COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 680

THE EXACT DISTRIBUTION OF ZELLNER'S SUR

Peter C. B. Phillips

August 1983

THE EXACT DISTRIBUTION OF ZELLNER'S SUR

by

P. C. B. Phillips*

0. ABSTRACT

This paper derives the exact finite sample distribution of the two-stage GLS (generalized least squares) estimator in a multivariate linear model with general linear parameter restrictions. This includes the seemingly unrelated regression (SUR) model as a special case and generalizes presently known exact results for the latter system. The usual classical assumptions are made concerning nonrandom exogenous variables and normally distributed errors.

*My thanks go to Glena Ames for her skill and effort in typing the manuscript of this paper.

1. INTRODUCTION

In the early 1960's Zellner [8] developed a two-stage GLS estimator for the coefficients in a linear multivariate system that is now popularly known as the SUR model. This two-stage procedure has since been used in many empirical applications. GLS also forms the basis of other commonly used estimators both in linear models with heteroscedastic or autocorrelated errors and in simultaneous equation systems where it leads to three stage least squares (3SLS). In spite of extensive research and perhaps surprisingly in view of the popularity of GLS methods in empirical work, the exact finite sample distribution of the SUR estimator is known only in highly specialized cases. These cases effectively restrict attention to two equation systems and models with orthogonal regressors [2]. Existing distribution theory is even more limited in the case of other commonly used GLS estimators, such as the two-stage estimator in linear models with heteroscedastic errors. Here, only low order moment formulae are known and then only in the simplest two sample setting.

The research underlying the present paper is motivated by the deficiencies outlined above. Our initial object of study was the exact distribution of the SUR estimator in the general case. But the methods we have developed open the way to an exact distribution theory for econometric estimators in a much wider setting than the SUR model. The present paper will derive the exact finite sample distribution of the two-stage GLS estimator in the multivariate linear model subject to general linear parameter restrictions. This generalizes all presently known distribution theory for the SUR models itself. Two important specializations of our results will be illustrated in detail.

A sequel to the present paper [5] will extend the distribution theory given here to other commonly used estimators of the GLS genre. Later work by the author will use the methods we have developed to open up an exact distribution theory for a wider class of econometric estimators.

2. THE MODEL AND NOTATION

We will work with the model

$$(1) \qquad y_t = Ax_t + u_t \ , \quad t = 1, \ldots, T$$

where y_t is an $n \times 1$ vector of endogenous variables, x_t is an $m \times 1$ vector of non-random exogenous variables and the u_t $(t = 1, \ldots, T)$ are i.i.d. $N(0, \Sigma)$ with non-singular covariance matrix Σ. We write (1) as $Y' = AX' + U'$ where the data matrices are assembled in columns as in $Y' = [y_1, \ldots, y_T]$; and we assume that X has full rank m.

The coefficient matrix A in (1) is assumed to be parameterized in the form

$$(2) \qquad \mathrm{vec}(A) = S\alpha - s$$

where vec( ) denotes vectorization by rows, S is an $nm \times q$ matrix whose elements are known constants and whose rank is q and s is a vector of known constants. In (2) α is taken as the $(q \times 1)$ vector of basic parameters.

The model given by (1) and (2) includes the SUR models as a special case as well as Malinvaud's general linear model [3] which allows for the same parameters to occur in more than one equation. Moreover, the model (1) and explicit parameterization (2) are formally equivalent to the same model (1) with the coefficient matrix A subject to $p = nm-q$ general linear restrictions

$$(3) \qquad D\,\mathrm{vec}(A) = d$$

where D is a $p \times nm$ matrix of rank p and d is a $p \times 1$ vector. All of our results apply to the restricted regression model (1) and (3) upon appropriate symbolic translation. We will therefore confine our attention in what follows to the explicit parameterization (2).

