Applied and
Computational
Mathematics
Division

NISTIR 89–4102

Center for Computing and Applied Mathematics

The Computation and Use of the Asymptotic Covariance Matrix for Measurement Error Models

Paul T. Boggs and Janet R. Donaldson

June 1989

U.S. DEPARTMENT OF COMMERCE
National Institute of Standards and Technology
Gaithersburg, MD 20899

The Computation and Use of the Asymptotic Covariance Matrix for Measurement Error Models *

Paul T. Boggs Janet R. Donaldson

June 26, 1989

Abstract

Measurement error models assume that errors occur in both the response and predictor variables. In using these models, it is of interest to compute confidence regions and intervals for the model parameters. Fuller [1987] provides an asymptotic form for the covariance matrix that can be used to construct approximate confidence regions and intervals. We discuss the solution of the minimization problem resulting from the use of a measurement error model, and we develop a procedure for accurately computing the covariance matrix. We then assess via a Monte Carlo study the quality of the confidence regions and intervals constructed from this matrix.

Keywords: confidence intervals, confidence regions, covariance matrix, errors in variables, measurement error models, Monte Carlo study, ordinary least squares, orthogonal distance regression, simultaneous equations models.

1. Introduction

Parameter estimation and data fitting are among the most common activities in science, with the ordinary least squares criterion being by far the most frequently used. The emergence over the last decade of high quality software for finding the

*Contribution of the National Institute of Standards and Technology and not subject to copyright in the United States.

ordinary least squares solution for nonlinear functions has allowed researchers to consider more realistic data fitting and parameter estimation models for many situations. Until recently, however, researchers have not had procedures available that would take into account the fact that errors in the observation of the predictor variables, x_i, are frequently significant with respect to the errors in the observations of the response variables, y_i. Such errors can make a substantial difference in resulting estimations. (See, e.g., [Boggs *et al.*, 1988].)

To be specific, ordinary least squares problems arise when the actual (or true) value of the response variable, denoted by a superscript a, i.e., y_i^a, is observed with some actual but unknowable error ϵ_i^a, while the predictor variables are observed without error, i.e., $x_i = x_i^a$. Now if we assume that

$$\begin{aligned} y_i &= y_i^a - \epsilon_i^a \\ &= f(x_i^a; \beta^a) - \epsilon_i^a \qquad i = 1, \ldots, n, \end{aligned}$$

where β^a denotes the actual value of the vector of model parameters, then the ordinary least squares criterion, which minimizes the sum of the squares of the estimates of the errors in y_i, can be applied to obtain an estimate of β^a.

If there are also significant actual, but unknowable, errors δ_i^a in the predictor variables, so that $x_i = x_i^a - \delta_i^a$, then a generalization of the ordinary least squares criterion is required, since, in this case the model becomes

$$y_i = f(x_i + \delta_i^a; \beta^a) - \epsilon_i^a \qquad i = 1, \ldots, n.$$

This problem goes under various names, including errors in variables, generalized least squares, orthogonal distance regression, and measurement error models. We prefer measurement error models in deference to the book of Fuller [1987] that presents the definitive modern treatment of the problem. We also use the term orthogonal distance regression since, as we show in §2, it is a useful geometric description of the problem actually solved.

As in the ordinary least squares case, when using measurement error models one is frequently interested in constructing confidence regions and/or confidence intervals for the model parameters. To this end, Fuller derives the asymptotic form of the covariance matrix and uses it in several examples. It is well known, however, that for nonlinear models in general, and for measurement error models in particular, confidence regions and intervals constructed using the covariance matrix are only approximate.

In this paper we discuss a stable and efficient numerical computation of the covariance matrix of the estimators of the parameters of measurement error models, and the use of this matrix to construct confidence regions and/or confidence

intervals. The quality of confidence regions and intervals for ordinary nonlinear least squares estimators was discussed by Donaldson and Schnabel [1987]. In a similar manner, we attempt to assess the quality of such regions and intervals for measurement error models via a Monte Carlo study.

In §2 we give the details of the measurement error model and the formulation of the minimization problem to be solved. We briefly review the solution of this problem using the numerically stable and efficient algorithm provided by Boggs *et al.* [1987], and describe the features of ODRPACK [Boggs *et al.*, 1989], a high quality, public domain implementation of this algorithm. ODRPACK has been successfully used to solve measurement error problems at a number of sites, and is used in the Monte Carlo study described in the last section.

In §3 we review the basis for using the covariance matrix for estimating confidence regions and intervals. We also show how the asymptotic formula of Fuller can be derived from the linearization of a related ordinary least squares problem. We then point out the potential problems associated with using this linearization, and elaborate on these problems for the special case at hand. In §4 we show how the covariance matrix can be efficiently computed in a numerically stable manner, as has been done in ODRPACK. Finally, in §5, we present the description and results of our Monte Carlo study of the accuracy of the confidence regions and intervals obtained using the covariance matrix. These results suggest a strategy for selecting appropriate weights to improve the accuracy of the resulting confidence intervals for β^a and δ_i^a.

The results of our Monte Carlo study are in accordance with those of Donaldson and Schnabel [1987] who show that even for ordinary nonlinear least squares, confidence regions and intervals constructed using the covariance matrix are less accurate than other approximate, but more computationally expensive methods. Despite its potential inaccuracy, the covariance matrix is frequently used to construct confidence regions and intervals for measurement error models because the resulting regions and intervals are inexpensive to compute, they are often adequate, and they are familiar to practitioners. Caution must be exercised when using such regions and intervals, however, since the validity of the approximation will depend on the nonlinearity of the model, the variance and distribution of the errors, and the data itself. When more reliable intervals and regions are required, other more accurate methods should be used. (See, e.g., Donaldson and Schnabel [1987], and Efron [1985].)

2. Measurement Error Models and Orthogonal Distance Regression

In this section, we derive the measurement error problem as a generalization of the ordinary nonlinear least squares problem. We then briefly discuss its efficient solution and the ODRPACK software package that incorporates this procedure.

The data fitting problem that we consider is composed of an observed data set (x_i, y_i), $i = 1, \ldots, n$, and a model that is purported to explain the relationship of the response variable $y_i \in \Re^1$ to the predictor variables $x_i \in \Re^m$. We assume that the response variable is given as a function of the predictor variables and a set of parameters $\beta \in \Re^p$, i.e.,

$$y_i^a = f(x_i^a; \beta^a) \qquad i = 1, \ldots, n,$$

where f is a smooth function that can be either linear or nonlinear in x_i and β, and the superscript a denotes the actual (or true) value of the corresponding quantity.

Now if we suppose that the observations y_i contain actual, but unknown, additive errors $\epsilon_i^a \in \Re^1$, while the predictor variables are still known exactly, i.e., $x_i = x_i^a$, then y_i satisfies

$$y_i = f(x_i; \beta^a) - \epsilon_i^a \qquad i = 1, \ldots, n.$$

One can therefore estimate β^a by finding that β which minimizes the sum of squares

$$S(\beta) \equiv F(\beta)^{\mathrm{T}} W F(\beta),$$

where $F(\beta)$ is the vector valued function with ith element equal to $f(x_i; \beta) - y_i$, $F(\beta)^{\mathrm{T}}$ denotes the transpose of $F(\beta)$, and $W \in \Re^{n \times n}$ is the diagonal matrix with (i,i)th component equal to w_i^2. (The weights w_i can be used to modify the contribution to $S(\beta)$ of the various observations (x_i, y_i), possibly because the n observations have unequal precision.) Thus the least squares approximation $\hat{\beta}$ to β^a solves the *ordinary least squares* minimization problem

$$\min_{\beta} S(\beta) = \min_{\beta} \sum_{i=1}^{n} w_i^2 \left[(f(x_i; \beta) - y_i)^2 \right]. \tag{2.1}$$

That is, $\hat{\beta}$ minimizes the sum of the squared distances in the y direction from the observation (x_i, y_i) to the point on the curve $f(x; \beta)$.

