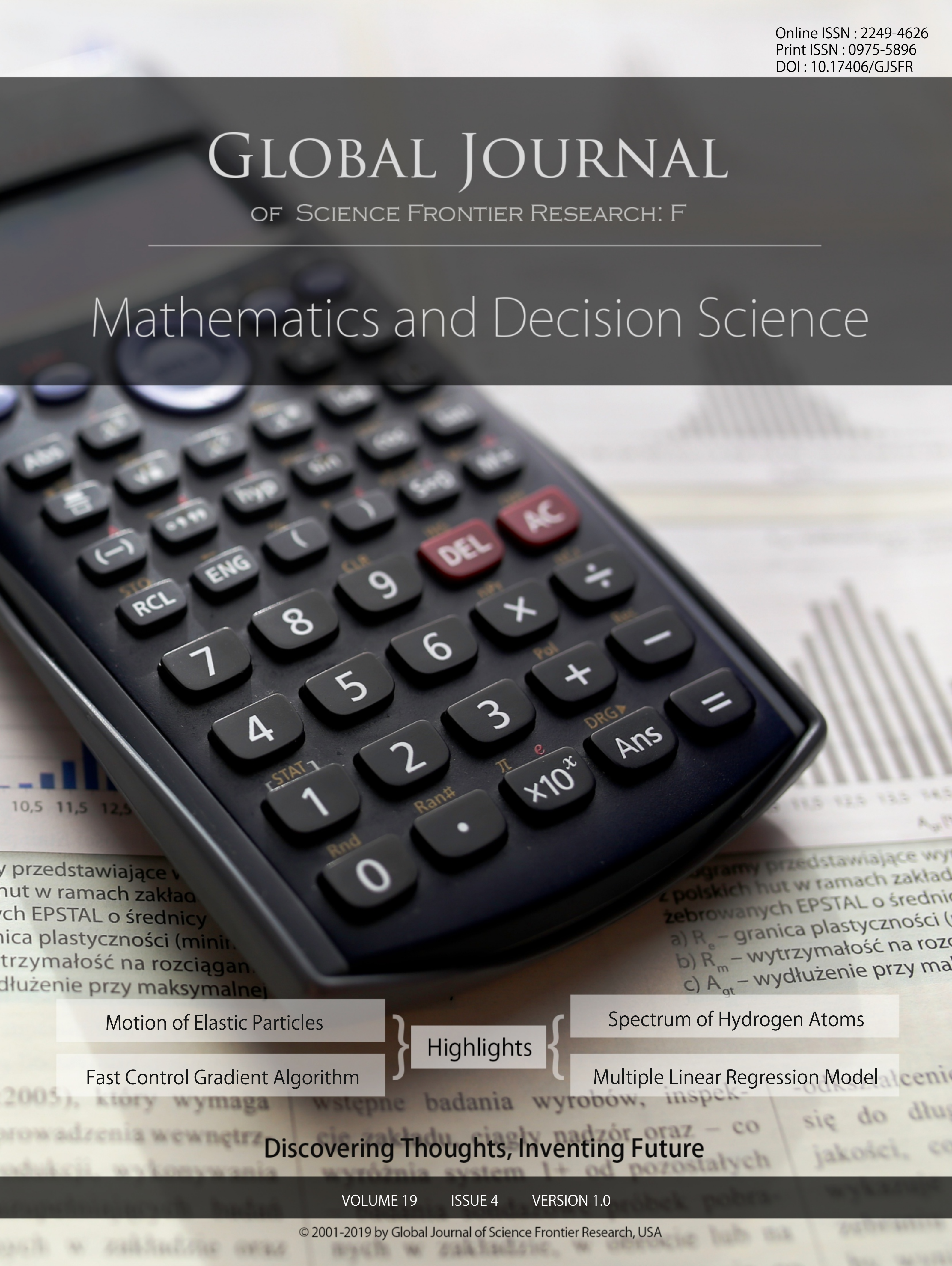
Online ISSN : 2249-4626
Print ISSN : 0975-5896
DOI : 10.17406/GJSFR

Global Journal

of Science Frontier Research: F

Mathematics and Decision Science



Highlights

- Motion of Elastic Particles
- Fast Control Gradient Algorithm
- Spectrum of Hydrogen Atoms
- Multiple Linear Regression Model

Discovering Thoughts, Inventing Future

VOLUME 19 ISSUE 4 VERSION 1.0

Global Journal of Science Frontier Research: F
Mathematics & Decision Sciences

Global Journal of Science Frontier Research: F
Mathematics & Decision Sciences

Volume 19 Issue 4 (Ver. 1.0)

Open Association of Research Society

Incorporation No.: 0423089
License No.: 42125/022010/1186
Registration No.: 430374
Import-Export Code: 1109007027
Employer Identification Number (EIN):
USA Tax ID: 98-0673427

Global Journals Inc.

*(A Delaware USA Incorporation with "Good Standing"; **Reg. Number: 0423089**)*

Sponsors: *Open Association of Research Society*
Open Scientific Standards

Publisher's Headquarters office

Global Journals® Headquarters
945th Concord Streets,
Framingham Massachusetts Pin: 01701,
United States of America
USA Toll Free: +001-888-839-7392
USA Toll Free Fax: +001-888-839-7392

Offset Typesetting

Global Journals Incorporated
2nd, Lansdowne, Lansdowne Rd., Croydon-Surrey,
Pin: CR9 2ER, United Kingdom

Packaging & Continental Dispatching

Global Journals Pvt Ltd
E-3130 Sudama Nagar, Near Gopur Square,
Indore, M.P., Pin:452009, India

Find a correspondence nodal officer near you

To find nodal officer of your country, please email us at *local@globaljournals.org*

eContacts

Press Inquiries: *press@globaljournals.org*
Investor Inquiries: *investors@globaljournals.org*
Technical Support: *technology@globaljournals.org*
Media & Releases: *media@globaljournals.org*

Pricing (Excluding Air Parcel Charges):

Yearly Subscription (Personal & Institutional)
250 USD (B/W) & 350 USD (Color)

Contents of the Issue

Global Journal of Science Frontier Research: F
Mathematics and Decision Sciences
Volume 19 Issue 4 Version 1.0 Year 2019
Type : Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals
Online ISSN: 2249-4626 & Print ISSN: 0975-5896

Fast Control Gradient Algorithm for Simple and Multiple Linear Regression Model

By Abdelkrim El Mouatasim
Ibn Zohr University

Abstract- Information is steadily increasing and hungry for knowledge. As the data grows, the world moves in on hunting knowledge with the help of analytics in the big data era. Flood data arising from diverse fields are described for automated learning technique of data analysis is intended as a machine learning, like classiffication and regression, which is a statistical method of predictive analysis.

We proposed in this paper, gradient method with control step and Nestrove step called fast control gradient (FCG) algorithm for multiple linear regression (MLR), the quadratic convergence rate $o(k^2)$ of FCG algorithm are proved. FCG algorithm are applicate to a real dataset of wine quality for simple linear regression and dataset of combined cycle power plan (CCPP) for multiple linear regression. The numerical experiment, show that our approach FCG algorithm is faster than gradient descent (GD) algorithm.

Keywords: *gradient algorithm, learning rate, linear regression model and mean square error.*

GJSFR-F Classification: *MSC 2010: 62J05*

FASTCONTROLGRADIENTALGORITHMFORSIMPLEANDMULTIPLELINEARREGRESSIONMODEL

Strictly as per the compliance and regulations of:

Ref
11. Ever Y.K, Dimililer K. and Sekeroglu B., Comparison of Machine Learning Techniques for Prediction Problems. AISC 927, pp. 713-723, 2019.

Fast Control Gradient Algorithm for Simple and Multiple Linear Regression Model

Abdelkrim El Mouatasim

Abstract- Information is steadily increasing and hungry for knowledge. As the data grows, the world moves in on hunting knowledge with the help of analytics in the big data era. Flood data arising from diverse fields are described for automated learning technique of data analysis is intended as a machine learning, like classiffication and regression, which is a statistical method of predictive analysis.

We proposed in this paper, gradient method with control step and Nestrove step called fast control gradient (FCG) algorithm for multiple linear regression (MLR), the quadratic convergence rate o(k^2) of FCG algorithm are proved. FCG algorithm are applicate to a real dataset of wine quality for simple linear regression and dataset of combined cycle power plan (CCPP) for multiple linear regression. The numerical experiment, show that our approach FCG algorithm is faster than gradient descent (GD) algorithm.

Keywords: gradient algorithm, learning rate, linear regression model and mean square error.

I. Introduction

Machine learning is automated learning technique consisting of certain methods that define patterns to predict potential data, or to execute critical decision-making in uncertainties.

Deep learning [13], face recognition [24] and autism spectrum disorder detection [5, 20] are some of the phrases defined by machine learning accurately. This refers to the transformation to complete a function related to artificial intelligence.

The designed model must maintain the features of the algorithm that rely on modifications in its environment and calculate the necessary actions by predicting the effects [3].

Predicting a real value (the target variable that we're trying to predict is continuous) which is known as regression problem, most common problem researched in machine learning [16].

For this reason, machine learning methods are used to control the system response to predict a digital target feature or real value.

Many real-life problems such as regression problems can be solved using machine learning methods to develop prediction models [11], such as water-demand forecasting [1].

Regression analysis is one of the fundamental scientific investigation tools enabling functional identification of the relationship between independent and dependent variables [12]. In classical regression analysis both independent and dependent variables are given as real numbers. However, in many real-life situations, where the complexity of the physical system dictates adoption from a more general view, regression variables are given as non-numerical entities such as linguistic variables. Unfortunately, such realistic situations are beyond the scope of classical regression analysis [2].

Linear regression analysis can be defined as the modeling approach for examining the linear relationship between a dependent variable Y and:

Author: Ibn Zohr University, Faculty of Polydisciplinary Ouarzazate (FPO), B.P. 284, Ouarzazate 45800, Morocco.
e-mail: a.elmouatasim@uiz.ac.ma

- an independent variable X called simple regression, or
- a set of independent variables $X_1, X_2, \ldots, X_k$ called multiple regression,

that can explain the amount of variation in Y.

Data inherent error. When data appear, rigorous error techniques must be used to fit the "best" curve with the data. Otherwise, prediction of intermediate values or value derivatives may produce unsatisfactory results. Visual inspection may be used to fit the "best" line through data points, but this method is very subjective. Some criteria should be established as a basis for suitability. One criterion is to derive a curve that reduces the discrepancy between the data points and the curve. The ordinary least squares (OLS) is one method to achieve this goal.

While using the OLS criterion as a performance index we let the MLR problem as unconstrained optimization problem, that can solved by using normal equation [8, 9], GD algorithm, the process which uses cost function on gradients and learning rate for minimizing the complexity in computing mean square error.

However, numerical experiments and theoretical analysis have shown that this learning rate results as a rule in slow convergence and further development in line with stochastic gradient descent algorithm [14] and control subgradient algorithm [7, 10]. On the other hand gradient algorithms still demonstrated quite satisfactory performance under special conditions, for example when $\mathbf{f}^* = \mathbf{f}(\mathbf{x}^*)$ is known or well-estimated so special learning rate control rule can be engaged.

In this paper, we propose a FCG algorithm a fast iterative algorithm for MLR with independents and dependent variables. We discuss the complexity convergence of FCG the proposed algorithm and present some promising results of numerical experiments application to CCPP. We compared our approach FCG algorithm with GD algorithm. These algorithms are implemented in this paper using python programming tool for analyzing CCPP dataset [3, 21]:

> In an effort to find accurate and effective ways to predict hourly electric power production, this study uses a 6-year dataset (2006-2011) whose data points correspond to the average sensor measurements per hour when the plant is set to work with full load. Input features are relative humidity, ambient temperature and ambient pressure which are known to be the main factors in gas turbines as well as vacuum exhaust measured from steam turbine.

This paper is organized in as drafted below in section 1 introduction. Section 2 notations and assumptions. Section 3 linear regression. In section 4 gradient algorithms. And finally the results are discussed in section 5 and section 6 concludes.

II. Notations and Assumptions

First, we establish notation for future use:

- features x_i is the input variables,
- target y_i is the output variable that we are trying to predict,
- training example is a pair (x_i, y_i),
- training set is the dataset that we'll be using to learn a list of n training examples

$$\{(x_i, y_i) : i = 1, \ldots, n\},$$

Note that the superscript $'i'$ in the above notation is simply an index into the dataset.
We will also use the following notation:

- $E = \Re^n$ is a finite dimensional Euclidean space of input variables,
- x^t the transpose of $x = (x_1, x_2, ..., x_n) \in E$;

Ref

14. Lakshmi J.V.N., Stochastic Gradient Descent using Linear Regression with Python. International Journal of Advanced Engineering Research and Applications, 2(8): 519–524, 2016.

Global Journal of Science Frontier Research (F) Volume XIX Issue IV Version I Year 2019

- $||x||_2 = \sqrt{x^t x} = \sqrt{(x_1^2 + ... + x_n^2)}$ is the Euclidean norm of x;
- $< x, y >= x^t y$ is the inner product and corresponding norm $\|x\|_2$.

To describe the MLR slightly more formally, given a training set, our objective is to learn a function $h : E \longrightarrow \Re$ called a hypothesis, so that $h(x)$ is an optimal predictor for the corresponding value of y.

We consider the following basic problem of convex optimization

$$\min_{\mathbf{x} \in \mathbf{E}} \mathbf{f}(\mathbf{x}), \tag{1}$$

where $\mathbf{f} : E \longrightarrow \Re$ is a smooth convex function of type $C^{1,1}$, i.e., continuously differentiable with Lipschitz continuous gradient L:

$$\|\nabla \mathbf{f}(\mathbf{x}) - \nabla \mathbf{f}(\mathbf{y})\| \leq L\|\mathbf{x} - \mathbf{y}\| \quad \forall \mathbf{x}, \mathbf{y} \in E,$$

and $\nabla \mathbf{f}(\mathbf{x})$ is the gradient of $\mathbf{f}(\mathbf{x})$.

It is assumed that solution $\mathbf{x}^*$ of (1) exists.

Typically computational algorithms for (1) rely on the use of gradient oracles which provide at arbitrary point $\mathbf{x}$ the value of objective function $\mathbf{f}(\mathbf{x})$ and some gradient $\mathbf{g}$ of $\mathbf{f}(\mathbf{x})$, then $\mathbf{g} = \nabla \mathbf{f}(\mathbf{x})$.

III. Linear Regression

Regression analysis is a statistical technique for estimating the relationship among variables which have reason and result relation. Main focus of univariate regression is analyses the relationship between a dependent variables $X_1, \ldots, X_n$ and one independent variable Y and formulates the linear relation equation between dependent and independent variable.

a) Simple linear regression

The simple linear regression model is the simplest regression model in which we have only one predictor X. This model, which is common in practice, is written as

$$Y_i = b + aX_i + \epsilon_i, \qquad i = 1, \ldots, n, \tag{2}$$

where

- Y_i, X_i are the values of the response and predictor variables in the trial, respectively;
- the unknown parameters: a is called the intercept, and b is the slope of the line;
- ϵ_i are usually assumed to be iid (error) from $N(0, \sigma_\epsilon^2)$ specially for inference purposes (see for instance [18]).

Considering n pairs $(X_1, Y_1), \ldots, (X_n, Y_n)$ of observations of independent and dependent variables. The ordinary least squares (OLS) method used for the parameter estimation minimizes the sum of squared errors $\sum \epsilon_i^2$ for the sample.

The least squares criterion

$$\mathbf{f}(\mathbf{b}, \mathbf{a}) = \|Y - (b + aX)\|_2^2$$

Solving the first optimality condition $\nabla \mathbf{f}(\mathbf{b}, \mathbf{a}) = 0$

Ref

18. A.A. Smadi and N.H. Abu-Afouna. On least squares estimation in a simple linear regression model with periodically correlated errors: A cautionary note. Austrian Journal of Statistics, v(41), n(3), 211-226, 2012.

which gives

$$\hat{b} = \bar{Y} - \hat{a}\bar{X}$$

and

$$\hat{a} = \frac{\sum_i (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_i (X_i - \bar{X})^2}.$$

Notes

b) *Multiple linear regression*

To perform MLR, we must decide how we're going to represent hypotheses h in a computer. Equation (3) displays the probabilistic linear equation of multiple regression.

$$Y = \alpha_0 + \alpha_1 X^1 + \cdots + \alpha_p X^p + \epsilon \tag{3}$$

where

- p is the number of independent variables,
- Y is the dependent variable,
- X^i are the independent variables,
- α_i are the parameters,
- ϵ is the error or residual term.

Subsequently, real numbers a_i is used in place of α_i and the deterministic prediction equation becomes as follows.

$$h(X) = Y = a_0 + a_1 X^1 + \cdots + a_p X^p.$$

To simplify our notation, we also introduce the convention of letting the intercept term $X^0 = 1$, so that

$$h(X) = \sum_{i=0}^{p} a_i X^i = a^t X,$$

The parameters $a_0, a_1, \ldots, a_p$ are evaluated by minimizing the cost function $\mathbf{f}(.)$ defined as a sum of squared errors.

The least squares criterion

$$\mathbf{f}(\mathbf{a}) = \|\mathbf{Y} - \mathbf{a}^{\mathbf{t}}\mathbf{X}\|_{\mathbf{2}}^{\mathbf{2}}.$$

c) *Normal equation*

The linear model can be written as

$$Y = X\alpha + \epsilon$$

where

- $Y \in \Re^n$: the vector of observations on the dependent variable,

$$Y = (Y_1, \ldots, Y_n)^T.$$

- $X \in \Re^n \times \Re^{p+1}$: the matrix consisting of a column of ones and p column vectors of the observations on the independent variables,

$$M = \begin{pmatrix} 1 & X_{11} & X_{12} & X_{13} & \dots & X_{1p} \\ 1 & X_{21} & X_{22} & X_{23} & \dots & X_{2p} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & X_{n1} & X_{n2} & X_{n3} & \dots & X_{np} \end{pmatrix}.$$

- $\alpha \in \Re^{p+1}$: the vector of parameters to be estimated,

$$\alpha = (\alpha_0, \alpha_1, \dots, \alpha_p)^T.$$

- $\epsilon \in \Re^n$: the vector of random errors,

$$\epsilon = (\epsilon_1, \epsilon_2, \dots, \epsilon_n)^T.$$

The vector α is a vector of unknown constants to be estimated from the data by $\hat{\alpha}$.

The normal equations [19] are written as

$$X^T X \hat{\alpha} = X^T Y.$$

If $X^T X$ has an inverse, then the unique solution of normal equations given by

$$\hat{\alpha} = (X^T X)^{-1}(X^T Y).$$

The vector $\hat{Y}$ of estimated means of the dependent variable Y for the values of the independent variables X_1, X_2, ..., X_n in the dataset is computed as

$$\hat{Y} = X\hat{\alpha}.$$

However, to express $\hat{Y}$ as a linear function of Y. Thus,

$$\hat{Y} = [X(X^T X)^{-1} X^T]Y.$$

IV. Gradient Methods

a) Gradient descent (GD)

The MLR algorithm use an optimization method to evaluate global optimum. The optimization method used in the MLR algorithm was GD algorithm, represented by equation (4), a well referenced artificial intelligence function to model a first order optimization algorithm that helps us to find a local minimum.

GD is an algorithm that is used to minimize a function, in this case, the cost function. The aim was to find a value of a which renders the lowest error and enables the cost function to reach a local minimum. Each iteration in this method aims to find a new value of a that yields a slightly lower error than the previous iteration. A learning rate is also used to control how large of a step we take downhill during each iteration.

Since GD algorithm starts with some initial a, and repeatedly performs the update:

$$a_i = a_i - \eta \frac{\partial}{\partial a_i} f(a), \qquad i = 0, \dots, n. \tag{4}$$

Here, η is a constant learning rate. This is a very natural algorithm that repeatedly takes a step in the direction of steepest decrease of f.

Since differentiation and combination are interchangeable we can calculate the gradient as a sum of discrete components thus avoiding unnecessary complications of analytical formulas for MLR.

Ref

19. J.O. Rawlings, S.G. Pantula and D.A. Dickey. Applied regression analysis: A research tool. Second Edition, Springer, 1998.

In order to implement this algorithm, we have to work out what is the partial derivative term on the right hand side. Let's first work it out for the case of if we have only one training example (x, y), so that we can neglect the sum in the definition of f. We have:

$$\begin{aligned} \frac{\partial}{\partial a_i} f(a) &= \frac{\partial}{\partial a_i} \frac{1}{2}(a^t X - Y)^2 \\ &= 2.\frac{1}{2}(a^t X - Y).\frac{\partial}{\partial a_i}(a^t X - Y) \\ &= (a^t X - Y).\frac{\partial}{\partial a_i}(\textstyle\sum_{j=0}^{p} a_j X^j) \\ &= (a^t X - Y)X^i \end{aligned}$$

For a single training example, this gives the training rule:

$$a_i = a_i - \eta(a^t X - Y)X^i.$$

To find a local minimum of a function using gradient descent, one takes steps proportional to the negative of the gradient (or of the approximate gradient) of the function at the current point. If instead one takes steps proportional to the positive of the gradient, one approaches a local maximum of that function; the procedure is then known as gradient ascent. In our case, we are looking for minimum value of the cost function, represented by Equation (5).

$$y(a) = f(a_0, a_1, \ldots, a_p) \tag{5}$$

Note that the value of the step size η is allowed to change at every iteration of the MLR algorithm (Equation 2), known as learning rate, of the optimization process. With certain assumptions on the function y and particular choices of η, convergence to a local minimum can be guaranteed. When the function y is convex, all local minimum are also global minimum, so in this case gradient descent can converge to the global solution.

b) Control gradient algorithms

The simplest algorithms for (1) are gradient methods of the kind

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \eta_k \mathbf{g}_k, \quad \mathbf{g}^k = \nabla \mathbf{f}(\mathbf{x}^k), \qquad k = 0, 1, ... \tag{6}$$

which were under intensive study since 1960's. It was shown that (6) converges under very mild conditions for learning rate η_k satisfying "divergence series" condition $\sum_k \eta_k = \infty, \quad \eta_k \to +0$.

Let us suppose a learning rate η_{k-1} was used at iteration $k-1$ and let us identify at x^k the relation between η_{k-1} and the learning rate providing minimization of f in the direction $-g^{k-1}$. To this end Uryasev used in [23] scalar products

$$u_k =< g^k, g^{k-1} >$$

It was heuristically suggested in [23] to correct learning rate according to the formula

$$\eta_{k+1} = \begin{cases} \eta_{decr}\eta_k, & \text{if} \qquad \mathbf{u}_{k+1} \leq 0, \\ \eta_{incr}\eta_k, & \text{otherwise,} \end{cases} \tag{7}$$

where:

- η_{decr} is a decreasing coefficient,
- η_{incr} is a increasing coefficient,
- $0 < \eta_{decr} < 1 < \eta_{incr}$ and $\eta_{incr}.\eta_{decr} < 1$.

