Level M Regression Models Examples

RSS

Show that these two forms of RSS are equivalent, i.e. show that

$$\begin{aligned} RSS &= \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \\ &= \mathbf{Y}^T\mathbf{Y} - \mathbf{Y}^T\mathbf{X}\hat{\boldsymbol{\beta}} \end{aligned}$$

Alzheimers Disease

Thirty-four sufferers from Alzheimer's disease have been studied. The data collected on each individual consist of a measure of cerebral blood flow (y_i) in the right temporal-occipital brain region, a score measuring performance on a cognitive test (x_{1i}) and age (x_{2i}). It was of interest to explore if flow could be estimated from score and age using the following model:

$$\text{Model 1}: \quad Y_i = \alpha + \beta x_{1i} + \gamma x_{2i} + \epsilon_i$$

The model can be written in vector-matrix form $E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$. The residual sum of squares, RSS for Model 1 above is 0.2099 and $\sum_{i=1}^{n}(y_i - \bar{y})^2 = 0.5993$.

a) Write down the design matrix $\mathbf{X}$ and $\mathbf{X}^T\mathbf{X}$ corresponding to this model.

b) Calculate R^2 and comment on this value.

Oxygen

As part of an investigation into oxygen starvation in babies during labour, the blood oxygen level of the fetal scalp and the mother were taken at the start of the well-defined second stage of labour, and the blood oxygen from the umbilical vein of the baby was taken at delivery. If the umbilical vein blood oxygen (Y) can be reasonably well predicted by the fetal scalp blood oxygen (x_1) and the maternal blood oxygen (x_2) then it may aid the obstetrician in gaining early warning of oxygen distress in the baby. Consider the model:

$$\text{Data: } (y_i, x_{1i}, x_{2i}); \quad i = 1, \ldots, 15$$

$$\text{Model: } Y_i = \alpha + \beta x_{1i} + \gamma x_{2i} + \epsilon_i$$

The model can be written in vector-matrix form $E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$.

$$\sum_{i=1}^{n} y_i = 315.6, \sum_{i=1}^{n} y_i^2 = 6775.83, \sum_{i=1}^{n} x_{1i}y_i = 13256.996, \sum_{i=1}^{n} x_{2i}y_i = 10655.224,$$

$$(\mathbf{X}^T\mathbf{X})^{-1} = \begin{pmatrix} 5.94801 & -0.0052074 & -0.169733 \\ -0.0052074 & 0.0034387 & -0.0041331 \\ -0.169733 & -0.0041331 & 0.010241 \end{pmatrix}$$

a) Use the summary statistics below to calculate $\hat{\alpha}$, $\hat{\beta}$, $\hat{\gamma}$ and RSS. Use all of the significant digits in your calculations.

b) Calculate R^2 and comment on the result.

Detergent

Engineers are experimenting with a new industrial process for manufacturing detergent. The amount y (kg) of detergent produced has been recorded when the process has been operated on 25 occasions, using different concentrations of three ingredients, x_1, x_2 and x_3. It is expected that x_1 and x_2 have the most influence on the amount produced and the following model has been fitted to the data:

$$\text{Model 1: } Y_i = \alpha + \beta x_{1i} + \gamma x_{2i} + \epsilon_i$$

Some summary statistics are:

$$RSS = 68.713, \sum_{i=1}^{n} (y_i - \bar{y})^2 = 382.731$$

The model can be written in vector-matrix form $E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$ where $\mathbf{X}$ denotes the design matrix with columns ordered as in the model above and $\hat{\boldsymbol{\beta}}$ and $(\mathbf{X}^T\mathbf{X})^{-1}$ are given below:

$$\hat{\boldsymbol{\beta}} = \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \\ \hat{\gamma} \end{pmatrix} = \begin{pmatrix} 4.952 \\ 0.704 \\ 0.939 \end{pmatrix}, \quad (\mathbf{X}^T\mathbf{X})^{-1} = \begin{pmatrix} 0.26943 & -0.02131 & -0.02302 \\ -0.02131 & 0.00520 & -0.00143 \\ -0.02302 & -0.00143 & 0.00625 \end{pmatrix}$$

a) Calculate R^2 for Model 1, which involves only x_1 and x_2.

b) If the following model:

$$\text{Model 2: } Y_i = \alpha + \beta x_{1i} + \gamma x_{2i} + \delta x_{3i} + \epsilon_i$$

is fitted to the data, the residual sum of squares is found to be 50.699.
Calculate R^2 for this model too.

c) Explain what these values of R^2 tell us about our ability to predict the amount of detergent produced from the concentration of ingredients x_1, x_2 and x_3. State an alternative measure that may be more appropriate in this context.