New formulations for recursive residuals as a diagnostic tool in the classical fixed-effects linear model of arbitrary rank

J.D. Godolphin*

Department of Mathematics, University of Surrey, Guildford, Surrey. GU2 7XH, U.K.

Abstract

The use of residuals for detecting departures from the assumptions of the linear model, whether full rank or not, has long been recognized as an important diagnostic tool, once it became computationally feasible to compute them in a straight forward way, and several contributions to the literature have focussed on their underlying properties and their effectiveness. The recursive residuals are attractive in Econometric applications where there is a natural ordering among the observations through time. New formulations for the recursive residuals are given in terms of the observation vector or the ordinary residuals, which do not require the computation of least squares parameter estimates and for which the transformation matrices are expressed wholly in terms of the rows of the Theil Z matrix. Illustrations of these new formulations are given.

Keywords: Generalised inverse; Least squares; Linear model specification; Matrix factorization; Outlier detection; Regression diagnostics; Residuals; Structural change; Transformation matrices.

**Email address:* j.godolphin@surrey.ac.uk

1 Introduction

The use of diagnostic procedures for detecting departures from model assumptions has long been an important part of model building in Econometrics and other sciences where linear models with fixed parameters are employed. The central role played by residuals in these procedures is well known. Much attention has focussed on the ordinary least squares (OLS) residuals for the routine testing of model fit and model assumptions. However, the recognition that the OLS residuals are not most suitable for many model-testing problems has led to the proposal of several other definitions of residual in the literature. Haslett and Haslett (2007) have clarified residual specification for the linear model by introducing three essentially different types of residual that categorize other definitions, and by establishing the algebraic relationships between these types. The residual specifications of Haslett and Haslett (2007) apply to the general linear model, in which the design matrix need not be full rank and the error covariance matrix need not be diagonal or even full rank so that, for example, models with fixed and random effects apply.

The OLS residuals are marginal type, being defined as the difference between actual and fitted observations. Haslett and Haslett (2007) suggest that residuals of the conditional type, which are expressed as the difference between actual and predicted observations, are important and arguably more fundamental than marginal residuals. A useful member of the conditional type is the class of recursive residuals, see Hedayat and Robson (1970) and Brown *et al.* (1975). Recursive residuals differ from full-conditional residuals, where the predicted observations are based on all other data, in that they are conditional only on a subset of the data which is usually interpreted as “past history”. They are therefore intended primarily for assessing linear models where there is a natural ordering among the n observations, usually measurements made over time although other sequential orderings are possible; see, in particular, Hawkins (1991) for comment on this point. It is argued that this form of conditioning is useful for detecting departures from model assumptions

at various points in the ordered sequence. For instance, Brown *et al.* (1975) reason that recursive residuals derived over time behave exactly as on the null hypothesis until a change in the model takes place when signs of this change should become apparent, unlike other methods where “one would expect the effects of the change to be spread over the full set of transformed residuals”.

Recursive residuals have received much attention in the literature. Hedayat and Robson (1970) considered a test that the disturbance variance depends on the regressors, whilst Brown *et al.* (1975) employed cusum or cusum squared statistics to test whether the parameters of a regression relation change over a period of time, and both tests were based on recursive residuals. Belsley *et al.* (1980) recommended the recursive residuals in preference to OLS residuals to investigate the presence of outliers even if the model does not change over time. Galpin and Hawkins (1984) proposed on the basis of their practical experience that recursive residuals should be calculated routinely for data sets in the full-rank case as a basis for diagnostic checking. Kianifard and Swallow (1996) gave a comprehensive review of recursive residuals and discussed in detail their use for studying influence and leverage. Further properties of recursive residuals are given by Sen (1982), Haslett (1985), Hawkins (1991) and Wright (1999). Loynes (1986) discussed the case where the matrix of explanatory variables is less than full rank; he proposed that the model should be reparameterised (see, for example, Zyskind (1967)) before computing recursive residuals and then the results from the full-rank case will apply. This case was also discussed by Clarke and Godolphin (1992) and Jammalamadaka and Sengupta (1999).

In this paper, attention is confined to the classical case of the linear model with fixed parameters where the model disturbances are mutually uncorrelated and possess a constant variance under the null hypothesis. The principal feature of the recursive residuals is that they are also uncorrelated with the same constant variance as the model disturbances. A complete set of n residuals with this property is not possible since the estimation of the model parameters generates a rank deficiency of some positive integer r so that only

$n - r$ degrees of freedom are associated with the residual set. For definiteness, it is customary for recursive residuals to be described so that it is the final $n - r$ residuals which are uncorrelated with constant variance and are in a one-to-one correspondence with the matching observations. Without a convention of this sort, the recursive residuals are not uniquely defined. It is possible to formulate exact tests based on the recursive residuals for structural and related changes to the latter, i.e. more recent, part of the series, although the absence of r residuals at the beginning of the data set is seen to be a disadvantage by some authors. Other unfavourable features of recursive residuals, by comparison with the well known BLUS residuals of Theil (1965), are discussed by Magnus and Sinha (2005).

In the simple linear regression model with an unknown intercept but without a slope, the observation vector $Y = [y_1 \ y_2 \ \dots \ y_n]'$ consists of n uncorrelated random variables with a common unknown mean and variance; then the matrix of explanatory variables is 1_n , a vector of n elements equal to unity, and the $n - 1$ recursive residuals are specified in terms of Y by the standard Helmert transformation; see Kendall and Stuart (1964, p.250) or Cox (1975). This convenient formulation was exploited by Cox (1975) who obtained general formulae for the asymptotic relative efficiency of the cusum test advocated by Brown *et al.* (1975) in circumstances where there is a single possible change point, in order to confirm useful properties of the test. However, Cox's hypotheses do not seem to have been investigated in detail for cases other than this simplest of the linear models. In more general models, recursive residuals are obtained sequentially by the indirect approach of deriving the least squares estimators of the unknown parameter vector from sub-vectors of Y and then computing the ordinary residuals in each case. Some assistance in this task is provided by the Plackett formula for updating the normal equations (see Brown *et al.*, 1975); and, although Nelder (1975) and Hawkins (1991) raise doubt about the numerical stability of these computations, it seems unlikely that this is a major problem with existing software. A difficulty may be apparent if the matrix of explanatory variables is less than full rank since there seem to be no guidelines in the literature for suggesting a preliminary

reparameterisation of the model before applying this procedure. The fact remains that this indirect approach of describing recursive residuals lacks the pedagogic feature of expressing the linear structure of these residuals in terms of Y for any models, whether full rank or not, which are more general than that considered by Cox (1975).

