ISSN 1995-0802, Lobachevskii Journal of Mathematics, 2012, Vol. 33, No. 3, pp. 284–292. © Pleiades Publishing, Ltd., 2012.

Few Remarks on the Geometry of the Uncentered Coefficient of Determination

Umberto Triacca[1,2*] **and Andrei Volodin**[1,2**]

[1]*Department of Economics, University of L'Aquila, P.le del Santuario 19, Roio Poggio, Aquila, 67040 Italy*

[2]*Department of Mathematics and Statistics, University of Regina, Regina, Saskatchewan, S4S 0A2 Canada*

Abstract—This paper investigates the geometrical properties of the uncentered coefficient of determination in linear regression analysis.

DOI: 10.1134/S1995080212030158

Keywords and phrases: *Uncentered coefficient of determination, Least squares, Vector space, Orthogonality.*

1. INTRODUCTION AND DEFINITIONS

An application of the Euclidean space geometry is a powerful tools for a linear regression study. Most of the ideas of the regression theory are simplified by understanding their n-dimensional geometry. A descriptive statistic that is widely used to determine how well a regression fits is the coefficient of determination. In this paper we investigate the geometrical properties of the uncentered coefficient of determination in linear regression analysis.

Consider a vector of "observations" $\mathbf{y}$ in n-dimensional Euclidean space E^n and a $n \times k$ "data matrix" $\mathbf{X} = [\mathbf{x}_1, ..., \mathbf{x}_k]$ each column $\mathbf{x}_i$ of which is a vector in E^n.

We can always write

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{u},$$

where $\beta \in \mathrm{E}^k$ and $\mathbf{u}$ is the difference between $\mathbf{y}$ and $\mathbf{X}\beta$. We are interested in finding the β such that $\mathbf{y}$ is as close as possible to $\mathbf{X}\beta$ in the sense that $\mathbf{u}$ is as short as possible.

This is the well known least squares (LS) problem. In order to discuss this problem in more formal terms, we briefly discuss definitions from linear algebra. Certainly, the reader knows them very well and our main goal here is to fix the notations.

If $M = \{\mathbf{v}_1, ..., \mathbf{v}_h\}$ is a subset of E^n, then the set of all $\alpha_1\mathbf{v}_1 + ... + \alpha_h\mathbf{v}_h$ ($\alpha_1, ..., \alpha_h$ arbitrary real numbers) is called the *linear manifold spanned by* M and is symbolized by sp(M). Let $\mathbf{A}$ be a $n \times m$ real matrix. The subspace of E^n spanned by the columns of $\mathbf{A}$, denoted with col($\mathbf{A}$), consists of all points $\mathbf{z}$ in E^n such that $\mathbf{z} = \mathbf{A}\delta$ where $\delta \in \mathrm{E}^m$. This subspace is called the *column space of matrix* $\mathbf{A}$. We note that $\mathrm{col}(\mathbf{X}) = sp\,(\{\mathbf{x}_1, ..., \mathbf{x}_k\})$.

The *inner product* of any two vectors, $\mathbf{v}$ and $\mathbf{w}$, in E^n, denoted by $(\mathbf{v}, \mathbf{w})$, is $(\mathbf{v}, \mathbf{w}) = v^T\mathbf{w} = \mathbf{w}^T\mathbf{v}$. The *length or norm* of a vector $\mathbf{w}$ in E^n, denoted by $||\mathbf{w}||$, is $||\mathbf{w}|| = \sqrt{\mathbf{w}^T\mathbf{w}}$.

The quantity θ defined by

$$\theta = \arccos\left\{\frac{\langle \mathbf{y}, \mathbf{x}\rangle}{||\mathbf{y}||\,||\mathbf{x}||}\right\}, \quad 0 \le \theta \le \pi$$

*E-mail: triacca@ec.univaq.it

**E-mail: andrei@math.uregina.ca

The work of A. Volodin is supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

is called the *angle* between the vectors $\mathbf{y}$ and $\mathbf{x}$ in E^n.

Two vectors $\mathbf{v}$ and $\mathbf{w}$ in E^n are said to be *orthogonal* if $\mathbf{v}^T\mathbf{w} = \mathbf{w}^T\mathbf{v} = 0$, we denote this by $\mathbf{v} \perp \mathbf{w}$. A $\mathbf{w}$ in E^n is *orthogonal to a subset* M of E^n, and we write $\mathbf{w} \perp M$, if $\mathbf{w} \perp \mathbf{v}$ for all $\mathbf{v} \in M$. Two subset M, N of E^n are *orthogonal*, and we write $M \perp N$, if $\mathbf{w} \perp \mathbf{v}$ for every $\mathbf{w} \in M$, $\mathbf{v} \in N$. If $M \subset \mathrm{E}^n$, then the set $M^\perp = \{\mathbf{v} \in \mathrm{E}^n; (\mathbf{v}, \mathbf{w}) = 0, \text{ for every } \mathbf{w} \in M\}$ is called the *orthogonal complement* of M.

