

# Geometric Views of Partial Correlation Coefficient in Regression Analysis

Bilin Zeng<sup>1</sup>, Kang Chen<sup>2</sup> & Cong Wang<sup>3</sup>

<sup>1</sup> Department of Mathematics, California State University, Bakersfield, California, United States

<sup>2</sup> Lee Kuan Yew School of Public Policy, National University of Singapore, Singapore

<sup>3</sup> Department of Finance, China Europe International Business School, Shanghai, China

Correspondence: Bilin Zeng, Department of Mathematics, California State University, Bakersfield, California, United States. Tel: 661-654-6519. E-mail: bzeng@csub.edu

Received: March 10, 2017 Accepted: March 31, 2017 Online Published: April 11, 2017

doi:10.5539/ijsp.v6n3p51

URL: <https://doi.org/10.5539/ijsp.v6n3p51>

## Abstract

By describing the geometric analogues of the concepts from various perspectives, this work aims to provide a richer and intuitive comprehension of the concept of partial correlation coefficient in the regression analysis, especially for beginning students. Based on a simple and strictly correct geometric framework, this article geometrically illustrates the concept of partial correlation coefficient in regression analysis from the views of the Frisch-Waugh-Lovell Theorem, partial F test statistics, and the comparisons with other levels of correlation coefficients. In our opinion, the geometric approach sheds lights on the regression analysis as it provides a richer and more concrete understanding for readers, especially for beginners. This paper can also be served as a supplementary reading material for serious beginners.

**Keywords:** Frisch-Waugh-Lovell Theorem, Geometry Interpretation, Partial Correlation Coefficient, Partial F test, Regression Analysis

## 1. Introduction

Since the subject of regression emerged in the late of 19<sup>th</sup> century, algebra has been widely used to express concepts and build up models in regression analysis. Most concepts in regression analysis are traditionally introduced in terms of algebraic equations and matrices. Projection arguments in the form of matrix algebras (Graybill, 1976; Kutner, Nachtsheim & Neter, 2004) are widely used as educational and research tools for high-dimensional modelling in the advanced studies of multiple regression analysis. However, sometimes, it is challenging for non-math majors or beginners to completely understand the matrix algebra approaches as it requires the solid pre-knowledge of linear algebra for readers. In fact, geometric interpretation is in reality more helpful than cumbersome algebraic equations and matrices in understanding regression concepts because its visual presentation is concrete (Margolis, 1979; Bring, 1996). An understanding of the geometrical aspects of elementary regression analysis may assist a student more effectively than elegantly derived formulas (Saville & Wood, 1986).

Despite its merits, geometry is seldom used in teaching regression analysis courses. Besides the long-time predominance of algebra since the 20<sup>th</sup> century, the resistance to abstraction is one of the primary reasons. To reduce the beginners' fear of abstraction, an intuitive display of regression model in a visible three-dimensional space is the key to open the door of geometric thinking. A multiple regression problem with two predictors is an ideal motivating example for introducing the geometric thinking of regression modeling to beginners as its layout perfectly fits into a visible space. Furthermore, it serves as an important step stone for understanding high dimensional linear regression models since all these basic results can be easily extended for a more general linear regression mode. By exploring vectors, triangles and projections, and drawing them clearly in such three-dimensional space, students do not have to delve into complicated algebraic calculations nor advanced matrix algebras.

Some classical textbooks introduce the simple geometric illustration of basic concepts in regression analysis including the least-square estimation and the simple correlation coefficient (Draper & Smith, 2014). But the geometrical interpretation of other important concepts such as the partial correlation coefficient, partial regression coefficient and partial test statistic are often left out. In fact, students usually experience difficulty in understanding these concepts in studies. The visualized geometric representation will be of great help for understanding the partial correlation coefficient and its related concepts. The views from different perspectives will provide richer understanding for students. Based on a simple and strictly correct  $E^3$ , the objective of this paper is to geometrically illustrate the concept of partial correlation coefficient from various perspectives. The rest of the paper is organized as follows. Section 2 first corrects a graphing problem in the

previous research by mapping the original  $n$ -dimensional space  $E^n$  to a  $E^3$  one without the loss of information. Section 3 geometrically interprets the partial correlation coefficients from the idea of the Frisch-Waugh-Lovell Theorem (Frisch & Waugh, 1933; Lovell, 1963), verifies its relationship with other types of correlation coefficients, and reveals its connection with the partial test statistics. Finally, conclusions and a brief discussion are given in section 4. Those who have finished an introductory course of econometrics or statistics would find it especially helpful when they examine these visual analogues sketched out in the paper.

## 2. Drawing Graphs on $E^3$

Consider a multiple regression model with two independent variables as follows:

$$v = \beta_1 u_1 + \beta_2 u_2 + w \quad (1)$$

where  $v$  contains  $n$  observations of response,  $u_1$  and  $u_2$  are two independent predictors,  $\beta_1 \in R^1$  and  $\beta_2 \in R^1$  are the regression coefficients,  $w$  is the error term, and  $n$  is the size of the sample. In geometry, each regression variable is considered as a vector in an  $n$ -dimensional space. For example,  $\vec{v} = (v_1, v_2, \dots, v_n)^T$  is the  $n$ -dimensional observation vector, two independent predictors  $\vec{u}_1$  and  $\vec{u}_2$  are both  $n$ -dimensional column vectors,  $\vec{w}$  is also a column vector in  $n$ -dimensional Euclidean space,  $E^n$ . For the ease of demonstration, in Figure 1, it is common practice to have an  $n$ -dimensional vector displayed in a 3-dimensional vector space (Bring, 1996). In Figure 1,  $\vec{v}$  (or  $\overrightarrow{OO'}$ ), is the orthogonal projection of  $\vec{v}$  on the plane spanned by  $\vec{u}_1$  and  $\vec{u}_2$  that denoted by  $Span(\vec{u}_1, \vec{u}_2)$ . Vector  $\overrightarrow{OA}$  is the orthogonal projection of  $\vec{v}$  on  $\vec{u}_1$ , and vector  $\overrightarrow{OB}$  is the perpendicular projection of  $\vec{v}$  on  $\vec{u}_2$ .

