Linear Regression to a Lower Order Model: Effects and Implications

M. L. Graham
Combat Control Systems Department

AD-A218 258

DTIC
S ELECTE D
FEB 21 1990
B

Naval Underwater Systems Center
Newport, Rhode Island / New London, Connecticut

PREFACE

This study was conducted under the Submarine Contact Management Task of the Combat Control Technology Block Program, Project No. RJ14F61, principal investigator K.F. Gong (Code 2221), NUSC block program manager N.A. Beck (Code 22101). The sponsoring activity is the Office of the Chief of Naval Research, program manager D.C. Houser (OCNR-232).

The technical reviewer for this report was K.F. Gong (Code 2221).

REVIEWED AND APPROVED: 15 NOVEMBER 1989

J.R. Short
Head, Combat Control Systems Department

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	15 November 1989	Final

4. TITLE AND SUBTITLE
Linear Regression to a Lower Order Model: Effects and Implications

5. FUNDING NUMBERS
PN RJ14F61

6. AUTHOR(S)
M.L. Graham

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Underwater Systems Center
Newport Laboratory
Newport, RI 02841-5047

8. PERFORMING ORGANIZATION REPORT NUMBER

TD 6848

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of the Chief of Naval Research
OCNR-232
Washington, DC 20362

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

Linear regression is used extensively in the fields of science, engineering, and business. Because data-gathering processes can be complex, show random effects, or be unknown, oftentimes the regression model used only approximates the actual process model. This report analyzes the effects of using a reduced-order process model in linear regression. In particular, the relationship of the Taylor series coefficients to the regression parameters is discussed. Relationships are generated to equate regression and Taylor series parameters, and an error analysis is performed to compare the effects of noise and modeling errors.

14. SUBJECT TERMS
Linear regression
Lower order model

15. NUMBER OF PAGES
15

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR

NSN 7540-01-280-5500

Standard Form 298 (Rev 2-89)

TABLE OF CONTENTS

LINEAR REGRESSION TO A LOWER ORDER MODEL:
EFFECTS AND IMPLICATIONS

INTRODUCTION

Linear regression is used extensively in the fields of science, engineering, and business. In many instances, because the data-generating process can be complex, exhibit random effects, or be unknown, the regression model used only approximates the actual process model. For these reasons, regression models of reduced order are often used. Such models are designed to be as accurate as necessary for their intended application without retaining unnecessary and noise-sensitive, higher order terms.

When regression is used for noise suppression, the order of the regression model required can be a function of the noise level encountered. Under high-noise conditions, the errors incurred by using a reduced-order regression model can be negligible when compared with the noise uncertainty. Under low-noise conditions, however, the errors can become significant.

When the regression parameters are to be related to process state parameters for subsequent processing, it is very important that the relationship be properly formed. Use of a reduced-order regression model can produce unexpected biases, which can be accounted for in the state relation if interpreted properly. Such a problem was encountered when hierarchical processing was applied to the contact motion analysis problem. In this application, a bearing sequence, which is related to the state by means of the arctangent function, was characterized by a second-order regression model. When relating the regression coefficients to the state parameters, the bias resulting from the nonzero derivatives of the arctangent function proved significant under low-noise conditions. Application of the analysis presented in this report allowed for bias compensation and the concomitant improvement in estimator performance.

This report presents a brief review of linear regression and an analysis of the effects of using a reduced-order regression model. In addition, a parallel analysis using the Householder transformation, which provides a convenient computational scheme for bias compensation, is discussed.

1

LEAST-SQUARES REGRESSION

A process of some fixed order can be represented by a finite Taylor series about a chosen point. If noise-free measurements are available, the Taylor series coefficients can be calculated exactly provided the number of measurements is at least equal to the order of the process. In this case, the system of measurement equations, with the Taylor series coefficients as unknowns, forms a complete set of linear equations. When the number of measurements is greater than the order of the process, the system is over determined, and redundant equations can be ignored. The problem becomes more complex when measurement noise is present.

