Journal of Applied Mathematics and Physics, 2023, 11, 421-427
https://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

Development of a Quantitative Prediction Support System Using the Linear Regression Method

Jeremie Ndikumagenge, Vercus Ntirandekura

Center of Research in Infrastructure, Environment and Technology (CRIET), University of Burundi, Bujumbura, Burundi
Email: jeremie.ndikumagenge@ub.edu.bi, ntirandekuraczi@gmail.com

How to cite this paper: Ndikumagenge, J. and Ntirandekura, V. (2023) Development of a Quantitative Prediction Support System Using the Linear Regression Method. *Journal of Applied Mathematics and Physics*, **11**, 421-427.
https://doi.org/10.4236/jamp.2023.112024

Received: November 24, 2022
Accepted: February 10, 2023
Published: February 13, 2023

Abstract

The development of prediction supports is a critical step in information systems engineering in this era defined by the knowledge economy, the hub of which is big data. Currently, the lack of a predictive model, whether qualitative or quantitative, depending on a company's areas of intervention can handicap or weaken its competitive capacities, endangering its survival. In terms of quantitative prediction, depending on the efficacy criteria, a variety of methods and/or tools are available. The multiple linear regression method is one of the methods used for this purpose. A linear regression model is a regression model of an explained variable on one or more explanatory variables in which the function that links the explanatory variables to the explained variable has linear parameters. The purpose of this work is to demonstrate how to use multiple linear regressions, which is one aspect of decisional mathematics. The use of multiple linear regressions on random data, which can be replaced by real data collected by or from organizations, provides decision makers with reliable data knowledge. As a result, machine learning methods can provide decision makers with relevant and trustworthy data. The main goal of this article is therefore to define the objective function on which the influencing factors for its optimization will be defined using the linear regression method.

Keywords

Prediction, Linear Regression, Machine Learning, Least Squares Method

1. Introduction

In this digital age, improving a system's yields is accomplished by rationalizing the mobilized resources involved in a production process through the use of optimization methods and models. To accomplish this, specialists in various fields such as political economists, statisticians, actuaries, mathematicians, and others can make significant contributions to solving certain optimization challenges such us climate factors in agriculture harvesting. Proven optimization methods can be used for this purpose.

The emergence of new data concepts such as big data or voluminous and numerous data necessitates the development of new tools, as evidenced by the rise of optimization or/and classification. Multiple linear regression models, particularly parametric models, are frequently used in data analysis procedures. The linear regression model has a wide range of applications [1]. It enables us to perform analyses and make predictions in particular. As a result, if there is a strict linear relationship between the variable to be explained or target variable and the explanatory variable or predictive variable, the prediction of the value for the target variable is unequivocal when the value for the explanatory variable is known. The model's random error term is ignored, and the magnitude of this error provides the accuracy of the established estimation [2].

In order to achieve the main goal, the present work will employ linear regression and the least squares method as mathematical tools and equipment. Furthermore, Python language utilities will be solicited for parameter value determination before discussing the obtained results and emphasizing their novelty and potential implications.

2. Materials, Tools, Equipment and Methods

2.1. Material

The spreadsheet and Python language allow you to create a linear regression model and determine the values of the model's parameters by solving the system obtained by using the least squares method.

2.2. Tools and Equipment

Sums are calculated in Excel, while python language libraries like numpy help with numerical calculations when pandas are used during the model data loading process.

2.3. Methods

When applied to the linear regression model, the least squares method yields exact and correct results. The least squares method is a tool used in all observational sciences for error theory or purely algebraic estimation [3]. It solves the linear regression model equation by determining the values of the parameters. According to [the Gauss-Markov theorem], "for a linear model, if the errors are uncorrelated and have zero expectation together with variances equal, then the least squares estimator is the best linear unbiased estimator of the coefficients" [4].

In this present work, the least squares method is used in this work to define the objective function of the model, from which a system of equations is derived by calculating the partial derivatives with respect to the model's coefficients.

2.3.1. Mathematical Modeling

Linear regression models are classified into two types: 1) simple linear regression, which employs the traditional intercept slope form and requires a and b to be learned in order to make accurate predictions; and 2) multiple linear regression, which begins with the estimation of parameters involving an endogenous variable y and p number of exogenous variables x_j.

