

More on the orthogonal complement functions.

Robert Jennrich^{1*} and Albert Satorra²

¹Department of Mathematics, University of California, Los Angeles, USA

²Department of Mathematics, Universitat Pompeu Fabra, Barcelona, Spain

Abstract

Continuous orthogonal complement functions have had an interesting history in covariance structure analysis. They were used in a seminal paper by Browne in his development of a distribution-free goodness of fit test for an arbitrary covariance structure.

The proof of his main result Proposition 4 used a locally continuous orthogonal complement function, but because he failed to show such functions existed his proof was incomplete. In spite of the fact that his test had been used extensively, this problem was not noticed until 2013 when Jennrich and Satorra pointed out that his proof was incomplete and completed it by showing that locally continuous orthogonal complement functions exist. This was done using the implicit function theorem. A problem with the implicit function approach is that it does not give a formula for the locally continuous function produced. This problem was potentially solved by Browne and Shapiro who gave a very simple formula $F(X)$ for an orthogonal complement of X .

Unfortunately, they failed to prove that their function actually produced orthogonal complements. We will prove that given a $p \times q$ matrix X_0 with full column rank $q < p$, then $F(X)$ is an orthogonal complement of X for all X sufficiently close to X_0 . Much more is true. We will give a necessary and sufficient condition on X under which $F(X)$ is an orthogonal complement of X . We will also show the X for which $F(X)$ is not an orthogonal complement of X is a subset with Lebesgue measure zero in $R^{p \times q}$.

Keywords: Covariance structure analysis, Distribution free tests, Implicit function theorem, QR factorization.

Accepted on February 23, 2019

Introduction

The orthogonal complement of a $p \times q$ matrix X with $q < p$ and full column rank is a $p \times (p-q)$ matrix Y such that $[X, Y]$ is invertible. Note that Y must have full column rank. Jennrich and Satorra in Theorem 1 show how to compute an orthogonal complement Y of an arbitrary $p \times q$ matrix X with full column rank $q < p$ using the long form of a QR factorization. Unfortunately, Y is not a continuous function of X [1].

In a seminal paper Browne showed how to test the goodness of fit of an arbitrary covariance structure [2]. The proof of his main result Proposition 4 used a locally continuous orthogonal complement function, but because he failed to show such functions existed his proof was incomplete. In spite of the fact that his test had been used extensively this problem was not noticed until 2013 when Jennrich and Satorra pointed out that his proof was incomplete and completed it by showing that locally continuous orthogonal complement functions exist [1]. This was done using the implicit function theorem. A problem with the implicit function approach is that it does not give a formula for the function produced and Jennrich and Satorra conjectured their function could not be expressed by an explicit formula.

Browne and Shapiro say that this conjecture is incorrect by presenting an explicit formula for a locally continuous orthogonal complement function [3]. Browne and Shapiro in

slightly different notation state the following: Let X_0 be any $p \times q$ matrix with full column rank $q < p$ and X_0^c be any orthogonal complement of X_0 . Consider the following matrix valued function:

$$F(X) = \left[I_p - X(XX')^{-1}X' \right] X_0^c$$

Let N be any neighbourhood of X_0 such that X has full column rank for all $X \in N$. Browne and Shapiro say that on N , $F(X)$ is a well-defined continuous function and claim $F(X)$ is an orthogonal complement of X for all X in N . They, however, fail to show this. In particular they fail to show $F(X)$ has full column rank. As a consequence their claim is only a conjecture. We will begin by proving their conjecture.

Proof of the Browne-Shapiro conjecture

Lemma 1: Let X_0 be a $p \times q$ matrix of full column rank $q < p$. There is a neighbourhood N of X_0 that contains only full column rank matrices.