Ref

23. Uryas'ev S.P., New Variable-Metric Algorithms for Nondifferentiable Optimization Problems. Journal of optimization theory and applications: 71(2): 359–388, 1991.

Remark: The control learning rate in this section is equivalent to the second conditions of Wolfe:

$$< g^k, g^{k-1} >\geq c\|g^k\|^2$$

where c is real positive.

c) Fast control Gradient Algorithm (FCGA)

Notes

Gradient and Nesterov step

1. Select a point $\mathbf{y}_0 \in \mathbf{E}$. Put

$$k = 0, \quad \mathbf{b}_0 = 1, \quad \mathbf{x}^{-1} = \mathbf{y}_0.$$

2. kth iteration.
   a) Compute $\mathbf{g_k} = \nabla \mathbf{f}(\mathbf{y}_k)$
   b) Put

$$\begin{aligned} \mathbf{x}^k &= \mathbf{y}_k - \eta_k \mathbf{g}_k, \\ \mathbf{b}_{k+1} &= 0.5(1+\sqrt{4\mathbf{b}_k^2+1}\ ), \\ \mathbf{y}_{k+1} &= \mathbf{x}^k + (\tfrac{\mathbf{b}_k - 1}{\mathbf{b}_{k+1}})(\mathbf{x}^k - \mathbf{x}^{k-1}). \end{aligned} \tag{8}$$

The recalculation of the point $\mathbf{y}_k$ in(8) is done using a “ravine” step, and η_k is the control learning rate (7).

Remark: We assume that

$$\eta_k \leq \frac{1}{L} \tag{9}$$

i. *Convergence rate of FCGA*

Lemma 1

For any $x, y \in \Re^n$, we have

$$f(y) \leq f(x) + < \nabla f(x), y - x > + \frac{L}{2}\|y - x\|^2.$$

Proof 1:

For all $x, y \in \Re^n$, we have

$$\begin{aligned} f(y) = &\quad f(x) + \int_0^1 < \nabla f(x + t(y-x)), y - x > dt \\ = &\quad f(x) + < \nabla f(x), y - x > + \int_0^1 < \nabla f(x + t(y-x)) - \nabla f(x), y - x > dt. \end{aligned}$$

Therefore

$$\begin{aligned} f(y) - f(x) - < \nabla f(x), y - x > &\leq |f(y)_f(x) - < \nabla f(x), y - x > | \\ &= |\int_0^1 < \nabla f(x + t(y-x)) - \nabla f(x), y - x > dt| \\ &\leq \int_0^1 | < \nabla f(x + t(y-x)) - \nabla f(x), y - x > |dt \\ &\leq \int_0^1 \| < \nabla f(x + t(y-x)) - \nabla f(x)\|.\|y - x\|dt \\ &\leq \int_0^1 tL\|y - x\|^2 dt \end{aligned}$$

$$= \frac{L}{2}\|y-x\|^2.$$

Then lemma 1 hold.

Lemma 2

The sequence $(b_k)_{k\geq 1}$ generated by the scheme (8) satisfies the following:

$$\frac{k+1}{2} \leq b_k \quad \forall k \geq 1. \tag{10}$$

Notes

Proof 2:

We have

$$0.5(1+\sqrt{4b_{k-1}^2+1}) \geq 0.5(1+2b_{k-1}^2)$$

then

$$b_k \geq \frac{1}{2} + b_{k-1}$$

Therefore, we have $b_k \geq \frac{k+1}{2}$.

Lemma 3

Let $\{x^k, y_k\}$ be the sequence generated by the FCGA, then for any $x \in \Re^n$ we have

$$f(x) \geq f(x^k) + \frac{L}{2}\|x-x^k\|^2 - \frac{L}{2}\|x-y^k\|^2. \tag{11}$$

Proof 3:

By the convexity of f we have

$$f(x) \geq f(y_k) + < \nabla f(y_k), x-y_k > . \tag{12}$$

By lemma 1

$$f(x^k) \leq f(y_k) + < \nabla f(y_k), x^k-y_k > + \frac{L}{2}\|x^k-y_k\|^2 \tag{13}$$

Combining (12) and (13), we have

$$f(x) \geq f(x^k) + < \nabla f(y_k), x-x^k > - \frac{L}{2}\|y_k-x^k\|^2. \tag{14}$$

Since $x^k = y_k - \eta_k \nabla f(y_k)$, then

$$f(x) \geq f(x^k) + \frac{1}{\eta_k} < y_k-x^k, x-x^k > - \frac{L}{2}\|y_k-x^k\|^2. \tag{15}$$

Using (9), we obtain

$$f(x) \geq f(x^k) + L < y_k-x^k, x-x^k > - \frac{L}{2}\|y_k-x^k\|^2. \tag{16}$$

Since

$$< y_k-x^k, x-x^k > = \frac{1}{2}(\|y_k-x^k\|^2 + \|x-x^k\|^2 - \|x-y_k\|^2)$$

then from (16) the inequality (11) hold.

Theorem 1

If the sequence $\{x_k\}_{k\geq 0}$ is constructed by FCGA, then there exist a constant C such that

$$f(x_k) - f(x^*) \leq C/k^2. \tag{17}$$

Proof 4:

We denote

$$z_k = \frac{1}{b_{k+1}}x^* + (1 - \frac{1}{b_{k+1}})x^k. \tag{18}$$

Notes

We know that $\frac{1}{b_{k+1}} \in (0,1], \quad \forall k \geq 0$, *by the convexity of* f *we have*

$$f(z_k) \leq \frac{1}{b_{k+1}}f(x^*) + (1 - \frac{1}{b_{k+1}})f(x^k). \tag{19}$$

$(19) \times b_{k+1}^2$, *we get*

$$b_{k+1}^2 f(z_k) \leq b_{k+1}f(x^*) + (b_{k+1}^2 - b_{k+1})f(x^k) \tag{20}$$

With $b_{k+1}^2 - b_{k+1} = b_k^2$ *yield*

$$b_{k+1}^2 f(z_k) \leq b_{k+1}^2 f(x^*) - b_k^2 f(x^*) + b_k^2 f(x^k). \tag{21}$$

$(21) + b_{k+1}^2 f(x^{k+1})$, *we get*

$$b_{k+1}^2(f(x^{k+1}) - f(x^*)) - b_k^2(f(x^k) - f(x^*)) \leq b_{k+1}^2(f(x^{k+1}) - f(z_k)). \tag{22}$$

By lemma 3, we have

$$f(x^{k+1}) - f(z_k) \leq \frac{L}{2}\|z_k - y_{k+1}\|^2 - \frac{L}{2}\|z_k - x^{k+1}\|^2. \tag{23}$$

Then

$$b_{k+1}^2(f(x^{k+1}) - f(x^*)) - b_k^2(f(x^k) - f(x^*)) \leq b_{k+1}^2\tfrac{L}{2}(\|z_k - y_{k+1}\|^2 - \|z_k - x^{k+1}\|^2). \tag{24}$$

Let $p_k = x^{k-1} + b_k(x^k - x^{k-1})$, *then*

$$\begin{aligned} z_k - y_{k+1} &= \tfrac{1}{b_{k+1}}x^* + (1 - \tfrac{1}{b_{k+1}})x^k - x^k + (\tfrac{b_k - 1}{b_{k+1}})(x^k - x^{k-1}) \\ &= \tfrac{1}{b_{k+1}}(x^* - x^{k-1} + b_k(x^{k-1} - x^k)) \\ &= \tfrac{1}{b_{k+1}}(x^* - p_k)) \end{aligned} \tag{25}$$

and

$$\begin{aligned} z_k - x^{k+1} &= \tfrac{1}{b_{k+1}}x^* + (1 - \tfrac{1}{b_{k+1}})x^k - x^{k+1} \\ &= \tfrac{1}{b_{k+1}}(x^* - x^k + b_{k+1}(x^k - x^{k+1})) \\ &= \tfrac{1}{b_{k+1}}(x^* - p_{k+1}). \end{aligned} \tag{26}$$

Using equations (25) and (26) in (24), we have

$$b_{k+1}^2(f(x^{k+1}) - f(x^*)) - b_k^2(f(x^k) - f(x^*)) \leq \tfrac{L}{2}\|x^* - p_k\|^2 - \tfrac{L}{2}\|x^* - p_{k+1}\|^2.$$

Since $b_0 = 0$ *and* $p_0 = x^{-1} = x^0$*, then we get*

$$b_k^2(f(x^k) - f(x^*)) \leq \frac{L}{2}\|x^* - x^0\|^2.$$

Let $C = 2L\|x^* - x^0\|^2$*, using lemma 2 then theorem hold.*

V. Computational Experiment

All the experiments were carried out on a personal computer with an HP i3 CPU processor 1.80GHz, 4 Go RAM, x64, using Python 3.7 for Windows 8.1.

a) Simple linear regression

In this section, The dataset contains 1,599 samples of red wine [4], with each sample receiving an expert's quality score.

The input variables are Fixed Acidity, Volatile Acidity, Citric Acid, Residual Sugar, Chlorides, Free Sulfur Dioxide, Total Sulfur Dioxide, Density, see Figure 1.

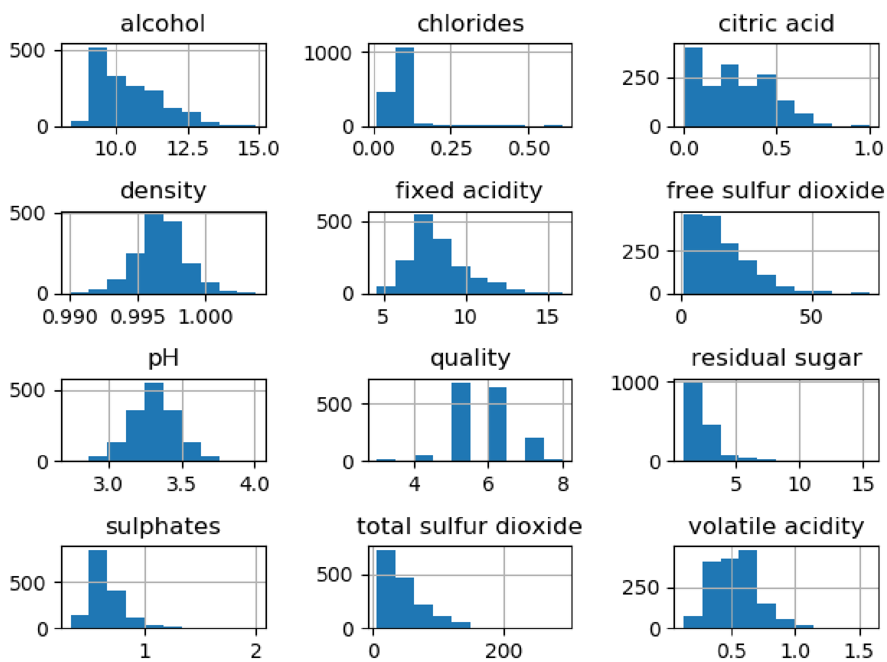


Figure 1: Histograms of the input data

The output variable is Quality, see Figure 2.

Simple Linear Regression is a very simple approach for supervised learning. It is used to make a prediction for the quantitative response, Y, from the the predictor variable, X. In this model, the target, Y, was the wine quality and the data inputs, X, were all other features of wine in the data file. For this model, the data was split into train and test parts; test fractions between 0.20 and 0.25

Ref

4. P. Cortez, A. Cerdeira, F. Almeida, T. Matos and J. Reis., Modeling wine preferences by data mining from physicochemical properties. Decision Support Systems, 47(4):547-553, 2009.

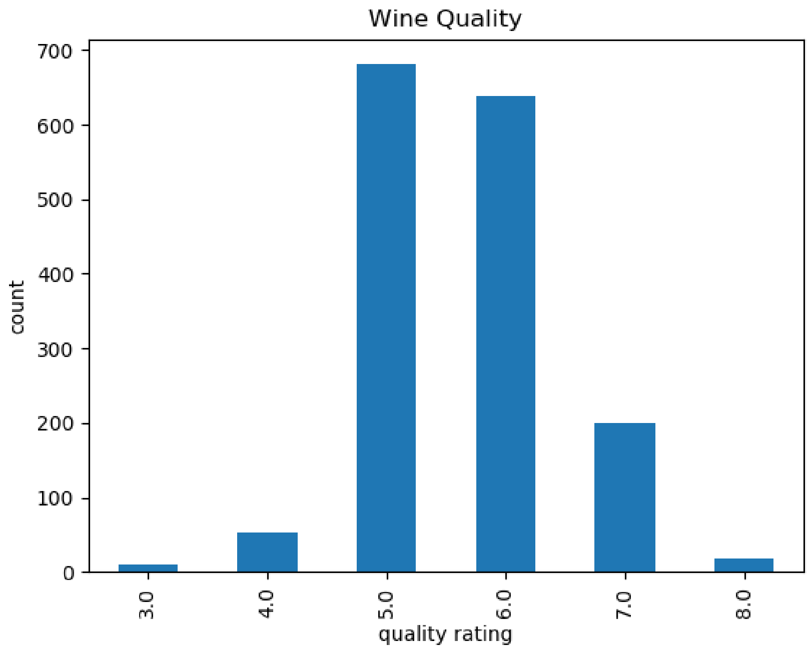


Figure 2: Histograms of the output data

were used. A linear model was then fit to the training part using the least squares approach and subsequently evaluated on both training and test data. Maximum likelihood weights were used, which gave the best linear fit between the processed inputs and the targets, see Figure 3.

The line in represents a simple linear regression of each predictor on the dependent variable. Sulphates and alcohol have a greater impact on quality than residual sugar and appeared to be the most influential.

To further explore simple linear regression, the GD algorithm and FCGA was considered to perform linear regression on our dataset. This model used the normalized inputs for optimization and 100 iterations were run to find the local minimum, we start by GD algorithm code with a small learning rate of 0.1 was used for a slow descent thereby increasing the accuracy of this approach. We used the minimal value of error fined by GD for stopping FCGA code.

A graph of the error function against the iterations was obtained and is depicted in Figure 4.

It can be seen in Figure 4 that after a certain point the error function reaches the local minimum. The value of theta that allowed the error function to reach this point was then used to create a predict function. The input variables from the data set were then used to compute the quality and the results obtained showed that this method rendered a test error of 0.646.

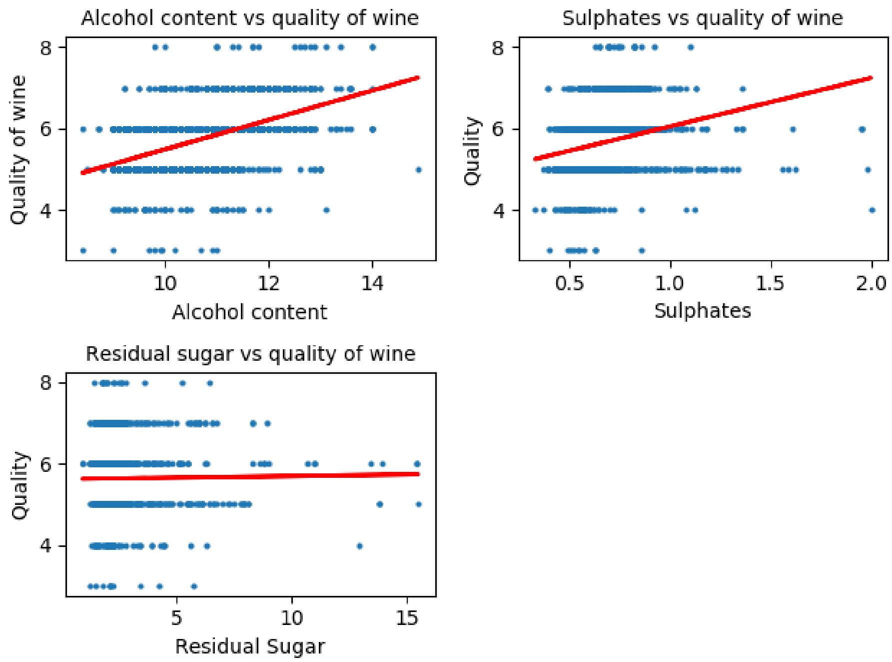


Figure 3: Examples of regression lines

b) Multiple linear regression

We terminate at iteration k if the relative norm of the difference between two consecutive estimates falls below **tolerance** $\approx 10^{-4}$

$$\frac{\|\mathbf{x}^k - \mathbf{x}^{k-1}\|}{\|\mathbf{x}^{k-1}\|} \leq \textbf{tolerance}.$$

In this section, we used the dataset CCPP [21] has four input variables and a target variable (5 attributes) of which 9568 data points excited were collected over a 6 years period $(2006 - 2011) \approx 72$ months ≈ 674 different days, from a combined cycle power plant, when the power plant was in working condition full load. Features consist of medium peripheral variables per hour ambient pressure (AP), relative humidity (RH), temperature (T), exhaust Vacuum (V) to predict the net hourly electrical energy output (EP) of the plant.

Table 1 we provides simple statistics of the dataset.

The correlation matrices can be found in figure 1

1. Power has a strong negative correlation with 'Ambient Temperature' and 'Exhaust Vacuum' of the Plant. So, it seems that as Temperature or Vacuum increases the Power output of the Plant decreases.

2. Similarly, Power has a positive correlation with Pressure and Humidity. It seems that on increasing Humidity and Pressure, the Power output will increase.

3. Vacuum has a strong positive correlation with Temperature of the plant and also Humidity has a small positive correlation with Pressure.

Ref

21. Tufekci P., Prediction of full load electrical power output of a base load operated combined cycle power plant using machine learning methods. Electrical Power and Energy Systems 60, 126140, 2014.

Notes

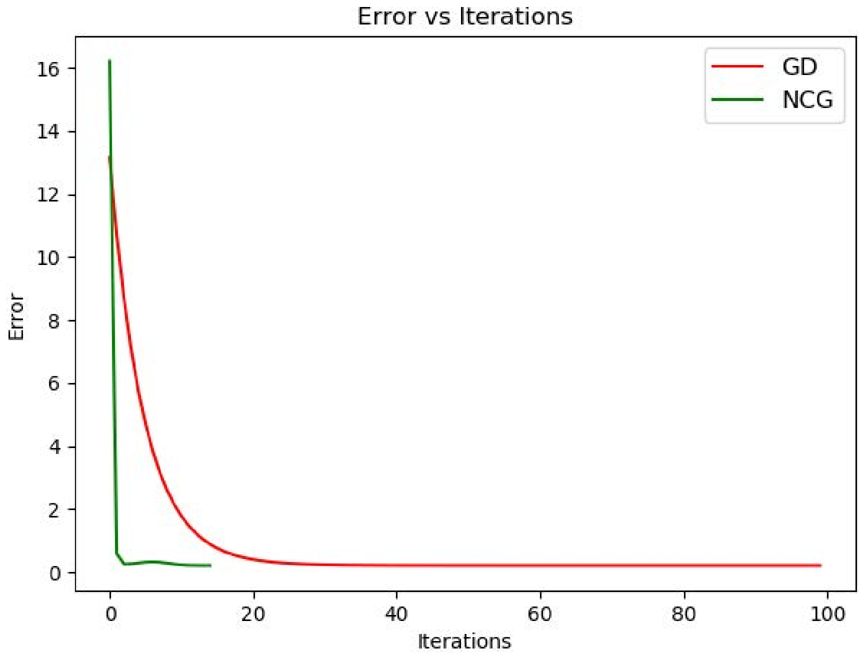


Figure 4: Plot of the error function against the number of iterations by performing linear regression using GD and FCGA

4. All the other non Power linkages have a negative correlation.

The split is done over a random shuffle of the rows of the dataset. Assumes each row in the data has the expected outcome in the last column.

The proportion (0.8) of data to keep in the training set. The rest (0.2) goes to the test set. Then 7654 cases were used for training models and the rest 1914 cases were testing data.

By using GD algorithm we found the following results:

The minimum $x^* = (454.391356, -14.7051025, -2.98027489, 0.402790279, -2.29464039)^t$

Training set $MSE = 20.913375142514127$

Training set $R^2 = 0.9283026158731529$

Test set $MSE = 20.202614118671825$

Test set $R^2 = 0.9302125967625947$

By using normal equation we found the following results:

The solution $x^* = (454.391356, -14.7530271, -2.94477983, 0.393686105, -2.31024216)^t$

Test set $MSE = 20.202667744391906$

Test set $R^2 = 0.930212411519253$

From figures 7 and 8, we show that the CGD algorithm is faster than GD algorithm,and NCGD algorithm is faster than GD algorithm and NCGD algorithm.

Table 1: Basic statistics of dataset CCPP

	T	V	AP	RH	PE
count	9568.000000	9568.000000	9568.000000	9568.000000	9568.000000
mean	19.651231	54.305804	1013.259078	73.308978	454.365009
std	7.452473	12.707893	5.938784	14.600269	17.066995
min	1.810000	25.360000	992.890000	25.560000	420.260000
25%	13.510000	41.740000	1009.100000	63.327500	439.750000
50%	20.345000	52.080000	1012.940000	74.975000	451.550000
75%	25.720000	66.540000	1017.260000	84.830000	468.430000
max	37.110000	81.560000	1033.300000	100.160000	495.760000

VI. Concluding Remarks

In this paper the gradient algorithm are accelerated by Nestrov learning rate and control step called NCG algorithm for solving minimum loss function of convex machine learning regression such as linear regression.

The convergence rate of NCG algorithm proved that it is quadratic convergence, and the comprising of numerical results for simple linear regression with wine quality dataset and for multiple linear regression with CCPP dataset, show that NCG algorithm is robust and faster than GD algorithm.

We can solving global optimization deep learning by stochastic perturbation [6] of gradient algorithm in next works.

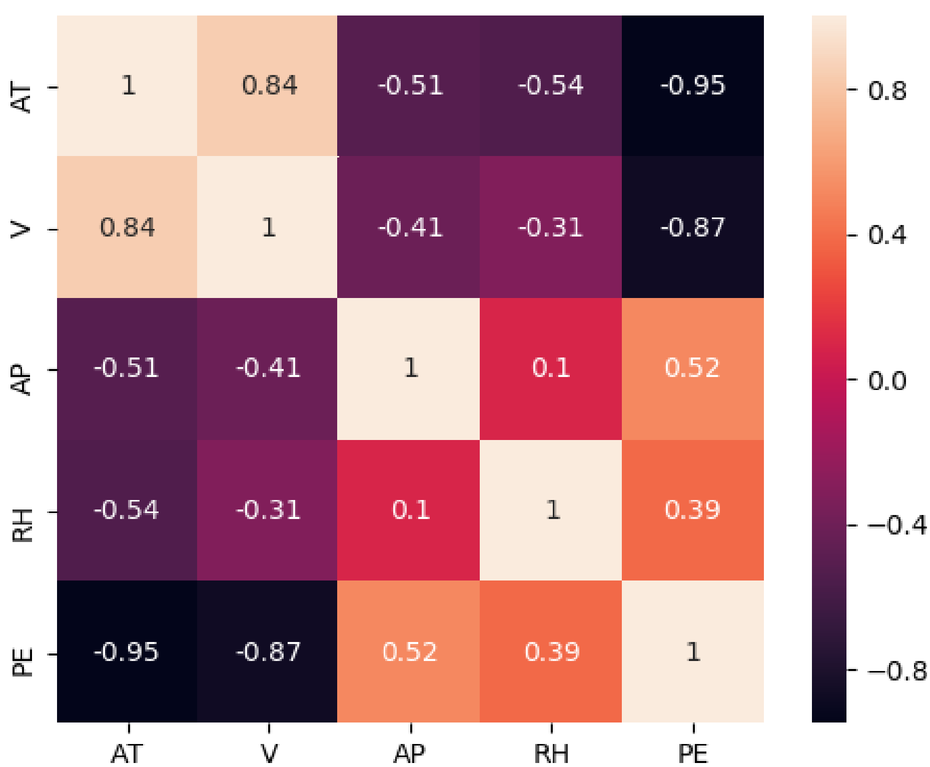


Figure 5: Correlation matrices

Ref

6. A. El Mouatasim, R. Ellaia and J.E. Souza de Cursi. Stochastic perturbation of reduced gradient & GRG methods for nonconvex programming problems. Journal of Applied Mathematics and Computation, 226, 198–211, 2014.