The most common method of solving an overdetermined system of equations in the presence of noise is to use the method of least squares (references 1 and 2). Consider the system of equations

$$Ax = b, \tag{1}$$

where x is the n × 1 vector of unknown coefficients, A is the m × n system matrix (m > n) created from the samples of the independent variable, and b is the m × 1 measurement vector. The least-squares solution is derived by first introducing the m × 1 error vector

$$e = b - Ax. \tag{2}$$

The objective in the least-squares technique is to minimize the squared magnitude of the error vector, so that

$$||e||^2 = e^T e = (b - Ax)^T(b - Ax), \tag{3}$$

giving the least-squares solution

$$\bar{x} = (A^T A)^{-1} A^T b = A^\# b. \tag{4}$$

The matrix $A^\# = (A^T A)^{-1} A^T$, which is the orthogonal projection matrix that projects an arbitrary vector into the subspace spanned by Ax, is called the generalized or pseudoinverse of the matrix A (references 1 and 2).

In the problem at hand, the rows of the matrix equation (Ax = b) are simply samples of the Taylor series representation, which results in the matrix A assuming a distinct form. To illustrate the form of A, consider the nth order Taylor series polynomial about the point t_0 given by

$$b(t) = x_0 + x_1(t - t_0) + x_2(t - t_0)^2 + \ldots + x_n(t - t_0)^n, \tag{5}$$

where

$$x_0 = b(t_0),$$

$$x_i = b^i(t_0)/(i)!, \quad i > 0, \tag{6}$$

with $b^i(t)$ being the ith derivative of b(t). With no loss in generality.

2

$t_0 = 0$ may be selected. For m samples of the function taken at points $t_1, t_2, \ldots, t_m$, the resulting set of equations is

$$b(t_1) = x_0 + x_1 t_1 + x_2 t_1^2 + \ldots + x_n t_1^n,$$

$$b(t_2) = x_0 + x_1 t_2 + x_2 t_2^2 + \ldots + x_n t_2^n, \qquad (7)$$

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

$$b(t_m) = x_0 + x_1 t_m + x_2 t_m^2 + \ldots + x_n t_m^n,$$

which may be written in the form of equation (1) with matrix A defined as

$$
A = \begin{bmatrix}
1 & t_1 & t_1^2 & \ldots & t_1^n \\
1 & t_2 & t_2^2 & \ldots & t_2^n \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
1 & t_m & t_m^2 & \ldots & t_m^n
\end{bmatrix}, \qquad (8)
$$

and

$$b = [b(t_1), b(t_2), \ldots, b(t_m)]^T.$$

An estimate of the nth order Taylor series coefficients x based on m noisy measurement samples $y_i = b(t_i) + w_i$, with w_i being the measurement noise, may be obtained by solving the least-squares equation (4) using the system matrix of equation (8) and $b = [y_1, y_2, \ldots, y_m]$.

The matrix A takes a particularly convenient form when the measurement points t_i are uniformly spaced and (again, with no loss in generality) are symmetrical about the point $t_0 = 0$; hence, $t_k = k\Delta t$ for $k = (-m/2, m/2)$. (Here, m is considered to be an even number for simplicity; similar results hold for m odd.) Now, the matrix A has the form

$$
A = \begin{bmatrix}
1 & (-m/2)\Delta t & [(-m/2)\Delta t]^2 & \ldots & [(-m/2)\Delta t]^n \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
1 & (+m/2)\Delta t & [(+m/2)\Delta t]^2 & \ldots & [(+m/2)\Delta t]^n
\end{bmatrix}, \qquad (9)
$$

and $A^T A$ is given by

$$
A = \begin{bmatrix}
m+1 & \Sigma(i\Delta t) & \Sigma(i\Delta t)^2 & \ldots & \Sigma(i\Delta t)^n \\
\Sigma(i\Delta t) & \vdots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\Sigma(i\Delta)^{n-1} & \ldots & \ldots & \ldots & \Sigma(i\Delta t)^{2n}
\end{bmatrix}, \qquad (10)
$$

where the summation is over $i = -m/2$ to $m/2$. Again, an estimate of the coefficient vector x may be obtained by solving the least-squares equation (4). Estimation using the least-squares criterion is nothing new; what is of interest here is the implication of using a model order for estimation that is lower than the actual process model order.

