SHIVAJI UNIVERSITY, KOLHAPUR

Estd. 1962
"A++" Accredited by NAAC (2021)
with CGPA 3.52

JOURNAL OF SHIVAJI UNIVERSITY : SCIENCE AND TECHNOLOGY

(Peer Reviewed Journal)

Volume-45, Issue-1 (January, 2019)

ISSN Science-0250-5347

JOURNAL OF SHIVAJI UNIVERSITY : SCIENCE AND TECHNOLOGY

(Peer Reviewed Journal)

Volume-45, Issue-1 (January, 2019)

ISSN Science -0250-5347

Estd. 1962
"A++" Accredited by NAAC (2021)
with CGPA 3.52

Shivaji University
Kolhapur 416 004 (INDIA)

JOURNAL OF SHIVAJI UNIVERSITY : SCIENCE AND TECHNOLOGY

(Peer Reviewed Journal)

Volume-45, Issue-1 (January, 2019)

ISSN Science -0250-5347

- Contact Details -
Prof. (Dr.) S. D. Delekar
Editor, Journal of Shivaji University : Science and Technology
and
Department of Chemistry
Shivaji University, Kolhapur 416 004 (India)
E-mail : editor.jscitech@unishivaji.ac.in
website : http://www.unishivaji.ac.in/journals/

Shivaji University, Kolhapur

JOURNAL OF SHIVAJI UNIVERSITY : SCIENCE AND TECHNOLOGY

(Peer Reviewed Journal)

Volume-45, Issue-1 (January, 2019)

ISSN Science -0250-5347

- Editor -

Prof. (Dr.) S. D. Delekar

Journal of Shivaji University :

Science and Technology

- Published by -

Dr. V. N. Shinde

Ag. Registrar,

Shivaji University, Kolhapur 416004.

- Printed by -

Shri. B. P. Patil

Superintend

Shivaji University, Kolhapur 416004.

The Editor, on behalf of the Editorial Board of the Journal of Shivaji University : Science and Technology, Vol. No. 45 (1) whises to express his thanks to the contributing authors and the experts for acting as referees for the papers included in the issue.

Published : July, 2023

Journal of Shivaji University: Science and Technology
Volume-45, Issue-1, (January, 2019)

INDEX

Sr. No.	Title of Research Article with Name of Author/s	Page No.
1.	**Ballistics** Smita D. Naik	1
2.	**Numerical Solution of Non-linear Impulsive Differential Equation by Simpson's $\frac{1}{3}^{rd}$ Rule** Indrayani Y. Sutar, Sandhya D. Kadam	12
3.	**Recent Development In Sturm–Liouville Theory** Pratiksha Demanna, Sarita H. Thakar	21
4.	**Numerical Simulations of Burgers' Equation by Cubic B-spline Galerkin Finite Element Method** Sunil S. Kumbhar	32
5.	**A Study of Gender Inequality According to Various Aspects** Prasad V. Shinde, Amol M. Ghadge	48
6.	**Modified Variable Selection Criterion in Presence of Multicollinearity** Ashok V. Dorugade	57
7.	**A Review on Nuisance Parameter Free Inferential Procedures for Shape-Scale and Location-Scale Family of Distributions** Kiran P. Patil, Hemangi V. Kulkarni	70
8.	**The Study of ExponentiatedGumbel Distribution and Related Inference Through Simulation** Chandrakant S. Kakade	81
9.	**On the Performance of Different Robust Criterion Functions based M-Estimators and RM-Estimators in the presence of Multicollinearity and Outliers** Nileshkumar H. Jadhav, Dattatray N. Kashid	94

Ballistics

Smita D. Naik[a,*]

Armament Research & Development Establishment, Pashan Pune 411021 (MS) India.

*Corresponding author: insdnaik@yahoo.com

Dynamics is a branch of Applied Mathematics which is the study of moving bodies. It includes effect of different causes like forces and moments affecting the motion. Ballistics is related to the motion of the bodies which move very fast. It covers motion of all types of projectiles like bullets, gravity bombs, rockets etc. Ballistics is a branch of Dynamics and hence that of Applied Mathematics.

The word 'BALLISTICS' has come from ba'llein , a Greek word, which means to throw. It came from Ballista (Fig 1), a machine used to throw iron balls for damage. It is defined as the science that deals with the motion, behavior, and effects of projectiles: the science or art of designing and hurling projectiles so as to achieve a desired performance.

Figure-1. Ballista.

Ballistics is classified according to the means used as Gun ballistics, Rocket ballistics, Torpedo dynamics, Under-water ballistics, Wound ballistics, Space dynamics. Each is a subject in itself and further divided into subclasses (Fig 2) as Internal or interior ballistics, External or exterior ballistics, Terminal ballistics, Intermediate ballistics and Experimental ballistics. Internal Ballistics deals with the motion of the projectile during launch. The study involves initiation of motion of a projectile and factors affectingit. Motion of a projectile and forces/moments arising due to the medium like air or water is studied in External ballistics/underwater Ballistics. Terminal

Ballistics is the study at the target end. When a projectile hits the target, damage to the target and rojectile is part of Terminal Ballistics. In anti-personal role, when the target is a human being, the analysis of wounds is covered in wound Ballistics. Experimental Ballistics talks about the experiments to be conducted for all these sub-branches and the findings.

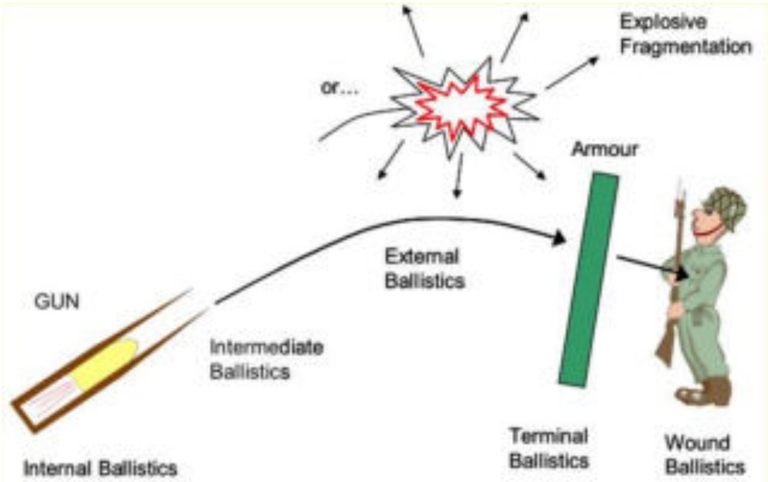


Figure-2. Classification of Ballistics

Ballistics study of a weapon system can be done with actual firing or with the modeling and simulation approach. Complete study is carried out with the help of five models: Gun Design, Target Definition, Gun Interior Ballistic, External Ballistic of Projectile and terminal Ballistic/ impact dynamics.

Target is defined in terms of its dimensions and strength. Target is classified as point target or area target depending on damage required. Damage criterion defines the critical points in Design of target.

Gun Design is defined to satisfy general requirements as, strong to meet the challenges of the enemy, capable of inflicting heavy damages to the target, easy to carry and handle and needs to be cost effective. Gun design model has input parameters as the design parameters of launcher/gun tube, projectile and propellant. Output required in terms of pressure inside the gun, velocity, range, drift, impact energy, damage to the defined target with the required accuracy. These are obtained with the help of Internal, External and Terminal Ballistics models.

To study **Internal Ballistics:** It is necessary to understand the processes taking place inside the gun which is called as Ballistic cycle (Fig.3). There are mainly two processes: Burning of the propellant and Motion of the Projectile (inside the barrel). These processes can be mathematically modeled with either lumped parameter or gas dynamics model. In lumped parameter model average properties of pressure, velocity and temperature are assumed and it results into pressure-space curve and velocity-space curve. Gas dynamics model gives the complete history of all properties with time and space. It is generally studied for boundary layer analysis of gas and solid phases and flame spread analysis.

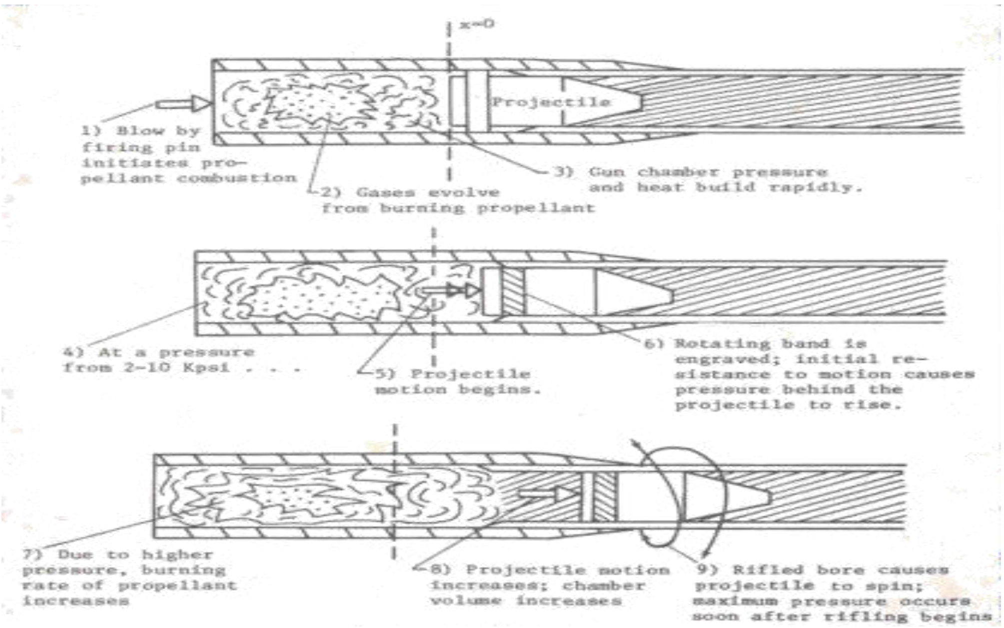


Figure-3. Ballistic cycle

Propellant is the source of energy provided to the system. Propellant consists of solid chemical grains which burn at a constant rate without the use of external oxygen. It follows Piobert's law of burning *i.e.* burning proceeds in parallel layers. Form function relation takes care of the shape and size parametersof the propellant defined in terms of form function constant. There are three types of burnings: degressive, progressive and neutral. It is related to form function constant.

Solution of the mathematical model gives pressure variation inside the barrel and muzzle velocity achieved by theprojectile (Fig.4).

The model consists of Variables – z, f , p, v, x ,TDesign Parameters:–

Propellant :θ , D, η , β , F , b , Te , ρ , Υ , cGun :– K_0 , A

Projectile :– m

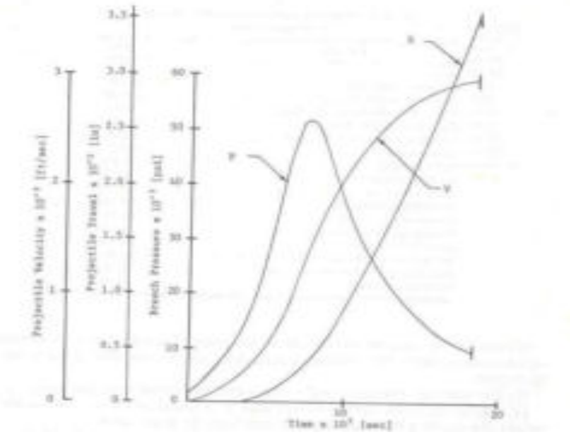


Figure-4. p-t, v-t curves

External ballistics is free flight dynamics of a projectile in a resisting medium (air). Initial flight conditions are governed by the projection and acceleration phase (Internal Ballistics). At the end of this phase projectile starts its uncontrolled flight with certain kinetic energy and attitude. The mathematical model is defined using Newton's Second Law with the external forces.

In External Ballistics study along with the design parameters few other aspects are also important.

These are:

1. Forces which influence the motion.
2. Stability of the projectile.
3. Trajectory modeling and analysis.

1. FORCES:

Forces which influence the motion of a projectile during the flight in air are – gravitational force, aerodynamic forces and forces due to rotation of earth.

1.1 GRAVITATIONAL FORCE

It is force of attraction between earth and projectile which creates pulling effect on the projectile towards centre of the earth. This effect produces acceleration denoted by 'g'. It varies inversely as the square of the distance from centre of earth. It is maximum at the pole and minimum at the equator.

1.2 **AERODYNAMIC FORCES**:

As a rigid body moves in the resisting medium, disturbances are created in the medium and in turn forces are generated which affect the motion. When the medium is air, these forces are called as aerodynamic forces. Eeffects responsible for the generation of aerodynamic forces are:

1. Viscous effect
2. Compressibility effect
3. Pressure effect

The Major aerodynamic forces and moments acting on a projectile in flight are of two types: Static and dynamic.

Static forces:

1. Drag D/axial force F_A - due to axial velocity
2. Lift L/normal force F_N - due to oblique motionStatic moments:

1. Over turning moment/yawing moment M - due to oblique motion
2. Spin driving moment - due to body asymmetries like canted finsDynamic forces:

1. Damping force S - due to cross spin
2. Magnus force

a. K - due to cross velocity and axial spin

b. Q - due to cross velocity (cross spin) and axial spin

Dynamic moments:

1. Spin damping moment I - due to axial spin
2. Magnus moment

a. T_K - due to cross velocity and axial spin

b. T_Q - due to cross velocity(cross spin) and axial spin

3. Damping moment H - due to cross spin

It is expressed as Aerodynamic force = $\frac{\Pi}{2} \rho v^2 d^2 C_F$. C_F is a constant called as Aerodynamic Coefficient. The aerodynamic coefficients are estimated using different ways like: Theoretic estimation using fluid dynamic theory, empirical estimation using data on similar projectile shapes, Wind tunnel testing, aeroballistic range testing or firing trials data.

Ballistic Coefficient- It is a constant defined from the design parameters of the projectile. It is also called as carrying capacity of a projectile.

1. Standard Ballistic Coefficient C_0

$$C_0 = \frac{m}{d^2 \, k_\sigma}$$

m – mass in lbs
d – caliber in inches
k_σ- shape and steadiness factor

2. Ballistic coefficient C

$$C = C_0 \, \frac{\rho_0}{\rho}$$

Retardation due to drag $R = \dfrac{\pi \, v^2 \rho \, C_D}{8C}$, $\quad C_D = C_{D_0} k_\sigma$

Larger the value of C, smaller is the retardation and in turn projectile covers more range. Remaining velocity of the projectile is also more for higher values of C. That is, a projectile having larger C can strike the target with higher velocity than a projectile having smaller value of C but fired with much higher muzzle velocity.

1.3 Forces due to rotation of Earth

For small ranges and low angle of launch, during trajectory computation generally Earth's rotation effect is ignored. For higher velocities and large angle trajectories it has to be included. The forces due to rotation ofEarth are

1. Centrifugal force- normal to Earth's axis
2. Coriolis force- It shifts the trajectory right which produces drift towards right side of the trajectory.

1.4 Thrust

Thrust is a force coming from within the rocket as a reaction to the burning of the propellant. It depends onthe mass flow rate of the propellant and the efflux velocity. It is expressed as

$$T = -V_{eff}\,\frac{dm}{dt}$$

The ability to adopt zero yaw attitude defines stable motion (Figure-5).

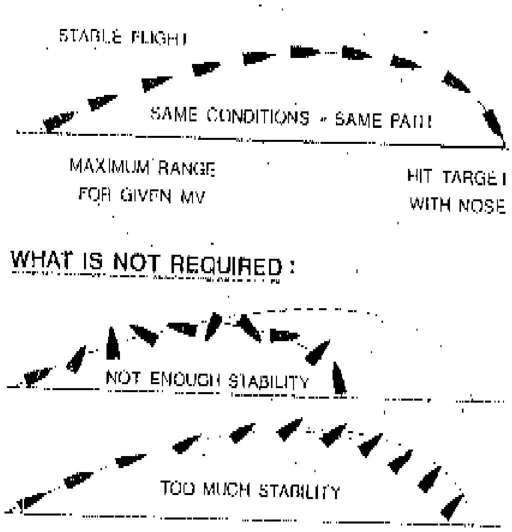


Figure-5. Stability of motion

Gun projectile is an axis symmetric body. Centre of pressure (C.P.) of static aerodynamic forces is ahead of centre of gravity (C.G.). Bodies having C.P. ahead of C.G. are statically unstable. The distance between C.P. and

C.G. is called static margin and is negative for statically unstable body. When C.P. is behind C.G., it is statically stable body and static margin is positive. Thus a gun projectile is statically unstable and has to be made stable during flight. The body is stabilized using spin or fins. Fins shift CP behind CG. Methods of stabilization:

1. Spin motion is imparted to projectile which makes projectile stable like spinning top-Gyroscopic or spin stabilization (Fig. 6(a)).
2. Mass of the projectile is so concentrated at the forward end as to move C.G. ahead of C.P. Projectile is provided with flat surfaces (fins) at the rear of the body-Aerodynamic or fin stabilization (Fig.6(b)).

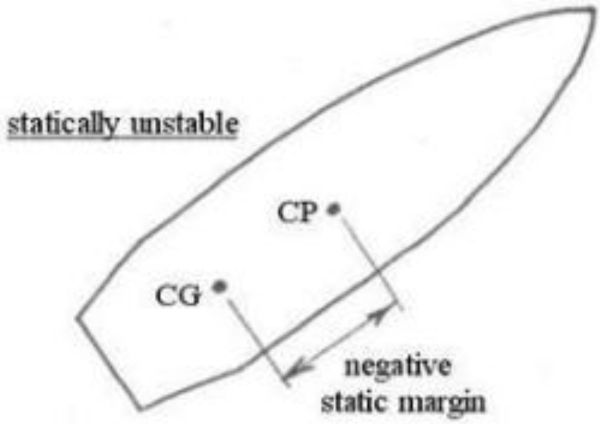


Figure-6(a). Spin stabilized projectile

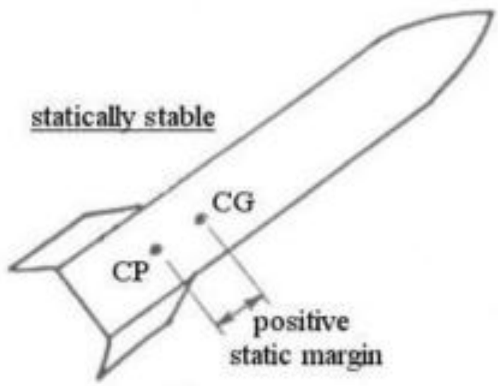


Figure-6(b). Fin stabilized projectile

Static stability relates to the initial response of a body when disturbed from equilibrium conditions.

The oscillatory motion damps out to minimum in short distance, then the projectile is dynamically stable (Fig. 7).

It relates to the time history of the subsequent motions following the initial response after being disturbed fromequilibrium conditions.

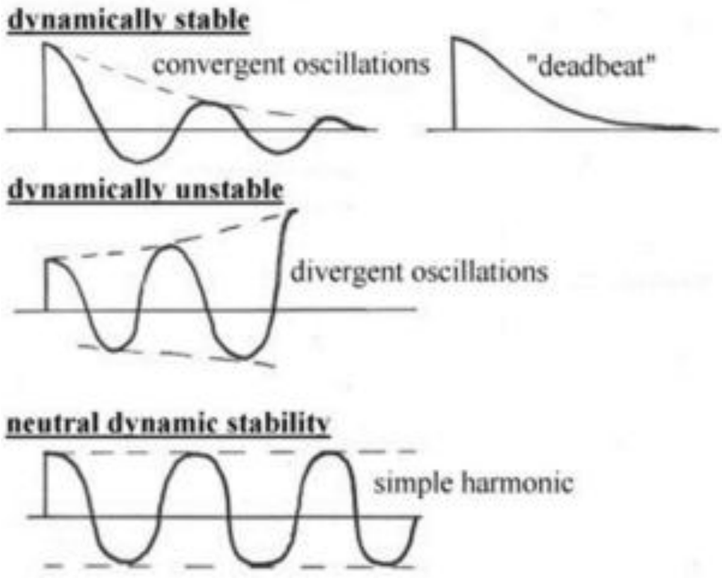


Figure-7. Stability during motion

3. TRAJECTORY

It is path taken by C.G. of the projectile. It gives knowledge of range, altitude, drift, remaining velocity, time of flight and slope. It is required to define proper frame of reference like Space fixed, earth fixed, body fixed and

The complete motion is studied with the help of 6-degree of freedom model, (6-DOF). Three scalar equations for linear motion (force equations) and Three for angular motion (moment equations). The mathematical model in the vector form is given

$$\frac{d\bar{F}}{dt} = \frac{\partial \bar{F}}{\partial t} + \bar{\Omega} \times \bar{F} = \sum \bar{F}$$

$$\frac{d\bar{H}}{dt} = \frac{\partial \bar{H}}{\partial t} + \bar{\Omega} \times \bar{H} = \sum \bar{H}$$

Important aspects considered in this study are stability of the projectile and trajectory computation. Factors affecting the trajectory and stability of the projectile during its motion are projectile parameters: mass, position of CP/CG, Moment of Inertia, Shape and surface design and forces and moments acting on the projectile.

Trajectory is computed using any of the models:

- ✓ *In vacuo* model
- ✓ Point mass model- Earth rotation effects, crosswind effects
- ✓ Modified point mass model- Equilibrium yaw, Magnus force/moment, spin damping couple

- ✓ 6 Degree of freedom model

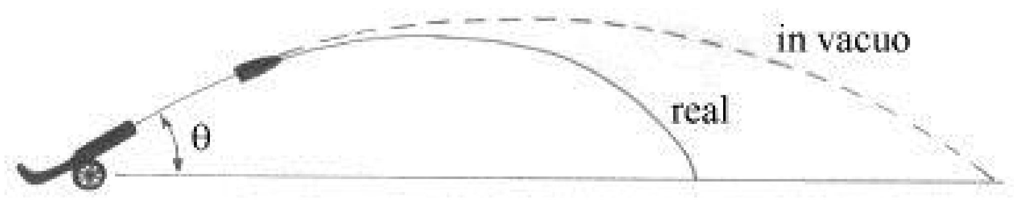


Figure-8. Trajectory of projectile.

Fig.8 consists of the trajectories obtained *in vacuo* and real which clearly shows that due to aerodynamic forces acting on the projectile the range reduces.

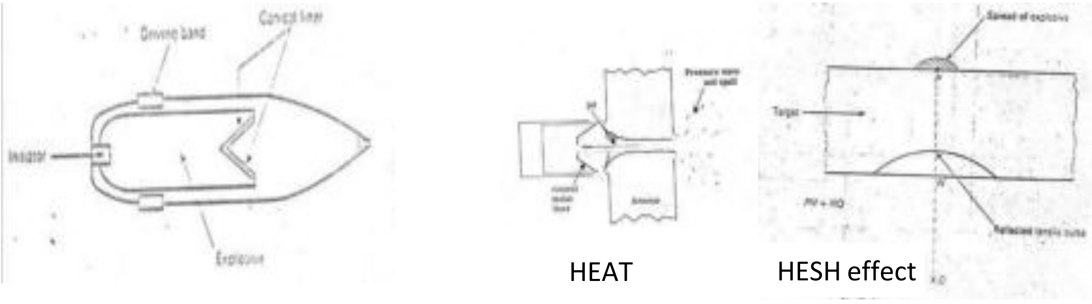
The model is selected for required study with given data and integrated numerically to get the trajectory. Simulation study is carried out to finalize the initial conditions in order to reduce the dispersion and increase the accuracy and consistency.

Terminal Ballistics

Projectile is designed to reach the terminal point: be in air (in the vicinity of the aircraft) or on the ground (in the vicinity of a structure, tank, bunker etc.) in desired orientation with a desired velocity and from here the terminal phase commences. Damage to the target is achieved in differentways : Kinetic energy of the penetrating bodies , chemical energy of a high explosive and a combination of the both. Damage to the target is classified as : scabbing, plugging, petalling and ductile failure. The projectile also gets damaged as: barrelling, shatter, lateral bending and compression. There are three types of KE shots: bullets, Long rod penetrator and Fragmentation shell with natural fragments or preformed fragments.

There are three types of chemical shots which use high explosive to damage the target viz. HEAT orshaped charge, HESH and Blast. The functioning of HE shots is shown in Fig.8.

In HEAT, shock energy is concentrated at a single point as a jet. In HESH, HE is spread over the target and detonated. The tensile strength of the target is overcome and the target material is broken. Blast warheads are used by detonating the HE for damaging the structure from outside or from inside.



HEAT

Figure-8. Terminal effects of HE shots

With help of the requirements and arget information, Gun design is decided and improved for required effect and accuracy.

Ballistics studies can be applied for analysis of various areas like safe ejection of the pilot in case of emergency exit, store separation from the aircraft, towed body dynamics for towing vehicles, safety/danger areas for firing ranges. An application of projectile ricochet analysis using mathematical modeling and simulation approach gives limiting conditions to define the safety zone.

References:

1. Philosophical Library, Interior Ballistics, HMSO, (1951).
2. Longdon, L.W. Text Book of Ballistics and Gunnery- volume 1, HMSO (1987).
3. Stifel, L. (1988), Progress in aeronautics and astronautics, 99.
4. Krier, H., Michael, J. A. (1979), American Institute of Aeronautics and Astronautics Journal, 66, 1-36.
5. Carlucci, D. E., Jacobson, S. S. Ballistics Theory And Design Of Guns And Ammunition, CRC Press (2008).

Numerical Solution of Non-linear Impulsive Differential Equation by Simpson's $\frac{1}{3}^{\mathrm{rd}}$ Rule

Indrayani Y. Sutar [a,*], Sandhyatai D. Kadam[a]
[a] Dr. D. Y. Patil Institute of Technology, Pimpri, Pune 411 018, (MS) India.
*Corresponding author: sutarindrayani01@gmail.com

ABSTRACT

Impulsive differential equations occur in many physical situations such as control theory, mechanics, epidemiology, pharmacokinetics etc. Finding solution of such differential equations using analytical methods is not always possible. Therefore, numerical methods can be employed to obtain approximate solutions of these differential equations. In the present paper, a new numerical method is proposed to obtain the solution of nonlinear impulsive differential equation with finite number of discontinuities. Integral term, involved in the mild solution impulsive differential equation, is approximated using Simpson's $\frac{1}{3}^{rd}$ rule and then splitting the solution by DGJ method. Further, error in the proposed method is computed. Furthermore, the results on error approximation and stability analysis are computed.

KEYWORDS

Impulsive differential equations, Simpson's $\frac{1}{3}^{rd}$ rule, Daftardar-Gejji and Jafari method, Numerical solution, Error.

***AMS Subject Classification*:** 34A12, 39B82, 47GXX, 65D30.

……………………………………………………………………………………….

1. INTRODUCTION

The theory of impulsive differential equation is an emerging area of research. Due to the nature of short-term perturbation of impulsive differential equations, whose duration is negligible as compared to the whole phenomenon, they are useful tools in modeling of many real-life problems that are subject to sudden changes in the state.

V Laxmikantham [8], D. Bainov and Simemnov [1] developed theory of impulsive differential equation. Many authors studied existence, uniqueness and qualitative properties of mild solutions of impulsive problems with the help of various fixed point theorems [3,6,11], measure of non-compactness, contraction principal etc. But many impulsive differential equations cannot be solved analytically or in some cases it is more confounded to tackle them.

Approximate solutions to differential equations are often computed using finite difference method resulting in estimated values for the solution of differential equation at some grid points. There are many methods for developing numerical approaches of these kinds of problems. The first stage is about substituting some values for the integral occurring in the solution of differential equation by any numerical quadrature based on grid point values.

V. Gejji and H. Jafari [2] developed an iterative method (Daftardar-Gejji and Jafri (DGJ) method) for solving functional differential equations. Jhinga A. and others [4], proposed a method for solving Volterra Integro differential equations using trapezoidal rule along with DGJ method. Authors also studied error and convergence of the proposed method.

To the best of our knowledge there is less contribution in the impulsive problems by the researchers.

In the present paper, we have studied impulsive nonlinear problem with finite number of discontinuities given by,

$$u'(x) \,\&= f(t,u(t))\ ,\ \ t \neq \tau_k\ , \qquad t \in [0,T] \tag{1}$$

$$u(x_0) = u_0\ , \tag{2}$$

$$\Delta\, u(\tau_k) = I_k\big(u(\tau_k)\big), \qquad k = 1,2,..m \tag{3}$$

where τ_k are moments of impulse and I_k is the sudden change of state at every τ_k.

2. PRELIMINARIES AND HYPOTHESES

Let X be a Banach space with the norm $\|\cdot\|$.

Let $PC([0,T],\ X) = \{u: [0,T] \rightarrow\ X \mid u(t)\}$ is piecewise continuous at $t \neq \tau_k$, left continuous at $t = \tau_k$, that is, $u(\tau_k^-) = \lim_{h\rightarrow 0^+} u(\tau_k - h) = u(\tau_k)$ and the right limit $u(\tau_k + 0)$ exists for $k = 1,2,\dots,m$. Clearly, $PC([0,T]\ ,\ X)$ is a Banach space with the supremum norm

$$\|u\|_{PC([0,T]\,,X)} = sup\{\|u(t)\| : t \in [0,T] \setminus \{\tau_1,\tau_2,\dots,\tau_m\}\}.$$

Definition 2.1: A function $u \in PC([0,T]\,,X)$ satisfying the equations

$$u(t) = u_0 + \int_0^t f\big(t,u(t)\big)dt + \sum_{0<\tau_k<t} I_k\, u(\tau_k), \quad t \in (0,T],$$

$$u(0) = u_0.$$

is said to be the mild solution of the initial value problem.

Definition 2.2: Let $u_0(h)$, $u_1(h), \cdots$ denote the approximation obtained by a given method using step size h then the method is said to be convergent if and only if

$$\lim_{x \to 0} |u_i(h) - u_i(x_i)| \; \to \; 0 \qquad for \;\; i = 1,2,3, \cdots, N \qquad (4)$$

as $h \to \; 0$ and $N \to \; \infty$.

Definition 2.3: A method is said to be of order k, if k is the largest number for which there exist a positive constant C such that

$$|u_i(h) - u_i(x_i)| \;\; \leq \; Ch^k, \qquad i = 0, 1,2,3, \cdots \; , N \; ; \forall \; h > 0 \qquad (5)$$

Hypothesis 1: Let $f: R \times \; X \to \; X$ be function, there exist a positive constant C such that

$$| \, f(t,y) \; - \; f(t,\xi)| \leq \; Ch \, | \, y - \xi \, | \, , \;\; \forall \, \xi, y \in \; X \, , \quad 0 \leq \; t \leq \; T$$

Hypothesis 2: Let $I_k: X \to \; X$ be function, there exist a positive constant h_k such that

$$|I_k(x) - I_k(y)| \leq \; h_k \, |x - y|, \qquad \forall \, x, y \; \in \; X$$

3. NUMERICAL METHOD

Consider an impulsive differential equation of the form

$$u'(x) = f\big(t, u(t)\big) \, , \qquad t \neq \tau_k \, , \;\; t \in \; [x_0, x \,] \qquad (6)$$

under the conditions

$$u(x_0) \; = \; u_0 \;\; , \qquad \Delta \, u(\tau_k) \; = \; I_k u(\tau_k). \qquad (7)$$

Solution of equation (6) along with the conditions stated in (7) is given in [6]

$$u(x) = u_0 \; + \; \int_{x_0}^{x} f\big(t, u(t)\big) dt \; + \sum_{x_0 < \tau_k < x} I_k u(\tau_k). \qquad (8)$$

In this section we apply Simpson's $\frac{1}{3}^{rd}$ rule, to the integral term in equation (8), to obtain its approximate solution. Interval is divided into equal parts with step length h. Thus we get

$$u\big(x_{j+1}\big) = \; u_{j+1} = u_j \; + \int_{x_j}^{x_j+h} \; f\big(t, u(t)\big) dt + \sum_{x_j \, < \, \tau_k \, < x_j + h} I_k u(\tau_k)$$

$$= \; S + \; u_j \;\; + \frac{h}{3} f\left(x_j + h, u\big(x_j + h\big)\right) \; + \frac{h}{3} \; f\left(x_j, u\big(x_j\big)\right)$$

$$+\frac{4h}{3}\sum_{i=1}^{\left\lfloor \frac{j}{2}-1\right\rfloor} f\big(x_{2i-1}, u(x_{2i-1})\big)$$

$$+\frac{2h}{3}\sum_{i=1}^{\left\lfloor \frac{j}{2}-1\right\rfloor} f\big(x_{2i}, u(x_{2i})\big), \qquad (9)$$

where, $S = \sum_{x_j < \tau_k < x_j + h} I_k u(\tau_k)$.

Approximating impulse term as

$$\Delta x(\tau_k) = x(\tau_k^+) - x(\tau_k^-) = I_k\big(x(\tau_k)\big),$$

$$I_k\big(x(\tau_k)\big) = \frac{B_k}{N_k}, \qquad (10)$$

where $B_k = h^k(\tau_{k+1} - \tau_k)$; $N_k \in \mathbb{N}$ and $x_j = jh$, $j = 0,1,\cdots T$.

$$\therefore \quad u\big(x_j + h\big) = u(jh + h) = u(j+1)h = u_{j+1}.$$

Then equation (9) becomes

$$u_{j+1} = S + u_j + \frac{h}{3}\big[f\big(x_{j+1}, u_{j+1}\big) + f\big(x_j, u_j\big)\big] + \frac{4h}{3}\sum_{i=1}^{\left\lfloor \frac{j}{2}-1\right\rfloor} f(x_{2i-1}, u_{2i-1})$$

$$+\frac{2h}{3}\sum_{i=1}^{\left\lfloor \frac{j}{2}-1\right\rfloor} f(x_{2i}, u_{2i}). \qquad (11)$$

Define

$$L(u) = S = \sum_{x_j < \tau_k < x_{j+h}} I_k u(\tau_k)$$

$$N(u) = \frac{h}{3} f\big(x_{j+1}, u_{j+1}\big)$$

$$g = u_j + \frac{h}{3} f\big(x_j, u_j\big) + \frac{4h}{3}\sum_{i=1}^{\left\lfloor \frac{j}{2}-1\right\rfloor} f(x_{2i-1}, u_{2i-1})$$

$$+\frac{2h}{3}\sum_{i=1}^{\left\lfloor \frac{j}{2}-1\right\rfloor} f(x_{2i}, u_{2i}). \qquad (12)$$

Where $L(u)$ and I_k are linear operators, $N(u)$ is a Nonlinear operator and g is a known function.