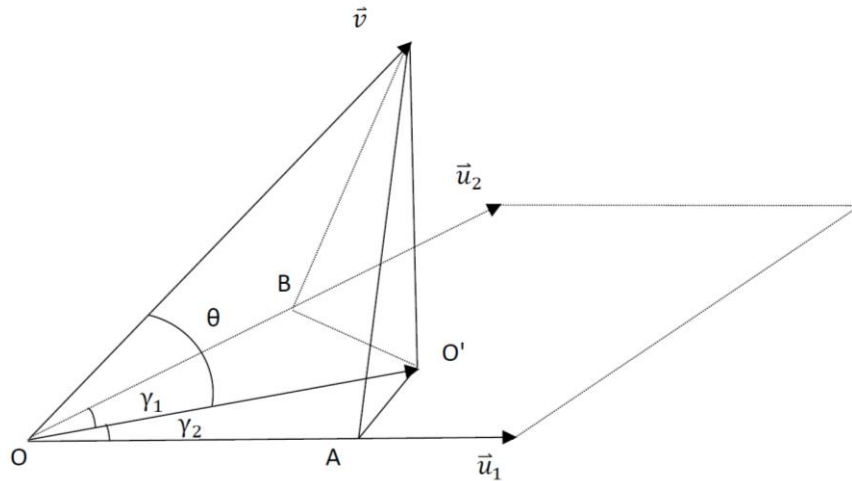


Figure 1. Geometric Interpretation of Least Squares Method in  $n$ -space.

However, the problem with Figure 1 is that it is displaying in a 3-dimensional space which contradicts with the fact that  $\vec{v}$ ,  $\vec{u}_1$  and  $\vec{u}_2$  are vectors in  $n$ -dimensional space  $E^n$ . According to Saville and Wood (1991), this demonstration is not strictly correct as in higher dimensions ( $n > 3$ ), vectors cannot be shown pictorially in a strictly correct manner. To solve this contradiction, we establish a transformation matrix  $A$  to map the  $n$ -dimensional vectors  $\vec{v}$ ,  $\vec{u}_1$  and  $\vec{u}_2$  onto the 3-dimensional space. This linear transformation keeps:

- The length of any vector unchanged;
- The angle between any two vectors unchanged.

After pre-multiplying  $A$  with both sides of the equation (1), it leads to the following model

$$y = \beta_1 x_1 + \beta_2 x_2 + \varepsilon \quad (2)$$

where  $y$ ,  $x_1$ , and  $x_2$  are vectors in the 3-dimensional Euclidean space  $E^3$ . At the same time, in the original  $n$ -dimensional space,  $\vec{v}$ ,  $\vec{u}_1$  and  $\vec{u}_2$  span a 3-dimensional subspace of  $E^n$ . This subspace has the same dimensions as  $E^3$ . On the other hand, we know that “any two finite Euclidean spaces are isomorphic if and only if they have the same number of dimensions.” Therefore, the subspace is isomorphic to  $E^3$ . This isomorphism ensures that the results we obtain from the new  $E^3$  are the same as those we get by analyzing the original  $E^n$ . Furthermore, this process ensures the strictness and correctness of drawing any high dimensional vectors, angles and triangles in a  $E^3$ . In this process, the regression coefficients  $\beta_i$  (for  $i = 1, 2$ ) is unchanged. This idea is summarized as:

Theorem 1: For any 3 linearly independent vectors  $\vec{v}$ ,  $\vec{u}_1$  and  $\vec{u}_2$  in  $E^n$ , there exists an orthogonal transformation  $A$

between  $\text{Span}(\vec{v}, \vec{u}_1, \vec{u}_2)$  and  $E^3$ , such that,

$$A\vec{u}_1 = \begin{pmatrix} |\vec{u}_1| \\ 0 \\ 0 \end{pmatrix} = \vec{x}_1, \quad A\vec{u}_2 = \begin{pmatrix} |\vec{u}_2| \cos \gamma \\ |\vec{u}_2| \sin \gamma \\ 0 \end{pmatrix} = \vec{x}_2, \quad \text{and} \quad A\vec{v} = \begin{pmatrix} |\vec{v}| \cos \theta \cos \gamma_1 \\ |\vec{v}| \cos \theta \sin \gamma_1 \\ |\vec{v}| \sin \theta \end{pmatrix} = \vec{y}$$

Where

$$A = \begin{pmatrix} \frac{\vec{u}_1^T}{|\vec{u}_1|} \\ \frac{(\vec{u}_2 - \hat{\vec{u}}_{21})^T}{|\vec{u}_2 - \hat{\vec{u}}_{21}|} \\ \frac{(\vec{v} - \hat{\vec{v}})^T}{|\vec{v} - \hat{\vec{v}}|} \end{pmatrix}.$$

In Theorem 1,  $\theta$  is the angle between vector  $\vec{v}$  and  $\hat{\vec{v}}$ ; the angle between vector  $\vec{u}_1$  and  $\hat{\vec{v}}$  is named as  $\gamma_1$ , the angle between  $\vec{u}_1$  and  $\vec{u}_2$  is  $\gamma$ ; and vector  $\hat{\vec{u}}_{21}$  is defined as the predictor of  $\vec{u}_2$  when regressing  $\vec{u}_2$  on  $\vec{u}_1$ . See appendix for the detailed proof. In the following sections, we focus on model (2).

