some title

January 7, 2018

Regression

Sample correlation

The most crucial part in defining correlation geometrically is definig the dot product as it enables to compute the length of a vecotr:

$$|\vec{a}| = \sqrt{\langle \vec{a}, \vec{a} \rangle}$$

and the angle between any two vectors:

$$\cos(\vec{a}, \vec{b}) = \frac{\langle \vec{a}, \vec{a} \rangle}{|\vec{a}||\vec{b}|}$$

Now we define scalar product of two vectors $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ and $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ as a sample covariation between them:

$$\langle x, y \rangle = \mathrm{sCov}(x, y)$$

The main characteristics of a vector are its length and direction. So, we introduce the length

$$\sqrt{\mathrm{sCov}(x, x)} = \sqrt{\mathrm{sVar}(x)} = \sigma_x$$

and the angle between two random vectors

$$\cos(x, y) = \frac{\mathrm{sCov}(x, y)}{\sqrt{\mathrm{sVar}(x)\,\mathrm{sVar}(y)}} = \mathrm{sCorr}(x, y)$$

Note that from the definition of the angle it follows that the sample correlation can range from -1 to 1.

$$\begin{aligned} \mathrm{sCorr}(x, y) &= \frac{\mathrm{sCov}(x, y)}{\sqrt{\mathrm{sVar}(x)\,\mathrm{sVar}(y)}} \\ &= \frac{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})\frac{1}{n-1}\sum_{i=1}^{n}(\hat{y}_i - \bar{\hat{y}})}} \end{aligned}$$

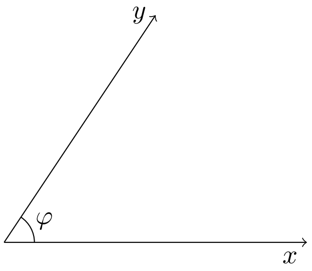


Figure 1: Vector x of length σ_x and random vector y of length σ_y, $\cos\varphi$ is the angle between x and y.

Another important geometrical tool is projection. Recall that for any two vectors the scalar product $\langle \vec{a}, \vec{b} \rangle$ can be interpreted as the length of projected $\vec{b}$ multuplied by the length of $\vec{a}$. The projection itself is $cos(\vec{a}, \vec{b})\vec{b}$. Same holds for the vectors sampled from some distribution. The projection of such a vector y onto $\{cx | c \in \mathbb{R}\}$ is $\hat{y} = \text{sCorr}(x, y) \cdot y$.

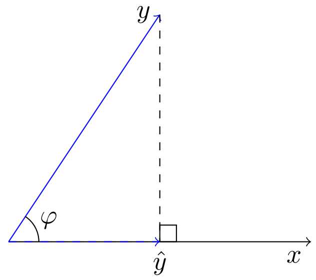


Figure 2: Vector y projected onto vector x.

Looking at Figure , we can interpret the square of sample correlation coefficient. Using the fact that $cos^2\varphi$ is the squared ratio of the leg adjacent to φ to hypotenuse, we can conclude that

$$\text{sCorr}^2(x, y) = \frac{\text{sVar}(\hat{y})}{\text{sVar}(y)}$$

as the variance of a vector is associated with the square of its length. Thus, the sample correlation coefficient squared shows the fraction of variance in y which can be explained with the most similar vector proportional to x.

Sample correlation when a constatnt vector added

Theorem 1. *Adding a vector of constants does not affect the sample correlation coefficient:*

$$\text{sCorr}(x + \alpha\mathbf{1}, y) = \text{sCorr}(x, y)$$

where $\alpha \in \mathbb{R}$.

$$\begin{aligned}\text{sCorr}(x + \alpha\mathbf{1}, y) &= \frac{\text{sCov}(x + \alpha\mathbf{1}, y)}{\text{sVar}(x + \alpha\mathbf{1})\,\text{sVar}(y)} \\ &= \frac{\text{sCov}(x, y)\,\text{sCov}(\alpha\mathbf{1}, y)}{\text{sVar}(x)\,\text{sVar}(y)} \\ &= \frac{\text{sCov}(x, y)}{\text{sVar}(x)\,\text{sVar}(y)} \\ &= \text{sCorr}(x, y)\end{aligned}$$