Notes

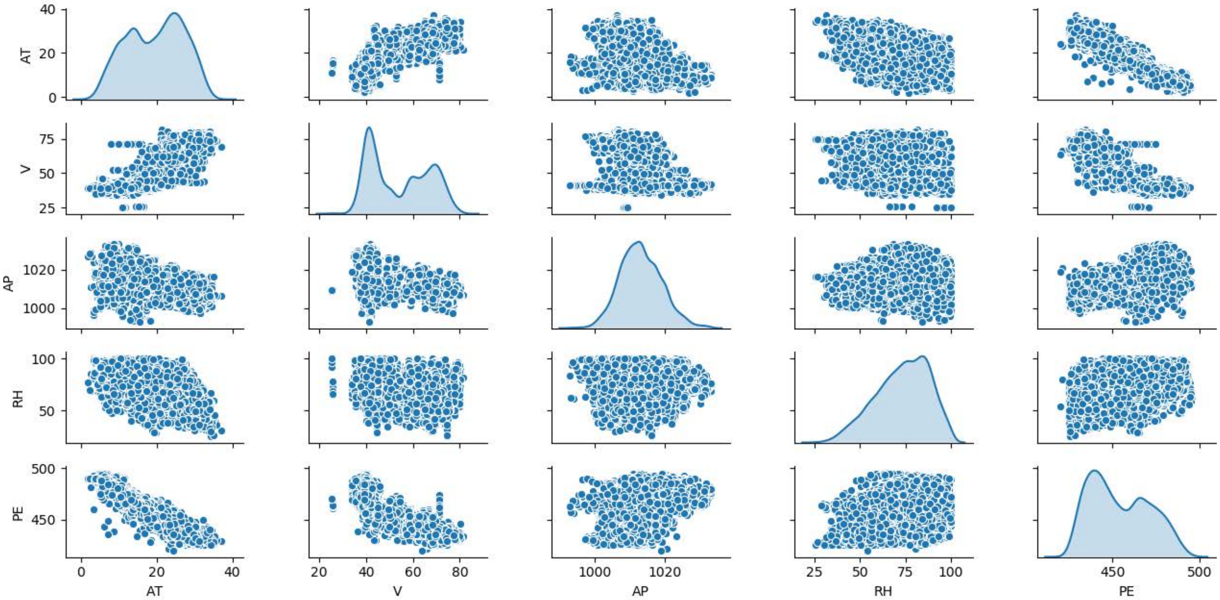


Figure 6: Plotting correlation

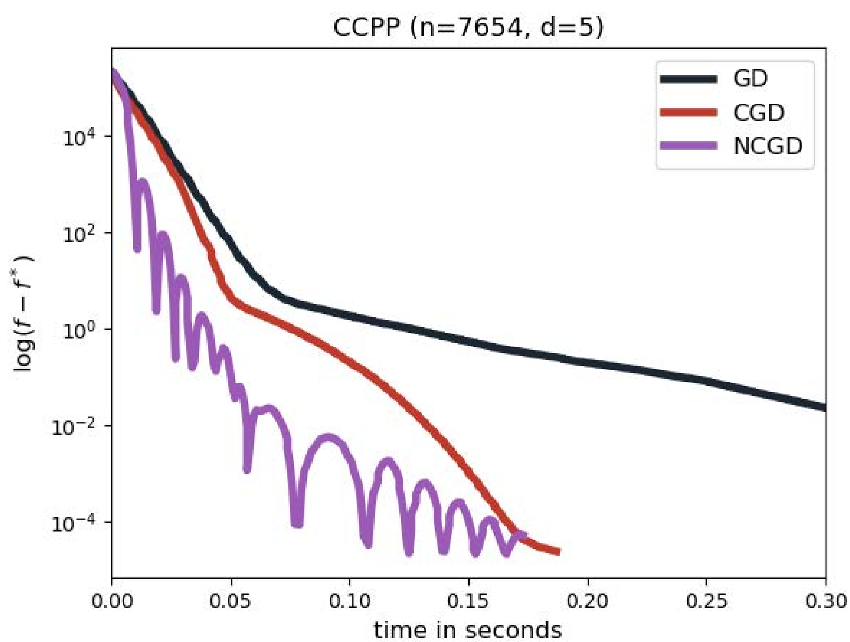


Figure 7: Comparing a time between GD,CGD and NCGD

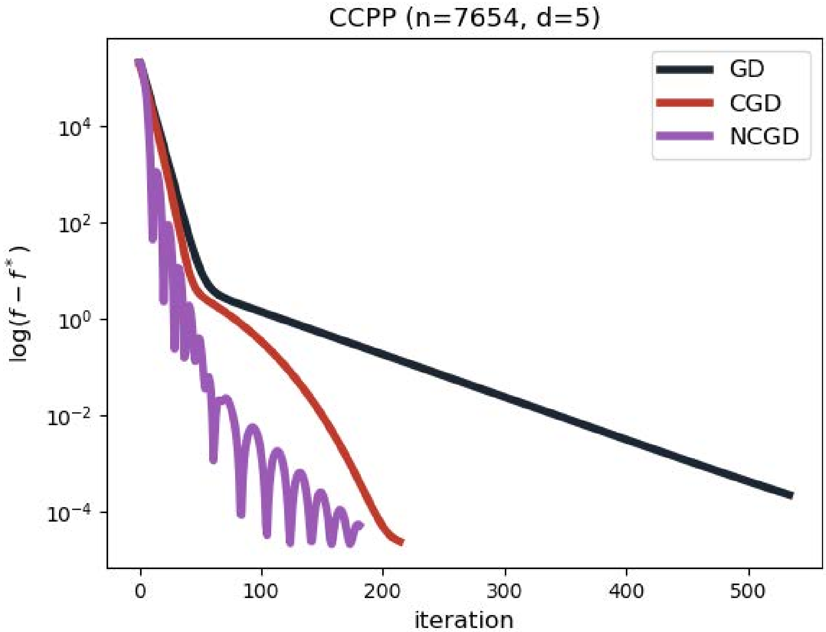


Figure 8: Comparing the iterations between GD,CGD and NCGD

Notes

References Références Referencias

1. Adamowski J. and Karapataki C., Comparison of Multivariate Regression and Arti˙cial Neural Networks for Peak Urban Water-Demand Forecasting: Evaluation of Different ANN Learning Algorithms. Journal of Hydrologic Engineering, 15:729–743, 2010.
2. Chatterjee S. and Simono˙ J.S., Handbook of regression analysis. Wiley, 2013.
3. Chen Z., Zou F., Liao L., Gao S., Zhang M. and Chun J., Prediction of Electrical Energy Output for Combined Cycle Power Plant with Di˙erent Regression Models. AISC 891, pp. 558–566, 2019.
4. P. Cortez, A. Cerdeira, F. Almeida, T. Matos and J. Reis., Modeling wine preferences by data mining from physicochemical properties. Decision Support Systems, 47(4):547-553, 2009.
5. Demirhan A., Performance of machine learning methods in determining the autism spectrum disorder cases. Mugla Journal of Science and Technology, 4(1), 79-84, 2018.
6. A. El Mouatasim, R. Ellaia and J.E. Souza de Cursi. Stochastic perturbation of reduced gradient & GRG methods for nonconvex programming problems. Journal of Applied Mathematics and Computation, 226, 198–211, 2014.
7. El Mouatasim, A. and Wakrim, M., Control Subgradient Algorithm for Image Regularization. Journal of Signal, Image and Video Processing (SIViP), 9, 275–283, 2015.
8. El Mouatasim, A., Simple and Multi Linear Regression Model of Verbs in Quran. American Journal of Computational Mathematics, 8(1): 68–77, 2017.
9. El Mouatasim, A., Regression analysis of a photovoltaic (PV) system in FPO. AIP Conference Proceedings 2056, 020008, 2018.
10. El Mouatasim, A., Control Proximal Gradient Algorithm for ℓ_1 Regularization Image. Journal of Signal, Image and Video Processing (SIViP), to appear, 2019.
11. Ever Y.K, Dimililer K. and Sekeroglu B., Comparison of Machine Learning Techniques for Prediction Problems. AISC 927, pp. 713-723, 2019.
12. Freund R.J., Wilson W.J. and Sa P., Regression analysis: statistical modeling of a response variable. Elsevier, 2006.
13. Huang K., Hussain A., Wang Q. and Zhang R., Deep learning: fundamentals, theory and applications. Springer, 2019.

Notes

14. Lakshmi J.V.N., Stochastic Gradient Descent using Linear Regression with Python. International Journal of Advanced Engineering Research and Applications, 2(8): 519–524, 2016.
15. Nesterov, Y.: Introduction lectures on convex optimization: A basic course. Vol. 87. Springer (2004).
16. Shi B. and Lyengar S.S., Mathematical Theories of Machine Learning: Theory and Applications. Springer, 2020.
17. Shinde A. and Raut P., Comparative Study of Regression Models and Deep Learning Models for Insurance Cost Prediction. AISC 940, pp. 11021111, 2020.
18. A.A. Smadi and N.H. Abu-Afouna. On least squares estimation in a simple linear regression model with periodically correlated errors: A cautionary note. Austrian Journal of Statistics, v(41), n(3), 211-226, 2012.
19. J.O. Rawlings, S.G. Pantula and D.A. Dickey. Applied regression analysis: A research tool. Second Edition, Springer, 1998.
20. Thabtah F., Machine learning in autistic spectrum disorder behavioral research: A review and ways forward. Informatics for health and social care,to appear, 2017.
21. Tufekci P., Prediction of full load electrical power output of a base load operated combined cycle power plant using machine learning methods. Electrical Power and Energy Systems 60, 126140, 2014.
22. Turkan Y.S., Aydogmus H.Y., and Erdal H., The prediction of the wind speed at different heights by machine learning methods. An International Journal of Optimization and Control: Theories & Applications, 6(2), pp.179–187, 2016.
23. Uryas'ev S.P., New Variable-Metric Algorithms for Nondifferentiable Optimization Problems. Journal of optimization theory and applications: 71(2): 359–388, 1991.
24. Xinhua L. and Qian Y., Face Recognition based on Deep Neural Network. International Journal of Signal Processing, Image Processing and Pattern Recognition Vol.8, No.10, pp:29-38, 2015.

Notes

This page is intentionally left blank

Global Journal of Science Frontier Research: F
Mathematics and Decision Sciences
Volume 19 Issue 4 Version 1.0 Year 2019
Type : Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals
Online ISSN: 2249-4626 & Print ISSN: 0975-5896

Motion of Elastic Particles and Spectrum of Hydrogen Atoms

By Zhong-Cheng Liang
Nanjing University of Posts and Telecommunications

Abstract- This article analyzes the spectral structure of hydrogen atoms according to the motion theory of elastic particles. The results demonstrate that optical radiation originates from the elastic vibration of atoms or molecules. The quantum state is the equilibrium feature of the classically statistical system, and the quantum transition is the process of conversion between different motion modes.

GJSFR-F Classification: MSC 2010: 78A10

MOTIONOFELASTICPARTICLESANDSPECTRUMOFHYDROGENATOMS

Strictly as per the compliance and regulations of:

Ref

1. G. Ludwig, *An Axiomatic Basis for Quantum Mechanics (Vol.2) Quantum Mechanics and Macrosystems* (Springer, Berlin, 1987). ISBN 978-3-642-71899-1.

Motion of Elastic Particles and Spectrum of Hydrogen Atoms

Zhong-Cheng Liang

Abstract- This article analyzes the spectral structure of hydrogen atoms according to the motion theory of elastic particles. The results demonstrate that optical radiation originates from the elastic vibration of atoms or molecules. The quantum state is the equilibrium feature of the classically statistical system, and the quantum transition is the process of conversion between different motion modes.

I. Introduction

Atomic spectra are discrete line spectra, which are information for understanding the composition of matters. In 1913, Niels Bohr proposed the hypothesis of the quantum state (stationary state) to explain the atomic stability and the hydrogen spectrum [1]. The success of Bohr's atom theory promoted the establishment of quantum mechanics. The hypothesis of the quantum state is the foundation of quantum mechanics. Although quantum mechanics has made significant achievements, the interpretation of the quantum state has always been controversial [1-3]. The essence of quantum has become the ultimate mystery of nature. Four years ago, the author put forward a physical theory based on the elastic particle model [4], made a unique interpretation to the nature of quantum by classical mechanics [4-6], and developed the theories of particle fields [4,7,8], motion states [4,5], and statistical thermodynamics [4,9]. This article analyzes the spectral structure of hydrogen atoms based on the motion state theory, and further proves the classical nature of the quantum state.

II. Theory of Elastic Particles

a) Elastic particle

Objects are particle systems composed of elastic particles [4-9]. Elastic particles are three-dimensional objects that have mass and volume, that can spin and deform. Electrons, protons, neutrons, atoms, and molecules are all elastic particles. Protons and electrons are indecomposable primary particles, while neutrons are composite particles containing one proton and one electron. The composition of any object is protons and electrons.

The spatial object structure is the nesting of particles at different levels [5,6]. Nucleus and electrons constitute atoms; atoms constitute molecules; molecules constitute supra molecule; and so on. The general model describing the nested structure is

Author: College of Electronic and Optical Engineering, Nanjing University of Posts and Telecommunications Nanjing 210046, China.
e-mail: zcliang@njupt.edu.cn

$$\text{Top-particle} \supseteq \text{meso-particles} \supseteq \text{base-particles} \supseteq \text{sub-particles} \quad (1)$$

A top-particle is the object to be studied, while base-particles are the statistical units of number conservation. An upper-level particle comprises all particles in lower-levels. The more levels of particles an object contains, the more complex the structure of the object is. The interaction of the elastic particles originates from two macroscopic constraints: the conservation of particle number, and the repulsion of particle volume. The elasticity of upper-level particles comes from the motion of lower-level particles. For example, the elasticity of atoms mainly comes from the movement of electrons outside the nucleus.

b) Spatial state

The spatial state of an elastic object includes position, posture, and profile. The position is represented by the position vector of the center-of-mass, the posture is represented by three principal inertia axes, and the profile is represented by three principal rotary inertias [5,6].

Let an object contain N particles, in which the particle P_i has mass M_i, and its position in the laboratory coordinate system $(O-XYZ)$ is

$$\boldsymbol{r}_i = r_s \cdot \tilde{\boldsymbol{r}}_i = \overrightarrow{OP_i} = (x_i, y_i, z_i). \quad (2)$$

Where r_s is the scale of space length, also known as space quantum. The position vector $(\boldsymbol{r}_c)$ of the center-of-mass (P_c) of the object is

$$\boldsymbol{r}_c = \overrightarrow{OP_c} = (x_c, y_c, z_c) = \sum_{i=1}^{N} \left(\frac{M_i}{M}\right) \boldsymbol{r}_i \,, M = \sum_{i=1}^{N} M_i \,. \quad (3)$$

The posture and profile of an object are determined by the inertial matrix. The inertial matrix of the N-particle system in the laboratory coordinate system is [10]

$$\boldsymbol{I} = \begin{pmatrix} I_{11} & -I_{12} & -I_{13} \\ -I_{21} & I_{22} & -I_{23} \\ -I_{31} & -I_{32} & I_{33} \end{pmatrix}; \quad I_s = M_s r_s^2 \,. \quad (4)$$

Where I_s is the scale of inertia. The elements of the inertial matrix are

$$I_{11} = \sum_{i=1}^{N} M_i (y_i^2 + z_i^2), \quad I_{22} = \sum_{i=1}^{N} M_i (z_i^2 + x_i^2), \quad I_{33} = \sum_{i=1}^{N} M_i (x_i^2 + y_i^2),$$

$$I_{12} = I_{21} = \sum_{i=1}^{N} M_i x_i y_i, \quad I_{31} = I_{13} = \sum_{i=1}^{N} M_i z_i x_i, \quad I_{23} = I_{32} = \sum_{i=1}^{N} M_i y_i z_i. \quad (5)$$

Inertial matrix is a real symmetric matrix. It has three real eigenvalues I_1, I_2, I_3 and three eigenvectors $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3$ according to the theory of linear algebra. We describe the profile by the eigenvalues of the inertial matrix, $\boldsymbol{I}_c = (I_1, I_2, I_3)$, and describe the posture by the angles between the eigenvectors and the laboratory coordinate axes, $\boldsymbol{\theta}_c = (\theta_1, \theta_2, \theta_3)$.

c) Motion energy

Motion is a process in which the spatial state of an object changes with time. Changes in position, posture, and profile are called translation, rotation, and vibration,

Ref

5. Z. C. Liang, "Motion, energy and state of body particle system", *Theoretical Physics*, 4, 66-84 (2019). DOI:10.22606/tp.2019.42003

respectively. The translation is the movement of the center-of-mass in the laboratory reference system, the rotation is the spin around the center-of-mass, and the vibration is the extension and contraction relative to the center-of-mass. Translation, rotation, and vibration are three independent motion modes, and each mode has three degrees of freedom. An elastic particle has 3×3 = 9 degrees of freedom, and an object composed of N particles has $9N$ degrees of freedom.

If the three mode energies of a particle are $H_{i\alpha}$, $L_{i\alpha}$, $K_{i\alpha}$ $(\alpha = 1,2,3)$, then the total energies of the three modes are [4-6]

$$H = \sum_{i=1}^{N}\sum_{\alpha=1}^{3} H_{i\alpha}\ ;\ \ H_{i\alpha} = \frac{V_i\chi_{i\alpha}^2}{2Y_{i\alpha}},$$
$$L = \sum_{i=1}^{N}\sum_{\alpha=1}^{3} L_{i\alpha}\ ;\ \ L_{i\alpha} = \frac{s_{i\alpha}^2}{2I_{i\alpha}}, \tag{6}$$
$$K = \sum_{i=1}^{N}\sum_{\alpha=1}^{3} K_{i\alpha}\ ;\ \ K_{i\alpha} = \frac{p_{i\alpha}^2}{2M_i}.$$

The total energy of each mode is the sum of $3N$ independent square terms and has positivity$(H > 0,\ L > 0,\ K > 0)$.

d) Energy space

The energy space is defined as the set of ordered array $\{\mathbb{E}^h, \mathbb{E}^l, \mathbb{E}^k\}$ [4-6]

$$\mathbb{E}^h = \langle H^h, L^h, K^h\rangle,\ \ \mathbb{E}^l = \langle L^l, K^l, H^l\rangle,\ \ \mathbb{E}^k = \langle K^k, H^k, L^k\rangle. \tag{7}$$

Where $x = h, l, k$ is the zone index. $\mathbb{E}^h, \mathbb{E}^l, \mathbb{E}^k$ is called gas zone, solid zone, and liquid zone, respectively.

The energy space is confined to the first octant (+,+,+) of the Cartesian space due to the positivity of motion energy. In the energy space, we defined the energy vector length (called entire energy) as

$$E = \sqrt{H^2 + L^2 + K^2}. \tag{8}$$

Three planes $\{H = K, K = L, L = H\}$ divide the energy space into six phases $\{\mathrm{B}[G_\pm^h], \mathrm{B}[G_\pm^l], \mathrm{B}[G_\pm^k]\}$. $\mathrm{B}[H] = \mathrm{B}[G_+^h] + \mathrm{B}[G_-^h]$ is the gas zone, $\mathrm{B}[L] = \mathrm{B}[G_+^l] + \mathrm{B}[G_-^l]$ is the solid zone, and $\mathrm{B}[K] = \mathrm{B}[G_+^k] + \mathrm{B}[G_-^k]$ is the liquid zone. There are six phase interfaces in the energy space, of which the J-type interfaces $\{\mathrm{S}[J_0^h], \mathrm{S}[J_0^l], \mathrm{S}[J_0^k]\}$ are the interfaces of zero potential energy, the G-type interfaces $\{\mathrm{S}[G_0^h], \mathrm{S}[G_0^l], \mathrm{S}[G_0^k]\}$ are the interfaces of zero chemical energy. The structure of the energy space is shown in Figure 1(a).

Ref

4. Z. C. Liang, *Physical Principles of Finite Particle System* (Scientific Research Publishing, Wuhan. 2015). DOI:10.13140/RG.2.1.2409.8004

Global Journal of Science Frontier Research (F) Volume XIX Issue IV Version I Year 2019

Notes

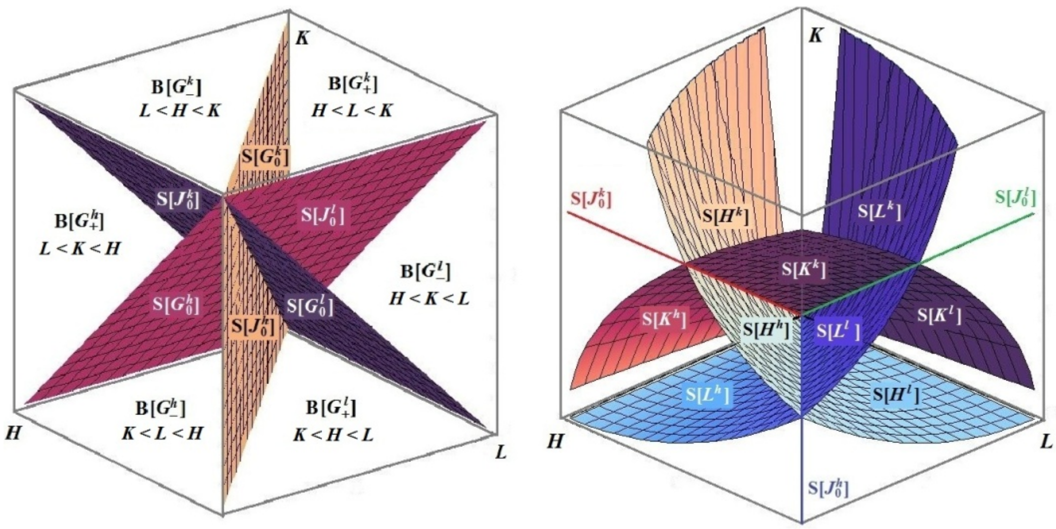


Figure 1: (a) Structure of energy space. (b) Equilibrium surfaces in energy space

e) Equilibrium surface

There are three equilibrium surfaces in the energy space: vibration surface $\mathrm{S}[H]$, rotation surface $\mathrm{S}[L]$, and translation surface $\mathrm{S}[K]$. Their corresponding equilibrium equations are

$$H=\sqrt{2LK},\quad L=\sqrt{2KH},\quad K=\sqrt{2HL}\,. \tag{9}$$

$\mathrm{S}[H]$, $\mathrm{S}[L]$ and $\mathrm{S}[K]$ represent vibration (radiative) equilibrium, rotation (magnetic) equilibrium, and translation (thermal) equilibrium, respectively. Each surface spans four phases and extends to three zones, as shown in Figure 1(b). A matrix describing the structure of the equilibrium surfaces is shown in Table 1. The diagonal elements of the matrix, $\{\mathrm{S}[H^h],\mathrm{S}[L^l],\mathrm{S}[K^k]\}$, are stable areas and the rest are excited areas.