Proof: Let X be an arbitrary $p \times q$ matrix and let $g(X) = \det(XX')$. Since $g(X)$ is continuous there is $\delta > 0$ such that

$$\|X - X_0\| < \delta \Rightarrow \|g(X) - g(X_0)\| < \|g(X_0)\|$$

It follows from the last inequality that $g(X) \neq 0$ and hence that $\det(XX') \neq 0$ and hence that X has full column rank for all X in the neighbourhood

$$N = \{X : \|X - X_0\| < \delta\}$$

Lemma 2: Let X_0 be a $p \times q$ matrix with full column rank $q < p$ and let X_0^c be any orthogonal complement of X_0 . Then

$$F(X) = [I_p - X(X'X)^{-1}X']X_0^c$$

is continuous and well defined for $X \in N$.

Proof: By Lemma 1 all $X \in N$ have full column rank. Thus $F(X)$ is well defined. It follows from the continuity of matrix multiplication and matrix inversion that $F(X)$ is continuous for all $X \in N$.

Theorem 1: Let X_0 be a $p \times q$ matrix with full column rank $q < p$. Let X_0^c be any orthogonal complement of X_0 . Then there is a neighbourhood N of X_0 such that X has full column rank for all $X \in N$ and

$$F(X) = [I - X(X'X)^{-1}X']X_0^c$$

is an orthogonal complement of X .

Proof: Note that $F(X)$ is $p \times (p-q)$ and $F(X)$ is orthogonal to X . It is sufficient to prove $F(X)$ has full column rank. Let $g(X) = \det(F(X)'F(X))$. Then $g(X)$ is a continuous and it follows from this that there is $\delta > 0$ such that

$$\|X - X_0\| < \delta \Rightarrow \|g(X) - g(X_0)\| < |g(X_0)|$$

It follows from the last inequality that $g(X) \neq 0$. Hence $\det(F(X)'F(X)) \neq 0$ and hence $F(X)$ has full column rank.

This proof is much simpler than that given by Jennrich and Satorra using the implicit function theorem [1]. In their paper Browne and Shapiro prove Browne's Proposition 4 without using their $F(X)$ formula. One could, however, use Theorem 1 because it asserts the existence of a locally continuous orthogonal complement function and this is all that is needed to fix Browne's proof of Proposition 4. There are now three proofs of Browne's Proposition 4, the one given by Jennrich and Satorra [1], the one given by Browne and Shapiro [3], and the one using Theorem 1.

A problem with using Theorem 1 is that for an X of interest one has no way of knowing if X is in N since the only thing we know about δ is that it is greater than zero. If necessary for an X of interest one can always compute an orthogonal complement of X using Theorem 1 of Jennrich and Satorra, but it will not be a continuous function of X [1].

See the following example which shows numerically that Theorem 1 of Jennrich and Satorra produces orthogonal complements as does Theorem 1 of this document and also suggests an interesting alternative method using $F(X)$ [1].

Theory:

Example:

Let

$$X = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Then X is a p by q matrix with $p = 4$ and $q = 2$. Theorem 1 of Jennrich and Satorra [1] says that if

$$X = QR$$

is the long form of a QR factorization of X , then the last $p-q$ columns of Q are an orthogonal complement of X . Computing a QR factorization of X gives

$$Q = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

The last $p-q$ columns of this are:

$$Y = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and this is an orthogonal complement X . This demonstrates numerically that Theorem 1 of Jennrich and Satorra [1] produces orthogonal complements.

Note that when $X = X_0$, X satisfies the assumptions of Theorem 1 and the computed value of

$$F(X) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

which is an orthogonal complement of X . Thus in this case at least $F(X)$ is an orthogonal complement of X as asserted by Theorem 1.

Let

$$X = \begin{pmatrix} 1.1650 & -0.6965 \\ 0.6268 & 1.6961 \\ 0.0751 & 0.0591 \\ 0.3516 & 1.7971 \end{pmatrix}$$

be a random matrix whose components are independent standard normal variables. Note that it has full column rank. The value of the Browne-Shapiro.

$$F(X) = \begin{pmatrix} -0.0418 & 0.1114 \\ -0.0304 & -0.4870 \\ 0.9969 & -0.0203 \\ -0.0203 & 0.5035 \end{pmatrix}$$

This has full column rank and hence is an orthogonal complement of X even though X may not satisfy the basic assumption of Theorem 1 above because X is simply a random matrix.

