



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/75829/>

Monograph:

Billings, S.A. (1979) The Equivalence of Two Least Squares Algorithms. Research Report. ACSE Report 93 . Department of Control Engineering, University of Sheffield, Mappin Street, Sheffield

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

The Equivalence of Two
Least Squares Algorithms

by

S. A. Billings, B.Eng., Ph.D., M.Inst.M.C., AMIEE

Department of Control Engineering
University of Sheffield
Mappin Street
Sheffield S1 3JD

Research Report No. 93

June 1979

Abstract

The instrumental variable estimator and the unbiased least squares algorithm developed by James, Souter and Dixon are shown to be asymptotically equivalent.

Introduction

Over the last decade numerous modifications of the conventional least squares algorithm have been developed to eliminate the bias in the parameter estimates which occurs when the system output is corrupted with correlated noise. Two of these algorithms commonly referred to as instrumental variables^{1,2} and suboptimal least squares³ are analysed and shown to be equivalent.

Consider an open-loop discrete time system described by the linear difference equation

$$\begin{aligned}
 y_t &= - \sum_{i=1}^n a_i y_{t-i} + \sum_{i=1}^n b_i u_{t-i} \\
 z_t &= y_t + v_t \quad (t = 1, 2, \dots, k) \\
 v_t + \sum_{i=1}^p c_i v_{t-i} &= \sum_{i=1}^p d_i \xi_{t-i} + \xi_t
 \end{aligned} \tag{1}$$

where u_t and y_t are the input and noise-free plant output at time t , z_t is the measured output corrupted by noise v_t , and ξ_t is a white noise sequence with zero mean.

For a sequence of k data points the system description can be expressed in matrix notation as

$$Z = \phi_{zu} \beta + W \tag{2}$$

where

$$\beta^T = [-a_1 \dots -a_n, b_1 \dots b_n]$$

$$\phi_{zu} = \begin{pmatrix} z_n & \dots & z_1 & u_n & \dots & u_1 \\ \vdots & & \vdots & \vdots & & \vdots \\ z_{k-1} & \dots & z_{k-n} & u_{k-1} & \dots & u_{k-n} \end{pmatrix}$$

$$Z^T = [z_{n+1} \cdot \cdot \cdot z_k]$$

$$W^T = [\omega_{n+1} \cdot \cdot \cdot \omega_k]$$

$$\text{and } \omega_t = v_t + \sum_{i=1}^n a_i v_{t-i}$$

The conventional least squares estimates are obtained by minimising

$$J = (Z - \phi_{zu}\beta)^T (Z - \phi_{zu}\beta) \quad (3)$$

$$\text{to yield } \hat{\beta} = (\phi_{zu}^T \phi_{zu})^{-1} \phi_{zu}^T Z \quad (4)$$

This estimate will be biased unless the elements of W reduce to a white noise sequence. James, Souter and Dixon³ developed a suboptimal least squares algorithm to eliminate this bias by defining a modified least squares objective function

$$J' = (Z - \phi_{zu}\beta)^T (Z - \phi_{zu}\beta) - \beta^T N \beta + 2\beta^T Q \quad (5)$$

where

$$N = \begin{bmatrix} \sum_{j=n}^{k-1} v_j v_j & \sum_{j=n}^{k-1} v_j v_{j-n+1} & \vdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{j=n}^{k-1} v_{j-n+1} v_j & \sum_{j=n}^{k-1} v_{j-n+1} v_{j-n+1} & \vdots & 0 \\ \hline 0 & 0 & \vdots & 0 \end{bmatrix} \quad (6)$$

and

$$Q^T = \begin{bmatrix} \sum_{j=n}^{k-1} v_j v_{j+1} & \cdot & \cdot & \sum_{j=n}^{k-1} v_{j-n+1} v_{j+1} & \vdots & 0 \end{bmatrix} \quad (7)$$

The last two terms in eqn (5) effectively subtract the bias associated with the conventional least squares estimate to yield

$$\hat{\beta}_{so} = (\phi_{zu}^T \phi_{zu} - N)^{-1} (\phi_{zu}^T Z - Q) \quad (8)$$

Provided $(\phi_{zu}^T \phi_{zu} - N)$ is positive definite eqn (8) gives an unbiased estimate of the system parameters.

The algorithm can be implemented³ by initially assuming that R and Q are zero and solving eqn (4) to give the conventional least squares estimate. The predicted output y_t and hence v_t can then be estimated, R and Q can be formed and $\hat{\beta}$ computed from eqn (8). Iterative updating of R, Q and $\hat{\beta}$ is continued until convergence is achieved.

Alternatively, unbiased estimates can be obtained using an instrumental variable estimator. Premultiplying eqn (2) by an instrument matrix X^T gives

$$X^T Z = X^T \phi_{zu} \beta_{IV} + X^T W \quad (9)$$

Providing the instrument matrix X^T is selected to have the following properties

$$\begin{aligned} \text{p.lim}[X^T W] &= 0 \\ \text{p.lim}[X^T \phi_{zu}] &\text{ positive definite} \end{aligned} \quad (10)$$

$$\text{then } \hat{\beta}_{IV} = (X^T \phi_{zu})^{-1} X^T Z \quad (11)$$

is an asymptotically unbiased estimate.