2.3.2. Model of Linear Regression

The equations x and y represent the simple linear regression equation and the multiple linear regression equation, respectively.

$$y = ax + b \tag{1}$$

$$Y_i = a_0 + a_1 x_{i,1} + a_2 x_{i,2} + a_3 x_{i,3} + \cdots + a_p x_{i,p} + \varepsilon_i \tag{2}$$

where Y_i is the i-th observation of variable y; $x_{i,j}$ is the i-th observation of variable j-th variable; ε_i is the model's error. It summarizes the missing information that would allow the values of y to be explained linearly using the p variables x_j.

To solve the regression problem, we must estimate p + 1 parameters, which leads to the equation number (3) Written as a matrix.

$$Y = Xa + \varepsilon \tag{3}$$

The dimensions of the matrices involved in the expression of equation 3 are as follows: for Y, its dimension is (n, 1), for X, it is (n, p + 1), for a, it is (p + 1, 1), and finally for its dimension is (n, 1).

The (n, p + 1)-dimensional matrix X contains all of the observations on the exogens, with the first column formed by the value 1 indicating the integration of the constant a_0 in the model equation.

$$\begin{pmatrix} 1 & x_{1,1} & \cdots & x_{1,p} \\ 1 & x_{2,1} & \cdots & x_{2,p} \\ 1 & x_{n,1} & \cdots & x_{n,p} \end{pmatrix}$$

2.3.3. Prediction Using Linear Regression

The linear regression model is used in prediction because of three key elements. The model data (dataset) contains the questions x and answers y for the problem to be solved. This data is used to generate a model represented by a mathematical function, with the coefficients of this function serving as the model's parameters. The cost function or objective function is the set of errors in the model on the data.

3. Results and Discussion

In the next article we plan to carry out tests of the designed support on climatic data in order to predict the harvestable quantities according to the influencing climatic factors. Thus, for practical reasons, the model data (dataset) used to determine the objective function will be taken from those provided by the Geographical Institute of Burundi (IGEEBU) in 2018.

3.1. Production Estimation Based on Weather Conditions

In this study, we used test data from a sampling provided by the Geographical Institute of Burundi as shown on Table 1.

The parameters *a*, *b*, *c*, *d*, *e*, *f*, *g*, *h*, *i*, *j*, and *k* are determined by applying the least squares method to the model, which is a formulated linear function.

$$f(x_i) = ax_1 + bx_2 + cx_3 + dx_4 + ex_5 + fx_6 + gx_7 + hx_8 + ix_9 + jx_{10} + k \tag{4}$$

To begin, let's use the least squares method on the model's linear function:

$$J(a,b,c,d,e,f,g,h,i,j) = \frac{1}{2m}\sum_{i=0}^{m}\left(f(x_i) - y^{(i)}\right)^2 \tag{5}$$

$$\begin{aligned} &J(a,b,c,d,e,f,g,h,i,j) \\ &= \frac{1}{2m}\sum_{i=0}^{m}\left(ax_1 + bx_2 + cx_3 + dx_4 + ex_5 + fx_6 + gx_7 + hx_8 + ix_9 + jx_{10} + k - y^{(i)}\right)^2 \end{aligned} \tag{6}$$

Calculating the partial derivatives in relation to the linear function coefficients yields the equations as shown on Table 2.

We can deduce the system of equations from these partial derivatives calculated with respect (7).