ESTIMATION TO A LOWER ORDER MODEL

When the process of interest is characterized by a Taylor series, for many practical applications, the higher order terms in the series are of significantly lower magnitude when compared with the lower order terms. It is often the case that the noise level will result in estimation errors that are comparable with the magnitude of these higher order items. For this reason, it is often desirable to model the system with a lower order, providing a more robust estimation of the desired coefficients. Also, in some instances, it may be possible for the Taylor series coefficients of the higher order terms to be expressed as functions of the lower order coefficients. Under such conditions, the lower order terms completely describe the system dynamics, even though there may be nonzero, higher order terms. Here again, it may be desirable to estimate the coefficients for the lowest order model to satisfactorily describe the system dynamics. However, care must be taken when interpreting the resulting coefficient estimates, because the basis functions of the Taylor series are not orthogonal. The interpretation becomes particularly important when the coefficient estimates are to be compared with predictions from a state parameter model.

To illustrate the effects of fitting to a lower order, consider the following process described by the fourth-order model:

$$y(t) = q_0 + q_1 t + q_2 t^2 + q_3 t^3 + q_4 t^4. \tag{11}$$

Then, consider the following case where the estimation is performed for a model of the second order:

$$Y(t) = x_0 + x_1 t + x_2 t^2. \tag{12}$$

Again, for simplicity, consider a symmetric time interval, which has been uniformly sampled. The impact of the reduced-order estimation can be evaluated by solving the least-squares equations under noise-free conditions. Under these conditions, the measurements are simply noise-free samples of the process function, which is shown as

$$b = \begin{bmatrix} q_0 + q_1(-m/2)\Delta t + \ldots + q_4[(-m/2)\Delta t]^4 \\ \vdots \quad \vdots \quad \vdots \quad \vdots \\ q_0 + q_1(0)\Delta t \quad + \ldots + q_4[(0)\Delta t]^4 \\ \vdots \quad \vdots \quad \vdots \quad \vdots \\ q_0 + q_1(m/2)\Delta t \quad + \ldots + q_4[(m/2)\Delta t]^4 \end{bmatrix}. \tag{13}$$

4

This may be written as $b = Qq$, where Q is the $(m \times 5)$ matrix in the form of equation (9) and q is the vector of coefficients from equation (11), such that

$$q = [q_0, q_1, q_2, q_3, q_4]. \tag{14}$$

From equations (9) and (12), it may be seen that the matrix A is given by

$$A = \begin{bmatrix} 1 & (-m/2)\Delta t & \ldots & [(-m/2)\Delta t]^2 \\ \vdots & \vdots & \vdots & \vdots \\ \cdot & (0)\Delta t & \cdot & [(0)\Delta t]^2 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 1 & (m/2)\Delta t & \ldots & [(m/2)\Delta t]^2 \end{bmatrix}. \tag{15}$$

From equation (10), $A^T A$ takes the form

$$A^T A = \begin{bmatrix} m+1 & 0 & 2(\Delta t)^2 \sum_{i=1}^{m/2} (i)^2 \\ 0 & 2(\Delta t)^2 \sum_{i=1}^{m/2} (i)^2 & 0 \\ 2(\Delta t)^2 \sum_{i=1}^{m/2} (i)^2 & 0 & 2(\Delta t)^4 \sum_{i=1}^{m/2} (i)^4 \end{bmatrix}, \tag{16}$$

and the product $A^T b$ takes the form

$$A^T b = A^T Q q, \tag{17}$$

$$A^T Q = \begin{bmatrix} m+1 & 0 & 2(\Delta t)^2 \sum_{i=1}^{m/2} (i)^2 & 0 & 2(\Delta t)^4 \sum_{i=1}^{m/2} (i)^4 \\ 0 & 2(\Delta t)^2 \sum_{i=1}^{m/2} (i)^2 & 0 & 2(\Delta t)^4 \sum_{i=1}^{m/2} (i)^4 & 0 \\ 2(\Delta t)^4 \sum_{i=1}^{m/2} (i)^2 & 0 & 2(\Delta t)^4 \sum_{i=1}^{m/2} (i)^4 & 0 & 2(\Delta t)^6 \sum_{i=1}^{m/2} (i)^6 \end{bmatrix}. \tag{18}$$