The explicit *measurement error model* results when we allow additive errors in both x_i and y_i. If we assume that $y_i = y_i^a - \epsilon_i^a$ and $x_i = x_i^a - \delta_i^a$, where $\delta_i^a \in \Re^m$ is the actual, but unknown, additive error in the observation x_i, then the observations satisfy

$$y_i = f(x_i + \delta_i^a; \beta^a) - \epsilon_i^a \qquad i = 1, \ldots, n. \tag{2.2}$$

The term *explicit* refers to the fact that y can be written directly as a function of x and β. The more general *implicit* problem, which has the form

$$\bar{f}(x_i + \delta_i^a, y_i + \epsilon_i^a; \beta^a) = 0,$$

is considered in [Fuller, 1987]. The implicit problem is computationally more difficult, and is not discussed further here. (See, e.g., Boggs *et al.*, 1987.)

When there are errors in both x_i and y_i, then it is reasonable to define the distance from the observation (x_i, y_i) to the curve $f(x; \beta)$ as the radius of the smallest circle centered at (x_i, y_i) that is tangent to the curve. If the point of tangency is $(x_i + \delta_i, y_i + \epsilon_i)$, then, by the Pythagorean theorem, this *orthogonal distance* is

$$r_i^2 = (f(x_i + \delta_i; \beta) - y_i)^2 + \delta_i^{\mathrm{T}} \delta_i. \tag{2.3}$$

The observations x_i and y_i can have unequal precision, however. We compensate for this by generalizing (2.3) to the *weighted orthogonal distance*, defining

$$\tilde{r}_i^2 = (f(x_i + \delta_i; \beta) - y_i)^2 + \delta_i^{\mathrm{T}} d_i^2 \delta_i,$$

where $d_i \in \Re^{m \times m}$, $i = 1, \ldots, n$, is a set of positive diagonal matrices that weight each individual component of δ_i.

We can then approximate β^a by finding that β which minimizes the sum of the squares of the $\tilde{r}_i$. That is, we solve

$$\min_{\beta, \delta} \sum_{i=1}^{n} w_i^2 \left[(f(x_i + \delta_i; \beta) - y_i)^2 + \delta_i^{\mathrm{T}} d_i^2 \delta_i \right], \tag{2.4}$$

where w_i, $i = 1, \ldots, n$, is again a set of nonnegative numbers that allows us to vary the contribution of the various observations to the sum of squares.

The orthogonal distance regression problem defined by (2.4) can also be expressed as an ordinary nonlinear least squares problem with $n + nm$ observations and $p + nm$ unknowns. We designate the unknowns of this ordinary least squares problem as $\eta^{\mathrm{T}} = (\beta^{\mathrm{T}}, \delta_1^{\mathrm{T}}, \ldots, \delta_n^{\mathrm{T}})$. The sum of squares to be minimized is therefore

$$S(\eta) \equiv G(\eta)^{\mathrm{T}} \Omega G(\eta) \tag{2.5}$$

where $G(\eta)$ is the vector valued function whose ith element is defined by

$$g_i(\eta) = \begin{cases} f(x_i + \delta_i; \beta) - y_i & i = 1, \ldots, n, \\ \eta_{p+i-n} & i = n+1, \ldots, n+nm, \end{cases} \tag{2.6}$$

and $\Omega \in \Re^{(n+nm)\times(n+nm)}$ is the diagonal weighting matrix given by

$$\Omega = \begin{bmatrix} W & \\ & D \end{bmatrix} \tag{2.7}$$

with $W \in \Re^{n\times n}$ the diagonal matrix with ith component w_i^2, and $D \in \Re^{nm\times nm}$ the diagonal matrix composed of the individual diagonal matrices $w_i^2 d_i^2$.

Boggs *et al.* [1987] have exploited the special structure of the first derivative of $G(\eta)$ with respect to η to create a trust-region, Levenberg-Marquardt algorithm for solving the orthogonal distance regression problem defined by (2.4). Their algorithm is both stable and efficient, requiring only $O(np^2)$ operations per iteration. A similar ordinary least squares algorithm applied to (2.5) would require $O(n(nm+p)^2)$ operations per iteration. Thus the time per iteration in the [Boggs *et al.*, 1987] algorithm grows linearly in n while it grows as n^3 in an ordinary least squares code applied to (2.5). The portable Fortran subroutine library ODRPACK [Boggs *et al.*, 1989] is an implementation of this algorithm. It is the first widely available, efficient package for solving the orthogonal distance regression problem. ODRPACK can therefore be used to solve much larger orthogonal distance regression problems than could be solved using ordinary nonlinear least squares software, even though both solutions are mathematically equivalent.

ODRPACK is designed to solve both ordinary least squares and orthogonal distance regression problems, handling many levels of user sophistication and problem difficulty.

- It is easy to use, providing two levels of user control of the computations, extensive error handling facilities, optional printed reports, and no size restrictions other than effective machine size.

- The necessary derivatives (Jacobian matrices) are approximated numerically if they are not supplied by the user. The correctness of user supplied derivatives can also be verified by the derivative checking procedure provided.

- Both weighted and unweighted analysis can be performed.

- Subsets of the unknowns can be treated as constants with their values held fixed at their input values, allowing the user to examine the results obtained by estimating subsets of the unknowns of a general model without rewriting the model subroutine.
- The covariance matrix and the standard errors for the model parameter estimators are optionally provided.
- The ODRPACK scaling algorithm automatically compensates for poorly scaled problems, in which the model parameters and/or unknown errors in the independent variables vary widely in magnitude.

ODRPACK is available free of charge from the authors.

3. Linearized Confidence Regions and Intervals

Confidence regions and confidence intervals are commonly computed in statistical applications to assess a bound on the expected difference between the estimated value and the actual (or true) value. A joint $(1-\alpha)$, $0 < \alpha < 1$, confidence region for all of the unknowns of the model can be informally defined as a region for the unknowns in which one expects the true value to lie with probability $(1-\alpha)$. That is, if the experiment used to generate the data were repeated a large number of times under the same conditions, and the unknowns estimated for each replication, then one would expect the true value of the unknowns to lie within the constructed confidence regions approximately $100(1-\alpha)\%$ of the time. A $(1-\alpha)$ confidence interval for an individual unknown can be defined analogously. Methods of constructing confidence regions and intervals that are statistically guaranteed to contain the true value $100(1-\alpha)\%$ of the time are called *exact*; all other methods are called *approximate*.

To understand the linearized confidence regions and intervals, we first review the ordinary linear least squares problem, where $f(x_i;\beta)$ is linear in the parameters β and the x_i are observed *without* error, i.e., $f(x_i;\beta) = (x_i^a)^{\mathrm{T}}\beta$, $i = 1,\ldots,n$. For such an ordinary linear model, if we assume that the errors $\epsilon \sim \mathrm{N}(0,(\sigma^a)^2W^{-1})$ with the actual residual standard deviation σ^a unknown, then an exact $100(1-\alpha)\%$ confidence region for β^a can be specified as the region that contains those values β for which

$$\frac{S(\beta)-S(\hat{\beta})}{\hat{\sigma}^2} \leq p\mathrm{F}_{p,n-p,1-\alpha}, \tag{3.8}$$

where $\hat{\sigma}^2$ is the estimated residual variance,

$$\hat{\sigma}^2 = \sum_{i=1}^{n} w_i^2 (f(x_i; \hat{\beta}) - y_i)^2 / (n - p),$$

and $F_{p,n-p,1-\alpha}$ is the $100(1-\alpha)\%$ percentage point for the F distribution with p and $n-p$ degrees of freedom. Similarly, an exact $100(1-\alpha)\%$ confidence interval for β_j^a can be specified as the interval between the values of β_j that maximize $(\beta_j - \hat{\beta}_j)^2$ subject to

$$\frac{S(\beta) - S(\hat{\beta})}{\hat{\sigma}^2} = t^2_{n-p,1-\alpha/2} = F_{1,n-p,1-\alpha}, \tag{3.9}$$

where $t^2_{n-p,1-\alpha/2}$ is the square of the $100(1-\alpha/2)\%$ percentage point of the t distribution with $n-p$ degrees of freedom.

