A Note on Geometric Interpretations of Regression Analysis

Kang Chen*, Cong Wang**, Bilin Zeng***,

*Lee Kuan Yew School of Public Policy, National University of Singapore, Singapore
**Department of Finance, Chinese University of Hong Kong, Hong Kong
***Department of Mathematics, California State University, Bakersfield, CA, U.S.A

Abstract
This article aims to systematically restate and interpret basic concepts and important properties of regression analysis under the geometric framework. We first propose a simple geometric proof for the Frisch-Waugh-Lovell Theorem which is then followed by developing geometric expressions of regression coefficients and partial correlation coefficients. In addition, we apply geometric approaches to prove and verify three formulas that display the relationship among simple, multiple and partial correlation coefficients. Moreover, by describing the geometric analogues of the regression concepts systematically, this article also illustrates that geometric approach can provide a better understanding of regression analysis. Although these formulas are well known in the basic statistics textbooks and have already been proved by algebraic methods, we are among the first who give the geometric proof.

Keywords
Regression Analysis, Geometry Interpretation, Frisch-Waugh-Lovell Theorem, Partial Correlation Coefficient

1. INTRODUCTION

Regression analysis is not only a widely used statistical tool but also a very important curriculum subject for both graduate and undergraduate students majoring in statistics, economics, politics, psychology and sociology. Regression analysis is traditionally presented in algebraic forms, especially in equations and matrices. Since the subject of regression emerged in the late of 19^{th} century, algebra has been widely used to express concepts and build up models in regression analysis. When taking the course of regression methods, college students also use textbooks that express regression analysis totally in algebraic equations and matrix forms. However, all concepts of regression analysis can be visualized by applying a few principles of geometry (Bryant 1984). Statisticians have shown that many key concepts in regression analysis, including the method of least squares, regression coefficients, simple and partial correlation coefficients, have direct visual analogues in the geometry of vectors.

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) pointed out that geometry seems to be the natural way to emphasize the unity of the fundamental ideas. Furthermore, regression concepts and techniques can be explained more simply and clearly in a geometric way than in sophisticated algebraic equations and complicated matrix forms (Herr 1980). Bring (1994) also advocated using geometry to present basic regression concepts to serious beginning students because it can definitely improve a student's comprehension of basic concepts. By exploring vectors, triangles and projections, and drawing them clearly in a three-dimensional space, students do not have to delve into complicated algebraic calculations. Therefore, an understanding of the geometrical aspects of elementary regression analysis may sometimes assist a student more effectively than elegantly derived formulas (Saville and Wood 1986).

*** Corresponding author. Email address: bzeng@csub.edu

Despite its merits, geometry is seldom used in regression analysis, except for the use of scatter plots (Marks 1982). The long-time predominance of algebra since the 20^{th} century might be one of the primary reasons. Besides, the formalist philosophical stance of mathematics, which means strictness in formal exposition, makes algebra a more popular instrument than geometry in academic research (Davis and Hersh 1981). On the other hand, some statisticians and professors who teach statistics in universities have given other reasons for the unpopularity of geometry in research and in teaching of regression analysis in particular. Herr (1980), for instance, reviewed selected statistical papers with a geometric slant, from R. A. Fisher's (1915) to William Kruskal's (1975), and concluded that the relative unpopularity of the geometric approach is "not due to its inherent inferiority but rather to a combination of inertia, poor exposition, and a resistance to abstraction". Davis (1981) had a similar opinion and emphasized that to fully appreciate the analytic geometric approach and to be able to use it effectively in research requires statisticians with a talent for abstract thought. The resistance to abstraction is probably the most important reason why geometric approach is so seldom used. Moreover, that the relevant material has not been published at an appropriate level is perhaps another reason why the geometric approach is seldom used in teaching regression analysis. Thus, although some ideas have been expressed in terms of geometry, they are pitched at such a high mathematical level that neither students nor professors would be likely to embrace them (Bryant 1984).

Previous work has been done to show the geometric interpretations of the method of least squares, regression coefficients, simple, multiple and partial correlation coefficients. However, these are far from enough. Box, Hunter, and Hunter (1978) applied vector geometry in one of their statistical texts intended for the first course in the design of regression analysis experiments. Draper and Smith (1980) used two chapters to discuss the geometry of least squares. From the statistical modeling point of view, Marks (1982) gave a distribution-based geometric interpretation of the correlation coefficient which was built upon an assumption of bivariate normal distribution on regression variables. This geometric interpretation is more consistent with the statistical model for the data. However, the result might be invalid if distribution assumption is violated. Morris L. Eaton (1983) wrote a book, using vector space approach to present a version of multivariate statistical theory. In the book, the author not only used random vectors to analyze normal distribution and linear models, but also combined the vector space method and invariance together to solve multivariate problems. However, the mathematics used in the book is in a graduate course level and difficult for a beginning student to read. Saville and Wood (1986) advocated a method for teaching statistics using n-dimensional geometry in the Teacher's Corner, and published two books on this subject (Saville and Wood 1991, 1996). In both of their books, the set of nonorthogonal predictor vectors are converted into an orthogonal sequence for the ease of their work. Thomas and O'Quigley (1993) gave a geometric interpretation of partial correlation from the perspective of spherical triangles model. And more recently, Bring (1996) showed the power of the geometric approach in studying regression by demonstrating the geometric view of some important concepts such as, least square, standardized regression coefficient, and R-Square. There are some projective arguments in the literature which require more knowledge of advanced linear algebra for beginning readers. In geometry, each regression variable is considered as a vector in an n-dimensional space, where n is the size of the sample. For the ease of demonstration, one usually has an n-dimensional vector displayed in a 3-dimensional vector space. This is a common practice in most existing literatures. However, this is not strictly correct as in higher dimensions ($n>3$), vectors cannot be shown pictorially in a strictly correct manner (Saville and Wood 1991).

The objective of this paper is to geometrically restate and interpret basic regression concepts and properties under a simple and strictly correct 3-dimensional vector space, without enforcing any distribution assumptions on regression variables or making any transformation of non-orthogonal regression vectors. The rest of the paper is organized as follows. In section 2, we first correct a graphing problem in the previous research by mapping the original n-dimensional space E^n to a 3-dimensional one without the loss of information. This process ensures that we can draw any high dimensional vectors, angles and triangles in a strictly correct manner. In addition to the previous studies, in section 3, we use geometry to prove the Frisch-Waugh-Lovell Theorem (Frisch and Waugh 1933; Lovell 1963) and provide another four geometric expressions of regression coefficients based on this 3-dimensional geometric framework. Moreover, we find another geometric interpretation of partial correlation coefficients. Finally in section 4, we use geometric approaches to prove and verify three formulas that display the relationship among simple, multiple and partial correlation coefficients. Those who have finished an introductory course of econometrics would find it especially helpful when they examine these visual analogues sketched out in the paper.

