diagram 2.

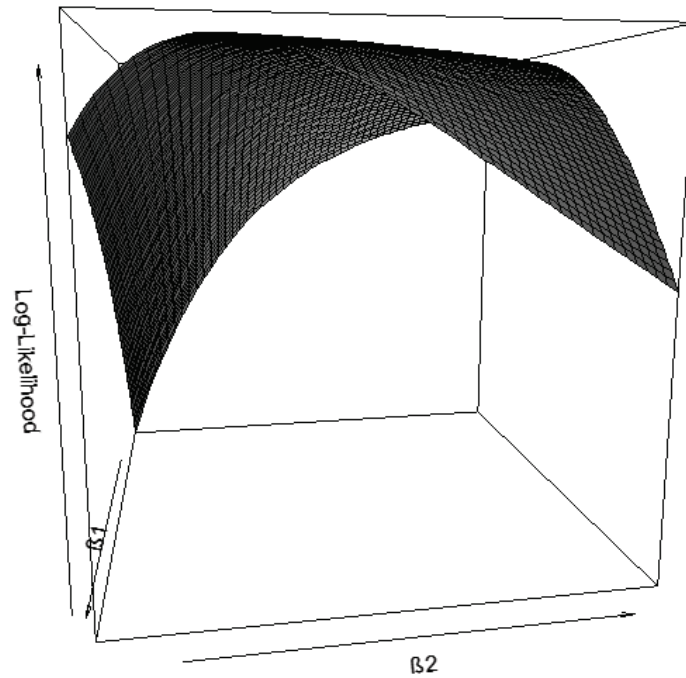


Diagram 2. Log-Likelihood for BOD Model and Data

As discussed in Bates and Watts (1988), p. 203, the non-standard behaviour of this model can be observed by considering the model as  $\beta_2 \rightarrow \infty$ . In this case the model converges to  $\beta_1$  which is estimated by  $\bar{y}$ . It follows that likelihood regions are open at confidence levels above 94% in the  $\beta_2$  direction. The first order derivatives are given here by an  $n$  by 2 matrix;

$$F = [(1 - e^{-\beta_2 x_i}), \beta_1 x_i e^{-\beta_2 x_i}]$$

for  $i = 1, \dots, n$ . The related 2 by 2 by  $n$  second order Hessian matrix is given by;

$$H = \begin{bmatrix} 0 & x_i e^{-\beta_2 x_i} \\ x_i e^{-\beta_2 x_i} & -\beta_1 x_i^2 e^{-\beta_2 x_i} \end{bmatrix}$$

where each  $h_{jk}$ ,  $j = 1, 2$  and  $k = 1, 2$  is an  $n$ -dimensional vector. The 0 value denotes a linear aspect to the model in certain directions. with the regression model here sometimes termed as partially linear.

The maximum likelihood or least squares value for  $(\beta_1, \beta_2)$  is given by  $(\hat{\beta}_1 = 19.143, \hat{\beta}_2 = 0.5311)$  with  $s = 6.498$  on 4 degrees of freedom. The Bates and Watts overall maximum relative curvature is given by 0.184 and can be interpreted versus an F-statistic based cutoff of 0.49, implying overall curvature effects exist in some direction. Figure 2 shows the presence of curvature primarily in the  $\beta_2$  direction as expected.

Three global null hypotheses are considered here to examine the effects of nonlinearity at various points on the regression surface. We take  $\beta_1 = 18$  for all three, a minor perturbation from the observed  $\hat{\beta}_1$  least squares value, and  $\beta_2 = 1.5, 3.5, 4.0$  respectively. This gives a set of values for which the intrinsic curvature increases sequentially. The applied results of the ANOVA based testing in Tables 2 and 3 are summarized in Table 5. These include the uncorrected linear test, the test for curvature effect and the curvature corrected global test.

For the initial hypothesis  $(\beta_1, \beta_2) = (18, 1.5)$ , the linear approximation based regression  $SS$  gives a non-significant result. The test for curvature is also non-significant and the adjusted regression does not attain significance. For the hypothesis  $(\beta_1, \beta_2) = (18, 3.5)$  the intrinsic curvature alters ( $\rho = 8.79$  to  $\rho = 25.58$ ). While the regression  $SS$  based on linear approximation alone remains non-significant, the curvature test is close to significant and the curvature adjusted regression is also close to significant. For the third hypothesis  $(\beta_1, \beta_2) = (18, 4.0)$  the relative curvature is  $\rho = 49.79$ , the curvature test remains close to significant and the curvature adjusted regression  $SS$  becomes significant. The  $SSE$  value also alters as the hypothesized value moves away from the least squares value, as does the overall  $SST$  or squared length of  $(\mathbf{y} - \eta(\beta_0))$ , reflecting the importance of the relative positioning of the response  $y$  in relation to the regression surface  $\eta(\beta)$ .