Proof. Firstly, we project vectors x and y onto $Lin^{\perp}(\mathbf{1})$ in order to get $x^c = x - \bar{x}$ and $y^c = y - \bar{y}$ ('c' stands for 'centred'). It can be shown that the matrix corresponding to projecting onto the line spanned by a vector of all ones has the following form

$$\frac{\mathbf{1}^T\mathbf{1}}{\mathbf{1}\mathbf{1}^T} = \frac{\begin{pmatrix}1 & \dots & 1\end{pmatrix}\begin{pmatrix}1 \\ \dots \\ 1\end{pmatrix}}{\sum_{i=1}^{n} 1} = \begin{pmatrix}\frac{1}{n} & \cdots & \frac{1}{n} \\ \vdots & \ddots & \vdots \\ \frac{1}{n} & \cdots & \frac{1}{n}\end{pmatrix}$$

Thus, projecting onto an orthogonal subspace is equivalent to substracting the projected vector, i.e., the vector of averages, from the original one.

Also note that the angle φ between original and centred vectors remains the same. The results of this step is shown in Figure .

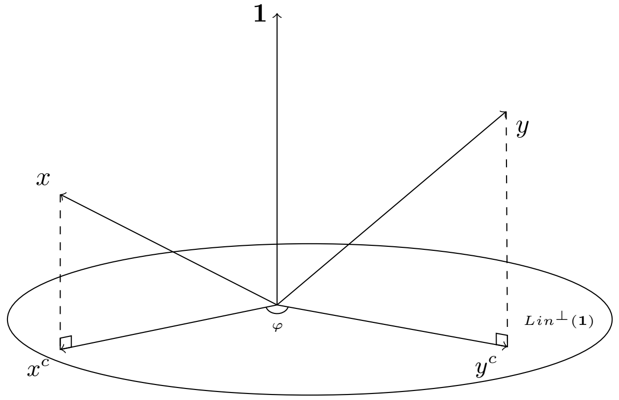


Figure 3: Centred vectors x^c and y^c

Then we need to derive a new vector $\tilde{x}$ with constants added to each component. Geometrically adding a vector of costants means adding a vector of all ones scaled by $\alpha \in \mathbb{R}$, i.e., $\alpha\mathbf{1}$. Then the new vector $\tilde{x}$ can be broken up into a sum of $\alpha\mathbf{1}$ and βx, $\alpha, \beta \in \mathbb{R}$, which can be seen in Figure . After that we will project this new vector $\tilde{x}$ onto $Lin^{\perp}(\mathbf{1})$. By the properties of projection it is of no difference whether to project the whole vector $\tilde{x}$ or project its parts $\alpha\mathbf{1}$ and βx — the result is the same. So, while βx is projected onto a span of x^c, the projection of $\alpha\mathbf{1}$ onto the orthohgonal space $Lin^{\perp}(\mathbf{1})$ yields zero as demonstrated in Figure . Moreover, it follows that the angle between $\tilde{x}$ and y is still φ.

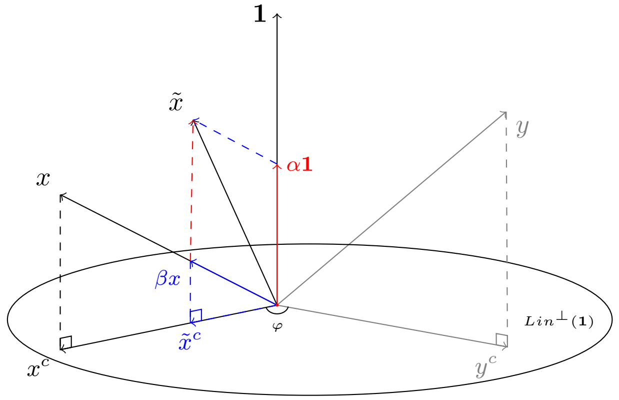


Figure 4: Decomposition and projection of $\tilde{x}$

Finally, putting everything together we finish the proof:

$$\text{sCorr}(x + \alpha\mathbf{1}, y) = \text{sCorr}(x, y)$$

□

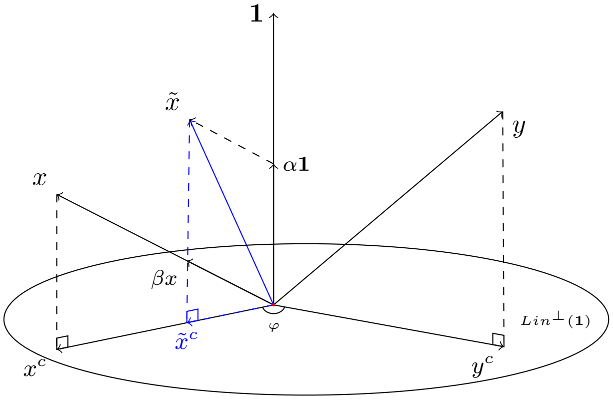


Figure 5: $\text{sCorr}(x + \alpha\mathbf{1}, y) = \text{sCorr}(x, y)$ as the corresponding angles are equal.