2. DRAWING GRAPHS ON E^3

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

$$v = \beta_1 u_1 + \beta_2 u_2 + w,$$

where $v = (v_1,\ v_2, \ldots, v_n)^T$ is the observed response vector, two independent predictors u_1 and u_2 are both n-dimensional column vectors, $\beta_1 \in R^1$ and $\beta_2 \in R^1$ are the regression coefficients, w is the error term, and n is the sample size. In terms of vector geometry, $\vec{v}$, $\vec{u}_1$, $\vec{u}_2$ and $\vec{w}$ can be regarded as vectors in n-dimensional Euclidean space, E^n. When using geometry to analyze this model, most scholars normally draw the diagram in Figure 1 directly. In Figure 1, $\overrightarrow{OO'}$ is the perpendicular projection of $\vec{v}$ on $Span\ (\vec{u}_1, \vec{u}_2)$, the plane spanned by $\vec{u}_1$ and $\vec{u}_2$. $\overrightarrow{OO'}$ is also named as $\vec{\hat{v}}$. Vector $\overrightarrow{OA}$ is the perpendicular projection of $\vec{v}$ on $\vec{u}_1$; vector $\overrightarrow{OB}$ is the perpendicular projection of $\vec{v}$ on $\vec{u}_2$. Then, $\overrightarrow{AO'} \perp \vec{u}_1$, $\overrightarrow{BO'} \perp \vec{u}_2$. The angle between vector $\vec{v}$ and $\vec{\hat{v}}$ is θ; the angle between vector $\vec{u}_1$ and $\vec{\hat{v}}$ is γ_1; the angle between vector $\vec{u}_2$ and $\vec{\hat{v}}$ is γ_2. Finally, the angle between $\vec{u}_1$ and $\vec{u}_2$ is γ.

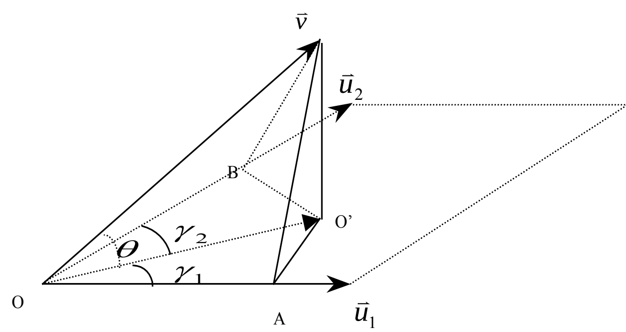

Figure 1. Geometric Interpretation of Least Squares Method in n-space.

However, the problem with Figure 1 is that it is displaying in a 3-dimensional space. And this contradicts with the fact that $\vec{v}$, $\vec{u}_1$ and $\vec{u}_2$ are vectors in n-space E^n. According to Saville and Wood (1991), in higher (n>3) dimensions, vectors cannot be shown pictorially in a strictly correct manner. Therefore, it is not strictly correct to use Figure 1 to analyze regressions. 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-dimentional space. This linear transformation keeps:

1) The length of any vector unchanged;
2) The angle between any two vectors unchanged.

When pre-multiplying A with $\vec{v}$, $\vec{u}_1$, $\vec{u}_2$, and $\vec{w}$, we obtain $\vec{y}$, $\vec{x}_1$, $\vec{x}_2$ and $\vec{\varepsilon}$ respectively. That is

$$\underset{n\times1}{Av} = A\underset{n\times2}{u}\beta + A\underset{n\times1}{w}$$
$$\downarrow \quad \downarrow \quad \downarrow$$
$$\underset{3\times1}{y} = \underset{3\times2}{x}\beta + \underset{3\times1}{\varepsilon}$$

where $\vec{y}$, $\vec{x}_1$, and $\vec{x}_2$ 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, the new vectors $\vec{y}$, $\vec{x}_1$ and $\vec{x}_2$ can be shown pictorially in a strictly correct manner. In this process, $\hat{\beta}_i$ $(i = 1{,}2)$ is unchanged. The 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 $Span(\vec{v}, \vec{u}_1, \vec{u}_2)$ and E^3, such that,

$$A\vec{u}_1 = \begin{pmatrix} |\vec{u}_1| \\ 0 \\ 0 \end{pmatrix} = \vec{x}_1, \; A\vec{u}_2 = \begin{pmatrix} |\vec{u}_2|\cos\gamma \\ |\vec{u}_2|\sin\gamma \\ 0 \end{pmatrix} = \vec{x}_2, \textit{ and } A\vec{v} = \begin{pmatrix} |\vec{v}|\cos\theta\cos\gamma_1 \\ |\vec{v}|\cos\theta\sin\gamma_1 \\ |\vec{v}|\sin\theta \end{pmatrix} = \vec{y} \qquad \textit{where} \quad A = \begin{pmatrix} \frac{\vec{u}_1^T}{|\vec{u}_1|} \\ \frac{(\vec{u}_2 - \hat{\vec{u}}_{21})^T}{|\vec{u}_2 - \hat{\vec{u}}_{21}|} \\ \frac{(\vec{v} - \hat{\vec{v}})^T}{|\vec{v} - \hat{\vec{v}}|} \end{pmatrix}.$$

In this theorem, vector $\vec{\hat{u}}_{21}$ is the predictor of $\vec{u}_2$ when regressing $\vec{u}_2$ on $\vec{u}_1$. Hence $(\vec{u}_2 - \vec{\hat{u}}_{21})$ is the residual vector from regressing $\vec{u}_2$ on $\vec{u}_1$ and thus is perpendicular to $\vec{u}_1$. $(\vec{v} - \hat{\vec{v}})$ is the residual vector from regressing $\vec{v}$ on $\vec{u}_1$ and $\vec{u}_2$, and is perpendicular to $Span(\vec{u}_1, \vec{u}_2)$. And we also define the angle between $\vec{u}_1$ and $\vec{v}$ as θ_1; the angle between $\vec{u}_2$ and $\vec{v}$ as θ_2.
To prove theorem 1, we need to prove

1. Any pairs of row vectors in A are orthogonal vectors;
2. $A\vec{u}_1 = \vec{x}_1$, $A\vec{u}_2 = \vec{x}_2$, $A\vec{v} = \vec{y}$ and A preserves (a) the length of any vector (b) the angle between any two vectors.