The sums in equations (16) and (18) may be replaced by their closed-form equivalents and for large m approximated by the most significant term. The resulting approximations follow, i.e.,

$$A^T A = \begin{bmatrix} m & 0 & \dfrac{(\Delta t)^2 m^3}{12} \\[2ex] 0 & \dfrac{(\Delta t)^2 m^3}{12} & 0 \\[2ex] \dfrac{(\Delta t)^2 m^3}{12} & 0 & \dfrac{(\Delta t)^4 m^5}{80} \end{bmatrix}, \tag{19}$$

$$A^T Q = \begin{bmatrix} m & 0 & \dfrac{(\Delta t)^2 m^3}{12} & 0 & \dfrac{(\Delta t)^4 m^5}{80} \\[2ex] 0 & \dfrac{(\Delta t)^2 m^3}{12} & 0 & \dfrac{(\Delta t)^4 m^5}{80} & 0 \\[2ex] \dfrac{(\Delta t)^2 m^3}{12} & 0 & \dfrac{(\Delta t)^4 m^5}{80} & 0 & \dfrac{(\Delta t)^6 m^7}{448} \end{bmatrix}, \tag{20}$$

and

$$(A^T A)^{-1} A^T Q = \begin{bmatrix} 1 & 0 & 0 & 0 & \dfrac{-3(\Delta t)^4 m^4}{560} \\[2ex] 0 & 1 & 0 & \dfrac{3(\Delta t)^2 m^2}{20} & 0 \\[2ex] 0 & 0 & 1 & 0 & \dfrac{3(\Delta t)^2 m^2}{14} \end{bmatrix}. \tag{21}$$

From the resulting least-squares solution for x

$$\hat{x} = (A^T A)^{-1} A^T Q q, \tag{22}$$

it may be seen that the estimates of the components of x are not exactly equal to the corresponding components of q. As an example, look at the estimate of x_0, such that

$$\hat{x}_0 = q_0 - \left[\frac{3(\Delta t)^4 m^4}{560} \right] q_4. \tag{23}$$

The use of a reduced-order estimation model produces a bias on the Taylor series coefficient estimates.

To evaluate whether this biasing effect is significant, one must look at its magnitude relative to the estimated noise variance. From the Cramer-Rao inequality, it may be seen that the covariance of an unbiased estimator is bounded below by the inverse of the Fisher information matrix (reference 2). While x_0 is a biased estimate of q_0, it is an unbiased estimate of the parameter for which the right-hand side of equation (23) is an approximation. This is also the case for the other elements of the vector x. The inverse of the Fisher information matrix, $\sigma_w^2 (A^T A)^{-1}$ for homoskedastic noise with variance σ_w^2 (reference 2), is a lower bound to the covariance on x, so that

6

$$\sigma_w^2 (A^T A)^{-1} = \sigma_w^2 \begin{bmatrix} \dfrac{9}{4m} & 0 & \dfrac{-15}{(\Delta t)^2 m^3} \\ 0 & \dfrac{12}{(\Delta t)^2 m^3} & 0 \\ \dfrac{-15}{(\Delta t)^2 m^3} & 0 & \dfrac{180}{(\Delta t)^4 m^5} \end{bmatrix}. \tag{24}$$

It may be seen that $\sigma_{x0}^2 > (9/4m)\sigma_w^2$ and, from equation (23), the bias error due to the reduced-order fit is

$$e_b = \left[\frac{3(\Delta t)^4 m^4 q_4}{560} \right]. \tag{25}$$

The bias effect becomes significant when the corresponding standard deviation is on the order of the bias magnitude

$$\sigma_w \sim \frac{(\Delta t)^4 m^{9/2} q_4}{280}. \tag{26}$$

This gives the noise level or, conversely, the magnitude of the process parameter q_4, for which the biasing effect of using a lower order model becomes significant.

