Qishi Quiz 2

I. Question 7

Consider linear regression of Y on feature X_1, X_2: Model1-(Y, X_2), $R^2 = 0.1$; Model2-(Y, X_2), $R^2 = 0.2$; Model3-(Y, X_1, X_2), calculate the range of R^2 of Model3.

A. Solution

In a univariate regression:

$$R^2 = \rho^2. \tag{1}$$

e.g., in a simple linear regression with one feature: $y = a + bx$, $R^2 = \rho^2_{x,y}$.
For Model1-(Y, X_1), we have

$$R^2 = 0.1 = \rho^2_{y,x_1} \tag{2}$$

For Model2-(Y, X_2), we have

$$R^2 = 0.2 = \rho^2_{y,x_2} \tag{3}$$

The correlation matrix of (Y, X_1, X_2) is positive semi-definite, where

$$Corr(Y, X_1, X_2) = \begin{pmatrix} 1 & \sqrt{0.1} & \sqrt{0.2} \\ \sqrt{0.1} & 1 & \rho_{1,2} \\ \sqrt{0.2} & \rho_{1,2} & 1 \end{pmatrix}. \tag{4}$$

Hence,

$$\rho_{1,2} - 0.2\sqrt{2}\rho_{1,2} - 0.7 \leq 0. \tag{5}$$

Thus, we have

$$-0.5\sqrt{2} \leq \rho_{1,2} \leq 0.7\sqrt{2}. \tag{6}$$

Since Y and X_1 are positively correlated and Y and X_2 are positively correlated, we have X_1 and X_2 are positively correlated, thus,

$$0 \leq \rho_{1,2} \leq 0.7\sqrt{2} \tag{7}$$

In regression $y = a + bx_1 + cx_2$, we have

$$R^2 = \frac{\rho^2_{y,1} + \rho^2_{y,2} - 2\rho_{y,1}\rho_{y,2}\rho_{1,2}}{1 - \rho^2_{1,2}}. \tag{8}$$

where $\rho_{y,1}$ is the correlation of y and x_1, $\rho_{y,2}$ is the correlation of y and x_2, and $\rho_{1,2}$ is the correlation of x_1 and x_2. Since we have $\rho^2_{y,1} = 0.1$ and $\rho^2_{y,2} = 0.2$, we have,

$$R^2 = \frac{0.3 - 0.2\sqrt{2}\rho_{1,2}}{1 - \rho^2_{1,2}}. \tag{9}$$

Let $R^2 = x$, we have

$$\rho_{1,2} = \frac{0.1\sqrt{2} \pm \sqrt{0.02 + x^2 - 0.3x}}{x}. \tag{10}$$

Since $x = R^2 \geq 0.2$, as $R_{y,x_2} = 0.2$, we have,

$$x^2 - 0.3x + 0.02 \geq 0. \tag{11}$$

From (6) and (10), we have

$$0 \leq \frac{0.1\sqrt{2} \pm \sqrt{0.02 + x^2 - 0.3x}}{x} \leq 0.7\sqrt{2}. \tag{12}$$

First, let us take a look at

$$0 \leq \frac{0.1\sqrt{2} \pm \sqrt{0.02 + x^2 - 0.3x}}{x}. \tag{13}$$

Since $x \geq 0$, we have for sure that

$$\frac{0.1\sqrt{2} + \sqrt{0.02 + x^2 - 0.3x}}{x} \geq 0. \tag{14}$$

From

$$0.1\sqrt{2} - \sqrt{0.02 + x^2 - 0.3x} \geq 0, \tag{15}$$

we can get

$$0 \leq x \leq 0.3 \tag{16}$$

Now let us take a look at

$$\frac{0.1\sqrt{2} \pm \sqrt{0.02 + x^2 - 0.3x}}{x} \leq 0.7\sqrt{2} \tag{17}$$

From

$$\frac{0.1\sqrt{2} + \sqrt{0.02 + x^2 - 0.3x}}{x} \leq 0.7\sqrt{2}, \tag{18}$$

we have

$$\sqrt{0.02 + x^2 - 0.3x} \leq 0.7\sqrt{2} - 0.1\sqrt{2}, \tag{19}$$

where the right hand side of (19) ≥ 0 since $x \geq 0.2$. Thus,

$$0 \leq x \leq 1. \tag{20}$$

From

$$\frac{0.1\sqrt{2} - \sqrt{0.02 + x^2 - 0.3x}}{x} \leq 0.7\sqrt{2}, \tag{21}$$

we have

$$-\sqrt{0.02 + x^2 - 0.3x} \leq 0.7\sqrt{2}x - 0.1\sqrt{2}, \tag{22}$$

where (22) is always true since $x \geq 0.2$ and thus $0.7\sqrt{2}x - 0.1\sqrt{2} \geq 0$. Combining (16) and (20), we have

$$0 \leq x \leq 0.3. \tag{23}$$