In general, the contours of constant likelihood represented by $S(\beta) - S(\hat{\beta})$ are expensive to compute, requiring the evaluation of $f(x_i; \beta)$ at a large number of points. For ordinary linear least squares, however, the region specified by (3.8) is mathematically equivalent to the region specified by

$$(\beta - \hat{\beta})^{\mathrm{T}} V^{-1} (\beta - \hat{\beta}) \leq p F_{p,n-p,1-\alpha}, \tag{3.10}$$

while the interval specified by (3.9) is equivalent to that specified by

$$\left| \beta_j - \hat{\beta}_j \right| \leq V_{jj}^{1/2} t_{n-p,1-\alpha/2} \tag{3.11}$$

where $V \in \Re^{p \times p}$ is the estimated covariance matrix for the parameter estimators $\hat{\beta}$, and $V_{jj}^{1/2}$ is the square root of its (j,j)th element. The covariance matrix for a linear model is easily computed using

$$V = \hat{\sigma}^2 [X^{\mathrm{T}} W X]^{-1},$$

where $X \in \Re^{n \times m}$ is the matrix with ith row x_i^{T}, i.e., the Jacobian matrix with (i,j)th component equal to $\partial f(x_i; \beta)/\partial \beta_j$. (We assume that X is of full rank, so that $X^{\mathrm{T}} W X$ is nonsingular.) Consequently, for ordinary linear least squares models, exact confidence regions and intervals can be easily constructed using (3.10), (3.11) and the $p \times p$ covariance matrix.

When $f(x; \beta)$ is nonlinear, construction of exact confidence regions and confidence intervals is much more difficult, and so approximate methods are frequently

used. It is possible to use the likelihood contours specified by (3.8) and (3.9) directly to construct confidence regions and intervals that are fairly accurate, although they are not exact except under special conditions (see, e.g., Donaldson and Schnabel [1987] or Draper and Smith [1981]). These *likelihood* regions and intervals have several computational disadvantages, however. First, both the confidence regions and confidence intervals produced using likelihood contours can be disjoint and unbounded. Second, as noted above, this method is computationally expensive, since it requires that $f(x_i;\beta)$ be evaluated at a sufficient number of points to allow the likelihood contours to be constructed. Finally, when the data arrays are large, it can be awkward to publish the information necessary to reconstruct the confidence region because this information cannot necessarily be succinctly summarized.

Because of the computational disadvantages of the likelihood method, the less accurate, but more easily computed, linearization method is the most widely used of the approximate methods for ordinary nonlinear functions. The *linearization method* assumes that the nonlinear function can be adequately approximated at the solution by a linear model. Under this assumption, and the assumption that $\epsilon \sim \mathrm{N}(0,(\sigma^a)^2W^{-1})$, we can construct the covariance matrix

$$\hat{V}_{ols} = \hat{\sigma}^2[F'(\hat{\beta})^{\mathrm{T}}WF'(\hat{\beta})]^{-1},$$

where $\hat{\sigma}^2 = F(\hat{\beta})^{\mathrm{T}}WF(\hat{\beta})/(n-p)$, and $F'(\hat{\beta}) \in \Re^{n\times p}$ is the Jacobian matrix with (i,j)th element equal to $\partial f(x_i;\beta)/\partial\beta_j$ evaluated at $\hat{\beta}$. (See, e.g., Bard [1974], p. 176-178.) We again assume that $F'(\hat{\beta})$ is of full rank, so that $F'(\hat{\beta})^{\mathrm{T}}WF'(\hat{\beta})$ is nonsingular.

Using $\hat{V}_{ols}$, (3.10) and (3.11), we can construct approximate "linearized" confidence regions and intervals for the parameters of an ordinary nonlinear model. The adequacy of these approximations, however, will depend on how well the linearized model approximates the actual function over the region defined by the linearized confidence region and confidence intervals. This, in turn, depends on the nonlinearity of $f(x;\beta)$ [Bates and Watts, 1980], and the size of σ^a. Donaldson and Schnabel [1987] have shown that linearized confidence intervals appear to be reasonably good in practice, while linearized confidence regions can be very inadequate.

The linearized confidence regions and intervals for the βs and the δs estimated by orthogonal distance regression are the same as the regions and intervals that are obtained when the orthogonal distance regression problem is solved as a $p+nm$ parameter ordinary nonlinear least squares problem (see §2). As above, we assume

that that the nonlinear function can be adequately approximated at the solution by a linear model. If $(\epsilon^{\mathrm{T}}, \delta_1^{\mathrm{T}}, \ldots, \delta_n^{\mathrm{T}})^{\mathrm{T}} \sim \mathrm{N}(0, (\sigma^a)^2\Omega^{-1})$, then the covariance matrix for the estimators $\hat{\eta}$ is

$$\hat{V}_{odr} = \hat{\sigma}^2[G'(\hat{\eta})^{\mathrm{T}}\Omega G'(\hat{\eta})]^{-1}, \tag{3.12}$$

where $\hat{\sigma}^2 = G(\hat{\eta})^{\mathrm{T}}\Omega G(\hat{\eta})/(n-p)$, and $G'(\hat{\eta}) \in \Re^{(n+nm)\times(p+nm)}$ is the Jacobian matrix with (i,j)th element equal to $\partial g_i(\eta)/\partial\eta_j$ evaluated at $\hat{\eta}$. Using $\hat{V}_{odr}$, we can construct linearized confidence regions for η^a using (3.10) and linearized confidence intervals for η_j^a using (3.11). In the next section, we show how the covariance matrix defined by (3.12) can be computed in a numerically stable way.

For ordinary nonlinear least squares, the linearization method is asymptotically correct as $n \to \infty$. (See, e.g., Jennrich [1969].) For the orthogonal distance regression problem, this method has been shown to be asymptotically correct as $\sigma^a \to 0$ [Fuller, 1987]. The difference between the conditions of asymptotic correctness can be explained by the fact that, as the number of observations increases in the orthogonal distance regression problem, we do not obtain additional information for δ_i. Thus, for orthogonal distance regression problems, we would expect the portion of the covariance matrix concerned with β to yield linearized regions and intervals as accurate as those computed for ordinary nonlinear least squares problems, while we would expect the regions and intervals for δ_i^a to be less accurate.

Note also that $\hat{V}_{odr}$ is dependent upon the weight matrix Ω, which must be assumed to be correct, and cannot be confirmed from the orthogonal distance regression results. Errors in the w_i and d_i that form Ω will have an adverse affect on the accuracy of $\hat{V}_{odr}$ and its component parts. In §5, we present the results of a Monte Carlo experiment examining the accuracy of the linearized confidence intervals. The results indicate that the confidence regions and intervals for δ_i^a are not as accurate as those for β^a. These results also show that errors in Ω can have an adverse affect on both confidence regions and intervals.

4. Computing the Covariance Matrix $\hat{V}_{odr}$

The most straightforward computation of a quantity is often not the most numerically stable. Although $\hat{V}_{odr}$ is defined as

$$\hat{V}_{odr} = \hat{\sigma}^2[G'(\hat{\eta})^{\mathrm{T}}\Omega G'(\hat{\eta})]^{-1},$$

we would not compute it by first calculating $G'(\hat{\eta})^{\mathrm{T}}\Omega G'(\hat{\eta})$ and then inverting the resulting $(p + nm) \times (p + nm)$ matrix because such a procedure would introduce unnecessary numerical errors that could severely jeopardize the accuracy of $\hat{V}_{odr}$. In this section, we present a numerically stable and efficient method for constructing $\hat{V}_{odr}$.

For any $A \in \Re^{n\times p}$ with linearly independent columns, it is generally recommended that the matrix $[A^{\mathrm{T}}A]^{-1}$ be computed by first constructing $A = QR$, where $Q \in \Re^{n\times n}$ has orthonormal columns, and $R \in \Re^{n\times p}$ is upper triangular with positive diagonal elements. If we let $\tilde{R} \in \Re^{p\times p}$ be the upper $p \times p$ portion of R, then $A^{\mathrm{T}}A = R^{\mathrm{T}}R = \tilde{R}^{\mathrm{T}}\tilde{R}$ and $[A^{\mathrm{T}}A]^{-1} = \tilde{R}^{-1}(\tilde{R}^{-1})^{\mathrm{T}}$. Since $\tilde{R}$ is triangular, its inverse can be accurately computed, thus allowing an accurate computation of $[A^{\mathrm{T}}A]^{-1}$. (See, e.g., [Dongarra *et al.*, 1979].)