### 3. Geometric Interpretations: From Different Perspectives

#### 3.1 The Basic Idea of Correlation Coefficients

A correlation coefficient is a quantitative measurement of the linear association between two variables of interests. In regression analysis, there are three classical correlation coefficients: the simple, partial and multiple correlation coefficients. In geometry, it is commonly known that the behavior of all three types of correlation coefficients can be expressed in the form of the cosine functions.

First of all, the *simple correlation coefficient* between any two variables can be expressed as the cosine of the angle between the two vectors that represent the variables. For example, the simple correlation coefficient between  $y$  and  $x_1$ , denoted by  $r_{yx_1}$ , can be depicted as the cosine of  $\theta_1$ , where  $\theta_1$  is the angle between  $\vec{x}_1$  and  $\vec{y}$  in Figure 2. It is clear that the magnitude of the correlation coefficient depends on the angle between two variable vectors of interests. The closer one vector is to the other (or its opposite vector), the stronger linear relationship it suggests. Particularly, there is no linear association when two variable vectors are perpendicular to each other.

The *multiple correlation coefficient* between  $y$  and  $x_1$  and  $x_2$ , denoted by  $r_{y, x_1 x_2}$ , is used to measure the goodness-of-fit of for a linear regression model. Similarly, in geometry, it is shown as the cosine of the angle between the response vector  $\vec{y}$  and the estimation space that spanned by  $\vec{x}_1$  and  $\vec{x}_2$ . In another word,

$$r_{y, x_1 x_2} = \pm \frac{|\vec{\hat{y}}|}{|\vec{y}|} = \cos \theta \left( r_{y, x_1 x_2} \geq 0, \text{ if } \theta \leq \frac{\pi}{2} \text{ or } \theta \geq \frac{3\pi}{2}; r_{y, x_1 x_2} < 0 \text{ otherwise} \right)$$

where the least square estimation vector  $\vec{\hat{y}}$  (or  $\overrightarrow{OO'}$ ) is achieved by orthogonally projecting the observation vector  $\vec{y}$  onto the  $x_1 x_2$  plane, and  $\theta$  is the angle between vector  $\vec{y}$  and  $\vec{\hat{y}}$  in Figure 2. In a general linear regression model with  $k$  predictors (where  $k > 2$ ), the multiple correlation coefficient is the cosine of the angle between the response vector  $\vec{y}$  and the estimation space that spanned by all predictors.

On the other hand, the *partial correlation coefficient* between  $y$  and  $x_1$ , denoted by  $r_{yx_1 \cdot x_2}$ , is defined in such a way that it measures the effect of  $x_1$  on  $y$  where  $x_2$  is not accounted for in the model. Conceptually, it is calculated by eliminating the linear effect of  $x_2$  on  $y$  as well as the linear effect of  $x_2$  on  $x_1$ . To purify  $y$  and  $x_1$  of the linear influence of  $x_2$ , we can first regress  $x_1$  on  $x_2$  and obtain the residual  $\hat{\epsilon}_1$  (i.e.  $\overrightarrow{DE}$  in Figure 2). Next, we regress  $y$  on  $x_2$  to obtain the second residual  $\hat{\epsilon}_2$  (i.e.  $\overrightarrow{BG}$  in Figure 2). Then given  $x_2$ , the partial correlation coefficient between  $y$  and  $x_1$ , can be obtained as the simple correlation coefficient between  $\hat{\epsilon}_2$  and  $\hat{\epsilon}_1$ , or equivalently in Figure 2,

$$r_{yx_1 \cdot x_2} = \cos \phi_1 \quad (3)$$

where  $\phi_1$  is the angle between  $\overrightarrow{BG}$  and  $\overrightarrow{DE}$ . This result leads to the following proposition.

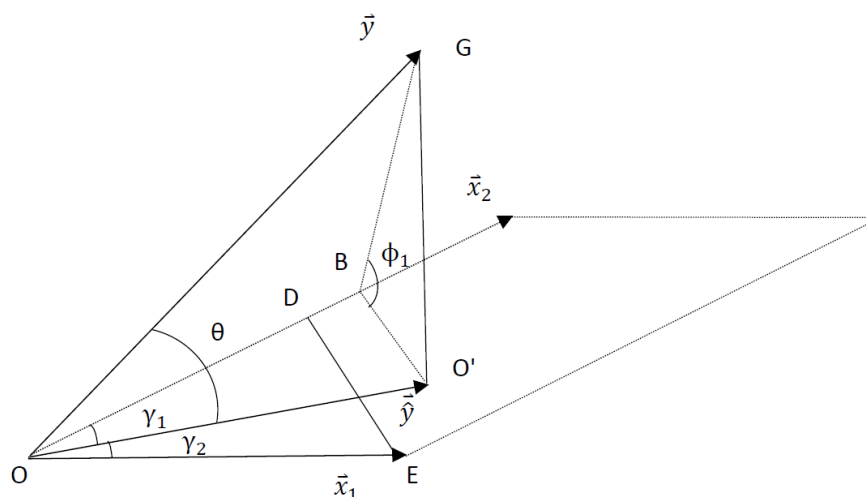


Figure 2. Geometric Interpretation of Partial Correlation Coefficients

**Proposition 1:** Given that  $x_2$  is retained in the model, the partial correlation coefficient between  $y$  and  $x_1$  is the cosine of the angle between the subspace that spanned by  $\vec{y}$  and  $\vec{x}_2$ , and the estimation space that spanned by  $\vec{x}_1$  and  $\vec{x}_2$ .