Thus equation (11) becomes

$$
\begin{aligned}
u_{j+1} = {} & S + u_j + \frac{h}{3} f(x_j, u_j) + \frac{4h}{3} \sum_{i=1}^{\left\lfloor \frac{j}{2} - 1 \right\rfloor} f(x_{2i-1}, u_{2i-1}) + \frac{2h}{3} \sum_{i=1}^{\left\lfloor \frac{j}{2} - 1 \right\rfloor} f(x_{2i}, u_{2i}) \\
& + \frac{h}{3} f\left(x_{j+1}, \left\{ S + u_j + \frac{h}{3} f(x_j, u_j) + \frac{4h}{3} \sum_{i=1}^{\left\lfloor \frac{j}{2} - 1 \right\rfloor} f(x_{2i-1}, u_{2i-1}) \right.\right. \\
& \left.\left. + \frac{2h}{3} \sum_{i=1}^{\left\lfloor \frac{j}{2} - 1 \right\rfloor} f(x_{2i}, u_{2i}) + \frac{h}{3} f(x_{j+1}, u_{j+1}) \right\} \right)
\end{aligned}
$$

$$
\begin{aligned}
= {} & S + u_j + \frac{h}{3} f(x_j, u_j) + \frac{4h}{3} \sum_{i=1}^{\left\lfloor \frac{j}{2} - 1 \right\rfloor} f(x_{2i-1}, u_{2i-1}) + \frac{2h}{3} \sum_{i=1}^{\left\lfloor \frac{j}{2} - 1 \right\rfloor} f(x_{2i}, u_{2i}) \\
& + \frac{h}{3} f\left(x_{j+1}, \left\{ S + u_j + \frac{h}{3} f(x_j, u_j) + \frac{4h}{3} \sum_{i=1}^{\left\lfloor \frac{j}{2} - 1 \right\rfloor} f(x_{2i-1}, u_{2i-1}) \right.\right. \\
& + \frac{2h}{3} \sum_{i=1}^{\left\lfloor \frac{j}{2} - 1 \right\rfloor} f(x_{2i}, u_{2i}) \\
& + \frac{h}{3} f\left(x_{j+1}, S + u_j + \frac{h}{3} f(x_j, u_j) + \frac{4h}{3} \sum_{i=1}^{\left\lfloor \frac{j}{2} - 1 \right\rfloor} f(x_{2i-1}, u_{2i-1}) \right. \\
& \left.\left.\left. + \frac{2h}{3} \sum_{i=1}^{\left\lfloor \frac{j}{2} - 1 \right\rfloor} f(x_{2i}, u_{2i}) \right) \right\} \right). \qquad (13)
\end{aligned}
$$

Let

$$N_1 = S + u_j + \frac{h}{3} f(x_j, u_j) + \frac{4h}{3} \sum_{i=1}^{\left\lfloor \frac{j}{2} - 1 \right\rfloor} f(x_{2i-1}, u_{2i-1})$$

$$+ \frac{2h}{3} \sum_{i=1}^{\left\lfloor \frac{j}{2} - 1 \right\rfloor} f(x_{2i}, u_{2i}), \qquad (14)$$

and $N_2 = N_1 + \frac{h}{3} f(x_{j+1}, N_1)$. (15)

With these notations equation (13) can be written as

$$u_{j+1} = N_1 + \frac{h}{3} f(x_{j+1}, N_2) \qquad (16)$$

Equation (16) represents an approximate solution of equation (8).

4. ERROR ANALYSIS

Theorem Assume that f satisfies the hypothesis 1 for any positive constant C, then the proposed numerical method is of fifth order.

Proof: Let u_{j+1} is an approximation to $u(x_{j+1})$. From equations (9),(14) and (16) we obtain

$$\begin{aligned}
|u(x_{j+1}) - u_{j+1}| &= \frac{h}{3} \left| f(x_{j+1}, u_{j+1}) - f(x_{j+1}, N_2) \right| + O(h^5) \\
&\le \frac{Ch^2}{3} \left| u_{j+1} - N_2 \right| + O(h^5) \\
&\le \frac{Ch^2}{3} \left| N_1 + \frac{h}{3} f(x_{j+1}, N_2) - N_2 \right| + O(h^5) \\
&\le \frac{Ch^3}{9} \left| f(x_{j+1}, N_2) - f(x_{j+1}, N_1) \right| + O(h^5) \\
&\le \frac{Ch^4}{9} |N_2 - N_1| + O(h^5) \\
&\le \frac{Ch^5}{27} \left| f(x_{j+1}, N_1) \right| + O(h^5).
\end{aligned}$$

Hence, the proposed method is of order 5.

COROLLARY: The numerical method (16) is convergent.

Proof: By error analysis result and definitions 2.2 and 2.3 numerical method (16) is convergent.

5. STABILITY ANALYSIS

Theorem Let u_{j+1}^{n} and v_{j+1}^{n} be the two n^{th} approximate numerical solutions of given impulsive differential equations, satisfying hypotheses 1 and 2, are stable if and only if $C_\alpha \geq 0$.

Proof: Let u_{j+1}^{n} and v_{j+1}^{n} be the two approximate solutions. To establish the stability, consider,

$$
\begin{aligned}
|u_{j+1}^{n} - v_{j+1}^{n}| &= \Big| N_1 + \frac{h}{3} f(x_{j+1}, N_2) - N_1' - \frac{h}{3} f(x_{j+1}, N_2') \Big| \\
&\leq |N_1 - N_1'| + \frac{h}{3} |f(x_{j+1}, N_2) - f(x_{j+1}, N_2')| \\
&\leq |N_1 - N_1'| + \frac{Ch^2}{3} |N_2 - N_2'| \\
&\leq |N_1 - N_1'| + \frac{Ch^2}{3} \Big| N_1 + \frac{h}{3} f(x_{j+1}, N_1) - N_1' - \frac{h}{3} f(x_{j+1}, N_1') \Big| \\
&\leq |N_1 - N_1'| + \frac{Ch^2}{3} |N_1 - N_1'| + \frac{Ch^2}{9} |f(x_{j+1}, N_1) - f(x_{j+1}, N_1')| \\
&\leq |N_1 - N_1'| + \frac{Ch^2}{3} |N_1 - N_1'| + \frac{C^2h^3}{9} |N_1 - N_1'|
\end{aligned}
$$

$$
|u_{j+1}^{n} - v_{j+1}^{n}| \leq \left(1 + \frac{Ch^2}{3} + \frac{C^2h^3}{9} \right) |N_1 - N_1'|. \tag{17}
$$

Using the definition of N_1 and Lipschitz's condition we get

$$
\begin{aligned}
|N_1 - N_1'| \leq\ & |S - S'| + |u_j - v_j| + \frac{h}{3} |f(x_j, u_j) - f(x_j, v_j)| \\
& + \frac{4h}{3} \sum |f(x_{2j+1}, u_{2j+1}) - f(x_{2j+1}, v_{2j+1})| \\
& + \frac{2h}{3} \sum |f(x_{2j}, u_{2j}) - f(x_{2j}, v_{2j})| \\
\leq\ & h_k |u_k - v_k| + |u_j - v_j| + \frac{h^2}{3} |u_j - v_j|
\end{aligned}
$$

$$+\frac{4h^2}{3}\sum|u_{2j-1} - v_{2j-1}| + \frac{2h^2}{3}\sum|u_{2j} - v_{2j}|$$

$$\leq h_k|u_0 - v_0| + |u_0 - v_0| + \frac{h^2}{3}|u_0 - v_0|$$

$$+\frac{4h^2}{3}\sum|u_0 - v_0| + \frac{2h^2}{3}\sum|u_0 - v_0|$$

$$\leq \left(1 + h_k + \frac{7h^2}{3}\right)|u_0 - v_0|. \tag{18}$$

Substituting equation (18) in equation (17) we get

$$|u_{j+1}^n - v_{j+1}^n| \leq C_\alpha |u_0 - v_0|, \tag{19}$$

where

$$c_\alpha = \left(1 + \frac{Ch^2}{3} + \frac{C^2h^3}{9}\right)\left(1 + h_k + \frac{7h^2}{3}\right) \geq 0.$$

Hence two approximate solutions of given impulsive differential equations are stable.

6. RESULTS AND DISCUSSION

In this paper we have considered impulsive differential equation given by equation (7). An integral term in the solution of impulsive differential equation is approximated using Simpson's $\frac{1}{3}^{rd}$ rule. It is proved that the proposed method is of order five. Stability and convergence of the method is also studied.

REFERENCES

1. Bainov D., Simeonov, P., Impulsive Differential Equations: Periodic Solutions and Applications, CRC Press, Boca Raton (1993).
2. Daftardar-Gejji, V., Jafari, H, (2006) Journal of Mathematical Analysis and Applications, 316, pp. 753–763.
3. Jain, R.S., Reddy, B.S., Kadam, S.D., (2018) Arabian Journal of Mathematics 7, pp. 273–279.
4. Jhinga, A, Patade, J., Daftardar-Gejji, V., (2020) arXiv:2009.11571v1 [math.NA]
5. Jin,Y., Chol, C., Ae, P., Song, J., Lu, G, (2018) J. Nonlinear Sciences and Applications, 11, pp. 1129–1148.

6. Kadam, S.D., Jain, R.S., Reddy, B.S., (2016) Open Journal of Applied and Theoretical Mathematics (OJATM), 2, No. 2455-7102, pp. 1027-1040.
7. Kumar, M., Jhinga, A. Daftardar-Gejji, V., (2020) International Journal of Applied and Computational Mathematics, 6:26.
8. Lakshmikantham, V, Bainov, D. D., Simeonov, P.S., Theory of Impulsive Differential Equations, World Scientific, Singapore (1889).
9. Liu, J. H., (1999) Dynamics of Continuous Discrete and Impulsive Systems, 6, pp. 77–85.
10. Ran, X. J., Liu,M. Z., Zhu,Q. Y., (2008) Mathematical and Computer Modelling, 48, pp. 46-55.
11. Shamsidah, N., Hamzah, B. A., Mamat, M B., Kavikumar, J., Lee, S C., Ahmad, N., (2010) Applied Mathematical Sciences, 4, no. 65, pp. 3219 - 3232.

RECENT DEVELOPMENT IN STURM–LIOUVILLE THEORY

Pratiksha Demanna[a,*], Sarita H. Thakar[a]
[a]Department of Mathematics, Shivaji University, Kolhapur 416 004 (MS) India.
*Corresponding author: pratikshademanna@gmail.com

ABSTRACT

A review of developments in Sturm-Liouville theory is presented. An overview of pioneering work of Sturm, Liouville, Weyl, Dixon, Stone and Titchmarsh is presented. Sturm-Liouville problems with separated and coupled boundary conditions are discussed. Haupt and Richardson's extension of Sturm-Liouville problems with indefinite weight is given. Further extension of Sturm-Liouville theory with discontinuous weight function, transmission condition, eigenparameter dependent boundary conditions is presented.

KEYWORDS

Sturm-Liouville theory, Eigenvalues, Eigenfunctions, eigenfunction expansions.

AMS Subject Classification: 34B24, 34L10, 34L15

……………………………………………………………………………………………….

1. INTRODUCTION

Swiss mathematician Jaçques Sturm (1803–1855) and French mathematician Joseph Liouville (1809–1882) independently worked on the second order differential equation in their notation

$$-\frac{d}{dt}\left(K\frac{dV}{dt}\right) + lV = rgV \quad \text{on} \quad [x, X] \tag{1.1}$$

with separated boundary conditions

$$\frac{dV}{dt} - hV = 0 \quad \text{for} \quad t = x \tag{1.2}$$

$$\frac{dV}{dt} + HV = 0 \quad \text{for} \quad t = X \tag{1.3}$$

K, l, g are positive functions on $[x, X]$, h and H are given positive numbers and r is a real–valued parameter. The zeros r of the transcendental equation

$$F(r) = 0 \tag{1.4}$$

for which boundary value problem (BVP) (1.1–1.2) has a nontrivial solution is called eigenvalue and the corresponding nontrivial solution is called eigenfunction.

Before the work of Sturm and Liouville the solution of differential equations are obtained as analytic expressions. They were among the first to identify the need of discovering the properties of solutions directly from the equation.

Though both of them worked on the same problem (1.1–1.2), there is a noteworthy disparity in their approach. Sturm focused on the qualitative behavior of eigenfunctions while Liouville is directed to eigenfunction expansion in Fourier series. They share some results in common viz. orthogonality theorem, reality of eigenvalues and determination of Fourier series coefficients.

Sturm in his first three papers [1–3] proved that the transcendental equation (1.4) has infinite number of real simple roots which are positive. In the joint work of Sturm and Liouville [4] these roots are arranged in increasing order of magnitude as $r_1, r_2, r_3, \cdots$ and $V_1, V_2, V_3, \cdots$ are associated eigenfunctions then these eigenfunctions are orthogonal w.r.t. weight function g. Moreover, given a function f defined on the interval $[x, X]$ follows the series expansion

$$f(x) = \sum_{n=1}^{\infty} a_n V_n(x)$$

where

$$a_n = \frac{\int_x^X g(y)V_n(y)f(y)dy}{\int_x^X g(y)V_n^2(y)dy}.$$

For more details see [84].

The widely used modern notation for Sturm–Liouville differential equation involving the notion of quasi–derivative is

$$(-p(x)y'(x))' + q(x)y(x) = \lambda\omega(x)y(x) \text{ for all } x \in (a, b) \qquad (1.5)$$

where

(i) $-\infty \leq a < b \leq \infty$,

(ii) $\frac{1}{p}$, q, $\omega : [a, b] \to \mathbb{R}$ $\frac{1}{p}$, q, $\omega \in L^1_{loc}$ (Local Lebesgue integration space)

(iii) $\omega(x) > 0$

(iv) $\lambda \in \mathbb{C}$

Definition 1.1. Problem (1.5) satisfying (i–iv) with two separated end point conditions

$$\alpha_1 y(a) + \alpha_2 y'(a) = 0, \ \beta_1 y(b) + \beta_2 y'(b) = 0$$

is called a regular Sturm––Liouville problem (SLP). Otherwise, it is singular.

Definition 1.2. Problem (1.5) with $P(a) = p(b)$ satisfying (i–iv) subjected to boundary conditions

$$y(a) = y(b), \ y'(a) = y'(b)$$

is called periodic Sturm–Liouville problem.

In the year 1910, Herman Weyl [5] started the investigation of singular SLP. He considered equation (1.5) with the following restrictions on the coefficients:
(i) $p(x)$, q, $\omega : [0, \infty) \to \mathbb{R}$, $\omega(x) = 1$ for $x \in [0, \infty)$
(ii)$p(x)$, $q(x) \in [0, \infty)$, $p(x) > 0$ on $[0, \infty)$
(iii)$\lambda \in \mathbb{C}$.
The proof of the general theorem for unbounded operators in Hilbert space by Von Neumann [6] and Stone [7] together with the fundamental work of Titchmarsh [13] created a great motivation for the stronger searches into the spectral theory of Sturm–Liouville operators.

Another significant development in Sturm–Liouville theory was done by A. C. Dixon in the year 1912. Dixon replaced the continuity conditions on the coefficients p, q, ω by the Lebesgue integrability conditions. The paper [9] considers equation (1.5) with the following assumptions:
(i) The interval $(a, b) \subseteq \mathbb{R}$ is compact, p^{-1}, q, $\omega : [a, b] \to \mathbb{R}$.
(ii) p^{-1}, q, $\omega \in L^1[a, b]$, p, $\omega > 0$ a. e. on $[a, b]$.
In this article, Dixon derived the existence of solutions and expansion theorem with the above restrictions on the coefficients.

From 1927 onwards John Von Neumann and M. H. Stone worked independently on the unbounded linear operators in the Hilbert space. In the year 1932, Stone published the book [7] which deals with the properties of Sturm–Liouville differential operator in Hilbert function spaces. In the modern notation, Stone studied equation (1.5) with:
(i) $(a, b) \subseteq \mathbb{R}$, $-\infty \leq a < b \leq \infty$.
(ii) $\omega(x) = 1$ for all $x \in (a, b)$.
(iii) p, $q : (a, b) \to \mathbb{R}$, p^{-1}, $q \in L^1_{loc}(a, b)$.

Titchmarsh's [10–12] work on regular and singular Sturm–Liouville differential equation is again a remarkable contribution to the field. He applied theory of function of single complex variable to study the Sturm–Liouville boundary value problems. His work [13, 14] for singular SLP plays an important role in the eigenfunction expansion for the singular case.

In the operator theoretic development of Sturm–Liouville theory after Stone [7], J. Weidmann [27], Naimark [29], Akhiezer–Glazman [26], Hellwig [85], K. Jörgens [86] made a remarkable contribution. In 40's a new tool i.e. transformation operator is introduced to enrich the field. Povzner [15] constructed transformation operator for arbitrary Sturm–Liouville equations, later he applied it to obtain the eigenfunction expansion for Sturm–Liouville equations with decreasing potential. Marchenko [16] used the transformation operators to study inverse problems and to derive the asymptotic behavior of singular

SLP. The details of Marchenko's work can be found in his monograph [16]. I. M. Gelfand and B. M. Levinton utilized transformation operators to prove the equiconvergence theorem.

After the operator theoretic development of Sturm–Liouville theory, the problem was essentially solving the eigenvalue problem for the differential operator

$$Ty = \lambda \omega y \tag{1.6}$$

with separated boundary conditions

$$AY(a) + BY(b) = 0 \quad \text{where} \quad A, B \in M_{2\times 2}(\mathbb{C}), Y = \begin{bmatrix} y \\ py' \end{bmatrix} \tag{1.7}$$

where $T = \frac{d}{dx}\left(-p(x)\frac{d}{dx} + q(x)\right)$. For the function f satisfying separated boundary conditions, let

$$D = \{f \in L^2(a,b) : f, pf' \in AC[a,b], \omega^{-1}(Tf) \in L^2(a,b)\}.$$

Associated with operator T the quadratic forms are defined as

$$Ly = \langle Ty, y\rangle \text{ and } Ry = \langle \omega y, y\rangle.$$

O. Haupt [17] and R. Richardson [28] were the first to notice that the spectrum of the problem (1.5–1.7) rely upon the definiteness conditions of the forms Ly and Ry.

2. Right Definite Sturm–Liouville Problems

Definition 2.1. If the form Ry is definite on D i.e. either $Ry > 0$ for all $y \neq 0$ in D or $Ry < 0$ for all $y \neq 0$ in D, then equation (1.5) with separated boundary conditions (1.7) is called right definite.

Hilbert and his school described such problems as orthogonal. When either $\omega > 0$ or $\omega < 0$ a.e. in (a, b) the problem (1.5), (1.7) is right definite. The spectrum of right definite Sturm–Liouville problems (RDSLP) with separated boundary conditions is real, $\lambda_n \to \infty$, $n \to \infty$ (see [24]). The eigenvalues are simple and the eigenfunction $y_n(x)$ corresponding to each eigenvalue λ_n has exaclty n zeros in (a, b) (see [19, 20]). The finiteness of negative spctrum of RDSLP is given by Mingarelli [87].

3. Left Definite Sturm–Liouville Problems

Definition 3.1. If the form Ly is definite on D for all $y \neq 0$ in D then the problem (1.5), (1.7) is left definite (LD).

The problem was termed as polar by Hilbert and his school. For such problems two sequence of eigenvalues $\{\lambda_n^{\pm}\}$ exists where $\lambda_n^{\pm} \pm \infty$. The eigenvalues can be numbered by the index set $\mathbb{Z} = \{\cdots, -2, -1, 0, 1, 2, \cdots\}$ such that

$$\cdots, \lambda_{-2}, \lambda_{-1}, \lambda_0, \lambda_1, \lambda_2, \cdots$$

and for every $n \in \mathbb{Z}$, the eigenfunction associated with eigenvalue λ_n has exactly $|n|$ zeros in the interval (a, b). For more details on LD SLP refer [18, 23–25]. Kong, Wu and Zettl [22] defined left definite SLP in terms of RD SLP.

4. Indefinite/ Non–definite Sturm–Liouville Problems

Definition 4.1. When neither the form Ly nor Ry is definite on D, the problem (1.5), (1.7) is called indefinite.

Initially Haupt [17] and later Richardson [28] extended the work of Sturm and Liouville by considering the sign changing (indefinite) weight function. Further in the year 1915, Haupt in his article [30] refined the results obtained in [17]. The first version of oscillation theorem was given by Haupt [30] whereas the final form of oscillation theorem was due to Richardson [18].

The spectrum of indefinite Sturm–Liouville problem (1.5), (1.7) is discrete, with doubly infinite sequence of real eigenvalue and has atmost finite and even number of complex eigenvalues [18, 23, 31]. Sufficient condition for the existence of non–real eigenvalue was given by Mingarelli [32]. Bound on the number of non–real eigenvalues of indefinite SLP in terms of number of negative eigenvalues of the corresponding RD SLP is established in [32].

In constrast with the RD and LD SLP the spectrum of indefinite SLP is not monotone. As a result the eigenfunction corresponding to the eigenvalue λ_0 can have any number of zeros. In relation with such nature of real spectrum Mingarelli [23] defined two types of indices namely Richardson Index (n_R)and Haupt Index (n_H), motivated by the work of Haupt and Richardson.

For each real eigenvalue there exists two numbers λ^+ and λ^- called as Richardson numbers [33]. Mingarelli noted that in the right definite case $\lambda^+ = \lambda^- = -\infty$ whereas in the left definite case $\lambda^+ = \lambda^- = 0$. Atkinson and Jabon [33] were the first to solve the acual example of indefinite Sturm–Liouville with sign changing weight function

$$\omega(x) = \begin{cases} -1 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

and obtained the lower bound on the Richardson numbers efficiently. The following theorem is due to Richardson [18] known as Richardson oscillation theorem.

Theorem 4.1. Let ω be continuous and not vanish identically in any right neighborhood of $x = a$. If $\omega(x)$ changes its sign precisely once in (a, b), then the roots of the real and imaginary parts ϕ and ψ of any non-real eigenfunction $u = \phi + i\psi$, corresponding to a non-real eigenvalue, separate one another (or interlace).

Kikonko and Mingarelli [37] extended the study of regular indefinite SLP when the weight function $\omega(x)$ changes sign twice. In this article the lower bound on the Richardson number λ^+ is determined. The detailed numerical study of SLP with Dirichlet boundary condition in two turning point case was done by Kikonko [34]. Kikonko noted that the Richardson oscillation theorem fails in two turning point case. In the same article Kikonko and mingarelli discussed the difference between the number of zeros of real and imaginary parts associated with complex eigenvalue when Richardson oscillation theorem fails (see [34]).

When non–real eigenvalue exists, the question of obtaining the bound on real and imaginary parts of these values was raised by Mingarelli. He was the first to notice that no bounds on these eigenvalues are obtained interms of the coefficient function q and w. This question was solved by Mingarelli [40] by using Green's function argument. Further it was answered by Qi, Xie and Chen [39], Behrndt, Chen, Philip and Qi [44, 45], Xie and Qi [42], Behrndt, Philip and Trunk [43] etc. applied L^2–estimates together with quadratic form argument and Krein space theory. Behrndt, Chen, Qi [41], Kikonko, Mingarelli [36] improved the bound on the real and imaginary part of a non–real eigenvalue corresponding to the bound obtained by Behrndt and etal [44]. Moreover Kokonko and Mingarelli [36] derived the lower bound for the eigenvalue of smallest modulus.

In comparison with SLP with separated boundary conditions much less is known about SLP with coupled/periodic boundary conditions. For periodic boundary conditions there may exists geometrically double eigenvalues. Oscillatory properties of regular SLP with periodic boundary conditions are described in [19, 83]. Kong, Wu and Zettl [22] analyzed left definite SLP with periodic boundary conditions. It is noted that eigevalue of left definite SLP may be geometrically simple or double.

Upto 2019, there is no work done on indefinite SLP with coupled boundary condition. There is need to determine the multiplicity of eigenvalues. Moreover, it is interesting to see whether the Richardson Oscillation theorem holds or not in one turning point case. All these questions are examined by Sarita Thakar and Pratiksha Demanna [47]. Moreover, the necessary and sufficient condition for the existece of double eigenvalue when the potential function $q(x)$ is constant is determined.

In the early years, the SLP with discontinuous coefficient or discontinuities in the solution or its derivative at an interior point attracted many researchers, so the new findings have been built up in this area and enhanced the field in many aspects.

Sturm–Liouville BVP with discontinuous coefficient arises in mathematical physics, natural sciences, geophysics and other fields of engineering for e.g. modelling toroidal vibrations and free oscillations of the earth [51, 52], analysis of one dimensional photonic crystals [53, 54]. Nabiev and Amirov [50] considered SLP (1.5) with $p(x) = 1$ and

$$\omega(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq c \\ \alpha^2 & \text{if } c \leq x \leq \pi,\ 0 < \alpha \neq 1,\ \alpha \in \mathbb{R} \end{cases}$$

subjected to separated boundary conditions. In this article spectral properties and eigenfunction expansion theorem is established for the problem under consideration. O. Akcay [88] considered SLP with discontinuous coefficient as well as discontinuity conditions inside the interval. In this study asymptotic nature of eigevalues, properties of kernel function and eigenfunction expansion theorem are discussed. Recently Mukhtarov and Ayedemir [59] examined the spectral properties and asymptotic behavior of discontinuous SLP subjected to periodic boundary condition.

In the classical Sturm–Liouville problem the eigenvalue parameter appears linearly in the differential equation. However the problems are noticed where the eigenvalue parameter occurs both in the differential equation and boundary conditions. In 1976, Fulton noted that the Titchmarsh's analysis [13] for regular SLP on a finite interval is also applicable to the regular SLP containing the eigenparameter in the boundary condition at one end but not both. The operator theoretic formulation for such problems was proposed by Walter [62] utilized by Fulton to illustrate the eigenfunction expansion theorem. Eigenvalues of such problems are studied in [60–67]. Belinsky and Dauer [68] obtained the Rayleigh–Ritz formula for eigenvalues.

In recent time more and more researchers are attracted to discontinuous Sturm–Liouville problems with eigenparameter dependent boundary and transmission conditions due to its applications in physics. Such type of problems appears in heat and mass transfer [55], vibrating string problems when the string was loaded additionally with point masses [55], thermal conduction problem for a thin laminated plate [56], diffraction problems [90]. SLP with eigenparameter dependent boundary conditions and transmission condition condition at one interior is studied in [73–76, 79, 82]. The study of Sturm–Liouville problems with eigenparameter dependent boundary conditions and finite number of transmission conditions is extended in [70–72, 77, 78].

REFERENCES

[1] Sturm, C. (1829) imprimerie de A. Firmin Dodot, rue Jacob, 24.

[2] Sturm, C. (1836) Joural of Mathematics Pures Applications, 1, 106–186.

[3] Sturm, C. (1836) Joural of Mathematics Pures Applications, 1, 373–444.

[4] Sturm, C. and Liouville, J. (1837) Joural of Mathematics Pures Applications, 2, 220-223.

[5] Weyl, H. (1910) Mathematische Annalen, 68, 220–269.

[6] Neumann, J. (1929) Mathematische Annalen, 102, 49–131.

[7] Stone, M. H., In American Mathematical Society Colloquium Publications (1932).

[8] Titchmarsh, E. C., Eigenfunction expansions I, In Oxford University Press (1962).

[9] Dixon, A. C. (1912) Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character, 211, 411–432.

[10] Titchmarsh, E. C. (1941) The Quarterly Journal of Mathematics, 12 , 33–50.

[11] Titchmarsh, E. C. (1941) The Quarterly Journal of Mathematics, 12, 89–107.

[12] Titchmarsh, E.C. (1941) The Quarterly Journal of Mathematics 12 154–166.

[13] Titchmarsh, E.C. Eigenfunction expansions I, Oxford University Press, second edition, (1962).

[14] Titchmarsh E.C. Eigenfunction expansions II, Oxford University Press, (1958).

[15] Povzner, A. (1948) Matematicheskii Sbornik, 65, 3–52.

[16] Marchenko, V. A. Sturm–Liouville Operators and Applications, American Mathematical Society, 373 (2011).

[17] Haupt, O. Untersuchungen über Oszillationstheoreme, Teubner (Leipzig), (1911).

[18] Richardson, R. (1918) American Journal of Mathematics, 40, 283–316.

[19] Coddington, E. A., Levinson N. Theory of ordinary differential equations, McGraw-Hill, New York /London/Toronto (1955).

[20] Atkinson, F. V., Discrete and continuous boundary value problems, Academic Press, New York /London (1964).

[21] Mingarelli, A. B., Atkinson, F. V. Asymptotics of the number of zeros and of the eigenvalues of general weighted Sturm–Liouville problems, Walter de Gruyter, Berlin/New York Berlin, New York (1987) 380–393.

[22] Kong, Q., Wu H., Zettl, A. (2001) Left–definite Sturm–Liouville problems, Journal Differential Equations, 177, 1–26.

[23] Mingarelli, A. B. (1986) A survey of the regular weighted Sturm–Liouville problem–The non-definite case. Applied Differential Equations, World Scientific, Singapore and Philadelphia, 109–137.

[24] Zettl, A., Sturm–liouville theory, American Mathematical Society, 121 (2012).

[25] Zettl, A., Recent Developments in Sturm–Liouville Theory, de Gruyter (2021).

[26] Akhiezer, I., Glazman, I. M., Theory of linear operators in Hilbert space, Pitman–London and Scottish Academic Press, Edinburgh (1981).

[27] Weidmann, J., Linear operators in Hilbert spaces, Springer Science & Business Media, 68 (2012).

[28] Richardson, R. (1912) Transactions of the American Mathematical Society, 13, 22–34.

[29] Naimark, M. A., Linear differential operators II. Ungar Publishing Company, New York (1968).

[30] Haupt, O.(1915) Mathematische Annalen, 76, 67–104.

[31] Allegretto, W., Mingarelli, A. B. (1989) Journal für die reine und angewandte Mathematik, 398, 1–24.

[32] Mingarelli, A. B. (1982) In Ordinary and partial differential equations, Lecture notes in Mathematics, Springer, Berlin-New York, 519–528

[33] Atkinson, F. V., Jabon, D. (1984) workshop on Spectral Theory of Sturm-Liouville Differential Operators, Argonne National Laboratory, Argonne, Illinois 60439.

[34] Kikonko, M. (2016) On a non–definite Sturm-Liouville problem in the two turninpoint case–analysis and numerical results. Joural of Applied Mathematics and Physics, 4, 1787–1810.

[35] Kikonko, M., Mingarelli A. B., Bounds on the real and Imaginary parts of non-real eigenvalues of non-definite Sturm-Liouville problems. J. Differential Equations, 261, no. 11, 6221–6232.

[36] Kikonko, M., Mingarelli, A. B. (2016) International Journal Differential Equations DOI:10.1155/2016/7396951.

[37] , M., Mingarelli, A. B. (2013), Journal of Applied Mathematics and Computations, 219, 9508–9515.

[38] Qi, J., Chen, S. (2014) Joural of Spectral Theory, 4, no. 1, 53–63.

[39] Qi, J., Xie, B. and Chen, S. (2016) Proceeding of American Mathematical Society, 144, 547–559.

[40] Mingarelli, A. B. (1988) Proceeding International Conference on Theory and Applications of Differential Equations, 222–228.

[41] Behrndt, J., Chen, S., and Qi, J. (2014) Proceedings of the Royal Society of Edinburgh Section A, 144, 1113–1126.

[42] Xie, B., Qi J. (2013) Journal of Differential Equations , 255, 2291–2301.

[43] Behrndt, J., Philipp, F. and Trunk, C. (2013) Mathematische Annalen, 357,185–213.

[44] Behrndt, J., Chen, S., Philipp, F., and Qi, J. (2013) Proceedings in Applied Mathematics and Mechanics, 13, 525–526.

[45] Behrndt, J., Chen, S., Philipp, F., and Qi, J. (2014) Proceeding Royal Society Edinburgh Section A, 144(6), 1113–1126.

[46] Ince E. L., Ordinary differential equations, Courier Corporation, (1956).

[47] Thakar Sarita, Demanna Pratiksha (2022) Journal of Indian Mathematical Society, 89 (1–2), 182–197.

[48] Akhmedova, E. N., Huseynov, H. M. (2003) On eigenvalues and eigenfunctions of one class of Sturm–Liouville operators with discontinuous coefficient. Trans. NAS Azerb. 23, 7–18.

[49] Mamedov, K. R., Cetinkaya, F. A. (2013) Boundary Value Problems, 183, 16.

[50] Nabiev, A. A., Amirov, R. K.(2013) Mathematical Methods in Applied Sciences, 36, 1685–1700.

[51] Anderssen, R. S. (1997) Geophysical Journal Royal Astronomical Society; 50:303–309.

[52] Lapwood, F. R., Usami T.(1981) Cambridge University Press: Cambridge, 1981.

[53] Morozov, G. V., Sprung, D. W. L. (2011) Europhysics Letters, 96(5), 54005.

[54] Nusinsky, I., Hardy, A. A., (2006) Physical review B, 73(12), 125104.

[55] Tikhonov, A. N., Samarskii, A. A., Equations of Mathematical Physics, Oxford and New York, Pergamon, 1963.

[56] Titeux, I. and Yakubov, S., Application of Abstract Differential Equations to some Mechanical Problems, Kluwer Academic Publishers, Dordrecht, Boston, London, 2003.

[57] Amirov, R. KH. (2006) Journal of Mathematical Analysis and Applications 317(1), 163–176.

[58] Aydemir, K., Mukhtarov, O. Sh. (2013) arXiv preprint arXiv:1312.3052 .

[59] Mukhtarov, O. Sh., Aydemir K. (2021) Mathematical Methods in the Applied Science,s 44(18), 14664–14676.

[60] Kapustin N.Y., Moisseev E.I. (1997) Spectral Problem with the Spectral Parameter in the Boundary Condition, Differential Equations, 33, 115–119.

[61] Shkalikov, A. A. (1983) Boundary Problems for Ordinary Differential Equations with Parameter in the Boundary Conditions, Translation from Trudy Seminara imeni I.G.Petrovskogo, 9, 190–229.

[62] Walter, J. (1973) Mathematische Zeitschrift, 133, 301–312.

[63] Schneider, A. (194) Mathematische Zeitschrift 136, 163–167.

[64] Hinton, D.B. (1979) An Expansion Theorem for an Eigenvalue Problem with Eigenvalue Parameter in the Boundary Condition, Quarterly Journal of Mathematics, 30, 33–42.

[65] Russakovskij, E.M. (1975) Funktsional'nyi Analiz i ego Prilozheniya, 9, 358–359.

[66] Binding, P.A., Browne, P.J., Siddighi, K. (1999) Proceedings of the Edinburgh Mathematical Society, 37(1), 57–72.

[67] Mamedov, Kh.R. (1997) Spectral Theory of Operators and Its Applications, 11, 117–121 (in Russian).

[68] Belinskiy, B.P., Dauer, J.P. (1997) Spectral Theory and Computational Methods of Sturm-Liouville Problems, University of Tennessee, 183–196.

[69] Mamedov, Kh.R. (2005) On a Basic Problem for a Second Order Differental Equation with a Discontinuous Coefficient and a Spectral Parameter in the Boundary Conditions, Geometry Integrability and Quantization, 7, 218–225.

[70] Şen, E., Mukhtarov O. Sh., and Orucoglu K. (2016) Iranian Journal of Science and Technology, Transactions A: Science, 40(1), 1–8.

[71] Şen, E., Mukhtarov, O. Sh. (2016) Mediterranean Journal of Mathematics, 13(1), 153–170.

[72] Kadakal, M., Mukhtarov, O. Sh. (2007) Computers & Mathematics with Applications 54 :1–12, 1367–1379.

[73] Mukhtarov, O. Sh., Kadakal, M. , and Muhtarov, F. Ş (2004) Reports on mathematical Physics, 54(1), 41–56.

[74] Mukhtarov O.Sh., Kadakal M.(2005) Some spectral properties of one Sturm–Liouville type problem with discontinuous weight, Siberian Mathematical Journal, 46(4), 681–694.

[75] Şen, E. (2013) Physics Research International 2013.

[76] Kadakal, M., Mukhtarov, O. Sh. (2006) Acta Mathematica Sinica, 22(5), 1519–1528.

[77] Yang, Q., Wang, W. (2011) Mathematical Methods in the Applied Sciences, 34(4), 373–383.

[78] Yang, Q., Wang, W. (2012) Mathematical Sciences 6(1), 1–9.

[79] Akcay, O. (2018) Acta Mathematica Scientia, 38, 1195–1213.

[80] Fulton, C. (1977) Proceedings of the Royal Society of Edinburgh Section A: Mathematics, 77, 293–308.

[81] Akdoğan, Z., Demirci, M., and Mukhtarov, O. Sh. (2007) Mathematical Methods in the Applied Sciences, 30, 1719–1738.

[82] Tunç, E., Muhtarov, O. Sh. (2004) , Applied mathematics and Computation, 157, 347–355.

[83] Ince E. L., Ordinary Differential Equations, Dover Publications, INC. New York, (1956).

[84] Lützen, In the emergence of Sturm-Liouville theory, Archive for history of exact sciences 29 (1984) 309–376.

[85] Hellwig, G., Differential operators of mathematical physics, Addison–Wesley, London, (1964).