The computation of $\hat{V}_{odr}$ can be further improved, however, since analysis of $G'(\eta)$ shows that it has the special structure

$$\begin{bmatrix} J & U \\ 0 & I \end{bmatrix}$$

where $J = F'(\hat{\beta})$, and $U \in \Re^{n\times nm}$ is the "staircase" matrix

$$\begin{bmatrix} u_{1,1} \cdots u_{1,m} & & & & \\ & u_{2,m+1} \cdots u_{2,2m} & & & \\ & & \vdots & & \\ & & & u_{n-1,1+m(n-2)} \cdots u_{n-1,m(n-1)} & \\ & & & & u_{n,1+m(n-1)} \cdots u_{n,nm} \end{bmatrix}$$

with

$$\begin{aligned} u_{i,j} &= \partial g_i(\eta)/\partial \eta_{p+j} \\ &= \begin{cases} \partial f(x_i + \delta_i; \beta)/\partial \delta_{i,j-(i-1)m} & \text{if } 1 + (i-1)m \le j \le im \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

for $i = 1, \ldots, n$ and $j = 1, \ldots, nm$. Thus,

$$\begin{aligned} \hat{V}_{odr} &= \hat{\sigma}^2[G'(\hat{\eta})^{\mathrm{T}}\Omega G'(\hat{\eta})]^{-1} \\ &= \hat{\sigma}^2 \left[\begin{bmatrix} J^{\mathrm{T}} & 0 \\ U^{\mathrm{T}} & I \end{bmatrix} \begin{bmatrix} W & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} J & U \\ 0 & I \end{bmatrix} \right]^{-1} \\ &= \hat{\sigma}^2 \begin{bmatrix} J^{\mathrm{T}}WJ & J^{\mathrm{T}}WU \\ U^{\mathrm{T}}WJ & U^{\mathrm{T}}WU + D \end{bmatrix}^{-1}. \end{aligned}$$

We can partition $\hat{V}_{odr}$ as

$$\hat{V}_{odr} = \begin{bmatrix} \hat{V}_\beta & \hat{V}_{\beta\delta} \\ \hat{V}_{\delta\beta} & \hat{V}_\delta \end{bmatrix} \tag{4.13}$$

where $\hat{V}_\beta \in \Re^{p\times p}$ is the covariance matrix of the estimated $\hat{\beta}$s, $\hat{V}_\delta \in \Re^{nm\times nm}$ is the covariance matrix of the estimated $\hat{\delta}$s, and $\hat{V}_{\beta\delta} = \hat{V}_{\delta\beta}^{\mathrm{T}} \in \Re^{p\times nm}$ gives covariances between the βs and the δs. The component parts of $\hat{V}_{odr}$ are thus

$$\begin{aligned}
\hat{V}_\beta &= \hat{\sigma}^2 \left[J^{\mathrm{T}} \left(W - WU \left[U^{\mathrm{T}} W U + D \right]^{-1} U^{\mathrm{T}} W \right) J \right]^{-1} \\
\hat{V}_{\beta\delta} &= -\hat{\sigma}^2 \hat{V}_\beta (J^{\mathrm{T}} W U) \left[U^{\mathrm{T}} W U + D \right]^{-1} \\
\hat{V}_{\delta\beta} &= \hat{V}_{\beta\delta}^{\mathrm{T}} \\
\hat{V}_\delta &= -\hat{\sigma}^2 \left[U^{\mathrm{T}} W U + D \right]^{-1} \left(I + (U^{\mathrm{T}} W J) \hat{V}_\beta (J^{\mathrm{T}} W U) \left[U^{\mathrm{T}} W U + D \right]^{-1} \right).
\end{aligned}$$

Boggs *et al.* [1987] show how the structural properties of the matrices appearing in (4.13) can be exploited to compute $\hat{V}_{odr}$ accurately and efficiently. In particular, they define

$$P^{-1} \equiv [U^{\mathrm{T}} W U + D]^{-1}$$

and

$$\omega_i \equiv \sum_{j=1}^{nm} \frac{u_{i,j}^2}{D_{j,j}} \qquad i = 1, \ldots, n.$$

They then show that

$$P^{-1} = D^{-1} - D^{-1} U^{\mathrm{T}} W^{1/2} M W^{1/2} U D^{-1} \tag{4.14}$$

where $M \in \Re^{n\times n}$ is the diagonal matrix defined by

$$M \equiv \mathrm{diag}\left\{ \left[\frac{1}{1+\omega_i} \right],\ i = 1, \ldots, n \right\}.$$

Because D and W are diagonal, P^{-1} can be easily computed.

Boggs *et al.* [1987] also show that

$$W - WU[U^{\mathrm{T}} W U + D^{\mathrm{T}} D]^{-1} U^{\mathrm{T}} W = M.$$

Thus,

$$\begin{aligned} \hat{V}_\beta &= \hat{\sigma}^2 \left[J^T \left(W - WU \left[U^T W U + D \right]^{-1} U^T W \right) J \right]^{-1} \\ &= \left[M^{1/2} J^T J M^{1/2} \right]^{-1} \end{aligned}$$

can be stably computed using the factorization techniques described at the beginning of this section. This result and (4.14) allow the efficient formation of $\hat{V}_{\beta\delta}$, $\hat{V}_{\delta\beta}$ and $\hat{V}_\delta$.

The covariance matrix $\hat{V}_\beta$ of the estimators $\hat{\beta}$ provided by ODRPACK is computed using the above technique; users are seldom interested in $\hat{V}_\delta$ or $\hat{V}_{\beta\delta}$. If necessary, however, the full covariance matrix $\hat{V}_{odr}$ for all of the estimators $\hat{\eta}$ either can be computed using the above equations, or can be "automatically" obtained from most ordinary least squares software (including ODRPACK) by solving the orthogonal distance regression problem as the ordinary least squares problem defined by (2.6).

5. Computational Experience

Given the ability to define and construct linearized confidence regions and intervals for the parameters of an orthogonal distance regression problem, it is reasonable to ask how good these regions and intervals are. In this section, we present the results of a Monte Carlo study that indicates that, at least in some cases, linearized confidence regions and intervals are quite good.

As we noted earlier, if a given experiment were repeated a large number of times under the same conditions, and the unknowns and their associated confidence intervals and regions estimated for each replication, then we would expect the true value of the unknowns to lie within the constructed confidence regions and intervals approximately $100(1 - \alpha)\%$ of the time. Most experiments, however, cannot be repeated a large number of times. In addition, even if the experiment were repeated a large number of times, the true values of the parameters are not generally known. Therefore, in practice it is almost impossible to assess the reliability of confidence intervals or regions except in the context of a Monte Carlo experiment, where the data are created by computer simulation.

A Monte Carlo experiment allows us to examine the properties of confidence regions and intervals for a given problem. For such an experiment, we define the *observed coverage*, $\hat{\gamma}_\alpha$, of a constructed confidence interval or region as the

percentage of the time that true value did lie within the interval or region for the parameter. The *nominal coverage* of such a region or interval is $100(1-\alpha)\%$. When the number of realizations of the data is large, then the observed coverage will reflect the actual (or true) coverage, γ_α^a, of the given region or interval. The actual coverage may or may not be the same as the nominal coverage, however. By comparing γ_α with $100(1-\alpha)\%$, we can thus assess the quality of an approximate confidence interval or region.

Donaldson and Schnabel [1987] examined linearized confidence intervals and regions for a number of ordinary nonlinear least squares models and data sets. They found that the linearization method is not always adequate. Their results showed that, although there were many examples in which the linearization method's coverage differed from nominal by only a very small amount, there were also many cases in which the observed coverage was far lower than nominal.