Table 1: Structure of equilibrium surfaces

	$\mathbf{B}[\boldsymbol{H}]$	$\mathbf{B}[\boldsymbol{L}]$	$\mathbf{B}[\boldsymbol{K}]$
$\mathrm{S}[H]$	$\mathrm{S}[H^h]$	$\mathrm{S}[H^l]$	$\mathrm{S}[H^k]$
$\mathrm{S}[L]$	$\mathrm{S}[L^h]$	$\mathrm{S}[L^l]$	$\mathrm{S}[L^k]$
$\mathrm{S}[K]$	$\mathrm{S}[K^h]$	$\mathrm{S}[K^l]$	$\mathrm{S}[K^k]$

In Table 2, we list the energy names and parameter definitions on the three equilibrium surfaces. Where $\{X,Y,Z\}$ is motion energy. $\{E,Q,\ J,G\}$ is auxiliary energy. $\{a,b\}$ is order parameter, which has the relation $2ab=1$. The relations $E=\sqrt{X^2+Y^2+Z^2}=Y+Z$ determine the equilibrium equation $X=\sqrt{2YZ}$.

Table 2: Energy names and parameter definition on equilibrium surfaces

Equilibrium surface	**Definition**	$\mathbf{S}[\boldsymbol{H}]$	$\mathbf{S}[\boldsymbol{L}]$	$\mathbf{S}[\boldsymbol{K}]$
Equilibrium equation	$X=\sqrt{2YZ}$	$H=\sqrt{2l}$	$L=\sqrt{2}$	$K=\sqrt{2HL}$
Major energy	X	H	L	K
Ahead energy	Y	L	K	H

Back energy	Z	K	H	L
Ahead parameter	$a = Y/X$	L/H	K/L	H/K
Back parameter	$b = Z/X$	K/H	H/L	L/K
Entire energy	$E = Y + Z$	$L + K$	$K + H$	$H + L$
Thermal energy	$Q = Z + X$	$K + H$	$H + L$	$L + K$
Potential energy	$J = Y - X$	$L - H$	$K - L$	$H - K$
Chemical energy	$G = Z - Y$	$K - L$	$H - K$	$L - H$
Energy quantum	X_s	$H_s = hv$	$L_s = lz$	$K_s = kT$

III. Quantum State and Transition

a) *Energy quantum*

According to the principle of objectivity [4-9], any physical quantity $\boldsymbol{x}$ can be expressed by the product of scale x_s and digit $\tilde{\boldsymbol{x}}$, namely $\boldsymbol{x} = x_s \cdot \tilde{\boldsymbol{x}}$. The scale is the identifier and metric of the physical quantity, and its essence is quantum. The energy quanta of the *N*-particle system are defined as the statistical average of three-mode energies as

$$H_s = H/N = hv,\ L_s = L/N = lz,\ K_s = K/N = kT. \tag{10}$$

In SI system, v is the vibration intensity (frequency) with unit hertz (Hz), z is the rotation intensity (magnetic induction) with unit tesla (T), and T is the translation intensity (thermodynamic temperature) with unit kelvin (K). Taking the energy unit as joule (J), then, $h = 6.6260693 \times 10^{-34}\,\mathrm{J \cdot Hz^{-1}}$ is Planck constant, $l = 9.2740095 \times 10^{-24}\,\mathrm{J \cdot T^{-1}}$ is Bohr magneton constant, and $k = 1.3806506 \times 10^{-23}\,\mathrm{J \cdot K^{-1}}$ is Boltzmann constant.

b) *Quantum state*

On an equilibrium surface, the digits $\{\tilde{X}, \tilde{Y}, \tilde{Z}\}$ must be integers when using the scale of major energy $\{X_s\}$ [4-6]. The quantum state is defined as the state whose digits are integers on an equilibrium surface. According to this definition, we have the following theorem.

Quantum state theorem. Quantum state $\{\tilde{X}, \tilde{Y}, \tilde{Z}\}$ is the positive integer solution of the algebraic equations $\{X^2 = 2YZ,\ Y^2 = 2ZX,\ Z^2 = 2XY\}$.

The quantum state on equilibrium surface S[X] can be easily determined by algebraic equation $\tilde{X}^2 = 2\tilde{Y}\tilde{Z}$. For example, Table 3 lists the quantum states of $\tilde{Z} = 1$. In the table, the ahead parameter is $a = \tilde{Y}/\tilde{X}$.

Table 3: Quantum state of $\tilde{Z} = 1$ on equilibrium surface S[X]

$\tilde{X}$	2	4	6	8	10	12	14	16	18	20	…
$\tilde{Y}$	2	8	18	32	50	72	98	128	162	200	…
$\tilde{Z}$	1	1	1	1	1	1	1	1	1	1	…
a	1	2	3	4	5	6	7	8	9	10	…

All quantum states on the surface S[X] can be plotted in the YZ plane with Z as the abscissa and Y as the ordinate, as shown in Figure 2. We marked the four straight

Ref

9. Z. C. Liang, "Cluster ensemble statistics of body particle system", *New Horizons in Mathematical Physics* **3**, 53-73 (2019). DOI:10.22606/nhmp.2019.32002

lines of $a = 1,2,3,4$ in the plot for easy identification. The order parameters outside the straight lines are rational numbers that are not integers.

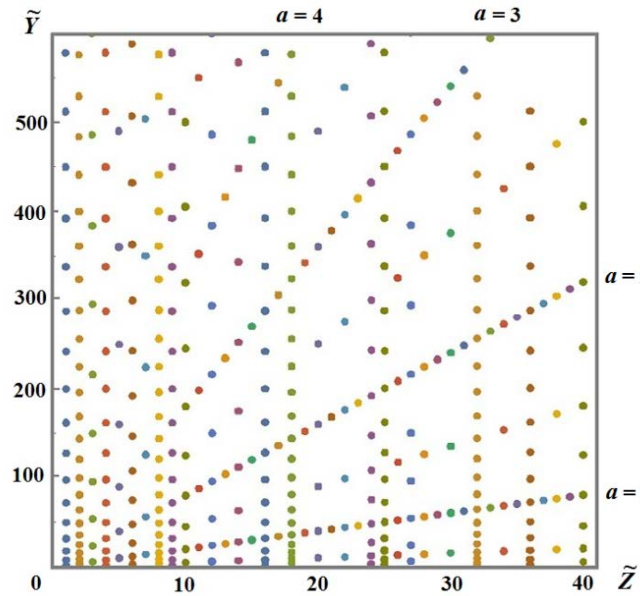


Figure 2: The plot of quantum state on equilibrium surface $S[X]$

c) *Ground and excited state*

Let's consider the quantum state in thermal equilibrium surface $S[K]$. In such case, it has $X_s = kT$; $\tilde{X} = \tilde{K}, \tilde{Y} = \tilde{H}, \tilde{Z} = \tilde{L}$. The quantum state equation is $\tilde{K}^2 = 2\tilde{H}\tilde{L}$, and the ahead parameter is $a = \tilde{H}/\tilde{K}$. Table 4 gives the quantum states of $\tilde{L} = 1\sim10$ on $S[K]$ surface. In the table, $S[K^h]$ represents the excited state of vibration. $L[J_0^k]$ is the intersection of $S[K]$ and $S[H]$, representing a stable state. The first column with $\tilde{L} = 1$ is the ground state.

Table 4: Quantum states of $\tilde{L} = 1\sim10$ with ground state $\tilde{L} = 1$ on $S[K]$ surface

State	$L[J_0^k]$	$S[K^h]$									
$\tilde{K}$	2	8	18	32	50	72	98	128	162	200	…
$\tilde{H}$	2	16	54	128	250	432	686	1024	1458	2000	…
$\tilde{L}$	1	2	3	4	5	6	7	8	9	10	…
a	1	2	3	4	5	6	7	8	9	10	…

d) *Quantum transition*

The quantum transition (or quantum jump) is a process of continuous conversion between quantum states. According to the continuous equilibrium equation $K^2 = 2HL$, the motion energy on $S[K]$ can be expressed by the ahead parameter $a = H/K$ as

$$K(a) = 2aL(a), H(a) = 2a^2L(a), L = L(a)\,. \tag{11}$$

Therefore, we express the rotation energy as

$$L(a) = \frac{H(a)}{2a^2}\,. \tag{12}$$

After the transition from a high-energy state (a') to a low-energy state (a), part of energy is converted to vibration radiation and lost. We express the lost rotation energy by the difference of the vibration frequency as

Notes

Global Journal of Science Frontier Research (F) Volume XIX Issue IV Version I Year 2019

$$v_{a'a} = \frac{L(a) - L(a')}{h} = \frac{H(a)}{2ha^2} - \frac{H(a')}{2ha'^2}. \tag{13}$$

The frequency difference of lost energy is called emission spectrum, which can be decomposed into two terms

$$v_{a'a} = \bar{v}_{a'a} - \Delta v_{a'a};\ \bar{v}_{a'a} = \frac{H(a)}{2h}\left(\frac{1}{a^2} - \frac{1}{a'^2}\right),\ \Delta v_{a'a} = \frac{\Delta H_{a'a}}{2ha'^2}. \tag{14}$$

$\bar{v}_{a'a} > 0$ is the spectral frequency caused by the change of ahead parameter. $\Delta v_{a'a} > 0$ is the spectral line-width caused by the loss of vibration energy, $\Delta H_{a'a} = H(a') - H(a)$. We can obtain the absorption spectrum ($v_{aa'}$) from the emission spectrum ($v_{a'a}$) by switching a and a', and changing the signs of frequency and line-width.

IV. Spectrum of Hydrogen Atoms

a) *Spectral frequency*

According to Table 4, we obtain the first three series of the emission spectrum as

$$\bar{v}_{a'1} = \frac{H(1)}{2h}\left(\frac{1}{1^2} - \frac{1}{a'^2}\right);\quad a' = 2,3,4,\cdots \tag{15a}$$

$$\bar{v}_{a'2} = \frac{H(2)}{2h}\left(\frac{1}{2^2} - \frac{1}{a'^2}\right);\quad a' = 3,4,5,\cdots \tag{15b}$$

$$\bar{v}_{a'3} = \frac{H(3)}{2h}\left(\frac{1}{3^2} - \frac{1}{a'^2}\right);\quad a' = 4,5,6,\cdots \tag{15c}$$

Compared with Rydberg formula of hydrogen atoms [1], spectral frequencies $\bar{v}_{a'1}, \bar{v}_{a'2}, \bar{v}_{a'3}$ are Lyman series, Balmer series, and Paschen series, respectively. Therefore, the emission spectrum with ground state $\tilde{L} = 1$ is the spectrum of hydrogen atoms.

The spectral frequency of hydrogen atoms can be generally expressed as

$$\bar{v}_{a'a} = \frac{H(a)}{2h}\left(\frac{1}{a^2} - \frac{1}{a'^2}\right),\ (a = 1,2,3,\cdots;\ a' = a+1, a+2, a+3,\cdots) \tag{16}$$

According to formula (16), we know that the ahead parameter is the principal quantum number in Bohr's atomic theory. From Table 4, we find that all three quantum numbers $\{\tilde{H}, \tilde{L}, \tilde{K}\}$ are changed after the transition. In other words, all three-mode energies $\{H, L, K\}$are continuously changing during the transition.

b) *Limit frequency*

Substituting $a' \to \infty$ into formula (16) obtains the limit frequency of the a -serial spectrum.

$$\bar{v}_{\infty a} = \frac{H(a)}{2ha^2} = \frac{L(a)}{h} = \frac{\tilde{L}(a)kT}{h} \tag{17}$$

For examples, the limit frequencies of the Lyman series and the Balmer series are $\bar{v}_{\infty 1} = kT/h$ and $\bar{v}_{\infty 2} = 2kT/h$, respectively. According to formula (17), $\bar{v}_{\infty a}$ is a variable that depends on $H(a)$. The limit frequency given by the Rydberg formula of hydrogen is $\bar{v}_{\infty a} = cR_{\mathrm{H}}/a^2$, where $R_{\mathrm{H}} = 10967758\ \mathrm{m}^{-1}$ is Rydberg constant [1].

Ref

1. G. Ludwig, *An Axiomatic Basis for Quantum Mechanics (Vol.2) Quantum Mechanics and Macrosystems* (Springer, Berlin, 1987). ISBN 978-3-642-71899-1.

Compared with (17), R_{H} should be $H(a)/(2hc)$. Therefore, regarding R_{H} as constant is an approximation.

We can express the spectral frequency by the limit frequency as

$$\bar{v}_{a'a} = \bar{v}_{\infty a}\left(1 - \frac{a^2}{a'^2}\right) \tag{18}$$

For example, the spectral frequency of the Lyman series is

$$\bar{v}_{a'1} = v_{\infty 1}\left(1 - \frac{1}{a'^2}\right), \quad a' = 2,3,4,\cdots \tag{19}$$

c) Spectral line-width

We have a differential relation in the continuous transition according to the expression of spectral line-width in formula (14)

$$\frac{d(v_{a'a})}{d(H_{a'a})} = \frac{1}{2ha'^2} \tag{20}$$

Roughly, the line-width is inversely proportional to the square of the ahead parameter.

The quantum transition is a continuous conversion between different motion modes. Because energy conversion takes time, the quantum transition is not instantaneous. The transition time $(\Delta t_{a'a})$ is represented by the reciprocal of spectral line-width $(\Delta v_{a'a})$ as

$$\Delta t_{a'a} = \frac{1}{\Delta v_{a'a}} = \frac{\Delta\lambda_{a'a}}{c}. \tag{21}$$

Where $\Delta\lambda_{a'a}$ is the spectral line-width in terms of wavelength. For example, a spectral line-width $\Delta\lambda_{a'a} = 1\mathrm{nm}$ corresponds to a transition time $\Delta t_{a'a} \approx 3 \times 10^{-18}\mathrm{sec}$.

d) Hydrogen-like spectrum

Table 5 gives the quantum states of $\tilde{L} = 2\sim11$ on $\mathrm{S}[K]$ surface. Where $\mathrm{L}[J_0^l]$ is the intersection of $\mathrm{S}[L]$ and $\mathrm{S}[K]$. $\mathrm{L}[J_0^l]$ and $\mathrm{L}[J_0^k]$ are both stable states. In this case, $\mathrm{L}[J_0^l]$ is ground state as $\tilde{L} = 2$ is the lowest rotation energy, and $\mathrm{L}[J_0^k]$ is meta-stable state as it has higher rotation energy.

Table 5: Quantum states of $\tilde{L} = 2\sim11$ with ground state $\tilde{L} = 2$ on $\mathrm{S}[K]$ surface

State	$\mathrm{L}[J_0^l]$	$\mathrm{L}[J_0^k]$	$\mathrm{S}[K^h]$								
$\tilde{K}$	2	6	12	20	30	42	56	72	90	110	…
$\tilde{H}$	1	6	18	40	75	126	196	288	405	550	…
$\tilde{L}$	2	3	4	5	6	7	8	9	10	11	…
a	1/2	1	3/2	2	5/2	3	7/2	4	9/2	5	…

From Table 5, we obtain the first three series of emission spectrum

$$\bar{v}_{a'\frac{1}{2}} = \frac{H(1/2)}{2h}\left[\frac{1}{(1/2)^2} - \frac{1}{a'^2}\right]; \quad a' = \frac{2}{2},\frac{3}{2},\frac{4}{2},\cdots \tag{22a}$$

$$\bar{v}_{a'\frac{2}{2}} = \frac{H(2/2)}{2h}\left[\frac{1}{(2/2)^2} - \frac{1}{a'^2}\right]; \quad a' = \frac{3}{2},\frac{4}{2},\frac{5}{2},\cdots \tag{22a}$$

$$\bar{v}_{a'\frac{3}{2}} = \frac{H(3/2)}{2h}\left[\frac{1}{(3/2)^2} - \frac{1}{a'^2}\right]; \quad a' = \frac{4}{2},\frac{5}{2},\frac{6}{2},\cdots \tag{22a}$$

Notes

Notes

Unlike the case of hydrogen atoms, the order parameter contains half integers, which is the spectral feature of hydrogen-like ions. Hydrogen-like spectral frequency can be uniformly expressed as

$$\bar{\nu}_{a'a} = \frac{H(a)}{2h}\left(\frac{1}{a^2} - \frac{1}{a'^2}\right), \left(a = \frac{1}{2}, \frac{2}{2}, \frac{3}{2}, \cdots;\ a' = a + \frac{1}{2}, a + \frac{2}{2}, a + \frac{3}{2}, \cdots\right) \qquad (23)$$

V. Conclusion

The elastic particle theory shows that particles have three independent motion modes of translation, rotation, and vibration. The three-mode energies $\{H, L, K\}$ form a Cartesian energy space. The energy quanta $\{K_s = kT,\ L_s = lz,\ H_s = hv\}$ are the statistical averages of $\{H, L, K\}$. There are three equilibrium surfaces in the energy space, respectively representing thermal equilibrium, magnetic equilibrium, and radiative equilibrium. Quantum states $\{\tilde{H}, \tilde{L}, \tilde{K}\}$ are those on the equilibrium surface in which the energy takes integer.

The application of particle motion theory to atomic system shows that atoms are elastic particles and the optical radiation originates from the elastic vibration of atoms or molecules. The transition between quantum states produces a line spectrum. The main characteristics of the optical spectrum, including spectral frequency, limit frequency, spectral line-width, and transition time, are predicted theoretically, which are consistent with the existing observations. The results demonstrate that the quantum state is the equilibrium property of particle statistics, the quantum transition is the conversion process of particle motion modes, and the quantum randomness originates from the classical statistics of elastic particles.

Acknowledgements

This research is partially supported by NSFC (No.61775102 and No.61905117).

References Références Referencias

1. G. Ludwig, *An Axiomatic Basis for Quantum Mechanics (Vol.2) Quantum Mechanics and Macrosystems* (Springer, Berlin, 1987). ISBN 978-3-642-71899-1.
2. H. Edward. *Cosmology: The Science of the Universe* (Cambridge University Press, Cambridge, 2000). ISBN 978-0-521-66148-5
3. Z. K. Minev, S. O. Mundhada, S. Shankar, *et al.* “To catch and reverse a quantum jump mid-flight”, *Nature* (Online), 570(7760), June 2019. DOI: 10.1038/s41586-019-1287-z
4. Z. C. Liang, *Physical Principles of Finite Particle System* (Scientific Research Publishing, Wuhan. 2015). DOI:10.13140/RG.2.1.2409.8004
5. Z. C. Liang, “Motion, energy and state of body particle system”, *Theoretical Physics*, **4**, 66-84 (2019). DOI:10.22606/tp.2019.42003
6. Z. C. Liang, “Modeling of real particles”, 8th International Conference on Mathematical Modeling in Physical Science, August 26-29, 2019, Bratislava, Slovakia.
7. Z. C. Liang, “The origin of gravitation and electromagnetism”, *Theoretical Physics*, **4**, 85-102 (2019). DOI:10.22606/tp.2019.42004

8. Z. C. Liang, "Essence of light: particle, field and interaction", in *Proc. SPIE*. **10755**, 1075501-10755014 (2018). DOI:10.1117/12.2316422.
9. Z. C. Liang, "Cluster ensemble statistics of body particle system", *New Horizons in Mathematical Physics* **3**, 53-73 (2019). DOI:10.22606/nhmp.2019.32002
10. S. N. Jin and Y.L. Ma, *Theoretical Mechanics* (Higher Education Press, Beijing. 2002).

Notes

Global Journal of Science Frontier Research: F
Mathematics and Decision Sciences
Volume 19 Issue 4 Version 1.0 Year 2019
Type : Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals
Online ISSN: 2249-4626 & Print ISSN: 0975-5896

Conjugate Fourier Series of (N, p, q) Summability of Approximation Theory F

By Sanjay Mukherjee & A J Khan
MATS University

Abstract- In this paper the author has obtained a theorem that is:

The monotonic non- increasing sequence of real constant that the conjugate Fourier series is almost (N, p, q) summable to

$$\|\tilde{t}_{n,m} - \tilde{f}\| = o(1)$$

Keywords: nörlund summability, conjugate fourier series, summability means, monotonic non-increasing sequence.

GJSFR-F Classification: *MSC 2010: 42A16*

CONJUGATEFOURIERSERIESOFN PQSUMMABILITYOFAPPROXIMATIONTHEORYF

Strictly as per the compliance and regulations of:

Open Association of Research Society, USA
Research | Diversity | Ethics

Conjugate Fourier Series of (N, p, q) Summability of Approximation Theory F

Sanjay Mukherjee [α] & A J Khan [σ]

Abstract- In this paper the author has obtained a theorem that is:

The monotonic non- increasing sequence of real constant that the conjugate Fourier series is almost (N, p, q) summable to

$$\|\tilde{t}_{n,m} - \tilde{f}\| = o(1)$$

Keywords: nörlund summability, conjugate fourier series, summability means, monotonic non-increasing sequence.

I. Introduction

Lorentz [3], for the first time in 1048, defined almost convergence of a bounded sequence. It is easy to see that a convergent sequence is almost convergent [4]. . Here, almost generalized Nörlund summability of method is considered. In 1913, Hardy [1] established $(c, \alpha), \alpha > 0$ summability of the series.Later on in 1948, harmonic summability which is weaker then summability $(c, \alpha), \alpha > 0$ of the series was discussed by Siddiqi[8]. The generalization of Siddiqi has been obtained by several workers, for example, Singh [9, 10], Iyengar[2], Pati[5]' Tripathi[11], Rajagopal[7] for Norlund mean. In an attempt to make an advance study in this direction, in the present paper, a theorem on almost generalized Nörlund summability of conjugate series of Fourier series has been obtained.