Sample correlation coefficient in simple linear regression

Theorem 2. *A linear regression model with one explanatory variable and constant term*

$$y = \beta_1 + \beta_2 x + \varepsilon$$

has the property

$$\text{sCorr}(y, \hat{y}) = sign(\hat{\beta}_2)\,\text{sCorr}(y, x)$$

Proof. Firstly, we consider the case when $\hat{\beta}_2 > 0$ so the main picuture is of the form depicted in Figure . It has been shown earlier that the correlation coefficient represents the angle betweem two random vectors. So in order to complete the proof we need to find the appropriate angles and compare them.

Assuming the underlying relationship between x and y to be

$$y_i = \beta_1 + \beta_2 x_i + \varepsilon_i \quad i = 1, \ldots, n$$

where ε_i is an error term the following holds

$$\begin{aligned}
\text{sCorr}(y, \hat{y}) &= \frac{\text{sCov}(y)\,\text{sCov}(\hat{y})}{\sqrt{\text{sVar}(y)\,\text{sVar}(\hat{y})}} \\
&= \frac{\text{sCov}(y)\,\text{sCov}(\hat{\beta}_1 + \hat{\beta}_2 x)}{\sqrt{\text{sVar}(y)\,\text{sVar}(\hat{\beta}_1 + \hat{\beta}_2 x)}} \\
&= \frac{\text{sCov}(y)\,\text{sCov}(\hat{\beta}_2 x)}{\sqrt{\text{sVar}(y)\,\text{sVar}(\hat{\beta}_2 x)}} \\
&= \frac{\hat{\beta}_2\,\text{sCov}(y)\,\text{sCov}(x)}{|\hat{\beta}_2|\sqrt{\text{sVar}(y)\,\text{sVar}(x)}} \\
&= sign(\hat{\beta}_2)\frac{\text{sCov}(y)\,\text{sCov}(x)}{\sqrt{\text{sVar}(y)\,\text{sVar}(x)}}
\end{aligned}$$

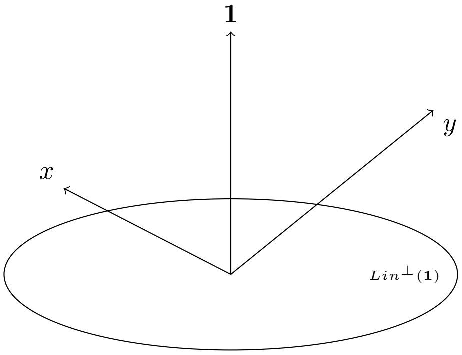
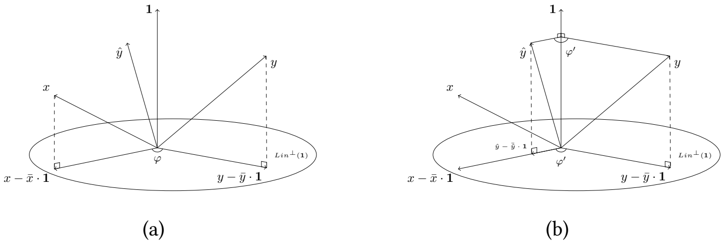


Figure 6: Vectors x, y and $\mathbf{1}$

However, it seems to be difficult to compare the angles in the three dimensional space. That is why we start with projecting both x and y onto the space perpendicular to the vector of all ones $\mathbf{1}$ as shown in Figure 7(a). We denote this space as $Lin^{\perp}(\mathbf{1})$. The resulting vectors are $x - \bar{x} \cdot \mathbf{1}$ and $y - \bar{y} \cdot \mathbf{1}$ respectively since projection of any vector $\vec{a}$ on the line given by a vector of all ones yields the vector of averages $\vec{\bar{a}}$.

In order to get the angle between y and $\hat{y}$ we should start with regressing y on $Lin(x, \mathbf{1})$. Then the only thing thing left is to project $\hat{y}$ onto $Lin^{\perp}\mathbf{1}$ since the y vector has already been projected. Note that the projected $\hat{y}$ falls onto tha span of vector $x - \bar{x} \cdot \mathbf{1}$ as it can be decomposed into a sum $ax + b\mathbf{1}$ where $a, b \in \mathbb{R}$. ax is projected in the same way as x and $b\mathbf{1}$ yields zero when projected onto the orthogonal space. The result of this step is shown in Figure 7(b).