[86] Jörgens, K., Rellich, F.,Eigenwerttheorie gewöhnlicher Differentialgleichungen, edited by J. Weidmann, Springer, Heidelberg, (1976).

[87] Jörgens, K. Spectral Theory of Second-order Ordinary Differential Operators: Lectures Delivered at Aarhus Universitet, 2, (1964).

[88] Akcay, O. (2019) Mediterranean Journal of Mathematics, 16, 1–17.

[89] Haberman, R., Elementary applied partial differential equations, Englewood Cliffs, NJ: Prentice Hall, 987, (1983).

[90] Agranovich, M. S., Katsenelenbaum, B. Z., Sivov, A. N., and Voitovich, N. N., Generalized method of eigen oscillations in diffraction theory, (1999).

Numerical Simulations of Burgers' Equation by Cubic B-spline Galerkin Finite Element Method

Sunil S. Kumbhar [a],*

[a]Department of Mathematics, Shivaji University, Kolhapur 416 004 (MS) India.
Corresponding author: sunilmath25@gmail.com

ABSTRACT

In this paper we have constructed cubic B-splines based Galerkin finite element method (FEM) to compute approximate solutions of one dimensional non-homogeneous Burgers' equation (NBE). Initially Euler's implicit technique is used to obtain time discretization of NBE. Galerkin FEM is then applied to this discretized form. Stability of the present method is studied by using von Neumann analysis. The applicability and accuracy of this method is demonstrated by comparing computed numerical solutions of some test examples by the proposed method with the exact and numerical solutions available in the literature.

KEYWORDS

Non-homogeneous Burgers' equation, Cubic B-Splines, Galerkin finite element method, Von neumann analysis.

AMS Subject Classification

65M60, 65N30.

1 INTRODUCTION

In 1915 the Burgers' equation is firstly introduced by Harry Batman [2] and then in 1948 it was taken by J.M. Burger as a model of turbulent fluid motion [3]. Burgers' equation is a nonlinear partial differential equation (PDE). The solution of this equation exhibits shock wave behavior for very small value of viscosity coefficient [22]. It is a one dimensional form of Navier-Stokes equation. This equation appears in Fluid Dynamics, Gas Dynamics, Nonlinear Acoustics, Traffic Flow, etc. In this paper, we consider the one dimensional non-homogeneous Burgers' equation,

$$u_t + uu_x - \nu u_{xx} = F(x,t), \quad a \le x \le b, \quad t > 0, \tag{1}$$

with the initial condition

$$u(x,0) = f(x), \quad a \le x \le b \tag{2}$$

and the boundary conditions,

$$u(a,t) = g_1(t), \quad u(b,t) = g_2(t), \quad t > 0, \tag{3}$$

where $\nu > 0$ is the coefficient of kinematic viscosity and $f(x), g_1(t), g_2(t), F(x,t)$ are known functions. Equation (1) with $F(x,t) = 0$ is Burgers' equation. It is parabolic ($\nu \neq 0$) or hyperbolic ($\nu = 0$) in nature. In the literature it is found that there are various numerical methods which have been constructed. The Galerkin FEM based on cubic B-splines is constructed in [20] to obtain numerical solutions of Burgers' equation. This method is implicit and unconditionally stable. Galerkin finite element methods based on quadratic and cubic B-splines are constructed for equivalent system of partial differential equations in [5]. Cubic B-spline and modified cubic B-spline collocation methods are constructed in [4] and [13] respectively. In [6] and [12] least square algorithms with cubic and quadratic B-splines have been set up for Burgers equation. In [1] Burgers' equation is converted to a system of nonlinear ordinary differential equations by method of discretization in time and space variables and then Quadratic B-spline Galerkin FEM is employed on the resulting system. Weighted avarage differential quadrature method is developed in [8]. Collocation method based on Eulers implicit technique and Haar wavelets is constructed [9]. In [19] authors applied biorthogonal wavelet technique to obtain solutions of Burgers' equation.
The Galerkin FEM for (1-3) have been constructed in [22]. In this paper Taylor's series expansion is used to construct second order explicit scheme and then Galerkin FEM based on cubic B splines is set up. The numerical solutions of system (1-3) based on multi quadratic quasi-interpolation operator and radial basis function network schemes are obtained in [18]. In these methods the solution or its space derivative is quasi interpolated by using Hardy basis functions. Both the methods are conditionally stable. Stability of both the methods depends upon the shape parameters and the number of collocation points.
In the present paper Galerkin FEM have been constructed to simulate system (1-3). This method is based on cubic B-splines. For very small value of viscosity parameter (i.e. for large Reynolds number) shock is being observed in the solution of the Burgers' equation. It is important to develop the numerical techniques that will produce accurate solutions in the neighborhood of shock. Many scientists have chosen B-splines as approximation functions for the numerical solutions of Burgers' equation [1,4–6,12–14,20,22] and Benjamin-Bona-Mahony-Burgers' equation [11]. In the present work we made an attempt to obtain numerical solutions of system (1-3) via Euler time discretization. The cubic B-spline Galerkin FEM is then set up for this time discretized equation. In order to handle nonlinear term, its quasilinearization has been considered for the construction of the method.
We organize this paper in the following manner. In section 2 the Eulers implicit method is used to discretize (1) in time and then nonlinear term is linearised by quasilinearization technique. Galerkin finite element method based on cubic B splines is then applied to construct a solution. Von Neumann Stability analysis of the corresponding linearized method is discussed in section 3. Numerical solutions of some test examples obtained by proposed method are reported in section 4. These solutions are compared with exact solutions and numerical solutions available in the literature. The concluding remarks are given in section 5.

2 METHOD OF SOLUTION

We take uniform partition of the domain $[a,b]$ as $a = x_0 < x_1 < x_2 < \ldots < x_N = b$ into N number of finite elements with step length $h = \frac{b-a}{N}$ and $x_j = x_0 + jh \;; j = 0,1,2,\cdots,N$. Using Eulers time descretization, equation (1) takes the form

$$\frac{u^{n+1} - u^n}{\Delta t} = [\nu u_{xx} - u u_x + F(x,t)]^{n+1}.$$

On simplification, above equation reduces to

$$u^{n+1} - \Delta t \left[\nu u_{xx}^{n+1} - (u u_x)^{n+1} + F(x, t_{n+1})\right] = u^n, \tag{4}$$

where $t_n = t_0 + n\Delta t$, $u^n = u(x,t_n)$, $u_x^n = u_x(x,t_n)$ and $u_{xx}^n = u_{xx}(x,t_n)$. The truncation error (T.E.) in (4) is given by

$$
\begin{aligned}
T.E. &= PDE - FDE \\
&= [u_t^{n+1} + u^{n+1}u_x^{n+1} - \nu u_{xx}^{n+1} - F(x, t_{n+1})] - [\frac{u^{n+1} - u^n}{\Delta t} - \nu u_{xx}^{n+1} + u^{n+1}u_x^{n+1} - F(x, t_{n+1})] \\
&= \frac{\Delta t}{2} u_{tt}^{n+1} + \cdots
\end{aligned}
$$

Therefore (4) is consistent with equation (1) and is of order one in time domain. Assume that the solution $u(x, t)$ of the equation (1) is of the form

$$u(x,t) = \sum_{m=-1}^{N+1} \delta_m(t)\phi_m(x), \tag{5}$$

where $\delta_m(t)$; $m = -1, 0, 1, \cdots, N+1$ are the time dependent functions to be determined and $\phi_m(x)$; $m = -1, 0, 1, \cdots, N+1$ are cubic B-splines [15] given by

$$\phi_m(x) = \frac{1}{h^3}\begin{cases} (x - x_{m-2})^3, & [x_{m-2}, x_{m-1}] \\ h^3 + 3h^2(x - x_{m-1}) + 3h(x - x_{m-1})^2 - 3(x - x_{m-1})^3, & [x_{m-1}, x_m] \\ h^3 + 3h^2(x_{m+1} - x) + 3h(x_{m+1} - x)^2 - 3(x_{m+1} - x)^3, & [x_m, x_{m+1}] \\ (x_{m+2} - x)^3, & [x_{m+1}, x_{m+2}] \\ 0, & o.w. \end{cases} \tag{6}$$

Using boundary conditions (3) we obtain

$$\delta_{-1}(t) = g_1(t) - 4\delta_0(t) - \delta_1(t), \tag{7}$$

$$\delta_{N+1}(t) = g_2(t) - \delta_{N-1}(t) - 4\delta_N(t). \tag{8}$$

The solution given by equation (5) now becomes

$$u(x,t) = g_1(t)\phi_{-1}(x) + g_2(t)\phi_{N+1}(x) + \sum_{i=0}^{N} \delta_i(t)B_i(x), \tag{9}$$

where
$B_0(x) = \phi_0(x) - 4\phi_{-1}(x)$, $B_1(x) = \phi_1(x) - \phi_{-1}(x)$,
$B_j(x) = \phi_j(x)$; for $j = 2, 3, \cdots N-2$,
$B_{N-1}(x) = \phi_{N-1}(x) - \phi_{N+1}(x)$, $B_N(x) = \phi_N(x) - 4\phi_{N+1}(x)$.
From equation (9) we have

$$u_x(x,t) = g_1(t)\phi'_{-1}(x) + g_2(t)\phi'_{N+1}(x) + \sum_{i=0}^{N} \delta_i(t)B'_i(x), \tag{10}$$

$$u_{xx}(x,t) = g_1(t)\phi''_{-1}(x) + g_2(t)\phi''_{N+1}(x) + \sum_{i=0}^{N} \delta_i(t)B''_i(x). \tag{11}$$

Define

$$h_i(x,t) = \begin{cases} g_1(t)\left[\phi_{-1}(x)B_i(x)\right]'; i = 0, 1, 2, \\ g_2(t)\left[\phi_{N+1}(x)B_i(x)\right]'; i = N-2, N-1, N, \\ 0; \quad \text{otherwise}, \end{cases} \tag{12}$$

$$
\begin{aligned}
R_1(x, t_n, t_{n+1}) &= \left[g_1(t_n) - g_1(t_{n+1})\right]\phi_{-1}(x) + \Delta t \nu g_1(t_{n+1})\phi''_{-1} \\
&\quad + \left(\Delta t g_1^2(t_n) - 2\Delta t g_1(t_n)g_1(t_{n+1})\right)\phi_{-1}\phi'_{-1},
\end{aligned} \tag{13}
$$

$$S_N(x,t_n,t_{n+1}) = [g_2(t_n) - g_2(t_{n+1})]\phi_{N+1}(x) + \Delta t\nu g_2(t_{n+1})\phi''_{N+1}$$
$$+ \left(\Delta t g_2^2(t_n) - 2\Delta t g_2(t_n) g_2(t_{n+1})\right)\phi_{N+1}\phi'_{N+1}. \tag{14}$$

The nonlinear term $(uu_x)^{n+1}$ in (4) is linearized by quasilinearization technique

$$(uu_x)^{n+1} = u^n u_x^{n+1} + u^{n+1}u_x^n - u^n u_x^n. \tag{15}$$

Using equations (9)-(15) we write equation (4) element wise as follows. On the element $[x_0, x_1]$ equation (4) becomes

$$\sum_{i=0}^{2}\left[B_i(x) - \Delta t\left(\nu B_i''(x) - h_i(x,t_n) - \sum_{j=0}^{2}\delta_j(t_n)(B_iB_j)'\right)\right]\delta_i(t_{n+1})$$
$$= \sum_{i=0}^{2}\left[B_i(x) + \Delta t\left(\sum_{j=0}^{2}\delta_j(t_n)B_iB_j' + [h_i(x,t_n) - h_i(x,t_{n+1})]\right)\right]\delta_i(t_n)$$
$$+ R_1(x,t_n,t_{n+1}) + \Delta t F(x,t_{n+1}). \tag{16}$$

The term $R_1(x,t_n,t_{n+1})$ in equation (16) is contribution of non-zero function $\phi_{-1}(x)$. On the element $[x_l, x_{l+1}]$ for $l = 1, 2, ..., N-2$, equation (4) reduces to the following form.

$$\sum_{i=l-1}^{l+2}\left[B_i(x) - \Delta t\left(\nu B_i''(x) - \sum_{j=l-1}^{l+2}\delta_j(t_n)(B_iB_j)'\right)\right]\delta_i(t_{n+1})$$
$$= \sum_{i=l-1}^{l+2}\left[B_i(x) + \Delta t\sum_{j=l-1}^{l+2}\delta_j(t_n)B_iB_j'\right]\delta_i(t_n) + \Delta t F(x,t_{n+1}). \tag{17}$$

On $[x_{N-1}, x_N]$ equation (4) takes the form

$$\sum_{i=N-2}^{N}\left[B_i(x) - \Delta t\left(\nu B_i''(x) - h_i(x,t_n) - \sum_{j=N-2}^{N}\delta_j(t_n)(B_iB_j)'\right)\right]\delta_i(t_{n+1})$$
$$= \sum_{i=N-2}^{N}\left[B_i(x) + \Delta t\left(\sum_{j=N-2}^{N}\delta_j(t_n)B_iB_j' + [h_i(x,t_n) - h_i(x,t_{n+1})]\right)\right]\delta_i(t_n)$$
$$+ S_N(x,t_n,t_{n+1}) + \Delta t F(x,t_{n+1}). \tag{18}$$

The term $S_N(x,t_n,t_{n+1})$ in equation (18) is contribution of non-zero function $\phi_{N+1}(x)$. Now we obtain Galerkin weak formulation element wise. Multiplying (16) by the weight function $B_k(x); k = 0,1,2$ and integrating on the interval $[x_0, x_1]$ we get

$$[\mathbf{A_1} + \Delta t(\nu \mathbf{C_1} + \mathbf{h_1^n} - \mathbf{B_1})]\cdot \boldsymbol{\delta}_1^{n+1}$$
$$= \left[\mathbf{A_1} + \Delta t\left(\mathbf{D_1} + \mathbf{h_1^n} - \mathbf{h_1^{n+1}}\right)\right]\cdot \boldsymbol{\delta}_1^n + \mathbf{R_1^n} + \Delta t \mathbf{F_1^{n+1}}, \tag{19}$$

where $\boldsymbol{\delta}_1^n = (\delta_0^n, \delta_1^n, \delta_2^n)^T$,

$$R_1^n = \frac{h[g_1(n\Delta t) - g_1((n+1)\Delta t)]}{140}\begin{bmatrix}49\\40\\1\end{bmatrix} + \frac{\nu\Delta t g_1((n+1)\Delta t)}{10h}\begin{bmatrix}51\\54\\3\end{bmatrix}$$

$$+\frac{\Delta t\left[2g_1(n\Delta t)g_1((n+1)\Delta t) - g_1^2(n\Delta t)\right]}{168}\begin{bmatrix}97\\70\\1\end{bmatrix},$$

$$F_1^{n+1} = \begin{bmatrix}\int_{x_0}^{x_1} F(x,(n+1)\Delta t)B_0(x)dx\\ \int_{x_0}^{x_1} F(x,(n+1)\Delta t)B_1(x)dx\\ \int_{x_0}^{x_1} F(x,(n+1)\Delta t)B_2(x)dx\end{bmatrix}.$$

$$A_1 = \frac{h}{140}\begin{bmatrix}476 & 644 & 56\\ 644 & 1088 & 128\\ 56 & 128 & 20\end{bmatrix}, \quad C_1 = \frac{1}{10h}\begin{bmatrix}222 & 108 & -24\\ 108 & 192 & 24\\ -24 & 24 & 18\end{bmatrix},$$

$$h_1^n = \frac{-g_1(n\Delta t)}{840}\begin{bmatrix}1235 & 1586 & 89\\ 758 & 1244 & 98\\ -1 & 26 & 5\end{bmatrix},$$

and for $i,j = 1,2,3$; $(i,j)^{th}$ elements of matrices B_1 and D_1 are given by

$$(B_1)_{ij} = \left(\int_{x_0}^{x_1} B_{j-1}B_0B'_{i-1}dx, \int_{x_0}^{x_1} B_{j-1}B_1B'_{i-1}dx, \int_{x_0}^{x_1} B_{j-1}B_2B'_{i-1}dx\right)\cdot\delta_1^n,$$

$$(D_1)_{ij} = \left(\int_{x_0}^{x_1} B_{j-1}B'_0B_{i-1}dx, \int_{x_0}^{x_1} B_{j-1}B'_1B_{i-1}dx, \int_{x_0}^{x_1} B_{j-1}B'_2B_{i-1}dx\right)\cdot\delta_1^n.$$

Multiply (17) by the weight function $B_k(x)$; $k = l-1, l, l+1, l+2$ and integrating on the interval $[x_l, x_{l+1}]$ we obtain

$$[A_{l+1} + \Delta t\,(\nu C_{l+1} - B_{l+1})]\cdot\delta_{l+1}^{n+1} = [A_{l+1} + \Delta t D_{l+1}]\cdot\delta_{l+1}^n + \Delta t F_{l+1}^{n+1}, \tag{20}$$

where for $l = 1,2,\cdots,N-2$; $\delta_{l+1}^n = (\delta_{l-1}^n, \delta_l^n, \delta_{l+1}^n, \delta_{l+2}^n)^T$,

$$F_{l+1}^{n+1} = \begin{bmatrix}\int_{x_l}^{x_{l+1}} F(x,(n+1)\Delta t)B_{l-1}(x)dx\\ \int_{x_l}^{x_{l+1}} F(x,(n+1)\Delta t)B_l(x)dx\\ \int_{x_l}^{x_{l+1}} F(x,(n+1)\Delta t)B_{l+1}(x)dx\\ \int_{x_l}^{x_{l+1}} F(x,(n+1)\Delta t)B_{l+2}(x)dx\end{bmatrix},$$

$$A_{l+1} = \frac{h}{140}\begin{bmatrix}20 & 129 & 60 & 1\\ 129 & 1188 & 933 & 60\\ 60 & 933 & 1188 & 129\\ 1 & 60 & 129 & 20\end{bmatrix}, \quad C_{l+1} = \frac{1}{10h}\begin{bmatrix}18 & 21 & -36 & -3\\ 21 & 102 & -87 & -36\\ -36 & -87 & 102 & 21\\ -3 & -36 & 21 & 18\end{bmatrix}$$

and for $i, j = 1, 2, 3, 4$; $(i, j)^{th}$ element of matrices $\boldsymbol{B_{l+1}}$ and $\boldsymbol{D_{l+1}}$ are computed by

$$(\boldsymbol{B_{l+1}})_{ij} = \left(\int_{x_l}^{x_{l+1}} B_{j+l-2} B_{l-1} B'_{i+l-2} dx, \int_{x_l}^{x_{l+1}} B_{j+l-2} B_l B'_{i+l-2} dx, \right.$$
$$\left. \int_{x_l}^{x_{l+1}} B_{j+l-2} B_{l+1} B'_{i+l-2} dx, \int_{x_l}^{x_{l+1}} B_{j+l-2} B_{l+2} B'_{i+l-2} dx \right) \cdot \delta^{\boldsymbol{n}}_{\boldsymbol{l+1}}.$$

$$(\boldsymbol{D_{l+1}})_{ij} = \left(\int_{x_l}^{x_{l+1}} B_{j+l-2} B'_{l-1} B_{i+l-2} dx, \int_{x_l}^{x_{l+1}} B_{j+l-2} B'_l B_{i+l-2} dx, \right.$$
$$\left. \int_{x_l}^{x_{l+1}} B_{j+l-2} B'_{l+1} B_{i+l-2} dx, \int_{x_l}^{x_{l+1}} B_{j+l-2} B'_{l+2} B_{i+l-2} dx \right) \cdot \delta^{\boldsymbol{n}}_{\boldsymbol{l+1}}.$$

On multiplying (18) by the weight function $B_k(x); k = N-2, N-1, N$ and integrating on the interval $[x_{N-1}, x_N]$ we get

$$\begin{aligned} &[\boldsymbol{A_N} + \Delta t\,(\nu \boldsymbol{C_N} + \boldsymbol{h^n_N} - \boldsymbol{B_N})] \cdot \boldsymbol{\delta^{n+1}_N} \\ &\quad = \left[\boldsymbol{A_N} + \Delta t\left(\boldsymbol{D_N} + \boldsymbol{h^n_N} - \boldsymbol{h^{n+1}_N}\right)\right] \cdot \boldsymbol{\delta^n_N} + \boldsymbol{S^n_N} + \Delta t \boldsymbol{F^{n+1}_N}, \end{aligned} \tag{21}$$

where $\boldsymbol{\delta^n_N} = (\delta^n_{N-2}, \delta^n_{N-1}, \delta^n_N)^T$,

$$\boldsymbol{S^n_N} = \frac{h(g_2(n\Delta t) - g_2((n+1)\Delta t))}{140} \begin{bmatrix} 1 \\ 40 \\ 49 \end{bmatrix} + \frac{\nu \Delta t g_2((n+1)\Delta t)}{10h} \begin{bmatrix} 3 \\ 54 \\ 51 \end{bmatrix}$$

$$- \frac{\Delta t \left[2 g_2(n\Delta t) g_2((n+1)\Delta t) - g_2^2(n\Delta t)\right]}{168} \begin{bmatrix} 1 \\ 70 \\ 97 \end{bmatrix},$$

$$\boldsymbol{F^{n+1}_N} = \Delta t \begin{bmatrix} \int_{x_{N-1}}^{x_N} F(x, (n+1)\Delta t) B_{N-2}(x) dx \\ \int_{x_{N-1}}^{x_N} F(x, (n+1)\Delta t) B_{N-1}(x) dx \\ \int_{x_{N-1}}^{x_N} F(x, (n+1)\Delta t) B_N(x) dx \end{bmatrix},$$

$$\boldsymbol{A_N} = \frac{h}{140} \begin{bmatrix} 20 & 128 & 56 \\ 128 & 1088 & 644 \\ 56 & 644 & 476 \end{bmatrix}, \quad \boldsymbol{C_N} = \frac{1}{10h} \begin{bmatrix} 18 & 24 & -24 \\ 24 & 192 & 108 \\ -24 & 108 & 222 \end{bmatrix},$$

$$\boldsymbol{h^n_N} = \frac{g_2(n\Delta t)}{840} \begin{bmatrix} 5 & 26 & -1 \\ 98 & 1244 & 758 \\ 89 & 1586 & 1235 \end{bmatrix},$$

and for $i, j = 1, 2, 3; (i, j)^{th}$ element of matrices $\boldsymbol{B_N}$ and $\boldsymbol{D_N}$ are given by

$$(\boldsymbol{B_N})_{ij} = \left(\int_{x_{N-1}}^{x_N} B_{j+N-3} B_{N-2} B'_{i+N-3} dx, \int_{x_{N-1}}^{x_N} B_{j+N-3} B_{N-1} B'_{i+N-3} dx, \int_{x_{N-1}}^{x_N} B_{j+N-3} B_N B'_{i+N-3} dx \right) \cdot \boldsymbol{\delta_N^n},$$

$$(\boldsymbol{D_N})_{ij} = \left(\int_{x_{N-1}}^{x_N} B_{j+N-3} B'_{N-2} B_{i+N-3} dx, \int_{x_{N-1}}^{x_N} B_{j+N-3} B'_{N-1} B_{i+N-3} dx, \int_{x_{N-1}}^{x_N} B_{j+N-3} B'_N B_{i+N-3} dx \right) \cdot \boldsymbol{\delta_N^n}.$$

Computation of matrices $\boldsymbol{C_l}$ and $\boldsymbol{B_l}$ for $l = 1, 2, \cdots, N$ is done using integration by parts. Since for $l = 0, 1, \cdots, N-1$ the overall contribution of the terms $B_i(x)B_j(x)B_k(x)|_{x_l}^{x_{l+1}}$ and $B'_i(x)B_k(x)|_{x_l}^{x_{l+1}}$ vanishes in the assembled system, we have excluded them from the final expression. Since $\phi_{-1}(x)$ is zero on $[x_l, x_{l+1}]; l = 1, 2, \cdots, N-1$, $\boldsymbol{R_k^n}; k = 2, 3, ..., N$ are zero vectors. Similarly $\phi_{N+1}(x)$ is zero on $[x_l, x_{l+1}]; l = 0, 1, ..., N-2$, and therefore $\boldsymbol{S_k^n}; k = 1, 2, ..., N-1$ are zero vectors. Also for $k = 2, 3, \cdots N-1$; $\boldsymbol{h_k^n}$ are zero matrices. Combining the contributions from elemental equations (19), (20) and (21) in usual way we obtain the following $(N+1) \times (N+1)$ system.

$$[\boldsymbol{A} + \Delta t (\nu \boldsymbol{C} + \boldsymbol{h^n} - \boldsymbol{B})] \cdot \boldsymbol{\delta^{n+1}} = \left[\boldsymbol{A} + \Delta t \left(\boldsymbol{D} + \boldsymbol{h^n} - \boldsymbol{h^{n+1}}\right)\right] \cdot \boldsymbol{\delta^n} + \boldsymbol{R^n} + \boldsymbol{S^n} + \Delta t \boldsymbol{F^{n+1}}, \quad (22)$$

3 STABILITY ANALYSIS

The stability of linearized scheme corresponding to constructed scheme (22) is analyzed by von Neumann analysis. The linearized form of (22) is obtained by assuming that the solution u in the nonlinear term uu_x is locally constant and is equal to U. Thus the linear system corresponding to scheme (22) is obtained as follows

$$\left[\boldsymbol{A} + U\Delta t \mathbf{B}^* + \nu \Delta t \boldsymbol{C}\right] \cdot \delta^{n+1} = \left[\boldsymbol{A} + U \Delta t \mathbf{B}^*\right] \cdot \delta^n + \boldsymbol{R}^n + \boldsymbol{S}^n + \boldsymbol{F}^{n+1}, \quad (23)$$

where $\mathbf{B}^*$ is obtained by combining contributions from $\int_{x_l}^{x_{l+1}} B'_i(x) B_j(x) dx$ in usual way. The error equation corresponding to the above equation is

$$\left[\boldsymbol{A} + U\Delta t \mathbf{B}^* + \nu \Delta t \boldsymbol{C}\right] \cdot \epsilon^{n+1} = \left[\boldsymbol{A} + U \Delta t \mathbf{B}^*\right] \cdot \epsilon^n, \quad (24)$$

where, ϵ^n is error in the solution at $t = t_n$. Matrices $\boldsymbol{A}$, $\boldsymbol{B}^*$ and $\boldsymbol{C}$ are septadiagonal matrices and general row of these matrices are

$$\boldsymbol{A} : \frac{h}{140}(1, 120, 1191, 2416, 1191, 120, 1)$$
$$\boldsymbol{B}^* : \frac{1}{20}(1, 56, 245, 0, -245, -56, -1)$$
$$\boldsymbol{C} : \frac{-1}{10h}(3, 72, 45, -240, 45, 72, 3)$$

The l^{th} error equation in (24) is

$$\begin{aligned} &\alpha_1 \epsilon_{l-3}^{n+1} + \alpha_2 \epsilon_{l-2}^{n+1} + \alpha_3 \epsilon_{l-1}^{n+1} + \alpha_4 \epsilon_l^{n+1} + \alpha_5 \epsilon_{l+1}^{n+1} + \alpha_6 \epsilon_{l+2}^{n+1} + \alpha_7 \epsilon_{l+3}^{n+1} \\ &= \alpha_8 \epsilon_{l-3}^n + \alpha_9 \epsilon_{l-2}^n + \alpha_{10} \epsilon_{l-1}^n + \alpha_{11} \epsilon_l^n + \alpha_{12} \epsilon_{l+1}^n + \alpha_{13} \epsilon_{l+2}^n + \alpha_{14} \epsilon_{l+3}^n, \end{aligned} \quad (25)$$

where ϵ_j^n is the error in δ_j at $t = n\Delta t$, $0 \leq j \leq N, 4 \leq l \leq N-3$,
$\alpha_1 = r_1 + r_2 - 3r_3$, $\alpha_2 = 120r_1 + 56r_2 - 72r_3$, $\alpha_3 = 1191r_1 + 245r_2 - 45r_3$, $\alpha_4 = 2416r_1 + 240r_3$,
$\alpha_5 = 1191r_1 - 245r_2 - 45r_3$, $\alpha_6 = 120r_1 - 56r_2 - 72r_3$, $\alpha_7 = r_1 - r_2 - 3r_3$, $\alpha_8 = r_1 + r_2$,
$\alpha_9 = 120r_1 + 56r_2$, $\alpha_{10} = 1191r_1 + 245r_2$, $\alpha_{11} = 2416r_1$, $\alpha_{12} = 1191r_1 - 245r_2$,
$\alpha_{13} = 120r_1 - 56r_2$, $\alpha_{14} = r_1 - r_2$, $r_1 = \frac{h}{140}$, $r_2 = \frac{U\Delta t}{20}$, $r_3 = \frac{\nu\Delta t}{10h}$.
The Fourier growth factor is defined as

$$\epsilon_l^n = \xi^n e^{ilkh}, \tag{26}$$

where k is mode number and h is the length of finite element. Using (26) equation (25) gives

$$[(a+b) + i(c + 240r_3)]\xi = [a + ic], \tag{27}$$

where

$$\begin{aligned}
a &= [2\cos(3kh) + 240\cos(2kh) + 2382\cos(kh) + 2416]r_1, \\
b &= [-6\cos(3kh) - 144\cos(2kh) - 90\cos(kh) + 240]r_3, \\
c &= 2416r_1 - [2\sin(3kh) + 112\sin(2kh) + 490\sin(kh)]r_2.
\end{aligned}$$

From (27) the amplification factor

$$\xi = \frac{a + ic}{(a + ic) + (b + i240r_3)}.$$

It is observed that if $(b + 240r_3)^2 + 2ab + 480r_3(c - b) \geq 0$ then $|\xi| \leq 1$ and hence linearized scheme (23) is conditionally stable. Since method is consistent and stable, by Lax equivalence theorem the method is convergent.

4 NUMERICAL RESULTS AND DISCUSSIONS

In this section we illustrate some test examples to support the method. Mathematica 10.0 software is used to compute numerical solution and error in the solution. The L_2 and L_∞ errors are defined as,

$$\begin{aligned}
L_2 &= \sqrt{h\sum_{j=0}^{N} |U_j^{exact} - u_j^{numer}|^2}, \\
L_\infty &= \max_{1\leq j\leq N} |U_j^{exact} - u_j^{numer}|.
\end{aligned}$$

where, U_j^{exact} and u_j^{numer} are exact and numerical solutions at $x = x_j$ respectively.
Example 1: In this test example we consider equation (1) with the initial condition $u(x, 0) = \sin(\pi x)$, boundary conditions $u(0, t) = 0$, $u(1, t) = 0$ and $F(x, t) = 0$. The exact solution of this problem is

$$u(x,t) = 2\pi\nu \frac{\sum_{n=1}^{\infty} a_n e^{-n^2\pi^2\nu t} n \sin(n\pi x)}{a_0 + \sum_{n=1}^{\infty} a_n e^{-n^2\pi^2\nu t} n \cos(n\pi x)},$$

where the Fourier coefficients a_n; $n = 0, 1, 2, \cdots$ are given by

$$\begin{aligned}
a_0 &= \int_0^1 e^{-(2\pi\nu)^{-1}(1-\cos(\pi x))} dx, \\
a_n &= 2\int_0^1 e^{-(2\pi\nu)^{-1}(1-\cos(\pi x))} \cos(n\pi x) dx.
\end{aligned}$$

The numerical solutions obtained by the proposed method, solutions obtained in [9] and the exact solution for $\nu = 0.003$ are shown in Table 1. The numerical solutions and contour for $\nu = 0.1, \Delta t = 0.001, N = 40$ are plotted in Fig 1 (a) and Fig 1 (b) respectively. Obtained numerical solutions by present method for $\nu = 0.001, \Delta t = 0.001, N = 400$ are depicted in Fig 2. It is seen that proposed method has ability to capture shocks. From Table 1 it is observed that the produced solutions by the proposed method are less accurate than solutions obtained in [9] and exact solutions. It is seen in Fig 1 that the physical behavior of obtained solution is correct.

Example 2: In this example we consider (1) with initial condition $u(x,0) = 4x(1-x)$ and boundary conditions $u(0,t) = 0$; $u(1,t) = 0$ with $F(x,t) = 0$. The numerical solution obtained by the proposed method, solution obtained in [1, 12, 13, 22] and exact solution for $\nu = 0.01$ are listed in Table 2. It is observed that the numerical solutions obtained by the proposed method are better than the solutions obtained in [12] even for small value of N but they are less accurate than solutions obtained in [1,13,22]. The numerical solutions for $\nu = 0.1, \Delta t = 0.001, N = 80$ and $\nu = 0.01, \Delta t = 0.001, N = 160$ are depicted in Fig 3.

Example 3: In this example we consider (1) with the initial condition $u(x,0) = \frac{2\nu\pi\sin(\pi x)}{\alpha+\cos(\pi x)}$; $\alpha > 1$, boundary conditions $u(0,t) = 0$; $u(1,t) = 0$ and $F(x,t) = 0$. The exact solution of this problem is

$$u(x,t) = \frac{2\nu\pi e^{-\pi^2\nu t}\sin(\pi x)}{\alpha + e^{-\pi^2\nu t}\cos(\pi x)}.$$

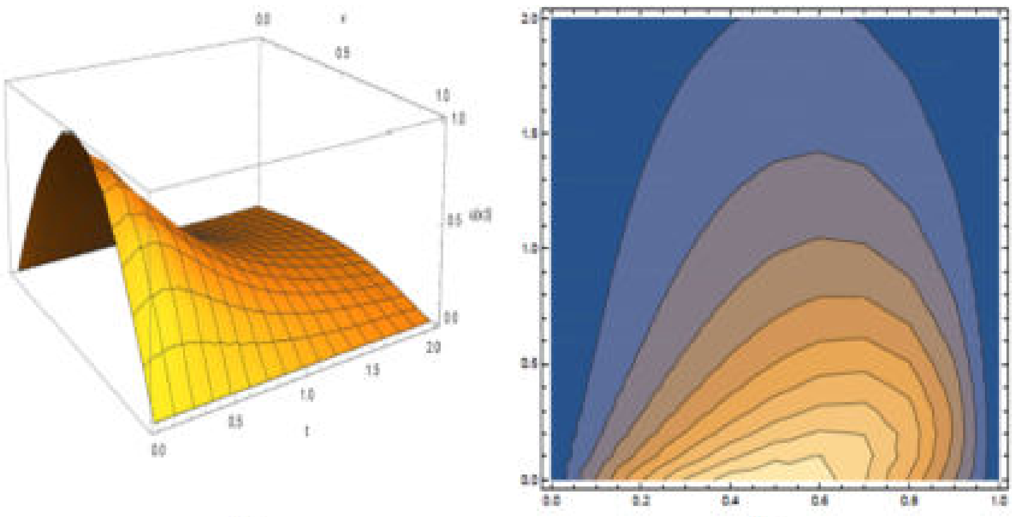
The numerical solutions and L_2, L_∞ errors for $\alpha = 2, h = 0.025, \nu = 1.0, 0.5, 0.2, 0.1$ and $\Delta t = 0.0001$ at $t = 0.001$ are presented in Table 3 and Table 4. From Table 3 and Table 4, it is observed that the numerical solutions obtained by the present method are slightly less accurate than solutions obtained in [22]. The plots of absolute error and numerical solution for $\alpha = 2, \nu = 0.01, \Delta t = 0.001$ and $N = 100$ are shown in Fig 4 (a) and Fig 4 (b) respectively. L_∞ errors are obtained for $\alpha = 3, \nu = 0.3, \Delta t = 0.001, N = 80$ and are presented in Table 5. It is seen that L_∞ error decreases as time increases and hence error is bounded. Thus proposed method produces stable solutions.