We would not expect an exhaustive study of orthogonal distance regression problems to produce results that were different than those found by Donaldson and Schnabel [1987]. Thus we do not attempt such a large scale study. Here we present the results of a Monte Carlo study of only four data sets. They demonstrate that the linearized confidence intervals and regions can be quite reasonable when Ω is known, but that when Ω must be approximated, the results may not be adequate.

Our first example is from Fuller [1987, example 3.2.2, p. 230-238]. The data (x_i, y_i) are the percent saturation of nitrogen gas in a brine solution forced into the pores of sandstone, and the observed compressional wave velocity of ultrasonic signals propagated through the sandstone, respectively. These data are assumed to be modeled by

$$f(x_i + \delta_i; \beta) = \beta_1 + \beta_2(e^{\beta_3(x_i+\delta_i)} - 1)^2.$$

Fuller analyzed the original data assuming a measurement error model with $D^a = I$. For our Monte Carlo experiment, we assigned

$$\begin{aligned}
\beta^a &= (1264.65, -54.02, -0.0879)^{\mathrm{T}}, \\
X^a &= (0.0, 0.0, 5.0, 7.0, 7.5, 10.0, 16.0, 26.0, 30.0, 34.0, 34.5, 100.0)^{\mathrm{T}}, \\
w_i^a &= 1, \; i = 1, \ldots, 12 \\
d_i^a &= 1, \; i = 1, \ldots, 12 \\
\sigma^a &= \sqrt{2.38}\,.
\end{aligned}$$

The values selected for β^a and σ^a are those estimated by Fuller using the original data with $\Omega = I$. The values x_i^a, $i = 1, \ldots, n$, are the observations from the

original experiment. Fuller notes that it is reasonable to believe that the error variance for $x = 0$ and $x = 100$ is smaller than the error variances for the remaining observations. For our Monte Carlo experiment, we assume that x_1, x_2 and x_{12} are observed without error, and thus fix $\delta_1^a = \delta_2^a = \delta_{12}^a = 0$.

Our second example is from Ratkowsky [1983, example 6.11, p. 119-120]. The response variable purports to represent resistance of a thermistor and the predictor variable temperature. Ratkowsky, however, notes that since the resistance of a thermistor increases with temperature, the response variable probably represents conductance. The model used to describe the original data is

$$f(x_i;\beta) = -\beta_1 + \frac{\beta_2}{x_i + \beta_3}.$$

The analysis by Ratkowsky assumed that there was no error in the response variable; for our results, we assume a measurement error model with $d_i^a = 1/10$, $i = 1,\ldots,n$. For this example, we assigned

$$\begin{aligned}
\beta^a &= (5.0, 6150.0, 350.0)^{\mathrm{T}}, \\
X^a &= (45 + 5i, i = 1,\ldots,16)^{\mathrm{T}}, \\
w_i^a &= 1,\ i = 1,\ldots,16 \\
d_i^a &= 1/10,\ i = 1,\ldots,16 \\
\sigma^a &= 0.0002\ .
\end{aligned}$$

The values x_i^a, $i = 1,\ldots,n$, are the observations from the original experiment. The other values are approximately those obtained by Ratkowsky in his analysis.

The third example is problem E, chapter 10 of Draper and Smith [1981, p. 518-519]. This example models the relationship between pressure and temperature in saturated steam using

$$f(x_i;\beta) = \beta_1 \cdot 10^{\beta_2 x_i/(\beta_3 + x_i)}.$$

Draper and Smith assumed that there was no error in the temperature observations; for our results, we assume a measurement error model with $d_i^a = 10$, $i = 1,\ldots,n$. We assigned

$$\begin{aligned}
\beta^a &= (4.18, 6.91, 205.0)^{\mathrm{T}}, \\
X^a &= (0, 10, 20, 30, 40, 50, 60, 70, 80, 85, 90, 95, 100, 105)^{\mathrm{T}}, \\
w_i^a &= 1,\ i = 1,\ldots,14 \\
d_i^a &= 10,\ i = 1,\ldots,14 \\
\sigma^a &= 1.2\ .
\end{aligned}$$

The values x_i^a, $i = 1, \ldots, n$, are the observations from the original experiment. The other values are approximately those obtained using the measurement error model for the original data with $d_i^a = 10$.

The data for the fourth example were collected as part of a phychophysical experiment to evaluate the ability of human subjects to perceive a visual signal as a function of the intensity of the signal. The predictor variable, x_i, represents the signal intensity and the response variable, y_i, is the fraction of the total number of trials during which a particular subject correctly identified the presence of the signal. Each signal level was repeated 80 times. A sigmoidal curve belonging to the family

$$f(x_i;\beta) = \frac{\beta_1}{[1 + e^{\beta_2 - \beta_3(x_i + \delta_i)}]^{\beta_4}}$$

is used to relate y to x. We assigned

$$\begin{aligned}
\beta^a &= (0.936, 3.400, 339.370, 0.954)^{\mathrm{T}}, \\
X^a &= (0.003, 0.007, 0.008, 0.010, 0.015, 0.026, 0.038, 0.060, 0.065)^{\mathrm{T}}, \\
w_i^a &= (y_i^a \cdot (1 - y_i^a)/80)^{-1/2}, \quad i = 1, \ldots, 9 \\
d_i^a &= \frac{30\,(y_i^a \cdot (1 - y_i^a)/80)^{1/2}}{x_i^a}, \quad i = 1, \ldots, 9 \\
\sigma^a &= 1.0\ .
\end{aligned}$$

The values x_i^a, $i = 1, \ldots, n$, are the observed values from the original experiment. The weights w_i are calculated as the inverses of the standard deviations of the y_i. The standard deviations of the errors in the measurements of the various signal levels are known from experience to be proportional to the value of the signal itself with a proportionality constant of 1/30. The d_i^a are computed accordingly.

These 4 examples are plotted in figures 1 thru 4, respectively. The graphs display $f(x;\beta^a)$ evalutated over the range of the values x_i^a, $i = 1, \ldots, n$. The n points (x_i^a, y_i^a) are indicated by the "dots" on each curve.

For each of these models we construct 500 sets of "observed" data (x_i, y_i) using

$$\begin{aligned}
x_i &= x_i^a - \delta_i^a \qquad i = 1, \ldots, n \\
y_i &= f(x_i^a;\beta^a) - \epsilon_i^a \qquad i = 1, \ldots, n.
\end{aligned}$$

The errors $(\epsilon_1^a, \ldots, \epsilon_n^a, \delta_1^a, \ldots, \delta_n^a) \sim \mathrm{N}(0, (\sigma^a)^2[\Omega^a]^{-1})$, are generated using the Marsaglia and Tsang [1984] pseudonormal random number algorithm as implemented by James Blue and David Kahaner of the National Institute of Standards

and Technology. We construct Ω^a using (2.7) and

$$\begin{aligned} W^a &= \text{diag}\{(w_i^a)^2,\ i = 1, \ldots, n\} \\ D^a &= \text{diag}\{(w_i^a d_i^a)^2,\ i = 1, \ldots, n\}. \end{aligned}$$

For each of the 500 realizations of the data, we solve for $\hat{\eta} = (\hat{\beta}_1, \ldots, \hat{\beta}_p, \hat{\delta}_1, \ldots, \hat{\delta}_n)^{\mathrm{T}}$ using the orthogonal distance regression software library ODRPACK [Boggs *et al.*, 1989] and the ordinary least squares representation of the problem. (Recall that ODRPACK only computes the covariance matrix for the estimators $\hat{\beta}$ when solving an orthogonal distance regression problem. Using the ordinary least squares representation of the problem allows us to easily obtain the covariance matrix for the all of the estimators $\hat{\eta}$.) The computations are performed in double precision Fortran on a Sun Workstation.

Initially, η is set to $\eta^a = (\beta_1^a, \ldots, \beta_p^a, \delta_1^a, \ldots, \delta_n^a)^{\mathrm{T}}$. This is reasonable, since we are interested in assessing the reliability of the confidence intervals and not in ODRPACK's ability to obtain a solution quickly. Default values are used for all ODRPACK arguments, except for the maximum number of iterations, which is set to 100. The Jacobian matrices are computed using finite differences. We expect that use of analytic derivatives would produce a slight improvement in the results reported here.