The purpose of the present paper is to derive two formulations for recursive residuals which are expressed directly in terms of the observation vector Y or, alternatively, which are expressed solely in terms of a component of the OLS residuals. Neither representation requires the computation of any least squares parameter estimates and all coefficients are given in closed form, where no matrix inversion is required in their derivation. These results apply whether the design matrix of explanatory variables has full rank or not, and then r is just the rank of this design matrix. The results therefore possess simple computational properties which parallel the approach of Kontoghiorghes and Foschi (2000) and Kontoghiorghes (2004), who gave methods for deriving a set of (ordinary) residuals for seemingly unrelated regressions which do not require any matrix inversion. Furthermore, the two residual transformations presented here possess the interesting feature of being expressed wholly in terms of the rows of the Z matrix of Theil (1965), which simplifies considerably the recursive residual formulations.

Section 2 of the paper presents and derives two lemmas which provide invariant results for the linear model. Sections 3 and 4 derive the two recursive residual representations and two illustrations of these formulations are given. Section 5 specifies two algorithms for computing the recursive residuals which exhibit straightforward computational properties, and a study of the computational complexity of these methods is presented.

2 Preliminaries

The linear model for the observation vector Y can be written as

$$Y = X\beta + \epsilon, \tag{2.1}$$

where β is a $k \times 1$ vector of unknown parameters, X is $n \times k$ matrix of explanatory variables with $\text{rank}(X) = r \leq k$ and ϵ is, under the null hypothesis, a vector of uncorrelated random variables with mean zero and variance σ^2 . It is well known that the ordinary least squares (OLS) residual vector $\hat{\epsilon}$ is given by

$$\hat{\epsilon} = (I - H)Y = (I - H)\epsilon, \quad (2.2)$$

where the hat matrix $H = X(X'X)^-X'$ is the projection operator for $\mathcal{C}(X)$, the column space of X , which is specified uniquely for any generalized inverse $(X'X)^-$ of $X'X$. Also

$$E[\hat{\epsilon}] = 0 \quad \text{and} \quad \text{Var}[\hat{\epsilon}] = \sigma^2(I - H),$$

i.e. the OLS residual has expectation zero, but the individual elements of $\hat{\epsilon}$ are dependent and have variable variances which depend on X .

Godolphin (2006) suggested that many properties of the model (2.1) can be represented in full by the $(n - r) \times r$ matrix

$$Z = X_*X'_0(X_0X'_0)^{-1}, \quad (2.3)$$

where X_0 and X_* are $r \times k$ and $(n - r) \times k$ components of X , i.e.

$$X = \begin{bmatrix} X_0 \\ X_* \end{bmatrix} = \begin{bmatrix} I \\ Z \end{bmatrix} X_0, \quad (2.4)$$

here assuming that X_0 has full rank r . The partition (2.4) was formulated originally, in the full rank case, by Theil (1965) but Z did not feature explicitly in seminal works of Hedayat and Robson (1970) or Brown *et al.* (1975). Godolphin (2006) showed that all rank reducing observation sets of Types I-III can be determined from an examination of the matrix (2.3). Therefore it is interesting to ask if a useful representation for the recursive residuals can be found from Z . In particular it should be asked if the decomposition (2.4) is likely to be at least as informative as other factorizations of X , such as the singular value decomposition, which have been advocated by many workers in this field.

The normal equations are given by

$$X'X\hat{\beta} = X'Y \quad (2.5)$$

which are consistent but singular if $r \neq k$, implying that there are many possible solutions for $\hat{\beta}$. Rao (1962) has shown that the general solution of equation (2.5) is given by

$$\hat{\beta} = GX'Y + (I - GX'X)\xi, \quad (2.6)$$

where ξ is an arbitrary $k \times 1$ vector and G is a generalized inverse for $X'X$. See the discussion of Harville (1997, §12.2). The definition for G to be a generalized inverse for $X'X$ is that G satisfies the equation $X'XGX'X = X'X$. From the partition (2.4) of X this becomes

$$X'_0(I + Z'Z)X_0GX'_0(I + Z'Z)X_0 = X'_0(I + Z'Z)X_0.$$

Pre-multiplying both sides by $(I + Z'Z)^{-1}(X_0X'_0)^{-1}X_0$ and post-multiplying both sides by $X'_0(X_0X'_0)^{-1}(I + Z'Z)^{-1}$ we obtain

$$X_0GX'_0 = (I + Z'Z)^{-1}.$$

Clearly this argument may be reversed. Thus we may appeal to the result of Rao (1962, Lemma 1) to state three equivalent conditions for G to be a generalized inverse for $X'X$.

LEMMA 1. *If the design matrix is partitioned by (2.4) and if $Z = X_*X'_0(X_0X'_0)^{-1}$ then the following three conditions are equivalent for G to be a generalized inverse for $X'X$:*

- (i) *For any q for which $X'X\beta = q$ is consistent then $\beta = Gq$ is a solution;*
- (ii) $X'XGX'X = X'X$;
- (iii) $X_0GX'_0 = (I + Z'Z)^{-1}$.

Part (iii) of Lemma 1 appears to be a new result which may be useful if the partition (2.4) is known. It can be shown from Lemma 1 that the vector $X_0\hat{\beta}$ is invariant with respect to the choice of generalized inverse for $X'X$ and it is possible to obtain a useful expression for it which is represented solely in terms of the matrix Z . We have

LEMMA 2. The value of $X_0\widehat{\beta}$ is independent of the particular choice for G , which was used to derive $\widehat{\beta}$, and is given by

$$X_0\widehat{\beta} = (I + Z'Z)^{-1} \begin{bmatrix} I & Z' \end{bmatrix} Y. \quad (2.7)$$

Proof: From Lemma 1 (iii) and the general solution (2.6) it follows that any estimator given by $X_0\widehat{\beta}$ is necessarily of the form

$$\begin{aligned} X_0\widehat{\beta} &= X_0 (GX'Y + (I - GX'X) \xi) \\ &= X_0 GX'_0 \begin{bmatrix} I & Z' \end{bmatrix} Y + (X_0 - X_0 GX'_0 (I + Z'Z) X_0) \xi \\ &= (I + Z'Z)^{-1} \begin{bmatrix} I & Z' \end{bmatrix} Y. \end{aligned}$$

For given X_0 , Lemma 2 demonstrates that $X_0\widehat{\beta}$ is invariant for any solution, $\widehat{\beta}$, of the normal equations (2.5). The value of $X_0\widehat{\beta}$ has the linear form AY , as expected, however it is interesting that the coefficient matrix A is expressed entirely in terms of Z , defined by (2.3), which is consistent with results of Godolphin (2006) referred to previously.

Rao (1973, p. 223) has shown that the parametric combination $\lambda'\beta$ is estimable if and only if $\lambda \in \mathcal{C}(X')$, the column space of X' . It is clear from (2.4) that $\mathcal{C}(X'_0) = \mathcal{C}(X')$, i.e. $\lambda'\beta$ is estimable if and only if $\lambda' = c'X_0$ for some $r \times 1$ vector c . In particular, the parametric vector $X_0\beta$ is estimable and it has unbiased estimator given by (2.7).

3 Uncorrelated Recursive Residual Specification

Denote the first r observations by the vector $y_0 = [y_1 \ y_2 \ \cdots \ y_r]'$, which is referred to as the *core observation set*. Also note explicitly the $n - r$ columns of the matrix Z' by writing $Z' = [\zeta_1 \ \zeta_2 \ \cdots \ \zeta_{n-r}]$. Let $K_0 = I$ and define the $r \times r$ matrix K_{i-1} for $i \geq 2$ by

$$K_{i-1} = \left(I + \sum_{j=1}^{i-1} \zeta_j \zeta_j' \right)^{-1}, \quad (3.1)$$

so that, in particular, we have $K_{n-r} = (I + Z'Z)^{-1}$.

We proceed initially in the same way as Brown *et al* (1975) for the full-rank model. Define formally, for each $i = 1, \dots, n - r$, the statistic u_i which is the OLS residual, i.e. the least-squares estimator of ϵ_{r+i} , based on $r + i - 1$ observations only, given by

$$u_i = y_{r+i} - (i\text{th row of } X_*)\widehat{\beta}_{r+i-1}. \quad (3.2)$$

Here the subscript j on $\widehat{\beta}_j$ means that only the first j observations are used for calculating a least-squares estimator for β . It follows from (2.4) that the i th row of X_* is given by $\zeta'_i X_0$. Furthermore, it follows from the same argument as used in the proof of Lemma 2 that $X_0\widehat{\beta}_{r+i-1}$ is specified uniquely and independently of the choice of generalized inverse used to yield $\widehat{\beta}_{r+i-1}$. Therefore, whether $r = k$ or $r < k$, the expression (3.2) becomes

$$u_1 = y_{r+1} - \zeta'_1 K_0 y_0 = y_{r+1} - \zeta'_1 y_0 \quad (3.3)$$

$$u_i = y_{r+i} - \zeta'_i K_{i-1} \left(y_0 + \sum_{j=1}^{i-1} y_{r+j} \zeta_j \right) \quad (i = 2, \dots, n - r). \quad (3.4)$$

From this expression for u_i it is possible to show directly that $u_1, u_2, \dots, u_{n-r}$ comprise a set of uncorrelated random variables with mean zero. We require the positive quantities

$$v_i^2 = 1 + \zeta'_i K_{i-1} \zeta_i \quad (i = 1, 2, \dots, n - r). \quad (3.5)$$

THEOREM 1. *The random variables $u_1, u_2, \dots, u_{n-r}$ are uncorrelated with expectation $E[u_i] = 0$ and variance $\text{Var}[u_i] = (1 + \zeta'_i K_{i-1} \zeta_i) \sigma^2 = v_i^2 \sigma^2$.*

Proof: The $(r + i)$ th observation y_{r+i} can be written as ϵ_{r+i} plus the i th row of X_* multiplied by β . This becomes $y_{r+i} = \zeta'_i X_0 \beta + \epsilon_{r+i}$. Similar expressions can be found for the other observations; furthermore, we have

$$y_0 + \sum_{j=1}^{i-1} y_{r+j} \zeta_j = [I \ \zeta_1 \ \zeta_2 \ \cdots \ \zeta_{i-1}] \left\{ \begin{bmatrix} X_0 \\ \zeta'_1 X_0 \\ \vdots \\ \zeta'_{i-1} X_0 \end{bmatrix} \beta + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_{r+i-1} \end{bmatrix} \right\},$$

therefore if $i \geq 2$ it follows from (3.4) and the representation (3.1) that the random variable u_i can be written as

$$\begin{aligned}
u_i &= \zeta_i' X_0 \beta + \epsilon_{r+i} - \zeta_i' K_{i-1} [I \ \zeta_1 \ \zeta_2 \ \cdots \ \zeta_{i-1}] \left\{ \begin{bmatrix} I \\ \zeta_1' \\ \vdots \\ \zeta_{i-1}' \end{bmatrix} X_0 \beta + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_{r+i-1} \end{bmatrix} \right\} \\
&= \epsilon_{r+i} - \zeta_i' K_{i-1} [I \ \zeta_1 \ \zeta_2 \ \cdots \ \zeta_{i-1}] \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_{r+i-1} \end{bmatrix}.
\end{aligned}$$

Therefore $E[u_i] = 0$ and $\text{Var}[u_i] = (1 + \zeta_i' K_{i-1} \zeta_i) \sigma^2 = \sigma^2 v_i^2$. Furthermore, if $j > i$ then the covariance $\text{Cov}(u_i, u_j) = E[u_i u_j]$ is given by

$$\begin{aligned}
\text{Cov}(u_i, u_j) &= E \left[-\zeta_j' K_{j-1} [I \ \zeta_1 \ \zeta_2 \ \cdots \ \zeta_{j-1}] \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_{r+j-1} \end{bmatrix} \epsilon_{r+i} \right] + \\
&E \left[\zeta_i' K_{i-1} [I \ \zeta_1 \ \zeta_2 \ \cdots \ \zeta_{i-1}] \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_{r+i-1} \end{bmatrix} [\epsilon_1 \ \epsilon_2 \ \cdots \ \epsilon_{r+j-1}] \begin{bmatrix} I \\ \zeta_1' \\ \vdots \\ \zeta_{j-1}' \end{bmatrix} K_{j-1} \zeta_j \right],
\end{aligned}$$

consequently $\text{Cov}(u_i, u_j)$ is given by

$$\text{Cov}(u_i, u_j) = -\zeta_j' K_{j-1} \zeta_i \sigma^2 + \zeta_i' K_{i-1} K_{i-1}^{-1} K_{j-1} \zeta_j \sigma^2 = 0.$$

This proves the theorem. Note that it is a trivial matter to define uncorrelated zero-mean random variables with constant variance σ^2 by $u_i^R = u_i / v_i$ where $v_i = (1 + \zeta_i' K_{i-1} \zeta_i)^{\frac{1}{2}}$ from (3.5); then a result corresponding to Theorem 1 holds also for the random variables $\{u_i^R\}$.

The proof of Theorem 1 also makes it possible to specify the recursive residuals as linear combinations of the elements of the observation vector Y . It is evident from equation (3.3) and (3.4) that the weights are expressed entirely in terms of the $n - r$ columns of Z' . The formulation (3.4) seems to require the inversion of $n - r$ different $r \times r$ matrices to give $K_1, K_2, \dots, K_{n-r}$; however, no inversion is needed since $K_0 = I$ and K_i can be obtained as an additive adjustment to K_{i-1} for each $i = 1, 2, \dots, n - r$, as the next result shows.

LEMMA 3. *Let the integer i satisfy $1 \leq i \leq n - r$ and let the matrix K_{i-1} , defined by (3.1), be given. Then K_i can be obtained by the following three steps:*

- (i) *Form the vector $h_i = K_{i-1}\zeta_i$.*
- (ii) *Form the ‘residual variance’ $v_i^2 = 1 + h_i'\zeta_i$.*
- (iii) *Form $K_i = K_{i-1} - \frac{1}{v_i^2} h_i h_i'$.*

The verification of Lemma 3 is straightforward, it being necessary only to post-multiply the right side of (iii) by $K_{i-1}^{-1} + \zeta_i \zeta_i'$ to obtain the $r \times r$ identity matrix.

Let the vectors of random variables u_i and u_i^R be denoted u and u^R respectively, i.e. write $u = [u_1 \ u_2 \ \dots \ u_{n-r}]'$ and $u^R = [u_1^R \ u_2^R \ \dots \ u_{n-r}^R]'$. Then we have established the following result.

THEOREM 2. *The uncorrelated random variables u_i are defined in terms of Y by $u = CY$, with C partitioned as $C = [C_0 \ C_*]$ such that the $(n-r) \times r$ matrix C_0 and the $(n-r) \times (n-r)$ lower triangular matrix C_* are given by*

$$C_0 = \begin{bmatrix} -h'_1 \\ -h'_2 \\ \vdots \\ -h'_{n-r} \end{bmatrix} \quad \text{and} \quad C_* = \begin{bmatrix} 1 & & & 0 \\ -h'_2 \zeta_1 & 1 & & \\ \vdots & \vdots & \ddots & \\ -h'_{n-r} \zeta_1 & -h'_{n-r} \zeta_2 & \dots & 1 \end{bmatrix}, \quad (3.6)$$

where ζ_i is the i th column of Z' , h_i is given by Lemma 3 and K_{i-1} is defined by (3.1).

The recursive residuals u_i^R are defined in terms of the observation vector Y by $u^R = C_R Y$, where $C_R = \text{Diag}[v_1^{-1} \ v_2^{-1} \ \dots \ v_{n-r}^{-1}] C$ and C is given by (3.6).

Illustrations

1. Consider the simplest case where $k = r = 1$. Then $Z = 1_{n-1}$ so that all $\zeta_i = 1$ whilst $K_i = 1/(i+1)$. Hence (3.3) gives $u_1 = y_2 - y_1$ and (3.4) gives

$$u_i = y_{i+1} - \frac{y_1 + y_2 + \dots + y_i}{i} = y_{i+1} - \bar{y}_{(i)}, \quad \text{hence } u_i^R = \sqrt{\frac{i}{i+1}} u_i, \quad (i = 1, \dots, n-2),$$

in this case, yielding the recursive residuals as the Helmert transformation of Y ; confer Kendall and Stuart (1963, §11.3) or Cox (1975).

2. An interesting second illustration is given by the simple regression model where $y_i = \alpha + \beta x_i + \epsilon_i$ for $i = 1, 2, \dots, n$. Thus $k = r = 2$ and

$$X_0 = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \end{bmatrix} \quad \text{and} \quad X_* = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_3 & x_4 & \dots & x_n \end{bmatrix}',$$

assuming $x_1 \neq x_2$, so that the transpose of $Z = X_* X_0' (X_0 X_0')^{-1}$ is given by

$$Z' = \frac{1}{x_2 - x_1} \begin{bmatrix} x_2 - x_3 & x_2 - x_4 & \dots & x_2 - x_n \\ x_3 - x_1 & x_4 - x_1 & \dots & x_n - x_1 \end{bmatrix} = \begin{bmatrix} \zeta_1 & \zeta_2 & \dots & \zeta_{n-2} \end{bmatrix}.$$

In this case K_{m-2} is given for any $m \geq 3$ from (3.1) as

$$K_{m-2} = \frac{1}{m s_m^2} \begin{bmatrix} d_{1m} & d_{3m} \\ d_{3m} & d_{2m} \end{bmatrix},$$

where $s_m^2 = \sum_{j=1}^m (x_j - \bar{x}_{(m)})^2$ is the residual sum of squares based on the first m observations and the individual terms d_{1m} , d_{2m} and d_{3m} are found to be

$$d_{im} = \sum_{j=1}^m (x_i - x_j)^2, \quad (i = 1, 2), \quad \text{and} \quad d_{3m} = \sum_{j=1}^m (x_1 - x_j)(x_2 - x_j).$$