Example 4: Consider equation (1) with the initial condition $u(x,1) = \frac{x}{1+e^{\frac{1}{4\nu}(x^2-\frac{1}{4})}}$ and boundary conditions $u(0,t) = u(1.2,t) = 0$ and $F(x,t) = 0$. The exact solution of this example is $u(x,t) = \frac{(\frac{x}{t})}{1+(\frac{t}{t_0})^{\frac{1}{2}}e^{\frac{x^2}{4\nu t}}}$ where $t_0 = e^{\frac{1}{8\nu}}$. The comparison of numerical solutions with exact solutions for $h = 0.005, \nu = 0.005$ and $\Delta t = 0.001$ is given in the Table 6. The L_2 and L_∞ errors are computed for $\nu = 0.005, h = 0.005$ and $\Delta t = 0.001$ at different time levels and their comparison with [13, 22] is shown in Table 7. Tables 6-7 shows that present method produces accurate solutions even for small value of ν but for very small value of Δt.

Example 5: Consider Burgers' equation (1) with $F(x,t) = \frac{kx}{(2\beta t+1)^2}$; $k > 0, \beta \geq 0$ and the initial condition $u(x,0) = kx$; $k > \beta$. The exact solution of this problem for $\nu = 1$ is obtained by Rao and Yadav [16] as follows

$$u(x,t) = \frac{A_0 x}{(2\beta t+1)}, \quad \text{where} \quad A_0 = \beta + \sqrt{\beta^2 + k}.$$

The boundary conditions are considered from the exact solution. In the computation region $[-1,1]$, L_2 and L_∞ errors in the numerical solutions obtained by proposed method for $k = 5$, $\beta = 2$ are listed in Table 8 and are compared with errors in [22]. It is seen that numerical solutions obtained by the present method are less accurate than solutions given in [22].



(a) (b)

Figure 1: (a) Numerical solutions of Ex 1. for $\nu = 0.1, \Delta t = 0.001, N = 40$, (b) Contour plot of Ex 1. for $\nu = 0.1, \Delta t = 0.001, N = 40, 0 \leq x \leq 1, 0 \leq t \leq 2$.

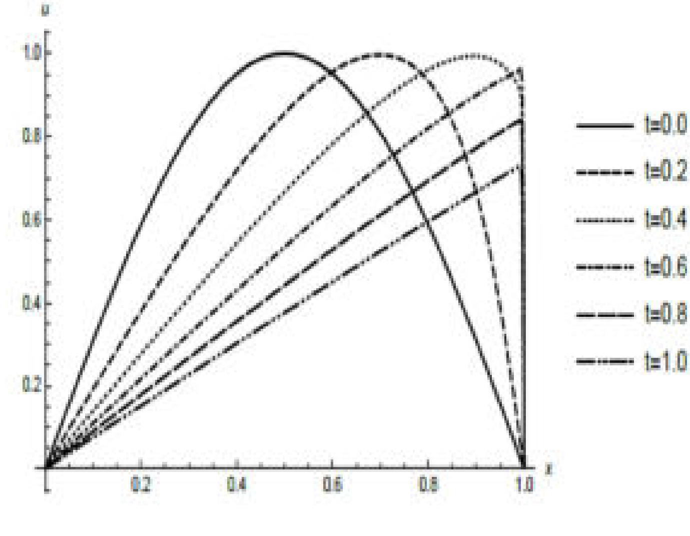


Figure 2: Numerical solutions of Ex 1 for $\nu = 0.001, \Delta t = 0.001, N = 400$.

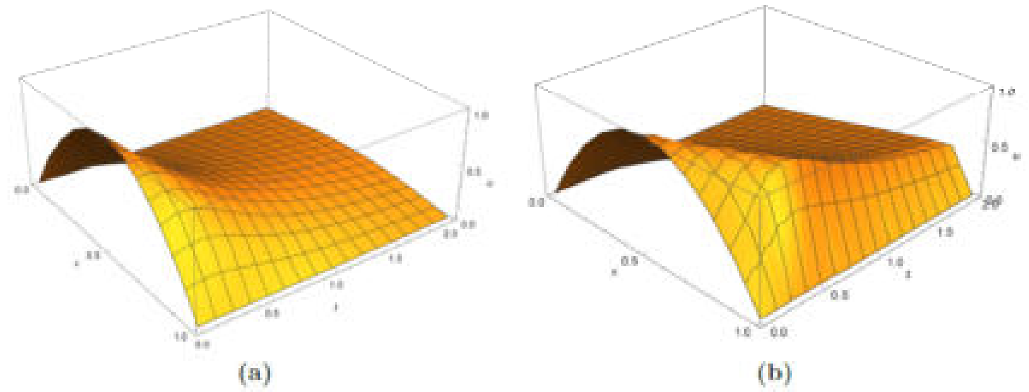


Figure 3: Numerical solutions of Ex 2 (a) for $\nu = 0.1, \Delta t = 0.001$ and $N = 80$, (b) for $\nu = 0.01, \Delta t = 0.001$ and $N = 160$.

Table 1: Comparison of numerical and exact solutions for $\nu = 0.003$, $\Delta t = 0.001$ and $N = 64$ of Ex 1.

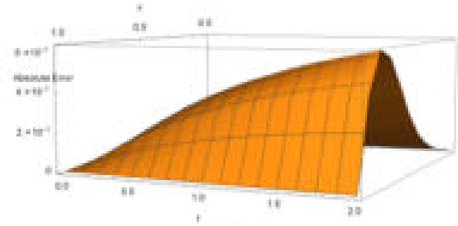
x	t	[9]	Present	Exact
0.25	1	0.18902	0.18922	0.18901
	5	0.04698	0.04701	0.04698
	10	0.02422	0.02423	0.02422
	15	0.01631	0.01632	0.01631
0.50	1	0.37623	0.37659	0.37619
	5	0.09396	0.09400	0.09395
	10	0.04844	0.04846	0.04843
	15	0.03263	0.03264	0.03263
0.75	1	0.55928	0.55967	0.55924
	5	0.14092	0.14099	0.14095
	10	0.07261	0.07263	0.07260
	15	0.04839	0.04840	0.04841

5 CONCLUSION

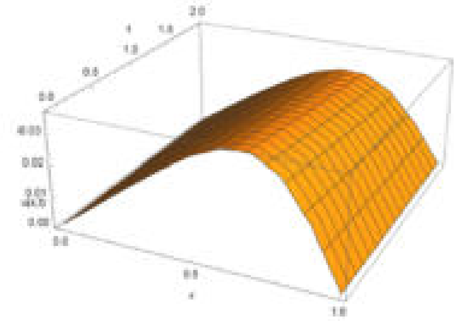
The cubic B-spline Galerkin finite element method is successfully implemented to the one dimensional non-homogeneous Burgers' equation. Burgers' equation is discretized in time by using Eulers implicit technique and then Galerkin finite element method is constructed for the discretized equation. The nonlinear term is linearized by quasilinearization technique. Five numerical test examples are solved to support the proposed method. It is seen that the method is reliable for solving the one dimensional non-homogeneous Burgers' equation but computational cost is more.

ACKNOWLEDGMENT

Author greatly acknowledges Research Grant received from **Shivaji University, Kolhapur** under Research Initiation Scheme (University Reference No.: SU/C&U.D.Section/70/214 dated 7 May 2018) for the present work. Author sincerely acknowledges and thanks to Prof. S. H. Thakar for the valuable discussion. Author extends sincere thanks to reviewer for careful and thoroughly reading of the paper and for valuable suggestions and comments which indeed improve the quality of the paper.



(a)



(b)

Figure 4: (a) Absolute error in numerical solutions of Ex 3 for $\alpha = 2, \nu = 0.01, \Delta t = 0.001$ and $N = 100$, (b) Numerical solutions of Ex 3 for $\alpha = 2, \nu = 0.01, \Delta t = 0.001$ and $N = 100$.

Table 2: Comparison of numerical and exact solution for $\nu = 0.01$ Ex 2.

x	t	[13] $\Delta t = 0.001$ $N = 40$	[1] $\Delta t = 0.0001$ $N = 40$	[12] $\Delta t = 0.0001$ $N = 80$	[22] $\Delta t = 0.001$ $N = 40$	Present $\Delta t = 0.0001$ $N = 60$	Exact
0.25	0.4	0.36225	0.36225	0.36218	0.36226	0.36230	0.36226
	0.6	0.28202	0.28199	0.28197	0.28204	0.28207	0.28204
	0.8	0.23044	0.23039	0.23040	0.23045	0.23048	0.23045
	1.0	0.19468	0.19463	0.19465	0.19469	0.19471	0.19469
	3.0	0.07613	0.07611	0.07613	0.07613	0.07614	0.07613
0.50	0.4	0.68368	0.68371	0.68364	0.68368	0.68371	0.68368
	0.6	0.54832	0.54835	0.54829	0.54832	0.54836	0.54832
	0.8	0.45371	0.45374	0.45368	0.45371	0.45376	0.45371
	1.0	0.38567	0.38568	0.38564	0.38568	0.38571	0.38568
	3.0	0.15218	0.15216	0.15217	0.15218	0.15219	0.15218
0.75	0.4	0.92052	0.92047	0.92047	0.92044	0.92045	0.92050
	0.6	0.78300	0.78302	0.78297	0.78288	0.78301	0.78299
	0.8	0.66272	0.66276	0.66270	0.66267	0.66276	0.66272
	1.0	0.56932	0.56936	0.56930	0.56931	0.56936	0.56933
	3.0	0.22782	0.22773	0.22773	0.22774	0.22776	0.22774

Table 3: Comparison of numerical and exact solutions for $\alpha = 2, h = 0.025$ and $\Delta t = 0.0001$ at $t = 0.001$ of Ex 3.

x	$\nu = 1$			$\nu = 0.5$		
	Present	[22]	Exact	Present	[22]	Exact
0.1	0.653545	0.653544	0.653544	0.327870	0.327870	0.327870
0.2	1.305540	1.305533	1.305534	0.655069	0.655069	0.655069
0.3	1.949370	1.949363	1.949364	0.978413	0.978412	0.978413
0.4	2.565930	2.565924	2.565925	1.288460	1.288464	1.288463
0.5	3.110750	3.110738	3.110739	1.563070	1.563063	1.563064
0.6	3.492890	3.492665	3.492866	1.756650	1.756642	1.756642
0.7	3.549640	3.549595	3.549595	1.787210	1.787206	1.787206
0.8	3.050200	3.050138	3.050134	1.537700	1.537696	1.537696
0.9	1.816710	1.816666	1.816660	0.916869	0.916863	0.916860
$L_\infty \times 10^4$	0.623	0.056	-	0.099	0.030	-
$L_2 \times 10^4$	0.306	0.021	-	0.045	0.011	-

Table 4: Comparison of numerical and exact solutions for $\alpha = 2, h = 0.025$ and $\Delta t = 0.0001$ at $t = 0.001$ of Ex 3.

x	$\nu = 0.2$			$\nu = 0.1$		
	Present	[22]	Exact	Present	[22]	Exact
0.1	0.131412	0.131412	0.131412	0.065750	0.065750	0.065750
0.2	0.262581	0.262581	0.262581	0.131383	0.131383	0.131383
0.3	0.392262	0.392262	0.392262	0.196281	0.196281	0.196281
0.4	0.516709	0.516709	0.516710	0.258576	0.258576	0.258576
0.5	0.627079	0.627079	0.627079	0.313850	0.313849	0.313849
0.6	0.705120	0.705120	0.705120	0.352972	0.352972	0.352972
0.7	0.717883	0.717882	0.717882	0.359443	0.359443	0.359443
0.8	0.618138	0.618137	0.618136	0.309581	0.309581	0.309580
0.9	0.368815	0.368815	0.368814	0.184754	0.184754	0.184754
$L_\infty \times 10^5$	0.164	0.123	-	0.068	0.063	-
$L_2 \times 10^6$	0.642	0.454	-	0.251	0.229	-

Table 6: Comparison of numerical and exact solutions for $\nu = 0.005$ and $h = 0.005$ of Ex 4.

x	t	Present $\Delta t = 0.0001$	[22] $\Delta t = 0.001$	Exact
0.2	1.7	0.1176490	0.1176452	0.1176452
	2.5	0.0800019	0.0799990	0.0799990
	3.0	0.0666683	0.0666658	0.0666658
	3.5	0.0571422	0.0571422	0.0571422
0.4	1.7	0.2351750	0.2351677	0.2351677
	2.5	0.1599830	0.1599769	0.1599769
	3.0	0.1333260	0.1333209	0.1333209
	3.5	0.1142820	0.1142779	0.1142779
0.6	1.7	0.2958750	0.2959101	0.2959097
	2.5	0.2381270	0.2381207	0.2381207
	3.0	0.1994870	0.1994806	0.1994805
	3.5	0.1712300	0.1712242	0.1712242
0.8	1.7	0.0006480	0.0006465	0.0006465
	2.5	0.1020790	0.1020955	0.1020957
	3.0	0.2088120	0.2088360	0.2088359
	3.5	0.2145830	0.2145869	0.2145869

Table 5: L_∞ errors for $\alpha = 3, \nu = 0.3, \Delta t = 0.001$ and $N = 80$ of Ex 3.

t	$L_\infty \times 10^3$	t	$L_\infty \times 10^3$	t	$L_\infty \times 10^3$	t	$L_\infty \times 10^3$
0.1	0.297	0.6	0.293	1.1	0.119	1.6	0.039
0.2	0.379	0.7	0.252	1.2	0.097	1.7	0.031
0.3	0.389	0.8	0.213	1.3	0.078	1.8	0.024
0.4	0.368	0.9	0.178	1.4	0.062	1.9	0.019
0.5	0.334	1.0	0.146	1.5	0.050	2.0	0.015

Table 7: Comparison of L_2 and L_∞ errors for $\nu = 0.005$ of Ex 4.

t	Present $\Delta t = 0.0001, h = 0.005$		[13] $\Delta t = 0.001, h = 0.005$		[22] $\Delta t = 0.001, h = 0.005$	
	$L_\infty \times 10^4$	$L_2 \times 10^4$	$L_\infty \times 10^4$	$L_2 \times 10^4$	$L_\infty \times 10^4$	$L_2 \times 10^4$
1.7	0.457	0.117	0.994	0.252	0.006	0.0017
2.5	0.350	0.100	0.549	0.151	0.002	0.0008
3.0	0.299	0.090	0.414	0.118	0.023	0.0029
3.5	0.572	0.111	0.486	0.117	0.572	0.0754

Table 8: Comparison of L_2 and L_∞ errors for $k = 5, \beta = 2, \Delta t = 0.01$ and $N = 10$ of Ex 5.

	L_∞		L_2	
	$t = 5$	$t = 10$	$t = 5$	$t = 10$
Present	5.80×10^{-6}	7.61×10^{-7}	5.89×10^{-6}	7.73×10^{-7}
[22]	2.811×10^{-9}	1.872×10^{-10}	2.854×10^{-9}	1.901×10^{-10}

References

[1] Aksan, E.N. (2006) Applied Mathematics and Computation, 174, 884-896.

[2] Bateman, H. (1915) Monthly Weather Review, 43, 163–170.

[3] Burgers', J. M. (1948) Advances in Applied Mechanics, 1, 171–199.

[4] Dag, I., Irk, D., Saka B. (2005) Applied Mathematics and Computation, 163, 199-211.

[5] Dag, I., Saka, B., Boz, A. (2005) Applied Mathematics and Computation, 166, 506-522.

[6] Dhawan, S., Kumar, S. (2011) Proceedings of the 13th IASME/WSEAS international conference on Mathematical Methods and Computational Techniques in Electrical Engineering conference on Applied Computing, 40-43.

[7] Jiwari R. (2012) Computer Physics Communications, 183, 2413–2423.

[8] Jiwari, R., Mittal, R.C., Sharma K. K. (2013) Applied Mathematics and Computation, 219, 6680-6691.

[9] Jiwari R. (2015) Computer Physics Communications, 188, 59–67.

[10] Karabutov, A. A., Lapshin, E. A., Rudenko, O. V. (1976) Journal of Experimental and Theoretical Physics, 44, 58-63.

[11] Kumbhar, S. S. (2023) Mathematics in Engineering, Science and Aerospace, 14, 19-32.

[12] Kutluay, S., Esen, A., Dag, I. (2004) Journal of Computational and Applied Mathematics, 167, 21-33.

[13] Mittal, R.C., Jain, R. (2012) Applied Mathematics and Computation, 218, 7839-7855.

[14] Mittal, R.C., Rohila, R. (2018) International Journal of Applied and Computational Mathematics, 4, 1-15.

[15] Prenter, P. M., Splines and Variational Methods, Wiley, (1975).

[16] Rao, Ch. S., Yadav, M. K. (2010) International Journal of Nonlinear Sciences and Numerical Simulation, 10, 141-145.

[17] Rathish Kumar, B. V., Mehra, M. (2005) BIT Numerical Mathematics, 45, 543-560.

[18] Sarboland, M., Aminataei, A. (2014) Journal of Applied Mathematics, 1-15.

[19] Shiralashetti, S.C., Angadi, L.M., Deshi, A.B. (2019) International Journal of Computational Materials Science and Engineering, 8 , p.1850030.

[20] Soliman, A. A. (2012) Abstract and Applied Analysis, 1-15.

[21] Tannehill, J. C., Anderson D. A., Pletcher, R. H., Computational Fluid Mechanics and Heat Transfer, Taylor and Francis, Philadelphia, (2010).

[22] Thakar, S., Kumbhar, S. S. (2019) Journal of Mathematical Modeling, 7, 445-467.

A Study of Gender Inequality According to Various Aspects

Prasad V. Shinde[a], Amol M. Ghadge[a]*
[a]Biostatistician, Research Guidance Cell
Dr. D. Y. Patil Medical College, Hospital and Research Institute, Kolhapur.
*Corresponding author: amolghadge410@gmail.com

ABSTRACT
Gender inequality has been a social issue in India for centuries. Indian society has always been discriminating amongst men and women. Gender equality is just an oral statement said by the influential people of the society but never seen in practice. Be it a household work, labour work or politics, women have always been underestimated about their efficiency and capabilities to work equally in comparison to men. In this paper, Authors studied that the gender inequality in accordance to male female ratio specifically related to birth rate, life expectancy, work participation and literacy.

KEYWORDS
Inequality, literacy, Work force participation, Life expectancy, Interpolation and extrapolation.

…………………………………………………………………………………………….

1. INTRODUCTION

Gender equality is not only a fundamental human right, but a necessary foundation to create a peaceful, prosperous and sustainable world to live in Gender inequality means that men and women are may be equal or not equal and it affects an individual's living experience. Gender discrimination has been a social issue in India for years and years. That in many parts of India, the birth of a girl child is not welcomed and it is a known fact. The report titled "Women, business and the law" revealed the world is moving to equal legal rights to both genders. Only 6 countries in the world give women the same legal work rights as men and prove to be gender equal [1]. Still women in many parts of the country have less access to education than men or are being denied from taking education. Many times, it has been seen that they are not even allowed to finish their primary education, which is a basic human right. Gender inequality can be seen even at the workplaces. Still women are being harassed mentally and physically by the higher authorities or people of the opposite gender at their work place. According to research from the World Bank, it has been observed that over one million women don't have any kind of legal protection against the domestic violence [2].

According to the United Nation Development, there are 17 goals for sustainable development. The fifth goal is gender inequality [3]. Ending all discrimination against women and girls is not only a basic human right, is need of an hour; it's proven that empowering women and girls can help in economic growth and development of the country. The number of schools going girls now is more as compared to the last 15 years, and most regions have reached gender parity in primary education. But still there are more women in the labour market, large inequalities in some regions, with women who are denied the same work rights as men. Still women have to face sexual violence and exploitation, the unequal division of unpaid care and domestic work, and discrimination in public office which can be huge barriers for women to work in safe and secure environment. Climate change and disasters continue to have a disproportionate effect on women and children, as do conflict and migration. It is very important to give women equal rights land and property, sexual and reproductive health, and to technology and the internet. Today there are more women in public office than ever before, but encouraging more women leaders will help achieve greater gender equality.

2. METHODS

The data were obtained from the census of India 1991 and *www.censusindia.net* for 2001 and 2011[4]. The interpolation and extrapolation method are used to fill gaps and predict future values in the time series data[5]. To obtain the data for inter-censual year linear interpolation has been used. We don't have the data for the census year 2021 so we have been used extrapolation method for 2012 to 2019. Female literacy is the percentage of female literates with respect to total literate population similarly for male literacy. Female workforce participation is referred to percentage of female workforce participation with respect to total female work participation and similar for male work participation. Life expectancy of male is the number of years a male person can expect to live since birth and same for female life expectancy. The data of life expectancy is taken from World Bank (macro trends)[6]. Unpaired t-test is used to test the hypothesis.

Table-1. Census Data

Year	Literacy		Workforce participation		Life expectancy	
	Male	Female	Male	Female	Male	Female
1991	64.13	39.29	51.1	22.1	59	59.7
2001	75.26	53.67	51.8	25.6	62.3	64.6
2011	82.24	65.46	53.2	25.6	65.8	69.3

Table-2. Interpolation and extrapolation of the census data

Year	Literacy		Workforce participation		Life expectancy	
	Male	Female	Male	Female	Male	Female
1991	64.13	39.29	51.1	22.1	59	59.7
1992	65.243	40.728	51.17	22.45	59.4	60.4
1993	66.356	42.166	51.24	22.8	59.7	60.9
1994	67.469	43.604	51.31	23.15	60.1	61.4
1995	68.582	45.042	51.38	23.5	60.4	61.8
Year	Literacy		Workforce participation		Life expectancy	
	Male	Female	Male	Female	Male	Female
1996	69.695	46.48	51.45	23.85	60.6	62.2
1997	70.808	47.918	51.52	24.2	60.8	62.3
1998	71.921	49.356	51.59	24.55	61.2	62.7
1999	73.034	50.794	51.66	24.9	61.4	63.3
2000	74.147	52.232	51.73	25.25	61.9	64
2001	75.26	53.67	51.8	25.6	62.3	64.6
2002	75.958	54.849	51.94	25.6	62.8	65.2
2003	76.656	56.028	52.08	25.6	63.1	65.6
2004	77.354	57.207	52.22	25.6	63.5	66.1
2005	78.052	58.386	52.36	25.6	63.7	66.5
2006	78.75	59.565	52.5	25.6	64	66.9
2007	79.448	60.744	52.64	25.6	64.3	67.2
2008	80.146	61.923	52.78	25.6	64.6	67.7
2009	80.844	63.102	52.92	25.6	64.9	68.2
2010	81.542	64.281	53.06	25.6	65.4	68.8
2011	83.441	66.21	52.997	26.613	65.8	69.3

2012	84.259	67.464	53.117	26.714	66.4	69.6
2013	85.065	68.71	53.239	26.805	65.65	69.15
2014	85.86	69.95	53.362	26.887	65.87	69.45
2015	86.648	71.185	53.487	26.963	66.1	69.77
2016	87.432	72.418	53.613	27.036	66.33	70.1
2017	88.217	73.651	53.738	27.109	66.57	70.45
2018	89.007	74.888	53.862	27.187	66.82	70.8
2019	89.809	76.132	53.985	27.275	67.06	71.15

4. DATA ANALYSIS

4.1. Graphical Representation of Data

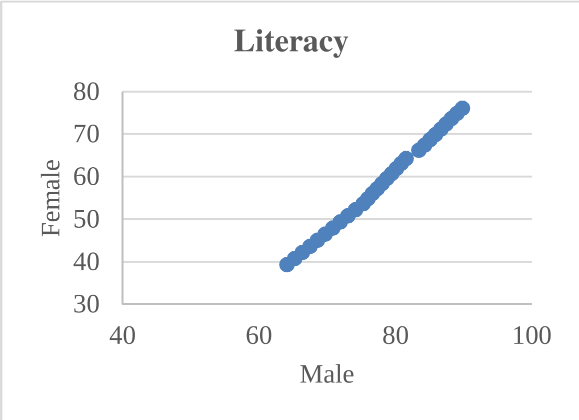


Figure-1. Relation between Male and Female Literacy

The Relation between Male and Female literacy is positive correlated. There is not perfect positive correlation, hence we cannot say there is an equality in terms of literacy.

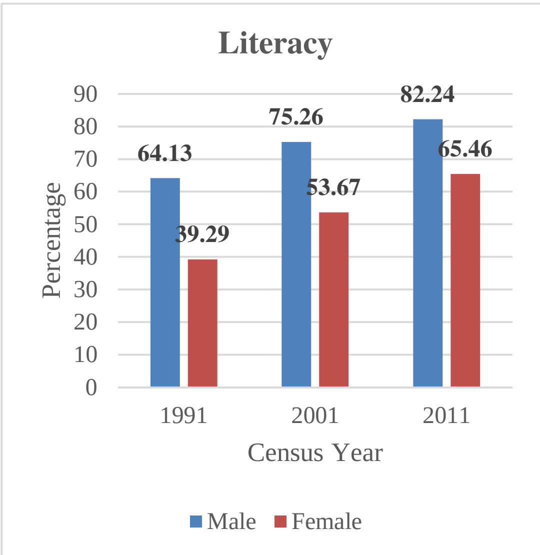


Figure-2. Census year wise literacy in Male and female

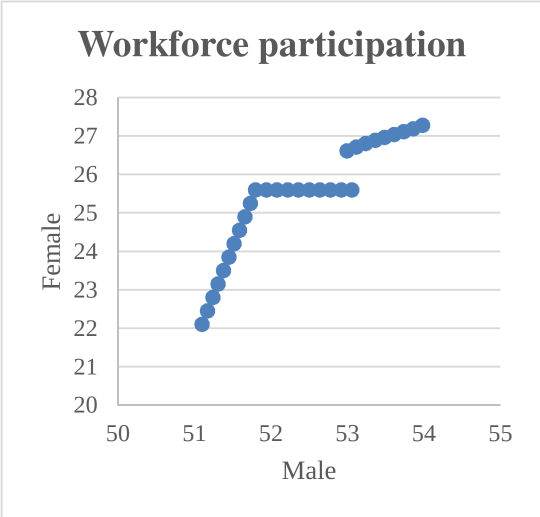


Figure-3. Relation between Male and female Workforce participation

The Relation between Male and Female work participation is weak positive correlated. There is not perfect positive correlation, hence we cannot say there is an equality in terms of work force participation.

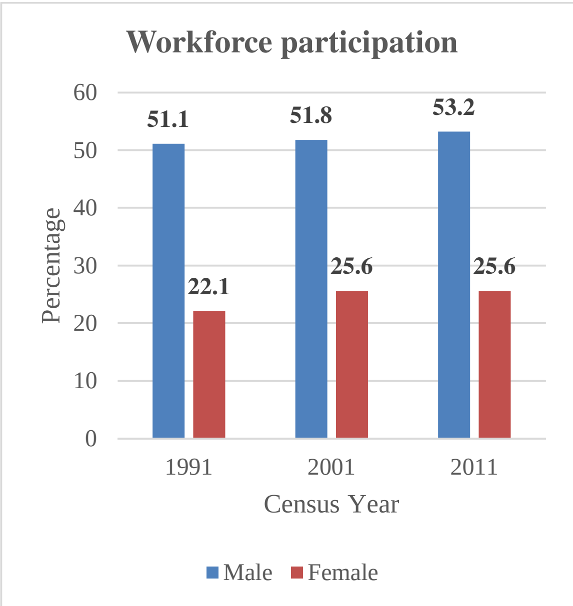


Figure-4. Census year wise Male and female workforce participation

Figure-5. Relation between Life expectancy in Male and Female

The Relation between Male and Female life expectancy is weak positive correlated. There is not perfect positive correlation, hence we cannot say there is an equality in terms of life expectancy.

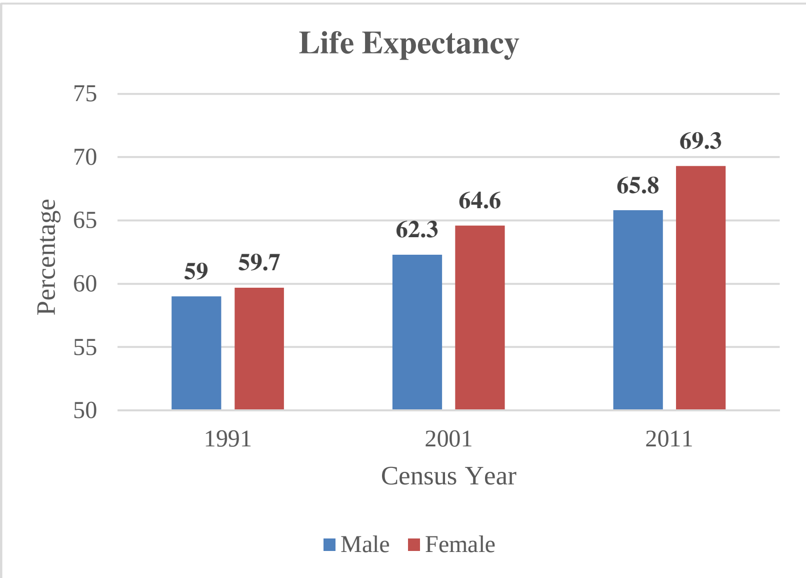


Figure-6. Census year wise Life expectancy in Male and Female

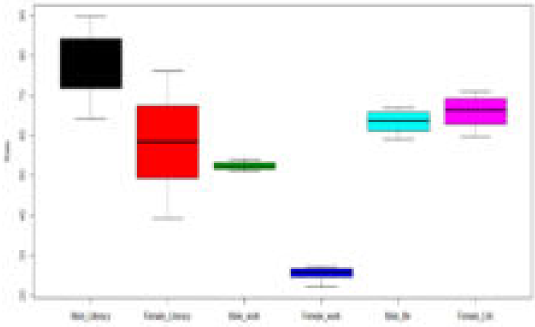


Figure-7. Box plot

This figure shows that there is more variation in the male literacy and female literacy. There is little bit variation in the male workforce participation and female workforce participation. There is variation in the male life expectancy and female life expectancy.

4.2. Descriptive Statistics

Table-3. Descriptive Statistics for the variables Literacy, Workforce Participation and Life expectancy

Descriptive statistics	Literacy		Workforce participation		Life expectancy	
	Male	Female	Male	Female	Male	Female
Mean	77.76	58.21	52.41	25.36	63.44	66.04
Standard deviation	7.65	11.09	0.91	1.49	2.55	3.55
Coefficient Variation	9.83	19.06	1.74	5.88	4.02	5.376
correlation	0.99		0.91		0.99	

4.3. Testing Of Hypothesis

Table-4. Table for Unpaired t-test and Decision

Hypothesis	Average		Calculated t value	Table value	P value	Decision
	Male	Female				
H_{01}	77.76	58.21	7.81	2.048	0.00	Reject H_{01}
H_{02}	52.41	25.36	6.77	2.048	0.00	Reject H_{02}
H_{03}	63.44	66.04	-3.20	2.048	0.00	Reject H_{03}

5. CONCLUSION

The gender equality is most important. It effects on social development of society as well as country. There are different aspects to check the gender equality. So in this paper we study three aspects of gender equality such as literacy, work force participation and life expectancy. The average male and female literacy are 77.76 and 58.21 respectively. The average male and female work force participation are 52.41 and 25.36 respectively. The average male and female life expectancy are 63.44 and 66.04 respectively. The males are more consistent than females in the all aspects of our study.

From the table for unpaired t test, that male literacy and female literacy are significant. For the hypothesis second, we can conclude that male literacy and female work participation are significant. For the hypothesis third, we can conclude that male literacy and female life expectancy are significant.

Male and Female literacy both differ significantly. Box-plot shows as little bit variation in literacy, small change with workforce participation and very little change in life expectancy. Male literacy and female life expectancy are significant. Female literacy is consistent as compare to other factor. Male literacy and female work participation are significant. Male literacy and female literacy are perfect correlated. Male and female literacy, male and female work participation are differing so gender inequity is seen now a day.

REFERENCES

1. https://wbl.worldbank.org/en/past-reports
2. https://www.worldbank.org/en/news/press-release/2018/02/01/more-than-1-billion-women-lack-legal-protection-against-domestic-sexual-violence-finds-world-bank-study
3. https://sdgs.un.org/goals
4. *www.censusindia.net* accessed
5. Brubacher. S.R., Tunnicliffe Wilson G., (1976) Journal of the Royal Statistical Society: Series C, 25, 107–116.
6. https://www.macrotrends.net/

Modified Variable Selection Criterion in Presence of Multicollinearity

Ashok V. Dorugade[a]*
[a]Department of Statistics, Y C Mahavidyalaya Halkarni, Chandgad 416 552 (MS) India.
*Corresponding author: adorugade@rediffmail.com

ABSTRACT
In this article, we consider the problem of variable selection in linear regression when multicollinearity is present in the data. The selection criterion (Rp) is based on the ordinary ridge estimator and it gives satisfactory results than the method based on Least Squares estimator. Performance of Rp* is studied over the Mallow's Cp criterion used for variable selection, for various combinations of ridge estimators and biasing constants.*

KEYWORDS
Linear regression, Subset selection, Ridge estimator, Multicollinearity.

1. INTRODUCTION

Consider the following linear regression model:

$$Y = X\beta + \varepsilon \qquad (1)$$

where, Y is a n × 1 vector of responses, X is a n × k full column rank matrix of n observations on k-1 explanatory variables, β is a k × 1 vector of unknown parameters, ε is a n × 1 vector of disturbance assumed to be distributed with mean vector 0 and variance covariance matrix $\sigma^2 I$, and I is an identity matrix of order n × n. Assume that the covariates $\boldsymbol{x}_i$'s and response variable Y are standardized in such a way that $X'X$ is a non-singular correlation matrix and XY is the correlation between X and Y. We assume that two or more variables in X are nearly linearly dependent. Therefore, the model in (1) suffers from the problem of multicollinearity. In estimating the regression coefficients $\boldsymbol{\beta}$, the ordinary least squares (OLS) estimator, $\hat{\beta} = (X'X)^{-1}X'Y$ the most common method, is unbiased. However, it may still have a large mean squared error (MSE) when the multicollinearity in the design matrix $\boldsymbol{X}$ causes unstable solutions.

One of the most frequently used statistical procedures is variable selection in regression. Variable selection is useful for two reasons: variance reduction and simplicity. A number of variable selection methods have been introduced in recent years. Among the classical variable selection methods such as Mallow's Cp (Mallow's, [1]), most are based on the OLS estimator. Due to poor performance, OLS estimators are sensitive to the presence of multicollinearity. Consequently, variable selection methods based on OLS estimator in turn leads to the inappropriate variable selections. In an effort to over come the problem of OLS with multicollinear data, widely used method of ridge regression, proposed by Hoerl and Kennard [2]. There are various types of ridge estimators including the ridge regression (RR) estimator (Hoerl and Kennard, [2]), Jackknifed ridge regression (JRR) estimator (Hinkley [3]) and Modified Jackknifed ridge regression (MJR) estimator (Batah et al. [4]) used for estimation of regression coefficients. Recently, Dorugade and Kashid [5] proposed the generalized ordinary Jackknife ridge regression (GOJR) estimator ($\hat{\beta}_{GJR}$) having the better performance than other ridge estimators. In ridge regression, selection of biasing constant is important. Using ridge regression, numerous articles have been written for suggesting different ways of estimating the biasing constants including (r_{HKB}, Hoerl and Kennard, [6]), (r_{LW}, Lawless and Wang, [7]), (r_{HMO}, Masuo Nomura, [8]), (r_{KS}, Khalaf and Shukur, [9]). Dorugade and Kashid [10] gives alternative method for determining biasing constant (r_D) and shown the better performance of k_D over other methods. In the presence of multicollinearity standard variable selection algorithms fail to select an adequate subset. Dorugade and Kashid [11] proposed variable selection criterion (Rp) in linear regression based on ridge estimator when multicollinearity is present in the data and shown that Rp gives satisfactory results than Cp-criterion.