<u>Proof :</u>

Let $A = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}$, then, $\alpha_1 = \frac{\vec{u}_1^T}{|\vec{u}_1|}$, $\alpha_2 = \frac{(\vec{u}_2 - \hat{\vec{u}}_{21})^T}{|\vec{u}_2 - \hat{\vec{u}}_{21}|}$, $\alpha_3 = \frac{(\vec{v} - \hat{\vec{v}})^T}{|\vec{v} - \hat{\vec{v}}|}$, $\quad \alpha_1^T\,\alpha_1 = \frac{\vec{u}_1}{|\vec{u}_1|}\frac{\vec{u}_1^T}{|\vec{u}_1|} = \frac{|\vec{u}_1|^2}{|\vec{u}_1|^2} = 1$;

$$\alpha_2^T\,\alpha_2 = \frac{(\vec{u}_2 - \hat{\vec{u}}_{21})}{|\vec{u}_2 - \hat{\vec{u}}_{21}|}\frac{(\vec{u}_2 - \hat{\vec{u}}_{21})^T}{|\vec{u}_2 - \hat{\vec{u}}_{21}|} = \frac{|\vec{u}_2 - \hat{\vec{u}}_{21}|^2}{|\vec{u}_2 - \hat{\vec{u}}_{21}|^2} = 1; \qquad \alpha_3^T\,\alpha_3 = \frac{(\vec{v} - \hat{\vec{v}})}{|\vec{v} - \hat{\vec{v}}|}\frac{(\vec{v} - \hat{\vec{v}})^T}{|\vec{v} - \hat{\vec{v}}|} = \frac{|\vec{v} - \hat{\vec{v}}|^2}{|\vec{v} - \hat{\vec{v}}|^2} = 1.$$

On the other hand,

$$\alpha_1^T\,\alpha_2 = \frac{\vec{u}_1}{|\vec{u}_1|}\frac{(\vec{u}_2 - \hat{\vec{u}}_{21})^T}{|\vec{u}_2 - \hat{\vec{u}}_{21}|} = 0; \; \alpha_1^T\,\alpha_3 = \frac{\vec{u}_1}{|\vec{u}_1|}\frac{(\vec{v} - \hat{\vec{v}})^T}{|\vec{v} - \hat{\vec{v}}|} = 0 \text{ and } \alpha_2^T\,\alpha_3 = \frac{(\vec{u}_2 - \hat{\vec{u}}_{21})}{|\vec{u}_2 - \hat{\vec{u}}_{21}|}\frac{(\vec{v} - \hat{\vec{v}})^T}{|\vec{v} - \hat{\vec{v}}|} = 0$$

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

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

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

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

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

On the other hand,

$$|\vec{x}_1|=\sqrt{\vec{x}_1^T\vec{x}_1}=\sqrt{|\vec{u}_1|^2+0+0}=|\vec{u}_1|,\ |\vec{x}_2|=\sqrt{\vec{x}_2^T\vec{x}_2}=\sqrt{|\vec{u}_2|^2\cos^2\gamma+|\vec{u}_2|^2\sin^2\gamma+0}=\sqrt{|\vec{u}_2^2|}=|\vec{u}_2|$$

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

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

$$<\vec{x}_1,\vec{x}_2>=\vec{x}_1^T\vec{x}_2=|\vec{u}_1||\vec{u}_2|\cos\gamma+0+0=|\vec{u}_1||\vec{u}_2|\cos\gamma=<\vec{u}_1,\vec{u}_2>$$

$$<\vec{x}_1,\vec{y}>=\vec{x}_1^T\vec{y}=|\vec{u}_1||\vec{v}|\cos\theta\cos\gamma_1+0+0=|\vec{u}_1||\vec{v}|\cos\theta\cos\gamma_1=|\vec{u}_1||\vec{v}|\cos\theta_1=<\vec{u}_1,\vec{v}>$$

$$<\vec{x}_2,\vec{y}>=\vec{x}_2^T\vec{y}=|\vec{u}_2||\vec{v}|\cos\theta\cos\gamma\cos\gamma_1+|\vec{u}_2||\vec{v}|\cos\theta\sin\gamma\sin\gamma_1+0$$

$$=|\vec{u}_2||\vec{v}|\cos\theta\cos(\gamma-\gamma_1)=|\vec{u}_2||\vec{v}|\cos\theta\cos\gamma_2=|\vec{u}_2||\vec{v}|\cos\theta_2=<\vec{u}_2,\vec{v}>$$

Therefore, A also keeps the inner product, or equivalently, the angle between any two vectors unchanged. Thus, theorem 1 is proved. In the following sections, our analysis is built upon this E^3 framework and thus our vectors can be pictured in a strictly correct manner.

3. GEOMETRIC INTERPRETATIONS OF BASIC CONCEPTS IN REGRESSION

In this section, we focus on the two-independent variable case: $v=\beta_1u_1+\beta_2u_2+w$.

Based on theorem 1, this n-dimensional case is transferred to the following 3-dimensional case

$$y=\beta_1x_1+\beta_2x_2+\varepsilon,$$

3.1 Regression Coefficients

Figure 3 demonstrates how to get the estimated regression coefficients. On the x_1x_2 plane, we need to focus on parallelogram $OA_1O^{'}B_1$. It has been concluded that the estimated regression coefficients are the proportions of the vectors $\vec{x}_1$ and $\vec{x}_2$, that, when are added, give the vector $\vec{\hat{y}}$.

Since $\overrightarrow{OA_1}+\overrightarrow{OB_1}=\vec{\hat{y}}$, we can easily see that: $\hat{\beta}_1=\frac{|\overrightarrow{OA_1}|}{|\vec{x}_1|}$, $\hat{\beta}_2=\frac{|\overrightarrow{OB_1}|}{|\vec{x}_2|}$

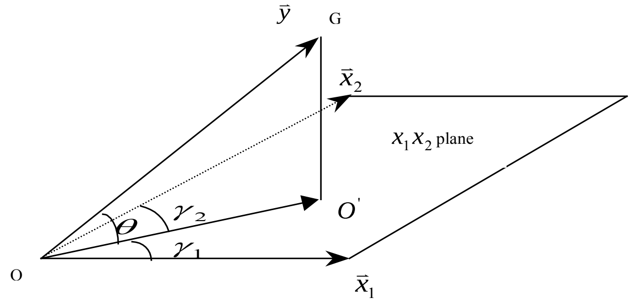

Figure 2. Geometric Interpretation of Least Square Method in Multiple Regression Model.