The choice of instruments has been investigated by several authors including Joseph, Lewis, Tou⁴, and Young¹. However, Wong and Polak² showed that optimal instrumental variables exist. They

formed X^T by replacing z_t in ϕ_{zu} by the predicted output y_t which is estimated using an auxiliary model and the parameter estimates of the previous iteration.

Hence selecting the instrument matrix as

$$X^T = \phi_{yu}^T = (\phi_{zu} - \phi_{vo})^T \quad (12)$$

the instrumental variable estimate may be expressed as

$$\hat{\beta}_{IV} = (\phi_{zu}^T \phi_{zu} - \phi_{vo}^T \phi_{zu})^{-1} (\phi_{zu}^T Z - \phi_{vo}^T Z) \quad (13)$$

where

$$\phi_{vo} = \begin{pmatrix} v_n & \dots & v_1 & 0 & \dots & 0 \\ \vdots & & & \vdots & & \vdots \\ \vdots & & & \vdots & & \vdots \\ \vdots & & & \vdots & & \vdots \\ v_{k-1} & \dots & v_{k-n} & 0 & \dots & 0 \end{pmatrix}$$

Consider the asymptotic properties of the instrumental variable estimator for an increasing number of observations $k \rightarrow \infty$. Taking the limit-in-probability and applying Slutsky's theorem⁵ yields

$$\begin{aligned} p.\lim_{k \rightarrow \infty} \hat{\beta}_{IV} &= (p.\lim_{k \rightarrow \infty} \frac{\phi_{zu}^T \phi_{zu}}{k} - p.\lim_{k \rightarrow \infty} \frac{\phi_{vo}^T \phi_{zu}}{k})^{-1} \\ &\quad \cdot (p.\lim_{k \rightarrow \infty} \frac{\phi_{zu}^T Z}{k} - p.\lim_{k \rightarrow \infty} \frac{\phi_{vo}^T Z}{k}) \end{aligned} \quad (14)$$

From eqn (12)

$$p.\lim_{k \rightarrow \infty} \frac{\phi_{vo}^T \phi_{zu}}{k} = p.\lim_{k \rightarrow \infty} \frac{\phi_{vo}^T (\phi_{yu} + \phi_{vo})}{k} = p.\lim_{k \rightarrow \infty} \frac{N}{k} \quad (15)$$

and

$$p.\lim_{k \rightarrow \infty} \frac{\phi_{vo}^T Z}{k} = p.\lim_{k \rightarrow \infty} \frac{\phi_{vo}^T (Y+V)}{k} = p.\lim_{k \rightarrow \infty} \frac{Q}{k} \quad (16)$$

The instrumental variable estimate, eqn (14) thus asymptotically reduces to

$$\begin{aligned} \text{p.lim}_{k \rightarrow \infty} \hat{\beta}_{IV} &= \left(\text{p.lim}_{k \rightarrow \infty} \frac{\phi^T \phi zu zu}{k} - \text{p.lim}_{k \rightarrow \infty} \frac{N}{k} \right)^{-1} \\ &\cdot \left(\text{p.lim}_{k \rightarrow \infty} \frac{\phi^T zu}{k} - \text{p.lim}_{k \rightarrow \infty} \frac{Q}{k} \right) \end{aligned} \quad (17)$$

which is of exactly the same form as the suboptimal least squares estimate when the limit-in-probability is taken. The two algorithms are therefore asymptotically identical and instrumental variables can be interpreted in terms of the modified least squares cost function defined in eqn (5).

The algorithm of James, Souter and Dixon³ appears to be more efficient computationally requiring only $(5n-1)k$ asymptotic multiplications at the second and succeeding iterations compared to $(9n-1)k$ for instrumental variables. The two algorithms have been compared using both simulated and industrial data by Clarke⁶.

References

- 1) Young, P.C.: An instrumental variable method for real-time identification of a noisy process; Automatica, 6, pp.271-287, 1970.
- 2) Wong, K.Y., Polak, E.: Identification of linear discrete time systems using the instrumental variable method; IEEE Trans. Auto. Control, AC-12, pp.707-718, 1967.
- 3) James, F.N., Souter, P., Dixon, D.C.: Suboptimal estimation of the parameters of discrete systems in the presence of correlated noise; Electronics Lett., 8, No.16, pp.411-412, 1972.

- 4) Joseph, P., Lewis, J., Tou, J.: Plant identification in the presence of disturbances and application to digital adaptive systems; AIEE Trans., 80, pp.18-24, 1961.
- 5) Goldberger, A.S.: Econometric Theory; Wiley, 1964.
- 6) Clarke, D.W.: Experimental comparison of identification methods; 5th UKAC Conf., Bath, pp.1-9, 1973.