The GLS estimator of α is given by

$$(4) \qquad \hat{\alpha} = \{S'(\Sigma^{-1} \otimes X'X)S\}^{-1}\{S'(\Sigma^{-1} \otimes X')\mathrm{vec}(Y') + S'(\Sigma^{-1} \otimes X'X)s\} \ .$$

The two-stage estimator of α is obtained by replacing Σ in (4) by an estimate that is typically based on the residuals of a preliminary least squares regression on (1). We take the estimate

$$(5) \qquad \Sigma^* = (T-m)^{-1}Y'(I - P_X)Y \ , \quad P_X = X(X'X)^{-1}X'$$

from an unrestricted regression. The corresponding two-stage estimate of α we will denote by α^* . The error in this estimate satisfies:

$$(6) \qquad \alpha^* - \alpha = \{S'(\Sigma^{*-1} \otimes X'X)S\}^{-1}\{S'(\Sigma^{*-1} \otimes X')\mathrm{vec}(U')\} \ .$$

3. THE EXACT DISTRIBUTION OF α^*

Define $M = T^{-1}X'X$, $p = \mathrm{vec}(U'X/T)$ and $D = Y'(I - P_X)Y$. We write the error in the estimator α^* given by (6) in the generic form

$$(7) \qquad \alpha^* - \alpha = e(p,D) \ .$$

Our approach is to work with the conditional distribution of e given D and then average over the distribution of D to achieve the marginal probability density function (p.d.f.) of e.

Since p is $N(0, \Sigma \otimes M/T)$ the conditional p.d.f. of e given D is

$$(8) \qquad \mathrm{pdf}(e|D) = \frac{T^{q/2}\exp\left\{-\frac{T}{2}e'[B(\Sigma \otimes M)B']^{-1}e\right\}}{(2\pi)^{q/2}(\det[B(\Sigma \otimes M)B'])^{1/2}}$$

where

$$(9) \qquad B = [S'(D^{-1} \otimes M)S]^{-1}[S'(D^{-1} \otimes I)] \ .$$

The matrix D is central Wishart with p.d.f.

$$(10) \qquad \mathrm{pdf}(D) = \frac{\mathrm{etr}\left(-\frac{1}{2}\Sigma^{-1}D\right)(\det D)^{(T-m-n-1)/2}}{2^{n(T-m)/2}\Gamma_n\left(\frac{T-m}{2}\right)(\det \Sigma)^{(T-m)/2}} \ .$$

It follows that the unconditional p.d.f. of e is given by the integral

$$\mathrm{pdf}(e) = \frac{T^{q/2}}{(2\pi)^{q/2}2^{n(T-m)/2}\Gamma_n\left(\frac{T-m}{2}\right)(\det \Sigma)^{(T-m)/2}}$$

$$\int_{D>0} \frac{\mathrm{etr}\left(-\frac{1}{2}\Sigma^{-1}D\right)\exp\left\{-\frac{T}{2}e'[B(\Sigma \otimes M)B']^{-1}e\right\}(\det D)^{(T-m-n-1)/2}dD}{(\det[B(\Sigma \otimes M)B'])^{1/2}}$$

$$(11) \quad = \frac{(T/2\pi)^{q/2}}{2^{n(T-m)/2}\Gamma_n\left(\frac{T-m}{2}\right)(\det \Sigma)^{(T-m)/2}} \sum_{j=0}^{\infty} \frac{\left(-\frac{T}{2}\right)^j}{j!}$$

$$\int_{D>0} \frac{\mathrm{etr}\left(-\frac{1}{2}\Sigma^{-1}D\right)(\det D)^{(T-m-n-1)/2}(e'[B(\Sigma \otimes M)B]^{-1}e)^j \, dD}{(\det[B(\Sigma \otimes M)B']^{1/2}}$$

where term by term integration of the series is justified by uniform convergence.

We now decompose the matrix

$$(12)\qquad B(\Sigma \otimes M)B' = [S'(D^{-1} \otimes M)S]^{-1}[S'(D^{-1}\Sigma D^{-1} \otimes M)S][S'(D^{-1} \otimes M)S]^{-1}$$

$$= [S'(D_a \otimes M)S]^{-1}[S'(D_a\Sigma D_a \otimes M)S][S'(D_a \otimes M)S]^{-1}$$

where the suffix a is used to indicate the adjoint of the associated matrix. The integral we need to evaluate in (11) has the following form:

$$(13)\qquad \int_{D>0} \operatorname{etr}\left(-\tfrac{1}{2}\Sigma^{-1}D\right)(\det D)^{(T-m-n-1)/2}\left\{e'[S'(D_a \otimes M)S][S'(D_a\Sigma D_a \otimes M)S]_a[S'(D_a \otimes M)S]e\right\}^j$$

$$\cdot \det[S'(D_a \otimes M)S]\left\{\det[S'(D_a\Sigma D_a \otimes M)S]\right\}^{-j-\frac{1}{2}}dD\ .$$