This is one geometric interpretation of regression coefficients and it is the most commonly known. However, we know that regression coefficients can also be estimated by other procedures that can disclose the true meaning of the partial regression coefficients. These procedures are summarized as Frisch-Waugh-Lovell Theorem. The Frisch-Waugh-Lovell theorem is a well-known result in econometrics which was named after the econometricians Ragnar Frisch, Frederick Waugh, and Michael Lovell. As an alternative to the direct application of least squares, Ragnar Frisch and F.V. Waugh (1993) first demonstrated a two-step trend removal procedure and proved a remarkable property of the method of least squares (Lovell 2008). According to the Frisch-Waugh-Lovell theorem, we can first regress x_1 on x_2 and obtain the residual $\hat{\varepsilon}_1$. Then we regress y on x_2 and obtain the second residual $\hat{\varepsilon}_2$. Frisch-Waugh-Lovell Theorem shows that if we regress $\hat{\varepsilon}_2$ on $\hat{\varepsilon}_1$, the estimated regression coefficient is just $\hat{\beta}_1$, the partial regression coefficient of x_1 from regressing y on x_1 and x_2 simultaneously.

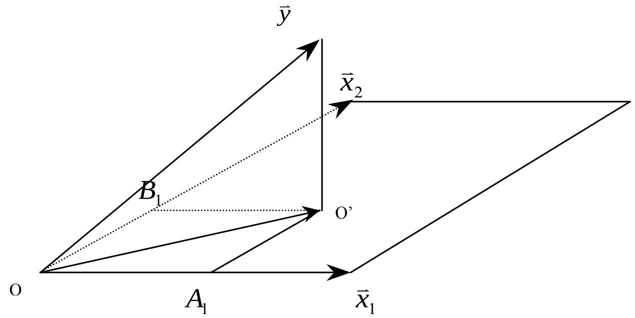

Figure 3. Geometric Interpretation of Estimated Regression Coefficients from OLS.

The theorem was originally proved by using Cramer's Rule (Frisch and Waugh 1933). Then, Lovell (1963) developed a more general result which was proved in terms of matrix algebra (1963) and algebra (2008), respectively. Davidson and MacKinnon (1999), and Sosa Escudero (2001) also showed geometric representations of this theorem in their work. Although the algebraic proof of Frisch-Waugh-Lovell Theorem is tedious, a geometric proof can be simple and straightforward. The process to proof this result geometrically is as follows.

In Figure 4, vector $\overrightarrow{OD}$ is the perpendicular projection of x_1 on x_2, or $\hat{\vec{x}}_{12}$. The corresponding residual vector $(\vec{x}_1 - \hat{\vec{x}}_{12})$ is $\overrightarrow{DE}$. Vector $\overrightarrow{OB}$ is the perpendicular projection of $\vec{y}$ on $\vec{x}_2$, or $\hat{\vec{y}}_2$. And the residual vector $(\vec{y} - \hat{\vec{y}}_2)$ is $\overrightarrow{BG}$. It is not difficult to find that vector $\overrightarrow{BO'}$ is the perpendicular projection of $\overrightarrow{BG}$ on $\overrightarrow{DE}$ ($\overrightarrow{BO'}$ is also the perpendicular projection of $\overrightarrow{BG}$ on the x_1x_2 plane). On the x_1x_2 plane, consider a vector $\overrightarrow{OA_3}$ such that $\overrightarrow{OA_3}$ // $\overrightarrow{DE}$ and $\left|\overrightarrow{OA_3}\right| = \left|\overrightarrow{DE}\right|$, then the residual vector $(\vec{x}_1 - \hat{\vec{x}}_{12})$ can also be represented by $\overrightarrow{OA_3}$. Suppose that there is a vector $\overrightarrow{A_2O'}$ that is parallel to $\vec{x}_2$ on the x_1x_2 plane, then we can find that $\overrightarrow{OA_2} = \overrightarrow{BO'}$ in the $OA_2O'B$ rectangular. Also clearly, the intersection between $\overrightarrow{OA_2}$ and $\vec{x}_2$ is the point A_1. On the other hand, the right-angled triangle ΔOA_2A_1 is similar to the right-angled triangle ΔOA_3E. According to the properties of similar triangles, we have: $\frac{\left|\overrightarrow{OA_2}\right|}{\left|\overrightarrow{OA_3}\right|} = \frac{\left|\overrightarrow{OA_1}\right|}{\left|\vec{x}_1\right|} = \hat{\beta}_1$. Equivalently, $\frac{\left|\overrightarrow{BO'}\right|}{\left|\overrightarrow{DE}\right|} = \hat{\beta}_1$.

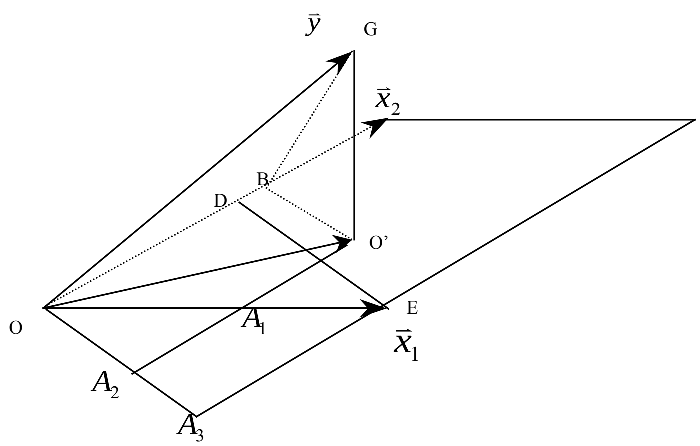

Figure 4. Geometric Interpretation of Frisch-Waugh-Lovell Theorem

Furthermore, the result of Frisch-Waugh-Lovell theorem can be restated geometrically in the following proposition.

Proposition 1: The estimated regression coefficient $\hat{\beta}_1$ is equal to the simple regression coefficient obtained by projecting the residual vector $(\vec{y} - \hat{\vec{y}}_2)$ on the residual vector $(\vec{x}_1 - \hat{\vec{x}}_{12})$.

On the other hand, vector $\vec{y}$ can be decomposed into two orthogonal components, vector $\hat{\vec{y}}_2$ and vector $(\vec{y} - \hat{\vec{y}}_2)$. Since vector $\hat{\vec{y}}_2$ (or $\overrightarrow{OB}$) is perpendicular to $\overrightarrow{DE}$, the projection of $\vec{y}$ on $\overrightarrow{DE}$ is equal to the projection of its second component $(\vec{y} - \hat{\vec{y}}_2)$ (or $\overrightarrow{BG}$) on $\overrightarrow{DE}$. Therefore, we develop another geometric procedure to calculate $\hat{\beta}_1$. And this leads to another result of Frisch-Waugh-Lovell theorem.

Proposition 2: The estimated regression coefficient $\hat{\beta}_1$ is equal to the simple regression coefficient obtained by projecting the observation vector $\vec{y}$ on the residual vector $(\vec{x}_1 - \vec{\hat{x}}_{12})$.