$\mathbf{1}$
$\hat{y}$
y
x
φ
$Lin^{\perp}(\mathbf{1})$
$x - \bar{x} \cdot \mathbf{1}$
$y - \bar{y} \cdot \mathbf{1}$
$\mathbf{1}$
$\hat{y}$
φ'
y
x
$\hat{y} - \bar{\hat{y}} \cdot \mathbf{1}$
φ'
$Lin^{\perp}(\mathbf{1})$
$x - \bar{x} \cdot \mathbf{1}$
$y - \bar{y} \cdot \mathbf{1}$

(a)

(b)

Figure 7: (a): 'Centred' x and y, i.e., projected onto $Lin^{\perp}(\mathbf{1})$; (b): 'Centred' $\hat{y}$, i.e., projected onto $Lin^{\perp}(\mathbf{1})$.

Since the projection of $\hat{y}$ lies exactly on the span of vector $x - \bar{x} \cdot \mathbf{1}$, we can conclude that $cos\varphi = \cos\varphi'$ and to put it another way sCorr$(x, y) =$ sCorr$(y, \hat{y})$.

Now consider the case when $\hat{\beta}_2 < 0$. Note that the sign of β_1 does not influence the correlation coefficient sign. The only difference is that now $\hat{y}$ is projected onto the span of $x - \bar{x} \cdot \mathbf{1}$ and not on this vector itself while the projections of x and y remain the same. Looking at Figure we deduce that the angle betwween y and $\hat{y}$ is compelement to the angle between x and y. Using trigonometric properties, we simplify $\cos(180^\circ - \varphi) = -\cos\varphi$ which in turn implies sCorr$(x, y) = -$ sCorr$(y, \hat{y})$.

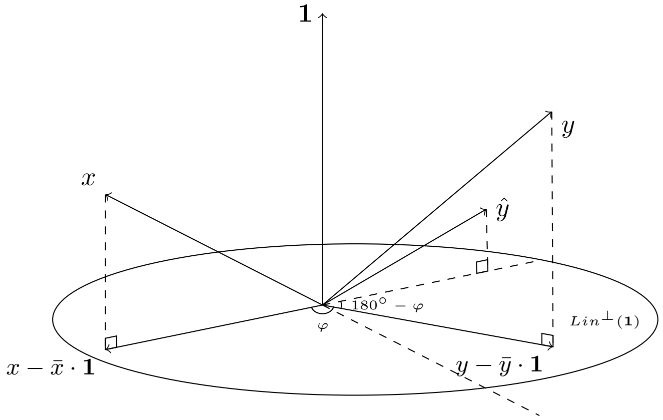

$\mathbf{1}$
y
x
$\hat{y}$
$180^\circ - \varphi$
φ
$Lin^{\perp}(\mathbf{1})$
$x - \bar{x} \cdot \mathbf{1}$
$y - \bar{y} \cdot \mathbf{1}$

Figure 8: Case of $\beta_2 < 0$.

□

RSS + ESS = TSS

Theorem 3. *A linear regression model with n observations and k explanatory variables including a constant unit vector*

$$y = X\beta + \varepsilon$$

has the following property

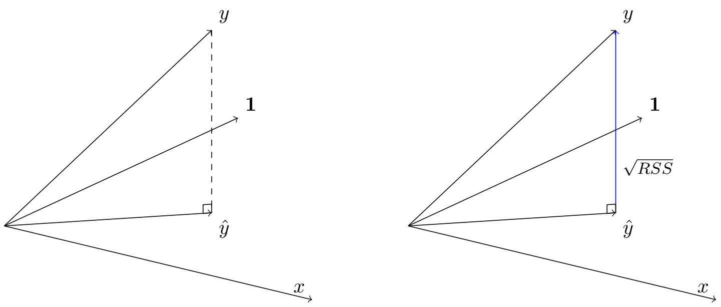
$$RSS + ESS = TSS$$

where $RSS = \|y - \hat{y}\|_2^2$, $ESS = \|\hat{y} - \bar{y}\|_2^2$, $TSS = \|y - \bar{y}\|_2^2$.

Proof. The proof will be presented for the case of two regressor x and $\mathbf{1}$ in order for the picture to be clear. However, the same logic applies for the case of k regressors.