In this article we develop variable selection criterion Rp*. It is proposed by computing Rp statistic based on estimator $\hat{\beta}_{GJR}$ which is determined using biasing constant 'k_D'. The Rp* is compared with Cp and Rp-statistic computed by using other biasing constants and ridge estimators. Also performance of Rp* is evaluated for real and simulated datasets exhibits with multicollinearity. The rest of the article is structured as follows:

In Section 2, we describe different ridge estimators and biasing constants. In Section 3, we present Rp-statistic and developed Rp* criterion for variable selection. Performance of Rp* is evaluated as compare to different combinations of ridge estimator and biasing constants through real and simulated data sets in Section 4. Performance of Rp* for different choice of $\hat{\sigma}^2$ and also model selection ability for Rp* evaluated through simulation study in the same section. Article ends with some summary points.

2. PARAMETER ESTIMATION METHODS AND BIASING CONSTANTS

Consider the linear regression model as given in (1). Let $\wedge$ and T be the matrices of eigen values and eigen vectors of $X'X$, respectively, satisfying $T'X'XT = \wedge$ = diagonal $(\lambda_1, \lambda_2, ..., \lambda_{k-1})$, where λi being the i[th] eigen value of $X'X$ and $T'T = TT' = I_{k-1}$. We obtain the equivalent model

$$Y = Z\alpha + \varepsilon, \qquad (2)$$

where Z = XT , it implies that $Z'Z = \wedge$, and $\alpha = T'\beta$ (see Montgomery et al., [12]) Then Ordinary least square (OLS) estimator of α is given by $\hat{\alpha} = (Z'Z)^{-1}Z'Y = \wedge^{-1} Z'Y$. (3)

Therefore, OLS estimator of β is given by

$$\hat{\beta} = T\hat{\alpha}.$$

For the model (2), we get the ORR, OJR, MOJR and GOJR estimators of α are given respectively by Hoerl and Kennard, [6], Hinkley [3] ,Batah et al. [4] and Dorugade and Kashid [5].

The ordinary ridge regression estimator (ORR) of α as

$$\hat{\beta}_R = T\left(I - rA_r^{-1}\right)\hat{\alpha} \qquad (4)$$

Similarly, the ordinary Jackknifed ridge estimator (OJR) of α is

$$\hat{\beta}_{JR} = T\left(I - r^2 A_r^{-2}\right)\hat{\alpha} \qquad (5)$$

Modified ordinary Jackknife ridge estimator (MOJR) of α is

$$\hat{\beta}_{MJR} = T\left(I - r^2 A_r^{-2}\right)\left(\mathrm{I} - \mathrm{r}\mathrm{A}_\mathrm{r}^{-1}\right)\hat{\alpha} \qquad (6)$$

and generalized ordinary Jackknifed ridge regression estimator (GOJR) and it can be written as $\hat{\beta}_{GJR} = T\left(I - r^2 A_r^{-2}\right)\left(\mathrm{I} - \mathrm{r}\mathrm{A}_\mathrm{r}^{-1}\right)^{\mathrm{S}}\hat{\alpha}$ $\mathrm{s} \geq 0$ (7)

where $A_r = \left(\wedge + rI_p\right)$ and 'r' be the biasing constant.

Determination of biasing constant

Many researchers have suggested various methods for determining the ridge parameter. In further study, we have used some of the wellknown methods available for the determination of biasing constant.

(1) $r_{HKB} = \dfrac{(k-1)\hat{\sigma}^2}{\hat{\alpha}'\hat{\alpha}}$ (Hoerl, Kennard, [6]) (8)

(2) $r_{LW} = \dfrac{(k-1)\hat{\sigma}^2}{\sum_{i=1}^{k-1}\lambda_i\hat{\alpha}_i^2}$ (Lawless and Wang, [7]) (9)

(3) $r_{HMO} = (k-1)\hat{\sigma}^2 \Big/ \sum_{i=1}^{(k-1)}\left[\hat{\alpha}_i^2 \Big/ \left\{1 + \left(1 + \lambda_i\left(\hat{\alpha}_i^2/\hat{\sigma}^2\right)^{1/2}\right)\right\}\right] \quad i = 1,2,\ldots,k-1.$

(Masuo Nomura, [8]) (10)

(4) $r_{KS} = (\lambda_{\max}\hat{\sigma}^2) \big/ ((n-k-2)\hat{\sigma}^2 + \lambda_{\max}\hat{\alpha}^2{}_{\max}$

(Khalaf and Shukur, [9]) (11)

(5) $r_D = \max\left(0, \frac{(k-1)\hat{\sigma}^2}{\hat{\alpha}'\hat{\alpha}} - \frac{1}{n(VIF_i)_{\max}}\right)$

(Dorugade and Kashid, [10]) (12)

where $VIF_i = \frac{1}{1-R_i^2}$ $i = 1,2,...,k-1$ is variance inflation factor of i[th] regressor. and $\hat{\sigma}^2$ is the OLS estimator of σ^2 i.e. $\hat{\sigma}^2 = \frac{Y'Y - \hat{\alpha}'Z'Y}{n-k}$.

3. PROPOSED CRITERION AND STEPWISE PROCEDURE FOR SUBSET SELECTION

Using the fitted regression equation based on the full model, we have the predicted values of Y which depends on the full set of size k-1.

$$\hat{Y}_{jk} = X_j'\hat{\beta}_{GJR} \qquad j = 1, 2, 3,\ldots,n. \qquad (13)$$

where $X_j' = (1, X_{j1}, X_{j2},...,X_{jk-1})$.

Now assume that a sub model 'A' based on a subset of p-1 predictor variables (p < k) is fitted to the data. The underlying model is given by

$$Y = X_A\beta_A + \varepsilon .$$

where X_A is an n × p matrix of the observations on p-1 predictors and β_A is a p × 1 vector of the regression coefficients based on the fitted submodel. We have the predicted values of Y as

$$\hat{Y}_{jp} = X_j'\hat{\beta}_{GJR} \qquad j = 1, 2, 3,\ldots,n. \qquad (14)$$

where $X_j' = (1, X_{j1}, X_{j2},...,X_{jp-1})$.

Where, $\hat{\beta}_{GJR}$ in (13) and (14) is computed using biasing constant r_D for full and subset model respectively.

We propose the new subset selection criterion Rp* on the similar line of Rp-criterion. It is defined as follows:

3.1 Definition

The Rp* statistic is defined as

$$\text{Rp*} = \frac{\sum_{j=1}^{n}\left(\hat{Y}_{ik} - \hat{Y}_{ip}\right)^2}{\sigma^2} - tr(H_R' H_R) + tr(H_{RA}' H_{RA}) + p \qquad (15)$$

where, $H_R = X(X'X + r_D I)^{-1}X'$ and $H_{RA} = X_A(X_A'X_A + r_D I)^{-1}X_A'$. p is the number of parameters in of the subset model . σ^2 is replaced by its suitable estimate (see Section 4).

Stepwise Procedure for Subset Selection

Here, we present steps actually involved in subset selection procedure.

Step- 1 Standardize regressor variables (X) and response variable (Y) in such way that $X^{'}X$ and $X^{'}Y$ are in the correlation forms.

Step- 2 Convert the model Y = Xβ + ε into the canonical form as Y = Z$\alpha+\varepsilon$.

Step- 3 Determine ridge parameter r_D using the model Y = Zα $+\varepsilon$.

Step- 4 Find the Generalized Ordinary Jackknife Ridge Regression Estimator (GJR) $\hat{\alpha}_{GJR}$ of α using ridge parameter obtained in Step- 3.

Step- 5 Convert the ridge estimator into the standardized form and finally, translate into the original form. It is denoted as $\hat{\beta}_{GJR}$.

Step- 6 Repeat Step 2 to Step 5 and Compute the predicted value $\hat{Y}_{ik}$ and $\hat{Y}_{ip}$ for full and all possible subset models respectively.

Step- 7 Compute the proposed statistic Rp* for all possible subsets.

Step- 8 Select a subset of minimum size, for which Rp*close to p.

4. COMPARATIVE STUDY

In this section, we compare and evaluate the performance of Rp*-statistic through simulation study. The simulation study is divided into three different parts:

A. Comparison between Rp and Rp*.

B. Performance of Rp* statistic for various estimators of σ^2.

C. Correct model selection ability of Rp*, Rp and Cp.

Part A:

We compare the performance of the proposed procedure Rp* with Rp-statistic by considering two numerical examples. We have used Hald Cement data and simulated data. The ridge regression estimators $\hat{\beta}_R$, $\hat{\beta}_{JR}$, $\hat{\beta}_{MJR}$ and ridge parameters r_{HKB}, r_{LW}, r_{HMO}, r_{KS} and r_D are used for computing the value of Rp for all possible subsets**.**

Example 4.1 Hald Cement Data: In this example, we use Hald cement data (Montgomery et al., [12]. The values of Rp and Rp* are computed for all possible subsets and reported in Table 4.1(a) and 4.1(b).

From the results mentioned in Table- 4.1 (a and b), it is clear that, Rp-statistic for various ridge parameters and ridge estimators and Rp* statistic agree for the same subset $\{X_1, X_2\}$. It indicates that, performance of both the methods is same for subset selection in the presence of multicollinearity.

Example 4.2 We have generated random sample from N_3 (0, Σ) on X_1, X_2 and X_3, and random error variable (ε) is generated from normal with mean 0 and variance 15.

where

$$\Sigma = \begin{bmatrix} 1 & 0.67 & 0.99 \\ 0.67 & 1 & 0.698 \\ 0.99 & 0.698 & 1 \end{bmatrix}$$

Response variable Y is generated using the following model.

$$Y = 5+3X_1 + 2X_2 + \varepsilon .$$

The values of Rp and Rp* are computed for all subset models and reported in Table-4.2(a) and 4.2(b).

From the Tables 4.2 (a and b), Rp* and Rp (obtained using various ridge estimators and ridge parameters) pick up the same subset $\{X_1, X_2\}$. Rp* is close to p when subset model is adequate as compared to other method.

Part B:

Performance of Rp* -statistic using various estimators of σ^2

We have used four different types of estimator of σ^2, which are based on the LS estimator ($\hat{\beta}$) and GJR estimator ($\hat{\beta}_{GJR}$) of β. These are given below.

1. $\hat{\sigma}_1{}^2 = (Y - X\hat{\beta})^{'}(Y - X\hat{\beta}) \Big/ (n-k)$.
2. $\hat{\sigma}_2{}^2 = (Y - X\hat{\beta}_{GJR})^{'}(Y - X\hat{\beta}_{GJR}) \Big/ (n-k)$.
3. $\hat{\sigma}_3{}^2 = (Y - X\hat{\beta}_{GJR})^{'}(Y - X\hat{\beta}_{GJR}) \Big/ (n-k-2)$.
4. $\hat{\sigma}_4{}^2 = (Y - X\hat{\beta}_{GJR})^{'}(Y - X\hat{\beta}_{GJR}) \Big/ (n - 2tr(H_R) + tr(H_R^{'} H_R))$.

where, k is the number of parameters.

We will use these estimators of σ^2 in Rp*-statistic for evaluating influence of these estimators on Rp*. For this study, we have considered following example.

Example 4.3 Here we have used Hald Cement data applied in Example 4.1. We have calculated the values of $\hat{\sigma}_i^2$ (i =1, 2, 3, 4). The Rp*-statistic is obtained using all values of $\hat{\sigma}_i^2$ (i =1, 2, 3, 4) for all possible subset models. The values of $\hat{\sigma}^2$ are given below:

$\hat{\sigma}_1^2$ = 5.98295, $\hat{\sigma}_2^2$ = 6.20300, $\hat{\sigma}_3^2$ = 8.27067 and $\hat{\sigma}_4^2$ = 4.98359.

The values of Rp*-statistic are presented in the following table.

From Table- 4.3, we observe that Rp* statistic selects the same subset {X_1, X_2} for all $\hat{\sigma}_i^2$. Therefore, we suggest that, one of the $\hat{\sigma}_i^2$'s (i =1, 2, 3, 4) can be used for computing Rp*.

Part C:

In this part, we study the performance of Rp*. Performance is evaluated in terms of number of times it selects a model correctly and incorrectly. The simulation study is carried out for different models. Here we have used r_{HKB} and r_D in the determination of $\hat{\beta}_R$. In this study, the performance of the proposed statistic Rp*, Rp (computed using $\hat{\beta}_R$) and Cp is evaluated for different subset models for different sample size (n) and variance of the error variable (σ^2). The values n and σ^2 are taken randomly. We have used various ridge estimators and ridge parameters.

The details about the submodel specification, sample size and variance of the error variable (ε) are given below.

where, $\Sigma = \begin{bmatrix} 1 & 0.67 & 0.99 \\ 0.67 & 1 & 0.698 \\ 0.99 & 0.698 & 1 \end{bmatrix}$ and

$$\Sigma_1 = \begin{bmatrix} 1 & 0.2290 & -0.8240 & -0.2450 \\ 0.2290 & 1 & -0.139 & -0.973 \\ -0.8240 & -0.139 & 1 & -0.030 \\ -0.2450 & -0.973 & -0.030 & 1 \end{bmatrix}$$

We have generated 1000 samples of size n from each model. Based on each sample, the values of Rp*, Rp, and Cp were computed for all possible subsets. Thereafter, the number of times a criterion selects a correct model and incorrect model is counted. The results are expressed in percentage and values are reported in Table -4.5.

From above simulation study, it can be seen that Rp* selects a correct model 81% for model I, 78% for model II, 80% for model III and 78% for model IV. The Rp-statistic with ridge parameter 'r_D' selects 80%, 75%, 70% and 72% for model I, II, III and IV respectively. Therefore, above study indicated that the performance of Rp* is better than Rp and Cp.

5. SUMMARY

Suggested criterion in this article for variable selection gives satisfactory results than the method based on LS estimator of β. In this article, we have shown that how the suggested criterion can be used to select subset of variables when several regressors are highly correlated to each other. The proposed method selects an appropriate subset of variables in the same situation.

REFERENCES

1. Mallows.C.L. (1973) Technometrics, 15, 661-675.
2. Hoerl, A.E., Kennard, R.W. (1970). Technometrics, 12, 55-67.
3. Hinkley (1977). Technometrics, 19, 285-292.
4. Bath, R., Gore. (2008) Surveys in Mathematics and its Applications, 24,157-174.
5. Dorugade, A.V., Kashid, D.N. (2011) InterStat.
6. Hoerl, A.E., Kennard, R.W., Baldwin, K.F. (1975) Communications in Statistics, 4, 105-123.
7. Lawless J.F., Wang P. A (1976) Communications in Statistics-Theory and Methods, 14, 1589- 1604.
8. Masuo N. (1988) Communications in Statistics - Simulation and Computation, 17, 729-743.
9. Khalaf G., Shukur G. (2005) Communications in Statistics-Theory and Methods, 34, 1177-1182.
10. Dorugade, A.V. and Kashid, D.N. (2010) Applied Mathematical Sciences, 4(9), 447 – 456.
11. Dorugade, A.V., Kashid, D.N. (2010) Journal of Statistical Computation and Simulation, 80(11), 1211–1224.
12. Montgomery D.C., Peck E.A., Vining, G.G. Introduction to Linear Regression Analysis, edited by Peter Bloomfield, Noel A. C Cressie, John Wiley and Sons, New York (2001) pp. 363, 366.

Table-4.1(a). The values of Cp, Rp for different combinations of ridge parameter and ridge estimators and Rp*.

Model	Cp	Rp									
		$\hat{\beta}_R$					$\hat{\beta}_{JR}$				
		r_{HKB}	r_{LW}	r_{HMO}	r_{KS}	r_D	r_{HKB}	r_{LW}	r_{HMO}	r_{KS}	r_D
(1)	202.549	199.383	200.77	178.335	201.316	199.47	201.754	202.011	201.149	202.42	201.768
(2)	142.486	138.468	139.994	119.687	141.21	138.565	142.21	142.439	138.985	142.716	142.224
(3)	315.154	309.287	312.218	268.886	311.391	309.476	313.149	313.687	311.02	314.382	313.179
(4)	138.731	135.481	136.643	119.775	137.542	135.554	138.094	138.316	136.382	138.641	138.106
(12)	**2.678**	**3.57**	**3.617**	**3.736**	**3.637**	**3.574**	**3.755**	**3.711**	**3.486**	**3.581**	**3.753**
(13)	198.095	200.545	206.885	180.354	198.996	200.618	198.112	199.909	203.746	198.063	198.119
(14)	5.496	6.874	6.663	9.55	8.117	6.862	6.216	6.176	7.161	6.081	6.213
(23)	62.438	59.723	60.79	54.708	61.707	59.794	62.748	62.883	59.638	62.978	62.757
(24)	138.226	135.191	135.278	131.284	136.277	135.277	136.981	137.243	137.071	137.608	136.994
(34)	22.373	21.066	21.597	18.529	23.148	21.1	22.638	22.715	20.778	22.757	22.644
(123)	3.041	4.292	4.228	4.518	4.127	4.292	4.178	4.116	4.271	3.968	4.175
(124)	3.018	3.712	3.777	4.728	3.73	3.715	3.835	3.87	3.459	3.776	3.838
(134)	3.497	4.202	4.179	5.259	8.607	4.202	4.153	4.137	3.921	3.636	4.153
(234)	7.337	7.15	9.8	8.449	14.091	7.124	7.264	7.534	7.426	8.782	7.268
(1234)	5	5	5	5	5	5	5	5	5	5	5

(i,j,k,…) indicates the variable $X_i, X_j, X_k, \ldots$ in the model

Table-4.1(b). The values of Rp for different combinations of ridge parameter and ridge estimators and Rp*.

Model	Rp									Rp*
	$\hat{\beta}_{MJR}$					$\hat{\beta}_{GJR}$				
	r_{HKB}	r_{LW}	r_{HMO}	r_{KS}	r_D	r_{HKB}	r_{LW}	r_{HMO}	r_{KS}	
(1)	199.029	200.389	177.542	200.761	199.116	196.641	200.364	147.853	200.104	154.403
(2)	138.053	139.625	117.722	140.836	138.154	133.944	137.562	98.826	139.587	105.239
(3)	309.672	312.631	268.769	310.456	309.863	309.354	316.446	213.703	308.179	242.927
(4)	135.098	136.256	118.327	137.11	135.172	132.004	134.895	99.83	136.287	103.682
(12)	**3.59**	**3.66**	**3.841**	**3.718**	**3.595**	**3.283**	**3.426**	**5.428**	**3.588**	**3.044**
(13)	203.271	212.566	181.762	199.307	203.336	208.703	224.445	142.875	203.463	164.152
(14)	6.909	6.675	10.29	8.378	6.894	7.418	7.045	12.4	14.157	6.26
(23)	59.462	60.6	55.712	61.563	59.538	56.718	58.922	57.493	60.328	45.028
(24)	135.065	134.921	137.391	135.87	135.153	135.235	133.507	130.275	135.017	106.44
(34)	20.909	21.476	18.315	23.184	20.945	19.249	20.355	20.312	25.915	15.602
(123)	4.337	4.286	4.616	4.218	4.338	4.341	4.296	5.1	4.261	4.283
(124)	3.678	3.774	5.189	3.761	3.682	3.84	3.857	6.743	3.801	3.784
(134)	4.196	4.173	5.579	9.811	4.196	4.256	4.198	8.13	24.36	4.204
(234)	7.081	11.149	9.453	18.122	7.044	7.259	16.362	10.19	26.669	6.442
(1234)	5	5	5	5	5	5	5	5	5	5

(i,j,k,…) indicates the variable X_i,X_j, X_k,… in the model

Table-4.2(a). The values of Rp for different combinations of ridge parameter and ridge estimators and Rp*.

Model	Rp							
	$\hat{\beta}_R$				$\hat{\beta}_{JR}$			
	r_{HKB}	r_{LW}	r_{KS}	r_D	r_{HKB}	r_{LW}	r_{KS}	r_D
(1)	10.5551	10.6329	10.382	10.5414	10.3357	10.409	10.2025	10.3225
(2)	48.1117	47.5435	49.2021	48.2145	49.5303	49.1588	50.0668	49.5924
(3)	9.4408	9.1305	10.2528	9.5044	10.4371	10.0353	11.2044	10.5119
(12)	**2.7449**	**2.9033**	**2.5769**	**2.7282**	**2.5745**	**2.6518**	**2.4861**	**2.5633**
(13)	10.2279	10.278	10.6022	10.2511	10.7024	10.8115	11.1458	10.7325
(23)	4.5073	4.3358	5.1067	4.5516	5.294	5.0014	5.9003	5.3513
(123)	4	4	4	4	4	4	4	4

(i,j,k,…) indicates the variable $X_i, X_j, X_k, \ldots$ in the model

Table-4.2(b). The values of Rp for different combinations of ridge parameter and ridge estimators and Rp*.

Model	Rp							Rp*
	$\hat{\beta}_{MJR}$				$\hat{\beta}_{GJR}$			
	r_{HKB}	r_{LW}	r_{KS}	r_D	r_{HKB}	r_{LW}	r_{KS}	
(1)	10.6565	10.7448	10.4387	10.6401	10.6757	10.7632	10.3883	10.2874
(2)	47.383	46.689	48.8845	47.5166	45.9144	45.2366	47.5944	44.3272
(3)	8.9054	8.5671	9.928	8.9795	8.1577	7.9804	8.9328	7.9222
(12)	**2.8985**	**3.0953**	**2.6374**	**2.8745**	**3.0713**	**3.2568**	**2.6546**	**3.0327**
(13)	9.9351	9.9482	10.3736	9.9621	9.589	9.6235	9.8658	9.3089
(23)	4.1562	3.9818	4.8692	4.2042	3.7745	3.6633	4.1577	3.7526
(123)	4	4	4	4	4	4	4	4

(i,j,k,…) indicates the variable $X_i, X_j, X_k, \ldots$ in the model

Table-4.3. Values of Rp* for $\hat{\sigma}_i^2$

Model	$\hat{\sigma}_1^2$	$\hat{\sigma}_2^2$	$\hat{\sigma}_3^2$	$\hat{\sigma}_4^2$
(1)	204.121	154.403	115.804	215.405
(2)	139.105	105.239	78.946	146.791
(3)	321.226	242.927	182.14	338.998
(4)	137.046	103.682	77.78	144.619
(12)	**3.336**	**3.044**	**2.817**	**3.402**
(13)	216.581	164.152	123.449	228.481
(14)	7.59	6.26	5.227	7.892
(23)	58.884	45.028	34.271	62.029
(24)	140.397	106.44	80.078	148.105
(34)	19.949	15.602	12.227	20.936
(123)	4.353	4.283	4.229	4.368
(124)	3.848	3.784	3.734	3.862
(134)	4.263	4.204	4.158	4.276
(234)	7.233	6.442	5.828	7.412
(1234)	5	5	5	5

(i,j,k,…) indicates the variable $X_i, X_j, X_k, \ldots$ in the model

Table-4.4. Submodel Specifications with sample size and error variable

Model	Sample Size (n)	Submodel specification	Error variable (ε) generated from	Predictors generated from
I	25	$Y = 5+3X_1 + 2X_2 + \varepsilon$	$N(0,15)$	Σ
II	50	$Y = 2X_1 + X_3 + \varepsilon$	$N(0,1)$	Σ
III	25	$Y = 20+X_3 +6 X_4 + \varepsilon$	$N(0,12)$	Σ_1
IV	75	$Y = 3X_2 +8X_4 +3 X_5 + \varepsilon$	$N(0,5)$	Σ_1

Table-4.5. Model selection ability (in %) of Rp*, Rp and Cp

Model	Model status	Cp	Rp		Rp*
			r_{HKB}	r_D	
I	Correct	38	57	80	**81**
	Incorrect	62	43	20	19
II	Correct	40	63	75	**78**
	Incorrect	60	37	25	22
III	Correct	30	65	70	**80**
	Incorrect	70	35	30	20
IV	Correct	38	68	72	**78**
	Incorrect	62	32	28	22

A Review on Nuisance Parameter Free Inferential Procedures for Shape-Scale and Location-Scale Family of Distributions

Kiran P. Patil*[a], Hemangi V. Kulkarni[b]

[a]Department of Statistics, Anandibai Raorane
Arts, Commerce and Science College Vaibhavwadi 416 810, Sindhudurg, (MS) India.
[b]Department of Statistics, Shivaji University, Kolhapur 416 004, (MS) India.
*Corresponding author: kiran.patil9020@gmail.com

ABSTRACT

A variety of parametric and non-parametric inferential procedures are available to study inference on the parameter of interest in the presence of nuisance parameters, but majority of these are constrained by certain limitations, as for example depicted through a variety of examples by Berger (1999). Also, small deviations from the underlying assumptions might often cause biased statistical inference, especially in small to moderate size samples. Additionally, existence of the nuisance parameters also disturbs the statistical properties of the estimation procedures of the parameter of interest. This motivates us to take brief review on improved or efficient and unified superior nuisance parameter-free (invariant) inferential procedures under shape-scale and location-scale family of distributions.

KEYWORDS

Generalized variable approach, Maximal scale invariant Estimator, Integrated likelihood, Profile likelihood.

…………………………………………………………………………………..............

1. INTRODUCTION

Lifetime data are often well modelled by distributions belonging to shape-scale and location-scale families of distributions and are widely used in almost every discipline, see for example Kulkarni and Powar (2010, 2011), Patil and Kulkarni (2011), Jones (2015), Powar and Kulkarni (2015), Sengupta et. al. (2015), Rigby et. al. (2005, 2019) and Maswadah (2013, 2022).[1-3] The characteristics of a dataset can be measured through the measures of central tendency, dispersion, skewness, and kurtosis, which are usually well-defined functions of the shape, scale, and location parameters. In this context, we review some efficient or improved inferential procedures for shape-scale and location-scale families.[4] The widely applicable shape-scale families for monitoring lifetime data include the important skewed distributions like Gamma distribution, Weibull distribution, Generalized exponential distribution, Pareto distribution, Log-

logistic, Log-normal distribution, Hyperbolic distribution, Exponentiated exponential, among others. The shape scale family of distributions is characterized by the probability density function (PDF) of the form:

$$g_1\big(x\big|(a,b)\big) = \frac{1}{a} f_1\left(\frac{x}{a}, b\right), \quad a, b, x > 0.$$

where a and b are the scale and shape parameters respectively, $f_1(.,b)$ being a function of only one parameter, namely the shape parameter b.

Distributions belonging to the location-scale family are used in hydrology, bio-statistics, various industrial and analytical fields, among others[5]. Normal, Logistic, Laplace, shifted exponential, Extreme value distribution are some popular members of the location-scale family, among others[6].

The PDF of a random variable Y from a location-scale family of distributions is characterized by density function of the form:

$$g_2\left(\frac{y-\mu}{\sigma}\right) = \frac{1}{\sigma} f_2\left(\frac{y-\mu}{\sigma}\right), \quad y, \mu \in \mathcal{R}, \sigma > 0.$$

where μ and σ are the location and scale parameters respectively, and $f_2(z)$ is the probability density function of the standard random variable Z having location parameter zero and scale parameter one[7-8].

This article aims to review improved inferential procedures, including point estimation, interval estimation, and hypothesis testing, related to distributions belonging to the location-scale and shape-scale families. Improved inference in the case of point estimation is often related to the reduction of bias and variability of the concerned estimator, while for the case of interval estimation and testing of the hypotheses it concerns the attainment of nominal level, increased coverage probability, and elevated powers, respectively[9-11].

Though often nuisance parameters are absolutely essential for better modeling of the data, most often, existence of one or more nuisance parameters adversely impacts the performance of inference procedures for the parameters of interest. Existence of nuisance parameters may produce their adverse impact in a variety of ways, e.g., increased standard errors of point estimators, volumes/ lengths/ area of confidence region/intervals or rate of convergence of asymptotic properties of the parameters of interest among others[10]. A way-out is an attempt for reducing their impact using some well-known likelihood-based techniques, including conditional likelihood, integrated, profile or pseudo-likelihood function, and their modifications, or through the use of pivot or generalized pivot quantities with completely known probability distributions or circumventing the existence of nuisance parameters through the tricky use of invariance principle[11].

Marginal and conditional likelihoods handle the problem by ignoring some of the data (marginalization) or by ignoring their variability (conditioning). When the number of nuisance parameters are large, then marginalization and conditioning are pretty complex, and sacrifice a sizeable information[12].

In this article, emphasis relies on the procedures eliminating of the impact of nuisance parameters through the invariance principle and generalized variable approach, which are expected to result in more efficient inference procedures by use of the entire data without losing any details [13].

The invariance principle is used to circumvent the effect of the nuisance parameters, making use of their property of being invariant under a group of transformations. The maximal scale invariant inference under a shape-scale family developed by Kulkarni and Patil (2018) turned out to be much efficient than classical procedures for the commonly encountered distributions enjoying the scale invariance property [14]. The generalized variable approach is another efficient tool for exact nuisance-parameters-free parametric inference in certain parametric families. The generalized variable approach is based on the generalized extreme region of a test, the generalization of a data-based extreme region of a test, which depends on the observed data and may involve all the parameters, where the associated p-value is independent of the nuisance parameters [15-16].

In this article, the improved inferences for the inferential problems including point estimation, one sample test and interval estimation for the parameter of interest under the shape-scale family of distributions, stress- strength reliability estimation for the exponentiated-scale family of distributions, test for two-sample comparison for two independent mixed continuous location- scale or some non-location-scale populations and test for homogeneity of variances among several location-scale populations are reviewed[17-19].

In more general set-up, some basic definitions in the generalized pivotal approach are given in the following subsection.

2. PRELIMINARIES

2.1. The Generalized Variable Approach

Tsui and Weerahandi (1989) introduced the concept of generalized p-values which is based on the generalized pivot quantity (GPQ) and generalized test variable (GTV)[20]. Let $\boldsymbol{X}$ be a random variable with cumulative distribution function (CDF) $F_{\xi}(.)$, where $\boldsymbol{\xi} = (\boldsymbol{\theta}, \boldsymbol{\delta})$ is an unknown parameter vector and $F_{\xi}(.)$ is a member of the shape-scale or location-scale family of distributions. Suppose the interest lies in the parameter $\boldsymbol{\theta}$ while $\boldsymbol{\delta}$ is the nuisance parameter. A GPQ for $\boldsymbol{\theta}$, GTV and generalized p-value (GPV) for testing a one-sided hypothesis $H_0: \boldsymbol{\theta} \leq \boldsymbol{\theta}_0$ verses $H_1: \boldsymbol{\theta} > \boldsymbol{\theta}_0$ is defined below:

Definition 1: Generalized pivot quantity (GPQ)

The GPQ $G_{\boldsymbol{\theta}} = \psi(\boldsymbol{X};\ \boldsymbol{x}, \boldsymbol{\xi})$ for $\boldsymbol{\theta}$ is a random quantity that satisfies following two conditions:

i. The distribution of $G_{\boldsymbol{\theta}}$ for given $\boldsymbol{X} = \boldsymbol{x}$ is free from any unknown parameters.

ii. The value of $G_{\boldsymbol{\theta}} = \psi(\boldsymbol{X};\ \boldsymbol{x}, \boldsymbol{\xi})$ at $\boldsymbol{X} = \boldsymbol{x}$ does not depend on any unknown parameter, other than $\boldsymbol{\theta}$. For most of the cases, $G_{\boldsymbol{\theta}} = \boldsymbol{\theta}$ at $\boldsymbol{X} = \boldsymbol{x}$.

The following invariance property of GPQs is an easy consequence of its definition:

Preposition 1: Invariance property of GPQ

If $G_{\boldsymbol{\theta}}$ is a GPQ for $\boldsymbol{\theta}$, then for any function π, $\pi(G_{\boldsymbol{\theta}})$ is GPQ for $\pi(\boldsymbol{\theta})$.

Definition 2: Generalized test variable (GTV)

A random quantity $\tau_{\boldsymbol{\theta}} = T\ (\boldsymbol{X};\ \boldsymbol{x}, \boldsymbol{\xi})$ is said to be GTV for the parameter of interest $\boldsymbol{\theta}$ if it satisfies following three properties:

i. The probability distribution of $\tau_{\boldsymbol{\theta}}$ is free from any unknown parameters.

ii. The value of $\tau_{\boldsymbol{\theta}} = T\ (\boldsymbol{X};\ \boldsymbol{x}, \boldsymbol{\xi})$ at $\mathbf{X} = \mathbf{x}$ does not depend on any unknown parameter, other than $\boldsymbol{\theta}$.

iii. For fixed $\mathbf{x}$, the probability $P\ (T\ (\boldsymbol{X};\ \boldsymbol{x}, \boldsymbol{\xi}) \geq t|\boldsymbol{\theta})$, for all $t \geq 0$ is non-decreasing in $\boldsymbol{\theta}$.

Preposition 2 : Connection between GPQ and GTV

If $G_{\boldsymbol{\theta}}$ is a GPQ for $\boldsymbol{\theta}$, then $\tau_{\boldsymbol{\theta}} = G_{\boldsymbol{\theta}} - \boldsymbol{\theta}$ is a GTV for $\boldsymbol{\theta}$ (Weerahandi (1995)).

Definition 3 : Generalized p-value (GPV)

Based on the GTV defined in Definition 2 and Preposition 2, the generalized

p-value for testing H_0 mentioned above is defined by

$p = Sup_{\boldsymbol{\theta}\ \epsilon\ H_0} P(T\ (\boldsymbol{X};\ \boldsymbol{x}, \boldsymbol{\theta}, \delta) \geq t)$, were, $t = T\ (\boldsymbol{x};\ \boldsymbol{x}, \boldsymbol{\theta}, \delta)$

$p = P(T\ (\boldsymbol{X};\ \boldsymbol{x}, \boldsymbol{\theta_0}, \delta) \geq t)$,on account of property iii of *Definition 2*.

2.2. The Invariance Principle

If $\mathbf{X}$ is a random variable having density function $f(\boldsymbol{x}, \boldsymbol{\theta})$, $\boldsymbol{\theta}\epsilon\boldsymbol{\Theta}$ and G be a group of transformation on the space of values of $\mathbf{X}$ then:

i. ϕ is invariant under G if $\phi\big(g(\boldsymbol{x})\big) = \phi(\boldsymbol{x})$ for all $\boldsymbol{x}$ and all $g\epsilon G$.

ii. $T(\boldsymbol{x})$ is maximal invariant under G if $T(\boldsymbol{x_1}) = T(\boldsymbol{x_2}) \Longrightarrow \boldsymbol{x_1} = g(\boldsymbol{x_2})$ for some $g \epsilon G$.

Where $\boldsymbol{x}$ is observed value of **X.**

2.2.1 Location Invariant

Let $x = (x_1, x_2, \dots, x_n)$, be the random sample from location family with location parameter μ and G be the group transformation then

$g(x) = (x_1 + \mu, x_2 + \mu, \dots., x_n + \mu)$, $\quad -\infty < \mu < \infty$, then

$$T(x) = T\big(g(x)\big) = (x_n - x_1, \dots, x_n - x_{n-1}).$$

is called as maximal location invariant estimator.

2.2.2 Scale Invariant

Let $x = (x_1, x_2, \dots, x_n)$, be the random sample from scale family with scale parameter σ and G be the group transformation then

$g(x) = (\sigma x_1, \sigma x_2, \dots., \sigma x_n)$, $\quad -\infty < \mu < \infty$, then

$$T(x) = T\big(g(x)\big) = (\frac{x_n}{x_1}, \frac{x_1}{x_2}, \dots, \frac{x_{n-1}}{x_n}).$$

$T(x)$ is maximal scale invariant estimator.

2.2.3 Location-Scale Invariant

Let $x = (x_1, x_2, \dots, x_n)$, be the random sample from location-scale family with location parameter μ and scale parameter σ. Let G be the group transformation then

$g(x) = (\sigma(x_1 + \mu), \sigma(x_2 + \mu), \dots., \sigma(x_n + \mu))$, $\quad -\infty < \mu < \infty$, then

$$T(x) = T\big(g(x)\big) = (\frac{x_n - x_{n-1}}{x_2 - x_1}, \frac{x_{n-1} - x_{n-2}}{x_3 - x_2}, \dots, \frac{x_2 - x_1}{x_n - x_{n-1}}, \frac{x_1 - x_n}{x_n - x_1}).$$

$T(x)$ is maximal location-scale invariant estimator.