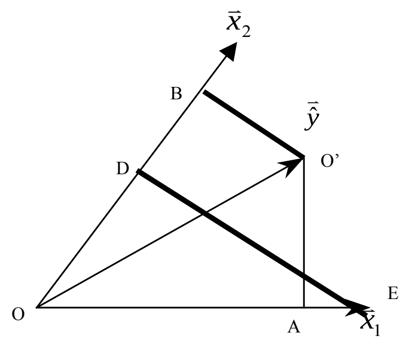

Figure 5. Two-Dimensional-Estimation Space.

Proposition 1 and 2 are proposed based on the analysis in E^3. If we completely center on the 2-dimentisonal x_1x_2 plane, we will find that the geometric interpretations of Frisch-Waugh-Lovell theorem are still valid. Refer to Figure 5, what we need to do is to change the residual vector $\overrightarrow{BG}$ to $\overrightarrow{BO'}$ in proposition 1 and replace the observation vector $\vec{y}$ by $\vec{\hat{y}}$ in proposition 2.

Figure 5 shows the Frisch-Waugh-Lovell theorem on the 2-dimentional x_1x_2 plane. Vector $\overrightarrow{DE}$ is still the residual vector from projecting $\vec{x}_1$ on $\vec{x}_2$. Vector $\overrightarrow{BO'}$ can be considered the residual vector from projecting $\vec{\hat{y}}$ on $\vec{x}_2$. We already know that $\hat{\beta}_1$ is equal to the ratio of the length of $\overrightarrow{BO'}$ to the length of $\overrightarrow{DE}$. This leads to the following proposition:

Proposition 3: The estimated regression coefficient $\hat{\beta}_1$ is the ratio of the length of the residual vector from projecting $\vec{\hat{y}}$ on $\vec{x}_2$ to the length of the residual vector from projecting $\vec{x}_1$ on $\vec{x}_2$.

With this proposition, we can easily estimate $\hat{\beta}_1$ without drawing the parallelogram $OA_1O'B_1$. Instead, we only need to find the length of the two residual vectors. This conclusion simplifies traditional way of estimating $\hat{\beta}_1$ geometrically.

Similarly, vector $\vec{\hat{y}}$ can be decomposed into two orthogonal components, vector $\vec{\hat{y}}_2$ and vector $(\vec{\hat{y}} - \vec{\hat{y}}_2)$. Since vector $\vec{\hat{y}}_2$ (or $\overrightarrow{OB}$) is perpendicular to $\overrightarrow{DE}$, the projection of $\vec{\hat{y}}$ on $\overrightarrow{DE}$ is equal to the projection of its second component $(\vec{\hat{y}} - \vec{\hat{y}}_2)$ (or $\overrightarrow{BG}$) on $\overrightarrow{DE}$. Therefore, we can also say that:

Proposition 4: The estimated regression coefficient $\hat{\beta}_1$ is equal to the simple regression coefficient obtained by projecting the vector $\vec{y}$ on the residual vector $(\vec{x}_1 - \hat{\vec{x}}_{12})$.

All the propositions above demonstrate the geometric interpretations of the regression coefficient β_1 from different points of view. These visual expressions facilitate the understanding of Frisch-Waugh-Lovell theorem and thus provide a richer interpretation of regression coefficients for readers.

3.2 R^2, the multiple and simple correlation coefficients

In the algebraic method, the coefficient of determination that is used to measure the goodness-of-fit of a model, denoted by R^2, is defined as $R^2 = \frac{\text{SSR}}{\text{SST}}$. In Figure 2, this is the ratio of two sides of the right-angled triangle $OO'G$. That is, $R^2 = \frac{|\hat{\vec{y}}|^2}{|\vec{y}|^2}$.

In algebra, the multiple correlation coefficient between y and x_1 and x_2 is $r_{y \cdot x_1 x_2} = \pm\sqrt{R^2}$. Based on a commonly known result of the geometric interpretation of R^2, we have

$$r_{y \cdot x_1 x_2} = \pm\frac{|\hat{\vec{y}}|}{|\vec{y}|} = \cos\theta (r_{y \cdot x_1 x_2} \geq 0, when \theta \leq 90^\circ or \theta \geq 270^\circ ; r_{y \cdot x_1 x_2} < 0, when 90^\circ < \theta < 270^\circ)$$

where θ is the angle between vector $\vec{y}$ and $\hat{\vec{y}}$, or the angle between vector $\vec{y}$ and the $x_1 x_2$ plane.

This result can be extended into general cases. Suppose α is the angle between two vectors with same dimensions, vector $\vec{z}_1$ and vector $\vec{z}_2$. Then inner product of z_1 and z_2 is

$$< z_1, z_2 >= z_1^T z_2 = |\vec{z}_1||\vec{z}_2| \cos\alpha$$

Then, we have,

$$\cos\alpha = \frac{z_1^T z_2}{|\vec{z}_1||\vec{z}_2|} = \frac{\sum z_1 z_2}{\sqrt{\sum z_1{}^2 \sum z_2{}^2}} = \frac{\text{cov}(z_1, z_2)}{\sqrt{\text{var}(z_1) \cdot \text{var}(z_2)}}$$

= Simple correlation coefficient between $\vec{z}_1$ and $\vec{z}_2$.

It has been shown that the simple correlation coefficient between any two variables can be represented as the cosine of the angle between the two vectors representing the variables.

3.3 Partial Correlation Coefficients

The partial correlation coefficient between y and x_1 is defined in such a way that it measures the effect of x_1 on y which is not accounted for by the x_2 in the model. Therefore, 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 $\hat{\varepsilon}_1$.

Then we regress y on x_2 and obtain the second residual $\hat{\varepsilon}_2$. Finally, we can say that simple correlation coefficient between $\hat{\varepsilon}_2$ and $\hat{\varepsilon}_1$ is the partial correlation coefficient between y and x_1.

Like the multiple and simple correlations, the partial correlation coefficient can also be represented by cosine of some angle. This result was concluded as: partial correlation coefficient between y and x_1 is the cosine of the angle between the "residual vectors", namely, $(\vec{y} - \vec{\hat{y}}_2)$ and $(\vec{x}_1 - \vec{\hat{x}}_{12})$, the components of $\vec{y}$ and $\vec{x}_1$ orthogonal to $\vec{x}_2$. In Figure 4, we can say that vector $\overrightarrow{BG}$ and $\overrightarrow{DE}$ represent the purified y and x_1, or $\hat{\varepsilon}_2$ and $\hat{\varepsilon}_1$. The simple correlation coefficient between $\hat{\varepsilon}_2$ and $\hat{\varepsilon}_1$ is the cosine of the angle between vector $\overrightarrow{BG}$ and $\overrightarrow{DE}$. On the other hand, vector $\overrightarrow{BO'}$, the residual vector from projecting $\vec{\hat{y}}$ on $\vec{x}_2$, is also perpendicular to $\vec{x}_2$ and on the same plane with $\overrightarrow{DE}$. Therefore, we have $\overrightarrow{DE} \mathbin{//} \overrightarrow{BO'}$. And we can say that the angle between $\overrightarrow{BG}$ and $\overrightarrow{DE}$ is the angle between $\overrightarrow{BG}$ and $\overrightarrow{BO'}$, denoted by ϕ_1.