We start with depicting the vectors $y \in \mathbb{R}^{n-2}$ and $x, \mathbf{1} \in \mathbb{R}^2$. Then we project y onto $Lin(x, \mathbf{1})$ and obtain $\hat{y}$ which is shown in Figure 9(a).

From this picture we can immediately derive $\sqrt{RSS}$ as by definition this is the squared difference between y and $\hat{y}$.



(a)

(b)

Figure 9: (a): Vectors $y \in \mathbb{R}^{n-2}$ and $\hat{y} \in Lin(x, \mathbf{1})$; (b): Residual sum of squares.

So as to visualize ESS and TSS we first need to visualize vector of averages $\bar{y}$. Geometrically this means projecting a vector onto a line spanned by vector $\mathbf{1}$.

Now we both project y and $\hat{y}$ onto $\mathbf{1}$ and following the definition obtain $\sqrt{TSS}$ as the difference vector $y - \bar{y}$ and $\sqrt{ESS}$ as the vector $\hat{y} - \bar{y}$.

The final step is to put everything together. Note that since $y - \hat{y}$ is perpendicular to $Lin(x, \mathbf{1})$ it is also perpendicular to $\hat{y} - \bar{y}$ and $\mathbf{1}$ as these vectoros are in $Lin(x, \mathbf{1})$. Then, applying the theorem of three perpendiculars we conclude that the foot of vector $y - \bar{y}$ is the same point as the foot of the vector $\hat{y} - \bar{y}$. Thus, we obtain a right angle triangle and can apply the Pythagorean theorem for the catheti $\sqrt{RSS}$ and $\sqrt{ESS}$ and the hypotenuse $\sqrt{TSS}$:

$$(\sqrt{RSS})^2 + (\sqrt{ESS})^2 = (\sqrt{TSS})^2$$

□

Consider a regresion model with n observations and k explanatory variables including a constant unit vector

$$y = X\beta + \varepsilon$$

The OLS estimator for the vector of coefficients β is

$$\hat{\beta} = (X^T X)^{-1} X^T y$$

and the residual vector is

$$\begin{aligned} \hat{e} &= y - \hat{y} \\ &= y - X\hat{\beta} \\ &= y - X(X^T X)^{-1} X^T y \end{aligned}$$

Then we define residual sum of squares (RSS), explained sum of squares (ESS) and total sum of squares (TSS) as follows:

$$\begin{aligned} RSS &= \|y - \hat{y}\|_2^2 \\ ESS &= \|\hat{y} - \bar{y}\|_2^2 \\ TSS &= \|y - \bar{y}\| \end{aligned}$$

Disclosing parentheses and using the fact that $\hat{y}^T y = \hat{y}^T \hat{y}$

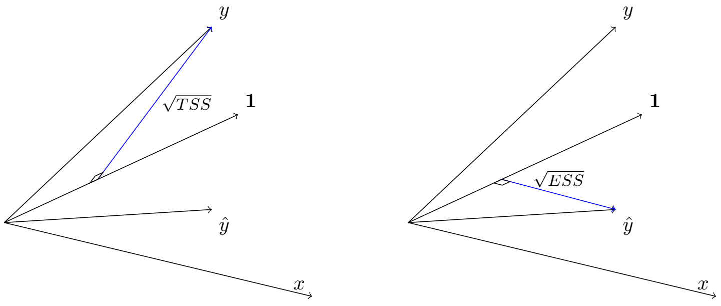
$$\begin{aligned} \hat{y}^T y &= \beta^T X^T y \\ &= y^T X(X^T X)^{-1} X^T y \\ \hat{y}^T \hat{y} &= \beta^T X^T X\beta \\ &= y^T X(X^T X)^{-1} X^T X(X^T X)^{-1} X^T y \\ &= y^T X(X^T X)^{-1} X^T y \end{aligned}$$

we obtain

$$\begin{aligned} RSS &= y^T y - \hat{y}^T \hat{y} \\ ESS &= \hat{y}^T \hat{y} - \hat{y}^T \bar{y} + \bar{y}^T \bar{y} \\ TSS &= y^T y - 2y^T \bar{y} + \bar{y}^T \bar{y} \end{aligned}$$

When putting everything together all the terms cancel out which proves

$$ESS + RSS = TSS$$



(a)

(b)

Figure 10: (a): Total sum of squares; (b): Explained sum of squares.