The *distance* between two vectors, $\mathbf{v}$ and $\mathbf{w}$, in E^n, denoted by $d(\mathbf{v}, \mathbf{w})$, is $d(\mathbf{v}, \mathbf{w}) = ||\mathbf{v} - \mathbf{w}||$. For any $\mathbf{v}$ in E^n and any subset $S \subset \mathrm{E}^n$, the *distance* $d(\mathbf{v}, S)$, is defined by

$$d(\mathbf{v}, S) = \inf\{||\mathbf{v} - \mathbf{w}||;\ \mathbf{w} \in S\}.$$

Let $M = \{\mathbf{v}_1, ..., \mathbf{v}_h\}$ be a subset of E^n. For a given $\mathbf{v} \in \mathrm{E}^n$, we define the *orthogonal projection* of $\mathbf{v}$ on $\mathrm{sp}(M)$, denoted by $\mathrm{P}(\mathbf{v}|\mathrm{sp}(M))$ as the unique element of $\mathrm{sp}(M)$ such that $||\mathbf{v} - \mathrm{P}(\mathbf{v}|\mathrm{sp}(M))|| \leq ||\mathbf{v} - \mathbf{w}||$ for any $\mathbf{w} \in \mathrm{sp}(M)$.

Let V be a vector space over a field $\mathbf{R}$. A *linear functional* (or *linear form*) on V is a mapping $T\colon V \to \mathbf{R}$ satisfying the following two conditions: $T(x + y) = T(x) + T(y)$ and $T(cx) = cT(x)$ for any $c \in \mathbf{R}$, $x, y \in V$. We say that $T\colon V \to \mathbf{R}$ is a *bounded linear functional* if there exists an $M \in \mathbf{R}^+$ such that $|T(x)| \leq M\,||x||$ for all $x \in V$. The quantity

$$||T|| = \sup_x \frac{|T(x)|}{||x||} = \sup_{||x||=1} |T(x)|$$

is called a *norm* of the functional T.

With the preliminaries accounted for, now we are ready to formulate the LS problem as follows:

$$\min_\beta ||\mathbf{y} - \mathbf{X}\beta||. \tag{1}$$

The following formula provides the well known LS Solution Property: $\hat{\beta}$ solves the LS problem (1.1) if and only if

$$\mathbf{X}^T(\mathbf{y} - \mathbf{X}\hat{\beta}) = \mathbf{0}. \tag{2}$$

The proof of this theorem can be found in Björck, 1996, p.5, for example.

The condition (1.2) says that the vector $\mathbf{y} - \mathbf{X}\hat{\beta}$ must be orthogonal to all of the columns of $\mathbf{X}$ and hence to any vectors that lies in the space spanned by those columns, that is $(\mathbf{y} - \mathbf{X}\hat{\beta}) \perp \mathrm{col}(\mathbf{X})$.

As a corollary to we obtain the LS problem solution: If $rank(\mathbf{X}) = k$ then (1.1) has a unique LS solution, given by

$$\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}. \tag{3}$$

In the following we assume that $rank(\mathbf{X}) = k$.

If we substitute the right-hand side of (1.3) for $\hat{\beta}$ into $\mathbf{X}\hat{\beta}$, we obtain

$$\mathbf{X}\hat{\beta} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \equiv \mathbf{P}_\mathbf{X}\mathbf{y}.$$

This equation defines the $n \times n$ matrix $\mathbf{P}_\mathbf{X} \equiv \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ also know as *hat matrix*. The hat matrix is the orthogonal projector onto $\mathrm{col}(\mathbf{X})$, that is, $\mathbf{P}_\mathbf{X}\mathbf{y} = \mathrm{P}(\mathbf{y}|\mathrm{sp}(\{\mathbf{x}_1, ..., \mathbf{x}_k\}))$.

Through this paper we will use $\mathbf{P}$ subscripted by matrix expression to denote the matrix that projects onto the subspace spanned by the columns of that matrix expression. The matrix that projects orthogonally the vector $\mathbf{y}$ on $\mathrm{col}(\mathbf{X})^\perp$ (the orthogonal complement of $\mathrm{col}(\mathbf{X})$) is

$$\mathbf{M}_\mathbf{X} \equiv \mathbf{I} - \mathbf{P}_\mathbf{X} \equiv \mathbf{I} - \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$$

where $\mathbf{I}$ is the $n \times n$ identity matrix. We have that

$$\mathbf{P}_\mathbf{X} + \mathbf{M}_\mathbf{X} = \mathbf{I}.$$

Thus

$$\mathbf{y} = \mathbf{P_X y} + \mathbf{M_X y}$$

with $\mathbf{P_X y} \perp \mathbf{M_X y}$. By the Pythagorean theorem, we have that

$$||\mathbf{y}||^2 = ||\mathbf{P_X y}||^2 + ||\mathbf{M_X y}||^2$$

Definition. The *uncentered coefficient of determination* R_u^2 is defined as

$$\mathrm{R}_u^2 = \frac{||\mathbf{P_X} y||^2}{||\mathbf{y}||^2}.$$

Note that the sum of squares

$$||\mathbf{y}||^2 = \mathbf{y}^T \mathbf{y}$$

is a measure of the variability of the observations. On the other hand we have that

$$||\mathbf{P_X y}||^2 = \beta^T \mathbf{X}^T \mathbf{X} \beta.$$

Thus, R_u^2 has the interpretation of the fraction of the variation of the observations that is attributable to the variation in the data matrix.