Therefore the transformation $u = CY$ is such that the components of the pseudo lower triangular matrix $C = ((\pi_{k\ell}))$ are given by (3.6) as

$$C_0 = \begin{bmatrix} -\pi_{11} & -\pi_{12} \\ -\pi_{21} & -\pi_{22} \\ -\pi_{31} & -\pi_{32} \\ \vdots & \\ -\pi_{n-2,1} & -\pi_{n-2,2} \end{bmatrix} \quad \text{and} \quad C_* = \begin{bmatrix} 1 & & 0 \\ -\pi_{23} & 1 & \\ -\pi_{33} & -\pi_{34} & \\ \vdots & \vdots & \ddots \\ -\pi_{n-2,3} & -\pi_{n-2,4} & \dots & 1 \end{bmatrix}, \quad (3.7)$$

where the coefficients $\pi_{k\ell}$ ($1 \leq k \leq (n-2)$, $1 \leq \ell \leq (k+1)$) are given by

$$\pi_{k\ell} = \frac{1}{(k+1)s_{k+1}^2} \sum_{\substack{i=1 \\ i \neq \ell}}^{k+1} (x_i - x_\ell)(x_i - x_{k+1}).$$

The vector of recursive residuals u^R is obtained from matrices (3.7) as $u^R = C_R Y$, where $C_R = \text{Diag}[v_1^{-1} \ v_2^{-1} \ \dots \ v_{n-2}^{-1}] C$ and $v_i^2 = 1 + \zeta_i' K_{i-1} \zeta_i$ is given by the weighted ratio of component sums of squares, i.e.

$$v_i^2 = \frac{(i+2)s_{i+2}^2}{(i+1)s_{i+1}^2}, \quad (i = 1, 2, \dots, n-2).$$

The corresponding recursive residuals

$$u_i^R = \frac{1}{v_i} \left(y_{i+2} - \sum_{k=1}^{i+1} \pi_{ki} y_k \right) \quad (i = 1, 2, \dots, n-2), \quad (3.8)$$

are expressed in convenient form for the investigation of hypotheses about the simple regression model, where there is a natural ordering of the data. This follows because such an alternative hypothesis would typically be stipulated in terms of a suspected departure from an assumption concerning y_i for all i greater than some change point m (say), so long as m is appreciably bigger than 2. Examples of such suspected departure from assumptions include a change in one or both of the intercept α and slope β , or a change in variance of y_i for all $i = m+1, \dots, n$. Using the representation (3.8) the behaviour of the u_i are determined directly from the behaviour of the y_i under a suspected departure from assumption and this should enable properties of tests based on the recursive residuals to be evaluated.

4 Relation of the Recursive Residual to OLS Residual

In many practical situations the ordinary least-squares residual (2.2) will be available to the practitioner on a routine basis and it is interesting to ask how the recursive residual u^R is related to the OLS residual $\hat{\epsilon}$. It follows from (3.6) that C_* post-multiplied by Z is

$$\begin{aligned} C_* Z &= \begin{bmatrix} 1 & & & 0 \\ -\zeta'_2 K_1 \zeta_1 & 1 & & \\ \vdots & \vdots & \ddots & \\ -\zeta'_{n-r} K_{n-r-1} \zeta_1 & -\zeta'_{n-r} K_{n-r-1} \zeta_2 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \zeta'_1 \\ \zeta'_2 \\ \vdots \\ \zeta'_{n-r} \end{bmatrix} \\ &= \begin{bmatrix} \zeta'_1 \\ \zeta'_2 (I - K_1 \zeta_1 \zeta'_1) \\ \vdots \\ \zeta'_{n-r} \{I - K_{n-r-1} (\zeta_1 \zeta'_1 + \cdots + \zeta_{n-r-1} \zeta'_{n-r-1})\} \end{bmatrix} = \begin{bmatrix} \zeta'_1 \\ \zeta'_2 K_1 \\ \vdots \\ \zeta'_{n-r} K_{n-r-1} \end{bmatrix}, \end{aligned}$$

where we have used the result

$$K_i(I + \zeta_1 \zeta'_1 + \cdots + \zeta_i \zeta'_i) = I \implies K_i = I - K_i(\zeta_1 \zeta'_1 + \cdots + \zeta_i \zeta'_i).$$

Therefore $C_* Z = -C_0$. Consequently $C = [-C_* Z \ C_*] = C_*[-Z \ I] = C_* J$ say, where $J = [-Z \ I]$. Further, it is evident from (2.4) that $JX = 0$ and that J has full row rank $n - r$, where r is the dimension of $\mathcal{C}(X)$. Therefore the columns of J' are orthogonal to $\mathcal{C}(X)$, the column space of X , and they form a basis for the complementary space $\mathcal{C}(X)^\perp$. Thus $\mathcal{C}(X)^\perp = \mathcal{C}(J')$. Therefore the complementary space $\mathcal{C}(X)^\perp$ has the property that the first r components of any member of this space is obtained by pre-multiplying its final $n - r$ components by $-Z'$. Since $\hat{\epsilon} = (I - H)Y \in \mathcal{C}(X)^\perp$ then it follows in particular that

$$JY = J(I - H)Y = JJ'(JJ')^{-1}JY = JJ'\hat{\epsilon}_* \quad (4.1)$$

where $\hat{\epsilon}_*$ denotes the final $n - r$ components of $\hat{\epsilon}$.