To illustrate that the use of a reduced-order model may be advantageous from a minimum variance perspective, look at the Cramer-Rao bound for a first-order model, when

$$\sigma_w^2 (A^T A)^{-1} = \sigma_w^2 \begin{bmatrix} (1/m) & 0 \\ 0 & \dfrac{12}{(\Delta t)^2 m^3} \end{bmatrix}. \tag{27}$$

It may be seen that the variance on the x_0 parameter for a first-order model is four-ninths of the corresponding second-order model estimate. This illustrates the fact that, while using a reduced-order model may introduce a bias, it can provide an estimate with significantly lower variance. In situations where the bias can be accounted for, or under relatively high-noise conditions, it is most dvantageous to use a model of minimal order.

BIAS COMPENSATION VIA THE HOUSEHOLDER TRANSFORMATION

Use of the pseudoinverse for the solution to the normal equations is convenient for the derivation of an analytic expression for noise-free regression coefficients. However, the analysis uses the assumption of a constant data rate and is further simplified by using approximate expressions for the sums involved. For large or poorly conditioned systems, the matrix inversion required for the pseudoinverse solution can be computationally intensive and problematical (reference 2). To avoid these difficulties,

numerical techniques have been developed that eliminate the need for an explicit matrix inversion. One technique commonly employed is the Householder transformation (references 2 and 3).

The Householder tranformation is a numerical technique used to effect an orthogonal transformation on the regression model. Without delving into details of the Householder procedure, the results of interest may be obtained by looking at the transformation matrix to which it effectively applies. The Householder matrix H transforms the regression model into the upper triangular form

$$HA = \begin{bmatrix} S_x \\ 0 \end{bmatrix}, \quad Hb = \begin{bmatrix} \xi_x \\ \xi_e \end{bmatrix}, \tag{28}$$

where S_x is an $n \times n$ upper triangular matrix, ξ_x is an $n \times 1$ vector, and ξ_e is an $(m - n) \times 1$ vector. Using the orthogonal property of the Householder matrix $(H^T H = I)$, the least-squares criterion may be rewritten as

$$||e||^2 = e^T e = (b - Ax)^T (b - Ax),$$

$$= (b - Ax)^T H^T H (b - Ax),$$

$$= (Hb - HAx)^T (Hb - HAx),$$

$$= \left(\begin{bmatrix} \xi_x \\ \xi_e \end{bmatrix} - \begin{bmatrix} S_x \\ 0 \end{bmatrix} x \right)^T \left(\begin{bmatrix} \xi_x \\ \xi_e \end{bmatrix} - \begin{bmatrix} S_x \\ 0 \end{bmatrix} x \right), \tag{29}$$

or

$$||e||^2 = (\xi_x - S_x x)^T (\xi_x - S_x x) + \xi_e^T \xi_e. \tag{30}$$

The second term in equation (30) is independent of x; hence, the least-squares solution is found by choosing the x that causes the first term to vanish, i.e.,

$$\dot{x} = S_x^{-1} \xi_x. \tag{31}$$

Because S_x is an upper triangular matrix, the inversion involved in equation (31) is straightforward.

Again, the order of the process model is considered to be greater than the regression model. Under noise-free conditions, the measurement vector is generated by $b = Qq$, where the matrix Q is of the form of equation (8), and it can be noted that

$$Q = [A : \bar{A}], \tag{32}$$

where A is the matrix for the lower order regression model and $\bar{A}$ contains the higher order components. Applying the Householder transformation to the resulting regression model makes the squared error

8

$$\|e\|^2 = (Hb - HAx)^T(Hb - HAx),$$

$$= \left(\begin{bmatrix} S_x & S_x \\ 0 & S_c \end{bmatrix} q - \begin{bmatrix} S_x \\ 0 \end{bmatrix} x \right)^T \left(\begin{bmatrix} S_x & S_x \\ 0 & S_e \end{bmatrix} q - \begin{bmatrix} S_x \\ 0 \end{bmatrix} \right)$$

$$= ([S_x : S_x]q - S_x x)^T ([S_x : S_x]q - S_x x)$$

$$+ ([0 : S_e]q)^T ([0 : S_e]q). \tag{33}$$

Again, the second term is independent of x and represents the magnitude of the squared error resulting from the use of a reduced-order regression model. The least-squares solution is obtained by choosing x so that the first term vanishes, i.e.,

$$\hat{x} = S_x^{-1}[S_x : S_x]q,$$

$$= [I : S_x^{-1}S_x]q. \tag{34}$$

The matrix in equation (34) is a general form of that found in equation (22), with none of the approximations or assumptions of constant data rate or symmetric interval. It is readily produced as a byproduct of the Householder procedure and is not a function of the measurement noise.