II. Definitions and Notations

Let $\sum a_n$ be an infinite series with $\{S_n\}$ as the sequence of its nth partial sums.
Lorentz [3] has given the following definition.
A bounded sequence$\{S_n\}$ is said to be almost convergent to a limit S, if

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_{v=m}^{n+m} S_v = S, \text{ uniformly with respect to m.} \qquad (2.1)$$

Let $\{p_n\}$ and $\{q_n\}$ be the two sequences of non-zero real constants such that

$$P_n = p_0 + p_1 + \cdots p_n, \qquad P_{-1} = p_{-1} = 0 \qquad (2.2a)$$

$$Q_n = q_0 + q_1 + \cdots q_n, \qquad Q_{-1} = q_{-1} = 0 \qquad (2.2b)$$

Author α σ: Research Scholar, MATS University, Raipur (C.G.), India. e-mails: sanjaymukherjeeruma@gmail.com, khanaj@matsuniversity.ac.in

Ref

3. G. G. Lorentz, A contribution to the theory of divergent series, Acta Math. 80 (1948), 167-190. MR10, 367e.

Given two sequences $\{p_n\}$ and $\{q_n\}$, convolution p*q is defined by

$$R_n = (p * q)_n = \sum_{k=0}^{n} p_k q_{n-k} \tag{2.3}$$

It is familiar and can be easily verified that the operation of convolution is commutative and associative, and

$$(p * 1)_n = \sum_{k=0}^{n} p_k \tag{2.4}$$

The series $\sum a_n$ or the sequence $\{S_n\}$ is said to be almost generalized Nörlund (N,p,q) (Qureshi[6]) summable to S, if

$$t_{n,m} = \frac{1}{R_n} \sum_{v=0}^{n} p_{n-v}\, q_v S_{v,m} \tag{2.5}$$

Tends to S, as $n \to \infty$, uniformly with respect to m, where

$$S_{v,m} = \frac{1}{v+1} \sum_{k=m}^{v+m} S_k \tag{2.6}$$

Let f(t) be a periodic function with period 2π and integrable in the sense over an interval $(-\pi, \pi)$.
Let its Fourier series be given by

$$f(t) \sim \frac{1}{2} a_0 + \sum_{n=1}^{\infty} (a_n \cos nt + b_n \sin nt) = \frac{1}{2} a_0 + \sum_{n=1}^{\infty} A_n(t) \tag{2.7}$$

And then the conjugate series of (2.7) is given by

$$\sum_{n=1}^{\infty} (a_n \sin nt - b_n \cos nt) = \sum_{n=1}^{\infty} B_n(t) \tag{2.8}$$

Let $\{p_n\}$ be a nonnegative non-increasing generating sequence for (N, p_n) method such that

$$P_n = P(n) = p_0 + p_1 + p_2 + \cdots p_n \to \infty, as\ n \to \infty \tag{2.9}$$

Particular Cases:

a) Almost (N,p,q) method reduces to almost Nörlund method (N, p_n) if $q_n = 1$ for all n.

b) Almost (N,p,q) method reduces to almost Riesz method $(\overline{N}, q_n)$ if $p_n = 1$ for all n.

c) In the special case when $p_n = \binom{n+\propto+1}{\propto-1}, \alpha > 0$, the method (N, p_n) reduces to the well known method of summability $(C, \propto)$.

d) $p_n = \frac{1}{n+1}$ of the Nörlund mean is known as harmonic mean and is written as (N,1/(n+1)).

Following notations will be used:

$$\emptyset(t) = f(x+t) + f(x-t) - 2f(x)$$

$$\psi(t) = f(x+t) - f(x-t)$$

$$\Phi(t) = \int_0^t |\phi(u)|\, du$$

Notes

$$\Psi(t) = \int_0^t |\psi(u)|\, du$$

$$\tau = \left[\frac{1}{t}\right] = The\ integral\ part\ of\ \frac{1}{t}$$

Notes

III. Known Theorem

If f(x) is periodic and belongs to the class $Lip(\alpha, p) for\ 0 < \alpha \leq 1$, if the sequence $\{p_n\}$ is as defined in (2.9) with other requirements there in and if

$$\int_1^n \left(\frac{(p(y)^q)}{y^{q\alpha+2-\delta q-q}}\right) = o\left(\frac{p(n)}{n^{\frac{\alpha-1}{q-\delta-1}}}\right) \tag{3.1}$$

Then

$$\left\|\tilde{t}_n - \tilde{f}\right\|_p = o\left(\frac{1}{n^{\frac{\alpha-1}{p}}}\right) \tag{3.2}$$

where $\tilde{t_n}$ are the (N, p_n) means of the series (2.8) and 1/p +1/q =1 such that $1 \leq p$

IV. Main Theorem

Our object of this paper is to prove the following theorem.

Theorem: The monotonic non- increasing sequence of real constant of the conjugate Fourier series is (N,p,q) summable to

$$\left\|\tilde{t}_{n,m} - \tilde{f}\right\| = o(1)$$

V. Lemmas

For the proof of theorem 4, the following lemmas are required

Lemma 5.1: For $0 < t < \frac{1}{(n+m)}$, we have

$$\left|N_{n,m}(t)\right| = o(n+m)$$

Proof: For $0 < t < \frac{1}{(n+m)}$, we have

$$\left|N_{n,m}(t)\right| = \frac{1}{2\pi R_n}\left|\sum_{v=0}^{n} p_{n-v}q_v \frac{sin(v+1)(t/2)\{cos(v+2m+1)(t/2) - cos(t/2)\}}{(v+1)sin^2(t/2)}\right|$$

$$= \frac{1}{2\pi R_n}\left|\sum_{v=0}^{n} p_{n-v}q_v \frac{sin(v+1)\left(\frac{t}{2}\right)\left\{2sin\left(\frac{(v+2m+2)}{2}\right)\left(\frac{t}{2}\right)sin\left(\frac{(v+2m)}{2}\right)\left(\frac{t}{2}\right)\right\}}{(v+1)sin^2\left(\frac{t}{2}\right)}\right|$$

$$\leq \frac{1}{2\pi R_n}\left|\sum_{v=0}^{n} p_{n-v}q_v \frac{(v+1)sin\left(\frac{t}{2}\right)\left\{2sin\left(\frac{(v+2m+2)}{2}\right)\left(\frac{t}{2}\right)sin\left(\frac{(v+2m)}{2}\right)\left(\frac{t}{2}\right)\right\}}{(v+1)sin^2\left(\frac{t}{2}\right)}\right|$$

$$\leq \frac{1}{2\pi R_n}\left|\sum_{v=0}^{n} p_{n-v}q_v \frac{2((v+2m+2)/2)\left\{sin\left(\frac{t}{2}\right) sin\left(\frac{(v+2m)}{2}\right)\left(\frac{t}{2}\right)\right\}}{(v+1)sin\left(\frac{t}{2}\right)}\right|$$

$$= \frac{1}{2\pi R_n}\left\{\sum_{v=0}^{n} p_{n-v}q_v\right\}(n+2m+2)$$

$$= O(n+m)\frac{1}{R_n}\sum_{v=0}^{n} p_{n-v}q_v$$

$$|N_{n,m}(t)| = o(n+m)$$

Lemma 5.2: For $\frac{1}{(n+m)} < t < \pi$, we have

$$|\bar{N}_{n,m}(t)| = o\left(\frac{1}{t^2 n}\right)$$

Proof: For $\frac{1}{(n+m)} < t < \pi$, we have

$$\bar{N}_{n,m}(t) = \frac{1}{2\pi R_n}\sum_{v=0}^{n} p_{n-v}q_v \frac{cos\,(m+(v+1)/2)t sin\,((v+1)/2)t}{(v+1)sin^2(t/2)}$$

$$|\bar{N}_{n,m}(t)| \leq \frac{1}{2\pi R_n}\sum_{v=0}^{n} p_{n-v}q_v \frac{cos(m+(v+1)/2)t sin((v+1)/2)t}{(v+1)sin^2(t/2)}$$

$$\leq \frac{1}{2\pi R_n}\sum_{v=0}^{n} p_{n-v}q_v \frac{1}{(v+1)sin^2(t/2)}$$

$$= o\left(\frac{1}{t^2}\right)\frac{1}{R_n}\sum_{v=0}^{n}\left(\frac{p_{n-v}q_v}{(v+1)}\right)$$

$$|\bar{N}_{n,m}(t)| = o\left(\frac{1}{t^2 n}\right)$$

VI. Proof of the Theorem (4)

Let $S_k(x)$ denote the nth partial sum of the series (2.8).
Then we have

$$S_k(x) = \frac{1}{2\pi}\int_0^{\pi} \frac{cos\,(k+(1/2))t - cos\,(t/2)}{sin\,(t/2)}\psi(t)dt \quad (6.1)$$

$$= \frac{1}{2\pi}\int_0^{\pi} \frac{cos\,(k+(1/2))t}{sin\,(t/2)}\psi(t)dt - \frac{1}{2\pi}\int_0^{\pi} cot\left(\frac{t}{2}\right)\psi(t)dt \quad (6.2)$$

Now, by using (2.6) we get

$$S_{v,m} = \frac{1}{v+1}\sum_{k=m}^{v+m}\left\{\frac{1}{2\pi}\int_0^{\pi}\frac{cos\,(k+(1/2))t}{sin\,(t/2)}\psi(t)dt - \frac{1}{2\pi}\int_0^{\pi} cot(t/2)\,\psi(t)dt\right\} \quad (6.3)$$

So that by using (2.5), we have

Notes

Notes

$$t_{n,m} = \frac{1}{R_n}\sum_{v=0}^{n} p_{n-v}q_v \left\{\frac{1}{2\pi}\int_0^{\pi}\sum_{k=m}^{v+m}\frac{cos\big(k+(1/2)\big)t}{sin(t/2)}\psi(t)dt - \frac{1}{2\pi}\int_0^{\pi} cot(t/2)\,\psi(t)dt\right\}$$

$$\|\tilde{t}_{n,m} - \tilde{f}\| = \frac{1}{R_n}\sum_{0}^{\pi} p_{n-v}q_v \frac{1}{2\pi(v+1)}\int_0^{\pi}\sum_{k=m}^{v+m}\frac{cos\big(k+(1/2)\big)t}{sin(t/2)}\psi(t)dt$$

$$= \frac{1}{2\pi R_n}\sum_{v=0}^{n} p_{n-v}q_v \int_0^{\pi}\frac{sin(v+m+1)t - sinmt}{2(v+1)sin^2(t/2)}\psi(t)dt$$

$$= \int_0^{\pi}\left\{\frac{1}{2\pi R_n}\sum_{v=0}^{n} p_{n-v}q_v \frac{cos(v+2m+1)(t/2)sin(v+1)(t/2)}{(v+1)sin^2(t/2)}\right\}\psi(t)dt$$

$$= \int_0^{\pi} \overline{N}_{n,m}\,(t)\psi(t)dt$$

$$= \left\{\int_0^{1/(n+m)} + \int_{1/(n+m)}^{1/(n+m)^{\delta}} + \int_{1/(n+m)^{\delta}}^{\pi}\right\}\overline{N}_{n,m}(t)\psi(t)dt = I_1 + I_2 + I_3 \tag{6.4}$$

First we consider,

$$I_1 = \int_0^{1/(n+m)} \overline{N}_{n,m}\,(t)\psi(t)dt$$

$$= \int_0^{1/(n+m)} \frac{1}{2\pi R_n}\sum_{v=0}^{n} p_{n-v}q_v \frac{cos(v+2m+1)(t/2)sin(v+1)(t/2)}{(v+1)sin^2(t/2)}\psi(t)dt$$

$$= \int_0^{1/(n+m)} \frac{1}{2\pi R_n}\sum_{v=0}^{n} p_{n-v}q_v \frac{sin(v+1)(t/2)\{cos(v+2m+1)(t/2) - cos(t/2)\}}{(v+1)sin^2(t/2)}\psi(t)dt$$

$$+ \int_0^{1/(n+m)} \frac{1}{2\pi R_n}\sum_{v=0}^{n} p_{n-v}q_v \frac{sin(v+1)(t/2)cot(t/2)}{(v+1)sin(t/2)}\psi(t)dt$$

$$= I_{1.1} + I_{1.2} \tag{6.5}$$

Now

$$|I_{1.1}| \le \int_0^{1/(n+m)} \frac{1}{2\pi R_n}\left|\sum_{v=0}^{n} p_{n-v}q_v \frac{sin(v+1)(t/2)\{cos(v+2m+1)(t/2) - cos(t/2)\}}{(v+1)sin^2(t/2)}\right| |\psi(t)|dt$$

$$= \int_0^{1/(n+m)} \left|\overline{N}_{n,m}(t)\right||\psi(t)|dt$$

$= o(n+m)\int_0^{1/(n+m)}|\psi(t)|\,dt$ by Lemma 5.1

$$= o(n+m)o\left[\frac{\alpha(n+m)}{(n+m)R_{n+m}}\right]$$

$$= o\left[\frac{1}{log(n+m)}\right]$$

$= o(1)$, as $n \to \infty$, uniformly with respect to m (6.6)

Next, for $0 \le t \le \frac{1}{(n+m)}$

$$|I_{1.2}| \le \int_0^{1/(n+m)} \frac{1}{2\pi R_n}\sum_{v=0}^{n} p_{n-v}q_v \frac{sin(v+1)(t/2)cot(t/2)}{(v+1)sin(t/2)}\psi(t)dt$$

$$\le \frac{1}{2\pi R_n}\sum_{v=0}^{n} p_{n-v}q_v \int_0^{1/(n+m)} \frac{(v+1)sin(t/2)cot(t/2)}{(v+1)sin(t/2)}\psi(t)dt$$

$$= \frac{1}{2\pi}\int_0^{1/(n+m)} cot(t/2)\psi(t)\,dt$$

Since the conjugate function exists, therefore

$= o(1), as\ n \to \infty$, uniformly with respect to m (6.7)

Thus from (6.6) and (6.7), we get

$I_1 = o(1), as\ n \to \infty$, uniformly with respect to m (6.8)

Now, we get

$$|I_2| \le \int_{1/(n+m)}^{1/(n+m)^{\delta}} \left|\bar{N}_{n,m}(t)\right||\psi(t)|\,dt$$

$= o\int_{1/(n+m)}^{1/(n+m)^{\delta}} \frac{|\psi(t)|}{t^2 n}dt$ by Lemma 5.2

$$= o\left(\frac{1}{n}\right)\int_{1/(n+m)}^{1/(n+m)^{\delta}} \frac{|\psi(t)|}{t^2}dt$$

$$= o\left(\frac{1}{n}\right)o(n)$$

$I_2 = o(1), as\ n \to \infty$, uniformly with respect to m (6.9)

Finally, we have

Notes

Notes

$$|I_3| \leq \int_{1/(n+m)^\delta}^{\pi} \frac{1}{2\pi R_n} \sum_{v=0}^{n} p_{n-v} q_v \left| \frac{cos(v+2m+1)(t/2) sin(v+1)(t/2)}{(v+1) sin^2(t/2)} \right| |\psi(t)| dt$$

$$= \int_{1/(n+m)^\delta}^{\pi} \frac{1}{2\pi R_n} \sum_{v=0}^{n} p_{n-v} q_v \left| \frac{sin(v+m+1)(t) - sinmt}{(v+1) sin^2(t/2)} \right| |\psi(t)| dt$$

$$= \frac{1}{2\pi R_n} \sum_{v=0}^{n} p_{n-v} q_v \left[\int_{1/(n+m)^\delta}^{\pi} \left| \frac{sin(v+m+1)(t)}{(v+1) sin^2(t/2)} \right| |\psi(t)| dt + \int_{1/(n+m)^\delta}^{\pi} \left| \frac{sinmt}{(v+1) sin^2(t/2)} \right| |\psi(t)| dt \right]$$

$$= I_{3.1} + I_{3.2} \tag{6.10}$$

Now, by using second mean value theorem, we have

$$|I_{3.1}| \leq \frac{1}{2\pi R_n} \sum_{v=0}^{n} \frac{p_{n-v} q_v}{2(v+1) 2 sin^2(1/(n+m)^\delta)} \int_{1/(n+m)^\delta}^{\epsilon} |sin(v+m+1)| \, |\psi(t)| dt$$

Where $\frac{\mathbf{1}}{(\boldsymbol{n}+\boldsymbol{m})^{\boldsymbol{\delta}}} \leq \epsilon \leq \boldsymbol{\pi}. \mathbf{0} \leq \boldsymbol{\delta} \leq \frac{\mathbf{1}}{\mathbf{2}}$

$$= o\left(\frac{1}{n}\right) (n+m)^{2\delta} \left(\frac{1/2(n+m)^\delta}{sin(1/2(n+m)^\delta)} \right)^2 \int_{1/(n+m)^\delta}^{\epsilon} |\psi(t)| dt$$

$I_{3.1} = o(1), as\ n \to \infty$, uniformly with respect to m (6.11)

Now

$$|I_{3.2}| \leq \int_{1/(n+m)^\delta}^{\pi} \frac{1}{2\pi R_n} \sum_{v=0}^{n} p_{n-v} q_v \left| \frac{sinmt}{(v+1) sin^2(t/2)} \right| |\psi(t)| dt$$

$$\leq \frac{1}{2 sin^2(1/2(n+m)^\delta)} \int_{1/(n+m)^\delta}^{\epsilon} |\psi(t)| dt$$

Where $\frac{\mathbf{1}}{(\boldsymbol{n}+\boldsymbol{m})^{\boldsymbol{\delta}}} \leq \epsilon \leq \boldsymbol{\pi}. \mathbf{0} \leq \boldsymbol{\delta} \leq \frac{\mathbf{1}}{\mathbf{2}}$

$I_{3.2} = o(1), as\ n \to \infty$, uniformly with respect to m (6.12)

Now combining (6.8), (6.9) and (6.12), we get

$$\left\| \tilde{t}_{n,m} - \tilde{f} \right\| = o(1)$$

Thus completes the theorem

VII. Conclusion

If $\{p_n\}$ and $\{q_n\}$ be the monotonic non- increasing sequence of real constant such that the conjugate Fourier series is almost (N,p,q) summable then

$$\left\|\tilde{t}_{n,m} - \tilde{f}\right\| = o(1)$$

Acknowledgement

All great works done requires either active or passive support of some person, surrounding and one's close quarters.

Thus is it not hard to conclude how active assistance from senior could positively impact the execution of the paper. I am highly thankful to our faculty head Dr. A.J. Khan Sir for his active guidance throughout the completion of the paper.

I must express my very profound gratitude to my parents Smt. Uma Mukherjee, Mr.Salil Kumar Mukherjee and to my in-law's parents Smt. Bina Pani Majumder, Mr.Priyalal Majumder for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. I also thank my better half Smt. Ruma Mukherjee and my son Saptarshi Mukherjee who has given mental support throughout my work. This accomplishment would not have been possible without them.

Last but not the least, I would also want to extend my appreciation to those who could not be mentioned here but have well played their role in inspire me behind the certain.

Notes

References Références Referencias

1. G. H. Hardy, On the summability of Fourier series, Proc. London Math. Soc. (2) 12 (1913), 365-372.
2. K. S. K. Iyengar, A Tauberian theorem and its application to convergence of Fourier series, Proc. Indian Acad. Sci. , Sect. A.18 (1943), 81-87. MR 5,65e.Zbl 060.18405.
3. G. G. Lorentz, A contribution to the theory of divergent series, Acta Math. 80 (1948), 167-190. MR10, 367e.
4. S. M. Mazhar and A. H. Siddiqi, On the almost summability of a trigonometric sequence, Acta Math.Acad. Sci.Hungar. 20 (1969), 21-24. MR 39#1897.Zbl 174.10106.
5. T.Pati, A generalization of a theorem of Iyengar of Harmonic summability of Fourier series, Indian J. Math. 3 (1961), 85-90. MR 27#527. Zbl 142.31801.
6. K. Qureshi, On the degree of approximation of a periodic function f by almost Nörlund means, Tamkang J. Math. 12 (1981), no. 1, 35-38. MR 85h: 42002. Zbl 502.42002.
7. C. T. Rajagopal,On the Nörlund summability of Fourier series, Proc. Cambridge Philos. Soc. 59 (1963),47-53. MR 27#530 .Zbl 117.29501.
8. J. A. Siddiqi , On the Harmonic summabilityof Fourier series, Proc. Indian Acad. Sci. , Sect. A.28 (1948), 527-531. MR 10,369a.
9. T. Singh, On Nörlund summability of Fourier series and its conjugate series, Proc. Nat.Inst. Sci. India Part., A.29 (1963), 65-73. MR 27#5082, Zbl 141.07005.
10. T.Singh, Nörlund summability of Fourier series and its conjugate series, Ann. Mat. Pura Appl. (4) 64 (1964), 123-132. MR 29#3824.Zbl 141.07101.
11. L. M. Tripathi, On almost Nörlund summability of conjugate series of a Fourier series, Vijnana Parishad Anusandhan Patrika 27 (1984), no. 2 175-181.

Global Journal of Science Frontier Research: F
Mathematics and Decision Sciences
Volume 19 Issue 4 Version 1.0 Year 2019
Type : Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals
Online ISSN: 2249-4626 & Print ISSN: 0975-5896

An Alternative Method of Detecting Outlier in Multivariate Data using Covariance Matrix

By Obafemi, O. S. & Alabi, N. O.
The federal polythechnic Ado

Abstract- In the Multivariate data analysis, the detection of outliers is important and necessary though this may be difficult and can pose a problem to the analyst. When a set of data is contaminated, the values obtained from such set of data are distorted and the results meaningless. In this work we present a simple multivariate outlier detection procedure using a robust estimator for variance-covariance matrix by using the best units from the available data set that satisfied the three predetermined optimality criteria, selected from all possible combinations of sub-sample obtained. The proposed estimator used is the variance-covariance estimator of the best unit multiplied by a constant. It is observed that, the proposed method combined the efficiencies of the classical and the existing robust (MCD and MVE) of being able to signal when there are few and multiple outliers in multivariate data.

Keywords: *outliers, robust estimator, multivariate data, signal probability, false alarm, hotelling T^2.*

GJSFR-F Classification: *MSC 2010: 97K80*

ANALTERNATIVEMETHODOFDETECTINGOUTLIERINMULTIVARIATEDATAUSINGCOVARIANCEMATRIX

eJournal Print Journal epaper open standard 3d virtual journal

Notes

An Alternative Method of Detecting Outlier in Multivariate Data using Covariance Matrix

Obafemi, O. S. [α] & Alabi, N. O. [σ]

Abstract- In the Multivariate data analysis, the detection of outliers is important and necessary though this may be difficult and can pose a problem to the analyst. When a set of data is contaminated, the values obtained from such set of data are distorted and the results meaningless. In this work we present a simple multivariate outlier detection procedure using a robust estimator for variance-covariance matrix by using the best units from the available data set that satisfied the three predetermined optimality criteria, selected from all possible combinations of sub-sample obtained. The proposed estimator used is the variance-covariance estimator of the best unit multiplied by a constant. It is observed that, the proposed method combined the efficiencies of the classical and the existing robust (MCD and MVE) of being able to signal when there are few and multiple outliers in multivariate data.

Keywords: outliers, robust estimator, multivariate data, signal probability, false alarm, hotelling T^2.