The following is the simple geometrical interpretation of the uncentered coefficient of determination. We show that R_u^2 is equal to the square of the cosine of the angle between $\mathbf{y}$ and $\mathbf{X}\hat{\beta}$. In fact, we have

$$\cos(\theta) = \frac{\left\langle \mathbf{y}, \mathbf{X}\hat{\beta} \right\rangle}{||\mathbf{y}|| \left|\left|\mathbf{X}\hat{\beta}\right|\right|} = \frac{\langle \mathbf{y}, \mathbf{P_X y} \rangle}{||\mathbf{y}|| \left|\left|\mathbf{X}\hat{\beta}\right|\right|} = \frac{\langle \mathbf{P_X y} + \mathbf{M_X y}, \mathbf{P_X y} \rangle}{||\mathbf{y}|| \left|\left|\mathbf{X}\hat{\beta}\right|\right|} = \frac{\langle \mathbf{P_X y}, \mathbf{P_X y} \rangle}{||\mathbf{y}|| \left|\left|\mathbf{X}\hat{\beta}\right|\right|}$$

$$= \frac{||\mathbf{P_X y}||^2}{||\mathbf{y}|| \left|\left|\mathbf{X}\hat{\beta}\right|\right|} = \frac{||\mathbf{P_X y}||}{||\mathbf{y}||} = \mathrm{R}_u.$$

The next proposition provides an interpretation of R_u as a bounded linear functional.

Proposition. There exists a bounded linear functional $T \colon \mathrm{E}^n \to \mathbf{R}$ with $||T|| = 1/\,||\mathbf{y}||$ such that $\mathrm{R}_u^2 = [T(\mathbf{y})]^2$.

Proof. Consider the functional $T \colon \mathrm{E}^n \to \mathbf{R}$ defined by

$$T(\mathbf{v}) = \frac{\langle \mathbf{v}, \mathbf{P_X y} \rangle}{||\mathbf{y}||\,||\mathbf{P_X y}||}.$$

We have that T is a linear functional because for any α, $\beta \in \mathbf{R}$, $\mathbf{v}_1, \mathbf{v}_2 \in \mathrm{E}^n$

$$T(\alpha \mathbf{v}_1 + \beta \mathbf{v}_2) = \alpha \frac{\langle \mathbf{v}_1, \mathbf{P_X y} \rangle}{||\mathbf{y}||\,||\mathbf{P_X y}||} + \beta \frac{\langle \mathbf{v}_2, \mathbf{P_X y} \rangle}{||\mathbf{y}||\,||\mathbf{P_X y}||} = \alpha T(\mathbf{v}_1) + \beta T(\mathbf{v}_2).$$

Since for all $\mathbf{v} \in \mathrm{E}^n : |\langle \mathbf{v}, \mathbf{P_X y} \rangle| \leq ||\mathbf{v}||\,||\mathbf{P_X y}||$, we have that

$$\frac{|\langle \mathbf{v}, \mathbf{P_X y} \rangle|}{||\mathbf{P_X y}||} \leq ||\mathbf{v}||\,,$$

and hence

$$|T(\mathbf{v})| = \frac{|\langle \mathbf{v}, \mathbf{P_X y} \rangle|}{||\mathbf{y}||\,||\mathbf{P_X y}||} \leq \frac{1}{||\mathbf{y}||}\,||\mathbf{v}||\,,$$

which means that T is a bounded linear functional. The norm of T is

$$||T|| = \sup_{\mathbf{v}} \frac{|T(\mathbf{v})|}{||\mathbf{v}||} \leq \frac{1}{||\mathbf{y}||}.$$

Next, because $|\langle \mathbf{P_X y}, \mathbf{P_X y}\rangle| = ||\mathbf{P_X y}||^2$ and $|T(\mathbf{P_X y})| = ||\mathbf{P_X y}|| / ||\mathbf{y}||$, we conclude that $||T|| = 1/\,||\mathbf{y}||$.

Finally, we note that

$$T(\mathbf{y}) = \frac{\langle \mathbf{y}, \mathbf{P_X y}\rangle}{||\mathbf{y}||\,||\mathbf{P_X y}||} = \mathrm{R}_u.$$

2. MAIN RESULTS

Obviously, $0 \le \mathrm{R}_u^2 \le 1$. What is the geometrical interpretation of the condition $\mathrm{R}_u^2 = 0$? The first simple interpretation according to the fact proved just above is that $\mathrm{R}_u^2 = 0$ if and only if $\mathbf{y}$ and $\mathbf{X}\hat{\beta}$ are orthogonal.

Now, we provide two more geometrical characterizations of the condition $\mathrm{R}_u^2 = 0$. In order to do this we consider the QR-decomposition of $\mathbf{X}$, that is

$$\mathbf{X} = \mathbf{QR},$$

where $\mathbf{R}$ is an $m \times n$ upper triangular matrix and $\mathbf{Q} = [\mathbf{q}_1, ..., \mathbf{q}_m]$ is an orthogonal $n \times m$ matrix, i.e., one satisfying $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$. The condition $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$ implies that the columns, $\mathbf{q}_i$, of $\mathbf{Q}$ are orthonormal, that is $||\mathbf{q}_i|| = 1$ and $\langle \mathbf{q}_i, \mathbf{q}_j\rangle = 0$ for $i \neq j$.