APPLICATION TO HIERARCHICAL ESTIMATION

The results of the previous sections become particularly relevant when the reduced-order regression is performed as the first stage in a hierarchical estimation procedure. As such, locally estimated regression parameters serve as "pseudomeasurements" for a second stage of estimation, which combines the local parameters from multiple data segments to provide a global state estimate. As stated earlier, it may be advantageous to perform a minimal-order local estimation to minimize noise effects. When these local estimates are used as pseudomeasurements, the global measurement model takes the form of equation (34), i.e.,

$$b(\theta) = [I : S_x^{-1}S_x]q(\theta), \tag{35}$$

where $b(\theta)$ is the predicted pseudomeasurement that corresponds to the solution $\hat{x}$ provided by the local regression. The vector $q(\theta)$ comprises the Taylor series coefficients based on the global state estimate θ. Use of the measurement model of equation (35) limits the modeling bias errors to a value determined by the order to which $q(\theta)$ is extended, while retaining the noise characteristics associated with the regression to the order of $b(\theta)$.

These developments were applied to the nonlinear state estimation problem using data segmentation and compression (reference 4). Here, bearing data, which are related to the state through the arctangent function, are character-ized on a segment by a second-order polynomial. For each segment, the result-ing estimates are used as the input to a second stage of processing that

9

performs the global nonlinear state estimation. For the problem at hand, only bearing and its first two derivatives $(\beta, \dot\beta, \ddot\beta)$ are independent, and all higher time derivatives of bearing can be written in terms of $\beta, \dot\beta, \ddot\beta$. The use of a second-order model to characterize the bearing curve results in biased estimates of the Taylor series coefficients due to the nonzero higher derivatives. However, when calculating the predicted polynomial coefficients $b(\Theta)$ from the current state estimate Θ, it is possible to use the correction terms of equation (35) by simply using the predicted value of $q(\Theta)$. That is, the components of q through q_4 are retained (corresponding to bearing time derivatives up to $\beta^{(4)}$). This results in a bias compensation that is accurate to the fourth order, while the regression is carried to only a second-order model and, hence, has lower variance than a fourth-order regression. The resulting estimation algorithm performs well under both high-noise and low-noise conditions.

SUMMARY

Because the actual data-gathering process can be complex, exhibit random effects, or be unknown, it is often useful to implement a minimal-order model when estimating process parameters. In cases where robustness and accuracy are of interest and the estimation algorithm must operate under both high-noise and low-noise conditions, it may be necessary to account for the biasing effects of higher order terms on the estimates of lower order parameters. While the biasing can be unnoticeable under high-noise conditions or for short observation intervals, it can account for a significant percentage of the errors under low noise or for long observation intervals. This circumstance has been illustrated for the case of a linear-regression model in additive Gaussian noise. The resulting analysis proved useful in bias compensation for a hierarchical estimation technique.

REFERENCES

1. W.L. Brogan, *Modern Control Theory*, Prentice-Hall Inc., Englewood Cliffs, NJ, 1982.

2. H.W. Sorenson, *Parameter Estimation, Principles, and Problems*, Marcel Dekker Inc., New York, 1980.

3. A.O. Steinhardt, "Householder Transforms in Signal Processing," *IEEE Proceedings for Acoustics, Speech, and Signal Processing (ASSP)*, July 1988.

4. A.G. Lindgren, K.F. Gong, and M.L. Graham. "Nonlinear Parameter Estimation with Segmented Data: Trajectory Estimation with Biased Measurements," *Proceedings of the 19th IEEE Asilomar Conference on Signals, Systems and Computers*, November 1985.

INITIAL DISTRIBUTION LIST

Addressee	No. of Copies
CNR (OCNR-00, OCNR-11, OCNR-232 (D. Houser (2)), OCNR-122)	5
NRL	1
NOSC (Code 632 (R. Moore))	2
DTIC	2
CNA	1