I. Introduction

The presence of outliers can distort the values of estimators arbitrarily and render the results meaningless (Obafemi and Oyeyemi 2018). In literature, it has been opined that Outliers in multivariate data are more difficult to detect than outliers in univariate data, since simple graphical methods can be used to detecting univariate outliers, which is impossible in multivariate data. Also, multivariate data come from many sources apart from the normal population. There could be outliers due to changes of location in random directions for each outlier, there could be a cluster of outliers due to location shift in a particular direction, there could be multiple clusters of outliers in different directions, there could be outliers with the same location as proper data but with more variability, outlier can also be due to shift in some of the elements of the location vector but not all of them (Rocke and Woodruff, 1996).

Rocke and woodruff, (1996) affirmed that the most problematic type of multivariate outliers detection are those clean data that have the same variance – covariance matrix. Barnet and Lewies (1994) argued that the moments used in describing data are often influenced by outliers.

Majorly, most rules adopted mean ±2 standard deviations from the observation as the outliers, which are identified for "clean" data, or at least no distinction is made between outliers and extremes of a distribution. The basis for multivariate outlier detection is the Mahalanobis distance incorporated into the standard method of robust estimation of the parameter estimates. This is compared with critical value of χ^2

Author α: Department of Mathematics and Statistics, The federal polythechnic Ado- Ekiti, Nigeria. e-mail: obafemisamuel22@gmail.com
Author σ: Department of Mathematics and Statistics, The federal polythechnic, Ilaro, Ogun state, Nigeria.
e-mail: nurudene.alabi@gmail.com

distribution Rousseeuw and van Zomeren (1990). Thus values above the rejection level may not always be outliers; they could still be among the data distribution.

To reduce the multivariate detection problems, Gnanadesikan and Kettering (1972) proposed a set of univariate solution by looking at projections of the data onto some direction. They chose the direction of maximum variability of data and, therefore, they suggested obtaining the principal components of the data and search for outliers in these directions. This method provides the correct solution when outliers are situated close to the directions of the principal components; in general case, this may fail to identify outliers.

Notes

An alternative approach developed by Maronna (1976) is to use robust location and scale estimators. He studied affine equivariant M estimators for covariance matrices, and campbell (1980) proposed using the Mahalanobis distance computed using M estimators for mean and covariance matrix. Stahel (1981) and Donoho (1982) proposed that to solve the dimensionality problem, by computing the weights for the robust estimators from the projections of the data onto some directions, these directions were chosen to maximize distances based on univariate location and scale estimators, and the optimal values for the distances could also be used to weigh each point in the computation of a robust covariance matrix. Rousseeuw (1985) proposed a different procedure based on the computation of the ellipsoid with the least volume or with the smallest covariance determinant that would encompass at least half of the data points.

Recently, many studies frequently involve a large number of variables and observations due to the availability of computer software, which, makes the computation easier and faster but does not account for the detection of outlying observation. In micro array studies, most researchers often work with a large number of variables even with few data at time, and these portend danger of the presence of contaminated observations since it will take time if outlying observations are to be detected, in such large data set. In most cases, they carry out their classifying analysis without taken note of outliers and such a classification may be invalid.

In p-dimensional multivariate normal data, both the location and shape parameters are the most concerned issue. The location is the mean vector which denotes a point in the multi-dimensional space and scatter or shape is the variance–covariance matrix of the dimensional space. In multivariate data, it is assumed that the data follow well-behaved statistical distribution. The Independent Standard Multivariate data are usually assumed to be normally distributed with zero (0) mean and units variance. Though, the assumption may not hold when the characteristics of the data complicate or confound both estimation and hypothesis testing, Jackson and Chen (2004). A principal factor leading to such problems is the influence of outliers.

II. Effect of Outliers in Multivariate Quality Control Charts

An outlier is an observation that deviates so much from other observations as to arouse suspicion that it was generated by different mechanism as defined in statistical quality control concepts (Hawkins, 1980). Outlier has been known to have a strong influence on resulting estimates and cause any out-of-control observations to remain undetected. By using the univariate or the multivariate method, outliers can be detected. When there are more than one outliers, the detection situations become more difficult due to masking and swamping (Rousseeuw and van Zomeren, 1990). When we fail to detect the outliers, masking occurs while swamping occurs when observations are incorrectly declared as outliers.

Notes

Outliers can heavily influence the estimation of the scatter matrix and subsequently, the parameters or statistics that are needed to be derived from it. Therefore a robust estimate of scatter matrix that would not be affected by outliers is required to obtain valid results (Hubert and Engelen, 2007).

Control charts are the most popular tools and techniques used in statistical process control (SPC) to monitor the quality characteristics of products and services in organizations and industries. In many of these industrial processes, it is frequently required to monitor several quality characteristics at the same time, such quality characteristics may include weight, degree of hardness, thickness, width and length of tablets (Liu, 1995). For the fact that the quality characteristics of these products are clearly correlated, the separate univariate control charts for monitoring such quality characteristics may not be good enough to detect outliers and changes in the overall quality of the products, therefore it is desirable to have a control charts that can measure and monitor these characteristics simultaneously, the multivariate control charts tend to be the most appropriate tools applicable in such situations. The simultaneous nature of the control scheme and the correlation structure between the qualities characteristics are taken into consideration by these control charts (Alt, 1985).

III. Statement of the Problem

A determination of appropriate critical value for the detection of outliers in univariate or multivariate is as a result of two major subjective elements. These are whether to investigate at all and, if so for how many outliers are to be tested, Collett and Lewis (1976) opined that failing to test might render the apparent significance levels invalid. The most harmful types of outliers, especially if there are several of them, may affect the estimated model so much "in their direction" and bring about poor inferences. In the light of the above-stated problems, the study proposes an alternative methods of detecting outliers, which is deterministic, robust, and also attempt to compare it with existing methods.

IV. Scope of the Study

The identification of multivariate outliers is particularly difficult, a variety of methods have been developed for detecting single point outliers which, when applied to groups of contaminated data, it leads to problems of "masking". Robust high-breakdown estimators overcome the masking effect, also allow for high tolerance of "bad" data. On the contrary, most of the robust statistics have a breakdown at a fraction 1/(p+1) of contaminated data, where p is the dimension. Therefore, high-breakdown estimators are particularly useful in high dimensional sets.

Different methods have been offered by the literature as well as feasible algorithms for their computation. The minimum volume Ellipsoid and the Minimum Covariance Determinant estimator are the most widely known among them.

With the later having better statistical properties than the former, however lack of a fast and efficient algorithm has made its use limited. The FSA (Feasible solution Algorithm) proposed by Hawkin (1994) is computationally heavy and relatively slow: the fast algorithm of Rousseeuw and Drissen (1999) solves problems of speed, and the forward search for the MCD by Aderson (1994) applies a simple but efficient criterion. These three aforementioned, are the main algorithm as developed for the computation of MCD estimate.

Robust methods allow us to find estimates for both the location and the scatter of a multivariate cloud according to robustness criteria and to detect groups of outliers at the same time. The study examined the classical and robust estimator of detecting outlier with respect to location and scatter.

V. Methodology

a) The Proposed Alternative Estimator for Outlier Detection

Given $y_1, y_2, \ldots, y_p$ for multivariate normal, i.e. $Y_p \sim N_p(\mu, \Sigma)$ where Σ is positive definite. The proposed method of estimating the parameter μ *and* Σ focused more on the eigen roots of the variance-covariance matrix. Given a p-dimensional multivariate normal data Y_{pxm} with m observation $\{y_i\}_{i=1}^{m}$, the interest here is to obtain a subset of $\{y_i\}_{i=1}^{m}$ of size k =p+1 that satisfy some criteria stated below:

$$C_A = least\{A(\lambda_{ij}), j = 1,2,3,,, C_k^m\}, \text{ where } A(\lambda_{ij}) = \frac{\sum_{i=1}^{p} \lambda_{ij}}{p}$$

$$C_H = least\{H(\lambda_{ij}), j = 1,2,3,,, C_k^m\}, \text{ where } H(\lambda_{ij}) = \frac{p}{\sum_{i=1}^{p} \frac{1}{\lambda_{ij}}}$$

C_G = $least\{G(\lambda_{ij}), j = 1,2,3,,, C_k^m\}$, where $G(\lambda_{ij}) = \sqrt[p]{\prod_{1=1}^{p} \lambda_{ij}}$ where $A(\lambda_{ij}), H(\lambda_{ij})$, and $G(\lambda_{ij})$ are the arithmetic, harmonic and geometric means of $\lambda's$ respectively and $\lambda's$ are the eigen- roots obtained from the covariance matrix

A sample of size k from m is therefore drawn that will give C_{p+1}^{m} possible subsets of size p + 1. The variance-covariance matrix Σ_j is therefore estimated as $\Sigma_j = \frac{1}{p+1}(y_j - \bar{y}_j)(y_j - \bar{y}_j)^T$.

For each of the pxp matrix Σ_j, the eigen-values $\lambda_{j1}, \lambda_{j2}, \ldots, \lambda_{jp}$ are obtained. from the eigen-roots, the Arithmetic, the harmonic and the geometric mean of the eigen-value denoted by A, H, and G respectively, from which the above optimality criteria are defined.

The objective here is to obtain data points whose variance–covariance matrix will satisfy at least two of the criteria taking into consideration when the variance-covariance matrix is from uncorrelated variables and correlated variables.

The resulting covariance matrix will be inflated or deflated to accommodate good data points within the observed data.

b) Algorithm for the Proposed Estimator

Given a P-dimensional multivariate normal data Y_{pxn} with n observations, $\{y_i\}_{i=1}^{n}$.

1. Decompose the data using singular value decomposition(SVD)

Notes

Notes

2. From the n observations from the above matrix, take a subsample of size k = p + 1, C_k^n times
3. For each sample of p + 1, obtain the three optimality criteria $\{C_A, C_H, C_G\}$ of the eigen- roots of the matrix
4. Seek the sample points that satisfy at least two of the optimality criteria.
5. Obtain the classical mean vector and variance-covariance matrix;
   $\bar{y}_k = \frac{1}{p+1}\sum_{i\in k} y_i$ and $S_k = \frac{1}{p}\sum_{i\in k}(y_i - \bar{y})(y_i - \bar{y})^T$ respectively.
6. Use the estimates to obtain the Mahalanobis distances;
   $d_j^2(i) = (y_i - \bar{y}_J)^T S_J^{-1}(y_i - \bar{y}_J)$, i = 1, 2, 3, …, n
7. The Mahalanobis distances are then ordered such that $d_1^2 \le d_2^2 \le d_3^2 \le, \ldots, \le d_n^2$
8. The p+2 points that correspond to the first p+2 ordered distances are picked to estimate the new estimates of mean vector and variance-covariance matrix.
9. Steps 5 to 7 are repeated until the selected sample points are h, where $h = \frac{n+p+1}{2}$.

The classical variance-covariance matrix of the h points is the robust estimate of the vector scatter matrix given as;

$$S_{proposed} = \frac{1}{h-1}\sum_{i\in h}(y_i - \bar{y}_{proposed})(y_i - \bar{y}_{proposed})^T \cdot (\chi^2_{j,0.025})^{\frac{1}{p}}$$

where $(\chi^2_{p,0.025})^{\frac{1}{p}}$ is a correcting factor with p as the dimension, $h = \frac{(n+p+1)}{2}$.

c) Comparison Of The Methods By Application To Some Multivariate Techniques Via Data Simulation

The Classical, MVE, MCD, and the proposed methods are applied to multivariate techniques and compared to determine their performances.

I. *The Simulation Study*

The Monte Carlo method of simulation is adopted to generate multivariate data set for comparing the proposed method of estimation with the other three methods. The simulation series considered the bivariate and tri-variate normal distribution. Sample size n=30 with contaminations of 1, 3, and 7 data point are considered. Each run consisted of 1000 iterations of size n. The control limit was determined such that the signal probability and false alarm, i.e. type I error (α) were based on the assumption of the Non-Centrality Parameter NCP= $E[\mu - \mu_0]^T \Sigma^{-1} E[\mu - \mu_0]$ to be the measure of severity of a shift to the out-of-control mean vector $\underset{-}{\mu}$ from the in-control mean vector μ_0 because the signal probability depends on the in-control mean vector $\underset{-}{\mu_0}$ or the variance-covariance $\sum$. We considered the mean vector=, $\begin{bmatrix}0\\0\end{bmatrix}$, $\begin{bmatrix}0\\0\\0\end{bmatrix}$ and the variance co-variance $=\begin{bmatrix}2 & 2\\2 & 4\end{bmatrix}$, and $\begin{bmatrix}4 & 1 & 3\\1 & 6 & 2\\3 & 2 & 8\end{bmatrix}$, for the bivariate and tri-variate cases, respectively. The simulated upper control limits were determined from 1000 simulation such that all the methods considered had an overall false alarm probability of 0.05. The limits were obtained by generating 1000 data set for n=30 and p=2, p=3. The HotellingT^2 statistic, T_i^2 were computed for i = 1,2,,, n. The maximum value was recorded and the 95th percentile of the maximum value of Hotelling T^2 for j = 1,2,,,,1000 was taken to be the upper control limits for the control chart. The values obtained are 9.02, 16.29, 16.29, and 15.42 for the normally distributed variables for Classical, MCD, MVE and Proposed methods, respectively. The lower control limit is normally set to be zero.

K (k=1, 3, and 7) outliers are randomly generated among the n (n=30) observations once the control limits are set. To generate the outliers, the process means vector was changed from $\mu=\mu_0$ to $\mu=\mu_1$ to obtain the given value of non- centrality parameter. The resulted probability of valid signal and the probability of false alarmed were compared.

Tables 1a - 6b showed the estimated signal probabilities and probabilities of false alarm for different non- centrality parameter values (NCP= 1, 2, 3, 4, 5, 6, 7, 8, 9, 10)

Notes

Table 1a: Results of Signal probability with multivariate normal distribution when p=2

Signal Probability when there is 1 outlier				
NCP	classical	mcd	mve	Proposed
1	0.000	0.000	0.000	0.000
2	0.000	0.000	0.000	0.000
3	0.000	0.000	0.000	0.000
4	0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000
6	0.000	0.000	0.000	0.000
7	0.000	0.000	0.000	0.000
8	0.000	0.000	0.000	0.000
9	0.000	0.000	0.000	0.000
10	0.000	0.000	0.000	0.000

Table 1b: Results of false alarm with multivariate normal distribution when p=2

Probability of false alarm when there is 1 outlier				
NCP	classical	mcd	mve	Proposed
1	0.0000	0.1035	0.1035	0.0689
2	0.0000	0.1035	0.1035	0.0689
3	0.0000	0.1035	0.1035	0.0689
4	0.0000	0.1035	0.1035	0.0689
5	0.0000	0.1035	0.1035	0.0689
6	0.0000	0.1035	0.1035	0.0689
7	0.0000	0.1035	0.1035	0.0689
8	0.0000	0.1035	0.1035	0.0689
9	0.0000	0.1035	0.1035	0.0689
10	0.0000	0.1035	0.1035	0.0689

Notes

Table 2a: Signal probability with multivariate normal distribution when p=2 with outlier equal 3

Signal Probability when there are 3 outliers				
NCP	classical	mcd	mve	Proposed
1	0.000	0.000	0.000	0.000
2	0.000	0.000	0.000	0.000
3	0.000	0.000	0.000	0.000
4	0.000	0.000	0.000	0.333
5	0.000	0.000	0.000	0.667
6	0.000	0.333	0.333	0.667
7	0.000	0.667	0.667	0.667
8	0.333	0.667	0.667	1.000
9	0.333	1.000	1.000	1.000
10	0.333	1.000	1.000	1.000

Table 2b: Probability of false alarm when p=2 with outlier equal 3

Probability of false alarm when there are 3 outliers				
NCP	classical	mcd	mve	Proposed
1	0.0370	0.0370	0.0370	0.0000
2	0.0370	0.0370	0.0370	0.0000
3	0.0370	0.0370	0.0370	0.0000
4	0.0370	0.0370	0.0370	0.0000
5	0.0370	0.0370	0.0370	0.0000
6	0.0370	0.0370	0.0370	0.0000
7	0.0370	0.0370	0.0370	0.0000
8	0.0370	0.0370	0.0370	0.0000
9	0.0370	0.0370	0.0370	0.0000
10	0.0370	0.0370	0.0370	0.0000

Table 3a: Signal probability and false alarm when p=2

Signal Probability when there are 7 outliers				
NCP	classical	mcd	mve	Proposed
1	0.0000	0.0000	0.0000	0.0000
2	0.0000	0.0000	0.0000	0.0000
3	0.0000	0.1429	0.1429	0.1429
4	0.0000	0.1429	0.1429	0.2857
5	0.0000	0.5714	0.5714	0.7143
6	0.0000	0.5714	0.5714	0.7143
7	0.0000	0.8571	0.8571	0.7143
8	0.0000	0.8571	0.8571	0.8571
9	0.0000	0.8571	0.8571	0.8571
10	0.0000	1.0000	1.0000	0.8571

Table 3b: Probability of false alarm when p=2

Probability of false alarm when there are 7 outliers				
NCP	classical	mcd	mve	Proposed
1	0.0000	0.0000	0.0000	0.0000
2	0.0000	0.0000	0.0000	0.0000
3	0.0000	0.0000	0.0000	0.0000
4	0.0000	0.0000	0.0000	0.0000
5	0.0000	0.0000	0.0000	0.0000
6	0.0000	0.0000	0.0000	0.0435
7	0.0000	0.0435	0.0435	0.0000
8	0.0000	0.0870	0.0870	0.0000
9	0.0000	0.1304	0.1304	0.0000
10	0.0000	0.1304	0.1304	0.0435

Notes

Table 4a: Signal probability when p=3

Signal Probability when there is 1 outlier				
NCP	classical	mcd	mve	Proposed
1	0.0000	0.0000	0.0000	0.0000
2	0.0000	0.0000	0.0000	0.0000
3	0.0000	0.0000	0.0000	0.0000
4	0.0000	0.0000	0.0000	0.0000
5	0.0000	0.0000	0.0000	0.0000
6	0.0000	0.0000	0.0000	0.0000
7	0.0000	0.0000	0.0000	0.0000
8	0.0000	0.0000	0.0000	0.0000
9	0.0000	0.0000	0.0000	0.0000
10	0.0000	0.0000	0.0000	0.0000

Table 4b: Probability of false alarm when p=3

Probability of false alarm when there is 1 outlier				
NCP	classical	mcd	mve	Proposed
1	0.0345	0.1379	0.1379	0.0690
2	0.0345	0.1379	0.1379	0.0690
3	0.0345	0.1379	0.1379	0.0690
4	0.0345	0.1379	0.1379	0.0690
5	0.0345	0.1379	0.1379	0.0690
6	0.0345	0.1379	0.1379	0.0690
7	0.0345	0.1379	0.1379	0.0690
8	0.0345	0.1379	0.1379	0.0690
9	0.0345	0.1379	0.1379	0.0690
10	0.0345	0.1379	0.1379	0.0690

Notes

Table 5a: Signal probability when p=3

Signal Probability when there are 3 outliers				
NCP	classical	mcd	mve	Proposed
1	0.0000	0.0000	0.0000	0.0000
2	0.0000	0.0000	0.0000	0.0000
3	0.0000	0.0000	0.0000	0.0000
4	0.0000	0.3333	0.6667	0.0000
5	0.0000	0.3333	0.3333	0.3333
6	0.0000	0.3333	0.3333	0.3333
7	0.0000	0.3333	0.3333	0.3333
8	0.3333	1.0000	1.0000	1.0000
9	0.3333	1.0000	1.0000	1.0000
10	0.3333	1.0000	1.0000	1.0000

Table 5b: Probability of false alarm when p=3

Probability of false alarm when there are 3 outliers				
NCP	classical	mcd	mve	Proposed
1	0.0370	0.0370	0.0000	0.0370
2	0.0370	0.0370	0.0741	0.0000
3	0.0000	0.0370	0.0370	0.1111
4	0.0000	0.0370	0.0370	0.0000
5	0.0000	0.0000	0.0000	0.0000
6	0.0000	0.0370	0.0370	0.0000
7	0.0000	0.0370	0.0370	0.0000
8	0.0000	0.0000	0.0000	0.0370
9	0.0000	0.0000	0.0000	0.0000
10	0.0000	0.0000	0.0000	0.0000

Table 6a: Signal probability when p=3

Signal Probability when there are 7 outliers				
NCP	Classical	mcd	mve	Proposed
1	0.000	0.1429	0.1429	0.1429
2	0.000	0.1429	0.000	0.0000
3	0.000	0.1429	0.1429	0.0000
4	0.000	0.1429	0.1429	0.1429
5	0.000	0.2857	0.0000	0.1429
6	0.000	0.4286	0.0000	0.0000
7	0.000	0.7143	0.2857	0.0000
8	0.000	0.8571	0.2857	0.2827
9	0.000	1.0000	0.5714	0.2857
10	0.000	1.0000	0.7143	0.4286

Table 6b: Probability of false alarm when p=3

Probability of false alarm when there are 7 outliers				
NCP	Classical	mcd	mve	Proposed
1	0.0000	0.0870	0.0870	0.0870
2	0.0000	0.0870	0.0000	0.0000
3	0.0000	0.0435	0.0000	0.0000
4	0.0000	0.0435	0.0000	0.0000
5	0.0000	0.0435	0.0000	0.0000
6	0.0000	0.0435	0.0000	0.0000
7	0.0000	0.0435	0.0000	0.0000
8	0.0000	0.0870	0.0000	0.0000
9	0.0000	0.0870	0.0000	0.0000
10	0.0000	0.1304	0.0000	0.0000

VI. Observations

From Table 1a, when there is only one outlier with p=2, all the four methods failed to detect the single outlier irrespective of the magnitude of the outlier. For the false alarm, all the methods raised false alarm except the classical while the false alarm of the proposed method is smaller than the one raised by both MVE and MCD methods.

When the numbers of outliers are three (3) as shown in Table 2a and 2b, all the four methods detected the outliers though at different levels of the outliers' magnitude. The proposed started detecting the outliers when NCP = 4 with the probability of 0.333 and the probability of 1 was attained when NCP = 8. Both MCD and MVE did not detect any outlier until when NCP = 6 with the probability of .333 and attained probability of 1 when NCP = 9. The Classical method performed poorly as it started detecting the outlier when NCP = 8 and did not attain the probability of 1 throughout the range of NCP used in the simulation. Also, all the methods gave the same false alarm though with a small probability of 0.037 irrespective of the magnitude of the outliers except the proposed method, which has zero probability of false alarm.

When the number of outliers was further increased to seven, the classical method shows no presence of outliers at all the level of magnitude, while the other three robust methods indicated the presence of outliers with signal probabilities of 0.1429 at NCP=3, this increased gradually though at different values till the level of magnitude is 10. The MCD and MVE maintained the same values of signal all through the level and attained the probability of 1 at NCP=10, the proposed remain 0.8571 signal probability at NCP=10. The MCD and MVE only indicated no false alarm within the first and sixth level of magnitude, while the classical and the proposed method maintained no false alarm all through the levels (Table 3a and b).

From table 4a, when there is one outlier, with p=3 all the three methods failed to signal for the presence of outlier irrespective of the magnitude of the outlier's level. For the false alarm, all the methods, Classical, MCD, MVE and Proposed false alarm of 0.0345, 0.1379, 0.1379 and 0.0690 respectively which was constant at all the levels of outlier's magnitude, with the least signal raised by classical methods.