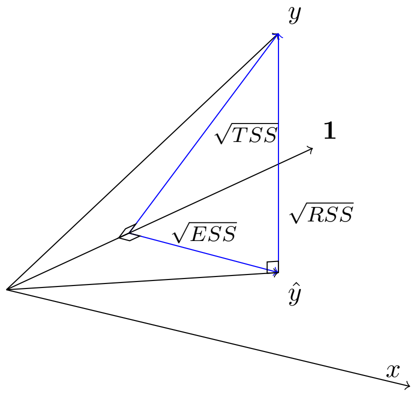


Figure 11: $(\sqrt{RSS})^2 + (\sqrt{ESS})^2 = (\sqrt{TSS})^2$

Determination coefficient

Theorem 4. *A linear regression model with n observations and k explanatory variables including a constant unit vector*

$$y = X\beta + \varepsilon$$

has the following property

$$R^2 = \mathrm{sCorr}^2(y, \hat{y})$$

Proof. Proving this theorem geometrically means showing that the determination coefficient can be interpreted as some squared angle which happens to be eqaul to the squared angle betwen y and $\hat{y}$.

Consider Figure from the previous proof. It was shown there that the vectors $y - \bar{y}$, $y - \hat{y}$ and $\hat{y} - \bar{y}$ form a right triangle. Having defined the determination coefficient as

$$R^2 = \frac{ESS}{TSS}$$

we conclude that its geometric interpretaion is

$$\sqrt{R^2} = \frac{\sqrt{ESS}}{\sqrt{TSS}} = \cos\varphi$$

$$\begin{aligned}
\mathrm{sCorr}^2(y, \hat{y}) &= \left(\frac{\mathrm{sCov}(y, \hat{y})}{\sqrt{\mathrm{sVar}(y)\,\mathrm{sVar}(\hat{y})}}\right)^2 \\
&= \frac{\mathrm{sCov}(y, \hat{y})\,\mathrm{sCov}(y, \hat{y})}{\mathrm{sVar}(y)\,\mathrm{sVar}(\hat{y})} \\
&= \frac{\mathrm{sCov}(\hat{y} + e, \hat{y})\,\mathrm{sCov}(\hat{y} + e, \hat{y})}{\mathrm{sVar}(y)\,\mathrm{sVar}(\hat{y})} \\
&= \frac{(\mathrm{sCov}(\hat{y}, \hat{y}) + \mathrm{sCov}(e, \hat{y}))(\mathrm{sCov}(\hat{y}, \hat{y}) + \mathrm{sCov}}{\mathrm{sVar}(y)\,\mathrm{sVar}(\hat{y})} \\
&= \frac{\mathrm{sVar}(\hat{y})\,\mathrm{sVar}(\hat{y})}{\mathrm{sVar}(y)\,\mathrm{sVar}(\hat{y})} \\
&= \frac{\mathrm{sVar}(\hat{y})}{\mathrm{sVar}(y)} \\
&= \frac{ESS}{TSS} \\
&= R^2
\end{aligned}$$

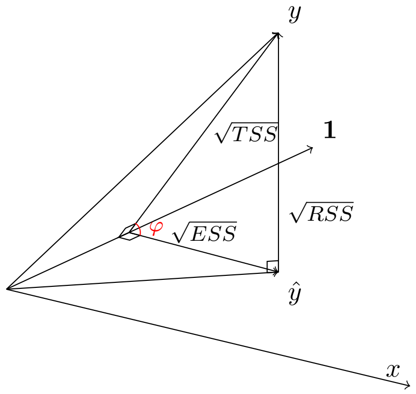


Figure 12: Determination coefficient as squared cos φ

as shown in Figure .

Recall that the sample correlation coefficient two vectors was defined earlier as the angle between these two vectors. Thus, we conclude that sCorr$(y, \hat{y})$ is the angle between y and $\hat{y}$ which is also eqaul to cos φ. Finally, squaring both sides, we obtain

$$R^2 = \text{sCorr}^2(y, \hat{y})$$

□

Regression line and point of averages

Theorem 5. *The point of averages lies on the estimated regression line.*

Proof. For the geometrical proof it suffices to show that $\hat{y}$ is a linear combination of the regressors, which is true by construction, and that $\frac{1}{n}\sum_{i=1}^{n}\hat{y}_i = \frac{1}{n}\sum_{i=1}^{n} y$. In order for the pictures to be more clear the proof will be presented for the case of two regressors.

The first step is regressing y on $Lin(\mathbf{1}, x)$. As shown in Figure 13(a), we obtain $\hat{y}$ as a linear combination of $\mathbf{1}$ and x. The next step is to regress both y and $\hat{y}$ on $\mathbf{1}$ which results in $\bar{y}$ and $\bar{\hat{y}}$ correspondingly. By the theorem of three perpendiculars, $\bar{y} = \bar{\hat{y}}$ which is shown in Figure 13(b).