Here we use $r_{yx_1 \cdot x_2}$ to stand for the partial correlation coefficient between y and x_1. That is,

$$r_{yx_1 \cdot x_2} = \frac{\left|\vec{\hat{y}} - \vec{\hat{y}}_2\right|}{\left|\vec{y} - \vec{\hat{y}}_2\right|} = \cos\phi_1$$

To get the partial correlation between y and x_2, simply switch the subscripts of x vectors. If we define the relevant angle as ϕ_2, we have $r_{yx_2 \cdot x_1} = \cos\phi_2$.

If examining Figure 4 further, we can find that ϕ_1 is actually the angle between two planes, Span ($\vec{y}$, $\vec{x}_2$) (i.e. yx_2 plane) and Span ($\vec{x}_1$, $\vec{x}_2$) (i.e. x_1x_2 plane). That is because $\overrightarrow{BG}$ is perpendicular to $\overrightarrow{OB}$ and $\overrightarrow{BO'}$ is also perpendicular to $\overrightarrow{OB}$. This finding leads to the following proposition:

Proposition 5: Partial correlation coefficient between y and x_1 is the cosine of the angle between the subspace spanned by $\vec{y}$ and $\vec{x}_2$, and the subspace spanned by $\vec{x}_1$ and $\vec{x}_2$.

Suppose that the angle between vector $\vec{y}$ and vector $\vec{x}_1$ be θ_1, the angle between vector $\vec{y}$ and vector $\vec{x}_2$ be θ_2, the angle between vector $\vec{\hat{y}}$ and vector $\vec{x}_1$ be γ_1, and the angle between vector $\vec{\hat{y}}$ and vector $\vec{x}_2$ be γ_2 (γ_1 and γ_2 are on the x_1x_2 plane). Then the angle between vector $\vec{x}_1$ and $\vec{x}_2$ is γ. We also define the projection of $\vec{y}$ on $\vec{x}_1$ as $\vec{\hat{y}}_1$ and the projection of $\vec{x}_2$ on $\vec{x}_1$ as $\vec{\hat{x}}_{21}$. The residual vectors are hence $(\vec{y} - \vec{\hat{y}}_1)$ and $(\vec{x}_2 - \vec{\hat{x}}_{21})$, respectively.

3.4 Geometric Expressions for Multiple, Simple and Partial Correlations

Table 1 summarize the geometric expressions for multiple, simple and partial correlations in Figure 2 and Figure 4.

Table 1. Summary for Geometric Expressions for Simple, Partial and Multiple Correlations.

Multiple and Simple Correlations	Geometric Expressions
Multiple correlation between y and x_1 and x_2: $r_{y\cdot x_1x_2}$	$\cos\theta = \frac{\lvert\vec{\hat{y}}\rvert}{\lvert\vec{y}\rvert}$
Simple correlation between y and x_1: r_{yx_1}	$\cos\theta_1 = \frac{\lvert\vec{\hat{y}}_1\rvert}{\lvert\vec{y}\rvert}$
Simple correlation between y and x_2: r_{yx_2}	$\cos\theta_2 = \frac{\lvert\vec{\hat{y}}_2\rvert}{\lvert\vec{y}\rvert}$
Simple correlation between $\hat{y}$ and x_1: $r_{\hat{y}x_1}$	$\cos\gamma_1 = \frac{\lvert\vec{\hat{y}}_1\rvert}{\lvert\vec{\hat{y}}\rvert}$
Simple correlation between $\hat{y}$ and x_2: $r_{\hat{y}x_2}$	$\cos\gamma_2 = \frac{\lvert\vec{\hat{y}}_2\rvert}{\lvert\vec{\hat{y}}\rvert}$
Simple correlation between x_1 and x_2: $r_{x_1x_2}$	$\cos\gamma = \frac{\lvert\vec{\hat{x}}_{12}\rvert}{\lvert\vec{x}_1\rvert} = \frac{\lvert\vec{\hat{x}}_{21}\rvert}{\lvert\vec{x}_2\rvert}$
Partial correlation between y and x_1, given x_2: $r_{yx_1\cdot x_2}$	$\cos\phi_1 = \frac{\lvert\vec{\hat{y}} - \vec{\hat{y}}_2\rvert}{\lvert\vec{y} - \vec{\hat{y}}_2\rvert}$
Partial correlation between y and x_2, given x_1: $r_{yx_2\cdot x_1}$	$\cos\phi_2 = \frac{\lvert\vec{\hat{y}} - \vec{\hat{y}}_1\rvert}{\lvert\vec{y} - \vec{\hat{y}}_1\rvert}$

3.5 Relationship among Simple, Partial, and Multiple Correlation Coefficients

In this section, we examine 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 other statisticians using algebras and matrices which can be found in basic statistics or econometrics texts. However, to the best of our knowledge, they are seldom proved by geometric approaches. Here, we will use geometry to prove the following three equations.

3.5.1 $(1-R^2) = (1-r_{yx_2}^2)(1-r_{yx_1\cdot x_2}^2)$

This result first appears in a basic text by Anderson (1958). Our geometric proof is as follows.

In Figure 6, $\Delta OO'G$, ΔOBG and $\Delta BO'G$ are all right-angled triangles. As described in the previous section, the angle between vectors $\vec{y}$ and $\vec{\hat{y}}$ ($\angle GOO'$) is named as θ, and $R^2 = \cos^2\theta$. The angle between vector $\vec{y}$ and $\vec{x}_2$ ($\angle GOB$) is named as θ_2, and $r_{yx_2}^2 = \cos^2\theta_2$. The angle between vector $\overrightarrow{BG}$ and $\overrightarrow{BO'}$ is named as ϕ_1, and $r_{yx_1\cdot x_2}^2 = \cos^2\phi_1$.

