Topic 2: Multiple correlation coefficients

Multiple correlation

Recall our discussion at the end of slide "Properties of multivariate normal", for the best prediction in mean squares sense in case of multivariate normality: If we want to predict a random variable Y that is correlated with p random variables (predictors) $X = (X_1 \quad X_2 \quad \cdots \quad X_p)^T$ by trying to minimise the expected value $\mathrm{E}(Y - g(\boldsymbol{X})|\boldsymbol{X} = \boldsymbol{x})^2$ the optimal solution (i.e. the regression function) was $g^*(\boldsymbol{X}) = \mathrm{E}(Y \mid \boldsymbol{X})$.

When the joint $(p+1)$-dimensional distribution of Y and $\boldsymbol{X}$ is **normal** this function was **linear** in $\boldsymbol{X}$. Given a specific realisation $\boldsymbol{x}$ of $\boldsymbol{X}$ it was given by $b + \boldsymbol{\sigma}_0^\top C^{-1}\boldsymbol{x}$ where $b = \mathrm{E}(Y) - \boldsymbol{\sigma}_0^\top C^{-1}\mathrm{E}(\boldsymbol{X}), C$ is the covariance matrix of the vector $\boldsymbol{X}$, $\boldsymbol{\sigma}_0$ is the vector of Covariances of Y with $X_i, i = 1, \ldots, p$. The vector $C^{-1}\boldsymbol{\sigma}_0 \in \mathbb{R}^p$ was the *vector of the regression coefficients.*

Now, let us **define** the multiple correlation coefficient between the random variable Y and the random vector $\boldsymbol{X} \in \mathbb{R}^p$ to be the maximum correlation between Y and *any linear combination* $\boldsymbol{\alpha}^\top \boldsymbol{X}, \boldsymbol{\alpha} \in \mathbb{R}^p$. This makes sense to look at the maximal correlation that we can get by trying to predict Y as a linear function of the predictors. The solution to this which also gives us an algorithm to calculate (and estimate) the multiple correlation coefficient is given in the next lemma.

Multiple correlation coefficient as ordinary correlation coefficient of transformed data

Lemma 2.1. *The multiple correlation coefficient is the ordinary correlation coefficient between Y and $\boldsymbol{\sigma}_0^\top C^{-1} X \equiv \beta^\top X$. (I.e., $\beta \equiv C^{-1}\boldsymbol{\sigma}_0$.)*

Coefficient of Determination From Lemma 2.1 the maximum correlation between Y and any linear combination $\boldsymbol{\alpha}^\top \boldsymbol{X}, \boldsymbol{\alpha} \in \mathbb{R}^p$ is $R = \sqrt{\frac{\boldsymbol{\beta}^{*\top} C \boldsymbol{\beta}^*}{\sigma_Y^2}}$. This is the multiple correlation coefficient. Its square R^2 is called *coefficient of determination.* Having in mind that $\boldsymbol{\beta}^* = C^{-1}\boldsymbol{\sigma}_0$ we see that $R = \sqrt{\frac{\boldsymbol{\sigma}_0^\top C^{-1}\boldsymbol{\sigma}_0}{\sigma_Y^2}}$.

If $\Sigma = \begin{pmatrix} \sigma_Y^2 & \boldsymbol{\sigma}_0^\top \\ \boldsymbol{\sigma}_0 & C \end{pmatrix} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$ is the partitioned covariance matrix of the $(p+1)$-dimensional vector $(Y, \boldsymbol{X})^\top$ then we know how to calculate the MLE of Σ by $\hat{\boldsymbol{\Sigma}} = \begin{pmatrix} \hat{\boldsymbol{\Sigma}}_{11} & \hat{\boldsymbol{\Sigma}}_{12} \\ \hat{\boldsymbol{\Sigma}}_{21} & \hat{\boldsymbol{\Sigma}}_{22} \end{pmatrix}$ so the MLE of R would be $\hat{R} = \sqrt{\frac{\hat{\Sigma}_{12}\hat{\Sigma}_{22}^{-1}\hat{\Sigma}_{21}}{\hat{\Sigma}_{11}}}$.