The next section reviews the literature related to the treatment for nuisance parameters.

3. LITERATURE REVIEW

There have been numerous articles addressing a systematic study of a variety of methods for eliminating nuisance parameters.

3.1. Likelihood Based Approach

A pseudo-likelihood or profile likelihood is obtained by replacing the nuisance parameters with their maximum likelihood estimators obtained by keeping the parameters of interest fixed. After fixing the interest parameters, the MLEs of nuisance parameters are expressed as functions of interest parameters and after replacing the nuisance parameters by these functions, the likelihood gets translated to a function of only interest parameters. This likelihood behaves similar to the classical likelihood. For the critical review and various aspects of pseudo or profile likelihood, we refer to Kalbfleish and Sprott (1989)[21], Gong and Samaniego (1981)[22], Fraser and Reid (1989)[23], Barndorff-Nielsen (1985)[24], Barndorff-Nielsen (1991)[25], Barndorff-Nielsen (1994)[26] and Severini (1998)[27].

Integrated likelihood approach is another way to eliminate nuisance parameters, For notable analytical results in this context we refer to Berger and Wolpert (1988), Berger et al. (1999), Severini (2000), and Severini (2010), among others. Notable novel recent inferential procedures based on integrated likelihood have been developed by SenGupta and Kulkarni (2018), Kulkarni and SenGupta (2021), Patil and Kulkarni (2022), and Kulkarni and Patil (2021) under directional and linear data[23-27].

3.2. Invariance Principle Approach:

Nuisance parameters free inference can also be based on an ancillary statistic, invariant or weighted average power criterion, and conditional probability as reported in Linnik and Technica (1968), Cox and Hinkley (1974), Engelhardt and Bain (1977), Andrews and Ploberger (1994), and Hansen (1996)[28].

Invariance principle can be coupled with appropriate data transformation to yield nuisance parameters free transformed likelihood that is purely function of the parameters of interest and the observed sample only. Zaigraev and Podraza-Karakulska (2008) addressed the maximal scale invariant estimation procedure for the shape parameter of gamma distribution. Kulkarni and Patil (2018a) derived maximal scale invariant inference for the shape parameter under shape-scale family of distributions[29].

Tsui and Weerahandi (1989) developed the concept of generalized test variable (GTV) and generalized p-value (GPV) for significance testing based on a suitable generalized extreme region where the p-value is independent of the nuisance parameters[30]. Exact statistical inference based on GTV, GPV, and generalized confidence interval (GCI) can be found in Weerahandi (1995). Hannig et al. (2006) identified an important subclass of generalized pivotal quantities (GPQ) which have asymptomatically correct frequentist coverage. Nkurnziza and Chen (2011) provide a systematic approach to construct GPQ, GCI, and GPV for a location-scale family of distributions[30].

The present work reviews univariate, two-sample, and multi-sample improved procedures that efficiently handle the nuisance parameters and the recommended procedures are given in the next section.

4. IMPROVED INFERENTIAL PROCEDURES

Kulkarni and Patil (2018a)[31] introduced the maximal scale-invariant estimation procedure for the shape parameter of the shape-scale family of distributions. The method for obtaining nuisance parameters-free likelihood for the shape parameter based on maximal scale-invariant transformation for eliminating the nuisance scale parameter is explained. The resulting likelihoods are functions of only the shape parameter of interest. The results are illustrated for popular shape-scale distributions, namely the Weibull, the Gamma and the Generalized exponential (GE) distribution under complete and type-II censored samples. The proposed maximal scale-invariant likelihood estimator (MSILE) for the shape parameter of interest, being based on a proper likelihood function enjoys all asymptotic properties under regular conditions[31].

A simulation study for the Weibull and Gamma distributions revealed an almost exact relationship between the bias of the MSILE and the maximum likelihood estimator (MLE). An improved, almost unbiased estimator (AUE) is proposed by exploiting this linearity. The extent of reduction in bias and mean square error (MSE) of the MLE, MSILE and AUE reveals the superiority of MSILE over MLE, and the superiority of AUE over MSILE and MLE for Weibull and Gamma distribution[32]. One-sample test and $100(1-\alpha)\%$ confidence interval for the shape parameter is developed, and performance is assessed with respect to the observed size of relevant test procedures, and coverage probability and average width of the associated confidence interval. Furthermore, the MLE of the scale parameter being a function of the shape parameter, is obtained by replacing the shape parameter with its MSILE. The performance of the resulting estimator was observed to be superior than its regular MLE[33].

The interval estimation for the stress-strength reliability (R) under the exponentiated-scale family of distributions is developed in the Patil and Kulkarni (2018)[34]. The exponentiated-scale family was introduced by Marshall and Olkin (2007), which is also known as resilience or frailty parameter family. The distributional form of resilience family is:

$$G\left(\frac{x}{\lambda},\alpha\right)=F^{\alpha}\left(\frac{x}{\lambda}\right),$$

α being a resilience parameter, while the distributional form of frailty family is:

$$\bar{G}\left(\frac{x}{\lambda},\alpha\right)=\bar{F}^{\alpha}\left(\frac{x}{\lambda}\right),$$

α being a frailty parameter, λ the scale parameter, and $F\,(.)$ is a known distribution function while $\bar{F}\,(.)$ is the corresponding survival function.

The stress–strength reliability $R = P\,(X_1 < X_2)$ where X_1 and X_2 represent the stress applied and strength of an equipment, respectively, plays a crucial role in setting warranty periods while launching new brands of a product, among other applications. Patil and Kulkarni (2018) address the issue of estimating R when X_1 and X_2 belong to the exponentiated scale family, which includes the popular Exponentiated-exponential distribution (EED) that has proven to be an excellent model for lifetime distributions. The cases of known/unknown and equal/unequal scale parameters are handled separately. For equal scale parameters of X_1 and X_2 the expression for R turns out to be purely function of the shape parameters. When the scale parameters are unequal the reliability R turns out to be a function of the underlying shape parameter and ratio of the scale parameters. For known scale parameter, a generalized pivot quantity for the shape parameter and R are developed. The interval estimates of R based on the proposed generalized pivot quantity exhibited uniformly best performance. For an unknown scale parameter, a maximum scale invariant likelihood estimator of the shape and an allied estimator of the scale are introduced. An extensive simulation-based comparison is performed among following five methods:

GPQ: Generalized pivotal quantity.

PBMSILE: A parametric bootstrap technique employed on MSILE.

PBMLE: A Parametric bootstrap technique employed on MLE.

NPBMSILE: A nonparametric bootstrap technique employed on MSILE.

NPBMLE: A nonparametric bootstrap technique employed on MLE.

The parametric bootstrap interval estimates of R based on the proposed maximum scale invariant likelihood estimator of the shape parameter exhibited best performance among others. An application in setting warranty periods is illustrated based on two real data sets[35].

Micro-array experiments are important fields in molecular biology where zero values mixed with a continuous outcome are frequently encountered leading to a mixed distribution with a clump at zero. Comparison of two mixed populations, for example of a control and a treated group; of two groups with different types of cancer, to name a few, are often encountered in these contexts. Fairly skewed distribution of the continuous part coupled with small sample sizes are issues of main concern to be attended for the quality of inference in such situations. However, popularly used non-parametric methods rely on asymptotic distribution of the underlying test statistics which are valid only under large sample sizes. Kulkarni and Patil (2018b) address the aforementioned issues via a newly proposed exact test for location-scale family distributions and GPQ based parametric test procedures for non-location-scale distributions. The proposed test procedure can be used under a best fitted continuous distribution. It consists of k+1 parts, where k is the number of parameters for a specific best fitting parametric model used for the continuous component. More specifically, the first part tests the equality of the proportions of zeros while the remaining k parts test the equality of the k corresponding individual parameters in the two populations under consideration. Note that the combined test is equivalent to testing equality of the two entire mixed populations under consideration. The k+1 parts and their combination produce an overall p-value for testing the combined hypothesis of equality of the two distributions. In order to account for the dependency among simultaneous testing of a large number of tests, we calibrate the observed p-values using the Benjamini–Hochberg (1995) procedure[36].

A simulation study is carried out for validation and performance evaluation of the proposed exact test for location-scale or log-location-scale family of distributions and GPQ based test for non-location-scale distributions. The proposed test is compared with the popular two-part (TP) test based on the type-I error and power of the tests. The TP test consists of two parts one is of testing equality of proportions of zeros and other non-parametric test comparing two continuous data sets. Different tests are used to compare the continuous part, namely Kolmogorov- Smirnov, t-test, Wilcoxon rank sum test, Ansari Bradley test, Sigel-Tukey test[37].

Simulation based assessment of the proposed exact test based on invariance principle for location-scale family distributions and GPQ based parametric test procedures for non-location-scale distributions showed their superior performance with respect to size and power in comparison to the above popular two-part tests, more prominently for small sample sizes[38].

A number of distributions including the Exponential, Extreme value, Normal, Double exponential, Inverse Gaussian, Weibull, Pareto, Log-Normal and Gamma distributions have been handled to illustrate the above testing procedure for microarray data. We could identify 1555 differentially expressed genes[39].

Future scope on RNA sequence count data analysis through the GPQ and GTV for Poison and Negative binomial parameters is discussed, and a generalized test procedure is suggested for two discrete populations in similar lines.

Patil and Kulkarni (2022) developed a unified approach for testing homogeneity of variances among k (k > 2) independent location-scale populations. The proposed test is based on a generalized test variable. The GPV for testing homogeneity of variances is obtained by constructing GPQs for the k distinct scale parameters of the k populations. The performance of the proposed test is assessed through an extensive simulation study on popular location-scale families in comparison to the existing tests. The proposed test is uniformly superior over existing popularly used parametric and non-parametric tests in terms of type-I errors and power function. A systematic study to assess the impact of the extent of kurtosis and skewness is made through simulation studies under the Generalized Normal and Skew Normal distributions respectively[40-41].

A uniformly implementable small sample integrated likelihood ratio test for one way and two-way ANOVA under heteroscedasticity and normality is developed by Patil and Kulkarni (2021) which has an asymptotic chi-square distribution up to second order accuracy. Simple ad hoc corrective adjustments recommended for improving the small sample distributional performance make the test usable even for very small group sizes. Empirical assessment of the test reveals that the test exhibits uniformly well-concentrated sizes at the desired level and the maximal power, particularly under very small size groups. In similar lines, Patil and Kulkarni (2022) develop a test for analysis of medians for Birnbaum–Saunders distributed response to assess the impact of two interacting factors on the median, where no any test available in the literature.

Ma et. al. (2022) studied the statistical inference on the location parameter vector in the multivariate skew-normal model with unknown scale parameter and known shape parameter. Based on the distribution of the generalized Hotelling's T^2 statistic, confidence regions and hypothesis tests on the location parameter μ are obtained[42].

5. RECOMMENDATIONS

The GPQ or Fiducial approach-based procedures or invariance-based procedures are recommended as the best alternative to classical or popularly used inferential procedures in the presence of nuisance parameters and often work well even under small sample sizes. A maximal scale invariant inference for shape and allied inference on scale parameter is a substitute for classical maximum likelihood point and interval estimation as well as testing problem under shape-scale and exponentiated-scale family of distributions. Generalized variable approach and a maximal scale invariant transformation-based inference is recommended for the stress-strength reliability under exponentiated-scale family of distributions. Exact test based on fiducial inference is recommended for Comparison of two continuous populations mixed with point mass at zero and to test the homogeneity of variances among several independent location-scale populations. When GPQ/invariance principle-based procedures are not available, among the likelihood-based procedures, the integrated likelihood principle works the best.

ACKNOWLEDGEMENT

Authors are very much thankful to associate editor and Reviewers for helpful comments for possible publication of this article.

REFERENCES

1. Andrews, D. W. K., Ploberger, W. (1994) Econometrica, 62, 1383–1414.

2. Barndorff-Nielsen, O. E. (1985), Ed. J. Lanke and G. Lindgren, 12, 25–38.
3. Barndorff-Nielsen, O. E. (1991) Likelihood theory. Statistical Theory and Modelling, 96, 232–264.
4. Barndorff-Nielsen, O. E. (1994) Journal of Royal Statistical Sociaty, 56, 125–140.
5. Benjamini, Y. and Hochberg, Y. (1995) Journal of the Royal Statistical Society, Series B, 57, 289–300.
6. Berger, J. O., Liseo, B., Wolpert, R. L. (1999) Statistical Science, 14, 01–28.
7. Berger, J. O. and Wolpert, R. L. (1988) Institute of Mathematical Statistics Lecture Notes - Monograph Series, 19–64.
8. Cox, D. R. and Hinkley, D. V. (1974) Theoretical Statistics. London: Chapman and Hall.
9. Engelhardt, M. and Bain, L. J. (1977) Technometrics,19, 77–81.
10. Fraser, D. A. S. and Reid, N. (1989) Biometrika,76, 477–488.
11. Gong, G. and Samaniego, F. J. (1981) Annals of Statistics, 11, 68–77.
12. Hannig, J., Iyer, H., and Patterson, P. (2006) Journal of the American Statistical Association, 101, 245–269.
13. Hansen, B. E. (1996) Econometrica, 137, 3537–3547.
14. Jones, M. C. (2015) International Statistical Review, 83, 175-192.
15. Kalbfleish, J. D. Sprott, D. A. (1989) Journal of Royal Statistical Society, 32,175–208.
16. Kulkarni, H. V., Patil, K. P. (2018) Journal of Statistical Computation and Simulation, 88, 2259–2272.
17. Kulkarni, H. V. and Patil, K. P. (2018) Mathematical Biosciences, 298, 19–28.
18. Kulkarni, H. V., Powar, S. K. (2010) Lifetime data Analysis, 16, 431–447.
19. Kulkarni, H. V., Powar, S. K. (2011) Journal of Probability and Statistics,1–10.
20. Kulkarni, H. V., Sengupta A. (2020 Statistica Sinica, 30, 1995-2021.
21. Kulkarni, H. V. and Sengupta A. (2021) International Statistical Review,90, 41-61.
22. Kulkarni, H. V., S. M. Patil. (2021) Advances in Statistical Analysis 105, 273-305.
23. Linnik, J. V. and Technica, S. (1968) American MathematicalSoc.
24. Ma, Z., Chen, Y. J., Wang, T., & Liu, J. (2020) Communications in Statistics-Simulation and Computation, 1-17.
25. Marshall, A. W., Olkin, I. (2007) Springer Series in Statistics.
26. Maswadah, M. (2013) Journal of Theoretical Statistics 42, 115-136.
27. Maswadah, M. (2022) Journal of Applied Statistics 49, 2825-2844.
28. Nkurnziza, S. and Chen, F. (2011). Sankhya B, 73, 218–240.

29. Patil, K. P., Kulkarni, H. V. (2018) Quality and Reliability Engineering International, 33, 1447–1453.
30. Patil, K. P., Kulkarni, H. V. (2022) Journal of Statistical Computation and Simulation, 00,01–27.
31. Patil, S. C., Kulkarni, H. V. (2011) Applied Mathematical Sciences, 5, 459–476.
32. Patil, S. M., H. V. Kulkarni. (2022) Journal of Applied Statistics, 8, 1-22.
33. Powar, S. K., Kulkarni, H. V. (2015) Stochastic Environmental Research and Risk Assessment, 29, 1691–1708.
34. Rigby, R. A., & Stasinopoulos, D. M. (2005) Journal of the Royal Statistical Society: Series C, 54, 507-554.
35. Rigby, R. A., Stasinopoulos, M. D., Heller, G. Z., De Bastiani, F. (2019) CRC press.
36. SenGupta, A., Kulkarni, H. V., Hubale, U. D. (2015) Environmental and Ecological Statistics, 22, 87–104.
37. Severini, T. A. (1998) Biometrika, 85, 507–522.
38. Severini, T. A. (2000) Likelihood methods in statistics, Oxford University Press, New York.
39. Severini, T. A. (2010) Biometrika, 97, 481–496.
40. Tsui, K., Weerahandi, S. (1989) Journal of the American Statistical Association, 84, 602–607.
41. Weerahandi, S. (1995) Springer Series in Statistics.
42. Zaigraev, A. Podraza-Karakulska, A. (2008) Statistics and Probability Letters, 78(3),286–29.

The Study of Exponentiated Gumbel Distribution and Related Inference Through Simulation

Chandrakant S. Kakade [a*]
[a]Anandibai Raorane Arts , Commerce and Science College, Vaibhavwadi 416 810 (MS) India.
[*]Corresponding author: cskboss08@gmail.com

ABSTRACT

Two parameter Exponentiated Gumbel (EG) distribution is a right skewed unimodal distribution. We discuss point and interval estimation of parameters of EG distribution by the method of maximum likelihood and provide an expression for the Fisher information matrix. A bootstrap method to obtain confidence interval is also discussed. Inference for R=P(Y<X) is provided when X and Y are independently but not identically EG distributed random variables. Testing for R based on exact and asymptotic distribution is discussed along with simulation study.

KEYWORDS

Maximum likelihood estimator, Fisher information matrix, uniformly minimum variance unbiased estimator and Bayes' estimator.

……………………………………………………………………………………………

1. INTRODUCTION

In literature, exponentiated family of distribution defined in two ways. If $F(x/\underline{\theta})$ is cumulative distribution function (c.d.f.) of base line distribution then by adding one more parameter (say α), the c.d.f. of exponentiated base line distribution is $G(x/\underline{\theta},\alpha)$ given by

(a) $G(x/\underline{\theta},\alpha)= [\ F(x/\underline{\theta})\ ]^{\alpha}$, $\alpha>0$, $\underline{\theta}\in\Theta$ and $x\in R$.

(b) $G(x/\underline{\theta},\alpha)= 1-[1- F(x/\underline{\theta})\ ]^{\alpha}$, $\alpha>0$, $\underline{\theta}\in\Theta$ and $x\in R$.

Gupta et al. (1998) introduced the Exponentiated Exponential (EE) distribution as a generalization of the standard Exponential distribution. The two parameter EE distribution associated with definition (a) above, have been studied in detail by Gupta and Kundu (2001) which is a sub-model of the Exponentiated Weibull distribution, introduced by Mudholkar and Shrivastava (1993). S. Nadarajah (2006) introduced Exponentiated Gumbel (EG) distribution using (b) above.

The cumulative distribution function of the EG distribution is defined by

$$F(x;\alpha,\sigma) = (G(x,\sigma))^{\alpha} = \left(\exp\left(-e^{-\frac{x}{\sigma}} \right) \right)^{\alpha}$$

$,\alpha,\ \sigma>0\ ,\ -\infty<x<\infty$ (1.1)

which is simply the α^{th} power of c.d.f. of the Gumbel distribution.

The Probability density function (p.d.f.) corresponding to (1.1) is

$$f(x;\alpha,\sigma) = \frac{\alpha}{\sigma} \left(\exp\left(-e^{-\frac{x}{\sigma}} \right) \right)^{\alpha} e^{-\frac{x}{\sigma}}$$

$,\ \alpha,\ \sigma>0\ ,\ -\infty<x<\infty$ (1.2)

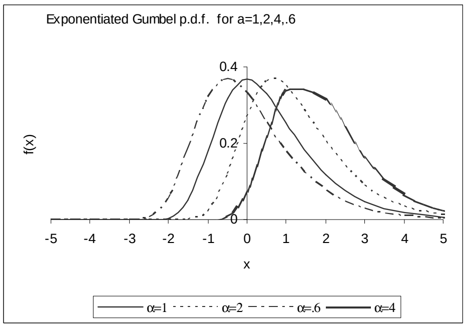


Figure-1. Probability density function.

We shall write x ~ EG(α,σ) to denote an absolutely continuous random variable X having the EG distribution with shape and scale parameters are α and σ respectively whose p.d.f. is given by (1.2). The shapes of p.d.f. for EG distribution with scale parameter σ=1 and various values of parameter α (=1, 2, 4, 0.6) are shown in the above Figures. Fig. 1 shows that it is an unimodal and right skewed density function.

2. MAXIMUM LIKELIHOOD ESTIMATOR AND THE FISHER INFORMATION MATRIX

Suppose $X_1,X_2,\ldots.X_n$ is a random sample from EG(α,σ). Therefore, the log-likelihood function L for the observed sample is

$$L = n\ln\alpha - n\ln\sigma - \frac{1}{\sigma}\sum_{i=1}^{n} x_i - \alpha\sum_{i=1}^{n} e^{-\frac{x_i}{\sigma}} \quad (2.1)$$

Therefore, to obtain the MLE's of α and σ , either we can maximize (2.1) directly with respect to α and σ or we can solve the non-linear normal equations which are

$$\frac{\partial L}{\partial \alpha} = \frac{n}{\alpha} - \sum_{i=1}^{n} e^{-\frac{x_i}{\sigma}} = 0 \qquad (2.2)$$

$$\frac{\partial L}{\partial \sigma} = \frac{-n}{\sigma} + \frac{1}{\sigma^2}\sum_{i=1}^{n} x_i - \frac{\alpha}{\sigma^2}\sum_{i=1}^{n} x_i\, e^{-\frac{x_i}{\sigma}} = 0 \qquad (2.3)$$

From (2.2), we obtain the MLE's of α as a function of σ, say $\hat{\alpha}(\sigma)$ as

$$\hat{\alpha}(\sigma) = \frac{n}{\sum_{i=1}^{n} e^{-\frac{x_i}{\sigma}}} \qquad (2.4)$$

Case 1: If the scale parameter is known (say σ=1), the MLE of the parameter α can be obtained directly from (2.4).

Lemma (2.1): For known scale parameter (say σ=1) the p.d.f. of $\hat{\alpha}$ is

$$f_Y(y,\alpha) = \frac{1}{n!\ \ \alpha}\left(\frac{n\alpha}{y}\right)^{n+1} e^{-\frac{n\alpha}{y}} \ , \ y>0 \qquad (2.5)$$

Proof : Suppose $W = \left(-2\alpha \sum \ln(\exp(-e^{-x_i}))\right)$ then W has chi-square distribution with 2n d.f., since $\left(\exp(-e^{-x_i})\right)^{\alpha}$ is c.d.f. of standard EG distribution and follows uniform distribution over (0,1). Let $Y = \frac{2n\alpha}{W}$, then c.d.f. of Y is given as

$$P(Y \le y) = P\left(\frac{2n\alpha}{W} \le y\right) = 1 - P\left(W \le \frac{2n\alpha}{y}\right) \qquad (2.6)$$

Using Chi-square distribution , the p.d.f. corresponding to (2.6) is

$$f_Y(y,\alpha) = \frac{1}{n!\,\alpha}\left(\frac{n\alpha}{y}\right)^{n+1} e^{-\frac{n\alpha}{y}} \ ,y>0$$

Lemma (2.2): For known scale parameter (say σ=1) , the 100(1-δ)% confidence interval of α is given by

$$\left(\frac{Y}{2n}\chi^2_{2n,\delta/2},\quad \frac{Y}{2n}\chi^2_{2n,1-\delta/2}\right).$$

Case 2: If both the parameters are unknown, first the estimate of the scale parameter can be obtained by using maximum likelihood estimation method

$$L\,(\hat{\alpha}(\sigma),\sigma) = C - n\ln\sum_{i=1}^{n} e^{-\frac{x_i}{\sigma}} - n\ln\sigma - \frac{1}{\sigma}\sum_{i=1}^{n} x_i \qquad (2.7)$$

With respect to σ. Here C is a constant independent of σ. Once $\hat{\sigma}$ is obtained, $\hat{\alpha}$ can be obtained from (2.4) as $\hat{\alpha}(\sigma)$. Therefore, it reduces the two-dimensional problem to a one-dimensional problem.

In this situation we use the asymptotic normality result to obtain the asymptotic confidence interval. We can state the result as follows.

$\sqrt{n}(\hat{\theta}-\theta) \rightarrow N_2(0, I^{-1}(\theta))$ where I(θ) is the Fisher Information matrix.

$$I(\theta) = \frac{-1}{n}\begin{bmatrix} E\left(\frac{\partial^2 L}{\partial\alpha^2}\right) & E\left(\frac{\partial^2 L}{\partial\alpha\partial\sigma}\right) \\ E\left(\frac{\partial^2 L}{\partial\sigma\partial\alpha}\right) & E\left(\frac{\partial^2 L}{\partial\sigma^2}\right) \end{bmatrix}$$ and $\hat{\theta} = (\hat{\alpha}, \hat{\sigma})$, θ=(α,σ),

$$E\left(\frac{\partial^2 L}{\partial\alpha^2}\right) = \frac{-n}{\alpha^2}, \qquad E\left(\frac{\partial^2 L}{\partial\alpha\partial\sigma}\right) = \frac{\alpha}{\sigma 2^{\alpha}}\sum_{i=1}^{n} E(\ln u_i),$$

$$E\left(\frac{\partial^2 L}{\partial\sigma^2}\right) = \frac{n}{\sigma^2} - \frac{2}{\sigma^2}\sum_{i=1}^{n} E(\ln v_i) - \frac{2\alpha^2}{2^{\alpha}\sigma^2}\sum_{i=1}^{n} E(\ln u_i) - \frac{\alpha^2}{2^{\alpha}\sigma^2}\sum_{i=1}^{n} E(\ln u_i)^2$$

where u_i and v_i has gamma distribution with parameters (2,α) and (1,α) respectively. Since θ is unknown, $I^{-1}(\theta)$ is estimated by replacing θ with its MLE and this can be used to obtain the asymptotic confidence intervals of α and σ.

2.1. Bootstrap Confidence Interval:

In this subsection, we propose a percentile bootstrap method (Efron, 1982) for constructing confidence interval of α and σ which is as follows.

Step-1: Generate random samples $x_1, x_2, \ldots.. x_n$ from EG(α,σ) and compute $\hat{\alpha}$ and $\hat{\sigma}$ using maximum likelihood method.

Step-2: Using $\hat{\alpha}$ and $\hat{\sigma}$ generate a bootstrap sample $x_1^*, x_2^*, \ldots, x_n^*$ from ES($\hat{\alpha}$, $\hat{\sigma}$). Based on bootstrap samples compute bootstrap estimate $\hat{\alpha}^*$ and $\hat{\sigma}^*$.

Step-3: Repeat step-2 NBOOT times (usually NBOOT=1000).

Step-4: Compute cumulative distribution function of $\hat{\alpha}^*$ and $\hat{\sigma}^*$, say H(x) and G(x) respectively , where H(x)= P($\hat{\alpha}^* \le$ x) and $\hat{\alpha}_{Boot-p}(x) = H^{-1}(x)$ and G(x)= P($\hat{\sigma}^* \le$ x) and $\hat{\sigma}_{Boot-p}(x) = G^{-1}(x)$ for a given x. The approximate 100(1-δ)% bootstrap confidence intervals for α and σ are given by

$\left(\hat{\alpha}_{Boot-p}(\delta/2)\, ,\, \hat{\alpha}_{Boot-p}(1-\delta/2)\right)$ and $\left(\hat{\sigma}_{Boot-p}(\delta/2)\, ,\, \hat{\sigma}_{Boot-p}(1-\delta/2)\right)$ respectively.

3. POINT AND INTERVAL ESTIMATION OF R

Now we consider the problem of estimating R=P(Y<X) when X and Y are independent EG random variables with shape, scale parameters α , σ and β , σ respectively then R=P(Y<X) $= \dfrac{\alpha}{\alpha+\beta}$

Case 1: When scale parameter σ is unknown.

Suppose $X_1, X_2, \ldots X_n$ is a random sample from EG(α,σ) and $Y_1, Y_2, \ldots Y_m$ is a random sample from EG(β,σ). Therefore, the log-likelihood function L of α, β and σ for the observed sample is

$$L = n\ln\alpha - \alpha\sum_{i=1}^{n} e^{-\frac{x_i}{\sigma}} + m\ln\beta - \beta\sum_{j=1}^{m} e^{-\frac{y_j}{\sigma}} - (m+n)\ln\sigma - \frac{\left(\sum_{i=1}^{n} x_i + \sum_{j=1}^{m} y_j\right)}{\sigma}, \qquad (3.1)$$

hence MLE's of α and β as $\hat{\alpha} = \dfrac{n}{\sum_{i=1}^{n} \exp(-\frac{x_i}{\hat{\sigma}})}$ and $\hat{\beta} = \dfrac{m}{\sum_{j=1}^{m} \exp(-\frac{y_j}{\hat{\sigma}})}$.

Therefore, the MLE of R namely $\hat{R}_1$ is given by

$$\hat{R}_1 = \frac{\hat{\alpha}}{\hat{\alpha}+\hat{\beta}} \qquad (3.2)$$

Now to obtain asymptotic distribution of R, we first obtain the asymptotic distribution of $(\hat{\alpha}, \hat{\beta}, \hat{\sigma})$. Based on the asymptotic distribution of $\hat{R}$, we obtain asymptotic confidence interval of R. Let us denote the Fisher Information matrix of (α,β,σ) as I(α,β,σ) where

$$I(\alpha,\beta,\sigma) = -\begin{bmatrix} E\left(\frac{\partial^2 L}{\partial\alpha^2}\right) & E\left(\frac{\partial^2 L}{\partial\alpha\partial\beta}\right) & E\left(\frac{\partial^2 L}{\partial\alpha\partial\sigma}\right) \\ E\left(\frac{\partial^2 L}{\partial\beta\partial\alpha}\right) & E\left(\frac{\partial^2 L}{\partial\beta^2}\right) & E\left(\frac{\partial^2 L}{\partial\beta\partial\sigma}\right) \\ E\left(\frac{\partial^2 L}{\partial\sigma\partial\alpha}\right) & E\left(\frac{\partial^2 L}{\partial\sigma\partial\beta}\right) & E\left(\frac{\partial^2 L}{\partial\sigma^2}\right) \end{bmatrix}$$

$$= \begin{bmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{bmatrix} \quad \text{say.}$$

Moreover $E\left(\frac{\partial^2 L}{\partial\alpha^2}\right) = -\frac{n}{\alpha^2}$ and $E\left(\frac{\partial^2 L}{\partial\beta^2}\right) = -\frac{m}{\beta^2}$, $E\left(\frac{\partial^2 L}{\partial\alpha\partial\beta}\right) = E\left(\frac{\partial^2 L}{\partial\beta\partial\alpha}\right) = 0$,

$$E\left(\frac{\partial^2 L}{\partial\alpha\partial\sigma}\right) = \frac{\alpha}{\sigma\, 2^{\alpha}}\sum_{i=1}^{n} E(\ln u_i) = E\left(\frac{\partial^2 L}{\partial\sigma\partial\alpha}\right) \quad E\left(\frac{\partial^2 L}{\partial\beta\partial\sigma}\right) = \frac{\beta}{\sigma\, 2^{\beta}}\sum_{j=1}^{m} E(\ln v_j) = E\left(\frac{\partial^2 L}{\partial\sigma\partial\alpha}\right)$$

$$E\left(\frac{\partial^2 L}{\partial\sigma^2}\right) = \frac{n}{\sigma^2} - \frac{2}{\sigma^2}\sum_{i=1}^{n} E(\ln w_i) - \frac{\alpha^2}{\sigma^2\, 2^{\alpha-1}}\sum_{i=1}^{n} E(\ln u_i) - \frac{\alpha^2}{\sigma^2\, 2^{\alpha}}\sum_{i=1}^{n} E(\ln u_i)^2$$

$$+\frac{m}{\sigma^2} - \frac{2}{\sigma^2}\sum_{j=1}^{m} E(\ln z_j) - \frac{\beta^2}{\sigma^2\, 2^{\beta-1}}\sum_{j=1}^{m} E(\ln v_j) - \frac{\beta^2}{\sigma^2\, 2^{\beta}}\sum_{j=1}^{m} E(\ln v_j)^2$$

where u_i and v_j has gamma $(2,\alpha)$ and $(2,\beta)$ and w_i and z_j has exponential α and β distribution respectively.

Theorem 1: As m, n $\to\infty$ and $\frac{m}{n} \to p$ then

$$\left((\hat\alpha-\alpha),(\hat\beta-\beta),(\hat\sigma-\sigma)\right) \to N_3(0, A(\alpha,\beta,\sigma)),$$

where $A(\alpha,\beta,\theta) = \begin{bmatrix} a_{11} & 0 & a_{13} \\ 0 & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ and elements of A(α ,β, σ) are the corresponding elements of the inverse of the Fisher Information matrix I(α, β, σ).

Proof : Proof follows from asymptotic properties of MLEs under regularity conditions and multivariate central limit theorem.

Theorem 2: As m, n $\to\infty$ and $\frac{n}{m} \to p$ then $\sqrt{n}(\hat R - R) \to N(0, B)$, where

$$B=\frac{1}{u(\alpha+\beta)^4}\left(\beta^2\left(a_{22}a_{33}-a_{23}^2\right)-2\alpha\beta\sqrt{p}a_{23}a_{31}+\alpha^2 p\left(a_{11}a_{33}-a_{13}^2\right)\right)$$

and $u=a_{11}a_{22}a_{33}-a_{11}a_{23}a_{32}-a_{13}a_{22}a_{31}.$

Proof : Proof follows from invariance property of CAN estimator under continuous transformation, and omitted for brevity.

Using Theorem 2, we can obtain asymptotic confidence interval of R as

$$\left(\hat{R}-Z_{1-\delta/2}\frac{\sqrt{\hat{B}}}{\sqrt{n}},\quad \hat{R}+Z_{1-\delta/2}\frac{\sqrt{\hat{B}}}{\sqrt{n}}\right) \qquad (3.3)$$

Remark (3.1): To estimate variance B, the empirical Fisher's information matrix and MLEs of α, β and σ may be used. However simulation study due to Kundu and Gupta (2005) for EE distribution indicates that confidence interval defined in (3.3) has comparatively low coverage probability. They have suggested bootstrap method to get a better confidence interval with respect to coverage probability.

Bootstrap confidence interval:

Step-1: Generate random samples $x_1,x_2,\ldots..x_n$ from ES(α,σ) and $y_1,y_2,\ldots.y_m$ from ES(β,σ) and compute $\hat{\alpha},\hat{\beta}$ and $\hat{\sigma}$ using maximum likelihood method.

Step-2: Using $\hat{\alpha}$ and $\hat{\sigma}$ generate a bootstrap sample $x_1^*,x_2^*,\ldots.,x_n^*$ from ES($\hat{\alpha}$, $\hat{\sigma}$) and similarly using $\hat{\beta}$ and $\hat{\sigma}$ generate a bootstrap sample $y_1^*,y_2^*,\ldots.,y_m^*$ from ES($\hat{\beta}$, $\hat{\sigma}$). Based on these bootstrap samples compute bootstrap estimate of R,

$\hat{R}^*=\dfrac{\hat{\alpha}^*}{\hat{\alpha}^*+\hat{\beta}^*}$, where $\hat{\alpha}^*$ and $\hat{\beta}^*$ are the MLEs of α and β obtained from the corresponding bootstrap samples.

Step-3: Repeat step-2 NBOOT times (usually NBOOT=1000).

Step-4: Compute cumulative distribution function of $\hat{R}^*$, say H(x), where

H(x)= P($\hat{R}^*$ ≤ x) and $\hat{R}_{Boot-p}(x)=H^{-1}(x)$ for a given x. The approximate 100(1-δ)% bootstrap confidence interval is given by

$$\left(\hat{R}_{Boot-p}(\delta/2)\,,\,\hat{R}_{Boot-p}(1-\delta/2)\right) \qquad (3.4)$$

Case 2: When scale parameter σ is known.