When this is repeated 1000 times in every case $F(X)$ is an orthogonal complement of X . This suggests that $F(X)$ is an orthogonal complement of X with high probability.

We show below that much more is true. We show $F(X)$ fails to be an orthogonal complement of X only for X in subset of Lebesgue measure zero in $R^{p \times q}$.

Note that when $X = X_0^c$, $F(X) = 0$ and hence $F(X)$ is not an orthogonal complement of X for all X with full column rank.

A necessary and sufficient condition for $F(X)$ to be an orthogonal complement of X .

To simplify the development that follows let us begin by looking at the simplest possible case when $p = 2$ and $q = 1$. Assume X is

not in the column space of X_0^c . Then in Figure 1 clearly, $F(X)$ is a non-zero vector and has full column rank. Thus $F(X)$ is an orthogonal complement of X for all X not in the column space X_0^c . If X is in the column space of X_0^c , then $F(X)=0$ and hence $F(X)$ is not an orthogonal complement of X . Note this happens only on a set that has Labesgue measure zero in R^2 . We will show that this happens for arbitrary p by $q < p$ matrices X .

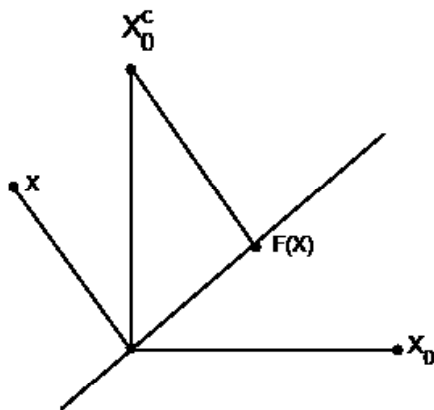


Figure 1. Showing an orthogonal complement of X .

We begin with the following theorem,

Theorem 2: Let $F(X)$ be the Browne-Shapiro formula. That $[X, X_0^c]$ is invertible is a necessary and sufficient condition for $F(X)$ to be an orthogonal complement of X .

Proof: Assume $[X, X_0^c]$ has full column rank. Assume $F(X)$ does not have full column rank. Then

$$F(X)a=0$$

for some vector $a \neq 0$. Thus

$$(I - X(X'X)^{-1}X')X_0^c a = 0$$

It follows that $X_0^c a$ is in the column space of X . Thus

$$X_0^c a = Xb$$

for some vector b . Thus there is a vector in the column space of X that is in the column space of X_0^c . It follows that $[X, X_0^c]$ does not have full column rank. This contradiction proves $F(X)$ has full column rank when $[X, X_0^c]$ has full column rank. This is a sufficient condition for $F(X)$ to be an orthogonal complement of X .

If $[X, X_0^c]$ does not have full column rank, then $Xa = X_0^c b$ for two vectors a and b not both zero. If $a = 0$, then $b = 0$. This implies $a \neq 0$ which in turn implies $b \neq 0$. Now

$$F(X)a = (I - X(X'X)^{-1}X')X_0^c a = (I - X(X'X)^{-1}X')Xb = 0$$

which implies that $F(X)$ does not have full column rank. Thus $F(X)$ is not an orthogonal complement of X when $[X, X_0^c]$ does not have full column rank. Thus $F(X)$ is an orthogonal complement of X if and only if $[X, X_0^c]$ has full column rank.

How often does $F(X)$ fail to be an orthogonal complement of X ?

Lemma 2: A polynomial in n variables is either identically 0 or its roots have Labesgue measure zero in R^n .