We introduce an $n \times n$ matrix W of auxiliary variables and using (A12), (A13) and (A14) in the Appendix we deduce that (13) equals

$$(14)\qquad \Bigg[\left\{e'[S'(\partial W_a \otimes M)S][S'(\partial W_a\Sigma\partial W_a \otimes M)S]_a[S'(\partial W_a \otimes M)S]e\right\}^j$$

$$\cdot \det[S'(\partial W_a \otimes M)S]\left\{\det[S'(\partial W_a\Sigma\partial W_a \otimes M)S]\right\}^{-j-\frac{1}{2}}$$

$$\cdot \int_{D>0} \operatorname{etr}\left[-\left(\tfrac{1}{2}\Sigma^{-1} - W\right)D\right](\det D)^{(T-m-n-1)/2}dD\Bigg]_{W=0}$$

In this expression ∂W_a denotes the adjoint of the matrix operator $\partial W = \partial/\partial W$ (in the notation of the Appendix) and the fractional matrix operator that appears in (14) is defined as in (A9) in the Appendix.

We evaluate the multivariate gamma integral in (14) and from (11) we then deduce:

$$(15)\quad \mathrm{pdf}(e) = \frac{(T/2\pi)^{q/2}}{2^{n(T-m)/2}\Gamma_n\left(\frac{T-m}{2}\right)(\det\ \Sigma)^{(T-m)/2}}$$

$$\cdot \sum_{j=0}^{\infty} \frac{\left(-\frac{T}{2}\right)^j}{j!}\left[\left\{e'[S'(\partial W_a \otimes M)S][S'(\partial W_a \Sigma \partial W_a \otimes M)S]_a[S'(\partial W_a \otimes M)S]e\right\}^j\right.$$

$$\cdot \det[S'(\partial W_a \otimes M)S]\left\{\det[S'(\partial W_a \Sigma \partial W_a \otimes M)S]\right\}^{-j-\frac{1}{2}}$$

$$\left.\cdot \Gamma_n\left(\frac{T-m}{2}\right)\left[\det\left(\frac{1}{2}\Sigma^{-1} - W\right)\right]^{-(T-m)/2}\right]_{W=0}$$

$$(16)\quad = \left(\frac{T}{2\pi}\right)^{q/2} \sum_{j=0}^{\infty} \frac{\left(-\frac{T}{2}\right)^j}{j!}\left[\left\{e'[S'(\partial W_a \otimes M)S][S'(\partial W_a \Sigma \partial W_a \otimes M)S]_a[S'(\partial W_a \otimes M)S]e\right\}^j\right.$$

$$\cdot \det[S'(\partial W_a \otimes M)S]\left\{\det[S'(\partial W_a \Sigma \partial W_a \otimes M)S]\right\}^{-j-\frac{1}{2}}$$

$$\left.\cdot [\det(I - 2\Sigma W)]^{-(T-m)/2}\right]_{W=0} .$$

In generalized operator notation this expression (16) for the p.d.f. of $e = \alpha^* - \alpha$ may be written more simply as:

$$(17)\quad \mathrm{pdf}(e) = \frac{T^{q/2}\exp\left\{-\frac{T}{2}e'G(\partial W_a)^{-1}e\right\}}{(2\pi)^{q/2}[\det\ G(\partial W_a)]^{1/2}}[\det(I - 2\Sigma W)]^{-(T-m)/2}\Bigg|_{W=0}$$

where

$$G(\partial W_a) = [S'(\partial W_a \otimes M)S]^{-1}[S'(\partial W_a \Sigma \partial W_a \otimes M)S][S'(\partial W_a \otimes M)S]^{-1} .$$

4. MARGINAL DISTRIBUTIONS

Let F be an $f \times q$ matrix of known constants of full rank f ($<q$) and consider the marginal distribution of $g = Fe$. The joint marginal p.d.f. of g is deduced by the same sequence of operations as those developed above for the full dimensional case. The final result corresponding to (16) is:

$$\text{(18)}\quad \mathrm{pdf}(g) = \left(\frac{T}{2\pi}\right)^{f/2} \sum_{j=0}^{\infty} \frac{\left(-\frac{T}{2}\right)^j}{j!} \Bigg[\Big\{g'[F\{S'(\partial W_a \otimes M)S\}_a\{S'(\partial W_a \Sigma \partial W_a \otimes M)S\}$$