We note that

$$\mathbf{P_X} \equiv \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T = \mathbf{QR}(\mathbf{R}^T\mathbf{Q}^T\mathbf{QR})^{-1}\mathbf{R}^T\mathbf{Q}^T = \mathbf{QR}(\mathbf{R}^T\mathbf{R})^{-1}\mathbf{R}^T\mathbf{Q}^T$$
$$= \mathbf{QRR}^{-1}(\mathbf{R}^{-1})^T\mathbf{R}^T\mathbf{Q}^T = \mathbf{QQ}^T.$$

Hence

$$\mathrm{R}_u^2 = \frac{||\mathbf{QQ}^T\mathbf{y}||^2}{||\mathbf{y}||^2}.$$

Now we can provide geometrical characterizations of the “extreme” cases: three conditions for $\mathrm{R}_u^2 = 0$ (Theorems 1, 2, and 3) and one for $\mathrm{R}_u^2 = 1$ (Theorem 5). We also prove a sufficient condition for $\mathrm{R}_u^2 = 1$ (Theorem 4). Next, we provide a result when the coefficient of determination can be calculated by a simple geometrical formula, that is, as the unit minus the square distance between the response vector $\mathbf{y}$ and the column vectors of matrix $\mathbf{Q}$ (Theorem 6).

Theorem 1. *The condition $R_u^2 = 0$ is satisfied if and only if $d(\mathbf{y}, \mathrm{col}(\mathbf{Q})) = ||\mathbf{y}||$.*

Proof. ($\Rightarrow$) We note that

$$d(\mathbf{y}, \mathrm{col}(\mathbf{Q})) = \inf\{||\mathbf{y} - \mathbf{q}||; \quad \mathbf{q} \in \mathrm{col}(\mathbf{Q})\} = ||\mathbf{y} - \mathbf{P_Q}y||.$$

If $\mathrm{R}_u^2 = 0$, then $||\mathbf{QQ}^T\mathbf{y}|| = 0$ and hence $\mathbf{y} \perp \mathrm{col}(\mathbf{Q})$. Now, if $\mathbf{y} \perp \mathrm{col}(\mathbf{Q})$ we have that $\mathbf{P_Q y} = \mathbf{0}$ and hence $\inf\{||\mathbf{y} - \mathbf{q}||; \mathbf{q} \in \mathrm{col}(\mathbf{Q})\} = ||\mathbf{y}||$. Thus we can conclude that $d(\mathbf{y}, \mathrm{col}(\mathbf{Q})) = ||\mathbf{y}||$.

($\Leftarrow$) If $\mathrm{R}_u^2 \neq 0$, then $||\mathbf{QQ}^T\mathbf{y}|| \neq 0$ and hence there exists $\mathbf{q} \in \mathrm{col}(\mathbf{Q})$ such that $\langle \mathbf{y}, \mathbf{q}\rangle \neq 0$. Thus $\langle \mathbf{y}, \mathbf{q}\rangle^2 > 0$. Now, we note that $\Delta = 4\langle \mathbf{y}, \mathbf{q}\rangle^2$ is the discriminant of the polynomial in λ

$$P(\lambda) = ||\mathbf{q}||^2\lambda^2 - 2\langle \mathbf{y}, \mathbf{q}\rangle\lambda.$$

Since $||\mathbf{q}|| > 0$ and $\Delta > 0$ we have that

$$||\mathbf{q}||^2\lambda^2 - 2\langle \mathbf{y}, \mathbf{q}\rangle\lambda < 0 \quad \forall \lambda \in (\lambda_1, \lambda_2),$$

where

$$\lambda_1 = \begin{cases} 0 \text{ if } \langle \mathbf{y}, \mathbf{q}\rangle > 0 \\ 2\langle \mathbf{y}, \mathbf{q}\rangle / ||\mathbf{q}|| \text{ if } \langle \mathbf{y}, \mathbf{q}\rangle < 0 \end{cases}$$

and

$$\lambda_2 = \begin{cases} 2\langle \mathbf{y}, \mathbf{q}\rangle / ||\mathbf{q}|| & \text{if } \langle \mathbf{y}, \mathbf{q}\rangle > 0 \\ 0 \text{ if } \langle \mathbf{y}, \mathbf{q}\rangle < 0 \end{cases} .$$

Now we remind that

$$||\mathbf{y} - \lambda\mathbf{q}||^2 = \langle \mathbf{y}-\lambda\mathbf{q}, \mathbf{y}-\lambda\mathbf{q}\rangle = \langle \mathbf{y}, \mathbf{y}\rangle - 2\lambda \langle \mathbf{y}, \mathbf{q}\rangle + \lambda^2 \langle \mathbf{q}, \mathbf{q}\rangle = ||\mathbf{y}||^2 - 2\lambda \langle \mathbf{y}, \mathbf{q}\rangle + ||\mathbf{q}||^2 \lambda^2$$

and hence

$$||\mathbf{q}||^2 \lambda^2 - 2\lambda \langle \mathbf{y}, \mathbf{q}\rangle = ||\mathbf{y} - \lambda\mathbf{q}||^2 - ||\mathbf{y}||^2 .$$

Thus the condition

$$||\mathbf{y} - \lambda\mathbf{q}|| < ||\mathbf{y}||$$

is equivalent to

$$||\mathbf{q}||^2 \lambda^2 - 2\lambda \langle \mathbf{y}, \mathbf{q}\rangle < 0.$$

Hence, we can conclude that

$$||\mathbf{y} - \lambda\mathbf{q}|| < ||\mathbf{y}|| \text{ for all } \lambda \in (\lambda_1, \lambda_2).$$

This implies that $d(\mathbf{y}, \text{col}(\mathbf{Q})) < ||\mathbf{y}||$.

