# Covariance of the Wishart Distribution with Applications to Regression

Ronald Christensen\*

April 30, 2015

### Abstract

We discuss the covariance matrix of the Wishart and the derivation of the covariance matrix of a regression estimate discussed in Tarpey et al. (2014). We also examine the relationship between inequalities considered in Cook, Forzani, and Rothman (2015) and Tarpey et al.'s response to that comment.

Keywords:Vector space, Wishart distribution, Wishart covariance.

<sup>\*</sup>Ronald Christensen is a Professor in the Department of Mathematics and of Statistics, University of New Mexico, Albuquerque, NM, 87131.

An appendix at the end contains results taken from Christensen (2011, Section B.6) on the algebra of using Kronecker products and Vec operators. The notation and concepts are in accord with Chistensen (2001, 2011). The first section introduces the vector space of  $p \times p$  matrices. The second section discusses the Wishart covariance matrix. The third section applies the results on the Wishart to the regression problem of Tarpey et al. (2014). The final section examines the relationship between inequalities considered in Cook, Forzani, and Rothman (2015) and Tarpey et al.'s response to that comment.

#### **1** The Vector Space of $p \times p$ Matrices

Clearly,  $p \times p$  matrices form a vector space under matrix addition and scalar multiplication. Note that transposing a matrix constitutes a linear transformation on the space of  $p \times p$  matrices. In other words, the transpose operator T(P) = P' is a linear transformation on the matrices.

Write the *p* dimensional identity matrix in terms of its columns,

$$I_p = [e_1, \cdots, e_p]$$

A basis for the  $p \times p$  matrices (that will turn out to be orthonormal) is  $e_i e'_j$ , i = 1, ..., p, j = 1, ..., p. For a matrix P with elements  $\rho_{ij}$ ,

$$P = \sum_{i=1}^{p} \sum_{j=1}^{p} \rho_{ij} e_i e'_j$$

The rank of the space is  $p^2$ .

Symmetric matrices are closed under matrix addition and scalar multiplication so they form a subspace with (eventually an orthogonal) basis  $(e_i e'_j + e_j e'_i)$ , i = 1, ..., p,  $j \le i$ . The rank of the subspace is p(p + 1)/2.

Matrix multiplication is defined by  $AP = \sum_{i=1}^{p} \sum_{j=1}^{p} \left( \sum_{k=1}^{p} a_{ik} \rho_{kj} \right) e_i e'_j$ .

In statistics a key reason for using vectors is that we like to look at linear combinations of the vector's elements. We can accomplish that by using an inner product on the vector space but for

most people it is probably easier to see what is going on if we turn  $p \times p$  matrices into column vectors in  $\mathbf{R}^{p^2}$ .

There are two ways to do this. We will identify the matrix P with

$$\operatorname{Vec}(P) = \sum_{i=1}^{p} \sum_{j=1}^{p} \rho_{ij}[e_j \otimes e_i] = \operatorname{Vec}\left(\sum_{i=1}^{p} \sum_{j=1}^{p} \rho_{ij}e_ie'_j\right).$$

Eaton (1982) (implicity) identifies P with  $\operatorname{Vec}(P') = \sum_{i=1}^{p} \sum_{j=1}^{p} \rho_{ij} [e_i \otimes e_j]$  which works just as well and may have notational advantages. It is important to keep track of which representation you use. For example, if the rows of an  $n \times p$  matrix Y are independent  $N_p(\mu, \Sigma)$  vectors,  $\operatorname{Cov}[\operatorname{Vec}(Y)] = [\Sigma \otimes I_n]$  whereas  $\operatorname{Cov}[\operatorname{Vec}(Y')] = [I_n \otimes \Sigma]$ , which is what Eaton calls  $\operatorname{Cov}(Y)$ .

A linear combination of two matrix vectors, say *P* and *A*, is

$$\operatorname{Vec}(P)'\operatorname{Vec}(A) = \operatorname{tr}(P'A).$$

This defines the standard Euclidean inner product on  $\mathbf{R}^{p^2}$  and a natural inner product on the  $p \times p$  matrices. Defining an inner product defines the concepts of vector length and orthogonality.