$$\cdot\{S'(\partial W_a \otimes M)S\}_a F']_a g\Big\}^j (\det[S'(\partial W_a \otimes M)S])^{f+2j}$$

$$\cdot\left(\det\left[F\{S'(\partial W_a \otimes M)S\}_a\{S'(\partial W_a \Sigma \partial W_a \otimes M)S\}\{S'(\partial W_a \otimes M)S\}_a F'\right]\right)^{-j-\frac{1}{2}}$$

$$[\det(I - 2\Sigma W)]^{-(T-m)/2}\Bigg]_{W=0} ;$$

or in generalized operator form:

$$\text{(19)}\quad \mathrm{pdf}(g) = \frac{T^{f/2}\exp\left\{-\frac{T}{2}g'H(\partial W_a)^{-1}g\right\}}{(2\pi)^{f/2}[\det H(\partial W_a)]^{1/2}}[\det(I - 2\Sigma W)]^{-(T-m)/2}\Bigg|_{W=0}$$

where

$$H(\partial W_a) = F[S'(\partial W_a \otimes M)S]^{-1}[S'(\partial W_a \Sigma \partial W_a \otimes M)S][S'(\partial W_a \otimes M)S]^{-1}F' .$$

5. SPECIALIZATIONS

5.1. The unrestricted model

In this case $\alpha = \text{vec}(A)$ and α^* is the unrestricted least squares estimator. To reduce the general expression for the density (16) we note that $(\partial W_a)(\partial W_a)_a = I(\det \partial W)^{n-1}$ and thereby

$$(\partial W_a \otimes M)(\partial W_a \Sigma \partial W_a \otimes M)_a (\partial W_a \otimes M)$$

$$= (\partial W_a \otimes M)\Big((\partial W_a)_a \Sigma_a (\partial W_a)_a \otimes M_a\Big)(\partial W_a \otimes M)[\det(\partial W_a \Sigma \partial W_a)]^{m-1}(\det M)^{n-1}$$

$$= (\Sigma_a \otimes MM_a M)(\det \partial W)^{2n-2}(\det \partial W_a)^{2m-2}(\det \Sigma)^{m-1}(\det M)^{n-1}$$

$$= (\Sigma^{-1} \otimes M)(\det \Sigma)^m (\det M)^n (\det \partial W)^{2m(n-1)} \; .$$

Additionally,

$$\det(\partial W_a \otimes M)\Big\{\det(\partial W_a \Sigma \partial W_a \otimes M)\Big\}^{-j-\frac{1}{2}}$$

$$= (\det M)^n \Big\{\det(\Sigma \otimes M)\Big\}^{-j-\frac{1}{2}} (\det \partial W_a)^{m-2m(j+\frac{1}{2})}$$

$$= (\det M)^n \Big\{\det(\Sigma \otimes M)\Big\}^{-j-\frac{1}{2}} (\det \partial W)^{m(n-1)-2m(n-1)(j+\frac{1}{2})} \; .$$

Using these reductions in (16) we deduce that:

$$\text{pdf}(e) = \left(\frac{T}{2\pi}\right)^{q/2} \sum_{j=0}^{\infty} \frac{\left(-\frac{T}{2}\right)^j}{j!} \left\{e'(\Sigma^{-1} \otimes M)e\right\}^j \Big[(\det \Sigma)^{mj} (\det M)^{nj} (\det \partial W)^{2m(n-1)j} (\det M)^n$$

$$\cdot \left\{\det(\Sigma \otimes M)\right\}^{-j-\frac{1}{2}} (\det \partial W)^{m(n-1)-2m(n-1)(j+\frac{1}{2})} \left\{\det(I - 2\Sigma W)\right\}^{-(T-m)/2}\Big]_{W=0}$$

$$= \left(\frac{T}{2\pi}\right)^{q/2} \sum_{j=0}^{\infty} \frac{\left(-\frac{T}{2}\right)^j}{j!} \left\{e'(\Sigma^{-1} \otimes M)e\right\}^j (\det \Sigma)^{-m/2} (\det M)^{n/2}$$

$$(20) \qquad = \frac{T^{q/2} \exp\left\{-\frac{T}{2} e'(\Sigma^{-1} \otimes M)e\right\}}{(2\pi)^{q/2} [\det(\Sigma \otimes M^{-1})]^{1/2}} , \quad q = nm .$$

Thus, the joint density of $e = \alpha^* - \alpha$ reduces to the well known multivariate $N(0, \Sigma \otimes (X'X)^{-1})$.

