Faculty of Technology
Natural Sciences and Maritime Sciences
University of South-Eastern Norway
DDiR, January 29, 2020

On Ordinary Least Squares (OLS) regression

1 Ordinary Least Squares (OLS) regression

Consider a linear regression model of the form

$$Y = XB + E, \tag{1}$$

where $Y \in \Re^{N \times m}$ is a matrix with measured observations, $X \in \Re^{N \times r}$ is a matrix with known measured quantities. X often consists of regression variables or regressors. the elements in Y is often called regressed variables. $E \in \Re^{N \times m}$ is a matrix with unknown variables or equation errors, noise or errors due to modeling errors.

Consider the performance index

$$\begin{aligned} V(B) &= ||E||_F^2 = \text{tr}(E^T E) \\ &= \text{tr}((Y - XB)^T (Y - XB)) \\ &= \text{tr}((Y^T Y - 2Y^T XB + B^T X^T XB) \end{aligned} \tag{2}$$

This is a convex optimization problem and the solution is found from

$$\frac{\partial V(B)}{\partial B} = -2X^T Y + 2X^T XB = 0, \tag{3}$$

witch gives

$$B_{OLS} = (X^T X)^{-1} X^T Y. \tag{4}$$

The predicted output matrix $\bar{Y}$ is then given by

$$\bar{Y} = XB_{OLS} = X(X^T X)^{-1} X^T Y, \tag{5}$$

and the prediction error

$$\varepsilon = Y - \bar{Y}. \tag{6}$$

2 Examples

Example 2.1 *Given pressure and temperature measurements of saturated steam in a tank (p_k, T_k) $\forall$ $k = 1, 2, \ldots N$. A relationship between temperature and pressure may be expressed with the Clausius Clapeyron equation*

$$p = c_1 e^{c_2 T_k}. \tag{7}$$

A non-linear equation may often simply be expressed as an linear equation using the natural logarithm operator $ln()$, i.e.,

$$ln(p) = ln(c_1) + ln(e^{c_2 T_k}) = ln(c_1) + c_2 T. \tag{8}$$

We have here used that $lne^c = c$ and $ln(AB) = ln(A) + ln(B)$.

References

Syllabus: Prediction Error Method (PEM)