We can now write the transpose operator, which is linear, as a  $p^2 \times p^2$  matrix,

$$T = \sum_{i=1}^{p} \sum_{j=1}^{p} [e_j \otimes e_i] [e_i \otimes e_j]'.$$

Note that

$$T\operatorname{Vec}(P) = \left[\sum_{i=1}^{p}\sum_{j=1}^{p}[e_{j}\otimes e_{i}][e_{i}\otimes e_{j}]'\right]\left[\sum_{h=1}^{p}\sum_{k=1}^{p}\rho_{hk}[e_{k}\otimes e_{h}]\right] = \sum_{i=1}^{p}\sum_{j=1}^{p}\rho_{ji}[e_{j}\otimes e_{i}] = \operatorname{Vec}(P').$$

This is because  $[e_i \otimes e_j]'[e_k \otimes e_h] = 1$  when i = k and j = h but is 0 otherwise. In a similar vein, the identity matrix can be written as

$$I_{p^2} = \sum_{i=1}^p \sum_{j=1}^p [e_i \otimes e_j] [e_i \otimes e_j]'$$

and of course we must have  $TT = I_{p^2}$ .

### 2 The Wishart Covariance Matrix

Let  $y_1, \ldots, y_n$  be independent  $N_p(0, \Sigma)$  vectors and let

$$Y = \begin{bmatrix} y_1' \\ \vdots \\ y_n' \end{bmatrix}$$

As mentioned earlier,  $\operatorname{Cov} [\operatorname{Vec}(Y)] = [\Sigma \otimes I_p]$  whereas  $\operatorname{Cov} [\operatorname{Vec}(Y')] = [I_p \otimes \Sigma]$ , which is what Eaton (1982) calls  $\operatorname{Cov}(Y)$ .

Define  $W \equiv Y'Y$  so that  $W \sim W_p(n, \Sigma)$ . The standard result for the covariance matrix of a Wishart is

$$\operatorname{Cov}\left[\operatorname{Vec}(W)\right] = n[\Sigma \otimes \Sigma]\left[I_{p^2} + T\right]$$
(1)

where *T* is the transpose operator, i.e.,  $T \operatorname{Vec}(P) = \operatorname{Vec}(P')$ .

Because *W* is symmetric, i.e.  $w_{ij} = w_{ji}$ ,

$$\operatorname{Vec}(P)'\operatorname{Vec}(W) = \sum_{i=1}^{p} \sum_{j=1}^{p} \rho_{ij} w_{ij} = \sum_{i=1}^{p} \sum_{j=1}^{p} \rho_{ji} w_{ji} = \sum_{i=1}^{p} \sum_{j=1}^{p} \rho_{ji} w_{ij} = \operatorname{Vec}(P')'\operatorname{Vec}(W)$$

which implies that

$$\operatorname{Vec}\left[\frac{1}{2}(P+P')\right]'\operatorname{Vec}(W) = \operatorname{Vec}(P)'\operatorname{Vec}(W) = \operatorname{Vec}(P')'\operatorname{Vec}(W)$$

When considering linear combinations of the elements of W, without loss of generality, we can assume that P = P', in which case

$$\operatorname{Var}\left[\operatorname{Vec}(P)'\operatorname{Vec}(W)\right] = \operatorname{Vec}(P)'[\Sigma \otimes \Sigma]\left[I_{p^2} + T\right]\operatorname{Vec}(P) = 2\operatorname{Vec}(P)'[\Sigma \otimes \Sigma]\operatorname{Vec}(P),$$

which is the result in Eaton (1982). Moreover,

$$\operatorname{Cov}\left[\operatorname{Vec}(A)'\operatorname{Vec}(W), \operatorname{Vec}(P)'\operatorname{Vec}(W)\right] = \operatorname{Vec}(A)'[\Sigma \otimes \Sigma]\left[I_{p^2} + T\right]\operatorname{Vec}(P) = 2\operatorname{Vec}(A)'[\Sigma \otimes \Sigma]\operatorname{Vec}(P)$$

when either P or A is symmetric.