From table 5a and b when three outliers were introduced, all the three methods started detecting outliers, though at a different level of outlier's magnitude. The Classical had its first signal of 0.333 when NCP=8, The MCD, and MVE started detecting outliers with probabilities of 0.333 and 0.667 respectively when NCP=4 while

Notes

the Proposed method gave its first signal of 0.333 when the NCP=5. All the three robust methods attain the probability signal of 1 at NCP=8, 9 and 10. For the false alarm, the classical MCD and Proposed methods raise a false alarm of 0.0370 at NCP=1 while the MVE method raises no alarm at that same level of magnitude but raises a false alarm of 0.0741 when NCP=2. The Classical, however, raises no more alarm from NCP=3 to 10. The Proposed method only alarm again at NCP=3 and 8, while the MCD and MVE did not signal at NCP=5, 8, 9 and 10.

Notes

When the number of outliers was increased to 7, the classical method failed to detect outliers at all the levels of magnitude; the three other methods started detecting outliers at NCP=1. The MCD method gave a signal at all the levels of magnitude, but the MVE did not give a signal of the presence of outliers at NCP=2, 5 and 6 while the Proposed method also failed to signal at NCP=2,3,6 and 7. For the false alarm, the Classical method did not give any false alarm at all the level of magnitude while the other methods gave a false alarm at one level or the other. However the MVE only gave a false alarm at NCP=1, and the Modified gave a false alarm at NCP=1, 4 and 5 (Table 6a and b)

VII. Conclusion and Recommendation

In signaling the presence of outlier, the proposed method performed comparably well with the other existing and widely used robust methods and performed better than the classical method when the number of outliers injected is high. In most cases, the proposed method performed better than the other methods in terms of raising false alarms except in some few cases at all level of NCP when the outliers injected is high. The method performed better than the other two robust methods when the outliers are few or single.

The classical method of estimation is only efficient in detecting outlier when no or very few number of outlier is present in the data set while the other two robust methods study in this work is efficient when there is the presence of multiple outliers in the data set.

However, the proposed robust method performed better and more efficient in the two extreme cases. The efficiencies of the classical and the existing and widely used robust method (MVE and MCD) of estimation is combined by the proposed robust method in terms of outlier detection. Generally, if the presence of outliers in multivariate data set cannot be ascertained, that is; if there is no information regarding the number of outlier in the multivariate data set as far as the analyst is concern, it is recommended therefore, that the proposed robust method of detection be used in detecting the outliers that may be present in the data set.

References Références Referencias

1. Alt, F.B. (1985). Multivariate Quality control in S.Kotz and Johnson (Eds), The Ecyclopedia of statistical Sciences, 6, 110-122. New York: John Wiley & Sons.
2. Anderson,T.W. (1994): "*An Introduction to Multivariate Statisticcal Analysis*" (2nd Edn) New York John Wiley and Sons.
3. Bernet and Lewis, (1994): "*Outliers in statistical Data*". *3rd Edn.* Wiley, New York, USA.
4. Collett, D. and Lewis,T (1976): "*The subjective nature of outlier rejection procedures*". Appi. Statist. 25, 228-237.

5. Donoho, D.C. (1982): *"Breakdown properties of multivariate location estimators"* unpublished Ph.D. qualifying paper, Harvard University Department of Statistics.
6. Gnanadesikan, R. and Kettenring, J.R. (1972): "*Robust Estimates, Residuals, and outliers Detection with multiresponse data"* Biometrics, 28, 81-124
7. Hawkins, D.M. (1980): *"Identification of Outliers".* New York, Chapman and Hall, LTD.
8. Hawkins, D.M. (1994): "*The feasible Solution Algorithm For the Minimum Covariance Estimator in Multivariate Data".* Computational statistics and data analysis. 17(2), 197-210.
9. Hubert, M and Engelen, S. (2007): "*Fast Cross –validation of High- breakdown Resampling Algorithms for PCA".* Computational Statistics & data Analysis, 51(10) 5013-5024.
10. Jackson D.A. and Chen, Y. (2004): *"Robust Principal Component Analysis and outlier detection with ecological data".* Eviornemetics, 15(2), 159-169.
11. Liu,R.V. (1995): "*Control Charts for Multivariate Processes".* Journal of the American Statistical Association, 90 (432), 1380-1387
12. Maronna, R.A. (1976): "*Robust M- Estimators of multivariate location and scatter"* the annals of Statistics, 4, 51-67.
13. Obafemi, O.S. and Oyeyemi, G.M. (2018): "*Alternative estimator for multivariate location and scatter matrix in the presence of outlier"* Annals. Computer Science Series. 16^{th} Tome 2^{nd} fasc. 130-136
14. Rocke, D. M. and woodruff, D. C. (1996): "*Identification of outliers in multivariate Data".* Journal of American Association, 91, 1047 – 1061.
15. Rousseeuw (1985): *"multivariate estimators with high breakdown points"* in Mathematical Statistics and it application (VolB) pp.283-297.
16. Rousseeuw, P. J. and Van Driessen, K. (1999): "*A fast algorithm for the minimum covariance determinant estimator"*. Technometrics. 41, 212-233
17. Rousseeuw P. J. and Van Zomeren B. C. (1990): "*unmasking multivariate ouliers and leverage points."* Journal of the American statistical Association. Vol. 85 (411), pp. 633-651.
18. Stahel, W.A.(1981): *"Robuste schatzungen: Infinitestimate optimalitat and schatzungen von kovarianzmatizen"* unpublished Ph.D. thesis gdgennoscisch Technische Hochschule, Zurich.

Notes

Global Journals Guidelines Handbook 2019

www.GlobalJournals.org

Fellows

FELLOW OF ASSOCIATION OF RESEARCH SOCIETY IN SCIENCE (FARSS)

Global Journals Incorporate (USA) is accredited by Open Association of Research Society (OARS), U.S.A and in turn, awards "FARSS" title to individuals. The 'FARSS' title is accorded to a selected professional after the approval of the Editor-in-Chief/Editorial Board Members/Dean.

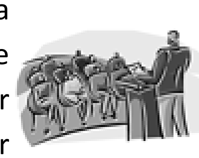
- The "FARSS" is a dignified title which is accorded to a person's name viz. Dr. John E. Hall, Ph.D., FARSS or William Walldroff, M.S., FARSS.

FARSS accrediting is an honor. It authenticates your research activities. After recognition as FARSB, you can add 'FARSS' title with your name as you use this recognition as additional suffix to your status. This will definitely enhance and add more value and repute to your name. You may use it on your professional Counseling Materials such as CV, Resume, and Visiting Card etc.

The following benefits can be availed by you only for next three years from the date of certification:

FARSS designated members are entitled to avail a 40% discount while publishing their research papers (of a single author) with Global Journals Incorporation (USA), if the same is accepted by Editorial Board/Peer Reviewers. If you are a main author or co-author in case of multiple authors, you will be entitled to avail discount of 10%.

Once FARSB title is accorded, the Fellow is authorized to organize a symposium/seminar/conference on behalf of Global Journal Incorporation (USA). The Fellow can also participate in conference/seminar/symposium organized by another institution as representative of Global Journal. In both the cases, it is mandatory for him to discuss with us and obtain our consent.

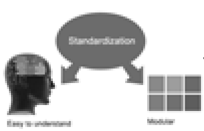


You may join as member of the Editorial Board of Global Journals Incorporation (USA) after successful completion of three years as Fellow and as Peer Reviewer. In addition, it is also desirable that you should organize seminar/symposium/conference at least once.

We shall provide you intimation regarding launching of e-version of journal of your stream time to time.This may be utilized in your library for the enrichment of knowledge of your students as well as it can also be helpful for the concerned faculty members.

The FARSS can go through standards of OARS. You can also play vital role if you have any suggestions so that proper amendment can take place to improve the same for the benefit of entire research community.

As FARSS, you will be given a renowned, secure and free professional email address with 100 GB of space e.g. johnhall@globaljournals.org. This will include Webmail, Spam Assassin, Email Forwarders,Auto-Responders, Email Delivery Route tracing, etc.



The FARSS will be eligible for a free application of standardization of their researches. Standardization of research will be subject to acceptability within stipulated norms as the next step after publishing in a journal. We shall depute a team of specialized research professionals who will render their services for elevating your researches to next higher level, which is worldwide open standardization.

The FARSS member can apply for grading and certification of standards of their educational and Institutional Degrees to Open Association of Research, Society U.S.A. Once you are designated as FARSS, you may send us a scanned copy of all of your credentials. OARS will verify, grade and certify them. This will be based on your academic records, quality of research papers published by you, and some more criteria. After certification of all your credentials by OARS, they will be published on your Fellow Profile link on website https://associationofresearch.org which will be helpful to upgrade the dignity.

The FARSS members can avail the benefits of free research podcasting in Global Research Radio with their research documents. After publishing the work, (including published elsewhere worldwide with proper authorization) you can upload your research paper with your recorded voice or you can utilize chargeable services of our professional RJs to record your paper in their voice on request.

The FARSS member also entitled to get the benefits of free research podcasting of their research documents through video clips. We can also streamline your conference videos and display your slides/ online slides and online research video clips at reasonable charges, on request.

The FARSS is eligible to earn from sales proceeds of his/her researches/reference/review Books or literature, while publishing with Global Journals. The FARSS can decide whether he/she would like to publish his/her research in a closed manner. In this case, whenever readers purchase that individual research paper for reading, maximum 60% of its profit earned as royalty by Global Journals, will be credited to his/her bank account. The entire entitled amount will be credited to his/her bank account exceeding limit of minimum fixed balance. There is no minimum time limit for collection. The FARSS member can decide its price and we can help in making the right decision.

The FARSS member is eligible to join as a paid peer reviewer at Global Journals Incorporation (USA) and can get remuneration of 15% of author fees, taken from the author of a respective paper. After reviewing 5 or more papers you can request to transfer the amount to your bank account.

MEMBER OF ASSOCIATION OF RESEARCH SOCIETY IN SCIENCE (MARSS)

The ' MARSS ' title is accorded to a selected professional after the approval of the Editor-in-Chief / Editorial Board Members/Dean.

The “MARSS” is a dignified ornament which is accorded to a person’s name viz. Dr. John E. Hall, Ph.D., MARSS or William Walldroff, M.S., MARSS.

MARSS accrediting is an honor. It authenticates your research activities. After becoming MARSS, you can add 'MARSS' title with your name as you use this recognition as additional suffix to your status. This will definitely enhance and add more value and repute to your name. You may use it on your professional Counseling Materials such as CV, Resume, Visiting Card and Name Plate etc.

The following benefitscan be availed by you only for next three years from the date of certification.

MARSS designated members are entitled to avail a 25% discount while publishing their research papers (of a single author) in Global Journals Inc., if the same is accepted by our Editorial Board and Peer Reviewers. If you are a main author or co-author of a group of authors, you will get discount of 10%.

As MARSS, you will be given a renowned, secure and free professional email address with 30 GB of space e.g. johnhall@globaljournals.org. This will include Webmail, Spam Assassin, Email Forwarders,Auto-Responders, Email Delivery Route tracing, etc.

We shall provide you intimation regarding launching of e-version of journal of your stream time to time.This may be utilized in your library for the enrichment of knowledge of your students as well as it can also be helpful for the concerned faculty members.

The MARSS member can apply for approval, grading and certification of standards of their educational and Institutional Degrees to Open Association of Research, Society U.S.A.

Once you are designated as MARSS, you may send us a scanned copy of all of your credentials. OARS will verify, grade and certify them. This will be based on your academic records, quality of research papers published by you, and some more criteria.

It is mandatory to read all terms and conditions carefully.

Institutional Fellow of Global Journals Incorporation (USA)-OARS (USA)

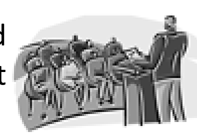
Global Journals Incorporation (USA) is accredited by Open Association of Research Society, U.S.A (OARS) and in turn, affiliates research institutions as "Institutional Fellow of Open Association of Research Society" (IFOARS).

The "FARSC" is a dignified title which is accorded to a person's name viz. Dr. John E. Hall, Ph.D., FARSC or William Walldroff, M.S., FARSC.

The IFOARS institution is entitled to form a Board comprised of one Chairperson and three to five board members preferably from different streams. The Board will be recognized as "Institutional Board of Open Association of Research Society"-(IBOARS).

The Institute will be entitled to following benefits:

The IBOARS can initially review research papers of their institute and recommend them to publish with respective journal of Global Journals. It can also review the papers of other institutions after obtaining our consent. The second review will be done by peer reviewer of Global Journals Incorporation (USA) The Board is at liberty to appoint a peer reviewer with the approval of chairperson after consulting us.
The author fees of such paper may be waived off up to 40%.



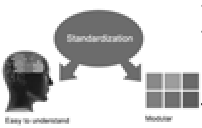
The Global Journals Incorporation (USA) at its discretion can also refer double blind peer reviewed paper at their end to the board for the verification and to get recommendation for final stage of acceptance of publication.

The IBOARS can organize symposium/seminar/conference in their country on behalf of Global Journals Incorporation (USA)-OARS (USA). The terms and conditions can be discussed separately.

The Board can also play vital role by exploring and giving valuable suggestions regarding the Standards of "Open Association of Research Society, U.S.A (OARS)" so that proper amendment can take place for the benefit of entire research community. We shall provide details of particular standard only on receipt of request from the Board.

The board members can also join us as Individual Fellow with 40% discount on total fees applicable to Individual Fellow. They will be entitled to avail all the benefits as declared. Please visit Individual Fellow-sub menu of GlobalJournals.org to have more relevant details.

We shall provide you intimation regarding launching of e-version of journal of your stream time to time. This may be utilized in your library for the enrichment of knowledge of your students as well as it can also be helpful for the concerned faculty members.



After nomination of your institution as “Institutional Fellow” and constantly functioning successfully for one year, we can consider giving recognition to your institute to function as Regional/Zonal office on our behalf.
The board can also take up the additional allied activities for betterment after our consultation.

The following entitlements are applicable to individual Fellows:

Open Association of Research Society, U.S.A (OARS) By-laws states that an individual Fellow may use the designations as applicable, or the corresponding initials. The Credentials of individual Fellow and Associate designations signify that the individual has gained knowledge of the fundamental concepts. One is magnanimous and proficient in an expertise course covering the professional code of conduct, and follows recognized standards of practice.

Open Association of Research Society (US)/ Global Journals Incorporation (USA), as described in Corporate Statements, are educational, research publishing and professional membership organizations. Achieving our individual Fellow or Associate status is based mainly on meeting stated educational research requirements.

Disbursement of 40% Royalty earned through Global Journals : Researcher = 50%, Peer Reviewer = 37.50%, Institution = 12.50% E.g. Out of 40%, the 20% benefit should be passed on to researcher, 15 % benefit towards remuneration should be given to a reviewer and remaining 5% is to be retained by the institution.

We shall provide print version of 12 issues of any three journals [as per your requirement] out of our 38 journals worth $ 2376 USD.

Other:

The individual Fellow and Associate designations accredited by Open Association of Research Society (US) credentials signify guarantees following achievements:

- The professional accredited with Fellow honor, is entitled to various benefits viz. name, fame, honor, regular flow of income, secured bright future, social status etc.

- In addition to above, if one is single author, then entitled to 40% discount on publishing research paper and can get 10%discount if one is co-author or main author among group of authors.
- The Fellow can organize symposium/seminar/conference on behalf of Global Journals Incorporation (USA) and he/she can also attend the same organized by other institutes on behalf of Global Journals.
- The Fellow can become member of Editorial Board Member after completing 3yrs.
- The Fellow can earn 60% of sales proceeds from the sale of reference/review books/literature/publishing of research paper.
- Fellow can also join as paid peer reviewer and earn 15% remuneration of author charges and can also get an opportunity to join as member of the Editorial Board of Global Journals Incorporation (USA)
- • This individual has learned the basic methods of applying those concepts and techniques to common challenging situations. This individual has further demonstrated an in–depth understanding of the application of suitable techniques to a particular area of research practice.

Note :

"

- In future, if the board feels the necessity to change any board member, the same can be done with the consent of the chairperson along with anyone board member without our approval.
- In case, the chairperson needs to be replaced then consent of 2/3rd board members are required and they are also required to jointly pass the resolution copy of which should be sent to us. In such case, it will be compulsory to obtain our approval before replacement.
- In case of "Difference of Opinion [if any]" among the Board members, our decision will be final and binding to everyone. "

Preferred Author Guidelines

We accept the manuscript submissions in any standard (generic) format.

We typeset manuscripts using advanced typesetting tools like Adobe In Design, CorelDraw, TeXnicCenter, and TeXStudio. We usually recommend authors submit their research using any standard format they are comfortable with, and let Global Journals do the rest.

Alternatively, you can download our basic template from https://globaljournals.org/Template.zip

Authors should submit their complete paper/article, including text illustrations, graphics, conclusions, artwork, and tables. Authors who are not able to submit manuscript using the form above can email the manuscript department at submit@globaljournals.org or get in touch with chiefeditor@globaljournals.org if they wish to send the abstract before submission.

Before and during Submission

Authors must ensure the information provided during the submission of a paper is authentic. Please go through the following checklist before submitting:

1. Authors must go through the complete author guideline and understand and *agree to Global Journals' ethics and code of conduct,* along with author responsibilities.
2. Authors must accept the privacy policy, terms, and conditions of Global Journals.
3. Ensure corresponding author's email address and postal address are accurate and reachable.
4. Manuscript to be submitted must include keywords, an abstract, a paper title, co-author(s') names and details (email address, name, phone number, and institution), figures and illustrations in vector format including appropriate captions, tables, including titles and footnotes, a conclusion, results, acknowledgments and references.
5. Authors should submit paper in a ZIP archive if any supplementary files are required along with the paper.
6. Proper permissions must be acquired for the use of any copyrighted material.
7. Manuscript submitted *must not have been submitted or published elsewhere* and all authors must be aware of the submission.

Declaration of Conflicts of Interest

It is required for authors to declare all financial, institutional, and personal relationships with other individuals and organizations that could influence (bias) their research.

Policy on Plagiarism

Plagiarism is not acceptable in Global Journals submissions at all.

Plagiarized content will not be considered for publication. We reserve the right to inform authors' institutions about plagiarism detected either before or after publication. If plagiarism is identified, we will follow COPE guidelines:

Authors are solely responsible for all the plagiarism that is found. The author must not fabricate, falsify or plagiarize existing research data. The following, if copied, will be considered plagiarism:

- Words (language)
- Ideas
- Findings
- Writings
- Diagrams
- Graphs
- Illustrations
- Lectures

- Printed material
- Graphic representations
- Computer programs
- Electronic material
- Any other original work

Authorship Policies

Global Journals follows the definition of authorship set up by the Open Association of Research Society, USA. According to its guidelines, authorship criteria must be based on:

1. Substantial contributions to the conception and acquisition of data, analysis, and interpretation of findings.
2. Drafting the paper and revising it critically regarding important academic content.
3. Final approval of the version of the paper to be published.

Changes in Authorship

The corresponding author should mention the name and complete details of all co-authors during submission and in manuscript. We support addition, rearrangement, manipulation, and deletions in authors list till the early view publication of the journal. We expect that corresponding author will notify all co-authors of submission. We follow COPE guidelines for changes in authorship.

Copyright

During submission of the manuscript, the author is confirming an exclusive license agreement with Global Journals which gives Global Journals the authority to reproduce, reuse, and republish authors' research. We also believe in flexible copyright terms where copyright may remain with authors/employers/institutions as well. Contact your editor after acceptance to choose your copyright policy. You may follow this form for copyright transfers.

Appealing Decisions

Unless specified in the notification, the Editorial Board’s decision on publication of the paper is final and cannot be appealed before making the major change in the manuscript.

Acknowledgments

Contributors to the research other than authors credited should be mentioned in Acknowledgments. The source of funding for the research can be included. Suppliers of resources may be mentioned along with their addresses.

Declaration of funding sources

Global Journals is in partnership with various universities, laboratories, and other institutions worldwide in the research domain. Authors are requested to disclose their source of funding during every stage of their research, such as making analysis, performing laboratory operations, computing data, and using institutional resources, from writing an article to its submission. This will also help authors to get reimbursements by requesting an open access publication letter from Global Journals and submitting to the respective funding source.

Preparing your Manuscript

Authors can submit papers and articles in an acceptable file format: MS Word (doc, docx), LaTeX (.tex, .zip or .rar including all of your files), Adobe PDF (.pdf), rich text format (.rtf), simple text document (.txt), Open Document Text (.odt), and Apple Pages (.pages). Our professional layout editors will format the entire paper according to our official guidelines. This is one of the highlights of publishing with Global Journals—authors should not be concerned about the formatting of their paper. Global Journals accepts articles and manuscripts in every major language, be it Spanish, Chinese, Japanese, Portuguese, Russian, French, German, Dutch, Italian, Greek, or any other national language, but the title, subtitle, and abstract should be in English. This will facilitate indexing and the pre-peer review process.

The following is the official style and template developed for publication of a research paper. Authors are not required to follow this style during the submission of the paper. It is just for reference purposes.

Manuscript Style Instruction (Optional)

- Microsoft Word Document Setting Instructions.
- Font type of all text should be Swis721 Lt BT.
- Page size: 8.27" x 11'", left margin: 0.65, right margin: 0.65, bottom margin: 0.75.
- Paper title should be in one column of font size 24.
- Author name in font size of 11 in one column.
- Abstract: font size 9 with the word "Abstract" in bold italics.
- Main text: font size 10 with two justified columns.
- Two columns with equal column width of 3.38 and spacing of 0.2.
- First character must be three lines drop-capped.
- The paragraph before spacing of 1 pt and after of 0 pt.
- Line spacing of 1 pt.
- Large images must be in one column.
- The names of first main headings (Heading 1) must be in Roman font, capital letters, and font size of 10.
- The names of second main headings (Heading 2) must not include numbers and must be in italics with a font size of 10.

Structure and Format of Manuscript

The recommended size of an original research paper is under 15,000 words and review papers under 7,000 words. Research articles should be less than 10,000 words. Research papers are usually longer than review papers. Review papers are reports of significant research (typically less than 7,000 words, including tables, figures, and references)

A research paper must include:

a) A title which should be relevant to the theme of the paper.
b) A summary, known as an abstract (less than 150 words), containing the major results and conclusions.
c) Up to 10 keywords that precisely identify the paper's subject, purpose, and focus.
d) An introduction, giving fundamental background objectives.
e) Resources and techniques with sufficient complete experimental details (wherever possible by reference) to permit repetition, sources of information must be given, and numerical methods must be specified by reference.
f) Results which should be presented concisely by well-designed tables and figures.
g) Suitable statistical data should also be given.
h) All data must have been gathered with attention to numerical detail in the planning stage.