To calculate the partial correlation between  $y$  and  $x_2$ , one just needs to simply switch the subscripts of  $x$  vectors. If we define the relevant angle as  $\phi_2$ , we have  $r_{yx_2 \cdot x_1} = \cos \phi_2$ .

### 3.2 Partial Correlation Coefficient and The Frisch-Waugh-Lovell Theorem

One way to view the partial correlation coefficient is from the perspective of partial regression coefficient, since both terms are used to evaluate the contribution of the interested variable alone given the remaining variables in the regression model. In fact, the idea of Frisch-Waugh-Lovell theorem (Frisch & Waugh, 1933; Lovell, 1963), a well-known econometric theorem that is proposed to estimate the partial regression coefficient, is also helpful for understanding the geometric interpretation in equation (3) since its two-step trend removal procedure discloses the true meaning of partial correlation coefficient. As an alternative to the direct application of least squares, Frisch-Waugh-Lovell (FWL) Theorem shows that for the regression model

$$Y = X_1\beta_1 + X_2\beta_2 + \varepsilon \quad (4)$$

where  $X_1$  and  $X_2$  are  $n \times k_1$  and  $n \times k_2$  design matrices, respectively, the estimate of  $\beta_1$ , the  $k_1 \times 1$  coefficient vector for  $X_1$ , will be the same as its estimate from a modified regression model

$$M_{X_2}Y = M_{X_2}X_1\beta_1 + M_{X_2}\varepsilon \quad (5)$$

where  $M_{X_2} = I - X_2(X_2^T X_2)^{-1} X_2^T$  projects onto the orthogonal complement of the column space of  $X_2$ . In a multiple regression with two-predictor setting (2), the second residual  $\hat{\varepsilon}_2$  can be obtained by projecting  $Y$  onto the orthogonal complement of the column space of  $X_2$ . That is,  $\hat{\varepsilon}_2 = M_{X_2} Y$ . The first residual  $\hat{\varepsilon}_1$  also yields to  $M_{X_2} X_1$ . According to the FWL theorem, the partial regression coefficient of  $x_1$  (i.e.  $\hat{\beta}_1$ ) from regressing  $y$  on  $x_1$  and  $x_2$  simultaneously is simply obtained by regressing  $\hat{\varepsilon}_2$  on  $\hat{\varepsilon}_1$ . Consequently, the partial correlation coefficient  $r_{y x_1 \cdot x_2}$  yields the simple correlation coefficient between  $\hat{\varepsilon}_1$  and  $\hat{\varepsilon}_2$  in the modified regression model (5).

### 3.3 Relationship among Simple, Partial, and Multiple Correlation Coefficients

Another way to examine the partial correlation coefficient is by looking into its relationship with other types of correlation coefficients. This section geometrically studies and verifies the relationship among simple, partial and multiple correlation coefficients. Three classical equations are chosen to characterize the relationship among three types of correlation coefficients. These equations have been proved by using algebras and matrices which can be found in basic statistics or econometrics texts. However, in this section, the process of deriving these equations totally relies on simple geometric techniques that are not only easy to grasp but also provide readers with richer understanding of correlation coefficients from a different perspective.

$$(I) \quad (1 - R^2) = (1 - r_{yx_2}^2)(1 - r_{yx_1, x_2}^2) \quad (6)$$

This result first appears in a basic text by Anderson (1958). Here, the coefficient of determination  $R^2$  is simply the square of multiple correlation coefficient. Based on the geometric expressions summarized from the previous section, in Figure 2, it is clear that the left side of (6) yields to

$$1 - r_{y \cdot x_1 x_2}^2 = (1 - \cos^2 \theta) = \sin^2 \theta = \frac{|\overrightarrow{O'G}|}{|\overrightarrow{OG}|}$$

Similarly, the right side of (6) yields to

$$\sin^2 \theta_2 \sin^2 \phi_1 = \frac{|\overrightarrow{BG}|}{|\overrightarrow{OG}|} \frac{|\overrightarrow{O'G}|}{|\overrightarrow{BG}|} = \frac{|\overrightarrow{O'G}|}{|\overrightarrow{OG}|}$$

where  $\theta_2$  is the angle between  $\vec{x}_2$  and  $\vec{y}$ . Then (6) is concluded.

Symmetrically,

$$(1 - R^2) = (1 - r_{yx_1}^2)(1 - r_{yx_2 \cdot x_1}^2) \quad (7)$$

Equation (6) and (7) can also be simplified as

$$R^2 = r_{yx_1}^2 + (1 - r_{yx_1}^2)r_{yx_2 \cdot x_1}^2 \quad (8)$$

and

$$R^2 = r_{yx_2}^2 + (1 - r_{yx_2}^2)r_{yx_1 \cdot x_2}^2 \quad (9)$$

respectively. Equation (8) states that the proportion of the variation in  $y$  explained by  $x_1$  and  $x_2$  jointly is the sum of two parts: the part explained by  $x_1$  alone (i.e.  $r_{yx_1}^2$ ) and the part not explained by  $x_1$  (i.e.  $1 - r_{yx_1}^2$ ) times the proportion that is explained by  $x_2$  after eliminating the influence of  $x_1$  (Gujarati, 1995).