We now characterize the covariance matrix of Vec(W) via

$$\operatorname{Vec}(A)'[\Sigma \otimes \Sigma] [I_{p^2} + T] \operatorname{Vec}(P) = \sum_{i=1}^{p} \sum_{j=1}^{p} \sum_{h=1}^{p} \sum_{k=1}^{p} [a_{ij}\rho_{hk} + a_{ij}\rho_{kh}]\sigma_{jk}\sigma_{ih}$$
(2)

for arbitrary A and P (not necessarily symmetric). Alternatively,

$$\operatorname{Vec}(A)'[\Sigma \otimes \Sigma] \left[ I_{p^2} + T \right] \operatorname{Vec}(P) = \sum_{i=1}^p \sum_{j=1}^p \sum_{h=1}^p \sum_{k=1}^p a_{ij} \rho_{hk} [\sigma_{jk} \sigma_{ih} + \sigma_{jh} \sigma_{ik}]$$
(3)

The proofs of these results appear at the end of the section.

If  $\operatorname{Cov}[\operatorname{Vec}(W)] = n[\Sigma \otimes \Sigma] [I_{p^2} + T]$ , then  $[\Sigma \otimes \Sigma] [I_{p^2} + T]$  must be nonnegative definite, even though that fact is my no means obvious from the structure of the matrix. Just for completeness, we provide a direct proof.

**Proposition** The matrix  $[\Sigma \otimes \Sigma] [I_{p^2} + T]$  is nonnegative (positive) definite if  $\Sigma$  is nonnegative (positive) definite.

The proof is deferred to the end of the section.

We now present the proofs of equations (2) and (3). Equation (2) follows from

$$\operatorname{Vec}(A)'[\Sigma \otimes \Sigma]\operatorname{Vec}(P) = \sum_{i=1}^{p} \sum_{j=1}^{p} \sum_{h=1}^{p} \sum_{k=1}^{p} a_{ij} [e_{j} \otimes e_{i}]'[\Sigma \otimes \Sigma] \rho_{hk} [e_{k} \otimes e_{h}]$$
$$= \sum_{i=1}^{p} \sum_{j=1}^{p} \sum_{h=1}^{p} \sum_{k=1}^{p} a_{ij} \rho_{hk} [e'_{j} \Sigma e_{k} \otimes e'_{i} \Sigma e_{h}]$$
$$= \sum_{i=1}^{p} \sum_{j=1}^{p} \sum_{h=1}^{p} \sum_{k=1}^{p} a_{ij} \rho_{hk} \sigma_{jk} \sigma_{ih}$$
(4)

and

$$\operatorname{Vec}(A)'[\Sigma \otimes \Sigma]T \operatorname{Vec}(P) = \operatorname{Vec}(A)'[\Sigma \otimes \Sigma]\operatorname{Vec}(P')$$
  
$$= \sum_{i=1}^{p} \sum_{j=1}^{p} \sum_{h=1}^{p} \sum_{k=1}^{p} a_{ij}[e_{j} \otimes e_{i}]'[\Sigma \otimes \Sigma]\rho_{kh}[e_{k} \otimes e_{h}]$$
  
$$= \sum_{i=1}^{p} \sum_{j=1}^{p} \sum_{h=1}^{p} \sum_{k=1}^{p} a_{ij}\rho_{kh}[e'_{j}\Sigma e_{k} \otimes e'_{i}\Sigma e_{h}]$$
  
$$= \sum_{i=1}^{p} \sum_{j=1}^{p} \sum_{h=1}^{p} \sum_{k=1}^{p} a_{ij}\rho_{kh}\sigma_{jk}\sigma_{ih}.$$

Equation (3) follows from

$$\sum_{i=1}^{p} \sum_{j=1}^{p} \sum_{h=1}^{p} \sum_{k=1}^{p} [a_{ij}\rho_{hk} + a_{ij}\rho_{kh}]\sigma_{jk}\sigma_{ih}$$

$$= \sum_{i=1}^{p} \sum_{j=1}^{p} \sum_{h=1}^{p} \sum_{k=1}^{p} a_{ij}\rho_{hk}\sigma_{jk}\sigma_{ih} + \sum_{i=1}^{p} \sum_{j=1}^{p} \sum_{h=1}^{p} \sum_{k=1}^{p} a_{ij}\rho_{kh}\sigma_{jk}\sigma_{ih}$$

$$= \sum_{i=1}^{p} \sum_{j=1}^{p} \sum_{h=1}^{p} \sum_{k=1}^{p} a_{ij}\rho_{hk}\sigma_{jk}\sigma_{ih} + \sum_{i=1}^{p} \sum_{j=1}^{p} \sum_{h=1}^{p} \sum_{k=1}^{p} a_{ij}\rho_{hk}\sigma_{jh}\sigma_{ik}$$

$$= \sum_{i=1}^{p} \sum_{j=1}^{p} \sum_{h=1}^{p} \sum_{k=1}^{p} a_{ij}\rho_{hk}[\sigma_{jk}\sigma_{ih} + \sigma_{jh}\sigma_{ik}].$$