5.2. The Zellner model with pairwise orthogonal regressors

We specialize the multivariate system (1) and (2) to the Zellner model

$$(21) \qquad y_{k\cdot} = X_k \beta_k + u_{k\cdot} , \quad k = 1, \ldots, n$$

where $y_{k\cdot}$ is the observation vector on the k^{th} dependent variable, X_k is a $T \times L_k$ observation matrix on L_k regressors, β_k is a vector of coefficients and $u_{k\cdot}$ is a vector of normally distributed serially independent errors. The covariance matrix of $u' = (u'_{1\cdot}, \ldots, u'_{n\cdot})$ is $\Sigma \otimes I$.

When we stack the model (21) and set $\alpha' = (\beta'_1, \ldots, \beta'_n)$ the Zellner SUR estimator of α is α^* and its exact distribution is given by (17). The marginal distribution of β^*_k , the SUR estimator of the subvector of the coefficients in the k^{th} equation can be deduced directly from (18).

To relate our results to the existing literature we now assume pairwise orthogonal regressors across the equations of (21), so that $X'_i X_j = 0$ for $i \neq j$. We may concentrate on the first equation of the system without loss of generality and set $k = 1$, $e_{11} = \beta_1^* - \beta_1$ and $M_{11} = T^{-1}X_1'X_1$. From the general expression (18) the joint marginal density of e_{11} is found to be:

$$(22)\quad \mathrm{pdf}(e_{11}) = \left(\frac{T}{2\pi}\right)^{L_1/2} (\det M_{11})^{1/2} \sum_{j=0}^{\infty} \frac{\left(-\frac{T}{2}\right)^j}{j!}$$

$$\cdot (e_{11}'M_{11}e_{11})^j \left[\frac{(\partial W_a)_{11}^{2j+L_1}}{(\partial W_a \Sigma \partial W_a)_{11}^{j+L_1/2}} [\det(I - 2\Sigma W)]^{-(T-m)/2} \right]_{W=0} .$$

Simple manipulations now verify that

$$(\partial W_a)_{11}^{2j+L_1} (\partial W_a \Sigma \partial W_a)_{11}^{-j-L_1/2} [\det(I - 2\Sigma W)]^{-(T-m)/2} \Big|_{W=0}$$

$$= \frac{1}{\Gamma_n\left(\frac{T-m}{2}\right)} \int_{S>0} \mathrm{etr}(-S)(\det S)^{\frac{T-m}{2} - \frac{n+1}{2}} (c'S^{-1}c)^{2j+L_1} (c'S^{-2}c)^{-j-L_1/2} dS$$

$$= \frac{(c'c)^{j+L_1/2}}{\Gamma_n\left(\frac{T-m}{2}\right)} \int_{S>0} \mathrm{etr}(-S)(\det S)^{\frac{T-m}{2} - \frac{n+1}{2}} (S^{-1})_{11}^{2j+L_1} (S^{-2})_{11}^{-j-L_1/2} dS$$

$$(23)\quad = \frac{(c'c)^{j+L_1/2} \Gamma\left(\frac{T-m}{2} + \frac{1}{2}\right) \Gamma\left(\frac{T-m}{2} + \frac{L_1}{2} - \frac{n}{2} + j + 1\right)}{\Gamma\left(\frac{T-m}{2} - \frac{n}{2} + 1\right) \Gamma\left(\frac{T-m}{2} + \frac{L_1}{2} + \frac{1}{2} + j\right)} .$$

where c' denotes the first row of $\Sigma^{-1/2}$. The final step in the argument leading to (23) is an interesting exercise for the careful reader. If we write Σ in partitioned form as

$$\Sigma = \left[\begin{array}{c|c} \sigma_{11} & \sigma_{12} \\ \hline \sigma_{21} & \Sigma_{22} \end{array} \right]$$

and set $\sigma_{11\cdot 2} = \mathrm{var}(u_{1t}|u_{2t} \ldots u_{nt}) = \sigma_{11} - \sigma_{12}\Sigma_{22}^{-1}\sigma_{21}$ we find that $c'c = \sigma^{11} = (\sigma_{11} - \sigma_{12}\Sigma_{22}^{-1}\sigma_{21})^{-1} = \sigma_{11\cdot 2}^{-1}$. We may now deduce from (22) and (23) the following expression for the joint marginal density of e_{11} :