These considerations serve as useful, detailed diagnostics for curvature related effects and can be applied to any specified  $\eta(\beta_0)$  value.

Table 5. ANOVA Tables for Specified Global Hypotheses with BOD Model and Data

Table 5a: ( $\beta_1 = 18, \beta_2 = 1.5$ )					
Source	df	SS	MS	F	p-value
Regression	2	62.58	31.29	4.14	0.11
Residual	4	30.23	7.56		
Curvature	1	5.24	5.24	0.63	0.49
Regression + Curvature	3	62.58 + 5.24	22.60	2.71	0.22
Residual*(w/o Curvature)	3	24.99	8.33		
$R^2$	6	92.81			

Table 5b: ( $\beta_1 = 18, \beta_2 = 3.5$ )					
Source	df	SS	MS	F	p-value
Regression	2	106.15	53.08	4.19	0.105
Residual	4	50.73	12.68		
Curvature	1	35.40	35.40	6.93	0.078
Regression + Curvature	3	106.15 + 35.40	47.18	9.23	0.05
Residual*(w/o Curvature)	3	15.33	5.11		
$R^2$	6	156.88			

Table 5c: ( $\beta_1 = 18, \beta_2 = 4.0$ )					
Source	df	SS	MS	F	p-value
Regression	2	108.27	54.14	4.11	0.11
Residual	4	52.73	13.18		
Curvature	1	38.20	38.20	7.89	0.068
Regression + Curvature	3	108.27 + 38.20	48.82	10.09	0.044
Residual*(w/o Curvature)	3	14.53	4.84		
$R^2$	6	161.00			

### 3.1 Example 2

A simulated dataset in Bates and Watts (1980) is used to examine the Michaelis-Menten model given by;

$$y_i = \frac{\beta_1 x_i}{(\beta_2 + x_i)} + \varepsilon_i$$

where  $\varepsilon_i$  are *i.i.d.*  $N(0, \sigma^2)$ . The data is given in Table 6.

Table 6. Michaelis-Menten Simulated Data

y	76	47	97	107	123	139	159	152	191	201	207	200
x	.02	.02	.06	.06	.11	.11	.22	.22	.56	.56	1.10	1.10

The first order derivatives are given here by an  $n$  by 2 matrix;

$$F = \left[ \frac{x_i}{(\beta_2 + x_i)}, \frac{-\beta_1 x_i}{(\beta_2 + x_i)^2} \right]$$

for  $i = 1, \dots, n$ . The related 2 by 2 by  $n$  second order Hessian matrix is given by;

$$H = \begin{bmatrix} 0 & \frac{-x_i}{(\beta_2 + x_i)^2} \\ \frac{-x_i}{(\beta_2 + x_i)^2} & \frac{2\beta_1 x_i}{(\beta_2 + x_i)^3} \end{bmatrix}$$

where each  $h_{jk}$ ,  $j = 1, 2$  and  $k = 1, 2$  is an  $n$ -dimensional vector. The parameter estimates are given by  $(\widehat{\beta}_1, \widehat{\beta}_2) = (212.68, 0.064)$ . The hypotheses value  $(\beta_1, \beta_2) = (200, 0.3)$  is examined which is relatively far from the *m.l.e.* value. We obtain Table 7.

Table 7. ANOVA Table for Specified Global Hypothesis

$(\beta_1 = 200, \beta_2 = 0.3)$					
Source	df	SS	MS	F	p-value
Regression	2	4190.56	2095.28	13.15	0.002
Residual	10	1593.74	159.37		
Curvature	1	767.05	767.05	8.35	0.018
Regression + Curvature	3	4190.56 + 767.05	1652.54	17.99	0.0004
Residual*(w/o Curvature)	9	826.69	91.85		
$R^n$	12	5784.30			

The *SSR* element here is clearly significant. We can also see that there is a small but significant relative curvature effect for the  $\eta(\mathbf{x}_i, \beta_0)$  value considered and this only moderately corrects the significant regression element, implying that the intrinsic curvature effect need not be useful even when it exists. This analysis can also be applied for various values and regions of interest along the regression surface. It is worth noting that the Michaelis-Menten model can be re-expressed and re-parameterized as;

$$y_i = \frac{\beta_1 x_i}{(\beta_2 + x_i)} = \frac{x_i}{((\beta_2/\beta_1) + (x_i/\beta_1))}$$

$$1/y_i = 1/\beta_1 + (\beta_2/\beta_1)(1/x_i)$$

Letting  $y_i^* = 1/y_i$  and  $x_i^* = 1/x_i$  this has a linear form. Note that this implies the restrictions  $\theta_1 > 0$ ,  $y_i > 0$  and  $x_i > 0$ , which may not be acceptable in a given scientific setting, limiting re-expression of the regression model.