$$(II) \quad r_{yx_1} = r_{y \cdot x_1 x_2} r_{\hat{y} x_1} \quad (10)$$

This equation builds up a connection between the simple correlation and multiple correlation. In Figure 2, section 3.1 suggests that  $r_{yx_1} = \cos \theta_1 = \frac{|\overrightarrow{OB}|}{|\overrightarrow{OG}|}$ . Note that

$$r_{y \cdot x_1 x_2} r_{\hat{y} x_1} = \cos \theta \cos \gamma_1 = \frac{|\overrightarrow{OO'}|}{|\overrightarrow{OG}|} \cdot \frac{|\overrightarrow{OB}|}{|\overrightarrow{OO'}|} = \frac{|\overrightarrow{OB}|}{|\overrightarrow{OG}|}$$

where  $\gamma_1$  denotes the angle between vector  $\vec{x}_1$  and  $\vec{\hat{y}}$ . Equation (10) is therefore clearly obtained. Symmetrically,

$$r_{yx_2} = r_{y \cdot x_1 x_2} r_{\hat{y} x_2} \quad (11)$$

$$(III) \quad r_{yx_1 \cdot x_2} = \frac{r_{yx_1} - r_{yx_2} r_{x_1 x_2}}{\sqrt{(1 - r_{yx_2}^2)(1 - r_{x_1 x_2}^2)}} \quad (12)$$

This is a well-known algebraic formula for developing partial correlation coefficient from simple correlation coefficients. The spherical triangles method (Thomas & O'quigley, 1993) shows that equation (12) is identical to the formula of spherical trigonometry, which is illustrating but advanced for some readers. Alternatively, this section proposes a simple version of its geometric proof.

Proof: According to the equations (6) and (10), we can conclude that

$$\cos \theta_1 = \cos \theta \cdot \cos \gamma_1 \quad (13)$$

And

$$\cos \theta_2 = \cos \theta \cdot \cos \gamma_2 \quad (14)$$

where  $\gamma_2$  is the angle between vector  $\vec{x}_2$  and  $\vec{\hat{y}}$ . In Figure 2,

$$\cos \phi_1 = \frac{|\overrightarrow{BO'}|}{|\overrightarrow{BG}|} = \frac{|\vec{\hat{y}}| \sin \gamma_2}{|\vec{y}| \sin \theta_2} = \frac{\cos \theta \sin \gamma_2}{\sin \theta_2} \quad (15)$$

On the other hand, since  $\gamma_1 = \gamma - \gamma_2$ , then

$$\begin{aligned}
\cos \theta_1 &= \cos \theta \cdot \cos \gamma_1 \\
&= \cos \theta \cdot \cos(\gamma - \gamma_2) \\
&= \cos \gamma \cos \theta \cos \gamma_2 + \sin \gamma \cos \theta \sin \gamma_2
\end{aligned} \tag{16}$$

Substitute (14) into (16),

$$\cos \theta \sin \gamma_2 = \frac{\cos \theta_1 - \cos \theta_2 \cos \gamma}{\sin \gamma} \tag{17}$$

Substitute (17) into (15),

$$\begin{aligned}
\cos \phi_1 &= \frac{\cos \theta_1 - \cos \theta_2 \cos \gamma}{\sin \theta_2 \sin \gamma} \\
&= \frac{\cos \theta_1 - \cos \theta_2 \cos \gamma}{\sqrt{(1 - \cos^2 \theta_2)} \sqrt{(1 - \cos^2 \gamma)}}
\end{aligned}$$

Therefore, equation (12) is concluded.

### 3.4 Partial Correlation Coefficient and Partial Test Statistic

Similar to the concept of partial correlation coefficient, given the effect of a set of removed controlling variables, the partial F test is used to measure the importance of the interested variable alone from the perspective of hypothesis testing. This section discusses the geometric expression of the partial F test and discovers its geometric connection with partial test statistics.

In general, consider a multiple regression model with  $k$  predictors

$$v = \beta_1 u_1 + \beta_2 u_2 + \cdots + \beta_k u_k + w \tag{18}$$

where  $\vec{v}$ ,  $\vec{u}_1$ ,  $\vec{u}_2 \dots \vec{u}_k$  are  $n$ -dimensional column vectors in  $E^n$  and the regression coefficients  $\beta_i \in R^1$  for  $i = 1, 2, \dots, k$ . Assume that we are interested in testing the statistical significance of the first predictor  $u_1$ . That is, test  $H_0: \beta_1 = 0$ . In algebraic methods, the corresponding test statistic  $t_1$  is given by

$$t_1^2 = F_{1,d} = \frac{\text{Extra sum of squares per dimension}}{\text{Residual sum of squares per dimension}} = \frac{(SSE_R - SSE_{UR})/1}{SSE_{UR}/d} \tag{19}$$

Here,  $SSE_{UR}$  is the residual sum of squares of the unrestricted model that regresses  $v$  on  $u_1, u_2, \dots, u_k$  while  $SSE_R$  is the residual sum of squares the restricted model that regresses  $v$  on  $u_2, \dots, u_k$  only, leaving out  $u_1$ . Thus,  $(SSE_R - SSE_{UR})$  is the extra sum of squares when omitting  $u_1$  in the model. In the geometric approach, the number of degrees of freedom, denoted by  $d$ , is the number of dimensions in which the vector is free to move. Note that the dimension of  $\text{span}(\vec{v}, \vec{u}_1, \dots, \vec{u}_k)$  is  $(n - 1)$ , while the dimension of  $\text{span}(\vec{u}_1, \dots, \vec{u}_k)$  is  $k$ . Consequently, the dimension of the subspace in which the residual vector is free to move is  $(n - k - 1)$ .