The covariance matrix is dependent upon Ω. Clearly, however, the W and D that make up Ω are not always known. Of particular interest for orthogonal distance regression problems is the case where the values of $d_i, i = 1, \ldots, n$, used to determine D are only approximate. In addition to reporting the observed coverage for Ω constructed using d_i^a and w_i^a, we therefore also report the coverage observed when the 500 replications are solved using Ω constructed with $d_i = d_i^a/10$, $d_i = d_i^a/2$, $d_i = 2d_i^a$, $d_i = 10d_i^a$, and $d_i = \infty d_i^a$, the latter indicating an ordinary least squares solution in which all values of δ_i are forced to zero.

The observed coverages for these problems are shown in Tables 1, 2, 3, and 4. The confidence region coverage when d_i is correct is surprisingly good when compared with that observed by Donaldson and Schnabel [1987]. We conjecture that this is due to our choice of examples, and is not a property of orthogonal distance regression in general. In their study, Donaldson and Schnabel frequently found that the observed coverage for linearized confidence regions was less than 80% of the expected nominal coverages, a difference that many, if not most, users would find unacceptable. Our results show a significant degradation in the coverage of the linearized confidence regions when d_i is incorrect by even a factor of 2.

These tables indicate that the confidence interval coverage when d_i is correct is also very good. In addition, these results illustrate that, even when d_i is known only to within a factor of 2, the confidence intervals for β^a are still quite good. When d_i is under-estimated by a factor of 10, however, we see a significant degradation for two of our examples.

There is a significant degradation in the confidence interval coverage for δ_i^a when d_i is not known. As we expected, our results show that the confidence intervals for β^a are in general more reliable than those for δ_i^a.

One surprising result is that, for confidence intervals for β^a, over-estimation of d_i is preferable to under-estimation. For δ_i^a, the opposite is true. We believe that this occurs because when d_i is over-estimated, we overly restrict the size of δ_i and thus prevent $\hat{\delta}_i$ from being "close enough" to δ_i^a to allow the confidence intervals and regions to include δ_i^a. When we under-estimate d_i, on the other hand, we artificially reduce the size of the residual variance, $\hat{\sigma}^2$, and thus the size of the covariance matrix and the resulting confidence intervals for β^a.

We conclude from this small study that for at least some orthogonal distance regression problems, confidence regions and intervals constructed using the covariance matrix do have some validity, especially when d_i is known at least to within a factor of 2. We recognize, however, that for other problems such intervals and regions may be very inaccurate. We also recognize that there is nothing better that can be easily reported in their place. Thus, the linearization method will continue to be the most frequently implemented method in production least squares software. We therefore advocate the use of linearized confidence intervals and regions for measurement error problems in the same spirit, and with the same caveats, that they are used for ordinary nonlinear least squares problems.

Acknowledgement. The authors thank H. Iyer, Colorado State University, for many useful discussions and comments concerning this work, and for providing the fourth example for our Monte Carlo study.

References

[1] Bard, Y. (1974), *Nonlinear Parameter Estimation,* New York: Academic Press.

[2] Bates, D. M., and Watts, D. G. (1980), "Relative Curvature Measures of Nonlinearity," *Journal of the Royal Statistical Society*, Ser. B, 42, 1-25

[3] Boggs, P. T., Byrd, R. H., and Schnabel, R. B. (1987), "A Stable and Efficient Algorithm for Nonlinear Orthogonal Distance Regression," *SIAM Journal of Scientific and Statatistical Computing*, 8(6), 1052-1078.

[4] Boggs, P. T., Donaldson, J. R., Byrd, R. H. and Schnabel, R. B. (1989), "ODRPACK — Software for Weighted Orthogonal Distance Regression," to appear in *ACM Transactions on Mathematical Software.*

[5] Boggs, P. T., Donaldson, J. R., Schnabel, R. B., and Spiegelman, C. H. (1988), "A Computational Examination of Orthogonal Distance Regression," *Journal of Econometrics*, 38(1/2), 169-201.

[6] Donaldson, J. R., and Schnabel, R. B. (1987), "Computational Experience with Confidence Regions and Confidence Intervals for Nonlinear Least Squares," *Technometrics*, 29(1), 67-82.

[7] Dongarra, J. J., Moler, C. B., Bunch, J. R., and Stewart, G. W. (1979), *LINPACK Users' Guide*, Philadelphia: SIAM.

[8] Draper, N. R., and Smith, H. (1981), *Applied Regression Analysis, Second Edition*, New York: John Wiley and Sons.

[9] Efron, B. (1985), *The Jackknife, the Bootstrap and Other Resampling Plans,* Philadelphia: SIAM.

[10] Fuller, W. A. (1987), *Measurement Error Models*, New York: John Wiley and Sons.

[11] Jennrich, R. I. (1969), "Asymptotic Properties of Non-linear Least Squares Estimators," *Annals of Mathematical Statistics*, 40, 633-643.

[12] Ratkowsky, D. A. (1983), *Nonlinear Regression Modeling*, New York: Marcel Dekker.

Table 1: Fuller Example 3.2.2

	Observed Coverage for Nominal 95% Confidence Regions and Intervals					
	$d_i = \frac{1}{10}d_i^a$	$d_i = \frac{1}{2}d_i^a$	$d_i = d_i^a$	$d_i = 2d_i^a$	$d_i = 10d_i^a$	$d_i = \infty d_i^a$ (OLS)
c.r. for η^a	16.8	78.4	93.6	86.6	0.0	—
c.r. for β^a	42.8	85.0	93.8	96.6	95.4	95.2
c.r. for δ^a	18.4	82.6	93.4	81.2	0.0	—
c.i. for β_1^a	50.0	84.4	93.8	97.2	97.4	97.4
c.i. for β_2^a	51.4	85.8	94.6	96.0	96.8	96.6
c.i. for β_3^a	89.4	92.2	93.8	94.4	95.2	95.6
c.i. for δ_3^a	68.8	89.4	96.0	92.4	25.4	—
c.i. for δ_4^a	77.8	89.6	93.6	92.4	33.0	—
c.i. for δ_5^a	80.0	89.8	94.4	91.8	27.6	—
c.i. for δ_6^a	85.0	91.2	97.0	93.0	31.0	—
c.i. for δ_7^a	88.2	92.6	94.4	87.4	27.8	—
c.i. for δ_8^a	93.0	93.8	92.2	83.6	25.8	—
c.i. for δ_9^a	89.0	96.2	95.8	85.8	28.6	—
c.i. for δ_{10}^a	90.0	94.6	93.4	82.8	31.2	—
c.i. for δ_{11}^a	89.0	95.4	93.4	85.6	30.6	—

Table 2: Ratkowsky Example 6.11

	Observed Coverage for Nominal 95% Confidence Regions and Intervals					
	$d_i = \frac{1}{10}d_i^a$	$d_i = \frac{1}{2}d_i^a$	$d_i = d_i^a$	$d_i = 2d_i^a$	$d_i = 10d_i^a$	$d_i = \infty d_i^a$ (OLS)
c.r. for η^a	0.0	99.6	96.2	28.0	0.0	—
c.r. for β^a	94.6	93.6	93.4	93.4	93.4	93.4
c.r. for δ^a	0.0	100.0	96.0	21.4	0.0	—
c.i. for β_1^a	95.4	95.2	95.2	95.0	95.0	95.0
c.i. for β_2^a	95.4	95.2	95.2	95.0	95.0	95.0
c.i. for β_3^a	95.4	95.2	95.2	95.0	95.0	95.0
c.i. for δ_1^a	97.8	99.6	95.0	69.4	14.4	—
c.i. for δ_2^a	87.8	99.0	96.2	74.2	18.2	—
c.i. for δ_3^a	82.8	99.8	94.6	69.2	17.6	—
c.i. for δ_4^a	73.0	99.0	95.2	70.0	15.4	—
c.i. for δ_5^a	73.8	99.4	93.2	70.8	17.4	—
c.i. for δ_6^a	71.4	99.2	94.4	74.4	16.4	—
c.i. for δ_7^a	75.0	99.2	95.4	75.0	15.4	—
c.i. for δ_8^a	73.8	99.6	95.2	73.8	17.6	—
c.i. for δ_9^a	72.2	99.2	95.2	72.6	18.2	—
c.i. for δ_{10}^a	75.4	99.6	96.4	70.6	19.0	—
c.i. for δ_{11}^a	69.4	98.2	94.2	74.2	18.8	—
c.i. for δ_{12}^a	69.8	99.0	94.4	72.6	20.6	—
c.i. for δ_{13}^a	69.6	99.0	95.8	68.2	16.8	—
c.i. for δ_{14}^a	76.2	99.0	94.0	71.0	14.6	—
c.i. for δ_{15}^a	84.0	99.8	95.4	73.0	14.4	—
c.i. for δ_{16}^a	95.6	99.6	95.0	70.6	17.0	—