□

Frisch–Waugh–Lovell theorem

Theorem 6. *Consider regression*

$$y = X_1\beta_1 + X_2\beta_2 + u \tag{1}$$

where $X_{n\times k} = [X_1 X_2]$, i.e. X_1 consists of first k_1 columns of X and X_2 consists of remaining k_2 columns of X, β_1 and β_2 are comfortable, i.e. $k_1 \times 1$ and $k_2 \times 1$ vectors. Consider another regresion

$$M_1 y = M_1 X_2 \beta_2 + M_1 u \tag{2}$$

If the regression contains the intercept, the following equation holds:

$$\begin{aligned}\hat{y} &= X\hat{\beta}\\ &= X(X^TX)^{-1}X^Ty\\ &= X(X^TX)^{-1}X^TX\beta + X(X^TX)^{-1}X^T\varepsilon\end{aligned}$$

Premultiplying both sides by X^T, we obtain:

$$\begin{aligned}X^T\hat{y} &= X^TX(X^TX)^{-1}X^TX\beta\\ &\quad + X^TX(X^TX)^{-1}X^T\varepsilon\\ &= X^TX\beta + X^T\varepsilon\end{aligned}$$

This is a system of equations. The first row of X^T is $\mathbf{1}$ vector, so we can write out the first equation:

$$\sum_{i=1}^{n}\hat{y}_i = \sum_{i=1}^{n}\sum_{j=1}^{k} x_{ij}\beta_j$$

From the first equation in the system

$$X^T\hat{y} = X^Ty$$

we obtain

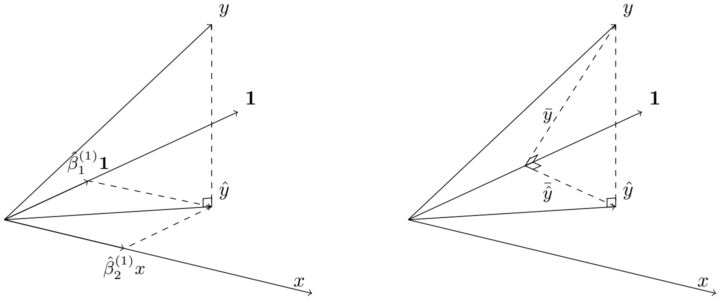
$$\sum_{i=1}^{n}\hat{y}_i = \sum_{i=1}^{n} y$$

And this finishes the proof:

$$\frac{1}{n}\sum_{i=1}^{n} y = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{k} x_{ij}\beta_j$$

From regresison 2 we get the following estimator:

$$\begin{aligned}\hat{\beta}_2 &= ((M_1X_2)^TM_1X_2)^{-1}(M_1X_2)^TM_1y\\ &= (X_2^TM_1^TM_1X_2)^{-1}X_2^TM_1^TM_1y\\ &= (X_2^TM_1X_2)^{-1}X_2^TM_1y\end{aligned}$$



(a)

(b)

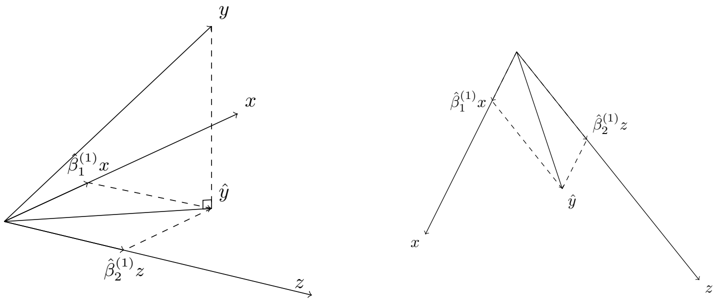
Figure 13: (a): Regression of y on $Lin(\mathbb{1}, x)$; (b): Regression of y and $\hat{y}$ on $\mathbb{1}$.

where $M_1 = I - P_1$ projects onto the orthogonal complement of the column space of X_1 and $P_1 = X_1(X_1^T X_1)^{-1}X_1^T$ is the projection onto the column space of X_1. Then the estimate of β_2 from regression 1 will be the same as the estimate from regression 2.

Proof. Geometrical proof will be presented for the following model:

$$y_i = \beta_1 x_i + \beta_2 z_i + u_i \tag{3}$$

We start with regression 'all-at-once' and will distinct its coefficients with index (1). The only step in obtaining $\beta_1^{(1)}$ is regressing y on $Lin(x, z)$ and then expanding $\hat{y}$ as a linear combination of basis vectors x and z, which is shown in Figure 14(a). Figure 14(b) depicts $Lin(x, z)$.