Design has been recognized to be essential to experiments for a considerable time, and the editor has decided that any paper that appears not to have adequate numerical treatments of the data will be returned unrefereed.

i) Discussion should cover implications and consequences and not just recapitulate the results; conclusions should also be summarized.
j) There should be brief acknowledgments.
k) There ought to be references in the conventional format. Global Journals recommends APA format.

Authors should carefully consider the preparation of papers to ensure that they communicate effectively. Papers are much more likely to be accepted if they are carefully designed and laid out, contain few or no errors, are summarizing, and follow instructions. They will also be published with much fewer delays than those that require much technical and editorial correction.

The Editorial Board reserves the right to make literary corrections and suggestions to improve brevity.

Format Structure

It is necessary that authors take care in submitting a manuscript that is written in simple language and adheres to published guidelines.

All manuscripts submitted to Global Journals should include:

Title

The title page must carry an informative title that reflects the content, a running title (less than 45 characters together with spaces), names of the authors and co-authors, and the place(s) where the work was carried out.

Author details

The full postal address of any related author(s) must be specified.

Abstract

The abstract is the foundation of the research paper. It should be clear and concise and must contain the objective of the paper and inferences drawn. It is advised to not include big mathematical equations or complicated jargon.

Many researchers searching for information online will use search engines such as Google, Yahoo or others. By optimizing your paper for search engines, you will amplify the chance of someone finding it. In turn, this will make it more likely to be viewed and cited in further works. Global Journals has compiled these guidelines to facilitate you to maximize the web-friendliness of the most public part of your paper.

Keywords

A major lynchpin of research work for the writing of research papers is the keyword search, which one will employ to find both library and internet resources. Up to eleven keywords or very brief phrases have to be given to help data retrieval, mining, and indexing.

One must be persistent and creative in using keywords. An effective keyword search requires a strategy: planning of a list of possible keywords and phrases to try.

Choice of the main keywords is the first tool of writing a research paper. Research paper writing is an art. Keyword search should be as strategic as possible.

One should start brainstorming lists of potential keywords before even beginning searching. Think about the most important concepts related to research work. Ask, "What words would a source have to include to be truly valuable in a research paper?" Then consider synonyms for the important words.

It may take the discovery of only one important paper to steer in the right keyword direction because, in most databases, the keywords under which a research paper is abstracted are listed with the paper.

Numerical Methods

Numerical methods used should be transparent and, where appropriate, supported by references.

Abbreviations

Authors must list all the abbreviations used in the paper at the end of the paper or in a separate table before using them.

Formulas and equations

Authors are advised to submit any mathematical equation using either MathJax, KaTeX, or LaTeX, or in a very high-quality image.

Tables, Figures, and Figure Legends

Tables: Tables should be cautiously designed, uncrowned, and include only essential data. Each must have an Arabic number, e.g., Table 4, a self-explanatory caption, and be on a separate sheet. Authors must submit tables in an editable format and not as images. References to these tables (if any) must be mentioned accurately.

Figures

Figures are supposed to be submitted as separate files. Always include a citation in the text for each figure using Arabic numbers, e.g., Fig. 4. Artwork must be submitted online in vector electronic form or by emailing it.

Preparation of Eletronic Figures for Publication

Although low-quality images are sufficient for review purposes, print publication requires high-quality images to prevent the final product being blurred or fuzzy. Submit (possibly by e-mail) EPS (line art) or TIFF (halftone/ photographs) files only. MS PowerPoint and Word Graphics are unsuitable for printed pictures. Avoid using pixel-oriented software. Scans (TIFF only) should have a resolution of at least 350 dpi (halftone) or 700 to 1100 dpi (line drawings). Please give the data for figures in black and white or submit a Color Work Agreement form. EPS files must be saved with fonts embedded (and with a TIFF preview, if possible).

For scanned images, the scanning resolution at final image size ought to be as follows to ensure good reproduction: line art: >650 dpi; halftones (including gel photographs): >350 dpi; figures containing both halftone and line images: >650 dpi.

Color charges: Authors are advised to pay the full cost for the reproduction of their color artwork. Hence, please note that if there is color artwork in your manuscript when it is accepted for publication, we would require you to complete and return a Color Work Agreement form before your paper can be published. Also, you can email your editor to remove the color fee after acceptance of the paper.

Tips for Writing a Good Quality Science Frontier Research Paper

Techniques for writing a good quality Science Frontier Research paper:

1. Choosing the topic: In most cases, the topic is selected by the interests of the author, but it can also be suggested by the guides. You can have several topics, and then judge which you are most comfortable with. This may be done by asking several questions of yourself, like "Will I be able to carry out a search in this area? Will I find all necessary resources to accomplish the search? Will I be able to find all information in this field area?" If the answer to this type of question is "yes," then you ought to choose that topic. In most cases, you may have to conduct surveys and visit several places. Also, you might have to do a lot of work to find all the rises and falls of the various data on that subject. Sometimes, detailed information plays a vital role, instead of short information. Evaluators are human: The first thing to remember is that evaluators are also human beings. They are not only meant for rejecting a paper. They are here to evaluate your paper. So present your best aspect.

2. Think like evaluators: If you are in confusion or getting demotivated because your paper may not be accepted by the evaluators, then think, and try to evaluate your paper like an evaluator. Try to understand what an evaluator wants in your research paper, and you will automatically have your answer. Make blueprints of paper: The outline is the plan or framework that will help you to arrange your thoughts. It will make your paper logical. But remember that all points of your outline must be related to the topic you have chosen.

3. Ask your guides: If you are having any difficulty with your research, then do not hesitate to share your difficulty with your guide (if you have one). They will surely help you out and resolve your doubts. If you can't clarify what exactly you require for your work, then ask your supervisor to help you with an alternative. He or she might also provide you with a list of essential readings.

4. Use of computer is recommended: As you are doing research in the field of science frontier then this point is quite obvious. Use right software: Always use good quality software packages. If you are not capable of judging good software, then you can lose the quality of your paper unknowingly. There are various programs available to help you which you can get through the internet.

5. Use the internet for help: An excellent start for your paper is using Google. It is a wondrous search engine, where you can have your doubts resolved. You may also read some answers for the frequent question of how to write your research paper or find a model research paper. You can download books from the internet. If you have all the required books, place importance on reading, selecting, and analyzing the specified information. Then sketch out your research paper. Use big pictures: You may use encyclopedias like Wikipedia to get pictures with the best resolution. At Global Journals, you should strictly follow here.

6. Bookmarks are useful: When you read any book or magazine, you generally use bookmarks, right? It is a good habit which helps to not lose your continuity. You should always use bookmarks while searching on the internet also, which will make your search easier.

7. Revise what you wrote: When you write anything, always read it, summarize it, and then finalize it.

8. Make every effort: Make every effort to mention what you are going to write in your paper. That means always have a good start. Try to mention everything in the introduction—what is the need for a particular research paper. Polish your work with good writing skills and always give an evaluator what he wants. Make backups: When you are going to do any important thing like making a research paper, you should always have backup copies of it either on your computer or on paper. This protects you from losing any portion of your important data.

9. Produce good diagrams of your own: Always try to include good charts or diagrams in your paper to improve quality. Using several unnecessary diagrams will degrade the quality of your paper by creating a hodgepodge. So always try to include diagrams which were made by you to improve the readability of your paper. Use of direct quotes: When you do research relevant to literature, history, or current affairs, then use of quotes becomes essential, but if the study is relevant to science, use of quotes is not preferable.

10. Use proper verb tense: Use proper verb tenses in your paper. Use past tense to present those events that have happened. Use present tense to indicate events that are going on. Use future tense to indicate events that will happen in the future. Use of wrong tenses will confuse the evaluator. Avoid sentences that are incomplete.

11. Pick a good study spot: Always try to pick a spot for your research which is quiet. Not every spot is good for studying.

12. Know what you know: Always try to know what you know by making objectives, otherwise you will be confused and unable to achieve your target.

13. Use good grammar: Always use good grammar and words that will have a positive impact on the evaluator; use of good vocabulary does not mean using tough words which the evaluator has to find in a dictionary. Do not fragment sentences. Eliminate one-word sentences. Do not ever use a big word when a smaller one would suffice.

Verbs have to be in agreement with their subjects. In a research paper, do not start sentences with conjunctions or finish them with prepositions. When writing formally, it is advisable to never split an infinitive because someone will (wrongly) complain. Avoid clichés like a disease. Always shun irritating alliteration. Use language which is simple and straightforward. Put together a neat summary.

14. Arrangement of information: Each section of the main body should start with an opening sentence, and there should be a changeover at the end of the section. Give only valid and powerful arguments for your topic. You may also maintain your arguments with records.

15. Never start at the last minute: Always allow enough time for research work. Leaving everything to the last minute will degrade your paper and spoil your work.

16. Multitasking in research is not good: Doing several things at the same time is a bad habit in the case of research activity. Research is an area where everything has a particular time slot. Divide your research work into parts, and do a particular part in a particular time slot.

17. Never copy others' work: Never copy others' work and give it your name because if the evaluator has seen it anywhere, you will be in trouble. Take proper rest and food: No matter how many hours you spend on your research activity, if you are not taking care of your health, then all your efforts will have been in vain. For quality research, take proper rest and food.

18. Go to seminars: Attend seminars if the topic is relevant to your research area. Utilize all your resources.

19. Refresh your mind after intervals: Try to give your mind a rest by listening to soft music or sleeping in intervals. This will also improve your memory. Acquire colleagues: Always try to acquire colleagues. No matter how sharp you are, if you acquire colleagues, they can give you ideas which will be helpful to your research.

20. Think technically: Always think technically. If anything happens, search for its reasons, benefits, and demerits. Think and then print: When you go to print your paper, check that tables are not split, headings are not detached from their descriptions, and page sequence is maintained.

21. Adding unnecessary information: Do not add unnecessary information like "I have used MS Excel to draw graphs." Irrelevant and inappropriate material is superfluous. Foreign terminology and phrases are not apropos. One should never take a broad view. Analogy is like feathers on a snake. Use words properly, regardless of how others use them. Remove quotations. Puns are for kids, not grunt readers. Never oversimplify: When adding material to your research paper, never go for oversimplification; this will definitely irritate the evaluator. Be specific. Never use rhythmic redundancies. Contractions shouldn't be used in a research paper. Comparisons are as terrible as clichés. Give up ampersands, abbreviations, and so on. Remove commas that are not necessary. Parenthetical words should be between brackets or commas. Understatement is always the best way to put forward earth-shaking thoughts. Give a detailed literary review.

22. Report concluded results: Use concluded results. From raw data, filter the results, and then conclude your studies based on measurements and observations taken. An appropriate number of decimal places should be used. Parenthetical remarks are prohibited here. Proofread carefully at the final stage. At the end, give an outline to your arguments. Spot perspectives of further study of the subject. Justify your conclusion at the bottom sufficiently, which will probably include examples.

23. Upon conclusion: Once you have concluded your research, the next most important step is to present your findings. Presentation is extremely important as it is the definite medium though which your research is going to be in print for the rest of the crowd. Care should be taken to categorize your thoughts well and present them in a logical and neat manner. A good quality research paper format is essential because it serves to highlight your research paper and bring to light all necessary aspects of your research.

Informal Guidelines of Research Paper Writing

Key points to remember:

- Submit all work in its final form.
- Write your paper in the form which is presented in the guidelines using the template.
- Please note the criteria peer reviewers will use for grading the final paper.

Final points:

One purpose of organizing a research paper is to let people interpret your efforts selectively. The journal requires the following sections, submitted in the order listed, with each section starting on a new page:

The introduction: This will be compiled from reference matter and reflect the design processes or outline of basis that directed you to make a study. As you carry out the process of study, the method and process section will be constructed like that. The results segment will show related statistics in nearly sequential order and direct reviewers to similar intellectual paths throughout the data that you gathered to carry out your study.

The discussion section:

This will provide understanding of the data and projections as to the implications of the results. The use of good quality references throughout the paper will give the effort trustworthiness by representing an alertness to prior workings.

Writing a research paper is not an easy job, no matter how trouble-free the actual research or concept. Practice, excellent preparation, and controlled record-keeping are the only means to make straightforward progression.

General style:

Specific editorial column necessities for compliance of a manuscript will always take over from directions in these general guidelines.

To make a paper clear: Adhere to recommended page limits.

Mistakes to avoid:

- Insertion of a title at the foot of a page with subsequent text on the next page.
- Separating a table, chart, or figure—confine each to a single page.
- Submitting a manuscript with pages out of sequence.
- In every section of your document, use standard writing style, including articles ("a" and "the").
- Keep paying attention to the topic of the paper.
- Use paragraphs to split each significant point (excluding the abstract).
- Align the primary line of each section.
- Present your points in sound order.
- Use present tense to report well-accepted matters.
- Use past tense to describe specific results.
- Do not use familiar wording; don't address the reviewer directly. Don't use slang or superlatives.
- Avoid use of extra pictures—include only those figures essential to presenting results.

Title page:

Choose a revealing title. It should be short and include the name(s) and address(es) of all authors. It should not have acronyms or abbreviations or exceed two printed lines.

Abstract: This summary should be two hundred words or less. It should clearly and briefly explain the key findings reported in the manuscript and must have precise statistics. It should not have acronyms or abbreviations. It should be logical in itself. Do not cite references at this point.

An abstract is a brief, distinct paragraph summary of finished work or work in development. In a minute or less, a reviewer can be taught the foundation behind the study, common approaches to the problem, relevant results, and significant conclusions or new questions.

Write your summary when your paper is completed because how can you write the summary of anything which is not yet written? Wealth of terminology is very essential in abstract. Use comprehensive sentences, and do not sacrifice readability for brevity; you can maintain it succinctly by phrasing sentences so that they provide more than a lone rationale. The author can at this moment go straight to shortening the outcome. Sum up the study with the subsequent elements in any summary. Try to limit the initial two items to no more than one line each.

Reason for writing the article—theory, overall issue, purpose.

- Fundamental goal.
- To-the-point depiction of the research.
- Consequences, including definite statistics—if the consequences are quantitative in nature, account for this; results of any numerical analysis should be reported. Significant conclusions or questions that emerge from the research.

Approach:

- o Single section and succinct.
- o An outline of the job done is always written in past tense.
- o Concentrate on shortening results—limit background information to a verdict or two.
- o Exact spelling, clarity of sentences and phrases, and appropriate reporting of quantities (proper units, important statistics) are just as significant in an abstract as they are anywhere else.

Introduction:

The introduction should "introduce" the manuscript. The reviewer should be presented with sufficient background information to be capable of comprehending and calculating the purpose of your study without having to refer to other works. The basis for the study should be offered. Give the most important references, but avoid making a comprehensive appraisal of the topic. Describe the problem visibly. If the problem is not acknowledged in a logical, reasonable way, the reviewer will give no attention to your results. Speak in common terms about techniques used to explain the problem, if needed, but do not present any particulars about the protocols here.

The following approach can create a valuable beginning:

- o Explain the value (significance) of the study.
- o Defend the model—why did you employ this particular system or method? What is its compensation? Remark upon its appropriateness from an abstract point of view as well as pointing out sensible reasons for using it.
- o Present a justification. State your particular theory(-ies) or aim(s), and describe the logic that led you to choose them.
- o Briefly explain the study's tentative purpose and how it meets the declared objectives.

Approach:

Use past tense except for when referring to recognized facts. After all, the manuscript will be submitted after the entire job is done. Sort out your thoughts; manufacture one key point for every section. If you make the four points listed above, you will need at least four paragraphs. Present surrounding information only when it is necessary to support a situation. The reviewer does not desire to read everything you know about a topic. Shape the theory specifically—do not take a broad view.

As always, give awareness to spelling, simplicity, and correctness of sentences and phrases.

Procedures (methods and materials):

This part is supposed to be the easiest to carve if you have good skills. A soundly written procedures segment allows a capable scientist to replicate your results. Present precise information about your supplies. The suppliers and clarity of reagents can be helpful bits of information. Present methods in sequential order, but linked methodologies can be grouped as a segment. Be concise when relating the protocols. Attempt to give the least amount of information that would permit another capable scientist to replicate your outcome, but be cautious that vital information is integrated. The use of subheadings is suggested and ought to be synchronized with the results section.

When a technique is used that has been well-described in another section, mention the specific item describing the way, but draw the basic principle while stating the situation. The purpose is to show all particular resources and broad procedures so that another person may use some or all of the methods in one more study or referee the scientific value of your work. It is not to be a step-by-step report of the whole thing you did, nor is a methods section a set of orders.

Materials:

Materials may be reported in part of a section or else they may be recognized along with your measures.

Methods:

- o Report the method and not the particulars of each process that engaged the same methodology.
- o Describe the method entirely.
- o To be succinct, present methods under headings dedicated to specific dealings or groups of measures.
- o Simplify—detail how procedures were completed, not how they were performed on a particular day.
- o If well-known procedures were used, account for the procedure by name, possibly with a reference, and that's all.

Approach:

It is embarrassing to use vigorous voice when documenting methods without using first person, which would focus the reviewer's interest on the researcher rather than the job. As a result, when writing up the methods, most authors use third person passive voice.

Use standard style in this and every other part of the paper—avoid familiar lists, and use full sentences.

What to keep away from:

- o Resources and methods are not a set of information.
- o Skip all descriptive information and surroundings—save it for the argument.
- o Leave out information that is immaterial to a third party.

Results:

The principle of a results segment is to present and demonstrate your conclusion. Create this part as entirely objective details of the outcome, and save all understanding for the discussion.

The page length of this segment is set by the sum and types of data to be reported. Use statistics and tables, if suitable, to present consequences most efficiently.

You must clearly differentiate material which would usually be incorporated in a study editorial from any unprocessed data or additional appendix matter that would not be available. In fact, such matters should not be submitted at all except if requested by the instructor.

Content:

- Sum up your conclusions in text and demonstrate them, if suitable, with figures and tables.
- In the manuscript, explain each of your consequences, and point the reader to remarks that are most appropriate.
- Present a background, such as by describing the question that was addressed by creation of an exacting study.
- Explain results of control experiments and give remarks that are not accessible in a prescribed figure or table, if appropriate.
- Examine your data, then prepare the analyzed (transformed) data in the form of a figure (graph), table, or manuscript.

What to stay away from:

- Do not discuss or infer your outcome, report surrounding information, or try to explain anything.
- Do not include raw data or intermediate calculations in a research manuscript.
- Do not present similar data more than once.
- A manuscript should complement any figures or tables, not duplicate information.
- Never confuse figures with tables—there is a difference.

Approach:

As always, use past tense when you submit your results, and put the whole thing in a reasonable order.

Put figures and tables, appropriately numbered, in order at the end of the report.

If you desire, you may place your figures and tables properly within the text of your results section.

Figures and tables:

If you put figures and tables at the end of some details, make certain that they are visibly distinguished from any attached appendix materials, such as raw facts. Whatever the position, each table must be titled, numbered one after the other, and include a heading. All figures and tables must be divided from the text.

Discussion:

The discussion is expected to be the trickiest segment to write. A lot of papers submitted to the journal are discarded based on problems with the discussion. There is no rule for how long an argument should be.

Position your understanding of the outcome visibly to lead the reviewer through your conclusions, and then finish the paper with a summing up of the implications of the study. The purpose here is to offer an understanding of your results and support all of your conclusions, using facts from your research and generally accepted information, if suitable. The implication of results should be fully described.

Infer your data in the conversation in suitable depth. This means that when you clarify an observable fact, you must explain mechanisms that may account for the observation. If your results vary from your prospect, make clear why that may have happened. If your results agree, then explain the theory that the proof supported. It is never suitable to just state that the data approved the prospect, and let it drop at that. Make a decision as to whether each premise is supported or discarded or if you cannot make a conclusion with assurance. Do not just dismiss a study or part of a study as "uncertain."

Research papers are not acknowledged if the work is imperfect. Draw what conclusions you can based upon the results that you have, and take care of the study as a finished work.

- You may propose future guidelines, such as how an experiment might be personalized to accomplish a new idea.
- Give details of all of your remarks as much as possible, focusing on mechanisms.
- Make a decision as to whether the tentative design sufficiently addressed the theory and whether or not it was correctly restricted. Try to present substitute explanations if they are sensible alternatives.
- One piece of research will not counter an overall question, so maintain the large picture in mind. Where do you go next? The best studies unlock new avenues of study. What questions remain?
- Recommendations for detailed papers will offer supplementary suggestions.

Approach:

When you refer to information, differentiate data generated by your own studies from other available information. Present work done by specific persons (including you) in past tense.

Describe generally acknowledged facts and main beliefs in present tense.

The Administration Rules

Administration Rules to Be Strictly Followed before Submitting Your Research Paper to Global Journals Inc.

Please read the following rules and regulations carefully before submitting your research paper to Global Journals Inc. to avoid rejection.

Segment draft and final research paper: You have to strictly follow the template of a research paper, failing which your paper may get rejected. You are expected to write each part of the paper wholly on your own. The peer reviewers need to identify your own perspective of the concepts in your own terms. Please do not extract straight from any other source, and do not rephrase someone else's analysis. Do not allow anyone else to proofread your manuscript.

Written material: You may discuss this with your guides and key sources. Do not copy anyone else's paper, even if this is only imitation, otherwise it will be rejected on the grounds of plagiarism, which is illegal. Various methods to avoid plagiarism are strictly applied by us to every paper, and, if found guilty, you may be blacklisted, which could affect your career adversely. To guard yourself and others from possible illegal use, please do not permit anyone to use or even read your paper and file.

Criterion for Grading a Research Paper (Compilation) by Global Journals

Please note that following table is only a Grading of "Paper Compilation" and not on "Performed/Stated Research" whose grading solely depends on Individual Assigned Peer Reviewer and Editorial Board Member. These can be available only on request and after decision of Paper. This report will be the property of Global Journals.

Topics	Grades		
	A-B	**C-D**	**E-F**
Abstract	Clear and concise with appropriate content, Correct format. 200 words or below	Unclear summary and no specific data, Incorrect form Above 200 words	No specific data with ambiguous information Above 250 words
Introduction	Containing all background details with clear goal and appropriate details, flow specification, no grammar and spelling mistake, well organized sentence and paragraph, reference cited	Unclear and confusing data, appropriate format, grammar and spelling errors with unorganized matter	Out of place depth and content, hazy format
Methods and Procedures	Clear and to the point with well arranged paragraph, precision and accuracy of facts and figures, well organized subheads	Difficult to comprehend with embarrassed text, too much explanation but completed	Incorrect and unorganized structure with hazy meaning
Result	Well organized, Clear and specific, Correct units with precision, correct data, well structuring of paragraph, no grammar and spelling mistake	Complete and embarrassed text, difficult to comprehend	Irregular format with wrong facts and figures
Discussion	Well organized, meaningful specification, sound conclusion, logical and concise explanation, highly structured paragraph reference cited	Wordy, unclear conclusion, spurious	Conclusion is not cited, unorganized, difficult to comprehend
References	Complete and correct format, well organized	Beside the point, Incomplete	Wrong format and structuring

Index

save our planet

Global Journal of Science Frontier Research

Visit us on the Web at www.GlobalJournals.org | www.JournalofScience.org
or email us at helpdesk@globaljournals.org

ISSN 9755896