PROOF OF PROPOSITION. To see that the matrix is symmetric, consider the (r-1)p + s row and (v-1)p + w column of  $[\Sigma \otimes \Sigma] [I_{p^2} + T]$  which is

$$\begin{split} [e_r \otimes e_s]'[\Sigma \otimes \Sigma] \left[ I_{p^2} + T \right] \left[ e_v \otimes e_w \right] &= \left[ e_r \otimes e_s \right]'[\Sigma \otimes \Sigma] \left[ e_v \otimes e_w \right] + \left[ e_r \otimes e_s \right]'[\Sigma \otimes \Sigma] \left[ e_w \otimes e_v \right] \\ &= \left[ e_r' \Sigma e_v \otimes e_s' \Sigma e_w \right] + \left[ e_r' \Sigma e_w \otimes e_s' \Sigma e_v \right] \\ &= \sigma_{rv} \sigma_{sw} + \sigma_{rw} \sigma_{sv} \end{split}$$

This equals the (v-1)p + w row and (r-1)p + s column of  $[\Sigma \otimes \Sigma] [I_{p^2} + T]$  which is

$$[e_v \otimes e_w]'[\Sigma \otimes \Sigma] [I_{p^2} + T] [e_r \otimes e_s] = \sigma_{vr}\sigma_{ws} + \sigma_{wr}\sigma_{vs} = \sigma_{rv}\sigma_{sw} + \sigma_{rw}\sigma_{sv}$$

because  $\Sigma$  is symmetric.

The matrix is nonnegative definite if for any matrix *P*,

$$\operatorname{Vec}(P)'[\Sigma \otimes \Sigma][I_{p^2} + T]\operatorname{Vec}(P) \ge 0.$$
(5)

This occurs if and only if

$$\operatorname{Vec}(P)'[\Sigma \otimes \Sigma]\operatorname{Vec}(P) \ge |\operatorname{Vec}(P)'[\Sigma \otimes \Sigma]T\operatorname{Vec}(P)| = |\operatorname{Vec}(P)'[\Sigma \otimes \Sigma]\operatorname{Vec}(P')|.$$

However, by Cauchy-Schwartz,

$$\left\{\operatorname{Vec}(P)'[\Sigma \otimes \Sigma]\operatorname{Vec}(P)\right\} \left\{\operatorname{Vec}(P')'[\Sigma \otimes \Sigma]\operatorname{Vec}(P')\right\} \ge \left\{\operatorname{Vec}(P)'[\Sigma \otimes \Sigma]\operatorname{Vec}(P')\right\}^2$$

where equality can only occur if P = P' or  $\Sigma$  is singular. Shortly we will demonstrate that

$$\operatorname{Vec}(P)'[\Sigma \otimes \Sigma]\operatorname{Vec}(P) = \operatorname{Vec}(P')'[\Sigma \otimes \Sigma]\operatorname{Vec}(P')$$
(6)

which gives us

$$\left\{\operatorname{Vec}(P)'[\Sigma \otimes \Sigma]\operatorname{Vec}(P)\right\}^2 \ge \left\{\operatorname{Vec}(P)'[\Sigma \otimes \Sigma]\operatorname{Vec}(P')\right\}^2$$

and

$$\operatorname{Vec}(P)'[\Sigma \otimes \Sigma]\operatorname{Vec}(P) \ge |\operatorname{Vec}(P)'[\Sigma \otimes \Sigma]\operatorname{Vec}(P')|$$

as desired. Since equality can only occur when P is symmetric or  $\Sigma$  is singular, when  $\Sigma$  is positive definite, the inequality (5) is strict unless P = 0.