$$(24) \quad \mathrm{pdf}(e_{11}) = \left(\frac{T}{2\pi}\right)^{L_1/2} [\det(\sigma_{11\cdot 2}M_{11}^{-1})]^{-1/2}$$

$$\frac{\Gamma\left(\frac{T-m}{2} + \frac{1}{2}\right)}{\Gamma\left(\frac{T-m}{2} - \frac{n}{2} + 1\right)} \sum_{j=0}^{\infty} \frac{\Gamma\left(\frac{T-m}{2} + \frac{L_1}{2} - \frac{n}{2} + j + 1\right)}{\Gamma\left(\frac{T-m}{2} + \frac{L_1}{2} + \frac{1}{2} + j\right) j!} \left(- \frac{Te_{11}'M_{11}e_{11}}{2\sigma_{11\cdot 2}}\right)^j .$$

Upon translation of notation this is the expression found in [2] by direct methods.

6. FINAL REMARKS

The general formulae (16) and (18) may be used to deduce the corresponding asymptotic distributions in a simple way. We replace $[\det(I - 2\Sigma W)]^{-(T-m)/2}$ in (16) or (18) by an asymptotic approximation as $W \to 0$. It is simplest to use $[\det(I - 2\Sigma W)]^{-(T-m)/2} \sim \mathrm{etr}[(T-m)\ W]$. Upon evaluation, we see that (16) now yields the asymptotic $N(0,\ T^{-1}[S'(\Sigma^{-1} \otimes M)S]^{-1})$ approximation directly. Higher order asymptotics may be obtained in a similar way although the algebra is more complicated.

The calculus developed in the Appendix may be applied to a variety of other unsolved problems in econometric distribution theory, including systems estimators in simultaneous equations and Stein-like estimators. The author currently has some work on these problems underway.

APPENDIX: FRACTIONAL MATRIX CALCULUS

This appendix extends the theory of fractional operators in differential calculus to matrix spaces. Those readers who are unfamiliar with scalar fractional operators as a generalization of traditional differential and integral calculus may wish to refer to the reviews in [6] and [7]. Additionally, [6] provides an interesting historical survey which traces the theory of fractional differentiation to the work of Leibnitz, Liouville and Euler.

Let $\mathcal{A}$ be the space of $n \times n$ symmetric matrices, $O(n)$ the group of $n \times n$ orthogonal matrices and $\mathcal{C}$ the class of symmetric functions on $\mathcal{A}$. $\mathcal{C}$ is defined as the set of all complex analytic functions of $\mathcal{A}$ for which $f(X) = f(HXH')$ for all $H \in O(n)$ and where $X \in \mathcal{A}$. Since $f \in \mathcal{C}$ is a complex analytic function of the n elementary symmetric functions of X (viz. $\sigma_1 = tr(X)$, $\sigma_2 = \ldots$, $\sigma_n = \det X$), the domain of definition of f may be extended to all complex $n \times n$ matrices X for which $f(X)$ continues to be defined [1]. In what follows we let X be an arbitrary complex $n \times n$ matrix and we use the notation ∂X to denote the matrix operator $\partial/\partial X$.

<u>*Definition.*</u> *If* f *is a complex analytic function of* X *and* $\alpha > 0$ *we define the fractional matrix operator* $(\det \partial X)^{-\alpha}$ *by the integral*

$$(A1) \qquad (\det \partial X)^{-\alpha} f(X) = \frac{1}{\Gamma_n(\alpha)} \int_{S>0} f(X-S)(\det S)^{\alpha-(n+1)/2} dS$$

when it exists. The integral is taken over the set of positive definite matrices $S > 0$.

This definition is motivated by the observation that $(\det \partial X)^{-\alpha}$ may be formally considered as the operator

(A2) $$(\det \partial X)^{-\alpha} = \frac{1}{\Gamma_n(\alpha)} \int_{S>0} \operatorname{etr}(-\partial XS)(\det S)^{\alpha-(n+1)/2} dS$$

using the multivariate gamma integral [1]. We observe that since f is analytic

(A3) $$\operatorname{etr}(-\partial XS)f(X) = f(X-S)$$

leading directly to (A1).