Theorem 2. *The condition $R_u^2 = 0$ is satisfied if and only if*

$$d(\{\mathbf{q}_1, ..., \mathbf{q}_m\}, \text{sp}(\mathbf{y})) = 1.$$

Proof. ($\Rightarrow$) We note that

$$\{||\mathbf{q} - \mathbf{z}|| ; \quad \mathbf{q} \in \{\mathbf{q}_1, ..., \mathbf{q}_m\}, \quad \mathbf{z} \in \text{sp}(\mathbf{y})\} = \bigcup_{i=1}^{m} \{||\mathbf{q}_i - \mathbf{z}|| ; \quad \mathbf{z} \in \text{sp}(\mathbf{y})\}$$

and

$$\inf \{||\mathbf{q}_i - \mathbf{z}|| ; \quad \mathbf{z} \in \text{sp}(\mathbf{y})\} = ||\mathbf{q}_i - \text{P}(\mathbf{q}_i|\text{sp}(\mathbf{y}))|| .$$

If $\text{R}_u^2 = 0$, then $\left|\left|\mathbf{Q}\mathbf{Q}^T\mathbf{y}\right|\right| = 0$ and hence $\mathbf{q}_i \perp \text{sp}(\mathbf{y})$. But then we have that $\text{P}\mathbf{q}_i|\text{sp}(\mathbf{y}) = 0$ and hence $\inf\{||\mathbf{q}_i - \mathbf{z}||; \mathbf{z} \in \text{sp}(\mathbf{y})\} = ||q_i|| = 1$ for all i. We conclude that $d(\{\mathbf{q}_1, ..., \mathbf{q}_m\}, \text{sp}(\mathbf{y})) = 1$.

($\Leftarrow$) If $\text{R}_u^2 \neq 0$, then $\left|\left|\mathbf{Q}\mathbf{Q}^T\mathbf{y}\right|\right| \neq 0$ and hence there exist a $\mathbf{q} \in \{\mathbf{q}_1, ..., \mathbf{q}_m\}$ such that $\langle \mathbf{y}, \mathbf{q}\rangle \neq 0$. Thus $\langle \mathbf{y}, \mathbf{q}\rangle^2 > 0$. Now, we note that $\Delta = 4\langle \mathbf{y}, \mathbf{q}\rangle^2$ is the discriminant of the polynomial in λ

$$P(\lambda) = ||\mathbf{y}||^2 \lambda^2 - 2\langle \mathbf{y}, \mathbf{q}\rangle \lambda.$$

Since $||\mathbf{y}|| > 0$ and $\Delta > 0$ we have that

$$||\mathbf{y}||^2 \lambda^2 - 2\langle \mathbf{y}, \mathbf{q}\rangle \lambda < 0 \text{ for all } \lambda \in (\lambda_1, \lambda_2),$$

where

$$\lambda_1 = \begin{cases} 0 \text{ if } \langle \mathbf{y}, \mathbf{q}\rangle > 0 \\ 2\langle \mathbf{y}, \mathbf{q}\rangle / ||\mathbf{y}|| & \text{if } \langle \mathbf{y}, \mathbf{q}\rangle < 0 \end{cases}$$

and

$$\lambda_2 = \begin{cases} 2\langle \mathbf{y}, \mathbf{q}\rangle / ||\mathbf{y}|| & \text{if } \langle \mathbf{y}, \mathbf{q}\rangle > 0 \\ 0 \text{ if } \langle \mathbf{y}, \mathbf{q}\rangle < 0 \end{cases} .$$

Now we remind that

$$||\mathbf{q} - \lambda\mathbf{y}||^2 = \langle \mathbf{q}-\lambda\mathbf{y}, \mathbf{q}-\lambda\mathbf{y}\rangle = \langle \mathbf{q}, \mathbf{q}\rangle - 2\lambda \langle \mathbf{y}, \mathbf{q}\rangle + \lambda^2 \langle \mathbf{y}, \mathbf{y}\rangle = ||\mathbf{q}||^2 - 2\lambda \langle \mathbf{y}, \mathbf{q}\rangle + ||\mathbf{y}||^2 \lambda^2$$

and hence

$$||\mathbf{y}||^2 \lambda^2 - 2\lambda \langle \mathbf{y}, \mathbf{q}\rangle = ||\mathbf{q} - \lambda\mathbf{y}||^2 - ||\mathbf{q}||^2 .$$

Thus the condition

$$||\mathbf{q}-\lambda\mathbf{y}|| < ||\mathbf{q}||$$

is equivalent to

$$||\mathbf{y}||^2\,\lambda^2 - 2\lambda\,\langle\mathbf{y},\mathbf{q}\rangle < 0.$$