To see equation (6), apply equation (4) to get

$$\operatorname{Vec}(P)'[\Sigma \otimes \Sigma]\operatorname{Vec}(P) = \sum_{i=1}^{p} \sum_{j=1}^{p} \sum_{h=1}^{p} \sum_{k=1}^{p} \rho_{ij} \rho_{hk} \sigma_{jk} \sigma_{ih}$$

and

$$\operatorname{Vec}(P')'[\Sigma \otimes \Sigma]\operatorname{Vec}(P') = \sum_{i=1}^{p} \sum_{j=1}^{p} \sum_{h=1}^{p} \sum_{k=1}^{p} \rho_{ji}\rho_{kh}\sigma_{jk}\sigma_{ih}$$
$$= \sum_{i=1}^{p} \sum_{j=1}^{p} \sum_{h=1}^{p} \sum_{k=1}^{p} \rho_{ij}\rho_{hk}\sigma_{ih}\sigma_{jk}.$$

## 3 Applications to Tarpey et al. (2014)

Assume that the p + 1 dimensional vectors  $(y_i, x_{i1}, \dots, x_{ip})' \equiv (y_i, x'_i)'$  are independent and

$$\begin{bmatrix} y_i \\ x_i \end{bmatrix} \sim N_p \left( \begin{bmatrix} \mu_y \\ \mu_x \end{bmatrix}, \begin{bmatrix} \sigma_y^2 & \Psi_{yx} \\ \Psi_{xy} & \Psi_x \end{bmatrix} \right).$$

Write,

$$Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}; \qquad X = \begin{bmatrix} x'_1 \\ \vdots \\ x'_n \end{bmatrix}.$$

Consider a linear model conditional on X

$$Y = \alpha J + X\beta + e, \quad E(e) = 0, \quad \operatorname{Cov}(e) = \sigma^2 I_n$$

where *J* is a column of ones,  $\beta$  is a solution to

 $\Psi_x \beta = \Psi_{xy},$  $\alpha = \mu_y - \mu'_x \beta,$ 

and

$$\sigma^2 = \sigma_y^2 - \beta' \Psi_x \beta.$$

Also write

$$Y_c \equiv [I - (1/n)JJ']Y;$$
  $X_c \equiv [I - (1/n)JJ']X$ 

and the usual unbiased estimates of the variance and covariance parameters as

$$\begin{bmatrix} s_y^2 & S_{yx} \\ S_{xy} & S_x \end{bmatrix} = \frac{1}{n-1} \begin{bmatrix} Y_c'Y_c & Y'X_c \\ X_c'Y & X_c'X_c \end{bmatrix}.$$

The usual least squares estimate of  $\beta$  is

$$\hat{\beta} = S_x^{-1} S_{xy} = (X_c' X_c)^{-1} X_c' Y.$$

Tarpey et al. examine the estimate

$$\tilde{\beta} = \Psi_x^{-1} S_{xy} = \Psi_x^{-1} \frac{1}{n-1} X_c' Y.$$

It is relatively easy to see that the covariance matrix of  $\hat{\beta}$  is

$$\operatorname{Cov}(\hat{\beta}) = \frac{\sigma_y^2 - \beta' \Psi_x \beta}{n - p - 2} \Psi_x^{-1} = \frac{\sigma_y^2 \Psi_x^{-1} - \beta' \Psi_x \beta \Psi_x^{-1}}{n - p - 2}$$

(especially after seeing the derivation below). We establish at the end of this section that

$$\operatorname{Cov}(\tilde{\beta}) = \frac{1}{n-1} \left( \sigma_y^2 \Psi_x^{-1} + \beta \beta' \right).$$
(1)

Clearly, when n >> p, the variability of  $\tilde{\beta}$  will exceed that of  $\hat{\beta}$ .