From (4.1) we conclude that $u = CY = C_*JY = C_*JJ'\hat{\epsilon}_*$. For simplicity, write B for C_*JJ' . Then from formulation (3.6) this transformation matrix turns out to be

$$\begin{aligned}
B &= C_*JJ' = C_* + C_*ZZ' = C_* - C_0[\zeta_1 \ \zeta_2 \ \cdots \ \zeta_{n-r}] \\
&= C_* + \begin{bmatrix} \zeta'_1\zeta_1 & \zeta'_1\zeta_2 & \cdots & \zeta'_1\zeta_{n-r} \\ \zeta'_2K_1\zeta_1 & \zeta'_2K_1\zeta_2 & \cdots & \zeta'_2K_1\zeta_{n-r} \\ \vdots & \vdots & & \vdots \\ \zeta'_{n-r}K_{n-r-1}\zeta_1 & \zeta'_{n-r}K_{n-r-1}\zeta_2 & \cdots & \zeta'_{n-r}K_{n-r-1}\zeta_{n-r} \end{bmatrix}, \\
\text{hence } B &= \begin{bmatrix} v_1^2 & \zeta'_1\zeta_2 & \cdots & \zeta'_1\zeta_{n-r} \\ & v_2^2 & \cdots & \zeta'_2K_1\zeta_{n-r} \\ & & \ddots & \vdots \\ 0 & & & v_{n-r}^2 \end{bmatrix}, \tag{4.2}
\end{aligned}$$

and the following theorem is derived.

THEOREM 3. *The uncorrelated random variables u_i are defined in terms of $\hat{\epsilon}_*$, the final $n - r$ components of the OLS residual $\hat{\epsilon}$, by $u = B\hat{\epsilon}_*$, with B defined by equation (4.2).*

The recursive residuals u_i^R are defined in terms of $\hat{\epsilon}_$ by $u^R = B_R\hat{\epsilon}_*$, where $B_R = \text{Diag}[v_1^{-1} \ v_2^{-1} \ \cdots \ v_{n-r}^{-1}]B$ and B is defined by (4.2).*

Remark Theorems 2 and 3 enable the following matrix identity to be established.

$$[0 \ B](I - H) = C = [C_0 \ C_*], \tag{4.3}$$

where $H = X(X'X)^{-1}X'$ is the hat matrix, C_0 and C_* are defined by (3.7), where C_* is lower triangular, and B is defined by (4.2), where B is upper triangular. All elements of H , C_0 , C_* and B are defined wholly in terms of columns of Z' .

Illustration

1. Consider again the simplest case where $k = r = 1$. The final $n - 1$ components of the OLS residual are given by the mean squared deviations

$$\hat{\epsilon}_* = [y_2 - \bar{y}, y_3 - \bar{y}, \dots, y_n - \bar{y}]', \tag{4.4}$$

and it is required to find the transformation which maps the $n - 1$ statistics (4.4) to the Helmert statistics $u_i = y_{i+1} - (y_1 + y_2 + \dots + y_i)/i = y_{i+1} - \bar{y}_{(i)}$ obtained previously. From (4.2), the required transformation matrix B to yield the uncorrelated statistics u_i is given by

$$B = I + \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & \frac{1}{2} & \frac{1}{2} & \cdots & \frac{1}{2} \\ 0 & 0 & \frac{1}{3} & \cdots & \frac{1}{3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{1}{n-r} \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 & \cdots & 1 \\ 0 & \frac{3}{2} & \frac{1}{2} & \cdots & \frac{1}{2} \\ 0 & 0 & \frac{4}{3} & \cdots & \frac{1}{3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{n-r+1}{n-r} \end{bmatrix}.$$

The upper triangular transformation matrix multiplying $\hat{\epsilon}_*$ which is required to yield the Helmert statistics as recursive residuals u_i^R is B_R , which is given by

$$B_R = \text{Diag} \left[\frac{1}{\sqrt{2}} \quad \frac{\sqrt{2}}{\sqrt{3}} \quad \frac{\sqrt{3}}{\sqrt{4}} \quad \cdots \quad \frac{\sqrt{n-r}}{\sqrt{n-r+1}} \right] B = \begin{bmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \cdots & \frac{1}{\sqrt{2}} \\ 0 & \frac{\sqrt{3}}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \cdots & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{\sqrt{4}}{\sqrt{3}} & \cdots & \frac{1}{\sqrt{12}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{\sqrt{n-r+1}}{\sqrt{n-r}} \end{bmatrix}.$$

Furthermore, the familiar form for the Helmert transformation can be obtained directly from the matrix B , using the identity (4.3). When $k = r = 1$ the projection matrix $I - H$ has the simple form

$$I - H = I - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n',$$

therefore

$$C = [0 \ B](I - H) = \begin{bmatrix} -1 & 1 & 0 & 0 & \cdots & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 1 & 0 & \cdots & 0 \\ -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{n-r} & -\frac{1}{n-r} & -\frac{1}{n-r} & -\frac{1}{n-r} & \cdots & 1 \end{bmatrix},$$

as required.

5 Formulation of the Recursive Residual

5.1 Specification of Recursive Residual Algorithms

Theorems 2 and 3 provide elegant formulations for the recursive residual in the classical linear model of arbitrary rank. The transformation matrices C and B in the expressions $u = CY$ and $u = B\hat{\epsilon}_*$ given by (3.6) and (4.2), are expressed wholly in terms of the columns $\zeta_1, \zeta_2, \dots, \zeta_{n-r}$ of the matrix Z' , demonstrating that the task of formulating recursive least squares estimates of β is unnecessary. At first sight it may appear from equation (3.1) that there is an implied requirement to evaluate the $r \times r$ inverse matrices $K_1, K_2, \dots, K_{n-r-1}$; but the repeated use of Lemma 3 ensures that the K_i are evaluated without the need for recursive inversion of the matrices $(I + \sum_{j=1}^{i-1} \zeta_j \zeta_j')$. However, n can be very large in some econometric applications and, in such circumstances, it may be preferable to obtain the u^R vector by a recursive procedure rather than by the direct formulations (3.6) and (4.2). Two algorithms which employ such recursive procedures are given in what follows. Algorithm 1 applies when the basic information available is the observation set $y_1, y_2, \dots, y_n$, whereas Algorithm 2 can be used when the OLS residuals are also available and it is desired to convert the sub-vector $\hat{\epsilon}_*$ to the corresponding recursive residual vector.