Without loss of generality, we can assume that σ=1. Suppose $X_1, X_2, \ldots, X_n$ is a random sample from EG(α,1) and $Y_1, Y_2, \ldots, Y_m$ is a random sample from EG(β,1) and based on the samples we want to estimate R. Based on the above sample, it is clear that , the MLE of R namely $\hat{R}_2$ is given by $\hat{R}_2 = \frac{\hat{\alpha}}{\hat{\alpha}+\hat{\beta}}$ where

$\hat{\alpha} = \frac{n}{\sum_{i=1}^{n} \exp(-x_i)}$ and $\hat{\beta} = \frac{m}{\sum_{j=1}^{m} \exp(-y_j)}$

Lemma (3.1) : The p.d.f. of $\hat{R}_2$ is given by

$$f_{\hat{R}_2}(r) = \frac{\Gamma(m+n)}{\Gamma m \, \Gamma n} \left(\frac{n}{m}\right)^n \left(\frac{\alpha}{\beta}\right)^{n-1} \frac{\left(\frac{1-r}{r}\right)^{n-1}}{\left(1+\frac{n\alpha}{m\beta}\left(\frac{1-r}{r}\right)\right)^{m+n}} ,$$

0<r<1 (3.5)

Proof : $\hat{R}_2$ can be expressed as $\hat{R}_2 = \frac{1}{1+\frac{mW}{nV}}$ Where

$W = -\sum \ln\left(\exp(-e^{-x_i})\right)$ and $V = -\sum \ln\left(\exp(-e^{-y_j})\right)$. We see that -2αW and -2βV are two independent chi-square random variables with 2n and 2m degrees of freedom (d.f.) respectively. Therefore $\hat{R}_2$ can be rewritten as $\hat{R}_2 = \left(1+\frac{\beta}{\alpha}Z\right)^{-1}$, where Z $= \frac{-2\alpha W/2n}{-2\beta V/2m}$ has F distribution with (2n, 2m) degrees of freedom (d.f.). Therefore p.d.f. of $\hat{R}_2$ can be obtained easily and is as given in equation (3.7).

Lemma (3.2) : An exact 100(1-γ)% confidence interval of R is

$$\left(\left(1+F_{2m,2n;(1-\gamma/2)}\left(\frac{1}{\hat{R}_2}-1\right)\right)^{-1}, \left(1+F_{2m,2n;(\gamma/2)}\left(\frac{1}{\hat{R}_2}-1\right)\right)^{-1}\right) \qquad (3.6)$$

Lemma (3.3) : The asymptotic 100(1-γ)% confidence interval of R is

$$\left(\left(\hat{R}_2 - Z_{1-\gamma/2}\sqrt{\frac{m+n}{mn}} \quad \hat{R}_2(1-\hat{R}_2)\right), \left(\hat{R}_2 + Z_{1-\gamma/2}\sqrt{\frac{m+n}{mn}} \quad \hat{R}_2(1-\hat{R}_2)\right)\right) \qquad (3.7)$$

where $Z_{1-\gamma/2}$ is the $(1-\gamma/2)^{th}$ quantile of the standard normal distribution.

Proof : The MLE $\hat{R}_2$ is asymptotically normal with mean R and variance

$\sigma^2_{\hat{R}_1} = \sum_{i=1}^{2}\sum_{j=1}^{2} \frac{\partial R}{\partial \theta_i}\frac{\partial R}{\partial \theta_j} I^{-1}_{ij}$ where $(\theta_1,\theta_2) = (\alpha,\beta)$ and I^{-1}_{ij} is the $(i,j)^{th}$ element of the inverse of the Fisher's information matrix $I(\alpha,\beta)$ about the parameters (α,β) and

$I(\alpha,\beta) = -\begin{bmatrix} \frac{n}{\alpha^2} & 0 \\ 0 & \frac{m}{\beta^2} \end{bmatrix}$, (See Rao (1965)). It can be seen that, $\sigma^2_{\hat{R}_1} = \left(\frac{m+n}{mn}\right) R^2 (1-R)^2$.

Therefore the asymptotic 100(1-γ)% confidence interval of R can be obtained using standardized statistic as a pivotal quantity. We replace 'R' in the asymptotic variance by its MLE.

We perform some simulation experiments using percentile bootstrap method when scale parameter σ is unknown to observe the behavior of the MLE and confidence intervals for various sample sizes and for various values of (α, β). We consider the sample sizes (n, m)= (10,10), (10, 20), (20, 20), (20, 40), (40, 40) and the parameter values α= 2, σ=4 and β = 2, 3, 6 and 8. Average biases and mean squared errors (MSEs) of R are reported over 1000 replications for 1000 bootstrap samples. We compute 95% confidence intervals using (3.4) and estimate coverage percentages and average lengths of confidence interval. The results are reported in Table 1.

We also perform some simulation experiments when scale parameter σ is known (σ=1). We consider the sample sizes (n, m)= (10,10), (10, 20), (20, 20), (20, 40), (40, 40) and the parameter values α= 2 and β = 2, 3, 6 and 8. Average biases and mean squared errors (MSEs) of R are reported over 10000 replications. We compute 95% confidence intervals and estimate coverage percentages and average lengths of both asymptotic and exact confidence interval. The results are reported in Table 2.

Table-1. Biases, MSEs, Confidence Lengths and Coverage Percentages of C. I.

Sample size	**2**	**3**	**6**	**8**
(10, 10)	-0.0058(0.0131) 0.4273(93.00)	-0.0005 (0.0124) 0.4139 (93.00)	-0.0096 (0.0077) 0.3286 (90.70)	-0.0054 (0.0061) 0.2899 (91.40)
(10, 20)	0.0125 (0.0109) 0.3748 (92.40)	0.0095 (0.0097) 0.3672 (0.9410)	0.0088 (0.0067) 0.3052 (93.10)	0.0011 (0.0050) 0.2643 (92.90)

(20, 20)	-0.0018 (0.0067) 0.3120 (93.70)	-0.0018 (0.0070) 0.3013 (92.80)	-0.0044 (0.0046) 0.2454 (91.50)	-0.0062 (0.0031) 0.2144 (92.30)
(20, 40)	0.0057 (0.0050) 0.2706 (94.00)	0.0067 (0.0050) 0.2630 (93.50)	0.0031 (0.0033) 0.2175 (93.90)	-0.0001 (0.0026) 0.1909 (93.40)
(40, 40)	0.0012 (0.0033) 0.2205 (94.20)	-0.0032 (0.0031) 0.2134 (94.40)	-0.0049 (0.0021) 0.1762 (93.90)	-0.0028 (0.0016) 0.1567 (93.60)

(The first row represent the average biases and MSEs. Second row represent the average length, coverage percentages of the corresponding asymptotic bootstrap confidence interval.)

Table-2. Biases, MSEs, Confidence Lengths and Coverage Percentages of C. I.

Sample size	**2**	**3**	**6**	**8**
(10, 10)	- 0.0003(0.0119) 0.4174(91.47) 0.4058(94.83)	0.0033(0.0110) 0.4027(91.86) 0.3935(95.22)	0.0087(0.0073) 0.3237(91.77) 0.3258(95.30)	0.0098(0.0056) 0.2810(91.50) 0.2876(94.93)
(10, 20)	0.0042(0.0090) 0.3659(92.60) 0.3581(94.70)	0.0093(0.0086) 0.3542(92.30) 0.3507(94.46)	0.0120(0.0057) 0.2851(93.52) 0.2927(94.72)	0.0105(0.0043) 0.2459(92.90) 0.2572(94.74)
(20, 20)	- 0.0018(0.0060) 0.3024(93.45) 0.2977(95.02)	0.0016(0.0057) 0.2909(93.11) 0.2872(94.78)	0.0056(0.0037) 0.2313(93.15) 0.2323(94.61)	0.0045(0.0026) 0.1984(93.49) 0.2012(95.18)
(20, 40)	0.0025(0.0045) 0.2636(94.03) 0.2605(95.15)	0.0040(0.0043) 0.2539(93.83) 0.2527(94.98)	0.006290.0027) 0.2017994.220 0.2048(94.910	0.0056(0.0021) 0.1732(94.10) 0.1776(94.92)
(40, 40)	-	0.0016(0.0028)	0.0018(0.0018)	0.0021(0.0013)

	0.0009(0.0030) 0.2165(94.57) 0.2147(95.39)	0.2082(94.25) 0.2068(95.15)	0.1636(94.22) 0.1640(95.04)	0.1402(94.40) 0.1413(95.23)

(The first rows represent the average biases and the corresponding MSEs are reported within brackets. Second and third rows represent the average lengths and the corresponding coverage percentages of the asymptotic and exact confidence intervals respectively.)

Based on the proposed Bootstrap and exact method, the overall findings in Tables 1 and 2 are satisfactory. When sample sizes are increased, bias and MSE decrease for each parameter value, demonstrating the consistency of the method. In each case's coverage probability closely

approximates the confidence coefficient, and the average length of the confidence interval is small and finite.

4. TESTING OF HYPOTHESIS

The EG distribution is ordered with respect to the 'likelihood ratio' ordering ($X \leq_{lr} Y$). Since α and β both are unknown, it will be of interest to know whether $\alpha<\beta$ or not. We put this as a problem of hypothesis testing. We consider test for hypothesis $H_0:\alpha \leq \beta$ against $H_1 : \alpha > \beta$. Equivalently we can test H_0: $R \leq 0.5$ against H_1: $R > 0.5$. Using Lemma (3.3), an asymptotic test of size γ rejects the null hypothesis if, $\left(\hat{R}_2 - \frac{1}{2}\right) > \sqrt{\frac{m+n}{16\,mn}} \quad Z_{1-\gamma}$ (4.1)

where $Z_{1-\gamma}$ is the $(1-\gamma)^{th}$ quantile of the standard normal distribution. Also an exact test of size γ for the above problem, using lemma (3.2), rejects the null hypothesis if $\left(\frac{\hat{R}_2}{1-\hat{R}_2}\right) > F_{2n,2m;1-\gamma}$, (4.2) where $F_{2n,\,2m;\,1-\gamma}$ is the $(1-\gamma)^{th}$ quantile of F distribution with (2n, 2m) d.f. As an independent interest, we can also obtain an asymptotic and exact test of the desired size for alternatives H'_1: $R<0.5$ and H''_1: $R \neq 0.5$.

Through simulation study, comparison of power has been made for two test given in (5.1) and (5.2). The power was determined by generating 1000 random samples of sizes (n, m)=(10,10), (10,20), (20,20), (20,40) and (40,40). The results for the tests at the significance level γ=0.01 and 0.05 are presented in Table 3 and Table 4 respectively. P_1 and P_2 are referred to as power based on asymptotic and exact test as defined in (5.1) and (5.2) respectively.

Table 3 : Power of the test based on asymptotic and exact distribution of R, γ=0.01.

R	(10, 10)		(10, 20)		(20, 20)		(20, 40)		(40, 40)	
	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2
0.500	0.006	0.089	0.010	0.020	0.007	0.009	0.012	0.017	0.009	0.010
0.526	0.011	0.016	0.018	0.034	0.019	0.022	0.030	0.040	0.027	0.030
0.555	0.021	0.029	0.039	0.066	0.048	0.042	0.063	0.085	0.082	0.087
0.588	0.039	0.057	0.072	0.113	0.098	0.109	0.142	0.181	0.210	0.223
0.625	0.079	0.105	0.137	0.203	0.208	0.227	0.307	0.368	0.464	0.478
0.666	0.159	0.208	0.257	0.347	0.408	0.434	0.556	0.620	0.766	0.777
0.714	0.301	0.366	0.460	0.567	0.675	0.701	0.839	0.876	0.955	0.958
0.769	0.539	0.606	0.744	0.827	0.914	0.924	0.980	0.987	0.998	0.999
0.833	0.840	0.879	0.956	0.978	0.995	0.996	0.999	0.999	1	1
0.909	0.992	0.995	0.999	0.999	1	1	1	1	1	1

Table 4 : Power of the test based on asymptotic and exact distribution of R, γ=0.05.

R	(10, 10)		(10, 20)		(20, 20)		(20, 40)		(40, 40)	
	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2
0.500	0.046	0.050	0.057	0.073	0.046	0.047	0.057	0.065	0.047	0.047
0.526	0.069	0.073	0.088	0.109	0.085	0.087	0.114	0.129	0.116	0.116
0.555	0.119	0.125	0.148	0.179	0.171	0.175	0.207	0.229	0.257	0.258
0.588	0.178	0.189	0.240	0.278	0.293	0.298	0.370	0.402	0.479	0.482
0.625	0.282	0.293	0.376	0.422	0.479	0.485	0.591	0.622	0.728	0.730
0.666	0.431	0.444	0.564	0.610	0.697	0.702	0.815	0.837	0.920	0.921
0.714	0.627	0.639	0.765	0.804	0.881	0.884	0.959	0.967	0.992	0.992
0.769	0.830	0.840	0.931	0.945	0.981	0.982	0.997	0.998	1	1
0.833	0.966	0.968	0.995	0.996	0.999	0.999	1	1	1	1
0.909	0.999	0.999	1	1	1	1	1	1	1	1

It is observed from the simulation study that (i) both the tests perform well with respect to the power. (ii) Power of the test based on exact test is slightly higher than that of asymptotic test. (iii) Both the tests are consistent in the sense that as sample sizes increase, their power show improvement. (iv) A comparison with the usual nonparametric Wilcoxon Mann Whitney test for H_0: $P(Y<X)=0.5$ was made. It is found that parametric procedure (i.e., exact and asymptotic test) have better power than the more general WMW-test.

5. CONCLUSIONS

In this paper we estimate reliability R for Exponentiated Gumbel distribution with different shape parameters and same scale parameter. The performance of the MLE is quite satisfactory in terms of biases and MSEs. It is observed that when sample sizes increase the MSEs decreases. It verifies the consistency property of the MLE of R. The exact distribution of MLE of R is obtained and used for constructing confidence interval. The asymptotic confidence interval based on the MLE of R also works well for samples of sizes greater than or equal to 20. The exact as well as asymptotic test for testing reliability R has been given. The performances of both the tests are satisfactory with respect to the power than usual nonparametric Wilcoxon Mann Whitney test.

REFERENCES

1. Efron, B. (1982) CBMS-NSF Regional Conference series in Applied Mathematics, 38, SIAM, Philadelphia, PA.
2. Ferguson, T.S., 1967, Mathematical Statistics, A Decision Theoretic Approach, Academic Press, New York.
3. Gupta R.C., Gupta P.L. AND Gupta R.D. (1998) Communication Statistics-Theory and Methods, 27, 887-904.
4. Kundu D., Gupta R. D. (2005) Metrika, 6, 291-308.
5. Mudholkar G. S., Srivastava D., Freimer M. (1995) Technometrics, 37(4), 436-445.
6. Nadarajah S. (2005) Environmetrics, 17, 13-23.
7. Rao, C.R. (1965) New York: John Wiley and Sons, 2, 455-480.
8. Raqab, M.Z., Kundu, D. (2005) Communication in Statistics-Computations and Simulations, 34, 465-483

On the Performance of Different Robust Criterion Functions based M-Estimators and RM-Estimators in the presence of Multicollinearity and Outliers

Nileshkumar H. Jadhav[a,]*, Dattatray N. Kashidb[b]

[a]Deshbhakta Ratnappa Kumbhar College of Commerce, Kolhapur 416 002 (MS) India.
[b]Department of Statistics, Shivaji University, Kolhapur 416 004 (MS) India.
*Corresponding author: n.nil08@gmail.com

ABSTRACT

A simultaneous occurrence of multicollinearity and outliers is one of the important problems in regression analysis. It dramatically affects not only the least squares estimator (LSE) but also the ridge regression estimator (RRE) as well as M-estimator (ME). Consequently, the inference based on the LSE, RRE and ME gives misleading results. To deal with the problem of multicollinearity and outliers, Silvapulle (1991) proposed and studied the performance of Huber's robust criterion function-based ridge M-estimator (RME). However, there are various robust criterion functions available in the literature. In this article, we have obtained the ME and RME based on the different robust criterion functions. An extensive simulation study is performed to compare the ME and RME through mean squared error sense when data suffers from the problem of only multicollinearity, only outliers and both, multicollinearity and outliers.

KEYWORDS

Multicollinearity, Outliers, Ridge M-estimator, Robust criterion functions, MSE.

1. INTRODUCTION

In real-life data analysis, while applying a multiple linear regression model, the violations of classical assumptions like linearity, non-normality, independence of covariates are commonly occurring problems[1]. The occurrence of such data anomalies adversely affects the well-known and widely used least square estimation method.

The near linear dependency between the set of covariates known as collinearity or multicollinearity is one of the important problems while estimating the unknown model parameters. Many researchers have considered this problem and proposed various alternative biased estimation methods[2-3]. Some notable references are Hoerl and Kennard (1970a, b), Hoerl et al. (1975), Hocking et al. (1976), Liu (1993, 2003), Troskie and Chalton (1996), Alkhamisi and Shukur (2007), Al-Hassan (2010), Dorugade (2014). The ridge regression estimator (RRE) proposed by Hoerl and Kennard (1970a, b) is the most commonly used estimator when multicollinearity presents in the data[1-7].

The presence of outliers is also one of the important problems which occur more frequently in real examples. Various robust estimators are put forward by many researchers to handle the problem of outliers in the response variable. Some notable references are Huber (1964, 1972)[8], Hample et al. (1986)[9], Rousseeuw and Leroy (1987), Maronna et al. (2006) and Huber and Ronchetti (2009). The M-estimator (ME) based on Huber robust criterion function (see Huber, 1972)[10] is the most popular estimator which dampens the effect of outliers present in the response variable. In the literature, various robust criterion functions (see Holland et al., 1977; Montgomery, et al., 2003)[11-12] are available to obtain ME. The ME obtained using different robust criterion functions have their own advantages and disadvantages.

Some researchers have considered the simultaneous occurrence of outliers and multicollinearity in the data to propose alternative robust parameter estimation methods. Some notable references are Silvapulle (1991)[13], Arslon and Billor (2000), Jadhav and Kashid (2011, 2016)[14-15]. Silvapulle (1991) has considered Huber robust criterion-based ME instead of LSE in RRE to propose ridge M-estimator (RME). This RME tackles the simultaneous occurrence of multicollinearity and outliers in the data.

In this article, we have considered different robust criterion functions to develop ME and RME. A simulation study is carried out to evaluate the performance of the different ME and RME in the presence of only multicollinearity, only outliers and both, multicollinearity and outliers. The article is organized as follows[16].

In Section 2, we introduce a multiple linear regression model and review some existing estimators which are available in the literature to tackle the problem of multicollinearity and/or outliers. Also, we summarize the various robust criterion functions available in the literature. In Section 3, an extensive simulation study is carried out to evaluate the performance of the LSE, ME and proposed RME developed using different robust criterion functions. Section 4 considers the real data set to study the effect of simultaneous occurrence of multicollinearity and outlier on the different estimators. The article ends with a brief summary and overall conclusion in Section 5.

2. REGRESSION MODEL AND SOME ESTIMATORS

The multiple linear regression model is the most widely and commonly used regression technique to model the linear relationship between the variables. The form of multiple linear regression model can be given as

$$Y = X\beta + \vartheta \tag{1}$$

where Y is an $n \times 1$ vector of the response variable, X is an $n \times p$ matrix of covariates, $\beta = (\beta_1, \beta_2, \dots \beta_p)$ is a $p \times 1$ vector of unknown regression coefficients and ϑ is an $n \times 1$ vector of random errors supposed to follow a normal distribution with constant but unknown variance σ^2. Without loss of generality, we consider that the response variable Y and covariates X are standardized in such a way that the $X'X$ is in the form of a correlation matrix and $X'Y$ is the correlation vector between variables X and Y.

It is well known that the least squares estimator (LSE) is widely used to estimate the unknown model parameters. The form of LSE is given by

$$\hat{\beta}_{LSE} = (X'X)^{-1}X'Y \tag{2}$$

As the LSE is unbiased, the covariance and mean squared error (MSE) of the LSE is given by

$$\begin{aligned} \mathrm{Cov}(\hat{\beta}_{LSE}) &= \mathrm{Cov}((X'X)^{-1}X'Y) \\ &= \sigma^2 (X'X)^{-1} \end{aligned} \tag{3}$$

$$\begin{aligned} \mathrm{MSE}(\hat{\beta}_{LSE}) &= tr\left(\mathrm{Cov}(\hat{\beta}_{LSE})\right) \\ &= \sigma^2 \textstyle\sum_{j=1}^{p} \frac{1}{\lambda_j} \end{aligned} \tag{4}$$

where λ_j, $j = 1, 2, \ldots p$ are the eigenvalues of $X'X$ matrix. In the presence of multicollinearity, some of the λ_j's are too small and consequently, the MSE of LSE becomes large. Due to inflated MSE, the LSE may give unreliable and misleading results.

2.1. Ridge Regression Estimator (RRE) in the Presence of Multicollinearity

To tackle the problem of multicollinearity, Hoerl and Kennard (1970a, b) proposed ridge regression estimator (RRE). The RRE is widely used due to its optimality properties (Vinod and Ullah, 1981). The RRE is obtained by simply adding positive constant 'k' to the $(X'X)^{-1}$ matrix of LSE. Hence the form of RRE is given by

$$\begin{aligned} \hat{\beta}_{RRE} &= (X'X + kI)^{-1}X'Y \\ &= (X'X + kI)^{-1}X'X\hat{\beta}_{LSE} \end{aligned} \tag{5}$$

where $k > 0$ is a biasing constant known as shrinkage parameter. Various choices of shrinkage parameter (k) are available in the literature. The choice of k proposed by Hoerl, Kennard and Baldwin (1975) is widely used and it is given by

$$k = \frac{p\sigma^2}{\beta'\beta} \tag{6}$$

where σ^2 and β are the unknown model parameters to be replaced by their estimates based on the LSE.

2.2. M-estimator (ME) in the Presence of Outliers

To tackle the problem of outliers in the response variable, various robust estimation methods like M-estimator (ME), least median squares estimator (LMSE), least trimmed squares estimator (LTSE) (see Rousseeuw and Leroy, 1987) are available in the literature. The ME is the most popular estimator which is obtained by minimizing

$$\sum_{i=1}^{n} \rho\left(\frac{Y_i - X_i'\beta}{\sigma}\right) \tag{7}$$

where $\rho(\cdot)$ is any robust criterion function and σ is a scale parameter. After differentiating above equations partially with respect to each parameter β_j, we get p nonlinear equations of the form

$$\sum_{i=1}^{n} \psi\left(\frac{Y_i - X_i'\beta}{\sigma}\right) X_{ij} = 0,\ j = 1, 2, \dots, p \tag{8}$$

where $\psi(\cdot)$ is partial derivative of $\rho(\cdot)$ with respect to β (See Huber, 1972). To find the estimate of β, we solve the above nonlinear equations by using iterative reweighted least squares method. The flowchart given in Fig. 1 shows the process of estimation of ME. At convergence, the form of ME is given by

$$\hat{\beta}_{ME} = (X'WX)^{-1}X'WY \tag{9}$$

In the literature, various robust criterion functions are available to develop ME. The Table 1 represents some robust criterion functions (ρ) along with their first order derivatives (ψ), weights (W) and ranges (Holland et al., 1977). Among the different robust criterion functions, the Huber's robust criterion function is most popularly used.

2.3. Ridge M-estimator (RME) in the Presence of Outliers and Multicollinearity

To tackle the problem of simultaneous occurrence of outliers and multicollinearity, various robust alternative methods are available in the literature like ridge M-estimator (Silvapulle, 1991)[12], Liu-type M-estimators (Arslon and Billor, 2000)[17], jackknifed ridge M-estimator (Jadhav and Kashid, 2011)[18], linearized ridge M-estimator (Jadhav and Kashid, 2016)[19]. Among these, the ridge M-estimator (RME) proposed by silvapulle (1991)[20] is widely used. The form of RME is given by

$$\hat{\beta}_{RME} = (X'X + kI)^{-1}X'X\hat{\beta}_{ME} \tag{10}$$

where k is shrinkage parameter obtained robustly by replacing the unknown model parameters with their robust estimates in the expression of choice given by Hoerl, Kennard and Baldwin (1975)[21-26] that is, $k = ps^2/\left(\hat{\beta}_{ME}'\hat{\beta}_{ME}\right)$, $\hat{\beta}_{ME}$ denote the M-estimator of β and s is a robust estimate of σ obtained by using the formula $s = 1.4826\ \text{median}|r_i - \text{median}(r_i)|$, $r_i = \left(Y_i - X_i'\hat{\beta}_{ME}\right)$ (see Silvapulle, 1991).

Table-1. Robust criterion functions

Name	$\rho(r)$	$\psi(r)$	$W(r)$	Range
A (Andrews et. al., 1972)	$\begin{cases} A^2[1-cos(r/A)] \\ 2A^2 \end{cases}$	$\begin{cases} A\,sin(r/A) \\ 0 \end{cases}$	$\begin{cases} (r/A)^{-1}\,sin(r/A) \\ 0 \end{cases}$	$\begin{cases} \lvert r\rvert \le \pi A \\ \lvert r\rvert \ge \pi A \end{cases}$ $A = 1.339$
B (Beaton and Tukey, 1974)	$\begin{cases} (B^2/2)[1-[1-(r/B)^2]^3] \\ B^2/2 \end{cases}$	$\begin{cases} r[1-(r/B)^2]^2 \\ 0 \end{cases}$	$\begin{cases} [1-(r/B)^2]^2 \\ 0 \end{cases}$	$\begin{cases} \lvert r\rvert \le B \\ \lvert r\rvert \ge B \end{cases}$ $B = 4.685$
C (Cauchy or t likelihood)	$(C^2/2)\,log[1+(r/C)^2]$	$r[1+(r/C)^2]^{-1}$	$[1+(r/C)^2]^{-1}$	$C = 2.385$
F (Fair, 1974)	$F^2\big[\lvert r\rvert/F - log[1+\lvert r\rvert/F]\big]$	$r[1+\lvert r\rvert/F]^{-1}$	$[1+\lvert r\rvert/F]^{-1}$	$F = 1.400$
H (Huber 1964)	$\begin{cases} r^2/2 \\ H\lvert r\rvert - H^2/2 \end{cases}$	$\begin{cases} r \\ H * sign\,(r) \end{cases}$	$\begin{cases} 1 \\ H(r)^{-1} \end{cases}$	$\begin{cases} \lvert r\rvert \le H \\ \lvert r\rvert \ge H \end{cases}$ $H = 1.345$
L (Logistic)	$L^2 log[cosh(r/L)]$	$L\,tanh\,(r/L)$	$(r/L)^{-1}$ $tanh\,(r/L)$	$L = 1.205$
T (Hinich and Talwar, 1975)	$\begin{cases} r^2/2 \\ T^2/2 \end{cases}$	$\begin{cases} r \\ 0 \end{cases}$	$\begin{cases} 1 \\ 0 \end{cases}$	$\begin{cases} \lvert r\rvert \le T \\ \lvert r\rvert \ge T \end{cases}$ $T = 2.795$
W (Dennis and Welsch, 1976)	$(W^2/2)\,\big[1-exp[-(r/W)^2]\big]$	$r\,exp[-(r/W)^2]$	$exp[-(r/W)^2]$	$W = 2.985$

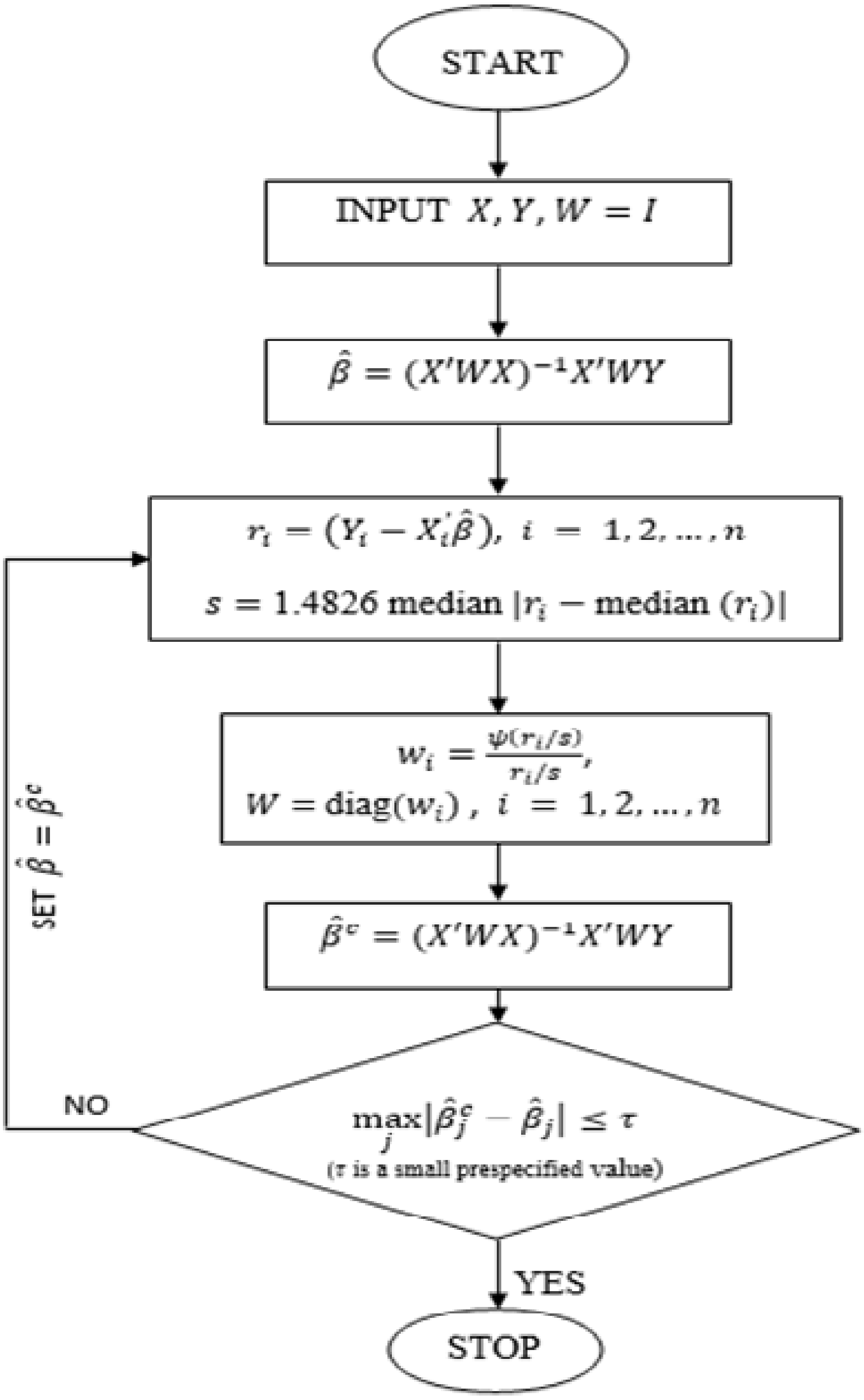


Figure-1. Flowchart showing the process of estimation of M-estimator (ME)

While estimating the ME or RME, various researchers have used Huber's robust criterion function. However, different choices of robust criterion functions are available in the literature. We have tabulated the same in Table 1. There are no significant contributions available in the literature to evaluate the performance of ME or RME developed using different robust criterion functions. By considering this perspective, in this article, we have developed ME and RME based on different robust criterion functions and the performance of the ME and RME is evaluated through MSE sense. The main approach of this article is to compare the performance of ME and RME developed using different robust criterion functions when data suffers from the problem of only multicollinearity, only outliers and simultaneous occurrence of outliers and multicollinearity. An extensive simulation study is carried out in the following section to evaluate the performance of the ME and RME developed using different robust criterion functions.

3. SIMULATION STUDY

In this section, we consider the simulation study to illustrate the performance of the different estimators. To evaluate the performance of an estimator (say $\hat{\beta}$), the Average MSE (AMSE) criterion is used. For different combinations of sample sizes (n), degree of multicollinearity (ρ) and error variance (σ^2), the experiment is repeated $10{,}000$ times and the AMSE of each estimator is obtained by using the formula

$$\text{AMSE} = \frac{1}{10000}\sum_{l=1}^{10000}\left\{\sum_{j=0}^{p}\left(\hat{\beta}_j - \beta_j\right)^2\right\} \tag{11}$$

where β_j denote the true j^{th} regression coefficient and $\hat{\beta}_j$ denote the estimate of β_j.

The one outlier, two outliers etc. in the response variable are introduced by multiplying actual value of Y by twenty corresponding to largest absolute residual, second largest absolute residual etc.

To distinguish between the ME and RME obtained using different robust criterion functions, the capital letters of robust criterion functions given in Table 1 are used. In the simulation study, LSE, RRE, ME with different robust criterion functions and RME with different robust criterion functions are considered to evaluate the performance in AMSE sense.

The simulation study is divided into three parts as follows.

1. Performance of LSE and different robust criterion functions based ME in the presence of outliers

2. Performance of LSE, RRE and different robust criterion functions based ME and RME in the presence of multicollinearity and one outlier

3. Performance of LSE, RRE and different robust criterion functions based ME and RME in the presence of multicollinearity and one and more than one outlier

3.1. Performance of LSE and different robust criterion functions based ME in the presence of outliers

In this subsection, we evaluate the performance of LSE and different robust criterion functions based ME through AMSE. The following regression models are used to generate n observations on the response variable Y as

Model I $Y_i = 0.3 + 0.2X_{i1} + 0.7X_{i2} + 0.4X_{i3} + 0.1X_{i4} + \vartheta_i\,, \; i = 1, 2, \dots, n$ (12)

Model II $Y_i = 5 + 2X_{i1} + 1X_{i2} + 4X_{i3} + 3X_{i4} + \vartheta_i, \quad i = 1, 2, \dots, n$ (13)

where $X_{ij} \sim N(0, 1), \; i = 1, 2, \dots, n, \; j = 1, 2, 3, 4, \vartheta_i \sim N(0, \sigma^2)$.

For $n = 30, 50, 100$ and $\sigma^2 = 1, 25, 100$, the experiment is repeated 10,000 times and the AMSE of LSE and ME based on different robust criterion functions is obtained for Model I and Model II and the results are reported in Table 2 and Table 3 respectively.

From Table 2 and Table 3, it is observed that:

- The AMSE of LSE is smaller than that of the other ME obtained using different robust criterion functions for all combinations of n and σ^2 with no outlier or zero outlier case. As soon as, the outlier introduced in the data, the AMSE of LSE increases considerably as compare to the AMSE of different robust criterion functions based ME.
- For $\sigma^2 = 100$ and $n = 100$, the AMSE of ME obtained using robust criterion function given by Fair (1974) (ME_F) is smaller for one, two and three outliers cases of both models.
- For $\sigma^2 = 1$, the AMSE of ME obtained using Cauchy robust criterion function is smaller than that of the others for Model I with all values of n and one and more than one outlier. However, the AMSE of ME obtained using robust criterion function given by Hinich and Talwar (1975) (ME_T) is smaller for Model II with all values of n and one, two and three outliers except for $n = 30$ and three outliers case of Model II.
- No single specific robust criterion function has better performance than the others for all combinations of n, σ^2 and the presence of different number of outliers.

3.2. Performance of LSE, RRE and different robust criterion functions based ME and RME in the presence of multicollinearity and one outlier

The simulation design given by McDonald and Galarneau (1975) is used to achieve the required degree of multicollinearity in the covariates as

$$X_{ij} = (1 - \rho^2)^{1/2} Z_{ij} + \rho Z_{i(p+1)}, \quad i = 1, 2, \dots, n, \; j = 1, 2, \dots, p \tag{14}$$

where Z_{ij}'s are independent standard normal pseudo-random numbers, ρ^2 is the correlation between any two covariate variables. The n observations on the response variable Y are generated using the following regression model

$$Y_i = 1 + 1X_{i1} + 1X_{i2} + 1X_{i3} + 1X_{i4} + \vartheta_i, \qquad i = 1, 2, \dots, n \tag{15}$$

where $\vartheta_i \sim N(0, \sigma^2)$. In this study, the simulation experiment is replicated 10000 times for $n = 30, 50, 100, \rho = 0.9, 0.99, 0.999, 0.9999$, $\sigma^2 = 1, 100$ and the AMSE of each estimator is obtained. The results of the simulation study for different n are reported in Table 4 to Table 6.