#### 4. Discussion

First stage modifications of sums of squares are used here to improve the global testing of parameters in a nonlinear regression model with potential intrinsic curvature effects. A one degree of freedom test for intrinsic curvature effect is also developed and the sums of squares adjustments presented in an easily interpreted ANOVA format. The local nature of intrinsic curvature is emphasized with a focus on the specific linear approximation and related observed direction of departure. The linear approximation here is taken at  $\beta = \beta_0$  not at the maximum likelihood value, giving the analysis a more focused local aspect. The use of ANOVA and focus on observed direction of departure provide clearly defined cutoffs for significant testing of the intrinsic curvature effect and corrected sums of squares.

Each individual model-data combination and selected null hypothesis may reflect intrinsic curvature effects which can correct and yield global test significance, even when they themselves are not formally significant. It is also the case that significant intrinsic curvature effect may only moderately correct the global test, implying that the intrinsic curvature effect need not be useful even when it exists. This was not considered in earlier work on intrinsic curvature corrections in nonlinear models which did not consider the local frame of reference provided by model-data combination and null hypothesis. Intrinsic curvature on its own does not tell the whole story and computing these across many examples is not particularly useful. Each model-data and null hypothesis combination must be analyzed on its own merits. Note that the approach given here can also be used as a diagnostic defining regions of intrinsic curvature on a regression surface in relation to the observed response.



## References

- Bates, D.M. and Watts, D.G. (1980). *Relative curvature measures of nonlinearity (with discussion)*. J. Roy. Statist. Soc. Ser. B, 42, 1-25.
- Bates, D.M. and Watts, D.G. (1988). *Nonlinear Regression Analysis and Its Applications*. John Wiley, New York. <http://dx.doi.org/10.1002/9780470316757>
- Beale E.M.L. (1960). *Confidence regions in nonlinear estimation*. J. Roy. Statist. Soc. Ser. B, 22, 41-76.
- Box, G.E.P. & Cox, D.R. (1964) *An analysis of transformations*, J. Roy. Statist. Soc. Ser. B, 26, 211-252.
- Box, G.E.P. & Cox, D.R. (1982) *An analysis of transformation revisited, rebutted*, Journal of the American Statistical Association, 77, 209-210. <http://dx.doi.org/10.1080/01621459.1982.10477788>
- Cook R.D. and Witmer J.A. (1985). *A Note on Parameter-Effects Curvature*. J. Am. Statist. Assoc., 80, 872 - 878. <http://dx.doi.org/10.1080/01621459.1985.10478196>
- Cox, D.R. and Hinkley, D.V. (1974). *Theoretical Statistics*. Chapman and Hall, London. <http://dx.doi.org/10.1007/978-1-4899-2887-0>
- Donaldson, J.R. and Schnabel, R.B. (1987). *Computational experience with confidence regions and confidence intervals for nonlinear least squares*. Technometrics, 29, 67-82. <http://dx.doi.org/10.1080/00401706.1987.10488184>
- Hoerl A.E. and Kennard R.W. (1970). *Ridge Regression: Biased Estimation for Nonorthogonal Problems*. Technometrics 12, 55-67. <http://dx.doi.org/10.1080/00401706.1970.10488634>
- Kreyszig, Erwin (1991). *Differential Geometry*. ISBN 0-486-66721-9.
- Kutner M.H., Nachtsheim C.J., Neter J., Li W. (2005). *Applied Linear Statistical Models*, 5th ed, McGraw-Hill, Irwin, New York.
- Luellen J.K., Shadish W.R. and Clark M.H. (2005). *Propensity Scores: An Introduction and Experimental Test*. Eval Rev 2005 29: 530-558. <http://dx.doi.org/10.1177/0193841X05275596>
- McCullagh P. and Nelder J.A. (1989). *Generalized Linear Models*, 2nd ed. Chapman and Hall, New York. <http://dx.doi.org/10.1007/978-1-4899-3242-6>
- Seber, G.A.F., and Wild, C.J. (1989). *Nonlinear Regression*. John Wiley, New York. <http://dx.doi.org/10.1002/0471725315>

## Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).