Hence, we can conclude that

$$||\mathbf{q}-\lambda\mathbf{y}|| < ||\mathbf{q}|| = 1 \text{ for all } \lambda\in(\lambda_1,\lambda_2).$$

This implies that $d\left(\{\mathbf{q}_1,...,\mathbf{q}_m\},\mathrm{sp}(\mathbf{y})\right) < 1.$

Corollary. The distance $d(\mathbf{y},\mathrm{col}(\mathbf{Q})) = ||\mathbf{y}||$ if and only if the distance

$$d\left(\{\mathbf{q}_1,...,\mathbf{q}_m\},\mathrm{sp}(\mathbf{y})\right) = 1.$$

Theorem 3. *The condition $R_u^2 = 0$ is satisfied if and only if* $d(\mathbf{y},\mathbf{q}_i) = \sqrt{1+||\mathbf{y}||^2}$ for all $i = 1,...,k$.

Proof. The condition $\mathrm{R}_u^2 = 0$ holds if an only if $\left|\left|\mathbf{Q}\mathbf{Q}^T\mathbf{y}\right|\right| = 0$, that is, $\mathbf{y}\perp\mathbf{q}_i$. Thus, by the Pythagorean theorem, the last condition is equivalent to

$$d(\mathbf{y},\mathbf{q}_i) = \sqrt{||\mathbf{q}_i|| + ||\mathbf{y}||^2}.$$

Since $||\mathbf{q}_i|| = 1$, we conclude that $d(\mathbf{y},\mathbf{q}_i) = \sqrt{1+||\mathbf{y}||^2}$.

Theorem 4. If the distance

$$d\left(\{\mathbf{q}_1,...,\mathbf{q}_m\},\mathrm{sp}(\mathbf{y})\right) = 0,$$

then $\mathrm{R}_u^2 = 1$.

Proof. First we note that if θ is an angle between vectors $\mathbf{q}_i$ and $\mathbf{y}$, then

$$\cos(\theta) = \frac{\langle\mathbf{q}_i,\mathbf{y}\rangle}{||\mathbf{y}||}.$$

Solving the right triangle we obtain

$$d\left(\mathbf{q}_i,\mathrm{sp}(\mathbf{y})\right) = \sqrt{||q_i||^2 - ||q_i||^2\cos^2(\theta)} = \sqrt{1-\frac{\langle\mathbf{q}_i,\mathbf{y}\rangle^2}{||\mathbf{y}||^2}} \quad i = 1,...,m,$$

and hence

$$d\left(\{\mathbf{q}_1,...,\mathbf{q}_m\},\mathrm{sp}(\mathbf{y})\right) = \min_{i=1,...,m}\sqrt{1-\frac{\langle\mathbf{q}_i,\mathbf{y}\rangle^2}{||\mathbf{y}||^2}}.$$

Now, the assumption $d\left(\{\mathbf{q}_1,...,\mathbf{q}_m\},\mathrm{sp}(\mathbf{y})\right) = 0$ is equivalent to the existence of a vector $\mathbf{q}_k$ such that

$$\frac{\langle\mathbf{q}_k,\mathbf{y}\rangle^2}{||\mathbf{y}||^2} = 1.$$

But then

$$\langle\mathbf{q}_k,\mathbf{y}\rangle^2 = ||\mathbf{y}||^2 = ||\mathbf{P}_\mathbf{X}\mathbf{y}+\mathbf{M}_\mathbf{X}\mathbf{y}||^2 = \left|\left|\mathbf{Q}\mathbf{Q}^T\mathbf{y}+\mathbf{M}_\mathbf{X}\mathbf{y}\right|\right|^2 = \left|\left|\mathbf{Q}\mathbf{Q}^T\mathbf{y}\right|\right|^2 + ||\mathbf{M}_\mathbf{X}\mathbf{y}||^2$$

Since

$$\mathbf{Q}\mathbf{Q}^T\mathbf{y} = \sum_{i=1}^m \langle\mathbf{q}_i,\mathbf{y}\rangle\,\mathbf{q}_i,$$

we have that

$$\left|\left|\mathbf{Q}\mathbf{Q}^T\mathbf{y}\right|\right|^2 = \sum_{i=1}^m \langle\mathbf{q}_i,\mathbf{y}\rangle^2.$$

Thus

$$\langle \mathbf{q}_k, \mathbf{y}\rangle^2 = \sum_{i=1}^{m} \langle \mathbf{q}_i, \mathbf{y}\rangle^2 + ||\mathbf{M}_\mathbf{X}\mathbf{y}||^2 = \langle \mathbf{q}_k, \mathbf{y}\rangle^2 + \sum_{\substack{i=1\\ i\neq k}}^{m} \langle \mathbf{q}_i, \mathbf{y}\rangle^2 + ||\mathbf{M}_\mathbf{X}\mathbf{y}||^2 .$$

Hence,

$$||\mathbf{M}_\mathbf{X}\mathbf{y}||^2 = 0,$$

that is,

$$\langle \mathbf{q}_k, y\rangle^2 = ||\mathbf{y}||^2 = \left|\left|\mathbf{Q}\mathbf{Q}^T\mathbf{y}\right|\right|^2 .$$

Since

$$\mathrm{R}_u^2 = \frac{\left|\left|\mathbf{Q}\mathbf{Q}^T\mathbf{y}\right|\right|^2}{||\mathbf{y}||^2},$$

the last equality is equivalent to $\mathrm{R}_u^2 = 1$.