3.1.1. Resultant 1: Gradient Descent Equation System

$$\begin{cases} a\sum x_1^2 + b\sum x_1x_2 + c\sum x_1x_3 + d\sum x_1x_4 + e\sum x_1x_5 + f\sum x_1x_6 + g\sum x_1x_7 + h\sum x_1x_8 + i\sum x_1x_9 + j\sum x_1x_{10} + k\sum x_1 = \sum yx_1 \\ a\sum x_1x_2 + b\sum x_2^2 + c\sum x_2x_3 + d\sum x_2x_4 + e\sum x_2x_5 + f\sum x_2x_6 + g\sum x_2x_7 + h\sum x_2x_8 + i\sum x_2x_9 + j\sum x_2x_{10} + k\sum x_2 = \sum yx_2 \\ a\sum x_1x_3 + b\sum x_3x_2 + c\sum x_3^2 + d\sum x_3x_4 + e\sum x_3x_5 + f\sum x_3x_6 + g\sum x_3x_7 + h\sum x_3x_8 + i\sum x_3x_9 + j\sum x_3x_{10} + k\sum x_3 = \sum yx_3 \\ a\sum x_1x_4 + b\sum x_4x_2 + c\sum x_4x_3 + d\sum x_4^2 + e\sum x_4x_5 + f\sum x_4x_6 + g\sum x_4x_7 + h\sum x_4x_8 + i\sum x_4x_9 + j\sum x_4x_{10} + k\sum x_4 = \sum yx_4 \\ a\sum x_1x_5 + b\sum x_5x_2 + c\sum x_5x_3 + d\sum x_5x_4 + e\sum x_5^2 + f\sum x_5x_6 + g\sum x_5x_7 + h\sum x_5x_8 + i\sum x_5x_9 + j\sum x_5x_{10} + k\sum x_5 = \sum yx_5 \\ a\sum x_1x_6 + b\sum x_6x_2 + c\sum x_6x_3 + d\sum x_6x_4 + e\sum x_5x_6 + f\sum x_6^2 + g\sum x_6x_7 + h\sum x_6x_8 + i\sum x_6x_9 + j\sum x_6x_{10} + k\sum x_6 = \sum yx_6 \\ a\sum x_1x_7 + b\sum x_7x_2 + c\sum x_7x_3 + d\sum x_7x_4 + e\sum x_5x_7 + f\sum x_7x_6 + g\sum x_7^2 + h\sum x_7x_8 + i\sum x_7x_9 + j\sum x_7x_{10} + k\sum x_7 = \sum yx_7 \\ a\sum x_1x_8 + b\sum x_8x_2 + c\sum x_8x_3 + d\sum x_8x_4 + e\sum x_5x_8 + f\sum x_8x_6 + g\sum x_8^2 + h\sum x_7x_8 + i\sum x_8x_9 + j\sum x_8x_{10} + k\sum x_8 = \sum yx_8 \\ a\sum x_1x_9 + b\sum x_9x_2 + c\sum x_9x_3 + d\sum x_9x_4 + e\sum x_5x_9 + f\sum x_9x_6 + g\sum x_9^2 + h\sum x_9x_8 + i\sum x_7x_9 + j\sum x_9x_{10} + k\sum x_9 = \sum yx_9 \\ a\sum x_{10}x_7 + b\sum x_{10}x_2 + c\sum x_{10}x_3 + d\sum x_{10}x_4 + e\sum x_5x_{10} + f\sum x_{10}x_6 + g\sum x_{10}^2 + h\sum x_{10}x_8 + i\sum x_{10}x_9 + j\sum x_7x_{10} + k\sum x_{10} = \sum yx_{10} \\ a\sum x_1 + b\sum x_2 + c\sum x_3 + d\sum x_4 + e\sum x_5 + f\sum x_6 + g\sum x_7 + h\sum x_8 + i\sum x_9 + j\sum x_{10} + mk = \sum y \end{cases} \tag{7}$$

The system of Equations (7) is shown in matrix form in system (8) below:

$$\begin{bmatrix} \sum x_1^2 & \sum x_1x_2 & \sum x_1x_3 & \sum x_1x_4 & \sum x_1x_5 & \sum x_1x_6 & \sum x_1x_7 & \sum x_1x_8 & \sum x_5x_8 & \sum x_5x_8 & \sum x_1 \\ \sum x_1x_2 & \sum x_2^2 & \sum x_3x_2 & \sum x_4x_2 & \sum x_5x_2 & \sum x_2x_6 & \sum x_2x_7 & \sum x_2x_8 & \sum x_5x_8 & \sum x_5x_8 & \sum x_2 \\ \sum x_1x_3 & \sum x_3x_2 & \sum x_3^2 & \sum x_3x_4 & \sum x_3x_5 & \sum x_3x_6 & \sum x_3x_7 & \sum x_3x_8 & \sum x_5x_8 & \sum x_5x_8 & \sum x_3 \\ \sum x_1x_4 & \sum x_4x_2 & \sum x_3x_4 & \sum x_4^2 & \sum x_4x_5 & \sum x_4x_6 & \sum x_4x_7 & \sum x_4x_8 & \sum x_5x_8 & \sum x_5x_8 & \sum x_4 \\ \sum x_1x_5 & \sum x_2x_5 & \sum x_3x_5 & \sum x_4x_5 & \sum x_5^2 & \sum x_5x_6 & \sum x_5x_7 & \sum x_5x_8 & \sum x_5x_8 & \sum x_5x_8 & \sum x_5 \\ \sum x_1x_6 & \sum x_2x_6 & \sum x_3x_6 & \sum x_4x_6 & \sum x_5x_6 & \sum x_6^2 & \sum x_6x_7 & \sum x_6x_8 & \sum x_5x_8 & \sum x_5x_8 & \sum x_6 \\ \sum x_1x_7 & \sum x_2x_7 & \sum x_3x_7 & \sum x_4x_7 & \sum x_5x_7 & \sum x_6x_7 & \sum x_7^2 & \sum x_7x_8 & \sum x_5x_8 & \sum x_5x_8 & \sum x_7 \\ \sum x_1x_8 & \sum x_2x_8 & \sum x_3x_8 & \sum x_4x_8 & \sum x_5x_8 & \sum x_6x_8 & \sum x_7x_8 & \sum x_8^2 & \sum x_9x_8 & \sum x_{10}x_8 & \sum x_8 \\ \sum x_1x_9 & \sum x_2x_9 & \sum x_3x_9 & \sum x_4x_9 & \sum x_5x_9 & \sum x_9x_6 & \sum x_9x_7 & \sum x_9x_8 & \sum x_9^2 & \sum x_{10}x_9 & \sum x_9 \\ \sum x_1x_{10} & \sum x_2x_{10} & \sum x_3x_{10} & \sum x_4x_{10} & \sum x_5x_{10} & \sum x_6x_{10} & \sum x_7x_{10} & \sum x_{10}x_8 & \sum x_{10}x_9 & \sum x_{10}^2 & \sum x_{10} \\ \sum x_1 & \sum x_2 & \sum x_3 & \sum x_4 & \sum x_5 & \sum x_6 & \sum x_7 & \sum x_8 & \sum x_9 & \sum x_{10} & m \end{bmatrix} * \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \\ i \\ j \\ k \end{bmatrix} = \begin{bmatrix} \sum yx_1 \\ \sum yx_2 \\ \sum yx_3 \\ \sum yx_4 \\ \sum yx_5 \\ \sum yx_6 \\ \sum yx_7 \\ \sum yx_8 \\ \sum yx_9 \\ \sum yx_{10} \\ \sum y \end{bmatrix} \tag{8}$$

Table 1. Dataset.

X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	Y
10.2	15	24	15	10.21	7.26	25	2.6	23	6.53	247.25
11.3	11.9	21	14	14.2	8.6	28	3.45	29	7.3	1147.569
5.148	4.89	17.6	47	11.6	1.15	29.54	5.9	36	8.95	958.25
1.575	2.56	3.75	5.94	2.145	1.055	39.95	7.5	45.457	17.5	856.545
8.7	1.75	5.45	6	7	5.96	35.25	9.54	31.015	11.57	915.75
9.5	2.65	7.15	9	19.015	6.25	27.97	10.25	25	12.15	715.685

X_1: The solar radiation Level, X_2: Water stress level, X_3: Temperature of the air, X_4: Soil depth, X_5: Temperature of the soil, X_6: Evaporation rate, X_7: Precipitation quantity, X_8: Wind speed, X_9: Soil Humidity, X_{10}: represents relative air Humidity, and Y: represents Production.

Table 2. Least square calculation.

Coefficients	Derivatives with regard to	Partial derivative expressions
a	$\frac{\partial J(a,b,c,d,e,f,g,h,i,j)}{\partial a}$	$\frac{1}{m}\sum_{i=0}^{m}\left(ax_1+bx_2+cx_3+dx_4+ex_5+fx_6+gx_7+hx_8+ix_9+jx_{10}+k-y^{(i)}\right)x_1$
b	$\frac{\partial J(a,b,c,d,e,f,g,h,i,j)}{\partial b}$	$\frac{1}{m}\sum_{i=0}^{m}\left(ax_1+bx_2+cx_3+dx_4+ex_5+fx_6+gx_7+hx_8+ix_9+jx_{10}+k-y^{(i)}\right)x_2$
c	$\frac{\partial J(a,b,c,d,e,f,g,h,i,j)}{\partial c}$	$\frac{1}{m}\sum_{i=0}^{m}\left(ax_1+bx_2+cx_3+dx_4+ex_5+fx_6+gx_7+hx_8+ix_9+jx_{10}+k-y^{(i)}\right)x_3$
d	$\frac{\partial J(a,b,c,d,e,f,g,h,i,j)}{\partial d}$	$\frac{1}{m}\sum_{i=0}^{m}\left(ax_1+bx_2+cx_3+dx_4+ex_5+fx_6+gx_7+hx_8+ix_9+jx_{10}+k-y^{(i)}\right)x_4$
e	$\frac{\partial J(a,b,c,d,e,f,g,h,i,j)}{\partial e}$	$\frac{1}{m}\sum_{i=0}^{m}\left(ax_1+bx_2+cx_3+dx_4+ex_5+fx_6+gx_7+hx_8+ix_9+jx_{10}+k-y^{(i)}\right)x_5$
f	$\frac{\partial J(a,b,c,d,e,f,g,h,i,j)}{\partial f}$	$\frac{1}{m}\sum_{i=0}^{m}\left(ax_1+bx_2+cx_3+dx_4+ex_5+fx_6+gx_7+hx_8+ix_9+jx_{10}+k-y^{(i)}\right)x_6$
g	$\frac{\partial J(a,b,c,d,e,f,g,h,i,j)}{\partial g}$	$\frac{1}{m}\sum_{i=0}^{m}\left(ax_1+bx_2+cx_3+dx_4+ex_5+fx_6+gx_7+hx_8+ix_9+jx_{10}+k-y^{(i)}\right)x_7$
h	$\frac{\partial J(a,b,c,d,e,f,g,h,i,j)}{\partial h}$	$\frac{1}{m}\sum_{i=0}^{m}\left(ax_1+bx_2+cx_3+dx_4+ex_5+fx_6+gx_7+hx_8+ix_9+jx_{10}+k-y^{(i)}\right)x_8$
i	$\frac{\partial J(a,b,c,d,e,f,g,h,i,j)}{\partial i}$	$\frac{1}{m}\sum_{i=0}^{m}\left(ax_1+bx_2+cx_3+dx_4+ex_5+fx_6+gx_7+hx_8+ix_9+jx_{10}+k-y^{(i)}\right)x_9$
j	$\frac{\partial J(a,b,c,d,e,f,g,h,i,j)}{\partial j}$	$\frac{1}{m}\sum_{i=0}^{m}\left(ax_1+bx_2+cx_3+dx_4+ex_5+fx_6+gx_7+hx_8+ix_9+jx_{10}+k-y^{(i)}\right)x_{10}$

3.1.2. Resultat 2: Factor Values or Climate Parameters

The application of the least squares method to the model's test data yields the effective values of the model's parameters as shown by the system results (9)

$$\begin{bmatrix} 426.652529 & 357.07572 & 693.95105 & 700.2115 & 569.2397 & 290.040825 & 1358.78317 & 288.0637 & 1326.553275 & 438.8171 & 46.423 \\ 357.07572 & 407.1607 & 734.049 & 670.9864 & 446.985 & 246.5568 & 1090.7306 & 171.9635 & 1103.03617 & 325.8305 & 38.75 \\ 693.95105 & 734.049 & 1421.6475 & 1600.525 & 929.551 & 456.20575 & 2249.8145 & 392.0955 & 2312.8455 & 683.094 & 78.95 \\ 700.2115 & 670.9864 & 1600.525 & 2782.2836 & 1123.026 & 381.6267 & 2855.913 & 558.64 & 3124.10458 & 903.52 & 96.94 \\ 569.2397 & 446.985 & 929.551 & 1123.026 & 855.6154 & 372.411325 & 1859.8063 & 421.74725 & 1854.215265 & 623.71105 & 64.17 \\ 290.040825 & 246.5568 & 456.20575 & 381.6267 & 372.411325 & 203.687225 & 883.32075 & 184.1644 & 846.836535 & 283.8375 & 30.275 \\ 1358.78317 & 1090.7306 & 2249.8145 & 2855.913 & 1859.8063 & 883.32075 & 5902.4975 & 1258.4885 & 60058.9759 & 2078.836 & 185.71 \\ 288.0637 & 171.9635 & 392.0955 & 558.64 & 421.74725 & 184.1644 & 1258.4885 & 305.7966 & 1265.3106 & 461.1333 & 39.24 \\ 1326.553275 & 1103.03617 & 2312.8455 & 3124.10458 & 1854.215265 & 846.836535 & 60058.9759 & 1265.3106 & 6319.269074 & 2142.18105 & 189.472 \\ 438.8171 & 325.8305 & 683.094 & 903.52 & 623.71105 & 283.8375 & 2078.836 & 461.1333 & 2142.18105 & 763.7708 & 64 \\ 46.423 & 38.75 & 78.95 & 96.94 & 64.17 & 30.275 & 185.71 & 39.24 & 189.472 & 64 & 6 \end{bmatrix} * \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \\ i \\ j \\ k \end{bmatrix} = \begin{bmatrix} 36537.64258 \\ 27742.54655 \\ 60218.178 \\ 81836.0083 \\ 51791.8916 \\ 23600.67213 \\ 153136.7567 \\ 32751.7518 \\ 158693.3283 \\ 52848.47145 \\ 4380762.654 \end{bmatrix} \tag{9}$$