The definition (A1) is a matrix generalization of the Weyl fractional integral [4]. Indeed, when X is a scalar we deduce from (A1) (with $D = d/dx$)

(A4) $$D^{-\alpha}f(x) = \frac{1}{\Gamma(\alpha)} \int_0^\infty f(x-s)s^{\alpha-1} ds = \frac{1}{\Gamma(\alpha)} \int_{-\infty}^x f(y)(x-y)^{\alpha-1} dy$$

which is one representation of the Weyl definition of a fractional integral. The Weyl integral (A4), when it exists, satisfies the familiar law of exponents $D^{-\alpha}D^{-\beta} = D^{-(\alpha+\beta)}$ for all α and β (with fractional differentiation being defined in terms of fractional integration to an appropriate order of a traditional derivative to the nearest integral order: see [4] for a full development).

(A1) may be used to define a fractional matrix derivative as follows. If $\mu > 0$ and $[\mu]$ = integral part of μ we set $\mu = n-\alpha$ where $n = [\mu] + 1$ and $0 < \alpha < 1$. Then the fractional matrix (differential) operator $(\det \partial X)^\mu$ is defined as:

(A5) $$(\det \partial X)^\mu = (\det \partial X)^{-\alpha}[(\det \partial X)^n f(X)] .$$

With these definitions we can verify that the fractional matrix operator satisfies the usual law of exponents. Writing $D_X = \det \partial X$ we have (for

$\alpha, \beta > 0$) :

$$D_X^{-\alpha}[D_X^{-\beta}f(X)] = D_X^{-\alpha}\frac{1}{\Gamma_n(\beta)}\int_{S>0} f(X-S)(\det S)^{\beta-(n+1)/2}dS$$

$$= \frac{1}{\Gamma_n(\alpha)\Gamma_n(\beta)}\int_{R>0}\int_{S>0} f(X-R-S)(\det S)^{\beta-(n+1)/2}dS(\det R)^{\alpha-(n+1)/2}dR$$

$$= \frac{1}{\Gamma_n(\alpha)\Gamma_n(\beta)}\int_{M>0} f(X-M)\int_{Q=0}^{M}[\det(M-Q)]^{\alpha-(n+1)/2}(\det Q)^{\beta-(n+1)/2}dQ$$

$$= \frac{1}{\Gamma_n(\alpha)\Gamma_n(\beta)}\int_{M>0} f(X-M)(\det M)^{\alpha+\beta-(n+1)/2}\int_0^I[\det(I-T)]^{\alpha-(n+1)/2}(\det T)^{\beta-(n+1)/2}dT$$

$$= \frac{1}{\Gamma_n(\alpha+\beta)}\int_{M>0} f(X-M)(\det M)^{\alpha+\beta-(n+1)/2}dM$$

$$= D_X^{-\alpha-\beta}f(X) .$$

To prove that the law holds for general indices as well as negative indices we use the argument that, if $\mu = n-\alpha$ and $\nu = m-\beta$ with n and m integer, then:

$$D_X^{\mu}[D_X^{\nu}f] = D_X^{-\alpha}D_X^{n}[D_X^{-\beta}D_X^{m}f] = D_X^{-\alpha}[D_X^{-\beta}D_X^{n+m}f] = D_X^{-\alpha-\beta}[D_X^{n+m}f] = D_X^{\mu+\nu}f .$$

The following examples illustrate the use of the fractional matrix operator on elementary functions of matrix argument:

(A6) $D_X^{\mu}\mathrm{etr}(AX) = \mathrm{etr}(AX)(\det A)^{\mu}$, all μ

(A7) $D_X^{\mu}[\det(I-X)]^{-\alpha} = \dfrac{\Gamma_n(\alpha-\mu)}{\Gamma_n(\alpha)}[\det(I-X)]^{-\alpha+\mu}$, $\mathrm{Re}(\alpha) > (n-1)/2$, $\mathrm{Re}(\alpha-\mu) > (n-1)/2$

(A8) $D_X\ {}_1F_1(\alpha,\beta;X) = \dfrac{\Gamma_n(\alpha-\mu)\Gamma_n(\beta-\mu)}{\Gamma_n(\alpha)\Gamma_n(\beta)}\ {}_1F_1(\alpha-\mu,\ \beta-\mu;\ X)$.