Table 3: Draper and Smith Problem 10.E

	Observed Coverage for Nominal 95% Confidence Regions and Intervals					
	$d_i = \frac{1}{10}d_i^a$	$d_i = \frac{1}{2}d_i^a$	$d_i = d_i^a$	$d_i = 2d_i^a$	$d_i = 10d_i^a$	$d_i = \infty d_i^a$ (OLS)
c.r. for η^a	13.2	85.0	92.4	72.2	0.0	—
c.r. for β^a	80.4	92.6	90.8	87.4	78.0	76.8
c.r. for δ^a	28.6	94.0	94.8	68.0	0.0	—
c.i. for β_1^a	71.8	89.8	94.8	96.0	96.6	96.4
c.i. for β_2^a	86.6	93.2	94.0	91.8	89.4	89.6
c.i. for β_3^a	82.2	92.6	95.0	93.6	93.0	92.8
c.i. for δ_1^a	100.0	98.8	94.4	82.2	29.2	—
c.i. for δ_2^a	99.8	99.2	93.2	82.2	30.0	—
c.i. for δ_3^a	92.6	99.2	94.6	78.8	25.0	—
c.i. for δ_4^a	77.2	98.2	94.4	79.4	25.4	—
c.i. for δ_5^a	82.6	97.0	94.4	79.8	26.2	—
c.i. for δ_6^a	93.8	97.4	95.4	81.0	25.8	—
c.i. for δ_7^a	96.8	96.0	96.0	83.4	25.8	—
c.i. for δ_8^a	96.2	97.2	95.6	83.2	27.6	—
c.i. for δ_9^a	98.2	97.0	96.0	85.6	23.8	—
c.i. for δ_{10}^a	98.2	96.4	95.8	88.2	27.8	—
c.i. for δ_{11}^a	99.4	96.0	95.4	90.2	28.0	—
c.i. for δ_{12}^a	99.6	97.4	94.0	87.0	26.8	—
c.i. for δ_{13}^a	100.0	98.6	95.2	85.6	25.2	—
c.i. for δ_{14}^a	99.8	98.4	94.8	86.6	26.8	—

Table 4: Psychophysical Example

	Observed Coverage for Nominal 95% Confidence Regions and Intervals					
	$d_i = \frac{1}{10} d_i^a$	$d_i = \frac{1}{2} d_i^a$	$d_i = d_i^a$	$d_i = 2d_i^a$	$d_i = 10 d_i^a$	$d_i = \infty d_i^a$ (OLS)
c.r. for η^a	46.6	59.4	57.4	37.6	0.0	—
c.r. for β^a	53.8	52.0	48.2	47.6	46.4	46.4
c.r. for δ^a	98.0	100.0	95.8	59.4	0.0	—
c.i. for β_1^a	86.0	94.8	95.0	95.2	95.4	95.4
c.i. for β_2^a	100.0	99.8	99.8	99.8	99.8	99.8
c.i. for β_3^a	93.4	96.6	95.8	95.4	95.4	95.4
c.i. for β_4^a	99.8	93.0	90.2	89.4	89.0	89.0
c.i. for δ_1^a	100.0	99.6	95.0	75.6	22.0	—
c.i. for δ_2^a	99.8	99.6	96.4	78.2	22.4	—
c.i. for δ_3^a	100.0	99.4	95.2	76.4	20.2	—
c.i. for δ_4^a	100.0	99.4	96.4	75.8	19.4	—
c.i. for δ_5^a	100.0	99.8	95.0	73.6	18.2	—
c.i. for δ_6^a	99.8	99.4	95.0	75.0	18.8	—
c.i. for δ_7^a	98.8	99.6	95.6	77.2	17.8	—
c.i. for δ_8^a	100.0	99.4	94.8	72.2	20.4	—
c.i. for δ_9^a	100.0	99.4	95.2	77.6	21.0	—