Figure 3 illustrates the geometric process of the partial test statistic  $t_1$ . In Figure 3, vector  $\vec{v}_1$  (i.e.  $\vec{OB}$ ) is the estimation vector from projecting the observation vector  $\vec{v}$  on the subspace  $\text{span}(\vec{u}_2, \dots, \vec{u}_k)$ . Vector  $\vec{\hat{v}}$  (i.e.  $\vec{OO'}$ ) is the estimation vector from projecting  $\vec{v}$  on the subspace  $\text{span}(\vec{u}_1, \dots, \vec{u}_k)$ . Vector  $\vec{O'G}$  and  $\vec{BG}$  represent the residual vectors from projecting  $\vec{v}$  on the subspace  $\text{span}(\vec{u}_1, \dots, \vec{u}_k)$  and the subspace  $\text{span}(\vec{u}_2, \dots, \vec{u}_k)$ , respectively. Therefore, it is easy to conclude that

$$F_{1,n-k-1} = \frac{|\vec{BG}|^2}{|\vec{O'G}|^2 / (n-k-1)} = (n - k - 1) \cot^2 \phi_1 \tag{20}$$

where  $\phi_1$  is the angle between the residual vector from the restricted model and the vector that obtained by subtracting the residual vector of restricted model from the residual vector of unrestricted model. Equation (20) discloses the geometric interpretation of partial F test statistic. A similar result is concluded in Siniksaran (2005) from a different way.

In the general regression model (18),  $\cos \phi_1$  is the partial correlation coefficient between  $v$  and  $u_1$ , given that  $u_2, \dots, u_k$  are retained in the model. Specially, consider model (2), when  $k = 2$ , the display in Figure 3 is consistent with that of Figure 2. Equation (20) is reduced to

$$\tan \phi_1 = \frac{\sqrt{n-3}}{t_1} \tag{21}$$

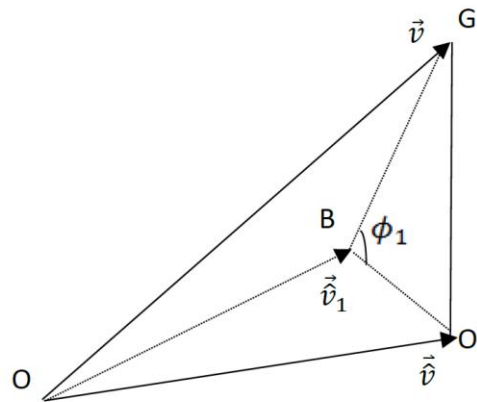


Figure 3 Geometric interpretation of t-statistics (3-D Graph)

According to equation (3) and trigonometric equation  $\cos^2 \phi_1 = \frac{1}{1+\tan^2 \phi_1}$ , a simple algebraic manipulation of (21) will lead to

$$r_{yx_1 \cdot x_2}^2 = \frac{t_1^2}{t_1^2 + (n-3)} \quad (22)$$

It is meaningful to study equation (22). On one hand, it furthermore discovers the relationship between test statistics and the partial correlation coefficients. On the other hand, it reveals the ideas for establishing the decision rule in the hypothesis testing. It is commonly known that the  $H_0: \beta_1 = 0$  should be rejected under significance level  $\alpha$  when  $t_1^2$  or  $F_{1, n-k-1}$  is large enough, say, larger than the critical value  $F_{1, (n-3), \alpha}$ . Based on equation (22), a large value of  $t_1^2$  or  $F_{1, n-k-1}$  implies a large value of  $r_{yx_1 \cdot x_2}$  which inherently suggests a stronger linear correlation between  $y$  and  $x_1$ , given that  $x_2$  is already retained in the model. Consequently, it is more likely that  $x_1$  is a significant addition to the prediction of  $y$ .

#### 4. Conclusion

To ensure the strictness of drawing  $n$ -dimensional vectors, angles and triangles into a  $E^3$ , this article first corrects an existing graphing problem in the current literature. This article also geometrically introduces and interprets the concept of partial correlation coefficient from the perspectives of FWL theorem, simple correlation coefficient, multiple correlation coefficient, and partial F test statistics. Unlike other pedagogical literatures in the regression studies, the geometric analogues of the regression concepts in a  $E^3$  do not require the prerequisite of advanced linear algebra nor abstruse abstract thinking. All the geometric proofs proposed in this article are concise and easy to follow. It is clear that the demonstration of the basic regression concepts in a visualized and familiar three-dimensional space facilitates the understanding of basic concepts for students and furthermore it serves as an important step stone for learning complex regression models in their future studies. It is worth to mention that the graphical displays in this paper are also suitable for a general multiple regression model with more than two predictors. And the results can be easily generalized. One just need to hold the vector of interested variable while having the other vector in the estimation space to represent the rest of predictors. Furthermore, for future work, other classical concepts and theorems in the regression analysis can also be visualized and interpreted in terms of a few principles of geometry through the geometric structure in this paper.