If we add one additional assumption to the assumptions of Theorem 4, we obtain the following criterion.

Theorem 5. *If there exists* $k, 1 \leq k \leq m$, *such that* $\sum_{\substack{i=1\\ i\neq k}}^{m} \langle \mathbf{q}_i, \mathbf{y}\rangle^2 = 0$, *then the condition* $R_u^2 = 1$ *holds if and only if the distance*

$$d\left(\{\mathbf{q}_1, ..., \mathbf{q}_m\}, \mathrm{sp}(\mathbf{y})\right) = 0.$$

Proof. In Theorem 4 we have shown that $d\left(\{\mathbf{q}_1, ..., \mathbf{q}_m\}, \mathrm{sp}(\mathbf{y})\right) = 0$ implies $\mathrm{R}_u^2 = 1$ and there exists $k, 1 \leq k \leq m$, such that $\sum_{\substack{i=1\\ i\neq k}}^{m} \langle \mathbf{q}_i, \mathbf{y}\rangle^2 = 0$.

Vice versa, if $\mathrm{R}_u^2 = 1$ and hence $||\mathbf{M}_\mathbf{X}\mathbf{y}||^2 = 0$ and there exists $k, 1 \leq k \leq m$, such that $\sum_{\substack{i=1\\ i\neq k}}^{m} \langle \mathbf{q}_i, \mathbf{y}\rangle^2 = 0$, then

$$d\left(\{\mathbf{q}_1, ..., \mathbf{q}_m\}, \mathrm{sp}(\mathbf{y})\right) = \min_{i=1,\ldots,m} \sqrt{1 - \frac{\langle \mathbf{q}_i, \mathbf{y}\rangle^2}{||\mathbf{y}||^2}} = \min_{i=1,\ldots,m} \sqrt{1 - \frac{\langle \mathbf{q}_i, \mathbf{Q}\mathbf{Q}^T\mathbf{y}\rangle^2}{||\mathbf{Q}\mathbf{Q}^T\mathbf{y}||^2}}$$

$$= \min_{i=1,\ldots,m} \sqrt{1 - \frac{\left\langle \mathbf{q}_i, \sum_{j=1}^{m} \langle \mathbf{q}_j, \mathbf{y}\rangle \mathbf{q}_j \right\rangle^2}{\sum_{j=1}^{m} \langle \mathbf{q}_j, \mathbf{y}\rangle^2}} = \min_{i=1,\ldots,m} \sqrt{1 - \frac{\left\langle \mathbf{q}_i, \sum_{j=1}^{m} \langle \mathbf{q}_j, \mathbf{y}\rangle \mathbf{q}_j \right\rangle^2}{\langle \mathbf{q}_k, \mathbf{y}\rangle^2 + \sum_{\substack{j=1\\ j\neq k}}^{m} \langle \mathbf{q}_j, \mathbf{y}\rangle^2}}$$

$$= \min_{i=1,\ldots,m} \sqrt{1 - \frac{\langle \mathbf{q}_i, \mathbf{y}\rangle^2}{\langle \mathbf{q}_k, \mathbf{y}\rangle^2}} = 0$$

Theorem 6. *If* $d\left(\{\mathbf{q}_1, ..., \mathbf{q}_m\}, sp(\mathbf{y})\right) < 1$, *then*

$$\mathrm{R}_u^2 = 1 - d^2\left(\{\mathbf{q}_1, ..., \mathbf{q}_m\}, \mathrm{sp}(\mathbf{y})\right)$$

if and only if there exists $k, 1 \leq k \leq m$, such that $\sum_{\substack{i=1\\ i\neq k}}^{m} \langle \mathbf{q}_i, \mathbf{y}\rangle^2 = 0$ and $\langle \mathbf{q}_k, \mathbf{y}\rangle \neq 0$.

Proof. ($\Rightarrow$) If

$$\mathrm{R}_u^2 = 1 - d^2\left(\{\mathbf{q}_1, ..., \mathbf{q}_m\}, \mathrm{sp}(\mathbf{y})\right),$$

then

$$1 = \mathrm{R}_u^2 + d^2\left(\{\mathbf{q}_1, ..., \mathbf{q}_m\}, \mathrm{sp}(\mathbf{y})\right) = \frac{||\mathbf{Q}\mathbf{Q}^T\mathbf{y}||^2}{||\mathbf{y}||^2} + \left(\min_{i=1,\ldots,m} \sqrt{1 - \frac{\langle \mathbf{q}_i, \mathbf{y}\rangle^2}{||\mathbf{y}||^2}}\right)^2$$

$$= \frac{\left|\left|\mathbf{Q}\mathbf{Q}^T\mathbf{y}\right|\right|^2}{||\mathbf{y}||^2} + 1 - \frac{\max_{i=1,\ldots,m} \langle \mathbf{q}_i, \mathbf{y}\rangle^2}{||\mathbf{y}||^2}.$$