It is frequently convenient to work with the adjoint of the matrix operator $\partial/\partial X$, which we will write in form $\partial X_a = \mathrm{adj}(\partial/\partial X)$. The fractional calculus can be extended to this operator as in (A2) and (A5). We define

$$\text{(A9)} \quad (\det \partial X_a)^{-\alpha} f(X) = \frac{1}{\Gamma_n(\alpha)} \int_{S>0} \{\mathrm{etr}(-\partial X_a S) f(X)\} (\det S)^{\alpha-(n+1)/2} dS \ , \quad \alpha > 0$$

provided the integral exists; and, for $\mu > 0$,

$$\text{(A10)} \quad (\det \partial X_a)^{\mu} f(X) = (\det \partial X_a)^{-\alpha} [(\det \partial X_a)^n f(X)] \ , \quad \mu = n-\alpha \ , \quad n \text{ integer.}$$

In our applications of this calculus $f(X)$ is often the elementary function $\mathrm{etr}(AX)$. Simple manipulations verify that for this function

$$\text{(A11)} \qquad (\det \partial X_a)^{\mu} \mathrm{etr}(AX) = \mathrm{etr}(AX)(\det A_a)^{\mu}$$

where $A_a = \mathrm{adj}(A)$. We also need to work with the operators $\det[S'(\partial X_a \otimes M)S]$ and $\det[S'(\partial X_a \Sigma \partial X_a \otimes M)S]$ where S is an $nm \times q$ matrix of rank q, Σ is $n \times n$ positive definite and M is $m \times m$ positive definite. Extensions of (A11) to these operators yield:

$$\begin{aligned} \text{(A12)} \quad \{\det[S'(\partial X_a \otimes M)S]\}^{\mu} \mathrm{etr}(AX) &= \mathrm{etr}(AX)\{\det[S'(A_a \otimes M)S]\} \\ &= \mathrm{etr}(AX)\{\det[S'(A^{-1} \otimes M)S]\}^{\mu} (\det A)^{\mu q} \end{aligned}$$

$$\begin{aligned} \text{(A13)} \quad \{\det[S'(\partial X_a \Sigma \partial X_a \otimes M)S]\}^{\mu} \mathrm{etr}(AX) &= \mathrm{etr}(AX)\{\det[S'(A_a \Sigma A_a \otimes M)S]\} \\ &= \mathrm{etr}(AX)\{\det[S'(A^{-1}\Sigma A^{-1} \otimes M)S]\}^{\mu} (\det A)^{2\mu q} \ . \end{aligned}$$

We also note that

$$g'[S'(\partial X_a \otimes M)S][S'(\partial X_a \Sigma \partial X_a \otimes M)S]_a [S'(\partial X_a \otimes M)S]g \ \mathrm{etr}(AX)$$

$$= \mathrm{etr}(AX) g'[S'(A_a \otimes M)S][S'(A_a \Sigma A_a \otimes M)S]_a [S'(A_a \otimes M)S]g$$

where g is any $q \times 1$ vector. Repeated use of this operator yields:

$$\text{(A14)} \quad \{g'[S'(\partial X_a \otimes M)S][S'(\partial X_a \Sigma \partial X_a \otimes M)S]_a [S'(\partial X_a \otimes M)S]g\}^j \mathrm{etr}(AX)$$

$$= \mathrm{etr}(AX)\{g'[S'(A_a \otimes M)S][S'(A_a \Sigma A_a \otimes M)S]_a [S'(A_a \otimes M)S]g\}^j \ .$$

REFERENCES

[1] Herz, C. W. "Bessel Functions of Matrix Argument," _Annals of Mathematics_, 1955, 61, 474-523.

[2] Kataoka, Y. "The Exact Finite Sample Distributions of Joint Least Squares Estimators for Seemingly Unrelated Regressions," _Economic Studies Quarterly_, 1974, 25, 36-44.

[3] Malinvaud, E. _Statistical Methods of Econometrics_, North-Holland: Amsterdam, 1980.

[4] Miller, K. S. "The Weyl Fractional Calculus" in B. Ross (ed.), _Fractional Calculus and Its Applications_, Springer-Verlag: Berlin, 1974.

[5] Phillips, P. C. B. "The Exact Distribution of Various Two Stage Estimators" (in preparation).

[6] Ross, B. _Fractional Calculus and Its Applications_, Springer-Verlag: Berlin, 1974.

[7] Spanier, J. and K. B. Oldham. _The Fractional Calculus_, Academic Press: New York, 1974.

[8] Zellner, A. "An Efficient Method of Estimating Seemingly Unrelated Regression Equations and Tests for Aggregation Bias," _Journal of the American Statistical Association_, 1962, 58, 348-368.