Proof: The proof is by mathematical induction on n . For $n=1$, $P(x)$ of degree d can have at most d roots, which gives us the base case. Now, assume that the theorem holds for all polynomials in $n-1$ variables. Let $\tilde{x}_2 = (x_2, \dots, x_n)$ and write

$$P(x) = P(x_1, \tilde{x}_2)$$

Let $I[A]$ be the indicator function of a set A . By Fubini's theorem,

$$\begin{aligned} \int I[P(x) = 0] dx &= \iint I[P(x_1, \tilde{x}_2) = 0] dx_1 d\tilde{x}_2 \\ &= \iint I[P(x_1, \tilde{x}_2) = 0] d\tilde{x}_2 dx_1 \\ &= \int (0) dx_1 = 0 \end{aligned}$$

The first two integral equalities follow from Fubini's Theorem. The last integral equality follows from the fact that given x_1 , $P(x_1, \tilde{x}_2)$ is a polynomial in $n-1$ variables and by the induction hypothesis this is identically zero for all \tilde{x}_2 or its zeros have Labesgue measure zero in R^{n-1} [4].

Comment: Lemma 2 is a simple very interesting result which apparently is not in the literature. It is not in any journals covered in Jstor. It is not in any of the five analysis books we own and according to Caron and Traynor it is not in any measure theory book.

Theorem 3: The X for which the Browne - Shapiro function $F(X)$ is not an orthogonal complement of X have Labesgue measure zero in $R^{p \times q}$.

Proof: $F(X)$ is not an orthogonal complement of X if and only if $[X, X_0^c]$ is singular. This happens if and only if $\det([X, X_0^c]'[X, X_0^c]) = 0$. As a function of X this is not identically zero because it is not zero when $X=X_0$. Moreover $\det([X, X_0^c]'[X, X_0^c])$ is a polynomial in $p \times q$ variables $x = \text{vec}(X)$ which is not identically zero. It follows from Lemma 2 that its zeros have Lebesgue measure zero in R^{pq} . Or equivalently the X such that $F(X)$ is not an orthogonal complement of X have Labesgue measure zero in $R^{p \times q}$.

Remark: Theorem 3 implies that $F(X)$ almost never fails to be an orthogonal complement of X . If X is a sample from a density, then $F(X)$ is an orthogonal complement of X with probability one.

Conclusion

Browne and Sharpiro [3] give a formula for computing orthogonal complements. Let X_0 be any $p \times q$ matrix with full column rank $q < p$ and X_0^c be any orthogonal complement of X_0 . Browne and Sharpiro function is

$$F(X) = [I_p - X(X'X)^{-1}X']X_0^c$$

They state that if N be any neighbourhood of X_0 such that X has full column rank for all $X \in N$ then $F(X)$ is an orthogonal complement of X for all X in N . They, however, fail to show this. In particular they fail to show $F(X)$ has full column rank. As a consequence their claim is only a conjecture. The main point of our paper is to prove their conjecture and much more. We prove that on all but a set of $p \times q$ matrices X with Labesque measure zero, $F(X)$ is an orthogonal complement of X . In particular if X

is a sample from a density, $F(X)$ is an orthogonal complement of X with probability one.

Theorem 1 is a proof for the Browne-Shapiro conjecture. We also prove additional results that should enhance considerably the practical relevance of the Browne-Shapiro function for an orthogonal complement. Theorem 2 gives a necessary and sufficient condition for $F(X)$ to be an orthogonal complement of X . Lemma 2 is a not well known, but very useful result about the roots of an arbitrary polynomial in n variables. If the polynomial is not identically zero, its roots have Lebesgue measure zero. Theorem 3 is our main result.

Acknowledgements

The authors would like to thank Ellen Jennrich for proof reading our paper and helping with its submission. The research of the second author is supported by grant EC02014-59885-P from the Spanish Ministry of Science and Innovation.

References

1. Jennrich RI, Satorra A. Continuous orthogonal complement

functions and distribution free goodness of fit tests for covariance structure analysis. *Psychometrika*. 2013;78: 445-552.

2. Browne MW. Asymptotically distribution-free methods for the analysis of covariance structures. *Psychometrika*. 1984;37: 62-83.
3. Browne MW, Shapiro A. A comment on the asymptotics of a distribution free goodness of fit test statistic. *Psychometrika*. 2013;78: 196-9.
4. Caron R, Traynor T. The zero set of a polynomial. 2005.

***Correspondence to:**

Robert Jennrich
Department of Mathematics
University of California
Los Angeles, USA
Tel: (310) 825-2207
E-mail: rij@stat.ucla.edu