ALGORITHM 1. *Given the $n - r$ columns $\zeta_1, \zeta_2, \dots, \zeta_{n-r}$; set $h_1 = \zeta_1$, $S_0 = y_0$ and $K_0 = I$, the $r \times r$ identity matrix.*

- (1) *Obtain $v_1^2 = 1 + h_1' \zeta_1$ and form the recursive residual $u_1^R = \frac{1}{v_1} (y_{r+1} - h_1' S_0)$. Form and store the $r \times 1$ vector $S_1 = S_0 + y_{r+1} \zeta_1$ and the $r \times r$ matrix $K_1 = I - h_1 h_1' / v_1^2$;*
 - (2) *Calculate $h_2 = K_1 \zeta_2$, obtain $v_2^2 = 1 + h_2' \zeta_2$ and form $u_2^R = \frac{1}{v_2} (y_{r+2} - h_2' S_1)$. Form and store the vector $S_2 = S_1 + y_{r+2} \zeta_2$ and the matrix $K_2 = K_1 - h_2 h_2' / v_2^2$;*
 - (3) *Calculate $h_3 = K_2 \zeta_3$, obtain $v_3^2 = 1 + h_3' \zeta_3$ and form $u_3^R = \frac{1}{v_3} (y_{r+3} - h_3' S_2)$. Form and store the vector $S_3 = S_2 + y_{r+3} \zeta_3$ and the matrix $K_3 = K_2 - h_3 h_3' / v_3^2$.*
- Continue in this fashion to yield the $n - r$ recursive residuals $u_1^R, u_2^R, \dots, u_{n-r}^R$.*

The other procedure is based on the premise that recursive residuals are to be computed from the elements of $\widehat{\epsilon}_*$, the final $n - r$ components of the OLS residual vector. Let

$$\widehat{\epsilon}_* = [\widehat{\epsilon}_{r+1} \ \widehat{\epsilon}_{r+2} \ \dots \ \widehat{\epsilon}_n]',$$

and define the $r \times 1$ vectors

$$T_i = \sum_{j=i}^{n-r} \widehat{\epsilon}_{r+j} \zeta_j \quad \text{for } i = 2, 3, \dots, (n - r).$$

ALGORITHM 2. *Given the $n - r$ columns $\zeta_1, \zeta_2, \dots, \zeta_{n-r}$; obtain $T_{n-r-1}, T_{n-r-2}, \dots, T_2$ iteratively, starting with $T_{n-r} = \widehat{\epsilon}_n \zeta_{n-r}$ and using $T_i = T_{i+1} + \widehat{\epsilon}_{r+i} \zeta_i$; set the vector $h_1 = \zeta_1$ and the matrix $K_0 = I$.*

(1) *Obtain $v_1^2 = 1 + h_1' \zeta_1$ and form the recursive residual $u_1^R = v_1 \widehat{\epsilon}_{r+1} + \frac{1}{v_1} h_1' T_2$. Form and store the $r \times r$ matrix $K_1 = I - h_1 h_1' / v_1^2$;*

(2) *Calculate the vector $h_2 = K_1 \zeta_2$, obtain $v_2^2 = 1 + h_2' \zeta_2$ and form $u_2^R = v_2 \widehat{\epsilon}_{r+2} + \frac{1}{v_2} h_2' T_3$. Form and store the $r \times r$ matrix $K_2 = I - h_2 h_2' / v_2^2$;*

(3) *Calculate the vector $h_3 = K_2 \zeta_3$, obtain $v_3^2 = 1 + h_3' \zeta_3$ and form $u_3^R = v_3 \widehat{\epsilon}_{r+3} + \frac{1}{v_3} h_3' T_4$. Form and store the $r \times r$ matrix $K_3 = I - h_3 h_3' / v_3^2$;*

Continue in this fashion to yield the $n - r - 1$ recursive residuals $u_1^R, u_2^R, \dots, u_{n-r-1}^R$ and then set $u_{n-r}^R = \widehat{\epsilon}_n$.

5.2 Computational Complexity

The speed with which calculations are performed by a numerical algorithm is usually related directly to the *computational complexity* of the algorithm, which is measured by the number of MDS (multiplication, divisional and square root) operations required; see for example (Heath, 2002). Both procedures described by Algorithms 1 and 2 involve $n - r$ iterations so their computational complexities can be assessed from the number of MDS operations required for each iteration. Fig. 1 specifies these counts for the two procedures and it turns out that the two computational complexity measures are virtually the same.

Therefore the methods for formulating the recursive residuals $u_1^R, u_2^R, \dots, u_{n-r}^R$ are similar in performance, whether the original observations $y_1, y_2, \dots, y_n$ or the final $n - r$ ordinary least squares residuals $\hat{\epsilon}_{r+1}, \hat{\epsilon}_{r+2}, \dots, \hat{\epsilon}_n$ are used.

Algorithm	Number of MDS operations for each iteration
Godolphin Algorithm 1	$\frac{3}{2}r^2 + \frac{9}{2}r + 2$
Godolphin Algorithm 2	$\frac{3}{2}r^2 + \frac{9}{2}r + 3$
Brown <i>et al.</i> (full rank case)	$\frac{3}{2}k^2 + \frac{11}{2}k + 3$

Fig. 1: MDS operations for three recursive residual algorithms

It is interesting to compare the procedures described by Algorithms 1 and 2 with the method for calculating recursive residuals which is in current use. This is achieved in part by assessing the computational complexity of the algorithm of Brown *et al.* (1975) using the recursive approach involving updating formulae as set out in Kianifard and Swallow (1996), even though the procedure only applies when X has full-rank so that $r = k$. The number of MDS operations at each iteration for the method of Brown *et al.* (1975) is given in Fig. 1. These calculations show that there is little to discriminate in performance between any of the algorithms in the full-rank case.