We find that in the right-angled triangle $\Delta OO'G$, $\sin\theta = \dfrac{|\overrightarrow{O'G}|}{|\overrightarrow{OG}|}$; in the right-angled triangle ΔOBG, $\sin\theta_2 = \dfrac{|\overrightarrow{BG}|}{|\overrightarrow{OG}|}$; in the right-angled triangle $\Delta BO'G$, $\sin\phi_1 = \dfrac{|\overrightarrow{O'G}|}{|\overrightarrow{BG}|}$;

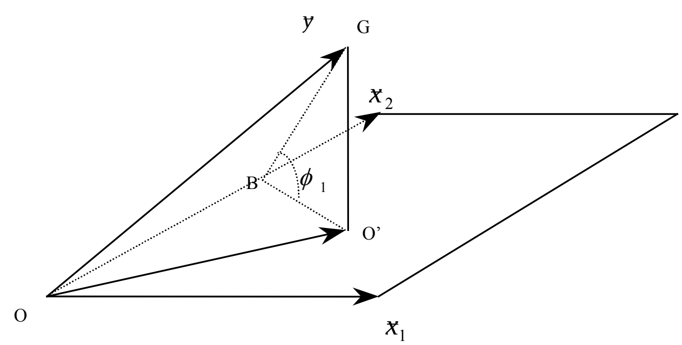

Figure 6. Relationship among Simple, Partial and Multiple Correlations.

Therefore, we have

$$\sin\theta_2 \cdot \sin\phi_1 = \frac{|\overrightarrow{BG}|}{|\overrightarrow{OG}|} \cdot \frac{|\overrightarrow{O'G}|}{|\overrightarrow{BG}|} = \frac{|\overrightarrow{O'G}|}{|\overrightarrow{OG}|} = \sin\theta, \qquad \text{or,} \qquad \sin\theta = \sin\theta_2 \cdot \sin\phi_1$$

Square both sides, we have, $\sin^2\theta = \sin^2\theta_2 \cdot \sin^2\phi_1$

Or, $(1-\cos^2\theta) = (1-\cos^2\theta_2)\cdot(1-\cos^2\phi_1)$

That is, $(1-R^2) = (1-r_{yx_2}^2)(1-r_{yx_1\cdot x_2}^2)$ (3.5-1-1)

Equation (3.5-1-1) discloses the relationship among the simple, partial and multiple correlation coefficients. And the process of deriving this equation totally relies on simple geometric method.

By symmetry, we have another equation, $(1-R^2) = (1-r_{yx_1}^2)(1-r_{yx_2\cdot x_1}^2)$ (3.5-1-2)

Another form of these two equations is $R^2 = r_{yx_1}^2 + (1-r_{yx_1}^2)r_{yx_2\cdot x_1}^2$ (3.5-1-3)

And, $R^2 = r_{yx_2}^2 + (1-r_{yx_2}^2)r_{yx_1\cdot x_2}^2$ (3.5-1-4)

Damodar N. Gujarati (1995) provides explanations for these two equations. Equation (3.5-1-3) 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 ($=r_{yx_1}^2$) and the part not explained by x_1 ($=1-r_{yx_1}^2$) times the proportion that is explained by x_2 after eliminating the influence of x_1. Equation (3.5-1-4), similarly, 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_2 alone ($=r^2_{yx_2}$) and the part not explained by x_2 (=1-$r^2_{yx_2}$) times the proportion that is explained by x_1 after eliminating the influence of x_2.

3.5.2 $r_{y\cdot x_2} = r_{y\cdot x_1 x_2} r_{\hat{y}\cdot x_2}$, and $r_{y\cdot x_1} = r_{y\cdot x_1 x_2} r_{\hat{y}\cdot x_1}$

In Figure 7, there exists another relationship among the three vertex angles, θ (the angle between vector $\vec{y}$ and $\vec{\hat{y}}$), θ_2 (the angle between vector $\vec{y}$ and vector $\vec{x}_2$), and γ_2 (the angle between vector $\vec{\hat{y}}$ and vector $\vec{x}_2$).

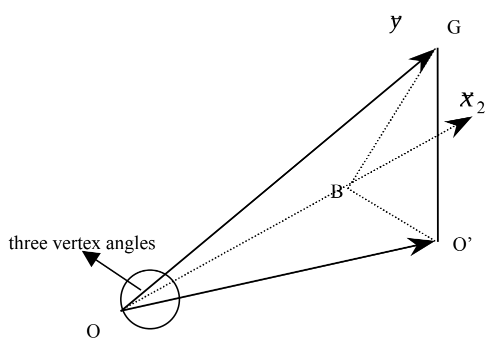

Figure 7. Relationship between Simple and Multiple Correlation.

In the right-angled triangle ΔOBG, $|\vec{OB}| = |\vec{y}|\cos\theta_2$; in the right-angled triangle $\Delta OBO'$, $|\vec{OB}| = |\vec{\hat{y}}|\cos\gamma_2$; in the right-angled triangle $\Delta OO'G$, $|\vec{\hat{y}}| = |\vec{y}|\cos\theta$.

Therefore, $|\vec{y}|\cos\theta_2 = |\vec{y}|\cos\theta\cos\gamma_2$.

Then, $\cos\theta_2 = \cos\theta\cos\gamma_2$, or, $r_{y\cdot x_2} = r_{y\cdot x_1 x_2} r_{\hat{y}\cdot x_2}$ (3.5-2-1)

By symmetry, we can obtain another equation, $r_{y\cdot x_1} = r_{y\cdot x_1 x_2} r_{\hat{y}\cdot x_1}$ (3.5-2-2)

3.5.3 $r_{yx_1\cdot x_2} = \dfrac{r_{yx_1} - r_{yx_2} r_{x_1x_2}}{\sqrt{(1-r^2_{yx_2})(1-r^2_{x_1x_2})}}$

This is a well-known algebraic formula to calculate partial correlation coefficient based on simple correlation coefficients. The algebraic proof can be easily found in a basic statistical text. But the proof is complicated. In fact, geometry is very helpful in interpreting this relation.

Kendall and Stuart (1973) developed a geometric proof of this formula. It was also discussed by Guy Thomas and John O'quigley in their paper "A geometric interpretation of partial correlation using spherical triangles" (1993). They used spherical triangles to give a geometric interpretation of this relationship. Their method is more illustrating because it shows that the above formula is identical with the formula of spherical trigonometry. However, to fully understand their interpretations, one need to have some geometry background in a higher level, specifically, the understanding of spherical trigonometry. Our proof here is simple and easy to understand.