The incorrect covariance matrix for  $\tilde{\beta}$  given (by me) in Tarpey et al. (2014) is actually an upper bound on  $\text{Cov}(\tilde{\beta})$ .

$$\frac{\sigma_y^2 + \Psi_{xy}' \Psi_x^{-1} \Psi_{xy}}{n-1} \Psi_x^{-1} \ge \frac{1}{n-1} \left( \sigma_y^2 \Psi_x^{-1} + \Psi_x^{-1} \Psi_{xy} \Psi_{xy}' \Psi_x^{-1} \right) = \frac{1}{n-1} \left( \sigma_y^2 \Psi_x^{-1} + \beta \beta' \right)$$

because

$$(\Psi'_{xy}\Psi_x^{-1}\Psi_{xy})\Psi_x^{-1} \ge \Psi_x^{-1}\Psi_{xy}\Psi'_{xy}\Psi_x^{-1}$$

which is true because for any vector *c*, by Cauchy-Schwartz,

$$(\Psi'_{xy}\Psi_x^{-1}\Psi_{xy})c'\Psi_x^{-1}c \ge c'\Psi_x^{-1}\Psi_{xy}\Psi'_{xy}\Psi_x^{-1}c = \left(\Psi'_{xy}\Psi_x^{-1}c\right)^2.$$

To establish (1) note that

$$Cov(\tilde{\beta}) = \frac{1}{(n-1)^2} \Psi^{-1} Cov(X'_c Y) \Psi^{-1}.$$
 (2)

We need to find

$$\operatorname{Cov}(X_c'Y) = E\left[\operatorname{Cov}(X_c'Y|X)\right] + \operatorname{Cov}\left[E(X_c'Y|X)\right]$$
$$= E\left[\sigma^2 X_c'X_c\right] + \operatorname{Cov}\left[X_c'X_c\beta\right].$$
(3)

Recall that  $X_c' X_c \sim W_p(n-1, \Psi_x)$ , so that

$$E(X'_{c}X_{c}) = (n-1)\Psi_{x}; \qquad \operatorname{Cov}\{\operatorname{Vec}[(X'_{c}X_{c})]\} = (n-1)(\Psi_{x} \otimes \Psi_{x})\{I_{p^{2}} + T\}.$$

It follows that

$$E\left\{\sigma^2 X'_c X_c\right\} = (n-1)[\sigma_y^2 - \beta' \Psi_x \beta] \Psi_x \tag{4}$$

and

$$\operatorname{Cov}[X'_{c}X_{c}\beta] = \operatorname{Cov}\{[\beta' \otimes I_{p}]\operatorname{Vec}(X'_{c}X_{c})\}\$$

$$= [\beta' \otimes I_{p}]\operatorname{Cov}\{\operatorname{Vec}(X'_{c}X_{c})\}[\beta \otimes I_{p}]\$$

$$= (n-1)[\beta' \otimes I_{p}][\Psi_{x} \otimes \Psi_{x}][\beta \otimes I_{p}] + (n-1)[\beta' \otimes I_{p}][\Psi_{x} \otimes \Psi_{x}]T[\beta \otimes I_{p}]\$$

$$= (n-1)(\beta'\Psi_{x}\beta)\Psi_{x} + (n-1)[\beta'\Psi_{x} \otimes \Psi_{x}]T[\beta \otimes I_{p}].$$
(5)

We will show below that

$$[\beta'\Psi_x \otimes \Psi_x]T[\beta \otimes I_p] = \Psi_x \beta \beta' \Psi_x, \tag{6}$$

so substituting (6) into (5), then (5) and (4) into (3), and then (3) into (2) gives (1).

To show that

$$[\beta'\Psi_x \otimes \Psi_x]T[\beta \otimes I_p] = \Psi_x \beta \beta' \Psi_x,$$

we show that for any r and s

$$e'_r[\beta'\Psi_x\otimes\Psi_x]T[\beta\otimes I_p]e_s = e'_r\Psi_x\beta\beta'\Psi_xe_s.$$

Write  $\Psi_x = [\psi_{x1}, \cdots, \psi_{xp}]$ , then

$$e_r'\Psi_x\beta\beta'\Psi_xe_s=\psi_{xr}'\beta\beta'\psi_{xs}.$$

On the other hand,

$$e'_{r}[\beta'\Psi_{x}\otimes\Psi_{x}]T[\beta\otimes I_{p}]e_{s} = [1\otimes e_{r}]'[\beta'\Psi_{x}\otimes\Psi_{x}]T[\beta\otimes I_{p}][1\otimes e_{s}]$$
$$= [\beta'\Psi_{x}\otimes e'_{r}\Psi_{x}]T[\beta\otimes e_{s}]$$
$$= [\beta'\Psi_{x}\otimes\psi'_{xr}][e_{s}\otimes\beta]$$
$$= [\beta'\Psi_{x}e_{s}\otimes\psi'_{xr}\beta]$$
$$= [\beta'\Psi_{x}e_{s}\otimes\psi'_{xr}\beta] = \beta'\psi_{xs}\psi'_{xr}\beta.$$