6 Concluding Remarks

Given the null hypothesis that the model disturbances are mutually uncorrelated with common variance σ^2 , Theorem 2 provides the transformation of Y which yields the random vector u^R whose elements have the same properties. In the notation of Haslett and Haslett (2007), u^R is a particular kind of conditional residual vector, where conditioning is on

“past history”, consequently these residuals are suited to models where there is a natural ordering such as time or the related situations considered by Hawkins (1991). Also, u^R can consist of only $n - r$ elements so it is usual for the core observation set to be the first r elements of Y . This is equivalent to choosing X_0 to be the first r rows of X and then the elements of u^R are in one-to-one correspondence with the final $n - r$ elements of Y , unless a minor rearrangement is required to ensure that X_0 has full rank. Theorem 3 gives a complementary expression for u^R in terms of the final $n - r$ elements of the OLS residual vector $\hat{\epsilon}$. Both of these formulations for u^R can be derived recursively using Algorithms 1 and 2. All of these specifications are couched in terms of the $n - r$ columns of Z' and they therefore have favourable computational implications, since no partial or recursive estimate of β is required and they do not depend on matrix inversion. It also seems likely that the representations suggested by Theorems 2 and 3 are conceptually simpler than existing recursive residual specifications so that methodological properties of the recursive residuals may be obtained more easily. Furthermore, these results apply to other residuals, such as those proposed by McGilchrist *et al* (1989) or Luger (2001), which differ from recursive residuals only in terms of their sign.

The derivation of the $n - r$ columns of Z' is an essential preliminary step for the formulation of recursive residuals and diagnostic procedures based upon them. This is, however, achieved in a straightforward way; see Godolphin (2006). Indeed, it is interesting to note that the Theil factorization (2.4) of the design matrix, on which these results are based, does not depend on the computation of eigenvalues or eigenvectors of $X'X$ as is required, for example, by the singular value decomposition; see Belsley *et al.* (1980, §3.2). It is suggested that the Z matrix be computed as part of the standard output since it may have additional value for examining the suitability of the model, in view of the definitive properties obtained for Z , both in the present paper and in the work of Godolphin (2006) and elsewhere, and it is reasonable to suppose that other useful properties of the Theil factorization may emerge from further consideration.

Acknowledgements

Grateful thanks are due to a referee for informed and helpful comments on an earlier draft of this paper.

References

- [1] Belsley, D.A., Kuh, E. and Welsch, R.E. (1980) *Regression Diagnostics* John Wiley, New York.
- [2] Brown, R.L., Durbin, J. and Evans, J.M. (1975). Techniques for testing the constancy of regression relationships over time. *J. R. Statist. Soc. B* 37, 149-192.
- [3] Clarke, B.R. and Godolphin, E.J. (1992) Uncorrelated residuals and an exact test for two variance components in experimental design. *Commun. Statist. Theory Meth.* 21, 2501-2526.
- [4] Cox, D.R. (1975) Contribution to Brown, R.L. et al. (1975) *J. Roy. Statist. Soc. B* 37, 163-164.
- [5] Galpin, J.S. and Hawkins, D.M. (1984) The use of recursive residuals in checking model fit in linear regression. *American Statistician* 38, 94-105.
- [6] Godolphin, J.D. (2006). The specification of rank reducing observation sets in experimental design. *Comp. Statist. Data Anal.* 51, 1862-1874.
- [7] Harville, D.A. (1997). *Matrix Algebra from a Statistician's Perspective* Springer, New York.
- [8] Haslett, J. and Haslett, S. (2007). The three basic types of residuals for a linear model. *Int. Statist. Rev.* 75, 1-24.
- [9] Haslett, S. (1985). Recursive estimation of the general linear model with dependent errors and multiple additional observations. *Aust. J. Statist.* 27, 183-188.
- [10] Hawkins, D.M. (1991). Diagnostics for use with regression recursive residuals. *Technometrics* 33, 221-234.

- [11] Heath, M.T. (2002). *Scientific Computing: A Survey* 2nd edition. McGraw-Hill, New York.
- [12] Hedayat, A. and Robson, D.S. (1970). Independent stepwise residuals for testing homoscedasticity. *J. Amer. Statist. Assoc.* 65, 1573-1581.
- [13] Jammalamadaka, S.R. and Sengupta, D. (1999). Changes in the general linear model: a unified approach. *Linear Algebra Appl.* 289, 225-242.
- [14] John, J.A. and Williams, E.R. (1995). *Cyclic and Computer Generated Designs* 2nd edition. London: Chapman and Hall.
- [15] Kendall, M.G. and Stuart, A. (1963). *The Advanced Theory of Statistics: Volume 1* 2nd edition. London: Griffin.
- [16] Kianifard, F. and Swallow, W.H. (1996). A review of the development and application of recursive residuals in linear models. *J. Amer. Statist. Assoc.* 91, 391-400.
- [17] Kontoghiorghes, E.J. (2004). Computational methods for modifying SURE models. *J. Computational Applied Mathematics* 162, 247-261.
- [18] Kontoghiorghes, E.J. and Foschi, P. (2000). Computationally efficient methods for solving SURE models. pp. 490-498 in *Lecture Notes in Computer Science* (L. Vulkov, J. Wasniewski and P. Yalamov, eds) Heidelberg: Springer.
- [19] Loynes, R.M. (1986) Recursive and related residuals *Statistica Neerlandica*, 40, 225-235.
- [20] Luger, R. (2001) A modified CUSUM test for orthogonal structural changes. *Econ. Letters*, 73, 301-306.
- [21] McGilchrist, C.A., Lianto, S. and Byron, D.M. (1989) Signed residuals. *Biometrics* 45, 237-246.
- [22] Magnus, J.R. and Sinha, A.K. (2005). On Theil's errors. *Econ. J.* 8, 39-54.
- [23] Nelder, J.A. (1975) Contribution to Brown, R.L. et al. (1975) *J. Roy. Statist. Soc. B* 37, 182-183.

- [24] Rao, C.R. (1962) A note on a generalized inverse of a matrix with applications to problems in mathematical statistics. *J. R. Statist. Soc. B* 24, 152-158.
- [25] Rao, C.R. (1973) *Linear Statistical Inference and its Applications*. Second edition. Wiley, New York.
- [26] Sen, P.K. (1982) Invariance principles for recursive residuals. *Ann. Statist.* 10, 307-312.
- [27] Theil, H. (1965). The analysis of disturbances in regression analysis. *J. Amer. Statist. Assoc.* 60, 1067-1079.
- [28] Wright, J.H. (1999). A new test for structural stability based on recursive residuals. *Oxford Bull. Econ. Statist.* 61, 109-119.
- [29] Zyskind, G. (1967) On canonical forms, non-negative covariance matrices and best and simple least squares linear estimators in linear models. *Ann. Math. Statist.* 38, 1092-1109.