It follows that

$$\left|\left|\mathbf{Q}\mathbf{Q}^T\mathbf{y}\right|\right|^2 = \max_{i=1,\ldots,m} \langle \mathbf{q}_i, \mathbf{y}\rangle^2 .$$

On the other hand

$$\left|\left|\mathbf{Q}\mathbf{Q}^T\mathbf{y}\right|\right|^2 = \sum_{i=1}^{m} \langle \mathbf{q}_i, \mathbf{y}\rangle^2 .$$

Thus

$$\sum_{i=1}^{m} \langle \mathbf{q}_i, \mathbf{y}\rangle^2 = \max_{i=1,\ldots,m} \langle \mathbf{q}_i, \mathbf{y}\rangle^2 .$$

Let

$$\langle \mathbf{q}_k, \mathbf{y}\rangle^2 = \max_{i=1,\ldots,m} \langle \mathbf{q}_i, \mathbf{y}\rangle^2 .$$

We have that

$$\sum_{i=1}^{m} \langle \mathbf{q}_i, \mathbf{y}\rangle^2 = \langle \mathbf{q}_k, \mathbf{y}\rangle^2 .$$

This implies that

$$\sum_{\substack{i=1\\ i\neq k}}^{m} \langle \mathbf{q}_i, \mathbf{y}\rangle^2 = 0.$$

Further, since by hypothesis

$$d\left(\{\mathbf{q}_1, ..., \mathbf{q}_m\}, \mathrm{sp}(\mathbf{y})\right) < 1,$$

we have that $\langle \mathbf{q}_k, \mathbf{y}\rangle \neq 0$.

($\Leftarrow$) If there exists $k, 1 \leq k \leq m$, such that $\sum_{\substack{i=1\\ i\neq k}}^{m} \langle \mathbf{q}_i, \mathbf{y}\rangle^2 = 0$ and $\langle \mathbf{q}_k, \mathbf{y}\rangle \neq 0$, then

$$\sum_{i=1}^{m} \langle \mathbf{q}_i, \mathbf{y}\rangle^2 = \langle \mathbf{q}_k, \mathbf{y}\rangle^2 = \max_{i=1,\ldots,m} \langle \mathbf{q}_i, \mathbf{y}\rangle^2 .$$

Since

$$\left|\left|\mathbf{Q}\mathbf{Q}^T\mathbf{y}\right|\right|^2 = \sum_{i=1}^{m} \langle \mathbf{q}_i, \mathbf{y}\rangle^2$$

we have

$$1 - \frac{\left|\left|\mathbf{Q}\mathbf{Q}^T\mathbf{y}\right|\right|^2}{||y||} = 1 - \frac{\max_{i=1,\ldots,m} \langle \mathbf{q}_i, \mathbf{y}\rangle^2}{||y||}.$$

The last equality is equivalent to

$$\mathrm{R}_u^2 = 1 - d^2\left(\{\mathbf{q}_1, ..., \mathbf{q}_m\}, \mathrm{sp}(\mathbf{y})\right).$$

Since $\sum_{\substack{i=1\\ i\neq k}}^{m} \langle \mathbf{q}_i, \mathbf{y}\rangle^2 = 0$ and $\langle \mathbf{q}_k, \mathbf{y}\rangle \neq 0$ holds if and only if $\langle \mathbf{q}_i, \mathbf{y}\rangle = 0$ for $i = 1, ..., m$, $i \neq k$ and $\langle \mathbf{q}_k, \mathbf{y}\rangle \neq 0$, Theorem 6 deals with the case when vector $\mathbf{y}$ is perpendicular to all vectors $\mathbf{q}_i$ except one vector $\mathbf{q}_k$, and vector $\mathbf{y}$ is *not* perpendicular to $\mathbf{q}_k$.

3. CONCLUSIONS

In this paper we have investigated the geometrical properties of the uncentered coefficient of determination in linear regression analysis. Theorems 1, 2, and 3 provide three geometrical characterizations of the condition $\mathrm{R}_u^2 = 0$. The first characterization concerns the distance between the vector $\mathbf{y}$ and the subspace $\mathrm{col}(\mathbf{Q})$. The second characterization concerns the distance between the subspace $\mathrm{sp}(\mathbf{y})$ and the set $\{\mathbf{q}_1, ..., \mathbf{q}_m\}$. It is important to note that $d(\mathbf{y}, \mathrm{col}(\mathbf{Q})) \leq ||\mathbf{y}||$ and $d(\{\mathbf{q}_1, ..., \mathbf{q}_m\}, \mathrm{sp}(\mathbf{y})) \leq 1$. Thus we can conclude that $\mathrm{R}_u^2 = 0$ if and only if these distances reach the largest possible values.

As it is expected as the negation of the case $\mathrm{R}_u^2 = 0$, the uncentered coefficient of determination reaches the highest value $\mathrm{R}_u^2 = 1$ in the case then the distance $d(\{\mathbf{q}_1, ..., \mathbf{q}_m\}, \mathrm{sp}(\mathbf{y}))$ is zero. Here the additional assumption of orthogonality of the response vector $\mathbf{y}$ to all column vectors of the matrix $\mathbf{Q}$ except one, plays an essential role.

REFERENCES

1. Å. Björck, *Numerical Methods for Least Squares Problems, xviii+* (SIAM, Philadelphiak, 1996).