## **4** Equivalence of Inequalities

We now demonstrate that inequality (4) in Cook et al. (2015) and inequality (1) in Tarpey et al.'s response are equivalent.

$$\frac{1+\mathcal{R}_p^2}{n-1} < \frac{1-\mathcal{R}_p^2}{n-p-2}$$

if and only if

$$2\mathcal{R}_p^2 n < (1 + \mathcal{R}_p^2)(p+2) - (1 - \mathcal{R}_p^2)$$

if and only if

$$2\mathcal{R}_p^2 n < (1+\mathcal{R}_p^2)(p+1) + 2\mathcal{R}_p^2$$

if and only if

$$2\mathcal{R}_p^2 n < (1 - \mathcal{R}_p^2)(p+1) + 2\mathcal{R}_p^2(p+1) + 2\mathcal{R}_p^2$$

if and only if

$$n < \left(\frac{1 - \mathcal{R}_p^2}{\mathcal{R}_p^2}\right) \left(\frac{p+1}{2}\right) + p + 2$$

as was to be shown.

## Appendix

Kronecker products and Vec operators are extremely useful in multivariate analysis and some approaches to variance component estimation. They are also often used in writing balanced ANOVA models. We now present their basic algebraic properties.

- 1. If the matrices are of conformable sizes,  $[A \otimes (B + C)] = [A \otimes B] + [A \otimes C]$ .
- 2. If the matrices are of conformable sizes,  $[(A + B) \otimes C] = [A \otimes C] + [B \otimes C]$ .
- 3. If *a* and *b* are scalars,  $ab[A \otimes B] = [aA \otimes bB]$ .
- 4. If the matrices are of conformable sizes,  $[A \otimes B][C \otimes D] = [AC \otimes BD]$ .
- 5. The transpose of a Kronecker product matrix is  $[A \otimes B]' = [A' \otimes B']$ .
- 6. The generalized inverse of a Kronecker product matrix is  $[A \otimes B]^- = [A^- \otimes B^-]$ .
- 7. For two vectors v and w,  $\operatorname{Vec}(vw') = w \otimes v$ .

- 8. For a matrix *W* and conformable matrices *A* and *B*,  $Vec(AWB') = [B \otimes A]Vec(W)$ .
- 9. For conformable matrices *A* and *B*, Vec(A)'Vec(B) = tr(A'B).
- 10. The Vec operator commutes with any matrix operation that is performed elementwise. For example,  $E{\operatorname{Vec}(W)} = \operatorname{Vec}{E(W)}$  when *W* is a random matrix. Similarly, for conformable matrices *A* and *B* and scalar  $\phi$ ,  $\operatorname{Vec}(A + B) = \operatorname{Vec}(A) + \operatorname{Vec}(B)$  and  $\operatorname{Vec}(\phi A) = \phi \operatorname{Vec}(A)$ .
- 11. If *A* and *B* are positive definite, then  $A \otimes B$  is positive definite.

#### References

- Christensen, Ronald (2001). Advanced Linear Modeling: Multivariate, Time Series, and Spatial Data; Nonparametric Regression and Response Surface Maximization, Second Edition. Springer, New York.
- Christensen, Ronald (2011). Plane Answers to Complex Questions: The Theory of Linear Models. Springer, New York.
- Cook, R. Dennis, Forzani, Liliana, and Rothman, Adam J. (2015). "Comment on Tarpey et al. (2014)." *The American Statistician*, **69**, ?
- Eaton, Morris L. (1983). *Multivariate Statistics: A Vector Space Approach*. John Wiley and Sons, New York. Reprinted in 2007 by IMS Lecture Notes Monograph Series.
- Tarpey, Thaddeus, Ogden, R. Todd, Petkova, Eva, and Christensen, Ronald (2014). "A Paradoxical Result in Estimating Regression Coefficients." *The American Statistician*, **68**, 271-276.