From Table 4 to Table 6, it is observed that:

- For without outlier case with any degree of multicollinearity and different sample sizes, the AMSE of the RRE is smaller than that of the other estimators. Hence the performance of RRE is good in the presence of only multicollinearity. As soon as we introduce the outlier in the response variable, the AMSE of RRE inflates and consequently, the RRE shows poor performance.
- In the presence of multicollinearity with and without outlier cases, the AMSE of ME obtained through different robust criterion functions is more than the AMSE of RME obtained through respective different robust criterion functions.
- For $\sigma^2 = 1$ with one outlier case, the AMSE of the RME obtained using Hinich and Talwar (1975) robust criterion function (RME_T) shows smaller AMSE than that of the other estimators for any degree of multicollinearity.
- For $\sigma^2 = 100$ with one outlier case, the AMSE of the RME obtained using the Logistic robust criterion function (RME_L) has smaller value except for $n = 100$ and $\rho = 0.9,\ 0.999, 0.9999$.

3.3. Performance of LSE, RRE and different robust criterion functions based ME and RME in the presence of multicollinearity and one and more than one outlier

In this subsection, the simulation design given in Subsection 3.2 is used to generate $n = 50$ observations on the response variable. The one outlier, two outliers and three outliers are introduced in the response variable by multiplying actual value of Y by twenty corresponding to largest absolute residual, second largest absolute residual and third largest absolute residual.

The AMSE of LSE, RRE, ME and RME based on different robust criterion functions are obtained for $n = 50,\ \rho = 0.9, 0.99, 0.999, 0.9999, \sigma^2 = 25$ with one outlier, two outliers and three outliers' cases and the results are reported in the Table 7.

From Table 7, it is seen that the AMSE of LSE, RRE and ME obtained using different robust criterion functions is more than that of the RME obtained using different robust criterion functions. The RME obtained using different robust criterion functions shows smaller AMSE value. The RME obtained using Logistic robust criterion function has smaller AMSE value than that of the other existing estimators when data suffers from the problem of multicollinearity with one and more than one outlier.

Table-2. The AMSE of LSE and different robust criterion functions based ME in the presence of outliers for Model I

	$n=30$			$n=50$			$n=100$		
	$\sigma^2=1$	$\sigma^2=25$	$\sigma^2=100$	$\sigma^2=1$	$\sigma^2=25$	$\sigma^2=100$	$\sigma^2=1$	$\sigma^2=25$	$\sigma^2=100$
					0 outlier				
LSE	1.1956	6.1765	21.4937	1.1106	3.7901	12.2885	1.0537	2.3065	6.3265
ME_A	1.2107	6.5578	23.0312	1.1175	3.9643	12.9630	1.0575	2.3788	6.6198
ME_B	1.2103	6.5486	22.9979	1.1174	3.9614	12.9540	1.0575	2.3784	6.6182
ME_C	1.2052	6.4025	22.4028	1.1163	3.9224	12.8282	1.0571	2.3727	6.5974
ME_F	1.2050	6.3890	22.3566	1.1161	3.9193	12.8134	1.0570	2.3721	6.5998
ME_H	1.2047	6.3873	22.3465	1.1162	3.9185	12.8142	1.0571	2.3719	6.5954
ME_L	1.2034	6.3732	22.2702	1.1160	3.9096	12.7813	1.0570	2.3712	6.5831
ME_T	1.1991	6.3347	21.9919	1.1148	3.9031	12.7242	1.0570	2.3649	6.5625
ME_W	1.2075	6.4733	22.6808	1.1168	3.9402	12.8842	1.0573	2.3753	6.6061
					1 outlier				
LSE	18.6007	270.9271	1070.9032	8.2675	120.1806	452.9613	3.3177	37.1397	143.0293
ME_A	1.2538	7.4147	26.4555	1.1375	4.3074	14.0782	1.0600	2.4508	6.8684
ME_B	1.2537	7.4121	26.4472	1.1375	4.3062	14.0764	1.0600	2.4508	6.8685
ME_C	1.2417	7.1615	25.4320	1.1318	4.2414	13.8467	1.0578	2.4430	6.8452
ME_F	1.2753	6.4530	22.4832	1.1432	3.9559	12.6340	1.0610	2.3652	6.5363
ME_H	1.2506	6.3868	22.2375	1.1355	3.9587	12.6703	1.0590	2.3703	6.5553
ME_L	1.2466	6.3780	22.2742	1.1345	3.9575	12.6696	1.0587	2.3677	6.5498
ME_T	1.2460	7.2428	25.7892	1.1338	4.2377	13.8241	1.0588	2.4230	6.7479
ME_W	1.2525	7.3957	26.3677	1.1373	4.3041	14.0752	1.0600	2.4533	6.8791
					2 outliers				
LSE	31.8878	393.8836	1534.0274	14.2929	181.8152	705.8253	5.3962	61.0072	232.7821
ME_A	1.2745	7.8281	27.7360	1.1370	4.5124	14.9391	1.0683	2.5179	7.1872
ME_B	1.2745	7.8260	27.7286	1.1369	4.5122	14.9375	1.0683	2.5179	7.1873
ME_C	1.2622	7.4410	26.3142	1.1276	4.3985	14.5278	1.0639	2.4982	7.1248
ME_F	1.3929	6.6215	23.1675	1.1657	3.9501	12.7759	1.0750	2.3644	6.5627
ME_H	1.3163	6.3927	22.1814	1.1449	3.9409	12.7236	1.0696	2.3705	6.5982
ME_L	1.3079	6.3925	22.1785	1.1425	3.9433	12.7341	1.0691	2.3705	6.5987
ME_T	1.2695	7.6457	27.1035	1.1343	4.4274	14.6480	1.0671	2.4852	7.0485
ME_W	1.2742	7.8246	27.7261	1.1367	4.5187	14.9545	1.0683	2.5215	7.2036
					3 outliers				
LSE	46.7004	473.9135	1804.5054	19.9245	228.1007	858.8484	7.2714	77.8247	300.1968
ME_A	1.2804	8.0381	28.5673	1.1444	4.5988	15.4098	1.0616	2.5873	7.3707
ME_B	1.2803	8.0352	28.5609	1.1443	4.5982	15.4085	1.0616	2.5873	7.3706
ME_C	1.2715	7.5114	26.6833	1.1326	4.4406	14.8308	1.0549	2.5549	7.2653
ME_F	1.6073	6.9292	24.2211	1.2117	3.9704	12.7422	1.0753	2.3759	6.5683
ME_H	1.4066	6.4085	22.1602	1.1721	3.9126	12.5784	1.0659	2.3828	6.5887
ME_L	1.3883	6.3923	22.1399	1.1669	3.9142	12.6046	1.0648	2.3844	6.5875
ME_T	1.2792	7.8748	28.4414	1.1418	4.5139	15.0884	1.0605	2.5494	7.2271
ME_W	1.2801	8.0404	28.5863	1.1440	4.6059	15.4406	1.0614	2.5916	7.3874

Table-3. The AMSE of LSE and different robust criterion functions based ME in the presence of Outliers for Model II

	$n=30$			$n=50$			$n=100$		
	$\sigma^2=1$	$\sigma^2=25$	$\sigma^2=100$	$\sigma^2=1$	$\sigma^2=25$	$\sigma^2=100$	$\sigma^2=1$	$\sigma^2=25$	$\sigma^2=100$
0 outlier									
LSE	1.1994	6.1115	21.3961	1.1107	3.8101	12.2386	1.0524	2.3285	6.2948
ME_A	1.2165	6.5186	22.8974	1.1175	3.9882	12.9934	1.0552	2.4106	6.5863
ME_B	1.2161	6.5116	22.8723	1.1174	3.9852	12.9812	1.0552	2.4103	6.5845
ME_C	1.2100	6.3806	22.3304	1.1162	3.9482	12.8238	1.0550	2.4041	6.5627
ME_F	1.2092	6.3646	22.2666	1.1160	3.9432	12.8029	1.0551	2.4033	6.5612
ME_H	1.2094	6.3653	22.2701	1.1161	3.9435	12.8059	1.0549	2.4033	6.5609
ME_L	1.2088	6.3456	22.2085	1.1156	3.9401	12.7795	1.0546	2.4001	6.5489
ME_T	1.2049	6.2392	21.9642	1.1150	3.9220	12.7097	1.0551	2.4004	6.5523
ME_W	1.2131	6.4525	22.5872	1.1167	3.9651	12.8937	1.0550	2.4069	6.5719
1 outlier									
LSE	161.3698	417.3674	1227.5165	61.9998	174.3699	500.6006	18.0252	51.2266	155.1078
ME_A	1.2629	7.3292	26.5304	1.1303	4.2872	14.1187	1.0584	2.4517	7.0268
ME_B	1.2628	7.3268	26.5229	1.1302	4.2868	14.1171	1.0584	2.4516	7.0269
ME_C	1.2662	7.1814	25.4666	1.1326	4.2395	13.8295	1.0597	2.4382	6.9820
ME_F	1.4636	7.1922	23.0958	1.2274	4.1928	12.7649	1.1011	2.4131	6.6516
ME_H	1.3734	6.8779	22.6969	1.1873	4.1149	12.7688	1.0847	2.4011	6.6810
ME_L	1.3612	6.8242	22.6624	1.1815	4.0964	12.7615	1.0820	2.3977	6.6749
ME_T	1.2566	7.1864	25.9202	1.1282	4.2197	13.8877	1.0580	2.4235	6.9275
ME_W	1.2622	7.3101	26.4386	1.1302	4.2876	14.1136	1.0585	2.4534	7.0359
2 outliers									
LSE	377.4864	751.0592	1885.7093	137.9229	314.0009	837.7166	39.1978	95.1373	263.0477
ME_A	1.2617	7.8365	28.0390	1.1384	4.4540	14.9573	1.0637	2.5204	7.1870
ME_B	1.2616	7.8345	28.0274	1.1384	4.4533	14.9544	1.0637	2.5204	7.1870
ME_C	1.2726	7.7281	26.5378	1.1441	4.3849	14.4739	1.0662	2.4952	7.0852
ME_F	1.7666	9.0764	24.8065	1.3488	4.5213	13.0976	1.1521	2.4917	6.5692
ME_H	1.5132	7.8799	23.2825	1.2600	4.2927	12.9350	1.1175	2.4529	6.5946
ME_L	1.4843	7.7350	23.1895	1.2475	4.2569	12.9158	1.1124	2.4473	6.5906
ME_T	1.2559	7.7127	27.3884	1.1353	4.3847	14.6567	1.0621	2.4900	7.0594
ME_W	1.2618	7.8294	27.9721	1.1387	4.4557	14.9715	1.0638	2.5233	7.2010
3 outliers									
LSE	633.6189	1067.9714	2410.9629	234.5688	449.6685	1095.4769	64.7617	135.9063	351.0168
ME_A	1.2713	8.0774	28.9696	1.1456	4.6061	15.4874	1.0656	2.5724	7.3990
ME_B	1.2713	8.0753	28.9579	1.1456	4.6057	15.4854	1.0656	2.5724	7.3987
ME_C	1.2874	8.0521	27.2627	1.1542	4.5386	14.8288	1.0695	2.5390	7.2383
ME_F	3.6413	12.6932	28.1444	1.4878	5.0297	13.5120	1.2024	2.5946	6.5973
ME_H	1.6999	9.1034	24.5104	1.3381	4.5745	13.0944	1.1483	2.5157	6.6073
ME_L	1.6479	8.8018	24.2736	1.3170	4.5093	13.0627	1.1402	2.5048	6.6059
ME_T	6.9569	8.2496	28.6810	1.1429	4.5388	15.1695	1.0638	2.5390	7.2593
ME_W	1.2719	8.0911	28.9530	1.1460	4.6134	15.5079	1.0659	2.5759	7.4145

Table-4. AMSE of LSE, RRE ME and RME obtained using different robust criterion functions for $n = 30$

	Without outliers				With one outlier			
ρ	0.9	0.99	0.999	0.9999	0.9	0.99	0.999	0.9999
					$\sigma^2 = 1$			
LSE	1.713	7.288	63.465	631.539	151.485	1392.441	13844.643	136365.238
RRE	1.482	3.430	17.011	152.679	38.402	285.214	2636.170	25676.256
ME_A	1.774	7.749	68.406	677.976	1.894	8.833	78.304	778.415
ME_B	1.773	7.741	68.312	677.131	1.894	8.828	78.263	778.029
ME_C	1.750	7.567	66.270	659.155	1.892	8.775	77.546	772.935
ME_F	1.747	7.549	66.051	657.183	2.048	9.600	85.474	848.939
ME_H	1.747	7.548	66.069	657.290	1.942	8.871	78.381	778.452
ME_L	1.746	7.526	65.893	655.488	1.927	8.763	77.340	769.514
ME_T	1.734	7.472	65.103	652.388	1.873	8.635	76.433	764.285
ME_W	1.762	7.651	67.271	668.022	1.891	8.802	77.933	776.297
RME_A	1.578	4.234	24.744	227.253	1.684	4.962	30.395	282.557
RME_B	1.577	4.228	24.670	226.703	1.683	4.958	30.355	282.196
RME_C	1.559	4.113	23.261	213.821	1.685	4.945	30.047	281.214
RME_F	1.558	4.124	23.277	213.670	1.829	5.546	35.214	332.878
RME_H	1.557	4.107	23.172	212.817	1.735	5.050	30.761	288.693
RME_L	1.549	4.027	22.511	207.103	1.715	4.903	29.420	276.618
RME_T	1.533	3.926	21.513	201.122	1.657	4.729	28.193	264.111
RME_W	1.568	4.166	23.926	220.103	1.682	4.946	30.166	281.533
					$\sigma^2 = 100$			
LSE	71.150	635.399	6309.948	62171.052	3717.824	32298.429	326455.412	3189295.807
RRE	20.705	156.897	1534.162	14913.509	791.471	5817.139	57769.521	551325.960
ME_A	76.718	682.393	6784.330	67582.283	88.635	784.520	7816.332	77854.621
ME_B	76.619	681.506	6774.984	67476.272	88.581	783.988	7810.620	77835.474
ME_C	74.508	662.322	6595.824	65527.316	84.913	754.403	7491.177	74882.114
ME_F	74.212	660.283	6575.657	65284.233	75.132	670.728	6653.431	65599.293
ME_H	74.278	660.358	6574.485	65315.724	74.124	662.088	6562.540	65134.877
ME_L	74.109	657.802	6552.966	65013.813	74.146	661.363	6559.707	65056.778
ME_T	73.224	655.759	6516.344	64184.962	85.725	765.262	7591.581	75814.077
ME_W	75.529	671.678	6688.823	66529.729	88.240	781.942	7779.404	77601.632
RME_A	29.491	230.533	2263.589	22718.111	36.610	290.158	2855.635	28414.177
RME_B	29.417	229.797	2256.435	22636.973	36.557	289.691	2850.547	28389.635
RME_C	27.945	216.344	2136.855	21227.988	34.363	273.193	2672.677	26744.680
RME_F	27.945	216.371	2133.894	21206.983	27.689	217.938	2127.522	20907.360
RME_H	27.839	215.387	2125.351	21114.147	27.601	217.168	2117.036	21002.880
RME_L	27.045	208.540	2059.458	20419.193	26.974	211.528	2063.036	20431.361
RME_T	25.889	202.201	1983.654	19459.365	33.519	267.534	2606.574	26111.666
RME_W	28.620	222.820	2201.366	21943.978	36.335	288.667	2832.361	28237.383

Table-5. AMSE of LSE, RRE ME and RME obtained using different robust criterion functions for $n = 50$

	Without outliers				With one outlier			
ρ	0.9	0.99	0.999	0.9999	0.9	0.99	0.999	0.9999
				$\sigma^2 = 1$				
LSE	1.388	4.436	35.279	347.054	60.195	519.631	5112.862	52164.533
RRE	1.289	2.646	10.204	84.377	17.437	111.811	1020.359	10761.119
ME_A	1.413	4.653	37.499	370.838	1.453	5.062	41.191	407.315
ME_B	1.413	4.650	37.460	370.337	1.453	5.061	41.186	407.164
ME_C	1.408	4.606	36.980	365.416	1.454	5.060	41.222	406.877
ME_F	1.407	4.602	36.911	364.928	1.509	5.277	43.122	423.278
ME_H	1.407	4.602	36.923	364.940	1.474	5.072	41.203	405.003
ME_L	1.406	4.592	36.847	363.969	1.468	5.034	40.869	401.675
ME_T	1.404	4.565	36.592	359.480	1.446	4.988	40.359	398.637
ME_W	1.410	4.625	37.196	367.487	1.453	5.062	41.196	407.078
RME_A	1.324	2.946	12.941	111.581	1.360	3.211	14.927	130.307
RME_B	1.324	2.943	12.912	111.199	1.360	3.210	14.922	130.148
RME_C	1.319	2.917	12.623	108.072	1.362	3.220	15.022	130.488
RME_F	1.319	2.920	12.615	108.140	1.415	3.399	16.279	142.245
RME_H	1.319	2.917	12.604	107.907	1.382	3.250	15.097	130.850
RME_L	1.316	2.889	12.395	105.860	1.375	3.199	14.703	126.958
RME_T	1.312	2.849	12.120	102.072	1.351	3.127	14.089	122.074
RME_W	1.321	2.927	12.742	109.277	1.360	3.214	14.948	130.128
				$\sigma^2 = 100$				
LSE	39.892	346.811	3452.683	33793.357	1618.081	13748.432	138432.106	1367440.751
RRE	12.445	85.443	832.234	7874.837	372.131	2661.419	26415.532	256428.775
ME_A	42.454	367.732	3675.145	35941.967	46.512	405.783	4045.994	39680.356
ME_B	42.413	367.450	3672.405	35907.605	46.499	405.683	4045.239	39668.640
ME_C	41.855	363.780	3631.637	35460.327	45.627	398.757	3975.466	38918.397
ME_F	41.730	363.559	3624.967	35374.367	41.558	362.529	3612.679	35236.315
ME_H	41.779	363.454	3626.647	35402.378	41.663	363.795	3625.710	35401.365
ME_L	41.742	362.200	3618.117	35376.028	41.691	363.274	3621.281	35438.925
ME_T	41.434	359.884	3599.965	35178.121	45.496	396.373	3963.095	38898.870
ME_W	42.107	365.344	3650.248	35659.769	46.469	405.687	4045.192	39642.505
RME_A	15.654	110.555	1091.822	10335.632	17.928	130.651	1285.689	12243.232
RME_B	15.620	110.343	1090.021	10309.958	17.915	130.550	1285.123	12232.576
RME_C	15.286	108.200	1066.668	10072.018	17.466	127.412	1253.056	11886.558
RME_F	15.272	108.521	1067.436	10082.493	15.086	107.235	1051.786	9950.884
RME_H	15.259	108.165	1065.365	10058.731	15.175	108.427	1064.044	10063.479
RME_L	15.035	105.869	1043.960	9866.914	14.982	106.587	1044.977	9900.576
RME_T	14.715	103.137	1022.353	9669.994	16.990	121.880	1209.640	11485.496
RME_W	15.420	109.008	1076.043	10162.312	17.910	130.671	1286.098	12225.762

Table-6. AMSE of LSE, RRE ME and RME obtained using different robust criterion functions for $n = 100$

	Without outliers				With one outlier			
ρ	0.9	0.99	0.999	0.9999	0.9	0.99	0.999	0.9999
				$\sigma^2 = 1$				
LSE	1.186	2.624	16.921	159.862	16.880	143.698	1419.197	14269.387
RRE	1.150	2.008	5.818	39.020	6.167	33.822	310.504	2975.673
ME_A	1.196	2.721	17.818	168.425	1.206	2.814	18.786	177.421
ME_B	1.196	2.720	17.813	168.379	1.206	2.814	18.786	177.430
ME_C	1.195	2.711	17.729	167.645	1.207	2.820	18.823	177.855
ME_F	1.195	2.710	17.725	167.595	1.226	2.858	19.021	179.774
ME_H	1.195	2.710	17.718	167.557	1.216	2.816	18.667	176.343
ME_L	1.194	2.706	17.696	167.376	1.214	2.804	18.584	175.764
ME_T	1.196	2.709	17.667	167.169	1.204	2.785	18.472	174.908
ME_W	1.195	2.715	17.766	167.977	1.206	2.816	18.809	177.692
RME_A	1.161	2.107	6.648	46.292	1.171	2.173	7.142	50.686
RME_B	1.161	2.107	6.644	46.260	1.171	2.173	7.142	50.687
RME_C	1.160	2.101	6.604	45.859	1.172	2.180	7.178	51.011
RME_F	1.160	2.102	6.611	45.852	1.191	2.217	7.338	52.217
RME_H	1.160	2.101	6.602	45.828	1.181	2.182	7.126	50.377
RME_L	1.159	2.094	6.545	45.440	1.178	2.169	7.033	49.716
RME_T	1.160	2.092	6.503	45.306	1.168	2.145	6.907	48.835
RME_W	1.160	2.103	6.618	46.026	1.171	2.175	7.159	50.852
				$\sigma^2 = 100$				
LSE	19.374	163.836	1581.939	16031.977	485.056	4384.449	42827.017	424174.632
RRE	6.750	40.901	368.324	3716.053	116.988	924.146	8607.990	84283.164
ME_A	20.437	172.737	1674.173	16958.090	21.559	180.899	1771.272	17775.257
ME_B	20.432	172.686	1673.585	16950.649	21.560	180.901	1771.346	17775.166
ME_C	20.349	172.003	1660.739	16850.174	21.453	180.082	1758.590	17670.261
ME_F	20.339	172.032	1656.706	16833.850	20.259	170.487	1651.471	16638.959
ME_H	20.339	171.935	1658.331	16836.124	20.379	171.076	1663.115	16741.038
ME_L	20.316	171.512	1659.402	16812.204	20.371	170.706	1665.503	16734.919
ME_T	20.233	171.034	1666.561	16825.035	21.185	177.577	1745.327	17508.300
ME_W	20.385	172.294	1667.234	16897.216	21.589	181.178	1772.507	17794.442
RME_A	7.789	48.846	446.532	4513.201	8.427	52.985	494.041	4911.195
RME_B	7.785	48.810	446.079	4508.045	8.427	52.982	494.024	4910.680
RME_C	7.743	48.496	439.082	4448.649	8.380	52.705	488.604	4863.160
RME_F	7.750	48.565	437.570	4436.354	7.689	47.750	433.558	4337.690
RME_H	7.742	48.490	438.096	4441.176	7.760	48.131	440.477	4397.138
RME_L	7.680	47.862	434.834	4402.581	7.705	47.531	437.937	4365.010
RME_T	7.600	47.316	437.720	4394.920	8.134	50.581	475.997	4704.227
RME_W	7.757	48.599	442.346	4475.118	8.449	53.169	494.891	4921.300

Table-7. AMSE of LSE, RRE ME and RME obtained using different robust criterion functions for $n = 50$ with multicollinearity and one and more than one outlier

Estimators	1 outlier	2 outliers	3 outliers	1 outlier	2 outliers	3 outliers
	$\rho = 0.9$			$\rho = 0.99$		
LSE	429.232	681.533	887.913	3910.242	6040.036	7790.447
RRE	102.285	146.668	183.379	783.243	1107.667	1320.946
ME_A	12.549	13.271	13.714	102.544	107.449	113.471
ME_B	12.547	13.269	13.712	102.511	107.434	113.456
ME_C	12.321	12.884	13.160	100.409	104.253	109.467
ME_F	11.347	11.479	11.821	91.753	93.028	97.868
ME_H	11.368	11.407	11.510	91.971	91.967	95.279
ME_L	11.346	11.400	11.506	91.965	91.815	95.009
ME_T	12.290	12.952	13.466	100.209	105.207	111.122
ME_W	12.544	13.283	13.727	102.406	107.503	113.661
RME_A	5.966	6.481	6.767	35.007	37.664	41.223
RME_B	5.963	6.479	6.765	34.976	37.648	41.209
RME_C	5.839	6.251	6.420	33.947	36.050	39.158
RME_F	5.235	5.300	5.546	29.049	29.771	32.002
RME_H	5.248	5.296	5.389	29.319	29.380	30.989
RME_L	5.169	5.226	5.324	28.849	28.836	30.308
RME_T	5.694	6.169	6.521	32.851	35.573	39.016
RME_W	5.970	6.498	6.782	34.941	37.728	41.420
	$\rho = 0.999$			$\rho = 0.9999$		
LSE	37915.185	60239.346	76044.160	377658.812	602087.286	762099.846
RRE	7233.452	10658.626	11995.423	73109.284	103112.711	120817.548
ME_A	1017.312	1060.452	1114.439	10043.999	10560.141	11075.158
ME_B	1017.077	1060.205	1114.228	10042.606	10559.166	11074.421
ME_C	998.590	1029.555	1068.266	9851.718	10238.498	10648.625
ME_F	915.241	921.960	938.951	9065.915	9099.276	9382.889
ME_H	916.455	910.887	918.395	9063.035	9009.942	9168.882
ME_L	913.763	908.929	918.783	9049.744	9021.569	9187.706
ME_T	993.117	1034.443	1086.330	9867.499	10323.164	10832.473
ME_W	1016.811	1061.686	1115.453	10038.424	10564.390	11097.481
RME_A	320.204	348.210	380.507	3140.625	3443.448	3772.520
RME_B	319.994	348.009	380.314	3139.224	3442.547	3771.404
RME_C	310.702	333.168	355.017	3052.308	3281.872	3554.506
RME_F	263.745	271.085	278.583	2626.989	2621.221	2808.262
RME_H	266.117	268.286	271.751	2636.898	2599.655	2728.275
RME_L	260.804	263.803	268.081	2585.465	2573.462	2695.232
RME_T	298.148	326.404	357.481	2970.362	3239.857	3563.930
RME_W	319.995	349.548	381.306	3140.817	3450.310	3793.217

4. REAL DATA APPLICATION: TOBACCO BLENDS DATA

In this section, we consider the real data set on tobacco blends given by Myers (1990)[28] to evaluate the performance of the LSE, RRE, ME and RME obtained using different robust criterion functions. The tobacco blends data contains 30 observations on the amount of heat evolved from tobacco during the smoking process (response variable, Y) and percentage concentration of four important components (covariates X_1, X_2, X_3 and X_4). The canonical form of the model is considered to model the data (Arslon and Billor, 2000; Jadhav and Kashid, 2016). Myers (1990), Arslon and Billor (2000)[29], Jadhav and Kashid (2016)[30] pointed out that, this data suffers from the simultaneous occurrence of outliers and multicollinearity. For this data, we estimate the LSE, RRE, ME and RME obtained using different robust criterion functions with their norm of the estimates and estimates of the respective shrinkage parameters. The results are reported in Table 8.

Table 8: Estimates of LSE, RRE ME and RME obtained using different robust criterion functions with norm of the estimates and estimates of respective shrinkage parameters

Estimators	Estimates				Estimate of Shrinkage Parameter	Norm of the Estimates
	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\alpha}_4$		
LSE	0.4857	-0.6727	-1.0746	1.4436	-	3.9272
RRE	0.4855	-0.6142	-0.8510	0.8097	0.0017	1.9927
ME_A	0.4894	-0.6601	-1.1181	-0.5346	-	2.2112
ME_B	0.4894	-0.6599	-1.1193	-0.5290	-	2.2076
ME_C	0.4886	-0.6559	-1.1720	0.2251	-	2.0932
ME_F	0.4872	-0.6754	-1.1789	0.8892	-	2.8741
ME_H	0.4864	-0.6509	-1.1805	0.3566	-	2.1812
ME_L	0.4873	-0.6639	-1.1787	0.6075	-	2.4366
ME_T	0.4858	-0.6831	-1.0584	-0.9285	-	2.6851
ME_W	0.4899	-0.6543	-1.1362	-0.3728	-	2.0979
RME_A	0.4892	-0.6141	-0.9267	-0.3310	0.0013	1.5847
RME_B	0.4892	-0.6141	-0.9285	-0.3281	0.0013	1.5863
RME_C	0.4885	-0.6344	-1.0722	0.1762	0.0006	1.8218
RME_F	0.4871	-0.6688	-1.1476	0.8223	0.0002	2.6778
RME_H	0.4863	-0.6323	-1.0919	0.2872	0.0005	1.9111
RME_L	0.4873	-0.6525	-1.1244	0.5312	0.0003	2.2098
RME_T	0.4857	-0.6364	-0.8805	-0.5796	0.0013	1.7522
RME_W	0.4897	-0.6149	-0.9659	-0.2444	0.0011	1.6105

Form Table 8, it can be seen that the simultaneous presence of multicollinearity and outliers affects the estimates as well as norm of the estimates. It is expected and observed that, the norm of LSE is larger than that of the other existing estimators. It is also observed that the norm of RME obtained using different robust criterion functions is smaller than that of ME obtained using respective robust criterion functions. Based on the norm of the estimate's criterion, the RME obtained using robust criterion function given by Andrews et al. (1972) shows smaller value than that of the other estimators.

5. SUMMARY AND CONCLUSIONS

In this article, we have compared the performance of the least squares estimator (LSE), ridge regression estimator (RRE) and M-estimator (ME) as well as ridge M-estimator (RME) obtained using different robust criterion functions. A real data set and simulation study were considered to evaluate the performance using the mean squared error (MSE) criterion. It is observed that the RME obtained using different robust criterion functions has smaller average MSE as compare to the other estimators. It seems that the no RME obtained using any specific robust criterion function shows uniformly better performance when the data suffers from the problem of simultaneous occurrence of multicollinearity and outliers. More specifically, for large error variance with large sample size, the RME obtained using Logistic robust criterion function shows smaller average MSE.

REFERENCES

1. Al-Hassan, Y. (2010) Journal of the Association of Arab Universities for Basic and Applied Sciences 9, 23–26.
2. Alkhamisi, M.A. and Shukur, G. (2007) Communications in Statistics–Simulation and Computation 36, 535–547.
3. Andrews, D., Bickel, P., Hampel, F., Huber, P., Rogers, W. and Tukey, J. (1972). Robust Estimates of Location: Survey and Advances. Princeton: Princeton University Press.
4. Arslan, O. and Billor, N. (2000) Journal of Applied Statistics, 27, 39–47.
5. Beaton, A.E. and Tukey, J.W. (1974) Technometrics, 16, 147–85.
6. Dennis, J.E. and Welsch, R.E. (1976) American Statistical Association,22, 83–87.
7. Dorugade, A.V. (2014) Journal of the Association of Arab Universities for Basic and Applied Sciences 15, 94–99.
8. Fair, R.C. (1974) Annals of Economic and Social Measurement, 2, 667–778.
9. Hampel, F.R., Ronchetti, E.M., Rousseeuvw, P.J. Stahel, W.A. (1986) Robust Statistics: The Approach Based on Influence Function. New York: Wiley, 1-77.
10. Hinich, M.J. and Talwar, P.P. (1975) Journal of the American Statistical Association 70, 113–19.
11. Hocking, R.R., Speed, F.M. and Lynn, M.J. (1976) Technometrics, 18(4), 425–437.

12. Hoerl, A.E. and Kennard, R.W. (1970) Ridge regression: biased estimation for nonorthogonal problems. Technometrics 12, 55–67.

13. Hoerl, A.E. and Kennard, R.W. (1970) Ridge regression: applications to nonorthogonal problems. Technometrics 12, 69–82.

14. Hoerl, A.E., Kennard, R.W. and Baldwin, K.F. (1975) Ridge regression: some simulations. Communication in Statistics 4,105–123.

15. Holland, P.W., Roy E. and Welsch, R.E. (1977) Communications in Statistics–Theory and Methods 6, 813–827.

16. Huber, P.J. (1964) Annals of Mathematical Statistics 35, 73-101.

17. Huber, P.J. (1972) Annals of Mathematical Statistics 43, 1041–1067.

18. Huber, P.J., Ronchetti, E.M. (2009) Robust Statistics. 2nd Edition, New York: John Wiley.

19. Jadhav, N.H., Kashid, D.N. (2011) Journal of Statistical Theory and Practice 45, 1001–1024.

20. Jadhav, N.H., Kashid, D.N. (2016) Communication in Statistics–Simulation and Computation 45, 1001–1024.

21. Liu, K. (1993) Communications in Statistics–Theory and Methods, 22, 393–402.

22. Liu, K. (2003) Communications in Statistics–Theory and Methods 32, 1009–1020.

23. Maronna, R.A., Martin, D.R. and Yohai, V.J. (2006) Robust Statistics: Theory and Methods. New York: John Wiley.

24. McDonald, G.C., Galarneau, D.I. (1975) Journal of the American Statistical Association 70, 407–416.

25. Montgomery, D.C., Peck, E.A., Vining, G.G. (2003) Introduction to Linear Regression Analysis. 3rd ed. New York: Wiley.

26. Myers, R. H. (1990). Classical and Modern Regression with Applications. 2nd ed. Boston: MA Duxbury.

27. Rousseeuw, P.J. and Leroy, A.M. (1987). Robust Regression and Outlier Detection. New York: Wiley.

28. Silvapulle, M.J. (1991) Australian Journal of Statistics 33, 319–333.

29. Troskie, C.G. and Chalton, D.O. (1996) South African Statistical Journal 30, 119–137.

30. Vinod, H.D., Ullah, A. (1981) Recent Advance in Regression Methods. New York: Marcel Dekker.

GUIDELINES FOR CONTRIBUTORS

1] **Journal of Shivaji University : Science & Technology** is the publication of Shivaji University, Kolhapur (Maharashtra, India), being published twice a year. It is an academic peer reviewed ISSN approved Journal.

2] The Journal welcomes research articles based on original research, review papers, research communication/notes by the faculty and research scholars working in various fields of Science and Technology. Articles/Papers can be submitted in English language only.

3] The length of manuscripts should not exceed 6000 words for review article, 5000 words for research article, 2000 words for research communication and 1000 words for research note.

4] The article/paper must accompany an **abstract not exceeding 200 words.**

5] During writing and submission of a manuscript, please follow the author guidelines arakble on the home page of Journal of Shivaji University: Science and Technology.

6] All the sources of literature referred to while writing the article/paper must be properly cited in the text. The serial numbers of End Notes, if any, must also be indicated within text at appropriate places.

7] The listing of references must follow the standard format.

8] Tables, charts, maps, figures etc. should be placed at appropriate places in the text of the article/paper and must be numbered serially with suitable headings. The tables should be referred to by their numbers within the text. Art-work for maps, figures and charts should be provided separately, if necessary.

9] Only articles evaluated and approved by the subject Experts/Referees/Departmental editorial board are considered for their publication in the Journal. The referees evaluate Article/ Paper drafts in term of structure and organization of paper/argument, originality of the research, appropriateness of abstract and introduction, literature review, methodology and data sources, data/evidence and conclusions, and quality of language.

10] The name of the author/co-author of the article being submitted should mentioned on the first page of the article/paper below title of that article. An asterisks (*) is used for defining corresponding author/s of a manuscript.

11] For any other information and for inquiries regarding submitted articles/papers preferably use e-mail communications. **(E-mail : editor.jscitech@unishivaji.ac.in)**

12] Manuscript should be sent to email of the Editor (**editor.jscitech@unishivaji.ac.in**) {or associate editor of the concerned subject}, Journal of Shivaji University: Science and Technology, Kolhapur with e-copy only (single M.S. word as well as PDF file).

Only research articles/review papers/research communication/notes prepared strictly in accordance with the above guidelines will be sent to referees for their evaluation and opinion about their acceptability or otherwise. We do not send back rejected articles.

शिवाजी विद्यापीठ कोल्हापूर

ज्ञानमेवामृतम्