Figure 1: Fuller Example 3.2.2

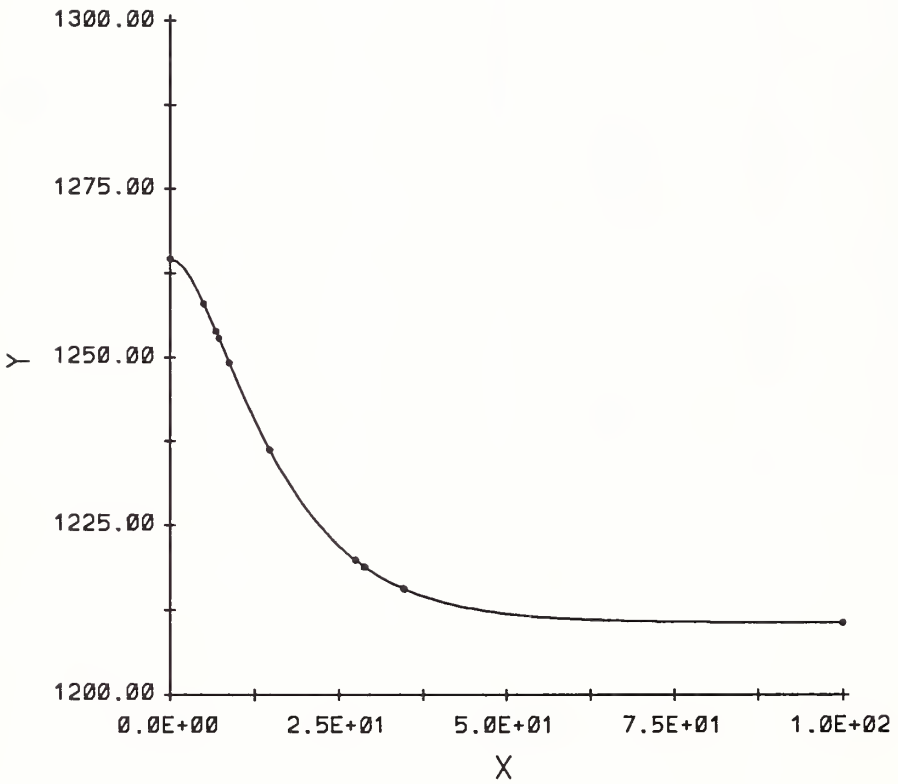


Figure 2: Ratkowksy Example 6.11

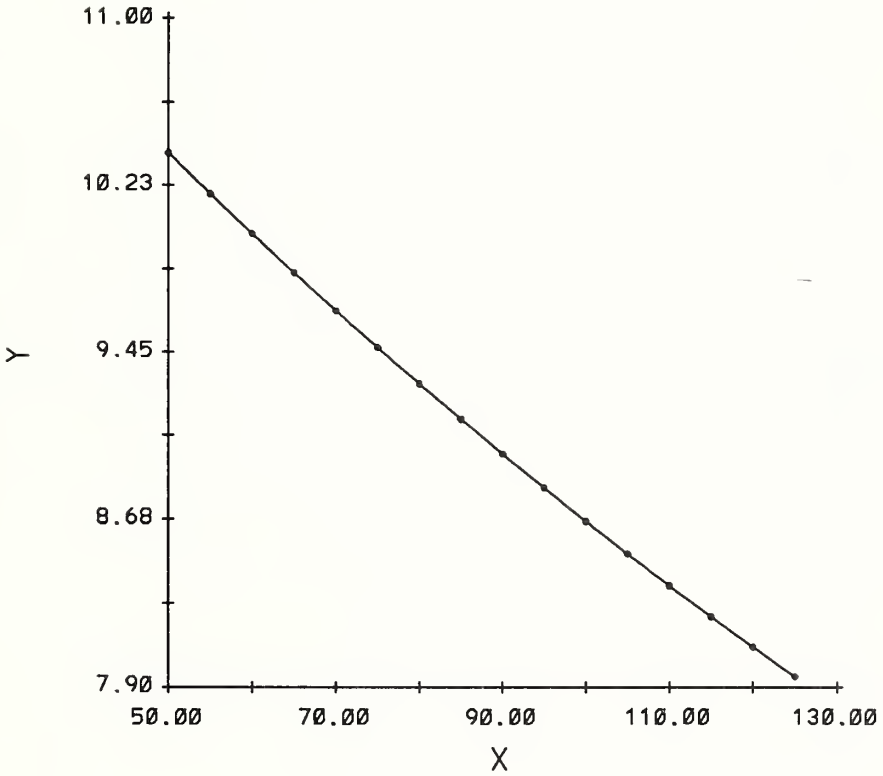


Figure 3: Draper and Smith Problem 10.E

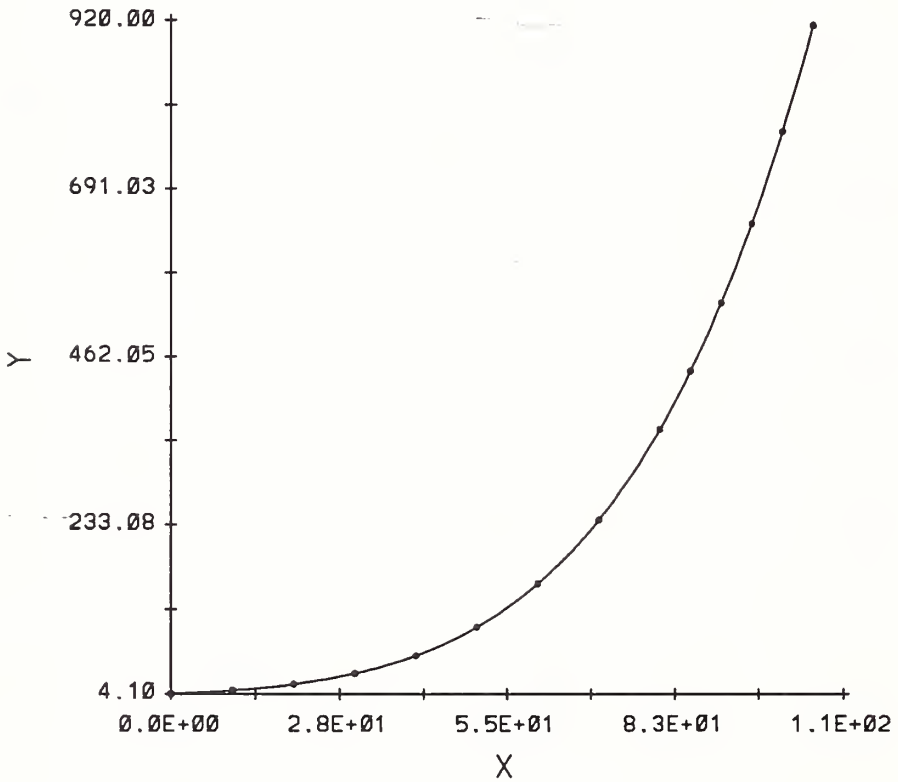
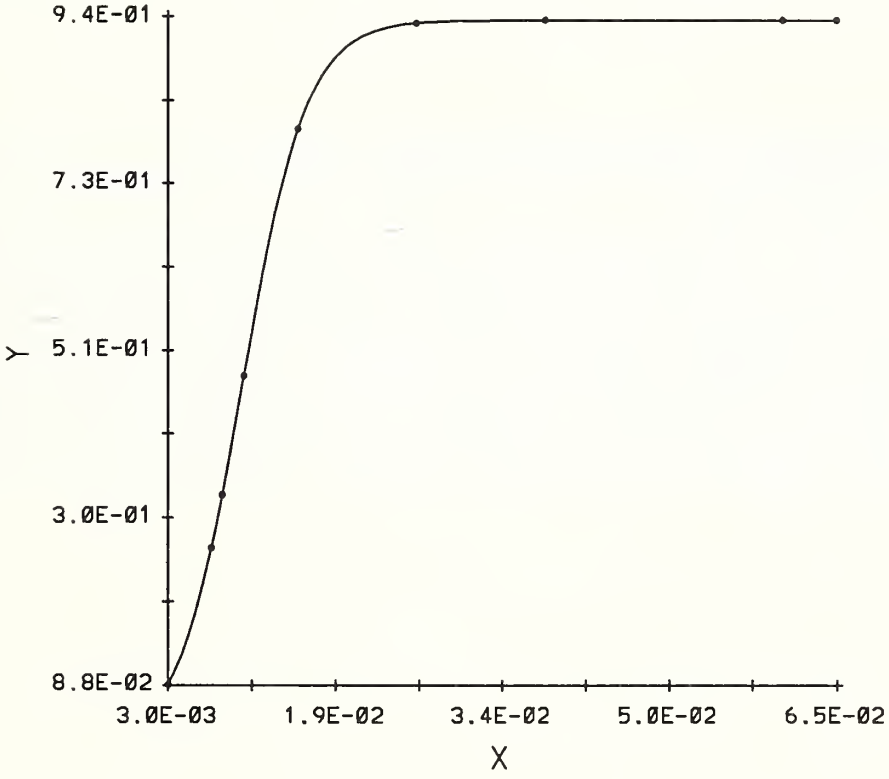


Figure 4: Psychophysical Example



U.S. DEPT. OF COMM. BIBLIOGRAPHIC DATA SHEET (See instructions)	1. PUBLICATION OR REPORT NO. NISTIR 89-4102	2. Performing Organ. Report No.	3. Publication Date June 1989

4. TITLE AND SUBTITLE

The Computation and Use of the Asymptotic Covariance Matrix for Measurement Error Models

5. AUTHOR(S)

Paul T. Boggs and Janet R. Donaldson

6. PERFORMING ORGANIZATION (If joint or other than NBS, see instructions)

NATIONAL BUREAU OF STANDARDS
U.S. DEPARTMENT OF COMMERCE
GAITHERSBURG, MD 20899

7. Contract/Grant No.

8. Type of Report & Period Covered

9. SPONSORING ORGANIZATION NAME AND COMPLETE ADDRESS (Street, City, State, ZIP)

10. SUPPLEMENTARY NOTES

☐ Document describes a computer program; SF-185, FIPS Software Summary, is attached.

11. ABSTRACT (A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here)

The measurement error model assumes that errors occur in both the response variables and the predictor variables. In using this model, it is of interest to compute confidence regions and intervals for the estimators of the model parameters. Fuller Fuller, 1987 provides an asymptotic form for the covariance matrix that can be used to construct approximate confidence regions and intervals. We discuss the solution of the minimization problem resulting from the use of the measurement error model, and we develop a procedure for accurately computing the covariance matrix. We then assess the quality of the confidence regions and intervals constructed from this matrix via a Monte Carlo study.

12. KEY WORDS (Six to twelve entries; alphabetical order; capitalize only proper names; and separate key words by semicolons)

confidence intervals; confidence regions; covariance matrix; errors in variables; measurement error models; Monte Carlo study; ordinary least squares; orthogonal distance regression; simultaneous equations models

13. AVAILABILITY

☒ Unlimited

☐ For Official Distribution. Do Not Release to NTIS

☐ Order From Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402.

☒ Order From National Technical Information Service (NTIS), Springfield, VA. 22161

14. NO. OF PRINTED PAGES

29

15. Price

$12.95