Interpretation of R

At the end of slide "Properties of multivariate normal", we derived the minimal value of the mean squared error when trying to predict Y by a linear function of the vector $\boldsymbol{X}$. It is achieved when using the regression function and the value itself was $\sigma_Y^2 - \boldsymbol{\sigma}_0^\top C^{-1} \boldsymbol{\sigma}_0$. The latter value can also be expressed by using the value of R. It is equal to $\sigma_Y^2(1 - R^2)$.

Thus, our conclusion is that when $R^2 = 0$ there is no predictive power at all. In the opposite extreme case, if $R^2 = 1$, it turns out that Y can be predicted without any error at all (it is a true linear function of $\boldsymbol{X}$).

Activity: Numerical example

Let $\mu = \begin{pmatrix} \mu_Y \\ \mu_{X_1} \\ \mu_{X_2} \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \\ 0 \end{pmatrix}$ and $\Sigma = \begin{pmatrix} 10 & 1 & -1 \\ 1 & 7 & 3 \\ -1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} \sigma_{YY} & \boldsymbol{\sigma}_0^\top \\ \boldsymbol{\sigma}_0 & \Sigma_{XX} \end{pmatrix}$.

Question

Calculate:

1. The best linear prediction of Y using X_1 and X_2
2. The multiple correlation coefficient $R^2_{Y.(X_1,X_2)}$
3. The mean squared error of the best linear predictor.

No response

Calculation of the coefficient of determination

Remark about the calculation of R^2

Sometimes, the *correlation matrix only* may be available. It can be shown that in that case the relation

$$1 - R^2 = \frac{1}{\rho^{YY}} \tag{2.2}$$

is the upper left-hand corner of the inverse of the *correlation matrix* $\boldsymbol{\rho} \in \mathcal{M}_{p+1,p+1}$ determined from Σ.

The following proof is not examinable.

We note that the relation $\boldsymbol{\rho} = V^{-\frac{1}{2}} \Sigma V^{-\frac{1}{2}}$ holds with

$$V = \begin{pmatrix} \sigma_y^2 & 0 & 0 & \cdots & 0 \\ 0 & c_{11} & 0 & \cdots & 0 \\ 0 & 0 & c_{22} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & c_{pp} \end{pmatrix}$$

One can use 2.2 to calculate R^2 by first calculating the right hand side in 2.2. To show Equality 2.2 we note that

$$1 - R^2 = \frac{\sigma_Y^2 - \boldsymbol{\sigma}_0^\top C^{-1} \boldsymbol{\sigma}_0}{\sigma_Y^2} = \frac{|C|}{|C|} \frac{\sigma_Y^2 - \boldsymbol{\sigma}_0^\top C^{-1} \boldsymbol{\sigma}_0}{\sigma_Y^2} = \frac{|\Sigma|}{|C| \sigma_Y^2}$$

But $\frac{|C|}{|\Sigma|} = \sigma^{YY}$, the entry in the first row and column of Σ^{-1}. (Recall from slide "Inverse matrices": $(X^{-1})_{ji} = \frac{|X_{-ij}|}{|X|} (-1)^{i+j}$.) Since $\boldsymbol{\rho}^{-1} = V^{\frac{1}{2}} \Sigma^{-1} V^{\frac{1}{2}}$, we see that $\rho^{YY} = \sigma^{YY} \sigma_Y^2$ holds. Therefore $1 - R^2 = \frac{1}{\rho^{YY}}$.

Demonstration: Total correlations

Please begin by watching the following demonstration by Dr Krivitsky.

Transcript

This demonstration can be completed using the provided RStudio or your own RStudio.

To complete this task select the 'Totcor_Example.demo.Rmd' in the 'Files' section of RStudio. Follow the demonstration contained within the RMD file.

If you choose to complete the example in your own RStudio, upload the following file:

Totcor_Example.demo.Rmd

The output of the RMD file is also displayed below: