IJFANS International Journal of Food and Nutritional Sciences

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 10, Spl Iss 1, 2021

A Study On Multiple Linear Regression Using Matrix Calculus

B. V. APPA RAO, Department of Physics, Koneru Lakshmaiah Education Foundation, Green Fields, Guntur District, Vaddeswaram, Andhra Pradesh 522502, India

Abstract Despite the availability of a plethora of imaginative tools in Applied Mathematics, the preeminent instrument that persists for mathematicians is the linear model. This model exhibits straightforward and seemingly restrictive attributes, including linearity, constancy of variance, normality, and independence. Linear models, along with their associated methodologies, are distinguished for their remarkable flexibility and potency. Given that nearly all advanced statistical techniques stem from generalizations of linear models, competence in this domain becomes a prerequisite for the comprehensive study of advanced statistical tools.

The central focus of this research article centers around the specific formulations of the Simple Linear Regression Model and the Multiple Linear Regression Model. It delves into the Least Squares Estimation (LSE) of their parameters and elucidates the properties inherent to LSE. Moreover, an innovative proof of the Gauss-Markov theorem is introduced, employing the Principles of Matrix Calculus as a foundational basis. Additionally, the article portrays the concept of Best Linear Unbiased Estimators (BLUE), underscoring its significance within the context of this study.

1. INTRODUCTION

A better understanding of a given phenomenon is facilitated by a model encompassing a theoretical framework. This model, a mathematical construction, gives rise to the observed data. The linear model serves as a foundational element in the education of both theoretical and applied mathematicians. The scientific method frequently serves as a guided approach to the process of learning.

Linear statistical models find extensive application within these processes and are utilized in diverse fields including the biological, physical, and social sciences, as well as business and engineering. While these mathematically constructed models may oversimplify complex real-world problems, they offer valuable approximations of the relationships among observations. The importance of accurate parameter estimators cannot be overstated, as they significantly enhance predictive performance. Scientists and engineers rely on estimated models to describe and synthesize observed data.

In 2017, B. Mahaboob et al. presented a paper where they utilized the principles of Matrix Calculus to estimate parameters of a CES production functional model. Similarly, in 2018, C. Narayana et al. explored misspecification and predictive accuracy of stochastic linear regression models. B. Mahaboob et al., in another research article in 2017, introduced computational techniques for the least squares estimator and maximum likelihood estimator through the application of Matrix Calculus. Their 2019 paper tackled estimation methods for the Cobb-Douglas production functional model.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 10, Spl Iss 1, 2021

Kushbukumari et al. provided an elucidation of basic concepts in Linear Regression Analysis in their 2018 paper, detailing calculations in software like SPSS and Excel. In the same year, Shrikant I. Bangdiwala comprehensively explained methods for fitting simple linear regression models in his article titled "Regression: simple linear explained". Further, in 2017, W. Superta et al. developed a numerical model utilizing nonlinear techniques to estimate thunderstorm activity.

For additional references, please consult sources [1]-[28].

2. SIMPLE LINEAR REGRESSION MODEL

In simple linear regression model an attempt can be made to model the relationship between the two variables by the form

$$Y = \alpha_0 + \alpha_1 X + \epsilon \,. \qquad (2.1)$$

where
Y=Response Variable (Dependent Variable)
X=Predictor Variable (Independent Variable)
ϵ=Random Variable (Error Term)

The linearity appearing in (2.1) is just an assumption. Some more assumptions namely distribution of independence of Y can be added.α_0 and α_1 can be estimated by making use of observed values of X and Y and some inferences namely confidence intervals and testing of hypothesis for β_0 and$\beta_1$can be made.The estimated model can be used to forecast the value of$Y$for a particular value of$X$.

2. MULTIPLE LINEAR REGRESSION MODEL

The dependent variable Y can be influenced by multiple independent variables. It is possible to formulate a linear model that establishes a relationship between Y and several predictors.

$$Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \ldots\ldots + \alpha_i X_i + \epsilon. \qquad (3.1)$$

The regression coefficients are denoted as arbitrary constants. The error term accounts for the random variation in Y that cannot be attributed to the predictors. The model presented in equation (3.1) exhibits linearity in the regression coefficients ('s) but not necessarily in the predictors. To derive estimates for the regression coefficients ('s) in equation (3.1), one can utilize a sample comprising 'n' observations on Y along with the corresponding X values. For the kth observation, the model can be expressed as follows

ISSN PRINT 2319 1775 Online 2320 7876

Research paper **© 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 10, Spl Iss 1, 2021**

(3.2) $$Y_k = \alpha_0 + \alpha_1 X_{k1} + \alpha_2 X_{k2} + ... + \alpha_i X_{ki} + \epsilon_k \,.$$

The assumptions for ϵ_k are as follows:

(i) $E(\epsilon_k) = 0$ for $k = 1, 2, ...m$, that is $E(Y_k) = \alpha_0 + \alpha_1 X_{k1} + \alpha_2 X_{k2} + ... + \alpha_i X_{ki}$.

(ii) $Var(\epsilon_k) = \sigma^2$ for $k = 1, 2, \ldots m$, that is $Var(Y_k) = \sigma^2$.

(iii) $Cov(\epsilon_k, \epsilon_l) = 0$ for all $k \neq l$, that is $Cov(Y_k, Y_l) = 0$.

(i) tells us that model is correct i.e., it is indeed linear, and
(ii) implies that variance of$Y$ is constant.
If all the above assumptions hold good the poor estimator will be achieved. If the normality assumption is inserted one can obtain MLE which can have excellent properties. For each of 'n' observations (3.2) can be depicted as

$$\begin{aligned} Y_1 &= \alpha_0 + \alpha_1 X_{11} + \alpha_2 X_{12} + ... + \alpha_i X_{li} + \epsilon_1 \\ Y_2 &= \alpha_0 + \alpha_1 X_{21} + \alpha_2 X_{22} + ... + \alpha_i X_{2i} + \epsilon_2 \\ & \ldots \\ Y_m &= \alpha_0 + \alpha_1 X_{m1} + \alpha_2 X_{m2} + ... + \alpha_i X_{mi} + \epsilon_m \,. \end{aligned}$$

Using matrices the above can be put as

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_m \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{11} & \ldots & X_{1i} \\ 1 & X_{21} & X_{22} & \ldots & X_{2i} \\ \vdots & \vdots & \ddots & \vdots & \\ 1 & X_{m1} & X_{m2} & \ldots & X_{mi} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_i \end{bmatrix} \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_m \end{bmatrix} .$$

(3.3) $$\bar{Y} = \bar{X}\bar{\alpha} + \bar{\epsilon}.$$

hence

$$\begin{aligned} E(\bar{\epsilon}) &= \bar{0} \text{ or } E(\bar{y}) = \bar{X}\bar{\alpha} \\ Cov(\bar{\epsilon}) &= \sigma^2 \bar{I} \text{ or } Cov(\bar{Y}) = \sigma^2 \bar{I} \,. \end{aligned}$$

X is a matrix of order $m \times (i+1)$ and assume that $m > i+1$ and $\rho(X) = i+1$. The α parameters in (2.1) or (3.3) are called regression coefficients. The partial derivative of

$$E(Y) = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \cdots \alpha_i X_i \,,$$

with respect to X_2 for example is α_2.Hence regression coefficients are sometimes referred to as partial regression coefficients. Thus α_2 indicates the change in $E(Y)$ with a unit change in X_2 when $x_1, X_2, X_3 \ldots X_i$ are fixed. α_2 depicts the influence of X_2 of $E(Y)$ in the presence of other$X$'s. This effect is different from the effect of X_2 on $E(Y)$ if the other X's are not appearing in the model. For instance α_0 and α_1

$$Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \epsilon$$

differ from α_0^* and α_1^* in

$$Y = \alpha_0^* + \alpha_1^* X_1 + \epsilon^* .$$

Roughly speaking if one is X is deleted then the parameters will be changed.

3. PRINCIPLE DUE TO GAUSS-MARKOV

The condition of unbiasedness is represented by the equation A X = I, as the relationship A X = I must be valid for all possible values of . The covariance matrix for the estimation AA is derived in the following manner

$$Cov(\bar{A}\bar{Y}) = \bar{A}(\sigma^2 \bar{I})\bar{A}^{-1} = \sigma^2 \bar{A}\bar{A}^{-1} .$$

The variance of the α_i 's are on the diagonal of $\sigma^2 \bar{A}\bar{A}^T$. So one have to choose $\bar{A}$ subject to $\bar{A}\bar{X} = \bar{I}$ and hence the diagonal elements of $\bar{A}\bar{A}^T$ are minimized. To relate $\bar{A}\bar{Y}$ to $\hat{\bar{\alpha}} = (\bar{X}^T\bar{X})^{-1}\bar{X}^T\bar{Y}$ one can add and subtract $(\bar{X}^T\bar{X})^{-1}\bar{X}^T$ to get

$$\bar{A}\bar{A}^T = [\bar{A} - (\bar{X}^T\bar{X})^{-1}\bar{X}^T + (\bar{X}^T\bar{X})^{-1}\bar{X}^T][\bar{A} - (\bar{X}^T\bar{X})^{-1}\bar{X}^T + (\bar{X}^T\bar{X})^{-1}\bar{X}^T]^T .$$

Expanding this in terms of $\bar{A} - (\bar{x}^T\bar{X})^{-1}\bar{X}^T$ and $(\bar{X}^T\bar{X})^{-1}\bar{X}^T$one can obtain 4 terms in which two become 0 according to $\bar{A}\bar{X} = \bar{I}$. Thus the output is

$$\bar{A}\bar{A}^T = [\bar{A} - \bar{X}^T\bar{X})^{-1}][\bar{A} - \bar{X}^T\bar{X})^{-1}\bar{X}^T]^T . \tag{6.1}$$

The matrix on the RHS of (6.1) is positive semi definite and its diagonal elements are 0. There elements can be made equal to 0 by taking $\bar{A}$ as $\bar{A} = (\bar{X}^T\bar{X})^{-1}\bar{X}^T$. The final minimum variance estimation of $\bar{\alpha}$ is $\bar{A}\bar{Y} = (\bar{X}^T\bar{X})^{-1}\bar{X}^T\bar{Y}$. This is nothing but the LSE of $\bar{\alpha}$.

4. CONCLUSIONS AND FUTURE RESEARCH

This research article outlines the specification of both simple and multiple linear regression models. It introduces a streamlined representation of the multiple linear regression model using Matrix Algebra. The process of least square estimation for regression coefficients is elucidated through Matrix Calculus. Additionally, the properties inherent to the Least Square Estimation (LSE) are established, accompanied by a novel proof of the Gauss-Markov Theorem.

Looking ahead to future research, the aforementioned concepts can also be applied to simple linear regression models. Furthermore, the exploration of additional concepts, such as Geometry Least Squares, Maximum Likelihood Estimation (MLE), and Generalized Least Squares, holds potential for further advancement and exploration within this field.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 10, Spl Iss 1, 2021

REFERENCES

[1] KUSHBUKUMARI, SUNITIYADAV: Linear Regression Analysis Study, Journal of the Prac tice of Cardiovascular Sciences., 4(1) (2018), 33-66

[2] S. I. BANGDIWALA: Regression: simplelinear,, International Journal of Safety Control and Safety Promotion, 25(1) (2018), 113-115.

[3] W. SUPARTA: UsingMultipleLinear Regression Model to Estimate Thunderstorm Activity, IOP Conference Series:Materials Science and Engineering, 185 (2017), ID012023.

[4] B. MAHABOOB: On Cobb-Douglas Production Function Model, AIP Conference Proceed ings Recent Trends in Pure and Applied Mathematics, 2019AIP Conf. Proc. 2177020040 1–020040-4; https://doi.org/10.1063/1.5135215, 2019.

[5] B.MAHABOOB: Criteria for Selection of Stochastic Linear Model Selection, AIP Confer ence Proceedings Recent Trends in Pure and Applied Mathematics, 2019 AIP Conf. Proc. 2177020040–1–020040-4; https://doi.org/10.1063/1.5135215, 2019.

[6] B. VENKATESWARLU: Application of DEA in Super Efficiency Estimation, International Journal of Scientific Technology and Research, 9(2) (2020), 4496-4499.

[7] B. VENKATESWARLU: Evaluating Different Types of Efficiency Stability Regions and their Infeasibility in DEA, International Journal of Scientific Technology and Research, 9(2) (2020), 3944-3949.

[8] B. VENKATESWARLU: Multi-Criteria Optimization,Techniques in DEA: Methods and Ap plications, International Journal of Scientific Technology and Research, 9(2) (2020), 509 515.

[9] B. MAHABOOB: An Evaluation in Generalized LSE of LinearizedStochasticStatisti cal Model with Non-Spherical Errors, AIP Conference Proceedings, Recent Trends in Pure and Applied Mathematics, (2019) AIP Conf. Proc. 2177, 020038-1-020038-5, https://doi.org/10.1063/1.5135213, 2019.

[10] B. MAHABOOB: On Misspecification Tests for Stochastic Linear, Regression Model, AIP Conference Proceedings Recent Trends in Pure and Applied Mathematics, 2019 AIP Conf. Proc. 2177020040–1–020040-4; https://doi.org/10.1063/1.5135215, 2019.

[11] Y. SOBHAN BABU, B. V .APPA RAO, T. SRINIVASA RAO: Realizability of matrix riccati type dynamical systems on time scales, International Journal of Civil Engineering and Technology, 8(12) (2017), 216-223.

[12] Y. SOBHAN BABU, B. V. APPA RAO, T. SRINIVASA RAO, K. A. S. N. V. PRASAD: Addaptive control and realizability on nabla settings, International Journal of Mechanical Engineering and Technology, 8(12) (2017), 236-324.

[13] CH. VASAVI, G. S. KUMAR, T. S. RAO, B. V. APPA RAO: Application of fuzzy differen tial equations for cooling problems, International Journal of Mechanical Engineering and Technology, 8(12) (2017), 712-721.

[14] T. S. RAO, G. S. KUMAR, CH. VASAVI, B. V. APPA RAO: On the controllability of fuzzy difference control systems, International Journal of Mechanical Engineering and Technology, 8(12) (2017), 723-732.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 10, Spl Iss 1, 2021

[15] B. BHASKARA RAMA SARMA, V. V. KUMAR, S. V. N. L. LALITHA: Modified estima tor for mean of log-normal population, Journal of Advanced Research in Dynamical and Control Systems, 10(2) (2018), 386-388