We obtain the following values of the following parameters after solving the system (9):

$$a = 15022653.083623783662915$$

$$b = 19087801.322295062243938$$

$$c = -19617686.517314746975898$$

$$d = 4433188.079017613083124$$

$$e = -0.037048308013294$$

$$f = -4477342.56212795432657$$

$$g = 293402.8806244044099$$

$$h = -10060668.647367989644408$$

$$i = -7.433182264782304$$

$$j = 2466614.860606360249221$$

$$k = 6.230437622139735$$

The solving system (9) returns the values of the final model's coefficients, as expressed:

$$\begin{aligned} & f\left(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9\right) \\ & = 15022653.083 x_1 + 19087801.322 x_2 - 19617686.517 x_3 + 4433188.079 x_4 \\ & \quad - 0.037 x_5 - 4477342.562 x_6 + 293402.880 x_7 - 10060668.647 x_8 \\ & \quad + 2466614.860 x_9 + 6.230 \end{aligned}$$

3.2. Discussion on the Obtained Results

Two results were obtained after applying the model to the study data (dataset).

1) A system of equations derived from study data using the law of the smallest squares and linear regression.

2) The values of the model's coefficients or parameters, which can be used to minimize or maximize the differences between the final and initial models.

3) The objective function found constitutes a quantitative prediction support which can be used in various fields to estimate the values of indicators of a given process involving and interacting quantifiable and countable input factors. For the last one, at the output, the results or products obtained are themselves also quantifiable, countable and optimal according to the case.

4) The determination of the influencing factors using the gradient descent method makes it possible to minimize or maximize the objective function which

ultimately can be used for prediction purposes.

A subsequent work will elucidate and investigate the avenues of application of this fourth result using case studies that trace real-world phenomena.

4. Conclusions

The objective function must be determined. Multiple linear regression allows for the determination of an objective function, which can then be optimized by adjusting the influencing factors. The precision of the influencing factors required to obtain an optimal yield has been obtained using the method of gradient descent and can be used for quantitative prediction processes or/and work.

The solution based on least squares methods coupled with multiple linear regression allowed for the determination of an objective function. The specification of influencing factors, combined with the use of gradient descent methods, transforms the latter into a tool, a support for quantitative prediction.

The use of a linear regression model, one of the artificial intelligence supervised learning methods, is what distinguishes this work from others. The work goes beyond the commonly used decision-making approaches. It focuses on prediction modeling for decision support systems in particular.

This final point will be addressed in future work. Future research will particularly concentrate on the specifications of the influencing factors of the objective function, as requested during the optimization process using the gradient descent method.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Etemadi, S. and Khashei, M. (2021) Etemadi Multiple Linear Regression. *Measurement*, **186**, 110080. https://doi.org/10.1016/j.measurement.2021.110080

[2] Schweppe, F.C. (1970) Power System Static-State Estimation. *IEEE Transactions on Power Apparatus and Systems*, 135.

[3] Helland, I.S. (1990) Partial Least Squares Regression and Statistical Models. *Scandinavian Journal of Statistics*, **17**, 97-114. https://www.jstor.org/stable/4616159.

[4] Lewis, P.T. (1966) A Generalization of the Gauss-Markov Theorem. *Journal of the American Statistical Association*, **61**, 1063-1066. https://doi.org/10.2307/2283200