#### References

- Anderson, T. W. (1958). An Introduction to Multivariate Statistical Analysis. *John Wiley & Sons, Inc.*, New York.
- Bring, J., (1996). A Geometric Approach to Compare Variables in a Regression Model. *The American Statistician*, 50(1), 57-62. <https://doi.org/10.1080/00031305.1996.10473543>
- Draper, N. R., & Smith, H. (2014). Applied regression analysis. *John Wiley & Sons*.
- Frisch, R., & Waugh, F. V. (1933). Partial time regressions as compared with individual trends. *Econometrica: Journal of the Econometric Society*, 387-401. <https://doi.org/10.2307/1907330>

- Graybill, F. A. (1976). Theory and applications of the linear model. Duxbury Thomson Learning.
- Gujarati, D. N. (1995). *Basic Econometrics* (3<sup>rd</sup> edition). McGraw-Hill.
- Kutner, M., Nachtsheim, C., & Neter, J. (2004). *Applied linear regression models*. McGraw-Hill/Irwin.
- Lovell, M. C. (1963). Seasonal adjustment of economic time series and multiple regression analysis. *Journal of the American Statistical Association*, 58(304), 993-1010. <https://doi.org/10.1080/01621459.1963.10480682>
- Margolis, M. S. (1979). Perpendicular Projections and Elementary Statistics. *The American Statistician*, 33(3), 131-135. <https://doi.org/10.1080/00031305.1979.10482677>
- Saville, D. J., & Wood, G. R. (1986). A Method for Teaching Statistics Using N-Dimensional Geometry. *The American Statistician*, 40(3), 205-214. <https://doi.org/10.2307/2684537>
- Saville, D. J., & Wood, G. R. (1991). Statistical Methods: The Geometric Approach. New York: Springer. <https://doi.org/10.1007/978-1-4612-0971-3>
- Siniksaran, E. (2005). On the geometry of F, Wald, LR, and LM tests in linear regression models. *Statistics*, 39(4), 287-299. <https://doi.org/10.1080/02331880500178521>
- Thomas, G., & O'quigley, J. (1993). A Geometric Interpretation of Partial Correlation Using Spherical Triangles. *The American Statistician*, 47(1), 30-32. <https://doi.org/10.1080/00031305.1993.10475929>

## Appendix

### Proof of Theorem 1

In Figure 1,  $\theta_1$  is the angle between  $\vec{u}_1$  and  $\vec{v}$ ;  $\theta_2$  is the angle between  $\vec{u}_2$  and  $\vec{v}$ ; the angle between vector  $\vec{u}_2$  and  $\vec{\hat{v}}$  is  $\gamma_2$ .

$$\text{Let } A = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}, \text{ then, } \alpha_1 = \frac{\vec{u}_1^T}{|\vec{u}_1|}, \alpha_2 = \frac{(\vec{u}_2 - \vec{\hat{u}}_{21})^T}{|\vec{u}_2 - \vec{\hat{u}}_{21}|}, \alpha_3 = \frac{(\vec{v} - \vec{\hat{v}})^T}{|\vec{v} - \vec{\hat{v}}|}.$$

$$\alpha_1^T \alpha_1 = \frac{\vec{u}_1}{|\vec{u}_1|} \frac{\vec{u}_1^T}{|\vec{u}_1|} = \frac{|\vec{u}_1|^2}{|\vec{u}_1|^2} = 1;$$

$$\alpha_2^T \alpha_2 = \frac{(\vec{u}_2 - \vec{\hat{u}}_{21})}{|\vec{u}_2 - \vec{\hat{u}}_{21}|} \frac{(\vec{u}_2 - \vec{\hat{u}}_{21})^T}{|\vec{u}_2 - \vec{\hat{u}}_{21}|} = \frac{|\vec{u}_2 - \vec{\hat{u}}_{21}|^2}{|\vec{u}_2 - \vec{\hat{u}}_{21}|^2} = 1;$$

$$\alpha_3^T \alpha_3 = \frac{(\vec{v} - \vec{\hat{v}})}{|\vec{v} - \vec{\hat{v}}|} \frac{(\vec{v} - \vec{\hat{v}})^T}{|\vec{v} - \vec{\hat{v}}|} = \frac{|\vec{v} - \vec{\hat{v}}|^2}{|\vec{v} - \vec{\hat{v}}|^2} = 1.$$

On the other hand,

$$\alpha_1^T \alpha_2 = \frac{\vec{u}_1}{|\vec{u}_1|} \frac{(\vec{u}_2 - \vec{\hat{u}}_{21})^T}{|\vec{u}_2 - \vec{\hat{u}}_{21}|} = 0 \quad (\because \vec{u}_1 \text{ and } (\vec{u}_2 - \vec{\hat{u}}_{21}) \text{ are perpendicular.})$$

$$\alpha_1^T \alpha_3 = \frac{\vec{u}_1}{|\vec{u}_1|} \frac{(\vec{v} - \vec{\hat{v}})^T}{|\vec{v} - \vec{\hat{v}}|} = 0 \quad (\because \vec{u}_1 \text{ and } (\vec{v} - \vec{\hat{v}}) \text{ are perpendicular.})$$



$$\alpha_2^T \alpha_3 = \frac{(\bar{u}_2 - \hat{u}_{21})}{|\bar{u}_2 - \hat{u}_{21}|} \frac{(\bar{v} - \hat{v})^T}{|\bar{v} - \hat{v}|} = 0 \quad (\because (\bar{u}_2 - \hat{u}_{21}) \text{ and } (\bar{v} - \hat{v}) \text{ are perpendicular.})$$

Therefore, we prove that any pairs of row vectors in  $A$  are orthogonal vectors.

The coordinates of  $\bar{x}_1$ ,  $\bar{x}_2$ , and  $\bar{y}$  in the 3-dimensional space are:

$$\bar{x}_1 = \begin{pmatrix} |\bar{u}_1| \\ 0 \\ 0 \end{pmatrix}, \bar{x}_2 = \begin{pmatrix} |\bar{u}_2| \cos \gamma \\ |\bar{u}_2| \sin \gamma \\ 0 \end{pmatrix}, \text{ and } \bar{y} = \begin{pmatrix} |\bar{v}| \cos \theta \cos \gamma_1 \\ |\bar{v}| \cos \theta \sin \gamma_1 \\ |\bar{v}| \sin \theta \end{pmatrix}, \text{ respectively.}$$

$$A\bar{u}_1 = \begin{pmatrix} \frac{\bar{u}_1^T \bar{u}_1}{|\bar{u}_1|} \\ \frac{(\bar{u}_2 - \hat{u}_{21})^T \bar{u}_1}{|\bar{u}_2 - \hat{u}_{21}|} \\ \frac{(\bar{v} - \hat{v})^T \bar{u}_1}{|\bar{v} - \hat{v}|} \end{pmatrix} = \begin{pmatrix} \frac{|\bar{u}_1|^2}{|\bar{u}_1|} \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} |\bar{u}_1| \\ 0 \\ 0 \end{pmatrix} = \bar{x}_1;$$

$$A\bar{u}_2 = \begin{pmatrix} \frac{\bar{u}_1^T \bar{u}_2}{|\bar{u}_1|} \\ \frac{(\bar{u}_2 - \hat{u}_{21})^T \bar{u}_2}{|\bar{u}_2 - \hat{u}_{21}|} \\ \frac{(\bar{v} - \hat{v})^T \bar{u}_2}{|\bar{v} - \hat{v}|} \end{pmatrix} = \begin{pmatrix} \frac{|\bar{u}_1| |\bar{u}_2| \cos \gamma}{|\bar{u}_1|} \\ \frac{|\bar{u}_2 - \hat{u}_{21}| |\bar{u}_2| \sin \gamma}{|\bar{u}_2 - \hat{u}_{21}|} \\ 0 \end{pmatrix} = \begin{pmatrix} |\bar{u}_2| \cos \gamma \\ |\bar{u}_2| \sin \gamma \\ 0 \end{pmatrix} = \bar{x}_2;$$

$$A\bar{v} = \begin{pmatrix} \frac{\bar{u}_1^T \bar{v}}{|\bar{u}_1|} \\ \frac{(\bar{u}_2 - \hat{u}_{21})^T \bar{v}}{|\bar{u}_2 - \hat{u}_{21}|} \\ \frac{(\bar{v} - \hat{v})^T \bar{v}}{|\bar{v} - \hat{v}|} \end{pmatrix} = \begin{pmatrix} \frac{|\bar{u}_1| |\bar{v}| \cos \theta \cos \gamma_1}{|\bar{u}_1|} \\ \frac{|\bar{u}_2 - \hat{u}_{21}| |\bar{v}| \cos \theta \sin \gamma_1}{|\bar{u}_2 - \hat{u}_{21}|} \\ \frac{|\bar{v} - \hat{v}| |\bar{v}| \sin \theta}{|\bar{v} - \hat{v}|} \end{pmatrix} = \begin{pmatrix} |\bar{v}| \cos \theta \cos \gamma_1 \\ |\bar{v}| \cos \theta \sin \gamma_1 \\ |\bar{v}| \sin \theta \end{pmatrix} = \bar{y}.$$

On the other hand,

$$|\bar{x}_1| = \sqrt{\bar{x}_1^T \bar{x}_1} = \sqrt{|\bar{u}_1|^2 + 0 + 0} = |\bar{u}_1|$$

$$|\bar{x}_2| = \sqrt{\bar{x}_2^T \bar{x}_2} = \sqrt{|\bar{u}_2|^2 \cos^2 \gamma + |\bar{u}_2|^2 \sin^2 \gamma + 0} = \sqrt{|\bar{u}_2|^2} = |\bar{u}_2|$$

$$|\bar{y}| = \sqrt{\bar{y}^T \bar{y}} = \sqrt{|\bar{v}|^2 \cos^2 \theta \cos^2 \gamma_1 + |\bar{v}|^2 \cos^2 \theta \sin^2 \gamma_1 + |\bar{v}|^2 \sin^2 \theta} = \sqrt{|\bar{v}|^2} = |\bar{v}|$$

Therefore,  $A$  keeps the length of any vector unchanged. In addition, we have,

$$\langle \bar{x}_1, \bar{x}_2 \rangle = \bar{x}_1^T \bar{x}_2 = |\bar{u}_1| |\bar{u}_2| \cos \gamma + 0 + 0 = |\bar{u}_1| |\bar{u}_2| \cos \gamma = \langle \bar{u}_1, \bar{u}_2 \rangle$$

$$\langle \bar{x}_1, \bar{y} \rangle = \bar{x}_1^T \bar{y} = |\bar{u}_1| |\bar{v}| \cos \theta \cos \gamma_1 + 0 + 0 = |\bar{u}_1| |\bar{v}| \cos \theta \cos \gamma_1 = |\bar{u}_1| |\bar{v}| \cos \theta_1 = \langle \bar{u}_1, \bar{v} \rangle$$

$$\langle \bar{x}_2, \bar{y} \rangle = \bar{x}_2^T \bar{y} = |\bar{u}_2| |\bar{v}| \cos \theta \cos \gamma \cos \gamma_1 + |\bar{u}_2| |\bar{v}| \cos \theta \sin \gamma \sin \gamma_1 + 0$$

$$= |\bar{u}_2| |\bar{v}| \cos \theta \cos(\gamma - \gamma_1) = |\bar{u}_2| |\bar{v}| \cos \theta \cos \gamma_2 = |\bar{u}_2| |\bar{v}| \cos \theta_2 = \langle \bar{u}_2, \bar{v} \rangle$$

Therefore,  $A$  also keeps the inner product, or equivalently, the angle between any two vectors unchanged. Thus, theorem 1 is proved.

### Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).