Proof: In section (3.5-2), we have obtained the results $r_{y\cdot x_2} = r_{y\cdot x_1x_2} r_{\hat{y}\cdot x_2}$ and $r_{y\cdot x_1} = r_{y\cdot x_1x_2} r_{\hat{y}\cdot x_1}$. Or equivalently, $\cos\theta_1 = \cos\theta\cdot\cos\gamma_1$ (3.5-3-1)

And, $\cos\theta_2 = \cos\theta\cdot\cos\gamma_2$ (3.5-3-2)

On the other hand, in the right-angled triangle $\Delta BO^{'}G$ in Figure 7, we have

$$\cos\phi_1 = \frac{\left|\vec{BO'}\right|}{\left|\vec{BG}\right|} = \frac{\left|\vec{y}\right|\sin\gamma_2}{\left|\vec{y}\right|\sin\theta_2} = \frac{\cos\theta\cdot\sin\gamma_2}{\sin\theta_2} \quad (3.5\text{-}3\text{-}3)$$

At the same time, since $\gamma_1 = \gamma - \gamma_2$, and then we have,

$$\begin{aligned}\cos\theta_1 &= \cos\theta\cdot\cos\gamma_1 \quad (3.5\text{-}3\text{-}4)\\ &= \cos\theta\cdot\cos(\gamma-\gamma_2)\\ &= \cos\theta\cdot(\cos\gamma\cos\gamma_2 - \sin\gamma\sin\gamma_2)\\ &= \cos\gamma\cos\theta\cos\gamma_2 - \sin\gamma\cos\theta\sin\gamma_2\end{aligned}$$

Substitute (3.5-3-2) into (3.5-3-4), we obtain, $\cos\theta\sin\gamma_2 = \dfrac{\cos\theta_1 - \cos\theta_2\cos\gamma}{\sin\gamma}$ (3.5-3-5)

Substitute (3.5-3-5) into (3.5-3-3), we obtain,

$$\cos\phi_1 = \frac{\cos\theta_1 - \cos\theta_2\cos\gamma}{\sin\theta_2\sin\gamma} \quad \text{or,} \quad \cos\phi_1 = \frac{\cos\theta_1 - \cos\theta_2\cos\gamma}{\sqrt{(1-\cos^2\theta_2)}\sqrt{(1-\cos^2\gamma)}}$$

Thus, we use geometry to prove $r_{yx_1\cdot x_2} = \dfrac{r_{yx_1} - r_{yx_2} r_{x_1x_2}}{\sqrt{(1-r_{yx_2}^2)(1-r_{x_1x_2}^2)}}$.

4. CONCLUSION

Regression analysis is traditionally presented in algebraic forms, especially in equations and matrices. Due to the predominance of algebra and the unpopularity of geometry, not much insightful work has been done to apply geometry in regression analysis. To correct a graphing problem in the previous research, the matrix established in this paper ensures the validity of drawing the n-dimensional vectors, angles and triangles into a 3-dimensional space. Based on this strictly correct 3-dimensional framework, we not only summarize the previous results about geometric interpretations of the least squares method, regression coefficients, simple, multiple and partial correlation coefficients, but also propose the geometric proofs to the Frisch-Waugh-Lovell Theorem, and three classical formulas that display the relationship among simple, multiple and partial correlation coefficients. Although these formulas are well known in the basic statistics texts and have already been proved with the use of algebraic methods, we are among the first to give the geometric proof. All these geometric proof proposed in this paper are concise and easy to understand for beginners. In fact, all concepts of regression analysis can be visualized by applying a few principles of geometry. In this paper, some fundamental theorems and concepts, such as Frisch-Waugh-Lovell and partial correlation coefficients, are restated from the perspective of geometry. In its sister paper, we will also provide geometric interpretations of principle component analysis and regression test statistics. In our opinion, the geometric approach sheds lights on the regression analysis as it provides a richer and concrete understanding for readers, especially for beginners.

REFERENCES

Anderson, T.W. (1958), "An Introduction to Multivariate Statistical Analysis," John Wiley & Sons, Inc., New York

Belov AG and Shchedrin BM. (2014). "Linear Estimation with Regressor Decomposition, " Computational Mathematics and Modeling, 25: 79-86

Box, G.E.P., Hunter, W.G., and Hunter, J. S. (1978), "Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building," New York: John Wiley.

Bring, J. (1994), "How to Standardize Regression Coefficients," The American Statistician, 48(3): 209-213

Bring, J., (1996), "A Geometric Approach to Compare Variables in a Regression Model," The American Statistician, 50:57-62

Bryant, P. (1984), "Geometry, Statistics, Probability: Variations on a Common Theme," The American Statistician, 38:38-48

Davidson, R., & MacKinnon, J. G. (2004). "Econometric theory and methods," New York: Oxford University Press, Vol. 5.

Davis, P. J., & Hersh, R. (1998), "The mathematical experience," Houghton Mifflin Harcourt.

Fisher (1915), "Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population," Biometrika, 10, 507–21.

Frisch, R., & Waugh, F. V. (1933), "Partial time regressions as compared with individual trends," Econometrica: Journal of the Econometric Society, 387-401.

Gujarati, D.N. (1995), "Basic Econometrics," 3rd edition. McGraw-Hill.

Herr, D.G. (1980), "On the History of the Use of Geometry in the General Linear Model," The American Statistician, 34:43-47

Kendal, M. G. and Stuart, A. (1973), "The Advanced Theory of Statistics," Vol. 1. London: C, Griffin.

Kruskal, William (1975), "The geometry of generalized inverses," Journal of the Royal Statistical Society. Series B (Methodological), 272-283.

Lovell, M. C. (1963), "Seasonal adjustment of economic time series and multiple regression analysis," Journal of the American Statistical Association, 58(304), 993-1010.

Lovell, M. C. (2008), "A simple proof of the FWL theorem," The Journal of Economic Education, 39(1), 88-91.

Margolis, M.S. (1979), "Perpendicular Projections and Elementary Statistics," The American Statistician, 33:131-135

Marks, Edmond (1982), "A Note on a Geometric Interpretation of the Correlation Coefficient," Journal of Educational and Behavioral Statistics, 7.3: 233-237

Morrison, D.F. (1983), "Applied Linear Statistical Methods," New Jersey: Prentice-Hall

Saville, D.J., and Wood, G.R. (1986), "A Method for Teaching Statistics Using N-Dimensional Geometry," The American Statistician, 40: 205-214

Saville, D. J., and Wood, G. R. (1991), "Statistical Methods: The Geometric Approach," New York: Springer-Verlag.

Saville, D. J., and Wood, G. R. (1996), "Statistical Methods: A Geometric Primer," New York: Springer-Verlag.

Sosa Escudero, W. (2001), "A geometric representation of the Frisch-Waugh-Lovell Theorem," Documentos de Trabajo.

Thomas, G. and O'quigley, J. (1993), "A Geometric Interpretation of Partial Correlation Using Spherical Triangles," The American Statistician, 47:30-32