(a)

(b)

Figure 14: (a): Regression of y on $Lin(x, z)$; (b): $Lin(x, z)$.

As for the model 2, where several regressions are performed consecutively, we start with regressing y on z, resulting in $\tilde{y}$, which we will refer to as "cleansed" y.

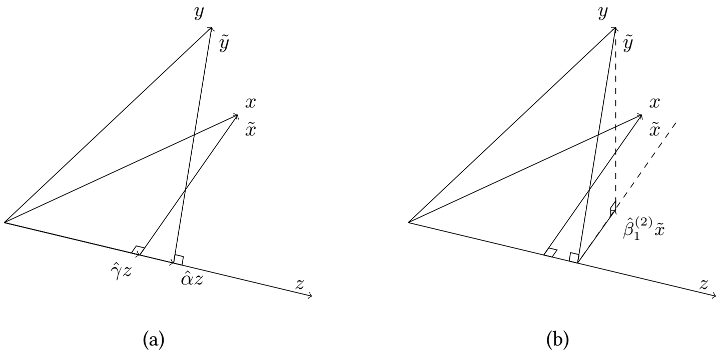
$$\begin{aligned} y &= \alpha z + \varepsilon \\ \hat{\alpha} &= \frac{y^T z}{z^T z} \\ \tilde{y} &= \hat{\varepsilon} = y - \frac{y^T z}{z^T z} z \end{aligned} \tag{4}$$

Following that, x is regressed on z, resulting in $\tilde{x}$ — “cleansed” x.

$$\begin{aligned} x &= \gamma z + \nu \\ \hat{\gamma} &= \frac{x^T z}{z^T z} \\ \tilde{x} &= \hat{\nu} = x - \frac{x^T z}{z^T z} z \end{aligned} \tag{5}$$

Geometric results of these two steps are presented in 15(a).

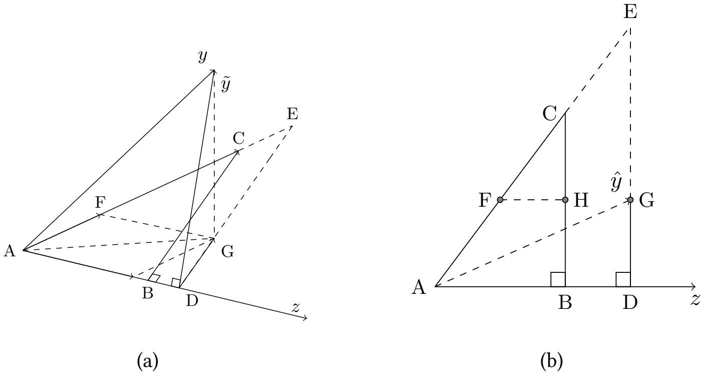
Finally, ‘cleansed’ y must be regressed on ‘cleansed’ x. However, it cannot be performed immediately as $\tilde{y}$ and $\tilde{x}$ are skew lines. So at first, we fix this problem by translation and after taht obtain $\hat{\beta}_1^{(2)}\tilde{x}$ (see Figure 15(b)).



(a) (b)

Figure 15: (a): Regression of y on z and of x on z; (b): Translation of $\tilde{x}$.

Now, let us picture all the results in one figure and mark some main points.



(a) (b)

Figure 16: (a): Point A stands for the origin, B — $\hat{\gamma}z$, C — x, D — $\hat{\alpha}z$, E — intersection of vector x and line parallel to $\tilde{x}$, F — $\hat{\beta}_1^{(1)}x$, G — $\hat{\beta}_1^{(2)}\tilde{x}$; (b): $Lin(x, z)$.

In Figure 16(b) segments AF and BH = DG stand for $\hat{\beta}_1^{(1)}x$ and $\hat{\beta}_1^{(2)}\tilde{x}$ respectively, while segments AC and BC represent x and $\tilde{x}$. Having two congruent angles, triangles ABC and FHC are simillar. Then, it follows:

$$\frac{AF}{AC} = \frac{BH}{BC} \Leftrightarrow \frac{\hat{\beta}_1^{(1)}x}{x} = \frac{\hat{\beta}_1^{(2)}\tilde{x}}{\tilde{x}} \Leftrightarrow \hat{\beta}_1^{(1)} = \hat{\beta}_1^{(2)}$$

□